# Introduction to Cloud Storage

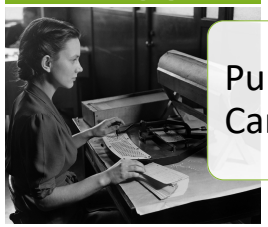
Harvesting, Storing, and Retrieving Data

# Pre-Cloud Storage: Traditional Data Storage Concepts

- Traditional storage has fast access speeds
- Security can be manually set up
- Users can recover the data anytime
- On-site backup and modification are easy.

# Pre-Cloud Storage: Traditional Data Storage

1795

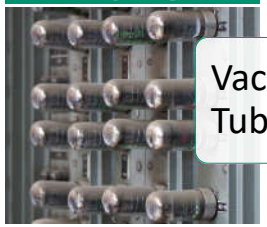


Punch Cards

1932

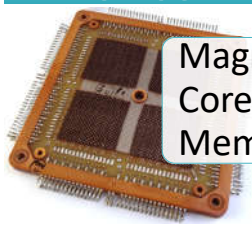
Magnetic Drum

1948



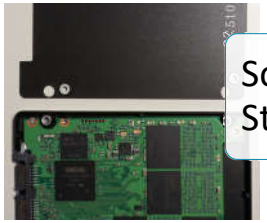
Vacuum Tube

1940s



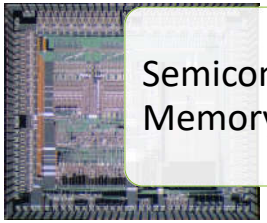
Magnetic Core Memory

1950



Solid State

1966



Semiconductor Memory

1969

8" Floppy Discs

1976

3.5" Floppy Discs

1980


CDs and DVDs

2000


Flash Drives

'12

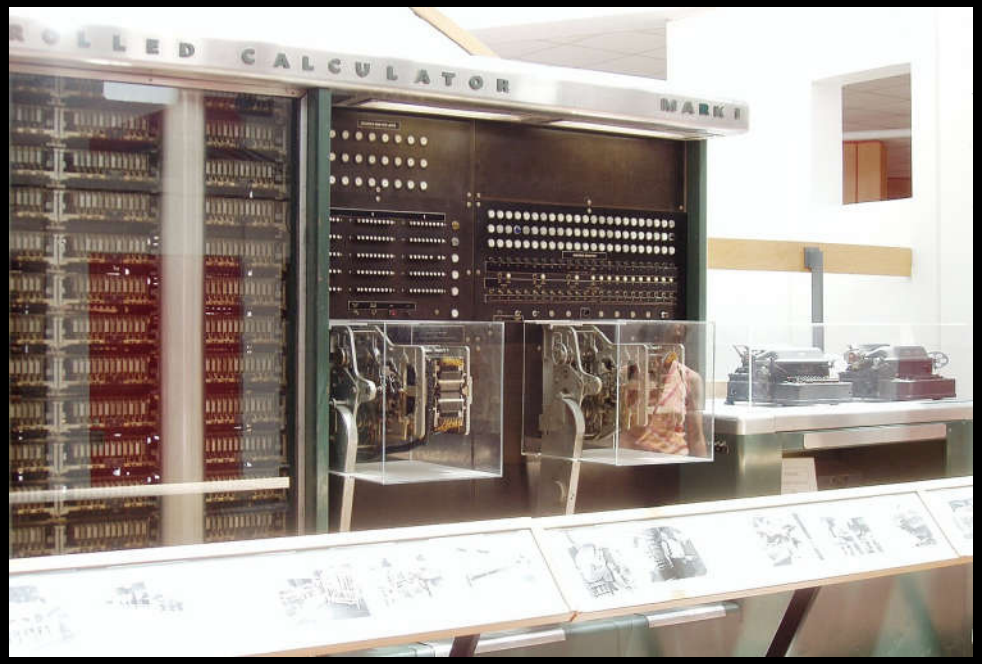
Data Silos

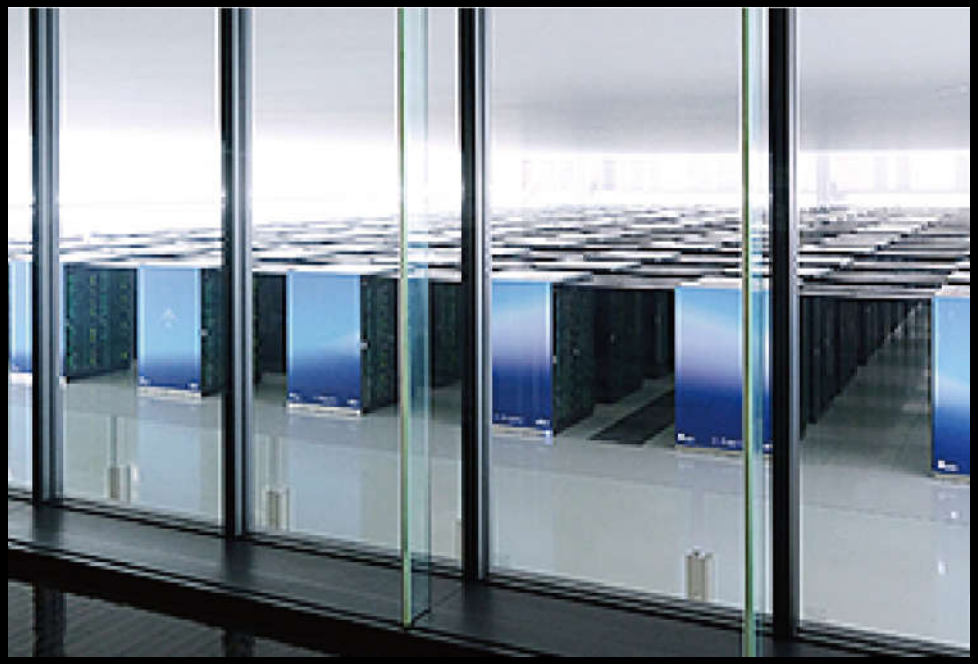
'12


Data Lakes

'12

Cloud Storage



## Pre-Cloud Storage: Traditional Data Warehouse Concepts

- Mainframes
- Supercomputers

# Mainframes: Yesterday and Today

- Still today:
  - 67 of the Fortune 100
  - 45 of the top 50 Banks
  - 8 of the top 10 Insurers and Telcos
  - 7 of the top 10 Retailers
  - 4 of the top 5 Airlines

- Benefits:
  - Improved overall operational efficiency
  - Improved security monitoring
  - Better adherence to regulatory and privacy compliance
  - Well-managed access (authentication and authorization)
  - Reduced risk to the business

# Supercomputers: Yesterday and Today

| | OS | CPU | Memory | Storage | FLOPS |
|---|---|---|---|---|---|
| CRAY1 | COS & UNICOS | 64-bit processor @ 80 MHz | 8.39 Megabyte | 303 Megabyte | 160 MFLOPS (mega FLOPS or $10^6$) |
| Fugaku | Custom Linux-based kernel | 158,976 nodes Fujitsu A64FX CPU (48+4 core) per node | HBM2 32 GiB/node | 1.6 TB NVMe SSD/16 nodes | 442 PFLOPS (peta FLOPS or $10^{15}$ ) |

