

Cloud Computing



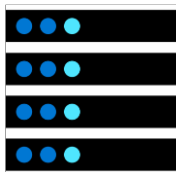
What is Cloud Computing?

"The practice of using a network of remote servers hosted on the Internet to store, manage, and process data, rather than a local server or a personal computer."

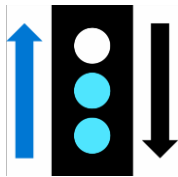
Oxford Dictionary

What is cloud computing?

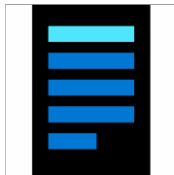
Cloud Computing is the delivery of computing services over the internet, enabling faster innovation, flexible resources, and economies of scale.



Compute



Networking



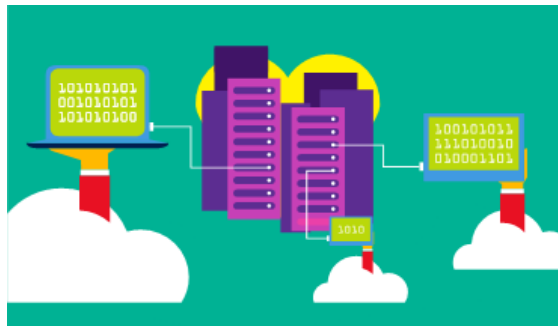
Storage



Analytics

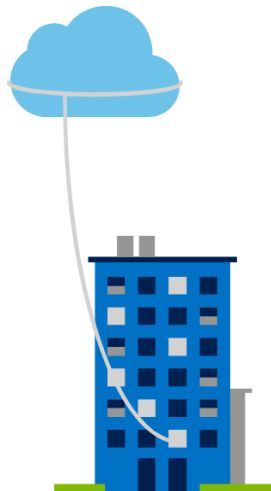
Public cloud

- Owned by cloud services or hosting provider.
- Provides resources and services to multiple organizations and users.
- Accessed via secure network connection (typically over the internet).



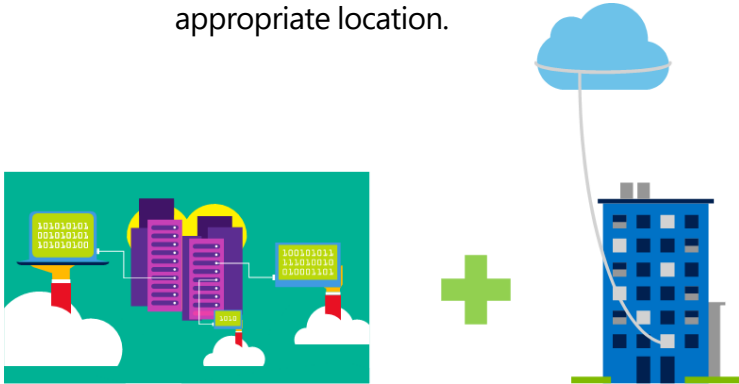
Private cloud

- Organizations create a cloud environment in their datacenter.
- Organization is responsible for operating the services they provide.
- Does not provide access to users outside of the organization.



Hybrid cloud

Combines **Public** and **Private** clouds to allow applications to run in the most appropriate location.



Cloud model comparison

Public Cloud

- No capital expenditures to scale up.
- Applications can be quickly provisioned and deprovisioned.
- Organizations pay only for what they use.

Private Cloud

- Hardware must be purchased for start-up and maintenance.
- Organizations have complete control over resources and security.
- Organizations are responsible for hardware maintenance and updates.

Hybrid Cloud

- Provides the most flexibility.
- Organizations determine where to run their applications.
- Organizations control security, compliance, or legal requirements.

Cloud benefits and considerations



Cloud Benefits

High availability

Fault tolerance

Scalability

Elasticity

Global reach

Customer latency capabilities

Agility

Predictive cost considerations

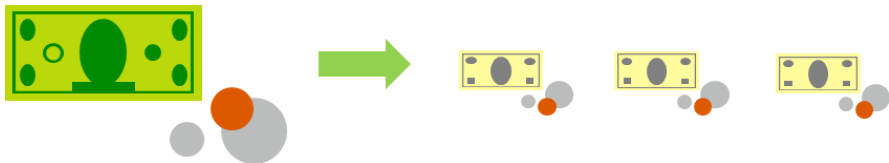
Compare CapEx vs. OpEx

Capital Expenditure (CapEx)

- The up-front spending of money on physical infrastructure.
- Costs from CapEx have a value that reduces over time.

Operational Expenditure (OpEx)

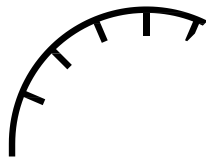
- The spending and billing of services or products as needed.
- Expenses are deducted in the same year.



Consumption-based model

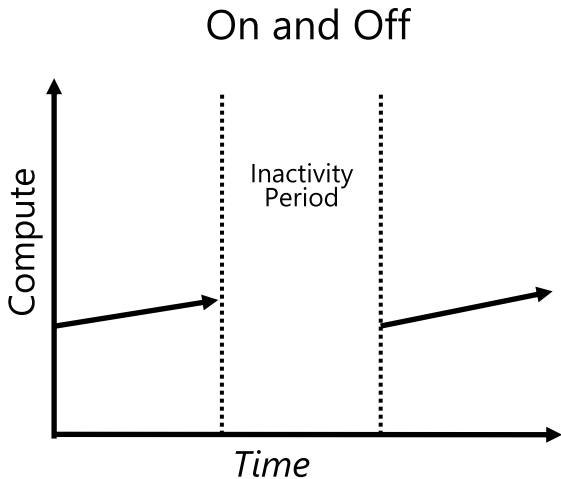
Cloud service providers operate on a consumption-based model, which means that end users only pay for the resources that they use. Whatever they use is what they pay for.

