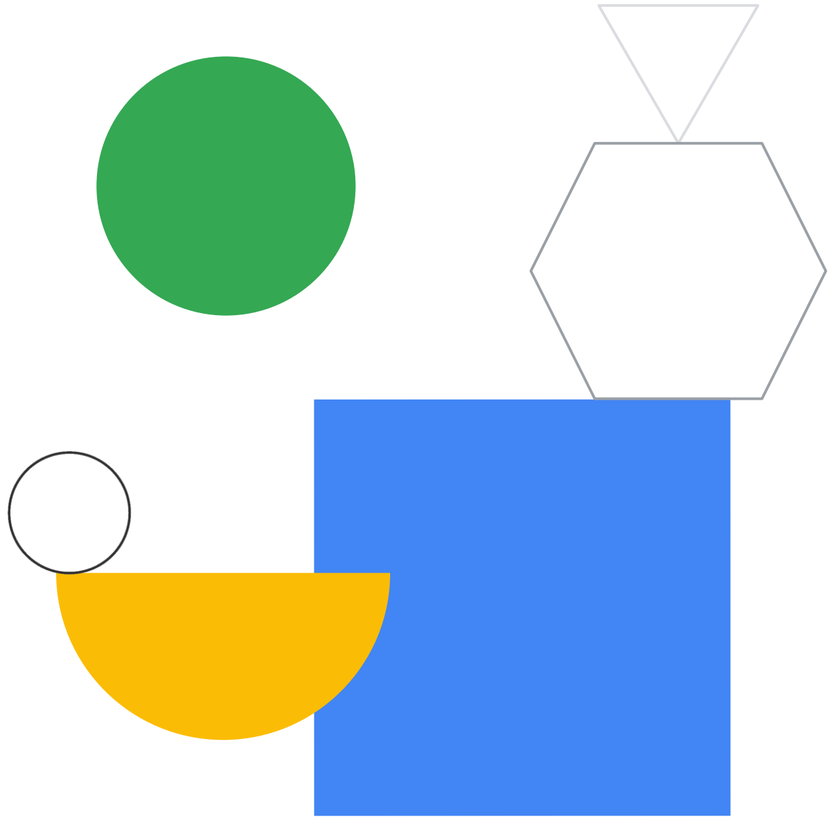
Google Cloud

Introduction to Cloud Architecture



Module 1 - Cloud basics
Google Cloud Computing Foundations

Objectives

01 Explore cloud computing.

02 Compare and contrast physical, virtual, and cloud architectures.

03 Differentiate IaaS, PaaS, and SaaS.

04 Be introduced to Google Cloud compute, storage, big data, and ML services.

05 Examine the Google network and how it powers cloud computing.

Agenda

01 Cloud computing

02 Cloud vs. traditional architecture

03 IaaS, PaaS, and SaaS

04 Google Cloud architecture

05 Quiz

06 Summary

Agenda

What is cloud computing?

US National Institute of Standards and Technology

Cloud computing

Cloud computing is a way of using information technology that has these five equally important traits.

01 Customers get computing resources that are on-demand and self-service

02 Customers get access to those resources over the internet, from anywhere

03 The provider of those resources allocates them to users out of that pool

04 The resources are elastic–which means they can increase or decrease as needed

05 Customers pay only for what they use, or reserve as they go

Cloud computing is a way of using information technology that has these five equally important traits.

On-demand self-service

01 Customers get computing resources that are on-demand and self-service

02 Customers get access to those resources over the internet, from anywhere

03 The provider of those resources allocates them to users out of that pool

04 The resources are elastic–which means they can increase or decrease as needed

05 Customers pay only for what they use, or reserve as they go

Cloud computing is a way of using information technology that has these five equally important traits.

Broad network access

01 Customers get computing resources that are on-demand and self-service

02 Customers get access to those resources over the internet, from anywhere

03 The provider of those resources allocates them to users out of that pool

04 The resources are elastic–which means they can increase or decrease as needed

05 Customers pay only for what they use, or reserve as they go

Cloud computing is a way of using information technology that has these five equally important traits.

01 Customers get computing resources that are on-demand and self-service

02 Customers get access to those resources over the internet, from anywhere

Resource pooling

03 The provider of those resources allocates them to users out of that pool

04 The resources are elastic–which means they can increase or decrease as needed

05 Customers pay only for what they use, or reserve as they go

Cloud computing is a way of using information technology that has these five equally important traits.

01 Customers get computing resources that are on-demand and self-service

02 Customers get access to those resources over the internet, from anywhere

03 The provider of those resources allocates them to users out of that pool

Rapid elasticity

04 The resources are elastic–which means they can increase or decrease as needed

05 Customers pay only for what they use, or reserve as they go

Cloud computing is a way of using information technology that has these five equally important traits.

01 Customers get computing resources that are on-demand and self-service

02 Customers get access to those resources over the internet, from anywhere

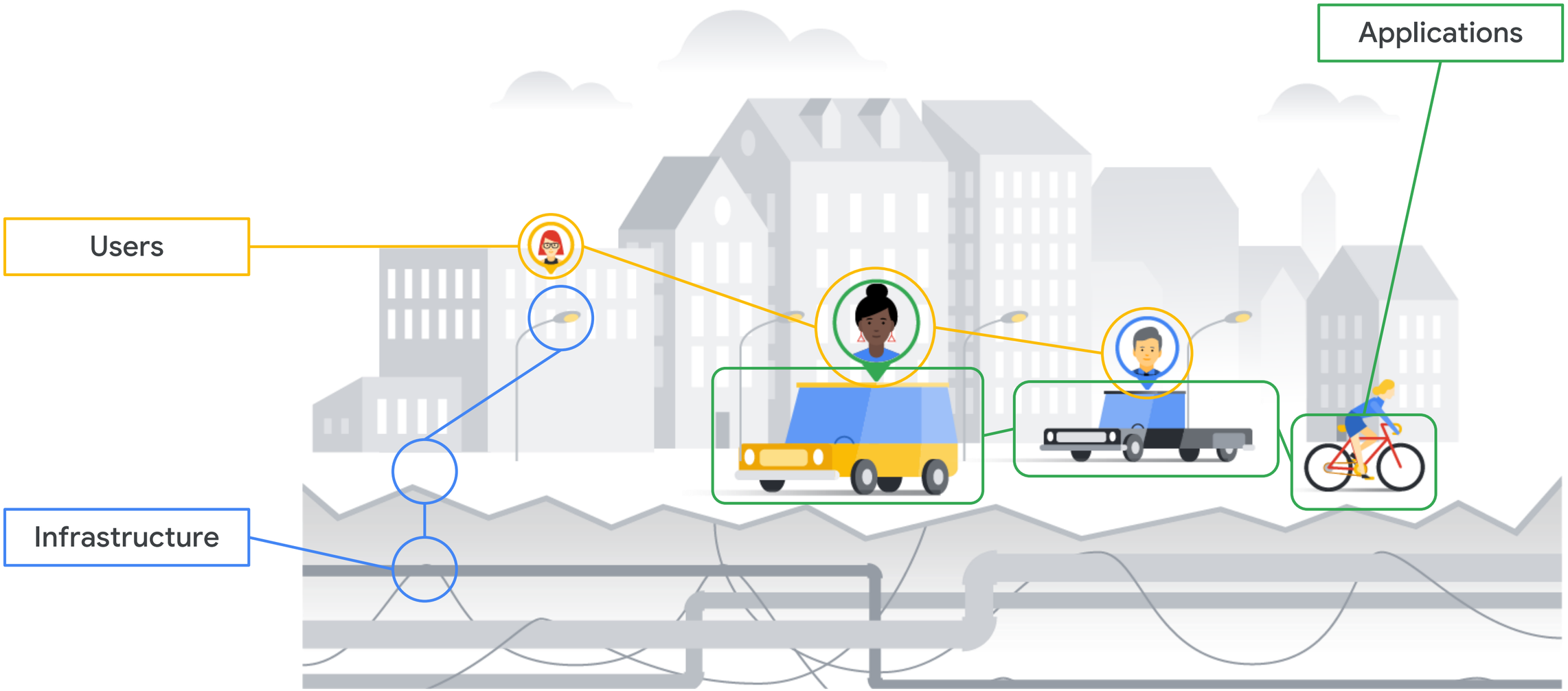
03 The provider of those resources allocates them to users out of that pool

04 The resources are elastic–which means they can increase or decrease as needed

Measured service

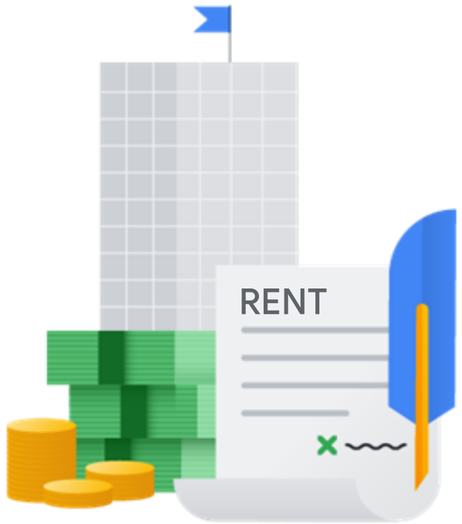
05 Customers pay only for what they use, or reserve as they go

IT infrastructure is like city infrastructure



Agenda

The history of cloud computing



First wave

Colocation

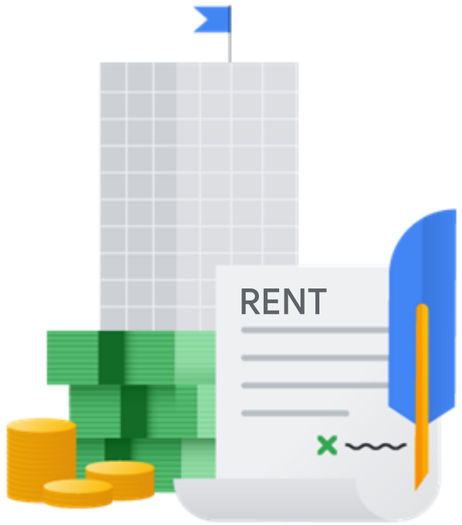
Server on-premises

You own everything.
It's yours to manage.

Data Centers

You pay for the hardware but rent the space. Still yours to manage.

