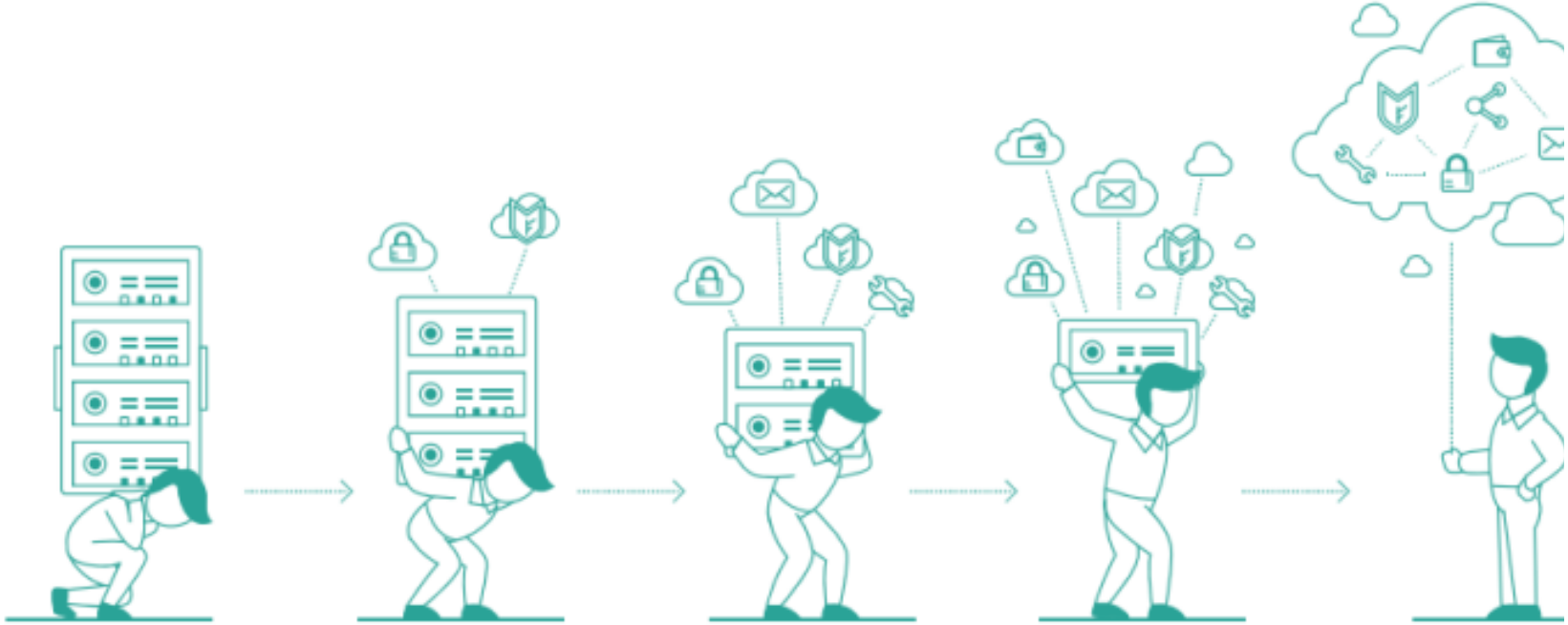
TEKsystems Global Services

Introduction to GCP and data in cloud!

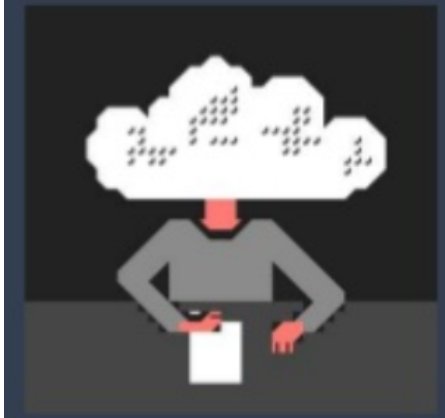
Presented by: Saranyan Mahadevan

Date of Presentation: 14 September 2020

The Evolution



What is Cloud Computing?



