Google Cloud Platform

An **IN ACTION** introduction

https://cloud.google.com

Carmine Spagnuolo, PhD

What is “cloud”?

The term “cloud” has been used in many different contexts and it has many different definitions, so it makes sense to define the term:

Cloud is a collection of services that helps developers focus on their project rather than on the infrastructure that powers it.

In more concrete terms, cloud services are things like Amazon Elastic Compute Cloud (EC2) or Google Compute Engine (GCE), which provide APIs to provision virtual servers (or services as storage and other computing tools), where customers pay per hour for the use of these servers.

What is Google Cloud Platform?

- There are many cloud providers out there, including Google, Amazon, Microsoft, Rackspace, DigitalOcean, and more.
- Each provides many similar products, the implementation and details of how these products work tends to vary quite a bit.
- Google Cloud Platform (often abbreviated as GCP) is a collection of products that allows the world to use some of Google's internal infrastructure.

What is Google Cloud Platform?

- This collection includes many things that are common across all cloud providers, such as on-demand virtual machines via Google Compute Engine or object storage for storing files via Google Cloud Storage.
- It also includes APIs to some of the more advanced Google- built technology, like Bigtable, Cloud Datastore, or Kubernetes.

Getting started with GCP

Signing up for GCP

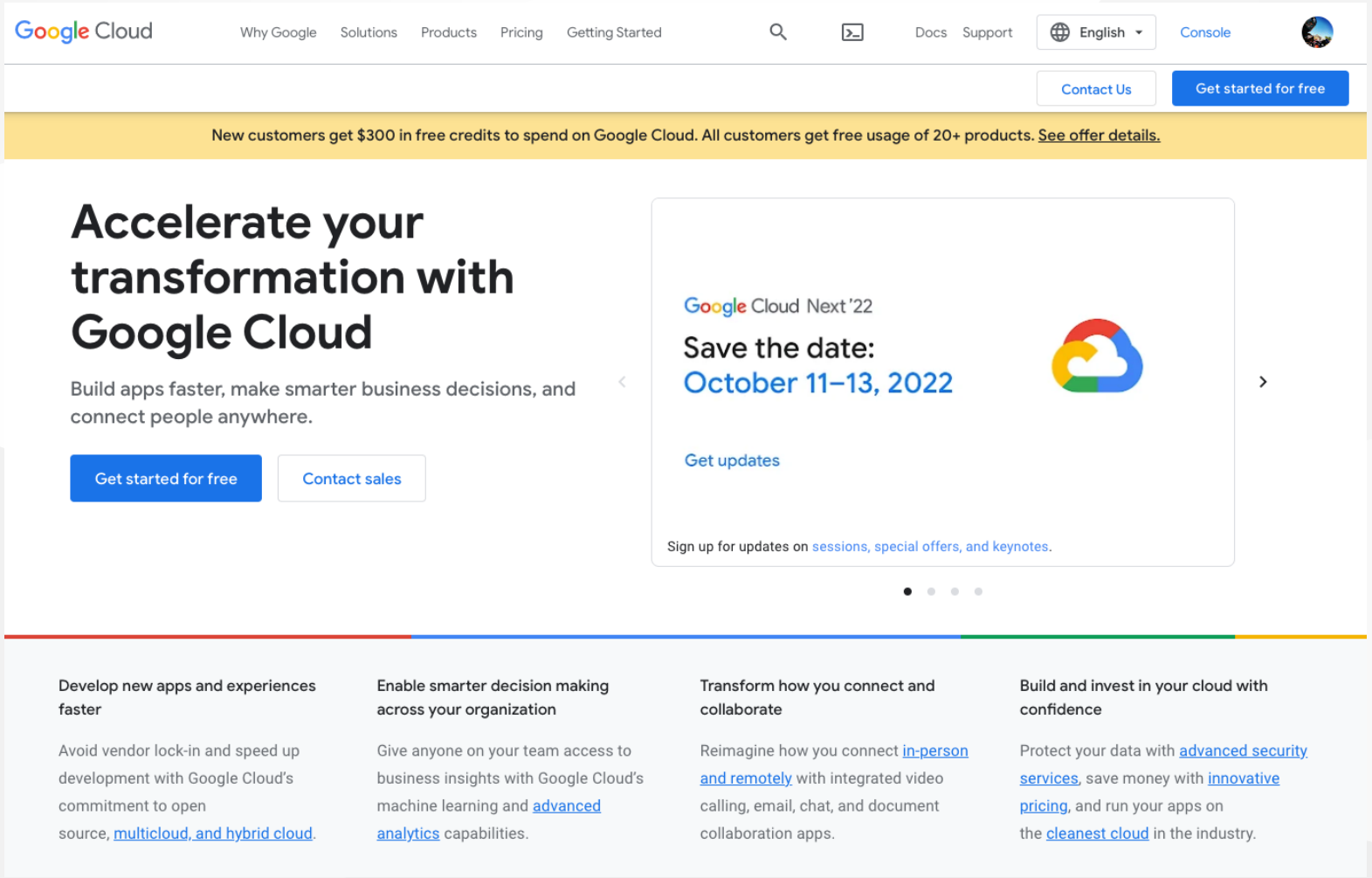
Before you can start using any of Google's Cloud services, you first need to sign up for an account.

1. **Instructions for Student Coupon Retrieval** ($50)
    - MS Teams - **Channel GCP - Account**
    - Students can request coupons from the URL and redeem them until: **1/7/2022**
    - Coupons Valid Through: **1/3/2023**

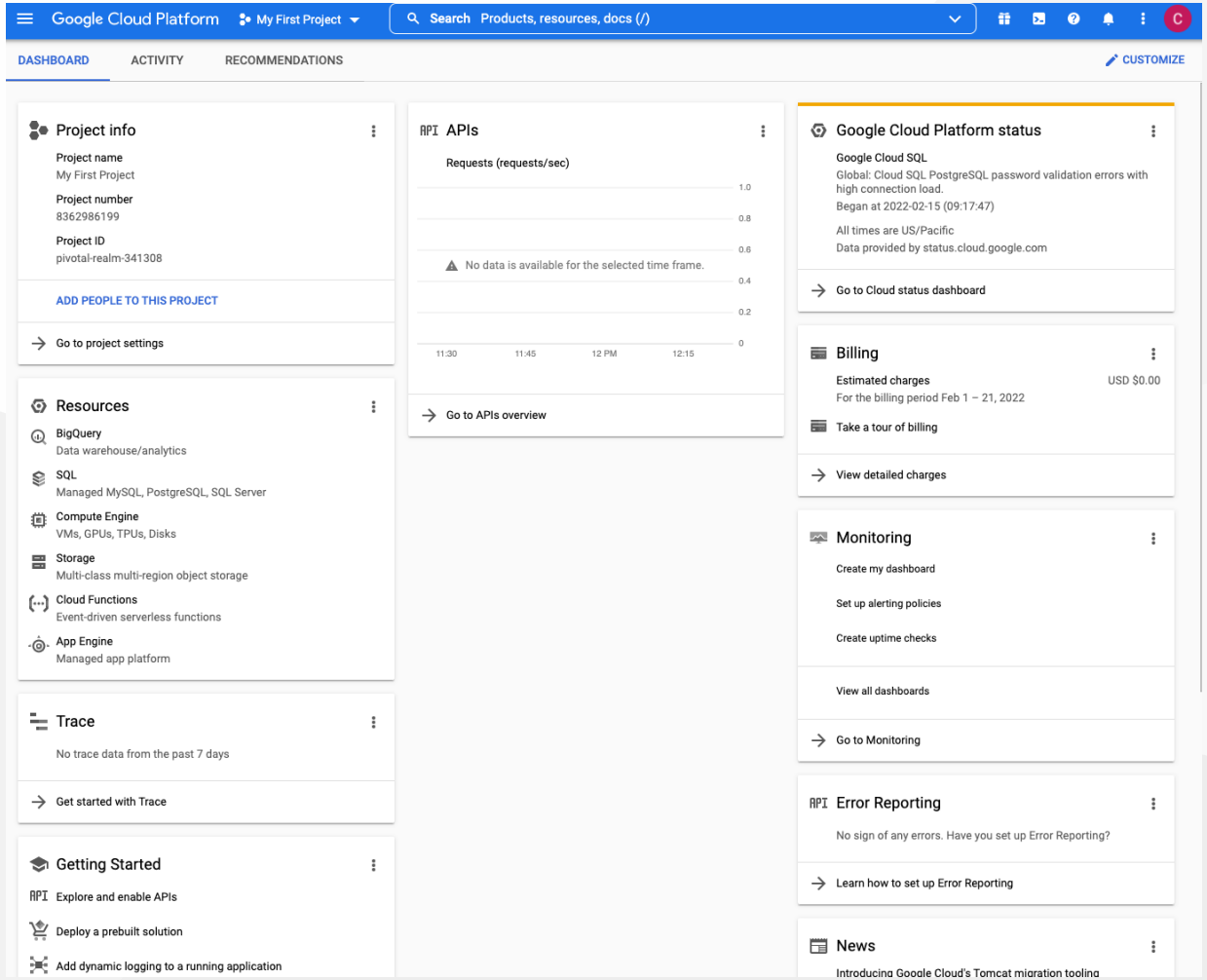
Getting started with GCP

2. Google is generous enough to give $300 worth of google cloud credit to the new google/gmail users free-trial 💳 required! 🛑

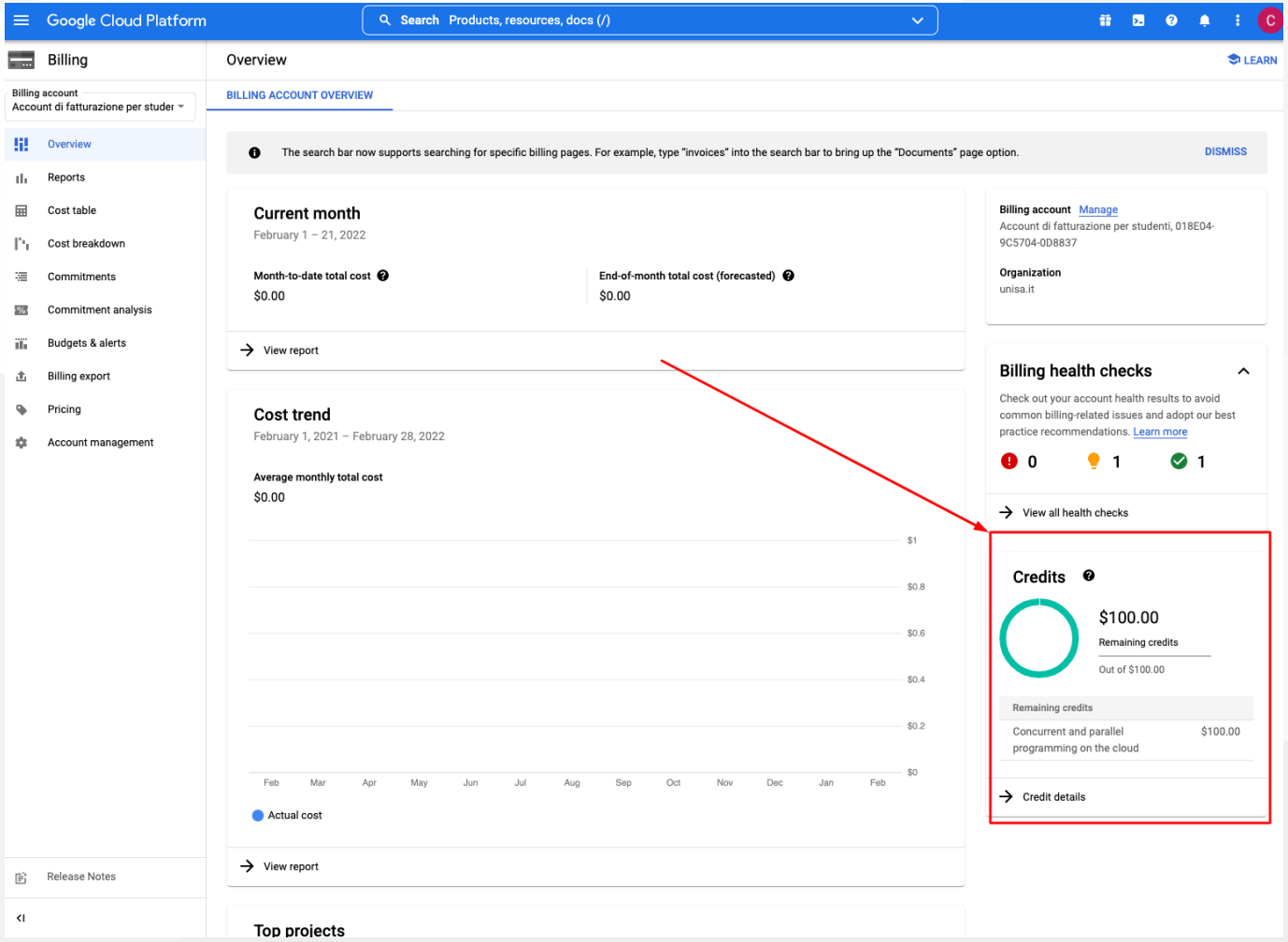
https://cloud.google.com



Exploring the console



Check "Billing"



Understanding projects

- When we first signed up for Google Cloud Platform, we learned that a new project is created automatically, and that projects have something to do with isolation, but what does this mean? And what are projects anyway?
- **Projects** are primarily a container for all the resources we create.
- For example, if we create a new VM, it will be “owned” by the parent project.
- Projects also act as a way of **isolating** things from one another, sort of like having a workspace for a specific purpose.

