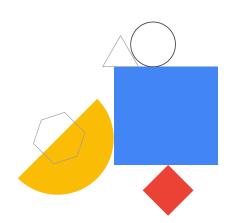
Google Cloud

Preparing for Your Professional Cloud Architect Journey



Course Workbook

Certification Exam Guide Sections

- 1 Designing and planning a cloud solution architecture
- 2 Managing and provisioning a solution infrastructure
- 3 Designing for security and compliance
- 4 Analyzing and optimizing technical and business processes
- 5 Managing implementation
- 6 Ensuring solution and operations reliability



Section 1: Designing and planning a cloud solution architecture

Cymbal Direct case study



Cymbal Direct case study - 01

Company overview

Cymbal Direct is an online direct-to-consumer
Chicago-based footwear and apparel retailer founded in 2008 and acquired by Cymbal Group in 2010. Cymbal Direct is a fair trade and B Corp certified, sustainability-focused company that works with cotton farmers to reinvest in their communities, a fact which appeals to Cymbal Direct's younger target market demographic.

Solution concept

- 1) The beta Delivery by Drone initiative enables licensed drone pilots to team up with Cymbal Direct to deliver shoes and sandals to customers via drone. DBD allows customers to place their orders and then get their shoes delivered in an expedited amount of time. The drones stream real-time video to their pilots, as well as their coordinates, so that customers can see the location of their shoes on a map.
- 2) Cymbal Direct wants to release official APIs for partners. APIs will be published in a controllable, versionable way, with the ability to track, secure and monetize.
- 3) A social integration service initiative which highlights images hashtagged with Cymbal Direct's products using machine learning to ensure images are appropriate. The social media highlighting service is currently proof-of-concept. Built by a developer in their own time after hours as an experiment, the service garnered a lot of excitement and interest, especially from the marketing team. During one of the internal demos, however, inappropriate images were included in the product gallery.

Existing technical environment

Cymbal Direct case study - 02

Delivery by Drone is an experiment by the supply chain and logistics team. Their core customer-facing application does order processing, showing the current status and location of their delivery. The drones connect via the cellular network. The drones use the drone API to receive commands and send real-time information and video about their location and status.

The existing technical environment includes:

- A website frontend and pilot and truck management systems run on Kubernetes
- Positional data for drone and truck location kept in MongoDB database clusters
- Drones connected to virtual machines using a stateful connection, streaming video via RMTP to the pilots and sending commands from the pilots to the drones

Purchase & Product APIs were developed over time as the business was being built. They were initially only intended to be used in-house, and not exposed to 3rd parties and partners.

- Many of the APIs are simply built into monolithic apps, and were not designed for partner integration, lacking functionality such as versioning.
- The majority of the APIs run on Ubuntu Linux VMs, and scaling has been somewhat difficult because of the use of virtual machines and monolithic architecture.
- APIs do not have a built-in mechanism for supporting multiple accounts and granting access is very limited as a result.

The social media highlighting service currently runs on a single virtual machine, and while it does work, it has some performance and scalability issues.

- SuSE linux
- MySQL DB
- Redis
- Python

Cymbal Direct case study - 03

Business requirements

- Easily scale to handle additional demand when needed and expand to more test markets.
- Streamline development for application modernization and new features/products
- Ensure that developers spend as much time on core business functionality as possible, and not have to worry about scalability wherever possible
- · Let partners order directly via API
- Deploy a production version of the social media highlighting service and ensure no inappropriate content

Technical requirements

- Move to managed services wherever possible
- Ensure that developers can deploy container-based workloads to testing and production environments in a highly scalable environment.
- Standardize on containers where possible, but also allow for existing virtualization infrastructure to run as-is without a re-write, so it can be slowly refactored over time
- Securely allow partner integration
- Stream IoT data from drones

Cymbal Direct case study - 04

Executive statement

Cymbal Direct has three areas of strategic focus: improving customer experience, leveraging analytics, and improving digital marketing. Cymbal Direct has experienced rapid growth and has had trouble meeting demand. The organization wants to implement solutions that will help scale services and personalize customer experiences. Cymbal Direct wants to be able to dynamically surge delivery during peak periods.

Cymbal Direct also wants to be able to facilitate large scale B2B orders and better predict customer demand and trends. The organization wants to ensure the security of its B2B partners' business plans and make it easier for those partners to integrate with Cymbal Direct's APIs to submit orders and specify customizations.

Cymbal Direct also wants to integrate social media and marketing applications into its platform. They would like to be able to highlight posts on social media platforms which feature Cymbal Direct products directly on their product pages, but are concerned about the possibility of having unsavory content shown to users accidentally.

Potential solutions

Cymbal Direct case study - 05



Existing environment

Website frontend, pilot, and truck management systems run on Kubernetes



Technical requirements (does it ...?)

- Move to managed services wherever possible
- Ensure that developers can deploy container based workloads to testing and production environments in a highly scalable environment.
- Standardize on containers where possible



Business requirements (does it ...?)

- Easily scale to handle additional demand when needed?
- Streamline development?



Proposed product/ solution

- Global external Application Load Balancer
- GKE in two regions
- Autoscaler
- Private cluster
- Separate projects for website / pilot / truck management -dev,test,staging for each
- Cloud Build
- Cloud Source Repository
- Artifact Registry
 Migration type: lift and shift
- Automation tooling: Terraform Firewall rules http/s
- Separate IAM roles for
- developers and devops Replace GKE with Cloud Run for website (future)



Cymbal Direct drones continuously send data during deliveries. You need to process and analyze the incoming telemetry data. After processing, the data should be retained, but it will only be accessed once every month or two. Your CIO has issued a directive to incorporate managed services wherever possible. You want a cost-effective solution to process the incoming streams of data.

What should you do?

- A. Ingest data with ClearBlade IoT Core, process it with Dataprep, and store it in a Coldline Cloud Storage bucket.
- B. Ingest data with ClearBlade IoT Core, and then publish to Pub/Sub. Use Dataflow to process the data, and store it in a Nearline Cloud Storage bucket.
- C. Ingest data with ClearBlade IoT Core, and then publish to Pub/Sub. Use BigQuery to process the data, and store it in a Standard Cloud Storage bucket.
- D. Ingest data with ClearBlade IoT Core, and then store it in BigQuery.



Customers need to have a good experience when accessing your web application so they will continue to use your service. You want to define key performance indicators (KPIs) to establish a service level objective (SLO).

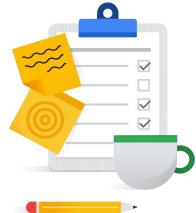
Which KPI could you use?

- A. Eighty-five percent of customers are satisfied users
- B. Eighty-five percent of requests succeed when aggregated over 1 minute
- C. Low latency for > 85% of requests when aggregated over 1 minute
- D. Eighty-five percent of requests are successful

Designing a solution infrastructure that meets business requirements

Resources to start your journey

Google Cloud Architecture Framework: System design SRE Books



Cymbal Direct developers have written a new application. Based on initial usage estimates, you decide to run the application on Compute Engine instances with 15 Gb of RAM and 4 CPUs. These instances store persistent data locally. After the application runs for several months, historical data indicates that the application requires 30 Gb of RAM. Cymbal Direct management wants you to make adjustments that will minimize costs.

What should you do?

- A. Stop the instance, and then use the command gcloud compute instances set-machine-type VM_NAME --machine-type e2-standard-8. Start the instance again.
- B. Stop the instance, and then use the command gcloud compute instances set-machine-type VM_NAME --machine-type e2-standard-8. Set the instance's metadata to: preemptible: true. Start the instance again.
- C. Stop the instance, and then use the command gcloud compute instances set-machine-type VM_NAME --machine-type 2-custom-4-30720. Start the instance again.
- D. Stop the instance, and then use the command gcloud compute instances set-machine-type VM_NAME --machine-type 2-custom-4-30720. Set the instance's metadata to: preemptible: true. Start the instance again.

