



# Google Cloud Digital Leader Exam Prep Guide

Exam Preparation for GCP Digital Leader Certification for NEW to  
Google Cloud & Non-technical Professionals

*January 2022*

Sarah Walker-Leptich | [www.walkrinthecloud.com](http://www.walkrinthecloud.com)

---

**Note from creator:** These notes have been adapted from the GCP [Digital Leader Exam Pro](#) course created by [Andrew Brown](#), as well as my own research and doc files. I highly recommend watching his free training videos, paying for his practice exams/console demos (\$24 USD), and reading through these notes before taking the exam.

**This exam prep was specifically designed with the exam guide before January 26 modifications. It still provides an in-depth and highly sufficient overview of Google Cloud that will prepare you for writing the updated exam.**

This is meant to be a community document. As of right now anyone can comment on the Google Doc if they see a mistake or want to make a correction. Permissions will be given to individuals who request to make edits, please always respect each other, even if you have different opinions.

<b>Chapter One   Cloud Computing Overview</b>	<b>4</b>
What is Cloud Computing?	4
The Evolution of Computing	4
Types of Environments	5
Benefits of Cloud Computing	6
Types of Cloud Computing	6
Shared Responsibility Model for Google Cloud	6
Shared Responsibility Rules	7
Shared Responsibility Model for Compute	7
Total Cost of Ownership (TCO)	9
Common Cloud Terminology	9
Global Infrastructure Terminology	11
Resource Scoping Terms	11
Cloud Interconnect	12
Latency and Lag	12
Computing Power for GCP	13
<b>Chapter Two   Google Cloud Tools and Products</b>	<b>13</b>
Google Cloud Tools	13
Projects and Folders	13
Google TAM	15
<b>Chapter Three   Google Cloud Adoption Framework</b>	<b>16</b>
Google Cloud Adoption Framework	16
GCAF Themes	16
GCAF Phases	17
GCAF Maturity Scale	18
GCAF Epics	18
GCAF Cloud Maturity Assessment	19
<b>Chapter Four   Google Cloud Core Products and Services</b>	<b>20</b>
Compute Services and Products	20
App Engine Environments	21
Kubernetes and Containers	22
Databases and Data Analytics	24
What is a Document store?	25
Example Questions:	27
Storage	28
AI & ML	29
Networking	32
API Management and Apigee	33

IoT	34
Developer Tools	34
Cloud Operations Suite	35
Migration and Anthos	36
Serverless	38
Firebase	39
Cloud Deployment Manager	39
Internal Services	40
<b>Chapter Five   Security and Identity Management</b>	<b>40</b>
Identity Management	41
Access Context Manager	41
Active Directory	41
Cloud Identity	42
Security	45
Google Cloud Enterprise Privacy Commitments	47
Google Cloud Trust Principles	48
<b>Chapter Six   Support</b>	<b>48</b>
Service Level Agreements (SLAs)	48
Support Plans & Services	49
<b>Chapter Seven   Billing</b>	<b>51</b>
Billing Health Checks and Alerts	51
Billing Services	52
<b>Chapter Eight   Pricing</b>	<b>53</b>
Free-Trial	53
Free-Tier	54
On-Demand	54
Commutated Use Discounts (CUDs)	54
Sustained Use Discounts (SUDs)	55
Preemptible VM Instances or Sole-Tenant Node Pricing	56
Flat-Rate Pricing	56
<b>Chapter Nine   Resource Hierarchy</b>	<b>57</b>
Environment-Oriented Hierarchy	57
Function-Oriented Hierarchy	57
Granular-Access-Oriented Hierarchy	57
<b>Chapter Ten   Practice Exam Q&amp;A</b>	<b>57</b>

# Chapter One | Cloud Computing Overview

## What is Cloud Computing?

The practice of using a network of remote services hosted on the internet to store, manage and process data rather than a local server or PC (personal computer).

On-premises	Public Cloud
<ul style="list-style-type: none"> <li>- You purchase and house your own servers</li> <li>- You have to hire people to manage the servers</li> <li>- You own the risk</li> </ul>	<ul style="list-style-type: none"> <li>- You rent servers</li> <li>- You rent experts/people</li> <li>- You are only responsible for the configuration and code</li> <li>- Shared responsibility model</li> </ul>

## The Evolution of Computing

Dedicated Server	Virtual Private Server	Shared Hosting	Cloud Hosting
One physical machine for a single business and a single app on site.	One physical machine dedicated to a single business, but runs multiple applications on a site on virtualized sub-machines.	One physical machine shared by hundreds of businesses, relies on most tenants under utilizing their resources.	Multiple physical machines that act as one with multiple cloud services.

### Dedicated Server (pros/cons)

- A physical server wholly utilized by a single customer
- Customers overpay for an underutilized server
- You can't vertical scale, you need manual migration
- Replacing a server is difficult
- You are limited by your host operating system (OS)
- Multiple apps can result in conflicts in resource sharing
- You have a 'guarantee' of security, privacy and full utility of underlying resources (guarantee is only based on the skills of your IT team)

### Virtual Machines (pros/cons)

- Can run multiple VMs (virtual machines) on one machine
- Hypervisor is the software layer that lets you run the VM
- A physical server shared by multiple customers
- Pay for a fraction of the server
- You will overpay for an underutilized VM
- Limited by your guest operating system
- Multiple apps on a single VM can result in conflicts in resource sharing
- Easy to export or import images for migration
- Easy to vertical or horizontal scale

### Containers (pros/cons)

- VM runs multiple containers
- You can maximize the utilization of the available capacity which is most cost effective
- Your containers share the same underlying OS so containers are more efficient than multiple VMs
- Multiple apps can run side by side without being limited to the same operating system requirements and will not cause conflicts during resource sharing

### Functions (pros/cons)

- Are managed VMs running managed containers (known as serverless compute)
- You upload a piece of code, choose the amount of memory and duration
- Only responsible for the code/data
- Very cost effective, only pay for the time code is running, VMs only run when there is code to be executed
- Cold start (booting up) is a con

## Types of Environments

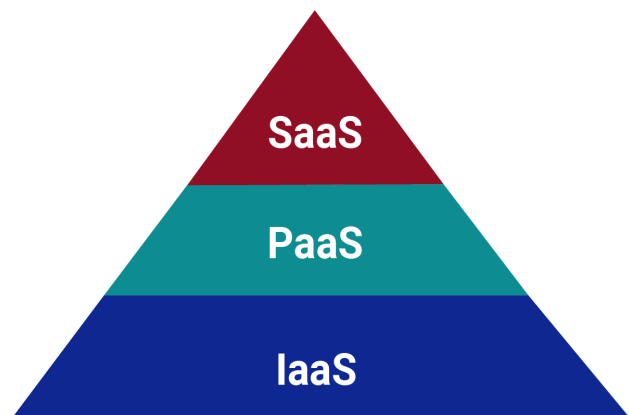
1. **On-premises** - sometimes called 'private cloud'
2. **Public Cloud** - sometimes called 'cloud native' (Azure, AWS, Google Cloud, etc)
3. **Hybrid** - mixture of of both on-premises and public cloud
4. **Multi Cloud** - mixture of different public clouds, sometimes called 'cross cloud'

## Benefits of Cloud Computing

1. **Cost Effective:** you pay for what you consume, no upfront cost. On-demand pricing or pay as you go with thousands of customers sharing the cost of resources.

2. **Global:** launch workloads anywhere in the world, just choose a region.
3. **Secure:** Cloud providers take care of the physical security. Cloud services are secure by default or you have the ability to configure access down to a granular level.
4. **Reliable:** data backup, disaster recovery, data replication and fault tolerance.
5. **Scalable:** Increase or decrease resources and services based on demand.
6. **Elastic:** automate scaling during spikes and drops in demand.
7. **Current:** the underlying hardware and managed software is patched, upgraded and replaced by the cloud provider without service interruptions.

## Types of Cloud Computing



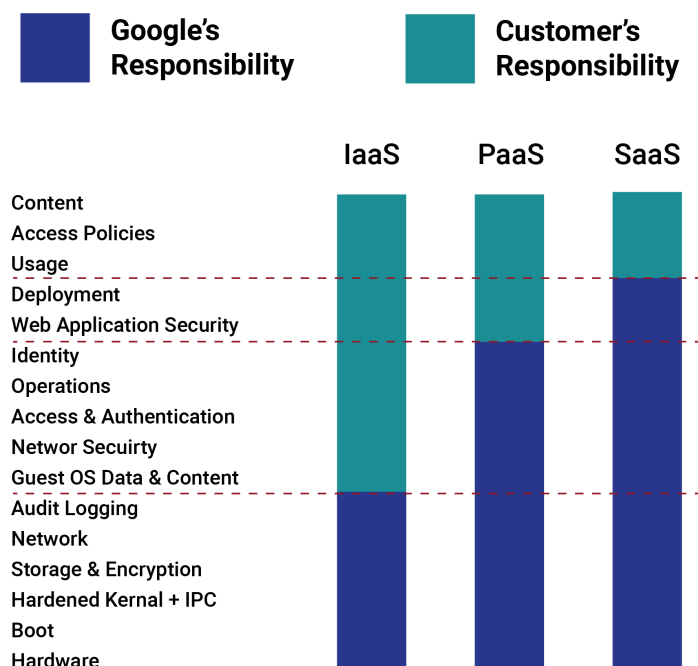
**Software as a Services (SaaS):** generally for end user customers. Examples include Gmail or Twitter.

**Platform as a Service (PaaS):** mainly for developers who don't want to worry about operating systems. Examples include Google AppEngine or AWS Elastic Beanstalk.

**Infrastructure as a Service (IaaS):** the cloud providers handle compute, networking, hardware etc. Examples are Azure, AWS, and Google Cloud.

## Shared Responsibility Model for Google Cloud

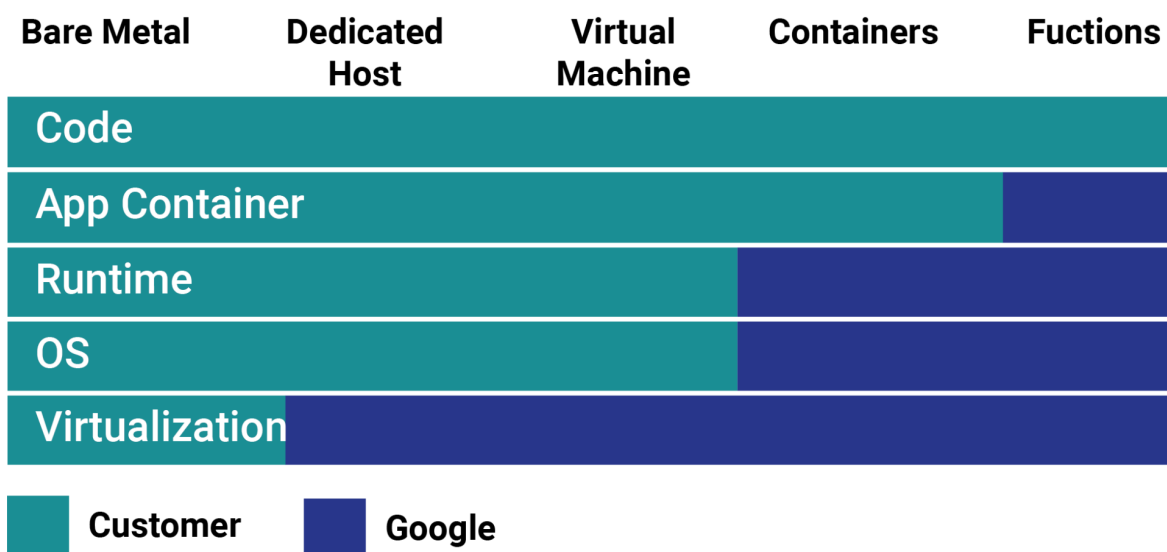
The chart below defines what the customer is responsible for VS what Google Cloud is responsible for.



## Shared Responsibility Rules

1. The **customer** is responsible for the data and the **configuration** of access controls that reside in Google Cloud.
2. The **customer** is responsible for the **configuration** of cloud services and granting access to users via permissions.
3. **Google** is generally responsible for the underlying infrastructure.
4. **If you can configure or store it, the customer is responsible.**
5. **If you cannot configure it, Google is responsible.**

## Shared Responsibility Model for Compute



Infrastructure as a Service (IaaS)		
<b>Bare Metal - Compute Engine</b>	<b>Virtual Machine - Computer Engine</b>	<b>Containers - Google Kubernetes Engine (GKE)</b>
<b>Customer</b> <ul style="list-style-type: none"> <li>- The host OS configuration</li> <li>- hypervisor</li> </ul> <b>Google</b> <ul style="list-style-type: none"> <li>- Physical machines</li> </ul>	<b>Customer</b> <ul style="list-style-type: none"> <li>- The guest OS configuration</li> <li>- Container runtime</li> </ul> <b>Google</b> <ul style="list-style-type: none"> <li>- Hypervisor, physical machine</li> </ul>	<b>Customer</b> <ul style="list-style-type: none"> <li>- Configuration of containers</li> <li>- Deployment of containers</li> <li>- Storage of containers</li> </ul> <b>Google</b> <ul style="list-style-type: none"> <li>- The OS, the hypervisor, container runtime</li> </ul>

Platform as a Service (PaaS)	Software as a Service (SaaS)	Function as a Service (FaaS)
Google App Engine	Google Doc/Microsoft Word	Google Cloud Functions
Customer	Customer	Customer

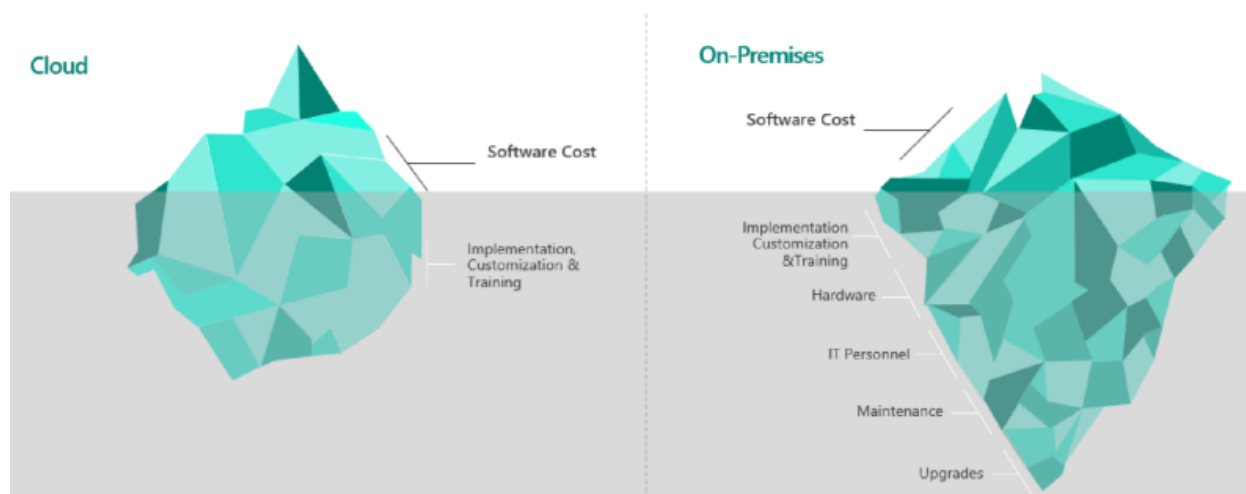


<ul style="list-style-type: none"> <li>- Uploading code</li> <li>- Some configuration of environment</li> <li>- Deployment strategy</li> <li>- Configuration of associated services</li> </ul> <p><b>Google</b></p> <ul style="list-style-type: none"> <li>- Servers, OS, networking, storage, security</li> </ul>	<ul style="list-style-type: none"> <li>- Contents of documents</li> <li>- Management of files</li> <li>- Configuration of sharing access</li> </ul> <p><b>Google</b></p> <ul style="list-style-type: none"> <li>- Servers, OS, networking, storage, security</li> </ul>	<ul style="list-style-type: none"> <li>- Upload your own code</li> </ul> <p><b>Google</b></p> <ul style="list-style-type: none"> <li>- Deployment, container runtime, networking, storage, security, physical machine (basically everything)</li> </ul> <p><b>Note: Function as a Service is sometimes called 'Serverless Computing'</b></p>
--	---	--

## Total Cost of Ownership (TCO)

**CapEx (on-premises):** money used by a company to acquire, upgrade, and maintain physical infrastructure assets such as data centers, people, buildings, electricity, technology, and equipment. Often hidden costs that we don't immediately think of, shown below in the iceberg analogy.

