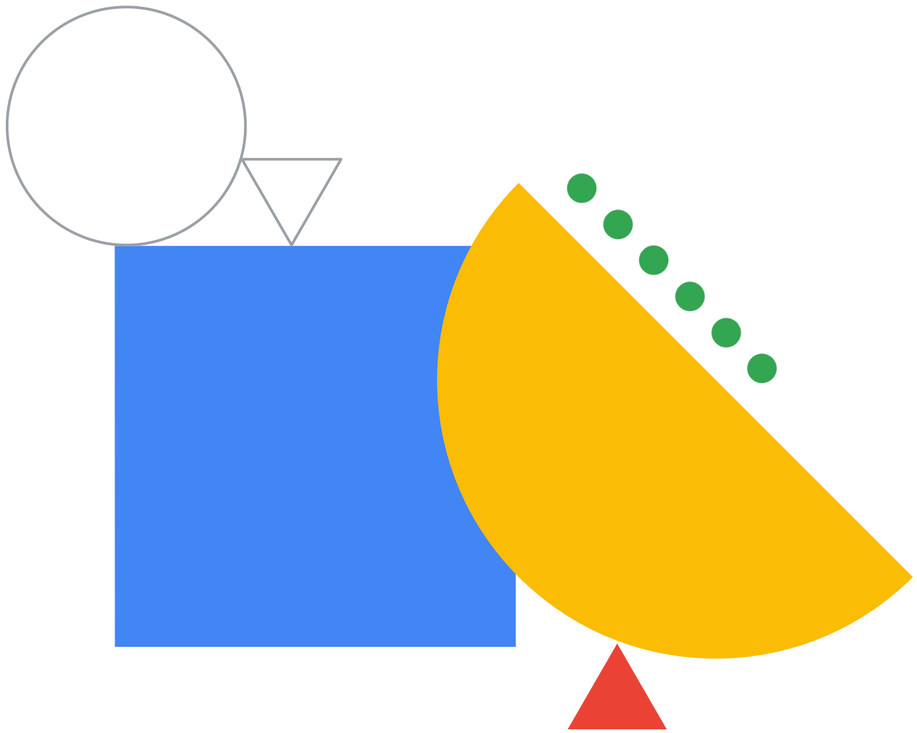
Google Cloud

Preparing for Your Associate Cloud Engineer Journey



Module 2: Planning and Configuring Cloud Solutions

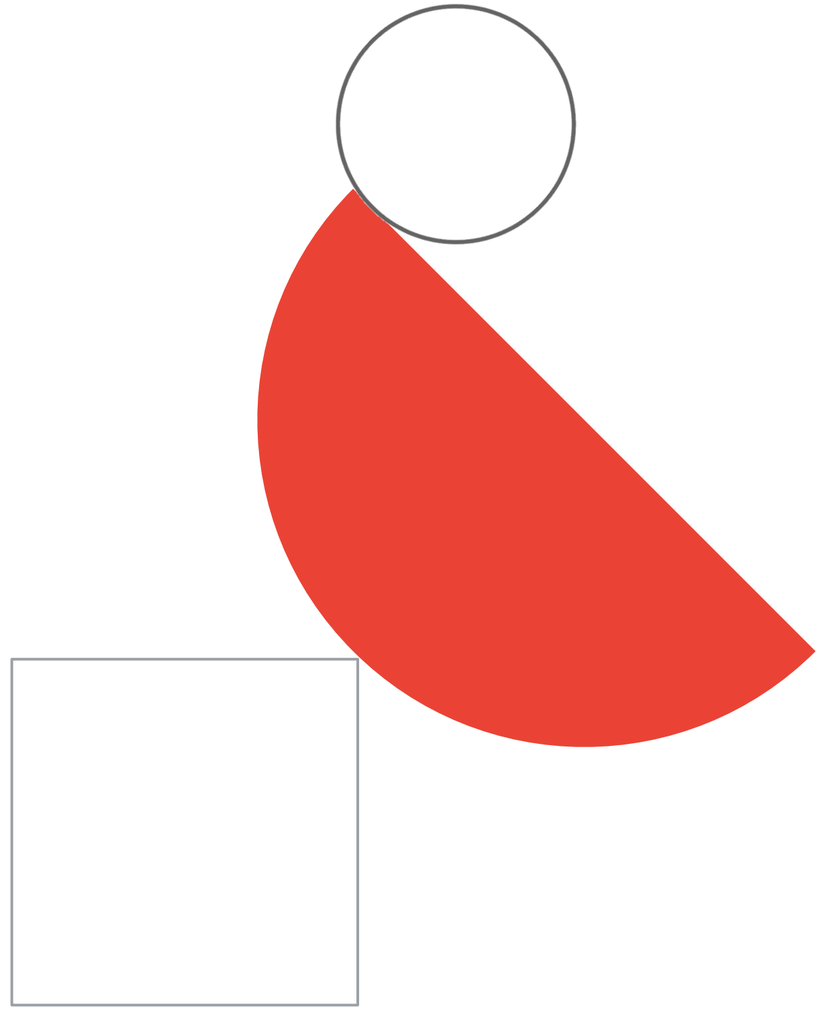
Module agenda

01 Selecting resources for Cymbal Superstore's cloud solutions

02 Diagnostic questions

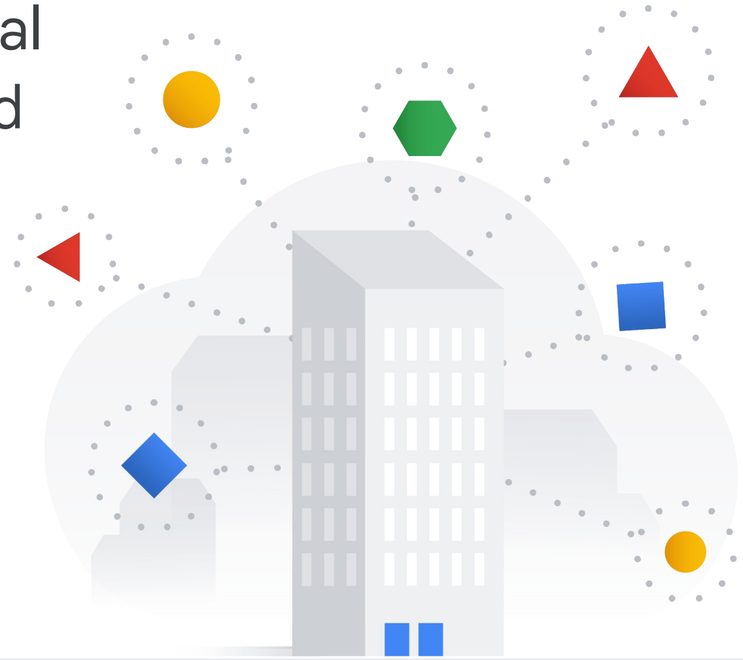
03 Review and study planning

Selecting resources for Cymbal Superstore's cloud solutions



The next step:

planning and configuring Cymbal Superstore's cloud solutions



- Planning and estimating Google Cloud pricing using the Pricing Calculator
- Planning and configuring compute resources
- Planning and configuring data storage options
- Planning and configuring network resources

Cymbal Superstore

Cymbal Superstore's existing applications

Ecommerce

Cymbal Superstore has an existing web application that provides an interface for customers to look at and order products.

Requirements:

- Compute: Container architecture
- Data: Relational backend
- Networking: Needs to be globally available
- Need analytical capabilities to inform marketing efforts

Transportation Management

Delivery services is becoming an important aspect of Cymbal Superstore's customer interactions. Cymbal Superstore would like to use Google Services to keep track of truck location.

Requirements:

- Dashboard of truck location in near real-time
- Analysis of truck mileage for preventive maintenance

Supply Chain

Cymbal Superstore has decided to migrate their legacy supply chain application to the cloud

Requirements:

- Available local to their HQ
- Currently implemented in virtual machines with a LINUX operating system and a LAMP stack

Cymbal Superstore's ecommerce solution

Ecommerce Requirements

- Compute: Container architecture
- Data: Relational backend
- Networking: Needs to be globally available
- Need analytical capabilities to inform marketing efforts

Ecommerce Cloud Solution

- Compute: Google Kubernetes Engine
- Data: Cloud Spanner
- Networking: External http(s) load balancing
- Feed historic sales data to BigQuery

Cymbal Superstore's transportation management solution

Transportation Management

Requirements

- Dashboard of truck location in near real-time
- Analysis of truck mileage for preventive maintenance

Transportation Management Cloud Solution

- Delivery: Pub/Sub
- Compute: Cloud Functions
- Orchestration: Dataflow
- Storage: Bigtable
- Network: Regional

Cymbal Superstore's supply chain solution

Supply Chain

Cymbal Superstore has decided to migrate their legacy supply chain application to the cloud

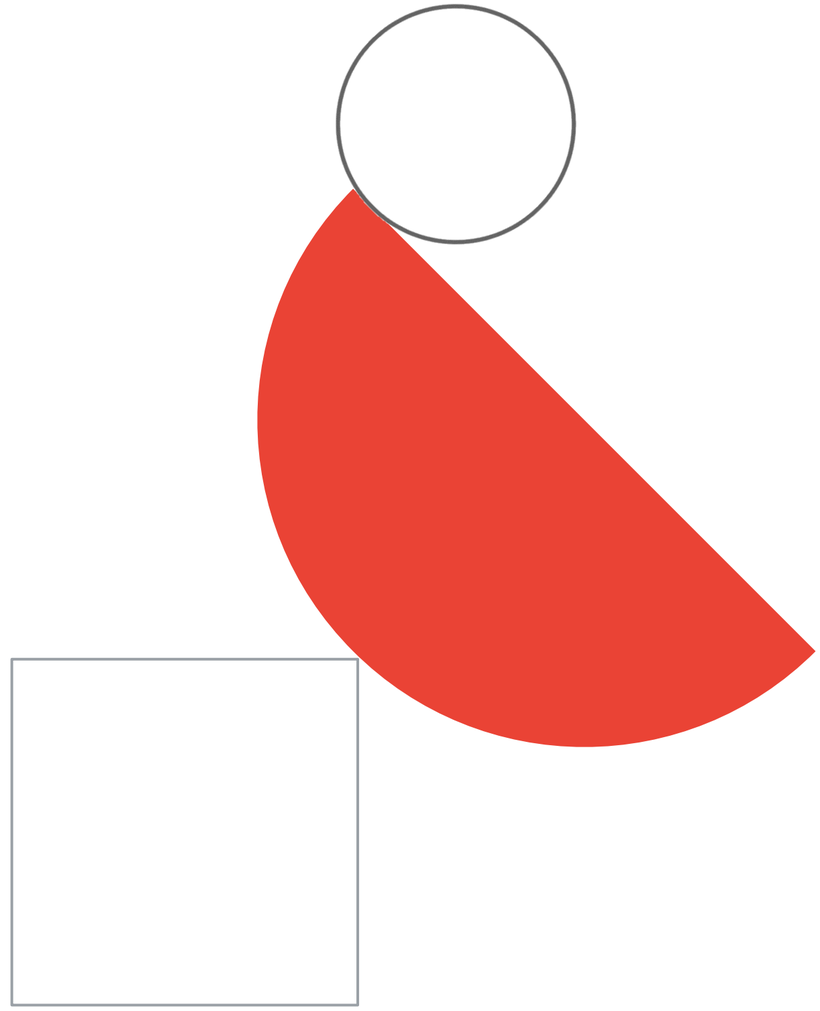
Requirements:

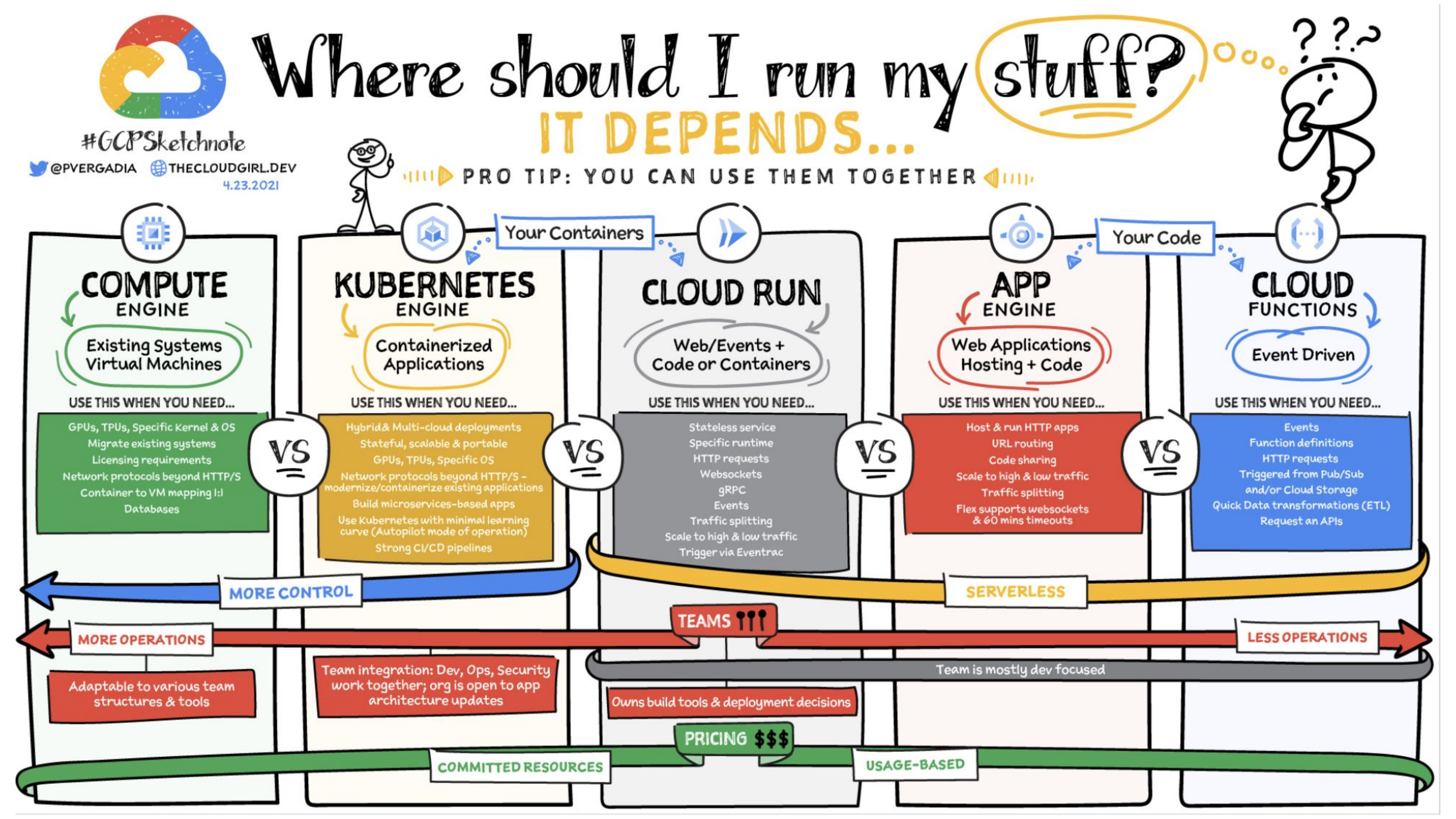
- Available local to their HQ
- Currently implemented in virtual machines with a LINUX operating system and a LAMP stack