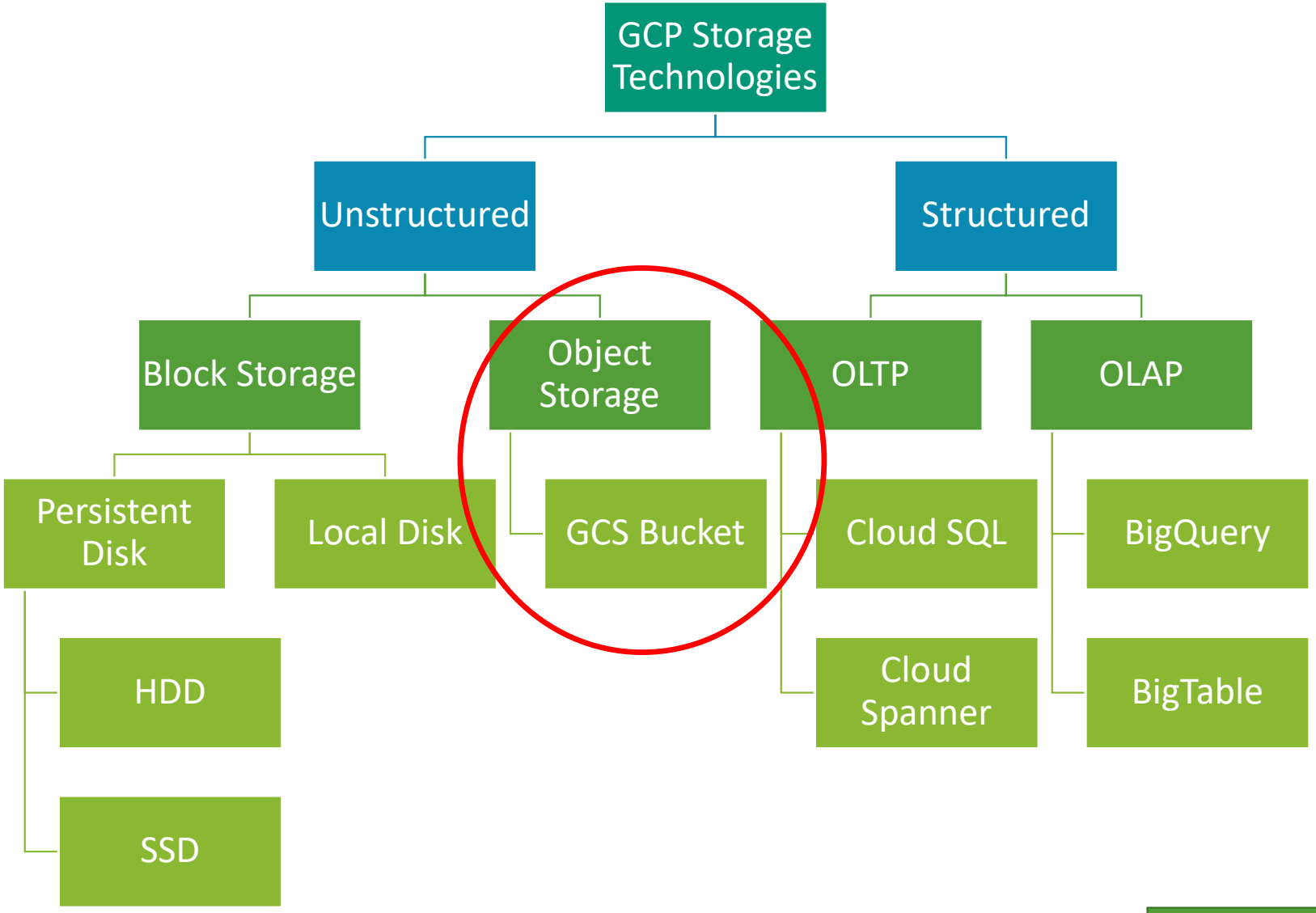
# Here in lies the need for storage in the cloud

We have discussed Big Data in an earlier lecture but how do we store Big Data?

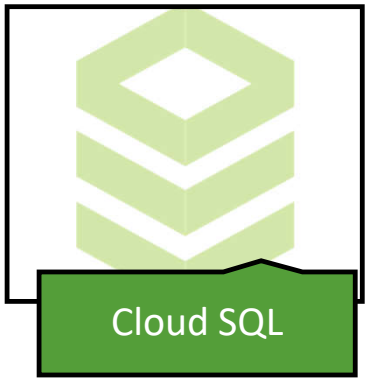
- Google gets over 3.5 billion searches daily
- In 2020, every person generated 1.7 megabytes of data every second
- 95% of businesses say that managing unstructured data is a problem
- 80 – 90% of the data generated today is unstructured
- Predicted that data creation will be more than 180 zettabytes by 2025
- Data interaction went up by 5000% from 2010 to 2020

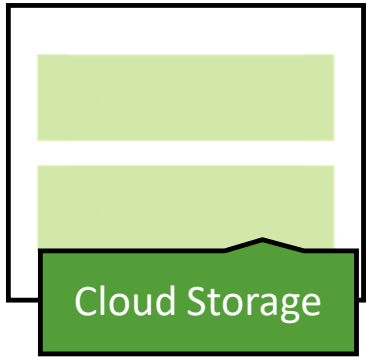
# Introduction to Google Cloud Storage

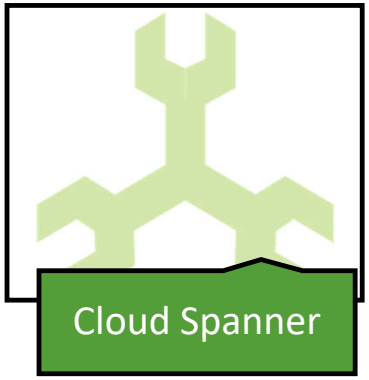
[redacted]

