**Module -1**

**Introduction to**

**Google Cloud Platform**



Sriram



Table of Contents

[1.1 Cloud Computing 3](#_Toc100414564)

[1.1.1 What is Cloud Computing? 3](#_Toc100414565)

[1.1.2 Growth of Cloud Computing 4](#_Toc100414566)

[1.1.3 Big Picture of Cloud Computing 4](#_Toc100414567)

[1.1.4 Advantages of Cloud Computing 7](#_Toc100414568)

[1.1.5 Cloud Providers 8](#_Toc100414569)

[1.1.6 Cloud Use Cases 9](#_Toc100414570)

[1.2 Getting Started with GCP 10](#_Toc100414571)

[1.2.1 Understanding GCP 10](#_Toc100414572)

[1.2.1 GCP Resource 18](#_Toc100414573)

[1.2.3 Resources scope 23](#_Toc100414574)

[1.2.4 Managing projects 25](#_Toc100414575)

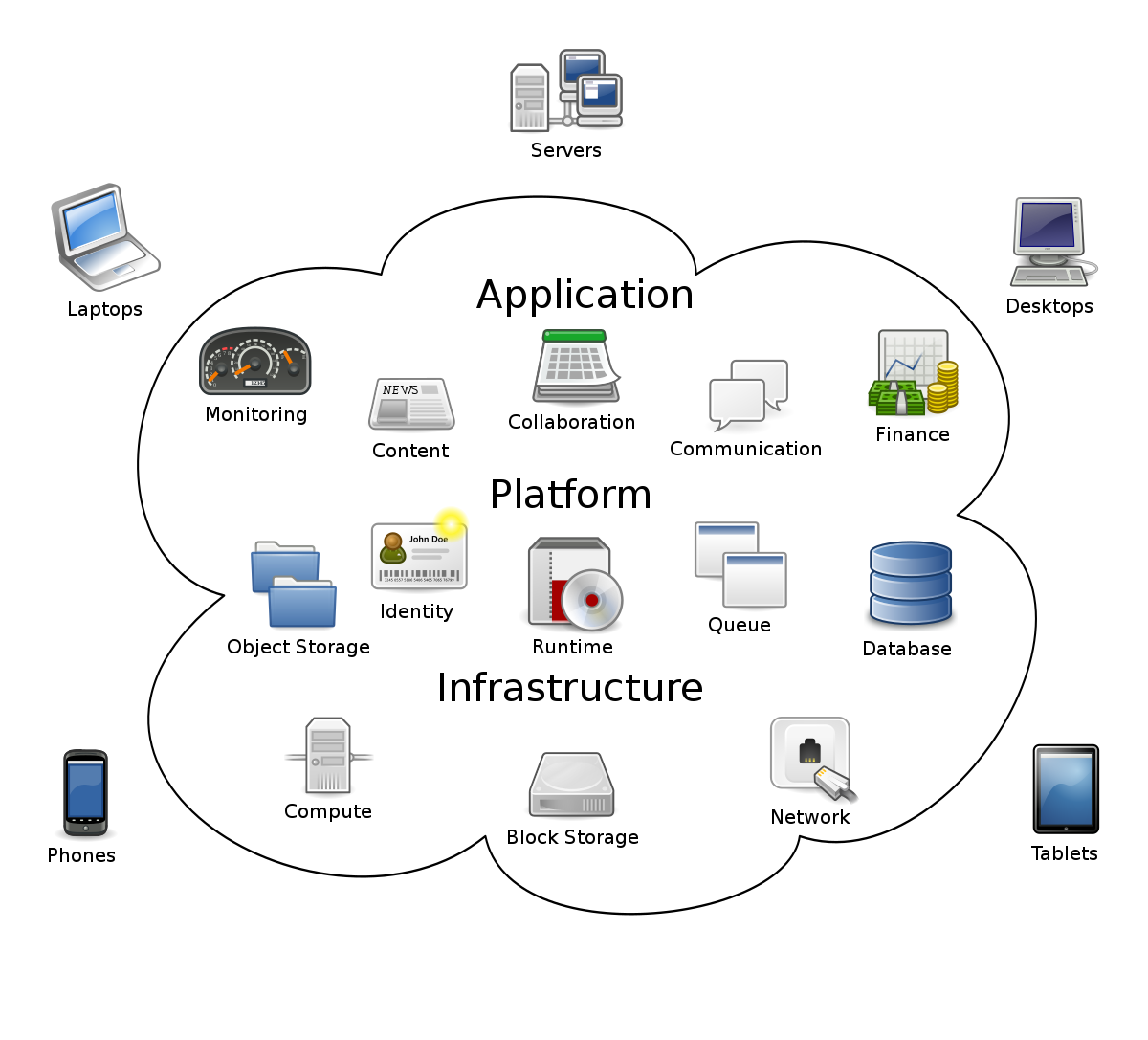
[1.2.5 Granting permissions 26](#_Toc100414576)

[1.2.6 Billing 27](#_Toc100414577)

# 1.1 Cloud Computing

## 1.1.1 What is Cloud Computing?

**Cloud** is a collection of computing resources which are accessed through the Internet and the software and databases that run on those servers



Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a

shared pool of configurable computing resources (for example, networks, servers, storage,

applications, and services) that can be rapidly provisioned and released with minimal management

effort or service provider interaction. This cloud model is composed of five essential characteristics,

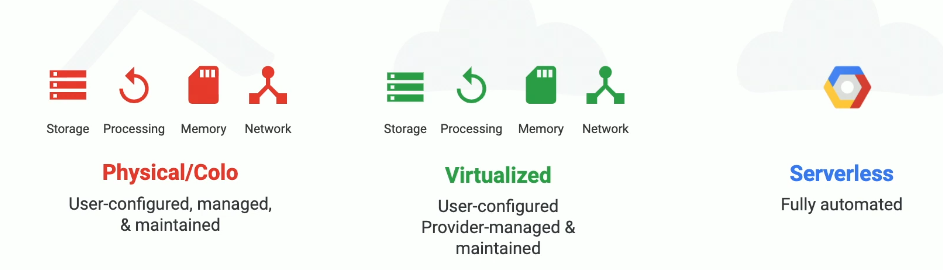
three service models, and four deployment models. – National Institute of Science and Technology