Understanding projects

- If you create new service account credentials (which we'll do later) inside one project, say `project-a`, those credentials have access to resources only inside `project-a` unless you explicitly grant more access.
- When running commands, those commands can access anything that you have access to inside the Cloud Console, which includes all of the projects you've created.

Using GCP

1. Cloud Console in the browser.
2. gcloud SDK
    - Integrate with APIs using Client Libraries for Java, Python, Node.js, Ruby, Go, .NET, and PHP.
    - Script or interact with cloud resources at scale using the Google Cloud CLI.
    - Accelerate local development with emulators for Pub/Sub, Spanner, Bigtable, and Datastore.

Using GCP

3. Google Cloud CLI lets you manage resources and services from the command line. It also contains service and data emulators to speed up local development.
4. Cloud Shell lets you code or use a terminal directly in the web-browser.
5. Cloud Code provides IDE extensions for VSCode and IntelliJ.

🚀 Please select your preferred method and install it for the next!

❗ Cloud Shell is a ready-to-use starting point for your learning.

Interacting with GCP

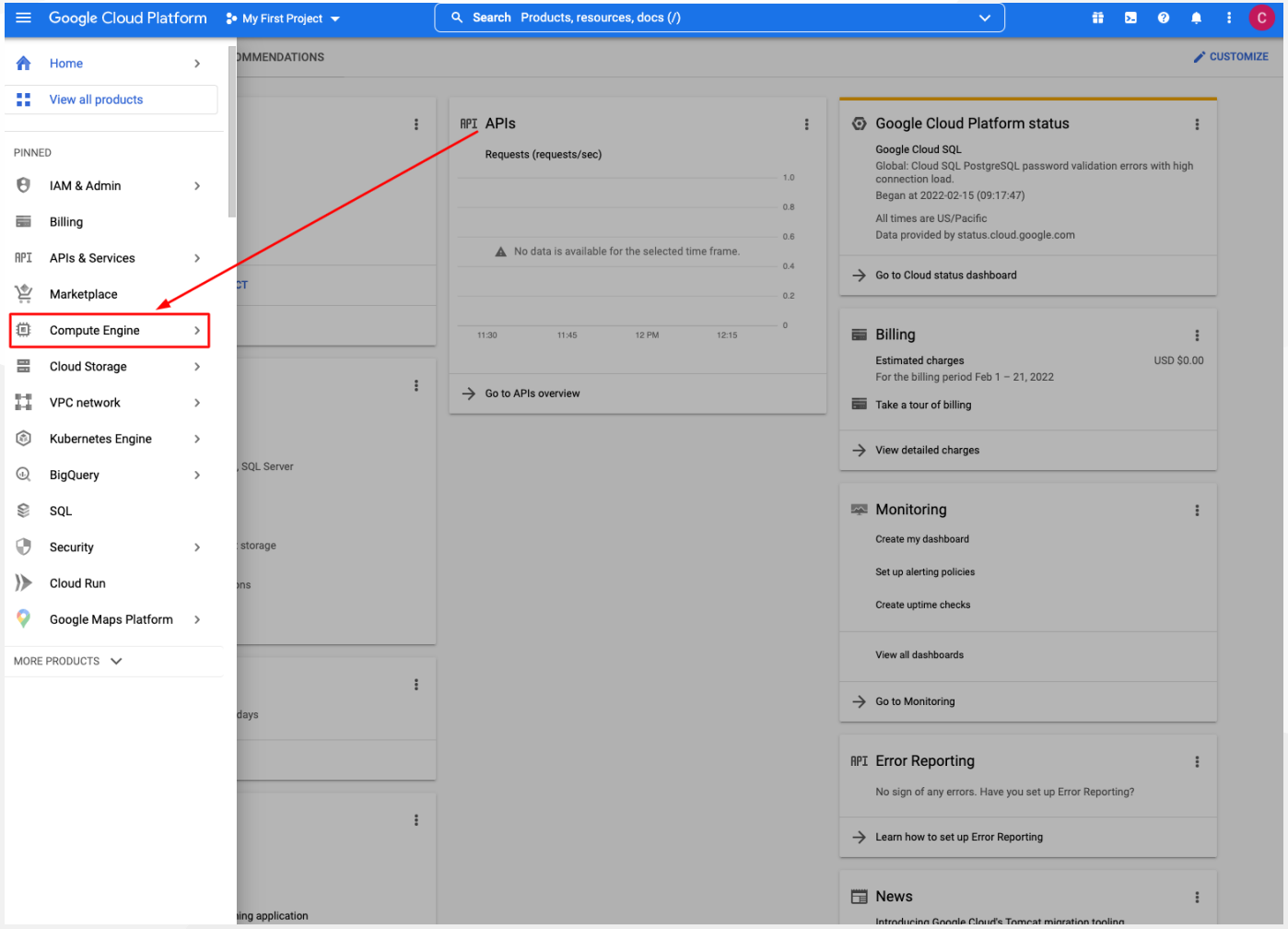
Let’s start:

- by launching a virtual machine in the cloud
- and then writing a script to terminate the virtual machine in JavaScript.

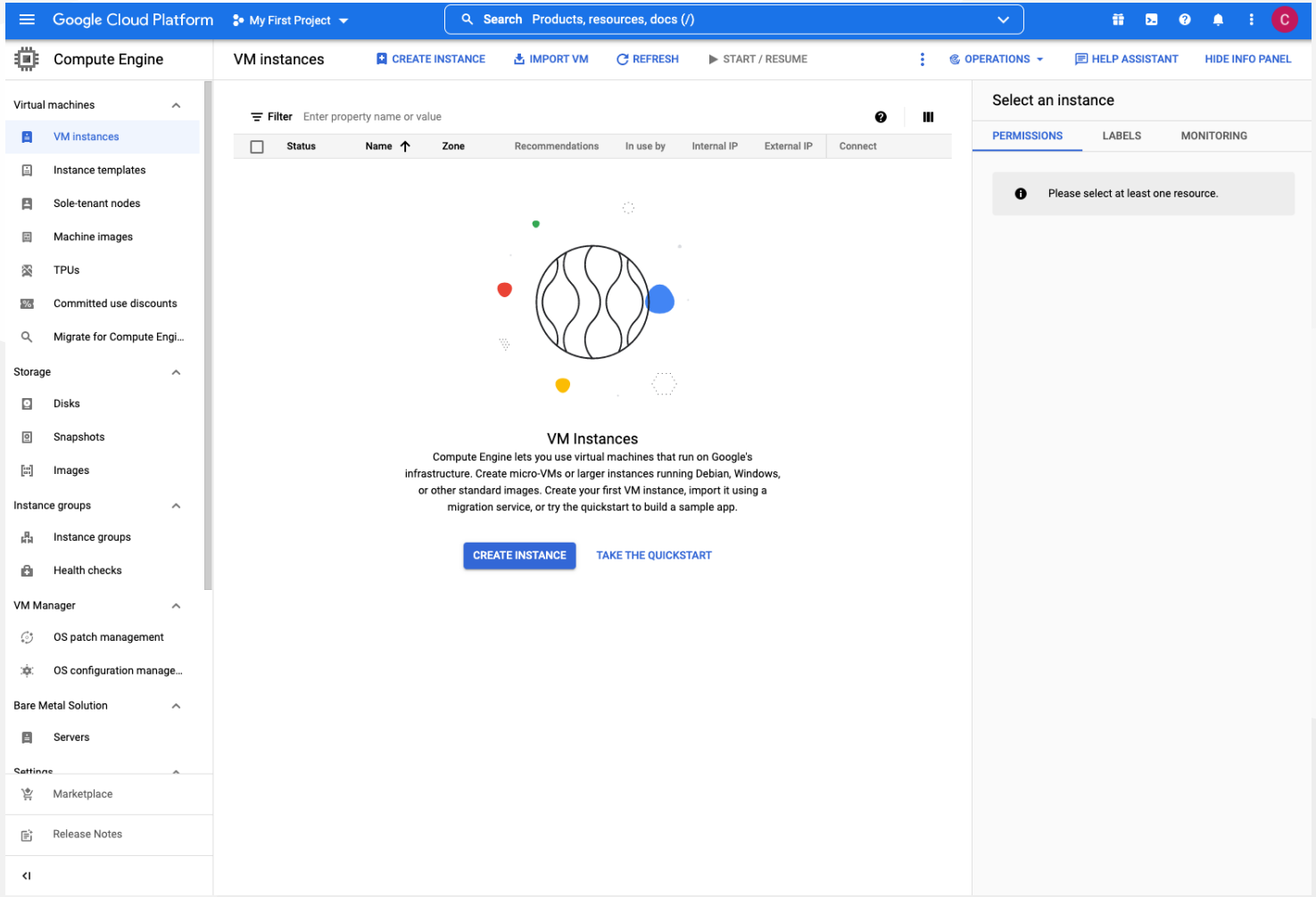
Prerequisites:

- access to cloud shell or gcloud ready in your dev environment.

