

GCP Assignment 1

1. What is the GCP project quota? If necessary, how GCP quota can be increased?

- GCP Project quotas **protect Google Cloud users from unforeseen spikes in usage.**
- However, as your usage of Google Cloud Platform increases, you can request an increase in your project quota.
- Quotas also help you with resource management: for instance, you can set your own caps on service usage within your Google-provided quota while developing and testing your applications.

GCP quota can be increased by following below steps:

1. In the Google Cloud Console, go to the **Quotas** page.
2. On the **Quotas** page, find the quota you want to increase in the **Limit name** column.
3. Select the checkbox to the left of your quota.
4. Click create **EDIT QUOTAS**. The **Quota changes** form displays.
5. In the **Quota changes** form, enter the increased quota that you want for your project in the **New limit** field.
6. Complete any additional fields in the form, and then click **DONE**.
7. Click **SUBMIT REQUEST**.

2. What is the definition of a virtual machine? Virtual machine types are offered in GCP. How to create a virtual computer in Google Cloud Platform.

- A virtual machine is a virtualized instance of a computer that can perform almost all of the same functions as a computer, including running applications and operating systems.

Virtual machine types are offered in GCP:

- Scale-out workloads (T2D)
T2D offers the best price-performance compared to general-purpose VMs from any of the leading public cloud vendors. It is the first instance type in the Tau VM family and comes in predefined shapes, with up to 60vCPUs per VM and 4GB of memory per vCPU. T2D is ideal for scale-out workloads like web servers, containerized microservices, media transcoding, and large scale java applications.
- General purpose workloads (E2, N2, N2D, N1)

E2, N2, N2D, and N1 are general-purpose machines offering a good balance of price and performance, and are suitable for a wide variety of common workloads including databases, development and testing environments, web applications, and mobile gaming. They support up to 224 vCPUs and 896 GB of memory.

- **Ultra-high memory (M2, M1)**
Memory-optimized machines offer the highest memory configurations with up to 12 TB for a single instance. They are well suited to memory-intensive workloads such as large in-memory databases like SAP HANA, and in-memory data analytics workloads.
- **Compute-intensive workloads (C2, C2D)**
Compute-optimized machines provide the highest performance per core on Compute Engine and are optimized for workloads such as high performance computing (HPC), game servers, and latency-sensitive API serving.
- **Most demanding applications and workloads (A2)**
Accelerator-optimized machines are based on the NVIDIA Ampere A100 Tensor Core GPU. Each A100 GPU offers up to 20x the compute performance compared to the previous generation GPU. These VMs are designed for your most demanding workloads such as machine learning and high performance computing.

create a virtual computer in Google Cloud Platform :

1. In the Google Cloud Console, go to the **VM instances** page.
2. Select your project and click **Continue**.
3. Click **Create instance**.
4. Specify a **Name** for your VM. For more information, see Resource naming convention.
5. Optional: Change the **Zone** for this VM. Compute Engine randomizes the list of zones within each region to encourage use across multiple zones.
6. Select a **Machine configuration** for your VM.
7. In the **Boot disk** section, click **Change**, and then do the following:
 - a. Select the **Custom Images** tab.
 - b. To select the image project, click **Select a project**, and then do the following:
 - i. Select the project that contains the image.