Designing a solution infrastructure that meets technical requirements

Resources to start your journey

Google Cloud Architecture Framework: System design



You are creating a new project. You plan to set up a Dedicated interconnect between two of your data centers in the near future and want to ensure that your resources are only deployed to the same regions where your data centers are located. You need to make sure that you don't have any overlapping IP addresses that could cause conflicts when you set up the interconnect. You want to use RFC 1918 class B address space.

What should you do?

- Create a new project, leave the default network in place, and then use the default
 10.x.x.x network range to create subnets in your desired regions.
- B. Create a new project, delete the default VPC network, set up an auto mode VPC network, and then use the default 10.x.x.x network range to create subnets in your desired regions.
- C. Create a new project, delete the default VPC network, set up a custom mode VPC network, and then use IP addresses in the 172.16.x.x address range to create subnets in your desired regions.
- D. Create a new project, delete the default VPC network, set up the network in custom mode, and then use IP addresses in the 192.168.x.x address range to create subnets in your desired zones. Use VPC Network Peering to connect the zones in the same region to create regional networks.

Cymbal Direct is working with Cymbal Retail, a separate, autonomous division of Cymbal with different staff, networking teams, and data center. Cymbal Direct and Cymbal Retail are not in the same Google Cloud organization. Cymbal Retail needs access to Cymbal Direct's web application for making bulk orders, but the application will not be available on the public internet. You want to ensure that Cymbal Retail has access to your application with low latency. You also want to avoid egress network charges if possible.

What should you do?

- A. Verify that the subnet range
 Cymbal Retail is using doesn't
 overlap with Cymbal Direct's subnet range, and then enable VPC
 Network Peering for the project.
- B. If Cymbal Retail does not have access to a Google Cloud data center, use Carrier Peering to connect the two networks.
- C. Specify Cymbal Direct's project as the Shared VPC host project, and then configure Cymbal Retail's project as a service project.
- D. Verify that the subnet Cymbal Retail is using has the same IP address range with Cymbal Direct's subnet range, and then enable VPC Network Peering for the project.



Cymbal Direct's employees will use Google Workspace. Your current on-premises network cannot meet the requirements to connect to Google's public infrastructure.

What should you do?

- A. Order a Dedicated Interconnect from a Google Cloud partner, and ensure that proper routes are configured.
- B. Connect the network to a Google point of presence, and enable Direct Peering.
- C. Order a Partner Interconnect from a Google Cloud partner, and ensure that proper routes are configured.
- D. Connect the on-premises network to Google's public infrastructure via a partner that supports Carrier Peering.



Cymbal Direct is evaluating database options to store the analytics data from its experimental drone deliveries. You're currently using a small cluster of MongoDB NoSQL database servers. You want to move to a managed NoSQL database service with consistent low latency that can scale throughput seamlessly and can handle the petabytes of data you expect after expanding to additional markets.

What should you do?

- A. Extract the data from MongoDB. Insert the data into Firestore using Datastore mode.
- B. Create a Bigtable instance, extract the data from MongoDB, and insert the data into Bigtable.
- C. Extract the data from MongoDB. Insert the data into Firestore using Native mode.
- D. Extract the data from MongoDB, and insert the data into BigQuery.



You are working with a client who is using Google Kubernetes Engine (GKE) to migrate applications from a virtual machine-based environment to a microservices-based architecture. Your client has a complex legacy application that stores a significant amount of data on the file system of its VM. You do not want to re-write the application to use an external service to store the file system data.

A. In Cloud Shell, create a YAML file defining your Deployment called deployment.yaml. Create a Deployment in GKE by running the command kubectl apply -f deployment.yaml

- B. In Cloud Shell, create a YAML file defining your Container called build.yaml. Create a Container in GKE by running the command gcloud builds submit -config build.yaml.
- C. In Cloud Shell, create a YAML file defining your StatefulSet called statefulset.yaml. Create a StatefulSet in GKE by running the command kubectl apply -f statefulset.yaml
- D. In Cloud Shell, create a YAML file defining your Pod called pod.yaml. Create a Pod in GKE by running the command kubectl apply -f pod.yaml

What should you do?

Designing network, storage, and compute resources

Resources to start your journey

Choose and manage compute | Architecture Framework | Google Cloud

<u>Design your network infrastructure | Architecture</u> <u>Framework | Google Cloud</u>

Select and implement a storage strategy | Architecture Framework | Google Cloud

Google Cloud documentation





You are working in a mixed environment of VMs and Kubernetes. Some of your resources are on-premises, and some are in Google Cloud. Using containers as a part of your CI/CD pipeline has sped up releases significantly. You want to start migrating some of those VMs to containers so you can get similar benefits. You want to automate the migration process where possible.

What should you do?

- A. Manually create a GKE cluster, and then use Migrate to Containers (Migrate for Anthos) to set up the cluster, import VMs, and convert them to containers.
- B. Use Migrate to Containers (Migrate for Anthos) to automate the creation of Compute Engine instances to import VMs and convert them to containers.
- C. Manually create a GKE cluster. Use Cloud Build to import VMs and convert them to containers.
- D. Use Migrate for Compute Engine to import VMs and convert them to containers.

1.4 Creating a migration plan

Resources to start your journey

Migrate for Anthos and GKE

Migration to Google Cloud: Choosing your migration path

Migrating to the cloud: a guide and checklist

Cloud Migration Products & Services

Application Migration | Google Cloud





Cymbal Direct has created a proof of concept for a social integration service that highlights images of its products from social media. The proof of concept is a monolithic application running on a single SuSE Linux virtual machine (VM). The current version requires increasing the VM's CPU and RAM in order to scale. You would like to refactor the VM so that you can scale out instead of scaling up.

What should you do?

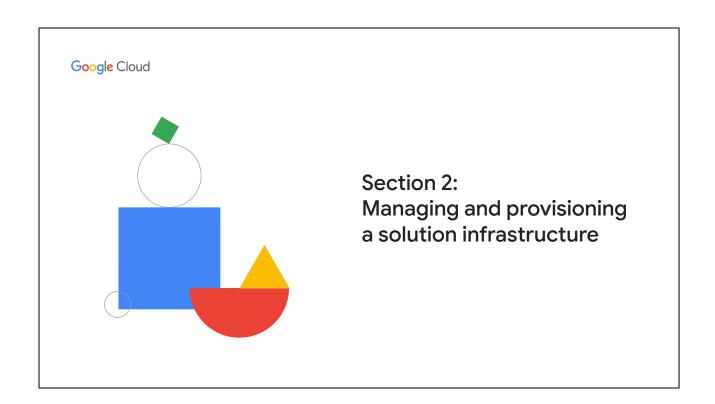
- A. Move the existing codebase and VM provisioning scripts to git, and attach external persistent volumes to the VMs.
- B. Make sure that the application declares any dependent requirements in a requirements.txt or equivalent statement so that they can be referenced in a startup script. Specify the startup script in a managed instance group template, and use an autoscaling policy.
- C. Make sure that the application declares any dependent requirements in a requirements.txt or equivalent statement so that they can be referenced in a startup script, and attach external persistent volumes to the VMs.
- D. Use containers instead of VMs, and use a GKE autoscaling deployment.

