



Mobile and Cloud Computing #1

Cloud Storage

Which storage type?

In Memory	Relational	NoSQL	Warehouse	Object	Block	File		
								
Cloud Memorystore	Cloud SQL	Cloud Spanner	Cloud Datastore, Firestore	Cloud Bigtable	BigQuery	Cloud Storage	Persistent Disk	Cloud Filestore
Good for: In memory datastore	Good for: Relational database service	Good for: Scalable relational database	Good for: Serverless NoSQL document	Good for: NoSQL key-value and wide-column	Good for: Enterprise DW	Good for: Unstructured data, objects or blobs	Good for: Local VM file storage	Good for: Lift/shift apps requiring file

Object

Key terms

Object

A piece of data. Identified by an **object name**.

Bucket

A container for **objects**. Identified by a **bucket name**.

Location

A region or multi-region where a **bucket** is located and its **objects** are stored.

Storage Class

Allows **objects** to be stored and served with different access frequencies, availability levels, and pricing profiles.

What is Cloud Storage?



**Highly available and simple
to use object storage**

Simple: a consistent API, latency, and speed across storage classes simplifies development and experience.

Reliable: Storage in geo-redundant locations offers 99.95% availability with transparent failover.

Cost-effective: Cold storage as low as \$0.004 per gb/month. Still accessible in ms and with geo-redundant options.

Secure: Data is encrypted at rest and in transit. WORM compliant options and integration with KMS and IAM.

Key features



Unified experience

Strongly consistent listing

Geo-redundant

Scalable to exabytes

Options for any use case

	Multi-regional	Regional	Nearline	Coldline
Common usage	For highest availability of frequently accessed data	For data accessed frequently within a region	For data access less than once a month	For data accessed roughly less than once a year
Customer use case examples	 Streaming videos  Images  Websites  Documents	 Video transcoding  Genomics  General data analytics and compute	 Serving rarely accessed docs  Backup	 Serve rarely used data  Movie archive  Disaster recovery

Common use cases

Content storage and delivery



Cloud Storage is ideal for low-latency, high-QPS content serving to users distributed across geographic regions. It's easy to stream audio or video directly to apps or websites.



Compute, analytics, and ML



Co-locate your compute and storage for higher throughput and low latency. Cloud Storage is strongly consistent, making it simpler to run analytics workloads.

Backup and archiving

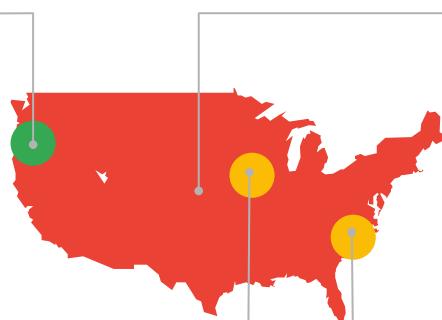


Low cost archive storage with ms latency to access content, geo-redundant options and a consistent API. Work with leading backup and DR partners to leverage Cloud Storage.

Choosing a location type

Regional

Your data is stored in a specific region with replication across availability zones in that region.



Multi-regional

Your data is distributed redundantly across US, EU or Asia.

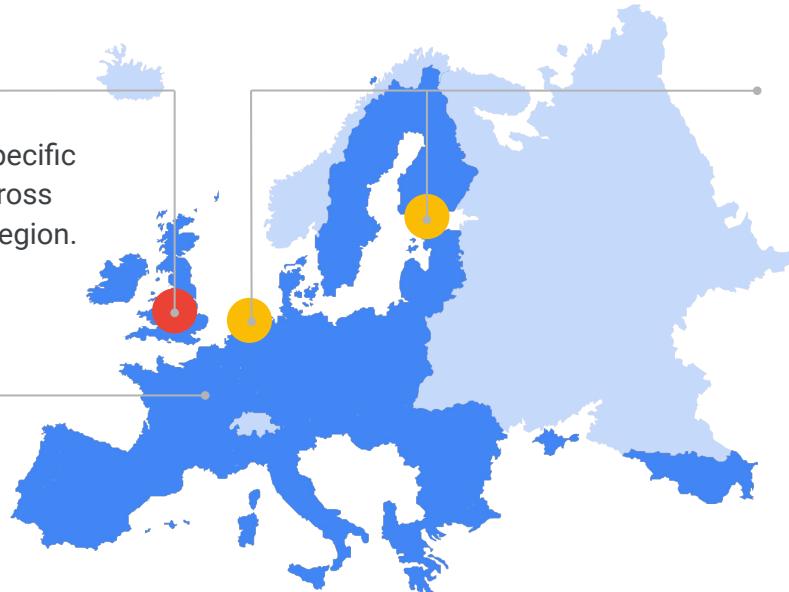
Dual-regional

Your data is replicated across a specific pair of regions.

Choosing a location type

Regional

Your data is stored in a specific region with replication across availability zones in that region.



Dual-regional

Your data is replicated across a specific pair of regions.

