Google Cloud

Partner Certification Academy

Professional Cloud Architect

pls-academy-pca-student-slides-2-2301

The information in this presentation is classified:

Google confidential & proprietary

⚠️ This presentation is shared with you under NDA.

- Do **not** record or take screenshots of this presentation.
- Do **not** share or otherwise distribute the information in this presentation with anyone **inside** or **outside** of your organization.

Thank you!

Google Cloud

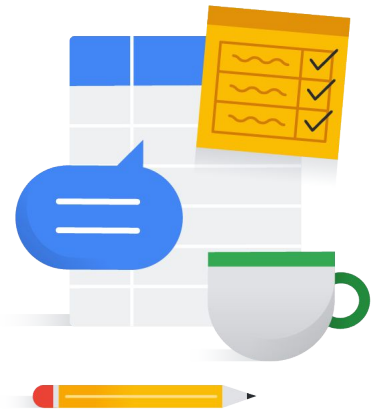
Session logistics

- When you have a question, please:
  - Click the Raise hand button in Google Meet.
  - Or add your question to the Q&A section of Google Meet.
  - Please note that answers may be deferred until the end of the session.
- These slides are available in the Student Lecture section of your Qwiklabs classroom.
- The session is **not recorded.**
- Google Meet does not have persistent chat.
  - If you get disconnected, you will lose the chat history.
  - Please copy any important URLs to a local text file as they appear in the chat.

Google Cloud

Program issues or concerns?

- Problems with **accessing** Cloud Skills Boost for Partners
  - partner-training@google.com
- Problems with **a lab** (locked out, etc.)
  - support@qwiklabs.com
- Problems with accessing Partner Advantage
  - https://support.google.com/googlecloud/topic/9198654



Google Cloud

Professional Cloud Architect (PCA)

Professional Cloud Architects enable organizations to leverage Google Cloud technologies. With a thorough understanding of cloud architecture and Google Cloud, they design, develop, and manage robust, secure, scalable, highly available, and dynamic solutions to drive business objectives.

The Professional Cloud Architect certification exam assesses your ability to:

- ✓ Design and plan a cloud solution architecture
- ✓ Analyze and optimize technical and business processes
- ✓ Manage and provision the cloud solution infrastructure
- ✓ Manage implementations of cloud architecture
- ✓ Design for security and compliance
- ✓ Ensure solution and operations reliability

Google Cloud https://cloud.google.com/certification/cloud-architect

What is a PCA supposed to know/do?

Professional Cloud Architect Certification Page
https://cloud.google.com/certification/cloud-architect

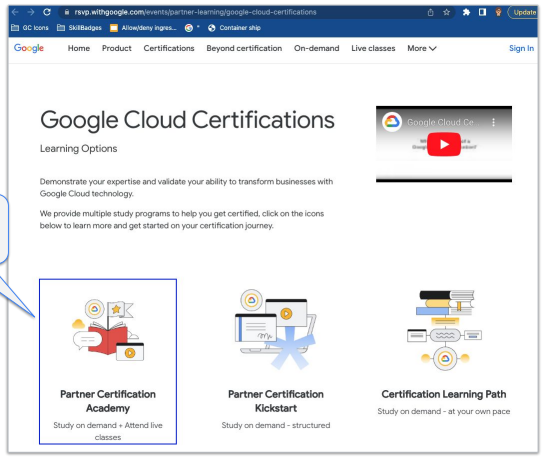
Professional Cloud Architect Exam Guide
https://cloud.google.com/certification/guides/professional-cloud-architect/

Professional Cloud Architect Sample Questions
https://docs.google.com/forms/d/e/1FAIpQLSdvf8Xq6m0kvyIoysdr8WZYCG32WHENStftiHTSdtW4ad2-0w/viewform

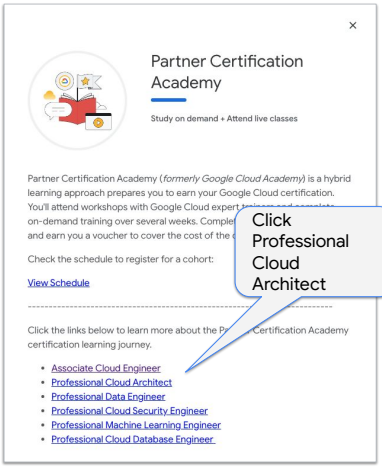
Google Cloud Adoption Framework + Whitepaper
https://cloud.google.com/adoption-framework
https://services.google.com/fh/files/misc/google_cloud_adoption_framework_whitepaper.pdf

Learning Path - Partner Certification Academy Website

Go to: https://rsvp.withgoogle.com/events/partner-learning/google-cloud-certifications



Click here



Click Professional Cloud Architect

Google Cloud

PARTNER CERTIFICATION ACADEMY

Professional Cloud Architect

On-demand Course
Hands-on Skill Badge
Hands-on Lab
Resource
Live Virtual Workshop
required to earn exam voucher

Click to register

Pre-Work

- Professional Cloud Architect
- Certification Learning Path: Professional Cloud Architect

Module 1

- A Tour of Google Cloud Hands-on Labs
- Preparing for your Professional Cloud Architect Journey (pre-work)
- Technical Workshop 1
- Technical Workshop 2
- Google Cloud Platform Fundamentals - Core Infrastructure
- Essential Cloud Infrastructure: Foundation
- Google Cloud security foundations guide (white paper)

Module 2

- Technical Workshop 3
- Technical Workshop 4
- Essential Cloud Infrastructure: Core Services
- Service Accounts and Roles: Fundamentals

Needed for Exam Voucher

Google Cloud

PARTNER CERTIFICATION ACADEMY

Professional Cloud Architect

On-demand Course
Hands-on Skill Badge
Hands-on Lab
Resource
Live Virtual Workshop
required to earn exam voucher

Click to register

Module 3

- Technical Workshop 5
- Technical Workshop 6
- Elastic Google Cloud Infrastructure: Scaling and Automation
- HTTP Load Balancer with Cloud Armor
- Create an Internal Load Balancer
- Network Tiers - Optimizing Network Spend
- Perform Foundational Infrastructure Tasks in Google Cloud
- Create and Manage Cloud Resources

Module 4

- Technical Workshop 7
- Technical Workshop 8
- Getting Started with Google Kubernetes Engine
- Setting up a Private Kubernetes Cluster
- Optimize Costs for Google Kubernetes Engine
- Deploy and Manage Cloud Environments with Google Cloud

Module 5

- Technical Workshop 9
- Technical Workshop 10
- Reliable Google Cloud Infrastructure: Design and Process
- Getting Started with Cloud KMS
- User Authentication: Identity-Aware Proxy
- Securing Virtual Machines using Beyond Corp Enterprise
- Set Up and Configure a Cloud Environment in Google Cloud
- Automating Infrastructure on Google Cloud with Terraform

Module 6

- Logging, Monitoring, and Observability in Google Cloud
- Technical Workshop 11
- Exam Readiness Workshop
- Cloud Architecture: Design, Implement, and Manage
- Reporting Application Metrics into Cloud Monitoring
- Monitoring Multiple Projects with Cloud Monitoring

Needed for Exam Voucher

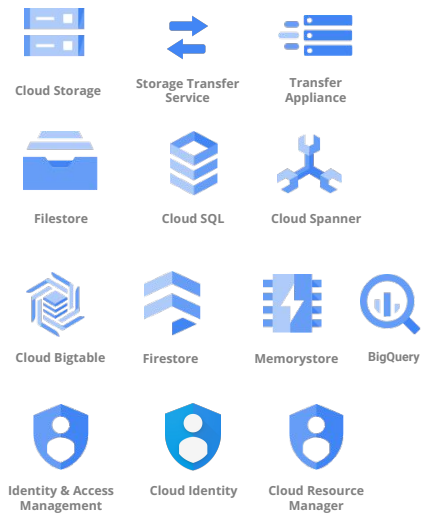
02

Key concepts in the on-demand content

Google Cloud

Topics in this module

- Storage options
  - Cloud Storage
    - Storage Transfer Options
  - Filestore
  - Cloud SQL
  - Spanner
  - Firestore
  - Bigtable
  - BigQuery
  - Memorystore
- Identity and Access Management (IAM)

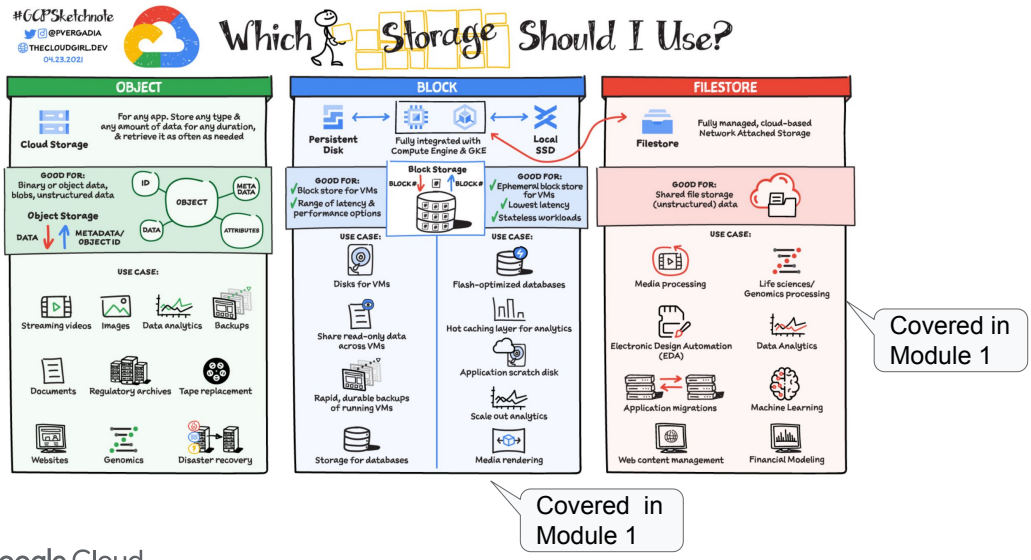


Google Cloud

Storage options in Google Cloud...

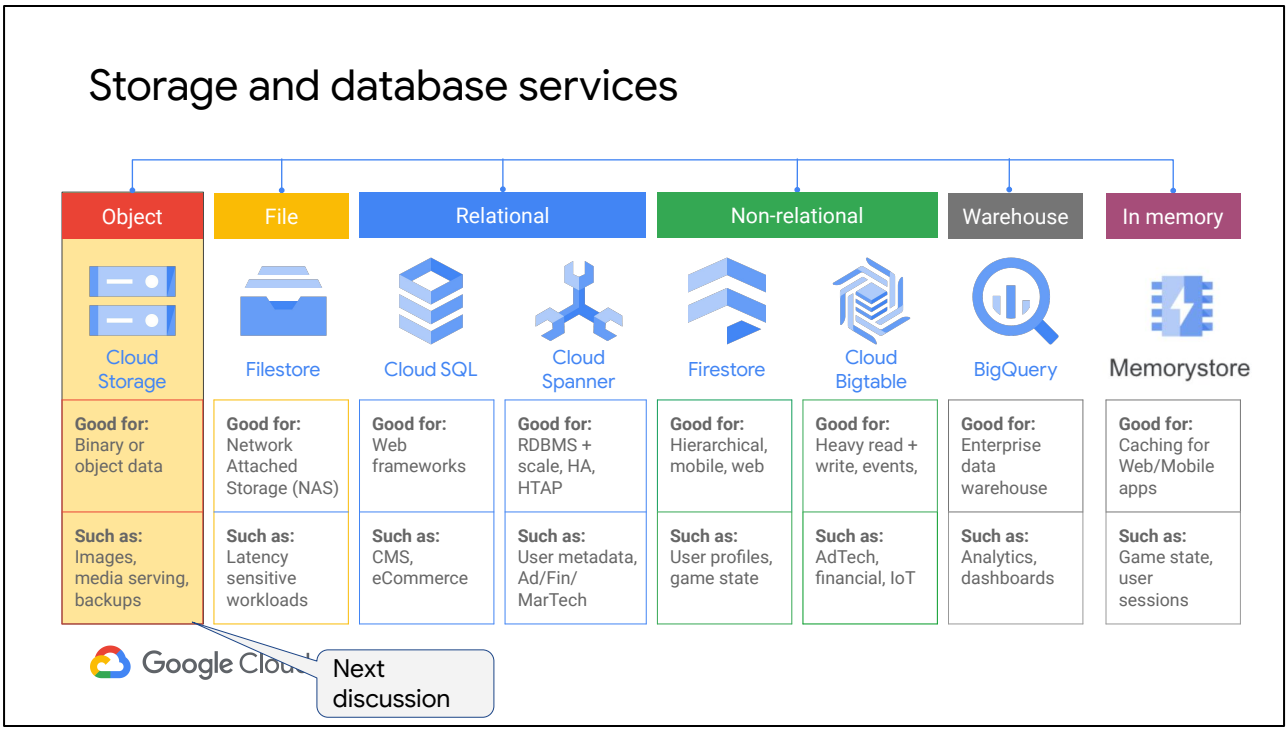
Google Cloud

Storage Overview



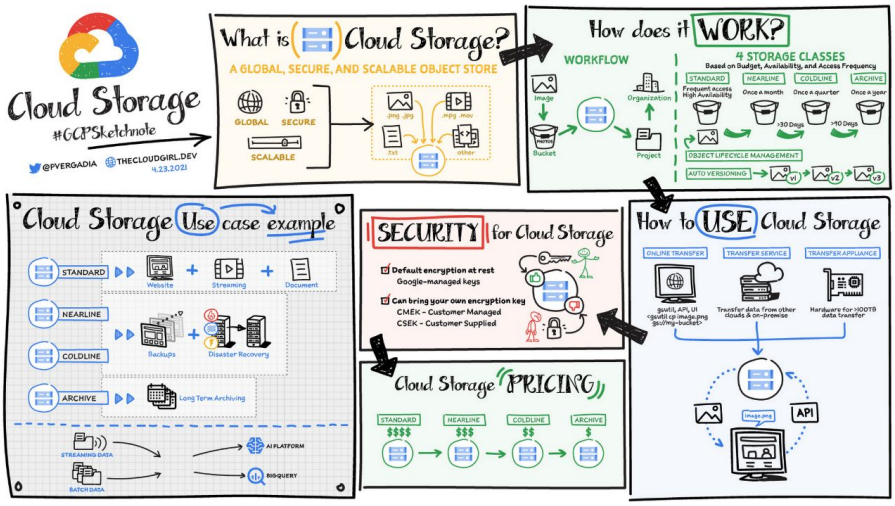
Google Cloud

https://cloud.google.com/blog/topics/developers-practitioners/map-storage-options-google-cloud



This table shows the storage and database services and highlights the storage service type (object, file, relational, non-relational, and data warehouse), what each service is good for, and intended use.

All you need to know about Cloud Storage

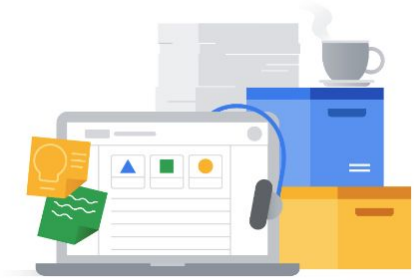


Google Cloud https://cloud.google.com/blog/topics/developers-practitioners/all-you-need-know-about-cloud-storage

Cloud Storage is a fully managed storage service

Binary large-object (BLOB) storage used for

- Online content
- Backup and archiving
- Storage of intermediate results
- And much more....

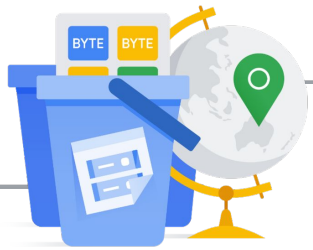


Google Cloud

Object storage for companies of all sizes
https://cloud.google.com/storage

Cloud Storage's primary use is whenever binary large-object storage (also known as a "BLOB") is needed for online content such as videos and photos, for backup and archived data, and for storage of intermediate results in processing workflows.

Files are organized into buckets



Geographic location

Unique name

Google Cloud

Cloud Storage files are organized into buckets. A bucket needs a globally-unique name and a specific geographic location for where it should be stored, and an ideal location for a bucket is where latency is minimized. For example, if most of your users are in Europe, you probably want to pick a European location so either a specific Google Cloud region in Europe, or else the EU multi-region.

There are four Cloud Storage classes

Storage Class	Standard	Nearline	Coldline	Archive
Data access frequency	Often	Once per 30 days	Once per 90 days	Once per year

Google Cloud

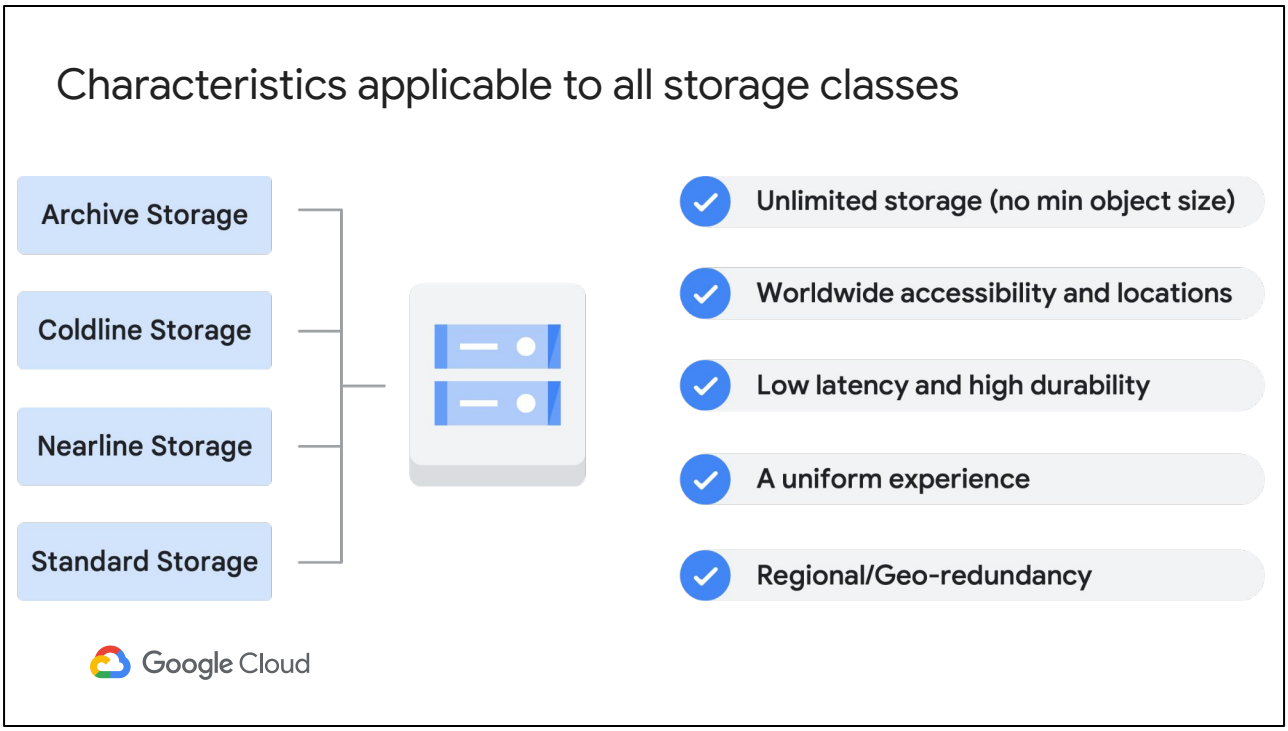
There are four primary storage classes in Cloud Storage and stored data is managed and billed according to which “class” it belongs.

The first is Standard Storage. Standard Storage is considered best for frequently accessed, or “hot,” data. It’s also great for data that is stored for only brief periods of time.

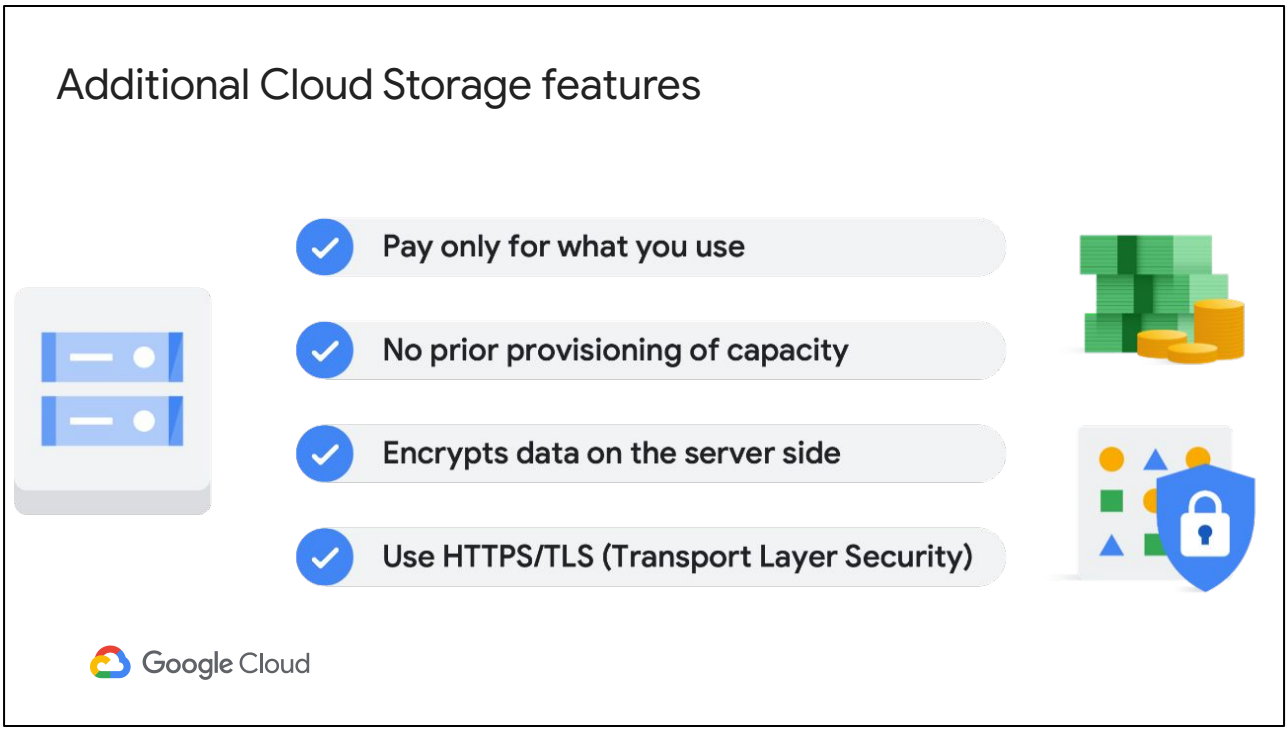
The second storage class is Nearline Storage. This is best for storing infrequently accessed data, like reading or modifying data once per month or less, on average. Examples might include data backups, long-tail multimedia content, or data archiving.

The third storage class is Coldline Storage. This is also a low-cost option for storing infrequently accessed data. However, as compared to Nearline Storage, Coldline Storage is meant for reading or modifying data, at most, once every 90 days.

The fourth storage class is Archive Storage. This is the lowest-cost option, used ideally for data archiving, online backup, and disaster recovery. It’s the best choice for data that you plan to access less than once a year, because it has higher costs for data access and operations and a 365-day minimum storage duration.

