# Index

[**Index**](#_vtsopnq8wpcz) **1**

[**Introduction to Cloud Computing**](#_lgpd4txpy7iz) **3**

[Understanding Cloud Concepts](#_bk0jv1xfshr5) 3

[What is Cloud Computing?.](#_9dvuhsz2u10f) 3

[Economics and Scalability](#_ezzs7l663fbs) 3

[Economy of Scale](#_244hfwliy2kw) 3

[Benefits of Cloud](#_o36wvk3rmu1f) 3

[Capex vs Opex](#_ghennrolgtd8) 3

[Scalability](#_b2971xyoucn6) 4

[Cloud Service Models](#_jw80ywg28bah) 4

[**Google Cloud Overview**](#_gq2og2fq7vkp) **6**

[Google Resources: Region/Zone Impact](#_p8prj2pr4xwn) 6

[Zonal Resources](#_dd7rbsm44hdj) 6

[Regional Resources](#_i4vo9h2q0wdt) 6

[Multi-Regional Resources](#_heowxshdvvvg) 6

[Google Services](#_7jw108uifzyw) 6

[Compute Services](#_ukkgc38mg1m8) 6

[Networking](#_5bedq5qlqpfh) 7

[Storage](#_ey8qmjvajb58) 7

[Identity and Security Services](#_i1w26966mhil) 7

[Big Data Services:](#_rzexy6rcqfgy) 8

[Machine Learning Services:](#_eludzw3ur5cz) 8

[Management Tools](#_5wqz4xf20ayf) 8

[Development Tools](#_6p9phrg8e9u0) 9

[URL For the Cloud Console](#_gkdhakwygtwo) 9

[Accessing GCP](#_ya5772r8yw8d) 9

[Cloud Platform Console](#_vkkuwcu2ytd8) 9

[Cloud Shell and SDK](#_h7hwj9ntvg7r) 10

[Cloud Console Mobile App](#_tg2cl6f6xnv) 10

[REST API](#_so4o1tlx1fx1) 10

[Install the Google Cloud SDK](#_8j0jcegcwsxe) 10

[Types of Cloud Services](#_sp7m5y6zo7vz) 11

[Core Cloud Services](#_6no99s8qz56o) 11

[Compute VirtualMachines](#_9bikvw1wfnq6) 11

[Storage Services](#_9fsekx3mxkq9) 12

[Networking Services :](#_fxkfnkejm0ze) 12

[Application Service:](#_1uiyb8u1tris) 12

[Other Specialised Services:](#_yw1ogn5ynnrt) 12

[**Project Accounts and Billing**](#_rifts8ha0q9m) **13**

[Google Cloud project Hierarchy Example](#_xdti7dcyetee) 13

[Project Identifiers](#_dvfgvwt9f9v7) 14

[Project-ID](#_bkfckgqct0gp) 15

[Project Name](#_8q6r56rnavwh) 15

[Project Number](#_ta6oswdbebt2) 15

[Reference Links](#_4alwd3ucg89r) 15

[Billing account](#_g69o3x1h8lv2) 15

[Steps to Create a Google Account](#_q3mjhw3w947m) 16

[Steps to Create a Organisation Google Account](#_pf12ko8ru2m5) 16

[Google Organizations](#_uvclw447n6oj) 16

[Create Google Project](#_tisnfzutntv3) 16

[Budgets and Alerts](#_tuzb5i73gpoi) 17

[**Identity and Access Management**](#_eooult3s7mby) **19**

[Who can do what on which resource](#_6r4wwg1qbuf5) 19

[Roles](#_e9twtp75t6h5) 19

[Role Inheritance](#_nfs9vflsn9et) 19

[Role Types](#_52j33dn3pa6k) 19

[Primitive Roles](#_6uso6nugdoh) 19

[Pre-Defined Roles (job specific)](#_4uon66js88lp) 20

[Custom Roles (specific roles and services we tailor make for individual people)](#_r4ywapxnqbll) 20

[Organisation Administrator](#_cbgug1tdc25m) 21

[Role Administrator](#_eglvcwwrqo6y) 21

[Assigning Roles](#_2hm58hp464ox) 22

[Create Service Account](#_y4eg3hl3s7wf) 22

# 

# Introduction to Cloud Computing

## Understanding Cloud Concepts

The transition has happened from a traditional datacenter running an operating system to running multiple virtual machine OS and Applications on the datacenter. Now it has evolved to cloud Azure, GCP, AWS

### What is Cloud Computing?.

1. It allows on demand self service without any human intervention needed to get resources
2. Broad Network Access: Access from anywhere (If I want them to be)
3. Resource Pooling: Provider shares resources to customers
4. Rapid Elasticity : Get more resources quickly, as needed.(auto scaling etc)
5. Measured Service : Pay only for what you consume (transparency)

It's about renting resources vs purchasing hardware.Cloud providers are responsible for the physical hardware and facilities necessary to execute your work like providing updates and they can choose to no longer provide a service.

### Economics and Scalability

#### Economy of Scale

Ability to do things more efficiently or at a lower-cost per unit when operating at a larger scale.

#### Benefits of Cloud

1. Pass on economies of sale to consumers
2. Acquire hardware at lower cost
3. Local Government Deals : to employ people locally, and lower electricity for data centers.

#### Capex vs Opex

Capital Expenditure(CapEx)

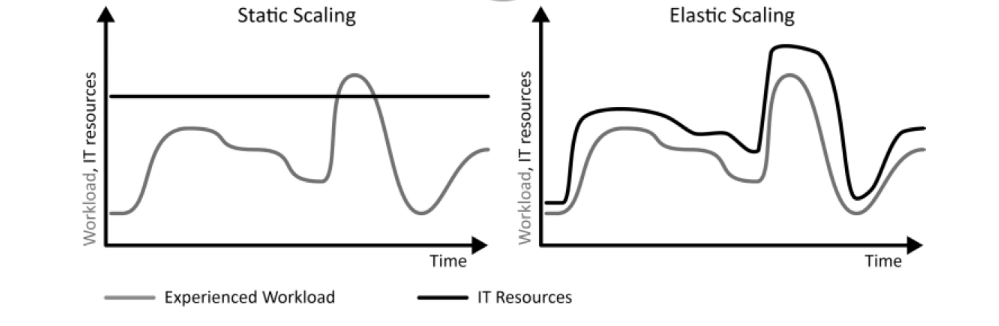
* Spending on Infrastructure is completed upfront
* Cost written off over a period of time
* Costs are Server Costs, Storage Cost, Network Costs, Backup and Archive Costs, Data Center Costs
* Benefits are predictability, cost effective when you consume the infrastructure quickly

Operational Expenditure(OpEx)

* No up-front cost
* Pay for service as you consume it
* Deduct from tax bill in same year as expense occurs
* Costs are Server Lease Costs Software and feature leases, usage/Demand cost scaling.
* Benefits are : Try and buy, Low initial costs, Demand Fluctuation.

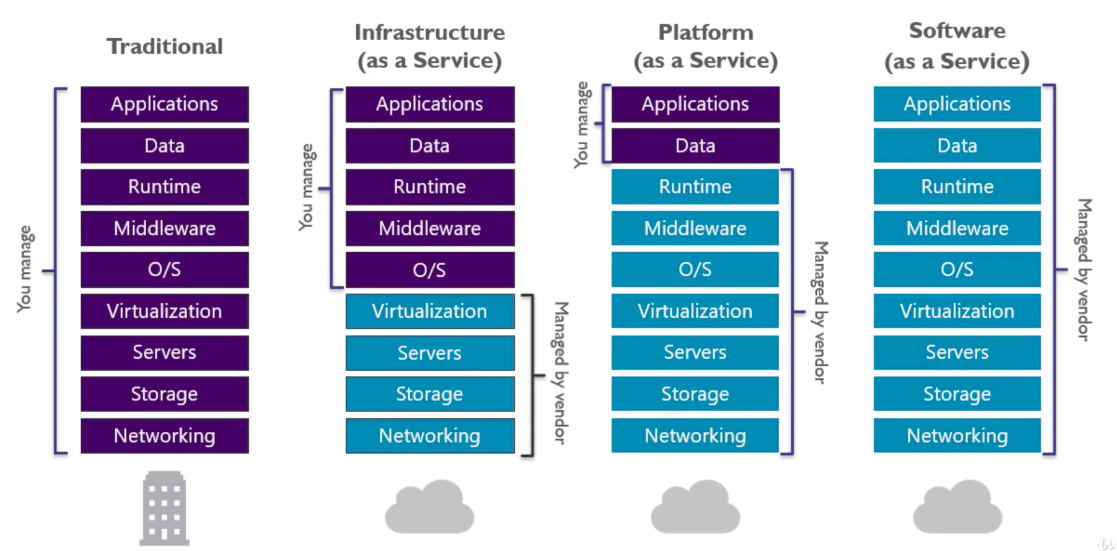
#### Scalability

* Increase or decrease resources based on workload demand,
* Vertical Scaling
  + Also known as scaling up
  + Add additional resources to increase the power of the workload
  + Eg: Add additional CPUs to a Virtual Machine
  + The machine itself is expanded
* Horizontal Scaling : Also known as scaling out , instead of making the machine bigger we will be adding more nodes.



