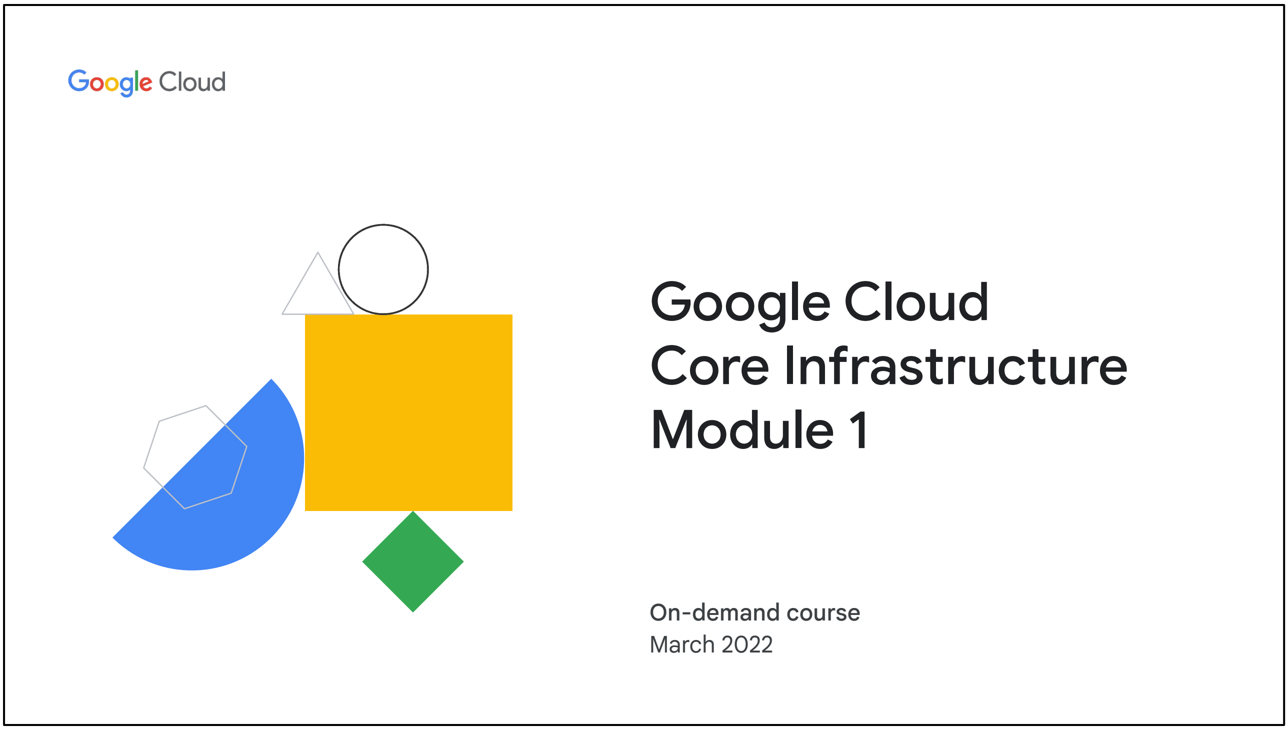
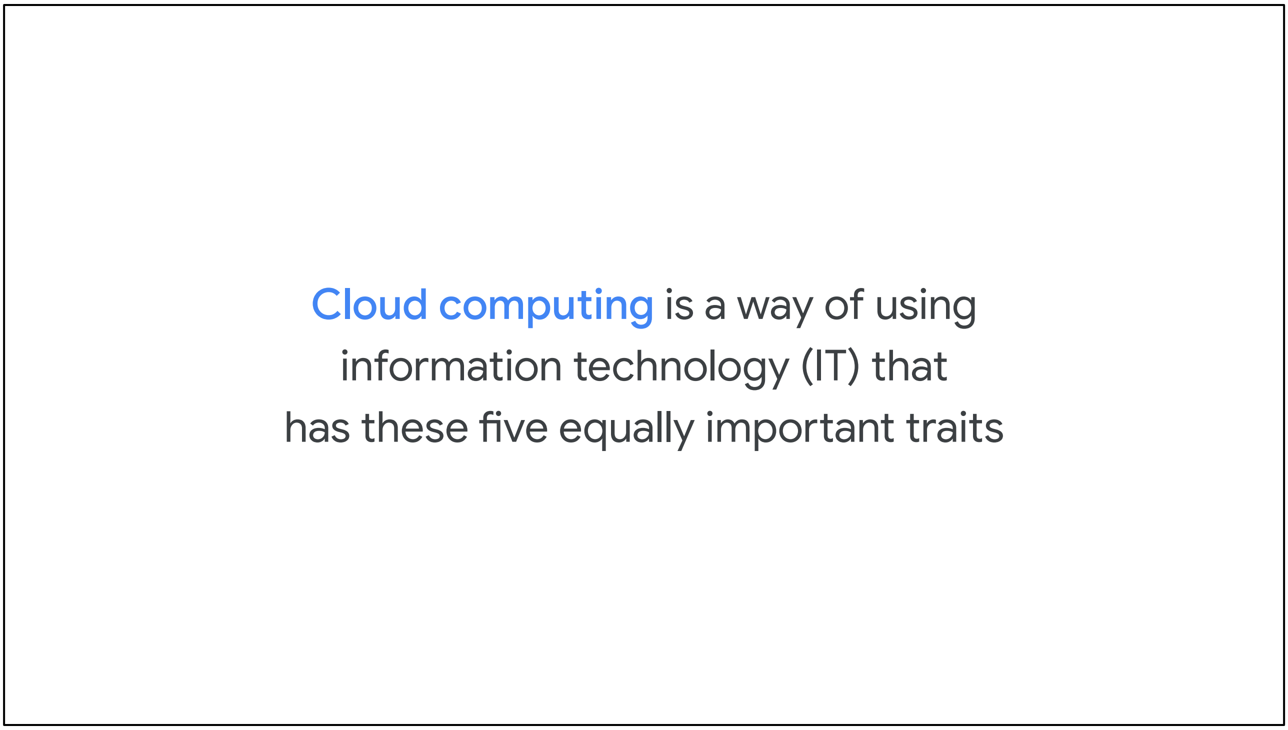


Google Cloud

Google Cloud Core Infrastructure Module 1

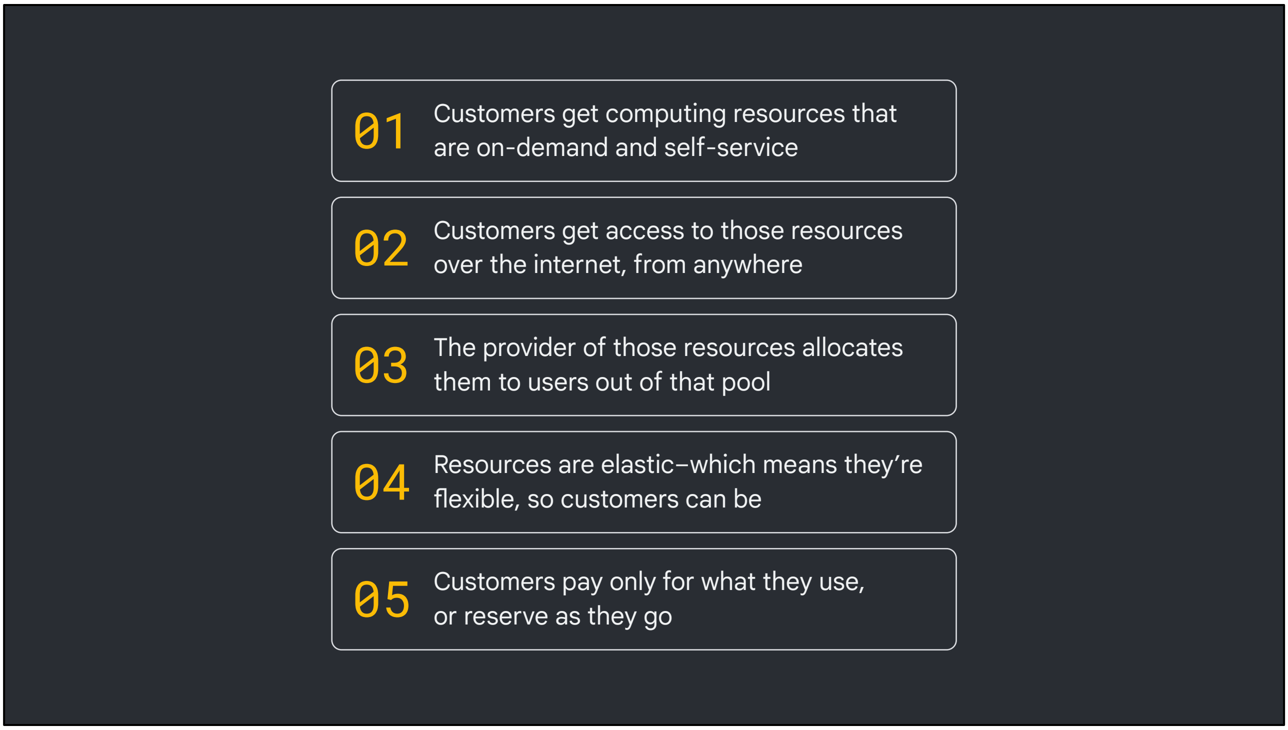
On-demand course

March 2022



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

Cloud computing is a way of using information technology (IT) 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 Resources are elastic–which means they’re flexible, so customers can be

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

- First, customers get computing resources that are on-demand and self-service. Through a web interface, users get the processing power, storage, and network they need with no need for human intervention.

- Second, customers get access to those resources over the internet, from anywhere they have a connection.

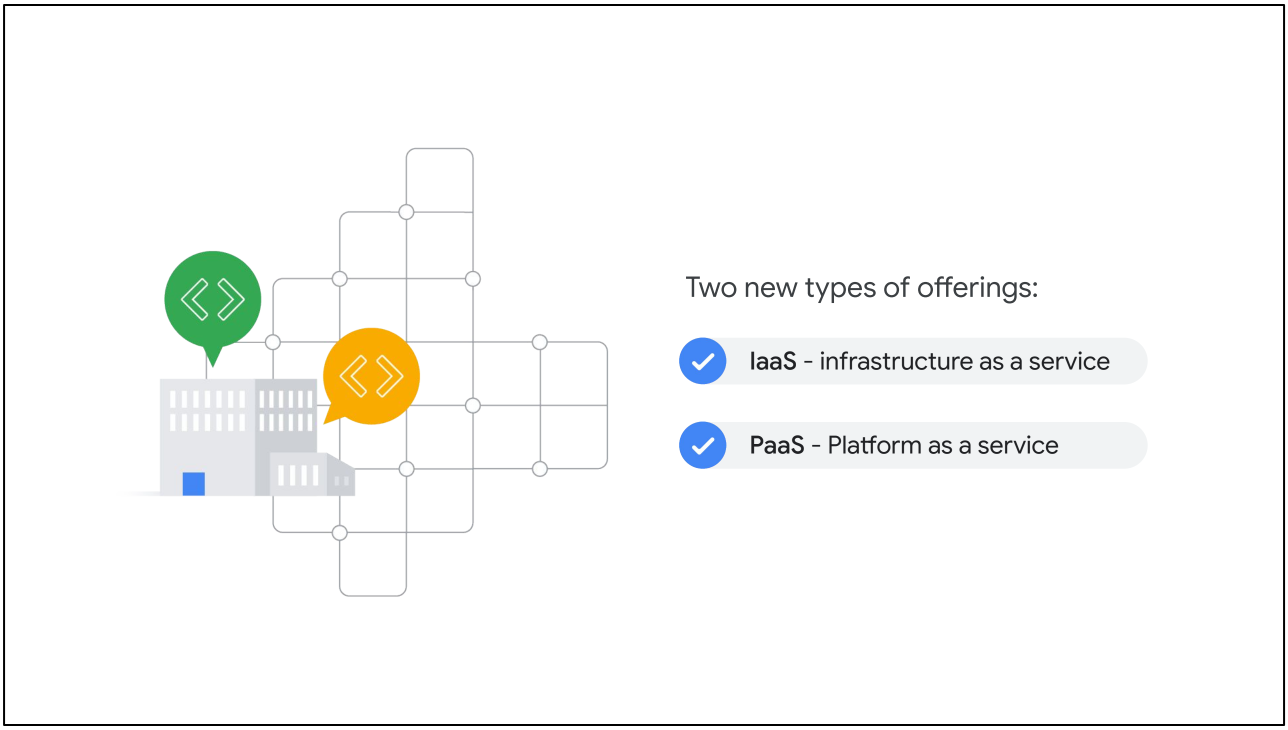
- Third, the cloud provider has a big pool of those resources and allocates them to users out of that pool. That allows the provider to buy in bulk and pass the savings on to the customers. Customers don't have to know or care about the exact physical location of those resources.

- Fourth, the resources are elastic–which means they’re flexible, so customers can be. If they need more resources they can get more, and quickly. If they need less, they can scale back.

- And finally, customers pay only for what they use, or reserve as they go. If they stop using resources, they stop paying.

Every company is, or will eventually become, a data company

This means that every company is, or will eventually become, a data company.