Supply Chain Cloud Solution

- Compute: Compute Engine
- Data: Cloud SQL
- Networking:
  - Internal to VPC for backing database
  - Regional External https access

Google Compute Engine (GCE)




#GCPSketchnote
@PVERGADIA
THECLOUDGIRL.DEV
4.23.2021
Where should I run my stuff?
IT DEPENDS...
PRO TIP: YOU CAN USE THEM TOGETHER
Your Containers
Your Code
COMPUTE ENGINE
Existing Systems Virtual Machines
USE THIS WHEN YOU NEED...
GPUs, TPUs, Specific Kernel & OS
Migrate existing systems
Licensing requirements
Network protocols beyond HTTP/S
Container to VM mapping 1:1
Databases
VS
KUBERNETES ENGINE
Containerized Applications
USE THIS WHEN YOU NEED...
Hybrid& Multi-cloud deployments
Stateful, scalable & portable
GPUs, TPUs, Specific OS
Network protocols beyond HTTP/S - modernize/containerize existing applications
Build microservices-based apps
Use Kubernetes with minimal learning curve (Autopilot mode of operation)
Strong CI/CD pipelines
VS
CLOUD RUN
Web/Events + Code or Containers
USE THIS WHEN YOU NEED...
Stateless service
Specific runtime
HTTP requests
Websockets
gRPC
Events
Traffic splitting
Scale to high & low traffic
Trigger via Eventrac
VS
APP ENGINE
Web Applications Hosting + Code
USE THIS WHEN YOU NEED...
Host & run HTTP apps
URL routing
Code sharing
Scale to high & low traffic
Traffic splitting
Flex supports websockets & 60 mins timeouts
VS
CLOUD FUNCTIONS
Event Driven
USE THIS WHEN YOU NEED...
Events
Function definitions
HTTP requests
Triggered from Pub/Sub and/or Cloud Storage
Quick Data transformations (ETL)
Request an APIs
MORE CONTROL
SERVERLESS
MORE OPERATIONS
LESS OPERATIONS
TEAMS
Adaptable to various team structures & tools
Team integration: Dev, Ops, Security work together; org is open to app architecture updates
Owns build tools & deployment decisions
Team is mostly dev focused
PRICING $$$
COMMITTED RESOURCES
USAGE-BASED

Google Compute Engine

Infrastructure as a Service (IaaS)

- vCPUs (cores) and Memory (RAM)
- Persistent disks
- Networking
- Linux or Windows

Exam Tips: *GCE is a basic IaaS service, but there are lots of details you're expected to know:*

- *Differences between PD images / snapshots / VM images.*
- *How to troubleshoot VM not booting up properly*
- *Custom image vs public image + startup scripts*
- *VM price differ between regions*
- *PDs are network-attach devices and - as such - consume VM bandwidth.*
- *VM network performance scales with # of vCPUs.*
- *etc...*

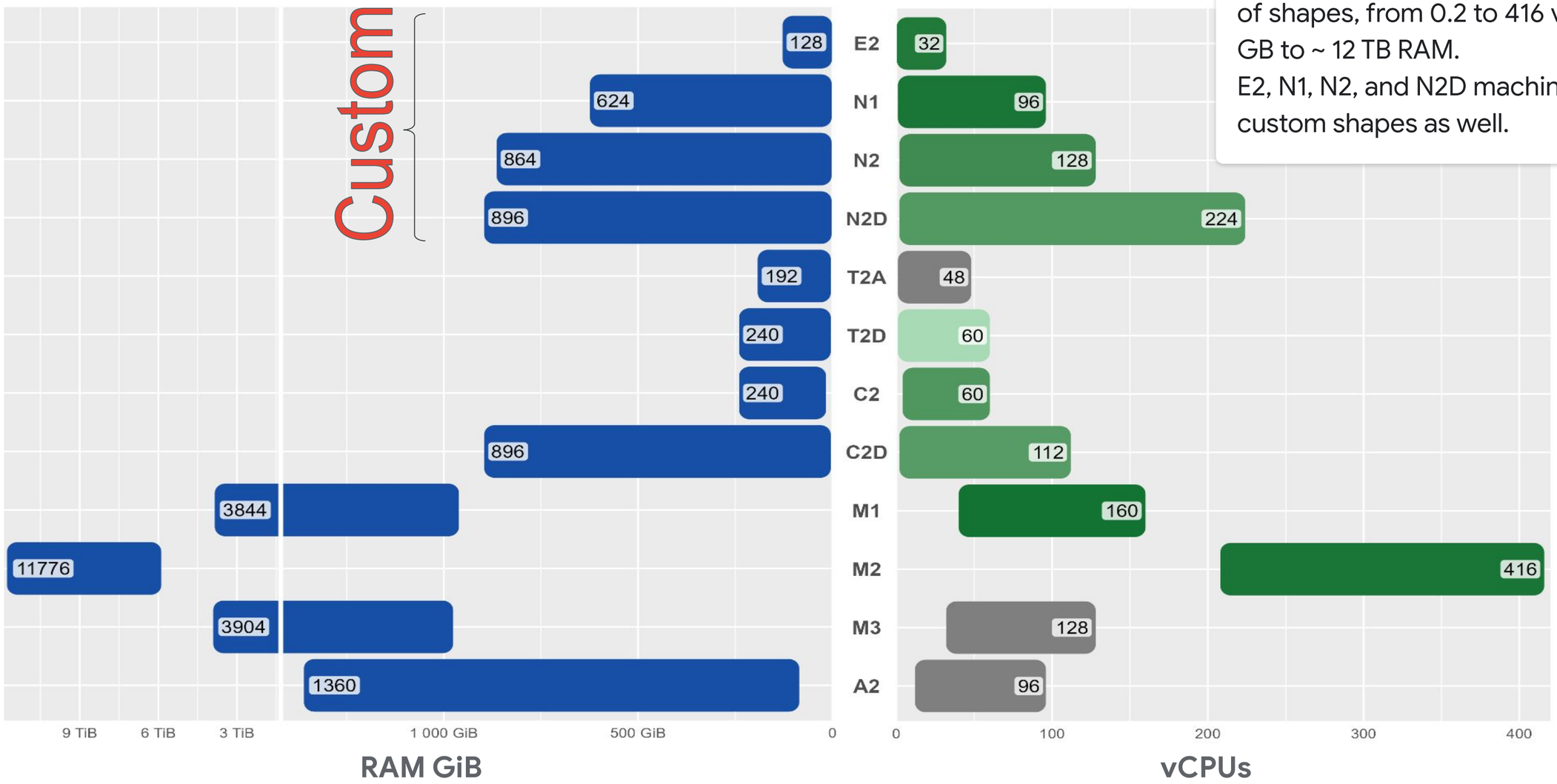
Compute Engine - how to differentiate between families?

Best TCO	Balanced	Scale-out Optimized	Workload-Optimized		
• Web Serving • Steady-state LOB apps • Dev & Test environments • Small prod environments	• Enterprise apps • Medium databases • Web & App Serving	• Scale-out Workloads • Web Serving • Containerized microservices	• EDA • HPC • Scientific Modeling • AAA Gaming	• SAP HANA • Largest in memory DBs • Real-time data analytics • In-memory cache	• ML • HPC • Massive parallelized computation
Cost savings a priority	Leading perf and perf/$	Best Perf/$ for scale out workloads	Highest performance CPUs	Most memory on Compute Engine	Highest performance GPUs
Cost-Optimized (E2)	**General Purpose** (N2 and N2D)	**ScaleOut optimized Tau (T2D, T2A)**	**Compute-Optimized** (C2, C2D)	**Memory-Optimized** (M1, M2, M3)	**Accelerator-Optimized** (A2)

Compute Engine: Max shapes by machine type

Exam Tip: *Custom machines can be used only for some VM families & up to 224vCPU/896 GB RAM*

Instances (VMs) are available in a wide range of shapes, from 0.2 to 416 vCPUs and from 1 GB to ~ 12 TB RAM.
E2, N1, N2, and N2D machine types provide custom shapes as well.

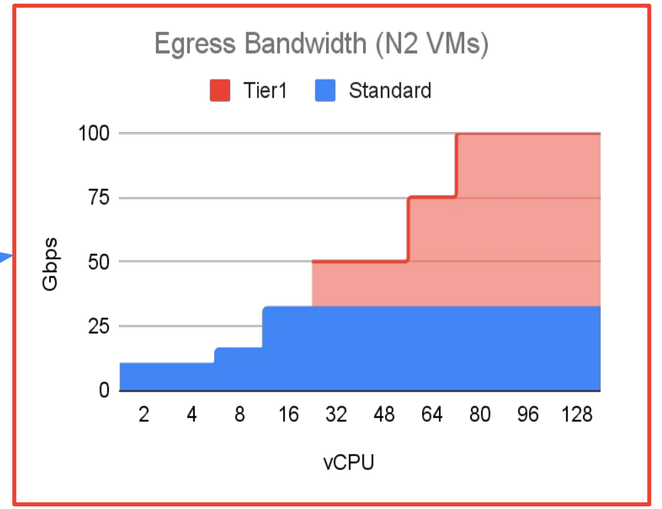


Compute Engine

Network perspective

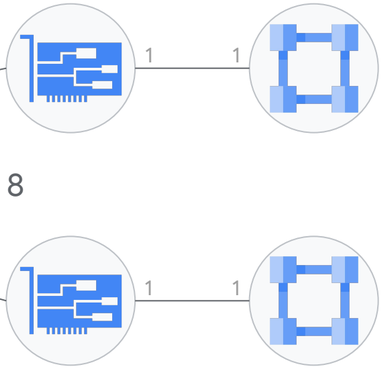
Exam Tips:

- *Network bandwidth limited & dependent on vCPU count (up to ~32Gbps for N2s + Tier1 extends further)*
- *You can expect the best network performance for traffic within the same zone, using internal IP addresses.*
- *Remember about multi-NIC VMs (up to 8)*
- *Storage is a network resource! => Network bandwidth shared between network AND disk activity*