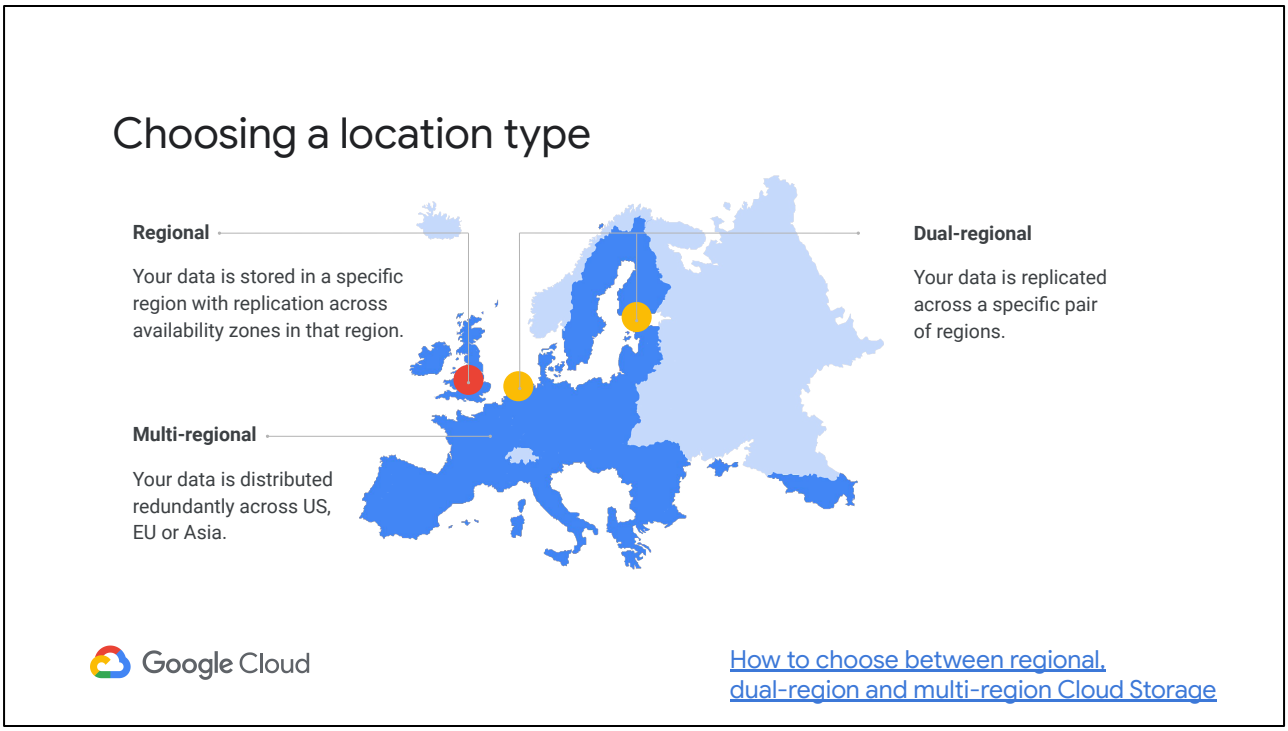


Although each of these four classes has differences, it's worth noting that several characteristics apply across all these storage classes.
These include:
●Unlimited storage with no minimum object size requirement,
●Worldwide accessibility and locations,
●Low latency and high durability,
●A uniform experience, which extends to security, tools, and APIs, and
●Geo-redundancy if data is stored in a multi-region or dual-region. So this means placing physical servers in geographically diverse data centers to protect against catastrophic events and natural disasters, and load-balancing traffic for optimal performance.



Cloud Storage has no minimum fee because you pay only for what you use, and prior provisioning of capacity isn't necessary.

And from a security perspective, Cloud Storage always encrypts data on the server side, before it's written to disk, at no additional charge. Data traveling between a customer's device and Google is encrypted by default using HTTPS/TLS (Transport Layer Security).



Bucket locations
https://cloud.google.com/storage/docs/locations

How to migrate Cloud Storage data from multi-region to regional
https://cloud.google.com/blog/products/storage-data-transfer/multi-region-google-cloud-storage-to-regional-data-migration

How to choose between regional, dual-region and multi-region Cloud Storage
https://cloud.google.com/blog/products/storage-data-transfer/choose-between-regional-dual-region-and-multi-region-cloud-storage/

Choosing a storage classes

	Standard	Nearline	Coldline	Archive
Use case	"Hot" data and/or stored for only brief periods of time like data-intensive computations	Infrequently accessed data like data backup, long-tail multimedia content, and data archiving	Infrequently accessed data that you read or modify at most once a quarter	Data archiving, online backup, and disaster recovery
Minimum storage duration*	None	30 days	90 days	365 days
Retrieval cost	None	$0.01 per GB	$0.02 per GB	$0.05 per GB
Availability SLA	99.95% (multi/dual) 99.90% (region)	99.90% (multi/dual) 99.00% (region)		None
Durability	99.999999999%			

Google Cloud

*Minimum storage duration = if delete file before x days, will still pay for x days

Cloud Storage has four storage classes: Standard, Nearline, Coldline and Archive and each of those storage classes provide 3 location types:

- A multi-region is a large geographic area, such as the United States, that contains two or more geographic places.
- A dual-region is a specific pair of regions, such as Finland and the Netherlands.
- A region is a specific geographic place, such as London.

Objects stored in a multi-region or dual-region are geo-redundant.

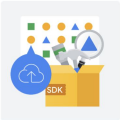
- **Standard** Storage is best for data that is frequently accessed ("hot" data) and/or stored for only brief periods of time. This is the most expensive storage class but it has no minimum storage duration and no retrieval cost. When used in a:
    - region, Standard Storage is appropriate for storing data in the same location as Google Kubernetes Engine clusters or Compute Engine instances that use the data. Co-locating your resources maximizes the performance for data-intensive computations and can reduce network charges.
    - dual-region, you still get optimized performance when accessing Google Cloud products that are located in one of the associated regions, but you also get the improved availability that comes from

    - storing data in geographically separate locations.
    - multi-region, Standard Storage is appropriate for storing data that is accessed around the world, such as serving website content, streaming videos, executing interactive workloads, or serving data supporting mobile and gaming applications.
- **Nearline** Storage is a low-cost, highly durable storage service for storing infrequently accessed data like data backup, long-tail multimedia content, and data archiving. Nearline Storage is a better choice than Standard Storage in scenarios where slightly lower availability, a 30-day minimum storage duration, and costs for data access are acceptable trade-offs for lowered at-rest storage costs.
- **Coldline** Storage is a very-low-cost, highly durable storage service for storing infrequently accessed data. Coldline Storage is a better choice than Standard Storage or Nearline Storage in scenarios where slightly lower availability, a 90-day minimum storage duration, and higher costs for data access are acceptable trade-offs for lowered at-rest storage costs.
- **Archive** Storage is the lowest-cost, highly durable storage service for data archiving, online backup, and disaster recovery. Unlike the "coldest" storage services offered by other Cloud providers, your data is available within milliseconds, not hours or days. Unlike other Cloud Storage storage classes, Archive Storage has no availability SLA, though the typical availability is comparable to Nearline Storage and Coldline Storage. Archive Storage also has higher costs for data access and operations, as well as a 365-day minimum storage duration. Archive Storage is the best choice for data that you plan to access less than once a year.

Let’s focus on durability and availability. All of these storage classes have 11 nines of durability, but what does that mean? Does that mean you have access to your files at all times? No, what that means is you won't lose data. You may not be able to access the data which is like going to your bank and saying well my money is in there, it's 11 nines durable. But when the bank is closed we don't have access to it which is the availability that differs between the storage classes and the location type.

Bringing data into Cloud Storage

- Online transfer
  - Drag and drop in the GC Console
  - Upload / download via the command line
    - gsutil/gcloud



Online transfer

- Storage Transfer Service
  - Create jobs to run once or on a scheduled bases
  - Transfer data from
    - Other clouds (AWS, Azure), URL, Posix filesystem
    - On-premise
    - Cloud Storage Bucket to Cloud Storage Bucket

Storage Transfer Service

- Transfer Appliance
  - Rackable, high-capacity storage server leased from Google
  - Used when the amount of data to be transferred would take too much time given network bandwidth between the customer location and Google

Transfer Appliance

Google Cloud

How long will it take to transfer data?

Cloud Storage - data transfer options:
https://cloud.google.com/architecture/migration-to-google-cloud-transferring-your-large-datasets#transfer-options

Online transfer
gsutil: https://cloud.google.com/storage/docs/gsutil
gcloud: https://cloud.google.com/sdk/gcloud/reference/storage

Storage Transfer Service
https://cloud.google.com/storage-transfer/docs/overview

Transfer Appliance
https://cloud.google.com/transfer-appliance

There are several ways to bring your data into Cloud Storage.

Many customers simply carry out their own online transfer using gsutil, which is the Cloud Storage command from the Cloud SDK. Data can also be moved in by using a drag and drop option in the Google Cloud console, if accessed through the Google Chrome web browser.

But what if you have to upload terabytes or even petabytes of data? **Storage Transfer Service** enables you to import large amounts of online data into Cloud Storage quickly

and cost-effectively. The Storage Transfer Service lets you schedule and manage batch transfers to Cloud Storage from another cloud provider, from a different Cloud Storage region, or from an HTTP(S) endpoint.

And then there is the **Transfer Appliance**, which is a rackable, high-capacity storage server that you lease from Google Cloud. You connect it to your network, load it with data, and then ship it to an upload facility where the data is uploaded to Cloud Storage. You can transfer up to a petabyte of data on a single appliance.

Transferring data into Google Cloud can be challenging

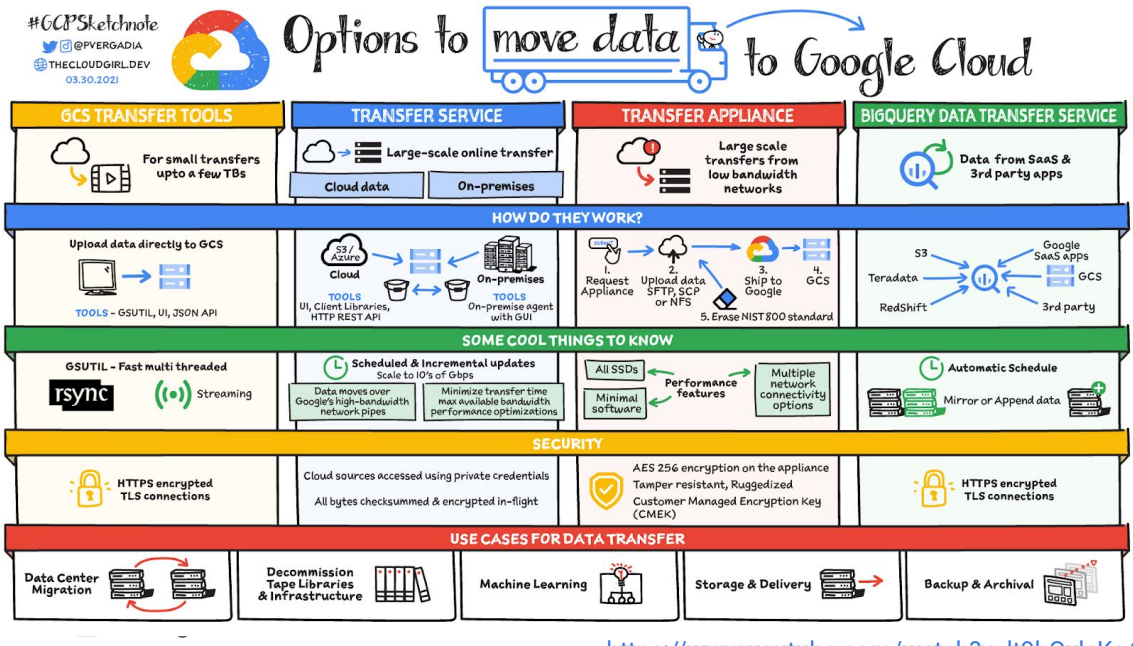
		1 Mbps	10 Mbps	100 Mbps	1 Gbps	10 Gbps	100 Gbps
	1 GB	3 hrs	18 mins	2 mins	11 secs	1 sec	.1 secs
	10 GB	30 hrs	3 hrs	18 mins	2 mins	11 secs	1 sec
	100 GB	12 days	30 hrs	3 hrs	18 mins	2 mins	11 secs
	1 TB	124 days	12 days	30 hrs	3 hrs	18 mins	2 mins
	10 TB	3 years	124 days	12 days	30 hrs	3 hrs	18 mins
Typical enterprise	100 TB	34 years	3 years	124 days	12 days	30 hrs	3 hrs
	1 PB	340 yrs	34 years	3 years	124 days	12 days	30 hrs
	10 PB	3.404 yrs	340 yrs	34 years	3 years	124 days	12 days
	100 PB	34,048 yr	3,404 yrs	340 yrs	34 years	3 years	124 days

23

Data transfer speeds
https://cloud.google.com/transfer-appliance/docs/4.0/overview#transfer-speeds

One of the data transfer barriers that all companies face is their available network. The rate at which we are creating data is accelerating far beyond our capacity to send it all over typical networks. So all companies battle with this

In working with our customers we've found the typical enterprise organization has about 10 petabytes of information and available network bandwidth of somewhere between 100 Mbps and 1 Gbps. This works out to somewhere between 3 and 34 years of ingress. Too long. Now all of that data has different storage requirements, use cases and will likely have many different eventual homes. Because of this organizations need data transfer options across this x and y axis.



https://www.youtube.com/watch?v=lt9bOxlsKs4

How to transfer data to Google Cloud
https://www.youtube.com/watch?v=lt9bOxlsKs4
https://cloud.google.com/blog/topics/developers-practitioners/how-transfer-your-data-google-cloud

Saving files to Cloud Storage - `gsutil`

- `gsutil` - Command line to upload/download files
  - For smaller amounts of data (<1TB) if have adequate bandwidth

  - Upload

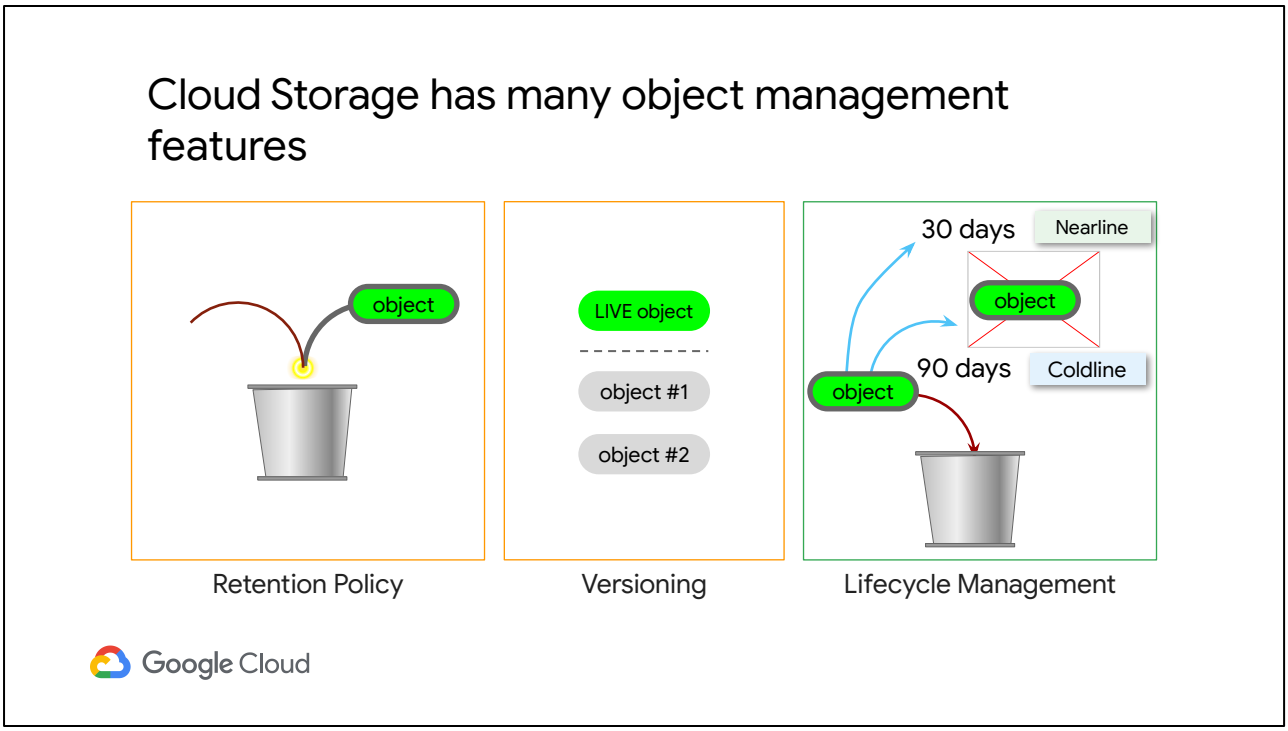
```
gsutil cp OBJECT_LOCATION gs://DESTINATION_BUCKET_NAME/
gsutil cp desktop/myfile.png gs://my-bucket/
```

  - Download

```
gsutil cp gs://BUCKET_NAME/OBJECT_NAME SAVE_TO_LOCATION
gsutil cp gs://my-bucket/* desktop/file-folder/
```

Google Cloud

New: `gcloud storage` commands were introduced in 2022



Best practices for Cloud Storage cost optimization
https://cloud.google.com/blog/products/storage-data-transfer/best-practices-for-cloud-storage-cost-optimization

Cloud Storage has many object management features. For example, you can set a retention policy on all objects in the bucket. For example, the objects should be expired after 30 days.

You can also use versioning, so that multiple versions of an object are tracked and available if necessary. You might even set up lifecycle management, to automatically move objects that haven't been accessed in 30 days to Nearline and after 90 days to Coldline.

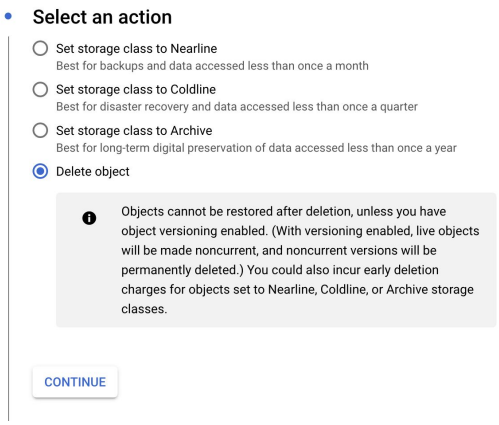
Options for controlling data lifecycles

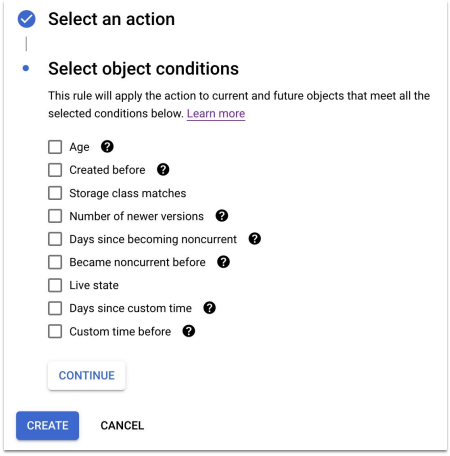
- A **retention policy** specifies a retention period to be placed on a bucket.
  - An object cannot be deleted or replaced until it reaches the specified age.
- **Object Versioning** can be enabled on a bucket in order to retain older versions of objects.
  - When the live version of an object is deleted or replaced, it becomes noncurrent
  - If a live object version is accidentally deleted, can restore the noncurrent version back to the live version.
  - Object Versioning increases storage costs, but can be mitigated by Lifecycle Management to delete older objects
- **Object Lifecycle Management** can be configured for a bucket, which provides automated control over deleting objects and changing storage classes

Google Cloud

Options for controlling data lifecycles
https://cloud.google.com/storage/docs/control-data-lifecycles

Set Lifecycle policy - Console





Google Cloud

Object Lifecycle Management
https://cloud.google.com/storage/docs/lifecycle

Storage and database services

Object	File	Relational		Non-relational		Warehouse	In memory
Cloud Storage	Filestore	Cloud SQL	Cloud Spanner	Firestore	Cloud Bigtable	BigQuery	Memorystore
Good for: Binary or object data	**Good for:** Network Attached Storage (NAS)	**Good for:** Web frameworks	**Good for:** RDBMS + scale, HA, HTAP	**Good for:** Hierarchical, mobile, web	**Good for:** Heavy read + write, events,	**Good for:** Enterprise data warehouse	**Good for:** Caching for Web/Mobile apps
Such as: Images, media serving, backups	**Such as:** Latency sensitive workloads	**Such as:** CMS, eCommerce	**Such as:** User metadata, Ad/Fin/ MarTech	**Such as:** User profiles, game state	**Such as:** AdTech, financial, IoT	**Such as:** Analytics, dashboards	**Such as:** Game state, user sessions

Google Clo

Next discussion

Filestore

Discussed in module 1

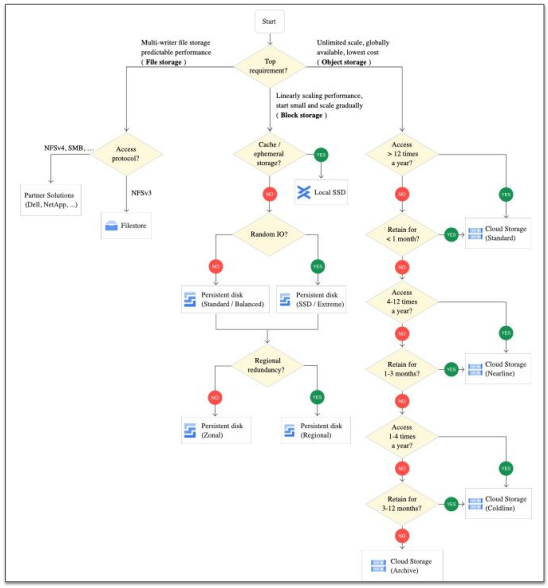
- High-performance, fully managed network attached storage (NAS) for files
  - Mount as file shares on Compute Engine instances
  - Used to store and serve files such as documents, images, videos, audio files, and other data
- Pay for what you use
- Capacity scales automatically scale based on demand
- Use cases:
  - Enterprise application migrations (SAP),
  - Media rendering where file shares are needed
  - Web content management

Google Cloud

YouTube video: https://www.youtube.com/watch?v=CUwpXqEitA0

Storage Options

- Recommended website **comparing object, block and file storage**:
  - [Designing an optimal storage strategy]



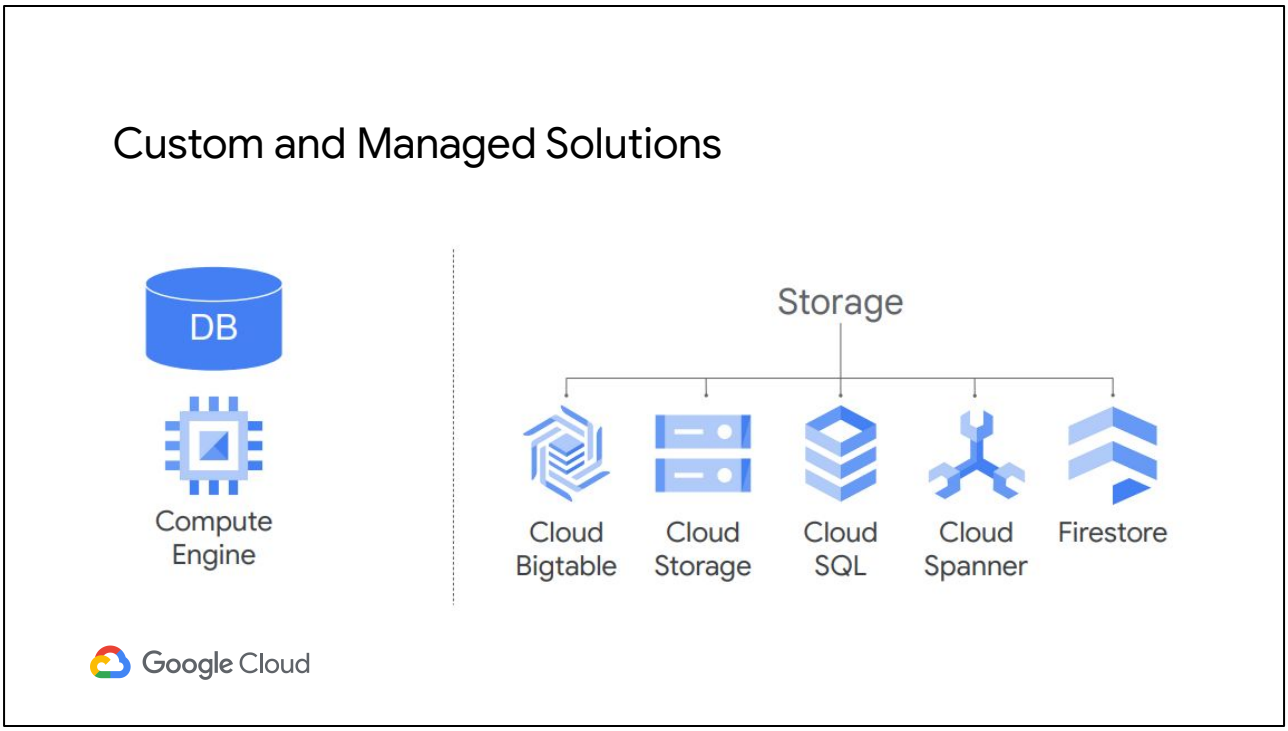
Google Cloud

Managed relational and non-relational databases

Object	File	Relational		Non-relational		Warehouse	In memory
Cloud Storage	Filestore	Cloud SQL	Cloud Spanner	Firestore	Cloud Bigtable	BigQuery	Memorystore
Good for: Binary or object data	**Good for:** Network Attached Storage (NAS)	**Good for:** Web frameworks	**Good for:** RDBMS + scale, HA, HTAP	**Good for:** Hierarchical, mobile, web	**Good for:** Heavy read + write, events,	**Good for:** Enterprise data warehouse	**Good for:** Caching for Web/Mobile apps
Such as: Images, media serving, backups	**Such as:** Latency sensitive workloads	**Such as:** CMS, eCommerce	**Such as:** User metadata, Ad/Fin/ MarTech	**Such as:** User profiles, game state	**Such as:** AdTech, financial, IoT	**Such as:** Analytics, dashboards	**Such as:** Game state, user sessions

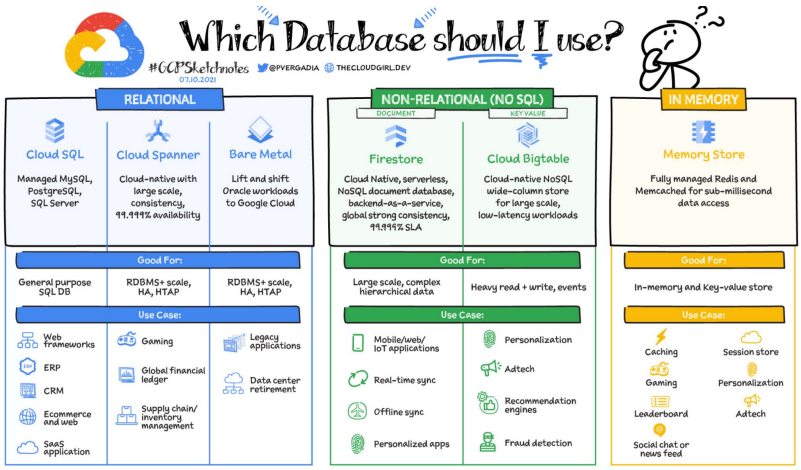
Google Cloud

Coming up



You can manage databases yourself by creating a VM and installing the database software. But you are responsible for everything - backups, patches, OS updates, etc. As an alternative, Google offers fully managed storage options where Google manages the underlying infrastructure

Database Options



https://cloud.google.com/blog/topics/developers-practitioners/your-google-cloud-database-options-explained

Google Cloud

Storage and database services

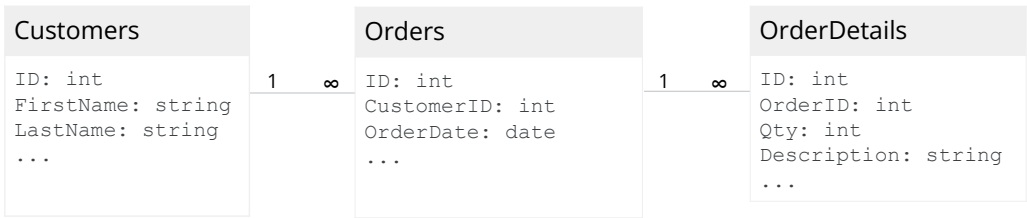
Object	File	Relational		Non-relational		Warehouse	In memory
Cloud Storage	Filestore	Cloud SQL	Cloud Spanner	Firestore	Cloud Bigtable	BigQuery	Memorystore
Good for: Binary or object data	**Good for:** Network Attached Storage (NAS)	**Good for:** Web frameworks	**Good for:** RDBMS + scale, HA, HTAP	**Good for:** Hierarchical, mobile, web	**Good for:** Heavy read + write, events,	**Good for:** Enterprise data warehouse	**Good for:** Caching for Web/Mobile apps
Such as: Images, media serving, backups	**Such as:** Latency sensitive workloads	**Such as:** CMS, eCommerce	**Such as:** User metadata, Ad/Fin/ MarTech	**Such as:** User profiles, game state	**Such as:** AdTech, financial, IoT	**Such as:** Analytics, dashboards	**Such as:** Game state, user sessions

Next discussion

Google Cloud

What does data in a relational database look like?