1.5 Envisioning future solution improvements

Resources to start your journey

<u>Twelve-factor app development on Google Cloud |</u>
<u>Cloud Architecture Center</u>







Cymbal Direct must meet compliance requirements. You need to ensure that employees with valid accounts cannot access their VPC network from locations outside of its secure corporate network, including from home. You also want a high degree of visibility into network traffic for auditing and forensics purposes.

What should you do?

- A. Ensure that all users install Cloud VPN. Enable VPC Flow Logs for the networks you need to monitor.
- B. Enable VPC Service Controls, define a network perimeter to restrict access to authorized networks, and enable VPC Flow Logs for the networks you need to monitor.
- C. Enable Identity-Aware Proxy (IAP) to allow users to access services securely. Use Google Cloud Observability to view audit logs for the networks you need to monitor.
- D. Enable VPC Service Controls, and use Google Cloud Observability to view audit logs for the networks you need to monitor.

You are working with a client who has built a secure messaging application. The application is open source and consists of two components. The first component is a web app, written in Go, which is used to register an account and authorize the user's IP address. The second is an encrypted chat protocol that uses TCP to talk to the backend chat servers running Debian. If the client's IP address doesn't match the registered IP address, the application is designed to terminate their session. The number of clients using the service varies greatly based on time of day, and the client wants to be able to easily scale as needed.

What should you do?

- Deploy the web application using the App Engine standard environment using a global external Application Load Balancer and a network endpoint group. Use an unmanaged instance group for the backend chat servers. Use an external network load balancer to load-balance traffic across the backend chat servers.
- B. Deploy the web application using the App Engine flexible environment using a global external Application Load Balancer and a network endpoint group. Use an unmanaged instance group for the backend chat servers. Use an external passthrough Network Load Balancer to load-balance traffic across the backend chat servers.
- C. Deploy the web application using the App Engine standard environment using a global external Application Load Balancer and a network endpoint group. Use a managed instance group for the backend chat servers. Use a global external proxy Network Load Balancer to load-balance traffic across the backend chat servers.
- D. Deploy the web application using the App Engine standard environment with a global external Application Load Balancer and a network endpoint group. Use a managed instance group for the backend chat servers. Use an external passthrough Network Load Balancer to load-balance traffic across the backend chat servers.



2.1 Configuring network topologies

Resources to start your journey

VPC network overview | Google Cloud

<u>Choosing a Network Connectivity product | Google Cloud</u> <u>Cloud VPN overview</u>

Best practices | Cloud Interconnect

Options for connecting to multiple VPC networks | Cloud Interconnect Best practices for enterprise organizations | Documentation | Google Cloud





Cymbal Direct's user account management app allows users to delete their accounts whenever they like. Cymbal Direct also has a very generous 60-day return policy for users. The customer service team wants to make sure that they can still refund or replace items for a customer even if the customer's account has been deleted.

What can you do to ensure that the customer service team has access to relevant account information?

- A. Temporarily disable the account for 30 days. Export account information to Cloud Storage, and enable lifecycle management to delete the data in 60 days.
- B. Ensure that the user clearly understands that after they delete their account, all their information will also be deleted. Remind them to download a copy of their order history and account information before deleting their account. Have the support agent copy any open or recent orders to a shared spreadsheet.
- C. Restore a previous copy of the user information database from a snapshot. Have a database administrator capture needed information about the customer.
- D. Disable the account. Export account information to Cloud Storage. Have the customer service team permanently delete the data after 30 days.

2.2 Configuring individual storage systems

Resources to start your journey

Select and implement a storage strategy | Architecture Framework | Google Cloud

Best practices for Cloud Storage

Enterprise tier | Filestore | Google Cloud

Design an optimal storage strategy for your cloud workload

Storage options | Compute Engine Documentation | Google Cloud

Cloud Storage Options | Google Cloud

Object storage vs block storage vs file storage: which should you choose? | Google Cloud Blog





Cymbal Direct wants to create a pipeline to automate the building of new application releases.

What sequence of steps should you use?

- A. Set up a source code repository. Run unit tests. Check in code. Deploy. Build a Docker container.
- B. Check in code. Set up a source code repository. Run unit tests. Deploy. Build a Docker container.
- C. Set up a source code repository. Check in code. Run unit tests. Build a Docker container. Deploy.
- D. Run unit tests. Deploy. Build a Docker container. Check in code. Set up a source code repository.



Your existing application runs on Ubuntu Linux VMs in an on-premises hypervisor. You want to deploy the application to Google Cloud with minimal refactoring.

What should you do?

- A. Set up a **Google Kubernetes Engine (GKE) cluster**, and then create a deployment with an autoscaler.
- B. Isolate the core features that the application provides. Use **Cloud Run** to deploy each feature independently as a microservice.
- C. Use Dedicated or Partner Interconnect to connect the on-premises network where your application is running to your VPC. Configure an endpoint for a global external Application Load Balancer that connects to the existing VMs.
- D. Write Terraform scripts to deploy the application as **Compute Engine instances**.



Cymbal Direct needs to use a tool to deploy its infrastructure. You want something that allows for repeatable deployment processes, uses a declarative language, and allows parallel deployment. You also want to deploy infrastructure as code on Google Cloud and other cloud providers.

A. Automate the deployment with Terraform scripts.

- B. Automate the deployment using scripts containing gcloud commands.
- Use Google Kubernetes Engine (GKE) to create deployments and manifests for your applications.
- D. Develop in Docker containers for portability and ease of deployment.

What should you do?



Cymbal Direct wants to allow partners to make orders programmatically, without having to speak on the phone with an agent.

What should you consider when designing the API?

- A. The API backend should be loosely coupled. Clients should not be required to know too many details of the services they use. REST APIs using gRPC should be used for all external APIs.
- B. The API backend should be tightly coupled. Clients should know a significant amount about the services they use. REST APIs using gRPC should be used for all external APIs.
- C. The API backend should be loosely coupled. Clients should not be required to know too many details of the services they use. For REST APIs, HTTP(S) is the most common protocol.
- D. The API backend should be tightly coupled. Clients should know a significant amount about the services they use. For REST APIs, HTTP(S) is the most common protocol used.



Cymbal Direct wants a layered approach to security when setting up Compute Engine instances.

What are some options you could use to make your Compute Engine instances more secure?

- A. Use labels to allow traffic only from certain sources and ports. Turn on Secure boot and vTPM.
- B. Use labels to allow traffic only from certain sources and ports. Use a Compute Engine service account.
- Use network tags to allow traffic only from certain sources and ports. Turn on Secure boot and vTPM.
- D. Use network tags to allow traffic only from certain sources and ports. Use a Compute Engine service account.



You have deployed your frontend web application in Kubernetes. Based on historical use, you need three pods to handle normal demand. Occasionally your load will roughly double. A load balancer is already in place.

How could you configure your environment to efficiently meet that demand?

- A. Edit your pod's configuration file and change the number of replicas to six.
- B. Edit your deployment's configuration file and change the number of replicas to six.
- C. Use the "kubectl autoscale" command to change the pod's maximum number of instances to six.
- D. Use the "kubectl autoscale" command to change the deployment's maximum number of instances to six.



You need to deploy a load balancer for a web-based application with multiple backends in different regions. You want to direct traffic to the backend closest to the end user, but also to different backends based on the URL the user is accessing.

Which of the following could be used to implement this?