Companies offering these computing services are called **Cloud providers** and typically charge for cloud computing services based on usage, similar to how you are billed for water or electricity at home. E.g.: AWS, AZURE, GOOGLE CLOUD, OCI, IBM Bluemix

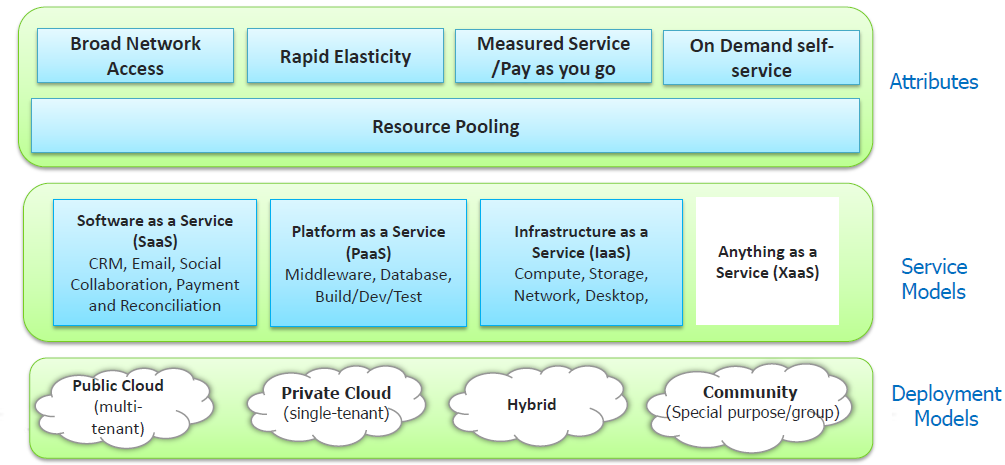
The primary reasons for moving to the cloud are: -

* You don’t need to maintain or administer any infrastructure
* It will never run out of capacity, since it is a virtually infinite
* You can access your cloud-based applications from anywhere, you just need a device which can connect to the internet

## 1.1.2 Growth of Cloud Computing



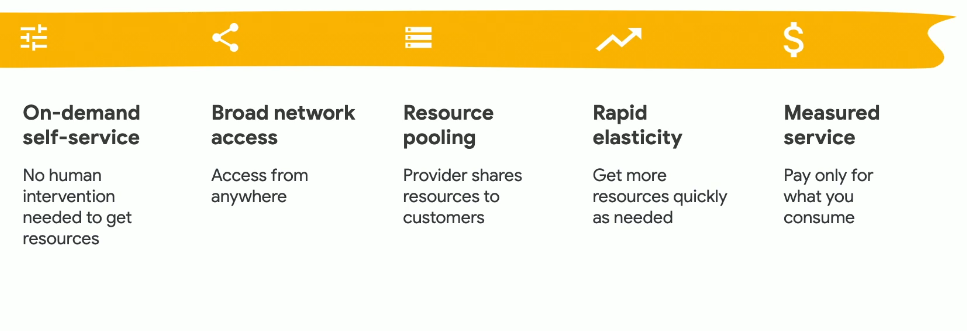
## 1.1.3 Big Picture of Cloud Computing



Five Attributes

The five essential attributes | characteristics of the cloud are as follows:

1. On-demand self-service: Services are provisioned automatically without manual provider intervention.
2. Broad network access: Resources are available through the network.
3. Resource pooling: Resources are pooled from a shared pool, giving the user a sense of location independence. For some of the resources, the location might be restricted.
4. Rapid elasticity: Services can be elastically provisioned and deprovisioned with capacity being managed by the provider.
5. Measured service: Resource usage is monitored and can be reported on.
6. When we look at GCP, it fulfills all of the five characteristics and fits into the public cloud deployment model.



When we look at GCP, it fulfills all of the five characteristics and fits into the public cloud deployment model.

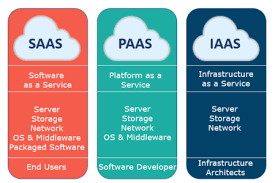
Four Deployment Models

The four deployment models are as follows:

1. Private cloud: Used by specific organizations, but can be managed by third parties
2. Public cloud: Used by the general public
3. Community cloud: Used by specific communities
4. Hybrid cloud: Composed of two or more different clouds



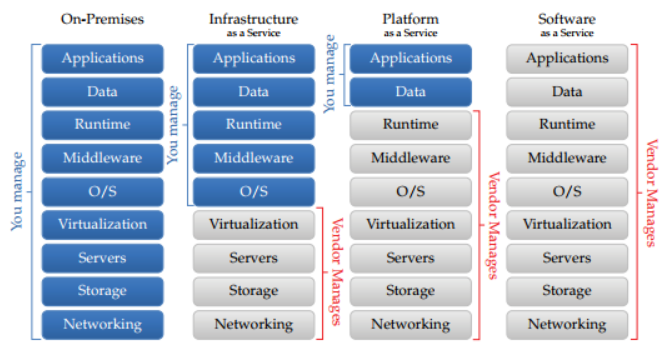
Three Serive Models



The three service models of cloud computing are: -

1. **Infrastructure as a Service** (**IaaS**) provides cloud infrastructure (Virtual Machine) in terms of hardware like memory, processor speed etc. Example: AWS, Azure, GCP
2. **Platform as a Service** (**PaaS**) provides cloud application platform (Programming Environment) for the developers. Example: Elastic Beanstalk, Google AppEngine, Heroku, Salesforce.com, Apache Stratos
3. **Software as a Service** (**SaaS**) provides cloud applications (Application Environment) which are used by the user directly without installing anything on the system. The application remains on the cloud and it can be saved and edited in there only. Example: Microsoft Office 365, CRM, Workday, NETSUITE, Athena Health Record Management, Intuit, Adobe Creative Cloud & Gmail