- Tables contain fields, indexes, and constraints
- Primary key ensures each row in a table is unique
- Relationships are constraints that ensure a parent row cannot be deleted if there are child rows in another table

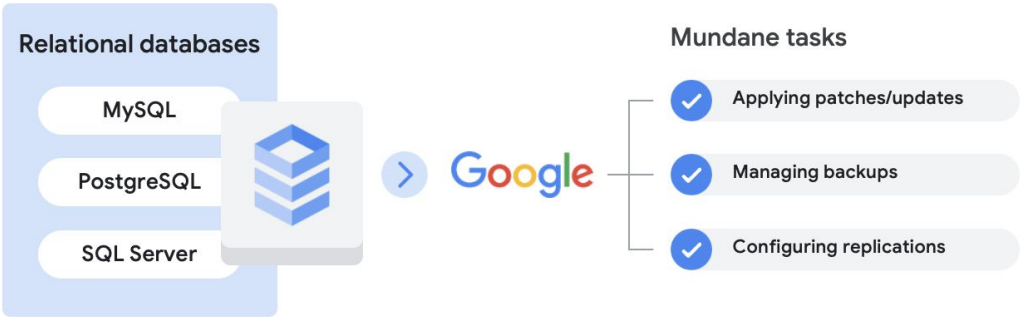


Google Cloud

This is a brief (and very, very simplistic) overview of a relational database.

Cloud SQL

A fully managed, cloud-based alternative to on-premise MySQL, PostgreSQL, and SQL Server databases



[The business value of Cloud SQL: how companies speed up deployments, lower costs and boost agility](https://cloud.google.com/blog/products/databases/the-business-value-of-cloud-sql/)

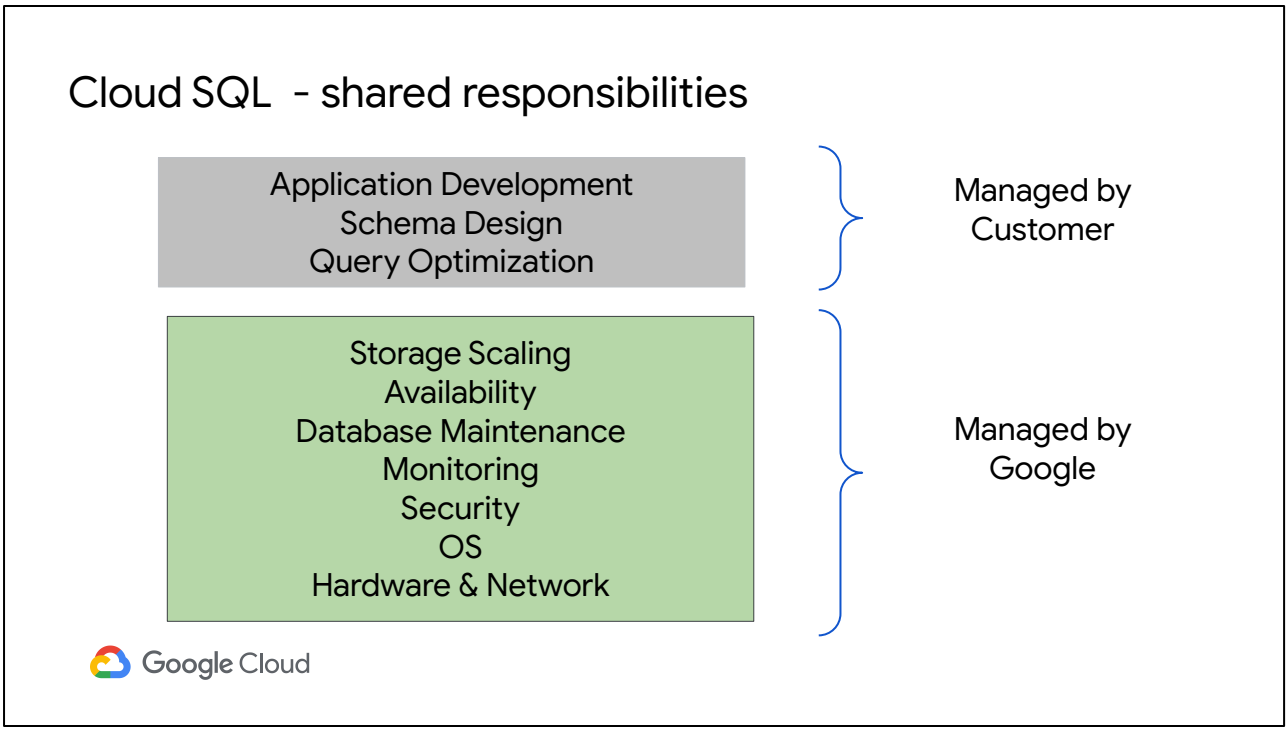
Google Cloud

The business value of Cloud SQL: how companies speed up deployments, lower costs and boost agility
https://cloud.google.com/blog/products/databases/the-business-value-of-cloud-sql/

Cloud SQL:
https://cloud.google.com/sql/docs/mysql/introduction

What is Cloud SQL
https://cloud.google.com/blog/topics/developers-practitioners/what-cloud-sql

Cloud SQL offers fully managed relational databases, including MySQL, PostgreSQL, and SQL Server as a service. It's designed to hand off mundane, but necessary and often time-consuming, tasks to Google—like applying patches and updates, managing backups, and configuring replications—so your focus can be on building great applications.



Google Cloud manages the infrastructure work while the customer focuses on their goals.

Cloud SQL supports a variety of use cases

- Storing and managing relational data such as customer information, product inventory, and financial transactions
- Backend database for web and mobile applications
- Migrating on-premises databases to the cloud
- Building and deploying data-driven applications with minimal setup and management overhead
- Replicating data for disaster recovery and high availability
    - Provides automatic failover for high availability, ensuring that databases are always available in case of an unexpected outage or hardware failure

Google Cloud

Cloud SQL - customer use case

- Qlue: Boosting intelligence about activity in Indonesian cities
- Applications include
  - Dashboard to track and resolve incidents of flooding, crime, fire, illegal rubbish dumping, and more
  - Monitor energy usage and traffic congestion in real time
  - Mobile app that enables citizens to report incidents

> "If we used a conventional Infrastructure-as-a Service offering, we would have had to dedicate a team to manage operating systems, libraries, services, and load balancing. With Google Cloud, we are able to focus purely on developing our applications and growing the business."
>
> *—Surya Darmadi, Chief Operating Officer, Qlue*

Google Cloud

Qlue: Boosting intelligence about activity in Indonesian cities
https://cloud.google.com/customers/qlue/

Note that Qlue uses multiple Google Cloud services, not just Cloud SQL. It's highly recommended that you understand the use case for all these services, and how they are used together to meet customer needs.

Scaling Cloud SQL Databases

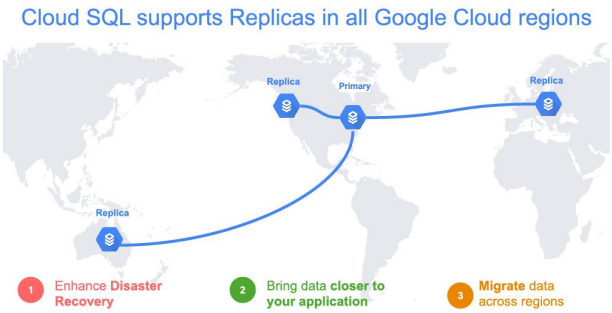
- Cloud SQL can scale in the following ways
  - **Vertical scaling:**
    - Increasing the amount of computing resources (such as CPU and memory) allocated to a single database instance.
    - Can be done with a few clicks in the Cloud Console, and requires no downtime.
  - **Horizontal scaling:**
    - Adding additional read replicas to distribute read workloads and improve performance.
    - Can be added on demand, and there's no limit to the number of replicas that can be added.

Google Cloud

Cloud SQL Read Replicas

Fully managed read replica in a different region(s) than that of the primary instance

- Enhance DR
- Bring data closer to your applications (performance and cost implications)
- Migrate data across regions
- Data and other changes on the primary instance are updated in almost real time on the read replicas



Google Cloud

Cloud SQL Read Replicas:
https://cloud.google.com/sql/docs/mysql/replication

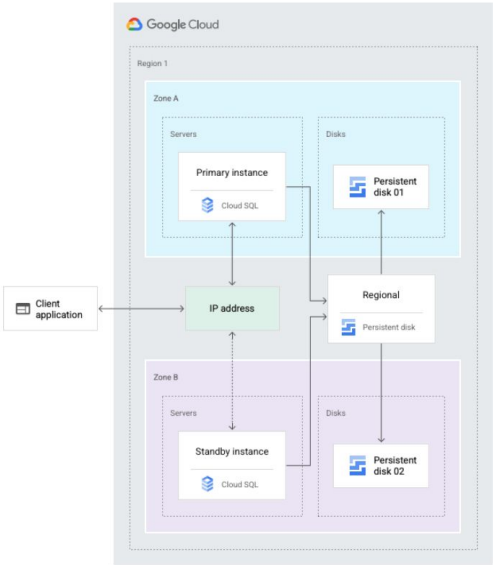
Cloud SQL read replicas provide

- Enhanced Disaster recovery. Can convert a read replica to the Primary (if the primary is no longer available)
- In terms of data closer to your app, you can have different connection strings for an application for different endpoints
    - For example,:
        - Read-Write connection strings will go to the primary
        - Read only connections hit a local read replica
- If reads comprise most of the use case for your app, then accessing the local replica saves on data transfer and egress charges since you are not accessing the primary sitting in the eastern United States in this example

Introducing cross-region replica for Cloud SQL (June 2, 2020)
https://cloud.google.com/blog/products/databases/introducing-cross-region-replica-for-cloud-sql

Cloud SQL - High Availability

- Provides automatic failover if a zone or instance become unavailable
- The primary instance is in one zone in a region
  - The failover instance is in another zone
- Synchronous replication is used to copy all writes from the primary disk to the replica



Google Cloud

https://cloud.google.com/sql/docs/mysql/high-availability

Storage and database services

Object	File	Relational		Non-relational		Warehouse	In memory
Cloud Storage	Filestore	Cloud SQL	Cloud Spanner	Firestore	Cloud Bigtable	BigQuery	Memorystore
Good for: Binary or object data	**Good for:** Network Attached Storage (NAS)	**Good for:** Web frameworks	**Good for:** RDBMS + scale, HA, HTAP	**Good for:** Hierarchical, mobile, web	**Good for:** Heavy read + write, events,	**Good for:** Enterprise data warehouse	**Good for:** Caching for Web/Mobile apps
Such as: Images, media serving, backups	**Such as:** Latency sensitive workloads	**Such as:** CMS, eCommerce	**Such as:** User metadata, Ad/Fin/ MarTech	**Such as:** User profiles, game state	**Such as:** AdTech, financial, IoT	**Such as:** Analytics, dashboards	**Such as:** Game state, user sessions

Google Cloud

Next discussion

Cloud Spanner

A enterprise-grade, globally distributed, externally consistent relational database having unlimited scalability and industry-leading 99.999% availability

- Powers Google's most popular, globally available products, like YouTube, Drive, and Gmail
- Capable of processing more than 1 billion queries per second at peak
- For any workload, large or small, that cannot tolerate downtime, and requires high availability
- Regional and multi-regional deployments
  - SLA: Multi-regional: 99.999%
  - SLA: Regional: 99.99%
- Supports ANSI standard SQL

Google Cloud

Cloud Spanner myths busted

Cloud Spanner:
https://cloud.google.com/blog/topics/developers-practitioners/what-cloud-spanner
https://cloud.google.com/spanner

Cloud Spanner myths busted:
https://cloud.google.com/blog/products/databases/cloud-spanner-myths-busted

Cloud Spanner is an enterprise-grade, globally distributed, externally consistent database that offers unlimited scalability and industry-leading 99.999% availability. It requires no maintenance windows and combines the benefits of relational databases with the unmatched scalability and availability of non-relational databases.

Spanner powers Google's most popular, globally available products, like YouTube, Drive, and Gmail, and it can processes more than 1 Billion queries per second at peak

But don't be mislead into thinking that Spanner is only for large, enterprise level applications. Many customers use Spanner for their smaller workloads (both in terms of transactions per second and storage size) for availability and scalability reasons. Spanner is appropriate for any customer that cannot tolerate downtime, and needs high availability for their applications

Limitless scalability and high availability is critical in many industry verticals such as gaming and retail, especially when a newly launched game goes viral and becomes

an overnight success or when a retailer has to handle a sudden surge in traffic due to a Black Friday/Cyber Monday sale.

Spanner offers the flexibility to interact with the database via a SQL dialect based on ANSI 2011 standard as well as via a REST or gRPC API interface, which are optimized for performance and ease-of-use. In addition to Spanner's interface, there is a PostgreSQL interface for Spanner, that leverages the ubiquity of PostgreSQL and provides development teams with an interface that they are familiar with. The PostgreSQL interface provides a rich subset of the open-source PostgreSQL SQL dialect, including common query syntax, functions, and operators. It also supports a core collection of open-source PostgreSQL data types, DDL syntax, and information schema views. You get the PostgreSQL familiarity, and relational semantics at Spanner scale.

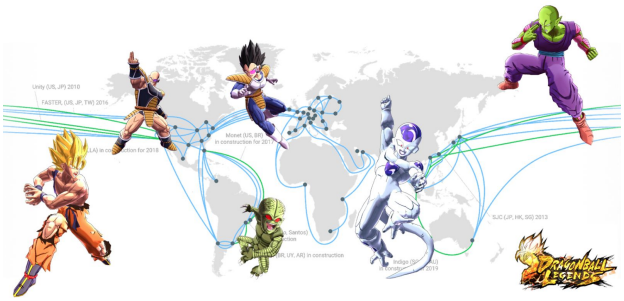
Cloud Spanner supports a variety of use cases

- Large-scale, multi-regional data storage
  - Can store and manage terabytes or petabytes of data across multiple regions
- Online transaction processing (OLTP) applications:
  - Can support OLTP workloads with low latency and high throughput, making it suitable for applications that require real-time data processing, such as e-commerce or financial services
- Banking and financial services
  - Spanner's high-availability and consistency guarantees make it well-suited for use cases in the financial services industry, such as stock trading or payment processing
- Geographically distributed data management
  - Spanner supports multiple geographic locations and provides a globally consistent view of data, making it ideal for use cases that require a database that can handle data distributed across multiple regions or continents

Google Cloud

Cloud Spanner - customer use case

- Dragon Ball Legends - mobile game from Bandai Namco Entertainment
- Requirements were:
  - Global backend that could scale with millions of players and still perform well.
  - Global reliable, low latency network to support multi-region player-versus-player battles
  - Real-time data analytics to measure and evaluate how people are playing the game and adjust it on-the-fly.



Behind the scenes with the Dragon Ball Legends GC backend

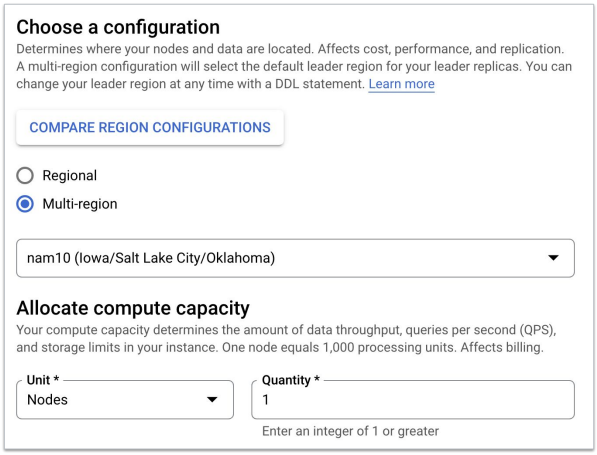
Google Cloud

This is another example where multiple Google Cloud services are used together to provide a solution.

Scaling Cloud Spanner

Scales out (horizontal scaling)

- Manually add nodes/processing units to support more data and users as needed
- Turn on autoscaling to automatically adjust the number of nodes in an instance to handle changing traffic patterns and load



Google Cloud

Autoscaling Cloud Spanner
https://cloud.google.com/architecture/autoscaling-cloud-spanner

Cloud SQL

- Max 64TB data
- High Availability via master / standby
- 99.95% SLA
- Vertically Scalable via reprovisioning
- Horizontally Scalable by read replicas
- Planned maintenance windows

Cloud Spanner

- 4TB / node data
- Distributed service - always available
- 99.99% SLA (regional)
- 99.999% SLA (multi-region)
- Horizontally Scalable
- No maintenance managed service

Google Cloud

Storage and database services

Object	File	Relational		Non-relational		Warehouse	In memory
Cloud Storage	Filestore	Cloud SQL	Cloud Spanner	Firestore	Cloud Bigtable	BigQuery	Memorystore
Good for: Binary or object data	**Good for:** Network Attached Storage (NAS)	**Good for:** Web frameworks	**Good for:** RDBMS + scale, HA, HTAP	**Good for:** Hierarchical, mobile, web	**Good for:** Heavy read + write, events,	**Good for:** Enterprise data warehouse	**Good for:** Caching for Web/Mobile apps
Such as: Images, media serving, backups	**Such as:** Latency sensitive workloads	**Such as:** CMS, eCommerce	**Such as:** User metadata, Ad/Fin/ MarTech	**Such as:** User profiles, game state	**Such as:** AdTech, financial, IoT	**Such as:** Analytics, dashboards	**Such as:** Game state, user sessions

Next discussion

Google Cloud

Types of NoSQL Database

- Key-value stores (**Cloud Memorystore**)
  - Data is stored in key-value pairs
  - Examples include Redis and SimpleDB
- Document stores (**Cloud Firestore**)
  - Data is stored in some standard format like XML or JSON
  - Nested and hierarchical data can be stored together
  - MongoDB, CouchDB, and DynamoDB are examples
- Wide-column stores (**Cloud Bigtable**)
  - Key identifies a row in a table
  - Columns can be different within each row
  - Cassandra and HBase are examples

Google Cloud

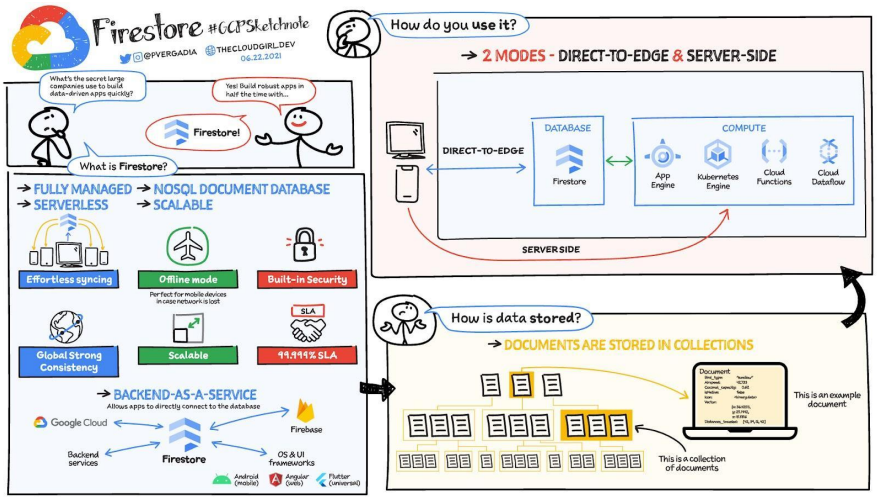
There are different types of NoSQL databases. (The Google Cloud services are in bold text in the above slide.)

Key-value stores are the simplest. These act like a map or dictionary. To save data, just specify a key and assign a value to that key.

Document stores allow you to store hierarchical data. So, instead of having an orders table and a details table, you can store an order in a single document. That document can have properties that represent the order along with an array of subdocuments representing the details. The data can be stored in the database as JSON or XML, or a binary format called BSON.

Wide-column stores have data in tables. Each row in the table has a unique value or key. Associated with that key can be any number of columns. Each column can have any type of data. There's no schema, so different rows in the same table can have different columns.

Cloud Firestore



Google Cloud https://cloud.google.com/blog/topics/developers-practitioners/all-you-need-know-about-firestore-cheatsheet

All you need to know about Firestore:
https://cloud.google.com/blog/topics/developers-practitioners/all-you-need-know-about-firestore-cheatsheet

Cloud Firestore

- Completely managed, document store, NoSQL Database
  - No administration, no maintenance, nothing to provision
- 1 GB per month free tier
- Indexes created for every property by default
  - Secondary indexes and composite indexes are supported
- Supports ACID* transactions
- For mobile, web, and IoT apps at global scale
  - Live synchronization and offline support
- Multi-region replication

Google Cloud

*ACID explained: https://en.wikipedia.org/wiki/ACID

Cloud Firestore is a fast, fully managed, serverless, cloud-native NoSQL document database

It simplifies storing, syncing, and querying data for your mobile, web, and IoT apps at global scale

Its client libraries provide live synchronization and offline support, and its security features and integrations with Firebase and Google Cloud accelerate building truly serverless apps.

Cloud Firestore also supports ACID transactions, so if any of the operations in the transaction fail and cannot be retried, the whole transaction will fail.

Also, with automatic multi-region replication and strong consistency, your data is safe and available, even when disasters strike.

Cloud Firestore allows you to run sophisticated queries against your NoSQL data without any degradation in performance. This gives you more flexibility in the way you structure your data.

Some features recently introduced are

- VPC Service Controls for Firestore. This allows you to define a perimeter to mitigate data exfiltration risks.