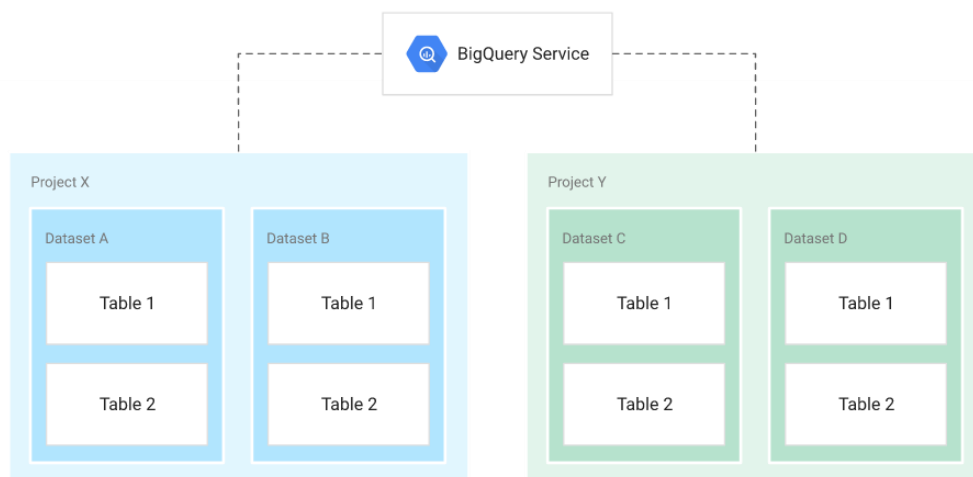
- ii. Click **Open**.
 - c. In the **Image** list, click the image that you want to import.
 - d. Select the type and size of your boot disk.
 - e. Optional: For advanced configuration options, click **Show advanced configuration**.
 - f. To confirm your boot disk options, click **Select**.
8. In the **Firewall** section, to permit HTTP or HTTPS traffic to the VM, select **Allow HTTP traffic** or **Allow HTTPS traffic**.
- a. The Cloud Console adds a network tag to your VM and creates the corresponding ingress firewall rule that allows all incoming traffic on tcp:80 (HTTP) or tcp:443 (HTTPS). The network tag associates the firewall rule with the VM. For more information, see Firewall rules overview in the Virtual Private Cloud documentation.
9. To create and start the VM, click **Create**.

3. What is Google Big Query, and how does it work? Replicate certain instances to demonstrate a use case.

- BigQuery is a web service from Google that is used for handling or analyzing big data. It is part of the Google Cloud Platform. As a NoOps (no operations) data analytics service, BigQuery offers users the ability to manage data using fast SQL-like queries for real-time analysis.

Working:

- BigQuery organizes data tables into units called datasets. These datasets are scoped to your GCP project. These multiple scopes — project, dataset, and table — helps you structure your information logically. You can use multiple datasets to separate tables pertaining to different analytical domains, and you can use project-level scoping to isolate datasets from each other according to your business needs.



Replication use cases

The following use cases apply for each type of replication.

Name	Primary	Replica	Benefits and use cases	More information
Read replica	Cloud SQL instance	Cloud SQL instance	<ul style="list-style-type: none"> Additional read capacity Analytics target 	<ul style="list-style-type: none"> Creating read replicas Managing read replicas Replication metrics Read replica indexes
Cross-region read replica	Cloud SQL instance	Cloud SQL instance	<ul style="list-style-type: none"> Additional read capacity Analytics target Additional disaster recovery capability Improve read performance Migrate data between regions 	<ul style="list-style-type: none"> Creating read replicas Managing read replicas Replication metrics
External read replica	Cloud SQL instance	MySQL instance external to Cloud SQL	<ul style="list-style-type: none"> Reduced latency for external connections 	<ul style="list-style-type: none"> Configuring external replicas

			<ul style="list-style-type: none"> Analytics target Migration path to other platforms 	
<ul style="list-style-type: none"> Replication from an external server 	MySQL instance external to Cloud SQL	Cloud SQL for MySQL instance	<ul style="list-style-type: none"> Migration path to Cloud SQL Data replication to Google Cloud Platform Analytics target 	<ul style="list-style-type: none"> <u>Replicating from an external server</u>

Before you can create a read replica of a primary Cloud SQL instance, the instance must meet the following requirements:

- Automated backups must be enabled.
- Binary logging must be enabled which requires point-in-time recovery to be enabled about the impact of these logs.
- At least one backup must have been created after binary logging was enabled.

4. What exactly is the Google Cloud SDK? List the numerous Google cloud SDK installation options.

- Google Cloud SDK is a set of tools for accessing Google public cloud platform in a secure way. It contains essential tools for maintaining, managing and monitoring Google Cloud Platform (GCP).