Comparison of On-Premise with Service Models



## 1.1.4 Advantages of Cloud Computing



* Pay-as-you-Go
* Increased Mobility, high availability, better Reliability, Scalability & Sustainability
* Focus on core business instead of spending time in IT Infrastructure
* Quick provisioning of services (Instead of building servers we will use services)
* Lower TCO (More Cost Effective) based on the utilization
* Less or No Capital Expenditure
* Easy to manage & Agile Deployment
* Brings high productivity with Less deployment time
* Auto or Dynamic Scaling based on the demand
* Shared Resources
* Device & Location Independent
* 24\*7 Support
* Environment Friendly

## 1.1.5 Cloud Providers

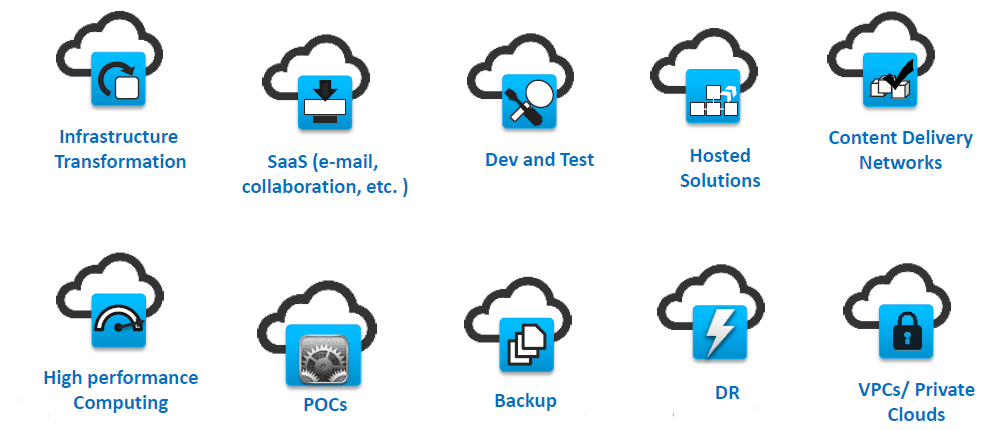
Public Cloud Providers

|  |  |  |
| --- | --- | --- |
| IaaS | SaaS | PaaS |
| * Amazon Web Services * MS Azure * Google Compute Engine * Alibaba Cloud * Rackspace * Digital ocean * Megha * TCS-Insta Compute | * SAP on AWS * OFFICE 365 * Maps * Facebook | * Elastic Beanstalk * Azure for .NET * developers. Google * Google app engine (GWT-RIA) * developers. Facebook * Cloud Foundry |

Open Source Cloud Providers

|  |  |
| --- | --- |
| **Open Source in IaaS** | **Open Source in PaaS** |

## 1.1.6 Cloud Use Cases

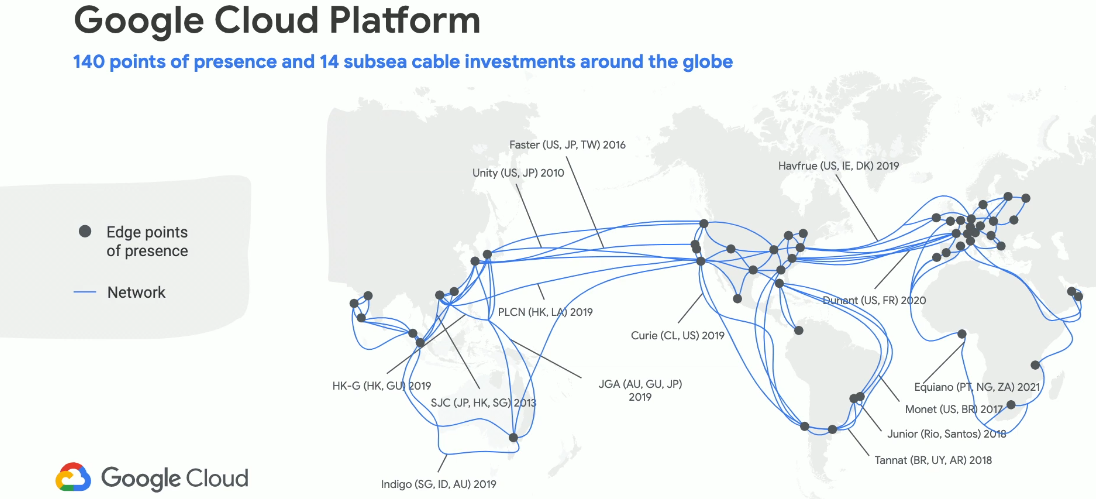


# 1.2 Getting Started with GCP

## 1.2.1 Understanding GCP

What is GCP?

* Google Cloud Platform, offered by Google, is a suite of cloud computing services that runs on the same infrastructure that Google uses internally for its end-user products, such as Google Search, Gmail, Google Drive & YouTube which is released in 2011
* Google Cloud Platform enables developers to build, test, and deploy applications on Google’s highly secure, reliable and scalable infrastructure.
* GCP is a collection of Google computing resources, which are offered via services, Provides many different cloud services such as Computing, Storage, Networking, Security, Big Data and Machine Learning
* Started as PaaS and now offers IaaS as well
* Support for various programming languages: Java, Python, Ruby, Go

****

What are the advantages of Google Cloud Platform?

* Provides Global Infrastructure, using private global fiber optic network
* Pay-as-you-Go Model
* Fully managed services & Fastest service provisioning
* Fastest and the best network among all the cloud providers
* GCP is Flexible, Scalable, Reliable Secured, highly scalable and provides high performance
* GCP is more cost-effective because it provides better pricing than other competitors with per second billing i.e. For more usage billing gets reduced
* GCP is highly secured, provides an excellent physical and digital security
* Supports Hybrid and Multi-cloud environment
* Open APIs compatibility with open-source services for Tensor Flow, Kubernetes & ForsetiSecurity
* Multi-vendor Friendly Technologies for Stack Driver, Bigdata, AI
* Support’s Server based & Serverless models
* Discount based on the usage
* Customized & Excellent Community Support

What are the differentiators in GCP?