- Better cost prediction
- Prices for individual resources and services are provided
- Billing is based on actual usage



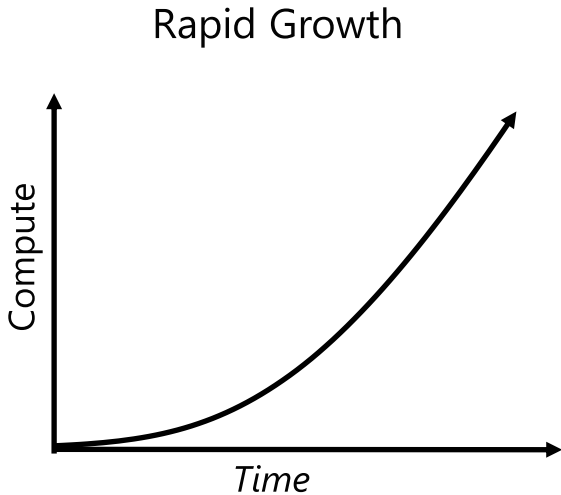
Typical Computing Pattern

- On & off workloads
 - Batch jobs
- Wasted Capacity
- Time to market can be cumbersome



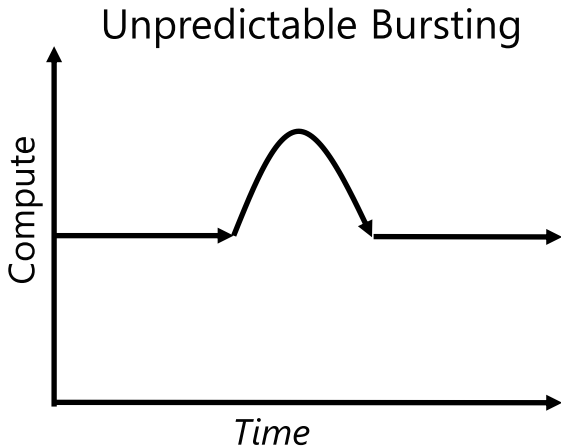
Typical Computing Pattern

- Rapidly growing company
- Major challenge for IT dept. to keep up with growth
- Potential loss of business opportunity
- Potential customer service problems



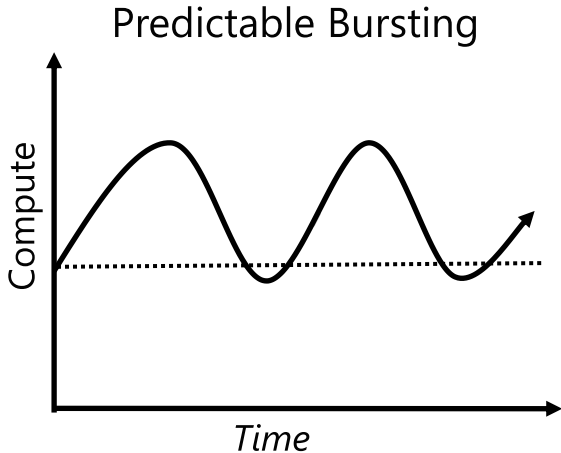
Typical Computing Pattern

- Unexpected peak in demand
- Loss of business opportunity
- Wasted capacity if demand wanes



Typical Computing Pattern

- Seasonal peaks and troughs
- Provisioning dilemma
 - Wasted capacity or
 - Loss of business



Cloud Computing Examples

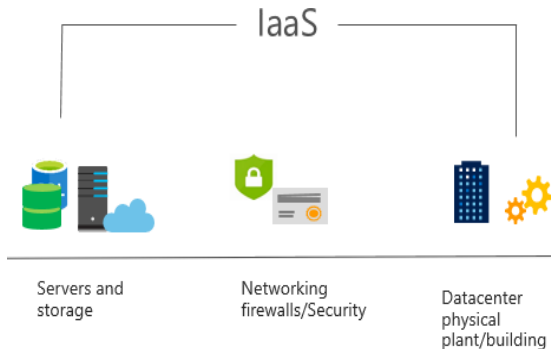
- A large enterprise quickly & economically deploys new internal applications to its distributed workforce.
- An e-commerce website accommodates sudden demand for a "hot" product caused by a viral buzz.
- A pharmaceutical research firm executes large-scale simulations using computing power provided by cloud vendors.
- A media company serves unlimited video, music, and other media to their worldwide customer base.

Cloud services



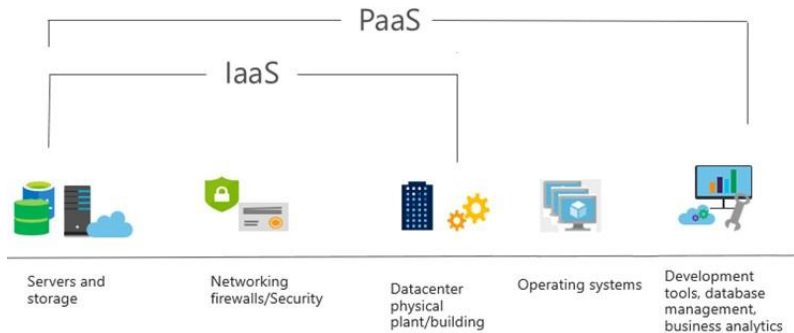
Infrastructure as a Service (IaaS)

Build pay-as-you-go IT infrastructure by renting servers, virtual machines, storage, networks, and operating systems from a cloud provider.



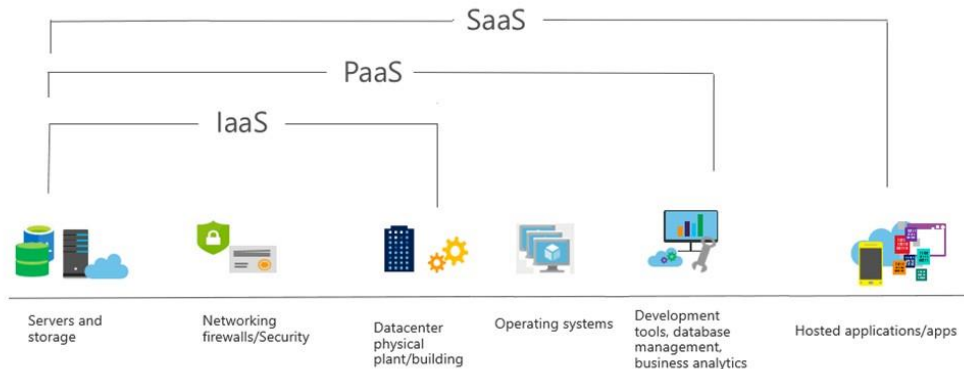
Platform as a Service (PaaS)

Provides environment for building, testing, and deploying software applications; without focusing on managing underlying infrastructure.