* Elasticity: Major pattern which benefits from cloud computing. As your workload changes, resources can be changes to compensate(up or down) like seasonal demands

### Cloud Service Models



# Google Cloud Overview

Every company is a *data company* according to Google.

Google has a massive network thanks to their massive search infrastructure which they expanded for GCP

<https://cloud.google.com/about>

Check this URL for the services and regions and features from this URL

## Google Resources: Region/Zone Impact

Regions are independant geographical locations, they consist of zones(data center in that region).Locations tend to have very low network latency and inside those locations zone is a deployment area for GCP resources in that region. Also a fault domain, for fault tolerant apps use 2 zones. So in case one goes down the other one can take over.

### Zonal Resources

Operate in one zone, if that zone becomes unavailable all Zonal Resources will be unavailable until the server comes back up. Eg: Google Compute Engine

### Regional Resources

Deployed with additional redundancy in that region. Higher availability relative to zonal resources

### Multi-Regional Resources

Few of the services have multi-region capability like storage services in the event of a disaster we can use the data from another region.

## Google Services

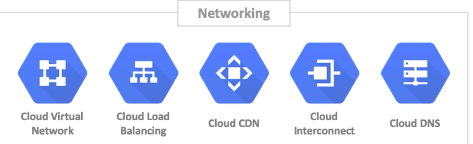
### Compute Services

Virtual Machines, containers etc. Virtual machines to run our applications.



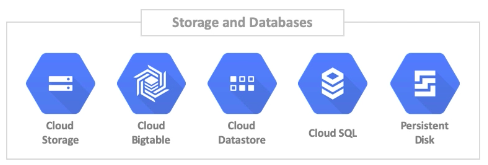
### Networking

Core infrastructure, cloud virtual networks load balances



### Storage

For Storage needs



### Identity and Security Services

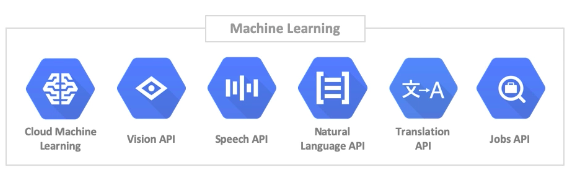
Security, role based access control



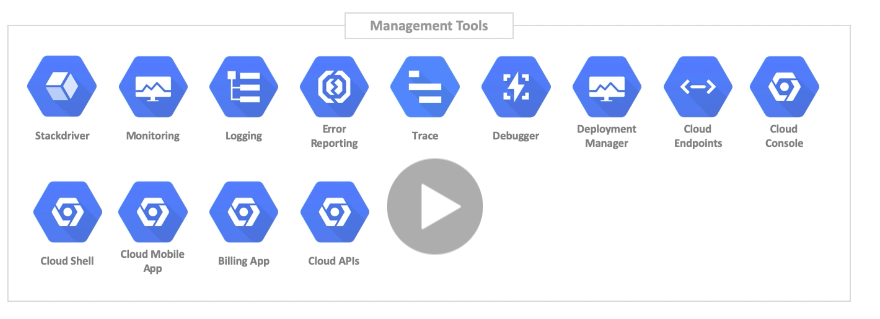
### Big Data Services:



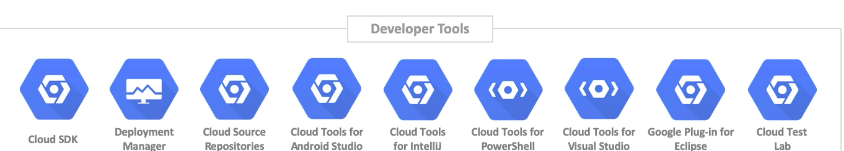
### Machine Learning Services:



### Management Tools



### Development Tools



## URL For the Cloud Console

<https://console.cloud.google.com>

We can go here to go and consume the services

Search for resources and product, there is also a filter options on the right of the screen

Create a virtual machine using Compute Engine, VM Instance. Couple of seconds to whip up a new one.

## Accessing GCP

### Cloud Platform Console

<https://console.cloud.google.com>

Accessed through the browse, this is the place we will be most often, looking around ,verifying this you did in the environment through the command line, nice graphical way to view the interface

### Cloud Shell and SDK

A command line utility to speed up the day to day. Command line shell in the console or install it locally

### Cloud Console Mobile App

View environment, verify, monitoring alert

### REST API

Build your own application. We use REST API calls to GCP to build something new.

## Install the Google Cloud SDK

Download it on the laptop or activate in the Cloud Platform Console

Go to cloud.google.com/sdk/docs

Google Cloud SDK is a set of tools that can be used to manage tools and applications hosted on GCP including gcloud, gsutil, and bq.

Download, authenticate and we can choose the project and choose the zone.

We can also do this using the Google Cloud Shell. In the cloud console.

## Types of Cloud Services

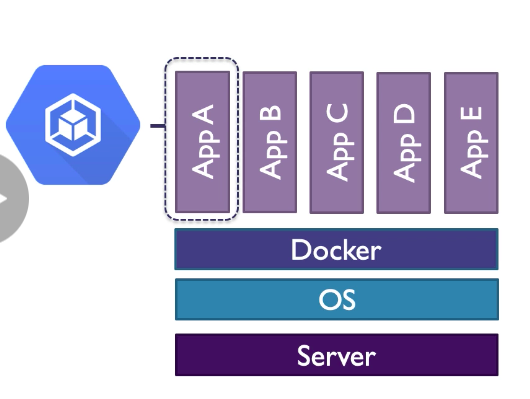
### Core Cloud Services

#### 

#### Compute VirtualMachines

Are things that are going to run our workloads like containers

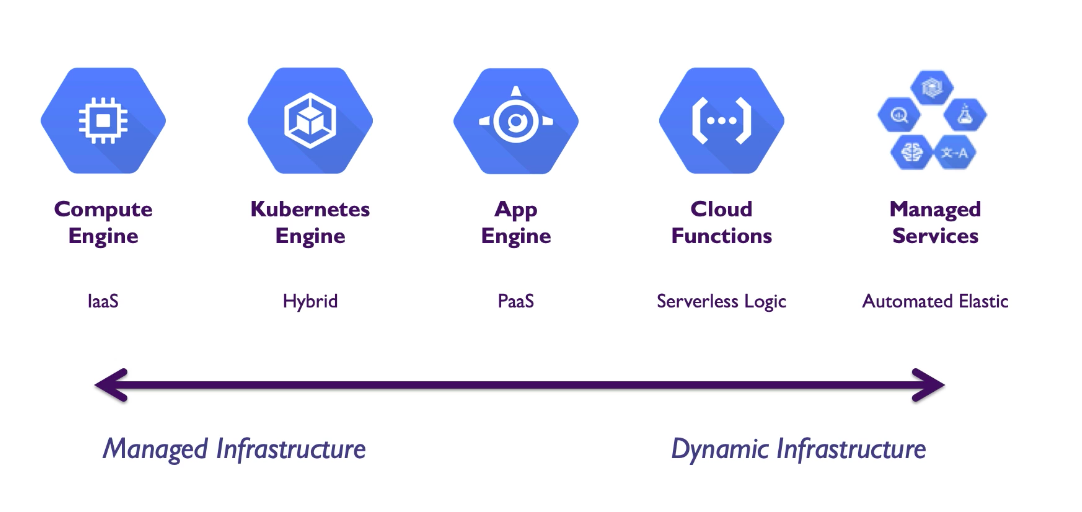
* Compute Engine : The App OS container. It is one OS instance
* Container Engine : Standardized packaging for software and dependencies. A way to isolate apps from each other. Works with Linux and Windows servers.