1. Download the Google Cloud CLI installer.

Alternatively, open a PowerShell terminal and run the following PowerShell commands:

```
(New-Object
Net.WebClient).DownloadFile("https://dl.google.com/dl/cloudsdk/channels/rapid/
GoogleCloudSDKInstaller.exe", "$env:Temp\GoogleCloudSDKInstaller.exe")
```

```
& $env:Temp\GoogleCloudSDKInstaller.exe
```

2. Launch the installer and follow the prompts. The installer is signed by Google LLC.

If you're using a screen reader, check the **Turn on screen reader mode** checkbox. This option configures gcloud to use status trackers instead of unicode spinners, display progress as a percentage, and flatten tables. For more information, see the Accessibility features guide.

3. Cloud SDK requires Python; supported versions are Python 3 (preferred, 3.5 to 3.8) and Python 2 (2.7.9 or later). By default, the Windows version of Cloud SDK comes bundled with Python 3 and Python 2. To use Cloud SDK, your operating system must be able to run a supported version of Python.

The installer installs all necessary dependencies, including the needed Python version. While Cloud SDK installs and manages Python 3 by default, you can use an existing Python installation if necessary by **unchecking** the option to Install Bundled Python. See gcloud topic startup to learn how to use an existing Python installation.

4. After installation is complete, the installer gives you the option to create Start Menu and Desktop shortcuts, start the Google Cloud CLI shell, and configure the gcloud CLI. Make sure that you leave the options to start the shell and configure your installation selected. The installer starts a terminal window and runs the gcloud init command.
5. The default installation doesn't include the App Engine extensions required to deploy an application using gcloud commands. These components can be installed using the gcloud CLI component manager.

Troubleshooting tips:

- If your installation is unsuccessful due to the find command not being recognized, ensure your PATH environment variable is set to include the folder containing find. Usually, this is C:\WINDOWS\system32;.
- If you uninstalled the gcloud CLI, you must reboot your system before installing the gcloud CLI again.
- If unzipping fails, run the installer as an administrator.

5. List the many cloud computing deployment models.

Different types of cloud computing deployment models are:

1. Public cloud
2. Private cloud
3. Hybrid cloud
4. Community cloud
5. Multi-cloud

Let us discuss them one by one:

1. Public Cloud

The public cloud makes it possible for anybody to access systems and services. The public cloud may be less secure as it is open for everyone. The public cloud is one in which cloud infrastructure services are provided over the internet to the general people or major industry groups. The

infrastructure in this cloud model is owned by the entity that delivers the cloud services, not by the consumer. It is a type of cloud hosting that allows customers and users to easily access systems and services. This form of cloud computing is an excellent example of cloud hosting, in which service providers supply services to a variety of customers. In this arrangement, storage backup and retrieval services are given for free, as a subscription, or on a per-use basis. Example: Google App Engine etc.

Advantages of the public cloud model:

- **Minimal Investment:** Because it is a pay-per-use service, there is no substantial upfront fee, making it excellent for enterprises that require immediate access to resources.
- **No setup cost:** The entire infrastructure is fully subsidized by the cloud service providers, thus there is no need to set up any hardware.
- **Infrastructure Management is not required:** Using the public cloud does not necessitate infrastructure management.
- **No maintenance:** The maintenance work is done by the service provider (Not users).
- **Dynamic Scalability:** To fulfill your company's needs, on-demand resources are accessible.

2. Private Cloud

The private cloud deployment model is the exact opposite of the public cloud deployment model. It's a one-on-one environment for a single user (customer). There is no need to share your hardware with anyone else. The distinction between private and public cloud is in how you handle all of the hardware. It is also called the "internal cloud" & it refers to the ability to access systems and services within a given border or organization. The cloud platform is implemented in a cloud-based secure environment that is protected by powerful firewalls and under the supervision of an organization's IT department.

The private cloud gives the greater flexibility of control over cloud resources.

Advantages of the private cloud model:

- **Better Control:** You are the sole owner of the property. You gain complete command over service integration, IT operations, policies, and user behaviour.
- **Data Security and Privacy:** It's suitable for storing corporate information to which only authorized staff have access. By segmenting resources within the same infrastructure, improved access and security can be achieved.
- **Supports Legacy Systems:** This approach is designed to work with legacy systems that are unable to access the public cloud.
- **Customization:** Unlike a public cloud deployment, a private cloud allows a company to tailor its solution to meet its specific needs.

