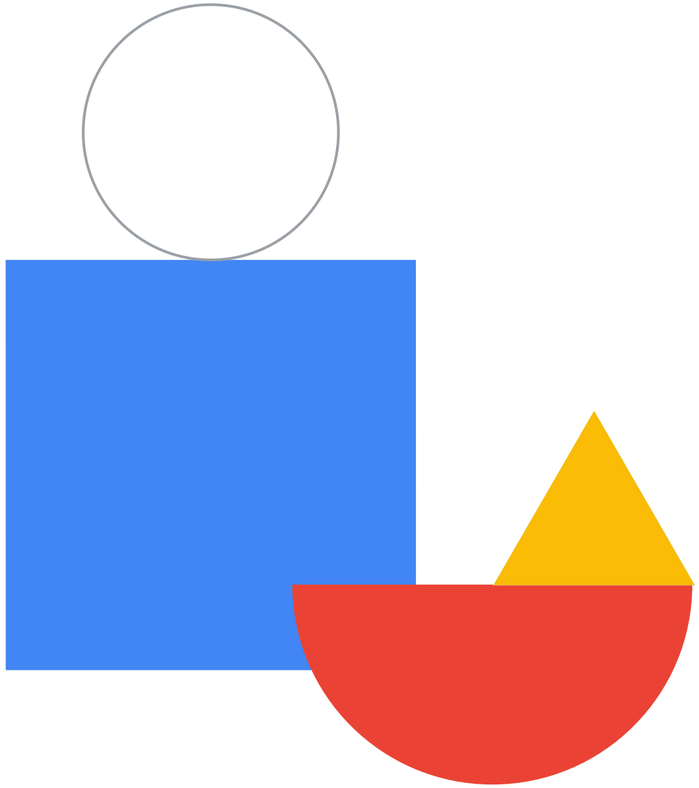
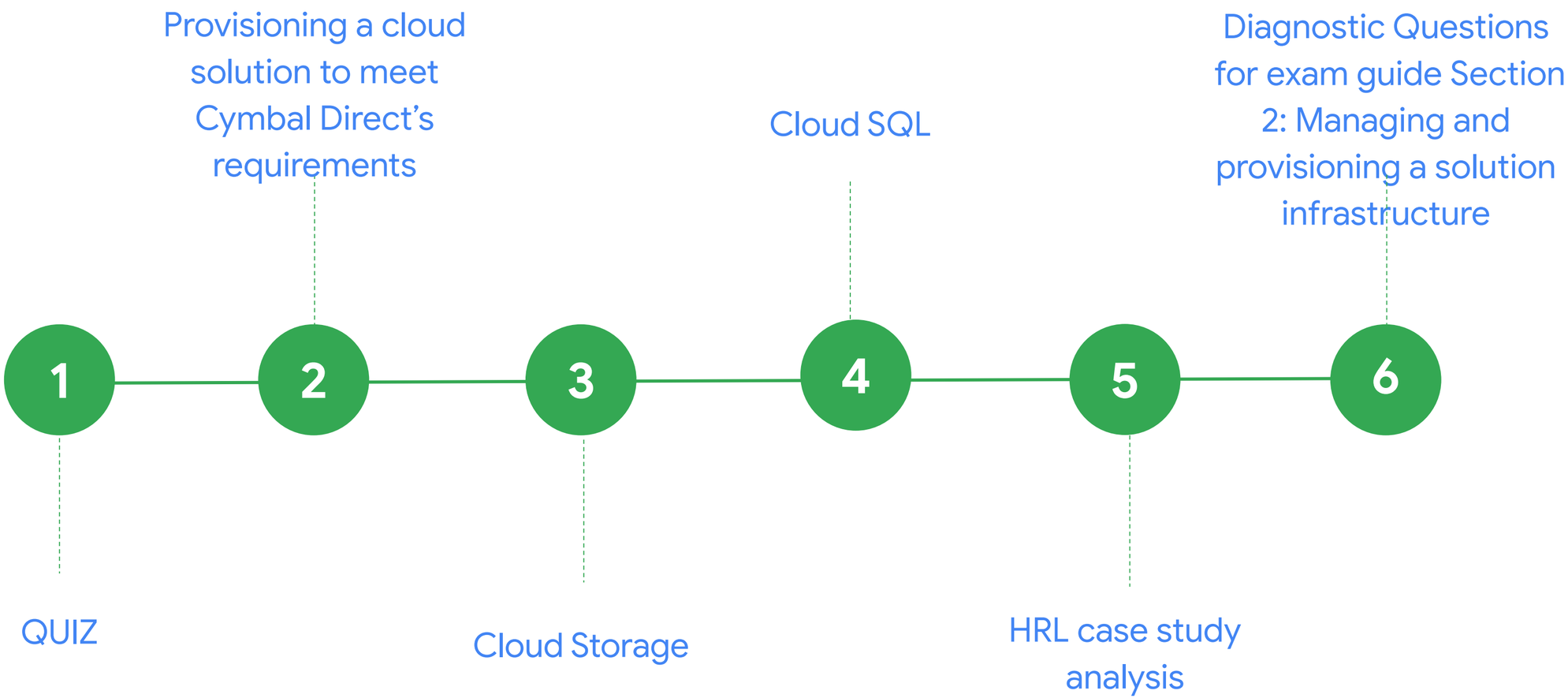
Google Cloud



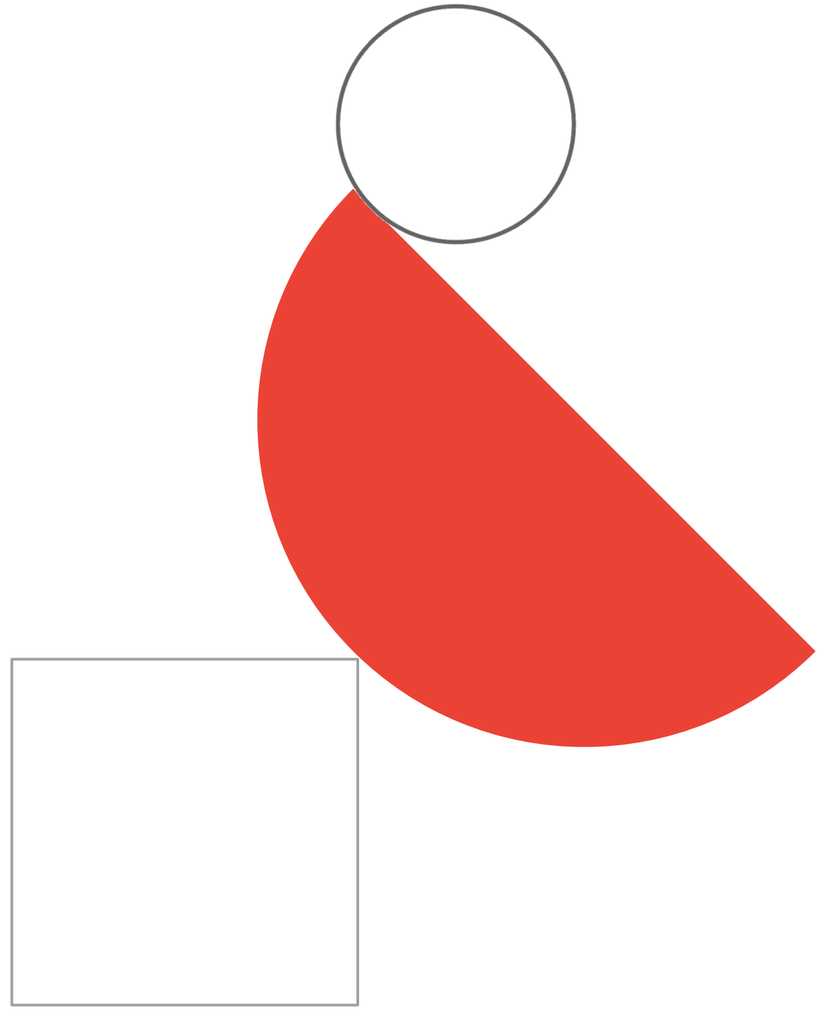
Preparing for Your Professional Cloud Architect Journey

Module 2: Managing and Provisioning a Solution Infrastructure

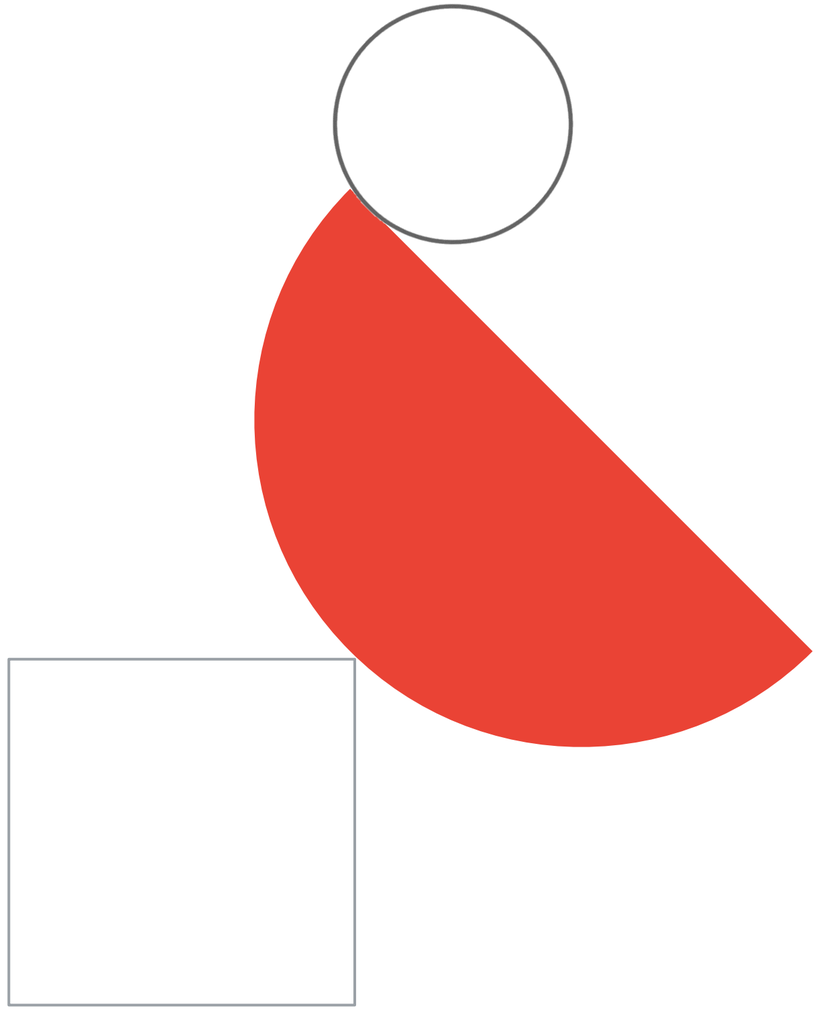
Week 3 agenda



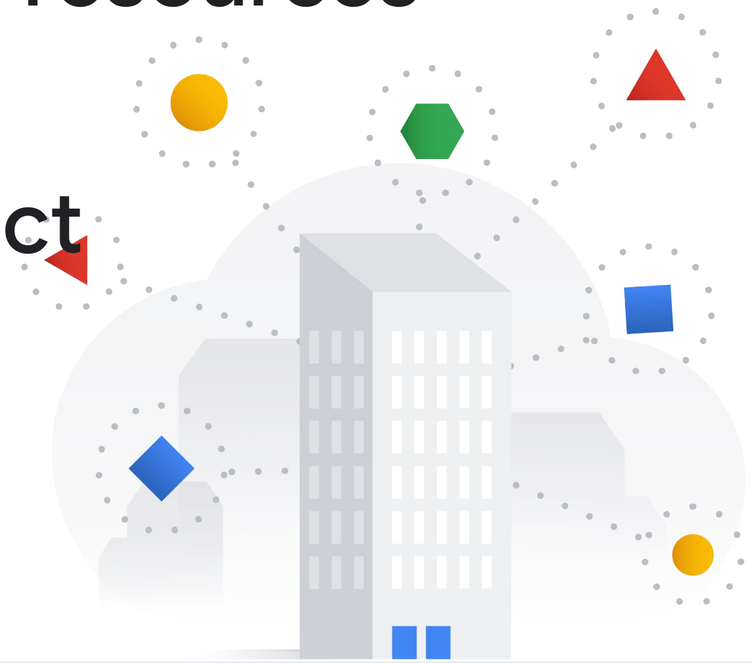
QUIZ time!



Provisioning a cloud solution to meet Cymbal Direct's requirements

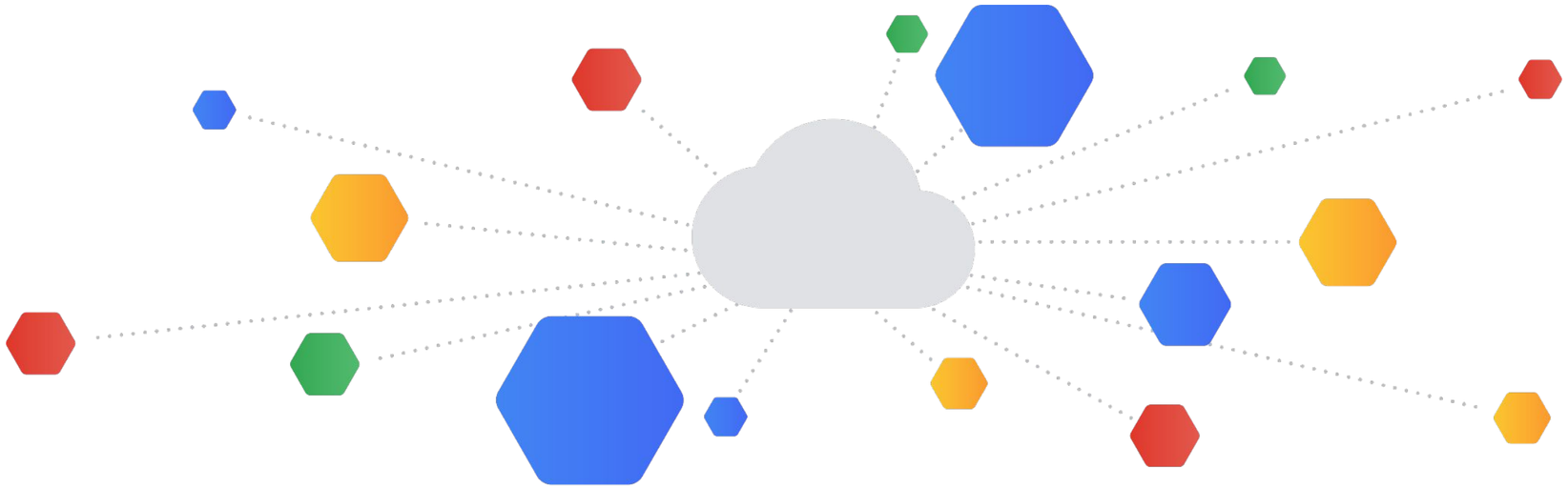


Your role in provisioning resources for Cymbal Direct



- Configuring network topologies
- Configuring individual storage systems
- Configuring compute systems

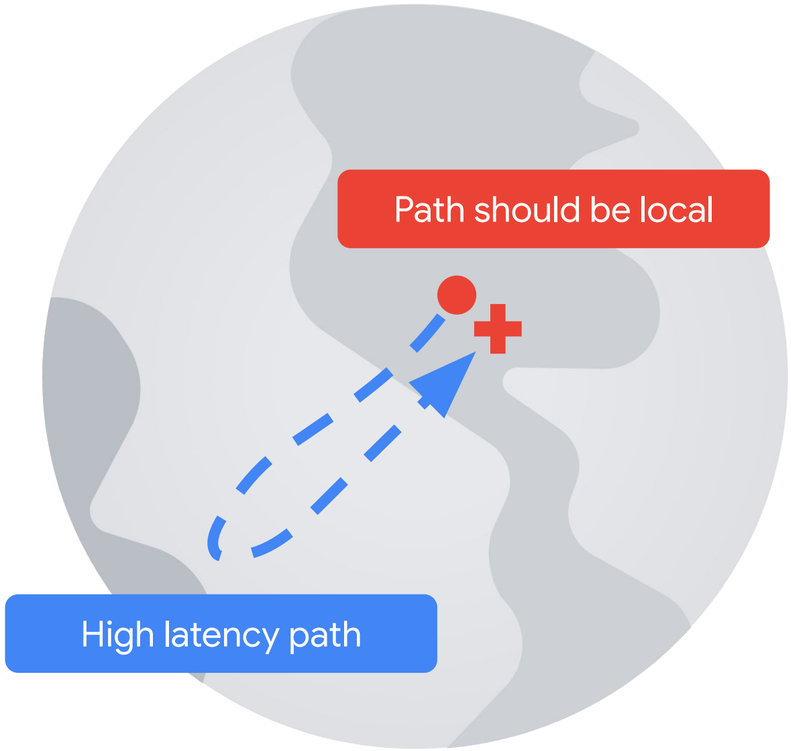
Cymbal Direct



What configuration choices will you make as a Professional Cloud Architect?

Issues with the Delivery by Drone Initiative

- Pilots complain about occasional unexpected latency.
- The service was initially set up using DNS load balancing.
- Some traffic is sent to resources further away.

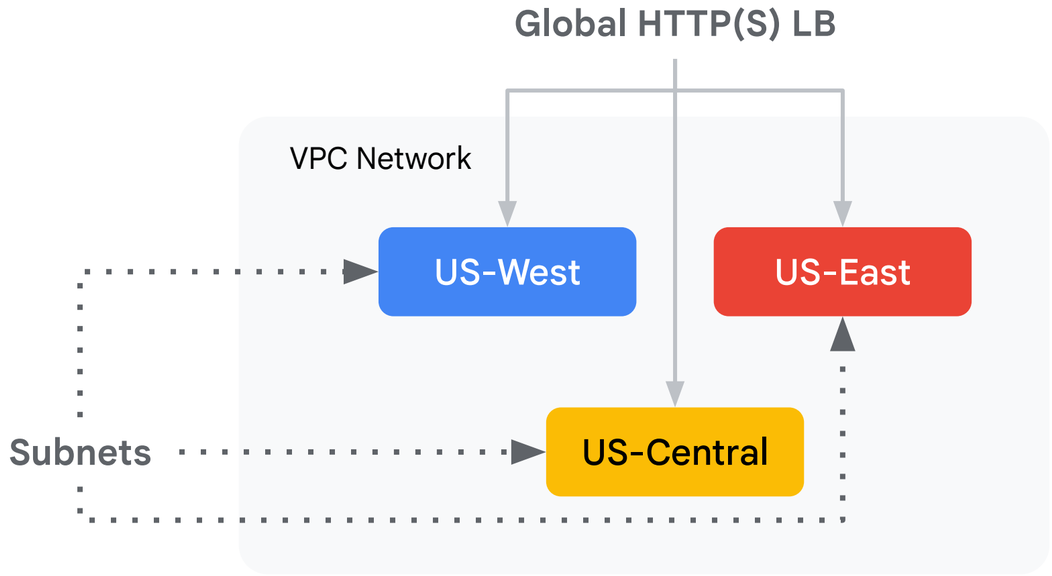


Exam Tip: *In many cases (both exam / case study related and real, it's not only about to migrate to the cloud as-is, but also: optimize processes which are not working as expected as of now.*

Provisioning networking for Delivery by Drone

You decided to:

- Move to a Global HTTP(S) Load Balancer
- Use a Custom Network
- Configure subnets only in nearby regions



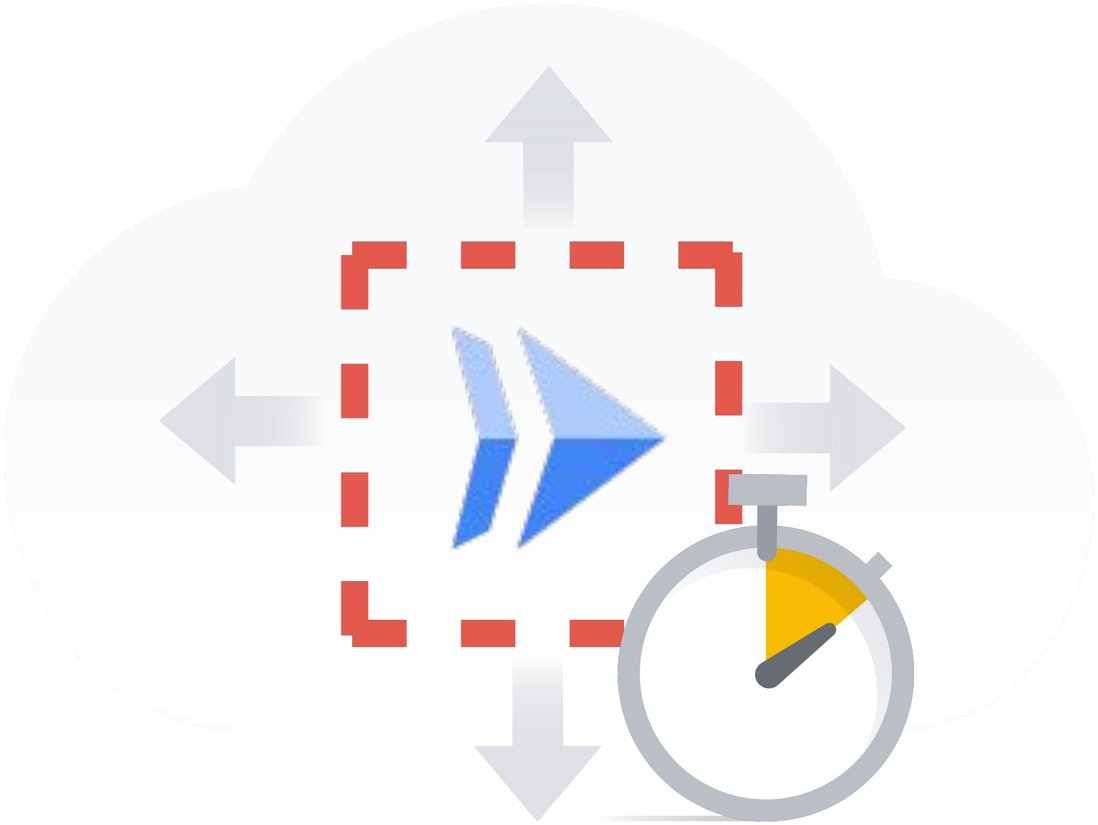
Configuring an organization policy constraint

Organization policy constraints | Resource Manager Documentation | Google Cloud

- You set `‘constraints/compute.skipDefaultNetworkCreation’` to `‘true’`
- You also configured `‘constraints/compute.restrictSharedVpcSubnetworks’` to subnets in regions near where the drones are being deployed for beta deliveries.