Cloud Computing is the use of hardware and software components to deliver a service to a network. Users can access these files and applications from any device that can access the internet

Automatic Software Integration

Backup and restore data

Unlimited Storage Capacity

Reliable

Cost efficient

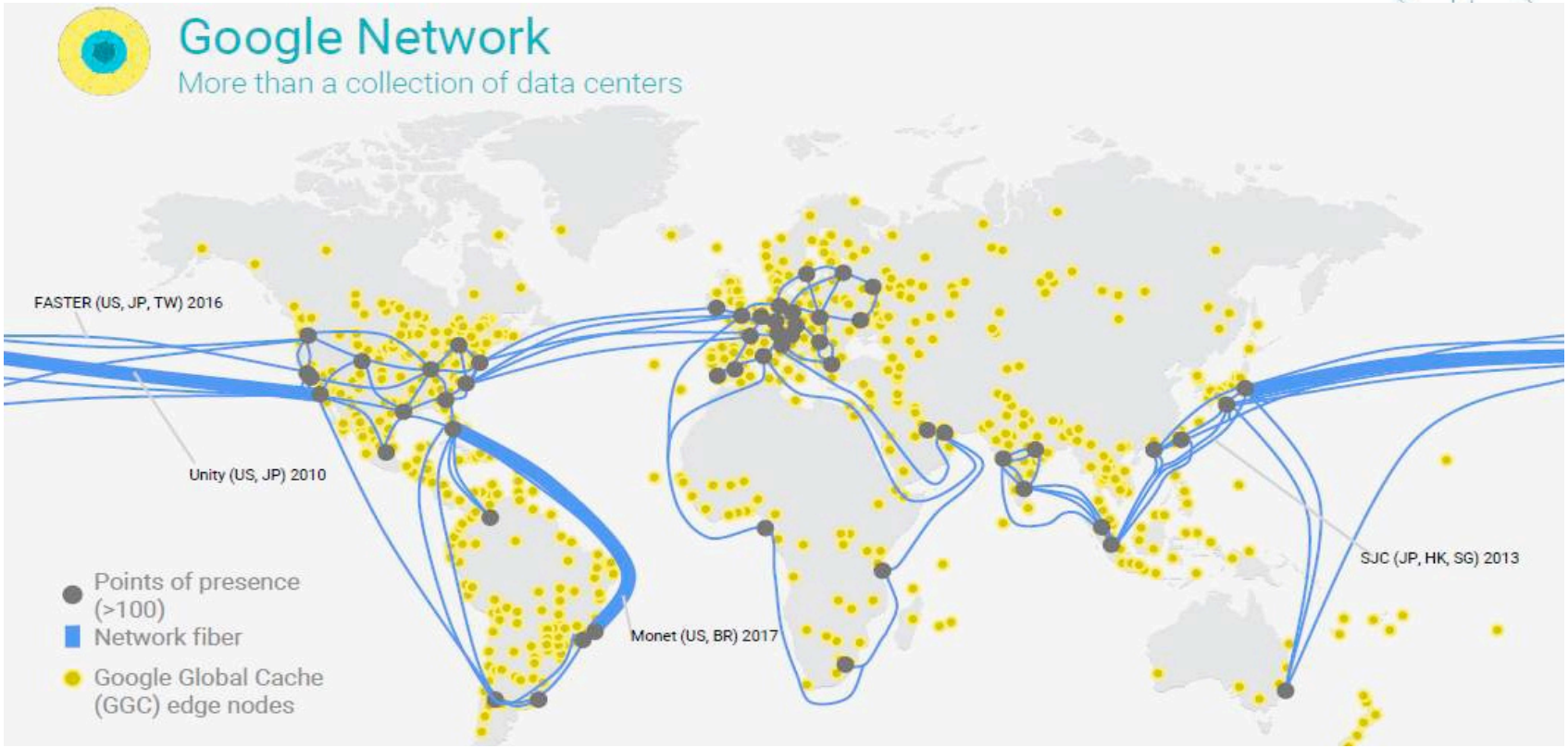
©TEKsystems | CHAMP 2021 | Private and Confidential

Cloud Players
rackspace
DigitalOcean
vmware
amazon
web services
terremark
IBM Cloud
Joyent
Microsoft Azure
Google Cloud Platform