3. Hybrid cloud

By bridging the public and private worlds with a layer of proprietary software, hybrid cloud computing gives the best of both worlds. With a hybrid solution, you may host the app in a safe environment while taking

advantage of the public cloud's cost savings. Organizations can move data and applications between different clouds using a combination of two or more cloud deployment methods, depending on their needs.

Advantages of the hybrid cloud model:

- **Flexibility and control:** Businesses with more flexibility can design personalized solutions that meet their particular needs.
- **Cost:** Because public clouds provide for scalability, you'll only be responsible for paying for the extra capacity if you require it.
- **Security:** Because data is properly separated, the chances of data theft by attackers are considerably reduced.

4. Community cloud

It allows systems and services to be accessible by a group of organizations. It is a distributed system that is created by integrating the services of different clouds to address the specific needs of a community, industry, or business. The infrastructure of the community could be shared between the organization which has shared concerns or tasks. It is generally managed by a third party or by the combination of one or more organizations in the community.

Advantages of the community cloud model:

- **Cost Effective:** It is cost-effective because the cloud is shared by multiple organizations or communities.
- **Security:** Community cloud provides better security.
- **Shared resources:** It allows you to share resources, infrastructure, etc. with multiple organizations.
- **Collaboration and data sharing:** It is suitable for both collaboration and data sharing.

5. Multi-cloud

We're talking about employing multiple cloud providers at the same time under this paradigm, as the name implies. It's similar to the hybrid cloud deployment approach, which combines public and private cloud resources. Instead of merging private and public clouds, multi-cloud uses many public clouds. Although public cloud providers provide numerous tools to improve the reliability of their services, mishaps still occur. It's quite rare that two distinct clouds would have an incident at the same moment. As a result, multi-cloud deployment improves the high availability of your services even more.

Advantages of a multi-cloud model:

- You can mix and match the best features of each cloud provider's services to suit the demands of your apps, workloads, and business by choosing different cloud providers.
- **Reduced Latency:** To reduce latency and improve user experience, you can choose cloud regions and zones that are close to your clients.
- **High availability of service:** It's quite rare that two distinct clouds would have an incident at the same moment. So, the multi-cloud deployment improves the high availability of your services.

6. Describe the Google cloud platform's security features.

Google Cloud Platform's security features are-

- 24/7/365 operations, device security detection and response from both internal and external threats.
- Data in-transit encrypted communication to and from Google's public cloud, including layered defence redundancies to protect customers from denial-of service (DoS) attacks.
- Identity protection and management through multiple authentication factor.
- Data at-rest storage security using encryption against unauthorized access and distribution for reliability.
- An entire hardware infrastructure created, built, controlled, and secured by Google including servers, networking equipment, and security chips.

7. What exactly is vertex AI? With some usage scenarios, implement vertex AI.

Vertex AI Workbench is the single environment for data scientists to complete all of their ML work, from experimentation, to deployment, to managing and monitoring models. It is a Jupyter-based fully managed, scalable, enterprise-ready compute infrastructure with security controls and user management capabilities.

USE CASES:

USE CASE

Data readiness

Vertex AI supports your data preparation process. You can ingest data from BigQuery and Cloud Storage and leverage Vertex AI Data Labeling to annotate high-quality training data and improve prediction accuracy.

USE CASE

Feature engineering

Use Vertex AI Feature Store, a fully managed rich feature repository, to serve, share, and reuse ML features; Vertex AI Experiments to track, analyze, and discover ML experiments for faster model selection; Vertex AI TensorBoard to visualize ML experiments; and Vertex AI Pipelines to simplify the MLOps process by streamlining the building and running of ML pipelines.

USE CASE

Training and hyperparameter tuning

Build state-of-the-art ML models without code by using AutoML to determine the optimal model architecture for your image, tabular, text, or video-prediction task, or build custom models using Notebooks. Vertex AI Training offers fully managed training services, and Vertex AI Vizier provides optimized hyperparameters for maximum predictive accuracy.