Most common Organization Policies

Policy Constraint	Description
compute.vmExternalIpAccess	A list of project/zone/instance names that are allowed to have external IP addresses and deny all others. Attempts to create any other VMs with an external IP address will fail.
compute.trustedImageProjects	A list of projects that contain trusted images that can be used as the basis for a VM and deny all others. Attempting to instantiate a VM with an image from another project is denied.
compute.skipDefaultNetworkCreation	Disables the creation of default VPC when creating a project. The default VPC uses auto mode subnetworks and includes default firewall rules which are often incompatible with production deployments.
iam.disableServiceAccountKeyCreation	This boolean constraint disables the creation of service account external keys where this constraint is set to `True`.
gcp.resourceLocations	This list constraint defines the set of locations where location-based GCP resources can be created. Policies for this constraint can specify multi-regions such as asia and europe, regions such as us-east1 or europe-west1, or individual zones such as europe-west1-b as allowed or denied locations.
sql.restrictPublicIp	This boolean constraint restricts configuring Public IP on Cloud SQL instances where this constraint is set to True. This constraint is not retroactive, Cloud SQL instances with existing Public IP access will still work even after this constraint is enforced. By default, Public IP access is allowed to Cloud SQL instances.
sql.disableDefaultEncryptionCreation	Restrict default Google-managed encryption on Cloud SQL instances
compute.requireShieldedVm	This boolean constraint, when set to True, requires that all new Compute Engine VM instances use Shielded disk images with Secure Boot, vTPM, and Integrity Monitoring options enabled. Secure Boot can be disabled after creation, if desired. Shielded VM features add verifiable integrity and exfiltration resistance to your VMs.



Cymbal Direct's compute configuration

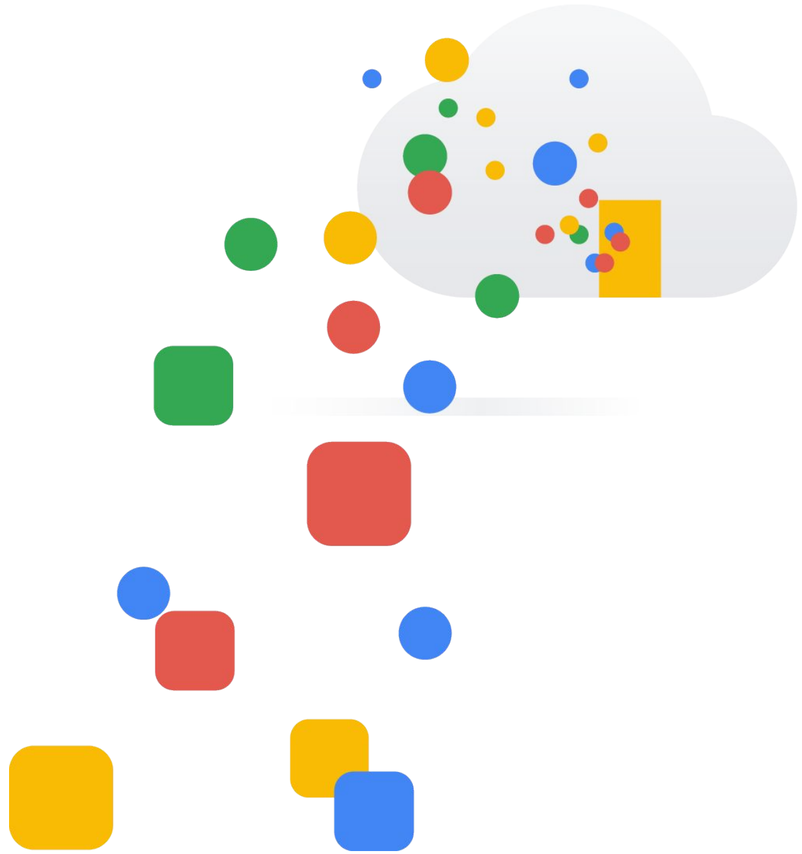
Cloud Run lets you scale with minimal latency.

You configure a high number of minimum instances to reduce cold starts.

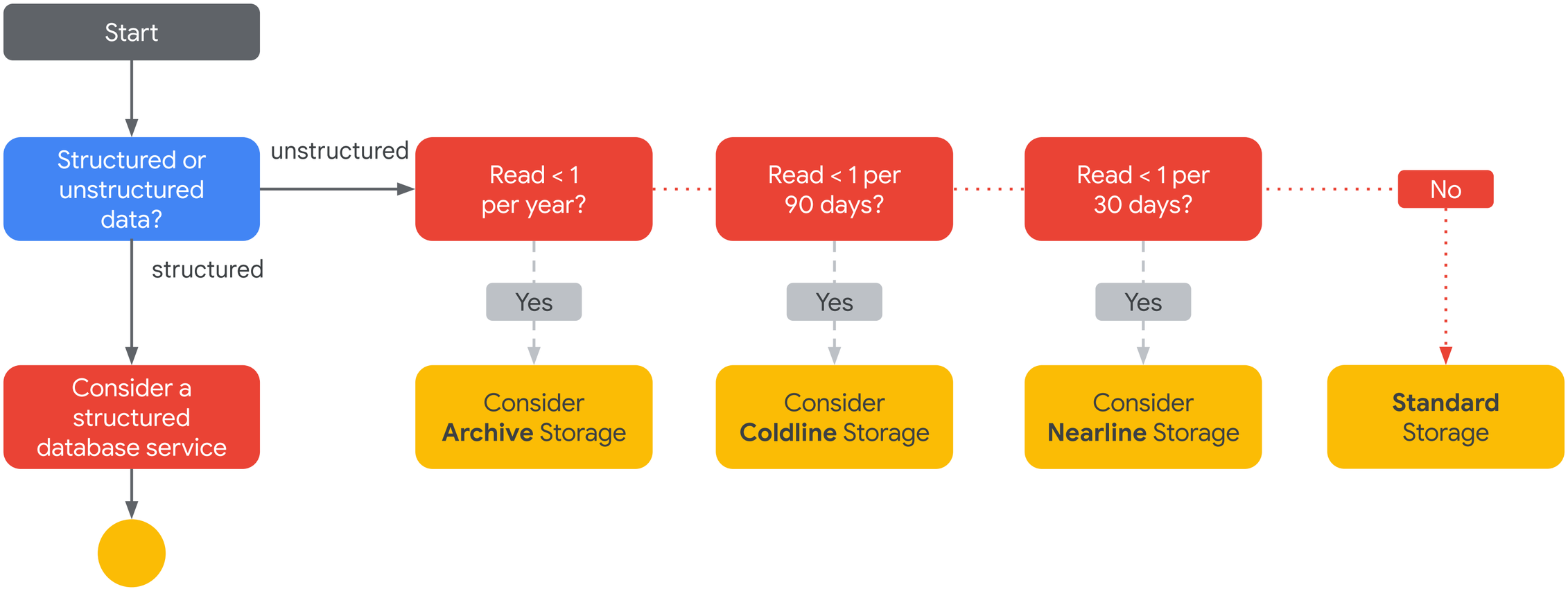
Configuration options for Cymbal Direct's other applications

For deployments that have stateful information such as databases, or where the developer prefers to preserve their existing workflow, you've chosen to leverage GKE.

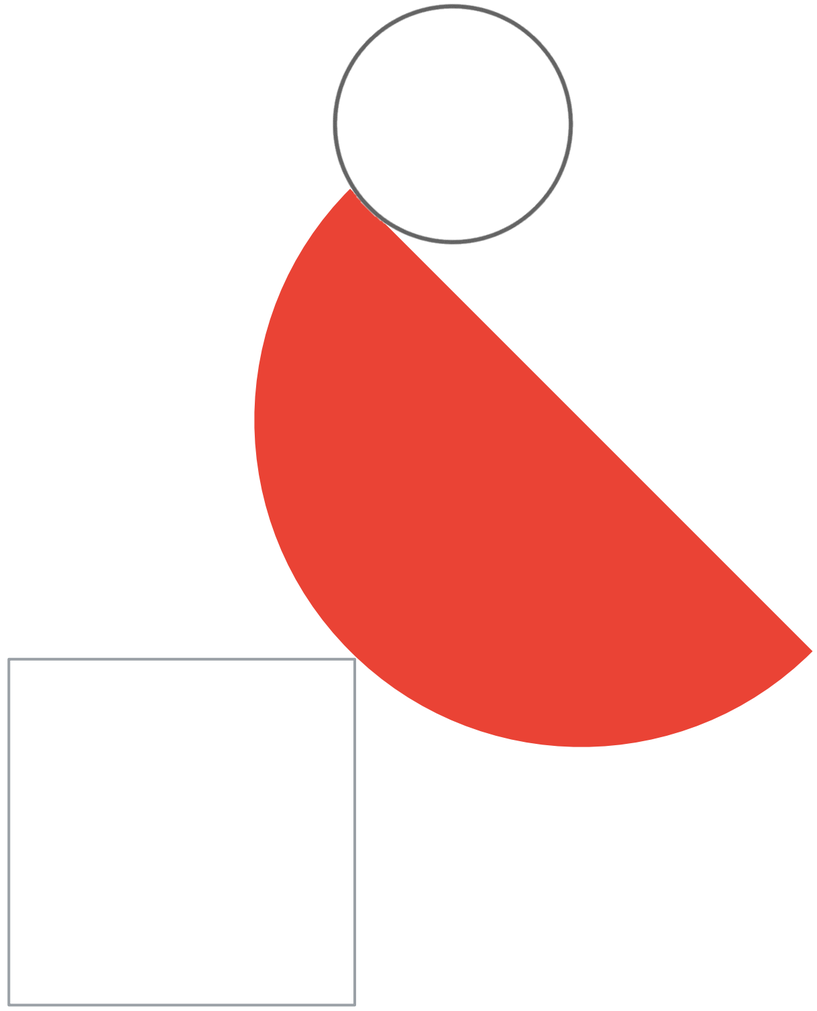
For their Compute Engine instances, you've decided to use managed instance groups with Shielded VMs.



Cymbal Direct's storage configuration



Google Cloud Storage (GCS)



Know these GCS features well!

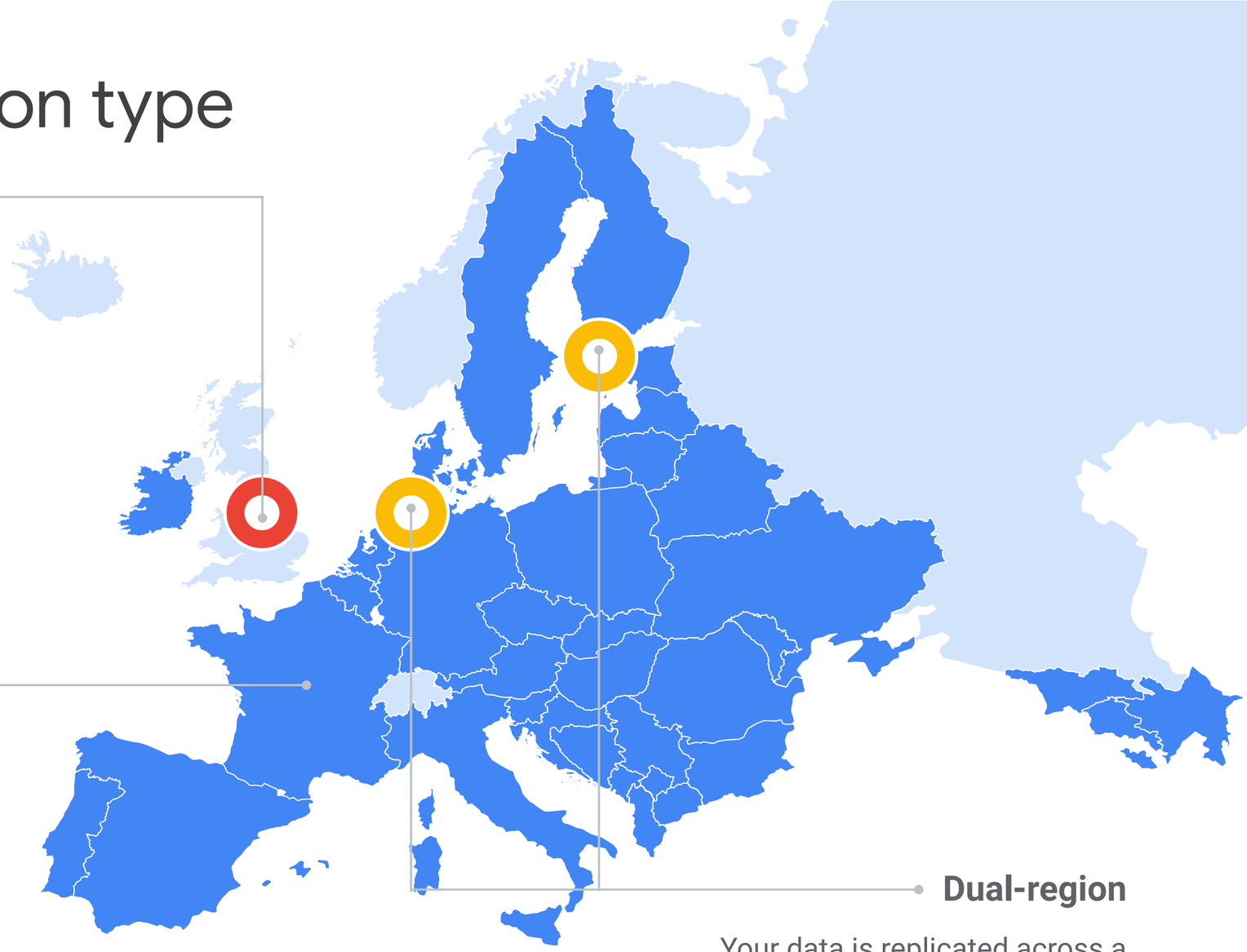
Object versioning and retention policies cannot be on at the same time. If you want to allow objects to be modified, choose object versioning. If you want to prevent deletions or changes to objects, set a retention policy.

- Controlling object lifecycle:
  - Retention policy (best for compliance)
  - Object Hold (prevent individual objects from being deleted)
  - Object Versioning (aka “automatic backups with retention policy; be aware of additional costs)
  - Object Lifecycle Management (aka “object TTL” / downgrade class to optimize costs)
- ACLs (read, write, full control on buckets or an object).
- Objects are immutable.
- Location constraints on buckets.
- Object change notifications (useful for automation, along with Pub/Sub).
- Resumable uploads (can restart from the last successful chunk).
- Strong consistency except for cached objects.
- Storage class set at object level (fine-grained performance/cost control without moving data to different buckets).
- Cloud Storage Triggers to handle events in Cloud Functions.
- Streaming uploads to GCS.
- For data encryption, GCS supports GMEK, CMEK and CSEK (most services do not support CSEK!)

GCS: Choosing a location type

Regional

Your data is stored in a specific region with replication across availability zones in that region. Good for colocating compute and storage for high performance (eg. data analytics).



Multi-Region

Your data is distributed redundantly across US, EU, or Asia. Good for serving content to end users and when you want automatic failover.

Dual-region

Your data is replicated across a specific pair of regions. Good for when you need colocated compute and storage and automatic DR.

Exam Tip:

- *Know the use-cases for each of those choices.*
- *Dual-region buckets are only available for selected regions!*

Google Cloud Storage: where to use each storage class

Standard		Nearline	Coldline	Archive
In multi-region locations for serving content globally.	**In regional locations for data accessed frequently or high throughput needs**	**For data access less than once a month**	**For data accessed roughly less than once a quarter**	**For long term retention**
Streaming videos	Video transcoding	Serving rarely accessed docs	Serve rarely used data	Regulatory archives
Images	Genomics	Backup	Movie archive	Tape replacement
Websites	General data analytics & compute		Disaster recovery	
Documents				

Google Cloud Storage: autoclass

Automatically transition objects to colder storage classes based on usage patterns and transition back to Standard on access.

Not too “intelligent” as of now.

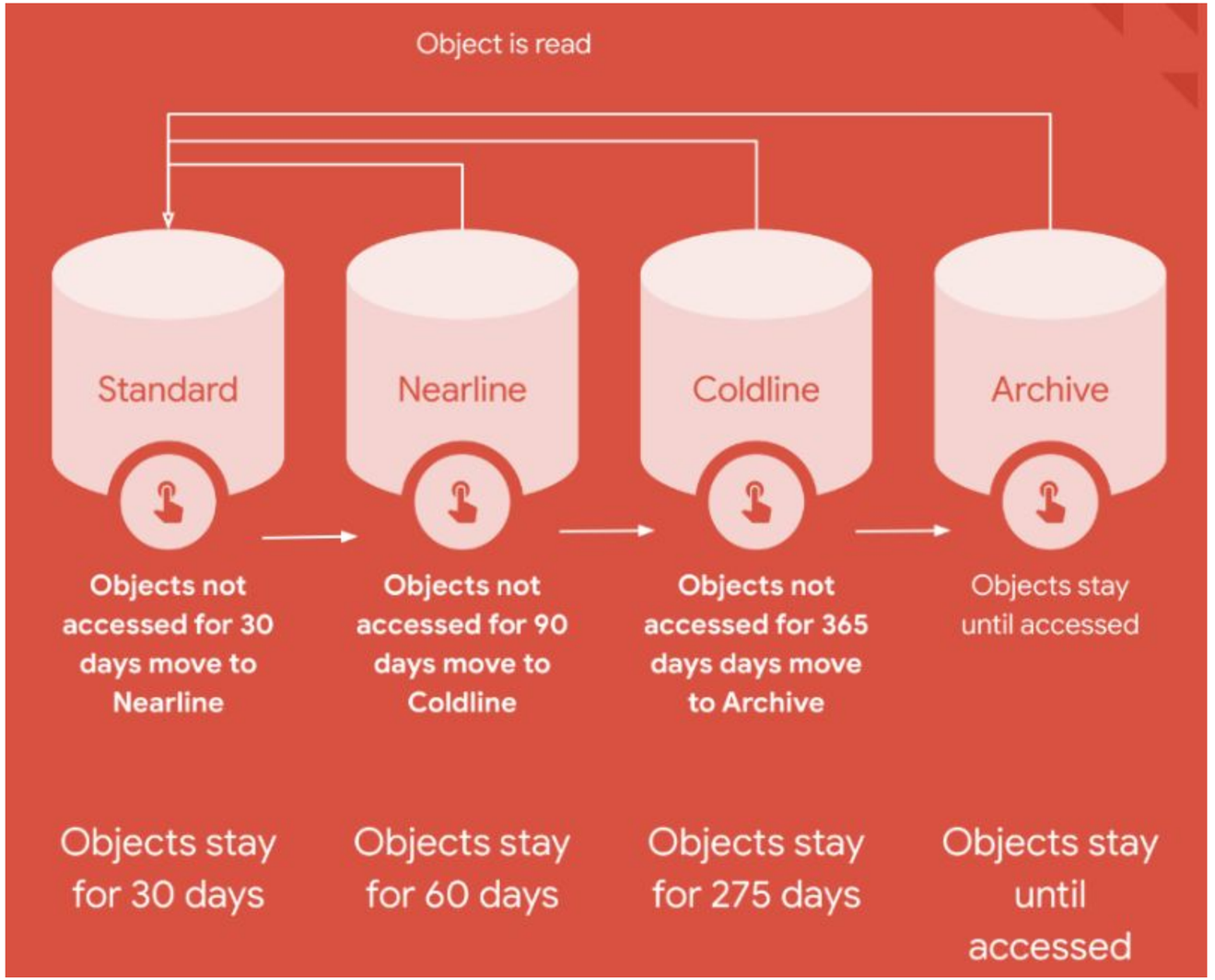
- **Choose a storage class for your data**

  A storage class sets costs for storage, retrieval, and operations, with minimal differences in uptime. Choose if you want objects to be managed automatically or specify a default storage class based on how long you plan to store your data and your workload or use case. Learn more

  (•) Autoclass
  Automatically transitions each object to hotter or colder storage based on object-level activity, to optimize for cost and latency. Recommended if usage frequency may be unpredictable. Can be changed to a default class at any time. Pricing details

  ( ) Set a default class
  Applies to all objects in your bucket unless you manually modify the class per object or set object lifecycle rules. Best when your usage is highly predictable. Can't be changed to Autoclass once the bucket is created.