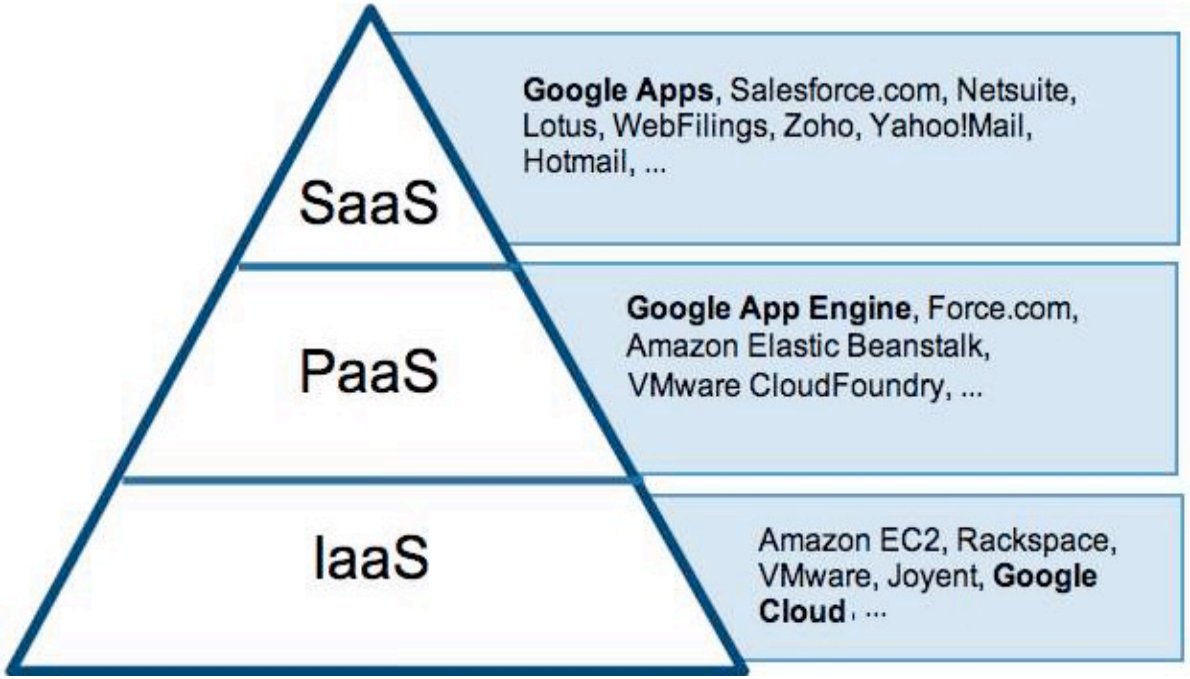
GCP is organized into regions and zones

- **Regions**: collections of zones
  - Specific geographical locations where you can run resources
  - Regions are interconnected using Google's global, meshed backbone network
- **Zones**: isolated deployment areas in a region
- Your resources can be regional, zonal, or in some cases multi-regional

©TEKsystems | CHAMP 2021 | Private and Confidential

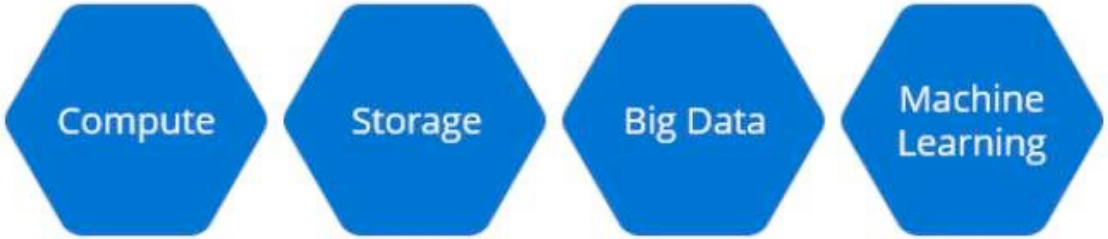

Google Network
More than a collection of data centers
FASTER (US, JP, TW) 2016
Unity (US, JP) 2010
Monet (US, BR) 2017
SJC (JP, HK, SG) 2013
Points of presence (>100)
Network fiber
Google Global Cache (GGC) edge nodes