Software as a Service (SaaS)

Users connect to and use cloud-based apps over the internet: for example, Microsoft Office 365, email, and calendars.



Cloud service comparison

IaaS

The most flexible cloud service.

You configure and manage the hardware for your application.

PaaS

Focus on application development.

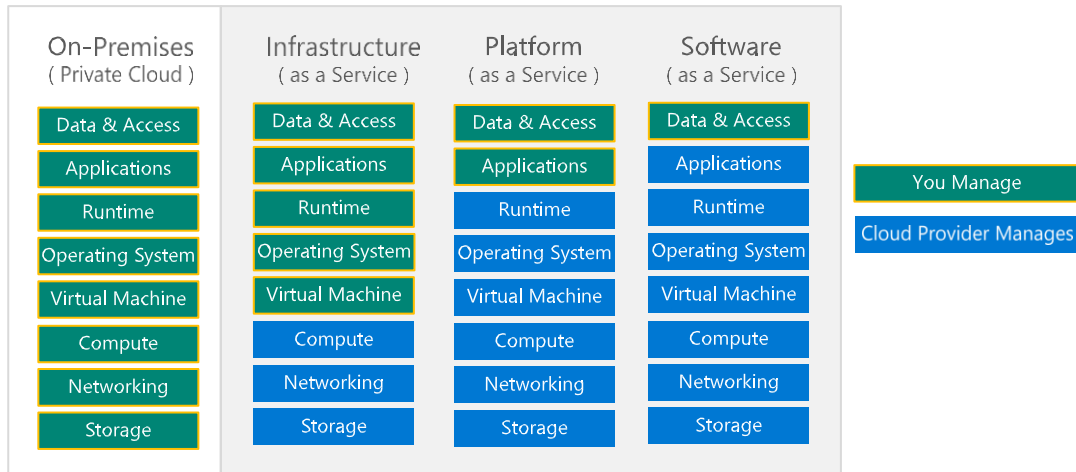
Platform management is handled by the cloud provider.

SaaS

Pay-as-you-go pricing model.

Users pay for the software they use on a subscription model.

Shared responsibility model



Cloud Vendor - Azure & AWS

Microsoft Azure and Amazon Web Services (AWS) offer broad and deep capabilities with global coverage

Category	Azure Service	AWS Service
Computing infrastructure	Virtual Machines	EC2
Object storage infrastructure	Blob Storage	S3
Networking	Virtual Network	Virtual Private Cloud
Relational database-as-a-service	SQL Database	RDS
NoSQL document database	DocumentDB	DynamoDB
Big data processing	HDInsight	Elastic MapReduce (EMR)
Visualization	Power BI	QuickSight

Cloud Vendor - Bluemix & Google

IBM Bluemix and Google Cloud each offer and deploy applications on highly-scalable and reliable infrastructure

Category	Bluemix	Google Service
Computing infrastructure	Virtual Server, Containers	Compute Engine
Object storage infrastructure	Object, Block Storage	Cloud Storage
Networking	Virtual Private Network	Cloud Virtual Network
Relational database-as-a-service	SQL Database	Cloud SQL
NoSQL document database	MongoDB	Cloud Datastore, Bigtable
Big data processing	Analytics for Apache Hadoop	BigQuery, Cloud Dataproc
Visualization		

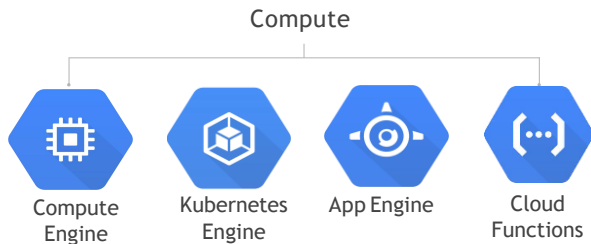
Vendor Lock-In

Companies that adopt cloud computing must be wary of potential vendor lock-in issues

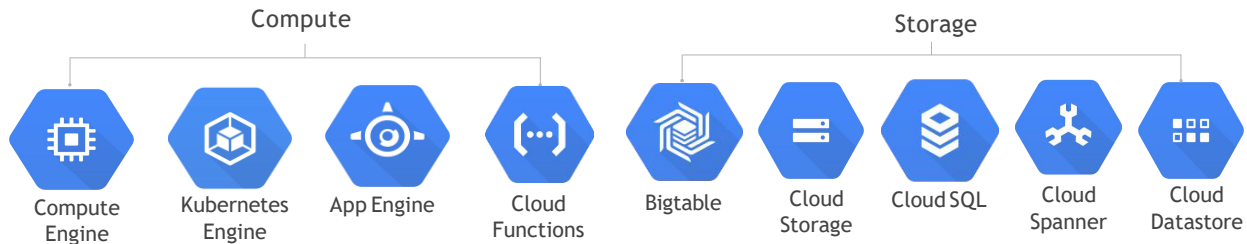
- Company's entire data is stored with a single vendor's cloud storage
- Company relies on a single vendor for all of its computations
- Changing vendors can be very costly

