

SET-1

1. Question

A company wants to build an application that stores images in a Cloud Storage bucket and wants to generate thumbnails as well as resize the images. They want to use a google managed service that can scale up and scale down to zero automatically with minimal effort. You have been asked to recommend a service. Which GCP service would you suggest?

- Google Compute Engine
- Google Kubernetes Engine
- **Cloud Functions**
- Google App Engine

Incorrect

Cloud Functions. is the right answer.

Cloud Functions is Google Cloud's event-driven serverless compute platform. It automatically scales based on the load and requires no additional configuration. You pay only for the resources used.

Ref: <https://cloud.google.com/functions>

While all other options i.e. Google Compute Engine, Google Kubernetes Engine, Google App Engine support autoscaling, it needs to be configured explicitly based on the load and is not as trivial as the scale up or scale down offered by Google's cloud functions.

2. Question

A team of data scientists infrequently needs to use a Google Kubernetes Engine (GKE) cluster that you manage. They require GPUs for some long-running, non-restartable jobs. You want to minimize cost. What should you do?

- Enable node auto-provisioning on the GKE cluster.
- Create a VerticalPodAutcaler for those workloads.
- Create a node pool with preemptible VMs and GPUs attached to those VMs.
- **Create a node pool of instances with GPUs, and enable autoscaling on this node pool with a minimum size of 1.**

Unattempted

Enable node auto-provisioning on the GKE cluster. is not right.

Node auto-provisioning automatically manages a set of node pools on the user's behalf.

Without Node auto-provisioning, GKE considers starting new nodes only from the set of user-created node pools. With node auto-provisioning, new node pools can be created and deleted automatically. This in no way helps us with our requirements.

Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/node-auto-provisioning>

Create a VerticalPodAutcaler for those workloads. is not right.

Vertical pod autoscaling (VPA) frees you from having to think about what values to specify for a container's CPU and memory requests. The autoscaler can recommend values for CPU and memory requests and limits, or it can automatically update the values. This doesn't help us with the GPU requirement. Moreover, due to Kubernetes limitations, the only way to modify the resource requests of a running Pod is to recreate the Pod. This has the negative effect of killing the non-restartable jobs which is undesirable.

<https://cloud.google.com/kubernetes-engine/docs/concepts/verticalpodautoscaler#overview>

Create a node pool with preemptible VMs and GPUs attached to those VMs. is not right.

You can use preemptible VMs in your GKE clusters or node pools to run batch or fault-tolerant jobs that are less sensitive to the ephemeral, non-guaranteed nature of preemptible VMs. Whereas we have long-running and non-restartable jobs so preemptible VMs aren't suitable for our requirement.

Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/preemptible-vms>

Create a node pool of instances with GPUs, and enable autoscaling on this node pool with a minimum size of 1. is the right answer.

A node pool is a group of nodes within a cluster that all have the same configuration. Our requirement is GPUs, so we create a node pool with GPU enabled and have the scientist's applications deployed to the cluster and use this node pool. At the same time, you want to minimize cost so you start with 1 instance and scale up as needed. It is important to note that the scale down needs to take into consideration if there are any running jobs otherwise the scale down may terminate the nonrestartable job.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/node-pools>

3. Question

An employee was terminated, but their access to Google Cloud Platform (GCP) was not removed until 2 weeks later. You need to find out this employee accessed any sensitive customer information after their termination. What should you do?

- View System Event Logs in Stackdriver. Search for the user's email as the principal.
- View System Event Logs in Stackdriver. Search for the service account associated with the user.
- **View Data Access audit logs in Stackdriver. Search for the user's email as the principal.**
- View the Admin Activity log in Stackdriver. Search for the service account associated with the user.

Unattempted

View the Admin Activity log in Stackdriver. Search for the service account associated with the user. is not right.

Admin Activity logs do not contain log entries for reading resource data. Admin Activity audit logs contain log entries for API calls or other administrative actions that modify the configuration or metadata of resources.

Ref: <https://cloud.google.com/logging/docs/audit#admin-activity>

View System Event Logs in Stackdriver. Search for the user's email as the principal. is not right.

System Event audit logs do not contain log entries for reading resource data. System Event audit logs contain log entries for Google Cloud administrative actions that modify the configuration of resources. System Event audit logs are generated by Google systems; they are not driven by direct user action.

Ref: <https://cloud.google.com/logging/docs/audit#system-event>

View System Event Logs in Stackdriver. Search for the service account associated with the user. is not right.

System Event audit logs do not contain log entries for reading resource data. System Event audit logs contain log entries for Google Cloud administrative actions that modify the configuration of resources. System Event audit logs are generated by Google systems; they are not driven by direct user action.

Ref: <https://cloud.google.com/logging/docs/audit#system-event>

View Data Access audit logs in Stackdriver. Search for the user's email as the principal. is the right answer.

Data Access audit logs contain API calls that read the configuration or metadata of resources, as well as user-driven API calls that create, modify, or read user-provided resource data.

Ref: <https://cloud.google.com/logging/docs/audit#data-access>

4. Question

An engineer from your team accidentally deployed several new versions of NodeJS application on Google App Engine Standard. You are concerned the new versions are serving traffic. You have been asked to produce a list of all the versions of the application that are receiving traffic as well the percent traffic split between them. What should you do?

- `gcloud app versions list --hide-no-traffic`
- `gcloud app versions list --show-traffic`
- `gcloud app versions list`
- `gcloud app versions list --traffic`

Unattempted

`gcloud app versions list`. is not right

This command lists all the versions of all services that are currently deployed to the App Engine server. While this list includes all versions that are receiving traffic, it also includes versions that are not receiving traffic.

Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/list>

`gcloud app versions list --traffic`. is not right

`gcloud app versions list` command does not support `--traffic` flag.

Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/list>

`gcloud app versions list --show-traffic`. is not right

`gcloud app versions list` command does not support `--show-traffic` flag.

Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/list>

`gcloud app versions list --hide-no-traffic`. is the right answer.

This command correctly lists just the versions that are receiving traffic by hiding versions that do not receive traffic. This is the only command that fits our requirements.

Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/list>

5. Question

An intern joined your team recently and needs access to Google Compute Engine in your sandbox project to explore various settings and spin up compute instances to test features. You have been asked to facilitate this. How should you give your intern access to compute engine without giving more permissions than is necessary?

- Grant Project Editor IAM role for sandbox project.
- Grant Compute Engine Admin Role for sandbox project.
- Create a shared VPC to enable the intern access Compute resources.
- **Grant Compute Engine Instance Admin Role for the sandbox project.**

Unattempted

Create a shared VPC to enable the intern access Compute resources. is not right.

Creating a shared VPC is not sufficient to grant intern access to compute resources.

Shared VPCs are primarily used by organizations to connect resources from multiple projects to a common Virtual Private Cloud (VPC) network, so that they can communicate with each other securely and efficiently using internal IPs from that network.

Ref: <https://cloud.google.com/vpc/docs/shared-vpc>

Grant Project Editor IAM role for sandbox project. is not right.

Project editor role grants all viewer permissions, plus permissions for actions that modify state, such as changing existing resources. While this role lets the intern explore compute engine settings and spin up compute instances, it grants more permissions than is needed. Our intern can modify any resource in the project.

https://cloud.google.com/iam/docs/understanding-roles#primitive_roles

Grant Compute Engine Admin Role for sandbox project. is not right.

Compute Engine Admin Role grants full control of all Compute Engine resources; including networks, load balancing, service accounts etc. While this role lets the intern explore compute engine settings and spin up compute instances, it grants more permissions than what is needed.

Ref: <https://cloud.google.com/compute/docs/access/iam#compute.storageAdmin>

Grant Compute Engine Instance Admin Role for the sandbox project. is the right answer.

Compute Engine Instance Admin Role grants full control of Compute Engine instances, instance groups, disks, snapshots, and images. It also provides read access to all Compute Engine networking resources. This provides just the required permissions to the intern.

Ref: <https://cloud.google.com/compute/docs/access/iam#compute.storageAdmin>

6. Question

Auditors visit your teams every 12 months and ask to review all the Google Cloud Identity and Access Management (Cloud IAM) policy changes in the previous 12 months. You want to streamline and expedite the analysis and audit process. What should you do?

- Enable Logging export to Google Cloud Storage (GCS) bucket and delegate access to the bucket
- **Enable Logging export to Google BigQuery and use ACLs and views to scope the data shared with the auditor**
- Create custom Google Stackdriver alerts and send them to the auditor
- Use Cloud Functions to transfer log entries to Google Cloud SQL and use ACLs and views to limit an auditor's view

Unattempted

Create custom Google Stackdriver alerts and send them in an email to the auditor. is not right.

Stackdriver Alerting gives timely awareness to problems in your cloud applications so you can resolve the problems quickly. Sending alerts to your auditor is not of much use during audits.

Ref: <https://cloud.google.com/monitoring/alerts>

Use Cloud Functions to transfer log entries to Google Cloud SQL and use ACLs and views to limit an auditor's view. is not right.

Using Cloud Functions to transfer log entries to Google Cloud SQL is expensive in comparison to audit logs export feature which exports logs to various destinations with minimal configuration.

Ref: <https://cloud.google.com/logging/docs/export/>

Auditors spend a lot of time reviewing log messages. And you want to expedite the audit process!! So you want to make it easier for the auditor to extract the information easily from the logs.

Between the two remaining options, the only difference is the log export sink destination

Ref: <https://cloud.google.com/logging/docs/export/>

One option exports to Google Cloud Storage (GCS) bucket whereas other exports to BigQuery. Querying information out of files in a bucket is much harder compared to querying information from BigQuery Dataset where it is as simple as running a job or set of jobs to extract just the required information and in the format required. By enabling

the auditor to run jobs in Big Queries, you streamline the log extraction process and the auditor can review the extracted logs much quicker. While as good as the other option (bucket) is, Enable Logging export to Google BigQuery and use ACLs and views to scope the data shared with the auditor is the right answer.

You need to configure log sinks before you can receive any logs, and you can't retroactively export logs that were written before the sink was created.

7. Question

Every employee of your company has a Google account. Your operational team needs to manage a large number of instances on the Compute Engine. Each member of this team needs only administrative access to the servers. Your security team wants to ensure that the deployment of credentials is operationally efficient and must be able to determine who accessed a given instance. What should you do?

- Ask each member of the team to generate a new SSH key pair and to send you their public key. Use a configuration management tool to deploy those keys on each instance.
- Ask each member of the team to generate a new SSH key pair and to add the public key to their Google account. Grant the "compute.osAdminLogin" role to the Google group corresponding to this team.
- Generate a new SSH key pair. Give the private key to each member of your team. Configure the public key in the metadata of each instance.
- Generate a new SSH key pair. Give the private key to each member of your team. Configure the public key as a project-wide public SSH key in your Cloud Platform project and allow project-wide public SSH keys on each instance.

Unattempted

Generate a new SSH key pair. Give the private key to each member of your team. Configure the public key in the metadata of each instance. is not right.
Reuse of a single SSH key pair by all employees is a very bad security practice as auditing becomes very impossible.

Generate a new SSH key pair. Give the private key to each member of your team. Configure the public key as a project-wide public SSH key in your Cloud Platform project and allow project-wide public SSH keys on each instance. is not right.

Reuse of a single SSH key pair by all employees is a very bad security practice as auditing becomes very impossible.

Ask each member of the team to generate a new SSH key pair and to send you their public key. Use a configuration management tool to deploy those keys on each instance. is not right.

While this can be done, it is not operationally efficient. Let's say a user leaves the company, you then have to remove their SSH key from all instances where it has been added (can't be removed at a single place). Similarly, when a user joins the company, you have to add their SSH key to all the instances. This is very tedious and not operationally efficient.

Ask each member of the team to generate a new SSH key pair and to add the public key to their Google account. Grant the "compute.osAdminLogin" role to the Google group corresponding to this team. is the right answer.

By letting users manage their own SSH key pair (and its rotation etc), you delete the operational burden of managing SSH keys to individual users. Secondly, granting compute.osAdminLogin grants the group administrator permissions (as opposed to granting compute.osLogin, which does not grant administrator permissions). Finally, managing provisioning and de-provisioning is as simple as adding or removing the user from the group.

OS Login lets you use Compute Engine IAM roles to efficiently manage SSH access to Linux instances and is an alternative to manually managing instance access by adding and removing SSH keys in the metadata. Before you can manage instance access using IAM roles, you must enable the OS Login feature by setting a metadata key-value pair in your project or in your instance's metadata: enable-oslogin=TRUE. After you enable OS Login on one or more instances in your project, those instances accept connections only from user accounts that have the necessary IAM roles in your project or organization. There are two predefined roles.

?roles/compute.osLogin, which does not grant administrator permissions

?roles/compute.osAdminLogin, which grants administrator permissions

At any point, to revoke user access to instances that are enabled to use OS Login, remove the user roles from that user account

Ref: https://cloud.google.com/compute/docs/instances/managing-instance-access#enable_oslogin

8. Question

For service discovery, you need to associate each of the Compute Engine instances of your VPC with an internal (DNS) record in a custom zone. You want to follow Google recommended practices. What should you do?

- Create a new VPC, block all external traffic with a firewall rule and create 2 Cloud DNS zones - a first zone in the new VPC and a second zone in the main VPC that is forwarding requests to the first Cloud DNS zone. Create records for each instance in the first zone.
- Deploy the BIND DNS server in the VPC, and create a Cloud DNS forwarding zone to forward the DNS requests to BIND. Create records for each instance in the BIND DNS server.
- Create a Cloud DNS zone, set its visibility to private and associate it with your VPC. Create records for each instance in that zone.
- Create your Compute Engine instances with custom hostnames.

Unattempted

Our requirements here are 1. Internal and 2. Custom Zone

Create your Compute Engine instances with custom hostnames. is not right.
This doesn't put them in a custom zone.

Deploy the BIND DNS server in the VPC, and create a Cloud DNS forwarding zone to forward the DNS requests to BIND. Create records for each instance in the BIND DNS server. is not right.

This might be possible but not something Google recommends. The Cloud DNS service offering from Google already offers these features so it is pointless installing a custom DNS server to do that.

Create a new VPC, block all external traffic with a firewall rule and create 2 Cloud DNS zones - a first zone in the new VPC and a second zone in the main VPC that is forwarding requests to the first Cloud DNS zone. Create records for each instance in the first zone. is not right.

This doesn't make any sense, moreover, the two VPCs can't communicate without VPC peering.

Ref: <https://cloud.google.com/dns/docs/overview#concepts>

Create a Cloud DNS zone, set its visibility to private and associate it with your VPC. Create records for each instance in that zone. is the right answer.
You should absolutely do this when you want internal DNS records in a custom zone.

Cloud DNS gives you the option of private zones and internal DNS names.

Ref: <https://cloud.google.com/dns/docs/overview#concepts>

9. Question

In Cloud Shell, your active gcloud configuration is as shown below.

```
$ gcloud config list  
[component_manager]  
disable_update_check = True  
[compute]  
gce_metadata_read_timeout_sec = 5  
zone = europe-west2-a  
[core]  
account = gcp-ace-lab-user@gmail.com  
disable_usage_reporting = False  
project = gcp-ace-lab-266520  
[metrics]
```

environment = devshell

You want to create two compute instances – one in europe-west2-a and another in europe-west2-b. What should you do? (Select 2)

- `gcloud compute instances create instance1 gcloud compute instances create instance2`
- `gcloud compute instances create instance1 gcloud config set compute/zone europe-west2-b gcloud compute instances create instance2`
- `gcloud compute instances create instance1 gcloud compute instances create instance2 --zone=europe-west2-b`
- `gcloud compute instances create instance1 gcloud config set zone europe-west2-b gcloud compute instances create instance2`
- `gcloud compute instances create instance1 gcloud configuration set compute/zone europe-west2-b gcloud compute instances create instance2`

Unattempted

`gcloud compute instances create instance1`
`gcloud compute instances create instance2.` is not right.

The default compute/zone property is set to europe-west2-a in the current gcloud configuration. Executing the two commands above would create two compute instances in the default zone i.e. europe-west2-a which doesn't satisfy our requirement.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/create>

```
gcloud compute instances create instance1
```

```
gcloud config set zone europe-west2-b
```

gcloud compute instances create instance2. is not right.

The approach is right but the syntax is wrong. gcloud config does not have a core/zone property. The syntax for this command is gcloud config set SECTION/PROPERTY VALUE. If SECTION is missing, SECTION is defaulted to core. We are effectively trying to run gcloud config set core/zone europe-west2-b but the core section doesn't have a property called zone, so this command fails.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/set>

```
gcloud compute instances create instance1
```

```
gcloud configuration set compute/zone europe-west2-b
```

gcloud compute instances create instance2. is not right.

Like above, the approach is right but the syntax is wrong. You want to set the default compute/zone property in gcloud configuration to europe-west2-b but it needs to be done via the command gcloud config set and not gcloud configuration set.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/set>

```
gcloud compute instances create instance1
```

```
gcloud config set compute/zone europe-west2-b
```

gcloud compute instances create instance2. is the right answer.

The default compute/zone property is europe-west2-a in the current gcloud configuration so executing the first gcloud compute instances create command creates the instance in europe-west2-a zone. Next, executing the gcloud config set compute/zone europe-west2-b changes the default compute/zone property in default configuration to europe-west2-b. Executing the second gcloud compute instances create command creates a compute instance in europe-west2-b which is what we want.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/set>

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/create>

```
gcloud compute instances create instance1
```

```
gcloud compute instances create instance2 -zone=europe-west2-b. is the right answer.
```

The default compute/zone property is europe-west2-a in the current gcloud configuration so executing the first gcloud compute instances create command creates the instance in europe-west2-a zone. Next, executing the second gcloud compute instances create command with -zone property creates a compute instance in provided zone i.e. europe-west2-b instead of using the default zone from the current active configuration.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/create>

10. Question

In Regional Storage buckets with object versioning enabled, what is the effect of deleting the live version of an object and deleting a noncurrent version of an object?

- 1. The live version becomes a noncurrent version. 2. The noncurrent version is deleted permanently.
- 1. The live version becomes a noncurrent version and a lifecycle rule is applied to delete after 30 days. 2. A lifecycle rule is applied on the noncurrent version to delete after 30 days.
- 1. The live version becomes a noncurrent version and a lifecycle rule is applied to transition to Nearline Storage after 30 days. 2. A lifecycle rule is applied on the noncurrent version to transition to Nearline Storage after 30 days.
- 1. The live version is deleted permanently. 2. The noncurrent version is deleted permanently.

Unattempted

1. The live version becomes a noncurrent version.
2. The noncurrent version is deleted permanently. is the right answer.

In buckets with object versioning enabled, deleting the live version of an object creates a noncurrent version while deleting a noncurrent version deletes that version permanently.

Ref: <https://cloud.google.com/storage/docs/lifecycle#actions>

11. Question

Several employees at your company have been creating projects with Cloud Platform and paying for it with their personal credit cards, which the company reimburses. The company wants to centralize all these projects under a single, new billing account. What should you do?

- Contact cloud-billing@google.com with your bank account details and request a corporate billing account for your company.
- In the Google Cloud Platform Console, create a new billing account and set up a payment method.
- In the Google Platform Console, go to the Resource Manager and move all projects to the root Organization.

- Create a ticket with Google Support and wait for their call to share your credit card details over the phone.

Unattempted

Contact cloud-billing@google.com with your bank account details and request a corporate billing account for your company. is not right.

That is not how we set up billing for the organization.

Ref: <https://cloud.google.com/billing/docs/concepts>

Create a ticket with Google Support and wait for their call to share your credit card details over the phone. is not right.

That is not how we set up billing for the organization.

Ref: <https://cloud.google.com/billing/docs/concepts>

In the Google Cloud Platform Console, create a new billing account and set up a payment method. is not right.

Unless all projects are modified to use the new billing account, this doesn't work.

Ref: <https://cloud.google.com/billing/docs/concepts>

In the Google Platform Console, go to the Resource Manager and move all projects to the root Organization. is the right answer.

If we move all projects under the root organization hierarchy, they need to use a billing account within the root organization. We can then consolidate all the costs under different billing accounts as needed e.g. per project, or one for dev work and another billing account for production usage, etc.

Ref: <https://cloud.google.com/billing/docs/concepts>

12. Question

The storage costs for your application logs have far exceeded the project budget. The logs are currently being retained indefinitely in the Cloud Storage bucket myapp-gcp-ace-logs. You have been asked to remove logs older than 90 days from your Cloud Storage bucket. You want to optimize ongoing Cloud Storage spend. What should you do?

- Write a script that runs `gsutil ls -l gs://myapp-gcp-ace-logs/**` to find and remove items older than 90 days. Schedule the script with cron.
- Write a script that runs `gsutil ls -lr gs://myapp-gcp-ace-logs/**` to find and remove items older than 90 days. Repeat this process every morning.

- Write a lifecycle management rule in XML and push it to the bucket with `gsutil lifecycle set config-xml-file`.
- Write a lifecycle management rule in JSON and push it to the bucket with `gsutil lifecycle set config-json-file`.

Unattempted

You write a lifecycle management rule in XML and push it to the bucket with `gsutil lifecycle set config-xml-file`. is not right.

`gsutil lifecycle set` enables you to set the lifecycle configuration on one or more buckets based on the configuration file provided. However, XML is not a valid supported type for the configuration file.

Ref: <https://cloud.google.com/storage/docs/gsutil/commands/lifecycle>

Write a script that runs `gsutil ls -lr gs://myapp-gcp-ace-logs/**` to find and remove items older than 90 days. Repeat this process every morning. is not right.

This manual approach is error-prone, time-consuming and expensive. GCP Cloud Storage provides lifecycle management rules that let you achieve this with minimal effort.

Write a script that runs `gsutil ls -l gs://myapp-gcp-ace-logs/**` to find and remove items older than 90 days. Schedule the script with cron. is not right.

This manual approach is error-prone, time-consuming and expensive. GCP Cloud Storage provides lifecycle management rules that let you achieve this with minimal effort.

Write a lifecycle management rule in JSON and push it to the bucket with `gsutil lifecycle set config-json-file`. is the right answer.

You can assign a lifecycle management configuration to a bucket. The configuration contains a set of rules which apply to current and future objects in the bucket. When an object meets the criteria of one of the rules, Cloud Storage automatically performs a specified action on the object. One of the supported actions is to Delete objects. You can set up a lifecycle management to delete objects older than 90 days. “`gsutil lifecycle set`” enables you to set the lifecycle configuration on the bucket based on the configuration file. JSON is the only supported type for the configuration file. The `config-json-file` specified on the command line should be a path to a local file containing the lifecycle configuration JSON document.

Ref: <https://cloud.google.com/storage/docs/gsutil/commands/lifecycle>

Ref: <https://cloud.google.com/storage/docs/lifecycle>

13. Question

Users of your application are complaining of slowness when loading the application. You realize the slowness is because the App Engine deployment serving the application is deployed in us-central whereas all users of this application are closest to europe-west3. You want to change the region of the App Engine application to europe-west3 to minimize latency. What's the best way to change the App Engine region?

- Create a new project and create an App Engine instance in europe-west3
- Contact Google Cloud Support and request the change.
- Use the gcloud app region set command and supply the name of the new region.
- From the console, under the App Engine page, click edit, and change the region drop-down.

Unattempted

Use the gcloud app region set command and supply the name of the new region. is not right.

gcloud app region command does not provide a set action. The only action gcloud app region command currently supports is list which lists the availability of flex and standard environments for each region.

Ref: <https://cloud.google.com/sdk/gcloud/reference/app/regions/list>

Contact Google Cloud Support and request the change. is not right.

Unfortunately, Google Cloud Support isn't of much use here as they would not be able to change the region of an App Engine Deployment. App engine is a regional service, which means the infrastructure that runs your app(s) is located in a specific region and is managed by Google to be redundantly available across all the zones within that region. Once an app engine deployment is created in a region, it can't be changed.

Ref: <https://cloud.google.com/appengine/docs/locations>

From the console, Click edit in App Engine dashboard page and change the region drop-down. is not right.

The settings mentioned in this option aren't available in the App Engine dashboard. App engine is a regional service. Once an app engine deployment is created in a region, it can't be changed. As shown in the screenshot below, Region is greyed out.

Create a new project and create an App Engine instance in europe-west3. is the right answer.

App engine is a regional service, which means the infrastructure that runs your app(s) is located in a specific region and is managed by Google to be redundantly available across all the zones within that region. Once an app engine deployment is created in a region, it

can't be changed. The only way is to create a new project and create an App Engine instance in europe-west3, send all user traffic to this instance and delete the app engine instance in us-central.

Ref: <https://cloud.google.com/appengine/docs/locations>

14. Question

You are analyzing Google Cloud Platform service costs from three separate projects. You want to use the information to create service costs estimates grouped by service type, daily and monthly, for the next six months using standard query syntax. What should you do?

- Export your bill to a BigQuery dataset and then write time window based SQL queries for analysis.
- Export your bill to a Cloud Storage bucket and then import into Cloud Bigtable for analysis.
- Export your bill to a Cloud Storage bucket and then import into Google Sheets for analysis
- Export your transactions to a local file and perform analysis with a suitable desktop tool.

Unattempted

Requirements

1. use query syntax
2. need the billing data of all three projects

Export your bill to a Cloud Storage bucket and then import into Cloud Bigtable for analysis. is not right.

BigTable is a NoSQL database and doesn't offer query syntax support.

Export your bill to a Cloud Storage bucket and then import into Google Sheets for analysis. is not right.

Google Sheets don't offer full support for query syntax. Moreover, export to Cloud Storage bucket captures a smaller dataset than export to BigQuery. For example, the exported billing data does not include resource labels or any invoice-level charges such as taxes accrued or adjustment memos.

Export your transactions to a local file and perform analysis with a suitable desktop tool. is not right.

Billing data can't be exported to a local file, it can only be exported to a BigQuery Dataset or Cloud Storage bucket.

Export your bill to a BigQuery dataset and then write time window based SQL queries for analysis. is the right answer.

You can export billing information from multiple projects into a BigQuery dataset. Unlike the export to Cloud Storage bucket, export to BigQuery dataset includes all information making it easy and straightforward to construct queries in BigQuery to estimate the cost. BigQuery supports Standard SQL so you can join tables and group by fields (labels in this case) as needed

Ref: <https://cloud.google.com/billing/docs/how-to/export-data-bigquery>.

15. Question

You are building a new version of an application hosted in an App Engine environment.

You want to test the new version with 1% of users before you completely switch your application over to the new version. What should you do?

- Deploy a new version of your application in Google Kubernetes Engine instead of App Engine and then use GCP Console to split traffic.
- Deploy a new version of your application in a Compute Engine instance instead of App Engine and then use GCP Console to split traffic.
- Deploy a new version as a separate app in App Engine. Then configure App Engine using GCP Console to split traffic between the two apps.
- Deploy a new version of your application in App Engine. Then go to App Engine settings in GCP Console and split traffic between the current version and newly deployed versions accordingly.

Unattempted

Deploy a new version of your application in Google Kubernetes Engine instead of App Engine and then use GCP Console to split traffic. is not right.

When you can achieve this natively in GCP app engine using versions, there is no need to do it outside App Engine.

Ref: <https://cloud.google.com/appengine/docs/standard/python/splitting-traffic>

Deploy a new version of your application in a Compute Engine instance instead of App Engine and then use GCP Console to split traffic. is not right.

When you can achieve this natively in GCP app engine using versions, there is no need to do it outside App Engine.

Ref: <https://cloud.google.com/appengine/docs/standard/python/splitting-traffic>

Deploy a new version as a separate app in App Engine. Then configure App Engine using GCP Console to split traffic between the two apps. is not right.

You can achieve this natively in GCP app engine using versions but App Engine doesn't let you split traffic between apps. If you need to do it between apps, you are probably looking at doing this at the load balancer layer or at the DNS layer – either increasing the cost/complexity or introduce other problems such as caching issues.

Ref: <https://cloud.google.com/appengine/docs/standard/python/splitting-traffic>

Deploy a new version of your application in App Engine. Then go to App Engine settings in GCP Console and split traffic between the current version and newly deployed versions accordingly. is the right answer.

GCP App Engine natively offers traffic splitting functionality between versions. You can use traffic splitting to specify a percentage distribution of traffic across two or more of the versions within a service. Splitting traffic allows you to conduct A/B testing between your versions and provides control over the pace when rolling out features.

Ref: <https://cloud.google.com/appengine/docs/standard/python/splitting-traffic>

16. Question

You are building a pipeline to process time-series data. Which Google Cloud Platform services should you put in boxes 1,2,3, and 4?

Larger image

- Firebase Messages, Cloud Pub/Sub, Cloud Spanner, BigQuery
- Cloud Pub/Sub, Cloud Storage, BigQuery, Cloud Bigtable
- Cloud Pub/Sub, Cloud Dataflow, Cloud Datastore, BigQuery
- **Cloud Pub/Sub, Cloud Dataflow, Cloud Bigtable, BigQuery**

Unattempted

For box 1 where you want to ingest time series data, your best bet is Cloud Pub/Sub. For box 2 where you want to process the data in pipelines, your best bet is Cloud Dataflow.

That leaves us with two remaining options, both have BigQuery as no 4. For (storage) 3, it is a choice between Bigtable and Datastore. Bigtable provides out of the box support for time series data. So using Bigtable for Storage is the right answer.

Ref: <https://cloud.google.com/bigtable/docs/schema-design-time-series>

The answer is Cloud Pub/Sub, Cloud Dataflow, Cloud Bigtable, BigQuery

17. Question

You are building a product on top of Google Kubernetes Engine (GKE). You have a single GKE cluster. For each of your customers, a Pod is running in that cluster, and your customers can run arbitrary code inside their Pod. You want to maximize the isolation between your customers' Pods. What should you do?

- Use Binary Authorization and whitelist only the container images used by your customers' Pods.
- Use the Container Analysis API to detect vulnerabilities in the containers used by your customers' Pods.
- Create a GKE node pool with a sandbox type configured to gvisor. Add the parameter runtimeClassName: gvisor to the specification of your customers' Pods.
- Use the cos_containerd image for your GKE nodes. Add a nodeSelector with the value cloud.google.com/gke-os-distribution: cos_containerd to the specification of your customers' Pods.

Unattempted

Use Binary Authorization and whitelist only the container images used by your customers' Pods. is not right.

Binary Authorization is a deploy-time security control that ensures only trusted container images are deployed on Google Kubernetes Engine (GKE). With Binary Authorization, you can require images to be signed by trusted authorities during the development process and then enforce signature validation when deploying. By enforcing validation, you can gain tighter control over your container environment by ensuring only verified images are integrated into the build-and-release process.

Ref: <https://cloud.google.com/binary-authorization>

Use the Container Analysis API to detect vulnerabilities in the containers used by your customers' Pods. is not right.

Container Analysis is a service that provides vulnerability scanning and metadata storage

for software artifacts. The scanning service performs vulnerability scans on images in Container Registry, then stores the resulting metadata and makes it available for consumption through an API. Metadata storage allows storing information from different sources, including vulnerability scanning, other Cloud services, and third-party providers.

Ref: <https://cloud.google.com/container-registry/docs/container-analysis>

Use the `cos_containerd` image for your GKE nodes. Add a `nodeSelector` with the value `cloud.google.com/gke-os-distribution: cos_containerd` to the specification of your customers' Pods. is not right.

The `cos_containerd` and `ubuntu_containerdimages` let you use containerd as the container runtime in your GKE cluster. This doesn't directly provide the isolation we require.

<https://cloud.google.com/kubernetes-engine/docs/concepts/using-containerd>

Create a GKE node pool with a `sandbox` type configured to `gvisor`. Add the parameter `runtimeClassName: gvisor` to the specification of your customers' Pods. is the right answer. GKE Sandbox provides an extra layer of security to prevent untrusted code from affecting the host kernel on your cluster nodes when containers in the Pod execute unknown or untrusted code. Multi-tenant clusters and clusters whose containers run untrusted workloads are more exposed to security vulnerabilities than other clusters. Examples include SaaS providers, web-hosting providers, or other organizations that allow their users to upload and run code. When you enable GKE Sandbox on a node pool, a sandbox is created for each Pod running on a node in that node pool. In addition, nodes running sandboxed Pods are prevented from accessing other Google Cloud services or cluster metadata. Each sandbox uses its own userspace kernel. With this in mind, you can make decisions about how to group your containers into Pods, based on the level of isolation you require and the characteristics of your applications.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/sandbox-pods>

18. Question

You are building an application that stores relational data from users. Users across the globe will use this application. Your CTO is concerned about the scaling requirements because the size of the user base is unknown. You need to implement a database solution that can scale with your user growth with minimum configuration changes. Which storage solution should you use?

- Cloud Datastore
- Cloud SQL

- Cloud Spanner
- Cloud Firestore

Unattempted

Our requirements are relational data, global users, scaling

Cloud Firestore is not right.

Cloud Firestore is not a relational database. Cloud Firestore is a flexible, scalable database for mobile, web, and server development from Firebase and Google Cloud Platform.

Ref: <https://firebase.google.com/docs/firestore>

Cloud Datastore is not right.

Cloud Datastore is not a relational database. Datastore is a NoSQL document database built for automatic scaling, high performance, and ease of application development

Ref: <https://cloud.google.com/datastore/docs/concepts/overview>

Cloud SQL is not right.

While Cloud SQL is a relational database, it does not offer infinite automated scaling with minimum configuration changes. Cloud SQL is a fully-managed database service that makes it easy to set up, maintain, manage, and administer your relational databases on Google Cloud Platform

Ref: <https://cloud.google.com/sql/docs>

Cloud Spanner is the right answer.

Cloud Spanner is a relational database and is highly scalable. Cloud Spanner is a highly scalable, enterprise-grade, globally-distributed, and strongly consistent database service built for the cloud specifically to combine the benefits of relational database structure with a non-relational horizontal scale. This combination delivers high-performance transactions and strong consistency across rows, regions, and continents with an industry-leading 99.999% availability SLA, no planned downtime, and enterprise-grade security

<https://cloud.google.com/spanner>

19. Question

You are building an application that will run in your data center. The application will use Google Cloud Platform (GCP) services like AutoML. You created a service account that has appropriate access to AutoML. You need to enable authentication to the APIs from your on-premises environment. What should you do?

- Use service account credentials in your on-premises application.
- Use gcloud to create a key file for the service account that has appropriate permissions.
- Set up direct interconnect between your data center and Google Cloud Platform to enable authentication for your on-premises applications.
- Go to the IAM & admin console, grant a user account permissions similar to the service account permissions, and use this user account for authentication from your data center.

Unattempted

Use service account credentials in your on-premises application. is not right.

Service accounts do not have passwords

Ref: <https://cloud.google.com/iam/docs/service-accounts>

Go to the IAM & admin console, grant a user account permissions similar to the service account permissions, and use this user account for authentication from your data center. is not right.

While granting Users a similar set of permissions lets them impersonate service accounts and access all resources the service account has access to, you should use a service account to represent a non-human user that needs to authenticate and be authorized to access data in Google APIs. Typically, service accounts are used in scenarios such as:

Running workloads on virtual machines (VMs).

Running workloads on on-premises workstations or data centers that call Google APIs.

Running workloads that are not tied to the lifecycle of a human user.

Your application assumes the identity of the service account to call Google APIs so that the users aren't directly involved.

Ref: <https://cloud.google.com/iam/docs/understanding-service-accounts>

Set up direct interconnect between your data center and Google Cloud Platform to enable authentication for your on-premises applications. is not right.

While setting up interconnect provides a direct physical connection between your on-premises network and Google's network, it doesn't directly help us authenticate our application running in the data center. You can configure Private Google Access for on-premises hosts by sending requests to restricted.googleapis.com and advertise a custom route on cloud router but this only lets you reach Google API and doesn't help with authentication.

Ref: <https://cloud.google.com/interconnect/docs/support/faq>

Use gcloud to create a key file for the service account that has appropriate permissions. is the right answer.

To use a service account outside of Google Cloud, such as on other platforms or on-

premises, you must first establish the identity of the service account. Public/private key pairs provide a secure way of accomplishing this goal. You can create a service account key using the Cloud Console, the gcloud tool, the serviceAccounts.keys.create() method, or one of the client libraries.

Ref: <https://cloud.google.com/iam/docs/creating-managing-service-account-keys>

20. Question

You are building an archival solution for your data warehouse and have selected Cloud Storage to archive your data. Your users need to be able to access this archived data once a quarter for some regulatory requirements. You want to select a cost-efficient option. Which storage option should you use?

- Coldline Storage
- Nearline Storage
- Regional Storage
- Multi-Regional Storage

Unattempted

Nearline Storage. is not right.

Nearline Storage is a low-cost, highly durable storage service for storing infrequently accessed data. Nearline Storage is ideal for data you plan to read or modify on average once per month or less. Nearline storage is more expensive than Coldline Storage which is more suitable for our requirements.

<https://cloud.google.com/storage/docs/storage-classes#nearline>

Regional Storage. is not right.

While this would certainly let you access your files once a quarter, it would be too expensive compared to Coldline storage which is more suitable for our requirement.

<https://cloud.google.com/storage/docs/storage-classes#standard>

Multi-Regional Storage. is not right.

While this would certainly let you access your files once a quarter, it would be too expensive compared to Coldline storage which is more suitable for our requirement.

<https://cloud.google.com/storage/docs/storage-classes#standard>

Coldline Storage. is the right answer.

Coldline Storage is a very-low-cost, highly durable storage service for storing infrequently

accessed data. Coldline Storage is ideal for data you plan to read or modify at most once a quarter. Since we have a requirement to access data once a quarter and want to go with the most cost-efficient option, we should select Coldline Storage.

Ref: <https://cloud.google.com/storage/docs/storage-classes#coldline>

21. Question

You are configuring service accounts for an application that spans multiple projects. Virtual machines (VMs) running in the web-applications project need access to BigQuery datasets in crm-databases project. You want to follow Google-recommended practices to give access to the service account in the web-applications project. What should you do?

- Grant project owner role on web-applications project to the service account in crm-databases project.
- Grant project owner role on crm-databases project to the service account in web-applications project.
- Grant project owner role on crm-databases project and bigquery.dataViewer role to the service account in web-applications.
- **Grant bigquery.dataViewer role on crm-databases project to the service account in web-applications.**

Unattempted

Grant project owner role on web-applications project to the service account in crm-databases project. is not right.

Our requirement is to identify the access needed for service account in the web-applications project, not the service account in crm-databases project

Grant project owner role on crm-databases project to the service account in web-applications project. is not right.

The primitive project owner role provides permissions to manage all resources within the project. For this scenario, the service account in the web-applications project needs access to BigQuery datasets in crm-databases project. Granting the project owner role would fall foul of least privilege principle.

Ref: <https://cloud.google.com/iam/docs/recommender-overview>

Grant project owner role on crm-databases project and bigquery.dataViewer role to the service account in web-applications. is not right.

The primitive project owner role provides permissions to manage all resources within the

project. For this scenario, the service account in the web-applications project needs access to BigQuery datasets in crm-databases project. Granting the project owner role would fall foul of least privilege principle.

Ref: <https://cloud.google.com/iam/docs/recommender-overview>

Grant `bigrquery.dataViewer` role on crm-databases project to the service account in web-applications. is the right answer.

`bigrquery.dataViewer` role provides permissions to read the dataset's metadata and list tables in the dataset as well as Read data and metadata from the dataset's tables. This is exactly what we need to fulfil this requirement and follows the least privilege principle.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#bigrquery-roles>

22. Question

You are creating a Google Kubernetes Engine (GKE) cluster with a cluster autoscaler feature enabled. You need to make sure that each node of the cluster will run a monitoring pod that sends container metrics to a third-party monitoring solution. What should you do?

- Deploy the monitoring pod in a StatefulSet object.
- Reference the monitoring pod in a Deployment object.
- Reference the monitoring pod in a cluster initializer at the GKE cluster creation time.
- Deploy the monitoring pod in a DaemonSet object.

Unattempted

Reference the monitoring pod in a Deployment object. is not right.

In our scenario, we need just 1 instance of the monitoring pod running on each node. Bundling the monitoring pod with a deployment object may result in multiple pod instances on the same node. In GKE, deployments represent a set of multiple, identical Pods with no unique identities. Deployment runs multiple replicas of your application and automatically replaces any instances that fail or become unresponsive. In this way, Deployments help ensure that one or more instances of your application are available to serve user requests.

<https://cloud.google.com/kubernetes-engine/docs/concepts/deployment>

Reference the monitoring pod in a cluster initializer at the GKE cluster creation time. is not right.

You can not use gcloud init to initialize a monitoring pod. gcloud initializer performs the following setup steps.

? Authorizes gcloud and other SDK tools to access Google Cloud Platform using your user account credentials, or from an account of your choosing whose credentials are already available.

? Sets up a new or existing configuration.

? Sets properties in that configuration, including the current project and optionally, the default Google Compute Engine region and zone you'd like to use.

Ref: <https://cloud.google.com/sdk/gcloud/reference/init>

Deploy the monitoring pod in a StatefulSet object. is not right.

In GKE, StatefulSets represents a set of Pods with unique, persistent identities and stable hostnames that GKE maintains regardless of where they are scheduled. The state information and other resilient data for any given StatefulSet Pod is maintained in persistent disk storage associated with the StatefulSet. The main purpose of StatefulSets is to set up persistent storage for pods that are deployed across multiple zones.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/statefulset>

Although persistent volumes can be used, they are limited to two zones and you'd have to get into node affinity if you want to use a persistent volume with a pod on a zone that is not covered by the persistent volumes zones.

See this for more information <https://kubernetes.io/docs/setup/best-practices/multiple-zones/>

Deploy the monitoring pod in a DaemonSet object. is the right answer.

In GKE, DaemonSets manage groups of replicated Pods and adhere to a one-Pod-per-node model, either across the entire cluster or a subset of nodes. As you add nodes to a node pool, DaemonSets automatically add Pods to the new nodes as needed. So, this is a perfect fit for our monitoring pod.

<https://cloud.google.com/kubernetes-engine/docs/concepts/daemonset>

DaemonSets are useful for deploying ongoing background tasks that you need to run on all or certain nodes, and which do not require user intervention. Examples of such tasks include storage daemons like ceph, log collection daemons like fluentd, and node monitoring daemons like collectd. For example, you could have DaemonSets for each type of daemon run on all of your nodes. Alternatively, you could run multiple DaemonSets for a single type of daemon, but have them use different configurations for different hardware types and resource needs.

23. Question

You are deploying an application to a Compute Engine VM in a managed instance group. The application must be running at all times, but only a single instance of the VM should run per GCP project. How should you configure the instance group?

- Set autoscaling to On, set the minimum number of instances to 1, and then set the maximum number of instances to 2.
- Set autoscaling to On, set the minimum number of instances to 1, and then set the maximum number of instances to 1.
- Set autoscaling to Off, set the minimum number of instances to 1, and then set the maximum number of instances to 1.
- Set autoscaling to Off, set the minimum number of instances to 1, and then set the maximum number of instances to 2.

Unattempted

Requirements

1. Since we need the application running at all times, we need a minimum 1 instance.
2. Only a single instance of the VM should run, we need a maximum 1 instance.
3. We want the application running at all times. If the VM crashes due to any underlying hardware failure, we want another instance to be added to MIG so that application can continue to serve requests. We can achieve this by enabling autoscaling.

The only option that satisfies these three is Set autoscaling to On, set the minimum number of instances to 1, and then set the maximum number of instances to 1.

Ref: <https://cloud.google.com/compute/docs/autoscaler>

24. Question

You are deploying an application to the App Engine. You want the number of instances to scale based on request rate. You need at least 3 unoccupied instances at all times. Which scaling type should you use?

- Basic Scaling with min_instances set to 3.
- Manual Scaling with 3 instances.
- Automatic Scaling with min_idle_instances set to 3.
- Basic Scaling with max_instances set to 3.

Unattempted

Manual Scaling with 3 instances. is not right.

Manual scaling uses resident instances that continuously run the specified number of instances regardless of the load level. This allows tasks such as complex initializations and applications that rely on the state of the memory over time. This does not autoscale based on the request rate so doesn't fit our requirements.

Ref: <https://cloud.google.com/appengine/docs/standard/python/how-instances-are-managed>

Basic Scaling with min_instances set to 3. is not right.

Basic scaling creates dynamic instances when your application receives requests. Each instance will be shut down when the app becomes idle. Basic scaling is ideal for work that is intermittent or driven by user activity. In absence of any load, the App engine may shut down all instances so it is not suitable for our requirement of "at least 3 instances at all times".

Ref: <https://cloud.google.com/appengine/docs/standard/python/how-instances-are-managed>

Basic Scaling with max_instances set to 3. is not right.

Basic scaling creates dynamic instances when your application receives requests. Each instance will be shut down when the app becomes idle. Basic scaling is ideal for work that is intermittent or driven by user activity. In absence of any load, the App engine may shut down all instances so it is not suitable for our requirement of "at least 3 instances at all times".

Ref: <https://cloud.google.com/appengine/docs/standard/python/how-instances-are-managed>

Automatic Scaling with min_idle_instances set to 3. is the right answer.

Automatic scaling creates dynamic instances based on request rate, response latencies, and other application metrics. However, if you specify the number of minimum idle instances, that specified number of instances run as resident instances while any additional instances are dynamic.

Ref: <https://cloud.google.com/appengine/docs/standard/python/how-instances-are-managed>

25. Question

You are designing an application that lets users upload and share photos. You expect your application to grow really fast and you are targeting a worldwide audience. You want to

delete uploaded photos after 30 days. You want to minimize costs while ensuring your application is highly available. Which GCP storage solution should you choose?

- Persistent SSD on VM instances.
- Cloud Filestore.
- **Multiregional Cloud Storage bucket.**
- Cloud Datastore database.

Unattempted

Cloud Datastore database. is not right.

Cloud Datastore is a NoSQL document database built for automatic scaling, high performance, and ease of application development. We want to store objects/files and Cloud Datastore is not a suitable storage option for such data.

Ref: <https://cloud.google.com/datastore/docs/concepts/overview>

Cloud Filestore. is not right.

Cloud Filestore is a managed file storage service based on NFSv3 protocol. While Cloud Filestore can be used to store images, Cloud Filestore is a zonal service and can not scale easily to support a worldwide audience. Also, Cloud Filestore costs a lot (10 times) more than some of the storage classes offered by Google Cloud Storage.

Ref: <https://cloud.google.com/filestore>, Ref: <https://cloud.google.com/storage/pricing>

Persistent SSD on VM instances. is not right.

Persistent SSD is a regional service and doesn't automatically scale to other regions to support a worldwide user base. Moreover, Persistent SSD disks are very expensive. A regional persistent SSD costs \$0.34 per GB per month. In comparison, Google Cloud Storage offers several storage classes that are significantly cheaper.

Ref: <https://cloud.google.com/persistent-disk>

Ref: <https://cloud.google.com/filestore/pricing>

Multiregional Cloud Storage bucket. is the right answer.

Cloud Storage allows world-wide storage and retrieval of any amount of data at any time. We don't need to set up auto-scaling ourselves. Cloud Storage autoscaling is managed by GCP. Cloud Storage is an object store so it is suitable for storing photos. Cloud Storage allows world-wide storage and retrieval so cater well to our worldwide audience. Cloud storage provides us lifecycle rules that can be configured to automatically delete objects older than 30 days. This also fits our requirements. Finally, Google Cloud Storage offers several storage classes such as Nearline Storage (\$0.01 per GB per Month) Coldline Storage (\$0.007 per GB per Month) and Archive Storage (\$0.004 per GB per month) which are significantly cheaper than any of the options above.

Ref: <https://cloud.google.com/storage/docs>
Ref: <https://cloud.google.com/storage/pricing>

26. Question

You are designing an application that uses WebSockets and HTTP sessions that are not distributed across the web servers. You want to ensure the application runs properly on Google Cloud Platform. What should you do?

- Meet with the cloud enablement team to discuss load balancer options.
- Redesign the application to use a distributed user session service that does not rely on WebSockets and HTTP sessions.
- Review the encryption requirements for WebSocket connections with the security team.
- Convert the WebSocket code to use HTTP streaming.

Unattempted

Google HTTP(S) Load Balancing has native support for the WebSocket protocol when you use HTTP or HTTPS, not HTTP/2, as the protocol to the backend.

Ref: https://cloud.google.com/load-balancing/docs/https#websocket_proxy_support

So the next possible step is to Meet with the cloud enablement team to discuss load balancer options.

We don't need to convert WebSocket code to use HTTP streaming or Redesign the application, as WebSocket support is offered by Google HTTP(S) Load Balancing. Reviewing the encryption requirements is a good idea but it has nothing to do with WebSockets.

27. Question

You are given a project with a single virtual private cloud (VPC) and a single subnet in the us-central1 region. There is a Compute Engine instance hosting an application in this subnet. You need to deploy a new instance in the same project in the europe-west1 region. This new instance needs access to the application. You want to follow Google-recommended practices. What should you do?

- 1. Create a VPC and a subnet in europe-west1. 2. Expose the application with an internal load balancer. 3. Create the new instance in the new subnet and use the load balancer's address as the endpoint.
- 1. Create a VPC and a subnet in europe-west1. 2. Peer the 2 VPCs. 3. Create the new instance in the new subnet and use the first instance's private address as the endpoint.
- 1. Create a subnet in the same VPC, in europe-west1. 2. Create the new instance in the new subnet and use the first instance subnet's private address as the endpoint.
- 1. Create a subnet in the same VPC, in europe-west1. 2. Use Cloud VPN to connect the two subnets. 3. Create the new instance in the new subnet and use the first instance's private address as the endpoint.

Unattempted

Our requirements are to connect the instance in europe-west1 region with the application running in us-central1 region following Google-recommended practices. The two instances are in the same project.

1. Create a VPC and a subnet in europe-west1.
2. Expose the application with an internal load balancer.
3. Create the new instance in the new subnet and use the load balancer's address as the endpoint. is not right.

We have two different VPCs. There is no mention of the CIDR range so let's assume the two subnets in two VPCs use different CIDR ranges. However, there is no communication route between the two VPCs. If we create an internal load balancer, that load balancer is not visible outside the VPC. So the new instance cannot connect to the load balancer's internal address.

Ref: <https://cloud.google.com/load-balancing/docs/internal>

1. Create a subnet in the same VPC, in europe-west1.
2. Use Cloud VPN to connect the two subnets.
3. Create the new instance in the new subnet and use the first instance's private address as the endpoint. is not right.

Cloud VPN securely connects your on-premises network to your Google Cloud (GCP) Virtual Private Cloud (VPC) network through an IPsec VPN connection. It is not meant to connect two subnets within the same VPC. Moreover, subnets within the same VPC can communicate with each other by setting up relevant firewall rules.

1. Create a VPC and a subnet in europe-west1.
2. Peer the 2 VPCs.
3. Create the new instance in the new subnet and use the first instance's private address as the endpoint. is not right.

Given that the new instance wants to access the application on the existing compute engine instance, these applications seem to be related so they should be within the same VPC. It is possible to have them in different VPCs and peer the VPCs but this is a lot of additional work and we can simplify this by choosing the option below (which is the answer)

1. Create a subnet in the same VPC, in europe-west1.
2. Create the new instance in the new subnet and use the first instance's private address as the endpoint. is the right answer.

We can create another subnet in the same VPC and this subnet is located in europe-west1.

We can then spin up a new instance in this subnet. We also have to set up a firewall rule to allow communication between the two subnets. All instances in the two subnets with the same VPC can communicate through the internal IP Address

Ref: <https://cloud.google.com/vpc>

28. Question

You are hosting an application on bare metal servers in your data center. The application needs access to Cloud Storage. However, security policies prevent the servers hosting the application from having public IP addresses or access to the internet. You want to follow Google recommended practices to provide the application with access to Cloud Storage. What should you do?

- Use nslookup to get the IP addresses for storage.googleapis.com Negotiate with the security team to be able to give public IP addresses to the servers. Only allow egress traffic from those servers to the IP addresses for storage.googleapis.com
- Using Cloud VPN, create a VPN tunnel to a Virtual Private Cloud (VPC) in Google Cloud Platform (GCP). In this VPC, create a Compute Engine instance and install the Squid proxy server on this instance. Configure your servers to use that instance as a proxy to access cloud storage
- Use Migrate for Compute Engine (formerly known as Velostrata) to migrate these servers to Compute Engine. Create an internal load balancer (ILB) that uses storage.googleapis.com as backend. Configure your new instances to use the ILB as a proxy
- Using Cloud VPN or Interconnect, create a tunnel to a VPC in GCP. Using Cloud Router to create a custom route advertisement for 199.36.153.4/30. Announce that network to your on-premises network through the VPN tunnel. In your on-premises

network, configure your DNS server to resolve *.googleapis.com as a CNAME to restricted.googleapis.com

Unattempted

Our requirement is to follow Google recommended practices to achieve the end result.

Configuring Private Google Access for On-Premises Hosts is best achieved by VPN/Interconnect + Advertise Routes + Use restricted Google IP Range.

Using Cloud VPN or Interconnect, create a tunnel to a VPC in GCP

Using Cloud Router to create a custom route advertisement for 199.36.153.4/30.

Announce that network to your on-premises network through the VPN tunnel.

In your on-premises network, configure your DNS server to resolve *.googleapis.com as a CNAME to restricted.googleapis.com is the right answer right, and it is what Google recommends.

Ref: <https://cloud.google.com/vpc/docs/configure-private-google-access-hybrid>

“You must configure routes so that Google API traffic is forwarded through your Cloud VPN or Cloud Interconnect connection, firewall rules on your on-premises firewall to allow the outgoing traffic, and DNS so that traffic to Google APIs resolves to the IP range you’ve added to your routes.”

“You can use Cloud Router Custom Route Advertisement to announce the Restricted Google APIs IP addresses through Cloud Router to your on-premises network. The Restricted Google APIs IP range is 199.36.153.4/30. While this is technically a public IP range, Google does not announce it publicly. This IP range is only accessible to hosts that can reach your Google Cloud projects through internal IP ranges, such as through a Cloud VPN or Cloud Interconnect connection.”

Without having a public IP address or access to the internet, the only way you could connect to cloud storage is if you have an internal route to it. So Negotiate with the security team to be able to give public IP addresses to the servers is not right.

Following “Google recommended practices” is synonymous with “using Google’s services” (Not quite, but it is – at least for the exam !!). So In this VPC, create a Compute Engine instance and install the Squid proxy server on this instance is not right.

Migrating the VM to Compute Engine is a bit drastic when Google says it is perfectly fine to have Hybrid Connectivity architectures <https://cloud.google.com/hybrid-connectivity>. So, Use Migrate for Compute Engine (formerly known as Velostrata) to migrate these servers to Compute Engine is not right.

29. Question

You are managing several Google Cloud Platform (GCP) projects and need access to all logs for the past 60 days. You want to be able to explore and quickly analyze the log contents. You want to follow Google-recommended practices to obtain the combined logs for all projects. What should you do?

- Navigate to Stackdriver Logging and select resource.labels.project_id="*"
- Create a Stackdriver Logging Export with a Sink destination to a BigQuery dataset. Configure the table expiration to 60 days.
- Create a Stackdriver Logging Export with a Sink destination to Cloud Storage. Create a lifecycle rule to delete objects after 60 days.
- Configure a Cloud Scheduler job to read from Stackdriver and store the logs in BigQuery. Configure the table expiration to 60 days.

Unattempted

Navigate to Stackdriver Logging and select resource.labels.project_id="*". is not right.
Log entries are held in Stackdriver Logging for a limited time known as the retention period – which is 30 days (default configuration). After that, the entries are deleted. To keep log entries longer, you need to export them outside of Stackdriver Logging by configuring log sinks.

<https://cloud.google.com/blog/products/gcp/best-practices-for-working-with-google-cloud-audit-logging>

Configure a Cloud Scheduler job to read from Stackdriver and store the logs in BigQuery. Configure the table expiration to 60 days. is not right.

While this works, it makes no sense to use Cloud Scheduler job to read from Stackdriver and store the logs in BigQuery when Google provides a feature (export sinks) that does exactly the same thing and works out of the box.

Ref: https://cloud.google.com/logging/docs/export/configure_export_v2

Create a Stackdriver Logging Export with a Sink destination to Cloud Storage. Create a lifecycle rule to delete objects after 60 days. is not right.

You can export logs by creating one or more sinks that include a logs query and an export destination. Supported destinations for exported log entries are Cloud Storage, BigQuery, and Pub/Sub.

Ref: https://cloud.google.com/logging/docs/export/configure_export_v2

Sinks are limited to exporting log entries from the exact resource in which the sink was created: a Google Cloud project, organization, folder, or billing account. If it makes it easier to exporting from all projects of an organization, you can create an aggregated sink that can export log entries from all the projects, folders, and billing accounts of a Google Cloud organization.

https://cloud.google.com/logging/docs/export/aggregated_sinks

Either way, we now have the data in Cloud Storage, but querying logs information from Cloud Storage is harder than Querying information from BigQuery dataset. For this reason, we should prefer Big Query over Cloud Storage.

Create a Stackdriver Logging Export with a Sink destination to a BigQuery dataset.

Configure the table expiration to 60 days. is the right answer.

You can export logs by creating one or more sinks that include a logs query and an export destination. Supported destinations for exported log entries are Cloud Storage, BigQuery, and Pub/Sub.

Ref: https://cloud.google.com/logging/docs/export/configure_export_v2

Sinks are limited to exporting log entries from the exact resource in which the sink was created: a Google Cloud project, organization, folder, or billing account. If it makes it easier to exporting from all projects of an organization, you can create an aggregated sink that can export log entries from all the projects, folders, and billing accounts of a Google Cloud organization.

https://cloud.google.com/logging/docs/export/aggregated_sinks

Either way, we now have the data in a BigQuery Dataset. Querying information from a Big Query dataset is easier and quicker than analyzing contents in Cloud Storage bucket. As our requirement is to “Quickly analyze the log contents”, we should prefer Big Query over Cloud Storage.

Also, You can control storage costs and optimize storage usage by setting the default table expiration for newly created tables in a dataset. If you set the property when the dataset is created, any table created in the dataset is deleted after the expiration period. If you set the property after the dataset is created, only new tables are deleted after the expiration period.

For example, if you set the default table expiration to 7 days, older data is automatically deleted after 1 week.

Ref: <https://cloud.google.com/bigquery/docs/best-practices-storage>

30. Question

You are migrating a mission critical on-premises application to cloud. The application requires 96 vCPUs to perform its task. You want to make sure the application runs in a similar environment on GCP. What should you do?

- When creating the VM, use machine type n1-standard-96.

- When creating the VM, use Intel Skylake as the CPU platform.
- Create the VM using Compute Engine default settings. Use gcloud to modify the running instance to have 96 vCPUs.
- Start the VM using Compute Engine default settings, and adjust as you go based on Rightsizing Recommendations.

Unattempted

Create the VM using Compute Engine default settings. Use gcloud to modify the running instance to have 96 vCPUs. is not right.

You can't increase the vCPUs to 96 without changing the machine type. While it is possible to set machine type using gcloud, this would mean downtime for the mission-critical application while the upgrade happens which is undesirable.

Ref: <https://cloud.google.com/compute/docs/instances/changing-machine-type-of-stopped-instance>

Start the VM using Compute Engine default settings, and adjust as you go based on Rightsizing Recommendations. is not right.

Since the application is mission-critical, we want to ensure that this application has all the required resources from the beginning. Starting with the default settings provisions a n1-standard-1 machine that has just 1 vCPU and our mission-critical application would be severely constrained for resources.

When creating the VM, use Intel Skylake as the CPU platform. is not right.

Intel Skylake is only offered in E2 machine types that are cost-optimized machine types and offer sizing between 2 to 16 vCPUs which is insufficient for our mission-critical application.

Ref: https://cloud.google.com/compute/docs/machine-types#e2_machine_types

When creating the VM, use machine type n1-standard-96. is the right answer.

n1-standard-96 offers 96 vCPUs and 624 GB of memory. This fits our requirements.

https://cloud.google.com/compute/docs/machine-types#n1_machine_type

31. Question

You are operating a Google Kubernetes Engine (GKE) cluster for your company where different teams can run non-production workloads. Your Machine Learning (ML) team needs access to Nvidia Tesla P100 GPUs to train their models. You want to minimize effort and cost. What should you do?

- Ask your ML team to add the "accelerator: gpu" annotation to their pod specification.
- Recreate all the nodes of the GKE cluster to enable GPUs on all of them.
- Create your own Kubernetes cluster on top of Compute Engine with nodes that have GPUs. Dedicate this cluster to your ML team.
- Add a new, GPU-enabled, node pool to the GKE cluster. Ask your ML team to add the cloud.google.com/gke-accelerator: nvidia-tesla-p100 nodeSelector to their pod specification.

Unattempted

Ask your ML team to add the "accelerator: gpu" annotation to their pod specification. is not right.

There are two issues with this approach. One – the syntax is invalid. Two – You cannot add GPUs to existing node pools.

Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/gpus>

Recreate all the nodes of the GKE cluster to enable GPUs on all of them. is not right.

There are two issues with this approach. One – recreating all nodes to enable GPUs makes the cluster very expensive. Only the ML team needs access to GPUs to train their models.

Recreating all nodes to enable GPUs helps your ML team use them but they are left unused for all other workloads yet cost you money. Two – Even though your nodes have GPUs enabled, you still have to modify pod specifications to request GPU. This step isn't performed in this option.

Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/gpus>

Create your own Kubernetes cluster on top of Compute Engine with nodes that have GPUs. Dedicate this cluster to your ML team. is not right.

While this works, it increases the cost as you now pay the Kubernetes cluster management fee for two clusters instead of one. GKE clusters accrue a management fee that is per cluster per hour, irrespective of cluster size or topology.

Ref: <https://cloud.google.com/kubernetes-engine/pricing>

Add a new, GPU-enabled, node pool to the GKE cluster. Ask your ML team to add the cloud.google.com/gke-accelerator: nvidia-tesla-p100 nodeSelector to their pod specification. is the right answer.

This is the most optimal solution. Rather than recreating all nodes, you create a new node pool with GPU enabled. You then modify the pod specification to target particular GPU types by adding node selector to your workload's Pod specification. YOu still have a single cluster so you pay Kubernetes cluster management fee for just one cluster thus minimizing the cost.

Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/gpus>

Ref: <https://cloud.google.com/kubernetes-engine/pricing>

Example:

```
apiVersion: v1
kind: Pod
metadata:
  name: my-gpu-pod
spec:
  containers:
    - name: my-gpu-container
      image: nvidia/cuda:10.0-runtime-ubuntu18.04
      command: ["/bin/bash"]
  resources:
    limits:
      nvidia.com/gpu: 2
  nodeSelector:
    cloud.google.com/gke-accelerator: nvidia-tesla-k80 # or nvidia-tesla-p100 or nvidia-tesla-p4 or nvidia-tesla-v100 or nvidia-tesla-t4
```

32. Question

You are running an application on multiple virtual machines within a managed instance group and have auto-scaling enabled. The autoscaling policy is configured so that additional instances are added to the group if the CPU utilization of instances goes above 80%. VMs are added until the instance group reaches its maximum limit of five VMs or until CPU utilization of instances lowers to 80%. The initial delay for HTTP health checks against the instances is set to 30 seconds. The virtual machine instances take around three minutes to become available for users. You observe that when the instance group auto-scales, it adds more instances than necessary to support the levels of end-user traffic. You want to properly maintain instance group sizes when autoscaling. What should you do?

- Decrease the maximum number of instances to 3.
- Increase the initial delay of the HTTP health check to 200 seconds.**
- Set the maximum number of instances to 1.
- Use a TCP health check instead of an HTTP health check.

Unattempted

Scenario

? Autoscaling is enabled and kicks off the scale-up
? Scaling policy is based on target CPU utilization of 80%
? The initial delay is 30 seconds

? VM startup time is 3 minutes.

? Auto-scaling creates more instances than necessary.

Set the maximum number of instances to 1. is not right.

Setting the maximum number of instances to 1 effectively limits the scale up to 1 instance which is undesirable as in this case we may still be struggling with the CPU usage but we can't scale up. Therefore this is not the right answer.

Decrease the maximum number of instances to 3. is not right.

Setting the maximum number of instances to 3 effectively limits the scale up to 3 instances which is undesirable as in this case we may still be struggling with the CPU usage but we can't scale up. Therefore this is not the right answer.

Use a TCP health check instead of an HTTP health check. is not right.

TCP health check is a legacy health check, whereas HTTP health check is more advanced and "non-legacy". It is possible a TCP health check might say the application is UP when it is not as it only listens on application servers TCP port and doesn't validate the application health through a HTTP check on its health endpoint. This results in the load balancer sending requests to the application server when it is still loading the application resulting in failures.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/health-checks/create/tcp>

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/health-checks/create/http>

Increase the initial delay of the HTTP health check to 200 seconds. is the right answer.

The reason why our autoscaling is adding more instances than needed is that it checks 30 seconds after launching the instance and at this point, the instance isn't up and isn't ready to serve traffic. So our autoscaling policy starts another instance – again checks this after 30 seconds and the cycle repeats until it gets to the maximum instances or the instances launched earlier are healthy and start processing traffic – which happens after 180 seconds (3 minutes). This can be easily rectified by adjusting the initial delay to be higher than the time it takes for the instance to become available for processing traffic. So setting this to 200 ensures that it waits until the instance is up (around 180-second mark) and then starts forwarding traffic to this instance. Even after a cool out period, if the CPU utilization is still high, the autoscaler can again scale up but this scale-up is genuine and is based on the actual load.

"Initial Delay Seconds" – This setting delays autohealing from potentially prematurely recreating the instance if the instance is in the process of starting up. The initial delay timer starts when the currentAction of the instance is VERIFYING.

Ref: <https://cloud.google.com/compute/docs/instance-groups/autohealing-instances-in-migs>

33. Question

You are setting up a Windows VM on Compute Engine and want to make sure you can log in to the VM via RDP. What should you do?

- After the VM has been created, use your Google Account credentials to log in into the VM.
- After the VM has been created, use `gcloud compute reset-windows-password` to retrieve the login credentials for the VM.
- When creating the VM, add metadata to the instance using 'windows-password' as the key and a password as the value.
- After the VM has been created, download the JSON private key for the default Compute Engine service account. Use the credentials in the JSON file to log in to the VM.

Unattempted

When creating the VM, add metadata to the instance using 'windows-password' as the key and a password as the value. is not right.

It is not possible to specify a windows password at the time of creating windows VM instance. You can generate Windows passwords using either the Google Cloud Console or the `gcloud` command-line tool. Alternatively, you can generate passwords programmatically with API commands but all these methods assume that you have an existing windows instance.

Ref: <https://cloud.google.com/compute/docs/instances/windows/creating-passwords-for-windows-instances#gcloud>

After the VM has been created, use your Google Account credentials to log in into the VM. is not right.

You can generate Windows passwords using either the Google Cloud Console or the `gcloud` command-line tool. Alternatively, you can generate passwords programmatically with API commands but you can't use your `gcloud` account credentials to log into the VM.

Ref: <https://cloud.google.com/compute/docs/instances/windows/creating-passwords-for-windows-instances#gcloud>

After the VM has been created, download the JSON private key for the default Compute Engine service account. Use the credentials in the JSON file to log in to the VM. is not right.

This is not a supported method of authentication for logging into the VM. You can generate Windows passwords using either the Google Cloud Console or the `gcloud`

command-line tool. Alternatively, you can generate passwords programmatically with API commands.

Ref: <https://cloud.google.com/compute/docs/instances/windows/creating-passwords-for-windows-instances#gcloud>

After the VM has been created, use `gcloud compute reset-windows-password` to retrieve the login credentials for the VM. is the right answer.

You can generate Windows passwords using either the Google Cloud Console or the `gcloud` command-line tool. This option uses the right syntax to reset the windows password.
`gcloud compute reset-windows-password windows-instance`

Ref: <https://cloud.google.com/compute/docs/instances/windows/creating-passwords-for-windows-instances#gcloud>

34. Question

You are the organization and billing administrator for your company. The engineering team has the Project Creator role at the organization level. You do not want the engineering team to be able to link projects to the billing account. Only the finance team should be able to link a project to a billing account, but they should not be able to make any other changes to projects. What should you do?

- Assign the engineering team the Billing Account User role on the billing account and the Project Billing Manager role on the organization.
- Assign the finance team only the Billing Account User role on the billing account.
- Assign the engineering team only the Billing Account User role on the billing account.
- Assign the finance team the Billing Account User role on the billing account and the Project Billing Manager role on the organization.

Unattempted

Assign the finance team only the Billing Account User role on the billing account. is not right.

In order to link a project to a billing account, you need the necessary roles at the project level as well as at the billing account level. In this scenario, we are granting just the Billing Account User role on the billing account to the Finance team which allows them to link projects to the billing account on which the role is granted. But we haven't granted them any role at the project level. So they would not be unable to link projects.

Assign the engineering team only the Billing Account User role on the billing account. is not right.

In order to link a project to a billing account, you need the necessary roles at the project level as well as at the billing account level. In this scenario, we are granting just the Billing Account User role on the billing account to the Engineering team which allows them to link projects to the billing account and our question clearly states we do not want to do that.

Assign the engineering team the Billing Account User role on the billing account and the Project Billing Manager role on the organization. is not right.

In order to link a project to a billing account, you need the necessary roles at the project level as well as at the billing account level. In this scenario, we are assigning the engineering team the Billing Account User role on the billing account which allows them to create new projects linked to the billing account on which the role is granted. We are also assigning them the Project Billing Manager role on the organization (trickles down to the project as well) which lets them attach the project to the billing account. But we don't want the engineering team to link projects to the billing account.

Assign the finance team the Billing Account User role on the billing account and the Project Billing Manager role on the organization. is the right answer.

In order to link a project to a billing account, you need the necessary roles at the project level as well as at the billing account level. In this scenario, we are assigning the finance team the Billing Account User role on the billing account which allows them to create new projects linked to the billing account on which the role is granted. We are also assigning them the Project Billing Manager role on the organization (trickles down to the project as well) which lets them attach the project to the billing account, but does not grant any rights over resources. This is exactly what we want.

35. Question

You are the project owner of a GCP project and want to delegate control to colleagues to manage buckets and files in Cloud Storage. You want to follow Google recommended practices. Which IAM roles should you grant your colleagues?

- Project Editor
- Storage Object Creator
- **Storage Admin**
- Storage Object Admin

Unattempted

Project Editor is not right. is not right.

Project editor is a primitive role that grants a lot more than what we need here. Google doesn't recommend using Primitive roles.

Ref: https://cloud.google.com/iam/docs/understanding-roles#primitive_role_definitions

All viewer permissions, plus permissions for actions that modify state, such as changing existing resources.

Storage Object Admin. is not right.

While this role grants full access to the objects, it does not grant access to the buckets so users of this role can not "manage buckets".

This role grants full control over objects, including listing, creating, viewing, and deleting objects.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#storage-roles>

Storage Object Creator. is not right.

This role allows users to create objects. It does not give permission to view, delete, or overwrite objects.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#storage-roles>

Storage Admin. is the right answer.

This role grants full control of buckets and objects. When applied to an individual bucket, control applies only to the specified bucket and objects within the bucket.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#storage-roles>

36. Question

You are using Container Registry to centrally store your company's container images in a separate project. In another project, you want to create a Google Kubernetes Engine (GKE) cluster. You want to ensure that Kubernetes can download images from Container Registry. What should you do?

- In the project where the images are stored, grant the Storage Object Viewer IAM role to the service account used by the Kubernetes nodes.
- When you create the GKE cluster, choose the Allow full access to all Cloud APIs option under 'Access scopes'.

- Create a service account, and give it access to Cloud Storage. Create a P12 key for this service account and use it as an imagePullSecrets in Kubernetes.
- Configure the ACLs on each image in Cloud Storage to give read-only access to the default Compute Engine service account.

Unattempted

Here's some info about where Container Registry stores images and how access is controlled.

Container Registry uses Cloud Storage buckets as the underlying storage for container images. You control access to your images by granting appropriate Cloud Storage permissions to a user, group, service account, or another identity. Cloud Storage permissions granted at the project level apply to all storage buckets in the project, not just the buckets used by Container Registry. To configure permissions specific to Container Registry, grant permissions on the storage bucket used by the registry. Container Registry ignores permissions set on individual objects within the storage bucket.

Ref: <https://cloud.google.com/container-registry/docs/access-control>

Configure the ACLs on each image in Cloud Storage to give read-only access to the default Compute Engine service account. is not right.

As mentioned above, Container Registry ignores permissions set on individual objects within the storage bucket so this isn't going to work.

Ref: <https://cloud.google.com/container-registry/docs/access-control>

When you create the GKE cluster, choose the Allow full access to all Cloud APIs option under 'Access scopes'. is not right.

Selecting Allow full access to all Cloud APIs does not provide access to GCR images in a different project. If the Google Kubernetes Engine cluster and the Container Registry storage bucket are in the same Google Cloud project, the Compute Engine default service account is configured with the appropriate permissions to push or pull images. But if the cluster is in a different project or if the VMs in the cluster use a different service account, you must grant the service account the appropriate permissions to access the storage bucket used by Container Registry.

Ref: <https://cloud.google.com/container-registry/docs/using-with-google-cloud-platform>

In this case, since there is no mention of a service account, we have to assume we are using a default service account that hasn't been provided permissions to access the storage bucket used by Container Registry in another project so the image pull isn't going to work.

You would end up with an error like:

```
Failed to pull image "gcr.io/kubernetes2-278322/simple-python-image": rpc error: code = Unknown desc = Error response from daemon: pull access denied for gcr.io/kubernetes2-278322/simple-python-image, repository does not exist or may require 'docker login'
```

Create a service account, and give it access to Cloud Storage. Create a P12 key for this service account and use it as an imagePullSecrets in Kubernetes. is not right.

It is technically possible to do it this way but using the JSON key and not P12 key as mentioned in this option. If you would like to understand how to do this, please look at these blogs.

Ref: <https://medium.com/hackernoon/today-i-learned-pull-docker-image-from-gcr-google-container-registry-in-any-non-gcp-kubernetes-5f8298f28969>

Ref: <https://medium.com/@michaelmorrissey/using-cross-project-gcr-images-in-gke-1ddc36de3d42>

Moreover, this approach is suitable for accessing GCR images in a non-Google Cloud Kubernetes environment. While it can be used in GKE too, it is not as secure as using Role Bindings since it involves downloading service account keys and setting them up as secret in Kubernetes.

In the project where the images are stored, grant the Storage Object Viewer IAM role to the service account used by the Kubernetes nodes. is the right answer.

Granting the storage object viewer IAM role in the project where images are stored to the service account used by the Kubernetes cluster ensures that the nodes in the cluster can Read Images from the storage bucket. It would be ideal to further restrict the role binding to provide access just to the Cloud Storage bucket that is used as the underlying storage for container images. This follows the principle of least privilege.

For more information about Storage Object Viewer IAM Role for GCR

refer: https://cloud.google.com/container-registry/docs/access-control#permissions_and_roles

37. Question

You are using Deployment Manager to create a Google Kubernetes Engine cluster. Using the same Deployment Manager deployment, you also want to create a DaemonSet in the kube-system namespace of the cluster. You want a solution that uses the fewest possible services. What should you do?

- Add the cluster's API as a new Type Provider in Deployment Manager, and use the new type to create the DaemonSet.
- Use the Deployment Manager Runtime Configurator to create a new Config resource that contains the DaemonSet definition.
- With Deployment Manager, create a Compute Engine instance with a startup script that uses kubectl to create the DaemonSet.

- In the cluster's definition in Deployment Manager, add a metadata that has kube-system as key and the DaemonSet manifest as value.

Unattempted

In the cluster's definition in Deployment Manager, add a metadata that has kube-system as key and the DaemonSet manifest as value. is not right.

Metadata entries are key-value pairs and do not influence this behavior.

Ref: <https://cloud.google.com/compute/docs/storing-retrieving-metadata>

With Deployment Manager, create a Compute Engine instance with a startup script that uses kubectl to create the DaemonSet. is not right.

It is possible to spin up a compute engine instance with a startup script that executes kubectl to create a DaemonSet deployment.

`kubectl apply -f https://k8s.io/examples/controllers/daemonset.yaml`

Ref: <https://kubernetes.io/docs/concepts/workloads/controllers/daemonset/>

But this involves using the compute engine service which is an additional service. Our requirement is to achieve using the fewest possible services and as you'll notice later, the correct answer uses fewer services.

Use the Deployment Manager Runtime Configurator to create a new Config resource that contains the DaemonSet definition. is not right.

You can configure the GKE nodes (provisioned by Deployment manager) to report their status to the Runtime Configurator, and when they are UP, you can run a task to create a DaemonSet. While this is possible, it involves one additional service – to run a task e.g. using Cloud Functions, etc. Our requirement is to achieve using the fewest possible services and as you'll notice later, the correct answer uses fewer services.

Here is some more info about Runtime Configurator. The Runtime Configurator feature lets you define and store data as a hierarchy of key-value pairs in Google Cloud Platform. You can use these key-value pairs as a way to:

1. Dynamically configure services
2. Communicate service states
3. Send notification of changes to data
4. Share information between multiple tiers of services

For example, imagine a scenario where you have a cluster of nodes that run a startup procedure. During startup, you can configure your nodes to report their status to the Runtime Configurator, and then have another application query the Runtime Configurator and run specific tasks based on the status of the nodes.

The Runtime Configurator also offers a Watcher service and a Waiter service. The Watcher service watches a specific key pair and returns when the value of the key pair changes, while the Waiter service waits for a specific end condition and returns a response once that end condition has been met.

Ref: <https://cloud.google.com/deployment-manager/runtime-configurator>

Add the cluster's API as a new Type Provider in Deployment Manager, and use the new type to create the DaemonSet. is the right answer.

A type provider exposes all resources of a third-party API to Deployment Manager as base types that you can use in your configurations. If you have a cluster running on Google Kubernetes Engine, you could add the cluster as a type provider and access the Kubernetes API using Deployment Manager. Using these inherited API, you can create a DaemonSet.

This option uses just the Deployment Manager to create a DaemonSet and is, therefore, the right answer.

Ref: <https://cloud.google.com/deployment-manager/docs/configuration/type-providers/creating-type-provider>

38. Question

You are using Google Kubernetes Engine with autoscaling enabled to host a new application. You want to expose this new application to the public, using HTTPS on a public IP address. What should you do?

- Create a Kubernetes Service of type NodePort for your application, and a Kubernetes Ingress to expose this Service via a Cloud Load Balancer.
- Create a Kubernetes Service of type ClusterIP for your application. Configure the public DNS name of your application using the IP of this Service.
- Create a Kubernetes Service of type NodePort to expose the application on port 443 of each node of the Kubernetes cluster. Configure the public DNS name of your application with the IP of every node of the cluster to achieve load-balancing.
- Create a HAProxy pod in the cluster to load-balance the traffic to all the pods of the application. Forward the public traffic to HAProxy with an iptable rule. Configure the DNS name of your application using the public IP of the nodeHAProxy is running on.

Unattempted

Create a Kubernetes Service of type ClusterIP for your application. Configure the public DNS name of your application using the IP of this Service. is not right.

Kubernetes Service of type ClusterIP exposes the Service on a cluster-internal IP. Choosing this value makes the Service only reachable from within the cluster so you can not route external traffic to this IP.

Ref: <https://kubernetes.io/docs/concepts/services-networking/service/>

Create a HAProxy pod in the cluster to load-balance the traffic to all the pods of the application. Forward the public traffic to HAProxy with an iptable rule. Configure the DNS name of your application using the public IP of the node HAProxy is running on. is not right.

HAProxy is a popular Kubernetes ingress controller. An Ingress object is an independent resource, apart from Service objects, that configures external access to a service's pods. Ingress Controllers still need a way to receive external traffic. This can be done by exposing the Ingress Controller as a Kubernetes service with either NodePort or LoadBalancer type. You can't use public IP of the node the HAProxy is running on as this may be running in any node in the Kubernetes Cluster and in most cases, these nodes do not have public IPs. They are meant to be private and the pods/deployments are accessed through Service objects.

Ref: <https://www.haproxy.com/blog/dissecting-the-haproxy-kubernetes-ingress-controller/>

Create a Kubernetes Service of type NodePort to expose the application on port 443 of each node of the Kubernetes cluster. Configure the public DNS name of your application with the IP of every node of the cluster to achieve load-balancing. is not right.

Kubernetes Service of type NodePort uses a port in the range 30000-32767. Assuming that all the nodes have public IP addresses, enabling NodePort would expose a port such as 32000 so the application is accessible on <https://IP:32000> which is not ideal. You want your application/website to be reachable directly on port 443. This also requires downstream clients to have awareness of all of your nodes' IP addresses, since they will need to connect to those addresses directly. In other words, they won't be able to connect to a single, proxied IP address. And this is against our requirement of "a public IP address".

Ref: <https://kubernetes.io/docs/concepts/services-networking/service/>

Ref: <https://www.haproxy.com/blog/dissecting-the-haproxy-kubernetes-ingress-controller/>

Create a Kubernetes Service of type NodePort for your application, and a Kubernetes Ingress to expose this Service via a Cloud Load Balancer. is the right answer.

This meets all our requirements. With (Global) Cloud Load Balancing, a single anycast IP front-ends all your backend instances in regions around the world. It provides cross-region load balancing, including automatic multi-region failover, which gently moves traffic in fractions if backends become unhealthy.

Ref: <https://cloud.google.com/load-balancing/>

The ingress accepts traffic from the cloud load balancer and can distribute the traffic across the pods in the cluster.

Ref: <https://kubernetes.io/docs/concepts/services-networking/ingress/>

39. Question

You are using multiple configurations for gcloud. You want to review the configured Kubernetes Engine cluster of an inactive configuration using the fewest possible steps. What should you do?

- Use gcloud config configurations describe to review the output.
- Use gcloud config configurations activate and gcloud config list to review the output.
- **Use kubectl config get-contexts to review the output.**
- Use kubectl config use-context and kubectl config view to review the output.

Unattempted

Our requirement is to get to the end goal with the fewest possible steps.

Use gcloud config configurations describe to review the output. is not right.
gcloud config configurations describe – describes a named configuration by listing its properties. This does not return any Kubernetes cluster details.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/describe>

Use gcloud config configurations activate and gcloud config list to review the output. is not right.

gcloud config configurations activate – activates an existing named configuration. This does not return any Kubernetes cluster details.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/activate>

Use kubectl config get-contexts to review the output. is the right answer.
kubectl config get-contexts displays a list of contexts as well as the clusters that use them.
Here's a sample output.

```
$ kubectl config get-contexts
CURRENT NAME CLUSTER
gke_kubernetes-260922_us-central1-a_standard-cluster-1 gke_kubernetes-260922_us-
central1-a_standard-cluster-1
gke_kubernetes-260922_us-central1-a_your-first-cluster-1 gke_kubernetes-260922_us-
central1-a_your-first-cluster-1
* gke_kubernetes-260922_us-central1_standard-cluster-1 gke_kubernetes-260922_us-
central1_standard-cluster-1
```

The output shows the clusters and the configurations they use. Using this information, it is possible to find out the cluster using the inactive configuration with just 1 step.

Use kubectl config use-context and kubectl config view to review the output. is not right. kubectl config use-context [my-cluster-name] is used to set the default context to [my-cluster-name]. But in order to do this, we first need a list of contexts and if you have multiple contexts, you'd need to execute kubectl config use-context [my-cluster-name] against each context. So that is at least 2+ steps. Further to that, the kubectl config view is used to get a full list of config. The output of the kubectl config view can be used to verify which clusters use what configuration but that is one additional step. Moreover, the output of the kubectl config view doesn't change much from one context to other – other than the current-context field. So our earlier steps of determining the contexts and using each context are of not much use. Though this can be used to achieve the same outcome, it involves more steps than the other option.

Here's a sample execution

Step 1: First get a list of contexts

```
kubectl config get-contexts -o=name  
gke_kubernetes-260922_us-central1-a_standard-cluster-1  
gke_kubernetes-260922_us-central1-a_your-first-cluster-1  
gke_kubernetes-260922_us-central1_standard-cluster-1
```

Step 2: Use each context and view the config.

```
kubectl config use-context gke_kubernetes-260922_us-central1-a_standard-cluster-1  
Switched to context "gke_kubernetes-260922_us-central1-a_standard-cluster-1".  
kubectl config view > 1.out (this saves the output in of config view in 1.out)
```

```
kubectl config use-context gke_kubernetes-260922_us-central1-a_your-first-cluster-1  
Switched to context "gke_kubernetes-260922_us-central1-a_your-first-cluster-1".  
kubectl config view > 2.out (this saves the output in of config view in 2.out)
```

```
kubectl config use-context gke_kubernetes-260922_us-central1_standard-cluster-1  
Switched to context "gke_kubernetes-260922_us-central1_standard-cluster-1".  
kubectl config view > 3.out (this saves the output in of config view in 3.out)
```

```
diff 1.out 2.out  
28c28  
< current-context: gke_kubernetes-260922_us-central1-a_standard-cluster-1 --- >  
current-context: gke_kubernetes-260922_us-central1-a_your-first-cluster-1
```

```
diff 2.out 3.out
28c28
< current-context: gke_kubernetes-260922_us-central1-a_your-first-cluster-1 --- >
current-context: gke_kubernetes-260922_us-central1_standard-cluster-1
```

Step 3: Determine the inactive configuration and the cluster using that configuration.

The config itself has details about the clusters and contexts as shown below.

```
$ kubectl config view
apiVersion: v1
clusters:
- cluster:
certificate-authority-data: DATA+OMITTED
server: https://35.222.130.166
name: gke_kubernetes-260922_us-central1-a_standard-cluster-1
- cluster:
certificate-authority-data: DATA+OMITTED
server: https://35.225.14.172
name: gke_kubernetes-260922_us-central1-a_your-first-cluster-1
- cluster:
certificate-authority-data: DATA+OMITTED
server: https://34.69.212.109
name: gke_kubernetes-260922_us-central1_standard-cluster-1
contexts:
- context:
cluster: gke_kubernetes-260922_us-central1-a_standard-cluster-1
user: gke_kubernetes-260922_us-central1-a_standard-cluster-1
name: gke_kubernetes-260922_us-central1-a_standard-cluster-1
- context:
cluster: gke_kubernetes-260922_us-central1-a_your-first-cluster-1
user: gke_kubernetes-260922_us-central1-a_your-first-cluster-1
name: gke_kubernetes-260922_us-central1-a_your-first-cluster-1
- context:
cluster: gke_kubernetes-260922_us-central1_standard-cluster-1
user: gke_kubernetes-260922_us-central1_standard-cluster-1
name: gke_kubernetes-260922_us-central1_standard-cluster-1
current-context: gke_kubernetes-260922_us-central1-a_standard-cluster-1
```

40. Question

You built an application on Google Cloud Platform that uses Cloud Spanner. The support team needs to monitor the environment but should not have access to the data. You need a streamlined solution to grant the correct permissions to your support team, and you want to follow Google recommended practices. What should you do?

- Add the support team group to the roles/spanner.database.reader role
- Add the support team group to the roles/stackdriver.accounts.viewer role
- **Add the support team group to the roles/monitoring.viewer role**
- Add the support team group to the roles/spanner.database.user role

Unattempted

Requirements –

1. Monitoring access but no data access
2. Streamlined solution
3. Google recommended practices (i.e. look for something out of the box).

roles/spanner.databaseReader provides permission to read from the Spanner database, execute SQL queries on the database, and view the schema. Since this provides read access to data, roles/spanner.databaseReader. is not right.

roles/spanner.databaseUser provides permission to read from and write to the Spanner database, execute SQL queries on the database, and view and update the schema. Since this provides both read and write access to data, roles/spanner.databaseUser. is not right.

roles/stackdriver.accounts.viewer read-only access to get and list information about Stackdriver account structure. Since this does not provide monitor access to Cloud Spanner, roles/stackdriver.accounts.viewer. is not right.

roles/monitoring.viewer provides read-only access to get and list information about all monitoring data and configurations. This role provides monitoring access and fits our requirements. roles/monitoring.viewer. is the right answer.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#cloud-spanner-roles>

41. Question

You create a Deployment with 2 replicas in a Google Kubernetes Engine cluster that has a single preemptible node pool. After a few minutes, you use kubectl to examine the status

of your Pod and observe that one of them is still in Pending status:

```
NAME READY STATUS RESTART AGE
myapp-deployment-58ddbbb995-lp86m 0/1 Pending 0 9m
myapp-deployment-58ddbbb995-qjpk 1/1 Running 0 9m
```

What is the most likely cause?

- The pending Pod's resource requests are too large to fit on a single node of the cluster.
- Too many Pods are already running in the cluster, and there are not enough resources left to schedule the pending Pod.
- The node pool is configured with a service account that does not have permission to pull the container image used by the pending Pod.
- The pending Pod was originally scheduled on a node that has been preempted between the creation of the Deployment and your verification of the Pod status. It is currently being rescheduled on a new node.

Unattempted

The pending Pod was originally scheduled on a node that has been preempted between the creation of the Deployment and your verification of the Pod status. It is currently being rescheduled on a new node. is not right.

Our question states that we provisioned a Google Kubernetes Engine cluster with a single preemptible node pool.

The node pool is configured with a service account that does not have permission to pull the container image used by the pending Pod. is not right.

If the node pool has permission issues when pulling the container image, the other pod would not be in Running status. And the status would have been ImagePullBackOff if there was a problem pulling the image.

The pending Pod's resource requests are too large to fit on a single node of the cluster. is not right.

If the resource requests in Pod specification are too large to fit on the node, the other pod would not be in Running status, i.e. both pods should have been in pending status if this was the case.

Ref: The pending Pod's resource requests are too large to fit on a single node of the cluster.

Too many Pods are already running in the cluster, and there are not enough resources left to schedule the pending Pod. is the right answer.

When you have a deployment with some pods in running and other pods in the pending state, more often than not it is a problem with resources on the nodes. Here's a sample

output of this use case. We see that the problem is with insufficient CPU on the Kubernetes nodes so we have to either enable auto-scaling or manually scale up the nodes.

```
kubectl describe pod myapp-deployment-58ddbbb995-lp86m
```

Events:

Type	Reason	Age	From	Message
Warning	FailedScheduling	28s (x4 over 3m1s)	default-scheduler	0/1 nodes are available: 1 Insufficient cpu.

42. Question

You create a new Google Kubernetes Engine (GKE) cluster and want to make sure that it always runs a supported and stable version of Kubernetes. What should you do?

- Select "Container-Optimized OS (cos)" as a node image for your GKE cluster.
- Select the latest available cluster version for your GKE cluster.
- **Enable the Node Auto-Upgrades feature for your GKE cluster.**
- Enable the Node Auto-Repair feature for your GKE cluster.

Unattempted

Select the latest available cluster version for your GKE cluster. is not right.
GKE's node auto-repair feature helps you keep the nodes in your cluster in a healthy, running state. When enabled, GKE makes periodic checks on the health state of each node in your cluster. If a node fails consecutive health checks over an extended time period, GKE initiates a repair process for that node.

Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/node-auto-repair>

Select the latest available cluster version for your GKE cluster. is not right.
We can certainly select the latest available cluster version at the time of GKE cluster provisioning, however, this does not automatically upgrade the cluster if new versions become available.

Select "Container-Optimized OS (cos)" as a node image for your GKE cluster. is not right.
Container-Optimized OS comes with the Docker container runtime and all Kubernetes components pre-installed for out of the box deployment, management, and orchestration of your containers. But these do not help with automatically upgrading GKE cluster

versions.

Ref: <https://cloud.google.com/container-optimized-os>

Enable the Node Auto-Upgrades feature for your GKE cluster. is the right answer.
Node auto-upgrades help you keep the nodes in your cluster up to date with the cluster master version when your master is updated on your behalf. When you create a new cluster or node pool with Google Cloud Console or the gcloud command, node auto-upgrade is enabled by default.

Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/node-auto-upgrades>

43. Question

You created a cluster.YAML file containing resources:

```
- name: cluster
  type: container.v1.cluster
  properties:
    zone: europe-west1-b
    cluster:
      description: "My GCP ACE cluster"
      initialNodeCount: 2
```

You want to use Cloud Deployment Manager to create this cluster in GKE. What should you do?

- `gcloud deployment-manager deployments create my-gcp-ace-cluster --config cluster.yaml`
- `gcloud deployment-manager deployments create my-gcp-ace-cluster --type container.v1.cluster --config cluster.yaml`
- `gcloud deployment-manager deployments apply my-gcp-ace-cluster --config cluster.yaml`
- `gcloud deployment-manager deployments apply my-gcp-ace-cluster --type container.v1.cluster --config cluster.yaml`

Unattempted

`gcloud deployment-manager deployments apply my-gcp-ace-cluster --config cluster.yaml`. is not right.

“`gcloud deployment-manager deployments`” doesn’t support action `apply`. With Google cloud in general, the action for creating is `create` and the action for retrieving is `list`. With Kubernetes resources, the corresponding actions are `apply` and `get` respectively.

Ref: <https://cloud.google.com/sdk/gcloud/reference/deployment-manager/deployments/create>

`gcloud deployment-manager deployments apply my-gcp-ace-cluster --type container.v1.cluster --config cluster.yaml`. is not right.

“`gcloud deployment-manager deployments`” doesn’t support action `apply`. With Google cloud in general, the action for creating is `create` and the action for retrieving is `list`. With Kubernetes resources, the corresponding actions are `apply` and `get` respectively.

Ref: <https://cloud.google.com/sdk/gcloud/reference/deployment-manager/deployments/create>

`gcloud deployment-manager deployments create my-gcp-ace-cluster --type container.v1.cluster --config cluster.yaml`. is not right.

“`gcloud deployment-manager deployments create`” creates deployments based on the configuration file. (Infrastructure as code). It doesn’t expect the parameter type passed to it directly and fails when executed with the type parameter.

Ref: <https://cloud.google.com/sdk/gcloud/reference/deployment-manager/deployments/create>

`gcloud deployment-manager deployments create my-gcp-ace-cluster --config cluster.yaml`. is the right answer.

“`gcloud deployment-manager deployments create`” creates deployments based on the configuration file. (Infrastructure as code). All the configuration related to the artifacts is in the configuration file. This command correctly creates a cluster based on the provided `cluster.yaml` configuration file.

Ref: <https://cloud.google.com/sdk/gcloud/reference/deployment-manager/deployments/create>

44. Question

You created a compute instance by running `gcloud compute instances create instance1`. You intended to create the instance in project `gcp-ace-proj-266520` but the instance got created in a different project. Your cloud shell `gcloud` configuration is as shown.

```
$ gcloud config list
[component_manager]
disable_update_check = True
[compute]
gce_metadata_read_timeout_sec = 5
zone = europe-west2-a
```

```
[core]
account = gcp-ace-lab-user@gmail.com
disable_usage_reporting = False
project = gcp-ace-lab-266520
```

[metrics]

```
environment = devshell
```

What should you do to delete the instance that was created in the wrong project and recreate it in gcp-ace-proj-266520 project?

- `gcloud compute instances delete instance1` `gcloud config set compute/project gcp-ace-proj-266520` `gcloud compute instances create instance1`
- `gcloud config set project gcp-ace-proj-266520` `gcloud compute instances recreate instance1 --previous-project gcp-ace-lab-266520`
- `gcloud compute instances delete instance1` `gcloud compute instances create instance1`
- `gcloud compute instances delete instance1` `gcloud config set project gcp-ace-proj-266520` `gcloud compute instances create instance1`

Unattempted

```
gcloud compute instances delete instance1
```

`gcloud compute instances create instance1.` is not right.

The default core/project property is set to gcp-ace-lab-266520 in our current configuration so the instance would have been created in this project. Running the first command to delete the instance correctly deletes it from this project but we haven't modified the core/project property before executing the second command so the instance is recreated in the same project which is not what we want.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/create>

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/delete>

```
gcloud config set project gcp-ace-proj-266520
```

`gcloud compute instances recreate instance1 --previous-project gcp-ace-lab-266520.` is not right.

`gcloud compute instances` command doesn't support recreate action. It supports create/delete which is what we are supposed to use for this requirement.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances>

```
gcloud compute instances delete instance1
```

```
gcloud config set compute/project gcp-ace-proj-266520
```

`gcloud compute instances create instance1.` is not right.

The approach is right but the syntax is wrong. gcloud config does not have a compute/project property. The project property is part of the core/ section as seen in the

output of gcloud configuration list in the question. In this scenario, we are trying to set compute/project property that doesn't exist in the compute section so the command fails.
Ref: <https://cloud.google.com/sdk/gcloud/reference/config/set>

```
gcloud compute instances delete instance1  
gcloud config set project gcp-ace-proj-266520  
gcloud compute instances create instance1. is the right answer.  
This sequence of commands correctly deletes the instance from gcp-ace-lab-266520  
which is the default project in the active gcloud configuration, then modifies the current  
configuration to set the default project to gcp-ace-proj-266520, and finally creates the  
instance in the project gcp-ace-proj-266520 which is the default project in active gcloud  
configuration at the time of running the command. This produces the intended outcome  
of deleting the instance from gcp-ace-lab-266520 project and recreating it in gcp-ace-  
prod-266520
```

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/set>
Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/create>
Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/delete>

45. Question

You created a Google Cloud Platform project with an App Engine application inside the project. You initially configured the application to be served from the us-central region. Now you want the application to be served from the asia-northeast1 region. What should you do?

- Create a new GCP project and create an App Engine application inside this new project. Specify asia-northeast1 as the region to serve your application.
- Create a second App Engine application in the existing GCP project and specify asia-northeast1 as the region to serve your application.
- Change the default region property setting in the existing GCP project to asia-northeast1.
- Change the region property setting in the existing App Engine application from us-central to asia-northeast1.

Unattempted

Change the default region property setting in the existing GCP project to asia-northeast1. is not right.
App Engine is regional, which means the infrastructure that runs your apps is located in a specific region and is managed by Google to be redundantly available across all the zones

within that region. You cannot change an app's region after you set it.

Ref: <https://cloud.google.com/appengine/docs/locations>

Change the region property setting in the existing App Engine application from us-central to asia-northeast1. is not right.

App Engine is regional, which means the infrastructure that runs your apps is located in a specific region and is managed by Google to be redundantly available across all the zones within that region. You cannot change an app's region after you set it.

Ref: <https://cloud.google.com/appengine/docs/locations>

Create a second App Engine application in the existing GCP project and specify asia-northeast1 as the region to serve your application. is not right.

App Engine is regional and you cannot change an app's region after you set it. You can deploy additional services in the App Engine but they will all be targeted to the same region.

Ref: <https://cloud.google.com/appengine/docs/locations>

Create a new GCP project and create an App Engine application inside this new project.

Specify asia-northeast1 as the region to serve your application. is the right answer.

App Engine is regional and you cannot change an app's region after you set it. Therefore, the only way to have an app run in another region is by creating a new project and targeting the app engine to run in the required region (asia-northeast1 in our case).

Ref: <https://cloud.google.com/appengine/docs/locations>

46. Question

You created a Kubernetes deployment by running `kubectl run nginx --image=nginx --labels="app=prod"`. Your Kubernetes cluster is also used by a number of other deployments. How can you find the identifier of the pods for this nginx deployment?

- `kubectl get deployments --output=pods`
- `gcloud get pods --selector="app=prod"`
- `gcloud list gke-deployments --filter={ pod }`
- `kubectl get pods -l "app=prod"`

Unattempted

`gcloud get pods --selector="app=prod"`. is not right.

You can not retrieve pods from the Kubernetes cluster by using `gcloud`. You can list pods

by using Kubernetes CLI – kubectl get pods.

Ref: <https://kubernetes.io/docs/tasks/access-application-cluster/list-all-running-container-images/>

gcloud list gke-deployments –filter={ pod }. is not right.

You can not retrieve pods from the Kubernetes cluster by using gcloud. You can list pods by using Kubernetes CLI – kubectl get pods.

Ref: <https://kubernetes.io/docs/tasks/access-application-cluster/list-all-running-container-images/>

kubectl get deployments –output=pods. is not right.

You can not list pods by listing Kubernetes deployments. You can list pods by using Kubernetes CLI – kubectl get pods.

Ref: <https://kubernetes.io/docs/tasks/access-application-cluster/list-all-running-container-images/>

kubectl get pods -l "app=prod". is the right answer.

This command correctly lists pods that have the label app=prod. When creating the deployment, we used the label app=prod so listing pods that have this label retrieve the pods belonging to nginx deployments. You can list pods by using Kubernetes CLI – kubectl get pods.

Ref: <https://kubernetes.io/docs/tasks/access-application-cluster/list-all-running-container-images/>

Ref: <https://kubernetes.io/docs/tasks/access-application-cluster/list-all-running-container-images/#list-containers-filtering-by-pod-label>

47. Question

You created a Kubernetes deployment by running kubectl run nginx –image=nginx –replicas=1. After a few days, you decided you no longer want this deployment. You identified the pod and deleted it by running kubectl delete pod. You noticed the pod got recreated.

```
$ kubectl get pods  
NAME READY STATUS RESTARTS AGE  
nginx-84748895c4-nqqmt 1/1 Running 0 9m41s
```

```
$ kubectl delete pod nginx-84748895c4-nqqmt  
pod "nginx-84748895c4-nqqmt" deleted
```

```
$ kubectl get pods
```

```
NAME READY STATUS RESTARTS AGE
```

```
nginx-84748895c4-k6bzl 1/1 Running 0 25s
```

What should you do to delete the deployment and avoid pod getting recreated?

- `kubectl delete nginx`
- `kubectl delete --deployment=nginx`
- `kubectl delete pod nginx-84748895c4-k6bzl --no-restart`
- `kubectl delete deployment nginx`

Unattempted

`kubectl delete pod nginx-84748895c4-k6bzl --no-restart.` is not right.

`kubectl delete pod` command does not support the flag `--no-restart`. The command fails to execute due to the presence of an invalid flag.

```
$ kubectl delete pod nginx-84748895c4-k6bzl --no-restart
```

```
Error: unknown flag: --no-restart
```

Ref: <https://kubernetes.io/docs/reference/kubectl/cheatsheet/#deleting-resources>

`kubectl delete --deployment=nginx.` is not right.

`kubectl delete` command does not support the parameter `--deployment`. The command fails to execute due to the presence of an invalid parameter.

```
$ kubectl delete --deployment=nginx
```

```
Error: unknown flag: --deployment
```

Ref: <https://kubernetes.io/docs/reference/kubectl/cheatsheet/#deleting-resources>

`kubectl delete nginx.` is not right.

We haven't provided the `kubectl delete` command information on what to delete, whether a pod, a service or a deployment. The command syntax is wrong and fails to execute.

```
$ kubectl delete nginx
```

```
error: resource(s) were provided, but no name, label selector, or -all flag specified
```

Ref: <https://kubernetes.io/docs/reference/kubectl/cheatsheet/#deleting-resources>

`kubectl delete deployment nginx.` is the right answer.

This command correctly deletes the deployment. Pods are managed by Kubernetes workloads (deployments). When a pod is deleted, the deployment detects the pod is unavailable and brings up another pod to maintain the replica count. The only way to delete the workload is by deleting the deployment itself using the `kubectl delete deployment` command.

```
$ kubectl delete deployment nginx
```

deployment.apps "nginx" deleted

Ref: <https://kubernetes.io/docs/reference/kubectl/cheatsheet/#deleting-resources>

48. Question

You created an instance of SQL Server 2017 on Compute Engine to test features in the new version. You want to connect to this instance using the fewest number of steps. What should you do?

- Set a Windows password in the GCP Console. Verify that a firewall rule for port 22 exists. Click the RDP button in the GCP Console and supply the credentials to log in.
- Set a Windows username and password in the GCP Console. Verify that a firewall rule for port 3389 exists. Click the RDP button in the GCP Console, and supply the credentials to log in.
- Install an RDP client on your desktop. Verify that a firewall rule for port 3389 exists.
- **Install an RDP client on your desktop. Set a Windows username and password in the GCP Console. Use the credentials to log in to the instance.**

Unattempted

Requirements – Connect to compute instance using fewest steps. The presence of SQL Server 2017 on the instance is a red herring and should be ignored as none of the options provided say anything about the database and all seem to revolve around RDP.

Install a RDP client on your desktop. Verify that a firewall rule for port 3389 exists. is not right.

Although opening port 3389 is essential for serving RDP traffic, we do not have the credentials to RDP so this isn't going to work.

Set a Windows password in the GCP Console. Verify that a firewall rule for port 22 exists. Click the RDP button in the GCP Console and supply the credentials to log in. is not right.
RDP uses port 3389 and not 22.

Ref: <https://cloud.google.com/compute/docs/troubleshooting/troubleshooting-rdp>

Set a Windows username and password in the GCP Console. Verify that a firewall rule for port 3389 exists. Click the RDP button in the GCP Console, and supply the credentials to log in. is not right.

While this option correctly sets the username and password on the console and verifies a firewall rule is set on port 3389 to allow RDP traffic, you can RDP from console unless you install Chrome RDP for Google Cloud Platform extension in order to RDP from the console. (See Chrome Desktop for GCP tab in <https://cloud.google.com/compute/docs/instances/connecting-to-instance#windows>). If we assume that installing Chrome RDP for Google Cloud Platform extension is carried out (even though not specified in the option), we end up executing more steps in this option to successfully RDP compare to the correct answer (below)

Install an RDP client on your desktop. Set a Windows username and password in the GCP Console. Use the credentials to log in to the instance. is the right answer.

This option correctly sets the username/password which is essential. In addition, the default VPC comes with port 3389 open to the public. The question doesn't explicitly state the compute engine is in a custom VPC so it is safe to assume we are using default VPC which has default RDP access open to the public. Finally, you install an RDP client on the desktop and use the credentials set up earlier to RDP to the server.

49. Question

You defined an instance template for a Python web application. When you deploy this application in Google Compute Engine, you want to ensure the service scales up and scales down automatically based on the number of HTTP requests. What should you do?

- 1. Create the necessary number of instances based on the instance template to handle peak user traffic. 2. Group the instances together in an unmanaged instance group. 3. Configure the instance group as the Backend Service of an External HTTP(S) load balancer.
- 1. Create an instance from the instance template. 2. Create an image from the instance's disk and export it to Cloud Storage. 3. Create an External HTTP(S) load balancer and add the Cloud Storage bucket as its backend service.
- 1. Create an unmanaged instance group from the instance template. 2. Configure autoscaling on the unmanaged instance group with a scaling policy based on HTTP traffic. 3. Configure the unmanaged instance group as the backend service of an Internal HTTP(S) load balancer.
- 1. Deploy your Python web application instance template to Google Cloud App Engine. 2. Configure autoscaling on the managed instance group with a scaling policy based on HTTP traffic.

- 1. Create a managed instance group from the instance template. 2. Configure autoscaling on the managed instance group with a scaling policy based on HTTP traffic. 3. Configure the instance group as the backend service of an External HTTP(S) load balancer.

Unattempted

1. Create an instance from the instance template.
2. Create an image from the instance's disk and export it to Cloud Storage.
3. Create an External HTTP(s) load balancer and add the Cloud Storage bucket as its backend service. is not right.

You can upload a custom image from instance's boot disk and export it to cloud storage.

<https://cloud.google.com/compute/docs/images/export-image>

However, this image in the Cloud Storage bucket is unable to handle traffic as it is not a running application. Cloud Storage can not serve requests of the custom image.

1. Create an unmanaged instance group from the instance template.
2. Configure autoscaling on the unmanaged instance group with a scaling policy based on HTTP traffic.
3. Configure the unmanaged instance group as the backend service of an Internal HTTP(S) load balancer. is not right.

An unmanaged instance group does not autoscale. An unmanaged instance group is a collection of virtual machines (VMs) that reside in a single zone, VPC network, and subnet. An unmanaged instance group is useful for grouping together VMs that require individual configuration settings or tuning.

Ref: <https://cloud.google.com/compute/docs/instance-groups/creating-groups-of-unmanaged-instances>

1. Create the necessary number of instances based on the instance template to handle peak user traffic.
2. Group the instances together in an unmanaged instance group.
3. Configure the instance group as the Backend Service of an External HTTP(S) load balancer. is not right.

An unmanaged instance group does not autoscale. Although we may have enough compute power to handle peak user traffic, it does not automatically scale down when the traffic goes down so it doesn't meet our requirements.

Ref: <https://cloud.google.com/compute/docs/instance-groups/creating-groups-of-unmanaged-instances>

1. Deploy your Python web application instance template to Google Cloud App Engine.
2. Configure autoscaling on the managed instance group with a scaling policy based on HTTP traffic. is not right.

You can not use compute engine instance templates to deploy applications to Google Cloud

App Engine. Google App Engine lets you deploy applications quickly by providing run time environments for many of the popular languages like Java, PHP, Node.js, Python, C#, .Net, Ruby, and Go. You have an option of using custom runtimes but using compute engine instance templates is not an option.

Ref: <https://cloud.google.com/appengine>

1. Create a managed instance group from the instance template.
2. Configure autoscaling on the managed instance group with a scaling policy based on HTTP traffic.
3. Configure the instance group as the backend service of an External HTTP(S) load balancer. is the right answer.

The auto-scaling capabilities of Managed instance groups let you automatically add or delete instances from a managed instance group based on increases or decreases in load – this can be set up by configuring scaling policies. In addition, you can configure External HTTP(S) load balancer to send traffic to the managed instance group. The External HTTP(S) load balancer tries to balance requests by using a round-robin algorithm and when the load increases beyond the threshold defined in the scaling policy, autoscaling kicks in and adds more nodes.

Ref: <https://cloud.google.com/load-balancing/docs/https>

Ref: <https://cloud.google.com/compute/docs/instance-groups/creating-groups-of-managed-instances>

50. Question

You deployed a new application inside your Google Kubernetes Engine cluster using the YAML file specified below.

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: myapp-deployment
spec:
  selector:
    matchLabels:
      app: myapp
  replicas: 2
  template:
    metadata:
      labels:
        app: myapp
    spec:
```

```

containers:
- name: myapp
image: myapp:1.1
ports:
- containerPort: 80
-
apiVersion: v1
kind: Service
metadata:
name: myapp-service
spec:
ports:
- port: 8000
targetPort: 80
protocol: TCP
selector:
app: myapp
You check the status of the deployed pods and notice that one of them is still in PENDING
status:
kubectl get pods -l app=myapp

```

NAME	READY	STATUS	RESTARTS	AGE
myapp-deployment-58ddbbb995-lp86m	0/1	Pending	0	9m
myapp-deployment-58ddbbb995-qjpkq	1/1	Running	0	9m

You want to find out why the pod is stuck in pending status. What should you do?

- Review details of the myapp-service Service object and check for error messages.
- Review details of the myapp-deployment Deployment object and check for error messages.
- Review details of myapp-deployment-58ddbbb995-lp86m Pod and check for warning messages.
- View logs of the container in myapp-deployment-58ddbbb995-lp86m pod and check for warning messages.

Unattempted

Review details of the myapp-service Service object and check for error messages. is not right.

The question states we have a problem with the deployment. Checking/Reviewing the status of the service object isn't of much use here.

View logs of the container in myapp-deployment-58ddbbb995-1p86m pod and check for warning messages. is not right.

Since the pod hasn't moved to Running state, the logs of the container would be empty.

So running

```
kubectl logs pod/myapp-deployment-58ddbbb995-1p86m
```

to check the logs of the pod isn't of much use.

Review details of the myapp-deployment Deployment object and check for error messages. is not right.

Describing the details of the deployment shows us how many of the pods are available and unavailable but does not show errors/warnings related to a specific pod.

Here's a sample output of this use case.

```
kubectl describe deployment myapp-deployment
```

```
Replicas: 3 desired | 3 updated | 3 total | 2 available | 1 unavailable
```

Events:

Type	Reason	Age	From	Message
---	---	---	---	---

```
Normal ScalingReplicaSet 4m54s deployment-controller Scaled up replica set myapp-deployment-869d88c75f to 3
```

Review details of myapp-deployment-58ddbbb995-1p86m Pod and check for warning messages. is the right answer.

Since the problem is with a specific pod, looking at the details of the pod is the best solution. When you have a deployment with some pods in running and other pods in Pending state, more often than not it is a problem with resources on the nodes. Here's a sample output of this use case. We see that the problem is with insufficient CPU on the Kubernetes nodes so we have to either enable auto-scaling or manually scale up the nodes.

```
kubectl describe pod myapp-deployment-58ddbbb995-1p86m
```

Events:

Type	Reason	Age	From	Message
---	---	---	---	---

```
Warning FailedScheduling 28s (x4 over 3m1s) default-scheduler 0/1 nodes are available: 1 Insufficient cpu.
```

51. Question

You deployed a number of services to Google App Engine Standard. The services are designed as microservices with several interdependencies between them. Most services have few version upgrades but some key services have over 20 version upgrades. You identified an issue with the service pt-createOrder and deployed a new version v3 for this service.

You are confident this works and want this new version to receive all traffic for the service. You want to minimize effort and ensure the availability of service. What should you do?

- Execute gcloud app versions stop v2 and gcloud app versions start v3
- Execute gcloud app versions stop v2 --service="pt-createOrder" and gcloud app versions start v3 --service="pt-createOrder"
- Execute gcloud app versions migrate v3
- **Execute gcloud app versions migrate v3 --service="pt-createOrder"**

Unattempted

Execute gcloud app versions migrate v3. is not right.

gcloud app versions migrate v3 migrates all services to version v3. In our scenario, we have multiple services with each service potentially being on a different version. We don't want to migrate all services to v3, instead, we only want to migrate the pt-createOrder service to v3. Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/migrate>

Execute gcloud app versions stop v2 --service="pt-createOrder" and gcloud app versions start v3 --service="pt-createOrder". is not right.

Stopping version v2 and starting version v3 for pt-createOrder service would result in v3 receiving all traffic for pt-createOrder. While this is the intended outcome, stopping version v2 before starting version v3 results in service being unavailable until v3 is ready to receive traffic. As we want to "ensure availability", this option is not suitable.

Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/migrate>

Execute gcloud app versions stop v2 and gcloud app versions start v3. is not right.

Stopping version v2 and starting version v3 would result in migrating all services to version v3 which is undesirable. We don't want to migrate all services to v3, instead, we only want to migrate the pt-createOrder service to v3. Moreover, stopping version v2 before starting version v3 results in service being unavailable until v3 is ready to receive traffic. As we want to "ensure availability", this option is not suitable.

Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/migrate>

Execute gcloud app versions migrate v3 --service="pt-createOrder". is the right answer. This command correctly migrates the service pt-createOrder to use version 3 and produces the intended outcome while minimizing effort and ensuring the availability of service.

Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/migrate>

52. Question

You deployed a workload to your GKE cluster by running the command `kubectl apply -f app.yaml`. You also enabled a LoadBalancer service to expose the deployment by running `kubectl apply -f service.yaml`. Your pods are struggling due to increased load so you decided to enable horizontal pod autoscaler by running `kubectl autoscale deployment [YOUR DEPLOYMENT] -cpu-percent=50 -min=1 -max=10`. You noticed the autoscaler has launched several new pods but the new pods have failed with the message “Insufficient cpu”. What should you do to resolve this issue?

- Use "gcloud container clusters resize" to add more nodes to the node pool.
- Use "kubectl container clusters resize" to add more nodes to the node pool.
- Edit the managed instance group of the cluster and enable autoscaling.
- Edit the managed instance group of the cluster and increase the number of VMs by 1.

Unattempted

Use “`kubectl container clusters resize`” to add more nodes to the node pool. is not right.
`kubectl` doesn’t support the command `kubectl container clusters resize`. You have to use `gcloud container clusters resize` to resize a cluster.

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/resize>

Edit the managed instance group of the cluster and increase the number of VMs by 1. is not right.

GKE Cluster does not use a managed instance group. Instead, the cluster master (control plan) handles the lifecycle of nodes in the node pools. The cluster master is responsible for managing the workloads’ lifecycle, scaling, and upgrades. The master also manages network and storage resources for those workloads.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/cluster-architecture>

Edit the managed instance group of the cluster and enable autoscaling. is not right.

GKE Cluster does not use a managed instance group. Instead, the cluster master (control plan) handles the lifecycle of nodes in the node pools. The cluster master is responsible for managing the workloads’ lifecycle, scaling, and upgrades. The master also manages network and storage resources for those workloads.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/cluster-architecture>

Use “`gcloud container clusters resize`” to add more nodes to the node pool. is the right answer.

Your pods are failing with “Insufficient cpu”. This is because the existing nodes in the node pool are maxed out, therefore, you need to add more nodes to your node pool. For such scenarios, enabling cluster autoscaling is ideal, however, this is not in any of the answer options. In the absence of cluster autoscaling, the next best approach is to add more nodes to the cluster manually. This is achieved by running the command gcloud container clusters resize which resizes an existing cluster for running containers.

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/resize>

53. Question

You deployed an App Engine application using gcloud app deploy, but it did not deploy to the intended project. You want to find out why this happened and where the application deployed. What should you do?

- Check the app YAML file for your application and check the project settings.
- Go to Deployment Manager and review settings for the deployment of application
- Check the web application XML file for your application and check project settings
- Go to Cloud Shell and run gcloud config list to review the Google Cloud configurations used for deployment.

Unattempted

Check the app YAML file for your application and check the project settings. is not right.
The Yaml file of application does not hold Google project information.

Check the web application XML file for your application and check project settings. is not right.

The web application file of the application does not hold Google project information.

Go to Deployment Manager and review settings for the deployment of the application. is not right.

Google Cloud Deployment Manager allows you to specify all the resources needed for your application in a declarative format using yaml. In this scenario, we haven't used Cloud Deployment Manager to deploy. The app was deployed using gcloud app deploy so this option is not right.

Ref: <https://cloud.google.com/deployment-manager>

Go to Cloud Shell and run gcloud config list to review the Google Cloud configurations used for deployment. is the right answer.

If the deployment was successful but it did not deploy to the intended project, it is likely that the gcloud app deploy command deployed the application to a different project. In the same gcloud shell, you can identify the current properties of the configuration by executing gcloud config list. This returns config properties such as project, account etc, as well as app-specific properties such as app/promote_by_default, app/stop_previous_version.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/list>

54. Question

You deployed an LDAP server on Compute Engine. You want to make sure it is reachable by external clients via TLS through port 636 using UDP. What should you do?

- Add the network tag allow-udp-636 to the VM instance running the LDAP server.
- Create a route called allow-udp-636 and set the next hop to be the VM instance running the LDAP server.
- Add a network tag of your choice to the instance. Create a firewall rule to allow ingress on UDP port 636 for that network tag.
- Add a network tag of your choice to the instance running the LDAP server. Create a firewall rule to allow egress on UDP port 636 for that network tag.

Unattempted

Create a route called allow-udp-636 and set the next hop to be the VM instance running the LDAP server. is not right.

Google Cloud routes define the paths that network traffic takes from a virtual machine (VM) instance to other destinations. These destinations can be inside your Google Cloud Virtual Private Cloud (VPC) network (for example, in another VM) or outside it. Routes aren't a suitable solution for our requirement as we need to enable EXTERNAL clients to reach our VM on port 636 using UDP.

Ref: <https://cloud.google.com/vpc/docs/routes>

Add the network tag allow-udp-636 to the VM instance running the LDAP server. is not right.

Tags enable you to make firewall rules and routes applicable to specific VM instances but allow-udp-636 is not a network tag that GCP provides. The default network tags provided by GCP are default-allow-icmp, default-allow-internal, default-allow-rdp and default-allow-ssh. In this scenario, we are assigning a tag to the instance with no network

rules so there would be no difference to behavior.

Ref: <https://cloud.google.com/vpc/docs/add-remove-network-tags>

Add a network tag of your choice to the instance running the LDAP server. Create a firewall rule to allow egress on UDP port 636 for that network tag. is not right.

We are interested in enabling inbound traffic to our VM whereas egress firewall rules control outgoing connections from target instances in your VPC network.

Ref: https://cloud.google.com/vpc/docs/firewalls#egress_cases

Add a network tag of your choice to the instance. Create a firewall rule to allow ingress on UDP port 636 for that network tag. is the right answer.

This fits all our requirements. Ingress firewall rules control incoming connections from a source to target instances in your VPC network. We can create an ingress firewall rule to allow UDP port 636 for a network tag. And when we assign this network tag to the instance, the firewall rule applies to the instances so traffic is accepted on port 636 using UDP. Although not specified in this option, it has to be assumed that the source for the firewall rule is set to 0.0.0.0/0 i.e. all IP ranges so that external clients are allowed to connect to this VM.

Ref: https://cloud.google.com/vpc/docs/firewalls#ingress_cases

55. Question

You deployed your application to a default node pool on the GKE cluster and you want to configure cluster autoscaling for this GKE cluster. For your application to be profitable, you must limit the number of Kubernetes nodes to 10. You want to start small and scale up as traffic increases and scale down when the traffic goes down. What should you do?

- Update existing GKE cluster to enable autoscaling by running the command `gcloud container clusters update [CLUSTER_NAME] --enable-autoscaling --min-nodes=1 --max-nodes=10`
- Create a new GKE cluster by running the command `gcloud container clusters create [CLUSTER_NAME] --enable-autoscaling --min-nodes=1 --max-nodes=10`. Redeploy your application
- To enable autoscaling, add a tag to the instances in the cluster by running the command `gcloud compute instances add-tags [INSTANCE] --tags=enable-autoscaling,min-nodes=1,max-nodes=10`

- Set up a stack driver alert to detect slowness in the application. When the alert is triggered, increase nodes in the cluster by running the command `gcloud container clusters resize CLUSTER_NAME --size`.

Unattempted

Set up a stack driver alert to detect slowness in the application. When the alert is triggered, increase nodes in the cluster by running the command `gcloud container clusters resize CLUSTER_NAME --size`. is not right.

The command `gcloud container clusters resize` command resizes an existing cluster for running containers. While it is possible to manually increase the number of nodes in the cluster by running the command, the scale-up is not automatic, it is a manual process. Also, there is no scale down so it doesn't fit our requirement of "scale up as traffic increases and scale down when the traffic goes down".

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/resize>

To enable autoscaling, add a tag to the instances in the cluster by running the command `gcloud compute instances add-tags [INSTANCE] --tags=enable-autoscaling,min-nodes=1,max-nodes=10`. is not right.

Autoscaling can not be enabled on the GKE cluster by adding tags on compute instances. Autoscaling can be enabled at the time of creating the cluster and can also be enabled for existing clusters by running one of the `gcloud container clusters` to create/update commands.

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/create>

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/update>

Create a new GKE cluster by running the command `gcloud container clusters create [CLUSTER_NAME] --enable-autoscaling --min-nodes=1 --max-nodes=10`. Redeploy your application. is not right.

The command `gcloud container clusters create` – creates a GKE cluster and the flag `--enable-autoscaling` enables autoscaling and the parameters `--min-nodes=1 --max-nodes=10` define the minimum and maximum number of nodes in the node pool. However, we want to configure cluster autoscaling for the existing GKE cluster; not create a new GKE cluster.

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/create>

Update existing GKE cluster to enable autoscaling by running the command `gcloud container clusters update [CLUSTER_NAME] --enable-autoscaling --min-nodes=1 --max-nodes=10`. is the right answer.

The command `gcloud container clusters update` – updates an existing GKE cluster. The flag `--enable-autoscaling` enables autoscaling and the parameters `--min-nodes=1 --max-nodes=10` define the minimum and maximum number of nodes in the node pool. This

enables cluster autoscaling which scales up and scales down the nodes automatically between 1 and 10 nodes in the node pool.

56. Question

You developed a web application that lets users upload and share images. You deployed this application in Google Compute Engine and you have configured Stackdriver Logging. Your application sometimes times out while uploading large images, and your application generates relevant error log entries that are ingested to Stackdriver Logging. You would now like to create alerts based on these metrics. You intend to add more compute resources manually when the number of failures exceeds a threshold. What should you do in order to alert based on these metrics with minimal effort?

- In Stackdriver logging, create a new logging metric with the required filters, edit the application code to set the metric value when needed, and create an alert in Stackdriver based on the new metric.
- Create a custom monitoring metric in code, edit the application code to set the metric value when needed, create an alert in Stackdriver based on the new metric.
- In Stackdriver Logging, create a custom monitoring metric from log data and create an alert in Stackdriver based on the new metric.
- Add the Stackdriver monitoring and logging agent to the instances running the code.

Unattempted

In Stackdriver logging, create a new logging metric with the required filters, edit the application code to set the metric value when needed, and create an alert in Stackdriver based on the new metric. is not right.

You don't need to edit the application code to send the metric values. The application already pushes error logs whenever the application times out. Since you already have the required entries in the Stackdriver logs, you don't need to edit the application code to send the metric values. You just need to create metrics from log data.

Ref: <https://cloud.google.com/logging>

Create a custom monitoring metric in code, edit the application code to set the metric value when needed, create an alert in Stackdriver based on the new metric. is not right. You don't create a custom monitoring metric in code. Stackdriver Logging allows you to easily create metrics from log data. Since the application already pushes error logs to Stackdriver Logging, we just need to create metrics from log data in Stackdriver Logging.

Ref: <https://cloud.google.com/logging>

Add the Stackdriver monitoring and logging agent to the instances running the code. is not right.

The Stackdriver Monitoring agent gathers system and application metrics from your VM instances and sends them to Monitoring. In order to make use of this approach, you need application metrics but our application doesn't generate metrics. It just logs errors whenever the upload times out and these are then ingested to Stackdriver logging. We can update our application to enable custom metrics for these scenarios, but that is a lot more work than creating metrics from log data in Stackdriver Logging

Ref: <https://cloud.google.com/logging>

In Stackdriver Logging, create a custom monitoring metric from log data and create an alert in Stackdriver based on the new metric. is the right answer.

Our application adds entries to error logs whenever the application times out during image upload and these logs are ingested to Stackdriver Logging. Since we already have the required data in logs, we just need to create metrics from this log data in Stackdriver Logging. And we can then set up an alert based on this metric. We can trigger an alert if the number of occurrences of the relevant error message is greater than a predefined value. Based on the alert, you can manually add more compute resources.

Ref: <https://cloud.google.com/logging>

57. Question

You developed an application that lets users upload statistical files and subsequently run analytics on this data. You chose to use Google Cloud Storage and BigQuery respectively for these requirements as they are highly available and scalable. You have a docker image for your application code, and you plan to deploy on your on-premises Kubernetes clusters. Your on-prem Kubernetes cluster needs to connect to Google Cloud Storage and BigQuery and you want to do this in a secure way following Google recommended practices. What should you do?

- Create a new service account, with editor permissions, generate and download a key. Use the key to authenticate inside the application.
- Use the default service account for App Engine, which already has the required permissions.
- Use the default service account for Compute Engine, which already has the required permissions.

- Create a new service account, grant it the least viable privileges to the required services, generate and download a JSON key. Use the JSON key to authenticate inside the application.

Unattempted

Use the default service account for Compute Engine, which already has the required permissions. is not right.

The Compute Engine default service account is created with the Cloud IAM project editor role

Ref: https://cloud.google.com/compute/docs/access/service-accounts#default_service_account

The project editor role includes all viewer permissions, plus permissions for actions that modify state, such as changing existing resources. Using a service account that is over-privileged falls foul of the principle of least privilege. Google recommends you enforce the principle of least privilege by ensuring that members have only the permissions that they actually need.

Ref: <https://cloud.google.com/iam/docs/understanding-roles>

Use the default service account for App Engine, which already has the required permissions. is not right.

App Engine default service account has the Editor role in the project (Same as the default service account for Compute Engine).

Ref: <https://cloud.google.com/appengine/docs/standard/python/service-account>

The project editor role includes all viewer permissions, plus permissions for actions that modify state, such as changing existing resources. Using a service account that is over-privileged falls foul of the principle of least privilege. Google recommends you enforce the principle of least privilege by ensuring that members have only the permissions that they actually need.

Ref: <https://cloud.google.com/iam/docs/understanding-roles>

Create a new service account, with editor permissions, generate and download a key. Use the key to authenticate inside the application. is not right.

The project editor role includes all viewer permissions, plus permissions for actions that modify state, such as changing existing resources. Using a service account that is over-privileged falls foul of the principle of least privilege. Google recommends you enforce the principle of least privilege by ensuring that members have only the permissions that they actually need.

Ref: <https://cloud.google.com/iam/docs/understanding-roles>

Create a new service account, grant it the least viable privileges to the required services, generate and download a JSON key. Use the JSON key to authenticate inside the application. is the right answer.

Using a new service account with just the least viable privileges for the required services follows the principle of least privilege. To use a service account outside of Google Cloud, such as on other platforms or on-premises, you must first establish the identity of the service account. Public/private key pairs provide a secure way of accomplishing this goal. Once you have the key, you can use it in your application to authenticate connections to Cloud Storage and BigQuery.

Ref: https://cloud.google.com/iam/docs/creating-managing-service-account-keys#creating_service_account_keys

Ref: <https://cloud.google.com/iam/docs/recommender-overview>

58. Question

You developed an application that reads objects from a cloud storage bucket. You followed GCP documentation and created a service account with just the permissions to read objects from the cloud storage bucket. However, when your application uses this service account, it fails to read objects from the bucket. You suspect this might be an issue with the permissions assigned to the service account. You would like to authenticate a gsutil session with the service account credentials, reproduce the issue yourself and identify the root cause. How can you authenticate gsutil with service account credentials?

- Create JSON keys for the service account and execute `gcloud auth activate-service-account --key-file [KEY_FILE]`
- Create JSON keys for the service account and execute `gcloud auth service-account --key-file [KEY_FILE]`
- Create JSON keys for the service account and execute `gcloud authenticate service-account --key-file [KEY_FILE]`
- Create JSON keys for the service account and execute `gcloud authenticate activate-service-account --key-file [KEY_FILE]`

Unattempted

Create JSON keys for the service account and execute `gcloud authenticate activate-service-account --key-file [KEY_FILE]`. is not right.

`gcloud` doesn't support using "authenticate" to grant/revoke credentials for Cloud SDK. The correct service is "auth".

Ref: <https://cloud.google.com/sdk/gcloud/reference/auth>

Create JSON keys for the service account and execute `gcloud authenticate service-account --key-file [KEY_FILE]`. is not right.

`gcloud` doesn't support using "authenticate" to grant/revoke credentials for Cloud SDK.

The correct service is “auth”.

Ref: <https://cloud.google.com/sdk/gcloud/reference/auth>

Create JSON keys for the service account and execute `gcloud auth service-account --key-file [KEY_FILE]`. is not right.

`gcloud auth` does not support `service-account` action. The correct action to authenticate a service account is `activate-service-account`.

Ref: <https://cloud.google.com/sdk/gcloud/reference/auth/activate-service-account>

Create JSON keys for the service account and execute `gcloud auth activate-service-account --key-file [KEY_FILE]`. is the right answer.

This command correctly authenticates access to Google Cloud Platform with a service account using its JSON key file. To allow `gcloud` (and other tools in Cloud SDK) to use service account credentials to make requests, use this command to import these credentials from a file that contains a private authorization key, and activate them for use in `gcloud`

Ref: <https://cloud.google.com/sdk/gcloud/reference/auth/activate-service-account>

59. Question

You developed an application to serve production users and you plan to use Cloud SQL to host user state data which is very critical for the application flow. You want to protect your user state data from zone failures. What should you do?

- Create a Failover replica in the same region but in a different zone.
- Create a Read replica in the same region but in a different zone.
- Configure High Availability (HA) for Cloud SQL and Create a Failover replica in the same region but in a different zone.
- Configure High Availability (HA) for Cloud SQL and Create a Failover replica in a different region

Unattempted

Create a Read replica in the same region but in a different zone. is not right.

Read replicas do not provide failover capability. To provide failover capability, you need to configure Cloud SQL Instance for High Availability.

Ref: <https://cloud.google.com/sql/docs/mysqlreplication>

Create a Read replica in a different region. is not right.

Read replicas do not provide failover capability. To provide failover capability, you need to configure Cloud SQL Instance for High Availability.

Ref: <https://cloud.google.com/sql/docs/mysql/replication>

Configure High Availability (HA) for Cloud SQL and Create a Failover replica in a different region. is not right.

A Cloud SQL instance configured for HA is called a regional instance because it's primary and secondary instances are in the same region. They are located in different zones but within the same region. It is not possible to create a Failover replica in a different region.

Ref: <https://cloud.google.com/sql/docs/mysql/high-availability>

Configure High Availability (HA) for Cloud SQL and Create a Failover replica in the same region but in a different zone. is the right answer.

If a HA-configured instance becomes unresponsive, Cloud SQL automatically switches to serving data from the standby instance. The HA configuration provides data redundancy.

A Cloud SQL instance configured for HA has instances in the primary zone (Master node) and secondary zone (standby/failover node) within the configured region. Through synchronous replication to each zone's persistent disk, all writes made to the primary instance are also made to the standby instance. If the primary goes down, the standby/failover node takes over and your data continues to be available to client applications.

Ref: <https://cloud.google.com/sql/docs/mysql/high-availability>

60. Question

You have 32 GB of data in a single file that you need to upload to a Nearline Storage bucket. The WAN connection you are using is rated at 1 Gbps, and you are the only one on the connection. You want to use as much of the rated 1 Gbps as possible to transfer the file rapidly. How should you upload the file?

- Use the GCP Console to transfer the file instead of gsutil.
- Change the storage class of the bucket from Nearline to Multi-Regional.
- **Enable parallel composite uploads using gsutil on the file transfer.**
- Decrease the TCP window size on the machine initiating the transfer.

Unattempted

Requirements – transfer the file rapidly, use as much of the rated 1 Gbps as possible

Use the GCP Console to transfer the file instead of gsutil. is not right.

GCP Console does not offer any specific features that help in improving the upload speed.

Decrease the TCP window size on the machine initiating the transfer. is not right.

By decreasing the TCP window size, you are reducing the chunks of data sent in the TCP window, and this has the effect of underutilizing your bandwidth and can slow down the upload.

Change the storage class of the bucket from Nearline to Multi-Regional. is not right.

Multi-Regional is not a storage class. It is a bucket location. You can transition between storage classes but that does not improve the upload speed.

<https://cloud.google.com/storage/docs/locations>

<https://cloud.google.com/storage/docs/storage-classes>

Enable parallel composite uploads using gsutil on the file transfer. is the right answer.

With cloud storage, Object composition can be used for uploading an object in parallel: you can divide your data into multiple chunks, upload each chunk to a distinct object in parallel, compose your final object, and delete any temporary source objects. This helps maximize your bandwidth usage and ensures the file is uploaded as fast as possible.

Ref: <https://cloud.google.com/storage/docs/composite-objects#uploads>

61. Question

You have a collection of audio/video files over 80GB each that you need to migrate to Google Cloud Storage. The files are in your on-premises data center. What migration method can you use to help speed up the transfer process?

- Use parallel uploads to break the file into smaller chunks then transfer it simultaneously.
- Use multithreaded uploads using the -m option.
- Use the Cloud Transfer Service to transfer.
- Start a recursive upload.

Unattempted

Use parallel uploads to break the file into smaller chunks then transfer it simultaneously. is the right answer.

With cloud storage, Object composition can be used for uploading an object in parallel: you can divide your data into multiple chunks, upload each chunk to a distinct object in

parallel, compose your final object, and delete any temporary source objects. This helps maximize your bandwidth usage and ensures the file is uploaded as fast as possible.

Ref: <https://cloud.google.com/storage/docs/composite-objects#uploads>

Use multithreaded uploads using the -m option. is not right.

Using the -m option lets you upload multiple files at the same time, but in our case, the individual files are over 80GB each. The best upload speed can be achieved by breaking the file into smaller chunks and transferring it simultaneously.

Use the Cloud Transfer Service to transfer. is not right.

Cloud Transfer Service is used for transferring massive amounts (in the range of petabytes of data) of data to the cloud. While nothing stops us from using Cloud Transfer Service to upload our files, it would be an overkill and very expensive.

Ref: <https://cloud.google.com/products/data-transfer>

Start a recursive upload. is not right.

In Google Cloud Storage, there is no such thing as a recursive upload.

62. Question

You have a compute engine instance running a production application. You want to receive an email when the instance consumes more than 90% of its CPU resources for more than 15 minutes. You want to use Google services. What should you do?

- 1. Create a Stackdriver Workspace and associate your GCP project with it.
 2. Write a script that monitors the CPU usage and sends it as a custom metric to Stackdriver.
 3. Create an uptime check for the instance in Stackdriver.
- 1. Create a consumer Gmail Account
 2. Write a script that monitors the CPU usage.
 3. When the CPU usage exceeds the threshold, have the script send an email using the Gmail account and smtp.gmail.com on port 25 as SMTP server.
- 1. Create a Stackdriver Workspace and associate your Google Cloud Platform (GCP) project with it
 2. Create an Alerting Policy in Stackdriver that uses the threshold as a trigger condition.
 3. Configure your email address in the notification channel.
- 1. In Stackdriver logging, create a logs based metric to extract the CPU usage by using a regular expression.
 2. In Stackdriver Monitoring, create an Alerting Policy based on this metric
 3. Configure your email address in the notification channel.

Unattempted

We want to use Google services. So that eliminates the two options where we Write a script. Why would we want to write a script when there is a Google service that does exactly that – with minimal configuration!!

Stackdriver logging does not log CPU usage. (Stackdriver monitoring does that) So that rules out the option In Stackdriver logging, create a logs based metric to extract the CPU usage by using a regular expression.

Ref: <https://cloud.google.com/logging/>

1. Create a Stackdriver Workspace and associate your Google Cloud Platform (GCP) project with it
2. Create an Alerting Policy in Stackdriver that uses the threshold as a trigger condition.
3. Configure your email address in the notification channel.

is the right answer.

A Workspace is a tool for monitoring resources contained in one or more Google Cloud projects or AWS accounts. In our case, we create a Stackdriver workspace and link our project to this workspace.

Ref: <https://cloud.google.com/monitoring/workspaces>

Stackdriver monitoring captures the CPU usage. By default, the Monitoring agent collects disk, CPU, network, and process metrics. You can also have the agent send custom metrics to Stackdriver monitoring.

Ref: <https://cloud.google.com/monitoring/>

You can then set up an alerting policy to alert with CPU utilization exceeds 90% for 15 minutes.

Ref: <https://cloud.google.com/monitoring/alerts/>. See here for an example of setting up an alerting policy on CPU load. In our case, we'd have to substitute the CPU load for the CPU utilization metric. <https://cloud.google.com/monitoring/quickstart-lamp>

Stack driver monitoring supports multiple notification options for triggering alerts; email is one of them. Ref: <https://cloud.google.com/monitoring/support/notification-options>

63. Question

You have a developer laptop with Cloud SDK installed on Ubuntu. The cloud SDK was installed from Google Cloud Ubuntu package repository. You want to test your application locally on your laptop with Cloud Datastore. What should you do?

- Create a Cloud Datastore index using `gcloud datastore indexes create`
- Install the `google-cloud-sdk-datastore-emulator` component using the `apt get install` command.
- Export Cloud Datastore data using `gcloud datastore export`
- **Install the `cloud-datastore-emulator` component using the `gcloud components install` command.**

Unattempted

Export Cloud Datastore data using `gcloud datastore export` is not right.
By all means, you can export a copy of all or a subset of entities from Google Cloud Datastore to another storage system such as Google Cloud Storage but your application is configured to connect to a Cloud Datastore instance, not another system that stores a raw dump of exported data. So this option is not right.

Create a Cloud Datastore index using `gcloud datastore indexes create`. is not right.
You could create an index but this doesn't help your application emulate connections to Cloud Datastore on your laptop. So this option is not right.

Install the `google-cloud-sdk-datastore-emulator` component using the `apt get install` command. is not right.
There is no such thing as `google-cloud-sdk-datastore-emulator`; and you don't install `gcloud` components using `apt get`. So this option is not right.

Install the `cloud-datastore-emulator` component using the `gcloud components install` command. is the right answer.
The Datastore emulator provides local emulation of the production Datastore environment.
You can use the emulator to develop and test your application locally
Ref: <https://cloud.google.com/datastore/docs/tools/datastore-emulator>

64. Question

You have a development project with appropriate IAM roles defined. You are creating a production project and want to have the same IAM roles on the new project, using the fewest possible steps. What should you do?

- Use gcloud iam roles copy and specify your organization as the destination organization.
- Use gcloud iam roles copy and specify the production project as the destination project.
- In the Google Cloud Platform Console, use the create role from role functionality.
- In the Google Cloud Platform Console, use the create role functionality and select all applicable permissions.

Unattempted

Our requirements are to create the same iam roles in a different (production) project with the fewest possible steps.

In the Google Cloud Platform Console, use the 'create role from role' functionality. is not right.

This creates a role in the same (development) project, not in the production project. So this doesn't meet our requirement to create same iam roles in production project

In the Google Cloud Platform Console, use the 'create role' functionality and select all applicable permissions. is not right.

This creates a role in the same (development) project, not in the production project. So this doesn't meet our requirement to create same iam roles in production project

Use gcloud iam roles copy and specify your organization as the destination organization. is not right.

We can optionally specify a destination organization but since our requirement is to copy the roles into "production project" (i.e. project, not organization), this option does not meet our requirement to create same iam roles in production project

Ref: <https://cloud.google.com/sdk/gcloud/reference/iam/roles/copy>

Use gcloud iam roles copy and specify the production project as the destination project. is the right answer.

This is the only option that fits our requirements. You copy the roles into the destination project using gcloud iam roles copy and by specifying the production project destination project.

\$gcloud iam roles copy -source "<>" -destination <> -dest-project <>

Ref: <https://cloud.google.com/sdk/gcloud/reference/iam/roles/copy>

65. Question

You have a Dockerfile that you need to deploy on Kubernetes Engine. What should you do?

- Use kubectl app deploy .
- Create a docker image from the Dockerfile and upload it to Container Registry. Create a Deployment YAML file to point to that image. Use kubectl to create the deployment with that file.
- Use gcloud app deploy .
- Create a docker image from the Dockerfile and upload it to Cloud Storage. Create a Deployment YAML file to point to that image. Use kubectl to create the deployment with that file.

Unattempted

Use kubectl app deploy . is not right.

kubectl does not accept app as a verb. Kubectl can deploy a configuration file using kubectl deploy.

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#apply>

Use gcloud app deploy . is not right.

gcloud app deploy – Deploys the local code and/or configuration of your app to App Engine. gcloud app deploy accepts a flag –image-url which is the docker image but it can't directly use a docker file.

Ref: <https://cloud.google.com/sdk/gcloud/reference/app/deploy>

Create a docker image from the Dockerfile and upload it to Cloud Storage. Create a Deployment YAML file to point to that image. Use kubectl to create the deployment with that file. is not right.

You can not upload a docker image to cloud storage. They can only be pushed to a Container Registry (e.g. GCR, Dockerhub etc.)

Ref: <https://cloud.google.com/container-registry/docs/pushing-and-pulling>

Create a docker image from the Dockerfile and upload it to Container Registry. Create a Deployment YAML file to point to that image. Use kubectl to create the deployment with that file. is the right answer.

Once you have a docker image, you can push it to the container register. You can then create a deployment YAML file pointing to this image and use kubectl apply -f to deploy this to the Kubernetes cluster. This assumes you already have a Kubernetes cluster and you gcloud environment is set up to talk to this container by executing gcloud container

```
clusters get-credentials -zone=
```

Ref: <https://cloud.google.com/container-registry/docs/pushing-and-pulling>

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#apply>

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/get-credentials>

66. Question

You have a Google Cloud Platform account with access to both production and development projects. You need to create an automated process to list all compute instances in development and production projects on a daily basis. What should you do?

- Create two configurations using gcloud config. Write a script that sets configurations as active, individually. For each configuration, use gcloud compute instances list to get a list of compute resources.
- Create two configurations using gsutil config. Write a script that sets configurations as active, individually. For each configuration, use gsutil compute instances list to get a list of compute resources.
- Go to Cloud Shell and export this information to Cloud Storage on a daily basis.
- Go to GCP Console and export this information to Cloud SQL on a daily basis.

Unattempted

Go to Cloud Shell and export this information to Cloud Storage on a daily basis. is not right.

You want an automated process but this is a manual activity that needs to be executed daily.

Go to GCP Console and export this information to Cloud SQL on a daily basis. is not right.
You want an automated process but this is a manual activity that needs to be executed daily.

Create two configurations using gsutil config. Write a script that sets configurations as active, individually. For each configuration, use gsutil compute instances list to get a list of compute resources. is not right.

The gsutil config command applies to users who have installed gsutil as a standalone tool and is used for obtaining access credentials for Cloud Storage and writes a boto/gsutil configuration file containing the obtained credentials along with a number of other

configuration-controllable values.

Ref: <https://cloud.google.com/storage/docs/gsutil/commands/config>

It is not used for creating Gcloud configurations. You use gcloud config to do that.

<https://cloud.google.com/sdk/gcloud/reference/config/configurations/create>

Create two configurations using gcloud config. Write a script that sets configurations as active, individually. For each configuration, use gcloud compute instances list to get a list of compute resources. is the right answer.

You can create two configurations – one for the development project and another for the production project. And you do that by running “gcloud config configurations create” command.

<https://cloud.google.com/sdk/gcloud/reference/config/configurations/create>

In your custom script, you can load these configurations one at a time and execute gcloud compute instances list to list Google Compute Engine instances in the project that is active in the gcloud configuration.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/list>

Once you have this information, you can export it in a suitable format to a suitable target e.g. export as CSV or export to Cloud Storage/BigQuery/SQL, etc.

67. Question

You have a large 5-TB AVRO file stored in a Cloud Storage bucket. Your analysts are proficient only in SQL and need access to the data stored in this file. You want to find a cost-effective way to complete their request as soon as possible. What should you do?

- Load data in Cloud Datastore and run a SQL query against it.
- Create a BigQuery table and load data in BigQuery. Run a SQL query on this table and drop this table after you complete your request.
- Create external tables in BigQuery that point to Cloud Storage buckets and run a SQL query on these external tables to complete your request.
- Create a Hadoop cluster and copy the AVRO file to NDFS by compressing it. Load the file in a hive table and provide access to your analysts so that they can run SQL queries.

Unattempted

Load data in Cloud Datastore and run a SQL query against it. is not right.

Datastore is a highly scalable NoSQL database and although it supports SQL like queries, it doesn't support SQL. Moreover, there is no out of the box way for transforming AVRO file from cloud storage into the Cloud Datastore entity. So we have to do in a bespoke way

which adds to our cost and time.

Ref: <https://cloud.google.com/datastore>

Create a Hadoop cluster and copy the AVRO file to NDFS by compressing it. Load the file in a hive table and provide access to your analysts so that they can run SQL queries. is not right.

Like Cloud Datastore, Hive doesn't directly support SQL, it provides HiveQL (HQL) which is SQL like. In addition, the process of creating a Hadoop cluster and getting the data eventually into a hive table is time-consuming and adds to our cost and time.

Create a BigQuery table and load data in BigQuery. Run a SQL query on this table and drop this table after you complete your request. is not right.

Like the above two, while it is possible to build a solution that transforms and loads data into the target, BigQuery in this case, is not a trivial process and involves cost and time. GCP provides an out of the box way to query AVRO files from Cloud Storage and this should be preferred.

Create external tables in BigQuery that point to Cloud Storage buckets and run a SQL query on these external tables to complete your request. is the right answer.

BigQuery supports querying Cloud Storage data in a number of formats such as CSV, JSON, AVRO, etc. You do this by creating a Big Query external table that points to a Cloud Storage data source (bucket). This solution works out of the box, involves minimal effort, minimal cost, and is quick.

<https://cloud.google.com/bigquery/external-data-cloud-storage>

68. Question

You have a Linux VM that must connect to Cloud SQL. You created a service account with the appropriate access rights. You want to make sure that the VM uses this service account instead of the default Compute Engine service account. What should you do?

- Download a JSON Private Key for the service account. On the Custom Metadata of the VM, add that JSON as the value for the key compute-engine-service-account
- Download a JSON Private Key for the service account. After creating the VM, ssh into the VM and save the JSON under ~/.gcloud/compute-engine-service-account.json

- When creating the VM via the web console, specify the service account under the 'Identity and API Access' section.

- Download a JSON Private Key for the service account. On the Project Metadata, add that JSON as the value for the key compute-engine-serviceaccount

Unattempted

When creating the VM via the web console, specify the service account under the 'Identity and API Access' section. is the right answer.

You can set the service account at the time of creating the compute instance. You can also update the service account used by the instance – this requires that you stop the instance first and then update the service account. Setting/Updating the service account can be done either via the web console or by executing gcloud command or by the REST API. See below an example for updating the service account through gcloud command.

```
gcloud compute instances set-service-account instance-1 --zone=us-central1-a --service-account=my-new-service-account@gcloud-gcp-ace-lab-266520.iam.gserviceaccount.com  
Updated [https://www.googleapis.com/compute/v1/projects/gcloud-gcp-ace-lab-266520/zones/us-central1-a/instances/instance-1].
```

Download a JSON Private Key for the service account. On the Project Metadata, add that JSON as the value for the key compute-engine-serviceaccount is not right.

While updating the service account for a compute instance can be done through the console, gcloud or the REST API, they don't do it based on the JSON Private Key.

Download a JSON Private Key for the service account. On the Custom Metadata of the VM, add that JSON as the value for the key compute-engine-service-account. is not right.

While updating the service account for a compute instance can be done through the console, gcloud or the REST API, they don't do it based on the JSON Private Key.

Download a JSON Private Key for the service account. After creating the VM, ssh into the VM and save the JSON under ~/.gcloud/compute-engine-service-account.json is not right. You can configure a VM to use a certain service account by providing the relevant JSON credentials file, but the procedure is different. Copying the JSON file to a specific path alone is not sufficient, moreover, the path mentioned is wrong as well. See below for a use case where a VM which is unable to list cloud storage buckets is updated to use a service account and it can then list the buckets.

Prior to using a service account. Use gsutil ls to list buckets and it fails.

```
$ gsutil ls
```

```
 ServiceException: 401 Anonymous caller does not have storage.buckets.list access to project 393066724129.
```

Within the VM, execute the command below to use the service account. (Assumes that you have created a service account that provides the necessary permissions and have copied it over the VM)

```
gcloud auth activate-service-account admin-service-account@gcloud-gcp-ace-  
266520.iam.gserviceaccount.com --key-file=~/compute-engine-service-account.json  
Activated service account credentials for: [admin-service-account@gcloud-gcp-ace-  
266520.iam.gserviceaccount.com]
```

The output above doesn't show this, but the credentials are written to the file /home/gcloud_gcp_ace_user/.config/gcloud/legacy_credentials/admin-service-account@gcloud-gcp-ace-266520.iam.gserviceaccount.com/adc.json

Now, use gsutil ls again to list buckets and it works.

```
$ gsutil ls  
gs://test-gcloud-gcp-ace-2020-bucket-1/  
gs://test-gcloud-gcp-ace-2020-bucket-2/
```

69. Question

You have a number of applications that have bursty workloads and are heavily dependent on topics to decouple publishing systems from consuming systems. Your company would like to go serverless to enable developers to focus on writing code without worrying about infrastructure. Your solution architect has already identified Cloud Pub/Sub as a suitable alternative for decoupling systems. You have been asked to identify a suitable GCP Serverless service that is easy to use with Cloud Pub/Sub. You want the ability to scale down to zero when there is no traffic in order to minimize costs. You want to follow Google recommended practices. What should you suggest?

- Cloud Run for Anthos
- **Cloud Functions**
- App Engine Standard
- Cloud Run

Unattempted

GCP serverless compute portfolio includes 4 services, which are all listed in the answer options. Our requirements are to identify a GCP serverless service that

1. Lets us scale down to 0
2. Integrates with Cloud Pub/Sub seamlessly

Cloud Run for Anthos. is not right.

Among the four options, App Engine Standard, Cloud Functions and Cloud Run can all scale down to zero. Cloud Run for Anthos can scale the pods down to zero but the number of nodes per cluster can not scale to zero so these nodes are billed in the absence of requests. This rules out Cloud Run for Anthos.

App Engine Standard. is not right.

App Engine Standard doesn't offer an out of the box integration with Cloud Pub/Sub. We can use the Cloud Client Library to send and receive Pub/Sub messages as described in the note below but the key point to note is the absence of out of the box integration with Cloud Pub/Sub so this rules out App Engine Standard

Ref: <https://cloud.google.com/appengine/docs/standard/nodejs/writing-and-responding-to-pub-sub-messages>

Cloud Run. is not right.

Cloud Run is an excellent product and integrates with Cloud Pub/Sub for several use cases. For example, every time a new .csv file is created inside a Cloud Storage bucket, an event is fired and delivered via a Pub/Sub subscription to a Cloud Run service. The Cloud Run service extracts data from the file and stores it as structured data into a BigQuery table.

Ref: <https://cloud.google.com/run#section-7>

At the same time, we want to follow Google recommended practices. Google doesn't list integration with Cloud Pub/Sub as a key feature of Cloud Run. Contrary to this, Google says "If you're building a simple API (a small set of functions to be accessed via HTTP or Cloud Pub/Sub), we recommend using Cloud Functions."

Cloud Functions. is the right answer.

Cloud Functions is Google Cloud's event-driven serverless compute platform that lets you run your code locally or in the cloud without having to provision servers. Cloud Functions scales up or down, so you pay only for compute resources you use. Cloud Functions have excellent integration with Cloud Pub/Sub, lets you scale down to zero and is recommended by Google as the ideal serverless platform to use when dependent on Cloud Pub/Sub.

"If you're building a simple API (a small set of functions to be accessed via HTTP or Cloud Pub/Sub), we recommend using Cloud Functions."

Ref: <https://cloud.google.com/serverless-options>

70. Question

You have a number of compute instances belonging to an unmanaged instances group. You need to SSH to one of the Compute Engine instances to run an ad hoc script. You've already authenticated gcloud, however, you don't have an SSH key deployed yet. In the fewest steps possible, what's the easiest way to SSH to the instance?

- Create a key with the ssh-keygen command. Upload the key to the instance. Run gcloud compute instances list to get the IP address of the instance, then use the ssh command.
- Run gcloud compute instances list to get the IP address of the instance, then use the ssh command.
- Create a key with the ssh-keygen command. Then use the gcloud compute ssh command.
- Use the gcloud compute ssh command.

Unattempted

Create a key with the ssh-keygen command. Upload the key to the instance. Run gcloud compute instances list to get the IP address of the instance, then use the ssh command. is not right.

This approach certainly works. You can create a key pair with ssh-keygen, update the instance metadata with the public key and SSH to the instance. But is it the easiest way to SSH to the instance with the fewest possible steps? Let's explore other options to decide (you will see that there is another option that does the same with less effort). You can find more information about this option

here: <https://cloud.google.com/compute/docs/instances/adding-removing-ssh-keys#block-project-keys>

Create a key with the ssh-keygen command. Then use the gcloud compute ssh command. is not right.

This works but is more work (having to create the key) than the answer. gcloud compute ssh ensures that the user's public SSH key is present in the project's metadata. If the user does not have a public SSH key, one is generated using ssh-keygen and added to the project's metadata.

Run gcloud compute instances list to get the IP address of the instance, then use the ssh command. is not right.

We can get the IP of the instance by executing the gcloud compute instances list but unless an SSH is generated and updated in project metadata, you would not be able to SSH to the instance. User access to a Linux instance through third-party tools is determined by which public SSH keys are available to the instance. You can control the public SSH keys that are available to a Linux instance by editing metadata, which is where your public SSH keys and related information are stored.

Ref: <https://cloud.google.com/compute/docs/instances/adding-removing-ssh-keys#block-project-keys>

Use the gcloud compute ssh command. is the right answer.

gcloud compute ssh ensures that the user's public SSH key is present in the project's metadata. If the user does not have a public SSH key, one is generated using ssh-keygen

and added to the project's metadata. This is similar to the other option where we copy the key explicitly to the project's metadata but here it is done automatically for us. There are also security benefits with this approach. When we use gcloud compute ssh to connect to Linux instances, we are adding a layer of security by storing your host keys as guest attributes. Storing SSH host keys as guest attributes improve the security of your connections by helping to protect against vulnerabilities such as man-in-the-middle (MITM) attacks. On the initial boot of a VM instance, if guest attributes are enabled, Compute Engine stores your generated host keys as guest attributes. Compute Engine then uses these host keys that were stored during the initial boot to verify all subsequent connections to the VM instance.

Ref: <https://cloud.google.com/compute/docs/instances/connecting-to-instance>

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/ssh>

SET-2

1. Question

You have a project for your App Engine application that serves a development environment. The required testing has succeeded and you want to create a new project to serve as your production environment. What should you do?

- Deploy your application again using gcloud and specify the project parameter with the new project name to create the new project.
- Use gcloud to create the new project and to copy the deployed application to the new project.
- Use gcloud to create the new project, and then deploy your application to the new project.
- Create a Deployment Manager configuration file that copies the current App Engine deployment into a new project.

Unattempted

Use gcloud to create the new project and to copy the deployed application to the new project. is not right.

You can use gcloud to create a new project but you can not copy a deployed application from one project to another. This feature is not offered by Google App Engine.

Create a Deployment Manager configuration file that copies the current App Engine deployment into a new project. is not right.

The deployment manager configuration file contains configuration about the resources that need to be created in Google cloud, however, it does not offer the feature to copy app engine deployment into a new project.

Deploy your application again using gcloud and specify the project parameter with the new project name to create the new project. is not right.

You can deploy using gcloud app deploy and target it to a different project using -project flag. However, you can only deploy to an existing project as the gcloud app deploy command is unable to create a new project if it doesn't already exist.

Use gcloud to create the new project, and then deploy your application to the new project. is the right answer.

You can deploy to a different project by using -project flag.

By default, the service is deployed the current project configured via:

```
$ gcloud config set core/project PROJECT
```

To override this value for a single deployment, use the `-project` flag:

```
$ gcloud app deploy ~/my_app/app.yaml --project=PROJECT
```

Ref: <https://cloud.google.com/sdk/gcloud/reference/app/deploy>

2. Question

You have a single binary application that you want to run on Google Cloud Platform. You decided to automatically scale the application based on underlying infrastructure CPU usage. Your organizational policies require you to use virtual machines directly. You need to ensure that the application scaling is operationally efficient and completed as quickly as possible. What should you do?

- Create a Google Kubernetes Engine cluster, and use horizontal pod autoscaling to scale the application.
- Create an instance template, and use the template in a managed instance group with autoscaling configured.
- Create an instance template, and use the template in a managed instance group that scales up and down based on the time of day.
- Use a set of third-party tools to build automation around scaling the application up and down, based on Stackdriver CPU usage monitoring

Unattempted

Our requirements are

1. Use Virtual Machines directly (i.e. not container-based)
2. Scale Automatically
3. Scaling is efficient & is quick

Create a Google Kubernetes Engine cluster, and use horizontal pod autoscaling to scale the application. is not right.

We want to use virtual machines directly. And although GKE uses virtual machines under the hood for its GKE cluster, the autoscaling is totally different – it uses scaling at VMs (cluster auto-scaling) as well as at pod level (horizontal and vertical pod autoscaling).

Create an instance template, and use the template in a managed instance group that scales up and down based on the time of day. is not right.

Scaling based on time of the day may be insufficient especially when there is a sudden surge of requests (causing high CPU utilization) or if the requests go down suddenly (resulting in low CPU usage). Our requirements state we need to scale automatically i.e.

we need autoscaling solution that scales up and down based on CPU usage which is indicative of the volume of requests processed but scaling based on time of the day is not indicative of the load (CPU) on the system and is therefore not right.

Use a set of third-party tools to build automation around scaling the application up and down, based on Stackdriver CPU usage monitoring. is not right.

While this can be done, it is not the most efficient solution when Google's own services offer this functionality and can do it more efficiently as they are all natively integrated.

Create an instance template, and use the template in a managed instance group with autoscaling configured. is the right answer.

Managed instance groups offer autoscaling capabilities that let you automatically add or delete instances from a managed instance group based on increases or decreases in load (CPU Utilization in this case). Autoscaling helps your apps gracefully handle increases in traffic and reduce costs when the need for resources is lower. You define the autoscaling policy and the autoscaler performs automatic scaling based on the measured load (CPU Utilization in this case). Autoscaling works by adding more instances to your instance group when there is more load (upscale), and deleting instances when the need for instances is lowered (downscale).

Ref: <https://cloud.google.com/compute/docs/autoscaler>

3. Question

You have a virtual machine that is currently configured with 2 vCPUs and 4 GB of memory. It is running out of memory. You want to upgrade the virtual machine to have 8GB of memory. What should you do?

- Stop the VM, increase the memory to 8 GB and start the VM
- Rely on live migration to move the workload to a machine with more memory.
- Stop the VM, change the machine type to n1-standard-2 and start the VM
- Use gcloud to add metadata to the VM. Set the key to required-memory-size and the value to 8 GB

Unattempted

Rely on live migration to move the workload to a machine with more memory. is not right.

Live migration migrates your running instances to another host in the same zone so that Google can perform maintenance such as a software or hardware update. It can not be used for changing machine type.

Ref: <https://cloud.google.com/compute/docs/instances/live-migration>

Use gcloud to add metadata to the VM. Set the key to required-memory-size and the value to 8 GB. is not right.

There is no such setting as required-memory-size.

Stop the VM, change the machine type to n1-standard-2 and start the VM. is not right. n1-standard-2 instance offers less than 8 GB (7.5 GB to be precise) so this falls short of the required memory.

Ref: <https://cloud.google.com/compute/docs/machine-types>

Stop the VM, increase the memory to 8 GB and start the VM. is the right answer.

In Google compute engine, if predefined machine types don't meet your needs, you can create an instance with custom virtualized hardware settings. Specifically, you can create an instance with a custom number of vCPUs and custom memory, effectively using a custom machine type. Custom machine types are ideal for the following scenarios:

1. Workloads that aren't a good fit for the predefined machine types that are available to you.
2. Workloads that require more processing power or more memory but don't need all of the upgrades that are provided by the next machine type level.

In our scenario, we only need a memory upgrade. Moving to a bigger instance would also bump up the CPU which we don't need so we have to use a custom machine type. It is not possible to change memory while the instance is running so you need to first stop the instance, change the memory and then start it again. See below a screenshot that shows how CPU/Memory can be customized for an instance that has been stopped.

Ref: <https://cloud.google.com/compute/docs/instances/creating-instance-with-custom-machine-type>

4. Question

You have a web application deployed as a managed instance group based on an instance template. You modified the startup script used in the instance template and would like the existing instances to pick up changes from the new startup scripts. Your web application is currently serving live web traffic. You want to propagate the startup script changes to all instances in the managed instances group while minimizing effort,

minimizing cost and ensuring that the available capacity does not decrease. What would you do?

- Delete instances in the managed instance group (MIG) one at a time and rely on auto-healing to provision an additional instance.
- Perform a rolling-action replace with max-unavailable set to 0 and max-surge set to 1
- Create a new managed instance group (MIG) based on a new template. Add the group to the backend service for the load balancer. When all instances in the new managed instance group are healthy, delete the old managed instance group
- Perform a rolling-action start-update with max-unavailable set to 1 and max-surge set to 0

Unattempted

Perform a rolling-action start-update with max-unavailable set to 1 and max-surge set to 0. is not right.

You can carry out a rolling action start update to fully replace the template by executing a command like

```
gcloud compute instance-groups managed rolling-action start-update instance-group-1  
--zone=us-central1-a --version template=instance-template-1 --canary-version  
template=instance-template-2,target-size=100%
```

which updates the instance-group-1 to use instance-template-2 instead of instance-template-1 and have instances created out of instance-template-2 serve 100% of traffic. However, the values specified for maxSurge and maxUnavailable mean that we will lose capacity which is against our requirements.

maxSurge specifies the maximum number of instances that can be created over the desired number of instances. If maxSurge is set to 0, the rolling update can not create additional instances and is forced to update existing instances. This results in a reduction in capacity and therefore does not satisfy our requirement to ensure that the available capacity does not decrease during the deployment.

maxUnavailable – specifies the maximum number of instances that can be unavailable during the update process. When maxUnavailable is set to 1, the rolling update updates 1 instance at a time. i.e. it takes 1 instance out of service, updates it, and puts it back into service. This results in a reduction in capacity while the instance is out of service. Example – if we have 10 instances in service, this combination of setting results in 1 instance at a time taken out of service for replacement while the remaining 9 continue to serve live traffic. That's a reduction of 10% in available capacity.

Ref: <https://kubernetes.io/docs/concepts/workloads/controllers/deployment/#max-unavailable>

Ref: <https://kubernetes.io/docs/concepts/workloads/controllers/deployment/#max-surge>

Create a new managed instance group (MIG) based on a new template. Add the group to the backend service for the load balancer. When all instances in the new managed instance group are healthy, delete the old managed instance group. is not right. While the end result is the same, we have a period of time where the traffic is served by instances from both the old managed instances group (MIG) which doubles our cost and increases effort and complexity.

Delete instances in the managed instance group (MIG) one at a time and rely on auto-healing to provision an additional instance. is not right.

While this would result in the same eventual outcome, there are two issues with this approach. First, deleting an instance one at a time would result in a reduction in capacity which is against our requirements. Secondly, deleting instances manually one at a time is error-prone and time-consuming. One of our requirements is to “minimize the effort” but deleting instances manually and relying on auto-healing health checks to provision them back is time-consuming and could take a lot of time depending on the number of instances in the MIG and the startup scripts executed during bootstrap.

Perform a rolling-action replace with max-unavailable set to 0 and max-surge set to 1. is the right answer.

This option achieves the outcome in the most optimal manner. The replace action is used to replace instances in a managed instance group. When maxUnavailable is set to 0, the rolling update can not take existing instances out of service. And when maxSurge is set to 1, we let the rolling update spin a single additional instance. The rolling update then puts the additional instance into service and takes one of the existing instances out of service for replacement. There is no reduction in capacity at any point in time.

Ref: <https://kubernetes.io/docs/concepts/workloads/controllers/deployment/#max-unavailable>

Ref: <https://kubernetes.io/docs/concepts/workloads/controllers/deployment/#max-surge>

Ref: <https://cloud.google.com/sdk/gcloud/reference/alpha/compute/instance-groups/managed/rolling-action/replace>

5. Question

You have a web application deployed as a managed instance group. You have a new version of the application to gradually deploy. Your web application is currently serving live web traffic. You want to ensure that the available capacity does not decrease during the deployment. What would you do?

- Create a new instance template with the new application version. Update the existing managed instance group with the new instance template. Delete the instances in the managed instance group to allow the managed instance group to recreate the instance using the new instance template.
- Perform a rolling-action start-update with maxSurge set to 0 and maxUnavailable set to 1
- Create a new managed instance group with an updated instance template. Add the group to the backend service for the load balancer. When all instances in the new managed instance group are healthy, delete the old managed instance group.
- Perform a rolling-action start-update with maxSurge set to 1 and maxUnavailable set to 0

Unattempted

Our requirements are

1. Deploy a new version gradually
2. Ensure available capacity does not decrease during deployment

Create a new managed instance group with an updated instance template. Add the group to the backend service for the load balancer. When all instances in the new managed instance group are healthy, delete the old managed instance group. is not right.
First of all instance templates can not be updated. So the phrase updated instance template rules out this option.

Ref: <https://cloud.google.com/compute/docs/instance-templates/>

Create a new instance template with the new application version. Update the existing managed instance group with the new instance template. Delete the instances in the managed instance group to allow the managed instance group to recreate the instance using the new instance template. is not right.

If we follow these steps, we end up with a full fleet of instances belonging to the new managed instances group (i.e. based on the new template) behind the load balancer, but our requirement to gradually deploy the new version is not met. In addition, deleting the existing instances of the managed instance group would almost certainly result in an outage to our application which is not desirable when we are serving live web traffic.

Perform a rolling-action start-update with maxSurge set to 0 and maxUnavailable set to 1 is not right.

maxSurge specifies the maximum number of instances that can be created over the desired number of instances. If maxSurge is set to 0, the rolling update can not create additional instances and is forced to update existing instances. This results in a reduction in capacity and therefore does not satisfy our requirement to ensure that the available capacity does not decrease during the deployment.

`maxUnavailable` – specifies the maximum number of instances that can be unavailable during the update process. When `maxUnavailable` is set to 1, the rolling update updates 1 instance at a time. i.e. it takes 1 instance out of service, updates it, and puts it back into service. This results in a reduction in capacity while the instance is out of service. Example – if we have 10 instances in service, this combination of setting results in 1 instance at a time taken out of service for an upgrade while the remaining 9 continue to serve live traffic. That's a reduction of 10% in available capacity and does not satisfy our requirement to ensure that the available capacity does not decrease during the deployment.

Perform a rolling-action start-update with `maxSurge` set to 1 and `maxUnavailable` set to 0 is the right answer.

This is the only option that satisfies our two requirements – deploying gradually and ensuring the available capacity does not decrease. When `maxUnavailable` is set to 0, the rolling update can not take existing instances out of service. And when `maxSurge` is set to 1, we let the rolling update spin a single additional instance. The rolling update then puts the additional instance into service and takes one of the existing instances out of service for the upgrade. There is no reduction in capacity at any point in time. And the rolling upgrade upgrades 1 instance at a time so we gradually deploy the new version. Example – if we have 10 instances in service, this combination of setting results in 1 additional instance put into service (resulting in 11 instances serving traffic), then a older instance taken out of service (resulting in 10 instances serving traffic) and puts the upgraded instance back into service (resulting in 11 instances serving traffic). The rolling upgrade continues updating the remaining 9 instances one at a time. Finally, when all 10 instances have been upgraded, the additional instance that is spun up is deleted. We still have 10 instances serving live traffic but now on the new version of code.

Ref: <https://kubernetes.io/docs/concepts/workloads/controllers/deployment/#max-unavailable>

Ref: <https://kubernetes.io/docs/concepts/workloads/controllers/deployment/#max-surge>

6. Question

You have a web application deployed as a managed instance group. You noticed some of the compute instances are running low on memory. You suspect this is due to JVM memory leak and you want to restart the compute instances to reclaim the leaked memory. Your web application is currently serving live web traffic. You want to ensure that the available capacity does not go below 80% at any time during the restarts and you want to do this at the earliest. What would you do?

- Stop instances in the managed instance group (MIG) one at a time and rely on autohealing to bring them back up.
- Perform a rolling-action replace with max-unavailable set to 20%.
- **Perform a rolling-action restart with max-unavailable set to 20%.**
- Perform a rolling-action reboot with max-surge set to 20%.

Unattempted

Perform a rolling-action reboot with max-surge set to 20%. is not right.
reboot is not a supported action for rolling updates. The supported actions are replace, restart, start-update and stop-proactive-update.

Ref: <https://cloud.google.com/sdk/gcloud/reference/beta/compute/instance-groups/managed/rolling-action>

Perform a rolling-action replace with max-unavailable set to 20%. is not right.
Performing a rolling-action replace – Replaces instances in a managed instance group. While this resolves the JVM memory leak issue, recreating the instances is a little drastic when the same result can be achieved with the simple restart action. One of our requirements is to “do this at the earliest” but recreating instances might take a lot of time depending on the number of instances and startup scripts; certainly more time than restart action.

Ref: <https://cloud.google.com/sdk/gcloud/reference/beta/compute/instance-groups/managed/rolling-action>

Stop instances in the managed instance group (MIG) one at a time and rely on autohealing to bring them back up. is not right.
While this would result in the same eventual outcome, it is manual, error-prone and time-consuming. One of our requirements is to “do this at the earliest” but stopping instances manually is time-consuming and could take a lot of time depending on the number of instances in the MIG. Also, relying on autohealing health checks to detect the failure and spin up the instance adds to the delay.

Perform a rolling-action restart with max-unavailable set to 20%. is the right answer.
This option achieves the outcome in the most optimal manner. The restart action restarts instances in a managed instance group. By performing a rolling restart with max-unavailable set to 20%, the rolling update restarts instances while ensuring there is at least 80% available capacity. The rolling update carries on restarting all the remaining instances until all instances in the MIG have been restarted.

Ref: <https://cloud.google.com/sdk/gcloud/reference/alpha/compute/instance-groups/managed/rolling-action/restart>

7. Question

You have a website hosted on App Engine standard environment. You want 1% of your users to see a new test version of the website. You want to minimize complexity. What should you do?

- Create a new App Engine application in the same project. Deploy the new version in that application. Configure your network load balancer to send 1% of the traffic to that new application.
- Create a new App Engine application in the same project. Deploy the new version in that application. Use the App Engine library to proxy 1% of the requests to the new version.
- Deploy the new version in the same application and use the `--migrate` option.
- Deploy the new version in the same application and use the `--splits` option to give a weight of 99 to the current version and a weight of 1 to the new version.

Unattempted

Deploy the new version in the same application and use the `--migrate` option. is not right.
`migrate` is not a valid flag for the `gcloud app deploy` command.

Ref: <https://cloud.google.com/sdk/gcloud/reference/app/deploy>

Also, `gcloud app versions migrate`, which is a valid command to migrate traffic from one version to another for a set of services, is not suitable either as we only want to send 1% traffic.

<https://cloud.google.com/sdk/gcloud/reference/app/versions/migrate>

Create a new App Engine application in the same project. Deploy the new version in that application. Use the App Engine library to proxy 1% of the requests to the new version. is not right.

While this can be done, we are increasing complexity and do not meet our requirement “minimize complexity”. There is an out of the box option in the app engine to split traffic in a seamless way.

Create a new App Engine application in the same project. Deploy the new version in that application. Configure your network load balancer to send 1% of the traffic to that new application. is not right.

Instances that participate as backend VMs for network load balancers must be running the appropriate Linux guest environment, Windows guest environment, or other processes that provide equivalent functionality. Network load balancer is not suitable for App Engine

standard environment which is container-based and provide us specific runtimes without any promise on the underlying guest environments.

Deploy the new version in the same application and use the `-splits` option to give a weight of 99 to the current version and a weight of 1 to the new version. is the right answer.

You can use traffic splitting to specify a percentage distribution of traffic across two or more of the versions within a service. Splitting traffic allows you to conduct A/B testing between your versions and provides control over the pace when rolling out features.

For this scenario, we can split the traffic as shown below, sending 1% to v2 and 99% to v1

by executing the command `gcloud app services set-traffic service1 -splits v2=1,v1=99`

Ref: <https://cloud.google.com/sdk/gcloud/reference/app/services/set-traffic>

8. Question

You have a workload running on Compute Engine that is critical to your business. You want to ensure that the data on the boot disk of this workload is backed up regularly. You need to be able to restore a backup as quickly as possible in case of disaster. You also want older backups to be cleaned automatically to save on cost. You want to follow Google-recommended practices. What should you do?

- Create a Cloud Function to create an instance template.
- **Create a snapshot schedule for the disk using the desired interval.**
- Create a cron job to create a new disk from the disk using gcloud.
- Create a Cloud Task to create an image and export it to Cloud Storage.

Unattempted

Create a Cloud Function to create an instance template. is not right.

This does not fulfil our requirement of backing up data on boot disk 'regularly'.

Create a cron job to create a new disk from the disk using gcloud. is not right.

Like above, this does not fulfil our requirement of backing up data on boot disk 'regularly'.

Create a Cloud Task to create an image and export it to Cloud Storage. is not right.

Like above, this does not fulfil our requirement of backing up data on boot disk 'regularly'.

Create a snapshot schedule for the disk using the desired interval. is the right answer.
Create snapshots to periodically back up data from your zonal persistent disks or regional persistent disks. To reduce the risk of unexpected data loss, consider the best practice of setting up a snapshot schedule to ensure your data is backed up on a regular schedule.

Ref: <https://cloud.google.com/compute/docs/disks/create-snapshots>

You can also delete snapshots on a schedule by defining a snapshot retention policy. A snapshot retention policy defines how long you want to keep your snapshots. If you choose to set up a snapshot retention policy, you must do so as part of your snapshot schedule.

Ref: https://cloud.google.com/compute/docs/disks/scheduled-snapshots#retention_policy

9. Question

You have an application deployed in a GKE Cluster as a Kubernetes workload with Daemon Sets. Your application has become very popular and is now struggling to cope up with increased traffic. You want to add more pods to your workload and want to ensure your cluster scales up and scales down automatically based on volume. What should you do?

- Perform a rolling update to modify machine type from n1-standard-2 to n1-standard-4.
- **Enable autoscaling on Kubernetes Engine.**
- Enable Horizontal Pod Autoscaling for the Kubernetes deployment.
- Create another identical Kubernetes workload and split traffic between the two workloads.

Unattempted

Enable Horizontal Pod Autoscaling for the Kubernetes deployment. is not right.
Horizontal Pod Autoscaling can not be enabled for Daemon Sets, this is because there is only one instance of a pod per node in the cluster. In a replica deployment, when Horizontal Pod Autoscaling scales up, it can add pods to the same node or another node within the cluster. Since there can only be one pod per node in the Daemon Set workload, Horizontal Pod Autoscaling is not supported with Daemon Sets.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/daemonset>

Create another identical Kubernetes cluster and split traffic between the two workloads. is not right.

Creating another identical Kubernetes cluster is going to double your costs; at the same time, there is no guarantee that this is enough to handle all the traffic. Finally, it doesn't satisfy our requirement of "cluster scales up and scales down automatically"

Perform a rolling update to modify machine type from n1-standard-2 to n1-standard-4. is not right.

While increasing the machine type from n1-standard-2 to n1-standard-4 gives the existing nodes more resources and processing power, we don't know if that would be enough to handle the increased volume of traffic. Also, it doesn't satisfy our requirement of "cluster scales up and scales down automatically"

Ref: <https://cloud.google.com/compute/docs/machine-types>

Enable autoscaling on Kubernetes Engine. is the right answer.

GKE's cluster autoscaler automatically resizes the number of nodes in a given node pool, based on the demands of your workloads. As you add nodes to a node pool, DaemonSets automatically add Pods to the new nodes as needed. DaemonSets attempt to adhere to a one-Pod-per-node model.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/cluster-autoscaler>

10. Question

You have an application on a general-purpose Compute Engine instance that is experiencing excessive disk read throttling on its Zonal SSD Persistent Disk. The application primarily reads large files from disk. The disk size is currently 350 GB. You want to provide the maximum amount of throughput while minimizing costs. What should you do?

- Increase the size of the disk to 1 TB.
- Increase the allocated CPU to the instance.
- Migrate to use a Local SSD on the instance.
- Migrate to use a Regional SSD on the instance.

Unattempted

Migrate to use a Regional SSD on the instance. is not right.

Migrating to a regional SSD would actually make it worse. At the time of writing, the Read IOPS for a Zonal standard persistent disks is 7,500 and the Read IOPS reduces to 3000 for a Regional standard persistent disks which reduces the throughput.

Ref: <https://cloud.google.com/compute/docs/disks/performance>

Increase the size of the disk to 1 TB. is not right.

The performance of SSD persistent disks scales with the size of the disk.

Ref: https://cloud.google.com/compute/docs/disks/performance#cpu_count_size

However, there is no guarantee that increasing the disk to 1 TB will increase the throughput in this scenario as disk performance also depends on the number of vCPUs on VM instance.

Ref: https://cloud.google.com/compute/docs/disks/performance#ssd_persistent_disk_performance_by_disk_size

Ref: <https://cloud.google.com/compute/docs/disks/performance#machine-type-disk-limits>

For example, consider a 1,000 GB SSD persistent disk attached to an instance with an N2 machine type and 4 vCPUs. The read limit based solely on the size of the disk is 30,000 IOPS. However, because the instance has 4 vCPUs, the read limit is restricted to 15,000 IOPS.

Increase the allocated CPU to the instance. is not right.

In Compute Engine, machine types are grouped and curated for different workloads. Each machine type is subject to specific persistent disk limits per vCPU. Increasing the vCPU count increases the Read IOPS

<https://cloud.google.com/compute/docs/disks/performance#machine-type-disk-limits>

However, there is no guarantee that increasing CPU will definitely increase the throughput in this scenario as disk performance could be limited by disk size.

Ref: https://cloud.google.com/compute/docs/disks/performance#ssd_persistent_disk_performance_by_disk_size

Ref: <https://cloud.google.com/compute/docs/disks/performance#machine-type-disk-limits>

For example, consider a 1,000 GB SSD persistent disk attached to an instance with an N2 machine type and 48 vCPUs.

The read limit based solely on the vCPU count is 60,000 IOPS. However, because the instance has 1000 GB SSD, the read limit is restricted to 30,000 IOPS.

Migrate to use a Local SSD on the instance. is the right answer.

Local SSDs are physically attached to the server that hosts your VM instance. Local SSDs have higher throughput and lower latency than standard persistent disks or SSD persistent disks. The performance gains from local SSDs require certain trade-offs in availability, durability, and flexibility. Because of these trade-offs, Local SSD storage isn't automatically replicated and all data on the local SSD might be lost if the instance terminates for any reason.

Ref: <https://cloud.google.com/compute/docs/disks#localssds>

Ref: https://cloud.google.com/compute/docs/disks/performance#type_comparison

11. Question

You have an application running in App Engine standard environment. You want to add a custom C# library to enhance the functionality of this application. However, C# isn't supported by App Engine standard. You want to maintain the serverless aspect of your application. What should you do? Choose 2 answers.

- Containerize your new application and deploy it to a Cloud Run on GKE environment.
- Containerize your new application and deploy it to a Cloud Run environment.
- Containerize your new application and deploy it to a App Engine flexible environment.
- Containerize your new application and deploy it to a Google Kubernetes Engine environment.
- Split your application into different functions. Deploy your application as separate cloud functions in Google Cloud Functions (GCP) environment.

Unattempted

App engine standard currently supports Python, Java, Node.js, PHP, Ruby and Go.

Ref: <https://cloud.google.com/appengine/docs/standard/>

The question already states C# isn't supported by App Engine. Our requirement is to ensure we maintain the serverless aspect of our application.

Split your application into different functions. Deploy your application as separate cloud functions in Google Cloud Functions (GCP) environment is not right.

Cloud Functions is a serverless platform where you can run the code in the cloud without having to provision servers. You split your application functionality into multiple functions, and each of these is defined as a cloud function. Cloud Functions don't support C#.

Supported runtimes are Python, Node.js and Go.

Ref: <https://cloud.google.com/functions>

Containerize your new application and deploy it to a App Engine flexible environment is not right.

While App Engine flexible lets us customize runtimes or provide our own runtime by supplying a custom Docker image or Dockerfile from the open-source community, it uses compute engine virtual machines so it is not serverless.

Ref: <https://cloud.google.com/appengine/docs/flexible/>

Containerize your new application and deploy it to a Google Kubernetes Engine environment. is not right.

GKE i.e. Google Kubernetes Clusters uses compute engine virtual machines so it is not

serverless.

Ref: <https://cloud.google.com/kubernetes-engine>

Containerize your new application and deploy it to a Cloud Run environment. is the right answer.

Cloud Run is a fully managed compute platform that automatically scales your stateless containers. Cloud Run is serverless: it abstracts away all infrastructure management, so you can focus on what matters most—building great applications. Run your containers in fully managed Cloud Run or on Anthos, which supports both Google Cloud and on-premises environments. Cloud Run is built upon an open standard, Knative, enabling the portability of your applications.

Ref: <https://cloud.google.com/run>

Containerize your new application and deploy it to a Cloud Run on GKE environment. is the right answer.

Cloud Run implements the Knative serving API, an open-source project to run serverless workloads on top of Kubernetes. That means you can deploy Cloud Run services anywhere Kubernetes runs. And if you need more control over your services (like access to GPU or more memory), you can also deploy these serverless containers in your own GKE cluster instead of using the fully managed environment. When using the fully managed environment, Cloud Run on GKE is serverless.

Ref: <https://cloud.google.com/blog/products/serverless/cloud-run-bringing-serverless-to-containers>

12. Question

You have an application running in Google Kubernetes Engine (GKE) with cluster autoscaling enabled. This application exposes a TCP endpoint. There are several replicas of the application. You have a Compute Engine instance in the same region but in another Virtual Private Cloud (VPC) called pt-network that has no overlapping CIDR range with the other VPC. The instance needs to connect to the application on GKE. You want to minimize effort. What should you do?

- 1. In GKE, create a service of type LoadBalancer that uses the application's pods as backend. 2. Set the service's externalTrafficPolicy to Cluster. 3. Configure the Compute Engine instance to use the address of the load balancer that has been assigned.

- 1. In GKE, create a service of type NodePort that uses the application's pods as backend. 2. Create a Compute Engine instance called proxy with 2 network interfaces one in VPC. 3. Use iptables on the instance to forward traffic from pt-network to the GKE nodes. 4. Configure the Compute Engine instance to use the address of proxy in pt-network as endpoint.
- 1. In GKE, create a Service of type LoadBalancer that uses the application's Pods as backend. 2. Add an annotation to this service cloud.google.com/load-balancer-type: Internal 3. Peer the two VPCs together 4. Configure the Compute Engine instance to use the address of the load balancer that has been created.
- 1. In GKE, create a Service of type LoadBalancer that uses the application's Pods as backend. 2. Add a Cloud Armor Security Policy to the load balancer that whitelists the internal IPs of the MIG's instances. 3. Configure the Compute Engine instance to use the address of the load balancer that has been created.

Unattempted

While it may be possible to set up the networking to let the compute engine instance in pt-network communicate with pods in the GKE cluster in multiple ways, we need to look for an option that minimizes effort. Generally speaking, this means using Google Cloud Platform services directly and configuring them to achieve the intended outcome; over setting up a service ourselves, installing/managing/upgrading it ourselves which is manual, error-prone, time-consuming and add to operational overhead.

1. In GKE, create a service of type LoadBalancer that uses the application's pods as backend.
2. Set the service's externalTrafficPolicy to Cluster.
3. Configure the Compute Engine instance to use the address of the load balancer that has been assigned. is not right.

In GKE, services are used to expose pods to the outside world. There are multiple types of services. The three common types are – NodePort, ClusterIP, and LoadBalancer (there are two more service types – ExternalName and Headless which are not relevant in this context). We do not want to create a Cluster IP as this is not accessible outside the cluster. And we do not want to create NodePort as this results in exposing a port on each node in the cluster; and as we have multiple replicas, this will result in them trying to open the same port on the nodes which fail. The compute engine instance in pt-network needs a single point of communication to reach GKE. This is achieved by creating a service of type LoadBalancer. This gives the service a public IP that is externally accessible.

Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/exposing-apps>

externalTrafficPolicy denotes how the service should route external traffic – including public access. Rather than trying to explain, I'll point you to a very good blog that does a great job of answering how this works. <https://www.asykim.com/blog/deep-dive-into-kubernetes-external-traffic-policies>

Since we have cluster autoscaling enabled, we can have more than 1 node and possibly multiple replicas running on each node. So externalTrafficPolicy set to Cluster plays well

with our requirement.

Finally, we configure the compute engine to use the (externally accessible) address of the load balancer.

So this certainly looks like an option, but is it the best option that minimizes effort? One of the disadvantages of this option is that it exposes the pods publicly by using a service of type LoadBalancer. We want our compute engine to talk to the pods, but do we really want to expose our pods to the whole world? Maybe not!! Let's look at the other options to find out if there is something more relevant and secure.

1. In GKE, create a service of type NodePort that uses the application's pods as backend.
2. Create a Compute Engine instance called proxy with 2 network interfaces one in VPC.
3. Use iptables on the instance to forward traffic from pt-network to the GKE nodes.
4. Configure the Compute Engine instance to use the address of proxy in pt-network as endpoint. is not right.

For reasons explained in the above option, we don't want to create a service of type NodePort. This opens up a port on each node for each replica (pod). If we choose to do this, the compute engine doesn't have a single point to contact. Instead, it would need to contact the GKE cluster nodes individually – and that is bound to have issues because we have autoscaling enabled and the nodes may scale up and scale down as per the scaling requirements. New nodes may have different IP addresses to the previous nodes, so unless the Compute engine is continuously supplied with the IP addresses of the nodes, it can't reach them. Moreover, we have multiple replicas and it is possible we might have multiple replicas of the pod on the same node in which case they all can't open the same node port – once a node port is opened by one replica (pod), it can't be used by other replicas on the same node. So this option can be ruled out without going into the rest of the answer.

1. In GKE, create a Service of type LoadBalancer that uses the application's Pods as backend.
2. Add a Cloud Armor Security Policy to the load balancer that whitelists the internal IPs of the MIG's instances.
3. Configure the Compute Engine instance to use the address of the load balancer that has been created. is not right.

Creating a service of type LoadBalancer and getting a Compute Engine instance to use the address of the load balancer is fine, but Cloud Armor is not required. You could certainly use Cloud Armor to set up a whitelist policy to only let traffic through from the compute engine instance, but hang on – this option says "MIG instances". We don't have a managed instance group. The question mentions a single instance but not MIG. If we were to assume the single instance is part of a MIG, i.e. a MIG with a single instance, this option works too. It is more secure than the first option discussed in the explanation but at the same time more expensive. Let's look at the other option to see if it provides a secure yet cost-effective way of achieving the same.

1. In GKE, create a Service of type LoadBalancer that uses the application's Pods as backend.
2. Add an annotation to this service cloud.google.com/load-balancer-type: Internal
3. Peer the two VPCs together
4. Configure the Compute Engine instance to use the address of the load balancer that has been created. is the right answer.

Creating a service of type LoadBalancer and getting a Compute Engine instance to use the address of the load balancer is fine. We covered this previously in the first option in the explanations section.

Adding the annotation cloud.google.com/load-balancer-type: Internal makes your cluster's services accessible to applications outside of your cluster that use the same VPC network and are located in the same Google Cloud region. So this improves security by not allowing public access, however, the compute engine is located in a different VPC so it can't access.

Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/internal-load-balancing>

But peering the VPCs together enables the compute engine to access the load balancer IP. And peering is possible because they do not use overlapping IP ranges. Peering essentially links up the two VPCs and resources inside the VPCs can communicate with each other as if they were all in a single VPC. More info about VPC

peering: <https://cloud.google.com/vpc/docs/vpc-peering>

So this option is the right answer. It provides a secure and cost-effective way of achieving our requirements. There are several valid answers but this option is more correct than the others.

13. Question

You have an application that looks for its licensing server on the IP 10.0.3.21. You need to deploy the licensing server on the Compute Engine. You do not want to change the configuration of the application and want the application to be able to reach the licensing server. What should you do?

- Use the IP 10.0.3.21 as a custom ephemeral IP address and assign it to the licensing server.
- Start the licensing server with an automatic ephemeral IP address, and then promote it to a static internal IP address.
- Reserve the IP 10.0.3.21 as a static public IP address using gcloud and assign it to the licensing server.
- Reserve the IP 10.0.3.21 as a static internal IP address using gcloud and assign it to the licensing server.

Unattempted

Reserve the IP 10.0.3.21 as a static public IP address using gcloud and assign it to the licensing server. is not right.

The private network range is defined by IETF (Ref: <https://tools.ietf.org/html/rfc1918>) and includes 10.0.0.0/8. So all IP Addresses from 10.0.0.0 to 10.255.255.255 belong to this internal IP range. As the IP of interest 10.0.3.21 falls within this range, it can not be reserved as a public IP Address.

Use the IP 10.0.3.21 as a custom ephemeral IP address and assign it to the licensing server. is not right.

Ephemeral IP address is the public IP Address assigned to compute instance. An ephemeral external IP address is an IP address that doesn't persist beyond the life of the resource. When you create an instance or forwarding rule without specifying an IP address, the resource is automatically assigned an ephemeral external IP address.

Ref: <https://cloud.google.com/compute/docs/ip-addresses#ephemeraladdress>

The private network range is defined by IETF (Ref: <https://tools.ietf.org/html/rfc1918>) and includes 10.0.0.0/8. So all IP Addresses from 10.0.0.0 to 10.255.255.255 belong to this internal IP range. As the IP of interest 10.0.3.21 falls within this range, it can not be used as a public IP Address (ephemeral IP is public).

Start the licensing server with an automatic ephemeral IP address, and then promote it to a static internal IP address. is not right.

When a compute instance is started with public IP, it gets an ephemeral IP address. An ephemeral external IP address is an IP address that doesn't persist beyond the life of the resource.

Ref: <https://cloud.google.com/compute/docs/ip-addresses#ephemeraladdress>

You can promote this ephemeral address into a Static IP address but this will be an external IP address and not an internal one.

Ref: https://cloud.google.com/compute/docs/ip-addresses/reserve-static-external-ip-address#promote_ephemeral_ip

Reserve the IP 10.0.3.21 as a static internal IP address using gcloud and assign it to the licensing server. is right.

This is the only option that lets us reserve IP 10.0.3.21 as a static internal IP address because it falls within the standard IP Address range as defined by IETF (Ref: <https://tools.ietf.org/html/rfc1918>). This includes the range 10.0.0.0/8 so all IP Addresses from 10.0.0.0 to 10.255.255.255 belong to this internal IP range. Since we can now reserve this IP Address as a static internal IP address, it can be assigned to the licensing server in the VPC so that the application is able to reach the licensing server.

14. Question

You have an application that receives SSL-encrypted TCP traffic on port 443. Clients for this application are located all over the world. You want to minimize latency for the clients. Which load balancing option should you use?

- HTTPS Load Balancer
- **Network Load Balancer**
- SSL Proxy Load Balancer
- Internal TCP/UDP Load Balancer. Add a firewall rule allowing ingress traffic from 0.0.0.0/0 on the target instances.

Unattempted

SSL Proxy Load Balancer. is not right.

Google says “SSL Proxy Load Balancing is intended for non-HTTP(S) traffic. For HTTP(S) traffic, we recommend that you use HTTP(S) Load Balancing.”

So this option can be ruled out.

Ref: <https://cloud.google.com/load-balancing/docs/ssl>

Internal TCP/UDP Load Balancer. Add a firewall rule allowing ingress traffic from 0.0.0.0/0 on the target instances. is not right.

Internal TCP Load Balancing is a regional load balancer that enables you to run and scale your services behind an internal load balancing IP address that is accessible only to your internal virtual machine (VM) instances. Since we need to serve public traffic, this load balancer is not suitable for us.

Ref: <https://cloud.google.com/load-balancing/docs/internal>

HTTPS Load Balancer. is not right.

The HTTPS load balancer terminates TLS in locations that are distributed globally, so as to minimize latency between clients and the load balancer. If you require geographic control over where TLS is terminated (which is our scenario with clients located all over the world), you should use Google Cloud Network Load Balancing instead, and terminate TLS on backends that are located in regions appropriate to your needs.

Ref: <https://cloud.google.com/load-balancing/docs/https#control-tls-termination>

Network Load Balancer. is the right answer.

Google Cloud external TCP/UDP Network Load Balancing (after this referred to as Network Load Balancing) is a regional, non-proxied load balancer. The network load balancer supports any and all ports. You can use Network Load Balancing to load balance

TCP and UDP traffic. Because the load balancer is a pass-through load balancer, your backends terminate the load-balanced TCP connection or UDP packets themselves. For example, you might run an HTTPS web server on your backends (which is our scenario) and use a Network Load Balancing to route requests to it, terminating TLS on your backends themselves.

Ref: <https://cloud.google.com/load-balancing/docs/network>

Also, the latency is minimized when using network load balancer. Because load balancing takes place in-region and traffic is merely forwarded, there is no significant latency impact compared with the no-load-balancer option.

Ref: https://cloud.google.com/load-balancing/docs/tutorials/optimize-app-latency#network_load_balancing

15. Question

You have an application that uses Cloud Spanner as a backend database. The application has a very predictable traffic pattern. You want to automatically scale up or down the number of Spanner nodes depending on traffic. What should you do?

- Create a cron job that runs on a scheduled basis to review stackdriver monitoring metrics, and then resize the Spanner instance accordingly.
- Create a Stackdriver alerting policy to send an alert to oncall SRE emails when Cloud Spanner CPU exceeds the threshold. SREs would scale resources up or down accordingly.
- Create a Stackdriver alerting policy to send an alert to Google Cloud Support email when Cloud Spanner CPU exceeds your threshold. Google support would scale resources up or down accordingly.
- Create a Stackdriver alerting policy to send an alert to webhook when Cloud Spanner CPU is over or under your threshold. Create a Cloud Function that listens to HTTP and resizes Spanner resources accordingly.

Unattempted

Create a cron job that runs on a scheduled basis to review stackdriver monitoring metrics, and then resize the Spanner instance accordingly. is not right.

While this works and does it automatically , it does not follow Google's recommended practices.

Ref: <https://cloud.google.com/spanner/docs/instances> “Note: You can scale the number of nodes in your instance based on the Cloud Monitoring metrics on CPU or storage utilization in conjunction with Cloud Functions.”

Create a Stackdriver alerting policy to send an alert to oncall SRE emails when Cloud Spanner CPU exceeds the threshold. SREs would scale resources up or down accordingly. is not right.

This does not follow Google's recommended practices.

Ref: <https://cloud.google.com/spanner/docs/instances> "Note: You can scale the number of nodes in your instance based on the Cloud Monitoring metrics on CPU or storage utilization in conjunction with Cloud Functions."

Create a Stackdriver alerting policy to send an alert to Google Cloud Support email when Cloud Spanner CPU exceeds your threshold. Google support would scale resources up or down accordingly. is not right.

This does not follow Google's recommended practices.

Ref: <https://cloud.google.com/spanner/docs/instances> "Note: You can scale the number of nodes in your instance based on the Cloud Monitoring metrics on CPU or storage utilization in conjunction with Cloud Functions."

Create a Stackdriver alerting policy to send an alert to webhook when Cloud Spanner CPU is over or under your threshold. Create a Cloud Function that listens to HTTP and resizes Spanner resources accordingly. is the right answer.

For scaling the number of nodes in Cloud spanner instance, Google recommends implementing this base on the Cloud Monitoring metrics on CPU or storage utilization in conjunction with Cloud Functions.

Ref: <https://cloud.google.com/spanner/docs/instances>

16. Question

You have an application that uses Cloud Spanner as a database backend to keep current state information about users. Cloud Bigtable logs all events triggered by users. You export Cloud Spanner data to Cloud Storage during daily backups. One of your analysts asks you to join data from Cloud Spanner and Cloud Bigtable for specific users. You want to complete this ad hoc request as efficiently as possible. What should you do?

- Create a dataflow job that copies data from Cloud Bigtable and Cloud Storage for specific users.
- Create a Cloud Dataproc cluster that runs a Spark job to extract data from Cloud Bigtable and Cloud Storage for specific users.

- Create two separate BigQuery external tables on Cloud Storage and Cloud Bigtable. Use the BigQuery console to join these tables through user fields, and apply appropriate filters.
- Create a dataflow job that copies data from Cloud Bigtable and Cloud Spanner for specific users.

Unattempted

Our requirements are to join user sessions with user events efficiently. We need to look for an option that is primarily a Google service and provides this feature out of the box or with minimal configuration.

Create two separate BigQuery external tables on Cloud Storage and Cloud Bigtable. Use the BigQuery console to join these tables through user fields, and apply appropriate filters is the right answer.

Big query lets you create tables that reference external data sources such as Bigtable and Cloud Storage. You can then join up these two tables through user fields and apply appropriate filters. You can achieve the end result with minimal configuration using this option.

Ref: <https://cloud.google.com/bigquery/external-data-sources>

Create a dataflow job that copies data from Cloud Bigtable and Cloud Spanner for specific users. is not right.

Cloud dataflow does not support Cloud Spanner. Cloud Dataflow SQL supports reading from Pub/Sub topics, Cloud Storage file sets, and BigQuery tables.

Create a Cloud Dataproc cluster that runs a Spark job to extract data from Cloud Bigtable and Cloud Storage for specific users. is not right.

While it is certainly possible to do this using a Spark job, it is complicated as we would have to come up with the code/logic to extract the data and certainly not straightforward.

Create a dataflow job that copies data from Cloud Bigtable and Cloud Storage for specific users. is not right.

This is possible but it is not as efficient as using Big Query.

Ref: <https://cloud.google.com/dataflow/docs/guides/sql/dataflow-sql-intro>

Here is some more documentation around this option, some of the issues are

1. Dataflow SQL expects CSV files in Cloud Storage filesets. CSV files must not contain a header row with column names; the first row in each CSV file is interpreted as a data record. – but our question doesn't say how the exported data is stored in cloud storage.
2. You can only run jobs in regions that have a Dataflow regional endpoint. Our question doesn't say which region. Ref: <https://cloud.google.com/dataflow/docs/concepts/regional->

endpoints.

3. Creating a Dataflow job can take several minutes – unlike Big Query external tables which can be created very easily.

Too many unknowns. Otherwise, this option is a good option.

Here is some more information if you'd like to get a better understanding of how to use Cloud Dataflow to achieve this result.

Cloud Dataflow SQL lets you use SQL queries to develop and run Dataflow jobs from the BigQuery web UI. You can join streams (such as Pub/Sub) and snapshotted datasets (such as BigQuery tables and Cloud Storage filesets); query your streams or static datasets with SQL by associating schemas with objects, such as tables, Cloud Storage filesets and Pub/Sub topics; and write your results into a BigQuery table for analysis and dashboarding.

Cloud Dataflow SQL supports multiple data sources including Cloud Storage and Big Query tables which are of interest for this scenario.

<https://cloud.google.com/dataflow/docs/guides/sql/data-sources-destinations>

17. Question

You have an instance group that you want to load balance. You want the load balancer to terminate the client SSL session. The instance group is used to serve a public web application over HTTPS. You want to follow Google recommended practices. What should you do?

- Configure an external TCP proxy load balancer.
- Configure an external SSL proxy load balancer.
- Configure an internal TCP load balancer.
- **Configure an HTTP(S) load balancer.**

Unattempted

Configure an internal TCP load balancer. is not right.

Internal TCP Load Balancing is a regional load balancer that enables you to run and scale your services behind an internal load balancing IP address that is accessible only to your internal virtual machine (VM) instances. Since we need to serve public traffic, this load balancer is not suitable for us.

Ref: <https://cloud.google.com/load-balancing/docs/internal>

Configure an external SSL proxy load balancer. is not right.

Google says “SSL Proxy Load Balancing is intended for non-HTTP(S) traffic. For HTTP(S) traffic, we recommend that you use HTTP(S) Load Balancing.”

So this option can be ruled out.

Ref: <https://cloud.google.com/load-balancing/docs/ssl>

Configure an external TCP proxy load balancer. is not right.

Google says "TCP Proxy Load Balancing is intended for non-HTTP traffic. For HTTP traffic, use HTTP Load Balancing instead. For proxied SSL traffic, use SSL Proxy Load Balancing." So this option can be ruled out.

Ref: <https://cloud.google.com/load-balancing/docs/tcp>

Configure an HTTP(S) load balancer. is the right answer.

This is the only option that fits all requirements. It can serve public traffic, can terminate SSL at the load balancer and follows google recommended practices.

? "The backends of a backend service can be either instance groups or network endpoint groups (NEG), but not a combination of both."

? "An external HTTP(S) load balancer distributes traffic from the internet"

? "The client SSL session terminates at the load balancer."

? "For HTTP traffic, use HTTP Load Balancing instead."

Ref: <https://cloud.google.com/load-balancing/docs/https>

18. Question

You have an object in a Cloud Storage bucket that you want to share with an external company. The object contains sensitive data. You want access to the content to be removed after four hours. The external company does not have a Google account to which you can grant specific user-based access privileges. You want to use the most secure method that requires the fewest steps. What should you do?

- Configure the storage bucket as a static website and furnish the object's URL to the company. Delete the object from the storage bucket after four hours.
- Create a new Cloud Storage bucket specifically for the external company to access. Copy the object to that bucket. Delete the bucket after four hours have passed.
- Create a signed URL with a four-hour expiration and share the URL with the company.
- Set object access to 'public' and use object lifecycle management to remove the object after four hours.

Unattempted

Set object access to ‘public’ and use object lifecycle management to remove the object after four hours. is not right.

While the external company can access the public objects from the bucket, it doesn’t stop bad actors from accessing the data as well. Since the data is “sensitive” and we want to follow a “secure method”, we shouldn’t do this.

Configure the storage bucket as a static website and furnish the object’s URL to the company. Delete the object from the storage bucket after four hours. is not right.

The static website is public by default. While the external company can access the objects from the static website, it doesn’t stop bad actors from accessing the data as well. Since the data is “sensitive” and we want to follow a “secure method”, we shouldn’t do this.

Create a new Cloud Storage bucket specifically for the external company to access. Copy the object to that bucket. Delete the bucket after four hours have passed. is not right.

Even if we were to create a separate bucket for the external company to access, since the company does not have a google account, the only way to have them access this separate bucket is by enabling public access which we can’t because of the nature of data (sensitive) and is against standard security practices.

Create a signed URL with a four-hour expiration and share the URL with the company. is the right answer.

This is the only option that fits all requirements. When we generate a signed URL, we can specify an expiry and only users with the signed URL can view/download the objects, and they don’t need a google account.

Ref: <https://cloud.google.com/storage/docs/access-control/signed-urls>

This page provides an overview of signed URLs, which you use to give time-limited resource access to anyone in possession of the URL, regardless of whether they have a Google account.

19. Question

You have annual audits every year and you need to provide external auditors access to the last 10 years of audit logs. You want to minimize the cost and operational overhead while following Google recommended practices. What should you do? (Select Three)

- Grant external auditors Storage Object Viewer role on the logs storage bucket.

- Set a custom retention of 10 years in Stackdriver logging and provide external auditors view access to Stackdriver Logs.
- Export audit logs to Cloud Storage via an audit log export sink.**
- Export audit logs to BigQuery via an audit log export sink.
- Export audit logs to Cloud Filestore via a Pub/Sub export sink.
- Configure a lifecycle management policy on the logs bucket to delete objects older than 10 years**

Unattempted

Export audit logs to Cloud Filestore via a Pub/Sub export sink. is not right.

Storing logs in Cloud Filestore is expensive. In Cloud Filestore, Standard Tier pricing costs \$0.2 per GB per month and Premium Tier pricing costs \$0.3 per GB per month. In comparison, Google Cloud Storage offers several storage classes that are significantly cheaper.

Ref: <https://cloud.google.com/bigquery/pricing>

Ref: <https://cloud.google.com/storage/pricing>

Set a custom retention of 10 years in Stackdriver logging and provide external auditors view access to Stackdriver Logs. is not right.

While it is possible to configure a custom retention period of 10 years in Stackdriver logging, storing logs in Stackdriver is expensive compared to Cloud Storage. Stackdriver charges \$0.01 per GB per month, whereas something like Cloud Storage Coldline Storage costs \$0.007 per GB per month (30% cheaper) and Cloud Storage Archive Storage costs 0.004 per GB per month (60% cheaper than Stackdriver)

Ref: <https://cloud.google.com/logging/docs/storage#pricing>

Ref: <https://cloud.google.com/storage/pricing>

Export audit logs to BigQuery via an audit log export sink. is not right.

Storing logs in BigQuery is expensive. In BigQuery, Active storage costs \$0.02 per GB per month and Long-term storage costs \$0.01 per GB per month. In comparison, Google Cloud Storage offers several storage classes that are significantly cheaper.

Ref: <https://cloud.google.com/bigquery/pricing>

Ref: <https://cloud.google.com/storage/pricing>

Export audit logs to Cloud Storage via an audit log export sink. is the right answer.

Among all the storage solutions offered by Google Cloud Platform, Cloud storage offers the best pricing for long term storage of logs. Google Cloud Storage offers several storage classes such as Nearline Storage (\$0.01 per GB per Month) Coldline Storage (\$0.007 per GB per Month) and Archive Storage (\$0.004 per GB per month) which are significantly cheaper than the storage options covered by the above options above.

Ref: <https://cloud.google.com/storage/pricing>

Grant external auditors Storage Object Viewer role on the logs storage bucket. is the right answer.

You can provide external auditors access to the logs in the bucket by granting the Storage Object Viewer role which allows them to read any object stored in any bucket.

Ref: <https://cloud.google.com/storage/docs/access-control/iam>

Configure a lifecycle management policy on the logs bucket to delete objects older than 10 years. is the right answer.

You need to archive log files for 10 years but you don't need log files older than 10 years. And since you also want to minimize costs, it is a good idea to set up a lifecycle management policy on the bucket to delete objects that are older than 10 years. Lifecycle management configuration is a set of rules which apply to current and future objects in the bucket. When an object meets the criteria of one of the rules, Cloud Storage automatically performs a specified action (delete in this case) on the object.

Ref: <https://cloud.google.com/storage/docs/lifecycle>

20. Question

You have asked your supplier to send you a purchase order and you want to enable them to upload the file to a cloud storage bucket within the next 4 hours. Your supplier does not have a Google account. You want to follow Google recommended practices. What should you do?

- Create a JSON key for the Default Compute Engine Service Account. Execute the command `gsutil signurl -m PUT -d 4h gs:///**`.
- Create a service account with just the permissions to upload files to the bucket. Create a JSON key for the service account. Execute the command `gsutil signurl -httpMethod PUT -d 4h gs:///**`.
- Create a service account with just the permissions to upload files to the bucket. Create a JSON key for the service account. Execute the command `gsutil signurl -m PUT -d 4h gs://po.pdf`.
- Create a service account with just the permissions to upload files to the bucket. Create a JSON key for the service account. Execute the command `gsutil signurl -d 4h gs://`.

Unattempted

Create a service account with just the permissions to upload files to the bucket. Create a JSON key for the service account. Execute the command `gsutil signurl -d 4h gs://`. is not right.

This command creates signed URLs for retrieving existing objects. This command does not specify a HTTP method and in the absence of one, the default HTTP method is GET.

Ref: <https://cloud.google.com/storage/docs/gsutil/commands/signurl>

Create a service account with just the permissions to upload files to the bucket. Create a JSON key for the service account. Execute the command `gsutil signurl -httpMethod PUT -d 4h gs:///**`. is not right.

`gsutil signurl` does not accept `-httpMethod` parameter.