* More Developer focused
* World largest and fastest network (40% of the world's internet traffic is carried by the)
* All Services are based on Global Scope not by regional
* Supports Bigdata & ML Services
* Pricing per second
* Agreed SLA
* Carbon Neutral
* Highly Secured (data is always encrypted with a choice of Google or customer-managed keys)

Why GCP?

* Very Fastest growing cloud computing platform on the planet
* Third Largest public cloud computing platform on the planet
* More and More organizations are outsourcing their IT to GCP
* The GCP certifications are the most popular IT certifications right now
* Top Paid certification according to Forbes
* GCP was named as a third leader in the “Iaas Magic Quadrant” – Gartner

What are the GCP locations?

GCP has a global footprint that includes North America, South America, Europe, Asia, and Australia. The locations are further split into regions and zones. our application should be located to provide low latency and high availability.



**Region**

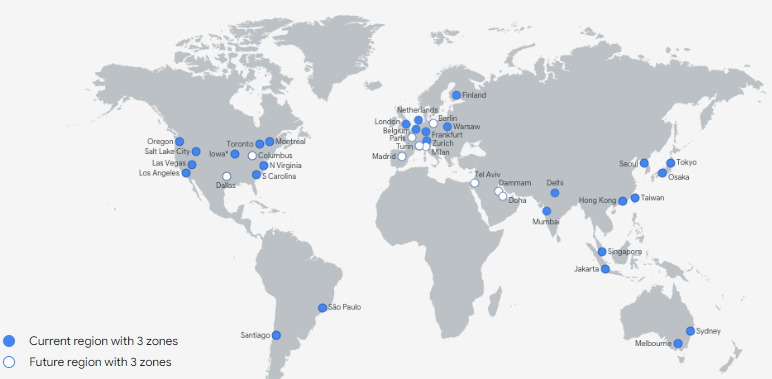
**A Region** is defined by Google as an independent geographic area that is divided into multiple zones. Locations within regions should have round-trip network latencies of under 1 ms in 95% of cases.

Regions are collections of zones. Zones have high-bandwidth, low-latency network connections to other zones in the same region. In order to deploy fault-tolerant applications that have high availability, Google recommends deploying applications across multiple zones and multiple regions. This helps protect against unexpected failures of components, up to and including a single zone or region. For example, the europe-west2 region denotes a region on the west coast of the Europe that has three zones: europe -west1-a, europe -west1-b, and europe -west1-c.

**Zone**

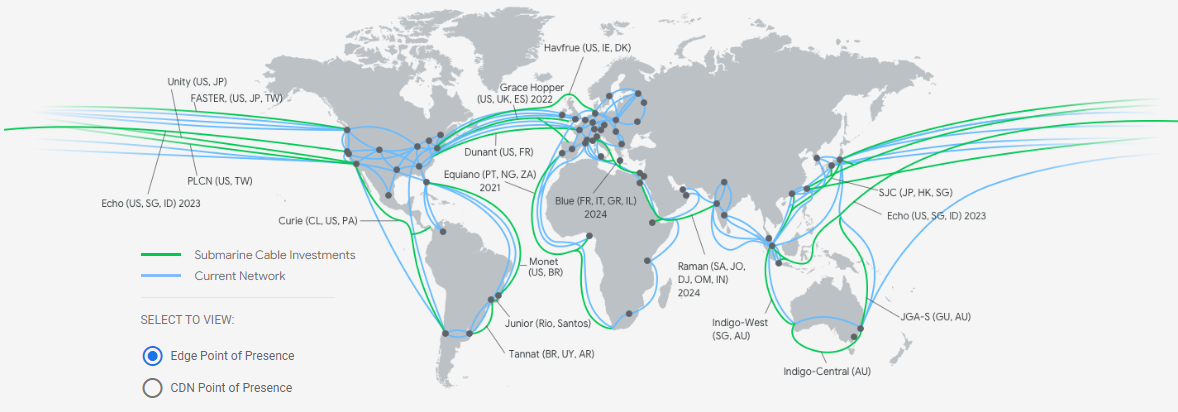
**A Zone** is a deployment area for GCP resources. Note that a zone does not correspond to a single data center; it can consist of multiple buildings. Even though a zone provides a certain amount of fault protections, a zone is considered a **single point of failure** (**SPOF**). Therefore, you should consider placing your application across multiple zones to provide fault tolerance.

A zone is a deployment area within a region. The fully-qualified name for a zone is made up of <region>-<zone>. For example, the fully qualified name for zone a in region us-central1 is us-central1-a. Depending on how widely you want to distribute your resources, create instances across multiple zones in multiple regions for redundancy.



**Network Edge Locations**

* Network edge locationsare connections to GCP services located in a particular metropolitan area.
* Edge Locations are endpoints for GCP which are used for caching content. Typically, this consists of Google Content Delivery Network (CDN). There are many more edge locations than regions. Currently there are over 100 edge locations. Based on the nearest edge location, customers can communicate and get data.
* Locations built to deliver cached data across the world. CDN utilizes this service for faster delivery to countries without AWS regions.
* Edge Location is an intermediate between the end users and servers to access the services from GCP.
* Edge Location is a small setup in different location to provide low latency connection by caching static content. Basically, it’s a Content Delivery Network and used with Google CDN.