Cloud Computing service levels



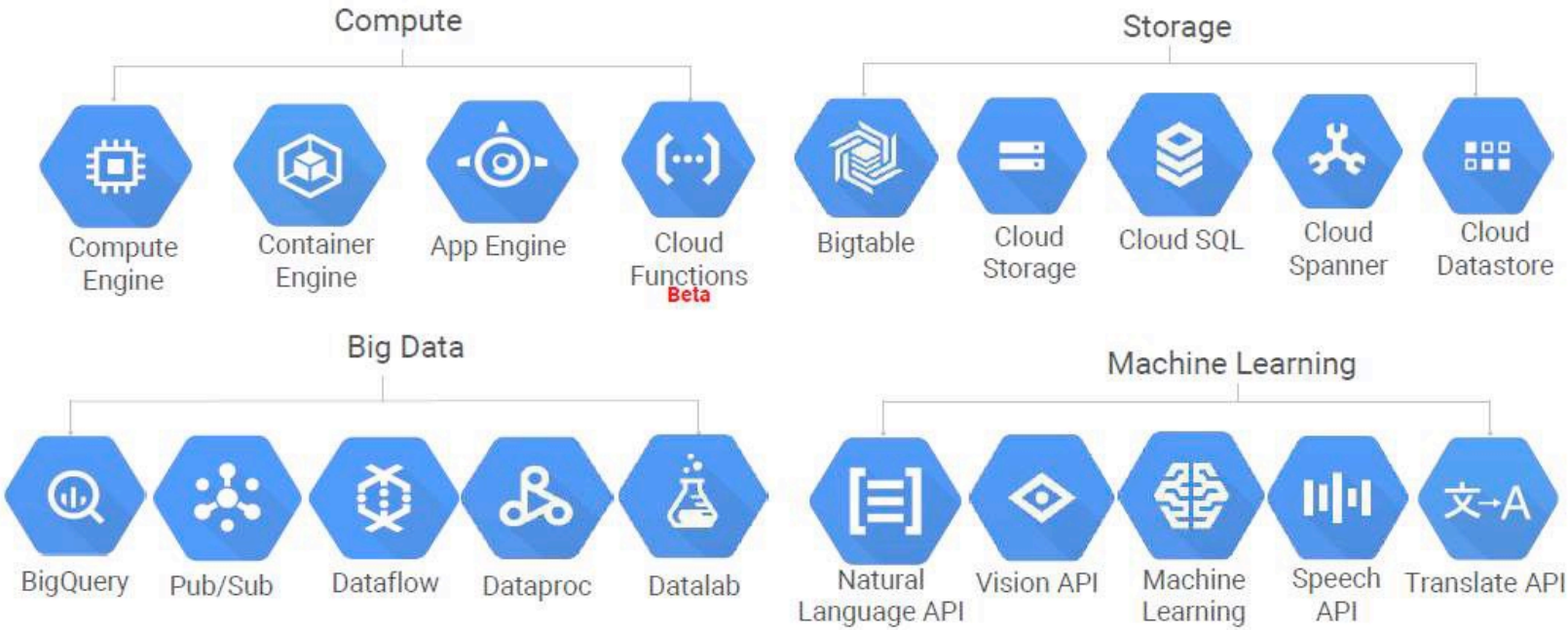
Google Cloud Platform

Products and services of GCP can be broadly categorized as **Compute**, **Storage**, **Big-data** and **Machine Learning**