* Container Register: They allow you to separate apps to share the OS Kernel and we use the container registry to store approved containers that we can then run
* Serverless: We can run code in an environment that does not require setting up VMs or Kubernetes clusters

GCP has 2 serverless options

* App Engine: Apps and containers that need to run for extended periods of times and they must conform to specific run times may have to run Python Stack or Java Stack
* Cloud Functions: On demand code runs for short processes



#### Storage Services

For Storage devices

1. Object Storage: Provide object storage, not stored in a conventional file system, Grouped into buckets and build security around them, Each URL is addressable by a URL.
2. File Storage: Network File System (NFS), Suitable for applications that require operating system type access to files. File system and directories are decoupled from VM No need to set up a machine with NFS share, . We can set this NFS service up and make it available to our virtual machines.
3. Block Storage: Fixed-size data structures that are organised in blocks. Attached to VMs in GCP, Persistent or ephemeral . Independent life cycle to that of VM. Support file system level access vs HTTP access to object storage.

#### Networking Services :

To communicate Network services must be defined.

#### Application Service:

App Engine and other marketplace services can run on top of compute

#### Other Specialised Services:

Speech API

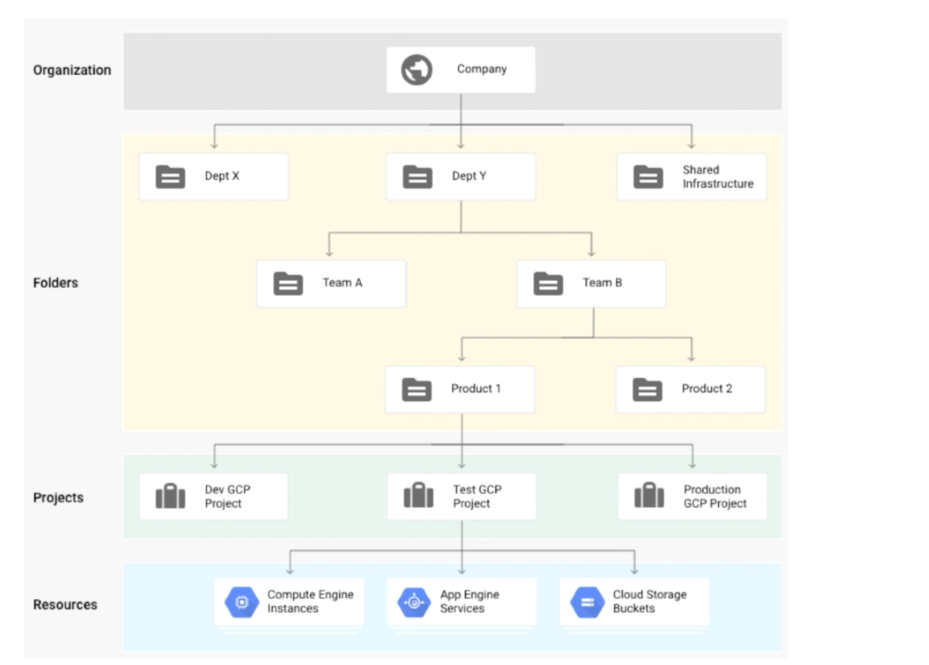
# 

# Project Accounts and Billing

Google Projects form the basis for creating , enabling and using ALL GCP services.

This includes managing APIs, enabling billing, adding/removing collaborators and managing permissions for GCP resources identities.

## Google Cloud project Hierarchy Example



.

All GCP resources are associated with a Project

* Organisation(Deloitte, or it could be a personal account)
* Folders (Various folders contain project or other folders or combination of both)

*To create folders you must have the Folder Admin or Folder Creator role at the parent level. If it is the first folder you need it at the organisational level.*

* Projects (Projects can be set up with organisation or without organisation)



Problem Associated with this is that when the employee account is deleted the project is irretrievably deleted.



With employees separately assigned to the project if another employee comes in, we can add them. The project must be migrated to the organisation itself.

## Project Identifiers

### Project-ID

This identifier should be globally unique and can be chosen by the user. GCP provides suggestions.

### Project Name

This identifier need not be unique, and it is chosen by you

### Project Number

This is globally unique and is assigned by GCP.

## Reference Links

|  |  |
| --- | --- |
| **Target** | **Link** |
| G Suite Admin Console (organization) | admin.google.com |
| G Suite Access | gsuite.google.com |
| Google Account Management | myaccount.google.com |
| Google Cloud Home | clooud.google.com |
| GCP Console | console.cloud.google.com |

## Billing account

A cloud Billing account defines who pays for a given set of GCP resources and it can be linked to one or more GCP projects.

Your project usage is charged to the linked Cloud Billing Account.

## Steps to Create a Google Account

1. Got to cloud.google.com , Go to the right top User Profile and click ‘Add Account’
2. You’re redirected
3. Enter your email and stuff or choose ‘’Create Account” , choose from ‘From myself’, ‘To manage my business’ and click Next
4. Fill the form
5. Enter phone number, backup email .
6. Create
7. Then Login
8. Then go to cloud.google.com and click lets ‘get started for free.’
9. Add the billing account. Select country , then type and details
10. Then tour around the console Top Right Profiles, feedback, cloud shell
11. The left side, the services and API’s

## Steps to Create a Organisation Google Account

1. Go to gsuite.google.com/signup
2. Fill the Form
3. Add the domain,
4. Enter business information
5. After it is set up
6. Admin.google.com to see Gsuite and the associated tools.

## Google Organizations

The differences between personal and organisation are different. You can create an organisational unit.

IAM -> Identity and Organisation -> Choose Organisation -> We can see the requests and we can migrate requests.

When creating a new project we can put the projects in the demo.

## Create Google Project

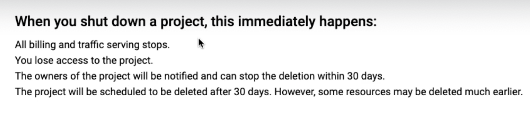
Go to the admin console, go to IAM & Admin, click manager resources.

This will show any projects that we already have that are in your account, and you can create folders here to organise resources here.

Create Project -> Project Name -> (Change project ID now if needed else it cannot be changed. -> Organisation -> Change location if needed) -> Create.

The project is spun up.

To delete a project



And then click shutdown.

We can choose it and restore it to restore it to.

Restored project, we need to go in a power on the VM and re-setup billing.

## Budgets and Alerts

Go to Google Console, Go to the project and Billing will be there on the right side and we can click detailed charges, we can also go to the left and then Billing. This brings up all the details of the billing account. Overview I can see the primary. And go to manage my billing accounts.

On clicking my Projects we can see the billing accounts that they are associated with the actions we can choose to disable and change billing.

If 1+ billing accounts are available you can change it here.

Bottom left -

-> Payment methods The cards are there

-> Account Management we can management account, like close the billing account.

Reporting

-> Cost Breakdown to see what you are billed for.

->Cost Table breaks down all the services I use, change month

Protection

-> Budgets and Alert

Create Budget

-> Name, Assign to projects, choose services in project, click next and set target amount based on a fixed month spend or last month spend. Also a chance to include credits in that cost.

Then we get the alert threshold. Percent, amount, Trigger we can send email or PUB/Sub notification system that we may have.

Select the alert at the bottom and click delete to delete that budget.

-> Billing Export

We can go there and export the billing as csv

->Commitment.