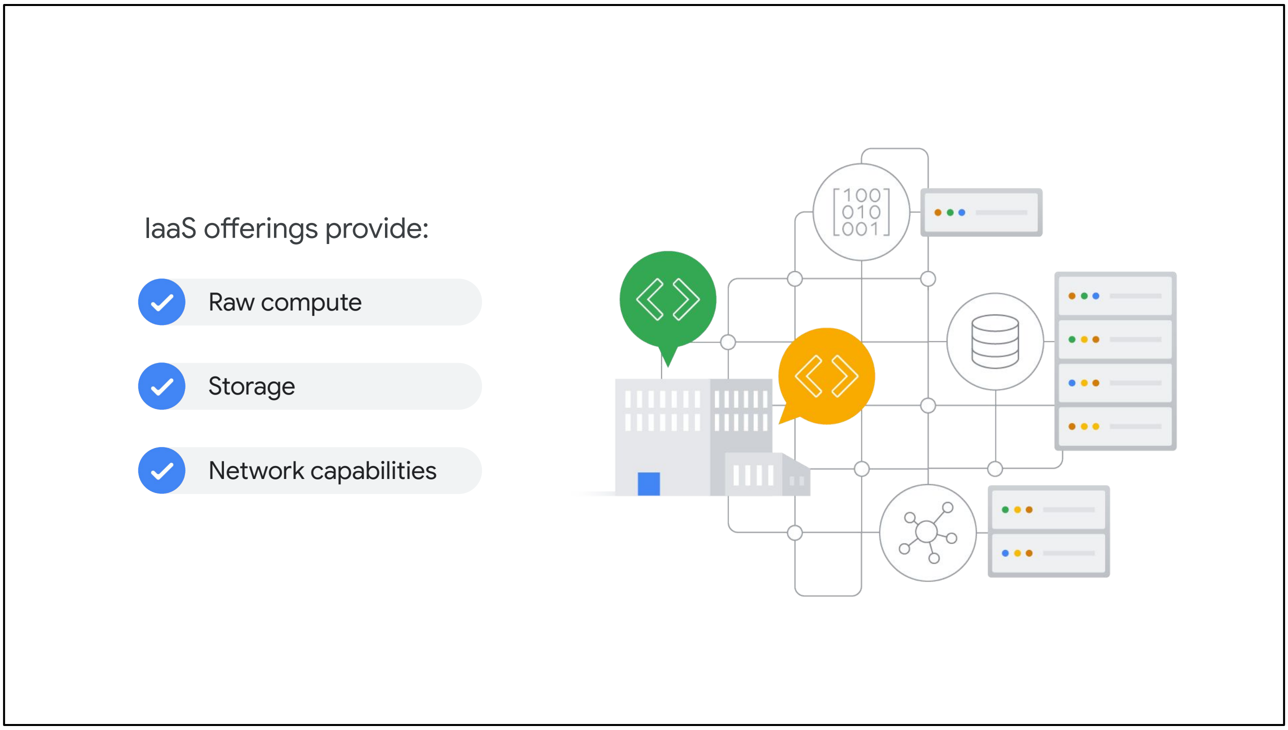


Two new types of offerings:

- ✓ **IaaS** - infrastructure as a service
- ✓ **PaaS** - Platform as a service

The move to virtualized data centers introduced customers to two new types of offerings:

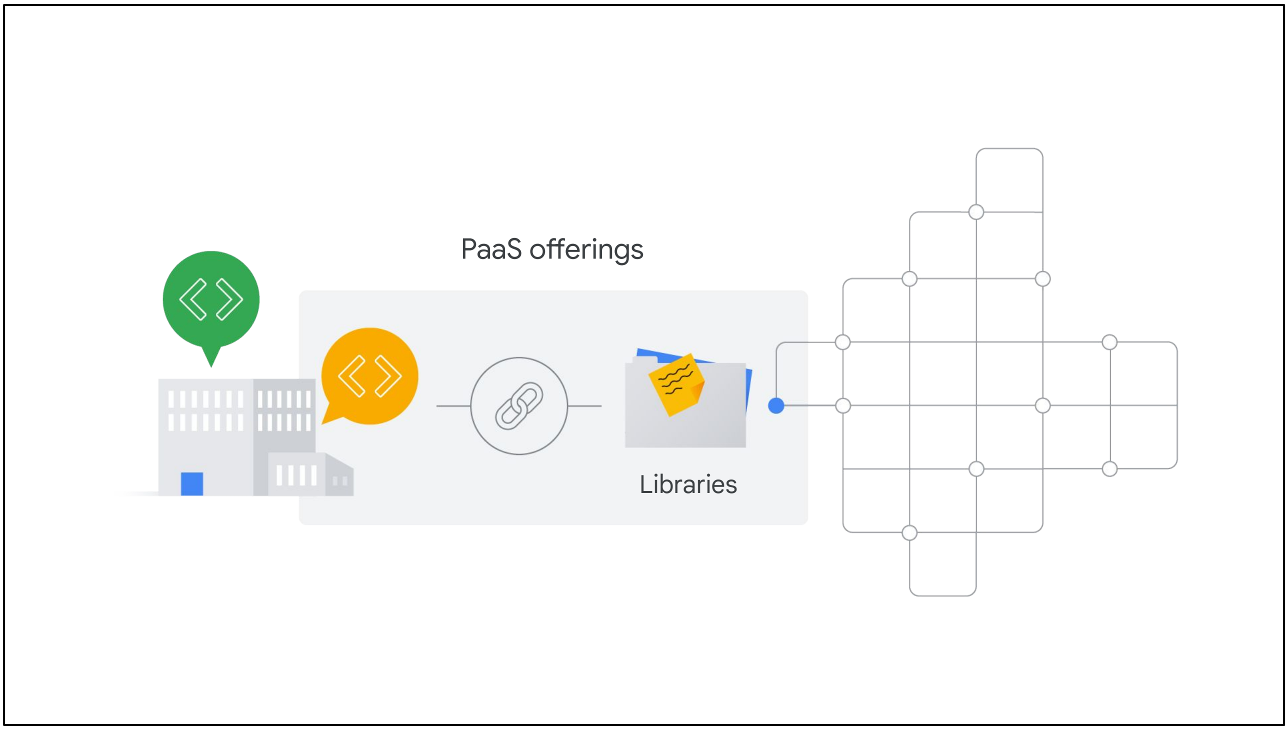
- **infrastructure as a service**, commonly referred to as IaaS, and
- **platform as a service**, or PaaS.



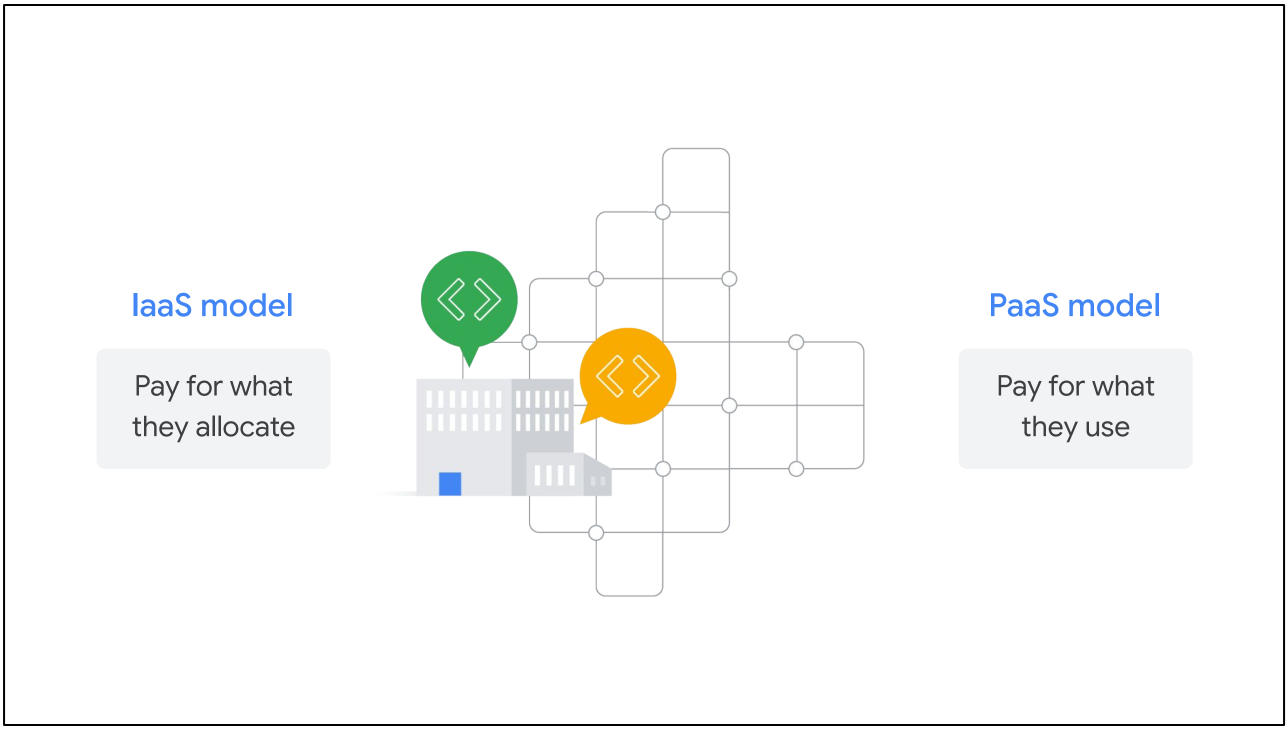
IaaS offerings provide:

- ✓ Raw compute
- ✓ Storage
- ✓ Network capabilities

IaaS offerings provide raw compute, storage, and network capabilities, organized virtually into resources that are similar to physical data centers.



PaaS offerings, in contrast, bind code to libraries that provide access to the infrastructure application needs. This allows more resources to be focused on application logic.



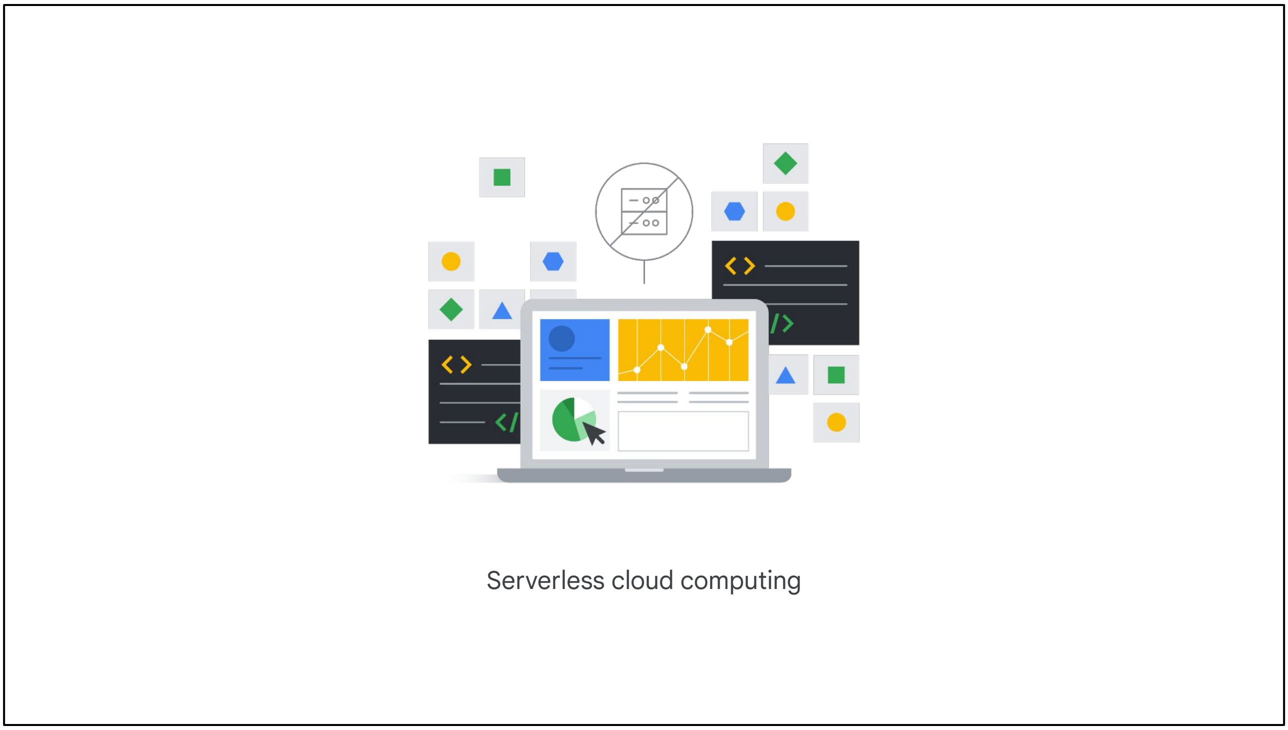
IaaS model

Pay for what they allocate

PaaS model

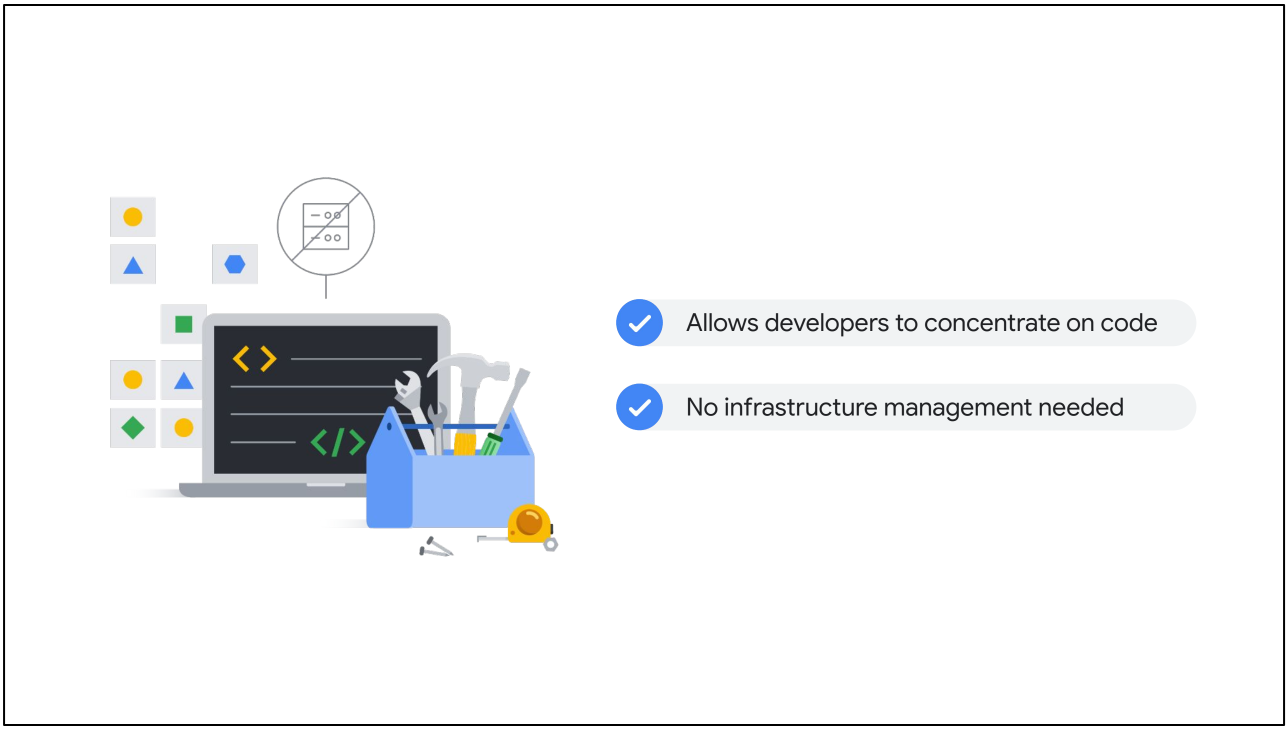
Pay for what they use

In the IaaS model, customers pay for the resources they allocate ahead of time; in the PaaS model, customers pay for the resources they actually use.