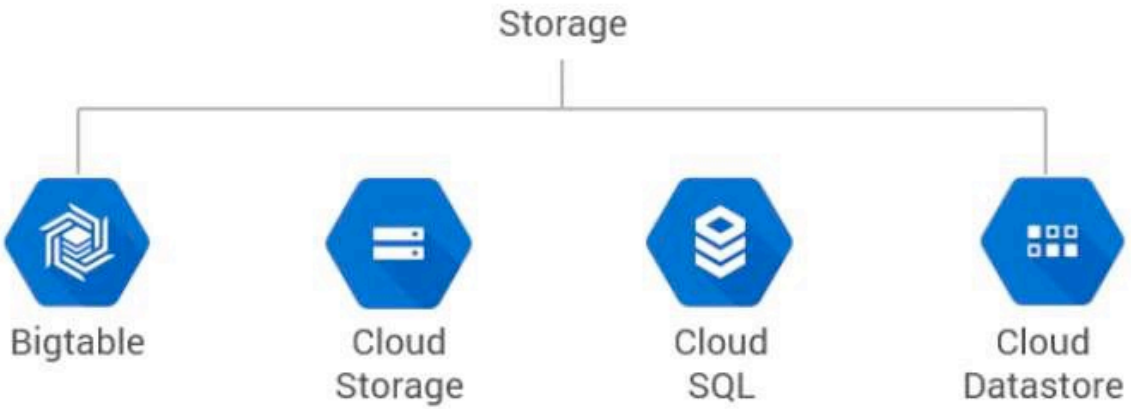


©TEKsystems | CHAMP 2021 | Private and Confidential

Google Cloud Platform



Google Cloud storage



Big-Data Services

BigQuery
Analytics data warehouse
Stream data at 100,000 rows per second

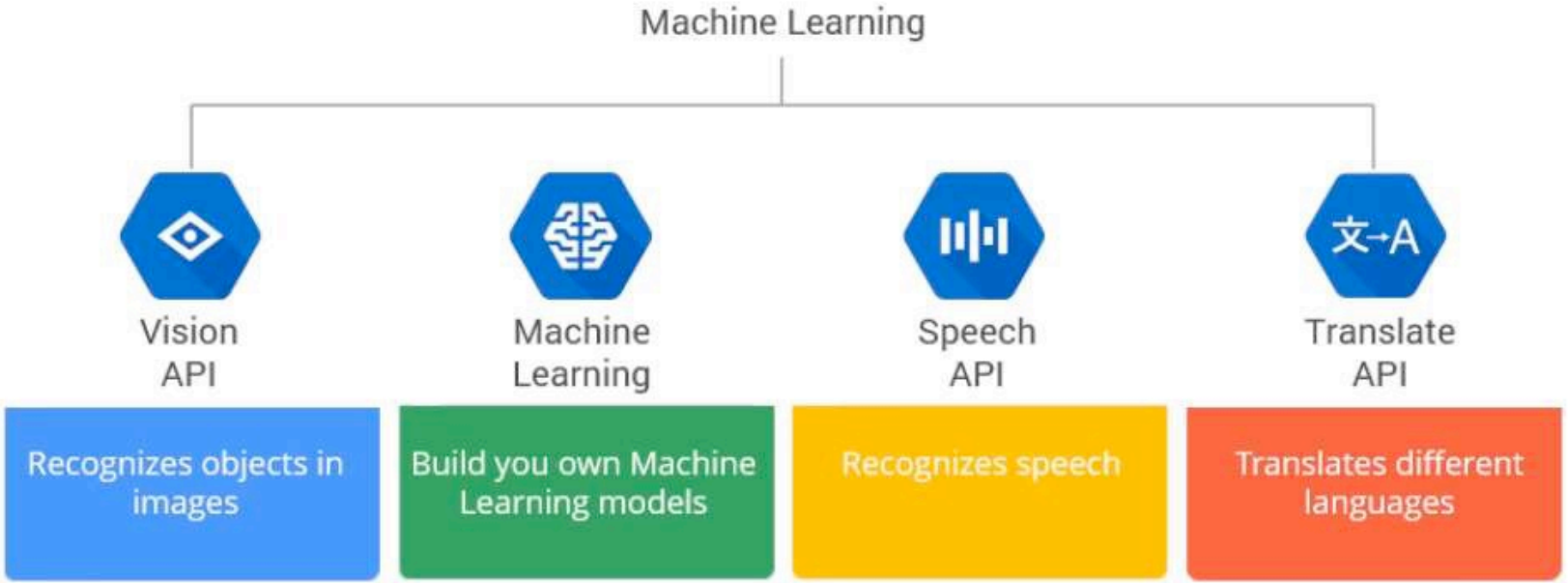
Dataflow
Stream and Batch processing of data
Unified programming model

Pub/Sub
Scalable & Reliable enterprise messaging middleware

Dataproc
Managed Hadoop, Spark, Pig and Hive at affordable pricing

©TEKsystems | CHAMP 2021 | Private and Confidential

Google Cloud storage



Interacting with cloud

Projects organize resources

- All Google Cloud Platform services are associated with a project that is used to:
  - Track resource and quota usage
  - Enable billing
  - Manage permissions and credentials
  - Enable services and APIs

Google Cloud Platform Console

- Centralized console for all project data
- Developer tools
  - Cloud Source Repositories
  - Cloud Shell
  - Test Lab (mobile app testing)
- Access to product APIs
- Manage, create projects