Google Cloud Platform

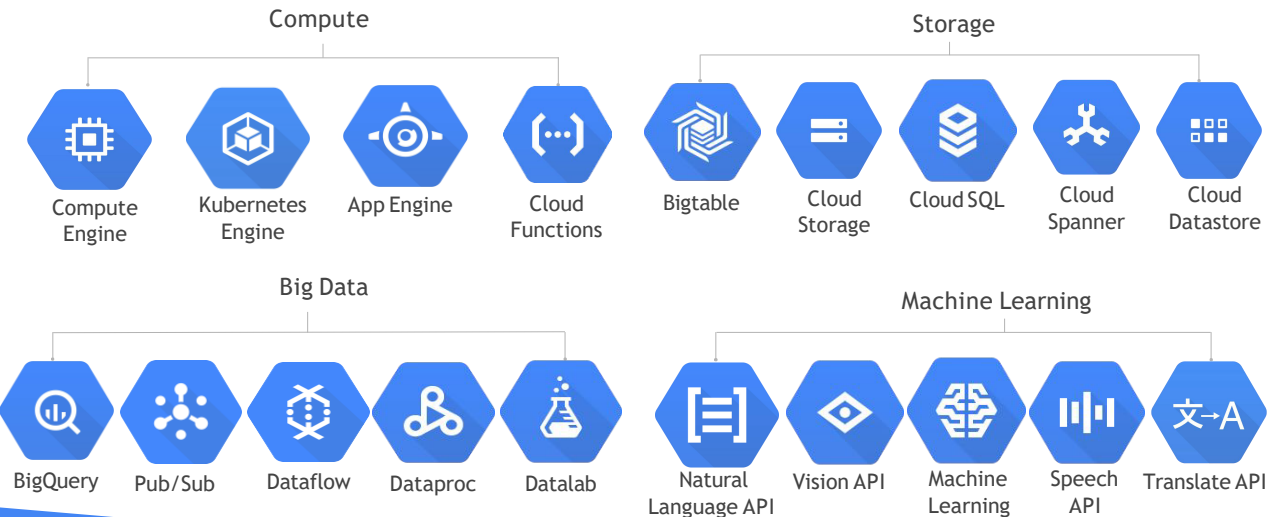
Google Cloud Platform offers a range of compute services



Google Cloud Platform offers a range of storage services

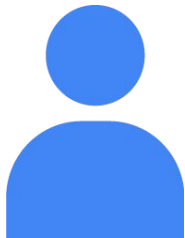


Google Cloud Platform offers services for getting value from data



Identity and Access Management

Google Cloud Identity and Access Management defines...



Who



can do what



on which resource

Who: IAM policies can apply to any of four types of principals



Who



Google account or Cloud Identity user

test@gmail.com

test@example.com



Service account

test@project_id.iam.gserviceaccount.com



Google group

test@googlegroups.com

G Suite

Cloud Identity or G Suite domain

example.com

Can do what: IAM roles are collections of related permissions



can do what



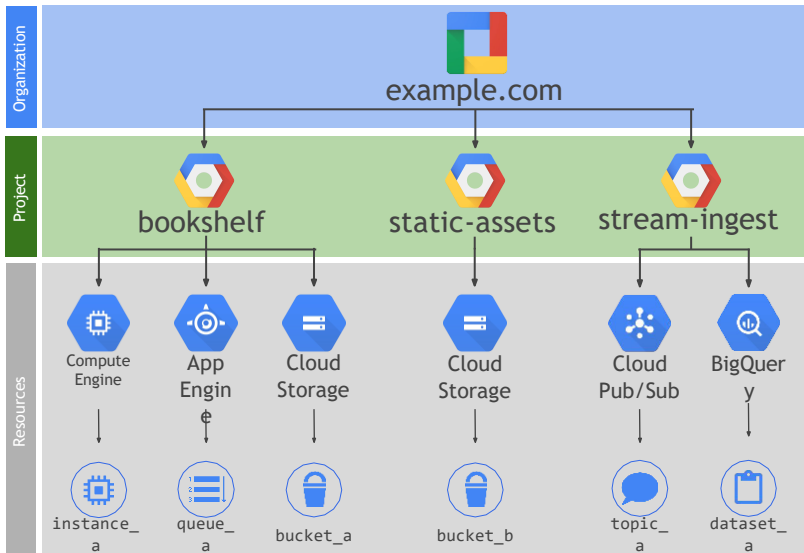
InstanceAdmin
Role

Service	Resource	Verb
<i>compute</i>	<i>instances</i>	<i>list</i>
<i>compute</i>	<i>instances</i>	<i>delete</i>
<i>compute</i>	<i>instances</i>	<i>start</i>
...		

On which resource: Users get roles on specific items in the hierarchy



on which resource



There are three types of IAM roles

Primitive



Predefined



Custom



IAM **primitive** roles apply across all GCP services in a project



can do what



on all resources

IAM primitive roles offer fixed, coarse-grained levels of access



Owner

- Invite members
- Remove members
- Delete projects
- And...



Editor

- Deploy applications
- Modify code
- Configure services
- And...