Multi-regional

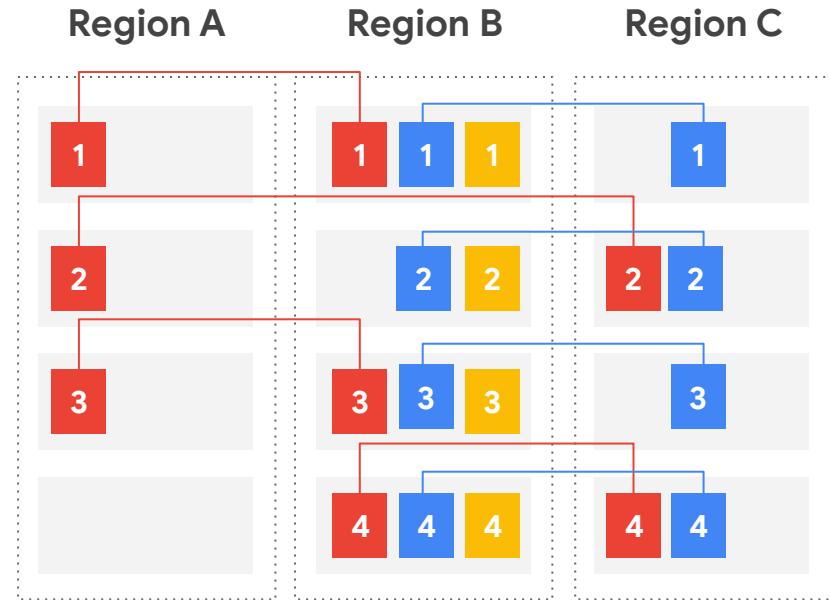
Your data is distributed redundantly across US, EU or Asia.

Data placement

Multi-region

Dual-region

Regional



Any regional failure is transparent to customers and data is automatically served from other region. No additional network fees for Multi-region or Dual region replication.

Multi-regional: Best for content serving

- **Geo-Redundant** across two or more regions
- **Simple pricing**, simple API, and simple business continuity
- **Highest level of availability**
- Optimally routes requests within Google Network
- Available in three Multi-regional Locations
- Most features are applicable to all other classes
- **Common use cases:**
 - Streaming videos and music
 - Serving images and website content



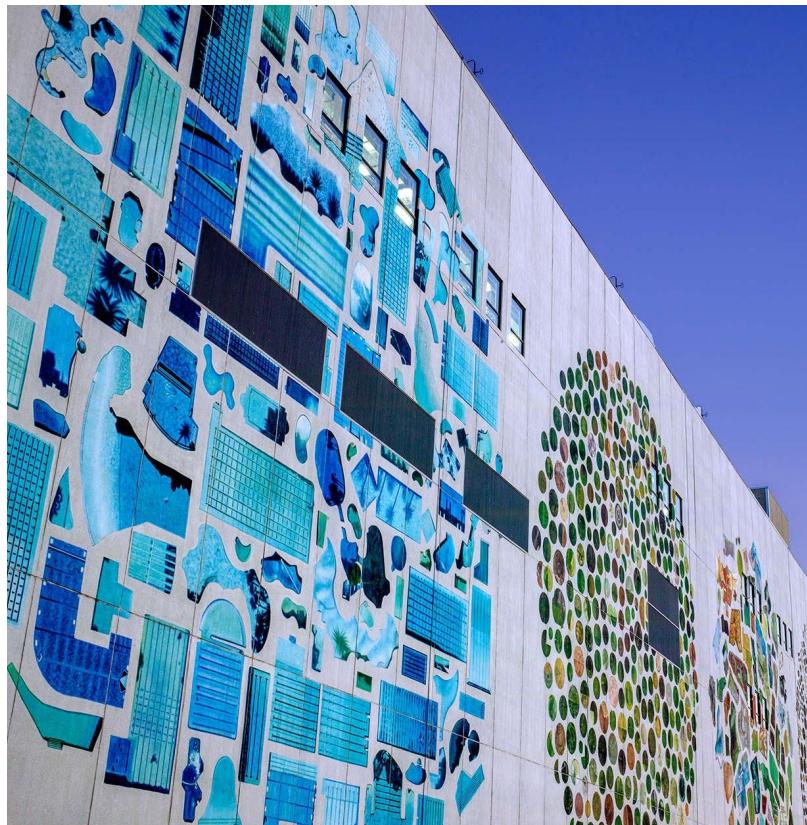
Multi-regional: Best for content serving

- **Simple API**
`gsutil mb -c multiRegional -l us gs://my_us_bucket
GET my_us_bucket.storage.googleapis.com/image.jpg`
- **Optimized for end-user latency**
 - Serves data from the closest available location to the end-user
 - Caches data in regions closest to the end-users
 - Has >100 points of presence for peering