**OpEx (cloud):** are the day-to-day expenses a company incurs to keep its business operational. Less hidden costs because infrastructure is managed by public cloud providers.



## Common Cloud Terminology

**GCP:** Google Cloud Platform - I will use both GCP and Google Cloud in this doc and I am referring to the exact same thing.

**Availability:** your ability to ensure a service remains available.

**Scalability:** your ability to grow rapidly or unimpeded .

**Elasticity:** your ability to shrink and grow to meet the demand.

**Disaster Recovery:** your ability to recover from a failure.

**Load Balancer:** a load balance allows you to evenly distribute traffic to multiple servers in one or more data centers.

**High Availability (HA):** ability for your service to remain available by ensuring there is no single point of failure and or ensure a certain level of performance.

**Highly Scalable:** your ability to increase your capacity based on the increasing demand of traffic, memory, and computer power.

**Vertical scaling:** or “scaling-up”, which is upgrading to a bigger server.

**Horizontal Scaling:** or “scaling-out”, which is adding more servers of the same size or “scaling-in”, which is removing servers.

**High Elasticity:** ability to automatically increase or decrease your capacity based on the current demand of traffic memory or computing power.

**NOTE:** the difference between high elasticity and highly scalable is that elasticity involves automation in either scaling or decreasing.

**Fail-over:** is when you have a plan to shift traffic to a redundant system in case the primary systems fail.

**High Durability:** your ability to recover from a disaster and to prevent the loss of data. Solutions that recover from a disaster are known as disaster recovery (DR).

**Managed Instance Group (MIG):** automatically increasing or decreasing the workload in response to demand or a defined schedule.

**Cloud DNS:** can detect a failing primary system and fail-over to a secondary system on **stand-by**.

**GovCloud:** isolated regions to run FedRAMP(Federal Risk and Authorization Management Program) workloads. Google doesn't have one and says the usual regions work at high security

**Computing Power:** the throughput measured at which a computer can complete a task.

**SLA:** service level agreement.

## Global Infrastructure Terminology

**Regions:** are independent geographic areas that consist of zones - smaller areas within a region.

**Zones:** physical locations made up of one or more data centers or hardware.

**Note:** as of Jan 2022, Google has 29 regions and 88 zones, roughly 3 zones per region. It is standard to have 3 zones per region. Data centers within a region will be isolated from each other (for example - different buildings, but close enough to provide low latency)

**Edge Networking:** is the practice of having compute and data storage resources as close as possible to the end user in order to deliver the lowest latency and to save bandwidth.

**Points of Presence (POP):** is an intermediate location between a GCP region and the end user. This location could be a third party datacenter or collection of hardware.

**Edge POP:** a location where a user can quickly enter (ingress) the GCP network for accelerated access to cloud resources.

**CDN POP:** a location to serve (egress) cached website files and assets so they load very fast for the end user.

**Cloud Media Edge:** a location specialized for the delivery of media such as video content.

## Resource Scoping Terms

**Zonal Resource:** a resource that resides in multiple zones in a single region.

**Multiple-regional Resource:** a resource that resides across multiple specified zones.

**Global Service:** a resource that resides globally and regions/zones are abstracted away.

**Internal Services:** a foundation service used by many other services. You don't interact with these services directly, they are managed by Google. Examples are Spanner, Colossus, Borg and Chubby.

**Data Residency:** the physical or geographical location of where an organization's data or cloud resources reside.

**Compliance Boundaries:** a regulatory compliance by a government or organization that describes where data and cloud resources are allowed to reside. For example, hospitals and private health data.

**Assured Workloads:** a feature that allows you to apply various 'security controls' based on attributes, personal support case ownership controls and encryption.

**Important note for exam:** in order to choose data residency, you need to update an "organization policy" called Resource Location and choose the allowed regions or multiregion.

## Cloud Interconnect

Cloud Interconnect provides direct physical connections between your on-premises network and Google's network. Cloud Interconnect enables you to transfer large amounts of data between networks, which can be more cost-effective than purchasing additional bandwidth over the public internet.

### Two Offerings | Dedicated or Partner

**Dedicated:** a direct physical connection between the on-premises network and Google's network through a co-location. **(Ideal for 10-200GB)**.

**Partner:** a direct physical connection between the on-premises network and Google's network through a trusted third-party. **(Ideal for 50MB-10GB)**

## Latency and Lag

**Latency:** the time delay between two physical systems.

**Lag:** the noticeable delay between the actions of input and the reactions of the server sent back to the client.

**Important for exam:**

**Inter-regional Latency:** triple digit milliseconds (500ms)

**Inter-zonal Latency:** double digit milliseconds (10ms)

## Computing Power for GCP

**CPU or Central Processing Unit:** CPU processors are found in Google's offering Compute Engine. These are basic CPUs and are normally Xeon.

**Tensor Computing:** TPUs or Tensor Processing Units are 50x faster than traditional CPUs, created by Google and are cooled with water. TPUs are great for AI/ML and the Google product is called Cloud TPU.

**Quantum Computing:** 100 million times faster. Will be essential to super artificial intelligence and Google's products are Google Quantum AI. The latest quantum computing chip is called Sycamore (2018), followed by Bristlecone (2017) and then Foxtail (2016)

## Chapter Two | Google Cloud Tools and Products

### Google Cloud Tools

**Google Cloud Console:** the GCP portal is a **web-based, unified console**, that you can **build, manage, and monitor** everything from a web browser.

**Cloud SDK:** SDK means software development kit and it is a collection of software development tools in one installable package. Cloud SDK is used to programmatically create, modify, delete or interact with GCP resources. Cloud SDK is offered in various languages: JAVA, Python, Nodejs, Ruby, Go, .Net and PHP.

**Cloud CLI:** CLI means command line interface and it processes commands to a computer program in the form of lines of text.

**Cloud Shell:** command line access for managing your infrastructure (basically the code editor in the GCP console for managing your infrastructure).

## Projects and Folders

**Note:** this section is important on the exam.

**Project:** a logical grouping of resources. A cloud resource must belong to a project. A project is made up of the settings, permissions, and other metadata.

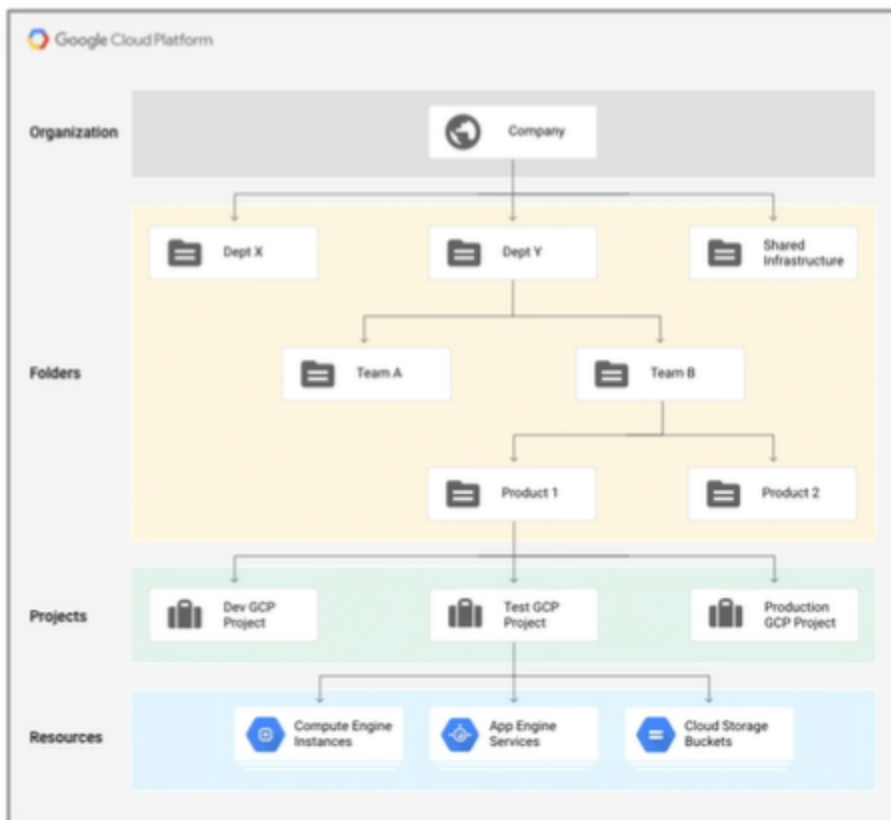
- A project can't access another project's resources unless you use Shared VPC\* (Virtual Private Cloud) or VPC Network Peering.
- Resources within a single project can work together easily, for example by communicating through an internal network, subject to the regions-and-zones rules.
- Each project ID is unique across Google Cloud.
- Once you have created a project, you can delete the project but its ID can never be used again.
- When billing is enabled, each project is associated with one billing account.
- Multiple projects can have their resource usage billed to the same account.
- A project serves as a namespace: this means every resource within each project must have a unique name, but you can usually reuse resource names if they are in separate projects

**Each project has the following:**

1. A project name, which you provide.
2. A project ID, which you can provide or Google Cloud can provide for you.
3. A project number, which Google Cloud provides

**Folders:** allow you to logical group multiple projects that share common IAM permissions. Folders are commonly used to isolate projects for different departments or for different environments.

**Important Note:** you do not need to have a folder, BUT you must have a project.



## Google TAM

Technical Account Manager (TAM) is a human resource **assigned to work with your organization when paying for Google Cloud's Premium Support**

A TAM can assist with Google Cloud Adoption Framework (next chapter) by performing a high-level assessment of your organization's cloud maturity and tell you how to prioritize your:

- training and change management programs
- partner relationships
- cloud operating model
- secure account configuration

## Chapter Three | Google Cloud Adoption Framework

### Google Cloud Adoption Framework

**Google Cloud Adoption Framework (GCAF)** is a guide from Google Cloud to help its customer with the following:

- Determining an organization's readiness to adopt Google Cloud
- Where to identify and fix any knowledge gaps
- How to develop new competencies

There are five main areas in the Google Cloud Adoption Framework guide:

1. **Themes** – Learn, Lead, Scale, Secure
2. **Maturity Phases** – Tactical, Strategic, Transformational
3. **Cloud Maturity Scale** – Matrix of Themes and Phases
4. **Epics** – Workstreams to scope and structure cloud adoption
5. **Programs** – Logical Grouping of Epics

Link to the [Google Cloud Adoption Framework](#).

## GCAF Themes

In order for your company to be successful in cloud adoption, Google Cloud recommends four themes to focus on and internally audit:

Learn	Lead	Scale	Secure
<p>The quality and scale of the learning programs you have in place to upskill your technical teams, and your ability to augment your IT staff with experienced partners.</p> <p><b>Questions to Ask:</b>  <i>Who is engaged? How widespread is that engagement? How concerted is the effort? How effective are the results?</i></p>	<p>The extent to which IT teams are supported by a mandate from leadership to migrate to cloud, and the degree to which the teams themselves are cross-functional, collaborative, and self-motivated.</p> <p><b>Questions to Ask:</b>  <i>How are the teams structured? Have they got executive sponsorship? How are cloud projects budgeted, governed, assessed?</i></p>	<p>The extent to which you use cloud-native services that reduce operational overhead and automate manual processes and policies.</p> <p><b>Questions to Ask:</b>  <i>How are cloud-based services provisioned? How is capacity for workloads allocated? How are application updates managed?</i></p>	<p>The capability to protect your services from unauthorized and inappropriate access with a multilayered, identity-centric security model. Dependent also on the advanced maturity of the other three themes.</p> <p><b>Questions to Ask:</b>  <i>What controls are in place? What technologies are used? What strategies govern the whole?</i></p>



## GCAF Phases

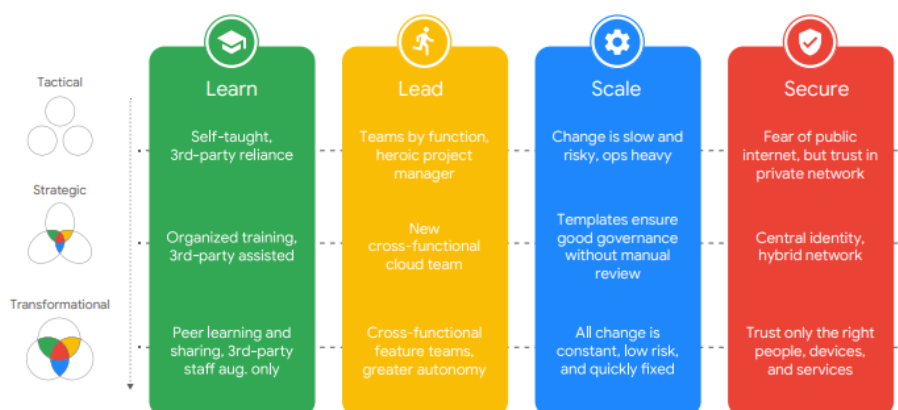
Google Cloud Adoption Framework also suggests that a businesses readiness for success in the cloud is determined by current business practices in each of these four themes. For each theme, those practices will fall into one of the following phases:

**Tactical (short term goals):** Individual workloads are in place, but no coherent plan encompassing all of them with a strategy for building out to the future. The focus is on reducing the cost of discrete systems and on getting to the cloud with minimal disruption. The wins are quick, but there is no provision for scale.