  - ( ) Standard
    Best for short-term storage and frequently accessed data
  - ( ) Nearline
    Best for backups and data accessed less than once a month
  - ( ) Coldline
    Best for disaster recovery and data accessed less than once a quarter
  - ( ) Archive
    Best for long-term digital preservation of data accessed less than once a year



Exam Tip: *autoclass is a relatively new feature (GA Q4 ‘22) and will not covered on the exam yet*

Best Practices on Storage Class Selection

Consider retention period and access frequency

Retention Period

Access Frequency	<1 mo	1–3 mo	3–12 mo	>12 mo
>12/yr	Standard	Standard	Standard	Standard
4–12/yr	Standard	Nearline	Nearline	Nearline
1–4/yr	Standard	Nearline	Coldline	Coldline
<1/yr	Standard	Nearline	Coldline	Archive

Exam Tips:

- *Each storage class has so-called "minimum storage duration", so when optimizing costs, you also need to validate if you'll need to keep your objects for at least this amount of time.*

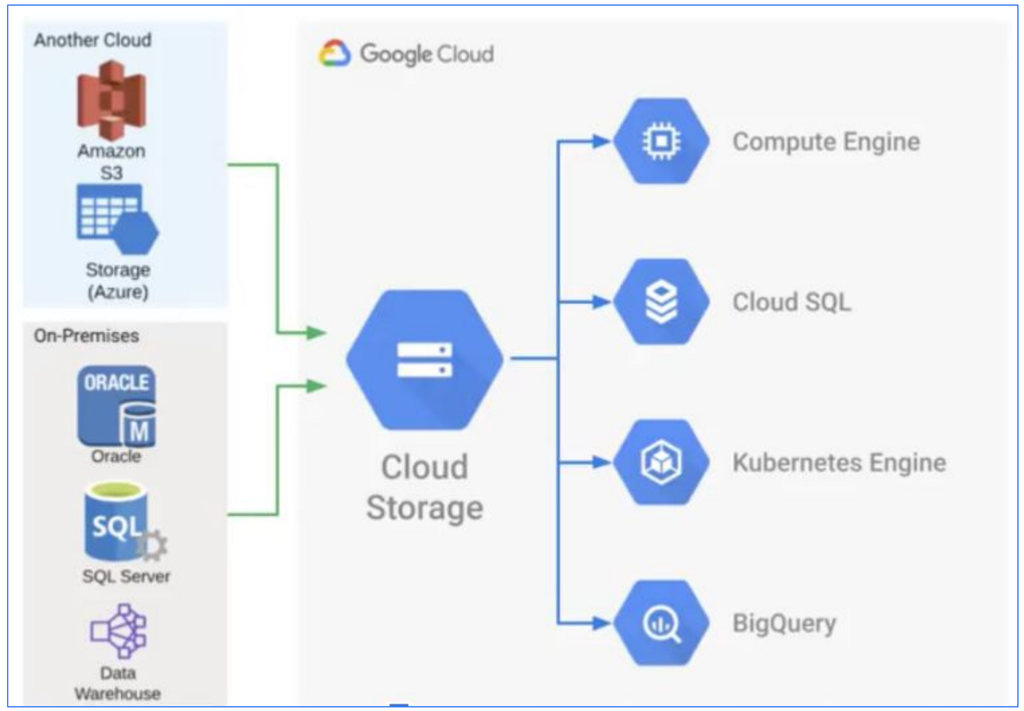
Choose the right tool to move data to GCS...

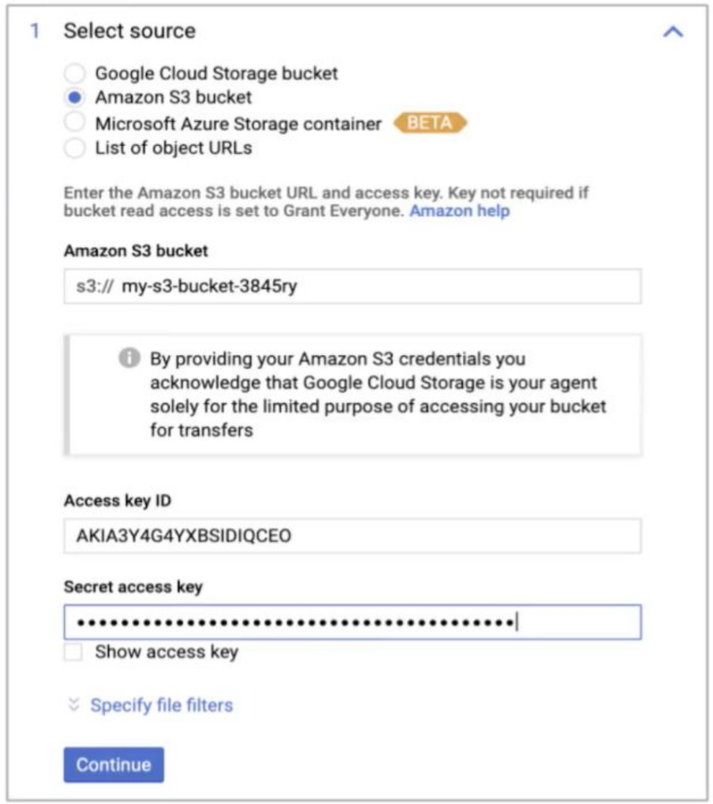
Where you're moving data from	Scenario	Suggested products
Another cloud provider (for example, Amazon Web Services or Microsoft Azure) to Google Cloud	—	Storage Transfer Service
Cloud Storage to Cloud Storage (two different buckets)	—	Storage Transfer Service
Your private data center to Google Cloud	Enough bandwidth to meet your project deadline for less than 1 TB of data	gsutil
Your private data center to Google Cloud	Enough bandwidth to meet your project deadline for more than 1 TB of data	Storage Transfer Service for on-premises data
Your private data center to Google Cloud	Not enough bandwidth to meet your project deadline	Transfer Appliance

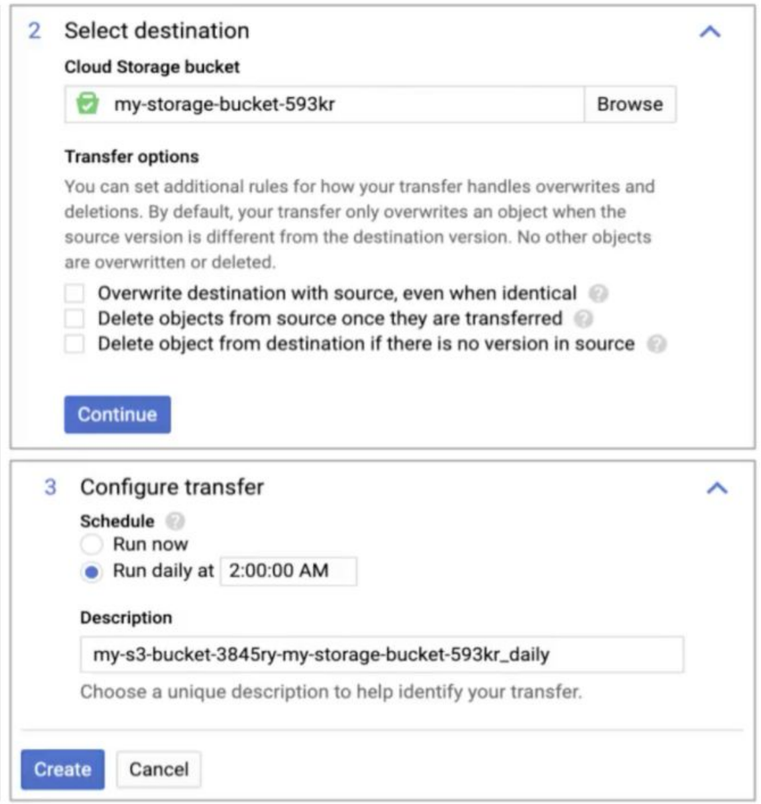
Exam Tips:

- *Depending on size and throughput, use gsutil / Transfer Service (low cost) / Transfer Appliance*
- ***When using Transfer Appliance, you need to execute so-called "rehydration" process*** *which will decrypt and uncompress before it's put to a destination bucket.*

Cloud Storage: Storage Transfer Service







Exam Tip: *Storage Transfer Service can be set up one-off (e.g., move bucket to new location) and recurring (e.g., back up from S3 to GCS with rsync-like semantics)*

GCS: most important ‘gsutil’ options

- -m: to perform a parallel (multi-threaded/multi-processing) copy;
- -J: applies gzip transport encoding to file uploads. For specifics file formats may cause longe uploads;
- -Z: Applies gzip content-encoding to file uploads prior to upload;
- -R, -r: Causes directories, buckets, and bucket subdirectories to be copied recursively.
- gsutil iam ch allUsers:objectViewer gs://my-awesome-bucket -> grant all users permission to read the objects stored in your bucket
- -o: Set/override values in the boto configuration value, in the format <section>:<name>=<value>, eg:
  - gsutil -o "GSUtil:max_upload_compression_buffer_size=8G” -m cp -j html,txt -r /local/source/dir gs://bucket/path

```
gsutil mb gs://my-bucket
gsutil cp -r ./on-prem-folder/ gs://my-bucket
gsutil ls -r gs://my-bucket
gsutil rm -r gs://my-bucket
gsutil -m rsync -r ./on-prem-folder/ gs://my-bucket
```

-m parameter enables multi-part, parallel uploads

Exam Tips:

- *“gsutil -m” option used to perform parallel uploads*
- *Be aware of .boto configuration file*

#GCPSketchnote
@PVERGADIA
THECLOUDGIRL.DEV
03.30.2021

Options to move data to Google Cloud

GCS TRANSFER TOOLS

For small transfers upto a few TBs

TRANSFER SERVICE

Large-scale online transfer

Cloud data | On-premises

TRANSFER APPLIANCE

Large scale transfers from low bandwidth networks

BIGQUERY DATA TRANSFER SERVICE

Data from SaaS & 3rd party apps

HOW DO THEY WORK?

Upload data directly to GCS

TOOLS - GSUTIL, UI, JSON API

S3 / Azure Cloud

TOOLS
UI, Client Libraries, HTTP REST API

On-premises

TOOLS
On-premise agent with GUI

1. Request Appliance
2. Upload data SFTP, SCP or NFS
3. Ship to Google
4. GCS
5. Erase NIST 800 standard

S3
Teradata
RedShift
Google SaaS apps
GCS
3rd party

SOME COOL THINGS TO KNOW

GSUTIL - Fast multi threaded

rsync

Streaming

Scheduled & Incremental updates
Scale to 10's of Gbps

Data moves over Google's high-bandwidth network pipes

Minimize transfer time max available bandwidth performance optimizations

All SSDs
Minimal software
Performance features
Multiple network connectivity options

Automatic Schedule

Mirror or Append data

SECURITY

HTTPS encrypted TLS connections

Cloud sources accessed using private credentials

All bytes checksummed & encrypted in-flight

AES 256 encryption on the appliance
Tamper resistant, Ruggedized
Customer Managed Encryption Key (CMEK)

HTTPS encrypted TLS connections

USE CASES FOR DATA TRANSFER

Data Center Migration

Decommission Tape Libraries & Infrastructure

Machine Learning

Storage & Delivery

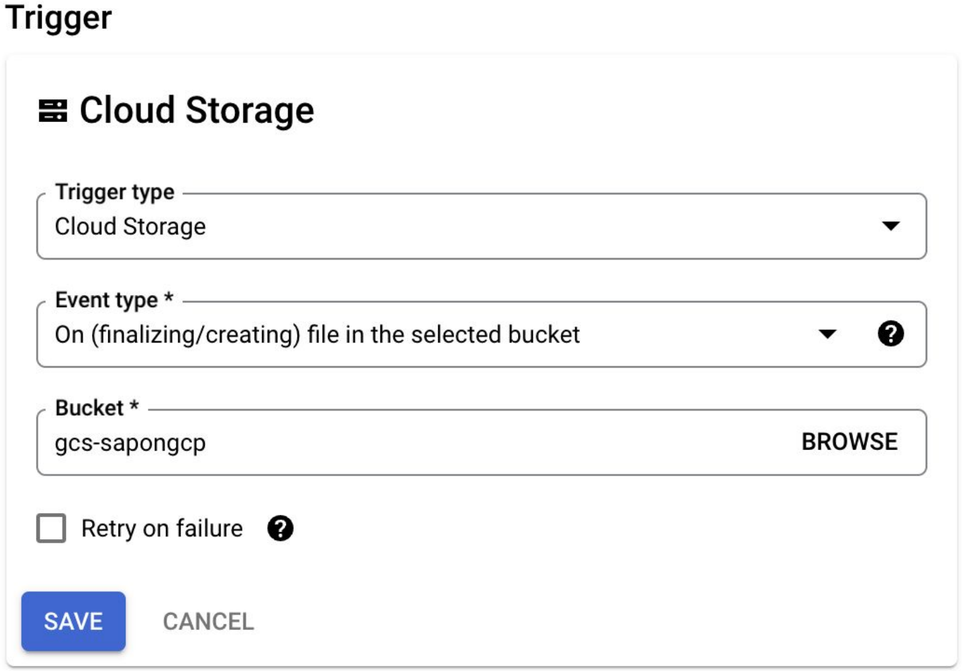
Backup & Archival

GCS: auto-reactions on object change:

Pub/Sub notifications vs object change notifications vs Cloud Functions

Exam Tips:

- *Each of those can be used to trigger some application-level action when an object is created / deleted / modified in a GCS bucket. The idea is the same for all three: instead of creating some recurring (eg. every 10 mins) job, have a trigger that will fire an application action whenever GCS action is finished.*
- *Cloud Functions can just "listen" to a specific GCS bucket*
- *Pub/Sub notifications are usually preferred. They send information about changes to objects in your buckets to Pub/Sub, where the information is added to a Pub/Sub topic of your choice in the form of messages.*
- *Object change notifications are usually NOT a preferred option (they're slower and less flexible than Pub/Sub)*



GCS: how to use CSEK?

Step-by-step procedure to manage objects in GCS bucket using CSEK::

1. Generate your own encryption key (make sure it's Base64-encoded AES-256 key)
2. Use this key to upload object in GCS:

   gcloud storage cp SOURCE_DATA gs://BUCKET_NAME/OBJECT_NAME --encryption-key=YOUR_ENCRYPTION_KEY

3. Use the same key when downloading the object:

   gcloud storage cp gs://BUCKET_NAME/OBJECT_NAME OBJECT_DESTINATION --decryption-keys=YOUR_ENCRYPTION_KEY

4. Make sure to regularly rotate encryption keys and update existing objects with the new one:

   gcloud storage objects update gs://BUCKET_NAME/OBJECT_NAME --encryption-key=NEW_KEY --decryption-keys=OLD_KEY

Exam Tips: *alternatively, use .boto configuration file. Make sure to know how to use it!*

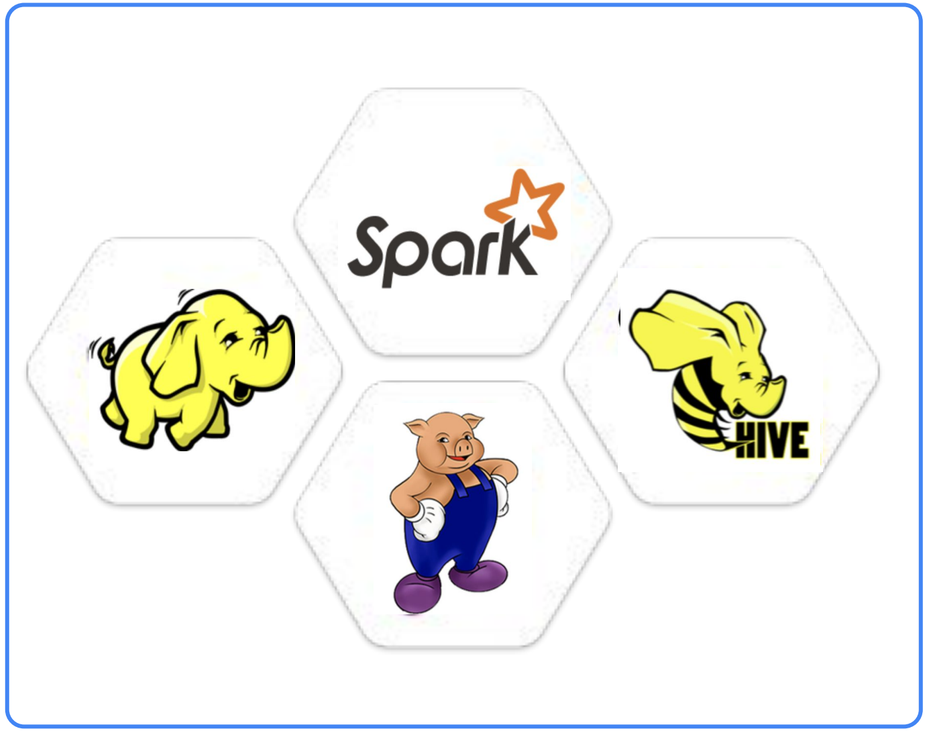
Cloud Storage: (most important) Organization Policy Constraints

Constraint	Description
Enforce Public Access Prevention	Secure your Cloud Storage data from public exposure by enforcing public access prevention. This governance policy prevents existing and future resources from being accessed via the public internet by disabling and blocking ACLs and IAM permissions that grant access to allUsers and allAuthenticatedUsers. Enforce this policy on the entire organization (recommended), specific projects, or specific folders to ensure no data is publicly exposed. constraints/storage.publicAccessPrevention
Retention policy duration in seconds	This list constraint defines the set of durations (in seconds) for retention policies that can be set on Cloud Storage buckets. constraints/storage.retentionPolicySeconds
Enforce uniform bucket-level access	This boolean constraint requires buckets to use uniform bucket-level access where this constraint is set to True. Enforcement of this constraint is not retroactive. Uniform bucket-level access disables the evaluation of ACLs assigned to Cloud Storage objects in the bucket. Consequently, only IAM policies grant access to objects in these buckets. constraints/storage.uniformBucketLevelAccess

HDFS to GCS: a common theme for Dataproc

Use Google Cloud Storage as your primary data source and sink

- Decouple storage and compute
- Un-silo your data
- Pros:
  - HDFS compatibility, lower costs, no storage management overhead, separate storage from compute, integration with other GCP services
- Cons:
  - Increased I/O variance and request latency (esp. Important with small files), GCS is immutable (no append / truncate)



Exam Tip: *Dataproc (and GCS usage instead of HDFS) is a common choice for customers migrating from existing, on-premises Hadoop ecosystem.*

Cloud Storage

#GCPSketchnote

@PVERGADIA THECLOUDGIRL.DEV

8.8.2020

What is Cloud Storage?