- A. The request is received by the global external Application Load Balancer. A global forwarding rule sends the request to a target proxy, which checks the URL map and selects the backend service. The backend service sends the request to Compute Engine instance groups in multiple regions.
- B. The request is matched by a URL map and then sent to a global external Application Load Balancer. A global forwarding rule sends the request to a target proxy, which selects a backend service. The backend service sends the request to Compute Engine instance groups in multiple regions.
- C. The request is received by the proxy Network Load Balancer, which uses a global forwarding rule to check the URL map, then sends the request to a backend service. The request is processed by Compute Engine instance groups in multiple regions.
- D. The request is matched by a URL map and then sent to a proxy Network Load Balancer. A global forwarding rule sends the request to a target proxy, which selects a backend service and sends the request to Compute Engine instance groups in multiple regions.

2.3 Configuring compute systems

Resources to start your journey

Choose a Compute Engine deployment strategy for your workload

Google Kubernetes Engine documentation

General development tips | Cloud Run Documentation

Choosing the right compute option in GCP: a decision tree | Google Cloud Blog

Google Kubernetes Engine vs Cloud Run: Which should you use?



Section 3:
Designing for security and compliance



Your client created an Identity and Access Management (IAM) resource hierarchy with Google Cloud when the company was a startup. Your client has grown and now has multiple departments and teams. You want to recommend a resource hierarchy that follows Google-recommended practices.

What should you do?

- A. Keep all resources in one project, and use a flat resource hierarchy to reduce complexity and simplify management.
- B. Keep all resources in one project, but change the resource hierarchy to reflect company organization.
- C. Use a flat resource hierarchy and multiple projects with established trust boundaries.
- D. Use multiple projects with established trust boundaries, and change the resource hierarchy to reflect company organization.



Cymbal Direct's social media app must run in a separate project from its APIs and web store. You want to use Identity and Access Management (IAM) to ensure a secure environment.

How should you set up IAM?

- A. Use separate service accounts for each component (social media app, APIs, and web store) with basic roles to grant access.
- B. Use one service account for all components (social media app, APIs, and web store) with basic roles to grant access.
- C. Use separate service accounts for each component (social media app, APIs, and web store) with predefined or custom roles to grant access.
- Use one service account for all components (social media app, APIs, and web store) with predefined or custom roles to grant access.



Michael is the owner/operator of "Zneeks," a retail shoe store that caters to sneaker aficionados. He regularly works with customers who order small batches of custom shoes. Michael is interested in using Cymbal Direct to manufacture and ship custom batches of shoes to these customers. Reasonably tech-savvy but not a developer, Michael likes using Cymbal Direct's partner purchase portal but wants the process to be easy. What is an example of a user story that could describe Michael's persona?

What is an example of a user story that could describe Michael's persona?

- A. As a shoe retailer, Michael wants to send Cymbal Direct custom purchase orders so that batches of custom shoes are sent to his customers.
- B. Michael is a tech-savvy owner/operator of a small business.
- Zneeks is a retail shoe store that caters to sneaker aficionados.
- D. Michael is reasonably tech-savvy but needs Cymbal Direct's partner purchase portal to be easy.



Cymbal Direct has an application running on a Compute Engine instance. You need to give the application access to several Google Cloud services. You do not want to keep any credentials on the VM instance itself.

What should you do?

- A. Create a service account for each of the services the VM needs to access. Associate the service accounts with the Compute Engine instance.
- B. Create a service account and assign it the project owner role, which enables access to any needed service.
- C. Create a service account for the instance. Use Access scopes to enable access to the required services.
- D. Create a service account with one or more predefined or custom roles, which give access to the required services.

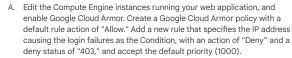
Cymbal Direct wants to use Identity and Access Management (IAM) to allow employees to have access to Google Cloud resources and services based on their job roles. Several employees are project managers and want to have some level of access to see what has been deployed. The security team wants to ensure that securing the environment and managing resources is simple so that it will scale.

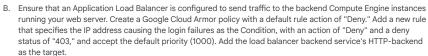
What approach should you use?

- A. Grant access by assigning custom roles to groups. Use multiple groups for better control. Give access as low in the hierarchy as possible to prevent the inheritance of too many abilities from a higher level.
- B. Grant access by assigning predefined roles to groups. Use multiple groups for better control. Give access as low in the hierarchy as possible to prevent the inheritance of too many abilities from a higher level.
- C. Give access directly to each individual for more granular control. Give access as low in the hierarchy as possible to prevent the inheritance of too many abilities from a higher level.
- D. Grant access by assigning predefined roles to groups. Use multiple groups for better control. Make sure you give out access to all the children in a hierarchy under the level needed, because child resources will not automatically inherit abilities.

You have several Compute Engine instances running NGINX and Tomcat for a web application. In your web server logs, many login failures come from a single IP address, which looks like a brute force attack.







- C. Ensure that an Application Load Balancer is configured to send traffic to the backend Compute Engine instances running your web server. Create a Google Cloud Armor policy with a default rule action of "Allow." Add a new rule that specifies the IP address causing the login failures as the Condition, with an action of "Deny" and a deny status of "403," and accept the default priority (1000). Add the load balancer backend service's HTTP-backend as the target.
- D. Ensure that an Application Load Balancer is configured to send traffic to your backend Compute Engine instances running your web server. Create a Google Cloud Armor policy using the instance's local firewall with a default rule action of "Allow." Add a new local firewall rule that specifies the IP address causing the login failures as the Condition, with an action of "Deny" and a deny status of "403," and accept the default priority (1000).



Cymbal Direct needs to make sure its new social media integration service can't be accessed directly from the public internet. You want to allow access only through the web frontend store.

How can you prevent access to the social media integration service from the outside world, but still allow access to the APIs of social media services? A. Remove external IP addresses from the VM instances running the social media service and place them in a private VPC behind Cloud NAT. Any SSH connection for should be done with Identity-Aware Proxy (IAP) host (jump box) after allowing SSH access from IAP or a corporate network.



- B. Limit access to the external IP addresses of the VM instances using firewall rules and place them in a private VPC behind Cloud NAT. Any SSH connection for management should be done with Identity-Aware Proxy (IAP) or a bastion host (jump box) after allowing SSH access from IAP or a corporate network.
- C. Limit access to the external IP addresses of the VM instances using a firewall rule to block all outbound traffic. Any SSH connection for management should be done with Identity-Aware Proxy (IAP) or a bastion host (jump box) after allowing SSH access from IAP or a corporate network.
- D. Remove external IP addresses from the VM instances running the social media service and place them in a private VPC behind Cloud NAT. Any SSH connection for management should be restricted to corporate network IP addresses by Google Cloud Armor.

Cymbal Direct is experiencing success using Google Cloud and you want to leverage tools to make your solutions more efficient. Erik, one of the original web developers, currently adds new products to your application manually. Erik has many responsibilities and requires a long lead time to add new products. You need to create a Cloud Run functions application to let Cymbal Direct employees add new products instead of waiting for Erik. However, you want to make sure that only authorized employees can use the application.

What should you do?

- Set up Cloud VPN between the corporate network and the Google Cloud project's VPC network. Allow users to connect to the Cloud Run functions instance.
- B. Use Google Cloud Armor to restrict access to the corporate network's external IP address. Configure firewall rules to allow only HTTP(S) access.
- C. Create a Google group and add authorized employees to it. Configure Identity-Aware Proxy (IAP) to the Cloud Run functions application as an HTTP-resource. Add the group as a principle with the role "Project Owner."
- D. Create a Google group and add authorized employees to it. Configure Identity-Aware Proxy (IAP) to the Cloud Run functions application as an HTTP-resource. Add the group as a principle with the role "IAP-secured Web App

You've recently created an internal Cloud Run application for developers in your organization. The application lets developers clone production Cloud SQL databases into a project specifically created to test code and deployments. Your previous process was to export a database to a Cloud Storage bucket, and then import the SQL dump into a legacy on-premises testing environment database with connectivity to Google Cloud via Cloud VPN. Management wants to incentivize using the new process with Cloud SQL for rapid testing and track how frequently rapid testing occurs.