- Firestore triggers for Cloud Functions. When certain events happen in Firestore, Cloud Functions can run in response.
- The same Key Visualizer service that's available for Bigtable and Spanner is also available for Firestore. It allows developers to quickly and visually identify performance issues

Firestore example data

An index is created for every property so that queries are extremely fast

Collections

Documents live in collections, which are simply containers for documents. For example, you could have a `users` collection to contain your various users, each represented by a document:

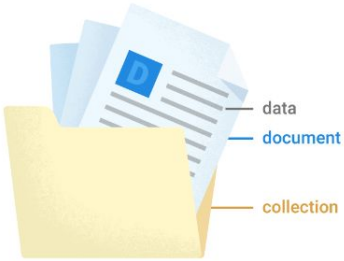
users

alovelace

```
first : "Ada"
last : "Lovelace"
born : 1815
```

aturing

```
first : "Alan"
last : "Turing"
born : 1912
```



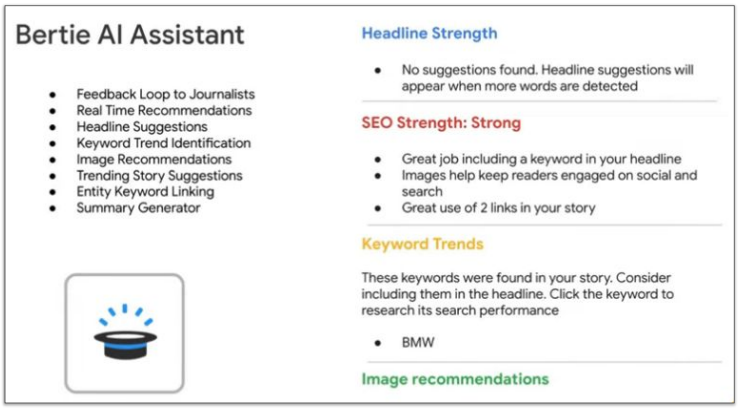
Cloud Firestore is schemaless, so you have complete freedom over what fields you put in each document and what data types you store in those fields. Documents within the same collection can all contain different fields or store different types of data in those fields. However, it's a good idea to use the same fields and data types across multiple documents, so that you can query the documents more easily.

Google Cloud

From: https://firebase.google.com/docs/firestore/data-model#hierarchical-data

Firebase - customer use case

- Forbes created Bertie - an AI assistant for journalists
- Journalists upload their content and Bertie provides feedback
  - Strength of the article's headline
  - Keywords needed for search engine optimization
  - Words to add to the headline to increase search performance



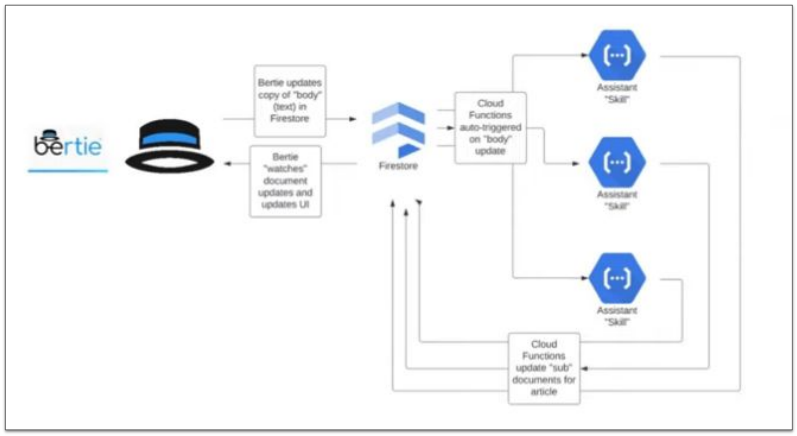
Google Cloud

YouTube video: https://www.youtube.com/watch?v=KVRxsRPhmoo

Forbes website (for anyone that wants to know more about the company)
https://www.forbes.com/

Forbes - Bertie AI Architecture

- Content is stored in Firestore
- Data updates trigger Cloud Functions
  - Each Cloud Function performs a different task
  - Results are written back to Firestore
  - Website is refreshed with the recommendations



Google Cloud

YouTube video: https://www.youtube.com/watch?v=KVRxsRPhmoo

Storage and database services

Object	File	Relational		Non-relational		Warehouse	In memory
Cloud Storage	Filestore	Cloud SQL	Cloud Spanner	Firestore	Cloud Bigtable	BigQuery	Memorystore
Good for: Binary or object data	**Good for:** Network Attached Storage (NAS)	**Good for:** Web frameworks	**Good for:** RDBMS + scale, HA, HTAP	**Good for:** Hierarchical, mobile, web	**Good for:** Heavy read + write, events,	**Good for:** Enterprise data warehouse	**Good for:** Caching for Web/Mobile apps
Such as: Images, media serving, backups	**Such as:** Latency sensitive workloads	**Such as:** CMS, eCommerce	**Such as:** User metadata, Ad/Fin/ MarTech	**Such as:** User profiles, game state	**Such as:** AdTech, financial, IoT	**Such as:** Analytics, dashboards	**Such as:** Game state, user sessions

Google Cloud

Next discussion

Cloud Bigtable

Highly scalable NoSQL database that can handle billions of rows and petabytes of data, making it ideal for use cases that require large-scale data storage and processing with up to 99.999% availability

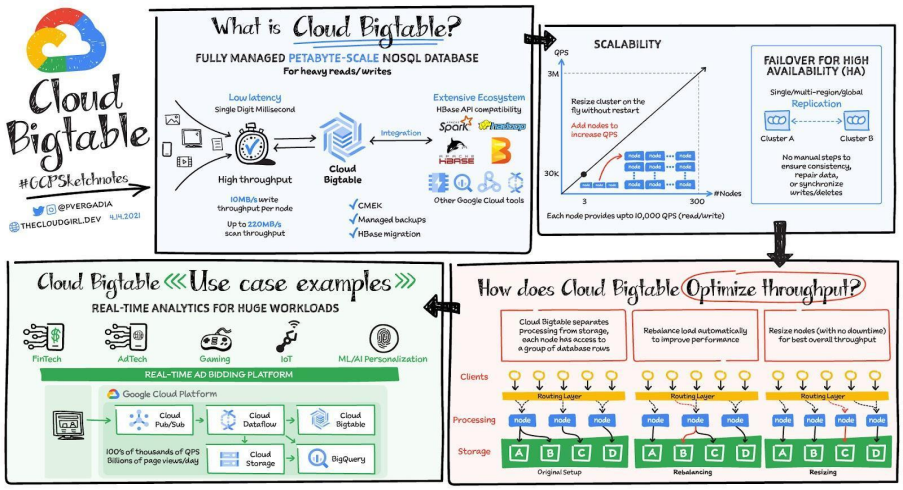
- High throughput at low latency
  - Ideal for use cases that require real-time data processing and analysis
- Integrates easily with big data tools
  - Same API as HBase
  - Allows on-premises HBase applications to be easily migrated
- Consistent sub-10ms latency

Google Cloud

Bigtable:
https://cloud.google.com/bigtable

- Bigtable is a fully managed, scalable NoSQL database service for large analytical and operational workloads with up to 99.999% availability
- Bigtable is built with proven infrastructure that powers Google products used by billions such as Search and Maps.
- Bigtable is ideal for storing very large amounts of data in a key-value store and supports high read and write throughput at low latency for fast access to large amounts of data.
- It is a Fully managed service that integrates easily with big data tools like Hadoop, Dataflow, and Dataproc. Plus, support for the open source HBase API standard makes it easy for development teams to get started
- Some of the features are
  - Autoscaling: Autoscaling helps prevent over-provisioning or under-provisioning by letting Cloud Bigtable automatically add or remove nodes to a cluster when usage changes.
    - In addition, metrics are available to help you understand how autoscaling is working.
  - You can use customer managed encryption keys (CMEK) in Cloud Bigtable instances, including ones that are replicated across multiple regions.
  - App profile cluster groups let you route an app profile's traffic to a subset of an instance's clusters.

Cloud Bigtable



Google Cloud

https://cloud.google.com/blog/topics/developers-practitioners/how-big-cloud-bigtable

How BIG is Cloud Bigtable:
https://cloud.google.com/blog/topics/developers-practitioners/how-big-cloud-bigtable

Cloud Bigtable example use cases

Healthcare	Patient data and other health-related information
Financial Services	Storage of large amounts of transaction, risk management, and compliance data
Retail	Storage of data related to customer behavior which can be used to generate product recommendations
Gaming	Player data for game analytics
IoT	Data generated by connected devices and sensors
Logs and Metrics	Storage of log data and metrics for analysis
Time series data	Storage of resource consumption like CPU and memory usage over time for multiple servers

Google Cloud

Cloud Bigtable example data

One key column

Max size of column is 256 MB/ea

Data that's likely to be accessed via the same request are grouped into column families

Column Families

Row Key

	Flight_Information					Aircraft_Information			
	Origin	Destination	Departure	Arrival	Passengers	Capacity	Make	Model	Age
ATL#arrival#20190321-1121	ATL	LON	20190321-0311	20190321-1121	158	162	B	737	18
ATL#arrival#20190321-1201	ATL	MEX	20190321-0821	20190321-1201	187	189	B	737	8
ATL#arrival#20190321-1716	ATL	YVR	20190321-1014	20190321-1716	201	259	B	757	23

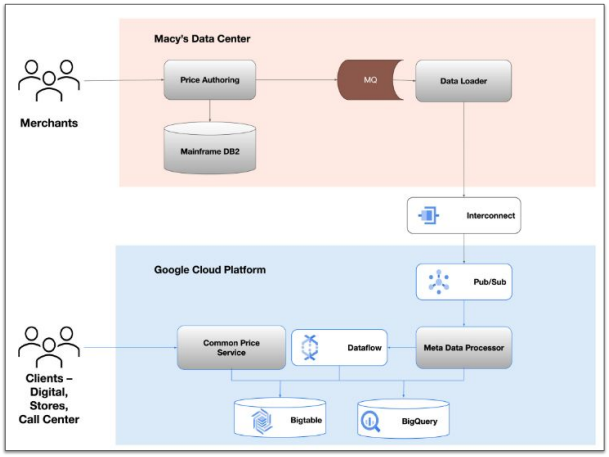
Rows can have millions of columns

Google Cloud

Cloud Bigtable provides Column Families. By accessing the Column Family, you can pull some of the data you need without pulling all of the data from the row or having to search for it and assemble it. This makes access more efficient.

Bigtable - customer use case

- Macy's sells a wide range of merchandise, including apparel and accessories (men's, women's and children's), cosmetics, home furnishings and other consumer goods
  - 700+ stores across the US
- Bigtable provides data for the pricing system
  - Access pattern entails finding an item's ticket price based on a given division, location, and the universal price code
  - Can search millions of product codes within single digit milliseconds



https://cloud.google.com/blog/products/databases/how-macys-enhances-customer-experience-google-cloud-services

Google Cloud

About Macy's
https://www.macysinc.com/about

How Macy's enhances the customer experience with Google Cloud services
https://cloud.google.com/blog/products/databases/how-macys-enhances-customer-experience-google-cloud-services

Storage and database services

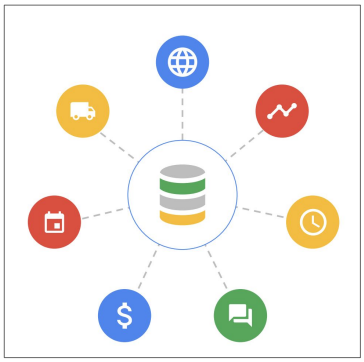
Object	File	Relational		Non-relational		Warehouse	In memory
Cloud Storage	Filestore	Cloud SQL	Cloud Spanner	Firestore	Cloud Bigtable	BigQuery	Memorystore
Good for: Binary or object data	**Good for:** Network Attached Storage (NAS)	**Good for:** Web frameworks	**Good for:** RDBMS + scale, HA, HTAP	**Good for:** Hierarchical, mobile, web	**Good for:** Heavy read + write, events,	**Good for:** Enterprise data warehouse	**Good for:** Caching for Web/Mobile apps
Such as: Images, media serving, backups	**Such as:** Latency sensitive workloads	**Such as:** CMS, eCommerce	**Such as:** User metadata, Ad/Fin/ MarTech	**Such as:** User profiles, game state	**Such as:** AdTech, financial, IoT	**Such as:** Analytics, dashboards	**Such as:** Game state, user sessions

Google Cloud

Next discussion

What is a Data Warehouse?

- Allows data from multiple sources to be combined and analyzed
  - Historical archive of data
- Data sources could be:
  - Relational databases
  - Logs
  - Web data
  - Etc.
- Optimized for analytical processing
  - Can handle large amounts of data and complex queries, and is well-suited for reporting and data analysis



A data warehouse is a central hub for business data. Different types of data can be transformed and consolidated into the warehouse for analysis

Google Cloud

BigQuery

- Fully managed, serverless, highly scalable data warehouse
- Multi-cloud capabilities using standard SQL
- Processes multi-terabytes of data in minutes
- Automatic high availability
- Supports federated queries
  - Cloud SQL & Cloud Spanner
  - Cloud Bigtable
  - Files in Cloud Storage
- Use cases:
  - Near real-time analytics of streaming data to predict business outcomes with built-in machine learning, geospatial analysis and more
  - Analysis of historical data

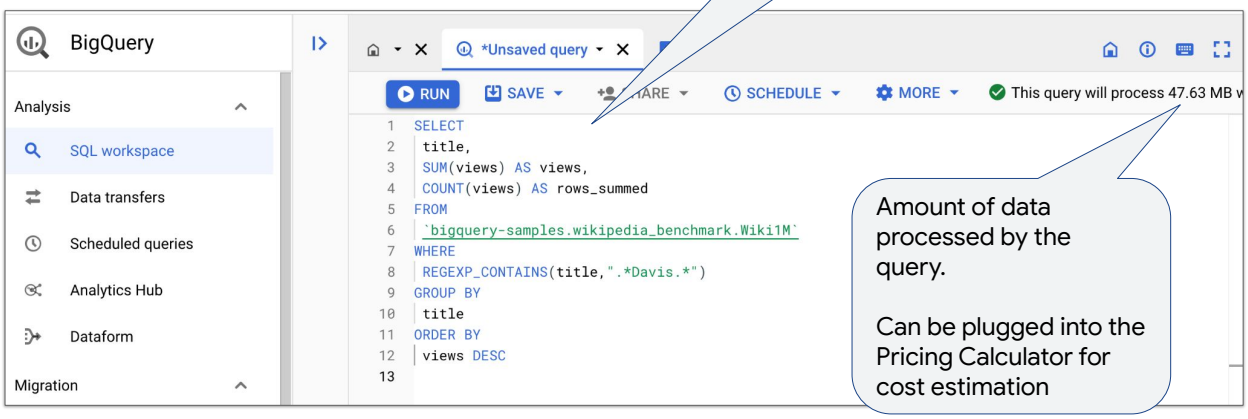
BigQuery

Google Cloud

Cloud data warehouse to power your data-driven innovation
https://cloud.google.com/bigquery

BigQuery User Interface

Can run queries in the console or schedule them to run later



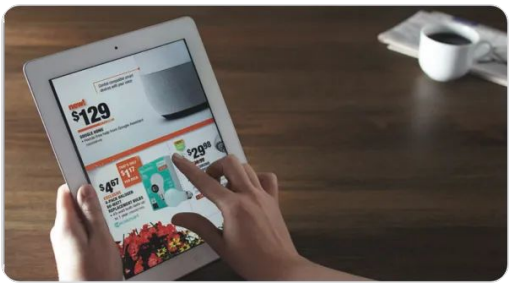
Amount of data processed by the query.

Can be plugged into the Pricing Calculator for cost estimation

Google Cloud

BigQuery - customer use case

- The Home Depot is the world's largest home improvement retailer
  - 2,300 stores in North America + online retail
  - Annual sales > $100 billion
- BigQuery provides timely data to help keep 50,000+ items stocked at over 2,000 locations, to ensure website availability, and provide relevant information through the call center
- No two Home Depots are alike, and the stock in each has to be managed at maximum efficiency.
  - Migrating to Google Cloud, THD's engineers built one of the industry's most efficient stock replenishment systems

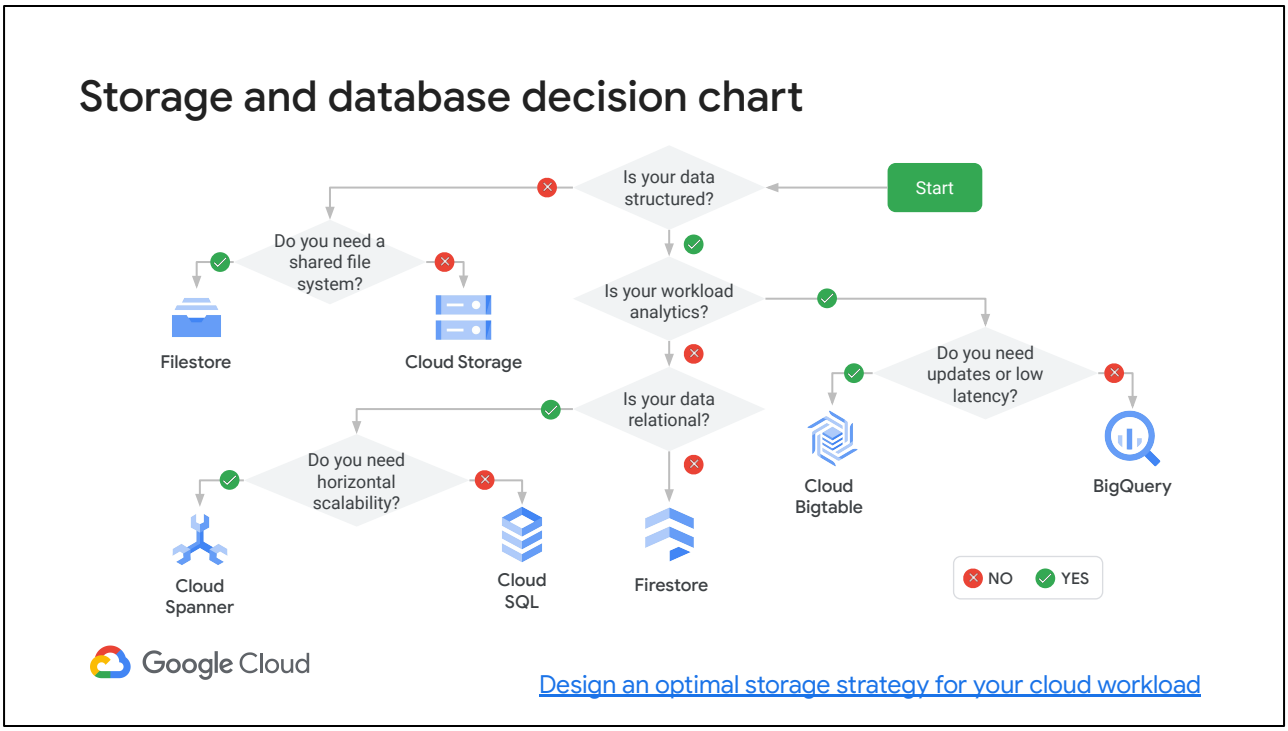


The Home Depot's data-driven focus on customer success

Google Cloud

About The Home Depot
https://corporate.homedepot.com/page/about-us

The Home Depot's data-driven focus on customer success
https://cloud.google.com/customers/featured/the-home-depot



Design an optimal storage strategy for your cloud workload
https://cloud.google.com/architecture/storage-advisor

Let's summarize the services in this module with this decision chart:

- First, ask yourself: Is your data structured, and will it need to be accessed using its structured data format? If the answer is no, then ask yourself if you need a shared file system. If you do, then choose Filestore.
- If you don't, then choose Cloud Storage.
- If your data is structured and needs to be accessed in this way, than ask yourself, does your workload focus on analytics? If it does, you will want to choose Cloud Bigtable or BigQuery, depending on your latency and update needs.
- Otherwise, check whether your data is relational. If it's not relational, choose Firestore.
- If it is relational, you will want to choose Cloud SQL or Cloud Spanner, depending on your need for horizontal scalability.

Depending on your application, you might use one or several of these services to get the job done. For more information on how to choose between these different services, please refer to the following two links:

https://cloud.google.com/storage-options/
https://cloud.google.com/products/databases/

Storage and database services

Object	File	Relational		Non-relational		Warehouse	In memory
Cloud Storage	Filestore	Cloud SQL	Cloud Spanner	Firestore	Cloud Bigtable	BigQuery	Memorystore
Good for: Binary or object data	**Good for:** Network Attached Storage (NAS)	**Good for:** Web frameworks	**Good for:** RDBMS + scale, HA, HTAP	**Good for:** Hierarchical, mobile, web	**Good for:** Heavy read + write, events,	**Good for:** Enterprise data warehouse	**Good for:** Caching for Web/Mobile apps
Such as: Images, media serving, backups	**Such as:** Latency sensitive workloads	**Such as:** CMS, eCommerce	**Such as:** User metadata, Ad/Fin/ MarTech	**Such as:** User profiles, game state	**Such as:** AdTech, financial, IoT	**Such as:** Analytics, dashboards	**Such as:** Game state, user sessions

Google Cloud

Next discussion

Memorystore

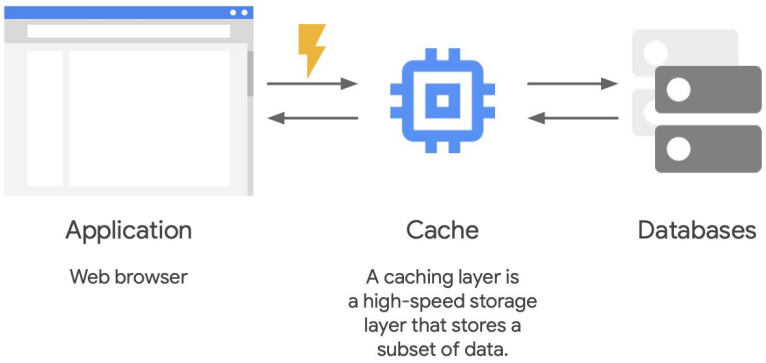
- Fully managed implementation of the open source in-memory databases Redis and Memcached
- High availability, failover, patching and monitoring
- Sub-millisecond latency
- Instances up to 300 GB
- Network throughput of 12 Gbps
- Use cases:
  - Lift and shift of Redis, Memcached
  - Anytime need a managed service for cached data

Memorystore

Google Cloud

Memorystore
https://cloud.google.com/memorystore

In-memory caching



Benefit

- Reduce latency
- Reduce back-end load

Main use case

- Gaming leaderboard
- Real-time application
- Social Media

Google Cloud

```
gcloud redis instances create my-redis-instance
gcloud memcache instances create my-memcache-instance
```

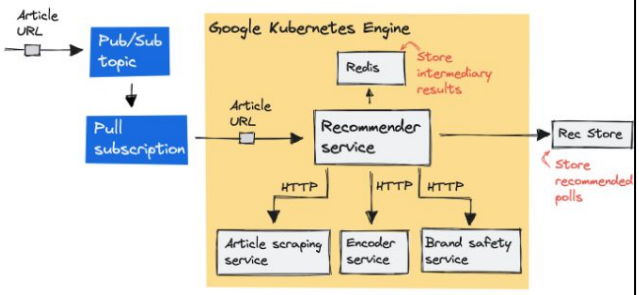
When to use Cloud Memorystore

Application domain	Game	Social Media	IoT	Analytics
Cloud Memorystore -driven capabilities	Leaderboard, advertisement	Chat, news feed	Real-time transactions, sensor analytics	Read-heavy dashboard

Google Cloud

Memorystore - customer use case

- Opinary creates polls that appear alongside news articles on various sites around the world
  - Machine learning is used to decide which poll to display by which article
- The polls let users share their opinion with one click and see how they compare to other readers.
- Publishers benefit by increased reader retention, and increased subscriptions
- Advertisers benefit from high-performing interaction with their audiences



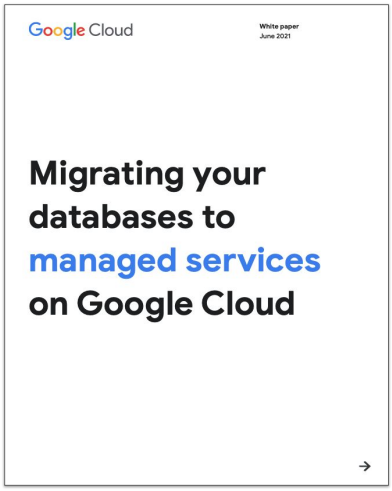
Opinary generates recommendations faster on Cloud Run

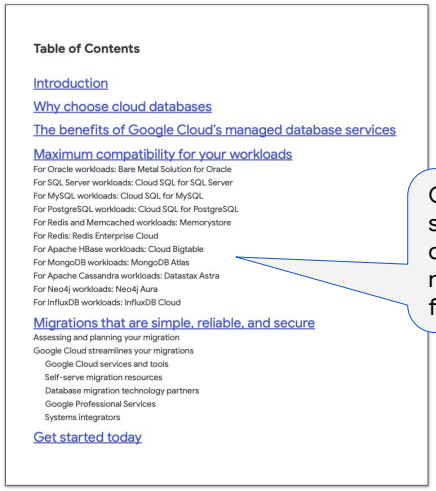
Google Cloud

Opinary generates recommendations faster on Cloud Run
https://cloud.google.com/blog/topics/developers-practitioners/opinary-generates-recommendations-faster-cloud-run/

More about Opinary (if interested)
https://opinary.com/

Google white paper - database migration



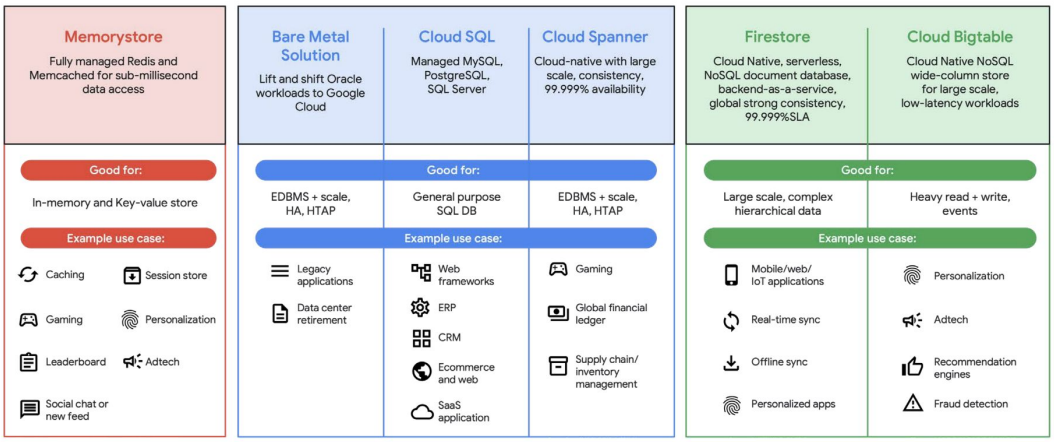


Covers the services discussed in this module plus a few more

Migrating Databases

Google Cloud

Which database should I use?



https://services.google.com/fh/files/misc/guide_to_google_cloud_databases.pdf

Google Cloud

Summary: Choosing the correct database

What are the requirements?