If we’re committing to resources like if we know we have resources for the next 2-3 years.Reservations in AWS and Azure. This is where we analyse that

# Identity and Access Management

## Who can do what on which resource

Who: Google Account, cloud identities, service account, Google Group

## Roles

A role is a set of permissions which can be assigned to users by assigning that role to the user

Least privileges should be given.

## Role Inheritance

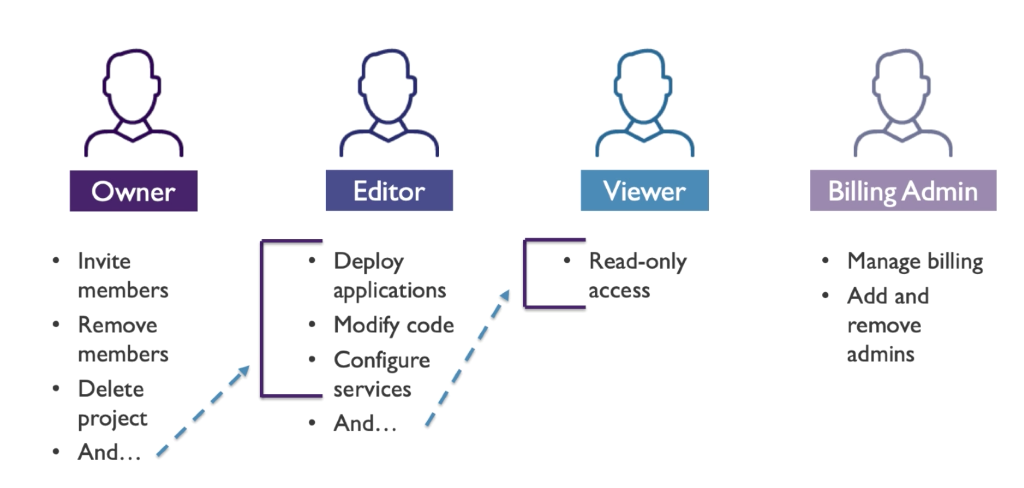
We can apply it at

* Organisation level
* Folder level
* Project level
* Resources Level

## Role Types

### Primitive Roles

(loosely defined generic role apply at a high level) (owner, editor, viewer)



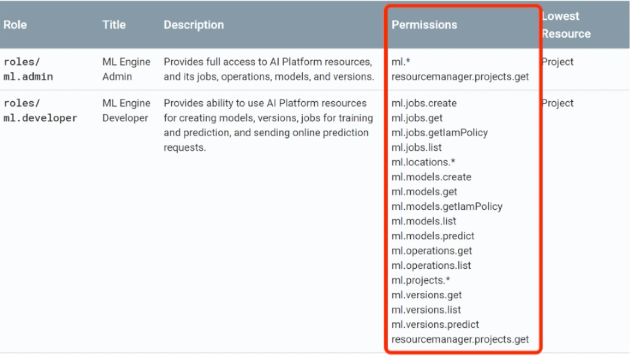
Owner can do the roles of Owner + Editor

Editor roles are Editor + Viewer

Viewer can only read code

### Pre-Defined Roles (job specific)

You can provide granular access to specific services that are managed by GCP, These roles are created by Default in GCP



The specific permissions make up the role

### Custom Roles (specific roles and services we tailor make for individual people)

Custom Roles are created by combining one or more of the cloud IAM permissions.

Permissions allow users to perform platform specific actions on the Google Cloud Platform Resources.

Permissions are represented in the form

<service>.<resource>.<verb>

Compute.instance.list

Compute.instance.stop

## Organisation Administrator

If you have an organisation associated with your GCP account, the Organisation Role Administrator allows you to administer all custom roles in your organisation.

This role is granted only on the organisation level

Only Organisation Administrator can grant a Organisation Administrator Role



## Role Administrator

Administer all customers roles for your organisation

Use this role if we don't have an organisation

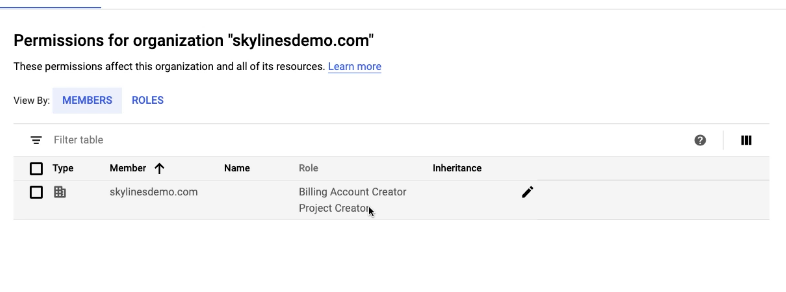
Role can be granted only by project or organisation owners



## Assigning Roles

Go to google cloud console -> Left Side -> IAM & Admin -> Identity and Organisation -> Set Permissions

You can go



CLick Add+ button on the top add member details, -> Select Roles , Resources -> Folder Admins for Folder privileges

We can remove the roles here as well.

## Create Service Account

Add roles for services account creation

Service accounts -> Service Accounts Admin

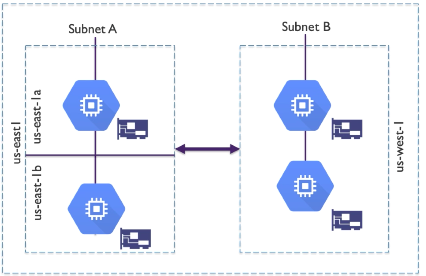
IAM -> Security Admin

Select Project -> IAM -> Service Accounts -> Service Accounts Details -> Give it Roles -> Grant users access to this service account.

# Networking

## VPC

Software version of physical networks that link resources in a project.



1. VPC’s are global
2. Subnets are Zonal and can extend across zones in the same region
3. Subnets are expandable (not shrink)
4. Firewall rules can be applied across zones

# Subnet Address Space:

Reserved Address :

***Example Subnet - 10.10.0.0/16***

***Subnet Address - 10.10.0.0***

***Gateway Address - 10.10.0.1***

***2nd to last Address - 10.10.255.244***

***Broadcast Address - 10.10.255.255***

These addresses are reserved and cannot be used.

Avoid large subnets to reduce chance of CIDR range collisions, better to start small and expand

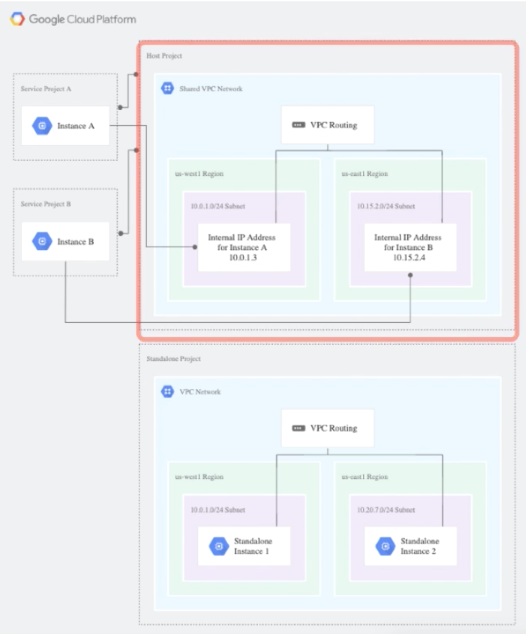
GCP Subnets are expandable(auto expandable from /20 to /16)

Check out http://www.subnet-calculator.com

## VPC Features

1. Transparent : They are transparent (Option to turn on Flow Logs)
2. Routes: You can modify the routes
3. Hybrid Connectivity: You can connect your GCP network to your on premises data center
4. Packet Mirroring: Security team, sometimes use it to turn it on and stuff.