```
$ gsutil signurl -d 4h -httpMethod PUT keys.json gs://gcp-ace-lab-255520/*
```

`CommandException: Incorrect option(s) specified. Usage:`

The HTTP method can be provided through `-m` flag.

Ref: <https://cloud.google.com/storage/docs/gsutil/commands/signurl>

Create a JSON key for the Default Compute Engine Service Account. Execute the command `gsutil signurl -m PUT -d 4h gs:///**`. is not right.

Using the default compute engine service account violates the principle of least privilege. The recommended approach is to create a service account with just the right permissions needed and create JSON keys for this service account to use with `gsutil signurl` command.

Create a service account with just the permissions to upload files to the bucket. Create a JSON key for the service account. Execute the command `gsutil signurl -m PUT -d 4h gs:///po.pdf`. is the right answer.

This command correctly creates a signed url that is valid for 4 hours and allows PUT (through the `-m` flag) operations on the file `po.pdf` in the bucket. The supplier can then use the signed URL to upload a file to this bucket within 4 hours.

Ref: <https://cloud.google.com/storage/docs/gsutil/commands/signurl>

21. Question

You have been asked to create a new Kubernetes Cluster on Google Kubernetes Engine that can autoscale the number of worker nodes as well as pods. What should you do?
(Select 2)

- Create a GKE cluster and enable autoscaling on Kubernetes Engine.
- Create Compute Engine instances for the workers and the master and install Kubernetes. Rely on Kubernetes to create additional Compute Engine instances when needed.

- Create a GKE cluster and enable autoscaling on the instance group of the cluster.
- Configure a Compute Engine instance as a worker and add it to an unmanaged instance group. Add a load balancer to the instance group and rely on the load balancer to create additional Compute Engine instances when needed.

- **Enable Horizontal Pod Autoscaling for the Kubernetes deployment.**

Unattempted

Create a GKE cluster and enable autoscaling on the instance group of the cluster. is not right.

GKE's cluster auto-scaler automatically resizes the number of nodes in a given node pool, based on the demands of your workloads. However, we should not enable Compute Engine autoscaling for managed instance groups for the cluster nodes. GKE's cluster auto-scaler is separate from Compute Engine autoscaling.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/cluster-autoscaler>

Configure a Compute Engine instance as a worker and add it to an unmanaged instance group. Add a load balancer to the instance group and rely on the load balancer to create additional Compute Engine instances when needed. is not right.

When using GKE to manage your Kubernetes clusters, you can not add manually created compute instances to the worker node pool. A node pool is a group of nodes within a cluster that all have the same configuration. Node pools use a NodeConfig specification.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/node-pools>

Moreover, Unmanaged instance groups do not autoscale. An unmanaged instance group is simply a collection of virtual machines (VMs) that reside in a single zone, VPC network, and subnet. An unmanaged instance group is useful for grouping together VMs that require individual configuration settings or tuning.

Ref: <https://cloud.google.com/compute/docs/instance-groups/creating-groups-of-unmanaged-instances>

Create Compute Engine instances for the workers and the master and install Kubernetes. Rely on Kubernetes to create additional Compute Engine instances when needed. is not right.

When using Google Kubernetes Engine, you can not install master node separately. The cluster master runs the Kubernetes control plane processes, including the Kubernetes API server, scheduler, and core resource controllers. The master's lifecycle is managed by GKE when you create or delete a cluster.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/cluster-architecture>

Also, you can not add manually created compute instances to the worker node pool. A node pool is a group of nodes within a cluster that all have the same configuration. Node pools use a NodeConfig specification.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/node-pools>

Create a GKE cluster and enable autoscaling on Kubernetes Engine. is the right answer. GKE's cluster autoscaler automatically resizes the number of nodes in a given node pool, based on the demands of your workloads. You don't need to manually add or remove nodes or over-provision your node pools. Instead, you specify a minimum and maximum size for the node pool, and the rest is automatic. When demand is high, cluster autoscaler adds nodes to the node pool. When demand is low, cluster autoscaler scales back down to a minimum size that you designate. This can increase the availability of your workloads when you need it while controlling costs.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/cluster-autoscaler>

Enable Horizontal Pod Autoscaling for the kubernetes deployment. is the right answer. Horizontal Pod Autoscaler scales up and scales down your Kubernetes workload by automatically increasing or decreasing the number of Pods in response to the workload's CPU or memory consumption, or in response to custom metrics reported from within Kubernetes or external metrics from sources outside of your cluster. Horizontal Pod Autoscaling cannot be used for workloads that cannot be scaled, such as DaemonSets.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/horizontalpodautoscaler>

22. Question

You have been asked to migrate a docker application from datacenter to cloud. Your solution architect has suggested uploading docker images to GCR in one project and running an application in a GKE cluster in a separate project. You want to store images in the project img-278322 and run the application in the project prod-278986. You want to tag the image as acme_track_n_trace:v1. You want to follow Google-recommended practices. What should you do?

- Run gcloud builds submit --tag gcr.io/img-278322/acme_track_n_trace
- Run gcloud builds submit --tag gcr.io/img-278322/acme_track_n_trace:v1
- Run gcloud builds submit --tag gcr.io/prod-278986/acme_track_n_trace
- Run gcloud builds submit --tag gcr.io/prod-278986/acme_track_n_trace:v1

Unattempted

Run gcloud builds submit -tag gcr.io/img-278322/acme_track_n_trace. is not right. This command tags the image as acme_track_n_trace:latest but we want to tag the image as acme_track_n_trace:v1.

Ref: <https://cloud.google.com/sdk/gcloud/reference/builds/submit>

Run `gcloud builds submit -tag gcr.io/prod-278986/acme_track_n_trace`. is not right.
This command tags the image as `acme_track_n_trace:latest` but we want to tag the image as `acme_track_n_trace:v1`. This command also upload the image to the wrong project.
Ref: <https://cloud.google.com/sdk/gcloud/reference/builds/submit>

Run `gcloud builds submit -tag gcr.io/prod-278986/acme_track_n_trace:v1`. is not right.
This command also upload the image to the wrong project.
Ref: <https://cloud.google.com/sdk/gcloud/reference/builds/submit>

Run `gcloud builds submit -tag gcr.io/img-278322/acme_track_n_trace:v1`. is the right answer.
This command correctly tags the image as `acme_track_n_trace:v1` and uploads the image to the `img-278322` project.
Ref: <https://cloud.google.com/sdk/gcloud/reference/builds/submit>

23. Question

You have designed a solution on Google Cloud Platform (GCP) that uses multiple GCP products. Your company has asked you to estimate the costs of the solution. You need to provide estimates for the monthly total cost. What should you do?

- For each GCP product in the solution, review the pricing details on the products pricing page. Use the pricing calculator to total the monthly costs for each GCP product.
- For each GCP product in the solution, review the pricing details on the products pricing page. Create a Google Sheet that summarizes the expected monthly costs for each product.
- Provision the solution on GCP. Leave the solution provisioned for 1 week. Navigate to the Billing Report page in the Google Cloud Platform Console. Multiply the 1 week cost to determine the monthly costs.
- Provision the solution on GCP. Leave the solution provisioned for 1 week. Use Stackdriver to determine the provisioned and used resource amounts. Multiply the 1 week cost to determine the monthly costs.

Unattempted

Provision the solution on GCP. Leave the solution provisioned for 1 week. Use Stackdriver to determine the provisioned and used resource amounts. Multiply the 1 week cost to determine the monthly costs. is not right.

By provisioning the solution on GCP, you are going to incur costs. Our requirement is to just estimate the costs and this can be done by using Google Cloud Pricing Calculator.

Ref: <https://cloud.google.com/products/calculator>

Provision the solution on GCP. Leave the solution provisioned for 1 week. Use Stackdriver to determine the provisioned and used resource amounts. Multiply the 1 week cost to determine the monthly costs. is not right.

By provisioning the solution on GCP, you are going to incur costs. Our requirement is to just estimate the costs and this can be done by using Google Cloud Pricing Calculator.

Ref: <https://cloud.google.com/products/calculator>

For each GCP product in the solution, review the pricing details on the products pricing page. Create a Google Sheet that summarizes the expected monthly costs for each product. is not right.

This would certainly work but is a manual task. Why use this when you can use Google Cloud Pricing Calculator to achieve the save?

Ref: <https://cloud.google.com/products/calculator>

For each GCP product in the solution, review the pricing details on the products pricing page. Use the pricing calculator to total the monthly costs for each GCP product. is the right answer.

You can use the Google Cloud Pricing Calculator to total the estimated monthly costs for each GCP product. You don't incur any charges for doing so.

Ref: <https://cloud.google.com/products/calculator>

24. Question

You have files in a Cloud Storage bucket that you need to share with your suppliers. You want to restrict the time that the files are available to your suppliers to 1 hour. You want to follow Google recommended practices. What should you do?

- Create a service account with just the permissions to access files in the bucket. Create a JSON key for the service account. Execute the command `gsutil signurl -p 60m gs://[Bucket]`.
- Create a service account with just the permissions to access files in the bucket. Create a JSON key for the service account. Execute the command `gsutil signurl -d 1h gs://[Bucket]/**`.

- Create a JSON key for the Default Compute Engine Service Account. Execute the command `gsutil signurl -t 60m gs:///*.*`.
- Create a service account with just the permissions to access files in the bucket. Create a JSON key for the service account. Execute the command `gsutil signurl -m 1h gs:///*.`

Unattempted

Create a JSON key for the Default Compute Engine Service Account. Execute the command `gsutil signurl -t 60m gs:///*.*` is not right.

`gsutil signurl` does not support `-t` flag. Executing the command with `-t` flag fails as shown.
`$ gsutil signurl -t 60m keys.json gs://gcp-ace-lab-255520/*.*`

`CommandException: Incorrect option(s) specified. Usage:`

Ref: <https://cloud.google.com/storage/docs/gsutil/commands/signurl>

Also, using the default compute engine service account violates the principle of least privilege. The recommended approach is to create a service account with just the right permissions needed and create JSON keys for this service account to use with `gsutil signurl` command.

Create a service account with just the permissions to access files in the bucket. Create a JSON key for the service account. Execute the command `gsutil signurl -p 60m gs://.` is not right.

With `gsutil signurl`, `-p` is used to specify the key store password instead of prompting for the password. It can not be used to pass a time value. Executing the command with `-p` flag fails as shown.

`$ gsutil signurl -p 60m keys.json gs://gcp-ace-lab-255520/*.*`

`TypeError: Last argument must be a byte string or a callable.`

Ref: <https://cloud.google.com/storage/docs/gsutil/commands/signurl>

Create a service account with just the permissions to access files in the bucket. Create a JSON key for the service account. Execute the command `gsutil signurl -m 1h gs:///*.` is not right.

With `gsutil signurl`, `-m` is used to specify the operation e.g. PUT/GET etc. Executing the command with `-m` flag fails as shown.

`$ gsutil signurl -m 1h keys.json gs://gcp-ace-lab-255520/*.*`

`CommandException: HTTP method must be one of[GET|HEAD|PUT|DELETE|RESUMABLE]`

Ref: <https://cloud.google.com/storage/docs/gsutil/commands/signurl>

Create a service account with just the permissions to access files in the bucket. Create a JSON key for the service account. Execute the command `gsutil signurl -d 1h gs:///**.` is the right answer.

This command correctly specifies the duration that the signed url should be valid for by using the `-d` flag. The default is 1 hour so omitting the `-d` flag would have also resulted in

the same outcome. Times may be specified with no suffix (default hours), or with s = seconds, m = minutes, h = hours, d = days. The max duration allowed is 7d.

Ref: <https://cloud.google.com/storage/docs/gsutil/commands/signurl>

25. Question

You have one GCP account running in your default region and zone and another account running in a non-default region and zone. You want to start a new Compute Engine instance in these two Google Cloud Platform accounts using the command line interface. What should you do?

- Activate two configurations using gcloud configurations activate [NAME]. Run gcloud config list to start the Compute Engine instances.
- Activate two configurations using gcloud configurations activate [NAME]. Run gcloud configurations list to start the Compute Engine instances.
- Create two configurations using gcloud config configurations create [NAME]. Run gcloud config configurations activate [NAME] to switch between accounts when running the commands to start the Compute Engine instances.
- Create two configurations using gcloud config configurations create [NAME]. Run gcloud configurations list to start the Compute Engine instances.

Unattempted

Create two configurations using gcloud config configurations create [NAME]. Run gcloud configurations list to start the Compute Engine instances. is not right.
gcloud configurations list is an invalid command. To list the existing named configurations, you need to execute gcloud config configurations list but this does not start the compute engine instances.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/list>

Activate two configurations using gcloud configurations activate [NAME]. Run gcloud configurations list to start the Compute Engine instances. is not right.
gcloud configurations list is an invalid command. To list the existing named configurations, you need to execute gcloud config configurations list but this does not start the compute engine instances.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/list>

Activate two configurations using gcloud configurations activate [NAME]. Run gcloud config list to start the Compute Engine instances. is not right.

`gcloud configurations activate [NAME]` activates an existing named configuration. It can't be used to activate two configurations at the same time. Moreover, `gcloud config list` lists Cloud SDK properties for the currently active configuration. It does not start the Compute Engine instances.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/activate>

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/list>

Create two configurations using `gcloud config configurations create [NAME]`. Run `gcloud config configurations activate [NAME]` to switch between accounts when running the commands to start the Compute Engine instances. is the right answer.

Each `gcloud` configuration has a 1 to 1 relationship with the region (if a region is defined). Since we have two different regions, we would need to create two separate configurations using `gcloud config configurations create`

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/create>

Secondly, you can activate each configuration independently by running `gcloud config configurations activate [NAME]`

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/activate>

Finally, while each configuration is active, you can run the `gcloud compute instances start [NAME]` command to start the instance in the configuration's region.

<https://cloud.google.com/sdk/gcloud/reference/compute/instances/start>

26. Question

You have one project called ptech-sa where you manage all your service accounts. You want to be able to use a service account from this project to take snapshots of VMs running in another project called ptech-vm. What should you do?

- When creating the VMs, set the service account's API scope for Compute Engine to read/write.
- Grant the service account the IAM Role of Compute Storage Admin in the project called ptech-vm.
- Download the private key from the service account, and add it to each VMs custom metadata.

- Download the private key from the service account, and add the private key to each VM's SSH keys.

Unattempted

Download the private key from the service account, and add it to each VMs custom metadata. is not right.

Adding service accounts private key (JSON file) to VMs custom metadata has no effect. Metadata entries are key-value pairs and do not influence any other behavior.

Ref: <https://cloud.google.com/compute/docs/storing-retrieving-metadata>

Download the private key from the service account, and add the private key to each VM's SSH keys. is not right.

Adding service accounts private key to the VMs SSH keys does not influence any other behavior. SSH keys are used for SSHing to the instance.

<https://cloud.google.com/compute/docs/instances/adding-removing-ssh-keys>

When creating the VMs, set the service account's API scope for Compute Engine to read/write. is not right.

The scopes can be modified only when using compute engine default service account.

Ref: https://cloud.google.com/compute/docs/access/service-accounts#default_service_account

See the screenshot below.

The scopes can not be modified when using a non-default service account. See the screenshot below.

Since we want to use service accounts from another project, it is safe to say they are not the default compute service accounts of this project and hence it is not possible to customize the scopes.

Grant the service account the IAM Role of Compute Storage Admin in the project called ptech-vm. is the right answer.

Compute Storage Admin role provides permissions to create, modify, and delete disks, images, and snapshots. If the service account in ptech-sa is granted the IAM Role of Compute Storage Admin in the project called ptech-vm, it can take snapshots and carry out other activities as defined by the role.

Ref: <https://cloud.google.com/compute/docs/access/iam#compute.storageAdmin>

27. Question

You have production and test workloads that you want to deploy on Compute Engine. Production VMs need to be in a different subnet than the test VMs. All the VMs must be

able to reach each other over internal IP without creating additional routes. You need to set up VPC and the 2 subnets. Which configuration meets these requirements?

- Create 2 custom VPCs, each with a single subnet. Create each subnet in a different region and with a different CIDR range.
- Create a single custom VPC with 2 subnets. Create each subnet in a different region and with a different CIDR range.
- Create 2 custom VPCs, each with a single subnet. Create each subnet in the same region and with the same CIDR range.
- Create a single custom VPC with 2 subnets. Create each subnet in the same region and with the same CIDR range.

Unattempted

Create 2 custom VPCs, each with a single subnet. Create each subnet in a different region and with a different CIDR range. is not right.

We need to get our requirements working with 1 VPC, not 2 !!

Create 2 custom VPCs, each with a single subnet. Create each subnet in the same region and with the same CIDR range. is not right.

We need to get our requirements working with 1 VPC, not 2 !!

Create a single custom VPC with 2 subnets. Create each subnet in the same region and with the same CIDR range. is not right.

We can not create two subnets in one VPC with the same CIDR range. “Primary and secondary ranges for subnets cannot overlap with any allocated range, any primary or secondary range of another subnet in the same network, or any IP ranges of subnets in peered networks.” Ref: <https://cloud.google.com/vpc/docs/using-vpc#subnet-rules>

Create a single custom VPC with 2 subnets. Create each subnet in a different region and with a different CIDR range. is the right answer.

When we create subnets in the same VPC with different CIDR ranges, they can communicate automatically within VPC. “Resources within a VPC network can communicate with one another by using internal (private) IPv4 addresses, subject to applicable network firewall rules”

Ref: <https://cloud.google.com/vpc/docs/vpc>

28. Question

You have sensitive data stored in three Cloud Storage buckets and have enabled data access logging. You want to verify activities for a particular user for these buckets, using the fewest possible steps. You need to verify the addition of metadata labels and which files have been viewed from those buckets. What should you do?

- View the bucket in the Storage section of the GCP Console.
- Using the GCP Console, filter the Activity log to view the information.
- **Using the GCP Console, filter the Stackdriver log to view the information.**
- Create a trace in Stackdriver to view the information.

Unattempted

Our requirements are – sensitive data, verify access, fewest possible steps.

Using the GCP Console, filter the Activity log to view the information. is not right.
Since data access logging is enabled, you can see relevant log entries in both activity Logs as well as stack driver logs. However, verifying what has been viewed/updated is not straightforward in activity logs. Activity logs display a list of all actions and you can restrict this down to a user and further filter by specifying Data access as the Activity types and GCS Bucket as the Resource type. But that is the extent of the filter functionality in Activity logs. It is not possible to restrict the activity logs to just the three buckets that we are interested in. Secondly, it is not possible to restrict the activity logs to just the gets and updates. So we'd have to go through the full list to identify activities of interest before verifying them which is a manual process and can be time taking depending on the number of activities in the list.

Ref: <https://cloud.google.com/storage/docs/audit-logs>

View the bucket in the Storage section of the GCP Console. is not right.
The bucket page in the GCP console does not show the logs.

Create a trace in Stackdriver to view the information. is not right.
Stackdriver trace is not supported on google cloud. Stackdriver Trace runs on Linux in the following environments: Compute Engine, Google Kubernetes Engine (GKE), App Engine flexible environment, App Engine standard environment.

Ref: <https://cloud.google.com/trace/docs/overview>

Using the GCP Console, filter the Stackdriver log to view the information. is the right answer.

Data access logs is already enabled, so we already record all API calls that read the configuration or metadata of resources, as well as user-driven API calls that create, modify, or read user-provided resource data. Data Access audit logs do not record the

data-access operations on resources that are publicly shared (available to All Users or All Authenticated Users) or that can be accessed without logging into Google Cloud.

Since we are dealing with sensitive data, it is safe to assume that these buckets are not publicly shared and therefore enabling Data access logging logs all data-access operations on resources. These logs are sent to Stackdriver where they can be viewed by applying a suitable filter.

Unlike activity logs, retrieving the required information to verify is easier and quicker through Stackdriver as you can apply filters such as

```
resource.type="gcs_bucket"  
(resource.labels.bucket_name="gcp-ace-lab-255520" OR  
resource.labels.bucket_name="gcp-ace-lab-255521" OR  
resource.labels.bucket_name="gcp-ace-lab-255522")  
(protoPayload.methodName="storage.objects.get" OR  
protoPayload.methodName="storage.objects.update")  
protoPayload.authenticationInfo.principalEmail="test.gcp.labs.user@gmail.com"
```

and query just the gets and updates, for specific buckets for a specific user. This involves fewer steps and is more efficient.

Data access logging is not enabled by default and needs to be enabled explicitly. The screenshot below shows a screenshot for enabling the data access logging for Google Cloud Storage.

29. Question

You have successfully created a development environment in a project for an application. This application uses Compute Engine and Cloud SQL. Now, you need to create a production environment for this application. The security team has forbidden the existence of network routes between these 2 environments, and asks you to follow Google-recommended practices. What should you do?

- Create a new project, enable the Compute Engine and Cloud SQL APIs in that project, and replicate the setup you have created in the development environment.

- Create a new production subnet in the existing VPC and a new production Cloud SQL instance in your existing project, and deploy your application using those resources.
- Create a new project, modify your existing VPC to be a Shared VPC, share that VPC with your new project, and replicate the setup you have in the development environment in that new project, in the Shared VPC.
- Ask the security team to grant you the Project Editor role in an existing production project used by another division of your company. Once they grant you that role, replicate the setup you have in the development environment in that project.

Unattempted

Create a new project, modify your existing VPC to be a Shared VPC, share that VPC with your new project, and replicate the setup you have in the development environment in that new project, in the Shared VPC. is not right.

A Shared VPC allows an organization to connect resources from multiple projects to a common Virtual Private Cloud (VPC) network so that they can communicate with each other securely and efficiently using internal IPs from that network. This goes totally against the recommendations of the security team.

Ref: <https://cloud.google.com/vpc/docs/shared-vpc>

Create a new production subnet in the existing VPC and a new production Cloud SQL instance in your existing project, and deploy your application using those resources. is not right.

You can't achieve complete isolation between development and production environments. When configuration access in Cloud SQL, while you can grant any application access to a Cloud SQL instance by authorizing the public IP addresses that the application uses to connect, you can not specify a private network (for example, 10.x.x.x) as an authorized network. The compute engine instances use their private IP addresses to reach out to Cloud SQL and because of the above limitation, we can't prevent the development compute engine reach out to production MySQL and vice versa. Since the security team has forbidden the existence of network routes between these 2 environments, having the production and development environments in a single project is not an option.

<https://cloud.google.com/sql/docs/mysql/connect-external-app#appaccessIP>

Ask the security team to grant you the Project Editor role in an existing production project used by another division of your company. Once they grant you that role, replicate the setup you have in the development environment in that project. is not right. While this would technically isolate the development environment from the production environment, your production application is running in a project that is also hosting production applications of another division of your company. This goes against Google's recommended practices. You can use folders to isolate requirements for different departments and teams in the parent organization. And you have separate projects under

the folders so as per Google recommendations we should be deploying the production application to a separate project that is just for one company division/department.
Ref: <https://cloud.google.com/docs/enterprise/best-practices-for-enterprise-organizations#define-hierarchy>

Create a new project, enable the Compute Engine and Cloud SQL APIs in that project, and replicate the setup you have created in the development environment. is the right answer.

This aligns with Google's recommended practices. By creating a new project, we achieve complete isolation between development and production environments; as well as isolate this production application from production applications of other departments.

Ref: <https://cloud.google.com/docs/enterprise/best-practices-for-enterprise-organizations#define-hierarchy>

30. Question

You have three gcloud configurations – one for each of development, test and production projects. You want to list all the configurations and switch to a new configuration. With the fewest steps possible, what's the fastest way to switch to the correct configuration?

- To list configurations - `gcloud config list` To activate a configuration - `gcloud config activate`.
- To list configurations - `gcloud configurations list` To activate a configuration - `gcloud configurations activate`
- To list configurations - `gcloud configurations list` To activate a configuration - `gcloud config activate`.
- To list configurations - `gcloud config configurations list` To activate a configuration - `gcloud config configurations activate`.

Unattempted

To list configurations – `gcloud configurations list`

To activate a configuration – `gcloud configurations activate`. is not right.

`gcloud configurations list` does not list configurations. To list existing configurations, you need to execute `gcloud config configurations list`.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/list>

`gcloud configurations activate` does not activate a named configuration. To activate a configuration, you need to execute `gcloud config configurations activate`

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/activate>

To list configurations – gcloud config list

To activate a configuration – gcloud config activate. is not right.

gcloud config list does not list configurations. It lists the properties of the existing configuration. To list existing configurations, you need to execute gcloud config configurations list.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/list>

gcloud config activate does not activate a named configuration. To activate a configuration, you need to execute gcloud config configurations activate

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/activate>

To list configurations – gcloud configurations list

To activate a configuration – gcloud config activate. is not right.

gcloud configurations list does not list configurations. To list existing configurations, you need to execute gcloud config configurations list.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/list>

gcloud config activate does not activate a named configuration. To activate a configuration, you need to execute gcloud config configurations activate

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/activate>

To list configurations – gcloud config configurations list

To activate a configuration – gcloud config configurations activate. is the right answer.
The two commands together achieve the intended outcome. gcloud config configurations list – lists existing named configurations and gcloud config configurations activate – activates an existing named configuration

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/list>

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/activate>

See an example below

```
$ gcloud config configurations list
```

NAME	IS_ACTIVE	ACCOUNT	PROJECT	DEFAULT_ZONE	DEFAULT_REGION
dev-configuration	False	gcp-ace-lab-dev			
prod-configuration	False	gcp-ace-lab-prod			
test-configuration	True	gcp-ace-lab-test			

```
$ gcloud config configurations activate prod-configuration
```

Activated [prod-configuration].

```
$ gcloud config configurations list
```

NAME	IS_ACTIVE	ACCOUNT	PROJECT	DEFAULT_ZONE	DEFAULT_REGION
dev-configuration	False	gcp-ace-lab-dev			
prod-configuration	True	gcp-ace-lab-prod			
test-configuration	False	gcp-ace-lab-test			

31. Question

You have two compute instances in the same VPC but in different regions. You can SSH from one instance to another instance using their external IP address but not their internal IP address. What could be the reason for SSH failing on the internal IP address?

- The internal IP address is disabled.
- The combination of compute instance network tags and VPC firewall rules allow SSH from 0.0.0.0 but denies SSH from the VPC subnets IP range.
- The compute instances are not using the right cross-region SSH IAM permissions
- The compute instances have a static IP for their internal IP.

Unattempted

The compute instances have a static IP for their internal IP. is not right.
Static internal IPs shouldn't be a reason for failed SSH connections. With all networking set up correctly, SSH works fine on Static internal IPs.

Ref: <https://cloud.google.com/compute/docs/ip-addresses#networkaddresses>

The internal IP address is disabled. is not right.

Every compute instance has one or more internal IP addresses so this option is not correct.

The compute instances are not using the right cross-region SSH IAM permissions. is not right.

There is no such thing as cross region SSH IAM permissions.

The combination of compute instance network tags and VPC firewall rules allow SSH from 0.0.0.0 but denies SSH from the VPC subnets IP range. is the right answer.

The combination of compute instance network tags and VPC firewall rules can certainly result in SSH traffic being allowed on the external IP range but disabled from subnets IP range. The firewall rule can be configured to allow SSH traffic from 0.0.0.0/0 but deny traffic from the VPC range e.g. 10.0.0.0/8. In this case, all SSH traffic from within the VPC is denied but external SSH traffic (i.e. on external IP) is allowed.

Ref: <https://cloud.google.com/vpc/docs/using-f火walls>

32. Question

You have two compute instances in the same VPC but in different regions. You can SSH from one instance to another instance using their internal IP address but not their external IP address. What could be the reason for SSH failing on external IP address?

- The combination of compute instance network tags and VPC firewall rules only allow SSH from the subnets IP range.
- The external IP address is disabled.
- The compute instances have a static IP for their external IP.
- The compute instances are not using the right cross-region IAM permissions

Unattempted

The compute instances have a static IP for their external IP. is not right.
Not having a static IP is not a reason for failed SSH connections. When the firewall rules are set up correctly, SSH works fine on compute instances having ephemeral IP Address.

The external IP address is disabled. is not right.

Our question states SSH doesn't work on external IP addresses so it is safe to assume they already have an external IP. Therefore, this option is not correct.

The compute instances are not using the right cross-region IAM permissions. is not right.

There is no such thing as cross region SSH IAM permissions.

The combination of compute instance network tags and VPC firewall rules only allow SSH from the subnets IP range. is the right answer.

The combination of compute instance network tags and VPC firewall rules can certainly result in SSH traffic being allowed from only subnets IP range. The firewall rule can be configured to allow SSH traffic from just the VPC range e.g. 10.0.0.0/8. In this scenario, all SSH traffic from within the VPC is accepted but external SSH traffic is blocked.

Ref: <https://cloud.google.com/vpc/docs/using-firewalls>

33. Question

You have two Kubernetes resource configuration files.

deployments.yaml – creates a deployment

service.yaml – sets up a LoadBalancer service to expose the pods.

You don't have a GKE cluster in the development project and you need to provision one. Which of the commands fail with an error in Cloud Shell when you are attempting to create a GKE cluster and deploy the YAML configuration files to create a deployment and service. (Select Two)

- `gcloud container clusters create cluster-1 --zone=us-central1-a gcloud container clusters get-credentials cluster-1 --zone=us-central1-a kubectl apply -f [deployment.yaml,service.yaml]`
- `gcloud container clusters create cluster-1 --zone=us-central1-a gcloud container clusters get-credentials cluster-1 --zone=us-central1-a kubectl apply -f deployment.yaml&&service.yaml`
- `gcloud config set compute/zone us-central1-a gcloud container clusters create cluster-1 gcloud container clusters get-credentials cluster-1 --zone=us-central1-a kubectl apply -f deployment.yaml kubectl apply -f service.yaml`
- `gcloud container clusters create cluster-1 --zone=us-central1-a gcloud container clusters get-credentials cluster-1 --zone=us-central1-a kubectl apply -f deployment.yaml kubectl apply -f service.yaml`
- `gcloud container clusters create cluster-1 --zone=us-central1-a gcloud container clusters get-credentials cluster-1 --zone=us-central1-a kubectl apply -f deployment.yaml,service.yaml`

Unattempted

`gcloud container clusters create cluster-1 --zone=us-central1-a
gcloud container clusters get-credentials cluster-1 --zone=us-central1-a
kubectl apply -f deployment.yaml
kubectl apply -f service.yaml.` is not right (i.e. commands executes successfully)
You create a cluster by running `gcloud container clusters create` command. You then fetch credentials for a running cluster by running `gcloud container clusters get-credentials` command. Finally, you apply the kubernetes resource configuration by running `kubectl apply -f`

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/create>

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/get-credentials>

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#apply>

`gcloud container clusters create cluster-1 --zone=us-central1-a
gcloud container clusters get-credentials cluster-1 --zone=us-central1-a
kubectl apply -f deployment.yaml,service.yaml.` is not right (i.e. commands executes successfully)

Like above, but the only difference is that both configurations are applied in the same statement. With `kubectl apply`, you can apply the configuration from a single file or

multiple files or even a complete directory.

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#apply>

gcloud config set compute/zone us-central1-a

gcloud container clusters create cluster-1

gcloud container clusters get-credentials cluster-1 --zone=us-central1-a

kubectl apply -f deployment.yaml

kubectl apply -f service.yaml. is not right (i.e. command executes successfully)

Like above, but the only difference is in how the compute zone is set. In this scenario, you set the zone us-central1-a as the default zone so when you don't pass a zone property to the gcloud container clusters create command, it takes the default zone which is us-central1-a.

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#apply>

gcloud container clusters create cluster-1 --zone=us-central1-a

gcloud container clusters get-credentials cluster-1 --zone=us-central1-a

kubectl apply -f [deployment.yaml,service.yaml]. is the right answer (i.e. commands fail)

kubectl apply can apply the configuration from a single file or multiple files or even a complete directory. When applying configuration from multiple files, the file names need to be separated by a comma. In this scenario, the filenames are passed as a list and Kubernetes treats the list as literal so looks for

files “[deployment.yaml]” and “[service.yaml]” which it doesn't find.

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#apply>

gcloud container clusters create cluster-1 --zone=us-central1-a

gcloud container clusters get-credentials cluster-1 --zone=us-central1-a

kubectl apply -f deployment.yaml&&service.yaml. is the right answer (i.e. commands fail)

kubectl apply can apply the configuration from a single file or multiple files or even a complete directory. When applying configuration from multiple files, the file names need to be separated by a comma. In this scenario, the filenames are separated by && and Kubernetes treats the && as literal so it looks for the

file “deployment.yaml&&service.yaml” which it doesn't find.

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#apply>

34. Question

You have two Kubernetes resource configuration files.

deployments.yaml – creates a deployment

service.yaml – sets up a LoadBalancer service to expose the pods.

You don't have a GKE cluster in the development project and you need to provision one. Which of the commands below would you run in Cloud Shell to create a GKE cluster and deploy the YAML configuration files to create a deployment and service?

- `gcloud container clusters create cluster-1 --zone=us-central1-a`
`gcloud container clusters get-credentials cluster-1 --zone=us-central1-a`
`kubectl create -f deployment.yaml`
`kubectl create -f service.yaml`
- `gcloud container clusters create cluster-1 --zone=us-central1-a`
`gcloud container clusters get-credentials cluster-1 --zone=us-central1-a`
`kubectl apply -f deployment.yaml`
`kubectl apply -f service.yaml`
- `gcloud container clusters create cluster-1 --zone=us-central1-a`
`gcloud container clusters get-credentials cluster-1 --zone=us-central1-a`
`gcloud gke apply -f deployment.yaml`
`gcloud gke apply -f service.yaml`
- `kubectl container clusters create cluster-1 --zone=us-central1-a`
`kubectl container clusters get-credentials cluster-1 --zone=us-central1-a`
`kubectl apply -f deployment.yaml`
`kubectl apply -f service.yaml.`

Unattempted

`kubectl container clusters create cluster-1 --zone=us-central1-a`
`kubectl container clusters get-credentials cluster-1 --zone=us-central1-a`
`kubectl apply -f deployment.yaml`
`kubectl apply -f service.yaml.`

is not right.

`kubectl` doesn't support `kubectl container clusters create` command. `kubectl` can not be used to create GKE clusters. To create a GKE cluster, you need to execute `gcloud container clusters create` command.

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/create>

`gcloud container clusters create cluster-1 --zone=us-central1-a`
`gcloud container clusters get-credentials cluster-1 --zone=us-central1-a`
`kubectl create -f deployment.yaml`
`kubectl create -f service.yaml.`

is not right.

`kubectl` doesn't support `kubectl create` command. The YAML file contains the cluster resource configuration. You don't create the configuration, instead, you apply the configuration to the cluster. The configuration can be applied by running `kubectl apply` command

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#apply>

```
gcloud container clusters create cluster-1 --zone=us-central1-a  
gcloud container clusters get-credentials cluster-1 --zone=us-central1-a  
gcloud gke apply -f deployment.yaml  
gcloud gke apply -f service.yaml.
```

is not right.

gcloud doesn't support gcloud gke apply command. The YAML file contains the cluster resource configuration. You don't create the configuration, instead, you apply the configuration to the cluster. The configuration can be applied by running kubectl apply command

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#apply>

```
gcloud container clusters create cluster-1 --zone=us-central1-a  
gcloud container clusters get-credentials cluster-1 --zone=us-central1-a  
kubectl apply -f deployment.yaml  
kubectl apply -f service.yaml.
```

is the right answer.

You create a cluster by running gcloud container clusters create command. You then fetch credentials for a running cluster by running gcloud container clusters get-credentials command. Finally, you apply the Kubernetes resource configuration by running kubectl apply -f

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/create>

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/get-credentials>

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#apply>

35. Question

You have two workloads on GKE (Google Kubernetes Engine) – create-order and dispatch-order. create-order handles the creation of customer orders, and dispatch-order handles dispatching orders to your shipping partner. Both create-order and dispatch-order workloads have cluster autoscaling enabled. The create-order deployment needs to access (i.e. invoke web service of) dispatch-order deployment. dispatch-order deployment cannot be exposed publicly. How should you define the services?

- Create a Service of type NodePort for dispatch-order and an Ingress Resource for that Service. Have create-order use the Ingress IP address.

- Create a Service of type LoadBalancer for dispatch-order and an Ingress Resource for that Service. Have create-order use the Ingress IP address.
- Create a Service of type LoadBalancer for dispatch-order. Have create-order use the Service IP address.
- Create a Service of type ClusterIP for dispatch-order. Have create-order use the Service IP address.

Unattempted

Create a Service of type LoadBalancer for dispatch-order. Have create-order use the Service IP address. is not right.

When you create a Service of type LoadBalancer, the Google Cloud controller configures a network load balancer that is publicly available. Since we don't want our service to be publicly available, we shouldn't create a Service of type LoadBalancer

Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/exposing-apps>

Create a Service of type LoadBalancer for dispatch-order and an Ingress Resource for that Service. Have create-order use the Ingress IP address. is not right.

When you create a Service of type LoadBalancer, the Google Cloud controller configures a network load balancer that is publicly available. Since we don't want our service to be publicly available, we shouldn't create a Service of type LoadBalancer

Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/exposing-apps>

Create a Service of type NodePort for dispatch-order and an Ingress Resource for that Service. Have create-order use the Ingress IP address. is not right.

Exposees the Service on each Node's IP at a static port (the NodePort). If the compute instance has public connectivity, the dispatch-order can be accessed publicly which is undesirable. Secondly, dispatch-order has auto-scaling enabled so we shouldn't create a service of NodePort. If autoscaler spins up another pod on the node, it fails to initialize as the port on the node is already taken by an existing pod on the same node.

Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/exposing-apps>

Create a Service of type ClusterIP for dispatch-order. Have create-order use the Service IP address. is the right answer.

ClusterIP exposes the Service on a cluster-internal IP that is only reachable within the cluster. This satisfies our requirement that dispatch-order shouldn't be publicly accessible. create-order which is also located in the same GKE cluster can now access the ClusterIP of the service to reach dispatch-order.

Ref: <https://kubernetes.io/docs/concepts/services-networking/service/>

36. Question

You host a production application in Google Compute Engine in the us-central1-a zone. Your application needs to be available 24*7 all through the year. The application suffered an outage recently due to a Compute Engine outage in the zone hosting your application. Your application is also susceptible to slowness during peak usage. You have been asked for a recommendation on how to modify the infrastructure to implement a cost-effective and scalable solution that can withstand zone failures. What would you recommend?

- Use Managed instance groups with instances in a single zone. Enable Autoscaling on the Managed instance group.
- Use Managed instance groups with preemptible instances across multiple zones. Enable Autoscaling on the Managed instance group.
- **Use Managed instance groups across multiple zones. Enable Autoscaling on the Managed instance group.**
- Use Unmanaged instance groups across multiple zones. Enable Autoscaling on the Unmanaged instance group.

Unattempted

Use Managed instance groups with preemptible instances across multiple zones. Enable Autoscaling on the Managed instance group. is not right.
A preemptible VM runs at a much lower price than normal instances and is cost-effective. However, Compute Engine might terminate (preempt) these instances if it requires access to those resources for other tasks. Preemptible instances are not suitable for production applications that need to be available 24*7.

Ref: <https://cloud.google.com/compute/docs/instances/preemptible>

Use Unmanaged instance groups across multiple zones. Enable Autoscaling on the Unmanaged instance group. is not right.

Unmanaged instance groups do not autoscale. An unmanaged instance group is simply a collection of virtual machines (VMs) that reside in a single zone, VPC network, and subnet. An unmanaged instance group is useful for grouping together VMs that require individual configuration settings or tuning.

Ref: <https://cloud.google.com/compute/docs/instance-groups/creating-groups-of-unmanaged-instances>

Use Managed instance groups with instances in a single zone. Enable Autoscaling on the Managed instance group. is not right.

While enabling auto-scaling is a good idea, autoscaling would spin up instances in the same zone. Should there be a zone failure, all instances of the managed instance group would be unreachable and cause the application to be unreachable. Google recommends

you distribute your resources across multiple zones to tolerate outages.

Ref: <https://cloud.google.com/compute/docs/regions-zones>

Use Managed instance groups across multiple zones. Enable Autoscaling on the Managed instance group. is the right answer.

Distribute your resources across multiple zones and regions to tolerate outages. Google designs zones to be independent of each other: a zone usually has power, cooling, networking, and control planes that are isolated from other zones, and most single failure events will affect only a single zone. Thus, if a zone becomes unavailable, you can transfer traffic to another zone in the same region to keep your services running.

Ref: <https://cloud.google.com/compute/docs/regions-zones>

In addition, a managed instance group (MIG) contains offers auto-scaling capabilities that let you automatically add or delete instances from a managed instance group based on increases or decreases in load. Autoscaling helps your apps gracefully handle increases in traffic and reduce costs when the need for resources is lower. Autoscaling works by adding more instances to your instance group when there is more load (upscale), and deleting instances when the need for instances is lowered (downscale).

Ref: <https://cloud.google.com/compute/docs/autoscaler/>

37. Question

You host a static website in Cloud Storage. Recently you began to include links to PDF files on this site. Currently, when users click on links to these PDF files, their browser prompts them to save the file to their machine locally. However, you want the clicked PDF files to be displayed within the browser window directly without prompting the user to save the files locally. What should you do?

- Set Content-Type metadata to application/pdf on the PDF file objects
- Enable Cloud CDN on the website frontend.
- Add a label to the storage bucket with a key of Content-Type and a value of application/pdf.
- Enable Share publicly on the PDF file objects

Unattempted

Set Content-Type metadata to application/pdf on the PDF file objects is the right answer. Content-Type allows browsers to render the object properly. If the browser prompts users to save files to their machine, it is likely the browser does not see the Content-Type as application/pdf. Setting this would ensure the browser displays PDF files within the browser instead of popping up a download dialog.

Ref: https://cloud.google.com/storage/docs/gsutil/addlhelp/WorkingWithObjectMetadata#content-type_1

Enable Cloud CDN on the website frontend. is not right.

CDN helps with caching content at the edge but doesn't help the browser in displaying pdf files.

Enable Share publicly on the PDF file objects. is not right.

The fact that the browser lets users download the file suggests the browser is able to reach out and download the file. Sharing publicly wouldn't make any difference.

Add a label to the storage bucket with a key of Content-Type and a value of application/pdf. is not right.

Bucket labels are key: value metadata pairs that allow you to group your buckets along with other Google Cloud resources such as virtual machine instances and persistent disks. They don't determine the file's content type.

38. Question

You installed Stackdriver Logging agent on all compute instances. You now need to forward logs from all Compute Engine instances to a BigQuery dataset called pt-logs. You want to minimize cost. What should you do?

- 1. Give the BigQuery Data Editor role on the pt-logs dataset to the service accounts used by your instances. 2. Update your instances' metadata to add the following value: logs-destination: bq://pt-logs.
- 1. In Stackdriver Logging, create a logs export with a Cloud Pub/Sub topic called logs as a sink. 2. Create a Cloud Function that is triggered by messages in the logs topic. 3. Configure that Cloud Function to drop logs that are not from Compute Engine and to insert Compute Engine logs in the pt-logs dataset.
- 1. In Stackdriver Logging, create a filter to view only Compute Engine logs. 2. Click Create Export. 3. Choose BigQuery as Sink Service, and the pt-logs dataset as Sink Destination.
- 1. Create a Cloud Function that has the BigQuery User role on the pt-logs dataset. 2. Configure this Cloud Function to create a BigQuery Job that executes this query: `INSERT INTO dataset.pt-logs (timestamp, log) SELECT timestamp, log FROM`

`compute.logs WHERE timestamp > DATE_SUB(CURRENT_DATE(), INTERVAL 1 DAY)` 3.

Use Cloud Scheduler to trigger this Cloud Function once a day.

Unattempted

1. Give the BigQuery Data Editor role on the pt-logs dataset to the service accounts used by your instances.
2. Update your instances' metadata to add the following value: logs-destination: bq://pt-logs. is not right.

Among other things, roles/bigquery.dataEditor lets you Create, update, get, and delete the dataset's tables. However, setting a metadata tag logs-destination to bq://pt-logs has no effect on how the logs are generated or forwarded. The stack driver agent is already installed so the logs are forwarded to stack driver logging and not to the BigQuery dataset. Metadata entries are key-value pairs and do not influence this behavior.

Ref: <https://cloud.google.com/compute/docs/storing-retrieving-metadata>

1. In Stackdriver Logging, create a logs export with a Cloud Pub/Sub topic called logs as a sink.
2. Create a Cloud Function that is triggered by messages in the logs topic.
3. Configure that Cloud Function to drop logs that are not from Compute Engine and to insert Compute Engine logs in the pt-logs dataset. is not right.

While the end result meets our requirement, this option involves more steps, it is inefficient and expensive. Triggering a cloud function for each log message and then dropping messages that are not relevant (i.e. not compute engine logs) is inefficient. We are paying for cloud function execution for all log entries when we are only interested in compute engine logs. Secondly, triggering a cloud function and then have that insert into the BigQuery dataset is also inefficient and expensive when the same can be achieved directly by configuring BigQuery as the sink destination – we don't pay for cloud function executions. Using this option, we are unnecessarily paying for Cloud Pub/Sub and Cloud Functions.

Ref: https://cloud.google.com/logging/docs/export/configure_export_v2

Ref: <https://cloud.google.com/logging/docs/view/advanced-queries>

1. Create a Cloud Function that has the BigQuery User role on the pt-logs dataset.

2. Configure this Cloud Function to create a BigQuery Job that executes this query:

`INSERT INTO dataset.pt-logs (timestamp, log)`

`SELECT timestamp, log FROM compute.logs`

`WHERE timestamp > DATE_SUB(CURRENT_DATE(), INTERVAL 1 DAY)`

3. Use Cloud Scheduler to trigger this Cloud Function once a day. is not right.

The role roles/bigquery.user provides permissions to run jobs, including queries, within the project. A cloud function with this role can execute queries in BigQuery, however, the logs are not available in BigQuery in compute.logs so you can not query compute engine logs by running `SELECT timestamp, log FROM compute.logs`.

Ref: <https://cloud.google.com/bigquery/docs/access-control>

1. In Stack driver Logging, create a filter to view only Compute Engine logs.

2. Click Create Export.

3. Choose BigQuery as Sink Service, and the pt-logs dataset as Sink Destination. is the right answer.

In stack driver logging, it is possible to create a filter to just query the compute engine logs which is what we are interested in.

Ref: <https://cloud.google.com/logging/docs/view/advanced-queries>

You can then export these logs into a sink that has the BigQuery dataset configured as the destination.

https://cloud.google.com/logging/docs/export/configure_export_v2

This way, just the logs that we need are exported to BigQuery. This option is the most efficient of all options and uses features provided by GCP out of the box.

39. Question

You need a dynamic way of provisioning VMs on the Compute Engine. The exact specifications will be in a dedicated configuration file. You want to follow Google recommended practices. Which method should you use?

- Unmanaged Instance Group
- Deployment Manager
- Managed Instance Group
- Cloud Composer

Unattempted

Unmanaged Instance Group. is not right.

Unmanaged instance groups let you load balance across a fleet of VMs that you manage yourself. But it doesn't help with dynamically provisioning VMs.

Ref: https://cloud.google.com/compute/docs/instance-groups#unmanaged_instance_groups

Cloud Composer. is not right.

Cloud Composer is a fully managed workflow orchestration service that empowers you to author, schedule, and monitor pipelines that span across clouds and on-premises data centers. Cloud Composer is deeply integrated within the Google Cloud Platform, giving users the ability to orchestrate their full pipeline. Cloud Composer has robust, built-in integration with many products, including BigQuery, Cloud Dataflow, Cloud Dataproc, Cloud Datastore, Cloud Storage, Cloud Pub/Sub, and AI Platform.

Ref: <https://cloud.google.com/composer>

Managed Instance Group. is not right.

Managed instance groups (MIGs) let you operate apps on multiple identical VMs. You can make your workloads scalable and highly available by taking advantage of automated MIG services, including autoscaling, autohealing, regional (multiple zones) deployment, and automatic updating. While MIG dynamically provisions virtual machines based on scaling policy, it doesn't satisfy our requirement of "dedicated configuration file"

Ref: https://cloud.google.com/compute/docs/instance-groups#managed_instance_groups

Deployment Manager. is the right answer.

Google Cloud Deployment Manager allows you to specify all the resources needed for your application in a declarative format using YAML. You can also use Python or Jinja2 templates to parameterize the configuration and allow reuse of common deployment paradigms such as a load-balanced, auto-scaled instance group. You can deploy many resources at one time, in parallel. Using the deployment manager, you can apply a Python/Jinja2 template to create a MIG/auto-scaling policy that dynamically provisions VM. And our other requirement of "dedicated configuration file" is also met. Using the deployment manager for provisioning results in a repeatable deployment process. By creating configuration files that define the resources, the process of creating those resources can be repeated over and over with consistent results. Google recommends we script our infrastructure and deploy using Deployment Manager

Ref: <https://cloud.google.com/deployment-manager>

40. Question

You need to assign a Cloud Identity and Access Management (Cloud IAM) role to an external auditor. The auditor needs to have permissions to review your Google Cloud Platform (GCP) Audit Logs and also to review your Data Access logs. What should you do?

- Assign the auditor the IAM role roles/logging.privateLogViewer. Perform the export of logs to Cloud Storage.
- Assign the auditor the IAM role roles/logging.privateLogViewer. Direct the auditor to also review the logs for changes to Cloud IAM policy.
- Assign the auditor's IAM user to a custom role that has logging.privateLogEntries.list permission. Perform the export of logs to Cloud Storage.
- Assign the auditor's IAM user to a custom role that has logging.privateLogEntries.list permission. Direct the auditor to also review the logs for changes to Cloud IAM policy.

Unattempted

Google Cloud provides Cloud Audit Logs, which is an integral part of Cloud Logging. It consists of two log streams for each project: Admin Activity and Data Access, which are generated by Google Cloud services to help you answer the question of “who did what, where, and when?” within your Google Cloud projects.

Ref: https://cloud.google.com/iam/docs/job-functions/auditing#scenario_external_auditors

To view Admin Activity audit logs, you must have one of the following Cloud IAM roles in the project that contains your audit logs:

- Project Owner, Project Editor, or Project Viewer.
- The Logging Logs Viewer role.
- A custom Cloud IAM role with the `logging.logEntries.list` Cloud IAM permission.

https://cloud.google.com/iam/docs/audit-logging#audit_log_permissions

To view Data Access audit logs, you must have one of the following roles in the project that contains your audit logs:

- Project Owner.
- Logging's Private Logs Viewer role.
- A custom Cloud IAM role with the `logging.privateLogEntries.list` Cloud IAM permission.

https://cloud.google.com/iam/docs/audit-logging#audit_log_permissions

Assign the auditor's IAM user to a custom role that has `logging.privateLogEntries.list` permission. Perform the export of logs to Cloud Storage. is not right.

`logging.privateLogEntries.list` provides permissions to view Data Access audit logs but this does not grant permissions to view Admin activity logs.

Ref: https://cloud.google.com/logging/docs/access-control#console_permissions

Assign the auditor's IAM user to a custom role that has `logging.privateLogEntries.list` permission. Direct the auditor to also review the logs for changes to Cloud IAM policy. is not right.

`logging.privateLogEntries.list` provides permissions to view Data Access audit logs but this does not grant permissions to view Admin activity logs.

Ref: https://cloud.google.com/logging/docs/access-control#console_permissions

Assign the auditor the IAM role `roles/logging.privateLogViewer`. Perform the export of logs to Cloud Storage. is not right.

`roles/logging.privateLogViewer` is the right role and lets the auditor review admin activity and data access logs but exporting logs to Cloud Storage indicates that we want the auditor to review logs from Cloud Storage and not the logs within Cloud Logging console.

In this scenario, unless the auditor is assigned a role that lets them access the relevant cloud storage buckets, they wouldn't be able to view log information in the buckets.

Assign the auditor the IAM role roles/logging.privateLogViewer. Direct the auditor to also review the logs for changes to Cloud IAM policy. is the right answer.

roles/logging.privateLogViewer (Private Logs Viewer) includes everything from roles/logging.viewer, plus the ability to read Access Transparency logs and Data Access audit logs. This lets the auditor review the admin activity and data access logs in Cloud Logging console.

Ref: <https://cloud.google.com/logging/docs/access-control>

41. Question

You need to configure IAM access audit logging in BigQuery for external auditors. You want to follow Google-recommended practices. What should you do?

- Add the auditors group to two new custom IAM roles.
- Add the auditor user accounts to the 'logging.viewer' and 'bigQuery.dataViewer' predefined IAM roles.
- Add the auditor user accounts to two new custom IAM roles.
- Add the auditors group to the 'logging.viewer' and 'bigQuery.dataViewer' predefined IAM roles.

Unattempted

Add the auditor user accounts to the 'logging.viewer' and 'bigQuery.dataViewer' predefined IAM roles. is not right.

Since auditing happens several times a year, we don't want to repeat the process of granting multiple roles to multiple users every time. Instead, we want to define a group with the required grants (a one time task) and assign this group to the auditor users during the time of the audit.

Add the auditor user accounts to two new custom IAM roles. is not right.

Google already provides roles that fit the external auditing requirements so we don't need to create custom roles. Nothing stops us from creating custom IAM roles to achieve the same purpose, but this doesn't follow "Google-recommended practices"

Add the auditors group to two new custom IAM roles. is not right.

Google already provides roles that fit the external auditing requirements so we don't need to create custom roles. Nothing stops us from creating custom IAM roles to achieve the same purpose, but this doesn't follow "Google-recommended practices"

Add the auditors group to the ‘logging.viewer’ and ‘bigQuery.dataViewer’ predefined IAM roles. is the right answer.

For external auditors, Google recommends we grant logging.viewer and bigquery.dataViewer roles. Since auditing happens several times a year to review the organization’s audit logs, it is recommended we create a group with these grants and assign the group to auditor user accounts during the time of the audit.

42. Question

You need to create a custom IAM role for use with a GCP service. All permissions in the role must be suitable for production use. You also want to clearly share with your organization the status of the custom role. This will be the first version of the custom role. What should you do?

- Use permissions in your role that use the SUPPORTED support level for role permissions. Set the role stage to ALPHA while testing the role permissions.
- Use permissions in your role that use the SUPPORTED support level for role permissions. Set the role stage to BETA while testing the role permissions.
- Use permissions in your role that use the TESTING support level for role permissions. Set the role stage to ALPHA while testing the role permissions.
- Use permissions in your role that use the TESTING support level for role permissions. Set the role stage to BETA while testing the role permissions.

Unattempted

When setting support levels for permissions in custom roles, you can set to one of SUPPORTED, TESTING or NOT_SUPPORTED.

SUPPORTED -The permission is fully supported in custom roles.

TESTING – The permission is being tested to check its compatibility with custom roles.

You can include the permission in custom roles, but you might see unexpected behavior.

Not recommended for production use.

Ref: <https://cloud.google.com/iam/docs/custom-roles-permissions-support>

Since we want the role to be suitable for production use, we need “SUPPORTED” and not “TESTING”.

In terms of role stage, the stage transitions from ALPHA → BETA → GA

Ref: https://cloud.google.com/iam/docs/understanding-custom-roles#testing_and_developing

Since this is the first version of custom role, we start with “ALPHA”.

The only option that satisfies “ALPHA” stage with “SUPPORTED” support level is Use permissions in your role that use the SUPPORTED support level for role permissions. Set the role stage to ALPHA while testing the role permissions.

43. Question

You need to create a custom VPC with a single subnet. The subnet's range must be as large as possible. Which range should you use?

- 10.0.0.0/8
- 192.168.0.0/16
- 172.16.0.0/12
- 0.0.0.0/0

Unattempted

The private network range is defined by IETF (Ref: <https://tools.ietf.org/html/rfc1918>) and adhered to by all cloud providers. The supported internal IP Address ranges are

1. 24-bit block 10.0.0.0/8 (16777216 IP Addresses)
2. 20-bit block 172.16.0.0/12 (1048576 IP Addresses)
3. 16-bit block 192.168.0.0/16 (65536 IP Addresses)

10.0.0.0/8 gives you the largest range – 16777216 IP Addresses.

44. Question

You need to create a new billing account and then link it with an existing Google Cloud Platform project. What should you do?

- Verify that you are Project Billing Manager for the GCP project. Update the existing project to link it to the existing billing account.
- Verify that you are Billing Administrator for the billing account. Update the existing project to link it to the existing billing account.

- Verify that you are Project Billing Manager for the GCP project. Create a new billing account and link the new billing account to the existing project.
- Verify that you are Billing Administrator for the billing account. Create a new project and link the new project to the existing billing account.

Unattempted

Our requirement is to link an existing google cloud project with a new billing account.

Verify that you are Billing Administrator for the billing account. Create a new project and link the new project to the existing billing account. is not right.

We do not need to create a new project.

Verify that you are Project Billing Manager for the GCP project. Update the existing project to link it to the existing billing account. is not right.

We want to link the project with a new billing account so is not right.

Verify that you are Billing Administrator for the billing account. Update the existing project to link it to the existing billing account. is not right.

We want to link the project with a new billing account so is not right.

Verify that you are Project Billing Manager for the GCP project. Create a new billing account and link the new billing account to the existing project. is the right answer.
The purpose of Project Billing Manager is to Link/unlink the project to/from a billing account. It is granted at the organization or project level. Project Billing Manager role allows a user to attach the project to the billing account, but does not grant any rights over resources. Project Owners can use this role to allow someone else to manage the billing for the project without granting them resource access.

Ref: <https://cloud.google.com/billing/docs/how-to/billing-access>

45. Question

You need to create an autoscaling managed instance group for an HTTPS web application. You want to make sure that unhealthy VMs are recreated. What should you do?

- In the Instance Template, add the label health-check.
- In the Instance Template, add a startup script that sends a heartbeat to the metadata server.

- Select Multi-Zone instead of Single-Zone when creating the Managed Instance Group.
- Create a health check on port 443 and use that when creating the Managed Instance Group.

Unattempted

Our requirement is to ensure unhealthy VMs are recreated.

Select Multi-Zone instead of Single-Zone when creating the Managed Instance Group. is not right.

You can create two types of MIGs: A zonal MIG, which deploys instances to a single zone and a regional MIG, which deploys instances to multiple zones across the same region. However, this doesn't help with recreating unhealthy VMs.

Ref: <https://cloud.google.com/compute/docs/instance-groups>

In the Instance Template, add the label health-check. is not right.

Metadata entries are key-value pairs and do not influence any other behavior.

Ref: <https://cloud.google.com/compute/docs/storing-retrieving-metadata>

In the Instance Template, add a startup script that sends a heartbeat to the metadata server. is not right.

The startup script is executed only at the time of startup. Whereas we need something like a liveness check that monitors the status of the server periodically to identify if the VM is unhealthy. So this is not going to work.

Ref: <https://cloud.google.com/compute/docs/startupscript>

Create a health check on port 443 and use that when creating the Managed Instance Group. is the right answer.

To improve the availability of your application and to verify that your application is responding, you can configure an auto-healing policy for your managed instance group (MIG). An auto-healing policy relies on an application-based health check to verify that an application is responding as expected. If the auto healer determines that an application isn't responding, the managed instance group automatically recreates that instance. Since our application is a HTTPS web application, we need to set up our health check on port 443 which is the standard port for HTTPS.

Ref: <https://cloud.google.com/compute/docs/instance-groups/autohealing-instances-in-migs>

46. Question

You need to deploy an application, which is packaged in a container image, in a new project. The application exposes an HTTP endpoint and receives very few requests per day. You want to minimize costs. What should you do?

- Deploy the container on Cloud Run.
- Deploy the container on Cloud Run on GKE.
- Deploy the container on App Engine Flexible.
- Deploy the container on Google Kubernetes Engine, with cluster autoscaling and horizontal pod autoscaling enabled.

Unattempted

Deploy the container on Cloud Run on GKE. is not right.

Cloud Run on GKE can scale the number of pods to zero. The number of nodes per cluster cannot scale to zero and these nodes are billed in the absence of requests.

Ref: <https://cloud.google.com/serverless-options>

Deploy the container on Google Kubernetes Engine, with cluster autoscaling and horizontal pod autoscaling enabled. is not right.

Like above, while you can set up the pod autoscaler to scale back the pods to zero, the number of nodes per cluster cannot scale to zero and these nodes are billed in the absence of requests. If you specify the minimum node pool size of zero nodes, an idle node pool can scale down completely. However, at least one node must always be available in the cluster to run system Pods.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/cluster-autoscaler>

Deploy the container on App Engine Flexible. is not right.

App Engine flexible environment instances are Compute Engine virtual machines. This means you can't truly scale down to zero and compute instances are billed in the absence of requests.

Ref: <https://cloud.google.com/appengine/docs/flexible>

Deploy the container on Cloud Run. is the right answer.

Cloud Run is a fully managed compute platform that automatically scales your stateless containers. Cloud Run is serverless. Cloud Run abstracts away all infrastructure management. It automatically scales up and down from zero depending on traffic almost instantaneously. Cloud Run only charges you for the exact resources you use.

Ref: <https://cloud.google.com/run>

47. Question

You need to enable traffic between multiple groups of Compute Engine instances that are currently running two different GCP projects. Each group of Compute Engine instances is running in its own VPC. What should you do?

- Verify that both projects are in a GCP Organization. Create a new VPC and add all instances.
- Verify that both projects are in a GCP Organization. Share the VPC from one project and request that the Compute Engine instances in the other project use this shared VPC.
- Verify that you are the Project Administrator of both projects. Create two new VPCs and add all instances.
- Verify that you are the Project Administrator of both projects. Create a new VPC and add all instances.

Unattempted

Verify that both projects are in a GCP Organization. Share the VPC from one project and request that the Compute Engine instances in the other project use this shared VPC. is the right answer.

All other options make no sense. Shared VPC allows an organization to connect resources from multiple projects to a common Virtual Private Cloud (VPC) network so that they can communicate with each other securely and efficiently using internal IPs from that network. When you use Shared VPC, you designate a project as a host project and attach one or more other service projects to it.

Ref: <https://cloud.google.com/vpc/docs/shared-vpc>

All other options make no sense.

48. Question

You need to grant access for three users so that they can view and edit table data on a Cloud Spanner instance. What should you do?

- Run gcloud iam roles describe roles/spanner.databaseUser. Add the users to the role.
- Run gcloud iam roles describe roles/spanner.viewer -- project my-gcp-ace-project. Add the users to a new group. Add the group to the role.

- Run gcloud iam roles describe roles/spanner.databaseUser. Add the users to a new group. Add the group to the role.
- Run gcloud iam roles describe roles/spanner.viewer -- project my-gcp-ace-project. Add the users to the role.

Unattempted

Our requirements

1. View and Edit table data
2. 3 users (i.e. multiple users)

3 users should give us the idea that we do not want to assign roles/permissions at the user level and instead want to do it based on groups so that we can create one group with all the required permissions and all such users who need this access can be assigned to the group.

Ref: <https://cloud.google.com/iam/docs/reference/rest/v1/Policy#Binding>

Ref: <https://cloud.google.com/iam/docs/granting-changing-revoking-access#granting-gcloud-manual>

Run gcloud iam roles describe roles/spanner.databaseUser. Add the users to the role. is not right.

We are looking for an option that assigns users to a group (in order to maximise reuse and minimize maintenance overhead). This option assigns users to the role so is not the right answer.

Run gcloud iam roles describe roles/spanner.viewer — project my-gcp-ace-project. Add the users to the role. is not right.

We are looking for an option that assigns users to a group (in order to maximise reuse and minimize maintenance overhead). This option assigns users to the role so is not the right answer.

Run gcloud iam roles describe roles/spanner.viewer — project my-gcp-ace-project. Add the users to a new group. Add the group to the role. is not right.

Adding users to a group and granting the role to the group is the right way forward. But the role used in this option is spanner.viewer which allows viewing all Cloud Spanner instances (but cannot modify instances), and allows viewing all Cloud Spanner databases (but cannot modify databases and cannot read from databases). Since we required edit access as well, this option is not right.

Ref: <https://cloud.google.com/spanner/docs/iam>

Run gcloud iam roles describe roles/spanner.databaseUser. Add the users to a new group. Add the group to the role. is the right answer.

Adding users to a group and granting the role to the group is the right way forward. In addition, we assign the role `spanner.databaseUser` which allows Read from and write to the Cloud Spanner database; execute SQL queries on the database, including DML and Partitioned DML; and View and update schema for the database. This is the only option that grants the right role to a group and assigns users to the group.

49. Question

You need to host an application on a Compute Engine instance in a project shared with other teams. You want to prevent the other teams from accidentally causing downtime on that application. Which feature should you use?

- Use a Shielded VM.
- Use a Preemptible VM.
- Use a sole-tenant node.
- **Enable deletion protection on the instance.**

Unattempted

Use a Shielded VM. is not right.

Shielded VMs are virtual machines (VMs) on Google Cloud hardened by a set of security controls that help defend against rootkits and boot kits. Using Shielded VMs helps protect enterprise workloads from threats like remote attacks, privilege escalation, and malicious insiders. But shielded VMs don't offer protection for accidental termination of the instance.

Ref: <https://cloud.google.com/shielded-vm>

Use a Preemptible VM. is not right.

A preemptible VM is an instance that you can create and run at a much lower price than normal instances. However, Compute Engine might terminate (preempt) these instances if it requires access to those resources for other tasks. Preemptible instances are excess Compute Engine capacity, so their availability varies with usage. Preemptible VMs don't offer protection for accidental termination of the instance.

Ref: <https://cloud.google.com/compute/docs/instances/preemptible>

Use a sole-tenant node. is not right.

Sole-tenancy lets you have exclusive access to a sole-tenant node, which is a physical Compute Engine server that is dedicated to hosting only your project's VMs. Use sole-tenant nodes to keep your VMs physically separated from VMs in other projects, or to group your VMs together on the same host hardware. Sole-tenant nodes don't offer

protection for accidental termination of the instance.

Ref: <https://cloud.google.com/compute/docs/nodes>

Enable deletion protection on the instance. is the right answer.

As part of your workload, there might be certain VM instances that are critical to running your application or services, such as an instance running a SQL server, a server used as a license manager, and so on. These VM instances might need to stay running indefinitely so you need a way to protect these VMs from being deleted. By setting the deletionProtection flag, a VM instance can be protected from accidental deletion. If a user attempts to delete a VM instance for which you have set the deletionProtection flag, the request fails. Only a user that has been granted a role with compute.instances.create permission can reset the flag to allow the resource to be deleted.

Ref: <https://cloud.google.com/compute/docs/instances/preventing-accidental-vm-deletion>

50. Question

You need to manage multiple Google Cloud Platform (GCP) projects in the fewest steps possible. You want to configure the Google Cloud SDK command line interface (CLI) so that you can easily manage multiple GCP projects. What should you?

- 1. Create a configuration for each project you need to manage. 2. Activate the appropriate configuration when you work with each of your assigned GCP projects.
- 1. Create a configuration for each project you need to manage. 2. Use gcloud init to update the configuration values when you need to work with a non-default project
- 1. Use the default configuration for one project you need to manage. 2. Activate the appropriate configuration when you work with each of your assigned GCP projects.
- 1. Use the default configuration for one project you need to manage. 2. Use gcloud init to update the configuration values when you need to work with a non-default project.

Unattempted

1. Create a configuration for each project you need to manage.
2. Activate the appropriate configuration when you work with each of your assigned GCP projects. is the right answer.

`gcloud` configurations enable you to manage multiple projects in `gcloud` cli using the fewest possible steps,

Ref: <https://cloud.google.com/sdk/gcloud/reference/config>

For example, we have two projects

```
$ gcloud projects list  
PROJECT_ID NAME PROJECT_NUMBER  
project-1-278333 project-1-278333 85524215451  
project-2-278333 project-2-278333 25349885274
```

We create configuration for each project. For project-2-278333,

```
$ gcloud config configurations create project-1-config  
$ gcloud config set project project-1-278333
```

And for project-2-278333,

```
$ gcloud config configurations create project-2-config  
$ gcloud config set project project-2-278333
```

We now have two configurations, one for each project.

```
$ gcloud config configurations list  
NAME IS_ACTIVE ACCOUNT PROJECT COMPUTE_DEFAULT_ZONE  
COMPUTE_DEFAULT_REGION  
cloudshell-4453 False  
project-1-config False project-1-278333  
project-2-config True project-2-278333
```

To activate configuration for project-1,

```
$ gcloud config configurations activate project-1-config  
Activated [project-1-config].  
$ gcloud config get-value project  
Your active configuration is: [project-1-config]  
project-1-278333
```

To activate configuration for project-2,

```
$ gcloud config configurations activate project-2-config  
Activated [project-2-config].  
$ gcloud config get-value project  
Your active configuration is: [project-2-config]  
project-2-278333
```

51. Question

You need to monitor resources that are distributed over different projects in Google Cloud Platform. You want to consolidate reporting under the same Stackdriver Monitoring dashboard. What should you do?

- Configure a single Stackdriver account, and link all projects to the same account.
- Use Shared VPC to connect all projects, and link Stackdriver to one of the projects.
- Configure a single Stackdriver account for one of the projects. In Stackdriver, create a Group and add the other project names as criteria for that Group.
- For each project, create a Stackdriver account. In each project, create a service account for that project and grant it the role of Stackdriver Account Editor in all other projects.

Unattempted

Use Shared VPC to connect all projects, and link Stackdriver to one of the projects. is not right.

Linking Stackdriver to one project brings metrics from that project alone. A Shared VPC allows an organization to connect resources from multiple projects to a common Virtual Private Cloud (VPC) network so that they can communicate with each other securely and efficiently using internal IPs from that network. But it does not help in linking all projects to a single Stackdriver workspace/account.

Ref: <https://cloud.google.com/vpc/docs/shared-vpc>

For each project, create a Stackdriver account. In each project, create a service account for that project and grant it the role of Stackdriver Account Editor in all other projects. is not right.

Stackdriver monitoring does not use roles to gather monitoring information from the project. Instead, the Stackdriver Monitoring agent, which is a collectd-based daemon, gathers system and application metrics from virtual machine instances and sends them to Monitoring. In this case, as each project is linked to a separate Stackdriver account, it is not possible to have a consolidated view of all monitoring.

Ref: <https://cloud.google.com/monitoring/agent>

Configure a single Stackdriver account for one of the projects. In Stackdriver, create a Group and add the other project names as criteria for that Group. is not right.

As the other projects are not linked to the stack driver, they can't be monitored.

Moreover, you can not add projects to Stackdriver groups. Groups provide a mechanism for alerting on the behavior of a set of resources, rather than on individual resources. For example, you can create an alerting policy that is triggered if some number of resources in the group violates a particular condition (for example, CPU load), rather than having each resource inform you of violations individually.

Ref: <https://cloud.google.com/monitoring/groups>

Configure a single Stackdriver account, and link all projects to the same account. is the right answer.

You can monitor resources of different projects in a single Stackdriver account by creating a Stackdriver workspace. A Stackdriver workspace is a tool for monitoring resources contained in one or more Google Cloud projects or AWS accounts. Each Workspace can have between 1 and 100 monitored projects, including Google Cloud projects and AWS accounts. A Workspace accesses metric data from its monitored projects, but the metric data and log entries remain in the individual projects.

Ref: <https://cloud.google.com/monitoring/workspaces>

52. Question

You need to produce a list of the enabled Google Cloud Platform APIs for a GCP project using the gcloud command line in the Cloud Shell. The project name is my-project. What should you do?

- Run gcloud projects list to get the project ID, and then run gcloud services list --project .
- Run gcloud init to set the current project to my-project, and then run gcloud services list --available.
- Run gcloud info to view the account value, and then run gcloud services list --account .
- Run gcloud projects describe to verify the project value, and then run gcloud services list --available.

Unattempted

Run gcloud init to set the current project to my-project, and then run gcloud services list --available. is not right.

--available return the services available to the project to enable and not the services that are enabled. This list will include any services that the project has already enabled plus the services that the project can enable.

Ref: <https://cloud.google.com/sdk/gcloud/reference/services/list>

Also, to set the current project, you need to use gcloud config set project Ref: <https://cloud.google.com/sdk/gcloud/reference/config/set>
gcloud init is used for initializing or reinitializing gcloud configurations.
<https://cloud.google.com/sdk/gcloud/reference/init>

Run gcloud info to view the account value, and then run gcloud services list --account . is not right.

We aren't passing any project id to the command so it would fail with the error shown below. (n.b. it is possible this command succeeds if you have an active gcloud configuration that has set the project so rather than accepting value from --project parameter, the command would obtain the project info from the gcloud configuration. The command shown below is run when no configuration is active).

gcloud services list --account

Errors with the following error.

ERROR: (gcloud.services.list) The project property is set to the empty string, which is invalid.

To set your project, run:

\$ gcloud config set project PROJECT_ID

or to unset it, run:

\$ gcloud config unset project

Run gcloud projects describe to verify the project value, and then run gcloud services list --available. is not right.

--available return the services available to the project to enable and not the services that are enabled. This list will include any services that the project has already enabled plus the services that the project can enable.

Ref: <https://cloud.google.com/sdk/gcloud/reference/services/list>

Run gcloud projects list to get the project ID, and then run gcloud services list --project . is the right answer.

For the gcloud services list command, --enabled is the default.

So running

gcloud services list --project is the same as running

gcloud services list --project --enabled

which would get all the enabled services for the project.

53. Question

You need to provide a cost estimate for a Kubernetes cluster using the GCP pricing calculator for Kubernetes. Your workload requires high IOPS, and you will also be using disk snapshots. You start by entering the number of nodes, average hours, and average days. What should you do next?

- Fill in local SSD. Fill in persistent disk storage and snapshot storage.
- Fill in local SSD. Add estimated cost for cluster management.
- Select Add GPUs. Fill in persistent disk storage and snapshot storage.
- Select Add GPUs. Add estimated cost for cluster management.

Unattempted

Fill in local SSD. Add estimated cost for cluster management. is not right.
You don't need to add an estimated cost for cluster management as the calculator automatically applies this. At the time of writing this, GKE clusters accrue a management fee of \$0.10 per cluster per hour, irrespective of cluster size or topology. One zonal cluster per billing account is free. GKE cluster management fees do not apply to Anthos GKE clusters.

Ref: <https://cloud.google.com/kubernetes-engine/pricing>

Select Add GPUs. Add estimated cost for cluster management. is not right.
You don't need to add an estimated cost for cluster management as the calculator automatically applies this. At the time of writing this, GKE clusters accrue a management fee of \$0.10 per cluster per hour, irrespective of cluster size or topology. One zonal cluster per billing account is free. GKE cluster management fees do not apply to Anthos GKE clusters.

Ref: <https://cloud.google.com/kubernetes-engine/pricing>

Select Add GPUs. Fill in persistent disk storage and snapshot storage. is not right.
GPUs don't help us with our requirement of high IOPS. Compute Engine provides graphics processing units (GPUs) that you can add to your virtual machine instances to accelerate specific workloads on your instances such as machine learning and data processing. But this doesn't help increase IOPS.

Ref: <https://cloud.google.com/compute/docs/gpus>

Fill in local SSD. Fill in persistent disk storage and snapshot storage. is the right answer.
The pricing calculator for Kubernetes Engine offers us the ability to add GPUs as well as specify Local SSD requirements for estimation. GPUs don't help us with our requirement of high IOPS but Local SSD does.

Ref: <https://cloud.google.com/products/calculator>

GKE offers always-encrypted local solid-state drive (SSD) block storage. Local SSDs are physically attached to the server that hosts the virtual machine instance for very high input/output operations per second (IOPS) and very low latency compared to persistent disks.

Ref: <https://cloud.google.com/kubernetes-engine>

Once you fill in the local SSD requirement, you can fill in persistent disk storage and snapshot storage.

54. Question

You need to reduce GCP service costs for a division of your company using the fewest possible steps. You need to turn off all configured services in an existing GCP project. What should you do?

- 1. Verify that you are assigned the Project Owners IAM role for this project. 2. Locate the project in the GCP console, click Shut down and then enter the project ID.
- 1. Verify that you are assigned the Project Owners IAM role for this project. 2. Switch to the project in the GCP console, locate the resources and delete them.
- 1. Verify that you are assigned the Organization Administrator IAM role for this project. 2. Locate the project in the GCP console, enter the project ID and then click Shut down.
- 1. Verify that you are assigned the Organization Administrator IAM role for this project. 2. Switch to the project in the GCP console, locate the resources and delete them.

Unattempted

1. Verify that you are assigned the Organization Administrator IAM role for this project.
2. Locate the project in the GCP console, enter the project ID and then click Shut down. is not right.

Organization Admin role provides permissions to get and list projects but not shutdown projects.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#resource-manager-roles>

1. Verify that you are assigned the Organization Administrator IAM role for this project.
2. Switch to the project in the GCP console, locate the resources and delete them. is not right.

Organization Admin role provides permissions to get and list projects but not delete

projects.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#resource-manager-roles>

1. Verify that you are assigned the Project Owner IAM role for this project.
2. Switch to the project in the GCP console, locate the resources and delete them. is not right.

The primitive Project Owner role provides permissionst to delete project

https://cloud.google.com/iam/docs/understanding-roles#primitive_roles

But locating all the resources and deleting them is a manual task, time consuming and error prone. Our goal is to accomplish the same but with fewest possible steps

1. Verify that you are assigned the Project Owner IAM role for this project.
2. Locate the project in the GCP console, click Shut down and then enter the project ID. is the right answer.

The primitive Project Owner role provides permissionst to delete project

https://cloud.google.com/iam/docs/understanding-roles#primitive_roles

You can shut down projects using the Cloud Console. When you shut down a project, this immediately happens: All billing and traffic serving stops, You lose access to the project, The owners of the project will be notified and can stop the deletion within 30 days, The project will be scheduled to be deleted after 30 days. However, some resources may be deleted much earlier.

Ref: https://cloud.google.com/resource-manager/docs/creating-managing-projects#shutting_down_projects

55. Question

You need to run an important query in BigQuery but expect it to return a lot of records.

You want to find out how much it will cost to run the query. You are using on-demand pricing. What should you do?

- Run a select count (*) to get an idea of how many records your query will look through. Then convert that number of rows to dollars using the Pricing Calculator.
- Arrange to switch to Flat-Rate pricing for this query, then move back to on-demand.
- Use the command line to run a dry run query to estimate the number of bytes returned. Then convert that bytes estimate to dollars using the Pricing Calculator.
- Use the command line to run a dry run query to estimate the number of bytes read. Then convert that bytes estimate to dollars using the Pricing Calculator.

Unattempted

Arrange to switch to Flat-Rate pricing for this query, then move back to on-demand. is not right.

The cost of acquiring a big query slot (associated with flat-rate pricing) is significantly higher than our requirement here to run a single important query or just to know how much it would cost to run that query. BigQuery offers flat-rate pricing for customers who prefer a stable monthly cost for queries rather than paying the on-demand price per TB of data processed. You enroll in flat-rate pricing by purchasing slot commitments, measured in BigQuery slots. Slot commitments start at 500 slots and the price starts from \$10000. Your queries consume this slot capacity, and you are not billed for bytes processed.

Ref: https://cloud.google.com/bigquery/pricing#flat_rate_pricing

Use the command line to run a dry run query to estimate the number of bytes returned.

Then convert that bytes estimate to dollars using the Pricing Calculator. is not right.

Under on-demand pricing, BigQuery doesn't charge for the query execution based on the output of the query (i.e. bytes returned) but on the number of bytes processed (also referred to as bytes read or bytes scanned) in order to arrive at the output of the query.

You are charged for the number of bytes processed whether the data is stored in BigQuery or in an external data source such as Cloud Storage, Google Drive, or Cloud Bigtable. On-demand pricing is based solely on usage. You are charged for the bytes scanned even if your query itself doesn't return any data.

Ref: <https://cloud.google.com/bigquery/pricing>

Run a select count (*) to get an idea of how many records your query will look through.

Then convert that number of rows to dollars using the Pricing Calculator. is not right.

This is not as practical as identifying the number of records your query will look through (i.e. scan/process) is not straightforward. Plus BigQuery supports external data sources such as Cloud Storage, Google Drive, or Cloud Bigtable; and the developer cost associated with identifying this information from various data sources is significant, not practical and sometimes not possible.

Use the command line to run a dry run query to estimate the number of bytes read. Then convert that bytes estimate to dollars using the Pricing Calculator. is the right answer.

BigQuery pricing is based on the number of bytes processed/read. Under on-demand pricing, BigQuery charges for queries by using one metric: the number of bytes processed (also referred to as bytes read). You are charged for the number of bytes processed whether the data is stored in BigQuery or in an external data source such as Cloud Storage, Google Drive, or Cloud Bigtable. On-demand pricing is based solely on usage.

Ref: <https://cloud.google.com/bigquery/pricing>

56. Question

You need to select and configure compute resources for a set of batch processing jobs. These jobs take around 2 hours to complete and are run nightly. You want to minimize service costs. What should you do?

- Select Compute Engine. Use VM instance types that support micro bursting.
- Select GKE. Use a single node cluster with a small instance type
- Select Compute Engine. Use preemptible VM instances of the appropriate standard machine types.
- Select GKE. Use a three-node cluster with micro instance types.

Unattempted

Requirements – achieve end goal while minimizing service costs.

Select GKE. Use a single node cluster with a small instance type is not right.
We do not know if a small instance is capable of handling all the batch volume. Plus this is not the most cost-effective of the options.

Select GKE. Use a three-node cluster with micro instance types is not right.
We do not know if three micro instances are capable of handling all the batch volume.
Plus this is not the most cost-effective of the options.

Select Compute Engine. Use VM instance types that support micro bursting is not right.
We can use an instance that supports micro bursting but we have a job that runs for 2 hours. Bursting is suitable for short periods.

Select Compute Engine. Use preemptible VM instances of the appropriate standard machine types is the right answer.
We minimize the cost by selecting a preemptible instance of the appropriate type. If the preemptible instance is terminated, the next nightly run picks up the unprocessed volume.

57. Question

You need to set up a policy so that videos stored in a specific Cloud Storage Regional bucket are moved to Coldline after 90 days and then deleted after one year from their creation. How should you set up the policy?

- Use Cloud Storage Object Lifecycle Management using Age conditions with SetStorageClass and Delete actions. Set the SetStorageClass action to 90 days and the Delete action to 365 days.

- Use gsutil rewrite and set the Delete action to 365 days.
- Use gsutil rewrite and set the Delete action to 275 days (365 - 90).
- Use Cloud Storage Object Lifecycle Management using Age conditions with SetStorageClass and Delete actions. Set the SetStorageClass action to 90 days and the Delete action to 275 days (365 - 90)

Unattempted

Use gsutil rewrite and set the Delete action to 275 days (365 - 90). is not right.
gsutil rewrite is used to change the storage class of objects within a bucket through overwriting the object. It does not support Delete action.

Ref: <https://cloud.google.com/storage/docs/changing-storage-classes>

Use gsutil rewrite and set the Delete action to 365 days. is not right.
gsutil rewrite is used to change the storage class of objects within a bucket through overwriting the object. It does not support Delete action.

Ref: <https://cloud.google.com/storage/docs/changing-storage-classes>

Use Cloud Storage Object Lifecycle Management using Age conditions with SetStorageClass and Delete actions. Set the SetStorageClass action to 90 days and the Delete action to 275 days (365 - 90). is not right.

Object Lifecycle Management does not rewrite an object when changing its storage class. This means that when an object is transitioned to Nearline Storage, Coldline Storage, or Archive Storage using the SetStorageClass feature, any subsequent early deletion and associated charges are based on the original creation time of the object, regardless of when the storage class changed.

If however, the change of storage class is done manually using a rewrite, the creation time of the objects is the new creation time since they are rewritten. In such a case, you would need to apply a lifecycle delete action of 275 days.

Ref: <https://cloud.google.com/storage/docs/lifecycle>

Use Cloud Storage Object Lifecycle Management using Age conditions with SetStorageClass and Delete actions. Set the SetStorageClass action to 90 days and the Delete action to 365 days. is the right answer.

Object Lifecycle Management does not rewrite an object when changing its storage class. This means that when an object is transitioned to Nearline Storage, Coldline Storage, or Archive Storage using the SetStorageClass feature, any subsequent early deletion and associated charges are based on the original creation time of the object, regardless of when the storage class changed.

Ref: <https://cloud.google.com/storage/docs/lifecycle>

58. Question

You need to set up permissions for a set of Compute Engine instances to enable them to write data into a particular Cloud Storage bucket. You want to follow Google-recommended practices. What should you do?

- Create a service account with an access scope. Use the access scope '<https://www.googleapis.com/auth/cloud-platform>'.
- Create a service account with an access scope. Use the access scope 'https://www.googleapis.com/auth/devstorage.write_only'.
- Create a service account and add it to the IAM role 'storage.objectAdmin' for that bucket.
- Create a service account and add it to the IAM role 'storage.objectCreator' for that bucket.

Unattempted

Our requirements are

1. Google recommended practices
2. Multiple compute engine instances to write data to a bucket.

Create a service account with an access scope. Use the access scope 'https://www.googleapis.com/auth/devstorage.write_only'. is not right.
There is no scope called "write_only".

Ref: <https://cloud.google.com/storage/docs/authentication>

Create a service account and add it to the IAM role 'storage.objectCreator' for that bucket. is not right.

You can't add a service account to a role. The relationship is the other way round. You grant roles to the service account. See below a screenshot of the role.

Ref: <https://cloud.google.com/iam/docs/granting-roles-to-service-accounts>

Create a service account and add it to the IAM role 'storage.objectAdmin' for that bucket. is not right.

You can't add a service account to a role. The relationship is the other way round. You grant roles to the service account.

Ref: <https://cloud.google.com/iam/docs/granting-roles-to-service-accounts>

Create a service account with an access scope. Use the access scope '<https://www.googleapis.com/auth/cloud-platform>'. is the right answer.

cloud-platform role lets you view and manage data across all Google Cloud services. For Cloud Storage, this is the same as devstorage.full-control which allows full control over data, including the ability to modify IAM policies.

Ref: <https://cloud.google.com/storage/docs/authentication>

59. Question

You need to trigger a budget alert for Compute Engine charges on one of the three Google Cloud Platform projects that you manage. All three projects are linked to a single billing account. What should you do?

- Verify that you have Billing Account Administrator role. Select the associated billing account and create a budget and alert for the appropriate project.
- Verify that you have Billing Account Administrator role. Select the associated billing account and create a budget and a custom alert.
- Verify that you have Project Billing Manager role. Select the associated billing account and create a budget for the appropriate project.
- Verify that you have Project Billing Manager role. Select the associated billing account and create a budget and a custom alert.

Unattempted

Verify that you have Project Billing Manager role. Select the associated billing account and create a budget for the appropriate project. is not right.

Project Billing Manager role allows a user to attach the project to the billing account, but does not grant any rights over resources. This role does not provide user permissions to view spending, create budgets and alerts.

Ref: <https://cloud.google.com/billing/docs/how-to/billing-access#overview-of-cloud-billing-roles-in-cloud-iam>

Verify that you have Project Billing Manager role. Select the associated billing account and create a budget and a custom alert. is not right.

Project Billing Manager role allows a user to attach the project to the billing account, but does not grant any rights over resources. This role does not provide user permissions to view spending, create budgets and alerts.

Ref: <https://cloud.google.com/billing/docs/how-to/billing-access#overview-of-cloud-billing-roles-in-cloud-iam>

Verify that you have Billing Account Administrator role. Select the associated billing account and create a budget and a custom alert. is not right.

Billing Account Administrator role enables a user to view spend and set budget alerts. But the budget here isn't scoped to the single project that we are interested in. Since the single billing account is linked to all three projects, this results in budget alerts being triggered for Compute Engine usage on all three projects – which is against our requirements.

Ref: <https://cloud.google.com/billing/docs/how-to/billing-access#overview-of-cloud-billing-roles-in-cloud-iam>

Ref: <https://cloud.google.com/billing/docs/how-to/budgets#budget-scope>

Verify that you have Billing Account Administrator role. Select the associated billing account and create a budget and alert for the appropriate project. is the right answer.

Billing Account Administrator role enables a user to view spend and set budget alerts. In addition, the budget here is scoped to a single project. Therefore, when the compute engine spend exceeds the budget threshold in the project, we send an alert, and this only works for the scoped project, and not all projects linked to the billing account.

Ref: <https://cloud.google.com/billing/docs/how-to/billing-access#overview-of-cloud-billing-roles-in-cloud-iam>

Ref: <https://cloud.google.com/billing/docs/how-to/budgets#budget-scope>

60. Question

You need to update a deployment in Deployment Manager without any resource downtime in the deployment. Which command should you use?

- `gcloud deployment-manager deployments update --config`
- `gcloud deployment-manager deployments create --config`
- `gcloud deployment-manager resources create --config`
- `gcloud deployment-manager resources update --config`

Unattempted

gcloud deployment-manager resources create -config . is not right.

gcloud deployment-manager resources command does not support the action create. The supported actions are describe and list. So this option is not right.

Ref: <https://cloud.google.com/sdk/gcloud/reference/deployment-manager/resources>

gcloud deployment-manager resources update -config . is not right.

gcloud deployment-manager resources command does not support the action update. The supported actions are describe and list. So this option is not right.

Ref: <https://cloud.google.com/sdk/gcloud/reference/deployment-manager/resources>

gcloud deployment-manager deployments create -config . is not right.

gcloud deployment-manager deployments create – creates a deployment but we want to update a deployment. So this option is not right.

Ref: <https://cloud.google.com/sdk/gcloud/reference/deployment-manager/deployments/create>

gcloud deployment-manager deployments update -config . is the right answer.

gcloud deployment-manager deployments update – updates a deployment based on a provided config file and fits our requirement.

<https://cloud.google.com/sdk/gcloud/reference/deployment-manager/deployments/update>

61. Question

You plan to deploy an application on an autoscaled managed instances group. The application uses a tomcat server and runs on port 8080. You want to access the application on <https://www.example.com>. You want to follow Google recommended practices. What services would you use?

- Google Domains, Cloud DNS private zone, HTTP(S) Load Balancer
- Google Domains, Cloud DNS private zone, SSL Proxy Load Balancer
- Google DNS, Google CDN, SSL Proxy Load Balancer
- **Google Domains, Cloud DNS, HTTP(S) Load Balancer**

Unattempted

To serve traffic on <https://www.example.com>, we have to first own the domain example.com. We can use Google Domains service to register a domain.

Ref: <https://domains.google/>

Once we own example.com domain, we need to create a zone <http://www.example.com>. We can use Cloud DNS, which is a scalable, reliable, and managed authoritative Domain Name System (DNS) to create a DNS zone.

Ref: <https://cloud.google.com/dns>

Once the <http://www.example.com> zone is set up, we need to create a DNS (A) record to point to the public IP of the Load Balancer. This is also carried out in Cloud DNS.

Finally, we need a load balancer to front the autoscaled managed instances group. Google recommends we use HTTP(S) Load Balancer for this requirement as “SSL Proxy Load Balancing is intended for non-HTTP(S) traffic. For HTTP(S) traffic, we recommend that you use HTTP(S) Load Balancing.”

Ref: <https://cloud.google.com/load-balancing/docs/ssl>

So the correct answer is Google Domains, Cloud DNS, HTTP(S) Load Balancer

62. Question

You ran the following commands to create two compute instances.

```
gcloud compute instances create instance1
```

```
gcloud compute instances create instance2
```

Both compute instances were created in europe-west2-a zone but you want to create them in other zones. Your active gcloud configuration is as shown below.

```
$ gcloud config list
[component_manager]
disable_update_check = True
[compute]
gce_metadata_read_timeout_sec = 5
zone = europe-west2-a
[core]
account = gcp-ace-lab-user@gmail.com
disable_usage_reporting = False
project = gcp-ace-lab-266520
[metrics]
environment = devshell
```

You want to modify the gcloud configuration such that you are prompted for a zone when you execute the create instance commands above. What should you do?

- `gcloud config unset compute/zone`
- `gcloud config set zone ""`
- `gcloud config set compute/zone ""`
- `gcloud config unset zone`

Unattempted

`gcloud config unset zone.` is not right.

`gcloud config` does not have a `core/zone` property. The syntax for this command is `gcloud config unset SECTION/PROPERTY`. If `SECTION` is missing from the command, `SECTION` is defaulted to `core`. We are effectively trying to run the command `gcloud config unset core/zone` but the `core` section doesn't have a property called `zone`, so this command fails.

`$ gcloud config unset zone`

`ERROR: (gcloud.config.unset) Section [core] has no property [zone].`

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/unset>

`gcloud config set zone "".` is not right.

`gcloud config` does not have a `core/zone` property. The syntax for this command is `gcloud config set SECTION/PROPERTY VALUE`. If `SECTION` is missing, `SECTION` is defaulted to `core`. We are effectively trying to run `gcloud config set core/zone ""` but the `core` section doesn't have a property called `zone`, so this command fails.

`$ gcloud config set zone ""`

`ERROR: (gcloud.config.unset) Section [core] has no property [zone].`

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/set>

`gcloud config set compute/zone "".` is not right.

This command uses the correct syntax but it doesn't unset the `compute/zone` property correctly. Instead it sets it to `""` in `gcloud` configuration. When the `gcloud compute instances create` command runs, it picks the `zone` value from this configuration property which is `""` and attempts to create an instance in `""` zone and fails because zone `""` doesn't exist. `gcloud` doesn't treat `""` zone as an unset value. The `zone` must be explicitly unset if it is to be removed from the configuration.

`$ gcloud config set compute/zone ""`

`$ gcloud compute instances create instance1`

`Zone: Expected type (,) for field id, found projects/compute-challenge-lab-266520/zones/ (type)`

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/set>

`gcloud config unset compute/zone`. is the right answer.

This command uses the correct syntax and correctly unsets the zone in `gcloud` configuration. The next time `gcloud compute instances create` command runs, it knows there is no default zone defined in `gcloud` configuration and therefore prompts for a zone before the instance can be created.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/unset>

63. Question

You recently deployed a new application in Google App Engine to serve production traffic. After analyzing logs for various user flows, you uncovered several issues in your application code and have developed a fix to address the issues. Parts of your proposed fix could not be validated in the pre-production environment by your testing team as some of the scenarios can only be validated by an end-user with access to specific data in your production environment. In the company's weekly Change Approval Board meeting, concerns were raised that the fix could possibly take down the application. It was unanimously agreed that while the fix is risky, it is a necessary change to the application. You have been asked to suggest a solution that minimizes the impact of the change going wrong. You also want to minimize costs. What should you do?

- Deploy a new version of the application, and use traffic splitting to send a small percentage of traffic to it.
- Set up a second Google App Engine service, and then update a subset of clients to hit the new service.
- Deploy the new application version temporarily, capture logs and then roll it back to the previous version.
- Create a second Google App Engine project with the new application code, and onboard users gradually to the new application.

Unattempted

Deploy the new application version temporarily, capture logs and then roll it back to the previous version. is not right.

Deploying a new application version and promoting it would result in your new version serving all production traffic. If the code fix doesn't work as expected, it would result in the application becoming unreachable to all users. This is a risky approach and should be avoided.

Create a second Google App Engine project with the new application code, and onboard users gradually to the new application. is not right.

You want to minimize costs. This approach effectively doubles your costs as you have to pay for two identical environments until all users are moved over to the new application. There is an additional overhead of manually onboarding users to the new application which could be expensive as well as time-consuming.

Set up a second Google App Engine service, and then update a subset of clients to hit the new service. is not right.

It is not straightforward to update a set of clients to hit the new service. When users access an App Engine service, they use an endpoint like `https://SERVICE_ID-dot-PROJECT_ID.REGION_ID.r.appspot.com`. Introducing a new service introduces a new URL and getting your users to use the new URL is possible but involves effort and coordination. If you want to mask these differences to the end-user, then you have to make changes in the DNS and use a weighted algorithm to split the traffic between the two services based on the weights assigned.

Ref: <https://cloud.google.com/appengine/docs/standard/python/splitting-traffic>

Ref: <https://cloud.google.com/appengine/docs/standard/python/an-overview-of-app-engine>

This approach also has the drawback of doubling your costs until all users are moved over to the new service.

Deploy a new version of the application, and use traffic splitting to send a small percentage of traffic to it. is the right answer.

This option minimizes the risk to the application while also minimizing the complexity and cost. When you deploy a new version to App Engine, you can choose not to promote it to serve live traffic. Instead, you could set up traffic splitting to split traffic between the two versions – this can all be done within Google App Engine. Once you send a small portion of traffic to the new version, you can analyze logs to identify if the fix has worked as expected. If the fix hasn't worked, you can update your traffic splitting configuration to send all traffic back to the old version. If you are happy your fix has worked, you can send more traffic to the new version or move all user traffic to the new version and delete the old version.

Ref: <https://cloud.google.com/appengine/docs/standard/python/splitting-traffic>

Ref: <https://cloud.google.com/appengine/docs/standard/python/an-overview-of-app-engine>

64. Question

You recently deployed a new version of an application to App Engine and then discovered a bug in the release. You need to immediately revert to the prior version of the application. What should you do?

- On the App Engine page of the GCP Console, select the application that needs to be reverted and click Revert.
- On the App Engine Versions page of the GCP Console, route 100% of the traffic to the previous version.
- Deploy the original version as a separate application. Then go to App Engine settings and split traffic between applications so that the original version serves 100% of the requests.
- Run `gcloud app restore`.

Unattempted

Run `gcloud app restore`. is not right.
restore action is not supported by `gcloud app` command.
Ref: <https://cloud.google.com/sdk/gcloud/reference/app/deploy>

On the App Engine page of the GCP Console, select the application that needs to be reverted and click Revert. is not right.

Revert option is not present on the App Engine page of the GCP Console.

Deploy the original version as a separate application. Then go to App Engine settings and split traffic between applications so that the original version serves 100% of the requests. is not right.

Each application in the app engine is different and it is not possible to split traffic between applications in App Engine. You can use traffic splitting to specify a percentage distribution of traffic across two or more of the versions within a service but not across applications.

Ref: <https://cloud.google.com/appengine/docs/standard/python/splitting-traffic>

On the App Engine Versions page of the GCP Console, route 100% of the traffic to the previous version. is the right answer

You can roll back to a previous version in the app engine GCP console. Go back to the list of versions and check the box next to the version that you want to receive all traffic and click the MAKE DEFAULT button located above the list. Traffic immediately switches over to the selected version.

Ref: <https://cloud.google.com/community/tutorials/how-to-roll-your-app-engine-managed-vms-app-back-to-a-previous-version-part-1>

65. Question

You significantly changed a complex deployment manager template and want to confirm that the dependencies of all defined resources are properly met before committing it to the project. You want the most rapid feedback on your changes. What would you do?

- Use granular logging statements within the Deployment Manager template authored in Python.
- Monitor activity of the Deployment Manager executing on Stackdriver logging page of the GCP console.
- Execute the Deployment manager template against a separate project with the same configuration, and monitor for failures
- Execute the Deployment Manager template using the --preview option in the same project, and observe the status of interdependent resources

Unattempted

Requirements – confirm dependencies, rapid feedback.

Use granular logging statements within the Deployment Manager template authored in Python. is not right.

Deployment Manager doesn't provide the ability to set granular logging statements. Moreover, if that was possible the logging statements wouldn't be written to a log file until the template is applied and it is already too late as the template is applied and we haven't had a chance to confirm that the dependencies of all defined resources are properly met

Monitor activity of the Deployment Manager executing on Stackdriver logging page of the GCP console. is not right.

This doesn't give us a chance to confirm that the dependencies of all defined resources are properly met before executing it.

Execute the Deployment manager template against a separate project with the same configuration, and monitor for failures. is not right.

While we can identify whether dependencies are met by monitoring the failures, it is not rapid. We need rapid feedback on changes and we want that before changes are committed (i.e. applied) to the project

Execute the Deployment Manager template using the `-preview` option in the same project, and observe the status of interdependent resources. Is the right answer.

After we have written a configuration file, we can preview the configuration before you create a deployment. Previewing a configuration lets you see the resources that

Deployment Manager would create but does not actually instantiate any actual resources.

In `gcloud` command-line, you use the `create` sub-command with the `-preview` flag to preview configuration changes.

Ref: <https://cloud.google.com/deployment-manager>

66. Question

You want to add a new auditor to a Google Cloud Platform project. The auditor should be allowed to read, but not modify, all project items. How should you configure the auditor's permissions?

- Create a custom role with view-only project permissions. Add the user's account to the custom role.
- Create a custom role with view-only service permissions. Add the user's account to the custom role.
- Select the built-in IAM project Viewer role. Add the user's account to this role.
- Select the built-in IAM service Viewer role. Add the user's account to this role.

Unattempted

Select the built-in IAM project Viewer role. Add the user's account to this role. Is the right answer

The primitive role roles/viewer provides read access to all resources in the project. The permissions in this role are limited to Get and list access for all resources. As we have an out of the box role that exactly fits our requirement, we should use this.

Ref: <https://cloud.google.com/resource-manager/docs/access-control-project>

It is advisable to use the existing GCP provided roles over creating custom roles with similar permissions as this becomes a maintenance overhead. If GCP modifies how permissions are handled or adds/removes permissions, the default GCP provided roles are automatically updated by Google whereas if they were custom roles, the responsibility is with us and this adds to the operational overhead and needs to be avoided.

67. Question

You want to configure 10 Compute Engine instances for availability when maintenance occurs. Your requirements state that these instances should attempt to automatically restart if they crash. Also, the instances should be highly available including during system maintenance. What should you do?

- Create an instance group for the instances. Verify that the Advanced creation options setting for do not retry machine creation is set to off.
- Create an instance group for the instances. Set the Autohealing health check to healthy (HTTP).
- Create an instance template for the instances. Set Automatic Restart to off. Set On-host maintenance to Terminate VM instances. Add the instance template to an instance group.
- Create an instance template for the instances. Set the Automatic Restart to on. Set the On-host maintenance to Migrate VM instance. Add the instance template to an instance group.

Unattempted

Requirements

1. 10 instances – indicates we need to look for MIG (Managed Instances Group) where we can configure healing/scaling settings. All options talk about creating an instance group so this point isn't of much use, unfortunately.
2. Highly available during system maintenance – indicates we need to look for Live Migration.
3. Automatically restart on crash – indicates we need to look for options that enable automatic restarts.

Create an instance template for the instances.

Set Automatic Restart to off.

Set On-host maintenance to Terminate VM instances.

Add the instance template to an instance group. is not right.

If Automatic Restart is off, then the compute engine instances are not automatically restarted. This results in loss of capacity and if GCP decides to start system maintenance on all instances at the same time, all instances are down and this does not meet our requirement "Highly available during system maintenance" so this option is not right.

Create an instance group for the instances.

Set the Autohealing health check to healthy (HTTP). is not right.

While auto-healing helps with the recreation of VM instances when needed, it doesn't live-migrate the instances so our requirement of "highly available including during system maintenance" is not met. More info about Autohealing - Auto-healing allows the recreation of VM instances when needed. You can use a health check to recreate a VM instance if the health check finds it unresponsive. If you don't select a health check, Compute Engine will recreate VM instances only when they're not running.

Ref: https://cloud.google.com/compute/docs/instance-groups/?hl=en_GB#managed_instance_groups_and_autohealing

Create an instance group for the instance.

Verify that the Advanced creation options setting for do not retry machine creation is set to off. is not right.

Like above - this option doesn't live-migrate the instances so our requirement of "highly available including during system maintenance" is not met.

Create an instance template for the instances.

Set the Automatic Restart to on.

Set the On-host maintenance to Migrate VM instance.

Add the instance template to an instance group. is the right option.

Enabling automatic restart ensures that compute engine instances are automatically restarted when they crash. And Enabling "Migrate VM Instance" enables live migrates i.e. compute instances are migrated during system maintenance and remain running during the migration.

Automatic Restart - If your instance is set to terminate when there is a maintenance event, or if your instance crashes because of an underlying hardware issue, you can set up Compute Engine to automatically restart the instance by setting the automaticRestart field to true. This setting does not apply if the instance is taken offline through a user action, such as calling sudo shutdown, or during a zone outage.

Ref: <https://cloud.google.com/compute/docs/instances/setting-instance-scheduling-options#autorestart>

Enabling the Migrate VM Instance option migrates your instance away from an infrastructure maintenance event, and your instance remains running during the migration. Your instance might experience a short period of decreased performance, although generally, most instances should not notice any difference. This is ideal for instances that require constant uptime and can tolerate a short period of decreased performance.

Ref: https://cloud.google.com/compute/docs/instances/setting-instance-scheduling-options#live_migrate

68. Question

You want to configure a cost-effective solution for archiving objects in a Cloud Storage bucket. Noncurrent versions should be archived after 30 days. Non-current versions are accessed once a month for reporting. This archived objects are also occasionally updated at month-end. What should you do?

- Add a bucket lifecycle rule that archives noncurrent versions after 30 days to Coldline Storage.
- Add a bucket lifecycle rule that archives noncurrent versions after 30 days to Nearline Storage.
- Add a bucket lifecycle rule that archives objects from regional storage after 30 days to Coldline Storage.
- Add a bucket lifecycle rule that archives objects from regional storage after 30 days to Nearline Storage.

Unattempted

We don't know what the current storage class is. In the absence of this information and considering the 4 options provided, it is safe to assume that objects are currently in Regional or Multi-Regional buckets. We want to archive noncurrent versions after 30 days and you need to read and modify on average once per month

Add a bucket lifecycle rule that archives noncurrent versions after 30 days to Coldline Storage. is not right.

Coldline Storage is ideal for data you plan to read or modify at most once a quarter. Our requirement states we need to read or modify every month so Coldline Storage is not an ideal storage class for our requirement.

Ref: <https://cloud.google.com/storage/docs/storage-classes#coldline>

Add a bucket lifecycle rule that archives objects from regional storage after 30 days to Coldline Storage. is not right.

Coldline Storage is ideal for data you plan to read or modify at most once a quarter. Our requirement states we need to read or modify every month so Coldline Storage is not an ideal storage class for our requirement. Moreover, we don't want to archive live versions, we want to archive just the noncurrent versions.

Ref: <https://cloud.google.com/storage/docs/storage-classes#coldline>

Add a bucket lifecycle rule that archives objects from regional storage after 30 days to Nearline Storage. is not right.

While Nearline Storage is ideal for data you plan to read or modify on average once per month or less, we don't want to archive live versions, we want to archive just the

noncurrent versions.

Ref: <https://cloud.google.com/storage/docs/storage-classes#nearline>

Add a bucket lifecycle rule that archives noncurrent versions after 30 days to Nearline Storage. is the right answer.

Nearline Storage is ideal for data you plan to read or modify on average once per month or less. And this option archives just the noncurrent versions which is what we want to do.

<https://cloud.google.com/storage/docs/storage-classes#nearline>

69. Question

You want to configure an SSH connection to a single Compute Engine instance for users in the dev1 group. This instance is the only resource in this particular Google Cloud Platform project that the dev1 users should be able to connect to. What should you do?

- Set metadata to enable-oslogin=true for the instance. Grant the dev1 group the compute.osLogin role. Direct them to use the Cloud Shell to ssh to that instance.
- Set metadata to enable-oslogin=true for the instance. Set the service account to no service account for that instance. Direct them to use the Cloud Shell to ssh to that instance.
- Enable block project wide keys for the instance. Generate an SSH key for each user in the dev1 group. Distribute the keys to dev1 users and direct them to use their third-party tools to connect.
- Enable block project wide keys for the instance. Generate an SSH key and associate the key with that instance. Distribute the key to dev1 users and direct them to use their third-party tools to connect.

Unattempted

You have multiple ways to connect to instances. More information can be found here: <https://cloud.google.com/compute/docs/instances/access-overview>

Enable block project wide keys for the instance. Generate an SSH key for each user in the dev1 group. Distribute the keys to dev1 users and direct them to use their third-party tools to connect. is not right.

Generating SSH keys for users is fine but unless the SSH keys are added to the instance, users would not be able to SSH to the server. If you need your instance to ignore project-wide public SSH keys and use only the instance-level keys, you can block project-wide

public SSH keys from the instance. This allows only users whose public SSH key is stored in instance-level metadata to access the instance.

Ref: <https://cloud.google.com/compute/docs/instances/adding-removing-ssh-keys#edit-ssh-metadata>

Enable block project wide keys for the instance. Generate an SSH key and associate the key with that instance. Distribute the key to dev1 users and direct them to use their third-party tools to connect. is not right.

While this is possible, sharing SSH keys is a strict NO from a security point of view as this breaks auditing. Should one of the developers create a disaster (either accidental or malicious), your security admin would be unable to identify which of the users in dev1 group caused the issue.

Set metadata to enable-oslogin=true for the instance. Set the service account to no service account for that instance. Direct them to use the Cloud Shell to ssh to that instance. is not right.

After you enable OS Login on one or more instances in your project, those VMs accept connections only from user accounts that have the necessary IAM roles in your project or organization. In this case, since we have not granted either of these roles – roles/compute.osLogin or roles/compute.osAdminLogin role, users of dev1 group can't SSH to the server.

Ref: https://cloud.google.com/compute/docs/instances/managing-instance-access#configure_users

Set metadata to enable-oslogin=true for the instance. Grant the dev1 group the compute.osLogin role. Direct them to use the Cloud Shell to ssh to that instance. is the right answer.

After you enable OS Login on one or more instances in your project, those VMs accept connections only from user accounts that have the necessary IAM roles in your project or organization. In this case, we are granting the group compute.osLogin which lets them log in as non-administrator account. And since we are directing them to use Cloud Shell to ssh, we don't need to add their SSH keys to the instance metadata.

Ref: https://cloud.google.com/compute/docs/instances/managing-instance-access#configure_users

Ref: https://cloud.google.com/compute/docs/instances/managing-instance-access#add_oslogin_keys

70. Question

You want to configure auto-healing for network load balancer for a group of Compute Engine instances that run in multiple zones using the fewest possible steps. You need to configure the recreation of VMs if they are unresponsive after 3 attempts of 10 seconds each. What should you do?

- Create a HTTP load balancer with a backend configuration that references an existing instance group. Define a balancing mode and set the maximum RPS to 10
- Create a HTTP load balancer with a backend configuration that references an existing instance group. Set the health check to healthy (HTTP)
- Create a managed instance group. Verify that the auto-scaling setting is on.
- Create a managed instance group. Set the Autohealing health check to healthy (HTTP)

SET-3

1. Question

You want to create a Google Cloud Storage regional bucket logs-archive in the Los Angeles region (us-west2). You want to use Coldline storage class to minimize costs and you want to retain files for 10 years. Which of the following commands should you run to create this bucket?

- gsutil mb -l us-west2 -s nearline --retention 10y gs://logs-archive
- gsutil mb -l los-angeles -s coldline --retention 10m gs://logs-archive
- gsutil mb -l us-west2 -s coldline --retention 10m gs://logs-archive
- gsutil mb -l us-west2 -s coldline --retention 10y gs://logs-archive

Correct

gsutil mb -l us-west2 -s nearline --retention 10y gs://logs-archive. is not right.

This command creates a bucket that uses nearline storage class whereas we want to use Coldline storage class.

Ref: [https://cloud.google.com/storage/docs/gsutil/commands\(mb](https://cloud.google.com/storage/docs/gsutil/commands(mb)

gsutil mb -l los-angeles -s coldline --retention 10m gs://logs-archive. is not right.

This command uses los-angeles as the location but los-angeles is not a supported region name. The region name for Los Angeles is us-west-2.

Ref: <https://cloud.google.com/storage/docs/locations>

gsutil mb -l us-west2 -s coldline --retention 10m gs://logs-archive. is not right.

This command creates a bucket with retention set to 10 months whereas we want to retain the objects for 10 years.

Ref: [https://cloud.google.com/storage/docs/gsutil/commands\(mb](https://cloud.google.com/storage/docs/gsutil/commands(mb)

gsutil mb -l us-west2 -s coldline --retention 10y gs://logs-archive. is the right answer.

This command correctly creates a bucket in Los Angeles, uses Coldline storage class and retains objects for 10 years.

Ref: [https://cloud.google.com/storage/docs/gsutil/commands\(mb](https://cloud.google.com/storage/docs/gsutil/commands(mb)

2. Question

You want to create a new role and grant it to the SME team. The new role should provide your SME team BigQuery Job User and Cloud Bigtable User roles on all projects in the organization. You want to minimize operational overhead. You want to follow Google recommended practices. How should you create the new role?

- Execute command gcloud iam combineroles --global to combine the 2 roles into a new custom role and grant them globally to SME team group.
- In GCP Console under IAM Roles, select both roles and combine them into a new custom role. Grant the role to the SME team group at the organization level.
- In GCP Console under IAM Roles, select both roles and combine them into a new custom role. Grant the role to the SME team group at project. Use gcloud iam promote-role to promote the role to all other projects and grant the role in each project to the SME team group.
- In GCP Console under IAM Roles, select both roles and combine them into a new custom role. Grant the role to the SME team group at project. Repeat this step for each project.

Unattempted

We want to create a new role and grant it to a team. Since you want to minimize operational overhead, we need to grant it to a group – so that new users who join the team just need to be added to the group and they inherit all the permissions. Also, this team needs to have the role for all projects in the organization. And since we want to minimize the operational overhead, we need to grant it at the organization level so that all current projects, as well as future projects, have the role granted to them.

In GCP Console under IAM Roles, select both roles and combine them into a new custom role. Grant the role to the SME team group at project. Repeat this step for each project. is not right.
Repeating the step for all projects is a manual, error-prone and time-consuming task. Also, if any projects were to be created in the future, we have to repeat the same process again. This increases operational overhead.

In GCP Console under IAM Roles, select both roles and combine them into a new custom role. Grant the role to the SME team group at project. Use gcloud iam promote-role to promote the role to all other projects and grant the role in each project to the SME team group. is not right.
Repeating the step for all projects is a manual, error-prone and time-consuming task. Also, if any projects were to be created in the future, we have to repeat the same process again. This increases operational overhead.

Execute command gcloud iam combine-roles --global to combine the 2 roles into a new custom role and grant them globally to all. is not right.
There are several issues with this. gcloud iam command doesn't support the action combine-roles.
Secondly, we don't want to grant the roles globally. We want to grant them to the SME team and no one else.

In GCP Console under IAM Roles, select both roles and combine them into a new custom role. Grant the role to the SME team group at the organization level. is the right answer.
This correctly creates the role and assigns the role to the group at the organization. When any new users join the team, the only additional task is to add them to the group. Also, when a new project is created under the organization, no additional human intervention is needed. Since the role is granted at the organization level, it automatically is granted to all the current and future projects belonging to the organization.

3. 3. Question

You want to deploy a python application to an autoscaled managed instance group on Compute Engine. You want to use GCP deployment manager to do this. What is the fastest way to get the application onto the instances without introducing undue complexity?

- Include a startup script to bootstrap the python application when creating an instance template by running gcloud compute instance-templates create app-template --metadata-from-file startup-script-url=/scripts/install_app.sh
- Once the instance starts up, connect over SSH and install the application.
- Include a startup script to bootstrap the python application when creating an instance template by running gcloud compute instance-templates create app-template --metadata-from-file startup-script=/scripts/install_app.sh
- Include a startup script to bootstrap the python application when creating an instance template by running gcloud compute instance-templates create app-template --startup-script=/scripts/install_app.sh

Unattempted

Include a startup script to bootstrap the python application when creating instance template by running gcloud compute instance-templates create app-template --startup-script=/scripts/install_app.sh. is not right.

gcloud compute instance-templates create command does not accept a flag called `--startup-script`. While creating compute engine images, the startup script can be provided through a special metadata key called `startup-script` which specifies a script that will be executed by the instances once they start running. For convenience, `--metadata-from-file` can be used to pull the value from a file.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instance-templates/create>

Include a startup script to bootstrap the python application when creating instance template by running `gcloud compute instance-templates create app-template --metadata-from-file startup-script-url=/scripts/install_app.sh`. is not right.

`startup-script-url` is to be used when contents of the script need to be pulled from a publicly-accessible location on the web. But in this scenario, we are passing the location of the script on the filesystem which doesn't work and the command errors out.

```
$ gcloud compute instance-templates create app-template --metadata-from-file startup-script-url=/scripts/install_app.sh
```

```
ERROR: (gcloud.compute.instance-templates.create) Unable to read file [/scripts/install_app.sh]: [Errno 2] No such file or directory: '/scripts/install_app.sh'
```

Once the instance starts up, connect over SSH and install the application. is not right.

The managed instances group has auto-scaling enabled. If we are to connect over SSH and install the application, we have to repeat this task on all current instances and on future instances the autoscaler adds to the group. This process is manual, error-prone, time consuming and should be avoided.

Include a startup script to bootstrap the python application when creating instance template by running `gcloud compute instance-templates create app-template --metadata-from-file startup-script=/scripts/install_app.sh`. is the right answer.

This command correctly provides the startup script using the flag `metadata-from-file` and providing a valid `startup-script` value. When creating compute engine images, the startup script can be provided through a special metadata key called `startup-script` which specifies a script that will be executed by the instances once they start running. For convenience, `--metadata-from-file` can be used to pull the value from a file.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instance-templates/create>

4. 4. Question

You want to deploy an application on Cloud Run that processes messages from a Cloud Pub/Sub topic. You want to follow Google-recommended practices. What should you do?

- 1. Create a Cloud Function that uses a Cloud Pub/Sub trigger on that topic. 2. Call your application on Cloud Run from the Cloud Function for every message.
- 1. Grant the Pub/Sub Subscriber role to the service account used by Cloud Run. 2. Create a Cloud Pub/Sub subscription for that topic. 3. Make your application pull messages from that subscription.
- 1. Create a service account. 2. Give the Cloud Run Invoker role to that service account for your Cloud Run application. 3. Create a Cloud Pub/Sub subscription that uses that service account and uses your Cloud Run application as the push endpoint.
- 1. Deploy your application on Cloud Run on GKE with the connectivity set to Internal. 2. Create a Cloud Pub/Sub subscription for that topic. 3. In the same Google Kubernetes Engine cluster as your application, deploy a container that takes the messages and sends them to your application.

Unattempted

1. Create a Cloud Function that uses a Cloud Pub/Sub trigger on that topic.

2. Call your application on Cloud Run from the Cloud Function for every message. is not right.

Both Cloud functions and Cloud Run are serverless offerings from GCP and they are both capable of integrating with Cloud Pub/Sub. It is pointless to invoking Cloud Function from Cloud Run.

- Grant the Pub/Sub Subscriber role to the service account used by Cloud Run.
 - Create a Cloud Pub/Sub subscription for that topic.
 - Make your application pull messages from that subscription. is not right.
- You need to provide Cloud Run Invoker role to that service account for your Cloud Run application.
- Ref: <https://cloud.google.com/run/docs/tutorials/pubsub>
- Deploy your application on Cloud Run on GKE with the connectivity set to Internal.
 - Create a Cloud Pub/Sub subscription for that topic.
 - In the same Google Kubernetes Engine cluster as your application, deploy a container that takes the messages and sends them to your application. is not right.
- Like above, you need cloud Run Invoker role on the service account.
- Ref: <https://cloud.google.com/run/docs/tutorials/pubsub>
- Also, our question states the application on Cloud Run processes messages from a Cloud Pub/Sub topic; whereas in this option, we are utilizing a separate container to process messages from the topic. So this doesn't satisfy our requirements.
- Create a service account.
 - Give the Cloud Run Invoker role to that service account for your Cloud Run application.
 - Create a Cloud Pub/Sub subscription that uses that service account and uses your Cloud Run application as the push endpoint. is the right answer.
- This exact process is described in
- <https://cloud.google.com/run/docs/tutorials/pubsub>
- You create a service account.
- ```
gcloud iam service-accounts create cloud-run-pubsub-invoker \
--display-name "Cloud Run Pub/Sub Invoker"
```
- You then give the invoker service account permission to invoke your service:
- ```
gcloud run services add-iam-policy-binding pubsub-tutorial \
--member=serviceAccount:cloud-run-pubsub-invoker@PROJECT_ID.iam.gserviceaccount.com \
--role=roles/run.invoker
```
- And finally, you create a Pub/Sub subscription with the service account:
- ```
gcloud pubsub subscriptions create myRunSubscription --topic myRunTopic \
--push-endpoint=SERVICE-URL/ \
--push-auth-service-account=cloud-run-pubsub-invoker@PROJECT_ID.iam.gserviceaccount
```

## 5. 5. Question

You want to ensure the boot disk of a preemptible instance is persisted for re-use. How should you provision the gcloud compute instance to ensure your requirement is met.

- `gcloud compute instances create [INSTANCE_NAME] --preemptible --no-boot-disk-auto-delete`
- `gcloud compute instances create [INSTANCE_NAME] --preemptible --boot-disk-auto-delete=no`
- `gcloud compute instances create [INSTANCE_NAME] --no-auto-delete`
- `gcloud compute instances create [INSTANCE_NAME] --preemptible. The flag --boot-disk-auto-delete is disabled by default.`

Unattempted

`gcloud compute instances create [INSTANCE_NAME] --preemptible --boot-disk-auto-delete=no.` is not right.  
`gcloud compute instances create` doesn't provide a parameter called boot-disk-auto-delete. It does have a flag by the same name. `--boot-disk-auto-delete` is enabled by default. It enables automatic deletion of boot disks when the instances are deleted. Use `--no-boot-disk-auto-delete` to disable.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/create>

`gcloud compute instances create [INSTANCE_NAME] --preemptible. --boot-disk-auto-delete` flag is disabled by default. is not right.  
`--boot-disk-auto-delete` is enabled by default. It enables automatic deletion of boot disks when the instances are deleted. Use `--no-boot-disk-auto-delete` to disable.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/create>

gcloud compute instances create [INSTANCE\_NAME] --no-auto-delete. is not right.  
gcloud compute instances create doesn't provide a flag called no-auto-delete  
Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/create>  
gcloud compute instances create [INSTANCE\_NAME] --preemptible --no-boot-disk-auto-delete. is the right answer.  
Use --no-boot-disk-auto-delete to disable automatic deletion of boot disks when the instances are deleted. --boot-disk-auto-delete flag is enabled by default. It enables automatic deletion of boot disks when the instances are deleted. In order to prevent automatic deletion, we have to specify --no-boot-disk-auto-delete flag.  
Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/create>

## 6. 6. Question

You want to find a list of regions and the prebuilt images offered by Google Compute Engine. Which commands should you execute to retrieve this information?

- gcloud compute regions list gcloud images list
- gcloud compute regions list gcloud compute images list
- gcloud regions list gcloud images list
- gcloud regions list gcloud compute images list

Unattempted

gcloud regions list.  
gcloud images list. is not right.

The correct command to list compute regions is gcloud compute regions list.  
Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/regions/list>

The correct command to list compute images is gcloud compute images list.  
Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/images/list>

gcloud compute regions list  
gcloud images list. is not right.

The correct command to list compute images is gcloud compute images list.  
Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/images/list>

gcloud regions list

gcloud compute images list. is not right.

The correct command to list compute regions is gcloud compute regions list.  
Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/regions/list>

gcloud compute regions list

gcloud compute images list. is the right answer.

Both the commands correctly retrieve images and regions offered by Google Compute Engine

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/regions/list>

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/images/list>

## 7. 7. Question

You want to find out when users were added to Cloud Spanner Identity Access Management (IAM) roles on your Google Cloud Platform (GCP) project. What should you do in the GCP Console?

- Open the Cloud Spanner console to review configurations.
- Open the IAM & admin console to review IAM policies for Cloud Spanner roles.
- Go to the Stackdriver Monitoring console and review information for Cloud Spanner.
- Go to the Stackdriver Logging console, review admin activity logs, and filter them for Cloud Spanner IAM roles.

Unattempted

Go to the Stackdriver Monitoring console and review information for Cloud Spanner. is not right.  
Monitoring collects metrics, events, and metadata from Google Cloud and lets you generate insights via dashboards, charts, and alerts. It can't provide information on when a role has been granted to a user.

Ref: <https://cloud.google.com/monitoring/docs>

Open the IAM & admin console to review IAM policies for Cloud Spanner roles. is not right.  
You can't find the role bindings and the timestamps in the policies.

<https://cloud.google.com/iam/docs/overview>

Open the Cloud Spanner console to review configurations. is not right.  
You manage cloud spanner instances in the console but you can't check when a role has been granted to a user.

Ref: <https://cloud.google.com/spanner/docs/quickstart-console>

Go to the Stackdriver Logging console, review admin activity logs, and filter them for Cloud Spanner IAM roles. is the right answer.

Admin Activity audit logs contain log entries for API calls or other administrative actions that modify the configuration or metadata of resources. For example, these logs record when users create VM instances or change Cloud Identity and Access Management permissions. Admin Activity audit logs are always written; you can't configure or disable them. There is no charge for your Admin Activity audit logs.

Ref: <https://cloud.google.com/logging/docs/audit#admin-activity>

See below a screenshot from GCP console showing this in action.

Among other things, the payload contains

```
{
action: "ADD"
role: "roles/spanner.admin"
member: "user:testuser@gmail.com"
}
```

## 8. Question

You want to ingest and analyze large volumes of stream data from sensors in real-time, matching the high speeds of IoT data to track normal and abnormal behavior. You want to run it through a data processing pipeline and store the results. Finally, you want to enable customers to build dashboards and drive analytics on their data in real-time. What services should you use for this task?

- Cloud Pub/Sub, Cloud Dataflow, BigQuery
- Cloud Pub/Sub, Cloud Dataflow, Cloud Dataprep
- Stackdriver, Cloud Dataflow, BigQuery
- Cloud Pub/Sub, Cloud Dataflow, Cloud Dataproc

Unattempted

You want to ingest large volumes of streaming data at high speeds. So you need to use Cloud Pub/Sub. Cloud Pub/Sub provides a simple and reliable staging location for your event data on its journey towards processing, storage, and analysis. Cloud Pub/Sub is serverless and you can ingest events at any scale.

Ref: <https://cloud.google.com/pubsub>

Next, you want to analyze this data. Cloud Dataflow is a fully managed streaming analytics service that minimizes latency, processing time, and cost through auto scaling and batch processing. Dataflow enables fast, simplified streaming data pipeline development with lower data latency.

Ref: <https://cloud.google.com/dataflow>

Next, you want to store these results. BigQuery is an ideal place to store these results as BigQuery supports the querying of streaming data in real-time. This assists in real-time predictive analytics.

Ref: <https://cloud.google.com/bigquery>

Therefore the correct answer is Cloud Pub/Sub, Cloud Dataflow, BigQuery

Here's more information from Google docs about the Stream analytics use case. Google recommends we use Dataflow along with Pub/Sub and BigQuery.

<https://cloud.google.com/dataflow#section-6>

Google's stream analytics makes data more organized, useful, and accessible from the instant it's generated. Built on Dataflow along with Pub/Sub and BigQuery, our streaming solution provisions the resources you need to ingest, process, and analyze fluctuating volumes of real-time data for real-time business insights. This abstracted provisioning reduces complexity and makes stream analytics accessible to both data analysts and data engineers.

and

<https://cloud.google.com/solutions/stream-analytics>

Ingest, process, and analyze event streams in real time. Stream analytics from Google Cloud makes data more organized, useful, and accessible from the instant it's generated. Built on the auto scaling infrastructure of Pub/Sub, Dataflow, and BigQuery, our streaming solution provisions the resources you need to ingest, process, and analyze fluctuating volumes of real-time data for real-time business insights.

## 9. Question

You want to list all the compute instances in zones us-central1-b and europe-west1-d. Which of the commands below should you run to retrieve this information?

- gcloud compute instances list --filter="zone:( us-central1-b )" and gcloud compute instances list --filter="zone:( europe-west1-d )" and combine the results.
- gcloud compute instances get --filter="zone:( us-central1-b )" and gcloud compute instances list --filter="zone:( europe-west1-d )" and combine the results.
- gcloud compute instances get --filter="zone:( us-central1-b europe-west1-d )"
- gcloud compute instances list --filter="zone:( us-central1-b europe-west1-d )"

Unattempted

gcloud compute instances get --filter="zone:( us-central1-b europe-west1-d )". is not right.  
gcloud compute instances command does not support get action.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances>

gcloud compute instances get --filter="zone:( us-central1-b )" and gcloud compute instances list --filter="zone:( europe-west1-d )" and combine the results. is not right.

gcloud compute instances command does not support get action.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances>

gcloud compute instances list --filter="zone:( us-central1-b )" and gcloud compute instances list --filter="zone:( europe-west1-d )" and combine the results. is not right.

The first command retrieves compute instances from us-central1-b and the second command retrieves compute instances from europe-west1-d. The output from the two statements can be combined to create a full list of instances from us-central1-b and europe-west1-d, however, this is not efficient as it is a manual activity. Moreover, gcloud already provides the ability to list and filter on multiple zones in a single command.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/list>

gcloud compute instances list --filter="zone:( us-central1-b europe-west1-d )". is the right answer.  
gcloud compute instances list -- lists Google Compute Engine instances. The output includes internal as well as external IP addresses. The filter expression --filter="zone:( us-central1-b europe-west1-d )" is used to filter instances from zones us-central1-b and europe-west1-d.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/list>

Here's a sample output of the command.

\$gcloud compute instances list

| NAME                                     | ZONE          | MACHINE_TYPE  | PREEMPTIBLE | INTERNAL_IP    | EXTERNAL_IP | STATUS  |
|------------------------------------------|---------------|---------------|-------------|----------------|-------------|---------|
| gke-cluster-1-default-pool-8c599c87-16g9 | us-central1-a | n1-standard-1 | 10.128.0.8  | 35.184.212.227 |             | RUNNING |
| gke-cluster-1-default-pool-8c599c87-36xh | us-central1-b | n1-standard-1 | 10.129.0.2  | 34.68.254.220  |             | RUNNING |
| gke-cluster-1-default-pool-8c599c87-lprq | us-central1-c | n1-standard-1 | 10.130.0.13 | 35.224.96.151  |             | RUNNING |

```
$gcloud compute instances list --filter="zone:(us-central1-b europe-west1-d)"
NAME ZONE MACHINE_TYPE PREEMPTIBLE INTERNAL_IP EXTERNAL_IP STATUS
gke-cluster-1-default-pool-8c599c87-36xhus-central1-bn1-standard-110.129.0.234.68.254.220RUNNING
```

## 10. 10. Question

You want to list all the internal and external IP addresses of all compute instances. Which of the commands below should you run to retrieve this information?

- `gcloud compute instances list.`
- `gcloud compute networks list-ip.`
- `gcloud compute networks list.`
- `gcloud compute instances list-ip.`

Unattempted

`gcloud compute instances list-ip.` is not right.

“`gcloud compute instances`” doesn’t support the action `list-ip`.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/list>

`gcloud compute networks list-ip.` is not right.

“`gcloud compute networks`” doesn’t support the action `list-ip`.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/networks/list>

`gcloud compute networks list.` is not right.

“`gcloud compute networks list`” doesn’t list the IP addresses. It is used for listing Google Compute Engine networks (i.e. VPCs)

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/networks/list>

Here’s a sample output of the command.

```
$ gcloud compute networks list
```

```
NAME SUBNET_MODE BGP_ROUTING_MODE IPV4_RANGE GATEWAY_IPV4
```

```
default AUTO REGIONAL
```

```
test-vpc CUSTOM REGIONAL
```

`gcloud compute instances list.` is the right answer

`gcloud compute instances list` – lists Google Compute Engine instances. The output includes internal as well as external IP addresses.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/list>

Here’s a sample output of the command.

```
$ gcloud compute instances list
```

```
NAME ZONE MACHINE_TYPE PREEMPTIBLE INTERNAL_IP EXTERNAL_IP STATUS
```

```
gke-cluster-1-default-pool-8c599c87-16g9 us-central1-a n1-standard-1 10.128.0.8 35.184.212.227
```

```
RUNNING
```

```
gke-cluster-1-default-pool-8c599c87-36xh us-central1-a n1-standard-1 10.128.0.6 34.68.254.220 RUNNING
```

```
gke-cluster-1-default-pool-8c599c87-lprq us-central1-a n1-standard-1 10.128.0.7 35.224.96.151 RUNNING
```

## 11. 11. Question

You want to migrate an application from Google App Engine Standard to Google App Engine Flex. Your application is currently serving live traffic and you want to ensure everything is working in Google App Engine Flex before migrating all traffic. You want to minimize effort and ensure the availability of service. What should you do?

- 1. Set env: flex in app.yaml 2. `gcloud app deploy --version=[NEW_VERSION]` 3. Validate [NEW\_VERSION] in App Engine Flex 4. `gcloud app versions migrate [NEW_VERSION]`
- 1. Set env: app-engine-flex in app.yaml 2. `gcloud app deploy --no-promote --version=[NEW_VERSION]` 3. Validate [NEW\_VERSION] in App Engine Flex 4. `gcloud app versions start [NEW_VERSION]`

- 1. Set env: app-engine-flex in app.yaml 2. gcloud app deploy --version=[NEW\_VERSION] 3. Validate [NEW\_VERSION] in App Engine Flex 4. gcloud app versions start [NEW\_VERSION]
- 1. Set env: flex in app.yaml 2. gcloud app deploy --no-promote --version=[NEW\_VERSION] 3. Validate [NEW\_VERSION] in App Engine Flex 4. gcloud app versions migrate [NEW\_VERSION]

Unattempted

1. Set env: flex in app.yaml
2. gcloud app deploy --version=[NEW\_VERSION]
3. Validate [NEW\_VERSION] in App Engine Flex
4. gcloud app versions migrate [NEW\_VERSION]. is not right.  
Executing gcloud app deploy --version=[NEW\_VERSION] without --no-promote would deploy the new version and immediately promote it to serve traffic. We don't want this version to receive traffic as we would like to validate the version first before sending it traffic.  
Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/migrate>
1. Set env: app-engine-flex in app.yaml
2. gcloud app deploy --version=[NEW\_VERSION]
3. Validate [NEW\_VERSION] in App Engine Flex
4. gcloud app versions start [NEW\_VERSION] is not right.  
env: app-engine-flex is an invalid setting. The correct syntax for using the flex engine is env: flex. Also, Executing gcloud app deploy --version=[NEW\_VERSION] without --no-promote would deploy the new version and immediately promote it to serve traffic. We don't want this version to receive traffic as we would like to validate the version first before sending it traffic.  
Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/migrate>
1. Set env: app-engine-flex in app.yaml
2. gcloud app deploy --no-promote --version=[NEW\_VERSION]
3. Validate [NEW\_VERSION] in App Engine Flex
4. gcloud app versions start [NEW\_VERSION] is not right.  
env: app-engine-flex is an invalid setting. The correct syntax for using the flex engine is env: flex.  
Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/migrate>
1. Set env: flex in app.yaml
2. gcloud app deploy --no-promote --version=[NEW\_VERSION]
3. Validate [NEW\_VERSION] in App Engine Flex
4. gcloud app versions migrate [NEW\_VERSION] is the right answer.  
These commands together achieve the end goal while satisfying our requirements. Setting env: flex in app.yaml and executing gcloud app deploy --no-promote --version=[NEW\_VERSION] results in a new version deployed to flex engine. but the new version is not configured to serve traffic. We take the opportunity to review this version before migrating it to serve live traffic by running gcloud app versions migrate [NEW\_VERSION]  
Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/migrate>  
Ref: <https://cloud.google.com/sdk/gcloud/reference/app/deploy>

## 12. 12. Question

You want to persist logs for 10 years to comply with regulatory requirements. You want to follow Google recommended practices. Which Google Cloud Storage class should you use?

- Archive storage class
- Nearline storage class
- Coldline storage class
- Standard storage class

Unattempted

In April 2019, Google introduced a new storage class “Archive storage class” is the lowest-cost, highly durable storage service for data archiving, online backup, and disaster recovery. Google previously recommended you use Coldline storage class but the recommendation has since been updated to “Coldline Storage is ideal for data you plan to read or modify at most once a quarter. Note, however, that for data being kept entirely for backup or archiving purposes, Archive Storage is more cost-effective, as it offers the lowest storage costs.”

Ref: <https://cloud.google.com/storage/docs/storage-classes#archive>

Ref: <https://cloud.google.com/storage/docs/storage-classes#coldline>

So the correct answer is Archive storage class.

### 13. 13. Question

You want to reduce storage costs for infrequently accessed data. The data will still be accessed approximately once a month and data older than 2 years is no longer needed. What should you do to reduce storage costs? (Select 2)

- Store infrequently accessed data in a Nearline bucket.
- Set an Object Lifecycle Management policy to delete data older than 2 years.
- Set an Object Lifecycle Management policy to change the storage class to Coldline for data older than 2 years.
- Store infrequently accessed data in a Multi-Regional bucket.
- Set an Object Lifecycle Management policy to change the storage class to Archive for data older than 2 years.

Unattempted

Set an Object Lifecycle Management policy to change the storage class to Coldline for data older than 2 years. is not right.

Data older than 2 years is not needed so there is no point in transitioning the data to Coldline. The data needs to be deleted.

Set an Object Lifecycle Management policy to change the storage class to Archive for data older than 2 years. is not right.

Data older than 2 years is not needed so there is no point in transitioning the data to Archive. The data needs to be deleted.

Store infrequently accessed data in a Multi-Regional bucket. is not right.

While infrequently accessed data can be stored in Multi-Regional bucket, there are several other storage classes offered by Google Cloud Storage that are primarily aimed at storing infrequently accessed data and cost less. Multi-Region buckets are primarily used for achieving geo-redundancy.

Ref: <https://cloud.google.com/storage/docs/locations>

Set an Object Lifecycle Management policy to delete data older than 2 years. is the right answer.

Since you don't need data older than 2 years, deleting such data is the right approach. You can set a lifecycle policy to automatically delete objects older than 2 years. The policy is valid on current as well as future objects and doesn't need any human intervention.

Ref: <https://cloud.google.com/storage/docs/lifecycle>

Store infrequently accessed data in a Nearline bucket. is the right answer.

Nearline Storage is a low-cost, highly durable storage service for storing infrequently accessed data.

Nearline Storage is ideal for data you plan to read or modify on average once per month or less.

Ref: <https://cloud.google.com/storage/docs/storage-classes#nearline>

### 14. 14. Question

You want to run a single caching HTTP reverse proxy on GCP for a latency-sensitive website. This specific reverse proxy consumes almost no CPU. You want to have a 30-GB in-memory cache and need an additional 2 GB of memory for the rest of the processes. You want to minimize costs. How should you run this reverse proxy?

- Create a Cloud Memorystore for Redis instance with 32-GB capacity.
  - Run it on Compute Engine and choose a custom instance type with 6 vCPUs and 32 GB of memory.
  - Package it in a container image, and run it on Kubernetes Engine, using n1-standard-32 instances as nodes.
  - Run it on Compute Engine, choose the instance type n1-standard-1, and add an SSD persistent disk of 32 GB.

Unattempted

#### Requirements

1. latency sensitive
2. 30 GB in-memory cache
3. 2 GB for rest of processes
4. Cost-effective

Run it on Compute Engine, choose the instance type n1-standard-1, and add an SSD persistent disk of 32 GB. is not right.

Fetching data from disk is slower compared to fetching from in-memory. Our requirements state we need 30GB in-memory cache for a latency-sensitive website and a compute engine with disk can't provide in-memory cache.

Run it on Compute Engine and choose a custom instance type with 6 vCPUs and 32 GB of memory. is not right.

While this option provides us with 32 GB of memory, a part of it used by the compute engine operating system as well as the reverse proxy process leaving us with less than 32GB which does not satisfy our requirements. In addition, the reverse proxy consumes almost no CPU so having 6vCPUs is a waste of resources and money.

Package it in a container image, and run it on Kubernetes Engine, using n1-standard-32 instances as nodes. is not right.

Without going into details of the feasibility of this option, let's assume for now that this option is possible. But this option is quite expensive. At the time of writing, just the compute cost for a n1-standard-32 instance is \$1.5200 per hour in the Iowa region.

Ref: <https://cloud.google.com/compute/all-pricing>

In comparison, the cost of GCP Cloud Memorystore which is \$0.023 per GB-hr which is \$0.736 for 32GB per hour. Ref: <https://cloud.google.com/memorystore>

Create a Cloud Memorystore for Redis instance with 32-GB capacity. is the right answer.

This is the only option that fits the requirements. Cloud Memorystore is a fully managed in-memory data store service for Redis built on scalable, secure, and highly available infrastructure managed by Google.

Use Memorystore to build application caches that provide sub-millisecond data access.

Ref: <https://cloud.google.com/memorystore>

Memorystore for Redis instance pricing is charged per GB-hour and you can scale as needed. You can also specify eviction (maxmemory) policies to restrict the rest of processes to 2GB or the reverse proxy to 30GB or both; you can select a suitable maxmemory policy to handle scenarios when memory is full.

Ref: [https://cloud.google.com/memorystore/docs/reference/redis-configs#maxmemory\\_policies](https://cloud.google.com/memorystore/docs/reference/redis-configs#maxmemory_policies)

15. 15. Question

You want to select and configure a cost-effective solution for relational data on Google Cloud Platform. You are working with a small set of operational data in one geographic location. You need to support point-in-time recovery. What should you do?

- Select Cloud Spanner. Set up your instance with 2 nodes.
- Select Cloud SQL (MySQL). Verify that the enable binary logging option is selected.
- Select Cloud SQL (MySQL). Select the create failover replicas option.
- Select Cloud Spanner. Set up your instance as multi-regional.

Unattempted

#### Requirements

1. Cost effective
2. Relational Data
3. Small set of data
4. One location
5. Point in time recovery

Select Cloud Spanner. Set up your instance with 2 nodes. is not right.

Cloud spanner is a massively scalable, fully managed, relational database service for regional and global application data. Cloud spanner is expensive compared to Cloud SQL. We have a small set of data and we want to be cost-effective, so Cloud Spanner doesn't fit these requirements. Furthermore, Cloud Spanner does not offer a "Point in time" recovery feature.

Ref: <https://cloud.google.com/spanner>

Select Cloud Spanner. Set up your instance as multi-regional. is not right.

Cloud spanner is a massively scalable, fully managed, relational database service for regional and global application data. Cloud spanner is expensive compared to Cloud SQL. We don't have a requirement for more than "one geographic location" and we also have a small set of data and we want to be cost-effective, so Cloud Spanner doesn't fit these requirements. Furthermore, Cloud Spanner does not offer a "Point in time" recovery feature.

Ref: <https://cloud.google.com/spanner>

Select Cloud SQL (MySQL). Select the create failover replicas option. is not right.

Cloud SQL can easily handle small sets of relational data and is cost-effective compared to Cloud Spanner. But This option does not enable point in time recovery so our requirement to support point-in-time recovery is not met.

Ref: <https://cloud.google.com/sql/docs/mysql>

Select Cloud SQL (MySQL). Verify that the enable binary logging option is selected. is the right answer  
Cloud SQL can easily handle small sets of relational data and is cost-effective compared to Cloud Spanner. And by enabling binary logging, we can enable point-in-time recovery which fits our requirement.

You must enable binary logging to use point-in-time recovery. Point-in-time recovery helps you recover an instance to a specific point in time. For example, if an error causes a loss of data, you can recover a database to its state before the error occurred.

Ref: <https://cloud.google.com/sql/docs/mysql/backup-recovery/backups#tips-pitr>

#### 16. Question

You want to select and configure a solution for storing and archiving data on Google Cloud Platform. You need to support compliance objectives for data from one geographic location. This data is archived after 30 days and needs to be accessed annually. What should you do?

- Select Multi-Regional Storage. Add a bucket lifecycle rule that archives data after 30 days to Coldline Storage.
- Select Multi-Regional Storage. Add a bucket lifecycle rule that archives data after 30 days to Nearline Storage.

- Select Regional Storage. Add a bucket lifecycle rule that archives data after 30 days to Nearline Storage.
- Select Regional Storage. Add a bucket lifecycle rule that archives data after 30 days to Coldline Storage.

Unattempted

Our requirements are one region, archival after 30 days and data to be accessed annually.

Select Multi-Regional Storage. Add a bucket lifecycle rule that archives data after 30 days to Coldline Storage. is not right.

When used in a multi-region, Standard Storage is appropriate for storing data that is accessed around the world, such as serving website content, streaming videos, executing interactive workloads, or serving data supporting mobile and gaming applications. This is against our requirement of one region, moreover, this is expensive compared to standard Regional storage.

Ref: <https://cloud.google.com/storage/docs/storage-classes#standard>

Select Multi-Regional Storage. Add a bucket lifecycle rule that archives data after 30 days to Nearline Storage. is not right.

When used in a multi-region, Standard Storage is appropriate for storing data that is accessed around the world, such as serving website content, streaming videos, executing interactive workloads, or serving data supporting mobile and gaming applications. This is against our requirement of one region, moreover, this is expensive compared to standard Regional storage.

Ref: <https://cloud.google.com/storage/docs/storage-classes#standard>

Select Regional Storage. Add a bucket lifecycle rule that archives data after 30 days to Nearline Storage. is not right.

While selecting Regional Storage is the right choice, archiving to Nearline is not the most optimal. We have a requirement to access data annually whereas Nearline Storage is ideal for data you plan to read or modify on average once per month or less. Nearline storage is more expensive than Coldline Storage which is more suitable for our requirements.

<https://cloud.google.com/storage/docs/storage-classes#nearline>

Select Regional Storage. Add a bucket lifecycle rule that archives data after 30 days to Coldline Storage. is the right answer.

Regional Storage is the right fit for our requirements (one geographic region) and archiving to Coldline storage is the most cost-efficient solution. Coldline Storage is a very-low-cost, highly durable storage service for storing infrequently accessed data. Coldline Storage is ideal for data you plan to read or modify at most once a quarter. Since we have a requirement to access data once a quarter and want to go with the most cost-efficient option, we should select Coldline Storage.

Ref: <https://cloud.google.com/storage/docs/storage-classes#coldline>

## 17. Question

You want to send and consume Cloud Pub/Sub messages from your App Engine application. The Cloud Pub/Sub API is currently disabled. You will use a service account to authenticate your application to the API. You want to make sure your application can use Cloud Pub/Sub. What should you do?

- Rely on the automatic enablement of the Cloud Pub/Sub API when the Service Account accesses it.
- Enable the Cloud Pub/Sub API in the API Library on the GCP Console.
- Grant the App Engine default service account the role of Cloud Pub/Sub Admin. Have your application enable the API on the first connection to Cloud Pub/Sub.
- Use the Deployment Manager to deploy your application. Rely on the automatic enablement of all APIs used by the application being deployed.

Unattempted

#### Requirements

1. We need to enable Cloud Pub/Sub API
2. Get our application to use the service account.

Grant the App Engine default service account the role of Cloud Pub/Sub Admin. Have your application enable the API on the first connection to Cloud Pub/Sub. is not right.

APIs are not automatically enabled on the first connection to the service (Cloud Pub/Sub in this scenario).

APIs can be enabled through Google Cloud Console, gcloud command-line and REST API.

See <https://cloud.google.com/service-usage/docs/enable-disable> for more information.

Rely on the automatic enablement of the Cloud Pub/Sub API when the Service Account accesses it. is not right.

There is no such thing as automatic enablement of the APIs when the service (Cloud Pub/Sub in this scenario) is accessed. APIs can be enabled through Google Cloud Console, gcloud command-line and REST API. See <https://cloud.google.com/service-usage/docs/enable-disable> for more information.

Use the Deployment Manager to deploy your application. Rely on the automatic enablement of all APIs used by the application being deployed. is not right.

There is no such thing as automatic enablement of the APIs (Cloud Pub/Sub in this scenario) is accessed. APIs can be enabled through Google Cloud Console, gcloud command-line and REST API.

See <https://cloud.google.com/service-usage/docs/enable-disable> for more information.

Enable the Cloud Pub/Sub API in the API Library on the GCP Console. is the right answer.

For most operational use cases, the simplest way to enable and disable services is to use the Google Cloud Console. you need to create scripts, you can also use the gcloud command-line interface. If you need to program against the Service Usage API, we recommend that you use one of our provided client libraries  
Ref: <https://cloud.google.com/service-usage/docs/enable-disable>

Secondly, after you create an App Engine application, the App Engine default service account is created and used as the identity of the App Engine service. The App Engine default service account is associated with your Cloud project and executes tasks on behalf of your apps running in App Engine. By default, the App Engine default service account has the Editor role in the project so this already has the permissions to push/pull/receive messages from Cloud Pub/Sub

#### 18. 18. Question

You want to serve files under the URL <https://www.my-new-gcp-ace-website.com/static/> from Cloud Storage. In addition, the URL <https://www.my-new-gcp-ace-website.com/app/> should be handled by a Compute Engine managed instance group (MIG). You want to follow Google recommended practices. How should you configure load balancing?

- 1. Deploy HAProxy in a second MIG and configure it to route /app/ to the first MIG and /static/ to your Cloud Storage bucket. 2. Create a network Load Balancer in front of the HAProxy MIG 3. Configure www.my-new-gcp-ace-website.com as an A record pointing to the address of the load balancer
- 1. Create a HTTPS Load Balancer in front of the MIG 2. In Cloud DNS in the my-new-gcp-ace-website.com zone, create a TXT record for \_app\_.\_routes\_.www.my-new-gcp-ace-website.com containing the address of the load balancer. 3. Create another TXT record for \_static\_.\_routes\_.www.my-new-gcp-ace-website.com containing the URL of your Cloud Storage bucket.
- 1. Configure www.my-new-gcp-ace-website.com as a CNAME pointing to storage.googleapis.com 2. Create a HTTPS Load Balancer in front of the MIG 3. IN the app folder of your Cloud Storage Bucket, add a file called redirect containing the address of the load balancer.
- 1. Create a HTTPS Load Balancer 2. Create a backend service associated with the MIG and route /app/ to the backend service 3. Create a backend bucket associated with your Cloud Storage Bucket, and route /static/ to the backend bucket 4. Configure www.my-new-gcp-ace-website.com as an A record pointing to the address of the load balancer.

Unattempted

Our requirement here is to serve content from two backends while following Google recommended practices.

Let's look at each of the options

1. Configure www.my-new-gcp-ace-website.com as a CNAME pointing to storage.googleapis.com
2. Create a HTTPS Load Balancer in front of the MIG
3. In the app folder of your Cloud Storage Bucket, add a file called redirect containing the address of the load balancer. is not right.

We can create a CNAME www.my-new-gcp-ace-website.com pointing to storage.googleapis.com, however, the cloud storage bucket does not support routing requests to a load balancer based on routing information in a file in the app folder. So this option doesn't work.

1. Deploy HAProxy in a second MIG and configure it to route /app/ to the first MIG and /static/ to your Cloud Storage bucket.
2. Create a network Load Balancer in front of the HAProxy MIG
3. Configure www.my-new-gcp-ace-website.com as an A record pointing to the address of the load balancer is not right.

This could possibly work, but we want to follow Google recommended practices and why deploy and manage HAProxy when there might be some other Google product that does exactly the same with minimal configuration (there is !!)?

1. Create a HTTPS Load Balancer in front of the MIG
2. In Cloud DNS in the example.com zone, create a TXT record for \_app\_.\_routes\_.www.my-new-gcp-ace-website.com containing the address of the load balancer.
3. Create another TXT record for \_static\_.\_routes\_.www.my-new-gcp-ace-website.com containing the URL of your Cloud Storage bucket. is not right.

TXT records are used to verify the domain and TXT records can also hold any arbitrary text but the DNS providers don't use the text in these TXT records for routing.

Ref: <https://cloud.google.com/dns/records>

Ref: <https://support.google.com/cloudidentity/answer/183895?hl=en>

1. Create a HTTPS Load Balancer
2. Create a backend service associated with the MIG and route /app/ to the backend service
3. Create a backend bucket associated with your Cloud Storage Bucket, and route /static/ to the backend bucket
4. Configure www.my-new-gcp-ace-website.com as an A record pointing to the address of the load balancer. is the right answer.

Since we need to send requests to multiple backends, Cloud DNS can't alone help us. We need Cloud HTTPS Load Balancer – it's URL maps (a fancy name for path-based routing) helps distribute traffic to backends based on the path information. Ref <https://cloud.google.com/load-balancing/docs/url-map>

Traffic received by Cloud HTTPS Load Balancer can be configured to send all requests on /app path to the MIG group; and requests on /static/ path to the bucket.

Ref Adding MIG as backend service- [https://cloud.google.com/load-balancing/docs/backend-service#backend\\_services\\_and\\_autoscaled\\_managed\\_instance\\_groups](https://cloud.google.com/load-balancing/docs/backend-service#backend_services_and_autoscaled_managed_instance_groups).

Ref Adding a backend bucket(s) – <https://cloud.google.com/load-balancing/docs/https/adding-backend-buckets-to-load-balancers>

The Load Balancer has a public IP address. But we want to instead access on www.my-new-gcp-ace-website.com, so we configure this as an A Record in our DNS provider. So this option is the right answer.  
Ref: <https://cloud.google.com/dns/records>.

## 19. Question

You want to use Google Cloud Storage to host a static website on <http://www.example.com> for your staff. You created a bucket example-static-website and uploaded index.html and css files to it. You turned on static website hosting on the bucket and set up a CNAME record on <http://www.example.com> to point to c.storage.googleapis.com. You access the static website by navigating to <http://www.example.com> in the browser but your index page is not displayed. What should you do?

- In example.com zone, delete the existing CNAME record and set up an A record instead to point to c.storage.googleapis.com.
- In example.com zone, modify the CNAME record to c.storage.googleapis.com/example-static-website.
- Reload the Cloud Storage static website server to load the objects.
- Delete the existing bucket, create a new bucket with the name [www.example.com](http://www.example.com) and upload the html/css files.

Unattempted

In example.com zone, modify the CNAME record to c.storage.googleapis.com/example-static-website. is not right.

CNAME records cannot contain paths. There is nothing wrong with the current CNAME record.

In example.com zone, delete the existing CNAME record and set up an A record instead to point to c.storage.googleapis.com. is not right.

A records cannot use hostnames. A records use IP Addresses.

Reload the Cloud Storage static website server to load the objects. is not right.

There is no such thing as a Cloud Storage static website server. All infrastructure that underpins the static websites is handled by Google Cloud Platform.

Delete the existing bucket, create a new bucket with the name [www.example.com](http://www.example.com) and upload the html/css files. is the right answer.

We need to create a bucket whose name matches the CNAME you created for your domain. For example, if you added a CNAME record pointing <http://www.example.com> to c.storage.googleapis.com., then create a bucket with the name “www.example.com”.A CNAME record is a type of DNS record. It directs traffic that requests a URL from your domain to the resources you want to serve, in this case, objects in your Cloud Storage buckets. For <http://www.example.com>, the CNAME record might contain the following information:

NAME TYPE DATA

<http://www.example.com> CNAME c.storage.googleapis.com.

Ref: <https://cloud.google.com/storage/docs/hosting-static-website>

## 20. 20. Question

You want to verify the IAM users and roles assigned within a GCP project named my-project. What should you do?

- Run gcloud iam service-accounts list. Review the output section.
- Navigate to the project and then to the IAM section in the GCP Console. Review the members and roles.
- Navigate to the project and then to the Roles section in the GCP Console. Review the roles and status.
- Run gcloud iam roles list. Review the output section.

Unattempted

Requirements – verify users (i.e. IAM members) and roles.

Run gcloud iam roles list. Review the output section. is not right.  
gcloud iam roles list lists the roles but does not list the users (i.e. IAM members)

Run gcloud iam service-accounts list. Review the output section. is not right.  
gcloud iam service-accounts list lists the service accounts which are users (i.e. IAM members) but it ignores other users that are not service accounts e.g. users in GSuite domain, or groups etc.

Navigate to the project and then to the Roles section in the GCP Console. Review the roles and status. is not right.

This allows us to review the roles but not users. See the screenshot below.

Navigate to the project and then to the IAM section in the GCP Console. Review the members and roles. is the right answer.

This is the only option that lets us view roles as well as users (members).

Ref: <https://cloud.google.com/iam/docs/overview>

See the screenshot below.

A member can be a Google Account (for end-users), a service account (for apps and virtual machines), a Google group, or a G Suite or Cloud Identity domain that can access a resource. The identity of a member is an email address associated with a user, service account, or Google group; or a domain name associated with G Suite or Cloud Identity domains

## 21. Question

You've deployed a microservice called myapp1 to a Google Kubernetes Engine cluster using the YAML file specified below:

Larger image

You need to refactor this configuration so that the database password is not stored in plain text. You want to follow Google-recommended practices. What should you do?

- Store the database password in a file inside a Kubernetes persistent volume, and use a persistent volume claim to mount the volume to the container.
- Store the database password inside a ConfigMap object. Modify the YAML file to populate the DB\_PASSWORD environment variable from the ConfigMap.
- Store the database password inside the Docker image of the container, not in the YAML file.
- Store the database password inside a Secret object. Modify the YAML file to populate the DB\_PASSWORD environment variable from the Secret.

Unattempted

Store the database password inside the Docker image of the container, not in the YAML file. is not right.  
Baking passwords into Docker images is a very bad idea. Anyone who spins up a container from this image has access to the password.

Store the database password inside a ConfigMap object. Modify the YAML file to populate the DB\_PASSWORD environment variable from the ConfigMap. is not right.

ConfigMaps are useful for storing and sharing non-sensitive, unencrypted configuration information. To use sensitive information in your clusters, you must use Secrets.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/configmap>

Store the database password in a file inside a Kubernetes persistent volume, and use a persistent volume claim to mount the volume to the container. is not right.

Persistent volumes should not be used for storing sensitive information. PersistentVolume resources are used to manage durable storage in a cluster and PersistentVolumeClaim is a request for and claim to a PersistentVolume resource.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/persistent-volumes>

Store the database password inside a Secret object. Modify the YAML file to populate the DB\_PASSWORD environment variable from the Secret. is the right answer.

In GKE, you can create a secret to hold the password; and then use the secret as an environment variable in the YAML file.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/secret>

You can create a secret using kubectl create secret generic passwords –from-literal myapp1\_db\_password=t0ugh2guess!

And you can then modify the YAML file to reference this secret as shown below.

## 22. 22. Question

Your company collects and stores CCTV footage videos in raw format in Google Cloud Storage. Within the first 30 days, the footage is processed regularly for detecting patterns such as threat/object/face detection and suspicious behavior detection. You want to minimize the cost of storing all the data in Google Cloud. How should you store the videos?

- Use Google Cloud Nearline Storage for the first 30 days, and use lifecycle rules to transition to Coldline Storage.
- Use Google Cloud Regional Storage for the first 30 days, and use lifecycle rules to transition to Coldline Storage
- Use Google Cloud Regional Storage for the first 30 days, and use lifecycle rules to transition to Nearline Storage.
- Use Google Cloud Regional Storage for the first 30 days, and then move videos to Google Persistent Disk.

Unattempted

Footage is processed regularly within the first 30 days and is rarely used after that. So we need to store the videos for the first 30 days in a storage class that supports economic retrieval (for processing) or at no cost, and then transition the videos to a cheaper storage after 30 days.

Use Google Cloud Regional Storage for the first 30 days, and use lifecycle rules to transition to Nearline Storage. is not right.

Transitioning the data to Nearline Storage is a good idea as Nearline Storage costs less than standard storage, is highly durable for storing infrequently accessed data and a better choice than Standard Storage in scenarios where slightly lower availability is an acceptable trade-off for lower at-rest storage costs.

Ref: <https://cloud.google.com/storage/docs/storage-classes#nearline>

However, we do not have a requirement to access the data after 30 days; and there are storage classes that are cheaper than nearline storage, so it is not a suitable option.

Ref: <https://cloud.google.com/storage/pricing#storage-pricing>

Use Google Cloud Regional Storage for the first 30 days, and then move videos to Google Persistent Disk. is not right.

Persistent disk pricing is almost double that of standard storage class in Google Cloud Storage service. Plus the persistent disk can only be accessed when attached to another service such as compute engine, GKE, etc making this option very expensive.

Ref: <https://cloud.google.com/storage/pricing#storage-pricing>

Ref: <https://cloud.google.com/compute/disks-image-pricing#persistentdisk>

Use Google Cloud Nearline Storage for the first 30 days, and use lifecycle rules to transition to Coldline Storage. is not right.

Nearline storage class is suitable for storing infrequently accessed data and has costs associated with retrieval. Since the footage is processed regularly within the first 30 days, data retrieval costs may far outweigh the savings made by using nearline storage over standard storage class.

Ref: <https://cloud.google.com/storage/docs/storage-classes#nearline>

Ref: <https://cloud.google.com/storage/pricing#archival-pricing>

Use Google Cloud Regional Storage for the first 30 days, and use lifecycle rules to transition to Coldline Storage. is the right answer.

We save the videos initially in Regional Storage (Standard) which does not have retrieval charges so we do not pay for accessing data within the first 30 days during which the videos are accessed frequently. We only pay for the standard storage costs. After 30 days, we transition the CCTV footage videos to Coldline storage which is a very-low-cost, highly durable storage service for storing infrequently accessed data. Coldline Storage is a better choice than Standard Storage or Nearline Storage in scenarios where slightly lower availability, a 90-day minimum storage duration, and higher costs for data access are acceptable trade-offs for lowered at-rest storage costs. Coldline storage class is cheaper than Nearline storage class.

Ref: <https://cloud.google.com/storage/docs/storage-classes#standard>

Ref: <https://cloud.google.com/storage/docs/storage-classes#coldline>

### 23. Question

Your company has a 3-tier solution running on Compute Engine. The configuration of the current infrastructure is shown below.

Larger image

Each tier has a service account that is associated with all instances within it. You need to enable communication on TCP port 8080 between tiers as follows:

- Instances in tier #1 must communicate with tier #2.
- Instances in tier #2 must communicate with tier #3.

What should you do?

- 1. Create an ingress firewall rule with the following settings: - Targets: all instances with tier #2 service account - Source filter: all instances with tier #1 service account - Protocols: allow TCP: 8080 2. Create an ingress firewall rule with the following settings: - Targets: all instances with tier #3 service account - Source filter: all instances with tier #2 service account - Protocols: allow TCP: 8080
- 1. Create an egress firewall rule with the following settings: - Targets: all instances - Source filter: IP ranges (with the range set to 10.0.2.0/24) - Protocols: allow TCP: 8080 2. Create an egress firewall rule with the following settings: - Targets: all instances - Source filter: IP ranges (with the range set to 10.0.1.0/24) - Protocols: allow TCP: 8080
- 1. Create an ingress firewall rule with the following settings: - Targets: all instances with tier #2 service account - Source filter: all instances with tier #1 service account - Protocols: allow all 2. Create an ingress firewall rule with the following settings: - Targets: all instances with tier #3 service account - Source filter: all instances with tier #2 service account - Protocols: allow all
- 1. Create an ingress firewall rule with the following settings: - Targets: all instances - Source filter: IP ranges (with the range set to 10.0.2.0/24) - Protocols: allow all 2. Create an ingress firewall rule with the following settings: - Targets: all instances - Source filter: IP ranges (with the range set to 10.0.1.0/24) - Protocols: allow all

Unattempted

This resembles a standard 3 tier architecture – web, application, and database; where the web tier can talk to just the application tier; and the application tier can talk to both the web and database tier. The database tier only accepts requests from the application tier and not the web tier.

We want to ensure that Tier 1 can communicate with Tier 2, and Tier 2 can communicate with Tier 3.

1. Create an egress firewall rule with the following settings:
  - Targets: all instances
  - Source filter: IP ranges (with the range set to 10.0.2.0/24)
  - Protocols: allow TCP: 8080
2. Create an egress firewall rule with the following settings:
  - Targets: all instances
  - Source filter: IP ranges (with the range set to 10.0.1.0/24)
  - Protocols: allow TCP: 8080.

is not right.

We are creating egress rules here which allow outbound communication but not ingress rules which are for inbound traffic.

1. Create an ingress firewall rule with the following settings:

- Targets: all instances
  - Source filter: IP ranges (with the range set to 10.0.2.0/24)
  - Protocols: allow all
2. Create an ingress firewall rule with the following settings:
- Targets: all instances
  - Source filter: IP ranges (with the range set to 10.0.1.0/24)
  - Protocols: allow all.

is not right.

If we create an ingress firewall rule with the settings

- Targets: all instances
- Source filter: IP ranges (with the range set to 10.0.1.0/24)
- Protocols: allow all.

then, we are allowing Tier 1 (10.0.1.0/24) access to all instances – including Tier 3 (10.0.3.0/24) which is not desirable.

1. Create an ingress firewall rule with the following settings:

- Targets: all instances with tier #2 service account
- Source filter: all instances with tier #1 service account
- Protocols: allow all

2. Create an ingress firewall rule with the following settings:

- Targets: all instances with tier #3 service account
- Source filter: all instances with tier #2 service account
- Protocols: allow all.

is not right.

The first firewall rule ensures that all instances with tier #2 service account i.e. all instances in Subnet Tier #2 (10.0.2.0/24) can be reached from all instances with tier #1 service account i.e. all instances in Subnet Tier #1 (10.0.1.0/24), on all ports. Similarly, the second firewall rule ensures that all instances with tier #3 service account i.e. all instances in Subnet Tier #3 (10.0.3.0/24) can be reached from all instances with tier #2 service account i.e. all instances in Subnet Tier #2 (10.0.2.0/24), on all ports. Though this matches our requirements, we are opening all ports instead of port 8080 which is our requirement. While this solution works, it is not as secure as the other option (see below)

1. Create an ingress firewall rule with the following settings:

- Targets: all instances with tier #2 service account
- Source filter: all instances with tier #1 service account
- Protocols: allow TCP:8080

2. Create an ingress firewall rule with the following settings:

- Targets: all instances with tier #3 service account
- Source filter: all instances with tier #2 service account
- Protocols: allow TCP: 8080.

is the right answer.

The first firewall rule ensures that all instances with tier #2 service account i.e. all instances in Subnet Tier #2 (10.0.2.0/24) can be reached from all instances with tier #1 service account i.e. all instances in Subnet Tier #1 (10.0.1.0/24), on port 8080. Similarly, the second firewall rule ensures that all instances with tier #3 service account i.e. all instances in Subnet Tier #3 (10.0.3.0/24) can be reached from all instances with tier #2 service account i.e. all instances in Subnet Tier #2 (10.0.2.0/24), on port 8080. This matches our requirements.

#### 24. 24. Question

Your company has a Google Cloud Platform project that uses BigQuery for data warehousing. Your data science team changes frequently and has few members. You need to allow members of this team to perform queries. You want to follow Google-recommended practices. What should you do?

- 1. Create a dedicated Google group in Cloud Identity. 2. Add each data scientist's user account to the group. 3. Assign the BigQuery jobUser role to the group.
- 1. Create a dedicated Google group in Cloud Identity. 2. Add each data scientist's user account to the group. 3. Assign the BigQuery dataViewer user role to the group.
- 1. Create an IAM entry for each data scientist's user account. 2. Assign the BigQuery jobUser role to the group.
- 1. Create an IAM entry for each data scientist's user account. 2. Assign the BigQuery dataViewer user role to the group.

Unattempted

1. Create an IAM entry for each data scientist's user account.

2. Assign the BigQuery dataViewer user role to the group. is not right.

dataViewer provides permissions to Read the dataset's metadata and to list tables in the dataset, and read data and metadata from the dataset's tables. But it does not provide permissions to run jobs, including queries within the project.

Ref: <https://cloud.google.com/bigquery/docs/access-control>

1. Create a dedicated Google group in Cloud Identity.

2. Add each data scientist's user account to the group.

3. Assign the BigQuery dataViewer user role to the group. is not right.

dataViewer provides permissions to Read the dataset's metadata and to list tables in the dataset, and read data and metadata from the dataset's tables. But it does not provide permissions to run jobs, including queries within the project.

Ref: <https://cloud.google.com/bigquery/docs/access-control>

1. Create an IAM entry for each data scientist's user account.

2. Assign the BigQuery jobUser role to the group. is not right.

jobUser is the right role. It provides permissions to run jobs, including queries, within the project. But given that our data science team changes frequently, we do not want to go through this lengthy provisioning and de-provisioning process. Instead, we should be using groups so that provisioning and de-provisioning is as simple as adding/removing the user to/from the group. Google Groups are a convenient way to apply an access policy to a collection of users

Ref: <https://cloud.google.com/bigquery/docs/access-control>

Ref: [https://cloud.google.com/iam/docs/overview#google\\_group](https://cloud.google.com/iam/docs/overview#google_group)

1. Create a dedicated Google group in Cloud Identity.

2. Add each data scientist's user account to the group.

3. Assign the BigQuery jobUser role to the group. is the right answer.

This is the only option that follows Google recommended practices and meets our requirements. jobUser is the right role. It provides permissions to run jobs, including queries, within the project.

And we want to use a group and grant the group all the necessary roles so that whenever a user joins or leaves, they can be provided access to run big query jobs by simply adding them to the group or removing from the group respectively. Google Groups are a convenient way to apply an access policy to a collection of users

Ref: <https://cloud.google.com/bigquery/docs/access-control>

Ref: [https://cloud.google.com/iam/docs/overview#google\\_group](https://cloud.google.com/iam/docs/overview#google_group)

#### 25. 25. Question

Your company has a large quantity of unstructured data in different file formats. You want to perform ETL transformations on the data. You need to make the data accessible on Google Cloud so it can be processed by a Dataflow job. What should you do?

- Upload the data to BigQuery using the bq command line tool.
- Upload the data to Cloud Storage using the gsutil command line tool.
- Upload the data into Cloud SQL using the import function in the console.
- Upload the data into Cloud Spanner using the import function in the console.

Unattempted

The key to answering this question is “unstructured data”.

Upload the data to BigQuery using the bq command line tool. is not right.

The bq load command is used to load data in BigQuery from a local data source i.e. local file but the data has to be in a structured format.