A GLOBAL, SECURE AND SCALABLE OBJECT STORE

GLOBAL SECURE SCALABLE

.png .jpg .mpg .mov .txt other

How does it WORK?

WORKFLOW

Image → Bucket (PHOTOS) → Project → Organization

4 STORAGE CLASSES

Based on Budget, Availability and Access Frequency

STANDARD	NEARLINE	COLDLINE	ARCHIVE
Frequent access High Availability	Once a month	Once a quarter	Once a year

New Version >30 Days >90 Days

OBJECT LIFECYCLE MANAGEMENT

AUTO VERSIONING → v1 → v2 → v3

Cloud Storage Use case example

STANDARD → Website + Streaming + Document

NEARLINE → Backups + Disaster Recovery

COLDLINE → Backups + Disaster Recovery

ARCHIVE → Long Term Archiving

STREAMING DATA, BATCH DATA → AI PLATFORM, BIGQUERY

SECURITY for Cloud Storage

- ☑ Encryption at rest
- ☑ Bring your own encryption key
  - CMEK - Customer Managed
  - CSEK - Customer Supplied

Cloud Storage PRICING

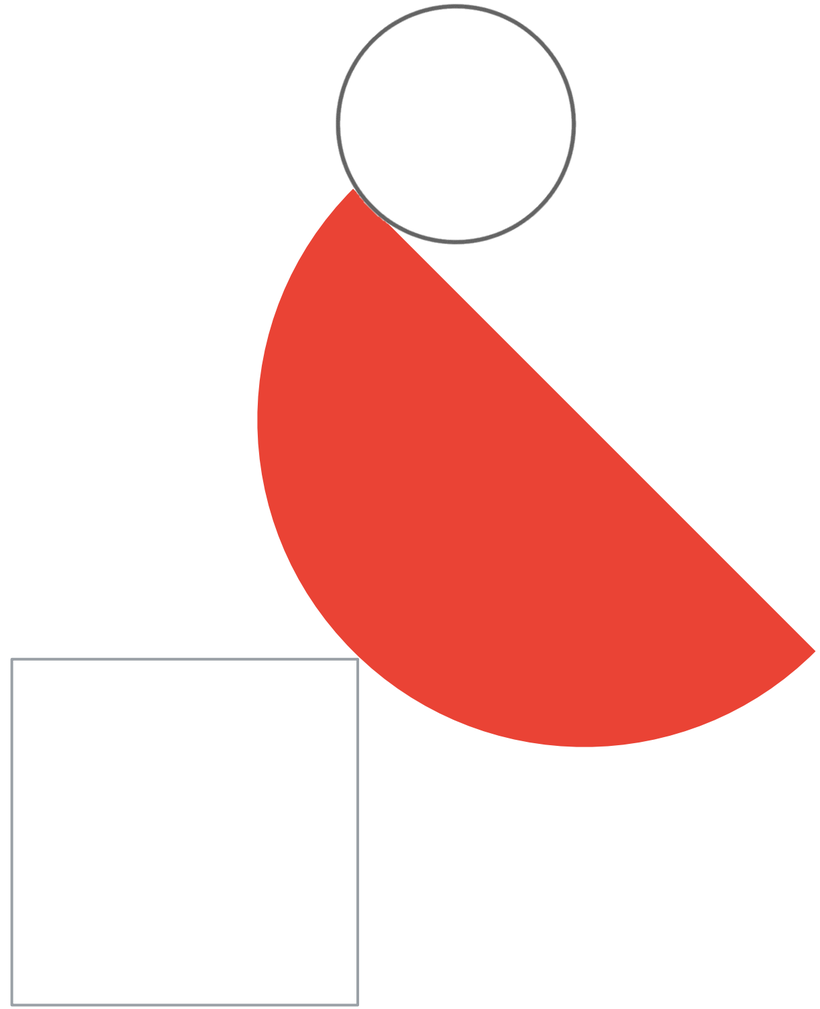
STANDARD	NEARLINE	COLDLINE	ARCHIVE
$$$$	$$$	$$	$
Automatic Redundancy Frequent Access			

How to USE Cloud Storage

ONLINE TRANSFER	TRANSFER SERVICE	TRANSFER APPLIANCE
gsutil, API, UI <gsutil cp image.png gs://my-bucket>	Transfer data from other clouds & on-premise	Hardware for >100tb data transfer

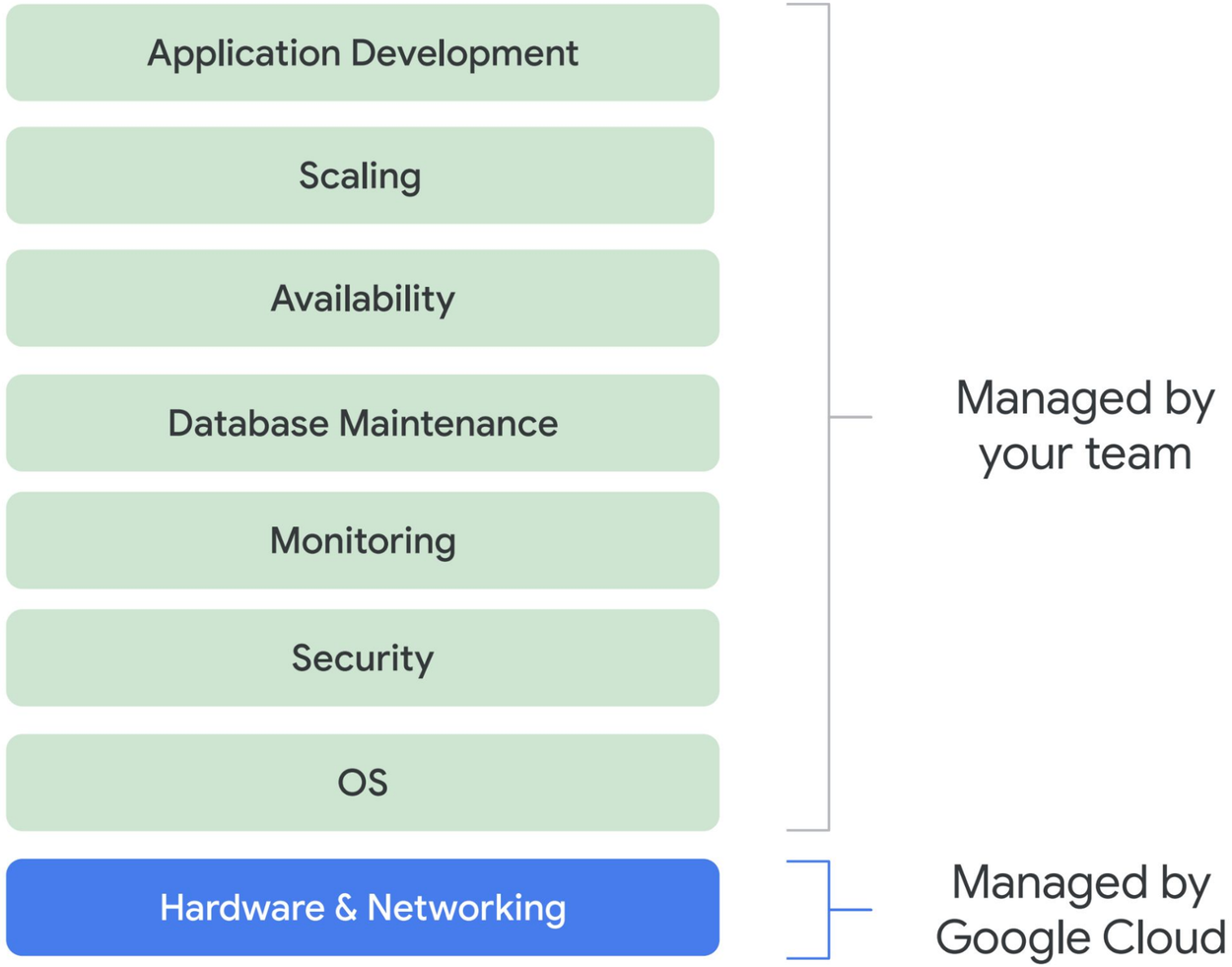
image.png API

Cloud SQL



Managed databases: the “why”

Self-managed DBs on GCE VMs



Exam Tip: *custom OS images / startup scripts / products from GCP Marketplace etc...*

I can automate the installation, but you still need to...

- Provision resources
- Install database engine
- Configure it
- Patch & update
- Handle high availability
- Implement & manage backup & restore
- Resize when needed
- Implement monitoring
- ... and so on

Exam Tip: *Hence* ***self-managed databases are NOT a preferred option from the exam perspective****... unless you have a valid reason.*

Managed database services.

Exam Tip: *using managed services is usually a preferred approach from exam perspective (unless you're running into some constraints / have special needs.*

Application Development

Managed by your team

Scaling

Availability

Database Maintenance

Monitoring

Security

OS

Hardware & Networking

99.999% SLA

Online scaling

Easy global replication

Automatic sharding

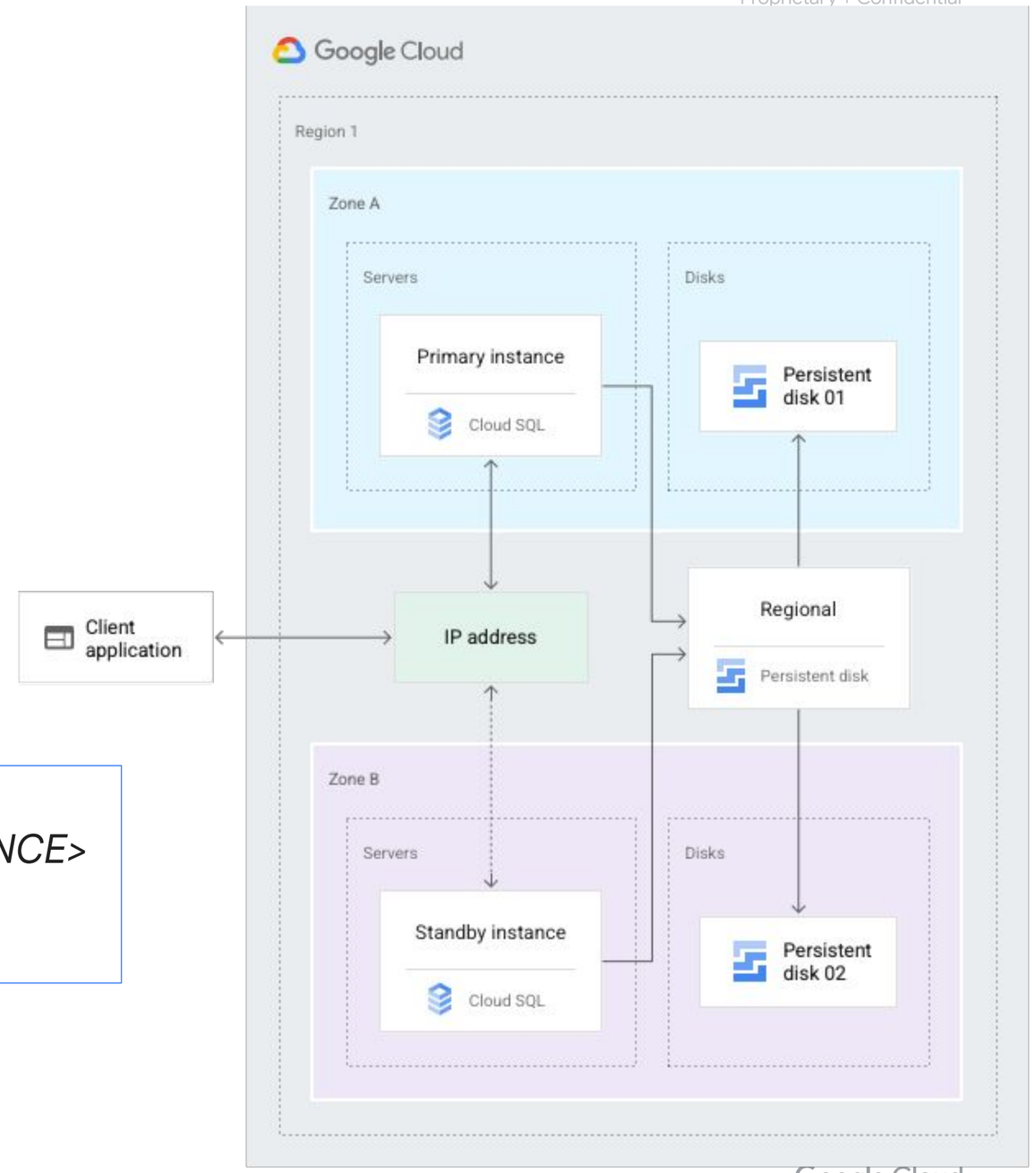
Automatic failure recovery

Built into cloud-native databases

Cloud SQL High availability

- Primary and secondary in different zones within the configured region
- Synchronous replication to each zone's persistent disk
- If heartbeat of the primary instance is unavailable for ~60 sec → automatic failover
- The persistent disk is then attached to the standby instance
- Less than 3 min of unavailability, the same IP address for the client application

Exam Tip: *Know how to:*
- *Initiate failover: gcloud sql instances failover <PRIM_INSTANCE>*
- *Check if instance is / isn't set for HA: gcloud sql instances describe (availabilityType = REGIONAL / ZONAL)*

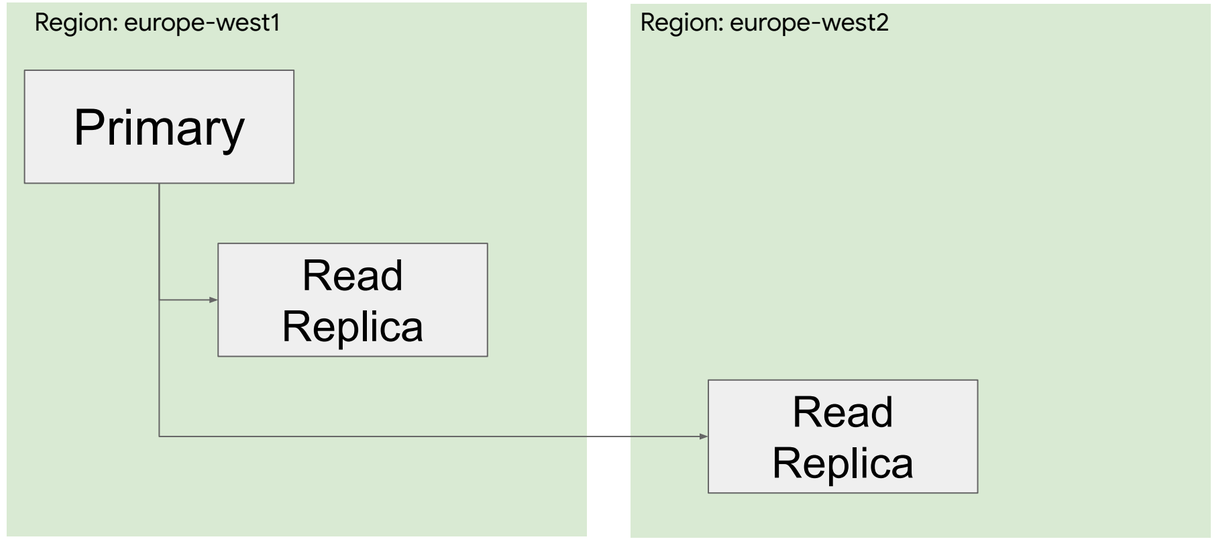


MySQL PostgreSQL Microsoft SQL Server

Cloud SQL: Read Replicas

Use cases: Disaster Recovery / offload analytics workloads / migrate between platforms or regions

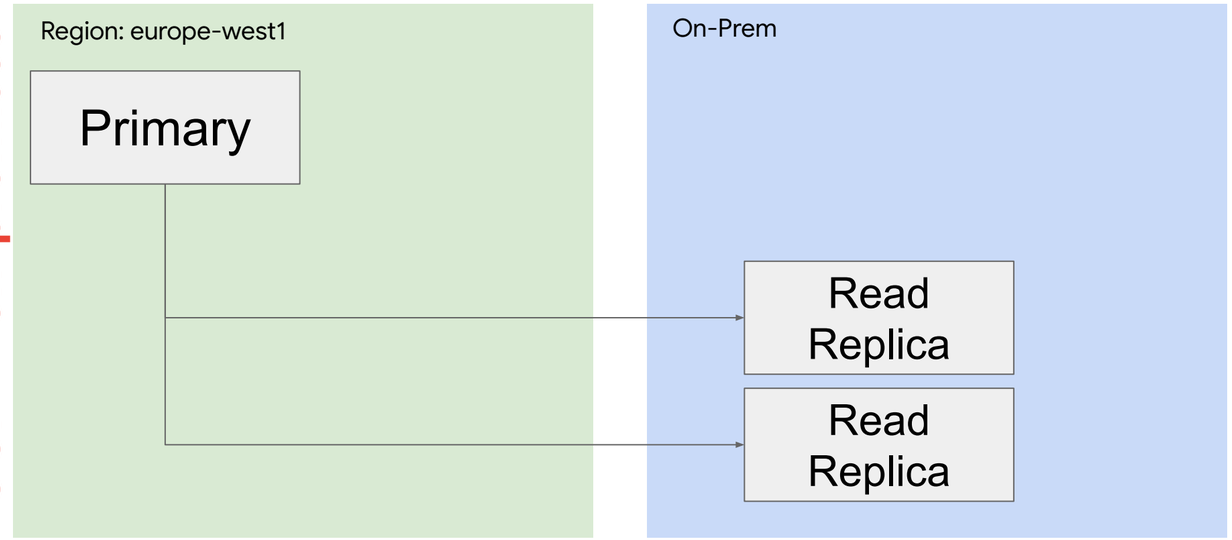
Read Replica
GCP -> GCP



Benefits & Use Cases

- Additional Read capacity (read only)
- Analytics target (adding secondary indexes)
- Read replicas can be different machine types than primary (never less vcpus for postgres)
- Settings of primary are propagated to replicas incl. root pwd &user table changes
- No load balancing between replicas
- Mysql Parallel replication (read replica side)

External Read Replica
GCP -> on-premises



Benefits & Use Cases

- Reduce latency for external connection
- Analytics target
- Migration path to other platforms
- In case of e.g. network outage on-prem the replication lag might be too large and replicas need to be recreated

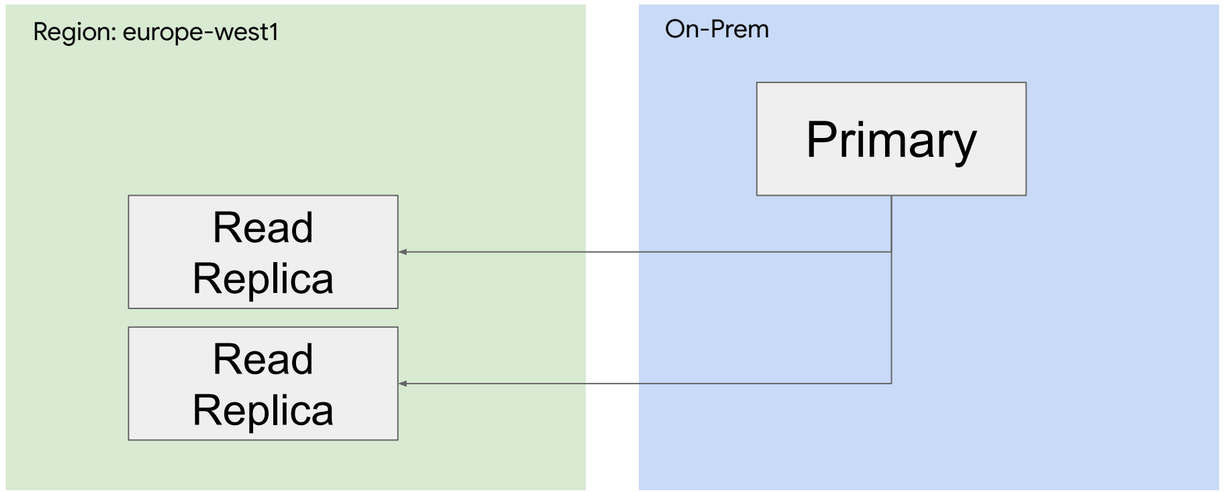
Exam Tip: *Docus on replicating TO external server*