## Shared VPC



The Host project utilises a shared VPC network rather than a VPC network used below. Then when creating a service projects (service project A, Service project b) they consume the subnets of the host network

Shared VPCs allow an organisation to connect projects to a common VPC network (Hub)

Shared VPCs contain shared networks

Eligible resources from service projects can use subnets in the shared VPC network

## Create a VPC and Subnet

1. Create a new project -> A default VPC is created on creating a new project
2. Click the Hamburger Icon
3. Scroll down to networking section -> VPC networks
4. Default VPC is automatically added
5. On top change ‘MODE’ from Auto, Switch to Custom Mode (cannot undo this action)
6. On top click ‘+CREATE VPC NETWORK’
7. Enter a name, Enter a description, Create choose between Automatic (to create a subnet in each region) or manual
8. In manual, enter the name, REgions, IP address Range, Private Google access(Private VMs can access Google services without assigning external IP addresses), Flow Logs (doesn’t affect performance, but some systems generate a large amount of logs, which can increase cost in StackDriver) Then click Done
9. You can also Add Subnets (you need to add a new subnet range)
10. Dynamic Routing can be turned off or on.
    1. Regional (Cloud Routers will only learn routes in the region where they are created)
    2. Global (dynamically learned routes to and from all regions with a single VPN or interconnect and Cloud Router). DO NOT USE THIS if there is an internal load balancer with a dedicated interconnect or a VPN on this network
11. DNS server policy: You can create a new server policy, Give it a name , you can use Google’s DNS or use alternate DNS servers. Enter the IP address and enter enable.
12. Then click create.

You can click the subnet and edit it. Increase the IP address range, increase form /16 to /8 maybe

Click VPC to add the new subnets, click edit DNS server policy or the dynamic routing mode.

VPC network details, things like static IP addresses are visible, or add a reserved static internal IP addresses, we can add a static IP address like a dedicated active directory server.

Firewall Rules: It can be added

Routes: Default routes can be added

VPC Network Peering: Peer this VPC with a Peer

Private Service connection: Add the private service connection.

## IP Addressing and DNS

### DNS is GCP

We have the option between a Google-provided DNS and Customer DNS Server.

Google-provided DNS

We’re basically consuming the DNS service provided by GCP. If you’re starting out you will be fine with this

Customer DNS Server

IaaS with DNS : VM where we install DNS, AD-DNS

Infoblox Virtual Appliance (on Site)

### Internal IP Address

We have an option of an Internal IP Address that we control, Allocated from subnet via DHCP (Dynamic Host Configuration Protocol). We also have the option of statically allocating IP.

*DHCP lease default is 24 hours.*

VM name is registered automatically to internal DNS

2 types on internal IP Address **Static** or **Ephemeral**

**Static :** Assigned until explicitly released (They remain attached to stopped instances until removed). Select if you want the VM to keep its IP addresses. Eg active\_directory

**Ephemeral:** Deleting and re-creating instances will change the IP, Stopping or restarting will not

1. May choose automatic or custom when selecting ephemeral options.

### Internal DNS Naming

[INSTANCE\_NAME].[ZONE].c.[PROJECT\_ID].internal

[INSTANCE\_NAME] : is the name of the instance

[ZONE]: is the zone where your instance is located

[PROJECT\_ID]: is the project to which the instance belongs.

### External IP Addresses

Can be assigned to instances or a forwarding rule

Needed to communicate to the internet, resources outside of VPC.

Only resources with an external IP can send and receive traffic directly to and from outside the network.

### Static and Ephemeral External Addresses

#### Static External

Long term assignment

Remains attached to stopped compute instances until they are explicitly removed

Regional and Global options (Global reserved for global forward rules used by global load balancers)

#### Ephemeral External

Remain in place only until the instances are stopped, restarted or instance is terminated.

New IPs are assigned when instances are started again.

There is a limited amount of external IP, so we need it to be reclaimed.

### Hybrid Connectivity

On-Premises to GCP Connectivity

Cloud VPN

IPsec VPN Connection , Tunnel between GCP and ON premises

Requires Cloud VPN Gateway service (Regional)

Cloud Router required for *BGP(Border Gateway Protocol)*

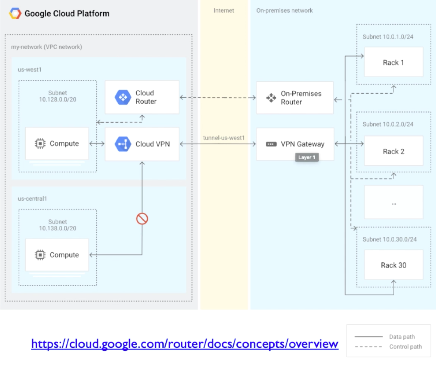
Cloud Interconnect (Direct connectivity, without using Internet)

Dedicated or partner-connected service upto 50 Gbps

### Static vs Dynamic Routing

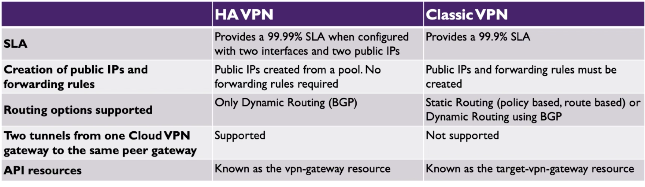
With static routes, you must create or maintain a **routing table.** A topology change on either network requires you to **manually update the static routes**

With **Cloud Router,** you can use **BGP** to exchange routing information between networks. Instead of manually configuring static routes, networks automatically and rapidly discover topology changes through BGP. This method of exchanging routes through BGP is **dynamic routing.**

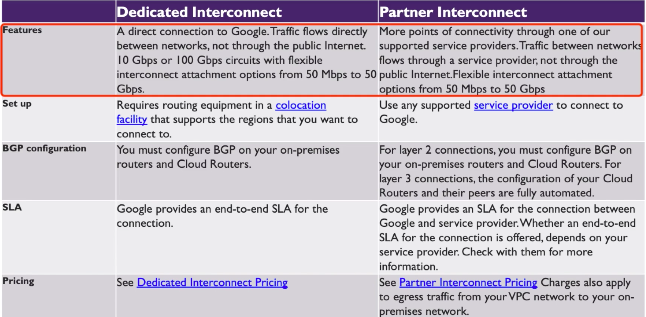
****

The on-premise router contacts the Cloud Router and are able to understand the network topologies.

VPN Comparison



### Cloud Interconnect



SLA - SERVICE LEVEL CONNECTIVITY

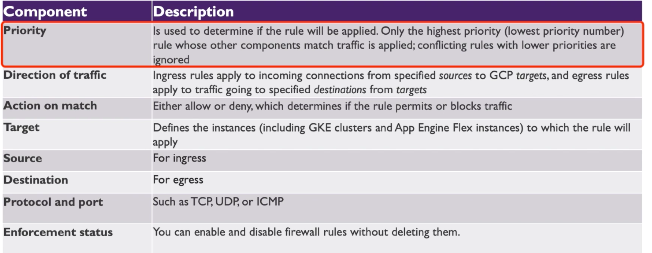
### Firewall Rules

Every VPC network functions as a distributed firewall

Firewall rules are defined at the network level and connections are allowed or denied on a per-instance basis

GCP Firewall rules exist not only between your instances and other networks but also between individual instances within the same network. Eg PORT denial etc

### Firewall Rule Components



### Rule Priority

Rules are enforced based on priority

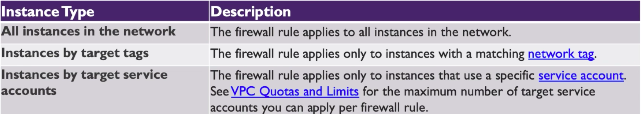
Range from 0 to 65535 (Default : 1000)

Lower numbers have higher priority

### Firewall Rule Creation

1. Specify network priority of Rule
2. Choose direction(ingress or egress)
3. Choose action (allow or deny)
4. Specify target of the rule (all instances, target tags, service account)
5. Source or destination filter
6. Define protocols and ports

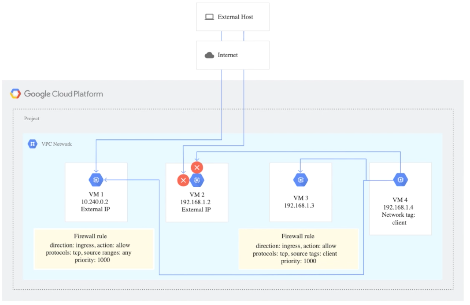
### Target



Default Rules in Default Network

*These rules are only included in the default network, but you can create your own rules that allow these types of traffic in your other networks*