```
bq --location=LOCATION load \
--source_format=FORMAT \
PROJECT_ID:DATASET.TABLE \
PATH_TO_SOURCE \
SCHEMA
```

where

schema: a valid schema. The schema can be a local JSON file, or it can be typed inline as part of the command. You can also use the --autodetect flag instead of supplying a schema definition.

Ref: <https://cloud.google.com/bigquery/docs/loading-data-local#bq>

Upload the data into Cloud SQL using the import function in the console. is not right.

Fully managed relational database service for MySQL, PostgreSQL, and SQL Server. As this is relational database, it is for structured data and not fit for unstructured data.

Ref: <https://cloud.google.com/sql>

Upload the data into Cloud Spanner using the import function in the console. is not right.

Cloud Spanner is the first scalable, enterprise-grade, globally-distributed, and strongly consistent database service built for the cloud specifically to combine the benefits of relational database structure with non-relational horizontal scale. Although Google claims Cloud Spanner is the best of the relational and non-relational worlds, it also says “With Cloud Spanner, you get the best of relational database structure and non-relational database scale and performance with external strong consistency across rows, regions, and continents.”. Cloud spanner is for structured data and not fit for unstructured data.

Ref: <https://cloud.google.com/spanner>

Upload the data to Cloud Storage using the gsutil command line tool. is the right answer.

Cloud storage imposes no such restrictions, you can store large quantities of unstructured data in different file formats. Cloud Storage provides globally unified, scalable, and highly durable object storage for developers and enterprises. In addition, Dataflow can query Cloud Storage files as described in this article Ref: <https://cloud.google.com/dataflow/docs/guides/sql/data-sources-destinations#querying-gcs-filessets>

## 26. Question

Your company has a number of GCP projects that are managed by the respective project teams. Your expenditure of all GCP projects combined has exceeded your operational expenditure budget. At a review meeting, it has been agreed that your finance team should be able to set budgets and view the current charges for all projects in the organization but not view the project resources; and your developers should be able to see the Google Cloud Platform billing charges for only their own projects as well as view resources within the project. You want to follow Google recommended practices to set up IAM roles and permissions. What should you do?

- Add the finance team to the Viewer role for the Project. Add the developers to the Security Reviewer role for each of the billing accounts.
- Add the developers and finance managers to the Viewer role for the Project.

- Add the finance team to the Billing Account Administrator role for each of the billing accounts that they need to manage. Add the developers to the Viewer role for the Project.
- Add the finance team to the default IAM Owner role. Add the developers to a custom role that allows them to see their own spend only.

Unattempted

Add the finance team to the default IAM Owner role. Add the developers to a custom role that allows them to see their own spend only. is not right.

Granting your finance team the default IAM role provides them permissions to manage roles and permissions for a project and subsequently use that to assign them the permissions to view/edit resources in all projects. This is against our requirements. Also, you can write a custom role that lets developers view their project spend but they are missing permissions to view project resources.

Ref: [https://cloud.google.com/iam/docs/understanding-roles#primitive\\_roles](https://cloud.google.com/iam/docs/understanding-roles#primitive_roles)

Add the developers and finance managers to the Viewer role for the Project. is not right.

Granting your finance team the Project viewer role lets them view resources in all projects and doesn't let them set budgets – both are against our requirements.

Ref: [https://cloud.google.com/iam/docs/understanding-roles#primitive\\_roles](https://cloud.google.com/iam/docs/understanding-roles#primitive_roles)

Add the finance team to the Viewer role on all projects. Add the developers to the Security Reviewer role for each of the billing accounts. is not right.

Granting your finance team the Project viewer role lets them view resources in all projects which is against our requirements. Also, the security Reviewer role enables the developers to view custom roles but doesn't let them view the project's costs or project resources.

Ref: [https://cloud.google.com/iam/docs/understanding-roles#primitive\\_roles](https://cloud.google.com/iam/docs/understanding-roles#primitive_roles)

Add the finance team to the Billing Account Administrator role for each of the billing accounts that they need to manage. Add the developers to the Viewer role for the Project. is the right answer.

Billing Account Administrator role is an owner role for a billing account. It provides permissions to manage payment instruments, configure billing exports, view cost information, set budgets, link and unlink projects and manage other user roles on the billing account.

Ref: <https://cloud.google.com/billing/docs/how-to/billing-access>

Project viewer role provides permissions for read-only actions that do not affect the state, such as viewing (but not modifying) existing resources or data; including viewing the billing charges for the project.

[https://cloud.google.com/iam/docs/understanding-roles#primitive\\_roles](https://cloud.google.com/iam/docs/understanding-roles#primitive_roles)

## 27. Question

Your company has a third-party single sign-on (SSO) identity provider that supports Security Assertion Markup Language (SAML) integration with service providers. Your company has users in Cloud Identity and requires them to authenticate using your company's SSO provider. What should you do?

- In Cloud Identity, set up SSO with Google as an identity provider to access custom SAML apps.
- In Cloud Identity, set up SSO with a third-party identity provider with Google as a service provider.
- Obtain OAuth 2.0 credentials, configure the user consent screen, and set up OAuth 2.0 for Mobile & Desktop Apps.
- Obtain OAuth 2.0 credentials, configure the user consent screen, and set up OAuth 2.0 for Web Server Applications.

Unattempted

In Cloud Identity, set up SSO with Google as an identity provider to access custom SAML apps. is not right.  
The question states that you want to use the company's existing Identity provider for SSO, not Google.  
Moreover, your users are in Cloud Identity and not in a GSuite domain so they don't have GSuite Gmail

accounts and therefore can not sign in through Google.

Ref: <https://cloud.google.com/identity/solutions/enable-sso>

Obtain OAuth 2.0 credentials, configure the user consent screen, and set up OAuth 2.0 for Mobile & Desktop Apps. is not right.

OAuth 2.0 credentials are needed for an OAuth 2.0 flow, not SAML flow. See <https://oauth.net/2/> for more information about OAuth 2.0 which is quite a popular protocol for SSO. When you sign in to a 3rd party website using Facebook/Twitter/Google, it uses OAuth 2.0 behind the scenes.

Ref: <https://cloud.google.com/identity/solutions/enable-sso>

Obtain OAuth 2.0 credentials, configure the user consent screen, and set up OAuth 2.0 for Web Server Applications. is not right.

OAuth 2.0 credentials are needed for an OAuth 2.0 flow, not SAML flow. See <https://oauth.net/2/> for more information about OAuth 2.0 which is quite a popular protocol for SSO. When you sign in to a 3rd party website using Facebook/Twitter/Google, it uses OAuth 2.0 behind the scenes.

Ref: <https://cloud.google.com/identity/solutions/enable-sso>

In Cloud Identity, set up SSO with a third-party identity provider with Google as a service provider. is the right answer.

This is the only possible option. You configure applications (service providers) to accept SAML assertions from the company's existing identity provider and users in Cloud Identity can sign in to various applications through the third-party single sign-on (SSO) identity provider. It is important to note that user authentication occurs in the third-party IdP so the absence of a Gmail login is not an issue for signing in.

Ref: <https://cloud.google.com/identity/solutions/enable-sso>

If you have a third-party IdP, you can still configure SSO for third-party apps in the Cloud Identity catalog. User authentication occurs in the third-party IdP, and Cloud Identity manages the cloud apps.

To use Cloud Identity for SSO, your users need Cloud Identity accounts. They sign in through your third-party IdP or using a password on their Cloud Identity accounts.

## 28. Question

Your company has an App Engine application that needs to store stateful data in a proper storage service. Your data is non-relational data. You do not expect the database size to grow beyond 10 GB and you need to have the ability to scale down to zero to avoid unnecessary costs. Which storage service should you use?

- Cloud SQL
- Cloud Datastore
- Cloud Bigtable
- Cloud Dataproc

Unattempted

Cloud SQL. is not right.

Cloud SQL is not suitable for non-relational data. Cloud SQL is a fully-managed database service that makes it easy to set up, maintain, manage, and administer your relational databases on Google Cloud Platform

Ref: <https://cloud.google.com/sql/docs>

Cloud Dataproc. is not right.

Cloud Dataproc is a fast, easy-to-use, fully managed cloud service for running Apache Spark and Apache Hadoop clusters in a simple, cost-efficient way. It is not a database.

Ref: <https://cloud.google.com/dataproc>

Cloud Bigtable. is not right.

Bigtable is a petabyte-scale, massively scalable, fully managed NoSQL database service for large analytical and operational workloads. Cloud Bigtable is overkill for our database which is just 10 GB. Also, Cloud Bigtable can't be scaled down to 0, as there is always a cost with the node, SSD/HDD storage etc.

Ref: <https://cloud.google.com/bigtable>

Cloud Datastore. is the right answer.

Cloud Datastore is a highly-scalable NoSQL database. Cloud Datastore scales seamlessly and automatically with your data, allowing applications to maintain high performance as they receive more traffic;

automatically scales back when the traffic reduces.

Ref: <https://cloud.google.com/datastore/>

## 29. 29. Question

Your company has an existing GCP organization with hundreds of projects and a billing account. Your company recently acquired another company that also has hundreds of projects and its own billing account. You would like to consolidate all GCP costs of both GCP organizations onto a single invoice and you would like to do this as soon as possible. What should you do?

- Link the acquired company's projects to your company's billing account.
- Configure the acquired company's billing account and your company's billing account to export the billing data into the same BigQuery dataset.
- Migrate the acquired company's projects into your company's GCP organization. Link the migrated projects to your company's billing account.
- Create a new GCP organization and a new billing account. Migrate the acquired company's projects and your company's projects into the new GCP organization and link the projects to the new billing account.

Unattempted

Configure the acquired company's billing account and your company's billing account to export the billing data into the same BigQuery dataset. is not right.

Cloud Billing export to BigQuery enables you to export detailed Google Cloud billing data (such as usage and cost estimate data) automatically throughout the day to a BigQuery dataset that you specify. Then you can access your Cloud Billing data from BigQuery for detailed analysis or use a tool like Google Data Studio to visualize your data. Exporting billing data from both the GCP organizations into a single BigQuery dataset can help you have a single view of the billing information, but it doesn't result in a consolidated invoice, which is our requirement.

Migrate the acquired company's projects into your company's GCP organization. Link the migrated projects to your company's billing account. is not right.

While the result is what we need, migrating projects from the acquired company into your company's GCP organization is not straightforward and takes a lot of time. Our requirements state we would like to do this as soon as possible but this option isn't quick.

Create a new GCP organization and a new billing account. Migrate the acquired company's projects and your company's projects into the new GCP organization and link the projects to the new billing account. is not right.

While the result is what we need, migrating projects from both organizations into a new single organization is not straightforward and takes a lot of time. Our requirements state we would like to do this as soon as possible but this option isn't quick.

Link the acquired company's projects to your company's billing account. is the right answer.

This option is the quickest that lets us achieve our end requirement of having all GCP billing in a single invoice. Linking the acquired company's projects to your company's billing account can be very quick and can be scripted using gcloud.

Ref: <https://cloud.google.com/logging/docs/reference/tools/gcloud-logging>

## 30. 30. Question

Your company has chosen to go serverless to enable developers to focus on writing code without worrying about infrastructure. You have been asked to identify a GCP Serverless service that does not limit your developers to specific runtimes. In addition, some of the applications need WebSockets support. What should you suggest?

- Cloud Run
- Cloud Run for Anthos
- App Engine Standard
- Cloud Functions

Unattempted

App Engine Standard. is not right.

Google App Engine Standard offers a limited number of runtimes – Java, Node.js, Python, Go, PHP and Ruby; and at the same time doesn't offer support for Websockets.

Ref: <https://cloud.google.com/appengine/docs/standard>

Cloud Functions. is not right.

Like Google App Engine Standard, Cloud functions offer a limited number of runtimes – Node.js, Python, Go and Java; and doesn't offer support for Websockets.

Ref: <https://cloud.google.com/blog/products/application-development/your-favorite-runtimes-now-generally-available-on-cloud-functions>

Cloud Run. is not right.

Cloud Run lets you run stateless containers in a fully managed environment. As this is container-based, we are not limited to specific runtimes. Developers can write code using their favorite languages (Go, Python, Java, C#, PHP, Ruby, Node.js, Shell, and others). However, Cloud Run does not support Websockets.

Ref: <https://cloud.google.com/run>

Cloud Run for Anthos. is the right answer.

Cloud Run for Anthos leverage Kubernetes and serverless together using Cloud Run integrated with Anthos. As this is container-based, we are not limited to specific runtimes. Developers can write code using their favorite languages (Go, Python, Java, C#, PHP, Ruby, Node.js, Shell, and others). Cloud Run for Anthos is the only serverless GCP offering that supports WebSockets.

<https://cloud.google.com/serverless-options>

### 31. 31. Question

Your company has migrated most of the data center VMs to Google Compute Engine. The remaining VMs in the data center host legacy applications that are due to be decommissioned soon and your company has decided to retain them in the datacenter. Due to a change in the business operational model, you need to introduce changes to one of the legacy applications to read files from Google Cloud Storage. However, your data center does not have access to the internet and your company doesn't want to invest in setting up internet access as the data center is due to be turned off soon. Your data center has a partner interconnect to GCP. You wish to route traffic from your datacenter to Google Storage through partner interconnect. What should you do?

- 1. In on-premises DNS configuration, map storage.cloud.google.com to restricted.googleapis.com, which resolves to the 199.36.153.4/30. 2. Configure Cloud Router to advertise the 199.36.153.4/30 IP address range through the Cloud VPN tunnel. 3. Add a custom static route to the VPC network to direct traffic with the destination 199.36.153.4/30 to the default internet gateway. 4. Created a Cloud DNS managed private zone for storage.cloud.google.com that maps to 199.36.153.4/30 and authorize the zone for use by VPC network
- 1. In on-premises DNS configuration, map storage.cloud.google.com to restricted.googleapis.com, which resolves to the 199.36.153.4/30. 2. Configure Cloud Router to advertise the 199.36.153.4/30 IP address range through the Cloud VPN tunnel. 3. Created a Cloud DNS managed public zone for storage.cloud.google.com that maps to 199.36.153.4/30 and authorize the zone for use by VPC network
- 1. In on-premises DNS configuration, map \*.googleapis.com to restricted.googleapis.com, which resolves to the 199.36.153.4/30. 2. Configure Cloud Router to advertise the 199.36.153.4/30 IP address range through the Cloud VPN tunnel. 3. Created a Cloud DNS managed public zone for \*.googleapis.com that maps to 199.36.153.4/30 and authorize the zone for use by VPC network

1. In on-premises DNS configuration, map \*.googleapis.com to restricted.googleapis.com, which resolves to the 199.36.153.4/30. 2. Configure Cloud Router to advertise the 199.36.153.4/30 IP address range through the Cloud VPN tunnel. 3. Add a custom static route to the VPC network to direct traffic with the destination 199.36.153.4/30 to the default internet gateway. 4. Created a Cloud DNS managed private zone for \*.googleapis.com that maps to 199.36.153.4/30 and authorize the zone for use by VPC network

Unattempted

While Google APIs are accessible on \*.googleapis.com, to restrict Private Google Access within a service perimeter to only VPC Service Controls supported Google APIs and services, hosts must send their requests to the restricted.googleapis.com domain name instead of \*.googleapis.com. The restricted.googleapis.com domain resolves to a VIP (virtual IP address) range 199.36.153.4/30. This IP address range is not announced to the Internet. If you require access to other Google APIs and services that aren't supported by VPC Service Controls, you can use 199.36.153.8/30 (private.googleapis.com). However, we recommend that you use restricted.googleapis.com, which integrates with VPC Service Controls and mitigates data exfiltration risks. In either case, VPC Service Controls service perimeters are always enforced on APIs and services that support VPC Service Controls.

Ref: <https://cloud.google.com/vpc-service-controls/docs/set-up-private-connectivity>

This rules out the two options that map storage.cloud.google.com to restricted.googleapis.com.

The main differences between the remaining two options are

1. Static route in the VPC network.
2. Public/Private zone.

According to Google's guide on setting up private connectivity, in order to configure a route to restricted.googleapis.com within the VPC, we need to create a static route whose destination is 199.36.153.4/30 and whose next hop is the default Internet gateway.

So, the right answer is

1. In on-premises DNS configuration, map \*.googleapis.com to restricted.googleapis.com, which resolves to the 199.36.153.4/30.
2. Configure Cloud Router to advertise the 199.36.153.4/30 IP address range through the Cloud VPN tunnel.
3. Add a custom static route to the VPC network to direct traffic with the destination 199.36.153.4/30 to the default internet gateway.
4. Created a Cloud DNS managed private zone for \*.googleapis.com that maps to 199.36.153.4/30 and authorize the zone for use by VPC network

Here's more information about how to set up private connectivity to Google's services through VPC.  
Ref: <https://cloud.google.com/vpc/docs/private-access-options#private-vips>

In the following example, the on-premises network is connected to a VPC network through a Cloud VPN tunnel. Traffic from on-premises hosts to Google APIs travels through the tunnel to the VPC network. After traffic reaches the VPC network, it is sent through a route that uses the default internet gateway as its next hop. The next hop allows traffic to leave the VPC network and be delivered to restricted.googleapis.com (199.36.153.4/30).

? The on-premises DNS configuration maps \*.googleapis.com requests to restricted.googleapis.com, which resolves to the 199.36.153.4/30.

? Cloud Router has been configured to advertise the 199.36.153.4/30 IP address range through the Cloud VPN tunnel by using a custom route advertisement. Traffic going to Google APIs is routed through the tunnel to the VPC network.

? A custom static route was added to the VPC network that directs traffic with the destination

199.36.153.4/30 to the default internet gateway (as the next hop). Google then routes traffic to the appropriate API or service.

If you created a Cloud DNS managed private zone for \*.googleapis.com that maps to 199.36.153.4/30 and have authorized that zone for use by your VPC network, requests to anything in the googleapis.com domain are sent to the IP addresses that are used by restricted.googleapis.com

### 32. 32. Question

Your company hosts a number of applications in Google Cloud and requires that log messages from all applications be archived for 10 years to comply with local regulatory requirements. Which approach should you use?

- 1. Enable Stackdriver Logging API 2. Configure web applications to send logs to Stackdriver  
3. Export logs to Google Cloud Storage
- Grant the security team access to the logs in each Project
- 1. Enable Stackdriver Logging API 2. Configure web applications to send logs to Stackdriver  
3. Export logs to BigQuery
- 1. Enable Stackdriver Logging API 2. Configure web applications to send logs to Stackdriver

Unattempted

Grant the security team access to the logs in each Project. is not right.

Granting the security team access to the logs in each Project doesn't guarantee log retention. If the security team is to come up with a manual process to copy all the logs files into another archival source, the ongoing operational costs can be huge.

1. Enable Stackdriver Logging API

2. Configure web applications to send logs to Stackdriver. is not right.

In Stackdriver, application logs are retained by default for just 30 days after which they are purged.

Ref: <https://cloud.google.com/logging/quotas>

While it is possible to configure a custom retention period of 10 years, storing logs in Stackdriver is very expensive compared to Cloud Storage. Stackdriver charges \$.01 per GB per month, whereas something like Cloud Storage Coldline Storage costs \$.0007 per GB per month (30% cheaper) and Cloud Storage Archive Storage costs 0.004 per GB per month (60% cheaper than Stackdriver)

Ref: <https://cloud.google.com/logging/docs/storage#pricing>

Ref: <https://cloud.google.com/storage/pricing>

The difference between the remaining two options is whether we store the logs in BigQuery or Google Cloud Storage.

1. Enable Stackdriver Logging API

2. Configure web applications to send logs to Stackdriver

3. Export logs to BigQuery. is not right.

While enabling Stackdriver Logging API and having the applications send logs to stack driver is a good start, exporting and storing logs in BigQuery is fairly expensive. In BigQuery, Active storage costs \$.02 per GB per month and Long-term storage costs \$.01 per GB per month. In comparison, Google Cloud Storage offers several storage classes that are significantly cheaper.

Ref: <https://cloud.google.com/bigquery/pricing>

Ref: <https://cloud.google.com/storage/pricing>

1. Enable Stackdriver Logging API

2. Configure web applications to send logs to Stackdriver

3. Export logs to Google Cloud Storage. is the right answer.

Google Cloud Storage offers several storage classes such as Nearline Storage (\$.01 per GB per Month) Coldline Storage (\$.0007 per GB per Month) and Archive Storage (\$.0004 per GB per month) which are significantly cheaper than the storage options covered by the above options above.

Ref: <https://cloud.google.com/storage/pricing>

### 33. 33. Question

Your company implemented BigQuery as an enterprise data warehouse. Users from multiple business units run queries on this data warehouse. However, you notice that query costs for BigQuery are very high, and you need to control costs. Which two methods should you use? (Choose two.)

- Split the users from business units to multiple projects.
- Apply a user- or project-level custom query quota for BigQuery data warehouse.
- Create separate copies of your BigQuery data warehouse for each business unit.
- Split your BigQuery data warehouse into multiple data warehouses for each business unit.
- Change your BigQuery query model from on-demand to flat rate. Apply the appropriate number of slots to each Project.

Unattempted

Once your data is loaded into BigQuery, you are charged for storing it. Storage pricing is based on the amount of data stored in your tables when it is uncompressed. BigQuery doesn't charge for the query execution based on the output of the query (i.e. bytes returned) but on the number of bytes processed (also referred to as bytes read or bytes scanned) in order to arrive at the output of the query. You are charged for the number of bytes processed whether the data is stored in BigQuery or in an external data source such as Cloud Storage, Google Drive, or Cloud Bigtable. On-demand pricing is based solely on usage. You are charged for the bytes scanned even if your query itself doesn't return any data.

Ref: <https://cloud.google.com/bigquery/pricing>

Split the users from business units to multiple projects. is not right.

The bytes scanned is not expected to go down by splitting the users into multiple projects so this wouldn't reduce/control the costs.

Ref: <https://cloud.google.com/bigquery/pricing>

Split your BigQuery data warehouse into multiple data warehouses for each business unit. is not right.

The bytes scanned is not expected to go down by splitting the BigQuery warehouse into two so this wouldn't reduce/control the costs either.

Ref: <https://cloud.google.com/bigquery/pricing>

Create separate copies of your BigQuery data warehouse for each business unit. is not right.

Creating separate copies of the BigQuery data warehouse for each business unit is going to increase your costs. Not only is this expected to reduce the bytes scanned, but this is also going to increase the storage costs as we are now storing double the amount of data.

Ref: <https://cloud.google.com/bigquery/pricing>

Apply a user- or project-level custom query quota for BigQuery data warehouse. is the right answer.

BigQuery limits the maximum rate of incoming requests and enforces appropriate quotas on a per-project basis. You can set various limits to control costs such as Concurrent rate limit for interactive queries, Concurrent rate limit for interactive queries against Bigtable external data sources, Concurrent rate limit for legacy SQL queries that contain UDFs, Cross-region federated querying, Daily query size limit, etc.

<https://cloud.google.com/bigquery/quotas>

Change your BigQuery query model from on-demand to flat rate. Apply the appropriate number of slots to each Project. is the right answer.

This pricing option is best for customers who desire cost predictability. Flat-rate customers purchase dedicated resources for query processing and are not charged for individual queries. BigQuery offers flat-rate pricing for customers who prefer a stable cost for queries rather than paying the on-demand price per TB of data processed. You can choose to use flat-rate pricing using BigQuery Reservations. When you enroll in flat-rate pricing, you purchase slot commitments – dedicated query processing capacity, measured in BigQuery slots. Your queries consume this capacity, and you are not billed for bytes processed. If your capacity demands exceed your committed capacity, BigQuery will queue up slots, and you will not be charged additional fees.

Ref: [https://cloud.google.com/bigquery/pricing#flat\\_rate\\_pricing](https://cloud.google.com/bigquery/pricing#flat_rate_pricing)

### 34. 34. Question

Your company is moving all corporate applications to Google Cloud Platform. The security team wants detailed visibility of all GCP projects in the organization. You provision the Google Cloud Resource Manager and set up yourself as the org admin. What Google Cloud Identity and Access Management (Cloud IAM) roles should you give to the security team?

- Grant roles/resourcemanager.organizationViewer and roles/viewer.
- Grant roles/resourcemanager.organizationViewer and roles/owner.
- Grant roles/owner, roles/networkmanagement.admin.
- Grant roles/resourcemanager.organizationAdmin and roles/browser.

Unattempted

The security team needs detailed visibility of all GCP projects in the organization so they should be able to view all the projects in the organization as well as view all resources within these projects.

Grant roles/resourcemanager.organizationViewer and roles/owner. is not right.  
roles/resourcemanager.organizationViewer role provides permissions to see the organization in the Cloud Console without having access to view all resources in the organization.  
roles/owner provides permissions to manage roles and permissions for a project and all resources within the project and set up billing for a project.  
Neither of the roles give the security team visibility of the projects in the organization.  
Ref: <https://cloud.google.com/resource-manager/docs/access-control-org>  
Ref: <https://cloud.google.com/iam/docs/understanding-roles#organization-policy-roles>  
Grant roles/resourcemanager.organizationAdmin and roles/browser. is not right.  
roles/resourcemanager.organizationAdmin provides access to administer all resources belonging to the organization. This doesn't follow the least privilege principle. Our security team needs detailed visibility i.e. read-only access but should not be able to administer resources..  
Ref: <https://cloud.google.com/iam/docs/understanding-roles#organization-policy-roles>  
Grant roles/owner, roles/networkmanagement.admin. is not right.  
roles/owner provides permissions to manage roles and permissions for a project and all resources within the project and set up billing for a project.  
roles/networkmanagement.admin provides full access to Cloud Network Management resources.  
Neither of the roles give the security team visibility of the projects in the organization.  
Ref: <https://cloud.google.com/resource-manager/docs/access-control-org>  
Ref: <https://cloud.google.com/iam/docs/understanding-roles#organization-policy-roles>  
Grant roles/resourcemanager.organizationViewer and roles/viewer. is the right answer.  
roles/viewer provides permissions to view existing resources or data.  
roles/resourcemanager.organizationViewer provides access to view an organization.  
With the two roles, the security team can view the organization including all the projects and folders; as well as view all the resources within the projects.  
Ref: <https://cloud.google.com/resource-manager/docs/access-control-org>  
Ref: <https://cloud.google.com/iam/docs/understanding-roles#organization-policy-roles>

### 35. Question

Your company is moving from an on-premises environment to Google Cloud Platform (GCP). You have multiple development teams that use Cassandra environments as backend databases. They all need a development environment that is isolated from other Cassandra instances. You want to move to GCP quickly and with minimal support effort. What should you do?

- 1. Build an instruction guide to install Cassandra on GCP. 2. Make the instruction guide accessible to your developers.
- 1. Advise your developers to go to Cloud Marketplace. 2. Ask the developers to launch a Cassandra image for their development work.

- 1. Build a Cassandra Compute Engine instance and take a snapshot of it. 2. Use the snapshot to create instances for your developers.
- 1. Build a Cassandra Compute Engine instance and take a snapshot of it. 2. Upload the snapshot to Cloud Storage and make it accessible to your developers. 3. Build instructions to create a Compute Engine instance from the snapshot so that developers can do it themselves.

Unattempted

1. Build an instruction guide to install Cassandra on GCP.

2. Make the instruction guide accessible to your developers. is not right.

There is a very simple and straightforward way to deploy Cassandra as a Service, called Astra, on the Google Cloud Marketplace. You don't need to come up with an installation guide and ask your developers to do it.

Ref: <https://cloud.google.com/blog/products/databases/open-source-cassandra-now-managed-on-google-cloud>

Ref: <https://console.cloud.google.com/marketplace/details/click-to-deploy-images/cassandra?filter=price:free&filter=category:database&id=25ca0967-cd8e-419e-b554-fe32e87f04be&pli=1>

1. Build a Cassandra Compute Engine instance and take a snapshot of it.

2. Use the snapshot to create instances for your developers. is not right.

Like above, there is a very simple and straightforward way to deploy Cassandra as a Service, called Astra, on the Google Cloud Marketplace. You don't need to do this in a convoluted way.

Ref: <https://cloud.google.com/blog/products/databases/open-source-cassandra-now-managed-on-google-cloud>

Ref: <https://console.cloud.google.com/marketplace/details/click-to-deploy-images/cassandra?filter=price:free&filter=category:database&id=25ca0967-cd8e-419e-b554-fe32e87f04be&pli=1>

1. Build a Cassandra Compute Engine instance and take a snapshot of it.

2. Upload the snapshot to Cloud Storage and make it accessible to your developers.

3. Build instructions to create a Compute Engine instance from the snapshot so that developers can do it themselves. is not right.

Like above, there is a very simple and straightforward way to deploy Cassandra as a Service, called Astra, on the Google Cloud Marketplace. You don't need to do this in a convoluted way.

Ref: <https://cloud.google.com/blog/products/databases/open-source-cassandra-now-managed-on-google-cloud>

Ref: <https://console.cloud.google.com/marketplace/details/click-to-deploy-images/cassandra?filter=price:free&filter=category:database&id=25ca0967-cd8e-419e-b554-fe32e87f04be&pli=1>

1. Advise your developers to go to Cloud Marketplace.

2. Ask the developers to launch a Cassandra image for their development work. is the right answer.

You can deploy Cassandra as a Service, called Astra, on the Google Cloud Marketplace. Not only do you get a unified bill for all GCP services, your Developers can now create Cassandra clusters on Google Cloud in minutes and build applications with Cassandra as a database as a service without the operational overhead of managing Cassandra. Each instance is deployed to a separate set of VM instances (at the time of writing this, 3 x VM instance: 4 vCPUs + 26 GB memory (n1-highmem-4) + 10-GB Boot Disk) which are all isolated from the VM instances for other Cassandra deployments.

Ref: <https://cloud.google.com/blog/products/databases/open-source-cassandra-now-managed-on-google-cloud>

Ref: <https://console.cloud.google.com/marketplace/details/click-to-deploy-images/cassandra?filter=price:free&filter=category:database&id=25ca0967-cd8e-419e-b554-fe32e87f04be&pli=1>

## 36. Question

Your Company is planning to migrate all Java web applications to Google App Engine. However, you still want to continue using your on-premise database. How can you set up the app engine to communicate with your on-premise database while minimizing effort?

- Setup the application using App Engine Flexible environment with Cloud Router to connect to an on-premise database.
- Setup the application using App Engine Standard environment with Cloud VPN to connect to an on-premise database.
- Setup the application using App Engine Standard environment with Cloud Router to connect to an on-premise database.
- Setup the application using App Engine Flexible environment with Cloud VPN to connect to an on-premise database.

Unattempted

Setup the application using App Engine Standard environment with Cloud Router to connect to an on-premise database. is not right.

Cloud router by itself is not sufficient to connect VPC to an on-premise network. Cloud Router enables you to dynamically exchange routes between your Virtual Private Cloud (VPC) and on-premises networks by using Border Gateway Protocol (BGP).

Ref: <https://cloud.google.com/router>

Setup the application using App Engine Flexible environment with Cloud Router to connect to an on-premise database. is not right.

Cloud router by itself is not sufficient to connect VPC to an on-premise network. Cloud Router enables you to dynamically exchange routes between your Virtual Private Cloud (VPC) and on-premises networks by using Border Gateway Protocol (BGP).

Ref: <https://cloud.google.com/router>

Setup the application using App Engine Flexible environment with Cloud VPN to connect to an on-premise database. is not right.

You need Cloud VPN to connect VPC to an on-premise network.

Ref: <https://cloud.google.com/vpn/docs/concepts/overview>

However, we don't have a requirement to run docker containers and App Engine Standard already supports the requirements of our existing web applications, we should avoid using App Engine Flexible. Converting to a container model involves effort and we want to minimize effort.

Setup the application using App Engine Standard environment with Cloud VPN to connect to an on-premise database. is the right answer.

You need Cloud VPN to connect VPC to an on-premise network.

Ref: <https://cloud.google.com/vpn/docs/concepts/overview>

And since we don't have a requirement to run docker containers, App Engine Standard already supports the requirements of our existing web applications – Java runtime environment, so we should use App Engine Standard

### 37. Question

Your company owns a web application that lets users post travel stories. You began noticing errors in logs for a specific Deployment. The deployment is responsible for translating a post from one language to another. You've narrowed the issue down to a specific container named "msg-translator-22" that is throwing the errors. You are unable to reproduce the error in any other environment, and none of the other containers serving the deployment have this issue. You would like to connect to this container to figure out the root cause. What steps would allow you to run commands against the msg-translator-22?

- Use the kubectl run msg-translator-22 /bin/ bash command to run a shell on that container.
- Use the kubectl exec -it -- /bin/bash command to run a shell on that container.
- Use the kubectl run command to run a shell on that container.
- Use the kubectl exec -it msg-translator-22 -- /bin/bash command to run a shell on that container.

Unattempted

Use the kubectl run command to run a shell on that container. is not right.

kubectl run creates and runs a deployment. It creates a deployment or a job to manage the created container(s). It is not possible to use kubectl run to connect to an existing container.

<https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#run>

Use the kubectl run msg-translator-22 /bin/ bash command to run a shell on that container. is not right.

kubectl run creates and runs a deployment. It creates a deployment or a job to manage the created container(s). It is not possible to use kubectl run to connect to an existing container.

<https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#run>

Use the kubectl exec -it — /bin/bash command to run a shell on that container. is not right.

While kubectl exec is used to execute a command in a container, the command above doesn't quite work because we haven't passed to it the identifier of the container.

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#exec>

Use the kubectl exec -it msg-translator-22 — /bin/bash command to run a shell on that container. is the right answer.

kubectl exec is used to execute a command in a container. We pass the container name msg-translator-22 so kubectl exec knows which container to connect to. And we pass the command /bin/bash to it, so it starts a shell on the container and we can then run custom commands and identify the root cause of the issue.

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#exec>

### 38. Question

Your company plans to store sensitive PII data in a cloud storage bucket. Your compliance department doesn't like encrypting sensitive PII data with Google-managed keys and has asked you to ensure the new objects uploaded to this bucket are encrypted by customer-managed encryption keys. What should you do? (Select Three)

- In the bucket advanced settings, select the Customer-supplied key and then select a Cloud KMS encryption key.
- Use gsutil with --encryption-key=[ENCRYPTION\_KEY] when uploading objects to the bucket.
- Use gsutil with -o "GSUtil:encryption\_key=[KEY\_RESOURCE]" when uploading objects to the bucket.
- In the bucket advanced settings, select the Customer-managed key and then select a Cloud KMS encryption key.
- Modify .boto configuration to include encryption\_key = [KEY\_RESOURCE] when uploading objects to bucket

Unattempted

In the bucket advanced settings, select the Customer-supplied key and then select a Cloud KMS encryption key. is not right.

The customer-supplied key is not an option when selecting the encryption method in the console.

Use gsutil with --encryption-key=[ENCRYPTION\_KEY] when uploading objects to the bucket. is not right. gsutil doesn't accept the flag --encryption-key. gsutil can be set up to use an encryption key by modifying boto configuration or by specifying a top-level -o flag but neither of these is included in this option.

Ref: <https://cloud.google.com/storage/docs/encryption/using-customer-managed-keys>

In the bucket advanced settings, select the Customer-managed key and then select a Cloud KMS encryption key. is the right answer.

Our compliance department wants us to use customer-managed encryption keys. We can select Customer-Managed radio and provide a cloud KMS encryption key to encrypt objects with the customer-managed key. This fit our requirements.

Use gsutil with -o "GSUtil:encryption\_key=[KEY\_RESOURCE]" when uploading objects to the bucket. is the right answer.

We can have gsutil use an encryption key by using the -o top-level flag: -o "GSUtil:encryption\_key=[KEY\_RESOURCE]".

Ref: <https://cloud.google.com/storage/docs/encryption/using-customer-managed-keys#add-object-key>  
Modify .boto configuration to include encryption\_key = [KEY\_RESOURCE] when uploading objects to bucket. is the right answer.

As an alternative to the -o top-level flag, gsutil can also use an encryption key if .boto configuration is modified to specify the encryption key.

encryption\_key = [KEY\_RESOURCE]

Ref: <https://cloud.google.com/storage/docs/encryption/using-customer-managed-keys#add-object-key>

### 39. 39. Question

Your company plans to store sensitive PII data in a cloud storage bucket. Your compliance department has asked you to ensure the objects in this bucket are encrypted by customer-managed encryption keys. What should you do?

- In the bucket advanced settings, select Google-managed key and then select a Cloud KMS encryption key.
- Recreate the bucket to use a Customer-managed key. Encryption can only be specified at the time of bucket creation.
- In the bucket advanced settings, select Customer-supplied key and then select a Cloud KMS encryption key.
- In the bucket advanced settings, select Customer-managed key and then select a Cloud KMS encryption key.

Unattempted

In the bucket advanced settings, select Customer-supplied key and then select a Cloud KMS encryption key. is not right.

Customer-Supplied key is not an option when selecting the encryption method in the console. Moreover, we want to use customer managed encryption keys and not customer supplied encryption keys. This does not fit our requirements.

In the bucket advanced settings, select Google-managed key and then select a Cloud KMS encryption key. is not right.

While Google-managed key is an option when selecting the encryption method in console, we want to use customer managed encryption keys and not Google Managed encryption keys. This does not fit our requirements.

Recreate the bucket to use a Customer-managed key. Encryption can only be specified at the time of bucket creation. is not right.

Bucket encryption can be changed at any time. The bucket doesn't have to be recreated to change encryption.

Ref: <https://cloud.google.com/storage/docs/encryption/using-customer-managed-keys#add-default-key>

In the bucket advanced settings, select Customer-managed key and then select a Cloud KMS encryption key. is the right answer.

This option correctly selects the Customer-managed key and then the key to use which satisfies our requirement. See the screenshot below for reference.

Ref: <https://cloud.google.com/storage/docs/encryption/using-customer-managed-keys#add-default-key>

### 40. 40. Question

Your company procured a license for a third-party cloud-based document signing system for the procurement team. All members of the procurement team need to sign in with the same service account.

Your security team prohibits sharing service account passwords. You have been asked to recommend a solution that lets the procurement team login as the service account in the document signing system but without the team knowing the service account password. What should you do?

- Ask the third-party provider to enable SAML for the application and set the credentials to the service account credentials.
- Ask the third-party provider to enable OAuth 2.0 for the application and set the credentials to the service account credentials.
- Have a single person from the procurement team access document signing system with the service account credentials.
- Register the application as a password vaulted app and set the credentials to the service account credentials.

Unattempted

Ask the third-party provider to enable SAML for the application and set the credentials to always use service account credentials. is not right.

The application may or may not support SAML. Moreover, you can't set the credentials in SAML integration to always use a particular account. The authentication is carried out by IdP such as GSuite or a third-party identity provider. Since users are prohibited from sharing the service account credentials, they wouldn't be able to sign in through the IdP with the service account credentials.

Ask the third-party provider to enable OAuth 2.0 for the application and set the credentials to always use service account credentials. is not right.

The application may or may not support OAuth 2.0. Moreover, you can't set the credentials in SAML integration to always use a particular account. The authentication is carried out by IdP such as GSuite or a third-party identity provider. Since users are prohibited from sharing the service account credentials, they wouldn't be able to sign in through the IdP with the service account credentials.

Have a single person from the procurement team access document signing system with the service account credentials. is not right.

While this would prevent password reuse, it goes against our requirements and results in a single person dependency.

Register the application as a password vaulted app and set the credentials to the service account credentials. is the right answer.

As a G Suite or Cloud Identity administrator, the password vaulted apps service enables you to manage access to some of the apps that don't support federation and that are available to users on the User Dashboard. The password vaulted apps service saves login credential sets for applications and assigns those credential sets to users through group association. When a user has access to one of these applications through a group, they can sign in to the application through the user dashboard, or they can sign in directly from the specific application. This functionality is possible by leveraging Chrome or Firefox extensions/plugins. When adding an app to the password vaulted apps service, you can search and choose from the available web-based applications in the app library, or you can add a custom app. You can then manage usernames and passwords safely while providing users in your organization with quick one-click access to all of the apps they already use.

Ref: <https://support.google.com/cloudidentity/answer/9178974?hl=en>

#### 41. 41. Question

Your company publishes large files on an Apache web server that runs on a Compute Engine instance. The Apache web server is not the only application running in the project. You want to receive an email when the egress network costs for the server exceed 100 dollars for the current month as measured by Google Cloud Platform (GCP). What should you do?

- Set up a budget alert on the project with an amount of 100 dollars, a threshold of 100%, and notification type of email.
- Set up a budget alert on the billing account with an amount of 100 dollars, a threshold of 100%, and notification type of email.
- Export the billing data to BigQuery. Create a Cloud Function that uses BigQuery to sum the egress network costs of the exported billing data for the Apache web server for the current month and sends an email if it is over 100 dollars. Schedule the Cloud Function using Cloud Scheduler to run hourly.
- Use the Stackdriver Logging Agent to export the Apache web server logs to Stackdriver Logging. Create a Cloud Function that uses BigQuery to parse the HTTP response log data in Stackdriver for the current month and sends an email if the size of all HTTP responses, multiplied by current GCP egress prices, totals over 100 dollars. Schedule the Cloud Function using Cloud Scheduler to run hourly.

Unattempted

Set up a budget alert on the project with an amount of 100 dollars, a threshold of 100%, and notification type of email. is not right.

This budget alert is defined for the project which means it includes all costs and not just the egress network costs – which goes against our requirements; and it also contains costs across all applications and not just the Compute Engine instance containing the Apache web server. While it is possible to set budget scope to include the Product (i.e. Google Compute Engine) and a label that uniquely identifies the specific compute engine instance, the option doesn't mention this.

Ref: <https://cloud.google.com/billing/docs/how-to/budgets#budget-scope>

Set up a budget alert on the billing account with an amount of 100 dollars, a threshold of 100%, and notification type of email. is not right.

Like above, but worse as this budget alert includes costs from all projects linked to the billing account. And like above, while it is possible to scope an alert down to Project/Product/Labels, the option doesn't mention this.

Ref: <https://cloud.google.com/billing/docs/how-to/budgets#budget-scope>

Use the Stackdriver Logging Agent to export the Apache web server logs to Stackdriver Logging. Create a Cloud Function that uses BigQuery to parse the HTTP response log data in Stackdriver for the current month and sends an email if the size of all HTTP responses, multiplied by current GCP egress prices, totals over 100 dollars. Schedule the Cloud Function using Cloud Scheduler to run hourly. is not right.

You can't arrive at the exact egress costs with this approach. You can configure apache logs to include the response object size.

Ref: <https://httpd.apache.org/docs/1.3/logs.html#common>

And you can then do what this option says to arrive at the combined size of all the responses but this is not 100% accurate as it does not include header sizes. Even if we assume the header size is insignificant compare to the large files published on apache web server, our question asks us to do this the Google way "as measured by Google Cloud Platform (GCP)". GCP does not look at the response sizes in the Apache log files to determine the egress costs. The GCP egress calculator takes into consideration the source and destination (source = the region that hosts the Compute Engine instance running Apache Web Server; and the destination is the destination region of the packet). The egress cost is different for different destinations as shown in this pricing reference.

Ref: [https://cloud.google.com/vpc/network-pricing#internet\\_egress](https://cloud.google.com/vpc/network-pricing#internet_egress)

The Apache logs do not give you the destination information and without this information, you can't accurately calculate the egress costs.

Export the billing data to BigQuery. Create a Cloud Function that uses BigQuery to sum the egress network costs of the exported billing data for the Apache web server for the current month and sends an email if it is over 100 dollars. Schedule the Cloud Function using Cloud Scheduler to run hourly. is the right answer.

This is the only option that satisfies our requirement. We do it the Google way by (re)using the Billing Data that GCP uses. And we scope down the costs to just egress network costs for the apache web server.

Finally, we schedule this to run hourly and send an email if the costs exceed 100 dollars.

Your company recently migrated all infrastructure to Google Cloud Platform (GCP) and you want to use Google Cloud Build to build all container images. You want to store the build logs in Google Cloud Storage. You also have a requirement to push the images to Google Container Registry. You wrote a cloud build YAML configuration file with the following contents.

steps:

```
- name: 'gcr.io/cloud-builders/docker'
args: ['build', '-t', 'gcr.io/[PROJECT_ID]/[IMAGE_NAME]', '.']
images: ['gcr.io/[PROJECT_ID]/[IMAGE_NAME]']
```

How should you execute Cloud build to satisfy these requirements?

- Execute gcloud builds run --config=[CONFIG\_FILE\_PATH] --gcs-log-dir=[GCS\_LOG\_DIR] [SOURCE]
- Execute gcloud builds submit --config=[CONFIG\_FILE\_PATH] --gcs-log-dir=[GCS\_LOG\_DIR] [SOURCE]
- Execute gcloud builds submit --config=[CONFIG\_FILE\_PATH] [SOURCE]
- Execute gcloud builds push --config=[CONFIG\_FILE\_PATH] [SOURCE]

Unattempted

Execute gcloud builds push --config=[CONFIG\_FILE\_PATH] [SOURCE]. is not right.

gcloud builds command does not support push operation. The correct operation to build images and push them to gcr is submit.

Ref: <https://cloud.google.com/sdk/gcloud/reference/builds/submit>

Execute gcloud builds run --config=[CONFIG\_FILE\_PATH] --gcs-log-dir=[GCS\_LOG\_DIR] [SOURCE]. is not right.

gcloud builds command does not support run operation. The correct operation to build images and push them to gcr is submit.

Ref: <https://cloud.google.com/sdk/gcloud/reference/builds/submit>

Execute gcloud builds submit --config=[CONFIG\_FILE\_PATH] [SOURCE]. is not right.

This command correctly builds the container image and pushes the image to GCR (Google Container Registry) but doesn't upload the build logs to Google Cloud Storage which is one of our requirements.

Ref: <https://cloud.google.com/sdk/gcloud/reference/builds/submit>

Ref: <https://cloud.google.com/cloud-build/docs/building/build-containers>

Execute gcloud builds submit --config=[CONFIG\_FILE\_PATH] --gcs-log-dir=[GCS\_LOG\_DIR] [SOURCE]. is the right answer.

This command correctly builds the container image, pushes the image to GCR (Google Container Registry) and uploads the build logs to Google Cloud Storage.

--config flag specifies the YAML or JSON file to use as the build configuration file.

--gcs-log-dir specifies the directory in Google Cloud Storage to hold build logs.

[SOURCE] is the location of the source to build. The location can be a directory on a local disk or a gzipped archive file (.tar.gz) in Google Cloud Storage.

Ref: <https://cloud.google.com/sdk/gcloud/reference/builds/submit>

Ref: <https://cloud.google.com/cloud-build/docs/building/build-containers>

#### 43. Question

Your company runs a very successful web platform and has accumulated 3 petabytes of customer activity data in sharded MySQL database located in your datacenter. Due to storage limitations in your on-premise data center, your company has decided to move this data to GCP. The data must be available all through the day. Your business analysts, who have experience of using a SQL Interface, have asked for a seamless transition. How should you store the data so that availability is ensured while optimizing the ease of analysis for the business analysts?

- Import data into Google Cloud SQL.
- Import flat files into Google Cloud Storage.
- Import data into Google Cloud Datastore.
- Import data into Google BigQuery.

Unattempted

Import data into Google Cloud SQL. is not right.

Cloud SQL is a fully-managed relational database service. It supports MySQL so the migration of data from your data center to cloud can be straightforward but Google Cloud SQL cannot handle petabyte-scale data. The current second-generation instances limit the storage to approximately 30TB.

Ref: <https://cloud.google.com/sql#overview>

Ref: <https://cloud.google.com/sql/docs/quotas>

Import flat files into Google Cloud Storage. is not right.

Cloud Storage is a service for storing objects in Google Cloud. You store objects in containers called buckets. You could export the MySQL data into files and import them into Google Cloud Storage, but it doesn't offer an SQL Interface to run queries/reports.

Ref: <https://cloud.google.com/storage/docs/introduction>

Import data into Google Cloud Datastore. is not right.

Your business analysts are already familiar with SQL Interface so we need a service that supports SQL. However, Cloud Datastore is a NoSQL document database. Cloud Datastore doesn't support SQL (it supports GQL which is similar to SQL, but not identical).

Ref: [https://cloud.google.com/datastore/docs/reference/gql\\_reference](https://cloud.google.com/datastore/docs/reference/gql_reference)

Ref: <https://cloud.google.com/datastore/docs/concepts/overview>

Import data into Google BigQuery. is the right answer.

Bigquery is a petabyte-scale serverless, highly scalable, and cost-effective cloud data warehouse that offers blazing-fast speeds, and with zero operational overhead. BigQuery supports a standard SQL dialect that is ANSI:2011 compliant, which reduces the impact and enables a seamless transition for your business analysts.

Ref: <https://cloud.google.com/bigquery>

#### 44. Question

Your company set up a complex organizational structure on Google Could Platform. The structure includes hundreds of folders and projects. Only a few team members should be able to view the hierarchical structure. You need to assign minimum permissions to these team members and you want to follow Google recommended practices. What should you do?

- Add the users to roles/browser role.
- Add the users to roles/iam.roleViewer role.
- Add the users to a group and add this group to roles/browser role.
- Add the users to a group and add this group to roles/iam.roleViewer role.

Unattempted

Add the users to roles/iam.roleViewer role. is not right.

roles/iam.roleViewer provides read access to all custom roles in the project and doesn't satisfy our requirement of viewing organization hierarchy.

Ref: <https://cloud.google.com/iam/docs/understanding-roles>

Add the users to a group and add this group to roles/iam.roleViewer role. is not right.

roles/iam.roleViewer provides read access to all custom roles in the project and doesn't satisfy our requirement of viewing organization hierarchy.

Ref: <https://cloud.google.com/iam/docs/understanding-roles>

Add the users to roles/browser role. is not right.

roles/browser provides read access to browse the hierarchy for a project, including the folder, organization, and Cloud IAM policy. This role doesn't include permission to view resources in the project. Although this is the role we require, you want to follow Google recommended practices which means we should instead add a group to the role and add users to the group instead of granting the role individually to users.

Ref: <https://cloud.google.com/iam/docs/understanding-roles>

Add the users to a group and add this group to roles/browser role. is the right answer.

roles/browser Read access to browse the hierarchy for a project, including the folder, organization, and Cloud IAM policy. This role doesn't include permission to view resources in the project. And you follow Google recommended practices by adding users to the group and group to the role. Groups are a convenient way to apply an access policy to a collection of users. You can grant and change access controls for a whole group at once instead of granting or changing access controls one at a time for individual users or service accounts. You can also easily add members to and remove members from a Google group instead of updating a Cloud IAM policy to add or remove users.

Ref: <https://cloud.google.com/iam/docs/understanding-roles>

Ref: <https://cloud.google.com/iam/docs/overview>

#### 45. 45. Question

Your company stores customer PII data in Cloud Storage buckets. A subset of this data is regularly imported into a BigQuery dataset to carry out analytics. You want to make sure the access to this bucket is strictly controlled. Your analytics team needs read access on the bucket so that they can import data in BigQuery. Your operations team needs read/write access to both the bucket and BigQuery dataset to add Customer PII data of new customers on an ongoing basis. Your Data Vigilance officers need Administrator access to the Storage bucket and BigQuery dataset. You want to follow Google recommended practices. What should you do?

- Create 3 custom IAM roles with appropriate permissions for the access levels needed for Cloud Storage and BigQuery. Add your users to the appropriate roles.
- At the Project level, add your Data Vigilance officers user accounts to the Owner role, add your operations team user accounts to the Editor role, and add your analytics team user accounts to the Viewer role.
- At the Organization level, add your Data Vigilance officers user accounts to the Owner role, add your operations team user accounts to the Editor role, and add your analytics team user accounts to the Viewer role.
- Use the appropriate predefined IAM roles for each of the access levels needed for Cloud Storage and BigQuery. Add your users to those roles for each of the services.

Unattempted

At the Organization level, add your Data Vigilance officers user accounts to the Owner role, add your operations team user accounts to the Editor role, and add your analytics team user accounts to the Viewer role. is not right.

Google recommends we apply the security principle of least privilege, where we grant only necessary permissions to access specific resources.

Ref: <https://cloud.google.com/iam/docs/overview>

Providing these primitive roles at the organization levels grants them permissions on all resources in all projects under the organization which violates the security principle of least privilege.

Ref: <https://cloud.google.com/iam/docs/understanding-roles>

At the Project level, add your Data Vigilance officers user accounts to the Owner role, add your operations team user accounts to the Editor role, and add your analytics team user accounts to the Viewer role. is not right.

Google recommends we apply the security principle of least privilege, where we grant only necessary permissions to access specific resources.

Ref: <https://cloud.google.com/iam/docs/overview>

Providing these primitive roles at the project level grants them permissions on all resources in the project

which violates the security principle of least privilege.

Ref: <https://cloud.google.com/iam/docs/understanding-roles>

Create 3 custom IAM roles with appropriate permissions for the access levels needed for Cloud Storage and BigQuery. Add your users to the appropriate roles. is not right.

While this has the intended outcome, it is not very efficient particularly when there are predefined roles that can be used. Secondly, if Google adds/modifies permissions for these services in the future, we would have to update our roles to reflect the modifications. This results in operational overhead and increases costs.

Ref: <https://cloud.google.com/storage/docs/access-control/iam-roles#primitive-roles-intrinsic>

Ref: <https://cloud.google.com/bigquery/docs/access-control>

Use the appropriate predefined IAM roles for each of the access levels needed for Cloud Storage and BigQuery. Add your users to those roles for each of the services. is the right answer.

For Google Cloud Storage service, Google provides predefined roles roles/owner, roles/editor, roles/viewer that match the access levels we need.

Similarly, Google provides the roles roles/bigquery.dataViewer, roles/bigquery.dataOwner, roles/bigquery.admin that match the access levels we need.

We can assign these predefined IAM roles to the respective users. Should Google add/modify permissions for these services in the future, we don't need to modify the roles above as Google does this for us; and this helps future proof our solution.

Ref: <https://cloud.google.com/storage/docs/access-control/iam-roles#primitive-roles-intrinsic>

Ref: <https://cloud.google.com/bigquery/docs/access-control>

#### 46. 46. Question

Your company stores sensitive PII data in a cloud storage bucket. The objects are currently encrypted by Google-managed keys. Your compliance department has asked you to ensure all current and future objects in this bucket are encrypted by customer-managed encryption keys. You want to minimize effort. What should you do?

- 1. In the bucket advanced settings, select the Customer-managed key and then select a Cloud KMS encryption key. 2. Rewrite all existing objects using gsutil rewrite to encrypt them with the new Customer-managed key.
- 1. In the bucket advanced settings, select the customer-supplied key and then select a Cloud KMS encryption key. 2. Delete all existing objects and upload them again so they use the new customer-supplied key for encryption.
- 1. Rewrite all existing objects using gsutil rewrite to encrypt them with the new Customer-managed key. 2. In the bucket advanced settings, select the Customer-managed key and then select a Cloud KMS encryption key.
- 1. In the bucket advanced settings, select the Customer-managed key and then select a Cloud KMS encryption key. 2. Existing objects encrypted by Google-managed keys can still be decrypted by the new Customer-managed key.

Unattempted

1. In the bucket advanced settings, select the Customer-managed key and then select a Cloud KMS encryption key.

2. Existing objects encrypted by Google-managed keys can still be decrypted by the new Customer-managed key. is not right.

While changing the bucket encryption to use the Customer-managed key ensures all new objects use this key, existing objects are still encrypted by the Google-managed key. This doesn't satisfy our compliance requirements. Moreover, the customer managed key can't decrypt objects created by Google-managed keys.

Ref: <https://cloud.google.com/storage/docs/encryption/using-customer-managed-keys#add-default-key>

1. In the bucket advanced settings, select the customer-supplied key and then select a Cloud KMS encryption key.

2. Delete all existing objects and upload them again so they use the new customer-supplied key for encryption. is not right.

The customer-supplied key is not an option when selecting the encryption method in the console. Moreover,

we want to use customer-managed encryption keys and not customer-supplied encryption keys. This does not fit our requirements.

1. Rewrite all existing objects using gsutil rewrite to encrypt them with the new Customer-managed key.
2. In the bucket advanced settings, select the Customer-managed key and then select a Cloud KMS encryption key. is not right.

While changing the bucket encryption to use the Customer-managed key ensures all new objects use this key, rewriting existing objects before changing the bucket encryption would result in the objects being encrypted by the encryption method in use at that point – which is still Google-managed.

1. In the bucket advanced settings, select the Customer-managed key and then select a Cloud KMS encryption key.
2. Rewrite all existing objects using gsutil rewrite to encrypt them with the new Customer-managed key. is the right answer.

Changing the bucket encryption to use the Customer-managed key ensures all new objects use this key. Now that bucket encryption is changed to use the Customer-managed key, rewrite all existing objects using gsutil rewrite results in objects being encrypted by the new Customer-managed key. This is the only option that satisfies our requirements.

Ref: <https://cloud.google.com/storage/docs/encryption/using-customer-managed-keys#add-default-key>

#### 47. Question

Your company uses a legacy application that still relies on the legacy LDAP protocol to authenticate. Your company plans to migrate this application to cloud and is looking for a cost effective solution while minimizing any developer effort. What should you do?

- Modify the legacy application to use SAML and ask users to sign in through Gmail.
- Modify the legacy application to use OAuth 2.0 and ask users to sign in through Gmail.
- Use secure LDAP to authenticate the legacy application and ask users to sign in through Gmail.
- Synchronize data within your LDAP server with Google Cloud Directory Sync.

Unattempted

Modify the legacy application to use SAML and ask users to sign in through Gmail. is not right.  
Modifying a legacy application to use SAML can be quite challenging. In any case, this is a time consuming and error-prone task and is very expensive.

Modify the legacy application to use OAuth 2.0 and ask users to sign in through Gmail. is not right.  
Modifying a legacy application to use OAuth 2.0 can be quite challenging. In any case, this is a time consuming and error-prone task and is very expensive.

Synchronize data within your LDAP server with Google Cloud Directory Sync. is not right.  
This can be done but this isn't going to help with the legacy LDAP protocol authentication unless the application is modified to work with either Cloud Identity or GSuite. And your company is looking for a cost-effective solution while minimizing developer effort so this isn't suitable.

Use secure LDAP to authenticate the legacy application and ask users to sign in through Gmail. is the right answer.

Secure LDAP enables authentication, authorization, and user/group lookups for LDAP-based apps and IT infrastructure. Secure LDAP uses the same user directory for both SaaS and LDAP-based applications, so people can use the same Cloud Identity credentials they use to log in to services like G Suite and other SaaS apps as they do to log into traditional applications. Applications and IT infrastructure that use LDAP

can be simply configured to leverage Cloud Identity's secure LDAP service instead of an existing legacy identity system—end-users don't have to change how they access their apps.

Ref: <https://cloud.google.com/blog/products/identity-security/cloud-identity-now-provides-access-to-traditional-apps-with-secure-ldap>

#### 48. 48. Question

Your company uses BigQuery for data warehousing. Over time, many different business units in your company have created 1000+ datasets across hundreds of projects. Your CIO wants you to examine all datasets to find tables that contain an employee\_ssn column. You want to minimize effort in performing this task. What should you do?

- Go to Data Catalog and search for employee\_ssn in the search box.
- Write a shell script that uses the bq command line tool to loop through all the projects in your organization.
- Write a script that loops through all the projects in your organization and runs a query on INFORMATION\_SCHEMA.COLUMNS view to find the employee\_ssn column.
- Write a Cloud Dataflow job that loops through all the projects in your organization and runs a query on INFORMATION\_SCHEMA.COLUMNS view to find employee\_ssn column.

Unattempted

Go to Data Catalog and search for employee\_ssn in the search box. is the right answer.

Data Catalog is a fully managed and scalable metadata management service that empowers organizations to quickly discover, understand, and manage all their data. It offers a simple and easy-to-use search interface for data discovery, a flexible and powerful cataloging system for capturing both technical and business metadata, and a strong security and compliance foundation with Cloud Data Loss Prevention (DLP) and Cloud Identity and Access Management (IAM) integrations. The service automatically ingests technical metadata for BigQuery and Cloud Pub/Sub and allows customers to capture business metadata in schematized format via tags, custom APIs, and the UI, offering a simple and efficient way to catalog their data assets. You can perform a search for data assets from the Data Catalog home page in the Google Cloud Console.

See <https://cloud.google.com/data-catalog/docs/how-to/search> for example.

All other options are manual, error-prone, time-consuming, and should be avoided.

#### 49. 49. Question

Your company uses Cloud Storage to store application backup files for disaster recovery purposes. You want to follow Google recommended practices. Which storage option should you use?

- Coldline Storage
- Multi-Regional Storage
- Regional Storage
- Nearline Storage

Unattempted

The ideal answer to this would have been Archive Storage but that is not one of the options.

Archive Storage is the lowest-cost, highly durable storage service for data archiving, online backup, and disaster recovery. Your data is available within milliseconds, not hours or days.

<https://cloud.google.com/storage/docs/storage-classes#archive>

In the absence of Archive Storage, the next best option for storing backups is Coldline Storage. Coldline Storage is a very-low-cost, highly durable storage service for storing infrequently accessed data. Coldline Storage is a better choice than Standard Storage or Nearline Storage in scenarios where slightly lower availability, a 90-day minimum storage duration, and higher costs for data access are acceptable trade-offs for lowered at-rest storage costs. Coldline Storage is ideal for data you plan to read or modify at most once a quarter.

Ref: <https://cloud.google.com/storage/docs/storage-classes#coldline>

Although Nearline, Regional and Multi-Regional can also be used to store the backups, they are expensive in comparison and Google recommends we use Coldline for backups.

More information about Nearline: <https://cloud.google.com/storage/docs/storage-classes#nearline>

More information about Standard/Regional: <https://cloud.google.com/storage/docs/storage-classes#standard>

More information about Standard/Multi-Regional: <https://cloud.google.com/storage/docs/storage-classes#standard>

## 50. 50. Question

Your company wants to move 200 TB of your website clickstream logs from your on-premise data center to Google Cloud Platform. These logs need to be retained in GCP for compliance requirements. Your business analysts also want to run analytics on these logs to understand user click behavior on your website. Which of the below would enable you to meet these requirements? (Select Two)

- Load logs into Google BigQuery.
- Load logs into Google Cloud SQL.
- Import logs into Google Stackdriver.
- Insert logs into Google Cloud Bigtable.
- Upload log files into Google Cloud Storage.

Unattempted

Load logs into Google Cloud SQL. is not right.

Cloud SQL is a fully-managed relational database service. Storing logs in Google Cloud SQL is very expensive. Cloud SQL doesn't help us with analytics. Moreover, Google Cloud Platform offers several storage classes in Google Cloud Storage that are more apt for storing logs at a much cheaper cost.

Ref: <https://cloud.google.com/sql/docs>

Ref: <https://cloud.google.com/sql/pricing#sql-storage-networking-prices>

Ref: <https://cloud.google.com/storage/pricing>

Import logs into Google Stackdriver. is not right.

You can push custom logs to Stackdriver and set custom retention periods to store the logs for longer durations. However, Stackdriver doesn't help us with analytics. You could create a sink and export data into Cloud BigQuery for analytics but that is more work. Moreover, Google Cloud Platform offers several storage classes in Google Cloud Storage that are more apt for storing logs at a much cheaper cost.

Ref: <https://cloud.google.com/logging>

Ref: <https://cloud.google.com/storage/pricing>

Insert logs into Google Cloud Bigtable. is not right.

Cloud Bigtable is a petabyte-scale, fully managed NoSQL database service for large analytical and operational workloads. Cloud Bigtable can not run analytics by itself. (But when combined with other services to ingest, process, analyze and present, it can help drive analytics) – see the diagram below. So this option is not right.

Ref: <https://cloud.google.com/bigtable/>

Upload log files into Google Cloud Storage. is the right answer.

Google Cloud Platform offers several storage classes in Google Cloud Storage that are suitable for storing/archiving logs at a reasonable cost.

GCP recommends you use

1. Standard storage class if you need to access objects frequently
2. Nearline storage class if you access infrequently i.e. once a month
3. Coldline storage class if you access even less frequently e.g. once a quarter

4. Archive storage for logs archival.

Ref: <https://cloud.google.com/storage/docs/storage-classes>

Load logs into Google BigQuery. is the right answer.

By loading logs into Google BigQuery, you can securely run and share analytical insights in your organization with a few clicks. BigQuery's high-speed streaming insertion API provides a powerful foundation for real-time analytics, making your latest business data immediately available for analysis.

Ref: <https://cloud.google.com/bigquery#marketing-analytics>

51. 51. Question

Your company wants to move all documents from a secure internal NAS drive to a Google Cloud Storage (GCS) bucket. The data contains personally identifiable information (PII) and sensitive customer information. Your company tax auditors need access to some of these documents. What security strategy would you recommend on GCS?

- Grant no Google Cloud Identity and Access Management (Cloud IAM) roles to users, and use granular ACLs on the bucket.
- Grant IAM read-only access to users, and use default ACLs on the bucket.
- Create randomized bucket and object names. Enable public access, but only provide specific file URLs to people who do not have Google accounts and need access.
- Use signed URLs to generate time-bound access to objects.

Unattempted

Use signed URLs to generate time-bound access to objects. is not right.

When dealing with sensitive customer information such as PII, using signed URLs is not a great idea as anyone with access to the URL has access to PII data. Signed URLs provide time-limited resource access to anyone in possession of the URL, regardless of whether they have a Google account. With PII Data, we want to be sure who has access and signed URLs don't guarantee that.

Ref: <https://cloud.google.com/storage/docs/access-control/signed-urls>

Grant IAM read-only access to users, and use default ACLs on the bucket. is not right.

We do not need to grant all IAM read-only access to this sensitive data. Just the users who need access to sensitive/PII data should be provided access to this data.

Create randomized bucket and object names. Enable public access, but only provide specific file URLs to people who do not have Google accounts and need access. is not right.

Enabling public access to the buckets and objects makes them visible to everyone. There are a number of scanning tools out in the market with the sole purpose of identifying buckets/objects that can be reached publicly. Should one of these tools be used by a bad actor to find out our public bucket/objects, it would result in a security breach.

Grant no Google Cloud Identity and Access Management (Cloud IAM) roles to users, and use granular ACLs on the bucket. is the right answer.

We start with no explicit access to any of the IAM users, and the bucket ACLs can then control which users can access what objects. This is the most secure way of ensuring just the people who require access to the bucket are provided with access. We block everyone from accessing the bucket and explicitly provided access to specific users through ACLs.

52. 52. Question

Your company's infrastructure is on-premises, but all machines are running at maximum capacity. You want to burst to Google Cloud. The workloads on Google Cloud must be able to directly communicate to the workloads on-premises using a private IP range. What should you do?

- In Google Cloud, configure the VPC as a host for Shared VPC.
- In Google Cloud, configure the VPC for VPC Network Peering.
- Create bastion hosts both in your on-premises environment and on Google Cloud. Configure both as proxy servers using their public IP addresses.
- Set up Cloud VPN between the infrastructure on-premises and Google Cloud.

Unattempted

In Google Cloud, configure the VPC as a host for Shared VPC. is not right.

Shared VPC allows an organization to connect resources from multiple projects to a common Virtual Private Cloud (VPC) network so that they can communicate with each other securely and efficiently using internal IPs from that network. When you use Shared VPC, you designate a project as a host project and attach one or more other service projects to it. This in no way helps us connect to our on-premises network.

Ref: <https://cloud.google.com/vpc/docs/shared-vpc>

In Google Cloud, configure the VPC for VPC Network Peering. is not right.

Google Cloud VPC Network Peering allows internal IP address connectivity across two Virtual Private Cloud (VPC) networks regardless of whether they belong to the same project or the same organization. VPC Network Peering enables you to connect VPC networks so that workloads in different VPC networks can communicate internally. Traffic stays within Google's network and doesn't traverse the public internet. This doesn't help us connect to our on-premises network.

Ref: <https://cloud.google.com/vpc/docs/vpc-peering>

Create bastion hosts both in your on-premises environment and on Google Cloud. Configure both as proxy servers using their public IP addresses. is not right.

Bastion hosts provide an external facing point of entry into a network containing private network instances. Bastion hosts are primarily for end users so they can connect to an instance that does not have an external IP address through a bastion host.

Ref: <https://cloud.google.com/compute/docs/instances/connecting-advanced>

Set up Cloud VPN between the infrastructure on-premises and Google Cloud. is the right answer.

Cloud VPN securely connects your on-premises network to your Google Cloud (GCP) Virtual Private Cloud (VPC) network through an IPsec VPN connection.

Ref: <https://cloud.google.com/vpn/docs/concepts/overview>

### 53. Question

Your company's test suite is a custom C++ application that runs tests each day on Linux virtual machines. The full test suite takes several hours to complete, running on a limited number of on-premises servers reserved for testing. Your company wants to move the testing infrastructure to Google Cloud Platform. Your company wants to reduce the amount of time it takes to fully test a change to the system while changing the tests as little as possible. Your project manager has asked you to suggest suitable services in Google Cloud and you want to follow Google recommended practices. What should you do?

- Use Google App Engine and Google Stackdriver for logging.
- Use Google Compute Engine unmanaged instance groups with a Network Load Balancer.
- Use Google Compute Engine managed instance groups and autoscaling.
- Use Google Cloud Dataproc and run Apache Hadoop jobs to process each test.

Unattempted

Use Google Compute Engine unmanaged instance groups with a Network Load Balancer. is not right.

An unmanaged instance group is a collection of virtual machines (VMs) that reside in a single zone, VPC network, and subnet. An unmanaged instance group is useful for grouping together VMs that require individual configuration settings or tuning. Unmanaged instance group does not autoscale, so it does not

help reduce the amount of time it takes to fully test a change to the system.

Ref: <https://cloud.google.com/compute/docs/instance-groups/creating-groups-of-unmanaged-instances>

Use Google App Engine and Google Stackdriver for logging. is not right.

App Engine supports many popular languages like Java, PHP, Node.js, Python, C#, .Net, Ruby, and Go. However, C++ isn't supported by App Engine.

Ref: <https://cloud.google.com/appengine>

Use Google Cloud Dataproc and run Apache Hadoop jobs to process each test. is not right.

Cloud Dataproc is a fast, easy-to-use, fully managed cloud service for running Apache Spark and Apache Hadoop clusters in a simpler, more cost-efficient way. While Dataproc is very efficient at processing ETL and Big Data pipelines, it is not as suitable for running a ruby application that runs tests each day.

Ref: <https://cloud.google.com/dataproc>

Use Google Compute Engine managed instance groups and autoscaling. is the right answer.

A managed instance group (MIG) contains identical virtual machine (VM) instances that are based on an instance template. MIGs support auto-healing, load balancing, autoscaling, and auto-updating. Managed instance groups offer auto-scaling capabilities that let you automatically add or delete instances from a managed instance group based on increases or decreases in load. Autoscaling helps your apps gracefully handle increases in traffic and reduce costs when the need for resources is lower. Autoscaling works by adding more instances to your instance group when there is more load (upscale), and deleting instances when the need for instances is lowered (downscale).

Ref: <https://cloud.google.com/compute/docs/autoscaler/>

#### 54. Question

Your compliance team requested all audit logs are stored for 10 years and to allow access for external auditors to view. You want to follow Google recommended practices. What should you do? (Choose two)

- Create an account for auditors to have view access to Stackdriver Logging.
- Export audit logs to Cloud Storage via an export sink.
- Export audit logs to BigQuery via an export sink.
- Generate a signed URL to the Stackdriver export destination for auditors to access.
- Export audit logs to Splunk via a Pub/Sub export sink.

Unattempted

Create an account for auditors to have view access to Stackdriver Logging. is not right.

While it is possible to configure a custom retention period of 10 years in Stackdriver logging, storing logs in Stackdriver is expensive compared to Cloud Storage. Stackdriver charges \$0.01 per GB per month, whereas something like Cloud Storage Coldline Storage costs \$0.007 per GB per month (30% cheaper) and Cloud Storage Archive Storage costs 0.004 per GB per month (60% cheaper than Stackdriver)

Ref: <https://cloud.google.com/logging/docs/storage#pricing>

Ref: <https://cloud.google.com/storage/pricing>

Export audit logs to BigQuery via an export sink. is not right.

Storing logs in BigQuery is expensive. In BigQuery, Active storage costs \$0.02 per GB per month and Long-term storage costs \$0.01 per GB per month. In comparison, Google Cloud Storage offers several storage classes that are significantly cheaper.

Ref: <https://cloud.google.com/bigquery/pricing>

Ref: <https://cloud.google.com/storage/pricing>

Export audit logs to Cloud Filestore via a Pub/Sub export sink. is not right.

Storing logs in Cloud Filestore is expensive. In Cloud Filestore, Standard Tier pricing costs \$0.2 per GB per month and Premium Tier pricing costs \$0.3 per GB per month. In comparison, Google Cloud Storage offers several storage classes that are significantly cheaper.

Ref: <https://cloud.google.com/bigquery/pricing>

Ref: <https://cloud.google.com/storage/pricing>

Export audit logs to Cloud Storage via an export sink. is the right answer.

Among all the storage solutions offered by Google Cloud Platform, Cloud storage offers the best pricing for long term storage of logs. Google Cloud Storage offers several storage classes such as Nearline Storage (\$0.01 per GB per Month) Coldline Storage (\$0.007 per GB per Month) and Archive Storage (\$0.004 per GB

per month) which are significantly cheaper than the storage options covered by the above options above.

Ref: <https://cloud.google.com/storage/pricing>

Generate a signed URL to the Stackdriver export destination for auditors to access. is the right answer.

In Google Cloud Storage, you can generate a signed URL to provide limited permission and time to make a request. Anyone who possesses it can use the signed URL to perform specified actions, such as reading an object, within a specified period of time.

In our scenario, we do not need to create accounts for our auditors to provide access to logs in Cloud Storage. Instead, we can generate them signed URLs which are time-bound and lets them access/download log files.

Ref: <https://cloud.google.com/storage/docs/access-control/signed-urls>

## 55. 55. Question

Your customer has implemented a solution that uses Cloud Spanner and notices some read latency-related performance issues on one table. This table is accessed only by their users using a primary key. The table schema is shown below.

```
CREATE TABLE Persons {
 person_id INT64 NOT NULL, // sequential number based on number of registrations
 account_creation_date DATE, // system date
 birthdate DATE, // customer birthdate
 firstname STRING (255), // first name
 lastname STRING (255), // last name
 profile_picture BYTES (255) // profile picture
} PRIMARY KEY (person_id)
```

You want to resolve the issue. What should you do?

- Remove the profile\_picture field from the table.
- Add a secondary index on the person\_id column.
- Change the primary key to not have monotonically increasing values.
- Create a secondary index using the following Data Definition Language (DDL):
- CREATE INDEX person\_id\_ix ON Persons ( person\_id, firstname, lastname ) STORING ( profile\_picture )

Unattempted

Change the primary key to not have monotonically increasing values. is the right answer.

You should be careful when choosing a primary key to not accidentally create hotspots in your database.

One cause of hotspots is having a column whose value monotonically increases as the first key part because this results in all inserts occurring at the end of your keyspace. This pattern is undesirable because Cloud Spanner divides data among servers by key ranges, which means all your inserts will be directed at a single server that will end up doing all the work.

Ref: <https://cloud.google.com/spanner/docs/schema-design#primary-key-prevent-hotspots>

All other options make no sense. The problem is with the monotonically increasing values in the primary key and removing profile\_picture or adding a secondary index isn't going to alleviate the problem.

## 56. 56. Question

Your development team needs a new Jenkins server for their project. You need to deploy the server using the fewest steps possible. What should you do?

- Create a new Compute Engine instance and install Jenkins through the command-line interface.
- Download and deploy the Jenkins Java WAR to App Engine Standard.
- Use GCP Marketplace to launch the Jenkins solution.

- Create a Kubernetes cluster on Compute Engine and create a deployment with the Jenkins Docker image.

Unattempted

Create a new Compute Engine instance and install Jenkins through the command line interface. is not right.  
While this can be done, this involves a lot more work than installing the Jenkins server through App Engine.

Create a Kubernetes cluster on Compute Engine and create a deployment with the Jenkins Docker image. is not right.

While this can be done, this involves a lot more work than installing the Jenkins server through App Engine.

Download and deploy the Jenkins Java WAR to App Engine Standard. is not right.

While this is possible, we need to ensure App Engine is enabled, we then need to download the Java project/WAR, and run gcloud app deploy to set up a Jenkins server. This involves more steps than spinning up an instance from GCP Marketplace.

Ref: <https://cloud.google.com/appengine/docs/standard/java/tools/uploadinganapp>

Ref: <https://cloud.google.com/solutions/using-jenkins-for-distributed-builds-on-compute-engine>

Use GCP Marketplace to launch the Jenkins solution. is the right answer.

The simplest way to launch a Jenkins server is from GCP Market place. GCP market place has a number of builds available for Jenkins: <https://console.cloud.google.com/marketplace/browse?q=jenkins>. All you need to do is spin up an instance from a suitable market place build and you have a Jenkins server in a few minutes with just a few clicks.

## 57. Question

Your finance team wants to view the billing report for your projects. You want to make sure that the finance team does not get additional permissions to the project. What should you do?

- Add the group for the finance team to roles/billing.user role.
- Add the group for the finance team to roles/billing.admin role.
- Add the group for the finance team to roles/billing.viewer role.
- Add the group for the finance team to roles/billing.projectManager role.

Unattempted

Add the group for the finance team to roles/billing.user role. is not right.

This role has very restricted permissions, so you can grant it broadly, typically in combination with Project Creator. These two roles allow a user to create new projects linked to the billing account on which the role is granted.

Ref: <https://cloud.google.com/billing/docs/how-to/billing-access>

Add the group for the finance team to roles/billing.admin role. is not right.

This role is an owner role for a billing account. Use it to manage payment instruments, configure billing exports, view cost information, link and unlink projects, and manage other user roles on the billing account.

Ref: <https://cloud.google.com/billing/docs/how-to/billing-access>

Add the group for the finance team to roles/billing.projectManager role. is not right.

This role allows a user to attach the project to the billing account, but does not grant any rights over resources. Project Owners can use this role to allow someone else to manage the billing for the project without granting them resource access.

Ref: <https://cloud.google.com/billing/docs/how-to/billing-access>

Add the group for the finance team to roles/billing.viewer role. is the right answer.

Billing Account Viewer access would usually be granted to finance teams, it provides access to spending information but does not confer the right to link or unlink projects or otherwise manage the properties of the

billing account.

Ref: <https://cloud.google.com/billing/docs/how-to/billing-access>

## 58. Question

Your managed instance group raised an alert stating that new instance creation has failed to create new instances. You need to maintain the number of running instances specified by the template to be able to process expected application traffic. What should you do?

- Create an instance template that contains valid syntax that will be used by the instance group. Delete any persistent disks with the same name as instance names.
- Create an instance template that contains valid syntax that will be used by the instance group. Verify that the instance name and persistent disk name values are not the same in the template.
- Verify that the instance template being used by the instance group contains valid syntax. Delete any persistent disks with the same name as instance names. Set the disks.autoDelete property to true in the instance template.
- Delete the current instance template and replace it with a new instance template. Verify that the instance name and persistent disk name values are not the same in the template. Set the disks.autoDelete property to true in the instance template.

Unattempted

Verify that the instance template being used by the instance group contains valid syntax. Delete any persistent disks with the same name as instance names. Set the disks.autoDelete property to true in the instance template. is the right answer.

As described in this article, "My managed instance group keeps failing to create a VM. What's going on?"  
<https://cloud.google.com/compute/docs/instance-groups/creating-groups-of-managed-instances#troubleshooting>

The likely causes are

1. A persistent disk already exists with the same name as VM Instance
2. disks.autoDelete option is set to false
3. instance template might be invalid

Therefore, we need to ensure that instance template is valid, disks.autoDelete is turned on, and that there are no existing persistent disks with the same name as VM instance.

## 59. Question

Your management has asked an external auditor to review all the resources in a specific project. The security team has enabled the Organization Policy called Domain Restricted Sharing on the organization node by specifying only your Cloud Identity domain. You want the auditor to only be able to view, but not modify, the resources in that project. What should you do?

- Ask the auditor for their Google account, and give them the Viewer role on the project.
- Ask the auditor for their Google account, and give them the Security Reviewer role on the project.
- Create a temporary account for the auditor in Cloud Identity, and give that account the Viewer role on the project.
- Create a temporary account for the auditor in Cloud Identity, and give that account the Security Reviewer role on the project.

Unattempted

Ask the auditor for their Google account, and give them the Viewer role on the project. is not right.  
Since the auditor's account is not part of your company's Cloud Identity domain, the auditor can not access resources from your GCP projects. Domain restriction constraint can be used in organization policies to limit resource sharing based on domain. This constraint allows you to restrict the set of identities that are allowed to be used in Cloud Identity and Access Management policies. In this scenario, since we are restricting based on the Cloud Identity domain, only the users in the Cloud Identity domain can access GCP services.

<https://cloud.google.com/resource-manager/docs/organization-policy/restricting-domains>

Ask the auditor for their Google account, and give them the Security Reviewer role on the project. is not right.

Since the auditor's account is not part of your company's Cloud Identity domain, the auditor can not access resources from your GCP projects. Domain restriction constraint can be used in organization policies to limit resource sharing based on domain. This constraint allows you to restrict the set of identities that are allowed to be used in Cloud Identity and Access Management policies. In this scenario, since we are restricting based on the Cloud Identity domain, only the users in the Cloud Identity domain can access GCP services.

<https://cloud.google.com/resource-manager/docs/organization-policy/restricting-domains>

Create a temporary account for the auditor in Cloud Identity, and give that account the Security Reviewer role on the project. is not right.

Creating a temporary account for the auditor in your cloud identity is the right approach as this makes the auditor part of the Cloud identity domain and the organization policy in place lets the auditor access resources. However, the role granted here is not suitable, it provides permissions to list all resources and Cloud IAM policies. Note that list permissions only allow you to list but not view resources. You need to get permission to view the resources.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#iam-roles>

Create a temporary account for the auditor in Cloud Identity, and give that account the Viewer role on the project. is the right answer.

The primitive viewer role provides permissions for read-only actions that do not affect the state, such as viewing (but not modifying) existing resources or data. This fits our requirements.

In addition, adding the auditor to Cloud Identity ensures that Organization Policy for Domain Restricted Sharing doesn't block them from accessing resources.

Ref: [https://cloud.google.com/iam/docs/understanding-roles#primitive\\_role\\_definitions](https://cloud.google.com/iam/docs/understanding-roles#primitive_role_definitions)

## 60. Question

Your networks team has set up Google compute network as shown below. In addition, firewall rules in the VPC network have been configured to allow egress to 0.0.0.0/0

Larger image

Which instances have access to Google APIs and Services such as Google Cloud Storage?

- VM A1, VM A2, VM B1
- VM A1, VM A2, VM B1, VM B2
- VM A1, VM A2
- VM A1, VM A2, VM B2

Unattempted

VM A1 can access Google APIs and services, including Cloud Storage because its network interface is located in subnet-a, which has Private Google Access enabled. Private Google Access applies to the instance because it only has an internal IP address.

VM B1 cannot access Google APIs and services because it only has an internal IP address and Private Google Access is disabled for subnet-b.

VM A2 and VM B2 can both access Google APIs and services, including Cloud Storage, because they each have external IP addresses. Private Google Access has no effect on whether or not these instances can access Google APIs and services because both have external IP addresses.

So the correct answer is VM A1, VM A2, VM B2

Ref: <https://cloud.google.com/vpc/docs/private-access-options#example>

## 61. Question

Your operations team has configured a lifecycle management rule on a bucket. The bucket is multi-regional and has versioning enabled. Which of the following statement accurately reflects the following lifecycle config?

```
{
 "rule": [
 {
 "action": {
 "type": "Delete"
 },
 "condition": {
 "age": 60,
 "isLive": false
 }
 },
 {
 "action": {
 "type": "SetStorageClass",
 "storageClass": "COLDLINE"
 },
 "condition": {
 "age": 366,
 "matchesStorageClass": "MULTI_REGIONAL"
 }
 }
]
}
```

- Move objects to Coldline Storage after 366 days if the storage class in Multi-regional First rule has no effect on the bucket.
- Archive objects older than 60 days and move objects to Coldline Storage after 366 days if the storage class in Multi-regional.
- Delete objects older than 60 days and move objects to Coldline Storage after 366 days if the storage class in Multi-regional.
- Delete archived objects older than 60 days and move objects to Coldline Storage after 366 days if the storage class is Multi-regional.

Unattempted

Archive objects older than 60 days and move objects to Coldline Storage after 366 days if the storage class in Multi-regional. is not right.

The action has "type": "Delete" which means we want to Delete, not archive.

Ref: <https://cloud.google.com/storage/docs/managing-lifecycles>

Delete objects older than 60 days and move objects to Coldline Storage after 366 days if the storage class in Multi-regional. is not right.

We want to delete objects as indicated by the action, however, we don't want to delete all objects older than

60 days. We only want to delete archived objects as indicated by "isLive":false condition

Ref: <https://cloud.google.com/storage/docs/managing-lifecycles>

Move objects to Coldline Storage after 366 days if the storage class is Multi-regional. First rule has no effect on the bucket. is not right.

The first rule certainly has an effect. It deletes archived objects older than 60 days.

Delete archived objects older than 60 days and move objects to Coldline Storage after 366 days if the storage class is Multi-regional. is the right answer.

The first part of the rule: The action has "type":"Delete" which means we want to Delete. "isLive":false condition means we are looking for objects that are not Live i.e. objects that are archived. Together, it means we want to delete archived objects older than 60 days. Note that if an object is deleted, it cannot be undeleted. Take care in setting up your lifecycle rules so that you do not cause more data to be deleted than you intend.

Ref: <https://cloud.google.com/storage/docs/managing-lifecycles>

The second part of the rule: The action indicates we want to set storage class to Coldline. The condition is true if the existing storage class is multi-regional and the age of the object is 366 days or over. Together it means we want to set the storage class to Coldline if existing storage class is multi-regional and age of the object is 366 days or over

## 62. 62. Question

Your organization has a dedicated person who creates and manages all service accounts for Google Cloud projects. You need to assign this person the minimum role for projects. What should you do?

- Add the user to roles/iam.roleAdmin role.
- Add the user to roles/iam.securityAdmin role.
- Add the user to roles/iam.serviceAccountUser role.
- Add the user to roles/iam.serviceAccountAdmin role.

Unattempted

Add the user to roles/iam.roleAdmin role. is not right.

roles/iam.roleAdmin is an administrator role that provides access to all custom roles in the project. This doesn't include permissions needed to manage service accounts.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#roles-roles>

Add the user to roles/iam.securityAdmin role. is not right.

roles/iam.securityAdmin role is a Security admin role, with permissions to get and set any IAM policy. This role is too broad i.e. includes too many permissions and goes against the principle of least privilege.

Moreover, although this role provides iam.serviceAccounts.get/list, it doesn't provide iam.serviceAccounts.create, iam.serviceAccounts.delete and iam.serviceAccounts.update permissions that are needed for managing service accounts.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#iam-roles>

Add the user to roles/iam.serviceAccountUser role. is not right.

roles/iam.serviceAccountUser is a service Account User role which is used for running operations as the service account. This role does not provide the permissions iam.serviceAccounts.create, iam.serviceAccounts.delete, iam.serviceAccounts.update, iam.serviceAccounts.get/list which are required for managing service accounts.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#service-accounts-roles>

Add the user to roles/iam.serviceAccountAdmin role. is the right answer.

roles/iam.serviceAccountAdmin is a Service Account Admin role that lets you Create and manage service accounts. This grants all the required permissions for managing service accounts (iam.serviceAccounts.create iam.serviceAccounts.delete, iam.serviceAccounts.update, iam.serviceAccounts.get/list etc) and therefore fits our requirements.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#service-accounts-roles>

## 63. 63. Question

Your organization has strict requirements to control access to Google Cloud projects. You need to enable your Site Reliability Engineers (SREs) to approve requests from the Google Cloud support team when an SRE opens a support case. You want to follow Google-recommended practices. What should you do?

- Add your SREs to roles/iam.roleAdmin role.
- Add your SREs to roles/accessapproval.approver role.
- Add your SREs to a group and then add this group to roles/iam.roleAdmin role.
- Add your SREs to a group and then add this group to roles/accessapproval.approver role.

Unattempted

Add your SREs to roles/iam.roleAdmin role. is not right.

roles/iam.roleAdmin provides access to all custom roles in the project. This doesn't fit our requirement of SREs being able to approve requests.

Add your SREs to a group and then add this group to roles/iam.roleAdmin role. is not right.

roles/iam.roleAdmin provides access to all custom roles in the project. This doesn't fit our requirement of SREs being able to approve requests.

Add your SREs to roles/accessapproval.approver role. is not right.

roles/accessapproval.approver is an Access Approval Approver role and provides the ability to view or act on access approval requests and view configuration. Although this is the role we require, you want to follow Google recommended practices which means we should instead add the group to the role and add users to the group instead of granting the role individually to users.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#access-approval-roles>

Ref: <https://cloud.google.com/iam/docs/overview>

Add your SREs to a group and then add this group to roles/accessapproval.approver role. is the right answer.

roles/accessapproval.approver is an Access Approval Approver role and provides the ability to view or act on access approval requests and view configuration. And you follow Google recommended practices by adding users to the group and group to the role. Groups are a convenient way to apply an access policy to a collection of users. You can grant and change access controls for a whole group at once instead of granting or changing access controls one at a time for individual users or service accounts. You can also easily add members to and remove members from a Google group instead of updating a Cloud IAM policy to add or remove users.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#access-approval-roles>

Ref: <https://cloud.google.com/iam/docs/overview>

#### 64. Question

Your organization has user identities in Active Directory. Your organization wants to use Active Directory as their source of truth for identities. Your organization wants to have full control over the Google accounts used by employees for all Google services, including your Google Cloud Platform (GCP) organization. What should you do?

- Use Google Cloud Directory Sync (GCDS) to synchronize users into Cloud Identity.
- Use the cloud Identity APIs and write a script to synchronize users to Cloud Identity.
- Export users from Active Directory as a CSV and import them to Cloud Identity via the Admin Console.
- Ask each employee to create a Google account using self signup. Require that each employee use their company email address and password.

Unattempted

Use the cloud Identity APIs and write a script to synchronize users to Cloud Identity. is not right.  
You could do this, but this process is manual, error-prone, time-consuming, and should be avoided especially when there is a service/tool that does it out of the box with minimal configuration.

Export users from Active Directory as a CSV and import them to Cloud Identity via the Admin Console. is not right.

You could do this, but like above this process is manual, error-prone, time-consuming, and should be avoided especially when there is a service/tool that does it out of the box with minimal configuration.

Ask each employee to create a Google account using self signup. Require that each employee use their company email address and password. is not right.

If you let employees create accounts, your organization no longer has full control over the Google accounts used. This approach has several other issues with respect to creating/managing user accounts and should be avoided.

Use Google Cloud Directory Sync (GCDS) to synchronize users into Cloud Identity. is the right answer.  
Since we already have user identities in Active Directory, it makes sense to reuse this directory as the source of truth for identities. But for GCP, you need identities either in G Suite or Google Cloud Identity. Cloud Directory Sync is a tool that enables you to synchronize users, groups, and other data from an Active Directory/LDAP service to their Google Cloud domain directory. This performs a one-way synchronization and ensures Cloud Identity users match that of your Active Directory. This also helps with our requirement of the organization having full control over the accounts used by employees.

Ref: <https://tools.google.com/dlpage/dirsync/>

Ref: [https://support.google.com/a/answer/106368?hl=en#:~:text=With%20Google%20Cloud%20Directory%20Sync,files\)](https://support.google.com/a/answer/106368?hl=en#:~:text=With%20Google%20Cloud%20Directory%20Sync,files)) to your Google Account.

## 65. Question

Your organization is a financial company that needs to store audit log files for 3 years. Your organization has hundreds of Google Cloud projects. You need to implement a cost-effective approach for log file retention. What should you do?

- Create an export to the sink that saves logs from Cloud Audit to a Coldline Storage bucket.
- Write a custom script that uses logging API to copy the logs from Stackdriver logs to BigQuery.
- Export these logs to Cloud Pub/Sub and write a Cloud Dataflow pipeline to store logs to Cloud SQL.
- Create an export to the sink that saves logs from Cloud Audit to BigQuery.

Unattempted

Create an export to the sink that saves logs from Cloud Audit to BigQuery. is not right.

You can export logs into BigQuery by creating one or more sinks that include a logs query and an export destination (big query). However, this option is very expensive compared to the cost of Cloud Storage.

Ref: [https://cloud.google.com/logging/docs/export/configure\\_export\\_v2](https://cloud.google.com/logging/docs/export/configure_export_v2)

Write a custom script that uses logging API to copy the logs from Stackdriver logs to BigQuery. is not right.  
Stackdriver already offers sink exports that let you copy logs from Stackdriver logs to BigQuery. While BigQuery is already quite expensive compared to Cloud Storage, coming up with a custom script and maintaining it to copy the logs from Stackdriver logs to BigQuery is going to add to the cost. This option is very inefficient and expensive.

Export these logs to Cloud Pub/Sub and write a Cloud Dataflow pipeline to store logs to Cloud SQL. is not right.

Cloud SQL is primarily used for storing relational data. Storing huge quantities of logs in Cloud SQL is very expensive compared to Cloud Storage. And add to it the fact that you also need to pay for Cloud Pub/Sub and Cloud Dataflow pipeline, and this option gets very expensive very soon.

Create an export to the sink that saves logs from Cloud Audit to a Coldline Storage bucket. is the right answer.

Coldline Storage is the perfect service to store audit logs from all the projects and is very cost-efficient as well. Coldline Storage is a very-low-cost, highly durable storage service for storing infrequently accessed data. Coldline Storage is a better choice than Standard Storage or Nearline Storage in scenarios where slightly lower availability, a 90-day minimum storage duration, and higher costs for data access are acceptable trade-offs for lowered at-rest storage costs. Coldline Storage is ideal for data you plan to read or modify at most once a quarter.

Ref: <https://cloud.google.com/storage/docs/storage-classes#coldline>

## 66. Question

Your organization is planning the infrastructure for a new large-scale application that will need to store anything between 200 TB to a petabyte of data in NoSQL format for Low-latency read/write and High-throughput analytics. Which storage option should you use?

- Cloud SQL.
- Cloud Bigtable.
- Cloud Datastore.
- Cloud Spanner.

Unattempted

Cloud Spanner. is not right.

Cloud Spanner is not a NoSQL database. Cloud SQL is a fully-managed relational database service.

Ref: <https://cloud.google.com/sql/docs>

Cloud SQL. is not right.

Cloud SQL is not a NoSQL database. Cloud Spanner is a highly scalable, enterprise-grade, globally-distributed, and strongly consistent relational database service

Ref: <https://cloud.google.com/spanner>

Cloud Datastore. is not right.

While Cloud Datastore is a highly scalable NoSQL database, it can't handle petabyte-scale data.

<https://cloud.google.com/datastore>

Cloud Bigtable. is the right answer.

Cloud Bigtable is a petabyte-scale, fully managed NoSQL database service for large analytical and operational workloads.

Ref: <https://cloud.google.com/bigtable/>

## 67. Question

Your organization is planning to deploy a Python web application to Google Cloud. The web application uses a custom Linux distribution and you want to minimize rework. The web application underpins an important website that is accessible to the customers globally. You have been asked to design a solution that scales to meet demand. What would you recommend to fulfill this requirement? (Select Two)

- HTTP(S) Load Balance.
- App Engine Standard environment
- Cloud Functions

- Managed Instance Group on Compute Engine
- Network Load Balance

Unattempted

Requirements are – use custom Linux distro, global access, auto scale.

Cloud Functions. is not right.

Cloud Functions is a serverless compute platform. You can not use a custom Linux distribution with Cloud Functions. Ref: <https://cloud.google.com/functions>

App Engine Standard environment. is not right.

The App Engine Standard Environment is based on container instances running on Google's infrastructure. Containers are preconfigured with one of several available runtimes such as Python, Java, NodeJS, PHP, Ruby, GO etc. It is not possible to specify a custom Linux distribution with App Engine Standard.

Ref: <https://cloud.google.com/appengine/docs/standard>

Network Load Balance. is not right.

The external (TCP/UDP) Network Load Balancing is a regional load balancer. Since we need to cater to a global user base, this load balancer is not suitable.

Ref: <https://cloud.google.com/load-balancing/docs/network>

HTTP(S) Load Balancer. is the right answer.

HTTP(S) Load Balancing is a global service (when the Premium Network Service Tier is used). We can create backend services in more than one region and have them all serviced by the same global load balancer

Ref: <https://cloud.google.com/load-balancing/docs/https>

Managed Instance Group on Compute Engine. is the right answer.

Managed instance groups (MIGs) maintain the high availability of your applications by proactively keeping your virtual machine (VM) instances available. An autohealing policy on the MIG relies on an application-based health check to verify that an application is responding as expected. If the auto-healer determines that an application isn't responding, the managed instance group automatically recreates that instance.

Ref: <https://cloud.google.com/compute/docs/instance-groups>

## 68. Question

Your organization needs to grant users access to query datasets in BigQuery but prevent them from accidentally deleting the datasets. You want a solution that follows Google-recommended practices. What should you do?

- Add users to roles/bigquery.user role only, instead of roles/bigquery.dataOwner.
- Add users to roles/bigquery.dataEditor role only, instead of roles/bigquery.dataOwner.
- Create a custom role by removing delete permissions, and add users to that role only.
- Create a custom role by removing delete permissions. Add users to the group, and then add the group to the custom role.

Unattempted

Add users to roles/bigquery.dataEditor role only, instead of roles/bigquery.dataOwner. is not right. roles/bigquery.dataEditor is a BigQuery Data Editor role which when applied to a dataset provides permissions to read the dataset's metadata and to list tables in the dataset; Create, update, get, and delete the dataset's tables. When applied at the project or organization level, this role can also create new datasets. We want to grant users access to query but not modify/delete.

Create a custom role by removing delete permissions, and add users to that role only. is not right.  
This might work but this is a manual, error-prone, time-consuming, and adds to operational overhead. If GCP provides a primitive role that is fit for purpose, this should be preferred over creating custom roles.

Create a custom role by removing delete permissions. Add users to the group, and then add the group to the custom role. is not right.

This might work but like above this is a manual, error-prone, time-consuming, and adds to operational overhead. If GCP provides a primitive role that is fit for purpose, this should be preferred over creating custom roles.

Add users to roles/bigquery.user role only, instead of roles/bigquery.dataOwner. is the right answer.  
roles/bigquery.user is a BigQuery User role which when applied to a project provides the ability to run jobs, including queries, within the project. A member with this role can enumerate their own jobs, cancel their own jobs, and enumerate datasets within a project.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#bigquery-roles>

#### 69. 69. Question

Your organization processes a very high volume of timestamped IoT data. The total volume can be several petabytes. The data needs to be written and changed at a high speed. You want to use the most performant storage option for your data. Which product should you use?

- BigQuery
- Cloud Bigtable
- Cloud Storage
- Cloud Datastore

Unattempted

Our requirement is to write/update a very high volume of data at a high speed. Performance is our primary concern, not cost.

Cloud Bigtable is the right answer.

Cloud Bigtable is Google's flagship product for ingest and analyze large volumes of time series data from sensors in real-time, matching the high speeds of IoT data to track normal and abnormal behavior.

Ref: <https://cloud.google.com/bigtable/>

While all other options are capable of storing high volumes of the order of petabytes, they are not as efficient as Bigtable at processing IoT time-series data.

#### 70. 70. Question

Your organization uses G Suite for communication and collaboration. All users in your organization have a G Suite account. You want to grant some G Suite users access to your Cloud Platform project. What should you do?

- Enable Cloud Identity in the GCP Console for your domain.
- Grant them the required IAM roles using their G Suite email address.
- Create a CSV sheet with all users' email addresses. Use the gcloud command line tool to convert them into Google Cloud Platform accounts.

- In the G Suite console, add the users to a special group called cloud-console-users@yourdomain.com. Rely on the default behavior of the Cloud Platform to grant users access if they are members of this group.

Unattempted

Grant them the required IAM roles using their G Suite email address. is the right answer.

You can use Cloud Identity or G Suite to create and manage users in GCP

Ref: <https://cloud.google.com/iam/docs/faq>

Since all users in organization already have a G Suite account, we should grant the roles to their G Suite email addresses for users that need access to GCP services.

## SET-4

### 1. Question

Your projects incurred more costs than you expected last month. Your research reveals that a development GKE container emitted a huge number of logs, which resulted in higher costs. You want to disable the logs quickly using the minimum number of steps. What should you do?

- 1. Go to the GKE console, and delete existing clusters. 2. Recreate a new cluster. 3. Clear the option to enable legacy Stackdriver Monitoring.
- Go to the Logs ingestion window in Stackdriver Logging, and disable the log source for the GKE container resource.**
- Go to the Logs ingestion window in Stackdriver Logging, and disable the log source for the GKE Cluster Operations resource.
- 1. Go to the GKE console, and delete existing clusters. 2. Recreate a new cluster. 3. Clear the option to enable legacy Stackdriver Logging.

### Unattempted

1. Go to the GKE console, and delete existing clusters.
  2. Recreate a new cluster.
  3. Clear the option to enable legacy Stackdriver Logging. is not right.  
Our requirement is to disable the logs ingested from the GKE container. We don't need to delete the existing cluster and create a new one.
- 
1. Go to the GKE console, and delete existing clusters.
  2. Recreate a new cluster.
  3. Clear the option to enable legacy Stackdriver Monitoring. is not right.  
Our requirement is to disable the logs ingested from the GKE container. We don't need to delete the existing cluster and create a new one.

Go to the Logs ingestion window in Stackdriver Logging, and disable the log source for the GKE Cluster Operations resource. is not right.  
Our requirement is to disable the logs ingested from GKE container, not the complete GKE Cluster Operations resource.

Go to the Logs ingestion window in Stackdriver Logging, and disable the log source for the GKE container resource. is the right answer.  
We want to disable logs from a specific GKE container and this is the only option that does that.  
More information about logs exclusions: <https://cloud.google.com/logging/docs/exclusions>

### 2. Question

Your team is working towards using the desired state configuration for your application deployed on the GKE cluster. You have YAML files for the Kubernetes Deployment and Service objects. Your application is designed to have 2 pods, which is defined by the replicas parameter in app-deployment.yaml. Your service uses GKE Load Balancer which is defined in app-service.yaml

You created the Kubernetes resources by running

kubectl apply -f app-deployment.yaml

kubectl apply -f app-service.yaml

Your deployment is now serving live traffic but is suffering from performance issues. You want to increase the number of replicas to 5. What should you do in order to update the replicas in existing Kubernetes deployment objects?

- Disregard the YAML file. Use the kubectl scale command to scale the replicas to 5. `kubectl scale --replicas=5 -f app-deployment.yaml`
- Disregard the YAML file. Enable autoscaling on the deployment to trigger on CPU usage and set max pods to 5. `kubectl autoscale myapp --max=5 --cpu-percent=80`
- Modify the current configuration of the deployment by using kubectl edit to open the YAML file of the current configuration, modify and save the configuration. `kubectl edit deployment/app-deployment -o yaml --save-config`
- Edit the number of replicas in the YAML file and rerun the kubectl apply. `kubectl apply -f app-deployment.yaml`

**Unattempted**

Disregard the YAML file. Use the kubectl scale command to scale the replicas to 5. `kubectl scale --replicas=5 -f app-deployment.yaml`. is not right.

While the outcome is the same, this approach doesn't update the change in the desired state configuration (YAML file). If you were to make some changes in your app-deployment.yaml and apply it, the update would scale back the replicas to 2. This is undesirable.

Ref: <https://kubernetes.io/docs/concepts/workloads/controllers/deployment/#scaling-a-deployment>

Disregard the YAML file. Enable autoscaling on the deployment to trigger on CPU usage and set minimum pods as well as maximum pods to 5. `kubectl autoscale myapp --min=5 --max=5 --cpu-percent=80`. is not right.

While the outcome is the same, this approach doesn't update the change in the desired state configuration (YAML file). If you were to make some changes in your app-deployment.yaml and apply it, the update would scale back the replicas to 2. This is undesirable.

Ref: <https://kubernetes.io/blog/2016/07/autoscaling-in-kubernetes/>

Modify the current configuration of the deployment by using kubectl edit to open the YAML file of the current configuration, modify and save the configuration. `kubectl edit deployment/app-deployment -o yaml --save-config`. is not right.

Like the above, the outcome is the same. This is equivalent to first getting the resource, editing it in a text editor, and then applying the resource with the updated version. This approach doesn't update the replicas change in our local YAML file. If you were to make some changes in your local app-deployment.yaml and apply it, the update would scale back the replicas to 2. This is undesirable.

Ref: <https://kubernetes.io/docs/concepts/cluster-administration/manage-deployment/#in-place-updates-of-resources>

Edit the number of replicas in the YAML file and rerun the kubectl apply. `kubectl apply -f app-deployment.yaml`. is the right answer.

This is the only approach that guarantees that you use desired state configuration. By updating the YAML file to have 5 replicas and applying it using

kubectl apply, you are preserving the intended state of Kubernetes cluster in the YAML file.

Ref: <https://kubernetes.io/docs/concepts/cluster-administration/manage-deployment/#in-place-updates-of-resources>

3. 3. Question

Your team uses Splunk for centralized logging and you have a number of reports and dashboards based on the logs in Splunk. You want to install Splunk forwarder on all nodes of your new Kubernetes Engine Autoscaled Cluster. The Splunk forwarder forwards the logs to a centralized Splunk Server. You want to minimize operational overhead. What is the best way to install Splunk Forwarder on all nodes in the cluster?

- Include the forwarder agent in a DaemonSet deployment.**
- Use Deployment Manager to orchestrate the deployment of forwarder agents on all nodes.
- Include the forwarder agent in a StatefulSet deployment.
- SSH to each node and run a script to install the forwarder agent.

**Unattempted**

SSH to each node and run a script to install the forwarder agent. is not right. While this can be done, this approach does not scale. Every time the Kubernetes cluster autoscaling adds a new node, we have to SSH to the instance and run the script which is manual, possibly error-prone and adds operational overhead. We need to look for a solution that automates this task.

Include the forwarder agent in a StatefulSet deployment. is not right.

In GKE, StatefulSets represents a set of Pods with unique, persistent identities and stable hostnames that GKE maintains regardless of where they are scheduled. The main purpose of StatefulSets is to set up persistent storage for pods that are deployed across multiple zones. StatefulSets are not suitable for installing the forwarder agent nor do they provide us the ability to install forwarder agents.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/statefulset>

Use Deployment Manager to orchestrate the deployment of forwarder agents on all nodes. is not right.

You can use a deployment manager to create a number of GCP resources including GKE Cluster but you can not use it to create Kubernetes deployments or apply configuration files.

Ref: <https://cloud.google.com/deployment-manager/docs/fundamentals>

Include the forwarder agent in a DaemonSet deployment. is the right answer.

In GKE, DaemonSets manage groups of replicated Pods and adhere to a one-Pod-per-node model, either across the entire cluster or a subset of nodes. As you add nodes to a node pool, DaemonSets automatically add Pods to the new nodes. So by configuring the pod to use Splunk forwarder agent image and with some minimal configuration (e.g. identifying which logs need to be forwarded), you can automate the installation and configuration of Splunk forwarder agent on each GKE cluster node.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/daemonset>

4. 4. Question

Your VMs are running in a subnet that has a subnet mask of 255.255.255.240. The current subnet has no more free IP addresses and you require an additional 10 IP addresses for new VMs. The existing and new VMs should all be able to reach each other without additional routes. What should you do?

- Use gcloud to expand the IP range of the current subnet.
- Delete the subnet, and recreate it using a wider range of IP addresses.
- Create a new project. Use Shared VPC to share the current network with the new project.
- Create a new subnet with the same starting IP but a wider range to overwrite the current subnet.

**Unattempted**

Use gcloud to expand the IP range of the current subnet. is the right answer.

Subnet mask of the existing subnet is 255.255.255.240 which means the max possible address in are 16. So the net prefix is /28 i.e. 4 bits free so 2 to the power of 4 is 16 IP Addresses.

As per IETF (Ref: <https://tools.ietf.org/html/rfc1918>), the supported internal IP Address ranges are

1. 24-bit block 10.0.0.0/8 (16777216 IP Addresses)
2. 20-bit block 172.16.0.0/12 (1048576 IP Addresses)
3. 16-bit block 192.168.0.0/16 (65536 IP Addresses)

A prefix of 28 is a very small subnet and could be in any of the ranges above; and all ranges have scope to accommodate a higher prefix.

A prefix of 27 gives you 32 IP Addresses i.e. 16 IP address more and we just need 10 more. So expanding the subnet to a prefix of 27 should give us the required capacity. And GCP lets you do exactly that running a gcloud command <https://cloud.google.com/sdk/gcloud/reference/compute/networks/subnets/expand-ip-range>

gcloud compute networks subnets expand-ip-range --region= --prefix-length=27

5. 5. Question

You've created a Kubernetes engine cluster named "my-gcp-ace-proj-1", which has a cluster pool named my-gcp-ace-primary-node-pool. You want to increase the number of nodes within your cluster pool from 10 to 20 to meet capacity demands. What is the command to change the number of nodes in your pool?

- gcloud container clusters update my-gcp-ace-proj-1 --node-pool my-gcp-ace-primary-node-pool --num-nodes 20
- gcloud container clusters resize my-gcp-ace-proj-1 --node-pool my-gcp-ace-primary-node-pool --num-nodes 20
- kubectl container clusters update my-gcp-ace-proj-1 --node-pool my-gcp-ace-primary-node-pool --num-nodes 20

- gcloud container clusters resize my-gcp-ace-proj-1 --node-pool my-gcp-ace-primary-node-pool --new-size 20

**Unattempted**

kubectl container clusters update my-gcp-ace-proj-1 --node-pool my-gcp-ace-primary-node-pool --num-nodes 20. is not right.

kubectl does not accept container as an operation.

Ref: <https://kubernetes.io/docs/reference/kubectl/overview/#operations>

gcloud container clusters update my-gcp-ace-proj-1 --node-pool my-gcp-ace-primary-node-pool --num-nodes 20. is not right.

gcloud container clusters update can not be used to specify the number of nodes. It can be used to specify the node locations, but not the number of nodes.

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/update>

gcloud container clusters resize my-gcp-ace-proj-1 --node-pool my-gcp-ace-primary-node-pool --new-size 20. is not right.

gcloud container clusters resize command does not support the parameter new-size. While --size can be used to resize the cluster node pool, use of --size is discouraged as this is a deprecated parameter. “The --size flag is now deprecated. Please use --num-nodes instead.”

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/resize>

gcloud container clusters resize my-gcp-ace-proj-1 --node-pool my-gcp-ace-primary-node-pool --num-nodes 20. is the right answer

gcloud container clusters resize can be used to specify the number of nodes using the --num-nodes parameter which is the target number of nodes in the cluster.

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/resize>

6. 6. Question

You are designing a large distributed application with 30 microservices. Each of your distributed microservices needs to connect to a database back-end. You want to store the credentials securely. Where should you store the credentials?

- A. In the source code
- B. In an environment variable
- C. In a secret management system
- D. In a config file that has restricted access through ACLs

**Unattempted**

Correct answer is C as it is a recommended practice to store the credentials in a secret management system such as KMS. Applications often require access to small pieces of sensitive data at build or run time. These pieces of data are often referred to as secrets. Secrets are similar in concept to configuration files, but are generally more sensitive, as they may grant access to additional data, such as user data.

Refer GCP documentation – Authentication Managing Credentials

Best practices for managing credentials

Credentials provide access to sensitive data. The following practices help protect access to these resources.

Do not embed secrets related to authentication in source code, such as API keys,

OAuth tokens, and service account credentials. You can use an environment variable pointing to credentials outside of the application's source code, such as Cloud Key Management Service.

Do use different credentials in different contexts, such as in testing and production environments.

Do transfer credentials only over HTTPS to prevent a third party from intercepting your credentials. Never transfer in clear text or as part of the URL.

Never embed long-lived credentials into your client-side app. For example, do not embed service account credentials into a mobile app. Client-side apps can be examined and credentials can easily be found and used by a third party.

Do revoke a token if you no longer need it.

Options A, B & D are wrong as they are not recommended and does not provide security.

## 7. Question

Your company's test suite is a custom C++ application that runs tests throughout each day on Linux virtual machines. The full test suite takes several hours to complete, running on a limited number of on-premises servers reserved for testing. Your company wants to move the testing infrastructure to the cloud, to reduce the amount of time it takes to fully test a change to the system, while changing the tests as little as possible. Which cloud infrastructure should you recommend?

- A. Google Compute Engine unmanaged instance groups and Network Load Balancer.
- B. Google Compute Engine managed instance groups with auto-scaling.
- C. Google Cloud Dataproc to run Apache Hadoop jobs to process each test.
- D. Google App Engine with Google Stackdriver for logging.

### Unattempted

Correct answer is B as Google Compute Engine managed instance group can help the testing application to scale to reduce the amount of time to run tests.

Refer GCP documentation – Instance groups

A managed instance group uses an instance template to create a group of identical instances. You control a managed instance group as a single entity. If you wanted to make changes to instances that are part of a managed instance group, you would make the change to the whole instance group. Because managed instance groups contain identical instances, they offer the following features.

When your applications require additional compute resources, managed instance groups can automatically scale the number of instances in the group.

Managed instance groups work with load balancing services to distribute traffic to all of the instances in the group.

If an instance in the group stops, crashes, or is deleted by an action other than the instance groups commands, the managed instance group automatically recreates the instance so it can resume its processing tasks. The recreated instance uses the same name and the same instance template as the previous

instance, even if the group references a different instance template. Managed instance groups can automatically identify and recreate unhealthy instances in a group to ensure that all of the instances are running optimally. The managed instance group updater allows you to easily deploy new versions of software to instances in your managed instance groups, while controlling the speed and scope of deployment as well as the level of disruption to your service. Option A is wrong as unmanaged group does not scale. Option C is wrong as Dataproc is for big data batch jobs. Option D is wrong as App Engine standard does not support C++ application and the testing application needs to be dockerized to be used with flexible engine.

8. 8. Question

Your company collects and stores security camera footage in Google Cloud Storage. Within the first 30 days, footage is processed regularly for threat detection, object detection, trend analysis, and suspicious behavior detection. You want to minimize the cost of storing all the data. How should you store the videos?

- A. Use Google Cloud Regional Storage for the first 30 days, and then move to Coldline Storage.
- B. Use Google Cloud Nearline Storage for the first 30 days, and then move to Coldline Storage.
- C. Use Google Cloud Regional Storage for the first 30 days, and then move to Nearline Storage.
- D. Use Google Cloud Regional Storage for the first 30 days, and then move to Google Persistent Disk.

**Unattempted**

Correct answer is A as the data is accessed frequently within the first 30 days, using Google Cloud Regional Storage will enable the most cost-effective solution for storing and accessing the data. For videos older than 30 days, Google Cloud Coldline Storage offers the most cost-effective solution since it won't be accessed.

Refer GCP documentation – Cloud Storage – Storage Classes

Option B is wrong as while Google Cloud Coldline storage is cost-effective for long-term video storage, Google Cloud Nearline Storage would not be an effective solution for the first 30 days as the data is expected to be accessed frequently.

Option C is wrong as while Google Cloud Regional Storage is the most cost-effective solution for the first 30 days, Google Cloud Nearline Storage is not cost effective for long-term storage.

Option D is wrong as while Google Cloud Regional Storage is the most cost-effective solution for the first 30 days, storing the data on Google Cloud Persistent Disk would not be cost-effective for long term storage.

9. 9. Question

Your company processes high volumes of IoT data that are time-stamped. The total data volume can be several petabytes. The data needs to be written and

changed at a high speed. You want to use the most performant storage option for your data. Which product should you use?

- A. Cloud Datastore
- B. Cloud Storage
- C. Cloud Bigtable
- D. BigQuery

**Unattempted**

Correct answer is C as Cloud Bigtable is the most performant storage option to work with IoT and time series data. Google Cloud Bigtable is a fast, fully managed, highly-scalable NoSQL database service. It is designed for the collection and retention of data from 1TB to hundreds of PB.

Refer GCP documentation – Bigtable Time series data

Option A is wrong as Cloud Datastore is not the most performant product for frequent writes or timestamp-based queries.

Option B is wrong as Cloud Storage is designed for object storage not for this type of data ingestion and collection.

Option D is wrong as BigQuery is more of an a scalable, fully managed enterprise data warehousing solution and not ideal fast changing data.

#### 10. Question

Your company is planning the infrastructure for a new large-scale application that will need to store over 100 TB or a petabyte of data in NoSQL format for Low-latency read/write and High-throughput analytics. Which storage option should you use?

- A. Cloud Bigtable
- B. Cloud Spanner
- C. Cloud SQL
- D. Cloud Datastore

**Unattempted**

Correct answer is A as Bigtable is an ideal solution to provide low latency, high throughput data processing storage option with analytics

Refer GCP documentation – Storage Options

Cloud Bigtable logo

A scalable, fully managed NoSQL wide-column database that is suitable for both low-latency single-point lookups and precalculated analytics.

Low-latency read/write access IoT, finance, adtech High-throughput data processing Personalization, recommendations Time series support Monitoring Geospatial datasets Graphs

Options B & C are wrong as they are relational databases

Option D is wrong as Cloud Datastore is not ideal for analytics.

#### 11. Question

A company wants building an application stores images in a Cloud Storage bucket and want to generate thumbnails as well resize the images. They want to use managed service which will help them scale automatically from zero to scale and back to zero. Which GCP service satisfies the requirement?

- A. Google Compute Engine
- B. Google Kubernetes Engine
- C. Google App Engine
- D. Cloud Functions

**Unattempted**

Correct answer is D as Cloud Functions can help automatically scale as per the demand, with no invocations if no demand.

Refer GCP documentation – Cloud Functions

Google Cloud Functions is a serverless execution environment for building and connecting cloud services. With Cloud Functions you write simple, single-purpose functions that are attached to events emitted from your cloud infrastructure and services. Your function is triggered when an event being watched is fired. Your code executes in a fully managed environment. There is no need to provision any infrastructure or worry about managing any servers.

Cloud Functions removes the work of managing servers, configuring software, updating frameworks, and patching operating systems. The software and infrastructure are fully managed by Google so that you just add code.

Furthermore, provisioning of resources happens automatically in response to events. This means that a function can scale from a few invocations a day to many millions of invocations without any work from you.

Options A, B & C are wrong as they need to be configured to scale down and would need warm up time to scale back again as compared to Cloud Functions.

## 12. Question

Your company is planning on deploying a web application to Google Cloud hosted on a custom Linux distribution. Your website will be accessible globally and needs to scale to meet demand. Choose all of the components that will be necessary to achieve this goal. (Select TWO)

- A. App Engine Standard environment
- B. HTTP Load Balancer
- C. Managed Instance Group on Compute Engine
- D. Network Load Balancer

**Unattempted**

Correct answers are B & C

Option B as only HTTP load balancer support global access.

Option C as the requirement is to support custom Linux distribution, only Compute Engine supports the same.

Refer GCP documentation – Load Balancing

HTTP(S) load balancing can balance HTTP and HTTPS traffic across multiple backend instances, across multiple regions. Your entire app is available via a single global IP address, resulting in a simplified DNS setup. HTTP(S) load balancing is scalable, fault-tolerant, requires no pre-warming, and enables content-based load balancing. For HTTPS traffic, it provides SSL termination and load balancing.

Option A is wrong as App Engine does not support custom linux distribution.

Option D is wrong as Network load balancer does not support global access.

### 13. 13. Question

Your customer is moving their corporate applications to Google Cloud Platform. The security team wants detailed visibility of all projects in the organization. You provision the Google Cloud Resource Manager and set up yourself as the org admin. What Google Cloud Identity and Access Management (Cloud IAM) roles should you give to the security team?

- A. Org viewer, project owner
- B. Org viewer, project viewer
- C. Org admin, project browser
- D. Project owner, network admin

**Unattempted**

Correct answer is B as the security team only needs visibility to the projects, project viewer provides the same with the best practice of least privilege.

Refer GCP documentation – Organization & Project access control

Option A is wrong as project owner will provide access however it does not align with the best practice of least privilege.

Option C is wrong as org admin does not align with the best practice of least privilege.

Option D is wrong as the user needs to be provided organization viewer access to see the organization.

### 14. 14. Question

Auditors visit your teams every 12 months and ask to review all the Google Cloud Identity and Access Management (Cloud IAM) policy changes in the previous 12 months. You want to streamline and expedite the analysis and audit process.

What should you do?

- A. Create custom Google Stackdriver alerts and send them to the auditor
- B. Enable Logging export to Google BigQuery and use ACLs and views to scope the data shared with the auditor
- C. Use Cloud Functions to transfer log entries to Google Cloud SQL and use ACLs and views to limit an auditor's view
- D. Enable Google Cloud Storage (GCS) log export to audit logs into a GCS bucket and delegate access to the bucket

### Unattempted

Correct answer is B as BigQuery is a good storage option with analysis capability. Also, the access to the data can be controlled using ACLs and Views. BigQuery uses access control lists (ACLs) to manage permissions on projects and datasets.

BigQuery is a petabyte-scale analytics data warehouse that you can use to run SQL queries over vast amounts of data in near realtime.

Giving a view access to a dataset is also known as creating an authorized view in BigQuery. An authorized view allows you to share query results with particular users and groups without giving them access to the underlying tables. You can also use the view's SQL query to restrict the columns (fields) the users are able to query. In this tutorial, you create an authorized view.

Option A is wrong as alerts are real time and auditor do not need them.

Option C is wrong as Cloud SQL is not ideal for storage of log files and cannot be controlled through ACLs.

Option D is wrong as Cloud Storage is a good storage option but does not provide direct analytics capabilities.

### 15. Question

Your App Engine application needs to store stateful data in a proper storage service. Your data is non-relational database data. You do not expect the database size to grow beyond 10 GB and you need to have the ability to scale down to zero to avoid unnecessary costs. Which storage service should you use?

- A. Cloud Bigtable
- B. Cloud Dataproc
- C. Cloud SQL
- D. Cloud Datastore

### Unattempted

Correct answer is D as Cloud Datastore provides a scalable, fully managed NoSQL document database for your web and mobile applications.

Cloud Datastore A scalable, fully managed NoSQL document database for your web and mobile applications. Semistructured application data User profiles

Hierarchical data Product catalogs Durable key-value data Game state

Option A is wrong as Bigtable is not an ideal storage option for state management. Cloud Bigtable A scalable, fully managed NoSQL wide-column database that is suitable for both low-latency single-point lookups and precalculated analytics. Low-latency read/write access IoT, finance, adtech High-throughput data processing Personalization, recommendations Time series support Monitoring Geospatial datasets Graphs

Option B is wrong as Dataproc is not a storage solution. Cloud Dataproc is a fast, easy-to-use, fully-managed cloud service for running Apache Spark and Apache Hadoop clusters in a simpler, more cost-efficient way.

Option C is wrong as you need to define a capacity while provisioning a database.

Cloud SQL A fully managed MySQL and PostgreSQL database service that is built on the strength and reliability of Google's infrastructure. Web frameworks

Websites, blogs, and content management systems (CMS) Structured data  
Business intelligence (BI) applications  
OLTP workloads ERP, CRM, and ecommerce applications Geospatial application

16. 16. Question

You have a collection of media files over 50GB each that you need to migrate to Google Cloud Storage. The files are in your on-premises data center. What migration method can you use to help speed up the transfer process?

- A. Use multi-threaded uploads using the -m option.
- B. Use parallel uploads to break the file into smaller chunks then transfer it simultaneously.
- C. Use the Cloud Transfer Service to transfer.
- D. Start a recursive upload.

**Unattempted**

Correct answer is B as gsutil provide object composition or parallel upload to handle upload of larger files.

Refer GCP documentation – Optimizing for Cloud Storage Performance

More efficient large file uploads

The gsutil utility can also automatically use object composition to perform uploads in parallel for large, local files that you want to upload to Cloud Storage. It splits a large file into component pieces, uploads them in parallel and then recomposes them once they're in the cloud (and deletes the temporary components it created locally).

You can enable this by setting the `parallel\_composite\_upload\_threshold` option on gsutil (or, updating your .boto file, like the console output suggests).

```
gsutil -o GSUtil:parallel_composite_upload_threshold=150M cp ./localbigfile
gs://your-bucket
```

Where "localbigfile" is a file larger than 150MB. This divides up your data into chunks ~150MB and uploads them in parallel, increasing upload performance. Option A is wrong as multi-threaded options is best suited for uploading multiple files to better utilize the bandwidth.

Option C is wrong as Cloud Transfer service cannot handle uploads from on-premises data center.

Option D is wrong as recursive upload helps handle folders and subfolders.

17. 17. Question

A Company is planning the migration of their web application to Google App Engine. However, they would still continue to use their on-premises database. How can they setup application?

- A. Setup the application using App Engine Standard environment with Cloud VPN to connect to database
- B. Setup the application using App Engine Flexible environment with Cloud VPN to connect to database

- C. Setup the application using App Engine Standard environment with Cloud Router to connect to database
- D. Setup the application using App Engine Flexible environment with Cloud Router to connect to database

#### Unattempted

Correct answer is B as Google App Engine provides connectivity to on-premises using Cloud VPN.

Refer GCP documentation – App Engine Flexible Network Settings

Advanced network configuration

You can segment your Compute Engine network into subnetworks. This allows you to enable VPN scenarios, such as accessing databases within your corporate network.

To enable subnetworks for your App Engine application:

Create a custom subnet network.

Add the network name and subnetwork name to your app.yaml file, as specified above.

To establish a simple VPN based on static routing, create a gateway and a tunnel for a custom subnet network. Otherwise, see how to create other types of VPNs.

Option A is wrong as Google App Engine Standard cannot use Cloud VPN.

Options C & D are wrong as you need a Cloud VPN to connect to on-premises data center. Cloud Route support dynamic routing.

#### 18. Question

A lead software engineer tells you that his new application design uses websockets and HTTP sessions that are not distributed across the web servers. You want to help him ensure his application will run properly on Google Cloud Platform. What should you do?

- A. Help the engineer to convert his websocket code to use HTTP streaming.
- B. Review the encryption requirements for websocket connections with the security team.
- C. Meet with the cloud operations team and the engineer to discuss load balancer options.
- D. Help the engineer redesign the application to use a distributed user session service that does not rely on websockets and HTTP sessions.

#### Unattempted

Correct answer is C as the HTTP(S) load balancer in GCP handles websocket traffic natively. Backends that use WebSocket to communicate with clients can use the HTTP(S) load balancer as a front end, for scale and availability.

Refer GCP documentation – HTTP Load Balancer

HTTP(S) Load Balancing has native support for the WebSocket protocol.

Backends that use WebSocket to communicate with clients can use the HTTP(S) load balancer as a front end, for scale and availability. The load balancer does not need any additional configuration to proxy WebSocket connections.

The WebSocket protocol, which is defined in RFC 6455, provides a full-duplex

communication channel between clients and servers. The channel is initiated from an HTTP(S) request

Option A is wrong as there is no compelling reason to move away from websockets as part of a move to GCP.

Option B is wrong as this may be a good exercise anyway, but it doesn't really have any bearing on the GCP migration.

Option D is wrong as there is no compelling reason to move away from websockets as part of a move to GCP.

#### 19. 19. Question

Your customer is moving their storage product to Google Cloud Storage (GCS). The data contains personally identifiable information (PII) and sensitive customer information. What security strategy should you use for GCS?

- A. Use signed URLs to generate time bound access to objects.
- B. Grant IAM read-only access to users, and use default ACLs on the bucket.
- C. Grant no Google Cloud Identity and Access Management (Cloud IAM) roles to users, and use granular ACLs on the bucket.
- D. Create randomized bucket and object names. Enable public access, but only provide specific file URLs to people who do not have Google accounts and need access.

**Unattempted**

Correct answer is C as this grants the least privilege required to access the data and minimizes the risk of accidentally granting access to the wrong people.

Refer GCP documentation – Cloud Storage Access Control

Option A is wrong as Signed URLs could potentially be leaked as anyone who gets access to the URL can access the data.

Option B is wrong as this is needlessly permissive, users only require one permission in order to get access.

Option D is wrong as this is security through obscurity, also known as no security at all.

#### 20. 20. Question

You've created a Kubernetes engine cluster named "project-1", which has a cluster pool named 'primary-node-pool'. You've realized that you need more total nodes within your cluster pool to meet capacity demands from 10 to 20. What is the command to change the number of nodes in your pool?

- A. gcloud container clusters update project-1 --node pool 'primary-node-pool' --num-nodes 20
- B. gcloud container clusters resize project-1 --node pool 'primary-node-pool' --size 20
- C. gcloud container clusters resize project-1 --node pool 'primary-node-pool' --num-nodes 20

- D. kubectl container clusters update project-1 --node pool 'primary-node-pool' --num-nodes 20

**Unattempted**

Correct answer is B as the resize command with gcloud can be used to increase the nodes.

NOTE – The size flag has been renamed to num-nodes flag from 242.0.0 (2019-04-16)

#### Kubernetes Engine

Renamed –size flag of gcloud container clusters resize to –num-nodes. –size retained as an alias.

Disabled node auto-repair and node auto-upgrade by default when –enable-kubernetes-alpha flag is used to create clusters with Kubernetes alpha features enabled. Users may now create alpha clusters without specifying –no-enable-autorepair or –no-enable-autoupgrade flags. However, for creating new node pools in an existing alpha cluster, these two flags may still be required.

Refer GCP documentation – Resizing Kubernetes Cluster

gcloud container clusters resize [CLUSTER\_NAME] –node-pool [POOL\_NAME] –size [SIZE];

Option A is wrong as update command takes in the –max-nodes & –min-nodes flags which are defining the autoscaling. –num-nodes flag is not applicable.

Option C is wrong as –num-nodes is a wrong flag for cluster resize command.

Option D is wrong as kubectl command cannot be used for resizing the cluster.

## 21. Question

A Company is using Cloud SQL to host critical data. They want to enable high availability in case a complete zone goes down. How should you configure the same?

- A. Create a Read replica in the same region different zone
- B. Create a Read replica in the different region different zone
- C. Create a Failover replica in the same region different zone
- D. Create a Failover replica in the different region different zone

**Unattempted**

Correct answer is C as a failover replica helps provides High Availability for Cloud SQL. The failover replica must be in the same region as the primary instance.

Refer GCP documentation – Cloud SQL High Availability

The HA configuration, sometimes called a cluster, provides data redundancy. The configuration is made up of a primary instance (master) in the primary zone and a failover replica in the secondary zone. Through semisynchronous replication, all changes made to the primary instance's data and user tables are copied onto the failover replica. In the event of an instance or zone failure, this configuration reduces downtime, and your data continues to be available to client applications. The failover replica must be in the same region as the primary instance, but in a different zone.

Diagram overview of MySQL HA configuration. Described in text below.

Option A & B are wrong as Read replicas do not provide failover capability and just additional read capacity.

Option D is wrong as failover replica must be in the same region as the primary instance.

## 22. Question

Your application is hosted across multiple regions and consists of both relational database data and static images. Your database has over 10 TB of data. You want to use a single storage repository for each data type across all regions. Which two products would you choose for this task? (Choose two)

- A. Cloud Bigtable
- B. Cloud Spanner
- C. Cloud SQL
- D. Cloud Storage

**Unattempted**

Correct answers are B & D

Option B to store the relational data. As the data is over 10TB and need across region, Cloud Spanner is preferred over Cloud SQL.

Option D to store unstructured static images.

Refer GCP documentation – Storage Options

Option A is wrong as Bigtable is a NoSQL data storage and not suitable to store unstructured data as images and files.

Option C is wrong as Cloud SQL is regional and not a preferred option for data over 10TB.

## 23. Question

Your project has all its Compute Engine resources in the europe-west1 region. You want to set europe-west1 as the default region for gcloud commands. What should you do?

- A. Use Cloud Shell instead of the command line interface of your device. Launch Cloud Shell after you navigate to a resource in the europe-west1 region. The europe-west1 region will automatically become the default region.
- B. Use gcloud config set compute/region europe-west1 to set the default region for future gcloud commands.
- C. Use gcloud config set compute/zone europe-west1 to set the default region for future gcloud commands.
- D. Create a VPN from on-premises to a subnet in europe-west1, and use that connection when executing gcloud commands.

**Unattempted**

Correct answer is B as this will ensure that the relevant region is used when not overwritten by a command parameter.

Refer GCP documentation – Change default zone and region

You can manually choose a different zone or region without updating the

metadata server by setting these properties locally on your gcloud client.  
gcloud config compute/region REGION

Option A is wrong as Cloud Shell will not default to the location that it's launched from.

Option C is wrong as this command should be used to set a zone, not a region.

Option D is wrong as a VPN to a specific subnet does not have any effect on the gcloud command region.

#### 24. 24. Question

You have an application server running on Compute Engine in the europe-west1-d zone. You need to ensure high availability and replicate the server to the europe-west2-c zone using the fewest steps possible. What should you do?

- A. Create a snapshot from the disk. Create a disk from the snapshot in the europe-west2-c zone. Create a new VM with that disk.
- B. Create a snapshot from the disk. Create a disk from the snapshot in the europe-west1-d zone and then move the disk to europe-west2-c. Create a new VM with that disk.
- C. Use gcloud to copy the disk to the europe-west2-c zone. Create a new VM with that disk.
- D. Use gcloud compute instances move with parameter --destination-zone europe-west2-c to move the instance to the new zone.

#### Unattempted

Correct answer is A as the best way to create a replica of disk is to create a snapshot and create a disk from the snapshot in the zone.

Refer GCP documentation – Disks

Disks are zonal resources, so they reside in a particular zone for their entire lifetime. The contents of a disk can be moved to a different zone by snapshotting the disk (using gcloud compute disks snapshot) and creating a new disk using –source-snapshot in the desired zone. The contents of a disk can also be moved across project or zone by creating an image (using gcloud compute images create) and creating a new disk using –image in the desired project and/or zone. Option B is wrong as the approach is possible, but not with the fewest steps. Option C is wrong as gcloud cannot be used to copy the disk to different zone. Option D is wrong as it would move and not create a copy. gcloud compute disks move facilitates moving a Google Compute Engine disk volume from one zone to another. You cannot move a disk if it is attached to a running or stopped instance; use the gcloud compute instances move command instead.

#### 25. 25. Question

You need to estimate the annual cost of running a BigQuery query that is scheduled to run nightly. What should you do?

- A. Use gcloud query --dry\_run to determine the number of bytes read by the query. Use this number in the Pricing Calculator.

- B. Use bq query --dry\_run to determine the number of bytes read by the query. Use this number in the Pricing Calculator.**
- C. Use gcloud estimate to determine the amount billed for a single query. Multiply this amount by 365.
- D. Use bq estimate to determine the amount billed for a single query. Multiply this amount by 365.

**Unattempted**

Correct answer is B as this is the correct way to estimate the yearly BigQuery querying costs.

Refer GCP documentation – BigQuery Best Practices – Price your Query

Best practice: Before running queries, preview them to estimate costs.

Queries are billed according to the number of bytes read. To estimate costs before running a query use:

The query validator in the GCP Console or the classic web UI

The –dry\_run flag in the CLI

The dryRun parameter when submitting a query job using the API

The Google Cloud Platform Pricing Calculator

Option A is wrong as you should use “bq”, not “gcloud”, to estimate the amount of bytes read.

Option C is wrong as you should use “bq”, not “gcloud”, to work with BigQuery.

Option D is wrong as this will not give the amount billed for a query.

## 26. Question

You work in a small company where everyone should be able to view all resources of a specific project. You want to grant them access following Google's recommended practices. What should you do?

- A. Create a script that uses gcloud projects add-iam-policy-binding for all users' email addresses and the Project Viewer role.
- B. Create a script that uses gcloud iam roles create for all users' email addresses and the Project Viewer role.
- C. Create a new Google Group and add all users to the group. Use gcloud projects add-iam-policy-binding with the Project Viewer role and Group email address.**
- D. Create a new Google Group and add all members to the group. Use gcloud iam roles create with the Project Viewer role and Group email address.

**Unattempted**

Correct answer is C as Google recommends to use groups where possible.

Refer GCP documentation – gcloud IAM

Option A is wrong as groups are recommended over individual assignments.

Option B is wrong as this command is to create roles, not to assign them.

Option D is wrong as this command is to create roles, not to assign them.

## 27. Question

Your developers are trying to select the best compute service to run a static website. They have a dozen HTML pages, a few JavaScript files, and some CSS. They need the site to be highly available for the few weeks it is running. They also have a limited budget. What is the best service to use to run the site?

- A. Kubernetes Engine
- B. Compute Engine
- C. Cloud Storage
- D. App Engine

**Unattempted**

Correct answer is C as the website is static and needs to be hosted with high availability and limited budget, Cloud Storage would be an ideal choice.

Refer GCP documentation – Cloud Storage Static Website

To host a static site in Cloud Storage, you need to create a Cloud Storage bucket, upload the content, and test your new site. You can serve your data directly from storage.googleapis.com, or you can verify that you own your domain and use your domain name. Either way, you'll get consistent, fast delivery from global edge caches.

You can create your static web pages however you choose. For example, you could hand-author pages by using HTML and CSS. You can use a static-site generator, such as Jekyll, Ghost, or Hugo, to create the content. Static-site generators make it easier for you to create a static website by letting you author in markdown, and providing templates and tools. Site generators generally provide a local web server that you can use to preview your content.

After your static site is working, you can update the static pages by using any process you like. That process could be as straightforward as hand-copying an updated page to the bucket. You might choose to use a more automated approach, such as storing your content on GitHub and then using a webhook to run a script that updates the bucket. An even more advanced system might use a continuous-integration /continuous-delivery (CI/CD) tool, such as Jenkins, to update the content in the bucket. Jenkins has a Cloud Storage plugin that provides a Google Cloud Storage Uploader post-build step to publish build artifacts to Cloud Storage.

If you have a web application that needs to serve static content or user-uploaded static media, using Cloud Storage can be a cost-effective and efficient way to host and serve this content, while reducing the amount of dynamic requests to your web application.

Options A, B & D are wrong as they would be an expensive option as compared to Cloud Storage hosting.

## 28. Question

You have an autoscaled managed instance group that is set to scale based on CPU utilization of 60%. There are currently 3 instances in the instance group. You're connected to one of the instances and notice that the CPU usage is a 70%. However, the instance group isn't starting up another instance. What's the most likely reason?

- A. The autoscaler is disabled.
- B. The autoscaler takes 60 seconds before creating a new instance.
- C. The load balancer doesn't recognize the instance as healthy.
- D. The average CPU for the entire instance group is below 60%.

**Unattempted**

Correct answer is D as the Auto Scaler checks for the average CPU utilization across the instances and is not done on the basis of a single instance.

Refer GCP documentation – Auto Scaler – CPU based Scaling

You can autoscale based on the average CPU utilization of a managed instance group. Using this policy tells the autoscaler to collect the CPU utilization of the instances in the group and determine whether it needs to scale. You set the target CPU utilization the autoscaler should maintain and the autoscaler will work to maintain that level.

The autoscaler treats the target CPU utilization level as a fraction of the average use of all vCPUs over time in the instance group. If the average usage of your total vCPUs exceeds the target utilization, the autoscaler will add more virtual machines. For example, setting a 0.75 target utilization tells the autoscaler to maintain an average usage of 75% among all vCPUs in the instance group.

Option A is wrong as the group is set to CPU utilization already, it is not disabled. Option B is wrong as Auto Scaler takes action immediately if the target is hit.

Option C is wrong as if the instance is marked unhealthy it would not serve any traffic and might be replaced.

## 29. Question

You are required to fire a query on large amount of data stored in BigQuery. You know the query is expected to return a large amount of data. How would you estimate the cost for the query?

- A. Using Command line, use the --dry\_run option on BigQuery to determine the amount of bytes read, and then use the price calculator to determine the cost.
- B. Using Command line, use the --dry\_run option on BigQuery to determine the amount of bytes returned, and then use the price calculator to determine the cost.
- C. Using Command line, use the --dry\_run option on BigQuery to determine the amount of time taken, and then use the price calculator to determine the cost.
- D. Using Command line, use the --dry\_run option on BigQuery to determine the total amount of table data in bytes, as it would be a full scan, and then use the price calculator to determine the cost.

**Unattempted**

Correct answer is A as the –dry-run option can be used to price your queries before they are actually fired. The Query returns the bytes read, which can then be used with the Pricing Calculator to estimate the query cost.

Refer GCP documentation – BigQuery Best Practices

Price your queries before running them

Best practice: Before running queries, preview them to estimate costs.

Queries are billed according to the number of bytes read. To estimate costs before running a query use:

The query validator in the GCP Console or the classic web UI

The --dry\_run flag in the CLI

The dryRun parameter when submitting a query job using the API

The Google Cloud Platform Pricing Calculator

Options B, C are wrong as the estimation needs to be done on the bytes read by the query and not returned or time taken.

Option D is wrong as it the bytes read would depend on the query and would not always a full table scan.

### 30. 30. Question

Your company wants to host confidential documents in Cloud Storage. Due to compliance requirements, there is a need for the data to be highly available and resilient even in case of a regional outage. Which storage classes help meet the requirement?

- A. Nearline
- B. Standard
- C. Multi-Regional
- D. Dual-Regional
- E. Regional

**Unattempted**

Correct answers are A & C as Multi-Regional and Nearline storage classes provide multi-region geo-redundant deployment, which can sustain regional failure.

Refer GCP documentation – Cloud Storage Classes

Multi-Regional Storage is geo-redundant.

The geo-redundancy of Nearline Storage data is determined by the type of location in which it is stored: Nearline Storage data stored in multi-regional locations is redundant across multiple regions, providing higher availability than Nearline Storage data stored in regional locations.

Data that is geo-redundant is stored redundantly in at least two separate geographic places separated by at least 100 miles. Objects stored in multi-regional locations are geo-redundant, regardless of their storage class.

Geo-redundancy occurs asynchronously, but all Cloud Storage data is redundant within at least one geographic place as soon as you upload it.

Geo-redundancy ensures maximum availability of your data, even in the event of large-scale disruptions, such as natural disasters. For a dual-regional location, geo-redundancy is achieved using two specific regional locations. For other multi-regional locations, geo-redundancy is achieved using any combination of data centers within the specified multi-region, which may include data centers that are not explicitly available as regional locations.

Options B & D are wrong as they do not exist

Option E is wrong as Regional storage class is not geo-redundant. Data stored in a narrow geographic region and Redundancy is across availability zones

31. 31. Question

Your company needs to backup data for disaster recovery scenarios store all the backup data. This data would be required only in the event of a disaster and won't be accessed otherwise. What is the best default storage class?

- A. Multi-regional
- B. Coldline
- C. Regional
- D. Nearline

**Unattempted**

Correct answer is B as Coldline storage is an ideal solution for disaster recovery data given its rarity of access.

Refer GCP documentation – Cloud Storage Classes

Google Cloud Storage Coldline is a very-low-cost, highly durable storage service for data archiving, online backup, and disaster recovery. Unlike other “cold” storage services, your data is available within milliseconds, not hours or days. Coldline Storage is the best choice for data that you plan to access at most once a year, due to its slightly lower availability, 90-day minimum storage duration, costs for data access, and higher per-operation costs. For example: Cold Data Storage – Infrequently accessed data, such as data stored for legal or regulatory reasons, can be stored at low cost as Coldline Storage, and be available when you need it.

Disaster recovery – In the event of a disaster recovery event, recovery time is key. Cloud Storage provides low latency access to data stored as Coldline Storage.

The geo-redundancy of Coldline Storage data is determined by the type of location in which it is stored: Coldline Storage data stored in multi-regional locations is redundant across multiple regions, providing higher availability than Coldline Storage data stored in regional locations.

Options A, C & D are wrong as they are not suited for infrequently accessed data, as disaster does not happen periodically but rarely.

32. 32. Question

Your company needs to backup data for disaster recovery scenarios store all the backup data. You are required to perform monthly disaster recovery drills, as a part of compliance. What is the best default storage class?

- A. Multi-regional
- B. Coldline
- C. Regional
- D. Nearline

### Unattempted

Correct answer is D as the data needs to be access monthly only, Nearline is the ideal solution for data storage.

Refer GCP documentation – Cloud Storage Classes

Google Cloud Storage Nearline is a low-cost, highly durable storage service for storing infrequently accessed data. Nearline Storage is a better choice than Multi-Regional Storage or Regional Storage in scenarios where slightly lower availability, a 30-day minimum storage duration, and costs for data access are acceptable trade-offs for lowered storage costs.

Nearline Storage is ideal for data you plan to read or modify on average once a month or less. For example, if you want to continuously add files to Cloud Storage and plan to access those files once a month for analysis, Nearline Storage is a great choice.

Nearline Storage is also appropriate for data backup, disaster recovery, and archival storage. Note, however, that for data accessed less frequently than once a year, Coldline Storage is the most cost-effective choice, as it offers the lowest storage costs.

The geo-redundancy of Nearline Storage data is determined by the type of location in which it is stored: Nearline Storage data stored in multi-regional locations is redundant across multiple regions, providing higher availability than Nearline Storage data stored in regional locations.

Options A, B & C are wrong as they are not ideal for data that is only accessed once monthly.

### 33. Question

Your developers are trying to connect to an Ubuntu server over SSH to diagnose some errors. However, the connection times out. Which command should help solve the problem?

- A. gcloud compute firewall-rules create open-ssh --network \$NETWORK --allow tcp:22
- B. gcloud compute firewall-rules create open-ssh
- C. gcloud compute firewall-rules create open-ssh --network \$NETWORK --deny tcp:22
- D. gcloud compute firewall-rules create open-ssh --network \$NETWORK --allow tcp:3389

### Unattempted

Correct answer is A as gcloud compute firewall-rules create is used to create firewall rules to allow/deny incoming/outgoing traffic.

Refer GCP documentation – Cloud SDK Firewall Rules – Create  
–allow=PROTOCOL[:PORT[-PORT]],[...]

A list of protocols and ports whose traffic will be allowed.

The protocols allowed over this connection. This can be the (case-sensitive) string values tcp, udp, icmp, esp, ah, sctp, or any IP protocol number. An IP-based protocol must be specified for each rule. The rule applies only to specified protocol.

For port-based protocols – tcp, udp, and sctp – a list of destination ports or port

ranges to which the rule applies may optionally be specified. If no port or port range is specified, the rule applies to all destination ports.

The ICMP protocol is supported, but there is no support for configuring ICMP packet filtering by ICMP code.

For example, to create a rule that allows TCP traffic through port 80 and ICMP traffic:

```
gcloud compute firewall-rules create MY-RULE --allow tcp:80,icmp
```

To create a rule that allows TCP traffic from port 20000 to 25000:

```
gcloud compute firewall-rules create MY-RULE --allow tcp:20000-25000
```

To create a rule that allows all TCP traffic:

```
gcloud compute firewall-rules create MY-RULE --allow tcp
```

Option B is wrong as the command would result in error.

ERROR: (gcloud.compute.firewall-rules.create) Exactly one of (-action | -allow) must be specified.

Option C is wrong as deny rule would prevent SSH login.

Option D is wrong as the port 3389 is for RDP and not for SSH.

#### 34. 34. Question

You're working on creating a script that can extract the IP address of a Kubernetes Service. Your coworker sent you a code snippet that they had saved. Which one is the best starting point for your code?

- A. kubectl get svc -o filtered-  
json='{"items":[]}.status.loadBalancer.ingress[0].ip'
- B. **kubectl get svc -o  
jsonpath='{"items":[]}.status.loadBalancer.ingress[0].ip'**
- C. kubectl get svc -o html
- D. kubectl get svc

**Unattempted**

Correct answer is B as kubectl get svc can be used to the data, and jsonpath can be used to the parse the data.

Refer Kubernetes documentation – Kubenetus IO & Tutorials

```
$ kubectl get services NAME CLUSTER-IP EXTERNAL-IP PORT(S) kubernetes
10.0.0.1 443/TCP bootcamp 10.3.245.61 104.155.111.170 8080/TCP
```

To access the services, use the external IP and the application port e.g. like this:

```
$ export EXTERNAL_IP=$(kubectl get service bootcamp -
output=jsonpath='{"status.loadBalancer.ingress[0].ip"}') $ export PORT=$(kubectl
get services --output=jsonpath='{"items[0].spec.ports[0].port"}') $ curl
"$EXTERNAL_IP:$PORT" Hello Kubernetes bootcamp! | Running on: bootcamp-
390780338-2fhnk | v=1
```

#### 35. 35. Question

Your team needs to set up a new Jenkins instance as quickly as possible. What's the best way to get it up-and-running?

- A. Use Google's Managed Jenkins Service.
- B. Deploy the jar file to a Compute Engine instance.



### C. Install with Cloud Launcher



- D. Create a Deployment Manager template and deploy it.

**Unattempted**

Correct answer is C as Cloud Launcher provides

Refer GCP documentation – Marketplace (Formerly Cloud Launcher)

GCP Marketplace offers ready-to-go development stacks, solutions, and services to accelerate development. So you spend less time installing and more time developing.

Deploy production-grade solutions in a few clicks

Single bill for all your GCP and 3rd party services

Manage solutions using Deployment Manager

Notifications when a security update is available

Direct access to partner support

Option A is wrong as there is no Google's Managed Jenkins Service.

Option B is wrong as hosting on the compute engine is still a manual step.

Option D is wrong as Deployment Manager would take time to build and deploy.

### 36. Question

You have a Cloud Storage bucket that needs to host static web assets with a dozen HTML pages, a few JavaScript files, and some CSS. How do you make the bucket public?



- A. Set allAuthenticatedUsers to have the Storage Object Viewer role.



- B. Check the make public box on the GCP Console for the bucket



- C. Set allUsers to have the Storage Object Viewer role.**



- D. gsutil make-public gs://bucket-name

**Unattempted**

Correct answer is C as the bucket can be shared by providing the Storage Object Viewer access to allUsers.

Refer GCP documentation – Cloud Storage Sharing files

You can either make all files in your bucket publicly accessible, or you can set individual objects to be accessible through your website. Generally, making all files in your bucket accessible is easier and faster.

To make all files accessible, follow the Cloud Storage guide for making groups of objects publicly readable.

To make individual files accessible, follow the Cloud Storage guide for making individual objects publicly readable.

If you choose to control the accessibility of individual files, you can set the default object ACL for your bucket so that subsequent files uploaded to your bucket are shared by default.

1. Open the Cloud Storage browser in the Google Cloud Platform Console.
2. In the list of buckets, click on the name of the bucket that contains the object you want to make public, and navigate to the object if it's in a subdirectory.
3. Click the drop-down menu associated with the object that you want to make public. The drop-down menu appears as three vertical dots to the far right of the object's row.

4. Select Edit permissions from the drop-down menu.  
5. In the overlay that appears, click the + Add item button.  
6. Add a permission for allUsers.  
Select User for the Entity.  
Enter allUsers for the Name.  
Select Reader for the Access.  
7. Click Save.  
Option A is wrong as access needs to be provided to allUsers to make it public and there is no allAuthenticatedUsers option.  
Option B is wrong as there is no make public option with GCP Console.  
Option D is wrong as there is no make public option with gsutil command.

### 37. Question

Your company has been running their marketing application on App Engine app for a few weeks with Autoscaling, and it's been performing well. However, the marketing team is planning on a massive campaign, and they expect a lot of burst traffic. How would you go about ensuring there are always 3 idle instances?

- A. Set the min\_instances property in the app.yaml
- B. Switch to manual scaling and use the burst\_traffic\_protection property to True in the app.yaml.
- C. Set the min\_idle\_instances property in the app.yaml.
- D. Switch to manual scaling and use the idle\_instance\_count property in the app.yaml.

### Unattempted

Correct answer is C as min\_idle\_instances property can be set to have minimum idle instances which would be always running.

Refer GCP documentation – App Engine Scaling & app.yaml Reference  
Auto scaling services use dynamic instances that get created based on request rate, response latencies, and other application metrics. However, if you specify a number of minimum idle instances, that specified number of instances run as resident instances while any additional instances are dynamic.

#### min\_idle\_instances

The number of instances to be kept running and ready to serve traffic. Note that you are charged for the number of instances specified whether they are receiving traffic or not. This setting only applies to the version that receives most of the traffic. Keep the following in mind:

A low minimum helps keep your running costs down during idle periods, but means that fewer instances might be immediately available to respond to a sudden load spike.

A high minimum allows you to prime the application for rapid spikes in request load. App Engine keeps the minimum number of instances running to serve incoming requests. You are charged for the number of instances specified, whether or not they are handling requests. For this feature to function properly, you must make sure that warmup requests are enabled and that your application handles warmup requests.

If you set a minimum number of idle instances, pending latency will have less effect on your application's performance. Because App Engine keeps idle

instances in reserve, it is unlikely that requests will enter the pending queue except in exceptionally high load spikes. You will need to test your application and expected traffic volume to determine the ideal number of instances to keep in reserve.

Option A is wrong as min\_instances applies to dynamic scaling. Also, number of instances

Options B & D are wrong as manual scaling would not provide the minimal running instances.

### 38. 38. Question

Your team has some new functionality that they want to roll out slowly so they can monitor for errors. The change contains some significant changes to the user interface. You've chosen to use traffic splitting to perform a canary deployment. You're going to start by rolling out the code to 15% of your users. How should you go about setting up traffic splitting with the user getting the same experience?

- A. Deploy the new version. Split the traffic using an IP or cookie based distribution.
- B. Use the gcloud app deploy command with the distribution flag to deploy and split the traffic in one command.
- C. Deploy the new version using the no-promote flag. Split the traffic using a random distribution.
- D. Deploy the new version using the no-promote flag. Split the traffic using Cookie.

**Unattempted**

Correct answer is D as the application needs to be promoted using the –no-promote parameter to avoid the new version getting all the 100% traffic. Once the application is deployed and tested, the traffic can be split using the Cookie approach to maintain User experience.

Refer GCP documentation – Splitting Traffic

When you have specified two or more versions for splitting, you must choose whether to split traffic by using either an IP address or HTTP cookie. It's easier to set up an IP address split, but a cookie split is more precise.

Options A & B are wrong as deploying the new version would configure it to receive all the traffic.

Option C is wrong as random distribution would not help maintain user experience.

### 39. 39. Question

Your company has decided to store data files in Cloud Storage. The data would be hosted in a regional bucket to start with. You need to configure Cloud Storage lifecycle rule to move the data for archival after 30 days and delete the data after a year. Which two actions should you take?

- A. Create a Cloud Storage lifecycle rule with Age: 30, Storage Class: Standard, and Action: Set to Coldline, and create a second GCS life-cycle rule with Age: 365, Storage Class: Coldline, and Action: Delete.
- B. Create a Cloud Storage lifecycle rule with Age: 30, Storage Class: Standard, and Action: Set to Coldline, and create a second GCS life-cycle rule with Age: 275, Storage Class: Coldline, and Action: Delete.
- C. Create a Cloud Storage lifecycle rule with Age: 30, Storage Class: Standard, and Action: Set to Nearline, and create a second GCS life-cycle rule with Age: 365, Storage Class: Nearline, and Action: Delete.
- D. Create a Cloud Storage lifecycle rule with Age: 30, Storage Class: Standard, and Action: Set to Nearline, and create a second GCS life-cycle rule with Age: 275, Storage Class: Nearline, and Action: Delete.

#### Unattempted

Correct answer is A as there are 2 actions needed. First archival after 30 days, which can be done by SetStorageClass action to Coldline. Second delete the data after an year, which can be done by delete action with Age 365 days. The Age condition is measured from the object's creation time.

Refer GCP documentation – Cloud Storage Lifecycle Management

Age: This condition is satisfied when an object reaches the specified age (in days). Age is measured from the object's creation time. For example, if an object's creation time is 2019/01/10 10:00 UTC and the Age condition is 10 days, then the condition is satisfied for the object on and after 2019/01/20 10:00 UTC. This is true even if the object becomes archived through object versioning sometime after its creation.

Option B is wrong as the Age needs to be set to 365 as its relative to the object creation date and not changed date.

Options C & D are wrong Nearline storage class is not an ideal storage class for archival

#### 40. Question

You've been tasked with getting all of your team's public SSH keys onto all of the instances of a particular project. You've collected them all. With the fewest steps possible, what is the simplest way to get the keys deployed?

- A. Add all of the keys into a file that's formatted according to the requirements. Use the gcloud compute instances add-metadata command to upload the keys to each instance
- B. Add all of the keys into a file that's formatted according to the requirements. Use the gcloud compute project-info add-metadata command to upload the keys.
- C. Use the gcloud compute ssh command to upload all the keys
- D. Format all of the keys as needed and then, using the user interface, upload each key one at a time.

#### Unattempted

Correct answer is B as project wide SSH keys can help provide users access to all the instances. The keys can be added or removed using the instance metadata.

Refer GCP documentation – Project wide SSH keys

Use project-wide public SSH keys to give users general access to a Linux instance. Project-wide public SSH keys give users access to all of the Linux instances in a project that allow project-wide public SSH keys. If an instance blocks project-wide public SSH keys, a user cannot use their project-wide public SSH key to connect to the instance unless the same public SSH key is also added to instance metadata.

`gcloud compute project-info add-metadata --metadata-from-file ssh-keys=[LIST_PATH]`

Option A is wrong as the gcloud compute instances provides only specific instance level access.

Option C is wrong as gcloud compute ssh is a thin wrapper around the ssh(1) command that takes care of authentication and the translation of the instance name into an IP address. It can be used to ssh to the instance.

Option D is wrong as there is no user interface to upload the keys.

#### 41. Question

Your developers have been thoroughly logging everything that happens in the API. The API allows end users to request the data as JSON, XML, CSV, and XLS. Supporting all of these formats is taking a lot of developer effort.

Management would like to start tracking which options are used over the next month. Without modifying the code, what's the fastest way to be able to report on this data at the end of the month?

- A. Create a custom counter logging metric that uses a regex to extract the data format into a label. At the end of the month, use the metric viewer to see the group by the label.
- B. Create a log sink that filters for rows that mention the data format. Export that to BigQuery, and run a query at the end of the month.
- C. Create a custom monitoring metric in code and edit the API code to set the metric each time the API is called.
- D. Export the logs to excel, and search for the different fields.

#### Unattempted

Correct answer is A as custom user defined log based metrics can be created on the logs already logged. These metrics can be used at the end of the month to check the stats for API call per format to gain insights.

Refer GCP documentation – Stackdriver logging – Log based metrics

User-defined (logs-based) metrics are created by a user on a project. They count the number of log entries that match a given filter, or keep track of particular values within the matching log entries.

Option B is wrong as the solution is possible but not the fastest as compared to log based metric.

Option C is wrong as it required a code change.

Option D is wrong as its more manual effort and not scalable.

#### 42. 42. Question

You've created a new firewall rule to allow incoming traffic on port 22, using a target tag of "dev-ssh". You tried to connect to one of your instances, and you're still unable to connect. What steps do you need to take to resolve the problem?

- A. Run the gcloud firewall-rules refresh command, as they need to be reloaded.
- B. Use source tags in place of the target tags.
- C. Reboot the instances for the firewall rule to take effect.
- D. Apply a network tag of dev-ssh to the instance you're trying to connect into and test again.

**Unattempted**

Correct answer is D as the firewall needs to be associated with the instance for the instance to follow the firewall rules. The association can be performed by applying the network tag "dev-ssh" to the instance.

Refer GCP documentation – VPC Network Tags

Network tags are text attributes you can add to Compute Engine virtual machine (VM) instances. Tags allow you to make firewall rules and routes applicable to specific VM instances.

You can only add network tags to VM instances or instance templates. You cannot tag other GCP resources. You can assign network tags to new instances at creation time, or you can edit the set of assigned tags at any time later.

Network tags can be edited without stopping an instance.

Option A is wrong as firewalls will associate through network tags reflect immediately and do not require any refresh.

Option B is wrong as Firewall needs to associate with target tags, which dictate the instances.

Option C is wrong as instances do not need to be rebooted and it's at the network level with no changes in the instances.

#### 43. 43. Question

You're migrating an on-premises application to Google Cloud. The application uses a component that requires a licensing server. The license server has the IP address 10.28.0.10. You want to deploy the application without making any changes to the code or configuration. How should you go about deploying the application?

- A. Create a subnet with a CIDR range of 10.28.0.0/29. Reserve a static internal IP address of 10.28.0.10. Assign the static address to the license server instance.
- B. Create a subnet with a CIDR range of 10.28.0.0/28. Reserve a static external IP address of 10.28.0.10. Assign the static address to the license server instance.
- C. Create a subnet with a CIDR range of 10.28.0.0/10. Reserve a static external IP address of 10.28.0.10. Assign the static address to the license server instance.

- D. Create a subnet with a CIDR range of 10.28.0.0/28. Reserve a static internal IP address of 10.28.0.10. Assign the static address to the license server instance.

**Unattempted**

Correct answer is D as the IP is internal it can be reserved using the static internal IP address, which blocks it and prevents it from getting allocated to other resource.

Refer GCP documentation – Compute Network Addresses

In Compute Engine, each VM instance can have multiple network interfaces.

Each interface can have one external IP address, one primary internal IP address, and one or more secondary internal IP addresses. Forwarding rules can have external IP addresses for external load balancing or internal addresses for internal load balancing.

Static internal IPs provide the ability to reserve internal IP addresses from the private RFC 1918 IP range configured in the subnet, then assign those reserved internal addresses to resources as needed. Reserving an internal IP address takes that address out of the dynamic allocation pool and prevents it from being used for automatic allocations. Reserving static internal IP addresses requires specific IAM permissions so that only authorized users can reserve a static internal IP address.

Option A is wrong as the 10.28.0.0/29 CIDR provides only 8 IP addresses and would not include 10.28.0.10.

Options B & C are wrong as the IP address is RFC 1918 address and needs to be an internal static IP address.

#### 44. Question

You've been running App Engine applications in a Standard Environment for a few weeks. With several successful deployments, you've just deployed a broken version, and the developers have gone home for the day. What is the fastest way to get the site back into a functioning state?

- A. Use the gcloud app deployments revert command.
- B. Use the gcloud app deployments rollback command.
- C. In GCP console, click Traffic Splitting and direct 100% of the traffic to the previous version.
- D. In GCP console, click the Rollback button on the versions page.

**Unattempted**

Correct answer is C as the best approach is the revert by the traffic to a previous deployed version.

Refer GCP documentation – Migrating & Splitting Traffic

Traffic splitting distributes a percentage of traffic to versions of your application. You can split traffic to move 100% of traffic to a single version or to route percentages of traffic to multiple versions. Splitting traffic to two or more versions allows you to conduct A/B testing between your versions and provides control over the pace when rolling out features.

Options A & B are as gcloud app command does not provide rollback and revert

feature

Option D is wrong as GCP console does not provide the ability to rollback.

45. Question

You have a 20 GB file that you need to securely share with some contractors. They need it as fast as possible. Which steps would get them the file quickly and securely?

- A. Set up a VPC with a custom subnet. Create a subnet tunnel. Upload the file to a network share. Grant the contractors temporary access.
- B. Using composite objects and parallel uploads to upload the file to Cloud Storage quickly. Then generate a signed URL and securely share it with the contractors.
- C. Upload the file to Bigtable using the bulk data import tool. Then provide the contractors with read access to the database.
- D. Upload the file to Cloud Storage. Grant the allAuthenticated users token view permissions.

Unattempted

Correct answer is B as the composite parallel upload can help upload the file quickly to Cloud Storage. Signed urls can be used to quickly and securely share the files with third party.

Refer GCP documentation – Cloud Storage Signed URLs

Signed URLs provide a way to give time-limited read or write access to anyone in possession of the URL, regardless of whether they have a Google account. In some scenarios, you might not want to require your users to have a Google account in order to access Cloud Storage, but you still want to control access using your application-specific logic. The typical way to address this use case is to provide a signed URL to a user, which gives the user read, write, or delete access to that resource for a limited time. Anyone who knows the URL can access the resource until the URL expires. You specify the expiration time in the query string to be signed.

Option A is wrong as it is not a quick solution, but a cumbersome solution.

Option C is wrong as Bigtable is not an ideal storage for files.

Option D is wrong as All Authenticated access would provide access to anyone who is authenticated with a Google account. The special scope identifier for all Google account holders is allAuthenticatedUser

46. Question

You're using a self-serve Billing Account to pay for your 2 projects. Your billing threshold is set to \$1000.00 and between the two projects you're spending roughly 50 dollars per day. It has been 18 days since you were last charged. Given the above data, when will you likely be charged next?

- A. On the first day of the next month.
- B. In 2 days when you'll hit your billing threshold.

- C. On the thirtieth day of the month.
- D. In 12 days, making it 30 days since the previous payment.

**Unattempted**

Correct answer is B as the billing is either monthly or the threshold, whichever comes first. As with average \$50 per day and 18 days passed the \$1000 threshold would hit in 2 days and so would be the billing.

Refer GCP documentation – Cloud Storage Billing

Your costs are charged automatically in one of two ways, whichever comes first:  
A regular monthly cycle (monthly billing)

When your account has accrued a certain amount of charges (threshold billing)  
Options A & D are wrong as the billing would not be triggered in 12 days as the threshold would be hit first.

Option C is wrong as there is no such fixed date.

#### 47. Question

Your company has created a new billing account and needs to move the projects to the billing account. What roles are needed to change the billing account?  
(Select two)

- A. Project Billing manager
- B. Project Owner
- C. Billing Account Billing administrator
- D. Billing Account Manager
- E. Project Editor

**Unattempted**

Correct answers are B & C as To change the billing account for an existing project, you must be an owner on the project and a billing administrator on the destination billing account.

Refer GCP documentation – Project Change Billing Account

#### 48. Question

You have deployed an application using Deployment manager. You want to update the deployment with minimal downtime. How can you achieve the same?

- A. gcloud deployment-manager deployments create
- B. gcloud deployment-manager deployments update
- C. gcloud deployment-manager resources create
- D. gcloud deployment-manager resources update

**Unattempted**

Correct answer is B as gcloud deployment-manager deployments update can be used to update the existing deployment.

Refer GCP documentation – Deployment Manager Update Deployment  
After you have created a deployment, you can update it as your application or service changes. You can use Deployment Manager to update a deployment by:  
Adding or removing resources from a deployment.

Updating the properties of existing resources in a deployment.

A single update can contain any combination of these changes. For example, you can make changes to the properties of existing resources and add new resources in the same request. You update your deployment by following these steps:

1. Make changes to or create a configuration file with the changes you want.
2. Optionally, pick the policies to use for your updates or use the default policies.
3. Make the update request to Deployment Manager.

`gcloud deployment-manager deployments update example-deployment`

Option A is wrong as `gcloud deployment-manager deployments create` is used to create deployment.

Options C & D are wrong as `resources` is not a valid parameter.

#### 49. Question

You did a deployment for App Engine using `gcloud app deploy`. However, checking the intended project you do not find the deployment and seems the application was deployed in the wrong project. How do you find out which project the application was deployed to?

- A. Check `app.yaml` for the project
- B. Check `application web.xml` for the project
- C. Run `gcloud config list` to check for the project
- D. Check `index.yaml` for the project

**Unattempted**

Correct answer is C as By default, the `deploy` command generates a unique ID for the version that you deploy, deploys the version to the GCP project you configured the `gcloud` tool to use, and routes all traffic to the new version. The project can be checked using the `gcloud config list` command.

Refer GCP documentation – App Engine Deploying Application

`gcloud app deploy app.yaml index.yaml`

Optional flags:

Include the `--project` flag to specify an alternate GCP Console project ID to what you initialized as the default in the `gcloud` tool. Example: `--project [YOUR_PROJECT_ID]`

Include the `-v` flag to specify a version ID, otherwise one is generated for you.

Example: `-v [YOUR_VERSION_ID]`

Options A, B & D are wrong as they do not provide the ability to set the project.

#### 50. Question

Your company has appointed external auditors for auditing the security of your setup. They want to check all the users and roles configured. What would be the best way to check the users and roles?

- A. Ask auditors to check using gcloud iam roles list command
- B. Ask auditors to check using gcloud iam service-accounts list command
- C. Ask Auditors to navigate to the IAM page and check member and roles section
- D. Ask Auditors to navigate to the IAM page section and check roles and status section

#### Unattempted

Correct answer is C as the auditor can check all the members and roles created for the project from the IAM page listing the members and roles.

Option A is wrong as the gcloud iam roles list command would only list roles.

Option B is wrong as the gcloud iam service-accounts list command would only list services accounts.

Option D is wrong as the roles menu only displays the predefined or custom roles and their status.

#### 51. Question

Your project manager wants to delegate the responsibility to manage files and buckets for Cloud Storage to his team members. Considering the principle of least privilege, which role should you assign to the team members?

- A. roles/storage.objectAdmin
- B. roles/storage.admin
- C. roles/storage.objectCreator
- D. roles/owner

#### Unattempted

Correct answer is B as roles/storage.admin would provide the team members full control of buckets and objects. When applied to an individual bucket, control applies only to the specified bucket and objects within the bucket.

Refer GCP documentation – Cloud Storage IAM Roles

Options A & C are wrong as they do not provide sufficient privileges to manage buckets.

Option D is wrong as it provides more privileges than required.

#### 52. Question

Your company is designing an application, which would interact with Cloud Spanner. The application should have the ability to view and edit Cloud Spanner tables. Considering the principle of least privilege, which role should you assign to the team members?

- A. roles/spanner.viewer
- B. roles/spanner.databaseUser

- C. roles/spanner.databaseReader
- D. roles/spanner.databaseAdmin

### Unattempted

Correct answer is B as roles/spanner.databaseUser is a machine only roles and provides the ability to read and write to database.

Recommended to grant at the databaselevel. A principal with this role can:

Read from and write to the Cloud Spanner database.

Execute SQL queries on the database, including DML and Partitioned DML.

View and update schema for the database.

Refer GCP documentation – Spanner IAM Roles

Options A & D are wrong as they are person role and either provide more or less privileges than required.

Option C is wrong as it provides only read permissions.

### 53. Question

A Company is using Cloud SQL to host critical data. They want to enable Point In Time recovery (PIT) to be able to recover the instance to a specific point in. How should you configure the same?

- A. Create a Read replica for the instance
- B. Switch to Spanner 3 node cluster
- C. Create a Failover replica for the instance
- D. Enable Binary logging and backups for the instance

### Unattempted

Correct answer is D as for performing Point In Time recovery for the Cloud SQL, you should enabled backups and binary logging.

Refer GCP documentation – Cloud SQL Point In Time Recovery

Point-in-time recovery enables you to recover an instance to a specific point in time. A point-in-time recovery always creates a new instance; you cannot perform a point-in-time recovery to an existing instance.

Before completing this task, you must have:

Binary logging and backups enabled for the instance, with continuous binary logs since the last backup before the event you want to recover from. For more information, see Enabling binary logging.

A binary log file name and the position of the event you want to recover from (that event and all events that came after it will not be reflected in the new instance).

Options A & C are wrong Read and Failover replicas do not aid in Point In Recovery.

Option B is wrong as it is not required to switch to Cloud Spanner.

### 54. Question

Your organization requires that log from all applications be archived for 10 years as a part of compliance. Which approach should you use?

- A. Configure Stackdriver Monitoring for all Projects, and export to BigQuery
- B. Configure Stackdriver Monitoring for all Projects with the default retention policies
- C. Configure Stackdriver Monitoring for all Projects, and export to Google Cloud Storage
- D. Grant the security team access to the logs in each Project

**Unattempted**

Correct answer is C as Stackdriver monitoring metrics can be exported to BigQuery or Google Cloud Storage. As the logs need to be archived, GCS is a better option.

Refer GCP documentation – Stackdriver

Stackdriver Logging provides you with the ability to filter, search, and view logs from your cloud and open source application services. Allows you to define metrics based on log contents that are incorporated into dashboards and alerts. Enables you to export logs to BigQuery, Google Cloud Storage, and Pub/Sub. Option A is wrong as BigQuery would be a better storage option for analytics capability.

Option B is wrong as Stackdriver cannot retain data for 5 year. Refer Stackdriver data retention

Option D is wrong as project logs are maintained in Stackdriver and it has limited data retention capability.

#### 55. Question

You are running an application in Google App Engine that is serving production traffic. You want to deploy a risky but necessary change to the application. It could take down your service if not properly coded. During development of the application, you realized that it can only be properly tested by live user traffic. How should you test the feature?

- A. Deploy the new application version temporarily, and then roll it back.
- B. Create a second project with the new app in isolation, and onboard users.
- C. Set up a second Google App Engine service, and then update a subset of clients to hit the new service.
- D. Deploy a new version of the application, and use traffic splitting to send a small percentage of traffic to it.

**Unattempted**

Correct answer is D as deploying a new version without assigning it as the default version will not create downtime for the application. Using traffic splitting allows for easily redirecting a small amount of traffic to the new version and can also be quickly reverted without application downtime.

Refer GCP documentation – App Engine Splitting Traffic

Traffic migration smoothly switches request routing, gradually moving traffic from

the versions currently receiving traffic to one or more versions that you specify. Traffic splitting distributes a percentage of traffic to versions of your application. You can split traffic to move 100% of traffic to a single version or to route percentages of traffic to multiple versions. Splitting traffic to two or more versions allows you to conduct A/B testing between your versions and provides control over the pace when rolling out features.

Option A is wrong as deploying the application version as default requires moving all traffic to the new version. This could impact all users and disable the service. Option B is wrong as deploying a second project requires data synchronization and having an external traffic splitting solution to direct traffic to the new application. While this is possible, with Google App Engine, these manual steps are not required.

Option C is wrong as App Engine services are intended for hosting different service logic. Using different services would require manual configuration of the consumers of services to be aware of the deployment process and manage from the consumer side who is accessing which service.

#### 56. Question

Using principle of least privilege and allowing for maximum automation, what steps can you take to store audit logs for long-term access and to allow access for external auditors to view? (Choose two)

- A. Generate a signed URL to the Stackdriver export destination for auditors to access.**
- B. Create an account for auditors to have view access to Stackdriver Logging.
- C. Export audit logs to Cloud Storage via an export sink.**
- D. Export audit logs to BigQuery via an export sink.

**Unattempted**

Correct answers are A & C as Stackdriver logging allows export to Cloud Storage which can be used for long term access and exposed to external auditors using signed urls.

Refer GCP documentation – Stackdriver logging export

Stackdriver Logging provides an operational datastore for logs and provides rich export capabilities. You might export your logs for several reasons, such as retaining logs for long-term storage (months or years) to meet compliance requirements or for running data analytics against the metrics extracted from the logs. Stackdriver Logging can export to Cloud Storage, BigQuery, and Cloud Pub/Sub.

Option B is wrong as Stackdriver logging does not support long term retention of logs

Option D is wrong as BigQuery can be used to export logs and retain for long term, however the access can be provided to only GCP users and not external auditors.

#### 57. Question

You created an update for your application on App Engine. You want to deploy the update without impacting your users. You want to be able to roll back as quickly as possible if it fails. What should you do?

- A. Delete the current version of your application. Deploy the update using the same version identifier as the deleted version.
- B. Notify your users of an upcoming maintenance window. Deploy the update in that maintenance window.
- C. Deploy the update as the same version that is currently running.
- D. Deploy the update as a new version. Migrate traffic from the current version to the new version.

**Unattempted**

Correct answer is D as the deployment can be done seamlessly by deploying a new version and migrating the traffic gradually from the old version to the new version. If any issue is encountered, the traffic can be migrated 100% to the old version.

Refer GCP documentation – App Engine Migrating Traffic

Manage how much traffic is received by a version of your application by migrating or splitting traffic.

Traffic migration smoothly switches request routing, gradually moving traffic from the versions currently receiving traffic to one or more versions that you specify. Traffic splitting distributes a percentage of traffic to versions of your application. You can split traffic to move 100% of traffic to a single version or to route percentages of traffic to multiple versions. Splitting traffic to two or more versions allows you to conduct A/B testing between your versions and provides control over the pace when rolling out features.

Options A & B are wrong as there is a downtime involved.

Option C is wrong as it would not allow an easier rollback in case of any issues.

58. **58. Question**

Using the principle of least privilege, your colleague Bob needs to be able to create new instances on Compute Engine in project ‘Project A’. How should you give him access without giving more permissions than is necessary?

- A. Give Bob Compute Engine Instance Admin Role for Project A.
- B. Give Bob Compute Engine Admin Role for Project A.
- C. Create a shared VPC that Bob can access Compute resources from.
- D. Give Bob Project Editor IAM role for Project A.

**Unattempted**

Correct answer is A as the access needs to be given only to create instances, the user should be given compute instance admin role, which provides the least privilege.

Refer GCP documentation – Compute IAM

roles/compute.instanceAdmin.v1

roles/compute.admin

Options B & D are wrong as it gives more permission than required

Option C is wrong as shared VPC does not give permissions to create instances to the user.

## 59. Question

You need to create a new Kubernetes Cluster on Google Cloud Platform that can autoscale the number of worker nodes. What should you do?

- A. Create a cluster on Kubernetes Engine and enable autoscaling on Kubernetes Engine.
- B. Create a cluster on Kubernetes Engine and enable autoscaling on the instance group of the cluster.
- C. Configure a Compute Engine instance as a worker and add it to an unmanaged instance group. Add a load balancer to the instance group and rely on the load balancer to create additional Compute Engine instances when needed.
- D. Create Compute Engine instances for the workers and the master and install Kubernetes. Rely on Kubernetes to create additional Compute Engine instances when needed.

**Unattempted**

Correct answer is A as Kubernetes cluster provides auto scaling feature which can be enabled on the cluster engine.

Refer GCP documentation – Kubernetes Cluster Autoscaler

GKE's cluster autoscaler automatically resizes clusters based on the demands of the workloads you want to run. With autoscaling enabled, GKE automatically adds a new node to your cluster if you've created new Pods that don't have enough capacity to run; conversely, if a node in your cluster is underutilized and its Pods can be run on other nodes, GKE can delete the node.

Cluster autoscaling allows you to pay only for resources that are needed at any given moment, and to automatically get additional resources when demand increases.

Option B is wrong as auto scaling is not configured on instance group.

Option C is wrong as unmanaged group cannot be scaled.

Option D is wrong as you don't manage kubernetes using compute engine.

## 60. Question

You are creating a solution to remove backup files older than 90 days from your backup Cloud Storage bucket. You want to optimize ongoing Cloud Storage spend. What should you do?

- A. Write a lifecycle management rule in XML and push it to the bucket with gsutil
- B. Write a lifecycle management rule in JSON and push it to the bucket with gsutil

- C. Schedule a cron script using gsutil ls -lr gs://backups/\*\* to find and remove items older than 90 days
- D. Schedule a cron script using gsutil ls -l gs://backups/\*\* to find and remove items older than 90 days and schedule it with cron

### Unattempted

Correct answer is B as the object lifecycle in Cloud Storage can be automatically controlled using a JSON document defining the rules.

Refer GCP documentation – gsutil lifecycle

Sets the lifecycle configuration on one or more buckets. The config-json-file specified on the command line should be a path to a local file containing the lifecycle configuration JSON document.

Option A is wrong as XML is not supported by the gsutil command. It works with direct REST APIs only.

Options C & D are wrong as it is quite cumbersome to list the objects, calculate the age and then delete the objects.

## 61. Question

You are working on a project with two compliance requirements. The first requirement states that your developers should be able to see the Google Cloud Platform billing charges for only their own projects. The second requirement states that your finance team members can set budgets and view the current charges for all projects in the organization. The finance team should not be able to view the project contents. You want to set permissions. What should you do?

- A. Add the finance team members to the default IAM Owner role. Add the developers to a custom role that allows them to see their own spend only.
- B. Add the finance team members to the Billing Administrator role for each of the billing accounts that they need to manage. Add the developers to the Viewer role for the Project.
- C. Add the developers and finance managers to the Viewer role for the Project.
- D. Add the finance team to the Viewer role for the Project. Add the developers to the Security Reviewer role for each of the billing accounts.

### Unattempted

Correct answer is B as there are 2 requirements, Finance team able to set budgets on project but not view project contents and developers able to only view billing charges of their projects. Finance with Billing Administrator role can set budgets and Developer with viewer role can view billing charges aligning with the principle of least privileges.

Refer GCP documentation – IAM Billing

Option A is wrong as GCP recommends using pre-defined roles instead of using primitive roles and custom roles.

Option C is wrong as viewer role to finance would not provide them the ability to set budgets.

Option D is wrong as viewer role to finance would not provide them the ability to

set budgets. Also, Security Reviewer role enables the ability to view custom roles but not administer them for the developers which they don't need.

62. 62. Question

Using principal of least privilege and allowing for maximum automation, what steps can you take to store audit logs for long-term access and to allow access for external auditors to view? (Select Two)

- A. Create account for auditors to have view access to Stackdriver Logging.
- B. Export audit logs to Cloud Storage via an export sink.
- C. Export audit logs to BigQuery via an export sink.
- D. Create account for auditors to have view access to export storage bucket with the Storage Object Viewer role.

Unattempted

Correct answers are B & D as the best approach for providing long term access with least privilege would be to store the data in Cloud Storage and provide the Storage Object viewer role.

Refer GCP documentation – Stackdriver Logging Export

Exporting involves writing a filter that selects the log entries you want to export, and choosing a destination in Cloud Storage, BigQuery, or Cloud Pub/Sub. The filter and destination are held in an object called a sink. Sinks can be created in projects, organizations, folders, and billing accounts.

roles/storage.objectViewer

Can also list the objects in a bucket.

Option A is wrong as Stackdriver does not provide long term data retention.

Option C is wrong as the data can be stored in BigQuery, however if it is required for analysis. Also the users need to be given limited access to the dataset, which is missing.

63. 63. Question

Your company has a set of compute engine instances that would be hosting production-based applications. These applications would be running 24x7 throughout the year. You need to implement the cost-effective, scalable and high availability solution even if a zone fails. How would you design the solution?

- A. Use Managed instance groups with preemptible instances across multiple zones
- B. Use Managed instance groups across multiple zones
- C. Use managed instance groups with instances in a single zone
- D. Use Unmanaged instance groups across multiple zones

Unattempted

Correct answer is B as it would provide a highly available solution in case a zone goes down and managed instance groups would provide the scalability.

Refer GCP documentation – Managed Instance Groups

A managed instance group uses an instance template to create a group of identical instances. You control a managed instance group as a single entity. If you wanted to make changes to instances that are part of a managed instance group, you would make the change to the whole instance group. Because managed instance groups contain identical instances, they offer the following features.

When your applications require additional compute resources, managed instance groups can automatically scale the number of instances in the group.

Managed instance groups work with load balancing services to distribute traffic to all of the instances in the group.

If an instance in the group stops, crashes, or is deleted by an action other than the instance groups commands, the managed instance group automatically recreates the instance so it can resume its processing tasks. The recreated instance uses the same name and the same instance template as the previous instance, even if the group references a different instance template.

Managed instance groups can automatically identify and recreate unhealthy instances in a group to ensure that all of the instances are running optimally.

The managed instance group updater allows you to easily deploy new versions of software to instances in your managed instance groups, while controlling the speed and scope of deployment as well as the level of disruption to your service.

Option A is wrong as preemptible instances, although cost-effective, are not suitable for production load.

Option C is wrong as deployment in a single zone does not provide high availability.

Option D is wrong as unmanaged instance group does not provide scalability. Unmanaged instance groups are groups of dissimilar instances that you can arbitrarily add and remove from the group. Unmanaged instance groups do not offer autoscaling, rolling update support, or the use of instance templates so Google recommends creating managed instance groups whenever possible. Use unmanaged instance groups only if you need to apply load balancing to your pre-existing configurations or to groups of dissimilar instances.

#### 64. 64. Question

Your company wants to reduce cost on infrequently accessed data by moving it to the cloud. The data will still be accessed approximately once a month to refresh historical charts. In addition, data older than 5 years is no longer needed. How should you store and manage the data?

- A. In Google Cloud Storage and stored in a Multi-Regional bucket. Set an Object Lifecycle Management policy to delete data older than 5 years.
- B. In Google Cloud Storage and stored in a Multi-Regional bucket. Set an Object Lifecycle Management policy to change the storage class to Coldline for data older than 5 years.
- C. In Google Cloud Storage and stored in a Nearline bucket. Set an Object Lifecycle Management policy to delete data older than 5 years.

- D. In Google Cloud Storage and stored in a Nearline bucket. Set an Object Lifecycle Management policy to change the storage class to Coldline for data older than 5 years.

#### Unattempted

Correct answer is C as the access pattern fits Nearline storage class requirements and Nearline is a more cost-effective storage approach than Multi-Regional. The object lifecycle management policy to delete data is correct versus changing the storage class to Coldline as the data is no longer needed.

Refer GCP documentation – Cloud Storage – Storage Classes

Options A & B are wrong as Multi-Regional storage class is not an ideal storage option with infrequent access.

Option D is wrong as changing the storage class to Coldline is incorrect as the data is no longer required after 5 years.

#### 65. 65. Question

You are creating a single preemptible VM instance named “preempt” to be used as scratch space for a single workload. If your VM is preempted, you need to ensure that disk contents can be re-used. Which gcloud command would you use to create this instance?

- A. gcloud compute instances create preempt --preemptible --no-boot-disk-auto-delete
- B. gcloud compute instances create preempt --preemptible --boot-disk-auto-delete=no
- C. gcloud compute instances create preempt --preemptible
- D. gcloud compute instances create preempt --no-auto-delete

#### Unattempted

Correct answer is A as the preemptible instances need to be created you need to pass the –preemptible flag and as disk contents need not be deleted, –no-boot-disk-auto-delete flag needs to be passed.

Refer GCP documentation – Command line

–boot-disk-auto-delete : Automatically delete boot disks when their instances are deleted. Enabled by default, use –no-boot-disk-auto-delete to disable.

–preemptible : If provided, instances will be preemptible and time-limited.

Instances may be preempted to free up resources for standard VM instances, and will only be able to run for a limited amount of time. Preemptible instances can not be restarted and will not migrate.

Option B is wrong as the parameter for disk retention is wrong.

Option C is wrong as the disk would be deleted when the instance terminates.

Option D is wrong as it would not create a preemptible instance.

#### 66. 66. Question

You have a definition for an instance template that contains a web application. You are asked to deploy the application so that it can scale based on the HTTP traffic it receives. What should you do?

- A. Create a VM from the instance template. Create a custom image from the VM's disk. Export the image to Cloud Storage. Create an HTTP load balancer and add the Cloud Storage bucket as its backend service.
- B. Create an unmanaged instance group based on the instance template. Configure autoscaling based on HTTP traffic and configure the instance group as the backend service of an HTTP load balancer.
- C. Create a managed instance group based on the instance template. Configure autoscaling based on HTTP traffic and configure the instance group as the backend service of an HTTP load balancer.
- D. Create the necessary number of instances required for peak user traffic based on the instance template. Create an unmanaged instance group and add the instances to that instance group. Configure the instance group as the Backend Service of an HTTP load balancer.

### Unattempted

Correct answer is C as the instance template can be used with the managed instance group to define autoscaling to scale as per demand, which can then be exposed through a load balancer as a backend service

Refer GCP documentation – Load Balancing & Autoscaling

Google Cloud Platform (GCP) offers load balancing and autoscaling for groups of instances.

GCP offers server-side load balancing so you can distribute incoming traffic across multiple virtual machine instances. Load balancing provides the following benefits:

Scale your application

Support heavy traffic

Detect and automatically remove unhealthy virtual machine instances using health checks. Instances that become healthy again are automatically re-added.

Route traffic to the closest virtual machine

Compute Engine offers autoscaling to automatically add or remove virtual machines from an instance group based on increases or decreases in load. This allows your applications to gracefully handle increases in traffic and reduces cost when the need for resources is lower. You just define the autoscaling policy and the autoscaler performs automatic scaling based on the measured load.

Option A is wrong as the application is not exposed but only the static image.

Option B is wrong as instance template cannot be used with an unmanaged instance group for scaling.

Option D is wrong as unmanaged instance groups do not offer autoscaling.

### 67. Question

A Company is using Cloud SQL to host critical data. They want to enable high availability in case a complete zone goes down. How should you configure the same?

- A. Create a Read replica in the same region different zone
- B. Create a Read replica in the different region different zone



**C. Create a Failover replica in the same region different zone**



D. Create a Failover replica in the different region different zone

**Unattempted**

Correct answer is C as a failover replica helps provides High Availability for Cloud SQL. The failover replica must be in the same region as the primary instance.

Refer GCP documentation – Cloud SQL High Availability

The HA configuration, sometimes called a cluster, provides data redundancy. The configuration is made up of a primary instance (master) in the primary zone and a failover replica in the secondary zone. Through semisynchronous replication, all changes made to the primary instance's data and user tables are copied onto the failover replica. In the event of an instance or zone failure, this configuration reduces downtime, and your data continues to be available to client applications. The failover replica must be in the same region as the primary instance, but in a different zone.

Diagram overview of MySQL HA configuration. Described in text below.

Option A & B are wrong as Read replicas do not provide failover capability and just additional read capacity.

Option D is wrong as failover replica must be in the same region as the primary instance.

#### 68. Question

You're writing a Python application and want your application to run in a sandboxed managed environment with the ability to scale up in seconds to account for huge spikes in demand. Which service should you host your application on?



A. Compute Engine



B. App Engine Flexible Environment



C. Kubernetes Engine



**D. App Engine Standard Environment**

**Unattempted**

Correct answer is D as the App Engine Standard Environment provides rapid scaling as compared to App Engine Flexible Environment and is ideal for applications requiring quick start times and handle sudden and extreme spikes. Refer GCP documentation – App Engine Environments

When to choose the standard environment

Application instances run in a sandbox, using the runtime environment of a supported language listed below.

Applications that need to deal with rapid scaling.

Experiences sudden and extreme spikes of traffic which require immediate scaling.

When to choose the flexible environment

Application instances run within Docker containers on Compute Engine virtual machines (VM).

Applications that receive consistent traffic, experience regular traffic fluctuations, or meet the parameters for scaling up and down gradually.

69. 69. Question

You are a project owner and need your co-worker to deploy a new version of your application to App Engine. You want to follow Google's recommended practices. Which IAM roles should you grant your co-worker?

- A. Project Editor
- B. App Engine Service Admin
- C. App Engine Deployer
- D. App Engine Code Viewer

**Unattempted**

Correct answer is C as App Engine Deployer gives write access only to create a new version.

Refer GCP documentation – App Engine Access Control

App Engine Deployer

/roles/appengine.deployer

Read-only access to all application configuration and settings.

Write access only to create a new version; cannot modify existing versions other than deleting versions that are not receiving traffic. Cannot configure traffic to a version.

Option A is wrong as this access is too wide, and Google recommends least-privilege. Also Google recommends predefined roles instead of primitive roles like Project Editor.

Option B is wrong as is not correct because although it gives write access to module-level and version-level settings, users cannot deploy a new version.

Option D is wrong as is not correct because this is read-only access.

70. 70. Question

You developed a new application for App Engine and are ready to deploy it to production. You need to estimate the costs of running your application on Google Cloud Platform as accurately as possible. What should you do?

- A. Create a YAML file with the expected usage. Pass this file to the gcloud app estimate command to get an accurate estimation.
- B. Multiply the costs of your application when it was in development by the number of expected users to get an accurate estimation.
- C. Use the pricing calculator for App Engine to get an accurate estimation of the expected charges.
- D. Create a ticket with Google Cloud Billing Support to get an accurate estimation.

**Unattempted**

Correct answer is C as this is the proper way to estimate charges.

Refer GCP documentation – GCP Price Calculator

Option A is wrong as that command will generate an error and not give you an estimation on workloads.

Option B is wrong as this does not result in an accurate estimation.  
Option D is wrong as billing support is available to help you set up billing and understand invoices, not to make estimations.

## SET-5

### 1. Question

You are creating a Kubernetes Engine cluster to deploy multiple pods inside the cluster. All container logs must be stored in BigQuery for later analysis. You want to follow Google-recommended practices. Which two approaches can you take?

- A. Turn on Stackdriver Logging during the Kubernetes Engine cluster creation.
- B. Turn on Stackdriver Monitoring during the Kubernetes Engine cluster creation.
- C. Develop a custom add-on that uses Cloud Logging API and BigQuery API. Deploy the add-on to your Kubernetes Engine cluster.
- D. Use the Stackdriver Logging export feature to create a sink to Cloud Storage. Create a Cloud Dataflow job that imports log files from Cloud Storage to BigQuery.
- E. Use the Stackdriver Logging export feature to create a sink to BigQuery. Specify a filter expression to export log records related to your Kubernetes Engine cluster only.

Unattempted

Correct answers are A & E

Option A as creating a cluster with Stackdriver Logging option will enable all the container logs to be stored in Stackdriver Logging.

Option E as Stackdriver Logging support exporting logs to BigQuery by creating sinks

Refer GCP documentation – Kubernetes logging

Option B is wrong as creating a cluster with Stackdriver Monitoring option will enable monitoring metrics to be gathered, but it has nothing to do with logging.

Option C is wrong as even if you can develop a Kubernetes addon that will send logs to BigQuery, this is not a Google-recommended practice.

Option D is wrong as this is not a Google recommended practice.

### 2. Question

Your company has a mission-critical application that serves users globally. You need to select a transactional and relational data storage system for this application. Which two products should you choose?

- A. BigQuery
- B. Cloud SQL
- C. Cloud Spanner

- D. Cloud Bigtable
- E. Cloud Datastore

**Unattempted**

Correct answers are B & C

Option B as because Cloud SQL is a relational and transactional database in the list.

Option C as Spanner is a relational and transactional database in the list.

Refer GCP documentation – Storage Options

Option A is wrong as BigQuery is not a transactional system.

Option D is wrong as Cloud Bigtable provides transactional support but it's not relational.

Option E is wrong as Datastore is not a relational data storage system.

3. 3. Question

You want to find out who in your organization has Owner access to a project called "my-project". What should you do?

- A. In the Google Cloud Platform Console, go to the IAM page for your organization and apply the filter Role:Owner.
- B. In the Google Cloud Platform Console, go to the IAM page for your project and apply the filter Role:Owner.
- C. Use gcloud iam list-grantable-role --project my-project from your Terminal.
- D. Use gcloud iam list-grantable-role from Cloud Shell on the project page.

**Unattempted**

Correct answer is B as this shows you the Owners of the project.

Option A is wrong as it will give the org-wide owners, but you are interested in the project owners, which could be different.

Option C is wrong as this command is to list grantable roles for a resource, but does not return who has a specific role.

Option D is wrong as this command is to list grantable roles for a resource, but does not return who has a specific role.

4. 4. Question

You need to verify the assigned permissions in a custom IAM role. What should you do

- A. Use the GCP Console, IAM section to view the information.
- B. Use the gcloud init command to view the information.
- C. Use the GCP Console, Security section to view the information.
- D. Use the GCP Console, API section to view the information.

### Unattempted

Correct answer is A as this is the correct console area to view permission assigned to a custom role in a particular project.

Refer GCP documentation – IAM Custom Rules

Option B is wrong as gcloud init will not provide the information required.

Options C and D are wrong as these are not the correct areas to view this information

### 5. Question

You have an App Engine application serving as your front-end. It's going to publish messages to Pub/Sub. The Pub/Sub API hasn't been enabled yet. What is the fastest way to enable the API?

- A. Use a service account with the Pub/Sub Admin role to auto-enable the API.
- B. Enable the API in the Console.**
- C. Application's in App Engine don't require external APIs to be enabled.
- D. The API will be enabled the first time the code attempts to access Pub/Sub.

### Unattempted

Correct answer is B as the simplest way to enable an API for the project is using the GCP console.

Refer GCP documentation – Enable/Disable APIs

The simplest way to enable an API for your project is to use the GCP Console, though you can also enable an API using gcloud or using the Service Usage API. You can find out more about these options in the Service Usage API docs.

To enable an API for your project using the console:

1. Go to the GCP Console API Library.
2. From the projects list, select a project or create a new one.
3. In the API Library, select the API you want to enable. If you need help finding the API, use the search field and/or the filters.
4. On the API page, click ENABLE.

Option A is wrong as providing the Pub/Sub Admin role does not provide the access to enable API.

Enabling an API requires the following two Cloud Identity and Access Management permissions:

1. The servicemanagement.services.bind permission on the service to enable. This permission is present for all users for public services. For private services, you must share the service with the user who needs to enable it.
2. The serviceusage.services.enable permission on the project to enable the service on. This permission is present in the Editor role as well as in the Service Usage Admin role.

Option C is wrong as all applications need the API to be enabled before they can use it.

Option D is wrong as the API is not enabled and it needs to be enabled.

6. 6. Question

Your team is working on designing an IoT solution. There are thousands of devices that need to send periodic time series data for processing. Which services should be used to ingest and store the data?

- A. Pub/Sub, Datastore
- B. Pub/Sub, Dataproc
- C. Dataproc, Bigtable
- D. Pub/Sub, Bigtable

Unattempted

Correct answer is D as Pub/Sub is ideal for ingestion and Bigtable for time series data storage.

Refer GCP documentation – IoT Overview

#### Ingestion

Google Cloud Pub/Sub provides a globally durable message ingestion service. By creating topics for streams or channels, you can enable different components of your application to subscribe to specific streams of data without needing to construct subscriber-specific channels on each device. Cloud Pub/Sub also natively connects to other Cloud Platform services, helping you to connect ingestion, data pipelines, and storage systems.

Cloud Pub/Sub can act like a shock absorber and rate leveller for both incoming data streams and application architecture changes. Many devices have limited ability to store and retry sending telemetry data. Cloud Pub/Sub scales to handle data spikes that can occur when swarms of devices respond to events in the physical world, and buffers these spikes to help isolate them from applications monitoring the data.

#### Time Series dashboards with Cloud Bigtable

Certain types of data need to be quickly sliceable along known indexes and dimensions for updating core visualizations and user interfaces. Cloud Bigtable provides a low-latency and high-throughput database for NoSQL data. Cloud Bigtable provides a good place to drive heavily used visualizations and queries, where the questions are already well understood and you need to absorb or serve at high volumes.

Compared to BigQuery, Cloud Bigtable works better for queries that act on rows or groups of consecutive rows, because Cloud Bigtable stores data by using a row-based format. Compared to Cloud Bigtable, BigQuery is a better choice for queries that require data aggregation.

Option A is wrong as Datastore is not an ideal solution for time series IoT data storage.

Options B & C are wrong as Dataproc is not an ideal ingestion service for IoT solution. Also the storage is HDFS based.

7. 7. Question

Your development team has asked you to set up an external TCP load balancer with SSL offload. Which load balancer should you use?

- A. SSL proxy

- B. HTTP load balancer
- C. TCP proxy
- D. HTTPS load balancer

**Unattempted**

Correct answer is A as SSL proxy support TCP traffic with an ability to SSL offload.

Refer GCP documentation – Choosing Load Balancer

Google Cloud SSL Proxy Load Balancing terminates user SSL (TLS) connections at the load balancing layer, then balances the connections across your instances using the SSL or TCP protocols. Cloud SSL proxy is intended for non-HTTP(S) traffic. For HTTP(S) traffic, HTTP(S) load balancing is recommended instead. SSL Proxy Load Balancing supports both IPv4 and IPv6 addresses for client traffic. Client IPv6 requests are terminated at the load balancing layer, then proxied over IPv4 to your backends.

Options B & D are wrong as they are recommended for HTTP or HTTPS traffic only

Option C is wrong as TCP proxy does not support SSL offload.

**8. Question**

Your company wants to host confidential documents in Cloud Storage. Due to compliance requirements, there is a need for the data to be highly available and resilient even in case of a regional outage. Which storage classes help meet the requirement?

- A. Standard
- B. Regional
- C. Coldline
- D. Dual-Regional
- E. Multi-Regional

**Unattempted**

Correct answers are C & E as Multi-Regional and Coldline storage classes provide multi-region geo-redundant deployment, which can sustain regional failure.

Refer GCP documentation – Cloud Storage Classes

Multi-Regional Storage is geo-redundant.

The geo-redundancy of Coldline Storage data is determined by the type of location in which it is stored: Coldline Storage data stored in multi-regional locations is redundant across multiple regions, providing higher availability than Coldline Storage data stored in regional locations.

Data that is geo-redundant is stored redundantly in at least two separate geographic places separated by at least 100 miles. Objects stored in multi-regional locations are geo-redundant, regardless of their storage class.

Geo-redundancy occurs asynchronously, but all Cloud Storage data is redundant within at least one geographic place as soon as you upload it.

Geo-redundancy ensures maximum availability of your data, even in the event of

large-scale disruptions, such as natural disasters. For a dual-regional location, geo-redundancy is achieved using two specific regional locations. For other multi-regional locations, geo-redundancy is achieved using any combination of data centers within the specified multi-region, which may include data centers that are not explicitly available as regional locations.

Options A & D are wrong as they do not exist

Option B is wrong as Regional storage class is not geo-redundant. Data stored in a narrow geographic region and Redundancy is across availability zones

#### 9. 9. Question

Your manager needs you to test out the latest version of MS-SQL on a Windows instance. You've created the VM and need to connect into the instance. What steps should you follow to connect to the instance?

- A. Generate a Windows password in the console, then use a client capable of communicating via RDP and provide the credentials.
- B. Generate a Windows password in the console, and then use the RDP button to connect in through the console.
- C. Connect in with your own RDP client using your Google Cloud username and password.
- D. From the console click the SSH button to automatically connect.

**Unattempted**

Correct answer is A as connecting to Windows instance involves installation of the RDP client. GCP does not provide RDP client and it needs to be installed.

Generate Windows instance password to connect to the instance.

Refer GCP documentation – Windows Connecting to Instance

Option B is wrong as GCP Console does not have a direct RDP connectivity.

Option C is wrong as a separate windows password needs to be generated.

Google Cloud username password cannot be used.

Option D is wrong as you cannot connect to Windows instance using SSH.

#### 10. 10. Question

You need to create a new development Kubernetes cluster with 3 nodes. The cluster will be named project-1-cluster. Which of the following truncated commands will create a cluster?

- A. gcloud container clusters create project-1-cluster --num-nodes 3
- B. kubectl clusters create project-1-cluster 3
- C. kubectl clusters create project-1-cluster --num-nodes 3
- D. gcloud container clusters create project-1-cluster 3

**Unattempted**

Correct answer is A as Kubernetes cluster can be created using the gcloud command only, with the cluster name and –num-nodes parameter.  
Refer GCP documentation – Kubernetes Create Cluster  
gcloud container clusters create my-regional-cluster –num-nodes 2 \ –region us-west1  
Options B & C are wrong as kubectl cannot be used to create Kubernetes cluster.  
Option D is wrong as the 3 parameter is invalid and needs to follow a parameter.

## 11. Question

Your security team wants to be able to audit network traffic inside of your network. What's the best way to ensure they have access to the data they need?

- A. Disable flow logs.
- B. Enable flow logs.
- C. Enable VPC Network logs
- D. Add a firewall capture filter.

**Unattempted**

Correct answer is B as VPC Flow logs track all the network flows and needs to be enabled.

Refer GCP documentation – VPC Flow logs

VPC Flow Logs record a sample of network flows sent from and received by VM instances. These logs can be used for network monitoring, forensics, real-time security analysis, and expense optimization.

Flow logs are aggregated by connection, at 5-second intervals, from Compute Engine VMs and exported in real time. By subscribing to Cloud Pub/Sub, you can analyze flow logs using real-time streaming APIs.

Option A is wrong as the VPC logs need to be enabled and are disabled by default.

Option C is wrong as there are no VPC Network logs.

Option D is wrong as there is no firewall capture filter.

## 12. Question

You have a Cloud Storage bucket that needs to host static web assets with a dozen HTML pages, a few JavaScript files, and some CSS. How do you make the bucket public?

- A. Check the make public box on the GCP Console for the bucket
- B. gsutil iam ch allAuthenticatedUsers:objectViewer gs://bucket-name
- C. gsutil make-public gs://bucket-name
- D. gsutil iam ch allUsers:objectViewer gs://bucket-name

**Unattempted**

Correct answer is D as the bucket can be shared by providing the Storage Object Viewer access to allUsers.

Refer GCP documentation – Cloud Storage Sharing files

You can either make all files in your bucket publicly accessible, or you can set

individual objects to be accessible through your website. Generally, making all files in your bucket accessible is easier and faster.

To make all files accessible, follow the Cloud Storage guide for making groups of objects publicly readable.

To make individual files accessible, follow the Cloud Storage guide for making individual objects publicly readable.

If you choose to control the accessibility of individual files, you can set the default object ACL for your bucket so that subsequent files uploaded to your bucket are shared by default.

Use the gsutil acl ch command, replacing [VALUES\_IN\_BRACKETS] with the appropriate values:

```
gsutil acl ch -u AllUsers:R gs://[BUCKET_NAME]/[OBJECT_NAME]
```

Option A is wrong as there is no make public option with GCP Console.

Option B is wrong as access needs to be provided to allUsers to make it public and there is no allAuthenticatedUsers option.

Option C is wrong as there is no make public option with gsutil command.

### 13. 13. Question

You've created a new Compute Engine instance in zone us-central1-b. When you tried to attach the GPU that your data engineers requested, you're getting an error. What is the most likely cause of the error?

- A. Your instance isn't running with the correct scopes to allow GPUs.
- B. The GPU is not supported for your OS.
- C. Your instance isn't running with the default compute engine service account.
- D. The desired GPU doesn't exist in that zone.

**Unattempted**

Correct answer is D as GPU availability varies for region to region and zone to zone. One GPU available in one region/zone is not guarantee to be available in other region/zone.

Refer GCP documentation – GPUs

Option A is wrong as access scope for compute engine does not control GPU attachment with the Compute Engine.

Option B is wrong as GPUs can be attached to any OS and machine type.

Option C is wrong as access scope for compute engine does not control GPU attachment with the Compute Engine.

### 14. 14. Question

Your data team is working on some new machine learning models. They're generating several files per day that they want to store in a regional bucket. They mostly focus on the files from the last week. However, they want to keep all the files just to be safe and if needed, would be referred once in a month. With the fewest steps possible, what's the best way to lower the storage costs?

- A. Create a Cloud Function triggered when objects are added to a bucket. Look at the date on all the files and move it to Nearline storage if it's older than a week.
- B. Create a Cloud Function triggered when objects are added to a bucket. Look at the date on all the files and move it to Coldline storage if it's older than a week.
- C. Create a lifecycle policy to switch the objects older than a week to Coldline storage.
- D. Create a lifecycle policy to switch the objects older than a week to Nearline storage.

**Unattempted**

Correct answer is D as the files are required for a week and then would be needed for only once in a month access, Nearline storage would be an ideal storage to save cost. The transition of the object can be handled easily using Object Lifecycle Management.

Refer GCP documentation – Cloud Storage Lifecycle Management

You can assign a lifecycle management configuration to a bucket. The configuration contains a set of rules which apply to current and future objects in the bucket. When an object meets the criteria of one of the rules, Cloud Storage automatically performs a specified action on the object. Here are some example use cases:

Downgrade the storage class of objects older than 365 days to Coldline Storage.  
Delete objects created before January 1, 2013.

Keep only the 3 most recent versions of each object in a bucket with versioning enabled.

Option C is wrong as the files are needed once in a month, Coldline storage would not be a cost effective option.

Options A & B are wrong as the transition can be handled easily using Object Lifecycle management.

## 15. Question

Your company wants to setup a virtual private cloud network. They want to configure a single Subnet within the VPC with maximum range of available. Which CIDR block would you choose?

- A. 0.0.0.0/0
- B. 10.0.0.0/8
- C. 172.16.0.0/12
- D. 192.168.0.0/16

**Unattempted**

Correct answer is B as you can assign a standard private CIDR blocks (192.168.0.0/16, 172.16.0.0/12, 10.0.0.0/8) to VPC and their subsets as the IP address range of a VPC.

CIDR block Number of available private IPs

192.168.0.0/16 65,532

172.16.0.0/12 1,048,572

10.0.0.0/8 16,777,212

Refer GCP documentation – VPC Subnet IP ranges

Option A is wrong as it is not an allowed RFC 1918 CIDR range allowed.

Options C & D are wrong as they provide less private IPs compared to CIDR

10.0.0.0/8

## 16. Question

You've been tasked with getting all of your team's public SSH keys onto to a specific Bastion host instance of a particular project. You've collected them all. With the fewest steps possible, what is the simplest way to get the keys deployed?

- A. Add all of the keys into a file that's formatted according to the requirements. Use the gcloud compute instances add-metadata command to upload the keys to each instance
- B. Add all of the keys into a file that's formatted according to the requirements. Use the gcloud compute project-info add-metadata command to upload the keys.
- C. Use the gcloud compute ssh command to upload all the keys
- D. Format all of the keys as needed and then, using the user interface, upload each key one at a time.

**Unattempted**

Correct answer is A as instance specific SSH keys can help provide users access to the specific bastion host. The keys can be added or removed using the instance metadata.

Refer GCP documentation – Instance level SSH keys

Instance-level public SSH keys give users access to a specific Linux instance.

Users with instance-level public SSH keys can access a Linux instance even if it blocks project-wide public SSH keys.

gcloud compute instances add-metadata [INSTANCE\_NAME] --metadata-from-file ssh-keys=[LIST\_PATH]

Option B is wrong as the gcloud compute project-info provides access to all the instances within a project.

Option C is wrong as gcloud compute ssh is a thin wrapper around the ssh(1) command that takes care of authentication and the translation of the instance name into an IP address. It can be used to ssh to the instance.

Option D is wrong as there is no user interface to upload the keys.

## 17. Question

You're migrating an on-premises application to Google Cloud. The application uses a component that requires a licensing server. The license server has the IP address 10.28.0.10. You want to deploy the application without making any changes to the code or configuration. How should you go about deploying the application?

- A. Create a subnet with a CIDR range of 10.28.0.0/28. Reserve a static internal IP address of 10.28.0.10. Assign the static address to the license server instance.
- B. Create a subnet with a CIDR range of 10.28.0.0/28. Reserve a static external IP address of 10.28.0.10. Assign the static address to the license server instance.
- C. Create a subnet with a CIDR range of 10.28.0.0/28. Reserve an ephemeral internal IP address of 10.28.0.10. Assign the static address to the license server instance.
- D. Create a subnet with a CIDR range of 10.28.0.0/28. Reserve an ephemeral external IP address of 10.28.0.10. Assign the static address to the license server instance.

### Unattempted

Correct answer is A as the IP is internal it can be reserved using the static internal IP address, which blocks it and prevents it from getting allocated to other resource.

Refer GCP documentation – Compute Network Addresses

In Compute Engine, each VM instance can have multiple network interfaces.

Each interface can have one external IP address, one primary internal IP address, and one or more secondary internal IP addresses. Forwarding rules can have external IP addresses for external load balancing or internal addresses for internal load balancing.

Static internal IPs provide the ability to reserve internal IP addresses from the private RFC 1918 IP range configured in the subnet, then assign those reserved internal addresses to resources as needed. Reserving an internal IP address takes that address out of the dynamic allocation pool and prevents it from being used for automatic allocations. Reserving static internal IP addresses requires specific IAM permissions so that only authorized users can reserve a static internal IP address.

With the ability to reserve static internal IP addresses, you can always use the same IP address for the same resource even if you have to delete and recreate the resource.

Option C is wrong as Ephemeral internal IP addresses remain attached to a VM instance only until the VM is stopped and restarted or the instance is terminated. If an instance is stopped, any ephemeral internal IP addresses assigned to the instance are released back into the network pool. When a stopped instance is started again, a new ephemeral internal IP address is assigned to the instance. Options B & D are wrong as the IP address is RFC 1918 address and needs to be an internal static IP address.

### 18. Question

You've setup and tested several custom roles in your development project. What is the fastest way to create the same roles for your new production project?

- A. Recreate them in the new project.
- B. Use the gcloud iam copy roles command and set the destination project.

- C. In GCP console, select the roles and click the Export button.
- D. Use the gcloud iam roles copy command and set the destination project.

**Unattempted**

Correct answer is D as Cloud SDK gcloud iam roles copy can be used to copy the roles to different organization or project.

Refer GCP documentation – Cloud SDK IAM Copy Role

gcloud iam roles copy – create a role from an existing role

–dest-organization=DEST\_ORGANIZATION (The organization of the destination role)

–dest-project=DEST\_PROJECT (The project of the destination role)

#### 19. Question

You have been tasked to grant access to sensitive files to external auditors for a limited time period of 4 hours only. The files should not be strictly available after 4 hours. Adhering to Google best practices, how would you efficiently share the file?

- A. Host a website on Compute Engine instance and expose the files using Public DNS and share the URL with the auditors. Bring down the instance after 4 hours.
- B. Host a website on App Engine instance and expose the files using Public DNS and share the URL with the auditors. Bring down the instance after 4 hours.
- C. Store the file in Cloud Storage. Generate a signed URL with 4 hours expiry and share it with the auditors.
- D. Store the file in Cloud Storage. Grant the allUsers access to the file share it with the auditors. Remove allUsers access after 4 hours.

**Unattempted**

Correct answer is C as the file can be stored in Cloud Storage and Signed urls can be used to quickly and securely share the files with third party.

Refer GCP documentation – Cloud Storage Signed URLs

Signed URLs provide a way to give time-limited read or write access to anyone in possession of the URL, regardless of whether they have a Google account. In some scenarios, you might not want to require your users to have a Google account in order to access Cloud Storage, but you still want to control access using your application-specific logic. The typical way to address this use case is to provide a signed URL to a user, which gives the user read, write, or delete access to that resource for a limited time. Anyone who knows the URL can access the resource until the URL expires. You specify the expiration time in the query string to be signed.

Options A & B are wrong as it is not a quick solution, but a manual effort to host, share and stop the solution.

Option D is wrong as All Users is not a secure way to share data and it would be marked public.

## 20. 20. Question

A member of the finance team informed you that one of the projects is using the old billing account. What steps should you take to resolve the problem?

- A. Go to the Project page; expand the Billing tile; select the Billing Account option; select the correct billing account and save.
- B. Go to the Billing page; view the list of projects; find the project in question and select Change billing account; select the correct billing account and save.
- C. Delete the project and recreate it with the correct billing account.
- D. Submit a support ticket requesting the change.

**Unattempted**

Correct answer is B as for changing the billing account you have to select the project and change the billing account.

Refer GCP documentation – Change Billing Account

To change the billing account for an existing project, you must be an owner on the project and a billing administrator on the destination billing account.

To change the billing account:

1. Go to the Google Cloud Platform Console.
2. Open the console left side menu and select Billing.
3. If you have more than one billing account, you'll be prompted to select Go to linked billing account to manage the current project's billing.
4. Under Projects linked to this billing account, locate the name of the project that you want to change billing for, and then click the menu next to it.
5. Select Change billing account, then choose the desired destination billing account.
6. Click Set account.

Option A is wrong as billing account cannot be changed from Project page.

Option C is wrong as the project need not be deleted.

Option D is wrong as Google support does not handle the changes and it is users responsibility.

## 21. 21. Question

Your billing department has asked you to help them track spending against a specific billing account. They've indicated that they prefer to use Excel to create their reports so that they don't need to learn new tools. Which export option would work best for them?

- A. BigQuery Export
- B. File Export with JSON
- C. SQL Export
- D. File Export with CSV

**Unattempted**

Correct answer is D as Cloud Billing allows export of the billing data as flat files in CSV and JSON format. As the billing department wants to use Excel to create their reports, CSV would be a ideal option.

Refer GCP documentation – Cloud Billing Export Billing Data

To access a detailed breakdown of your charges, you can export your daily usage and cost estimates automatically to a CSV or JSON file stored in a Google Cloud Storage bucket you specify. You can then access the data via the Cloud Storage API, CLI tool, or Google Cloud Platform Console.

Usage data is labeled with the project number and resource type. You use ACLs on your Cloud Storage bucket to control who can access this data.

Options A, B, & C are wrong as they do not support Excel directly and would need conversions.

## 22. Question

A company wants to setup a template for deploying resources. They want the provisioning to be dynamic with the specifications in configuration files. Which of the following service would be ideal for this requirement?

- A. Cloud Composer
- B. Deployment Manager
- C. Cloud Scheduler
- D. Cloud Deployer

**Unattempted**

Correct answer is B as Deployment Manager provide Infrastructure as a Code capability.

Refer GCP documentation – Deployment Manager

Google Cloud Deployment Manager allows you to specify all the resources needed for your application in a declarative format using yaml. You can also use Python or Jinja2 templates to parameterize the configuration and allow reuse of common deployment paradigms such as a load balanced, auto-scaled instance group. Treat your configuration as code and perform repeatable deployments.

Option A is wrong as Cloud Composer is a fully managed workflow orchestration service that empowers you to author, schedule, and monitor pipelines that span across clouds and on-premises data centers.

Option C is wrong as Cloud Scheduler is a fully managed enterprise-grade cron job scheduler. It allows you to schedule virtually any job, including batch, big data jobs, cloud infrastructure operations, and more.

Option D is wrong as Cloud Deployer is not a valid service.

## 23. Question

Your project manager wants to delegate the responsibility to upload objects to Cloud Storage buckets to his team members. Considering the principle of least privilege, which role should you assign to the team members?

- A. roles/storage.objectAdmin
- B. roles/storage.objectViewer

- C. roles/storage.objectCreator

- D. roles/storage.admin

**Unattempted**

Correct answer is C as roles/storage.objectCreator allows users to create objects. Does not give permission to view, delete, or overwrite objects.

Refer GCP documentation – Cloud Storage IAM Roles

Options B is wrong as roles/storage.objectViewer role does not provide sufficient privileges to manage buckets.

Options A & D are wrong as it provides more privileges than required.

#### 24. 24. Question

Your company needs to create a new Kubernetes Cluster on Google Cloud Platform. As a security requirement, they want to upgrade the nodes to the latest stable version of Kubernetes with no manual intervention. How should the Kubernetes cluster be configured?

- A. Always use the latest version while creating the cluster

- B. Enable node auto-repairing

- C. Enable node auto-upgrades

- D. Apply security patches on the nodes as they are released

**Unattempted**

Correct answer is C as the Kubernetes cluster can be configured for node auto-upgrades to update them to the latest stable version of Kubernetes.

Refer GCP documentation – Kubernetes Auto Upgrades

Node auto-upgrades help you keep the nodes in your cluster up to date with the latest stable version of Kubernetes. Auto-Updates use the same update mechanism as manual node upgrades.

Some benefits of using auto-upgrades:

Lower management overhead: You don't have to manually track and update to the latest version of Kubernetes.

Better security: Sometimes new binaries are released to fix a security issue. With auto-upgrades, GKE automatically ensures that security updates are applied and kept up to date.

Ease of use: Provides a simple way to keep your nodes up to date with the latest Kubernetes features.

Node pools with auto-upgrades enabled are automatically scheduled for upgrades when a new stable Kubernetes version becomes available. When the upgrade is performed, nodes are drained and re-created to match the current cluster master version. Modifications on the boot disk of a node VM do not persist across node re-creations. To preserve modifications across node re-creation, use a DaemonSet.

Option A is wrong as this would not take into account any latest updates.

Option B is wrong as auto repairing helps in keeping nodes healthy and does not handle upgrades.

Option D is wrong as it is a manual effort and not feasible.

## 25. 25. Question

You have created an App engine application in the us-central region. However, you found out the network team has configured all the VPN connections in the asia-east2 region, which are not possible to move. How can you change the location efficiently?

- A. Change the region in app.yaml and redeploy
- B. From App Engine console, change the region of the application
- C. Change the region in application.xml within the application and redeploy
- D. Create a new project in the asia-east2 region and create app engine in the project

**Unattempted**

Correct answer is D as app engine is a regional resource, it needs to be redeployed to the different region.

Refer GCP documentation – App Engine locations

App Engine is regional, which means the infrastructure that runs your apps is located in a specific region and is managed by Google to be redundantly available across all the zones within that region.

Meeting your latency, availability, or durability requirements are primary factors for selecting the region where your apps are run. You can generally select the region nearest to your app's users but you should consider the location of the other GCP products and services that are used by your app. Using services across multiple locations can affect your app's latency as well as pricing  
You cannot change an app's region after you set it.

Options A, B & C are wrong as one the region it set for the app engine it cannot be modified.

## 26. 26. Question

Your team needs to set up a MongoDB instance as quickly as possible. You don't know how to install it and what configuration files are needed. What's the best way to get it up-and-running quickly?

- A. Use Cloud Memorystore
- B. Learn and deploy MongoDB to a Compute Engine instance.
- C. Install with Cloud Launcher Marketplace
- D. Create a Deployment Manager template and deploy it.

**Unattempted**

Correct answer is C as Cloud Launcher provides out of box deployments that are completely transparent to you and can be done in no time.

Refer GCP documentation – Cloud Launcher

GCP Marketplace offers ready-to-go development stacks, solutions, and services to accelerate development. So you spend less time installing and more time developing.

Deploy production-grade solutions in a few clicks  
Single bill for all your GCP and 3rd party services  
Manage solutions using Deployment Manager  
Notifications when a security update is available  
Direct access to partner support  
Option A is wrong as Cloud Memorystore is Redis compliant and not an alternative for MongoDB  
Option B is wrong as hosting on the compute engine is still a manual step and would require time.  
Option D is wrong as Deployment Manager would take time to build and deploy.

## 27. Question

Your company wants to setup Production and Test environment. They want to use different subjects and the key requirement is that the VMs must be able to communicate with each other using internal IPs no additional routes configured. How can the solution be designed?

- A. Configure a single VPC with 2 subnets having the same CIDR range hosted in the same region
- B. Configure a single VPC with 2 subnets having the different CIDR range hosted in the different region
- C. Configure 2 VPCs with 1 subnet each having the same CIDR range hosted in the same region
- D. Configure 2 VPCs with 1 subnet each having the different CIDR range hosted in the different region

**Unattempted**

Correct answer is B as the VMs need to be able to communicate using private IPs they should be hosted in the same VPC. The Subnets can be in any region, however they should have non-overlapping CIDR range.

Refer GCP documentation – VPC Intra VPC reqs

The system-generated subnet routes define the paths for sending traffic among instances within the network using internal (private) IP addresses. For one instance to be able to communicate with another, appropriate firewall rules must also be configured because every network has an implied deny firewall rule for ingress traffic.

Option A is wrong as CIDR range cannot overlap.

Options C & D are wrong as VMs in subnet in different VPC cannot communicate with each other using private IPs.

## 28. Question

Your company is hosting their static website on Cloud Storage. You have implemented a change to add PDF files to the website. However, when the user clicks on the PDF file link it downloads the PDF instead of opening it within the browser. What would you change to fix the issue?

- A. Set content-type as object metadata to application/octet-stream on the files
- B. Set content-type as object metadata to application/pdf on the files**
- C. Set content-type as object metadata to application/octet-stream on the bucket
- D. Set content-type as object metadata to application/pdf on the bucket

**Unattempted**

Correct answer is B as the browser needs the correct content-type to be able to interpret and render the file correctly. The content-type can be set on object metadata and should be set to application/pdf.

Refer GCP documentation – Cloud Storage Object Metadata Content-Type

The most commonly set metadata is Content-Type (also known as MIME type), which allows browsers to render the object properly. All objects have a value specified in their Content-Type metadata, but this value does not have to match the underlying type of the object. For example, if the Content-Type is not specified by the uploader and cannot be determined, it is set to application/octet-stream or application/x-www-form-urlencoded, depending on how you uploaded the object.

Option A is wrong the content type needs to be set to application/pdf

Options C & D are wrong as the metadata should be set on the objects and not on the bucket.

## 29. Question

You currently are running an application on a machine type with 2 vCPUs and 4gb RAM. However, recently there have been plenty of memory problems. How to increase the memory of the application with minimal downtime?

- A. In GCP console, upgrade the memory of the Compute Engine instance
- B. Use gcloud compute instances increase-memory to increase the memory
- C. Use Live migration to move to machine type with higher memory**
- D. Use Live migration to move to machine type with higher CPU

**Unattempted**

Correct answer is C as Live migration would help migrate the instance to an machine-type with higher memory with minimal to no downtime.

Refer GCP documentation – Live Migration

Compute Engine offers live migration to keep your virtual machine instances running even when a host system event occurs, such as a software or hardware update. Compute Engine live migrates your running instances to another host in the same zone rather than requiring your VMs to be rebooted. This allows Google

to perform maintenance that is integral to keeping infrastructure protected and reliable without interrupting any of your VMs.

Live migration keeps your instances running during:

- Regular infrastructure maintenance and upgrades.
- Network and power grid maintenance in the data centers.
- Failed hardware such as memory, CPU, network interface cards, disks, power, and so on. This is done on a best-effort basis; if a hardware fails completely or otherwise prevents live migration, the VM crashes and restarts automatically and a hostError is logged.
- Host OS and BIOS upgrades.
- Security-related updates, with the need to respond quickly.
- System configuration changes, including changing the size of the host root partition, for storage of the host image and packages.

Live migration does not change any attributes or properties of the VM itself. The live migration process just transfers a running VM from one host machine to another host machine within the same zone. All VM properties and attributes remain unchanged, including internal and external IP addresses, instance metadata, block storage data and volumes, OS and application state, network settings, network connections, and so on.

Options A & B are wrong as the memory cannot be increased for an instance from console or command line

Option D is wrong the live migration needs to be done to an instance type with higher CPU.

### 30. 30. Question

Your billing department has asked you to help them track spending against a specific billing account. They've indicated that they prefer SQL querying to create their reports so that they don't need to learn new tools. The data should be as latest as possible. Which export option would work best for them?

- A. File Export with JSON and load to Cloud SQL and provide Cloud SQL access to billing department
- B. Create a sink to BigQuery and provide BigQuery access to billing department
- C. Create a sink to Cloud SQL and provide Cloud SQL access to billing department
- D. File Export with CSV and load to Cloud SQL and provide Cloud SQL access to billing department

**Unattempted**

Correct answer is B as Billing data can be automatically exported to BigQuery and BigQuery provides the SQL interface for the billing department to query the data.

Refer GCP documentation – Cloud Billing Export BigQuery

Tools for monitoring, analyzing and optimizing cost have become an important part of managing development. Billing export to BigQuery enables you to export your daily usage and cost estimates automatically throughout the day to a BigQuery dataset you specify. You can then access your billing data from BigQuery. You can also use this export method to export data to a JSON file.

Options A & D are wrong as it would need manual exporting and loading the data to Cloud SQL.

Option C is wrong as Billing does not export to Cloud SQL

### 31. Question

Your company hosts multiple applications on Compute Engine instances. They want the instances to be resilient to any Host maintenance activities performed on the instance. How would you configure the instances?

- A. Set automaticRestart availability policy to true
- B. Set automaticRestart availability policy to false
- C. Set onHostMaintenance availability policy to migrate instances
- D. Set onHostMaintenance availability policy to terminate instances

**Unattempted**

Correct answer is C as onHostMaintenance availability policy determines how the instance reacts to the host maintenance events.

Refer GCP documentation – Instance Scheduling Options

A VM instance's availability policy determines how it behaves when an event occurs that requires Google to move your VM to a different host machine. For example, you can choose to keep your VM instances running while Compute Engine live migrates them to another host or you can choose to terminate your instances instead. You can update an instance's availability policy at any time to control how you want your VM instances to behave.

You can change an instance's availability policy by configuring the following two settings:

The VM instance's maintenance behavior, which determines whether the instance is live migrated or terminated when there is a maintenance event.

The instance's restart behavior, which determines whether the instance automatically restarts if it crashes or gets terminated.

The default maintenance behavior for instances is to live migrate, but you can change the behavior to terminate your instance during maintenance events instead.

Configure an instance's maintenance behavior and automatic restart setting using the onHostMaintenance and automaticRestart properties. All instances are configured with default values unless you explicitly specify otherwise.

onHostMaintenance: Determines the behavior when a maintenance event occurs that might cause your instance to reboot.

[Default] migrate, which causes Compute Engine to live migrate an instance when there is a maintenance event.

terminate, which terminates an instance instead of migrating it.

automaticRestart: Determines the behavior when an instance crashes or is terminated by the system.

[Default] true, so Compute Engine restarts an instance if the instance crashes or is terminated.

false, so Compute Engine does not restart an instance if the instance crashes or is terminated.

Options A & B are wrong as automaticRestart does not apply to host

maintenance event.

Option D is wrong as the onHostMaintenance needs to be set to migrate the instance as termination would lead to loss of instance.

### 32. Question

Your company wants to try out the cloud with low risk. They want to archive approximately 100 TB of their log data to the cloud and test the analytics features available to them there, while also retaining that data as a long-term disaster recovery backup. Which two steps should they take? (Choose two answers)

- A. Load logs into Google BigQuery.
- B. Load logs into Google Cloud SQL.
- C. Import logs into Google Stackdriver.
- D. Insert logs into Google Cloud Bigtable.
- E. Upload log files into Google Cloud Storage.

Unattempted

Correct answers are A & E as Google Cloud Storage can provide long term archival option and BigQuery provides analytics capabilities.

Option B is wrong as Cloud SQL is relational database and does not support the capacity required as well as not suitable for long term archival storage.

Option C is wrong as Stackdriver is a monitoring, logging, alerting and debugging tool. It is not ideal for long term retention of data and does not provide analytics capabilities.

Option D is wrong as Bigtable is a NoSQL solution and can be used for analytics. However it is ideal for data with low latency access and is expensive.

### 33. Question

Your company wants to reduce cost on infrequently accessed data by moving it to the cloud. The data will still be accessed approximately once a month to refresh historical charts. In addition, data older than 5 years needs to be archived for 5 years for compliance reasons. How should you store and manage the data?

- A. In Google Cloud Storage and stored in a Multi-Regional bucket. Set an Object Lifecycle Management policy to delete data older than 5 years.
- B. In Google Cloud Storage and stored in a Multi-Regional bucket. Set an Object Lifecycle Management policy to change the storage class to Coldline for data older than 5 years.
- C. In Google Cloud Storage and stored in a Nearline bucket. Set an Object Lifecycle Management policy to delete data older than 5 years.
- D. In Google Cloud Storage and stored in a Nearline bucket. Set an Object Lifecycle Management policy to change the storage class to Coldline for data older than 5 years.

Unattempted

Correct answer is D as the access pattern fits Nearline storage class requirements and Nearline is a more cost-effective storage approach than Multi-Regional. The object lifecycle management policy to move data to Coldline is ideal for archival.

Refer GCP documentation – Cloud Storage – Storage Classes

Options A & B are wrong as Multi-Regional storage class is not an ideal storage option with infrequent access.

Option C is wrong as the data is required for compliance it cannot be deleted and needs to be moved to the Coldline storage.

#### 34. 34. Question

Your company plans to migrate a multi-petabyte data set to the cloud. The data set must be available 24hrs a day. Your business analysts have experience only with using a SQL interface. How should you store the data to optimize it for ease of analysis?

- A. Load data into Google BigQuery.
- B. Insert data into Google Cloud SQL.
- C. Put flat files into Google Cloud Storage.
- D. Stream data into Google Cloud Datastore.

**Unattempted**

Correct answer is A as BigQuery is the only of these Google products that supports an SQL interface and a high enough SLA (99.9%) to make it readily available.

Option B is wrong as Cloud SQL cannot support multi-petabyte data. Storage limit for Cloud SQL is 10TB

Option C is wrong as Cloud Storage does not provide SQL interface.

Option D is wrong as Datastore does not provide a SQL interface and is a NoSQL solution.

#### 35. 35. Question

You have a Kubernetes cluster with 1 node-pool. The cluster receives a lot of traffic and needs to grow. You decide to add a node. What should you do?

- A. Use gcloud container clusters resize with the desired number of nodes.
- B. Use kubectl container clusters resize with the desired number of nodes.
- C. Edit the managed instance group of the cluster and increase the number of VMs by 1.
- D. Edit the managed instance group of the cluster and enable autoscaling.

**Unattempted**

Correct answer is A as the kubernetes cluster can be resized using the gcloud command.

Refer GCP documentation – Resizing Kubernetes Cluster

```
gcloud container clusters resize [CLUSTER_NAME] --node-pool [POOL_NAME] \--size [SIZE]
```

Option B is wrong as kubernetes cluster cannot be resized using the kubectl command

Options C & D are wrong as the managed instance groups should be changed manually.

### 36. Question

What is the command for creating a storage bucket that has once per month access and is named ‘archive\_bucket’?

- A. gsutil rm -coldline gs://archive\_bucket
- B. gsutil mb -c coldline gs://archive\_bucket
- C. **gsutil mb -c nearline gs://archive\_bucket**
- D. gsutil mb gs://archive\_bucket

**Unattempted**

Correct answer is C as the data needs to be accessed on monthly basis Nearline is an ideal storage class. Also gsutil needs -c parameter to pass the class.

Refer GCP documentation – Storage Classes

Nearline – Data you do not expect to access frequently (i.e., no more than once per month). Ideal for back-up and serving long-tail multimedia content.

Option A is wrong as rm is the wrong parameter and removes the data.

Option B is wrong as coldline is not suited for data that needs monthly access.

Option D is wrong as by default, gsutil would create a regional bucket.

### 37. Question

You need to take streaming data from thousands of Internet of Things (IoT) devices, ingest it, run it through a processing pipeline, and store it for analysis. You want to run SQL queries against your data for analysis. What services in which order should you use for this task?

- A. Cloud Dataflow, Cloud Pub/Sub, BigQuery
- B. Cloud Pub/Sub, Cloud Dataflow, Cloud Dataproc
- C. **Cloud Pub/Sub, Cloud Dataflow, BigQuery**
- D. App Engine, Cloud Dataflow, BigQuery

**Unattempted**

Correct answer is C as the need to ingest it, transform and store the Cloud Pub/Sub, Cloud Dataflow, BigQuery is ideal stack to handle the IoT data.

Refer GCP documentation – IoT

Google Cloud Pub/Sub provides a globally durable message ingestion service. By creating topics for streams or channels, you can enable different components

of your application to subscribe to specific streams of data without needing to construct subscriber-specific channels on each device. Cloud Pub/Sub also natively connects to other Cloud Platform services, helping you to connect ingestion, data pipelines, and storage systems.

Google Cloud Dataflow provides the open Apache Beam programming model as a managed service for processing data in multiple ways, including batch operations, extract-transform-load (ETL) patterns, and continuous, streaming computation. Cloud Dataflow can be particularly useful for managing the high-volume data processing pipelines required for IoT scenarios. Cloud Dataflow is also designed to integrate seamlessly with the other Cloud Platform services you choose for your pipeline.

Google BigQuery provides a fully managed data warehouse with a familiar SQL interface, so you can store your IoT data alongside any of your other enterprise analytics and logs. The performance and cost of BigQuery means you might keep your valuable data longer, instead of deleting it just to save disk space.

Sample Arch – Mobile Gaming Analysis Telemetry

Processing game client and game server events in real time

Option A is wrong as the stack is correct, however the order is not correct.

Option B is wrong as Dataproc is not an ideal tool for analysis. Cloud Dataproc is a fast, easy-to-use, fully-managed cloud service for running Apache Spark and Apache Hadoop clusters in a simpler, more cost-efficient way.

Option D is wrong as App Engine is not an ideal ingestion tool to handle IoT data.

### 38. Question

Your application has a large international audience and runs stateless virtual machines within a managed instance group across multiple locations. One feature of the application lets users upload files and share them with other users. Files must be available for 30 days; after that, they are removed from the system entirely. Which storage solution should you choose?

- A. A Cloud Datastore database.
- B. A multi-regional Cloud Storage bucket.
- C. Persistent SSD on virtual machine instances.
- D. A managed instance group of Filestore servers.

**Unattempted**

Correct answer is B as the key storage requirements is it being global, allow lifecycle management and sharing capability. Cloud Storage is an ideal choice as it can be configured to be multi-regional, have lifecycle management rules to auto delete the files after 30 days and share them with others.

Option A is wrong Datastore is a NoSQL solution and not ideal for unstructured data.

Option C is wrong as SSD disks are ephemeral storage option for virtual machines.

Option D is wrong as disks are regional and not ideal storage option for content that needs to be shared.

### 39. Question

Your company wants to track whether someone is present in a meeting room reserved for a scheduled meeting. There are 1000 meeting rooms across 5 offices on 3 continents. Each room is equipped with a motion sensor that reports its status every second. The data from the motion detector includes only a sensor ID and several different discrete items of information. Analysts will use this data, together with information about account owners and office locations. Which database type should you use?

- A. Flat file
- B. NoSQL
- C. Relational
- D. Blobstore

**Unattempted**

Correct answer is B as NoSQL like Bigtable and Datastore solution is an ideal solution to store sensor ID and several different discrete items of information. It also provides an ability to join with other data. Datastore can also be configured to store data in multi-region locations.

Refer GCP documentation – Storage Options

Option A is wrong as flat file is not an ideal storage option. It is not scalable.  
Option C is wrong as relational database like Cloud SQL is not an ideal solution to store schema less data.

Option D is wrong as blob storage like Cloud Storage is not an ideal solution to store, analyze schema less data and join with other sources.

#### 40. 40. Question

You have data stored in a Cloud Storage dataset and also in a BigQuery dataset. You need to secure the data and provide 3 different types of access levels for your Google Cloud Platform users: administrator, read/write, and read-only. You want to follow Google-recommended practices. What should you do?

- A. Create 3 custom IAM roles with appropriate policies for the access levels needed for Cloud Storage and BigQuery. Add your users to the appropriate roles.
- B. At the Organization level, add your administrator user accounts to the Owner role, add your read/write user accounts to the Editor role, and add your read-only user accounts to the Viewer role.
- C. At the Project level, add your administrator user accounts to the Owner role, add your read/write user accounts to the Editor role, and add your read-only user accounts to the Viewer role.
- D. Use the appropriate pre-defined IAM roles for each of the access levels needed for Cloud Storage and BigQuery. Add your users to those roles for each of the services.

**Unattempted**

Correct answer is D as Google best practice is to use pre-defined rules over legacy primitive and custom roles. Pre-defined roles can help grant fine grained control per service.

Refer GCP documentation – IAM Overview

Primitive roles: The roles historically available in the Google Cloud Platform

Console will continue to work. These are the Owner, Editor, and Viewer roles.

Predefined roles: Predefined roles are the Cloud IAM roles that give finer-grained access control than the primitive roles. For example, the predefined role Pub/Sub Publisher (roles/pubsub.publisher) provides access to only publish messages to a Cloud Pub/Sub topic.

Custom roles: Roles that you create to tailor permissions to the needs of your organization when predefined roles don't meet your needs.

What is the difference between primitive roles and predefined roles?

Primitive roles are the legacy Owner, Editor, and Viewer roles. IAM provides predefined roles, which enable more granular access than the primitive roles. Grant predefined roles to identities when possible, so you only give the least amount of access necessary to access your resources.

When would I use primitive roles?

Use primitive roles in the following scenarios:

When the GCP service does not provide a predefined role. See the predefined roles table for a list of all available predefined roles.

When you want to grant broader permissions for a project. This often happens when you're granting permissions in development or test environments.

When you need to allow a member to modify permissions for a project, you'll want to grant them the owner role because only owners have the permission to grant access to other users for projects.

When you work in a small team where the team members don't need granular permissions.

Option A is wrong as you should use custom roles only if predefined roles are not available.

Options B & C are wrong Google does not recommend using primitive roles which do not allow fine grained access control. Also primitive roles are applied at project or service resource levels

#### 41. Question

You have created a Kubernetes deployment, called Deployment-A, with 3 replicas on your cluster. Another deployment, called Deployment-B, needs access to Deployment-A. You cannot expose Deployment-A outside of the cluster. What should you do?

- A. Create a Service of type NodePort for Deployment A and an Ingress Resource for that Service. Have Deployment B use the Ingress IP address.
- B. Create a Service of type LoadBalancer for Deployment A. Have Deployment B use the Service IP address.
- C. Create a Service of type LoadBalancer for Deployment A and an Ingress Resource for that Service. Have Deployment B use the Ingress IP address.

- D. Create a Service of type ClusterIP for Deployment A. Have Deployment B use the Service IP address.

Unattempted

Correct answer is D as this exposes the service on a cluster-internal IP address. Choosing this method makes the service reachable only from within the cluster.

Refer GCP documentation – Kubernetes Networking

Option A is wrong as this exposes Deployment A over the public internet.

Option B is wrong as LoadBalancer will expose the service publicly.

Option C is wrong as this exposes the service externally using a cloud provider's load balancer, and Ingress can work only with nodeport, not LoadBalancer.

#### 42. 42. Question

You want to create a new role for your colleagues that will apply to all current and future projects created in your organization. The role should have the permissions of the BigQuery Job User and Cloud Bigtable User roles. You want to follow Google's recommended practices. How should you create the new role?

- A. Use gcloud iam combine-roles --global to combine the 2 roles into a new custom role.
- B. For one of your projects, in the Google Cloud Platform Console under Roles, select both roles and combine them into a new custom role. Use gcloud iam promote-role to promote the role from a project role to an organization role.
- C. For all projects, in the Google Cloud Platform Console under Roles, select both roles and combine them into a new custom role.
- D. For your organization, in the Google Cloud Platform Console under Roles, select both roles and combine them into a new custom role.

Unattempted

Correct answer is D as this creates a new role with the combined permissions on the organization level.

Option A is wrong as this does not create a new role.

Option B is wrong as gcloud cannot promote a role to org level.

Option C is wrong as it's recommended to define the role on the organization level. Also, the role will not be applied on new projects.

#### 43. 43. Question

Your team uses a third-party monitoring solution. They've asked you to deploy it to all nodes in your Kubernetes Engine Cluster. What's the best way to do that?

- A. Connect to each node via SSH and install the monitoring solution.
- B. Deploy the monitoring pod as a StatefulSet.
- C. Deploy the monitoring pod as a DaemonSet.

- D. Use Deployment Manager to deploy the monitoring solution.

**Unattempted**

Correct answer is C as Daemon set helps deploy applications or tools that you need to run on all the nodes.

Refer GCP documentation – Kubernetes Engine Daemon Set

Like other workload objects, DaemonSets manage groups of replicated Pods. However, DaemonSets attempt to adhere to a one-Pod-per-node model, either across the entire cluster or a subset of nodes. As you add nodes to a node pool, DaemonSets automatically add Pods to the new nodes as needed.

DaemonSets use a Pod template, which contains a specification for its Pods. The Pod specification determines how each Pod should look: what applications should run inside its containers, which volumes it should mount, its labels and selectors, and more.

DaemonSets are useful for deploying ongoing background tasks that you need to run on all or certain nodes, and which do not require user intervention. Examples of such tasks include storage daemons like ceph, log collection daemons like fluentd, and node monitoring daemons like collectd.

For example, you could have DaemonSets for each type of daemon run on all of your nodes. Alternatively, you could run multiple DaemonSets for a single type of daemon, but have them use different configurations for different hardware types and resource needs.

Option A is wrong as it is not a viable option.

Option B is wrong as Stateful set is useful for maintaining state. StatefulSets represent a set of [Pods] with unique, persistent identities and stable hostnames that GKE maintains regardless of where they are scheduled. The state information and other resilient data for any given StatefulSet Pod is maintained in persistent disk storage associated with the StatefulSet.

Option D is wrong as Deployment manager does not control Pods.

#### 44. Question

You're attempting to deploy a new instance that uses the centos 7 family. You can't recall the exact name of the family. Which command could you use to determine the family names?

- A. gcloud compute instances list
- B. gcloud compute images show-families
- C. gcloud compute instances show-families
- D. **gcloud compute images list**

**Unattempted**

Correct answer is D as family names are image attributes.

Refer GCP documentation – Cloud SDK Compute Images List & Image Families  
Image families simplify the process of managing images in your project by grouping related images together and making it easy to roll forward and roll back between specific image versions. An image family always points to the latest version of an image that is not deprecated. Most public images are grouped into an image families. For example, the debian-9image family in the debian-cloud project always points to the most recent Debian 9 image.

You can add your own images to an image family when you create a custom image. The image family points to the most recent image that you added to that family. Because the image family never points to a deprecated image, rolling the image family back to a previous image version is as simple as deprecating the most recent image in that family.

Options A, B & C are wrong as they do not help retrieve the image family.

#### 45. Question

Your security team has asked you to present them some numbers based on the logs that are exported to BigQuery. Due to the team structure, your manager has asked you to determine how much the query will cost. What's the best way to determine the cost?

- A. It's not possible to estimate the cost of a query.
- B. Create the query and execute the query in cost estimation mode
- C. Create the query and use the `--dry_run` option to determine the amount of data read, and then use the price calculator to determine the cost.
- D. Use the BigQuery index viewer to determine how many records you'll be reading.

**Unattempted**

Correct answer is C as the `--dry-run` option can be used to price your queries before they are actually fired. The Query returns the bytes read, which can then be used with the Pricing Calculator to estimate the query cost.

Refer GCP documentation – BigQuery Best Practices

Price your queries before running them

Best practice: Before running queries, preview them to estimate costs.

Queries are billed according to the number of bytes read. To estimate costs before running a query use:

The query validator in the GCP Console or the classic web UI

The `--dry_run` flag in the CLI

The `dryRun` parameter when submitting a query job using the API

The Google Cloud Platform Pricing Calculator

Options A, B & D are wrong as they are not valid options.

#### 46. Question

Your development team has asked you to set up load balancer with SSL termination. The website would be using HTTPS protocol. Which load balancer should you use?

- A. SSL proxy
- B. HTTP load balancer
- C. TCP proxy
- D. HTTPS load balancer

### Unattempted

Correct answer is D as HTTPS load balancer supports the HTTPS traffic with the SSL termination ability.

Refer GCP documentation – Choosing Load Balancer

An HTTPS load balancer has the same basic structure as an HTTP load balancer (described above), but differs in the following ways:

An HTTPS load balancer uses a target HTTPS proxy instead of a target HTTP proxy.

An HTTPS load balancer requires at least one signed SSL certificate installed on the target HTTPS proxy for the load balancer. You can use Google-managed or self-managed SSL certificates.

The client SSL session terminates at the load balancer.

HTTPS load balancers support the QUIC transport layer protocol.

Option A is wrong as SSL proxy is not recommended for HTTPS traffic.

Google Cloud SSL Proxy Load Balancing terminates user SSL (TLS) connections at the load balancing layer, then balances the connections across your instances using the SSL or TCP protocols. Cloud SSL proxy is intended for non-HTTP(S) traffic. For HTTP(S) traffic, HTTP(S) load balancing is recommended instead.

SSL Proxy Load Balancing supports both IPv4 and IPv6 addresses for client traffic. Client IPv6 requests are terminated at the load balancing layer, then proxied over IPv4 to your backends.

Option B is wrong as HTTP load balancer does not support SSL termination.

Option C is wrong as TCP proxy does not support SSL offload and not recommended for HTTP/S traffic.

### 47. Question

You've created a bucket to store some data archives for compliance. The data isn't likely to need to be viewed. However, you need to store it for at least 7 years. What is the best default storage class?

- A. Multi-regional
- B. Coldline
- C. Regional
- D. Nearline

### Unattempted

Correct answer is B as Coldline storage is an ideal solution for archival of infrequently accessed data at low cost.

Refer GCP documentation – Cloud Storage Classes

Google Cloud Storage Coldline is a very-low-cost, highly durable storage service for data archiving, online backup, and disaster recovery. Unlike other “cold” storage services, your data is available within milliseconds, not hours or days.

Coldline Storage is the best choice for data that you plan to access at most once a year, due to its slightly lower availability, 90-day minimum storage duration, costs for data access, and higher per-operation costs. For example:

Cold Data Storage – Infrequently accessed data, such as data stored for legal or regulatory reasons, can be stored at low cost as Coldline Storage, and be available when you need it.

Disaster recovery – In the event of a disaster recovery event, recovery time is key. Cloud Storage provides low latency access to data stored as Coldline Storage.

The geo-redundancy of Coldline Storage data is determined by the type of location in which it is stored: Coldline Storage data stored in multi-regional locations is redundant across multiple regions, providing higher availability than Coldline Storage data stored in regional locations.

Options A, C & D are wrong as they are not suited for archival data.

#### 48. Question

You have installed an SQL server on a windows instance. You want to connect to the instance. What steps should you follow to connect to the instance with fewest steps?

- A. Generate Windows user and password. Check security group for 3389 firewall rule. Use RDP option from GCP Console to connect
- B. Generate Windows password. Check security group for 22 firewall rule. Use RDP option from GCP Console to connect
- C. Generate Windows user and password. Check security group for 22 firewall rule. Install RDP Client to connect
- D. Generate Windows password. Check security group for 3389 firewall rule. Install RDP Client to connect

Unattempted

Correct answer is D as connecting to Windows instance involves installation of the RDP client. GCP does not provide RDP client and it needs to be installed. Generate Windows instance password to connect to the instance and the RDP port is 3389

Refer GCP documentation – Windows Connecting to Instance

Options A & B are wrong as you need an external client and connect connect directly from GCP console.

Options B & C are wrong as 22 port is for SSH.

#### 49. Question

Your team has been working on building a web application. The plan is to deploy to Kubernetes. You currently have a Dockerfile that works locally. How can you get the application deployed to Kubernetes?

- A. Use kubectl to push the convert the Dockerfile into a deployment.
- B. Use docker to create a container image, save the image to Cloud Storage, deploy the uploaded image to Kubernetes with kubectl.
- C. Use kubectl apply to push the Dockerfile to Kubernetes.
- D. Use docker to create a container image, push it to the Google Container Registry, deploy the uploaded image to Kubernetes with kubectl.

### Unattempted

Correct answer is D as the correct steps are to create the container image and push it to Google Container Registry and deploy the image to Kubernetes with Kubectl.

Refer GCP documentation – Kubernetes Engine Deploy

To package and deploy your application on GKE, you must:

1. Package your app into a Docker image
2. Run the container locally on your machine (optional)
3. Upload the image to a registry
4. Create a container cluster
5. Deploy your app to the cluster
6. Expose your app to the Internet
7. Scale up your deployment
8. Deploy a new version of your app

Option A is wrong as kubectl cannot convert the Dockerfile to deployment.

Option B is wrong as Cloud Storage is not Docker image repository.

Option C is wrong as kubectl cannot push Dockerfile to Kubernetes and it does not result into deployment.

### 50. Question

You've created the code for a Cloud Function that will respond to HTTP triggers and return some data in JSON format. You have the code locally; it's tested and working. Which command can you use to create the function inside Google Cloud?

- A. gcloud functions deploy
- B. gcloud function create
- C. gcloud functions create
- D. gcloud function deploy

### Unattempted

Correct answer is A as the code can be deployed using gcloud functions deploy command.

Refer GCP documentation – Cloud Functions Deploy

Deployments work by uploading an archive containing your function's source code to a Google Cloud Storage bucket. You can deploy Cloud Functions from your local machine or from your GitHub or Bitbucket source repository (via Cloud Source Repositories).

Using the gcloud command-line tool, deploy your function from the directory containing your function code with the gcloud functions deploy command:  
gcloud functions deploy NAME --runtime RUNTIME TRIGGER [FLAGS...]

### 51. Question

Your data team is working on some new machine learning models. They are generating several output files per day that they want to store in a regional bucket. They focus on the output files from the last month. The output files older

than a month needs to be cleaned up. With the fewest steps possible, what's the best way to implement the solution?

- A. Create a lifecycle policy to switch the objects older than a month to Coldline storage.
- B. Create a lifecycle policy to delete the objects older than a month.**
- C. Create a Cloud Function triggered when objects are added to a bucket. Look at the date on all the files and delete it, if it's older than a month.
- D. Create a Cloud Function triggered when objects are added to a bucket. Look at the date on all the files and move it to Coldline storage if it's older than a month.

#### Unattempted

Correct answer is B as the files are not needed anymore they can be deleted and need not be stored. The transition of the object can be handled easily using Object Lifecycle Management.

Refer GCP documentation – Cloud Storage Lifecycle Management

You can assign a lifecycle management configuration to a bucket. The configuration contains a set of rules which apply to current and future objects in the bucket. When an object meets the criteria of one of the rules, Cloud Storage automatically performs a specified action on the object. Here are some example use cases:

Downgrade the storage class of objects older than 365 days to Coldline Storage.  
Delete objects created before January 1, 2013.

Keep only the 3 most recent versions of each object in a bucket with versioning enabled.

Option A is wrong as the files are not needed anymore they can be deleted.

Options C & D are wrong as the transition can be handled easily using Object Lifecycle management.

#### 52. Question

You've been tasked with getting all of only the operations team's public SSH keys onto a specific Bastion host instance of a particular project. Currently Project wide access has already been granted to all the instances within the projects. With the fewest steps possible, how do you block or override the project level access on the Bastion host?

- A. Use the gcloud compute instances add-metadata [INSTANCE\_NAME] --metadata block-project-ssh-keys=TRUE command to block the access**
- B. Use the gcloud compute instances add-metadata [INSTANCE\_NAME] --metadata block-project-ssh-keys=FALSE command to block the access.

- C. Use the gcloud compute project-info add-metadata [INSTANCE\_NAME] --metadata block-project-ssh-keys=FALSE command to block the access.
- D. Project wide SSH access cannot be overridden or blocked and needs to be removed.

#### Unattempted

Correct answer is A as the project wide SSH access can be blocked by using the –metadata block-project-ssh-keys=TRUE

Refer GCP documentation – Compute Block Project Keys

If you need your instance to ignore project-wide public SSH keys and use only the instance-level keys, you can block project-wide public SSH keys from the instance. This will only allow users whose public SSH key is stored in instance-level metadata to access the instance. If you want your instance to use both project-wide and instance-level public SSH keys, set the instance metadata to allow project-wide SSH keys. This will allow any user whose public SSH key is stored in project-wide or instance-level metadata to access the instance.

gcloud compute instances add-metadata [INSTANCE\_NAME] –metadata block-project-ssh-keys=TRUE

Option B is wrong as the –metadata block-project-ssh-keys parameter needs to be set to TRUE

Option C is wrong as the command needs to be execute at the instance level.

Option D is wrong as project wide SSH key access can be blocked.

#### 53. Question

You're migrating an on-premises application to Google Cloud. The application uses a component that requires a licensing server. The license server has the IP address 10.28.0.10. You want to deploy the application without making any changes to the code or configuration. How should you go about deploying the application?

- A. Create a subnet with a CIDR range of 10.28.0.0/31. Reserve a static internal IP address of 10.28.0.10. Assign the static address to the license server instance.
- B. Create a subnet with a CIDR range of 10.28.0.0/30. Reserve a static internal IP address of 10.28.0.10. Assign the static address to the license server instance.
- C. Create a subnet with a CIDR range of 10.28.0.0/29. Reserve a static internal IP address of 10.28.0.10. Assign the static address to the license server instance.
- D. Create a subnet with a CIDR range of 10.28.0.0/28. Reserve a static internal IP address of 10.28.0.10. Assign the static address to the license server instance.

#### Unattempted

Correct answer is D as only the CIDR range 10.28.0.0/28 would include the 10.28.0.10 address. It provides 16 ip addresses i.e. 10.28.0.0 to 10.28.0.15

Option A is wrong as 10.28.0.0/31 CIDR range provides 2 ip addresses i.e.

10.28.0.0 to 10.28.0.1

Option B is wrong as 10.28.0.0/30 CIDR range provides 4 ip addresses i.e.

10.28.0.0 to 10.28.0.3

Option C is wrong as 10.28.0.0/29 CIDR range provides 8 ip addresses i.e.

10.28.0.0 to 10.28.0.7

#### 54. Question

While looking at your application's source code in your private Github repo, you've noticed that a service account key has been committed to git. What steps should you take next?

- A. Delete the project and create a new one.
- B. Do nothing. Git is fine for keys if the repo is private.
- C. **Revoke the key, remove the key from Git, purge the Git history to remove all traces of the file, ensure the key is added to the .gitignore file.**
- D. Contact Google Cloud Support

**Unattempted**

Correct answer is C as all the traces of the keys needs to removed and add the key to .gitignore file.

Option A is wrong as deleting project does not remove the keys from Git.

Option B is wrong as it is bad practice to store keys in Git, irrespective of private repo.

Option D is wrong as Google Cloud support cannot help.

#### 55. Question

You need to help a developer install the App Engine Go extensions. However, you've forgotten the exact name of the component. Which command could you run to show all of the available options?

- A. gcloud config list
- B. gcloud component list
- C. gcloud config components list
- D. **gcloud components list**

**Unattempted**

Correct answer is D as gcloud components list provides the list of components with the installation status.

Refer GCP documentation – Cloud SDK Components List

gcloud components list – list the status of all Cloud SDK components

This command lists all the available components in the Cloud SDK. For each component, the command lists the following information:

Status on your local workstation: not installed, installed (and up to date), and update available (installed, but not up to date)

Name of the component (a description)

ID of the component (used to refer to the component in other [gcloud components] commands)  
Size of the component  
In addition, if the --show-versions flag is specified, the command lists the currently installed version (if any) and the latest available version of each individual component.  
Options A & C are wrong as config helps view and edit Cloud SDK properties. It does not provide components detail.  
Option B is wrong as it is not a valid command.

56. Question

Your finance team is working with the engineering team to try and determine your spending for each service by day and month across all projects used by the billing account. What is the easiest and most flexible way to aggregate and analyze the data?

- A. Export the data for the billing account(s) involved to a JSON File; use a Cloud Function to listen for a new file in the Storage bucket; code the function to analyze the service data for the desired projects, by day and month.
- B. Export the data for the billing account(s) involved to BigQuery; then use BigQuery to analyze the service data for the desired projects, by day and month.
- C. Export the data for the billing account(s) to File, import the files into a SQL database; and then use BigQuery to analyze the service data for the desired projects, by day and month.
- D. Use the built-in reports, which already show this data.

Unattempted

Correct answer is B as the billing data can be exported to BigQuery for running daily and monthly to calculate spending across services.

Refer GCP documentation – Cloud Billing Export to BigQuery  
Tools for monitoring, analyzing and optimizing cost have become an important part of managing development. Billing export to BigQuery enables you to export your daily usage and cost estimates automatically throughout the day to a BigQuery dataset you specify. You can then access your billing data from BigQuery. You can also use this export method to export data to a JSON file.  
Options A & C are wrong as they are not easy and flexible.  
Option D is wrong as there are no built-in reports.

57. Question

A company wants to deploy their application using Deployment Manager. However, they want to understand how the changes will affect before implementing the updated. How can the company achieve the same?

- A. Use Deployment Manager Validate Deployment feature
- B. Use Deployment Manager Dry Run feature



### C. Use Deployment Manager Preview feature



### D. Use Deployment Manager Snapshot feature

**Unattempted**

Correct answer is C as Deployment Manager provides the preview feature to check on what resources would be created.

Refer GCP documentation – Deployment Manager Preview

After you have written a configuration file, you can preview the configuration before you create a deployment. Previewing a configuration lets you see the resources that Deployment Manager would create but does not actually instantiate any actual resources. The Deployment Manager service previews the configuration by:

1. Expanding the full configuration, including any templates.
2. Creating a deployment and “shell” resources.

You can preview your configuration by using the preview query parameter when making an insert() request.

```
gcloud deployment-manager deployments create example-deployment --config configuration-file.yaml \ --preview
```

58. **58. Question**

Your company needs to create a new Kubernetes Cluster on Google Cloud Platform. They want the nodes to be configured for resiliency and high availability with no manual intervention. How should the Kubernetes cluster be configured?



### A. Enable auto-healing for the managed instance groups



### B. Enable auto-upgrades for the nodes



### **C. Enable auto-repairing for the nodes**



### D. Enable auto-healing for the nodes

**Unattempted**

Correct answer is C as the resiliency and high availability can be increased using the node auto-repair feature, which would allow Kubernetes engine to replace unhealthy nodes.

Refer GCP documentation – Kubernetes Auto-Repairing

GKE’s node auto-repair feature helps you keep the nodes in your cluster in a healthy, running state. When enabled, GKE makes periodic checks on the health state of each node in your cluster. If a node fails consecutive health checks over an extended time period, GKE initiates a repair process for that node.

Option A is wrong as this cannot be implemented for the Kubernetes cluster.

Option B is wrong as auto-upgrades are to upgrade the node version to the latest stable Kubernetes version.

Option D is wrong as there is no auto-healing feature.

59. **59. Question**

You have created an App engine application in the development environment. The testing for the application has been successful. You want to move the

application to production environment. How can you deploy the application with minimal steps?

- A. Activate the production config, perform app engine deploy
- B. Perform app engine deploy using the --project parameter**
- C. Clone the app engine application to the production environment
- D. Change the project parameter in app.yaml and redeploy

**Unattempted**

Correct answer is B as the gcloud app deploy allows the --project parameter to be passed to override the project that the app engine application needs to be deployed to.

Refer GCP documentation – Cloud SDK

-project=PROJECT\_ID

The Google Cloud Platform project name to use for this invocation. If omitted, then the current project is assumed; the current project can be listed using gcloud config list --format='text(core.project)' and can be set using gcloud config set project PROJECTID. Overrides the default core/projectproperty value for this command invocation.

Option A is wrong as it is a two step process, although a valid solution

Option C is wrong as Clone of the application is possible

Option D is wrong app.yaml does not control the project it is deployed to.

60. **60. Question**

Your company hosts multiple applications on Compute Engine instances. They want the instances to be resilient to any instance crashes or system termination. How would you configure the instances?

- A. Set automaticRestart availability policy to true**
- B. Set automaticRestart availability policy to false
- C. Set onHostMaintenance availability policy to migrate instances
- D. Set onHostMaintenance availability policy to terminate instances

**Unattempted**

Correct answer is A as automaticRestart availability policy determines how the instance reacts to the crashes and system termination and should be set to true to restart the instance.

Refer GCP documentation – Instance Scheduling Options

A VM instance's availability policy determines how it behaves when an event occurs that requires Google to move your VM to a different host machine. For example, you can choose to keep your VM instances running while Compute Engine live migrates them to another host or you can choose to terminate your instances instead. You can update an instance's availability policy at any time to control how you want your VM instances to behave.

You can change an instance's availability policy by configuring the following two settings:

The VM instance's maintenance behavior, which determines whether the instance is live migrated or terminated when there is a maintenance event. The instance's restart behavior, which determines whether the instance automatically restarts if it crashes or gets terminated. The default maintenance behavior for instances is to live migrate, but you can change the behavior to terminate your instance during maintenance events instead. Configure an instance's maintenance behavior and automatic restart setting using the onHostMaintenance and automaticRestart properties. All instances are configured with default values unless you explicitly specify otherwise. onHostMaintenance: Determines the behavior when a maintenance event occurs that might cause your instance to reboot. [Default] migrate, which causes Compute Engine to live migrate an instance when there is a maintenance event. terminate, which terminates an instance instead of migrating it. automaticRestart: Determines the behavior when an instance crashes or is terminated by the system. [Default] true, so Compute Engine restarts an instance if the instance crashes or is terminated. false, so Compute Engine does not restart an instance if the instance crashes or is terminated. Option B is wrong as automaticRestart availability policy should be set to true. Options C & D are wrong as the onHostMaintenance does not apply to crashes or system termination.

## 61. Question

Your organization requires that metrics from all applications be retained for 5 years for future analysis in possible legal proceedings. Which approach should you use?

- A. Grant the security team access to the logs in each Project
- B. Configure Stackdriver Monitoring for all Projects, and export to BigQuery
- C. Configure Stackdriver Monitoring for all Projects with the default retention policies
- D. Configure Stackdriver Monitoring for all Projects, and export to Google Cloud Storage

**Unattempted**

Correct answer is B as Stackdriver monitoring metrics can be exported to BigQuery or Google Cloud Storage. However as the need is for future analysis, BigQuery is a better option.

Refer GCP documentation – Stackdriver

Stackdriver Logging provides you with the ability to filter, search, and view logs from your cloud and open source application services. Allows you to define metrics based on log contents that are incorporated into dashboards and alerts. Enables you to export logs to BigQuery, Google Cloud Storage, and Pub/Sub.

Option A is wrong as project logs are maintained in Stackdriver and it has limited data retention capability.

Option C is wrong as Stackdriver cannot retain data for 5 year. Refer Stackdriver data retention

Option D is wrong as Google Cloud Storage does not provide analytics capability.

62. **62. Question**

A recent software update to a static e-commerce website running on Google Cloud has caused the website to crash for several hours. The CTO decides that all critical changes must now have a back-out/roll-back plan. The website is deployed Cloud Storage and critical changes are frequent. Which action should you take to implement the back-out/roll-back plan?

- A. Create a Nearline copy for the website static data files stored in Google Cloud Storage.
- B. Enable object versioning on the website's static data files stored in Google Cloud Storage.**
- C. Enable Google Cloud Deployment Manager (CDM) on the project, and define each change with a new CDM template.
- D. Create a snapshot of each VM prior to an update, and recover the VM from the snapshot in case of a new version failure.

**Unattempted**

Correct answers are B as this is a seamless way to ensure the last known good version of the static content is always available.

Option A is wrong as this copy process is unreliable and makes it tricky to keep things in sync, it also doesn't provide a way to rollback once a bad version of the data has been written to the copy.

Option C is wrong as this would add a great deal of overhead to the process and would cause conflicts in association between different Deployment Manager deployments which could lead to unexpected behavior if an old version is changed.

Option D is wrong as this approach doesn't scale well, there is a lot of management work involved.

63. **63. Question**

A user wants to install a tool on the Cloud Shell. The tool should be available across sessions. Where should the user install the tool?

- A. /bin
- B. /usr/local/bin
- C. /google/scripts
- D. ~/bin**

**Unattempted**

Correct answer is D as only HOME directory is persisted across sessions.

Refer GCP documentation – Cloud Shell

Cloud Shell provisions 5 GB of free persistent disk storage mounted as

your \$HOME directory on the virtual machine instance. This storage is on a per-user basis and is available across projects. Unlike the instance itself, this storage does not time out on inactivity. All files you store in your home directory, including installed software, scripts and user configuration files like .bashrc and .vimrc, persist between sessions. Your \$HOME directory is private to you and cannot be accessed by other users.

64. 64. Question

Your company has hosted their critical application on Compute Engine managed instance groups. They want the instances to be configured for resiliency and high availability with no manual intervention. How should the managed instance group be configured?

- A. Enable auto-repairing for the managed instance groups
- B. Enable auto-updating for the managed instance groups
- C. Enable auto-restarts for the managed instance groups
- D. Enable auto-healing for the managed instance groups

Unattempted

Correct answer is D as Managed Instance Groups provide AutoHealing feature, which performs a health check and if the application is not responding the instance is automatically recreated.

Refer GCP documentation – Managed Instance Groups

Autohealing — You can also set up an autohealing policy that relies on an application-based health check, which periodically verifies that your application is responding as expected on each of the MIG's instances. If an application is not responding on an instance, that instance is automatically recreated. Checking that an application responds is more precise than simply verifying that an instance is up and running.

Managed instance groups maintain high availability of your applications by proactively keeping your instances available, which means in RUNNING state. A managed instance group will automatically recreate an instance that is not RUNNING. However, relying only on instance state may not be sufficient. You may want to recreate instances when an application freezes, crashes, or runs out of memory.

Application-based autohealing improves application availability by relying on a health checking signal that detects application-specific issues such as freezing, crashing, or overloading. If a health check determines that an application has failed on an instance, the group automatically recreates that instance.

Options A & C are wrong as these features are not available.

Option B is wrong as auto-updating helps deploy new versions of software to instances in a managed instance group. The rollout of an update happens automatically based on your specifications: you can control the speed and scope of the update rollout in order to minimize disruptions to your application. You can optionally perform partial rollouts which allows for canary testing.

65. 65. Question

Your company has deployed their application on managed instance groups, which is served through a network load balancer. They want to enable health checks for the instances. How do you configure the health checks?

- A. Perform the health check using HTTPS by hosting a basic web server
- B. Perform the health check using HTTP by hosting a basic web server
- C. Perform the health check using TCP
- D. Update Managed Instance groups to send a periodic ping to the network load balancer

**Unattempted**

Correct answer is B as Network Load Balancer does not support TCP health checks and hence HTTP health checks need to be performed. You can run a basic web server on each instance for health checks.

Refer GCP documentation – Network Load Balancer Health Checks

Health checks ensure that Compute Engine forwards new connections only to instances that are up and ready to receive them. Compute Engine sends health check requests to each instance at the specified frequency; once an instance exceeds its allowed number of health check failures, it is no longer considered an eligible instance for receiving new traffic. Existing connections will not be actively terminated which allows instances to shut down gracefully and to close TCP connections.

The health checker continues to query unhealthy instances, and returns an instance to the pool when the specified number of successful checks is met. If all instances are marked as UNHEALTHY, the load balancer directs new traffic to all existing instances.

Network Load Balancing relies on legacy HTTP Health checks for determining instance health. Even if your service does not use HTTP, you'll need to at least run a basic web server on each instance that the health check system can query. Option A is wrong as the traffic is not secured, HTTPS health checks are not needed.

Option C is wrong as Network Load Balancer does not support TCP health checks.

Option D is wrong as instances do not need to send any traffic to Network Load Balancer.

## 66. Question

You need to deploy an update to an application in Google App Engine. The update is risky, but it can only be tested in a live environment. What is the best way to introduce the update to minimize risk?

- A. Deploy the application temporarily and be prepared to pull it back if needed.
- B. Warn users that a new app version may have issues and provide a way to contact you if there are problems.

- C. Deploy a new version of the application but use traffic splitting to only direct a small number of users to the new version.
- D. Create a new project with the new app version, and then redirect users to the new version.

#### Unattempted

Correct answer is C as deploying a new version without assigning it as the default version will not create downtime for the application. Using traffic splitting allows for easily redirecting a small amount of traffic to the new version and can also be quickly reverted without application downtime.

Refer GCP documentation – App Engine Splitting Traffic

Traffic migration smoothly switches request routing, gradually moving traffic from the versions currently receiving traffic to one or more versions that you specify. Traffic splitting distributes a percentage of traffic to versions of your application. You can split traffic to move 100% of traffic to a single version or to route percentages of traffic to multiple versions. Splitting traffic to two or more versions allows you to conduct A/B testing between your versions and provides control over the pace when rolling out features.

Option A is wrong as deploying the application version as default requires moving all traffic to the new version. This could impact all users and disable the service. Option B is wrong as this is not a recommended practice and it impacts user experience.

Option D is wrong as App Engine services are intended for hosting different service logic. Using different services would require manual configuration of the consumers of services to be aware of the deployment process and manage from the consumer side who is accessing which service.

#### 67. Question

Your company has reserved a monthly budget for your project. You want to be informed automatically of your project spend so that you can take action when you approach the limit. What should you do?

- B. Create a budget alert for desired percentages such as 50%, 90%, and 100% of your total monthly budget.
- A. Link a credit card with a monthly limit equal to your budget.
- C. In App Engine Settings, set a daily budget at the rate of 1/30 of your monthly budget.
- D. In the GCP Console, configure billing export to BigQuery. Create a saved view that queries your total spend.

#### Unattempted

Correct answer is B as Budget Alerts allow you configure thresholds and if crossed alerts are automatically triggered.

Refer GCP documentation – Billing Budgets Alerts

To help you with project planning and controlling costs, you can set a budget alert. Setting a budget alert lets you track how your spend is growing toward a particular amount.

You can apply budget alerts to either a billing account or a project, and you can

set the budget alert at a specific amount or match it to the previous month's spend. The alerts will be sent to billing administrators and billing account users when spending exceeds a percentage of your budget.

Option A is wrong as linked card does not alert. The charges would still increase as per the usage.

Option C is wrong as App Engine does not have budget settings.

Option D is wrong as the solution would not trigger automatic alerts and the checks would not be immediate as well.

#### 68. 68. Question

Your company plans to archive data to Cloud Storage, which would be needed only in case of any compliance issues, or Audits. What is the command for creating the storage bucket with rare access and named 'archive\_bucket'?

- A. gsutil rm -coldline gs://archive\_bucket
- B. **gsutil mb -c coldline gs://archive\_bucket**
- C. gsutil mb -c nearline gs://archive\_bucket
- D. gsutil mb gs://archive\_bucket

**Unattempted**

Correct answer is B as the data would be rarely accessed, Coldline is an ideal storage class. Also gsutil needs -c parameter to pass the class.

Refer GCP documentation – Storage Classes

Coldline – Data you expect to access infrequently (i.e., no more than once per year). Typically this is for disaster recovery, or data that is archived and may or may not be needed at some future time

Option A is wrong as rm is the wrong parameter and removes the data.

Option C is wrong as Nearline is not suited for data that needs rare access.

Option D is wrong as by default, gsutil would create a regional bucket.

#### 69. 69. Question

An application that relies on Cloud SQL to read infrequently changing data is predicted to grow dramatically. How can you increase capacity for more read-only clients?

- A. Configure high availability on the master node
- B. Establish an external replica in the customer's data center
- C. Use backups so you can restore if there's an outage
- D. **Configure read replicas.**

**Unattempted**

Correct answer is D as read replicas can help handle the read traffic reducing the load from the primary database.

Refer GCP documentation – Cloud SQL Replication Options

Cloud SQL provides the ability to replicate a master instance to one or more read replicas. A read replica is a copy of the master that reflects changes to the

master instance in almost real time.

Option A is wrong as high availability is for failover and not for performance.

Option B is wrong as external replica is not recommended for scaling as it needs to be maintained and the network established for replication.

Option C is wrong as backups are more to restore the database in case of any outage.

#### 70. Question

You've been asked to add a new IAM member and grant them access to run some queries on BigQuery. Considering Google recommended best practices and the principle of least privilege, how would you assign the access?

- A. Create a custom role with roles/bigquery.dataViewer and roles/bigquery.jobUser roles; assign custom role to the users
- B. Create a custom role with roles/bigquery.dataViewer and roles/bigquery.jobUser roles; assign custom role to the group; add users to groups
- C. Assign roles/bigquery.dataViewer and roles/bigquery.jobUser roles to the users
- D.  
**Assign roles/bigquery.dataViewer and roles/bigquery.jobUser roles to a group; add users to groups**

**Unattempted**

Correct answer is D as the user would need

the roles/bigquery.dataViewer and roles/bigquery.jobUser to access and query the BigQuery tables inline with the least privilege. As per google best practices it is recommended to use predefined roles and create groups to control access to multiple users with same responsibility

Refer GCP documentation – IAM Best Practices

Use Cloud IAM to apply the security principle of least privilege, so you grant only the necessary access to your resources.

We recommend collecting users with the same responsibilities into groups and assigning Cloud IAM roles to the groups rather than to individual users. For example, you can create a “data scientist” group and assign appropriate roles to enable interaction with BigQuery and Cloud Storage. When a new data scientist joins your team, you can simply add them to the group and they will inherit the defined permissions.

Options A & B are wrong as the predefined roles can be assigned directly and there is not need to create custom roles.

Option C is wrong as it is recommended to create groups instead of using individual users.

SALE IS ON  | 12 HOURS LEFT | BUY 2 & GET ADDITIONAL 25% OFF | Use Coupon - BLACKFRIDAY



# SKILLCERTPRO

IT CERTIFICATION TRAININGS



Google Cloud / By SkillCertPro

## Practice Set 6

Your results are here!! for " Google Certified Associate Cloud Engineer Practice Test 6 "

0 of 69 questions answered correctly

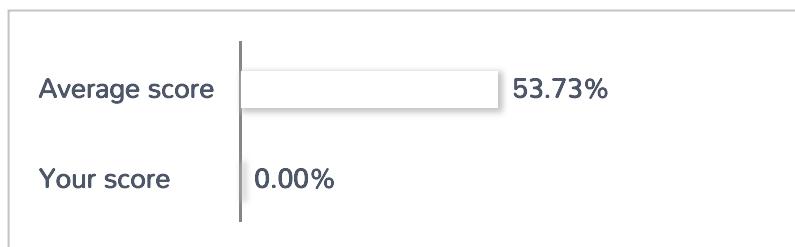
Your time: 00:00:44

Your Final Score is : 0

You have attempted : 0

Number of Correct Questions : 0 and scored 0

Number of Incorrect Questions : 0 and Negative marks 0



You can review your answers by clicking view questions.

**Important Note :** Open Reference Documentation Links in New Tab (Right Click and Open in New Tab).

[Restart Test](#)

[View Answers](#)

|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 |

|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 |
| 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 | 64 | 65 | 66 | 67 | 68 |
|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |

Correct   Incorrect

Review Question

Summary

## 1. Question

You've been asked to add a new IAM member and grant her access to run some queries on BigQuery. Considering the principle of least privilege, which role should you assign?

- A. roles/bigquery.dataEditor and roles/bigquery.jobUser
- B. roles/bigquery.dataViewer and roles/bigquery.user
- C. roles/bigquery.dataViewer and roles/bigquery.jobUser
- D. roles/bigquery.dataOwner and roles/bigquery.jobUser

### Unattempted

Correct answer is C as the user needs to only query the data, they should have access to view the dataset and query the dataset which would provided

by roles/bigquery.dataViewer and roles/bigquery.jobUser inline with the least privilege principle

Refer GCP documentation – BigQuery Access Control

Option A is wrong as roles/bigquery.dataEditor provides more than required privileges

Option B is wrong as roles/bigquery.user provides more than required privileges

Option D is wrong as roles/bigquery.dataOwner provides more than required privileges

## 2. Question

You have a managed instance group comprised of preemptible VM's. All of the VM's keep deleting and recreating themselves every minute. What is a possible cause of this behavior?

- A. Your zonal capacity is limited, causing all preemptible VM's to be shutdown to recover capacity.  
Try deploying your group to another zone.
- B. You have hit your instance quota for the region.
- C. Your managed instance group's VM's are toggled to only last 1 minute in preemptible settings.

- D. Your managed instance group's health check is repeatedly failing, either to a misconfigured health check or misconfigured firewall rules not allowing the health check to access the instances.

#### Unattempted

Correct answer is D as the instances (normal or preemptible) would be terminated and relaunched if the health check fails either due to application not configured properly or the instances firewall do not allow health check to happen.

Refer GCP documentation – Health Check concepts

GCP provides health check systems that connect to virtual machine (VM) instances on a configurable, periodic basis. Each connection attempt is called a probe. GCP records the success or failure of each probe.

Health checks and load balancers work together. Based on a configurable number of sequential successful or failed probes, GCP computes an overall health state for each VM in the load balancer. VMs that respond successfully for the configured number of times are considered healthy. VMs that fail to respond successfully for a separate number of times are unhealthy.

GCP uses the overall health state of each VM to determine its eligibility for receiving new requests. In addition to being able to configure probe frequency and health state thresholds, you can configure the criteria that define a successful probe.

### 3. Question

You write a Python script to connect to Google BigQuery from a Google Compute Engine virtual machine. The script is printing errors that it cannot connect to BigQuery. What should you do to fix the script?

- A. Install the latest BigQuery API client library for Python
- B. Run your script on a new virtual machine with the BigQuery access scope enabled
- C. Create a new service account with BigQuery access and execute your script with that user
- D. Install the bq component for gcloud with the command gcloud components install bq.

#### Unattempted

Correct answer is B as by default an instance is associated with default service account and default access scope, neither of which provides an access to BigQuery. While Service account is the recommended approach and Access scope are legacy, access scope still need to granted to the instance for applications to access the services. So enabling only the Service Account with role would not enable the script to access BigQuery.

Refer GCP documentation – Service Account

When you set up an instance to run as a service account, you determine the level of access the service account has by the IAM roles you grant to the service account. If the service account has no IAM roles, then no API methods can be run by the service account on that instance.

Furthermore, an instance's access scopes determine the default OAuth scopes for requests made through the gcloud tool and client libraries on the instance. As a result, access scopes potentially further limit access to API methods when authenticating through OAuth. However, they do not extend to other authentication protocols like gRPC.

The best practice is to set the full cloud-platform access scope on the instance, then securely limit the service account's access using IAM roles.

Essentially:

IAM restricts access to APIs based on the IAM roles that are granted to the service account.

Access scopes potentially further limit access to API methods when authenticating through OAuth.

You must set access scopes on the instance to authorize access.

While a service account's access level is determined by the IAM roles granted to the service account, an instance's access scopes determine the default OAuth scopes for requests made through the gcloud tool and client libraries on the instance. As a result, access scopes potentially further limit access to API methods when authenticating through OAuth.

Option A is wrong as it is an issue with connectivity to BigQuery and not a client version mismatch issue.

Option C is wrong as adding a service account would not work without having the access granted through access scope.

Option D is wrong as bq command is installed by default and not needed with the python client. It is for direct command line interaction with BigQuery.

#### 4. Question

You have created a Kubernetes engine cluster named 'project-1'. You've realized that you need to change the machine type for the cluster from n1-standard-1 to n1-standard-4. What is the command to make this change?

- A. Create a new node pool in the same cluster, and migrate the workload to the new pool.
- B. gcloud container clusters resize project-1 --machine-type n1-standard-4
- C. gcloud container clusters update project-1 --machine-type n1-standard-4
- D. gcloud container clusters migrate project-1 --machine-type n1-standard-4

#### Unattempted

Correct answer is A as the machine type for the cluster cannot be changed through commands. A new node pool with the updated machine type needs to be created and workload migrated to the new node pool.

Refer GCP documentation – Kubernetes Engine – Migrating Node Pools

A node pool is a subset of machines that all have the same configuration, including machine type (CPU and memory) authorization scopes. Node pools represent a subset of nodes within a cluster; a container cluster can contain one or more node pools.

When you need to change the machine profile of your Compute Engine cluster, you can create a new node pool and then migrate your workloads over to the new node pool.

To migrate your workloads without incurring downtime, you need to:

Mark the existing node pool as unschedulable.

Drain the workloads running on the existing node pool.

Delete the existing node pool.

## 5. Question

You need to have a backup/rollback plan in place for your application that is distributed across a large managed instance group. What is the preferred method for doing so?

- A. Use the Rolling Update feature to deploy/roll back versions with different managed instance group templates.
- B. Use the managed instance group snapshot function that is included in Compute Engine.
- C. Have each instance write critical application data to a Cloud Storage bucket.
- D. Schedule a cron job to take snapshots of each instance in the group.

### Unattempted

Correct answer is A as rolling update helps to apply the update on a controlled number of instances to maintain high availability and ability to rollback in case of any issues.

Refer GCP documentation – Updating Managed Instance Groups

A managed instance group contains one or more virtual machine instances that are controlled using an instance template. To update instances in a managed instance group, you can make update requests to the group as a whole, using the Managed Instance Group Updater feature.

The Managed Instance Group Updater allows you to easily deploy new versions of software to instances in your managed instance groups, while controlling the speed of deployment, the level of disruption to your service, and the scope of the update. The Updater offers two primary advantages:

The rollout of an update happens automatically to your specifications, without the need for additional user input after the initial request.

You can perform partial rollouts which allows for canary testing.

By allowing new software to be deployed inside an existing managed instance group, there is no need for you to reconfigure the instance group or reconnect load balancing, autoscaling, or autohealing each time new version of software is rolled out. Without the Updater, new software versions must be deployed either by creating a new managed instance group with a new software version, requiring additional set up each time, or through a manual, user-initiated, instance-by-instance recreate. Both of these approaches require significant manual steps throughout the process.

A rolling update is an update that is gradually applied to all instances in an instance group until all instances have been updated. You can control various aspects of a rolling update, such as how many

instances can be taken offline for the update, how long to wait between updating instances, whether the update affects all or just a portion of instances, and so on.

Options B, C & D are wrong as the key for scaling is to create stateless, disposable VMs to be able scale and have seamless deployment.

## 6. Question

You have a project using BigQuery. You want to list all BigQuery jobs for that project. You want to set this project as the default for the bq command-line tool. What should you do?

- A. Use gcloud config set project to set the default project
- B. Use bq config set project to set the default project.
- C. Use gcloud generate config-url to generate a URL to the Google Cloud Platform Console to set the default project.
- D. Use bq generate config-url to generate a URL to the Google Cloud Platform Console to set the default project.

### Unattempted

Correct answer is A as you need to use gcloud to manage the config/defaults.

Refer GCP documentation – Cloud SDK Config Set

`-project=PROJECT_ID`

The Google Cloud Platform project name to use for this invocation. If omitted, then the current project is assumed; the current project can be listed using `gcloud config list --format='text(core.project)'` and can be set using `gcloud config set project PROJECTID`. Overrides the default core/project property value for this command invocation.

Option B is wrong as the bq command-line tool assumes the gcloud configuration settings and can't be set through BigQuery.

Option C is wrong as entering this command will not achieve the desired result and will generate an error.

Option D is wrong as entering this command will not achieve the desired result and will generate an error.

## 7. Question

You're deploying an application to a Compute Engine instance, and it's going to need to make calls to read from Cloud Storage and Bigtable. You want to make sure you're following the principle of least privilege.

What's the easiest way to ensure the code can authenticate to the required Google Cloud APIs?

- A. Create a new user account with the required roles. Store the credentials in Cloud Key Management Service and download them to the instance in code.

- B. Use the default Compute Engine service account and set its scopes. Let the code find the default service account using Application Default Credentials.
- C. Create a new service account and key with the required limited permissions. Set the instance to use the new service account. Edit the code to use the service account key.
- D. Register the application with the Binary Registration Service and apply the required roles.

#### Unattempted

Correct answer is C as the best practice is to use a Service Account to grant the application the required access.

Refer GCP documentation – Service Accounts

A service account is a special type of Google account that belongs to your application or a virtual machine (VM), instead of to an individual end user. Your application assumes the identity of the service account to call Google APIs, so that the users aren't directly involved.

A service account is a special type of Google account that represents a Google Cloud service identity or app rather than an individual user. Like users and groups, service accounts can be assigned IAM roles to grant access to specific resources. Service accounts authenticate with a key rather than a password.

Google manages and rotates the service account keys for code running on GCP. We recommend that you use service accounts for server-to-server interactions.

Option A is wrong as it is not the recommended approach

Option B is wrong as the default Service Account does not have the required permissions.

Option D is wrong as there is Binary Registration service.

## 8. Question

You've been trying to deploy a container to Kubernetes; however, kubectl doesn't seem to be able to connect to the cluster. Of the following, what is the most likely cause and how can you fix it?

- A. The firewall rules are preventing the connection. Open up the firewall rules to allow traffic to port 1337.
- B. The kubeconfig is missing the credentials. Run the gcloud container clusters get-credentials command.
- C. The kubeconfig is missing the credentials. Run the gcloud container clusters auth login command.
- D. The firewall rules are preventing the connection. Open up the firewall rules to allow traffic to port 3682.

#### Unattempted

Correct answer is B as the connection is refused, the context needs to be set using the gcloud container clusters get-credentials command

Refer GCP documentation – Kubernetes Engine Troubleshooting

kubectl commands return “connection refused” error

Set the cluster context with the following command:

gcloud container clusters get-credentials [CLUSTER\_NAME]

If you are unsure of what to enter for CLUSTER\_NAME, use the following command to list your clusters:

gcloud container clusters list

Options A & D are wrong as only SSH access is required and it is automatically added.

Option C is wrong as auth login would be needed if the Resource was not found.

## 9. Question

Your engineers have hardcoded the database credentials to be used by application on Kubernetes Engine.

The YAML they're using looks similar to the following:

```
apiVersion: "extensions/v1beta1"
```

```
kind: "Deployment"
```

```
metadata:
```

```
name: "products-service"
```

```
namespace: "default"
```

```
labels:
```

```
app: "products-service"
```

```
spec:
```

```
replicas: 3
```

```
selector:
```

```
matchLabels:
```

```
app: "products-service"
```

```
template:
```

```
metadata:
```

```
labels:
```

```
app: "products-service"
```

```
spec:
```

```
containers:
```

```
– name: "products"
```

```
image: "gcr.io/find-seller-app-dev/products:latest"
```

```
env:
```

```
– name: "database_user"
```

```
value: "admin"
```

```
– name: "database_password"
```

```
value: "TheB3stP@ssW0rd"
```

What is Google's recommended best practice for working with sensitive information inside of Kubernetes?

- A. Store the credentials in a ConfigMap.

- B. Mount the credentials in a volume.
- C. Use an environment variable.
- D. Store the credentials in a Secret.

### Unattempted

Correct answer is D as the Kubernetes allows credentials to be stored in Secret, which can be used by the containers.

Refer GCP documentation – Kubernetes Secrets

Kubernetes offers the Secret resource type to store credentials inside the container cluster and use them in the applications deployed on GKE directly.

Kubernetes secret objects let you store and manage sensitive information, such as passwords, OAuth tokens, and ssh keys. Putting this information in a secret is safer and more flexible than putting it verbatim in a Pod Lifecycle definition or in a container image.

Option A is wrong as ConfigMaps bind configuration files, command-line arguments, environment variables, port numbers, and other configuration artifacts to your Pods' containers and system components at runtime. ConfigMaps allow you to separate your configurations from your Pods and components, which helps keep your workloads portable, makes their configurations easier to change and manage, and prevents hardcoding configuration data to Pod specifications.

Option B is wrong as credentials cannot be mounted in the volume.

Option C is wrong as environment variable does not secure the credentials.

## 10. Question

A SysOps admin has configured a lifecycle rule on an object versioning disabled multi-regional bucket.

Which of the following statement effect reflects the following lifecycle config?

```
{
 "rule": [
 {
 "action": {"type": "Delete"},
 "condition": {"age": 30, "isLive": false}
 },
 {
 "action": {"type": "SetStorageClass", "storageClass": "COLDLINE"},
 "condition": {"age": 365, "matchesStorageClass": "MULTI_REGIONAL"}
 }
]
}
```

- A. Archive objects older than 30 days and move objects to Coldline Storage after 365 days if the storage class in Multi-regional
- B. Delete objects older than 30 days and move objects to Coldline Storage after 365 days if the storage class in Multi-regional.
- C. Delete archived objects older than 30 days and move objects to Coldline Storage after 365 days if the storage class in Multi-regional.
- D. Move objects to Coldline Storage after 365 days if the storage class in Multi-regional First rule has no effect on the bucket.

### Unattempted

Correct answer is D.

First rule will delete any object if it has a age over 30 days and is not live (not the latest version).

However as the bucket is not versioning enabled it does not have any effect. Second rule will change the storage class of the live object from multi-regional to Coldline for objects with age over 365 days.

Refer GCP documentation – Object Lifecycle

The following conditions are supported for a lifecycle rule:

Age: This condition is satisfied when an object reaches the specified age (in days). Age is measured from the object's creation time. For example, if an object's creation time is 2019/01/10 10:00 UTC and the Age condition is 10 days, then the condition is satisfied for the object on and after 2019/01/20 10:00 UTC. This is true even if the object becomes archived through object versioning sometime after its creation.

CreatedBefore: This condition is satisfied when an object is created before midnight of the specified date in UTC.

IsLive: If the value is true, this lifecycle condition matches only live objects; if the value is false, it matches only archived objects. For the purposes of this condition, objects in non-versioned buckets are considered live.

MatchesStorageClass: This condition is satisfied when an object in the bucket is stored as the specified storage class. Generally, if you intend to use this condition on Multi-Regional Storage or Regional Storage objects, you should also include STANDARD and DURABLE\_REDUCED\_AVAILABILITY in the condition to ensure all objects of similar storage class are covered.

Option A is wrong as the first rule does not archive but deletes the archived objects, but does not have any impact on a versioning disabled bucket.

Option B is wrong as the first rule does not delete live objects but only archives objects.

Option C is wrong as first rule does not have any impact on a versioning disabled bucket.

## 11. Question

A SysOps admin has configured a lifecycle rule on an object versioning enabled multi-regional bucket.

Which of the following statement effect reflects the following lifecycle config?

```
{
 "rule": [
 {
 "action": {"type": "Delete"},
 "condition": {"age": 30, "isLive": false}
 },
 {
 "action": {"type": "SetStorageClass", "storageClass": "COLDLINE"},
 "condition": {"age": 365, "matchesStorageClass": "MULTI_REGIONAL"}
 }
]
}
```

- A. Archive objects older than 30 days and move objects to Coldline Storage after 365 days if the storage class in Multi-regional
- B. Delete objects older than 30 days and move objects to Coldline Storage after 365 days if the storage class in Multi-regional.
- C. Delete archived objects older than 30 days and move objects to Coldline Storage after 365 days if the storage class in Multi-regional.
- D. Move objects to Coldline Storage after 365 days if the storage class in Multi-regional First rule has no effect on the bucket.

### Unattempted

Correct answer is C.

First rule will delete any object if it has a age over 30 days and is not live (not the latest version). Second rule will change the storage class of the live object from multi-regional to Coldline for objects with age over 365 days.

Refer GCP documentation – Object Lifecycle

The following conditions are supported for a lifecycle rule:

Age: This condition is satisfied when an object reaches the specified age (in days). Age is measured from the object's creation time. For example, if an object's creation time is 2019/01/10 10:00 UTC and the Age condition is 10 days, then the condition is satisfied for the object on and after 2019/01/20 10:00 UTC. This is true even if the object becomes archived through object versioning sometime after its creation.

CreatedBefore: This condition is satisfied when an object is created before midnight of the specified date in UTC.

IsLive: If the value is true, this lifecycle condition matches only live objects; if the value is false, it matches only archived objects. For the purposes of this condition, objects in non-versioned buckets are

considered live.

**MatchesStorageClass:** This condition is satisfied when an object in the bucket is stored as the specified storage class. Generally, if you intend to use this condition on Multi-Regional Storage or Regional Storage objects, you should also include STANDARD and DURABLE\_REDUCED\_AVAILABILITY in the condition to ensure all objects of similar storage class are covered.

Option A is wrong as the first rule does not archive but deletes the archived objects.

Option B is wrong as the first rule does not delete live objects but only archives objects.

Option D is wrong as first rule applies to archived or not live objects.

## 12. Question

A SysOps admin has configured a lifecycle rule on an object versioning enabled multi-regional bucket.

Which of the following statement effect reflects the following lifecycle config?

```
{
 "rule": [
 {
 "action": {"type": "Delete"},
 "condition": {"age": 30, "isLive": false}
 },
 {
 "action": {"type": "SetStorageClass", "storageClass": "COLDLINE"},
 "condition": {"age": 365, "matchesStorageClass": "MULTI_REGIONAL"}
 }
]
}
```

- A. Archive objects older than 30 days and move objects to Coldline Storage after 365 days if the storage class in Multi-regional
- B. Delete objects older than 30 days and move objects to Coldline Storage after 365 days if the storage class in Multi-regional.
- C. Delete archived objects older than 30 days and move objects to Coldline Storage after 365 days if the storage class in Multi-regional.
- D. Move objects to Coldline Storage after 365 days if the storage class in Multi-regional First rule has no effect on the bucket.

### Unattempted

Correct answer is C.

First rule will delete any object if it has a age over 30 days and is not live (not the latest version). Second

rule will change the storage class of the live object from multi-regional to Coldline for objects with age over 365 days.

Refer GCP documentation – Object Lifecycle

The following conditions are supported for a lifecycle rule:

**Age:** This condition is satisfied when an object reaches the specified age (in days). Age is measured from the object's creation time. For example, if an object's creation time is 2019/01/10 10:00 UTC and the Age condition is 10 days, then the condition is satisfied for the object on and after 2019/01/20 10:00 UTC. This is true even if the object becomes archived through object versioning sometime after its creation.

**CreatedBefore:** This condition is satisfied when an object is created before midnight of the specified date in UTC.

**IsLive:** If the value is true, this lifecycle condition matches only live objects; if the value is false, it matches only archived objects. For the purposes of this condition, objects in non-versioned buckets are considered live.

**MatchesStorageClass:** This condition is satisfied when an object in the bucket is stored as the specified storage class. Generally, if you intend to use this condition on Multi-Regional Storage or Regional Storage objects, you should also include STANDARD and DURABLE\_REDUCED\_AVAILABILITY in the condition to ensure all objects of similar storage class are covered.

Option A is wrong as the first rule does not archive but deletes the archived objects.

Option B is wrong as the first rule does not delete live objects but only archives objects.

Option D is wrong as first rule applies to archived or not live objects.

### 13. Question

You want to enable your running Google Container Engine cluster to scale as demand for your application changes. What should you do?

- A. Add additional nodes to your Container Engine cluster using the following command: gcloud container clusters resize CLUSTER\_Name --size 10
- B. Add a tag to the instances in the cluster with the following command: gcloud compute instances add-tags INSTANCE --tags --enable-autoscaling max-nodes-10
- C. Update the existing Container Engine cluster with the following command: gcloud alpha container clusters update mycluster --enable-autoscaling --min-nodes=1 --max-nodes=10
- D. Create a new Container Engine cluster with the following command: gcloud alpha container clusters create mycluster --enable-autoscaling --min-nodes=1 --max-nodes=10 and redeploy your application

Unattempted

Correct answer is C as you need to update the cluster to enable auto scaling with min and max nodes to scale as per the demand.

Refer GCP documentation – Cluster Autoscaling

Option A is wrong as it would only increase the nodes.

Option B is wrong as the cluster needs to be updated and not the instances.

Option D is wrong as you do not need to create a new cluster and the existing cluster can be updated to enable auto scaling.

## 14. Question

You've set up an instance inside your new network and subnet. Your firewall rules are set to target all instances in your network with the following firewall rules.

NAME:deny-all | NETWORK:devnet | DIRECTION:INGRESS | PRIORITY:1000 | DENY:tcp:0-65535,udp:0-6553

NAME:open-ssh | NETWORK:devnet | DIRECTION:INGRESS | PRIORITY:5000 | ALLOW:tcp:22

However, when you attempt to connect to your instance via SSH, your connection is timing out. What is the most likely cause?

- A. SSH would be denied and would need instance reboot for the allow rule to take effect
- B. SSH key hasn't been uploaded to the instance.
- C. Firewall rule needs to be applied to the instance specifically.
- D. SSH would be denied as the deny rule overrides the allow

### Unattempted

Correct answer is D as the firewall rules are applied as per the priority and as the deny rule has the higher priority as compared to the allow rule, the SSH access is denied.

Refer GCP documentation – VPC Firewall Rules – Priority

The firewall rule priority is an integer from 0 to 65535, inclusive. Lower integers indicate higher priorities. If you do not specify a priority when creating a rule, it is assigned a priority of 1000.

The relative priority of a firewall rule determines if it is applicable when evaluated against others. The evaluation logic works as follows:

The highest priority rule applicable to a target for a given type of traffic takes precedence. Target specificity does not matter. For example, a higher priority ingress rule for certain ports and protocols intended for all targets overrides a similarly defined rule for the same ports and protocols intended for specific targets.

The highest priority rule applicable for a given protocol and port definition takes precedence, even when the protocol and port definition is more general. For example, a higher priority ingress rule allowing traffic for all protocols and ports intended for given targets overrides a lower priority ingress rule denying TCP 22 for the same targets.

A rule with a deny action overrides another with an allow action only if the two rules have the same priority. Using relative priorities, it is possible to build allowrules that override deny rules, and vice versa. Rules with the same priority and the same action have the same result. However, the rule that is used during the evaluation is indeterminate. Normally, it doesn't matter which rule is used except when you enable firewall rule logging. If you want your logs to show firewall rules being evaluated in a consistent and well-defined order, assign them unique priorities.

Option A is wrong as firewall rules are applied directly and do not require an instance restart.

Option B is wrong as SSH are autogenerated and transferred to the instance.

Option C is wrong as firewall are not applied to instance directly but through network tags.

## 15. Question

You've set up an instance inside your new network and subnet. You create firewall rules to target all instances in your network with the following firewall rules.

NAME:open-ssh | NETWORK:devnet | DIRECTION:INGRESS | PRIORITY:1000 | ALLOW:tcp:22

NAME:deny-all | NETWORK:devnet | DIRECTION:INGRESS | PRIORITY:5000 | DENY:tcp:0-65535,udp:0-6553

If you try to SSH to the instance, what would be the result?

- A. SSH would be denied and would need gcloud firewall refreshcommand for the allow rule to take effect.
- B. SSH would be allowed as the allow rule overrides the deny
- C. SSH would be denied as the deny rule overrides the allow
- D. SSH would be denied and would need instance reboot for the allow rule to take effect

### Unattempted

Correct answer is B as the firewall rules are applied as per the priority and as the allow rule has the higher priority as compared to the deny rule, the SSH access is allowed.

Refer GCP documentation – VPC Firewall Rules – Priority

The firewall rule priority is an integer from 0 to 65535, inclusive. Lower integers indicate higher priorities.

If you do not specify a priority when creating a rule, it is assigned a priority of 1000.

The relative priority of a firewall rule determines if it is applicable when evaluated against others. The evaluation logic works as follows:

The highest priority rule applicable to a target for a given type of traffic takes precedence. Target specificity does not matter. For example, a higher priority ingress rule for certain ports and protocols intended for all targets overrides a similarly defined rule for the same ports and protocols intended for specific targets.

The highest priority rule applicable for a given protocol and port definition takes precedence, even when the protocol and port definition is more general. For example, a higher priority ingress rule allowing traffic

for all protocols and ports intended for given targets overrides a lower priority ingress rule denying TCP 22 for the same targets.

A rule with a deny action overrides another with an allow action only if the two rules have the same priority. Using relative priorities, it is possible to build allowrules that override deny rules, and vice versa. Rules with the same priority and the same action have the same result. However, the rule that is used during the evaluation is indeterminate. Normally, it doesn't matter which rule is used except when you enable firewall rule logging. If you want your logs to show firewall rules being evaluated in a consistent and well-defined order, assign them unique priorities.

Options A, C & D are wrong the SSH access would be allowed.

## 16. Question

Your company has a number of internal backends that they do not want to be exposed to the public Internet. How can they reduce their external exposure while still allowing maintenance access to resources when working remotely?

- A. Remove the external IP address and use Cloud Shell to access internal-only resources
- B. Remove the external IP address and use a bastion host to access internal-only resources.
- C. Remove the external IP address and have remote employees dial into the company VPN connection for maintenance work.
- D. Hide the external IP address behind a load balancer and restrict load balancer access to the internal company network.

### Unattempted

Correct answer is B as it is a best practice to remove external ip address from the instances so that they are not reachable from the internet and have a Bastion host or Jump server to be able to login into the servers.

Refer GCP documentation – Bastion Hosts

Bastion hosts provide an external facing point of entry into a network containing private network instances. This host can provide a single point of fortification or audit and can be started and stopped to enable or disable inbound SSH communication from the Internet.

By using a bastion host, you can connect to an instance that does not have an external IP address. This approach allows you to connect to a development environment or manage the database instance for your external application, for example, without configuring additional firewall rules.

A complete hardening of a bastion host is outside the scope of this article, but some initial steps taken can include:

Limit the CIDR range of source IPs that can communicate with the bastion.

Configure firewall rules to allow SSH traffic to private instances from only the bastion host.

By default, SSH on instances is configured to use private keys for authentication. When using a bastion

host, you log into the bastion host first, and then into your target private instance. Because of this two-step login, which is why bastion hosts are sometimes called “jump servers,” you should use ssh-agent forwarding instead of storing the target machine’s private key on the bastion host as a way of reaching the target machine. You need to do this even if using the same key-pair for both bastion and target instances, as the bastion has direct access to only the public half of the key-pair.

## 17. Question

The development team has provided you with a Kubernetes Deployment file. You have no infrastructure yet and need to deploy the application. What should you do?

- A. Use gcloud to create a Kubernetes cluster. Use Deployment Manager to create the deployment.
- B. Use gcloud to create a Kubernetes cluster. Use kubectl to create the deployment.
- C. Use kubectl to create a Kubernetes cluster. Use Deployment Manager to create the deployment.
- D. Use kubectl to create a Kubernetes cluster. Use kubectl to create the deployment.

### Unattempted

Correct answer is B as you would need gcloud to create a kubernetes cluster. Once the cluster is created you can use kubectl to manage the deployments.

Refer GCP documentation – Kubernetes Cluster Tutorial

To create a cluster with the gcloud command-line tool, use the gcloud container clusters command:  
gcloud container clusters create hello-cluster –num-nodes=3

To deploy and manage applications on a GKE cluster, you must communicate with the Kubernetes cluster management system. You typically do this by using the kubectl command-line tool.

Kubernetes represents applications as Pods, which are units that represent a container (or group of tightly-coupled containers). The Pod is the smallest deployable unit in Kubernetes. In this tutorial, each Pod contains only your hello-app container.

The kubectl run command below causes Kubernetes to create a Deployment named hello-web on your cluster. The Deployment manages multiple copies of your application, called replicas, and schedules them to run on the individual nodes in your cluster. In this case, the Deployment will be running only one Pod of your application.

kubectl run hello-web –image=gcr.io/\${PROJECT\_ID}/hello-app:v1 –port 8080

Options A & C are wrong as you need kubectl to do a kubernetes deployment.

Options C & D are wrong as you need gcloud to create the kubernetes cluster.

## 18. Question

One of the microservices in your application has an intermittent performance problem. You have not observed the problem when it occurs but when it does, it triggers a particular burst of log lines. You want to debug a machine while the problem is occurring. What should you do?

- A. Log into one of the machines running the microservice and wait for the log storm.
- B. In the Stackdriver Error Reporting dashboard, look for a pattern in the times the problem occurs.
- C. Configure your microservice to send traces to Stackdriver Trace so you can find what is taking so long.
- D. Set up a log metric in Stackdriver Logging, and then set up an alert to notify you when the number of log lines increases past a threshold.

#### Unattempted

Correct answer is D as there is a burst of log lines you can set up a metric that identifies those lines.

Stackdriver will also allow you to set up a text, email or messaging alert that can notify promptly when the error is detected so you can hop onto the system to debug.

Option A is wrong as logging into an individual machine may not see the specific performance problem as multiple machines may be in the configuration and reducing the chances of interacting with an intermittent performance problem.

Option B is wrong as error reporting won't necessarily catch the log lines unless they are stack traces in the proper format. Additionally just because there is a pattern doesn't mean you will know exactly when and where to log in to debug.

Option C is wrong as trace may tell you where time is being spent but won't let you hone in on the exact host that the problem is occurring on because you generally only send samples of traces. There is also no alerting on traces to notify exactly when the problem is happening.

### 19. Question

You're writing a Java application with lots of threading and concurrency. You want your application to run in a sandboxed managed environment with the ability to perform SSH debugging to check on any thread dump for troubleshooting. Which service should you host your application on?

- A. Compute Engine
- B. App Engine Flexible Environment
- C. Cloud Functions
- D. App Engine Standard Environment

#### Unattempted

Correct answer is B as App Engine provides the managed service and Flexible environment supports the ability to perform SSH debugging.

Refer GCP documentation – App Engine Environments

Feature – SSH-debugging

Standard environment – No

### Flexible environment – Yes

Flexible environment instances are permitted to have higher CPU and memory limits than is possible with standard environment instances. This allows flexible instances to run applications that are more memory and CPU intensive. However, it may increase the likelihood of concurrency bugs due to the increase in threads within a single instance.

Developers can SSH to a flexible environment instance and obtain a thread dump to troubleshoot this type of problem.

Option A is wrong as Compute Engine does not provide managed service

Option C is wrong as Cloud Functions provides serverless event driven compute platform.

Option D is wrong as App Engine Standard environment does not provide SSH debugging

## 20. Question

Several employees at your company have been creating projects with Cloud Platform and paying for it with their personal credit cards, which the company reimburses. The company wants to centralize all these projects under a single, new billing account. What should you do?

- A. Contact cloud-billing@google.com with your bank account details and request a corporate billing account for your company.
- B. Create a ticket with Google Support and wait for their call to share your credit card details over the phone.
- C. In the Google Platform Console, go to the Resource Manager and move all projects to the root Organization.
- D. ?In the Google Cloud Platform Console, create a new billing account and set up a payment method.

### Unattempted

Correct answer is C as Google Cloud Resource Manager can help group the existing accounts under an Organization for centralized billing.

Refer GCP documentation – Resource Manager

Google Cloud Platform (GCP) customers need an easy way to centrally manage and control GCP resources, projects and billing accounts that belong to their organization. As companies grow, it becomes progressively difficult to keep track of an ever-increasing number of projects, created by multiple users, with different access control policies and linked to a variety of billing instruments. Google Cloud Resource Manager allows you to group resource containers under the Organization resource, providing full visibility, centralized ownership and unified management of your company's assets on GCP.

Options A & B are wrong as billing consolidation is User responsibility and GCP does not support it.

Option D is wrong as it would not centralize the billing under a single account.

## 21. Question

A company is hosting their Echo application on Google Cloud using Google Kubernetes Engine. The application is deployed with deployment echo-deployment exposed with echo-service. They have a new image that needs to be deployed for the application. How can the change be deployed with minimal downtime?

- A. Update image using kubectl set image deployment
- B. Delete the deployment and create a new deployment with the updated image
- C. Delete the service and create a new service with the updated image
- D. Update image in instance template and use rolling deployment of instance group with Kubernetes engine.

### Unattempted

Correct answer is A as the image can be directly updated using the kubectl command and Kubernetes Engine performs a rolling update.

Refer GCP documentation – Kubernetes Engine Rolling updates

You can perform a rolling update to update the images, configuration, labels, annotations, and resource limits/requests of the workloads in your clusters. Rolling updates incrementally replace your resource's Pods with new ones, which are then scheduled on nodes with available resources. Rolling updates are designed to update your workloads without downtime.

You can use kubectl set to make changes to an object's image, resources (compute resource such as CPU and memory), or selector fields.

For example, to update a Deployment from nginx version 1.7.9 to 1.9.1, run the following command:

```
kubectl set image deployment nginx nginx=nginx:1.9.1
```

The kubectl set image command updates the nginx image of the Deployment's Pods one at a time.

Option B is wrong as creating a new deployment would result in downtime.

Option C is wrong as service does not have a mapping of image.

Option D is wrong as Kubernetes Engine does not work with instance template and managed instance groups

## 22. Question

A member of the finance team informed you that one of the projects is using the old billing account. What steps should you take to resolve the problem?

- A. Go to the Project page; expand the Billing tile; select the Billing Account option; select the correct billing account and save.
- B. Go to the Billing page; view the list of projects; find the project in question and select Change billing account; select the correct billing account and save.

- C. Delete the project and recreate it with the correct billing account.
- D. Submit a support ticket requesting the change.

### Unattempted

Correct answer is B as the billing account for the project can be modified from the Billing page.

Refer GCP documentation – Billing Modify Project

If you are a billing administrator on only one billing account, new projects you create are automatically linked to your existing billing account. If you create and have access to multiple billing accounts, you can change the billing account a project is billed to. This article describes how to change the billing account for your project, as well as how to enable and disable billing for a project.

To change the billing account:

Sign in to the Google Cloud Platform Console.

Open the console navigation menu (menu) and select Billing.

If you have more than one billing account, you'll be prompted to select Go to linked billing account to manage the current project's billing.

From the Billing navigation menu, click Account management.

Under Projects linked to this billing account, locate the name of the project that you want to change billing for, and then click the menu (more\_vert) next to it.

Select Change billing, then choose the desired destination billing account.

## 23. Question

A company uses Cloud Storage for storing their critical data. As a part of compliance, the objects need to be encrypted using customer-supplied encryption keys. How should the object be handled to support customer-supplied encryption?

- A. Use gsutil with —encryption-key to pass the encryption key
- B. Use gsutil with GSUtil:encryption\_key=[YOUR\_ENCRYPTION\_KEY] to pass the encryption key
- C. Use gcloud config to define the encryption
- D. Create bucket with —encryption-key and use gsutil to upload files

### Unattempted

Correct answer is B as the customer supplied encryption key can be passed using the encryption\_key parameter.

Refer GCP documentation – Cloud Storage Encryption

Add the following option to the [GSUtil] section of your boto configuration file:

```
encryption_key = [YOUR_ENCRYPTION_KEY]
```

where [YOUR\_ENCRYPTION\_KEY] is the key for encrypting the uploaded file.

Note: You can alternatively include this information in each gsutil command by using the -o top level

flag: -o "GSUtil:encryption\_key=[YOUR\_ENCRYPTION\_KEY]".

Option A is wrong as the parameter is wrong. Parameter -o "GSUtil:encryption\_key=[YOUR\_ENCRYPTION\_KEY]" can be used.

Option C is wrong as encryption key cannot be defined using gcloud config.

Option D is wrong as encryption is not set on bucket and needs to be applied when the object is uploaded.

## 24. Question

The development team needs a regional MySQL database with point-in-time recovery for a new proof-of-concept application. What's the most inexpensive way to enable point-in-time recovery?

- A. Replicate to a Cloud Spanner database.
- B. Create a read replica in the same region.
- C. Enable binary logging.
- D. Create hourly back-ups.

### Unattempted

Correct answer is C as binary logging helps Point-in-time recovery.

Refer GCP documentation – Cloud SQL MySQL Point In Time Recovery

Point-in-time recovery enables you to recover an instance to a specific point in time. A point-in-time recovery always creates a new instance; you cannot perform a point-in-time recovery to an existing instance.

Options A & B are wrong as Read Replica and Cloud Spanner are not cost-effective options.

Option D is wrong as hourly back-ups does not meet the point-in-time requirement.

## 25. Question

Your application deployed on a Google Compute Engine virtual machine instance needs to connect to Google Cloud Pub/Sub. What is the best way to provision the access to the application?

- A. Whitelist Google Compute Engine virtual machine instance IP on the Cloud Pub/Sub firewall
- B. Build or leverage an OAuth-compatible access control system
- C. Create a new service account with no access and enable access scope to allow Cloud Pub/Sub access for the instance
- D. Create a new service account with Cloud Pub/Sub access and associate with the instance

### Unattempted

Correct answer is D as the VM needs to be granted permissions using the service account to be able to communicate with Cloud Pub/Sub.

Refer GCP documentation – Service Account Permissions

When you set up an instance to run as a service account, the level of access the service account has is determined by the combination of access scopes granted to the instance and IAM roles granted to the service account. You need to configure both access scopes and IAM roles to successfully set up an instance to run as a service account. Essentially:

Access scopes authorize the potential access that an instance has to API methods.

IAM restricts that access to the roles granted to the service account.

Option A is wrong as there is feature to whitelist IPs as firewalls only apply to Compute Engines.

Option B is wrong as service accounts handle the OAuth authentication and it is best suited for service to service communication.

Google OAuth 2.0 system supports server-to-server interactions such as those between a web application and a Google service. For this scenario you need a service account, which is an account that belongs to your application instead of to an individual end user. Your application calls Google APIs on behalf of the service account, so users aren't directly involved. This scenario is sometimes called "two-legged OAuth," or "2LO." (The related term "three-legged OAuth" refers to scenarios in which your application calls Google APIs on behalf of end users, and in which user consent is sometimes required.) Typically, an application uses a service account when the application uses Google APIs to work with its own data rather than a user's data. For example, an application that uses Google Cloud Datastore for data persistence would use a service account to authenticate its calls to the Google Cloud Datastore API. Option C is wrong as access scope and service account IAM role permissions are applied together with the more restrictive taking effect. With no permissions to the Service Account, the instance would not be able to communicate with the VM instance.

## 26. Question

Your company pushes batches of sensitive transaction data from its application server VMs to Cloud Pub/Sub for processing and storage. What is the Google-recommended way for your application to authenticate to the required Google Cloud services?

- A. Ensure that VM service accounts are granted the appropriate Cloud Pub/Sub IAM roles.
- B. Ensure that VM service accounts do not have access to Cloud Pub/Sub, and use VM access scopes to grant the appropriate Cloud Pub/Sub IAM roles.
- C. Generate an OAuth2 access token for accessing Cloud Pub/Sub, encrypt it, and store it in Cloud Storage for access from each VM.
- D. Create a gateway to Cloud Pub/Sub using a Cloud Function, and grant the Cloud Function service account the appropriate Cloud Pub/Sub IAM roles.

**Unattempted**

Correct answer is A as the VM needs to be granted permissions using the service account to be able to communicate with Cloud Pub/Sub.

Refer GCP documentation – Service Account Permissions

When you set up an instance to run as a service account, the level of access the service account has is determined by the combination of access scopes granted to the instance and IAM roles granted to the service account. You need to configure both access scopes and IAM roles to successfully set up an instance to run as a service account. Essentially:

Access scopes authorize the potential access that an instance has to API methods.

IAM restricts that access to the roles granted to the service account.

Google OAuth 2.0 system supports server-to-server interactions such as those between a web application and a Google service. For this scenario you need a service account, which is an account that belongs to your application instead of to an individual end user. Your application calls Google APIs on behalf of the service account, so users aren't directly involved. This scenario is sometimes called "two-legged OAuth," or "2LO." (The related term "three-legged OAuth" refers to scenarios in which your application calls Google APIs on behalf of end users, and in which user consent is sometimes required.) Typically, an application uses a service account when the application uses Google APIs to work with its own data rather than a user's data. For example, an application that uses Google Cloud Datastore for data persistence would use a service account to authenticate its calls to the Google Cloud Datastore API. Option B is wrong as access scope and service account IAM role permissions are applied together with the more restrictive taking effect. With no permissions to the Service Account, the instance would not be able to communicate with the VM instance.

Option C is wrong as service accounts handle the OAuth authentication and it is best suited for service to service communication.

Option D is wrong as there is no need for the gateway. Also, the VM and Cloud Function access needs to be handled.

## 27. Question

You can SSH into an instance from another instance in the same VPC by its internal IP address, but not its external IP address. What is one possible reason why this is so?

- A. The outgoing instance does not have correct permission granted to its service account.
- B. The external IP address is disabled.
- C. The firewall rule to allow SSH is restricted to the internal VPC.
- D. The receiving instance has an ephemeral address instead of a reserved address.

**Unattempted**

Correct answer is C as firewall rules need to be enabled for both the network and external network to be allowed to ssh into the instances.

Refer GCP documentation – VPC Firewalls

Google Cloud Platform (GCP) firewall rules let you allow or deny traffic to and from your virtual machine (VM) instances based on a configuration you specify. GCP firewall rules are applied at the virtual networking level, so they provide effective protection and traffic control regardless of the operating system your instances use.

Every VPC network functions as a distributed firewall. While firewall rules are defined at the network level, connections are allowed or denied on a per-instance basis. You can think of the GCP firewall rules as existing not only between your instances and other networks, but between individual instances within the same network

**Source IP ranges:** You can specify ranges of IP addresses as sources for packets. The ranges can include addresses inside your VPC network and those outside of it. Source IP ranges can be used to define sources both inside and outside of GCP.

## 28. Question

You have an application deployed on Kubernetes Engine using a Deployment named echo-deployment. The deployment is exposed using a Service called echo-service. You need to perform an update to the application with minimal downtime to the application. What should you do?

- A. Use the rolling update functionality of the Instance Group behind the Kubernetes cluster
- B. Update the deployment yaml file with the new container image. Use kubectl delete deployment/echo-deployment and kubectl create –f
- C. Use kubectl set image deployment/echo-deployment
- D. Update the service yaml file with the new container image. Use kubectl delete service/echoservice and kubectl create –f

## Unattempted

Correct answer is C as the image can be directly updated using the kubectl command and Kubernetes Engine performs a rolling update.

Refer GCP documentation – Kubernetes Engine Rolling updates

You can perform a rolling update to update the images, configuration, labels, annotations, and resource limits/requests of the workloads in your clusters. Rolling updates incrementally replace your resource's Pods with new ones, which are then scheduled on nodes with available resources. Rolling updates are designed to update your workloads without downtime.

You can use kubectl set to make changes to an object's image, resources (compute resource such as CPU and memory), or selector fields.

For example, to update a Deployment from nginx version 1.7.9 to 1.9.1, run the following command:

```
kubectl set image deployment nginx nginx=nginx:1.9.1
```

The kubectl set image command updates the nginx image of the Deployment's Pods one at a time.

Option A is wrong as you do not work with underlying managed instance groups. It is managed by Kubernetes.

Option B is wrong as creating a new deployment would result in downtime.

Option D is wrong as service does not have a mapping of image.

## 29. Question

You have created an App engine application in the us-central region. However, you found out the network team has configured all the VPN connections in the asia-east2 region, which are not possible to move. How can you change the location efficiently?

- A. Change the region in app.yaml and redeploy
- B. From App Engine console, change the region of the application
- C. Change the region in application.xml within the application and redeploy
- D. ?Create a new project in the asia-east2 region and create app engine in the project

### Unattempted

Correct answer is D as app engine is a regional resource, it needs to be redeployed to the different region.

Refer GCP documentation – App Engine locations

App Engine is regional, which means the infrastructure that runs your apps is located in a specific region and is managed by Google to be redundantly available across all the zones within that region.

Meeting your latency, availability, or durability requirements are primary factors for selecting the region where your apps are run. You can generally select the region nearest to your app's users but you should consider the location of the other GCP products and services that are used by your app. Using services across multiple locations can affect your app's latency as well as pricing

You cannot change an app's region after you set it.

Options A, B & C are wrong as once the region is set for the app engine it cannot be modified.

## 30. Question

You want to verify the IAM users and roles assigned within a GCP project named my-project. What should you do?

- A. Run gcloud iam roles list. Review the output section.
- B. Run gcloud iam service-accounts list. Review the output section.

- C. Navigate to the project and then to the IAM section in the GCP Console. Review the members and roles.
- D. Navigate to the project and then to the Roles section in the GCP Console. Review the roles and status.

#### Unattempted

Correct answer is C as IAM section provides the list of both Members and Roles.

Option A is wrong as it would provide information about the roles only.

Option B is wrong as it would provide only the service accounts.

Option D is wrong as it would provide information about the roles only.

### 31. Question

You have a Linux VM that must connect to Cloud SQL. You created a service account with the appropriate access rights. You want to make sure that the VM uses this service account instead of the default Compute Engine service account. What should you do?

- A. When creating the VM via the web console, specify the service account under the 'Identity and API Access' section.
- B. Download a JSON Private Key for the service account. On the Project Metadata, add that JSON as the value for the key compute-engine-service-account.
- C. Download a JSON Private Key for the service account. On the Custom Metadata of the VM, add that JSON as the value for the key compute-engine-service-account.
- D. Download a JSON Private Key for the service account. After creating the VM, ssh into the VM and save the JSON under `~/.gcloud/compute-engine-service-account.json`.

#### Unattempted

Correct answer is A as the service account can be specified to replace the default service account when the VM is created.

Refer GCP documentation – Compute Enable Service Accounts for Instances

After creating a new service account, you can create new virtual machine instances to run as the service account.

You can enable multiple virtual machine instances to use the same service account, but a virtual machine instance can only have one service account identity. If you assign the same service account to multiple virtual machine instances, any subsequent changes you make to the service account will affect instances using the service account. This includes any changes you make to the IAM roles granted to the service account. For example, if you remove a role, all instances using the service account will lose permissions granted by that role.

You can set up a new instance to run as a service account through the Google Cloud Platform Console,

the gcloud command-line tool, or directly through the API.

In the GCP Console, go to the VM Instances page.[GO TO THE VM INSTANCES PAGE](#)

Click Create instance.

On the Create a new instance page, fill in the properties for your instance.

In the Identity and API Access section, choose the service account you want to use from the dropdown list.

Click Create to create the instance.

Options B, C & D are wrong as the approaches would not work and replace the default service account.

## 32. Question

You have an instance group that you want to load balance. You want the load balancer to terminate the client SSL session. The instance group is used to serve a public web application over HTTPS. You want to follow Google-recommended practices. What should you do?

- A. Configure an HTTP(S) load balancer.
- B. Configure an internal TCP load balancer.
- C. Configure an external SSL proxy load balancer.
- D. Configure an external TCP proxy load balancer.

### Unattempted

Correct answer is A as HTTPS load balancer supports the HTTPS traffic with the SSL termination ability.

Refer GCP documentation – Choosing Load Balancer

An HTTPS load balancer has the same basic structure as an HTTP load balancer (described above), but differs in the following ways:

An HTTPS load balancer uses a target HTTPS proxy instead of a target HTTP proxy.

An HTTPS load balancer requires at least one signed SSL certificate installed on the target HTTPS proxy for the load balancer. You can use Google-managed or self-managed SSL certificates.

The client SSL session terminates at the load balancer.

HTTPS load balancers support the QUIC transport layer protocol.

Option B is wrong as internal TCP load balancer does not serve external public traffic.

Option C is wrong as SSL proxy is not recommended for HTTPS traffic.

Google Cloud SSL Proxy Load Balancing terminates user SSL (TLS) connections at the load balancing layer, then balances the connections across your instances using the SSL or TCP protocols. Cloud SSL proxy is intended for non-HTTP(S) traffic. For HTTP(S) traffic, HTTP(S) load balancing is recommended instead.

SSL Proxy Load Balancing supports both IPv4 and IPv6 addresses for client traffic. Client IPv6 requests are terminated at the load balancing layer, then proxied over IPv4 to your backends.

Option D is wrong as TCP proxy does not support SSL offload and not recommended for HTTP/S traffic.

### 33. Question

You need to set up a policy so that videos stored in a specific Cloud Storage Regional bucket are moved to Coldline after 90 days, and then deleted after one year from their creation. How should you set up the policy?

- A. Use Cloud Storage Object Lifecycle Management using Age conditions with SetStorageClass and Delete actions. Set the SetStorageClass action to 90 days and the Delete action to 275 days (365 – 90)
- B. Use Cloud Storage Object Lifecycle Management using Age conditions with SetStorageClass and Delete actions. Set the SetStorageClass action to 90 days and the Delete action to 365 days.
- C. Use gsutil rewrite and set the Delete action to 275 days (365-90).
- D. Use gsutil rewrite and set the Delete action to 365 days.

#### Unattempted

Correct answer is B as there are 2 actions needed. First archival after 90 days, which can be done by SetStorageClass action to Coldline. Second delete the data after a year, which can be done by delete action with Age 365 days. The Age condition is measured from the object's creation time.

Refer GCP documentation – Cloud Storage Lifecycle Management

Age: This condition is satisfied when an object reaches the specified age (in days). Age is measured from the object's creation time. For example, if an object's creation time is 2019/01/10 10:00 UTC and the Age condition is 10 days, then the condition is satisfied for the object on and after 2019/01/20 10:00 UTC. This is true even if the object becomes archived through object versioning sometime after its creation.

Option A is wrong as the Age needs to be set to 365 as its relative to the object creation date and not changed date.

Options C & D are wrong as gsutil rewrite can be used to change the storage class. However, it needs to be triggered and the solution does not handle archival of data.

### 34. Question

You have 32 GB of data in a single file that you need to upload to a Nearline Storage bucket. The WAN connection you are using is rated at 1 Gbps, and you are the only one on the connection. You want to use as much of the rated 1 Gbps as possible to transfer the file rapidly. How should you upload the file?

- A. Use the GCP Console to transfer the file instead of gsutil.
- B. Enable parallel composite uploads using gsutil on the file transfer.
- C. Decrease the TCP window size on the machine initiating the transfer.
- D. ?Change the storage class of the bucket from Nearline to Multi-Regional.

**Unattempted**

Correct answer is B as the bandwidth is good and its a single file, gsutil parallel composite uploads can be used to split the large file and upload in parallel.

Refer GCP documentation – Transferring Data to GCP & Storage Composite Objects

To support parallel uploads and limited append/edit functionality, Cloud Storage allows users to compose up to 32 existing objects into a new object without transferring additional object data.

Object composition can be used for uploading an object in parallel: simply divide your data into multiple chunks, upload each chunk to a distinct object in parallel, compose your final object, and delete any temporary objects.

gsutil tool is an open-source command-line utility available for Windows, Linux, and Mac.

Multi-threaded/processed: Useful when transferring large number of files.

Parallel composite uploads: Splits large files, transfers chunks in parallel, and composes at destination.

Retry: Applies to transient network failures and HTTP/429 and 5xx error codes.

Resumability: Resumes the transfer after an error.

Option A is wrong as it is not recommended for large files and it would do a sequential upload of a single file.

Options C & D are wrong as they would not help in improving the performance.

### 35. Question

You need to monitor resources that are distributed over different projects in Google Cloud Platform. You want to consolidate reporting under the same Stackdriver Monitoring dashboard. What should you do?

- A. Use Shared VPC to connect all projects, and link Stackdriver to one of the projects.
- B. For each project, create a Stackdriver account. In each project, create a service account for that project and grant it the role of Stackdriver Account Editor in all other projects.
- C. Configure a single Stackdriver account, and link all projects to the same account.
- D. Configure a single Stackdriver account for one of the projects. In Stackdriver, create a Group and add the other project names as criteria for that Group.

**Unattempted**

Correct answer is C as you can create a single Stackdriver account and add multiple projects to the same account.

Refer GCP documentation – Stackdriver Monitoring

A single Workspace can monitor any number of GCP projects or AWS accounts. The best-practice recommendation to create a multi-project Workspace is as follows:

Create a new GCP project. For instructions on creating a new GCP project, go to Before you begin.

Create a new Workspace for that project. For detailed steps, go to Creating a single-project Workspace.

Add GCP projects or AWS accounts to the Workspace. For details, go to Adding monitored projects.

Option A is wrong as Shared VPC would not allow consolidation of multiple project monitoring. Shared VPC allows an organization to connect resources from multiple projects to a common VPC network, so that they can communicate with each other securely and efficiently using internal IPs from that network.

Option B is wrong as you do not need to create stackdriver account for each project.

Option D is wrong as it is recommended to create a separate stackdriver account instead of an account for one of the project.

### 36. Question

You are deploying an application to a Compute Engine VM in a managed instance group. The application must be running at all times, but only a single instance of the VM should run per GCP project. How should you configure the instance group?

- A. Set autoscaling to On, set the minimum number of instances to 1, and then set the maximum number of instances to 1.
- B. Set autoscaling to Off, set the minimum number of instances to 1, and then set the maximum number of instances to 1.
- C. Set autoscaling to On, set the minimum number of instances to 1, and then set the maximum number of instances to 2.
- D. Set autoscaling to Off, set the minimum number of instances to 1, and then set the maximum number of instances to 2.

#### Unattempted

Correct answer is A as you can configure Auto Scaling with minimum and maximum 1, to ensure only 1 instance is running. Auto Scaling needs be configured with an Auto Scaling policy to detect the failure and create a new instance. Ideally, you can enable Auto Healing to recover the instance, however that is not covered in any answer option.

Refer GCP documentation – Compute Engine Auto Scaler

Managed instance groups offer autoscaling capabilities that allow you to automatically add or delete instances from a managed instance group based on increases or decreases in load. Autoscaling helps your applications gracefully handle increases in traffic and reduce costs when the need for resources is lower. You define the autoscaling policy and the autoscaler performs automatic scaling based on the measured load.

Autoscaling works by adding more instances to your instance group when there is more load (upscale), and deleting instances when the need for instances is lowered (downscale).

Option C is wrong as you need only 1 instance at a time, the maximum needs to be set to 1.

Options B & D are wrong as you need to enable autoscaling.

### 37. Question

You need to allow traffic from specific virtual machines in ‘subnet-a’ network access to machines in ‘subnet-b’ without giving the entirety of subnet-a access. How can you accomplish this?

- A. Create a firewall rule to allow traffic from resources with specific network tags, then assign the machines in subnet-a the same tags.
- B. Relocate the subnet-a machines to a different subnet and give the new subnet the needed access.
- C. Create a rule to deny all traffic to the entire subnet, then create a second rule with higher priority giving access to tagged VM's in subnet-a.
- D. You can only grant firewall access to an entire subnet and not individual VM's inside.

### Unattempted

Correct answer is A as Network tags allow more granular access based on individually tagged instances.

Refer GCP documentation – VPC Network Tags

Network tags are text attributes you can add to Compute Engine virtual machine (VM) instances. Tags allow you to make firewall rules and routes applicable to specific VM instances.

You can only add network tags to VM instances or instance templates. You cannot tag other GCP resources. You can assign network tags to new instances at creation time, or you can edit the set of assigned tags at any time later. Network tags can be edited without stopping an instance.

Network tags allow you to apply firewall rules and routes to a specific instance or set of instances:

You make a firewall rule applicable to specific instances by using target tags and source tags.

You make a route applicable to specific instances by using a tag.

Option B is wrong as this would give the entire subnet access which is against the requirements: allow traffic from specific virtual machines in ‘subnet-a’ network access to machines in ‘subnet-b’ without giving the entirety of subnet-a access.

Option C is wrong as an explicit deny is not needed as implicitly all traffic is allowed.

Option D is wrong as firewall access can be granted to individual instances.

## 38. Question

You want to send and consume Cloud Pub/Sub messages from your App Engine application. The Cloud Pub/Sub API is currently disabled. You will use a service account to authenticate your application to the API. You want to make sure your application can use Cloud Pub/Sub. What should you do?

- A. Enable the Cloud Pub/Sub API in the API Library on the GCP Console.
- B. Rely on the automatic enablement of the Cloud Pub/Sub API when the Service Account accesses it.
- C. Use Deployment Manager to deploy your application. Rely on the automatic enablement of all APIs used by the application being deployed.

- D. ?Grant the App Engine Default service account the role of Cloud Pub/Sub Admin. Have your application enable the API on the first connection to Cloud Pub/Sub.

### Unattempted

Correct answer is A as the standard method is to enable services in the Google Cloud Console. You can also enable services with the Cloud SDK CLI gcloud services enable pubsub.googleapis.com

Refer GCP documentation – Cloud Pub/Sub Quick Setup

Option B is wrong as Google Cloud Services are not automatically enabled when the service account accesses it. First, service accounts do not access APIs. Service accounts are used to obtain an OAuth Access Token (or Identity Token). These tokens are used to authorize APIs. Services are not automatically enabled with an API makes first access.

Option C is wrong as Deployment Manager does not automatically enable services. You can use Deployment Manager Resource Types to enable services. You must create a virtual resource for each API that you want enabled.

Option D is wrong as Cloud Pub/Sub Admin does not have permissions to enable services. To enable services the service account (or User Account) will need roles/serviceusage.serviceUsageAdmin or another role with the permission serviceusage.services.enable.

## 39. Question

You are using Cloud Shell and need to install a custom utility for use in a few weeks. Where can you store the file so it is in the default execution path and persists across sessions?

- A. Cloud Storage
- B. /google/scripts
- C. ~/bin
- D. ?/usr/local/bin

### Unattempted

Correct answer is C as only HOME directory is persisted across sessions.

Refer GCP documentation – Cloud Shell

Cloud Shell provisions 5 GB of free persistent disk storage mounted as your \$HOME directory on the virtual machine instance. This storage is on a per-user basis and is available across projects. Unlike the instance itself, this storage does not time out on inactivity. All files you store in your home directory, including installed software, scripts and user configuration files like .bashrc and .vimrc, persist between sessions. Your \$HOME directory is private to you and cannot be accessed by other users.

Options A, B & D are wrong as they are not persistent across sessions.

## 40. Question

You have a single binary application that you want to run on Google Cloud Platform. You decided to automatically scale the application based on underlying infrastructure CPU usage. Your organizational policies require you to use virtual machines directly. You need to ensure that the application scaling is operationally efficient and completed as quickly as possible. What should you do?

- A. Create a Google Kubernetes Engine cluster, and use horizontal pod autoscaling to scale the application.
- B. Create an instance template, and use the template in a managed instance group with autoscaling configured.
- C. Create an instance template, and use the template in a managed instance group that scales up and down based on the time of day.
- D. Use a set of third-party tools to build automation around scaling the application up and down, based on Stackdriver CPU usage monitoring.

#### Unattempted

Correct answer is B as a managed instance group can help use virtual machines directly and with autoscaling can scale as per the demand.

Refer GCP documentation – Managed Instance Groups AutoScaling

Managed instance groups offer autoscaling capabilities that allow you to automatically add or delete instances from a managed instance group based on increases or decreases in load. Autoscaling helps your applications gracefully handle increases in traffic and reduce costs when the need for resources is lower. You define the autoscaling policy and the autoscaler performs automatic scaling based on the measured load.

Autoscaling works by adding more instances to your instance group when there is more load (upscaling), and deleting instances when the need for instances is lowered (downscaling).

Option A is wrong as Google Kubernetes Engine cluster can support scaling, however it would not meet the requirement of using virtual machines directly.

Option C is wrong as scaling based on time does not effectively utilize the scaling as per the demand.

Option D is wrong as using external tools is the least preferred option.

#### 41. Question

You have a project for your App Engine application that serves a development environment. The required testing has succeeded and you want to create a new project to serve as your production environment.

What should you do?

- A. Use gcloud to create the new project, and then deploy your application to the new project.
- B. Use gcloud to create the new project and to copy the deployed application to the new project.

- C. Create a Deployment Manager configuration file that copies the current App Engine deployment into a new project.
- D. Deploy your application again using gcloud and specify the project parameter with the new project name to create the new project.

### Unattempted

Correct answer is A as gcloud can be used to create a new project and the gcloud app deploy can point to the new project.

Refer GCP documentation – GCloud App Deploy

`–project=PROJECT_ID`

The Google Cloud Platform project name to use for this invocation. If omitted, then the current project is assumed; the current project can be listed using gcloud config list `–format='text(core.project)'` and can be set using gcloud config set project PROJECTID.

`–project` and its fallback core/project property play two roles in the invocation. It specifies the project of the resource to operate on. It also specifies the project for API enablement check, quota, and billing. To specify a different project for quota and billing, use `–billing-project` or `billing/quota_project` property.

Option B is wrong as the option to use gcloud app cp does not exist.

Option C is wrong as Deployment Manager does not copy the application, but allows you to specify all the resources needed for your application in a declarative format using yaml

Option D is wrong as gcloud app deploy would not create a new project. The project should be created before usage.

## 42. Question

Your company uses Cloud Storage to store application backup files for disaster recovery purposes. You want to follow Google's recommended practices. Which storage option should you use?

- A. Multi-Regional Storage
- B. Regional Storage
- C. Nearline Storage
- D. ?Coldline Storage

### Unattempted

Correct answer is D as Coldline storage is an ideal solution for disaster recovery data given its rarity of access.

Refer GCP documentation – Cloud Storage Classes

Google Cloud Storage Coldline is a very-low-cost, highly durable storage service for data archiving, online backup, and disaster recovery. Unlike other “cold” storage services, your data is available within milliseconds, not hours or days.

Coldline Storage is the best choice for data that you plan to access at most once a year, due to its slightly lower availability, 90-day minimum storage duration, costs for data access, and higher per-operation costs. For example:

Cold Data Storage – Infrequently accessed data, such as data stored for legal or regulatory reasons, can be stored at low cost as Coldline Storage, and be available when you need it.

Disaster recovery – In the event of a disaster recovery event, recovery time is key. Cloud Storage provides low latency access to data stored as Coldline Storage.

The geo-redundancy of Coldline Storage data is determined by the type of location in which it is stored: Coldline Storage data stored in multi-regional locations is redundant across multiple regions, providing higher availability than Coldline Storage data stored in regional locations.

Options A, B & C are wrong as they are not suited for infrequently accessed data, as disaster does not happen periodically but rarely.

### 43. Question

You've deployed a microservice called myapp1 to a Google Kubernetes Engine cluster using the YAML file specified below:

```
apiVersion: apps/v1
kind: Deployment
metadata:
 name: myappl-deployment
spec:
 selector:
 matchLabels:
 app: myappl
 replicas: 2
 template:
 metadata:
 labels:
 app: myappl
 spec:
 containers:
 name: main-container
 image: gcr.io/my-company-repo/myapp1:1.4
 env:
 name: DS_PASSWORD
 value: "tOugh2guess!"
 ports:
 - containerPort: 8080
```

You need to refactor this configuration so that the database password is not stored in plain text. You want to follow Google-recommended practices. What should you do?

- A. Store the database password inside the Docker image of the container, not in the YAML file.
- B. Store the database password inside a Secret object. Modify the YAML file to populate the DB\_PASSWORD environment variable from the Secret.
- C. Store the database password inside a ConfigMap object. Modify the YAML file to populate the DB\_PASSWORD environment variable from the ConfigMap.
- D. ?Store the database password in a file inside a Kubernetes persistent volume, and use a persistent volume claim to mount the volume to the container.

#### Unattempted

Correct answer is B as Google Kubernetes Engine supports secret to store sensitive data such as database passwords and is a google recommended practice.

Refer GCP documentation – Kubernetes Engine Secret

Secrets are secure objects which store sensitive data, such as passwords, OAuth tokens, and SSH keys, in your clusters. Storing sensitive data in Secrets is more secure than plaintext ConfigMaps or in Pod specifications. Using Secrets gives you control over how sensitive data is used, and reduces the risk of exposing the data to unauthorized users.

Options A, C & D are wrong as others options are not secured and not recommended as best practice.

#### 44. Question

You created an instance of SQL Server 2017 on Compute Engine to test features in the new version. You want to connect to this instance using the fewest number of steps. What should you do?

- A. Install a RDP client on your desktop. Verify that a firewall rule for port 3389 exists.
- B. Install a RDP client in your desktop. Set a Windows username and password in the GCP Console. Use the credentials to log in to the instance.
- C. Set a Windows password in the GCP Console. Verify that a firewall rule for port 22 exists. Click the RDP button in the GCP Console and supply the credentials to log in.
- D. ?Set a Windows username and password in the GCP Console. Verify that a firewall rule for port 3389 exists. Click the RDP button in the GCP Console, and supply the credentials to log in.

#### Unattempted

Correct answer is A as it mentions fewest number of steps to connect to the instance. You can download the RDP Client and verify 3389 firewall is open. If the RDP asks for username and password, the instance is working.

Option B is wrong as it fails to mention the key requirement of port 3389 be opened.

Option C is wrong as RDP requires port 3389 to be opened.

Option D is wrong as you need an RDP client.

## 45. Question

You are building a pipeline to process time-series data. Which Google Cloud Platform services should you put in boxes 1,2,3, and 4?

- A. Cloud Pub/Sub, Cloud Dataflow, Cloud Datastore, BigQuery
- B. Firebase Messages, Cloud Pub/Sub, Cloud Spanner, BigQuery
- C. Cloud Pub/Sub, Cloud Storage, BigQuery, Cloud Bigtable
- D. Cloud Pub/Sub, Cloud Dataflow, Cloud Bigtable, BigQuery

### Unattempted

Correct answer is D as Cloud Pub/Sub for data ingestion, Dataflow for data handling and transformation, Bigtable for storage to provide low latency data access and BigQuery for analytics.

Refer GCP documentation – Time Series Dataflow

Cloud Pub/Sub. As well as performing ingestion, Cloud Pub/Sub can also act as the glue between the loosely coupled systems. You can send the processed data to other systems to consume; for example, you might send all correlations with more than the value of ABS(0.2) to other systems.

BigQuery. Place any data that you want to process or access later using a SQL interface into BigQuery.

Cloud Bigtable. Place any data that you want to use for low-latency storage, or where you might want to get at a very small subset of a larger dataset quickly (key lookups as well as range scans), in Cloud Bigtable.

Option A is wrong as Datastore is not an ideal solution to store large time series data.

Option B is wrong as Cloud Spanner is not an ideal solution for storage.

Option C is wrong as Cloud Storage is for storage and doesn't help handle and source data to storage and analytics.

## 46. Question

You are deploying an application to App Engine. You want the number of instances to scale based on request rate. You need at least 3 unoccupied instances at all times. Which scaling type should you use?

- A. Manual Scaling with 3 instances.
- B. Basic Scaling with min\_instances set to 3.
- C. Basic Scaling with max\_instances set to 3.
- D. Automatic Scaling with min\_idle\_instances set to 3.

### Unattempted

Correct answer is D as min\_idle\_instances property can be set to have minimum idle instances which would be always running.

Refer GCP documentation – App Engine Scaling & app.yaml Reference

Auto scaling services use dynamic instances that get created based on request rate, response latencies, and other application metrics. However, if you specify a number of minimum idle instances, that specified number of instances run as resident instances while any additional instances are dynamic.

#### min\_idle\_instances

The number of instances to be kept running and ready to serve traffic. Note that you are charged for the number of instances specified whether they are receiving traffic or not. This setting only applies to the version that receives most of the traffic. Keep the following in mind:

A low minimum helps keep your running costs down during idle periods, but means that fewer instances might be immediately available to respond to a sudden load spike.

A high minimum allows you to prime the application for rapid spikes in request load. App Engine keeps the minimum number of instances running to serve incoming requests. You are charged for the number of instances specified, whether or not they are handling requests. For this feature to function properly, you must make sure that warmup requests are enabled and that your application handles warmup requests.

If you set a minimum number of idle instances, pending latency will have less effect on your application's performance. Because App Engine keeps idle instances in reserve, it is unlikely that requests will enter the pending queue except in exceptionally high load spikes. You will need to test your application and expected traffic volume to determine the ideal number of instances to keep in reserve.

Option A is wrong as manual scaling would not provide the scaling based on the request rate and would need manual intervention.

Options B & C are wrong as basic scaling will not allow the scaling based on the request rate.

## 47. Question

You significantly changed a complex Deployment Manager template and want to confirm that the dependencies of all defined resources are properly met before committing it to the project. You want the most rapid feedback on your changes. What should you do?

- A. Use granular logging statements within a Deployment Manager template authored in Python.
- B. Monitor activity of the Deployment Manager execution on the Stackdriver Logging page of the GCP Console.
- C. Execute the Deployment Manager template against a separate project with the same configuration, and monitor for failures.
- D. Execute the Deployment Manager template using the --preview option in the same project, and observe the state of interdependent resources.

**Unattempted**

Correct answer is D as Deployment Manager provides the preview feature to check on what resources would be created.

Refer GCP documentation – Deployment Manager Preview

After you have written a configuration file, you can preview the configuration before you create a deployment. Previewing a configuration lets you see the resources that Deployment Manager would create but does not actually instantiate any actual resources. The Deployment Manager service previews the configuration by:

Expanding the full configuration, including any templates.

Creating a deployment and “shell” resources.

You can preview your configuration by using the preview query parameter when making an insert() request.