Viewer

- Read-only access



Billing administrator

- Manage billing
- Add and remove administrators

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

IAM **predefined** roles apply to a particular GCP service in a project

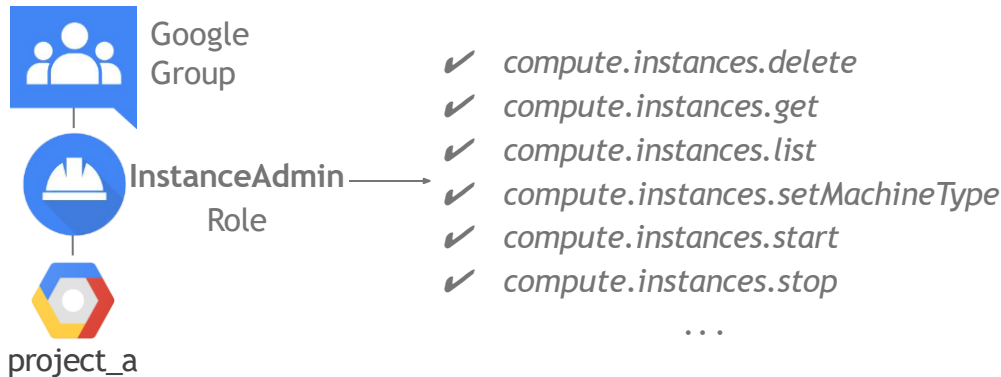


can do what

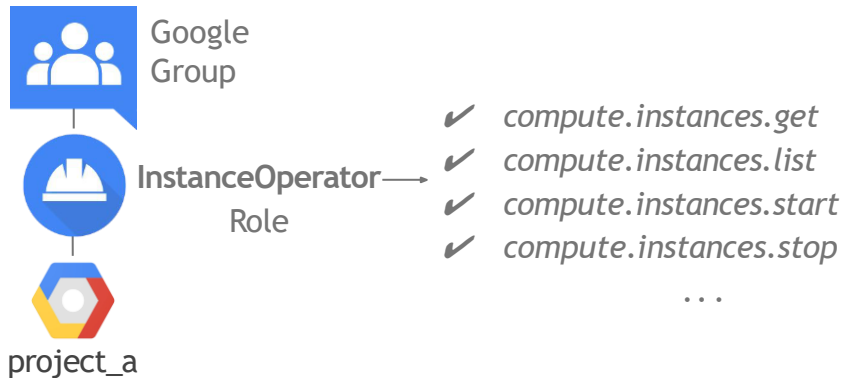


on Compute Engine resources
in this project, or folder, or org

IAM predefined roles offer more fine-grained permissions on particular services



IAM custom roles let you define a precise set of permissions



Service Accounts control server-to-server interactions

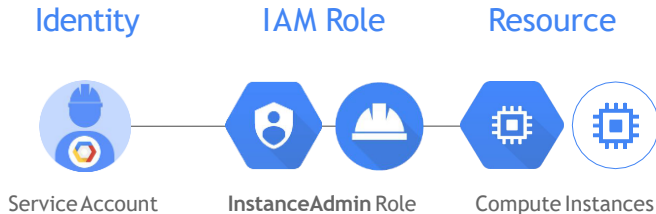
- Provide an identity for carrying out **server-to-server** interactions in a project
- Used to **authenticate** from one service to another
- Used to **control privileges** used by resources
 - So that applications can perform actions on behalf of authenticated end users
- Identified with an **email address**:

PROJECT_NUMBER-compute@developer.gserviceaccount.com

PROJECT_ID@appspot.gserviceaccount.com

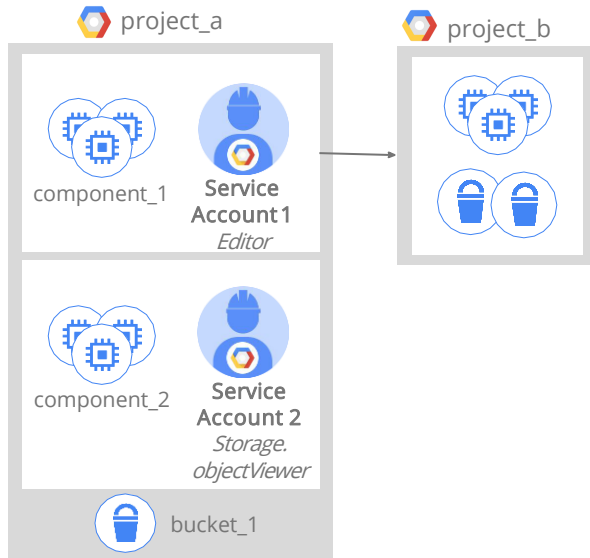
Service Accounts and IAM

- Service accounts authenticate using keys.
 - Google manages keys for Compute Engine and App Engine.
- You can assign a predefined or custom IAM role to the service account.



Example: Service Accounts and IAM

- VMs running component_1 are granted **Editor** access to project_b using *Service Account 1*.
- VMs running component_2 are granted **objectViewer** access to bucket_1 using *Service Account 2*.
- Service account permissions can be changed without recreating VMs.



Compute

Compute Engine offers managed virtual machines

- High CPU, high memory, standard and shared-core machine types
- Persistent disks
 - Standard, SSD, localSSD
 - Snapshots
- Resize disks with no downtime
- Instance metadata and startup scripts



Scale up or scale out with Compute Engine

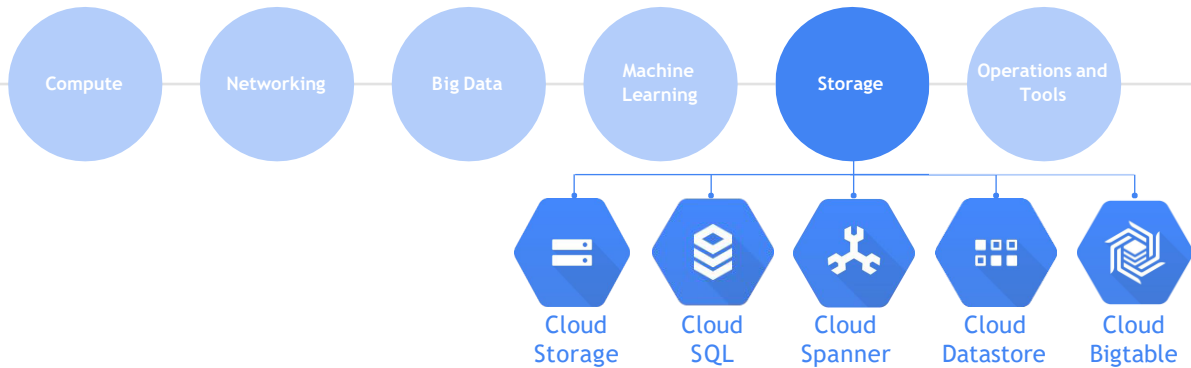


Use big VMs for memory- and compute-intensive applications



Use Autoscaling for resilient, scalable applications

Google Cloud Platform



Agenda

Cloud Storage

Cloud Bigtable

Cloud SQL and Cloud Spanner

Cloud Datastore

Comparing storage options

Quiz and Lab

Cloud Storage is binary large-object storage

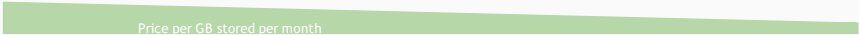
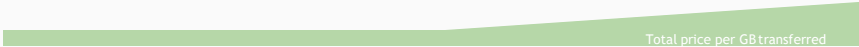
- High performance, internet-scale
- Simple administration
 - Does not require capacity management
- Data encryption at rest
- Data encryption in transit by default from Google to endpoint
- Online and offline import services are available



Your Cloud Storage files are organized into buckets

Bucket attributes	Bucket contents
Globally unique name	Files (in a flat namespace)
Storage class	
Location (region or multi-region)	
IAM policies or Access Control Lists	Access Control Lists
Object versioning setting	
Object lifecycle management rules	