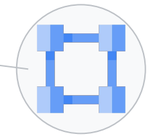
ingress

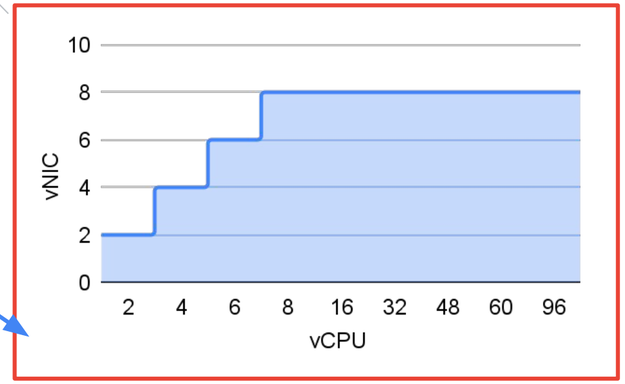
ingress

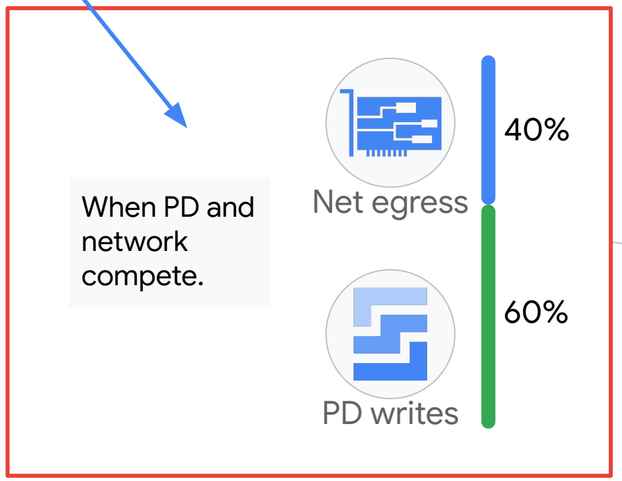


egress

egress



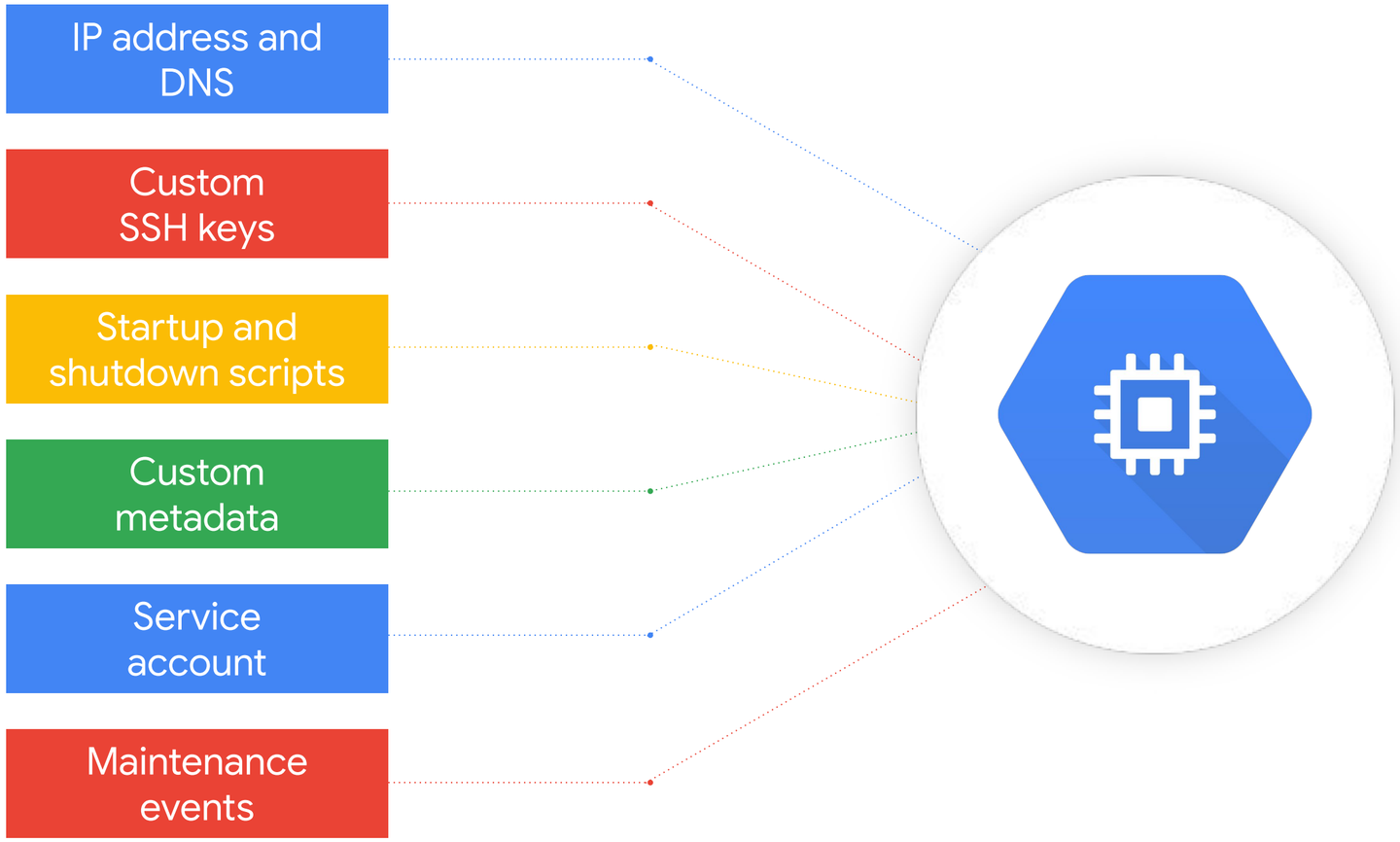




Compute Engine: Metadata Server

- The metadata server stores information about the instance or project.

  - Metadata request/response never leaves the physical host.
  - Metadata information is encrypted on the way to the virtual machine host.
  - Metadata server can generate a signed token for apps to verify the instance identity.



Compute Engine: Spot (Preemptible) VMs

Made for batch, fault-tolerant, and high throughput computing

Super-low-cost, short-term instances

- Up to 91% less than standard instances
- No maximum duration, may be preempted with 30-seconds notice (preemptible: max 24 hrs)
- Simple to use with graceful termination

Exam Tips:

- *Those use-cases usually pop up at the exam with regards to Spot VMs / Preemptibles.*
- *Can also be used in GKE clusters!*

Ideal for a variety of stateless, fault-tolerant workloads

- Genomics, pharmaceuticals
- Physics, math, computational chemistry
- Data processing (for example, with Hadoop or Cloud DataProc)
- Image handling, rendering, and media transcoding
- Monte Carlo simulations
- Financial services

Compute Engine: automate start & stop activities

Executed from metadata, either directly or from file:

- Startup:
  - gcloud compute instances create VM_NAME \
    --image-project=debian-cloud \
    --image-family=debian-10 \
    --metadata=startup-script='#! /bin/bash
    apt update
    apt -y install apache2
    cat <<EOF > /var/www/html/index.html
    <html><body> Linux startup script added directly. </body></html>
    EOF'

- Shutdown:
  - gcloud compute instances create example-instance --metadata-from-file=shutdown-script=FILE_PATH

To see output of startup/shutdown script:

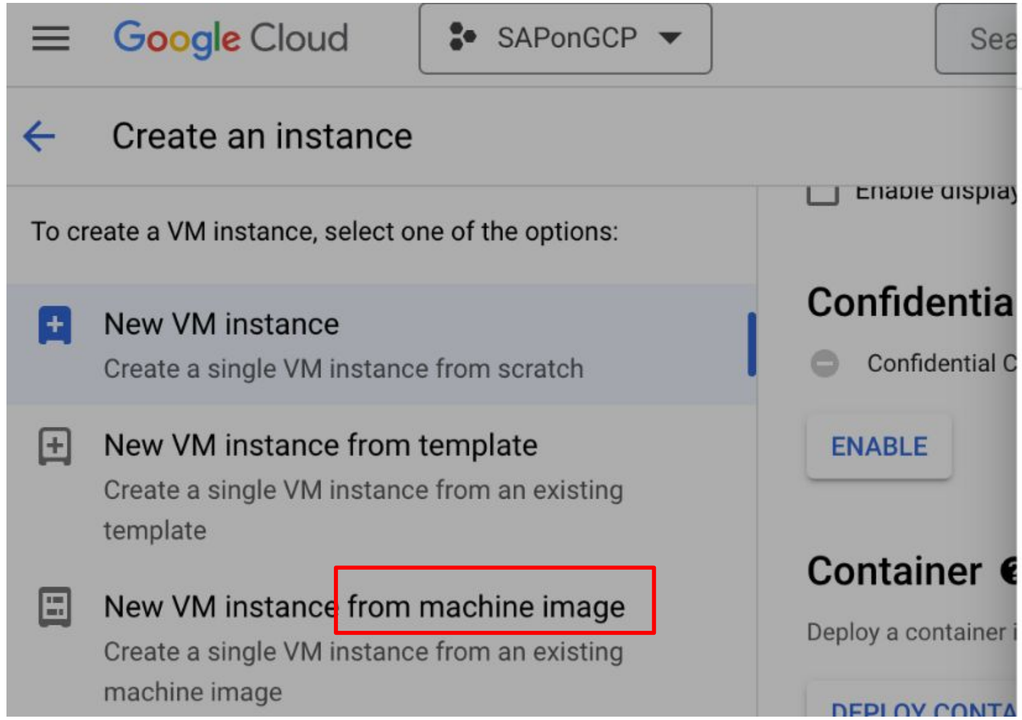
- gcloud compute instances create example-instance --metadata shutdown-script="#! /bin/bash
  > # Shuts down Apache server
  > /etc/init.d/apache2 stop"

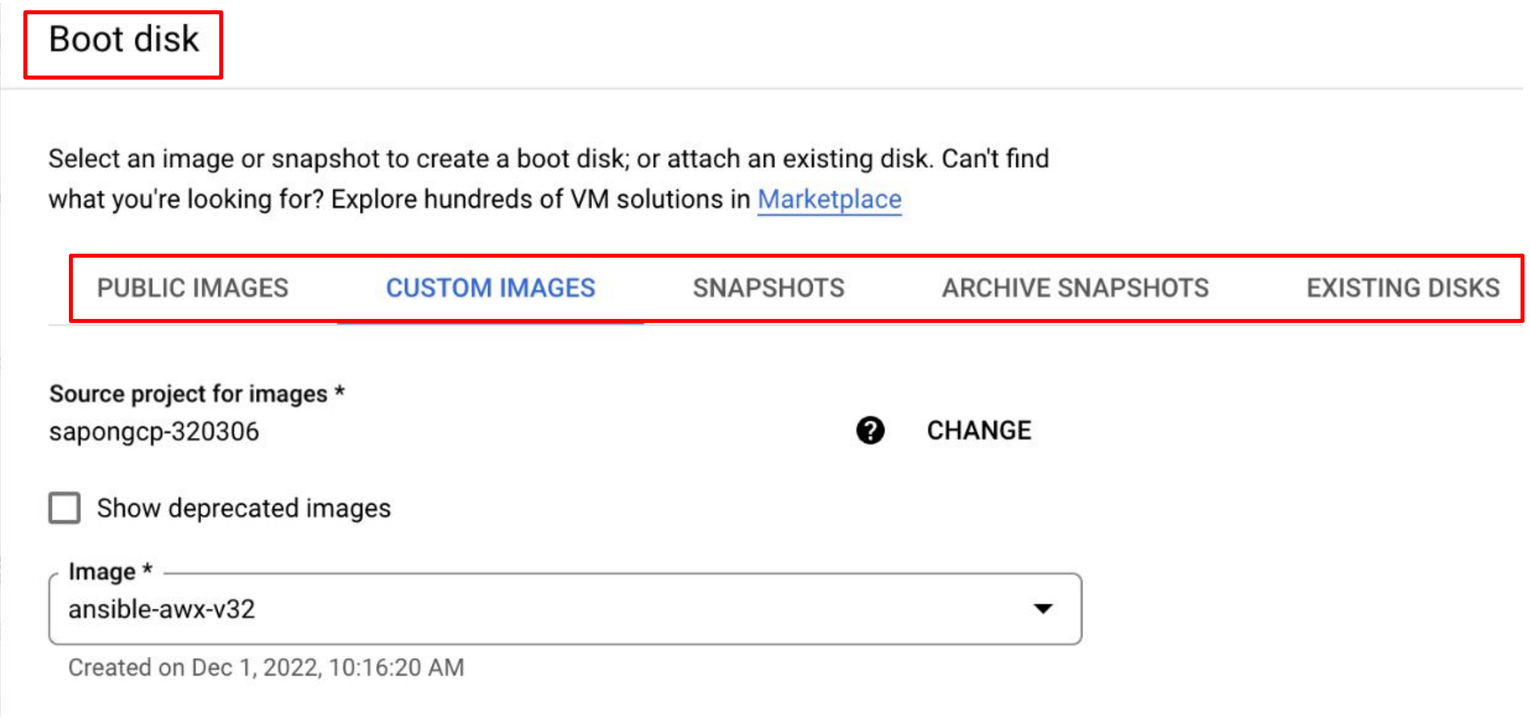
Exam Tips:

- *Startup / shutdown scripts are best-effort only!*
- *Startup / shutdown scripts are always run by root (Linux) / System (Windows)*
- *Shutdown scripts are especially useful for:*
  - *MIGs (to copy back processed data or logs before a VM goes down).*
  - *Spot / Preemptible VMs, which are much more vulnerable to be stopped.*
- *Startup / shutdown scripts can be set on VM or project (!!!) level -> will trigger for every VM. VM-level always take precedence (if exists, project-level script is not executed)*
- *Shutdown scripts have timeouts:*
  - *90s for standard instances*
  - *30s for Spot / Preemptible instances*

Compute Engine creation

public OS image vs custom OS image vs snapshot vs machine image





Exam Tips:

- *Custom images should be centralized and controlled from lifecycle perspective (know what are image families and image states*
- *Public / Custom OS image IS NOT the same as "machine image"*
- *You can create a VM based on all of those options (public / custom OS image, snapshot, existing disk, machine image)*
- *You can 'automate' post-processing with startup script, regardless of how boot disk was created.*

Shielded VMs

Exam Tips: *Using Shielded VMs is a best practice in GCP!*

Secure Boot	vTPM	Integrity Monitoring	Result/implications
ON	ON	ON	Most secure. Allows for use of vTPM for data encryption using vTPM protected key, Secure Boot to prevent malicious rootkits and bootkits, and Integrity Monitoring to alert to any changes in boot process. Secure Boot may not be compatible with customers drivers or other software.
OFF	ON	ON	Default when creating a GCP VM. Allows for use of vTPM for data encryption using vTPM protected key and Integrity Monitoring to alert to any changes in boot process. If customer has unsigned drivers or low level software this is the most secure option as Secure Boot would not be compatible.
OFF	OFF	OFF	Least secure. No benefits of Shielded VM. This is not recommended.