```
gcloud deployment-manager deployments create example-deployment \
--config configuration-file.yaml --preview
```

**48. Question**

You have a development project with appropriate IAM roles defined. You are creating a production project and want to have the same IAM roles on the new project, using the fewest possible steps. What should you do?

- A. Use gcloud iam roles copy and specify the production project as the destination project.
- B. Use gcloud iam roles copy and specify your organization as the destination organization.
- C. In the Google Cloud Platform Console, use the ‘create role from role’ functionality.
- D. ?In the Google Cloud Platform Console, use the ‘create role’ functionality and select all applicable permissions.

**Unattempted**

Correct answer is A as Cloud SDK gcloud iam roles copy can be used to copy the roles to different organization or project.

Refer GCP documentation – Cloud SDK IAM Copy Role

gcloud iam roles copy – create a role from an existing role

–dest-organization=DEST\_ORGANIZATION (The organization of the destination role)

–dest-project=DEST\_PROJECT (The project of the destination role)

Option B is wrong as the destination new project needs to be specified instead of the organization.

Option C is wrong as creating roles through GCP Console is cumbersome, time consuming and error prone.

Option D is wrong as it does not replicate the IAM roles permission.

## 49. Question

You have one GCP account running in your default region and zone and another account running in a non-default region and zone. You want to start a new Compute Engine instance in these two Google Cloud Platform accounts using the command line interface. What should you do?

- A. Create two configurations using gcloud config configurations create [NAME]. Run gcloud config configurations activate [NAME] to switch between accounts when running the commands to start the Compute Engine instances.
- B. Create two configurations using gcloud config configurations create [NAME]. Run gcloud config configurations list to start the Compute Engine instances.
- C. Activate two configurations using gcloud config configurations activate [NAME] . Run gcloud config configurations list to start the Compute Engine instances.
- D. Activate two configurations using gcloud config configurations activate [NAME] . Run gcloud config configurations list to start the Compute Engine instances.

### Unattempted

Correct answer is A as you can create different configurations for each account and create compute instances in each account by activating the respective account.

Refer GCP documentation – Configurations Create & Activate

Options B, C & D are wrong as gcloud config configurations list does not help create instances. It would only lists existing named configurations.

## 50. Question

You recently deployed a new version of an application to App Engine and then discovered a bug in the release. You need to immediately revert to the prior version of the application. What should you do?

- A. Run gcloud app restore.
- B. On the App Engine page of the GCP Console, select the application that needs to be reverted and click Revert.
- C. On the App Engine Versions page of the GCP Console, route 100% of the traffic to the previous version.
- D. Deploy the original version as a separate application. Then go to App Engine settings and split traffic between applications so that the original version serves 100% of the requests.

### Unattempted

Correct answer is C as you can migrate all the traffic back to the previous version.

Refer GCP documentation – App Engine Overview

Having multiple versions of your app within each service allows you to quickly switch between different versions of that app for rollbacks, testing, or other temporary events. You can route traffic to one or more specific versions of your app by migrating or splitting traffic.

Option A is wrong as gcloud app restore was used for backup and restore and has been deprecated.

Option B is wrong as there is no application revert functionality available.

Option D is wrong as App Engine maintains version and need not be redeployed.

## 51. Question

You need to select and configure compute resources for a set of batch processing jobs. These jobs take around 2 hours to complete and are run nightly. You want to minimize service costs. What should you do?

- A. Select Google Kubernetes Engine. Use a single-node cluster with a small instance type.
- B. Select Google Kubernetes Engine. Use a three-node cluster with micro instance type.
- C. Select Compute Engine. Use preemptible VM instances of the appropriate standard machine type.
- D. Select Compute Engine. Use VM instance types that support micro bursting.

### Unattempted

Correct answer is C as Compute Engine preemptible VMs are ideal for batch processing jobs and are able to run at a much lower price than standard instances.

Refer GCP documentation – Compute Engine Preemptible

A preemptible VM is an instance that you can create and run at a much lower price than normal instances. However, Compute Engine might terminate (preempt) these instances if it requires access to those resources for other tasks. Preemptible instances are excess Compute Engine capacity, so their availability varies with usage.

If your apps are fault-tolerant and can withstand possible instance preemptions, then preemptible instances can reduce your Compute Engine costs significantly. For example, batch processing jobs can run on preemptible instances. If some of those instances terminate during processing, the job slows but does not completely stop. Preemptible instances complete your batch processing tasks without placing additional workload on your existing instances and without requiring you to pay full price for additional normal instances.

Options A, B & D are wrong as they would require Compute Engine instances running, which is not a cost effective option for batch processing jobs.

## 52. Question

You are analyzing Google Cloud Platform service costs from three separate projects. You want to use this information to create service cost estimates by service type, daily and monthly, for the next six months using standard query syntax. What should you do?

- A. Export your bill to a Cloud Storage bucket and then import into Cloud Bigtable for analysis.
- B. Export your bill to a Cloud Storage bucket, and then import into Google Sheets for analysis.
- C. Export your transactions to a local file and perform analysis with a desktop tool.
- D. Export your bill to a BigQuery dataset, and then write time window-based SQL queries for analysis.

#### Unattempted

Correct answer is D as BigQuery provides an ideal storage option to store and query in standard SQL dialect.

BigQuery, Google's serverless, highly scalable enterprise data warehouse, is designed to make data analysts more productive with unmatched price-performance. Because there is no infrastructure to manage, you can focus on uncovering meaningful insights using familiar SQL without the need for a database administrator.

Option A is wrong Bigtable is a NoSQL solution and does not support SQL dialect.

Cloud Bigtable is Google's sparsely populated NoSQL database which can scale to billions of rows, thousands of columns, and petabytes of data. Cloud Bigtable has a data model similar to Apache HBase and provides an HBase-compatible client library.

Options B & C are wrong as Google Sheets and local file does not provide standard query syntax querying.

### 53. Question

You need a dynamic way of provisioning VMs on Compute Engine. The exact specifications will be in a dedicated configuration file. You want to follow Google's recommended practices. Which method should you use?

- A. Deployment Manager
- B. Cloud Composer
- C. Managed Instance Group
- D. Unmanaged Instance Group

#### Unattempted

Correct answer is A as Deployment Manager provide Infrastructure as a Code capability.

Refer GCP documentation – Deployment Manager

Google Cloud Deployment Manager allows you to specify all the resources needed for your application in a declarative format using yaml. You can also use Python or Jinja2 templates to parameterize the configuration and allow reuse of common deployment paradigms such as a load balanced, auto-scaled instance group. Treat your configuration as code and perform repeatable deployments.

Option B is wrong as Cloud Composer is a fully managed workflow orchestration service that empowers you to author, schedule, and monitor pipelines that span across clouds and on-premises data centers.

Options C & D are wrong as An instance group is a collection of VM instances that you can manage as a single entity.

Managed instance groups (MIGs) allow you to operate applications on multiple identical VMs. You can make your workloads scalable and highly available by taking advantage of automated MIG services, including: autoscaling, autohealing, regional (multi-zone) deployment, and auto-updating.

Unmanaged instance groups allow you to load balance across a fleet of VMs that you manage yourself.

#### 54. Question

Your team is developing a product catalog that allows end users to search and filter. The full catalog of products consists of about 500 products. The team doesn't have any experience with SQL, or schema migrations, so they're considering a NoSQL option. Which database service would work best?

- A. Cloud SQL
- B. Cloud Memorystore
- C. Bigtable
- D. Cloud Datastore

#### Unattempted

Correct answer is D as Cloud Datastore would provide the NoSQL option for storing the product catalog.

As the data is limited, it would be a good fit.

Option A is wrong as Cloud SQL is a relational SQL solution.

Option B is wrong as Cloud Memorystore for Redis provides a fully managed in-memory data store service built on scalable, secure, and highly available infrastructure managed by Google. Use Cloud Memorystore to build application caches that provides sub-millisecond data access. Cloud Memorystore is compatible with the Redis protocol, allowing easy migration with zero code changes.

Option C is wrong as although Bigtable provides a NoSQL solution, it is a petabyte-scale, fully managed NoSQL database service ideal for large analytical and operational workloads.

#### 55. Question

You're trying to provide temporary access to some files in a Cloud Storage bucket. You want to limit the time that the files are available to 10 minutes. With the fewest steps possible, what is the best way to generate a signed URL?

- A. Create a service account and JSON key. Use the gsutil signurl -t 10m command and pass in the JSON key and bucket.

- B. Create a service account and JSON key. Use the gsutil signurl -d 10m command and pass in the JSON key and bucket.
- C. Create a service account and JSON key. Use the gsutil signurl -p 10m command and pass in the JSON key and bucket.
- D. Create a service account and JSON key. Use the gsutil signurl -m 10m command and pass in the JSON key and bucket.

### Unattempted

Correct answer is B as signurl command will generate a signed URL that embeds authentication data so the URL can be used by someone who does not have a Google account. -d can help provide the time duration.

Refer GCP documentation – Cloud Storage gsutil signurl

gsutil signurl [-c] [-d] [-m] \

[-p] [-r] keystore-file url...

-m Specifies the HTTP method to be authorized for use with the signed url, default is GET. You may also specify RESUMABLE to create a signed resumable upload start URL. When using a signed URL to start a resumable upload session, you will need to specify the ‘x-goog-resumable:start’ header in the request or else signature validation will fail.

-d Specifies the duration that the signed url should be valid for, default duration is 1 hour. Times may be specified with no suffix (default hours), or with s = seconds, m = minutes, h = hours, d = days. This option may be specified multiple times, in which case the duration the link remains valid is the sum of all the duration options. The max duration allowed is 7d.

-c Specifies the content type for which the signed url is valid for.-p Specify the keystore password instead of prompting.

-r Specifies the region in which the resources for which you are creating signed URLs are stored. Default value is ‘auto’ which will cause gsutil to fetch the region for the resource. When auto-detecting the region, the current gsutil user’s credentials, not the credentials from the private-key-file, are used to fetch the bucket’s metadata. This option must be specified and not ‘auto’ when generating a signed URL to create a bucket.

## 56. Question

You’re about to deploy your team’s App Engine application. They’re using the Go runtime with a Standard Environment. Which command should you use to deploy the application?

- A. gcloud app deploy app.yaml
- B. gcloud app-engine apply app.yaml
- C. gcloud app apply app.yaml

- D. ?gcloud app-engine deploy app.yaml

#### Unattempted

Correct answer is A as gcloud app deploy provides an ability to deploy the local code and/or configuration of your app to App Engine.

Refer GCP documentation – gcloud app deploy

This command is used to deploy both code and configuration to the App Engine server. As an input it takes one or more DEPLOYABLES that should be uploaded. A DEPLOYABLE can be a service's .yaml file or a configuration's .yaml file.

Option C is wrong as gcloud app apply is not a valid command.

Options B & D are wrong as gcloud app-engine is not a valid command.

### 57. Question

You have a Windows server running on a custom network. There's an allow firewall rule with an IP filter of 0.0.0.0/0 with a protocol/port of tcp:3389. The logs on the instance show a constant stream of attempts from different IP addresses, trying to connect via RDP. You suspect this is a brute force attack. How might you change the firewall rule to stop this from happening and still enable access for legit users?

- A. Stop the instance.
- B. Deny all traffic to port 3389.
- C. Change the port that RDP is running on in the instance and change the port number in the firewall rule.
- D. Change the IP address range in the filter to only allow known IP addresses.

#### Unattempted

Correct answer is D as by using 0.0.0.0/0, you're opening the port to the internet. By whitelisting known IP addresses, it will block anyone not on the list.

Option A is wrong as it is not a viable solution for protect the instances.

Option B is wrong denying all traffic would block all.

Option C is wrong as it is not possible to change the default RDP port.

### 58. Question

You've found that your Linux server keeps running low on memory. It's currently using 8GB of memory, and it needs to be increased to 16. What is the simplest way to do that?

- A. Use the gcloud compute add-memory command to increase the memory.
- B. Use the Linux memincr command to increase the memory.

- C. Stop the instance and change the machine type.
- D. Create a new instance with the correct amount of memory.

#### Unattempted

Correct answer is C as you can increase the memory by changing the instance machine type.

Refer GCP documentation – Changing Machine Type

You can change the machine type of a stopped instance if it is not part of a managed instance group. If you need to change the machine type of instances within a managed instance group, read Updating managed instance groups.

Change the machine types of your instances if your existing machine type is not a good fit for the workloads you run on that instance. You can change the machine type of an instance to adjust the number of vCPUs and memory as your workload changes. For example, you can start an instance with a smaller machine during setup, development, and testing and change the instance to use a larger machine type when you are ready for production workloads.

Options A & B are wrong as the options are invalid.

Option D is wrong as the solution is valid, but it is not the simplest.

## 59. Question

You're working on setting up a cluster of virtual machines with GPUs to perform some 3D rendering for a customer. They're on a limited budget and are looking for ways to save money. What is the best solution for implementing this?

- A. Use an autoscaled managed instance group containing some preemptible instances.
- B. Use an unmanaged instance group with preemptible instances.
- C. Use App Engine with Flexible Environments.
- D. Use App Engine with Standard Environments.

#### Unattempted

Correct answer is A as Preemptible with managed instance groups would help add GPUs at a lower cost.

Refer GCP documentation – Compute Engine Preemptible

A preemptible VM is an instance that you can create and run at a much lower price than normal instances. However, Compute Engine might terminate (preempt) these instances if it requires access to those resources for other tasks. Preemptible instances are excess Compute Engine capacity, so their availability varies with usage.

If your apps are fault-tolerant and can withstand possible instance preemptions, then preemptible instances can reduce your Compute Engine costs significantly. For example, batch processing jobs can run on preemptible instances. If some of those instances terminate during processing, the job slows but does not completely stop. Preemptible instances complete your batch processing tasks without placing

additional workload on your existing instances and without requiring you to pay full price for additional normal instances.

You can add GPUs to your preemptible VM instances at lower preemptible prices for the GPUs. GPUs attached to preemptible instances work like normal GPUs but persist only for the life of the instance.

Preemptible instances with GPUs follow the same preemption process as all preemptible instances.

Option B is wrong as unmanaged instance group does not provide scaling.

Options C & D are wrong as GCP currently does not support GPUs for App Engine.

## 60. Question

Your coworker has helped you set up several configurations for gcloud. You've noticed that you're running commands against the wrong project. Being new to the company, you haven't yet memorized any of the projects. With the fewest steps possible, what's the fastest way to switch to the correct configuration?

- A. Run gcloud configurations list followed by gcloud configurations activate.
- B. Run gcloud config list followed by gcloud config activate.
- C. Run gcloud config configurations list followed by gcloud config configurations activate.
- D. Re-authenticate with the gcloud auth login command and select the correct configurations on login.

### Unattempted

Correct answer is C as gcloud config configurations list can help check for the existing configurations and activate can help switch to the configuration.

Refer GCP documentation – Cloud SDK gcloud config

gcloud config configurations list – lists existing named configurations

gcloud config configurations activate – activates an existing named configuration

Options A & B are wrong as they are invalid commands.

Option D is wrong as does not help to identify and activate configurations.

gcloud auth login – authorize gcloud to access the Cloud Platform with Google user credentials

Obtains access credentials for your user account via a web-based authorization flow. When this command completes successfully, it sets the active account in the current configuration to the account specified. If no configuration exists, it creates a configuration named default.

## 61. Question

You have an App Engine application running in us-east1. You've noticed 90% of your traffic comes from the West Coast. You'd like to change the region. What's the best way to change the App Engine region?

- A. Use the gcloud app region set command and supply the name of the new region.

- B. Contact Google Cloud Support and request the change.
- C. From the console, under the App Engine page, click edit, and change the region drop-down.
- D. ?Create a new project and create an App Engine instance in us-west2.

### Unattempted

Correct answer is D as app engine is a regional resource, it needs to be redeployed to the different region.

Refer GCP documentation – App Engine locations

App Engine is regional, which means the infrastructure that runs your apps is located in a specific region and is managed by Google to be redundantly available across all the zones within that region.

Meeting your latency, availability, or durability requirements are primary factors for selecting the region where your apps are run. You can generally select the region nearest to your app's users but you should consider the location of the other GCP products and services that are used by your app. Using services across multiple locations can affect your app's latency as well as pricing

You cannot change an app's region after you set it.

Options A, B & C are wrong as once the region is set for the app engine it cannot be modified.

## 62. Question

You're using Deployment Manager to deploy your application to an autoscaled, managed instance group on Compute Engine. The application is a single binary. What is the fastest way to get the binary onto the instance, without introducing undue complexity?

- A. When creating the instance template use the startup-script metadata key to bootstrap the application.
- B. When creating the instance template use the initialize-script metadata key to bootstrap the application.
- C. When creating the instance template, use the startup script metadata key to install Ansible. Have the instance run the play-book at startup to install the application.
- D. ?Once the instance starts up, connect over SSH and install the application.

### Unattempted

Correct answer is A as Instance Template can be specified startup-script to install/download the binary artifact.

Refer GCP documentation – Deployment Manager Startup Scripts

When you are deploying more complex configurations, you might have tens, hundreds, or even thousands of virtual machine instances. If you're familiar with Compute Engine, it's likely that you want to use startup scripts to help install or configure your instances automatically.

Using Deployment Manager, you can run the same startup scripts or add metadata to virtual machine instances in your deployment by specifying the metadata in your template or configuration.

To add metadata or startup scripts to your template, add the metadata property and the relevant metadata keys and values. For example, for specifying a startup script, the metadata key must be startup-script and the value would be the contents of your startup script.

Option B is wrong as initialize-script is not a valid option.

Option C is wrong as although the solution is valid, it introduces complexity.

Option D is wrong as it is cumbersome to do it for a autoscaled managed instance group.

### 63. Question

You've created a Pod using the kubectl run command. Now you're attempting to remove the Pod, and it keeps being recreated. Which command might help you as you attempt to remove the pod?

- A. gcloud container describe pods
- B. kubectl get pods
- C. kubectl get secrets
- D. kubectl get deployments

#### Unattempted

Correct answer is D as Pods would be recreated and you need to remove the deployment to remove the associated pods. kubectl get deployments would help get the list of deployments

Refer GCP documentation – Kubernetes Deployment

Deployments represent a set of multiple, identical Pods with no unique identities. A Deployment runs multiple replicas of your application and automatically replaces any instances that fail or become unresponsive. In this way, Deployments help ensure that one or more instances of your application are available to serve user requests. Deployments are managed by the Kubernetes Deployment controller.

Option A is wrong as it is not a valid command.

Option B is wrong as it would only provide the information for the pods.

Option C is wrong as it would provide information about the secrets.

### 64. Question

You're attempting to remove the zone property from the Compute Engine service, that was set with the incorrect value. Which command would accomplish your task?

- A. gcloud config unset compute/zone
- B. gcloud config unset zone
- C. gcloud config configurations unset compute/zone

- D. ?gcloud unset compute/zone

#### Unattempted

Correct answer is A as Zone can be corrected using the gcloud config unset compute/zone

Refer GCP documentation – gcloud config unset

To unset the zone property in the compute section, run:

gcloud config unset compute/zone

### 65. Question

You've seen some errors in the logs for a specific Deployment. You've narrowed the issue down to the Pod named "ad-generator" that is throwing the errors. Your engineers aren't able to reproduce the error in any other environment. They've told you that if they could just "connect into the container" for 5 minutes, they could figure out the root cause. What steps would allow them to run commands against the container?

- A. Use the kubectl exec -it ad-generator -- /bin/ bash command to run a shell on that container.
- B. Use the kubectl exec -it -- /bin/ bash command to run a shell on that container.
- C. Use the kubectl run command to run a shell on that container.
- D. Use the kubectl run ad-generator /bin/ bash command to run a shell on that container.

#### Unattempted

Correct answer is A as kubectl exec can help open a shell on the pod in an interactive mode.

Refer GCP documentation – Kubernetes Engine Troubleshooting

Connect to a running container

Open a shell to the Pod:

kubectl exec -it [POD\_NAME] — /bin/ bash

If there is more than one container in your Pod, add -c [CONTAINER\_NAME].

Now, you can run bash commands from the container: you can test the network or check if you have access to files or databases used by your application.

### 66. Question

Your team has been working towards using desired state configuration for your entire infrastructure, which is why they're excited to store the Kubernetes Deployments in YAML. You created a Kubernetes Deployment with the kubectl apply command and passed on a YAML file. You need to edit the number of replicas. What steps should you take to update the Deployment?

- A. Edit the number of replicas in the YAML file and rerun the kubectl apply.

- B. Edit the YAML and push it to Github so that the git triggers deploy the change.
- C. Disregard the YAML file. Use the kubectl scale command.
- D. Edit the number of replicas in the YAML file and run the kubectl set image command

### Unattempted

Correct answer is A as to set the desired state, the replicas of needs to be updated in the configuration file and changes applied.

Refer GCP documentation – Kubernetes Scaling Apps

Kubernetes uses the Deployment controller to deploy stateless applications as uniform, non-unique Pods. Deployments manage the desired state of your application: how many Pods should run your application, what version of the container image should run, what the Pods should be labelled, and so on. The desired state can be changed dynamically through updates to the Deployment's Pod specification.

You can use kubectl apply to apply a new configuration file to an existing controller object. kubectl apply is useful for making multiple changes to a resource, and may be useful for users who prefer to manage their resources in configuration files.

To scale using kubectl apply, the configuration file you supply should include a new number of replicas in the replicas field of the object's specification.

Options B & D are wrong they are not valid options to update the desired state.

Option C is wrong as kubectl scale disregards the configuration files, which is the key requirement.

kubectl scale lets your instantaneously change the number of replicas you want to run your application.

## 67. Question

Your developers have some application metrics that they're tracking. They'd like to be able to create alerts based on these metrics. What steps need to happen in order to alert based on these metrics?

- A. In the UI create a new logging metric with the required filters, edit the application code to set the metric value when needed, and create an alert in Stackdriver based on the new metric.
- B. Create a custom monitoring metric in code, edit the application code to set the metric value when needed, create an alert in Stackdriver based on the new metric.
- C. Add the Stackdriver monitoring and logging agent to the instances running the code.
- D. Create a custom monitoring metric in code, in the UI create a matching logging metric, and create an alert in Stackdriver based on the new metric.

### Unattempted

Correct answer is B as Stackdriver allows custom metrics, which can be used to create alerts.

Refer GCP documentation – Stackdriver Monitoring Custom Metrics

Custom metrics are metrics defined by users. Custom metrics use the same elements that the built-in

Stackdriver Monitoring metrics use:

A set of data points.

Metric-type information, which tells you what the data points represent.

Monitored-resource information, which tells you where the data points originated.

To use a custom metric, you must have a metric descriptor for your new metric type. Stackdriver Monitoring can create the metric descriptor for you automatically, or you can use the `metricDescriptors.create` API method to create it yourself.

To have Stackdriver Monitoring create the metric descriptor for you, you simply write time series data for your metric, and Stackdriver Monitoring creates a descriptor based on the data you are writing. There are limits to auto-creation, so it's helpful to know what information goes into a metric definition.

After you have a new custom metric descriptor, whether you or Monitoring created it, you can use the metric descriptor with the metric descriptor API methods and the time series API methods.

You can also create charts and alerts for your custom metric data.

Options A & D are wrong as you need to create monitoring metric and not logging metric

Option C is wrong as Stackdriver agent, by default, would not track custom metrics.

## 68. Question

Your developers have created an application that needs to be able to make calls to Cloud Storage and BigQuery. The code is going to run inside a container and will run on Kubernetes Engine and on-premises. What's the best way for them to authenticate to the Google Cloud services?

- A. Create a service account, grant it the least viable privileges to the required services, generate and download a key. Use the key to authenticate inside the application.
- B. Use the default service account for App Engine, which already has the required permissions.
- C. Use the default service account for Compute Engine, which already has the required permissions.
- D. Create a service account, with editor permissions, generate and download a key. Use the key to authenticate inside the application.

### Unattempted

Correct answer is A as Service accounts can be used by the application to authenticate and call the service APIs securely.

Refer GCP documentation – IAM Service Account

To use a service account outside of GCP, such as on other platforms or on-premises, you must first establish the identity of the service account. Public/private key pairs provide a secure way of accomplishing this goal.

When you create a key, your new public/private key pair is generated and downloaded to your machine; it serves as the only copy of the private key. You are responsible for storing the private key securely. Take note of its location and ensure the key is accessible to your application; it needs the key to make

authenticated API calls.

Options B & C are wrong as default service account does not provide the requirement permissions and would not be available for application deployed on on-premises.  
Option D is wrong as although the solution would work, however, it violates the principle of least privilege. Also, it would still require a service account key for the on-premises code.

## 69. Question

You need to connect to one of your Compute Engine instances using SSH. You've already authenticated gcloud, however, you don't have an SSH key deployed yet. In the fewest steps possible, what's the easiest way to connect to the app?

- A. Create a key with the ssh-keygen command. Upload the key to the instance. Run gcloud compute instances list to get the IP address of the instance, then use the ssh command.
- B. Use the gcloud compute ssh command.
- C. Create a key with the ssh-keygen command. Then use the gcloud compute ssh command.
- D. Run gcloud compute instances list to get the IP address of the instance, then use the ssh command.

### Unattempted

Correct answer is B as using gcloud compute ssh is the easiest and quickest way to use SSH. It would generate the keys and add to the project metadata to enable login.

Refer GCP documentation – gcloud compute ssh

gcloud compute ssh is a thin wrapper around the ssh(1) command that takes care of authentication and the translation of the instance name into an IP address.

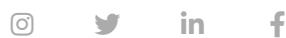
gcloud compute ssh ensures that the user's public SSH key is present in the project's metadata. If the user does not have a public SSH key, one is generated using ssh-keygen(1) (if the –quiet flag is given, the generated key will have an empty passphrase).

**Use Page numbers below to navigate to other  
practice tests**

Pages:

[← Previous Post](#)[Next Post →](#)

## Skillcertpro



### Quick Links

[ABOUT US](#)[FAQ](#)[BROWSE ALL PRACTICE TESTS](#)[CONTACT FORM](#)

### Important Links

[REFUND POLICY](#)[REFUND REQUEST](#)[TERMS & CONDITIONS](#)[PRIVACY POLICY](#)[Privacy Policy](#)

SALE IS ON  | 12 HOURS LEFT | BUY 2 & GET ADDITIONAL 25% OFF | Use Coupon - BLACKFRIDAY



# SKILLCERTPRO

IT CERTIFICATION TRAININGS



Google Cloud / By SkillCertPro

## Practice Set 7

Your results are here!! for " Google Certified Associate Cloud Engineer Practice Test 7 "

0 of 65 questions answered correctly

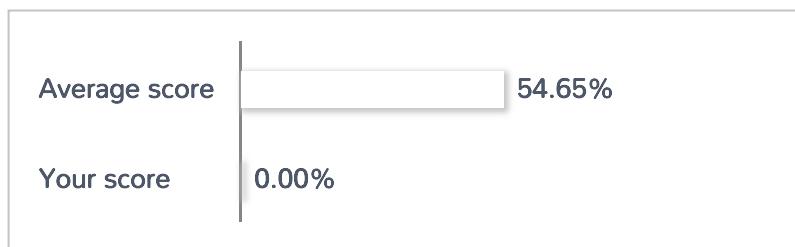
Your time: 00:00:15

Your Final Score is : 0

You have attempted : 0

Number of Correct Questions : 0 and scored 0

Number of Incorrect Questions : 0 and Negative marks 0



You can review your answers by clicking view questions.

**Important Note :** Open Reference Documentation Links in New Tab (Right Click and Open in New Tab).

[Restart Test](#)

[View Answers](#)

|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 |

|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 |
| 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 | 64 | 65 |    |    |    |

Correct Incorrect

Review Question

Summary

## 1. Question

Your projects incurred more costs than you expected last month. Your research reveals that a development GKE container emitted a huge number of logs, which resulted in higher costs. You want to disable the logs quickly using the minimum number of steps. What should you do?

- 1. Go to the GKE console, and delete existing clusters. 2. Recreate a new cluster. 3. Clear the option to enable legacy Stackdriver Monitoring.
- Go to the Logs ingestion window in Stackdriver Logging, and disable the log source for the GKE container resource.
- Go to the Logs ingestion window in Stackdriver Logging, and disable the log source for the GKE Cluster Operations resource.
- 1. Go to the GKE console, and delete existing clusters. 2. Recreate a new cluster. 3. Clear the option to enable legacy Stackdriver Logging.

### Unattempted

1. Go to the GKE console, and delete existing clusters.

2. Recreate a new cluster.

3. Clear the option to enable legacy Stackdriver Logging. is not right.

Our requirement is to disable the logs ingested from the GKE container. We don't need to delete the existing cluster and create a new one.

Go to the Logs ingestion window in Stackdriver Logging, and disable the log source for the GKE container resource. is the right answer.

We want to disable logs from a specific GKE container and this is the only option that does that.

More information about logs exclusions: <https://cloud.google.com/logging/docs/exclusions>

## 2. Question

Your team is working towards using the desired state configuration for your application deployed on the GKE cluster. You have YAML files for the Kubernetes Deployment and Service objects. Your application is

designed to have 2 pods, which is defined by the replicas parameter in app-deployment.yaml. Your service uses GKE Load Balancer which is defined in app-service.yaml

You created the Kubernetes resources by running

```
kubectl apply -f app-deployment.yaml
```

```
kubectl apply -f app-service.yaml
```

Your deployment is now serving live traffic but is suffering from performance issues. You want to increase the number of replicas to 5. What should you do in order to update the replicas in existing Kubernetes deployment objects?

- Disregard the YAML file. Use the kubectl scale command to scale the replicas to 5. `kubectl scale --replicas=5 -f app-deployment.yaml`
- Disregard the YAML file. Enable autoscaling on the deployment to trigger on CPU usage and set max pods to 5. `kubectl autoscale myapp --max=5 --cpu-percent=80`
- Modify the current configuration of the deployment by using kubectl edit to open the YAML file of the current configuration, modify and save the configuration. `kubectl edit deployment/app-deployment -o yaml --save-config`
- Edit the number of replicas in the YAML file and rerun the kubectl apply. `kubectl apply -f app-deployment.yaml`

#### Unattempted

Disregard the YAML file. Use the kubectl scale command to scale the replicas to 5. `kubectl scale replicas=5 -f app-deployment.yaml`. is not right.

While the outcome is the same, this approach doesn't update the change in the desired state configuration (YAML file). If you were to make some changes in your app-deployment.yaml and apply it, the update would scale back the replicas to 2. This is undesirable.

Ref: <https://kubernetes.io/docs/concepts/workloads/controllers/deployment/#scaling-a-deployment>

Disregard the YAML file. Enable autoscaling on the deployment to trigger on CPU usage and set minimum pods as well as maximum pods to 5. `kubectl autoscale myapp min=5 max=5 cpu-percent=80`. is not right.

While the outcome is the same, this approach doesn't update the change in the desired state configuration (YAML file). If you were to make some changes in your app-deployment.yaml and apply it, the update would scale back the replicas to 2. This is undesirable.

Ref: <https://kubernetes.io/blog/2016/07/autoscaling-in-kubernetes/>

Modify the current configuration of the deployment by using kubectl edit to open the YAML file of the current configuration, modify and save the configuration. kubectl edit deployment/app-deployment -o yaml save-config. is not right.

Like the above, the outcome is the same. This is equivalent to first getting the resource, editing it in a text editor, and then applying the resource with the updated version. This approach doesn't update the replicas change in our local YAML file. If you were to make some changes in your local app-deployment.yaml and apply it, the update would scale back the replicas to 2. This is undesirable.

Ref: <https://kubernetes.io/docs/concepts/cluster-administration/manage-deployment/#in-place-updates-of-resources>

Edit the number of replicas in the YAML file and rerun the kubectl apply. kubectl apply -f app-deployment.yaml. is the right answer.

This is the only approach that guarantees that you use desired state configuration. By updating the YAML file to have 5 replicas and applying it using kubectl apply, you are preserving the intended state of Kubernetes cluster in the YAML file.

Ref: <https://kubernetes.io/docs/concepts/cluster-administration/manage-deployment/#in-place-updates-of-resources>

### 3. Question

Your team uses Splunk for centralized logging and you have a number of reports and dashboards based on the logs in Splunk. You want to install Splunk forwarder on all nodes of your new Kubernetes Engine Autoscaled Cluster. The Splunk forwarder forwards the logs to a centralized Splunk Server. You want to minimize operational overhead. What is the best way to install Splunk Forwarder on all nodes in the cluster?

- Include the forwarder agent in a DaemonSet deployment.
- Use Deployment Manager to orchestrate the deployment of forwarder agents on all nodes.
- Include the forwarder agent in a StatefulSet deployment.
- SSH to each node and run a script to install the forwarder agent.

#### Unattempted

SSH to each node and run a script to install the forwarder agent. is not right.

While this can be done, this approach does not scale. Every time the Kubernetes cluster autoscaling adds a new node, we have to SSH to the instance and run the script which is manual, possibly error-prone and adds operational overhead. We need to look for a solution that automates this task.

Include the forwarder agent in a StatefulSet deployment. is not right.

In GKE, StatefulSets represents a set of Pods with unique, persistent identities and stable hostnames that GKE maintains regardless of where they are scheduled. The main purpose of StatefulSets is to set

up persistent storage for pods that are deployed across multiple zones. StatefulSets are not suitable for installing the forwarder agent nor do they provide us the ability to install forwarder agents.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/statefulset>

Use Deployment Manager to orchestrate the deployment of forwarder agents on all nodes. is not right.  
You can use a deployment manager to create a number of GCP resources including GKE Cluster but you can not use it to create Kubernetes deployments or apply configuration files.

Ref: <https://cloud.google.com/deployment-manager/docs/fundamentals>

Include the forwarder agent in a DaemonSet deployment. is the right answer.

In GKE, DaemonSets manage groups of replicated Pods and adhere to a one-Pod-per-node model, either across the entire cluster or a subset of nodes. As you add nodes to a node pool, DaemonSets automatically add Pods to the new nodes. So by configuring the pod to use Splunk forwarder agent image and with some minimal configuration (e.g. identifying which logs need to be forwarded), you can automate the installation and configuration of Splunk forwarder agent on each GKE cluster node.

Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/daemonset>

#### 4. Question

Your VMs are running in a subnet that has a subnet mask of 255.255.255.240. The current subnet has no more free IP addresses and you require an additional 10 IP addresses for new VMs. The existing and new VMs should all be able to reach each other without additional routes. What should you do?

- Use gcloud to expand the IP range of the current subnet.
- Delete the subnet, and recreate it using a wider range of IP addresses.
- Create a new project. Use Shared VPC to share the current network with the new project.
- Create a new subnet with the same starting IP but a wider range to overwrite the current subnet.

#### Unattempted

Use gcloud to expand the IP range of the current subnet. is the right answer.

Subnet mask of the existing subnet is 255.255.255.240 which means the max possible address in are 16. So the net prefix is /28 i.e. 4 bits free so 2 to the power of 4 is 16 IP Addresses.

As per IETF (Ref: <https://tools.ietf.org/html/rfc1918>), the supported internal IP Address ranges are

1. 24-bit block 10.0.0.0/8 (16777216 IP Addresses)
2. 20-bit block 172.16.0.0/12 (1048576 IP Addresses)
3. 16-bit block 192.168.0.0/16 (65536 IP Addresses)

A prefix of 28 is a very small subnet and could be in any of the ranges above; and all ranges have scope to accommodate a higher prefix.

A prefix of 27 gives you 32 IP Addresses i.e. 16 IP address more and we just need 10 more. So expanding the subnet to a prefix of 27 should give us the required capacity. And GCP lets you do exactly that running a gcloud command

<https://cloud.google.com/sdk/gcloud/reference/compute/networks/subnets/expand-ip-range>

```
gcloud compute networks subnets expand-ip-range --region= --prefix-length=27
```

## 5. Question

You've created a Kubernetes engine cluster named `my-gcp-ace-proj-1`, which has a cluster pool named `my-gcp-ace-primary-node-pool`. You want to increase the number of nodes within your cluster pool from 10 to 20 to meet capacity demands. What is the command to change the number of nodes in your pool?

- `gcloud container clusters update my-gcp-ace-proj-1 --node-pool my-gcp-ace-primary-node-pool --num-nodes 20`
- `gcloud container clusters resize my-gcp-ace-proj-1 --node-pool my-gcp-ace-primary-node-pool --num-nodes 20`
- `kubectl container clusters update my-gcp-ace-proj-1 --node-pool my-gcp-ace-primary-node-pool --num-nodes 20`
- `gcloud container clusters resize my-gcp-ace-proj-1 --node-pool my-gcp-ace-primary-node-pool --new-size 20`

### Unattempted

`kubectl container clusters update my-gcp-ace-proj-1 node-pool my-gcp-ace-primary-node-pool num-nodes 20.` is not right.

`kubectl` does not accept `container` as an operation.

Ref: <https://kubernetes.io/docs/reference/kubectl/overview/#operations>

`gcloud container clusters update my-gcp-ace-proj-1 node-pool my-gcp-ace-primary-node-pool num-nodes 20.` is not right.

`gcloud container clusters update` can not be used to specify the number of nodes. It can be used to specify the node locations, but not the number of nodes.

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/update>

`gcloud container clusters resize my-gcp-ace-proj-1 node-pool my-gcp-ace-primary-node-pool new-size 20.` is not right.

`gcloud container clusters resize` command does not support the parameter `new-size`. While `size` can be used to resize the cluster node pool, use of `size` is discouraged as this is a deprecated parameter. The

size flag is now deprecated. Please use num-nodes instead.

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/resize>

gcloud container clusters resize my-gcp-ace-proj-1 node-pool my-gcp-ace-primary-node-pool num-nodes

20. is the right answer

gcloud container clusters resize can be used to specify the number of nodes using the num-nodes parameter which is the target number of nodes in the cluster.

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/resize>

## 6. Question

You are designing a large distributed application with 30 microservices. Each of your distributed microservices needs to connect to a database back-end. You want to store the credentials securely. Where should you store the credentials?

- A. In the source code
- B. In an environment variable
- C. In a secret management system
- D. In a config file that has restricted access through ACLs

### Unattempted

Correct answer is C as it is a recommended practice to store the credentials in a secret management system such as KMS. Applications often require access to small pieces of sensitive data at build or run time. These pieces of data are often referred to as secrets. Secrets are similar in concept to configuration files, but are generally more sensitive, as they may grant access to additional data, such as user data.

Refer GCP documentation Authentication Managing Credentials

Best practices for managing credentials

Credentials provide access to sensitive data. The following practices help protect access to these resources.

Do not embed secrets related to authentication in source code, such as API keys, OAuth tokens, and service account credentials. You can use an environment variable pointing to credentials outside of the application's source code, such as Cloud Key Management Service.

Do use different credentials in different contexts, such as in testing and production environments.

Do transfer credentials only over HTTPS to prevent a third party from intercepting your credentials. Never transfer in clear text or as part of the URL.

Never embed long-lived credentials into your client-side app. For example, do not embed service account credentials into a mobile app. Client-side apps can be examined and credentials can easily be found and used by a third party.

Do revoke a token if you no longer need it.

Options A, B & D are wrong as they are not recommended and does not provide security.

## 7. Question

Your company's test suite is a custom C++ application that runs tests throughout each day on Linux virtual machines. The full test suite takes several hours to complete, running on a limited number of on-premises servers reserved for testing. Your company wants to move the testing infrastructure to the cloud, to reduce the amount of time it takes to fully test a change to the system, while changing the tests as little as possible. Which cloud infrastructure should you recommend?

- A. Google Compute Engine unmanaged instance groups and Network Load Balancer.
- B. Google Compute Engine managed instance groups with auto-scaling.
- C. Google Cloud Dataproc to run Apache Hadoop jobs to process each test.
- D. Google App Engine with Google Stackdriver for logging.

### Unattempted

Correct answer is B as Google Compute Engine managed instance group can help the testing application to scale to reduce the amount of time to run tests.

Refer GCP documentation Instance groups

A managed instance group uses an instance template to create a group of identical instances. You control a managed instance group as a single entity. If you wanted to make changes to instances that are part of a managed instance group, you would make the change to the whole instance group. Because managed instance groups contain identical instances, they offer the following features.

When your applications require additional compute resources, managed instance groups can automatically scale the number of instances in the group.

Managed instance groups work with load balancing services to distribute traffic to all of the instances in the group.

If an instance in the group stops, crashes, or is deleted by an action other than the instance groups commands, the managed instance group automatically recreates the instance so it can resume its processing tasks. The recreated instance uses the same name and the same instance template as the previous instance, even if the group references a different instance template.

Managed instance groups can automatically identify and recreate unhealthy instances in a group to ensure that all of the instances are running optimally.

The managed instance group updater allows you to easily deploy new versions of software to instances in your managed instance groups, while controlling the speed and scope of deployment as well as the level of disruption to your service.

Option A is wrong as unmanaged group does not scale.

Option C is wrong as Dataproc is for big data batch jobs.

Option D is wrong as App Engine standard does not support C++ application and the testing application needs to be dockerized to be used with flexible engine.

## 8. Question

Your company collects and stores security camera footage in Google Cloud Storage. Within the first 30 days, footage is processed regularly for threat detection, object detection, trend analysis, and suspicious behavior detection. You want to minimize the cost of storing all the data. How should you store the videos?

- A. Use Google Cloud Regional Storage for the first 30 days, and then move to Coldline Storage.
- B. Use Google Cloud Nearline Storage for the first 30 days, and then move to Coldline Storage.
- C. Use Google Cloud Regional Storage for the first 30 days, and then move to Nearline Storage.
- D. Use Google Cloud Regional Storage for the first 30 days, and then move to Google Persistent Disk.

### Unattempted

Correct answer is A as the data is accessed frequently within the first 30 days, using Google Cloud Regional Storage will enable the most cost-effective solution for storing and accessing the data. For videos older than 30 days, Google Cloud Coldline Storage offers the most cost-effective solution since it won't be accessed.

Refer GCP documentation [Cloud Storage](#) [Storage Classes](#)

Option B is wrong as while Google Cloud Coldline storage is cost-effective for long-term video storage, Google Cloud Nearline Storage would not be an effective solution for the first 30 days as the data is expected to be accessed frequently.

Option C is wrong as while Google Cloud Regional Storage is the most cost-effective solution for the first 30 days, Google Cloud Nearline Storage is not cost effective for long-term storage.

Option D is wrong as while Google Cloud Regional Storage is the most cost-effective solution for the first 30 days, storing the data on Google Cloud Persistent Disk would not be cost-effective for long term storage.

## 9. Question

Your company processes high volumes of IoT data that are time-stamped. The total data volume can be several petabytes. The data needs to be written and changed at a high speed. You want to use the most performant storage option for your data. Which product should you use?

- A. Cloud Datastore
- B. Cloud Storage
- C. Cloud Bigtable
- D. BigQuery

### Unattempted

Correct answer is C as Cloud Bigtable is the most performant storage option to work with IoT and time series data. Google Cloud Bigtable is a fast, fully managed, highly-scalable NoSQL database service. It is designed for the collection and retention of data from 1TB to hundreds of PB.

Refer GCP documentation Bigtable Time series data

Option A is wrong as Cloud Datastore is not the most performant product for frequent writes or timestamp-based queries.

Option B is wrong as Cloud Storage is designed for object storage not for this type of data ingestion and collection.

Option D is wrong as BigQuery is more of an a scalable, fully managed enterprise data warehousing solution and not ideal fast changing data.

## 10. Question

Your company is planning the infrastructure for a new large-scale application that will need to store over 100 TB or a petabyte of data in NoSQL format for Low-latency read/write and High-throughput analytics. Which storage option should you use?

- A. Cloud Bigtable
- B. Cloud Spanner
- C. Cloud SQL
- D. Cloud Datastore

#### Unattempted

Correct answer is A as Bigtable is an ideal solution to provide low latency, high throughput data processing storage option with analytics

Refer GCP documentation Storage Options

Cloud Bigtable logoCloud Bigtable

A scalable, fully managed NoSQL wide-column database that is suitable for both low-latency single-point lookups and precalculated analytics.

Low-latency read/write access IoT, finance, adtech High-throughput data processing Personalization, recommendations Time series support Monitoring Geospatial datasets Graphs

Options B & C are wrong as they are relational databases

Option D is wrong as Cloud Datastore is not ideal for analytics.

#### 11. Question

A company wants building an application stores images in a Cloud Storage bucket and want to generate thumbnails as well resize the images. They want to use managed service which will help them scale automatically from zero to scale and back to zero. Which GCP service satisfies the requirement?

- A. Google Compute Engine
- B. Google Kubernetes Engine
- C. Google App Engine
- D. Cloud Functions

#### Unattempted

Correct answer is D as Cloud Functions can help automatically scale as per the demand, with no invocations if no demand.

Refer GCP documentation Cloud Functions

Google Cloud Functions is a serverless execution environment for building and connecting cloud services. With Cloud Functions you write simple, single-purpose functions that are attached to events emitted from your cloud infrastructure and services. Your function is triggered when an event being watched is fired. Your code executes in a fully managed environment. There is no need to provision any infrastructure or worry about managing any servers.

Cloud Functions removes the work of managing servers, configuring software, updating frameworks, and patching operating systems. The software and infrastructure are fully managed by Google so that you just add code. Furthermore, provisioning of resources happens automatically in response to events. This means that a function can scale from a few invocations a day to many millions of invocations without any work from you.

Options A, B & C are wrong as they need to be configured to scale down and would need warm up time to scale back again as compared to Cloud Functions.

## 12. Question

Your company is planning on deploying a web application to Google Cloud hosted on a custom Linux distribution. Your website will be accessible globally and needs to scale to meet demand. Choose all of the components that will be necessary to achieve this goal. (Select TWO)

- A. App Engine Standard environment
- B. HTTP Load Balancer
- C. Managed Instance Group on Compute Engine
- D. Network Load Balancer

### Unattempted

Correct answers are B & C

Option B as only HTTP load balancer support global access.

Option C as the requirement is to support custom Linux distribution, only Compute Engine supports the same.

Refer GCP documentation Load Balancing

HTTP(S) load balancing can balance HTTP and HTTPS traffic across multiple backend instances, across multiple regions. Your entire app is available via a single global IP address, resulting in a simplified DNS setup. HTTP(S) load balancing is scalable, fault-tolerant, requires no pre-warming, and enables content-based load balancing. For HTTPS traffic, it provides SSL termination and load balancing.

Option A is wrong as App Engine does not support custom linux distribution.

Option D is wrong as Network load balancer does not support global access.

### 13. Question

Your customer is moving their corporate applications to Google Cloud Platform. The security team wants detailed visibility of all projects in the organization. You provision the Google Cloud Resource Manager and set up yourself as the org admin. What Google Cloud Identity and Access Management (Cloud IAM) roles should you give to the security team?

- A. Org viewer, project owner
- B. Org viewer, project viewer
- C. Org admin, project browser
- D. Project owner, network admin

#### Unattempted

Correct answer is B as the security team only needs visibility to the projects, project viewer provides the same with the best practice of least privilege.

Refer GCP documentation Organization & Project access control

Option A is wrong as project owner will provide access however it does not align with the best practice of least privilege.

Option C is wrong as org admin does not align with the best practice of least privilege.

Option D is wrong as the user needs to be provided organization viewer access to see the organization.

### 14. Question

Auditors visit your teams every 12 months and ask to review all the Google Cloud Identity and Access Management (Cloud IAM) policy changes in the previous 12 months. You want to streamline and expedite the analysis and audit process. What should you do?

- A. Create custom Google Stackdriver alerts and send them to the auditor
- B. Enable Logging export to Google BigQuery and use ACLs and views to scope the data shared with the auditor
- C. Use Cloud Functions to transfer log entries to Google Cloud SQL and use ACLs and views to limit an auditor's view
- D. Enable Google Cloud Storage (GCS) log export to audit logs into a GCS bucket and delegate access to the bucket

#### Unattempted

Correct answer is B as BigQuery is a good storage option with analysis capability. Also, the access to the data can be controlled using ACLs and Views.

BigQuery uses access control lists (ACLs) to manage permissions on projects and datasets.

BigQuery is a petabyte-scale analytics data warehouse that you can use to run SQL queries over vast amounts of data in near realtime.

Giving a view access to a dataset is also known as creating an authorized view in BigQuery. An authorized view allows you to share query results with particular users and groups without giving them access to the underlying tables. You can also use the view's SQL query to restrict the columns (fields) the users are able to query. In this tutorial, you create an authorized view.

Option A is wrong as alerts are real time and auditor do not need them.

Option C is wrong as Cloud SQL is not ideal for storage of log files and cannot be controlled through ACLs.

Option D is wrong as Cloud Storage is a good storage option but does not provide direct analytics capabilities.

### 15. Question

Your App Engine application needs to store stateful data in a proper storage service. Your data is non-relational database data. You do not expect the database size to grow beyond 10 GB and you need to have the ability to scale down to zero to avoid unnecessary costs. Which storage service should you use?

- A. Cloud Bigtable
- B. Cloud Dataproc
- C. Cloud SQL

D. Cloud Datastore**Unattempted**

Correct answer is D as Cloud Datastore provides a scalable, fully managed NoSQL document database for your web and mobile applications.

Cloud Datastore A scalable, fully managed NoSQL document database for your web and mobile applications. Semistructured application data User profiles Hierarchical data Product catalogs Durable key-value data Game state

Option A is wrong as Bigtable is not an ideal storage option for state management. Cloud Bigtable A scalable, fully managed NoSQL wide-column database that is suitable for both low-latency single-point lookups and precalculated analytics.Low-latency read/write access IoT, finance, adtech High-throughput data processing Personalization, recommendations Time series support Monitoring Geospatial datasets Graphs

Option B is wrong as Dataproc is not a storage solution. Cloud Dataproc is a fast, easy-to-use, fully-managed cloud service for running Apache Spark and Apache Hadoop clusters in a simpler, more cost-efficient way.

Option C is wrong as you need to define a capacity while provisioning a database.

Cloud SQL A fully managed MySQL and PostgreSQL database service that is built on the strength and reliability of Google's infrastructure. Web frameworks Websites, blogs, and content management systems (CMS) Structured data Business intelligence (BI) applications

OLTP workloads ERP, CRM, and ecommerce applications Geospatial application

## 16. Question

You have a collection of media files over 50GB each that you need to migrate to Google Cloud Storage. The files are in your on-premises data center. What migration method can you use to help speed up the transfer process?

- A. Use multi-threaded uploads using the -m option.
- B. Use parallel uploads to break the file into smaller chunks then transfer it simultaneously.
- C. Use the Cloud Transfer Service to transfer.
- D. Start a recursive upload.

**Unattempted**

Correct answer is B as gsutil provide object composition or parallel upload to handle upload of larger files.

Refer GCP documentation Optimizing for Cloud Storage Performance

More efficient large file uploads

The gsutil utility can also automatically use object composition to perform uploads in parallel for large, local files that you want to upload to Cloud Storage. It splits a large file into component pieces, uploads them in parallel and then recomposes them once they're in the cloud (and deletes the temporary components it created locally).

You can enable this by setting the `parallel\_composite\_upload\_threshold` option on gsutil (or, updating your .boto file, like the console output suggests).

```
gsutil -o GSUtil:parallel_composite_upload_threshold=150M cp ./localbigfile gs://your-bucket
```

Where `localbigfile` is a file larger than 150MB. This divides up your data into chunks ~ 150MB and uploads them in parallel, increasing upload performance.

Option A is wrong as multi-threaded options is best suited for uploading multiple files to better utilize the bandwidth.

Option C is wrong as Cloud Transfer service cannot handle uploads from on-premises data center.

Option D is wrong as recursive upload helps handle folders and subfolders.

## 17. Question

A Company is planning the migration of their web application to Google App Engine. However, they would still continue to use their on-premises database. How can they setup application?

- A. Setup the application using App Engine Standard environment with Cloud VPN to connect to database
- B. Setup the application using App Engine Flexible environment with Cloud VPN to connect to database
- C. Setup the application using App Engine Standard environment with Cloud Router to connect to database
- D. Setup the application using App Engine Flexible environment with Cloud Router to connect to database

Unattempted

Correct answer is B as Google App Engine provides connectivity to on-premises using Cloud VPN.

Refer GCP documentation App Engine Flexible Network Settings

Advanced network configuration

You can segment your Compute Engine network into subnetworks. This allows you to enable VPN scenarios, such as accessing databases within your corporate network.

To enable subnetworks for your App Engine application:

Create a custom subnet network.

Add the network name and subnetwork name to your app.yaml file, as specified above.

To establish a simple VPN based on static routing, create a gateway and a tunnel for a custom subnet network. Otherwise, see how to create other types of VPNs.

Option A is wrong as Google App Engine Standard cannot use Cloud VPN.

Options C & D are wrong as you need a Cloud VPN to connect to on-premises data center. Cloud Route support dynamic routing.

## 18. Question

A lead software engineer tells you that his new application design uses websockets and HTTP sessions that are not distributed across the web servers. You want to help him ensure his application will run properly on Google Cloud Platform. What should you do?

- A. Help the engineer to convert his websocket code to use HTTP streaming.
- B. Review the encryption requirements for websocket connections with the security team.
- C. Meet with the cloud operations team and the engineer to discuss load balancer options.
- D. Help the engineer redesign the application to use a distributed user session service that does not rely on websockets and HTTP sessions.

### Unattempted

Correct answer is C as the HTTP(S) load balancer in GCP handles websocket traffic natively. Backends that use WebSocket to communicate with clients can use the HTTP(S) load balancer as a front end, for scale and availability.

Refer GCP documentation HTTP Load Balancer

HTTP(S) Load Balancing has native support for the WebSocket protocol. Backends that use WebSocket to communicate with clients can use the HTTP(S) load balancer as a front end, for scale and availability. The load balancer does not need any additional configuration to proxy WebSocket connections.

The WebSocket protocol, which is defined in RFC 6455, provides a full-duplex communication channel between clients and servers. The channel is initiated from an HTTP(S) request

Option A is wrong as there is no compelling reason to move away from websockets as part of a move to GCP.

Option B is wrong as this may be a good exercise anyway, but it doesn't really have any bearing on the GCP migration.

Option D is wrong as there is no compelling reason to move away from websockets as part of a move to GCP.

## 19. Question

Your customer is moving their storage product to Google Cloud Storage (GCS). The data contains personally identifiable information (PII) and sensitive customer information. What security strategy should you use for GCS?

- A. Use signed URLs to generate time bound access to objects.
- B. Grant IAM read-only access to users, and use default ACLs on the bucket.
- C. Grant no Google Cloud Identity and Access Management (Cloud IAM) roles to users, and use granular ACLs on the bucket.
- D. Create randomized bucket and object names. Enable public access, but only provide specific file URLs to people who do not have Google accounts and need access.

### Unattempted

Correct answer is C as this grants the least privilege required to access the data and minimizes the risk of accidentally granting access to the wrong people.

Refer GCP documentation Cloud Storage Access Control

Option A is wrong as Signed URLs could potentially be leaked as anyone who gets access to the URL can access the data.

Option B is wrong as this is needlessly permissive, users only require one permission in order to get access.

Option D is wrong as this is security through obscurity, also known as no security at all.

## 20. Question

You've created a Kubernetes engine cluster named `project-1`, which has a cluster pool named `primary-node-pool`. You've realized that you need more total nodes within your cluster pool to meet capacity demands from 10 to 20. What is the command to change the number of nodes in your pool?

- A. gcloud container clusters update project-1 --node pool 'primary-node-pool' --num-nodes 20
- B. gcloud container clusters resize project-1 --node pool 'primary-node-pool' --size 20
- C. gcloud container clusters resize project-1 --node pool 'primary-node-pool' --num-nodes 20
- D. kubectl container clusters update project-1 --node pool 'primary-node-pool' --num-nodes 20

### Unattempted

Correct answer is B as the resize command with gcloud can be used to increase the nodes.

NOTE The size flag has been renamed to num-nodes flag from 242.0.0 (2019-04-16)

### Kubernetes Engine

Renamed `size` flag of `gcloud container clusters resize` to `num-nodes`. `size` retained as an alias.

Disabled node auto-repair and node auto-upgrade by default when `enable-kubernetes-alpha` flag is used to create clusters with Kubernetes alpha features enabled. Users may now create alpha clusters without specifying `no-enable-autorepair` or `no-enable-autoupgrade` flags. However, for creating new node pools in an existing alpha cluster, these two flags may still be required.

Refer GCP documentation Resizing Kubernetes Cluster

`gcloud container clusters resize [CLUSTER_NAME] node-pool [POOL_NAME] size [SIZE];`

Option A is wrong as update command takes in the `max-nodes` & `min-nodes` flags which are defining the autoscaling. `num-nodes` flag is not applicable.

Option C is wrong as `num-nodes` is a wrong flag for cluster resize command.

Option D is wrong as `kubectl` command cannot be used for resizing the cluster.

## 21. Question

A Company is using Cloud SQL to host critical data. They want to enable high availability in case a complete zone goes down. How should you configure the same?

- A. Create a Read replica in the same region different zone
- B. Create a Read replica in the different region different zone
- C. Create a Failover replica in the same region different zone
- D. Create a Failover replica in the different region different zone

#### Unattempted

Correct answer is C as a failover replica helps provides High Availability for Cloud SQL. The failover replica must be in the same region as the primary instance.

Refer GCP documentation Cloud SQL High Availability

The HA configuration, sometimes called a cluster, provides data redundancy. The configuration is made up of a primary instance (master) in the primary zone and a failover replica in the secondary zone.

Through semisynchronous replication, all changes made to the primary instance's data and user tables are copied onto the failover replica. In the event of an instance or zone failure, this configuration reduces downtime, and your data continues to be available to client applications.

The failover replica must be in the same region as the primary instance, but in a different zone.

Diagram overview of MySQL HA configuration. Described in text below.

Option A & B are wrong as Read replicas do not provide failover capability and just additional read capacity.

Option D is wrong as failover replica must be in the same region as the primary instance.

## 22. Question

Your application is hosted across multiple regions and consists of both relational database data and static images. Your database has over 10 TB of data. You want to use a single storage repository for each data type across all regions. Which two products would you choose for this task? (Choose two)

- A. Cloud Bigtable
- B. Cloud Spanner
- C. Cloud SQL
- D. Cloud Storage

**Unattempted**

Correct answers are B & D

Option B to store the relational data. As the data is over 10TB and need across region, Cloud Spanner is preferred over Cloud SQL.

Option D to store unstructured static images.

Refer GCP documentation Storage Options

Option A is wrong as Bigtable is a NoSQL data storage and not suitable to store unstructured data as images and files.

Option C is wrong as Cloud SQL is regional and not a preferred option for data over 10TB.

### 23. Question

Your project has all its Compute Engine resources in the europe-west1 region. You want to set europe-west1 as the default region for gcloud commands. What should you do?

- A. Use Cloud Shell instead of the command line interface of your device. Launch Cloud Shell after you navigate to a resource in the europe-west1 region. The europe-west1 region will automatically become the default region.
- B. Use `gcloud config set compute/region europe-west1` to set the default region for future gcloud commands.
- C. Use `gcloud config set compute/zone europe-west1` to set the default region for future gcloud commands.
- D. Create a VPN from on-premises to a subnet in europe-west1, and use that connection when executing gcloud commands.

**Unattempted**

Correct answer is B as this will ensure that the relevant region is used when not overwritten by a command parameter.

Refer GCP documentation Change default zone and region

You can manually choose a different zone or region without updating the metadata server by setting these properties locally on your gcloud client.

`gcloud config compute/region REGION`

Option A is wrong as Cloud Shell will not default to the location that it's launched from.

Option C is wrong as this command should be used to set a zone, not a region.

Option D is wrong as a VPN to a specific subnet does not have any effect on the gcloud command region.

## 24. Question

You have an application server running on Compute Engine in the europe-west1-d zone. You need to ensure high availability and replicate the server to the europe-west2-c zone using the fewest steps possible. What should you do?

- A. Create a snapshot from the disk. Create a disk from the snapshot in the europe-west2-c zone. Create a new VM with that disk.
- B. Create a snapshot from the disk. Create a disk from the snapshot in the europe-west1-d zone and then move the disk to europe-west2-c. Create a new VM with that disk.
- C. Use gcloud to copy the disk to the europe-west2-c zone. Create a new VM with that disk.
- D. Use gcloud compute instances move with parameter --destination-zone europe-west2-c to move the instance to the new zone.

### Unattempted

Correct answer is A as the best way to create a replica of disk is to create a snapshot and create a disk from the snapshot in the zone.

Refer GCP documentation Disks

Disks are zonal resources, so they reside in a particular zone for their entire lifetime. The contents of a disk can be moved to a different zone by snapshotting the disk (using gcloud compute disks snapshot) and creating a new disk using source-snapshot in the desired zone. The contents of a disk can also be moved across project or zone by creating an image (using gcloud compute images create) and creating a new disk using image in the desired project and/or zone.

Option B is wrong as the approach is possible, but not with the fewest steps.

Option C is wrong as gcloud cannot be used to copy the disk to different zone.

Option D is wrong as it would move and not create a copy. gcloud compute disks move facilitates moving a Google Compute Engine disk volume from one zone to another. You cannot move a disk if it is attached to a running or stopped instance; use the gcloud compute instances move command instead.

## 25. Question

You need to estimate the annual cost of running a BigQuery query that is scheduled to run nightly. What should you do?

- A. Use gcloud query --dry\_run to determine the number of bytes read by the query. Use this number in the Pricing Calculator.
- B. Use bq query --dry\_run to determine the number of bytes read by the query. Use this number in the Pricing Calculator.
- C. Use gcloud estimate to determine the amount billed for a single query. Multiply this amount by 365.
- D. Use bq estimate to determine the amount billed for a single query. Multiply this amount by 365.

### Unattempted

Correct answer is B as this is the correct way to estimate the yearly BigQuery querying costs.

Refer GCP documentation   BigQuery Best Practices   Price your Query

Best practice: Before running queries, preview them to estimate costs.

Queries are billed according to the number of bytes read. To estimate costs before running a query use:

The query validator in the GCP Console or the classic web UI

The dry\_run flag in the CLI

The dryRun parameter when submitting a query job using the API

The Google Cloud Platform Pricing Calculator

Option A is wrong as you should use `bq`, not `gcloud`, to estimate the amount of bytes read.

Option C is wrong as you should use `bq`, not `gcloud`, to work with BigQuery.

Option D is wrong as this will not give the amount billed for a query.

## 26. Question

You work in a small company where everyone should be able to view all resources of a specific project.

You want to grant them access following Google's recommended practices. What should you do?

- A. Create a script that uses gcloud projects add-iam-policy-binding for all users' email addresses and the Project Viewer role.
- B. Create a script that uses gcloud iam roles create for all users' email addresses and the Project Viewer role.
- C. Create a new Google Group and add all users to the group. Use gcloud projects add-iam-policy-binding with the Project Viewer role and Group email address.
- D. Create a new Google Group and add all members to the group. Use gcloud iam roles create with the Project Viewer role and Group email address.

#### Unattempted

Correct answer is C as Google recommends to use groups where possible.

Refer GCP documentation gcloud IAM

Option A is wrong as groups are recommended over individual assignments.

Option B is wrong as this command is to create roles, not to assign them.

Option D is wrong as this command is to create roles, not to assign them.

## 27. Question

Your developers are trying to select the best compute service to run a static website. They have a dozen HTML pages, a few JavaScript files, and some CSS. They need the site to be highly available for the few weeks it is running. They also have a limited budget. What is the best service to use to run the site?

- A. Kubernetes Engine
- B. Compute Engine
- C. Cloud Storage
- D. App Engine

#### Unattempted

Correct answer is C as the website is static and needs to be hosted with high availability and limited budget, Cloud Storage would be an ideal choice.

Refer GCP documentation Cloud Storage Static Website

To host a static site in Cloud Storage, you need to create a Cloud Storage bucket, upload the content, and test your new site. You can serve your data directly from storage.googleapis.com, or you can verify

that you own your domain and use your domain name. Either way, you'll get consistent, fast delivery from global edge caches.

You can create your static web pages however you choose. For example, you could hand-author pages by using HTML and CSS. You can use a static-site generator, such as Jekyll, Ghost, or Hugo, to create the content. Static-site generators make it easier for you to create a static website by letting you author in markdown, and providing templates and tools. Site generators generally provide a local web server that you can use to preview your content.

After your static site is working, you can update the static pages by using any process you like. That process could be as straightforward as hand-copying an updated page to the bucket. You might choose to use a more automated approach, such as storing your content on GitHub and then using a webhook to run a script that updates the bucket. An even more advanced system might use a continuous-integration /continuous-delivery (CI/CD) tool, such as Jenkins, to update the content in the bucket. Jenkins has a Cloud Storage plugin that provides a Google Cloud Storage Uploader post-build step to publish build artifacts to Cloud Storage.

If you have a web application that needs to serve static content or user-uploaded static media, using Cloud Storage can be a cost-effective and efficient way to host and serve this content, while reducing the amount of dynamic requests to your web application.

Options A, B & D are wrong as they would be an expensive option as compared to Cloud Storage hosting.

## 28. Question

You have an autoscaled managed instance group that is set to scale based on CPU utilization of 60%. There are currently 3 instances in the instance group. You're connected to one of the instances and notice that the CPU usage is at 70%. However, the instance group isn't starting up another instance. What's the most likely reason?

- A. The autoscaler is disabled.
- B. The autoscaler takes 60 seconds before creating a new instance.
- C. The load balancer doesn't recognize the instance as healthy.
- D. The average CPU for the entire instance group is below 60%.

### Unattempted

Correct answer is D as the Auto Scaler checks for the average CPU utilization across the instances and is not done on the basis of a single instance.

Refer GCP documentation Auto Scaler CPU based Scaling

You can autoscale based on the average CPU utilization of a managed instance group. Using this policy tells the autoscaler to collect the CPU utilization of the instances in the group and determine whether it needs to scale. You set the target CPU utilization the autoscaler should maintain and the autoscaler will work to maintain that level.

The autoscaler treats the target CPU utilization level as a fraction of the average use of all vCPUs over time in the instance group. If the average usage of your total vCPUs exceeds the target utilization, the autoscaler will add more virtual machines. For example, setting a 0.75 target utilization tells the autoscaler to maintain an average usage of 75% among all vCPUs in the instance group.

Option A is wrong as the group is set to CPU utilization already, it is not disabled.

Option B is wrong as Auto Scaler takes action immediately if the target is hit.

Option C is wrong as if the instance is marked unhealthy it would not serve any traffic and might be replaced.

## 29. Question

You are required to fire a query on large amount of data stored in BigQuery. You know the query is expected to return a large amount of data. How would you estimate the cost for the query?

- A. Using Command line, use the `--dry_run` option on BigQuery to determine the amount of bytes read, and then use the price calculator to determine the cost.
- B. Using Command line, use the `--dry_run` option on BigQuery to determine the amount of bytes returned, and then use the price calculator to determine the cost.
- C. Using Command line, use the `--dry_run` option on BigQuery to determine the amount of time taken, and then use the price calculator to determine the cost.
- D. Using Command line, use the `--dry_run` option on BigQuery to determine the total amount of table data in bytes, as it would be a full scan, and then use the price calculator to determine the cost.

### Unattempted

Correct answer is A as the `dry-run` option can be used to price your queries before they are actually fired. The Query returns the bytes read, which can then be used with the Pricing Calculator to estimate the query cost.

Refer GCP documentation BigQuery Best Practices

Price your queries before running them

Best practice: Before running queries, preview them to estimate costs.

Queries are billed according to the number of bytes read. To estimate costs before running a query use:

The query validator in the GCP Console or the classic web UI

The `dry_run` flag in the CLI

The `dryRun` parameter when submitting a query job using the API

The Google Cloud Platform Pricing Calculator

Options B, C are wrong as the estimation needs to be done on the bytes read by the query and not returned or time taken.

Option D is wrong as it the bytes read would depend on the query and would not always a full table scan.

### 30. Question

Your company wants to host confidential documents in Cloud Storage. Due to compliance requirements, there is a need for the data to be highly available and resilient even in case of a regional outage. Which storage classes help meet the requirement?

- A. Nearline
- B. Standard
- C. Multi-Regional
- D. Dual-Regional
- E. Regional

#### Unattempted

Correct answers are A & C as Multi-Regional and Nearline storage classes provide multi-region geo-redundant deployment, which can sustain regional failure.

Refer GCP documentation Cloud Storage Classes

Multi-Regional Storage is geo-redundant.

The geo-redundancy of Nearline Storage data is determined by the type of location in which it is stored: Nearline Storage data stored in multi-regional locations is redundant across multiple regions, providing

higher availability than Nearline Storage data stored in regional locations.

Data that is geo-redundant is stored redundantly in at least two separate geographic places separated by at least 100 miles. Objects stored in multi-regional locations are geo-redundant, regardless of their storage class.

Geo-redundancy occurs asynchronously, but all Cloud Storage data is redundant within at least one geographic place as soon as you upload it.

Geo-redundancy ensures maximum availability of your data, even in the event of large-scale disruptions, such as natural disasters. For a dual-regional location, geo-redundancy is achieved using two specific regional locations. For other multi-regional locations, geo-redundancy is achieved using any combination of data centers within the specified multi-region, which may include data centers that are not explicitly available as regional locations.

Options B & D are wrong as they do not exist

Option E is wrong as Regional storage class is not geo-redundant. Data stored in a narrow geographic region and Redundancy is across availability zones

### 31. Question

Your company needs to backup data for disaster recovery scenarios store all the backup data. This data would be required only in the event of a disaster and won't be accessed otherwise. What is the best default storage class?

- A. Multi-regional
- B. Coldline
- C. Regional
- D. Nearline

#### Unattempted

Correct answer is B as Coldline storage is an ideal solution for disaster recovery data given its rarity of access.

Refer GCP documentation Cloud Storage Classes

Google Cloud Storage Coldline is a very-low-cost, highly durable storage service for data archiving, online backup, and disaster recovery. Unlike other cold storage services, your data is available within milliseconds, not hours or days.

Coldline Storage is the best choice for data that you plan to access at most once a year, due to its slightly lower availability, 90-day minimum storage duration, costs for data access, and higher per-operation costs. For example:

**Cold Data Storage** Infrequently accessed data, such as data stored for legal or regulatory reasons, can be stored at low cost as Coldline Storage, and be available when you need it.

**Disaster recovery** In the event of a disaster recovery event, recovery time is key. Cloud Storage provides low latency access to data stored as Coldline Storage.

The geo-redundancy of Coldline Storage data is determined by the type of location in which it is stored: Coldline Storage data stored in multi-regional locations is redundant across multiple regions, providing higher availability than Coldline Storage data stored in regional locations.

Options A, C & D are wrong as they are not suited for infrequently accessed data, as disaster does not happen periodically but rarely.

## 32. Question

Your company needs to backup data for disaster recovery scenarios store all the backup data. You are required to perform monthly disaster recovery drills, as a part of compliance. What is the best default storage class?

- A. Multi-regional
- B. Coldline
- C. Regional
- D. Nearline

### Unattempted

Correct answer is D as the data needs to be access monthly only, Nearline is the ideal solution for data storage.

Refer GCP documentation Cloud Storage Classes

Google Cloud Storage Nearline is a low-cost, highly durable storage service for storing infrequently accessed data. Nearline Storage is a better choice than Multi-Regional Storage or Regional Storage in scenarios where slightly lower availability, a 30-day minimum storage duration, and costs for data access are acceptable trade-offs for lowered storage costs.

Nearline Storage is ideal for data you plan to read or modify on average once a month or less. For example, if you want to continuously add files to Cloud Storage and plan to access those files once a month for analysis, Nearline Storage is a great choice.

Nearline Storage is also appropriate for data backup, disaster recovery, and archival storage. Note, however, that for data accessed less frequently than once a year, Coldline Storage is the most cost-effective choice, as it offers the lowest storage costs.

The geo-redundancy of Nearline Storage data is determined by the type of location in which it is stored: Nearline Storage data stored in multi-regional locations is redundant across multiple regions, providing higher availability than Nearline Storage data stored in regional locations.

Options A, B & C are wrong as they are not ideal for data that is only accessed monthly.

### 33. Question

Your developers are trying to connect to an Ubuntu server over SSH to diagnose some errors. However, the connection times out. Which command should help solve the problem?

- A. gcloud compute firewall-rules create open-ssh --network \$NETWORK --allow tcp:22
- B. gcloud compute firewall-rules create open-ssh
- C. gcloud compute firewall-rules create open-ssh --network \$NETWORK --deny tcp:22
- D. gcloud compute firewall-rules create open-ssh --network \$NETWORK --allow tcp:3389

#### Unattempted

Correct answer is A as gcloud compute firewall-rules create is used to create firewall rules to allow/deny incoming/outgoing traffic.

Refer GCP documentation Cloud SDK Firewall Rules Create

allow=PROTOCOL[:PORT[-PORT]],[]

A list of protocols and ports whose traffic will be allowed.

The protocols allowed over this connection. This can be the (case-sensitive) string values tcp, udp, icmp, esp, ah, sctp, or any IP protocol number. An IP-based protocol must be specified for each rule. The rule applies only to specified protocol.

For port-based protocols tcp, udp, and sctp a list of destination ports or port ranges to which the rule applies may optionally be specified. If no port or port range is specified, the rule applies to all destination ports.

The ICMP protocol is supported, but there is no support for configuring ICMP packet filtering by ICMP code.

For example, to create a rule that allows TCP traffic through port 80 and ICMP traffic:

```
gcloud compute firewall-rules create MY-RULE allow tcp:80,icmp
```

To create a rule that allows TCP traffic from port 20000 to 25000:

```
gcloud compute firewall-rules create MY-RULE allow tcp:20000-25000
```

To create a rule that allows all TCP traffic:

```
gcloud compute firewall-rules create MY-RULE allow tcp
```

Option B is wrong as the command would result in error.

ERROR: (gcloud.compute.firewall-rules.create) Exactly one of ( action | allow) must be specified.

Option C is wrong as deny rule would prevent SSH login.

Option D is wrong as the port 3389 is for RDP and not for SSH.

### 34. Question

You're working on creating a script that can extract the IP address of a Kubernetes Service. Your coworker sent you a code snippet that they had saved. Which one is the best starting point for your code?

- A. kubectl get svc -o filtered-json='[.items[\*].status.loadBalancer.ingress[0].ip]'
- B. kubectl get svc -o jsonpath='[.items[\*].status.loadBalancer.ingress[0].ip]'
- C. kubectl get svc -o html
- D. kubectl get svc

#### Unattempted

Correct answer is B as kubectl get svc can be used to get the data, and jsonpath can be used to parse the data.

Refer Kubernetes documentation [Kubernetes IO & Tutorials](#)

```
$ kubectl get services
NAME CLUSTER-IP EXTERNAL-IP PORT(S) kubernetes 10.0.0.1 443/TCP
bootcamp 10.3.245.61 104.155.111.170 8080/TCP
```

To access the services, use the external IP and the application port e.g. like this:

```
$ export EXTERNAL_IP=$(kubectl get service bootcamp
output=jsonpath='{.status.loadBalancer.ingress[0].ip} ') $ export PORT=$(kubectl get services
output=jsonpath='{.items[0].spec.ports[0].port} ') $ curl $EXTERNAL_IP:$PORT Hello Kubernetes
bootcamp! | Running on: bootcamp-390780338-2fhnk | v=1
```

### 35. Question

Your team needs to set up a new Jenkins instance as quickly as possible. What's the best way to get it up-and-running?

- A. Use Google's Managed Jenkins Service.
- B. Deploy the jar file to a Compute Engine instance.
- C. Install with Cloud Launcher
- D. Create a Deployment Manager template and deploy it.

#### Unattempted

Correct answer is C as Cloud Launcher provides

Refer GCP documentation Marketplace (Formerly Cloud Launcher)

GCP Marketplace offers ready-to-go development stacks, solutions, and services to accelerate development. So you spend less time installing and more time developing.

Deploy production-grade solutions in a few clicks

Single bill for all your GCP and 3rd party services

Manage solutions using Deployment Manager

Notifications when a security update is available

Direct access to partner support

Option A is wrong as there is no Google's Managed Jenkins Service.

Option B is wrong as hosting on the compute engine is still a manual step.

Option D is wrong as Deployment Manager would take time to build and deploy.

## 36. Question

You have a Cloud Storage bucket that needs to host static web assets with a dozen HTML pages, a few JavaScript files, and some CSS. How do you make the bucket public?

- A. Set allAuthenticatedUsers to have the Storage Object Viewer role.
- B. Check the make public box on the GCP Console for the bucket
- C. Set allUsers to have the Storage Object Viewer role.
- D. gsutil make-public gs://bucket-name

### Unattempted

Correct answer is C as the bucket can be shared by providing the Storage Object Viewer access to allUsers.

Refer GCP documentation [Cloud Storage Sharing files](#)

You can either make all files in your bucket publicly accessible, or you can set individual objects to be accessible through your website. Generally, making all files in your bucket accessible is easier and faster.

To make all files accessible, follow the Cloud Storage guide for making groups of objects publicly readable.

To make individual files accessible, follow the Cloud Storage guide for making individual objects publicly readable.

If you choose to control the accessibility of individual files, you can set the default object ACL for your bucket so that subsequent files uploaded to your bucket are shared by default.

1. Open the Cloud Storage browser in the Google Cloud Platform Console.
2. In the list of buckets, click on the name of the bucket that contains the object you want to make public, and navigate to the object if it's in a subdirectory.
3. Click the drop-down menu associated with the object that you want to make public. The drop-down menu appears as three vertical dots to the far right of the object's row.
4. Select Edit permissions from the drop-down menu.
5. In the overlay that appears, click the + Add item button.
6. Add a permission for allUsers.

Select User for the Entity.

Enter allUsers for the Name.

Select Reader for the Access.

7. Click Save.

Option A is wrong as access needs to be provided to allUsers to make it public and there is no allAuthenticatedUsers option.

Option B is wrong as there is no make public option with GCP Console.

Option D is wrong as there is no make public option with gsutil command.

### 37. Question

Your company has been running their marketing application on App Engine app for a few weeks with Autoscaling, and it's been performing well. However, the marketing team is planning on a massive campaign, and they expect a lot of burst traffic. How would you go about ensuring there are always 3 idle instances?

- A. Set the min\_instances property in the app.yaml
- B. Switch to manual scaling and use the burst\_traffic\_protection property to True in the app.yaml.
- C. Set the min\_idle\_instances property in the app.yaml.
- D. Switch to manual scaling and use the idle\_instance\_count property in the app.yaml.

#### Unattempted

Correct answer is C as min\_idle\_instances property can be set to have minimum idle instances which would be always running.

Refer GCP documentation [App Engine Scaling & app.yaml Reference](#)

Auto scaling services use dynamic instances that get created based on request rate, response latencies, and other application metrics. However, if you specify a number of minimum idle instances, that specified number of instances run as resident instances while any additional instances are dynamic.

min\_idle\_instances

The number of instances to be kept running and ready to serve traffic. Note that you are charged for the number of instances specified whether they are receiving traffic or not. This setting only applies to the

version that receives most of the traffic. Keep the following in mind:

A low minimum helps keep your running costs down during idle periods, but means that fewer instances might be immediately available to respond to a sudden load spike.

A high minimum allows you to prime the application for rapid spikes in request load. App Engine keeps the minimum number of instances running to serve incoming requests. You are charged for the number of instances specified, whether or not they are handling requests. For this feature to function properly, you must make sure that warmup requests are enabled and that your application handles warmup requests.

If you set a minimum number of idle instances, pending latency will have less effect on your application's performance. Because App Engine keeps idle instances in reserve, it is unlikely that requests will enter the pending queue except in exceptionally high load spikes. You will need to test your application and expected traffic volume to determine the ideal number of instances to keep in reserve.

Option A is wrong as min\_instances applies to dynamic scaling. Also, number of instances

Options B & D are wrong as manual scaling would not provide the minimal running instances.

### 38. Question

Your team has some new functionality that they want to roll out slowly so they can monitor for errors. The change contains some significant changes to the user interface. You've chosen to use traffic splitting to perform a canary deployment. You're going to start by rolling out the code to 15% of your users. How should you go about setting up traffic splitting with the user getting the same experience?

- A. Deploy the new version. Split the traffic using an IP or cookie based distribution.
- B. Use the gcloud app deploy command with the distribution flag to deploy and split the traffic in one command.
- C. Deploy the new version using the no-promote flag. Split the traffic using a random distribution.
- D. Deploy the new version using the no-promote flag. Split the traffic using Cookie.

### Unattempted

Correct answer is D as the application needs to be promoted using the no-promote parameter to avoid the new version getting all the 100% traffic. Once the application is deployed and tested, the traffic can be split using the Cookie approach to maintain User experience.

Refer GCP documentation Splitting Traffic

When you have specified two or more versions for splitting, you must choose whether to split traffic by using either an IP address or HTTP cookie. It's easier to set up an IP address split, but a cookie split is more precise.

Options A & B are wrong as deploying the new version would configure it to receive all the traffic.

Option C is wrong as random distribution would not help maintain user experience.

### 39. Question

Your company has decided to store data files in Cloud Storage. The data would be hosted in a regional bucket to start with. You need to configure Cloud Storage lifecycle rule to move the data for archival after 30 days and delete the data after a year. Which two actions should you take?

- A. Create a Cloud Storage lifecycle rule with Age: 30, Storage Class: Standard, and Action: Set to Coldline, and create a second GCS life-cycle rule with Age: 365, Storage Class: Coldline, and Action: Delete.
- B. Create a Cloud Storage lifecycle rule with Age: 30, Storage Class: Standard, and Action: Set to Coldline, and create a second GCS life-cycle rule with Age: 275, Storage Class: Coldline, and Action: Delete.
- C. Create a Cloud Storage lifecycle rule with Age: 30, Storage Class: Standard, and Action: Set to Nearline, and create a second GCS life-cycle rule with Age: 365, Storage Class: Nearline, and Action: Delete.
- D. Create a Cloud Storage lifecycle rule with Age: 30, Storage Class: Standard, and Action: Set to Nearline, and create a second GCS life-cycle rule with Age: 275, Storage Class: Nearline, and Action: Delete.

### Unattempted

Correct answer is A as there are 2 actions needed. First archival after 30 days, which can be done by SetStorageClass action to Coldline. Second delete the data after a year, which can be done by delete action with Age 365 days. The Age condition is measured from the object's creation time.

Refer GCP documentation - Cloud Storage Lifecycle Management

Age: This condition is satisfied when an object reaches the specified age (in days). Age is measured from the object's creation time. For example, if an object's creation time is 2019/01/10 10:00 UTC and the Age condition is 10 days, then the condition is satisfied for the object on and after 2019/01/20 10:00 UTC. This is true even if the object becomes archived through object versioning sometime after its creation.

Option B is wrong as the Age needs to be set to 365 as its relative to the object creation date and not changed date.

Options C & D are wrong Nearline storage class is not an ideal storage class for archival

## 40. Question

You've been tasked with getting all of your team's public SSH keys onto all of the instances of a particular project. You've collected them all. With the fewest steps possible, what is the simplest way to get the keys deployed?

- A. Add all of the keys into a file that's formatted according to the requirements. Use the gcloud compute instances add-metadata command to upload the keys to each instance
- B. Add all of the keys into a file that's formatted according to the requirements. Use the gcloud compute project-info add-metadata command to upload the keys.
- C. Use the gcloud compute ssh command to upload all the keys
- D. Format all of the keys as needed and then, using the user interface, upload each key one at a time.

### Unattempted

Correct answer is B as project wide SSH keys can help provide users access to all the instances. The keys can be added or removed using the instance metadata.

Refer GCP documentation Project wide SSH keys

Use project-wide public SSH keys to give users general access to a Linux instance. Project-wide public SSH keys give users access to all of the Linux instances in a project that allow project-wide public SSH keys. If an instance blocks project-wide public SSH keys, a user cannot use their project-wide public SSH key to connect to the instance unless the same public SSH key is also added to instance metadata.

`gcloud compute project-info add-metadata metadata-from-file ssh-keys=[LIST_PATH]`

Option A is wrong as the gcloud compute instances provides only specific instance level access.

Option C is wrong as gcloud compute ssh is a thin wrapper around the ssh(1) command that takes care of authentication and the translation of the instance name into an IP address. It can be used to ssh to the instance.

Option D is wrong as there is no user interface to upload the keys.

## 41. Question

Your developers have been thoroughly logging everything that happens in the API. The API allows end users to request the data as JSON, XML, CSV, and XLS. Supporting all of these formats is taking a lot of developer effort. Management would like to start tracking which options are used over the next month. Without modifying the code, what's the fastest way to be able to report on this data at the end of the month?

- A. Create a custom counter logging metric that uses a regex to extract the data format into a label. At the end of the month, use the metric viewer to see the group by the label.
- B. Create a log sink that filters for rows that mention the data format. Export that to BigQuery, and run a query at the end of the month.
- C. Create a custom monitoring metric in code and edit the API code to set the metric each time the API is called.
- D. Export the logs to excel, and search for the different fields.

#### Unattempted

Correct answer is A as custom user defined log based metrics can be created on the logs already logged. These metrics can be used at the end of the month to check the stats for API call per format to gain insights.

Refer GCP documentation Stackdriver logging Log based metrics

User-defined (logs-based) metrics are created by a user on a project. They count the number of log entries that match a given filter, or keep track of particular values within the matching log entries.

Option B is wrong as the solution is possible but not the fastest as compared to log based metric.

Option C is wrong as it required a code change.

Option D is wrong as its more manual effort and not scalable.

## 42. Question

You've created a new firewall rule to allow incoming traffic on port 22, using a target tag of `dev-ssh`. You tried to connect to one of your instances, and you're still unable to connect. What steps do you need to take to resolve the problem?

- A. Run the `gcloud firewall-rules refresh` command, as they need to be reloaded
- B. Use source tags in place of the target tags.
- C. Reboot the instances for the firewall rule to take effect.

- D. Apply a network tag of dev-ssh to the instance you're trying to connect into and test again.

#### Unattempted

Correct answer is D as the firewall needs to be associated with the instance for the instance to follow the firewall rules. The association can be performed by applying the network tag dev-ssh to the instance.

Refer GCP documentation VPC Network Tags

Network tags are text attributes you can add to Compute Engine virtual machine (VM) instances. Tags allow you to make firewall rules and routes applicable to specific VM instances.

You can only add network tags to VM instances or instance templates. You cannot tag other GCP resources. You can assign network tags to new instances at creation time, or you can edit the set of assigned tags at any time later. Network tags can be edited without stopping an instance.

Option A is wrong as firewalls if associated through network tags reflect immediately and do not require any refresh.

Option B is wrong as Firewall needs to associate with target tags, which dictate the instances.

Option C is wrong as instances do not need to be rebooted and it's at the network level with no changes in the instances.

### 43. Question

You're migrating an on-premises application to Google Cloud. The application uses a component that requires a licensing server. The license server has the IP address 10.28.0.10. You want to deploy the application without making any changes to the code or configuration. How should you go about deploying the application?

- A. Create a subnet with a CIDR range of 10.28.0.0/29. Reserve a static internal IP address of 10.28.0.10. Assign the static address to the license server instance.
- B. Create a subnet with a CIDR range of 10.28.0.0/28. Reserve a static external IP address of 10.28.0.10. Assign the static address to the license server instance.
- C. Create a subnet with a CIDR range of 10.28.0.0/10. Reserve a static external IP address of 10.28.0.10. Assign the static address to the license server instance.
- D. Create a subnet with a CIDR range of 10.28.0.0/28. Reserve a static internal IP address of 10.28.0.10. Assign the static address to the license server instance.

**Unattempted**

Correct answer is D as the IP is internal it can be reserved using the static internal IP address, which blocks it and prevents it from getting allocated to other resource.

Refer GCP documentation Compute Network Addresses

In Compute Engine, each VM instance can have multiple network interfaces. Each interface can have one external IP address, one primary internal IP address, and one or more secondary internal IP addresses. Forwarding rules can have external IP addresses for external load balancing or internal addresses for internal load balancing.

Static internal IPs provide the ability to reserve internal IP addresses from the private RFC 1918 IP range configured in the subnet, then assign those reserved internal addresses to resources as needed.

Reserving an internal IP address takes that address out of the dynamic allocation pool and prevents it from being used for automatic allocations. Reserving static internal IP addresses requires specific IAM permissions so that only authorized users can reserve a static internal IP address.

Option A is wrong as the 10.28.0.0/29 CIDR provides only 8 IP addresses and would not include 10.28.0.10.

Options B & C are wrong as the IP address is RFC 1918 address and needs to be an internal static IP address.

**44. Question**

You've been running App Engine applications in a Standard Environment for a few weeks. With several successful deployments, you've just deployed a broken version, and the developers have gone home for the day. What is the fastest way to get the site back into a functioning state?

- A. Use the gcloud app deployments revert command.
- B. Use the gcloud app deployments rollback command.
- C. In GCP console, click Traffic Splitting and direct 100% of the traffic to the previous version.
- D. In GCP console, click the Rollback button on the versions page.

**Unattempted**

Correct answer is C as the best approach is the revert by the traffic to a previous deployed version.

Refer GCP documentation Migrating & Splitting Traffic

Traffic splitting distributes a percentage of traffic to versions of your application. You can split traffic to move 100% of traffic to a single version or to route percentages of traffic to multiple versions. Splitting traffic to two or more versions allows you to conduct A/B testing between your versions and provides control over the pace when rolling out features.

Options A & B are as gcloud app command does not provide rollback and revert feature

Option D is wrong as GCP console does not provide the ability to rollback.

## 45. Question

You have a 20 GB file that you need to securely share with some contractors. They need it as fast as possible. Which steps would get them the file quickly and securely?

- A. Set up a VPC with a custom subnet. Create a subnet tunnel. Upload the file to a network share. Grant the contractors temporary access.
- B. Using composite objects and parallel uploads to upload the file to Cloud Storage quickly. Then generate a signed URL and securely share it with the contractors.
- C. Upload the file to Bigtable using the bulk data import tool. Then provide the contractors with read access to the database.
- D. Upload the file to Cloud Storage. Grant the allAuthenticated users token view permissions.

### Unattempted

Correct answer is B as the composite parallel upload can help upload the file quickly to Cloud Storage.

Signed urls can be used to quickly and securely share the files with third party.

Refer GCP documentation Cloud Storage Signed URLs

Signed URLs provide a way to give time-limited read or write access to anyone in possession of the URL, regardless of whether they have a Google account

In some scenarios, you might not want to require your users to have a Google account in order to access Cloud Storage, but you still want to control access using your application-specific logic. The typical way to address this use case is to provide a signed URL to a user, which gives the user read, write, or delete access to that resource for a limited time. Anyone who knows the URL can access the resource until the URL expires. You specify the expiration time in the query string to be signed.

Option A is wrong as it is not a quick solution, but a cumbersome solution.

Option C is wrong as Bigtable is not an ideal storage for files.

Option D is wrong as All Authenticated access would provide access to anyone who is authenticated with a Google account. The special scope identifier for all Google account holders is allAuthenticatedUser

#### 46. Question

You're using a self-serve Billing Account to pay for your 2 projects. Your billing threshold is set to \$1000.00 and between the two projects you're spending roughly 50 dollars per day. It has been 18 days since you were last charged. Given the above data, when will you likely be charged next?

- A. On the first day of the next month.
- B. In 2 days when you'll hit your billing threshold.
- C. On the thirtieth day of the month.
- D. In 12 days, making it 30 days since the previous payment.

#### Unattempted

Correct answer is B as the billing is either monthly or the threshold, whichever comes first. As with average \$50 per day and 18 days passed the \$1000 threshold would hit in 2 days and so would be the billing.

Refer GCP documentation Cloud Storage Billing

Your costs are charged automatically in one of two ways, whichever comes first:

A regular monthly cycle (monthly billing)

When your account has accrued a certain amount of charges (threshold billing)

Options A & D are wrong as the billing would not be triggered in 12 days as the threshold would be hit first.

Option C is wrong as there is no such fixed date.

#### 47. Question

Your company has created a new billing account and needs to move the projects to the billing account.

What roles are needed to change the billing account? (Select two)

- A. Project Billing manager
- B. Project Owner

C. Billing Account Billing administrator

D. Billing Account Manager

E. Project Editor

### Unattempted

Correct answers are B & C as To change the billing account for an existing project, you must be an owner on the project and a billing administrator on the destination billing account.

Refer GCP documentation [Project Change Billing Account](#)

## 48. Question

You have deployed an application using Deployment manager. You want to update the deployment with minimal downtime. How can you achieve the same?

A. gcloud deployment-manager deployments create

B. gcloud deployment-manager deployments update

C. gcloud deployment-manager resources create

D. gcloud deployment-manager resources update

### Unattempted

Correct answer is B as gcloud deployment-manager deployments update can be used to update the existing deployment.

Refer GCP documentation [Deployment Manager Update Deployment](#)

After you have created a deployment, you can update it as your application or service changes. You can use Deployment Manager to update a deployment by:

Adding or removing resources from a deployment.

Updating the properties of existing resources in a deployment.

A single update can contain any combination of these changes. For example, you can make changes to the properties of existing resources and add new resources in the same request. You update your deployment by following these steps:

1. Make changes to or create a configuration file with the changes you want.

2. Optionally, pick the policies to use for your updates or use the default policies.

3. Make the update request to Deployment Manager.

gcloud deployment-manager deployments update example-deployment

Option A is wrong as gcloud deployment-manager deployments create is used to create deployment.

Options C & D are wrong as resources is not a valid parameter.

## 49. Question

You did a deployment for App Engine using gcloud app deploy. However, checking the intended project you do not find the deployment and seems the application was deployed in the wrong project. How do you find out which project the application was deployed to?

- A. Check app.yaml for the project
- B. Check application web.xml for the project
- C. Run gcloud config list to check for the project
- D. Check index.yaml for the project

### Unattempted

Correct answer is C as By default, the deploy command generates a unique ID for the version that you deploy, deploys the version to the GCP project you configured the gcloud tool to use, and routes all traffic to the new version. The project can be checked using the gcloud config list command.

Refer GCP documentation App Engine Deploying Application

gcloud app deploy app.yaml index.yaml

Optional flags:

Include the --project flag to specify an alternate GCP Console project ID to what you initialized as the default in the gcloud tool. Example: --project [YOUR\_PROJECT\_ID]

Include the -v flag to specify a version ID, otherwise one is generated for you. Example: -v [YOUR\_VERSION\_ID]

Options A, B & D are wrong as they do provide the ability to set the project.

## 50. Question

Your company has appointed external auditors for auditing the security of your setup. They want to check all the users and roles configured. What would be the best way to check the users and roles?

- A. Ask auditors to check using gcloud iam roles list command
- B. Ask auditors to check using gcloud iam service-accounts list command
- C. Ask Auditors to navigate to the IAM page and check member and roles section
- D. Ask Auditors to navigate to the IAM page section and check roles and status section

#### Unattempted

Correct answer is C as the auditor can check all the members and roles created for the project from the IAM page listing the members and roles.

Option A is wrong as the gcloud iam roles list command would only list roles.

Option B is wrong as the gcloud iam service-accounts list command would only list services accounts.

Option D is wrong as the roles menu only displays the predefined or custom roles and their status.

#### 51. Question

Your project manager wants to delegate the responsibility to manage files and buckets for Cloud Storage to his team members. Considering the principle of least privilege, which role should you assign to the team members?

- A. roles/storage.objectAdmin
- B. roles/storage.admin
- C. roles/storage.objectCreator
- D. roles/owner

#### Unattempted

Correct answer is B as roles/storage.admin would provide the team members full control of buckets and objects. When applied to an individual bucket, control applies only to the specified bucket and objects within the bucket.

Refer GCP documentation Cloud Storage IAM Roles

Options A & C are wrong as they do not provide sufficient privileges to manage buckets.

Option D is wrong as it provides more privileges than required.

## 52. Question

Your company is designing an application, which would interact with Cloud Spanner. The application should have the ability to view and edit Cloud Spanner tables. Considering the principle of least privilege, which role should you assign to the team members?

- A. roles/spanner.viewer
- B. roles/spanner.databaseUser
- C. roles/spanner.databaseReader
- D. roles/spanner.databaseAdmin

### Unattempted

Correct answer is B as roles/spanner.databaseUser is a machine only roles and provides the ability to read and write to database.

Recommended to grant at the databaselevel. A principal with this role can:

Read from and write to the Cloud Spanner database.

Execute SQL queries on the database, including DML and Partitioned DML.

View and update schema for the database.

Refer GCP documentation [Spanner IAM Roles](#)

Options A & D are wrong as they are person role and either provide more or less privileges than required.

Option C is wrong as it provides only read permissions.

## 53. Question

A Company is using Cloud SQL to host critical data. They want to enable Point In Time recovery (PIT) to be able to recover the instance to a specific point in. How should you configure the same?

- A. Create a Read replica for the instance
- B. Switch to Spanner 3 node cluster
- C. Create a Failover replica for the instance
- D. Enable Binary logging and backups for the instance

**Unattempted**

Correct answer is D as for performing Point In Time recovery for the Cloud SQL, you should enable backups and binary logging.

Refer GCP documentation [Cloud SQL Point In Time Recovery](#)

Point-in-time recovery enables you to recover an instance to a specific point in time. A point-in-time recovery always creates a new instance; you cannot perform a point-in-time recovery to an existing instance.

Before completing this task, you must have:

Binary logging and backups enabled for the instance, with continuous binary logs since the last backup before the event you want to recover from. For more information, see [Enabling binary logging](#).

A binary log file name and the position of the event you want to recover from (that event and all events that came after it will not be reflected in the new instance).

Options A & C are wrong Read and Failover replicas do not aid in Point In Recovery.

Option B is wrong as it is not required to switch to Cloud Spanner.

**54. Question**

Your organization requires that log from all applications be archived for 10 years as a part of compliance.

Which approach should you use?

- A. Configure Stackdriver Monitoring for all Projects, and export to BigQuery
- B. Configure Stackdriver Monitoring for all Projects with the default retention policies
- C. Configure Stackdriver Monitoring for all Projects, and export to Google Cloud Storage
- D. Grant the security team access to the logs in each Project

**Unattempted**

Correct answer is C as Stackdriver monitoring metrics can be exported to BigQuery or Google Cloud Storage. As the logs need to be archived, GCS is a better option.

Refer GCP documentation [Stackdriver](#)

Stackdriver Logging provides you with the ability to filter, search, and view logs from your cloud and open source application services. Allows you to define metrics based on log contents that are incorporated into

dashboards and alerts. Enables you to export logs to BigQuery, Google Cloud Storage, and Pub/Sub.

Option A is wrong as BigQuery would be a better storage option for analytics capability.

Option B is wrong as Stackdriver cannot retain data for 5 year. Refer Stackdriver data retention

Option D is wrong as project logs are maintained in Stackdriver and it has limited data retention capability.

## 55. Question

You are running an application in Google App Engine that is serving production traffic. You want to deploy a risky but necessary change to the application. It could take down your service if not properly coded. During development of the application, you realized that it can only be properly tested by live user traffic. How should you test the feature?

- A. Deploy the new application version temporarily, and then roll it back.
- B. Create a second project with the new app in isolation, and onboard users.
- C. Set up a second Google App Engine service, and then update a subset of clients to hit the new service.
- D. Deploy a new version of the application, and use traffic splitting to send a small percentage of traffic to it.

### Unattempted

Correct answer is D as deploying a new version without assigning it as the default version will not create downtime for the application. Using traffic splitting allows for easily redirecting a small amount of traffic to the new version and can also be quickly reverted without application downtime.

Refer GCP documentation [App Engine Splitting Traffic](#)

Traffic migration smoothly switches request routing, gradually moving traffic from the versions currently receiving traffic to one or more versions that you specify.

Traffic splitting distributes a percentage of traffic to versions of your application. You can split traffic to move 100% of traffic to a single version or to route percentages of traffic to multiple versions. Splitting traffic to two or more versions allows you to conduct A/B testing between your versions and provides control over the pace when rolling out features.

Option A is wrong as deploying the application version as default requires moving all traffic to the new version. This could impact all users and disable the service.

Option B is wrong as deploying a second project requires data synchronization and having an external traffic splitting solution to direct traffic to the new application. While this is possible, with Google App Engine, these manual steps are not required.

Option C is wrong as App Engine services are intended for hosting different service logic. Using different services would require manual configuration of the consumers of services to be aware of the deployment process and manage from the consumer side who is accessing which service.

## 56. Question

Using principal of least privilege and allowing for maximum automation, what steps can you take to store audit logs for long-term access and to allow access for external auditors to view? (Choose two)

- A. Generate a signed URL to the Stackdriver export destination for auditors to access.
- B. Create an account for auditors to have view access to Stackdriver Logging.
- C. Export audit logs to Cloud Storage via an export sink.
- D. Export audit logs to BigQuery via an export sink.

### Unattempted

Correct answers are A & C as Stackdriver logging allows export to Cloud Storage which can be used for long term access and exposed to external auditors using signed urls.

Refer GCP documentation Stackdriver logging export

Stackdriver Logging provides an operational datastore for logs and provides rich export capabilities. You might export your logs for several reasons, such as retaining logs for long-term storage (months or years) to meet compliance requirements or for running data analytics against the metrics extracted from the logs. Stackdriver Logging can export to Cloud Storage, BigQuery, and Cloud Pub/Sub.

Option B is wrong as Stackdriver logging does not support long term retention of logs

Option D is wrong as BigQuery can be used to export logs and retain for long term, however the access can be provided to only GCP users and not external auditors.

## 57. Question

You created an update for your application on App Engine. You want to deploy the update without impacting your users. You want to be able to roll back as quickly as possible if it fails. What should you do?

- A. Delete the current version of your application. Deploy the update using the same version identifier as the deleted version.
- B. Notify your users of an upcoming maintenance window. Deploy the update in that maintenance window.
- C. Deploy the update as the same version that is currently running.
- D. Deploy the update as a new version. Migrate traffic from the current version to the new version.

### Unattempted

Correct answer is D as the deployment can be done seamlessly by deploying a new version and migrating the traffic gradually from the old version to the new version. If any issue is encountered, the traffic can be migrated 100% to the old version.

Refer GCP documentation App Engine Migrating Traffic

Manage how much traffic is received by a version of your application by migrating or splitting traffic.

Traffic migration smoothly switches request routing, gradually moving traffic from the versions currently receiving traffic to one or more versions that you specify.

Traffic splitting distributes a percentage of traffic to versions of your application. You can split traffic to move 100% of traffic to a single version or to route percentages of traffic to multiple versions. Splitting traffic to two or more versions allows you to conduct A/B testing between your versions and provides control over the pace when rolling out features.

Options A & B are wrong as there is a downtime involved.

Option C is wrong as it would not allow an easier rollback in case of any issues.

## 58. Question

Using the principle of least privilege, your colleague Bob needs to be able to create new instances on Compute Engine in project Project A . How should you give him access without giving more permissions than is necessary?

- A. Give Bob Compute Engine Instance Admin Role for Project A.
- B. Give Bob Compute Engine Admin Role for Project A.
- C. Create a shared VPC that Bob can access Compute resources from.
- D. Give Bob Project Editor IAM role for Project A.

**Unattempted**

Correct answer is A as the access needs to be given only to create instances, the user should be given compute instance admin role, which provides the least privilege.

Refer GCP documentation Compute IAM

roles/compute.instanceAdmin.v1

roles/compute.admin

Options B & D are wrong as it gives more permission than required

Option C is wrong as shared VPC does not give permissions to create instances to the user.

**59. Question**

You need to create a new Kubernetes Cluster on Google Cloud Platform that can autoscale the number of worker nodes. What should you do?

- A. Create a cluster on Kubernetes Engine and enable autoscaling on Kubernetes Engine.
- B. Create a cluster on Kubernetes Engine and enable autoscaling on the instance group of the cluster.
- C. Configure a Compute Engine instance as a worker and add it to an unmanaged instance group. Add a load balancer to the instance group and rely on the load balancer to create additional Compute Engine instances when needed.
- D. Create Compute Engine instances for the workers and the master and install Kubernetes. Rely on Kubernetes to create additional Compute Engine instances when needed.

**Unattempted**

Correct answer is A as Kubernetes cluster provides auto scaling feature which can be enabled on the cluster engine.

Refer GCP documentation Kubernetes Cluster Autoscaler

GKE's cluster autoscaler automatically resizes clusters based on the demands of the workloads you want to run. With autoscaling enabled, GKE automatically adds a new node to your cluster if you've created new Pods that don't have enough capacity to run; conversely, if a node in your cluster is underutilized and its Pods can be run on other nodes, GKE can delete the node.

Cluster autoscaling allows you to pay only for resources that are needed at any given moment, and to automatically get additional resources when demand increases.

Option B is wrong as auto scaling is not configured on instance group.

Option C is wrong as unmanaged group cannot be scaled.

Option D is wrong as you don't manage kubernetes using compute engine.

## 60. Question

You are creating a solution to remove backup files older than 90 days from your backup Cloud Storage bucket. You want to optimize ongoing Cloud Storage spend. What should you do?

- A. Write a lifecycle management rule in XML and push it to the bucket with gsutil
- B. Write a lifecycle management rule in JSON and push it to the bucket with gsutil
- C. Schedule a cron script using gsutil ls -l gs://backups/\*\* to find and remove items older than 90 days
- D. Schedule a cron script using gsutil ls -l gs://backups/\*\* to find and remove items older than 90 days and schedule it with cron

### Unattempted

Correct answer is B as the object lifecycle in Cloud Storage can be automatically controlled using a JSON document defining the rules.

Refer GCP documentation [gsutil lifecycle](#)

Sets the lifecycle configuration on one or more buckets. The config-json-file specified on the command line should be a path to a local file containing the lifecycle configuration JSON document.

Option A is wrong as XML is not supported by the gsutil command. It works with direct REST APIs only.

Options C & D are wrong as it is quite cumbersome to list the objects, calculate the age and then delete the objects.

## 61. Question

You want to create a Google Cloud Storage regional bucket logs-archive in the Los Angeles region (us-west2). You want to use Coldline storage class to minimize costs and you want to retain files for 10 years. Which of the following commands should you run to create this bucket?

- gsutil mb -l us-west2 -s nearline --retention 10y gs://logs-archive
- gsutil mb -l los-angeles -s coldline --retention 10m gs://logs-archive

- `gsutil mb -l us-west2 -s coldline --retention 10m gs://logs-archive`
- `gsutil mb -l us-west2 -s coldline --retention 10y gs://logs-archive`

#### Unattempted

`gsutil mb -l us-west2 -s nearline retention 10y gs://logs-archive.` is not right.

This command creates a bucket that uses nearline storage class whereas we want to use Coldline storage class.

Ref: [https://cloud.google.com/storage/docs/gsutil/commands\(mb](https://cloud.google.com/storage/docs/gsutil/commands(mb)

`gsutil mb -l los-angeles -s coldline retention 10m gs://logs-archive.` is not right.

This command uses los-angeles as the location but los-angeles is not a supported region name. The region name for Los Angeles is us-west-2.

Ref: <https://cloud.google.com/storage/docs/locations>

`gsutil mb -l us-west2 -s coldline retention 10m gs://logs-archive.` is not right.

This command creates a bucket with retention set to 10 months whereas we want to retain the objects for 10 years.

Ref: [https://cloud.google.com/storage/docs/gsutil/commands\(mb](https://cloud.google.com/storage/docs/gsutil/commands(mb)

`gsutil mb -l us-west2 -s coldline retention 10y gs://logs-archive.` is the right answer.

This command correctly creates a bucket in Los Angeles, uses Coldline storage class and retains objects for 10 years.

Ref: [https://cloud.google.com/storage/docs/gsutil/commands\(mb](https://cloud.google.com/storage/docs/gsutil/commands(mb)

## 62. Question

You want to create a new role and grant it to the SME team. The new role should provide your SME team BigQuery Job User and Cloud Bigtable User roles on all projects in the organization. You want to minimize operational overhead. You want to follow Google recommended practices. How should you create the new role?

- Execute command `gcloud iam combineroles --global` to combine the 2 roles into a new custom role and grant them globally to SME team group.
- In GCP Console under IAM Roles, select both roles and combine them into a new custom role. Grant the role to the SME team group at the organization level.
- In GCP Console under IAM Roles, select both roles and combine them into a new custom role. Grant the role to the SME team group at project. Use `gcloud iam promote-role` to promote the role to all other projects and grant the role in each project to the SME team group.

- In GCP Console under IAM Roles, select both roles and combine them into a new custom role. Grant the role to the SME team group at project. Repeat this step for each project.

### Unattempted

We want to create a new role and grant it to a team. Since you want to minimize operational overhead, we need to grant it to a group so that new users who join the team just need to be added to the group and they inherit all the permissions. Also, this team needs to have the role for all projects in the organization. And since we want to minimize the operational overhead, we need to grant it at the organization level so that all current projects, as well as future projects, have the role granted to them. In GCP Console under IAM Roles, select both roles and combine them into a new custom role. Grant the role to the SME team group at project. Repeat this step for each project. is not right. ?Repeating the step for all projects is a manual, error-prone and time-consuming task. Also, if any projects were to be created in the future, we have to repeat the same process again. This increases operational overhead. In GCP Console under IAM Roles, select both roles and combine them into a new custom role. Grant the role to the SME team group at project. Use gcloud iam promote-role to promote the role to all other projects and grant the role in each project to the SME team group. is not right.

?Repeating the step for all projects is a manual, error-prone and time-consuming task. Also, if any projects were to be created in the future, we have to repeat the same process again. This increases operational overhead. Execute command gcloud iam combine-roles global to combine the 2 roles into a new custom role and grant them globally to all. is not right.

?There are several issues with this. gcloud iam command doesn't support the action combine-roles. Secondly, we don't want to grant the roles globally. We want to grant them to the SME team and no one else. In GCP Console under IAM Roles, select both roles and combine them into a new custom role. Grant the role to the SME team group at the organization level. is the right answer.

?This correctly creates the role and assigns the role to the group at the organization. When any new users join the team, the only additional task is to add them to the group. Also, when a new project is created under the organization, no additional human intervention is needed. Since the role is granted at the organization level, it automatically is granted to all the current and future projects belonging to the organization.

## 63. Question

You want to deploy a python application to an autoscaled managed instance group on Compute Engine. You want to use GCP deployment manager to do this. What is the fastest way to get the application onto the instances without introducing undue complexity?

- Include a startup script to bootstrap the python application when creating an instance template by running gcloud compute instance-templates create app-template --metadata-from-file startup-script-

url=/scripts/install\_app.sh

- Once the instance starts up, connect over SSH and install the application.
- Include a startup script to bootstrap the python application when creating an instance template by running gcloud compute instance-templates create app-template --metadata-from-file startup-script=/scripts/install\_app.sh**
- Include a startup script to bootstrap the python application when creating an instance template by running gcloud compute instance-templates create app-template --startup-script=/scripts/install\_app.sh

### Unattempted

Include a startup script to bootstrap the python application when creating instance template by running gcloud compute instance-templates create app-template startup-script=/scripts/install\_app.sh. is not right.

gcloud compute instance-templates create command does not accept a flag called startup-script. While creating compute engine images, the startup script can be provided through a special metadata key called startup-script which specifies a script that will be executed by the instances once they start running. For convenience, metadata-from-file can be used to pull the value from a file.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instance-templates/create>

Include a startup script to bootstrap the python application when creating instance template by running gcloud compute instance-templates create app-template metadata-from-file startup-script-url=/scripts/install\_app.sh. is not right.

startup-script-url is to be used when contents of the script need to be pulled from a publicly-accessible location on the web. But in this scenario, we are passing the location of the script on the filesystem which doesn't work and the command errors out.

\$ gcloud compute instance-templates create app-template metadata-from-file startup-script-url=/scripts/install\_app.sh

ERROR: (gcloud.compute.instance-templates.create) Unable to read file [/scripts/install\_app.sh]: [Errno 2]

No such file or directory: /scripts/install\_app.sh

Once the instance starts up, connect over SSH and install the application. is not right.

The managed instances group has auto-scaling enabled. If we are to connect over SSH and install the application, we have to repeat this task on all current instances and on future instances the autoscaler adds to the group. This process is manual, error-prone, time consuming and should be avoided.

Include a startup script to bootstrap the python application when creating instance template by running gcloud compute instance-templates create app-template metadata-from-file startup-script=/scripts/install\_app.sh. is the right answer.

This command correctly provides the startup script using the flag metadata-from-file and providing a valid startup-script value. When creating compute engine images, the startup script can be provided through a special metadata key called startup-script which specifies a script that will be executed by the instances

once they start running. For convenience, metadata-from-file can be used to pull the value from a file.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instance-templates/create>

#### 64. Question

You want to deploy an application on Cloud Run that processes messages from a Cloud Pub/Sub topic. You want to follow Google-recommended practices. What should you do?

- 1. Create a Cloud Function that uses a Cloud Pub/Sub trigger on that topic. 2. Call your application on Cloud Run from the Cloud Function for every message.
- 1. Grant the Pub/Sub Subscriber role to the service account used by Cloud Run. 2. Create a Cloud Pub/Sub subscription for that topic. 3. Make your application pull messages from that subscription.
- 1. Create a service account. 2. Give the Cloud Run Invoker role to that service account for your Cloud Run application. 3. Create a Cloud Pub/Sub subscription that uses that service account and uses your Cloud Run application as the push endpoint.
- 1. Deploy your application on Cloud Run on GKE with the connectivity set to Internal. 2. Create a Cloud Pub/Sub subscription for that topic. 3. In the same Google Kubernetes Engine cluster as your application, deploy a container that takes the messages and sends them to your application.

#### Unattempted

1. Create a Cloud Function that uses a Cloud Pub/Sub trigger on that topic.

2. Call your application on Cloud Run from the Cloud Function for every message. is not right.

Both Cloud functions and Cloud Run are serverless offerings from GCP and they are both capable of integrating with Cloud Pub/Sub. It is pointless to invoking Cloud Function from Cloud Run.

1. Grant the Pub/Sub Subscriber role to the service account used by Cloud Run.

2. Create a Cloud Pub/Sub subscription for that topic.

3. Make your application pull messages from that subscription. is not right.

You need to provide Cloud Run Invoker role to that service account for your Cloud Run application.

Ref: <https://cloud.google.com/run/docs/tutorials/pubsub>

1. Deploy your application on Cloud Run on GKE with the connectivity set to Internal.

2. Create a Cloud Pub/Sub subscription for that topic.

3. In the same Google Kubernetes Engine cluster as your application, deploy a container that takes the messages and sends them to your application. is not right.

Like above, you need cloud Run Invoker role on the service account.

Ref: <https://cloud.google.com/run/docs/tutorials/pubsub>

Also, our question states the application on Cloud Run processes messages from a Cloud Pub/Sub topic; whereas in this option, we are utilizing a separate container to process messages from the topic. So this doesn't satisfy our requirements.

1. Create a service account.
2. Give the Cloud Run Invoker role to that service account for your Cloud Run application.
3. Create a Cloud Pub/Sub subscription that uses that service account and uses your Cloud Run application as the push endpoint. is the right answer.

This exact process is described in

<https://cloud.google.com/run/docs/tutorials/pubsub>

You create a service account.

```
gcloud iam service-accounts create cloud-run-pubsub-invoker \
 display-name Cloud Run Pub/Sub Invoker
```

You then give the invoker service account permission to invoke your service:

```
gcloud run services add-iam-policy-binding pubsub-tutorial \
 member=serviceAccount:cloud-run-pubsub-invoker@PROJECT_ID.iam.gserviceaccount.com \
 role=roles/run.invoker
```

And finally, you create a Pub/Sub subscription with the service account:

```
gcloud pubsub subscriptions create myRunSubscription topic myRunTopic \
 push-endpoint=SERVICE-URL/ \
 push-auth-service-account=cloud-run-pubsub-invoker@PROJECT_ID.iam.gserviceaccount
```

## 65. Question

You want to ensure the boot disk of a preemptible instance is persisted for re-use. How should you provision the gcloud compute instance to ensure your requirement is met.

- gcloud compute instances create [INSTANCE\_NAME] --preemptible --no-boot-disk-auto-delete
- gcloud compute instances create [INSTANCE\_NAME] --preemptible --boot-disk-auto-delete=no
- gcloud compute instances create [INSTANCE\_NAME] --no-auto-delete
- gcloud compute instances create [INSTANCE\_NAME] --preemptible. The flag --boot-disk-auto-delete is disabled by default.

### Unattempted

gcloud compute instances create [INSTANCE\_NAME] preemptible boot-disk-auto-delete=no. is not right.

gcloud compute instances create doesn't provide a parameter called boot-disk-auto-delete. It does have a flag by the same name. boot-disk-auto-delete is enabled by default. It enables automatic deletion of boot disks when the instances are deleted. Use no-boot-disk-auto-delete to disable.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/create>

gcloud compute instances create [INSTANCE\_NAME] preemptible. boot-disk-auto-delete flag is disabled by default. is not right.

boot-disk-auto-delete is enabled by default. It enables automatic deletion of boot disks when the

instances are deleted. Use `no-boot-disk-auto-delete` to disable.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/create>

`gcloud compute instances create [INSTANCE_NAME] no-auto-delete`. is not right.

`gcloud compute instances create` doesn't provide a flag called `no-auto-delete`

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/create>

`gcloud compute instances create [INSTANCE_NAME] preemptible no-boot-disk-auto-delete`. is the right answer.

Use `no-boot-disk-auto-delete` to disable automatic deletion of boot disks when the instances are deleted. `boot-disk-auto-delete` flag is enabled by default. It enables automatic deletion of boot disks when the instances are deleted. In order to prevent automatic deletion, we have to specify `no-boot-disk-auto-delete` flag.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/create>

**Use Page numbers below to navigate to other practice tests**

Pages: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [11](#) [12](#) [13](#)

← Previous Post

Next Post →

We help you to succeed in your certification exams

We have helped over thousands of working professionals to achieve their certification goals with our practice tests.

Skillcertpro



## Quick Links

[ABOUT US](#)

[FAQ](#)

[BROWSE ALL PRACTICE TESTS](#)

[CONTACT FORM](#)

## Important Links

[REFUND POLICY](#)

[REFUND REQUEST](#)

[TERMS & CONDITIONS](#)

[PRIVACY POLICY](#)

[Privacy Policy](#)

SALE IS ON  | 12 HOURS LEFT | BUY 2 & GET ADDITIONAL 25% OFF | Use Coupon - BLACKFRIDAY



# SKILLCERTPRO

IT CERTIFICATION TRAININGS



Google Cloud / By SkillCertPro

## Practice Set 8

Your results are here!! for " Google Certified Associate Cloud Engineer Practice Test 8 "

0 of 65 questions answered correctly

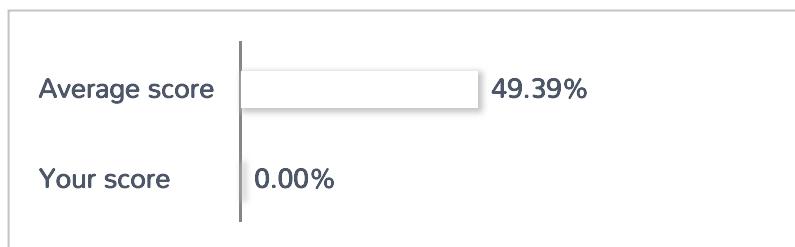
Your time: 00:00:24

Your Final Score is : 0

You have attempted : 0

Number of Correct Questions : 0 and scored 0

Number of Incorrect Questions : 0 and Negative marks 0



You can review your answers by clicking view questions.

**Important Note :** Open Reference Documentation Links in New Tab (Right Click and Open in New Tab).

[Restart Test](#)

[View Answers](#)

|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 |

|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 | 45 | 46 | 47 | 48 | 49 | 50 | 51 |
| 52 | 53 | 54 | 55 | 56 | 57 | 58 | 59 | 60 | 61 | 62 | 63 | 64 | 65 |    |    |    |



Correct   Incorrect

Review Question

Summary

## 1. Question

You want to find a list of regions and the prebuilt images offered by Google Compute Engine. Which commands should you execute to retrieve this information?

- gcloud compute regions list gcloud images list
- gcloud compute regions list gcloud compute images list
- gcloud regions list gcloud images list
- gcloud regions list gcloud compute images list

### Unattempted

gcloud regions list.

gcloud images list. is not right.

The correct command to list compute regions is gcloud compute regions list.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/regions/list>

The correct command to list compute images is gcloud compute images list.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/images/list>

gcloud compute regions list

gcloud images list. is not right.

The correct command to list compute images is gcloud compute images list.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/images/list>

gcloud regions list

gcloud compute images list. is not right.

The correct command to list compute regions is gcloud compute regions list.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/regions/list>

gcloud compute regions list

gcloud compute images list. is the right answer.

Both the commands correctly retrieve images and regions offered by Google Compute Engine

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/regions/list>

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/images/list>

## 2. Question

You want to find out when users were added to Cloud Spanner Identity Access Management (IAM) roles on your Google Cloud Platform (GCP) project. What should you do in the GCP Console?

- Open the Cloud Spanner console to review configurations.
- Open the IAM & admin console to review IAM policies for Cloud Spanner roles.
- Go to the Stackdriver Monitoring console and review information for Cloud Spanner.
- Go to the Stackdriver Logging console, review admin activity logs, and filter them for Cloud Spanner IAM roles.

### Unattempted

Go to the Stackdriver Monitoring console and review information for Cloud Spanner. is not right.

Monitoring collects metrics, events, and metadata from Google Cloud and lets you generate insights via dashboards, charts, and alerts. It can't provide information on when a role has been granted to a user.

Ref: <https://cloud.google.com/monitoring/docs>

Open the IAM & admin console to review IAM policies for Cloud Spanner roles. is not right.

You can't find the role bindings and the timestamps in the policies.

<https://cloud.google.com/iam/docs/overview>

Open the Cloud Spanner console to review configurations. is not right.

You manage cloud spanner instances in the console but you can't check when a role has been granted to a user.

Ref: <https://cloud.google.com/spanner/docs/quickstart-console>

Go to the Stackdriver Logging console, review admin activity logs, and filter them for Cloud Spanner IAM roles. is the right answer.

Admin Activity audit logs contain log entries for API calls or other administrative actions that modify the configuration or metadata of resources. For example, these logs record when users create VM instances or change Cloud Identity and Access Management permissions. Admin Activity audit logs are always written; you can't configure or disable them. There is no charge for your Admin Activity audit logs.

Ref: <https://cloud.google.com/logging/docs/audit#admin-activity>

See below a screenshot from GCP console showing this in action.

Among other things, the payload contains

{

action: ADD

```
role: roles/spanner.admin
member: user:testuser@gmail.com
}
```

### 3. Question

You want to ingest and analyze large volumes of stream data from sensors in real-time, matching the high speeds of IoT data to track normal and abnormal behavior. You want to run it through a data processing pipeline and store the results. Finally, you want to enable customers to build dashboards and drive analytics on their data in real-time. What services should you use for this task?

- Cloud Pub/Sub, Cloud Dataflow, BigQuery
- Cloud Pub/Sub, Cloud Dataflow, Cloud Dataprep
- Stackdriver, Cloud Dataflow, BigQuery
- Cloud Pub/Sub, Cloud Dataflow, Cloud Dataproc

#### Unattempted

You want to ingest large volumes of streaming data at high speeds. So you need to use Cloud Pub/Sub. Cloud Pub/Sub provides a simple and reliable staging location for your event data on its journey towards processing, storage, and analysis. Cloud Pub/Sub is serverless and you can ingest events at any scale.

Ref: <https://cloud.google.com/pubsub>

Next, you want to analyze this data. Cloud Dataflow is a fully managed streaming analytics service that minimizes latency, processing time, and cost through autoscaling and batch processing. Dataflow enables fast, simplified streaming data pipeline development with lower data latency.

Ref: <https://cloud.google.com/dataflow>

Next, you want to store these results. BigQuery is an ideal place to store these results as BigQuery supports the querying of streaming data in real-time. This assists in real-time predictive analytics.

Ref: <https://cloud.google.com/bigquery>

Therefore the correct answer is Cloud Pub/Sub, Cloud Dataflow, BigQuery

Here's more information from Google docs about the Stream analytics use case. Google recommends we use Dataflow along with Pub/Sub and BigQuery.

<https://cloud.google.com/dataflow#section-6>

Google's stream analytics makes data more organized, useful, and accessible from the instant it's generated. Built on Dataflow along with Pub/Sub and BigQuery, our streaming solution provisions the resources you need to ingest, process, and analyze fluctuating volumes of real-time data for real-time business insights. This abstracted provisioning reduces complexity and makes stream analytics

accessible to both data analysts and data engineers.

and

<https://cloud.google.com/solutions/stream-analytics>

Ingest, process, and analyze event streams in real time. Stream analytics from Google Cloud makes data more organized, useful, and accessible from the instant it's generated. Built on the autoscaling infrastructure of Pub/Sub, Dataflow, and BigQuery, our streaming solution provisions the resources you need to ingest, process, and analyze fluctuating volumes of real-time data for real-time business insights.

#### 4. Question

You want to list all the compute instances in zones us-central1-b and europe-west1-d. Which of the commands below should you run to retrieve this information?

- gcloud compute instances list --filter="zone:( us-central1-b )" and gcloud compute instances list --filter="zone:( europe-west1-d )" and combine the results.
- gcloud compute instances get --filter="zone:( us-central1-b )" and gcloud compute instances list --filter="zone:( europe-west1-d )" and combine the results.
- gcloud compute instances get --filter="zone:( us-central1-b europe-west1-d )"
- gcloud compute instances list --filter="zone:( us-central1-b europe-west1-d )"

#### Unattempted

gcloud compute instances get filter= zone:( us-central1-b europe-west1-d ) . is not right.

gcloud compute instances command does not support get action.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances>

gcloud compute instances get filter= zone:( us-central1-b ) and gcloud compute instances list filter= zone:( europe-west1-d ) and combine the results. is not right.

gcloud compute instances command does not support get action.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances>

gcloud compute instances list filter= zone:( us-central1-b ) and gcloud compute instances list filter= zone:( europe-west1-d ) and combine the results. is not right.

The first command retrieves compute instances from us-central1-b and the second command retrieves compute instances from europe-west1-d. The output from the two statements can be combined to create a full list of instances from us-central1-b and europe-west1-d, however, this is not efficient as it is a manual activity. Moreover, gcloud already provides the ability to list and filter on multiple zones in a single command.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/list>

gcloud compute instances list filter= zone:( us-central1-b europe-west1-d ) . is the right answer.

gcloud compute instances list lists Google Compute Engine instances. The output includes internal as well as external IP addresses. The filter expression filter= zone:( us-central1-b europe-west1-d ) is used to filter instances from zones us-central1-b and europe-west1-d.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/list>

Here s a sample output of the command.

\$gcloud compute instances list

| NAME                                     | ZONE          | MACHINE_TYPE  | PREEMPTIBLE | INTERNAL_IP    | EXTERNAL_IP | STATUS  |
|------------------------------------------|---------------|---------------|-------------|----------------|-------------|---------|
| gke-cluster-1-default-pool-8c599c87-16g9 | us-central1-a | n1-standard-1 | 10.128.0.8  | 35.184.212.227 |             | RUNNING |

\$gcloud compute instances list filter= zone:( us-central1-b europe-west1-d )

| NAME                                     | ZONE          | MACHINE_TYPE  | PREEMPTIBLE | INTERNAL_IP   | EXTERNAL_IP | STATUS  |
|------------------------------------------|---------------|---------------|-------------|---------------|-------------|---------|
| gke-cluster-1-default-pool-8c599c87-36xh | us-central1-b | n1-standard-1 | 10.129.0.2  | 34.68.254.220 |             | RUNNING |

## 5. Question

You want to list all the internal and external IP addresses of all compute instances. Which of the commands below should you run to retrieve this information?

- gcloud compute instances list.
- gcloud compute networks list-ip.
- gcloud compute networks list.
- gcloud compute instances list-ip.

### Unattempted

gcloud compute instances list-ip. is not right.

gcloud compute instances doesn t support the action list-ip.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/list>

gcloud compute networks list-ip. is not right.

gcloud compute networks doesn t support the action list-ip.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/networks/list>

gcloud compute networks list. is not right.

gcloud compute networks list doesn t list the IP addresses. It is used for listing Google Compute

Engine networks (i.e. VPCs)

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/networks/list>

Here's a sample output of the command.

```
$ gcloud compute networks list
```

```
NAME SUBNET_MODE BGP_ROUTING_MODE IPV4_RANGE GATEWAY_IPV4
default AUTO REGIONAL
test-vpc CUSTOM REGIONAL
```

gcloud compute instances list. is the right answer

gcloud compute instances list lists Google Compute Engine instances. The output includes internal as well as external IP addresses.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/list>

Here's a sample output of the command.

```
$ gcloud compute instances list
```

```
NAME ZONE MACHINE_TYPE PREEMPTIBLE INTERNAL_IP EXTERNAL_IP STATUS
gke-cluster-1-default-pool-8c599c87-16g9 us-central1-a n1-standard-1 10.128.0.8 35.184.212.227
RUNNING
gke-cluster-1-default-pool-8c599c87-36xh us-central1-a n1-standard-1 10.128.0.6 34.68.254.220
RUNNING
gke-cluster-1-default-pool-8c599c87-lprq us-central1-a n1-standard-1 10.128.0.7 35.224.96.151
RUNNING
```

## 6. Question

You want to migrate an application from Google App Engine Standard to Google App Engine Flex. Your application is currently serving live traffic and you want to ensure everything is working in Google App Engine Flex before migrating all traffic. You want to minimize effort and ensure the availability of service. What should you do?

- 1. Set env: flex in app.yaml 2. gcloud app deploy --version=[NEW\_VERSION] 3. Validate [NEW\_VERSION] in App Engine Flex 4. gcloud app versions migrate [NEW\_VERSION]
- 1. Set env: app-engine-flex in app.yaml 2. gcloud app deploy --no-promote --version=[NEW\_VERSION] 3. Validate [NEW\_VERSION] in App Engine Flex 4. gcloud app versions start [NEW\_VERSION]
- 1. Set env: app-engine-flex in app.yaml 2. gcloud app deploy --version=[NEW\_VERSION] 3. Validate [NEW\_VERSION] in App Engine Flex 4. gcloud app versions start [NEW\_VERSION]
- 1. Set env: flex in app.yaml 2. gcloud app deploy --no-promote --version=[NEW\_VERSION] 3. Validate [NEW\_VERSION] in App Engine Flex 4. gcloud app versions migrate [NEW\_VERSION]

Unattempted

1. Set env: flex in app.yaml
2. gcloud app deploy version=[NEW\_VERSION]
3. Validate [NEW\_VERSION] in App Engine Flex
4. gcloud app versions migrate [NEW\_VERSION]. is not right.

Executing gcloud app deploy version=[NEW\_VERSION] without no-promote would deploy the new version and immediately promote it to serve traffic. We don't want this version to receive traffic as we would like to validate the version first before sending it traffic.

Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/migrate>

1. Set env: app-engine-flex in app.yaml
2. gcloud app deploy version=[NEW\_VERSION]
3. Validate [NEW\_VERSION] in App Engine Flex
4. gcloud app versions start [NEW\_VERSION] is not right.

env: app-engine-flex is an invalid setting. The correct syntax for using the flex engine is env: flex. Also, Executing gcloud app deploy version=[NEW\_VERSION] without no-promote would deploy the new version and immediately promote it to serve traffic. We don't want this version to receive traffic as we would like to validate the version first before sending it traffic.

Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/migrate>

1. Set env: app-engine-flex in app.yaml
2. gcloud app deploy no-promote version=[NEW\_VERSION]
3. Validate [NEW\_VERSION] in App Engine Flex
4. gcloud app versions start [NEW\_VERSION] is not right.

env: app-engine-flex is an invalid setting. The correct syntax for using the flex engine is env: flex.

Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/migrate>

1. Set env: flex in app.yaml
2. gcloud app deploy no-promote version=[NEW\_VERSION]
3. Validate [NEW\_VERSION] in App Engine Flex
4. gcloud app versions migrate [NEW\_VERSION] is the right answer.

These commands together achieve the end goal while satisfying our requirements. Setting env: flex in app.yaml and executing gcloud app deploy no-promote version=[NEW\_VERSION] results in a new version deployed to flex engine. but the new version is not configured to serve traffic. We take the opportunity to review this version before migrating it to serve live traffic by running gcloud app versions migrate [NEW\_VERSION]

Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/migrate>

Ref: <https://cloud.google.com/sdk/gcloud/reference/app/deploy>

## 7. Question

You want to persist logs for 10 years to comply with regulatory requirements. You want to follow Google recommended practices. Which Google Cloud Storage class should you use?

- Archive storage class
- Nearline storage class
- Coldline storage class
- Standard storage class

#### Unattempted

In April 2019, Google introduced a new storage class Archive storage class is the lowest-cost, highly durable storage service for data archiving, online backup, and disaster recovery. Google previously recommended you use Coldline storage class but the recommendation has since been updated to Coldline Storage is ideal for data you plan to read or modify at most once a quarter. Note, however, that for data being kept entirely for backup or archiving purposes, Archive Storage is more cost-effective, as it offers the lowest storage costs.

Ref: <https://cloud.google.com/storage/docs/storage-classes#archive>

Ref: <https://cloud.google.com/storage/docs/storage-classes#coldline>

So the correct answer is Archive storage class.

## 8. Question

You want to reduce storage costs for infrequently accessed data. The data will still be accessed approximately once a month and data older than 2 years is no longer needed. What should you do to reduce storage costs? (Select 2)

- Store infrequently accessed data in a Nearline bucket.
- Set an Object Lifecycle Management policy to delete data older than 2 years.
- Set an Object Lifecycle Management policy to change the storage class to Coldline for data older than 2 years.
- Store infrequently accessed data in a Multi-Regional bucket.
- Set an Object Lifecycle Management policy to change the storage class to Archive for data older than 2 years.

#### Unattempted

Set an Object Lifecycle Management policy to change the storage class to Coldline for data older than 2 years. is not right.

Data older than 2 years is not needed so there is no point in transitioning the data to Coldline. The data needs to be deleted.

Set an Object Lifecycle Management policy to change the storage class to Archive for data older than 2 years. is not right.

Data older than 2 years is not needed so there is no point in transitioning the data to Archive. The data needs to be deleted.

Store infrequently accessed data in a Multi-Regional bucket. is not right.

While infrequently accessed data can be stored in Multi-Regional bucket, there are several other storage classes offered by Google Cloud Storage that are primarily aimed at storing infrequently accessed data and cost less. Multi-Region buckets are primarily used for achieving geo-redundancy.

Ref: <https://cloud.google.com/storage/docs/locations>

Set an Object Lifecycle Management policy to delete data older than 2 years. is the right answer.

Since you don't need data older than 2 years, deleting such data is the right approach. You can set a lifecycle policy to automatically delete objects older than 2 years. The policy is valid on current as well as future objects and doesn't need any human intervention.

Ref: <https://cloud.google.com/storage/docs/lifecycle>

Store infrequently accessed data in a Nearline bucket. is the right answer.

Nearline Storage is a low-cost, highly durable storage service for storing infrequently accessed data.

Nearline Storage is ideal for data you plan to read or modify on average once per month or less.

Ref: <https://cloud.google.com/storage/docs/storage-classes#nearline>

## 9. Question

You want to run a single caching HTTP reverse proxy on GCP for a latency-sensitive website. This specific reverse proxy consumes almost no CPU. You want to have a 30-GB in-memory cache and need an additional 2 GB of memory for the rest of the processes. You want to minimize costs. How should you run this reverse proxy?

- Create a Cloud Memorystore for Redis instance with 32-GB capacity.
- Run it on Compute Engine and choose a custom instance type with 6 vCPUs and 32 GB of memory.
- Package it in a container image, and run it on Kubernetes Engine, using n1-standard-32 instances as nodes.
- Run it on Compute Engine, choose the instance type n1-standard-1, and add an SSD persistent disk of 32 GB.

Unattempted

## Requirements

1. latency sensitive
2. 30 GB in-memory cache
3. 2 GB for rest of processes
4. Cost-effective

Run it on Compute Engine, choose the instance type n1-standard-1, and add an SSD persistent disk of 32 GB. is not right.

Fetching data from disk is slower compared to fetching from in-memory. Our requirements state we need 30GB in-memory cache for a latency-sensitive website and a compute engine with disk can't provide in-memory cache.

Run it on Compute Engine and choose a custom instance type with 6 vCPUs and 32 GB of memory. is not right.

While this option provides us with 32 GB of memory, a part of it used by the compute engine operating system as well as the reverse proxy process leaving us with less than 32GB which does not satisfy our requirements. In addition, the reverse proxy consumes almost no CPU so having 6vCPUs is a waste of resources and money.

Package it in a container image, and run it on Kubernetes Engine, using n1-standard-32 instances as nodes. is not right.

Without going into details of the feasibility of this option, let's assume for now that this option is possible. But this option is quite expensive. At the time of writing, just the compute cost for a n1-standard-32 instance is \$1.5200 per hour in the Iowa region.

Ref: <https://cloud.google.com/compute/all-pricing>

In comparison, the cost of GCP Cloud Memorystore which is \$0.023 per GB-hr which is \$0.736 for 32GB per hour. Ref: <https://cloud.google.com/memorystore>

Create a Cloud Memorystore for Redis instance with 32-GB capacity. is the right answer.

This is the only option that fits the requirements. Cloud Memorystore is a fully managed in-memory data store service for Redis built on scalable, secure, and highly available infrastructure managed by Google.

Use Memorystore to build application caches that provide sub-millisecond data access.

Ref: <https://cloud.google.com/memorystore>

Memorystore for Redis instance pricing is charged per GB-hour and you can scale as needed. You can also specify eviction (maxmemory) policies to restrict the rest of processes to 2GB or the reverse proxy to 30GB or both; you can select a suitable maxmemory policy to handle scenarios when memory is full.

Ref: [https://cloud.google.com/memorystore/docs/reference/redis-configs#maxmemory\\_policies](https://cloud.google.com/memorystore/docs/reference/redis-configs#maxmemory_policies)

## 10. Question

You want to select and configure a cost-effective solution for relational data on Google Cloud Platform. You are working with a small set of operational data in one geographic location. You need to support point-in-

time recovery. What should you do?

- Select Cloud Spanner. Set up your instance with 2 nodes.
- Select Cloud SQL (MySQL). Verify that the enable binary logging option is selected.
- Select Cloud SQL (MySQL). Select the create failover replicas option.
- Select Cloud Spanner. Set up your instance as multi-regional.

### Unattempted

#### Requirements

1. Cost effective
2. Relational Data
3. Small set of data
4. One location
5. Point in time recovery

Select Cloud Spanner. Set up your instance with 2 nodes. is not right.

Cloud spanner is a massively scalable, fully managed, relational database service for regional and global application data. Cloud spanner is expensive compared to Cloud SQL. We have a small set of data and we want to be cost-effective, so Cloud Spanner doesn't fit these requirements. Furthermore, Cloud Spanner does not offer a Point in time recovery feature.

Ref: <https://cloud.google.com/spanner>

Select Cloud Spanner. Set up your instance as multi-regional. is not right.

Cloud spanner is a massively scalable, fully managed, relational database service for regional and global application data. Cloud spanner is expensive compared to Cloud SQL. We don't have a requirement for more than one geographic location and we also have a small set of data and we want to be cost-effective, so Cloud Spanner doesn't fit these requirements. Furthermore, Cloud Spanner does not offer a Point in time recovery feature.

Ref: <https://cloud.google.com/spanner>

Select Cloud SQL (MySQL). Select the create failover replicas option. is not right.

Cloud SQL can easily handle small sets of relational data and is cost-effective compared to Cloud Spanner. But This option does not enable point in time recovery so our requirement to support point-in-time recovery is not met.

Ref: <https://cloud.google.com/sql/docs/mysql>

Select Cloud SQL (MySQL). Verify that the enable binary logging option is selected. is the right answer  
Cloud SQL can easily handle small sets of relational data and is cost-effective compared to Cloud Spanner. And by enabling binary logging, we can enable point-in-time recovery which fits our requirement.

You must enable binary logging to use point-in-time recovery. Point-in-time recovery helps you recover an instance to a specific point in time. For example, if an error causes a loss of data, you can recover a database to its state before the error occurred.

Ref: <https://cloud.google.com/sql/docs/mysql/backup-recovery/backups#tips-pitr>

## 11. Question

You want to select and configure a solution for storing and archiving data on Google Cloud Platform. You need to support compliance objectives for data from one geographic location. This data is archived after 30 days and needs to be accessed annually. What should you do?

- Select Multi-Regional Storage. Add a bucket lifecycle rule that archives data after 30 days to Coldline Storage.
- Select Multi-Regional Storage. Add a bucket lifecycle rule that archives data after 30 days to Nearline Storage.
- Select Regional Storage. Add a bucket lifecycle rule that archives data after 30 days to Nearline Storage.
- Select Regional Storage. Add a bucket lifecycle rule that archives data after 30 days to Coldline Storage.

### Unattempted

Our requirements are one region, archival after 30 days and data to be accessed annually.

Select Multi-Regional Storage. Add a bucket lifecycle rule that archives data after 30 days to Coldline Storage. is not right.

When used in a multi-region, Standard Storage is appropriate for storing data that is accessed around the world, such as serving website content, streaming videos, executing interactive workloads, or serving data supporting mobile and gaming applications. This is against our requirement of one region, moreover, this is expensive compared to standard Regional storage.

Ref: <https://cloud.google.com/storage/docs/storage-classes#standard>

Select Multi-Regional Storage. Add a bucket lifecycle rule that archives data after 30 days to Nearline Storage. is not right.

When used in a multi-region, Standard Storage is appropriate for storing data that is accessed around the world, such as serving website content, streaming videos, executing interactive workloads, or serving data supporting mobile and gaming applications. This is against our requirement of one region, moreover, this is expensive compared to standard Regional storage.

Ref: <https://cloud.google.com/storage/docs/storage-classes#standard>

Select Regional Storage. Add a bucket lifecycle rule that archives data after 30 days to Nearline Storage. is not right.

While selecting Regional Storage is the right choice, archiving to Nearline is not the most optimal. We have a requirement to access data annually whereas Nearline Storage is ideal for data you plan to read or modify on average once per month or less. Nearline storage is more expensive than Coldline Storage which is more suitable for our requirements.

<https://cloud.google.com/storage/docs/storage-classes#nearline>

Select Regional Storage. Add a bucket lifecycle rule that archives data after 30 days to Coldline Storage. is the right answer.

Regional Storage is the right fit for our requirements (one geographic region) and archiving to Coldline storage is the most cost-efficient solution. Coldline Storage is a very-low-cost, highly durable storage service for storing infrequently accessed data. Coldline Storage is ideal for data you plan to read or modify at most once a quarter. Since we have a requirement to access data once a quarter and want to go with the most cost-efficient option, we should select Coldline Storage.

Ref: <https://cloud.google.com/storage/docs/storage-classes#coldline>

## 12. Question

You want to send and consume Cloud Pub/Sub messages from your App Engine application. The Cloud Pub/Sub API is currently disabled. You will use a service account to authenticate your application to the API. You want to make sure your application can use Cloud Pub/Sub. What should you do?

- Rely on the automatic enablement of the Cloud Pub/Sub API when the Service Account accesses it.
- Enable the Cloud Pub/Sub API in the API Library on the GCP Console.**
- Grant the App Engine default service account the role of Cloud Pub/Sub Admin. Have your application enable the API on the first connection to Cloud Pub/Sub.
- Use the Deployment Manager to deploy your application. Rely on the automatic enablement of all APIs used by the application being deployed.

### Unattempted

Requirements

1. We need to enable Cloud Pub/Sub API
2. Get our application to use the service account.

Grant the App Engine default service account the role of Cloud Pub/Sub Admin. Have your application enable the API on the first connection to Cloud Pub/Sub. is not right.

APIs are not automatically enabled on the first connection to the service (Cloud Pub/Sub in this scenario).

APIs can be enabled through Google Cloud Console, gcloud command-line and REST API.

See <https://cloud.google.com/service-usage/docs/enable-disable> for more information.

Rely on the automatic enablement of the Cloud Pub/Sub API when the Service Account accesses it. is not right.

There is no such thing as automatic enablement of the APIs when the service (Cloud Pub/Sub in this scenario) is accessed. APIs can be enabled through Google Cloud Console, gcloud command-line and REST API. See <https://cloud.google.com/service-usage/docs/enable-disable> for more information.

Use the Deployment Manager to deploy your application. Rely on the automatic enablement of all APIs used by the application being deployed. is not right.

There is no such thing as automatic enablement of the APIs (Cloud Pub/Sub in this scenario) is accessed. APIs can be enabled through Google Cloud Console, gcloud command-line and REST API.

See <https://cloud.google.com/service-usage/docs/enable-disable> for more information.

Enable the Cloud Pub/Sub API in the API Library on the GCP Console. is the right answer.

For most operational use cases, the simplest way to enable and disable services is to use the Google Cloud Console. you need to create scripts, you can also use the gcloud command-line interface. If you need to program against the Service Usage API, we recommend that you use one of our provided client libraries

Ref: <https://cloud.google.com/service-usage/docs/enable-disable>

Secondly, after you create an App Engine application, the App Engine default service account is created and used as the identity of the App Engine service. The App Engine default service account is associated with your Cloud project and executes tasks on behalf of your apps running in App Engine. By default, the App Engine default service account has the Editor role in the project so this already has the permissions to push/pull/receive messages from Cloud Pub/Sub

### 13. Question

You want to serve files under the URL <https://www.my-new-gcp-ace-website.com/static/> from Cloud Storage. In addition, the URL <https://www.my-new-gcp-ace-website.com/app/> should be handled by a Compute Engine managed instance group (MIG). You want to follow Google recommended practices. How should you configure load balancing?

- 1. Deploy HAProxy in a second MIG and configure it to route /app/ to the first MIG and /static/ to your Cloud Storage bucket. 2. Create a network Load Balancer in front of the HAProxy MIG 3. Configure www.my-new-gcp-ace-website.com as an A record pointing to the address of the load balancer
- 1. Create a HTTPS Load Balancer in front of the MIG 2. In Cloud DNS in the my-new-gcp-ace-website.com zone, create a TXT record for \_app\_.\_routes\_.www.my-new-gcp-ace-website.com containing the address of the load balancer. 3. Create another TXT record for \_static\_.\_routes\_.www.my-new-gcp-ace-website.com containing the URL of your Cloud Storage bucket.
- 1. Configure www.my-new-gcp-ace-website.com as a CNAME pointing to storage.googleapis.com 2. Create a HTTPS Load Balancer in front of the MIG 3. IN the app folder of your Cloud Storage Bucket,

add a file called redirect containing the address of the load balancer.

1. Create a HTTPS Load Balancer  
2. Create a backend service associated with the MIG and route /app/ to the backend service  
3. Create a backend bucket associated with your Cloud Storage Bucket, and route /static/ to the backend bucket  
4. Configure www.my-new-gcp-ace-website.com as an A record pointing to the address of the load balancer.

### Unattempted

Our requirement here is to serve content from two backends while following Google recommended practices.

Let's look at each of the options

1. Configure <http://www.my-new-gcp-ace-website.com> as a CNAME pointing to storage.googleapis.com
2. Create a HTTPS Load Balancer in front of the MIG
3. In the app folder of your Cloud Storage Bucket, add a file called redirect containing the address of the load balancer. is not right.

We can create a CNAME <http://www.my-new-gcp-ace-website.com> pointing to storage.googleapis.com, however, the cloud storage bucket does not support routing requests to a load balancer based on routing information in a file in the app folder. So this option doesn't work.

1. Deploy HAProxy in a second MIG and configure it to route /app/ to the first MIG and /static/ to your Cloud Storage bucket.
2. Create a network Load Balancer in front of the HAProxy MIG
3. Configure <http://www.my-new-gcp-ace-website.com> as an A record pointing to the address of the load balancer is not right.

This could possibly work, but we want to follow Google recommended practices and why deploy and manage HAProxy when there might be some other Google product that does exactly the same with minimal configuration (there is !!)?

1. Create a HTTPS Load Balancer in front of the MIG
2. In Cloud DNS in the example.com zone, create a TXT record for \_app\_.\_routes\_.www.my-new-gcp-ace-website.com containing the address of the load balancer.
3. Create another TXT record for \_static\_.\_routes\_.www.my-new-gcp-ace-website.com containing the URL of your Cloud Storage bucket. is not right.

TXT records are used to verify the domain and TXT records can also hold any arbitrary text but the DNS providers don't use the text in these TXT records for routing.

Ref: <https://cloud.google.com/dns/records>

Ref: <https://support.google.com/cloudidentity/answer/183895?hl=en>

1. Create a HTTPS Load Balancer
2. Create a backend service associated with the MIG and route /app/ to the backend service

3. Create a backend bucket associated with your Cloud Storage Bucket, and route /static/ to the backend bucket
4. Configure <http://www.my-new-gcp-ace-website.com> as an A record pointing to the address of the load balancer. is the right answer.

Since we need to send requests to multiple backends, Cloud DNS can't alone help us. We need Cloud HTTPS Load Balancer it's URL maps (a fancy name for path-based routing) helps distribute traffic to backends based on the path information. Ref <https://cloud.google.com/load-balancing/docs/url-map> Traffic received by Cloud HTTPS Load Balancer can be configured to send all requests on /app path to the MIG group; and requests on /static/ path to the bucket.

Ref Adding MIG as backend service- [https://cloud.google.com/load-balancing/docs/backend-service#backend\\_services\\_and\\_autoscaled\\_managed\\_instance\\_groups](https://cloud.google.com/load-balancing/docs/backend-service#backend_services_and_autoscaled_managed_instance_groups).

Ref Adding a backend bucket(s) <https://cloud.google.com/load-balancing/docs/https/adding-backend-buckets-to-load-balancers>

The Load Balancer has a public IP address. But we want to instead access on <http://www.my-new-gcp-ace-website.com>, so we configure this as an A Record in our DNS provider. So this option is the right answer.

Ref: <https://cloud.google.com/dns/records>.

## 14. Question

You want to use Google Cloud Storage to host a static website on <http://www.example.com> for your staff. You created a bucket example-static-website and uploaded index.html and css files to it. You turned on static website hosting on the bucket and set up a CNAME record on <http://www.example.com> to point to c.storage.googleapis.com. You access the static website by navigating to <http://www.example.com> in the browser but your index page is not displayed. What should you do?

- In example.com zone, delete the existing CNAME record and set up an A record instead to point to c.storage.googleapis.com.
- In example.com zone, modify the CNAME record to c.storage.googleapis.com/example-static-website.
- Reload the Cloud Storage static website server to load the objects.
- Delete the existing bucket, create a new bucket with the name www.example.com and upload the html/css files.

### Unattempted

In example.com zone, modify the CNAME record to c.storage.googleapis.com/example-static-website. is not right.

CNAME records cannot contain paths. There is nothing wrong with the current CNAME record.

In example.com zone, delete the existing CNAME record and set up an A record instead to point to c.storage.googleapis.com. is not right.

A records cannot use hostnames. A records use IP Addresses.

Reload the Cloud Storage static website server to load the objects. is not right.

There is no such thing as a Cloud Storage static website server. All infrastructure that underpins the static websites is handled by Google Cloud Platform.

Delete the existing bucket, create a new bucket with the name http://www.example.com and upload the html/css files. is the right answer.

We need to create a bucket whose name matches the CNAME you created for your domain. For example, if you added a CNAME record pointing http://www.example.com to c.storage.googleapis.com., then create a bucket with the name www.example.com .A CNAME record is a type of DNS record. It directs traffic that requests a URL from your domain to the resources you want to serve, in this case, objects in your Cloud Storage buckets. For http://www.example.com, the CNAME record might contain the following information:

#### NAME TYPE DATA

<http://www.example.com> CNAME c.storage.googleapis.com.

Ref: <https://cloud.google.com/storage/docs/hosting-static-website>

## 15. Question

You want to verify the IAM users and roles assigned within a GCP project named my-project. What should you do?

- Run gcloud iam service-accounts list. Review the output section.
- Navigate to the project and then to the IAM section in the GCP Console. Review the members and roles.
- Navigate to the project and then to the Roles section in the GCP Console. Review the roles and status.
- Run gcloud iam roles list. Review the output section.

#### Unattempted

Requirements verify users (i.e. IAM members) and roles.

Run gcloud iam roles list. Review the output section. is not right.

gcloud iam roles list lists the roles but does not list the users (i.e. IAM members)

Run gcloud iam service-accounts list. Review the output section. is not right.

gcloud iam service-accounts list lists the service accounts which are users (i.e. IAM members) but it ignores other users that are not service accounts e.g. users in GSuite domain, or groups etc.

Navigate to the project and then to the Roles section in the GCP Console. Review the roles and status. is not right.

This allows us to review the roles but not users. See the screenshot below.

Navigate to the project and then to the IAM section in the GCP Console. Review the members and roles. is the right answer.

This is the only option that lets us view roles as well as users (members).

Ref: <https://cloud.google.com/iam/docs/overview>

See the screenshot below.

A member can be a Google Account (for end-users), a service account (for apps and virtual machines), a Google group, or a G Suite or Cloud Identity domain that can access a resource. The identity of a member is an email address associated with a user, service account, or Google group; or a domain name associated with G Suite or Cloud Identity domains

## 16. Question

Your company collects and stores CCTV footage videos in raw format in Google Cloud Storage. Within the first 30 days, the footage is processed regularly for detecting patterns such as threat/object/face detection and suspicious behavior detection. You want to minimize the cost of storing all the data in Google Cloud. How should you store the videos?

- Use Google Cloud Nearline Storage for the first 30 days, and use lifecycle rules to transition to Coldline Storage.
- Use Google Cloud Regional Storage for the first 30 days, and use lifecycle rules to transition to Coldline Storage.
- Use Google Cloud Regional Storage for the first 30 days, and use lifecycle rules to transition to Nearline Storage.
- Use Google Cloud Regional Storage for the first 30 days, and then move videos to Google Persistent Disk.

### Unattempted

Footage is processed regularly within the first 30 days and is rarely used after that. So we need to store the videos for the first 30 days in a storage class that supports economic retrieval (for processing) or at no cost, and then transition the videos to a cheaper storage after 30 days.

Use Google Cloud Regional Storage for the first 30 days, and use lifecycle rules to transition to Nearline Storage. is not right.

Transitioning the data to Nearline Storage is a good idea as Nearline Storage costs less than standard storage, is highly durable for storing infrequently accessed data and a better choice than Standard Storage in scenarios where slightly lower availability is an acceptable trade-off for lower at-rest storage costs. Ref: <https://cloud.google.com/storage/docs/storage-classes#nearline>

However, we do not have a requirement to access the data after 30 days; and there are storage classes that are cheaper than nearline storage, so it is not a suitable option.

Ref: <https://cloud.google.com/storage/pricing#storage-pricing>

Use Google Cloud Regional Storage for the first 30 days, and then move videos to Google Persistent Disk. is not right.

Persistent disk pricing is almost double that of standard storage class in Google Cloud Storage service.

Plus the persistent disk can only be accessed when attached to another service such as compute engine, GKE, etc making this option very expensive.

Ref: <https://cloud.google.com/storage/pricing#storage-pricing>

Ref: <https://cloud.google.com/compute/disks-image-pricing#persistentdisk>

Use Google Cloud Nearline Storage for the first 30 days, and use lifecycle rules to transition to Coldline Storage. is not right.

Nearline storage class is suitable for storing infrequently accessed data and has costs associated with retrieval. Since the footage is processed regularly within the first 30 days, data retrieval costs may far outweigh the savings made by using nearline storage over standard storage class.

Ref: <https://cloud.google.com/storage/docs/storage-classes#nearline>

Ref: <https://cloud.google.com/storage/pricing#archival-pricing>

Use Google Cloud Regional Storage for the first 30 days, and use lifecycle rules to transition to Coldline Storage. is the right answer.

We save the videos initially in Regional Storage (Standard) which does not have retrieval charges so we do not pay for accessing data within the first 30 days during which the videos are accessed frequently. We only pay for the standard storage costs. After 30 days, we transition the CCTV footage videos to Coldline storage which is a very-low-cost, highly durable storage service for storing infrequently accessed data. Coldline Storage is a better choice than Standard Storage or Nearline Storage in scenarios where slightly lower availability, a 90-day minimum storage duration, and higher costs for data access are acceptable trade-offs for lowered at-rest storage costs. Coldline storage class is cheaper than Nearline storage class.

Ref: <https://cloud.google.com/storage/docs/storage-classes#standard>

Ref: <https://cloud.google.com/storage/docs/storage-classes#coldline>

## 17. Question

Your company has a Google Cloud Platform project that uses BigQuery for data warehousing. Your data science team changes frequently and has few members. You need to allow members of this team to perform queries. You want to follow Google-recommended practices. What should you do?

- 1. Create a dedicated Google group in Cloud Identity. 2. Add each data scientist's user account to the group. 3. Assign the BigQuery jobUser role to the group.
- 1. Create a dedicated Google group in Cloud Identity. 2. Add each data scientist's user account to the group. 3. Assign the BigQuery dataViewer user role to the group.
- 1. Create an IAM entry for each data scientist's user account. 2. Assign the BigQuery jobUser role to the group.
- 1. Create an IAM entry for each data scientist's user account. 2. Assign the BigQuery dataViewer user role to the group.

#### Unattempted

1. Create an IAM entry for each data scientist's user account.

2. Assign the BigQuery dataViewer user role to the group. is not right.

dataViewer provides permissions to Read the dataset's metadata and to list tables in the dataset, and read data and metadata from the dataset's tables. But it does not provide permissions to run jobs, including queries within the project.

Ref: <https://cloud.google.com/bigquery/docs/access-control>

1. Create a dedicated Google group in Cloud Identity.

2. Add each data scientist's user account to the group.

3. Assign the BigQuery dataViewer user role to the group. is not right.

dataViewer provides permissions to Read the dataset's metadata and to list tables in the dataset, and read data and metadata from the dataset's tables. But it does not provide permissions to run jobs, including queries within the project.

Ref: <https://cloud.google.com/bigquery/docs/access-control>

1. Create an IAM entry for each data scientist's user account.

2. Assign the BigQuery jobUser role to the group. is not right.

jobUser is the right role. It provides permissions to run jobs, including queries, within the project. But given that our data science team changes frequently, we do not want to go through this lengthy provisioning and de-provisioning process. Instead, we should be using groups so that provisioning and de-provisioning is as simple as adding/removing the user to/from the group. Google Groups are a convenient way to apply an access policy to a collection of users

Ref: <https://cloud.google.com/bigquery/docs/access-control>

Ref: [https://cloud.google.com/iam/docs/overview#google\\_group](https://cloud.google.com/iam/docs/overview#google_group)

1. Create a dedicated Google group in Cloud Identity.
2. Add each data scientist's user account to the group.
3. Assign the BigQuery jobUser role to the group. is the right answer.

This is the only option that follows Google recommended practices and meets our requirements. jobUser is the right role. It provides permissions to run jobs, including queries, within the project.

And we want to use a group and grant the group all the necessary roles so that whenever a user joins or leaves, they can be provided access to run big query jobs by simply adding them to the group or removing from the group respectively. Google Groups are a convenient way to apply an access policy to a collection of users

Ref: <https://cloud.google.com/bigquery/docs/access-control>

Ref: [https://cloud.google.com/iam/docs/overview#google\\_group](https://cloud.google.com/iam/docs/overview#google_group)

## 18. Question

Your company has a large quantity of unstructured data in different file formats. You want to perform ETL transformations on the data. You need to make the data accessible on Google Cloud so it can be processed by a Dataflow job. What should you do?

- Upload the data to BigQuery using the bq command line tool.
- Upload the data to Cloud Storage using the gsutil command line tool.**
- Upload the data into Cloud SQL using the import function in the console.
- Upload the data into Cloud Spanner using the import function in the console.

### Unattempted

The key to answering this question is unstructured data .

Upload the data to BigQuery using the bq command line tool. is not right.

The bq load command is used to load data in BigQuery from a local data source i.e. local file but the data has to be in a structured format.