****

**Edge Cache**

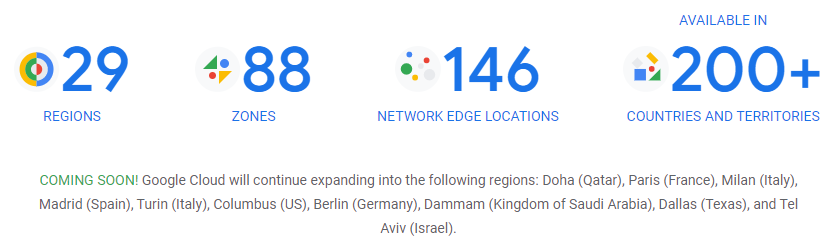
Edge Cache is used to store my frequently accessed data in the server

**Latency**

* Latency is the measure of time required for transfer of data from client to server and again back to client
* More the latency lower is the efficiency. Low latency, higher efficiency
* GCP has its clients throughout the world, so in order to reduce latency and avoid load on servers they make use of Edge Locations

**To know the latest information about GCP Location - Regions and Zones**

**Locations:** <https://cloud.google.com/about/locations>



**Regions & Zones :** <https://cloud.google.com/compute/docs/regions-zones>

|  |  |  |
| --- | --- | --- |
| Region | Zones | Location |
| asia-east1 | a, b, c | Changhua County, Taiwan |
| asia-east2 | a, b, c | Hong Kong |
| asia-northeast1 | a, b, c | Tokyo, Japan |
| asia-northeast2 | a, b, c | Osaka, Japan |
| asia-northeast3 | a, b, c | Seoul, South Korea |
| asia-south1 | a, b, c | Mumbai, India |
| asia-southeast1 | a, b, c | Jurong West, Singapore |
| australia-southeast1 | a, b, c | Sydney, Australia |
| europe-north1 | a, b, c | Hamina, Finland |
| europe-west1 | b, c, d | St. Ghislain, Belgium |
| europe-west2 | a, b, c | London, England, UK |
| europe-west3 | a, b, c | Frankfurt, Germany |
| europe-west4 | a, b, c | Eemshaven, Netherlands |
| europe-west6 | a, b, c | Zürich, Switzerland |
| northamerica-northeast1 | a, b, c | Montréal, Québec, Canada |
| southamerica-east1 | a, b, c | Osasco (São Paulo), Brazil |
| us-central1 | a, b, c, f | Council Bluffs, Iowa, USA |
| us-east1 | b, c, d | Moncks Corner, South Carolina, USA |
| us-east4 | a, b, c | Ashburn, Northern Virginia, USA |
| us-west1 | a, b, c | The Dalles, Oregon, USA |
| us-west2 | a, b, c | Los Angeles, California, USA |
| us-west3 | a, b, c | Salt Lake City, Utah, USA |

How GCP is committed to environmental responsibility?

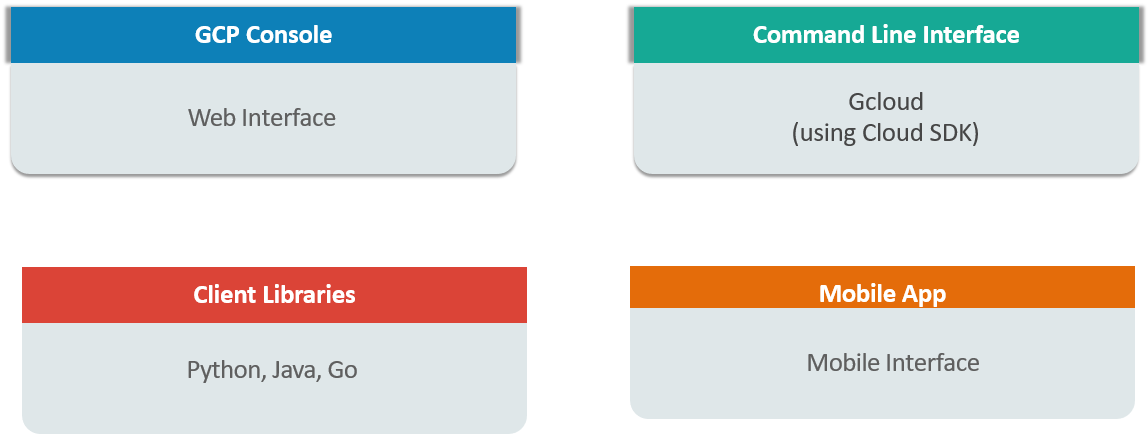
****

*First Carbon Neutral Cloud is GCP*

What are the several ways of accessing GCP?

There are 4 ways to interact with GCP which includes: -

* GCP Console (Web Interface)
* Command Line Interface – Gcloud (using cloud SDK) , Work with gcloud , gsutil & Bq
* Client Libraries – Python, Java, Go, Rest APIs -> To build the Custom Applications
* Cloud Console Mobile App – For iOS & Android Mobile Interface

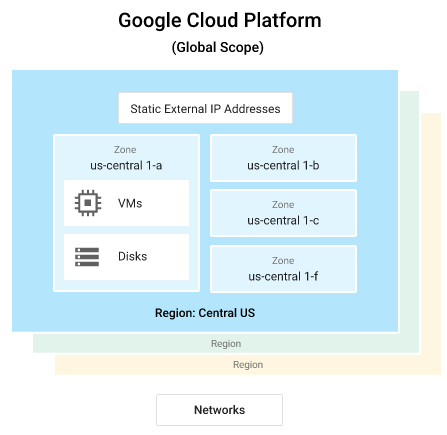


## 1.2.1 GCP Resource

GCP consists of a set of physical assets, such as computers and hard disk drives, and virtual resources, such as virtual machines (VMs), that are contained in Google's data centers around the globe referred to as **GCP Resources**.

The GCP resource hierarchy is organized as follows: -

* All resources (VMs, Storage Buckets, etc) are organized into **Projects**.
* These projects may be organized into folders, which can contain other **Folders**.
* All folders and projects can be brought together under an **organization** node.
* Project folders and organization nodes are where policies can be defined
* Policies are inherited downstream and dictate who can access what resources
* Every resource must belong to a project and every must have a billing account associate with it



**Resource hierarchy levels define the trust boundaries**

* Group your resources according to your **organization structure**
* Levels of the hierarchy provide trust boundaries and resource isolation