Regional: Best for data analytics

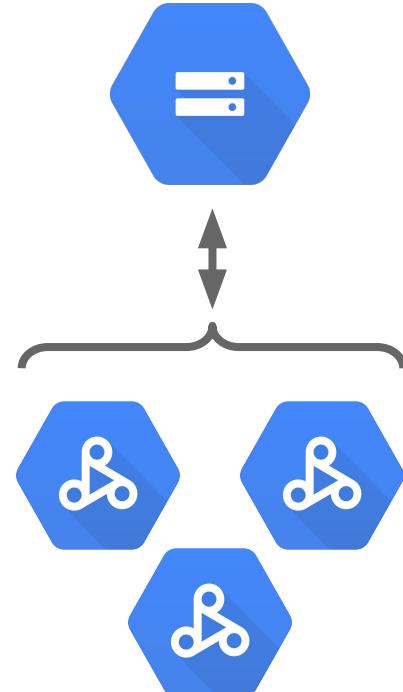
- Redundant within a region
- For data accessed frequently
- Ideal for:
 - Running analytics and compute within the same region
 - Content serving in geographic locations nearby, or origin shielding
- Available in every region, and will be in all new regions
- Common use cases:
 - Genomics analysis, data transcoding
 - Analytics for ecommerce and IoT
 - Integrated with Cloud DataProc, Cloud Machine Learning, BigQuery



Example

Using Hadoop with Cloud DataProc

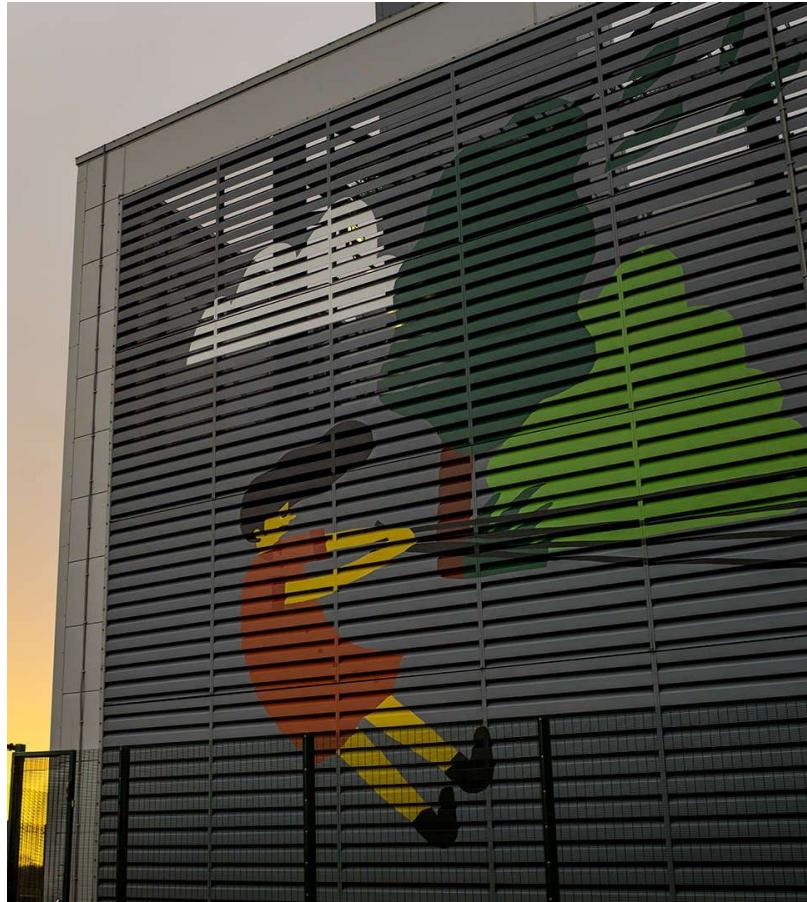
- Cloud DataProc is a managed Spark and Hadoop service
- Cloud Storage connector is HCFS compliant
- Solves many headaches with:
 - High availability, easy to scale, no management overhead
 - Persists outside of clusters, and available for other uses
- Easy code migration:
 - `textFile = sc.textFile("hdfs://...")`
 - `textFile = sc.textFile("gs://...")`
- cloud.google.com/dataproc/



Nearline:

Best for infrequently accessed data

- For data accessed infrequently
- Instant access speed.
- Same API as other storage classes
- *Common use cases:*
 - Serving long-tail content
 - Historic usage data for rare analysis
 - Back-up



Coldline: Best for cold data

- Cold storage you can actually use
- Instant access speed
- Same API as other storage classes
- *Common use cases:*
 - Archive
 - Multimedia source file storage
 - Disaster recovery



Coldline: Consistent API and performance

- The same API as for other storage
- GET/PUT coldline.storage.googleapis.com/image.jpg
- The same performance as other storage classes
- You can actually serve your data from Coldline when you need it
- gsutil perfdiag:



```
,  
  "DOWNLOAD_1048576": [  
    0.19501709938049316,  
    0.15822696685791016,  
    0.14728212356567383,  
    0.16345596313476562,  
    0.18323922157287598  
  ],
```

Tiering data maintains object redundancy



Predictable storage costs

Storage Growth Plan is a new way to commit to Cloud Storage that protects you from the cost volatility associated with your data storage behavior.

Fixed price storage with flexible growth.