- A. Use an ACL on the Cloud Storage bucket. Create a read-only group that only has viewer privileges, and ensure that the developers are in that group.
- B. Leave the ACLs on the Cloud Storage bucket as-is. Disable Cloud VPN, and have developers use Identity-Aware Proxy (IAP) to connect. Create an organization policy to enforce public access protection.
- C. Use predefined roles to restrict access to what the developers are allowed to do. Create a group for the developers, and associate the group with the Cloud SQL Viewer role. Remove the "cloudsql.instances.export" ability from the role.
- D. Create a custom role to restrict access to what developers are allowed to do. Create a group for the developers, and associate the group with your custom role. Ensure that the custom role does not have "cloudsql.instances.export."

3.1 Designing for security

Resources to start your journey

Google Cloud Architecture Framework: Security, privacy, and compliance

IAM best practice guides available now | Google Cloud Blog

<u>Using resource hierarchy for access control I IAM Documentation</u> Google Cloud

Chapter 18 - SRE Engagement Model

Service accounts | Compute Engine Documentation | Google Cloud

Google Cloud Armor overview

Private clusters | Kubernetes Engine Documentation | Google Cloud

<u>Understanding IAM custom roles | IAM Documentation | Google Cloud</u>





Your client is legally required to comply with the Payment Card Industry Data Security Standard (PCI-DSS). The client has formal audits already, but the audits are only done periodically. The client needs to monitor for common violations to meet those requirements more easily. The client does not want to replace audits but wants to engage in continuous compliance and catch violations early.

What would you recommend that this client do?

- A. Enable the Security Command Center (SCC) dashboard, asset discovery, and Security Health Analytics in the Premium tier. Export or view the PCI-DSS Report from the SCC dashboard's Compliance tab.
- B. Enable the Security Command Center (SCC) dashboard, asset discovery, and Security Health Analytics in the Standard tier. Export or view the PCI-DSS Report from the SCC dashboard's Compliance tab.
- C. Enable the Security Command Center (SCC) dashboard, asset discovery, and Security Health Analytics in the Premium tier. Export or view the PCI-DSS Report from the SCC dashboard's Vulnerabilities tab.
- D. Enable the Security Command Center (SCC) dashboard, asset discovery, and Security Health Analytics in the Standard tier. Export or view the PCI-DSS Report from the SCC dashboard's Vulnerabilities tab.

32 Designing for compliance

Resources to start your journey

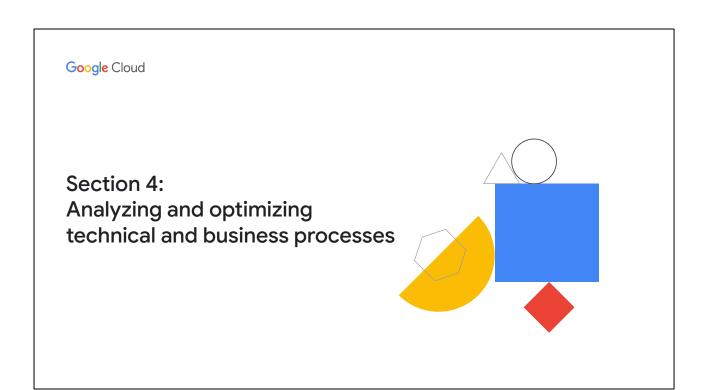
Manage compliance obligations | Architecture Framework | Google Cloud

Cloud Compliance & Regulations Resources

Assuring Compliance in the Cloud

Security Command Center | Google Cloud







You are asked to implement a lift and shift operation for Cymbal Direct's Social Media Highlighting service. You compose a Terraform configuration file to build all the necessary Google Cloud resources.

What is the next step in the Terraform workflow for this effort?

- A. Commit the configuration file to your software repository.
- B. Run terraform plan to verify the contents of the Terraform configuration file.
- C. Run terraform apply to deploy the resources described in the configuration file.
- D. Run terraform init to download the necessary provider modules.



You have implemented a manual CI/CD process for the container services required for the next implementation of the Cymbal Direct's Drone Delivery project. You want to automate the process.

What should you do?

- A. Implement and reference a source repository in your Cloud Build configuration file.
- B. Implement a build trigger that applies your build configuration when a new software update is committed to Cloud Source Repositories.
- C. Specify the name of your Container Registry in your Cloud Build configuration.
- D. Configure and push a manifest file into an environment repository in Cloud Source Repositories.



You have an application implemented on Compute Engine. You want to increase the durability of your application.

What should you do?

- A. Implement a scheduled snapshot on your Compute Engine instances.
- B. Implement a regional managed instance group.
- C. Monitor your application's usage metrics and implement autoscaling.
- D. Perform health checks on your Compute Engine instances.



Developers on your team frequently write new versions of the code for one of your applications. You want to automate the build process when updates are pushed to Cloud Source Repositories.

- A. Implement a Cloud Build configuration file with build steps.
- B. Implement a build trigger that references your repository and branch.
- C. Set proper permissions for Cloud Build to access deployment resources.
- D. Upload application updates and Cloud Build configuration files to Cloud Source Repositories.

What should you do?



Your development team used Cloud Source Repositories, Cloud Build, and Artifact Registry to successfully implement the build portion of an application's CI/CD process.. However, the deployment process is erroring out. Initial troubleshooting shows that the runtime environment does not have access to the build images. You need to advise the team on how to resolve the issue.

What could cause this problem?

- A. The runtime environment does not have permissions to the Artifact Registry in your current project.
- B. The runtime environment does not have permissions to Cloud Source Repositories in your current project.
- C. The Artifact Registry might be in a different project.
- D. You need to specify the Artifact Registry image by name.



You are implementing a disaster recovery plan for the cloud version of your drone solution. Sending videos to the pilots is crucial from an operational perspective.

What design pattern should you choose for this part of your architecture?

- A. Hot with a low recovery time objective (RTO)
- B. Warm with a high recovery time objective (RTO)
- C. Cold with a low recovery time objective (RTO)
- D. Hot with a high recovery time objective (RTO)



The number of requests received by your application is nearing the maximum specified in your design. You want to limit the number of incoming requests until the system can handle the workload.

What design pattern does this situation describe?

- A. Applying a circuit breaker
- B. Applying exponential backoff
- C. Increasing jitter
- D. Applying graceful degradation



The pilot subsystem in your Delivery by Drone service is critical to your service. You want to ensure that connections to the pilots can survive a VM outage without affecting connectivity.

What should you do?

- A. Configure proper startup scripts for your VMs.
- B. Deploy a load balancer to distribute traffic across multiple machines.
- C. Create persistent disk snapshots.
- D. Implement a managed instance group and load balancer..



Cymbal Direct wants to improve its drone pilot interface. You want to collect feedback on proposed changes from the community of pilots before rolling out updates systemwide.

What type of deployment pattern should you implement?

- A. You should implement canary testing.
- B. You should implement A/B testing.
- C. You should implement a blue/green deployment.
- D. You should implement an in-place release.

4.1 Analyzing and defining technical processes

Resources to start your journey

Securing the software development lifecycle with Cloud Build and SLSA

CI/CD with Google Cloud

Site Reliability Engineering

DevOps tech: Continuous testing | Google Cloud

Application deployment and testing strategies | Cloud Architecture Center

Chapter 17 - Testing for Reliability

Service Catalog documentation | Google Cloud

What is Disaster Recovery? | Google Cloud

API design guide





You want to establish procedures for testing the resilience of the delivery-by-drone solution.