The history of cloud computing



First wave

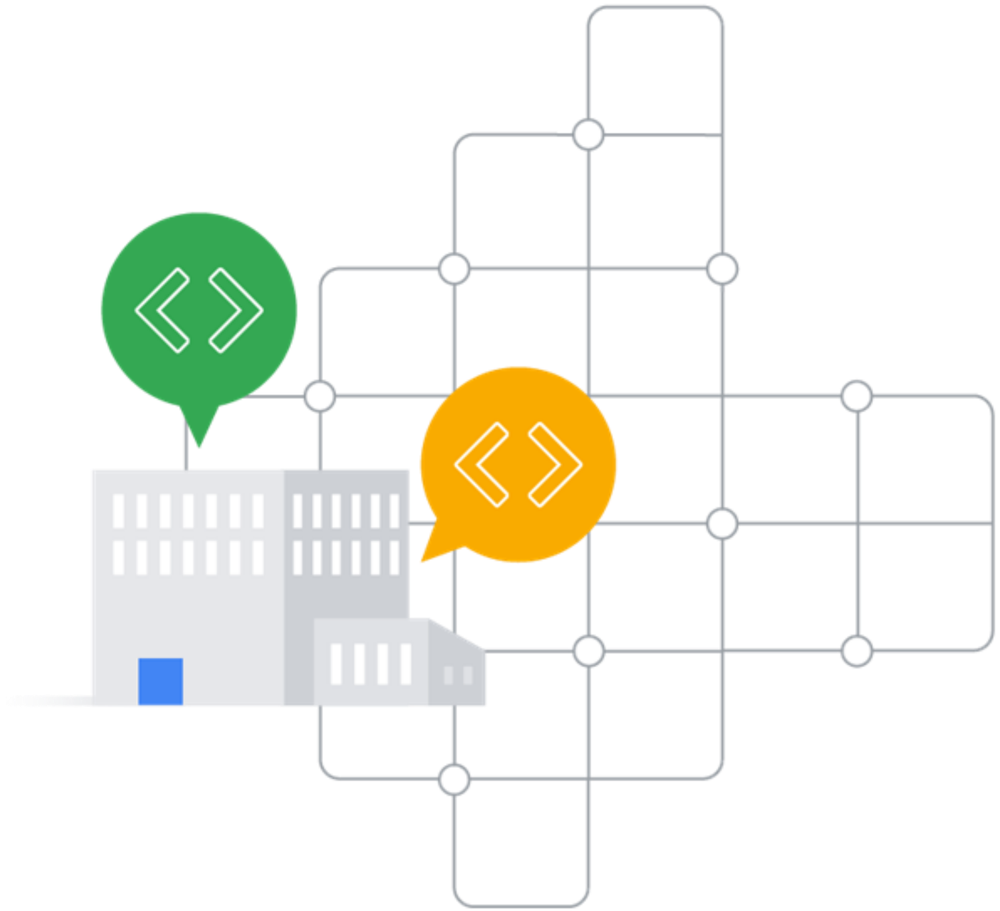
Colocation

Second wave

Virtualized data center

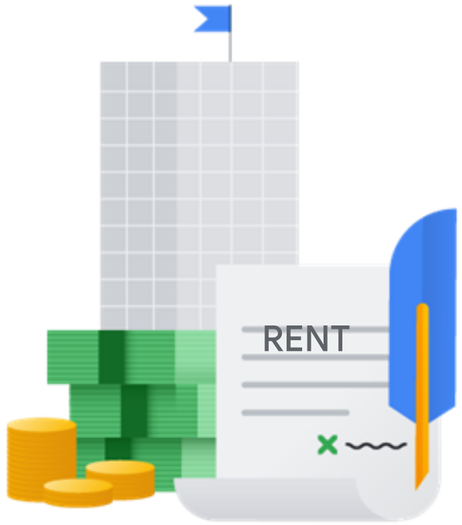
You rent hardware and space, but still control and configure virtual machines. Pay for what you provision.

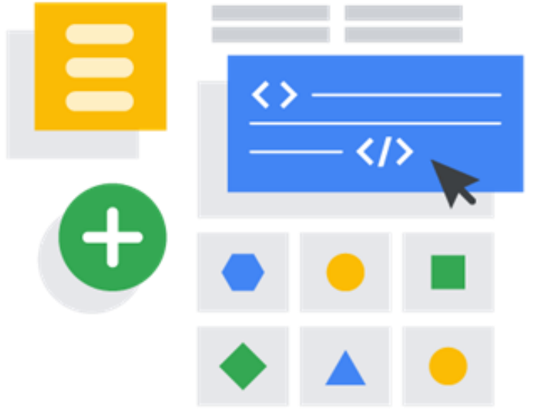
Enterprises still maintain the infrastructure



- User-controlled environment
- User-configured environment

The history of cloud computing





First wave

Colocation

Second wave

Virtualized data center

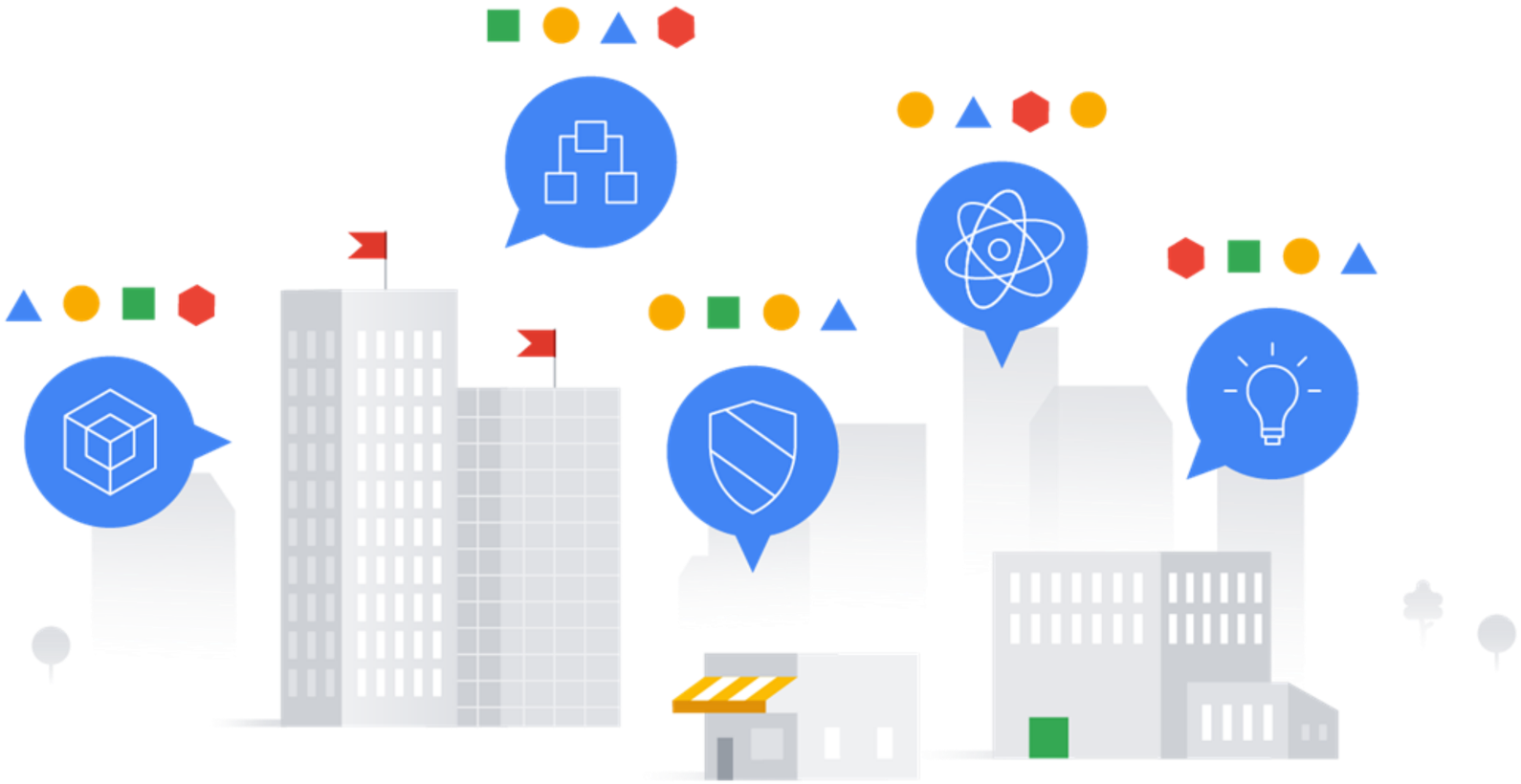
Third wave

Container-based architecture

Managed service

Completely elastic storage, processing, and machine learning so that you can invest your energy in great apps. Pay for what you use.