Cloud Storage classes today

	Multi-region	Regional	Nearline	Coldline
SLA	99.95%	99.9%	99%	99%
Storage Price*	\$0.026-\$0.036 per GB/month	\$0.02 per GB/month	\$0.01 per GB/month	\$0.004 per GB/month
Min. Storage Duration	0 days	0 days	30 days	90 days
Access Price	Free	Free	\$0.01 per GB	\$0.05 per GB
Use For ...	Highest throughput, lowest latency; frequently accessed data	High throughput, low latency; frequently accessed data	Lower cost; infrequently accessed data (< once per month)	Low cost; very infrequently accessed data (< once per year)

Cloud Storage classes

Storage class	Location type					
	region	multi-region		dual-region		
Standard	99.9% availability	2 ¢	99.95% availability	2.6 ¢	99.95% availability	3.6 ¢
Nearline	99%	1 ¢	99.9%	1 ¢	99.9%	2 ¢
Coldline	99%	.4 ¢	99.9%	.7 ¢	99.9%	.9 ¢
Archive	99%	.12 ¢	99.9%	.4 ¢	99.9%	.5 ¢

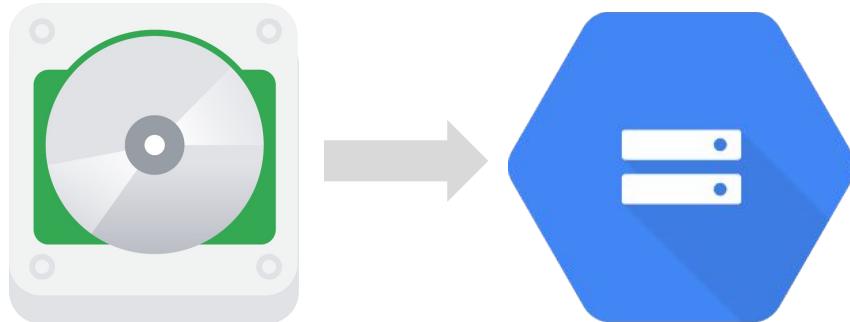


* For lowest-cost locations. Some locations are priced higher.

New Archive class of Cloud Storage

Replace your on-premises tape archives with the new Archive class of Cloud Storage. Coming soon.

Starting at **\$0.0012 per GB/per month** in regional locations.



Secure storage

100% of data is encrypted at rest and in transit

Always on
Just works



Default
Google manages keys
transparently

Always-on encryption for everything.
No choices here :-)

Customer-managed (CMEK)
You manage keys in
Google Cloud KMS

Cloud KMS lets you control key
creation, revocation & rotation.

Customer-supplied (CSEK)
You store keys outside
Google Cloud

Keep keys on your own premises, and
only supply them to Google when
doing an operation.

If you lose the keys, your data is
effectively cryptodeleted.

Digitizing “the morgue” archive

The digitized archive will inspire stories for [Past Tense](#), a body of coverage revisiting history through photos. Journalism by The New York Times, powered by technology from Google Cloud.

The image shows a laptop screen displaying a New York Times multimedia photo gallery. The header includes the menu icon, search bar, "MULTIMEDIA/PHOTOS" link, "The New York Times" logo, "GIFT THE TIMES" button, and "Account" dropdown. Below the header, the title "Past Tense" is displayed, followed by a subtitle "California: State of Change". A brief description states: "As we digitize some six million photo prints in our files — dating back more than 100 years — we are using those images to bring vivid narratives and compelling characters of the past to life." The main image in the gallery shows a group of people swimming in a large body of water. Below the main image are three smaller thumbnail images: one showing a street scene with flags, another showing a person in profile, and a third showing a group of people with the text "1981-1983 NEW YORK".

Cloud Storage: Qwik Start - Cloud Console

30 minutes

1 Credit



GSP073



Google Cloud Self-Paced Labs

Cloud Storage: Qwik Start - CLI/SDK

30 minutes

1 Credit

★★★★★ Rate Lab

GSP074



Google Cloud Self-Paced Labs

Block



Persistent Disk

Persistent Disk is **durable and high performance block storage** for the Google Cloud Platform.

Persistent Disk provides SSD and HDD storage which can be attached to instances running in either **Google Compute Engine or Google Kubernetes Engine**.

Storage volumes can be **transparently resized** (up to 64TB), quickly backed up, and offer the ability to support simultaneous readers.

Single disk for max VM performance - **no striping needed**.



Persistent Disk

	Local SSD	Standard Persistent Disk	SSD Persistent Disk
Target scenarios	In-memory databases High-performance scratch space	Large data processing workloads and some enterprise applications Genomics processing, video transcoding in GCE	High performance database and enterprise applications MySQL, SQL Server, Oracle
	Ephemeral storage	Persistent storage	Persistent storage
Features	Highest-performance (\$0.218 GB)	Cost sensitive (\$.04 GB)	Performance sensitive (\$0.17GB)
	Encryption 3TB	Encryption, Snapshots 64 TB, Disk Size sets performance (Attach larger VMS for max SSD performance)	

Performance expectations

	Standard Persistent Disk	SSD Persistent Disk
IOPs	0.75 read, 1.5 write IOPS / GB Max: 3k read, 15k write	30 IOPS / GB Max: 40k read, 30k write
Bandwidth	180 read MB/s, 120 MB/s write	32+ core VMs: 800 MB/s read, 400 MB/s write 16+ core VMs: 480 MB/s read, 240 MB/s write 01+ core VMs: 240 MB/s read, 240 MB/s write
Features	IOPs and B/w independently throttled IOPs measured at 4k block sizes Single disk delivers max performance	