**Strategic (mid term goals):** A broader vision governs individual workloads, which are designed and developed with an eye to future needs and scale. You have begun to embrace change, and the people and processes portion of the equation are now involved. IT teams are both efficient and effective, increasing the value of harnessing the cloud for your business operations.

**Transformational (long term goals):** With cloud operations functioning smoothly, you've turned your attention to integrating the data and insights garnered from working now in the cloud. Existing data is transparently shared. New data is collected and analyzed. The predictive and prescriptive analytics of machine learning applied. Your people and processes are being transformed, which further supports the technological changes. IT is no longer a cost center, but has become instead a partner to the business.

## GCAF Maturity Scale



When you evaluate the four themes in terms of the three phases, you get the Cloud Maturity Scale. In each of the themes, you can see what happens when you move from adopting new

technologies ad hoc, to working with them more and more strategically across the organization – which naturally means deeper, more comprehensive, and more consistent training for your people, which in turn means streamlined and updated processes, which in its turn drives innovation. The whole organization gradually transforms.

## GCAF Epics

Once you've determined where you are in your cloud maturity journey, it's time to move forward.



Tactical



Strategic



Transformational

To scope and structure your program of cloud adoption, you will implement a number of workstreams (which we call epics). The epics are defined so that they do not overlap, they are aligned to manageable groups of stakeholders, and they can be further broken down into individual user stories, making your program planning easier.

Here's a look at those epics within the familiar rubric of people, technology, and process. If you can do only a subset of the epics, focus on the ones in the colored segments. Those are the epics that align with Learn, Lead, Scale, and Secure; and so those are the epics that will define your journey to successful cloud adoption.



When you've determined *where* your organization is in the adoption process using the **Cloud Maturity Scale** then you need to define Epics.

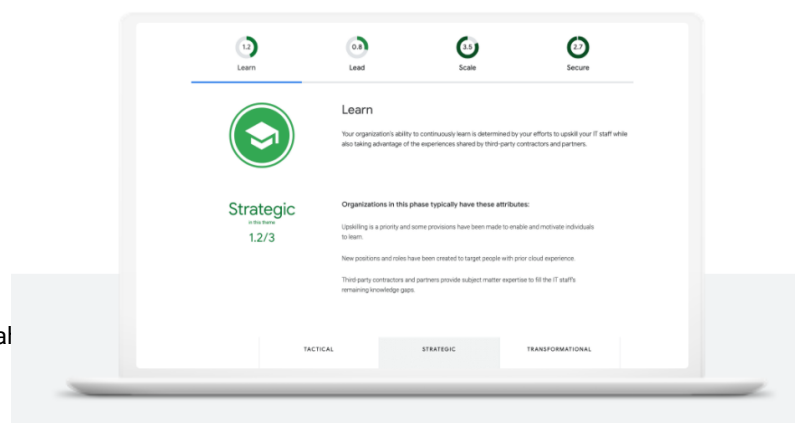
**Epics** are workstreams to scope and structure cloud adoption and they include everything in the venn diagram above.

- **epics** are defined so that they do not overlap
- they are aligned to manageable groups of stakeholders
- they can be further broken down into individual **user stories**

If you are limited in time and resources, focus on the epics in the **coloured segments** since these align with Learn, Lead, Scale, and Secure.

## GCAF Cloud Maturity Assessment

In order to determine where you are on the Cloud Maturity Scale, Google created a Cloud Maturity Assessment that you can practice filling out before the exam.



[Check out the Cloud Maturity Assessment.](#)



## Chapter Four | Google Cloud Core Products and Services

After you read this chapter - **VIEW THE [GOOGLE CLOUD PRODUCT FLASHCARDS](#)**



### Compute Services and Products

**Important Note:** read the deeper dive on App Engine below chart.

CATEGORY	PRODUCT TYPE	GCP PRODUCT	ICON	DESCRIPTION
Compute/ Serverless Computing	Serverless Applications Platform	App Engine		Serverless Application platform for apps and back ends. Build highly scalable applications on a fully managed serverless platform. Build your application in Node.js, Java, Ruby, C#, Go, Python, or PHP—or bring your own language runtime.
Compute	Compute Service (VMs in the Cloud)	Compute Engine		Virtual machines running in Google's data center. Secure and customizable compute service that lets you create and run virtual machines on Google's infrastructure.
Compute	VM Migration as a service	Migrate for Compute Engine	N/A	Server and virtual machine migration to Compute Engine. With Migrate for Compute Engine simple “as a service” interface within Cloud Console and flexible migration options. Migrate VM instances to Google Cloud from AWS, Azure, or VMWare VSphere.

Compute	Dedicated & Physical Compute Engine Server	Sole-Tenant Nodes	N/A	Sole-tenancy lets you have exclusive access to a sole-tenant node, which is a physical Compute Engine server that is dedicated to hosting only your project's VMs.
Compute	VMware app migration as a service	VMware Engine (GCVE)	N/A	Easily lift and shift your VMware-based applications to Google Cloud without changes to your apps, tools, or processes. The service provides all the hardware and VMware licenses you need to run in a dedicated VMware SDDC in Google Cloud.
Compute	Containers without infrastructure	Cloud Run		Develop and deploy highly scalable containerized applications on a fully managed serverless platform. <b>Run stateless containers on a fully managed environment or on Anthos.</b>
Compute	Core compute	Cloud GPUs		Add GPUs (graphics processing unit) to your workloads for machine learning, scientific computing, and 3D visualization.
Compute	Cloud cost optimization	Recommender	N/A	Optimize your Google Cloud usage with proactive, easily actionable recommendations
Serverless Computing	FaaS	Cloud Functions		Run your code with zero server management with this scalable, pay-as-you-go functions-as-a-service (FaaS) offering. <b>Scalable to Zero</b>
Serverless Computing	Workflow orchestration	Workflows		Orchestrate and automate Google Cloud and HTTP-based API services with serverless workflows.

## App Engine Environments

App Engine is a Platform as a Service (PaaS) for your application where you can quickly deploy and scale web-applications without worrying about the underlying infrastructure.

App Engine has two types of environments: Flexible and Standard. You can simultaneously use both environments for your application. App Engine is well suited to applications that are designed using a microservice architecture.

Standard (serverless compute)	Flexible (fully managed containers)
<ul style="list-style-type: none"> <li>- starts in seconds</li> <li>- Runs in a sandbox</li> <li>- designed for rapid scaling (sudden traffic spikes)</li> </ul>	<ul style="list-style-type: none"> <li>- starts in minutes</li> <li>- Runs within Docker Containers on Compute Engine (VMs)</li> <li>- designed for predictable and consistent traffic</li> </ul>

<ul style="list-style-type: none"> <li>- supports specific language versions, not custom runtime</li> <li>- can scale to zero instances (scale to zero)</li> <li>- pricing based on hours</li> <li>- cannot SSH to debug</li> <li>- no background processes</li> <li>- Flexible fully managed container</li> </ul>	<ul style="list-style-type: none"> <li>- supports generally any language version or run custom time</li> <li>- must have at least once instance running</li> <li>- Supports autoscaling</li> <li>- pricing based on vCPUs, Memory and Disks</li> <li>- can SSH to debug</li> <li>- can have background processes</li> </ul>
--	---

## Kubernetes and Containers

### Quick Facts about Kubernetes:

- The name Kubernetes originates from the Greek word κυβερνήτης, meaning helmsmen or steersmen, hence the wheel of a ship as the logo
- It was initially called Project Seven of Nine, based on the female Borg character in Star Trek, which is why there are seven spokes in the wheel
- Kubernetes can also be referred to as K8 (K + the eight letters of ubernete)
- It was designed by Google in 2014, but was donated to the Cloud Native Computing Foundation

### What are Kubernetes and containers?

**Containers:** is a standard unit of software that packages up code and all its dependencies, so the application runs quickly and reliably from one computing environment to another.”

**Container Management:** is the process of organizing, adding, or replacing large numbers of software containers. This leads to the need for container-orchestration, which is basically Kubernetes, a tool that automates scaling, management, and cloning of your containers plus much more.

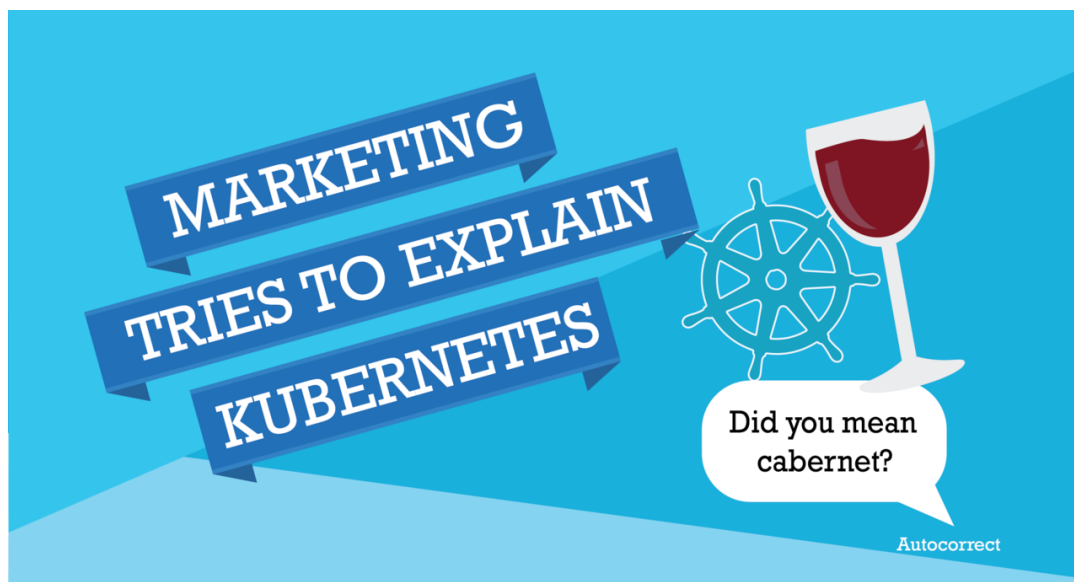
**Kubernetes:** is an open source container-orchestration platform for your applications that can be deployed on your on-premises or cloud environment. It can regulate your cloud costs, but most importantly it offers high-availability to your applications by being able to automatically spin up identical containers and offload them when no longer needed.

A unique component of kubernetes is pods. A **pod** is a group of containers with shared storage, network resources and other shared settings.

**Kubernetes is ideal for microservice-architectures where a company has tens to hundreds of services they need to manage. Best choice for when you have A LOT of services.**

To learn more, I wrote a very non-technical and easy to read article explaining Kubernetes to the average layperson:

<https://www.linkedin.com/pulse/marketing-girl-tries-explain-kubernetes-sarah-kymberlee-walker-/>



Containers	Container registry & security	Container Registry		Store, manage, and secure your container images. Deploy only trusted containers on GKE. Only multi-regional repositories.
Containers	Container registry & security	Artifact Registry		Considered as the next generation of container registry. A fully-managed service with support for both container images and non-container artifacts.
Containers	CI/CD	Cloud Build		Continuously build, test, and deploy containers using the Google Cloud infrastructure.
Containers	Containers without Infrastructure	Cloud Run		Develop and deploy highly scalable containerized applications on a fully managed serverless platform.
Containers	Container security	Container Security	N/A	Secure your container environment on Google Cloud, GKE, or Anthos.
Containers	CaaS	Google Kubernetes Engine (GKE)		Secured and managed Kubernetes service with four-way auto scaling and multi-cluster support.

Containers	AI Container Management	Deep Learning Containers	N/A	Take advantage of containers preconfigured with data science frameworks, libraries, and tools. Efficiently run batch jobs using Kubernetes.
Containers	Pre-built Containers	Kubernetes applications on Google Cloud Marketplace	N/A	Deploy pre built containerized apps.

## Databases and Data Analytics

### What is a database?

A database is a **data-store that stores semi-structured and structured data**. Databases have a rich set of functionality to either use a specialized language to query (retrieve data) specialized modeling strategies to optimize retrieval for different use cases.

### What is structured and unstructured data?

Structured data consists of clearly defined data types with patterns that make them easily searchable; while unstructured data – “everything else” – is composed of data that is usually not as easily searchable, including formats like audio, video, and social media postings.

**Important to Note:** The **General Data Protection Regulation (GDPR)** is a regulation in EU law on data protection and privacy in the European Union and the European Economic Area. It is considered a strict data policy regulation that many businesses should consider making this part of their data security strategy.

### Typical types of databases

Relational Databases	Non-relational Databases
<ul style="list-style-type: none"> <li>- Structured data that strongly represents tabular data (tables, rows and columns)</li> <li>- Row-oriented or Columnar-oriented</li> </ul>	<ul style="list-style-type: none"> <li>- Semi-structured that may or may not distantly resemble tabular data</li> </ul>

### What is a data warehouse?

A data warehouse is a relational datastore designed for **analytic workloads**, which is generally **column-oriented data-store**.



## What is a Document store?







A document store is a **NoSQL database that stores documents** as its primary data structure. A document could be an XML but more commonly is JSON or JSON-Like










## What is a Key/Value store?

A **key-value database** is a type of non-relational database (NoSQL) that uses a simple key-value method to store data. A key/value stores a **unique key** alongside a value. **Key values stores are dumb and fast.**

## What is a data lake?

A data lake is a centralized repository designed to store, process, and secure large amounts of structured, semistructured, and unstructured data. It can store data in its native format and process any variety of it, ignoring size limits.

Databases	NoSQL: Key-value	Cloud Bigtable		Run large analytical and operational workloads using this fully managed, scalable NoSQL database service.
Databases	NoSQL: Indexed	Datastore		A highly scalable NoSQL database for your web and mobile applications.
Databases	RDBMS	Cloud SQL		Manage relational data for MySQL, PostgreSQL, and SQL Server for workloads under 30TB.
Databases	Document data storage	Firestore		Easily develop rich applications using a fully managed, scalable, and serverless document database for both web and mobile apps. <b>Key feature: Able to store and sync data in realtime.</b>
Databases	SQL database migration	Database Migration Service	N/A	Migrate databases to Cloud SQL from on-premises, Compute Engine, and other clouds
Databases	In-memory data store	Memorystore		Reduce latency with scalable, secure, and highly available in-memory service.
Databases	RDBMS for global reach	Cloud Spanner		Fully managed relational database with massive scale, strong consistency worldwide, and up to 99.999% availability.