Serverless cloud computing

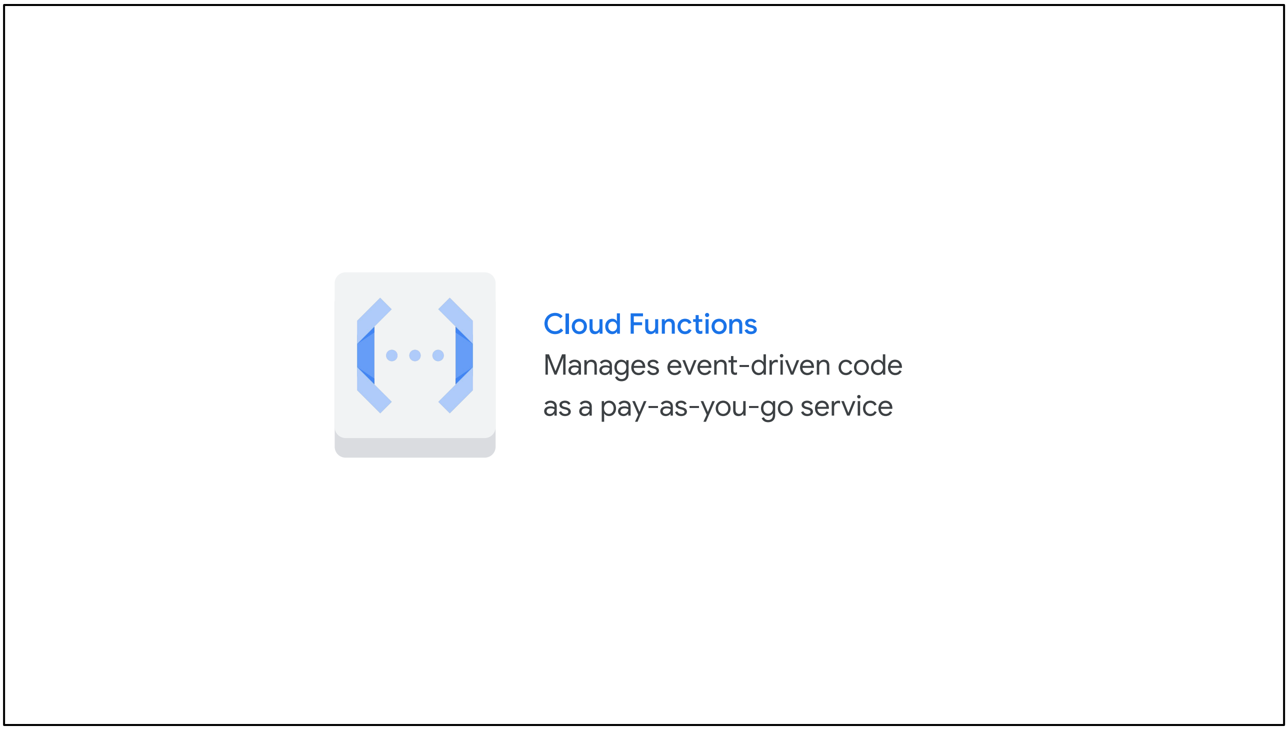
Serverless is yet another step in the evolution of cloud computing.



✓ Allows developers to concentrate on code

✓ No infrastructure management needed

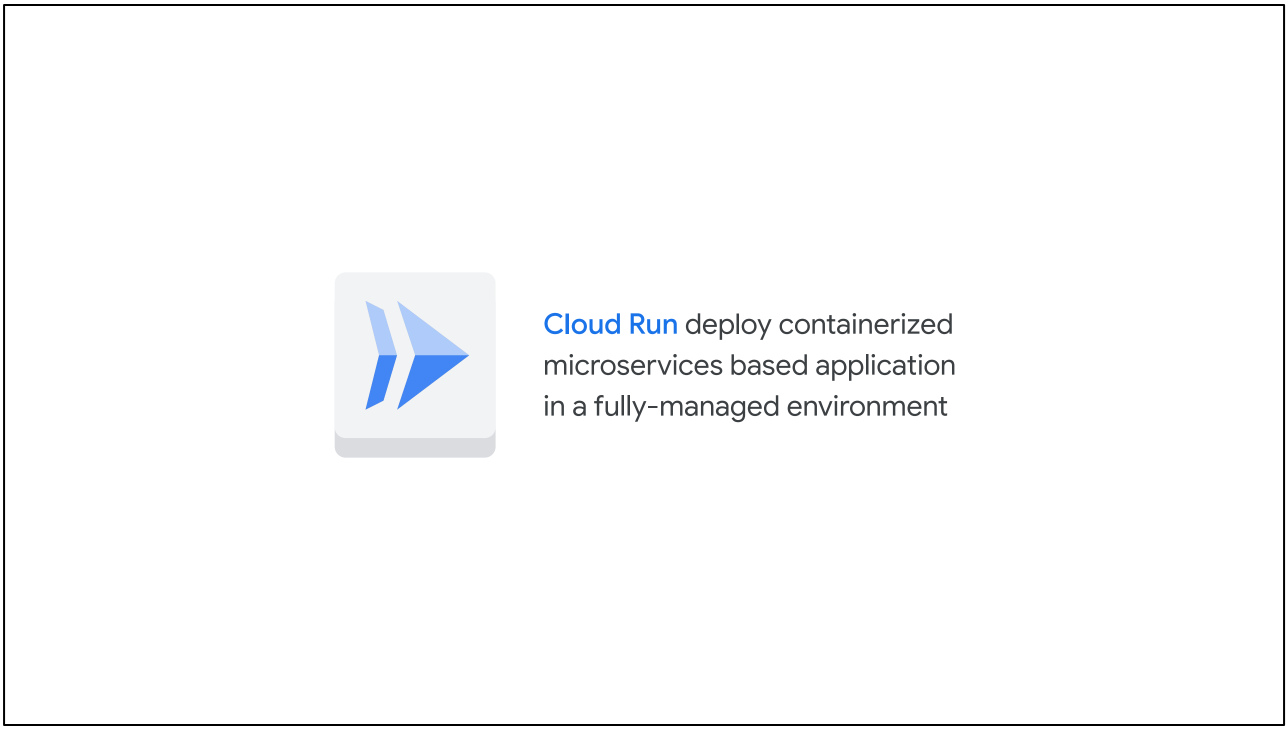
It allows developers to concentrate on their code, rather than on server configuration, by eliminating the need for any infrastructure management. Serverless technologies offered by Google



Cloud Functions

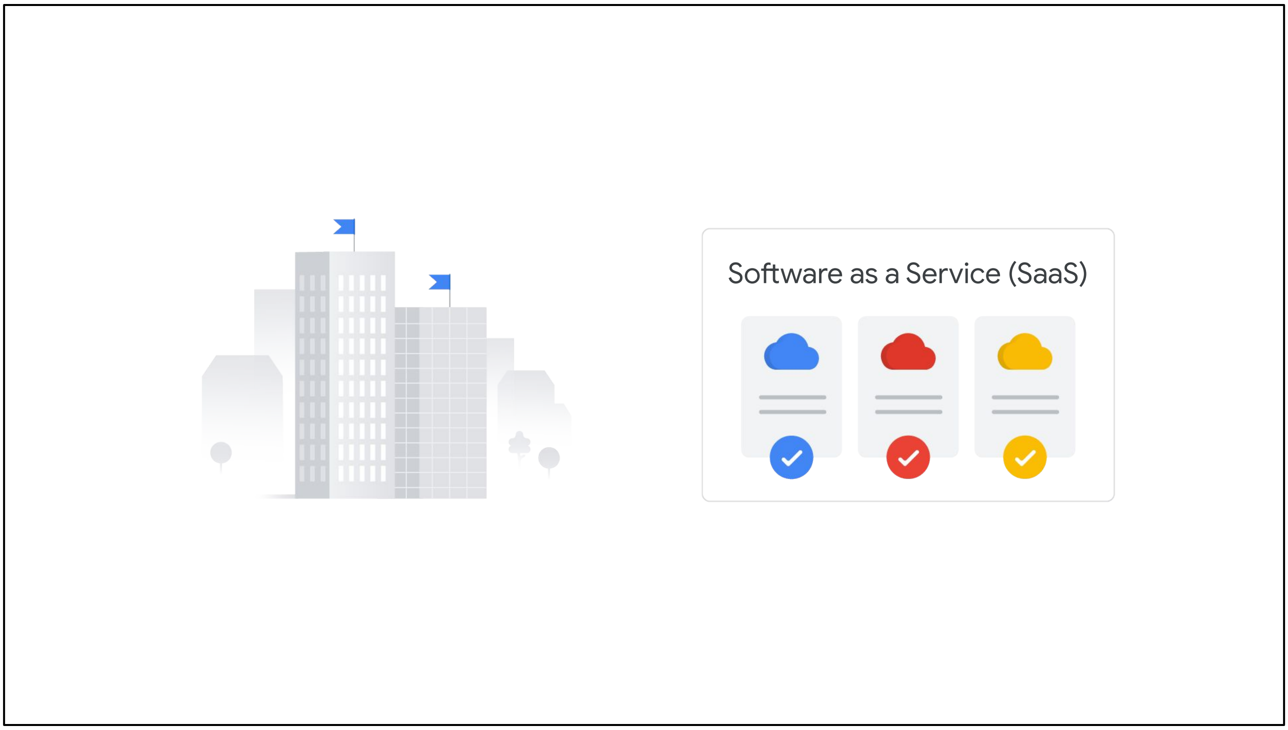
Manages event-driven code
as a pay-as-you-go service

include Cloud Functions, which manages event-driven code as a pay-as-you-go service, and



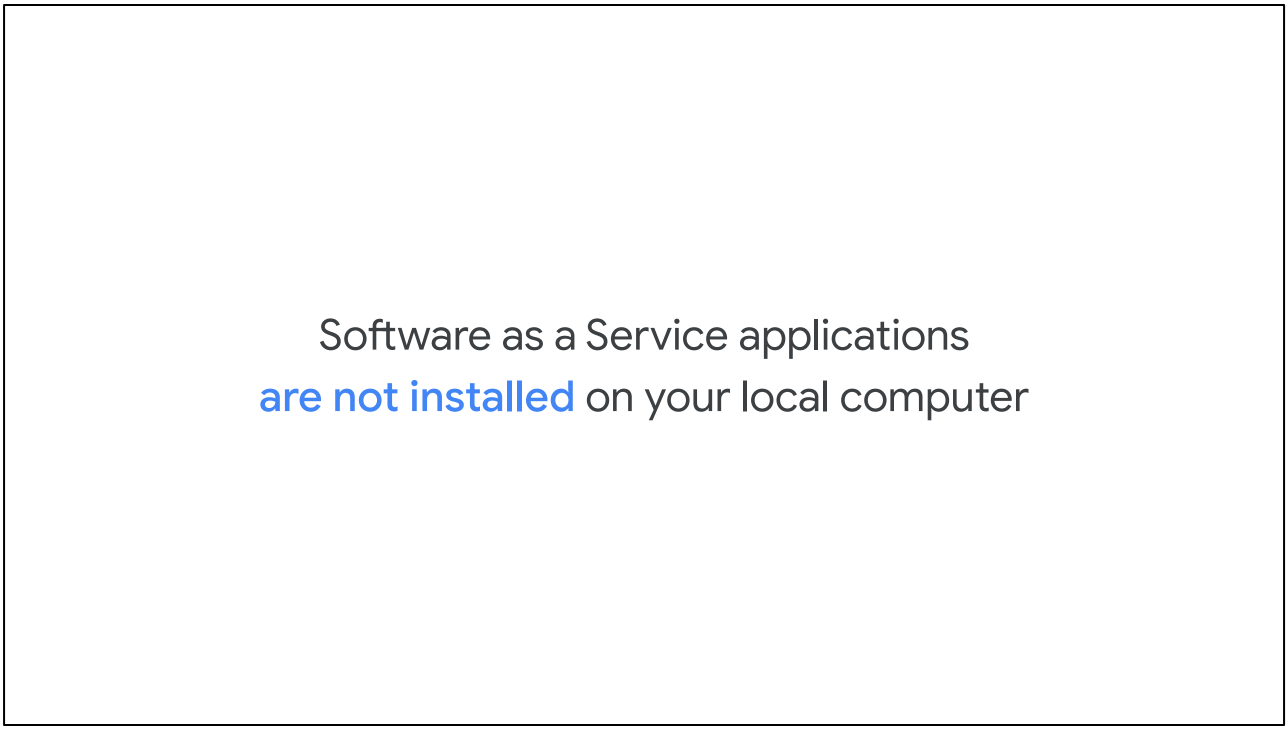
Cloud Run deploy containerized microservices based application in a fully-managed environment

Cloud Run, which allows customers to deploy their containerized microservices based application in a fully-managed environment.