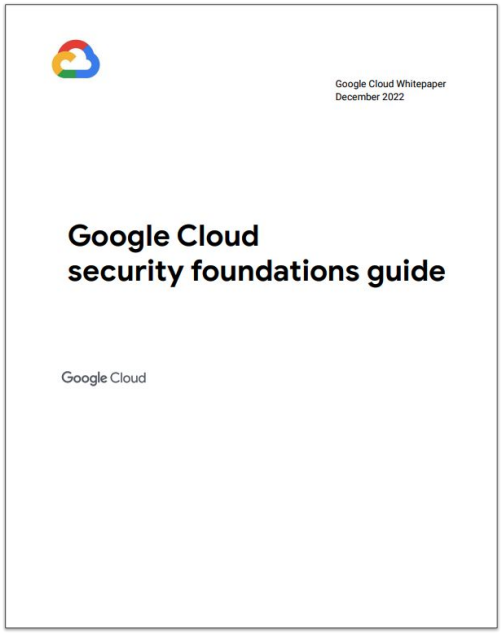
- Structured vs Unstructured
- Relational (SQL) vs Non-Relational (NoSQL)
- Transactional (OLTP) vs Analytical (OLAP)
- Fully Managed vs Requires Provisioning
- Global vs Regional

Google Cloud

IAM and Resource Hierarchy...

Google Cloud

Security Foundations Guide



Google Cloud

This is a highly recommended read for anyone working in Google Cloud.

Key items include:

I. Security Foundations Blueprint
II. Step-by-step guide

Google Cloud security foundations guide
Curated, opinionated guidance and accompanying automation that helps you build a secure starting point for your Google Cloud deployment.
https://services.google.com/fh/files/misc/google-cloud-security-foundations-guide.pdf

All Google Cloud white papers
https://cloud.google.com/whitepapers

From the exam guide

3.1 Designing for security. Considerations include:

- **Identity and access management (IAM)**
- **Resource hierarchy (organizations, folders, projects)**
- Data security (key management, encryption, secret management)
- **Separation of duties (SoD)**
- Security controls (e.g., auditing, VPC Service Controls, context aware access, **organization policy**)
- Managing customer-managed encryption keys with Cloud Key Management Service
- Remote access

Google Cloud

Principle of least privilege

Least privilege is a practice of granting a user only the minimal set of permissions required to perform a duty.

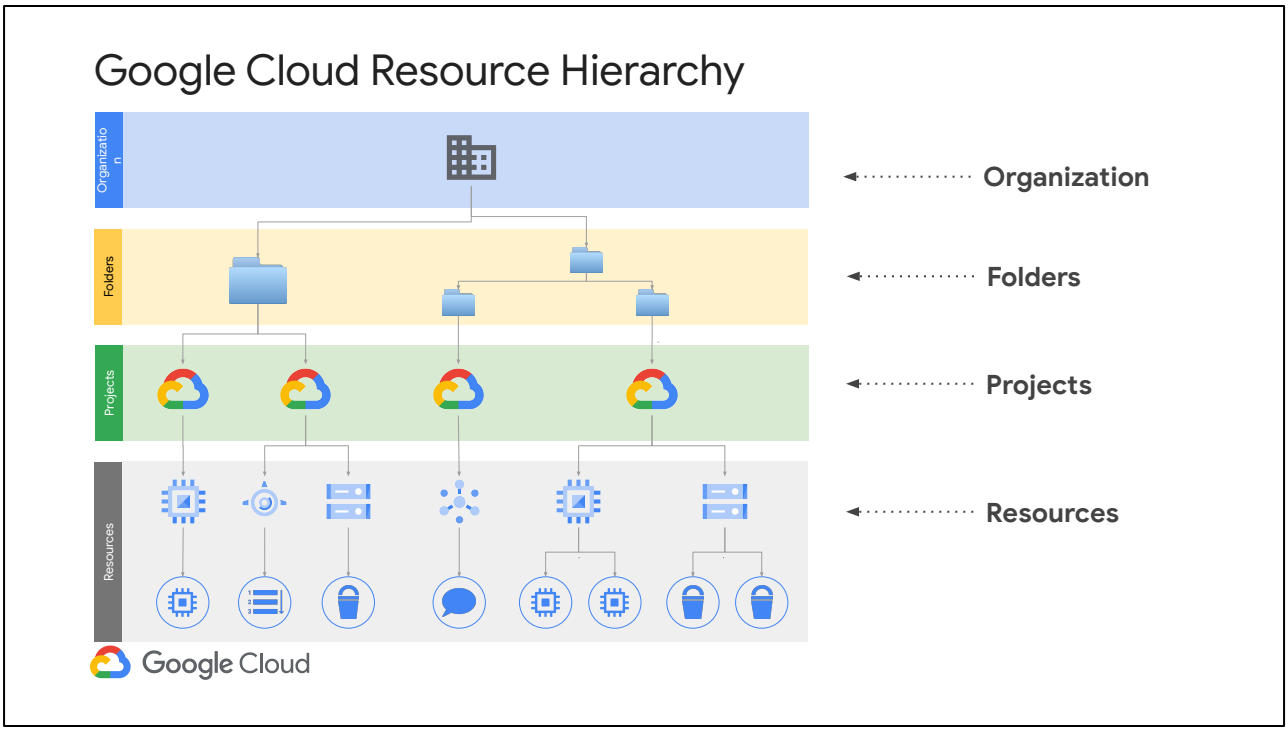
- Users should only be able to do the tasks that are required by their jobs
- This should also apply to machine instances and run-time processes

Google Cloud

Separation of duties

Examples	Implementation Suggestions
• No one person can change or delete data without being detected • No one person is in charge of designing, implementing, and reporting on sensitive systems • The people who write the code shouldn't deploy the code; those who deploy the code shouldn't be able to change it	• Use multiple projects to separate duties • Different people can be given different rights in different projects • Use folders to help organize projects

Google Cloud



Resource hierarchy
https://cloud.google.com/resource-manager/docs/cloud-platform-resource-hierarchy

Resource hierarchy and billing:
https://cloud.google.com/billing/docs/concepts

Organization and Folders

Organization is a registered domain	Folders can be added to the organization	Projects are added to either the organization or to a folder	Resources are added to projects
Rights can be granted at the organization level	Rights can be granted to folders Folders can contain other folders	Rights can be granted at the project level	Rights can be granted at the resource level

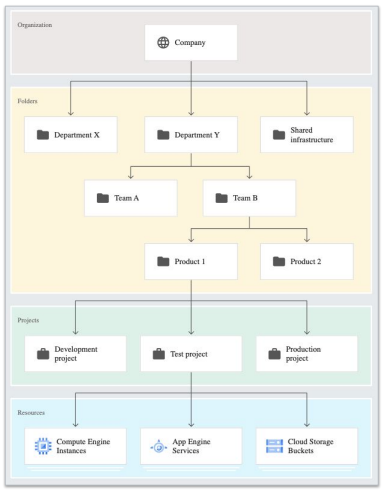
Google Cloud

The IAM Hierarchy is organized in four distinct levels. The levels are:

- Organization, which is a registered domain. Rights granted at the organization level are applied to the entire infrastructure.
- Folders are an optional level used to organize projects. Rights can be granted to folders and folders can contain other folders.
- Projects are accounts that can be added to an organization or folder. Rights granted at the project level are applied to all resources in a project.
- Resources are the products and services used within a project. Rights can also be granted at the resource level.

Folders provide a logical way to organize teams and projects

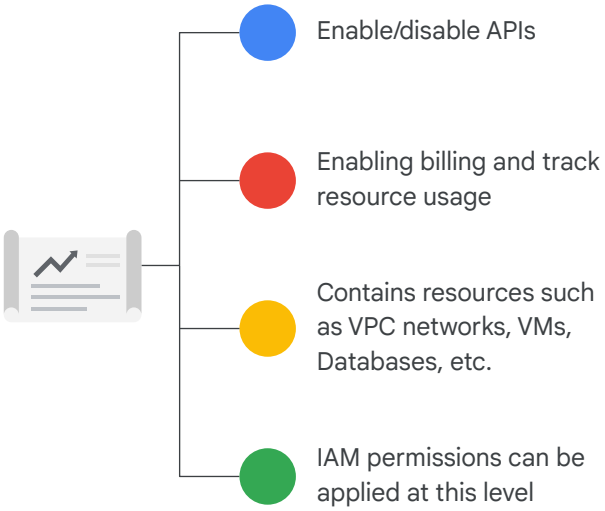
- Must have an Organization node in order to create folders
- Folders contain projects, other folders, or both
- Roles and Organization policies can be applied at the folder level (and also at the project/organization level)



Google Cloud

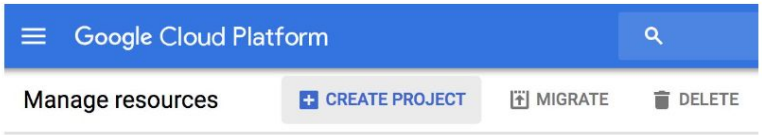
https://cloud.google.com/resource-manager/docs/cloud-platform-resource-hierarchy#folders

All products & services are associated with a project



Google Cloud

Creating a project



Project ID	Globally unique	Chosen by you	Immutable
Project name	Need not be unique	Chosen by you	Mutable
Project number	Globally unique	Assigned by Google Cloud	Immutable

Google Cloud

Creating and managing projects
https://cloud.google.com/resource-manager/docs/creating-managing-projects

Resources

- Resources are anything created when using services
- All resources are associated with a project
- Examples of resources are:
  - Virtual machines
  - Persistent disk
  - Cloud Storage buckets
  - BigQuery datasets and tables
  - Spanner databases
  - Kubernetes Engine clusters
- Must enable service-specific APIs before creating resources within a project

Google Cloud

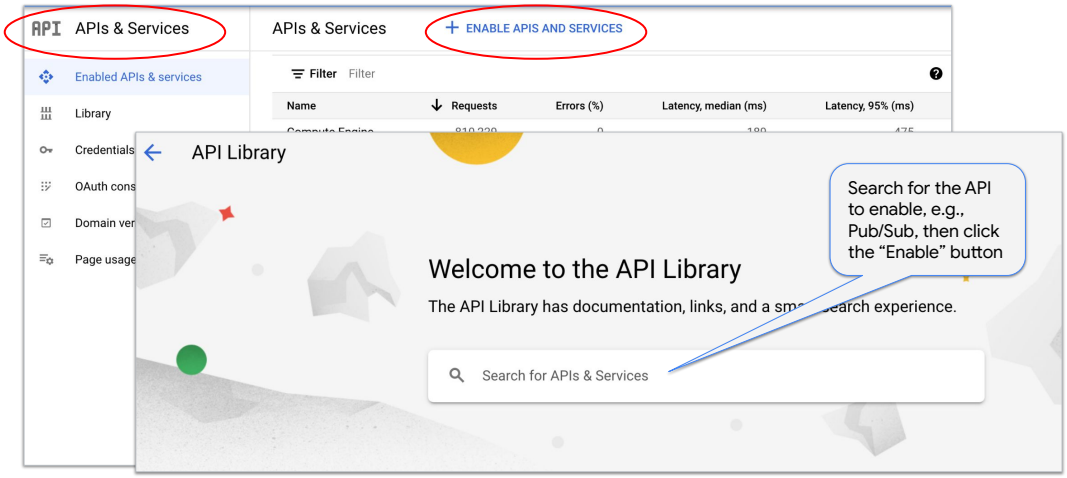
All resources must be associated with a project, so proper billing can be implemented.

Examples of resources include:

- Virtual machines
- Disks
- Cloud Storage buckets
- BigQuery datasets and tables
- Spanner databases
- Kubernetes Engine clusters

These are just a small sample

Enabling APIs - Console & Command Line



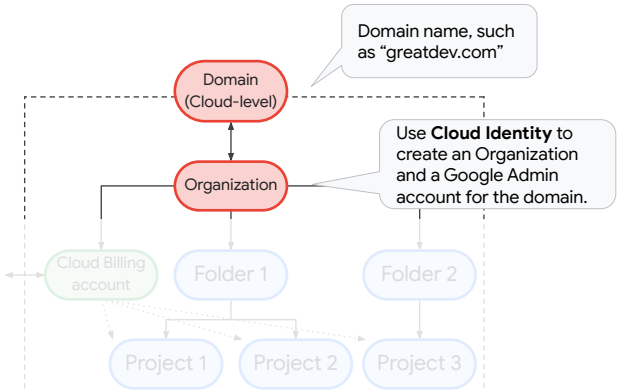
Google Cloud

```
gcloud services enable pubsub.googleapis.com
```

Showing both console and command line

How is an Organization Created?

- **Cloud Identity** manages the users and groups that have access to Google Cloud
  - Federated identities from Google Workspace and other identity providers, such as Active Directory and Azure Active Directory
    - Bring existing users/groups into Cloud Identity
    - Use Identity and Access Management (IAM) to manage access to Google Cloud resources



Google Cloud

Cloud Identity:
https://cloud.google.com/identity

Cloud Identity centrally manages users and groups for Google Workspace and Google Cloud

Google's Identity as a Service (IDaaS) solution

- Users and groups that are to be added to Google Cloud need accounts in Cloud Identity

Manually creating user accounts

- Add users individually using the Google Admin console
- Add several users at once by uploading their names in a CSV file

Focus of this module

Options for large organizations

- Use Google Cloud Directory Sync to synchronize user data in your existing LDAP directory with your Google account
- Use the Admin SDK Directory API to provision a large number of users with data from your existing LDAP directory, such as Microsoft® Active Directory®
  - Requires programming

Google Cloud

*Cloud Identity has advanced management features not covered in this module, e.g., mobile app management, 2-Step verification, etc.

Overview of Cloud Identity
https://cloud.google.com/identity/docs/overview

Features provided by Cloud Identity
https://cloud.google.com/identity/docs/editions

Creating users and groups
https://support.google.com/cloudidentity/answer/7332836?hl=en&ref_topic=7558418

Proprietary + Confidential

Google Cloud Directory Sync (GCDS)

Best practice

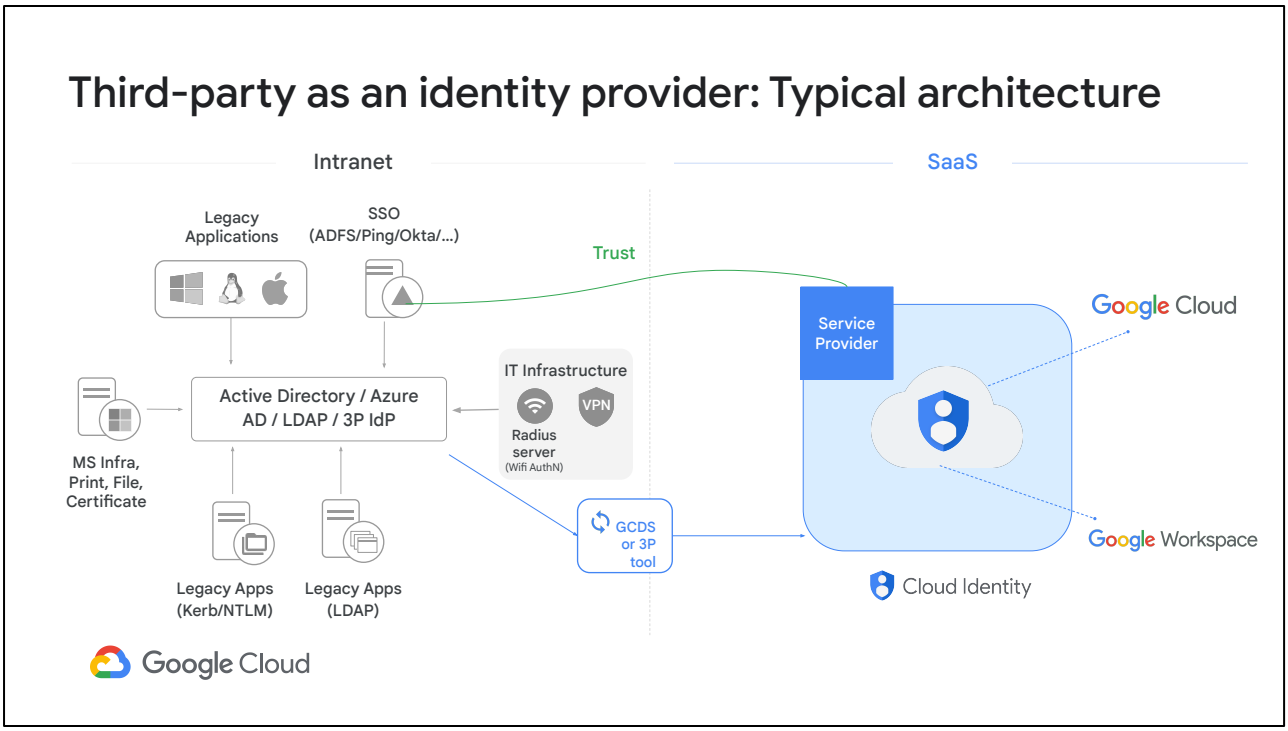
Provision from a single source of truth

- One-way synchronization of corporate data (no writing to LDAP system)
- Only synchronizes deltas for fastest possible provisioning
- Syncs all object types (users, aliases, profiles, groups)
- Utilizes Google APIs to provision all object types, the same APIs available to customers

Google Cloud

Google Cloud Directory Sync examples
https://cloud.google.com/community/tutorials/gcds-use-cases-common-and-complex

- Recommendation is to **provision from single source of truth** (i.e. single consolidated LDAP source). Potential multi-directory pitfalls:
  - Involves creating highly complex configuration files
  - Difficult to troubleshoot problems
  - Difficult to maintain configuration files
  - High probability of inadvertently deleting objects
- Can be installed on **Windows or Linux**
- Easy to schedule (**Task Scheduler / cron jobs**) or run manually
- Helps automate onboarding and off-boarding scenarios
- Supports **multiple authentication methods**



Google Cloud Directory Sync (GCDS) can be used to import users from a 3rd party Identity Provider into Cloud Identity

User will authenticate with the third-party IdP (AD Federation Service in this diagram, but it can also be Azure AD, Ping, OKTA, etc.)

Proprietary + Confidential

Controlling access

Authentication	Authorization	Auditing
Cloud Identity	Identity Access Management (IAM)	Google Cloud's Operations suite Cloud Audit Logs & Reports API

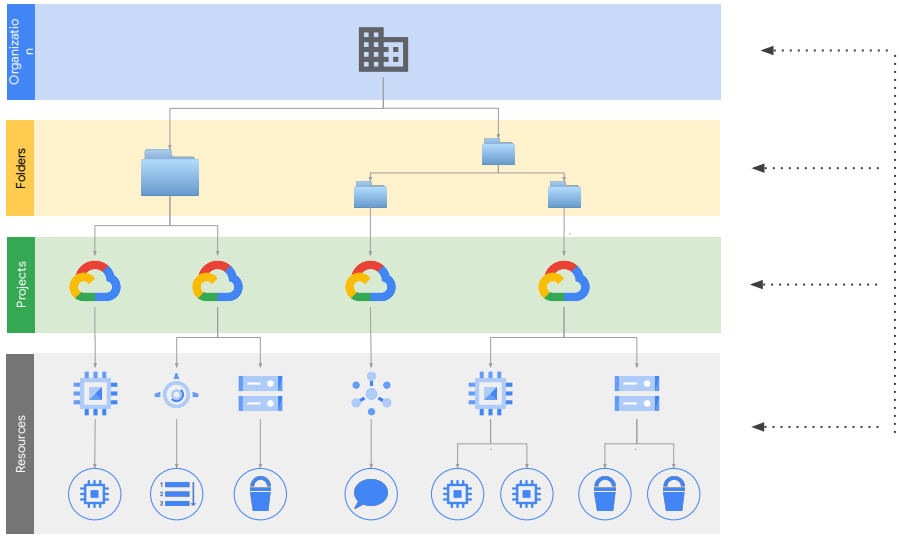
Google Cloud

- Authentication: identifying a user typically done through a process in which the user provides a private form of verification (for example, password, a key, etc.). We use Cloud Identity for AuthN

- Authorization: set of permissions that a user is allocated post authentication. A user can be authenticated (thus confirming his identity) but authentication, on its own, provides no set of permissions. We use IAM for AuthZ (and Cloud Identity for assigning Admin roles).

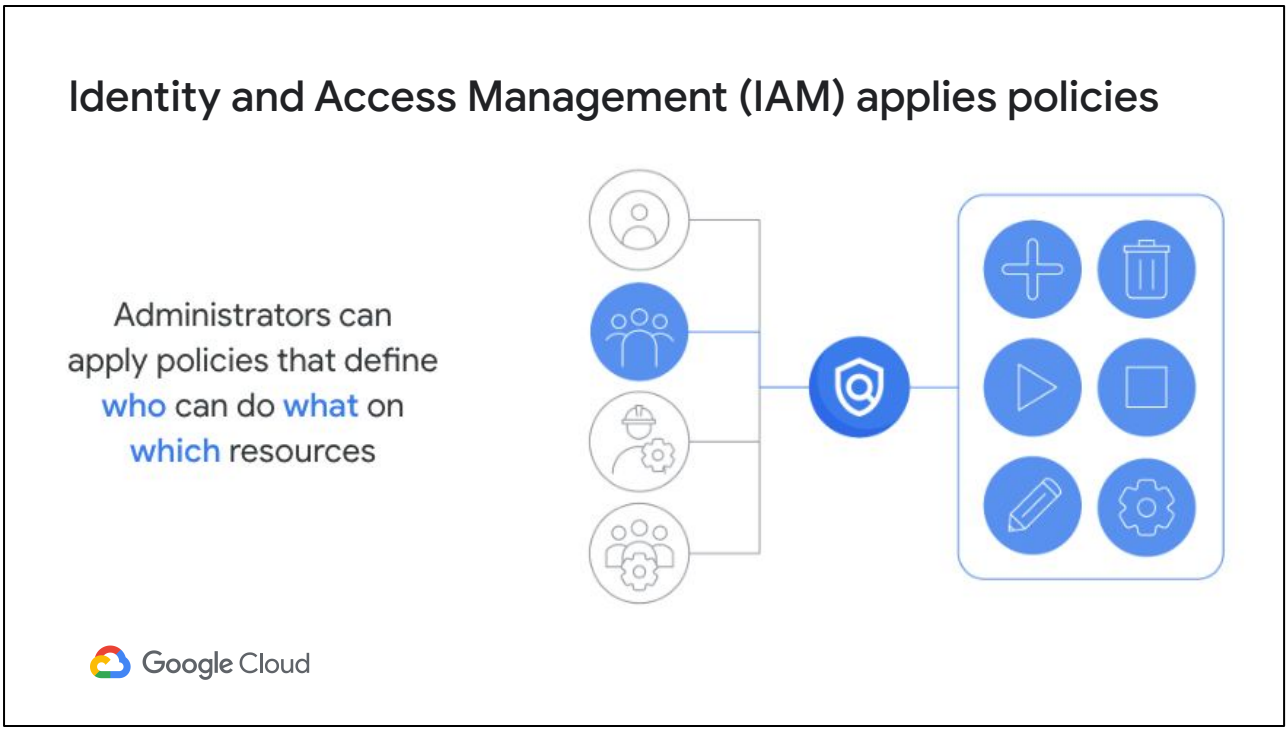
- Auditing: a form of monitoring the resources accessed or modified by a particular identity. We use Google Cloud operations suite (formerly known as Stackdriver), Cloud Audit Logs for auditing in Google Cloud, Reports API for auditing in Cloud Identity operations.