Third-wave cloud is available to Google customers

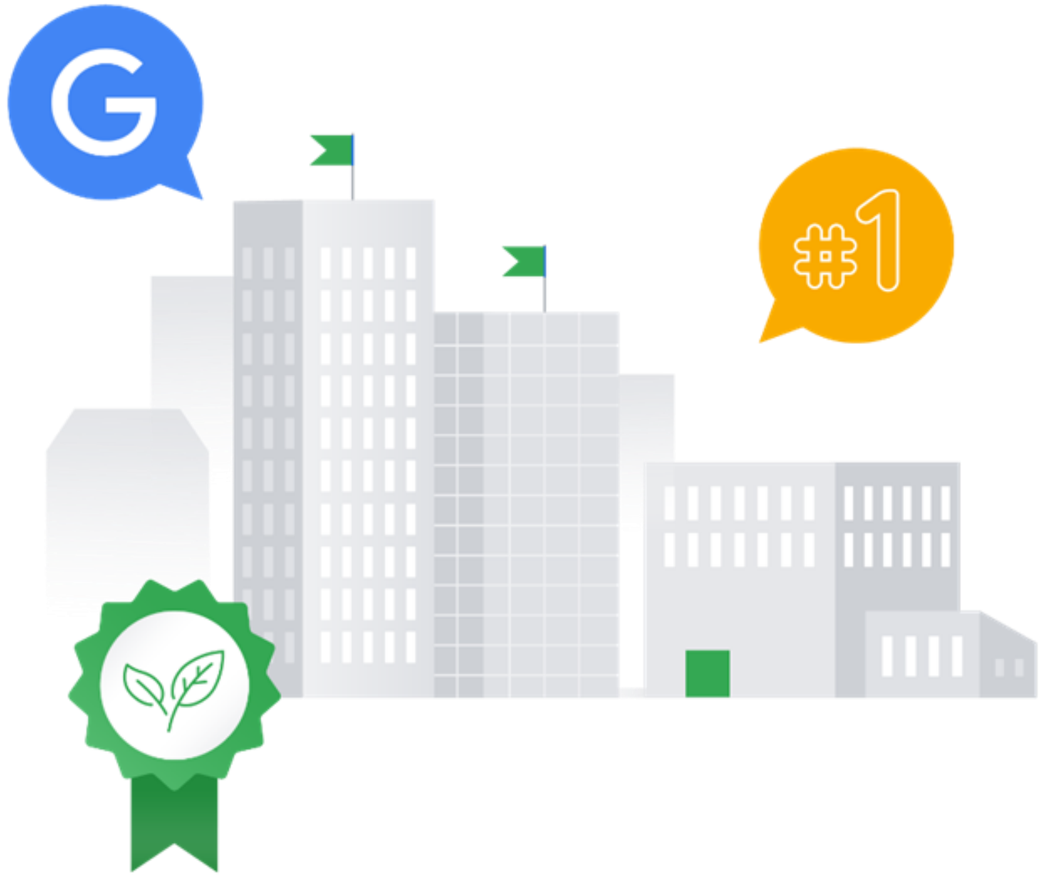


Data center energy consumption

2%

of the world's electricity

Google aims to improve efficiency and reduce waste

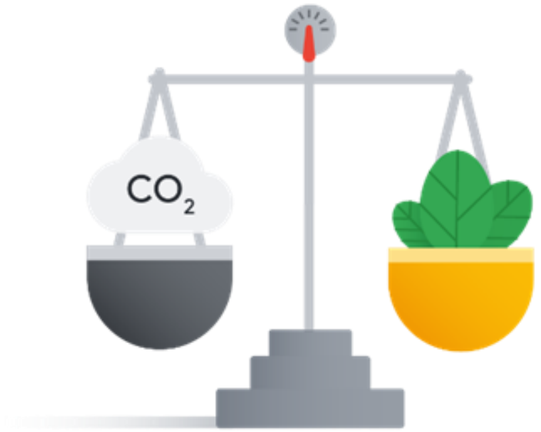


Google's data centers were the first to achieve ISO 14001 certification

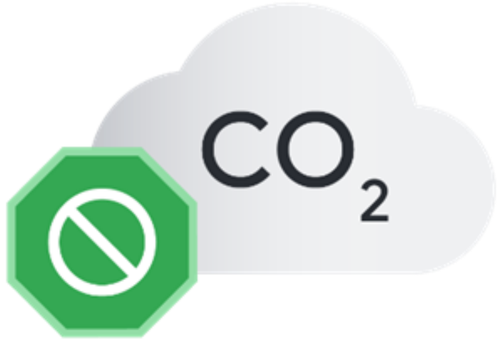
The data center cooling system in Finland is the first of its kind anywhere in the world.