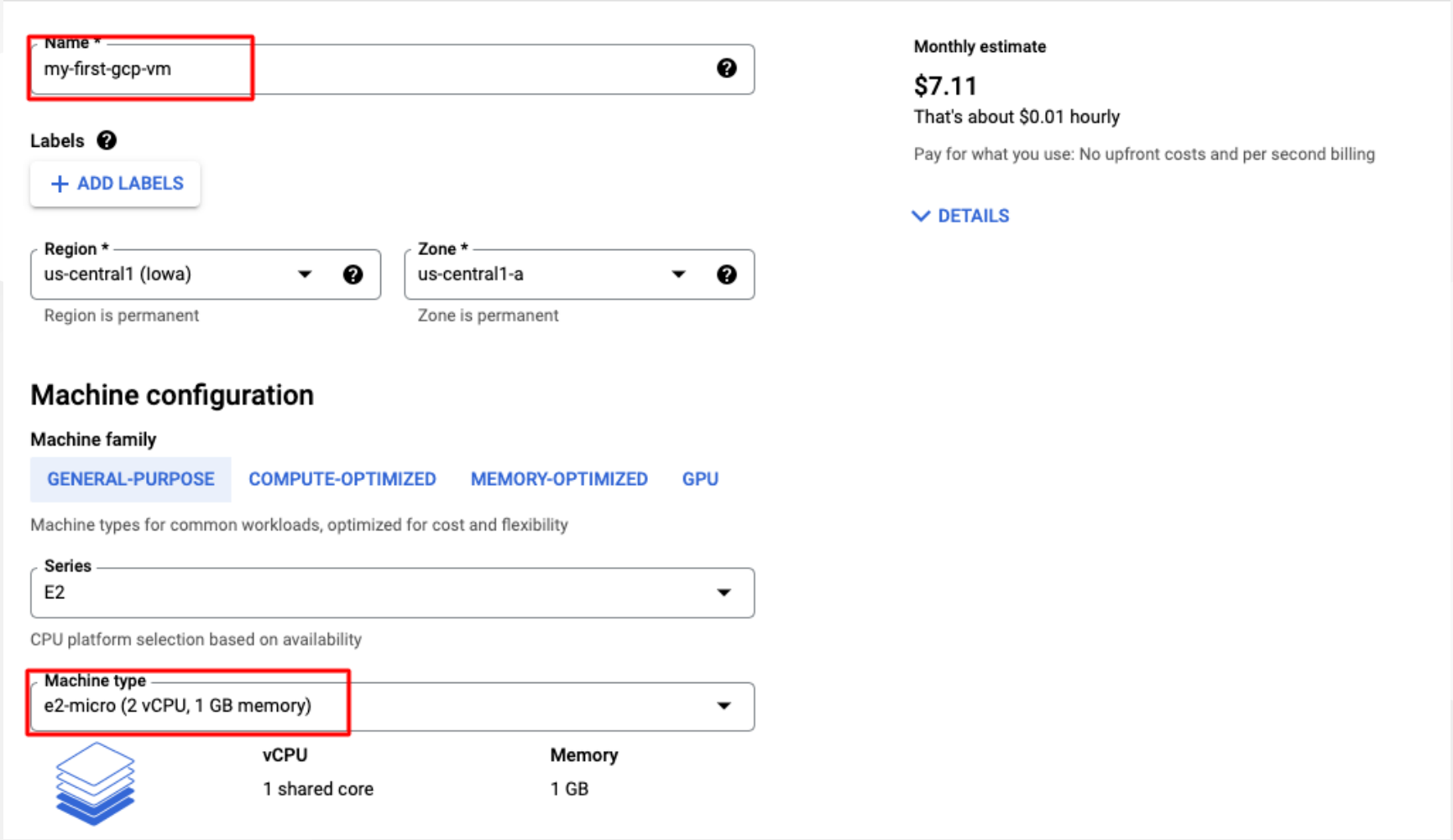
Google Compute Engine



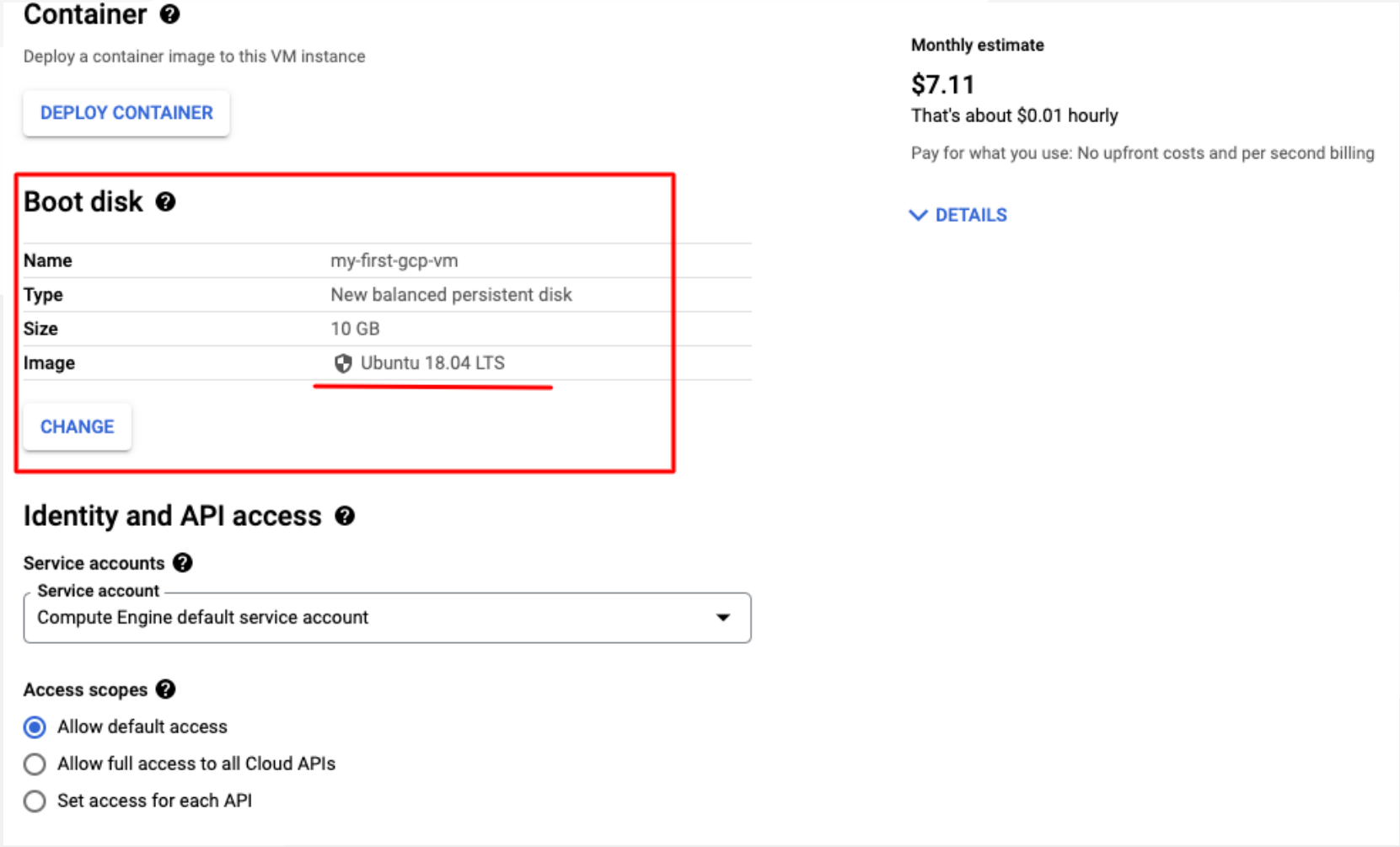
Google Compute Engine



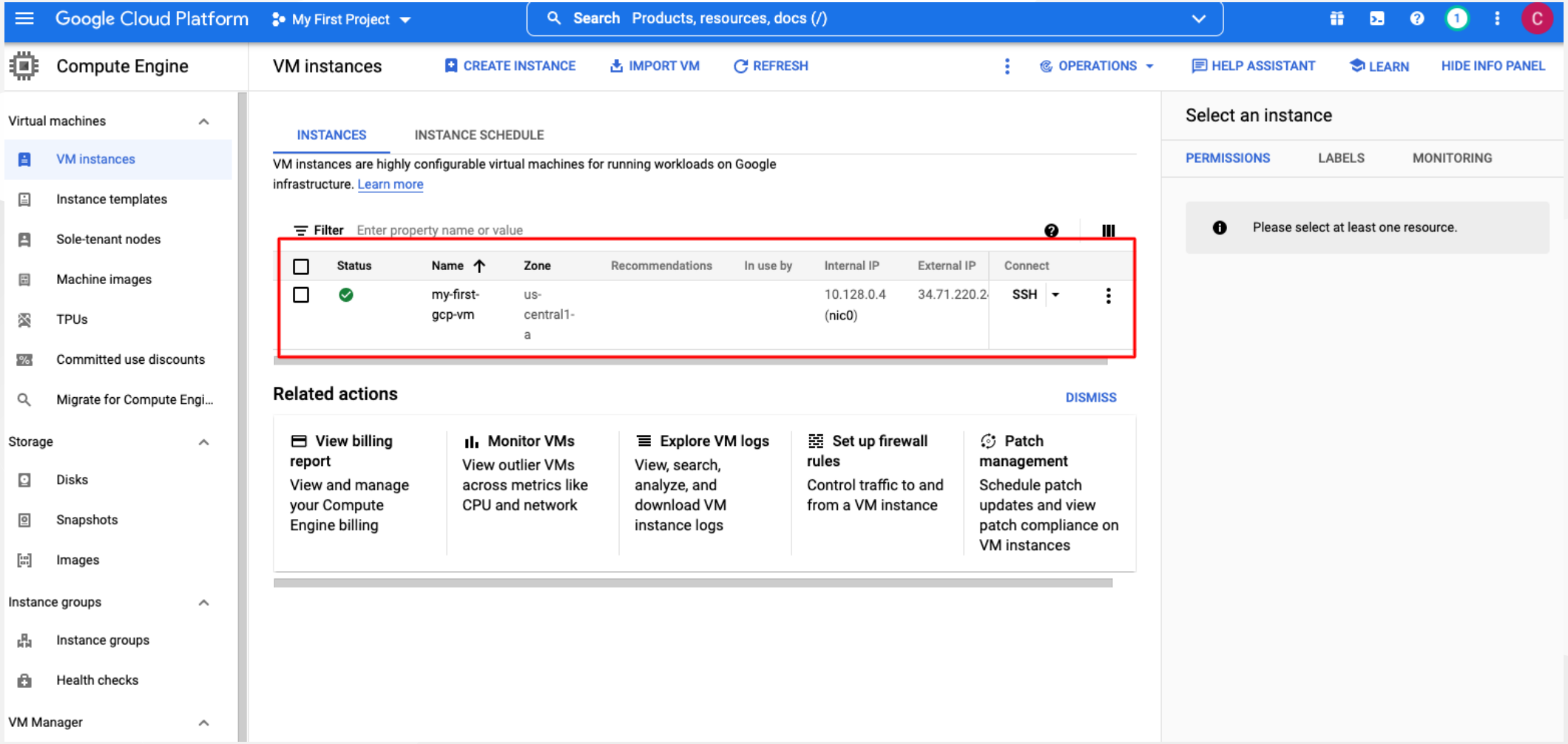
Create Instance: Name/Machine Type



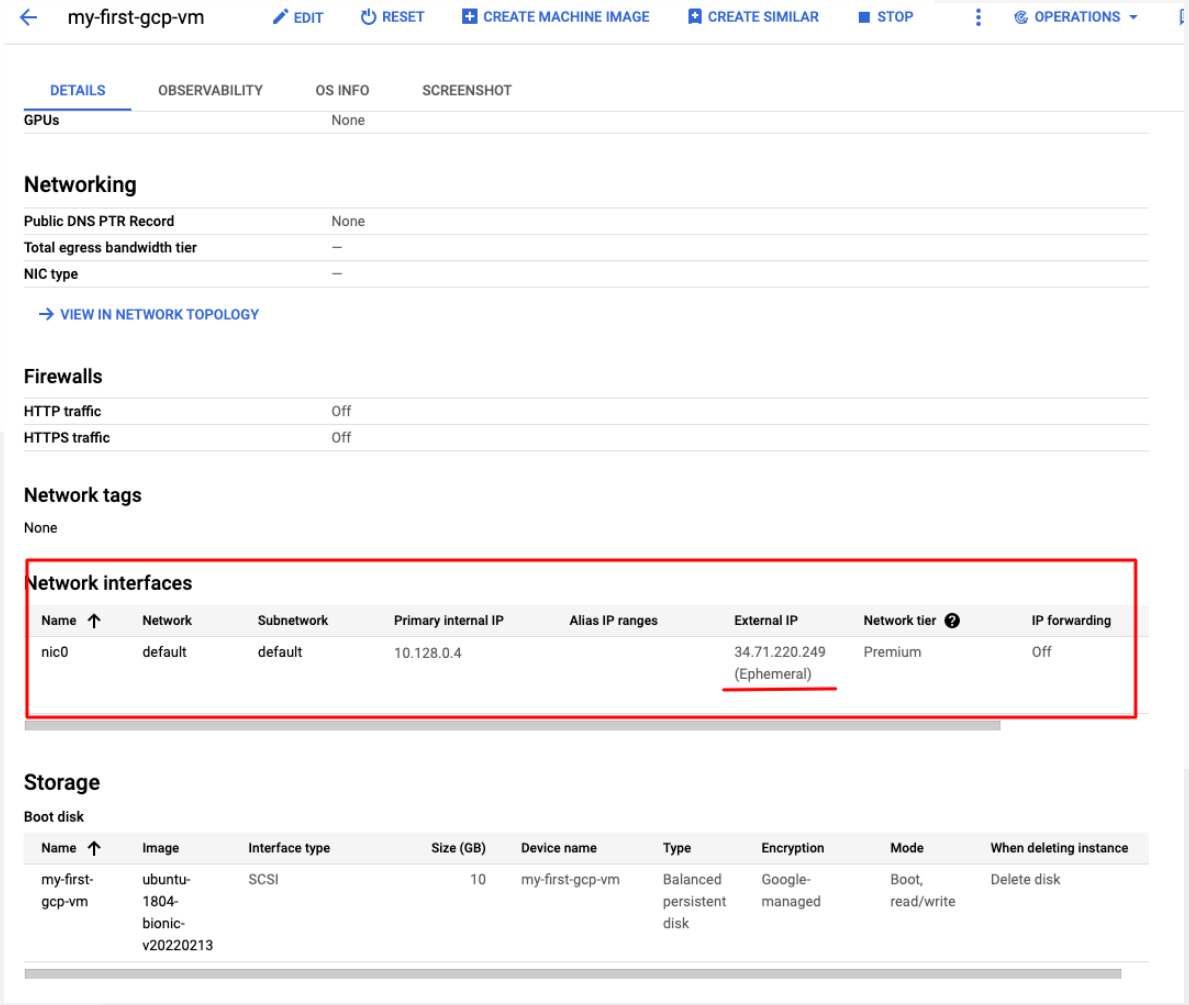
Create Instance: Boot disk Ubuntu



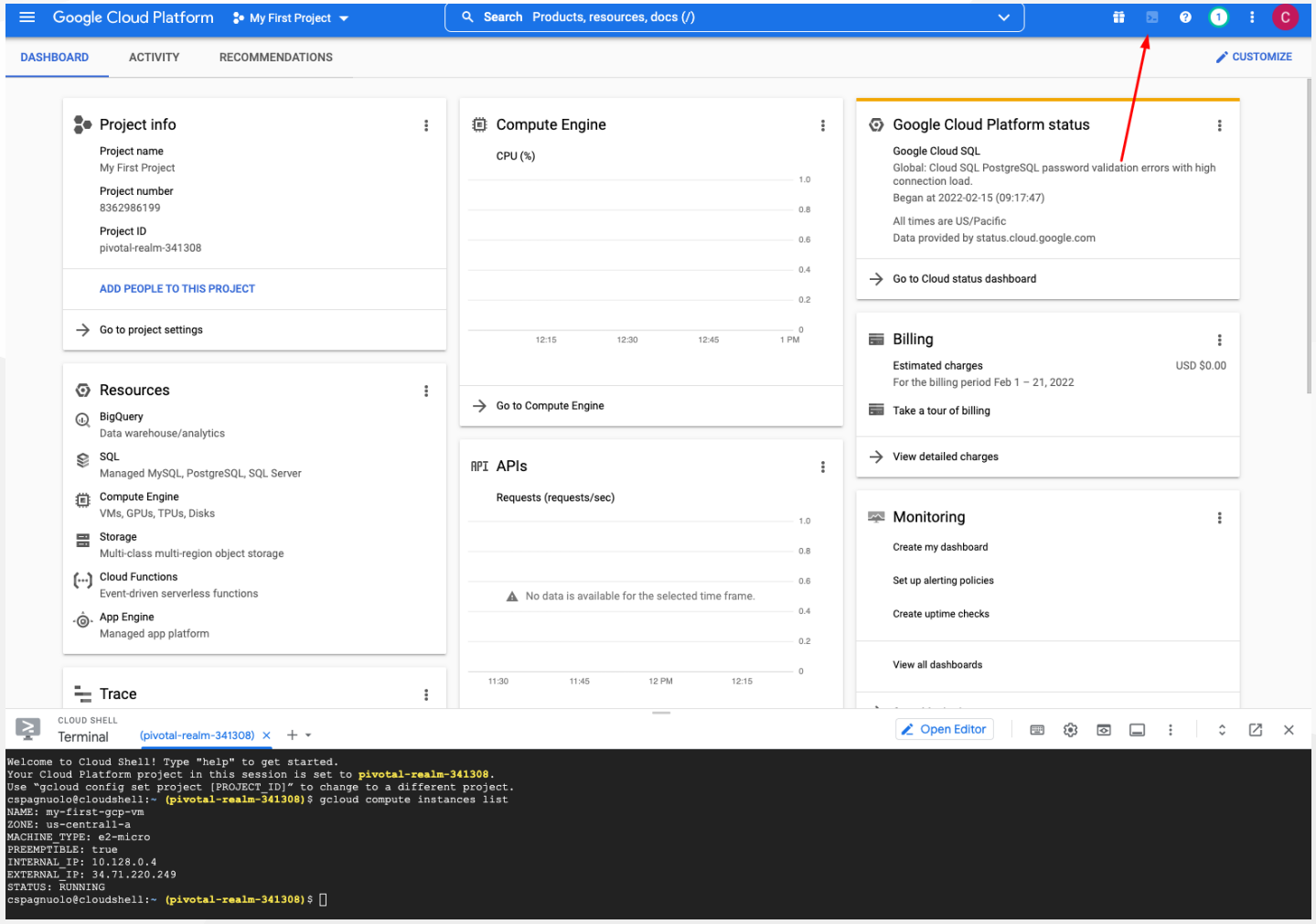
Your first VM is ready



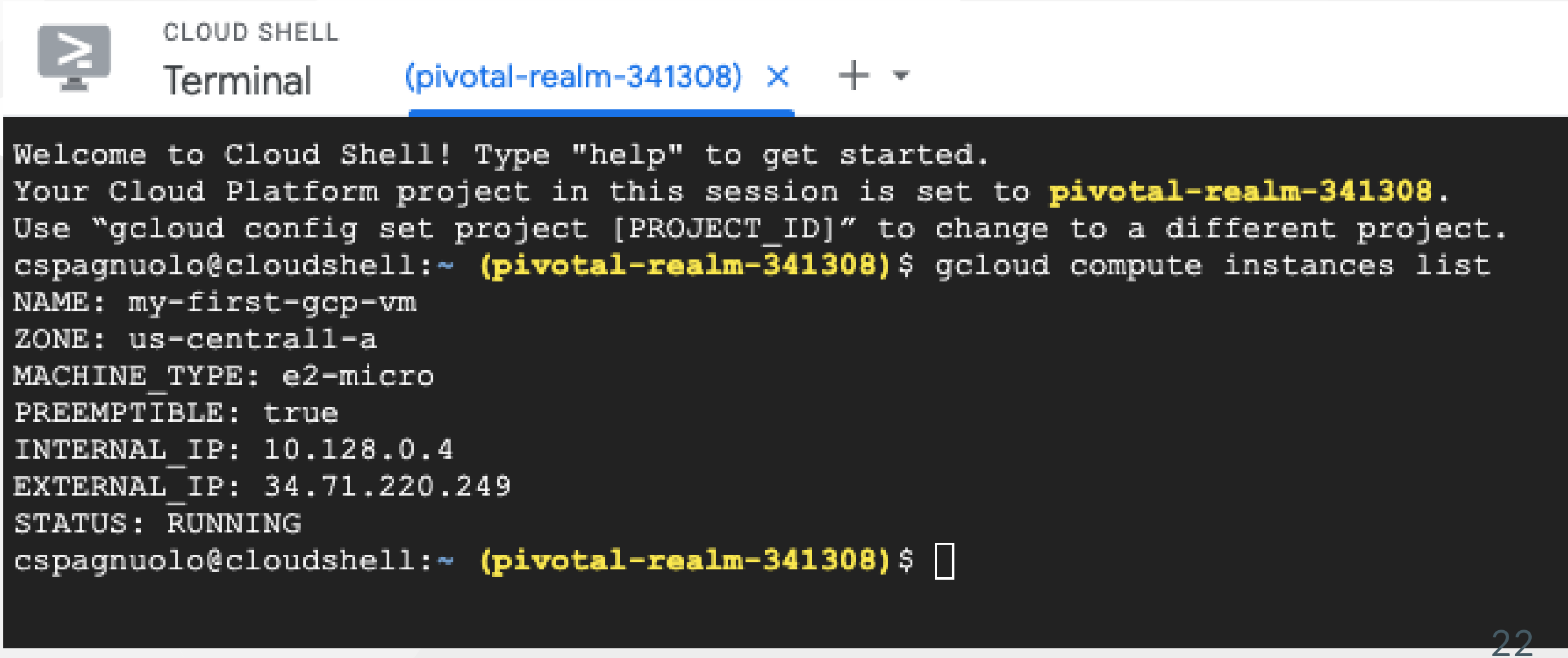
Check the External IP