Hierarchy inheritance



Google Cloud

Next discussion: Identity and Access Management



IAM overview:
https://cloud.google.com/iam/docs/overview

When an organization node contains lots of folders, projects, and resources, it's likely there is a need to restrict who has access to what.

To help with this task, administrators can use **Identity and Access Management**, or IAM. With IAM, administrators can apply policies that define *who can do what on which resources*.

- The "who" part of an IAM policy can be a Google account, a Google group, a service account, or Cloud Identity domain.
- The "can do what" part of an IAM policy is defined by a role. An IAM role is a collection of permissions. For example, to manage virtual machine instances in a project, you have to be able to create, delete, start, stop and change virtual machines. So these permissions are grouped together into a role to make them easier to understand and easier to manage.

Who ...?

Service Accounts are discussed later

The “who” part of an IAM policy can be a

- Google account
- Google group
- Service account
- Google Workspace or Cloud Identity domain

Also

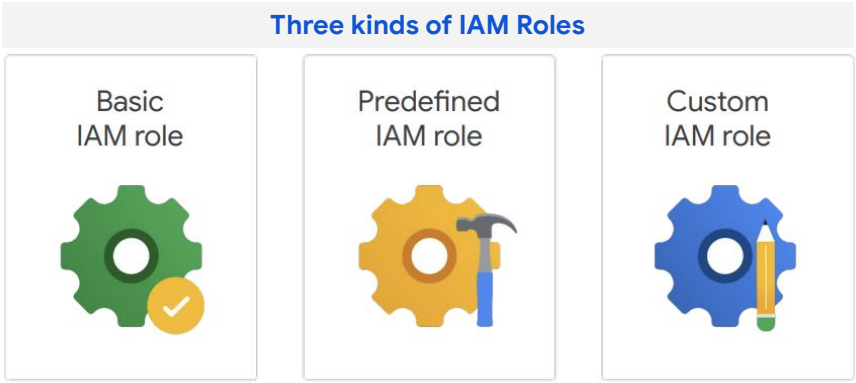
- allAuthenticatedUsers
- allUsers

Google Cloud

Concepts related to identity
https://cloud.google.com/iam/docs/overview#concepts_related_identity

... can do what ...?

- The “can do what” is defined by an IAM role.



Google Cloud

Role types
https://cloud.google.com/iam/docs/roles-overview#role-types

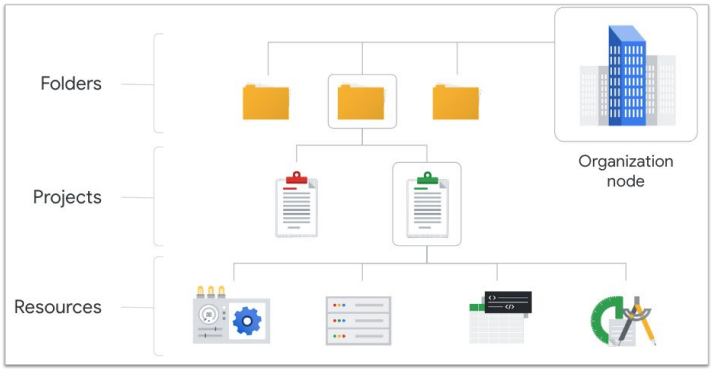
Basic roles, which include the Owner, Editor, and Viewer roles that existed prior to the introduction of IAM.

Predefined roles, which provide granular access for a specific service and are managed by Google Cloud.

Custom roles, which provide granular access according to a user-specified list of permissions.

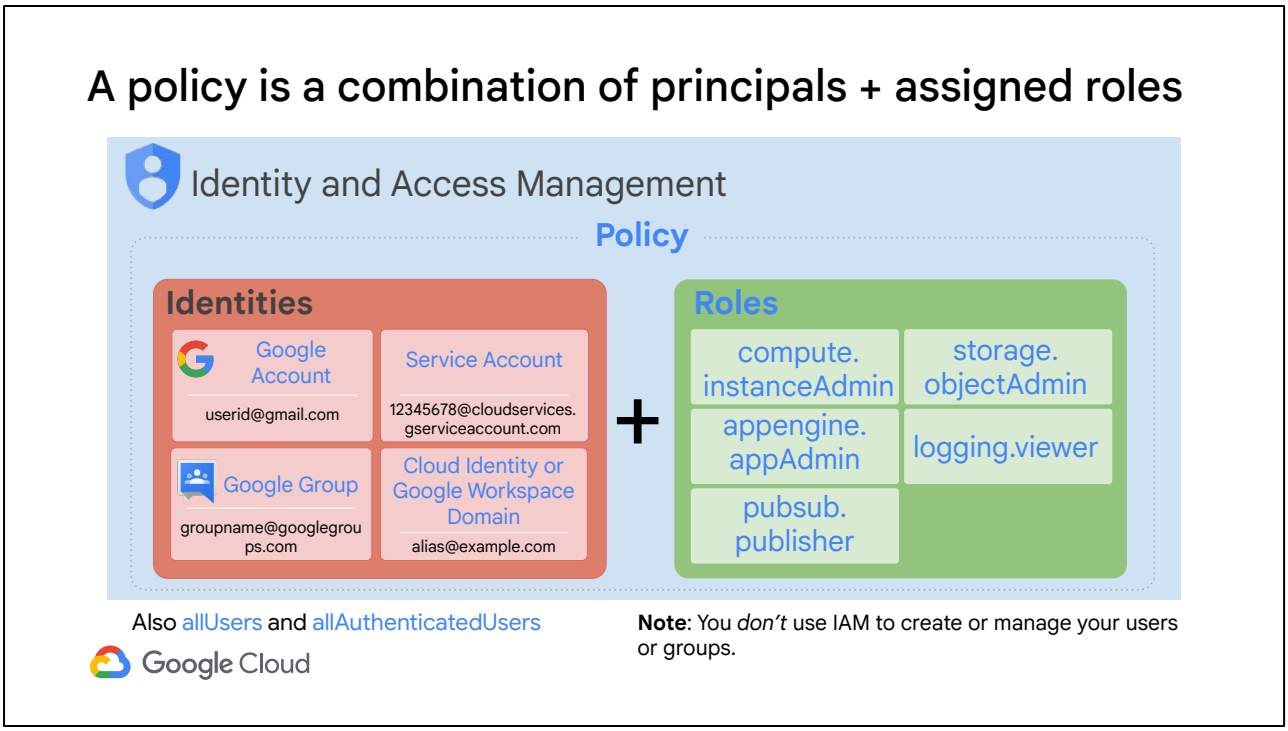
… on which resource?

- Can manage IAM at the project level, folder level or organization level*



*Note: Some Cloud Storage IAM is applied at the bucket level

Google Cloud



Policy
https://cloud.google.com/iam/docs/reference/rest/v1/Policy

There are five different types of members: Google Accounts, Service Accounts, Google Groups, Google Workspace domains, and Cloud Identity domains.

A Google account represents a developer, an administrator, or any other person who interacts with Google Cloud. Any email address that is associated with a Google account can be an identity, including gmail.com or other domains. New users can sign up for a Google account by going to the Google account signup page, without receiving mail through Gmail.

A service account is an account that belongs to your application instead of to an individual end user. When you run code that is hosted on Google Cloud, you specify the account that the code should run as. You can create as many service accounts as needed to represent the different logical components of your application.

A Google group is a named collection of Google accounts and service accounts. Every group has a unique email address that is associated with the group. Google groups are a convenient way to apply an access policy to a collection of users. You can grant and change access controls for a whole group at once instead of granting or changing access controls one-at-a-time for individual users or service accounts.

A Workspace domain represents a virtual group of all the Google accounts that have

been created in an organization's Workspace account. Workspace domains represent your organization's internet domain name, such as example.com, and when you add a user to your Workspace domain, a new Google account is created for the user inside this virtual group, such as username@example.com.

Google Cloud customers who are not Workspace customers can get these same capabilities through Cloud Identity. Cloud Identity lets you manage users and groups using the Google Admin Console, but you do not pay for or receive Workspace's collaboration products such as Gmail, Docs, Drive, and Calendar. Refer to the documentation for more information on Cloud Identity Editions.

Now it's important to note that you cannot use IAM to create or manage your users or groups. Instead, you can use Cloud Identity or Workspace to create and manage users.

… can do what …?

Revisiting this slide

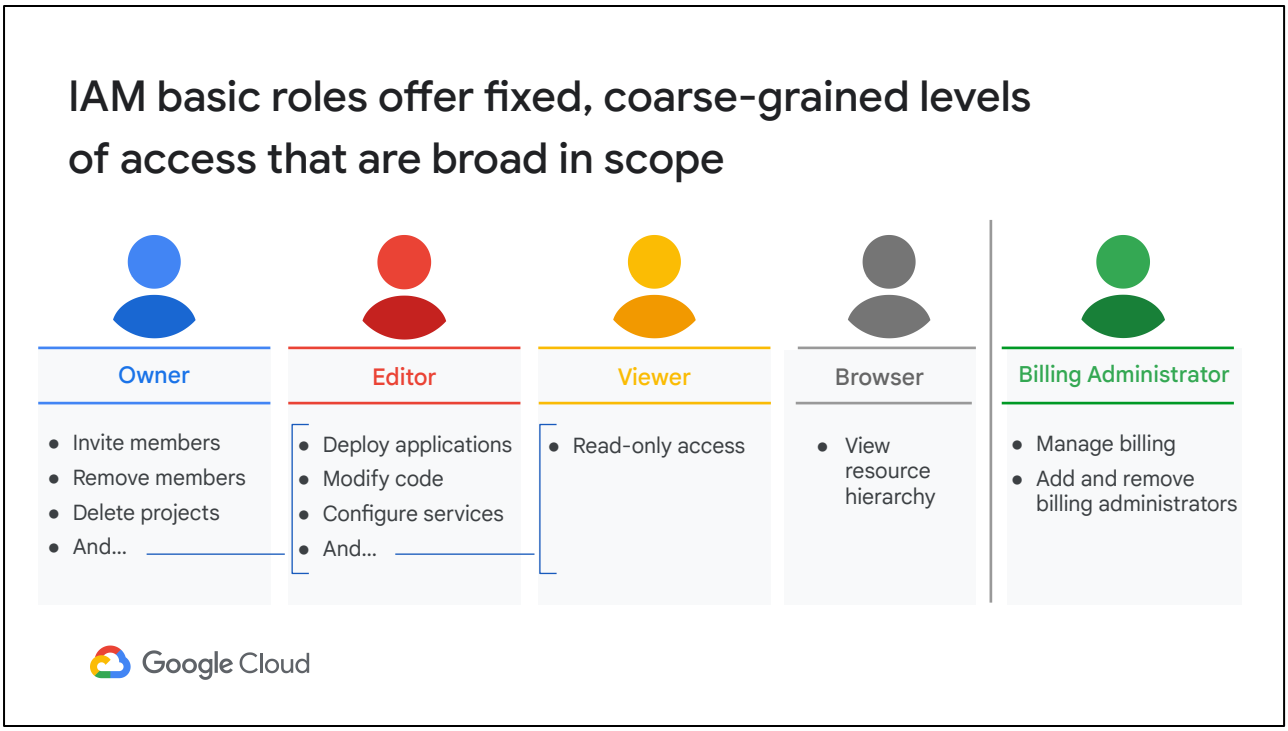
- The “can do what” is defined by an IAM role.

Basic roles discussed next

Three kinds of IAM Roles

Basic IAM role	Predefined IAM role	Custom IAM role

Google Cloud



IAM basic roles offer fixed, coarse-grained levels of access.

The basic roles are the Owner, Editor, Viewer, Browser and Billing Administrator roles.

- The owner has full administrative access. This includes the ability to add and remove members and delete projects.
- The editor role has modify and delete access. This allows a developer to deploy applications and modify or configure its resources.
- The viewer role has read-only access
- The browser role can view only the resource hierarchy

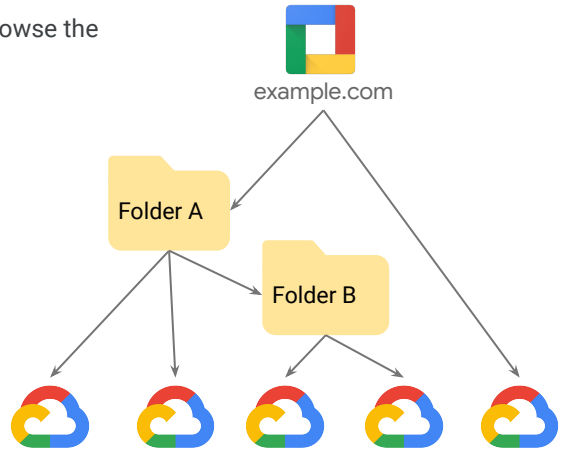
The Owner role includes the permissions of the Editor role, and the Editor role includes the permissions of the Viewer role.

There is also a billing administrator role to manage billing and add or remove administrators without the right to change the resources in the project.

Each project can have multiple owners, editors, viewers, and billing administrators.

The **Browser** basic role

This role provides read access to browse the hierarchy for a project, including the organization and folders.



Google Cloud

The predefined Browser role provides read-access to browse the **hierarchy** for a project, including the folder, organization, and IAM policy.

The Browser role does not include permission to view resources in the project.

… can do what …?

Revisiting this slide

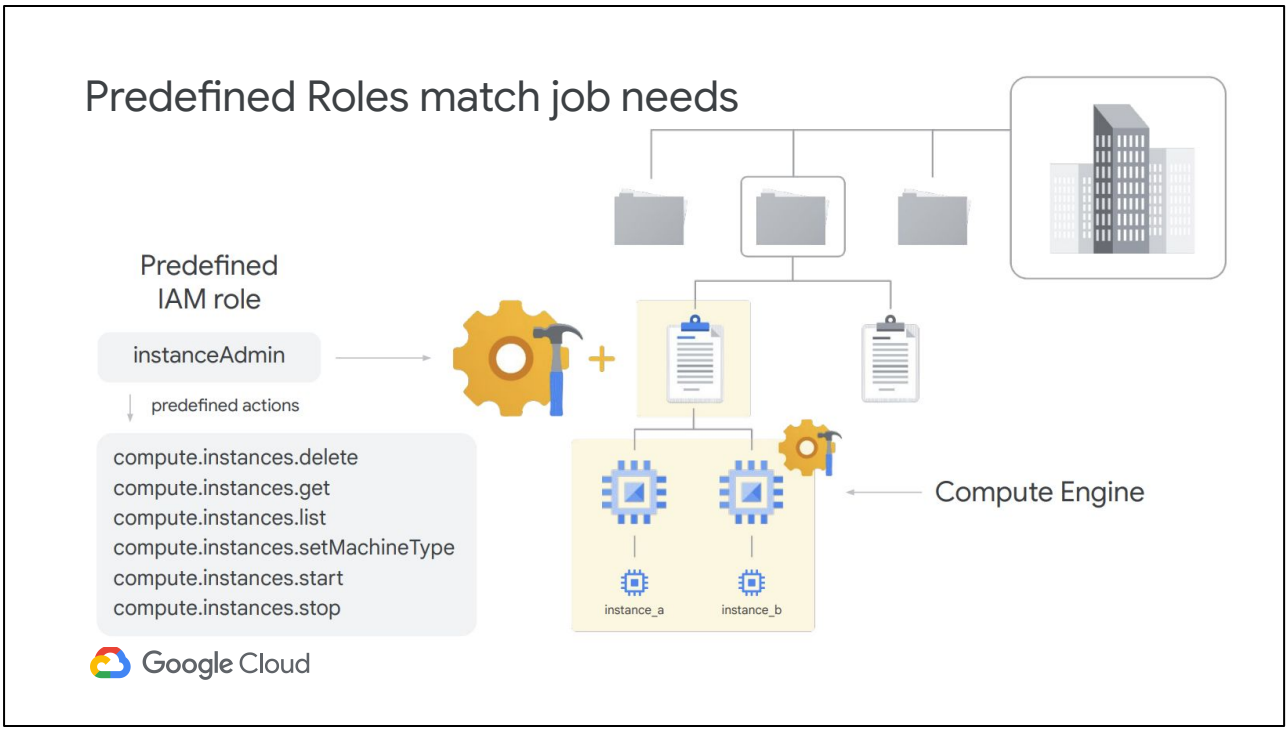
- The “can do what” is defined by an IAM role.

Three kinds of IAM Roles

Basic IAM role	Predefined IAM role	Custom IAM role

Predefined roles discussed next

Google Cloud



Google Cloud uses roles as a method to distribute permissions.

Members are assigned one or more roles. The roles a member is assigned will determine the permissions of a member.

A role is a list of permissions organized by a service.

This list of permissions correlate with Google Cloud API calls and are aligned based on job functions.

The example shown are some of the roles for Compute Engine Instance Admin

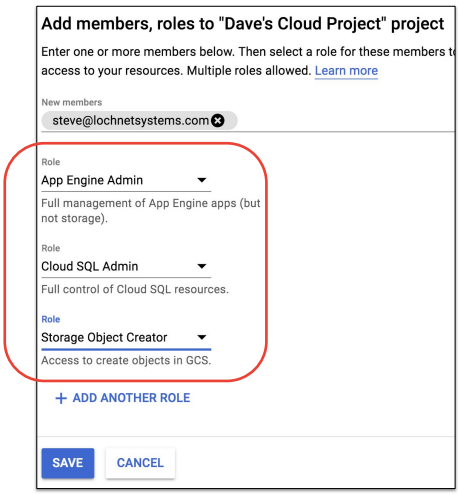
Using predefined roles

Google-created roles for each service

- Permissions defined for different job roles working with each service
- Permissions maintained by Google

Best practice is to avoid the basic roles and use predefined roles when adding members or service accounts

- Principle of least privilege



Google Cloud

Google provides roles for each service. There are hundreds of predefined roles in Google Cloud.

Predefined roles are based on different job roles when working with different services.

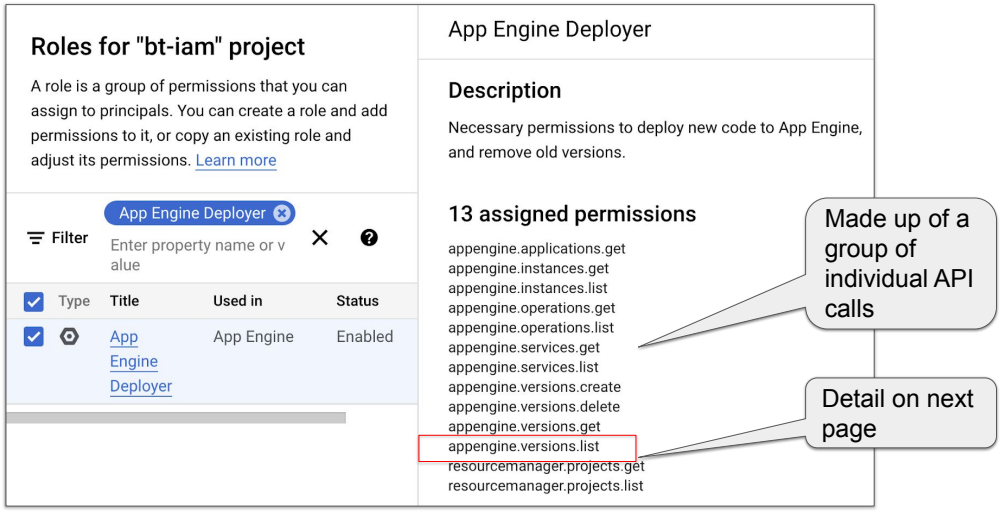
And are maintained by Google.

A Google best practice is to avoid using basic roles. Basic roles may not be specific enough.

Always follow the principle of least privilege.

Role permissions

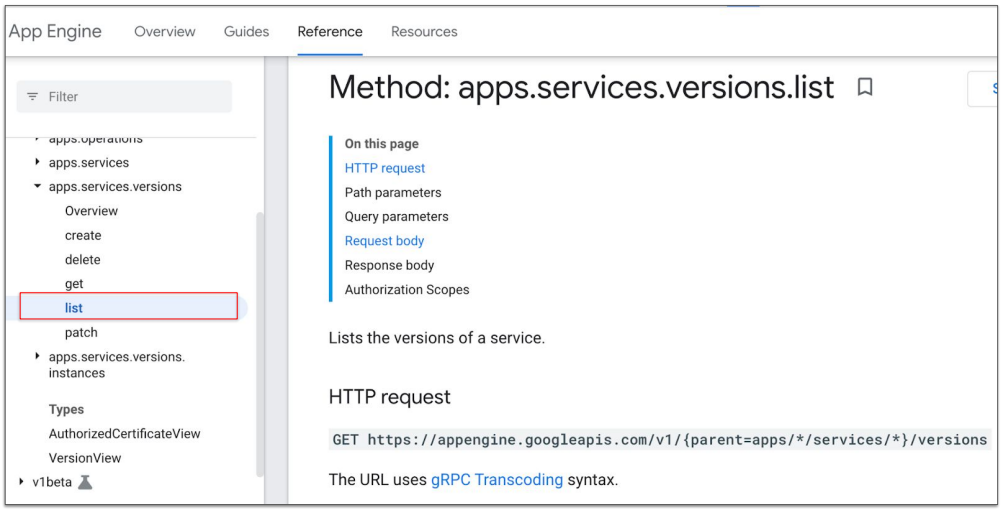
- Cloud Console > IAM & Admin > Roles



Google Cloud

IAM basic and predefined roles reference
https://cloud.google.com/iam/docs/understanding-roles

App Engine API Documentation



Google Cloud

https://cloud.google.com/appengine/docs/admin-api/reference/rest/v1/apps.services.versions/list

… can do what …?

Revisiting this slide

- The “can do what” is defined by an IAM role.

Three kinds of IAM Roles

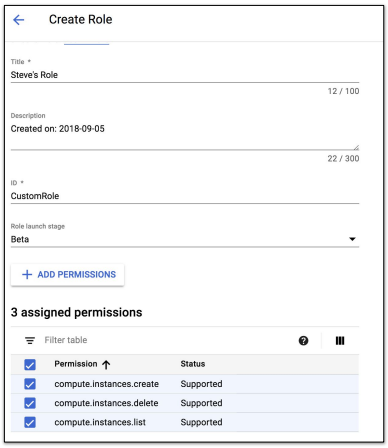
Basic IAM role	Predefined IAM role	Custom IAM role

Custom roles discussed next

Google Cloud

Custom Roles

- Roles that you create
  - Fine-grained control over permissions
- Can add any permissions you like
- Can create custom roles based on predefined roles and add /remove permissions
- Custom roles add operational overhead
  - You must maintain the permissions
- Applied at the project or organization level



Google Cloud

Custom roles are an option when there is not a predefined role available that provides the permissions you would like to group.

Custom roles give you fine-grained control over permissions by allowing you to add any permission you like to a role you create.

You can create custom roles by copying and modifying a predefined role or you can create one from scratch.

Custom roles add operational overhead, as you are responsible for maintaining the permissions of a custom role.

Custom roles can only be applied either to project level or organization level. They can't be applied to the folder level.

Deny Policies

- Introduced in 2022
- Inherited through the resource hierarchy just like IAM allow policies
- Attached to project, folder or organization
- Denies override grants further down the hierarchy
- Currently, must be created via command line

First create a deny policy and store it in a file

People in the dev@example.com group are not allowed to create or delete service account keys

```
{
  "deniedPrincipals": [
    "principalSet://goog/group/dev@example.com"
  ],
  "deniedPermissions": [
    "iam.googleapis.com/serviceAccountKeys.create",
    "Iam.googleapis.com/serviceAccountKeys.delete"
  ]
}
```

Next apply the deny policy

```
gcloud iam policies create POLICY_ID \
    --attachment-point=[proj-id|folder-id|org-id] \
    --kind=denypolicies \
    --policy-file=POLICY_FILE
```

Google Cloud

Deny policies
https://cloud.google.com/iam/docs/deny-overview

Who ...?