Google’s data center, Hamina, Finland

Google’s commitment to sustainability



100%



Founding decade

Carbon neutral

Second decade

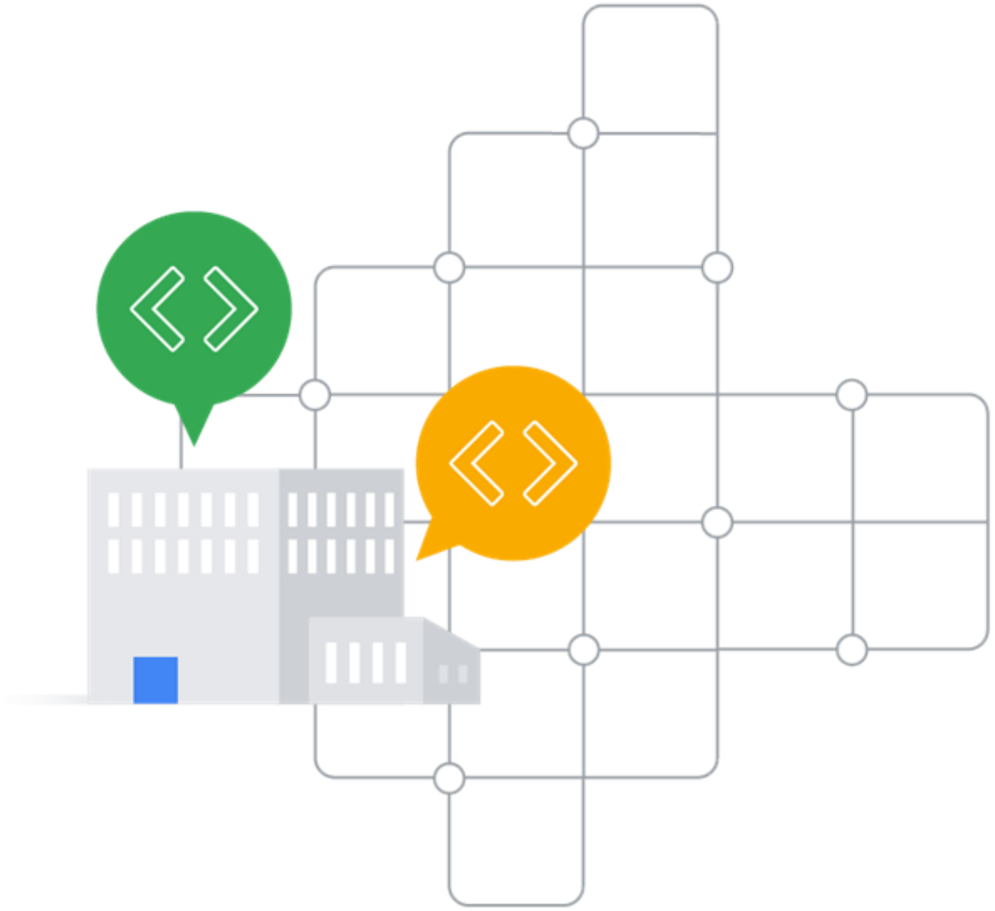
Renewable energy

2030

Carbon free

Agenda

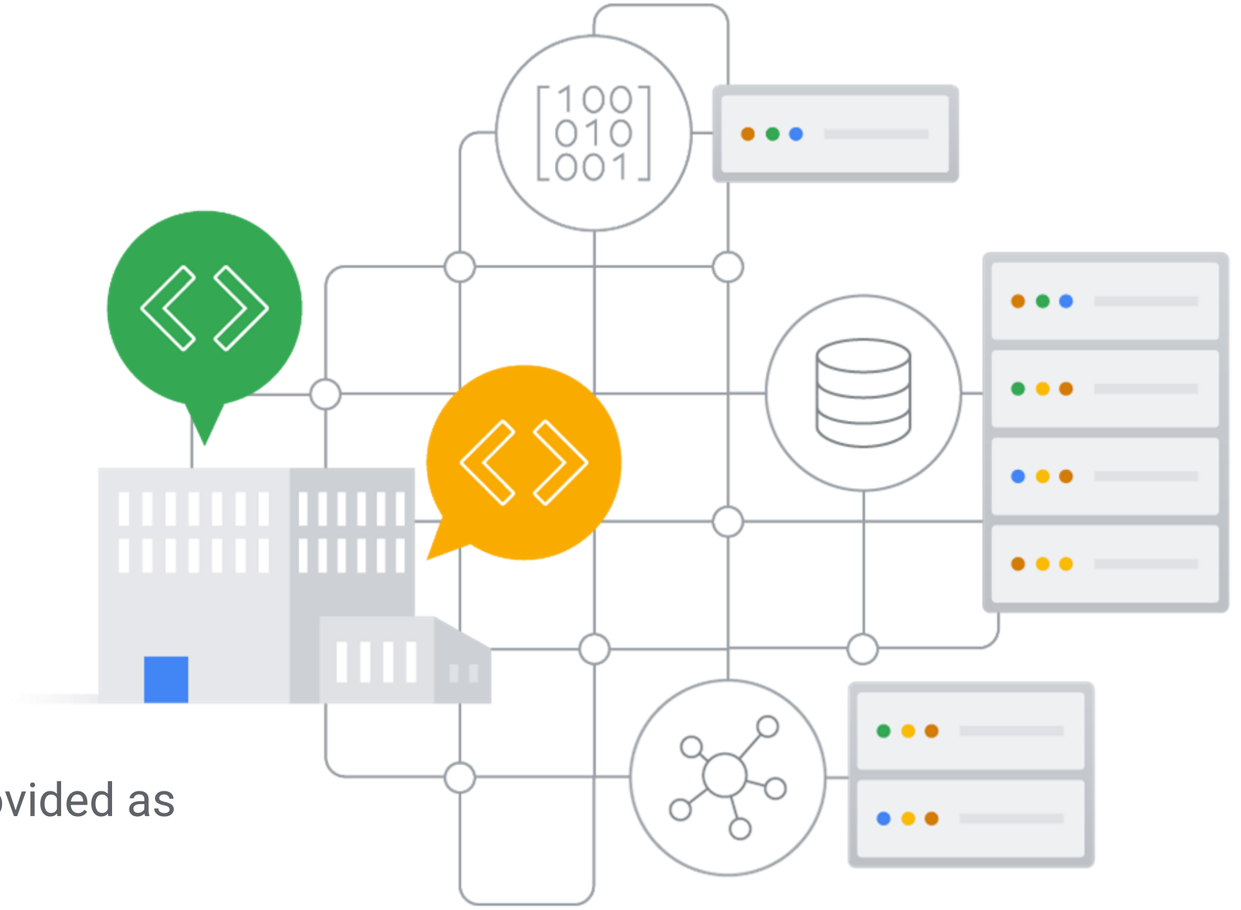
Cloud service offerings



- IaaS - Infrastructure as a service
- PaaS - Platform as a service

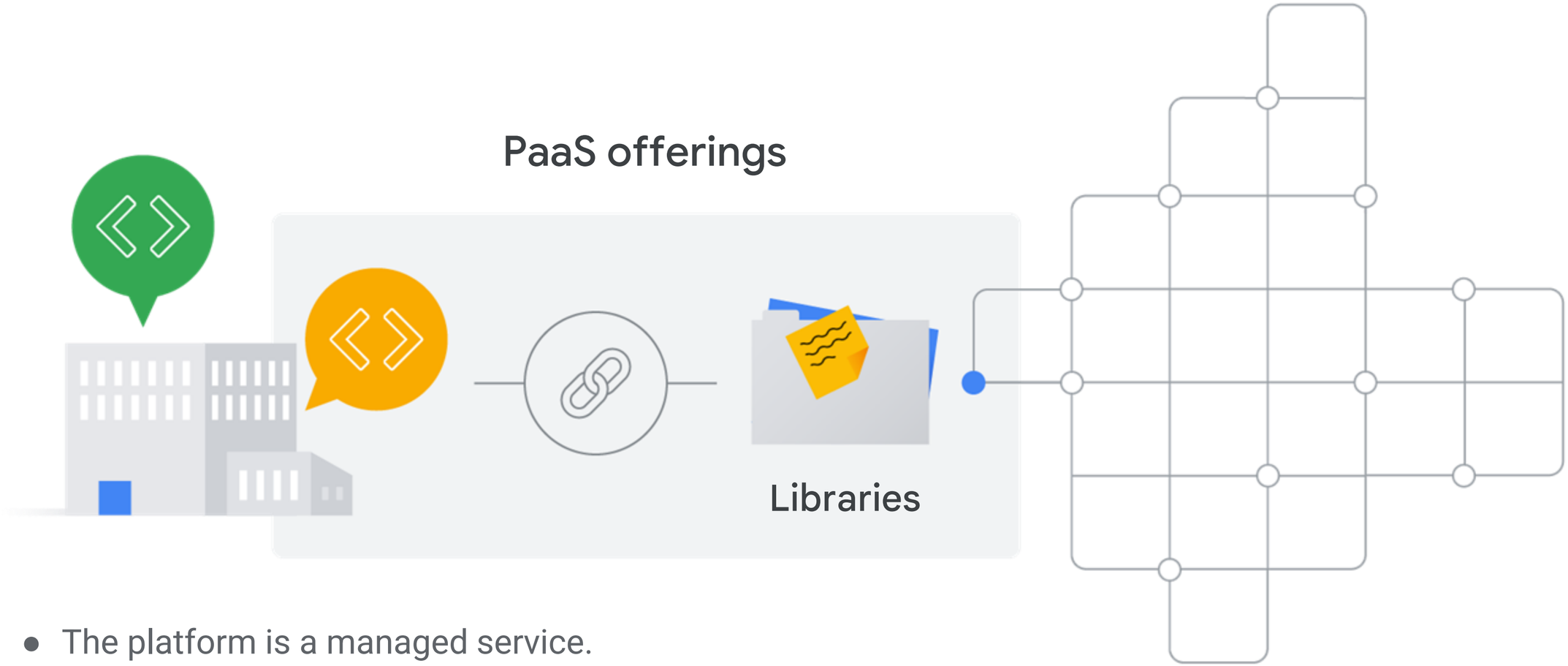
Infrastructure as a Service (IaaS)

- Raw compute
- Storage
- Network capabilities



- CPU, memory, storage, and networking are provided as a service.
- The user needs to manage the OS and the application.

Platform as a Service (PaaS)

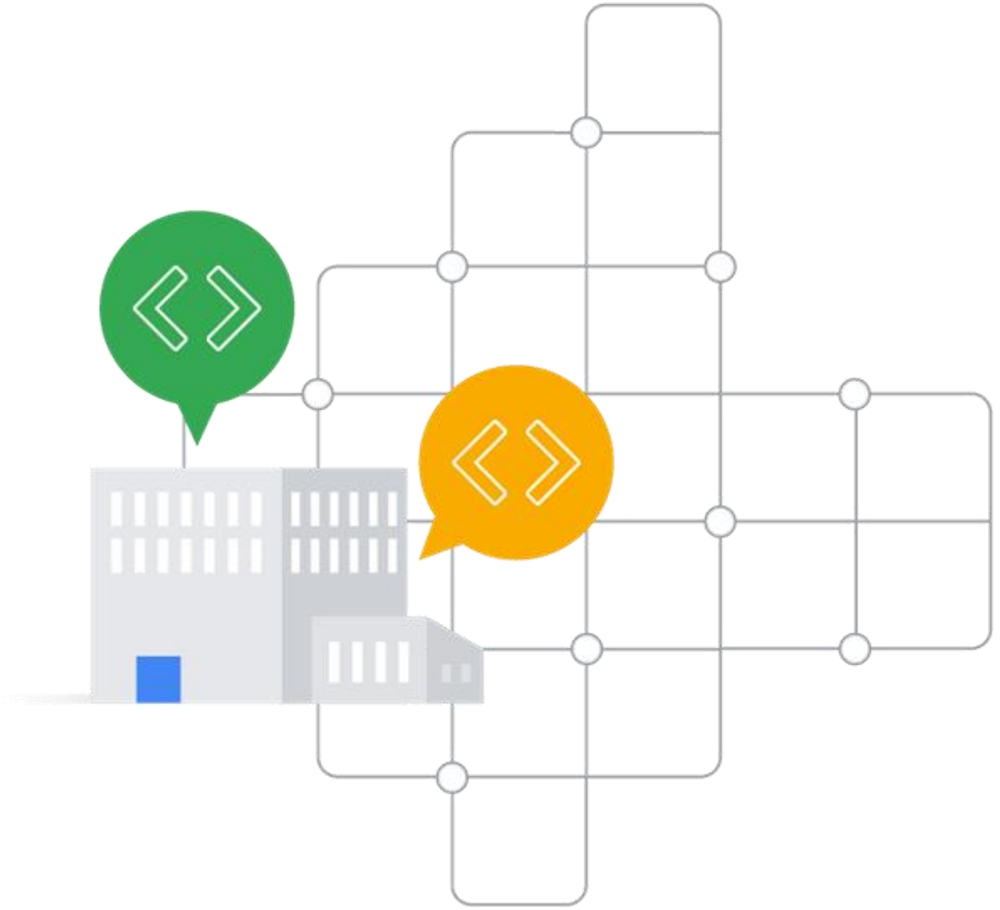


- The platform is a managed service.
- All the user provides is the application.