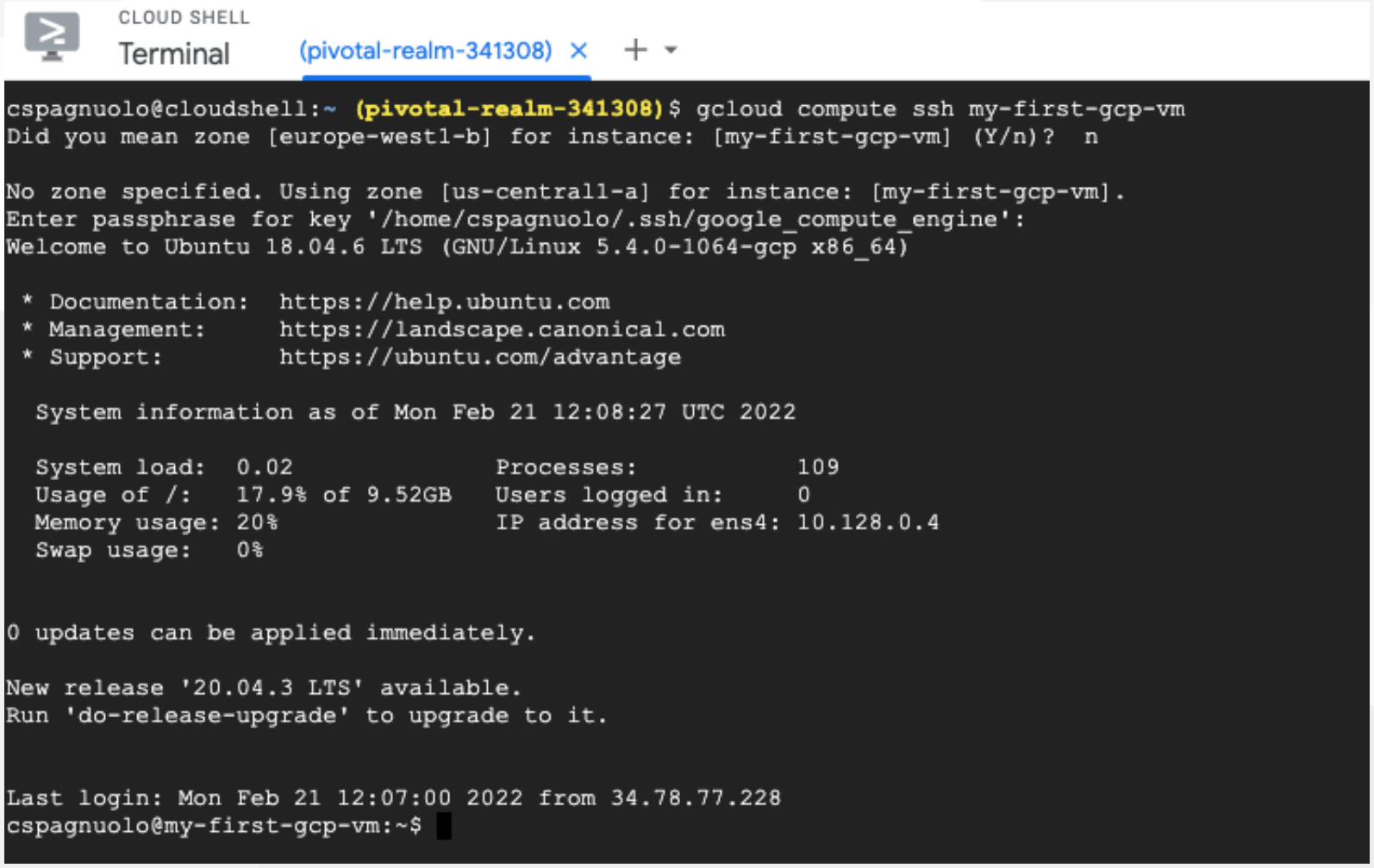
On the command line: gcloud



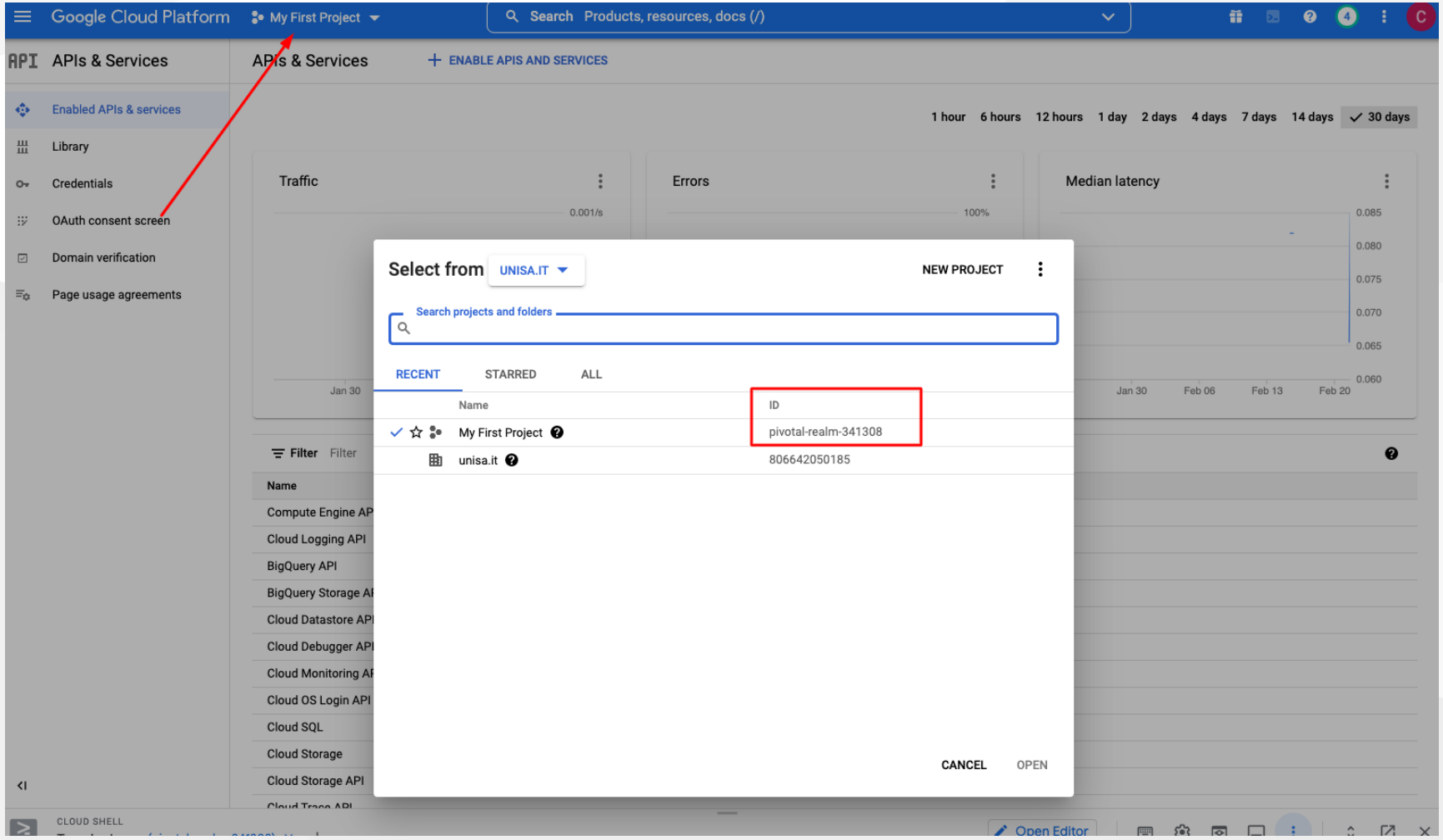
Ready for yuor first gcloud command



Connecting to your instance



Checking your Project ID



In your own code: google-cloud-*

- In this example we use JavaScript and NodeJS from the Cloud Shell.
- Before start install the required libs.

```
$ npm install --save @google-cloud/compute@0.7.1
```