How would you simulate a scalability issue?

- A. Block access to storage assets in one of your zones.
- B. Inject a bad health check for one or more of your resources.
- C. Load test your application to see how it responds.
- D. Block access to all resources in a zone.

4.3 Developing procedures to ensure reliability of solutions in production

Resources to start your journey

Site Reliability Engineering

Site Reliability Engineering (SRE) | Google Cloud

Patterns for scalable and resilient apps | Cloud Architecture Center

How to achieve a resilient IT strategy with Google Cloud

Patterns for scalable and resilient apps | Cloud Architecture Center

Disaster recovery planning guide | Cloud Architecture Center



Google Cloud

The diagnostic question that you just reviewed tested your knowledge of one aspect of developing procedures to ensure reliability of solutions in production. These are some links to learn more.

https://sre.google/

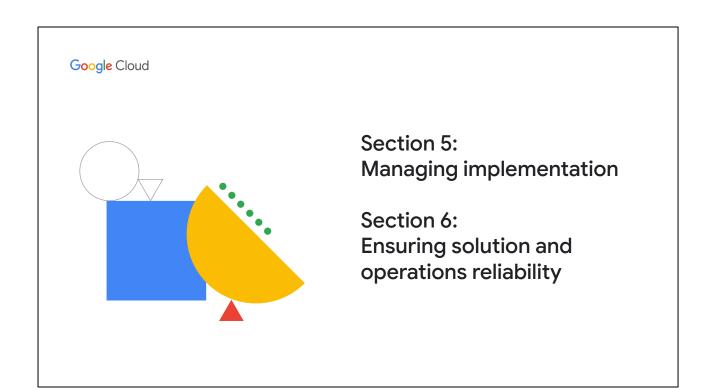
https://cloud.google.com/sre#section-6

https://cloud.google.com/architecture/scalable-and-resilient-apps

https://cloud.google.com/blog/topics/inside-google-cloud/rethinking-business-resilience-with-google-cloud

https://cloud.google.com/architecture/scalable-and-resilient-apps#test_your_resilience

https://cloud.google.com/architecture/dr-scenarios-planning-guide





Cymbal Direct is working on a social media integration service in Google Cloud. Mahesh is a non-technical manager who wants to ensure that the project doesn't exceed the budget and responds quickly to unexpected cost increases. You need to set up access and billing for the project.

What should you do?

- A. Assign the predefined Billing Account Administrator role to Mahesh. Create a project budget. Configure billing alerts to be sent to the Billing Administrator. Use resource quotas to cap how many resources can be deployed.
- B. Assign the predefined Billing Account Administrator role to Mahesh. Create a project budget. Configure billing alerts to be sent to the Project Owner. Use resource quotas to cap how much money can be spent.
- C. Use the predefined Billing Account Administrator role for the Billing Administrator group, and assign Mahesh to the group. Create a project budget. Configure billing alerts to be sent to the Billing Administrator. Use resource quotas to cap how many resources can be deployed.
- D. Use the predefined Billing Account Administrator role for the Billing Administrator group, and assign Mahesh to the group. Create a project budget. Configure billing alerts to be sent to the Billing Account Administrator. Use resource quotas to cap how much money can be spent.

Your organization is planning a disaster recovery (DR) strategy. Your stakeholders require a recovery time objective (RTO) of 0 and a recovery point objective (RPO) of 0 for zone outage. They require an RTO of 4 hours and an RPO of 1 hour for a regional outage. Your application consists of a web application and a backend MySQL database. You need the most efficient solution to meet your recovery KPIs.

What should you do?

- A. Use a global external Application Load Balancer. Deploy the web application as Compute Engine managed instance groups (MIG) in two regions, us-west and us-east. Configure the load balancer to use both backends. Use Cloud SQL with high availability (HA) enabled in us-east and a cross-region replica in us-west.
- B. Use a global external Application Load Balancer. Deploy the web application as Compute Engine managed instance groups (MIG) in two regions, us-west and us-east. Configure the load balancer to the us-east backend. Use Cloud SQL with high availability (HA) enabled in us-east and a cross-region replica in us-west. Manually promote the us-west Cloud SQL instance and change the load balancer backend to us-west.
- C. Use a global external Application Load Balancer. Deploy the web application as Compute Engine managed instance groups (MIG) in two regions, us-west and us-east. Configure the load balancer to use both backends. Use Cloud SQL with high availability (HA) enabled in us-east and back up the database every hour to a multi-region Cloud Storage bucket. Restore the data to a Cloud SQL database in us-west if there is a failure.
- D. Use a global external Application Load Balancer. Deploy the web application as Compute Engine managed instance groups (MIG) in two regions, us-west and us-east. Configure the load balancer to use both backends. Use Cloud SQL with high availability (HA) enabled in us-east and back up the database every hour to a multi-region Cloud Storage bucket. Restore the data to a Cloud SQL database in us-west if there is a failure and change the load balancer backend to us-west.

5.1 Advising development/operation team(s) to ensure successful deployment of the solution

Resources to start your journey

<u>Cloud Reference Architectures and Diagrams | Cloud Architecture</u> Center

What is DevOps? Research and Solutions | Google Cloud

Develop and deliver apps with Cloud Code, Cloud Build, Google Cloud Deploy, and GKE | Cloud Architecture Center

Google Cloud API design tips

DevOps tech: Continuous testing | Google Cloud

DevOps tech: Test data management | Google Cloud

Testing Overview | Cloud Run Functions Documentation

Database Migration Service | Google Cloud

Cloud Migration Products & Services



Your environment has multiple projects used for development and testing. Each project has a budget, and each developer has a budget. A personal budget overrun can cause a project budget overrun. Several developers are creating resources for testing as part of their CI/CD pipeline but are not deleting these resources after their tests are complete. If the compute resource fails during testing, the test can be run again. You want to reduce costs and notify the developer when a personal budget overrun causes a project budget overrun.

What should you do?

- A. Configure billing export to BigQuery. Create a Google Cloud budget for each project. Create a group for the developers in each project, and add them to the appropriate group. Create a notification channel for each group. Configure a billing alert to notify the group when their budget is exceeded. Modify the build scripts/pipeline to label all resources with the label "creator" set to the developer's email address. Use spot (preemptible) instances wherever possible.
- B. Configure billing export to BigQuery. Create a Google Cloud budget for each project. Configure a billing alert to notify billing admins and users when their budget is exceeded. Modify the build scripts/pipeline to label all resources with the label "creator" set to the developer's email address. Use spot (preemptible) instances wherever possible.
- C. Configure billing export to BigQuery. Create a Google Cloud budget for each project. Create a Pub/Sub topic for developer-budget-notifications. Create a Cloud Run function to notify the developer based on the labels. Modify the build scripts/pipeline to label all resources with the label "creator" set to the developer's email address. Use spot (preemptible) instances wherever possible.
- D. Configure billing export to BigQuery. Create a Google Cloud budget for each project. Create a Pub/Sub topic for developer-budget-notifications. Create a Cloud Run function to notify the developer based on the labels. Modify the build scripts/pipeline to label all resources with the label "creator" set to the developer's email address. Use spot (preemptible) instances wherever possible. Use Cloud Scheduler to delete resources older than 24 hours in each project.

5.2 Interacting with Google Cloud programmatically

Resources to start your journey

gcloud CLI overview | Google Cloud CLI Documentation

How Cloud Shell works

Google Cloud APIs

Testing apps locally with the emulator | Pub/Sub Documentation

Connect your app and start prototyping | Firebase Documentation

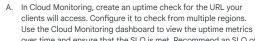
Use the emulator | Bigtable Documentation

Using the Spanner Emulator



Your client has adopted a multi-cloud strategy that uses a virtual machine-based infrastructure. The client's website serves users across the globe. The client needs a single dashboard view to monitor performance in their AWS and Google Cloud environments. Your client previously experienced an extended outage and wants to establish a monthly service level objective (SLO) of no outage longer than an hour.