```
bq location=LOCATION load \
source_format=FORMAT \
PROJECT_ID:DATASET.TABLE \
PATH_TO_SOURCE \
SCHEMA
```

where

schema: a valid schema. The schema can be a local JSON file, or it can be typed inline as part of the command. You can also use the autodetect flag instead of supplying a schema definition.

Ref: <https://cloud.google.com/bigquery/docs/loading-data-local#bq>

Upload the data into Cloud SQL using the import function in the console. is not right.

Fully managed relational database service for MySQL, PostgreSQL, and SQL Server. As this is relational database, it is for structured data and not fit for unstructured data.

Ref: <https://cloud.google.com/sql>

Upload the data into Cloud Spanner using the import function in the console. is not right.

Cloud Spanner is the first scalable, enterprise-grade, globally-distributed, and strongly consistent database service built for the cloud specifically to combine the benefits of relational database structure with non-relational horizontal scale. Although Google claims Cloud Spanner is the best of the relational and non-relational worlds, it also says With Cloud Spanner, you get the best of relational database structure and non-relational database scale and performance with external strong consistency across rows, regions, and continents. . Cloud spanner is for structured data and not fit for unstructured data.

Ref: <https://cloud.google.com/spanner>

Upload the data to Cloud Storage using the gsutil command line tool. is the right answer.

Cloud storage imposes no such restrictions, you can store large quantities of unstructured data in different file formats. Cloud Storage provides globally unified, scalable, and highly durable object storage for developers and enterprises. In addition, Dataflow can query Cloud Storage filesets as described in this article Ref: <https://cloud.google.com/dataflow/docs/guides/sql/data-sources-destinations#querying-gcs-filesets>

## 19. Question

Your company has a number of GCP projects that are managed by the respective project teams. Your expenditure of all GCP projects combined has exceeded your operational expenditure budget. At a review meeting, it has been agreed that your finance team should be able to set budgets and view the current charges for all projects in the organization but not view the project resources; and your developers should be able to see the Google Cloud Platform billing charges for only their own projects as well as view resources within the project. You want to follow Google recommended practices to set up IAM roles and permissions. What should you do?

- Add the finance team to the Viewer role for the Project. Add the developers to the Security Reviewer role for each of the billing accounts.
- Add the developers and finance managers to the Viewer role for the Project.
- Add the finance team to the Billing Account Administrator role for each of the billing accounts that they need to manage. Add the developers to the Viewer role for the Project.
- Add the finance team to the default IAM Owner role. Add the developers to a custom role that allows them to see their own spend only.

### Unattempted

Add the finance team to the default IAM Owner role. Add the developers to a custom role that allows them to see their own spend only. is not right.

Granting your finance team the default IAM role provides them permissions to manage roles and permissions for a project and subsequently use that to assign them the permissions to view/edit resources in all projects. This is against our requirements. Also, you can write a custom role that lets developers view their project spend but they are missing permissions to view project resources.

Ref: [https://cloud.google.com/iam/docs/understanding-roles#primitive\\_roles](https://cloud.google.com/iam/docs/understanding-roles#primitive_roles)

Add the developers and finance managers to the Viewer role for the Project. is not right.

Granting your finance team the Project viewer role lets them view resources in all projects and doesn't let them set budgets both are against our requirements.

Ref: [https://cloud.google.com/iam/docs/understanding-roles#primitive\\_roles](https://cloud.google.com/iam/docs/understanding-roles#primitive_roles)

Add the finance team to the Viewer role on all projects. Add the developers to the Security Reviewer role for each of the billing accounts. is not right.

Granting your finance team the Project viewer role lets them view resources in all projects which is against our requirements. Also, the security Reviewer role enables the developers to view custom roles but doesn't let them view the project's costs or project resources.

Ref: [https://cloud.google.com/iam/docs/understanding-roles#primitive\\_roles](https://cloud.google.com/iam/docs/understanding-roles#primitive_roles)

Add the finance team to the Billing Account Administrator role for each of the billing accounts that they need to manage. Add the developers to the Viewer role for the Project. is the right answer.

Billing Account Administrator role is an owner role for a billing account. It provides permissions to manage payment instruments, configure billing exports, view cost information, set budgets, link and unlink projects and manage other user roles on the billing account.

Ref: <https://cloud.google.com/billing/docs/how-to/billing-access>

Project viewer role provides permissions for read-only actions that do not affect the state, such as viewing (but not modifying) existing resources or data; including viewing the billing charges for the project.

[https://cloud.google.com/iam/docs/understanding-roles#primitive\\_roles](https://cloud.google.com/iam/docs/understanding-roles#primitive_roles)

## 20. Question

Your company has a third-party single sign-on (SSO) identity provider that supports Security Assertion Markup Language (SAML) integration with service providers. Your company has users in Cloud Identity and requires them to authenticate using your company's SSO provider. What should you do?

- In Cloud Identity, set up SSO with Google as an identity provider to access custom SAML apps.
- In Cloud Identity, set up SSO with a third-party identity provider with Google as a service provider.

- Obtain OAuth 2.0 credentials, configure the user consent screen, and set up OAuth 2.0 for Mobile & Desktop Apps.
- Obtain OAuth 2.0 credentials, configure the user consent screen, and set up OAuth 2.0 for Web Server Applications.

### Unattempted

In Cloud Identity, set up SSO with Google as an identity provider to access custom SAML apps. is not right.

The question states that you want to use the company's existing Identity provider for SSO, not Google. Moreover, your users are in Cloud Identity and not in a GSuite domain so they don't have GSuite Gmail accounts and therefore can not sign in through Google.

Ref: <https://cloud.google.com/identity/solutions/enable-sso>

Obtain OAuth 2.0 credentials, configure the user consent screen, and set up OAuth 2.0 for Mobile & Desktop Apps. is not right.

OAuth 2.0 credentials are needed for an OAuth 2.0 flow, not SAML flow. See <https://oauth.net/2/> for more information about OAuth 2.0 which is quite a popular protocol for SSO. When you sign in to a 3rd party website using Facebook/Twitter/Google, it uses OAuth 2.0 behind the scenes.

Ref: <https://cloud.google.com/identity/solutions/enable-sso>

Obtain OAuth 2.0 credentials, configure the user consent screen, and set up OAuth 2.0 for Web Server Applications. is not right.

OAuth 2.0 credentials are needed for an OAuth 2.0 flow, not SAML flow. See <https://oauth.net/2/> for more information about OAuth 2.0 which is quite a popular protocol for SSO. When you sign in to a 3rd party website using Facebook/Twitter/Google, it uses OAuth 2.0 behind the scenes.

Ref: <https://cloud.google.com/identity/solutions/enable-sso>

In Cloud Identity, set up SSO with a third-party identity provider with Google as a service provider. is the right answer.

This is the only possible option. You configure applications (service providers) to accept SAML assertions from the company's existing identity provider and users in Cloud Identity can sign in to various applications through the third-party single sign-on (SSO) identity provider. It is important to note that user authentication occurs in the third-party IdP so the absence of a Gmail login is not an issue for signing in.

Ref: <https://cloud.google.com/identity/solutions/enable-sso>

If you have a third-party IdP, you can still configure SSO for third-party apps in the Cloud Identity catalog. User authentication occurs in the third-party IdP, and Cloud Identity manages the cloud apps.

To use Cloud Identity for SSO, your users need Cloud Identity accounts. They sign in through your third-party IdP or using a password on their Cloud Identity accounts.

## 21. Question

Your company has an App Engine application that needs to store stateful data in a proper storage service. Your data is non-relational data. You do not expect the database size to grow beyond 10 GB and you need to have the ability to scale down to zero to avoid unnecessary costs. Which storage service should you use?

- Cloud SQL
- Cloud Datastore**
- Cloud Bigtable
- Cloud Dataproc

#### Unattempted

Cloud SQL. is not right.

Cloud SQL is not suitable for non-relational data. Cloud SQL is a fully-managed database service that makes it easy to set up, maintain, manage, and administer your relational databases on Google Cloud Platform

Ref: <https://cloud.google.com/sql/docs>

Cloud Dataproc. is not right.

Cloud Dataproc is a fast, easy-to-use, fully managed cloud service for running Apache Spark and Apache Hadoop clusters in a simple, cost-efficient way. It is not a database.

Ref: <https://cloud.google.com/dataproc>

Cloud Bigtable. is not right.

Bigtable is a petabyte-scale, massively scalable, fully managed NoSQL database service for large analytical and operational workloads. Cloud Bigtable is overkill for our database which is just 10 GB. Also, Cloud Bigtable can't be scaled down to 0, as there is always a cost with the node, SSD/HDD storage etc.

Ref: <https://cloud.google.com/bigtable>

Cloud Datastore. is the right answer.

Cloud Datastore is a highly-scalable NoSQL database. Cloud Datastore scales seamlessly and automatically with your data, allowing applications to maintain high performance as they receive more traffic; automatically scales back when the traffic reduces.

Ref: <https://cloud.google.com/datastore/>

## 22. Question

Your company has an existing GCP organization with hundreds of projects and a billing account. Your company recently acquired another company that also has hundreds of projects and its own billing account. You would like to consolidate all GCP costs of both GCP organizations onto a single invoice and you would like to do this as soon as possible. What should you do?

- Link the acquired company's projects to your company's billing account.
- Configure the acquired company's billing account and your company's billing account to export the billing data into the same BigQuery dataset.
- Migrate the acquired company's projects into your company's GCP organization. Link the migrated projects to your company's billing account.
- Create a new GCP organization and a new billing account. Migrate the acquired company's projects and your company's projects into the new GCP organization and link the projects to the new billing account.

### Unattempted

Configure the acquired company's billing account and your company's billing account to export the billing data into the same BigQuery dataset. is not right.

Cloud Billing export to BigQuery enables you to export detailed Google Cloud billing data (such as usage and cost estimate data) automatically throughout the day to a BigQuery dataset that you specify. Then you can access your Cloud Billing data from BigQuery for detailed analysis or use a tool like Google Data Studio to visualize your data. Exporting billing data from both the GCP organizations into a single BigQuery dataset can help you have a single view of the billing information, but it doesn't result in a consolidated invoice, which is our requirement.

Migrate the acquired company's projects into your company's GCP organization. Link the migrated projects to your company's billing account. is not right.

While the result is what we need, migrating projects from the acquired company into your company's GCP organization is not straightforward and takes a lot of time. Our requirements state we would like to do this as soon as possible but this option isn't quick.

Create a new GCP organization and a new billing account. Migrate the acquired company's projects and your company's projects into the new GCP organization and link the projects to the new billing account. is not right.

While the result is what we need, migrating projects from both organizations into a new single organization is not straightforward and takes a lot of time. Our requirements state we would like to do this as soon as possible but this option isn't quick.

Link the acquired company's projects to your company's billing account. is the right answer.

This option is the quickest that lets us achieve our end requirement of having all GCP billing in a single invoice. Linking the acquired company's projects to your company's billing account can be very quick and can be scripted using gcloud.

Ref: <https://cloud.google.com/logging/docs/reference/tools/gcloud-logging>

## 23. Question

Your company has chosen to go serverless to enable developers to focus on writing code without worrying about infrastructure. You have been asked to identify a GCP Serverless service that does not limit your developers to specific runtimes. In addition, some of the applications need WebSockets support. What should you suggest?

- Cloud Run
- Cloud Run for Anthos**
- App Engine Standard
- Cloud Functions

### Unattempted

App Engine Standard. is not right.

Google App Engine Standard offers a limited number of runtimes Java, Node.js, Python, Go, PHP and Ruby; and at the same time doesn't offer support for Websockets.

Ref: <https://cloud.google.com/appengine/docs/standard>

Cloud Functions. is not right.

Like Google App Engine Standard, Cloud functions offer a limited number of runtimes Node.js, Python, Go and Java; and doesn't offer support for Websockets.

Ref: <https://cloud.google.com/blog/products/application-development/your-favorite-runtimes-now-generally-available-on-cloud-functions>

Cloud Run. is not right.

Cloud Run lets you run stateless containers in a fully managed environment. As this is container-based, we are not limited to specific runtimes. Developers can write code using their favorite languages (Go, Python, Java, C#, PHP, Ruby, Node.js, Shell, and others). However, Cloud Run does not support Websockets.

Ref: <https://cloud.google.com/run>

Cloud Run for Anthos. is the right answer.

Cloud Run for Anthos leverage Kubernetes and serverless together using Cloud Run integrated with Anthos. As this is container-based, we are not limited to specific runtimes. Developers can write code using their favorite languages (Go, Python, Java, C#, PHP, Ruby, Node.js, Shell, and others). Cloud Run for Anthos is the only serverless GCP offering that supports WebSockets.

<https://cloud.google.com/serverless-options>

### 24. Question

Your company has migrated most of the data center VMs to Google Compute Engine. The remaining VMs in the data center host legacy applications that are due to be decommissioned soon and your company has

decided to retain them in the datacenter. Due to a change in the business operational model, you need to introduce changes to one of the legacy applications to read files from Google Cloud Storage. However, your data center does not have access to the internet and your company doesn't want to invest in setting up internet access as the data center is due to be turned off soon. Your data center has a partner interconnect to GCP. You wish to route traffic from your datacenter to Google Storage through partner interconnect. What should you do?

- 1. In on-premises DNS configuration, map storage.cloud.google.com to restricted.googleapis.com, which resolves to the 199.36.153.4/30. 2. Configure Cloud Router to advertise the 199.36.153.4/30 IP address range through the Cloud VPN tunnel. 3. Add a custom static route to the VPC network to direct traffic with the destination 199.36.153.4/30 to the default internet gateway. 4. Created a Cloud DNS managed private zone for storage.cloud.google.com that maps to 199.36.153.4/30 and authorize the zone for use by VPC network
- 1. In on-premises DNS configuration, map storage.cloud.google.com to restricted.googleapis.com, which resolves to the 199.36.153.4/30. 2. Configure Cloud Router to advertise the 199.36.153.4/30 IP address range through the Cloud VPN tunnel. 3. Created a Cloud DNS managed public zone for storage.cloud.google.com that maps to 199.36.153.4/30 and authorize the zone for use by VPC network
- 1. In on-premises DNS configuration, map \*.googleapis.com to restricted.googleapis.com, which resolves to the 199.36.153.4/30. 2. Configure Cloud Router to advertise the 199.36.153.4/30 IP address range through the Cloud VPN tunnel. 3. Created a Cloud DNS managed public zone for \*.googleapis.com that maps to 199.36.153.4/30 and authorize the zone for use by VPC network
- 1. In on-premises DNS configuration, map \*.googleapis.com to restricted.googleapis.com, which resolves to the 199.36.153.4/30. 2. Configure Cloud Router to advertise the 199.36.153.4/30 IP address range through the Cloud VPN tunnel. 3. Add a custom static route to the VPC network to direct traffic with the destination 199.36.153.4/30 to the default internet gateway. 4. Created a Cloud DNS managed private zone for \*.googleapis.com that maps to 199.36.153.4/30 and authorize the zone for use by VPC network

#### Unattempted

While Google APIs are accessible on \*.googleapis.com, to restrict Private Google Access within a service perimeter to only VPC Service Controls supported Google APIs and services, hosts must send their requests to the restricted.googleapis.com domain name instead of \*.googleapis.com. The restricted.googleapis.com domain resolves to a VIP (virtual IP address) range 199.36.153.4/30. This IP address range is not announced to the Internet. If you require access to other Google APIs and services that aren't supported by VPC Service Controls, you can use 199.36.153.8/30 (private.googleapis.com). However, we recommend that you use restricted.googleapis.com, which integrates with VPC Service Controls and mitigates data exfiltration risks. In either case, VPC Service Controls service perimeters are always enforced on APIs and services that support VPC Service Controls.

Ref: <https://cloud.google.com/vpc-service-controls/docs/set-up-private-connectivity>

This rules out the two options that map storage.cloud.google.com to restricted.googleapis.com.

The main differences between the remaining two options are

1. Static route in the VPC network.
2. Public/Private zone.

According to Google's guide on setting up private connectivity, in order to configure a route to restricted.googleapis.com within the VPC, we need to create a static route whose destination is 199.36.153.4/30 and whose next hop is the default Internet gateway.

So, the right answer is

1. In on-premises DNS configuration, map \*.googleapis.com to restricted.googleapis.com, which resolves to the 199.36.153.4/30.
2. Configure Cloud Router to advertise the 199.36.153.4/30 IP address range through the Cloud VPN tunnel.
3. Add a custom static route to the VPC network to direct traffic with the destination 199.36.153.4/30 to the default internet gateway.
4. Create a Cloud DNS managed private zone for \*.googleapis.com that maps to 199.36.153.4/30 and authorize the zone for use by VPC network

Here's more information about how to set up private connectivity to Google's services through VPC.

Ref: <https://cloud.google.com/vpc/docs/private-access-options#private-vips>

In the following example, the on-premises network is connected to a VPC network through a Cloud VPN tunnel. Traffic from on-premises hosts to Google APIs travels through the tunnel to the VPC network.

After traffic reaches the VPC network, it is sent through a route that uses the default internet gateway as its next hop. The next hop allows traffic to leave the VPC network and be delivered to restricted.googleapis.com (199.36.153.4/30).

? The on-premises DNS configuration maps \*.googleapis.com requests to restricted.googleapis.com, which resolves to the 199.36.153.4/30.

? Cloud Router has been configured to advertise the 199.36.153.4/30 IP address range through the Cloud VPN tunnel by using a custom route advertisement. Traffic going to Google APIs is routed through the tunnel to the VPC network.

? A custom static route was added to the VPC network that directs traffic with the destination 199.36.153.4/30 to the default internet gateway (as the next hop). Google then routes traffic to the appropriate API or service.

If you created a Cloud DNS managed private zone for \*.googleapis.com that maps to 199.36.153.4/30 and have authorized that zone for use by your VPC network, requests to anything in the googleapis.com domain are sent to the IP addresses that are used by restricted.googleapis.com

## 25. Question

Your company hosts a number of applications in Google Cloud and requires that log messages from all applications be archived for 10 years to comply with local regulatory requirements. Which approach should you use?

- 1. Enable Stackdriver Logging API 2. Configure web applications to send logs to Stackdriver 3. Export logs to Google Cloud Storage
- Grant the security team access to the logs in each Project
- 1. Enable Stackdriver Logging API 2. Configure web applications to send logs to Stackdriver 3. Export logs to BigQuery
- 1. Enable Stackdriver Logging API 2. Configure web applications to send logs to Stackdriver

### Unattempted

Grant the security team access to the logs in each Project. is not right.

Granting the security team access to the logs in each Project doesn't guarantee log retention. If the security team is to come up with a manual process to copy all the logs files into another archival source, the ongoing operational costs can be huge.

1. Enable Stackdriver Logging API

2. Configure web applications to send logs to Stackdriver. is not right.

In Stackdriver, application logs are retained by default for just 30 days after which they are purged.

Ref: <https://cloud.google.com/logging/quotas>

While it is possible to configure a custom retention period of 10 years, storing logs in Stackdriver is very expensive compared to Cloud Storage. Stackdriver charges \$.01 per GB per month, whereas something like Cloud Storage Coldline Storage costs \$0.007 per GB per month (30% cheaper) and Cloud Storage Archive Storage costs 0.004 per GB per month (60% cheaper than Stackdriver)

Ref: <https://cloud.google.com/logging/docs/storage#pricing>

Ref: <https://cloud.google.com/storage/pricing>

The difference between the remaining two options is whether we store the logs in BigQuery or Google Cloud Storage.

1. Enable Stackdriver Logging API

2. Configure web applications to send logs to Stackdriver

3. Export logs to BigQuery. is not right.

While enabling Stackdriver Logging API and having the applications send logs to stack driver is a good start, exporting and storing logs in BigQuery is fairly expensive. In BigQuery, Active storage costs \$0.02 per GB per month and Long-term storage costs \$0.01 per GB per month. In comparison, Google Cloud Storage offers several storage classes that are significantly cheaper.

Ref: <https://cloud.google.com/bigquery/pricing>

Ref: <https://cloud.google.com/storage/pricing>

1. Enable Stackdriver Logging API
2. Configure web applications to send logs to Stackdriver
3. Export logs to Google Cloud Storage. is the right answer.

Google Cloud Storage offers several storage classes such as Nearline Storage (\$0.01 per GB per Month) Coldline Storage (\$0.007 per GB per Month) and Archive Storage (\$0.004 per GB per month) which are significantly cheaper than the storage options covered by the above options above.

Ref: <https://cloud.google.com/storage/pricing>

## 26. Question

Your company implemented BigQuery as an enterprise data warehouse. Users from multiple business units run queries on this data warehouse. However, you notice that query costs for BigQuery are very high, and you need to control costs. Which two methods should you use? (Choose two.)

- Split the users from business units to multiple projects.
- Apply a user- or project-level custom query quota for BigQuery data warehouse.**
- Create separate copies of your BigQuery data warehouse for each business unit.
- Split your BigQuery data warehouse into multiple data warehouses for each business unit.
- Change your BigQuery query model from on-demand to flat rate. Apply the appropriate number of slots to each Project.**

### Unattempted

Once your data is loaded into BigQuery, you are charged for storing it. Storage pricing is based on the amount of data stored in your tables when it is uncompressed. BigQuery doesn't charge for the query execution based on the output of the query (i.e. bytes returned) but on the number of bytes processed (also referred to as bytes read or bytes scanned) in order to arrive at the output of the query. You are charged for the number of bytes processed whether the data is stored in BigQuery or in an external data source such as Cloud Storage, Google Drive, or Cloud Bigtable. On-demand pricing is based solely on usage. You are charged for the bytes scanned even if your query itself doesn't return any data.

Ref: <https://cloud.google.com/bigquery/pricing>

Split the users from business units to multiple projects. is not right.

The bytes scanned is not expected to go down by splitting the users into multiple projects so this wouldn't reduce/control the costs.

Ref: <https://cloud.google.com/bigquery/pricing>

Split your BigQuery data warehouse into multiple data warehouses for each business unit. is not right.

The bytes scanned is not expected to go down by splitting the BigQuery warehouse into two so this wouldn't reduce/control the costs either.

Ref: <https://cloud.google.com/bigquery/pricing>

Create separate copies of your BigQuery data warehouse for each business unit. is not right.

Creating separate copies of the BigQuery data warehouse for each business unit is going to increase your costs. Not only is this expected to reduce the bytes scanned, but this is also going to increase the storage costs as we are now storing double the amount of data.

Ref: <https://cloud.google.com/bigquery/pricing>

Apply a user- or project-level custom query quota for BigQuery data warehouse. is the right answer.

BigQuery limits the maximum rate of incoming requests and enforces appropriate quotas on a per-project basis. You can set various limits to control costs such as Concurrent rate limit for interactive queries, Concurrent rate limit for interactive queries against Bigtable external data sources, Concurrent rate limit for legacy SQL queries that contain UDFs, Cross-region federated querying, Daily query size limit, etc.

<https://cloud.google.com/bigquery/quotas>

Change your BigQuery query model from on-demand to flat rate. Apply the appropriate number of slots to each Project. is the right answer.

This pricing option is best for customers who desire cost predictability. Flat-rate customers purchase dedicated resources for query processing and are not charged for individual queries. BigQuery offers flat-rate pricing for customers who prefer a stable cost for queries rather than paying the on-demand price per TB of data processed. You can choose to use flat-rate pricing using BigQuery Reservations. When you enroll in flat-rate pricing, you purchase slot commitments – dedicated query processing capacity, measured in BigQuery slots. Your queries consume this capacity, and you are not billed for bytes processed. If your capacity demands exceed your committed capacity, BigQuery will queue up slots, and you will not be charged additional fees.

Ref: [https://cloud.google.com/bigquery/pricing#flat\\_rate\\_pricing](https://cloud.google.com/bigquery/pricing#flat_rate_pricing)

## 27. Question

Your company is moving all corporate applications to Google Cloud Platform. The security team wants detailed visibility of all GCP projects in the organization. You provision the Google Cloud Resource Manager and set up yourself as the org admin. What Google Cloud Identity and Access Management (Cloud IAM) roles should you give to the security team?

Grant roles/resourcemanager.organizationViewer and roles/viewer.

Grant roles/resourcemanager.organizationViewer and roles/owner.

- Grant roles/owner, roles/networkmanagement.admin.
- Grant roles/resourcemanager.organizationAdmin and roles/browser.

### Unattempted

The security team needs detailed visibility of all GCP projects in the organization so they should be able to view all the projects in the organization as well as view all resources within these projects.

Grant roles/resourcemanager.organizationViewer and roles/owner. is not right.

roles/resourcemanager.organizationViewer role provides permissions to see the organization in the Cloud Console without having access to view all resources in the organization.

roles/owner provides permissions to manage roles and permissions for a project and all resources within the project and set up billing for a project.

Neither of the roles give the security team visibility of the projects in the organization.

Ref: <https://cloud.google.com/resource-manager/docs/access-control-org>

Ref: <https://cloud.google.com/iam/docs/understanding-roles#organization-policy-roles>

Grant roles/resourcemanager.organizationAdmin and roles/browser. is not right.

roles/resourcemanager.organizationAdmin provides access to administer all resources belonging to the organization. This doesn't follow the least privilege principle. Our security team needs detailed visibility i.e. read-only access but should not be able to administer resources..

Ref: <https://cloud.google.com/iam/docs/understanding-roles#organization-policy-roles>

Grant roles/owner, roles/networkmanagement.admin. is not right.

roles/owner provides permissions to manage roles and permissions for a project and all resources within the project and set up billing for a project.

roles/networkmanagement.admin provides full access to Cloud Network Management resources.

Neither of the roles give the security team visibility of the projects in the organization.

Ref: <https://cloud.google.com/resource-manager/docs/access-control-org>

Ref: <https://cloud.google.com/iam/docs/understanding-roles#organization-policy-roles>

Grant roles/resourcemanager.organizationViewer and roles/viewer. is the right answer.

roles/viewer provides permissions to view existing resources or data.

roles/resourcemanager.organizationViewer provides access to view an organization.

With the two roles, the security team can view the organization including all the projects and folders; as well as view all the resources within the projects.

Ref: <https://cloud.google.com/resource-manager/docs/access-control-org>

Ref: <https://cloud.google.com/iam/docs/understanding-roles#organization-policy-roles>

## 28. Question

Your company is moving from an on-premises environment to Google Cloud Platform (GCP). You have multiple development teams that use Cassandra environments as backend databases. They all need a development environment that is isolated from other Cassandra instances. You want to move to GCP quickly and with minimal support effort. What should you do?

- 1. Build an instruction guide to install Cassandra on GCP. 2. Make the instruction guide accessible to your developers.
- 1. Advise your developers to go to Cloud Marketplace. 2. Ask the developers to launch a Cassandra image for their development work.
- 1. Build a Cassandra Compute Engine instance and take a snapshot of it. 2. Use the snapshot to create instances for your developers.
- 1. Build a Cassandra Compute Engine instance and take a snapshot of it. 2. Upload the snapshot to Cloud Storage and make it accessible to your developers. 3. Build instructions to create a Compute Engine instance from the snapshot so that developers can do it themselves.

#### Unattempted

1. Build an instruction guide to install Cassandra on GCP.
2. Make the instruction guide accessible to your developers. is not right.

There is a very simple and straightforward way to deploy Cassandra as a Service, called Astra, on the Google Cloud Marketplace. You don't need to come up with an installation guide and ask your developers to do it.

Ref: <https://cloud.google.com/blog/products/databases/open-source-cassandra-now-managed-on-google-cloud>

Ref: <https://console.cloud.google.com/marketplace/details/click-to-deploy-images/cassandra?filter=price:free&filter=category:database&id=25ca0967-cd8e-419e-b554-fe32e87f04be&pli=1>

1. Build a Cassandra Compute Engine instance and take a snapshot of it.
2. Use the snapshot to create instances for your developers. is not right.

Like above, there is a very simple and straightforward way to deploy Cassandra as a Service, called Astra, on the Google Cloud Marketplace. You don't need to do this in a convoluted way.

Ref: <https://cloud.google.com/blog/products/databases/open-source-cassandra-now-managed-on-google-cloud>

Ref: <https://console.cloud.google.com/marketplace/details/click-to-deploy-images/cassandra?filter=price:free&filter=category:database&id=25ca0967-cd8e-419e-b554-fe32e87f04be&pli=1>

1. Build a Cassandra Compute Engine instance and take a snapshot of it.
2. Upload the snapshot to Cloud Storage and make it accessible to your developers.
3. Build instructions to create a Compute Engine instance from the snapshot so that developers can do it themselves. is not right.

Like above, there is a very simple and straightforward way to deploy Cassandra as a Service, called Astra, on the Google Cloud Marketplace. You don't need to do this in a convoluted way.

Ref: <https://cloud.google.com/blog/products/databases/open-source-cassandra-now-managed-on-google-cloud>

Ref: <https://console.cloud.google.com/marketplace/details/click-to-deploy-images/cassandra?filter=price:free&filter=category:database&id=25ca0967-cd8e-419e-b554-fe32e87f04be&pli=1>

1. Advise your developers to go to Cloud Marketplace.

2. Ask the developers to launch a Cassandra image for their development work. is the right answer.

You can deploy Cassandra as a Service, called Astra, on the Google Cloud Marketplace. Not only do you get a unified bill for all GCP services, your Developers can now create Cassandra clusters on Google Cloud in minutes and build applications with Cassandra as a database as a service without the operational overhead of managing Cassandra. Each instance is deployed to a separate set of VM instances (at the time of writing this, 3 x VM instance: 4 vCPUs + 26 GB memory (n1-highmem-4) + 10-GB Boot Disk) which are all isolated from the VM instances for other Cassandra deployments.

Ref: <https://cloud.google.com/blog/products/databases/open-source-cassandra-now-managed-on-google-cloud>

Ref: <https://console.cloud.google.com/marketplace/details/click-to-deploy-images/cassandra?filter=price:free&filter=category:database&id=25ca0967-cd8e-419e-b554-fe32e87f04be&pli=1>

## 29. Question

Your Company is planning to migrate all Java web applications to Google App Engine. However, you still want to continue using your on-premise database. How can you set up the app engine to communicate with your on-premise database while minimizing effort?

- Setup the application using App Engine Flexible environment with Cloud Router to connect to an on-premise database.
- Setup the application using App Engine Standard environment with Cloud VPN to connect to an on-premise database.**
- Setup the application using App Engine Standard environment with Cloud Router to connect to an on-premise database.
- Setup the application using App Engine Flexible environment with Cloud VPN to connect to an on-premise database.

### Unattempted

Setup the application using App Engine Standard environment with Cloud Router to connect to an on-premise database. is not right.

Cloud router by itself is not sufficient to connect VPC to an on-premise network. Cloud Router enables you to dynamically exchange routes between your Virtual Private Cloud (VPC) and on-premises networks

by using Border Gateway Protocol (BGP).

Ref: <https://cloud.google.com/router>

Setup the application using App Engine Flexible environment with Cloud Router to connect to an on-premise database. is not right.

Cloud router by itself is not sufficient to connect VPC to an on-premise network. Cloud Router enables you to dynamically exchange routes between your Virtual Private Cloud (VPC) and on-premises networks by using Border Gateway Protocol (BGP).

Ref: <https://cloud.google.com/router>

Setup the application using App Engine Flexible environment with Cloud VPN to connect to an on-premise database. is not right.

You need Cloud VPN to connect VPC to an on-premise network.

Ref: <https://cloud.google.com/vpn/docs/concepts/overview>

However, we don't have a requirement to run docker containers and App Engine Standard already supports the requirements of our existing web applications, we should avoid using App Engine Flexible. Converting to a container model involves effort and we want to minimize effort.

Setup the application using App Engine Standard environment with Cloud VPN to connect to an on-premise database. is the right answer.

You need Cloud VPN to connect VPC to an on-premise network.

Ref: <https://cloud.google.com/vpn/docs/concepts/overview>

And since we don't have a requirement to run docker containers, App Engine Standard already supports the requirements of our existing web applications Java runtime environment, so we should use App Engine Standard

## 30. Question

Your company owns a web application that lets users post travel stories. You began noticing errors in logs for a specific Deployment. The deployment is responsible for translating a post from one language to another. You've narrowed the issue down to a specific container named `msg-translator-22` that is throwing the errors. You are unable to reproduce the error in any other environment, and none of the other containers serving the deployment have this issue. You would like to connect to this container to figure out the root cause. What steps would allow you to run commands against the `msg-translator-22`?

- Use the `kubectl run msg-translator-22 /bin/bash` command to run a shell on that container.
- Use the `kubectl exec -it -- /bin/bash` command to run a shell on that container.
- Use the `kubectl run` command to run a shell on that container.
- Use the `kubectl exec -it msg-translator-22 -- /bin/bash` command to run a shell on that container.

**Unattempted**

Use the kubectl run command to run a shell on that container. is not right.

kubectl run creates and runs a deployment. It creates a deployment or a job to manage the created container(s). It is not possible to use kubectl run to connect to an existing container.

<https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#run>

Use the kubectl run msg-translator-22 /bin/ bash command to run a shell on that container. is not right.

kubectl run creates and runs a deployment. It creates a deployment or a job to manage the created container(s). It is not possible to use kubectl run to connect to an existing container.

<https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#run>

Use the kubectl exec -it /bin/bash command to run a shell on that container. is not right.

While kubectl exec is used to execute a command in a container, the command above doesn't quite work because we haven't passed to it the identifier of the container.

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#exec>

Use the kubectl exec -it msg-translator-22 /bin/bash command to run a shell on that container. is the right answer.

kubectl exec is used to execute a command in a container. We pass the container name msg-translator-22 so kubectl exec knows which container to connect to. And we pass the command /bin/bash to it, so it starts a shell on the container and we can then run custom commands and identify the root cause of the issue.

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#exec>

### 31. Question

Your company plans to store sensitive PII data in a cloud storage bucket. Your compliance department doesn't like encrypting sensitive PII data with Google-managed keys and has asked you to ensure the new objects uploaded to this bucket are encrypted by customer-managed encryption keys. What should you do? (Select Three)

- In the bucket advanced settings, select the Customer-supplied key and then select a Cloud KMS encryption key.
- Use gsutil with --encryption-key=[ENCRYPTION\_KEY] when uploading objects to the bucket.
- Use gsutil with -o "GSUtil:encryption\_key=[KEY\_RESOURCE]" when uploading objects to the bucket.
- In the bucket advanced settings, select the Customer-managed key and then select a Cloud KMS encryption key.
- Modify .boto configuration to include encryption\_key = [KEY\_RESOURCE] when uploading objects to bucket.

**Unattempted**

In the bucket advanced settings, select the Customer-supplied key and then select a Cloud KMS encryption key. is not right.

The customer-supplied key is not an option when selecting the encryption method in the console.

Use gsutil with `encryption-key=[ENCRYPTION_KEY]` when uploading objects to the bucket. is not right. gsutil doesn't accept the flag `encryption-key`. gsutil can be set up to use an encryption key by modifying boto configuration or by specifying a top-level `-o` flag but neither of these is included in this option.

Ref: <https://cloud.google.com/storage/docs/encryption/using-customer-managed-keys>

In the bucket advanced settings, select the Customer-managed key and then select a Cloud KMS encryption key. is the right answer.

Our compliance department wants us to use customer-managed encryption keys. We can select Customer-Managed radio and provide a cloud KMS encryption key to encrypt objects with the customer-managed key. This fit our requirements.

Use gsutil with `-o GSUtil:encryption_key=[KEY_RESOURCE]` when uploading objects to the bucket. is the right answer.

We can have gsutil use an encryption key by using the `-o` top-level flag: `-o GSUtil:encryption_key=[KEY_RESOURCE]`.

Ref: <https://cloud.google.com/storage/docs/encryption/using-customer-managed-keys#add-object-key>

Modify .boto configuration to include `encryption_key = [KEY_RESOURCE]` when uploading objects to bucket. is the right answer.

As an alternative to the `-o` top-level flag, gsutil can also use an encryption key if .boto configuration is modified to specify the encryption key.

`encryption_key = [KEY_RESOURCE]`

Ref: <https://cloud.google.com/storage/docs/encryption/using-customer-managed-keys#add-object-key>

## 32. Question

Your company plans to store sensitive PII data in a cloud storage bucket. Your compliance department has asked you to ensure the objects in this bucket are encrypted by customer-managed encryption keys. What should you do?

- In the bucket advanced settings, select Google-managed key and then select a Cloud KMS encryption key.
- Recreate the bucket to use a Customer-managed key. Encryption can only be specified at the time of bucket creation.

- In the bucket advanced settings, select Customer-supplied key and then select a Cloud KMS encryption key.
- In the bucket advanced settings, select Customer-managed key and then select a Cloud KMS encryption key.

#### Unattempted

In the bucket advanced settings, select Customer-supplied key and then select a Cloud KMS encryption key. is not right.

Customer-Supplied key is not an option when selecting the encryption method in the console. Moreover, we want to use customer managed encryption keys and not customer supplied encryption keys. This does not fit our requirements.

In the bucket advanced settings, select Google-managed key and then select a Cloud KMS encryption key. is not right.

While Google-managed key is an option when selecting the encryption method in console, we want to use customer managed encryption keys and not Google Managed encryption keys. This does not fit our requirements.

Recreate the bucket to use a Customer-managed key. Encryption can only be specified at the time of bucket creation. is not right.

Bucket encryption can be changed at any time. The bucket doesn't have to be recreated to change encryption.

Ref: <https://cloud.google.com/storage/docs/encryption/using-customer-managed-keys#add-default-key>

In the bucket advanced settings, select Customer-managed key and then select a Cloud KMS encryption key. is the right answer.

This option correctly selects the Customer-managed key and then the key to use which satisfies our requirement. See the screenshot below for reference.

Ref: <https://cloud.google.com/storage/docs/encryption/using-customer-managed-keys#add-default-key>

### 33. Question

Your company procured a license for a third-party cloud-based document signing system for the procurement team. All members of the procurement team need to sign in with the same service account. Your security team prohibits sharing service account passwords. You have been asked to recommend a solution that lets the procurement team login as the service account in the document signing system but without the team knowing the service account password. What should you do?

- Ask the third-party provider to enable SAML for the application and set the credentials to the service account credentials.

- Ask the third-party provider to enable OAuth 2.0 for the application and set the credentials to the service account credentials.
- Have a single person from the procurement team access document signing system with the service account credentials.
- Register the application as a password vaulted app and set the credentials to the service account credentials.

#### Unattempted

Ask the third-party provider to enable SAML for the application and set the credentials to always use service account credentials. is not right.

The application may or may not support SAML. Moreover, you can't set the credentials in SAML integration to always use a particular account. The authentication is carried out by IdP such as GSuite or a third-party identity provider. Since users are prohibited from sharing the service account credentials, they wouldn't be able to sign in through the IdP with the service account credentials.

Ask the third-party provider to enable OAuth 2.0 for the application and set the credentials to always use service account credentials. is not right.

The application may or may not support OAuth 2.0. Moreover, you can't set the credentials in SAML integration to always use a particular account. The authentication is carried out by IdP such as GSuite or a third-party identity provider. Since users are prohibited from sharing the service account credentials, they wouldn't be able to sign in through the IdP with the service account credentials.

Have a single person from the procurement team access document signing system with the service account credentials. is not right.

While this would prevent password reuse, it goes against our requirements and results in a single person dependency.

Register the application as a password vaulted app and set the credentials to the service account credentials. is the right answer.

As a G Suite or Cloud Identity administrator, the password vaulted apps service enables you to manage access to some of the apps that don't support federation and that are available to users on the User Dashboard. The password vaulted apps service saves login credential sets for applications and assigns those credential sets to users through group association. When a user has access to one of these applications through a group, they can sign in to the application through the user dashboard, or they can sign in directly from the specific application. This functionality is possible by leveraging Chrome or Firefox extensions/plugins. When adding an app to the password vaulted apps service, you can search and choose from the available web-based applications in the app library, or you can add a custom app. You can then manage usernames and passwords safely while providing users in your organization with quick one-click access to all of the apps they already use.

Ref: <https://support.google.com/cloudidentity/answer/9178974?hl=en>

### 34. Question

Your company publishes large files on an Apache web server that runs on a Compute Engine instance. The Apache web server is not the only application running in the project. You want to receive an email when the egress network costs for the server exceed 100 dollars for the current month as measured by Google Cloud Platform (GCP). What should you do?

- Set up a budget alert on the project with an amount of 100 dollars, a threshold of 100%, and notification type of email.
- Set up a budget alert on the billing account with an amount of 100 dollars, a threshold of 100%, and notification type of email.
- Export the billing data to BigQuery. Create a Cloud Function that uses BigQuery to sum the egress network costs of the exported billing data for the Apache web server for the current month and sends an email if it is over 100 dollars. Schedule the Cloud Function using Cloud Scheduler to run hourly.
- Use the Stackdriver Logging Agent to export the Apache web server logs to Stackdriver Logging. Create a Cloud Function that uses BigQuery to parse the HTTP response log data in Stackdriver for the current month and sends an email if the size of all HTTP responses, multiplied by current GCP egress prices, totals over 100 dollars. Schedule the Cloud Function using Cloud Scheduler to run hourly.

#### Unattempted

Set up a budget alert on the project with an amount of 100 dollars, a threshold of 100%, and notification type of email. is not right.

This budget alert is defined for the project which means it includes all costs and not just the egress network costs which goes against our requirements; and it also contains costs across all applications and not just the Compute Engine instance containing the Apache web server. While it is possible to set budget scope to include the Product (i.e. Google Compute Engine) and a label that uniquely identifies the specific compute engine instance, the option doesn't mention this.

Ref: <https://cloud.google.com/billing/docs/how-to/budgets#budget-scope>

Set up a budget alert on the billing account with an amount of 100 dollars, a threshold of 100%, and notification type of email. is not right.

Like above, but worse as this budget alert includes costs from all projects linked to the billing account. And like above, while it is possible to scope an alert down to Project/Product/Labels, the option doesn't mention this.

Ref: <https://cloud.google.com/billing/docs/how-to/budgets#budget-scope>

Use the Stackdriver Logging Agent to export the Apache web server logs to Stackdriver Logging. Create a Cloud Function that uses BigQuery to parse the HTTP response log data in Stackdriver for the current month and sends an email if the size of all HTTP responses, multiplied by current GCP egress prices,

totals over 100 dollars. Schedule the Cloud Function using Cloud Scheduler to run hourly. is not right.  
You can't arrive at the exact egress costs with this approach. You can configure apache logs to include the response object size.

Ref: <https://httpd.apache.org/docs/1.3/logs.html#common>

And you can then do what this option says to arrive at the combined size of all the responses but this is not 100% accurate as it does not include header sizes. Even if we assume the header size is insignificant compare to the large files published on apache web server, our question asks us to do this the Google way as measured by Google Cloud Platform (GCP) . GCP does not look at the response sizes in the Apache log files to determine the egress costs. The GCP egress calculator takes into consideration the source and destination (source = the region that hosts the Compute Engine instance running Apache Web Server; and the destination is the destination region of the packet). The egress cost is different for different destinations as shown in this pricing reference.

Ref: [https://cloud.google.com/vpc/network-pricing#internet\\_egress](https://cloud.google.com/vpc/network-pricing#internet_egress)

The Apache logs do not give you the destination information and without this information, you can't accurately calculate the egress costs.

Export the billing data to BigQuery. Create a Cloud Function that uses BigQuery to sum the egress network costs of the exported billing data for the Apache web server for the current month and sends an email if it is over 100 dollars. Schedule the Cloud Function using Cloud Scheduler to run hourly. is the right answer.

This is the only option that satisfies our requirement. We do it the Google way by (re)using the Billing Data that GCP uses. And we scope down the costs to just egress network costs for the apache web server. Finally, we schedule this to run hourly and send an email if the costs exceed 100 dollars.

### 35. Question

Your company recently migrated all infrastructure to Google Cloud Platform (GCP) and you want to use Google Cloud Build to build all container images. You want to store the build logs in Google Cloud Storage. You also have a requirement to push the images to Google Container Registry. You wrote a cloud build YAML configuration file with the following contents.

steps:

```
name: gcr.io/cloud-builders/docker
```

```
args: [build , -t , gcr.io/[PROJECT_ID]/[IMAGE_NAME] , .]
```

```
images: [gcr.io/[PROJECT_ID]/[IMAGE_NAME]]
```

How should you execute Cloud build to satisfy these requirements?

- Execute gcloud builds run --config=[CONFIG\_FILE\_PATH] --gcs-log-dir=[GCS\_LOG\_DIR] [SOURCE]

- Execute gcloud builds submit --config=[CONFIG\_FILE\_PATH] --gcs-log-dir=[GCS\_LOG\_DIR] [SOURCE]
- Execute gcloud builds submit --config=[CONFIG\_FILE\_PATH] [SOURCE]
- Execute gcloud builds push --config=[CONFIG\_FILE\_PATH] [SOURCE]

### Unattempted

Execute gcloud builds push config=[CONFIG\_FILE\_PATH] [SOURCE]. is not right.

gcloud builds command does not support push operation. The correct operation to build images and push them to gcr is submit.

Ref: <https://cloud.google.com/sdk/gcloud/reference/builds/submit>

Execute gcloud builds run config=[CONFIG\_FILE\_PATH] gcs-log-dir=[GCS\_LOG\_DIR] [SOURCE]. is not right.

gcloud builds command does not support run operation. The correct operation to build images and push them to gcr is submit.

Ref: <https://cloud.google.com/sdk/gcloud/reference/builds/submit>

Execute gcloud builds submit config=[CONFIG\_FILE\_PATH] [SOURCE]. is not right.

This command correctly builds the container image and pushes the image to GCR (Google Container Registry) but doesn't upload the build logs to Google Cloud Storage which is one of our requirements.

Ref: <https://cloud.google.com/sdk/gcloud/reference/builds/submit>

Ref: <https://cloud.google.com/cloud-build/docs/building/build-containers>

Execute gcloud builds submit config=[CONFIG\_FILE\_PATH] gcs-log-dir=[GCS\_LOG\_DIR] [SOURCE]. is the right answer.

This command correctly builds the container image, pushes the image to GCR (Google Container Registry) and uploads the build logs to Google Cloud Storage.

config flag specifies the YAML or JSON file to use as the build configuration file.

gcs-log-dir specifies the directory in Google Cloud Storage to hold build logs.

[SOURCE] is the location of the source to build. The location can be a directory on a local disk or a gzipped archive file (.tar.gz) in Google Cloud Storage.

Ref: <https://cloud.google.com/sdk/gcloud/reference/builds/submit>

Ref: <https://cloud.google.com/cloud-build/docs/building/build-containers>

## 36. Question

Your company runs a very successful web platform and has accumulated 3 petabytes of customer activity data in sharded MySQL database located in your datacenter. Due to storage limitations in your on-premise

data center, your company has decided to move this data to GCP. The data must be available all through the day. Your business analysts, who have experience of using a SQL Interface, have asked for a seamless transition. How should you store the data so that availability is ensured while optimizing the ease of analysis for the business analysts?

- Import data into Google Cloud SQL.
- Import flat files into Google Cloud Storage.
- Import data into Google Cloud Datastore.
- Import data into Google BigQuery.

#### Unattempted

Import data into Google Cloud SQL. is not right.

Cloud SQL is a fully-managed relational database service. It supports MySQL so the migration of data from your data center to cloud can be straightforward but Google Cloud SQL cannot handle petabyte-scale data. The current second-generation instances limit the storage to approximately 30TB.

Ref: <https://cloud.google.com/sql#overview>

Ref: <https://cloud.google.com/sql/docs/quotas>

Import flat files into Google Cloud Storage. is not right.

Cloud Storage is a service for storing objects in Google Cloud. You store objects in containers called buckets. You could export the MySQL data into files and import them into Google Cloud Storage, but it doesn't offer an SQL Interface to run queries/reports.

Ref: <https://cloud.google.com/storage/docs/introduction>

Import data into Google Cloud Datastore. is not right.

Your business analysts are already familiar with SQL Interface so we need a service that supports SQL. However, Cloud Datastore is a NoSQL document database. Cloud Datastore doesn't support SQL (it supports GQL which is similar to SQL, but not identical).

Ref: [https://cloud.google.com/datastore/docs/reference/gql\\_reference](https://cloud.google.com/datastore/docs/reference/gql_reference)

Ref: <https://cloud.google.com/datastore/docs/concepts/overview>

Import data into Google BigQuery. is the right answer.

Bigquery is a petabyte-scale serverless, highly scalable, and cost-effective cloud data warehouse that offers blazing-fast speeds, and with zero operational overhead. BigQuery supports a standard SQL dialect that is ANSI:2011 compliant, which reduces the impact and enables a seamless transition for your business analysts.

Ref: <https://cloud.google.com/bigquery>

#### 37. Question

Your company set up a complex organizational structure on Google Could Platform. The structure includes hundreds of folders and projects. Only a few team members should be able to view the hierarchical structure. You need to assign minimum permissions to these team members and you want to follow Google recommended practices. What should you do?

- Add the users to roles/browser role.
- Add the users to roles/iam.roleViewer role.
- Add the users to a group and add this group to roles/browser role.
- Add the users to a group and add this group to roles/iam.roleViewer role.

#### Unattempted

Add the users to roles/iam.roleViewer role. is not right.

roles/iam.roleViewer provides read access to all custom roles in the project and doesn't satisfy our requirement of viewing organization hierarchy.

Ref: <https://cloud.google.com/iam/docs/understanding-roles>

Add the users to a group and add this group to roles/iam.roleViewer role. is not right.

roles/iam.roleViewer provides read access to all custom roles in the project and doesn't satisfy our requirement of viewing organization hierarchy.

Ref: <https://cloud.google.com/iam/docs/understanding-roles>

Add the users to roles/browser role. is not right.

roles/browser provides read access to browse the hierarchy for a project, including the folder, organization, and Cloud IAM policy. This role doesn't include permission to view resources in the project. Although this is the role we require, you want to follow Google recommended practices which means we should instead add a group to the role and add users to the group instead of granting the role individually to users.

Ref: <https://cloud.google.com/iam/docs/understanding-roles>

Add the users to a group and add this group to roles/browser role. is the right answer.

roles/browser Read access to browse the hierarchy for a project, including the folder, organization, and Cloud IAM policy. This role doesn't include permission to view resources in the project. And you follow Google recommended practices by adding users to the group and group to the role. Groups are a convenient way to apply an access policy to a collection of users. You can grant and change access controls for a whole group at once instead of granting or changing access controls one at a time for individual users or service accounts. You can also easily add members to and remove members from a Google group instead of updating a Cloud IAM policy to add or remove users.

Ref: <https://cloud.google.com/iam/docs/understanding-roles>

Ref: <https://cloud.google.com/iam/docs/overview>

### 38. Question

Your company stores customer PII data in Cloud Storage buckets. A subset of this data is regularly imported into a BigQuery dataset to carry out analytics. You want to make sure the access to this bucket is strictly controlled. Your analytics team needs read access on the bucket so that they can import data in BigQuery. Your operations team needs read/write access to both the bucket and BigQuery dataset to add Customer PII data of new customers on an ongoing basis. Your Data Vigilance officers need Administrator access to the Storage bucket and BigQuery dataset. You want to follow Google recommended practices. What should you do?

- Create 3 custom IAM roles with appropriate permissions for the access levels needed for Cloud Storage and BigQuery. Add your users to the appropriate roles.
- At the Project level, add your Data Vigilance officers user accounts to the Owner role, add your operations team user accounts to the Editor role, and add your analytics team user accounts to the Viewer role.
- At the Organization level, add your Data Vigilance officers user accounts to the Owner role, add your operations team user accounts to the Editor role, and add your analytics team user accounts to the Viewer role.
- Use the appropriate predefined IAM roles for each of the access levels needed for Cloud Storage and BigQuery. Add your users to those roles for each of the services.

#### Unattempted

At the Organization level, add your Data Vigilance officers user accounts to the Owner role, add your operations team user accounts to the Editor role, and add your analytics team user accounts to the Viewer role. is not right.

Google recommends we apply the security principle of least privilege, where we grant only necessary permissions to access specific resources.

Ref: <https://cloud.google.com/iam/docs/overview>

Providing these primitive roles at the organization levels grants them permissions on all resources in all projects under the organization which violates the security principle of least privilege.

Ref: <https://cloud.google.com/iam/docs/understanding-roles>

At the Project level, add your Data Vigilance officers user accounts to the Owner role, add your operations team user accounts to the Editor role, and add your analytics team user accounts to the Viewer role. is not right.

Google recommends we apply the security principle of least privilege, where we grant only necessary permissions to access specific resources.

Ref: <https://cloud.google.com/iam/docs/overview>

Providing these primitive roles at the project level grants them permissions on all resources in the project

which violates the security principle of least privilege.

Ref: <https://cloud.google.com/iam/docs/understanding-roles>

Create 3 custom IAM roles with appropriate permissions for the access levels needed for Cloud Storage and BigQuery. Add your users to the appropriate roles. is not right.

While this has the intended outcome, it is not very efficient particularly when there are predefined roles that can be used. Secondly, if Google adds/modifies permissions for these services in the future, we would have to update our roles to reflect the modifications. This results in operational overhead and increases costs.

Ref: <https://cloud.google.com/storage/docs/access-control/iam-roles#primitive-roles-intrinsic>

Ref: <https://cloud.google.com/bigquery/docs/access-control>

Use the appropriate predefined IAM roles for each of the access levels needed for Cloud Storage and BigQuery. Add your users to those roles for each of the services. is the right answer.

For Google Cloud Storage service, Google provides predefined roles roles/owner, roles/editor, roles/viewer that match the access levels we need.

Similarly, Google provides the roles roles/bigquery.dataViewer, roles/bigquery.dataOwner, roles/bigquery.admin that match the access levels we need.

We can assign these predefined IAM roles to the respective users. Should Google add/modify permissions for these services in the future, we don't need to modify the roles above as Google does this for us; and this helps future proof our solution.

Ref: <https://cloud.google.com/storage/docs/access-control/iam-roles#primitive-roles-intrinsic>

Ref: <https://cloud.google.com/bigquery/docs/access-control>

### 39. Question

Your company stores sensitive PII data in a cloud storage bucket. The objects are currently encrypted by Google-managed keys. Your compliance department has asked you to ensure all current and future objects in this bucket are encrypted by customer-managed encryption keys. You want to minimize effort. What should you do?

- 1. In the bucket advanced settings, select the Customer-managed key and then select a Cloud KMS encryption key. 2. Rewrite all existing objects using gsutil rewrite to encrypt them with the new Customer-managed key.
- 1. In the bucket advanced settings, select the customer-supplied key and then select a Cloud KMS encryption key. 2. Delete all existing objects and upload them again so they use the new customer-supplied key for encryption.
- 1. Rewrite all existing objects using gsutil rewrite to encrypt them with the new Customer-managed key. 2. In the bucket advanced settings, select the Customer-managed key and then select a Cloud KMS encryption key.

1. In the bucket advanced settings, select the Customer-managed key and then select a Cloud KMS encryption key. 2. Existing objects encrypted by Google-managed keys can still be decrypted by the new Customer-managed key.

### Unattempted

1. In the bucket advanced settings, select the Customer-managed key and then select a Cloud KMS encryption key.
2. Existing objects encrypted by Google-managed keys can still be decrypted by the new Customer-managed key. is not right.

While changing the bucket encryption to use the Customer-managed key ensures all new objects use this key, existing objects are still encrypted by the Google-managed key. This doesn't satisfy our compliance requirements. Moreover, the customer managed key can't decrypt objects created by Google-managed keys.

Ref: <https://cloud.google.com/storage/docs/encryption/using-customer-managed-keys#add-default-key>

1. In the bucket advanced settings, select the customer-supplied key and then select a Cloud KMS encryption key.
2. Delete all existing objects and upload them again so they use the new customer-supplied key for encryption. is not right.

The customer-supplied key is not an option when selecting the encryption method in the console.

Moreover, we want to use customer-managed encryption keys and not customer-supplied encryption keys. This does not fit our requirements.

1. Rewrite all existing objects using gsutil rewrite to encrypt them with the new Customer-managed key.
2. In the bucket advanced settings, select the Customer-managed key and then select a Cloud KMS encryption key. is not right.

While changing the bucket encryption to use the Customer-managed key ensures all new objects use this key, rewriting existing objects before changing the bucket encryption would result in the objects being encrypted by the encryption method in use at that point which is still Google-managed.

1. In the bucket advanced settings, select the Customer-managed key and then select a Cloud KMS encryption key.
2. Rewrite all existing objects using gsutil rewrite to encrypt them with the new Customer-managed key. is the right answer.

Changing the bucket encryption to use the Customer-managed key ensures all new objects use this key.

Now that bucket encryption is changed to use the Customer-managed key, rewrite all existing objects using gsutil rewrite results in objects being encrypted by the new Customer-managed key. This is the only option that satisfies our requirements.

Ref: <https://cloud.google.com/storage/docs/encryption/using-customer-managed-keys#add-default-key>

## 40. Question

Your company uses a legacy application that still relies on the legacy LDAP protocol to authenticate. Your company plans to migrate this application to cloud and is looking for a cost effective solution while minimizing any developer effort. What should you do?

- Modify the legacy application to use SAML and ask users to sign in through Gmail.
- Modify the legacy application to use OAuth 2.0 and ask users to sign in through Gmail.
- Use secure LDAP to authenticate the legacy application and ask users to sign in through Gmail.
- Synchronize data within your LDAP server with Google Cloud Directory Sync.

### Unattempted

Modify the legacy application to use SAML and ask users to sign in through Gmail. is not right.

Modifying a legacy application to use SAML can be quite challenging. In any case, this is a time consuming and error-prone task and is very expensive.

Modify the legacy application to use OAuth 2.0 and ask users to sign in through Gmail. is not right.

Modifying a legacy application to use OAuth 2.0 can be quite challenging. In any case, this is a time consuming and error-prone task and is very expensive.

Synchronize data within your LDAP server with Google Cloud Directory Sync. is not right.

This can be done but this isn't going to help with the legacy LDAP protocol authentication unless the application is modified to work with either Cloud Identity or G Suite. And your company is looking for a cost-effective solution while minimizing developer effort so this isn't suitable.

Use secure LDAP to authenticate the legacy application and ask users to sign in through Gmail. is the right answer.

Secure LDAP enables authentication, authorization, and user/group lookups for LDAP-based apps and IT infrastructure. Secure LDAP uses the same user directory for both SaaS and LDAP-based applications, so people can use the same Cloud Identity credentials they use to log in to services like G Suite and other SaaS apps as they do to log into traditional applications. Applications and IT infrastructure that use LDAP can be simply configured to leverage Cloud Identity's secure LDAP service instead of an existing legacy identity system end-users don't have to change how they access their apps.

Ref: <https://cloud.google.com/blog/products/identity-security/cloud-identity-now-provides-access-to-traditional-apps-with-secure-ldap>

## 41. Question

Your company uses BigQuery for data warehousing. Over time, many different business units in your company have created 1000+ datasets across hundreds of projects. Your CIO wants you to examine all

datasets to find tables that contain an employee\_ssn column. You want to minimize effort in performing this task. What should you do?

- Go to Data Catalog and search for employee\_ssn in the search box.
- Write a shell script that uses the bq command line tool to loop through all the projects in your organization.
- Write a script that loops through all the projects in your organization and runs a query on INFORMATION\_SCHEMA.COLUMNS view to find the employee\_ssn column.
- Write a Cloud Dataflow job that loops through all the projects in your organization and runs a query on INFORMATION\_SCHEMA.COLUMNS view to find employee\_ssn column.

#### Unattempted

Go to Data Catalog and search for employee\_ssn in the search box. is the right answer.

Data Catalog is a fully managed and scalable metadata management service that empowers organizations to quickly discover, understand, and manage all their data. It offers a simple and easy-to-use search interface for data discovery, a flexible and powerful cataloging system for capturing both technical and business metadata, and a strong security and compliance foundation with Cloud Data Loss Prevention (DLP) and Cloud Identity and Access Management (IAM) integrations. The service automatically ingests technical metadata for BigQuery and Cloud Pub/Sub and allows customers to capture business metadata in schematized format via tags, custom APIs, and the UI, offering a simple and efficient way to catalog their data assets. You can perform a search for data assets from the Data Catalog home page in the Google Cloud Console.

See <https://cloud.google.com/data-catalog/docs/how-to/search> for example.

All other options are manual, error-prone, time-consuming, and should be avoided.

#### 42. Question

Your company uses Cloud Storage to store application backup files for disaster recovery purposes. You want to follow Google recommended practices. Which storage option should you use?

- Coldline Storage
- Multi-Regional Storage
- Regional Storage
- Nearline Storage

#### Unattempted

The ideal answer to this would have been Archive Storage but that is not one of the options.

Archive Storage is the lowest-cost, highly durable storage service for data archiving, online backup, and disaster recovery. Your data is available within milliseconds, not hours or days.

<https://cloud.google.com/storage/docs/storage-classes#archive>

In the absence of Archive Storage, the next best option for storing backups is Coldline Storage.

Coldline Storage is a very-low-cost, highly durable storage service for storing infrequently accessed data. Coldline Storage is a better choice than Standard Storage or Nearline Storage in scenarios where slightly lower availability, a 90-day minimum storage duration, and higher costs for data access are acceptable trade-offs for lowered at-rest storage costs. Coldline Storage is ideal for data you plan to read or modify at most once a quarter.

Ref: <https://cloud.google.com/storage/docs/storage-classes#coldline>

Although Nearline, Regional and Multi-Regional can also be used to store the backups, they are expensive in comparison and Google recommends we use Coldline for backups.

More information about Nearline: <https://cloud.google.com/storage/docs/storage-classes#nearline>

More information about Standard/Regional: <https://cloud.google.com/storage/docs/storage-classes#standard>

More information about Standard/Multi-Regional: <https://cloud.google.com/storage/docs/storage-classes#standard>

### 43. Question

Your company wants to move 200 TB of your website clickstream logs from your on-premise data center to Google Cloud Platform. These logs need to be retained in GCP for compliance requirements. Your business analysts also want to run analytics on these logs to understand user click behavior on your website. Which of the below would enable you to meet these requirements? (Select Two)

- Load logs into Google BigQuery.
- Load logs into Google Cloud SQL.
- Import logs into Google Stackdriver.
- Insert logs into Google Cloud Bigtable.
- Upload log files into Google Cloud Storage.

#### Unattempted

Load logs into Google Cloud SQL. is not right.

Cloud SQL is a fully-managed relational database service. Storing logs in Google Cloud SQL is very expensive. Cloud SQL doesn't help us with analytics. Moreover, Google Cloud Platform offers several storage classes in Google Cloud Storage that are more apt for storing logs at a much cheaper cost.

Ref: <https://cloud.google.com/sql/docs>

Ref: <https://cloud.google.com/sql/pricing#sql-storage-networking-prices>

Ref: <https://cloud.google.com/storage/pricing>

Import logs into Google Stackdriver. is not right.

You can push custom logs to Stackdriver and set custom retention periods to store the logs for longer durations. However, Stackdriver doesn't help us with analytics. You could create a sink and export data into Cloud BigQuery for analytics but that is more work. Moreover, Google Cloud Platform offers several storage classes in Google Cloud Storage that are more apt for storing logs at a much cheaper cost.

Ref: <https://cloud.google.com/logging>

Ref: <https://cloud.google.com/storage/pricing>

Insert logs into Google Cloud Bigtable. is not right.

Cloud Bigtable is a petabyte-scale, fully managed NoSQL database service for large analytical and operational workloads. Cloud Bigtable can not run analytics by itself. (But when combined with other services to ingest, process, analyze and present, it can help drive analytics) see the diagram below. So this option is not right.

Ref: <https://cloud.google.com/bigtable/>

Upload log files into Google Cloud Storage. is the right answer.

Google Cloud Platform offers several storage classes in Google Cloud Storage that are suitable for storing/archiving logs at a reasonable cost.

GCP recommends you use

1. Standard storage class if you need to access objects frequently
2. Nearline storage class if you access infrequently i.e. once a month
3. Coldline storage class if you access even less frequently e.g. once a quarter
4. Archive storage for logs archival.

Ref: <https://cloud.google.com/storage/docs/storage-classes>

Load logs into Google BigQuery. is the right answer.

By loading logs into Google BigQuery, you can securely run and share analytical insights in your organization with a few clicks. BigQuery's high-speed streaming insertion API provides a powerful foundation for real-time analytics, making your latest business data immediately available for analysis.

Ref: <https://cloud.google.com/bigquery#marketing-analytics>

#### 44. Question

Your company wants to move all documents from a secure internal NAS drive to a Google Cloud Storage (GCS) bucket. The data contains personally identifiable information (PII) and sensitive customer information. Your company tax auditors need access to some of these documents. What security strategy would you recommend on GCS?

- Grant no Google Cloud Identity and Access Management (Cloud IAM) roles to users, and use granular ACLs on the bucket.
- Grant IAM read-only access to users, and use default ACLs on the bucket.
- Create randomized bucket and object names. Enable public access, but only provide specific file URLs to people who do not have Google accounts and need access.
- Use signed URLs to generate time-bound access to objects.

### Unattempted

Use signed URLs to generate time-bound access to objects. is not right.

When dealing with sensitive customer information such as PII, using signed URLs is not a great idea as anyone with access to the URL has access to PII data. Signed URLs provide time-limited resource access to anyone in possession of the URL, regardless of whether they have a Google account. With PII Data, we want to be sure who has access and signed URLs don't guarantee that.

Ref: <https://cloud.google.com/storage/docs/access-control/signed-urls>

Grant IAM read-only access to users, and use default ACLs on the bucket. is not right.

We do not need to grant all IAM read-only access to this sensitive data. Just the users who need access to sensitive/PII data should be provided access to this data.

Create randomized bucket and object names. Enable public access, but only provide specific file URLs to people who do not have Google accounts and need access. is not right.

Enabling public access to the buckets and objects makes them visible to everyone. There are a number of scanning tools out in the market with the sole purpose of identifying buckets/objects that can be reached publicly. Should one of these tools be used by a bad actor to find out our public bucket/objects, it would result in a security breach.

Grant no Google Cloud Identity and Access Management (Cloud IAM) roles to users, and use granular ACLs on the bucket. is the right answer.

We start with no explicit access to any of the IAM users, and the bucket ACLs can then control which users can access what objects. This is the most secure way of ensuring just the people who require access to the bucket are provided with access. We block everyone from accessing the bucket and explicitly provided access to specific users through ACLs.

## 45. Question

Your company's infrastructure is on-premises, but all machines are running at maximum capacity. You want to burst to Google Cloud. The workloads on Google Cloud must be able to directly communicate to the workloads on-premises using a private IP range. What should you do?

- In Google Cloud, configure the VPC as a host for Shared VPC.

- In Google Cloud, configure the VPC for VPC Network Peering.
- Create bastion hosts both in your on-premises environment and on Google Cloud. Configure both as proxy servers using their public IP addresses.
- Set up Cloud VPN between the infrastructure on-premises and Google Cloud.

### Unattempted

In Google Cloud, configure the VPC as a host for Shared VPC. is not right.

Shared VPC allows an organization to connect resources from multiple projects to a common Virtual Private Cloud (VPC) network so that they can communicate with each other securely and efficiently using internal IPs from that network. When you use Shared VPC, you designate a project as a host project and attach one or more other service projects to it. This in no way helps us connect to our on-premises network.

Ref: <https://cloud.google.com/vpc/docs/shared-vpc>

In Google Cloud, configure the VPC for VPC Network Peering. is not right.

Google Cloud VPC Network Peering allows internal IP address connectivity across two Virtual Private Cloud (VPC) networks regardless of whether they belong to the same project or the same organization. VPC Network Peering enables you to connect VPC networks so that workloads in different VPC networks can communicate internally. Traffic stays within Google's network and doesn't traverse the public internet. This doesn't help us connect to our on-premises network.

Ref: <https://cloud.google.com/vpc/docs/vpc-peering>

Create bastion hosts both in your on-premises environment and on Google Cloud. Configure both as proxy servers using their public IP addresses. is not right.

Bastion hosts provide an external facing point of entry into a network containing private network instances. Bastion hosts are primarily for end users so they can connect to an instance that does not have an external IP address through a bastion host.

Ref: <https://cloud.google.com/compute/docs/instances/connecting-advanced>

Set up Cloud VPN between the infrastructure on-premises and Google Cloud. is the right answer.

Cloud VPN securely connects your on-premises network to your Google Cloud (GCP) Virtual Private Cloud (VPC) network through an IPsec VPN connection.

Ref: <https://cloud.google.com/vpn/docs/concepts/overview>

## 46. Question

Your company's test suite is a custom C++ application that runs tests each day on Linux virtual machines. The full test suite takes several hours to complete, running on a limited number of on-premises servers reserved for testing. Your company wants to move the testing infrastructure to Google Cloud Platform. Your company wants to reduce the amount of time it takes to fully test a change to the system while

changing the tests as little as possible. Your project manager has asked you to suggest suitable services in Google Cloud and you want to follow Google recommended practices. What should you do?

- Use Google App Engine and Google Stackdriver for logging.
- Use Google Compute Engine unmanaged instance groups with a Network Load Balancer.
- Use Google Compute Engine managed instance groups and autoscaling.**
- Use Google Cloud Dataproc and run Apache Hadoop jobs to process each test.

#### Unattempted

Use Google Compute Engine unmanaged instance groups with a Network Load Balancer. is not right.  
An unmanaged instance group is a collection of virtual machines (VMs) that reside in a single zone, VPC network, and subnet. An unmanaged instance group is useful for grouping together VMs that require individual configuration settings or tuning. Unmanaged instance group does not autoscale, so it does not help reduce the amount of time it takes to fully test a change to the system.

Ref: <https://cloud.google.com/compute/docs/instance-groups/creating-groups-of-unmanaged-instances>

Use Google App Engine and Google Stackdriver for logging. is not right.

App Engine supports many popular languages like Java, PHP, Node.js, Python, C#, .Net, Ruby, and Go. However, C++ isn't supported by App Engine.

Ref: <https://cloud.google.com/appengine>

Use Google Cloud Dataproc and run Apache Hadoop jobs to process each test. is not right.

Cloud Dataproc is a fast, easy-to-use, fully managed cloud service for running Apache Spark and Apache Hadoop clusters in a simpler, more cost-efficient way. While Dataproc is very efficient at processing ETL and Big Data pipelines, it is not as suitable for running a ruby application that runs tests each day.

Ref: <https://cloud.google.com/dataproc>

Use Google Compute Engine managed instance groups and autoscaling. is the right answer.

A managed instance group (MIG) contains identical virtual machine (VM) instances that are based on an instance template. MIGs support auto-healing, load balancing, autoscaling, and auto-updating. Managed instance groups offer auto-scaling capabilities that let you automatically add or delete instances from a managed instance group based on increases or decreases in load. Autoscaling helps your apps gracefully handle increases in traffic and reduce costs when the need for resources is lower. Autoscaling works by adding more instances to your instance group when there is more load (upscale), and deleting instances when the need for instances is lowered (downscale).

Ref: <https://cloud.google.com/compute/docs/autoscaler/>

#### 47. Question

Your compliance team requested all audit logs are stored for 10 years and to allow access for external auditors to view. You want to follow Google recommended practices. What should you do? (Choose two)

- Create an account for auditors to have view access to Stackdriver Logging.
- Export audit logs to Cloud Storage via an export sink.**
- Export audit logs to BigQuery via an export sink.
- Generate a signed URL to the Stackdriver export destination for auditors to access.
- Export audit logs to Splunk via a Pub/Sub export sink.

#### Unattempted

Create an account for auditors to have view access to Stackdriver Logging. is not right.

While it is possible to configure a custom retention period of 10 years in Stackdriver logging, storing logs in Stackdriver is expensive compared to Cloud Storage. Stackdriver charges \$0.01 per GB per month, whereas something like Cloud Storage Coldline Storage costs \$0.007 per GB per month (30% cheaper) and Cloud Storage Archive Storage costs 0.004 per GB per month (60% cheaper than Stackdriver)

Ref: <https://cloud.google.com/logging/docs/storage#pricing>

Ref: <https://cloud.google.com/storage/pricing>

Export audit logs to BigQuery via an export sink. is not right.

Storing logs in BigQuery is expensive. In BigQuery, Active storage costs \$0.02 per GB per month and Long-term storage costs \$0.01 per GB per month. In comparison, Google Cloud Storage offers several storage classes that are significantly cheaper.

Ref: <https://cloud.google.com/bigquery/pricing>

Ref: <https://cloud.google.com/storage/pricing>

Export audit logs to Cloud Filestore via a Pub/Sub export sink. is not right.

Storing logs in Cloud Filestore is expensive. In Cloud Filestore, Standard Tier pricing costs \$0.2 per GB per month and Premium Tier pricing costs \$0.3 per GB per month. In comparison, Google Cloud Storage offers several storage classes that are significantly cheaper.

Ref: <https://cloud.google.com/bigquery/pricing>

Ref: <https://cloud.google.com/storage/pricing>

Export audit logs to Cloud Storage via an export sink. is the right answer.

Among all the storage solutions offered by Google Cloud Platform, Cloud storage offers the best pricing for long term storage of logs. Google Cloud Storage offers several storage classes such as Nearline Storage (\$0.01 per GB per Month) Coldline Storage (\$0.007 per GB per Month) and Archive Storage (\$0.004 per GB per month) which are significantly cheaper than the storage options covered by the above options above.

Ref: <https://cloud.google.com/storage/pricing>

Generate a signed URL to the Stackdriver export destination for auditors to access. is the right answer.  
In Google Cloud Storage, you can generate a signed URL to provide limited permission and time to make a request. Anyone who possesses it can use the signed URL to perform specified actions, such as reading an object, within a specified period of time.

In our scenario, we do not need to create accounts for our auditors to provide access to logs in Cloud Storage. Instead, we can generate them signed URLs which are time-bound and lets them access/download log files.

Ref: <https://cloud.google.com/storage/docs/access-control/signed-urls>

## 48. Question

Your customer has implemented a solution that uses Cloud Spanner and notices some read latency-related performance issues on one table. This table is accessed only by their users using a primary key. The table schema is shown below.

```
CREATE TABLE Persons {
```

```
 person_id INT64 NOT NULL, // sequential number based on number of registrations
```

```
 account_creation_date DATE, // system date
```

```
 birthdate DATE, // customer birthdate
```

```
 firstname STRING (255), // first name
```

```
 lastname STRING (255), // last name
```

```
 profile_picture BYTES (255) // profile picture
```

```
} PRIMARY KEY (person_id)
```

You want to resolve the issue. What should you do?

- Remove the profile\_picture field from the table.
- Add a secondary index on the person\_id column.
- Change the primary key to not have monotonically increasing values.
- Create a secondary index using the following Data Definition Language (DDL):
- CREATE INDEX person\_id\_ix ON Persons ( person\_id, firstname, lastname ) STORING ( profile\_picture )

**Unattempted**

Change the primary key to not have monotonically increasing values. is the right answer.

You should be careful when choosing a primary key to not accidentally create hotspots in your database.

One cause of hotspots is having a column whose value monotonically increases as the first key part because this results in all inserts occurring at the end of your keyspace. This pattern is undesirable because Cloud Spanner divides data among servers by key ranges, which means all your inserts will be directed at a single server that will end up doing all the work.

Ref: <https://cloud.google.com/spanner/docs/schema-design#primary-key-prevent-hotspots>

All other options make no sense. The problem is with the monotonically increasing values in the primary key and removing profile\_picture or adding a secondary index isn't going to alleviate the problem.

**49. Question**

Your development team needs a new Jenkins server for their project. You need to deploy the server using the fewest steps possible. What should you do?

- Create a new Compute Engine instance and install Jenkins through the command-line interface.
- Download and deploy the Jenkins Java WAR to App Engine Standard.
- Use GCP Marketplace to launch the Jenkins solution.**
- Create a Kubernetes cluster on Compute Engine and create a deployment with the Jenkins Docker image.

**Unattempted**

Create a new Compute Engine instance and install Jenkins through the command line interface. is not right.

While this can be done, this involves a lot more work than installing the Jenkins server through App Engine.

Create a Kubernetes cluster on Compute Engine and create a deployment with the Jenkins Docker image. is not right.

While this can be done, this involves a lot more work than installing the Jenkins server through App Engine.

Download and deploy the Jenkins Java WAR to App Engine Standard. is not right.

While this is possible, we need to ensure App Engine is enabled, we then need to download the Java project/WAR, and run gcloud app deploy to set up a Jenkins server. This involves more steps than spinning up an instance from GCP Marketplace.

Ref: <https://cloud.google.com/appengine/docs/standard/java/tools/uploadinganapp>

Ref: <https://cloud.google.com/solutions/using-jenkins-for-distributed-builds-on-compute-engine>

Use GCP Marketplace to launch the Jenkins solution. is the right answer.

The simplest way to launch a Jenkins server is from GCP Market place. GCP market place has a number of builds available for Jenkins: <https://console.cloud.google.com/marketplace/browse?q=jenkins>. All you need to do is spin up an instance from a suitable market place build and you have a Jenkins server in a few minutes with just a few clicks.

## 50. Question

Your finance team wants to view the billing report for your projects. You want to make sure that the finance team does not get additional permissions to the project. What should you do?

- Add the group for the finance team to roles/billing.user role.
- Add the group for the finance team to roles/billing.admin role.
- Add the group for the finance team to roles/billing.viewer role.
- Add the group for the finance team to roles/billing.projectManager role.

### Unattempted

Add the group for the finance team to roles/billing.user role. is not right.

This role has very restricted permissions, so you can grant it broadly, typically in combination with Project Creator. These two roles allow a user to create new projects linked to the billing account on which the role is granted.

Ref: <https://cloud.google.com/billing/docs/how-to/billing-access>

Add the group for the finance team to roles/billing.admin role. is not right.

This role is an owner role for a billing account. Use it to manage payment instruments, configure billing exports, view cost information, link and unlink projects, and manage other user roles on the billing account.

Ref: <https://cloud.google.com/billing/docs/how-to/billing-access>

Add the group for the finance team to roles/billing.projectManager role. is not right.

This role allows a user to attach the project to the billing account, but does not grant any rights over resources. Project Owners can use this role to allow someone else to manage the billing for the project without granting them resource access.

Ref: <https://cloud.google.com/billing/docs/how-to/billing-access>

Add the group for the finance team to roles/billing.viewer role. is the right answer.

Billing Account Viewer access would usually be granted to finance teams, it provides access to spending information but does not confer the right to link or unlink projects or otherwise manage the properties of the billing account.

Ref: <https://cloud.google.com/billing/docs/how-to/billing-access>

## 51. Question

Your managed instance group raised an alert stating that new instance creation has failed to create new instances. You need to maintain the number of running instances specified by the template to be able to process expected application traffic. What should you do?

- Create an instance template that contains valid syntax that will be used by the instance group. Delete any persistent disks with the same name as instance names.
- Create an instance template that contains valid syntax that will be used by the instance group. Verify that the instance name and persistent disk name values are not the same in the template.
- Verify that the instance template being used by the instance group contains valid syntax. Delete any persistent disks with the same name as instance names. Set the `disks.autoDelete` property to true in the instance template.
- Delete the current instance template and replace it with a new instance template. Verify that the instance name and persistent disk name values are not the same in the template. Set the `disks.autoDelete` property to true in the instance template.

### Unattempted

Verify that the instance template being used by the instance group contains valid syntax. Delete any persistent disks with the same name as instance names. Set the `disks.autoDelete` property to true in the instance template. is the right answer.

As described in this article, My managed instance group keeps failing to create a VM. What's going on?  
<https://cloud.google.com/compute/docs/instance-groups/creating-groups-of-managed-instances#troubleshooting>

The likely causes are

1. A persistent disk already exists with the same name as VM Instance
2. `disks.autoDelete` option is set to false
3. instance template might be invalid

Therefore, we need to ensure that instance template is valid, `disks.autoDelete` is turned on, and that there are no existing persistent disks with the same name as VM instance.

## 52. Question

Your management has asked an external auditor to review all the resources in a specific project. The security team has enabled the Organization Policy called Domain Restricted Sharing on the organization node by specifying only your Cloud Identity domain. You want the auditor to only be able to view, but not modify, the resources in that project. What should you do?

- Ask the auditor for their Google account, and give them the Viewer role on the project.
- Ask the auditor for their Google account, and give them the Security Reviewer role on the project.
- Create a temporary account for the auditor in Cloud Identity, and give that account the Viewer role on the project.
- Create a temporary account for the auditor in Cloud Identity, and give that account the Security Reviewer role on the project.

### Unattempted

Ask the auditor for their Google account, and give them the Viewer role on the project. is not right.  
Since the auditor s account is not part of your company s Cloud Identity domain, the auditor can not access resources from your GCP projects. Domain restriction constraint can be used in organization policies to limit resource sharing based on domain. This constraint allows you to restrict the set of identities that are allowed to be used in Cloud Identity and Access Management policies. In this scenario, since we are restricting based on the Cloud Identity domain, only the users in the Cloud Identity domain can access GCP services.

<https://cloud.google.com/resource-manager/docs/organization-policy/restricting-domains>

Ask the auditor for their Google account, and give them the Security Reviewer role on the project. is not right.

Since the auditor s account is not part of your company s Cloud Identity domain, the auditor can not access resources from your GCP projects. Domain restriction constraint can be used in organization policies to limit resource sharing based on domain. This constraint allows you to restrict the set of identities that are allowed to be used in Cloud Identity and Access Management policies. In this scenario, since we are restricting based on the Cloud Identity domain, only the users in the Cloud Identity domain can access GCP services.

<https://cloud.google.com/resource-manager/docs/organization-policy/restricting-domains>

Create a temporary account for the auditor in Cloud Identity, and give that account the Security Reviewer role on the project. is not right.

Creating a temporary account for the auditor in your cloud identity is the right approach as this makes the auditor part of the Cloud identity domain and the organization policy in place lets the auditor access resources. However, the role granted here is not suitable, it provides permissions to list all resources and Cloud IAM policies. Note that list permissions only allow you to list but not view resources. You need to get permission to view the resources.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#iam-roles>

Create a temporary account for the auditor in Cloud Identity, and give that account the Viewer role on the project. is the right answer.

The primitive viewer role provides permissions for read-only actions that do not affect the state, such as

viewing (but not modifying) existing resources or data. This fits our requirements.

In addition, adding the auditor to Cloud Identity ensures that Organization Policy for Domain Restricted Sharing doesn't block them from accessing resources.

Ref: [https://cloud.google.com/iam/docs/understanding-roles#primitive\\_role\\_definitions](https://cloud.google.com/iam/docs/understanding-roles#primitive_role_definitions)

### 53. Question

Your operations team has configured a lifecycle management rule on a bucket. The bucket is multi-regional and has versioning enabled. Which of the following statement accurately reflects the following lifecycle config?

```
{
```

```
rule :[
```

```
{
```

```
action :{
```

```
type : Delete
```

```
,
```

```
condition :{
```

```
age :60,
```

```
isLive :false
```

```
}
```

```
,
```

```
{
```

```
action :{
```

```
type : SetStorageClass ,
```

```
storageClass : COLDLINE
```

```
,
```

```
condition :{
```

```
 age :366,
```

```
 matchesStorageClass : MULTI_REGIONAL
```

```
}
```

```
}
```

```
]
```

```
}
```

- Move objects to Coldline Storage after 366 days if the storage class in Multi-regional First rule has no effect on the bucket.
- Archive objects older than 60 days and move objects to Coldline Storage after 366 days if the storage class in Multi-regional.
- Delete objects older than 60 days and move objects to Coldline Storage after 366 days if the storage class in Multi-regional.
- Delete archived objects older than 60 days and move objects to Coldline Storage after 366 days if the storage class is Multi-regional.

#### Unattempted

Archive objects older than 60 days and move objects to Coldline Storage after 366 days if the storage class in Multi-regional. is not right.

The action has type : Delete which means we want to Delete, not archive.

Ref: <https://cloud.google.com/storage/docs/managing-lifecycles>

Delete objects older than 60 days and move objects to Coldline Storage after 366 days if the storage class in Multi-regional. is not right.

We want to delete objects as indicated by the action, however, we don t want to delete all objects older than 60 days. We only want to delete archived objects as indicated by isLive :false condition

Ref: <https://cloud.google.com/storage/docs/managing-lifecycles>

Move objects to Coldline Storage after 366 days if the storage class in Multi-regional. First rule has no effect on the bucket. is not right.

The first rule certainly has an effect. It deletes archived objects older than 60 days.

Delete archived objects older than 60 days and move objects to Coldline Storage after 366 days if the storage class is Multi-regional. is the right answer.

The first part of the rule: The action has type : Delete which means we want to Delete. isLive :false condition means we are looking for objects that are not Live i.e. objects that are archived. Together, it means we want to delete archived objects older than 60 days. Note that if an object is deleted, it cannot be undeleted. Take care in setting up your lifecycle rules so that you do not cause more data to be deleted than you intend.

Ref: <https://cloud.google.com/storage/docs/managing-lifecycles>

The second part of the rule: The action indicates we want to set storage class to Coldline. The condition is true if the existing storage class is multi-regional and the age of the object is 366 days or over.

Together it means we want to set the storage class to Coldline if existing storage class is multi-regional and age of the object is 366 days or over

#### 54. Question

Your organization has a dedicated person who creates and manages all service accounts for Google Cloud projects. You need to assign this person the minimum role for projects. What should you do?

- Add the user to roles/iam.roleAdmin role.
- Add the user to roles/iam.securityAdmin role.
- Add the user to roles/iam.serviceAccountUser role.
- Add the user to roles/iam.serviceAccountAdmin role.

#### Unattempted

Add the user to roles/iam.roleAdmin role. is not right.

roles/iam.roleAdmin is an administrator role that provides access to all custom roles in the project. This doesn't include permissions needed to manage service accounts.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#roles-roles>

Add the user to roles/iam.securityAdmin role. is not right.

roles/iam.securityAdmin role is a Security admin role, with permissions to get and set any IAM policy.

This role is too broad i.e. includes too many permissions and goes against the principle of least privilege. Moreover, although this role provides iam.serviceAccounts.get/list, it doesn't provide iam.serviceAccounts.create, iam.serviceAccounts.delete and iam.serviceAccounts.update permissions that are needed for managing service accounts.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#iam-roles>

Add the user to roles/iam.serviceAccountUser role. is not right.

roles/iam.serviceAccountUser is a service Account User role which is used for running operations as the service account. This role does not provide the permissions iam.serviceAccounts.create,

iam.serviceAccounts.delete, iam.serviceAccounts.update, iam.serviceAccounts.get/list which are required for managing service accounts.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#service-accounts-roles>

Add the user to roles/iam.serviceAccountAdmin role. is the right answer.

roles/iam.serviceAccountAdmin is a Service Account Admin role that lets you Create and manage service accounts. This grants all the required permissions for managing service accounts (iam.serviceAccounts.create iam.serviceAccounts.delete, iam.serviceAccounts.update, iam.serviceAccounts.get/list etc) and therefore fits our requirements.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#service-accounts-roles>

## 55. Question

Your organization has strict requirements to control access to Google Cloud projects. You need to enable your Site Reliability Engineers (SREs) to approve requests from the Google Cloud support team when an SRE opens a support case. You want to follow Google-recommended practices. What should you do?

- Add your SREs to roles/iam.roleAdmin role.
- Add your SREs to roles/accessapproval.approver role.
- Add your SREs to a group and then add this group to roles/iam.roleAdmin role.
- Add your SREs to a group and then add this group to roles/accessapproval.approver role.

### Unattempted

Add your SREs to roles/iam.roleAdmin role. is not right.

roles/iam.roleAdmin provides access to all custom roles in the project. This doesn't fit our requirement of SREs being able to approve requests.

Add your SREs to a group and then add this group to roles/iam.roleAdmin role. is not right.

roles/iam.roleAdmin provides access to all custom roles in the project. This doesn't fit our requirement of SREs being able to approve requests.

Add your SREs to roles/accessapproval.approver role. is not right.

roles/accessapproval.approver is an Access Approval Approver role and provides the ability to view or act on access approval requests and view configuration. Although this is the role we require, you want to follow Google recommended practices which means we should instead add the group to the role and add users to the group instead of granting the role individually to users.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#access-approval-roles>

Ref: <https://cloud.google.com/iam/docs/overview>

Add your SREs to a group and then add this group to roles/accessapproval.approver role. is the right answer.

roles/accessapproval.approver is an Access Approval Approver role and provides the ability to view or act on access approval requests and view configuration. And you follow Google recommended practices by adding users to the group and group to the role. Groups are a convenient way to apply an access policy to a collection of users. You can grant and change access controls for a whole group at once instead of granting or changing access controls one at a time for individual users or service accounts. You can also easily add members to and remove members from a Google group instead of updating a Cloud IAM policy to add or remove users.

Ref: <https://cloud.google.com/iam/docs/understanding-roles#access-approval-roles>

Ref: <https://cloud.google.com/iam/docs/overview>

## 56. Question

A company wants to build an application that stores images in a Cloud Storage bucket and wants to generate thumbnails as well as resize the images. They want to use a google managed service that can scale up and scale down to zero automatically with minimal effort. You have been asked to recommend a service. Which GCP service would you suggest?

- Google Compute Engine
- Google Kubernetes Engine
- Cloud Functions**
- Google App Engine

### Unattempted

Cloud Functions. is the right answer.

?Cloud Functions is Google Cloud's event-driven serverless compute platform. It automatically scales based on the load and requires no additional configuration. You pay only for the resources used.

?Ref: <https://cloud.google.com/functions>

?While all other options i.e. Google Compute Engine, Google Kubernetes Engine, Google App Engine support autoscaling, it needs to be configured explicitly based on the load and is not as trivial as the scale up or scale down offered by Google's cloud functions.

## 57. Question

A team of data scientists infrequently needs to use a Google Kubernetes Engine (GKE) cluster that you manage. They require GPUs for some long-running, non-restartable jobs. You want to minimize cost. What

should you do?

- Enable node auto-provisioning on the GKE cluster.
- Create a VerticalPodAutscaler for those workloads.
- Create a node pool with preemptible VMs and GPUs attached to those VMs.
- Create a node pool of instances with GPUs, and enable autoscaling on this node pool with a minimum size of 1.**

### Unattempted

Enable node auto-provisioning on the GKE cluster. is not right.

?Node auto-provisioning automatically manages a set of node pools on the user's behalf. Without Node auto-provisioning, GKE considers starting new nodes only from the set of user-created node pools. With node auto-provisioning, new node pools can be created and deleted automatically. This in no way helps us with our requirements.

?Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/node-auto-provisioning> Create a VerticalPodAutscaler for those workloads. is not right.

?Vertical pod autoscaling (VPA) frees you from having to think about what values to specify for a container's CPU and memory requests. The autoscaler can recommend values for CPU and memory requests and limits, or it can automatically update the values. This doesn't help us with the GPU requirement. Moreover, due to Kubernetes limitations, the only way to modify the resource requests of a running Pod is to recreate the Pod. This has the negative effect of killing the non-restartable jobs which is undesirable.

?<https://cloud.google.com/kubernetes-engine/docs/concepts/verticalpodautoscaler#overview> Create a node pool with preemptible VMs and GPUs attached to those VMs. is not right.

?You can use preemptible VMs in your GKE clusters or node pools to run batch or fault-tolerant jobs that are less sensitive to the ephemeral, non-guaranteed nature of preemptible VMs. Whereas we have long-running and non-restartable jobs so preemptible VMs aren't suitable for our requirement.

?Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/preemptible-vms> Create a node pool of instances with GPUs, and enable autoscaling on this node pool with a minimum size of 1. is the right answer.

?A node pool is a group of nodes within a cluster that all have the same configuration. Our requirement is GPUs, so we create a node pool with GPU enabled and have the scientist's applications deployed to the cluster and use this node pool. At the same time, you want to minimize cost so you start with 1 instance

and scale up as needed. It is important to note that the scale down needs to take into consideration if there are any running jobs otherwise the scale down may terminate the nonrestartable job.

?Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/node-pools>

## 58. Question

An employee was terminated, but their access to Google Cloud Platform (GCP) was not removed until 2 weeks later. You need to find out this employee accessed any sensitive customer information after their termination. What should you do?

- View System Event Logs in Stackdriver. Search for the user's email as the principal.
- View System Event Logs in Stackdriver. Search for the service account associated with the user.
- View Data Access audit logs in Stackdriver. Search for the user's email as the principal.**
- View the Admin Activity log in Stackdriver. Search for the service account associated with the user.

### Unattempted

View the Admin Activity log in Stackdriver. Search for the service account associated with the user. is not right.

?Admin Activity logs do not contain log entries for reading resource data. Admin Activity audit logs contain log entries for API calls or other administrative actions that modify the configuration or metadata of resources.

?Ref: <https://cloud.google.com/logging/docs/audit#admin-activity> View System Event Logs in Stackdriver. Search for the user s email as the principal. is not right.

?System Event audit logs do not contain log entries for reading resource data. System Event audit logs contain log entries for Google Cloud administrative actions that modify the configuration of resources. System Event audit logs are generated by Google systems; they are not driven by direct user action.

?Ref: <https://cloud.google.com/logging/docs/audit#system-event> View System Event Logs in Stackdriver. Search for the service account associated with the user. is not right.

?System Event audit logs do not contain log entries for reading resource data. System Event audit logs contain log entries for Google Cloud administrative actions that modify the configuration of resources. System Event audit logs are generated by Google systems; they are not driven by direct user action.

?Ref: <https://cloud.google.com/logging/docs/audit#system-event> View Data Access audit logs in Stackdriver. Search for the user s email as the principal. is the right answer.

?Data Access audit logs contain API calls that read the configuration or metadata of resources, as well as user-driven API calls that create, modify, or read user-provided resource data.

?Ref: <https://cloud.google.com/logging/docs/audit#data-access>

## 59. Question

An engineer from your team accidentally deployed several new versions of NodeJS application on Google App Engine Standard. You are concerned the new versions are serving traffic. You have been asked to produce a list of all the versions of the application that are receiving traffic as well the percent traffic split between them. What should you do?

- gcloud app versions list --hide-no-traffic
- gcloud app versions list --show-traffic
- gcloud app versions list
- gcloud app versions list --traffic

### Unattempted

gcloud app versions list. is not right

?This command lists all the versions of all services that are currently deployed to the App Engine server. While this list includes all versions that are receiving traffic, it also includes versions that are not receiving traffic.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/list> gcloud app versions list traffic. is not right

?gcloud app versions list command does not support traffic flag.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/list> gcloud app versions list show-traffic. is not right

?gcloud app versions list command does not support show-traffic flag.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/list> gcloud app versions list hide-no-traffic. is the right answer.

?This command correctly lists just the versions that are receiving traffic by hiding versions that do not receive traffic. This is the only command that fits our requirements.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/list>

## 60. Question

An intern joined your team recently and needs access to Google Compute Engine in your sandbox project to explore various settings and spin up compute instances to test features. You have been asked to facilitate this. How should you give your intern access to compute engine without giving more permissions than is necessary?

- Grant Project Editor IAM role for sandbox project.
- Grant Compute Engine Admin Role for sandbox project.
- Create a shared VPC to enable the intern access Compute resources.
- Grant Compute Engine Instance Admin Role for the sandbox project.

### Unattempted

Create a shared VPC to enable the intern access Compute resources. is not right.

?Creating a shared VPC is not sufficient to grant intern access to compute resources. Shared VPCs are primarily used by organizations to connect resources from multiple projects to a common Virtual Private Cloud (VPC) network, so that they can communicate with each other securely and efficiently using internal IPs from that network.

?Ref: <https://cloud.google.com/vpc/docs/shared-vpc> Grant Project Editor IAM role for sandbox project. is not right.

?Project editor role grants all viewer permissions, plus permissions for actions that modify state, such as changing existing resources. While this role lets the intern explore compute engine settings and spin up compute instances, it grants more permissions than what is needed. Our intern can modify any resource in the project.

?[https://cloud.google.com/iam/docs/understanding-roles#primitive\\_roles](https://cloud.google.com/iam/docs/understanding-roles#primitive_roles) Grant Compute Engine Admin Role for sandbox project. is not right.

?Compute Engine Admin Role grants full control of all Compute Engine resources; including networks, load balancing, service accounts etc. While this role lets the intern explore compute engine settings and spin up compute instances, it grants more permissions than what is needed.

?Ref: <https://cloud.google.com/compute/docs/access/iam#compute.storageAdmin> Grant Compute Engine Instance Admin Role for the sandbox project. is the right answer.

?Compute Engine Instance Admin Role grants full control of Compute Engine instances, instance groups, disks, snapshots, and images. It also provides read access to all Compute Engine networking resources.

This provides just the required permissions to the intern.

?Ref: <https://cloud.google.com/compute/docs/access/iam#compute.storageAdmin>

## 61. Question

Auditors visit your teams every 12 months and ask to review all the Google Cloud Identity and Access Management (Cloud IAM) policy changes in the previous 12 months. You want to streamline and expedite the analysis and audit process. What should you do?

- Enable Logging export to Google Cloud Storage (GCS) bucket and delegate access to the bucket
- Enable Logging export to Google BigQuery and use ACLs and views to scope the data shared with the auditor**
- Create custom Google Stackdriver alerts and send them to the auditor
- Use Cloud Functions to transfer log entries to Google Cloud SQL and use ACLs and views to limit an auditor's view

### Unattempted

Create custom Google Stackdriver alerts and send them in an email to the auditor. is not right.

?Stackdriver Alerting gives timely awareness to problems in your cloud applications so you can resolve the problems quickly. Sending alerts to your auditor is not of much use during audits.

?Ref: <https://cloud.google.com/monitoring/alerts> Use Cloud Functions to transfer log entries to Google Cloud SQL and use ACLs and views to limit an auditor's view. is not right.

?Using Cloud Functions to transfer log entries to Google Cloud SQL is expensive in comparison to audit logs export feature which exports logs to various destinations with minimal configuration.

?Ref: <https://cloud.google.com/logging/docs/export/>

?Auditors spend a lot of time reviewing log messages. And you want to expedite the audit process!! So you want to make it easier for the auditor to extract the information easily from the logs. Between the two remaining options, the only difference is the log export sink destination

?Ref: <https://cloud.google.com/logging/docs/export/>

One option exports to Google Cloud Storage (GCS) bucket whereas other exports to BigQuery. Querying information out of files in a bucket is much harder compared to querying information from BigQuery Dataset where it is as simple as running a job or set of jobs to extract just the required information and in the format required. By enabling the auditor to run jobs in Big Queries, you streamline the log extraction

process and the auditor can review the extracted logs much quicker. While as good as the other option (bucket) is, Enable Logging export to Google BigQuery and use ACLs and views to scope the data shared with the auditor is the right answer. You need to configure log sinks before you can receive any logs, and you can't retroactively export logs that were written before the sink was created.

## 62. Question

Every employee of your company has a Google account. Your operational team needs to manage a large number of instances on the Compute Engine. Each member of this team needs only administrative access to the servers. Your security team wants to ensure that the deployment of credentials is operationally efficient and must be able to determine who accessed a given instance. What should you do?

- Ask each member of the team to generate a new SSH key pair and to send you their public key. Use a configuration management tool to deploy those keys on each instance.
- Ask each member of the team to generate a new SSH key pair and to add the public key to their Google account. Grant the "compute.osAdminLogin" role to the Google group corresponding to this team.
- Generate a new SSH key pair. Give the private key to each member of your team. Configure the public key in the metadata of each instance.
- Generate a new SSH key pair. Give the private key to each member of your team. Configure the public key as a project-wide public SSH key in your Cloud Platform project and allow project-wide public SSH keys on each instance.

### Unattempted

Generate a new SSH key pair. Give the private key to each member of your team. Configure the public key in the metadata of each instance. is not right.

?Reuse of a single SSH key pair by all employees is a very bad security practice as auditing becomes very impossible. Generate a new SSH key pair. Give the private key to each member of your team. Configure the public key as a project-wide public SSH key in your Cloud Platform project and allow project-wide public SSH keys on each instance. is not right.

?Reuse of a single SSH key pair by all employees is a very bad security practice as auditing becomes very impossible. Ask each member of the team to generate a new SSH key pair and to send you their public key. Use a configuration management tool to deploy those keys on each instance. is not right.

?While this can be done, it is not operationally efficient. Let's say a user leaves the company, you then have to remove their SSH key from all instances where it has been added (can't be removed at a single place). Similarly, when a user joins the company, you have to add their SSH key to all the instances. This is very tedious and not operationally efficient. Ask each member of the team to generate a new SSH key

pair and to add the public key to their Google account. Grant the `compute.osAdminLogin` role to the Google group corresponding to this team. is the right answer.

?By letting users manage their own SSH key pair (and its rotation etc), you reduce the operational burden of managing SSH keys to individual users. Secondly, granting `compute.osAdminLogin` grants the group administrator permissions (as opposed to granting `compute.osLogin`, which does not grant administrator permissions). Finally, managing provisioning and de-provisioning is as simple as adding or removing the user from the group. OS Login lets you use Compute Engine IAM roles to efficiently manage SSH access to Linux instances and is an alternative to manually managing instance access by adding and removing SSH keys in the metadata. Before you can manage instance access using IAM roles, you must enable the OS Login feature by setting a metadata key-value pair in your project or in your instance's metadata: `enable-oslogin=TRUE`. After you enable OS Login on one or more instances in your project, those instances accept connections only from user accounts that have the necessary IAM roles in your project or organization. There are two predefined roles.

?? `roles/compute.osLogin`, which does not grant administrator permissions

?? `roles/compute.osAdminLogin`, which grants administrator permissions

?At any point, to revoke user access to instances that are enabled to use OS Login, remove the user roles from that user account

?Ref: [https://cloud.google.com/compute/docs/instances/managing-instance-access#enable\\_oslogin](https://cloud.google.com/compute/docs/instances/managing-instance-access#enable_oslogin)

### 63. Question

For service discovery, you need to associate each of the Compute Engine instances of your VPC with an internal (DNS) record in a custom zone. You want to follow Google recommended practices. What should you do?

- Create a new VPC, block all external traffic with a firewall rule and create 2 Cloud DNS zones - a first zone in the new VPC and a second zone in the main VPC that is forwarding requests to the first Cloud DNS zone. Create records for each instance in the first zone.
- Deploy the BIND DNS server in the VPC, and create a Cloud DNS forwarding zone to forward the DNS requests to BIND. Create records for each instance in the BIND DNS server.
- Create a Cloud DNS zone, set its visibility to private and associate it with your VPC. Create records for each instance in that zone.**
- Create your Compute Engine instances with custom hostnames.

**Unattempted**

Our requirements here are 1. Internal and 2. Custom Zone Create your Compute Engine instances with custom hostnames. is not right.

?This doesn't put them in a custom zone. Deploy the BIND DNS server in the VPC, and create a Cloud DNS forwarding zone to forward the DNS requests to BIND. Create records for each instance in the BIND DNS server. is not right.

?This might be possible but not something Google recommends. The Cloud DNS service offering from Google already offers these features so it is pointless installing a custom DNS server to do that. Create a new VPC, block all external traffic with a firewall rule and create 2 Cloud DNS zones a first zone in the new VPC and a second zone in the main VPC that is forwarding requests to the first Cloud DNS zone. Create records for each instance in the first zone. is not right.

?This doesn't make any sense, moreover, the two VPCs can't communicate without VPC peering.

?Ref: <https://cloud.google.com/dns/docs/overview#concepts> Create a Cloud DNS zone, set its visibility to private and associate it with your VPC. Create records for each instance in that zone. is the right answer.

?You should absolutely do this when you want internal DNS records in a custom zone. Cloud DNS gives you the option of private zones and internal DNS names.

?Ref: <https://cloud.google.com/dns/docs/overview#concepts>

## 64. Question

In Cloud Shell, your active gcloud configuration is as shown below.

```
$ gcloud config list
```

```
[component_manager]
```

```
disable_update_check = True
```

```
[compute]
```

```
gce_metadata_read_timeout_sec = 5
```

```
zone = europe-west2-a
```

```
[core]
```

```
account = gcp-ace-lab-user@gmail.com
```

```
disable_usage_reporting = False
```

```
project = gcp-ace-lab-266520
```

```
[metrics]
```

```
environment = devshell
```

You want to create two compute instances one in europe-west2-a and another in europe-west2-b. What should you do? (Select 2)

- gcloud compute instances create instance1 gcloud compute instances create instance2
- gcloud compute instances create instance1 gcloud config set compute/zone europe-west2-b gcloud compute instances create instance2
- gcloud compute instances create instance1 gcloud compute instances create instance2 – zone=europe-west2-b
- gcloud compute instances create instance1 gcloud config set zone europe-west2-b gcloud compute instances create instance2
- gcloud compute instances create instance1 gcloud configuration set compute/zone europe-west2-b gcloud compute instances create instance2

### Unattempted

gcloud compute instances create instance1

?gcloud compute instances create instance2. is not right.

?The default compute/zone property is set to europe-west2-a in the current gcloud configuration.

Executing the two commands above would create two compute instances in the default zone i.e. europe-west2-a which doesn't satisfy our requirement.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/create> gcloud compute instances create instance1

?gcloud config set zone europe-west2-b

?gcloud compute instances create instance2. is not right.

?The approach is right but the syntax is wrong. gcloud config does not have a core/zone property. The syntax for this command is gcloud config set SECTION/PROPERTY VALUE. If SECTION is missing,

SECTION is defaulted to core. We are effectively trying to run gcloud config set core/zone europe-west2-b but the core section doesn't have a property called zone, so this command fails.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/config/set> gcloud compute instances create instance1

?gcloud configuration set compute/zone europe-west2-b

?gcloud compute instances create instance2. is not right.

?Like above, the approach is right but the syntax is wrong. You want to set the default compute/zone property in gcloud configuration to europe-west2-b but it needs to be done via the command gcloud config set and not gcloud configuration set.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/config/set> gcloud compute instances create instance1

?gcloud config set compute/zone europe-west2-b

?gcloud compute instances create instance2. is the right answer.

?The default compute/zone property is europe-west2-a in the current gcloud configuration so executing the first gcloud compute instances create command creates the instance in europe-west2-a zone. Next, executing the gcloud config set compute/zone europe-west2-b changes the default compute/zone property in default configuration to europe-west2-b. Executing the second gcloud compute instances create command creates a compute instance in europe-west2-b which is what we want.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/config/set>

?Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/create> gcloud compute instances create instance1

?gcloud compute instances create instance2 --zone=europe-west2-b. is the right answer.

?The default compute/zone property is europe-west2-a in the current gcloud configuration so executing the first gcloud compute instances create command creates the instance in europe-west2-a zone. Next, executing the second gcloud compute instances create command with --zone property creates a compute instance in provided zone i.e. europe-west2-b instead of using the default zone from the current active configuration.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/create>

## 65. Question

In Regional Storage buckets with object versioning enabled, what is the effect of deleting the live version of an object and deleting a noncurrent version of an object?

- 1. The live version becomes a noncurrent version. 2. The noncurrent version is deleted permanently.
- 1. The live version becomes a noncurrent version and a lifecycle rule is applied to delete after 30 days. 2. A lifecycle rule is applied on the noncurrent version to delete after 30 days.
- 1. The live version becomes a noncurrent version and a lifecycle rule is applied to transition to Nearline Storage after 30 days. 2. A lifecycle rule is applied on the noncurrent version to transition to Nearline Storage after 30 days.
- 1. The live version is deleted permanently. 2. The noncurrent version is deleted permanently.

### Unattempted

1. The live version becomes a noncurrent version.

?2. The noncurrent version is deleted permanently. is the right answer. In buckets with object versioning enabled, deleting the live version of an object creates a noncurrent version while deleting a noncurrent version deletes that version permanently.

?Ref: <https://cloud.google.com/storage/docs/lifecycle#actions>

**Use Page numbers below to navigate to other  
practice tests**

Pages:

[1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [11](#) [12](#) [13](#)

← Previous Post

Next Post →

We help you to succeed in your certification exams

We have helped over thousands of working professionals to achieve their certification goals with our practice tests.

## Skillcertpro



### Quick Links

[ABOUT US](#)

[FAQ](#)

[BROWSE ALL PRACTICE TESTS](#)

[CONTACT FORM](#)

### Important Links

[REFUND POLICY](#)

[REFUND REQUEST](#)

[TERMS & CONDITIONS](#)

[PRIVACY POLICY](#)

[Privacy Policy](#)

## SET-9

### 1. Question

Several employees at your company have been creating projects with Cloud Platform and paying for it with their personal credit cards, which the company reimburses. The company wants to centralize all these projects under a single, new billing account. What should you do?

- Contact cloud-billing@google.com with your bank account details and request a corporate billing account for your company.
- In the Google Cloud Platform Console, create a new billing account and set up a payment method.
- In the Google Platform Console, go to the Resource Manage and move all projects to the root Organization.
- Create a ticket with Google Support and wait for their call to share your credit card details over the phone.

**Incorrect**

Contact [cloud-billing@google.com](mailto:cloud-billing@google.com) with your bank account details and request a corporate billing account for your company. is not right.

?That is not how we set up billing for the organization.

?Ref: <https://cloud.google.com/billing/docs/concepts> Create a ticket with Google Support and wait for their call to share your credit card details over the phone. is not right.

?That is not how we set up billing for the organization.

?Ref: <https://cloud.google.com/billing/docs/concepts> In the Google Cloud Platform Console, create a new billing account and set up a payment method. is not right.

?Unless all projects are modified to use the new billing account, this doesn't work.

?Ref: <https://cloud.google.com/billing/docs/concepts> In the Google Platform Console, go to the Resource Manage and move all projects to the root Organization. is the right answer.

?If we move all projects under the root organization hierarchy, they need to use a billing account within the root organization. We can then consolidate all the costs under different billing accounts as needed e.g. per project, or one for dev work and another billing account for production usage, etc.

?Ref: <https://cloud.google.com/billing/docs/concepts>

### 2. Question

The storage costs for your application logs have far exceeded the project budget. The logs are currently being retained indefinitely in the Cloud Storage bucket

myapp-gcp-ace-logs. You have been asked to remove logs older than 90 days from your Cloud Storage bucket. You want to optimize ongoing Cloud Storage spend. What should you do?

- Write a script that runs gsutil ls -l gs://myapp-gcp-ace-logs/\*\* to find and remove items older than 90 days. Schedule the script with cron.
- Write a script that runs gsutil ls -lr gs://myapp-gcp-ace-logs/\*\* to find and remove items older than 90 days. Repeat this process every morning.
- Write a lifecycle management rule in XML and push it to the bucket with gsutil lifecycle set config-xml-file.
- Write a lifecycle management rule in JSON and push it to the bucket with gsutil lifecycle set config-json-file.**

#### Unattempted

You write a lifecycle management rule in XML and push it to the bucket with gsutil lifecycle set config-xml-file. is not right.

?gsutil lifecycle set enables you to set the lifecycle configuration on one or more buckets based on the configuration file provided. However, XML is not a valid supported type for the configuration file.

?Ref: <https://cloud.google.com/storage/docs/gsutil/commands/lifecycle> Write a script that runs gsutil ls -lr gs://myapp-gcp-ace-logs/\*\* to find and remove items older than 90 days. Repeat this process every morning. is not right.

?This manual approach is error-prone, time-consuming and expensive. GCP Cloud Storage provides lifecycle management rules that let you achieve this with minimal effort. Write a script that runs gsutil ls -l gs://myapp-gcp-ace-logs/\*\* to find and remove items older than 90 days. Schedule the script with cron. is not right.

?This manual approach is error-prone, time-consuming and expensive. GCP Cloud Storage provides lifecycle management rules that let you achieve this with minimal effort. Write a lifecycle management rule in JSON and push it to the bucket with gsutil lifecycle set config-json-file. is the right answer.

?You can assign a lifecycle management configuration to a bucket. The configuration contains a set of rules which apply to current and future objects in the bucket. When an object meets the criteria of one of the rules, Cloud Storage automatically performs a specified action on the object. One of the supported actions is to Delete objects. You can set up a lifecycle management to delete objects older than 90 days. “gsutil lifecycle set” enables you to set the lifecycle configuration on the bucket based on the configuration file. JSON is the only supported type for the configuration file. The config-json-file specified on the command line should be a path to a local file containing the lifecycle configuration JSON document.

?Ref: <https://cloud.google.com/storage/docs/gsutil/commands/lifecycle>

?Ref: <https://cloud.google.com/storage/docs/lifecycle>

3. 3. Question

Users of your application are complaining of slowness when loading the application. You realize the slowness is because the App Engine deployment serving the application is deployed in us-central whereas all users of this application are closest to europe-west3. You want to change the region of the App Engine application to europe-west3 to minimize latency. What's the best way to change the App Engine region?

- Create a new project and create an App Engine instance in europe-west3
- Contact Google Cloud Support and request the change.
- Use the gcloud app region set command and supply the name of the new region.
- From the console, under the App Engine page, click edit, and change the region drop-down.

**Unattempted**

Use the gcloud app region set command and supply the name of the new region. is not right.

?gcloud app region command does not provide a set action. The only action gcloud app region command currently supports is list which lists the availability of flex and standard environments for each region.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/app/regions/list> Contact Google Cloud Support and request the change. is not right.

?Unfortunately, Google Cloud Support isn't of much use here as they would not be able to change the region of an App Engine Deployment. App engine is a regional service, which means the infrastructure that runs your app(s) is located in a specific region and is managed by Google to be redundantly available across all the zones within that region. Once an app engine deployment is created in a region, it can't be changed.

?Ref: <https://cloud.google.com/appengine/docs/locations> From the console, Click edit in App Engine dashboard page and change the region drop-down. is not right.

?The settings mentioned in this option aren't available in the App Engine dashboard. App engine is a regional service. Once an app engine deployment is created in a region, it can't be changed. As shown in the screenshot below, Region is greyed out. Create a new project and create an App Engine instance in europe-west3. is the right answer.

?App engine is a regional service, which means the infrastructure that runs your app(s) is located in a specific region and is managed by Google to be redundantly available across all the zones within that region. Once an app engine deployment is created in a region, it can't be changed. The only way is to create

a new project and create an App Engine instance in europe-west3, send all user traffic to this instance and delete the app engine instance in us-central.

?Ref: <https://cloud.google.com/appengine/docs/locations>

4. 4. Question

You are analyzing Google Cloud Platform service costs from three separate projects. You want to use the information to create service costs estimates grouped by service type, daily and monthly, for the next six months using standard query syntax. What should you do?

- Export your bill to a BigQuery dataset and then write time window based SQL queries for analysis.
- Export your bill to a Cloud Storage bucket and then import into Cloud Bigtable for analysis.
- Export your bill to a Cloud Storage bucket and then import into Google Sheets for analysis
- Export your transactions to a local file and perform analysis with a suitable desktop tool.

**Unattempted**

Requirements

?1. use query syntax

?2. need the billing data of all three projects Export your bill to a Cloud Storage bucket and then import into Cloud Bigtable for analysis. is not right.

?BigTable is a NoSQL database and doesn't offer query syntax support. Export your bill to a Cloud Storage bucket and then import into Google Sheets for analysis. is not right.

?Google Sheets don't offer full support for query syntax. Moreover, export to Cloud Storage bucket captures a smaller dataset than export to BigQuery. For example, the exported billing data does not include resource labels or any invoice-level charges such as taxes accrued or adjustment memos. Export your transactions to a local file and perform analysis with a suitable desktop tool. is not right.

?Billing data can't be exported to a local file, it can only be exported to a BigQuery Dataset or Cloud Storage bucket. Export your bill to a BigQuery dataset and then write time window based SQL queries for analysis. is the right answer.

?You can export billing information from multiple projects into a BigQuery dataset. Unlike the export to Cloud Storage bucket, export to BigQuery dataset includes all information making it easy and straightforward to construct queries in

BigQuery to estimate the cost. BigQuery supports Standard SQL so you can join tables and group by fields (labels in this case) as needed

?Ref: <https://cloud.google.com/billing/docs/how-to/export-data-bigquery>.

5. Question

You are building a new version of an application hosted in an App Engine environment. You want to test the new version with 1% of users before you completely switch your application over to the new version. What should you do?

- Deploy a new version of your application in Google Kubernetes Engine instead of App Engine and then use GCP Console to split traffic.
- Deploy a new version of your application in a Compute Engine instance instead of App Engine and then use GCP Console to split traffic.
- Deploy a new version as a separate app in App Engine. Then configure App Engine using GCP Console to split traffic between the two apps.
- Deploy a new version of your application in App Engine. Then go to App Engine settings in GCP Console and split traffic between the current version and newly deployed versions accordingly.

Unattempted

Deploy a new version of your application in Google Kubernetes Engine instead of App Engine and then use GCP Console to split traffic. is not right.

?When you can achieve this natively in GCP app engine using versions, there is no need to do it outside App Engine.

?Ref: <https://cloud.google.com/appengine/docs/standard/python/splitting-traffic> Deploy a new version of your application in a Compute Engine instance instead of App Engine and then use GCP Console to split traffic. is not right.  
?When you can achieve this natively in GCP app engine using versions, there is no need to do it outside App Engine.

?Ref: <https://cloud.google.com/appengine/docs/standard/python/splitting-traffic> Deploy a new version as a separate app in App Engine. Then configure App Engine using GCP Console to split traffic between the two apps. is not right.

?You can achieve this natively in GCP app engine using versions but App Engine doesn't let you split traffic between apps. If you need to do it between apps, you are probably looking at doing this at the load balancer layer or at the DNS layer – either increasing the cost/complexity or introduce other problems such as caching issues.

?Ref: <https://cloud.google.com/appengine/docs/standard/python/splitting-traffic> Deploy a new version of your application in App Engine. Then go to App

Engine settings in GCP Console and split traffic between the current version and newly deployed versions accordingly. is the right answer.

?GCP App Engine natively offers traffic splitting functionality between versions. You can use traffic splitting to specify a percentage distribution of traffic across two or more of the versions within a service. Splitting traffic allows you to conduct A/B testing between your versions and provides control over the pace when rolling out features.

?Ref: <https://cloud.google.com/appengine/docs/standard/python/splitting-traffic>

## 6. Question

You are building a product on top of Google Kubernetes Engine (GKE). You have a single GKE cluster. For each of your customers, a Pod is running in that cluster, and your customers can run arbitrary code inside their Pod. You want to maximize the isolation between your customers' Pods. What should you do?

- Use Binary Authorization and whitelist only the container images used by your customers' Pods.
- Use the Container Analysis API to detect vulnerabilities in the containers used by your customers' Pods.
- Create a GKE node pool with a sandbox type configured to gvisor. Add the parameter runtimeClassName: gvisor to the specification of your customers' Pods.
- Use the cos\_containerd image for your GKE nodes. Add a nodeSelector with the value cloud.google.com/gke-os-distribution: cos\_containerd to the specification of your customers' Pods.

**Unattempted**

Use Binary Authorization and whitelist only the container images used by your customers' Pods. is not right.

?Binary Authorization is a deploy-time security control that ensures only trusted container images are deployed on Google Kubernetes Engine (GKE). With Binary Authorization, you can require images to be signed by trusted authorities during the development process and then enforce signature validation when deploying. By enforcing validation, you can gain tighter control over your container environment by ensuring only verified images are integrated into the build-and-release process.

?Ref: <https://cloud.google.com/binary-authorization> Use the Container Analysis API to detect vulnerabilities in the containers used by your customers' Pods. is not right.

?Container Analysis is a service that provides vulnerability scanning and metadata storage for software artifacts. The scanning service performs vulnerability scans on images in Container Registry, then stores the resulting metadata and makes it available for consumption through an API. Metadata storage allows storing information from different sources, including vulnerability scanning, other Cloud services, and third-party providers.

?Ref: <https://cloud.google.com/container-registry/docs/container-analysis> Use the cos\_containerd image for your GKE nodes. Add a nodeSelector with the value cloud.google.com/gke-os-distribution: cos\_containerd to the specification of your customers' Pods. is not right.

?The cos\_containerd and ubuntu\_containerdimages let you use containerd as the container runtime in your GKE cluster. This doesn't directly provide the isolation we require.

?<https://cloud.google.com/kubernetes-engine/docs/concepts/using-containerd> Create a GKE node pool with a sandbox type configured to gvisor. Add the parameter runtimeClassName: gvisor to the specification of your customers' Pods. is the right answer.

?GKE Sandbox provides an extra layer of security to prevent untrusted code from affecting the host kernel on your cluster nodes when containers in the Pod execute unknown or untrusted code. Multi-tenant clusters and clusters whose containers run untrusted workloads are more exposed to security vulnerabilities than other clusters. Examples include SaaS providers, web-hosting providers, or other organizations that allow their users to upload and run code. When you enable GKE Sandbox on a node pool, a sandbox is created for each Pod running on a node in that node pool. In addition, nodes running sandboxed Pods are prevented from accessing other Google Cloud services or cluster metadata. Each sandbox uses its own userspace kernel. With this in mind, you can make decisions about how to group your containers into Pods, based on the level of isolation you require and the characteristics of your applications.

?Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/sandbox-pods>

## 7. Question

You are building an application that stores relational data from users. Users across the globe will use this application. Your CTO is concerned about the scaling requirements because the size of the user base is unknown. You need to implement a database solution that can scale with your user growth with minimum configuration changes. Which storage solution should you use?

- Cloud Datastore
- Cloud SQL
- Cloud Spanner
- Cloud Firestore

**Unattempted**

Our requirements are relational data, global users, scaling Cloud Firestore is not right.

?Cloud Firestore is not a relational database. Cloud Firestore is a flexible, scalable database for mobile, web, and server development from Firebase and Google Cloud Platform.

?Ref: <https://firebase.google.com/docs/firestore> Cloud Datastore is not right.

?Cloud Datastore is not a relational database. Datastore is a NoSQL document database built for automatic scaling, high performance, and ease of application development

?Ref: <https://cloud.google.com/datastore/docs/concepts/overview> Cloud SQL is not right.

?While Cloud SQL is a relational database, it does not offer infinite automated scaling with minimum configuration changes. Cloud SQL is a fully-managed database service that makes it easy to set up, maintain, manage, and administer your relational databases on Google Cloud Platform

?Ref: <https://cloud.google.com/sql/docs> Cloud Spanner is the right answer.

?Cloud Spanner is a relational database and is highly scalable. Cloud Spanner is a highly scalable, enterprise-grade, globally-distributed, and strongly consistent database service built for the cloud specifically to combine the benefits of relational database structure with a non-relational horizontal scale. This combination delivers high-performance transactions and strong consistency across rows, regions, and continents with an industry-leading 99.999% availability SLA, no planned downtime, and enterprise-grade security

?<https://cloud.google.com/spanner>

8. 8. Question

You are building an application that will run in your data center. The application will use Google Cloud Platform (GCP) services like AutoML. You created a service account that has appropriate access to AutoML. You need to enable authentication to the APIs from your on-premises environment. What should you do?

- Use service account credentials in your on-premises application.
- Use gcloud to create a key file for the service account that has appropriate permissions.**
- Set up direct interconnect between your data center and Google Cloud Platform to enable authentication for your on-premises applications.
- Go to the IAM & admin console, grant a user account permissions similar to the service account permissions, and use this user account for authentication from your data center.

**Unattempted**

Use service account credentials in your on-premises application. is not right.

?Service accounts do not have passwords

?Ref: <https://cloud.google.com/iam/docs/service-accounts> Go to the IAM & admin console, grant a user account permissions similar to the service account permissions, and use this user account for authentication from your data center. is not right.

?While granting Users a similar set of permissions lets them impersonate service accounts and access all resources the service account has access to, you should use a service account to represent a non-human user that needs to authenticate and be authorized to access data in Google APIs. Typically, service accounts are used in scenarios such as:

?Running workloads on virtual machines (VMs).

?Running workloads on on-premises workstations or data centers that call Google APIs.

?Running workloads that are not tied to the lifecycle of a human user.

?Your application assumes the identity of the service account to call Google APIs so that the users aren't directly involved.

?Ref: <https://cloud.google.com/iam/docs/understanding-service-accounts> Set up direct interconnect between your data center and Google Cloud Platform to enable authentication for your on-premises applications. is not right.

?While setting up interconnect provides a direct physical connection between your on-premises network and Google's network, it doesn't directly help us authenticate our application running in the data center. You can configure Private Google Access for on-premises hosts by sending requests to restricted.googleapis.com and advertise a custom route on cloud router but this only lets you reach Google API and doesn't help with authentication.

?Ref: <https://cloud.google.com/interconnect/docs/support/faq> Use gcloud to create a key file for the service account that has appropriate permissions. is the right answer.

?To use a service account outside of Google Cloud, such as on other platforms or on-premises, you must first establish the identity of the service account. Public/private key pairs provide a secure way of accomplishing this goal. You can create a service account key using the Cloud Console, the gcloud tool, the serviceAccounts.keys.create() method, or one of the client libraries.

?Ref: <https://cloud.google.com/iam/docs/creating-managing-service-account-keys>

9. Question

You are building an archival solution for your data warehouse and have selected Cloud Storage to archive your data. Your users need to be able to access this archived data once a quarter for some regulatory requirements. You want to select a cost-efficient option. Which storage option should you use?

- Coldline Storage**
- Nearline Storage
- Regional Storage
- Multi-Regional Storage

## Unattempted

Nearline Storage. is not right.

?Nearline Storage is a low-cost, highly durable storage service for storing infrequently accessed data. Nearline Storage is ideal for data you plan to read or modify on average once per month or less. Nearline storage is more expensive than Coldline Storage which is more suitable for our requirements.

?<https://cloud.google.com/storage/docs/storage-classes#nearline> Regional Storage. is not right.

?While this would certainly let you access your files once a quarter, it would be too expensive compared to Coldline storage which is more suitable for our requirement.

?<https://cloud.google.com/storage/docs/storage-classes#standard> Multi-Regional Storage. is not right.

?While this would certainly let you access your files once a quarter, it would be too expensive compared to Coldline storage which is more suitable for our requirement.

?<https://cloud.google.com/storage/docs/storage-classes#standard> Coldline Storage. is the right answer.

?Coldline Storage is a very-low-cost, highly durable storage service for storing infrequently accessed data. Coldline Storage is ideal for data you plan to read or modify at most once a quarter. Since we have a requirement to access data once a quarter and want to go with the most cost-efficient option, we should select Coldline Storage.

?Ref: <https://cloud.google.com/storage/docs/storage-classes#coldline>

### 10. Question

You are configuring service accounts for an application that spans multiple projects. Virtual machines (VMs) running in the web-applications project need access to BigQuery datasets in crm-databases project. You want to follow Google-recommended practices to give access to the service account in the web-applications project. What should you do?

- Grant project owner role on web-applications project to the service account in crm-databases project.
- Grant project owner role on crm-databases project to the service account in web-applications project.
- Grant project owner role on crm-databases project and bigquery.dataViewer role to the service account in web-applications.
- Grant bigquery.dataViewer role on crm-databases project to the service account in web-applications.**

### Unattempted

Grant project owner role on web-applications project to the service account in crm-databases project. is not right.

?Our requirement is to identify the access needed for service account in the web-applications project, not the service account in crm-databases project Grant project owner role on crm-databases project to the service account in web-applications project. is not right.

?The primitive project owner role provides permissions to manage all resources within the project. For this scenario, the service account in the web-applications project needs access to BigQuery datasets in crm-databases project. Granting the project owner role would fall foul of least privilege principle.

?Ref: <https://cloud.google.com/iam/docs/recommender-overview> Grant project owner role on crm-databases project and bigquery.dataViewer role to the service account in web-applications. is not right.

?The primitive project owner role provides permissions to manage all resources within the project. For this scenario, the service account in the web-applications project needs access to BigQuery datasets in crm-databases project. Granting the project owner role would fall foul of least privilege principle.

?Ref: <https://cloud.google.com/iam/docs/recommender-overview> Grant bigquery.dataViewer role on crm-databases project to the service account in web-applications. is the right answer.

?bigquery.dataViewer role provides permissions to read the dataset's metadata and list tables in the dataset as well as Read data and metadata from the dataset's tables. This is exactly what we need to fulfil this requirement and follows the least privilege principle.

?Ref: <https://cloud.google.com/iam/docs/understanding-roles#bigquery-roles>

### 11. Question

You are creating a Google Kubernetes Engine (GKE) cluster with a cluster autoscaler feature enabled. You need to make sure that each node of the cluster will run a monitoring pod that sends container metrics to a third-party monitoring solution. What should you do?

- Deploy the monitoring pod in a StatefulSet object.
- Reference the monitoring pod in a Deployment object.
- Reference the monitoring pod in a cluster initializer at the GKE cluster creation time.

- Deploy the monitoring pod in a DaemonSet object.

### Unattempted

Reference the monitoring pod in a Deployment object. is not right.

?In our scenario, we need just 1 instance of the monitoring pod running on each node. Bundling the monitoring pod with a deployment object may result in multiple pod instances on the same node. In GKE, deployments represent a set of multiple, identical Pods with no unique identities. Deployment runs multiple replicas of your application and automatically replaces any instances that fail or become unresponsive. In this way, Deployments help ensure that one or more instances of your application are available to serve user requests.

?<https://cloud.google.com/kubernetes-engine/docs/concepts/deployment> Reference the monitoring pod in a cluster initializer at the GKE cluster creation time. is not right.

?You can not use gcloud init to initialize a monitoring pod. gcloud initializer performs the following setup steps.

?? Authorizes gcloud and other SDK tools to access Google Cloud Platform using your user account credentials, or from an account of your choosing whose credentials are already available.

?? Sets up a new or existing configuration.

?? Sets properties in that configuration, including the current project and optionally, the default Google Compute Engine region and zone you'd like to use.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/init> Deploy the monitoring pod in a StatefulSet object. is not right.

?In GKE, StatefulSets represents a set of Pods with unique, persistent identities and stable hostnames that GKE maintains regardless of where they are scheduled. The state information and other resilient data for any given StatefulSet Pod is maintained in persistent disk storage associated with the StatefulSet. The main purpose of StatefulSets is to set up persistent storage for pods that are deployed across multiple zones.

?Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/statefulset>

?Although persistent volumes can be used, they are limited to two zones and you'd have to get into node affinity if you want to use a persistent volume with a pod on a zone that is not covered by the persistent volumes zones.

?See this for more information <https://kubernetes.io/docs/setup/best-practices/multiple-zones/> Deploy the monitoring pod in a DaemonSet object. is the right answer.

?In GKE, DaemonSets manage groups of replicated Pods and adhere to a one-Pod-per-node model, either across the entire cluster or a subset of nodes. As you add nodes to a node pool, DaemonSets automatically add Pods to the new nodes as needed. So, this is a perfect fit for our monitoring pod.

?<https://cloud.google.com/kubernetes-engine/docs/concepts/daemonset>

?DaemonSets are useful for deploying ongoing background tasks that you need to run on all or certain nodes, and which do not require user intervention.

Examples of such tasks include storage daemons like ceph, log collection daemons like fluentd, and node monitoring daemons like collectd. For example, you could have DaemonSets for each type of daemon run on all of your nodes. Alternatively, you could run multiple DaemonSets for a single type of daemon, but have them use different configurations for different hardware types and resource needs.

## 12. Question

You are deploying an application to a Compute Engine VM in a managed instance group. The application must be running at all times, but only a single instance of the VM should run per GCP project. How should you configure the instance group?

- Set autoscaling to On, set the minimum number of instances to 1, and then set the maximum number of instances to 2.
- Set autoscaling to On, set the minimum number of instances to 1, and then set the maximum number of instances to 1.
- Set autoscaling to Off, set the minimum number of instances to 1, and then set the maximum number of instances to 1.
- Set autoscaling to Off, set the minimum number of instances to 1, and then set the maximum number of instances to 2.

**Unattempted**

### Requirements

?1. Since we need the application running at all times, we need a minimum 1 instance.

?2. Only a single instance of the VM should run, we need a maximum 1 instance.

?3. We want the application running at all times. If the VM crashes due to any underlying hardware failure, we want another instance to be added to MIG so that application can continue to serve requests. We can achieve this by enabling autoscaling. The only option that satisfies these three is Set autoscaling to On, set the minimum number of instances to 1, and then set the maximum number of instances to 1.

?Ref: <https://cloud.google.com/compute/docs/autoscaler>

## 13. Question

You are deploying an application to the App Engine. You want the number of instances to scale based on request rate. You need at least 3 unoccupied instances at all times. Which scaling type should you use?

- Basic Scaling with min\_instances set to 3.

- Manual Scaling with 3 instances.
- Automatic Scaling with min\_idle\_instances set to 3.
- Basic Scaling with max\_instances set to 3.

#### Unattempted

Manual Scaling with 3 instances. is not right.

?Manual scaling uses resident instances that continuously run the specified number of instances regardless of the load level. This allows tasks such as complex initializations and applications that rely on the state of the memory over time. This does not autoscale based on the request rate so doesn't fit our requirements.

?Ref: <https://cloud.google.com/appengine/docs/standard/python/how-instances-are-managed> Basic Scaling with min\_instances set to 3. is not right.

?Basic scaling creates dynamic instances when your application receives requests. Each instance will be shut down when the app becomes idle. Basic scaling is ideal for work that is intermittent or driven by user activity. In absence of any load, the App engine may shut down all instances so it is not suitable for our requirement of " at least 3 instances at all times" .

?Ref: <https://cloud.google.com/appengine/docs/standard/python/how-instances-are-managed> Basic Scaling with max\_instances set to 3. is not right.

?Basic scaling creates dynamic instances when your application receives requests. Each instance will be shut down when the app becomes idle. Basic scaling is ideal for work that is intermittent or driven by user activity. In absence of any load, the App engine may shut down all instances so it is not suitable for our requirement of " at least 3 instances at all times" .

?Ref: <https://cloud.google.com/appengine/docs/standard/python/how-instances-are-managed> Automatic Scaling with min\_idle\_instances set to 3. is the right answer.

?Automatic scaling creates dynamic instances based on request rate, response latencies, and other application metrics. However, if you specify the number of minimum idle instances, that specified number of instances run as resident instances while any additional instances are dynamic.

?Ref: <https://cloud.google.com/appengine/docs/standard/python/how-instances-are-managed>

#### 14. Question

You are designing an application that lets users upload and share photos. You expect your application to grow really fast and you are targeting a worldwide audience. You want to delete uploaded photos after 30 days. You want to minimize costs while ensuring your application is highly available. Which GCP storage solution should you choose?

- Persistent SSD on VM instances.

- Cloud Filestore.
- Multiregional Cloud Storage bucket.
- Cloud Datastore database.

#### Unattempted

Cloud Datastore database. is not right.

?Cloud Datastore is a NoSQL document database built for automatic scaling, high performance, and ease of application development. We want to store objects/files and Cloud Datastore is not a suitable storage option for such data.

?Ref: <https://cloud.google.com/datastore/docs/concepts/overview> Cloud Filestore. is not right.

?Cloud Filestore is a managed file storage service based on NFSv3 protocol. While Cloud Filestore can be used to store images, Cloud Filestore is a zonal service and can not scale easily to support a worldwide audience. Also, Cloud Filestore costs a lot (10 times) more than some of the storage classes offered by Google Cloud Storage.

?Ref: <https://cloud.google.com/filestore>,

Ref: <https://cloud.google.com/storage/pricing> Persistent SSD on VM instances. is not right.

?Persistent SSD is a regional service and doesn't automatically scale to other regions to support a worldwide user base. Moreover, Persistent SSD disks are very expensive. A regional persistent SSD costs \$0.34 per GB per month. In comparison, Google Cloud Storage offers several storage classes that are significantly cheaper.

?Ref: <https://cloud.google.com/persistent-disk>

?Ref: <https://cloud.google.com/filestore/pricing> Multiregional Cloud Storage bucket. is the right answer.

?Cloud Storage allows world-wide storage and retrieval of any amount of data at any time. We don't need to set up auto-scaling ourselves. Cloud Storage autoscaling is managed by GCP. Cloud Storage is an object store so it is suitable for storing photos. Cloud Storage allows world-wide storage and retrieval so cater well to our worldwide audience. Cloud storage provides us lifecycle rules that can be configured to automatically delete objects older than 30 days. This also fits our requirements. Finally, Google Cloud Storage offers several storage classes such as Nearline Storage (\$0.01 per GB per Month) Coldline Storage (\$0.007 per GB per Month) and Archive Storage (\$0.004 per GB per month) which are significantly cheaper than any of the options above.

?Ref: <https://cloud.google.com/storage/docs>

?Ref: <https://cloud.google.com/storage/pricing>

You are designing an application that uses WebSockets and HTTP sessions that are not distributed across the web servers. You want to ensure the application runs properly on Google Cloud Platform. What should you do?

- Meet with the cloud enablement team to discuss load balancer options.**
- Redesign the application to use a distributed user session service that does not rely on WebSockets and HTTP sessions.
- Review the encryption requirements for WebSocket connections with the security team.
- Convert the WebSocket code to use HTTP streaming.

**Unattempted**

Google HTTP(S) Load Balancing has native support for the WebSocket protocol when you use HTTP or HTTPS, not HTTP/2, as the protocol to the backend.

?Ref: [https://cloud.google.com/load-balancing/docs/https#websocket\\_proxy\\_support](https://cloud.google.com/load-balancing/docs/https#websocket_proxy_support)

?So the next possible step is to Meet with the cloud enablement team to discuss load balancer options.

?We don't need to convert WebSocket code to use HTTP streaming or Redesign the application, as WebSocket support is offered by Google HTTP(S) Load Balancing. Reviewing the encryption requirements is a good idea but it has nothing to do with WebSockets.

## 16. Question

You are given a project with a single virtual private cloud (VPC) and a single subnet in the us-central1 region. There is a Compute Engine instance hosting an application in this subnet. You need to deploy a new instance in the same project in the europe-west1 region. This new instance needs access to the application. You want to follow Google-recommended practices. What should you do?

- 1. Create a VPC and a subnet in europe-west1. 2. Expose the application with an internal load balancer. 3. Create the new instance in the new subnet and use the load balancer's address as the endpoint.
- 1. Create a VPC and a subnet in europe-west1. 2. Peer the 2 VPCs. 3. Create the new instance in the new subnet and use the first instance's private address as the endpoint.
- 1. Create a subnet in the same VPC, in europe-west1. 2. Create the new instance in the new subnet and use the first instance subnet's private address as the endpoint.
- 1. Create a subnet in the same VPC, in europe-west1. 2. Use Cloud VPN to connect the two subnets. 3. Create the new instance in the new subnet and use the first instance's private address as the endpoint.

## Unattempted

Our requirements are to connect the instance in europe-west1 region with the application running in us-central1 region following Google-recommended practices. The two instances are in the same project.

1. Create a VPC and a subnet in europe-west1.

?2. Expose the application with an internal load balancer.

?3. Create the new instance in the new subnet and use the load balancer's address as the endpoint. is not right.

?We have two different VPCs. There is no mention of the CIDR range so let's assume the two subnets in two VPCs use different CIDR ranges. However, there is no communication route between the two VPCs. If we create an internal load balancer, that load balancer is not visible outside the VPC. So the new instance cannot connect to the load balancer's internal address.

?Ref: <https://cloud.google.com/load-balancing/docs/internal>

1. Create a subnet in the same VPC, in europe-west1.

?2. Use Cloud VPN to connect the two subnets.

?3. Create the new instance in the new subnet and use the first instance's private address as the endpoint. is not right.

?Cloud VPN securely connects your on-premises network to your Google Cloud (GCP) Virtual Private Cloud (VPC) network through an IPsec VPN connection. It is not meant to connect two subnets within the same VPC. Moreover, subnets within the same VPC can communicate with each other by setting up relevant firewall rules.

1. Create a VPC and a subnet in europe-west1.

?2. Peer the 2 VPCs.

?3. Create the new instance in the new subnet and use the first instance's private address as the endpoint. is not right.

?Given that the new instance wants to access the application on the existing compute engine instance, these applications seem to be related so they should be within the same VPC. It is possible to have them in different VPCs and peer the VPCs but this is a lot of additional work and we can simplify this by choosing the option below (which is the answer)

1. Create a subnet in the same VPC, in europe-west1.

?2. Create the new instance in the new subnet and use the first instance subnet's private address as the endpoint. is the right answer.

?We can create another subnet in the same VPC and this subnet is located in europe-west1. We can then spin up a new instance in this subnet. We also have to set up a firewall rule to allow communication between the two subnets. All instances in the two subnets with the same VPC can communicate through the internal IP Address

?Ref: <https://cloud.google.com/vpc>

17. Question

You are hosting an application on bare metal servers in your data center. The application needs access to Cloud Storage. However, security policies prevent the servers hosting the application from having public IP addresses or access to the internet. You want to follow Google recommended practices to provide the application with access to Cloud Storage. What should you do?

- Use nslookup to get the IP addresses for storage.googleapis.com Negotiate with the security team to be able to give public IP addresses to the servers. Only allow egress traffic from those servers to the IP addresses for storage.googleapis.com
- Using Cloud VPN, create a VPN tunnel to a Virtual Private Cloud (VPC) in Google Cloud Platform (GCP). In this VPC, create a Compute Engine instance and install the Squid proxy server on this instance. Configure your servers to use that instance as a proxy to access cloud storage
- Use Migrate for Compute Engine (formerly known as Velostrata) to migrate these servers to Compute Engine. Create an internal load balancer (ILB) that uses storage.googleapis.com as backend. Configure your new instances to use the ILB as a proxy
- Using Cloud VPN or Interconnect, create a tunnel to a VPC in GCP. Using Cloud Router to create a custom route advertisement for 199.36.153.4/30. Announce that network to your on-premises network through the VPN tunnel. In your on-premises network, configure your DNS server to resolve \*.googleapis.com as a CNAME to restricted.googleapis.com

Unattempted

Our requirement is to follow Google recommended practices to achieve the end result. Configuring Private Google Access for On-Premises Hosts is best achieved by VPN/Interconnect + Advertise Routes + Use restricted Google IP Range.

?Using Cloud VPN or Interconnect, create a tunnel to a VPC in GCP

?Using Cloud Router to create a custom route advertisement for 199.36.153.4/30. Announce that network to your on-premises network through the VPN tunnel.

?In your on-premises network, configure your DNS server to resolve \*.googleapis.com as a CNAME to restricted.googleapis.com is the right answer right, and it is what Google recommends.

?Ref: <https://cloud.google.com/vpc/docs/configure-private-google-access-hybrid>

?“ You must configure routes so that Google API traffic is forwarded through your Cloud VPN or Cloud Interconnect connection, firewall rules on your on-premises firewall to allow the outgoing traffic, and DNS so that traffic to Google APIs resolves to the IP range you’ve added to your routes.”

?“ You can use Cloud Router Custom Route Advertisement to announce the Restricted Google APIs IP addresses through Cloud Router to your on-premises network. The Restricted Google APIs IP range is 199.36.153.4/30. While this is technically a public IP range, Google does not announce it publicly. This IP range is only accessible to hosts that can reach your Google Cloud projects through internal IP ranges, such as through a Cloud VPN or Cloud Interconnect connection.” Without having a public IP address or access to the internet, the only way you could connect to cloud storage is if you have an internal route to it. So Negotiate with the security team to be able to give public IP addresses to the servers is not right. Following “ Google recommended practices” is synonymous with “ using Google’s services” (Not quite, but it is – at least for the exam !!). So In this VPC, create a Compute Engine instance and install the Squid proxy server on this instance is not right. Migrating the VM to Compute Engine is a bit drastic when Google says it is perfectly fine to have Hybrid Connectivity architectures <https://cloud.google.com/hybrid-connectivity>. So, Use Migrate for Compute Engine (formerly known as Velostrata) to migrate these servers to Compute Engine is not right.

## 18. Question

You are managing several Google Cloud Platform (GCP) projects and need access to all logs for the past 60 days. You want to be able to explore and quickly analyze the log contents. You want to follow Google-recommended practices to obtain the combined logs for all projects. What should you do?

- Navigate to Stackdriver Logging and select resource.labels.project\_id="\*"
- Create a Stackdriver Logging Export with a Sink destination to a BigQuery dataset. Configure the table expiration to 60 days.
- Create a Stackdriver Logging Export with a Sink destination to Cloud Storage. Create a lifecycle rule to delete objects after 60 days.
- Configure a Cloud Scheduler job to read from Stackdriver and store the logs in BigQuery. Configure the table expiration to 60 days.

**Unattempted**

Navigate to Stackdriver Logging and select resource.labels.project\_id="\*" . is not right.

Log entries are held in Stackdriver Logging for a limited time known as the retention period – which is 30 days (default configuration). After that, the entries

are deleted. To keep log entries longer, you need to export them outside of Stackdriver Logging by configuring log sinks.

<https://cloud.google.com/blog/products/gcp/best-practices-for-working-with-google-cloud-audit-logging>

Configure a Cloud Scheduler job to read from Stackdriver and store the logs in BigQuery. Configure the table expiration to 60 days. is not right.

While this works, it makes no sense to use Cloud Scheduler job to read from Stackdriver and store the logs in BigQuery when Google provides a feature (export sinks) that does exactly the same thing and works out of the box.

Ref: [https://cloud.google.com/logging/docs/export/configure\\_export\\_v2](https://cloud.google.com/logging/docs/export/configure_export_v2)

Create a Stackdriver Logging Export with a Sink destination to Cloud Storage.

Create a lifecycle rule to delete objects after 60 days. is not right.

You can export logs by creating one or more sinks that include a logs query and an export destination. Supported destinations for exported log entries are Cloud Storage, BigQuery, and Pub/Sub.

Ref: [https://cloud.google.com/logging/docs/export/configure\\_export\\_v2](https://cloud.google.com/logging/docs/export/configure_export_v2)

Sinks are limited to exporting log entries from the exact resource in which the sink was created: a Google Cloud project, organization, folder, or billing account. If it makes it easier to exporting from all projects of an organization, you can create an aggregated sink that can export log entries from all the projects, folders, and billing accounts of a Google Cloud organization.

[https://cloud.google.com/logging/docs/export/aggregated\\_sinks](https://cloud.google.com/logging/docs/export/aggregated_sinks)

Either way, we now have the data in Cloud Storage, but querying logs information from Cloud Storage is harder than Querying information from BigQuery dataset.

For this reason, we should prefer Big Query over Cloud Storage.

Create a Stackdriver Logging Export with a Sink destination to a BigQuery dataset. Configure the table expiration to 60 days. is the right answer.

You can export logs by creating one or more sinks that include a logs query and an export destination. Supported destinations for exported log entries are Cloud Storage, BigQuery, and Pub/Sub.

Ref: [https://cloud.google.com/logging/docs/export/configure\\_export\\_v2](https://cloud.google.com/logging/docs/export/configure_export_v2)

Sinks are limited to exporting log entries from the exact resource in which the sink was created: a Google Cloud project, organization, folder, or billing account. If it makes it easier to exporting from all projects of an organization, you can create an aggregated sink that can export log entries from all the projects, folders, and billing accounts of a Google Cloud organization.

[https://cloud.google.com/logging/docs/export/aggregated\\_sinks](https://cloud.google.com/logging/docs/export/aggregated_sinks)

Either way, we now have the data in a BigQuery Dataset. Querying information from a Big Query dataset is easier and quicker than analyzing contents in Cloud Storage bucket. As our requirement is to “ Quickly analyze the log contents” , we should prefer Big Query over Cloud Storage.

Also, You can control storage costs and optimize storage usage by setting the default table expiration for newly created tables in a dataset. If you set the property when the dataset is created, any table created in the dataset is deleted after the expiration period. If you set the property after the dataset is created, only new tables are deleted after the expiration period.

For example, if you set the default table expiration to 7 days, older data is automatically deleted after 1 week.

Ref: <https://cloud.google.com/bigquery/docs/best-practices-storage>

You are migrating a mission critical on-premises application to cloud. The application requires 96 vCPUs to perform its task. You want to make sure the application runs in a similar environment on GCP. What should you do?

- When creating the VM, use machine type n1-standard-96.
- When creating the VM, use Intel Skylake as the CPU platform.
- Create the VM using Compute Engine default settings. Use gcloud to modify the running instance to have 96 vCPUs.
- Start the VM using Compute Engine default settings, and adjust as you go based on Rightsizing Recommendations.

#### Unattempted

Create the VM using Compute Engine default settings. Use gcloud to modify the running instance to have 96 vCPUs. is not right.

?You can't increase the vCPUs to 96 without changing the machine type. While it is possible to set machine type using gcloud, this would mean downtime for the mission-critical application while the upgrade happens which is undesirable.

?Ref: <https://cloud.google.com/compute/docs/instances/changing-machine-type-of-stopped-instance> Start the VM using Compute Engine default settings, and adjust as you go based on Rightsizing Recommendations. is not right.

?Since the application is mission-critical, we want to ensure that this application has all the required resources from the beginning. Starting with the default settings provisions a n1-standard-1 machine that has just 1 vCPU and our mission-critical application would be severely constrained for resources. When creating the VM, use Intel Skylake as the CPU platform. is not right.

?Intel Skylake is only offered in E2 machine types that are cost-optimized machine types and offer sizing between 2 to 16 vCPUs which is insufficient for our mission-critical application.

?Ref: [https://cloud.google.com/compute/docs/machine-types#e2\\_machine\\_types](https://cloud.google.com/compute/docs/machine-types#e2_machine_types) When creating the VM, use machine type n1-standard-96. is the right answer.

?n1-standard-96 offers 96 vCPUs and 624 GB of memory. This fits our requirements.

?[https://cloud.google.com/compute/docs/machine-types#n1\\_machine\\_type](https://cloud.google.com/compute/docs/machine-types#n1_machine_type)

#### 20. Question

You are operating a Google Kubernetes Engine (GKE) cluster for your company where different teams can run non-production workloads. Your Machine Learning (ML) team needs access to Nvidia Tesla P100 GPUs to train their models. You want to minimize effort and cost. What should you do?

- Ask your ML team to add the "accelerator: gpu" annotation to their pod specification.
- Recreate all the nodes of the GKE cluster to enable GPUs on all of them.
- Create your own Kubernetes cluster on top of Compute Engine with nodes that have GPUs. Dedicate this cluster to your ML team.
- Add a new, GPU-enabled, node pool to the GKE cluster. Ask your ML team to add the cloud.google.com/gke-accelerator: nvidia-tesla-p100 nodeSelector to their pod specification.

### Unattempted

Ask your ML team to add the “ accelerator: gpu” annotation to their pod specification. is not right.

There are two issues with this approach. One – the syntax is invalid. Two – You cannot add GPUs to existing node pools.

Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/gpus>

Recreate all the nodes of the GKE cluster to enable GPUs on all of them. is not right.

There are two issues with this approach. One – recreating all nodes to enable GPUs makes the cluster very expensive. Only the ML team needs access to GPUs to train their models. Recreating all nodes to enable GPUs helps your ML team use them but they are left unused for all other workloads yet cost you money. Two – Even though your nodes have GPUs enabled, you still have to modify pod specifications to request GPU. This step isn’ t performed in this option.

Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/gpus>

Create your own Kubernetes cluster on top of Compute Engine with nodes that have GPUs. Dedicate this cluster to your ML team. is not right.

While this works, it increases the cost as you now pay the Kubernetes cluster management fee for two clusters instead of one. GKE clusters accrue a management fee that is per cluster per hour, irrespective of cluster size or topology.

Ref: <https://cloud.google.com/kubernetes-engine/pricing>

Add a new, GPU-enabled, node pool to the GKE cluster. Ask your ML team to add the cloud.google.com/gke-accelerator: nvidia-tesla-p100 nodeSelector to their pod specification. is the right answer.

This is the most optimal solution. Rather than recreating all nodes, you create a new node pool with GPU enabled. You then modify the pod specification to target particular GPU types by adding node selector to your workload’ s Pod specification. YOu still have a single cluster so you pay Kubernetes cluster management fee for just one cluster thus minimizing the cost.

Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/gpus>

Ref: <https://cloud.google.com/kubernetes-engine/pricing>

Example:

apiVersion: v1

kind: Pod

```

metadata:
name: my-gpu-pod
spec:
containers:
- name: my-gpu-container
image: nvidia/cuda:10.0-runtime-ubuntu18.04
command: ["/bin/bash"]
resources:
limits:
nvidia.com/gpu: 2
nodeSelector:
cloud.google.com/gke-accelerator: nvidia-tesla-k80 # or nvidia-tesla-p100 or
nvidia-tesla-p4 or nvidia-tesla-v100 or nvidia-tesla-t4

```

## 21. Question

You are running an application on multiple virtual machines within a managed instance group and have auto-scaling enabled. The autoscaling policy is configured so that additional instances are added to the group if the CPU utilization of instances goes above 80%. VMs are added until the instance group reaches its maximum limit of five VMs or until CPU utilization of instances lowers to 80%. The initial delay for HTTP health checks against the instances is set to 30 seconds. The virtual machine instances take around three minutes to become available for users. You observe that when the instance group auto-scales, it adds more instances than necessary to support the levels of end-user traffic. You want to properly maintain instance group sizes when autoscaling. What should you do?

- Decrease the maximum number of instances to 3.
- Increase the initial delay of the HTTP health check to 200 seconds.
- Set the maximum number of instances to 1.
- Use a TCP health check instead of an HTTP health check.

**Unattempted**

### Scenario

- ? Autoscaling is enabled and kicks off the scale-up
- ? Scaling policy is based on target CPU utilization of 80%
- ? The initial delay is 30 seconds
- ? VM startup time is 3 minutes.
- ? Auto-scaling creates more instances than necessary.

Set the maximum number of instances to 1. is not right.

Setting the maximum number of instances to 1 effectively limits the scale up to 1 instance which is undesirable as in this case we may still be struggling with the CPU usage but we can't scale up. Therefore this is not the right answer.

Decrease the maximum number of instances to 3. is not right.

Setting the maximum number of instances to 3 effectively limits the scale up to 3

instances which is undesirable as in this case we may still be struggling with the CPU usage but we can't scale up. Therefore this is not the right answer.

Use a TCP health check instead of an HTTP health check. is not right.

TCP health check is a legacy health check, whereas HTTP health check is more advanced and “non-legacy”. It is possible a TCP health check might say the application is UP when it is not as it only listens on application servers TCP port and doesn't validate the application health through a HTTP check on its health endpoint. This results in the load balancer sending requests to the application server when it is still loading the application resulting in failures.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/health-checks/create/tcp>

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/health-checks/create/http>

Increase the initial delay of the HTTP health check to 200 seconds. is the right answer.

The reason why our autoscaling is adding more instances than needed is that it checks 30 seconds after launching the instance and at this point, the instance isn't up and isn't ready to serve traffic. So our autoscaling policy starts another instance – again checks this after 30 seconds and the cycle repeats until it gets to the maximum instances or the instances launched earlier are healthy and start processing traffic – which happens after 180 seconds (3 minutes). This can be easily rectified by adjusting the initial delay to be higher than the time it takes for the instance to become available for processing traffic.

So setting this to 200 ensures that it waits until the instance is up (around 180-second mark) and then starts forwarding traffic to this instance. Even after a cool out period, if the CPU utilization is still high, the autoscaler can again scale up but this scale-up is genuine and is based on the actual load.

“Initial Delay Seconds” – This setting delays autohealing from potentially prematurely recreating the instance if the instance is in the process of starting up. The initial delay timer starts when the currentAction of the instance is VERIFYING.

Ref: <https://cloud.google.com/compute/docs/instance-groups/autohealing-instances-in-migs>

## 22. Question

You are setting up a Windows VM on Compute Engine and want to make sure you can log in to the VM via RDP. What should you do?

- After the VM has been created, use your Google Account credentials to log in into the VM.
- After the VM has been created, use gcloud compute reset-windows-password to retrieve the login credentials for the VM.**
- When creating the VM, add metadata to the instance using 'windows-password' as the key and a password as the value.

- After the VM has been created, download the JSON private key for the default Compute Engine service account. Use the credentials in the JSON file to log in to the VM.

### Unattempted

When creating the VM, add metadata to the instance using ‘ windows-password’ as the key and a password as the value. is not right.

?It is not possible to specify a windows password at the time of creating windows VM instance. You can generate Windows passwords using either the Google Cloud Console or the gcloud command-line tool. Alternatively, you can generate passwords programmatically with API commands but all these methods assume that you have an existing windows instance.

?Ref: <https://cloud.google.com/compute/docs/instances/windows/creating-passwords-for-windows-instances#gcloud> After the VM has been created, use your Google Account credentials to log in into the VM. is not right.

?You can generate Windows passwords using either the Google Cloud Console or the gcloud command-line tool. Alternatively, you can generate passwords programmatically with API commands but you can’t use your gcloud account credentials to log into the VM.

?Ref: <https://cloud.google.com/compute/docs/instances/windows/creating-passwords-for-windows-instances#gcloud> After the VM has been created, download the JSON private key for the default Compute Engine service account. Use the credentials in the JSON file to log in to the VM. is not right.

?This is not a supported method of authentication for logging into the VM. You can generate Windows passwords using either the Google Cloud Console or the gcloud command-line tool. Alternatively, you can generate passwords programmatically with API commands.

?Ref: <https://cloud.google.com/compute/docs/instances/windows/creating-passwords-for-windows-instances#gcloud> After the VM has been created, use gcloud compute reset-windows-password to retrieve the login credentials for the VM. is the right answer.

?You can generate Windows passwords using either the Google Cloud Console or the gcloud command-line tool. This option uses the right syntax to reset the windows password.

?gcloud compute reset-windows-password windows-instance

?Ref: <https://cloud.google.com/compute/docs/instances/windows/creating-passwords-for-windows-instances#gcloud>

### 23. Question

You are the organization and billing administrator for your company. The engineering team has the Project Creator role at the organization level. You do not want the engineering team to be able to link projects to the billing account. Only the finance team should be able to link a project to a billing account, but

they should not be able to make any other changes to projects. What should you do?

- Assign the engineering team the Billing Account User role on the billing account and the Project Billing Manager role on the organization.
- Assign the finance team only the Billing Account User role on the billing account.
- Assign the engineering team only the Billing Account User role on the billing account.
- Assign the finance team the Billing Account User role on the billing account and the Project Billing Manager role on the organization.**

#### Unattempted

Assign the finance team only the Billing Account User role on the billing account. is not right.

?In order to link a project to a billing account, you need the necessary roles at the project level as well as at the billing account level. In this scenario, we are granting just the Billing Account User role on the billing account to the Finance team which allows them to link projects to the billing account on which the role is granted. But we haven't granted them any role at the project level. So they would not be unable to link projects. Assign the engineering team only the Billing Account User role on the billing account. is not right.

?In order to link a project to a billing account, you need the necessary roles at the project level as well as at the billing account level. In this scenario, we are granting just the Billing Account User role on the billing account to the Engineering team which allows them to link projects to the billing account and our question clearly states we do not want to do that. Assign the engineering team the Billing Account User role on the billing account and the Project Billing Manager role on the organization. is not right.

?In order to link a project to a billing account, you need the necessary roles at the project level as well as at the billing account level. In this scenario, we are assigning the engineering team the Billing Account User role on the billing account which allows them to create new projects linked to the billing account on which the role is granted. We are also assigning them the Project Billing Manager role on the organization (trickles down to the project as well) which lets them attach the project to the billing account. But we don't want the engineering team to link projects to the billing account. Assign the finance team the Billing Account User role on the billing account and the Project Billing Manager role on the organization. is the right answer.

?In order to link a project to a billing account, you need the necessary roles at the project level as well as at the billing account level. In this scenario, we are assigning the finance team the Billing Account User role on the billing account which allows them to create new projects linked to the billing account on which

the role is granted. We are also assigning them the Project Billing Manager role on the organization (trickles down to the project as well) which lets them attach the project to the billing account, but does not grant any rights over resources. This is exactly what we want.

#### 24. Question

You are the project owner of a GCP project and want to delegate control to colleagues to manage buckets and files in Cloud Storage. You want to follow Google recommended practices. Which IAM roles should you grant your colleagues?

- Project Editor
- Storage Object Creator
- Storage Admin
- Storage Object Admin

**Unattempted**

Project Editor is not right. is not right.

?Project editor is a primitive role that grants a lot more than what we need here. Google doesn't recommend using Primitive roles.

?Ref: [https://cloud.google.com/iam/docs/understanding-roles#primitive\\_role\\_definitions](https://cloud.google.com/iam/docs/understanding-roles#primitive_role_definitions)

?All viewer permissions, plus permissions for actions that modify state, such as changing existing resources. Storage Object Admin. is not right.

?While this role grants full access to the objects, it does not grant access to the buckets so users of this role can not “ manage buckets” .

?This role grants full control over objects, including listing, creating, viewing, and deleting objects.

?Ref: <https://cloud.google.com/iam/docs/understanding-roles#storage-roles> Storage Object Creator. is not right.

?This role allows users to create objects. It does not give permission to view, delete, or overwrite objects.

?Ref: <https://cloud.google.com/iam/docs/understanding-roles#storage-roles> Storage Admin. is the right answer.

?This role grants full control of buckets and objects. When applied to an individual bucket, control applies only to the specified bucket and objects within the bucket.

?Ref: <https://cloud.google.com/iam/docs/understanding-roles#storage-roles>

## 25. Question

You are using Container Registry to centrally store your company's container images in a separate project. In another project, you want to create a Google Kubernetes Engine (GKE) cluster. You want to ensure that Kubernetes can download images from Container Registry. What should you do?

- In the project where the images are stored, grant the Storage Object Viewer IAM role to the service account used by the Kubernetes nodes.
- When you create the GKE cluster, choose the Allow full access to all Cloud APIs option under 'Access scopes'.
- Create a service account, and give it access to Cloud Storage. Create a P12 key for this service account and use it as an imagePullSecrets in Kubernetes.
- Configure the ACLs on each image in Cloud Storage to give read-only access to the default Compute Engine service account.

### Unattempted

Here's some info about where Container Registry stores images and how access is controlled.

Container Registry uses Cloud Storage buckets as the underlying storage for container images. You control access to your images by granting appropriate Cloud Storage permissions to a user, group, service account, or another identity. Cloud Storage permissions granted at the project level apply to all storage buckets in the project, not just the buckets used by Container Registry. To configure permissions specific to Container Registry, grant permissions on the storage bucket used by the registry. Container Registry ignores permissions set on individual objects within the storage bucket.

Ref: <https://cloud.google.com/container-registry/docs/access-control>

Configure the ACLs on each image in Cloud Storage to give read-only access to the default Compute Engine service account. is not right.

As mentioned above, Container Registry ignores permissions set on individual objects within the storage bucket so this isn't going to work.

Ref: <https://cloud.google.com/container-registry/docs/access-control>

When you create the GKE cluster, choose the Allow full access to all Cloud APIs option under ' Access scopes' . is not right.

Selecting Allow full access to all Cloud APIs does not provide access to GCR images in a different project. If the Google Kubernetes Engine cluster and the Container Registry storage bucket are in the same Google Cloud project, the Compute Engine default service account is configured with the appropriate permissions to push or pull images. But if the cluster is in a different project or if the VMs in the cluster use a different service account, you must grant the service account the appropriate permissions to access the storage bucket used by Container Registry.

Ref: <https://cloud.google.com/container-registry/docs/using-with-google-cloud-platform>

In this case, since there is no mention of a service account, we have to assume we are using a default service account that hasn't been provided permissions to access the storage bucket used by Container Registry in another project so the image pull isn't going to work. You would end up with an error like:  
Failed to pull image " gcr.io/kubernetes2-278322/simple-python-image" : rpc error: code = Unknown desc = Error response from daemon: pull access denied for gcr.io/kubernetes2-278322/simple-python-image, repository does not exist or may require ' docker login'

Create a service account, and give it access to Cloud Storage. Create a P12 key for this service account and use it as an imagePullSecrets in Kubernetes. is not right.

It is technically possible to do it this way but using the JSON key and not P12 key as mentioned in this option. If you would like to understand how to do this, please look at these blogs.

Ref: <https://medium.com/hackeroon/today-i-learned-pull-docker-image-from-gcr-google-container-registry-in-any-non-gcp-kubernetes-5f8298f28969>

Ref: <https://medium.com/@michaelmorrissey/using-cross-project-gcr-images-in-gke-1ddc36de3d42>

Moreover, this approach is suitable for accessing GCR images in a non-Google Cloud Kubernetes environment. While it can be used in GKE too, it is not as secure as using Role Bindings since it involves downloading service account keys and setting them up as secret in Kubernetes.

In the project where the images are stored, grant the Storage Object Viewer IAM role to the service account used by the Kubernetes nodes. is the right answer.

Granting the storage object viewer IAM role in the project where images are stored to the service account used by the Kubernetes cluster ensures that the nodes in the cluster can Read Images from the storage bucket. It would be ideal to further restrict the role binding to provide access just to the Cloud Storage bucket that is used as the underlying storage for container images. This follows the principle of least privilege.

For more information about Storage Object Viewer IAM Role for GCR refer: [https://cloud.google.com/container-registry/docs/access-control#permissions\\_and\\_roles](https://cloud.google.com/container-registry/docs/access-control#permissions_and_roles)

## 26. Question

You are using Deployment Manager to create a Google Kubernetes Engine cluster. Using the same Deployment Manager deployment, you also want to create a DaemonSet in the kube-system namespace of the cluster. You want a solution that uses the fewest possible services. What should you do?

- Add the cluster's API as a new Type Provider in Deployment Manager, and use the new type to create the DaemonSet.
- Use the Deployment Manager Runtime Configurator to create a new Config resource that contains the DaemonSet definition.
- With Deployment Manager, create a Compute Engine instance with a startup script that uses kubectl to create the DaemonSet.

- In the cluster's definition in Deployment Manager, add a metadata that has kube-system as key and the DaemonSet manifest as value.

### Unattempted

In the cluster' s definition in Deployment Manager, add a metadata that has kube-system as key and the DaemonSet manifest as value. is not right.

?Metadata entries are key-value pairs and do not influence this behavior.

?Ref: <https://cloud.google.com/compute/docs/storing-retrieving-metadata> With Deployment Manager, create a Compute Engine instance with a startup script that uses kubectl to create the DaemonSet. is not right.

?It is possible to spin up a compute engine instance with a startup script that executes kubectl to create a DaemonSet deployment.

?kubectl apply -f <https://k8s.io/examples/controllers/daemonset.yaml>

?Ref: <https://kubernetes.io/docs/concepts/workloads/controllers/daemonset/>

?But this involves using the compute engine service which is an additional service. Our requirement is to achieve using the fewest possible services and as you' ll notice later, the correct answer uses fewer services. Use the Deployment Manager Runtime Configurator to create a new Config resource that contains the DaemonSet definition. is not right.

?You can configure the GKE nodes (provisioned by Deployment manager) to report their status to the Runtime Configurator, and when they are UP, you can run a task to create a DaemonSet. While this is possible, it involves one additional service – to run a task e.g. using Cloud Functions, etc. Our requirement is to achieve using the fewest possible services and as you' ll notice later, the correct answer uses fewer services.

?Here is some more info about Runtime Configurator. The Runtime Configurator feature lets you define and store data as a hierarchy of key-value pairs in Google Cloud Platform. You can use these key-value pairs as a way to:

?1. Dynamically configure services

?2. Communicate service states

?3. Send notification of changes to data

?4. Share information between multiple tiers of services

?For example, imagine a scenario where you have a cluster of nodes that run a startup procedure. During startup, you can configure your nodes to report their status to the Runtime Configurator, and then have another application query the Runtime Configurator and run specific tasks based on the status of the nodes.

?The Runtime Configurator also offers a Watcher service and a Waiter service. The Watcher service watches a specific key pair and returns when the value of the key pair changes, while the Waiter service waits for a specific end condition and returns a response once that end condition has been met.

?Ref: <https://cloud.google.com/deployment-manager/runtime-configurator> Add the cluster' s API as a new Type Provider in Deployment Manager, and use the new type to create the DaemonSet. is the right answer.

?A type provider exposes all resources of a third-party API to Deployment Manager as base types that you can use in your configurations. If you have a cluster running on Google Kubernetes Engine, you could add the cluster as a type provider and access the Kubernetes API using Deployment Manager. Using these inherited API, you can create a DaemonSet.

?This option uses just the Deployment Manager to create a DaemonSet and is, therefore, the right answer.

?Ref: <https://cloud.google.com/deployment-manager/docs/configuration/type-providers/creating-type-provider>

## 27. Question

You are using Google Kubernetes Engine with autoscaling enabled to host a new application. You want to expose this new application to the public, using HTTPS on a public IP address. What should you do?

- Create a Kubernetes Service of type NodePort for your application, and a Kubernetes Ingress to expose this Service via a Cloud Load Balancer.
- Create a Kubernetes Service of type ClusterIP for your application. Configure the public DNS name of your application using the IP of this Service.
- Create a Kubernetes Service of type NodePort to expose the application on port 443 of each node of the Kubernetes cluster. Configure the public DNS name of your application with the IP of every node of the cluster to achieve load-balancing.
- Create a HAProxy pod in the cluster to load-balance the traffic to all the pods of the application. Forward the public traffic to HAProxy with an iptable rule. Configure the DNS name of your application using the public IP of the nodeHAProxy is running on.

**Unattempted**

Create a Kubernetes Service of type ClusterIP for your application. Configure the public DNS name of your application using the IP of this Service. is not right.

?Kubernetes Service of type ClusterIP exposes the Service on a cluster-internal IP. Choosing this value makes the Service only reachable from within the cluster so you can not route external traffic to this IP.

?Ref: <https://kubernetes.io/docs/concepts/services-networking/service/> Create a HAProxy pod in the cluster to load-balance the traffic to all the pods of the application. Forward the public traffic to HAProxy with an iptable rule. Configure the DNS name of your application using the public IP of the nodeHAProxy is running on. is not right.

?HAProxy is a popular Kubernetes ingress controller. An Ingress object is an independent resource, apart from Service objects, that configures external access to a service's pods. Ingress Controllers still need a way to receive external traffic. This can be done by exposing the Ingress Controller as a Kubernetes service with either NodePort or LoadBalancer type. You can't use public IP of the node the HAProxy is running on as this may be running in any node in the Kubernetes Cluster and in most cases, these nodes do not have public IPs. They are meant to be private and the pods/deployments are accessed through Service objects.

?Ref: <https://www.haproxy.com/blog/dissecting-the-haproxy-kubernetes-ingress-controller/> Create a Kubernetes Service of type NodePort to expose the application on port 443 of each node of the Kubernetes cluster. Configure the public DNS name of your application with the IP of every node of the cluster to achieve load-balancing. is not right.

?Kubernetes Service of type NodePort uses a port in the range 30000-32767. Assuming that all the nodes have public IP addresses, enabling NodePort would expose a port such as 32000 so the application is accessible on <https://IP:32000> which is not ideal. You want your application/website to be reachable directly on port 443. This also requires downstream clients to have awareness of all of your nodes' IP addresses, since they will need to connect to those addresses directly. In other words, they won't be able to connect to a single, proxied IP address. And this is against our requirement of "a public IP address".

?Ref: <https://kubernetes.io/docs/concepts/services-networking/service/>  
?Ref: <https://www.haproxy.com/blog/dissecting-the-haproxy-kubernetes-ingress-controller/> Create a Kubernetes Service of type NodePort for your application, and a Kubernetes Ingress to expose this Service via a Cloud Load Balancer. is the right answer.

?This meets all our requirements. With (Global) Cloud Load Balancing, a single anycast IP front-ends all your backend instances in regions around the world. It provides cross-region load balancing, including automatic multi-region failover, which gently moves traffic in fractions if backends become unhealthy.

?Ref: <https://cloud.google.com/load-balancing/>

?The ingress accepts traffic from the cloud load balancer and can distribute the traffic across the pods in the cluster.

?Ref: <https://kubernetes.io/docs/concepts/services-networking/ingress/>

## 28. Question

You are using multiple configurations for gcloud. You want to review the configured Kubernetes Engine cluster of an inactive configuration using the fewest possible steps. What should you do?

- Use gcloud config configurations describe to review the output.
- Use gcloud config configurations activate and gcloud config list to review the output.
- Use kubectl config get-contexts to review the output.
- Use kubectl config use-context and kubectl config view to review the output.

### Unattempted

Our requirement is to get to the end goal with the fewest possible steps.

Use gcloud config configurations describe to review the output. is not right.  
gcloud config configurations describe – describes a named configuration by listing its properties. This does not return any Kubernetes cluster details.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/describe>

Use gcloud config configurations activate and gcloud config list to review the output. is not right.

gcloud config configurations activate – activates an existing named configuration. This does not return any Kubernetes cluster details.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/activate>

Use kubectl config get-contexts to review the output. is the right answer.  
kubectl config get-contexts displays a list of contexts as well as the clusters that use them. Here's a sample output.

```
$ kubectl config get-contexts
CURRENT NAME CLUSTER
gke_kubernetes-260922_us-central1-a_standard-cluster-1 gke_kubernetes-
260922_us-central1-a_standard-cluster-1
gke_kubernetes-260922_us-central1-a_your-first-cluster-1 gke_kubernetes-
260922_us-central1-a_your-first-cluster-1
* gke_kubernetes-260922_us-central1_standard-cluster-1 gke_kubernetes-
260922_us-central1_standard-cluster-1
```

The output shows the clusters and the configurations they use. Using this information, it is possible to find out the cluster using the inactive configuration with just 1 step.

Use kubectl config use-context and kubectl config view to review the output. is not right.

kubectl config use-context [my-cluster-name] is used to set the default context to [my-cluster-name]. But in order to do this, we first need a list of contexts and if you have multiple contexts, you'd need to execute kubectl config use-context [my-cluster-name] against each context. So that is at least 2+ steps. Further to that, the kubectl config view is used to get a full list of config. The output of the kubectl config view can be used to verify which clusters use what configuration but that is one additional step. Moreover, the output of the kubectl config view doesn't change much from one context to other – other than the current-context field. So our earlier steps of determining the contexts and using each context are

of not much use. Though this can be used to achieve the same outcome, it involves more steps than the other option.

Here's a sample execution

Step 1: First get a list of contexts

```
kubectl config get-contexts -o=name
gke_kubernetes-260922_us-central1-a_standard-cluster-1
gke_kubernetes-260922_us-central1-a_your-first-cluster-1
gke_kubernetes-260922_us-central1_standard-cluster-1
```

Step 2: Use each context and view the config.

```
kubectl config use-context gke_kubernetes-260922_us-central1-a_standard-cluster-1
Switched to context "gke_kubernetes-260922_us-central1-a_standard-cluster-1".
kubectl config view > 1.out (this saves the output of config view in 1.out)
```

```
kubectl config use-context gke_kubernetes-260922_us-central1-a_your-first-cluster-1
Switched to context "gke_kubernetes-260922_us-central1-a_your-first-cluster-1".
kubectl config view > 2.out (this saves the output of config view in 2.out)
```

```
kubectl config use-context gke_kubernetes-260922_us-central1_standard-cluster-1
Switched to context "gke_kubernetes-260922_us-central1_standard-cluster-1".
kubectl config view > 3.out (this saves the output of config view in 3.out)
```

diff 1.out 2.out

28c28

```
< current-context: gke_kubernetes-260922_us-central1-a_standard-cluster-1 --- >
current-context: gke_kubernetes-260922_us-central1-a_your-first-cluster-1
```

diff 2.out 3.out

28c28

```
< current-context: gke_kubernetes-260922_us-central1-a_your-first-cluster-1 --- >
current-context: gke_kubernetes-260922_us-central1_standard-cluster-1
```

Step 3: Determine the inactive configuration and the cluster using that configuration.

The config itself has details about the clusters and contexts as shown below.

```
$ kubectl config view
apiVersion: v1
clusters:
- cluster:
 certificate-authority-data: DATA+OMITTED
 server: https://35.222.130.166
 name: gke_kubernetes-260922_us-central1-a_standard-cluster-1
- cluster:
```

```
certificate-authority-data: DATA+OMITTED
server: https://35.225.14.172
name: gke_kubernetes-260922_us-central1-a_your-first-cluster-1
- cluster:
 certificate-authority-data: DATA+OMITTED
 server: https://34.69.212.109
 name: gke_kubernetes-260922_us-central1_standard-cluster-1
 contexts:
 - context:
 cluster: gke_kubernetes-260922_us-central1-a_standard-cluster-1
 user: gke_kubernetes-260922_us-central1-a_standard-cluster-1
 name: gke_kubernetes-260922_us-central1-a_standard-cluster-1
 - context:
 cluster: gke_kubernetes-260922_us-central1-a_your-first-cluster-1
 user: gke_kubernetes-260922_us-central1-a_your-first-cluster-1
 name: gke_kubernetes-260922_us-central1-a_your-first-cluster-1
 - context:
 cluster: gke_kubernetes-260922_us-central1_standard-cluster-1
 user: gke_kubernetes-260922_us-central1_standard-cluster-1
 name: gke_kubernetes-260922_us-central1_standard-cluster-1
current-context: gke_kubernetes-260922_us-central1-a_standard-cluster-1
```

29. 29. Question

You built an application on Google Cloud Platform that uses Cloud Spanner. The support team needs to monitor the environment but should not have access to the data. You need a streamlined solution to grant the correct permissions to your support team, and you want to follow Google recommended practices. What should you do?

- Add the support team group to the roles/spanner.database.reader role
- Add the support team group to the roles/stackdriver.accounts.viewer role
- Add the support team group to the roles/monitoring.viewer role
- Add the support team group to the roles/spanner.database.user role

**Unattempted**

Requirements –

?1. Monitoring access but no data access

?2. Streamlined solution

?3. Google recommended practices (i.e. look for something out of the box).  
roles/spanner.databaseReader provides permission to read from the Spanner database, execute SQL queries on the database, and view the schema. Since this provides read access to data, roles/spanner.databaseReader. is not right.  
roles/spanner.databaseUser provides permission to read from and write to the Spanner database, execute SQL queries on the database, and view and update

the schema. Since this provides both read and write access to data, roles/spanner.databaseUser. is not right. roles/stackdriver.accounts.viewer read-only access to get and list information about Stackdriver account structure. Since this does not provide monitor access to Cloud Spanner, roles/stackdriver.accounts.viewer. is not right. roles/monitoring.viewer provides read-only access to get and list information about all monitoring data and configurations. This role provides monitoring access and fits our requirements. roles/monitoring.viewer. is the right answer.

?Ref: <https://cloud.google.com/iam/docs/understanding-roles#cloud-spanner-roles>

### 30. 30. Question

You create a Deployment with 2 replicas in a Google Kubernetes Engine cluster that has a single preemptible node pool. After a few minutes, you use kubectl to examine the status of your Pod and observe that one of them is still in Pending status:

NAME READY STATUS RESTART AGE

myapp-deployment-58ddbbb995-lp86m 0/1 Pending 0 9m

myapp-deployment-58ddbbb995-qjpkd 1/1 Running 0 9m

What is the most likely cause?

- The pending Pod's resource requests are too large to fit on a single node of the cluster.
- Too many Pods are already running in the cluster, and there are not enough resources left to schedule the pending Pod.**
- The node pool is configured with a service account that does not have permission to pull the container image used by the pending Pod.
- The pending Pod was originally scheduled on a node that has been preempted between the creation of the Deployment and your verification of the Pod status. It is currently being rescheduled on a new node.

### Unattempted

The pending Pod was originally scheduled on a node that has been preempted between the creation of the Deployment and your verification of the Pod status. It is currently being rescheduled on a new node. is not right.

?Our question states that we provisioned a Google Kubernetes Engine cluster with a single preemptible node pool. The node pool is configured with a service account that does not have permission to pull the container image used by the pending Pod. is not right.

?If the node pool has permission issues when pulling the container image, the other pod would not be in Running status. And the status would have been ImagePullBackOff if there was a problem pulling the image. The pending Pod's resource requests are too large to fit on a single node of the cluster. is not right.

?If the resource requests in Pod specification are too large to fit on the node, the other pod would not be in Running status, i.e. both pods should have been in pending status if this was the case.

?Ref: The pending Pod's resource requests are too large to fit on a single node of the cluster. Too many Pods are already running in the cluster, and there are not enough resources left to schedule the pending Pod. is the right answer.

?When you have a deployment with some pods in running and other pods in the pending state, more often than not it is a problem with resources on the nodes. Here's a sample output of this use case. We see that the problem is with insufficient CPU on the Kubernetes nodes so we have to either enable auto-scaling or manually scale up the nodes.

?kubectl describe pod myapp-deployment-58ddbbb995-lp86m

?Events:

?Type Reason Age From Message

?— — — —

?Warning FailedScheduling 28s (x4 over 3m1s) default-scheduler 0/1 nodes are available: 1 Insufficient cpu.

### 31. Question

You create a new Google Kubernetes Engine (GKE) cluster and want to make sure that it always runs a supported and stable version of Kubernetes. What should you do?

- Select "Container-Optimized OS (cos)" as a node image for your GKE cluster.
- Select the latest available cluster version for your GKE cluster.
- Enable the Node Auto-Upgrades feature for your GKE cluster.**
- Enable the Node Auto-Repair feature for your GKE cluster.

**Unattempted**

Enable the Node Auto-Repair feature for your GKE cluster. is not right.

?GKE's node auto-repair feature helps you keep the nodes in your cluster in a healthy, running state. When enabled, GKE makes periodic checks on the health state of each node in your cluster. If a node fails consecutive health checks over an extended time period, GKE initiates a repair process for that node.

?Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/node-auto-repair> Select the latest available cluster version for your GKE cluster. is not right.

?We can certainly select the latest available cluster version at the time of GKE cluster provisioning, however, this does not automatically upgrade the cluster if new versions become available. Select “ Container-Optimized OS (cos)” as a node image for your GKE cluster. is not right.

?Container-Optimized OS comes with the Docker container runtime and all Kubernetes components pre-installed for out of the box deployment, management, and orchestration of your containers. But these do not help with automatically upgrading GKE cluster versions.

?Ref: <https://cloud.google.com/container-optimized-os> Enable the Node Auto-Updates feature for your GKE cluster. is the right answer.

?Node auto-upgrades help you keep the nodes in your cluster up to date with the cluster master version when your master is updated on your behalf. When you create a new cluster or node pool with Google Cloud Console or the gcloud command, node auto-upgrade is enabled by default.

?Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/node-auto-upgrades>

## 32. Question

You created a cluster.YAML file containing

resources:

– name: cluster

type: container.v1.cluster

properties:

zone: europe-west1-b

cluster:

description: “ My GCP ACE cluster”

initialNodeCount: 2

You want to use Cloud Deployment Manager to create this cluster in GKE. What should you do?

- gcloud deployment-manager deployments create my-gcp-ace-cluster --config cluster.yaml**
- gcloud deployment-manager deployments create my-gcp-ace-cluster --type container.v1.cluster --config cluster.yaml
- gcloud deployment-manager deployments apply my-gcp-ace-cluster --config cluster.yaml
- gcloud deployment-manager deployments apply my-gcp-ace-cluster --type container.v1.cluster --config cluster.yaml

**Unattempted**

gcloud deployment-manager deployments apply my-gcp-ace-cluster – config cluster.yaml. is not right.

?“ gcloud deployment-manager deployments” doesn’ t support action apply. With Google cloud in general, the action for creating is create and the action for retrieving is list. With Kubernetes resources, the corresponding actions are apply and get respectively.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/deployment-manager/deployments/create> gcloud deployment-manager deployments apply my-gcp-ace-cluster – type container.v1.cluster – config cluster.yaml. is not right.

?“ gcloud deployment-manager deployments” doesn’ t support action apply. With Google cloud in general, the action for creating is create and the action for retrieving is list. With Kubernetes resources, the corresponding actions are apply and get respectively.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/deployment-manager/deployments/create> gcloud deployment-manager deployments create my-gcp-ace-cluster – type container.v1.cluster – config cluster.yaml. is not right.

?“ gcloud deployment-manager deployments create” creates deployments based on the configuration file. (Infrastructure as code). It doesn’ t expect the parameter type passed to it directly and fails when executed with the type parameter.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/deployment-manager/deployments/create> gcloud deployment-manager deployments create my-gcp-ace-cluster – config cluster.yaml. is the right answer.

?“ gcloud deployment-manager deployments create” creates deployments based on the configuration file. (Infrastructure as code). All the configuration related to the artifacts is in the configuration file. This command correctly creates a cluster based on the provided cluster.yaml configuration file.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/deployment-manager/deployments/create>

33. 33. Question

You created a compute instance by running gcloud compute instances create instance1. You intended to create the instance in project gcp-ace-proj-266520 but the instance got created in a different project. Your cloud shell gcloud configuration is as shown.

```
$ gcloud config list
```

```
[component_manager]
```

```
disable_update_check = True
```

```
[compute]
```

```
gce_metadata_read_timeout_sec = 5
```

```
zone = europe-west2-a
```

```
[core]
```

```
account = gcp-ace-lab-user@gmail.com
```

```
disable_usage_reporting = False
```

```
project = gcp-ace-lab-266520
```

```
[metrics]
```

```
environment = devshell
```

What should you do to delete the instance that was created in the wrong project and recreate it in gcp-ace-proj-266520 project?

- gcloud compute instances delete instance1 gcloud config set compute/project gcp-ace-proj-266520 gcloud compute instances create instance1
- gcloud config set project gcp-ace-proj-266520 gcloud compute instances recreate instance1 --previous-project gcp-ace-lab-266520
- gcloud compute instances delete instance1 gcloud compute instances create instance1

○ `gcloud compute instances delete instance1` `gcloud config set project gcp-ace-proj-266520` `gcloud compute instances create instance1`

**Unattempted**

`gcloud compute instances delete instance1`

`gcloud compute instances create instance1.` is not right.

The default core/project property is set to gcp-ace-lab-266520 in our current configuration so the instance would have been created in this project. Running the first command to delete the instance correctly deletes it from this project but we haven't modified the core/project property before executing the second command so the instance is recreated in the same project which is not what we want.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/create>

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/delete>

`gcloud config set project gcp-ace-proj-266520`

`gcloud compute instances recreate instance1 – previous-project gcp-ace-lab-266520.` is not right.

`gcloud compute instances` command doesn't support recreate action. It supports create/delete which is what we are supposed to use for this requirement.

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances>

`gcloud compute instances delete instance1`

`gcloud config set compute/project gcp-ace-proj-266520`

`gcloud compute instances create instance1.` is not right.

The approach is right but the syntax is wrong. `gcloud config` does not have a compute/project property. The project property is part of the core/ section as seen in the output of `gcloud configuration list` in the question. In this scenario, we are trying to set compute/project property that doesn't exist in the compute section so the command fails.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/set>

`gcloud compute instances delete instance1`

`gcloud config set project gcp-ace-proj-266520`

`gcloud compute instances create instance1.` is the right answer.

This sequence of commands correctly deletes the instance from gcp-ace-lab-266520 which is the default project in the active `gcloud` configuration, then modifies the current configuration to set the default project to gcp-ace-proj-266520, and finally creates the instance in the project gcp-ace-proj-266520 which is the default project in active `gcloud` configuration at the time of running the command. This produces the intended outcome of deleting the instance from gcp-ace-lab-266520 project and recreating it in gcp-ace-prod-266520

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/set>

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/create>

Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/delete>

34. **34. Question**

You created a Google Cloud Platform project with an App Engine application inside the project. You initially configured the application to be served from the

us-central region. Now you want the application to be served from the asia-northeast1 region. What should you do?

- Create a new GCP project and create an App Engine application inside this new project. Specify asia-northeast1 as the region to serve your application.
- Create a second App Engine application in the existing GCP project and specify asia-northeast1 as the region to serve your application.
- Change the default region property setting in the existing GCP project to asia-northeast1.
- Change the region property setting in the existing App Engine application from us-central to asia-northeast1.

#### Unattempted

Change the default region property setting in the existing GCP project to asia-northeast1. is not right.

?App Engine is regional, which means the infrastructure that runs your apps is located in a specific region and is managed by Google to be redundantly available across all the zones within that region. You cannot change an app's region after you set it.

?Ref: <https://cloud.google.com/appengine/docs/locations> Change the region property setting in the existing App Engine application from us-central to asia-northeast1. is not right.

?App Engine is regional, which means the infrastructure that runs your apps is located in a specific region and is managed by Google to be redundantly available across all the zones within that region. You cannot change an app's region after you set it.

?Ref: <https://cloud.google.com/appengine/docs/locations> Create a second App Engine application in the existing GCP project and specify asia-northeast1 as the region to serve your application. is not right.

?App Engine is regional and you cannot change an app's region after you set it. You can deploy additional services in the App Engine but they will all be targeted to the same region.

?Ref: <https://cloud.google.com/appengine/docs/locations> Create a new GCP project and create an App Engine application inside this new project. Specify asia-northeast1 as the region to serve your application. is the right answer.

?App Engine is regional and you cannot change an app's region after you set it. Therefore, the only way to have an app run in another region is by creating a new project and targeting the app engine to run in the required region (asia-northeast1 in our case).

?Ref: <https://cloud.google.com/appengine/docs/locations>

### 35. 35. Question

You created a Kubernetes deployment by running `kubectl run nginx –image=nginx –labels=" app=prod"`. Your Kubernetes cluster is also used by a number of other deployments. How can you find the identifier of the pods for this nginx deployment?

- `kubectl get deployments --output=pods`
- `gcloud get pods --selector="app=prod"`
- `gcloud list gke-deployments --filter={ pod }`
- `kubectl get pods -l "app=prod"`

**Unattempted**

`gcloud get pods – selector=" app=prod" .` is not right.

?You can not retrieve pods from the Kubernetes cluster by using gcloud. You can list pods by using Kubernetes CLI – `kubectl get pods`.

?Ref: <https://kubernetes.io/docs/tasks/access-application-cluster/list-all-running-container-images/> `gcloud list gke-deployments – filter={ pod }`. is not right.

?You can not retrieve pods from the Kubernetes cluster by using gcloud. You can list pods by using Kubernetes CLI – `kubectl get pods`.

?Ref: <https://kubernetes.io/docs/tasks/access-application-cluster/list-all-running-container-images/> `kubectl get deployments – output=pods`. is not right.

?You can not list pods by listing Kubernetes deployments. You can list pods by using Kubernetes CLI – `kubectl get pods`.

?Ref: <https://kubernetes.io/docs/tasks/access-application-cluster/list-all-running-container-images/> `kubectl get pods -l " app=prod" .` is the right answer.

?This command correctly lists pods that have the label `app=prod`. When creating the deployment, we used the label `app=prod` so listing pods that have this label retrieve the pods belonging to nginx deployments. You can list pods by using Kubernetes CLI – `kubectl get pods`.

?Ref: <https://kubernetes.io/docs/tasks/access-application-cluster/list-all-running-container-images/>

?Ref: <https://kubernetes.io/docs/tasks/access-application-cluster/list-all-running-container-images/#list-containers-filtering-by-pod-label>

### 36. 36. Question

You created a Kubernetes deployment by running `kubectl run nginx –image=nginx –replicas=1`. After a few days, you decided you no longer want this deployment. You identified the pod and deleted it by running `kubectl delete pod`. You noticed the pod got recreated.

\$ `kubectl get pods`

NAME READY STATUS RESTARTS AGE

nginx-84748895c4-nqqmt 1/1 Running 0 9m41s

```
$ kubectl delete pod nginx-84748895c4-nqqmt
pod "nginx-84748895c4-nqqmt" deleted
```

```
$ kubectl get pods
NAME READY STATUS RESTARTS AGE
nginx-84748895c4-k6bz1 1/1 Running 0 25s
What should you do to delete the deployment and avoid pod getting recreated?
```

- `kubectl delete nginx`
- `kubectl delete --deployment=nginx`
- `kubectl delete pod nginx-84748895c4-k6bz1 --no-restart`
- `kubectl delete deployment nginx`

Unattempted

`kubectl delete pod nginx-84748895c4-k6bz1 --no-restart.` is not right.

?`kubectl delete pod` command does not support the flag – no-restart. The command fails to execute due to the presence of an invalid flag.

?`$ kubectl delete pod nginx-84748895c4-k6bz1 --no-restart`

?Error: unknown flag: – no-restart

?Ref: <https://kubernetes.io/docs/reference/kubectl/cheatsheet/#deleting-resources> `kubectl delete --deployment=nginx.` is not right.

?`kubectl delete` command does not support the parameter – deployment. The command fails to execute due to the presence of an invalid parameter.

?`$ kubectl delete --deployment=nginx`

?Error: unknown flag: – deployment

?Ref: <https://kubernetes.io/docs/reference/kubectl/cheatsheet/#deleting-resources> `kubectl delete nginx.` is not right.

?We haven't provided the kubectl delete command information on what to delete, whether a pod, a service or a deployment. The command syntax is wrong and fails to execute.

?`$ kubectl delete nginx`

?error: resource(s) were provided, but no name, label selector, or – all flag specified

?Ref: <https://kubernetes.io/docs/reference/kubectl/cheatsheet/#deleting-resources> `kubectl delete deployment nginx.` is the right answer.

?This command correctly deletes the deployment. Pods are managed by kubernetes workloads (deployments). When a pod is deleted, the deployment detects the pod is unavailable and brings up another pod to maintain the replica count. The only way to delete the workload is by deleting the deployment itself using the kubectl delete deployment command.

```
?$ kubectl delete deployment nginx
```

```
?deployment.apps "nginx" deleted
```

?Ref: <https://kubernetes.io/docs/reference/kubectl/cheatsheet/#deleting-resources>

### 37. Question

You created an instance of SQL Server 2017 on Compute Engine to test features in the new version. You want to connect to this instance using the fewest number of steps. What should you do?

- Set a Windows password in the GCP Console. Verify that a firewall rule for port 22 exists. Click the RDP button in the GCP Console and supply the credentials to log in.
- Set a Windows username and password in the GCP Console. Verify that a firewall rule for port 3389 exists. Click the RDP button in the GCP Console, and supply the credentials to log in.
- Install an RDP client on your desktop. Verify that a firewall rule for port 3389 exists.
- Install an RDP client on your desktop. Set a Windows username and password in the GCP Console. Use the credentials to log in to the instance.

### Unattempted

Requirements – Connect to compute instance using fewest steps. The presence of SQL Server 2017 on the instance is a red herring and should be ignored as none of the options provided say anything about the database and all seem to revolve around RDP. Install a RDP client on your desktop. Verify that a firewall rule for port 3389 exists. is not right.

?Although opening port 3389 is essential for serving RDP traffic, we do not have the credentials to RDP so this isn't going to work. Set a Windows password in the GCP Console. Verify that a firewall rule for port 22 exists. Click the RDP button in the GCP Console and supply the credentials to log in. is not right.

?RDP uses port 3389 and not 22.

?Ref: <https://cloud.google.com/compute/docs/troubleshooting/troubleshooting-rdp> Set a Windows username and password in the GCP Console. Verify that a

firewall rule for port 3389 exists. Click the RDP button in the GCP Console, and supply the credentials to log in. is not right.

?While this option correctly sets the username and password on the console and verifies a firewall rule is set on port 3389 to allow RDP traffic, you can RDP from console unless you install Chrome RDP for Google Cloud Platform extension in order to RDP from the console. (See Chrome Desktop for GCP tab in <https://cloud.google.com/compute/docs/instances/connecting-to-instance#windows>). If we assume that installing Chrome RDP for Google Cloud Platform extension is carried out (even though not specified in the option), we end up executing more steps in this option to successfully RDP compare to the correct answer (below) Install an RDP client on your desktop. Set a Windows username and password in the GCP Console. Use the credentials to log in to the instance. is the right answer.

?This option correctly sets the username/password which is essential. In addition, the default VPC comes with port 3389 open to the public. The question doesn't explicitly state the compute engine is in a custom VPC so it is safe to assume we are using default VPC which has default RDP access open to the public. Finally, you install an RDP client on the desktop and use the credentials set up earlier to RDP to the server.

### 38. Question

You defined an instance template for a Python web application. When you deploy this application in Google Compute Engine, you want to ensure the service scales up and scales down automatically based on the number of HTTP requests. What should you do?

- 1. Create the necessary number of instances based on the instance template to handle peak user traffic. 2. Group the instances together in an unmanaged instance group. 3. Configure the instance group as the Backend Service of an External HTTP(S) load balancer.
- 1. Create an instance from the instance template. 2. Create an image from the instance's disk and export it to Cloud Storage. 3. Create an External HTTP(s) load balancer and add the Cloud Storage bucket as its backend service.
- 1. Create an unmanaged instance group from the instance template. 2. Configure autoscaling on the unmanaged instance group with a scaling policy based on HTTP traffic. 3. Configure the unmanaged instance group as the backend service of an Internal HTTP(S) load balancer.
- 1. Deploy your Python web application instance template to Google Cloud App Engine. 2. Configure autoscaling on the managed instance group with a scaling policy based on HTTP traffic.
- 1. Create a managed instance group from the instance template. 2. Configure autoscaling on the managed instance group with a scaling policy based on HTTP traffic. 3. Configure the instance group as the backend service of an External HTTP(S) load balancer.

Unattempted

1. Create an instance from the instance template.
2. Create an image from the instance's disk and export it to Cloud Storage.
3. Create an External HTTP(S) load balancer and add the Cloud Storage bucket as its backend service. is not right.

You can upload a custom image from instance's boot disk and export it to cloud storage.

<https://cloud.google.com/compute/docs/images/export-image>

However, this image in the Cloud Storage bucket is unable to handle traffic as it is not a running application. Cloud Storage can not serve requests of the custom image.

1. Create an unmanaged instance group from the instance template.
2. Configure autoscaling on the unmanaged instance group with a scaling policy based on HTTP traffic.
3. Configure the unmanaged instance group as the backend service of an Internal HTTP(S) load balancer. is not right.

An unmanaged instance group does not autoscale. An unmanaged instance group is a collection of virtual machines (VMs) that reside in a single zone, VPC network, and subnet. An unmanaged instance group is useful for grouping together VMs that require individual configuration settings or tuning.

Ref: <https://cloud.google.com/compute/docs/instance-groups/creating-groups-of-unmanaged-instances>

1. Create the necessary number of instances based on the instance template to handle peak user traffic.
2. Group the instances together in an unmanaged instance group.
3. Configure the instance group as the Backend Service of an External HTTP(S) load balancer. is not right.

An unmanaged instance group does not autoscale. Although we may have enough compute power to handle peak user traffic, it does not automatically scale down when the traffic goes down so it doesn't meet our requirements.

Ref: <https://cloud.google.com/compute/docs/instance-groups/creating-groups-of-unmanaged-instances>

1. Deploy your Python web application instance template to Google Cloud App Engine.
2. Configure autoscaling on the managed instance group with a scaling policy based on HTTP traffic. is not right.

You can not use compute engine instance templates to deploy applications to Google Cloud App Engine. Google App Engine lets you deploy applications quickly by providing run time environments for many of the popular languages like Java, PHP, Node.js, Python, C#, .Net, Ruby, and Go. You have an option of using custom runtimes but using compute engine instance templates is not an option.

Ref: <https://cloud.google.com/appengine>

1. Create a managed instance group from the instance template.
2. Configure autoscaling on the managed instance group with a scaling policy based on HTTP traffic.
3. Configure the instance group as the backend service of an External HTTP(S) load balancer. is the right answer.

The auto-scaling capabilities of Managed instance groups let you automatically

add or delete instances from a managed instance group based on increases or decreases in load – this can be set up by configuring scaling policies. In addition, you can configure External HTTP(S) load balancer to send traffic to the managed instance group. The External HTTP(S) load balancer tries to balance requests by using a round-robin algorithm and when the load increases beyond the threshold defined in the scaling policy, autoscaling kicks in and adds more nodes.

Ref: <https://cloud.google.com/load-balancing/docs/https>

Ref: <https://cloud.google.com/compute/docs/instance-groups/creating-groups-of-managed-instances>

39. 39. Question

You deployed a new application inside your Google Kubernetes Engine cluster using the YAML file specified below.