In your own code: google-cloud-*

```
const gce = require('@google-cloud/compute')({
  projectId: 'your-project-id'
});
const zone = gce.zone('us-central1-a');
console.log('Getting your VMs...');
zone.getVMs().then((data) => {
  data[0].forEach((vm) => {
    console.log('Found a VM called', vm.name);
  });
  console.log('Done.');
});
```

In your own code: google-cloud-*

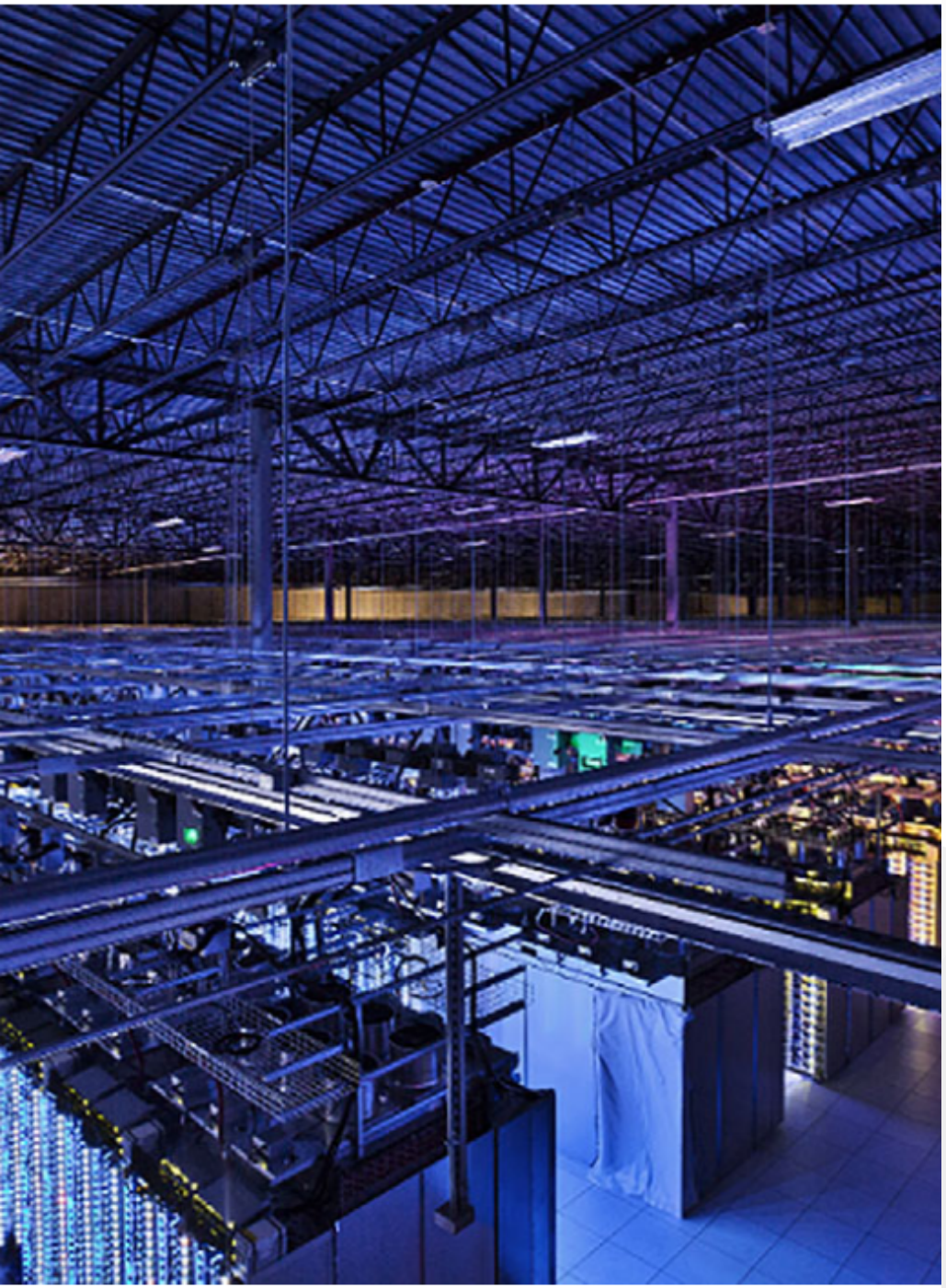
```
$ node script.js
Getting your VMs...
Found a VM called my-first-gcp-vm
Done.
```

In your own code: google-cloud-*

```
const gce = require('@google-cloud/compute')({
  projectId: 'your-project-id'
});
const zone = gce.zone('us-central1-a');
console.log('Getting your VMs...');
zone.getVMs().then((data) => {
  data[0].forEach((vm) => {
    console.log('Found a VM called', vm.name);
    console.log('Stopping', vm.name, '...');
    vm.stop((err, operation) => {
      operation.on('complete', (err) => {
        console.log('Stopped', vm.name);
}); });
});
});
```

In your own code: google-cloud-*

```
$ node script.js
Getting your VMs...
Found a VM called my-first-gcp-vm
Stopping my-first-gcp-vm ...
Stopped my-first-gcp-vm
```

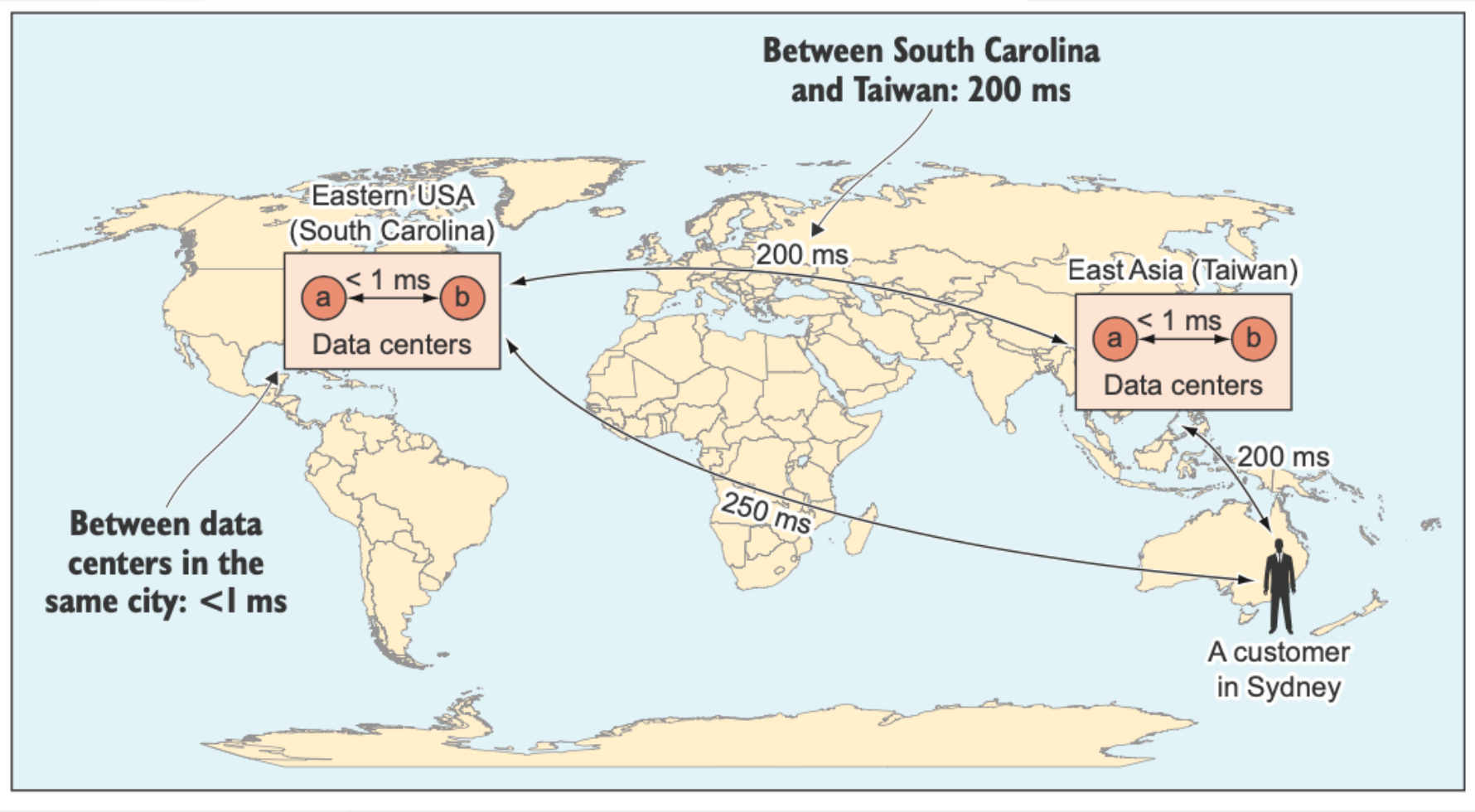


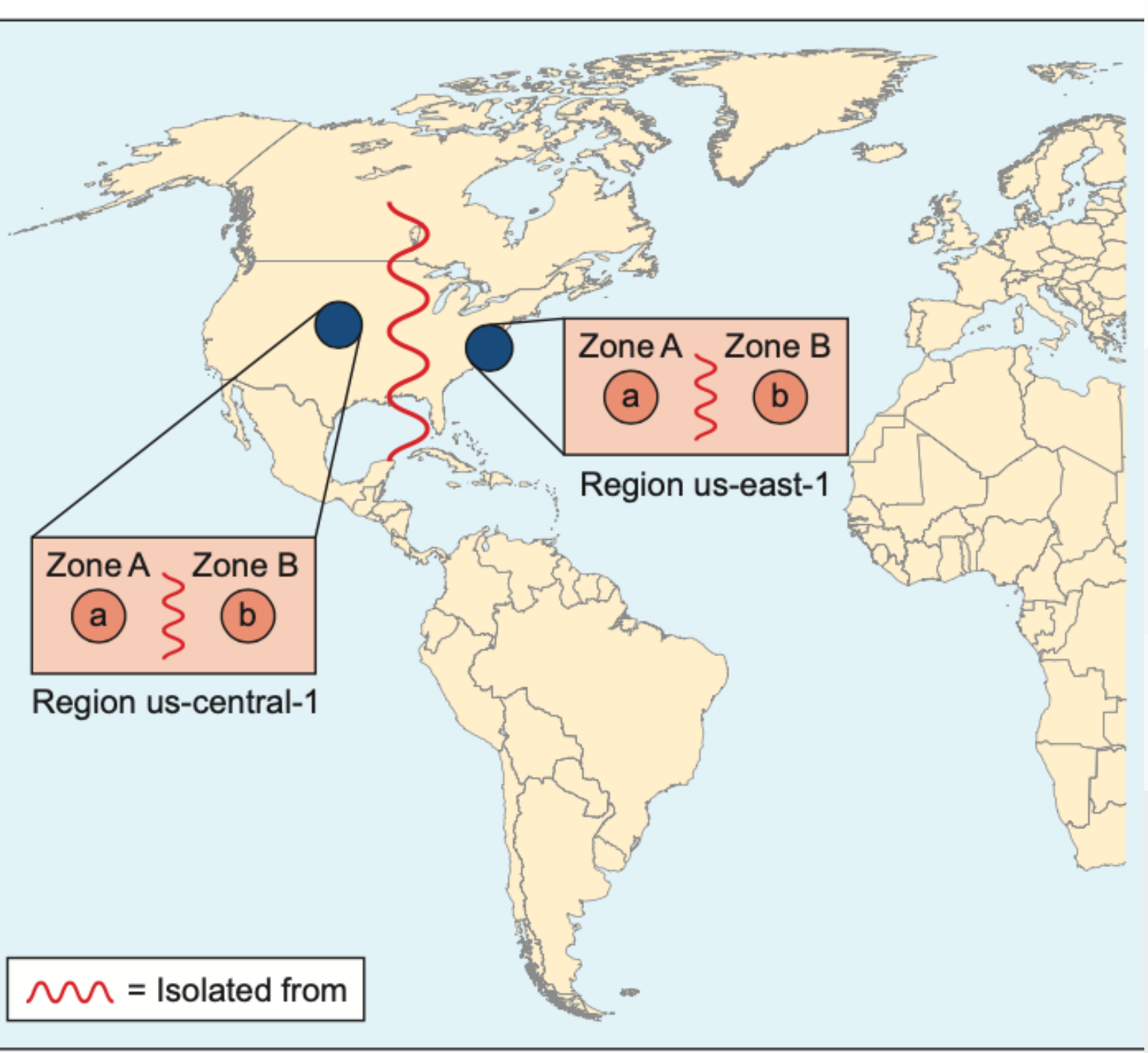
The cloud data center

- If you've ever paid for web hosting before, it's likely that the computer running as your web host was physically located in a data center.
- Cloud is similar to traditional hosting.
- Now we highlight some of the details of Google Cloud Platform's data centers.

Data center locations

Network latency to reach clients





- A **zone** is the smallest unit in which a resource can exist. A unique identifier for a particular facility like **us-east1-b**.
- A collection of zones is called a **region**, such as Council Bluffs, Iowa, USA.