Data Analytics	<b>SERVERLESS Data Warehouse</b>	BigQuery		Serverless, highly scalable, and cost-effective multi-cloud data warehouse designed for business agility. <b>Built-in ML.</b>
Data Analytics	<b>Open Source Processing</b>	Dataproc		<b>Deploy open-source data and analytics (batch) processing services (Apache Hadoop, Apache Spark, etc.) with improved efficiency and security.</b>
Data Analytics	<b>Business Intelligence</b>	Looker		Business intelligence software and big data analytics platform that helps you explore, analyze and share real-time business analytics easily.
Data Analytics	<b>Data Integration</b>	Cloud Data Fusion		<b>Quickly build and manage data pipelines using fully managed, code-free data integration with a graphical interface</b>
Data Analytics	<b>Data discovery and metadata management</b>	Data Catalog		Discover and understand your data using a fully managed and scalable data discovery and metadata management service.
Data Analytics	<b>Serverless data processing</b>	Dataflow		<b>Unify stream and batch data processing that's serverless, fast (real-time), and cost-effective.</b>
Data Analytics	<b>Workflow orchestration</b>	Cloud Composer		Create, schedule, monitor, and manage workflows using a fully managed orchestration service built on Apache Airflow.
Data Analytics	<b>Messaging</b>	Pub/Sub		Streaming analytics and data integration pipelines to ingest and distribute data. It is equally effective as a messaging- oriented middleware for service integration or as a queue to parallelize tasks.
Data Analytics	<b>Intelligent data preparation</b>	Dataprep		Dataprep by Trifacta is an intelligent data service for visually exploring, cleaning, and preparing structured and unstructured data for analysis, reporting, and machine learning.

### Important notes for databases and analytics for exams.

- Understand BigQuery, Bigtable and Spanner.
- Remember that Firestore has a realtime feature.
- Know the difference between Dataflow, Dataproc, and Cloud Data Fusion.

#### Dataproc *Open-source pipelines*

Perform batch processing, querying, and streaming using a managed **Apache Spark and Apache Hadoop service.**

#### Dataflow *Fully-Managed Pipelines*

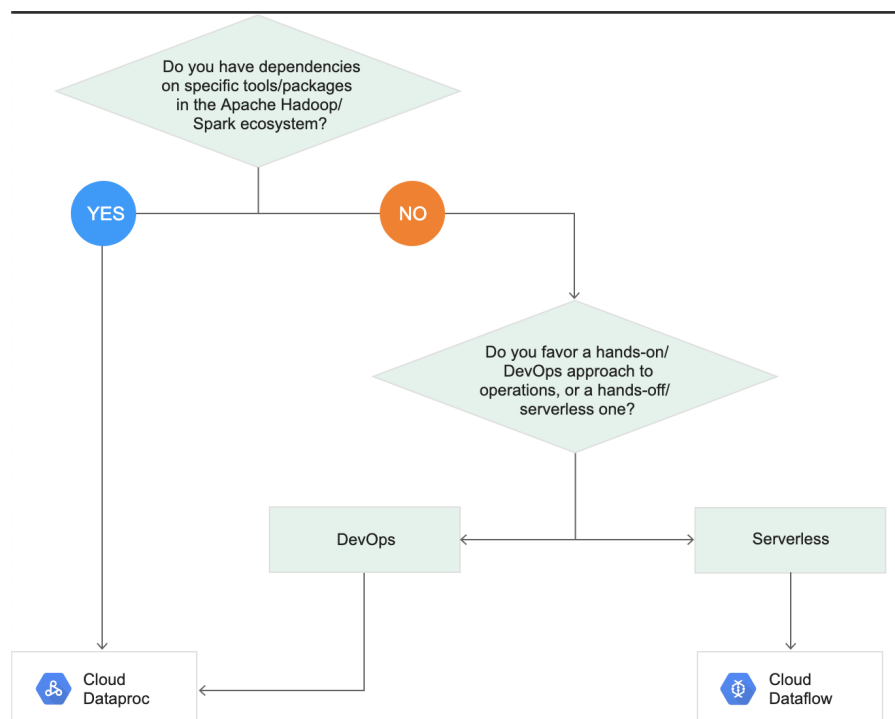
Uses Apache Beam. Fully-managed batch and streaming pipelines. You don't need to balancing work, scaling workers, or any other cluster management. **Remember fully-managed, or serverless.**

### Cloud Data Fusion *Visually build Pipelines*

**A no-code enterprise solution for building ETL pipelines via a drag-and-drop interface.** 150+ pre-configured connectors and transformation.

### Example Questions:

1. Your company is forecasting a sharp increase in the number and size of Apache Spark and Hadoop jobs being run on your local data center. You want to utilize the cloud to help you scale this upcoming demand with the least amount of operations work and code change. Which product should you use?
  - Google Cloud Dataflow
  - Google Cloud Dataproc**
  - Google Compute Engine
  - Google Container Engine
2. A startup plans to use a data processing platform, which supports both batch and streaming applications. They would prefer to have a hands-off/serverless data processing platform to start with. Which GCP service is suited for them?
  - Dataproc
  - Dataprep
  - Dataflow**
  - BigQuery



If you are interested in understanding more about the business impacts of data analytics as well as some of the biggest challenges in the data science community, [check out this PDF](#).



Data Science PDF Link:  
[https://www.walkrinthecloud.com/\\_files/ugd/b4e3ab\\_454701c030134079a662f96beb57d423.pdf](https://www.walkrinthecloud.com/_files/ugd/b4e3ab_454701c030134079a662f96beb57d423.pdf)

# Storage






## Cloud Storage (the main type of Google Cloud Storage):

**Standard Storage:** Good for “hot” data that’s accessed frequently, including websites, streaming videos, and mobile apps. **0 days retrieval**

**Nearline Storage:** Low cost. Good for data that can be stored for at least **30 days**, including data backup and long-tail multimedia content.

**Coldline Storage:** Very low cost. Good for data that can be stored for at least **90 days**, including disaster recovery.

**Archive Storage:** Lowest cost. Good for data that can be stored for at least **365 days**, including regulatory archives.

Storage	File storage	Cloud Storage		Store any amount of data and retrieve it as often as you'd like, using Google Cloud's object storage offering. Comes in four different types, Standard, Nearline, Coldline and Archive.
Storage	File storage	Filestore		High-performance, fully managed file storage.
Storage	Build   Firebase Apps	Cloud Storage for Firebase		Store and serve user-generated content from Firebase apps, such as photos or videos, including bandwidth-friendly transactions and automated ML, synced automatically in real time.
Storage	Block Storage	Persistent Disk		Store data from VM instances running in Compute Engine or GKE, Google Cloud's state-of-the-art block storage offering.
Storage	Storage Migration	Storage Transfer Service		Complete large-scale online data transfers from online and on-premises sources to Cloud Storage.
Backup & DR	Backup and disaster recovery	Actifio GO	N/A	Actifio GO is a Google Cloud backup and disaster recovery offering which enables powerful data protection for Google Cloud and hybrid workloads. Actifio GO supports Google workloads.

## AI & ML

The term artificial intelligence was coined in 1955 by John McCarthy, a math professor at Dartmouth College. Although AI has made significant advancements over the past 66 years, the most substantial improvement has not come from even the past ten years, but just in the last five years since 2016. There are a few different flavors of AI that you will need to understand for the exam.




**Artificial Intelligence (AI):** a broad field or term that describes any machine capable of acting autonomously.

**Machine Learning (ML):** a brand of AI where computers learn from data without using complex rules. ML is a way to use standard algorithms or models to analyze data to drive predictive insights and repeated decisions at scale. It teaches a computer how to solve problems by feeding it examples of the correct answers.

**Deep Learning (DL):** is a type of machine learning that imitates the way humans gain certain types of knowledge. While traditional machine learning algorithms are linear, deep learning algorithms are stacked in increasing complexity and abstraction hierarchy.

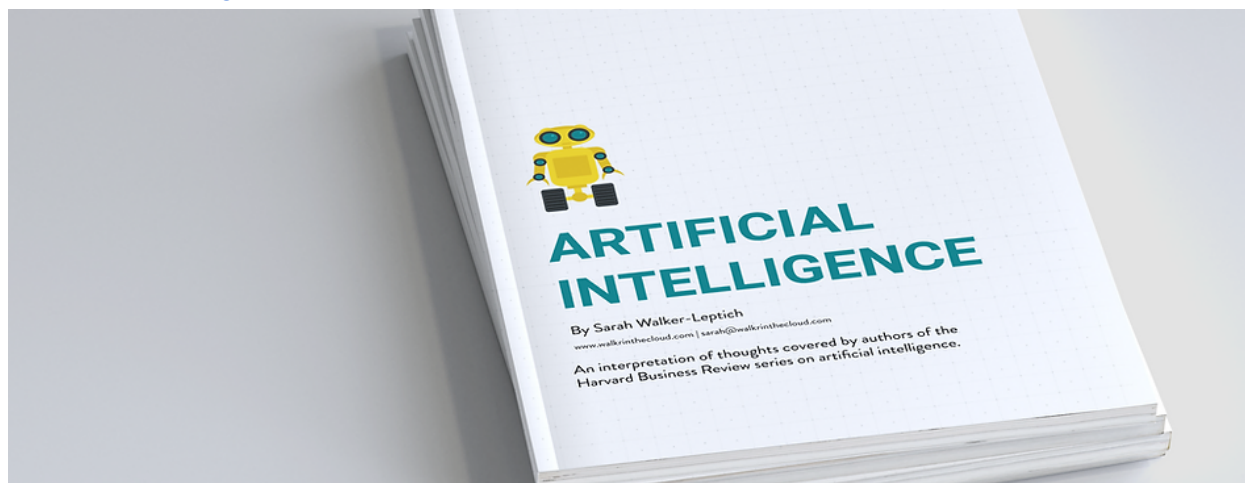
**Reinforcement Learning (RL):** an advanced ML technique that enables models to learn complex behaviors without labeled training data and then make short-term decisions while optimizing for longer-term goals.

**For the exam, familiarize yourself with Vertex AI, Google's all up AI/ML platform and know the benefits of AutoML (which is ML for teams with less experience). Tensorflow is also important to understand, it is an open-source software library for AI/ML. Please bet that there will be a question on either Vision AI or Video AI.**

AI and ML	Fully Managed AI Solution	Vertex AI		Build, deploy, and scale ML models faster, with pre-trained and custom tooling within a unified artificial intelligence platform.
AI and ML	Machine Learning	AutoML		Train high-quality custom machine learning models with minimal effort and machine learning expertise.
AI and ML	Machine Learning Library	Tensorflow		Open-source software library for machine learning and artificial intelligence.

AI and ML	Machine Learning	AutoML Tables		AutoML Tables enables your entire team to automatically build and deploy state-of-the-art machine learning models on structured data at massively increased speed and scale.
AI and ML	Containers	Deep Learning Containers	N/A	Preconfigured and optimized containers for deep learning environments.
AI and ML	AI Services	Vision AI		Derive insights from your images in the cloud or at the edge with AutoML Vision or use pre-trained Vision API models to detect emotion, understand text, and more.
AI and ML	AI Services	Video AI		Enable powerful content discovery and engaging video experiences
AI and ML	AI Services	Natural Language API		Derive insights from unstructured text using Google machine learning.
AI and ML	AI Services	Recommendations AI		Deliver highly personalized product recommendations at scale.
AI and ML	AI Services	Translation		Make your content and apps multilingual with fast, dynamic machine translation.
AI and ML	AI Services	Document AI	N/A	Automate data capture at scale to reduce document processing costs.
AI and ML	Conversational AI	Agent Assist	N/A	Deliver AI-powered conversations with human agents using Agent Assist, part of Google Cloud's Contact Center AI.
AI and ML	Conversational AI	Dialogflow		Lifelike conversational AI with state-of-the-art virtual agents.
AI and ML	Conversational AI	Text-to-Speech		Convert text into natural-sounding speech using an API powered by Google's AI technologies.
AI and ML	Conversational AI	Speech-to-Text		Accurately convert speech into text using an API powered by Google's AI technologies.

If you would like to understand more about the business impacts of AI, download the following PDF - [An interpretation of thoughts covered by authors of the Harvard Business Review series on artificial intelligence.](#)



## Networking

### What is Networking?

It's how you connect to Google Cloud.










Google Cloud offers a broad portfolio of networking services built on infrastructure that leverages automation, advanced AI, and programmability, enabling enterprises to connect, scale, secure, modernize and optimize their infrastructure.

From **high-performance options such as Dedicated Interconnect and Partner Interconnect**, to **Cloud VPN for lower volume needs**, and even direct and carrier peering options, Google Cloud Hybrid Connectivity has a solution for connecting your infrastructure to the cloud that fits your needs.

Not all of these will be on the exam, I've highlighted the important ones to memorize and have only included the important products for the exam in the [flashcards](#).

Network	Web Application Firewall	Cloud Armor		Help protect your services against DoS and web attacks.
Network	Content Delivery Network	Cloud CDN		Cache your content close to your users using Google's global network.



Network	Virtual Private Network	Cloud VPN		<b>Cloud Virtual Private Network. Securely extend your on-premises network to Google's network through an IPsec VPN tunnel.</b>
Network	Cloud Domains and DNS	Cloud DNS		Publish and manage your domain names using Google's reliable, resilient, low-latency DNS serving.
Network	Virtual Networks	Cloud VPC		<b>Virtual Private Cloud (VPC) is a logically isolated section of the Google Cloud Network where you can launch Google Cloud resources. You choose a range of IPs using the CIDR Range.</b>
Network	Load Balancer	Cloud Load Balancing		Scale and distribute app access with high-performance load balancing.
Network	Virtual Networks	Cloud NAT		Provision application instances without public IP addresses while allowing them to access the internet.
Network	Virtual Networks	Cloud Interconnect		<b>Low latency, high availability connections that enable you to reliably transfer data between your on-premises and Google Cloud Virtual Private Cloud (VPC) networks.</b>
Network	Service Mesh	Cloud Router		Dynamically exchange routes between your Google Cloud Virtual Private Cloud (VPC) network and your on-premises networks using Border Gateway Protocol (BGP).
Network	Service Mesh	Traffic Director		Deploy global load balancing across clusters and configure sophisticated traffic control policies for open service mesh.
Network	Network Monitoring	Network Intelligence Center	N/A	Use a single console for comprehensive network monitoring, verification, and optimization.
Network	Premium Networking	Network ServiceTiers		Optimize your network for performance or cost.
Network	Optimization	Network Telemetry	N/A	Track network flows for monitoring, forensics, real-time security analysis, and expense optimization.

## API Management and Apigee

### What is an API?

An application programming interface is a connection between computers or between computer programs. It is a type of software interface, offering a service to other pieces of software. In other words, it is a software intermediary that allows two applications to talk to each other.

## What is Apigee?

Apigee Corp was an API management and predictive analytics software provider that Google Cloud acquired in 2016. Apigee is also a founding member of the OpenAPI Initiative. OpenAPI specification is an open-sourced standard for writing declarative structure of an Application Programming Interface. All Cloud Service Providers (CSPs) will have a fully-managed API service offering known as an API Gateway that will typically support OpenAPI standard.