Saw this slide earlier

Next topic: Service accounts

The “who” part of an IAM policy can be a

- Google account
- Google group
- **Service account**
- Google Workspace or Cloud Identity domain

Also

- allAuthenticatedUsers
- allUsers

Google Cloud

There are two types of Google Service Accounts

Google-managed service accounts

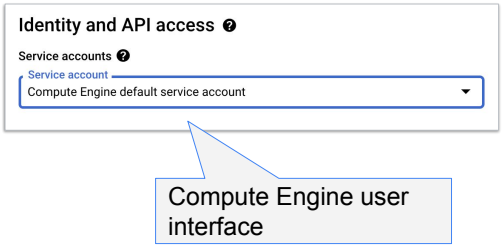
- Created and managed by Google
- Used by Google to perform operations on your behalf, for example
  - Moving a VM during a maintenance event
- These activities are logged to Cloud Logging for full transparency

User-managed service accounts

- Created and managed by someone with Service Account Admin or Owner IAM
  - Are given various IAM roles depending on the use case
- Attached to services, which limits those services to the actions allowed by the roles assigned to the service account

Service accounts are identities for Google Cloud services

- Special type of account intended to represent a non-human user that needs to authenticate and be authorized to access data
  - Are created *within* Google Cloud, unlike other Principals
- Attached to VMs or other services
  - Applications will only be allowed to perform actions allowed by the roles given to the service account



Google Cloud

Understanding Service Accounts

For example, a service account is created and granted the Compute Engine Instance Admin role. This service account is attached to VM-1. Any applications running in VM-1 would be able to create, modify, and delete other VMs.

Nice set of Youtube videos explaining service accounts:
What are service accounts:
https://www.youtube.com/watch?v=xXk1YlkKW_k&list=PLIivdWyY5sqIlPnZ7cvkg2Ck-8ZZ8TA5t&index=1

Creating, managing and retiring Service accounts:
https://www.youtube.com/watch?v=2TdLmI3G5Rc&list=PLIivdWyY5sqIlPnZ7cvkg2Ck-8ZZ8TA5t&index=2

How to secure your service accounts:
https://www.youtube.com/watch?v=CTcQNPWNRkE&list=PLIivdWyY5sqIlPnZ7cvkg2Ck-8ZZ8TA5t&index=3

Service account keys and impersonation:
https://www.youtube.com/watch?v=SDhMwyyd9_0&list=PLIivdWyY5sqIlPnZ7cvkg2Ck-8ZZ8TA5t&index=4

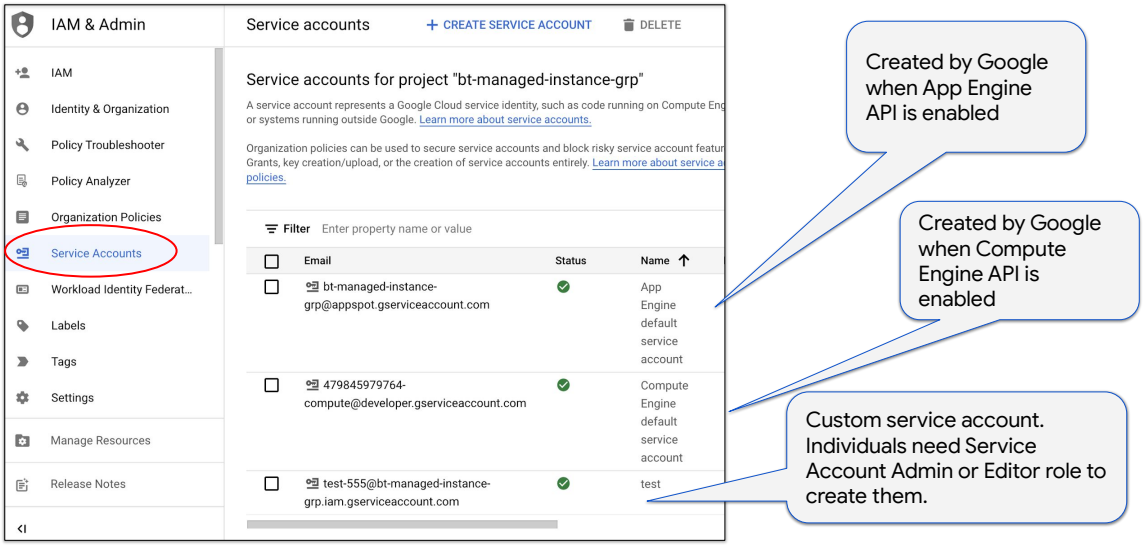
Service accounts in action:
https://www.youtube.com/watch?v=UhYqag7Xjhw&list=PLIivdWyY5sqIlPnZ7cvkg2Ck-8ZZ8TA5t&index=5

Service account use cases

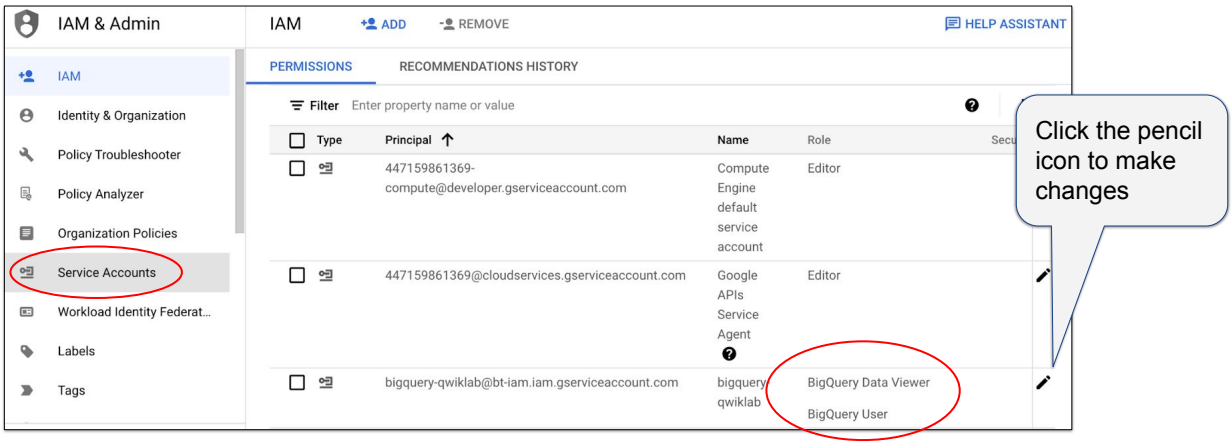
- Typically, service accounts are used in scenarios such as:
  - Running workloads on virtual machines (VMs)
    - E.g., Create a service account with permissions to query BigQuery
      - Attach it to a VM
      - Deploy an application onto the VM that submits SQL commands to BigQuery
  - Running workloads on on-premises workstations or data centers that call Google APIs.
    - E.g., Same example as above, but now the application is running on-premise
  - Running workloads which are not tied to the lifecycle of a human user.
    - E.g., Batch jobs that are scheduled to run periodically

Google Cloud

Viewing Service Accounts in the Console



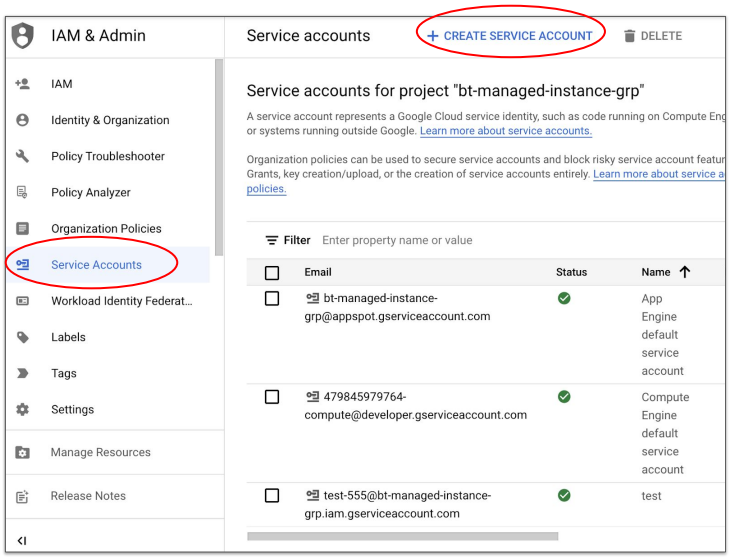
Viewing Service Accounts IAM in the Console



Google Cloud

Creating Service Accounts in the Console

IAM & Admin > Service Accounts> Create Service account



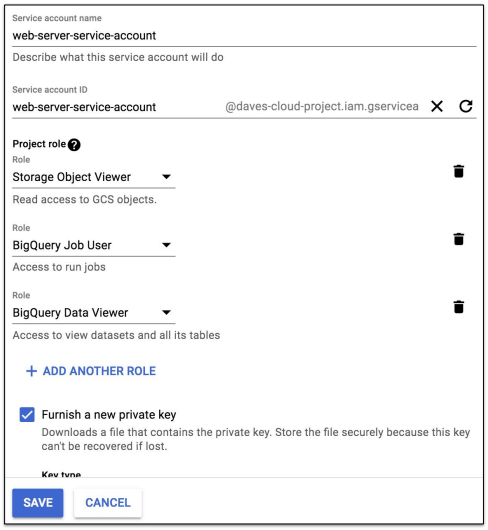
Google Cloud

Creating service accounts
https://cloud.google.com/iam/docs/creating-managing-service-accounts

Creating Service Accounts (continued)

Service account name

Add one or more roles



Here is an example of creating a service account.

First, give your service account a name. A best practice is to give the account a name that can be used to easily determine the purpose of the service account.

Next, you add one or more roles to associate with the service account. The roles will determine the permissions associated with the account.

If required, you can also create a JSON or P12 key to apply to applications and services.

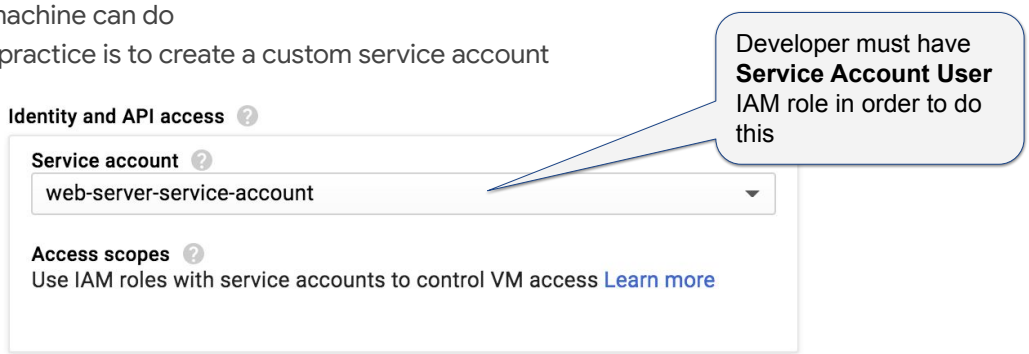
Attaching service accounts to resources

- Service accounts can be attached to multiple services in Google Cloud, including
  - Compute Engine (and Kubernetes Engine)
    - Default service account is created by Google when API is enabled
  - App Engine
    - Default service account is created by Google when API is enabled
  - Cloud Run
    - Uses the Compute Engine default service account
  - Cloud Functions
    - Uses the App Engine default service account
- Best practice: Create custom service accounts for all services

Google Cloud

Example: Assigning Service Accounts to VMs

- Default service account is a project editor which requires scopes to control what the machine can do
- Best practice is to create a custom service account



Google Cloud

Service accounts can be used to control the level of access a virtual machine can have to other Google Cloud services.

The default service account for a virtual machine gives it project editor privileges. However, the default service account can be overridden by using scopes to set the permissions per service. (Note: Access scopes are a legacy method of specifying authorization for your instance. As such, they will not be further discussed. https://cloud.google.com/compute/docs/access/service-accounts#accesscopesiam)

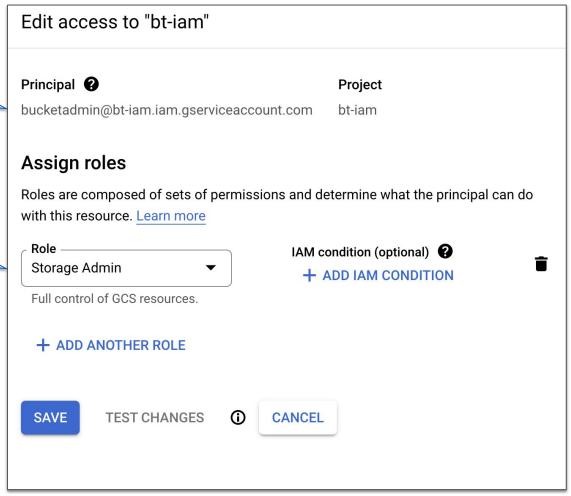
A better alternative is be to create a new service account, grant it the appropriate IAM roles, and select it when creating a virtual machine.

Service accounts are both principals and resources

Is a principal when roles are assigned to it

Service Account

Assigned role



Google Cloud

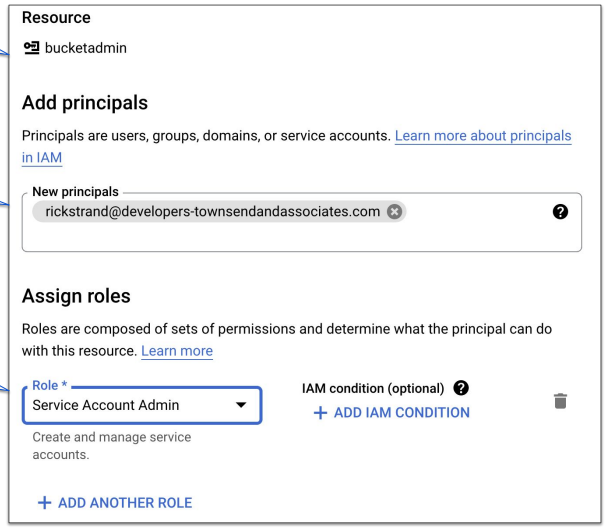
Service accounts are both principals and resources

Is a resource when users (principals) are given Service Account IAM roles to manage the service account in some way

Service Account

Principal

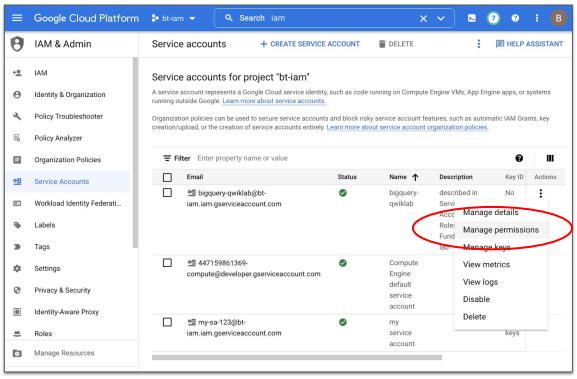
Service Account roles

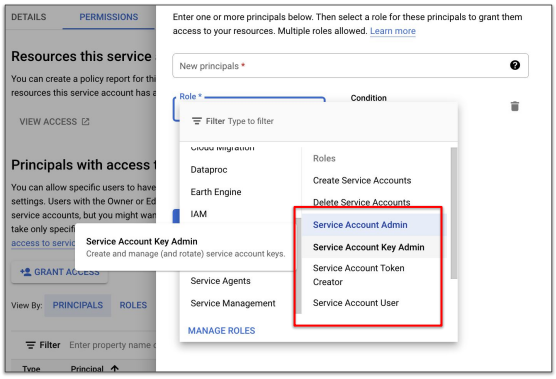


Google Cloud

Managing service account role assignments

Console > IAM & Admin > Service Accounts > Manage Permissions





Google Cloud

Service Account predefined roles (not a complete list)

- Assigned to principles
  - Service Account Admin
    - Create and manage service accounts
  - Service Account User
    - Can attach service account to resources (e.g., Compute Engine)
    - Can “impersonate” the service account and perform the tasks allowed by IAM given to the service account
  - Service Account Key Admin
    - Create and manage (and rotate) service account keys
    - Keys are used by applications external to Google Cloud
  - Service Account Token Creator
    - Short lived credentials represented as OAuth 2.0 access tokens, OpenID Connect ID tokens, self-signed JSON Web Tokens (JWTs), and self-signed binary objects (blobs)

Google Cloud

Accessing Google Cloud resources from an external application

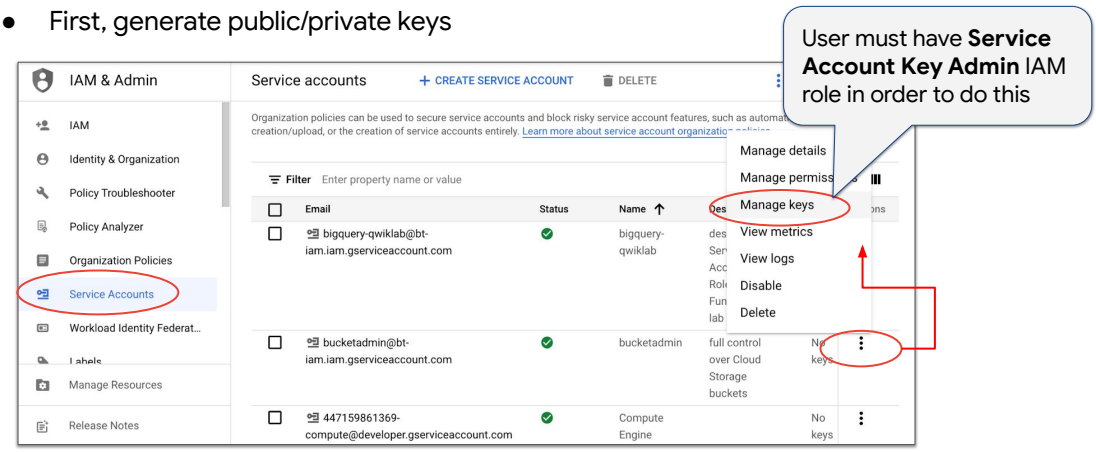
- Any user, group, application, etc. must authenticate prior to accessing Google Cloud services*
  - This also applies to external (e.g., on-premise) applications
    - They use service accounts for this purpose
- To authenticate as a service account, applications must use either use
  - Service Account Keys
  - Service Account Tokens
- Both are discussed next

Google Cloud

*Exception: When authentication is not required, e.g., a public website hosted in Google Cloud

Service Account keys can be used by external applications when authenticating to Google Cloud

- First, generate public/private keys

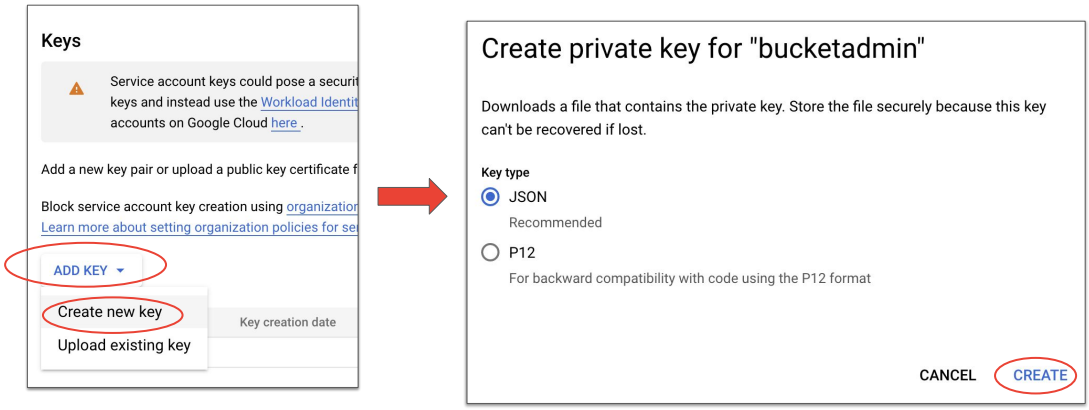


Google Cloud

Create and manage service account keys
https://cloud.google.com/iam/docs/creating-managing-service-account-keys

Service Account keys stored locally must be secured

- Next, download the private key
  - Public key is kept in Google Cloud
  - **Customer is responsible for storing the private key securely**



Google Cloud

Create and manage service account keys
https://cloud.google.com/iam/docs/creating-managing-service-account-keys

Using a Service Account key with code

- To configure gcloud to use a service account:

```
$ gcloud auth activate-service-account \
test-service-account@google.com \
        --key-file=/path/key.json --project=testproject
```

Path where the service account key is stored

- To use credentials in your code (*Python is this example*):
  - The SDK will automatically look for the environment variable value

```
import os
from google.cloud.bigquery.client import Client

os.environ['GOOGLE_APPLICATION_CREDENTIALS'] = 'path_to_json_file'
bq_client = Client()
```

Google Cloud

The top example shows how to configure gcloud to use a service account by linking gcloud to a key that was created.

The bottom example shows how to use your keys via code. In this example, Python.

The SDK will automatically look for the value in the GOOGLE_APPLICATION_CREDENTIALS environment variable.

Using short-lived credentials with a Service Account

- While Service Account keys work for application authentication, they are not the preferred method
  - The private Service Account **key must be secured by the customer** after downloading
    - An unauthorized person with access to the key can authenticate as the Service Account, and do whatever the Service Account's IAM roles allows
  - Service account **keys have an unlimited lifetime**
- The **better alternative is to create short-lived credentials via Service Account tokens**
  - Principals (including Service Accounts) must have **Service Account Token Creator** role in order to create tokens
  - Different types of token are supported
    - OAuth 2.0 access tokens, OpenID Connect ID tokens, self-signed JSON Web Tokens (JWTs), and self-signed binary objects (blobs)

Google Cloud

Best practices for managing service account keys
https://cloud.google.com/iam/docs/best-practices-for-managing-service-account-key

Create short-lived credentials for a service account
https://cloud.google.com/iam/docs/create-short-lived-credentials-direct

Types of tokens

- OAuth 2.0 access tokens
  - Used by an application to authenticate to Google APIs
- OpenID Connect (OIDC) tokens
  - Used when
    - Accessing a Cloud Run service
    - Invoking a Cloud Function
    - Authenticating a user to an application secured by Identity-Aware Proxy (IAP)
    - Making a request to an API deployed with API Gateway or Cloud Endpoints
- Self-signed JSON Web Tokens (JWTs)
  - Used to authenticate communication between services in a microservices architecture
  - Used by an application to authenticate to Google APIs (e.g. FireStore)
- Self-signed binary objects (blobs)
  - Used to securely identify the issuer of a request to a Google Cloud Storage bucket

Google Cloud

Typical token lifespan is 1 hour, but can vary by type

This is a Developer topic, not an Architect

Shown for completeness

Further discussion of creating tokens is outside of the scope of the Architect exam

Token types
https://cloud.google.com/docs/authentication/token-types

Using OAuth 2.0 to Access Google APIs
https://developers.google.com/identity/protocols/oauth2

Generate an ID token by impersonating a service account
https://cloud.google.com/docs/authentication/get-id-token#impersonation

Allowing principals to impersonate service accounts

- Impersonation allows principles to indirectly access all the resources that the service account can access.
- **The roles that allow impersonation are**
  - **Service Account User**
    - This role also allows a principal to attach the service account to a resource, e.g., a VM
  - **Service Account Token Creator**
    - Can impersonate the service account and create tokens
  - **Workload identity User**
    - Can impersonate service accounts from GKE workloads
      - Kubernetes service accounts are not the same as IAM service accounts
      - Create IAM service accounts with roles to access Google Cloud resources
      - Map the Kubernetes service account to the IAM service account

Google Cloud

Managing service account impersonation
https://cloud.google.com/iam/docs/impersonating-service-accounts

Impersonation use case

- Service account impersonation is useful for delegating permissions to perform certain tasks to service accounts, rather than granting them to individual users
  - For example, a developer needs to troubleshoot a Cloud Run deployment
    - The developer has no IAM roles for Cloud Run
    - A service account exists with the Cloud Run Admin role
    - Allow the developer to “act-as” the service account by granting the **Service Account Token Creator** role on that service account
      - Add a condition to limit to the grant to a certain time period if desired
    - The developer can impersonate the service account for troubleshooting
    - When done, the developer reverts back to their normal IAM roles

Google Cloud

Managing service account impersonation
https://cloud.google.com/iam/docs/impersonating-service-accounts

Example: Using Service Account Token Creator role

- Create a Service Account with Cloud Storage Admin role

```
gcloud iam service-accounts create storage-admin \
    --display-name="Storge Admin"
gcloud projects add-iam-policy-binding bt-iam \
    --member=serviceAccount:storage-admin@bt-iam.iam.gserviceaccount.com \
    --role=roles/storage.admin
```

- Give the Service Account Token Creator to a user

```
gcloud beta iam service-accounts add-iam-policy-binding \
   storage-admin@bt-iam.iam.gserviceaccount.com \
   --member=user:caliagusta@developers-townsendandassociates.com \
   --role=roles/iam.serviceAccountTokenCreator
```

- Sign into the command line as the user. Can they create a bucket?
  - No, unless IAM exists elsewhere which allows them to do so

```
gcloud storage buckets create gs://cali-bucket2
```

This should fail

Google Cloud

Example: Using Service Account Token Creator role - continued

- Impersonate the Service Account and create a bucket

```
gcloud storage buckets create gs://cali-bucket4 \
    --impersonate-service-account=storage-admin@bt-iam.iam.gserviceaccount.com
```

No service account key needed

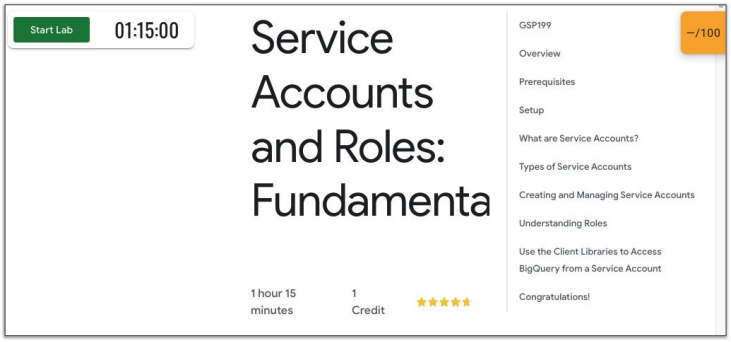
User must have the **Service Account Token Creator** role

Google Cloud

Suggested lab: Service Accounts and Roles: Fundamentals

In this lab you will

- Create a service account
  - IAM - BigQuery Data Viewer & BigQueryUser
- Attach it to a VM
- Create and run Python app to run a SQL query in BigQuery



https://partner.cloudskillsboost.google/catalog_lab/956

Google Cloud

VPC Firewall rules (revisiting the topic)...