Choosing among Cloud Storage classes

	Multi-regional	Regional	Nearline	Coldline
Intended for data that is...	Most frequently accessed	Accessed frequently within a region	Accessed less than once a month	Accessed less than once a year
Availability SLA	99.95%	99.90%	99.00%	99.00%
Access APIs	<i>Consistent APIs</i>			
Access time	<i>Millisecond access</i>			
Storage price				
Retrieval price				
Use cases	Content storage and delivery	In-region analytics, transcoding	Long-tail content, backups	Archiving, disaster recovery

There are several ways to bring data into Cloud Storage



Online transfer

Self-managed copies using command-line tools or drag-and-drop



Storage Transfer Service

Scheduled, managed batch transfers



Transfer Appliance ^{Beta}

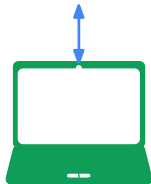
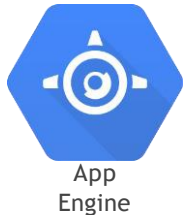
Rackable appliances to securely ship your data

Cloud Storage works with other GCP services

Import
and
export
tables



Object
storage,
logs, and
Datastore
backups



Startup
scripts,
images,
and
general
object
storage



Import
and
export
tables

Agenda

Cloud Storage

Cloud Bigtable

Cloud SQL and Cloud Spanner

Cloud Datastore

Comparing storage options

Integrations with other services

Quiz and Lab

Cloud Bigtable is managed NoSQL

- Fully managed NoSQL, wide-column database service for terabyte applications
- Integrated
 - Accessed using HBase API
 - Native compatibility with bigdata, Hadoop ecosystems

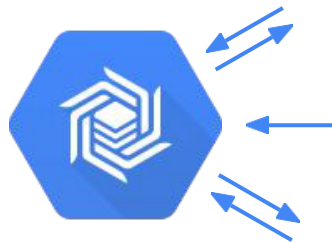


Why choose Cloud Bigtable?

- Replicated storage
- Data encryption in-flight and at rest
- Role-based ACLs
- Drives major applications such as Google Analytics and Gmail



Bigtable Access Patterns



Cloud Bigtable

Application API

Data can be read from and written to Cloud Bigtable through a data service layer like Managed VMs, the HBase REST Server, or a Java Server using the HBase client. Typically this will be to serve data to applications, dashboards, and dataservices.

Streaming

Data can be streamed in (written event by event) through a variety of popular stream processing frameworks like Cloud Dataflow Streaming, Spark Streaming, and Storm.

Batch Processing

Data can be read from and written to Cloud Bigtable through batch processes like Hadoop MapReduce, Dataflow, or Spark. Often, summarized or newly calculated data is written back to Cloud Bigtable or to a downstream database.

Agenda

Cloud Storage

Cloud Bigtable

Cloud SQL and Cloud Spanner

Cloud Datastore

Comparing storage options

Quiz and Lab

Cloud SQL is a managed RDBMS

- Offers MySQL and PostgreSQL databases as a service
- Automatic replication
- Managed backups
- Vertical scaling (read and write)
- Horizontal scaling (read)
- Google security



Cloud SQL can be used with other GCP services



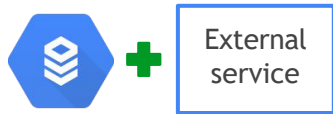
Cloud SQL can be used with App Engine using standard drivers.

You can configure a Cloud SQL instance to follow an App Engine application.



Compute Engine instances can be authorized to access Cloud SQL instances using an external IP address.

Cloud SQL instances can be configured with a preferred zone.



Cloud SQL can be used with external applications and clients.

Standard tools can be used to administer databases.

External read replicas can be configured.

Cloud Spanner is a horizontally scalable RDBMS

Cloud Spanner supports:

- Automatic replication
- Strong global consistency
- Managed instances with high availability
- SQL (ANSI 2011 with extensions)



Agenda

Cloud Storage

Cloud Bigtable

Cloud SQL and Cloud Spanner

Cloud Datastore

Comparing storage options

Quiz and Lab

Cloud Datastore is a horizontally scalable NoSQL DB

- [NoSQL](#) designed for application backends
- Fully managed
 - Uses a distributed architecture to automatically manage scaling
- Built-in redundancy
- Supports [ACID](#) transactions



Google Cloud Datastore: benefits

- Schemaless access
 - No need to think about underlying data structure
- Local development tools
- Includes a free daily quota
- Access from anywhere through a [RESTful interface](#)



Agenda

Cloud Storage

Cloud Bigtable

Cloud SQL and Cloud Spanner

Cloud Datastore

Comparing storage options

Integrations with other services

Quiz and Lab

Comparing storage options: technical details

	Cloud Datastore	Bigtable	Cloud Storage	Cloud SQL	Cloud Spanner	BigQuery
Type	NoSQL document	NoSQL wide column	Blobstore	Relational SQL for OLTP	Relational SQL for OLTP	Relational SQL for OLAP
Transactions	Yes	Single-row	No	Yes	Yes	No
Complex queries	No	No	No	Yes	Yes	Yes
Capacity	Terabytes+	Petabytes+	Petabytes+	Up to ~10 TB	Petabytes	Petabytes+
Unit size	1 MB/entity	~10 MB/cell ~100 MB/row	5 TB/object	Determined by DB engine	10,240 MiB/row	10 MB/row

Comparing storage options: use cases

	Cloud Datastore	Cloud Bigtable	Cloud Storage	Cloud SQL	Cloud Spanner	BigQuery
Type	NoSQL document	NoSQL wide column	Blobstore	Relational SQL for OLTP	Relational SQL for OLTP	Relational SQL for OLAP
Best for	Semi-structured application data, durable key-value data	“Flat” data, Heavy read/write, events, analytical data	Structured and unstructured binary or object data	Web frameworks, existing applications	Large-scale database applications (> ~2 TB)	Interactive querying, offline analytics
Use cases	Getting started, App Engine applications	AdTech, Financial and IoT data	Images, large media files, backups	User credentials, customer orders	Whenever high I/O, global consistency is needed	Data warehousing