API Management	API Gateway	Apigee API Platform		Develop, secure, deploy, and monitor your APIs everywhere. Expensive, but has many advanced features including, analytics, monetization, Apigee Sense, Apigee Hybrid and Cloud Healthcare API.
API Management	Feature of Apigee API Platform	API Analytics		Get insight into operational and business metrics for your APIs.
API Management	Feature of Apigee API Platform	API Monetization		Realize value from your APIs with a flexible, easy-to-use solution.
API Management	Feature of Apigee API Platform	Apigee Sense		Add intelligent behavior detection to protect APIs from attacks.
API Management	Feature of Apigee API Platform	Apigee Hybrid	N/A	Manage APIs on-premises, on Google Cloud, or in a hybrid environment.
API Management	Feature of Apigee API Platform	Cloud Healthcare API	N/A	Help secure APIs that power actionable healthcare insights.
API Management	Cheaper API Gateway	Cloud Endpoints		Develop, deploy, and manage APIs on Google Cloud. Cheap and simple, good integrations with App Engine.
API Management	Feature of Cloud Endpoints	Developer Portal		Create a lightweight portal that enables developers and API teams, using a turnkey self-service platform.

## IoT

By far the most extensive product list thus far in this study guide, with exactly one service, IoT Core. Which is a service to manage your connected devices in your network,

IoT	Device Management	IoT Core		IoT Core is a fully managed service that allows you to easily and securely connect, manage, and ingest data from millions of globally dispersed devices. IoT Core, in combination with other services on Google Cloud, provides a complete solution for collecting, processing,
-----	-------------------	----------	--	---

				analyzing, and visualizing IoT data in real time to support improved operational efficiency.
--	--	--	--	--

## Developer Tools

There are loads of Developer Tools and they will most likely not be on the exam, or if they are, we have covered them in other sections. Aside from the main ones, there are many others including ones exclusively for Firebase (which we will talk about later).

**Here are the main ones:**

Dev Tools	Container registry	Artifact Registry		Store, manage, and secure container images and language packages.
Dev Tools	CI/CD	Cloud Build		Continuously build, test, and deploy containers, Java archives, and more using the Google Cloud infrastructure.
Dev Tools	Cloud Development IDE Plugin	Cloud Code		Extend your IDE with tools to write, debug, and deploy Kubernetes applications.
Dev Tools	Job Scheduling	Cloud Scheduler		Schedule batch jobs, big data jobs, and cloud infrastructure operations using a fully managed cron job service.
Dev Tools	Command-line Interface (CLI)	Cloud SDK		Install a command-line interface to script and manage Google Cloud products from your own computer.
Dev Tools	Repository	Cloud Source Repositories		Manage code and extend your Git workflow by connecting to Cloud Build, App Engine, Cloud Logging, Cloud Monitoring, Pub/Sub, and more.
Dev Tools	Execution Control	Cloud Tasks		Asynchronously execute, dispatch, and deliver distributed tasks.
Dev Tools	Container Registry	Container Registry		Store, manage, and secure your Docker container images.

## Cloud Operations Suite

Google's Operations Suite allows you to **monitor, log, trace, and profile** your apps and services. Not extremely important for the exam, but understand Cloud Monitoring.

Operations	Monitoring	Cloud Monitoring		Cloud Monitoring provides visibility into the performance, availability, and overall health of cloud-powered applications.
Operations	Monitoring	Service Level Monitoring		Define and measure availability, performance, and other service levels for cloud-powered applications.
Operations	Logging	Cloud Logging		Store, search, analyze, monitor, and alert on log data and events from Google Cloud and AWS.
Operations	Error Handling	Error Reporting		Identify and understand application errors.
Operations	Performance Tracing	Cloud Trace		Find performance bottlenecks in production for applications.
Operations	Debugging	Cloud Debugger		Investigate code behavior in production. for applications.
Operations	Profiling	Cloud Profiler		Continuously gather performance information using a low-impact CPU and heap profiling service.

## Migration and Anthos

### Types of Migration

There are **three types** of migrations from **on-premise to the cloud**:

Lift and Shift	Improve and Move	Rip and Replace
<ul style="list-style-type: none"> <li>- <b>Little to no modification</b></li> <li>- Taking least advantage of the cloud</li> <li>- Fastest migration strategy</li> <li>- Easier to Implement</li> <li>- Limited Cloud Benefits</li> </ul>	<ul style="list-style-type: none"> <li>- Refactor your existing application</li> <li>- Take advantage of most of the cloud offerings</li> <li>- Slow migration process</li> </ul>	<ul style="list-style-type: none"> <li>- Rebuild your app from scratch</li> <li>- Take advantage of the maximum value of cloud offerings</li> <li>- Can take the longest amount of time</li> <li>- Labour Intensive, Full Cloud Benefits</li> </ul>
<b>Ideal when:</b>	<b>Ideal when:</b>	<b>Ideal when:</b>







<ul style="list-style-type: none"> <li>- a workload can operate as-is in the target environment</li> <li>- little or no business need for change</li> </ul>	<ul style="list-style-type: none"> <li>- architecture or infrastructure of an app isn't supported in the target environment</li> <li>- a major update to the workload is necessary</li> </ul>	<ul style="list-style-type: none"> <li>- current app isn't meeting your goals</li> <li>- you want to remove legacy technical debt</li> </ul>
<p><b>Considerations:</b></p> <ul style="list-style-type: none"> <li>- Requires the least amount of time because the amount of refactoring is kept to a minimum</li> <li>- The team can continue to use the same set of tools and skills that they were using before</li> <li>- Doesn't take full advantage of cloud platform features:</li> <li>- horizontal scalability</li> <li>- fine-grained pricing</li> <li>- highly managed services</li> </ul>	<p><b>Considerations:</b></p> <ul style="list-style-type: none"> <li>- take longer than lift and shift migrations</li> <li>- must be refactored in order for the app to migrate</li> <li>- extra time and effort as part of the life cycle of the app</li> <li>- requires that you learn new skills</li> </ul>	<p><b>Considerations:</b></p> <ul style="list-style-type: none"> <li>- Requires the most amount of time to develop</li> <li>- Requires the most amount of learning</li> </ul>

## The Four Paths to Migration

1. **Assess.** perform a thorough assessment and discovery of your existing environment in order to understand your app and environment inventory, identify app dependencies and requirements, perform total cost of ownership calculations, and establish app performance benchmarks.
2. **Plan.** create the basic cloud infrastructure for your workloads to live in and plan how you will move apps. This planning includes identity management, organization and project structure, networking, sorting your apps, and developing a prioritized migration strategy.
3. **Deploy.** design, implement and execute a deployment process to move workloads to Google Cloud. You might also have to refine your cloud infrastructure to deal with new needs.
4. **Optimize.** begin to take full advantage of cloud-native technologies and capabilities to expand your business's potential to things such as performance, scalability, disaster recovery, costs, training, as well as opening the doors to machine learning and artificial intelligence integrations for your app.

## Google Cloud Migration Services Overview

Migration	SQL Database Migration	Database Migration Service (DMS)		<p>Migrating open-source relational databases like SQL. Serverless, easy, minimal downtime migrations to Cloud SQL.</p>
-----------	------------------------	----------------------------------	--	---

Migration	Data Transfer Service	BigQuery Data Transfer Service		Automate scheduled data movement into BigQuery using a fully managed data import service.
Migration	VM or Server Migration	Migration for Compute Engine		Migrate servers and VMs from on-premises or another cloud to Compute Engine.
Migration	Container Migration	Migrate for Anthos		Migrate VMs from on-premises or other clouds directly into containers in GKE.
Migration	Virtual Storage Migration	Cloud Storage Transfer Service		Transfer data between cloud storage services such as AWS S3 and Cloud Storage.
Migration	Physical Storage Migration	Transfer Appliance		Ship large volumes of data to Google Cloud using rackable storage or large physical drives.
Migration	Multi-cloud Managed Application Platform	Anthos		Modern application management platform used for managing hybrid architectures that span from Google Cloud to another public cloud (AWS and Azure) or on-premise data centers running VMWare.

## Anthos

### Anthos Migrate VS Migrate for Compute Engine

**Migrate for Compute Engine** enables you to migrate (Lift and Shift) your virtual machines (VMs), with minor automatic modifications, from your source environment to Google Compute Engine

- continuously replicates disk data from the source VMs to Google Cloud
- no downtime on the source via transfer
- quickly **clone and test** a migrated VM **test clones**
- easily perform all migration tasks within the Google Cloud Console

**Migrate for Anthos and Google Kubernetes Engine (GKE)** is a tool to move and automatically convert workloads directly into containers in Google Kubernetes Engine (GKE) and Anthos

With Migrate for Anthos, you can migrate your VMs from supported source platforms to:

- Google Kubernetes Engine (GKE)
- Anthos
- Anthos clusters on VMware
- Anthos clusters on AWS

## Storage Transfer Service Vs Transfer Application

### Serverless

#### What is Serverless?

Serverless architectures are fully-managed services that automatically scale, are highly available, durable, and secure by default.

Abstracts away the underlying infrastructure and are billed based on the execution of your business task. **Pay-for-Value** (you don't pay for idle servers ).

Serverless can **Scale-to-Zero** meaning when not in use they cost nothing.

#### Examples of Google Cloud Serverless Products:

- **Cloud Functions** Function as a Service
- **Cloud Run** Serverless containers
- **App Engine** Platform as a Service
- **Knative** Serverless K8 containers
- **Workflows** Serverless State Machine
- **BigQuery** Serverless Data-Warehouse
- **Cloud Storage** Serverless Storage

### Firestore

**Firestore** is an alternative to Google Cloud for users who want to focus on building and deploying their applications in a highly opinionated framework. It is a separate product that Google Cloud bought in 2014, although it runs on Google Cloud and can be either managed from the Firestore platform OR the Google Cloud Console.

**Firestore** is Google's fully-managed platform for rapidly developing and deploying web and mobile applications. It is a **Platform as a Service** utilizing *Serverless technology*. Firestore offers the following services and features:

- Cloud Firestore
- Machine Learning
- Cloud Functions
- Authentication- Hosting
- Cloud Storage
- Realtime Database
- Crashlytics
- Performance Monitoring
- Test Lab
- Google Analytics
- In-App Messaging
- Predictions
- A/B Testing
- Cloud Messaging
- Remote Config
- Dynamic Links

## Cloud Deployment Manager

### What is Infrastructure as Code?

Infrastructure as code is the process of managing and provisioning computer data centers through machine-readable definition files, rather than physical hardware configuration or interactive configuration tools.

**Infrastructure as Code (IaC)** is the **process of managing and provisioning cloud services** through machine-readable definition files (eg, YAML, JSON files) rather than manual configuration

**Google Cloud Deployment Manager is an infrastructure deployment service** that automates the creation and management of Google Cloud resources. Write flexible template and configuration files and use them to create deployments that have a variety of Google Cloud services, such as Cloud Storage, Compute Engine, and Cloud SQL, configured to work together.

One of the most well known and recognizable IaCs is Hashicorp [Terraform](#).



## Internal Services

These are Google Cloud's internal services. Internal services are the underlying infrastructure to many Google cloud services. These are not services you can use as a consumer, it is what Google uses on the backend. These also most likely won't be direct questions on the exam, but knowing these could help you eliminate wrong answers,

**Spanner:** Globally-consistent, scalable relational database. Cloud Spanner is the external offering of this service.



**Borg:** A cluster manager that runs hundreds of thousands of jobs, from many thousands of different applications, across a number of clusters each with up to tens of thousands of machines.

**Chubby:** A distributed lock manager (DLM) is a service that temporarily prevents files and records from being used by another user or operation on a Virtual Machine

**Colossus:** Cluster-level file system, successor to the Google File System (GFS) provides the underlying infrastructure for all Google Cloud storage services, from Firestore to Cloud SQL to Filestore, and Cloud Storage.

## Chapter Five | Security and Identity Management

Google Cloud prides itself on being highly secure and operating under a zero-trust model meaning that it operates under the idea that no one should have any more access than exactly what they need to perform their job.

### Identity Management

#### What is Identity Management?

Identity management, also known as identity and access management, is a framework of policies and technologies to ensure that the right users have the appropriate access to technology resources

For the exam, it is important to focus on the following services, Identity and Access Management, Cloud Identity, Identity-Aware Proxy BeyondCorp Enterprise, and Managed Service for Microsoft Active Identity.

**BeyondCorp:** The Zero Trust model operates on the principle of “**trust no one, verify everything.**” By shifting access controls from the network perimeter to individual users, BeyondCorp enables secure work from virtually any location without the need for a traditional VPN.

BeyondCorp itself is just a collection of identity, access, and security services to meet Zero Trust model requirements, and **BeyondCorp Enterprise is a subscription service and it is built into the Chrome Browser.**

#### Access Context Manager

Access Context Manager allows Google Cloud organization admins to define fine-grained, attribute-based access control for projects and resources in Google Cloud. Access Context Manager keeps mobile workforces utilizing Bring-Your-Own-Devices (BYOD) secure.

You create an access policy and determine what level of access based on attributes such as:

- Device Type
- Operating System
- IP Address
- User Identity

## Active Directory

### What is Active Directory?

Microsoft introduced Active Directory Domain Services in Windows 2000 to give organizations the ability to manage multiple on-premises infrastructure components and systems using a single identity per user.

Active Directory Services consist of multiple directory services

### What is a Directory Service?

A directory service is a critical component of a network operating system.

A directory service maps the names of network resources to their network addresses. A directory service is a shared information infrastructure for locating, managing, administering, and organizing resources like Volumes, Folders, Files, Printers, Users, Groups, Devices, Telephone numbers, and other objects.

### Managed Service for Microsoft Active Directory

Managed Service for Microsoft Active Directory (AD) is an Active Directory hosted on the Google Cloud Platform.

## Cloud Identity

Cloud Identity is an Identity as a Service (IDaaS) that centrally manages users and groups **BETWEEN** Google Cloud, Active Directory, Azure AD ETC. **THEN** you would use Identity and Access Management (IAM) to manage access to Google Cloud resources for each Cloud Identity Account.