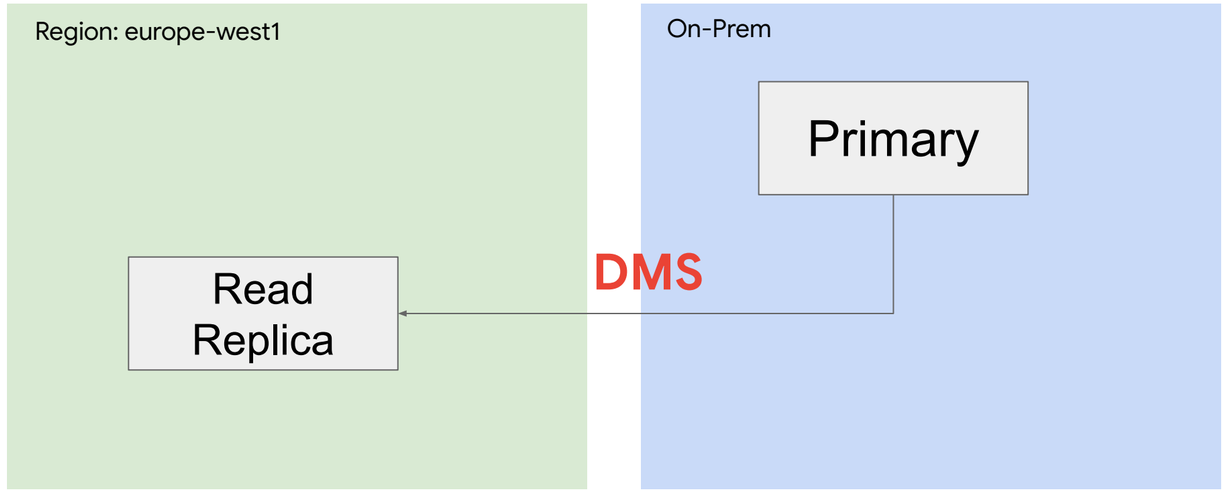
Cloud SQL: Read Replicas

Use cases: Disaster Recovery / offload analytics workloads / migrate between platforms or regions

Replication from external server

On-premises -> GCP





MYSQL Benefits & Use Cases

- Migration path to Cloud SQL with minimum downtime
- Data replication to GCP
- Offloading admin overhead of replicas to GCP
- Analytics target
- Parallel replication (read replica)

POSTGRES - DMS

- Use DMS to replicate from an external DB Server to a Cloud SQL Read replica (One-off Migration or Continuous cdc replication)
- DB Source can be self-managed DB (on-prem or IaaS), Aws Rds, Aurora, Cloud SQL

Exam Tip: *Docus on replicating FROM external server*

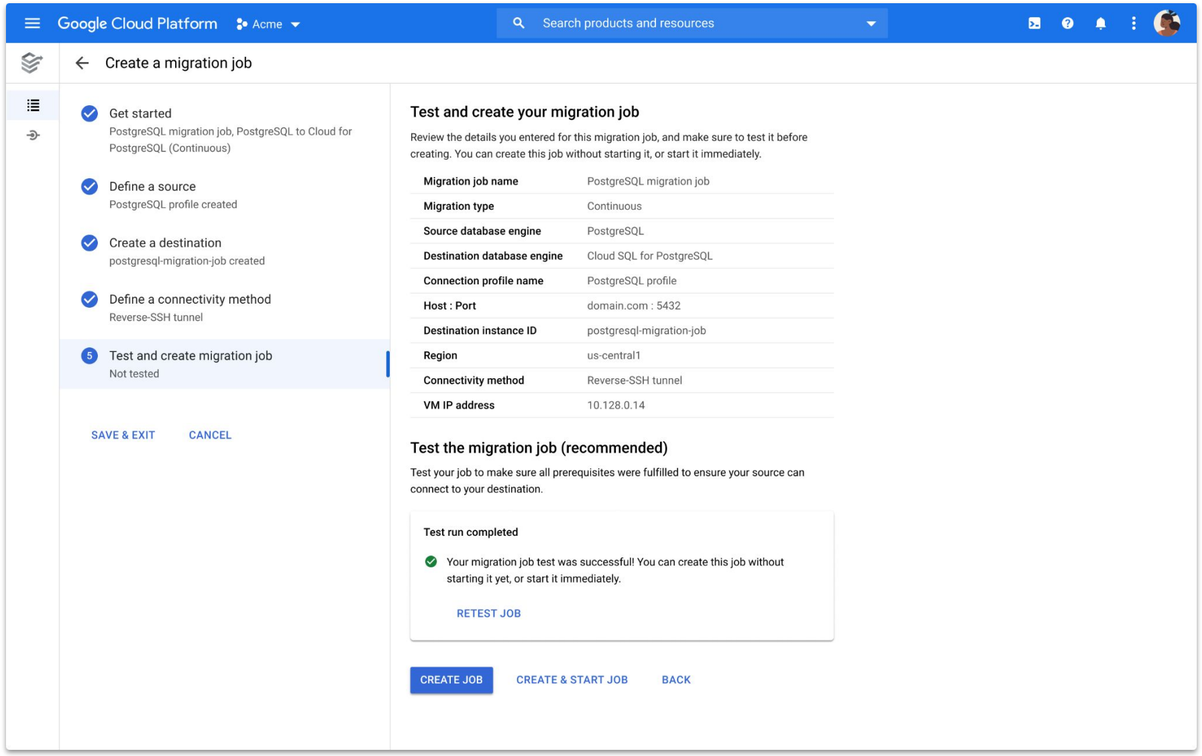
Simplified migration with Data Migration Service (DMS)

Continuous migration path to Cloud SQL with minimal downtime

Simplest way to migrate to Cloud SQL:

- No migration servers to manage
- Baked-in configuration support
- Native replication method

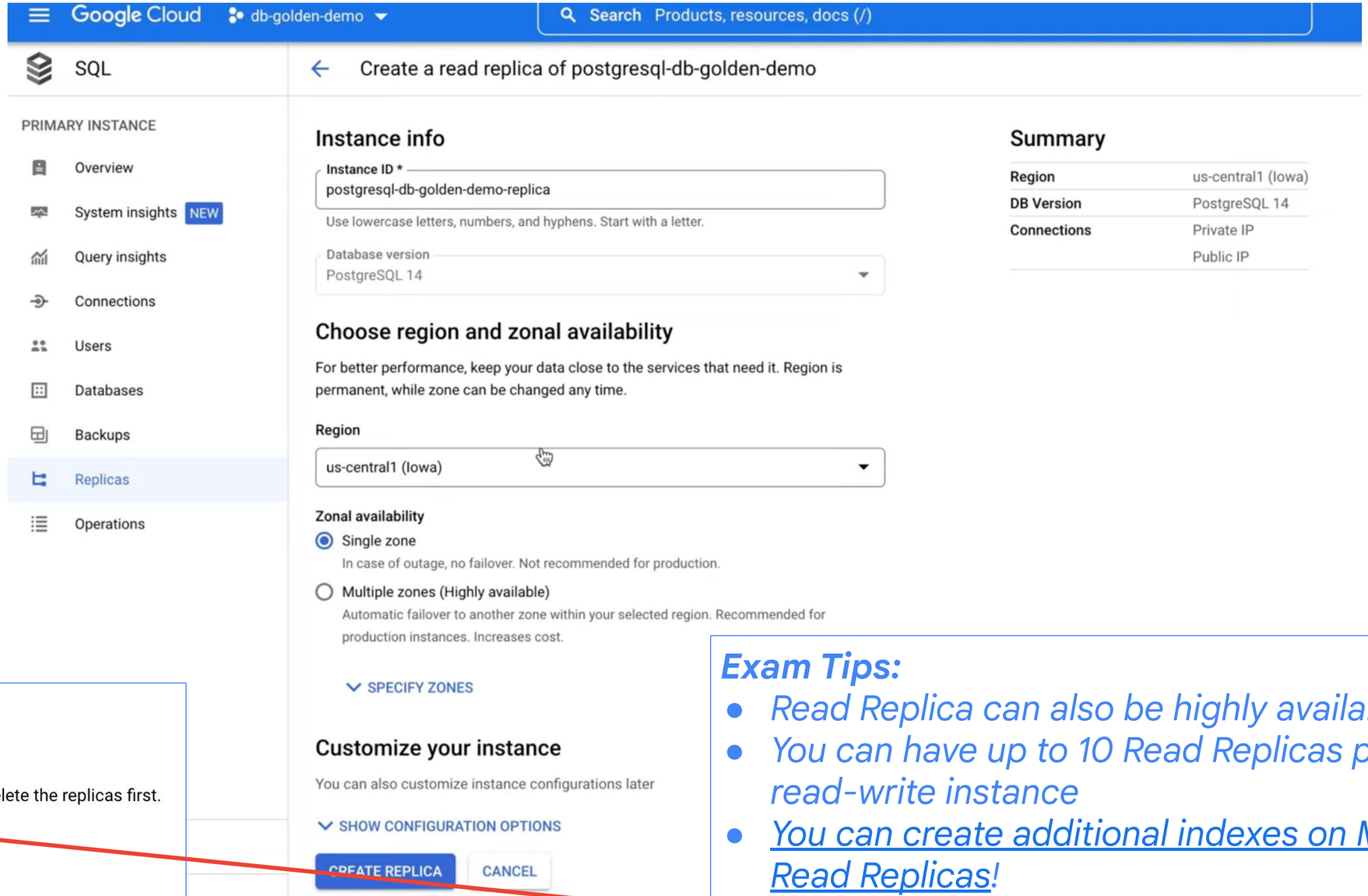
Secure connection options for encrypted data and using private networking

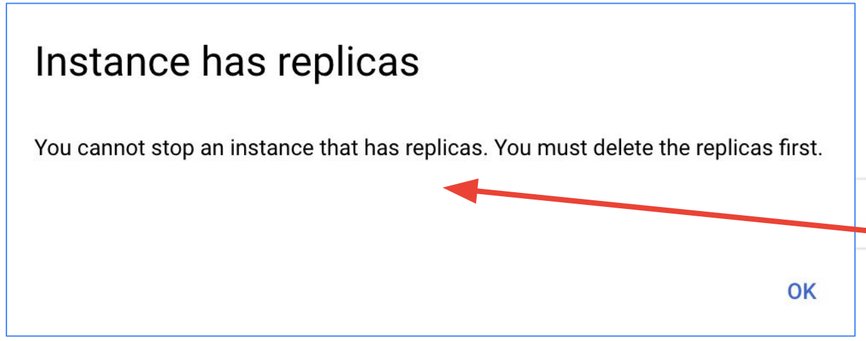


Exam Tip: *Make sure to be familiar with this overview: Database Migration Service.*

Cloud SQL DR with cross-region Read Replicas

How to create





Exam Tips:

- *Read Replica can also be highly available.*
- *You can have up to 10 Read Replicas per read-write instance*
- *You can create additional indexes on MySQL Read Replicas!*
- *!!!!*

Google Cloud

Cloud SQL: edit instance

Zone	Y	The possible values depend on the region.
Database version	N	Console string / API enum string **PostgreSQL 14** POSTGRES_14 **PostgreSQL 13** (default) POSTGRES_13 **PostgreSQL 12** POSTGRES_12 **PostgreSQL 11** POSTGRES_11 **PostgreSQL 10** POSTGRES_10 **PostgreSQL 9.6** POSTGRES_9_6
Set password policy	Y	Configured or not.
Private IP	After it is enabled, it cannot be disabled.	Enabled or disabled.
Public IP	Y	Enabled or disabled.
Authorized networks	Y	If Public IP is enabled, IP addresses authorized to connect to the instance. You can also specify this value as an IP address range, in CIDR notation.
Private path for Google Cloud services	Y	Enabled or disabled.
Machine type	Y	Select from Shared core, Lightweight, Standard (Most common), or High memory. Select the **Custom** radio button to create a custom machine type. Learn more

Exam Tips:

- *Can be done with command: gcloud sql instances patch INSTANCE_NAME –<setting_name> <value>*
- *Have a look at the parameter list, know the most important ones and focus if those can be changed AFTER instance creation or not.*
- *Storage (=regional PD) size CANNOT be reduced (just like with normal VMs)*

Cloud SQL - Performance & scaling

Machine Type

Choose a preset or customize your own. For better performance, choose a machine type with enough memory to hold your largest table.

Shared core

- (•) 1 vCPU, 0.614 GB
- ( ) 1 vCPU, 1.7 GB

...

- (•) Custom

vCPUs * 96 vCPU (1 - 96)

Memory * 624 GB (86.5 - 624)

Storage

Storage type

Choice is permanent. Storage type affects performance.

- ( ) SSD (Recommended)
  Most popular choice. Lower latency than HDD with higher QPS and data throughput.
- (•) HDD
  Lower performance than SSD with lower storage rates.

Storage capacity

10 - 65,536 GB. Higher capacity improves performance, up to the limits set by the machine type. Capacity can't be decreased later.

- (•) 10 GB

...

Storage

Storage type

Choice is permanent. Storage type affects performance.

- (•) SSD (Recommended)
  Most popular choice. Lower latency than HDD with higher QPS and data throughput.
- ( ) HDD
  Lower performance than SSD with lower storage rates.

Storage capacity

10 - 65,536 GB. Higher capacity improves performance, up to the limits set by the machine type. Capacity can't be decreased later.

- ( ) 10 GB
- ( ) 20 GB
- ( ) 100 GB
- ( ) 200 GB
- (•) Custom

65536 GB (10 - 65,536)

[x] Enable automatic storage increases

If enabled, whenever you are nearing capacity, storage will be incrementally (and permanently) increased. Learn more

Cloud SQL: GCP-native Backup and Restore

Types

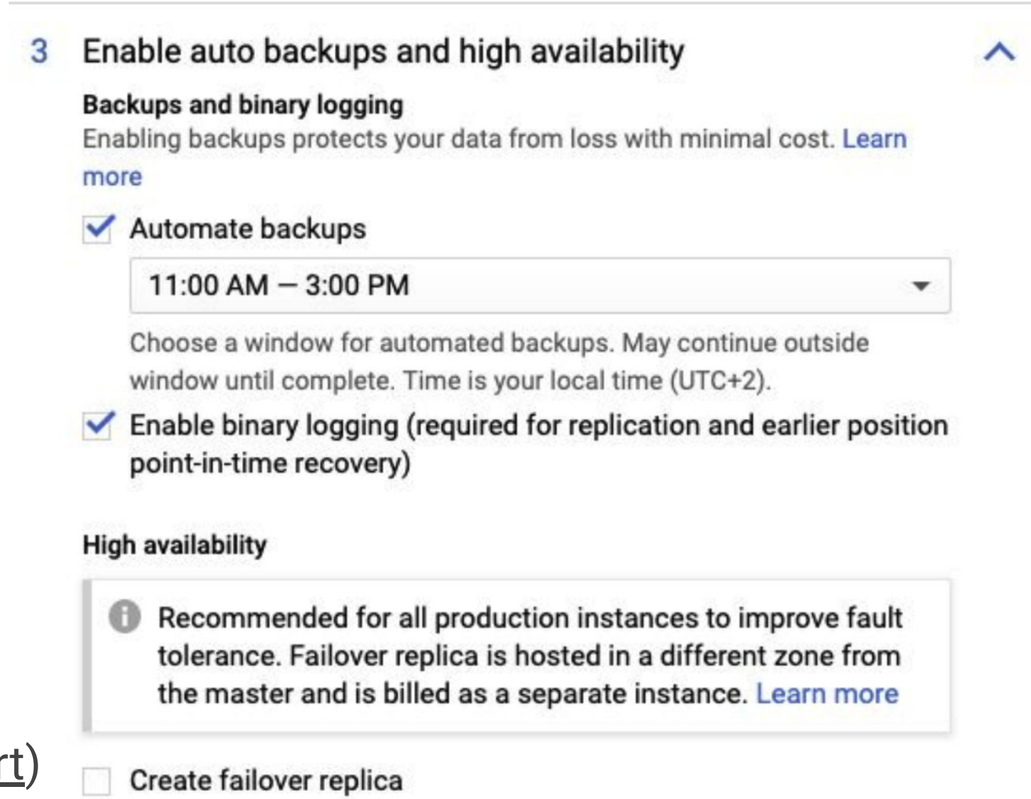
- **On-Demand**
  - Create disk-level snapshot backup at any time
  - Not deleted automatically

- **Automated**
  - 4 hour backup window (e.g. 11am - 3pm)
  - Schedule when instance has least activity
  - If data has not changed since last backup then no backup is taken

Characteristics

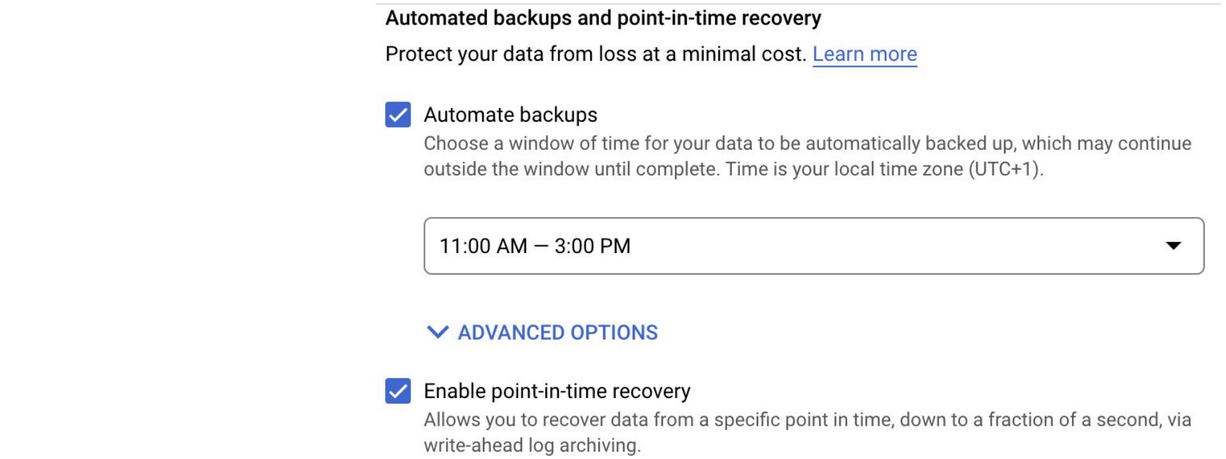
- Up to 365 daily automated backups for each instance. (Only up to 7 days for binlog/WAL files)
- Incremental Backups
- Storage used by backups is charged at a reduced rate. (see pricing)
- Backups can not be exported - only instance data (see doc for export)
- Backups are deleted after instance is deleted (Data export required to retain data; read replica for export without perf. impact)
- Backups are disk-level snapshots stored on GCS

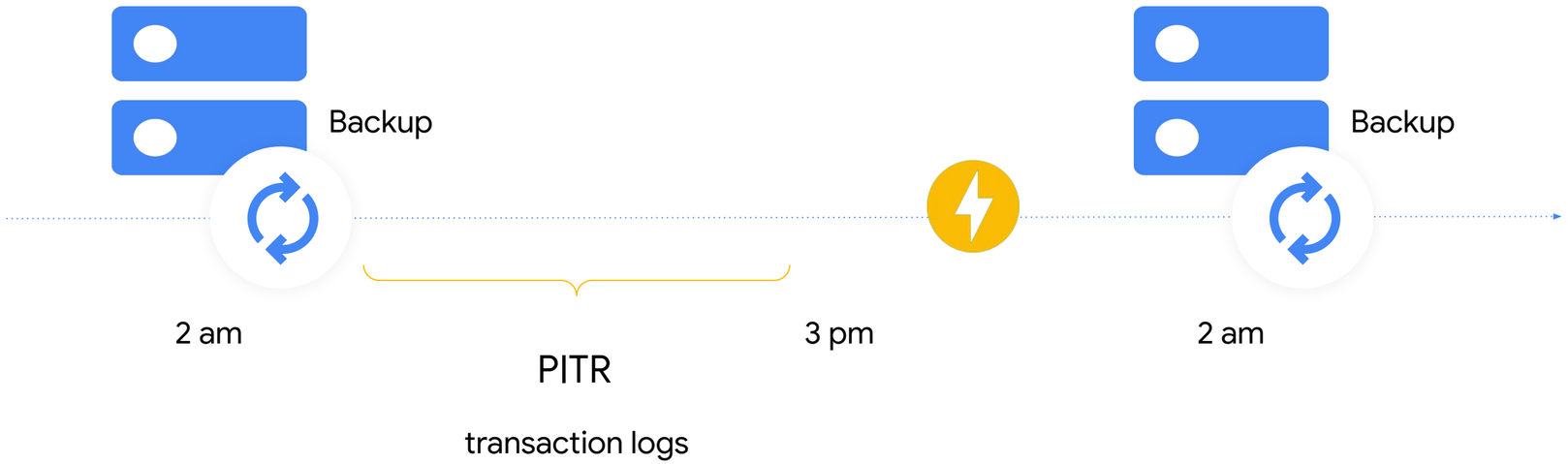
Exam Tip: *Cloud SQL backup window is NOT the same as maintenance window.*