Serverless Computing and Google Cloud Functions

Serverless

- **Definition**

- Serverless is a cloud computing execution model in which the cloud provider runs the server, and dynamically manages the allocation of machine resources. Pricing is based on the actual amount of resources consumed by an application, rather than on pre-purchased units of capacity.
- Generally, Serverless = FaaS + BaaS

- **Pros & Cons**

- Scale to **zero/infinite is possible** (Google: “from prototype to production to planet-scale”) → Maximum elasticity
- Cost-effective since there is **no fee for idle times**.
 - Flip side: infrequently-used code may suffer from greater latency (performance issues)
- Lack of standards → vendor lock-in



Serverless

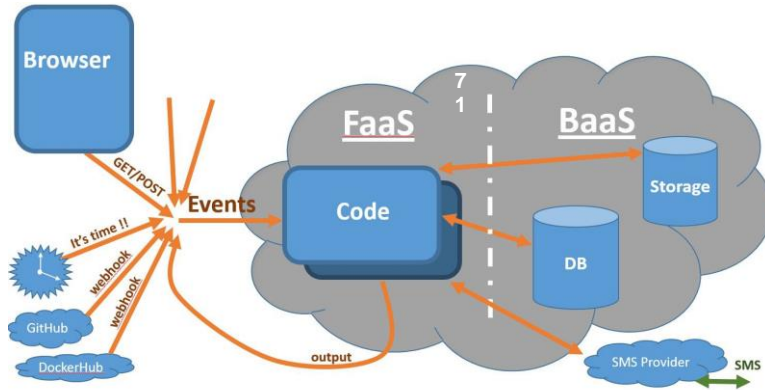
- **FaaS**

- FaaS is a form of serverless computing, where you execute certain functions of your app. in a abstracted computing env.
- Developers deploy an individual “function” (a piece of business logic). They are expected to start within milliseconds and process individual requests and then the process ends.
- Major providers (see <http://serverlesscalc.com/>)
 - AWS Lambda (Amazon)
 - Azure Functions (Microsoft)
 - Cloud Functions (Google)
 - IBM Cloud Functions (IBM)
 - Pivotal Function Service (VMware)
- PaaS vs. FaaS
 - PaaS simplifies dev./deployment process of applications and they run on server like a typical application once they deployed
 - FaaS provides the ability to deploy a single function (part of application) and **scale to zero** is possible since the rest is managed by the provider



Serverless

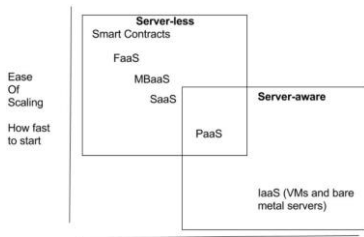
- Event-driven Architecture



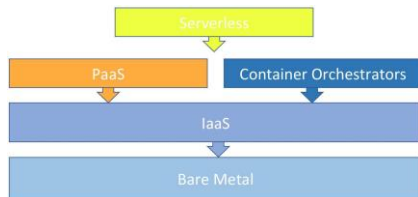
Serverless

• Use cases

- Serverless platforms are for **short-running, stateless computation and event-driven applications** which scales up and down instantly and automatically.
- General characteristics
 - latency tolerant, event-driven, short-lived, periodic
 - simply, where it doesn't make sense to pay for always-on services



Granularity - Average time-to-live



source: Cloud Computing resources abstraction in FaaS env. / Daniele Spiga / SOSC

2019

Serverless

• Use cases


good for
short-running
stateless
event-driven



-  Microservices
-  Mobile Backends
-  Bots, ML Inferencing
-  IoT
-  Modest Stream Processing
-  Service integration

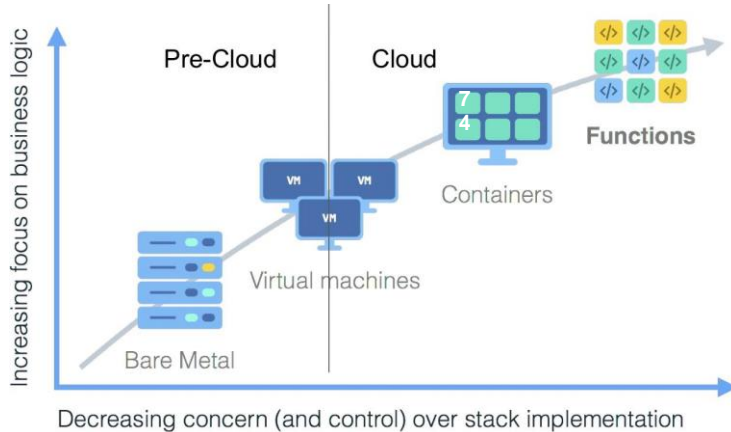
not good for
long-running
stateful
number crunching



-  Databases
-  Deep Learning Training
-  Heavy-Duty Stream Analytics
-  Numerical Simulation
-  Video Streaming

Serverless

- **Focus: Business logic vs. Tech stack**



Cloud Functions

- Create single-purpose functions that respond to events without a server or runtime
 - Event examples: New instance created, file added to Cloud Storage
- Written in Javascript; execute in managed Node.js environment on Google Cloud Platform



Big Data and Machine Learning in the Cloud

GCP Fundamentals: Core Infrastructure



Agenda

Google Cloud Big Data Platform

Google Cloud Machine Learning Platform

Quiz and Lab

Google Cloud's big data services are fully managed and scalable



Cloud
Dataproc

Managed
Hadoop
MapReduce,
Spark, Pig, and
Hive service



Cloud
Dataflow

Stream and
batch
processing;
unified and
simplified
pipelines



BigQuery

Analytics
database;
stream data at
100,000
rows per second



Cloud
Pub/Sub

Scalable and
flexible
enterprise
messaging



Cloud
Datalab

Interactive data
exploration

Cloud Dataproc is managed Hadoop

- Fast, easy, managed way to run Hadoop and Spark/Hive/Pig on GCP
- Create clusters in 90 seconds or less on average.
- Scale clusters up and down even when jobs are running.