### Cloud Identity Features and Versions

Device Management	Directory	Security
<p><b>Free</b></p> <ul style="list-style-type: none"> <li>- Basic Mobile Management</li> <li>- Device inventory</li> <li>- Basic passcode enforcement</li> <li>- Remote account wipe</li> <li>- Android</li> <li>- Apple® iOS®</li> </ul>	<p><b>Free</b></p> <ul style="list-style-type: none"> <li>- Basic directory management</li> <li>- Organizational units and groups (Unlimited)</li> <li>- Admin managed groups</li> <li>- Groups for Business</li> <li>- Google Cloud Directory Sync</li> <li>- Admin roles and privileges</li> <li>- Google Admin App for Android</li> <li>- Google Admin App for iOS</li> <li>- Admin SDK/API</li> <li>- Secure LDAP</li> </ul>	<p><b>Free</b></p> <ul style="list-style-type: none"> <li>- User security management</li> <li>- Self-service password recovery</li> <li>- 2-Step Verification (2SV) including security key management</li> <li>- 2SV enforcement controls with security key enforcement and management</li> <li>- Password management and strength alert</li> </ul>
<p><b>Premium</b></p> <ul style="list-style-type: none"> <li>- Advanced Mobile Management</li> <li>- Advanced passcode enforcement</li> <li>- Security policies</li> <li>- Application management</li> <li>- Network management</li> <li>- Remote device wipe</li> <li>- Reporting</li> <li>- Application auditing</li> <li>- Company-owned devices</li> <li>- Mobile audit</li> <li>- MDM rules</li> </ul>	<p><b>Premium</b></p> <ul style="list-style-type: none"> <li>- User lifecycle management (no user cap)</li> <li>- Secure LDAP</li> </ul>	<p><b>Premium</b></p> <ul style="list-style-type: none"> <li>- First-party session management</li> <li>- Google security center</li> </ul>
Single sign-on (SSO) and automated provisioning	Reporting	Service Level Agreements (SLAs)
<p><b>Free</b></p> <ul style="list-style-type: none"> <li>- Set up SSO using Google as an identity provider (IdP) to access a pre-integrated list of third-party SAML apps (Unlimited)</li> <li>- Set up SSO using Google as an IdP to access custom SAML apps</li> <li>- Set up SSO using a third-party IdP with Google as a service provider</li> </ul>	<p><b>Free</b></p> <ul style="list-style-type: none"> <li>- Admin, Login, SAML, Groups, Token audit logs</li> <li>- Security reports</li> <li>- SAML audit log</li> <li>- App reports</li> <li>- Account activity reports</li> </ul>	<ul style="list-style-type: none"> <li>- Premium has 99.9%</li> </ul>
<p><b>Premium</b></p>	<p><b>Premium</b></p>	

- Automated user provisioning	- Devices audit log - Auto export audit logs to BigQuery	
-------------------------------	---	--

### Identity Providers (IdP)

Identity Provider (IdP) is a system entity that creates, maintains, and manages identity information for principals and also provides authentication services to applications within a federation or distributed network.

A trusted provider of your user identity that lets you use to authenticate to access other services, but does not share passwords. Identity Providers could be: Facebook, Amazon, Google, Twitter, Github, LinkedIn.

### Single-Sign-On (SSO)

Single sign-on (SSO) is an authentication scheme that allows a user to log in with a single ID and password to different systems and software.

SSO allows IT departments to administrator a single identity that can access many machines and cloud services.

Login for SSO is seamless, where once a user is logged into their primary directory, as soon as they utilize this software, they are presented with a login screen.

### Lightweight Directory Access Protocol (LDAP)

Lightweight Directory Access Protocol (LDAP) is an open, vendor-neutral, industry standard application protocol for accessing and maintaining distributed directory information services over an Internet Protocol (IP) network. *This is a sign-in protocol mainly for things outside of web applications like Jenkins or Kubernetes or on-premises infrastructure tools.*

LDAP enables for same-sign on. The same sign-on allows the same userID and password to be used for multiple applications, but they have to enter it in every time they want to login.

Why use LDAP when SSO is more convenient?

- Most SSO systems are using LDAP.
- LDAP was not designed natively to work with web applications.
- Some systems only support integration with LDAP and not SSO





### Google Cloud Directory Sync

Google Cloud Directory Sync enables administrators to synchronize users, groups, and other data from an Active Directory/LDAP service to their Managed Service for Microsoft Active Directory within Google

## Cloud Billing IAM Roles

Cloud Billing lets you control which users have administrative and cost viewing permissions for specified resources by setting Identity and Access Management (IAM) policies on the resources

To grant or limit access to Cloud Billing, you can set an IAM policy at the organization level, the Cloud Billing account level, and/or the project level.

Identity and Access	IAM	Identity and Access Management		Establish fine-grained identity and access management for Google Cloud resources.
Identity and Access	IAM	Cloud Identity	N/A	Easily manage user identities, devices, and applications from one console.
Identity and Access	IAM	Identity Platform	N/A	Add Google-grade identity and access management to your apps.
Identity and Access	Zero-trust	BeyondCorp Enterprise		A zero-trust solution that enables secure access with integrated threat and data protection.
Identity and Access	IAM	Identity-Aware Proxy		Use identity and context to guard access to your applications and VMs. Application level access instead of relying on network-level firewalls.
Identity and Access	IAM	Managed Service for Microsoft Active Directory	N/A	Use a highly available, hardened service running Microsoft Active Directory (AD).
Identity and Access	Resource Monitoring	Resource Manager	N/A	Hierarchically manage resources on Google Cloud.
Identity and Access	Secrets Management	Security Enforcement		Enforce the use of security keys to help prevent account takeovers.
Identity and Access	IAM	Titan Security Keys	N/A	A physical USB Key. Titan Security Keys help prevent phishing and keep out anyone who shouldn't have access to your online accounts, including Gmail.

## Security

## Google's Secure-by-Design Infrastructure and Encryption

Google has a global scale technical infrastructure designed to provide security through the entire information processing life cycle at Google. This infrastructure provides secure deployment of services, secure storage of data with end user privacy safeguards, secure communications between services, secure and private communication with customers over the internet, and safe operation by administrators.

Google uses this infrastructure to build its internet services, including both consumer services such as Search, Gmail, and Photos, and enterprise services such as G Suite and Google Cloud.

The security of the infrastructure is designed in progressive layers starting from the physical security of data centers, continuing on to the security of the hardware and software that underlie the infrastructure, and finally, the technical constraints and processes in place to support operational security.

Google invests heavily in securing its infrastructure with many hundreds of engineers dedicated to security and privacy distributed across all of Google, including many who are recognized industry authorities.

For the exam, it is important to focus on the following services, Cloud Data Loss Prevention and Security Command Center.

**Security Command Center** is a centralized security and risk management platform for your Google Cloud resources. It is a single tool that offers a variety of security features including:

### Asset discovery and inventory

Inventory and historical information about your google cloud resources

### Threat detection

Audits your cloud resources for security vulnerability

### Threat prevention

Fix security misconfiguration with single-click remediation

**Cloud Data Loss Prevention** is a fully managed service designed to help you discover, classify, and protect your most sensitive data. Sensitive data includes **Personally Identifiable Information (PII)** like birthdates, government IDs, addresses etc and **Personally Protected Health Information (PHI)** which is any health related information stored from patients from their practice.

Security	Cloud provider access management	Access Transparency	N/A	Get visibility over your cloud provider through near real-time logs.
Security	Container Security	Binary Authorization	N/A	Deploy only trusted containers on Kubernetes Engine.
Security	Resource Monitoring	Cloud Asset Inventory	N/A	View, monitor, and analyze Google Cloud and Anthos assets across projects and services.
Security	Audit Logging	Cloud Audit Logs	N/A	Gain visibility into who did what, when, and where for all user activity on Google Cloud.
Security	DLP	Cloud Data Loss Prevention		<b>Discover and redact sensitive data.</b>
Security	Security and Risk Management	Security Command Center		<b>Understand your security and data attack surface.</b>
Security	VM Security	Shielded VMs	N/A	Deploy hardened virtual machines on Google Cloud.
Security	Exfiltration Prevention	VPC Service Controls	N/A	Protect sensitive data in Google Cloud services using security perimeters. Functions like a firewall for GCP APIs.
Security	Incident Monitoring	Incident Response Management	N/A	Improve your incident median time to mitigation.

If you are interested in learning more about the business impacts of cybersecurity, download the PDF - [Top Ten Takaways from the Harvard Business Review on Cybersecurity](#).



## Google Cloud Enterprise Privacy Commitments

Google Cloud Enterprise Privacy Commitments describe how we protect the privacy of Google Cloud Platform and Google Workspace customers

### **1) You control your data**

Customer data is your data, not Google's. We only process your data according to your agreement(s).

### **2) We never use your data for ads targeting**

We do not process your customer data to create ads profiles or improve Google Ads products.

### **3) We are transparent about data collection and use**

We're committed to transparency, compliance with regulations like the GDPR, and privacy best practices.

### **4) We never sell customer data or service data**

We never sell customer data or service data to third parties.

### **5) Security and privacy are primary design criteria for all of our products**

Prioritizing the privacy of our customers means protecting the data you trust us with. We build the strongest security technologies into our products. Google provides resources on privacy regulations such as the LGPD, GDPR, CCPA, the Australian Privacy Act, My Number Act, and PIPEDA, among others.

## Google Cloud Trust Principles

1. You own your data, not Google
2. Google does not sell customer data to third parties
3. Google Cloud does not use customer data for advertising
4. All customer data is encrypted by default
5. We guard against insider access to your data
6. We never give any government entity "backdoor" access
7. Our privacy practices are audited against international standards



## Chapter Six | Support

### Service Level Agreements (SLAs)

#### **What is a Service Level Agreement (SLA)?**

A SLA is a formal commitment about the expected level of service between a customer and provider. When a service level is not met and if the Customer meets its obligations under the SLA, Customer will be eligible to receive the compensation eg. Financial or Service Credits

#### **What is a Service Level Indicator (SLI)?**

A metric/measurement that indicates what measure of performance a customer is receiving at a given time. A SLI metric could be uptime, performance, availability, throughput, latency, error rate, durability, correctness.

#### **What is a Service Level Objective (SLO)?**

The objective that the provider has agreed to meet. SLOs are represented as a specific target percentage over a period of time.

Availability SLA of 99.99% in a period of 3 months

## Support Plans & Services

Basic Support	Standard Support	Enhanced Support	Premium Support
	Unlimited access to support		
Billing Support: Case (Email), Phone, and Chat			
	P2: 4-hour response	P1: 1-hour response	P1: 15min response
	Technical Support: Case (Email)	Technical Support: Case (Email) and Phone	
	8/5 response for high-impact issues	24/7 response for high-impact and critical issues	
	Only English Support	English, Japanese, Mandarin Chinese, and Korean	
		<ul style="list-style-type: none"> <li>Active Assist Recommender API</li> <li>Third Party Support</li> <li>Cloud Support API</li> <li>Technical support escalation</li> </ul>	
		<ul style="list-style-type: none"> <li>Access to purchase <b>Technical Account Advisor Service (TAAS)</b></li> </ul>	<ul style="list-style-type: none"> <li><b>Technical Account Manager (TAM)</b> </li> <li>Event management service</li> <li>Operational Health Reviews</li> <li>Customer Aware Support</li> <li>New Product previews</li> <li>Training credits</li> <li>Access to purchase <b>Mission Critical Services</b></li> <li>Access to purchase <b>Assured Support</b></li> </ul>
<b>FREE</b>	<b>\$29</b> per month + 3% net spend	<b>\$500</b> per month + 3% net spend	<b>\$\$\$\$</b> Contact Sales

### Active Assist

Active Assist, a portfolio of intelligent tools and capabilities to help actively assist you in managing complexity in your cloud operations.

Helps with 3 key activities:

- making proactive improvements to your cloud with smart recommendations
- preventing mistakes from happening in the first place by giving you a better analysis
- helping you figure out why something went wrong by using intuitive troubleshooting tools

### Technical Account Advisory Service

Technical Account Advisor Service (TAAS) provides both proactive guidance and reactive support to help you succeed with your Cloud journey.

TAAS deliver the following services:

- Guided onboarding to help you get started with Enhanced Support and set up your operations with Google Cloud.
- Best practices and additional support for your most critical cases, including proactive monitoring and guidance on case escalation.
- Monthly, quarterly, and yearly reviews to assess your operational health across Google Cloud and deliver recommendations for improving your usage of Enhanced Support.

- Recommended training paths and courses tailored to your organization's needs.

When you purchase TAAS, you pay a monthly fee, with a minimum 1-year contract. After the first year, your contract is month-to-month.

### **TAM (Technical Account Manager)**

As a Premium Support customer, you are assigned a named Technical Account Manager (TAM). Technical Account Managers are trusted technical advisors that focus on operational rigor, platform health, and architectural stability for your organization.

Your Technical Account Manager supports and guides you in the following ways:

- Assists you with onboarding to Premium Support.
- Assesses your cloud maturity and works with you to create an adoption roadmap and operating model.
- Advises on best practices for using Google Cloud.
- Delivers frequent Operational Health Reviews.
- Connects you with Google technical experts, such as Product Managers and Support Engineers.
- Works with you on support cases and case escalations. For high-priority cases, your TAM analyzes the incident and identifies root causes.
- 

By default, you receive 8 hours per week of foundational technical account management services. If you require more assistance, you can purchase additional TAM services

### **Operational Health Reviews**

Operational Health Reviews are essentially quarterly business reviews (QBRs) with your TAM that help you measure your progress and proactively address blockers to your goals with Google Cloud.

The reviews serve as a regular touchpoint with your TAM where you can discuss various topics related to your Customer Care experience, including:

- The efficiency of your cloud operations, including support trends.
- Analysis of trends in operational metrics.
- Incidents, case escalations, and outages.
- Tracking of open cases.
- Status reports for high-priority Cloud projects.

### **Event Management Service**

Premium Support's Event Management Service for planned peak events, such as a product launch or major sales event. With this service, Customer Care partners with your team to create a plan and provide guidance throughout the event.

With Event Management Service, your team is supported with the following tasks:

- Preparing your systems for key moments and heavy workloads.
- Running disaster tests to proactively resolve potential issues.
- Developing and implementing a faster path to resolution to reduce the impact of any issues that might occur.
- 

After the event, your TAM works with you to review the outcomes and make recommendations for future events.

### **Training Credits**

With Premium Support, you receive training credits for the Google Cloud Qwiklabs that you can distribute to users in your organization. Qwiklabs is a demo environment or lab that your team can explore Google Cloud on without accidentally increasing your monthly cloud bill. Your TAM identifies learning opportunities and indicates which training resources can be most beneficial to your organization. With this training, your developers have the resources to find answers quickly and test out ideas in safe environments.

For each 1-year contract with Premium Support, you receive 6,250 credits.