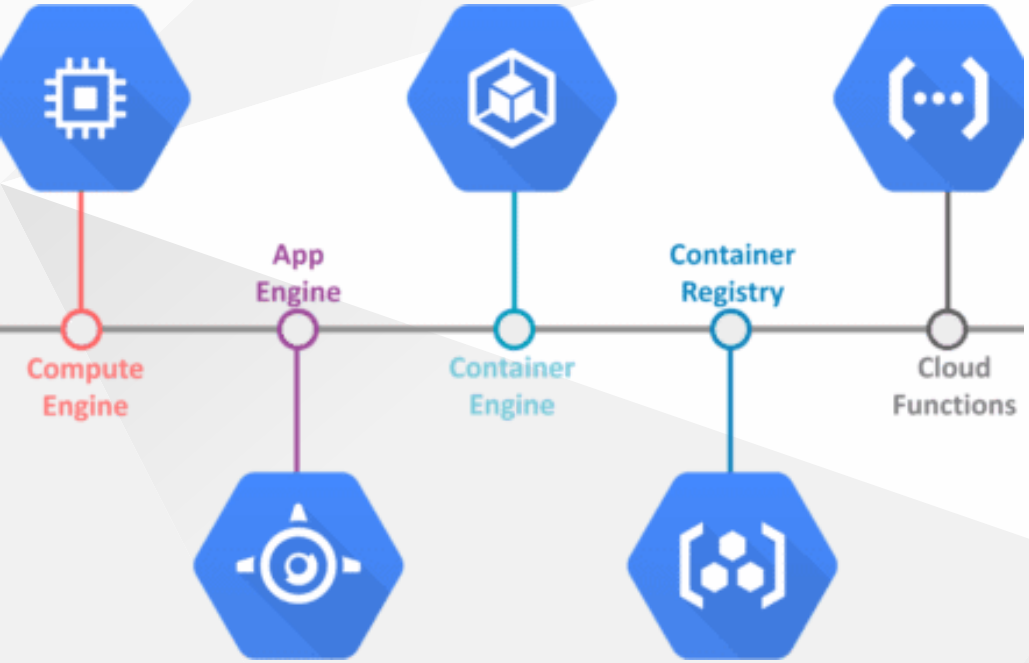
Google Cloud Platform

Services Overview

Google Cloud Platform services

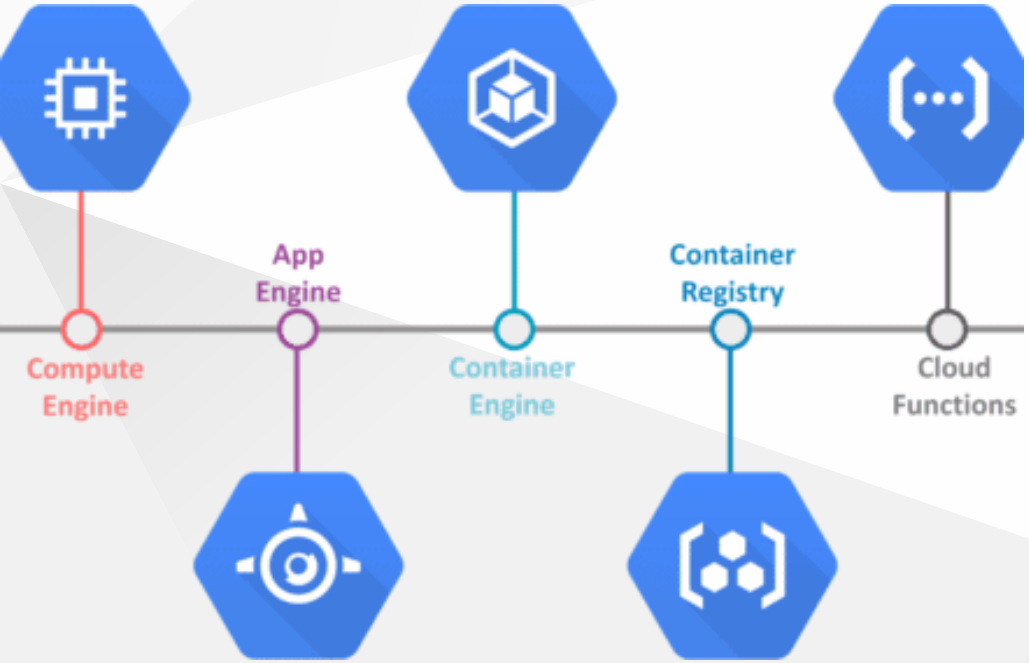
COMPUTE	STORAGE/DATABASES	NETWORKING	BIG DATA/IoT	MACHINE LEARNING
Compute Engine	Cloud Storage	Virtual Private Cloud (VPC)	BigQuery	Cloud Machine Learning Engine
App Engine	Cloud SQL	Cloud Load Balancing	Cloud Dataflow	Cloud Jobs API
Container Engine	Cloud Bigtable	Cloud CDN	Cloud Dataproc	Cloud Natural Language API
Cloud Functions	Cloud Spanner	Cloud Interconnect	Cloud Datalab	Cloud Speech API
	Cloud Datastore	Cloud DNS	Cloud Dataprep	Cloud Translation API
	Persistent Disk		Cloud Pub/Sub	Cloud Vision API
	Data Transfer		Genomics	Cloud Video Intelligence
			Google Data Studio	
			Cloud IoT Core	

Computing Service



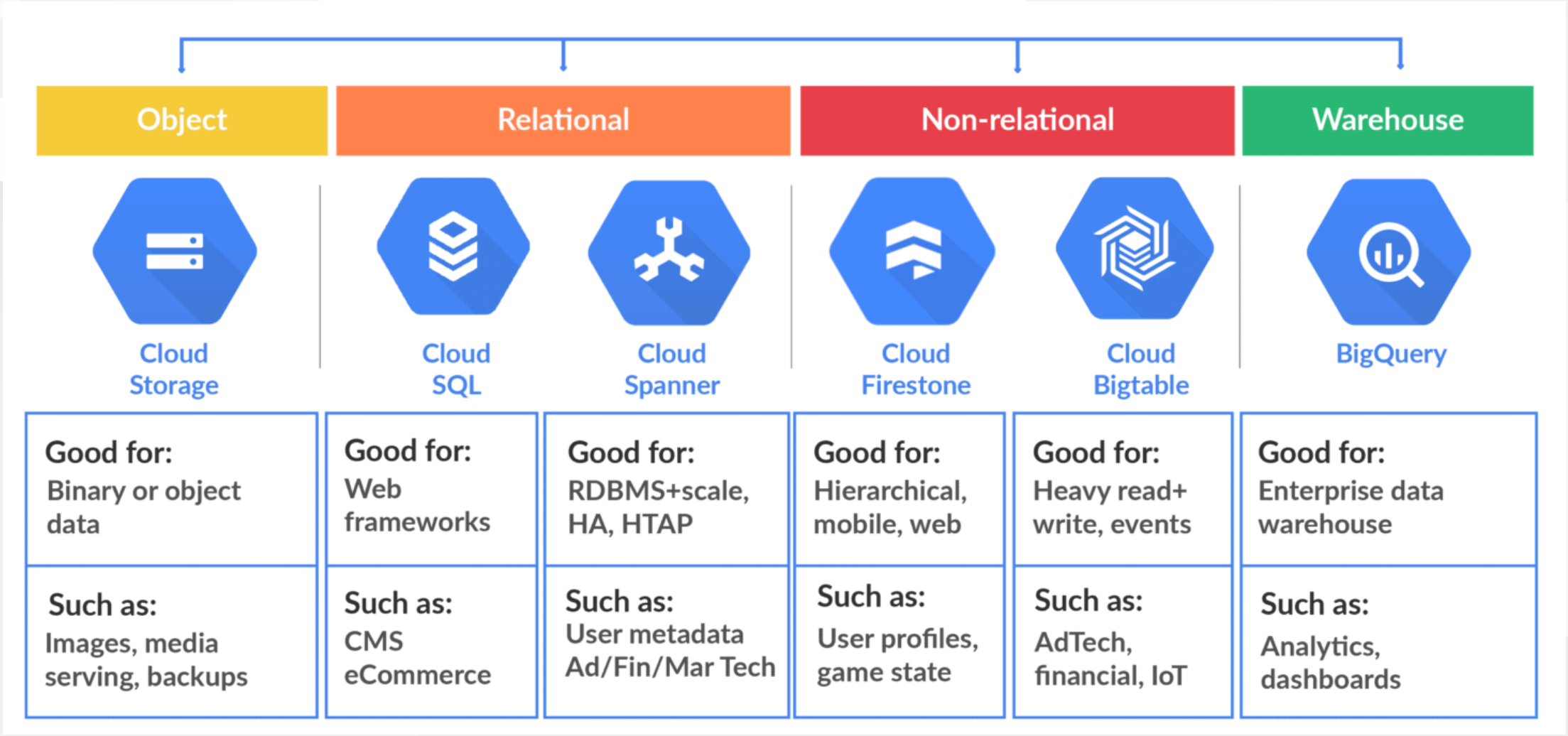
- **Compute engine** is IaaS (Infrastructure As A Service).
- **Google App Engine** is PaaS (Platform As A Service) used for building scalable web applications and IoT backends.
- **Cloud Functions** is a functions as a service (FaaS).

Computing Service



- **Google Kubernetes Engine/ Container Engine** is a strong Cluster Manager and balanced system for running Docker containers.
- **Google Cloud Container Registry** is a private Docker repository to manage Docker Images.

Storage And Database Services



Storage And Database Services

- **Cloud Storage** is a service for storing objects in Google Cloud. An object is an immutable piece of data consisting of a file of any format.
- **Cloud SQL** fully-managed database service (relational).
- **Cloud Bigtable** is a high-performance NoSQL Big Data database service.
- **Cloud Datastore** is a NoSQL database that stores data in JSON documents (similar to MongoDB).

Storage And Database Services

- **Persistent Disk** is a service that provides SSD and HDD storage that can be attached to instances running in either Compute Engine or Container Engine.
- **Cloud Spanner** is a managed globally distributed relational database with ACID transactions, strong consistency, SQL semantics, horizontal scaling, and high availability features.

Network Services

Cloud Virtual Network

Cloud Load Balancing

Cloud CDN

Cloud Interconnect

Cloud DNS

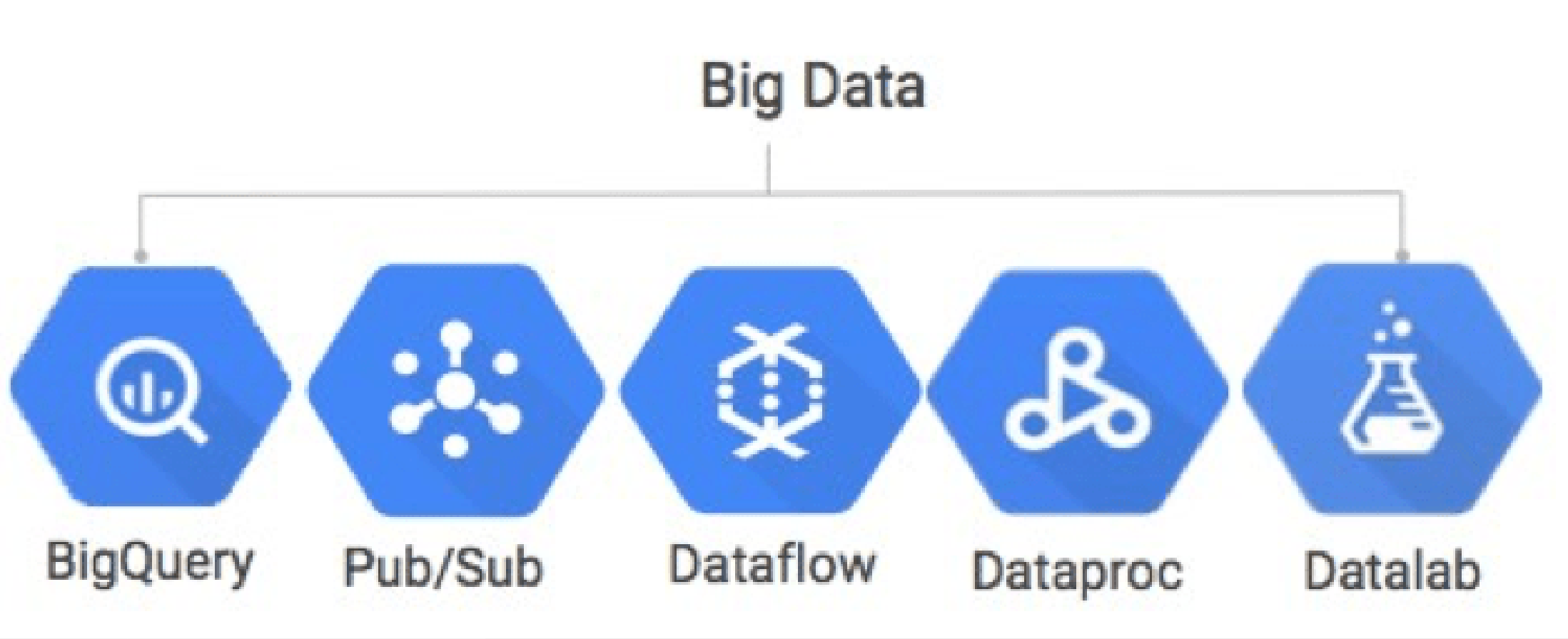
Network Services

- **Virtual Private Cloud (VPC)** provides a set of networking services that are used by VM instances.
- **Cloud Load Balancing** is a process of distributing workloads across many computing resources.
- **Cloud DNS (Domain Naming System)** is a scalable, reliable, and managed authoritative DNS service running on the same infrastructure as Google.