Why use Cloud Dataproc?

- Easily migrate on-premises Hadoop jobs to the cloud.
- Quickly analyze data (like log data) stored in Cloud Storage; create a cluster in 90 seconds or less on average, and then delete it immediately.
- Use Spark/Spark SQL to quickly perform data mining and analysis.
- Use Spark Machine Learning Libraries (MLlib) to run classification algorithms.

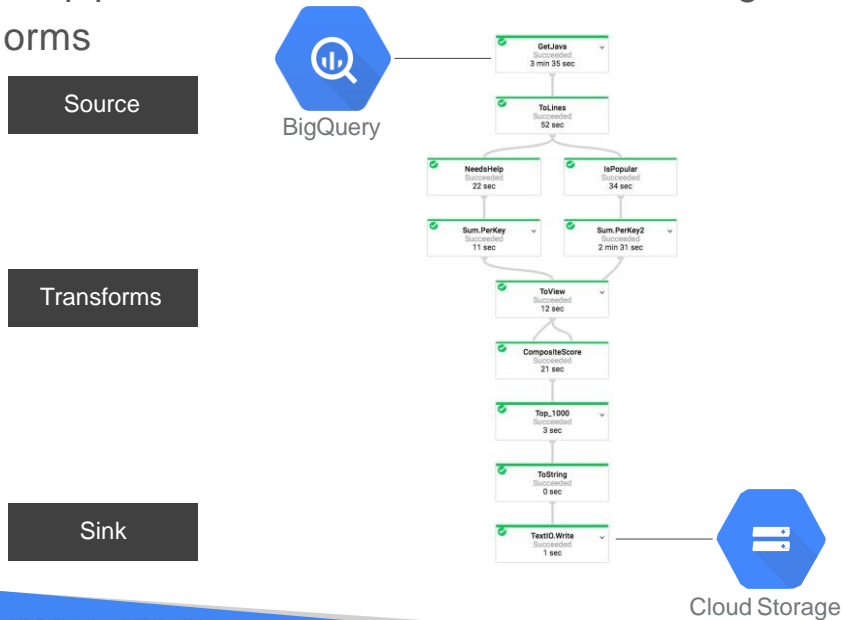


Cloud Dataflow offers managed datapipelines

- Processes data using Compute Engine instances.
 - Clusters are sized for you
 - Automated scaling, no instance provisioning required
- Write code once and get *batch and streaming*.
 - Transform-based programming model



Dataflow pipelines flow data from a source through transforms



Why use Cloud Dataflow?

- *ETL* (extract/transform/load) pipelines to move, filter, enrich, shape data
- *Data analysis*: batch computation or continuous computation using streaming
- *Orchestration*: create pipelines that coordinate services, including external services
- Integrates with GCP services like Cloud Storage, Cloud Pub/Sub, BigQuery, and Bigtable
 - Open source [Java](#) and [Python](#) SDKs



BigQuery is a fully managed data warehouse

- Provides near real-time interactive analysis of massive datasets (hundreds of TBs)
- Query using SQL syntax (SQL2011)
- No cluster maintenance is required.



BigQuery runs on Google's high-performance infrastructure

- Compute and storage are separated with a terabit network in between
- You only pay for storage and processing used
- Automatic discount for long-term data storage



Cloud Pub/Sub is scalable, reliable messaging

- Supports many-to-many asynchronous messaging
 - Application components make push/pull subscriptions to topics
- Includes support for offline consumers
- Based on proven Google technologies
- Integrates with Cloud Dataflow for data processing pipelines



Why use Cloud Pub/Sub?

- Building block for data ingestion in Dataflow, Internet of Things (IoT), Marketing Analytics
- Foundation for Dataflow streaming
- Push notifications for cloud-based applications
- Connect applications across Google Cloud Platform (push/pull between Compute Engine and App Engine)



Cloud Datalab offers interactive data exploration

- Interactive tool for large-scale data exploration, transformation, analysis, and visualization
- Integrated, open source
 - Built on Jupyter (formerly IPython)



Why use Cloud Datalab?

- Create and manage code, documentation, results, and visualizations in intuitive notebook format.
 - Use Google Charts or matplotlib for easy visualizations.
- Analyze data in BigQuery, Compute Engine, and Cloud Storage using Python, SQL, and JavaScript.
- Easily deploy models to BigQuery.



Agenda

Google Cloud Big Data Platform

Google Cloud Machine Learning Platform

Quiz and Lab

Machine Learning APIs enable apps that see, hear, and understand



Cloud Machine Learning Platform



Open source tool to build and run neural network models

- Wide platform support: CPU or GPU; mobile, server, or cloud



Cloud ML

Fully managed machine learning service

- Familiar notebook-based developer experience
- Optimized for Google infrastructure; integrates with BigQuery and Cloud Storage



Machine Learning APIs

Pre-trained machine learning models built by Google

- Speech: Stream results in real time, detects 80 languages
- Vision: Identify objects, landmarks, text, and content
- Translate: Language translation including detection
- Natural language: Structure, meaning of text

Why use the Cloud Machine Learning platform?

For structured data



Classification and regression



Recommendation



Anomaly detection

For unstructured data



Image and video analytics



Text analytics

Cloud Vision API

- Analyze images with a simple REST API
 - Logo detection, label detection, etc
- With the Cloud Vision API, you can:
 - Gain insight from images
 - Detect inappropriate content
 - Analyze sentiment
 - Extract text



Cloud Speech API

- Recognizes over 80 languages and variants
- Can return text in realtime
- Highly accurate, even in noisy environments
- Access from any device
- Powered by Google's machine learning



Cloud Natural Language API

- Uses machine learning models to reveal structure and meaning of text.
- Extract information about items mentioned in text documents, news articles, and blog posts.
- Analyze text uploaded in request or integrate with Cloud Storage.



Cloud Translation API

- Translate arbitrary strings between thousands of language pairs
- Programmatically detect a document's language
- Support for dozens of languages



Cloud Video Intelligence API

- Annotate the contents of videos
- Detect scene changes
- Flag inappropriate content
- Support for a variety of video formats