1. Default-allow-internal : Allow internal traffic
2. Default-allow-ssh : Allow ssh connections (port 22)
3. Default-allow-rdp : Allow RDP connections from windows server
4. Default-allow-icmp : You can do pings on your machines if you want to as well



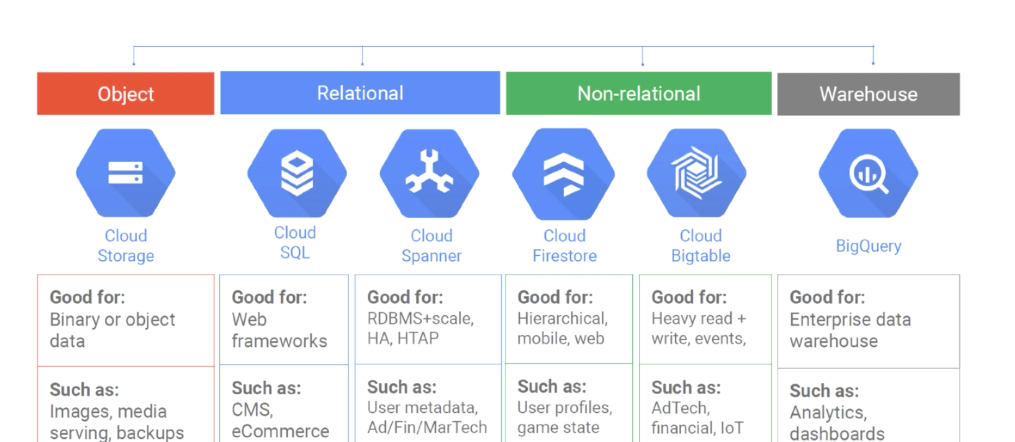
## 

# GCP Storage

Types of data that are stored in GCP include

1. Structured Data :
   1. Adheres to a schema
   2. All data has same field or properties
   3. Stored in a database tables with rows and columns (eg SQL )
   4. Relies on Keys to indicate how one rows in a table, relates to data in another row of another table
   5. Referred to as Relational Data
2. Semi-Structured Data : NoSQL
   1. Doesn’t fit neatly into table, rows or columns
   2. Use tags and keys to organise and provide the hierarchy for the data
   3. Often referred to as NoSQL or non-relational data
3. Unstructured Data :
   1. No structure whatsoever
   2. No restrictions on the data it can hold.
   3. Example, a object can store a PDF, JSON, etc
   4. Enterprises are struggling to manage and tap into the insights from their unstructured data (Big Query), (Data Warehouses)

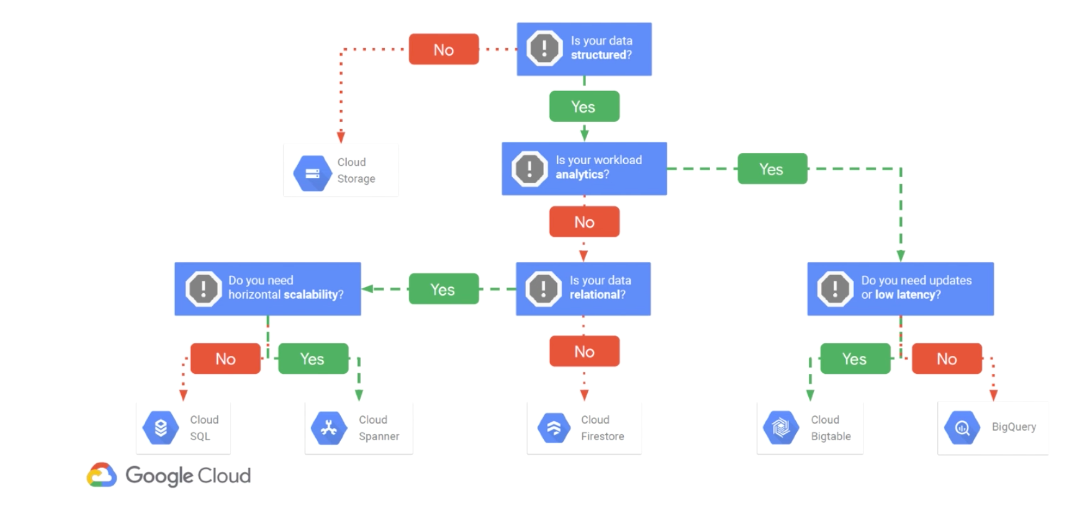
## Storage and DB Services



Cloud Storage is a low object store like throwing stuff in a bucket.

Relational : CloudSpanner is a hyperscale Relational Data Base

## Storage Decisions



## Cloud Storage Bucket Access

We can use the cloud storage using the below methods.

Cloud IAM: Legacy Roles (ACLs)

Access Control Lists

Signed URLs

## Demo:

Hamburger Menu -> Storage -> Create Bucket ->

**Name bucket**

Choose where to store the data(Region lowest latency), multi-region(can’t choose), Dual-region(pairs)

// costs are visible in the right

Choose default storage class (Standard, Nearline(backups), Coldline(archives))

**Access Control**

Fine Grained(Specify access to Individual objects by using object level permissions (ACL) in addition to your bucket-level permissions)

Uniform (Ensure Uniform access to all objects in a bucket by using only bucket level permissions)

**Retention Policy**

Requester Pays: On

The Object Permissions can be edited on the bucket level

# 

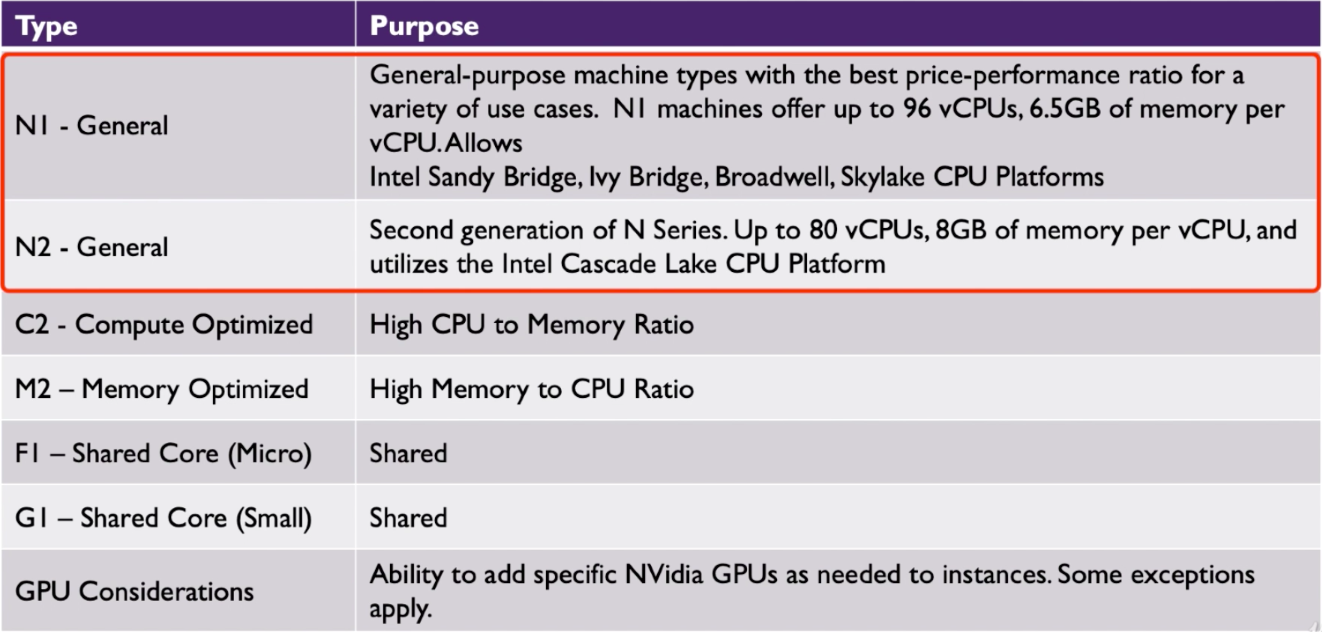
# Google Compute Engine (Virtual Machines)

## Machine Type

A Machine Type is a set of virtualized hardware resources available to a virtual machine instance, including the system memory size, Virtual CPU (vCPU) count and persistent disk limits.