What should you do?





- C. Authorize access to your Google Cloud project from AWS with a service account. Install the monitoring agent on AWS EC2 (virtual machines) and Compute Engine instances. Use Cloud Monitoring to create dashboards that use the performance metrics from virtual machines to ensure that the SLO is met.
- D. Create a new project to use as an AWS connector project. Authorize access to the project from AWS with a service account. Install the monitoring agent on AWS EC2 (virtual machines) and Compute Engine instances. Use Cloud Monitoring to create dashboards that use the performance metrics from virtual machines to ensure that the SLO is met.



Cymbal Direct uses a proprietary service to manage on-call rotation and alerting. The on-call rotation service has an API for integration. Cymbal Direct wants to monitor its environment for service availability and ensure that the correct person is notified.

What should you do?

- A. Ensure that VPC firewall rules allow access from the IP addresses used by Google Cloud's uptime-check servers. Create a Pub/Sub topic for alerting as a monitoring notification channel in Google Cloud Observability. Create an uptime check for the appropriate resource's internal IP address, with an alerting policy set to use the Pub/Sub topic. Create a Cloud Run function that subscribes to the Pub/Sub topic to send the alert to the on-call API.
- B. Ensure that VPC firewall rules allow access from the IP addresses used by Google Cloud's uptime-check servers. Create a Pub/Sub topic for alerting as a monitoring notification channel in Google Cloud Observability. Create an uptime check for the appropriate resource's external IP address, with an alerting policy set to use the Pub/Sub topic. Create a Cloud Run function that subscribes to the Pub/Sub topic to send the alert to the on-call API.
- C. Ensure that VPC firewall rules allow access from the on-call API. Create a Cloud Run function to send the alert to the on-call API. Add Cloud Run functions as a monitoring notification channel in Google Cloud Observability. Create an uptime check for the appropriate resource's external IP address, with an alerting policy set to use the Cloud Run function.
- D. Ensure that VPC firewall rules allow access from the IP addresses used by Google Cloud's uptime-check servers. Add the URL for the on-call rotation API as a monitoring notification channel in Google Cloud Observability. Create an uptime check for the appropriate resource's internal IP address, with an alerting policy set to use the API.



Cymbal Direct releases new versions of its drone delivery software every 1.5 to 2 months. Although most releases are successful, you have experienced three problematic releases that made drone delivery unavailable while software developers rolled back the release. You want to increase the reliability of software releases and prevent similar problems in the future.

What should you do?

- A. Adopt a "waterfall" development process. Maintain the current release schedule. Ensure that documentation explains how all the features interact. Ensure that the entire application is tested in a staging environment before the release. Ensure that the process to roll back the release is documented. Use Cloud Monitoring, Cloud Logging, and Cloud Alerting to ensure visibility.
- B. Adopt a "waterfall" development process. Maintain the current release schedule. Ensure that documentation explains how all the features interact. Automate testing of the application. Ensure that the process to roll back the release is well documented. Use Cloud Monitoring, Cloud Logging, and Cloud Alerting to ensure visibility.
- C. Adopt an "agile" development process. Maintain the current release schedule. Automate build processes from a source repository. Automate testing after the build process. Use Cloud Monitoring, Cloud Logging, and Cloud Alerting to ensure visibility. Deploy the previous version if problems are detected and you need to roll back.
- D. Adopt an "agile" development process. Reduce the time between releases as much as possible. Automate the build process from a source repository, which includes versioning and self-testing. Use Cloud Monitoring, Cloud Logging, and Cloud Alerting to ensure visibility. Use a canary deployment to detect issues that could cause rollback.



6.3 Diagnostic Question 07

Cymbal Direct's warehouse and inventory system was written in Java. The system uses a microservices architecture in GKE and is instrumented with Zipkin. Seemingly at random, a request will be 5-10 times slower than others. The development team tried to reproduce the problem in testing, but failed to determine the cause of the issue.

What should you do?

- A. Create metrics in Cloud Monitoring for your microservices to test whether they are intermittently unavailable or slow to respond to HTTPS requests. Use Cloud Profiler to determine which functions/methods in your application's code use the most system resources. Use Cloud Trace to identify slow requests and determine which microservices/calls take the most time to respond.
- B. Create metrics in Cloud Monitoring for your microservices to test whether they are intermittently unavailable or slow to respond to HTTPS requests. Use Cloud Trace to determine which functions/methods in your application's code use the most system resources. Use Cloud Profiler to identify slow requests and determine which microservices/calls take the most time to respond.
- C. Use Error Reporting to test whether your microservices are intermittently unavailable or slow to respond to HTTPS requests. Use Cloud Profiler to determine which functions/methods in your application's code use the most system resources. Use Cloud Trace to identify slow requests and determine which microservices/calls take the most time to respond.
- D. Use Error Reporting to test whether your microservices are intermittently unavailable or slow to respond to HTTPS requests. Use Cloud Trace to determine which functions/methods in your application's code Use the most system resources. Use Cloud Profiler to identify slow requests and determine which microservices/calls take the most time to respond.



6.4 Diagnostic Question 08

Cymbal Direct has a new social media integration service that pulls images of its products from social media sites and displays them in a gallery of customer images on your online store. You receive an alert from Cloud Monitoring at 3:34 AM on Saturday. The store is still online, but the gallery does not appear. The CPU utilization is 30% higher than expected on the VMs running the service, which causes the managed instance group (MIG) to scale to the maximum number of instances. You verify that the issue is real by checking the site and by checking the incidents timeline.







- B. Check the incident documentation or labels to determine the on-call contact. Appoint an incident commander, and open a chat channel, or conference call for emergency response. Investigate and resolve the issue by increasing the maximum number of instances in the MIG, and verify that this resolves the issue. Mark the incident as closed.
- C. Increase the maximum number of instances in the MIG and verify that this resolves the issue. Check the incident documentation or labels to determine the on-call contact. Appoint an incident commander, and open a chat channel, or conference call for emergency response. Investigate and resolve the root cause of the issue. Write a blameless post-mortem and identify steps to prevent the issue, to ensure a culture of continuous improvement.
- D. Verify the high CPU is not user impacting, increase the maximum number of instances in the MIG and verify that this resolves the issue.

6.4 Diagnostic Question 09

You need to adopt Site Reliability Engineering principles and increase visibility into your environment. You want to minimize management overhead and reduce noise generated by the information being collected. You also want to streamline the process of reacting to analyzing and improving your environment, and to ensure that only trusted container images are deployed to production.

What should you do?

- A. Adopt Google Cloud Observability to gain visibility into the environment. Use Cloud Trace for distributed tracing, Cloud Logging for logging, and Cloud Monitoring for monitoring, alerting, and dashboards. Only page the on-call contact about novel issues or events that haven't been seen before. Use GNU Privacy Guard (GPG) to check container image signatures and ensure that only signed containers are deployed.
- B. Adopt Google Cloud Observability to gain visibility into the environment. Use Cloud Trace for distributed tracing, Cloud Logging for logging, and Cloud Monitoring for monitoring, alerting, and dashboards. Page the on-call contact when issues that affect resources in the environment are detected. Use GPG to check container image signatures and ensure that only signed containers are deployed.
- C. Adopt Google Cloud Observability to gain visibility into the environment. Use Cloud Trace for distributed tracing, Cloud Logging for logging, and Cloud Monitoring for monitoring, alerting, and dashboards. Only page the on-call contact about novel issues that violate a SLO or events that haven't been seen before. Use Binary Authorization to ensure that only signed container images are deployed.
- D. Adopt Google Cloud Observability to gain visibility into the environment. Use Cloud Trace for distributed tracing, Cloud Logging for logging, and Cloud Monitoring for monitoring, alerting, and dashboards. Page the on-call contact when issues that affect resources in the environment are detected. Use Binary Authorization to ensure that only signed container images are deployed.