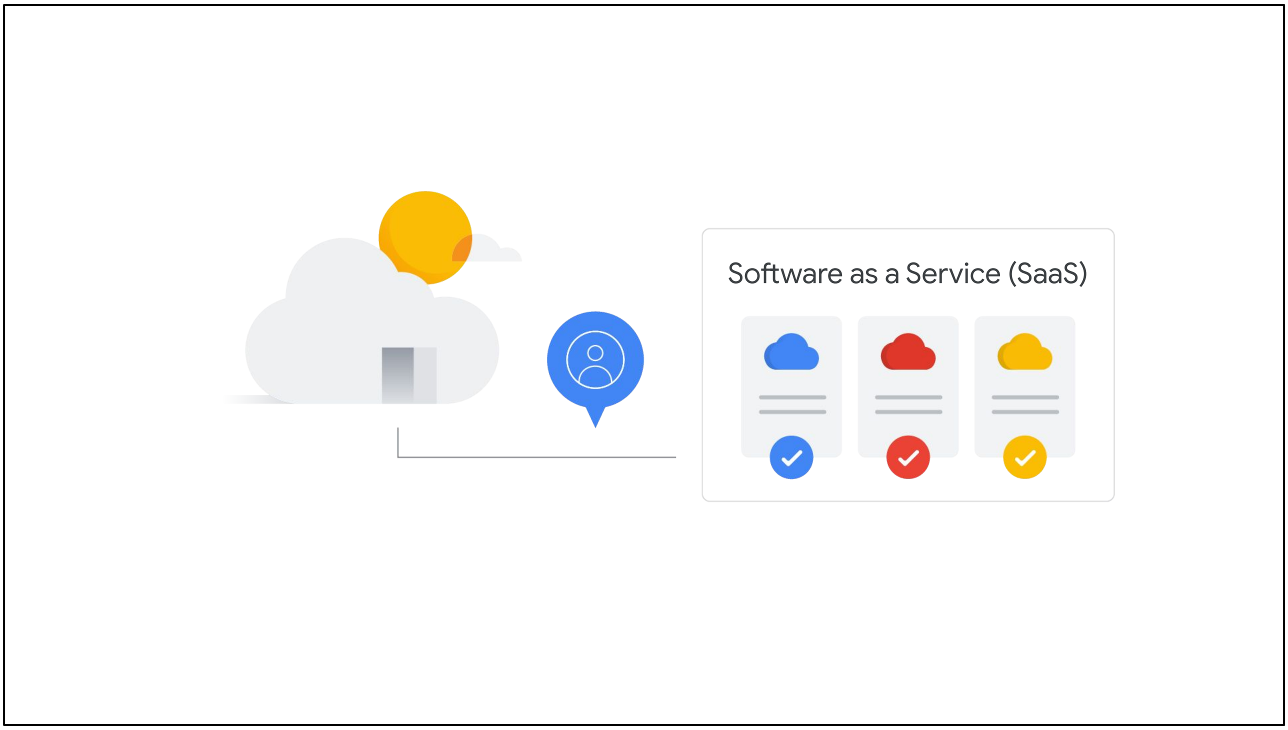
Software as a Service (SaaS)

While it’s outside the scope of *this* course, you might have heard about software as a service, SaaS, and wondered what it is and how it fits into the Cloud ecosphere.



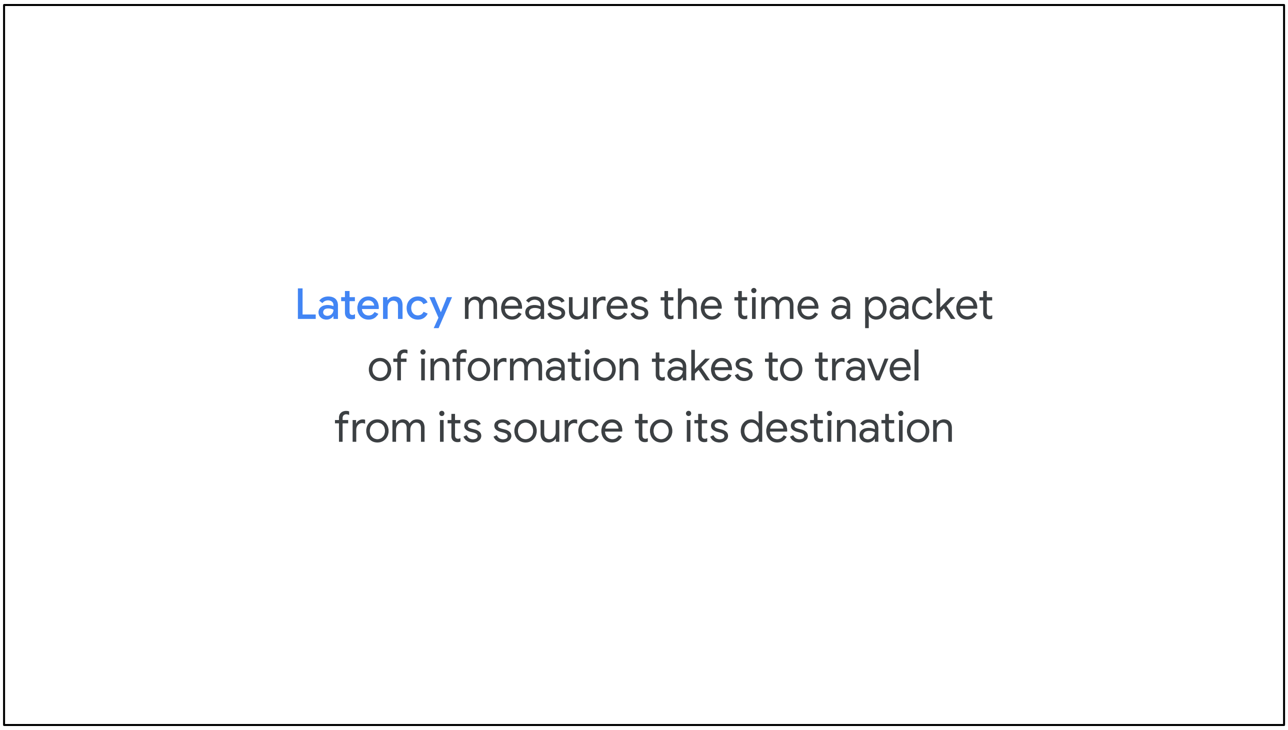
Software as a Service applications
are not installed on your local computer

Software as a Service applications are not installed on your local computer.



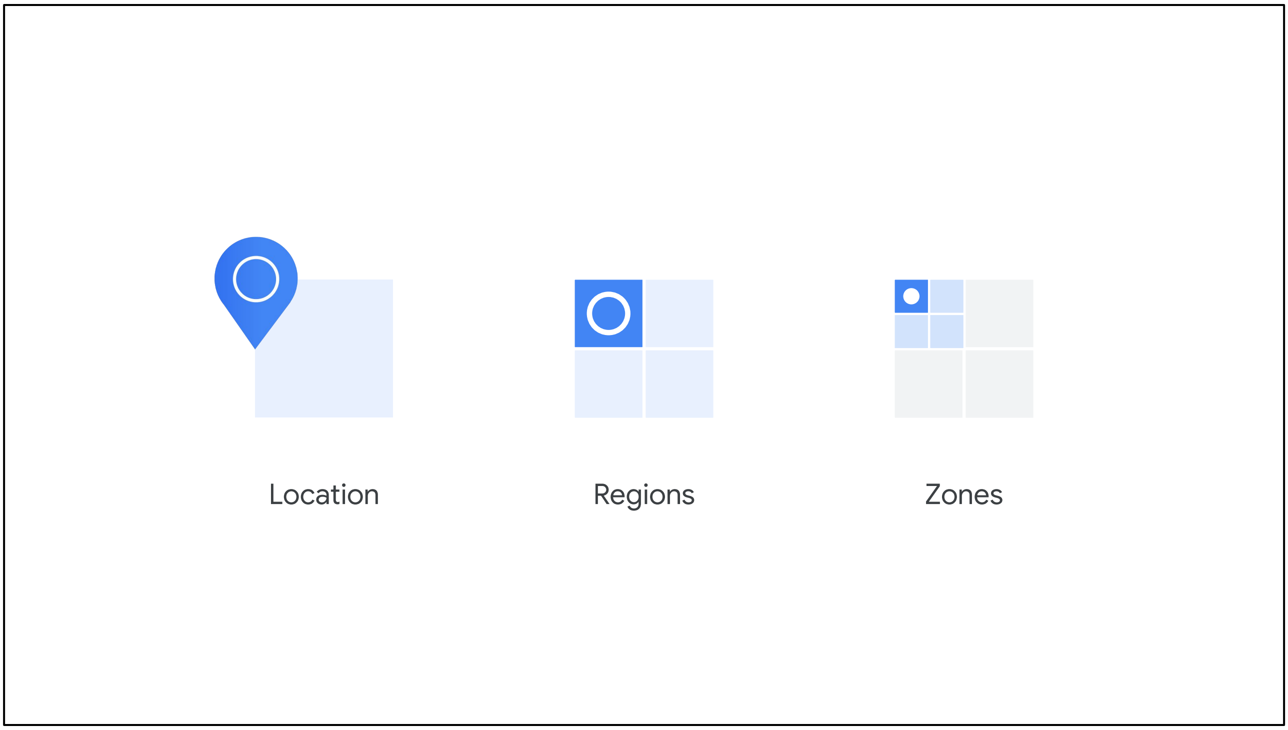
Software as a Service (SaaS)

Instead, they run in the cloud as a service and are consumed directly over the internet by end users.



Latency measures the time a packet of information takes to travel from its source to its destination

which measures the time a packet of information takes to travel from its source to its destination.

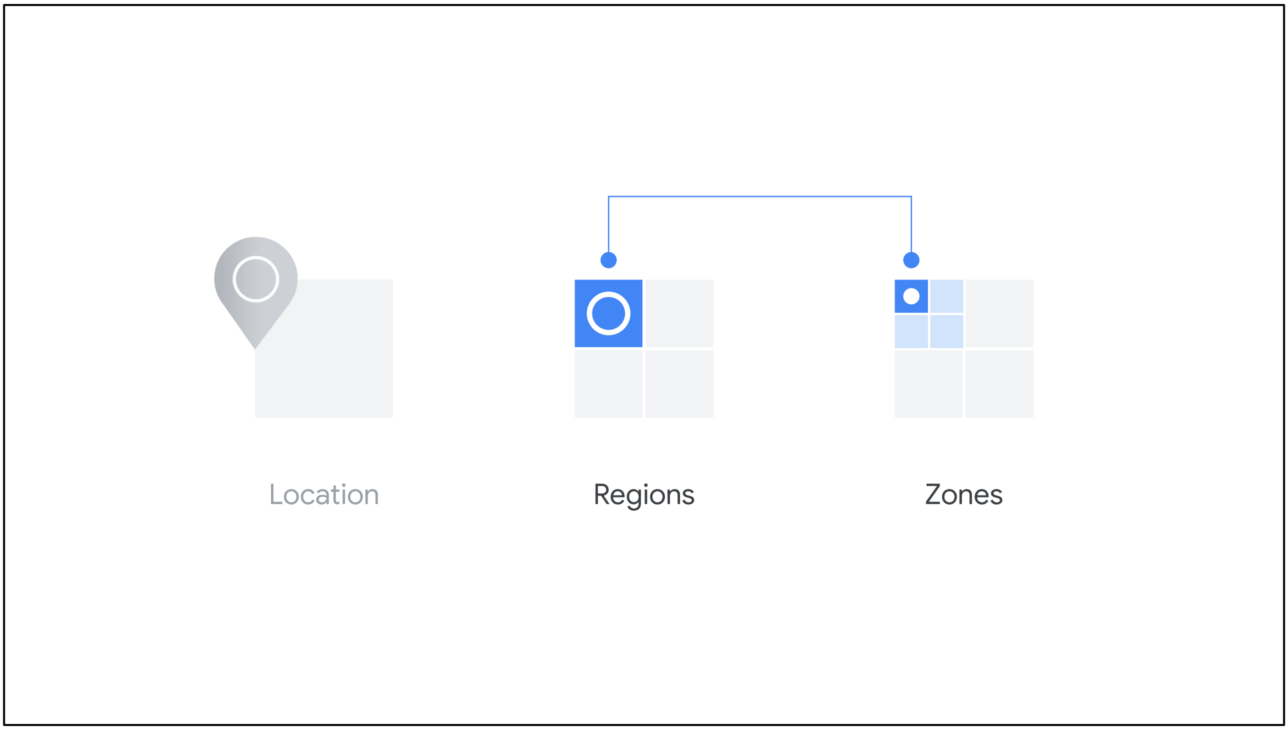


Location

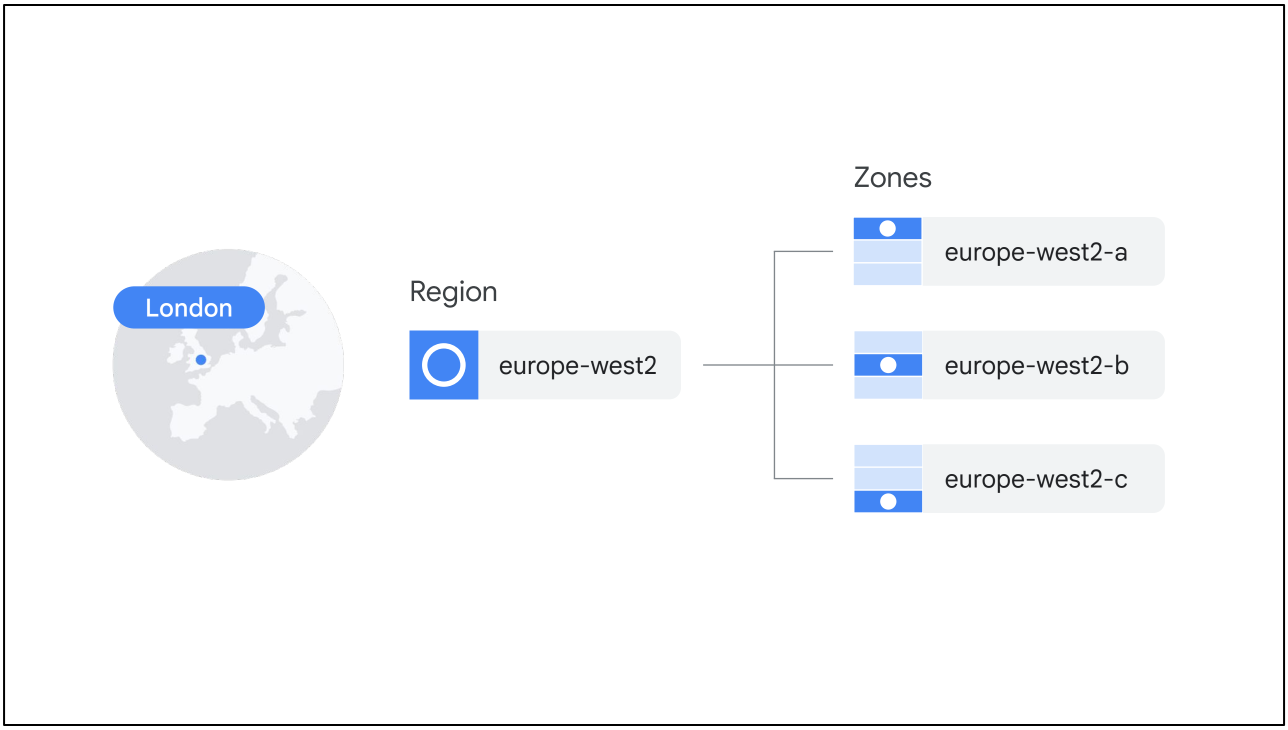
Regions

Zones

Each of these locations is divided into several different **regions** and **zones**.