Organizations

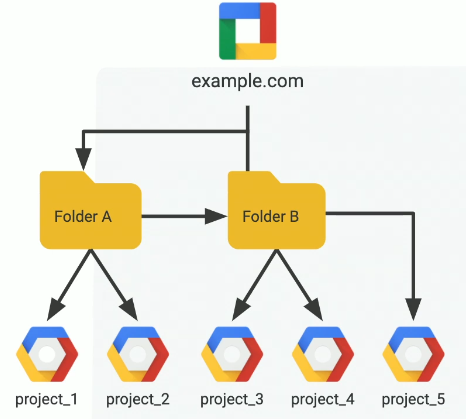
GCP consists of containers such as organizations, folders, and projects to hierarchically group your resources. This allows you to manage their configuration and access control. The resources can be managed programmatically using APIs. Google also provides tools such as Google Cloud Console and command-line utilities, which are wrappers around the API calls. Let's now have a look at the hierarchy presented in the following diagram and familiarize ourselves with each of the resources:



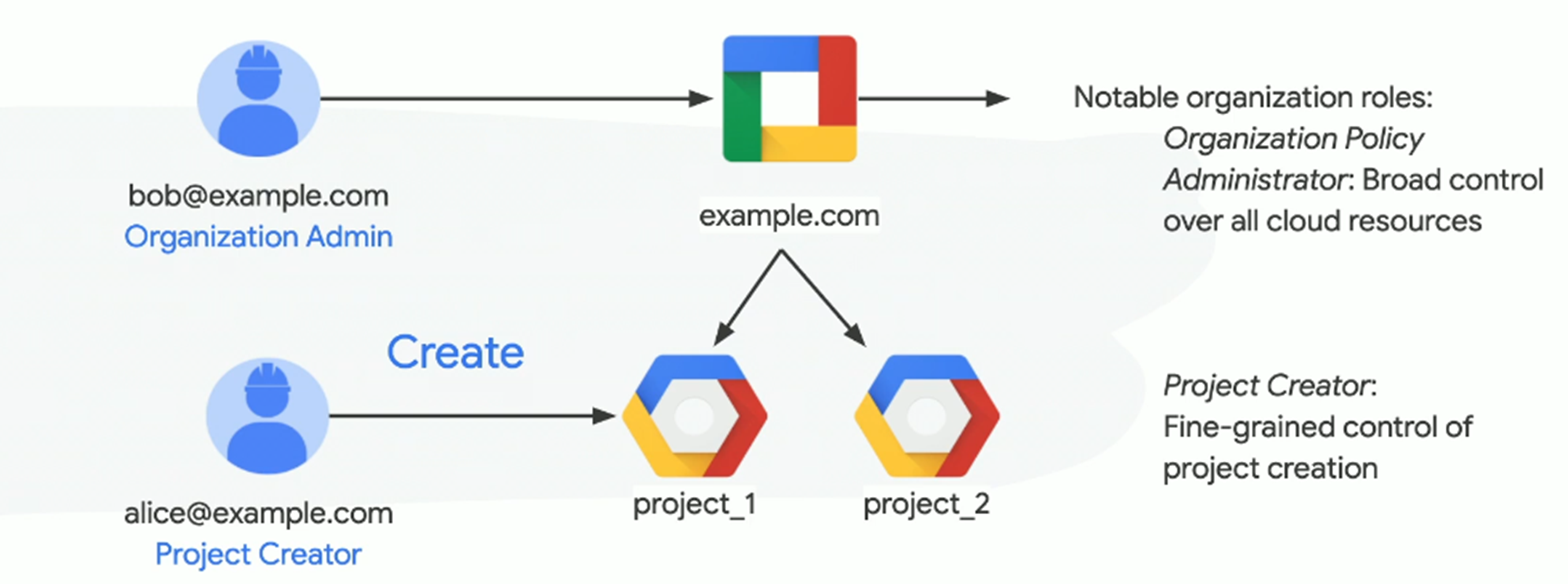
From the above resource manager hierarchy diagram, we have an Organization that can be mapped to a company. Next, we have Folders that can represent a company's departments. Next, we have Projects, which further divide the actual company projects or environments, such as development and production. Finally, under Projects, we have GCP Resources.

Folders (Offers flexible management)

* Folders group projects under an organization
* Folders can contain projects, other folders, or both Use folders to assign policies
* Folders are logical containers that can group projects or other folders. They can be used to assign IAM policies. Again, the use of folders is optional and is available only when an organization resource exists. The use case for using folders is to group projects that will use the same IAM policies.

****

Organization Node organize into Projects

****

GCP Resources with Project Association

In GCP everything is a resource. GCP services that you manage are associated with the project

In GCP Project, we can manage resources such as

* Track Resource and Quota Usage
* Enable Billing
* Managing Permissions and Credentials
* Enable Services & API
* Enable Different Devices



Projects

Projects are the smallest logical containers that group resources. Every resource within GCP needs to belong to exactly one project. Each project is managed separately, and IAM roles can be assigned per project to control the access in a fine-grained way.

All GCP services are associated with the Project. The Project attributes are:-

* **Project ID:** Should be Worldwide unique while creation and cannot be changed
* **Project Name**: Need not be unique, can be changed anytime
* **Project Number:** Worldwide unique, assigned by Google Cloud Platform and cannot be changed

Using the project ID to identify your project. To manage resources within GCP, you will always need to identify which project they belong to by either project ID or project number. You can create multiple projects, but there is a quota that limits the number of projects per account. If you reach the quota, you will need to submit a request to extend it.