Cloud SQL: Point-in-time recovery

recover an instance to a specific point in time



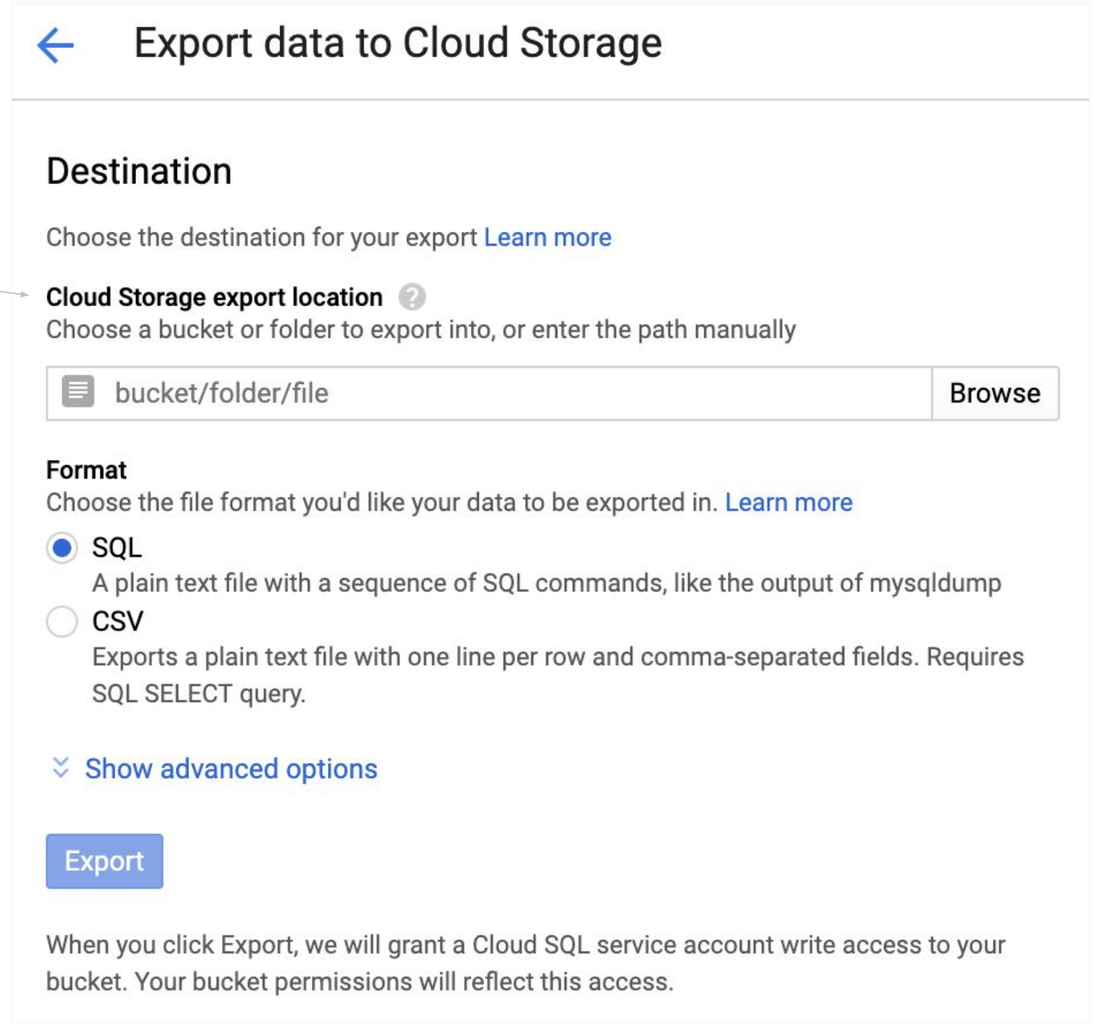


Cloud SQL: Data Export

Export (to GCS)

- ➔ SQL - high fidelity
- ➔ CSV - database-agnostic

Note:
Export to different buckets in different project is also supported if you have granted the service account with write permissions on this bucket.



Exam Tip: *For regular & automatic Cloud SQL exports, use Cloud Functions and Cloud Scheduler.*

Connection Options (external apps)

Exam Tip: *Make sure to know when to use which pattern: Cloud SQL Connection options*

Exam Tip: *Common solution to questions about connectivity from a GKE cluster to Cloud SQL*

Connection option	Secure, encrypted?	More information	Notes
Public IP address with SSL	Yes	• Configuring SSL for Instances • Configuring access for IP connections • Connect mysql client using SSL	SSL certificate management required.
Public IP address without SSL	No	• Configuring access for IP connections	Not recommended for production instances.
Cloud SQL Proxy	Yes	• Connecting from an external application using the Cloud SQL Proxy • Connecting mysql Client Using the Cloud SQL Proxy • About the Cloud SQL Proxy	
Cloud SQL Proxy Docker image	Yes	• Connecting mysql Client Using the Cloud SQL Proxy Docker Image • About the Cloud SQL Proxy	
JDBC Socket Library	Yes	• External connections with Java • JDBC socket factory GitHub page	Java programming language only.
Go Proxy Library	Yes	• External connections with Go • Cloud SQL Proxy GitHub page	Go programming language only.
Cloud Shell	No	• Using the mysql client in the Cloud Shell	Uses the Cloud SQL Proxy to easily connect from the Google Cloud Console. Best for quick administration tasks requiring the `mysql` command-line tool.
Apps Script	Yes	• External connections with Apps Script • Apps Script sample GitHub page	Apps Script can connect to external databases through the JDBC service, a wrapper around the standard Java Database Connectivity technology.

Cloud SQL

IP address assignment

Private IP:

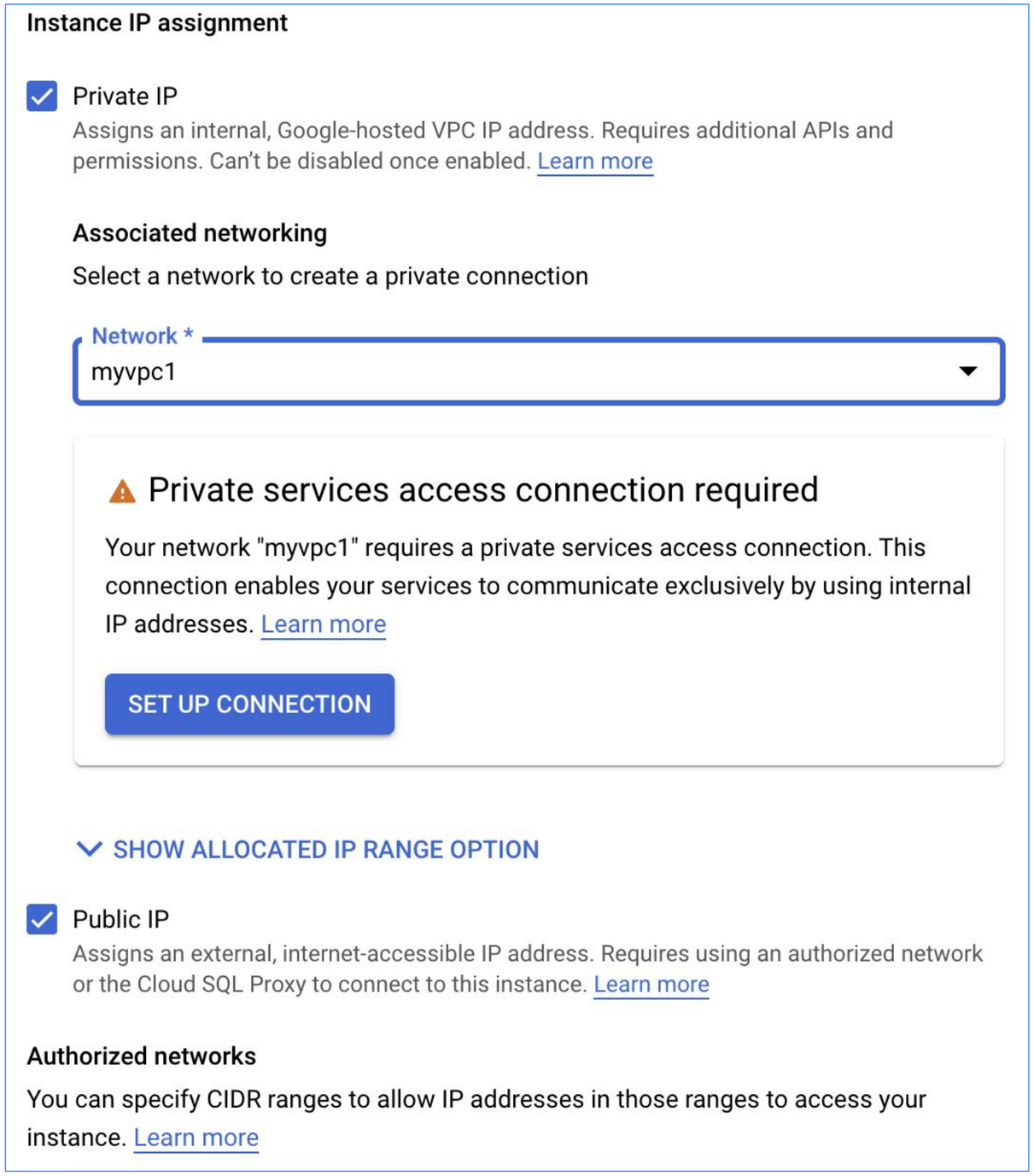
- Preferred when client is coming from resource with internal visibility (not necessarily from the same VPC!
- IPv4 address accessible from VPC
- Connections **may** be configured to use Cloud SQL proxy or self-managed SSL certificates
- Low latency and increased security

Public IP:

- IPv4 address accessible from the public network
- Connections **must** be authorized using either the Cloud SQL Auth proxy or authorized networks

Exam Tip: *Cloud SQL instances can have* ***both*** *a public and a private IP address. If private IP address is configured, Private Service Access (technically: VPC peering) is configured underneath.*

MUST-WATCH VIDEO

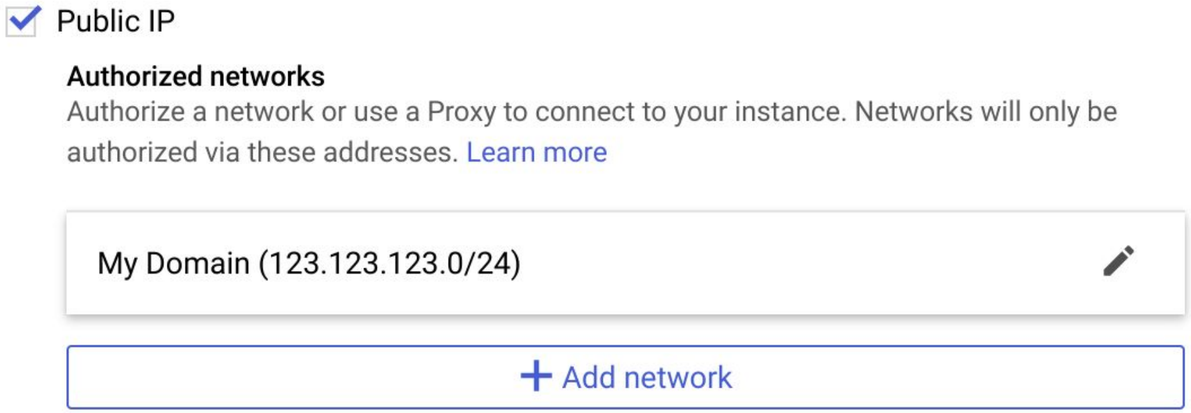


Cloud SQL: Public IP & risk mitigation

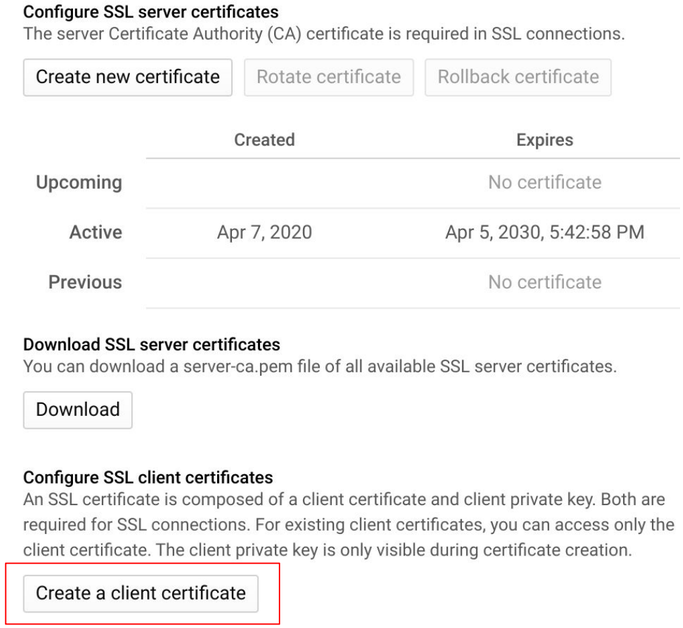
Access by public IP, without SSL : maximum risk for your data !!

Risk mitigation options:

Set an authorized network domain



Use SSL Certificate for transit data




```
psql "sslmode=verify-ca sslrootcert=server-ca.pem \
      sslcert=client-cert.pem sslkey=client-key.pem \
      hostaddr=01.23.45.67 \
      user=postgres dbname=postgres"
```

Cloud SQL: Private IP

Characteristics

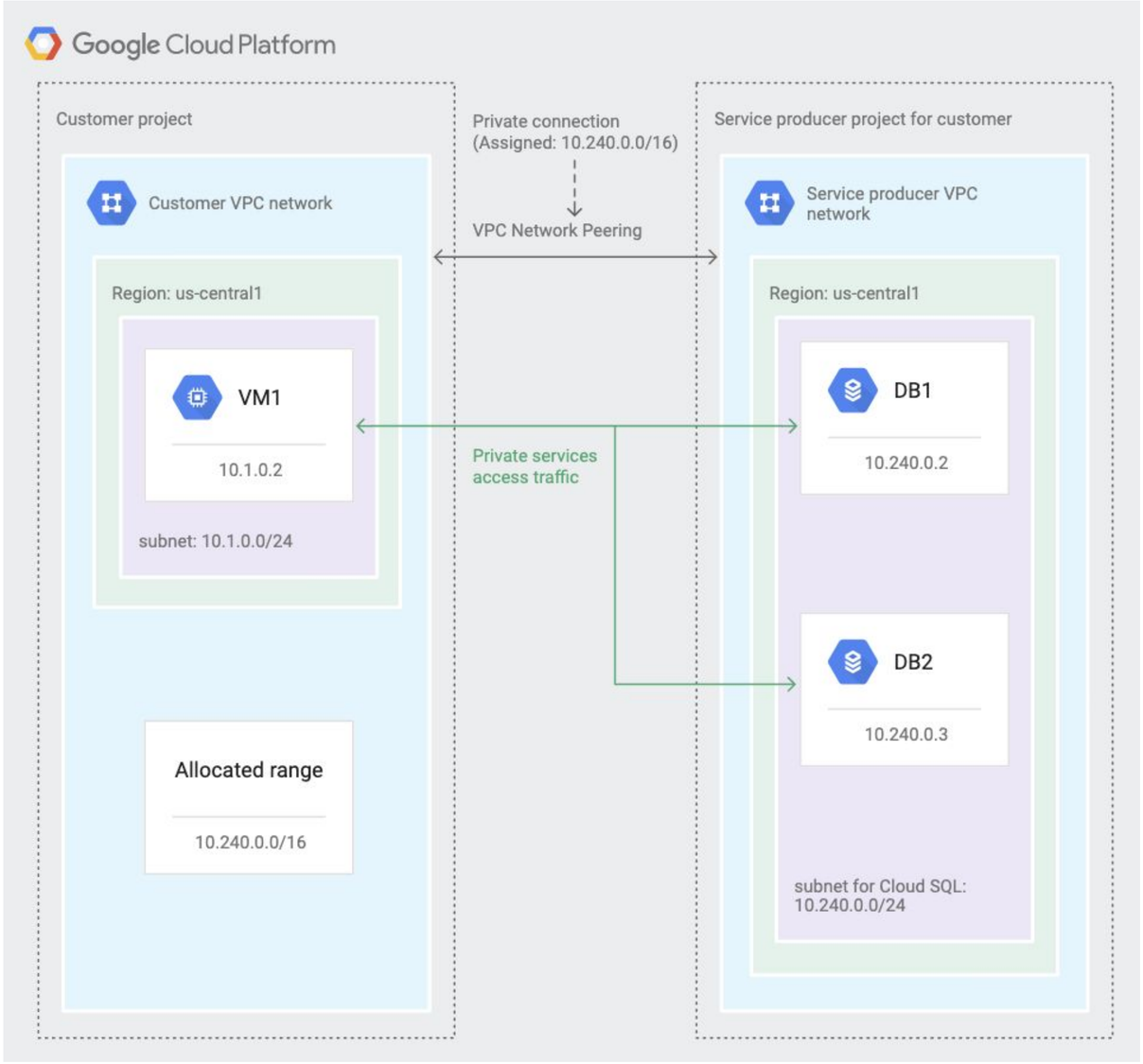
- Allows VM instances in your VPC network to use internal IP addresses to reach the service resources that have internal IP addresses
  → i.e. Connect CloudSQL to GCE or GKE instances
- **GCP uses network peering to create connection**

Benefits

- **Lower network latency**
  Best performance
- **Improved network security**
  Traffic is never exposed to public internet
- **Lower egress cost**
  Regular network pricing still applies to all traffic.

Limitations

- **Max 25 Peering connection per VPC network**
- **Transitive peering is not supported.**
- **Subnet in Peered VPC cannot overlap with CloudSQL**



Cloud SQL

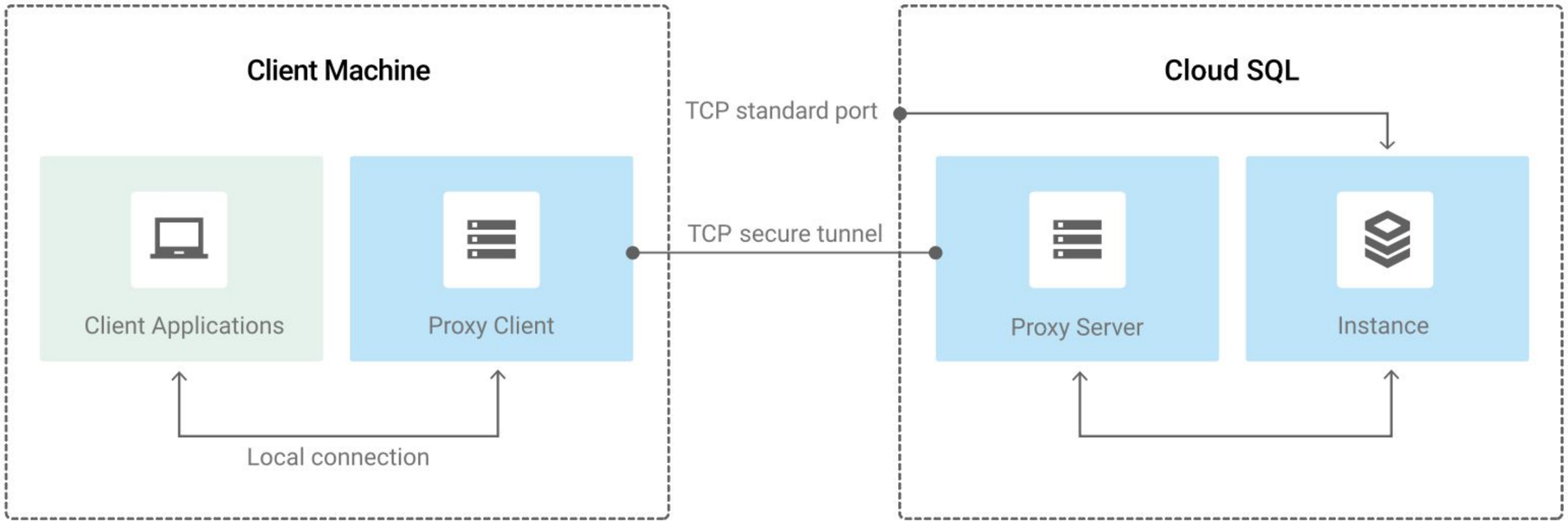
Access authorization -> Cloud SQL Auth proxy details

How the Cloud SQL Auth proxy works?