```
apiVersion: apps/v1
```

```
kind: Deployment
```

```
metadata:
```

```
 name: myapp-deployment
```

```
spec:
```

```
 selector:
```

```
 matchLabels:
```

```
 app: myapp
```

```
 replicas: 2
```

```
 template:
```

```
 metadata:
```

```
 labels:
```

```
 app: myapp
```

```
 spec:
```

```
 containers:
```

```
 - name: myapp
```

```
 image: myapp:1.1
```

```
 ports:
```

```
 - containerPort: 80
```

```
—
```

```
apiVersion: v1
```

```
kind: Service
```

```
metadata:
```

```
 name: myapp-service
```

```
spec:
```

```
 ports:
```

```
 - port: 8000
```

```
 targetPort: 80
```

```
 protocol: TCP
```

```
 selector:
```

```
 app: myapp
```

You check the status of the deployed pods and notice that one of them is still in PENDING status:

```
kubectl get pods -l app=myapp
```

| NAME                              | READY | STATUS  | RESTARTS | AGE |
|-----------------------------------|-------|---------|----------|-----|
| myapp-deployment-58ddbbb995-lp86m | 0/1   | Pending | 0        | 9m  |

myapp-deployment-58ddbbb995-qjpk 1/1 Running 0m

You want to find out why the pod is stuck in pending status. What should you do?

- Review details of the myapp-service Service object and check for error messages.
- Review details of the myapp-deployment Deployment object and check for error messages.
- Review details of myapp-deployment-58ddbbb995-lp86m Pod and check for warning messages.**
- View logs of the container in myapp-deployment-58ddbbb995-lp86m pod and check for warning messages.

#### Unattempted

Review details of the myapp-service Service object and check for error messages. is not right.

The question states we have a problem with the deployment.

Checking/Reviewing the status of the service object isn't of much use here.

View logs of the container in myapp-deployment-58ddbbb995-lp86m pod and check for warning messages. is not right.

Since the pod hasn't moved to Running state, the logs of the container would be empty. So running

kubectl logs pod/myapp-deployment-58ddbbb995-lp86m  
to check the logs of the pod isn't of much use.

Review details of the myapp-deployment Deployment object and check for error messages. is not right.

Describing the details of the deployment shows us how many of the pods are available and unavailable but does not show errors/warnings related to a specific pod.

Here's a sample output of this use case.

kubectl describe deployment myapp-deployment

Replicas: 3 desired | 3 updated | 3 total | 2 available | 1 unavailable

Events:

Type Reason Age From Message

Normal Scaling ReplicaSet 4m54s deployment-controller Scaled up replica set myapp-deployment-869d88c75f to 3

Review details of myapp-deployment-58ddbbb995-lp86m Pod and check for warning messages. is the right answer.

Since the problem is with a specific pod, looking at the details of the pod is the best solution. When you have a deployment with some pods in running and other pods in Pending state, more often than not it is a problem with resources on the nodes. Here's a sample output of this use case. We see that the problem is with insufficient CPU on the Kubernetes nodes so we have to either enable auto-scaling or manually scale up the nodes.

kubectl describe pod myapp-deployment-58ddbbb995-lp86m

Events:

| Type | Reason | Age | From | Message |
|------|--------|-----|------|---------|
|------|--------|-----|------|---------|

Warning FailedScheduling 28s (x4 over 3m1s) default-scheduler 0/1 nodes are available: 1 Insufficient cpu.

40. Question

You deployed a number of services to Google App Engine Standard. The services are designed as microservices with several interdependencies between them. Most services have few version upgrades but some key services have over 20 version upgrades. You identified an issue with the service pt-createOrder and deployed a new version v3 for this service. You are confident this works and want this new version to receive all traffic for the service. You want to minimize effort and ensure the availability of service. What should you do?

- Execute gcloud app versions stop v2 and gcloud app versions start v3
- Execute gcloud app versions stop v2 --service="pt-createOrder" and gcloud app versions start v3 --service="pt-createOrder"
- Execute gcloud app versions migrate v3
- Execute gcloud app versions migrate v3 --service="pt-createOrder"

Unattempted

Execute gcloud app versions migrate v3. is not right.

?gcloud app versions migrate v3 migrates all services to version v3. In our scenario, we have multiple services with each service potentially being on a different version. We don't want to migrate all services to v3, instead, we only want to migrate the pt-createOrder service to v3.

Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/migrate> Execute gcloud app versions stop v2 –service=" pt-createOrder" and gcloud app versions start v3 –service=" pt-createOrder". is not right.

?Stopping version v2 and starting version v3 for pt-createOrder service would result in v3 receiving all traffic for pt-createOrder. While this is the intended outcome, stopping version v2 before starting version v3 results in service being unavailable until v3 is ready to receive traffic. As we want to " ensure availability" , this option is not suitable.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/migrate> Execute gcloud app versions stop v2 and gcloud app versions start v3. is not right.

?Stopping version v2 and starting version v3 would result in migrating all services to version v3 which is undesirable. We don't want to migrate all services to v3, instead, we only want to migrate the pt-createOrder service to v3. Moreover, stopping version v2 before starting version v3 results in service being unavailable until v3 is ready to receive traffic. As we want to " ensure availability" , this option is not suitable.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/migrate> Execute gcloud app versions migrate v3 – service="pt-createOrder". is the right answer.

?This command correctly migrates the service pt-createOrder to use version 3 and produces the intended outcome while minimizing effort and ensuring the availability of service.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/app/versions/migrate>

#### 41. Question

You deployed a workload to your GKE cluster by running the command kubectl apply -f app.yaml. You also enabled a LoadBalancer service to expose the deployment by running kubectl apply -f service.yaml. Your pods are struggling due to increased load so you decided to enable horizontal pod autoscaler by running kubectl autoscale deployment [YOUR DEPLOYMENT] – cpu-percent=50 – min=1 – max=10. You noticed the autoscaler has launched several new pods but the new pods have failed with the message “ Insufficient cpu” . What should you do to resolve this issue?

- Use "gcloud container clusters resize" to add more nodes to the node pool.
- Use "kubectl container clusters resize" to add more nodes to the node pool.
- Edit the managed instance group of the cluster and enable autoscaling.
- Edit the managed instance group of the cluster and increase the number of VMs by 1.

**Unattempted**

Use “ kubectl container clusters resize” to add more nodes to the node pool. is not right.

?kubectl doesn't support the command kubectl container clusters resize. You have to use gcloud container clusters resize to resize a cluster.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/resize> Edit the managed instance group of the cluster and increase the number of VMs by 1. is not right.

?GKE Cluster does not use a managed instance group. Instead, the cluster master (control plane) handles the lifecycle of nodes in the node pools. The cluster master is responsible for managing the workloads' lifecycle, scaling, and upgrades. The master also manages network and storage resources for those workloads.

?Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/cluster-architecture> Edit the managed instance group of the cluster and enable autoscaling. is not right.

?GKE Cluster does not use a managed instance group. Instead, the cluster master (control plane) handles the lifecycle of nodes in the node pools. The cluster master is responsible for managing the workloads' lifecycle, scaling, and upgrades. The master also manages network and storage resources for those workloads.

?Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/cluster-architecture> Use " gcloud container clusters resize" to add more nodes to the node pool. is the right answer.

?Your pods are failing with " Insufficient cpu" . This is because the existing nodes in the node pool are maxed out, therefore, you need to add more nodes to your node pool. For such scenarios, enabling cluster autoscaling is ideal, however, this is not in any of the answer options. In the absence of cluster autoscaling, the next best approach is to add more nodes to the cluster manually. This is achieved by running the command gcloud container clusters resize which resizes an existing cluster for running containers.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/resize>

## 42. Question

You deployed an App Engine application using gcloud app deploy, but it did not deploy to the intended project. You want to find out why this happened and where the application deployed. What should you do?

- Check the app YAML file for your application and check the project settings.
- Go to Deployment Manager and review settings for the deployment of application
- Check the web application XML file for your application and check project settings
- Go to Cloud Shell and run gcloud config list to review the Google Cloud configurations used for deployment.

**Unattempted**

Check the app YAML file for your application and check the project settings. is not right.

?The Yaml file of application does not hold Google project information. Check the web application XML file for your application and check project settings. is not right.

?The web application file of the application does not hold Google project information. Go to Deployment Manager and review settings for the deployment of the application. is not right.

?Google Cloud Deployment Manager allows you to specify all the resources needed for your application in a declarative format using yaml. In this scenario,

we haven't used Cloud Deployment Manager to deploy. The app was deployed using gcloud app deploy so this option is not right.

?Ref: <https://cloud.google.com/deployment-manager> Go to Cloud Shell and run gcloud config list to review the Google Cloud configurations used for deployment. is the right answer.

?If the deployment was successful but it did not deploy to the intended project, it is likely that the gcloud app deploy command deployed the application to a different project. In the same gcloud shell, you can identify the current properties of the configuration by executing gcloud config list. This returns config properties such as project, account etc, as well as app-specific properties such as app/promote\_by\_default, app/stop\_previous\_version.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/config/list>

### 43. Question

You deployed an LDAP server on Compute Engine. You want to make sure it is reachable by external clients via TLS through port 636 using UDP. What should you do?

- Add the network tag allow-udp-636 to the VM instance running the LDAP server.
- Create a route called allow-udp-636 and set the next hop to be the VM instance running the LDAP server.
- Add a network tag of your choice to the instance. Create a firewall rule to allow ingress on UDP port 636 for that network tag.
- Add a network tag of your choice to the instance running the LDAP server. Create a firewall rule to allow egress on UDP port 636 for that network tag.

### Unattempted

Create a route called allow-udp-636 and set the next hop to be the VM instance running the LDAP server. is not right.

?Google Cloud routes define the paths that network traffic takes from a virtual machine (VM) instance to other destinations. These destinations can be inside your Google Cloud Virtual Private Cloud (VPC) network (for example, in another VM) or outside it. Routes aren't a suitable solution for our requirement as we need to enable EXTERNAL clients to reach our VM on port 636 using UDP.

?Ref: <https://cloud.google.com/vpc/docs/routes> Add the network tag allow-udp-636 to the VM instance running the LDAP server. is not right.

?Tags enable you to make firewall rules and routes applicable to specific VM instances but allow-udp-636 is not a network tag that GCP provides. The default network tags provided by GCP are default-allow-icmp, default-allow-internal, default-allow-rdp and default-allow-ssh. In this scenario, we are assigning a tag to the instance with no network rules so there would be no difference to behavior.

?Ref: <https://cloud.google.com/vpc/docs/add-remove-network-tags> Add a network tag of your choice to the instance running the LDAP server. Create a firewall rule to allow egress on UDP port 636 for that network tag. is not right.

?We are interested in enabling inbound traffic to our VM whereas egress firewall rules control outgoing connections from target instances in your VPC network.

?Ref: [https://cloud.google.com/vpc/docs/firewalls#egress\\_cases](https://cloud.google.com/vpc/docs/firewalls#egress_cases) Add a network tag of your choice to the instance. Create a firewall rule to allow ingress on UDP port 636 for that network tag. is the right answer.

?This fits all our requirements. Ingress firewall rules control incoming connections from a source to target instances in your VPC network. We can create an ingress firewall rule to allow UDP port 636 for a network tag. And when we assign this network tag to the instance, the firewall rule applies to the instances so traffic is accepted on port 636 using UDP. Although not specified in this option, it has to be assumed that the source for the firewall rule is set to 0.0.0.0/0 i.e. all IP ranges so that external clients are allowed to connect to this VM.

?Ref: [https://cloud.google.com/vpc/docs/firewalls#ingress\\_cases](https://cloud.google.com/vpc/docs/firewalls#ingress_cases)

#### 44. Question

You deployed your application to a default node pool on the GKE cluster and you want to configure cluster autoscaling for this GKE cluster. For your application to be profitable, you must limit the number of Kubernetes nodes to 10. You want to start small and scale up as traffic increases and scale down when the traffic goes down. What should you do?

- Update existing GKE cluster to enable autoscaling by running the command gcloud container clusters update [CLUSTER\_NAME] --enable-autoscaling --min-nodes=1 --max-nodes=10**
- Create a new GKE cluster by running the command gcloud container clusters create [CLUSTER\_NAME] --enable-autoscaling --min-nodes=1 --max-nodes=10. Redeploy your application
- To enable autoscaling, add a tag to the instances in the cluster by running the command gcloud compute instances add-tags [INSTANCE] --tags=enable-autoscaling,min-nodes=1,max-nodes=10
- Set up a stack driver alert to detect slowness in the application. When the alert is triggered, increase nodes in the cluster by running the command gcloud container clusters resize CLUSTER\_Name --size .

#### Unattempted

Set up a stack driver alert to detect slowness in the application. When the alert is triggered, increase nodes in the cluster by running the command gcloud container clusters resize CLUSTER\_Name – size . is not right.

?The command gcloud container clusters resize command resizes an existing cluster for running containers. While it is possible to manually increase the number of nodes in the cluster by running the command, the scale-up is not automatic, it is a manual process. Also, there is no scale down so it doesn't fit

our requirement of “ scale up as traffic increases and scale down when the traffic goes down” .

?Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/resize> To enable autoscaling, add a tag to the instances in the cluster by running the command gcloud compute instances add-tags [INSTANCE] –tags=enable-autoscaling,min-nodes=1,max-nodes=10. is not right.

?Autoscaling can not be enabled on the GKE cluster by adding tags on compute instances. Autoscaling can be enabled at the time of creating the cluster and can also be enabled for existing clusters by running one of the gcloud container clusters to create/update commands.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/create>

?Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/update> Create a new GKE cluster by running the command gcloud container clusters create [CLUSTER\_NAME] –enable-autoscaling – min-nodes=1 – max-nodes=10. Redeploy your application. is not right.

?The command gcloud container clusters create – creates a GKE cluster and the flag –enable-autoscaling enables autoscaling and the parameters –min-nodes=1 –max-nodes=10 define the minimum and maximum number of nodes in the node pool. However, we want to configure cluster autoscaling for the existing GKE cluster; not create a new GKE cluster.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/create> Update existing GKE cluster to enable autoscaling by running the command gcloud container clusters update [CLUSTER\_NAME] –enable-autoscaling – min-nodes=1 – max-nodes=10. is the right answer.

?The command gcloud container clusters update – updates an existing GKE cluster. The flag –enable-autoscaling enables autoscaling and the parameters –min-nodes=1 –max-nodes=10 define the minimum and maximum number of nodes in the node pool. This enables cluster autoscaling which scales up and scales down the nodes automatically between 1 and 10 nodes in the node pool.

#### 45. Question

You developed a web application that lets users upload and share images. You deployed this application in Google Compute Engine and you have configured Stackdriver Logging. Your application sometimes times out while uploading large images, and your application generates relevant error log entries that are ingested to Stackdriver Logging. You would now like to create alerts based on these metrics. You intend to add more compute resources manually when the number of failures exceeds a threshold. What should you do in order to alert based on these metrics with minimal effort?

- In Stackdriver logging, create a new logging metric with the required filters, edit the application code to set the metric value when needed, and create an alert in Stackdriver based on the new metric.
- Create a custom monitoring metric in code, edit the application code to set the metric value when needed, create an alert in Stackdriver based on the new metric.

- In Stackdriver Logging, create a custom monitoring metric from log data and create an alert in Stackdriver based on the new metric.
- Add the Stackdriver monitoring and logging agent to the instances running the code.

#### Unattempted

In Stackdriver logging, create a new logging metric with the required filters, edit the application code to set the metric value when needed, and create an alert in Stackdriver based on the new metric. is not right.

?You don't need to edit the application code to send the metric values. The application already pushes error logs whenever the application times out. Since you already have the required entries in the Stackdriver logs, you don't need to edit the application code to send the metric values. You just need to create metrics from log data.

?Ref: <https://cloud.google.com/logging> Create a custom monitoring metric in code, edit the application code to set the metric value when needed, create an alert in Stackdriver based on the new metric. is not right.

?You don't create a custom monitoring metric in code. Stackdriver Logging allows you to easily create metrics from log data. Since the application already pushes error logs to Stackdriver Logging, we just need to create metrics from log data in Stackdriver Logging.

?Ref: <https://cloud.google.com/logging> Add the Stackdriver monitoring and logging agent to the instances running the code. is not right.

?The Stackdriver Monitoring agent gathers system and application metrics from your VM instances and sends them to Monitoring. In order to make use of this approach, you need application metrics but our application doesn't generate metrics. It just logs errors whenever the upload times out and these are then ingested to Stackdriver logging. We can update our application to enable custom metrics for these scenarios, but that is a lot more work than creating metrics from log data in Stackdriver Logging

?Ref: <https://cloud.google.com/logging> In Stackdriver Logging, create a custom monitoring metric from log data and create an alert in Stackdriver based on the new metric. is the right answer.

?Our application adds entries to error logs whenever the application times out during image upload and these logs are ingested to Stackdriver Logging. Since we already have the required data in logs, we just need to create metrics from this log data in Stackdriver Logging. And we can then set up an alert based on this metric. We can trigger an alert if the number of occurrences of the relevant error message is greater than a predefined value. Based on the alert, you can manually add more compute resources.

?Ref: <https://cloud.google.com/logging>

You developed an application that lets users upload statistical files and subsequently run analytics on this data. You chose to use Google Cloud Storage and BigQuery respectively for these requirements as they are highly available and scalable. You have a docker image for your application code, and you plan to deploy on your on-premises Kubernetes clusters. Your on-prem Kubernetes cluster needs to connect to Google Cloud Storage and BigQuery and you want to do this in a secure way following Google recommended practices. What should you do?

- Create a new service account, with editor permissions, generate and download a key. Use the key to authenticate inside the application.
- Use the default service account for App Engine, which already has the required permissions.
- Use the default service account for Compute Engine, which already has the required permissions.
- Create a new service account, grant it the least viable privileges to the required services, generate and download a JSON key. Use the JSON key to authenticate inside the application.**

#### Unattempted

Use the default service account for Compute Engine, which already has the required permissions. is not right.

The Compute Engine default service account is created with the Cloud IAM project editor role

Ref: [https://cloud.google.com/compute/docs/access/service-accounts#default\\_service\\_account](https://cloud.google.com/compute/docs/access/service-accounts#default_service_account)

The project editor role includes all viewer permissions, plus permissions for actions that modify state, such as changing existing resources. Using a service account that is over-privileged falls foul of the principle of least privilege. Google recommends you enforce the principle of least privilege by ensuring that members have only the permissions that they actually need.

Ref: <https://cloud.google.com/iam/docs/understanding-roles>

Use the default service account for App Engine, which already has the required permissions. is not right.

App Engine default service account has the Editor role in the project (Same as the default service account for Compute Engine).

Ref: <https://cloud.google.com/appengine/docs/standard/python/service-account>

The project editor role includes all viewer permissions, plus permissions for actions that modify state, such as changing existing resources. Using a service account that is over-privileged falls foul of the principle of least privilege. Google recommends you enforce the principle of least privilege by ensuring that members have only the permissions that they actually need.

Ref: <https://cloud.google.com/iam/docs/understanding-roles>

Create a new service account, with editor permissions, generate and download a key. Use the key to authenticate inside the application. is not right.

The project editor role includes all viewer permissions, plus permissions for actions that modify state, such as changing existing resources. Using a service

account that is over-privileged falls foul of the principle of least privilege. Google recommends you enforce the principle of least privilege by ensuring that members have only the permissions that they actually need.  
Ref: <https://cloud.google.com/iam/docs/understanding-roles>

Create a new service account, grant it the least viable privileges to the required services, generate and download a JSON key. Use the JSON key to authenticate inside the application. is the right answer.

Using a new service account with just the least viable privileges for the required services follows the principle of least privilege. To use a service account outside of Google Cloud, such as on other platforms or on-premises, you must first establish the identity of the service account. Public/private key pairs provide a secure way of accomplishing this goal. Once you have the key, you can use it in your application to authenticate connections to Cloud Storage and BigQuery.  
Ref: [https://cloud.google.com/iam/docs/creating-managing-service-account-keys#creating\\_service\\_account\\_keys](https://cloud.google.com/iam/docs/creating-managing-service-account-keys#creating_service_account_keys)

Ref: <https://cloud.google.com/iam/docs/recommender-overview>

#### 47. Question

You developed an application that reads objects from a cloud storage bucket. You followed GCP documentation and created a service account with just the permissions to read objects from the cloud storage bucket. However, when your application uses this service account, it fails to read objects from the bucket. You suspect this might be an issue with the permissions assigned to the service account. You would like to authenticate a gsutil session with the service account credentials, reproduce the issue yourself and identify the root cause. How can you authenticate gsutil with service account credentials?

- Create JSON keys for the service account and execute gcloud auth activate-service-account --key-file [KEY\_FILE]
- Create JSON keys for the service account and execute gcloud auth service-account --key-file [KEY\_FILE]
- Create JSON keys for the service account and execute gcloud authenticate service-account --key-file [KEY\_FILE]
- Create JSON keys for the service account and execute gcloud authenticate activate-service-account --key-file [KEY\_FILE]

**Unattempted**

Create JSON keys for the service account and execute gcloud authenticate activate-service-account – key-file [KEY\_FILE]. is not right.

?gcloud doesn' t support using “ authenticate” to grant/revoke credentials for Cloud SDK. The correct service is “ auth” .

?Ref: <https://cloud.google.com/sdk/gcloud/reference/auth> Create JSON keys for the service account and execute gcloud authenticate service-account – key-file [KEY\_FILE]. is not right.

?gcloud doesn't support using "authenticate" to grant/revoke credentials for Cloud SDK. The correct service is "auth".

?Ref: <https://cloud.google.com/sdk/gcloud/reference/auth> Create JSON keys for the service account and execute gcloud auth service-account --key-file [KEY\_FILE]. is not right.

?gcloud auth does not support service-account action. The correct action to authenticate a service account is activate-service-account.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/auth/activate-service-account> Create JSON keys for the service account and execute gcloud auth activate-service-account --key-file [KEY\_FILE]. is the right answer.

?This command correctly authenticates access to Google Cloud Platform with a service account using its JSON key file. To allow gcloud (and other tools in Cloud SDK) to use service account credentials to make requests, use this command to import these credentials from a file that contains a private authorization key, and activate them for use in gcloud

?Ref: <https://cloud.google.com/sdk/gcloud/reference/auth/activate-service-account>

#### 48. Question

You developed an application to serve production users and you plan to use Cloud SQL to host user state data which is very critical for the application flow. You want to protect your user state data from zone failures. What should you do?

- Create a Failover replica in the same region but in a different zone.
- Create a Read replica in the same region but in a different zone.
- Configure High Availability (HA) for Cloud SQL and Create a Failover replica in the same region but in a different zone.
- Configure High Availability (HA) for Cloud SQL and Create a Failover replica in a different region

**Unattempted**

Create a Read replica in the same region but in a different zone. is not right.

?Read replicas do not provide failover capability. To provide failover capability, you need to configure Cloud SQL Instance for High Availability.

?Ref: <https://cloud.google.com/sql/docs/mysqlreplication> Create a Read replica in a different region. is not right.

?Read replicas do not provide failover capability. To provide failover capability, you need to configure Cloud SQL Instance for High Availability.

?Ref: <https://cloud.google.com/sql/docs/mysqlreplication> Configure High Availability (HA) for Cloud SQL and Create a Failover replica in a different region. is not right.

?A Cloud SQL instance configured for HA is called a regional instance because it's primary and secondary instances are in the same region. They are located in different zones but within the same region. It is not possible to create a Failover replica in a different region.

?Ref: <https://cloud.google.com/sql/docs/mysql/high-availability> Configure High Availability (HA) for Cloud SQL and Create a Failover replica in the same region but in a different zone. is the right answer.

?If a HA-configured instance becomes unresponsive, Cloud SQL automatically switches to serving data from the standby instance. The HA configuration provides data redundancy. A Cloud SQL instance configured for HA has instances in the primary zone (Master node) and secondary zone (standby/failover node) within the configured region. Through synchronous replication to each zone's persistent disk, all writes made to the primary instance are also made to the standby instance. If the primary goes down, the standby/failover node takes over and your data continues to be available to client applications.

?Ref: <https://cloud.google.com/sql/docs/mysql/high-availability>

#### 49. Question

You have 32 GB of data in a single file that you need to upload to a Nearline Storage bucket. The WAN connection you are using is rated at 1 Gbps, and you are the only one on the connection. You want to use as much of the rated 1 Gbps as possible to transfer the file rapidly. How should you upload the file?

- Use the GCP Console to transfer the file instead of gsutil.
- Change the storage class of the bucket from Nearline to Multi-Regional.
- Enable parallel composite uploads using gsutil on the file transfer.**
- Decrease the TCP window size on the machine initiating the transfer.

#### Unattempted

Requirements – transfer the file rapidly, use as much of the rated 1 Gbps as possible Use the GCP Console to transfer the file instead of gsutil. is not right.

?GCP Console does not offer any specific features that help in improving the upload speed. Decrease the TCP window size on the machine initiating the transfer. is not right.

?By decreasing the TCP window size, you are reducing the chunks of data sent in the TCP window, and this has the effect of underutilizing your bandwidth and can slow down the upload. Change the storage class of the bucket from Nearline to Multi-Regional. is not right.

?Multi-Regional is not a storage class. It is a bucket location. You can transition between storage classes but that does not improve the upload speed.

?<https://cloud.google.com/storage/docs/locations>

?<https://cloud.google.com/storage/docs/storage-classes> Enable parallel composite uploads using gsutil on the file transfer. is the right answer.

?With cloud storage, Object composition can be used for uploading an object in parallel: you can divide your data into multiple chunks, upload each chunk to a distinct object in parallel, compose your final object, and delete any temporary source objects. This helps maximize your bandwidth usage and ensures the file is uploaded as fast as possible.

?Ref: <https://cloud.google.com/storage/docs/composite-objects#uploads>

50. Question

You have a collection of audio/video files over 80GB each that you need to migrate to Google Cloud Storage. The files are in your on-premises data center. What migration method can you use to help speed up the transfer process?

- Use parallel uploads to break the file into smaller chunks then transfer it simultaneously.
- Use multithreaded uploads using the -m option.
- Use the Cloud Transfer Service to transfer.
- Start a recursive upload.

**Unattempted**

Use parallel uploads to break the file into smaller chunks then transfer it simultaneously. is the right answer.

?With cloud storage, Object composition can be used for uploading an object in parallel: you can divide your data into multiple chunks, upload each chunk to a distinct object in parallel, compose your final object, and delete any temporary source objects. This helps maximize your bandwidth usage and ensures the file is uploaded as fast as possible.

?Ref: <https://cloud.google.com/storage/docs/composite-objects#uploads> Use multithreaded uploads using the -m option. is not right.

?Using the -m option lets you upload multiple files at the same time, but in our case, the individual files are over 80GB each. The best upload speed can be achieved by breaking the file into smaller chunks and transferring it simultaneously. Use the Cloud Transfer Service to transfer. is not right.

?Cloud Transfer Service is used for transferring massive amounts (in the range of petabytes of data) of data to the cloud. While nothing stops us from using Cloud Transfer Service to upload our files, it would be an overkill and very expensive.

?Ref: <https://cloud.google.com/products/data-transfer> Start a recursive upload. is not right.

?In Google Cloud Storage, there is no such thing as a recursive upload.

### 51. Question

You have a compute engine instance running a production application. You want to receive an email when the instance consumes more than 90% of its CPU resources for more than 15 minutes. You want to use Google services. What should you do?

- 1. Create a Stackdriver Workspace and associate your GCP project with it. 2. Write a script that monitors the CPU usage and sends it as a custom metric to Stackdriver 3. Create an uptime check for the instance in Stackdriver.
- 1. Create a consumer Gmail Account 2. Write a script that monitors the CPU usage. 3. When the CPU usage exceeds the threshold, have the script send an email using the Gmail account and smtp.gmail.com on port 25 as SMTP server.
- 1. Create a Stackdriver Workspace and associate your Google Cloud Platform (GCP) project with it 2. Create an Alerting Policy in Stackdriver that uses the threshold as a trigger condition. 3. Configure your email address in the notification channel.
- 1. In Stackdriver logging, create a logs based metric to extract the CPU usage by using a regular expression. 2. In Stackdriver Monitoring, create an Alerting Policy based on this metric 3. Configure your email address in the notification channel.

### Unattempted

We want to use Google services. So that eliminates the two options where we Write a script. Why would we want to write a script when there is a Google service that does exactly that – with minimal configuration!! Stackdriver logging does not log CPU usage. (Stackdriver monitoring does that) So that rules out the option In Stackdriver logging, create a logs based metric to extract the CPU usage by using a regular expression.

?Ref: <https://cloud.google.com/logging/> 1. Create a Stackdriver Workspace and associate your Google Cloud Platform (GCP) project with it  
?2. Create an Alerting Policy in Stackdriver that uses the threshold as a trigger condition.

?3. Configure your email address in the notification channel.

?is the right answer. A Workspace is a tool for monitoring resources contained in one or more Google Cloud projects or AWS accounts. In our case, we create a Stackdriver workspace and link our project to this workspace.

?Ref: <https://cloud.google.com/monitoring/workspaces> Stackdriver monitoring captures the CPU usage. By default, the Monitoring agent collects disk, CPU, network, and process metrics. You can also have the agent send custom metrics to Stackdriver monitoring.

?Ref: <https://cloud.google.com/monitoring/> You can then set up an alerting policy to alert with CPU utilization exceeds 90% for 15 minutes.

?Ref: <https://cloud.google.com/monitoring/alerts/>. See here for an example of setting up an alerting policy on CPU load. In our case, we'd have to substitute the CPU load for the CPU utilization

metric. <https://cloud.google.com/monitoring/quickstart-lamp> Stack driver monitoring supports multiple notification options for triggering alerts; email is one of them. Ref: <https://cloud.google.com/monitoring/support/notification-options>

## 52. Question

You have a developer laptop with Cloud SDK installed on Ubuntu. The cloud SDK was installed from Google Cloud Ubuntu package repository. You want to test your application locally on your laptop with Cloud Datastore. What should you do?

- Create a Cloud Datastore index using gcloud datastore indexes create
- Install the google-cloud-sdk-datastore-emulator component using the apt get install command.
- Export Cloud Datastore data using gcloud datastore export
- Install the cloud-datastore-emulator component using the gcloud components install command.**

**Unattempted**

Export Cloud Datastore data using gcloud datastore export is not right.

?By all means, you can export a copy of all or a subset of entities from Google Cloud Datastore to another storage system such as Google Cloud Storage but your application is configured to connect to a Cloud Datastore instance, not another system that stores a raw dump of exported data. So this option is not right. Create a Cloud Datastore index using gcloud datastore indexes create. is not right.

?You could create an index but this doesn't help your application emulate connections to Cloud Datastore on your laptop. So this option is not right. Install the google-cloud-sdk-datastore-emulator component using the apt get install command. is not right.

?There is no such thing as google-cloud-sdk-datastore-emulator; and you don't install gcloud components using apt get. So this option is not right. Install the cloud-datastore-emulator component using the gcloud components install command. is the right answer.

?The Datastore emulator provides local emulation of the production Datastore environment. You can use the emulator to develop and test your application locally

?Ref: <https://cloud.google.com/datastore/docs/tools/datastore-emulator>

### 53. Question

You have a development project with appropriate IAM roles defined. You are creating a production project and want to have the same IAM roles on the new project, using the fewest possible steps. What should you do?

- Use gcloud iam roles copy and specify your organization as the destination organization.
- Use gcloud iam roles copy and specify the production project as the destination project.
- In the Google Cloud Platform Console, use the create role from role functionality.
- In the Google Cloud Platform Console, use the create role functionality and select all applicable permissions.

#### Unattempted

Our requirements are to create the same iam roles in a different (production) project with the fewest possible steps. In the Google Cloud Platform Console, use the ‘create role from role’ functionality. is not right.

?This creates a role in the same (development) project, not in the production project. So this doesn’t meet our requirement to create same iam roles in production project In the Google Cloud Platform Console, use the ‘create role’ functionality and select all applicable permissions. is not right.

?This creates a role in the same (development) project, not in the production project. So this doesn’t meet our requirement to create same iam roles in production project Use gcloud iam roles copy and specify your organization as the destination organization. is not right.

?We can optionally specify a destination organization but since our requirement is to copy the roles into “production project” (i.e. project, not organization), this option does not meet our requirement to create same iam roles in production project

?Ref: <https://cloud.google.com/sdk/gcloud/reference/iam/roles/copy> Use gcloud iam roles copy and specify the production project as the destination project. is the right answer.

?This is the only option that fits our requirements. You copy the roles into the destination project using gcloud iam roles copy and by specifying the production project destination project.

```
?$gcloud iam roles copy --source ">" --destination > --dest-project >
```

?Ref: <https://cloud.google.com/sdk/gcloud/reference/iam/roles/copy>

#### 54. Question

You have a Dockerfile that you need to deploy on Kubernetes Engine. What should you do?

- Use kubectl app deploy .
- Create a docker image from the Dockerfile and upload it to Container Registry. Create a Deployment YAML file to point to that image. Use kubectl to create the deployment with that file.
- Use gcloud app deploy .
- Create a docker image from the Dockerfile and upload it to Cloud Storage. Create a Deployment YAML file to point to that image. Use kubectl to create the deployment with that file.

**Unattempted**

Use kubectl app deploy . is not right.

?kubectl does not accept app as a verb. Kubectl can deploy a configuration file using kubectl deploy.

?Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#apply> Use gcloud app deploy . is not right.

?gcloud app deploy – Deploys the local code and/or configuration of your app to App Engine. gcloud app deploy accepts a flag –image-url which is the docker image but it can't directly use a docker file.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/app/deploy> Create a docker image from the Dockerfile and upload it to Cloud Storage. Create a Deployment YAML file to point to that image. Use kubectl to create the deployment with that file. is not right.

?You can not upload a docker image to cloud storage. They can only be pushed to a Container Registry (e.g. GCR, Dockerhub etc.)

?Ref: <https://cloud.google.com/container-registry/docs/pushing-and-pulling> Create a docker image from the Dockerfile and upload it to Container Registry. Create a Deployment YAML file to point to that image. Use kubectl to create the deployment with that file. is the right answer.

?Once you have a docker image, you can push it to the container register. You can then create a deployment YAML file pointing to this image and use kubectl apply -f to deploy this to the Kubernetes cluster. This assumes you already have a Kubernetes cluster and you gcloud environment is set up to talk to this container by executing gcloud container clusters get-credentials –zone=

?Ref: <https://cloud.google.com/container-registry/docs/pushing-and-pulling>

?Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#apply>

?Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/get-credentials>

## 55. Question

You have a Google Cloud Platform account with access to both production and development projects. You need to create an automated process to list all compute instances in development and production projects on a daily basis. What should you do?

- Create two configurations using gcloud config. Write a script that sets configurations as active, individually. For each configuration, use gcloud compute instances list to get a list of compute resources.
- Create two configurations using gsutil config. Write a script that sets configurations as active, individually. For each configuration, use gsutil compute instances list to get a list of compute resources.
- Go to Cloud Shell and export this information to Cloud Storage on a daily basis.
- Go to GCP Console and export this information to Cloud SQL on a daily basis.

### Unattempted

Go to Cloud Shell and export this information to Cloud Storage on a daily basis. is not right.

?You want an automated process but this is a manual activity that needs to be executed daily. Go to GCP Console and export this information to Cloud SQL on a daily basis. is not right.

?You want an automated process but this is a manual activity that needs to be executed daily. Create two configurations using gsutil config. Write a script that sets configurations as active, individually. For each configuration, use gsutil compute instances list to get a list of compute resources. is not right.

?The gsutil config command applies to users who have installed gsutil as a standalone tool and is used for obtaining access credentials for Cloud Storage and writes a boto/gsutil configuration file containing the obtained credentials along with a number of other configuration-controllable values.

?Ref: <https://cloud.google.com/storage/docs/gsutil/commands/config>

?It is not used for creating Gcloud configurations. You use gcloud config to do that.

?<https://cloud.google.com/sdk/gcloud/reference/config/configurations/create> Create two configurations using gcloud config. Write a script that sets configurations

as active, individually. For each configuration, use gcloud compute instances list to get a list of compute resources. is the right answer.

?You can create two configurations – one for the development project and another for the production project. And you do that by running “ gcloud config configurations create” command.

?<https://cloud.google.com/sdk/gcloud/reference/config/configurations/create>

?In your custom script, you can load these configurations one at a time and execute gcloud compute instances list to list Google Compute Engine instances in the project that is active in the gcloud configuration.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/instances/list>

?Once you have this information, you can export it in a suitable format to a suitable target e.g. export as CSV or export to Cloud Storage/BigQuery/SQL, etc.

## 56. Question

You have a large 5-TB AVRO file stored in a Cloud Storage bucket. Your analysts are proficient only in SQL and need access to the data stored in this file. You want to find a cost-effective way to complete their request as soon as possible. What should you do?

- Load data in Cloud Datastore and run a SQL query against it.
- Create a BigQuery table and load data in BigQuery. Run a SQL query on this table and drop this table after you complete your request.
- Create external tables in BigQuery that point to Cloud Storage buckets and run a SQL query on these external tables to complete your request.**
- Create a Hadoop cluster and copy the AVRO file to NDFS by compressing it. Load the file in a hive table and provide access to your analysts so that they can run SQL queries.

**Unattempted**

Load data in Cloud Datastore and run a SQL query against it. is not right.

?Datastore is a highly scalable NoSQL database and although it supports SQL like queries, it doesn’ t support SQL. Moreover, there is no out of the box way for transforming AVRO file from cloud storage into the Cloud Datastore entity. So we have to do in a bespoke way which adds to our cost and time.

?Ref: <https://cloud.google.com/datastore> Create a Hadoop cluster and copy the AVRO file to NDFS by compressing it. Load the file in a hive table and provide access to your analysts so that they can run SQL queries. is not right.

?Like Cloud Datastore, Hive doesn’ t directly support SQL, it provides HiveQL (HQL) which is SQL like. In addition, the process of creating a Hadoop cluster and getting the data eventually into a hive table is time-consuming and adds to our cost and time. Create a BigQuery table and load data in BigQuery. Run a

SQL query on this table and drop this table after you complete your request. is not right.

?Like the above two, while it is possible to build a solution that transforms and loads data into the target, BigQuery in this case, is not a trivial process and involves cost and time. GCP provides an out of the box way to query AVRO files from Cloud Storage and this should be preferred. Create external tables in BigQuery that point to Cloud Storage buckets and run a SQL query on these external tables to complete your request. is the right answer.

?BigQuery supports querying Cloud Storage data in a number of formats such as CSV, JSON, AVRO, etc. You do this by creating a Big Query external table that points to a Cloud Storage data source (bucket). This solution works out of the box, involves minimal effort, minimal cost, and is quick.

?<https://cloud.google.com/bigquery/external-data-cloud-storage>

## 57. Question

You have a Linux VM that must connect to Cloud SQL. You created a service account with the appropriate access rights. You want to make sure that the VM uses this service account instead of the default Compute Engine service account. What should you do?

- Download a JSON Private Key for the service account. On the Custom Metadata of the VM, add that JSON as the value for the key compute-engine-service-account
- Download a JSON Private Key for the service account. After creating the VM, ssh into the VM and save the JSON under ~/.gcloud/compute-engine-service-account.json
- When creating the VM via the web console, specify the service account under the ' Identity and API Access' section.
- Download a JSON Private Key for the service account. On the Project Metadata, add that JSON as the value for the key compute-engine-serviceaccount

## Unattempted

When creating the VM via the web console, specify the service account under the ' Identity and API Access' section. is the right answer.

You can set the service account at the time of creating the compute instance.

You can also update the service account used by the instance – this requires that you stop the instance first and then update the service account.

Setting/Updating the service account can be done either via the web console or by executing gcloud command or by the REST API. See below an example for updating the service account through gcloud command.

```
gcloud compute instances set-service-account instance-1 --zone=us-central1-a
--service-account=my-new-service-account@gcloud-gcp-ace-lab-
266520.iam.gserviceaccount.com
```

Updated [https://www.googleapis.com/compute/v1/projects/gcloud-gcp-ace-lab-266520/zones/us-central1-a/instances/instance-1].

Download a JSON Private Key for the service account. On the Project Metadata, add that JSON as the value for the key compute-engine-serviceaccount is not right.

While updating the service account for a compute instance can be done through the console, gcloud or the REST API, they don't do it based on the JSON Private Key.

Download a JSON Private Key for the service account. On the Custom Metadata of the VM, add that JSON as the value for the key compute-engine-service-account. is not right.

While updating the service account for a compute instance can be done through the console, gcloud or the REST API, they don't do it based on the JSON Private Key.

Download a JSON Private Key for the service account. After creating the VM, ssh into the VM and save the JSON under ~/.gcloud/compute-engine-service-account.json is not right.

You can configure a VM to use a certain service account by providing the relevant JSON credentials file, but the procedure is different. Copying the JSON file to a specific path alone is not sufficient, moreover, the path mentioned is wrong as well. See below for a use case where a VM which is unable to list cloud storage buckets is updated to use a service account and it can then list the buckets.

Prior to using a service account. Use gsutil ls to list buckets and it fails.

```
$ gsutil ls
```

```
ServiceException: 401 Anonymous caller does not have storage.buckets.list access to project 393066724129.
```

Within the VM, execute the command below to use the service account.

(Assumes that you have created a service account that provides the necessary permissions and have copied it over the VM)

```
gcloud auth activate-service-account admin-service-account@gcloud-gcp-ace-266520.iam.gserviceaccount.com --key-file=~/compute-engine-service-account.json
```

Activated service account credentials for: [admin-service-account@gcloud-gcp-ace-266520.iam.gserviceaccount.com]

The output above doesn't show this, but the credentials are written to the file /home/gcloud\_gcp\_ace\_user/.config/gcloud/legacy\_credentials/admin-service-account@gcloud-gcp-ace-266520.iam.gserviceaccount.com/adc.json

Now, use gsutil ls again to list buckets and it works.

```
$ gsutil ls
```

```
gs://test-gcloud-gcp-ace-2020-bucket-1/
gs://test-gcloud-gcp-ace-2020-bucket-2/
```

## 58. Question

You have a number of applications that have bursty workloads and are heavily dependent on topics to decouple publishing systems from consuming systems. Your company would like to go serverless to enable developers to focus on

writing code without worrying about infrastructure. Your solution architect has already identified Cloud Pub/Sub as a suitable alternative for decoupling systems. You have been asked to identify a suitable GCP Serverless service that is easy to use with Cloud Pub/Sub. You want the ability to scale down to zero when there is no traffic in order to minimize costs. You want to follow Google recommended practices. What should you suggest?

- Cloud Run for Anthos
- Cloud Functions**
- App Engine Standard
- Cloud Run

**Unattempted**

GCP serverless compute portfolio includes 4 services, which are all listed in the answer options. Our requirements are to identify a GCP serverless service that

?1. Lets us scale down to 0

?2. Integrates with Cloud Pub/Sub seamlessly Cloud Run for Anthos. is not right.

?Among the four options, App Engine Standard, Cloud Functions and Cloud Run can all scale down to zero. Cloud Run for Anthos can scale the pods down to zero but the number of nodes per cluster can not scale to zero so these nodes are billed in the absence of requests. This rules out Cloud Run for Anthos. App Engine Standard. is not right.

?App Engine Standard doesn't offer an out of the box integration with Cloud Pub/Sub. We can use the Cloud Client Library to send and receive Pub/Sub messages as described in the note below but the key point to note is the absence of out of the box integration with Cloud Pub/Sub so this rules out App Engine Standard

?Ref: <https://cloud.google.com/appengine/docs/standard/nodejs/writing-and-responding-to-pub-sub-messages> Cloud Run. is not right.

?Cloud Run is an excellent product and integrates with Cloud Pub/Sub for several use cases. For example, every time a new .csv file is created inside a Cloud Storage bucket, an event is fired and delivered via a Pub/Sub subscription to a Cloud Run service. The Cloud Run service extracts data from the file and stores it as structured data into a BigQuery table.

?Ref: <https://cloud.google.com/run#section-7>

?At the same time, we want to follow Google recommended practices. Google doesn't list integration with Cloud Pub/Sub as a key feature of Cloud Run. Contrary to this, Google says " If you're building a simple API (a small set of

functions to be accessed via HTTP or Cloud Pub/Sub), we recommend using Cloud Functions.” Cloud Functions. is the right answer.

?Cloud Functions is Google Cloud’ s event-driven serverless compute platform that lets you run your code locally or in the cloud without having to provision servers. Cloud Functions scales up or down, so you pay only for compute resources you use. Cloud Functions have excellent integration with Cloud Pub/Sub, lets you scale down to zero and is recommended by Google as the ideal serverless platform to use when dependent on Cloud Pub/Sub.

?“ If you’ re building a simple API (a small set of functions to be accessed via HTTP or Cloud Pub/Sub), we recommend using Cloud Functions.”

?Ref: <https://cloud.google.com/serverless-options>

#### 59. Question

You have a number of compute instances belonging to an unmanaged instances group. You need to SSH to one of the Compute Engine instances to run an ad hoc script. You’ ve already authenticated gcloud, however, you don’ t have an SSH key deployed yet. In the fewest steps possible, what’ s the easiest way to SSH to the instance?

- Create a key with the ssh-keygen command. Upload the key to the instance. Run gcloud compute instances list to get the IP address of the instance, then use the ssh command.
- Run gcloud compute instances list to get the IP address of the instance, then use the ssh command.
- Create a key with the ssh-keygen command. Then use the gcloud compute ssh command.
- Use the gcloud compute ssh command.

Unattempted

Create a key with the ssh-keygen command. Upload the key to the instance. Run gcloud compute instances list to get the IP address of the instance, then use the ssh command. is not right.

?This approach certainly works. You can create a key pair with ssh-keygen, update the instance metadata with the public key and SSH to the instance. But is it the easiest way to SSH to the instance with the fewest possible steps? Let’ s explore other options to decide (you will see that there is another option that does the same with less effort). You can find more information about this option here: <https://cloud.google.com/compute/docs/instances/adding-removing-ssh-keys#block-project-keys> Create a key with the ssh-keygen command. Then use the gcloud compute ssh command. is not right.

?This works but is more work (having to create the key) than the answer. gcloud compute ssh ensures that the user’ s public SSH key is present in the project’ s metadata. If the user does not have a public SSH key, one is generated using ssh-keygen and added to the project’ s metadata. Run gcloud compute instances

list to get the IP address of the instance, then use the ssh command. is not right.

?We can get the IP of the instance by executing the gcloud compute instances list but unless an SSH is generated and updated in project metadata, you would not be able to SSH to the instance. User access to a Linux instance through third-party tools is determined by which public SSH keys are available to the instance. You can control the public SSH keys that are available to a Linux instance by editing metadata, which is where your public SSH keys and related information are stored.

?Ref: <https://cloud.google.com/compute/docs/instances/adding-removing-ssh-keys#block-project-keys> Use the gcloud compute ssh command. is the right answer.

?gcloud compute ssh ensures that the user's public SSH key is present in the project's metadata. If the user does not have a public SSH key, one is generated using ssh-keygen and added to the project's metadata. This is similar to the other option where we copy the key explicitly to the project's metadata but here it is done automatically for us. There are also security benefits with this approach. When we use gcloud compute ssh to connect to Linux instances, we are adding a layer of security by storing your host keys as guest attributes. Storing SSH host keys as guest attributes improve the security of your connections by helping to protect against vulnerabilities such as man-in-the-middle (MITM) attacks. On the initial boot of a VM instance, if guest attributes are enabled, Compute Engine stores your generated host keys as guest attributes. Compute Engine then uses these host keys that were stored during the initial boot to verify all subsequent connections to the VM instance.

?Ref: <https://cloud.google.com/compute/docs/instances/connecting-to-instance>

?Ref: <https://cloud.google.com/sdk/gcloud/reference/compute/ssh>

#### 60. Question

You need to create a custom VPC with a single subnet. The subnet's range must be as large as possible. Which range should you use?

- 10.0.0.0/8
- 192.168.0.0/16
- 172.16.0.0/12
- 0.0.0.0/0

**Unattempted**

The private network range is defined by IETF  
(Ref: <https://tools.ietf.org/html/rfc1918>) and adhered to by all cloud providers. The supported internal IP Address ranges are

1. 24-bit block 10.0.0.0/8 (16777216 IP Addresses)
2. 20-bit block 172.16.0.0/12 (1048576 IP Addresses)
3. 16-bit block 192.168.0.0/16 (65536 IP Addresses)

10.0.0.0/8 gives you the largest range – 16777216 IP Addresses.

## 61. Question

You have a project for your App Engine application that serves a development environment. The required testing has succeeded and you want to create a new project to serve as your production environment. What should you do?

- Deploy your application again using gcloud and specify the project parameter with the new project name to create the new project.
- Use gcloud to create the new project and to copy the deployed application to the new project.
- Use gcloud to create the new project, and then deploy your application to the new project.
- Create a Deployment Manager configuration file that copies the current App Engine deployment into a new project.

### Unattempted

Use gcloud to create the new project and to copy the deployed application to the new project. is not right.

?You can use gcloud to create a new project but you can not copy a deployed application from one project to another. This feature is not offered by Google App Engine. Create a Deployment Manager configuration file that copies the current App Engine deployment into a new project. is not right.

?The deployment manager configuration file contains configuration about the resources that need to be created in Google cloud, however, it does not offer the feature to copy app engine deployment into a new project. Deploy your application again using gcloud and specify the project parameter with the new project name to create the new project. is not right.

?You can deploy using gcloud app deploy and target it to a different project using –project flag. However, you can only deploy to an existing project as the gcloud app deploy command is unable to create a new project if it doesn't already exist. Use gcloud to create the new project, and then deploy your application to the new project. is the right answer.

?You can deploy to a different project by using –project flag.

?By default, the service is deployed the current project configured via:

?\$ gcloud config set core/project PROJECT

?To override this value for a single deployment, use the –project flag:

```
?$ gcloud app deploy ~/my_app/app.yaml --project=PROJECT
```

?Ref: <https://cloud.google.com/sdk/gcloud/reference/app/deploy>

## 62. Question

You have a single binary application that you want to run on Google Cloud Platform. You decided to automatically scale the application based on underlying infrastructure CPU usage. Your organizational policies require you to use virtual machines directly. You need to ensure that the application scaling is operationally efficient and completed as quickly as possible. What should you do?

- Create a Google Kubernetes Engine cluster, and use horizontal pod autoscaling to scale the application.
- Create an instance template, and use the template in a managed instance group with autoscaling configured.
- Create an instance template, and use the template in a managed instance group that scales up and down based on the time of day.
- Use a set of third-party tools to build automation around scaling the application up and down, based on Stackdriver CPU usage monitoring

**Unattempted**

Our requirements are

?1. Use Virtual Machines directly (i.e. not container-based)

?2. Scale Automatically

?3. Scaling is efficient & is quick Create a Google Kubernetes Engine cluster, and use horizontal pod autoscaling to scale the application. is not right.

?We want to use virtual machines directly. And although GKE uses virtual machines under the hood for its GKE cluster, the autoscaling is totally different – it uses scaling at VMs (cluster auto-scaling) as well as at pod level (horizontal and vertical pod autoscaling). Create an instance template, and use the template in a managed instance group that scales up and down based on the time of day. is not right.

?Scaling based on time of the day may be insufficient especially when there is a sudden surge of requests (causing high CPU utilization) or if the requests go down suddenly (resulting in low CPU usage). Our requirements state we need to scale automatically i.e. we need autoscaling solution that scales up and down based on CPU usage which is indicative of the volume of requests processed but scaling based on time of the day is not indicative of the load (CPU) on the system and is therefore not right. Use a set of third-party tools to build automation around

scaling the application up and down, based on Stackdriver CPU usage monitoring. is not right.

?While this can be done, it is not the most efficient solution when Google's own services offer this functionality and can do it more efficiently as they are all natively integrated. Create an instance template, and use the template in a managed instance group with autoscaling configured. is the right answer.

?Managed instance groups offer autoscaling capabilities that let you automatically add or delete instances from a managed instance group based on increases or decreases in load (CPU Utilization in this case). Autoscaling helps your apps gracefully handle increases in traffic and reduce costs when the need for resources is lower. You define the autoscaling policy and the autoscaler performs automatic scaling based on the measured load (CPU Utilization in this case). Autoscaling works by adding more instances to your instance group when there is more load (upscale), and deleting instances when the need for instances is lowered (downscale).

?Ref: <https://cloud.google.com/compute/docs/autoscaler>

### 63. Question

You have a virtual machine that is currently configured with 2 vCPUs and 4 GB of memory. It is running out of memory. You want to upgrade the virtual machine to have 8GB of memory. What should you do?

- Stop the VM, increase the memory to 8 GB and start the VM
  - Rely on live migration to move the workload to a machine with more memory.
  - Stop the VM, change the machine type to n1-standard-2 and start the VM
  - Use gcloud to add metadata to the VM. Set the key to required-memory-size and the value to 8 GB

**Unattempted**

Rely on live migration to move the workload to a machine with more memory. is not right.

?Live migration migrates your running instances to another host in the same zone so that Google can perform maintenance such as a software or hardware update. It can not be used for changing machine type.

?Ref: <https://cloud.google.com/compute/instances/live-migration> Use gcloud to add metadata to the VM. Set the key to required-memory-size and the value to 8 GB. is not right.

?There is no such setting as required-memory-size. Stop the VM, change the machine type to n1-standard-2 and start the VM. is not right.

?n1-standard-2 instance offers less than 8 GB (7.5 GB to be precise) so this falls short of the required memory.

?Ref: <https://cloud.google.com/compute/docs/machine-types> Stop the VM, increase the memory to 8 GB and start the VM. is the right answer.

?In Google compute engine, if predefined machine types don't meet your needs, you can create an instance with custom virtualized hardware settings. Specifically, you can create an instance with a custom number of vCPUs and custom memory, effectively using a custom machine type. Custom machine types are ideal for the following scenarios:

?1. Workloads that aren't a good fit for the predefined machine types that are available to you.

?2. Workloads that require more processing power or more memory but don't need all of the upgrades that are provided by the next machine type level.

?In our scenario, we only need a memory upgrade. Moving to a bigger instance would also bump up the CPU which we don't need so we have to use a custom machine type. It is not possible to change memory while the instance is running so you need to first stop the instance, change the memory and then start it again. See below a screenshot that shows how CPU/Memory can be customized for an instance that has been stopped.

?Ref: <https://cloud.google.com/compute/docs/instances/creating-instance-with-custom-machine-type>

#### 64. Question

You have a web application deployed as a managed instance group based on an instance template. You modified the startup script used in the instance template and would like the existing instances to pick up changes from the new startup scripts. Your web application is currently serving live web traffic. You want to propagate the startup script changes to all instances in the managed instances group while minimizing effort, minimizing cost and ensuring that the available capacity does not decrease. What would you do?

- Delete instances in the managed instance group (MIG) one at a time and rely on auto-healing to provision an additional instance.
- Perform a rolling-action replace with max-unavailable set to 0 and max-surge set to 1**
- Create a new managed instance group (MIG) based on a new template. Add the group to the backend service for the load balancer. When all instances in the new managed instance group are healthy, delete the old managed instance group
- Perform a rolling-action start-update with max-unavailable set to 1 and max-surge set to 0

## Unattempted

Perform a rolling-action start-update with max-unavailable set to 1 and max-surge set to 0. is not right.

?You can carry out a rolling action start update to fully replace the template by executing a command like

```
?gcloud compute instance-groups managed rolling-action start-update instance-group-1 --zone=us-central1-a --version template=instance-template-1 --canary-version template=instance-template-2,target-size=100%
```

?which updates the instance-group-1 to use instance-template-2 instead of instance-template-1 and have instances created out of instance-template-2 serve 100% of traffic.

?However, the values specified for maxSurge and maxUnavailable mean that we will lose capacity which is against our requirements.

?maxSurge specifies the maximum number of instances that can be created over the desired number of instances. If maxSurge is set to 0, the rolling update can not create additional instances and is forced to update existing instances. This results in a reduction in capacity and therefore does not satisfy our requirement to ensure that the available capacity does not decrease during the deployment.

?maxUnavailable – specifies the maximum number of instances that can be unavailable during the update process. When maxUnavailable is set to 1, the rolling update updates 1 instance at a time. i.e. it takes 1 instance out of service, updates it, and puts it back into service. This results in a reduction in capacity while the instance is out of service. Example – if we have 10 instances in service, this combination of setting results in 1 instance at a time taken out of service for replacement while the remaining 9 continue to serve live traffic. That's a reduction of 10% in available capacity.

?Ref: <https://kubernetes.io/docs/concepts/workloads/controllers/deployment/#max-unavailable>

?Ref: <https://kubernetes.io/docs/concepts/workloads/controllers/deployment/#max-surge> Create a new managed instance group (MIG) based on a new template. Add the group to the backend service for the load balancer. When all instances in the new managed instance group are healthy, delete the old managed instance group. is not right.

?While the end result is the same, we have a period of time where the traffic is served by instances from both the old managed instances group (MIG) which doubles our cost and increases effort and complexity. Delete instances in the managed instance group (MIG) one at a time and rely on auto-healing to provision an additional instance. is not right.

?While this would result in the same eventual outcome, there are two issues with this approach. First, deleting an instance one at a time would result in a reduction in capacity which is against our requirements. Secondly, deleting instances manually one at a time is error-prone and time-consuming. One of our requirements is to “ minimize the effort” but deleting instances manually and relying on auto-healing health checks to provision them back is time-consuming and could take a lot of time depending on the number of instances in the MIG and the startup scripts executed during bootstrap. Perform a rolling-action replace with max-unavailable set to 0 and max-surge set to 1. is the right answer.

?This option achieves the outcome in the most optimal manner. The replace action is used to replace instances in a managed instance group. When maxUnavailable is set to 0, the rolling update can not take existing instances out of service. And when maxSurge is set to 1, we let the rolling update spin a single additional instance. The rolling update then puts the additional instance into service and takes one of the existing instances out of service for replacement. There is no reduction in capacity at any point in time.

Ref: <https://kubernetes.io/docs/concepts/workloads/controllers/deployment/#max-unavailable>

Ref: <https://kubernetes.io/docs/concepts/workloads/controllers/deployment/#max-surge>

Ref: <https://cloud.google.com/sdk/gcloud/reference/alpha/compute/instance-groups/managed/rolling-action/replace>

## 65. Question

You have a web application deployed as a managed instance group. You have a new version of the application to gradually deploy. Your web application is currently serving live web traffic. You want to ensure that the available capacity does not decrease during the deployment. What would you do?

- Create a new instance template with the new application version. Update the existing managed instance group with the new instance template. Delete the instances in the managed instance group to allow the managed instance group to recreate the instance using the new instance template.
- Perform a rolling-action start-update with maxSurge set to 0 and maxUnavailable set to 1
- Create a new managed instance group with an updated instance template. Add the group to the backend service for the load balancer. When all instances in the new managed instance group are healthy, delete the old managed instance group.
- Perform a rolling-action start-update with maxSurge set to 1 and maxUnavailable set to 0

**Unattempted**

Our requirements are

?1. Deploy a new version gradually

?2. Ensure available capacity does not decrease during deployment Create a new managed instance group with an updated instance template. Add the group to the backend service for the load balancer. When all instances in the new managed instance group are healthy, delete the old managed instance group. is not right.

?First of all instance templates can not be updated. So the phrase updated instance template rules out this option.

?Ref: <https://cloud.google.com/compute/docs/instance-templates/> Create a new instance template with the new application version. Update the existing managed instance group with the new instance template. Delete the instances in the managed instance group to allow the managed instance group to recreate the instance using the new instance template. is not right.

?If we follow these steps, we end up with a full fleet of instances belonging to the new managed instances group (i.e. based on the new template) behind the load balancer, but our requirement to gradually deploy the new version is not met. In addition, deleting the existing instances of the managed instance group would almost certainly result in an outage to our application which is not desirable when we are serving live web traffic. Perform a rolling-action start-update with maxSurge set to 0 and maxUnavailable set to 1 is not right.

?maxSurge specifies the maximum number of instances that can be created over the desired number of instances. If maxSurge is set to 0, the rolling update can not create additional instances and is forced to update existing instances. This results in a reduction in capacity and therefore does not satisfy our requirement to ensure that the available capacity does not decrease during the deployment.

?maxUnavailable – specifies the maximum number of instances that can be unavailable during the update process. When maxUnavailable is set to 1, the rolling update updates 1 instance at a time. i.e. it takes 1 instance out of service, updates it, and puts it back into service. This results in a reduction in capacity while the instance is out of service. Example – if we have 10 instances in service, this combination of setting results in 1 instance at a time taken out of service for an upgrade while the remaining 9 continue to serve live traffic. That's a reduction of 10% in available capacity and does not satisfy our requirement to ensure that the available capacity does not decrease during the deployment. Perform a rolling-action start-update with maxSurge set to 1 and maxUnavailable set to 0 is the right answer.

?This is the only option that satisfies our two requirements – deploying gradually and ensuring the available capacity does not decrease. When maxUnavailable is set to 0, the rolling update can not take existing instances out of service. And when maxSurge is set to 1, we let the rolling update spin a single additional instance. The rolling update then puts the additional instance into service and takes one of the existing instances out of service for the upgrade. There is no reduction in capacity at any point in time. And the rolling upgrade upgrades 1 instance at a time so we gradually deploy the new version. Example – if we have 10 instances in service, this combination of setting results in 1 additional instance

put into service (resulting in 11 instances serving traffic), then takes an older instance out of service (resulting in 10 instances serving traffic) and puts the upgraded instance back into service (resulting in 11 instances serving traffic). The rolling upgrade continues updating the remaining 9 instances one at a time. Finally, when all 10 instances have been upgraded, the additional instance that is spun up is deleted. We still have 10 instances serving live traffic but now on the new version of code.

?Ref: <https://kubernetes.io/docs/concepts/workloads/controllers/deployment/#max-unavailable>

?Ref: <https://kubernetes.io/docs/concepts/workloads/controllers/deployment/#max-surge>

# SET-10

## 1. Question

You have a web application deployed as a managed instance group. You noticed some of the compute instances are running low on memory. You suspect this is due to JVM memory leak and you want to restart the compute instances to reclaim the leaked memory. Your web application is currently serving live web traffic. You want to ensure that the available capacity does not go below 80% at any time during the restarts and you want to do this at the earliest. What would you do?

- Stop instances in the managed instance group (MIG) one at a time and rely on autohealing to bring them back up.
- Perform a rolling-action replace with max-unavailable set to 20%.
- Perform a rolling-action restart with max-unavailable set to 20%.
- Perform a rolling-action reboot with max-surge set to 20%.

**Incorrect**

Perform a rolling-action reboot with max-surge set to 20%. is not right.

?reboot is not a supported action for rolling updates. The supported actions are replace, restart, start-update and stop-proactive-update.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/beta/compute/instance-groups/managed/rolling-action> Perform a rolling-action replace with max-unavailable set to 20%. is not right.

?Performing a rolling-action replace – Replaces instances in a managed instance group. While this resolves the JVM memory leak issue, recreating the instances is a little drastic when the same result can be achieved with the simple restart action. One of our requirements is to “do this at the earliest” but recreating instances might take a lot of time depending on the number of instances and startup scripts; certainly more time than restart action.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/beta/compute/instance-groups/managed/rolling-action> Stop instances in the managed instance group (MIG) one at a time and rely on autohealing to bring them back up. is not right.

?While this would result in the same eventual outcome, it is manual, error-prone and time-consuming. One of our requirements is to “do this at the earliest” but stopping instances manually is time-consuming and could take a lot of time depending on the number of instances in the MIG. Also, relying on autohealing health checks to detect the failure and spin up the instance adds to the delay. Perform a rolling-action restart with max-unavailable set to 20%. is the right answer.

?This option achieves the outcome in the most optimal manner. The restart action restarts instances in a managed instance group. By performing a rolling

restart with max-unavailable set to 20%, the rolling update restarts instances while ensuring there is at least 80% available capacity. The rolling update carries on restarting all the remaining instances until all instances in the MIG have been restarted.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/alpha/compute/instance-groups/managed/rolling-action/restart>

2. 2. Question

You have a website hosted on App Engine standard environment. You want 1% of your users to see a new test version of the website. You want to minimize complexity. What should you do?

- Create a new App Engine application in the same project. Deploy the new version in that application. Configure your network load balancer to send 1% of the traffic to that new application.
- Create a new App Engine application in the same project. Deploy the new version in that application. Use the App Engine library to proxy 1% of the requests to the new version.
- Deploy the new version in the same application and use the --migrate option.
- Deploy the new version in the same application and use the --splits option to give a weight of 99 to the current version and a weight of 1 to the new version.

Unattempted

Deploy the new version in the same application and use the –migrate option. is not right.

?migrate is not a valid flag for the gcloud app deploy command.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/app/deploy>

?Also, gcloud app versions migrate, which is a valid command to migrate traffic from one version to another for a set of services, is not suitable either as we only want to send 1% traffic.

?<https://cloud.google.com/sdk/gcloud/reference/app/versions/migrate> Create a new App Engine application in the same project. Deploy the new version in that application. Use the App Engine library to proxy 1% of the requests to the new version. is not right.

?While this can be done, we are increasing complexity and do not meet our requirement “ minimize complexity” . There is an out of the box option in the app engine to split traffic in a seamless way. Create a new App Engine application in the same project. Deploy the new version in that application. Configure your network load balancer to send 1% of the traffic to that new application. is not right.

?Instances that participate as backend VMs for network load balancers must be running the appropriate Linux guest environment, Windows guest environment, or other processes that provide equivalent functionality. Network load balancer is not suitable for App Engine standard environment which is container-based and provide us specific runtimes without any promise on the underlying guest environments. Deploy the new version in the same application and use the –splits option to give a weight of 99 to the current version and a weight of 1 to the new version. is the right answer.

?You can use traffic splitting to specify a percentage distribution of traffic across two or more of the versions within a service. Splitting traffic allows you to conduct A/B testing between your versions and provides control over the pace when rolling out features.

?For this scenario, we can split the traffic as shown below, sending 1% to v2 and 99% to v1

?by executing the command gcloud app services set-traffic service1 –splits v2=1,v1=99

?Ref: <https://cloud.google.com/sdk/gcloud/reference/app/services/set-traffic>

### 3. Question

You have a workload running on Compute Engine that is critical to your business. You want to ensure that the data on the boot disk of this workload is backed up regularly. You need to be able to restore a backup as quickly as possible in case of disaster. You also want older backups to be cleaned automatically to save on cost. You want to follow Google-recommended practices. What should you do?

- Create a Cloud Function to create an instance template.
- Create a snapshot schedule for the disk using the desired interval.
- Create a cron job to create a new disk from the disk using gcloud.
- Create a Cloud Task to create an image and export it to Cloud Storage.

**Unattempted**

Create a Cloud Function to create an instance template. is not right.

?This does not fulfil our requirement of backing up data on boot disk ‘regularly’ . Create a cron job to create a new disk from the disk using gcloud. is not right.

?Like above, this does not fulfil our requirement of backing up data on boot disk ‘regularly’ . Create a Cloud Task to create an image and export it to Cloud Storage. is not right.

?Like above, this does not fulfil our requirement of backing up data on boot disk ‘regularly’. Create a snapshot schedule for the disk using the desired interval. is the right answer.

?Create snapshots to periodically back up data from your zonal persistent disks or regional persistent disks. To reduce the risk of unexpected data loss, consider the best practice of setting up a snapshot schedule to ensure your data is backed up on a regular schedule.

?Ref: <https://cloud.google.com/compute/docs/disks/create-snapshots>

?You can also delete snapshots on a schedule by defining a snapshot retention policy. A snapshot retention policy defines how long you want to keep your snapshots. If you choose to set up a snapshot retention policy, you must do so as part of your snapshot schedule.

?Ref: [https://cloud.google.com/compute/docs/disks/scheduled-snapshots#retention\\_policy](https://cloud.google.com/compute/docs/disks/scheduled-snapshots#retention_policy)

4. 4. Question

You have an application deployed in a GKE Cluster as a Kubernetes workload with Daemon Sets. Your application has become very popular and is now struggling to cope up with increased traffic. You want to add more pods to your workload and want to ensure your cluster scales up and scales down automatically based on volume. What should you do?

- Perform a rolling update to modify machine type from n1-standard-2 to n1-standard-4.
- Enable autoscaling on Kubernetes Engine.**
- Enable Horizontal Pod Autoscaling for the Kubernetes deployment.
- Create another identical Kubernetes workload and split traffic between the two workloads.

**Unattempted**

Enable Horizontal Pod Autoscaling for the Kubernetes deployment. is not right.

?Horizontal Pod Autoscaling can not be enabled for Daemon Sets, this is because there is only one instance of a pod per node in the cluster. In a replica deployment, when Horizontal Pod Autoscaling scales up, it can add pods to the same node or another node within the cluster. Since there can only be one pod per node in the Daemon Set workload, Horizontal Pod Autoscaling is not supported with Daemon Sets.

?Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/daemonset> Create another identical Kubernetes cluster and split traffic between the two workloads. is not right.

?Creating another identical Kubernetes cluster is going to double your costs; at the same time, there is no guarantee that this is enough to handle all the traffic. Finally, it doesn't satisfy our requirement of "cluster scales up and scales down automatically". Perform a rolling update to modify machine type from n1-standard-2 to n1-standard-4. is not right.

?While increasing the machine type from n1-standard-2 to n1-standard-4 gives the existing nodes more resources and processing power, we don't know if that would be enough to handle the increased volume of traffic. Also, it doesn't satisfy our requirement of "cluster scales up and scales down automatically".

?Ref: <https://cloud.google.com/compute/docs/machine-types> Enable autoscaling on Kubernetes Engine. is the right answer.

?GKE's cluster autoscaler automatically resizes the number of nodes in a given node pool, based on the demands of your workloads. As you add nodes to a node pool, DaemonSets automatically add Pods to the new nodes as needed. DaemonSets attempt to adhere to a one-Pod-per-node model.

?Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/cluster-autoscaler>

## 5. Question

You have an application on a general-purpose Compute Engine instance that is experiencing excessive disk read throttling on its Zonal SSD Persistent Disk. The application primarily reads large files from disk. The disk size is currently 350 GB. You want to provide the maximum amount of throughput while minimizing costs. What should you do?

- Increase the size of the disk to 1 TB.
- Increase the allocated CPU to the instance.
- Migrate to use a Local SSD on the instance.**
- Migrate to use a Regional SSD on the instance.

**Unattempted**

Migrate to use a Regional SSD on the instance. is not right.

Migrating to a regional SSD would actually make it worse. At the time of writing, the Read IOPS for a Zonal standard persistent disks is 7,500 and the Read IOPS reduces to 3000 for a Regional standard persistent disks which reduces the throughput.

Ref: <https://cloud.google.com/compute/docs/disks/performance>

Increase the size of the disk to 1 TB. is not right.

The performance of SSD persistent disks scales with the size of the disk.

Ref: [https://cloud.google.com/compute/docs/disks/performance#cpu\\_count\\_size](https://cloud.google.com/compute/docs/disks/performance#cpu_count_size) However, there is no guarantee that increasing the disk to 1 TB will increase the throughput in this scenario as disk performance also depends on the number of vCPUs on VM instance.

Ref: [https://cloud.google.com/compute/docs/disks/performance#ssd\\_persistent\\_d](https://cloud.google.com/compute/docs/disks/performance#ssd_persistent_d)

### disk\_performance\_by\_disk\_size

Ref: <https://cloud.google.com/compute/docs/disks/performance#machine-type-disk-limits>

For example, consider a 1,000 GB SSD persistent disk attached to an instance with an N2 machine type and 4 vCPUs. The read limit based solely on the size of the disk is 30,000 IOPS. However, because the instance has 4 vCPUs, the read limit is restricted to 15,000 IOPS.

Increase the allocated CPU to the instance. is not right.

In Compute Engine, machine types are grouped and curated for different workloads. Each machine type is subject to specific persistent disk limits per vCPU. Increasing the vCPU count increases the Read IOPS

<https://cloud.google.com/compute/docs/disks/performance#machine-type-disk-limits>

However, there is no guarantee that increasing CPU will definitely increase the throughput in this scenario as disk performance could be limited by disk size.

Ref: [https://cloud.google.com/compute/docs/disks/performance#ssd\\_persistent\\_disk\\_performance\\_by\\_disk\\_size](https://cloud.google.com/compute/docs/disks/performance#ssd_persistent_disk_performance_by_disk_size)

Ref: <https://cloud.google.com/compute/docs/disks/performance#machine-type-disk-limits>

For example, consider a 1,000 GB SSD persistent disk attached to an instance with an N2 machine type and 48 vCPUs.

The read limit based solely on the vCPU count is 60,000 IOPS. However, because the instance has 1000 GB SSD, the read limit is restricted to 30,000 IOPS.

Migrate to use a Local SSD on the instance. is the right answer.

Local SSDs are physically attached to the server that hosts your VM instance. Local SSDs have higher throughput and lower latency than standard persistent disks or SSD persistent disks. The performance gains from local SSDs require certain trade-offs in availability, durability, and flexibility. Because of these trade-offs, Local SSD storage isn't automatically replicated and all data on the local SSD might be lost if the instance terminates for any reason.

Ref: <https://cloud.google.com/compute/docs/disks#localssds>

Ref: [https://cloud.google.com/compute/docs/disks/performance#type\\_comparison](https://cloud.google.com/compute/docs/disks/performance#type_comparison)

## 6. Question

You have an application running in App Engine standard environment. You want to add a custom C# library to enhance the functionality of this application.

However, C# isn't supported by App Engine standard. You want to maintain the serverless aspect of your application. What should you do? Choose 2 answers.

- Containerize your new application and deploy it to a Cloud Run on GKE environment.
- Containerize your new application and deploy it to a Cloud Run environment.
- Containerize your new application and deploy it to a App Engine flexible environment.

- Containerize your new application and deploy it to a Google Kubernetes Engine environment.
- Split your application into different functions. Deploy your application as separate cloud functions in Google Cloud Functions (GCP) environment.

### Unattempted

App engine standard currently supports Python, Java, Node.js, PHP, Ruby and Go.

?Ref: <https://cloud.google.com/appengine/docs/standard/>

?The question already states C# isn't supported by App Engine. Our requirement is to ensure we maintain the serverless aspect of our application. Split your application into different functions. Deploy your application as separate cloud functions in Google Cloud Functions (GCP) environment is not right.

?Cloud Functions is a serverless platform where you can run the code in the cloud without having to provision servers. You split your application functionality into multiple functions, and each of these is defined as a cloud function. Cloud Functions don't support C#. Supported runtimes are Python, Node.js and Go.

?Ref: <https://cloud.google.com/functions> Containerize your new application and deploy it to a App Engine flexible environment is not right.

?While App Engine flexible lets us customize runtimes or provide our own runtime by supplying a custom Docker image or Dockerfile from the open-source community, it uses compute engine virtual machines so it is not serverless.

Ref: <https://cloud.google.com/appengine/docs/flexible/> Containerize your new application and deploy it to a Google Kubernetes Engine environment. is not right.

?GKE i.e. Google Kubernetes Clusters uses compute engine virtual machines so it is not serverless.

?Ref: <https://cloud.google.com/kubernetes-engine> Containerize your new application and deploy it to a Cloud Run environment. is the right answer.

?Cloud Run is a fully managed compute platform that automatically scales your stateless containers. Cloud Run is serverless: it abstracts away all infrastructure management, so you can focus on what matters most—building great applications. Run your containers in fully managed Cloud Run or on Anthos, which supports both Google Cloud and on-premises environments. Cloud Run is built upon an open standard, Knative, enabling the portability of your applications.

?Ref: <https://cloud.google.com/run> Containerize your new application and deploy it to a Cloud Run on GKE environment. is the right answer.

?Cloud Run implements the Knative serving API, an open-source project to run serverless workloads on top of Kubernetes. That means you can deploy Cloud Run services anywhere Kubernetes runs. And if you need more control over your services (like access to GPU or more memory), you can also deploy these serverless containers in your own GKE cluster instead of using the fully managed

environment. When using the fully managed environment, Cloud Run on GKE is serverless.

Ref: <https://cloud.google.com/blog/products/serverless/cloud-run-bringing-serverless-to-containers>

7. 7. Question

You have an application running in Google Kubernetes Engine (GKE) with cluster autoscaling enabled. This application exposes a TCP endpoint. There are several replicas of the application. You have a Compute Engine instance in the same region but in another Virtual Private Cloud (VPC) called pt-network that has no overlapping CIDR range with the other VPC. The instance needs to connect to the application on GKE. You want to minimize effort. What should you do?

- 1. In GKE, create a service of type LoadBalancer that uses the application' s pods as backend. 2. Set the service' s externalTrafficPolicy to Cluster. 3. Configure the Compute Engine instance to use the address of the load balancer that has been assigned.
- 1. In GKE, create a service of type NodePort that uses the application' s pods as backend. 2. Create a Compute Engine instance called proxy with 2 network interfaces one in VPC. 3. Use iptables on the instance to forward traffic from pt-network to the GKE nodes. 4. Configure the Compute Engine instance to use the address of proxy in pt-network as endpoint.
- 1. In GKE, create a Service of type LoadBalancer that uses the application' s Pods as backend. 2. Add an annotation to this service `cloud.google.com/load-balancer-type: Internal` 3. Peer the two VPCs together 4. Configure the Compute Engine instance to use the address of the load balancer that has been created.
- 1. In GKE, create a Service of type LoadBalancer that uses the application' s Pods as backend. 2. Add a Cloud Armor Security Policy to the load balancer that whitelists the internal IPs of the MIG' s instances. 3. Configure the Compute Engine instance to use the address of the load balancer that has been created.

**Unattempted**

While it may be possible to set up the networking to let the compute engine instance in pt-network communicate with pods in the GKE cluster in multiple ways, we need to look for an option that minimizes effort. Generally speaking, this means using Google Cloud Platform services directly and configuring them to achieve the intended outcome; over setting up a service ourselves, installing/managing/upgrading it ourselves which is manual, error-prone, time-consuming and add to operational overhead.

1. In GKE, create a service of type LoadBalancer that uses the application' s pods as backend.
2. Set the service' s externalTrafficPolicy to Cluster.
3. Configure the Compute Engine instance to use the address of the load balancer that has been assigned. is not right.

In GKE, services are used to expose pods to the outside world. There are multiple types of services. The three common types are – NodePort, ClusterIP, and LoadBalancer (there are two more service types – ExternalName and Headless which are not relevant in this context). We do not want to create a Cluster IP as this is not accessible outside the cluster. And we do not want to create NodePort as this results in exposing a port on each node in the cluster; and as we have multiple replicas, this will result in them trying to open the same port on the nodes which fail. The compute engine instance in pt-network needs a single point of communication to reach GKE. This is achieved by creating a service of type LoadBalancer. This gives the service a public IP that is externally accessible.

Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/exposing-apps>  
externalTrafficPolicy denotes how the service should route external traffic – including public access. Rather than trying to explain, I'll point you to a very good blog that does a great job of answering how this works. <https://www.asykim.com/blog/deep-dive-into-kubernetes-external-traffic-policies>

Since we have cluster autoscaling enabled, we can have more than 1 node and possibly multiple replicas running on each node. So externalTrafficPolicy set to Cluster plays well with our requirement.

Finally, we configure the compute engine to use the (externally accessible) address of the load balancer.

So this certainly looks like an option, but is it the best option that minimizes effort? One of the disadvantages of this option is that it exposes the pods publicly by using a service of type LoadBalancer. We want our compute engine to talk to the pods, but do we really want to expose our pods to the whole world? Maybe not!! Let's look at the other options to find out if there is something more relevant and secure.

1. In GKE, create a service of type NodePort that uses the application's pods as backend.
2. Create a Compute Engine instance called proxy with 2 network interfaces one in VPC.
3. Use iptables on the instance to forward traffic from pt-network to the GKE nodes.
4. Configure the Compute Engine instance to use the address of proxy in pt-network as endpoint. is not right.

For reasons explained in the above option, we don't want to create a service of type NodePort. This opens up a port on each node for each replica (pod). If we choose to do this, the compute engine doesn't have a single point to contact. Instead, it would need to contact the GKE cluster nodes individually – and that is bound to have issues because we have autoscaling enabled and the nodes may scale up and scale down as per the scaling requirements. New nodes may have different IP addresses to the previous nodes, so unless the Compute engine is continuously supplied with the IP addresses of the nodes, it can't reach them. Moreover, we have multiple replicas and it is possible we might have multiple replicas of the pod on the same node in which case they all can't open the same node port – once a node port is opened by one replica (pod), it can't be used by other replicas on the same node. So this option can be ruled out without going into the rest of the answer.

1. In GKE, create a Service of type LoadBalancer that uses the application's Pods as backend.

2. Add a Cloud Armor Security Policy to the load balancer that whitelists the internal IPs of the MIG's instances.  
3. Configure the Compute Engine instance to use the address of the load balancer that has been created. is not right.  
Creating a service of type LoadBalancer and getting a Compute Engine instance to use the address of the load balancer is fine, but Cloud Armor is not required. You could certainly use Cloud Armor to set up a whitelist policy to only let traffic through from the compute engine instance, but hang on – this option says “MIG instances”. We don't have a managed instance group. The question mentions a single instance but not MIG. If we were to assume the single instance is part of a MIG, i.e. a MIG with a single instance, this option works too. It is more secure than the first option discussed in the explanation but at the same time more expensive. Let's look at the other option to see if it provides a secure yet cost-effective way of achieving the same.

1. In GKE, create a Service of type LoadBalancer that uses the application's Pods as backend.  
2. Add an annotation to this service cloud.google.com/load-balancer-type: Internal  
3. Peer the two VPCs together  
4. Configure the Compute Engine instance to use the address of the load balancer that has been created. is the right answer.  
Creating a service of type LoadBalancer and getting a Compute Engine instance to use the address of the load balancer is fine. We covered this previously in the first option in the explanations section.  
Adding the annotation cloud.google.com/load-balancer-type: Internal makes your cluster's services accessible to applications outside of your cluster that use the same VPC network and are located in the same Google Cloud region. So this improves security by not allowing public access, however, the compute engine is located in a different VPC so it can't access.  
Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/internal-load-balancing>  
But peering the VPCs together enables the compute engine to access the load balancer IP. And peering is possible because they do not use overlapping IP ranges. Peering essentially links up the two VPCs and resources inside the VPCs can communicate with each other as if they were all in a single VPC. More info about VPC peering: <https://cloud.google.com/vpc/docs/vpc-peering>  
So this option is the right answer. It provides a secure and cost-effective way of achieving our requirements. There are several valid answers but this option is more correct than the others.

8. Question

You have an application that looks for its licensing server on the IP 10.0.3.21. You need to deploy the licensing server on the Compute Engine. You do not want to change the configuration of the application and want the application to be able to reach the licensing server. What should you do?

- Use the IP 10.0.3.21 as a custom ephemeral IP address and assign it to the licensing server.
- Start the licensing server with an automatic ephemeral IP address, and then promote it to a static internal IP address.

- Reserve the IP 10.0.3.21 as a static public IP address using gcloud and assign it to the licensing server.
- Reserve the IP 10.0.3.21 as a static internal IP address using gcloud and assign it to the licensing server.

### Unattempted

Reserve the IP 10.0.3.21 as a static public IP address using gcloud and assign it to the licensing server. is not right.

?The private network range is defined by IETF  
(Ref: <https://tools.ietf.org/html/rfc1918>) and includes 10.0.0.0/8. So all IP Addresses from 10.0.0.0 to 10.255.255.255 belong to this internal IP range. As the IP of interest 10.0.3.21 falls within this range, it can not be reserved as a public IP Address. Use the IP 10.0.3.21 as a custom ephemeral IP address and assign it to the licensing server. is not right.

?Ephemeral IP address is the public IP Address assigned to compute instance. An ephemeral external IP address is an IP address that doesn't persist beyond the life of the resource. When you create an instance or forwarding rule without specifying an IP address, the resource is automatically assigned an ephemeral external IP address.

?Ref: <https://cloud.google.com/compute/docs/ip-addresses#ephemeraladdress>

?The private network range is defined by IETF  
(Ref: <https://tools.ietf.org/html/rfc1918>) and includes 10.0.0.0/8. So all IP Addresses from 10.0.0.0 to 10.255.255.255 belong to this internal IP range. As the IP of interest 10.0.3.21 falls within this range, it can not be used as a public IP Address (ephemeral IP is public). Start the licensing server with an automatic ephemeral IP address, and then promote it to a static internal IP address. is not right.

?When a compute instance is started with public IP, it gets an ephemeral IP address. An ephemeral external IP address is an IP address that doesn't persist beyond the life of the resource.

?Ref: <https://cloud.google.com/compute/docs/ip-addresses#ephemeraladdress>

?You can promote this ephemeral address into a Static IP address but this will be an external IP address and not an internal one.

?Ref: [https://cloud.google.com/compute/docs/ip-addresses/reserve-static-external-ip-address#promote\\_ephemeral\\_ip](https://cloud.google.com/compute/docs/ip-addresses/reserve-static-external-ip-address#promote_ephemeral_ip) Reserve the IP 10.0.3.21 as a static internal IP address using gcloud and assign it to the licensing server. is right.

?This is the only option that lets us reserve IP 10.0.3.21 as a static internal IP address because it falls within the standard IP Address range as defined by IETF  
(Ref: <https://tools.ietf.org/html/rfc1918>). This includes the range 10.0.0.0/8 so all IP Addresses from 10.0.0.0 to 10.255.255.255 belong to this internal IP range. Since we can now reserve this IP Address as a static internal IP address, it can be assigned to the licensing server in the VPC so that the application is able to reach the licensing server.

9. 9. Question

You have an application that receives SSL-encrypted TCP traffic on port 443. Clients for this application are located all over the world. You want to minimize latency for the clients. Which load balancing option should you use?

- HTTPS Load Balancer
- Network Load Balancer
- SSL Proxy Load Balancer
- Internal TCP/UDP Load Balancer. Add a firewall rule allowing ingress traffic from 0.0.0.0/0 on the target instances.

**Unattempted**

SSL Proxy Load Balancer. is not right.

?Google says “ SSL Proxy Load Balancing is intended for non-HTTP(S) traffic. For HTTP(S) traffic, we recommend that you use HTTP(S) Load Balancing.”

?So this option can be ruled out.

?Ref: <https://cloud.google.com/load-balancing/docs/ssl> Internal TCP/UDP Load Balancer. Add a firewall rule allowing ingress traffic from 0.0.0.0/0 on the target instances. is not right.

?Internal TCP Load Balancing is a regional load balancer that enables you to run and scale your services behind an internal load balancing IP address that is accessible only to your internal virtual machine (VM) instances. Since we need to serve public traffic, this load balancer is not suitable for us.

?Ref: <https://cloud.google.com/load-balancing/docs/internal> HTTPS Load Balancer. is not right.

?The HTTPS load balancer terminates TLS in locations that are distributed globally, so as to minimize latency between clients and the load balancer. If you require geographic control over where TLS is terminated (which is our scenario with clients located all over the world), you should use Google Cloud Network Load Balancing instead, and terminate TLS on backends that are located in regions appropriate to your needs.

?Ref: <https://cloud.google.com/load-balancing/docs/https#control-tls-termination> Network Load Balancer. is the right answer.

?Google Cloud external TCP/UDP Network Load Balancing (after this referred to as Network Load Balancing) is a regional, non-proxied load balancer. The network load balancer supports any and all ports. You can use Network Load Balancing to load balance TCP and UDP traffic. Because the load balancer is a pass-through load balancer, your backends terminate the load-balanced TCP connection or UDP packets themselves. For example, you might run an HTTPS web server on your backends (which is our scenario) and use a Network Load Balancing to route requests to it, terminating TLS on your backends themselves.

?Ref: <https://cloud.google.com/load-balancing/docs/network>

?Also, the latency is minimized when using network load balancer. Because load balancing takes place in-region and traffic is merely forwarded, there is no significant latency impact compared with the no-load-balancer option.

?Ref: [https://cloud.google.com/load-balancing/docs/tutorials/optimize-app-latency#network\\_load\\_balancing](https://cloud.google.com/load-balancing/docs/tutorials/optimize-app-latency#network_load_balancing)

#### 10. Question

You have an application that uses Cloud Spanner as a backend database. The application has a very predictable traffic pattern. You want to automatically scale up or down the number of Spanner nodes depending on traffic. What should you do?

- Create a cron job that runs on a scheduled basis to review stackdriver monitoring metrics, and then resize the Spanner instance accordingly.
- Create a Stackdriver alerting policy to send an alert to oncall SRE emails when Cloud Spanner CPU exceeds the threshold. SREs would scale resources up or down accordingly.
- Create a Stackdriver alerting policy to send an alert to Google Cloud Support email when Cloud Spanner CPU exceeds your threshold. Google support would scale resources up or down accordingly.
- Create a Stackdriver alerting policy to send an alert to webhook when Cloud Spanner CPU is over or under your threshold. Create a Cloud Function that listens to HTTP and resizes Spanner resources accordingly.

#### Unattempted

Create a cron job that runs on a scheduled basis to review stackdriver monitoring metrics, and then resize the Spanner instance accordingly. is not right.

?While this works and does it automatically , it does not follow Google' s recommended practices.

?Ref: <https://cloud.google.com/spanner/docs/instances> “ Note: You can scale the number of nodes in your instance based on the Cloud Monitoring metrics on CPU or storage utilization in conjunction with Cloud Functions.” Create a Stackdriver alerting policy to send an alert to oncall SRE emails when Cloud Spanner CPU exceeds the threshold. SREs would scale resources up or down accordingly. is not right.

?This does not follow Google' s recommended practices.

?Ref: <https://cloud.google.com/spanner/docs/instances> “ Note: You can scale the number of nodes in your instance based on the Cloud Monitoring metrics on CPU or storage utilization in conjunction with Cloud Functions.” Create a Stackdriver alerting policy to send an alert to Google Cloud Support email when Cloud Spanner CPU exceeds your threshold. Google support would scale resources up or down accordingly. is not right.

?This does not follow Google's recommended practices.

?Ref: <https://cloud.google.com/spanner/docs/instances> “ Note: You can scale the number of nodes in your instance based on the Cloud Monitoring metrics on CPU or storage utilization in conjunction with Cloud Functions.” Create a Stackdriver alerting policy to send an alert to webhook when Cloud Spanner CPU is over or under your threshold. Create a Cloud Function that listens to HTTP and resizes Spanner resources accordingly. is the right answer.

?For scaling the number of nodes in Cloud spanner instance, Google recommends implementing this base on the Cloud Monitoring metrics on CPU or storage utilization in conjunction with Cloud Functions.

?Ref: <https://cloud.google.com/spanner/docs/instances>

#### 11. Question

You have an application that uses Cloud Spanner as a database backend to keep current state information about users. Cloud Bigtable logs all events triggered by users. You export Cloud Spanner data to Cloud Storage during daily backups. One of your analysts asks you to join data from Cloud Spanner and Cloud Bigtable for specific users. You want to complete this ad hoc request as efficiently as possible. What should you do?

- Create a dataflow job that copies data from Cloud Bigtable and Cloud Storage for specific users.
- Create a Cloud Dataproc cluster that runs a Spark job to extract data from Cloud Bigtable and Cloud Storage for specific users.
- Create two separate BigQuery external tables on Cloud Storage and Cloud Bigtable. Use the BigQuery console to join these tables through user fields, and apply appropriate filters.**
- Create a dataflow job that copies data from Cloud Bigtable and Cloud Spanner for specific users.

#### Unattempted

Our requirements are to join user sessions with user events efficiently. We need to look for an option that is primarily a Google service and provides this feature out of the box or with minimal configuration.

?Create two separate BigQuery external tables on Cloud Storage and Cloud Bigtable. Use the BigQuery console to join these tables through user fields, and apply appropriate filters is the right answer.

?Big query lets you create tables that reference external data sources such as Bigtable and Cloud Storage. You can then join up these two tables through user fields and apply appropriate filters. You can achieve the end result with minimal configuration using this option.

?Ref: <https://cloud.google.com/bigquery/external-data-sources> Create a dataflow job that copies data from Cloud Bigtable and Cloud Spanner for specific users. is not right.

?Cloud dataflow does not support Cloud Spanner. Cloud Dataflow SQL supports reading from Pub/Sub topics, Cloud Storage file sets, and BigQuery tables.

Create a Cloud Dataproc cluster that runs a Spark job to extract data from Cloud Bigtable and Cloud Storage for specific users. is not right.

?While it is certainly possible to do this using a Spark job, it is complicated as we would have to come up with the code/logic to extract the data and certainly not straightforward. Create a dataflow job that copies data from Cloud Bigtable and Cloud Storage for specific users. is not right.

?This is possible but it is not as efficient as using Big Query.

?Ref: <https://cloud.google.com/dataflow/docs/guides/sql/dataflow-sql-intro> Here is some more documentation around this option, some of the issues are

?1. Dataflow SQL expects CSV files in Cloud Storage filesets. CSV files must not contain a header row with column names; the first row in each CSV file is interpreted as a data record. – but our question doesn't say how the exported data is stored in cloud storage.

?2. You can only run jobs in regions that have a Dataflow regional endpoint. Our question doesn't say which region.

Ref: <https://cloud.google.com/dataflow/docs/concepts/regional-endpoints>.

?3. Creating a Dataflow job can take several minutes – unlike Big Query external tables which can be created very easily.

?Too many unknowns. Otherwise, this option is a good option.

?Here is some more information if you'd like to get a better understanding of how to use Cloud Dataflow to achieve this result.

?Cloud Dataflow SQL lets you use SQL queries to develop and run Dataflow jobs from the BigQuery web UI. You can join streams (such as Pub/Sub) and snapshotted datasets (such as BigQuery tables and Cloud Storage filesets); query your streams or static datasets with SQL by associating schemas with objects, such as tables, Cloud Storage filesets and Pub/Sub topics; and write your results into a BigQuery table for analysis and dashboarding.

?Cloud Dataflow SQL supports multiple data sources including Cloud Storage and Big Query tables which are of interest for this scenario.

?<https://cloud.google.com/dataflow/docs/guides/sql/data-sources-destinations>

## 12. Question

You have an instance group that you want to load balance. You want the load balancer to terminate the client SSL session. The instance group is used to serve

a public web application over HTTPS. You want to follow Google recommended practices. What should you do?

- Configure an external TCP proxy load balancer.
- Configure an external SSL proxy load balancer.
- Configure an internal TCP load balancer.
- Configure an HTTP(S) load balancer.

**Unattempted**

Configure an internal TCP load balancer. is not right.

?Internal TCP Load Balancing is a regional load balancer that enables you to run and scale your services behind an internal load balancing IP address that is accessible only to your internal virtual machine (VM) instances. Since we need to serve public traffic, this load balancer is not suitable for us.

?Ref: <https://cloud.google.com/load-balancing/docs/internal> Configure an external SSL proxy load balancer. is not right.

?Google says “ SSL Proxy Load Balancing is intended for non-HTTP(S) traffic. For HTTP(S) traffic, we recommend that you use HTTP(S) Load Balancing.”

?So this option can be ruled out.

?Ref: <https://cloud.google.com/load-balancing/docs/ssl> Configure an external TCP proxy load balancer. is not right.

?Google says “ TCP Proxy Load Balancing is intended for non-HTTP traffic. For HTTP traffic, use HTTP Load Balancing instead. For proxied SSL traffic, use SSL Proxy Load Balancing.” So this option can be ruled out.

?Ref: <https://cloud.google.com/load-balancing/docs/tcp> Configure an HTTP(S) load balancer. is the right answer.

?This is the only option that fits all requirements. It can serve public traffic, can terminate SSL at the load balancer and follows google recommended practices.

?? “ The backends of a backend service can be either instance groups or network endpoint groups (NEGs), but not a combination of both.”

?? “ An external HTTP(S) load balancer distributes traffic from the internet”

?? “ The client SSL session terminates at the load balancer.”

?? “ For HTTP traffic, use HTTP Load Balancing instead.”

?Ref: <https://cloud.google.com/load-balancing/docs/https>

### 13. 13. Question

You have an object in a Cloud Storage bucket that you want to share with an external company. The object contains sensitive data. You want access to the content to be removed after four hours. The external company does not have a Google account to which you can grant specific user-based access privileges. You want to use the most secure method that requires the fewest steps. What should you do?

- Configure the storage bucket as a static website and furnish the object's URL to the company. Delete the object from the storage bucket after four hours.
- Create a new Cloud Storage bucket specifically for the external company to access. Copy the object to that bucket. Delete the bucket after four hours have passed.
- Create a signed URL with a four-hour expiration and share the URL with the company.**
- Set object access to 'public' and use object lifecycle management to remove the object after four hours.

#### Unattempted

Set object access to ' public' and use object lifecycle management to remove the object after four hours. is not right.

?While the external company can access the public objects from the bucket, it doesn't stop bad actors from accessing the data as well. Since the data is " sensitive" and we want to follow a " secure method" , we shouldn't do this. Configure the storage bucket as a static website and furnish the object' s URL to the company. Delete the object from the storage bucket after four hours. is not right.

?The static website is public by default. While the external company can access the objects from the static website, it doesn't stop bad actors from accessing the data as well. Since the data is " sensitive" and we want to follow a " secure method" , we shouldn't do this. Create a new Cloud Storage bucket specifically for the external company to access. Copy the object to that bucket. Delete the bucket after four hours have passed. is not right.

?Even if we were to create a separate bucket for the external company to access, since the company does not have a google account, the only way to have them access this separate bucket is by enabling public access which we can't because of the nature of data (sensitive) and is against standard security practices. Create a signed URL with a four-hour expiration and share the URL with the company. is the right answer.

?This is the only option that fits all requirements. When we generate a signed URL, we can specify an expiry and only users with the signed URL can view/download the objects, and they don't need a google account.

?Ref: <https://cloud.google.com/storage/docs/access-control/signed-urls>

?This page provides an overview of signed URLs, which you use to give time-limited resource access to anyone in possession of the URL, regardless of whether they have a Google account.

#### 14. 14. Question

You have annual audits every year and you need to provide external auditors access to the last 10 years of audit logs. You want to minimize the cost and operational overhead while following Google recommended practices. What should you do? (Select Three)

- Grant external auditors Storage Object Viewer role on the logs storage bucket.**
- Set a custom retention of 10 years in Stackdriver logging and provide external auditors view access to Stackdriver Logs.
- Export audit logs to Cloud Storage via an audit log export sink.**
- Export audit logs to BigQuery via an audit log export sink.
- Export audit logs to Cloud Filestore via a Pub/Sub export sink.
- Configure a lifecycle management policy on the logs bucket to delete objects older than 10 years**

**Unattempted**

Export audit logs to Cloud Filestore via a Pub/Sub export sink. is not right.  
Storing logs in Cloud Filestore is expensive. In Cloud Filestore, Standard Tier pricing costs \$0.2 per GB per month and Premium Tier pricing costs \$0.3 per GB per month. In comparison, Google Cloud Storage offers several storage classes that are significantly cheaper.

Ref: <https://cloud.google.com/bigquery/pricing>

Ref: <https://cloud.google.com/storage/pricing>

Set a custom retention of 10 years in Stackdriver logging and provide external auditors view access to Stackdriver Logs. is not right.

While it is possible to configure a custom retention period of 10 years in Stackdriver logging, storing logs in Stackdriver is expensive compared to Cloud Storage. Stackdriver charges \$0.01 per GB per month, whereas something like Cloud Storage Coldline Storage costs \$0.007 per GB per month (30% cheaper) and Cloud Storage Archive Storage costs 0.004 per GB per month (60% cheaper than Stackdriver)

Ref: <https://cloud.google.com/logging/docs/storage#pricing>

Ref: <https://cloud.google.com/storage/pricing>

Export audit logs to BigQuery via an audit log export sink. is not right.

Storing logs in BigQuery is expensive. In BigQuery, Active storage costs \$0.02 per GB per month and Long-term storage costs \$0.01 per GB per month. In comparison, Google Cloud Storage offers several storage classes that are significantly cheaper.

Ref: <https://cloud.google.com/bigquery/pricing>

Ref: <https://cloud.google.com/storage/pricing>

Export audit logs to Cloud Storage via an audit log export sink. is the right answer.

Among all the storage solutions offered by Google Cloud Platform, Cloud storage offers the best pricing for long term storage of logs. Google Cloud Storage offers several storage classes such as Nearline Storage (\$0.01 per GB per Month) Coldline Storage (\$0.007 per GB per Month) and Archive Storage (\$0.004 per GB per month) which are significantly cheaper than the storage options covered by the above options above.

Ref: <https://cloud.google.com/storage/pricing>

Grant external auditors Storage Object Viewer role on the logs storage bucket. is the right answer.

You can provide external auditors access to the logs in the bucket by granting the Storage Object Viewer role which allows them to read any object stored in any bucket.

Ref: <https://cloud.google.com/storage/docs/access-control/iam>

Configure a lifecycle management policy on the logs bucket to delete objects older than 10 years. is the right answer.

You need to archive log files for 10 years but you don't need log files older than 10 years. And since you also want to minimize costs, it is a good idea to set up a lifecycle management policy on the bucket to delete objects that are older than 10 years. Lifecycle management configuration is a set of rules which apply to current and future objects in the bucket. When an object meets the criteria of one of the rules, Cloud Storage automatically performs a specified action (delete in this case) on the object.

Ref: <https://cloud.google.com/storage/docs/lifecycle>

## 15. Question

You have asked your supplier to send you a purchase order and you want to enable them to upload the file to a cloud storage bucket within the next 4 hours. Your supplier does not have a Google account. You want to follow Google recommended practices. What should you do?

- Create a JSON key for the Default Compute Engine Service Account. Execute the command `gsutil signurl -m PUT -d 4h gs:///**`.
- Create a service account with just the permissions to upload files to the bucket. Create a JSON key for the service account. Execute the command `gsutil signurl -httpMethod PUT -d 4h gs:///**`.
- Create a service account with just the permissions to upload files to the bucket. Create a JSON key for the service account. Execute the command `gsutil signurl -m PUT -d 4h gs:///po.pdf`.
- Create a service account with just the permissions to upload files to the bucket. Create a JSON key for the service account. Execute the command `gsutil signurl -d 4h gs:///`.

## Unattempted

Create a service account with just the permissions to upload files to the bucket. Create a JSON key for the service account. Execute the command gsutil signurl -d 4h gs:/// is not right.

?This command creates signed URLs for retrieving existing objects. This command does not specify a HTTP method and in the absence of one, the default HTTP method is GET.

?Ref: <https://cloud.google.com/storage/docs/gsutil/commands/signurl> Create a service account with just the permissions to upload files to the bucket. Create a JSON key for the service account. Execute the command gsutil signurl -httpMethod PUT -d 4h gs:///\*\*. is not right.

?gsutil signurl does not accept -httpMethod parameter.

```
?$ gsutil signurl -d 4h -httpMethod PUT keys.json gs://gcp-ace-lab-255520/*
```

?CommandException: Incorrect option(s) specified. Usage:

?The HTTP method can be provided through -m flag.

?Ref: <https://cloud.google.com/storage/docs/gsutil/commands/signurl> Create a JSON key for the Default Compute Engine Service Account. Execute the command gsutil signurl -m PUT -d 4h gs:///\*\*. is not right.

?Using the default compute engine service account violates the principle of least privilege. The recommended approach is to create a service account with just the right permissions needed and create JSON keys for this service account to use with gsutil signurl command. Create a service account with just the permissions to upload files to the bucket. Create a JSON key for the service account. Execute the command gsutil signurl -m PUT -d 4h gs:///po.pdf. is the right answer.

?This command correctly creates a signed url that is valid for 4 hours and allows PUT (through the -m flag) operations on the file po.pdf in the bucket. The supplier can then use the signed URL to upload a file to this bucket within 4 hours.

?Ref: <https://cloud.google.com/storage/docs/gsutil/commands/signurl>

## 16. Question

You have been asked to create a new Kubernetes Cluster on Google Kubernetes Engine that can autoscale the number of worker nodes as well as pods. What should you do? (Select 2)

- 

Create a GKE cluster and enable autoscaling on Kubernetes Engine.

- Create Compute Engine instances for the workers and the master and install Kubernetes. Rely on Kubernetes to create additional Compute Engine instances when needed.
- Create a GKE cluster and enable autoscaling on the instance group of the cluster.
- Configure a Compute Engine instance as a worker and add it to an unmanaged instance group. Add a load balancer to the instance group and rely on the load balancer to create additional Compute Engine instances when needed.
- **Enable Horizontal Pod Autoscaling for the Kubernetes deployment.**

### Unattempted

Create a GKE cluster and enable autoscaling on the instance group of the cluster. is not right.

?GKE' s cluster auto-scaler automatically resizes the number of nodes in a given node pool, based on the demands of your workloads. However, we should not enable Compute Engine autoscaling for managed instance groups for the cluster nodes. GKE' s cluster auto-scaler is separate from Compute Engine autoscaling.

?Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/cluster-autoscaler> Configure a Compute Engine instance as a worker and add it to an unmanaged instance group. Add a load balancer to the instance group and rely on the load balancer to create additional Compute Engine instances when needed. is not right.

?When using GKE to manage your Kubernetes clusters, you can not add manually created compute instances to the worker node pool. A node pool is a group of nodes within a cluster that all have the same configuration. Node pools use a NodeConfig specification.

?Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/node-pools> ?Moreover, Unmanaged instance groups do not autoscale. An unmanaged instance group is simply a collection of virtual machines (VMs) that reside in a single zone, VPC network, and subnet. An unmanaged instance group is useful for grouping together VMs that require individual configuration settings or tuning.

?Ref: <https://cloud.google.com/compute/docs/instance-groups/creating-groups-of-unmanaged-instances> Create Compute Engine instances for the workers and the master and install Kubernetes. Rely on Kubernetes to create additional Compute Engine instances when needed. is not right.

?When using Google Kubernetes Engine, you can not install master node separately. The cluster master runs the Kubernetes control plane processes, including the Kubernetes API server, scheduler, and core resource controllers. The master' s lifecycle is managed by GKE when you create or delete a cluster.

?Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/cluster-architecture>

?Also, you can not add manually created compute instances to the worker node pool. A node pool is a group of nodes within a cluster that all have the same configuration. Node pools use a NodeConfig specification.

?Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/node-pools> Create a GKE cluster and enable autoscaling on Kubernetes Engine. is the right answer.

?GKE' s cluster autoscaler automatically resizes the number of nodes in a given node pool, based on the demands of your workloads. You don' t need to manually add or remove nodes or over-provision your node pools. Instead, you specify a minimum and maximum size for the node pool, and the rest is automatic. When demand is high, cluster autoscaler adds nodes to the node pool. When demand is low, cluster autoscaler scales back down to a minimum size that you designate. This can increase the availability of your workloads when you need it while controlling costs.

?Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/cluster-autoscaler> Enable Horizontal Pod Autoscaling for the kubernetes deployment. is the right answer.

?Horizontal Pod Autoscaler scales up and scales down your Kubernetes workload by automatically increasing or decreasing the number of Pods in response to the workload' s CPU or memory consumption, or in response to custom metrics reported from within Kubernetes or external metrics from sources outside of your cluster. Horizontal Pod Autoscaling cannot be used for workloads that cannot be scaled, such as DaemonSets.

?Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/horizontalpodautoscaler>

## 17. Question

You have two workloads on GKE (Google Kubernetes Engine) – create-order and dispatch-order. create-order handles the creation of customer orders, and dispatch-order handles dispatching orders to your shipping partner. Both create-order and dispatch-order workloads have cluster autoscaling enabled. The create-order deployment needs to access (i.e. invoke web service of) dispatch-order deployment. dispatch-order deployment cannot be exposed publicly. How should you define the services?

- Create a Service of type NodePort for dispatch-order and an Ingress Resource for that Service. Have create-order use the Ingress IP address.
- Create a Service of type LoadBalancer for dispatch-order and an Ingress Resource for that Service. Have create-order use the Ingress IP address.
- Create a Service of type LoadBalancer for dispatch-order. Have create-order use the Service IP address.



- Create a Service of type ClusterIP for dispatch-order. Have create-order use the Service IP address.

Unattempted

Create a Service of type LoadBalancer for dispatch-order. Have create-order use the Service IP address. is not right.

?When you create a Service of type LoadBalancer, the Google Cloud controller configures a network load balancer that is publicly available. Since we don't want our service to be publicly available, we shouldn't create a Service of type LoadBalancer

?Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/exposing-apps> Create a Service of type LoadBalancer for dispatch-order and an Ingress Resource for that Service. Have create-order use the Ingress IP address. is not right.

?When you create a Service of type LoadBalancer, the Google Cloud controller configures a network load balancer that is publicly available. Since we don't want our service to be publicly available, we shouldn't create a Service of type LoadBalancer

?Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/exposing-apps> Create a Service of type NodePort for dispatch-order and an Ingress Resource for that Service. Have create-order use the Ingress IP address. is not right.

?Exposes the Service on each Node's IP at a static port (the NodePort). If the compute instance has public connectivity, the dispatch-order can be accessed publicly which is undesirable. Secondly, dispatch-order has auto-scaling enabled so we shouldn't create a service of NodePort. If autoscaler spins up another pod on the node, it fails to initialize as the port on the node is already taken by an existing pod on the same node.

?Ref: <https://cloud.google.com/kubernetes-engine/docs/how-to/exposing-apps> Create a Service of type ClusterIP for dispatch-order. Have create-order use the Service IP address. is the right answer.

?ClusterIP exposes the Service on a cluster-internal IP that is only reachable within the cluster. This satisfies our requirement that dispatch-order shouldn't be publicly accessible. create-order which is also located in the same GKE cluster can now access the ClusterIP of the service to reach dispatch-order.

?Ref: <https://kubernetes.io/docs/concepts/services-networking/service/>

## 18. Question

You host a production application in Google Compute Engine in the us-central1-a zone. Your application needs to be available 24\*7 all through the year. The application suffered an outage recently due to a Compute Engine outage in the zone hosting your application. Your application is also susceptible to slowness during peak usage. You have been asked for a recommendation on how to modify the infrastructure to implement a cost-effective and scalable solution that can withstand zone failures. What would you recommend?

- Use Managed instance groups with instances in a single zone. Enable Autoscaling on the Managed instance group.
- Use Managed instance groups with preemptible instances across multiple zones. Enable Autoscaling on the Managed instance group.
- Use Managed instance groups across multiple zones. Enable Autoscaling on the Managed instance group.**
- Use Unmanaged instance groups across multiple zones. Enable Autoscaling on the Unmanaged instance group.

### Unattempted

Use Managed instance groups with preemptible instances across multiple zones. Enable Autoscaling on the Managed instance group. is not right.

A preemptible VM runs at a much lower price than normal instances and is cost-effective. However, Compute Engine might terminate (preempt) these instances if it requires access to those resources for other tasks. Preemptible instances are not suitable for production applications that need to be available 24\*7.

Ref: <https://cloud.google.com/compute/docs/instances/preemptible>

Use Unmanaged instance groups across multiple zones. Enable Autoscaling on the Unmanaged instance group. is not right.

Unmanaged instance groups do not autoscale. An unmanaged instance group is simply a collection of virtual machines (VMs) that reside in a single zone, VPC network, and subnet. An unmanaged instance group is useful for grouping together VMs that require individual configuration settings or tuning.

Ref: <https://cloud.google.com/compute/docs/instance-groups/creating-groups-of-unmanaged-instances>

Use Managed instance groups with instances in a single zone. Enable Autoscaling on the Managed instance group. is not right.

While enabling auto-scaling is a good idea, autoscaling would spin up instances in the same zone. Should there be a zone failure, all instances of the managed instance group would be unreachable and cause the application to be unreachable. Google recommends you distribute your resources across multiple zones to tolerate outages.

Ref: <https://cloud.google.com/compute/docs/regions-zones>

Use Managed instance groups across multiple zones. Enable Autoscaling on the Managed instance group. is the right answer.

Distribute your resources across multiple zones and regions to tolerate outages. Google designs zones to be independent of each other: a zone usually has power, cooling, networking, and control planes that are isolated from other zones, and most single failure events will affect only a single zone. Thus, if a zone becomes unavailable, you can transfer traffic to another zone in the same region to keep your services running.

Ref: <https://cloud.google.com/compute/docs/regions-zones>

In addition, a managed instance group (MIG) contains offers auto-scaling capabilities that let you automatically add or delete instances from a managed instance group based on increases or decreases in load. Autoscaling helps your apps gracefully handle increases in traffic and reduce costs when the need for

resources is lower. Autoscaling works by adding more instances to your instance group when there is more load (upscaling), and deleting instances when the need for instances is lowered (downscaling).

Ref: <https://cloud.google.com/compute/docs/autoscaler/>

#### 19. 19. Question

You host a static website in Cloud Storage. Recently you began to include links to PDF files on this site. Currently, when users click on links to these PDF files, their browser prompts them to save the file to their machine locally. However, you want the clicked PDF files to be displayed within the browser window directly without prompting the user to save the files locally. What should you do?

- Set Content-Type metadata to application/pdf on the PDF file objects
- Enable Cloud CDN on the website frontend.
- Add a label to the storage bucket with a key of Content-Type and a value of application/pdf.
- Enable Share publicly on the PDF file objects

**Unattempted**

Set Content-Type metadata to application/pdf on the PDF file objects is the right answer.

?Content-Type allows browsers to render the object properly. If the browser prompts users to save files to their machine, it is likely the browser does not see the Content-Type as application/pdf. Setting this would ensure the browser displays PDF files within the browser instead of popping up a download dialog.

?Ref: [https://cloud.google.com/storage/docs/gsutil/addlhelp/WorkingWithObjectMetadata#content-type\\_1](https://cloud.google.com/storage/docs/gsutil/addlhelp/WorkingWithObjectMetadata#content-type_1) Enable Cloud CDN on the website frontend. is not right.

?CDN helps with caching content at the edge but doesn't help the browser in displaying pdf files. Enable Share publicly on the PDF file objects. is not right.

?The fact that the browser lets users download the file suggests the browser is able to reach out and download the file. Sharing publicly wouldn't make any difference. Add a label to the storage bucket with a key of Content-Type and a value of application/pdf. is not right.

?Bucket labels are key: value metadata pairs that allow you to group your buckets along with other Google Cloud resources such as virtual machine instances and persistent disks. They don't determine the file's content type.

#### 20. 20. Question

You installed Stackdriver Logging agent on all compute instances. You now need to forward logs from all Compute Engine instances to a BigQuery dataset called pt-logs. You want to minimize cost. What should you do?

- 1. Give the BigQuery Data Editor role on the pt-logs dataset to the service accounts used by your instances. 2. Update your instances' metadata to add the following value: logs-destination: bq://pt-logs.
- 1. In Stackdriver Logging, create a logs export with a Cloud Pub/Sub topic called logs as a sink. 2. Create a Cloud Function that is triggered by messages in the logs topic. 3. Configure that Cloud Function to drop logs that are not from Compute Engine and to insert Compute Engine logs in the pt-logs dataset.
- **1. In Stackdriver Logging, create a filter to view only Compute Engine logs. 2. Click Create Export. 3. Choose BigQuery as Sink Service, and the pt-logs dataset as Sink Destination.**
- 1. Create a Cloud Function that has the BigQuery User role on the pt-logs dataset. 2. Configure this Cloud Function to create a BigQuery Job that executes this query: `INSERT INTO dataset.pt-logs (timestamp, log) SELECT timestamp, log FROM compute.logs WHERE timestamp > DATE_SUB(CURRENT_DATE(), INTERVAL 1 DAY)` 3. Use Cloud Scheduler to trigger this Cloud Function once a day.

### Unattempted

1. Give the BigQuery Data Editor role on the pt-logs dataset to the service accounts used by your instances.
2. Update your instances' metadata to add the following value: logs-destination: bq://pt-logs. is not right.

Among other things, roles/bigquery.dataEditor lets you Create, update, get, and delete the dataset's tables. However, setting a metadata tag logs-destination to bq://pt-logs has no effect on how the logs are generated or forwarded. The stack driver agent is already installed so the logs are forwarded to stack driver logging and not to the BigQuery dataset. Metadata entries are key-value pairs and do not influence this behavior.

Ref: <https://cloud.google.com/compute/docs/storing-retrieving-metadata>

1. In Stackdriver Logging, create a logs export with a Cloud Pub/Sub topic called logs as a sink.
2. Create a Cloud Function that is triggered by messages in the logs topic.
3. Configure that Cloud Function to drop logs that are not from Compute Engine and to insert Compute Engine logs in the pt-logs dataset. is not right.

While the end result meets our requirement, this option involves more steps, it is inefficient and expensive. Triggering a cloud function for each log message and then dropping messages that are not relevant (i.e. not compute engine logs) is inefficient. We are paying for cloud function execution for all log entries when we are only interested in compute engine logs. Secondly, triggering a cloud function and then have that insert into the BigQuery dataset is also inefficient and expensive when the same can be achieved directly by configuring BigQuery as the sink destination – we don't pay for cloud function executions. Using this option, we are unnecessarily paying for Cloud Pub/Sub and Cloud Functions.

Ref: [https://cloud.google.com/logging/docs/export/configure\\_export\\_v2](https://cloud.google.com/logging/docs/export/configure_export_v2)  
Ref: <https://cloud.google.com/logging/docs/view/advanced-queries>

1. Create a Cloud Function that has the BigQuery User role on the pt-logs dataset.

2. Configure this Cloud Function to create a BigQuery Job that executes this query:

```
INSERT INTO dataset.pt-logs (timestamp, log)
SELECT timestamp, log FROM compute.logs
WHERE timestamp > DATE_SUB(CURRENT_DATE(), INTERVAL 1 DAY)
```

3. Use Cloud Scheduler to trigger this Cloud Function once a day. is not right.  
The role roles/bigquery.user provides permissions to run jobs, including queries, within the project. A cloud function with this role can execute queries in BigQuery, however, the logs are not available in BigQuery in compute.logs so you can not query compute engine logs by running SELECT timestamp, log FROM compute.logs.

Ref: <https://cloud.google.com/bigquery/docs/access-control>

1. In Stack driver Logging, create a filter to view only Compute Engine logs.

2. Click Create Export.

3. Choose BigQuery as Sink Service, and the pt-logs dataset as Sink Destination. is the right answer.

In stack driver logging, it is possible to create a filter to just query the compute engine logs which is what we are interested in.

Ref: <https://cloud.google.com/logging/docs/view/advanced-queries>

You can then export these logs into a sink that has the BigQuery dataset configured as the destination.

[https://cloud.google.com/logging/docs/export/configure\\_export\\_v2](https://cloud.google.com/logging/docs/export/configure_export_v2)

This way, just the logs that we need are exported to BigQuery. This option is the most efficient of all options and uses features provided by GCP out of the box.

## 21. Question

You need a dynamic way of provisioning VMs on the Compute Engine. The exact specifications will be in a dedicated configuration file. You want to follow Google recommended practices. Which method should you use?

- Unmanaged Instance Group
- Deployment Manager
- Managed Instance Group
- Cloud Composer

**Unattempted**

Unmanaged Instance Group. is not right.

?Unmanaged instance groups let you load balance across a fleet of VMs that you manage yourself. But it doesn't help with dynamically provisioning VMs.

?Ref: [https://cloud.google.com/compute/docs/instance-groups#unmanaged\\_instance\\_groups](https://cloud.google.com/compute/docs/instance-groups#unmanaged_instance_groups) Cloud Composer. is not right.  
?Cloud Composer is a fully managed workflow orchestration service that empowers you to author, schedule, and monitor pipelines that span across clouds and on-premises data centers. Cloud Composer is deeply integrated within the Google Cloud Platform, giving users the ability to orchestrate their full pipeline. Cloud Composer has robust, built-in integration with many products, including BigQuery, Cloud Dataflow, Cloud Dataproc, Cloud Datastore, Cloud Storage, Cloud Pub/Sub, and AI Platform.

?Ref: <https://cloud.google.com/composer> Managed Instance Group. is not right.

?Managed instance groups (MIGs) let you operate apps on multiple identical VMs. You can make your workloads scalable and highly available by taking advantage of automated MIG services, including autoscaling, autohealing, regional (multiple zones) deployment, and automatic updating. While MIG dynamically provisions virtual machines based on scaling policy, it doesn't satisfy our requirement of "dedicated configuration file"

?Ref: [https://cloud.google.com/compute/docs/instance-groups#managed\\_instance\\_groups](https://cloud.google.com/compute/docs/instance-groups#managed_instance_groups) Deployment Manager. is the right answer.  
?Google Cloud Deployment Manager allows you to specify all the resources needed for your application in a declarative format using YAML. You can also use Python or Jinja2 templates to parameterize the configuration and allow reuse of common deployment paradigms such as a load-balanced, auto-scaled instance group. You can deploy many resources at one time, in parallel. Using the deployment manager, you can apply a Python/Jinja2 template to create a MIG/auto-scaling policy that dynamically provisions VM. And our other requirement of "dedicated configuration file" is also met. Using the deployment manager for provisioning results in a repeatable deployment process. By creating configuration files that define the resources, the process of creating those resources can be repeated over and over with consistent results. Google recommends we script our infrastructure and deploy using Deployment Manager

?Ref: <https://cloud.google.com/deployment-manager>

## 22. Question

You need to assign a Cloud Identity and Access Management (Cloud IAM) role to an external auditor. The auditor needs to have permissions to review your Google Cloud Platform (GCP) Audit Logs and also to review your Data Access logs. What should you do?

- Assign the auditor the IAM role roles/logging.privateLogViewer. Perform the export of logs to Cloud Storage.
- Assign the auditor the IAM role roles/logging.privateLogViewer. Direct the auditor to also review the logs for changes to Cloud IAM policy.**

- Assign the auditor's IAM user to a custom role that has `logging.privateLogEntries.list` permission. Perform the export of logs to Cloud Storage.
  - Assign the auditor's IAM user to a custom role that has `logging.privateLogEntries.list` permission. Direct the auditor to also review the logs for changes to Cloud IAM policy.

### Unattempted

Google Cloud provides Cloud Audit Logs, which is an integral part of Cloud Logging. It consists of two log streams for each project: Admin Activity and Data Access, which are generated by Google Cloud services to help you answer the question of “ who did what, where, and when?” within your Google Cloud projects.

Ref: [https://cloud.google.com/iam/docs/job-functions/auditing#scenario\\_external\\_auditors](https://cloud.google.com/iam/docs/job-functions/auditing#scenario_external_auditors)

To view Admin Activity audit logs, you must have one of the following Cloud IAM roles in the project that contains your audit logs:

- Project Owner, Project Editor, or Project Viewer.
- The Logging Logs Viewer role.
- A custom Cloud IAM role with the `logging.logEntries.list` Cloud IAM permission.

[https://cloud.google.com/iam/docs/audit-logging#audit\\_log\\_permissions](https://cloud.google.com/iam/docs/audit-logging#audit_log_permissions)

To view Data Access audit logs, you must have one of the following roles in the project that contains your audit logs:

- Project Owner.
- Logging' s Private Logs Viewer role.
- A custom Cloud IAM role with the `logging.privateLogEntries.list` Cloud IAM permission.

[https://cloud.google.com/iam/docs/audit-logging#audit\\_log\\_permissions](https://cloud.google.com/iam/docs/audit-logging#audit_log_permissions)

---

Assign the auditor' s IAM user to a custom role that has `logging.privateLogEntries.list` permission. Perform the export of logs to Cloud Storage. is not right.

`logging.privateLogEntries.list` provides permissions to view Data Access audit logs but this does not grant permissions to view Admin activity logs.

Ref: [https://cloud.google.com/logging/docs/access-control#console\\_permissions](https://cloud.google.com/logging/docs/access-control#console_permissions)

Assign the auditor' s IAM user to a custom role that has

`logging.privateLogEntries.list` permission. Direct the auditor to also review the logs for changes to Cloud IAM policy. is not right.

`logging.privateLogEntries.list` provides permissions to view Data Access audit logs but this does not grant permissions to view Admin activity logs.

Ref: [https://cloud.google.com/logging/docs/access-control#console\\_permissions](https://cloud.google.com/logging/docs/access-control#console_permissions)

Assign the auditor the IAM role `roles/logging.privateLogViewer`. Perform the export of logs to Cloud Storage. is not right.

`roles/logging.privateLogViewer` is the right role and lets the auditor review admin activity and data access logs but exporting logs to Cloud Storage indicates that we want the auditor to review logs from Cloud Storage and not the logs within Cloud Logging console. In this scenario, unless the auditor is assigned a role that

lets them access the relevant cloud storage buckets, they wouldn't be able to view log information in the buckets.

Assign the auditor the IAM role roles/logging.privateLogViewer. Direct the auditor to also review the logs for changes to Cloud IAM policy. is the right answer. roles/logging.privateLogViewer (Private Logs Viewer) includes everything from roles/logging.viewer, plus the ability to read Access Transparency logs and Data Access audit logs. This lets the auditor review the admin activity and data access logs in Cloud Logging console.

Ref: <https://cloud.google.com/logging/docs/access-control>

### 23. Question

You need to configure IAM access audit logging in BigQuery for external auditors. You want to follow Google-recommended practices. What should you do?

- Add the auditors group to two new custom IAM roles.
- Add the auditor user accounts to the 'logging.viewer' and 'bigQuery.dataViewer' predefined IAM roles.
- Add the auditor user accounts to two new custom IAM roles.
- Add the auditors group to the 'logging.viewer' and 'bigQuery.dataViewer' predefined IAM roles.

**Unattempted**

Add the auditor user accounts to the ' logging.viewer' and ' bigQuery.dataViewer' predefined IAM roles. is not right.

?Since auditing happens several times a year, we don't want to repeat the process of granting multiple roles to multiple users every time. Instead, we want to define a group with the required grants (a one time task) and assign this group to the auditor users during the time of the audit. Add the auditor user accounts to two new custom IAM roles. is not right.

?Google already provides roles that fit the external auditing requirements so we don't need to create custom roles. Nothing stops us from creating custom IAM roles to achieve the same purpose, but this doesn't follow " Google-recommended practices" Add the auditors group to two new custom IAM roles. is not right.

?Google already provides roles that fit the external auditing requirements so we don't need to create custom roles. Nothing stops us from creating custom IAM roles to achieve the same purpose, but this doesn't follow " Google-recommended practices" Add the auditors group to the ' logging.viewer' and ' bigQuery.dataViewer' predefined IAM roles. is the right answer.

?For external auditors, Google recommends we grant logging.viewer and bigquery.dataViewer roles. Since auditing happens several times a year to review

the organization's audit logs, it is recommended we create a group with these grants and assign the group to auditor user accounts during the time of the audit.

#### 24. Question

You need to create a custom IAM role for use with a GCP service. All permissions in the role must be suitable for production use. You also want to clearly share with your organization the status of the custom role. This will be the first version of the custom role. What should you do?

- Use permissions in your role that use the SUPPORTED support level for role permissions. Set the role stage to ALPHA while testing the role permissions.
- Use permissions in your role that use the SUPPORTED support level for role permissions. Set the role stage to BETA while testing the role permissions.
- Use permissions in your role that use the TESTING support level for role permissions. Set the role stage to ALPHA while testing the role permissions.
- Use permissions in your role that use the TESTING support level for role permissions. Set the role stage to BETA while testing the role permissions.

#### Unattempted

When setting support levels for permissions in custom roles, you can set to one of SUPPORTED, TESTING or NOT\_SUPPORTED.

?SUPPORTED -The permission is fully supported in custom roles.

?TESTING – The permission is being tested to check its compatibility with custom roles. You can include the permission in custom roles, but you might see unexpected behavior. Not recommended for production use.

?Ref: <https://cloud.google.com/iam/docs/custom-roles-permissions-support>

?Since we want the role to be suitable for production use, we need “ SUPPORTED” and not “ TESTING” . In terms of role stage, the stage transitions from ALPHA → BETA → GA

?Ref: [https://cloud.google.com/iam/docs/understanding-custom-roles#testing\\_and\\_deployment](https://cloud.google.com/iam/docs/understanding-custom-roles#testing_and_deployment)

?Since this is the first version of custom role, we start with “ ALPHA” . The only option that satisfies “ ALPHA” stage with “ SUPPORTED” support level is

?Use permissions in your role that use the SUPPORTED support level for role permissions. Set the role stage to ALPHA while testing the role permissions.

#### 25. Question

You need to create a custom VPC with a single subnet. The subnet's range must be as large as possible. Which range should you use?

- 10.0.0.0/8
- 192.168.0.0/16
- 172.16.0.0/12
- 0.0.0.0/0

**Unattempted**

The private network range is defined by IETF

(Ref: <https://tools.ietf.org/html/rfc1918>) and adhered to by all cloud providers. The supported internal IP Address ranges are 1. 24-bit block 10.0.0.0/8 (16777216 IP Addresses)

?2. 20-bit block 172.16.0.0/12 (1048576 IP Addresses)

?3. 16-bit block 192.168.0.0/16 (65536 IP Addresses) 10.0.0.0/8 gives you the largest range – 16777216 IP Addresses.

## 26. Question

You need to create a new billing account and then link it with an existing Google Cloud Platform project. What should you do?

- Verify that you are Project Billing Manager for the GCP project. Update the existing project to link it to the existing billing account.
- Verify that you are Billing Administrator for the billing account. Update the existing project to link it to the existing billing account.
- Verify that you are Project Billing Manager for the GCP project. Create a new billing account and link the new billing account to the existing project.
- Verify that you are Billing Administrator for the billing account. Create a new project and link the new project to the existing billing account.

**Unattempted**

Our requirement is to link an existing google cloud project with a new billing account. Verify that you are Billing Administrator for the billing account. Create a new project and link the new project to the existing billing account. is not right.

?We do not need to create a new project. Verify that you are Project Billing Manager for the GCP project. Update the existing project to link it to the existing billing account. is not right.

?We want to link the project with a new billing account so is not right. Verify that you are Billing Administrator for the billing account. Update the existing project to link it to the existing billing account. is not right.

?We want to link the project with a new billing account so is not right. Verify that you are Project Billing Manager for the GCP project. Create a new billing account and link the new billing account to the existing project. is the right answer.

?The purpose of Project Billing Manager is to Link/unlink the project to/from a billing account. It is granted at the organization or project level. Project Billing Manager role allows a user to attach the project to the billing account, but does not grant any rights over resources. Project Owners can use this role to allow someone else to manage the billing for the project without granting them resource access.

?Ref: <https://cloud.google.com/billing/docs/how-to/billing-access>

## 27. Question

You need to create an autoscaling managed instance group for an HTTPS web application. You want to make sure that unhealthy VMs are recreated. What should you do?

- In the Instance Template, add the label health-check.
- In the Instance Template, add a startup script that sends a heartbeat to the metadata server.
- Select Multi-Zone instead of Single-Zone when creating the Managed Instance Group.
- Create a health check on port 443 and use that when creating the Managed Instance Group.

**Unattempted**

Our requirement is to ensure unhealthy VMs are recreated. Select Multi-Zone instead of Single-Zone when creating the Managed Instance Group. is not right.

?You can create two types of MIGs: A zonal MIG, which deploys instances to a single zone and a regional MIG, which deploys instances to multiple zones across the same region. However, this doesn't help with recreating unhealthy VMs.

?Ref: <https://cloud.google.com/compute/docs/instance-groups> In the Instance Template, add the label health-check. is not right.

?Metadata entries are key-value pairs and do not influence any other behavior.

?Ref: <https://cloud.google.com/compute/docs/storing-retrieving-metadata> In the Instance Template, add a startup script that sends a heartbeat to the metadata server. is not right.

?The startup script is executed only at the time of startup. Whereas we need something like a liveness check that monitors the status of the server periodically to identify if the VM is unhealthy. So this is not going to work.

?Ref: <https://cloud.google.com/compute/docs/startupscript> Create a health check on port 443 and use that when creating the Managed Instance Group. is the right answer.

?To improve the availability of your application and to verify that your application is responding, you can configure an auto-healing policy for your managed instance group (MIG). An auto-healing policy relies on an application-based health check to verify that an application is responding as expected. If the auto healer determines that an application isn't responding, the managed instance group automatically recreates that instance. Since our application is a HTTPS web application, we need to set up our health check on port 443 which is the standard port for HTTPS.

?Ref: <https://cloud.google.com/compute/docs/instance-groups/autohealing-instances-in-migs>

## 28. Question

You need to deploy an application, which is packaged in a container image, in a new project. The application exposes an HTTP endpoint and receives very few requests per day. You want to minimize costs. What should you do?

- Deploy the container on Cloud Run.
- Deploy the container on Cloud Run on GKE.
- Deploy the container on App Engine Flexible.
- Deploy the container on Google Kubernetes Engine, with cluster autoscaling and horizontal pod autoscaling enabled.

**Unattempted**

Deploy the container on Cloud Run on GKE. is not right.

?Cloud Run on GKE can scale the number of pods to zero. The number of nodes per cluster cannot scale to zero and these nodes are billed in the absence of requests.

?Ref: <https://cloud.google.com/serverless-options> Deploy the container on Google Kubernetes Engine, with cluster autoscaling and horizontal pod autoscaling enabled. is not right.

?Like above, while you can set up the pod autoscaler to scale back the pods to zero, the number of nodes per cluster cannot scale to zero and these nodes are billed in the absence of requests. If you specify the minimum node pool size of zero nodes, an idle node pool can scale down completely. However, at least one node must always be available in the cluster to run system Pods.

?Ref: <https://cloud.google.com/kubernetes-engine/docs/concepts/cluster-autoscaler> Deploy the container on App Engine Flexible. is not right.

?App Engine flexible environment instances are Compute Engine virtual machines. This means you can't truly scale down to zero and compute instances are billed in the absence of requests.

?Ref: <https://cloud.google.com/appengine/docs/flexible> Deploy the container on Cloud Run. is the right answer.

?Cloud Run is a fully managed compute platform that automatically scales your stateless containers. Cloud Run is serverless. Cloud Run abstracts away all infrastructure management. It automatically scales up and down from zero depending on traffic almost instantaneously. Cloud Run only charges you for the exact resources you use.

?Ref: <https://cloud.google.com/run>

#### 29. 29. Question

You need to enable traffic between multiple groups of Compute Engine instances that are currently running two different GCP projects. Each group of Compute Engine instances is running in its own VPC. What should you do?

- Verify that both projects are in a GCP Organization. Create a new VPC and add all instances.
- Verify that both projects are in a GCP Organization. Share the VPC from one project and request that the Compute Engine instances in the other project use this shared VPC.**
- Verify that you are the Project Administrator of both projects. Create two new VPCs and add all instances.
- Verify that you are the Project Administrator of both projects. Create a new VPC and add all instances.

#### Unattempted

Verify that both projects are in a GCP Organization. Share the VPC from one project and request that the Compute Engine instances in the other project use this shared VPC. is the right answer.

?All other options make no sense. Shared VPC allows an organization to connect resources from multiple projects to a common Virtual Private Cloud (VPC) network so that they can communicate with each other securely and efficiently using internal IPs from that network. When you use Shared VPC, you designate a project as a host project and attach one or more other service projects to it.

?Ref: <https://cloud.google.com/vpc/docs/shared-vpc>

?All other options make no sense.

#### 30. 30. Question

You need to grant access for three users so that they can view and edit table data on a Cloud Spanner instance. What should you do?

- Run gcloud iam roles describe roles/spanner.databaseUser. Add the users to the role.

- Run gcloud iam roles describe roles/spanner.viewer -- project my-gcp-ace-project. Add the users to a new group. Add the group to the role.
- Run gcloud iam roles describe roles/spanner.databaseUser. Add the users to a new group. Add the group to the role.**
- Run gcloud iam roles describe roles/spanner.viewer -- project my-gcp-ace-project. Add the users to the role.

### Unattempted

Our requirements

1. View and Edit table data
2. 3 users (i.e. multiple users)

3 users should give us the idea that we do not want to assign roles/permissions at the user level and instead want to do it based on groups so that we can create one group with all the required permissions and all such users who need this access can be assigned to the group.

Ref: <https://cloud.google.com/iam/docs/reference/rest/v1/Policy#Binding>

Ref: <https://cloud.google.com/iam/docs/granting-changing-revoking-access#granting-gcloud-manual>

Run gcloud iam roles describe roles/spanner.databaseUser. Add the users to the role. is not right.

We are looking for an option that assigns users to a group (in order to maximise reuse and minimize maintenance overhead). This option assigns users to the role so is not the right answer.

Run gcloud iam roles describe roles/spanner.viewer —project my-gcp-ace-project. Add the users to the role. is not right.

We are looking for an option that assigns users to a group (in order to maximise reuse and minimize maintenance overhead). This option assigns users to the role so is not the right answer.

Run gcloud iam roles describe roles/spanner.viewer —project my-gcp-ace-project. Add the users to a new group. Add the group to the role. is not right. Adding users to a group and granting the role to the group is the right way forward. But the role used in this option is spanner.viewer which allows viewing all Cloud Spanner instances (but cannot modify instances), and allows viewing all Cloud Spanner databases (but cannot modify databases and cannot read from databases). Since we required edit access as well, this option is not right.

Ref: <https://cloud.google.com/spanner/docs/iam>

Run gcloud iam roles describe roles/spanner.databaseUser. Add the users to a new group. Add the group to the role. is the right answer.

Adding users to a group and granting the role to the group is the right way forward. In addition, we assign the role spanner.databaseUser which allows Read from and write to the Cloud Spanner database; execute SQL queries on the database, including DML and Partitioned DML; and View and update schema for

the database. This is the only option that grants the right role to a group and assigns users to the group.

### 31. Question

You need to host an application on a Compute Engine instance in a project shared with other teams. You want to prevent the other teams from accidentally causing downtime on that application. Which feature should you use?

- Use a Shielded VM.
- Use a Preemptible VM.
- Use a sole-tenant node.
- Enable deletion protection on the instance.

**Unattempted**

Use a Shielded VM. is not right.

Shielded VMs are virtual machines (VMs) on Google Cloud hardened by a set of security controls that help defend against rootkits and boot kits. Using Shielded VMs helps protect enterprise workloads from threats like remote attacks, privilege escalation, and malicious insiders. But shielded VMs don't offer protection for accidental termination of the instance.

Ref: <https://cloud.google.com/shielded-vm>

Use a Preemptible VM. is not right.

A preemptible VM is an instance that you can create and run at a much lower price than normal instances. However, Compute Engine might terminate (preempt) these instances if it requires access to those resources for other tasks. Preemptible instances are excess Compute Engine capacity, so their availability varies with usage. Preemptible VMs don't offer protection for accidental termination of the instance.

Ref: <https://cloud.google.com/compute/docs/instances/preemptible>

Use a sole-tenant node. is not right.

Sole-tenancy lets you have exclusive access to a sole-tenant node, which is a physical Compute Engine server that is dedicated to hosting only your project's VMs. Use sole-tenant nodes to keep your VMs physically separated from VMs in other projects, or to group your VMs together on the same host hardware. Sole-tenant nodes don't offer protection for accidental termination of the instance.

Ref: <https://cloud.google.com/compute/docs/nodes>

Enable deletion protection on the instance. is the right answer.

As part of your workload, there might be certain VM instances that are critical to running your application or services, such as an instance running a SQL server, a server used as a license manager, and so on. These VM instances might need to stay running indefinitely so you need a way to protect these VMs from being deleted. By setting the deletionProtection flag, a VM instance can be protected from accidental deletion. If a user attempts to delete a VM instance for which you have set the deletionProtection flag, the request fails. Only a user that has been granted a role with compute.instances.create permission can reset the flag to

allow the resource to be deleted.

Ref: <https://cloud.google.com/compute/docs/instances/preventing-accidental-vm-deletion>

### 32. Question

You need to manage multiple Google Cloud Platform (GCP) projects in the fewest steps possible. You want to configure the Google Cloud SDK command line interface (CLI) so that you can easily manage multiple GCP projects. What should you?

- 1. Create a configuration for each project you need to manage. 2. Activate the appropriate configuration when you work with each of your assigned GCP projects.
- 1. Create a configuration for each project you need to manage. 2. Use gcloud init to update the configuration values when you need to work with a non-default project
- 1. Use the default configuration for one project you need to manage. 2. Activate the appropriate configuration when you work with each of your assigned GCP projects.
- 1. Use the default configuration for one project you need to manage. 2. Use gcloud init to update the configuration values when you need to work with a non-default project.

### Unattempted

1. Create a configuration for each project you need to manage.  
2. Activate the appropriate configuration when you work with each of your assigned GCP projects. is the right answer.

gcloud configurations enable you to manage multiple projects in gcloud cli using the fewest possible steps,

Ref: <https://cloud.google.com/sdk/gcloud/reference/config>

For example, we have two projects

```
$ gcloud projects list
PROJECT_ID NAME PROJECT_NUMBER
project-1-278333 project-1-278333 85524215451
project-2-278333 project-2-278333 25349885274
```

We create configuration for each project. For project-2-278333,

```
$ gcloud config configurations create project-1-config
```

```
$ gcloud config set project project-1-278333
```

And for project-2-278333,

```
$ gcloud config configurations create project-2-config
```

```
$ gcloud config set project project-2-278333
```

We now have two configurations, one for each project.

```
$ gcloud config configurations list
NAME IS_ACTIVE ACCOUNT PROJECT COMPUTE_DEFAULT_ZONE
COMPUTE_DEFAULT_REGION
cloudshell-4453 False
project-1-config False project-1-278333
project-2-config True project-2-278333
```

To activate configuration for project-1,

```
$ gcloud config configurations activate project-1-config
Activated [project-1-config].
```

```
$ gcloud config get-value project
```

```
Your active configuration is: [project-1-config]
project-1-278333
```

To activate configuration for project-2,

```
$ gcloud config configurations activate project-2-config
Activated [project-2-config].
```

```
$ gcloud config get-value project
```

```
Your active configuration is: [project-2-config]
project-2-278333
```

### 33. Question

You need to monitor resources that are distributed over different projects in Google Cloud Platform. You want to consolidate reporting under the same Stackdriver Monitoring dashboard. What should you do?

- Configure a single Stackdriver account, and link all projects to the same account.**
- Use Shared VPC to connect all projects, and link Stackdriver to one of the projects.
- Configure a single Stackdriver account for one of the projects. In Stackdriver, create a Group and add the other project names as criteria for that Group.
- For each project, create a Stackdriver account. In each project, create a service account for that project and grant it the role of Stackdriver Account Editor in all other projects.

**Unattempted**

Use Shared VPC to connect all projects, and link Stackdriver to one of the projects. is not right.

?Linking Stackdriver to one project brings metrics from that project alone. A Shared VPC allows an organization to connect resources from multiple projects to a common Virtual Private Cloud (VPC) network so that they can communicate with each other securely and efficiently using internal IPs from that network. But it does not help in linking all projects to a single Stackdriver workspace/account.

Ref: <https://cloud.google.com/vpc/docs/shared-vpc> For each project, create a

Stackdriver account. In each project, create a service account for that project and grant it the role of Stackdriver Account Editor in all other projects. is not right.  
?Stackdriver monitoring does not use roles to gather monitoring information from the project. Instead, the Stackdriver Monitoring agent, which is a collectd-based daemon, gathers system and application metrics from virtual machine instances and sends them to Monitoring. In this case, as each project is linked to a separate Stackdriver account, it is not possible to have a consolidated view of all monitoring. Ref: <https://cloud.google.com/monitoring/agent> Configure a single Stackdriver account for one of the projects. In Stackdriver, create a Group and add the other project names as criteria for that Group. is not right.

?As the other projects are not linked to the stack driver, they can't be monitored. Moreover, you can not add projects to Stackdriver groups. Groups provide a mechanism for alerting on the behavior of a set of resources, rather than on individual resources. For example, you can create an alerting policy that is triggered if some number of resources in the group violates a particular condition (for example, CPU load), rather than having each resource inform you of violations individually.

?Ref: <https://cloud.google.com/monitoring/groups> Configure a single Stackdriver account, and link all projects to the same account. is the right answer.

?You can monitor resources of different projects in a single Stackdriver account by creating a Stackdriver workspace. A Stackdriver workspace is a tool for monitoring resources contained in one or more Google Cloud projects or AWS accounts. Each Workspace can have between 1 and 100 monitored projects, including Google Cloud projects and AWS accounts. A Workspace accesses metric data from its monitored projects, but the metric data and log entries remain in the individual projects. Ref: <https://cloud.google.com/monitoring/workspaces>

#### 34. Question

You need to produce a list of the enabled Google Cloud Platform APIs for a GCP project using the gcloud command line in the Cloud Shell. The project name is my-project. What should you do?

- Run gcloud projects list to get the project ID, and then run gcloud services list --project .
- Run gcloud init to set the current project to my-project, and then run gcloud services list --available.
- Run gcloud info to view the account value, and then run gcloud services list --account .
- Run gcloud projects describe to verify the project value, and then run gcloud services list --available.

#### Unattempted

Run gcloud init to set the current project to my-project, and then run gcloud services list – available. is not right.

– available return the services available to the project to enable and not the services that are enabled. This list will include any services that the project has already enabled plus the services that the project can enable.

Ref: <https://cloud.google.com/sdk/gcloud/reference/services/list>

Also, to set the current project, you need to use gcloud config set project  
Ref: <https://cloud.google.com/sdk/gcloud/reference/config/set>  
gcloud init is used for initializing or reinitializing gcloud configurations.

<https://cloud.google.com/sdk/gcloud/reference/init>

Run gcloud info to view the account value, and then run gcloud services list –account . is not right.

We aren't passing any project id to the command so it would fail with the error shown below. (n.b. it is possible this command succeeds if you have an active gcloud configuration that has set the project so rather than accepting value from –project parameter, the command would obtain the project info from the gcloud configuration. The command shown below is run when no configuration is active).  
gcloud services list –account

Errors with the following error.

ERROR: (gcloud.services.list) The project property is set to the empty string, which is invalid.

To set your project, run:

\$ gcloud config set project PROJECT\_ID

or to unset it, run:

\$ gcloud config unset project

Run gcloud projects describe to verify the project value, and then run gcloud services list –available. is not right.

–available return the services available to the project to enable and not the services that are enabled. This list will include any services that the project has already enabled plus the services that the project can enable.

Ref: <https://cloud.google.com/sdk/gcloud/reference/services/list>

Run gcloud projects list to get the project ID, and then run gcloud services list –project . is the right answer.

For the gcloud services list command, –enabled is the default.

So running

gcloud services list –project is the same as running

gcloud services list –project –enabled

which would get all the enabled services for the project.

### 35. Question

You need to provide a cost estimate for a Kubernetes cluster using the GCP pricing calculator for Kubernetes. Your workload requires high IOPS, and you will also be using disk snapshots. You start by entering the number of nodes, average hours, and average days. What should you do next?

- Fill in local SSD. Fill in persistent disk storage and snapshot storage.**
- Fill in local SSD. Add estimated cost for cluster management.
- Select Add GPUs. Fill in persistent disk storage and snapshot storage.
- Select Add GPUs. Add estimated cost for cluster management.

**Unattempted**

Fill in local SSD. Add estimated cost for cluster management. is not right.  
You don't need to add an estimated cost for cluster management as the calculator automatically applies this. At the time of writing this, GKE clusters accrue a management fee of \$0.10 per cluster per hour, irrespective of cluster size or topology. One zonal cluster per billing account is free. GKE cluster management fees do not apply to Anthos GKE clusters.

Ref: <https://cloud.google.com/kubernetes-engine/pricing>

Select Add GPUs. Add estimated cost for cluster management. is not right.  
You don't need to add an estimated cost for cluster management as the calculator automatically applies this. At the time of writing this, GKE clusters accrue a management fee of \$0.10 per cluster per hour, irrespective of cluster size or topology. One zonal cluster per billing account is free. GKE cluster management fees do not apply to Anthos GKE clusters.

Ref: <https://cloud.google.com/kubernetes-engine/pricing>

Select Add GPUs. Fill in persistent disk storage and snapshot storage. is not right.

GPUs don't help us with our requirement of high IOPS. Compute Engine provides graphics processing units (GPUs) that you can add to your virtual machine instances to accelerate specific workloads on your instances such as machine learning and data processing. But this doesn't help increase IOPS.

Ref: <https://cloud.google.com/compute/docs/gpus>

Fill in local SSD. Fill in persistent disk storage and snapshot storage. is the right answer.

The pricing calculator for Kubernetes Engine offers us the ability to add GPUs as well as specify Local SSD requirements for estimation. GPUs don't help us with our requirement of high IOPS but Local SSD does.

Ref: <https://cloud.google.com/products/calculator>

GKE offers always-encrypted local solid-state drive (SSD) block storage. Local SSDs are physically attached to the server that hosts the virtual machine instance for very high input/output operations per second (IOPS) and very low latency compared to persistent disks.

Ref: <https://cloud.google.com/kubernetes-engine>

Once you fill in the local SSD requirement, you can fill in persistent disk storage and snapshot storage.

### 36. Question

You need to reduce GCP service costs for a division of your company using the fewest possible steps. You need to turn off all configured services in an existing GCP project. What should you do?

- 1. Verify that you are assigned the Project Owners IAM role for this project. 2. Locate the project in the GCP console, click Shut down and then enter the project ID.

- 1. Verify that you are assigned the Project Owners IAM role for this project. 2. Switch to the project in the GCP console, locate the resources and delete them.
- 1. Verify that you are assigned the Organization Administrator IAM role for this project. 2. Locate the project in the GCP console, enter the project ID and then click Shut down.
- 1. Verify that you are assigned the Organization Administrator IAM role for this project. 2. Switch to the project in the GCP console, locate the resources and delete them.

### Unattempted

1. Verify that you are assigned the Organization Administrator IAM role for this project.  
2. Locate the project in the GCP console, enter the project ID and then click Shut down. is not right.  
Organization Admin role provides permissions to get and list projects but not shutdown projects.  
Ref: <https://cloud.google.com/iam/docs/understanding-roles#resource-manager-roles>

1. Verify that you are assigned the Organization Administrator IAM role for this project.  
2. Switch to the project in the GCP console, locate the resources and delete them. is not right.  
Organization Admin role provides permissions to get and list projects but not delete projects.  
Ref: <https://cloud.google.com/iam/docs/understanding-roles#resource-manager-roles>

1. Verify that you are assigned the Project Owner IAM role for this project.  
2. Switch to the project in the GCP console, locate the resources and delete them. is not right.

The primitive Project Owner role provides permissionst to delete project  
[https://cloud.google.com/iam/docs/understanding-roles#primitive\\_roles](https://cloud.google.com/iam/docs/understanding-roles#primitive_roles)

But locating all the resources and deleting them is a manual task, time consuming and error prone. Our goal is to accomplish the same but with fewest possible steps

1. Verify that you are assigned the Project Owner IAM role for this project.  
2. Locate the project in the GCP console, click Shut down and then enter the project ID. is the right answer.

The primitive Project Owner role provides permissionst to delete project  
[https://cloud.google.com/iam/docs/understanding-roles#primitive\\_roles](https://cloud.google.com/iam/docs/understanding-roles#primitive_roles)

You can shut down projects using the Cloud Console. When you shut down a project, this immediately happens: All billing and traffic serving stops, You lose access to the project, The owners of the project will be notified and can stop the deletion within 30 days, The project will be scheduled to be deleted after 30 days. However, some resources may be deleted much earlier.

Ref: [https://cloud.google.com/resource-manager/docs/creating-managing-projects#shutting\\_down\\_projects](https://cloud.google.com/resource-manager/docs/creating-managing-projects#shutting_down_projects)

### 37. Question

You need to run an important query in BigQuery but expect it to return a lot of records. You want to find out how much it will cost to run the query. You are using on-demand pricing. What should you do?

- Run a select count (\*) to get an idea of how many records your query will look through. Then convert that number of rows to dollars using the Pricing Calculator.
- Arrange to switch to Flat-Rate pricing for this query, then move back to on-demand.
- Use the command line to run a dry run query to estimate the number of bytes returned. Then convert that bytes estimate to dollars using the Pricing Calculator.
- Use the command line to run a dry run query to estimate the number of bytes read. Then convert that bytes estimate to dollars using the Pricing Calculator.

#### Unattempted

Arrange to switch to Flat-Rate pricing for this query, then move back to on-demand. is not right.

The cost of acquiring a big query slot (associated with flat-rate pricing) is significantly higher than our requirement here to run a single important query or just to know how much it would cost to run that query. BigQuery offers flat-rate pricing for customers who prefer a stable monthly cost for queries rather than paying the on-demand price per TB of data processed. You enroll in flat-rate pricing by purchasing slot commitments, measured in BigQuery slots. Slot commitments start at 500 slots and the price starts from \$10000. Your queries consume this slot capacity, and you are not billed for bytes processed.

Ref: [https://cloud.google.com/bigquery/pricing#flat\\_rate\\_pricing](https://cloud.google.com/bigquery/pricing#flat_rate_pricing)

Use the command line to run a dry run query to estimate the number of bytes returned. Then convert that bytes estimate to dollars using the Pricing Calculator. is not right.

Under on-demand pricing, BigQuery doesn't charge for the query execution based on the output of the query (i.e. bytes returned) but on the number of bytes processed (also referred to as bytes read or bytes scanned) in order to arrive at the output of the query. You are charged for the number of bytes processed whether the data is stored in BigQuery or in an external data source such as Cloud Storage, Google Drive, or Cloud Bigtable. On-demand pricing is based solely on usage. You are charged for the bytes scanned even if your query itself doesn't return any data.

Ref: <https://cloud.google.com/bigquery/pricing>

Run a select count (\*) to get an idea of how many records your query will look through. Then convert that number of rows to dollars using the Pricing Calculator. is not right.

This is not as practical as identifying the number of records your query will look through (i.e. scan/process) is not straightforward. Plus BigQuery supports

external data sources such as Cloud Storage, Google Drive, or Cloud Bigtable; and the developer cost associated with identifying this information from various data sources is significant, not practical and sometimes not possible.

Use the command line to run a dry run query to estimate the number of bytes read. Then convert that bytes estimate to dollars using the Pricing Calculator. is the right answer.

BigQuery pricing is based on the number of bytes processed/read. Under on-demand pricing, BigQuery charges for queries by using one metric: the number of bytes processed (also referred to as bytes read). You are charged for the number of bytes processed whether the data is stored in BigQuery or in an external data source such as Cloud Storage, Google Drive, or Cloud Bigtable. On-demand pricing is based solely on usage.

Ref: <https://cloud.google.com/bigquery/pricing>

### 38. Question

You need to select and configure compute resources for a set of batch processing jobs. These jobs take around 2 hours to complete and are run nightly. You want to minimize service costs. What should you do?

- Select Compute Engine. Use VM instance types that support micro bursting.
- Select GKE. Use a single node cluster with a small instance type
- Select Compute Engine. Use preemptible VM instances of the appropriate standard machine types.**
- Select GKE. Use a three-node cluster with micro instance types.

**Unattempted**

Requirements – achieve end goal while minimizing service costs. Select GKE. Use a single node cluster with a small instance type is not right.

?We do not know if a small instance is capable of handling all the batch volume. Plus this is not the most cost-effective of the options. Select GKE. Use a three-node cluster with micro instance types is not right.

?We do not know if three micro instances are capable of handling all the batch volume. Plus this is not the most cost-effective of the options. Select Compute Engine. Use VM instance types that support micro bursting is not right.

?We can use an instance that supports micro bursting but we have a job that runs for 2 hours. Bursting is suitable for short periods. Select Compute Engine. Use preemptible VM instances of the appropriate standard machine types is the right answer.

?We minimize the cost by selecting a preemptible instance of the appropriate type. If the preemptible instance is terminated, the next nightly run picks up the unprocessed volume.

39. 39. Question

You need to set up a policy so that videos stored in a specific Cloud Storage Regional bucket are moved to Coldline after 90 days and then deleted after one year from their creation. How should you set up the policy?

- Use Cloud Storage Object Lifecycle Management using Age conditions with SetStorageClass and Delete actions. Set the SetStorageClass action to 90 days and the Delete action to 365 days.**
- Use gsutil rewrite and set the Delete action to 365 days.
- Use gsutil rewrite and set the Delete action to 275 days (365-90).
- Use Cloud Storage Object Lifecycle Management using Age conditions with SetStorageClass and Delete actions. Set the SetStorageClass action to 90 days and the Delete action to 275 days (365 - 90)

**Unattempted**

Use gsutil rewrite and set the Delete action to 275 days (365-90). is not right.  
gsutil rewrite is used to change the storage class of objects within a bucket through overwriting the object. It does not support Delete action.

Ref: <https://cloud.google.com/storage/docs/changing-storage-classes>

Use gsutil rewrite and set the Delete action to 365 days. is not right.  
gsutil rewrite is used to change the storage class of objects within a bucket through overwriting the object. It does not support Delete action.

Ref: <https://cloud.google.com/storage/docs/changing-storage-classes>

Use Cloud Storage Object Lifecycle Management using Age conditions with SetStorageClass and Delete actions. Set the SetStorageClass action to 90 days and the Delete action to 275 days (365 – 90). is not right.

Object Lifecycle Management does not rewrite an object when changing its storage class. This means that when an object is transitioned to Nearline Storage, Coldline Storage, or Archive Storage using the SetStorageClass feature, any subsequent early deletion and associated charges are based on the original creation time of the object, regardless of when the storage class changed.

If however, the change of storage class is done manually using a rewrite, the creation time of the objects is the new creation time since they are rewritten. In such a case, you would need to apply a lifecycle delete action of 275 days.

Ref: <https://cloud.google.com/storage/docs/lifecycle>

Use Cloud Storage Object Lifecycle Management using Age conditions with SetStorageClass and Delete actions. Set the SetStorageClass action to 90 days and the Delete action to 365 days. is the right answer.

Object Lifecycle Management does not rewrite an object when changing its storage class. This means that when an object is transitioned to Nearline Storage, Coldline Storage, or Archive Storage using the SetStorageClass feature, any subsequent early deletion and associated charges are based on the original creation time of the object, regardless of when the storage class changed.

Ref: <https://cloud.google.com/storage/docs/lifecycle>

#### 40. 40. Question

You need to set up permissions for a set of Compute Engine instances to enable them to write data into a particular Cloud Storage bucket. You want to follow Google-recommended practices. What should you do?

- Create a service account with an access scope. Use the access scope '<https://www.googleapis.com/auth/cloud-platform>'.
- Create a service account with an access scope. Use the access scope '[https://www.googleapis.com/auth/devstorage.write\\_only](https://www.googleapis.com/auth/devstorage.write_only)'.
- Create a service account and add it to the IAM role 'storage.objectAdmin' for that bucket.
- Create a service account and add it to the IAM role 'storage.objectCreator' for that bucket.

#### Unattempted

Our requirements are

1. Google recommended practices
2. Multiple compute engine instances to write data to a bucket.

Create a service account with an access scope. Use the access scope '[https://www.googleapis.com/auth/devstorage.write\\_only](https://www.googleapis.com/auth/devstorage.write_only)' . is not right.

There is no scope called “ write\_only” .

Ref: <https://cloud.google.com/storage/docs/authentication>

Create a service account and add it to the IAM role ‘ storage.objectCreator’ for that bucket. is not right.

You can’t add a service account to a role. The relationship is the other way round. You grant roles to the service account. See below a screenshot of the role. Ref: <https://cloud.google.com/iam/docs/granting-roles-to-service-accounts>

Create a service account and add it to the IAM role ‘ storage.objectAdmin’ for that bucket. is not right.

You can’t add a service account to a role. The relationship is the other way round. You grant roles to the service account.

Ref: <https://cloud.google.com/iam/docs/granting-roles-to-service-accounts>

Create a service account with an access scope. Use the access scope '<https://www.googleapis.com/auth/cloud-platform>' . is the right answer. cloud-platform role lets you view and manage data across all Google Cloud services. For Cloud Storage, this is the same as devstorage.full-control which allows full control over data, including the ability to modify IAM policies.

Ref: <https://cloud.google.com/storage/docs/authentication>

#### 41. 41. Question

You need to trigger a budget alert for Compute Engine charges on one of the three Google Cloud Platform projects that you manage. All three projects are linked to a single billing account. What should you do?

- Verify that you have Billing Account Administrator role. Select the associated billing account and create a budget and alert for the appropriate project.
- Verify that you have Billing Account Administrator role. Select the associated billing account and create a budget and a custom alert.
- Verify that you have Project Billing Manager role. Select the associated billing account and create a budget for the appropriate project.
- Verify that you have Project Billing Manager role. Select the associated billing account and create a budget and a custom alert.

### Unattempted

Verify that you have Project Billing Manager role. Select the associated billing account and create a budget for the appropriate project. is not right.  
Project Billing Manager role allows a user to attach the project to the billing account, but does not grant any rights over resources. This role does not provide user permissions to view spending, create budgets and alerts.

Ref: <https://cloud.google.com/billing/docs/how-to/billing-access#overview-of-cloud-billing-roles-in-cloud-iam>

Verify that you have Project Billing Manager role. Select the associated billing account and create a budget and a custom alert. is not right.  
Project Billing Manager role allows a user to attach the project to the billing account, but does not grant any rights over resources. This role does not provide user permissions to view spending, create budgets and alerts.

Ref: <https://cloud.google.com/billing/docs/how-to/billing-access#overview-of-cloud-billing-roles-in-cloud-iam>

Verify that you have Billing Account Administrator role. Select the associated billing account and create a budget and a custom alert. is not right.  
Billing Account Administrator role enables a user to view spend and set budget alerts. But the budget here isn't scoped to the single project that we are interested in. Since the single billing account is linked to all three projects, this results in budget alerts being triggered for Compute Engine usage on all three projects – which is against our requirements.

Ref: <https://cloud.google.com/billing/docs/how-to/billing-access#overview-of-cloud-billing-roles-in-cloud-iam>

Ref: <https://cloud.google.com/billing/docs/how-to/budgets#budget-scope>

Verify that you have Billing Account Administrator role. Select the associated billing account and create a budget and alert for the appropriate project. is the right answer.

Billing Account Administrator role enables a user to view spend and set budget alerts. In addition, the budget here is scoped to a single project. Therefore, when the compute engine spend exceeds the budget threshold in the project, we send an alert, and this only works for the scoped project, and not all projects linked to the billing account.

Ref: <https://cloud.google.com/billing/docs/how-to/billing-access#overview-of-cloud-billing-roles-in-cloud-iam>

Ref: <https://cloud.google.com/billing/docs/how-to/budgets#budget-scope>

42. 42. Question

You need to update a deployment in Deployment Manager without any resource downtime in the deployment. Which command should you use?

- gcloud deployment-manager deployments update --config**
- gcloud deployment-manager deployments create --config
- gcloud deployment-manager resources create --config
- gcloud deployment-manager resources update --config

**Unattempted**

gcloud deployment-manager resources create – config . is not right.  
gcloud deployment-manager resources command does not support the action create. The supported actions are describe and list. So this option is not right.  
Ref: <https://cloud.google.com/sdk/gcloud/reference/deployment-manager/resources>

gcloud deployment-manager resources update – config . is not right.  
gcloud deployment-manager resources command does not support the action update. The supported actions are describe and list. So this option is not right.  
Ref: <https://cloud.google.com/sdk/gcloud/reference/deployment-manager/resources>

gcloud deployment-manager deployments create – config . is not right.  
gcloud deployment-manager deployments create – creates a deployment but we want to update a deployment. So this option is not right.  
Ref: <https://cloud.google.com/sdk/gcloud/reference/deployment-manager/deployments/create>

gcloud deployment-manager deployments update – config . is the right answer.  
gcloud deployment-manager deployments update – updates a deployment based on a provided config file and fits our requirement.

<https://cloud.google.com/sdk/gcloud/reference/deployment-manager/deployments/update>

43. 43. Question

You plan to deploy an application on an autoscaled managed instances group. The application uses a tomcat server and runs on port 8080. You want to access the application on <https://www.example.com>. You want to follow Google recommended practices. What services would you use?

- Google Domains, Cloud DNS private zone, HTTP(S) Load Balancer
- Google Domains, Cloud DNS private zone, SSL Proxy Load Balancer
- Google DNS, Google CDN, SSL Proxy Load Balancer

- Google Domains, Cloud DNS, HTTP(S) Load Balancer**

**Unattempted**

To serve traffic on <https://www.example.com>, we have to first own the domain example.com. We can use Google Domains service to register a domain.

?Ref: <https://domains.google/> Once we own example.com domain, we need to create a zone <http://www.example.com>. We can use Cloud DNS, which is a scalable, reliable, and managed authoritative Domain Name System (DNS) to create a DNS zone.

?Ref: <https://cloud.google.com/dns> Once the <http://www.example.com> zone is set up, we need to create a DNS (A) record to point to the public IP of the Load Balancer. This is also carried out in Cloud DNS. Finally, we need a load balancer to front the autoscaled managed instances group. Google recommends we use HTTP(S) Load Balancer for this requirement as “ SSL Proxy Load Balancing is intended for non-HTTP(S) traffic. For HTTP(S) traffic, we recommend that you use HTTP(S) Load Balancing.”

?Ref: <https://cloud.google.com/load-balancing/docs/ssl> So the correct answer is Google Domains, Cloud DNS, HTTP(S) Load Balancer

#### 44. Question

You ran the following commands to create two compute instances.

```
gcloud compute instances create instance1
```

```
gcloud compute instances create instance2
```

Both compute instances were created in europe-west2-a zone but you want to create them in other zones. Your active gcloud configuration is as shown below.

```
$ gcloud config list
```

```
[component_manager]
```

```
disable_update_check = True
```

```
[compute]
```

```
gce_metadata_read_timeout_sec = 5
```

```
zone = europe-west2-a
```

```
[core]
```

```
account = gcp-ace-lab-user@gmail.com
```

```
disable_usage_reporting = False
```

```
project = gcp-ace-lab-266520
```

```
[metrics]
```

environment = devshell

You want to modify the gcloud configuration such that you are prompted for a zone when you execute the create instance commands above. What should you do?

- gcloud config unset compute/zone**
- gcloud config set zone ""
- gcloud config set compute/zone ""
- gcloud config unset zone

#### Unattempted

gcloud config unset zone. is not right.

gcloud config does not have a core/zone property. The syntax for this command is gcloud config unset SECTION/PROPERTY. If SECTION is missing from the command, SECTION is defaulted to core. We are effectively trying to run the command gcloud config unset core/zone but the core section doesn't have a property called zone, so this command fails.

\$ gcloud config unset zone

ERROR: (gcloud.config.unset) Section [core] has no property [zone].

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/unset>

gcloud config set zone " ". is not right.

gcloud config does not have a core/zone property. The syntax for this command is gcloud config set SECTION/PROPERTY VALUE. If SECTION is missing, SECTION is defaulted to core. We are effectively trying to run gcloud config set core/zone " " but the core section doesn't have a property called zone, so this command fails.

\$ gcloud config set zone " "

ERROR: (gcloud.config.unset) Section [core] has no property [zone].

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/set>

gcloud config set compute/zone " ". is not right.

This command uses the correct syntax but it doesn't unset the compute/zone property correctly. Instead it sets it to " " in gcloud configuration. When the gcloud compute instances create command runs, it picks the zone value from this configuration property which is " " and attempts to create an instance in " " zone and fails because zone " " doesn't exist. gcloud doesn't treat " " zone as an unset value. The zone must be explicitly unset if it is to be removed from the configuration.

\$ gcloud config set compute/zone " "

\$ gcloud compute instances create instance1

Zone: Expected type (, ) for field id, found projects/compute-challenge-lab-266520/zones/ (type )

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/set>

gcloud config unset compute/zone. is the right answer.

This command uses the correct syntax and correctly unsets the zone in gcloud

configuration. The next time gcloud compute instances create command runs, it knows there is no default zone defined in gcloud configuration and therefore prompts for a zone before the instance can be created.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/unset>

#### 45. Question

You recently deployed a new application in Google App Engine to serve production traffic. After analyzing logs for various user flows, you uncovered several issues in your application code and have developed a fix to address the issues. Parts of your proposed fix could not be validated in the pre-production environment by your testing team as some of the scenarios can only be validated by an end-user with access to specific data in your production environment. In the company's weekly Change Approval Board meeting, concerns were raised that the fix could possibly take down the application. It was unanimously agreed that while the fix is risky, it is a necessary change to the application. You have been asked to suggest a solution that minimizes the impact of the change going wrong. You also want to minimize costs. What should you do?

- Deploy a new version of the application, and use traffic splitting to send a small percentage of traffic to it.
- Set up a second Google App Engine service, and then update a subset of clients to hit the new service.
- Deploy the new application version temporarily, capture logs and then roll it back to the previous version.
- Create a second Google App Engine project with the new application code, and onboard users gradually to the new application.

#### Unattempted

Deploy the new application version temporarily, capture logs and then roll it back to the previous version. is not right.

Deploying a new application version and promoting it would result in your new version serving all production traffic. If the code fix doesn't work as expected, it would result in the application becoming unreachable to all users. This is a risky approach and should be avoided.

Create a second Google App Engine project with the new application code, and onboard users gradually to the new application. is not right.

You want to minimize costs. This approach effectively doubles your costs as you have to pay for two identical environments until all users are moved over to the new application. There is an additional overhead of manually onboarding users to the new application which could be expensive as well as time-consuming.

Set up a second Google App Engine service, and then update a subset of clients to hit the new service. is not right.

It is not straightforward to update a set of clients to hit the new service. When users access an App Engine service, they use an endpoint like [https://SERVICE\\_ID-dot-PROJECT\\_ID.REGION\\_ID.r.appspot.com](https://SERVICE_ID-dot-PROJECT_ID.REGION_ID.r.appspot.com).

Introducing a new service introduces a new URL and getting your users to use

the new URL is possible but involves effort and coordination. If you want to mask these differences to the end-user, then you have to make changes in the DNS and use a weighted algorithm to split the traffic between the two services based on the weights assigned.

Ref: <https://cloud.google.com/appengine/docs/standard/python/splitting-traffic>

Ref: <https://cloud.google.com/appengine/docs/standard/python/an-overview-of-app-engine>

This approach also has the drawback of doubling your costs until all users are moved over to the new service.

Deploy a new version of the application, and use traffic splitting to send a small percentage of traffic to it. is the right answer.

This option minimizes the risk to the application while also minimizing the complexity and cost. When you deploy a new version to App Engine, you can choose not to promote it to serve live traffic. Instead, you could set up traffic splitting to split traffic between the two versions – this can all be done within Google App Engine. Once you send a small portion of traffic to the new version, you can analyze logs to identify if the fix has worked as expected. If the fix hasn't worked, you can update your traffic splitting configuration to send all traffic back to the old version. If you are happy your fix has worked, you can send more traffic to the new version or move all user traffic to the new version and delete the old version.

Ref: <https://cloud.google.com/appengine/docs/standard/python/splitting-traffic>

Ref: <https://cloud.google.com/appengine/docs/standard/python/an-overview-of-app-engine>

#### 46. Question

You recently deployed a new version of an application to App Engine and then discovered a bug in the release. You need to immediately revert to the prior version of the application. What should you do?

- On the App Engine page of the GCP Console, select the application that needs to be reverted and click Revert.
- On the App Engine Versions page of the GCP Console, route 100% of the traffic to the previous version.**
- Deploy the original version as a separate application. Then go to App Engine settings and split traffic between applications so that the original version serves 100% of the requests.
- Run gcloud app restore.

**Unattempted**

Run gcloud app restore. is not right.

restore action is not supported by gcloud app command.

Ref: <https://cloud.google.com/sdk/gcloud/reference/app/deploy>

On the App Engine page of the GCP Console, select the application that needs to be reverted and click Revert. is not right.

Revert option is not present on the App Engine page of the GCP Console.

Deploy the original version as a separate application. Then go to App Engine settings and split traffic between applications so that the original version serves 100% of the requests. is not right.

Each application in the app engine is different and it is not possible to split traffic between applications in App Engine. You can use traffic splitting to specify a percentage distribution of traffic across two or more of the versions within a service but not across applications.

Ref: <https://cloud.google.com/appengine/docs/standard/python/splitting-traffic>

On the App Engine Versions page of the GCP Console, route 100% of the traffic to the previous version. is the right answer

You can roll back to a previous version in the app engine GCP console. Go back to the list of versions and check the box next to the version that you want to receive all traffic and click the MAKE DEFAULT button located above the list.

Traffic immediately switches over to the selected version.

Ref: <https://cloud.google.com/community/tutorials/how-to-roll-your-app-engine-managed-vms-app-back-to-a-previous-version-part-1>

#### 47. Question

You significantly changed a complex deployment manager template and want to confirm that the dependencies of all defined resources are properly met before committing it to the project. You want the most rapid feedback on your changes. What would you do?

- Use granular logging statements within the Deployment Manager template authored in Python.
- Monitor activity of the Deployment Manager executing on Stackdriver logging page of the GCP console.
- Execute the Deployment manager template against a separate project with the same configuration, and monitor for failures
- Execute the Deployment Manager template using the --preview option in the same project, and observe the status of interdependent resources

Unattempted

Requirements – confirm dependencies, rapid feedback.

Use granular logging statements within the Deployment Manager template authored in Python. is not right.

Deployment Manager doesn't provide the ability to set granular logging statements. Moreover, if that was possible the logging statements wouldn't be written to a log file until the template is applied and it is already too late as the template is applied and we haven't had a chance to confirm that the dependencies of all defined resources are properly met

Monitor activity of the Deployment Manager executing on Stackdriver logging page of the GCP console. is not right.

This doesn't give us a chance to confirm that the dependencies of all defined resources are properly met before executing it.

Execute the Deployment manager template against a separate project with the same configuration, and monitor for failures. is not right.

While we can identify whether dependencies are met by monitoring the failures, it is not rapid. We need rapid feedback on changes and we want that before changes are committed (i.e. applied) to the project

Execute the Deployment Manager template using the – preview option in the same project, and observe the status of interdependent resources. is the right answer.

After we have written a configuration file, we can preview the configuration before you create a deployment. Previewing a configuration lets you see the resources that Deployment Manager would create but does not actually instantiate any actual resources. In gcloud command-line, you use the create sub-command with the – preview flag to preview configuration changes.

Ref: <https://cloud.google.com/deployment-manager>

#### 48. Question

You want to add a new auditor to a Google Cloud Platform project. The auditor should be allowed to read, but not modify, all project items. How should you configure the auditor's permissions?

- Create a custom role with view-only project permissions. Add the user's account to the custom role.
- Create a custom role with view-only service permissions. Add the user's account to the custom role.
- Select the built-in IAM project Viewer role. Add the user's account to this role.
- Select the built-in IAM service Viewer role. Add the user's account to this role.

**Unattempted**

Select the built-in IAM project Viewer role. Add the user's account to this role. Is the right answer The primitive role roles/viewer provides read access to all resources in the project. The permissions in this role are limited to Get and list access for all resources. As we have an out of the box role that exactly fits our requirement, we should use this.

?Ref: <https://cloud.google.com/resource-manager/docs/access-control-project>

?It is advisable to use the existing GCP provided roles over creating custom roles with similar permissions as this becomes a maintenance overhead. If GCP modifies how permissions are handled or adds/removes permissions, the default GCP provided roles are automatically updated by Google whereas if they were

custom roles, the responsibility is with us and this adds to the operational overhead and needs to be avoided.

#### 49. Question

You want to configure 10 Compute Engine instances for availability when maintenance occurs. Your requirements state that these instances should attempt to automatically restart if they crash. Also, the instances should be highly available including during system maintenance. What should you do?

- Create an instance group for the instance. Verify that the Advanced creation options setting for do not retry machine creation is set to off.
- Create an instance group for the instances. Set the Autohealing health check to healthy (HTTP).
- Create an instance template for the instances. Set Automatic Restart to off. Set On-host maintenance to Terminate VM instances. Add the instance template to an instance group.
- Create an instance template for the instances. Set the Automatic Restart to on. Set the On-host maintenance to Migrate VM instance. Add the instance template to an instance group.**

#### Unattempted

##### Requirements

1. 10 instances – indicates we need to look for MIG (Managed Instances Group) where we can configure healing/scaling settings. All options talk about creating an instance group so this point isn't of much use, unfortunately.
2. Highly available during system maintenance – indicates we need to look for Live Migration.
3. Automatically restart on crash – indicates we need to look for options that enable automatic restarts.

Create an instance template for the instances.

Set Automatic Restart to off.

Set On-host maintenance to Terminate VM instances.

Add the instance template to an instance group. is not right.

If Automatic Restart is off, then the compute engine instances are not automatically restarted. This results in loss of capacity and if GCP decides to start system maintenance on all instances at the same time, all instances are down and this does not meet our requirement “ Highly available during system maintenance” so this option is not right.

Create an instance group for the instances.

Set the Autohealing health check to healthy (HTTP). is not right.

While auto-healing helps with the recreation of VM instances when needed, it doesn't Live-migrate the instances so our requirement of “ highly available including during system maintenance” is not met. More info about Autohealing – Auto-healing allows the recreation of VM instances when needed. You can use a health check to recreate a VM instance if the health check finds it unresponsive. If you don't select a health check, Compute Engine will recreate VM instances

only when they're not running.

Ref: [https://cloud.google.com/compute/docs/instance-groups/?hl=en\\_GB#managed\\_instance\\_groups\\_and\\_autohealing](https://cloud.google.com/compute/docs/instance-groups/?hl=en_GB#managed_instance_groups_and_autohealing)

Create an instance group for the instance.

Verify that the Advanced creation options setting for do not retry machine creation is set to off. is not right.

Like above – this option doesn't Live-migrate the instances so our requirement of "highly available including during system maintenance" is not met.

Create an instance template for the instances.

Set the Automatic Restart to on.

Set the On-host maintenance to Migrate VM instance.

Add the instance template to an instance group. is the right option.

Enabling automatic restart ensures that compute engine instances are automatically restarted when they crash. And Enabling "Migrate VM Instance" enables live migrates i.e. compute instances are migrated during system maintenance and remain running during the migration.

Automatic Restart – If your instance is set to terminate when there is a maintenance event, or if your instance crashes because of an underlying hardware issue, you can set up Compute Engine to automatically restart the instance by setting the automaticRestart field to true. This setting does not apply if the instance is taken offline through a user action, such as calling sudo shutdown, or during a zone outage.

Ref: <https://cloud.google.com/compute/docs/instances/setting-instance-scheduling-options#autorestart>

Enabling the Migrate VM Instance option migrates your instance away from an infrastructure maintenance event, and your instance remains running during the migration. Your instance might experience a short period of decreased performance, although generally, most instances should not notice any difference. This is ideal for instances that require constant uptime and can tolerate a short period of decreased performance.

Ref: [https://cloud.google.com/compute/docs/instances/setting-instance-scheduling-options#live\\_migrate](https://cloud.google.com/compute/docs/instances/setting-instance-scheduling-options#live_migrate)

## 50. Question

You want to configure a cost-effective solution for archiving objects in a Cloud Storage bucket. Noncurrent versions should be archived after 30 days. Non-current versions are accessed once a month for reporting. This archived objects are also occasionally updated at month-end. What should you do?

- Add a bucket lifecycle rule that archives noncurrent versions after 30 days to Coldline Storage.
- Add a bucket lifecycle rule that archives noncurrent versions after 30 days to Nearline Storage.**
- Add a bucket lifecycle rule that archives objects from regional storage after 30 days to Coldline Storage.

- Add a bucket lifecycle rule that archives objects from regional storage after 30 days to Nearline Storage.

### Unattempted

We don't know what the current storage class is. In the absence of this information and considering the 4 options provided, it is safe to assume that objects are currently in Regional or Multi-Regional buckets. We want to archive noncurrent versions after 30 days and you need to read and modify on average once per month. Add a bucket lifecycle rule that archives noncurrent versions after 30 days to Coldline Storage. is not right.

?Coldline Storage is ideal for data you plan to read or modify at most once a quarter. Our requirement states we need to read or modify every month so Coldline Storage is not an ideal storage class for our requirement.

?Ref: <https://cloud.google.com/storage/docs/storage-classes#coldline> Add a bucket lifecycle rule that archives objects from regional storage after 30 days to Coldline Storage. is not right.

?Coldline Storage is ideal for data you plan to read or modify at most once a quarter. Our requirement states we need to read or modify every month so Coldline Storage is not an ideal storage class for our requirement. Moreover, we don't want to archive live versions, we want to archive just the noncurrent versions.

?Ref: <https://cloud.google.com/storage/docs/storage-classes#coldline> Add a bucket lifecycle rule that archives objects from regional storage after 30 days to Nearline Storage. is not right.

?While Nearline Storage is ideal for data you plan to read or modify on average once per month or less, we don't want to archive live versions, we want to archive just the noncurrent versions.

?Ref: <https://cloud.google.com/storage/docs/storage-classes#nearline> Add a bucket lifecycle rule that archives noncurrent versions after 30 days to Nearline Storage. is the right answer.

?Nearline Storage is ideal for data you plan to read or modify on average once per month or less. And this option archives just the noncurrent versions which is what we want to do.

?<https://cloud.google.com/storage/docs/storage-classes#nearline>

### 51. Question

You want to configure an SSH connection to a single Compute Engine instance for users in the dev1 group. This instance is the only resource in this particular Google Cloud Platform project that the dev1 users should be able to connect to. What should you do?

- Set metadata to enable-oslogin=true for the instance. Grant the dev1 group the compute.osLogin role. Direct them to use the Cloud Shell to ssh to that instance.

- Set metadata to enable-oslogin=true for the instance. Set the service account to no service account for that instance. Direct them to use the Cloud Shell to ssh to that instance.
- Enable block project wide keys for the instance. Generate an SSH key for each user in the dev1 group. Distribute the keys to dev1 users and direct them to use their third-party tools to connect.
- Enable block project wide keys for the instance. Generate an SSH key and associate the key with that instance. Distribute the key to dev1 users and direct them to use their third-party tools to connect.

### Unattempted

You have multiple ways to connect to instances. More information can be found here: <https://cloud.google.com/compute/docs/instances/access-overview>

Enable block project wide keys for the instance. Generate an SSH key for each user in the dev1 group. Distribute the keys to dev1 users and direct them to use their third-party tools to connect. is not right.

Generating SSH keys for users is fine but unless the SSH keys are added to the instance, users would not be able to SSH to the server. If you need your instance to ignore project-wide public SSH keys and use only the instance-level keys, you can block project-wide public SSH keys from the instance. This allows only users whose public SSH key is stored in instance-level metadata to access the instance.

Ref: <https://cloud.google.com/compute/docs/instances/adding-removing-ssh-keys#edit-ssh-metadata>

Enable block project wide keys for the instance. Generate an SSH key and associate the key with that instance. Distribute the key to dev1 users and direct them to use their third-party tools to connect. is not right.

While this is possible, sharing SSH keys is a strict NO from a security point of view as this breaks auditing. Should one of the developers create a disaster (either accidental or malicious), your security admin would be unable to identify which of the users in dev1 group caused the issue.

Set metadata to enable-oslogin=true for the instance. Set the service account to no service account for that instance. Direct them to use the Cloud Shell to ssh to that instance. is not right.

After you enable OS Login on one or more instances in your project, those VMs accept connections only from user accounts that have the necessary IAM roles in your project or organization. In this case, since we have not granted either of these roles – roles/compute.osLogin or roles/compute.osAdminLogin role, users of dev1 group can't SSH to the server.

Ref: [https://cloud.google.com/compute/docs/instances/managing-instance-access#configure\\_users](https://cloud.google.com/compute/docs/instances/managing-instance-access#configure_users)

Set metadata to enable-oslogin=true for the instance. Grant the dev1 group the compute.osLogin role. Direct them to use the Cloud Shell to ssh to that instance. is the right answer.

After you enable OS Login on one or more instances in your project, those VMs

accept connections only from user accounts that have the necessary IAM roles in your project or organization. In this case, we are granting the group compute.osLogin which lets them log in as non-administrator account. And since we are directing them to use Cloud Shell to ssh, we don't need to add their SSH keys to the instance metadata.

Ref: [https://cloud.google.com/compute/docs/instances/managing-instance-access#configure\\_users](https://cloud.google.com/compute/docs/instances/managing-instance-access#configure_users)

Ref: [https://cloud.google.com/compute/docs/instances/managing-instance-access#add\\_oslogin\\_keys](https://cloud.google.com/compute/docs/instances/managing-instance-access#add_oslogin_keys)

## 52. Question

You want to configure auto-healing for network load balancer for a group of Compute Engine instances that run in multiple zones using the fewest possible steps. You need to configure the recreation of VMs if they are unresponsive after 3 attempts of 10 seconds each. What should you do?

- Create a HTTP load balancer with a backend configuration that references an existing instance group. Define a balancing mode and set the maximum RPS to 10
- Create a HTTP load balancer with a backend configuration that references an existing instance group. Set the health check to healthy (HTTP)
- Create a managed instance group. Verify that the auto-scaling setting is on.
- Create a managed instance group. Set the Autohealing health check to healthy (HTTP)

**Unattempted**

Create a managed instance group. Verify that the auto-scaling setting is on. is not right.

While auto-scaling capabilities of Managed instance groups let you automatically add or delete instances from a managed instance group based on increases or decreases in load, they don't help you with re-creation should the VMs go unresponsive.

Ref: <https://cloud.google.com/compute/docs/autoscaler>

Create a HTTP load balancer with a backend configuration that references an existing instance group. Define a balancing mode and set the maximum RPS to 10 is not right.

You set RPS (Requests per Second) on load balancer when using RATE balancing mode. This has no effect on auto-healing.

Ref: <https://cloud.google.com/load-balancing/docs/https/>

Create a HTTP load balancer with a backend configuration that references an existing instance group. Set the health check to healthy (HTTP) is not right.

The health checks defined on the load balancer determine whether VM instances respond properly to traffic. This has no impact on auto-healing. It is important to note that the health checks defined on the load balancer are different to the

health checks defined on the auto-healing for managed instances group – see the explanation in the right answer for more information.

Ref: [https://cloud.google.com/load-balancing/docs/health-checks#create\\_a\\_health\\_check](https://cloud.google.com/load-balancing/docs/health-checks#create_a_health_check)

Create a managed instance group. Set the Autohealing health check to healthy (HTTP) is the right answer.

In order to enable auto-healing, you need to group the instances into a managed instance group. Managed instance groups (MIGs) maintain the high availability of your applications by proactively keeping your virtual machine (VM) instances available. An auto-healing policy on the MIG relies on an application-based health check to verify that an application is responding as expected. If the auto-healer determines that an application isn't responding, the managed instance group automatically recreates that instance.

It is important to use separate health checks for load balancing and for auto-healing. Health checks for load balancing can and should be more aggressive because these health checks determine whether an instance receives user traffic. You want to catch non-responsive instances quickly, so you can redirect traffic if necessary. In contrast, health checking for auto-healing causes Compute Engine to proactively replace failing instances, so this health check should be more conservative than a load balancing health check.

### 53. Question

You have been asked to migrate a docker application from datacenter to cloud. Your solution architect has suggested uploading docker images to GCR in one project and running an application in a GKE cluster in a separate project. You want to store images in the project img-278322 and run the application in the project prod-278986. You want to tag the image as acme\_track\_n\_trace:v1. You want to follow Google-recommended practices. What should you do?

- Run gcloud builds submit --tag gcr.io/img-278322/acme\_track\_n\_trace
- Run gcloud builds submit --tag gcr.io/img-278322/acme\_track\_n\_trace:v1**
- Run gcloud builds submit --tag gcr.io/prod-278986/acme\_track\_n\_trace
- Run gcloud builds submit --tag gcr.io/prod-278986/acme\_track\_n\_trace:v1

**Unattempted**

Run gcloud builds submit --tag gcr.io/img-278322/acme\_track\_n\_trace. is not right.

?This command tags the image as acme\_track\_n\_trace:latest but we want to tag the image as acme\_track\_n\_trace:v1.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/builds/submit> Run gcloud builds submit – tag gcr.io/prod-278986/acme\_track\_n\_trace. is not right.  
?This command tags the image as acme\_track\_n\_trace:latest but we want to tag the image as acme\_track\_n\_trace:v1. This command also upload the image to the wrong project.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/builds/submit> Run gcloud builds submit – tag gcr.io/prod-278986/acme\_track\_n\_trace:v1. is not right.  
?This command also upload the image to the wrong project.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/builds/submit> Run gcloud builds submit – tag gcr.io/img-278322/acme\_track\_n\_trace:v1. is the right answer.  
?This command correctly tags the image as acme\_track\_n\_trace:v1 and uploads the image to the img-278322 project.

?Ref: <https://cloud.google.com/sdk/gcloud/reference/builds/submit>

#### 54. Question

You have designed a solution on Google Cloud Platform (GCP) that uses multiple GCP products. Your company has asked you to estimate the costs of the solution. You need to provide estimates for the monthly total cost. What should you do?

- For each GCP product in the solution, review the pricing details on the products pricing page. Use the pricing calculator to total the monthly costs for each GCP product.
- For each GCP product in the solution, review the pricing details on the products pricing page. Create a Google Sheet that summarizes the expected monthly costs for each product.
- Provision the solution on GCP. Leave the solution provisioned for 1 week. Navigate to the Billing Report page in the Google Cloud Platform Console. Multiply the 1 week cost to determine the monthly costs.
- Provision the solution on GCP. Leave the solution provisioned for 1 week. Use Stackdriver to determine the provisioned and used resource amounts. Multiply the 1 week cost to determine the monthly costs.

#### Unattempted

Provision the solution on GCP. Leave the solution provisioned for 1 week. Use Stackdriver to determine the provisioned and used resource amounts. Multiply the 1 week cost to determine the monthly costs. is not right.

?By provisioning the solution on GCP, you are going to incur costs. Our requirement is to just estimate the costs and this can be done by using Google Cloud Pricing Calculator.

?Ref: <https://cloud.google.com/products/calculator> Provision the solution on GCP. Leave the solution provisioned for 1 week. Use Stackdriver to determine the provisioned and used resource amounts. Multiply the 1 week cost to determine the monthly costs. is not right.

?By provisioning the solution on GCP, you are going to incur costs. Our requirement is to just estimate the costs and this can be done by using Google Cloud Pricing Calculator.

?Ref: <https://cloud.google.com/products/calculator> For each GCP product in the solution, review the pricing details on the products pricing page. Create a Google Sheet that summarizes the expected monthly costs for each product. is not right.

?This would certainly work but is a manual task. Why use this when you can use Google Cloud Pricing Calculator to achieve the save?

?Ref: <https://cloud.google.com/products/calculator> For each GCP product in the solution, review the pricing details on the products pricing page. Use the pricing calculator to total the monthly costs for each GCP product. is the right answer.

?You can use the Google Cloud Pricing Calculator to total the estimated monthly costs for each GCP product. You don't incur any charges for doing so.

?Ref: <https://cloud.google.com/products/calculator>

## 55. Question

You have files in a Cloud Storage bucket that you need to share with your suppliers. You want to restrict the time that the files are available to your suppliers to 1 hour. You want to follow Google recommended practices. What should you do?

- Create a service account with just the permissions to access files in the bucket. Create a JSON key for the service account. Execute the command `gsutil signurl -p 60m gs:///`.
- Create a service account with just the permissions to access files in the bucket. Create a JSON key for the service account. Execute the command `gsutil signurl -d 1h gs:///**`.**
- Create a JSON key for the Default Compute Engine Service Account. Execute the command `gsutil signurl -t 60m gs:///*.*`.
- Create a service account with just the permissions to access files in the bucket. Create a JSON key for the service account. Execute the command `gsutil signurl -m 1h gs:///*`.

## Unattempted

Create a JSON key for the Default Compute Engine Service Account. Execute the command `gsutil signurl -t 60m gs:///*.*` is not right.  
`gsutil signurl` does not support `-t` flag. Executing the command with `-t` flag fails as shown.

```
$ gsutil signurl -t 60m keys.json gs://gcp-ace-lab-255520/*.*
```

CommandException: Incorrect option(s) specified. Usage:

Ref: <https://cloud.google.com/storage/docs/gsutil/commands/signurl>  
Also, using the default compute engine service account violates the principle of least privilege. The recommended approach is to create a service account with just the right permissions needed and create JSON keys for this service account to use with gsutil signurl command.

Create a service account with just the permissions to access files in the bucket. Create a JSON key for the service account. Execute the command gsutil signurl -p 60m gs:/// is not right.

With gsutil signurl, -p is used to specify the key store password instead of prompting for the password. It can not be used to pass a time value. Executing the command with -p flag fails as shown.