Payment models

IaaS

Pay for what they allocate



PaaS

Pay for what they use

The evolution of cloud computing

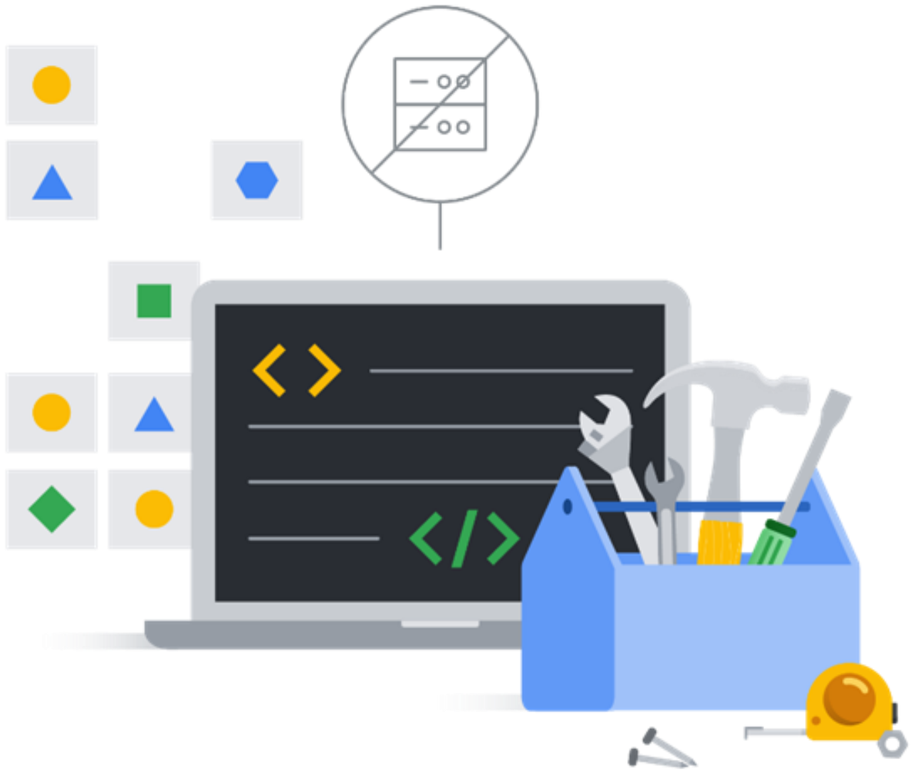
Managed resources

Managed services

Deliver products/services

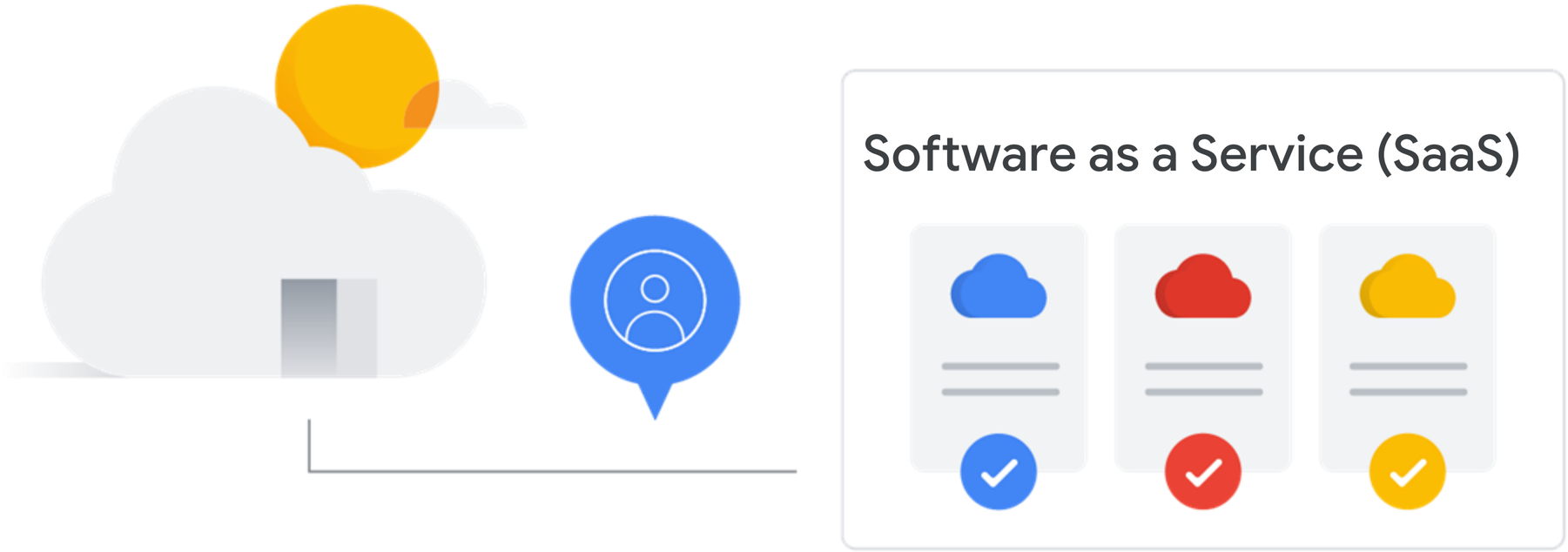
- More quickly
- More reliably

Serverless



- Allows developers to concentrate on code
- No infrastructure management needed

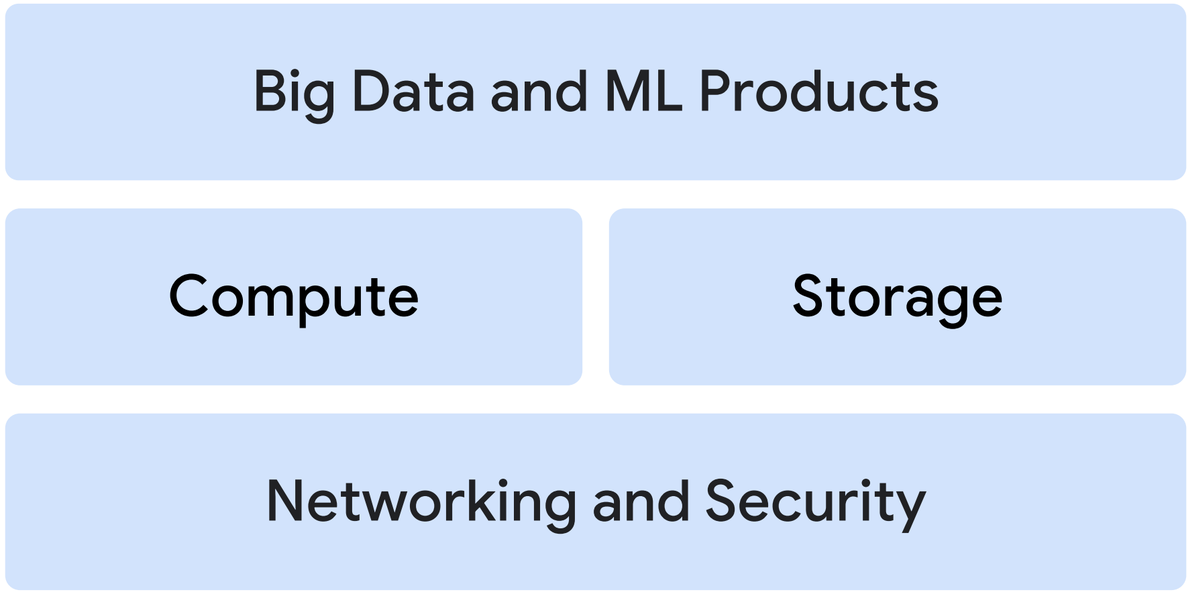
What about SaaS?



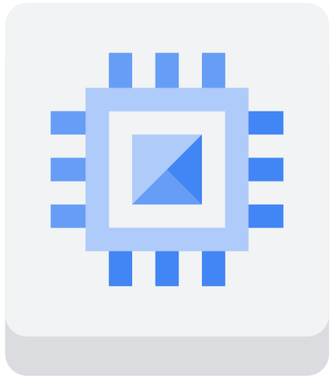
- The platform and software is provided as a service to the user.
- The user supplies the data.
- Example: Google Workspace (Gmail, Drive, Documents)