Project Roles and Permissions

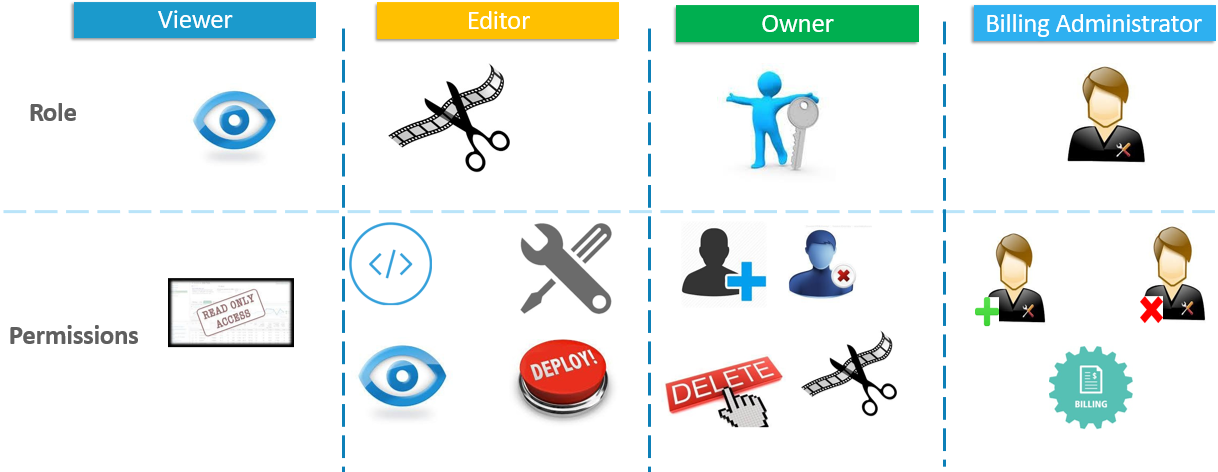
* A project can have multiple owners, editors, viewers and billing administrators
* Provide least permission to the user based on the project requirement
* Different Roles & Permissions are: -

**Viewer:** Read only access

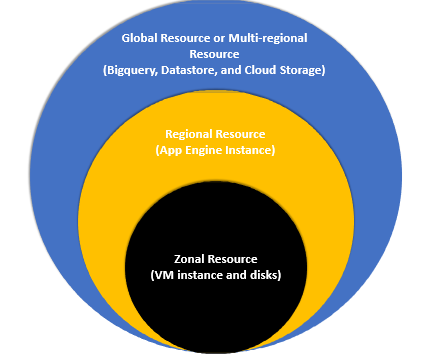
**Editor:** Deploy Applications, Modify Code, Configure Service

**Owner:** Invite Members, Remove Members, Delete Projects

**Billing Administrator:** Manage Billing, Add or Remove Admins



## 1.2.3 Resources scope



Global resources

Global resources are globally available within the same project and can be accessed from any zone.

These include the following objects:-

* **Addresses**: These are reserved external IP addresses and can be used by global load balancers.
* **Images**: These are either predefined or user customized. They can be used for provisioning VMs.
* **Snapshots**: Snapshots of a persistent disk allow the creation of new disks and VMs. Note that you can also expose a snapshot to a different project.
* **Instance templates**: These can be used for the creation of managed instance groups.
* **Virtual Private Cloud (VPC) networks**: These are virtual networks that you can connect your workloads to.
* **Firewall**: These are, in fact, defined per VPC, but are accessible globally.
* **Routes**: Routes allow you to direct your network traffic and are assigned to VPCs, but are also considered global.

Regional resources

Regional resources are accessible by other resources only within the same region. These include the following objects:-

**Addresses**: Static, external IP addresses can only be used by instances that are in the same region.

**Subnets**: These are associated with VPC networks and allow the assignment of IP addresses to VMs.

**Regional managed instance groups**: These allow you to scale groups of instances. The scope can be set to either regions or zones.

**Regional persistent disks**: These provide replicated, persistent storage to VM instances. They can also be shared between projects for the creation of snapshots and images, but not disk attachments.

Zonal resources

Zonal resources are only accessible by other resources within the same zone.

These include the following objects:-

**VM instances**: These reside in a particular zone.

**Zonal persistent disks**: These provide persistent storage to VM instances. They can also be shared as disks between projects for the creation of snapshots and images, but not disk attachments.

**Machine types**: These define the hardware configuration for your VM instances and are defined for any particular zone.

**Zonal managed instance groups**: These allow you to autoscale groups of instances. The scope can be set to either regions or zones.

## 1.2.4 Managing projects

**To create a new project, please follow the below steps:-**

1. Log in to the GCP console at https://console.cloud.google.com and click on the drop-down arrow next to the name of the project you are currently logged into. A Select a project window will pop up. Click on NEW PROJECT in the top-right corner.

2. Fill in the name and choose the billing account. You can attach the project to an organization or a folder. Choose the default billing account. In the following steps, we will show you how to create a new billing account and associate it with the project we are now creating. Click on the CREATE button, as shown in the following screenshot:

3. The new project has been created. You can now manage it from the GCP console:

4. To start using the GCP services, click on the hamburger icon. A menu will pop up. You can access all GCP services from here, as shown in the following screenshot:

## 1.2.5 Granting permissions

For assigning permissions to GCP Resources, first add a member and assign previously-defined roles to them. Roles are basically sets of permissions.

The following are the step-by-step instructions to grant permission:

1. To add a new member to your project, go to the IAM section of the IAM & admin pane.

2. Select the MEMBERS tab and click on ADD. Now, select a member and choose a role. Click on Save to confirm.

3. The user has been added and has the permissions as per the defined role.

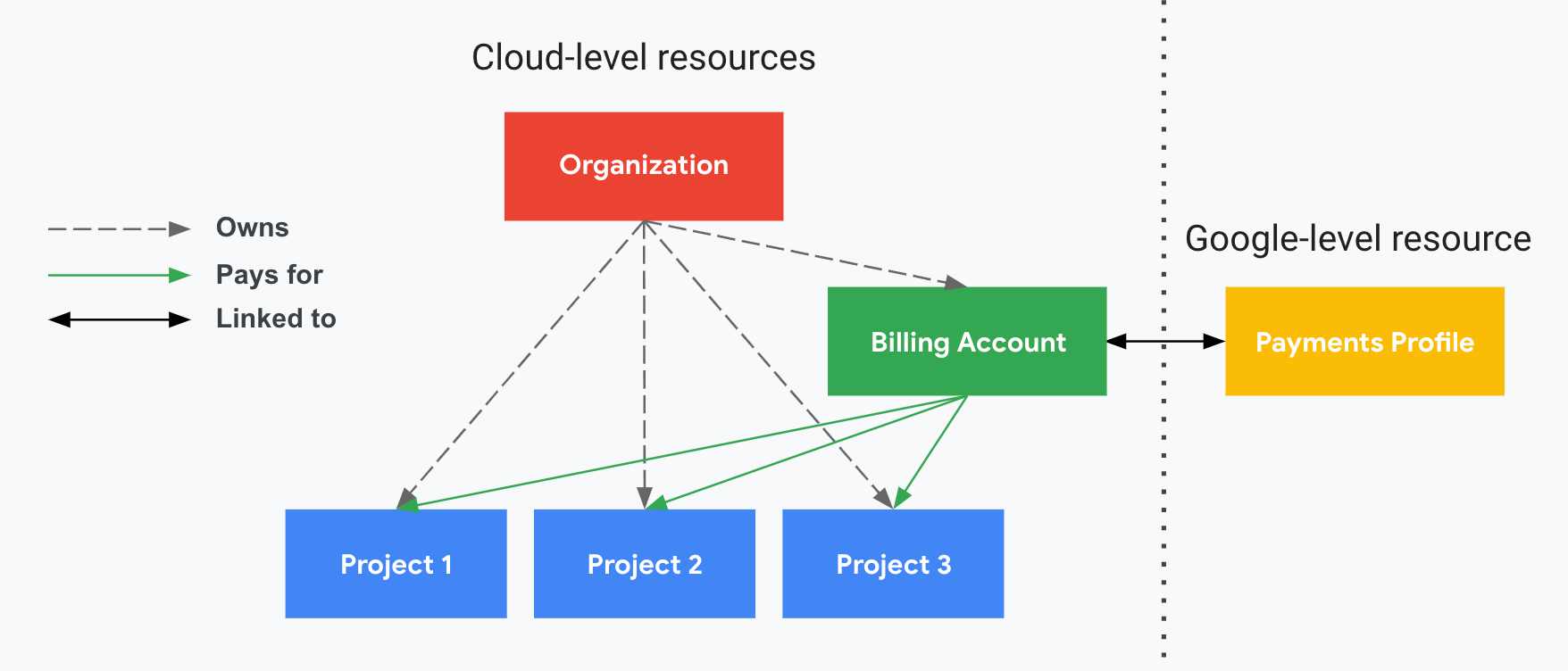
Another User been sent an invitation to join the project as an Owner. The triangle with an exclamation mark will be displayed until the invitation is accepted.

## 1.2.6 Billing

Depending on the company structure, we might have different requirements as regards billing. With GCP, you have the possibility of creating a single or multiple billing accounts. As shown in the following

diagram, the billing accounts can be associated with one or more projects. The actual payment details are created in the payment profiles that are attached to the billing account.

The following diagram shows the relationship of ownership and payment linkages for an organization.



Where

* **Ownership** refers to IAM permission inheritance
* **Payment linkages** define which Cloud Billing account pays for a given project

**GCP Pricing Model**

* Per Second Billing
* Free monthly quota for the respective services like GAE, Bigquery, …
* Usage based Charging for data storage, processing,
* Preemptible VM instances: Up to 75% discount
* Committed-use discounts: save up to 60%
* Sustained-use discounts: Up to 30%
* Custom machine types: save up to 50%
* Coldline: The archival storage with the speed of disk at the cost of tape
* All ingress traffic is free of cost. Egress traffic is charged

Managing billing accounts

To create a new billing account, follow these steps:-

1. Go to the GCP console and choose Billing from the left navigation pane

2. You will be presented with the existing billing accounts. Click on Create **Account**

3. In the next window, name your billing **Account**

4. Choose the **country and the currency** will be presented for you. Click on the **Confirm** button

5. Now, you can choose an existing payment profile or create a **new one**.

Note that we do not see any existing profiles. This is because my existing payment profiles are set for Polish PLN, while the new billing profile is set for USD. Fill in the customer information and scroll down to Payment method.

6. Fill in your payment details and click on the **Submit and enable** billing button

7. Now, your **billing account has been created** and you can manage it from the Billing window

Now, we can assign a project to the newly created billing account.

Assigning a project to a billing account

You can assign multiple projects to one billing account. In the following screenshot, you can see that we now have three billing accounts and multiple projects assigned to them. Our newly created billing account has no project assigned to it, so let's therefore move our GCP Cloud Architect project to that billing account, as follows:-

1. First, we need to click on **My billing accounts**, as this is where the **project** is attached

2. This brings us to the Overview page. Now, we click on the **three-dots icon next to the Project name** and choose **Change billing**

3. In the next window, we choose the billing account we want the project to be attached to and click on **SET ACCOUNT**

The project is now attached to its proper billing account.

Exporting billing

GCP allows you to export the billing information to a **BigQuery dataset** or a file in **a Cloud Storage bucket**. This can be useful for preparing reports and an analysis of the cost of your cloud consumption.

To perform the export, follow these steps:-

1. Go to Billing and choose Billing export. You will be presented with both options. Select the type of export, fill in the information about the dataset or bucket, and click on **SAVE**

2. After the data is exported to **BigQuery**, you can perform queries on it. As an example, you can check which service has generated the most costs

This information is very useful for creating all sorts of billing reports. On top of the reporting, we would also like to be informed upfront whether we are exceeding our budget. Let's have a look at how this can be done in the next section.

Budgets and alerts

Budgets and alerts can be set for each billing account or project. You can set up a specific threshold. Once the amount spent is higher than the defined threshold, billing administrators and billing account users will be notified.

This will not stop the usage of any services, and charges will continue to apply for the running resources. By default, there are three alert thresholds: 50%, 90%, and 100%. Both the number of thresholds and their values can be modified.

Billing account roles

The following list shows the roles that can be used to control the billing:

* **Billing Account Creator:** This is used for the initial billing setup,including signing up for GCP with a credit card.
* **Billing Account Administrator:** This is the owner of a billing account. This role is allowed to link and unlink projects and manage other users' roles for the billing account. This role can manage payment instruments, billing exports, and view cost information.
* **Billing Account User:** In combination with the project creator role, the Billing Account User role is allowed to create new projects linked to the billing account on which the role is granted.
* **Billing Account Viewer:** This role allows access to view the billing information. It can be used by the finance team.
* **Project Billing Manager:** This role enables the attachment of the project to a billing account without rights to resources.