- a local client running in the local environment + companion process running on the server.
- doesn't provide connection pooling, but can be paired with other connection pooling to increase efficiency.
- also available as a Docker container.

Exam Tips:

- *Cloud SQL Proxy is the recommended option even when connecting to Cloud SQL behind a Private IP (because of strong encryption and authentication using IAM)*



Cloud SQL

Access authentication

- Basic roles (should NOT be used!):
  - Owner (Full access and control for all Google Cloud resources)
  - Editor (Read-write access to all Google Cloud resources)
  - Viewer (Read-only access to all Google Cloud resources)

Common Cloud SQL IAM Roles:

Role (predefined)	Privileges	For who/which service
Cloud SQL Admin	Full control for all Cloud SQL resources.	DBA Team / DB owner
Cloud SQL Editor	Manage Cloud SQL resources. No ability to see or modify permissions, nor modify users or sslCerts. No ability to import data or restore from a backup, nor clone, delete, or promote instances. No ability to start or stop replicas. No ability to delete databases, replicas, or backups.	DB Operator
Cloud SQL Viewer	View all Cloud SQL resources (read-only)	Audit, Security, DevOps Team
Cloud SQL Client	Connectivity access to Cloud SQL instances from App Engine and the Cloud SQL Proxy. Not required for accessing an instance using IP addresses.	Apps service (AppEngine, CloudSQL Auth Proxy)

Exam Tip: *Understand permissions in predefined Cloud SQL roles: Admin / Editor / Viewer / Client.*

CloudSQL

#GCPSketchnotes

@PVERGADIA THECLOUDGIRL.DEV

What is CloudSQL?

FULLY MANAGED RELATIONAL DB FOR

MySQL PostgreSQL SQL Server

How does it WORK?

CREATE

region
+ zone
+ machine type (cpu & memory)
+ storage (SSD or HDD)
= Instance

- [x] Enable automatic storage increases

BACKUPS

- [x] Automate backups [time windows]
- [x] Enable point-in-time recovery

HIGH AVAILABILITY

Automatic failover to another zone or region

ZONE A
AUTOMATIC FAILOVERS
ZONE B

How do you USE it?

CONNECTIVITY WITH OTHER SERVICES

CloudSQL connects to App Engine, GKE, GCE, Cloud Run and Workstation

CloudSQL

Steps to MIGRATE

EXISTING MYSQL DB TO CLOUDSQL

1 2 3 4

1. PROVIDE details on your data source
2. CREATE a Cloud SQL read replica
3. SYNC read replica with source
4. PROMOTE read replica to primary instance (optional)

SECURITY for CloudSQL?

ENCRYPTED AT REST & IN MOTION

CloudSQL PRICING

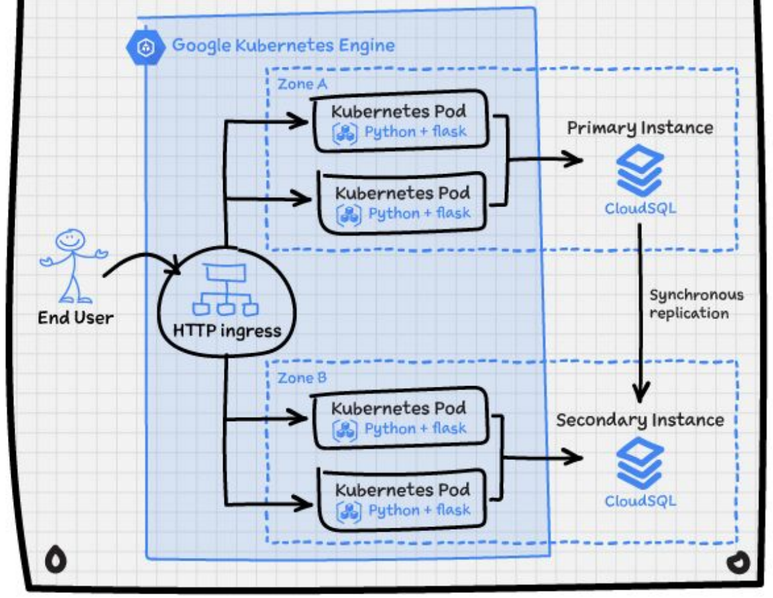
IN GENERAL PRICING IS A FACTOR OF:

Instance + CPU + Memory + Storage + Network Cost

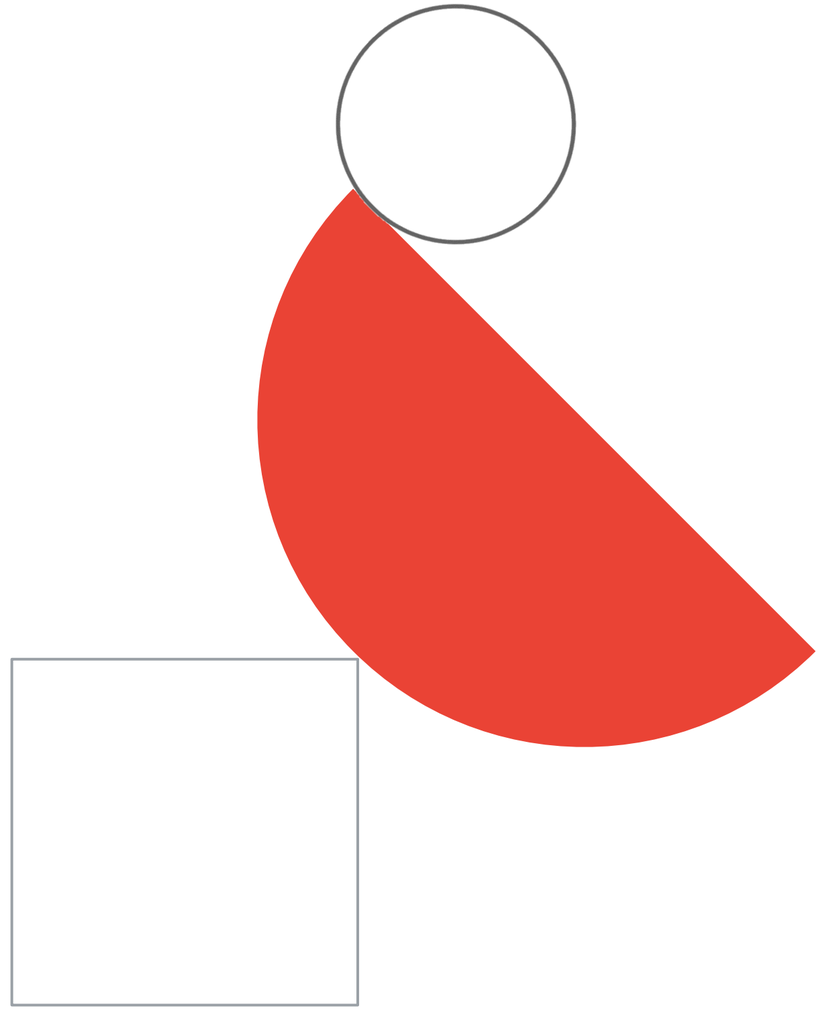
$

CloudSQL Use case (example)

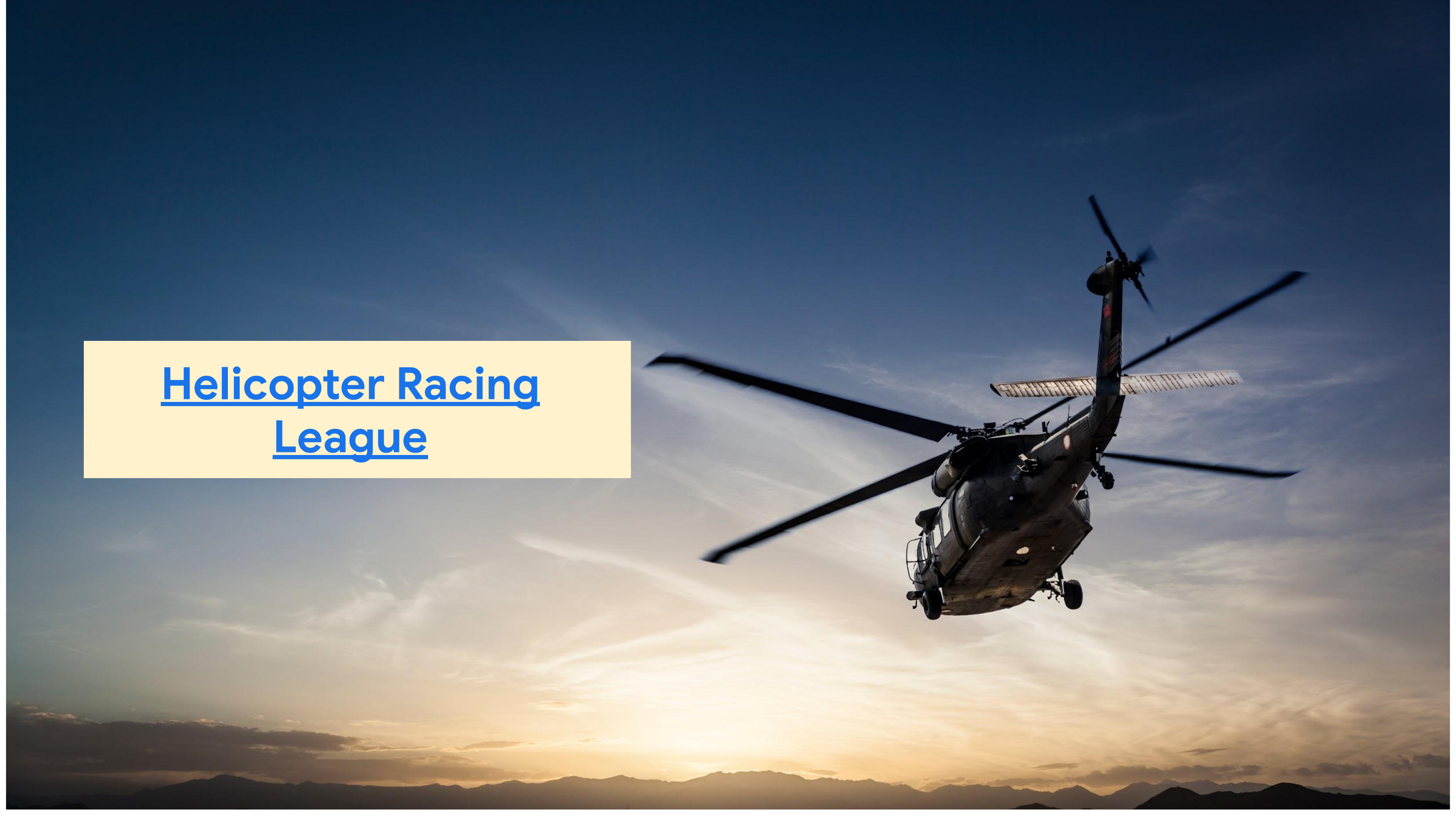
- [x] Online Transaction processing (OLTP)
- [x] Frequent queries
- [x] Fast response times



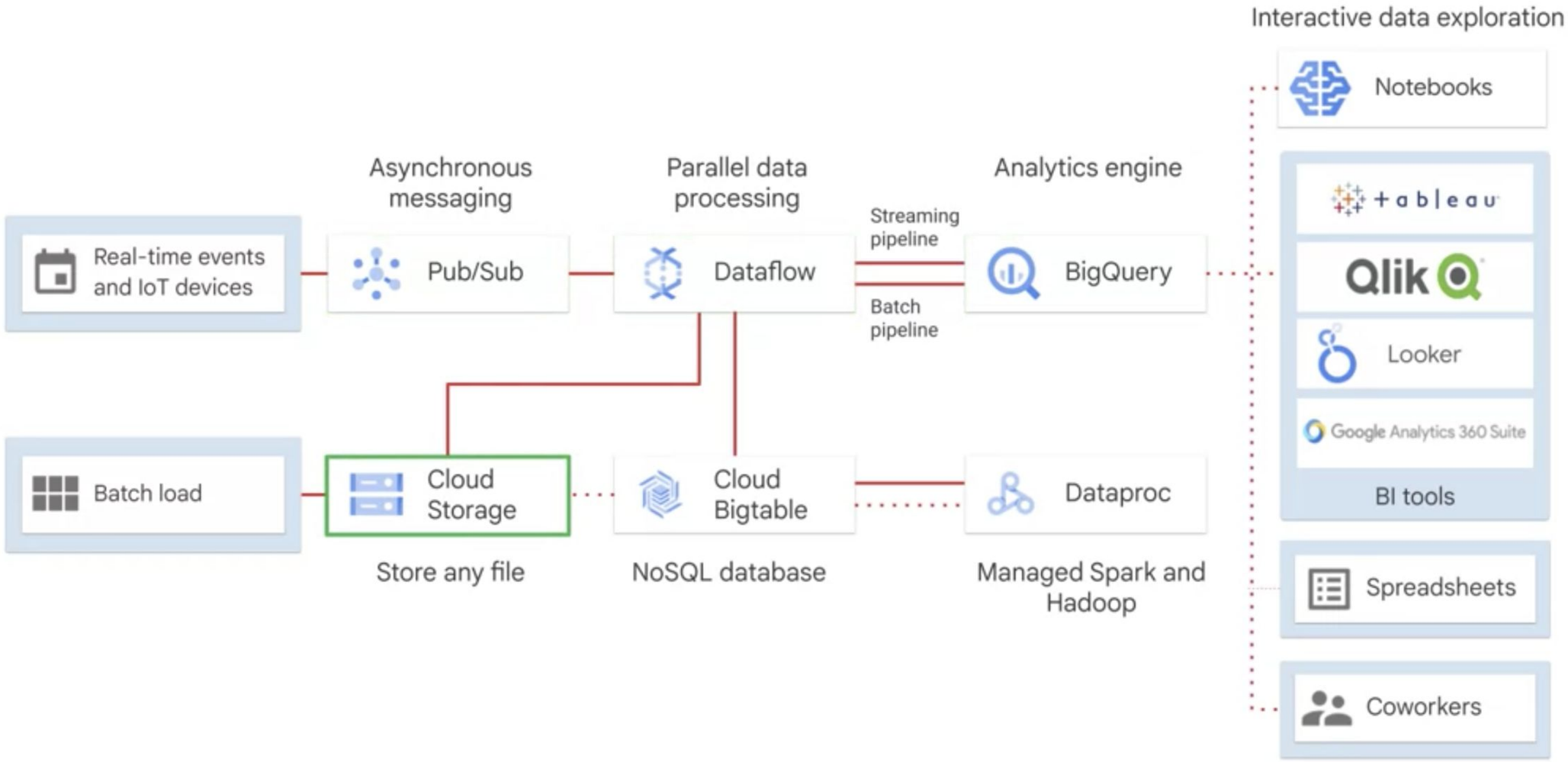
HRL case study analysis



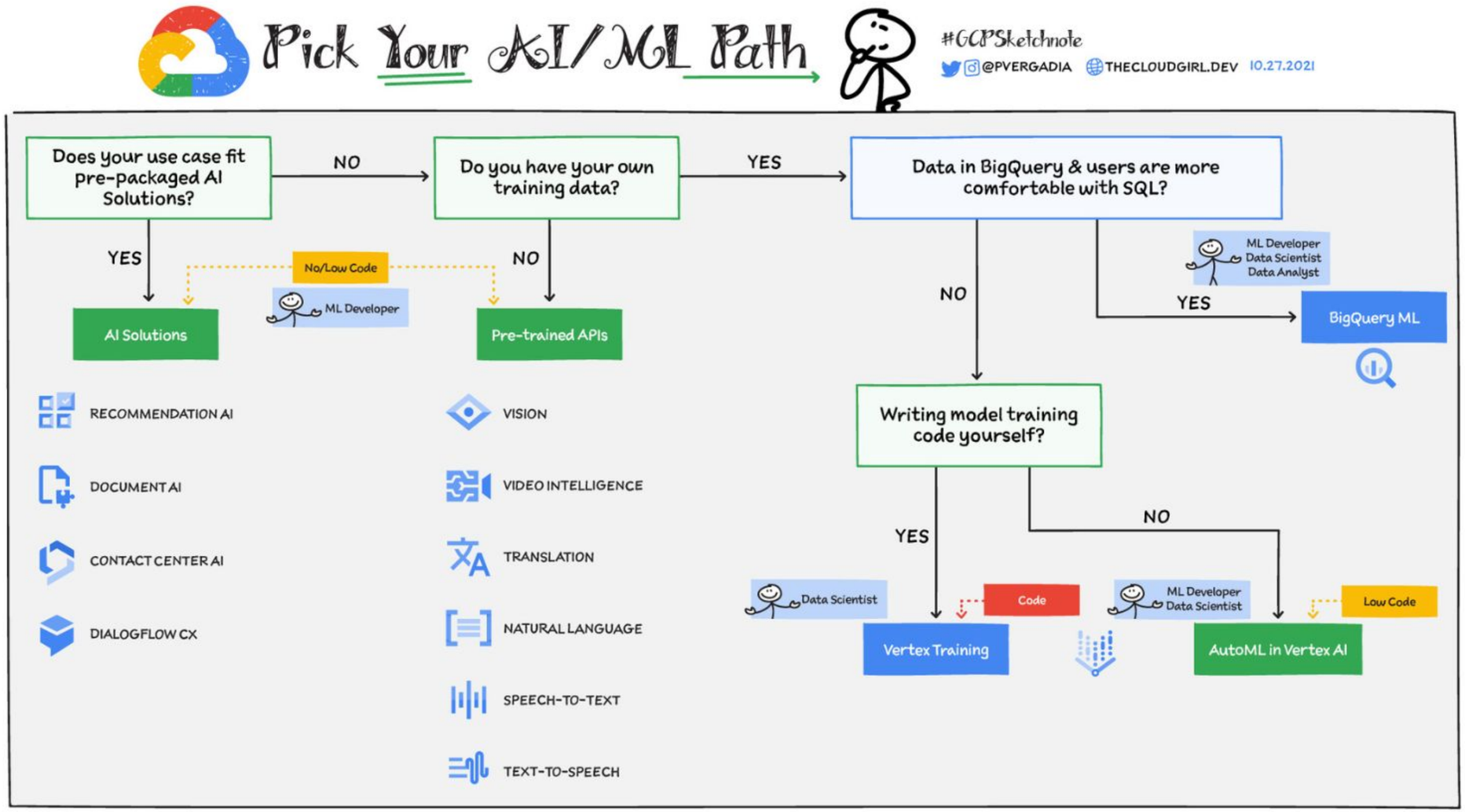
Helicopter Racing League



Modern serverless data management architecture



The Google Cloud machine learning spectrum



Exam Tip: *4 high-level ways to approach ML in GCP. All depends on particular use-case, but the approach should be like with compute: start from more managed / easier approaches and choose more advanced ones only if needed.*

Proposed Technical Solutions

- Processing environment:
  - GCE (VMs) for lift and shift, add managed instance groups for automatic scaling.
  - If they'd like to optimize for costs and scalability: containerize the app and deploy on GKE with multiple clusters in different regions.
    i. Use GKE multi-cluster capabilities or Anthos to manage/orchestrate workloads across clusters.
  - Transcoding API to perform video transcoding (from raw streams to different quality of ready to consume video files).
  - **Video Intelligence Streaming API with HLS Protocol for live streaming.**
- Connectivity between GCP and other hyperscalers: standard patterns.
- Multi-regional GCS buckets for increased availability and reduced latency to end users.
  - Use on-line Storage Transfer Service to migrate the data from other cloud provider to GCP.
- Global HTTPS Load Balancer + Cloud CDN (serving data from GCS buckets or from a different hyperscaler) + Cloud Armor.
- Preemptible VMs (maybe with GPUs or even TPUs); Spot VMs are “2nd generation of Preemptible VMs”.
- Telemetry
  - Native services (Cloud Operations Suite), maybe stronger focus on Cloud Trace (importance of latency) and VPC Flow Logs.
  - Network Intelligence Center for more insights into networking topology / connectivity / performance.
- **BigQuery for real-time analytics and data mart**
  - (if ingested data is structured) GCS bucket => BigQuery
  - (if ingested data needs transformations) GCS bucktes => (optionally if streaming) Pub/Sub => Dataflow => BigQuery
  - Data Studio or Looker for visualization and analytics
- AI/ML:
  - Enriching video (tags, labels, object detection etc): Video Intelligence API.
  - **Predicting race results, experimental forecasting: Vertex AI** with Jupyter Notebooks and options to deploy and expose the model to partners.
  - Viewers sentiment analysis: Natural Language API.
- **Expose APIs / ML models to partners: Apigee (monetization, rate limiting, merchandising revenue stream etc)**

[HRL case study] Diagnostic Question #1

For this question, refer to the Helicopter Racing League (HRL) case study. HRL is looking for a cost-effective approach for storing their race data such as telemetry. They want to keep all historical records, train models using only the previous season's data, and plan for data growth in terms of volume and information collected. You need to propose a data solution.