Network Services

- **Cloud CDN (Content Delivery Network)** is a geographically distributed network that consists of proxy servers and their data centers.
- **Google Cloud Interconnect** allows Cloud platform customers to connect to Google via enterprise-grade connections with higher availability and lower latency than their existing Internet connections.

Big Data Services



Big data Services

- **BigQuery** is a fully managed data analysis service that enables users to:
  - create custom schemas to organize data into tables and datasets;
  - load data from different sources;
  - use SQL like commands to query large data;
  - make use of web UI, command-line interface, or API;
  - manage and protect data.

Big Data Services

- **Cloud Dataflow** provides a set of managed services and a set of SDKs that can be used to perform batch and streaming data processing tasks.
- **Google Cloud Pub/Sub** is an asynchronous, serverless, large-scale, reliable real-time messaging service.
- **Cloud Dataproc** is a managed Spark and Hadoop service used to process big datasets using the powerful and open tools in the Apache big data ecosystem.
- **Cloud Datalab** i is an interactive Jupyter like notebook used to explore, collaborate, analyze, and visualize data.

Machine Learning Services

Custom ML models

TensorFlow

Machine Learning Engine

Pre-trained ML models

Vision API

Speech API

Jobs API

Natural Language API

Translation API

Video Intelligence API

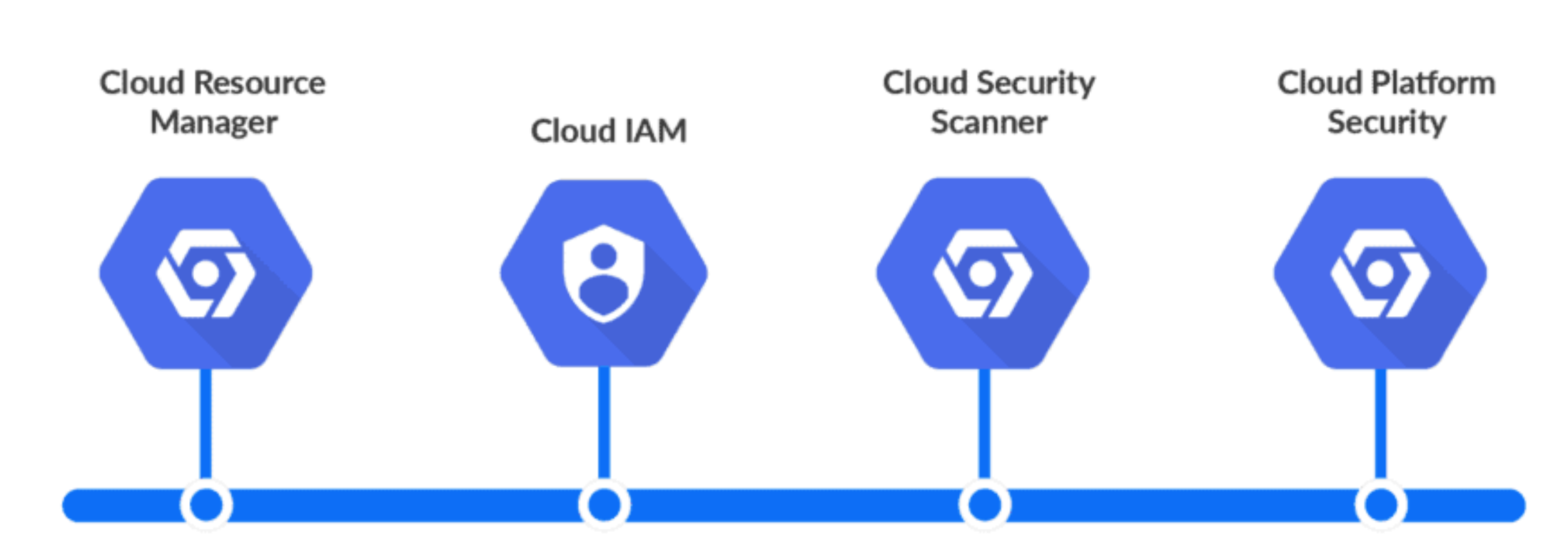
Machine Learning Services

- **Cloud Machine Learning** is a managed service that enables users to build Machine Learning models based on mainstream frameworks like TensorFlow.
- **Cloud AutoML** is a Machine Learning product that enables developers who are not that experienced in this field to provide train their high-quality models by Google's transfer learning and Neural Architecture Search.
- **Cloud Vision API** is a REST API used for image recognition and classification.

Machine Learning Services

- **Cloud Speech API** is a REST API that can be used to convert audio to text recognizing over 110 languages and variants, to support customer's global user base.
- **Cloud Natural Language API** allows users to add sentiment analysis, entity analysis, entity-sentiment analysis, content classification, and syntax analysis.
- **Translate API** allows the users to quickly translate source text into any of over a hundred supported languages. Language detection helps out in cases when the source language is unknown.

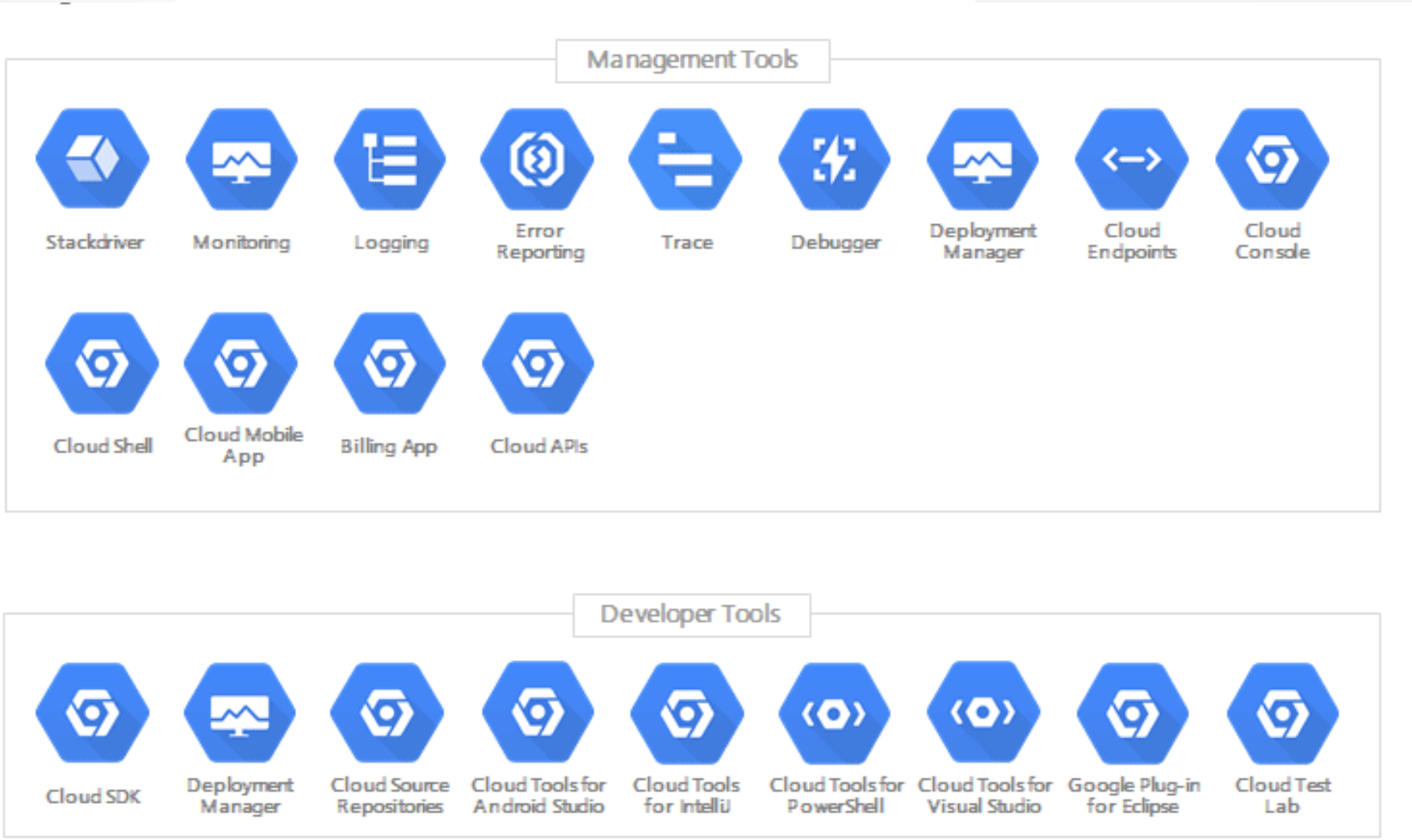
Identity And Security Services



Identity And Security Services

- **Google Cloud IAM** can be defined as a framework of policies and technologies for ensuring that authorized people in an enterprise have the appropriate access to technology and resources.
- **Cloud Resource Manager** is used for programmatically managing resource containers used for grouping and hierarchically organizing GCP resources.
- **Cloud Security Scanner** is a web security scanner for common vulnerabilities in App Engine applications, like cross-site-scripting (XSS), Flash injection, mixed content, and outdated or insecure libraries.

Management And Developer Tools



👀 In-depth view of

Google Compute Engine (GCE)

But first.. Launching your first (or second) VM

👍 If you don't have gcloud installed yet, check out https://cloud.google.com/sdk for instructions on how to get set up, or use the Cloud Shell.

- The first thing you'll need to do is authenticate using `gcloudauth login`.
- `gcloud config set project your-project-id-here`

But first.. Launching your first (or second) VM

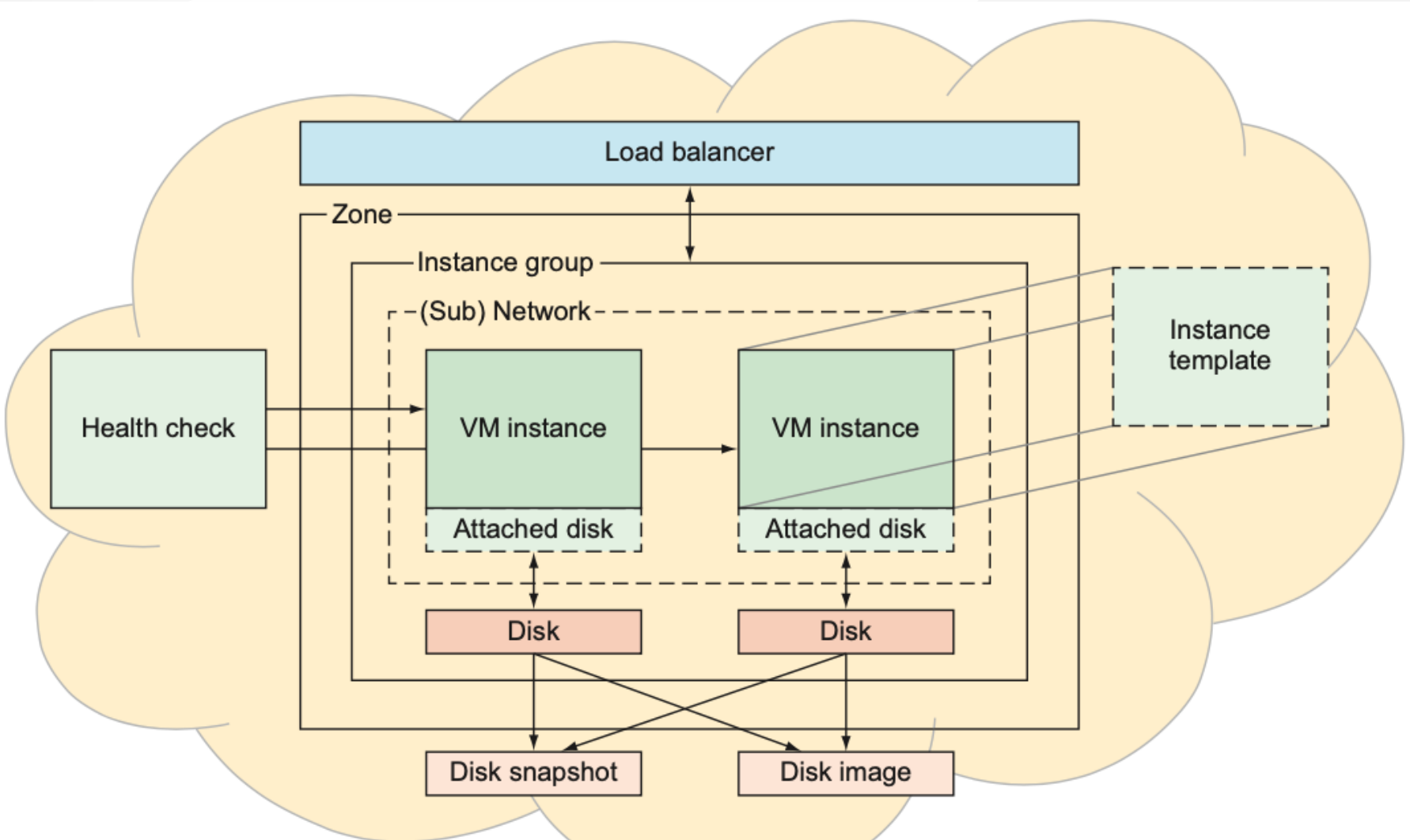
```
$ gcloud compute instances create test-instance-1 --zone us-central1-a
```