Regions represent independent geographic areas and are composed of zones.



Region

europe-west2

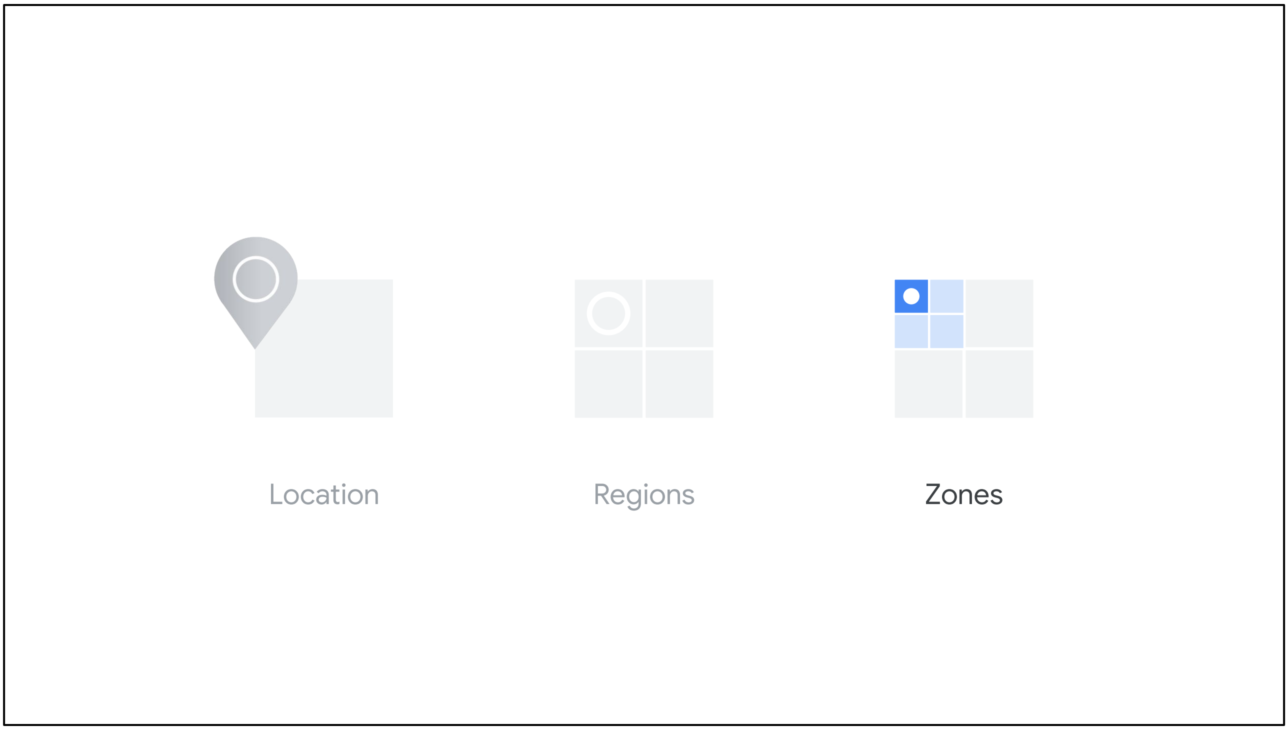
Zones

europe-west2-a

europe-west2-b

europe-west2-c

For example, London, or europe-west2, is a region that currently comprises three different zones.

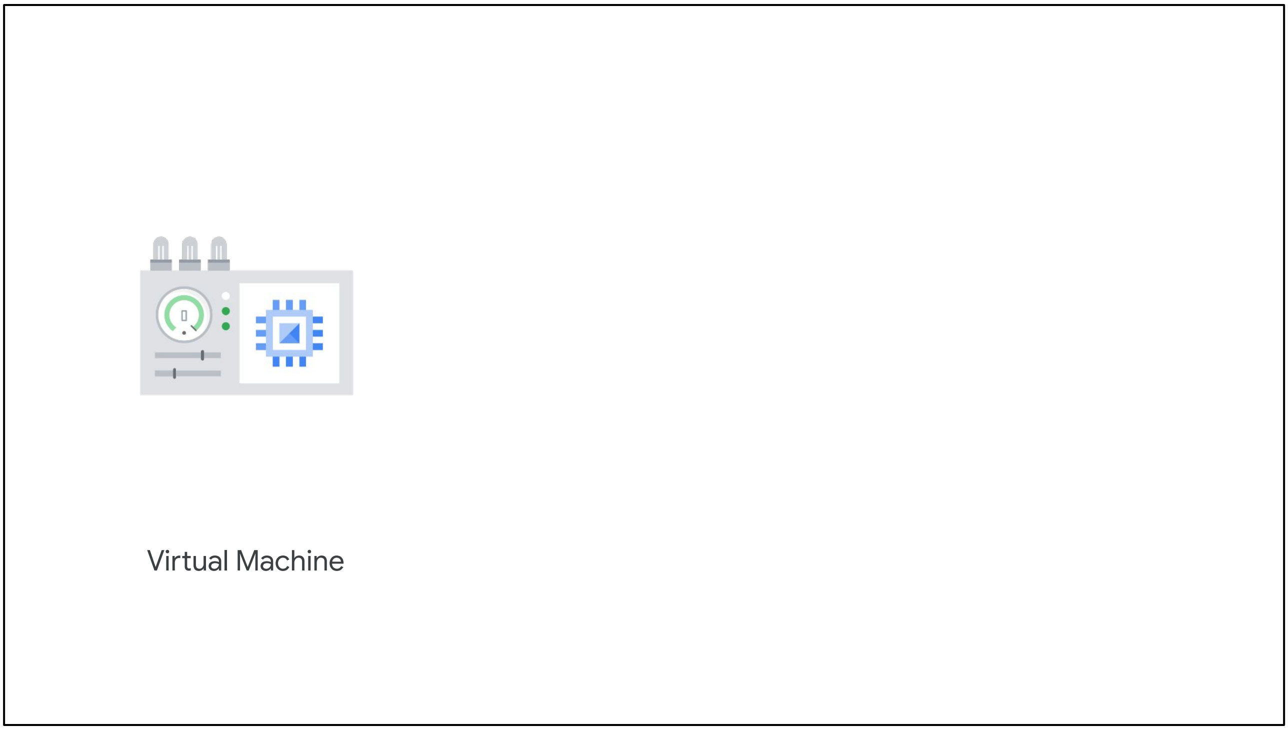


Location

Regions

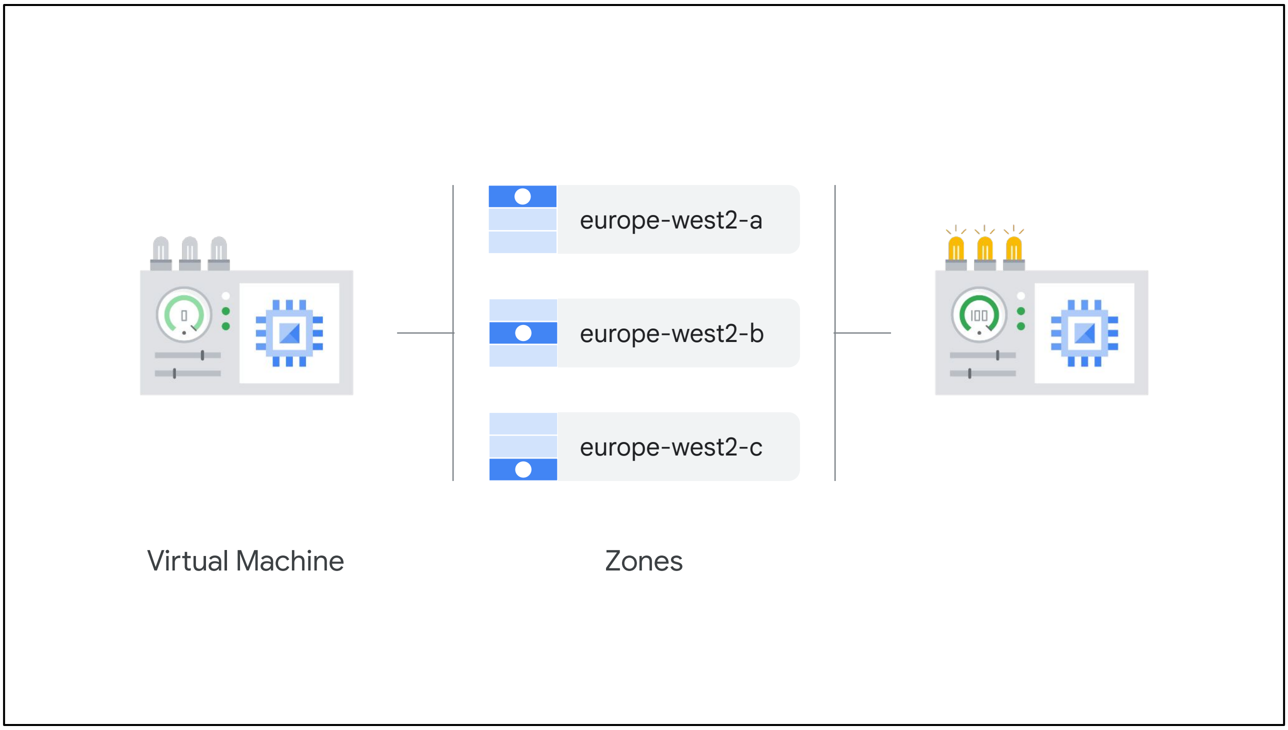
Zones

A zone is an area where Google Cloud resources are deployed.