Agenda

The Google Cloud infrastructure

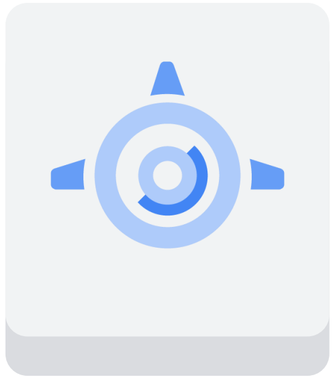


Google Cloud computing services

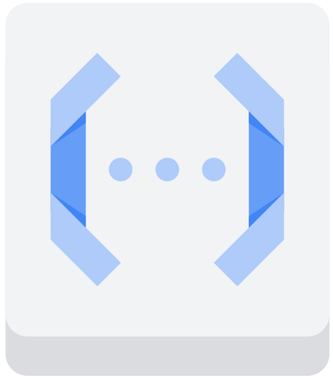


Compute Engine

Google Kubernetes Engine



App Engine



Cloud Functions

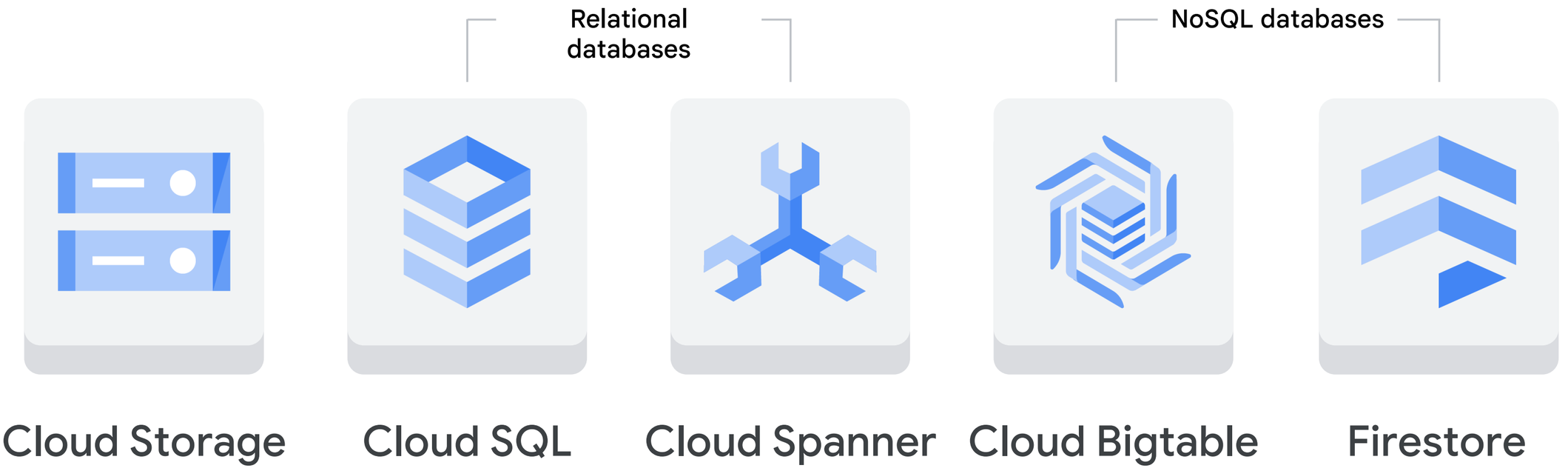
Cloud Run

IaaS	Hybrid	PaaS	Serverless logic	Automated elastic resources

*Toward **managed** infrastructure*

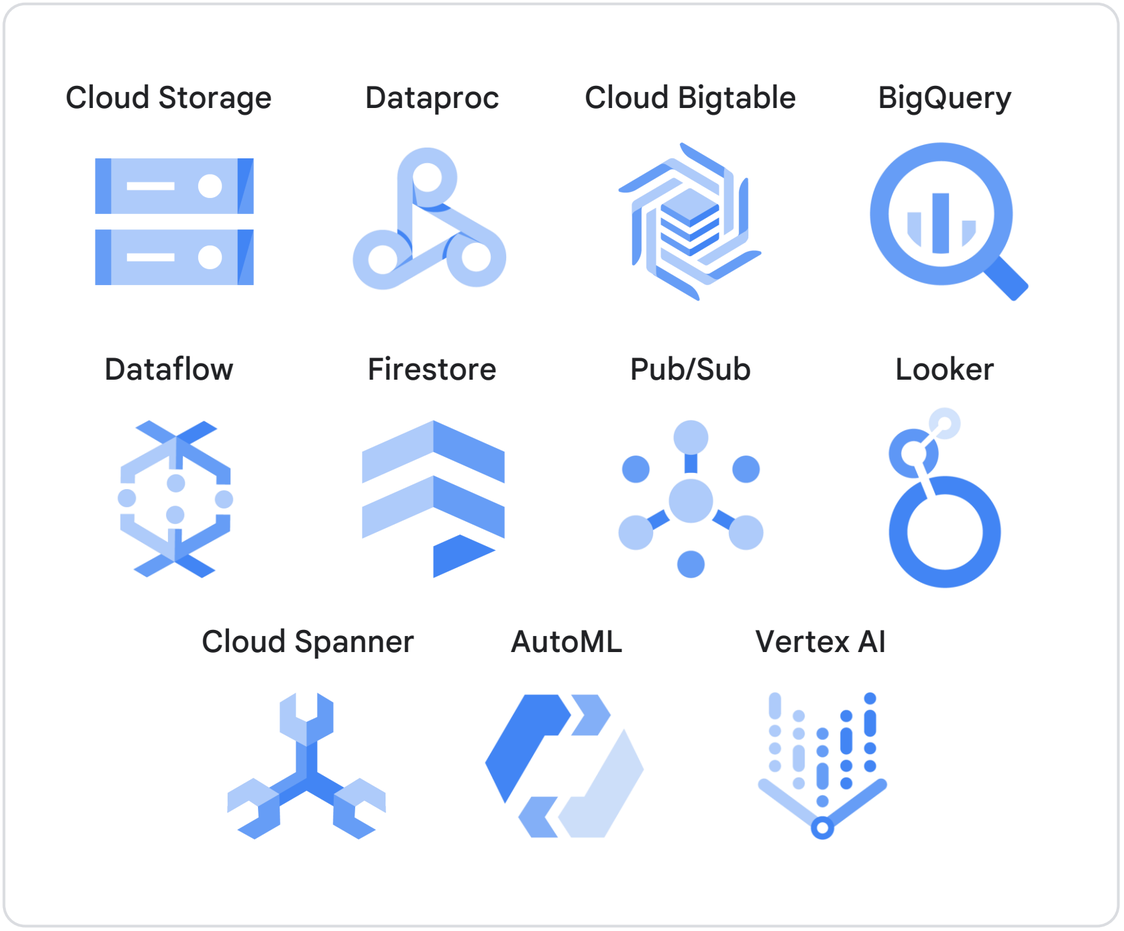
*Toward **dynamic** infrastructure*

Google Cloud storage services

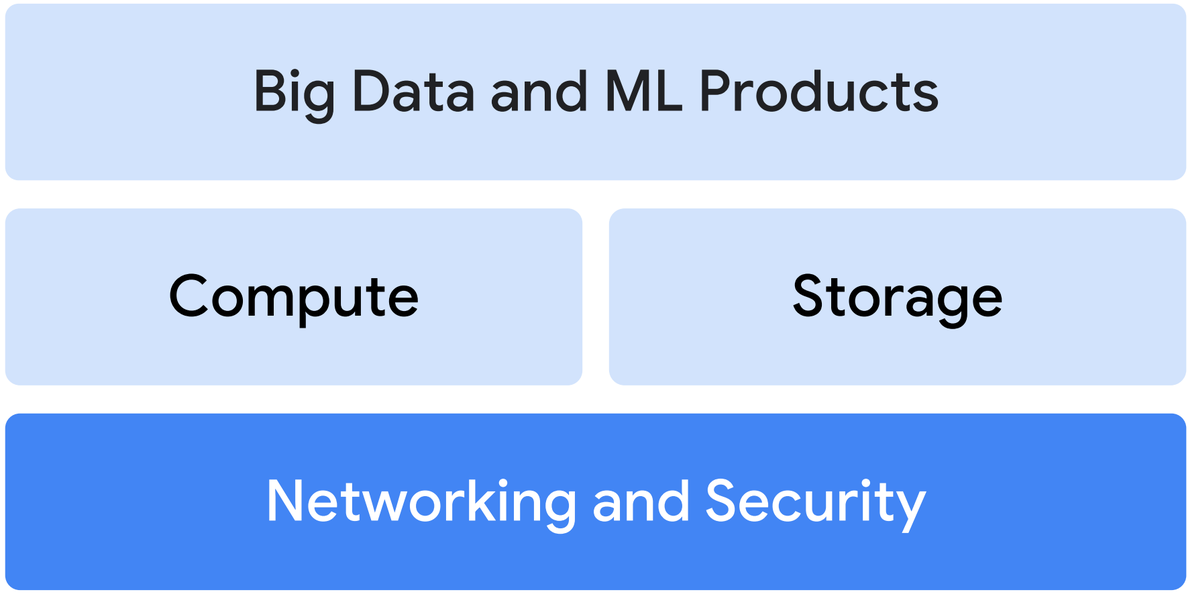


A robust big data and ML product line

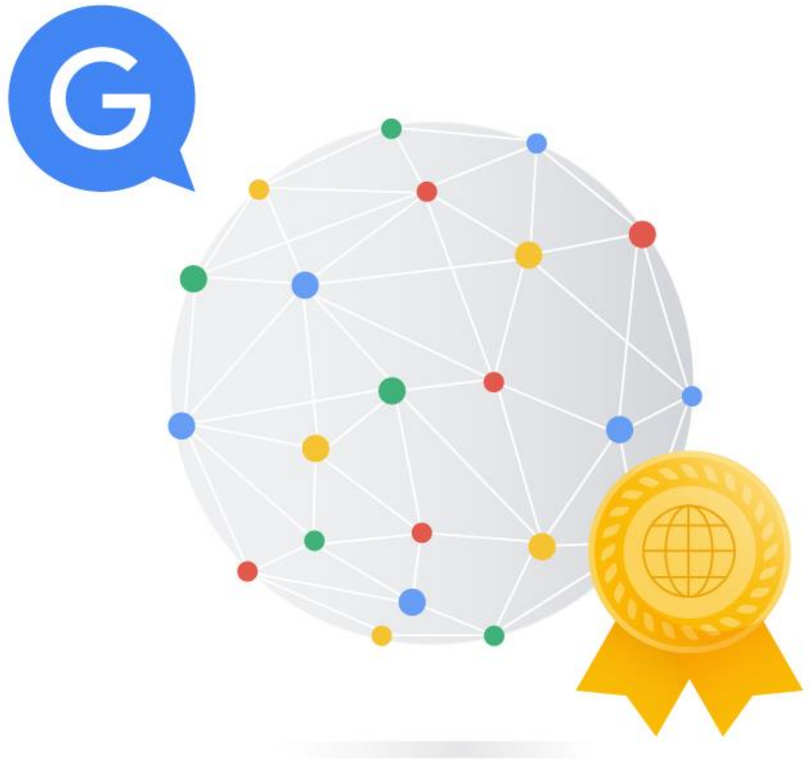
Google Cloud



The Google Cloud infrastructure



Largest network of its kind

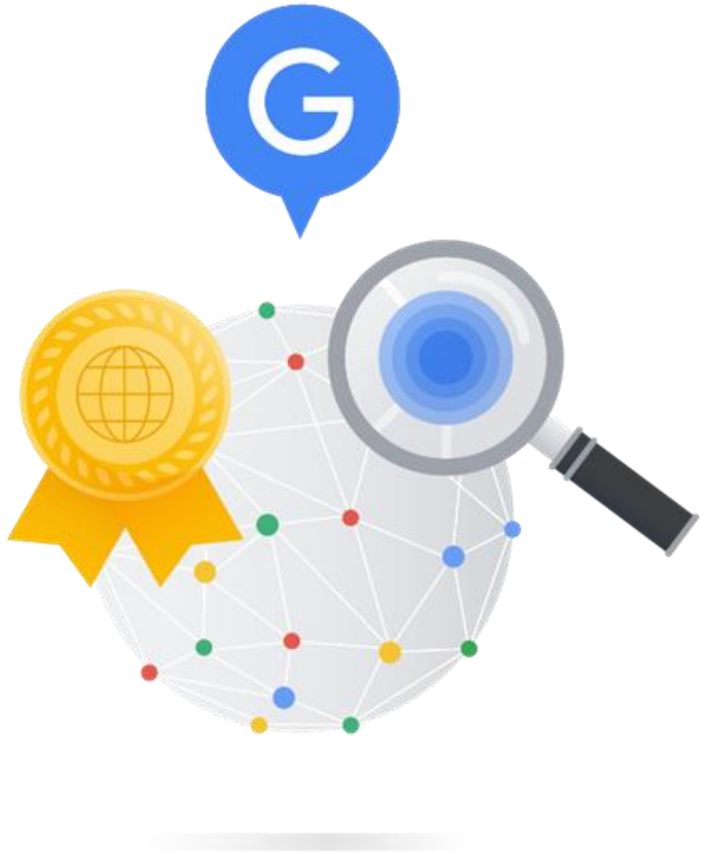


40%

of the world's internet traffic every day

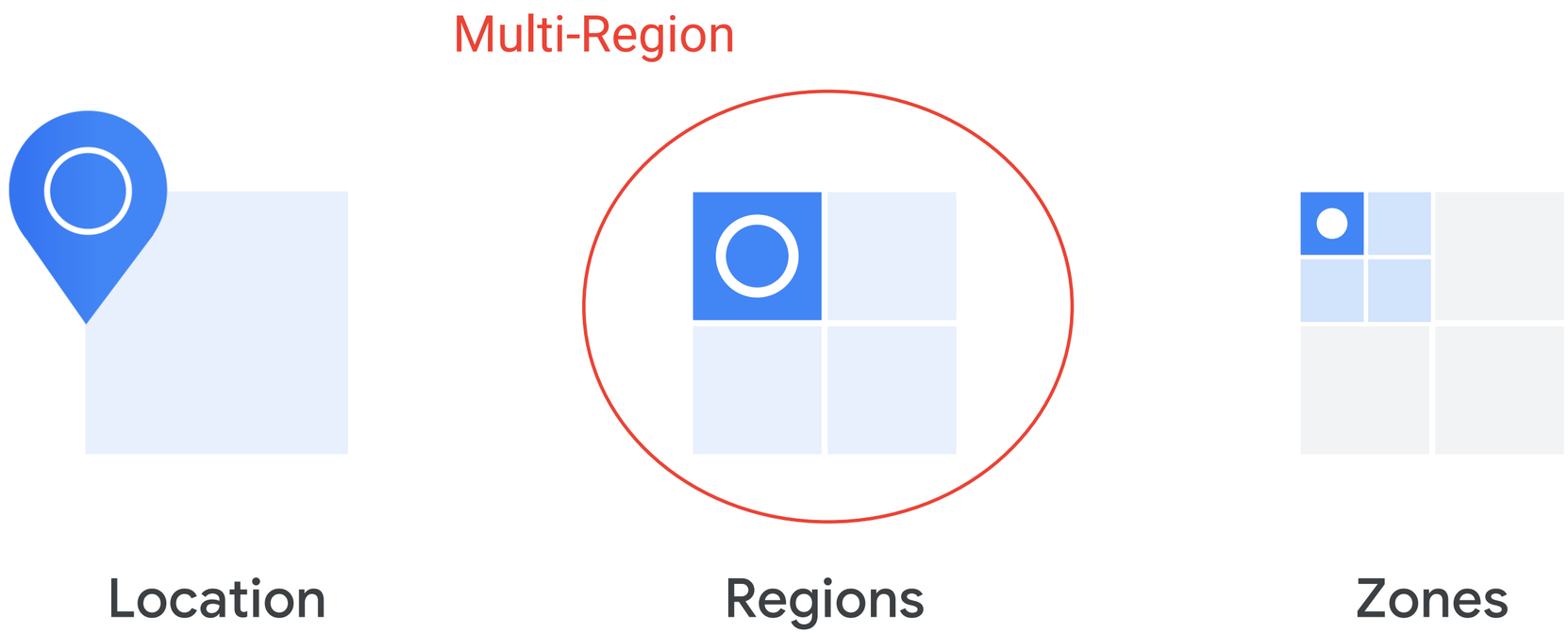
Designed for high throughput

- Highest possible throughput
- Lowest possible latencies
- 100+ content caching nodes worldwide
- High demand content is cached for quicker access