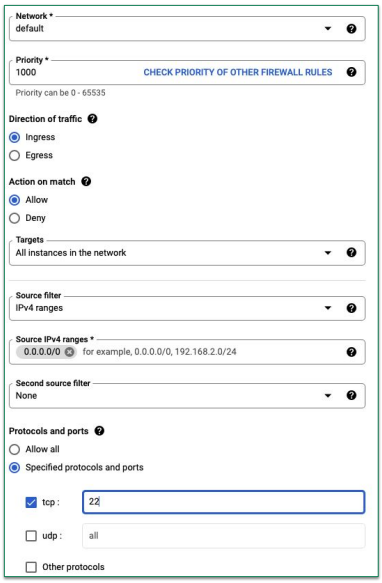
Google Cloud

Creating Firewall Rules

Saw this slide earlier

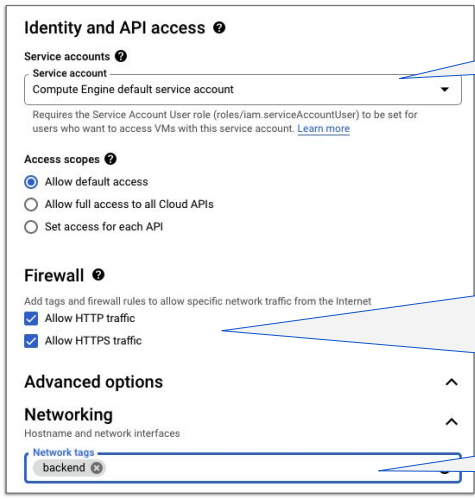
- When creating rules, specify
  - Source
    - Could be the internet (0.0.0.0/0 IP range)
    - Individual or ranges of IPv4 or IPv6 addresses
    - Could be VMs with specific network tags or service accounts
  - Target - Defines which VMs the rule applies to
    - All instances in the network
    - VMs with specific network tags
    - VM's with service accounts

Will revisit the last 2 after service account discussion



Google Cloud

Console - Creating a VM with network tags and a service account



Set the Service Account here

Developer needs a specific IAM role to attach a service account to a resource

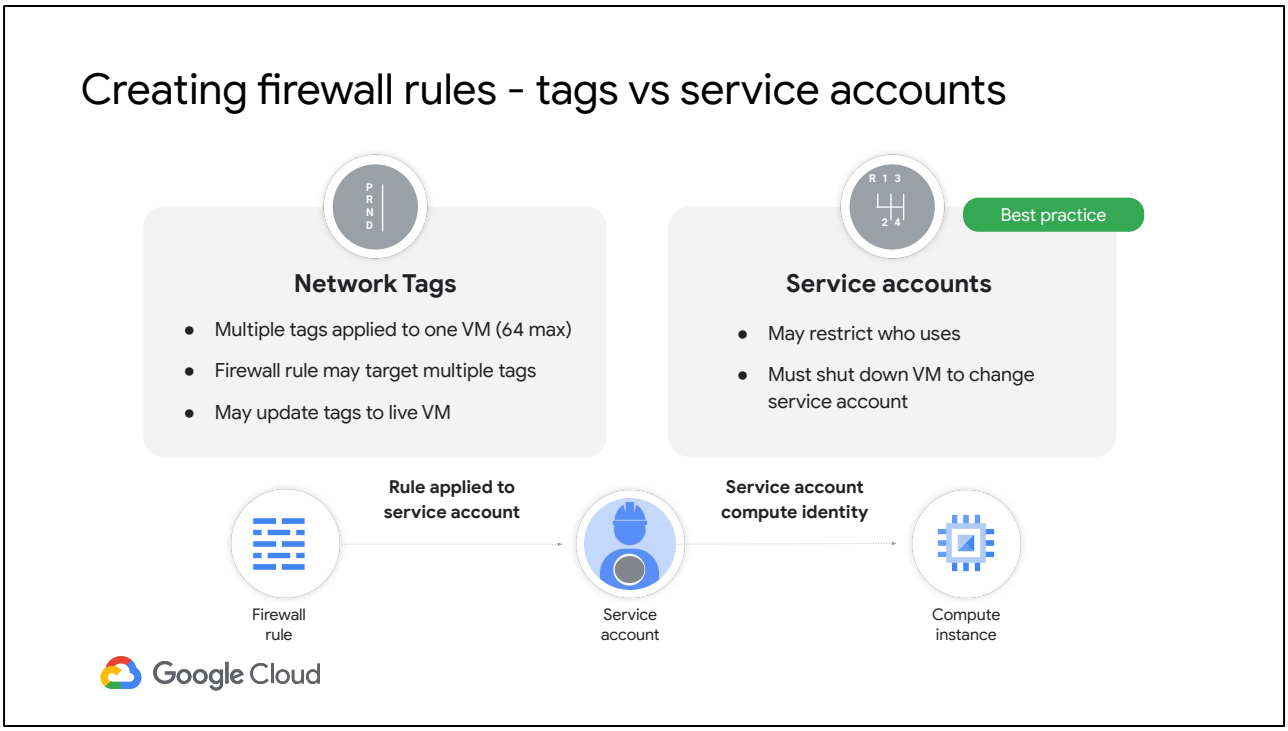
Clicking one (or both) of these boxes results in automatically generated firewall rule(s) that apply to VMs with the network tags:

- allow-http
- allow-https

These tags are automatically added to the VM upon creation.

Can manually add network tags. One VM can have multiple tags

Google Cloud



- While some firewall rules may apply to all VM instances, there will be others that you want to apply selected instances.
- For example, you may have
  - A front end Web application running in a Managed Instance Group, that should only be accessible via an external Load balancer proxy
  - And a Backend app that should only be accessible on specific ports by the Web application instances
- This can be done based on either network tags or service accounts. Which one should you use?

Network Tags vs Service Accounts

Network Tags

Are very flexible

- Tags can be updated while a VM is running, while a service account cannot
- A VM can have multiple tags, while a VM can have one service account only
- FW rules match if >= 1 tag matchesTags

Potential con:

- Not possible to control who can apply a specific network tag.
- Example: A HTTP ingress allow rule exists that applies to instances with 'webapp-http-allow' network tag
- Anyone with Compute Engine Admin IAM role can create VM's with this network tag.

```
gcloud compute firewall-rules create web-logdata \
    --network logging-network \
    --allow TCP:443 \
    --source-tags web-production \
    --target-tags log-data
```

Google Cloud

Network Tags vs Service Accounts

Network Tags

Are very flexible

- Tags can be updated while a VM is running, while a service account cannot
- A VM can have multiple tags, while a VM can have one service account only
- FW rules match if >= 1 tag matchesTags

Potential con:

- Not possible to control who can apply a specific network tag.
- Example: A HTTP ingress allow rule exists that applies to instances with 'webapp-http-allow' network tag
- Anyone with Compute Engine Admin IAM role can create VM's with this network tag.

Service accounts

Best practice from a security viewpoint

- Can control who can assign a given service account to instances
- Allows strict control over the instances a given rule will apply to.
- Goes hand in hand with IAM best practice - creating service accounts with minimal privileges for different applications and their components.

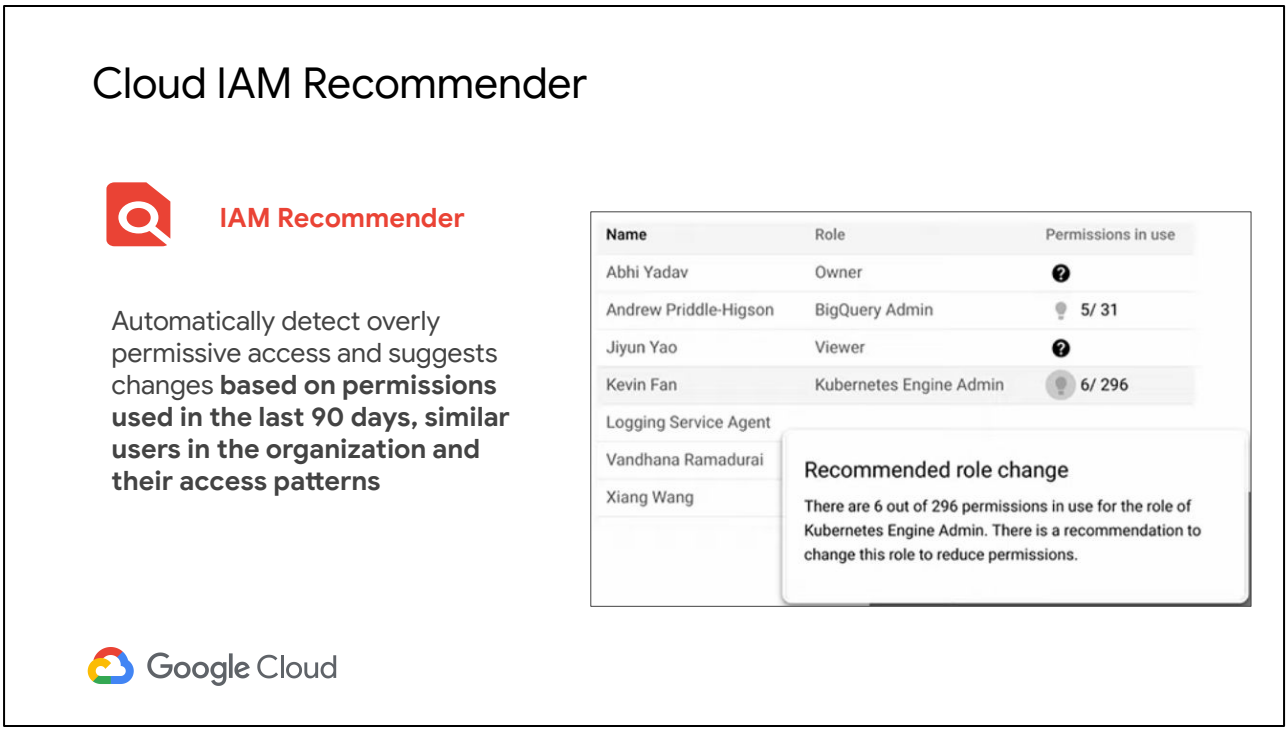
```
gcloud compute firewall-rules create frontend-to-backend \
--direction=INGRESS --network=default --allow TCP:443 \
--source-service-accounts=frontend@proj.iam.gserviceaccount.com \
--target-service-accounts=backend@proj.iam.gserviceaccount.com
```

Google Cloud

IAM Best Practices

- Separation of duties – use projects to control access to resources based on access requirements
  - Developer project allows the programmers access
  - Production project allows the operations people access
- Principle of least privilege – use roles to control what users can do
  - Avoid making everyone a Project Owner just because it's easy
- Use more restrictive rights at higher levels and provide access to lower levels and resources when required

Google Cloud



Recommender
https://cloud.google.com/recommender/docs

Recommender is a service that provides recommendations and insights for using resources on Google Cloud. These recommendations and insights are per-product or per-service, and are generated based on heuristic methods, machine learning, and current resource usage.

An example of a recommendation is one generated by the VM instance rightsizing recommender. In this case, the recommender generates recommendations based on system metrics collected by Cloud Monitoring over the last eight days. If the recommender detects that a VM instance is underutilized, it recommends changing the machine size to save cost.

Cloud IAM Recommender
Cloud IAM Recommender is one of Google Cloud's Recommendation Engines that is powered by ML behind the scenes.
https://cloud.google.com/iam/docs/recommender-overview

The IAM recommender helps you enforce the principle of least privilege by ensuring that members have only the permissions that they actually need

IAM uses Recommender to compare role grants with the permissions that each member used during the past 90 days. If you grant a role to a member, and the

member does not use all of that role's permissions, then the IAM recommender is likely to recommend that you revoke the role. If necessary, the IAM recommender also recommends less permissive roles as a replacement. This suggested replacement could be a new custom role, an existing custom role, or one or more predefined roles. Except in the case of recommendations for Google-managed service accounts, the IAM recommender never suggests a change that increases a member's level of access.

The IAM recommender also uses machine learning to identify permissions in a member's current role that the member is likely to need in the future, even if the member did not use those permissions in the past 90 days.

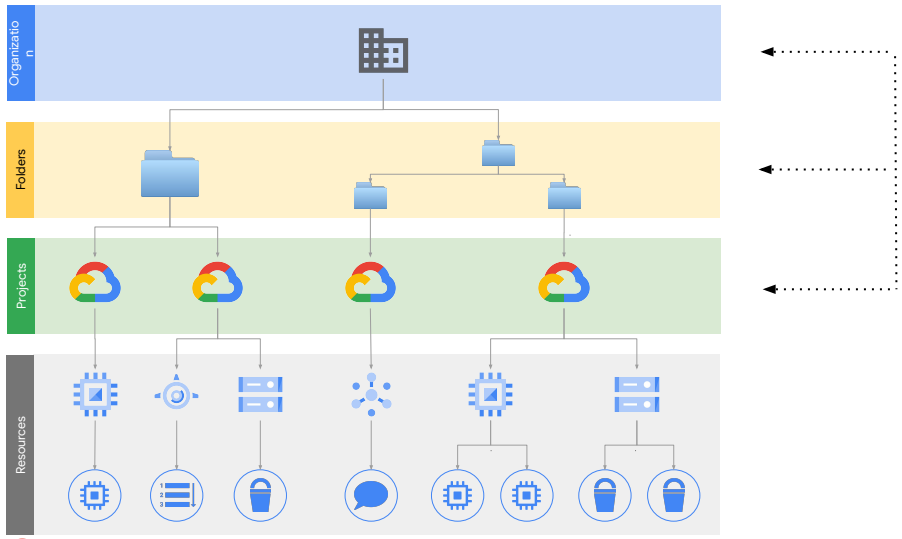
The IAM recommender does not apply recommendations automatically. Instead, you must review each recommendation, then either apply or dismiss the recommendation.

The IAM recommender evaluates only role grants that were made at the project, folder, or organization level, and that have existed for at least 90 days. It does not evaluate any of the following items:

- Role grants made below the project level; that is, role grants on service-specific resources within a project
- Conditional role grants
- Role grants for Google-managed service accounts for roles other than Owner, Editor, and Viewer
- Access controls that are separate from IAM

Google Cloud Blog Posts
Achieve least privilege with less effort using IAM Recommender (Aug 4, 2020)

Resource Hierarchy

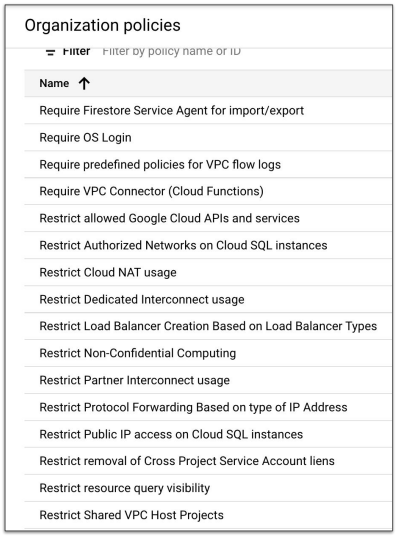


Google Cloud

Next discussion: Organization Policies

Organizational policies

- An organization policy is a **restriction** or constraint that you can set over the use of a service
- For example
  - Restrict what type of VMs developers can create
  - Restrict what Google regions resources can be created in
  - Restrict the use of public IPs to some VMs (or none)
- Are set at the organization level, folder level or project level
- Define and establish guardrails for your development teams to stay within compliance boundaries.



Google Cloud

Introduction to the Organization Policy Service

Organization policy constraints

- Configured with constraints
  - Particular type of restriction against either a Google Cloud service or a group of Google Cloud services
- Descendants of the targeted resource hierarchy node inherit the organization policy.

```
resource: "organizations/[ORG_ID]"
policy: {
  constraint: “constraints/compute.skipDefaultNetworkCreation”
  booleanPolicy: {
    enforced: true
  }
}
```

Google Cloud

In order to define an organization policy, you choose a constraint, which is a particular type of restriction against either a Google Cloud service or a group of Google Cloud services. You configure that constraint with your desired restrictions.

Descendants of the targeted resource hierarchy node inherit the organization policy. By applying an organization policy to the root organization node, you are able to effectively drive enforcement of that organization policy and configuration of restrictions across your organization.

The example on the slide shows how an organization policy that’s used to disable the creation of the default network

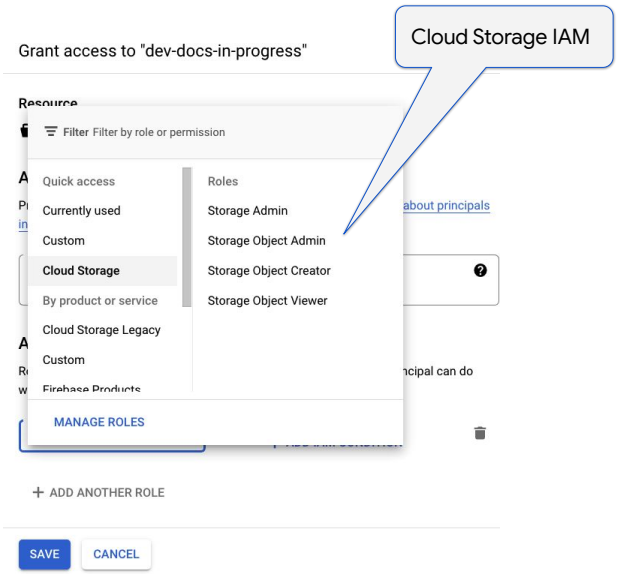
Cloud Storage security...

Google Cloud

Bucket Security

Permissions can be added to control access to buckets and objects

- Use IAM users and groups



Google Cloud

As with everything in Google Cloud, Cloud Storage security is managed using IAM roles and members.

You can add IAM users and groups to buckets or objects and assign roles to those members. The roles determine what permissions they have. Note that there are a couple built-in groups for assigning public access to Storage—these are allUsers and allAuthenticatedUsers.

Predefined IAM Storage Roles

- Roles can be added to a principle or service account at the organization, folder, project or bucket level
  - Bucket level - applies to that specific bucket only

Avoid using the legacy roles

Storage Admin	Storage	Enabled	
Storage Legacy Bucket Owner	Storage Legacy	Enabled	
Storage Legacy Bucket Reader	Storage Legacy	Enabled	
Storage Legacy Bucket Writer	Storage Legacy	Enabled	
Storage Legacy Object Owner	Storage Legacy	Enabled	
Storage Legacy Object Reader	Storage Legacy	Enabled	
Storage Object Admin	Storage	Enabled	
Storage Object Creator	Storage	Enabled	
Storage Object Viewer	Storage	Enabled	

Google Cloud

IAM permissions can be granted at the organization, folder, project or bucket level

Storage Role Permissions

Storage Object Admin	Storage Object Creator	Storage Object Viewer
Description Full control of storage objects. 9 assigned permissions: • resourcemanager.projects.get • resourcemanager.projects.list • storage.objects.create • storage.objects.delete • storage.objects.get • storage.objects.getIamPolicy • storage.objects.list • storage.objects.setIamPolicy • storage.objects.update	Description Access to create objects in storage. 3 assigned permissions: • resourcemanager.projects.get • resourcemanager.projects.list • storage.objects.create	Description Read access to storage objects. 4 assigned permissions: • resourcemanager.projects.get • resourcemanager.projects.list • storage.objects.get • storage.objects.list

Google Cloud

The Storage Object Viewer role grants read-only permissions to Storage. Storage Object Creator grants write permissions and Storage Object Admin grants full control over Cloud Storage Buckets and Objects.

Storage Object ACLs

Access Control Lists (ACLs) can be used to grant access to objects in buckets

Not best practice

ENTITY	NAME	ACCESS
Project	owners-411554854281	Owner
Project	editors-411554854281	Owner
Project	viewers-411554854281	Reader
User	storage-transfer-11367529508056	Owner
User	allUsers	Reader

+ Add item

Google Cloud

You can also grant read or read/write access to individual objects within a bucket using Access Control Lists. Google recommends to use IAM at the bucket level instead, for each of administration and security

Signed URLs

- Provide temporary access to buckets
  - Create a service account with rights to storage
  - Create a service account key
  - Use signurl command to create a URL that allows access to the resource
    - -d parameter is used to specify duration

```
gcloud iam service-accounts keys create ~/key.json --iam-account
storage-admin-sa@doug-demo-project.iam.gserviceaccount.com

gsutil signurl -d 10m ~/key.json gs://super-secure-bucket/noir.jpg
```

Google Cloud

Sometimes, you want to programatically grant temporary access to an object in a bucket. You can do this with a signed URL. Use the gsutil signurl command as shown on the screen. The -d parameter determines how long the signed URL works.

Signed URL Example Output

```
me@doug-demo-project:~$ gsutil signurl -d 10m ~/key.json gs://super-secure-bucket/noir.jpg
URL     HTTP Method     Expiration      Signed URL
gs://super-secure-bucket/noir.jpg       GET     2018-08-31 16:29:25     https://storage.google
apis.com/super-secure-bucket/noir.jpg?x-goog-signature=107d26e38f5c962296c26f4153a1cbeb61a84ac
a905009752e849f8f890de1f9a80e482da3bae562c7796389e12a8657a70c87860700149c4b2218c81ad3d57730cd3
5ced850b266cdffd84de01898ee8c807d742a85136e56f46d83c29ceb792bdd3a22adbe2e540ba27b0f565bbf8f31a
ee6ae61d6ae20968021d5a47c8d0aada43f2d32407f2977a4c7b4c66ef64ddd68bd6f6135936f847ace3530a968d72
63ff5e70f9fc39bf16fabbd472f63584a8d8c6b24b1f81859f1c5176b8e97580a6b4a7613ad76bfcdd403e6afc9a70
90a3e1b4cf95c7fb68142416af86ef5ef6bfab93c00492b307233180df9b3dfeefe1bb9a5bf81cb441f879ecc2e57c
def&x-goog-algorithm=GOOG4-RSA-SHA256&x-goog-credential=storage-admin-sa%40doug-demo-project.i
am.gserviceaccount.com%2F20180831%2Fus%2Fstorage%2Fgoog4_request&x-goog-date=20180831T201925Z&
x-goog-expires=600&x-goog-signedheaders=host
me@doug-demo-project:~$
```

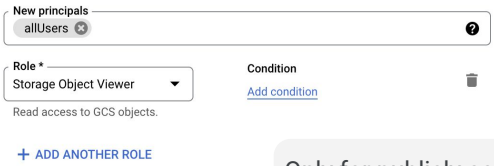
Google Cloud

A long URL will be generated. This URL provides access to the file in storage for the duration you specified in the command.

This is similar to what you might do when creating a shared link in Google Docs.

Making buckets public

To make a **bucket** public, grant `allUsers` the Storage Object Viewer role.



To make an **object** public, grant `allUsers` Reader access

Only for publicly accessible web content:
Use with caution!

Note: There is also an `allAuthenticatedUsers` role. This represents the principals who have identities within your Google Cloud domain

Google Cloud

In Cloud Storage it is possible to make entire buckets or individual objects public. An entire bucket can be made public by granting the **Storage Object Viewer** role to the **allUsers** group on the bucket.

To make individual objects in a bucket public, grant the **Reader** access to **allUsers** on the individual objects.

Be very careful on which objects or buckets you make public - use this option with **extreme caution**. Making buckets or objects public should only be done for publicly accessible web content.

Google Cloud