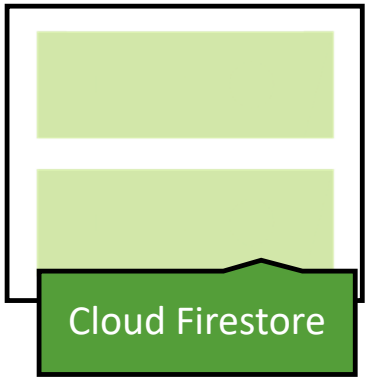

GCP Storage Technologies
Unstructured
Structured
Block Storage
Object Storage
OLTP
OLAP
Persistent Disk
Local Disk
GCS Bucket
Cloud SQL
BigQuery
HDD
Cloud Spanner
BigTable
SSD

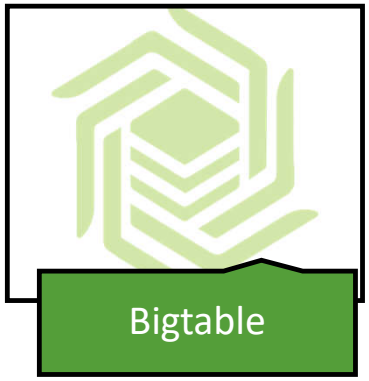
# Different GCP Storage Services

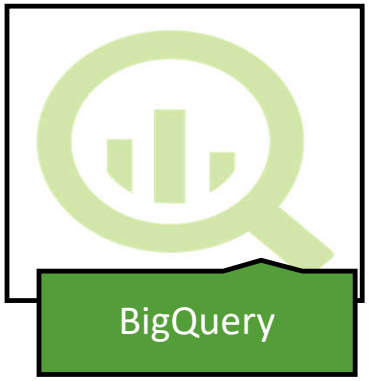


# What are the Cloud Storage Classes?

- Standard
- Nearline
- Coldline
- Archive

# Storage Classes

## Cloud Storage at a Glance

### High Frequency Access

#### Regional

- 99.9% Availability
- 99.999999999% Durability
- Lower Cost Storage
- No Retrieval Cost

#### Multi-Regional

- 99.95% Availability
- 99.999999999% Durability
- Geo Redundant
- No Retrieval Cost

### Low Frequency Access

#### Nearline

- 99.9% or 99.0% Availability
- 99.999999999% Durability
- Low-Cost Storage
- 30 day minimum

### Lowest Frequency Access

#### Coldline

- 99.9% or 99.0% Availability
- 99.999999999% Durability
- Low-Cost Storage
- 90 day minimum

# Understanding Cloud Storage

- Infinite Memory Storage
- Object Storage
- Pay for what you Need
- Files  Accessible around the World
- Low Maintenance

# Understanding GCP Cloud Storage

- Google uses buckets for storage
- Allows sharing across instances and zones
- Common Storage between on-premise and cloud

# Bucket Locations & Considerations

For data consumers, a suitable location strikes a balance between latency, availability, and bandwidth costs.

- If possible, choose the region that is convenient for your customers.
- If you need higher availability, consider using dual region.
- If the majority of your customers are geographically dispersed, use a multi-region.

| Continent | Region | |
|---|---|---|
| North America | NORTHAMERICA-NORTHEAST1 | Montreal |
| | US-CENTRAL1 | Iowa |
| | US-EAST1 | South Carolina |
| | US-EAST4 | Northern Virginia |
| | US-WEST1 | Oregon |
| | US-WEST2 | Los Angeles |
| | US-WEST3 | Salt Lake City |
| | US-WEST4 | Las Vegas |

# Bucket Naming Rules

- Contain only lowercase letters, numbers, dashes (-), and underscores (_)
- Spaces are not allowed.
- Must start and end with a number or letter.
- Must contain 3-63 characters.
- Cannot be represented as an IP address in dotted-decimal notation (for example, 192.168.5.4).
- Cannot begin with the "goog" prefix.
- Cannot contain "google" or close misspellings, such as "g00gle".

# Creating Buckets