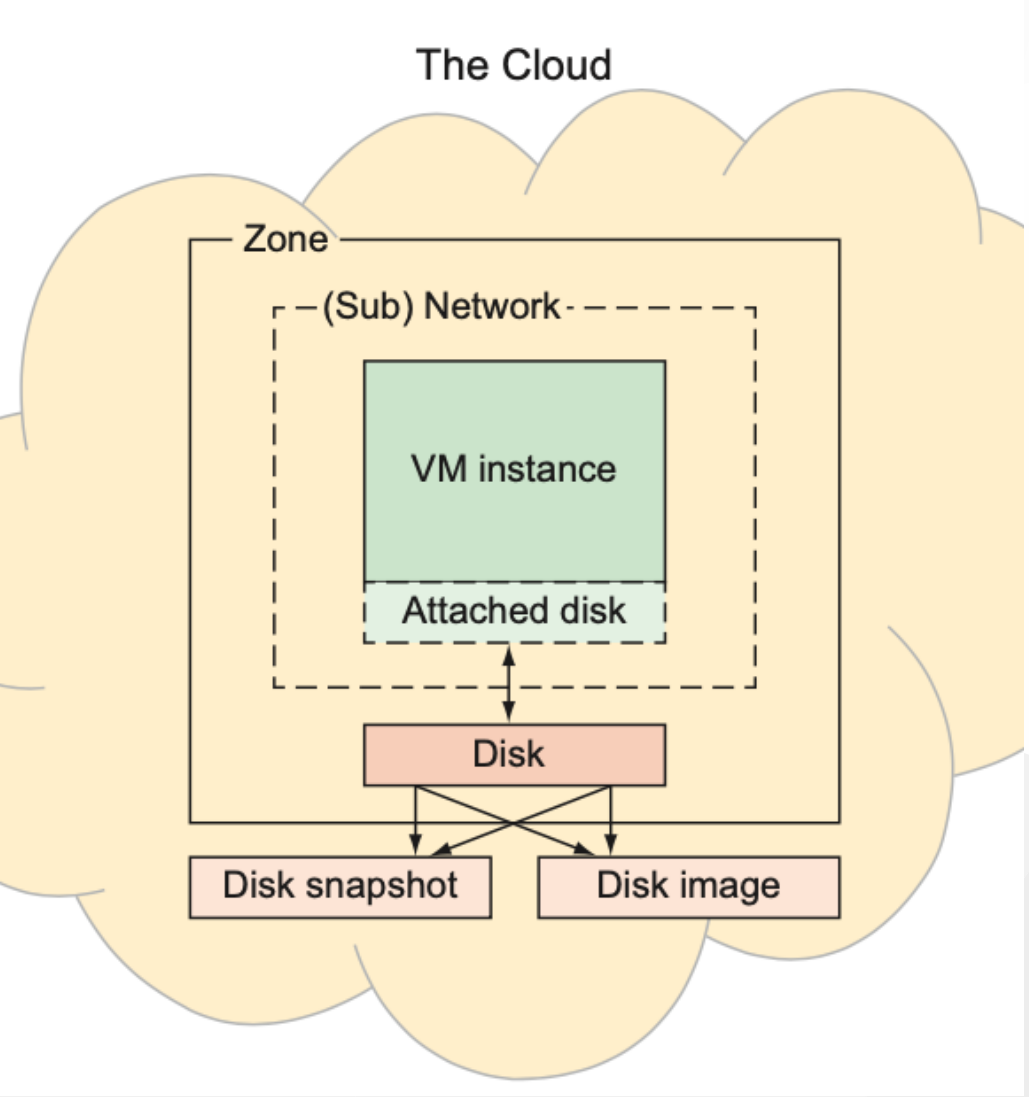
↪ connect to the instance over SSH

```
$ gcloud compute ssh --zone us-central1-a test-instance-1
```

💡 If you are looking for your instance also in the web console ➡️ use the refresh button.

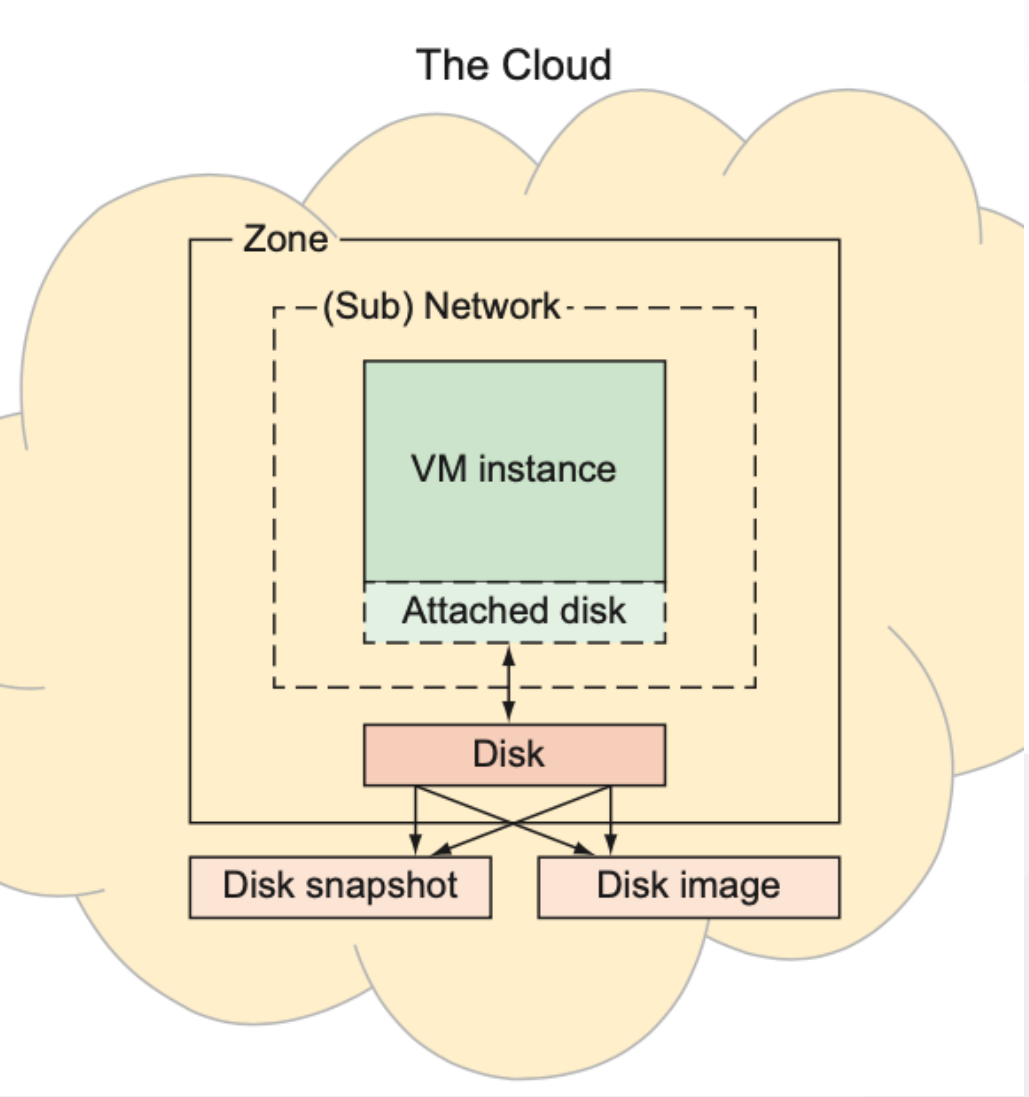
But … what happened?





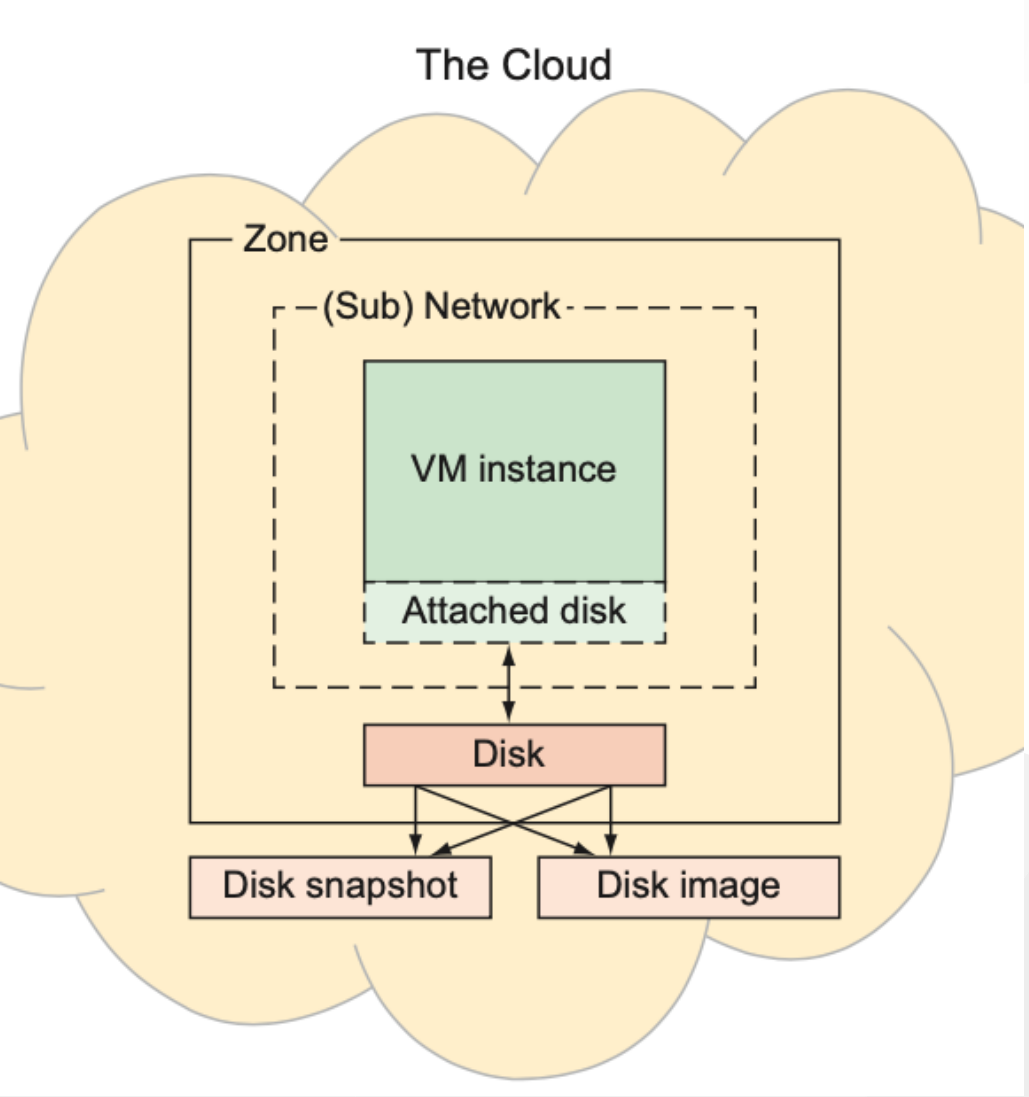
Block storage with Persistent Disks

- A persistent **disk** is a bit like an external hard drive.
- Cloud hosting providers came up with a storage service that looked and acted like a regular disk but was replicated and highly available.



Block storage with Persistent Disks

- A **snapshot** is a check-point for the data on your disk.
- It uses differential storage, storing only what's changed from one snapshot to the next.
- **Images** are similar to snapshots in that both can be used as the source of content when you create a new disk.



Block storage with Persistent Disks

- Every time you create a new VM from a base operating system, you're using an image under the hood.
- The primary difference is that an image doesn't rely on differential storage like snapshots do, which means it may be more expensive to keep around.

Understanding pricing

The basic features of GCE have straightforward prices, whereas some of the more advanced features can get complicated, and even more complicated when you consider an important discount available for sustained use.

You need to consider three factors for pricing with GCE:

1. Computing capacity using CPUs and memory.
2. Storage using persistent disks.
3. Network traffic leaving Google Cloud.

Computing capacity

- The most common way of using GCE is with a predefined instance type, such as `n1-standard-1`.
- By turning on an instance of a particu- lar predefined type, you're charged a specific amount every hour for the use of the computing capacity.
- That capacity is a set amount of CPU time, which is measured in vCPUs (a virtual CPU measurement), and memory, which is measured in GB.

Computing capacity

- Each predefined type has a specific number of vCPUs, a specific amount of memory, and a fixed hourly cost.
- Price list
- Compute Engine

Computing capacity: Spot VMs

- Spot VM are available at much lower price—a 60-91% discount—compared to the price of standard VMs.
- However, Compute Engine might preempt Spot VMs if it needs to reclaim those resources for other tasks.
- At this uncertain preemption time, Compute Engine either stops (default) or deletes your Spot VMs depending on your specified termination action for each VM.

- Google Cloud documentation

cspagnuolo@unisa.it