Google Cloud SDK

- SDK includes CLI tools for Cloud Platform products and services
  - gcloud, gsutil (Cloud Storage), bq (BigQuery)
- Available as Docker image
- Available via Cloud Shell
  - Containerized version of Cloud SDK running on Compute Engine instance

Machine Types

- Standard
- High-memory
- High-CPU
- Shared-core (small, non-resource intensive)

- Can attach GPU dies to most machine types

©TEKsystems | CHAMP 2021 | Private and Confidential

Storage Options & Big-data

Block storage for compute VMs - persistent disks or SSDs

Immutable blobs like video/images - Cloud Storage

OLTP - Cloud SQL or Cloud Spanner

NoSQL Documents like HTML/XML - Datastore

NoSQL Key-values - BigTable (~HBase)

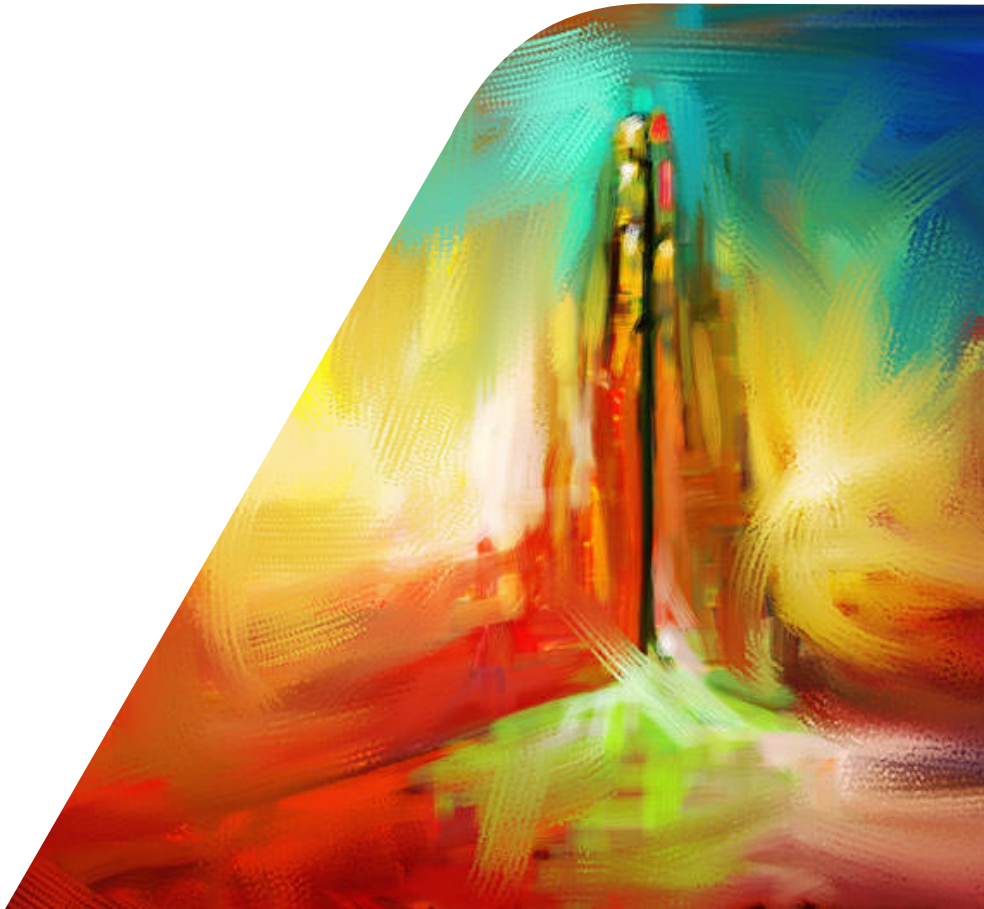
Getting data into Cloud Storage - Transfer service

©TEKsystems | CHAMP 2021 | Private and Confidential

Use cases

When you need	Use
Storage for Compute, Block Storage	Persistent (hard disks), SSD
Storing media, Blob Storage	Cloud Storage
SQL Interface atop file data	BigQuery
Document database, NoSQL	DataStore
Fast scanning, NoSQL	BigTable
Transaction Processing (OLTP)	Cloud SQL, Cloud Spanner
Analytics/Data Warehouse (OLAP)	BigQuery

©TEKsystems | CHAMP 2021 | Private and Confidential

TEKsystems
Own change

THANK YOU