Virtual Machine

For example, if you launch a virtual machine using Compute Engine



europe-west2-a

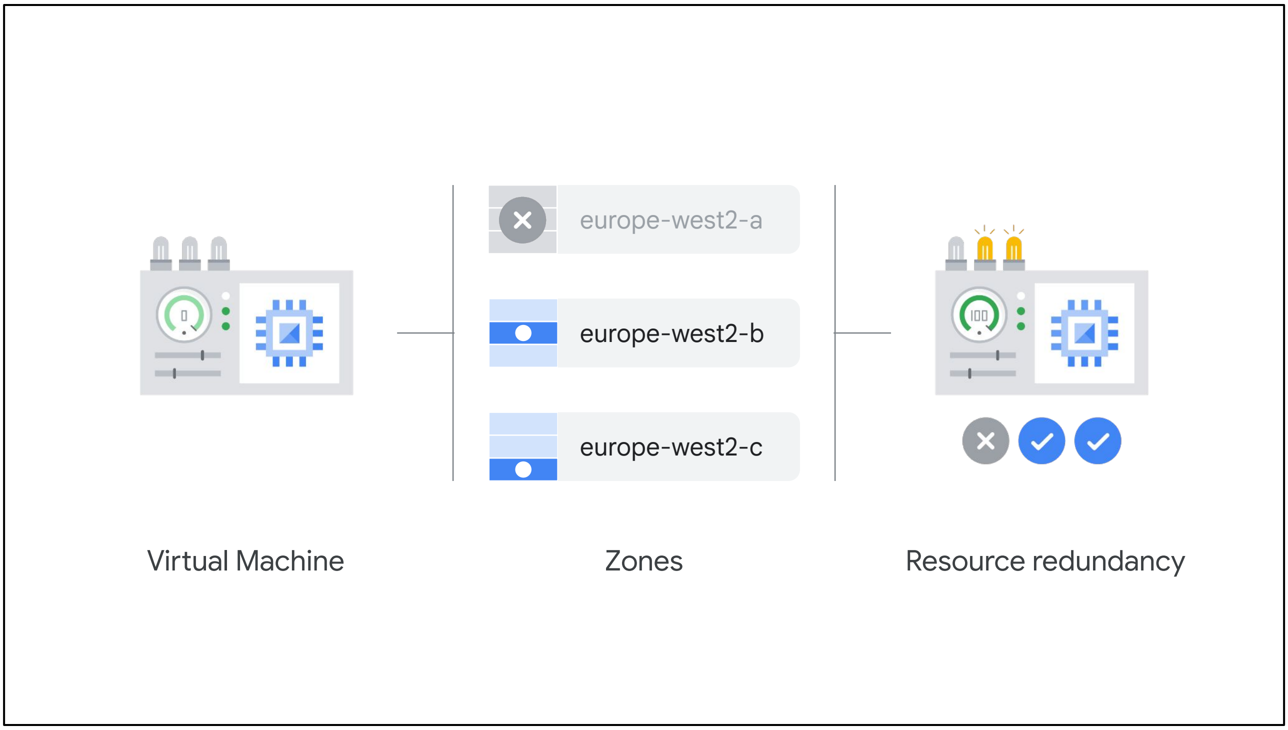
europe-west2-b

europe-west2-c

Virtual Machine

Zones

it will run in the zone that you specify



europe-west2-a

europe-west2-b

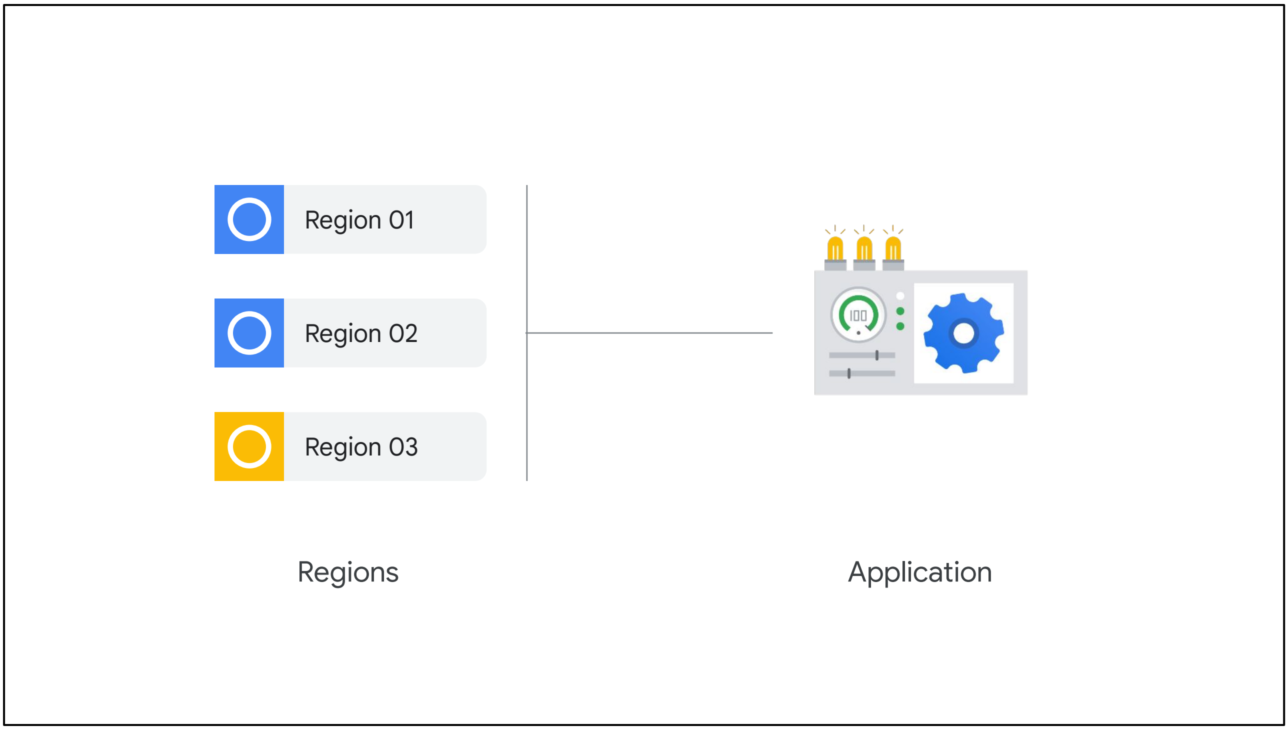
europe-west2-c

Virtual Machine

Zones

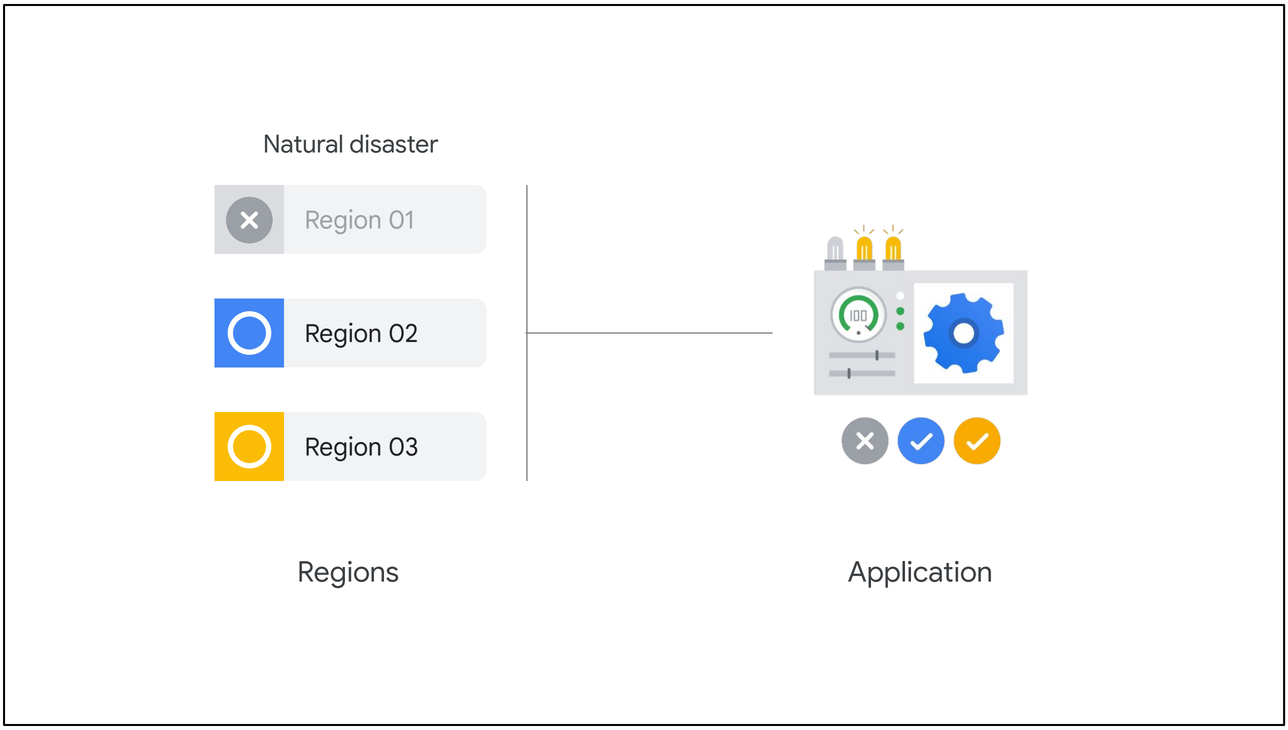
Resource redundancy

to ensure resource redundancy.

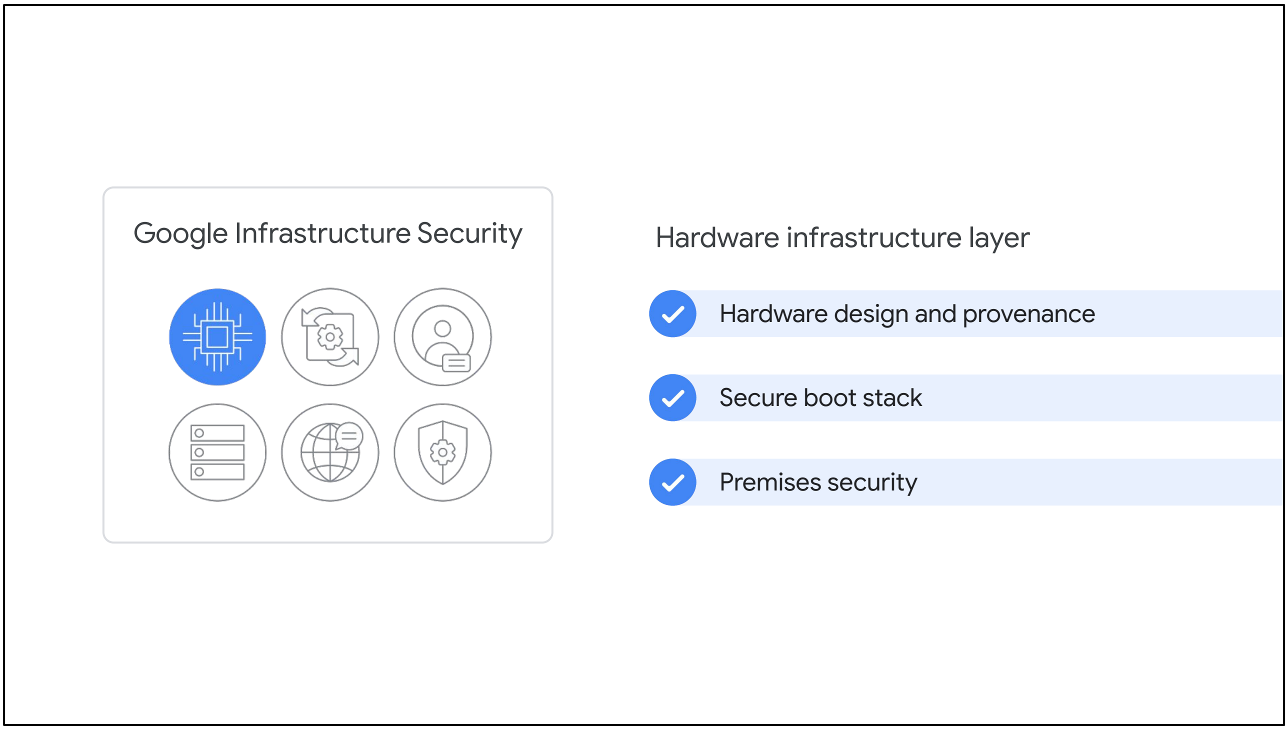


You can run resources in different regions.

This is useful for bringing applications closer to users around the world, and also for protection in case there are issues with an entire region,



say, due to a natural disaster.



Google Infrastructure Security

Hardware infrastructure layer

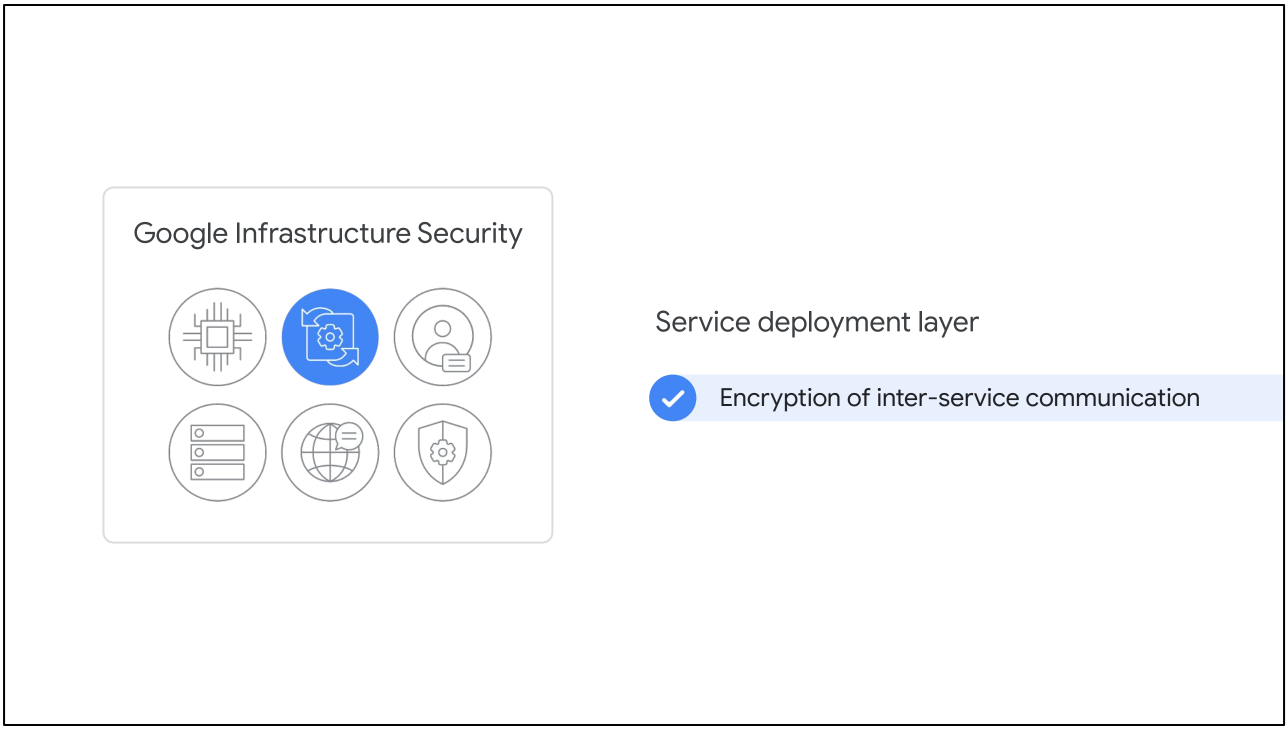
- ✓ Hardware design and provenance
- ✓ Secure boot stack
- ✓ Premises security

We begin with the **Hardware infrastructure** layer which comprises three key security features:

- The first is **hardware design and provenance**. Both the server boards and the networking equipment in Google data centers are custom-designed by Google. Google also designs custom chips, including a hardware security chip that's currently being deployed on both servers and peripherals.

- The next feature is a **secure boot stack**. Google server machines use a variety of technologies to ensure that they are booting the correct software stack, such as cryptographic signatures over the BIOS, bootloader, kernel, and base operating system image.