6.1 – 6.4 Ensuring solution and operations reliability

Resources to start your journey

Observability in Google Cloud documentation

Operations: Cloud Monitoring & Logging | Google Cloud

Continuous Delivery | Google Cloud

Concepts | Google Cloud Deploy

Adopting SLOs | Cloud Architecture Center



Analyzing a case study: Dress4Win

Dress4Win case study - 01

Company overview

Dress4Win is a web-based company that helps their users organize and manage their personal wardrobe using a web app and mobile application. The company also cultivates an active social network that connects their users with designers and retailers. They monetize their services through advertising, ecommerce, referrals, and a freemium app model. The application has grown from a few servers in the founder's garage to several hundred servers and appliances in a colocated data center. However, the capacity of their infrastructure is now insufficient for the application's rapid growth. Because of this growth and the company's desire to innovate faster, Dress4Win is committing to a full migration to a public cloud.

Solution concept

For the first phase of their migration to the cloud, Dress4Win is moving their development and test environments. They are also building a disaster recovery site, because their current infrastructure is at a single location. They are not sure which components of their architecture they can migrate as is and which components they need to change before migrating them.

Existing technical environment

Dress4Win case study - 02

The Dress4Win application is served out of a single data center location. All servers run Ubuntu LTS v16.04.

Databases:

MySQL. One server for user data, inventory, static data MySQL 5.7 8 core CPUs 128 GB of RAM 2x 5 TB HDD (RAID 1)

Storage appliances:

iSCSI for VM hosts Fibre channel SAN - MySQL databases 1 PB total storage; 400 TB available NAS - image storage, logs, backups 100 TB total storage; 35 TB available

Compute:

40 web application servers providing micro-services based APIs and static content
Tomcat - Java
Nginx
Four core CPUs
32 GB of RAM
20 Apache Hadoop/Spark servers:
Data analysis
Real-time trending calculations
Eight core CPUs
128 GB of RAM

4x 5 TB HDD (RAID 1)

Three RabbitMQ servers for messaging, social notifications, and events
Eight core CPUs
32GB of RAM
Miscellaneous servers:
Jenkins, monitoring, bastion hosts, security scanners
Eight core CPUs
32GB of RAM

Dress4Win case study - 03

Business requirements

- Build a reliable and reproducible environment with scaled parity of production
- Improve security by defining and adhering to a set of security and identity and access management (IAM) best practices for cloud
- Improve business agility and speed of innovation through rapid provisioning of new resources
- Analyze and optimize architecture for performance in the cloud

Technical requirements

- Easily create non-production environments in the cloud
- Implement an automation framework for provisioning resources in cloud
- Implement a continuous deployment process for deploying applications to the on-premises data center or cloud
- Support failover of the production environment to cloud during an emergency
- Encrypt data on the wire and at rest
- Support multiple private connections between the production data center and cloud environment

Dress4Win case study - 04

Executive statement

Our investors are concerned about our ability to scale and contain costs with our current infrastructure. They are also concerned that a competitor could use a public cloud platform to offset their up-front investment and free them to focus on developing better features. Our traffic patterns are highest in the mornings and weekend evenings; during other times, 80% of our capacity is sitting idle.

Our capital expenditure is now exceeding our quarterly projections. Migrating to the cloud will likely cause an initial increase in spending, but we expect to fully transition before our next hardware refresh cycle. Our total cost of ownership (TCO) analysis over the next five years for a public cloud strategy achieves a cost reduction between 30% and 50% over our current model.

Categorizing Objectives

Dress4Win case study - REF

Business requirements

- Build a reliable and reproducible environment with scaled parity of production
- Improve security by defining and adhering to a set of security and identity and access management (IAM) best practices for cloud
- Improve business agility and speed of innovation through rapid provisioning of new resources
- Analyze and optimize architecture for performance in the cloud

Itemized list of objectives

Technical requirements

- Easily create non-production environments in the cloud
- Implement an automation framework for provisioning resources in cloud
- Implement a continuous deployment process for deploying applications to the on-premises data center or cloud
- Support failover of the production environment to cloud during an emergency
- Encrypt data on the wire and at rest
- Support multiple private connections between the production data center and cloud environment.

Solution component

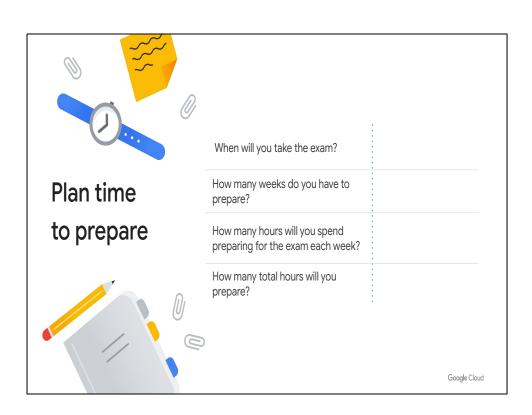
Databases:

- MySQL. One server for user data, inventory, static data
 - o MySQL 5.7
 - o 8 core CPUs
 - o 128 GB of RAM
 - o 2x 5 TB HDD (RAID 1)

Compute:

- 40 web application servers providing micro-services based APIs and static content
 - o Tomcat Java
 - Nginx
 - o Four core CPUs
 - o RAM

... more in actual case study



Weekly study plan

Now, consider what you've learned about your knowledge and skills through the diagnostic questions in this course. You should have a better understanding of what areas you need to focus on and what resources are available.

Use the template that follows to plan your study goals for each week. Consider:

- What exam guide section(s) or topic area(s) will you focus on?
- What courses (or specific modules) will help you learn more?
- What Skill Badges or labs will you work on for hands-on practice?
- What documentation links will you review?
- What additional resources will you use such as sample questions?
- What will you do to prepare for the case studies? You may do some or all of these study activities each week.

Duplicate the weekly template for the number of weeks in your individual preparation journey.



Weekly study template (example)

Area(s) of focus: Automating infrastructure with Terraform

Courses/modules to complete:

Elastic Google Cloud Infrastructure: Scaling and Automation M3 Reliable Google Cloud Infrastructure: Design and Process, M3

Skill Badges/labs to complete:

Build Infrastructure with Terraform on Google Cloud

Documentation to review:

 $\underline{\textbf{Using Recommendations for Infrastructure as Code | Recommender Documentation | Google}$

Cloud

Using Terraform with Google Cloud

 $\underline{\text{Managing infrastructure as code with Terraform, Cloud Build, and GitOps | Cloud Architecture}$

Center | Google Cloud

Additional study: Sample questions 1-3

Review case study 2 and search for relevant reference architectures

Google Cloud

If you've identified managed services as a particular area you need to study, you might choose to structure your study for a week to include targeted modules from the on-demand training, a related Skill Badge for hands-on practice, and documentation.

Alternatively, you might choose one week to complete an entire course, and another week to focus on a Skill Badge. You can determine the approach that fits your existing skillset.

Weekly study template		
Area(s) of focus:		
Courses/modules to complete:		
Skill Badges/labs to complete:		
Documentation to review:		
Additional study:		
		Google Cloud