Considering HRL business requirements and the goals expressed by CEO S. Hawke, what should you do?

A. Use Firestore for its scalable and flexible document-based database. Use collections to aggregate race data by season and event.

B. Use Cloud Spanner for its scalability and ability to version schemas with zero downtime. Split race data using season as a primary key.

C. Use BigQuery for its scalability and ability to add columns to a schema. Partition race data based on season.

D. Use Cloud SQL for its ability to automatically manage storage increases and compatibility with MySQL. Use separate database instances for each season.

[HRL case study] Diagnostic Question #1

For this question, refer to the Helicopter Racing League (HRL) case study. HRL is looking for a cost-effective approach for storing their race data such as telemetry. They want to keep all historical records, train models using only the previous season's data, and plan for data growth in terms of volume and information collected. You need to propose a data solution.

Considering HRL business requirements and the goals expressed by CEO S. Hawke, what should you do?

A. Use Firestore for its scalable and flexible document-based database. Use collections to aggregate race data by season and event.

B. Use Cloud Spanner for its scalability and ability to version schemas with zero downtime. Split race data using season as a primary key.

C. Use BigQuery for its scalability and ability to add columns to a schema. Partition race data based on season.

D. Use Cloud SQL for its ability to automatically manage storage increases and compatibility with MySQL. Use separate database instances for each season.

[HRL case study] Diagnostic Question #2

For this question, refer to the Helicopter Racing League (HRL) case study. A recent finance audit of cloud infrastructure noted an exceptionally high number of Compute Engine instances are allocated to do video encoding and transcoding. You suspect that these Virtual Machines are zombie machines that were not deleted after their workloads completed. You need to quickly get a list of which VM instances are idle.

What should you do?

A. Log into each Compute Engine instance and collect disk, CPU, memory, and network usage statistics for analysis.

B. Use the gcloud compute instances list to list the virtual machine instances that have the idle: true label set.

C. Use the gcloud recommender command to list the idle virtual machine instances.

D. From the Google Console, identify which Compute Engine instances in the managed instance groups are no longer responding to health check probes.

[HRL case study] Diagnostic Question #2

For this question, refer to the Helicopter Racing League (HRL) case study. A recent finance audit of cloud infrastructure noted an exceptionally high number of Compute Engine instances are allocated to do video encoding and transcoding. You suspect that these Virtual Machines are zombie machines that were not deleted after their workloads completed. You need to quickly get a list of which VM instances are idle.

What should you do?

A. Log into each Compute Engine instance and collect disk, CPU, memory, and network usage statistics for analysis.

B. Use the gcloud compute instances list to list the virtual machine instances that have the idle: true label set.

C. Use the gcloud recommender command to list the idle virtual machine instances.

D. From the Google Console, identify which Compute Engine instances in the managed instance groups are no longer responding to health check probes.

[HRL case study] Diagnostic Question #3

For this question, refer to the Helicopter Racing League (HRL) case study. Your team is in charge of creating a payment card data vault for card numbers used to bill tens of thousands of viewers, merchandise consumers, and season ticket holders. You need to implement a custom card tokenization service that meets the following requirements:

- It must provide low latency at minimal cost.
- It must be able to identify duplicate credit cards and must not store plaintext card numbers.
- It should support annual key rotation.

Which storage approach should you adopt for your tokenization service?

A. Store the card data in Secret Manager after running a query to identify duplicates.

B. Encrypt the card data with a deterministic algorithm stored in Firestore using Datastore mode.

C. Encrypt the card data with a deterministic algorithm and shard it across multiple Memorystore instances.

D. Use column-level encryption to store the data in Cloud SQL.

[HRL case study] Diagnostic Question #3

For this question, refer to the Helicopter Racing League (HRL) case study. Your team is in charge of creating a payment card data vault for card numbers used to bill tens of thousands of viewers, merchandise consumers, and season ticket holders. You need to implement a custom card tokenization service that meets the following requirements:

- It must provide low latency at minimal cost.
- It must be able to identify duplicate credit cards and must not store plaintext card numbers.
- It should support annual key rotation.

Which storage approach should you adopt for your tokenization service?

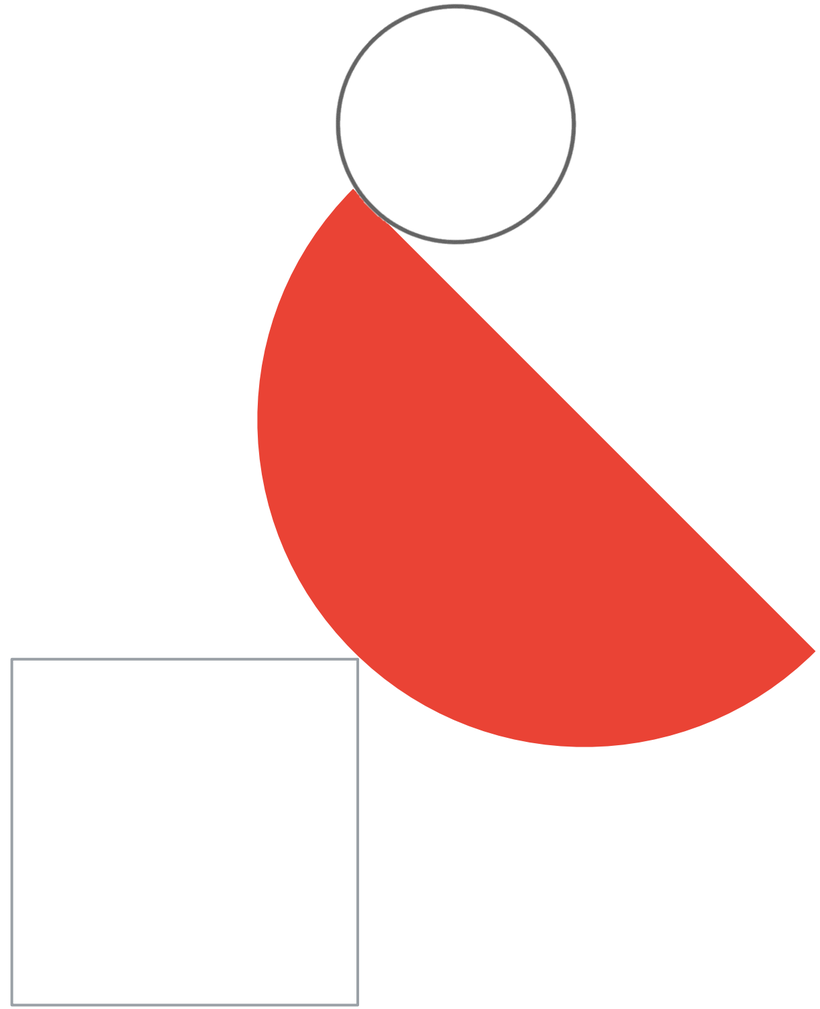
A. Store the card data in Secret Manager after running a query to identify duplicates.

B. Encrypt the card data with a deterministic algorithm stored in Firestore using Datastore mode.

C. Encrypt the card data with a deterministic algorithm and shard it across multiple Memorystore instances.

D. Use column-level encryption to store the data in Cloud SQL.

[optional] Links to useful materials



Optional materials 1

[READING]

- Boto configuration file | Cloud Storage
- Use customer-supplied encryption keys | Cloud Storage
- Object change notification | Cloud Storage
- Read about Database Migration Service.
- What are the options for connecting to Cloud SQL instance:
  a. https://cloud.google.com/sql/docs/mysql/connect-overview
  b. https://cloud.google.com/sql/docs/postgres/external-connection-methods
- How to choose optimal AI/ML path in GCP?

[VIDEOS]

- Cloud Networking 103 (Securing Network): Cloud OnAir: CE TV: Google Cloud Networking 103 - Securing your Network
- Google Cloud Storage options: Difference between object store, block store and file store | Google Cloud Storage options
- GCS Offline Transfer Appliance: Introducing Google Cloud's Transfer Appliance
- How to transfer data to GCS: How to transfer data to Google Cloud? #GCPSketchnote
- Authentication controls for Cloud Storage
- What's new with Cloud SQL

Optional materials 2

- [IMPORTANT TO KNOW] Different patterns for connecting to Cloud SQL: Cloud SQL: Concepts of Networking
- Great demo of how to centralize network management and set up Shared VPC in GCP: Level Up From Zero Episode 4: Shared VPC
- Accelerating cloud migrations with managed databases: Accelerating cloud migration with managed databases
- [Highly recommended] Choose your database on Google Cloud: Choose your database on Google Cloud
- Introducing Database Migration Service: Introducing Database Migration Service
- How to achieve high resiliency and availability with GCP: How to achieve high resiliency and availability with Google Cloud infrastructure
- Deploying MongoDB via GCP Marketplace: Deploying MongoDB from Google Cloud Marketplace
- Infrastructure as code with Terraform and Cloud Run: Infrastructure as code with Terraform and Cloud Run
- Build ETL Pipelines using Cloud Dataflow: Build ETL Pipelines using Cloud Dataflow

Optional materials 3

[PODCASTS]

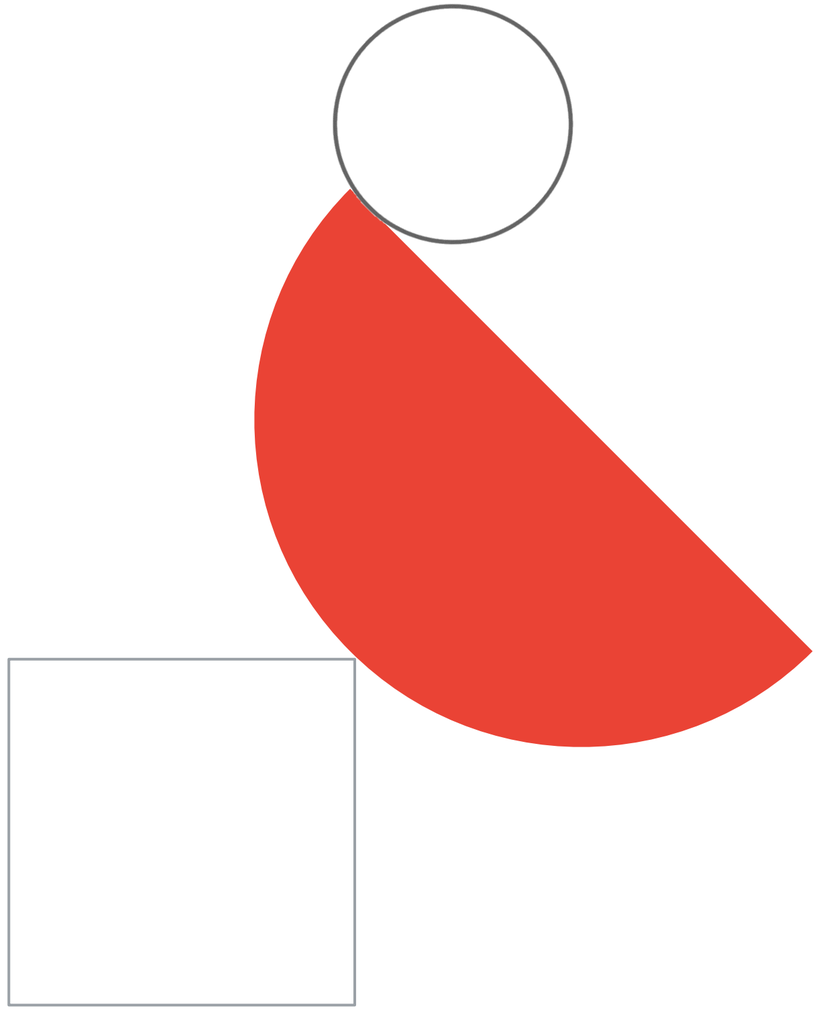
- Cloud SQL
- Database Migration Service
- Beam and Spark
- Cloud Dataflow

[DEEP DIVES]

- The battle of relational and non-relational databases | SQL vs NoSQL Explained: The battle of relational and non-relational databases | SQL vs NoSQL Explained
- [video] How to accelerate migration to GCP: Tools and services to accelerate your migration to Google Cloud
- 5 ways Google can help you succeed in the multicloud world.

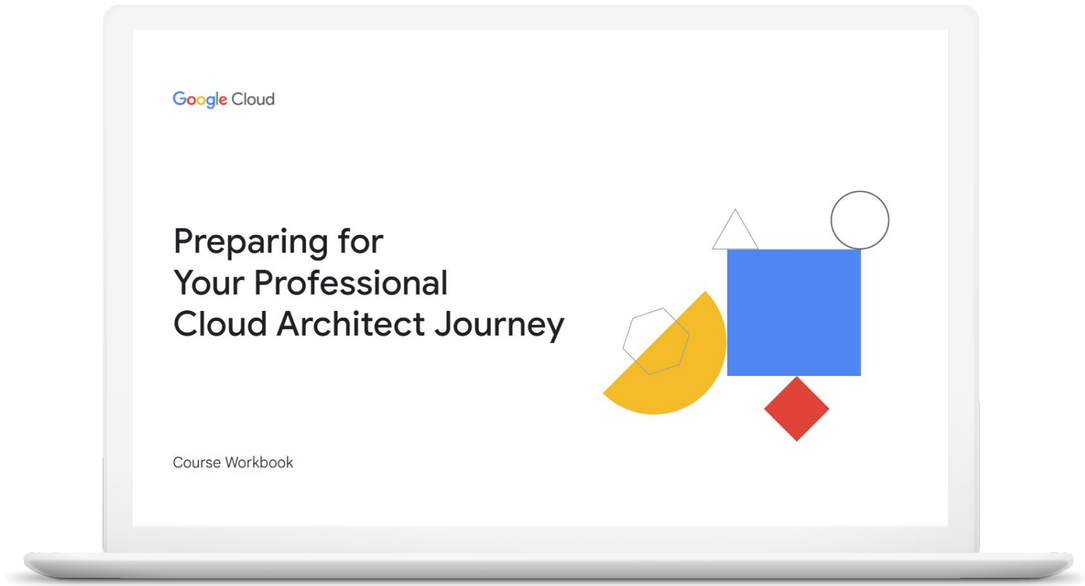
Diagnostic Questions

for Exam Guide Section 2: Managing and provisioning a solution infrastructure



PCA Exam Guide Section 2:

Managing and provisioning a solution infrastructure



2.1 Configuring network topologies

2.2 Configuring individual storage systems

2.3 Configuring compute systems

2.1 Configuring network topologies

Considerations include:

- Extending to on-premises environments (hybrid networking)
- Extending to a multicloud environment that may include Google Cloud to Google Cloud communication
- Security protection (e.g. intrusion protection, access control, firewalls)

2.1 Diagnostic Question 01 Discussion

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's operations suite to view audit logs for the networks you need to monitor.

D. Enable **VPC Service Controls**, and use **Google Cloud's operations suite** to view audit logs for the networks you need to monitor.

2.1 Diagnostic Question 01 Discussion

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's operations suite to view audit logs for the networks you need to monitor.

D. Enable **VPC Service Controls**, and use **Google Cloud's operations suite** to view audit logs for the networks you need to monitor.

2.1 Diagnostic Question 02 Discussion

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?

A. Deploy the web application using the **App Engine standard environment** using a global external HTTP(S) 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 HTTP(S) 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.

C. Deploy the web application using the **App Engine standard environment** using a global external HTTP(S) load balancer and a network endpoint group. **Use a managed instance group** for the backend chat servers. **Use a global SSL proxy 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 HTTP(S) load balancer and a network endpoint group. **Use a managed instance group** for the backend chat servers. Use an **external network load balancer to load-balance traffic** across the backend chat servers.

2.1 Diagnostic Question 02 Discussion

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?

A. Deploy the web application using the **App Engine standard environment** using a global external HTTP(S) 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 HTTP(S) 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.

C. Deploy the web application using the **App Engine standard environment** using a global external HTTP(S) load balancer and a network endpoint group. **Use a managed instance group** for the backend chat servers. **Use a global SSL proxy 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 HTTP(S) load balancer and a network endpoint group. **Use a managed instance group** for the backend chat servers. Use an **external 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]

[Choosing a Network Connectivity product | Google Cloud]

[Cloud VPN overview]

[Best practices | Cloud Interconnect]

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

2.2 Configuring individual storage systems

Considerations include:

- Data storage allocation
- Data processing/compute provisioning
- Security and access management
- Network configuration for data transfer and latency
- Data retention and data life cycle management
- Data growth planning

2.2 Diagnostic Question 03 Discussion

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 Diagnostic Question 03 Discussion

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

2.3 Configuring compute systems

Considerations include:

- Compute resource provisioning
- Compute volatility configuration (preemptible vs. standard)
- Network configuration for compute resources (Google Compute Engine, Google Kubernetes Engine, serverless networking)
- Infrastructure orchestration, resource configuration, and patch management
- Container orchestration

2.3 Diagnostic Question 04 Discussion

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.

2.3 Diagnostic Question 04 Discussion

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.

2.3 Diagnostic Question 05 Discussion

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 X or Partner Interconnect to **connect the on-premises network where your application is running to your VPC**. Configure an endpoint for a global external HTTP(S) load balancer that connects to the existing VMs.

D. Write Terraform scripts to deploy the application as **Compute Engine instances**.

2.3 Diagnostic Question 05 Discussion

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 X or Partner Interconnect to **connect the on-premises network where your application is running to your VPC**. Configure an endpoint for a global external HTTP(S) load balancer that connects to the existing VMs.

D. Write Terraform scripts to deploy the application as **Compute Engine instances**.

2.3 Diagnostic Question 06 Discussion

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.

What should you do?

A. Automate the deployment with **Terraform scripts**.

B. Automate the deployment using scripts containing **gcloud commands**.

C. Use **Google Kubernetes Engine (GKE)** to create deployments and manifests for your applications.

D. Develop in **Docker containers** for portability and ease of deployment.

2.3 Diagnostic Question 06 Discussion

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.

What should you do?

A. Automate the deployment with **Terraform scripts**.

B. Automate the deployment using scripts containing **gcloud commands**.

C. Use **Google Kubernetes Engine (GKE)** to create deployments and manifests for your applications.

D. Develop in **Docker containers** for portability and ease of deployment.

2.3 Diagnostic Question 07 Discussion

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.

2.3 Diagnostic Question 07 Discussion

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.

2.3 Diagnostic Question 08 Discussion

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.

C. 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.**

2.3 Diagnostic Question 08 Discussion

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.

C. 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.**

2.3 Diagnostic Question 09 Discussion

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.

2.3 Diagnostic Question 09 Discussion

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.

2.3 Diagnostic Question 10 Discussion

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 HTTP(S) 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 HTTP(S) 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 SSL proxy 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 **SSL proxy 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 Diagnostic Question 10 Discussion

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 HTTP(S) 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 HTTP(S) 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 SSL proxy 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 **SSL proxy 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?

Make sure to...

Enjoy the journey as much as the destination!