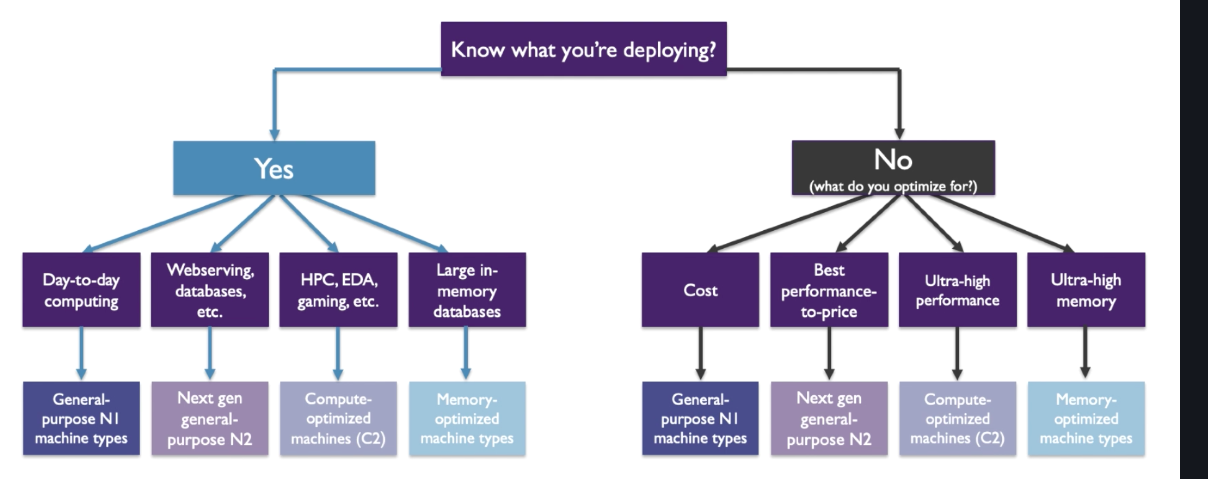
Machine types are grouped for different workloads and come in a variety of families. Within the families you can fine tune the specifications.

## Machine Type Families

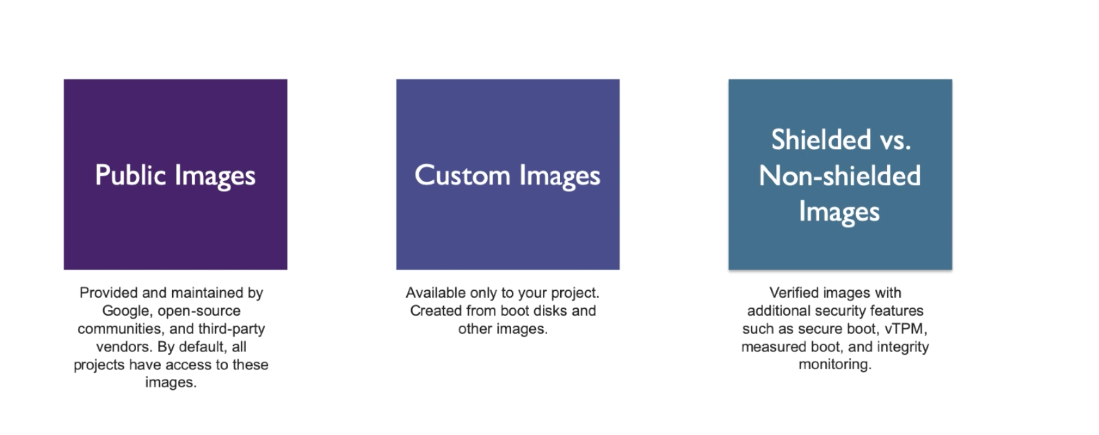


<https://cloud.google.com/compute/docs/machine-types>

## Compute Decision Tree



## VM Images



## 

## Instance Groups

Instead of Deploying Single Instances, we can also deploy groups of instances. We can deploy both Vertically and Horizontally.

## Considerations

* Cost
* Data Locality
* High Availability
  1. Consider preemptible machine for non HA use cases.
* Latency

## Compute Instance Pre-Defined Roles

Compute Engine Admin: Full control

Compute Engine Network Admin : Create, Modify and Delete most network resources

Compute Engine Security Admin: Can create, modify and delete SSL certificates and firewall rules

Compute Engine Viewer: Read only roles, can list the resources but can’t read the resources themselves.

### Google cloud shell commands

gcloud compute instances list

gcloud compute images list

gcloud compute machine-types list

## Deploy a VM

gcloud compute instances create <name>

You get the zone, y

The machine is created with default settings.

gcloud compute instances create skyline-vm-2 --boot-disk-size 200GB --machine-type n1-standard-2

Y

## Virtual Machine Management

Click into the virtual machines, we can edit the configurations at the top. We can also set the Windows password. Reset, (Hard Reset)

Create similar, (cloneish)

STOP shutdown the instance.

DELETE, the bootdisk will be there if you choose not to delete the bootdisk

# Google Kubernetes Engine(GKE)

## What is a Docker container?

1. It is a standardised packaging for software and dependencies.
2. It is a way to isolate apps from each other.
3. Works with Linux and Windows servers
4. Allows separate OS Apps to share the same kernel