USE CASE

Model serving

Vertex AI Prediction makes it easy to deploy models into production, for online serving via HTTP or batch prediction for bulk scoring. You can deploy custom models built on any framework (including TensorFlow, PyTorch, scikit or XGB) to Vertex AI Prediction, with built-in tooling to track your models' performance.

USE CASE

Model tuning and understanding

Get detailed model evaluation metrics and feature attributions, powered by Vertex Explainable AI. Vertex Explainable AI tells you how important each input feature is to your prediction. Available out of the box in AutoML Forecasting, Vertex AI Prediction, and Vertex AI Workbench.

USE CASE

Edge

Vertex AI Edge Manager (in experimental phase) is designed to facilitate seamless deployment and monitoring of edge inferences and automated processes with flexible APIs, to allow you to distribute AI across your private and public cloud infrastructure, on-premises data centers, and edge devices.

USE CASE

Model monitoring

Continuous monitoring offers easy and proactive monitoring of model performance over time for models deployed in the Vertex AI Prediction service. Continuous monitoring monitors signals for your model's predictive performance and alerts when the signals deviate, diagnose the cause of the deviation, and trigger model-retraining pipelines or collect relevant training data.

USE CASE

Model management

Vertex ML Metadata enables easier auditability and governance by automatically tracking inputs and outputs to all components in Vertex Pipelines for artifact, lineage, and execution tracking for your ML workflow. Track custom metadata directly from your code and query metadata using a Python SDK.

AutoML	Easily develop high-quality custom machine learning models without writing training routines. Powered by Google's state-of-the-art transfer learning and hyperparameter search technology.
Deep Learning VM Images	Instantiate a VM image containing the most popular AI frameworks on a Compute Engine instance without worrying about software compatibility.
Vertex AI Workbench	Vertex AI Workbench is the single environment for data scientists to complete all of their ML work, from experimentation, to deployment, to managing and monitoring models. It is a Jupyter-based fully managed, scalable, enterprise-ready compute infrastructure with security controls and user management capabilities.
Vertex AI Matching Engine	Massively scalable, low latency, and cost-efficient vector similarity matching service.
Vertex AI Data Labeling	Get highly accurate labels from human labelers for better machine learning models.

Vertex AI Deep Learning Containers	Quickly build and deploy models in a portable and consistent environment for all your AI applications.
Vertex AI Edge Manager	Seamlessly deploy and monitor edge inferences and automated processes with flexible APIs.
Vertex Explainable AI	Understand and build trust in your model predictions with robust, actionable explanations integrated into Vertex AI Prediction, AutoML Tables, and Vertex AI Workbench.
Vertex AI Feature Store	A fully managed rich feature repository for serving, sharing, and reusing ML features.
Vertex ML Metadata	Artifact, lineage, and execution tracking for ML workflows, with an easy-to-use Python SDK.
Vertex AI Model Monitoring	Automated alerts for data drift, concept drift, or other model performance incidents which may require supervision.

Vertex AI Neural Architecture Search	Build new model architectures targeting application-specific needs and optimize your existing model architectures for latency, memory, and power with this automated service powered by Google's leading AI research.
Vertex AI Pipelines	Build pipelines using TensorFlow Extended and Kubeflow Pipelines, and leverage Google Cloud's managed services to execute scalably and pay per use. Streamline your MLOps with detailed metadata tracking, continuous modeling, and triggered model retraining.
Vertex AI Prediction	Deploy models into production more easily with online serving via HTTP or batch prediction for bulk scoring. Vertex AI Prediction offers a unified framework to deploy custom models trained in TensorFlow, scikit or XGB, as well as BigQuery ML and AutoML models, and on a broad range of machine types and GPUs.
Vertex AI Tensorboard	This visualization and tracking tool for ML experimentation includes model graphs which display images, text, and audio data.

Vertex AI Training	Vertex AI Training provides a set of pre-built algorithms and allows users to bring their custom code to train models. A fully managed training service for users needing greater flexibility and customization or for users running training on-premises or another cloud environment.
Vertex AI Vizier	Optimized hyperparameters for maximum predictive accuracy.