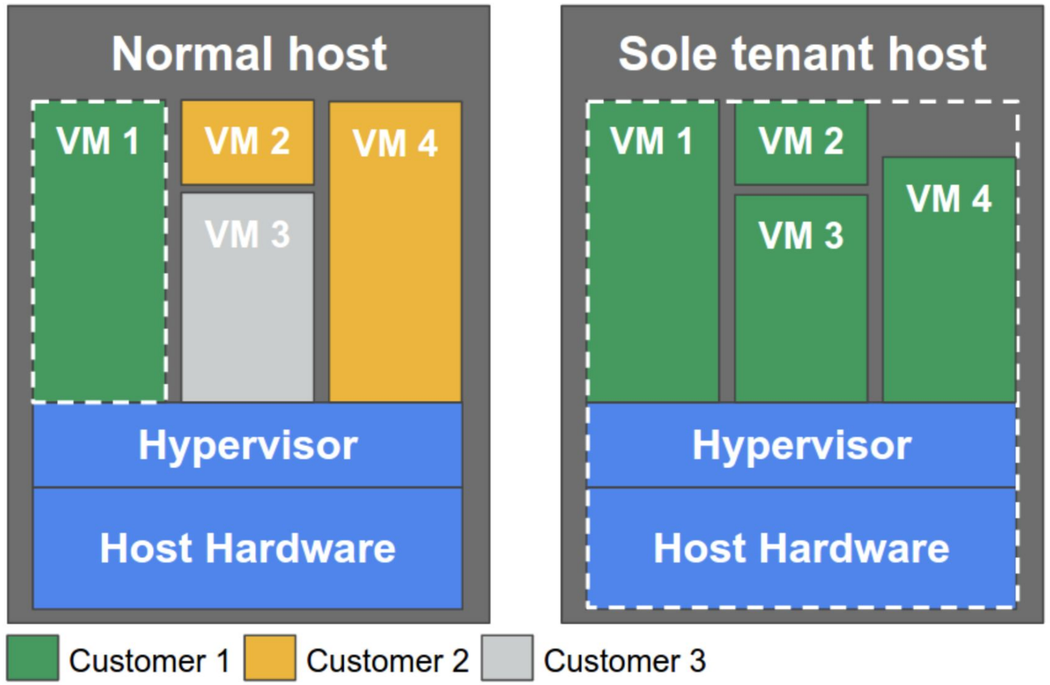
```
gcloud compute instances update instance name \
```

Feature	Flag to Turn On	Flag to Turn Off
Secure boot	--shielded-secure-boot	--no-shielded-secure-boot
vTPM (measure boot)	--shielded-vtpm	--no-shielded-vtpm
Integrity monitoring	--shielded-integrity-monitoring	--no-shielded-integrity-monitoring

Sole-Tenant Nodes

Regular VMs on regular machines, dedicated specifically to your workloads.

- Dedicated hardware
- Mix-and-match VMs to consume host resources
- Full access to host resources for 10% premium*

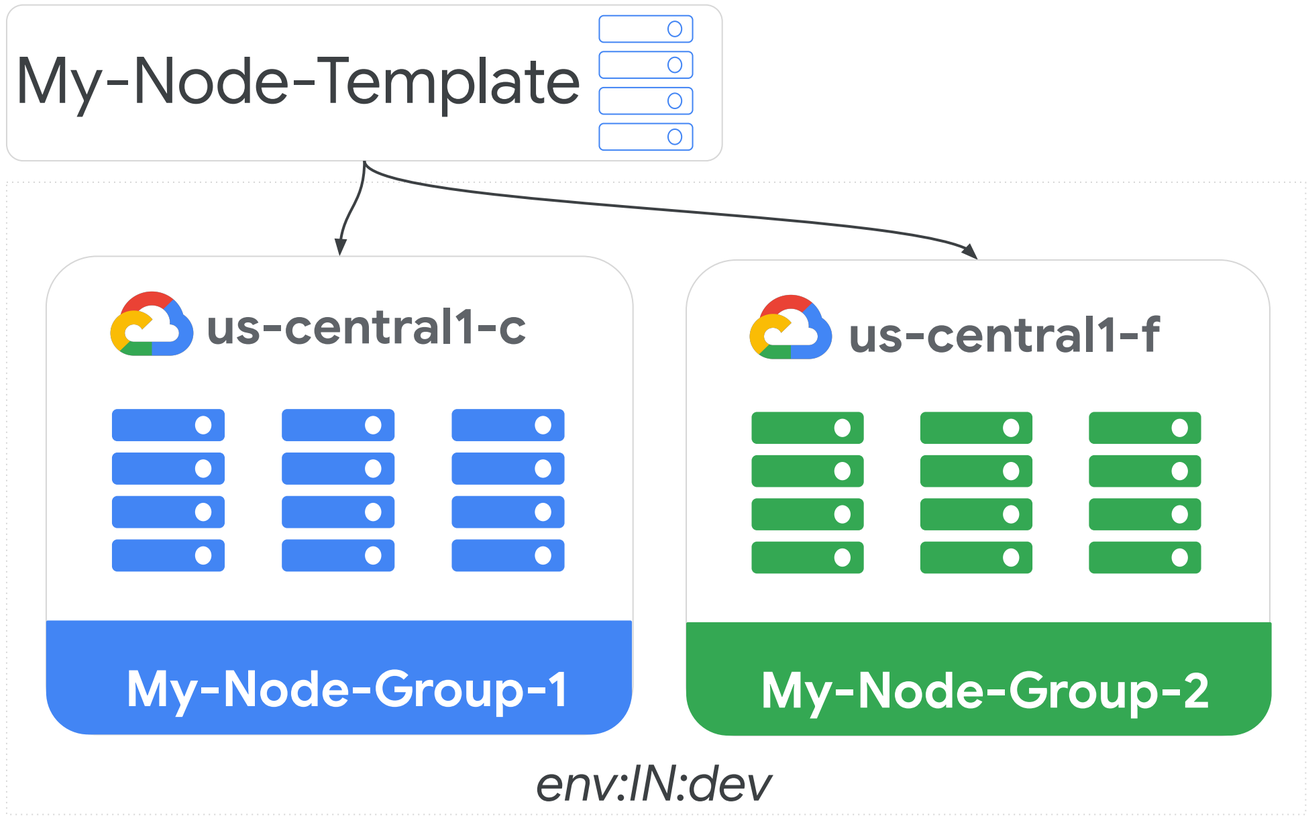


*10% Premium based on on-demand price

Quick Start for Sole-Tenant Nodes (1/2)

Each Sole-Tenant Node has a 1:1 mapping to a physical Host and represents a reserved host.

Step 1: Reserve Sole-Tenant Node(s)




```
// 1. CREATE NODE TEMPLATE
$ gcloud compute sole-tenancy \
node-templates create my-node-template \
--node-type n1-node-96-624 \
--region us-central1
```

```
// 2. CREATE NODE GROUP OF 3 NODES
$ gcloud compute sole-tenancy \
node-groups create my-node-group-1 \
--node-template my-node-template \
--target-size 3 \
--zone us-central1-c
```

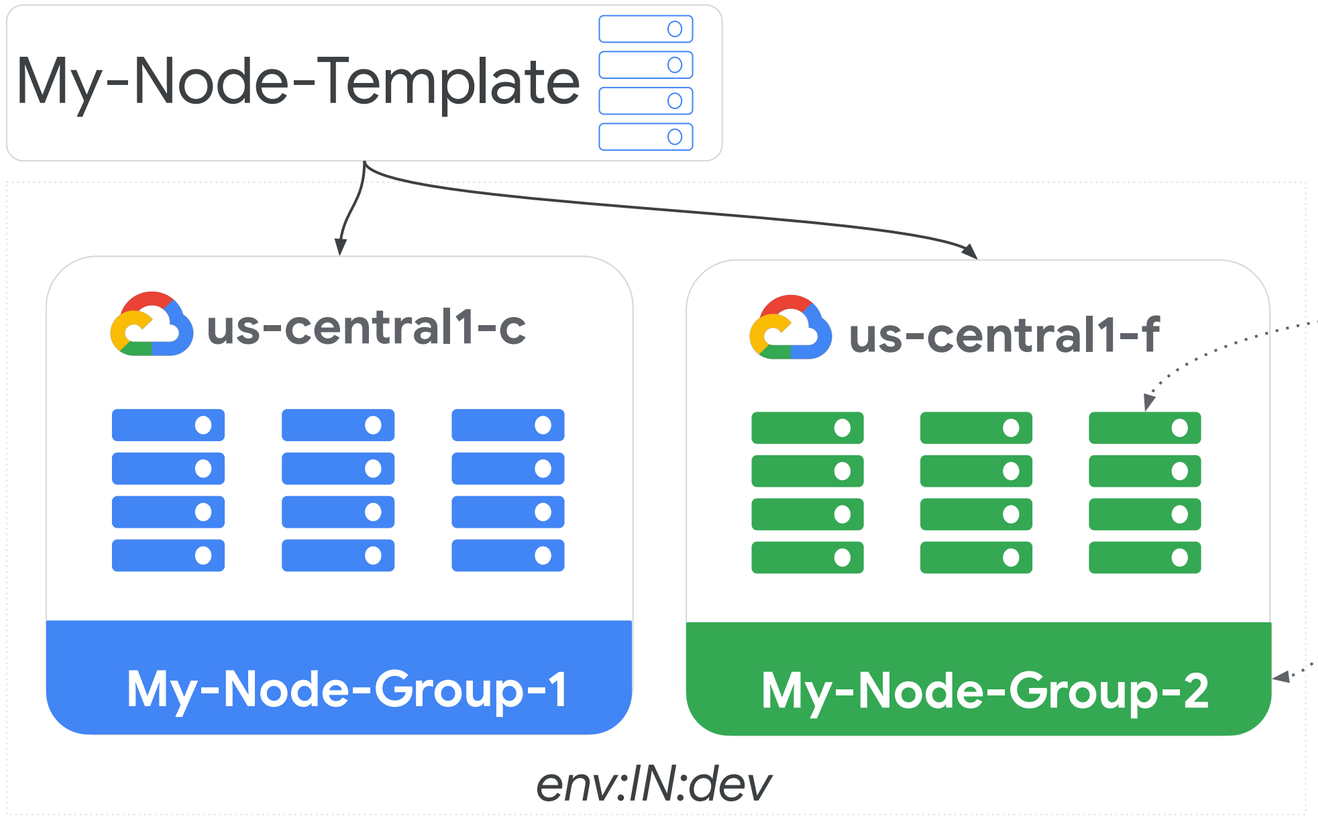
```
// 2b. [FOR ILLUSTRATION] CREATE ANOTHER
$ gcloud compute sole-tenancy \
node-groups create my-node-group-2 \
--node-template my-node-template \
--target-size 3 \
--zone us-central1-f
```

Google Cloud

Quick Start for Sole-Tenant Nodes (2/2)

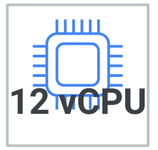
Exam Tips: *More info on provisioning VMs on sole-tenant nodes can be found here.*

Step 1: Reserve Sole-Tenant Node(s)



2. Schedule Instance(s)

VM



3 ways to schedule:

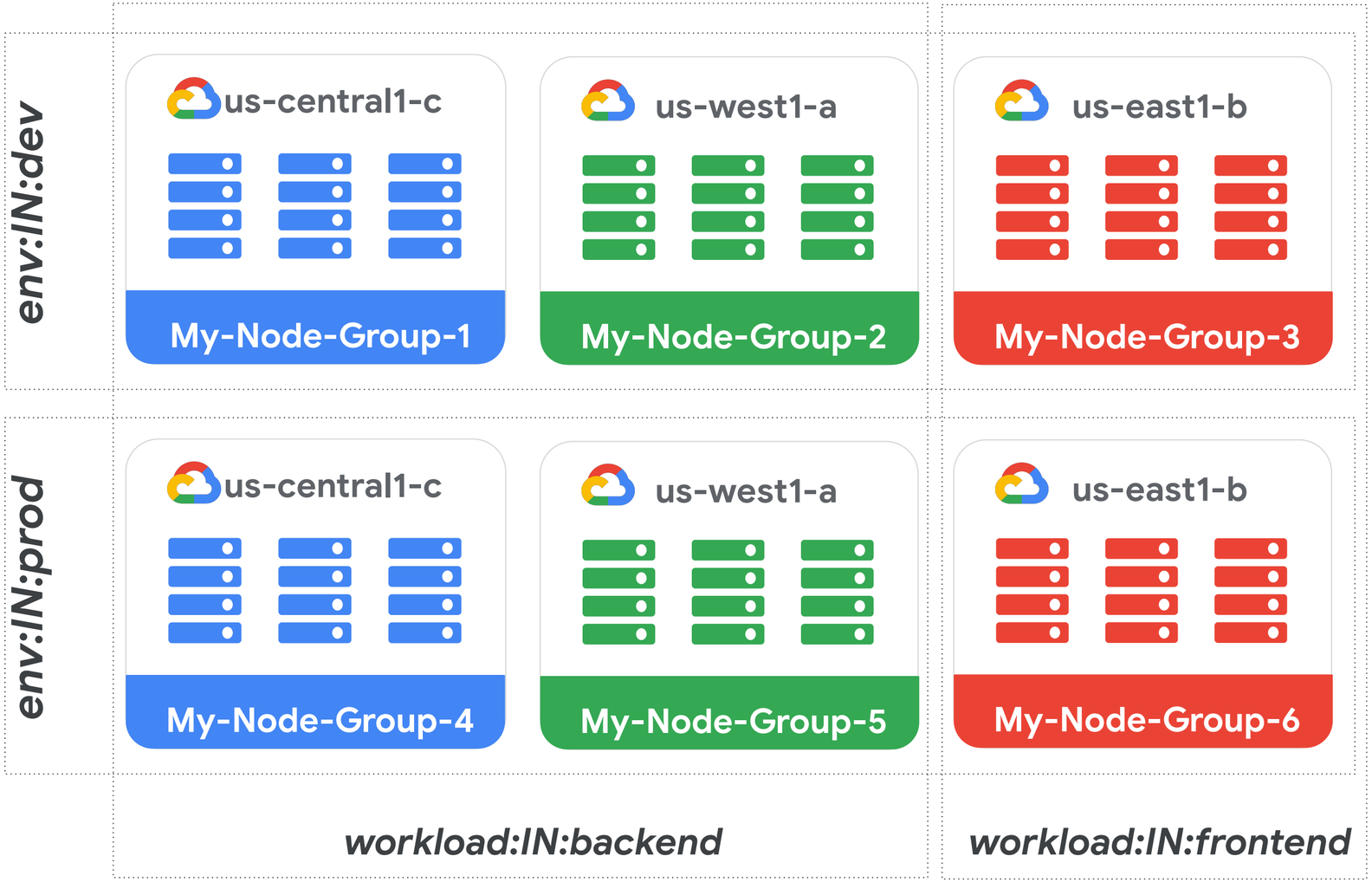
```
// SCHEDULE ONTO A SPECIFIC NODE
$ gcloud compute instances create \
INSTANCE_NAME --node=NODE_NAME
```

```
// SCHEDULE ON ANY NODE IN NODE GROUP
$ gcloud compute instances create \
INSTANCE_NAME \
--node-group=NODE_GROUP_NAME
```

```
// SCHEDULE ONTO ANY HOST WITH
// MATCHING LABELS
$ gcloud compute instances create
INSTANCE_NAME --zone ZONE
--node-affinity-file=FILE.json
```