```
$ gsutil signurl -p 60m keys.json gs://gcp-ace-lab-255520/*.*
```

TypeError: Last argument must be a byte string or a callable.

Ref: <https://cloud.google.com/storage/docs/gsutil/commands/signurl>

Create a service account with just the permissions to access files in the bucket. Create a JSON key for the service account. Execute the command gsutil signurl -m 1h gs:///\*. is not right.

With gsutil signurl, -m is used to specify the operation e.g. PUT/GET etc.

Executing the command with -m flag fails as shown.

```
$ gsutil signurl -m 1h keys.json gs://gcp-ace-lab-255520/*.*
```

CommandException: HTTP method must be one  
of[GET|HEAD|PUT|DELETE|RESUMABLE]

Ref: <https://cloud.google.com/storage/docs/gsutil/commands/signurl>

Create a service account with just the permissions to access files in the bucket. Create a JSON key for the service account. Execute the command gsutil signurl -d 1h gs:///\*\*. is the right answer.

This command correctly specifies the duration that the signed url should be valid for by using the -d flag. The default is 1 hour so omitting the -d flag would have also resulted in the same outcome. Times may be specified with no suffix (default hours), or with s = seconds, m = minutes, h = hours, d = days. The max duration allowed is 7d.

Ref: <https://cloud.google.com/storage/docs/gsutil/commands/signurl>

## 56. Question

You have one GCP account running in your default region and zone and another account running in a non-default region and zone. You want to start a new Compute Engine instance in these two Google Cloud Platform accounts using the command line interface. What should you do?

- Activate two configurations using gcloud configurations activate [NAME]. Run gcloud config list to start the Compute Engine instances.
- Activate two configurations using gcloud configurations activate [NAME]. Run gcloud configurations list to start the Compute Engine instances.

- Create two configurations using gcloud config configurations create [NAME]. Run gcloud config configurations activate [NAME] to switch between accounts when running the commands to start the Compute Engine instances.
- Create two configurations using gcloud config configurations create [NAME]. Run gcloud configurations list to start the Compute Engine instances.

### Unattempted

Create two configurations using gcloud config configurations create [NAME]. Run gcloud configurations list to start the Compute Engine instances. is not right.  
gcloud configurations list is an invalid command. To list the existing named configurations, you need to execute gcloud config configurations list but this does not start the compute engine instances.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/list>

Activate two configurations using gcloud configurations activate [NAME]. Run gcloud configurations list to start the Compute Engine instances. is not right.  
gcloud configurations list is an invalid command. To list the existing named configurations, you need to execute gcloud config configurations list but this does not start the compute engine instances.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/list>

Activate two configurations using gcloud configurations activate [NAME]. Run gcloud config list to start the Compute Engine instances. is not right.  
gcloud configurations activate [NAME] activates an existing named configuration. It can't be used to activate two configurations at the same time. Moreover, gcloud config list lists Cloud SDK properties for the currently active configuration. It does not start the Compute Engine instances.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/activate>

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/list>

Create two configurations using gcloud config configurations create [NAME]. Run gcloud config configurations activate [NAME] to switch between accounts when running the commands to start the Compute Engine instances. is the right answer.

Each gcloud configuration has a 1 to 1 relationship with the region (if a region is defined). Since we have two different regions, we would need to create two separate configurations using gcloud config configurations create

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/create>

Secondly, you can activate each configuration independently by running gcloud config configurations activate [NAME]

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/activate>

Finally, while each configuration is active, you can run the gcloud compute instances start [NAME] command to start the instance in the configuration's

region.

<https://cloud.google.com/sdk/gcloud/reference/compute/instances/start>

57. Question

You have one project called ptech-sa where you manage all your service accounts. You want to be able to use a service account from this project to take snapshots of VMs running in another project called ptech-vm. What should you do?

- When creating the VMs, set the service account's API scope for Compute Engine to read/write.
- Grant the service account the IAM Role of Compute Storage Admin in the project called ptech-vm.**
- Download the private key from the service account, and add it to each VMs custom metadata.
- Download the private key from the service account, and add the private key to each VM's SSH keys.

**Unattempted**

Download the private key from the service account, and add it to each VMs custom metadata. is not right.

?Adding service accounts private key (JSON file) to VMs custom metadata has no effect. Metadata entries are key-value pairs and do not influence any other behavior.

?Ref: <https://cloud.google.com/compute/docs/storing-retrieving-metadata> Download the private key from the service account, and add the private key to each VM' s SSH keys. is not right.

?Adding service accounts private key to the VMs SSH keys does not influence any other behavior. SSH keys are used for SSHing to the instance.

?<https://cloud.google.com/compute/docs/instances/adding-removing-ssh-keys> When creating the VMs, set the service account' s API scope for Compute Engine to read/write. is not right.

?The scopes can be modified only when using compute engine default service account.

?Ref: [https://cloud.google.com/compute/docs/access/service-accounts#default\\_service\\_account](https://cloud.google.com/compute/docs/access/service-accounts#default_service_account)

?See the screenshot below.

?The scopes can not be modified when using a non-default service account. See the screenshot below.

?Since we want to use service accounts from another project, it is safe to say they are not the default compute service accounts of this project and hence it is

not possible to customize the scopes. Grant the service account the IAM Role of Compute Storage Admin in the project called ptech-vm. is the right answer.

?Compute Storage Admin role provides permissions to create, modify, and delete disks, images, and snapshots. If the service account in ptech-sa is granted the IAM Role of Compute Storage Admin in the project called ptech-vm, it can take snapshots and carry out other activities as defined by the role.

?Ref: <https://cloud.google.com/compute/docs/access/iam#compute.storageAdmin>

#### 58. Question

You have production and test workloads that you want to deploy on Compute Engine. Production VMs need to be in a different subnet than the test VMs. All the VMs must be able to reach each other over internal IP without creating additional routes. You need to set up VPC and the 2 subnets. Which configuration meets these requirements?

- Create 2 custom VPCs, each with a single subnet. Create each subnet in a different region and with a different CIDR range.
- Create a single custom VPC with 2 subnets. Create each subnet in a different region and with a different CIDR range.**
- Create 2 custom VPCs, each with a single subnet. Create each subnet in the same region and with the same CIDR range.
- Create a single custom VPC with 2 subnets. Create each subnet in the same region and with the same CIDR range.

**Unattempted**

Create 2 custom VPCs, each with a single subnet. Create each subnet in a different region and with a different CIDR range. is not right.

?We need to get our requirements working with 1 VPC, not 2 !! Create 2 custom VPCs, each with a single subnet. Create each subnet in the same region and with the same CIDR range. is not right.

?We need to get our requirements working with 1 VPC, not 2 !! Create a single custom VPC with 2 subnets. Create each subnet in the same region and with the same CIDR range. is not right.

?We can not create two subnets in one VPC with the same CIDR range.  
“ Primary and secondary ranges for subnets cannot overlap with any allocated range, any primary or secondary range of another subnet in the same network, or any IP ranges of subnets in peered networks.”

Ref: <https://cloud.google.com/vpc/docs/using-vpc#subnet-rules> Create a single custom VPC with 2 subnets. Create each subnet in a different region and with a different CIDR range. is the right answer.

?When we create subnets in the same VPC with different CIDR ranges, they can communicate automatically within VPC. “ Resources within a VPC network can communicate with one another by using internal (private) IPv4 addresses, subject to applicable network firewall rules”

?Ref: <https://cloud.google.com/vpc/docs/vpc>

59. **59. Question**

You have sensitive data stored in three Cloud Storage buckets and have enabled data access logging. You want to verify activities for a particular user for these buckets, using the fewest possible steps. You need to verify the addition of metadata labels and which files have been viewed from those buckets. What should you do?

- View the bucket in the Storage section of the GCP Console.
- Using the GCP Console, filter the Activity log to view the information.
- Using the GCP Console, filter the Stackdriver log to view the information.**
- Create a trace in Stackdriver to view the information.

**Unattempted**

Our requirements are – sensitive data, verify access, fewest possible steps.

Using the GCP Console, filter the Activity log to view the information. is not right. Since data access logging is enabled, you can see relevant log entries in both activity Logs as well as stack driver logs. However, verifying what has been viewed/updated is not straightforward in activity logs. Activity logs display a list of all actions and you can restrict this down to a user and further filter by specifying Data access as the Activity types and GCS Bucket as the Resource type. But that is the extent of the filter functionality in Activity logs. It is not possible to restrict the activity logs to just the three buckets that we are interested in. Secondly, it is not possible to restrict the activity logs to just the gets and updates. So we'd have to go through the full list to identify activities of interest before verifying them which is a manual process and can be time taking depending on the number of activities in the list.

Ref: <https://cloud.google.com/storage/docs/audit-logs>

View the bucket in the Storage section of the GCP Console. is not right.  
The bucket page in the GCP console does not show the logs.

Create a trace in Stackdriver to view the information. is not right.  
Stackdriver trace is not supported on google cloud. Stackdriver Trace runs on Linux in the following environments: Compute Engine, Google Kubernetes Engine (GKE), App Engine flexible environment, App Engine standard environment.

Ref: <https://cloud.google.com/trace/docs/overview>

Using the GCP Console, filter the Stackdriver log to view the information. is the right answer.

Data access logs is already enabled, so we already record all API calls that read the configuration or metadata of resources, as well as user-driven API calls that create, modify, or read user-provided resource data. Data Access audit logs do not record the data-access operations on resources that are publicly shared (available to All Users or All Authenticated Users) or that can be accessed without logging into Google Cloud.

Since we are dealing with sensitive data, it is safe to assume that these buckets are not publicly shared and therefore enabling Data access logging logs all data-access operations on resources. These logs are sent to Stackdriver where they can be viewed by applying a suitable filter.

Unlike activity logs, retrieving the required information to verify is easier and quicker through Stackdriver as you can apply filters such as

```
resource.type=" gcs_bucket"
(resource.labels.bucket_name=" gcp-ace-lab-255520? OR
resource.labels.bucket_name=" gcp-ace-lab-255521? OR
resource.labels.bucket_name=" gcp-ace-lab-255522?)
(protoPayload.methodName=" storage.objects.get" OR
protoPayload.methodName=" storage.objects.update")
protoPayload.authenticationInfo.principalEmail=" test.gcp.labs.user@gmail.com"
```

and query just the gets and updates, for specific buckets for a specific user. This involves fewer steps and is more efficient.

Data access logging is not enabled by default and needs to be enabled explicitly. The screenshot below shows a screenshot for enabling the data access logging for Google Cloud Storage.

#### 60. Question

You have successfully created a development environment in a project for an application. This application uses Compute Engine and Cloud SQL. Now, you need to create a production environment for this application. The security team has forbidden the existence of network routes between these 2 environments, and asks you to follow Google-recommended practices. What should you do?

- Create a new project, enable the Compute Engine and Cloud SQL APIs in that project, and replicate the setup you have created in the development environment.
- Create a new production subnet in the existing VPC and a new production Cloud SQL instance in your existing project, and deploy your application using those resources.
- Create a new project, modify your existing VPC to be a Shared VPC, share that VPC with your new project, and replicate the setup you have in the development environment in that new project, in the Shared VPC.

- Ask the security team to grant you the Project Editor role in an existing production project used by another division of your company. Once they grant you that role, replicate the setup you have in the development environment in that project.

### Unattempted

Create a new project, modify your existing VPC to be a Shared VPC, share that VPC with your new project, and replicate the setup you have in the development environment in that new project, in the Shared VPC. is not right.

A Shared VPC allows an organization to connect resources from multiple projects to a common Virtual Private Cloud (VPC) network so that they can communicate with each other securely and efficiently using internal IPs from that network. This goes totally against the recommendations of the security team.

Ref: <https://cloud.google.com/vpc/docs/shared-vpc>

Create a new production subnet in the existing VPC and a new production Cloud SQL instance in your existing project, and deploy your application using those resources. is not right.

You can't achieve complete isolation between development and production environments. When configuration access in Cloud SQL, while you can grant any application access to a Cloud SQL instance by authorizing the public IP addresses that the application uses to connect, you can not specify a private network (for example, 10.x.x.x) as an authorized network. The compute engine instances use their private IP addresses to reach out to Cloud SQL and because of the above limitation, we can't prevent the development compute engine reach out to production MySQL and vice versa. Since the security team has forbidden the existence of network routes between these 2 environments, having the production and development environments in a single project is not an option.

<https://cloud.google.com/sql/docs/mysql/connect-external-app#appaccessIP>

Ask the security team to grant you the Project Editor role in an existing production project used by another division of your company. Once they grant you that role, replicate the setup you have in the development environment in that project. is not right.

While this would technically isolate the development environment from the production environment, your production application is running in a project that is also hosting production applications of another division of your company. This goes against Google's recommended practices. You can use folders to isolate requirements for different departments and teams in the parent organization. And you have separate projects under the folders so as per Google recommendations we should be deploying the production application to a separate project that is just for one company division/department.

Ref: <https://cloud.google.com/docs/enterprise/best-practices-for-enterprise-organizations#define-hierarchy>

Create a new project, enable the Compute Engine and Cloud SQL APIs in that project, and replicate the setup you have created in the development environment. is the right answer.

This aligns with Google's recommended practices. By creating a new project, we achieve complete isolation between development and production environments; as well as isolate this production application from production applications of other departments.

Ref: <https://cloud.google.com/docs/enterprise/best-practices-for-enterprise-organizations#define-hierarchy>

61. Question

You have three gcloud configurations – one for each of development, test and production projects. You want to list all the configurations and switch to a new configuration. With the fewest steps possible, what's the fastest way to switch to the correct configuration?

- To list configurations - gcloud config list To activate a configuration - gcloud config activate.
- To list configurations - gcloud configurations list To activate a configuration - gcloud configurations activate
- To list configurations - gcloud configurations list To activate a configuration - gcloud config activate.
- To list configurations - gcloud config configurations list To activate a configuration - gcloud config configurations activate.

**Unattempted**

To list configurations – gcloud configurations list

To activate a configuration – gcloud configurations activate. is not right.

gcloud configurations list does not list configurations. To list existing configurations, you need to execute gcloud config configurations list.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/list>  
gcloud configurations activate does not activate a named configuration. To activate a configuration, you need to execute gcloud config configurations activate

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/activate>

To list configurations – gcloud config list

To activate a configuration – gcloud config activate. is not right.

gcloud config list does not list configurations. It lists the properties of the existing configuration. To list existing configurations, you need to execute gcloud config configurations list.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/list>  
gcloud config activate does not activate a named configuration. To activate a configuration, you need to execute gcloud config configurations activate

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/activate>

To list configurations – gcloud configurations list

To activate a configuration – gcloud config activate. is not right.

gcloud configurations list does not list configurations. To list existing configurations, you need to execute gcloud config configurations list.

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/list>  
gcloud config activate does not activate a named configuration. To activate a configuration, you need to execute gcloud config configurations activate

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/activate>

To list configurations – gcloud config configurations list

To activate a configuration – gcloud config configurations activate. is the right answer.

The two commands together achieve the intended outcome. gcloud config configurations list – lists existing named configurations and gcloud config configurations activate – activates an existing named configuration

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/list>

Ref: <https://cloud.google.com/sdk/gcloud/reference/config/configurations/activate>

See an example below

```
$ gcloud config configurations list
```

| NAME               | IS_ACTIVE | ACCOUNT          | PROJECT | DEFAULT_ZONE | DEFAULT_REGION |
|--------------------|-----------|------------------|---------|--------------|----------------|
| dev-configuration  | False     | gcp-ace-lab-dev  |         |              |                |
| prod-configuration | False     | gcp-ace-lab-prod |         |              |                |
| test-configuration | True      | gcp-ace-lab-test |         |              |                |

```
$ gcloud config configurations activate prod-configuration
```

Activated [prod-configuration].

```
$ gcloud config configurations list
```

| NAME               | IS_ACTIVE | ACCOUNT          | PROJECT | DEFAULT_ZONE | DEFAULT_REGION |
|--------------------|-----------|------------------|---------|--------------|----------------|
| dev-configuration  | False     | gcp-ace-lab-dev  |         |              |                |
| prod-configuration | True      | gcp-ace-lab-prod |         |              |                |
| test-configuration | False     | gcp-ace-lab-test |         |              |                |

## 62. Question

You have two compute instances in the same VPC but in different regions. You can SSH from one instance to another instance using their external IP address but not their internal IP address. What could be the reason for SSH failing on the internal IP address?

- The internal IP address is disabled.
- The combination of compute instance network tags and VPC firewall rules allow SSH from 0.0.0.0 but denies SSH from the VPC subnets IP range.
- The compute instances are not using the right cross-region SSH IAM permissions
- The compute instances have a static IP for their internal IP.

### Unattempted

The compute instances have a static IP for their internal IP. is not right.

?Static internal IPs shouldn't be a reason for failed SSH connections. With all networking set up correctly, SSH works fine on Static internal IPs.

?Ref: <https://cloud.google.com/compute/docs/ip-addresses#networkaddresses>

The internal IP address is disabled. is not right.

?Every compute instance has one or more internal IP addresses so this option is not correct. The compute instances are not using the right cross-region SSH IAM permissions. is not right.

?There is no such thing as cross region SSH IAM permissions. The combination of compute instance network tags and VPC firewall rules allow SSH from 0.0.0.0 but denies SSH from the VPC subnets IP range. is the right answer.

?The combination of compute instance network tags and VPC firewall rules can certainly result in SSH traffic being allowed on the external IP range but disabled from subnets IP range. The firewall rule can be configured to allow SSH traffic from 0.0.0.0/0 but deny traffic from the VPC range e.g. 10.0.0.0/8. In this case, all SSH traffic from within the VPC is denied but external SSH traffic (i.e. on external IP) is allowed.

?Ref: <https://cloud.google.com/vpc/docs/using-firewalls>

### 63. Question

You have two compute instances in the same VPC but in different regions. You can SSH from one instance to another instance using their internal IP address but not their external IP address. What could be the reason for SSH failing on external IP address?

- The combination of compute instance network tags and VPC firewall rules only allow SSH from the subnets IP range.
- The external IP address is disabled.
- The compute instances have a static IP for their external IP.
- The compute instances are not using the right cross-region SSH IAM permissions

**Unattempted**

The compute instances have a static IP for their external IP. is not right.

?Not having a static IP is not a reason for failed SSH connections. When the firewall rules are set up correctly, SSH works fine on compute instances having ephemeral IP Address. The external IP address is disabled. is not right.

?Our question states SSH doesn't work on external IP addresses so it is safe to assume they already have an external IP. Therefore, this option is not correct. The compute instances are not using the right cross-region SSH IAM permissions. is not right.

?There is no such thing as cross region SSH IAM permissions. The combination of compute instance network tags and VPC firewall rules only allow SSH from the subnets IP range. is the right answer.

?The combination of compute instance network tags and VPC firewall rules can certainly result in SSH traffic being allowed from only subnets IP range. The firewall rule can be configured to allow SSH traffic from just the VPC range e.g. 10.0.0.0/8. In this scenario, all SSH traffic from within the VPC is accepted but external SSH traffic is blocked.

?Ref: <https://cloud.google.com/vpc/docs/using-firewalls>

64. Question

You have two Kubernetes resource configuration files.

deployments.yaml – creates a deployment

service.yaml – sets up a LoadBalancer service to expose the pods.

You don't have a GKE cluster in the development project and you need to provision one. Which of the commands fail with an error in Cloud Shell when you are attempting to create a GKE cluster and deploy the YAML configuration files to create a deployment and service. (Select Two)

- gcloud container clusters create cluster-1 --zone=us-central1-a  
gcloud container clusters get-credentials cluster-1 --zone=us-central1-a kubectl apply -f [deployment.yaml,service.yaml]**
- gcloud container clusters create cluster-1 --zone=us-central1-a  
gcloud container clusters get-credentials cluster-1 --zone=us-central1-a kubectl apply -f deployment.yaml&&service.yaml**
- gcloud config set compute/zone us-central1-a gcloud container clusters create cluster-1 gcloud container clusters get-credentials cluster-1 --zone=us-central1-a kubectl apply -f deployment.yaml kubectl apply -f service.yaml
- gcloud container clusters create cluster-1 --zone=us-central1-a gcloud container clusters get-credentials cluster-1 --zone=us-central1-a kubectl apply -f deployment.yaml kubectl apply -f service.yaml
- gcloud container clusters create cluster-1 --zone=us-central1-a gcloud container clusters get-credentials cluster-1 --zone=us-central1-a kubectl apply -f deployment.yaml,service.yaml

**Unattempted**

gcloud container clusters create cluster-1 – zone=us-central1-a  
gcloud container clusters get-credentials cluster-1 – zone=us-central1-a  
kubectl apply -f deployment.yaml

kubectl apply -f service.yaml. is not right (i.e. command executes successfully)  
You create a cluster by running gcloud container clusters create command. You then fetch credentials for a running cluster by running gcloud container clusters get-credentials command. Finally, you apply the kubernetes resource configuration by running kubectl apply -f

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/create>

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/get-credentials>

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#apply>

gcloud container clusters create cluster-1 – zone=us-central1-a  
gcloud container clusters get-credentials cluster-1 – zone=us-central1-a  
kubectl apply -f deployment.yaml,service.yaml. is not right (i.e. commands executes successfully)

Like above, but the only difference is that both configurations are applied in the same statement. With kubectl apply, you can apply the configuration from a single file or multiple files or even a complete directory.

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#apply>

gcloud config set compute/zone us-central1-a  
gcloud container clusters create cluster-1  
gcloud container clusters get-credentials cluster-1 – zone=us-central1-a  
kubectl apply -f deployment.YAML  
kubectl apply -f service.yaml. is not right (i.e. commands executes successfully)  
Like above, but the only difference is in how the compute zone is set. In this scenario, you set the zone us-central1-a as the default zone so when you don't pass a zone property to the gcloud container clusters create command, it takes the default zone which is us-central1-a.

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#apply>

gcloud container clusters create cluster-1 – zone=us-central1-a  
gcloud container clusters get-credentials cluster-1 – zone=us-central1-a  
kubectl apply -f [deployment.yaml,service.yaml]. is the right answer (i.e. commands fail)  
kubectl apply can apply the configuration from a single file or multiple files or even a complete directory. When applying configuration from multiple files, the file names need to be separated by a comma. In this scenario, the filenames are passed as a list and Kubernetes treats the list as literal so looks for files “[deployment.yaml]” and “[service.yaml]” which it doesn't find.

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#apply>

gcloud container clusters create cluster-1 – zone=us-central1-a  
gcloud container clusters get-credentials cluster-1 – zone=us-central1-a  
kubectl apply -f deployment.yaml&&service.yaml. is the right answer (i.e. commands fail)  
kubectl apply can apply the configuration from a single file or multiple files or even a complete directory. When applying configuration from multiple files, the file names need to be separated by a comma. In this scenario, the filenames are separated by && and kubernetes treats the && as literal so it looks for the file “deployment.yaml&&service.yaml” which it doesn't find.

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#apply>

## 65. 65. Question

You have two Kubernetes resource configuration files.

deployments.yaml – creates a deployment

service.YAML – sets up a LoadBalancer service to expose the pods.

You don't have a GKE cluster in the development project and you need to provision one. Which of the commands below would you run in Cloud Shell to create a GKE cluster and deploy the YAML configuration files to create a deployment and service?

- `gcloud container clusters create cluster-1 --zone=us-central1-a gcloud container clusters get-credentials cluster-1 --zone=us-central1-a kubectl create -f deployment.yaml kubectl create -f service.yaml`
- `gcloud container clusters create cluster-1 --zone=us-central1-a  
gcloud container clusters get-credentials cluster-1 --zone=us-central1-a kubectl apply -f deployment.yaml kubectl apply -f service.yaml`
- `gcloud container clusters create cluster-1 --zone=us-central1-a gcloud container clusters get-credentials cluster-1 --zone=us-central1-a gcloud gke apply -f deployment.yaml gcloud gke apply -f service.yaml`
- `kubectl container clusters create cluster-1 --zone=us-central1-a kubectl container clusters get-credentials cluster-1 --zone=us-central1-a kubectl apply -f deployment.yaml kubectl apply -f service.yaml`

### Unattempted

`kubectl container clusters create cluster-1 – zone=us-central1-a  
kubectl container clusters get-credentials cluster-1 – zone=us-central1-a  
kubectl apply -f deployment.yaml  
kubectl apply -f service.yaml.`

is not right.

`kubectl` doesn't support `kubectl container clusters create` command. `kubectl` can not be used to create GKE clusters. To create a GKE cluster, you need to execute `gcloud container clusters create` command.

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/create>

`gcloud container clusters create cluster-1 – zone=us-central1-a  
gcloud container clusters get-credentials cluster-1 – zone=us-central1-a  
kubectl create -f deployment.yaml  
kubectl create -f service.yaml.`

is not right.

`kubectl` doesn't support `kubectl create` command. The YAML file contains the cluster resource configuration. You don't create the configuration, instead, you apply the configuration to the cluster. The configuration can be applied by running `kubectl apply` command

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#apply>

```
gcloud container clusters create cluster-1 --zone=us-central1-a
gcloud container clusters get-credentials cluster-1 --zone=us-central1-a
gcloud gke apply -f deployment.yaml
gcloud gke apply -f service.yaml.
```

is not right.

gcloud doesn't support gcloud gke apply command. The YAML file contains the cluster resource configuration. You don't create the configuration, instead, you apply the configuration to the cluster. The configuration can be applied by running kubectl apply command

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#apply>

```
gcloud container clusters create cluster-1 --zone=us-central1-a
gcloud container clusters get-credentials cluster-1 --zone=us-central1-a
kubectl apply -f deployment.yaml
kubectl apply -f service.yaml.
```

is the right answer.

You create a cluster by running gcloud container clusters create command. You then fetch credentials for a running cluster by running gcloud container clusters get-credentials command. Finally, you apply the Kubernetes resource configuration by running kubectl apply -f

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/create>

Ref: <https://cloud.google.com/sdk/gcloud/reference/container/clusters/get-credentials>

Ref: <https://kubernetes.io/docs/reference/generated/kubectl/kubectl-commands#app>

# SET-11

## 1. Question

You are creating a Kubernetes Engine cluster to deploy multiple pods inside the cluster. All container logs must be stored in BigQuery for later analysis. You want to follow Google-recommended practices. Which two approaches can you take?

- A. Turn on Stackdriver Logging during the Kubernetes Engine cluster creation.
- B. Turn on Stackdriver Monitoring during the Kubernetes Engine cluster creation.
- C. Develop a custom add-on that uses Cloud Logging API and BigQuery API. Deploy the add-on to your Kubernetes Engine cluster.
- D. Use the Stackdriver Logging export feature to create a sink to Cloud Storage. Create a Cloud Dataflow job that imports log files from Cloud Storage to BigQuery.
- E. Use the Stackdriver Logging export feature to create a sink to BigQuery. Specify a filter expression to export log records related to your Kubernetes Engine cluster only.

**Incorrect**

Correct answers are A & E

Option A as creating a cluster with Stackdriver Logging option will enable all the container logs to be stored in Stackdriver Logging.

Option E as Stackdriver Logging support exporting logs to BigQuery by creating sinks

Refer GCP documentation – Kubernetes logging

Option B is wrong as creating a cluster with Stackdriver Monitoring option will enable monitoring metrics to be gathered, but it has nothing to do with logging.

Option C is wrong as even if you can develop a Kubernetes addon that will send logs to BigQuery, this is not a Google-recommended practice.

Option D is wrong as this is not a Google recommended practice.

## 2. Question

Your company has a mission-critical application that serves users globally. You need to select a transactional and relational data storage system for this application. Which two products should you choose?

- A. BigQuery
- B. Cloud SQL
- C. Cloud Spanner
- D. Cloud Bigtable
- E. Cloud Datastore

**Unattempted**

Correct answers are B & C

Option B as because Cloud SQL is a relational and transactional database in the list.

Option C as Spanner is a relational and transactional database in the list.

Refer GCP documentation – Storage Options

Option A is wrong as BigQuery is not a transactional system.

Option D is wrong as Cloud Bigtable provides transactional support but it's not relational.

Option E is wrong as Datastore is not a relational data storage system.

3. 3. Question

You want to find out who in your organization has Owner access to a project called “ my-project” . What should you do?

- A. In the Google Cloud Platform Console, go to the IAM page for your organization and apply the filter Role:Owner.
- B. In the Google Cloud Platform Console, go to the IAM page for your project and apply the filter Role:Owner.
- C. Use gcloud iam list-grantable-role --project my-project from your Terminal.
- D. Use gcloud iam list-grantable-role from Cloud Shell on the project page.

**Unattempted**

Correct answer is B as this shows you the Owners of the project.

Option A is wrong as it will give the org-wide owners, but you are interested in the project owners, which could be different.

Option C is wrong as this command is to list grantable roles for a resource, but does not return who has a specific role.

Option D is wrong as this command is to list grantable roles for a resource, but does not return who has a specific role.

4. 4. Question

You need to verify the assigned permissions in a custom IAM role. What should you do?

- A. Use the GCP Console, IAM section to view the information.
- B. Use the gcloud init command to view the information.
- C. Use the GCP Console, Security section to view the information.
- D. Use the GCP Console, API section to view the information.

**Unattempted**

Correct answer is A as this is the correct console area to view permission assigned to a custom role in a particular project.

Refer GCP documentation – IAM Custom Rules

Option B is wrong as gcloud init will not provide the information required.

Options C and D are wrong as these are not the correct areas to view this information

5. 5. Question

You have an App Engine application serving as your front-end. It's going to publish messages to Pub/Sub. The Pub/Sub API hasn't been enabled yet. What is the fastest way to enable the API?

- A. Use a service account with the Pub/Sub Admin role to auto-enable the API.
- B. Enable the API in the Console.
- C. Application's in App Engine don't require external APIs to be enabled.
- D. The API will be enabled the first time the code attempts to access Pub/Sub.

**Unattempted**

Correct answer is B as the simplest way to enable an API for the project is using the GCP console.

### Refer GCP documentation – Enable/Disable APIs

The simplest way to enable an API for your project is to use the GCP Console, though you can also enable an API using gcloud or using the Service Usage API. You can find out more about these options in the Service Usage API docs.

To enable an API for your project using the console:

1. Go to the GCP Console API Library.
2. From the projects list, select a project or create a new one.
3. In the API Library, select the API you want to enable. If you need help finding the API, use the search field and/or the filters.
4. On the API page, click ENABLE.

Option A is wrong as providing the Pub/Sub Admin role does not provide the access to enable API.

Enabling an API requires the following two Cloud Identity and Access Management permissions:

1. The servicemanagement.services.bind permission on the service to enable. This permission is present for all users for public services. For private services, you must share the service with the user who needs to enable it.
2. The serviceusage.services.enable permission on the project to enable the service on. This permission is present in the Editor role as well as in the Service Usage Admin role.

Option C is wrong as all applications need the API to be enabled before they can use it.

Option D is wrong as the API is not enabled and it needs to be enabled.

#### 6. Question

Your team is working on designing an IoT solution. There are thousands of devices that need to send periodic time series data for processing. Which services should be used to ingest and store the data?

- A. Pub/Sub, Datastore
- B. Pub/Sub, Dataproc
- C. Dataproc, Bigtable

 D. Pub/Sub, Bigtable

Unattempted

Correct answer is D as Pub/Sub is ideal for ingestion and Bigtable for time series data storage.

Refer GCP documentation – IoT Overview

### Ingestion

Google Cloud Pub/Sub provides a globally durable message ingestion service. By creating topics for streams or channels, you can enable different components of your application to subscribe to specific streams of data without needing to construct subscriber-specific channels on each device. Cloud Pub/Sub also natively connects to other Cloud Platform services, helping you to connect ingestion, data pipelines, and storage systems.

Cloud Pub/Sub can act like a shock absorber and rate leveller for both incoming data streams and application architecture changes. Many devices have limited ability to store and retry sending telemetry data. Cloud Pub/Sub scales to handle data spikes that can occur when swarms of devices respond to events in the physical world, and buffers these spikes to help isolate them from applications monitoring the data.

### Time Series dashboards with Cloud Bigtable

Certain types of data need to be quickly sliceable along known indexes and dimensions for updating core visualizations and user interfaces. Cloud Bigtable provides a low-latency and high-throughput database for NoSQL data. Cloud Bigtable provides a good place to drive heavily used visualizations and queries, where the questions are already well understood and you need to absorb or serve at high volumes.

Compared to BigQuery, Cloud Bigtable works better for queries that act on rows or groups of consecutive rows, because Cloud Bigtable stores data by using a row-based format. Compared to Cloud Bigtable, BigQuery is a better choice for queries that require data aggregation.

Option A is wrong as Datastore is not an ideal solution for time series IoT data storage.

Options B & C are wrong as Dataproc is not an ideal ingestion service for IoT solution. Also the storage is HDFS based.

7. 7. Question

Your development team has asked you to set up an external TCP load balancer with SSL offload. Which load balancer should you use?

- A. SSL proxy
- B. HTTP load balancer
- C. TCP proxy
- D. HTTPS load balancer

**Unattempted**

Correct answer is A as SSL proxy support TCP traffic with an ability to SSL offload.

Refer GCP documentation – Choosing Load Balancer

Google Cloud SSL Proxy Load Balancing terminates user SSL (TLS) connections at the load balancing layer, then balances the connections across your instances using the SSL or TCP protocols. Cloud SSL proxy is intended for non-HTTP(S) traffic. For HTTP(S) traffic, HTTP(S) load balancing is recommended instead.

SSL Proxy Load Balancing supports both IPv4 and IPv6 addresses for client traffic. Client IPv6 requests are terminated at the load balancing layer, then proxied over IPv4 to your backends.

Options B & D are wrong as they are recommended for HTTP or HTTPS traffic only

Option C is wrong as TCP proxy does not support SSL offload.

8. Question

Your company wants to host confidential documents in Cloud Storage. Due to compliance requirements, there is a need for the data to be highly available and resilient even in case of a regional outage. Which storage classes help meet the requirement?

- A. Standard
- B. Regional
- C. Coldline
- D. Dual-Regional
- E. Multi-Regional

**Unattempted**

Correct answers are C & E as Multi-Regional and Coldline storage classes provide multi-region geo-redundant deployment, which can sustain regional failure.

Refer GCP documentation – Cloud Storage Classes

Multi-Regional Storage is geo-redundant.

The geo-redundancy of Coldline Storage data is determined by the type of location in which it is stored: Coldline Storage data stored in multi-regional locations is redundant across multiple regions, providing higher availability than Coldline Storage data stored in regional locations.

Data that is geo-redundant is stored redundantly in at least two separate geographic places separated by at least 100 miles. Objects stored in multi-regional locations are geo-redundant, regardless of their storage class.

Geo-redundancy occurs asynchronously, but all Cloud Storage data is redundant within at least one geographic place as soon as you upload it.

Geo-redundancy ensures maximum availability of your data, even in the event of large-scale disruptions, such as natural disasters. For a dual-regional location, geo-redundancy is achieved using two specific regional locations. For other multi-regional locations, geo-redundancy is achieved using any combination of data centers within the specified multi-region, which may include data centers that are not explicitly available as regional locations.

Options A & D are wrong as they do not exist

Option B is wrong as Regional storage class is not geo-redundant. Data stored in a narrow geographic region and Redundancy is across availability zones

9. Question

Your manager needs you to test out the latest version of MS-SQL on a Windows instance. You've created the VM and need to connect into the instance. What steps should you follow to connect to the instance?

- A. Generate a Windows password in the console, then use a client capable of communicating via RDP and provide the credentials.
- B. Generate a Windows password in the console, and then use the RDP button to connect in through the console.
- C. Connect in with your own RDP client using your Google Cloud username and password.
- D. From the console click the SSH button to automatically connect.

**Unattempted**

Correct answer is A as connecting to Windows instance involves installation of the RDP client. GCP does not provide RDP client and it needs to be installed. Generate Windows instance password to connect to the instance.

Refer GCP documentation – Windows Connecting to Instance

Option B is wrong as GCP Console does not have a direct RDP connectivity.

Option C is wrong as a separate windows password needs to be generate. Google Cloud username password cannot be used.

Option D is wrong as you cannot connect to Windows instance using SSH.

#### 10. 10. Question

You need to create a new development Kubernetes cluster with 3 nodes. The cluster will be named project-1-cluster. Which of the following truncated commands will create a cluster?

- A. gcloud container clusters create project-1-cluster --num-nodes 3
- B. kubectl clusters create project-1-cluster 3
- C. kubectl clusters create project-1-cluster --num-nodes 3
- D. gcloud container clusters create project-1-cluster 3

**Unattempted**

Correct answer is A as Kubernetes cluster can be created using the gcloud command only, with the cluster name and – num-nodes parameter.

Refer GCP documentation – Kubernetes Create Cluster

gcloud container clusters create my-regional-cluster – num-nodes 2 \ – region us-west1

Options B & C are wrong as kubectl cannot be used to create Kubernetes cluster.

Option D is wrong as the 3 parameter is invalid and needs to follow a parameter.

#### 11. 11. Question

Your security team wants to be able to audit network traffic inside of your network. What's the best way to ensure they have access to the data they need?

- A. Disable flow logs.
- B. Enable flow logs.
- C. Enable VPC Network logs
- D. Add a firewall capture filter.

**Unattempted**

Correct answer is B as VPC Flow logs track all the network flows and needs to be enabled.

Refer GCP documentation – VPC Flow logs

VPC Flow Logs record a sample of network flows sent from and received by VM instances. These logs can be used for network monitoring, forensics, real-time security analysis, and expense optimization.

Flow logs are aggregated by connection, at 5-second intervals, from Compute Engine VMs and exported in real time. By subscribing to Cloud Pub/Sub, you can analyze flow logs using real-time streaming APIs.

Option A is wrong as the VPC logs need to be enabled and are disabled by default.

Option C is wrong as there is no VPC Network logs.

Option D is wrong as there is no firewall capture filter.

## 12. 12. Question

While looking at your application's source code in your private Github repo, you've noticed that a service account key has been committed to git. What steps should you take next?

- A. Delete the project and create a new one.
- B. Do nothing. Git is fine for keys if the repo is private.
- C. Revoke the key, remove the key from Git, purge the Git history to remove all traces of the file, ensure the key is added to the .gitignore file.
- D. Contact Google Cloud Support

**Unattempted**

Correct answer is C as all the traces of the keys need to be removed and add the key to .gitignore file.

Option A is wrong as deleting project does not remove the keys from Git.

Option B is wrong as it is bad practice to store keys in Git, irrespective of private repo.

Option D is wrong as Google Cloud support cannot help.

## 13. 13. Question

You need to help a developer install the App Engine Go extensions. However, you've forgotten the exact name of the component. Which command could you run to show all of the available options?

- A. gcloud config list
- B. gcloud component list
- C. gcloud config components list
- D. **gcloud components list**

**Unattempted**

Correct answer is D as gcloud components list provides the list of components with the installation status.

Refer GCP documentation – Cloud SDK Components List

gcloud components list – list the status of all Cloud SDK components

This command lists all the available components in the Cloud SDK. For each component, the command lists the following information:

Status on your local workstation: not installed, installed (and up to date), and update available (installed, but not up to date)

Name of the component (a description)

ID of the component (used to refer to the component in other [gcloud components] commands)

Size of the component

In addition, if the –show-versions flag is specified, the command lists the currently installed version (if any) and the latest available version of each individual component.

Options A & C are wrong as config helps view and edit Cloud SDK properties. It does not provide components detail.

Option B is wrong as it is not a valid command.

#### 14. 14. Question

Your finance team is working with the engineering team to try and determine your spending for each service by day and month across all projects used by the billing account. What is the easiest and most flexible way to aggregate and analyze the data?

- A. Export the data for the billing account(s) involved to a JSON File; use a Cloud Function to listen for a new file in the Storage bucket; code the function to analyze the service data for the desired projects, by day and month.
- B. Export the data for the billing account(s) involved to BigQuery; then use BigQuery to analyze the service data for the desired projects, by day and month.**
- C. Export the data for the billing account(s) to File, import the files into a SQL database; and then use BigQuery to analyze the service data for the desired projects, by day and month.
- D. Use the built-in reports, which already show this data.

**Unattempted**

Correct answer is B as the billing data can be exported to BigQuery for running daily and monthly to calculate spending across services.

Refer GCP documentation – Cloud Billing Export to BigQuery

Tools for monitoring, analyzing and optimizing cost have become an important part of managing development. Billing export to BigQuery enables you to export your daily usage and cost estimates automatically throughout the day to a BigQuery dataset you specify. You can then access your billing data from BigQuery. You can also use this export method to export data to a JSON file.

Options A & C are wrong as they are not easy and flexible.

Option D is wrong as there are no built-in reports.

#### 15. Question

A company wants to deploy their application using Deployment Manager. However, they want to understand how the changes will affect before implementing the update. How can the company achieve the same?

- A. Use Deployment Manager Validate Deployment feature
- B. Use Deployment Manager Dry Run feature
- C. Use Deployment Manager Preview feature**
- D. Use Deployment Manager Snapshot feature

**Unattempted**

Correct answer is C as Deployment Manager provides the preview feature to check on what resources would be created.

Refer GCP documentation – Deployment Manager Preview

After you have written a configuration file, you can preview the configuration before you create a deployment. Previewing a configuration lets you see the resources that Deployment Manager would create but does not actually instantiate any actual resources. The Deployment Manager service previews the configuration by:

1. Expanding the full configuration, including any templates.
2. Creating a deployment and “ shell” resources.

You can preview your configuration by using the preview query parameter when making an insert() request.

```
gcloud deployment-manager deployments create example-deployment --config configuration-file.yaml \ --preview
```

#### 16. Question

Your company needs to create a new Kubernetes Cluster on Google Cloud Platform. They want the nodes to be configured for resiliency and high availability with no manual intervention. How should the Kubernetes cluster be configured?

- A. Enable auto-healing for the managed instance groups
- B. Enable auto-upgrades for the nodes
- C. **Enable auto-repairing for the nodes**
- D. Enable auto-healing for the nodes

**Unattempted**

Correct answer is C as the resiliency and high availability can be increased using the node auto-repair feature, which would allow Kubernetes engine to replace unhealthy nodes.

Refer GCP documentation – Kubernetes Auto-Repairing

GKE’s node auto-repair feature helps you keep the nodes in your cluster in a healthy, running state. When enabled, GKE makes periodic checks on the health state of each node in your cluster. If a node fails consecutive health checks over an extended time period, GKE initiates a repair process for that node.

Option A is wrong as this cannot be implemented for the Kubernetes cluster.

Option B is wrong as auto-upgrades are to upgrade the node version to the latest stable Kubernetes version.

Option D is wrong as there is no auto-healing feature.

17. 17. Question

You have created an App engine application in the development environment. The testing for the application has been successful. You want to move the application to production environment. How can you deploy the application with minimal steps?

- A. Activate the production config, perform app engine deploy
- B. Perform app engine deploy using the --project parameter
- C. Clone the app engine application to the production environment
- D. Change the project parameter in app.yaml and redeploy

**Unattempted**

Correct answer is B as the gcloud app deploy allows the – project parameter to be passed to override the project that the app engine application needs to be deployed to.

Refer GCP documentation – Cloud SDK

-project=PROJECT\_ID

The Google Cloud Platform project name to use for this invocation. If omitted, then the current project is assumed; the current project can be listed using gcloud config list –format='text(core.project)' and can be set using gcloud config set project PROJECTID. Overrides the default core/projectproperty value for this command invocation.

Option A is wrong as it is a two step process, although a valid solution

Option C is wrong as Clone of the application is possible

Option D is wrong app.yaml does not control the project it is deployed to.

18. 18. Question

Your company hosts multiple applications on Compute Engine instances. They want the instances to be resilient to any instance crashes or system termination. How would you configure the instances?

- A. Set automaticRestart availability policy to true
- B. Set automaticRestart availability policy to false
- C. Set onHostMaintenance availability policy to migrate instances
- D. Set onHostMaintenance availability policy to terminate instances

**Unattempted**

Correct answer is A as automaticRestart availability policy determines how the instance reacts to the crashes and system termination and should be set to true to restart the instance.

Refer GCP documentation – Instance Scheduling Options

A VM instance's availability policy determines how it behaves when an event occurs that requires Google to move your VM to a different host machine. For example, you can choose to keep your VM instances running while Compute Engine live migrates them to another host or you can choose to terminate your instances instead. You can update an instance's availability policy at any time to control how you want your VM instances to behave.

You can change an instance's availability policy by configuring the following two settings:

The VM instance's maintenance behavior, which determines whether the instance is live migrated or terminated when there is a maintenance event.

The instance's restart behavior, which determines whether the instance automatically restarts if it crashes or gets terminated.

The default maintenance behavior for instances is to live migrate, but you can change the behavior to terminate your instance during maintenance events instead.

Configure an instance's maintenance behavior and automatic restart setting using the onHostMaintenance and automaticRestart properties. All instances are configured with default values unless you explicitly specify otherwise.

**onHostMaintenance:** Determines the behavior when a maintenance event occurs that might cause your instance to reboot.

[Default] `migrate`, which causes Compute Engine to live migrate an instance when there is a maintenance event.

`terminate`, which terminates an instance instead of migrating it.

**automaticRestart:** Determines the behavior when an instance crashes or is terminated by the system.

[Default] `true`, so Compute Engine restarts an instance if the instance crashes or is terminated.

`false`, so Compute Engine does not restart an instance if the instance crashes or is terminated.

Option B is wrong as automaticRestart availability policy should be set to true.

Options C & D are wrong as the onHostMaintenance does not apply to crashes or system termination.

#### 19. 19. Question

Your organization requires that metrics from all applications be retained for 5 years for future analysis in possible legal proceedings. Which approach should you use?

- A. Grant the security team access to the logs in each Project
- B. Configure Stackdriver Monitoring for all Projects, and export to BigQuery
- C. Configure Stackdriver Monitoring for all Projects with the default retention policies
- D. Configure Stackdriver Monitoring for all Projects, and export to Google Cloud Storage

**Unattempted**

Correct answer is B as Stackdriver monitoring metrics can be exported to BigQuery or Google Cloud Storage. However as the need is for future analysis, BigQuery is a better option.

Refer GCP documentation – Stackdriver

Stackdriver Logging provides you with the ability to filter, search, and view logs from your cloud and open source application services. Allows you to define metrics based on log contents that are incorporated into dashboards and alerts. Enables you to export logs to BigQuery, Google Cloud Storage, and Pub/Sub.

Option A is wrong as project logs are maintained in Stackdriver and it has limited data retention capability.

Option C is wrong as Stackdriver cannot retain data for 5 year. Refer Stackdriver data retention

Option D is wrong as Google Cloud Storage does not provide analytics capability.

#### 20. 20. Question

A recent software update to a static e-commerce website running on Google Cloud has caused the website to crash for several hours. The CTO decides that all critical changes must now have a back-out/roll-back plan. The website is deployed Cloud Storage and critical changes are frequent. Which action should you take to implement the back-out/roll-back plan?

- A. Create a Nearline copy for the website static data files stored in Google Cloud Storage.
- B. Enable object versioning on the website's static data files stored in Google Cloud Storage.**
- C. Enable Google Cloud Deployment Manager (CDM) on the project, and define each change with a new CDM template.
- D. Create a snapshot of each VM prior to an update, and recover the VM from the snapshot in case of a new version failure.

**Unattempted**

Correct answers are B as this is a seamless way to ensure the last known good version of the static content is always available.

Option A is wrong as this copy process is unreliable and makes it tricky to keep things in sync, it also doesn't provide a way to rollback once a bad version of the data has been written to the copy.

Option C is wrong as this would add a great deal of overhead to the process and would cause conflicts in association between different Deployment Manager deployments which could lead to unexpected behavior if an old version is changed.

Option D is wrong as this approach doesn't scale well, there is a lot of management work involved.

## 21. Question

A user wants to install a tool on the Cloud Shell. The tool should be available across sessions. Where should the user install the tool?

- A. /bin
- B. /usr/local/bin
- C. /google/scripts
- D. ~/bin**

**Unattempted**

Correct answer is D as only HOME directory is persisted across sessions.

Refer GCP documentation – Cloud Shell

Cloud Shell provisions 5 GB of free persistent disk storage mounted as your \$HOME directory on the virtual machine instance. This storage is on a per-user basis and is available across projects. Unlike the instance itself, this storage does not time out on inactivity. All files you store in your home directory, including installed software, scripts and user configuration files like .bashrc and .vimrc,

persist between sessions. Your \$HOME directory is private to you and cannot be accessed by other users.

22. Question

Your company has hosted their critical application on Compute Engine managed instance groups. They want the instances to be configured for resiliency and high availability with no manual intervention. How should the managed instance group be configured?

- A. Enable auto-repairing for the managed instance groups
- B. Enable auto-updating for the managed instance groups
- C. Enable auto-restarts for the managed instance groups
- D. Enable auto-healing for the managed instance groups

**Unattempted**

Correct answer is D as Managed Instance Groups provide AutoHealing feature, which performs a health check and if the application is not responding the instance is automatically recreated.

Refer GCP documentation – Managed Instance Groups

Autohealing — You can also set up an autohealing policy that relies on an application-based health check, which periodically verifies that your application is responding as expected on each of the MIG's instances. If an application is not responding on an instance, that instance is automatically recreated. Checking that an application responds is more precise than simply verifying that an instance is up and running.

Managed instance groups maintain high availability of your applications by proactively keeping your instances available, which means in RUNNING state. A managed instance group will automatically recreate an instance that is not RUNNING. However, relying only on instance state may not be sufficient. You may want to recreate instances when an application freezes, crashes, or runs out of memory.

Application-based autohealing improves application availability by relying on a health checking signal that detects application-specific issues such as freezing, crashing, or overloading. If a health check determines that an application has failed on an instance, the group automatically recreates that instance.

Options A & C are wrong as these features are not available.

Option B is wrong as auto-updating helps deploy new versions of software to instances in a managed instance group. The rollout of an update happens automatically based on your specifications: you can control the speed and scope

of the update rollout in order to minimize disruptions to your application. You can optionally perform partial rollouts which allows for canary testing.

23. Question

Your company has deployed their application on managed instance groups, which is served through a network load balancer. They want to enable health checks for the instances. How do you configure the health checks?

- A. Perform the health check using HTTPS by hosting a basic web server
- B. Perform the health check using HTTP by hosting a basic web server
- C. Perform the health check using TCP
- D. Update Managed Instance groups to send a periodic ping to the network load balancer

**Unattempted**

Correct answer is B as Network Load Balancer does not support TCP health checks and hence HTTP health checks need to be performed. You can run a basic web server on each instance for health checks.

Refer GCP documentation – Network Load Balancer Health Checks

Health checks ensure that Compute Engine forwards new connections only to instances that are up and ready to receive them. Compute Engine sends health check requests to each instance at the specified frequency; once an instance exceeds its allowed number of health check failures, it is no longer considered an eligible instance for receiving new traffic. Existing connections will not be actively terminated which allows instances to shut down gracefully and to close TCP connections.

The health checker continues to query unhealthy instances, and returns an instance to the pool when the specified number of successful checks is met. If all instances are marked as UNHEALTHY, the load balancer directs new traffic to all existing instances.

Network Load Balancing relies on legacy HTTP Health checks for determining instance health. Even if your service does not use HTTP, you'll need to at least run a basic web server on each instance that the health check system can query.

Option A is wrong as the traffic is not secured, HTTPS health checks are not needed.

Option C is wrong as Network Load Balancer does not support TCP health checks.

Option D is wrong as instances do not need to send any traffic to Network Load Balancer.

#### 24. 24. Question

You need to deploy an update to an application in Google App Engine. The update is risky, but it can only be tested in a live environment. What is the best way to introduce the update to minimize risk?

- A. Deploy the application temporarily and be prepared to pull it back if needed.
- B. Warn users that a new app version may have issues and provide a way to contact you if there are problems.
- C. Deploy a new version of the application but use traffic splitting to only direct a small number of users to the new version.
- D. Create a new project with the new app version, and then redirect users to the new version.

**Unattempted**

Correct answer is C as deploying a new version without assigning it as the default version will not create downtime for the application. Using traffic splitting allows for easily redirecting a small amount of traffic to the new version and can also be quickly reverted without application downtime.

Refer GCP documentation – App Engine Splitting Traffic

Traffic migration smoothly switches request routing, gradually moving traffic from the versions currently receiving traffic to one or more versions that you specify.

Traffic splitting distributes a percentage of traffic to versions of your application. You can split traffic to move 100% of traffic to a single version or to route percentages of traffic to multiple versions. Splitting traffic to two or more versions allows you to conduct A/B testing between your versions and provides control over the pace when rolling out features.

Option A is wrong as deploying the application version as default requires moving all traffic to the new version. This could impact all users and disable the service

Option B is wrong as this is not a recommended practice and it impacts user experience.

Option D is wrong as App Engine services are intended for hosting different service logic. Using different services would require manual configuration of the consumers of services to be aware of the deployment process and manage from the consumer side who is accessing which service.

#### 25. 25. Question

Your company has reserved a monthly budget for your project. You want to be informed automatically of your project spend so that you can take action when you approach the limit. What should you do?

- A. Link a credit card with a monthly limit equal to your budget.
- B. Create a budget alert for desired percentages such as 50%, 90%, and 100% of your total monthly budget.**
- C. In App Engine Settings, set a daily budget at the rate of 1/30 of your monthly budget.
- D. In the GCP Console, configure billing export to BigQuery. Create a saved view that queries your total spend.

**Unattempted**

Correct answer is B as Budget Alerts allow you configure thresholds and if crossed alerts are automatically triggered.

Refer GCP documentation – Billing Budgets Alerts

To help you with project planning and controlling costs, you can set a budget alert. Setting a budget alert lets you track how your spend is growing toward a particular amount.

You can apply budget alerts to either a billing account or a project, and you can set the budget alert at a specific amount or match it to the previous month's spend. The alerts will be sent to billing administrators and billing account users when spending exceeds a percentage of your budget.

Option A is wrong as linked card does not alert. The charges would still increase as per the usage.

Option C is wrong as App Engine does not have budget settings.

Option D is wrong as the solution would not trigger automatic alerts and the checks would not be immediate as well.

## 26. Question

Your company plans to archive data to Cloud Storage, which would be needed only in case of any compliance issues, or Audits. What is the command for creating the storage bucket with rare access and named 'archive\_bucket' ?

- A. gsutil rm -coldline gs://archive\_bucket
- B. gsutil mb -c coldline gs://archive\_bucket**
- C. gsutil mb -c nearline gs://archive\_bucket

- D. gsutil mb gs://archive\_bucket

**Unattempted**

Correct answer is B as the data would be rarely accessed, Coldline is an ideal storage class. Also gsutil needs -c parameter to pass the class.

Refer GCP documentation – Storage Classes

Coldline – Data you expect to access infrequently (i.e., no more than once per year). Typically this is for disaster recovery, or data that is archived and may or may not be needed at some future time

Option A is wrong as rm is the wrong parameter and removes the data.

Option C is wrong as Nearline is not suited for data that needs rare access.

Option D is wrong as by default, gsutil would create a regional bucket.

## 27. 27. Question

An application that relies on Cloud SQL to read infrequently changing data is predicted to grow dramatically. How can you increase capacity for more read-only clients?

- A. Configure high availability on the master node
- B. Establish an external replica in the customer's data center
- C. Use backups so you can restore if there's an outage
- D. Configure read replicas.

**Unattempted**

Correct answer is D as read replicas can help handle the read traffic reducing the load from the primary database.

Refer GCP documentation – Cloud SQL Replication Options

Cloud SQL provides the ability to replicate a master instance to one or more read replicas. A read replica is a copy of the master that reflects changes to the master instance in almost real time.

Option A is wrong as high availability is for failover and not for performance.

Option B is wrong as external replica is not recommended for scaling as it needs to be maintained and the network established for replication.

Option C is wrong as backups are more to restore the database in case of any outage.

## 28. Question

You've been asked to add a new IAM member and grant them access to run some queries on BigQuery. Considering Google recommended best practices and the principle of least privilege, how would you assign the access?

- A. Create a custom role with roles/bigquery.dataViewer and roles/bigquery.jobUser roles; assign custom role to the users
- B. Create a custom role with roles/bigquery.dataViewer and roles/bigquery.jobUser roles; assign custom role to the group; add users to groups
- C. Assign roles/bigquery.dataViewer and roles/bigquery.jobUser roles to the users
- D. Assign roles/bigquery.dataViewer and roles/bigquery.jobUser roles to a group; add users to groups

**Unattempted**

Correct answer is D as the user would need the roles/bigquery.dataViewer and roles/bigquery.jobUser to access and query the BigQuery tables inline with the least privilege. As per google best practices it is recommended to use predefined roles and create groups to control access to multiple users with same responsibility

Refer GCP documentation – IAM Best Practices

Use Cloud IAM to apply the security principle of least privilege, so you grant only the necessary access to your resources.

We recommend collecting users with the same responsibilities into groups and assigning Cloud IAM roles to the groups rather than to individual users. For example, you can create a " data scientist" group and assign appropriate roles to enable interaction with BigQuery and Cloud Storage. When a new data scientist joins your team, you can simply add them to the group and they will inherit the defined permissions.

Options A & B are wrong as the predefined roles can be assigned directly and there is not need to create custom roles.

Option C is wrong as it is recommended to create groups instead of using individual users.

## 29. Question

You have a Cloud Storage bucket that needs to host static web assets with a dozen HTML pages, a few JavaScript files, and some CSS. How do you make the bucket public?

- A. Check the make public box on the GCP Console for the bucket
- B. gsutil iam ch allAuthenticatedUsers:objectViewer gs://bucket-name
- C. gsutil make-public gs://bucket-name
- D. **gsutil iam ch allUsers:objectViewer gs://bucket-name**

**Unattempted**

Correct answer is D as the bucket can be shared by providing the Storage Object Viewer access to allUsers.

Refer GCP documentation – Cloud Storage Sharing files

You can either make all files in your bucket publicly accessible, or you can set individual objects to be accessible through your website. Generally, making all files in your bucket accessible is easier and faster.

To make all files accessible, follow the Cloud Storage guide for making groups of objects publicly readable.

To make individual files accessible, follow the Cloud Storage guide for making individual objects publicly readable.

If you choose to control the accessibility of individual files, you can set the default object ACL for your bucket so that subsequent files uploaded to your bucket are shared by default.

Use the gsutil acl ch command, replacing [VALUES\_IN\_BRACKETS] with the appropriate values:

```
gsutil acl ch -u AllUsers:R gs://[BUCKET_NAME]/[OBJECT_NAME]
```

Option A is wrong as there is no make public option with GCP Console.

Option B is wrong as access needs to be provided to allUsers to make it public and there is no allAuthenticatedUsers option.

Option C is wrong as there is no make public option with gsutil command.

30. 30. Question

You've created a new Compute Engine instance in zone us-central1-b. When you tried to attach the GPU that your data engineers requested, you're getting an error. What is the most likely cause of the error?

- A. Your instance isn't running with the correct scopes to allow GPUs.
- B. The GPU is not supported for your OS.
- C. Your instance isn't running with the default compute engine service account.
- D. The desired GPU doesn't exist in that zone.

**Unattempted**

Correct answer is D as GPU availability varies for region to region and zone to zone. One GPU available in one region/zone is not guarantee to be available in other region/zone.

Refer GCP documentation – GPUs

Option A is wrong as access scope for compute engine does not control GPU attachment with the Compute Engine.

Option B is wrong as GPUs can be attached to any OS and machine type.

Option C is wrong as access scope for compute engine does not control GPU attachment with the Compute Engine.

### 31. Question

Your data team is working on some new machine learning models. They're generating several files per day that they want to store in a regional bucket. They mostly focus on the files from the last week. However, they want to keep all the files just to be safe and if needed, would be referred once in a month. With the fewest steps possible, what's the best way to lower the storage costs?

- A. Create a Cloud Function triggered when objects are added to a bucket. Look at the date on all the files and move it to Nearline storage if it's older than a week.
- B. Create a Cloud Function triggered when objects are added to a bucket. Look at the date on all the files and move it to Coldline storage if it's older than a week.
- C. Create a lifecycle policy to switch the objects older than a week to Coldline storage.
- D. Create a lifecycle policy to switch the objects older than a week to Nearline storage.

**Unattempted**

Correct answer is D as the files are required for a week and then would be needed for only once in a month access, Nearline storage would be an ideal storage to save cost. The transition of the object can be handled easily using Object Lifecycle Management.

Refer GCP documentation – Cloud Storage Lifecycle Management

You can assign a lifecycle management configuration to a bucket. The configuration contains a set of rules which apply to current and future objects in the bucket. When an object meets the criteria of one of the rules, Cloud Storage automatically performs a specified action on the object. Here are some example use cases:

Downgrade the storage class of objects older than 365 days to Coldline Storage.

Delete objects created before January 1, 2013.

Keep only the 3 most recent versions of each object in a bucket with versioning enabled.

Option C is wrong as the files are needed once in a month, Coldline storage would not be a cost effective option.

Options A & B are wrong as the transition can be handled easily using Object Lifecycle management.

### 32. Question

Your company wants to setup a virtual private cloud network. They want to configure a single Subnet within the VPC with maximum range of available. Which CIDR block would you choose?

- A. 0.0.0.0/0
- B. 10.0.0.0/8
- C. 172.16.0.0/12
- D. 192.168.0.0/16

**Unattempted**

Correct answer is B as you can assign a standard private CIDR blocks (192.168.0.0/16, 172.16.0.0/12, 10.0.0.0/8) to VPC and their subsets as the IP address range of a VPC.

CIDR block Number of available private IPs

192.168.0.0/16 65,532

172.16.0.0/12 1,048,572

10.0.0.0/8 16,777,212

Refer GCP documentation – VPC Subnet IP ranges

Option A is wrong as it is not an allowed RFC 1918 CIDR range allowed.

Options C & D are wrong as they provide less private IPs compared to CIDR 10.0.0.0/8

33. 33. Question

You've been tasked with getting all of your team's public SSH keys onto a specific Bastion host instance of a particular project. You've collected them all. With the fewest steps possible, what is the simplest way to get the keys deployed?

- A. Add all of the keys into a file that's formatted according to the requirements. Use the gcloud compute instances add-metadata command to upload the keys to each instance
- B. Add all of the keys into a file that's formatted according to the requirements. Use the gcloud compute project-info add-metadata command to upload the keys.
- C. Use the gcloud compute ssh command to upload all the keys
- D. Format all of the keys as needed and then, using the user interface, upload each key one at a time.

**Unattempted**

Correct answer is A as instance specific SSH keys can help provide users access to the specific bastion host. The keys can be added or removed using the instance metadata.

Refer GCP documentation – Instance level SSH keys

Instance-level public SSH keys give users access to a specific Linux instance. Users with instance-level public SSH keys can access a Linux instance even if it blocks project-wide public SSH keys.

gcloud compute instances add-metadata [INSTANCE\_NAME] --metadata-from-file ssh-keys=[LIST\_PATH]

Option B is wrong as the gcloud compute project-info provides access to all the instances within a project.

Option C is wrong as gcloud compute ssh is a thin wrapper around the ssh(1) command that takes care of authentication and the translation of the instance name into an IP address. It can be used to ssh to the instance.

Option D is wrong as there is no user interface to upload the keys.

34. 34. Question

You're migrating an on-premises application to Google Cloud. The application uses a component that requires a licensing server. The license server has the IP address 10.28.0.10. You want to deploy the application without making any changes to the code or configuration. How should you go about deploying the application?

- A. Create a subnet with a CIDR range of 10.28.0.0/28. Reserve a static internal IP address of 10.28.0.10. Assign the static address to the license server instance.
- B. Create a subnet with a CIDR range of 10.28.0.0/28. Reserve a static external IP address of 10.28.0.10. Assign the static address to the license server instance.
- C. Create a subnet with a CIDR range of 10.28.0.0/28. Reserve an ephemeral internal IP address of 10.28.0.10. Assign the static address to the license server instance.
- D. Create a subnet with a CIDR range of 10.28.0.0/28. Reserve an ephemeral external IP address of 10.28.0.10. Assign the static address to the license server instance.

**Unattempted**

Correct answer is A as the IP is internal it can be reserved using the static internal IP address, which blocks it and prevents it from getting allocated to other resource.

Refer GCP documentation – Compute Network Addresses

In Compute Engine, each VM instance can have multiple network interfaces. Each interface can have one external IP address, one primary internal IP address, and one or more secondary internal IP addresses. Forwarding rules can have external IP addresses for external load balancing or internal addresses for internal load balancing.

Static internal IPs provide the ability to reserve internal IP addresses from the private RFC 1918 IP range configured in the subnet, then assign those reserved internal addresses to resources as needed. Reserving an internal IP address takes that address out of the dynamic allocation pool and prevents it from being used for automatic allocations. Reserving static internal IP addresses requires specific IAM permissions so that only authorized users can reserve a static internal IP address.

With the ability to reserve static internal IP addresses, you can always use the same IP address for the same resource even if you have to delete and recreate the resource.

Option C is wrong as Ephemeral internal IP addresses remain attached to a VM instance only until the VM is stopped and restarted or the instance is terminated. If an instance is stopped, any ephemeral internal IP addresses assigned to the instance are released back into the network pool. When a stopped instance is started again, a new ephemeral internal IP address is assigned to the instance.

Options B & D are wrong as the IP address is RFC 1918 address and needs to be an internal static IP address.

### 35. Question

You've setup and tested several custom roles in your development project. What is the fastest way to create the same roles for your new production project?

- A. Recreate them in the new project.
- B. Use the gcloud iam copy roles command and set the destination project.
- C. In GCP console, select the roles and click the Export button.
- D. Use the gcloud iam roles copy command and set the destination project.

**Unattempted**

Correct answer is D as Cloud SDK gcloud iam roles copy can be used to copy the roles to different organization or project.

Refer GCP documentation – Cloud SDK IAM Copy Role

gcloud iam roles copy – create a role from an existing role

– dest-organization=DEST\_ORGANIZATION (The organization of the destination role)

– dest-project=DEST\_PROJECT (The project of the destination role)

### 36. Question

You have been tasked to grant access to sensitive files to external auditors for a limited time period of 4 hours only. The files should not be strictly available after 4 hours. Adhering to Google best practices, how would you efficiently share the file?

- A. Host a website on Compute Engine instance and expose the files using Public DNS and share the URL with the auditors. Bring down the instance after 4 hours.
- B. Host a website on App Engine instance and expose the files using Public DNS and share the URL with the auditors. Bring down the instance after 4 hours.
- C. Store the file in Cloud Storage. Generate a signed URL with 4 hours expiry and share it with the auditors.
- D. Store the file in Cloud Storage. Grant the allUsers access to the file share it with the auditors. Remove allUsers access after 4 hours.

#### Unattempted

Correct answer is C as the file can be stored in Cloud Storage and Signed urls can be used to quickly and securely share the files with third party.

Refer GCP documentation – Cloud Storage Signed URLs

Signed URLs provide a way to give time-limited read or write access to anyone in possession of the URL, regardless of whether they have a Google account

In some scenarios, you might not want to require your users to have a Google account in order to access Cloud Storage, but you still want to control access using your application-specific logic. The typical way to address this use case is to provide a signed URL to a user, which gives the user read, write, or delete access to that resource for a limited time. Anyone who knows the URL can access the resource until the URL expires. You specify the expiration time in the query string to be signed.

Options A & B are wrong as it is not a quick solution, but a manual effort to host, share and stop the solution.

Option D is wrong as All Users is not a secure way to share data and it would be marked public.

#### 37. Question

A member of the finance team informed you that one of the projects is using the old billing account. What steps should you take to resolve the problem?

- A. Go to the Project page; expand the Billing tile; select the Billing Account option; select the correct billing account and save.
- B. Go to the Billing page; view the list of projects; find the project in question and select Change billing account; select the correct billing account and save.
- C. Delete the project and recreate it with the correct billing account.

- D. Submit a support ticket requesting the change.

### Unattempted

Correct answer is B as for changing the billing account you have to select the project and change the billing account.

Refer GCP documentation – Change Billing Account

To change the billing account for an existing project, you must be an owner on the project and a billing administrator on the destination billing account.

To change the billing account:

1. Go to the Google Cloud Platform Console.
2. Open the console left side menu and select Billing.
3. If you have more than one billing account, you'll be prompted to select Go to linked billing account to manage the current project's billing.
4. Under Projects linked to this billing account, locate the name of the project that you want to change billing for, and then click the menu next to it.
5. Select Change billing account, then choose the desired destination billing account.
6. Click Set account.

Option A is wrong as billing account cannot be changed from Project page.

Option C is wrong as the project need not be deleted.

Option D is wrong as Google support does not handle the changes and it is users responsibility.

### 38. Question

Your billing department has asked you to help them track spending against a specific billing account. They've indicated that they prefer to use Excel to create their reports so that they don't need to learn new tools. Which export option would work best for them?

- A. BigQuery Export
- B. File Export with JSON
- C. SQL Export



#### D. File Export with CSV

**Unattempted**

Correct answer is D as Cloud Billing allows export of the billing data as flat files in CSV and JSON format. As the billing department wants to use Excel to create their reports, CSV would be a ideal option.

Refer GCP documentation – Cloud Billing Export Billing Data

To access a detailed breakdown of your charges, you can export your daily usage and cost estimates automatically to a CSV or JSON file stored in a Google Cloud Storage bucket you specify. You can then access the data via the Cloud Storage API, CLI tool, or Google Cloud Platform Console.

Usage data is labeled with the project number and resource type. You use ACLs on your Cloud Storage bucket to control who can access this data.

Options A, B, & C are wrong as they do not support Excel directly and would need conversions.

39. 39. Question

A company wants to setup a template for deploying resources. They want the provisioning to be dynamic with the specifications in configuration files. Which of the following service would be ideal for this requirement?



A. Cloud Composer



#### B. Deployment Manager



C. Cloud Scheduler



D. Cloud Deployer

**Unattempted**

Correct answer is B as Deployment Manager provide Infrastructure as a Code capability.

Refer GCP documentation – Deployment Manager

Google Cloud Deployment Manager allows you to specify all the resources needed for your application in a declarative format using yaml. You can also use Python or Jinja2 templates to parameterize the configuration and allow reuse of common deployment paradigms such as a load balanced, auto-scaled instance group. Treat your configuration as code and perform repeatable deployments.

Option A is wrong as Cloud Composer is a fully managed workflow orchestration service that empowers you to author, schedule, and monitor pipelines that span across clouds and on-premises data centers.

Option C is wrong as Cloud Scheduler is a fully managed enterprise-grade cron job scheduler. It allows you to schedule virtually any job, including batch, big data jobs, cloud infrastructure operations, and more.

Option D is wrong as Cloud Deployer is not a valid service.

40. 40. Question

Your project manager wants to delegate the responsibility to upload objects to Cloud Storage buckets to his team members. Considering the principle of least privilege, which role should you assign to the team members?

- A. roles/storage.objectAdmin
- B. roles/storage.objectViewer
- C. roles/storage.objectCreator
- D. roles/storage.admin

**Unattempted**

Correct answer is C as roles/storage.objectCreator allows users to create objects. Does not give permission to view, delete, or overwrite objects.

Refer GCP documentation – Cloud Storage IAM Roles

Options B is wrong as roles/storage.objectViewer role does not provide sufficient privileges to manage buckets.

Options A & D are wrong as it provides more privileges than required.

41. 41. Question

Your company needs to create a new Kubernetes Cluster on Google Cloud Platform. As a security requirement, they want to upgrade the nodes to the latest stable version of Kubernetes with no manual intervention. How should the Kubernetes cluster be configured?

- A. Always use the latest version while creating the cluster
- B. Enable node auto-repairing
- C. Enable node auto-upgrades
- D. Apply security patches on the nodes as they are released

**Unattempted**

Correct answer is C as the Kubernetes cluster can be configured for node auto-upgrades to update them to the latest stable version of Kubernetes.

Refer GCP documentation – Kubernetes Auto Upgrades

Node auto-upgrades help you keep the nodes in your cluster up to date with the latest stable version of Kubernetes. Auto-Updates use the same update mechanism as manual node upgrades.

Some benefits of using auto-upgrades:

Lower management overhead: You don't have to manually track and update to the latest version of Kubernetes.

Better security: Sometimes new binaries are released to fix a security issue. With auto-upgrades, GKE automatically ensures that security updates are applied and kept up to date.

Ease of use: Provides a simple way to keep your nodes up to date with the latest Kubernetes features.

Node pools with auto-upgrades enabled are automatically scheduled for upgrades when a new stable Kubernetes version becomes available. When the upgrade is performed, nodes are drained and re-created to match the current cluster master version. Modifications on the boot disk of a node VM do not persist across node re-creations. To preserve modifications across node re-creation, use a DaemonSet.

Option A is wrong as this would not take into account any latest updates.

Option B is wrong as auto repairing helps in keeping nodes healthy and does not handle upgrades.

Option D is wrong as it is a manual effort and not feasible.

## 42. Question

You have created an App engine application in the us-central region. However, you found out the network team has configured all the VPN connections in the asia-east2 region, which are not possible to move. How can you change the location efficiently?

- A. Change the region in app.yaml and redeploy
- B. From App Engine console, change the region of the application
- C. Change the region in application.xml within the application and redeploy
- D. Create a new project in the asia-east2 region and create app engine in the project

Unattempted

Correct answer is D as app engine is a regional resource, it needs to be redeployed to the different region.

Refer GCP documentation – App Engine locations

App Engine is regional, which means the infrastructure that runs your apps is located in a specific region and is managed by Google to be redundantly available across all the zones within that region.

Meeting your latency, availability, or durability requirements are primary factors for selecting the region where your apps are run. You can generally select the region nearest to your app's users but you should consider the location of the other GCP products and services that are used by your app. Using services across multiple locations can affect your app's latency as well as pricing

You cannot change an app's region after you set it.

Options A, B & C are wrong as once the region is set for the app engine it cannot be modified.

### 43. Question

Your team needs to set up a MongoDB instance as quickly as possible. You don't know how to install it and what configuration files are needed. What's the best way to get it up-and-running quickly?

- A. Use Cloud Memorystore
- B. Learn and deploy MongoDB to a Compute Engine instance.
- C. Install with Cloud Launcher Marketplace
- D. Create a Deployment Manager template and deploy it.

**Unattempted**

Correct answer is C as Cloud Launcher provides out of box deployments that are completely transparent to you and can be done in no time.

Refer GCP documentation – Cloud Launcher

GCP Marketplace offers ready-to-go development stacks, solutions, and services to accelerate development. So you spend less time installing and more time developing.

Deploy production-grade solutions in a few clicks

Single bill for all your GCP and 3rd party services

Manage solutions using Deployment Manager

Notifications when a security update is available

Direct access to partner support

Option A is wrong as Cloud Memorystore is Redis compliant and not an alternative for MongoDB

Option B is wrong as hosting on the compute engine is still a manual step and would require time.

Option D is wrong as Deployment Manager would take time to build and deploy.

#### 44. Question

Your company wants to setup Production and Test environment. They want to use different subjects and the key requirement is that the VMs must be able to communicate with each other using internal IPs no additional routes configured. How can the solution be designed?

- A. Configure a single VPC with 2 subnets having the same CIDR range hosted in the same region
- B. Configure a single VPC with 2 subnets having the different CIDR range hosted in the different region
- C. Configure 2 VPCs with 1 subnet each having the same CIDR range hosted in the same region
- D. Configure 2 VPCs with 1 subnet each having the different CIDR range hosted in the different region

**Unattempted**

Correct answer is B as the VMs need to be able to communicate using private IPs they should be hosted in the same VPC. The Subnets can be in any region, however they should have non-overlapping CIDR range.

Refer GCP documentation – VPC Intra VPC reqs

The system-generated subnet routes define the paths for sending traffic among instances within the network using internal (private) IP addresses. For one instance to be able to communicate with another, appropriate firewall rules must also be configured because every network has an implied deny firewall rule for ingress traffic.

Option A is wrong as CIDR range cannot overlap.

Options C & D are wrong as VMs in subnet in different VPC cannot communicate with each other using private IPs.

45. Question

Your company is hosting their static website on Cloud Storage. You have implemented a change to add PDF files to the website. However, when the user clicks on the PDF file link it downloads the PDF instead of opening it within the browser. What would you change to fix the issue?

- A. Set content-type as object metadata to application/octet-stream on the files
- B. Set content-type as object metadata to application/pdf on the files**
- C. Set content-type as object metadata to application/octet-stream on the bucket
- D. Set content-type as object metadata to application/pdf on the bucket

**Unattempted**

Correct answer is B as the browser needs the correct content-type to be able to interpret and render the file correctly. The content-type can be set on object metadata and should be set to application/pdf.

Refer GCP documentation – Cloud Storage Object Metadata

Content-Type

The most commonly set metadata is Content-Type (also known as MIME type), which allows browsers to render the object properly. All objects have a value specified in their Content-Type metadata, but this value does not have to match the underlying type of the object. For example, if the Content-Type is not specified by the uploader and cannot be determined, it is set to application/octet-stream or application/x-www-form-urlencoded, depending on how you uploaded the object.

Option A is wrong the content type needs to be set to application/pdf

Options C & D are wrong as the metadata should be set on the objects and not on the bucket.

46. Question

You currently are running an application on a machine type with 2 vCPUs and 4gb RAM. However, recently there have been plenty of memory problems. How to increase the memory of the application with minimal downtime?

- A. In GCP console, upgrade the memory of the Compute Engine instance
- B. Use gcloud compute instances increase-memory to increase the memory
- C. Use Live migration to move to machine type with higher memory
- D. Use Live migration to move to machine type with higher CPU

### Unattempted

Correct answer is C as Live migration would help migrate the instance to an machine-type with higher memory with minimal to no downtime.

Refer GCP documentation – Live Migration

Compute Engine offers live migration to keep your virtual machine instances running even when a host system event occurs, such as a software or hardware update. Compute Engine live migrates your running instances to another host in the same zone rather than requiring your VMs to be rebooted. This allows Google to perform maintenance that is integral to keeping infrastructure protected and reliable without interrupting any of your VMs.

Live migration keeps your instances running during:

Regular infrastructure maintenance and upgrades.

Network and power grid maintenance in the data centers.

Failed hardware such as memory, CPU, network interface cards, disks, power, and so on. This is done on a best-effort basis; if a hardware fails completely or otherwise prevents live migration, the VM crashes and restarts automatically and a hostError is logged.

Host OS and BIOS upgrades.

Security-related updates, with the need to respond quickly.

System configuration changes, including changing the size of the host root partition, for storage of the host image and packages.

Live migration does not change any attributes or properties of the VM itself. The live migration process just transfers a running VM from one host machine to another host machine within the same zone. All VM properties and attributes remain unchanged, including internal and external IP addresses, instance metadata, block storage data and volumes, OS and application state, network settings, network connections, and so on.

Options A & B are wrong as the memory cannot be increased for an instance from console or command line

Option D is wrong the live migration needs to be done to an instance type with higher CPU.

#### 47. Question

Your billing department has asked you to help them track spending against a specific billing account. They've indicated that they prefer SQL querying to create their reports so that they don't need to learn new tools. The data should be as latest as possible. Which export option would work best for them?

- A. File Export with JSON and load to Cloud SQL and provide Cloud SQL access to billing department
- B. Create a sink to BigQuery and provide BigQuery access to billing department
- C. Create a sink to Cloud SQL and provide Cloud SQL access to billing department
- D. File Export with CSV and load to Cloud SQL and provide Cloud SQL access to billing department

**Unattempted**

Correct answer is B as Billing data can be automatically exported to BigQuery and BigQuery provides the SQL interface for the billing department to query the data.

Refer GCP documentation – Cloud Billing Export BigQuery

Tools for monitoring, analyzing and optimizing cost have become an important part of managing development. Billing export to BigQuery enables you to export your daily usage and cost estimates automatically throughout the day to a BigQuery dataset you specify. You can then access your billing data from BigQuery. You can also use this export method to export data to a JSON file.

Options A & D are wrong as it would need manual exporting and loading the data to Cloud SQL.

Option C is wrong as Billing does not export to Cloud SQL

#### 48. Question

Your company hosts multiple applications on Compute Engine instances. They want the instances to be resilient to any Host maintenance activities performed on the instance. How would you configure the instances?

- A. Set automaticRestart availability policy to true

- B. Set automaticRestart availability policy to false
- C. Set onHostMaintenance availability policy to migrate instances
- D. Set onHostMaintenance availability policy to terminate instances

**Unattempted**

Correct answer is C as onHostMaintenance availability policy determines how the instance reacts to the host maintenance events.

Refer GCP documentation – Instance Scheduling Options

A VM instance's availability policy determines how it behaves when an event occurs that requires Google to move your VM to a different host machine. For example, you can choose to keep your VM instances running while Compute Engine live migrates them to another host or you can choose to terminate your instances instead. You can update an instance's availability policy at any time to control how you want your VM instances to behave.

You can change an instance's availability policy by configuring the following two settings:

The VM instance's maintenance behavior, which determines whether the instance is live migrated or terminated when there is a maintenance event.

The instance's restart behavior, which determines whether the instance automatically restarts if it crashes or gets terminated.

The default maintenance behavior for instances is to live migrate, but you can change the behavior to terminate your instance during maintenance events instead.

Configure an instance's maintenance behavior and automatic restart setting using the onHostMaintenance and automaticRestart properties. All instances are configured with default values unless you explicitly specify otherwise.

**onHostMaintenance:** Determines the behavior when a maintenance event occurs that might cause your instance to reboot.

[Default] **migrate**, which causes Compute Engine to live migrate an instance when there is a maintenance event.

**terminate**, which terminates an instance instead of migrating it.

**automaticRestart:** Determines the behavior when an instance crashes or is terminated by the system.

[Default] true, so Compute Engine restarts an instance if the instance crashes or is terminated.

false, so Compute Engine does not restart an instance if the instance crashes or is terminated.

Options A & B are wrong as automaticRestart does not apply to host maintenance event.

Option D is wrong as the onHostMaintenance needs to be set to migrate the instance as termination would lead to loss of instance.

#### 49. Question

Your company wants to try out the cloud with low risk. They want to archive approximately 100 TB of their log data to the cloud and test the analytics features available to them there, while also retaining that data as a long-term disaster recovery backup. Which two steps should they take? (Choose two answers)

- A. Load logs into Google BigQuery.
- B. Load logs into Google Cloud SQL.
- C. Import logs into Google Stackdriver.
- D. Insert logs into Google Cloud Bigtable.
- E. Upload log files into Google Cloud Storage.

Unattempted

Correct answers are A & E as Google Cloud Storage can provide long term archival option and BigQuery provides analytics capabilities.

Option B is wrong as Cloud SQL is relational database and does not support the capacity required as well as not suitable for long term archival storage.

Option C is wrong as Stackdriver is a monitoring, logging, alerting and debugging tool. It is not ideal for long term retention of data and does not provide analytics capabilities.

Option D is wrong as Bigtable is a NoSQL solution and can be used for analytics. However it is ideal for data with low latency access and is expensive.

#### 50. Question

Your company wants to reduce cost on infrequently accessed data by moving it to the cloud. The data will still be accessed approximately once a month to refresh historical charts. In addition, data older than 5 years needs to be archived for 5 years for compliance reasons. How should you store and manage the data?

- A. In Google Cloud Storage and stored in a Multi-Regional bucket. Set an Object Lifecycle Management policy to delete data older than 5 years.
- B. In Google Cloud Storage and stored in a Multi-Regional bucket. Set an Object Lifecycle Management policy to change the storage class to Coldline for data older than 5 years.
- C. In Google Cloud Storage and stored in a Nearline bucket. Set an Object Lifecycle Management policy to delete data older than 5 years.
- D. In Google Cloud Storage and stored in a Nearline bucket. Set an Object Lifecycle Management policy to change the storage class to Coldline for data older than 5 years.

**Unattempted**

Correct answer is D as the access pattern fits Nearline storage class requirements and Nearline is a more cost-effective storage approach than Multi-Regional. The object lifecycle management policy to move data to Coldline is ideal for archival.

Refer GCP documentation – Cloud Storage – Storage Classes

Options A & B are wrong as Multi-Regional storage class is not an ideal storage option with infrequent access.

Option C is wrong as the data is required for compliance it cannot be deleted and needs to be moved to the Coldline storage.

## 51. Question

Your company plans to migrate a multi-petabyte data set to the cloud. The data set must be available 24hrs a day. Your business analysts have experience only with using a SQL interface. How should you store the data to optimize it for ease of analysis?

- A. Load data into Google BigQuery.
- B. Insert data into Google Cloud SQL.
- C. Put flat files into Google Cloud Storage.
- D. Stream data into Google Cloud Datastore.

**Unattempted**

Correct answer is A as BigQuery is the only of these Google products that supports an SQL interface and a high enough SLA (99.9%) to make it readily available.

Option B is wrong as Cloud SQL cannot support multi-petabyte data. Storage limit for Cloud SQL is 10TB

Option C is wrong as Cloud Storage does not provide SQL interface.

Option D is wrong as Datastore does not provide a SQL interface and is a NoSQL solution.

52. **52. Question**

You have a Kubernetes cluster with 1 node-pool. The cluster receives a lot of traffic and needs to grow. You decide to add a node. What should you do?

- A. Use gcloud container clusters resize with the desired number of nodes.
- B. Use kubectl container clusters resize with the desired number of nodes.
- C. Edit the managed instance group of the cluster and increase the number of VMs by 1.
- D. Edit the managed instance group of the cluster and enable autoscaling.

**Unattempted**

Correct answer is A as the kubernetes cluster can be resized using the gcloud command.

Refer GCP documentation – Resizing Kubernetes Cluster

gcloud container clusters resize [CLUSTER\_NAME] –node-pool [POOL\_NAME] \ –size [SIZE]

Option B is wrong as kubernetes cluster cannot be resized using the kubectl command

Options C & D are wrong as the managed instance groups should be changed manually.

53. **53. Question**

What is the command for creating a storage bucket that has once per month access and is named ‘ archive\_bucket’ ?

- A. gsutil rm -coldline gs://archive\_bucket
- B. gsutil mb -c coldline gs://archive\_bucket
- C. gsutil mb -c nearline gs://archive\_bucket
- D. gsutil mb gs://archive\_bucket

**Unattempted**

Correct answer is C as the data needs to be accessed on monthly basis Nearline is an ideal storage class. Also gsutil needs -c parameter to pass the class.

Refer GCP documentation – Storage Classes

Nearline – Data you do not expect to access frequently (i.e., no more than once per month). Ideal for back-up and serving long-tail multimedia content.

Option A is wrong as rm is the wrong parameter and removes the data.

Option B is wrong as coldline is not suited for data that needs monthly access.

Option D is wrong as by default, gsutil would create a regional bucket.

#### 54. Question

You need to take streaming data from thousands of Internet of Things (IoT) devices, ingest it, run it through a processing pipeline, and store it for analysis. You want to run SQL queries against your data for analysis. What services in which order should you use for this task?

- A. Cloud Dataflow, Cloud Pub/Sub, BigQuery
- B. Cloud Pub/Sub, Cloud Dataflow, Cloud Dataproc
- C. Cloud Pub/Sub, Cloud Dataflow, BigQuery
- D. App Engine, Cloud Dataflow, BigQuery

**Unattempted**

Correct answer is C as the need to ingest it, transform and store the Cloud Pub/Sub, Cloud Dataflow, BigQuery is ideal stack to handle the IoT data.

Refer GCP documentation – IoT

Google Cloud Pub/Sub provides a globally durable message ingestion service. By creating topics for streams or channels, you can enable different components of your application to subscribe to specific streams of data without needing to construct subscriber-specific channels on each device. Cloud Pub/Sub also natively connects to other Cloud Platform services, helping you to connect ingestion, data pipelines, and storage systems.

Google Cloud Dataflow provides the open Apache Beam programming model as a managed service for processing data in multiple ways, including batch operations, extract-transform-load (ETL) patterns, and continuous, streaming computation. Cloud Dataflow can be particularly useful for managing the high-volume data processing pipelines required for IoT scenarios. Cloud Dataflow is also designed to integrate seamlessly with the other Cloud Platform services you choose for your pipeline.

Google BigQuery provides a fully managed data warehouse with a familiar SQL interface, so you can store your IoT data alongside any of your other enterprise analytics and logs. The performance and cost of BigQuery means you might keep your valuable data longer, instead of deleting it just to save disk space.

Sample Arch – Mobile Gaming Analysis Telemetry

Processing game client and game server events in real time

Option A is wrong as the stack is correct, however the order is not correct.

Option B is wrong as Dataproc is not an ideal tool for analysis. Cloud Dataproc is a fast, easy-to-use, fully-managed cloud service for running Apache Spark and Apache Hadoop clusters in a simpler, more cost-efficient way.

Option D is wrong as App Engine is not an ideal ingestion tool to handle IoT data.

55. Question

Your application has a large international audience and runs stateless virtual machines within a managed instance group across multiple locations. One feature of the application lets users upload files and share them with other users. Files must be available for 30 days; after that, they are removed from the system entirely. Which storage solution should you choose?

- A. A Cloud Datastore database.
- B. A multi-regional Cloud Storage bucket.
- C. Persistent SSD on virtual machine instances.
- D. A managed instance group of Filestore servers.

**Unattempted**

Correct answer is B as the key storage requirements is it being global, allow lifecycle management and sharing capability. Cloud Storage is an ideal choice as it can be configured to be multi-regional, have lifecycle management rules to auto delete the files after 30 days and share them with others.

Option A is wrong Datastore is a NoSQL solution and not ideal for unstructured data.

Option C is wrong as SSD disks are ephemeral storage option for virtual machines.

Option D is wrong as disks are regional and not ideal storage option for content that needs to be shared.

56. Question

Your company wants to track whether someone is present in a meeting room reserved for a scheduled meeting. There are 1000 meeting rooms across 5 offices on 3 continents. Each room is equipped with a motion sensor that reports its status every second. The data from the motion detector includes only a sensor ID and several different discrete items of information. Analysts will use this data, together with information about account owners and office locations. Which database type should you use?

- A. Flat file
- B. NoSQL
- C. Relational
- D. Blobstore

**Unattempted**

Correct answer is B as NoSQL like Bigtable and Datastore solution is an ideal solution to store sensor ID and several different discrete items of information. It also provides an ability to join with other data. Datastore can also be configured to store data in multi-region locations.

Refer GCP documentation – Storage Options

Option A is wrong as flat file is not an ideal storage option. It is not scalable.

Option C is wrong as relational database like Cloud SQL is not an ideal solution to store schema less data.

Option D is wrong as blob storage like Cloud Storage is not an ideal solution to store, analyze schema less data and join with other sources.

#### 57. Question

You have data stored in a Cloud Storage dataset and also in a BigQuery dataset. You need to secure the data and provide 3 different types of access levels for your Google Cloud Platform users: administrator, read/write, and read-only. You want to follow Google-recommended practices. What should you do?

- A. Create 3 custom IAM roles with appropriate policies for the access levels needed for Cloud Storage and BigQuery. Add your users to the appropriate roles.
- B. At the Organization level, add your administrator user accounts to the Owner role, add your read/write user accounts to the Editor role, and add your read-only user accounts to the Viewer role.
- C. At the Project level, add your administrator user accounts to the Owner role, add your read/write user accounts to the Editor role, and add your read-only user accounts to the Viewer role.

- D. Use the appropriate pre-defined IAM roles for each of the access levels needed for Cloud Storage and BigQuery. Add your users to those roles for each of the services.

### Unattempted

Correct answer is D as Google best practice is to use pre-defined rules over legacy primitive and custom roles. Pre-defined roles can help grant fine grained control per service.

Refer GCP documentation – IAM Overview

**Primitive roles:** The roles historically available in the Google Cloud Platform Console will continue to work. These are the Owner, Editor, and Viewer roles.

**Predefined roles:** Predefined roles are the Cloud IAM roles that give finer-grained access control than the primitive roles. For example, the predefined role Pub/Sub Publisher (roles/pubsub.publisher) provides access to only publish messages to a Cloud Pub/Sub topic.

**Custom roles:** Roles that you create to tailor permissions to the needs of your organization when predefined roles don't meet your needs.

What is the difference between primitive roles and predefined roles?

Primitive roles are the legacy Owner, Editor, and Viewer roles. IAM provides predefined roles, which enable more granular access than the primitive roles. Grant predefined roles to identities when possible, so you only give the least amount of access necessary to access your resources.

When would I use primitive roles?

Use primitive roles in the following scenarios:

When the GCP service does not provide a predefined role. See the predefined roles table for a list of all available predefined roles.

When you want to grant broader permissions for a project. This often happens when you're granting permissions in development or test environments.

When you need to allow a member to modify permissions for a project, you'll want to grant them the owner role because only owners have the permission to grant access to other users for projects.

When you work in a small team where the team members don't need granular permissions.

Option A is wrong as you should use custom roles only if predefined roles are not available.

Options B & C are wrong Google does not recommend using primitive roles which do not allow fine grained access control. Also primitive roles are applied at project or service resource levels

58. **58. Question**

You have created a Kubernetes deployment, called Deployment-A, with 3 replicas on your cluster. Another deployment, called Deployment-B, needs access to Deployment-A. You cannot expose Deployment-A outside of the cluster. What should you do?

- A. Create a Service of type NodePort for Deployment A and an Ingress Resource for that Service. Have Deployment B use the Ingress IP address.
- B. Create a Service of type LoadBalancer for Deployment A. Have Deployment B use the Service IP address.
- C. Create a Service of type LoadBalancer for Deployment A and an Ingress Resource for that Service. Have Deployment B use the Ingress IP address.
- D. Create a Service of type ClusterIP for Deployment A. Have Deployment B use the Service IP address.

**Unattempted**

Correct answer is D as this exposes the service on a cluster-internal IP address. Choosing this method makes the service reachable only from within the cluster.

Refer GCP documentation – Kubernetes Networking

Option A is wrong as this exposes Deployment A over the public internet.

Option B is wrong as LoadBalancer will expose the service publicly.

Option C is wrong as this exposes the service externally using a cloud provider's load balancer, and Ingress can work only with nodeport, not LoadBalancer.

59. **59. Question**

You want to create a new role for your colleagues that will apply to all current and future projects created in your organization. The role should have the permissions of the BigQuery Job User and Cloud Bigtable User roles. You want to follow Google's recommended practices. How should you create the new role?

- A. Use gcloud iam combine-roles --global to combine the 2 roles into a new custom role.
- B. For one of your projects, in the Google Cloud Platform Console under Roles, select both roles and combine them into a new custom role. Use gcloud iam promote-role to promote the role from a project role to an organization role.
- C. For all projects, in the Google Cloud Platform Console under Roles, select both roles and combine them into a new custom role.
- D. For your organization, in the Google Cloud Platform Console under Roles, select both roles and combine them into a new custom role.

### Unattempted

Correct answer is D as this creates a new role with the combined permissions on the organization level.

Option A is wrong as this does not create a new role.

Option B is wrong as gcloud cannot promote a role to org level.

Option C is wrong as it's recommended to define the role on the organization level. Also, the role will not be applied on new projects.

### 60. Question

Your team uses a third-party monitoring solution. They've asked you to deploy it to all nodes in your Kubernetes Engine Cluster. What's the best way to do that?

- A. Connect to each node via SSH and install the monitoring solution.
- B. Deploy the monitoring pod as a StatefulSet.
- C. Deploy the monitoring pod as a DaemonSet.
- D. Use Deployment Manager to deploy the monitoring solution.

### Unattempted

Correct answer is C as Daemon set helps deploy applications or tools that you need to run on all the nodes.

Refer GCP documentation – Kubernetes Engine Daemon Set

Like other workload objects, DaemonSets manage groups of replicated Pods. However, DaemonSets attempt to adhere to a one-Pod-per-node model, either across the entire cluster or a subset of nodes. As you add nodes to a node pool, DaemonSets automatically add Pods to the new nodes as needed.

DaemonSets use a Pod template, which contains a specification for its Pods. The Pod specification determines how each Pod should look: what applications should run inside its containers, which volumes it should mount, its labels and selectors, and more.

DaemonSets are useful for deploying ongoing background tasks that you need to run on all or certain nodes, and which do not require user intervention. Examples of such tasks include storage daemons like ceph, log collection daemons like fluentd, and node monitoring daemons like collectd.

For example, you could have DaemonSets for each type of daemon run on all of your nodes. Alternatively, you could run multiple DaemonSets for a single type of daemon, but have them use different configurations for different hardware types and resource needs.

Option A is wrong as it is not a viable option.

Option B is wrong as StatefulSet is useful for maintaining state. StatefulSets represent a set of [Pods] with unique, persistent identities and stable hostnames that GKE maintains regardless of where they are scheduled. The state information and other resilient data for any given StatefulSet Pod is maintained in persistent disk storage associated with the StatefulSet.

Option D is wrong as Deployment manager does not control Pods.

## 61. Question

You've been asked to add a new IAM member and grant her access to run some queries on BigQuery. Considering the principle of least privilege, which role should you assign?

- A. roles/bigquery.dataEditor and roles/bigquery.jobUser
- B. roles/bigquery.dataViewer and roles/bigquery.user
- C. roles/bigquery.dataViewer and roles/bigquery.jobUser
- D. roles/bigquery.dataOwner and roles/bigquery.jobUser

**Unattempted**

Correct answer is C as the user needs to only query the data, they should have access to view the dataset and query the dataset which would be provided by roles/bigquery.dataViewer and roles/bigquery.jobUser inline with the least privilege principle

Refer GCP documentation – BigQuery Access Control

Option A is wrong as roles/bigquery.dataEditor provides more than required privileges

Option B is wrong as roles/bigquery.user provides more than required privileges

Option D is wrong as roles/bigquery.dataOwner provides more than required privileges

## 62. Question

You have a managed instance group comprised of preemptible VM's. All of the VM's keep deleting and recreating themselves every minute. What is a possible cause of this behavior?

- A. Your zonal capacity is limited, causing all preemptible VM's to be shutdown to recover capacity. Try deploying your group to another zone.
- B. You have hit your instance quota for the region.
- C. Your managed instance group's VM's are toggled to only last 1 minute in preemptible settings.
- D. Your managed instance group's health check is repeatedly failing, either to a misconfigured health check or misconfigured firewall rules not allowing the health check to access the instances.

Unattempted

Correct answer is D as the instances (normal or preemptible) would be terminated and relaunched if the health check fails either due to application not configured properly or the instances firewall do not allow health check to happen.

Refer GCP documentation – Health Check concepts

GCP provides health check systems that connect to virtual machine (VM) instances on a configurable, periodic basis. Each connection attempt is called a probe. GCP records the success or failure of each probe.

Health checks and load balancers work together. Based on a configurable number of sequential successful or failed probes, GCP computes an overall health state for each VM in the load balancer. VMs that respond successfully for the configured number of times are considered healthy. VMs that fail to respond successfully for a separate number of times are unhealthy.

GCP uses the overall health state of each VM to determine its eligibility for receiving new requests. In addition to being able to configure probe frequency and health state thresholds, you can configure the criteria that define a successful probe.

## 63. Question

You write a Python script to connect to Google BigQuery from a Google Compute Engine virtual machine. The script is printing errors that it cannot connect to BigQuery. What should you do to fix the script?

- A. Install the latest BigQuery API client library for Python
- B. Run your script on a new virtual machine with the BigQuery access scope enabled**
- C. Create a new service account with BigQuery access and execute your script with that user
- D. Install the bq component for gcloud with the command gcloud components install bq.

### Unattempted

Correct answer is B as by default an instance is associated with default service account and default access scope, neither of which provides an access to BigQuery. While Service account is the recommended approach and Access scope are legacy, access scope still need to be granted to the instance for applications to access the services. So enabling only the Service Account with role would not enable the script to access BigQuery.

Refer GCP documentation – Service Account

When you set up an instance to run as a service account, you determine the level of access the service account has by the IAM roles you grant to the service account. If the service account has no IAM roles, then no API methods can be run by the service account on that instance.

Furthermore, an instance's access scopes determine the default OAuth scopes for requests made through the gcloud tool and client libraries on the instance. As a result, access scopes potentially further limit access to API methods when authenticating through OAuth. However, they do not extend to other authentication protocols like gRPC.

The best practice is to set the full cloud-platform access scope on the instance, then securely limit the service account's access using IAM roles.

Essentially:

IAM restricts access to APIs based on the IAM roles that are granted to the service account.

Access scopes potentially further limit access to API methods when authenticating through OAuth.

You must set access scopes on the instance to authorize access.

While a service account's access level is determined by the IAM roles granted to the service account, an instance's access scopes determine the default OAuth scopes for requests made through the gcloud tool and client libraries on the

instance. As a result, access scopes potentially further limit access to API methods when authenticating through OAuth.

Option A is wrong as it is an issue with connectivity to BigQuery and not a client version mismatch issue.

Option C is wrong as adding a service account would not work without having the access granted through access scope.

Option D is wrong as bq command is installed by default and not needed with the python client. It is for direct command line interaction with BigQuery.

#### 64. Question

You have created a Kubernetes engine cluster named ‘ project-1’ . You’ ve realized that you need to change the machine type for the cluster from n1-standard-1 to n1-standard-4. What is the command to make this change?

- A. Create a new node pool in the same cluster, and migrate the workload to the new pool.
- B. gcloud container clusters resize project-1 --machine-type n1-standard-4
- C. gcloud container clusters update project-1 --machine-type n1-standard-4
- D. gcloud container clusters migrate project-1 --machine-type n1-standard-4

**Unattempted**

Correct answer is A as the machine type for the cluster cannot be changed through commands. A new node pool with the updated machine type needs to be created and workload migrated to the new node pool.

Refer GCP documentation – Kubernetes Engine – Migrating Node Pools

A node pool is a subset of machines that all have the same configuration, including machine type (CPU and memory) authorization scopes. Node pools represent a subset of nodes within a cluster; a container cluster can contain one or more node pools.

When you need to change the machine profile of your Compute Engine cluster, you can create a new node pool and then migrate your workloads over to the new node pool.

To migrate your workloads without incurring downtime, you need to:

Mark the existing node pool as unschedulable.

Drain the workloads running on the existing node pool.

Delete the existing node pool.

65. **65. Question**

You need to have a backup/rollback plan in place for your application that is distributed across a large managed instance group. What is the preferred method for doing so?

- A. Use the Rolling Update feature to deploy/roll back versions with different managed instance group templates.
- B. Use the managed instance group snapshot function that is included in Compute Engine.
- C. Have each instance write critical application data to a Cloud Storage bucket.
- D. Schedule a cron job to take snapshots of each instance in the group.

**Unattempted**

Correct answer is A as rolling update helps to apply the update on a controlled number of instances to maintain high availability and ability to rollback in case of any issues.

Refer GCP documentation – Updating Managed Instance Groups

A managed instance group contains one or more virtual machine instances that are controlled using an instance template. To update instances in a managed instance group, you can make update requests to the group as a whole, using the Managed Instance Group Updater feature.

The Managed Instance Group Updater allows you to easily deploy new versions of software to instances in your managed instance groups, while controlling the speed of deployment, the level of disruption to your service, and the scope of the update. The Updater offers two primary advantages:

The rollout of an update happens automatically to your specifications, without the need for additional user input after the initial request.

You can perform partial rollouts which allows for canary testing.

By allowing new software to be deployed inside an existing managed instance group, there is no need for you to reconfigure the instance group or reconnect load balancing, autoscaling, or autohealing each time new version of software is rolled out. Without the Updater, new software versions must be deployed either by creating a new managed instance group with a new software version, requiring additional set up each time, or through a manual, user-initiated,

instance-by-instance recreate. Both of these approaches require significant manual steps throughout the process.

A rolling update is an update that is gradually applied to all instances in an instance group until all instances have been updated. You can control various aspects of a rolling update, such as how many instances can be taken offline for the update, how long to wait between updating instances, whether the update affects all or just a portion of instances, and so on.

Options B, C & D are wrong as the key for scaling is to create stateless, disposable VMs to be able scale and have seamless deployment.

# SET-12

## 1. Question

You have a project using BigQuery. You want to list all BigQuery jobs for that project. You want to set this project as the default for the bq command-line tool. What should you do?

- A. Use gcloud config set project to set the default project
- B. Use bq config set project to set the default project.
- C. Use gcloud generate config-url to generate a URL to the Google Cloud Platform Console to set the default project.
- D. Use bq generate config-url to generate a URL to the Google Cloud Platform Console to set the default project.

**Correct**

Correct answer is A as you need to use gcloud to manage the config/defaults.

Refer GCP documentation – Cloud SDK Config Set

`–project=PROJECT_ID`

The Google Cloud Platform project name to use for this invocation. If omitted, then the current project is assumed; the current project can be listed using `gcloud config list –format='text(core.project)'` and can be set using `gcloud config set project PROJECTID`. Overrides the default core/project property value for this command invocation.

Option B is wrong as the bq command-line tool assumes the gcloud configuration settings and can't be set through BigQuery.

Option C is wrong as entering this command will not achieve the desired result and will generate an error.

Option D is wrong as entering this command will not achieve the desired result and will generate an error.

## 2. Question

You're deploying an application to a Compute Engine instance, and it's going to need to make calls to read from Cloud Storage and Bigtable. You want to make sure you're following the principle of least privilege. What's the easiest way to ensure the code can authenticate to the required Google Cloud APIs?

- A. Create a new user account with the required roles. Store the credentials in Cloud Key Management Service and download them to the instance in code.

- B. Use the default Compute Engine service account and set its scopes. Let the code find the default service account using Application Default Credentials.
- C. Create a new service account and key with the required limited permissions. Set the instance to use the new service account. Edit the code to use the service account key.
- D. Register the application with the Binary Registration Service and apply the required roles.

#### Unattempted

Correct answer is C as the best practice is to use a Service Account to grant the application the required access.

Refer GCP documentation – Service Accounts

A service account is a special type of Google account that belongs to your application or a virtual machine (VM), instead of to an individual end user. Your application assumes the identity of the service account to call Google APIs, so that the users aren't directly involved.

A service account is a special type of Google account that represents a Google Cloud service identity or app rather than an individual user. Like users and groups, service accounts can be assigned IAM roles to grant access to specific resources. Service accounts authenticate with a key rather than a password. Google manages and rotates the service account keys for code running on GCP. We recommend that you use service accounts for server-to-server interactions.

Option A is wrong as it is not the recommended approach

Option B is wrong as the default Service Account does not have the required permissions.

Option D is wrong as there is Binary Registration service.

### 3. Question

You've been trying to deploy a container to Kubernetes; however, kubectl doesn't seem to be able to connect to the cluster. Of the following, what is the most likely cause and how can you fix it?

- A. The firewall rules are preventing the connection. Open up the firewall rules to allow traffic to port 1337.
- B. The kubeconfig is missing the credentials. Run the gcloud container clusters get-credentials command.
- C. The kubeconfig is missing the credentials. Run the gcloud container clusters auth login command.

- D. The firewall rules are preventing the connection. Open up the firewall rules to allow traffic to port 3682.

### Unattempted

Correct answer is B as the connection is refused, the context needs to be set using the gcloud container clusters get-credentials command

Refer GCP documentation – Kubernetes Engine Troubleshooting

kubectl commands return “ connection refused” error

Set the cluster context with the following command:

gcloud container clusters get-credentials [CLUSTER\_NAME]

If you are unsure of what to enter for CLUSTER\_NAME, use the following command to list your clusters:

gcloud container clusters list

Options A & D are wrong as only SSH access is required and it is automatically added.

Option C is wrong as auth login would be needed if the Resource was not found.

#### 4. Question

Your engineers have hardcoded the database credentials to be used by application on Kubernetes Engine. The YAML they’re using looks similar to the following:

```
apiVersion: "extensions/v1beta1"
kind: "Deployment"
metadata:
 name: "products-service"
 namespace: "default"
 labels:
 app: "products-service"
 spec:
 replicas: 3
 selector:
 matchLabels:
 app: "products-service"
 template:
 metadata:
 labels:
 app: "products-service"
 spec:
```

containers:

```
– name: “ products”
image: “ gcr.io/find-seller-app-dev/products:latest”
env:
– name: “ database_user”
value: “ admin”
– name: “ database_password”
value: “ TheB3stP@ssW0rd”
```

What is Google's recommended best practice for working with sensitive information inside of Kubernetes?

- A. Store the credentials in a ConfigMap.
- B. Mount the credentials in a volume.
- C. Use an environment variable.
- D. Store the credentials in a Secret.

**Unattempted**

Correct answer is D as the Kubernetes allows credentials to be stored in Secret, which can be used by the containers.

Refer GCP documentation – Kubernetes Secrets

Kubernetes offers the Secret resource type to store credentials inside the container cluster and use them in the applications deployed on GKE directly.

Kubernetes secret objects let you store and manage sensitive information, such as passwords, OAuth tokens, and ssh keys. Putting this information in a secret is safer and more flexible than putting it verbatim in a Pod Lifecycle definition or in a container image.

Option A is wrong as ConfigMaps bind configuration files, command-line arguments, environment variables, port numbers, and other configuration artifacts to your Pods' containers and system components at runtime. ConfigMaps allow you to separate your configurations from your Pods and components, which helps keep your workloads portable, makes their configurations easier to change and manage, and prevents hardcoding configuration data to Pod specifications.

Option B is wrong as credentials cannot be mounted in the volume.

Option C is wrong as environment variable does not secure the credentials.

5. 5. Question

A SysOps admin has configured a lifecycle rule on an object versioning disabled multi-regional bucket. Which of the following statement effect reflects the following lifecycle config?

```

{
 "rule" :
 [
 {
 "action" : {"type" : "Delete"},

 "condition" : {"age" : 30, "isLive" : false}

 },
 {
 "action" : {"type" : "SetStorageClass", "storageClass" : "COLDLINE"},

 "condition" : {"age" : 365, "matchesStorageClass" : "MULTI_REGIONAL" }

 }
]
}

```

- A. Archive objects older than 30 days and move objects to Coldline Storage after 365 days if the storage class in Multi-regional
- B. Delete objects older than 30 days and move objects to Coldline Storage after 365 days if the storage class in Multi-regional.
- C. Delete archived objects older than 30 days and move objects to Coldline Storage after 365 days if the storage class in Multi-regional.
- D. Move objects to Coldline Storage after 365 days if the storage class in Multi-regional First rule has no effect on the bucket.

**Unattempted**

Correct answer is D.

First rule will delete any object if it has a age over 30 days and is not live (not the latest version). However as the bucket is not versioning enabled it does not have any effect. Second rule will change the storage class of the live object from multi-regional to Coldline for objects with age over 365 days.

Refer GCP documentation – Object Lifecycle

The following conditions are supported for a lifecycle rule:

**Age:** This condition is satisfied when an object reaches the specified age (in days). Age is measured from the object's creation time. For example, if an object's creation time is 2019/01/10 10:00 UTC and the Age condition is 10 days, then the condition is satisfied for the object on and after 2019/01/20 10:00 UTC. This is true even if the object becomes archived through object versioning sometime after its creation.

**CreatedBefore:** This condition is satisfied when an object is created before midnight of the specified date in UTC.

**IsLive:** If the value is true, this lifecycle condition matches only live objects; if the value is false, it matches only archived objects. For the purposes of this condition, objects in non-versioned buckets are considered live.

**MatchesStorageClass:** This condition is satisfied when an object in the bucket is stored as the specified storage class. Generally, if you intend to use this condition on Multi-Regional Storage or Regional Storage objects, you should also include STANDARD and DURABLE\_REDUCED\_AVAILABILITY in the condition to ensure all objects of similar storage class are covered.

Option A is wrong as the first rule does not archive but deletes the archived objects, but does not have any impact on a versioning disabled bucket.

Option B is wrong as the first rule does not delete live objects but only archives objects.

Option C is wrong as first rule does not have any impact on a versioning disabled bucket.

6. 6. Question

A SysOps admin has configured a lifecycle rule on an object versioning enabled multi-regional bucket. Which of the following statement effect reflects the following lifecycle config?

{

“rule” :

[

{

```

 "action": {"type": "Delete"},

 "condition": {"age": 30, "isLive": false}

},

{

 "action": {"type": "SetStorageClass", "storageClass": "COLDLINE"},

 "condition": {"age": 365, "matchesStorageClass": "MULTI_REGIONAL"}}

}
]
```

- A. Archive objects older than 30 days and move objects to Coldline Storage after 365 days if the storage class is Multi-regional
- B. Delete objects older than 30 days and move objects to Coldline Storage after 365 days if the storage class is Multi-regional.
- C. Delete archived objects older than 30 days and move objects to Coldline Storage after 365 days if the storage class is Multi-regional.
- D. Move objects to Coldline Storage after 365 days if the storage class is Multi-regional First rule has no effect on the bucket.

#### Unattempted

Correct answer is C.

First rule will delete any object if it has an age over 30 days and is not live (not the latest version). Second rule will change the storage class of the live object from multi-regional to Coldline for objects with age over 365 days.

Refer GCP documentation – Object Lifecycle

The following conditions are supported for a lifecycle rule:

Age: This condition is satisfied when an object reaches the specified age (in days). Age is measured from the object's creation time. For example, if an object's creation time is 2019/01/10 10:00 UTC and the Age condition is 10 days, then the condition is satisfied for the object on and after 2019/01/20 10:00

UTC. This is true even if the object becomes archived through object versioning sometime after its creation.

**CreatedBefore:** This condition is satisfied when an object is created before midnight of the specified date in UTC.

**IsLive:** If the value is true, this lifecycle condition matches only live objects; if the value is false, it matches only archived objects. For the purposes of this condition, objects in non-versioned buckets are considered live.

**MatchesStorageClass:** This condition is satisfied when an object in the bucket is stored as the specified storage class. Generally, if you intend to use this condition on Multi-Regional Storage or Regional Storage objects, you should also include STANDARD and DURABLE\_REDUCED\_AVAILABILITY in the condition to ensure all objects of similar storage class are covered.

Option A is wrong as the first rule does not archive but deletes the archived objects.

Option B is wrong as the first rule does not delete live objects but only archives objects.

Option D is wrong as first rule applies to archived or not live objects.

## 7. 7. Question

You want to deploy an application on Cloud Run that processes messages from a Cloud Pub/Sub topic. You want to follow Google-recommended practices. What should you do?

- 1. Create a Cloud Function that uses a Cloud Pub/Sub trigger on that topic. 2. Call your application on Cloud Run from the Cloud Function for every message.
- 1. Grant the Pub/Sub Subscriber role to the service account used by Cloud Run. 2. Create a Cloud Pub/Sub subscription for that topic. 3. Make your application pull messages from that subscription.
- 1. Create a service account. 2. Give the Cloud Run Invoker role to that service account for your Cloud Run application. 3. Create a Cloud Pub/Sub subscription that uses that service account and uses your Cloud Run application as the push endpoint.
- 1. Deploy your application on Cloud Run on GKE with the connectivity set to Internal. 2. Create a Cloud Pub/Sub subscription for that topic. 3. In the same Google Kubernetes Engine cluster as your application, deploy a container that takes the messages and sends them to your application.

**Unattempted**

1. Create a Cloud Function that uses a Cloud Pub/Sub trigger on that topic.
2. Call your application on Cloud Run from the Cloud Function for every message. is not right.

Both Cloud functions and Cloud Run are serverless offerings from GCP and they are both capable of integrating with Cloud Pub/Sub. It is pointless to invoking Cloud Function from Cloud Run.

1. Grant the Pub/Sub Subscriber role to the service account used by Cloud Run.
  2. Create a Cloud Pub/Sub subscription for that topic.
  3. Make your application pull messages from that subscription. is not right.
- You need to provide Cloud Run Invoker role to that service account for your Cloud Run application.

Ref: <https://cloud.google.com/run/docs/tutorials/pubsub>

1. Deploy your application on Cloud Run on GKE with the connectivity set to Internal.
2. Create a Cloud Pub/Sub subscription for that topic.
3. In the same Google Kubernetes Engine cluster as your application, deploy a container that takes the messages and sends them to your application. is not right.

Like above, you need cloud Run Invoker role on the service account.

Ref: <https://cloud.google.com/run/docs/tutorials/pubsub>

Also, our question states the application on Cloud Run processes messages from a Cloud Pub/Sub topic; whereas in this option, we are utilizing a separate container to process messages from the topic. So this doesn't satisfy our requirements.

1. Create a service account.
2. Give the Cloud Run Invoker role to that service account for your Cloud Run application.
3. Create a Cloud Pub/Sub subscription that uses that service account and uses your Cloud Run application as the push endpoint. is the right answer.

This exact process is described in

<https://cloud.google.com/run/docs/tutorials/pubsub>

You create a service account.

```
gcloud iam service-accounts create cloud-run-pubsub-invoker \
--display-name " Cloud Run Pub/Sub Invoker"
```

You then give the invoker service account permission to invoke your service:

```
gcloud run services add-iam-policy-binding pubsub-tutorial \
--member=serviceAccount:cloud-run-pubsub-
invoker@PROJECT_ID.iam.gserviceaccount.com \
--role=roles/run.invoker
```

And finally, you create a Pub/Sub subscription with the service account:

```
gcloud pubsub subscriptions create myRunSubscription --topic myRunTopic \
--push-endpoint=SERVICE-URL/ \
--push-auth-service-account=cloud-run-pubsub-
invoker@PROJECT_ID.iam.gserviceaccount
```

8. Question

You want to enable your running Google Container Engine cluster to scale as demand for your application changes. What should you do?

- A. Add additional nodes to your Container Engine cluster using the following command: gcloud container clusters resize CLUSTER\_Name --size 10
- B. Add a tag to the instances in the cluster with the following command: gcloud compute instances add-tags INSTANCE --tags --enable-autoscaling max-nodes-10
- C. Update the existing Container Engine cluster with the following command: gcloud alpha container clusters update mycluster --enable-autoscaling --min-nodes=1 --max-nodes=10
- D. Create a new Container Engine cluster with the following command: gcloud alpha container clusters create mycluster --enable-autoscaling --min-nodes=1 --max-nodes=10 and redeploy your application

**Unattempted**

Correct answer is C as you need to update the cluster to enable auto scaling with min and max nodes to scale as per the demand.

Refer GCP documentation – Cluster Autoscaling

Option A is wrong as it would only increase the nodes.

Option B is wrong as the cluster needs to updated and not the instances.

Option D is wrong as you do not need to create a new cluster and the existing cluster can be updated to enable auto scaling.

9. 9. Question

You've set up an instance inside your new network and subnet. Your firewall rules are set to target all instances in your network with the following firewall rules.

NAME:deny-all | NETWORK:devnet | DIRECTION:INGRESS | PRIORITY:1000 | DENY:tcp:0-65535,udp:0-6553

NAME:open-ssh | NETWORK:devnet | DIRECTION:INGRESS | PRIORITY:5000 | ALLOW:tcp:22

However, when you attempt to connect to your instance via SSH, your connection is timing out. What is the most likely cause?

- A. SSH would be denied and would need instance reboot for the allow rule to take effect

- B. SSH key hasn't been uploaded to the instance.
- C. Firewall rule needs to be applied to the instance specifically.
- D. SSH would be denied as the deny rule overrides the allow

**Unattempted**

Correct answer is D as the firewall rules are applied as per the priority and as the deny rule has the higher priority as compared to the allow rule, the SSH access is denied.

Refer GCP documentation – VPC Firewall Rules – Priority

The firewall rule priority is an integer from 0 to 65535, inclusive. Lower integers indicate higher priorities. If you do not specify a priority when creating a rule, it is assigned a priority of 1000.

The relative priority of a firewall rule determines if it is applicable when evaluated against others. The evaluation logic works as follows:

The highest priority rule applicable to a target for a given type of traffic takes precedence. Target specificity does not matter. For example, a higher priority ingress rule for certain ports and protocols intended for all targets overrides a similarly defined rule for the same ports and protocols intended for specific targets.

The highest priority rule applicable for a given protocol and port definition takes precedence, even when the protocol and port definition is more general. For example, a higher priority ingress rule allowing traffic for all protocols and ports intended for given targets overrides a lower priority ingress rule denying TCP 22 for the same targets.

A rule with a deny action overrides another with an allow action only if the two rules have the same priority. Using relative priorities, it is possible to build allowrules that override deny rules, and vice versa.

Rules with the same priority and the same action have the same result. However, the rule that is used during the evaluation is indeterminate. Normally, it doesn't matter which rule is used except when you enable firewall rule logging. If you want your logs to show firewall rules being evaluated in a consistent and well-defined order, assign them unique priorities.

Option A is wrong as firewall rules are applied directly and do not require an instance restart.

Option B is wrong as SSH are autogenerated and transferred to the instance.

Option C is wrong as firewall are not applied to instance directly but through network tags.

10. 10. Question

You've set up an instance inside your new network and subnet. You create firewall rules to target all instances in your network with the following firewall rules.

NAME:open-ssh | NETWORK:devnet | DIRECTION:INGRESS | PRIORITY:1000 | ALLOW:tcp:22

NAME:deny-all | NETWORK:devnet | DIRECTION:INGRESS | PRIORITY:5000 | DENY:tcp:0-65535,udp:0-6553

If you try to SSH to the instance, what would be the result?

- A. SSH would be denied and would need gcloud firewall refresh command for the allow rule to take effect.
- B. SSH would be allowed as the allow rule overrides the deny
- C. SSH would be denied as the deny rule overrides the allow
- D. SSH would be denied and would need instance reboot for the allow rule to take effect

**Unattempted**

Correct answer is B as the firewall rules are applied as per the priority and as the allow rule has the higher priority as compared to the deny rule, the SSH access is allowed.

Refer GCP documentation – VPC Firewall Rules – Priority

The firewall rule priority is an integer from 0 to 65535, inclusive. Lower integers indicate higher priorities. If you do not specify a priority when creating a rule, it is assigned a priority of 1000.

The relative priority of a firewall rule determines if it is applicable when evaluated against others. The evaluation logic works as follows:

The highest priority rule applicable to a target for a given type of traffic takes precedence. Target specificity does not matter. For example, a higher priority ingress rule for certain ports and protocols intended for all targets overrides a similarly defined rule for the same ports and protocols intended for specific targets.

The highest priority rule applicable for a given protocol and port definition takes precedence, even when the protocol and port definition is more general. For example, a higher priority ingress rule allowing traffic for all protocols and ports intended for given targets overrides a lower priority ingress rule denying TCP 22 for the same targets.

A rule with a deny action overrides another with an allow action only if the two rules have the same priority. Using relative priorities, it is possible to build allowrules that override deny rules, and vice versa.

Rules with the same priority and the same action have the same result. However, the rule that is used during the evaluation is indeterminate. Normally, it doesn't matter which rule is used except when you enable firewall rule logging. If you want your logs to show firewall rules being evaluated in a consistent and well-defined order, assign them unique priorities.

Options A, C & D are wrong the SSH access would be allowed.

## 11. Question

Your company has a number of internal backends that they do not want to be exposed to the public Internet. How can they reduce their external exposure while still allowing maintenance access to resources when working remotely?

- A. Remove the external IP address and use Cloud Shell to access internal-only resources
- B. Remove the external IP address and use a bastion host to access internal-only resources.**
- C. Remove the external IP address and have remote employees dial into the company VPN connection for maintenance work.
- D. Hide the external IP address behind a load balancer and restrict load balancer access to the internal company network.

**Unattempted**

Correct answer is B as it is a best practice to remove external ip address from the instances so that they are not reachable from the internet and have a Bastion host or Jump server to be able to login into the servers.

Refer GCP documentation – Bastion Hosts

Bastion hosts provide an external facing point of entry into a network containing private network instances. This host can provide a single point of fortification or audit and can be started and stopped to enable or disable inbound SSH communication from the Internet.

By using a bastion host, you can connect to an instance that does not have an external IP address. This approach allows you to connect to a development environment or manage the database instance for your external application, for example, without configuring additional firewall rules.

A complete hardening of a bastion host is outside the scope of this article, but some initial steps taken can include:

Limit the CIDR range of source IPs that can communicate with the bastion.

Configure firewall rules to allow SSH traffic to private instances from only the bastion host.

By default, SSH on instances is configured to use private keys for authentication. When using a bastion host, you log into the bastion host first, and then into your target private instance. Because of this two-step login, which is why bastion hosts are sometimes called “jump servers,” you should use ssh-agent forwarding instead of storing the target machine’s private key on the bastion host as a way of reaching the target machine. You need to do this even if using the same key-pair for both bastion and target instances, as the bastion has direct access to only the public half of the key-pair.

## 12. Question

The development team has provided you with a Kubernetes Deployment file. You have no infrastructure yet and need to deploy the application. What should you do?

- A. Use gcloud to create a Kubernetes cluster. Use Deployment Manager to create the deployment.
- B. Use gcloud to create a Kubernetes cluster. Use kubectl to create the deployment.
- C. Use kubectl to create a Kubernetes cluster. Use Deployment Manager to create the deployment.
- D. Use kubectl to create a Kubernetes cluster. Use kubectl to create the deployment.

**Unattempted**

Correct answer is B as you would need gcloud to create a kubernetes cluster. Once the cluster is created you can use kubectl to manage the deployments.

Refer GCP documentation – Kubernetes Cluster Tutorial

To create a cluster with the gcloud command-line tool, use the gcloud container clusters command:

```
gcloud container clusters create hello-cluster --num-nodes=3
```

To deploy and manage applications on a GKE cluster, you must communicate with the Kubernetes cluster management system. You typically do this by using the kubectl command-line tool.

Kubernetes represents applications as Pods, which are units that represent a container (or group of tightly-coupled containers). The Pod is the smallest

deployable unit in Kubernetes. In this tutorial, each Pod contains only your hello-app container.

The kubectl run command below causes Kubernetes to create a Deployment named hello-web on your cluster. The Deployment manages multiple copies of your application, called replicas, and schedules them to run on the individual nodes in your cluster. In this case, the Deployment will be running only one Pod of your application.

```
kubectl run hello-web – image=gcr.io/${PROJECT_ID}/hello-app:v1 – port 8080
```

Options A & C are wrong as you need kubectl to do a kubernetes deployment.

Options C & D are wrong as you need gcloud to create the kubernetes cluster.

### 13. 13. Question

One of the microservices in your application has an intermittent performance problem. You have not observed the problem when it occurs but when it does, it triggers a particular burst of log lines. You want to debug a machine while the problem is occurring. What should you do?

- A. Log into one of the machines running the microservice and wait for the log storm.
- B. In the Stackdriver Error Reporting dashboard, look for a pattern in the times the problem occurs.
- C. Configure your microservice to send traces to Stackdriver Trace so you can find what is taking so long.
- D. Set up a log metric in Stackdriver Logging, and then set up an alert to notify you when the number of log lines increases past a threshold.

**Unattempted**

Correct answer is D as there is a burst of log lines you can set up a metric that identifies those lines. Stackdriver will also allow you to set up a text, email or messaging alert that can notify promptly when the error is detected so you can hop onto the system to debug.

Option A is wrong as logging into an individual machine may not see the specific performance problem as multiple machines may be in the configuration and reducing the chances of interacting with an intermittent performance problem.

Option B is wrong as error reporting won't necessarily catch the log lines unless they are stack traces in the proper format. Additionally just because there is a pattern doesn't mean you will know exactly when and where to log in to debug.

Option C is wrong as trace may tell you where time is being spent but won't let you hone in on the exact host that the problem is occurring on because you generally only send samples of traces. There is also no alerting on traces to notify exactly when the problem is happening.

14. 14. Question

You're writing a Java application with lots of threading and concurrency. You want your application to run in a sandboxed managed environment with the ability to perform SSH debugging to check on any thread dump for troubleshooting. Which service should you host your application on?

- A. Compute Engine
- B. App Engine Flexible Environment
- C. Cloud Functions
- D. App Engine Standard Environment

**Unattempted**

Correct answer is B as App Engine provides the managed service and Flexible environment supports the ability to perform SSH debugging.

Refer GCP documentation – App Engine Environments

Feature – SSH-debugging

Standard environment – No

Flexible environment – Yes

Flexible environment instances are permitted to have higher CPU and memory limits than is possible with standard environment instances. This allows flexible instances to run applications that are more memory and CPU intensive. However, it may increase the likelihood of concurrency bugs due to the increase in threads within a single instance.

Developers can SSH to a flexible environment instance and obtain a thread dump to troubleshoot this type of problem.

Option A is wrong as Compute Engine does not provide managed service

Option C is wrong as Cloud Functions provides serverless event driven compute platform.

Option D is wrong as App Engine Standard environment does not provide SSH debugging

### 15. 15. Question

Several employees at your company have been creating projects with Cloud Platform and paying for it with their personal credit cards, which the company reimburses. The company wants to centralize all these projects under a single, new billing account. What should you do?

- A. Contact cloud-billing@google.com with your bank account details and request a corporate billing account for your company.
- B. Create a ticket with Google Support and wait for their call to share your credit card details over the phone.
- C. In the Google Platform Console, go to the Resource Manager and move all projects to the root Organization.
- D. In the Google Cloud Platform Console, create a new billing account and set up a payment method.

**Unattempted**

Correct answer is C as Google Cloud Resource Manager can help group the existing accounts under an Organization for centralized billing.

Refer GCP documentation – Resource Manager

Google Cloud Platform (GCP) customers need an easy way to centrally manage and control GCP resources, projects and billing accounts that belong to their organization. As companies grow, it becomes progressively difficult to keep track of an ever-increasing number of projects, created by multiple users, with different access control policies and linked to a variety of billing instruments. Google Cloud Resource Manager allows you to group resource containers under the Organization resource, providing full visibility, centralized ownership and unified management of your company's assets on GCP.

Options A & B are wrong as billing consolidation is User responsibility and GCP does not support it.

Option D is wrong as it would not centralize the billing under a single account.

### 16. 16. Question

A company is hosting their Echo application on Google Cloud using Google Kubernetes Engine. The application is deployed with deployment echo-deployment exposed with echo-service. They have a new image that needs to be deployed for the application. How can the change be deployed with minimal downtime?

- A. Update image using kubectl set image deployment
- B. Delete the deployment and create a new deployment with the updated image

- C. Delete the service and create a new service with the updated image
- D. Update image in instance template and use rolling deployment of instance group with Kubernetes engine.

**Unattempted**

Correct answer is A as the image can be directly updated using the kubectl command and Kubernetes Engine performs a rolling update.

Refer GCP documentation – Kubernetes Engine Rolling updates

You can perform a rolling update to update the images, configuration, labels, annotations, and resource limits/requests of the workloads in your clusters. Rolling updates incrementally replace your resource's Pods with new ones, which are then scheduled on nodes with available resources. Rolling updates are designed to update your workloads without downtime.

You can use kubectl set to make changes to an object's image, resources (compute resource such as CPU and memory), or selector fields.

For example, to update a Deployment from nginx version 1.7.9 to 1.9.1, run the following command:

```
kubectl set image deployment nginx nginx=nginx:1.9.1
```

The kubectl set image command updates the nginx image of the Deployment's Pods one at a time.

Option B is wrong as creating a new deployment would result in downtime.

Option C is wrong as service does not have a mapping of image.

Option D is wrong as Kubernetes Engine does not work with instance template and managed instance groups

**17. 17. Question**

A member of the finance team informed you that one of the projects is using the old billing account. What steps should you take to resolve the problem?

- A. Go to the Project page; expand the Billing tile; select the Billing Account option; select the correct billing account and save.
- B. Go to the Billing page; view the list of projects; find the project in question and select Change billing account; select the correct billing account and save.

- C. Delete the project and recreate it with the correct billing account.
- D. Submit a support ticket requesting the change.

### Unattempted

Correct answer is B as the billing account for the project can be modified from the Billing page.

Refer GCP documentation – Billing Modify Project

If you are a billing administrator on only one billing account, new projects you create are automatically linked to your existing billing account. If you create and have access to multiple billing accounts, you can change the billing account a project is billed to. This article describes how to change the billing account for your project, as well as how to enable and disable billing for a project.

To change the billing account:

Sign in to the Google Cloud Platform Console.

Open the console navigation menu (menu) and select Billing.

If you have more than one billing account, you'll be prompted to select Go to linked billing account to manage the current project's billing.

From the Billing navigation menu, click Account management.

Under Projects linked to this billing account, locate the name of the project that you want to change billing for, and then click the menu (more\_vert) next to it.

Select Change billing, then choose the desired destination billing account.

### 18. Question

A company uses Cloud Storage for storing their critical data. As a part of compliance, the objects need to be encrypted using customer-supplied encryption keys. How should the object be handled to support customer-supplied encryption?

- A. Use gsutil with —encryption-key to pass the encryption key
- B. Use gsutil with  
**GSUtil:encryption\_key=[YOUR\_ENCRYPTION\_KEY] to pass the encryption key**
- C. Use gcloud config to define the encryption
- D. Create bucket with —encryption-key and use gsutil to upload files

### Unattempted

Correct answer is B as the customer supplied encryption key can be passed using the encryption\_key parameter.

Refer GCP documentation – Cloud Storage Encryption

Add the following option to the [GSUtil] section of your boto configuration file:

encryption\_key = [YOUR\_ENCRYPTION\_KEY]

where [YOUR\_ENCRYPTION\_KEY] is the key for encrypting the uploaded file.

Note: You can alternatively include this information in each gsutil command by using the -o top level flag: -o  
“ GSUtil:encryption\_key=[YOUR\_ENCRYPTION\_KEY]” .

Option A is wrong as the parameter is wrong. Parameter -o  
“ GSUtil:encryption\_key=[YOUR\_ENCRYPTION\_KEY]” can be used.

Option C is wrong as encryption key cannot be defined using gcloud config.

Option D is wrong as encryption is not set on bucket and needs to applied when the object is uploaded.

### 19. Question

The development team needs a regional MySQL database with point-in-time recovery for a new proof-of-concept application. What's the most inexpensive way to enable point-in-time recovery?

- A. Replicate to a Cloud Spanner database.
- B. Create a read replica in the same region.
- C. Enable binary logging.
- D. Create hourly back-ups.

### Unattempted

Correct answer is C as binary logging helps Point-in-time recovery.

Refer GCP documentation – Cloud SQL MySQL Point In Time Recovery

Point-in-time recovery enables you to recover an instance to a specific point in time. A point-in-time recovery always creates a new instance; you cannot perform a point-in-time recovery to an existing instance.

Options A & B are wrong as Read Replica and Cloud Spanner are not cost-effective options.

Option D is wrong as hourly back-ups does not meet the point-in-time requirement.

20. 20. Question

Your application deployed on a Google Compute Engine virtual machine instance needs to connect to Google Cloud Pub/Sub. What is the best way to provision the access to the application?

- A. Whitelist Google Compute Engine virtual machine instance IP on the Cloud Pub/Sub firewall
- B. Build or leverage an OAuth-compatible access control system
- C. Create a new service account with no access and enable access scope to allow Cloud Pub/Sub access for the instance
- D. Create a new service account with Cloud Pub/Sub access and associate with the instance

Unattempted

Correct answer is D as the VM needs to be granted permissions using the service account to be able to communicate with Cloud Pub/Sub.

Refer GCP documentation – Service Account Permissions

When you set up an instance to run as a service account, the level of access the service account has is determined by the combination of access scopes granted to the instance and IAM roles granted to the service account. You need to configure both access scopes and IAM roles to successfully set up an instance to run as a service account. Essentially:

Access scopes authorize the potential access that an instance has to API methods.

IAM restricts that access to the roles granted to the service account.

Option A is wrong as there is feature to whitelist IPs as firewalls only apply to Compute Engines.

Option B is wrong as service accounts handle the OAuth authentication and it is best suited for service to service communication.

Google OAuth 2.0 system supports server-to-server interactions such as those between a web application and a Google service. For this scenario you need a service account, which is an account that belongs to your application instead of to

an individual end user. Your application calls Google APIs on behalf of the service account, so users aren't directly involved. This scenario is sometimes called "two-legged OAuth," or "2LO." (The related term "three-legged OAuth" refers to scenarios in which your application calls Google APIs on behalf of end users, and in which user consent is sometimes required.)

Typically, an application uses a service account when the application uses Google APIs to work with its own data rather than a user's data. For example, an application that uses Google Cloud Datastore for data persistence would use a service account to authenticate its calls to the Google Cloud Datastore API.

Option C is wrong as access scope and service account IAM role permissions are applied together with the more restrictive taking effect. With no permissions to the Service Account, the instance would not be able to communicate with the VM instance.

## 21. Question

Your company pushes batches of sensitive transaction data from its application server VMs to Cloud Pub/Sub for processing and storage. What is the Google-recommended way for your application to authenticate to the required Google Cloud services?

- A. Ensure that VM service accounts are granted the appropriate Cloud Pub/Sub IAM roles.
- B. Ensure that VM service accounts do not have access to Cloud Pub/Sub, and use VM access scopes to grant the appropriate Cloud Pub/Sub IAM roles.
- C. Generate an OAuth2 access token for accessing Cloud Pub/Sub, encrypt it, and store it in Cloud Storage for access from each VM.
- D. Create a gateway to Cloud Pub/Sub using a Cloud Function, and grant the Cloud Function service account the appropriate Cloud Pub/Sub IAM roles.

### Unattempted

Correct answer is A as the VM needs to be granted permissions using the service account to be able to communicate with Cloud Pub/Sub.

Refer GCP documentation – Service Account Permissions

When you set up an instance to run as a service account, the level of access the service account has is determined by the combination of access scopes granted to the instance and IAM roles granted to the service account. You need to configure both access scopes and IAM roles to successfully set up an instance to run as a service account. Essentially:

Access scopes authorize the potential access that an instance has to API methods.

IAM restricts that access to the roles granted to the service account.

Google OAuth 2.0 system supports server-to-server interactions such as those between a web application and a Google service. For this scenario you need a service account, which is an account that belongs to your application instead of to an individual end user. Your application calls Google APIs on behalf of the service account, so users aren't directly involved. This scenario is sometimes called "two-legged OAuth," or "2LO." (The related term "three-legged OAuth" refers to scenarios in which your application calls Google APIs on behalf of end users, and in which user consent is sometimes required.)

Typically, an application uses a service account when the application uses Google APIs to work with its own data rather than a user's data. For example, an application that uses Google Cloud Datastore for data persistence would use a service account to authenticate its calls to the Google Cloud Datastore API.

Option B is wrong as access scope and service account IAM role permissions are applied together with the more restrictive taking effect. With no permissions to the Service Account, the instance would not be able to communicate with the VM instance.

Option C is wrong as service accounts handle the OAuth authentication and it is best suited for service to service communication.

Option D is wrong as there is no need for the gateway. Also, the VM and Cloud Function access needs to be handled.

## 22. Question

You can SSH into an instance from another instance in the same VPC by its internal IP address, but not its external IP address. What is one possible reason why this is so?

- A. The outgoing instance does not have correct permission granted to its service account.
- B. The external IP address is disabled.
- C. The firewall rule to allow SSH is restricted to the internal VPC.
- D. The receiving instance has an ephemeral address instead of a reserved address.

**Unattempted**

Correct answer is C as firewall rules need to be enabled for within the network and external network to be allowed to ssh into the instances.

Refer GCP documentation – VPC Firewalls

Google Cloud Platform (GCP) firewall rules let you allow or deny traffic to and from your virtual machine (VM) instances based on a configuration you specify. GCP firewall rules are applied at the virtual networking level, so they provide effective protection and traffic control regardless of the operating system your instances use.

Every VPC network functions as a distributed firewall. While firewall rules are defined at the network level, connections are allowed or denied on a per-instance basis. You can think of the GCP firewall rules as existing not only between your instances and other networks, but between individual instances within the same network

Source IP ranges: You can specify ranges of IP addresses as sources for packets. The ranges can include addresses inside your VPC network and those outside of it. Source IP ranges can be used to define sources both inside and outside of GCP.

### 23. Question

You have an application deployed on Kubernetes Engine using a Deployment named echo-deployment. The deployment is exposed using a Service called echo-service. You need to perform an update to the application with minimal downtime to the application. What should you do?

- A. Use the rolling update functionality of the Instance Group behind the Kubernetes cluster
- B. Update the deployment yaml file with the new container image. Use kubectl delete deployment/echo-deployment and kubectl create –f
- C. Use kubectl set image deployment/echo-deployment
- D. Update the service yaml file which the new container image. Use kubectl delete service/echoservice and kubectl create –f

**Unattempted**

Correct answer is C as the image can be directly updated using the kubectl command and Kubernetes Engine performs a rolling update.

Refer GCP documentation – Kubernetes Engine Rolling updates

You can perform a rolling update to update the images, configuration, labels, annotations, and resource limits/requests of the workloads in your clusters. Rolling updates incrementally replace your resource's Pods with new ones, which are then scheduled on nodes with available resources. Rolling updates are designed to update your workloads without downtime.

You can use kubectl set to make changes to an object's image, resources (compute resource such as CPU and memory), or selector fields.

For example, to update a Deployment from nginx version 1.7.9 to 1.9.1, run the following command:

```
kubectl set image deployment nginx nginx=nginx:1.9.1
```

The kubectl set image command updates the nginx image of the Deployment's Pods one at a time.

Option A is wrong as you do not work with underlying managed instance groups. It is managed by Kubernetes.

Option B is wrong as creating a new deployment would result in downtime.

Option D is wrong as service does not have a mapping of image.

#### 24. Question

You have created an App engine application in the us-central region. However, you found out the network team has configured all the VPN connections in the asia-east2 region, which are not possible to move. How can you change the location efficiently?

- A. Change the region in app.yaml and redeploy
- B. From App Engine console, change the region of the application
- C. Change the region in application.xml within the application and redeploy
- D. Create a new project in the asia-east2 region and create app engine in the project

Unattempted

Correct answer is D as app engine is a regional resource, it needs to be redeployed to the different region.

Refer GCP documentation – App Engine locations

App Engine is regional, which means the infrastructure that runs your apps is located in a specific region and is managed by Google to be redundantly available across all the zones within that region.

Meeting your latency, availability, or durability requirements are primary factors for selecting the region where your apps are run. You can generally select the region nearest to your app's users but you should consider the location of the other GCP products and services that are used by your app. Using services across multiple locations can affect your app's latency as well as pricing

You cannot change an app's region after you set it.

Options A, B & C are wrong as once the region is set for the app engine it cannot be modified.

25. 25. Question

You want to verify the IAM users and roles assigned within a GCP project named my-project. What should you do?

- A. Run gcloud iam roles list. Review the output section.
- B. Run gcloud iam service-accounts list. Review the output section.
- C. Navigate to the project and then to the IAM section in the GCP Console. Review the members and roles.
- D. Navigate to the project and then to the Roles section in the GCP Console. Review the roles and status.

**Unattempted**

Correct answer is C as IAM section provides the list of both Members and Roles.

Option A is wrong as it would provide information about the roles only.

Option B is wrong as it would provide only the service accounts.

Option D is wrong as it would provide information about the roles only.

26. 26. Question

You have a Linux VM that must connect to Cloud SQL. You created a service account with the appropriate access rights. You want to make sure that the VM uses this service account instead of the default Compute Engine service account. What should you do?

- A. When creating the VM via the web console, specify the service account under the 'Identity and API Access' section.
- B. Download a JSON Private Key for the service account. On the Project Metadata, add that JSON as the value for the key compute-engine-service-account.
- C. Download a JSON Private Key for the service account. On the Custom Metadata of the VM, add that JSON as the value for the key compute-engine-service-account.
- D. Download a JSON Private Key for the service account. After creating the VM, ssh into the VM and save the JSON under ~/gcloud/compute-engine-service-account.json.

**Unattempted**

Correct answer is A as the service account can be specified to replace the default service account when the VM is created.

Refer GCP documentation – Compute Enable Service Accounts for Instances

After creating a new service account, you can create new virtual machine instances to run as the service account.

You can enable multiple virtual machine instances to use the same service account, but a virtual machine instance can only have one service account identity. If you assign the same service account to multiple virtual machine instances, any subsequent changes you make to the service account will affect instances using the service account. This includes any changes you make to the IAM roles granted to the service account. For example, if you remove a role, all instances using the service account will lose permissions granted by that role.

You can set up a new instance to run as a service account through the Google Cloud Platform Console, the gcloud command-line tool, or directly through the API.

In the GCP Console, go to the VM Instances page.[GO TO THE VM INSTANCES PAGE](#)

Click Create instance.

On the Create a new instance page, fill in the properties for your instance.

In the Identity and API Access section, choose the service account you want to use from the dropdown list.

Click Create to create the instance.

Options B, C & D are wrong as the approaches would not work and replace the default service account.

## 27. Question

You have an instance group that you want to load balance. You want the load balancer to terminate the client SSL session. The instance group is used to serve a public web application over HTTPS. You want to follow Google-recommended practices. What should you do?

- A. Configure an HTTP(S) load balancer.
- B. Configure an internal TCP load balancer.
- C. Configure an external SSL proxy load balancer.
- D. Configure an external TCP proxy load balancer.

Unattempted

Correct answer is A as HTTPS load balancer supports the HTTPS traffic with the SSL termination ability.

Refer GCP documentation – Choosing Load Balancer

An HTTPS load balancer has the same basic structure as an HTTP load balancer (described above), but differs in the following ways:

An HTTPS load balancer uses a target HTTPS proxy instead of a target HTTP proxy.

An HTTPS load balancer requires at least one signed SSL certificate installed on the target HTTPS proxy for the load balancer. You can use Google-managed or self-managed SSL certificates.

The client SSL session terminates at the load balancer.

HTTPS load balancers support the QUIC transport layer protocol.

Option B is wrong as internal TCP load balancer does not serve external public traffic.

Option C is wrong as SSL proxy is not recommended for HTTPS traffic.

Google Cloud SSL Proxy Load Balancing terminates user SSL (TLS) connections at the load balancing layer, then balances the connections across your instances using the SSL or TCP protocols. Cloud SSL proxy is intended for non-HTTP(S) traffic. For HTTP(S) traffic, HTTP(S) load balancing is recommended instead.

SSL Proxy Load Balancing supports both IPv4 and IPv6 addresses for client traffic. Client IPv6 requests are terminated at the load balancing layer, then proxied over IPv4 to your backends.

Option D is wrong as TCP proxy does not support SSL offload and not recommended for HTTP/S traffic.

## 28. Question

You need to set up a policy so that videos stored in a specific Cloud Storage Regional bucket are moved to Coldline after 90 days, and then deleted after one year from their creation. How should you set up the policy?

- A. Use Cloud Storage Object Lifecycle Management using Age conditions with SetStorageClass and Delete actions. Set the SetStorageClass action to 90 days and the Delete action to 275 days (365 – 90)

- B. Use Cloud Storage Object Lifecycle Management using Age conditions with SetStorageClass and Delete actions. Set the SetStorageClass action to 90 days and the Delete action to 365 days.**
- C. Use gsutil rewrite and set the Delete action to 275 days (365-90).
- D. Use gsutil rewrite and set the Delete action to 365 days.

**Unattempted**

Correct answer is B as there are 2 actions needed. First archival after 90 days, which can be done by SetStorageClass action to Coldline. Second delete the data after a year, which can be done by delete action with Age 365 days. The Age condition is measured from the object's creation time.

Refer GCP documentation – Cloud Storage Lifecycle Management

**Age:** This condition is satisfied when an object reaches the specified age (in days). Age is measured from the object's creation time. For example, if an object's creation time is 2019/01/10 10:00 UTC and the Age condition is 10 days, then the condition is satisfied for the object on and after 2019/01/20 10:00 UTC. This is true even if the object becomes archived through object versioning sometime after its creation.

Option A is wrong as the Age needs to be set to 365 as its relative to the object creation date and not changed date.

Options C & D are wrong as gsutil rewrite can be used to change the storage class. However, it needs to be triggered and the solution does not handle archival of data.

## 29. Question

You have 32 GB of data in a single file that you need to upload to a Nearline Storage bucket. The WAN connection you are using is rated at 1 Gbps, and you are the only one on the connection. You want to use as much of the rated 1 Gbps as possible to transfer the file rapidly. How should you upload the file?

- A. Use the GCP Console to transfer the file instead of gsutil.
- B. Enable parallel composite uploads using gsutil on the file transfer.**
- C. Decrease the TCP window size on the machine initiating the transfer.
- D. ?Change the storage class of the bucket from Nearline to Multi-Regional.

**Unattempted**

Correct answer is B as the bandwidth is good and its a single file, gsutil parallel composite uploads can be used to split the large file and upload in parallel.

Refer GCP documentation – Transferring Data to GCP & Storage Composite Objects

To support parallel uploads and limited append/edit functionality, Cloud Storage allows users to compose up to 32 existing objects into a new object without transferring additional object data.

Object composition can be used for uploading an object in parallel: simply divide your data into multiple chunks, upload each chunk to a distinct object in parallel, compose your final object, and delete any temporary objects.

gsutil tool is an open-source command-line utility available for Windows, Linux, and Mac.

Multi-threaded/processed: Useful when transferring large number of files.

Parallel composite uploads: Splits large files, transfers chunks in parallel, and composes at destination.

Retry: Applies to transient network failures and HTTP/429 and 5xx error codes.

Resumability: Resumes the transfer after an error.

Option A is wrong as it is not recommended for large files and it would do a sequential upload of a single file.

Options C & D are wrong as they would not help in improving the performance.

### 30. Question

You need to monitor resources that are distributed over different projects in Google Cloud Platform. You want to consolidate reporting under the same Stackdriver Monitoring dashboard. What should you do?

- A. Use Shared VPC to connect all projects, and link Stackdriver to one of the projects.
- B. For each project, create a Stackdriver account. In each project, create a service account for that project and grant it the role of Stackdriver Account Editor in all other projects.
- C. Configure a single Stackdriver account, and link all projects to the same account.
- D. Configure a single Stackdriver account for one of the projects. In Stackdriver, create a Group and add the other project names as criteria for that Group.

Unattempted

Correct answer is C as you can create a single Stackdriver account and add multiple projects to the same account.

Refer GCP documentation – Stackdriver Monitoring

A single Workspace can monitor any number of GCP projects or AWS accounts. The best-practice recommendation to create a multi-project Workspace is as follows:

Create a new GCP project. For instructions on creating a new GCP project, go to Before you begin.

Create a new Workspace for that project. For detailed steps, go to Creating a single-project Workspace.

Add GCP projects or AWS accounts to the Workspace. For details, go to Adding monitored projects.

Option A is wrong as Shared VPC would not allow consolidation of multiple project monitoring. Shared VPC allows an organization to connect resources from multiple projects to a common VPC network, so that they can communicate with each other securely and efficiently using internal IPs from that network.

Option B is wrong as you do not need to create stackdriver account for each project.

Option D is wrong as it is recommended to create a separate stackdriver account instead of an account for one of the project.

### 31. Question

You are deploying an application to a Compute Engine VM in a managed instance group. The application must be running at all times, but only a single instance of the VM should run per GCP project. How should you configure the instance group?

- A. Set autoscaling to On, set the minimum number of instances to 1, and then set the maximum number of instances to 1.
- B. Set autoscaling to Off, set the minimum number of instances to 1, and then set the maximum number of instances to 1.
- C. Set autoscaling to On, set the minimum number of instances to 1, and then set the maximum number of instances to 2.
- D. Set autoscaling to Off, set the minimum number of instances to 1, and then set the maximum number of instances to 2.

**Unattempted**

Correct answer is A as you can configure Auto Scaling with minimum and maximum 1, to ensure only 1 instance is running. Auto Scaling needs to be configured with an Auto Scaling policy to detect the failure and create a new instance. Ideally, you can enable Auto Healing to recover the instance, however that is not covered in any answer option.

Refer GCP documentation – Compute Engine Auto Scaler

Managed instance groups offer autoscaling capabilities that allow you to automatically add or delete instances from a managed instance group based on increases or decreases in load. Autoscaling helps your applications gracefully handle increases in traffic and reduce costs when the need for resources is lower. You define the autoscaling policy and the autoscaler performs automatic scaling based on the measured load.

Autoscaling works by adding more instances to your instance group when there is more load (upscale), and deleting instances when the need for instances is lowered (downscale).

Option C is wrong as you need only 1 instance at a time, the maximum needs to be set to 1.

Options B & D are wrong as you need to enable autoscaling.

### 32. 32. Question

You need to allow traffic from specific virtual machines in ‘ subnet-a’ network access to machines in ‘ subnet-b’ without giving the entirety of subnet-a access. How can you accomplish this?

- A. Create a firewall rule to allow traffic from resources with specific network tags, then assign the machines in subnet-a the same tags.
- B. Relocate the subnet-a machines to a different subnet and give the new subnet the needed access.
- C. Create a rule to deny all traffic to the entire subnet, then create a second rule with higher priority giving access to tagged VM's in subnet-a.
- D. You can only grant firewall access to an entire subnet and not individual VM's inside.

**Unattempted**

Correct answer is A as Network tags allow more granular access based on individually tagged instances.

Refer GCP documentation – VPC Network Tags

Network tags are text attributes you can add to Compute Engine virtual machine (VM) instances. Tags allow you to make firewall rules and routes applicable to specific VM instances.

You can only add network tags to VM instances or instance templates. You cannot tag other GCP resources. You can assign network tags to new instances at creation time, or you can edit the set of assigned tags at any time later. Network tags can be edited without stopping an instance.

Network tags allow you to apply firewall rules and routes to a specific instance or set of instances:

You make a firewall rule applicable to specific instances by using target tags and source tags.

You make a route applicable to specific instances by using a tag.

Option B is wrong as this would give the entire subnet access which is against the requirements: allow traffic from specific virtual machines in ‘ subnet-a’ network access to machines in ‘ subnet-b’ without giving the entirety of subnet-a access.

Option C is wrong as an explicit deny is not needed as implicitly all traffic is allowed.

Option D is wrong as firewall access can be granted to individual instances.

### 33. Question

You want to send and consume Cloud Pub/Sub messages from your App Engine application. The Cloud Pub/Sub API is currently disabled. You will use a service account to authenticate your application to the API. You want to make sure your application can use Cloud Pub/Sub. What should you do?

- A. Enable the Cloud Pub/Sub API in the API Library on the GCP Console.
- B. Rely on the automatic enablement of the Cloud Pub/Sub API when the Service Account accesses it.
- C. Use Deployment Manager to deploy your application. Rely on the automatic enablement of all APIs used by the application being deployed.
- D. Grant the App Engine Default service account the role of Cloud Pub/Sub Admin. Have your application enable the API on the first connection to Cloud Pub/Sub.

Unattempted

Correct answer is A as the standard method is to enable services in the Google Cloud Console. You can also enable services with the Cloud SDK CLI gcloud services enable pubsub.googleapis.com

Refer GCP documentation – Cloud Pub/Sub Quick Setup

Option B is wrong as Google Cloud Services are not automatically enabled when the service account accesses it. First, service accounts do not access APIs. Service accounts are used to obtain an OAuth Access Token (or Identity Token). These tokens are used to authorize APIs. Services are not automatically enabled with an API makes first access.

Option C is wrong as Deployment Manager does not automatically enable services. You can use Deployment Manager Resource Types to enable services. You must create a virtual resource for each API that you want enabled.

Option D is wrong as Cloud Pub/Sub Admin does not have permissions to enable services. To enable services the service account (or User Account) will need roles/serviceusage.serviceUsageAdmin or another role with the permission serviceusage.services.enable.

#### 34. Question

You are using Cloud Shell and need to install a custom utility for use in a few weeks. Where can you store the file so it is in the default execution path and persists across sessions?

- A. Cloud Storage
- B. /google/scripts
- C. ~/bin
- D. ?/usr/local/bin

**Unattempted**

Correct answer is C as only HOME directory is persisted across sessions.

Refer GCP documentation – Cloud Shell

Cloud Shell provisions 5 GB of free persistent disk storage mounted as your \$HOME directory on the virtual machine instance. This storage is on a per-user basis and is available across projects. Unlike the instance itself, this storage does not time out on inactivity. All files you store in your home directory, including installed software, scripts and user configuration files like .bashrc and .vimrc, persist between sessions. Your \$HOME directory is private to you and cannot be accessed by other users.

Options A, B & D are wrong as they are not persistent across sessions.

### 35. 35. Question

You have a single binary application that you want to run on Google Cloud Platform. You decided to automatically scale the application based on underlying infrastructure CPU usage. Your organizational policies require you to use virtual machines directly. You need to ensure that the application scaling is operationally efficient and completed as quickly as possible. What should you do?

- A. Create a Google Kubernetes Engine cluster, and use horizontal pod autoscaling to scale the application.
- B. Create an instance template, and use the template in a managed instance group with autoscaling configured.
- C. Create an instance template, and use the template in a managed instance group that scales up and down based on the time of day.
- D. Use a set of third-party tools to build automation around scaling the application up and down, based on Stackdriver CPU usage monitoring.

#### Unattempted

Correct answer is B as a managed instance group can help use virtual machines directly and with autoscaling can scaling as per the demand.

Refer GCP documentation – Managed Instance Groups AutoScaling

Managed instance groups offer autoscaling capabilities that allow you to automatically add or delete instances from a managed instance group based on increases or decreases in load. Autoscaling helps your applications gracefully handle increases in traffic and reduce costs when the need for resources is lower. You define the autoscaling policy and the autoscaler performs automatic scaling based on the measured load.

Autoscaling works by adding more instances to your instance group when there is more load (upscale), and deleting instances when the need for instances is lowered (downscale).

Option A is wrong as Google Kubernetes Engine cluster can support scaling, however it would not meet the requirement of using virtual machines directly.

Option C is wrong as scaling based on time does not effectively utilize the scaling as per the demand.

Option D is wrong as using external tools is the least preferred option.

### 36. 36. Question

You have a project for your App Engine application that serves a development environment. The required testing has succeeded and you want to create a new project to serve as your production environment. What should you do?

- A. Use gcloud to create the new project, and then deploy your application to the new project.
- B. Use gcloud to create the new project and to copy the deployed application to the new project.
- C. Create a Deployment Manager configuration file that copies the current App Engine deployment into a new project.
- D. Deploy your application again using gcloud and specify the project parameter with the new project name to create the new project.

### Unattempted

Correct answer is A as gcloud can be used to create a new project and the gcloud app deploy can point to the new project.

Refer GCP documentation – GCloud App Deploy

–project=PROJECT\_ID

The Google Cloud Platform project name to use for this invocation. If omitted, then the current project is assumed; the current project can be listed using gcloud config list –format='text(core.project)' and can be set using gcloud config set project PROJECTID.

–project and its fallback core/project property play two roles in the invocation. It specifies the project of the resource to operate on. It also specifies the project for API enablement check, quota, and billing. To specify a different project for quota and billing, use –billing-project or billing/quota\_project property.

Option B is wrong as the option to use gcloud app cp does not exist.

Option C is wrong as Deployment Manager does not copy the application, but allows you to specify all the resources needed for your application in a declarative format using yaml

Option D is wrong as gcloud app deploy would not create a new project. The project should be created before usage.

### 37. Question

Your company uses Cloud Storage to store application backup files for disaster recovery purposes. You want to follow Google's recommended practices. Which storage option should you use?

- A. Multi-Regional Storage
- B. Regional Storage
- C. Nearline Storage

○  D. Coldline Storage  
Unattempted

Correct answer is D as Coldline storage is an ideal solution for disaster recovery data given its rarity of access.

Refer GCP documentation – Cloud Storage Classes

Google Cloud Storage Coldline is a very-low-cost, highly durable storage service for data archiving, online backup, and disaster recovery. Unlike other “cold” storage services, your data is available within milliseconds, not hours or days.

Coldline Storage is the best choice for data that you plan to access at most once a year, due to its slightly lower availability, 90-day minimum storage duration, costs for data access, and higher per-operation costs. For example:

Cold Data Storage – Infrequently accessed data, such as data stored for legal or regulatory reasons, can be stored at low cost as Coldline Storage, and be available when you need it.

Disaster recovery – In the event of a disaster recovery event, recovery time is key. Cloud Storage provides low latency access to data stored as Coldline Storage.

The geo-redundancy of Coldline Storage data is determined by the type of location in which it is stored: Coldline Storage data stored in multi-regional locations is redundant across multiple regions, providing higher availability than Coldline Storage data stored in regional locations.

Options A, B & C are wrong as they are not suited for infrequently accessed data, as disaster does not happen periodically but rarely.

38. 38. Question

You’ ve deployed a microservice called myapp1 to a Google Kubernetes Engine cluster using the YAML file specified below:

```
apiVersion: apps/v1
kind: Deployment
metadata:
 name: myapp1-deployment
spec:
 selector:
 matchLabels:
 app: myapp1
 replicas: 2
 template:
 metadata:
 labels:
 app: myapp1
```

```
spec:
containers:
name: main-container
image: gcr.io/my-company-repo/myapp1:1.4
env:
name: DS_PASSWORD
value: "tOugh2guess!"
ports: - containerPort: 8080
```

You need to refactor this configuration so that the database password is not stored in plain text. You want to follow Google-recommended practices. What should you do?

- A. Store the database password inside the Docker image of the container, not in the YAML file.
- B. Store the database password inside a Secret object. Modify the YAML file to populate the DB\_PASSWORD environment variable from the Secret.
- C. Store the database password inside a ConfigMap object. Modify the YAML file to populate the DB\_PASSWORD environment variable from the ConfigMap.
- D. ?Store the database password in a file inside a Kubernetes persistent volume, and use a persistent volume claim to mount the volume to the container.

#### Unattempted

Correct answer is B as Google Kubernetes Engine supports secret to store sensitive data such as database passwords and is a google recommended practice.

Refer GCP documentation – Kubernetes Engine Secret

Secrets are secure objects which store sensitive data, such as passwords, OAuth tokens, and SSH keys, in your clusters. Storing sensitive data in Secrets is more secure than plaintext ConfigMaps or in Pod specifications. Using Secrets gives you control over how sensitive data is used, and reduces the risk of exposing the data to unauthorized users.

Options A, C & D are wrong as others options are not secured and not recommended as best practice.

#### 39. Question

You created an instance of SQL Server 2017 on Compute Engine to test features in the new version. You want to connect to this instance using the fewest number of steps. What should you do?

- A. Install a RDP client on your desktop. Verify that a firewall rule for port 3389 exists.

- B. Install a RDP client in your desktop. Set a Windows username and password in the GCP Console. Use the credentials to log in to the instance.
- C. Set a Windows password in the GCP Console. Verify that a firewall rule for port 22 exists. Click the RDP button in the GCP Console and supply the credentials to log in.
- D. Set a Windows username and password in the GCP Console. Verify that a firewall rule for port 3389 exists. Click the RDP button in the GCP Console, and supply the credentials to log in.

**Unattempted**

Correct answer is A as it mentions fewest number of steps to connect to the instance. You can download the RDP Client and verify 3389 firewall is open. If the RDP asks for username and password, the instance is working.

Option B is wrong as it fails to mention the key requirement of port 3389 being opened.

Option C is wrong as RDP requires port 3389 to be opened.

Option D is wrong as you need an RDP client.

#### 40. 40. Question

You are building a pipeline to process time-series data. Which Google Cloud Platform services should you put in boxes 1,2,3, and 4?

- A. Cloud Pub/Sub, Cloud Dataflow, Cloud Datastore, BigQuery
- B. Firebase Messages, Cloud Pub/Sub, Cloud Spanner, BigQuery
- C. Cloud Pub/Sub, Cloud Storage, BigQuery, Cloud Bigtable
- D. Cloud Pub/Sub, Cloud Dataflow, Cloud Bigtable, BigQuery

**Unattempted**

Correct answer is D as Cloud Pub/Sub for data ingestion, Dataflow for data handling and transformation, Bigtable for storage to provide low latency data access and BigQuery for analytics.

Refer GCP documentation – Time Series Dataflow

Cloud Pub/Sub. As well as performing ingestion, Cloud Pub/Sub can also act as the glue between the loosely coupled systems. You can send the processed data to other systems to consume; for example, you might send all correlations with more than the value of  $\text{ABS}(0.2)$  to other systems.

BigQuery. Place any data that you want to process or access later using a SQL interface into BigQuery.

Cloud Bigtable. Place any data that you want to use for low-latency storage, or where you might want to get at a very small subset of a larger dataset quickly (key lookups as well as range scans), in Cloud Bigtable.

Option A is wrong as Datastore is not an ideal solution to store large time series data.

Option B is wrong as Cloud Spanner is not an ideal solution for storage.

Option C is wrong as Cloud Storage is for storage and doesn't help handle and source data to storage and analytics.

#### 41. Question

You are deploying an application to App Engine. You want the number of instances to scale based on request rate. You need at least 3 unoccupied instances at all times. Which scaling type should you use?

- A. Manual Scaling with 3 instances.
- B. Basic Scaling with min\_instances set to 3.
- C. Basic Scaling with max\_instances set to 3.
- D. Automatic Scaling with min\_idle\_instances set to 3.

**Unattempted**

Correct answer is D as min\_idle\_instances property can be set to have minimum idle instances which would be always running.

Refer GCP documentation – App Engine Scaling & app.yaml Reference

Auto scaling services use dynamic instances that get created based on request rate, response latencies, and other application metrics. However, if you specify a number of minimum idle instances, that specified number of instances run as resident instances while any additional instances are dynamic.

min\_idle\_instances

The number of instances to be kept running and ready to serve traffic. Note that you are charged for the number of instances specified whether they are receiving traffic or not. This setting only applies to the version that receives most of the traffic. Keep the following in mind:

A low minimum helps keep your running costs down during idle periods, but means that fewer instances might be immediately available to respond to a sudden load spike.

A high minimum allows you to prime the application for rapid spikes in request load. App Engine keeps the minimum number of instances running to serve incoming requests. You are charged for the number of instances specified, whether or not they are handling requests. For this feature to function properly, you must make sure that warmup requests are enabled and that your application handles warmup requests.

If you set a minimum number of idle instances, pending latency will have less effect on your application's performance. Because App Engine keeps idle instances in reserve, it is unlikely that requests will enter the pending queue except in exceptionally high load spikes. You will need to test your application and expected traffic volume to determine the ideal number of instances to keep in reserve.

Option A is wrong as manual scaling would not provide the scaling based on the request rate and would need manual intervention.

Options B & C are wrong as basic scaling will not allow the scaling based on the request rate.

#### 42. Question

You significantly changed a complex Deployment Manager template and want to confirm that the dependencies of all defined resources are properly met before committing it to the project. You want the most rapid feedback on your changes. What should you do?

- A. Use granular logging statements within a Deployment Manager template authored in Python.
- B. Monitor activity of the Deployment Manager execution on the Stackdriver Logging page of the GCP Console.
- C. Execute the Deployment Manager template against a separate project with the same configuration, and monitor for failures.
- D. Execute the Deployment Manager template using the --preview option in the same project, and observe the state of interdependent resources.

**Unattempted**

Correct answer is D as Deployment Manager provides the preview feature to check on what resources would be created.

Refer GCP documentation – Deployment Manager Preview

After you have written a configuration file, you can preview the configuration before you create a deployment. Previewing a configuration lets you see the resources that Deployment Manager would create but does not actually instantiate any actual resources. The Deployment Manager service previews the configuration by:

Expanding the full configuration, including any templates.

Creating a deployment and “ shell” resources.

You can preview your configuration by using the preview query parameter when making an insert() request.

```
gcloud deployment-manager deployments create example-deployment \
- config configuration-file.yaml - preview
```

43. 43. Question

You have a development project with appropriate IAM roles defined. You are creating a production project and want to have the same IAM roles on the new project, using the fewest possible steps. What should you do?

- A. Use gcloud iam roles copy and specify the production project as the destination project.
- B. Use gcloud iam roles copy and specify your organization as the destination organization.
- C. In the Google Cloud Platform Console, use the ‘ create role from role’ functionality.
- D. ?In the Google Cloud Platform Console, use the ‘ create role’ functionality and select all applicable permissions.

**Unattempted**

Correct answer is A as Cloud SDK gcloud iam roles copy can be used to copy the roles to different organization or project.

Refer GCP documentation – Cloud SDK IAM Copy Role

gcloud iam roles copy – create a role from an existing role

– dest-organization=DEST\_ORGANIZATION (The organization of the destination role)

– dest-project=DEST\_PROJECT (The project of the destination role)

Option B is wrong as the destination new project needs to be specified instead of the organization.

Option C is wrong as creating roles through GCP Console is cumbersome, time consuming and error prone.

Option D is wrong as it does not replicate the IAM roles permission.

44. 44. Question

You have one GCP account running in your default region and zone and another account running in a non-default region and zone. You want to start a new Compute Engine instance in these two Google Cloud Platform accounts using the command line interface. What should you do?

- A. Create two configurations using gcloud config configurations create [NAME]. Run gcloud config configurations activate [NAME] to switch between accounts when running the commands to start the Compute Engine instances.
- B. Create two configurations using gcloud config configurations create [NAME]. Run gcloud config configurations list to start the Compute Engine instances.
- C. Activate two configurations using gcloud config configurations activate [NAME] . Run gcloud config configurations list to start the Compute Engine instances.
- D. Activate two configurations using gcloud config configurations activate [NAME] . Run gcloud config configurations list to start the Compute Engine instances.

**Unattempted**

Correct answer is A as you can create different configurations for each account and create compute instances in each account by activating the respective account.

Refer GCP documentation – Configurations Create & Activate

Options B, C & D are wrong as gcloud config configurations list does not help create instances. It would only lists existing named configurations.

45. 45. Question

You recently deployed a new version of an application to App Engine and then discovered a bug in the release. You need to immediately revert to the prior version of the application. What should you do?

- A. Run gcloud app restore.
- B. On the App Engine page of the GCP Console, select the application that needs to be reverted and click Revert.
- C. On the App Engine Versions page of the GCP Console, route 100% of the traffic to the previous version.
- D. ?Deploy the original version as a separate application. Then go to App Engine settings and split traffic between applications so that the original version serves 100% of the requests.

### Unattempted

Correct answer is C as you can migrate all the traffic back to the previous version.

Refer GCP documentation – App Engine Overview

Having multiple versions of your app within each service allows you to quickly switch between different versions of that app for rollbacks, testing, or other temporary events. You can route traffic to one or more specific versions of your app by migrating or splitting traffic.

Option A is wrong as gcloud app restore was used for backup and restore and has been deprecated.

Option B is wrong as there is no application revert functionality available.

Option D is wrong as App Engine maintains version and need not be redeployed.

### 46. Question

You need to select and configure compute resources for a set of batch processing jobs. These jobs take around 2 hours to complete and are run nightly. You want to minimize service costs. What should you do?

- A. Select Google Kubernetes Engine. Use a single-node cluster with a small instance type.
- B. Select Google Kubernetes Engine. Use a three-node cluster with micro instance type.
- C. Select Compute Engine. Use preemptible VM instances of the appropriate standard machine type.
- D. Select Compute Engine. Use VM instance types that support micro bursting.

### Unattempted

Correct answer is C as Compute Engine preemptible VMs are ideal for batch processing jobs and are able run at a much lower prices than standard instances.

Refer GCP documentation – Compute Engine Preemptible

A preemptible VM is an instance that you can create and run at a much lower price than normal instances. However, Compute Engine might terminate (preempt) these instances if it requires access to those resources for other tasks. Preemptible instances are excess Compute Engine capacity, so their availability varies with usage.

If your apps are fault-tolerant and can withstand possible instance preemptions, then preemptible instances can reduce your Compute Engine costs significantly. For example, batch processing jobs can run on preemptible instances. If some of those instances terminate during processing, the job slows but does not completely stop. Preemptible instances complete your batch processing tasks without placing additional workload on your existing instances and without requiring you to pay full price for additional normal instances.

Options A, B & D are wrong as they would Compute Engine instances running, which is not a cost effective option for batch processing jobs.

47. **47. Question**

You are analyzing Google Cloud Platform service costs from three separate projects. You want to use this information to create service cost estimates by service type, daily and monthly, for the next six months using standard query syntax. What should you do?

- A. Export your bill to a Cloud Storage bucket and then import into Cloud Bigtable for analysis.
- B. Export your bill to a Cloud Storage bucket, and then import into Google Sheets for analysis.
- C. Export your transactions to a local file and perform analysis with a desktop tool.
- D. Export your bill to a BigQuery dataset, and then write time window-based SQL queries for analysis.

**Unattempted**

Correct answer is D as BigQuery provides an ideal storage option to store and query in standard SQL dialect.

BigQuery, Google's serverless, highly scalable enterprise data warehouse, is designed to make data analysts more productive with unmatched price-performance. Because there is no infrastructure to manage, you can focus on uncovering meaningful insights using familiar SQL without the need for a database administrator.

Option A is wrong Bigtable is a NoSQL solution and does not support SQL dialect.

Cloud Bigtable is Google's sparsely populated NoSQL database which can scale to billions of rows, thousands of columns, and petabytes of data. Cloud Bigtable has a data model similar to Apache HBase and provides an HBase-compatible client library.

Options B & C are wrong as Google Sheets and local file does not provide standard query syntax querying.

48. 48. Question

You need a dynamic way of provisioning VMs on Compute Engine. The exact specifications will be in a dedicated configuration file. You want to follow Google's recommended practices. Which method should you use?

- A. Deployment Manager
- B. Cloud Composer
- C. Managed Instance Group
- D. Unmanaged Instance Group

**Unattempted**

Correct answer is A as Deployment Manager provide Infrastructure as a Code capability.

Refer GCP documentation – Deployment Manager

Google Cloud Deployment Manager allows you to specify all the resources needed for your application in a declarative format using yaml. You can also use Python or Jinja2 templates to parameterize the configuration and allow reuse of common deployment paradigms such as a load balanced, auto-scaled instance group. Treat your configuration as code and perform repeatable deployments.

Option B is wrong as Cloud Composer is a fully managed workflow orchestration service that empowers you to author, schedule, and monitor pipelines that span across clouds and on-premises data centers.

Options C & D are wrong as An instance group is a collection of VM instances that you can manage as a single entity.

Managed instance groups (MIGs) allow you to operate applications on multiple identical VMs. You can make your workloads scalable and highly available by taking advantage of automated MIG services, including: autoscaling, autohealing, regional (multi-zone) deployment, and auto-updating.

Unmanaged instance groups allow you to load balance across a fleet of VMs that you manage yourself.

49. 49. Question

Your team is developing a product catalog that allows end users to search and filter. The full catalog of products consists of about 500 products. The team doesn't have any experience with SQL, or schema migrations, so they're considering a NoSQL option. Which database service would work best?

- A. Cloud SQL
- B. Cloud Memorystore

- C. Bigtable
- D. Cloud Datastore

**Unattempted**

Correct answer is D as Cloud Datastore would provide the NoSQL option for storing the product catalog. As the data is limited, it would be a good fit.

Option A is wrong as Cloud SQL is a relational SQL solution.

Option B is wrong as Cloud Memorystore for Redis provides a fully managed in-memory data store service built on scalable, secure, and highly available infrastructure managed by Google. Use Cloud Memorystore to build application caches that provides sub-millisecond data access. Cloud Memorystore is compatible with the Redis protocol, allowing easy migration with zero code changes.

Option C is wrong as although Bigtable provides a NoSQL solution, it is a petabyte-scale, fully managed NoSQL database service ideal for large analytical and operational workloads.

50. 50. Question

You're trying to provide temporary access to some files in a Cloud Storage bucket. You want to limit the time that the files are available to 10 minutes. With the fewest steps possible, what is the best way to generate a signed URL?

- A. Create a service account and JSON key. Use the gsutil signurl -t 10m command and pass in the JSON key and bucket.
- B. Create a service account and JSON key. Use the gsutil signurl -d 10m command and pass in the JSON key and bucket.
- C. Create a service account and JSON key. Use the gsutil signurl -p 10m command and pass in the JSON key and bucket.
- D. ?Create a service account and JSON key. Use the gsutil signurl -m 10m command and pass in the JSON key and bucket.

**Unattempted**

Correct answer is B as signurl command will generate a signed URL that embeds authentication data so the URL can be used by someone who does not have a Google account. -d can help provide the time duration.

Refer GCP documentation – Cloud Storage gsutil signurl

```
gsutil signurl [-c] [-d] [-m] \
[-p] [-r] keystore-file url...
```

-m Specifies the HTTP method to be authorized for use with the signed url, default is GET. You may also specify RESUMABLE to create a signed resumable upload start URL. When using a signed URL to start a resumable upload session, you will need to specify the 'x-goog-resumable:start' header in the request or else signature validation will fail.

-d Specifies the duration that the signed url should be valid for, default duration is 1 hour. Times may be specified with no suffix (default hours), or with s = seconds, m = minutes, h = hours, d = days. This option may be specified multiple times, in which case the duration the link remains valid is the sum of all the duration options. The max duration allowed is 7d.

-c Specifies the content type for which the signed url is valid for.-p Specify the keystore password instead of prompting.

-r Specifies the region in which the resources for which you are creating signed URLs are stored. Default value is 'auto' which will cause gsutil to fetch the region for the resource. When auto-detecting the region, the current gsutil user's credentials, not the credentials from the private-key-file, are used to fetch the bucket's metadata. This option must be specified and not 'auto' when generating a signed URL to create a bucket.

## 51. Question

You're about to deploy your team's App Engine application. They're using the Go runtime with a Standard Environment. Which command should you use to deploy the application?

- A. gcloud app deploy app.yaml
- B. gcloud app-engine apply app.yaml
- C. gcloud app apply app.yaml
- D. ?gcloud app-engine deploy app.yaml

**Unattempted**

Correct answer is A as gcloud app deploy provides an ability to deploy the local code and/or configuration of your app to App Engine.

Refer GCP documentation – gcloud app deploy

This command is used to deploy both code and configuration to the App Engine server. As an input it takes one or more DEPLOYABLES that should be uploaded. A DEPLOYABLE can be a service's .yaml file or a configuration's .yaml file.

Option C is wrong as gcloud app apply is not a valid command.

Options B & D are wrong as gcloud app-engine is not a valid command.

## 52. 52. Question

You have a Windows server running on a custom network. There's an allow firewall rule with an IP filter of 0.0.0.0/0 with a protocol/port of tcp:3389. The logs on the instance show a constant stream of attempts from different IP addresses, trying to connect via RDP. You suspect this is a brute force attack. How might you change the firewall rule to stop this from happening and still enable access for legit users?

- A. Stop the instance.
- B. Deny all traffic to port 3389.
- C. Change the port that RDP is running on in the instance and change the port number in the firewall rule.
- D. Change the IP address range in the filter to only allow known IP addresses.

**Unattempted**

Correct answer is D as by using 0.0.0.0/0, you're opening the port to the internet. By whitelisting known IP addresses, it will block anyone not on the list.

Option A is wrong as it is not a viable solution for protecting the instances.

Option B is wrong denying all traffic would block all.

Option C is wrong as it is not possible to change the default RDP port.

## 53. 53. Question

You've found that your Linux server keeps running low on memory. It's currently using 8GB of memory, and it needs to be increased to 16. What is the simplest way to do that?

- A. Use the gcloud compute add-memory command to increase the memory.
- B. Use the Linux memincr command to increase the memory.
- C. Stop the instance and change the machine type.
- D. Create a new instance with the correct amount of memory.

**Unattempted**

Correct answer is C as you can increase the memory by changing the instance machine type.

Refer GCP documentation – Changing Machine Type

You can change the machine type of a stopped instance if it is not part of a managed instance group. If you need to change the machine type of instances within a managed instance group, read Updating managed instance groups.

Change the machine types of your instances if your existing machine type is not a good fit for the workloads you run on that instance. You can change the machine type of an instance to adjust the number of vCPUs and memory as your workload changes. For example, you can start an instance with a smaller machine during setup, development, and testing and change the instance to use a larger machine type when you are ready for production workloads.

Options A & B are wrong as the options are invalid.

Option D is wrong as the solution is valid, but it is not the simplest.

#### 54. Question

You're working on setting up a cluster of virtual machines with GPUs to perform some 3D rendering for a customer. They're on a limited budget and are looking for ways to save money. What is the best solution for implementing this?

- A. Use an autoscaled managed instance group containing some preemptible instances.
- B. Use an unmanaged instance group with preemptible instances.
- C. Use App Engine with Flexible Environments.
- D. Use App Engine with Standard Environments.

**Unattempted**

Correct answer is A as Preemptible with managed instance groups would help add GPUs at a lower cost.

Refer GCP documentation – Compute Engine Preemptible

A preemptible VM is an instance that you can create and run at a much lower price than normal instances. However, Compute Engine might terminate (preempt) these instances if it requires access to those resources for other tasks. Preemptible instances are excess Compute Engine capacity, so their availability varies with usage.

If your apps are fault-tolerant and can withstand possible instance preemptions, then preemptible instances can reduce your Compute Engine costs significantly. For example, batch processing jobs can run on preemptible instances. If some of those instances terminate during processing, the job slows but does not completely stop. Preemptible instances complete your batch processing tasks without placing additional workload on your existing instances and without requiring you to pay full price for additional normal instances.

You can add GPUs to your preemptible VM instances at lower preemptible prices for the GPUs. GPUs attached to preemptible instances work like normal GPUs but persist only for the life of the instance. Preemptible instances with GPUs follow the same preemption process as all preemptible instances.

Option B is wrong as unmanaged instance group does not provide scaling.

Options C & D are wrong as GCP currently does not support GPUs for App Engine.

##### 55. Question

Your coworker has helped you set up several configurations for gcloud. You've noticed that you're running commands against the wrong project. Being new to the company, you haven't yet memorized any of the projects. With the fewest steps possible, what's the fastest way to switch to the correct configuration?

- A. Run gcloud configurations list followed by gcloud configurations activate.
- B. Run gcloud config list followed by gcloud config activate.
- C. Run gcloud config configurations list followed by gcloud config configurations activate.
- D. Re-authenticate with the gcloud auth login command and select the correct configurations on login.

**Incorrect**

Correct answer is C as gcloud config configurations list can help check for the existing configurations and activate can help switch to the configuration.

Refer GCP documentation – Cloud SDK gcloud config

gcloud config configurations list – lists existing named configurations

gcloud config configurations activate – activates an existing named configuration

Options A & B are wrong as they are invalid commands.

Option D is wrong as does not help to identify and activate configurations.

gcloud auth login – authorize gcloud to access the Cloud Platform with Google user credentials

Obtains access credentials for your user account via a web-based authorization flow. When this command completes successfully, it sets the active account in the current configuration to the account specified. If no configuration exists, it creates a configuration named default.