File



3

Cloud Filestore



Our fully managed file storage service for applications that require a file system interface and a shared file system for data. Offering the **best performance predictability and price for performance.**

	Standard	Premium
Max throughput	180 MB/s at ≥ 1.5 TB	700 MB/s at ≥ 1.6 TB
Max IOPS	5,000 at ≥ 10 TB	30,000 at ≥ 1.3 TB
Everything else	Same	Same
Price	20¢ / GB / Month	30¢ / GB / Month

- Price and Performance is the only difference between the two tiers
- Within a tier, performance scales with provisioned capacity (similar to PD)
- Result: customers can optimize spend for their target workloads

Filestore use cases

Rendering	Application migrations	Web content management	Media processing	Home directories
Mount Filestore volumes on GCE instances, enabling visual effects artists to collaborate on the same file share. Burst compute to meet rendering demands.	Filestore can support a broad range of enterprise applications that need a shared file system interface to data. (ex. SAP)	Web developers creating websites and blogs that serve file content to their audience will find it easy to integrate Filestore with web software like Wordpress.	Graphic design, video and image editing and other media workflows use files as an input and files as the output. Filestore helps creators access shared storage to manipulate and produce large files.	Users across your organization probably need to access and share common data sets. You can host file content in Cloud Filestore and enable shared access to that data.

Cloud Filestore was easy to provision and mount, and reliable for the kind of workload we have. Having a POSIX file system that we can mount and use directly helps us speed-read our files, especially on new machines.

Charlie Rice,
Chief Technology Officer, ever.ai



“Wordpress hosts 30% of the world’s websites, so delivering a highly available and high performance Wordpress solution for our clients is critical to our business. Cloud Filestore enabled us to simply and natively integrate Wordpress on Kubernetes Engine.”

Ashley Maloney,
Lead DevOps Engineer at Jellyfish Online
Marketing



It took **one premium Cloud Filestore** instance to support a workload that previously required 16 powerful servers.

Abhishek Arya,
Information Security Engineer,
Google Chrome

Cloud Filestore: Qwik Start

45 minutes

1 Credit



GSP244



Google Cloud Self-Paced Labs

Transfer



You can do amazing things with your data



Optimize
infrastructure
costs and
achieve scale



Gain value
from data to
predict business
outcomes



Build new
apps and
experiences



Connect to
business
platforms of
services and
partners



Make teams
productive
with secure
mobile/devices

But amazing will only happen if you get your data here

0101



So much data



Too little bandwidth

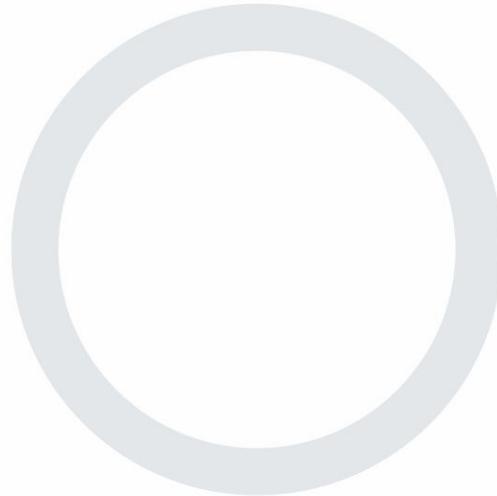


Checksumming,
encryption and
firewalls

Not enough time or
resources

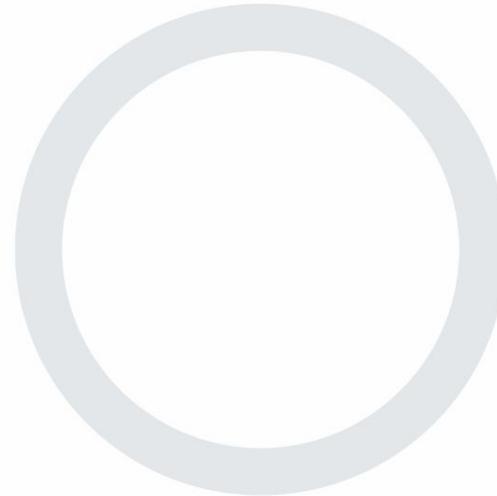
	1 Mbps	10 Mbps	100 Mbps	1 Gbps	10 Gbps	100 Gbps
1GB	3 hrs	18 mins	2 mins	11 secs	1 sec	.1 secs
10 GB	30 hrs	3 hrs	18 mins	2 mins	11 secs	1 sec
100 GB	12 days	30 hrs	3 hrs	18 mins	2 mins	11 secs
1TB	124 days	12 days	30 hrs	3 hrs	18 mins	2 mins
10 TB	3 years	124 days	12 days	30 hrs	3 hrs	18 mins
Typical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

Typical Network
100Mbps



0 Days

Transfer Appliance



0 Days

Transfer options to meet your needs

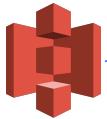


On-prem;
good network

Online transfers



Region
to region



Cloud storage transfer service

Partners



On-prem;
bad network;
tons of data

Transfer Appliance Beta

Partners



DBs;
data warehouse;
market data;
SaaS data

BigQuery transfer service

Partners

Google Cloud



Google Cloud



Transfer Appliance

Google's rackable,
high-capacity storage server
for shipping large amounts of
data faster than a network.
This is the 100TB version.