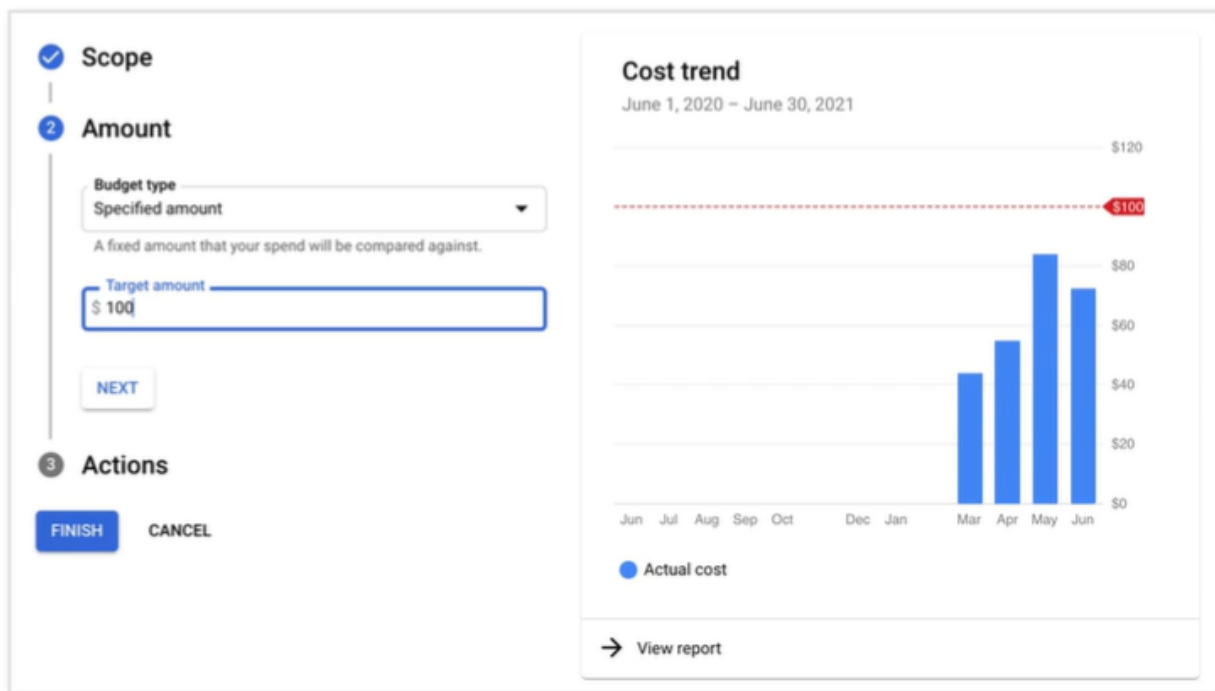
## Chapter Seven | Billing

There are two types of billing accounts, a Cloud Billing Account and Payments Profile. Payment Profile can have multiple Cloud Billing Accounts and is ideal for Managed Service Providers who manage multiple customers on Google Cloud. Each customer would have its own Cloud Billing Account under the Payment Profile. This can also be used to separate projects within one organization.

### Billing Health Checks and Alerts

**Billing Health Checks:** are recommendations to avoid common billing issues, like set up budget alerts, grant access to view billing accounts, or assign multiple billing accounts.

**Budget Alerts:** set up multiple alert thresholds to reduce spending surprises and unexpected cost overruns. They work by setting a budget and alerting you when you are close to spending that amount.



## Billing Services

**There are four main Billing Services for Google Cloud, and these are important to know for the exam. There will be questions asking you specific customer needs/wants in the below chart.**

Billing Reports	Cost Table Reports
<p>Use the billing report to view and analyze your Google Cloud usage costs using many selectable settings and filters. <b>Essentially a visual dashboard that analyzes your cloud spend.</b></p> <p>Configuring various views of the Cloud Billing report can help you answer questions like these:</p> <ul style="list-style-type: none"> <li>- How is my current month's Google Cloud spending trending?</li> <li>- What Google Cloud project cost the most last month?</li> <li>- What Google Cloud service (for example, Compute Engine or Cloud Storage) cost me the most?</li> <li>- What are my forecasted future costs based on historical trends?</li> <li>- How much am I spending by region?</li> <li>- What was the cost of resources with label X?</li> </ul>	<p>Use the cost table report to access and analyze the details of your invoices and statements. <b>Instead of a PDF invoice, this report breaks down the costs in an online table view.</b></p> <p>Because your generated invoice and statement PDFs only contain simplified, summarized views of your costs, the cost table report is available to provide invoice or statement cost details, such as the following:</p> <ul style="list-style-type: none"> <li>- Includes project-level cost details from your invoices and statements, including your tax costs broken out by project.</li> <li>- Includes additional details you might need, such as service IDs, SKU IDs, and project numbers.</li> <li>- The report view is customizable and downloadable to CSV.</li> </ul>

<ul style="list-style-type: none"> <li>- Your customized report views are saveable and shareable.</li> </ul>	
<b>Cost Breakdown Report</b>	<b>Pricing Report</b>
<p>Use the cost breakdown report for an <b>at-a-glance waterfall overview of your monthly costs and savings</b>.</p> <p><b>This report shows the following summarized view of monthly charges and credits:</b></p> <ul style="list-style-type: none"> <li>- The combined costs of your monthly Google Cloud usage at the on-demand rate, calculated using non-discounted list prices.</li> <li>- Savings realized on your invoice due to negotiated pricing (if applicable to your Cloud Billing account).</li> <li>- Savings earned on your invoice with usage-based credits, broken down by credit type (for example, committed use discounts, sustained use discounts, free tier usage).</li> <li>- Your invoice-level charges such as tax and adjustments (if any) applied for that invoice month.</li> </ul>	<p>Use the pricing table report to access SKU prices for Google's cloud services, including <b>Google Cloud, Google Maps Platform, and Google Workspace</b>, as of the date the report is viewed.</p> <p><b>This report shows the following pricing information:</b></p> <ul style="list-style-type: none"> <li>- Displays SKU prices specific to the selected Cloud Billing account.</li> <li>- If your Cloud Billing account has negotiated contract pricing, each SKU displays the list price, your contract price, and your effective discount.</li> <li>- If a SKU is subject to tiered pricing, each pricing tier for a SKU is listed as a separate row.</li> <li>- All the prices are shown in the currency of the selected billing account.</li> <li>- The report view is customizable and downloadable to CSV for offline analysis.</li> </ul>

## Chapter Eight | Pricing

Google Cloud offers various different pricing schemas that vary per service. **There are seven main types of Google Cloud pricing:**

### Free-Trial

New Google Cloud and Google Maps Platform users can take advantage of a 90-day trial period that includes \$300 in free Cloud Billing credits to explore and evaluate Google Cloud and Google Maps Platform products and services. You can use these credits toward one or a combination of products.

Trial Limitations:

- You can't add GPUs to your VM instances
- You can't request a quota increase
- You can't create VM instances that are based on Windows Server images.
- You need to verify a credit card or other payment methods to signup

- At end of the trial to continue using Google Cloud, you must upgrade to a paid Cloud Billing account.

## Free-Tier

All Google Cloud customers can use select Google Cloud products—like Compute Engine, Cloud Storage, and BigQuery—free of charge, within specified monthly usage limits. For Example, AutoML Translate lets you do 500,000 translated characters per month or you can do 1 TB of querying per month on Big Query. For more information, view the Google Cloud [Free Usage Limits](#).

## On-Demand

On-demand pricing is when you pay for a google cloud resource based on a consumption-based model. A consumption-based model means you only pay for what you use, based on a consumption metric:

- By time: hourly, minutes, seconds, milliseconds
- Can be multiplied by configuration variables: vCPUs and Mem
- By API calls: \$1 every 1000 transactions

### **On Demand is ideal for:**

- Low cost and flexible
- Only pay per hour
- Short-term, spiky, unpredictable workloads
- Cannot be interrupted
- For first time apps

## Committed Use Discounts (CUDs)

Committed Use Discounts (CUD) lets us commit to a contract for deeply discounted Virtual Machines on Google Compute Engine.

- Simple and flexible, and require no upfront costs
- Ideal for workloads with predictable resource needs
- You purchase compute resources (vCPUs, memory, GPUs, and local SSDs)
- Discounts apply to the aggregate number of vCPUs, memory, GPUs, and local SSDs within a region

- Not affected by changes to your instance's machine setup
- You commit for payment terms of 1 Year to 3 Years
  
- Purchase a committed use contract for a single project
  
- Purchase multiple contracts share across many projects by enabling Shared Discounts
  
- You are billed monthly for the resources you purchased for the duration of the term whether or not you use the services

## Sustained Use Discounts (SUDs)

Sustained use discounts are automatic discounts for running specific Compute Engine resources for a significant portion of the billing month.

Sustained use discounts apply to the following resources:

- GPU devices

The vCPUs and memory for:

- General-purpose custom and predefined machine types
- Compute-optimized machine types
- Memory-optimized machine types
- Sole-tenant nodes
- 10% premium cost even if the vCPUs and memory in those nodes are covered by committed use discounts

GPU devices

- Applied on incremental use after you reach certain usage thresholds
- You pay only for the number of minutes that you use an instance,
- Compute Engine automatically gives you the best price
- There's no reason to run an instance for longer than you need it.
- Automatically apply to VMs created by both Google Kubernetes Engine and Compute Engine

It does not apply to:

- VMs created using the App Engine flexible environment and Dataflow.
- E2 and A2 machine types.

## Preemptible VM Instances or Sole-Tenant Node Pricing

A sole-tenant node (single tenant VM) is a physical Compute Engine server that is dedicated to hosting only your project's VM instances.



When you create sole-tenant nodes, you are billed for all of the vCPU and memory resources on the sole-tenant nodes, plus a sole-tenancy premium, which is 10% of the cost of all of the underlying vCPU and memory resources.

- Sustained use discounts apply to this premium, but committed use discounts do not.
- After you create the node, you can place VMs on that node, and these VMs run for no additional cost.
- vCPUs and GB of memory are charged a minimum of 1 minute.
- After 1 minute of use, sole-tenant nodes are billed in 1 second increments

The price of a node type depends on the following:

- Number of vCPUs of the node type
- GBs of memory of the node type
- Region where you create the node

## Flat-Rate Pricing

**BigQuery offers flat-rate pricing for high-volume or enterprise customers who prefer a stable monthly cost for queries rather than paying the on-demand price per GB of data processed**

When you enroll in flat-rate pricing, you purchase dedicated query processing capacity, measured in BigQuery slots.

Your queries consume this capacity, and you are not billed for bytes processed. If your capacity demands exceed your committed capacity, BigQuery will queue up slots, and you will not be charged additional fees.

To enable flat-rate pricing, use BigQuery Reservations.

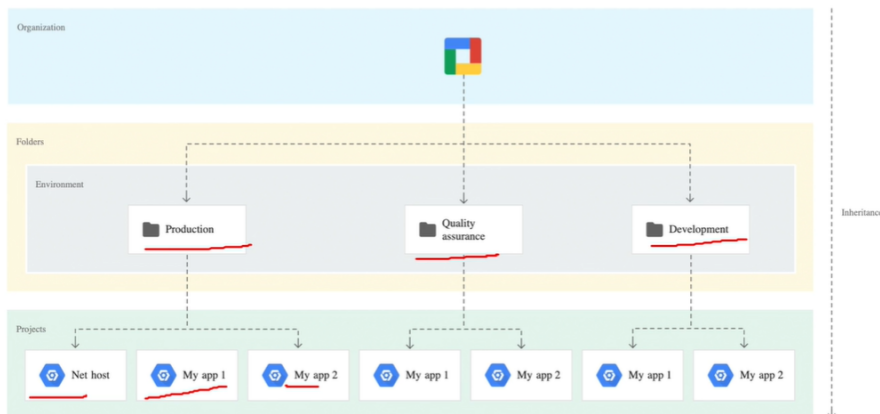
## Chapter Nine | Resource Hierarchy

This is essentially how you organize your Google Cloud resources and it allows you to provide users with specific access to certain projects or folders (workloads). Helps you set up permissions and billing accounts.

### Environment-Oriented Hierarchy

**Use this Hierarchy when:**

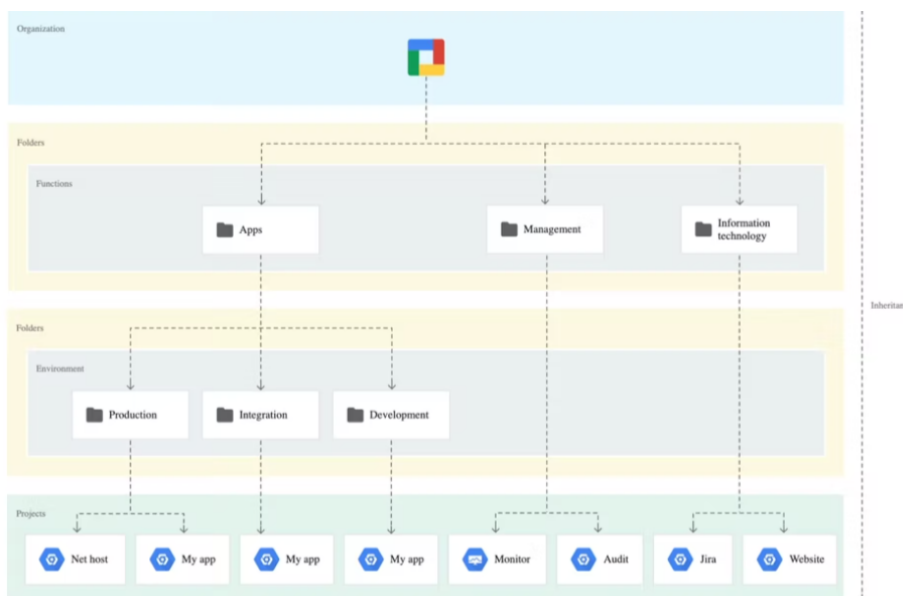
- You have one organization that contains one folder per environment simple to implement
- Can pose challenges if you have to deploy services that are shared by multiple environments



## Function-Oriented Hierarchy

### Use this Hierarchy when:

- one organization that contains one folder per business function
- Each business function folder can contain multiple environment folders
- multiple business functions are apps, management, and information technology
- more flexible compared to environment-oriented
- gives you the same environment separation
- allows you to deploy shared services