- This layer's final feature is **premises security**. Google designs and builds its own data centers, which incorporate multiple layers of physical security protections. Access to these data centers is limited to only a very small number of Google employees. Google additionally hosts some servers in third-party data centers, where we ensure that there are Google-controlled physical security measures on top of the security layers provided by the data center operator.



Google Infrastructure Security

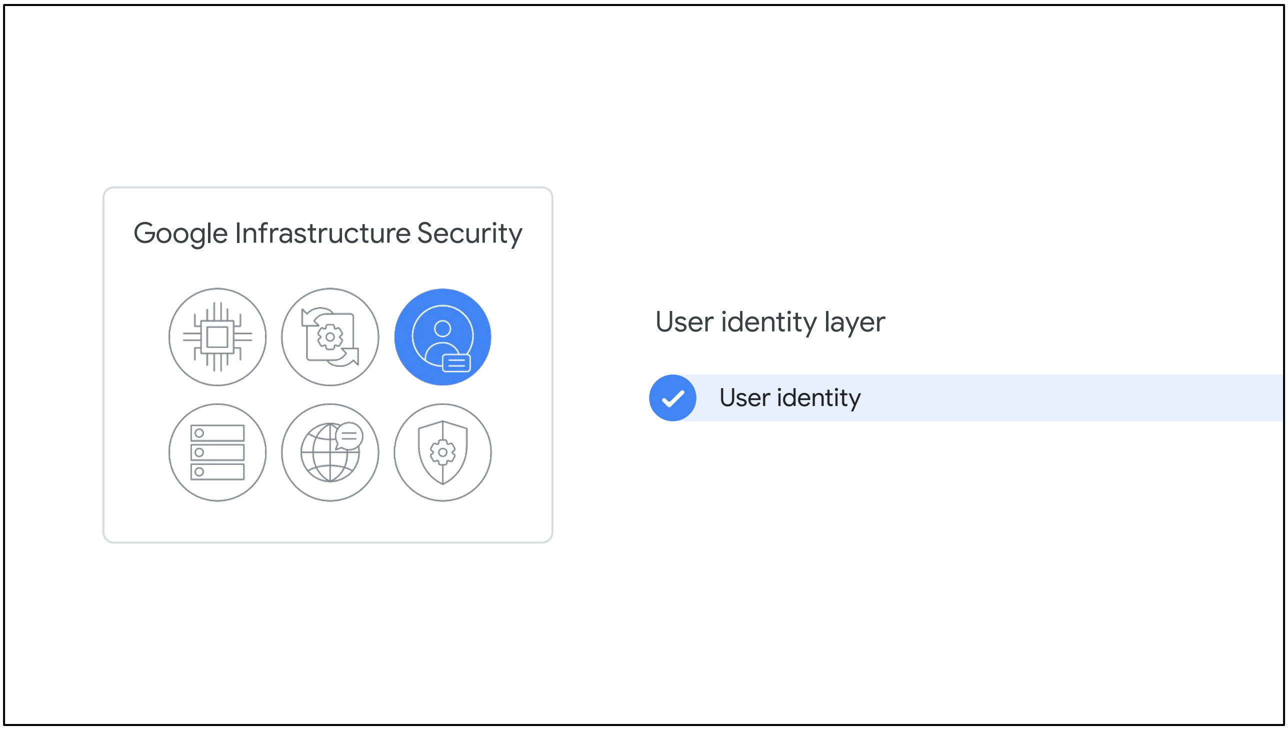
Service deployment layer

✓ Encryption of inter-service communication

Next is the **Service deployment** layer, where the key feature is encryption of inter-service communication.

Google's infrastructure provides cryptographic privacy and integrity for remote procedure call ("RPC") data on the network. Google's services communicate with each other using RPC calls. The infrastructure automatically encrypts all infrastructure RPC traffic that goes between data centers.

Google has started to deploy hardware cryptographic accelerators that will allow it to extend this default encryption to all infrastructure RPC traffic inside Google data centers.



Google Infrastructure Security

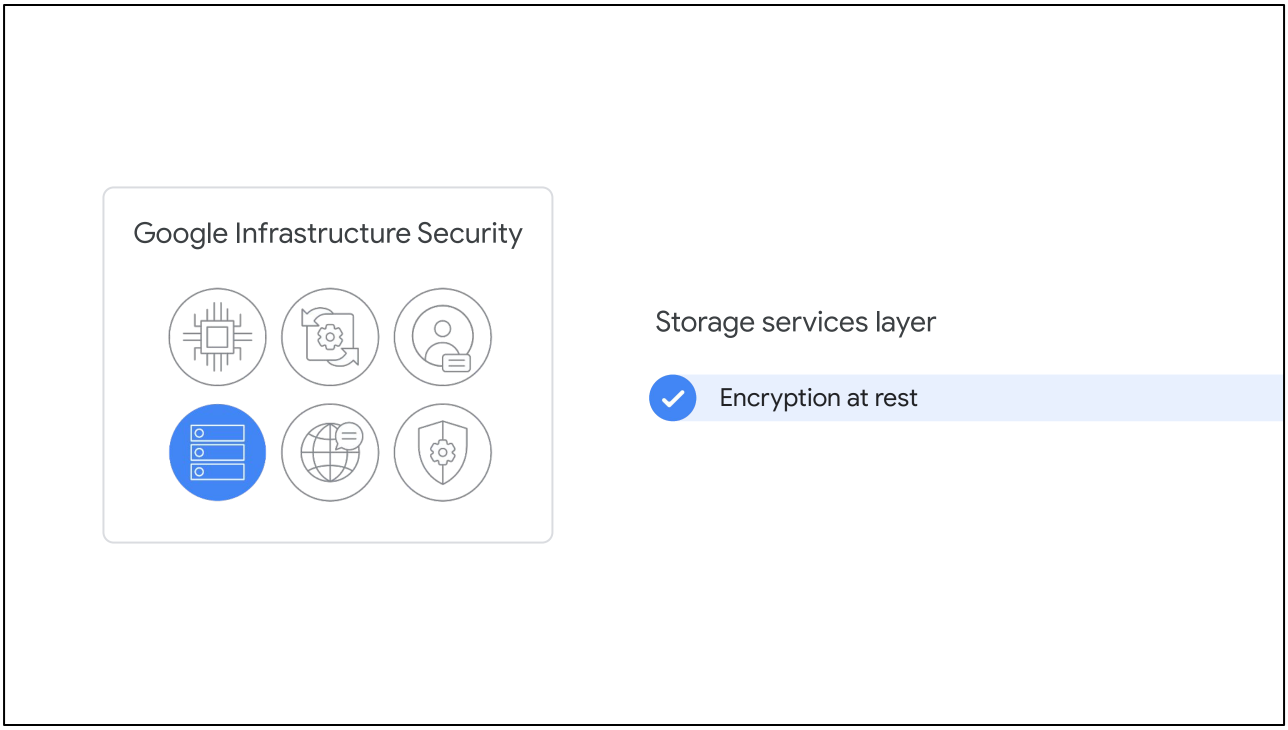
User identity layer

✓ User identity

Then we have the **User identity** layer.

Google’s central identity service, which usually manifests to end users as the Google login page, goes beyond asking for a simple username and password. The service also intelligently challenges users for additional information based on risk factors such as whether they have logged in from the same device or a similar location in the past.

Users can also employ secondary factors when signing in, including devices based on the Universal 2nd Factor (U2F) open standard.



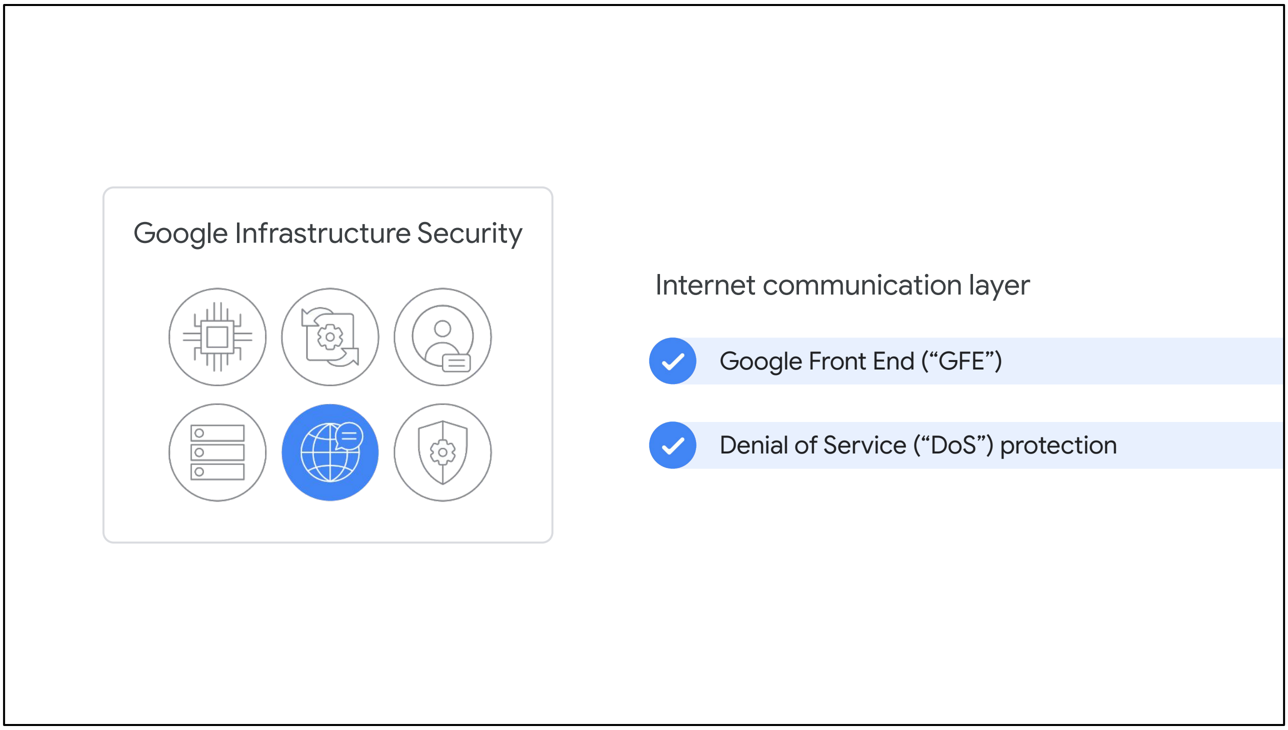
Google Infrastructure Security

Storage services layer

✓ Encryption at rest

On the **Storage services** layer we find the **encryption at rest** security feature.

Most applications at Google access physical storage (in other words, “file storage”) indirectly via storage services, and encryption using centrally managed keys is applied at the layer of these storage services. Google also enables hardware encryption support in hard drives and SSDs.



Google Infrastructure Security

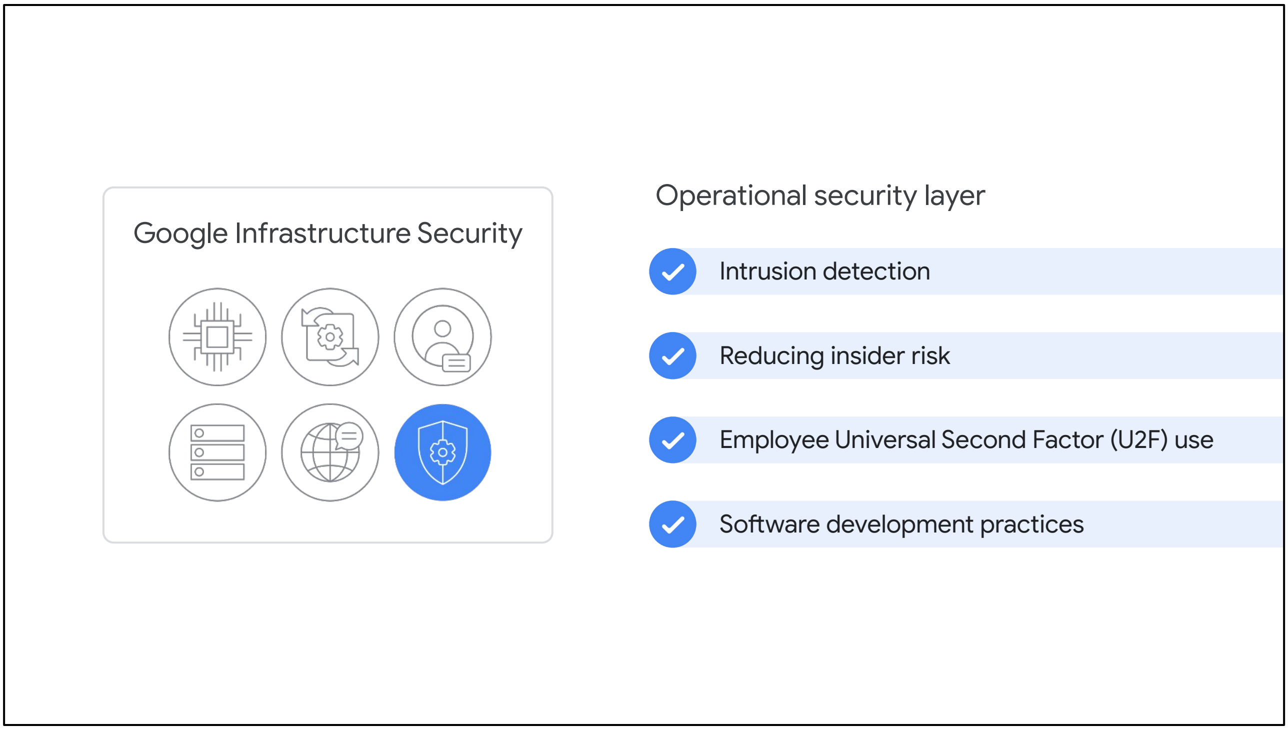
Internet communication layer

- ✓ Google Front End (“GFE”)
- ✓ Denial of Service (“DoS”) protection

The next layer up is the **Internet communication** layer, and this comprises two key security features.