Google Cloud

It also comes super-sized

With 480 TB of raw capacity



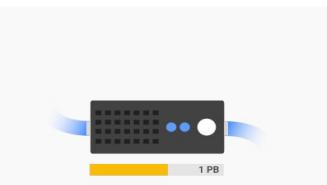
Transfer Appliance



STEP 1

Request a Transfer Appliance

Submit request online. We will work with you on an appropriate size and number.



STEP 2

Fill Appliance with up to a PB of Data

Rack appliance in your datacenter and connect to network. Capture data. It's encrypted.



STEP 3

Ship the Appliance Back to Google

Use the provided return label and ship the appliance back to Google.



STEP 4

Google Uploads Data to the Cloud

Google will notify you when your data is ready.

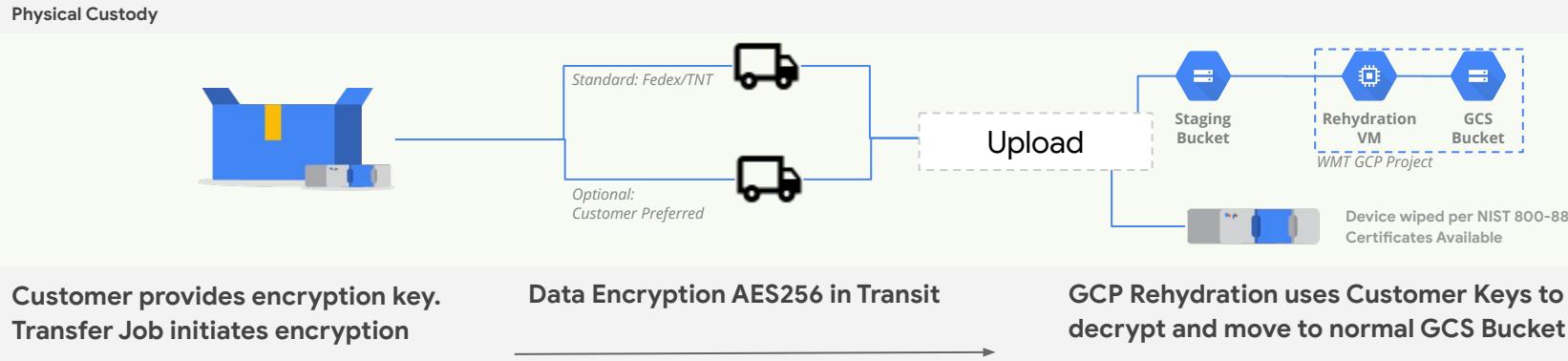


STEP 5

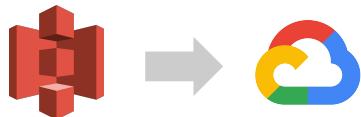
Access and Place Your Data

Log into your console to decrypt your data and choose a storage bucket.

Transfer Appliance



Storage Transfer Service



What is it?

- 3-step process to quickly transfer data from cloud to cloud

Good for:

- S3 to GCS
- GCS to GCS (Multi-regional Storage bucket to Regional Storage Bucket)
- Popular use cases
 - Backup data in S3 to GCS
 - Move data from S3 → GCS → BigQuery

Tips and tricks:

- Schedule one-time transfers
- Schedule periodic syncs (cron like) from data source to data sink with filters
- Schedule hourly syncs by creating 24 transfer jobs starting an hour apart from each other

Cloud Storage for Firebase

Store and serve apps content

Store and serve user-generated content like images, audio, and video directly in your mobile app, using the Firebase SDKs.

Resumable uploads and downloads are resilient to changes in network connectivity.

All data is accessible from the Cloud Storage APIs.



Cloud Storage for Firebase

Optimized for mobile content delivery



Granular security



Robust networking

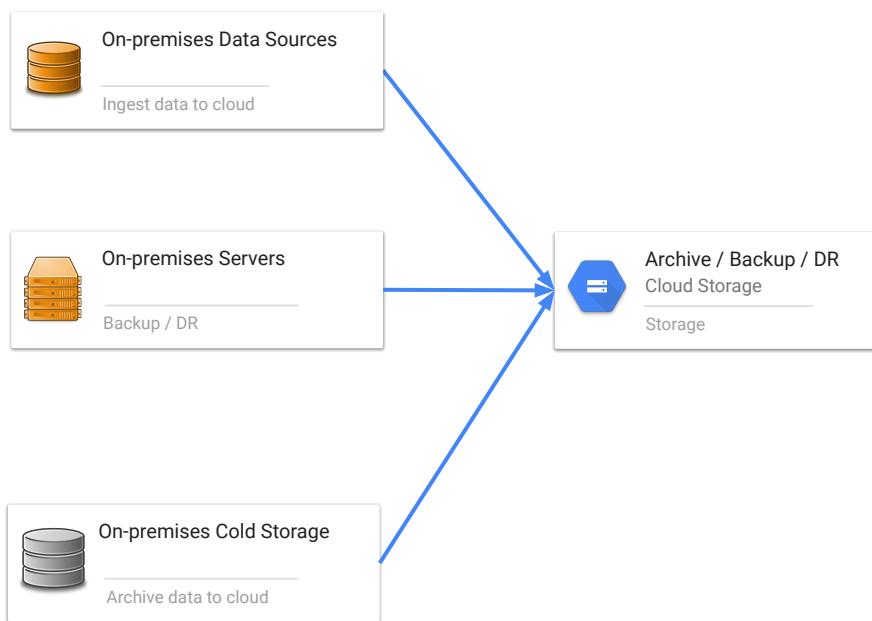


Google scale

More resources

Architectures

Cloud Storage: Data Protection (Archive, Backup/DR, Ingest)