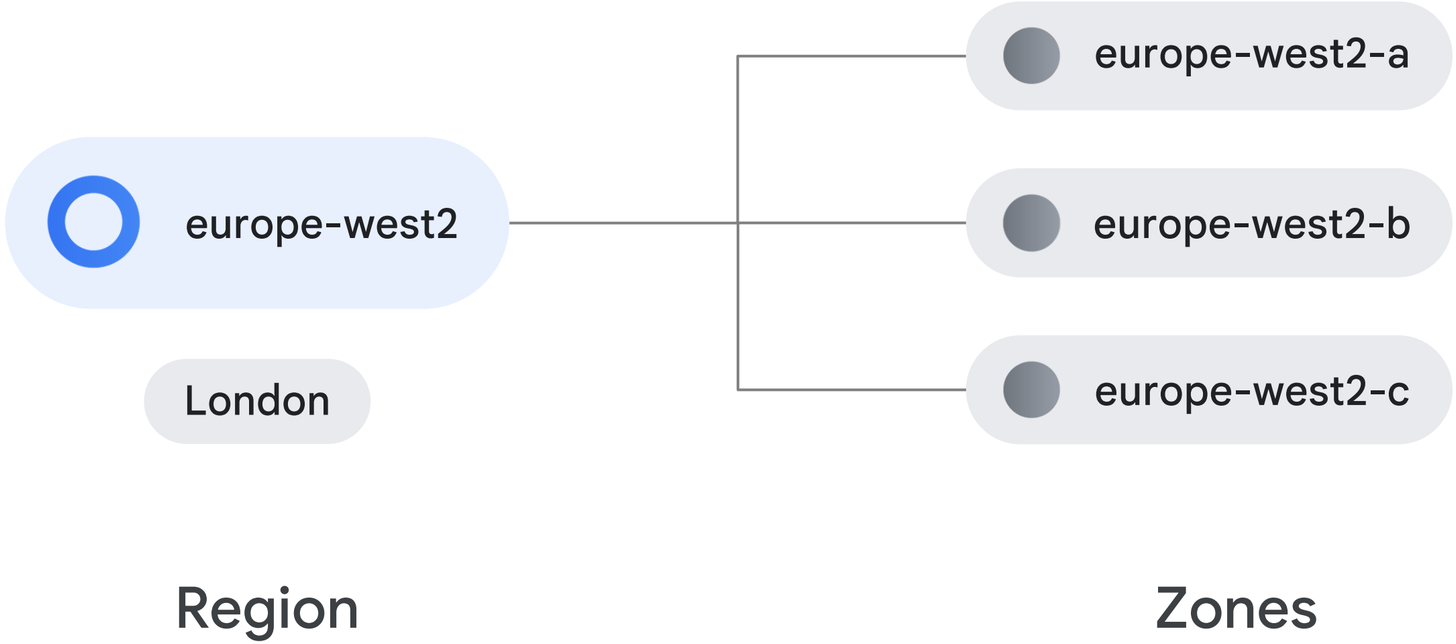


Infrastructure locations

Geographic locations contain regions and zones



Regions contain multiple zones



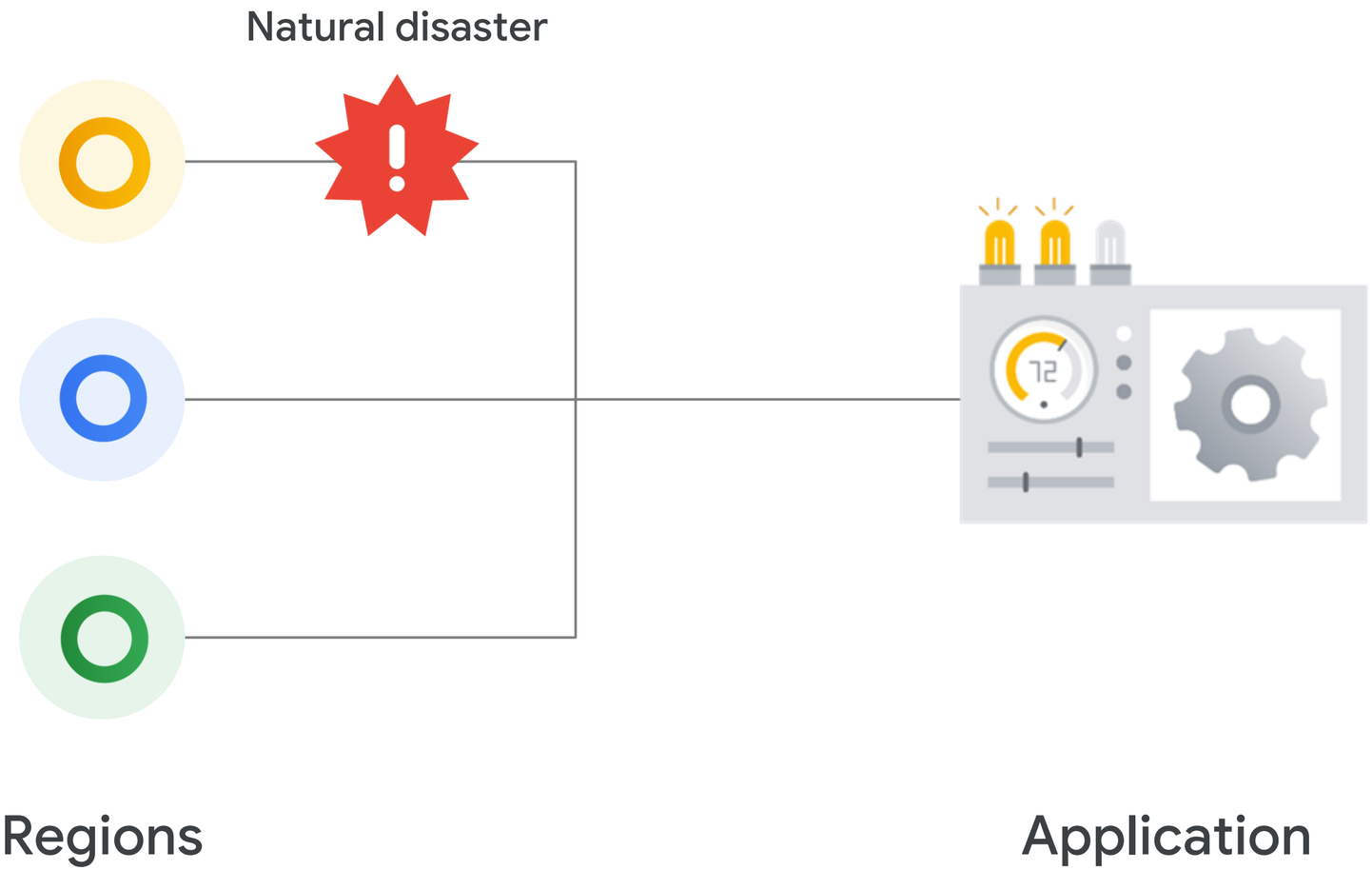
Zones are where Cloud resources are deployed

Virtual Machine

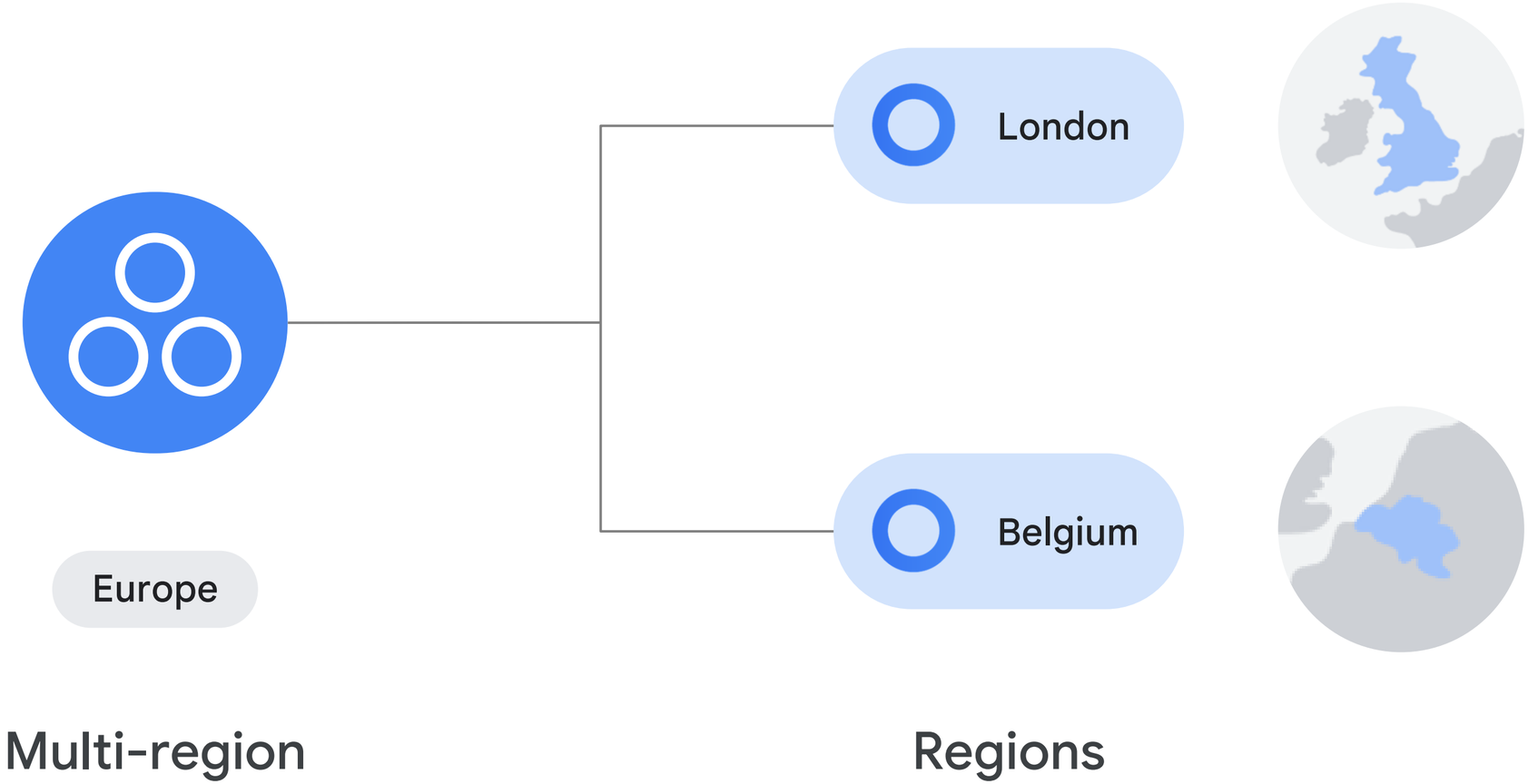
Zones

Resource redundancy

Resources can run in different regions



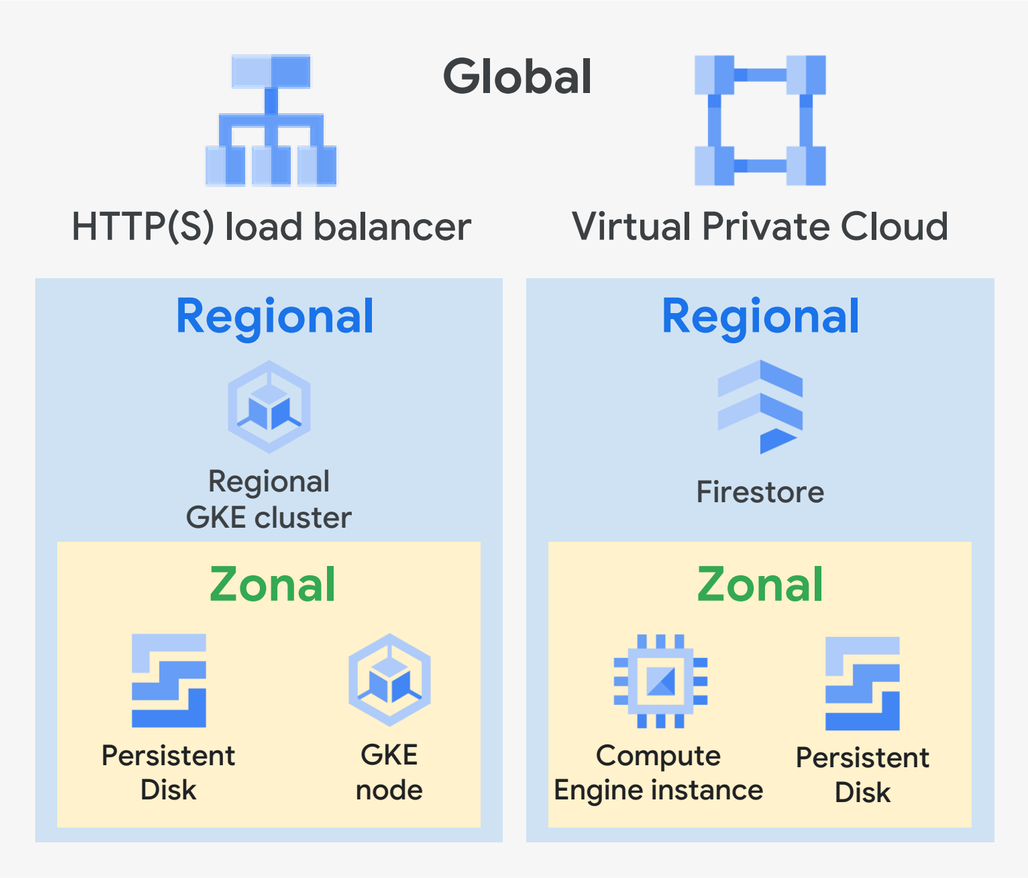
Some services can run in multiple geographic locations