Google Cloud

Sole-Tenant Nodes: Using Node Affinity Labels

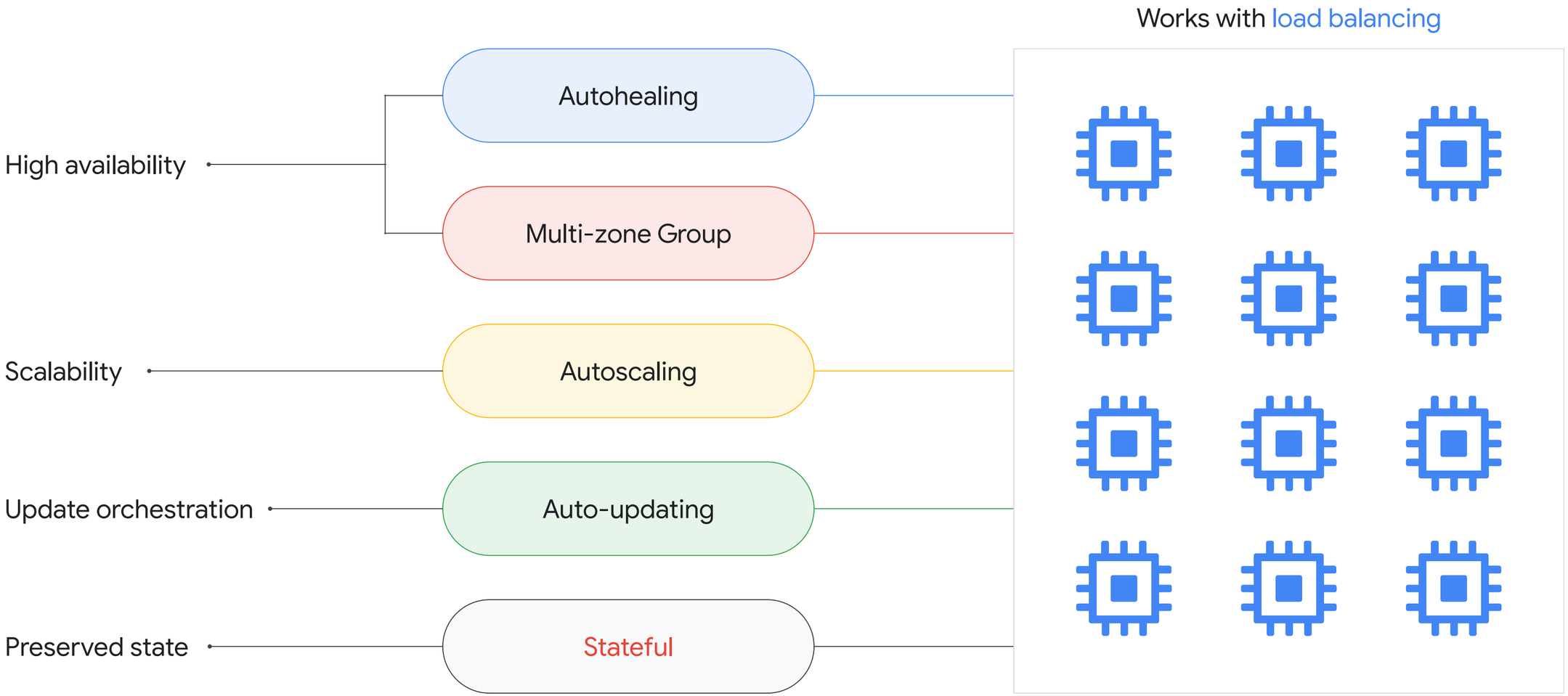



```
// SCHEDULE ONTO ANY HOST WITH
// MATCHING LABELS
$ gcloud compute instances create
INSTANCE_NAME --zone ZONE
--node-affinity-file=FILE.json
```

```
// SCHEDULE ONTO ANY HOST WITH
// IN MY-NODE-GROUP-1
$ gcloud compute instances create
INSTANCE_NAME --zone ZONE
--node-group my-node-group-1
```

Managed Instance Groups: Run VMs at Scale

Up to thousands of VMs



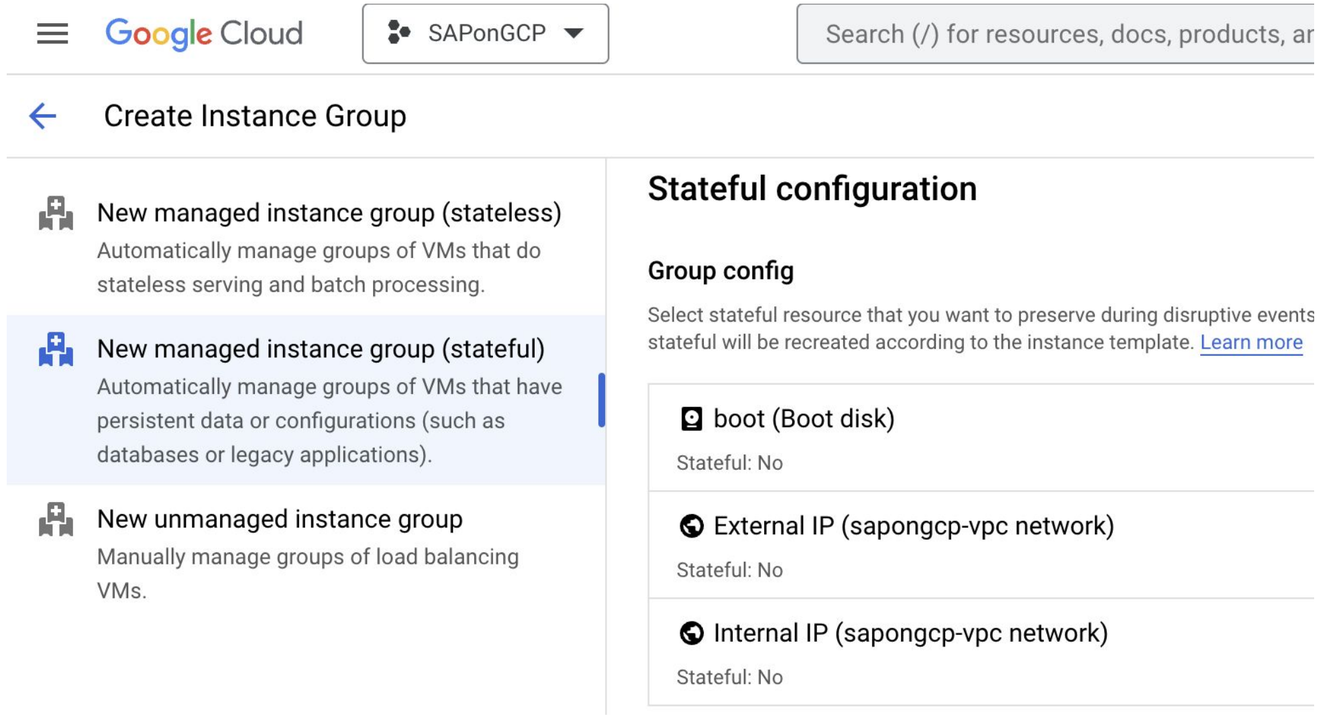
Exam Tips: *pros & cons of "ready" custom OS image vs public image + startup scripts*

Stateful vs stateless

And why stateless is usually preferred...

Exam Tips:

- *Here a look at this document.*
- *Prefer stateless. Use stateful only when necessary, eg:*
  - *Databases*
  - *Data processing apps (Kafka etc)*
  - *Legacy monoliths*



Stateful	Stateless (PREFERRED!)
Each server retains information about it's client sessions, such as the current state of an application or the content of a user shopping cart.	Server does not retain any information about the client sessions. Each request made by the client is treated as an independent transaction, and the server does not maintain any memory of previous requests.
Not perfect if we have multiple backends that can serve the requests...	Greater scalability and flexibility.
Can scale up easily. Can't scale down easily since each server keeps its' state.	Ability to scale up & down easily

Choosing instance groups for Compute Engine

Type of Instance Group		Properties of Instances	Feature
Unmanaged		Heterogeneous	
Managed		Homogeneous	Instance Templates Autoscaling
	Zonal	Same zone	Latency consistency
	Regional	Different zones	Reliability

Exam Tip:

- *Unmanaged are used to group EXISTING, different VMs under one "umbrella" and balance traffic to healthy ones only. For example, used in lift&shift migrations.*
- *You can't update existing instance template (need to create a new one)*
- *Know the difference between scale-out and scale-up!*

```
gcloud compute instance-groups managed create [*INSTANCE_GROUP_NAME*] \
    --size=[*SIZE*] \
    --template=[*INSTANCE_TEMPLATE_NAME*] \
    --zone=[*ZONE*]
```

MIG - Autoscaling

CPU Utilization

Treats the target CPU utilization level as a fraction of the average use of all vCPUs over time in the instance group

External HTTPS Capacity

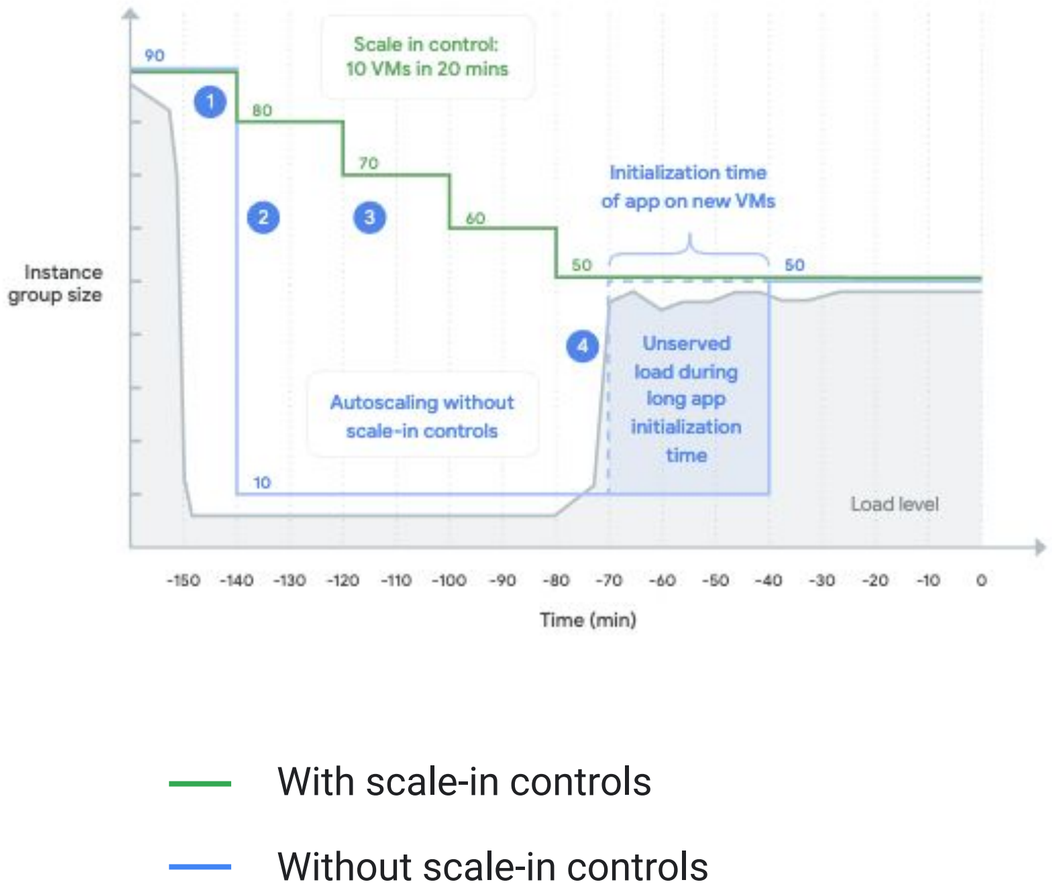
Autoscaling works with maximum backend utilization and maximum requests per second/instance