Ingest



Backup & DR

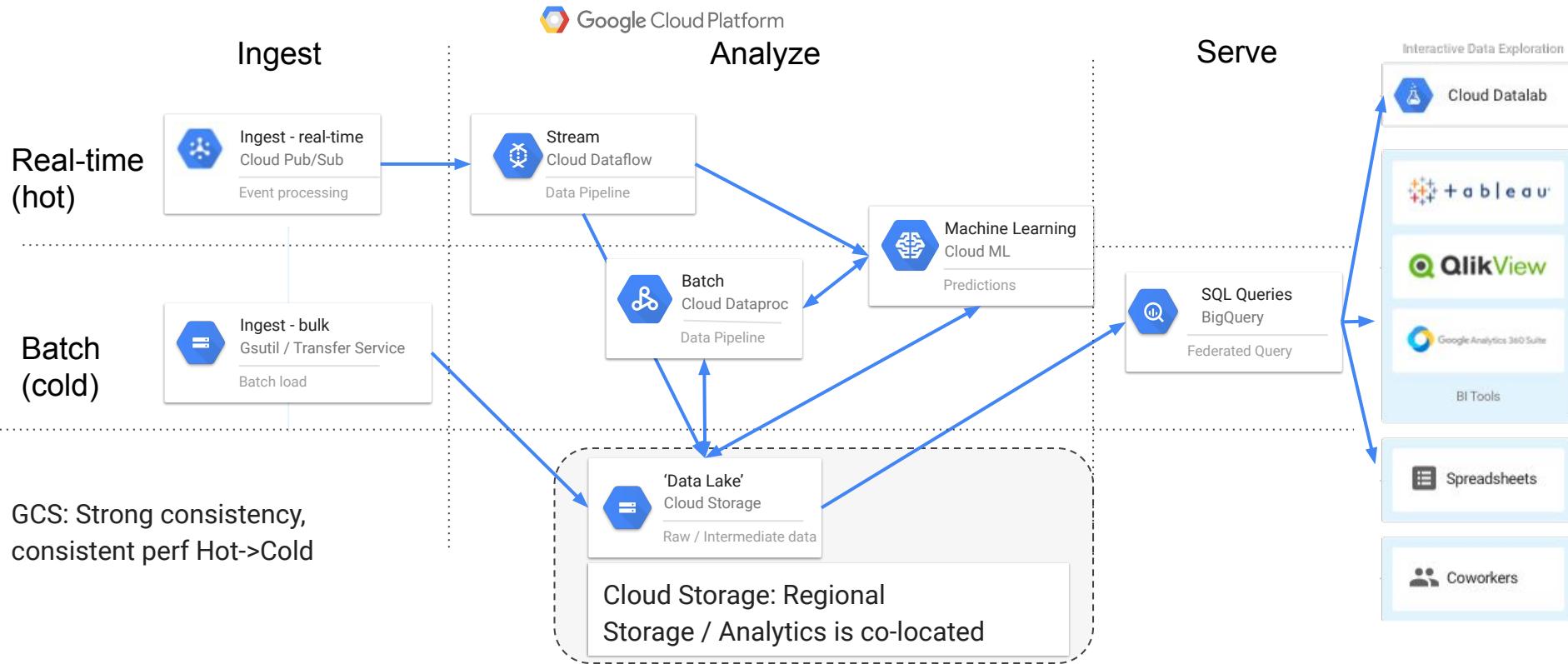


Tiering / Archive

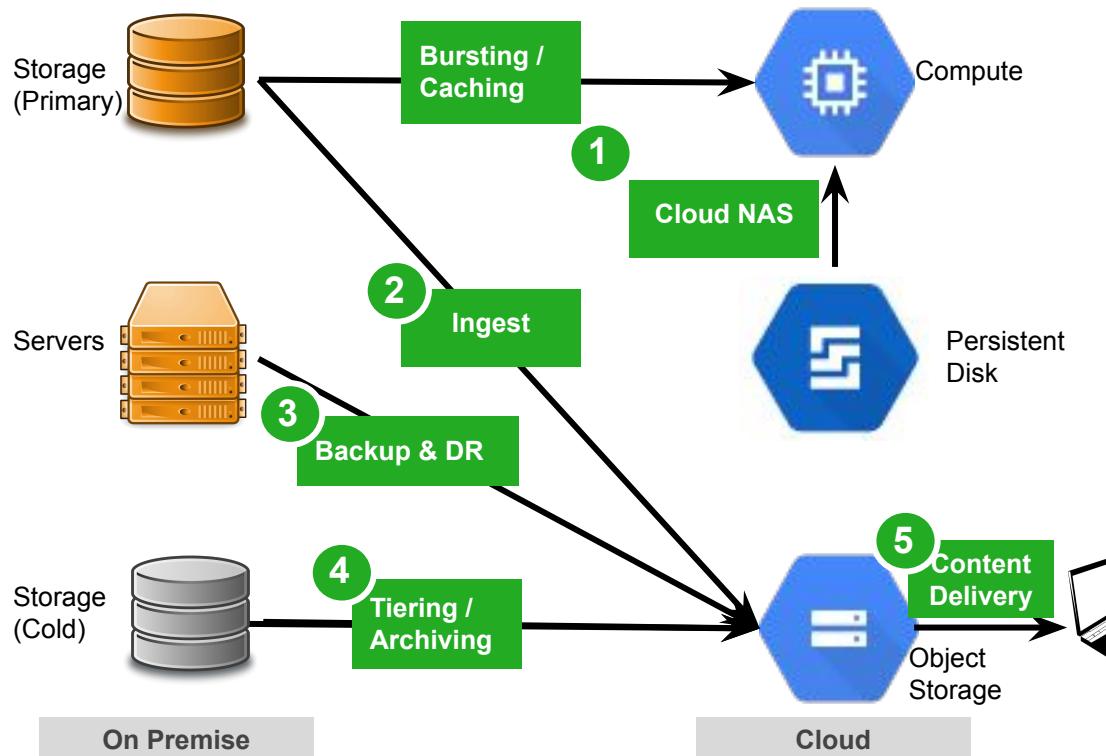


Cloud Storage - Analytics Pipeline

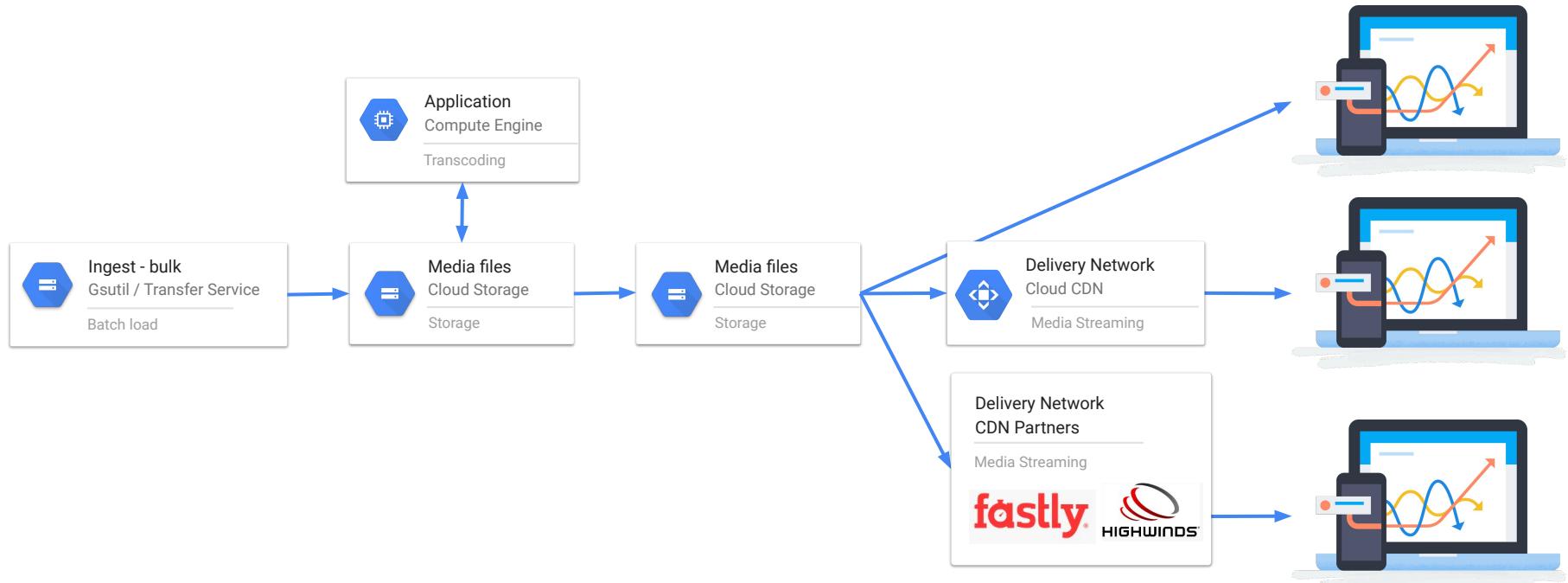
Omni-channel customer analysis (e.g. shopping, gaming, telco, finserv)



Hybrid use-cases: architecture and ecosystem



Cloud Storage - Content Storage and Delivery



App Dev: Storing Image and Video Files in Cloud Storage - Python

1 hour

7 Credits

★★★★★ Rate Lab

GSP185



Google Cloud Self-Paced Labs

Google Cloud Storage - Bucket Lock

40 minutes

5 Credits



GSP297



Google Cloud Self-Paced Labs



That's a wrap