88 Zones → 29 Regions

cloud.google.com/about/locations

Resources are organized both physically and logically



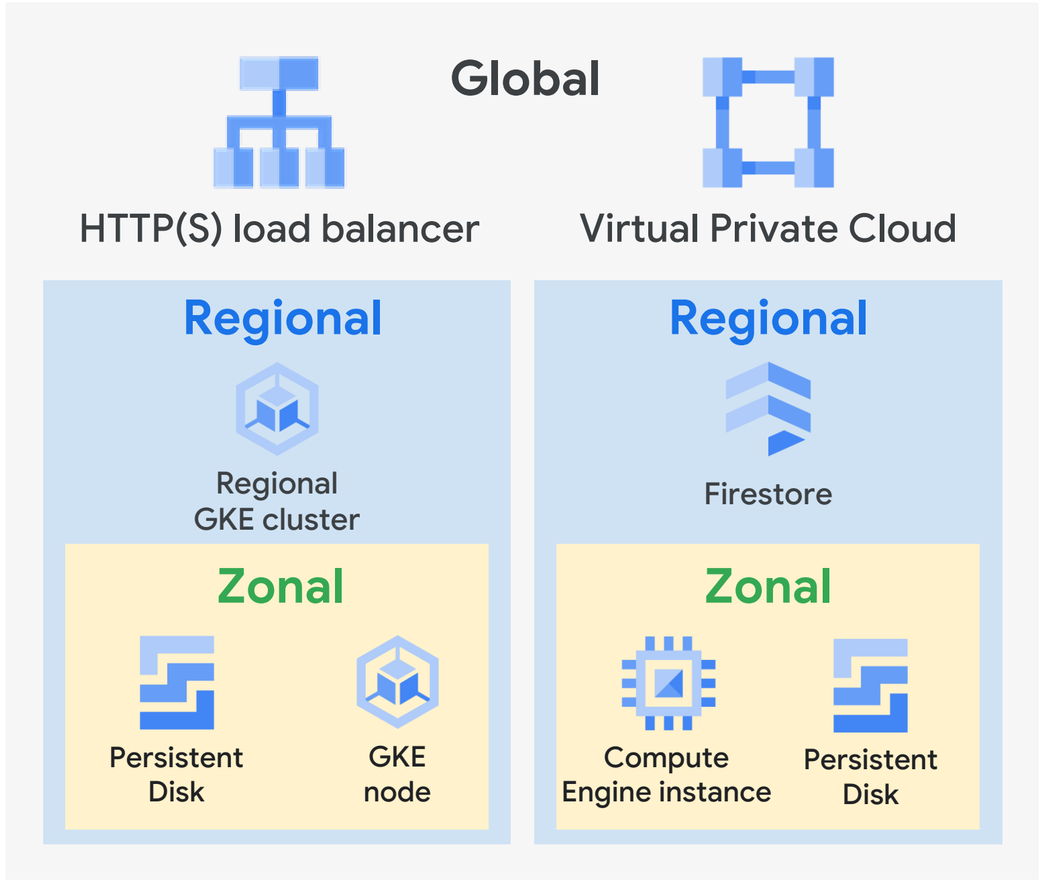
Physical organization

Resources are organized both physically and logically

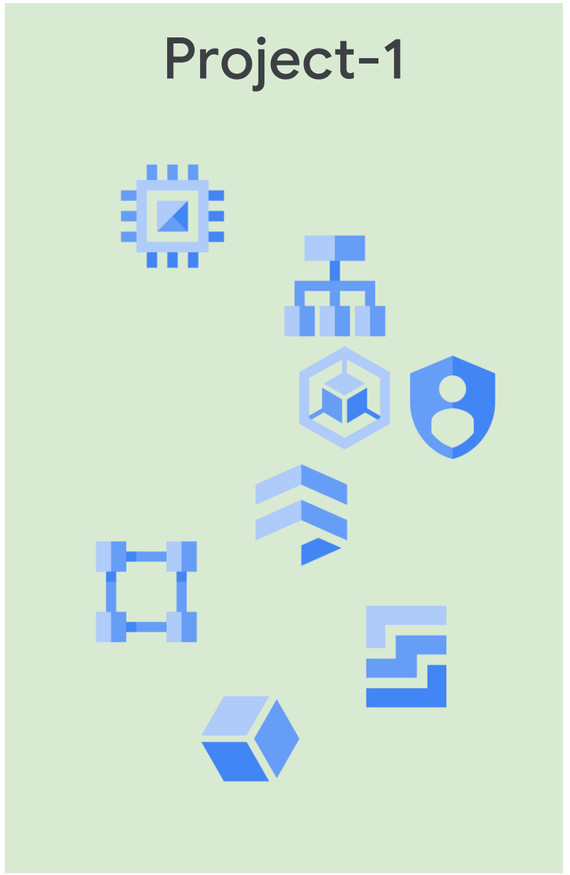
Global Resources

Regional resources

Zonal resources



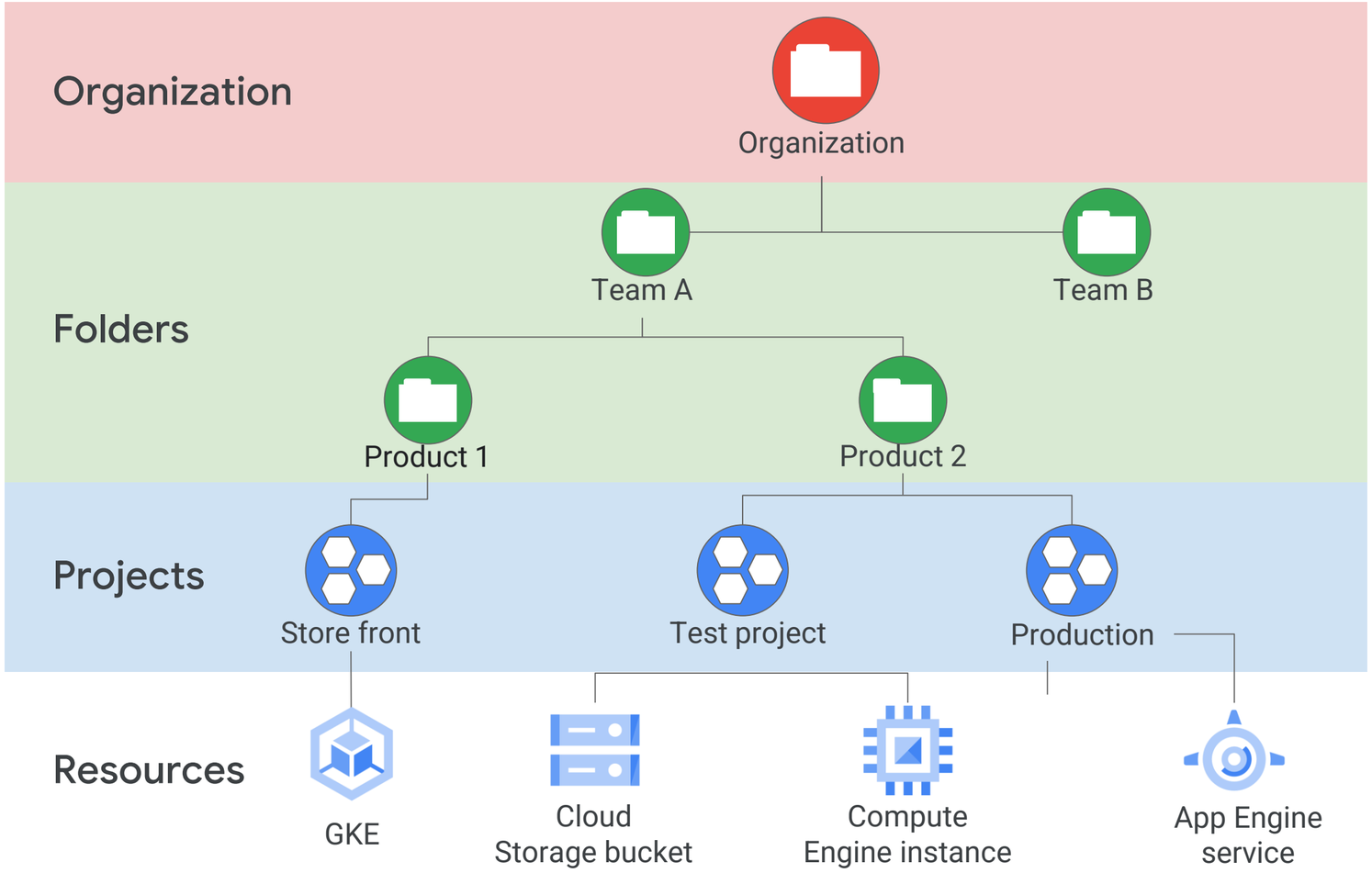
Physical organization



Logical organization

A **project** is the base-level organizing entity for creating and using resources and services and managing billing, APIs, and permissions.

Resources have hierarchy



An organization is the **root node** of a Google Cloud resource hierarchy

Folders to reflect the hierarchy of your enterprise and apply policies at the right levels in the enterprise

A **project** is the base-level organizing entity for creating and using resources and services and managing billing, APIs, and permissions.

Cloud Identity and Access Management (**Cloud IAM**) creates trust boundaries and resource isolation

Agenda

Agenda

Summary

- Explored cloud computing.
- Compared and contrasted physical, virtual, and cloud architectures.
- Differentiated IaaS, PaaS, and SaaS.
- Were introduced to Google Cloud compute, storage, big data, and ML services.
- Examined the Google network and how it powers cloud computing.