Cloud Monitoring Metrics

Per Instance or Per Group
Standard or custom metrics
Not for log-based metrics

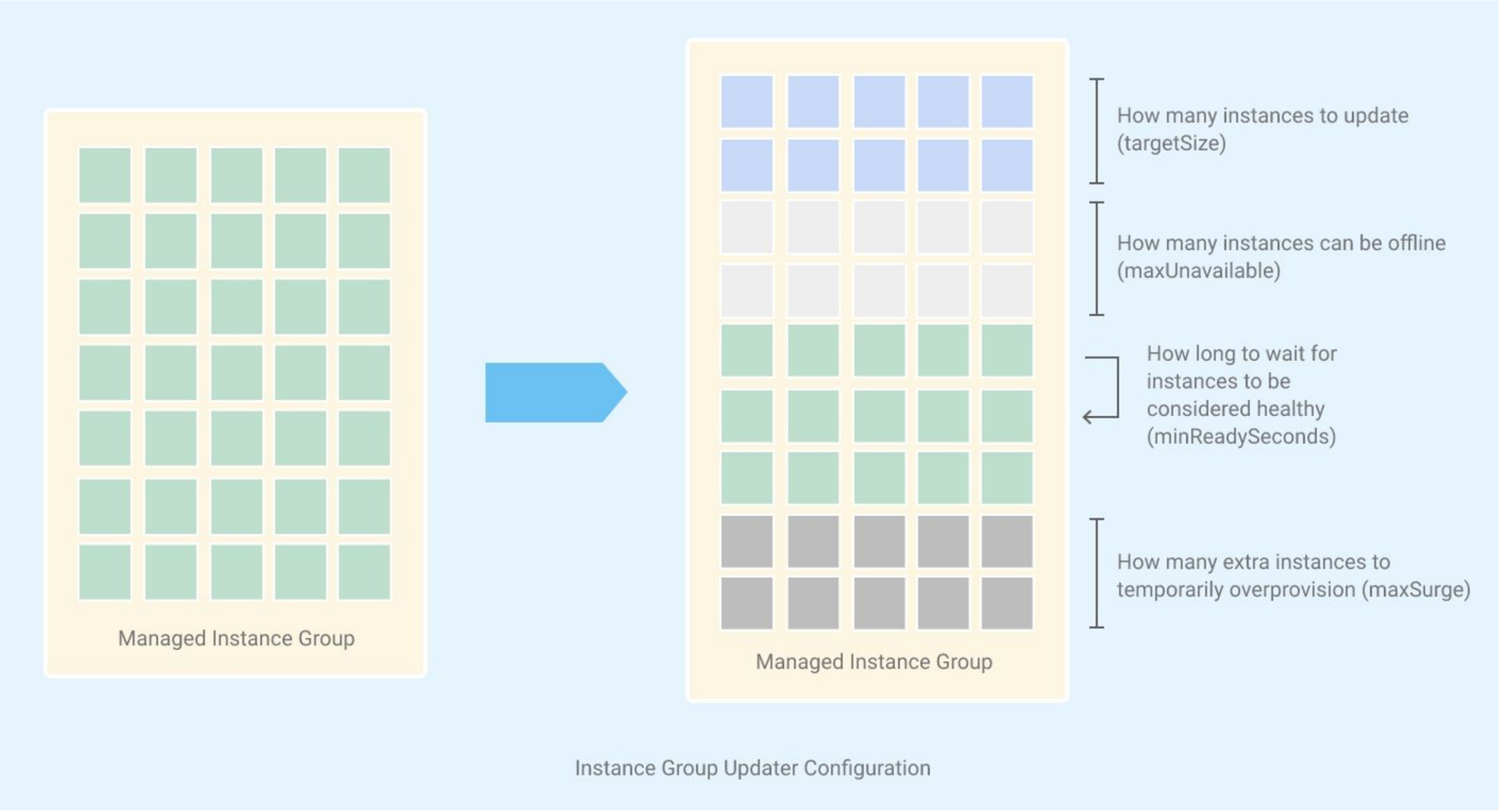
Schedules

Additional autoscaler
Up to 128 schedules
Min instances
Duration
Start time & Recurrence



Updating MIGs

= implementing new image versions



Exam Tip: *Know WELL how to rollout new versions to MIGs, incl. canary & rollback strategies*

Compute Engine: (most important) Organization Policy Constraints

Constraint	Description	Supported Prefixes
Disable VM serial port access	disables serial port access to Compute Engine VMs belonging to the organization, project, or folder constraints/compute.disableSerialPortAccess	"is:"
Disable SSH in browser	disables the SSH-in-browser tool in the Cloud Console. When enforced, the SSH-in-browser button is disabled. constraints/compute.disableSshInBrowser	"is:"
Require OS Login	enables OS Login on all newly created Projects. All VM instances created in new projects will have OS Login enabled. constraints/compute.requireOsLogin	"is:"
Shielded VMs	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. constraints/compute.requireShieldedVm	"is:"
Restrict VPC peering usage	list constraint defines the set of VPC networks that are allowed to be peered with the VPC networks belonging to this project, folder, or organization. constraints/compute.restrictVpcPeering	"is:", "under:"

Compute Engine: (most important) Organization Policy Constraints

Constraint	Description	Supported Prefixes
Skip default network creation	Skips the creation of the default network and related resources during Project resource creation if set to True. constraints/compute.skipDefaultNetworkCreation	"is:"
Define trusted image projects	defines the set of projects that can be used for image storage and disk instantiation for Compute Engine. constraints/compute.trustedImageProjects	"is:"
Define allowed external IPs for VM instances	Defines the set of Compute Engine VM instances that are allowed to use external IP addresses. constraints/compute.vmExternalIpAccess	"is:"

VM Pricing and cost optimization

Sustained Use Discounts (SUD)

Up to 30% savings on Compute Engine and Cloud SQL

Committed Use Discounts (CUD)

Up to 70% savings without upfront fees or instance-type lock-in

Spot / Preemptible VM instances

Up to 91% savings on workloads that can be interrupted, like data mining and data processing

Per second billing

Up to 38% savings by paying per second, not per hour

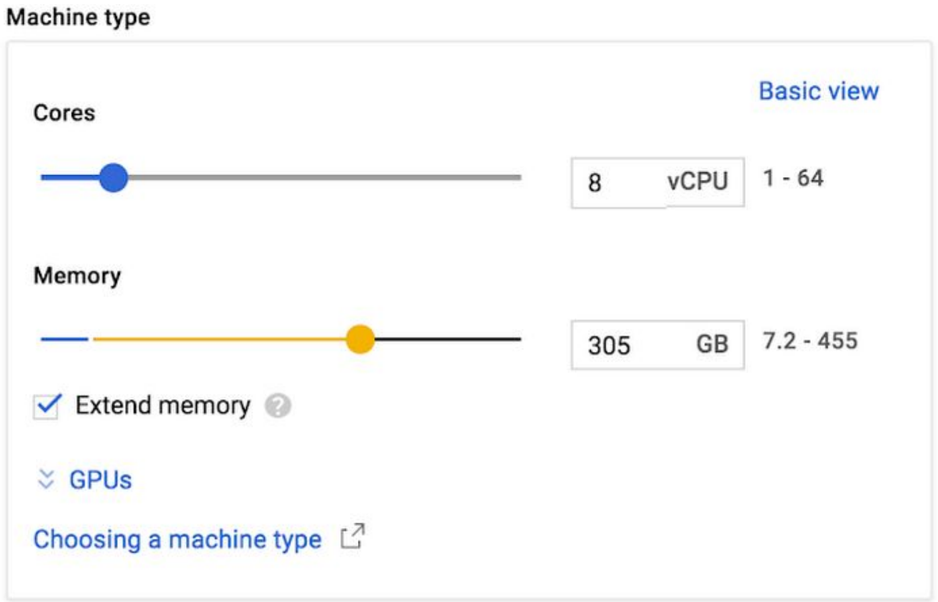
Network Service Tiers

Pick performance and get 70% more bandwidth than other clouds, or pick cost savings and save up to 9% compared to other clouds

Rightsizing (incl. choosing optimal GCE families) and Custom Machine Types

Exam Tips:

- Common pattern for optimization costs for unused PDs: you can create a snapshot, and delete the disk to reduce the maintenance cost of that disk by 35% to 92%.
- For premium OS, you're billed for license per vCPU per second.
- Bring Your Own License is an option for some OSes
- Use Extended memory to save on OS license costs.



Migrate for Compute Engine

Lift&Shift your VMWare, AWS, Azure workloads to GCE

- Purpose-built, enterprise-grade
- Migrate from on-prem or other clouds
- Proven at scale, having migrated customers w/ thousands of workloads
- Success across healthcare, energy, government, manufacturing, and more

Agentless

Nothing to install on source machines

Minimize complexity, reduce IT labor requirements by 5+ hours per server, keep migrations on track.

Streaming

Migrate storage while apps run in GCP

Eliminate long upfront data transfers and unpredictable maintenance windows, enabling fast time-to-cloud and reduced downtime.

Frictionless

Automate migration and in-cloud conversion

Reduce touch points for IT, provide uninterrupted experience for line of business owners and end users.

Persistent Disks best practices, tips & tricks

Exam Tips:

- Use "**--no-boot-disk-auto-delete**" parameter if you don't want boot / OS disk to be deleted if a VM gets deleted.
- CMEK and CSEK can be used to encrypt PDs. Have a look at how to use CSEK for a PD.
- Avoid using ext3 filesystems in Linux (poor performance under heavy write loads). Prefer ext4.
- You can share a PD across multiple VMs at the same time in **read-only mode**.
    - If read-write required, prefer a managed solution such as Filestore or utilize GCS.
- You can share a SSD PD across two N2 VMs at the same time in read-write mode. In Preview as of Q1 '23 -> should NOT be covered on the exam.
- To recover from a corrupted disk / Os not booting properly, follow this procedure.
    - On high-level, just attach the corrupted disk as non-boot disk to another VM and troubleshoot.
- For special use-cases (app needs a RAM disk with exceptionally low latency and high throughput and cannot just use the VM memory), you can create a tmpfs filesystem by allocating some VM memory as a RAM disk.
- If needed, you can attach 1-24 **local SSDs** (ephemeral = data is lost if VM stops; each 375 GB and physically attached to the server that hosts your VM instance). **Local SSDs are NOT the same as PD SSDs**!!! More information here.

Compute Engine

#GCPSketchnotes

@PVERGADIA THECLOUDGIRL.DEV

What is Compute Engine?

CUSTOMIZABLE VIRTUAL MACHINES IN GOOGLE CLOUD

THERE ARE 2 MACHINE TYPES

PREDEFINED
n1-standard-1 =
(1 vCPU, 3.75 GB memory)

CUSTOM
vCPU + Memory

CORES — 1 vCPU — 1-96

MEMORY — 3.75 GB — 1-6.5

THERE ARE 3 MACHINE TYPE FAMILIES

GENERAL PURPOSE Machine Type
- General Servers
- Websites
- Databases

COMPUTE OPTIMIZED Machine Type
- High Performance Computing
- Gaming
- Electronic Design Automation
- Single Threaded Applications

MEMORY OPTIMIZED Machine Type
- Large In-Memory Databases (SAP HANNA)
- In-Memory Analytics

Compute Engine Use case (example)

- FIG. 1 — Websites
- FIG. 2 — Databases
- FIG. 3 — Legacy Monolithic Apps
- FIG. 4 — Containers
- FIG. 5 — Windows Apps

Compute Engine PRICING

SUSTAINED USE SAVINGS
Automatic discounts for running VMs a significant portion of the month

COMMITTED USE DISCOUNT
Up to 57% savings with no up-front cost

PREEMPTIBLE VMs
Up to 80% savings and run batch jobs & fault-tolerant workloads

RIGHT SIZE RECOMMENDATIONS
Suggests resizing for efficiency and cost

How does it WORK??

CREATE

region
+ zone
+ machine type (cpu & memory)
= Instance

BACKUPS

- Automatic Scheduled snapshots
- Live migration — Keep apps running during maintenance

HIGH AVAILABILITY

- Automatic failover to another zone or region (ZONE A, AUTOMATIC FAILOVERS, ZONE B)
- MANAGED INSTANCE GROUPS (MIGs)
  - Autohealing
  - Load balancing
  - Automatic updates
  - Support for stateful workloads
  - Container support
- AUTOSCALER - 3 types of policies:
  1. CPU utilization = more than 60% → create new instance
  2. HTTP(S) load balancers service capacity — Requests per second or utilization
  3. Cloud monitoring metrics

Persistent Disk

#GCPSketchnotes

@PVERGADIA THECLOUDGIRL.DEV

3.29.2020

What is Persistent Disk?

HIGH PERFORMANCE BLOCK STORAGE

- ✓ Transparently Resize
- ✓ Quick backup

Persistent Disk

SSD/HDD storage

Compute Engine

Kubernetes Engine

- ✓ Supports Simultaneous readers

Which Block storage is right for your app?

LOWER COST ← → PERFORMANCE

Standard PD	Balanced PD	Performance (PD-SSD)	Extreme PD	Local SSD
HDD	SSD	SSD	SSD	Local SSD
Standard throughput	Best price per GB	Best price per IOPS	Best performance	Lowest Latency
Current 'cost sensitive'	Best fit for most workloads	Performance sensitive workloads	Uncompromising performance	For caches and hot layers

(How to pick) based on availability needs.