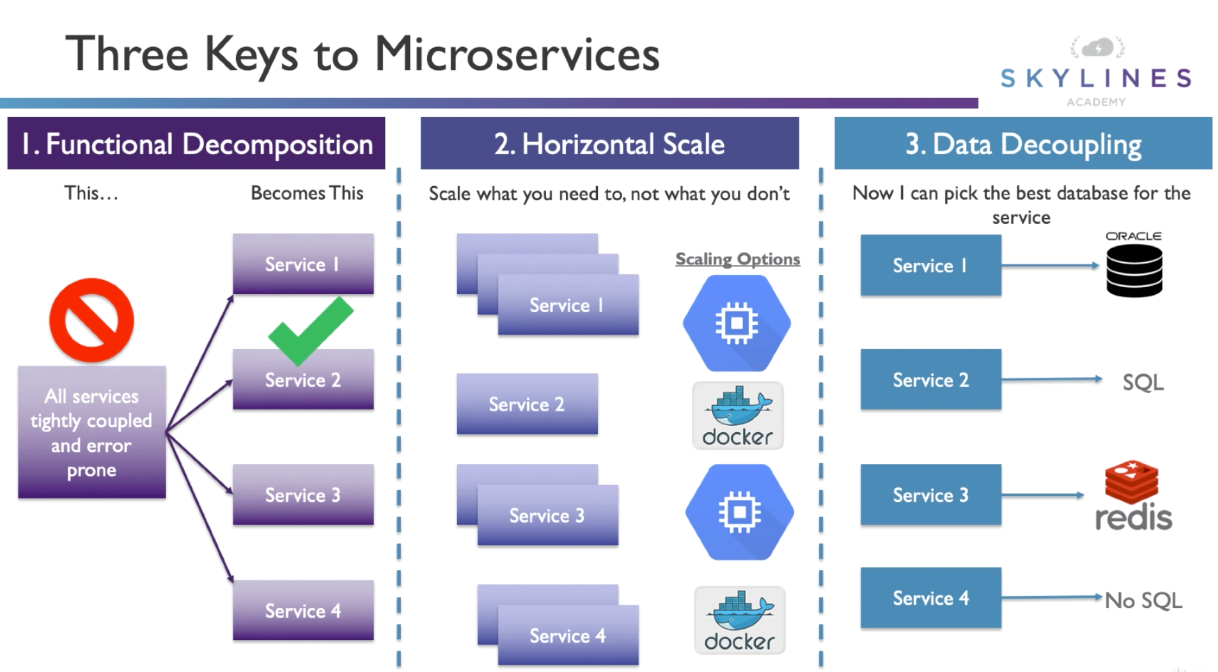
## Application Modernization

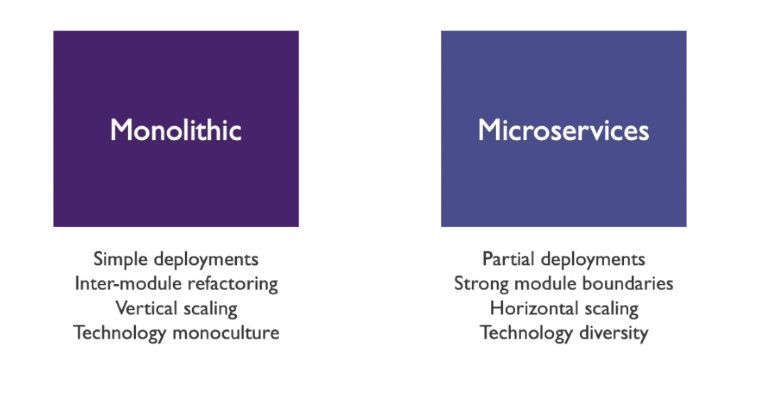
### Monolithic App Issues

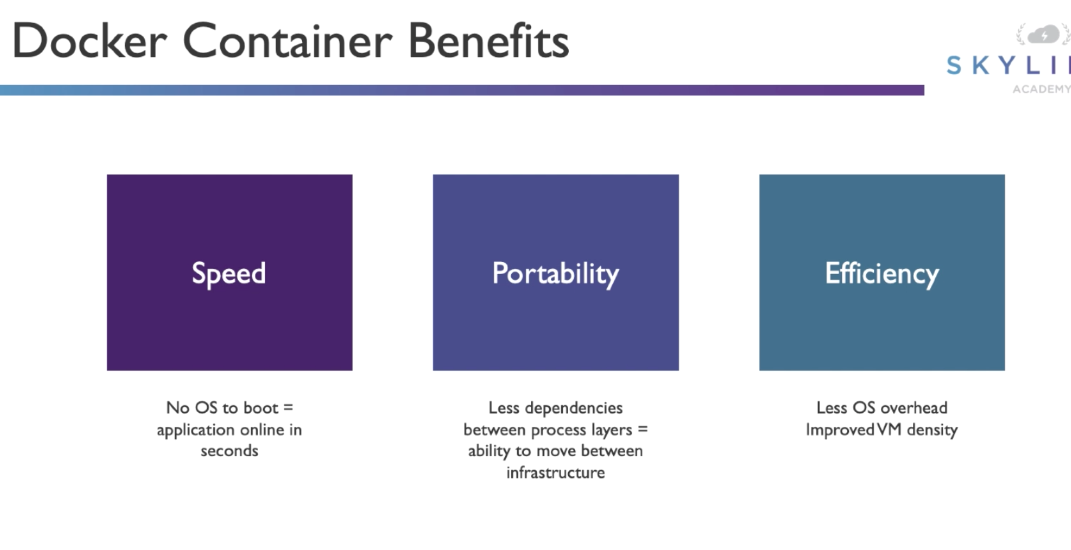
* Minor Code Requires full recompile and Testing
* Application becomes a single point of failure
* Application is difficult and expensive to scale

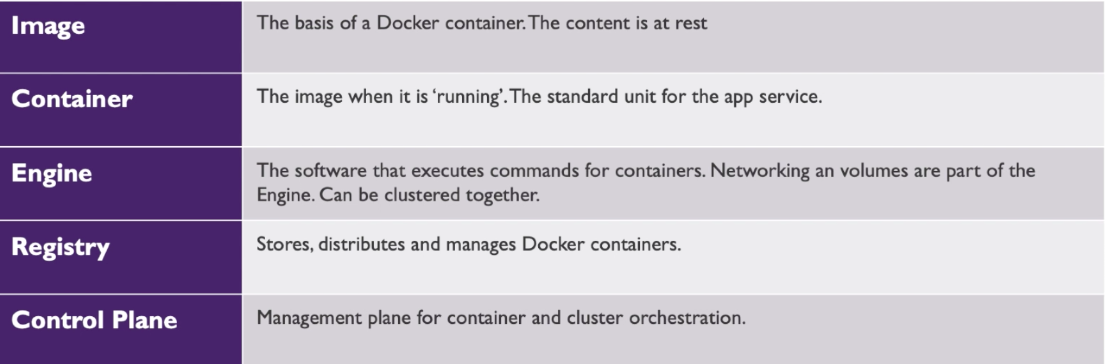
### Microservices

* Break the Application into separate Microservices
* 12 Factor Apps -> Make the apps independently scalable, stateless and highly availably by design.

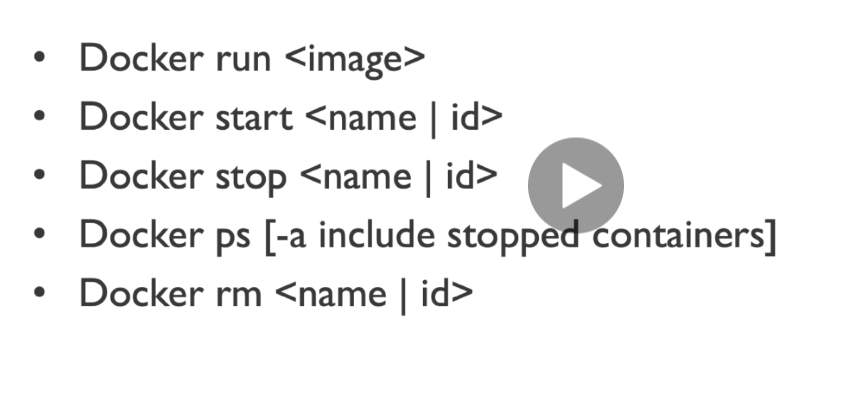








Docker commands

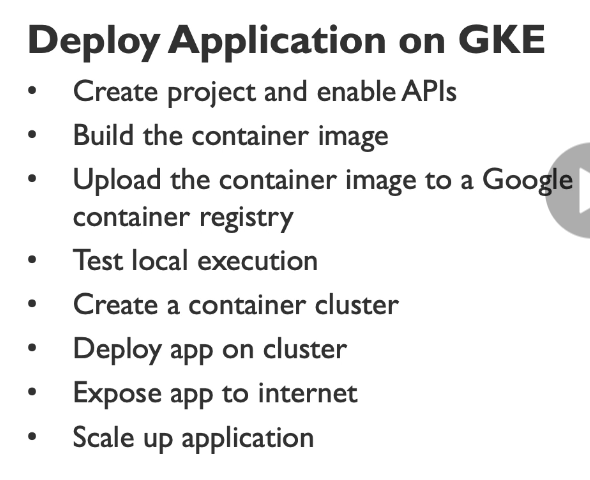


## Google Kubernetes Engine

Google Kubernetes Engine provides a managed environment for deploying, managing , and scaling your containerized applications using Google Infrastructure. The GKE environment consists of multiple machines grouped together to form a cluster.

GKE Features

1. **Load Balancing** for Compute Engine Instances
2. **Node pools** to designate subsets of node within the additional cluster for additional visibility
3. **Automatic scaling** for your clusters node instance count
4. **Automatic Upgrades** for your clusters node software
5. **Node auto-repair** to maintain node-health and availability
6. **Logging and Monitoring** with stackdriver for visibility into your cluster



Click explore and enable API’s from scrolling down which will take you to the API’s and services dashboard.

Enable the Kubernetes Engine API -> ENABLE

Container Registry, Google Container Registry API

Export PROJECT\_ID=<>

docker build -t gcr.io/$(PROJECT\_ID)/hello-app:v1 .

docker images

gcloud auth configure-docker

Docker push gcr.io/$(PROJECT\_ID)/hello-app:v1

Web preview

Docker run --rm -p 8080:8080 gcr.io/$(PROJECT\_ID)/hello-app:v1

gcloud config set project $(PROJECT\_ID)

gcloud config set compute/zone us-east1

gcloud container clusters create skylines-cluster --num-nodes = 2

gcloud compute instances list

Kubectl create deployment hello-web --image-gcr.io/$(PROJECT\_ID)/hello-app:v1

Kubectl get pods

Kubectl expose deployment hello-web --type=LoadBalancer --port 80 --target -port 8080

Kubectl get service