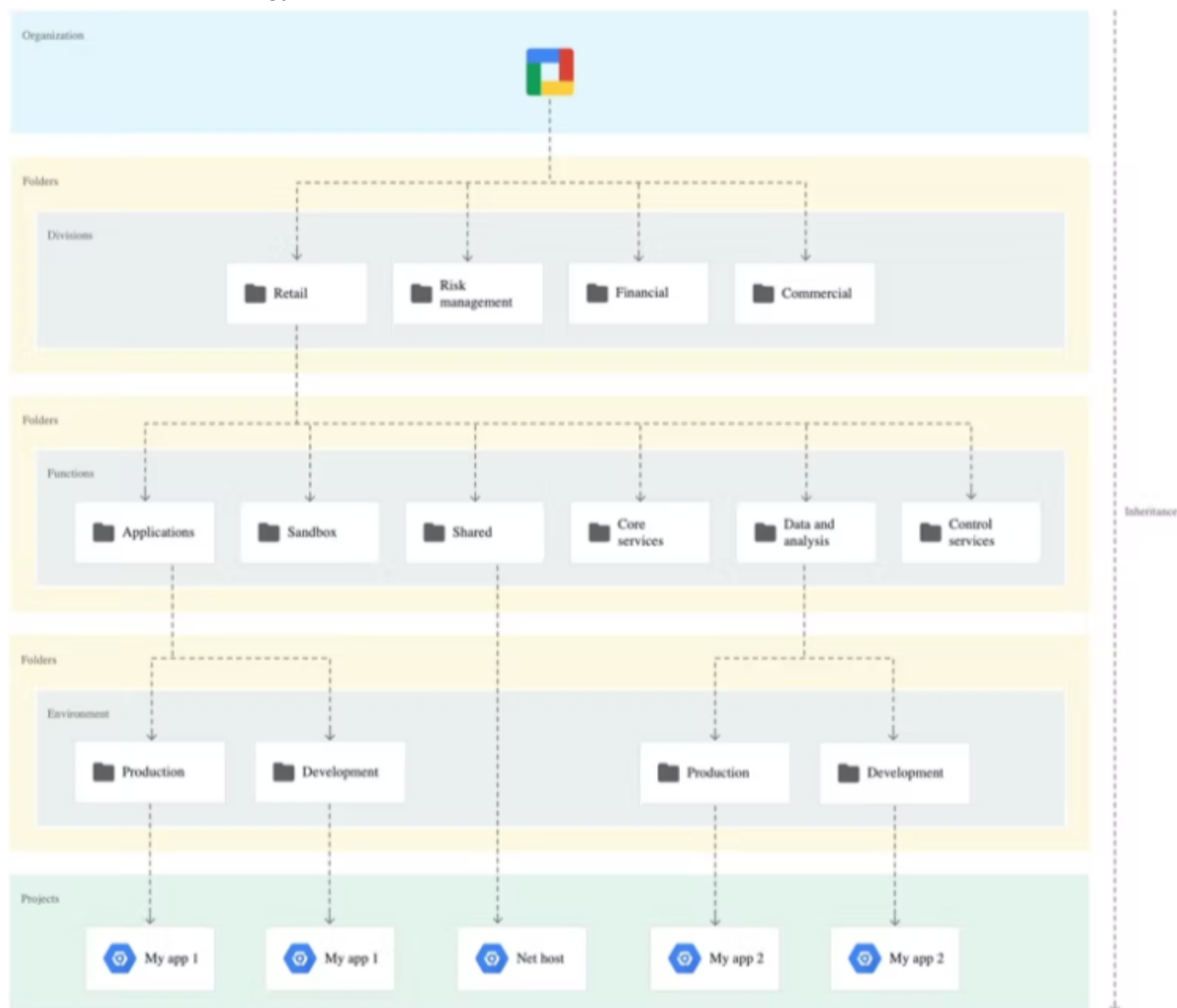
- function-oriented hierarchy is more complex to manage than an environment-oriented

- separate access by business

## Granular-Access-Oriented Hierarchy

### Use this Hierarchy when:

- one organization that contains one folder per business unit
- Each business unit folder can contain one folder per business function
- Each business function folder can contain one folder per environment
- most flexible and extensible option
- you need to spend a greater effort to manage the structure, roles, and permissions
- network topology is more complex



## Chapter Ten | Practice Exam Q&A

These are the same questions provided by [Google's practice exam](#). Try to block the answers and answer them on your own. Or read through the questions and answers and take Andrew Brown's practice exams.

1. An organization wants to move from a strategic cloud adoption maturity level to a transformational one. How should the organization change the way they scale?
  - A) Review changes manually.
  - B) Deploy changes programmatically.
  - C) Deploy changes when problems arise.

**Answer:**

A) is not correct because this is a tactical approach whereby changes are often high-risk and deployed infrequently.

**B) is correct because automation is a transformational approach which ensures changes are constant and low-risk.**

C) is not correct because this approach is not transformational because it does not mitigate problems from happening in the first place.

2. A video game organization has invested in cloud technology to generate insights from user behaviors. They want to ensure recommendations of games are aligned to players' interests. What may have prompted this business decision?
  - A) Customers expect a personalized experience.
  - B) Employees expect source code changes to be deployed faster.
  - C) Employees expect more predictable data management spending.
  - D) Customers expect faster time to market for games.

**Answer:**

**A) is correct because in the cloud era, users expect more personalization and customization.**

B) is not correct because though this may be true, it is unrelated to the business decision they made.

C) is not correct because generating user insights using data will not increase predictability in financial data spend; it is not directly related to the business decision.

D) is not correct because reduced time-to-market is unrelated to the business decision.

3. An organization has completely migrated all their infrastructure to the cloud to benefit from its agility. Now they want to innovate faster and achieve a higher return on investment. What should the organization do?

- A) Lower their service level objective (SLO).
- B) Modernize their applications.
- C) Manually provision all cloud infrastructure for increased control.
- D) Move to a hybrid architecture with some of their infrastructure on-premises.

**Answer:**

A) is not correct because this will increase the error budget but not their financial budget.

**B) is correct because this will enable the business to better serve their users.**

C) is not correct because this is a tactical business approach to scale that is slower, error prone and unlikely to increase innovation.

D) is not correct because returning to some on-premises infrastructure could burden teams and not free them up for innovation.

4. An organization has created an ecommerce website. What data on this website would be considered structured data?
- A) Product ratings score
  - B) Product reviews
  - C) Product photographs
  - D) Product descriptions

**Answer:**

**A) is correct because product ratings are structured because they are numerical scores.**

B) is not correct because product reviews are unstructured because they are text files.

C) is not correct because product photographs are unstructured because they are image files.

D) is not correct because product descriptions are unstructured because they are text files.

5. An organization wants to dynamically adjust its application to serve different user needs. What are the benefits of storing their data in the cloud for this use case?
- A) Data can be stored in archive for long term access
  - B) No data access management required
  - C) Automatic data cleaning and validation
  - D) Real-time data ingestion and analysis

**Answer:**

A) is not correct. Retaining user data long term in archives could support future application adjustments to suit user needs. However, data that is archived is not quickly accessible.

B) is not correct. Data access management isn't directly linked to the potential for an application to dynamically adjust. Furthermore, when moving into the cloud, organizations retain their data access responsibilities.

C) is not correct. Automatic data cleaning and validation would speed up data processing and so allow an application to adjust more quickly. However, organizations are responsible for the quality and validity of the data that they store.

**D) is correct. By storing their application data in the cloud the organization will be able to gather and analyze user behavior data in real-time. This will enable them to dynamically adjust their application for different user needs.**

6. An organization is looking for a data warehouse for analysis and reporting with seamless scaling. Which Google Cloud product or service should the organization use?

A) Cloud Storage

B) Dataflow

C) BigQuery

D) Looker

**Answer:**

A) is not correct because Cloud Storage is not a data warehouse but an object storage service.

B) is not correct because Dataflow is not a data warehouse but a data processing service.

**C) is correct because BigQuery is a data warehouse used for analytics.**

D) is not correct because Looker is not a data warehouse but a Business Intelligence platform.

7. An organization wants to search for and share plug-and-play AI components which can easily build ML services into their project. Which Google Cloud product should the organization use?

A) Recommendations AI

B) Document AI

C) AI Hub

D) Cloud Talent Solution

**Answer:**

A) is not correct. Recommendations AI delivers highly personalized product recommendations at scale.

- B) is not correct. Document AI uses AI to unlock insights from documents.  
**C) is correct because AI Hub is a hosted repository of plug-and-play AI components.**  
D) is not correct. Cloud Talent Solution uses AI with job search and talent acquisition capabilities.

8. An organization currently stores its data on-premises and they receive different levels of traffic on their website every month. How could the organization benefit from modernizing their infrastructure with cloud technology?

- A) 100% service availability.  
B) They can rely on the cloud provider for all website source code.  
C) They can shift from heavy operational expenditure to a capital expenditure model.  
D) Agile storage scalability.

**Answer:**

- A) is not correct. Organizations should only expect to achieve a maximum of 99.999% availability.  
B) is not correct. Organizations' developers maintain control over source code when infrastructure is modernized with the cloud.  
C) is not correct. Modernizing with the cloud creates a shift from capital expenditure to operational expenditure.  
**D) is correct. Organizations can scale in the cloud by paying for what they use, when they use it.**

9. An organization runs their application on a virtual machine, but every time they want to edit specific features, they have to bring the system offline to update the application. What would be a more appropriate solution for their app?

- A) Solid State Disk  
B) Containers  
C) Hypervisors  
D) GPUs

**Answer:**

- A) is not correct. Solid State Disks are only used to enhance the speed of storage and are not relevant to application architecture.  
**B) is correct because containers can compartmentalize applications which enables parts to be edited in isolation.**  
C) is not correct. Hypervisors enable virtual machines to efficiently run multiple applications at the same time, but do not compartmentalize the apps.

D) is not correct. GPUs are just processing power, using a graphical chip architecture.

10. An organization wants to scale their existing virtual machine architecture as quickly as possible. Why should the organization use VMware Engine?

- A) To migrate virtual machines to containers.
- B) To replatform virtual machines as they are.
- C) To deploy custom APIs seamlessly.
- D) To archive virtual machine instances.

**Answer:**

A) is not correct. VMware Engine does not migrate virtual machines to containers.

**B) is correct. VMware Engine helps migrate and run virtual machines in Google Cloud with minimal changes to the VM architecture.**

C) is not correct. VMware Engine does not deploy custom APIs.

D) is not correct. VMware Engine does not archive virtual machine instances.

11. An organization's applications run on an inflexible, on-premises architecture. The organization has decided to modernize their existing applications with the cloud. What may have prompted this business decision?

- A) Developers want to test ideas and experiment with more ease.
- B) Developers want cloud providers to take full control of their application performance.
- C) IT managers want cloud providers to automatically deploy their infrastructure.
- D) IT managers want to stop making gradual changes.

**Answer:**

**A) is correct. Modernizing applications means they can make alterations and innovate more easily.**

B) is not correct. This is not the responsibility of a cloud provider.

C) is not correct. This is not the responsibility of a cloud provider.

D) is not correct. According to DevOps, implementing gradual changes helps reduce cost of failure.

12. An organization with **hybrid cloud architecture** wants to build an application once and be able to run it both on-premises and in their public cloud. Which Google Cloud solution should the organization use?

- A) Cloud Functions
- B) Anthos
- C) App Engine
- D) Compute Engine



**Answer:**

A) is not correct. Cloud Functions cannot run on-premises.

**B) is correct because Anthos allows organizations to build an application once and run it anywhere. (the hint here was hybrid cloud)**

C) is not correct. App Engine cannot run on-premises.

D) is not correct. Compute Engine cannot run on-premises.

13. A retail store has discovered a cost-effective solution for creating self-service kiosks. They can use existing check-out hardware and purchase a virtual customer service application. Why do they also need an API?

A) To update the check-out hardware remotely.

B) To migrate all customer data for disaster recovery.

C) To connect the new application with the legacy system.

D) To connect the check-out hardware to the public cloud.

**Answer:**

A) is not correct. APIs don't update hardware.

B) is not correct. Disaster recovery is not relevant to the business decision.

**C) is correct. APIs can create new business value by connecting legacy systems (the checkout hardware) with new software (the virtual customer service application).**

D) is not correct. Confidential information should not be shared with the public cloud.

14. A bank wants to track the success of their existing ATM network, which has been modernized with APIs to instantly notify customers about their transfers. What is the benefit of using Apigee to achieve this goal?

A) It allows developers to connect the banking APIs with the public cloud.

B) It measures and tracks their total cost of ownership (TCO).

C) It has dashboards that chart dimensions and metrics to report on APIs.

D) It replicates banking APIs to create new business value.

**Answer:**

A) is not correct. Apigee does not share data with the public cloud.

B) is not correct. It measures and tracks the performance of their APIs, not TCO.

**C) is correct. Apigee includes analytics services which allow enterprises to report on various aspects of an API.**

D) is not correct. Apigee does not replicate APIs.

15. A large organization is struggling to manage their cloud costs effectively. They want to increase visibility into cloud costs. Which cost management approach should the organization use?

- A) Appoint a single person to monitor cloud spending across the organization.
- B) Review any cloud spending that exceeds the organization's error budget.
- C) Increase monitoring of on-premises infrastructure and services.
- D) Establish a partnership between finance, technology, and business teams.

**Answer:**

- A) is not correct. This approach might work for a small organization, but would be problematic for a large organization.
- B) is not correct. An error budget is not related to cost management. Instead it is the difference between an SLO (service level objective) and an SLA (service level agreement).
- C) is not correct. Cloud costs are not impacted by on-premises infrastructure and services.
- D) is correct because cross-team partnerships are part of the visibility cost management strategy.**

16. What issues can arise when organizations integrate third-party systems into their cloud infrastructure?
- A) Without sufficient security measures and regular checks, unsecured third-party systems can pose a threat to data security.
  - B) Third-party systems are less capable of addressing an organization's security requirements.
  - C) Over-reliance on third-party systems limits an organization's potential for innovation.
  - D) Third-party systems may not be powerful enough to run many critical business applications.

**Answer:**

- A) is correct because unsecured third-party systems are a cybersecurity threat.**
- B) is not correct. Third-party systems are not inherently less secure than cloud systems.
- C) is not correct. The use of third-party systems does not necessarily limit innovation.
- D) is not correct. Third-party systems are not inherently less powerful than cloud systems.

17. How does a least privilege resource access model contribute to cloud security?
- A) Only managers and other senior employees have cloud resource access.
  - B) Employees only have access to the cloud resources necessary for their job.
  - C) Employees may only access on-premises software with special permission.
  - D) Google is responsible for determining access to cloud resources.

**Answer:**

- A) is not correct. This access model could impede an organization's ability to operate in the cloud.

**B) is correct. This is the definition of a least privilege model.**

C) is not correct. Access to on-premises software is not directly related to cloud resource security.

D) is not correct. Organizations are responsible for determining access to cloud resources.

18. An organization wants to measure everything as part of its new DevOps philosophy. What should the organization measure?

A) The satisfaction and happiness of their employees.

B) The reliability and health of their systems.

C) The risk and reward of their investments.

D) The speed of their cloud adoption process.

**Answer:**

A) is not correct. This is not measured as part of DevOps.

**B) is correct. This is one of the five DevOps aims.**

C) is not correct. This is not measured as part of DevOps.

D) is not correct. This is not measured as part of DevOps.

19. An organization wants to introduce a new feature to their customer-facing application. They want to follow SRE best practices. What should the organization do?

A) Roll out the new feature to all customers so that no customer feels left out.

B) Postpone rolling out the new feature until they are 100% sure there will be no bugs.

C) Roll out the new feature to a subset of customers to test it out.

D) Postpone rolling out the new feature until toil automation is minimized.

**Answer:**

A) is not correct. This is the opposite of implementing gradual changes, which is one of the SRE best practices.

B) is not correct. This is an unrealistic expectation and goes against the SRE best practice of accepting failure as normal.

**C) is correct because implementing gradual changes is one of the SRE best practices.**

D) is not correct. Toil automation is a goal of SRE, not something to be avoided. Toil automation reduces the amount of manual, repetitive work.

20. An organization wants to evaluate the performance of their entire cloud infrastructure, including metrics like server uptime and response rate reports. Which Google Cloud tool should the organization use?

A) Cloud Profiler

B) Cloud Debugger

- C) Cloud Trace
- D) Cloud Monitoring

**Answer:**

A) is not correct. Cloud Profiler helps analyze application performance by gathering CPU usage and memory-allocation information.

B) is not correct. Cloud Debugger helps monitor application performance.

C) is not correct. Cloud Trace helps developers debug or fix and optimize their code.

**D) is correct because Cloud Monitoring enables users to monitor the performance of their entire cloud infrastructure.**