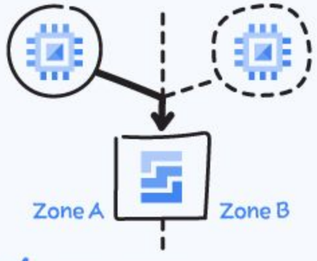
Local SSD

- ✓ Ephemeral
- ✓ Stateless workloads, or replication managed at app or database layer

Persistent Disk

- ✓ Durable, support snapshots
- ✓ Most workloads

Regional Persistent Disk

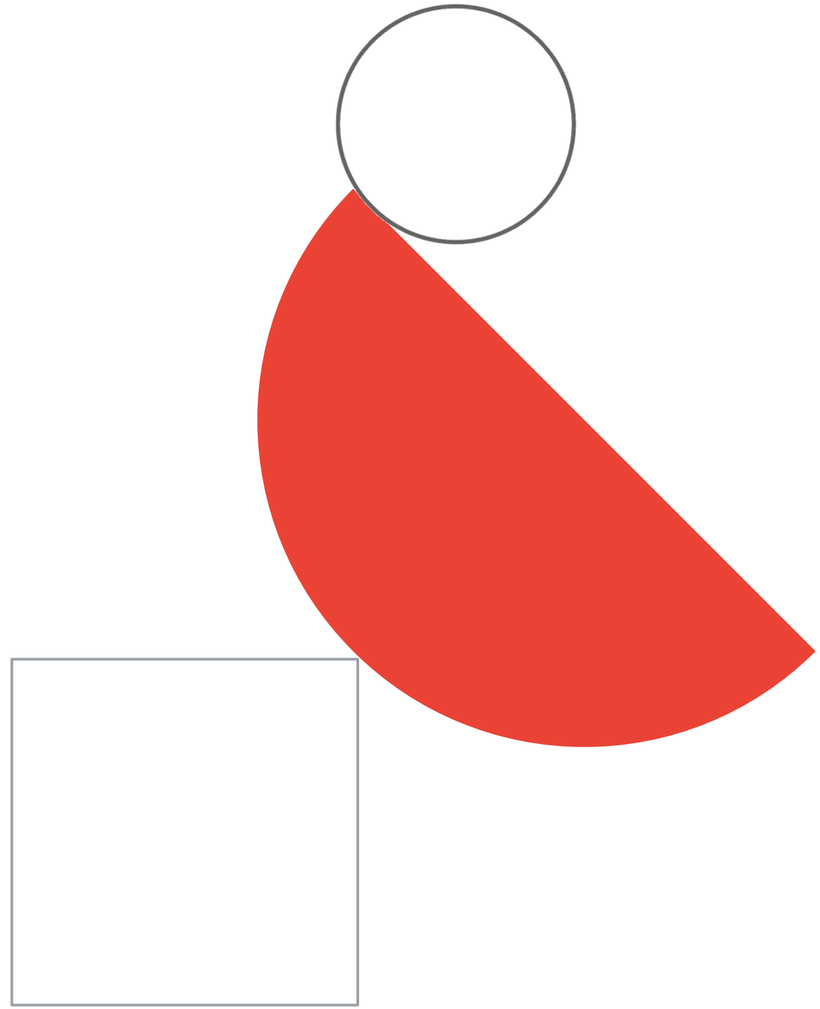


- ✓ Durable & Highly Available
- ✓ Mission-critical workloads with RPO/RTO near 0

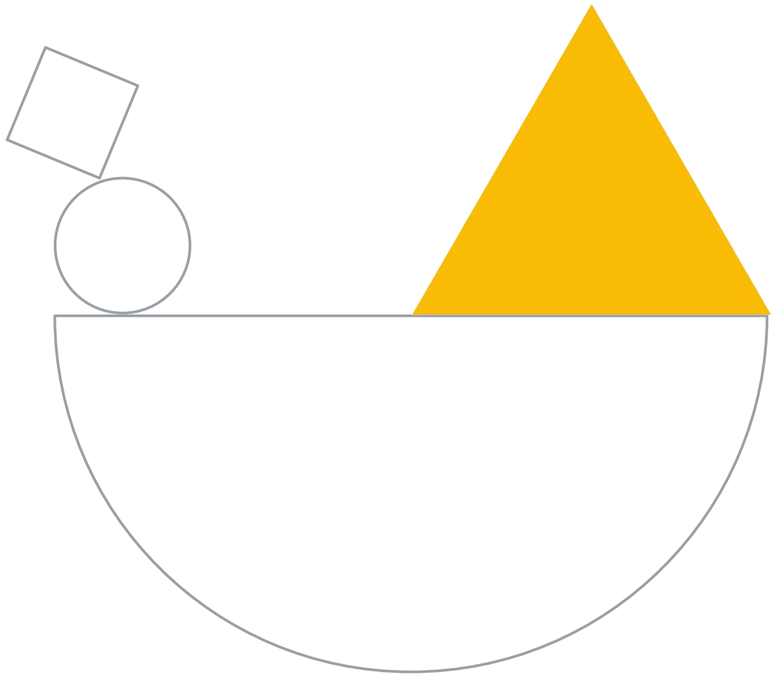
Persistent Disk «Use case example»

Standard PD	Balanced PD	Performance (PD-SSD)	Extreme PD	Local SSD
Cost sensitive workloads	Most enterprise apps	Most databases	SAP HANA	Scale out analytics
Scale out analytics (Hadoop, Kafka)	LOB apps	Persistent cache	Oracle	Media rendering
	Boot disks	Scale-out analytics	Largest in-memory DBs	Other use cases where ephemeral scratch space is required
	Webserving			

QUIZ time!



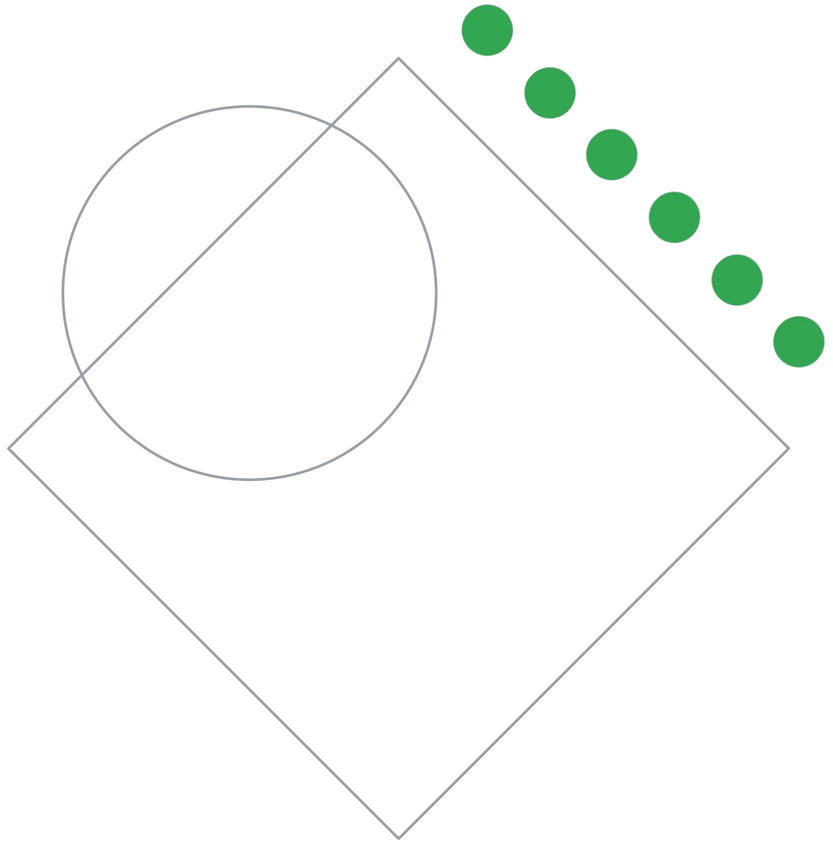
Diagnostic questions



Please complete the diagnostic questions now

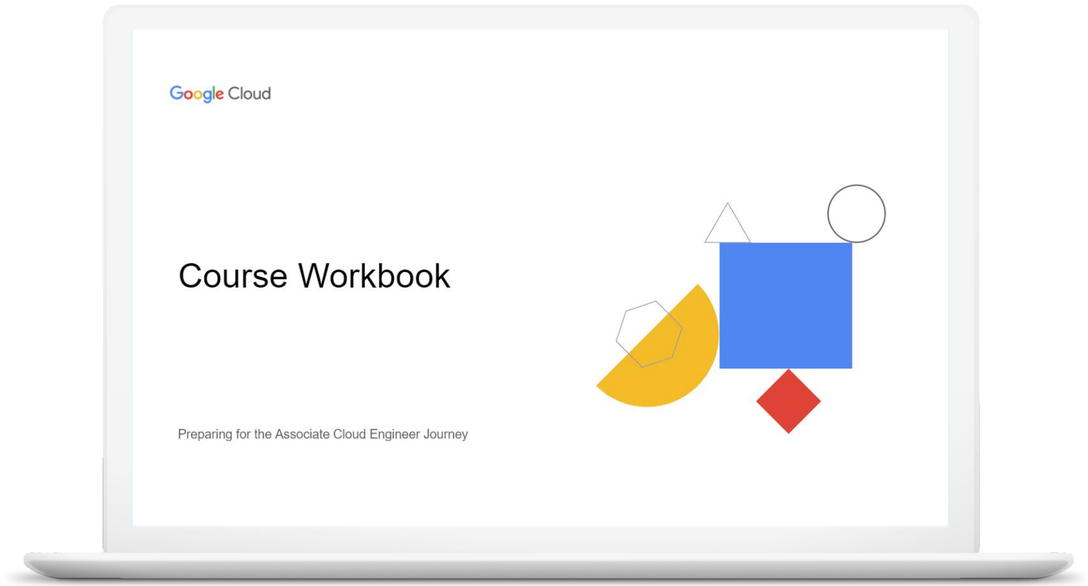
- Forms are provided for you to answer the diagnostic questions
- The instructor will provide you a link to the forms
- The diagnostic questions are also available in the workbook

Review and study planning



Your study plan:

Planning and configuring cloud solutions



2.1 Planning and estimating using the Pricing Calculator

2.2 Planning and configuring compute resources

2.3 Planning and configuring data storage options

2.4 Planning and configuring network resources

2.1 Planning and estimating using the Pricing Calculator

2.1 Diagnostic Question 01 Discussion

The projected amount of cloud storage required for Cymbal Superstore to enable users to post pictures for project reviews is 10 TB of immediate access storage in the US and 30 TB of storage for historical posts in a bucket located near Cymbal Superstore's headquarters. The contents of this bucket will need to be accessed once every 30 days. You want to estimate the cost of these storage resources to ensure this is economically feasible.

What should you do?

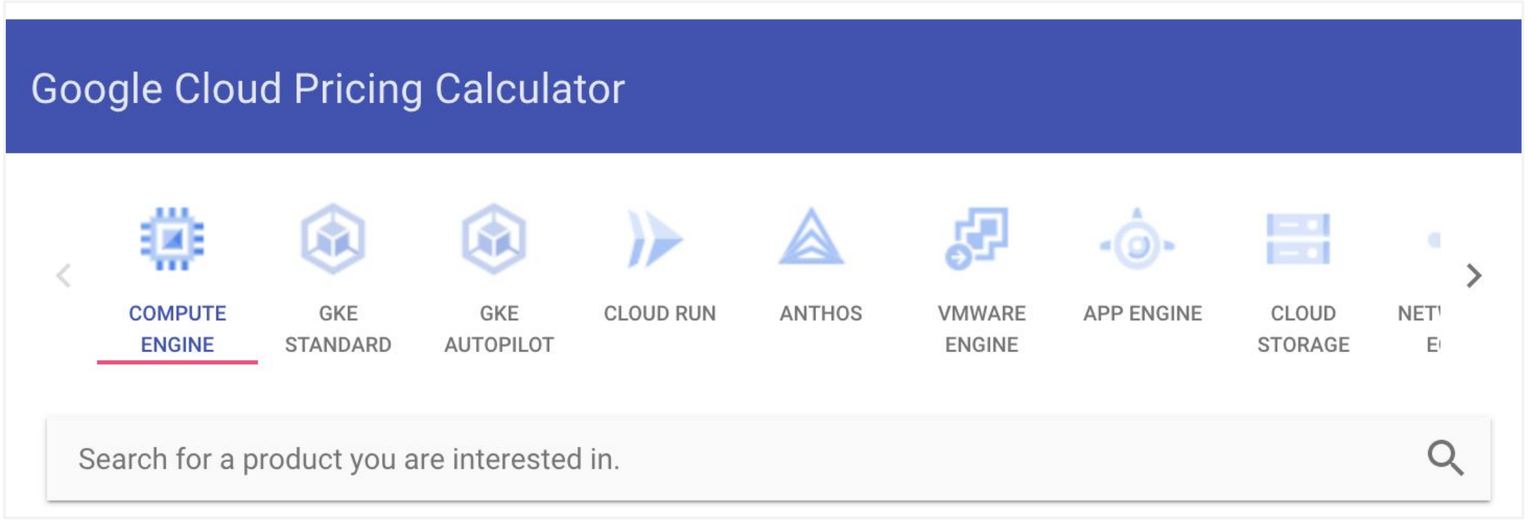
A. Use the pricing calculator to estimate the costs for 10 TB of regional standard storage, 30 TB of regional Coldline storage, and egress charges for reads from storage.

B. Use the pricing calculator to estimate the price for 10 TB of regional standard storage, 30 TB of regional Nearline storage, and ingress charges for posts to the bucket.

C. Use the pricing calculator to estimate the price for 10 TB of multi-region standard storage, 30 TB for regional Coldline storage, and ingress charges for posts to the bucket.

D. Use the pricing calculator to estimate the price for 10 TB of multi-region standard storage, 30 TB for regional Nearline, and egress charges for reads from the bucket.

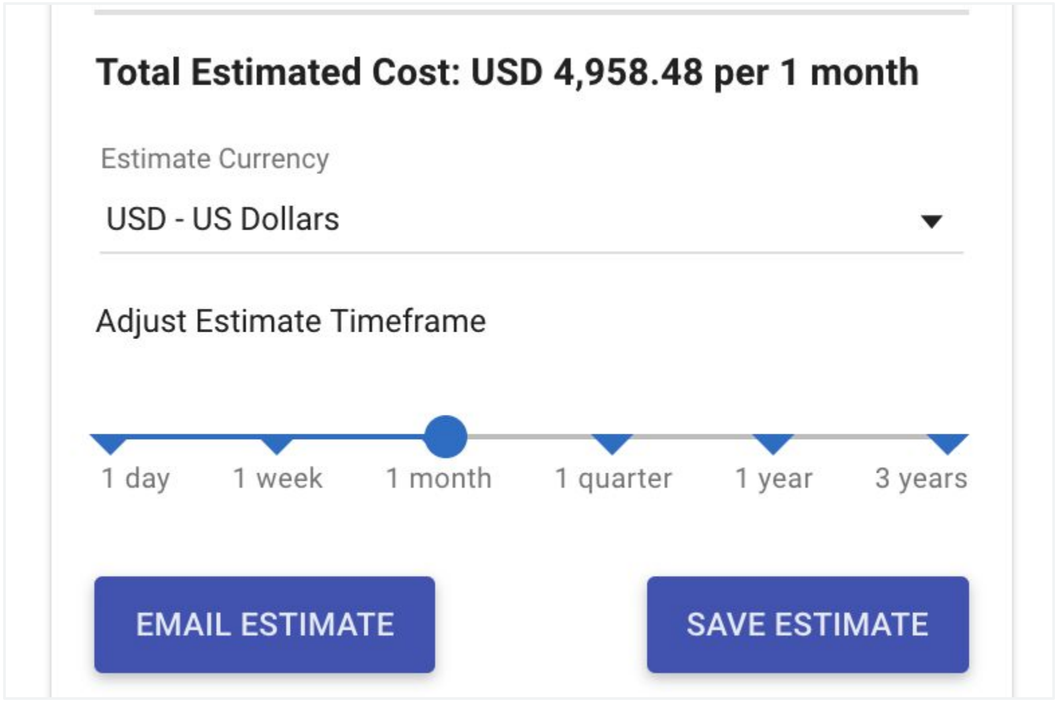
Pricing Calculator



https://cloud.google.com/products/calculator/

1. Select a product from scrolling list at top of the form
2. Form for each product will show expense variables
3. Enter target configuration in form
4. Submit each section to add to your overall estimate

Pricing Calculator



- Total estimated cost: daily, weekly, monthly, quarterly, yearly and 3-year increments
- Cost is only an estimate
- Question to ask: How closely your estimated usage matches your actual usage?
- Not a binding contract, just a planning tool

2.1 Planning and estimating using the Pricing Calculator

Courses

[Architecting with Google Compute Engine](#)

- M3 Virtual Machines
- M6 Resource Management

=

[Essential Google Cloud Infrastructure: Foundation](#)

- M3 Virtual Machines

[Essential Google Cloud Infrastructure: Core Services](#)

- M3 Resource Management

Documentation

[Google Cloud Pricing Calculator](#)

2.2 Planning and configuring compute resources

Considerations include:

- Selecting appropriate compute choices for a given workload

  (e.g., Compute Engine, Google Kubernetes Engine, Cloud Run, Cloud Functions)
- Using preemptible VMs and custom machine types as appropriate

2.2 Diagnostic Question 02 Discussion

Cymbal Superstore decides to migrate their supply chain application to Google Cloud. You need to configure specific operating system dependencies.

What should you do?

A. Implement an application using containers on Cloud Run.

B. Implement an application using code on App Engine.

C. Implement an application using containers on Google Kubernetes Engine.

D. Implement an application using virtual machines on Compute Engine.

Infrastructure as a service:

	Google Compute Engine	Google Kubernetes Engine
	Virtual machines running in Google's global data centers	Logical infrastructure powered by Kubernetes, the open source container orchestration system
Use When You Need...	• Complete control • Ability to make OS level changes • To be able to move to the cloud without rewriting your code • To use custom VM images	• No dependencies on a specific OS • Increased velocity and operability • To manage containers in production
Typical Use Cases	• Any workload requiring a specific OS or configuration • On-premises software that you want to run in the cloud	• Containerized workloads • Cloud-native distributed systems • Hybrid applications

Platform as a service:

	Google App Engine	Google Cloud Run	Google Cloud Functions
	Flexible, zero-ops platform for building apps	Deploy code or containers that listens for requests or events	Serverless execution environment for building and connecting cloud services
Use When You Need...	• To just focus on writing code • Developer velocity • To minimize operational overhead	• Scales to meet demand • Pay for what you use • Supports API endpoints	• For event-driven workloads • Scales to meet demand • Minimal configuration
Typical Use Cases	• Web sites • Apps (of course!) • Gaming back ends • IoT applications	• Web frameworks • Microservices	• Statistical analysis • Image thumbnail generation • Post a comment to a Slack channel after a GitHub commit

2.2 Diagnostic Question 03 Discussion

Cymbal Superstore decides to pilot a cloud application for their point of sale system in their flagship store. You want to focus on code and develop your solution quickly, and you want your code to be portable.

How do you proceed?

A. SSH into a Compute Engine VM and execute your code.

B. Package your code to a container image and post it to Cloud Run.

C. Implement a deployment manifest and run kubectl apply on it in Google Kubernetes Engine.

D. Code your solution in Cloud Functions.

2.2 Diagnostic Question 04 Discussion

An application running on a highly-customized version of Ubuntu needs to be migrated to Google Cloud. You need to do this in the least amount of time with minimal code changes.

How should you proceed?

A. Create Compute Engine Virtual Machines and migrate the app to that infrastructure.

B. Deploy the existing application to App Engine.

C. Deploy your application in a container image to Cloud Run.

D. Implement a Kubernetes cluster and create pods to enable your app.

2.2 Diagnostic Question 05 Discussion

You want to deploy a microservices application. You need full control of how you manage containers, reliability, and autoscaling, but don't want or need to manage the control plane.

Which compute option should you use?

A. Cloud Run
B. App Engine
C. Google Kubernetes Engine
D. Compute Engine

2.2 Planning and configuring compute resources

Courses

[Google Cloud Fundamentals: Core Infrastructure](#)

- M3 Virtual Machines in the Cloud
- M5 Containers in the Cloud
- M6 Applications in the Cloud

[Getting Started with Google Kubernetes Engine](#)

- M2 Introduction to Containers and Kubernetes

[Architecting with Google Compute Engine](#)

- M3 Virtual Machines

=

[Essential Google Cloud Infrastructure: Foundation](#)

- M3 Virtual Machines

Skill Badges

Google Cloud

[Set Up and Configure a Cloud Environment in Google Cloud Quest](#)

Documentation

[Choosing the right compute option in GCP: a decision tree](#)

[Application Hosting Options](#)

[Tutorials | Compute Engine Documentation](#)

2.3 Planning and configuring data storage options

Considerations include:

- Product choice

  (e.g., Cloud SQL, BigQuery, Firestore, Cloud Spanner, Cloud Bigtable)
- Choosing storage options

  (e.g., Zonal persistent disk, Regional balanced persistent disk, standard, Nearline, Coldline, Archive)

2.3 Diagnostic Question 06 Discussion

Cymbal Superstore needs to analyze whether they met quarterly sales projections. Analysts assigned to run this query are familiar with SQL.

A. BigQuery
B. Cloud SQL
C. Cloud Spanner
D. Cloud Firestore

What data solution should they implement?

Comparing Data Storage and Database Options

Relational		Non-relational		Object	Warehouse
 Cloud SQL	 Cloud Spanner	 Cloud Datastore	 Cloud Bigtable	 Cloud Storage	 BigQuery
Good for: Web frameworks	**Good for:** RDBMS+scale, HA, HTAP	**Good for:** Hierarchical, mobile, web	**Good for:** Heavy read + write, events	**Good for:** Binary or object data	**Good for:** Enterprise data warehouse
Such as: CMS, eCommerce	**Such as:** User metadata, Ad/Fin/MarTech	**Such as:** User profiles, Game State	**Such as:** AdTech, financial, IoT	**Such as:** Images, media serving, backups	**Such as:** Analytics, dashboards

2.3 Diagnostic Question 07 Discussion

Cymbal Superstore's supply chain application frequently analyzes large amounts of data to inform business processes and operational dashboards.

A. Multi-regional
B. Regional
C. Nearline
D. Coldline

What storage class would make sense for this use case?

Storage Classes and use cases summary

	Regional Lower cost	Multi-Regional Lower cost, geo-redundant	Nearline Very low storage cost, has data retrieval costs	Coldline Lowest storage cost of all, takes longer to retrieve, costs to retrieve data
Use When You Need...	• Lower cost per GB stored • Data stored in a narrow geographic region • Redundant across zones	• Redundant across regions	• Very low cost per GB stored • Higher per-operation costs • 30-day minimum storage duration	• Lowest cost per GB stored • Higher per-operation costs • 90-day minimum storage duration
Typical Use Cases	Storing frequently accessed data in the same region as your instances that use it, such as for data analytics.	Storing data that is frequently accessed around the world, such as website content, streaming videos or gaming content	Infrequently (i.e., no more than once per month) accessed data. Ideal for back-up and serving long-tail multimedia content.	Very infrequently accessed data - ie, once a year.Typically this is for disaster recovery, or for financial data that has to be kept for a certain length of time to meet regulatory needs.

2.3 Diagnostic Question 08 Discussion

Cymbal Superstore has a need to populate visual dashboards with historical time-based data. This is an analytical use-case.

Which two storage solutions could they use?

A. BigQuery
B. Cloud Storage
C. Cloud Firestore
D. Cloud SQL
E. Cloud Bigtable

Comparing storage options: use cases

	Firestore	Cloud Bigtable	Cloud Storage	Cloud SQL	Cloud Spanner	BigQuery
Type	NoSQL document	NoSQL wide column	Blobstore	Relational SQL for OLTP	Relational SQL for OLTP	Relational SQL for OLAP
Best for	Storing, syncing, and querying data	"Flat" data, Heavy read/write, events, analytical data	Structured and unstructured binary or object data	Web frameworks, existing applications	Large-scale database applications (> ~2 TB)	Interactive querying, offline analytics
Use cases	Mobile, web, and server development	AdTech, Financial and IoT data	Images, large media files, backups	User credentials, customer orders	Whenever high I/O, global consistency is needed	Data warehousing

2.3 Planning and configuring data storage options

Courses

[Google Cloud Fundamentals: Core Infrastructure](#)

- M4 Storage in the Cloud

[Architecting with Google Compute Engine](#)

- M5 Storage and Database Services

=

[Essential Google Cloud Infrastructure: Core Services](#)

- M2 Storage and Database Services

Skill Badges

Google Cloud

[Perform Foundational Infrastructure Tasks in Google Cloud Quest](#)

Documentation

[Cloud Storage Options](#)

[Storage classes](#)

[Data lifecycle | Cloud Architecture Center](#)

2.4 Planning and configuring network resources

Considerations include:

- Differentiating load balancing options
- Identifying resource locations in a network for availability
- Configuring Cloud DNS

2.4 Diagnostic Question 09 Discussion

Cymbal Superstore is piloting an update to its ecommerce app for the flagship store in Minneapolis, Minnesota. The app is implemented as a three-tier web service with traffic originating from the local area and resources dedicated for it in us-central1. You need to configure a secure, low-cost network load-balancing architecture for it.

How do you proceed?

A. Implement a premium tier pass-through external https load balancer connected to the web tier as the frontend and a regional internal load balancer between the web tier and backend.

B. Implement a proxied external TCP/UDP network load balancer connected to the web tier as the frontend and a premium network tier ssl load balancer between the web tier and the backend.

C. Configure a standard tier proxied external https load balancer connected to the web tier as a frontend and a regional internal load balancer between the web tier and the backend.

D. Configure a proxied SSL load balancer connected to the web tier as the frontend and a standard tier internal TCP/UDP load balancer between the web tier and the backend.

Google VPC offers a suite of load balancing options

Global HTTP(S)	Global SSL Proxy	Global TCP Proxy	Regional	Regional internal
Layer 7 load balancing based on load	Layer 4 load balancing of non-HTTPS SSL traffic based on lo ad	Layer 4 load balancing of non-SSL TCP traffic	Load balancing of any traffic (TCP, UDP)	Load balancing of traffic inside a VPC
Can route different URLs to different backends	Supported on specific port numbers	Supported on specific port numbers	Supported on any port number	Use for the internal tiers of multi-tier applications

2.4 Diagnostic Question 10 Discussion

What Google Cloud load balancing option runs at Layer 7 of the TCP stack?

A. Global http(s)
B. Global SSL Proxy
C. Global TCP Proxy
D. Regional Network

Summary of load balancers

Load balancer	Traffic type	Global/ Regional	External/ Internal	External ports for load balancing
HTTP(S)	HTTP or HTTPS	Global IPv4 IPv6	External	HTTP on 80 or 8080; HTTPS on 443
SSL Proxy	TCP with SSL offload			25, 43, 110, 143, 195, 443, 465, 587, 700, 993, 995, 1883, 5222
TCP Proxy	• TCP without SSL offload • Does not preserve client IP addresses			25, 43, 110, 143, 195, 443, 465, 587, 700, 993, 995, 1883, 5222
Network TCP/UDP	• TCP/UDP without SSL offload • Preserves client IP addresses	Regional IPv4		Any
Internal TCP/UDP	TCP or UDP		Internal	Any
Internal HTTP(S)	HTTP or HTTPS			HTTP on 80 or 8080; HTTPS on 443

2.4 Planning and configuring network resources

Courses

Google Cloud Fundamentals: Core Infrastructure

- M3 Virtual Machines in the Cloud

Architecting with Google Compute Engine

- M2 Virtual Networks
- M9 Load Balancing and Autoscaling

=

Essential Google Cloud Infrastructure: Foundation

- M2 Virtual Network

Elastic Google Cloud Infrastructure: Scaling and Automation

- M2 Load Balancing and Autoscaling

Documentation

Cloud Load Balancing overview

Cloud Load Balancing

Knowledge Check 1

Which storage class is designed for long term storage has a 365 day minimum storage agreement, and a lower storage price as compared to other storage types?

A. Standard Storage

B. Cold Line Storage

C. Nearline Storage

D. Archive storage

Knowledge Check 1

Which storage class is designed for long term storage has a 365 day minimum storage agreement, and a lower storage price as compared to other storage types?

A. Standard Storage

B. Cold Line Storage

C. Nearline Storage

D. Archive storage

Knowledge Check 2

Which serverless option is based on developing and executing small snippets of code?

A. Cloud Functions

B. Cloud Run

C. BigQuery

D. Dataflow

Knowledge Check 2

Which serverless option is based on developing and executing small snippets of code?

A. Cloud Functions

B. Cloud Run

C. BigQuery

D. Dataflow