- Google services that are being made available on the internet, register themselves with an infrastructure service called the **Google Front End**, which ensures that all TLS connections are ended using a public-private key pair and an X.509 certificate from a Certified Authority (CA), as well as following best practices such as supporting perfect forward secrecy. The GFE additionally applies protections against Denial of Service attacks.

- Also provided is **Denial of Service (“DoS”) protection**. The sheer scale of its infrastructure enables Google to simply absorb many DoS attacks. Google also has multi-tier, multi-layer DoS protections that further reduce the risk of any DoS impact on a service running behind a GFE.



Google Infrastructure Security

Operational security layer

- ✓ Intrusion detection
- ✓ Reducing insider risk
- ✓ Employee Universal Second Factor (U2F) use
- ✓ Software development practices

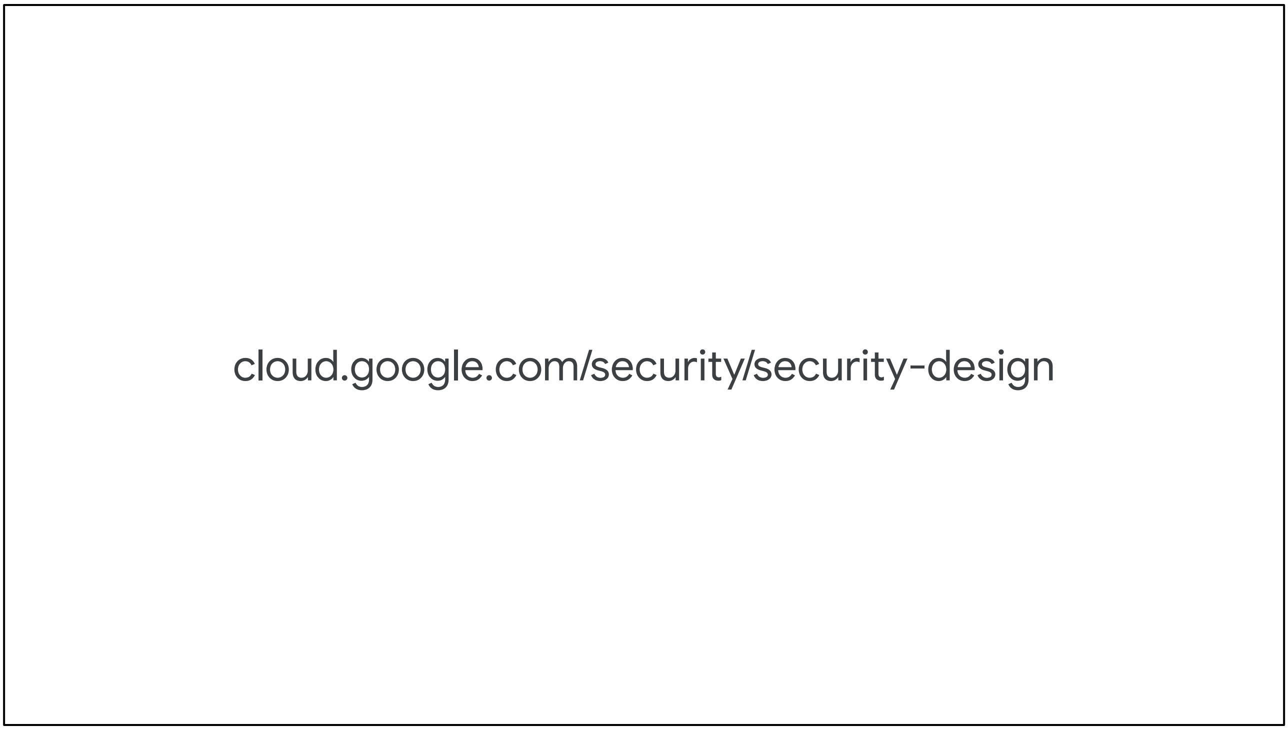
The final layer is Google's **Operational security** layer which provides four key features.

- First is **intrusion detection**. Rules and machine intelligence give Google's operational security teams warnings of possible incidents. Google conducts Red Team exercises to measure and improve the effectiveness of its detection and response mechanisms.

- Next is **reducing insider risk**. Google aggressively limits and actively monitors the activities of employees who have been granted administrative access to the infrastructure.

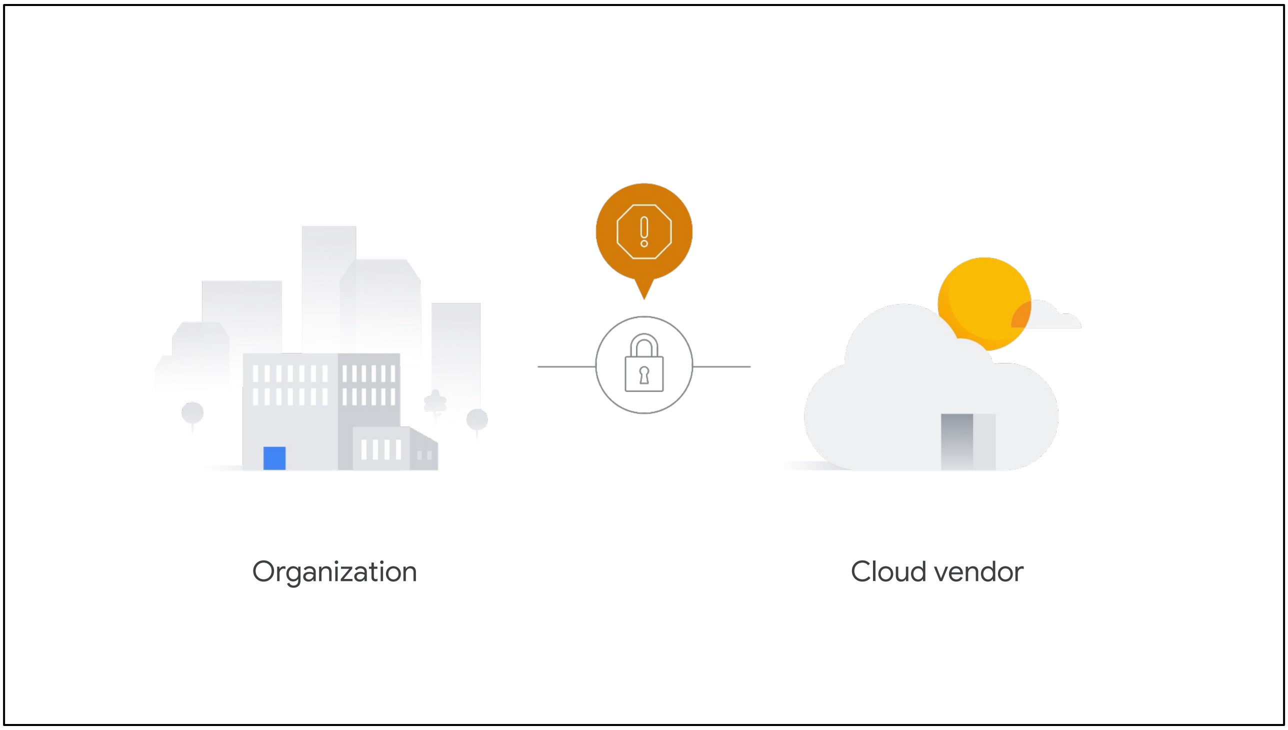
- Then there's **employee U2F use**. To guard against phishing attacks against Google employees, employee accounts require use of U2F-compatible Security Keys.

- Finally, there are stringent **software development practices**. Google employs central source control and requires two-party review of new code. Google also provides its developers libraries that prevent them from introducing certain classes of security bugs. Additionally, Google runs a Vulnerability Rewards Program where we pay anyone who is able to discover and inform us of bugs in our infrastructure or applications.



cloud.google.com/security/security-design

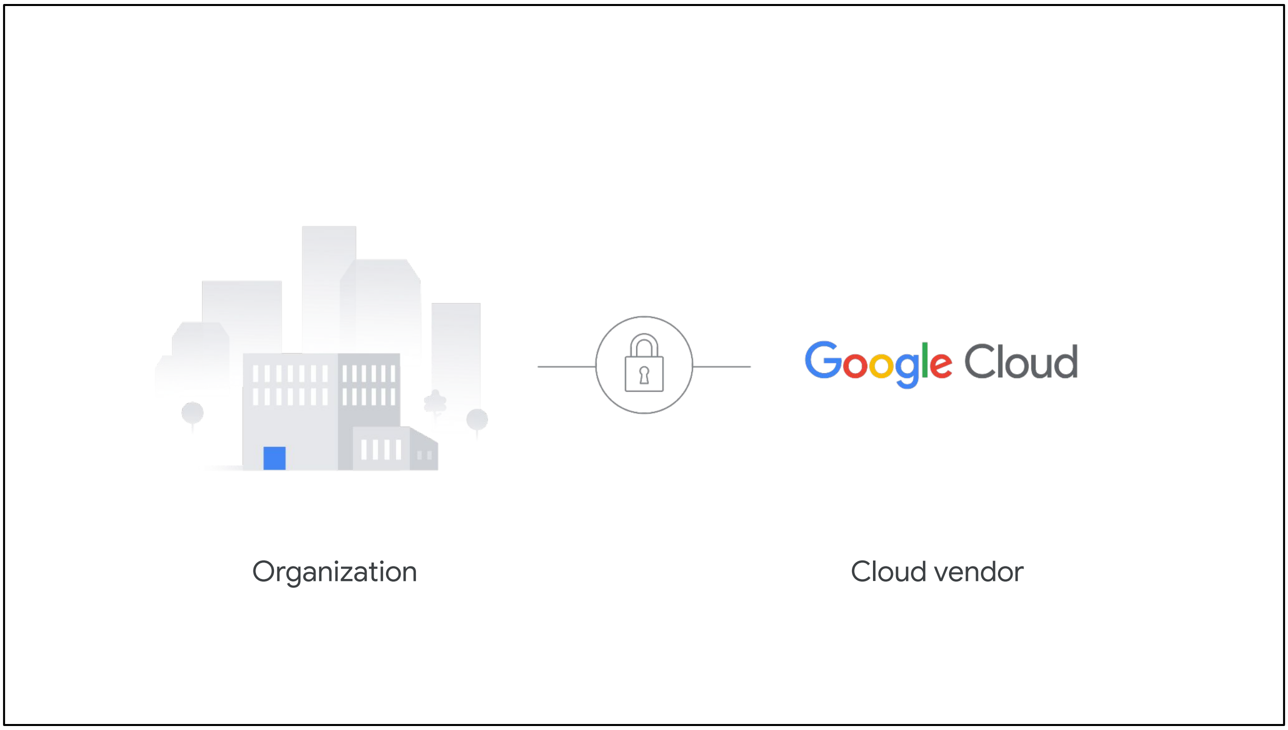
You can learn more about Google’s technical-infrastructure security at cloud.google.com/security/security-design.



Organization

Cloud vendor

Some organizations are afraid to bring their workloads to the cloud because they're afraid they'll get locked into a particular vendor.

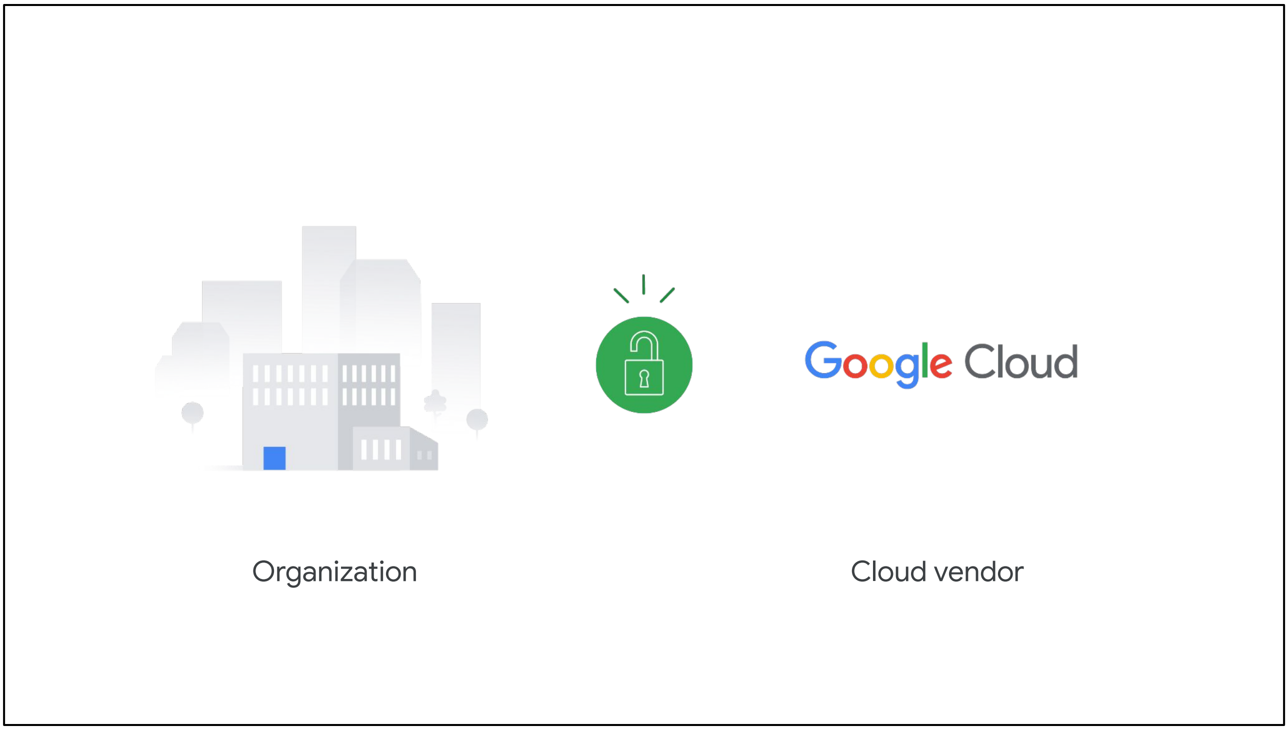


Google Cloud

Organization

Cloud vendor

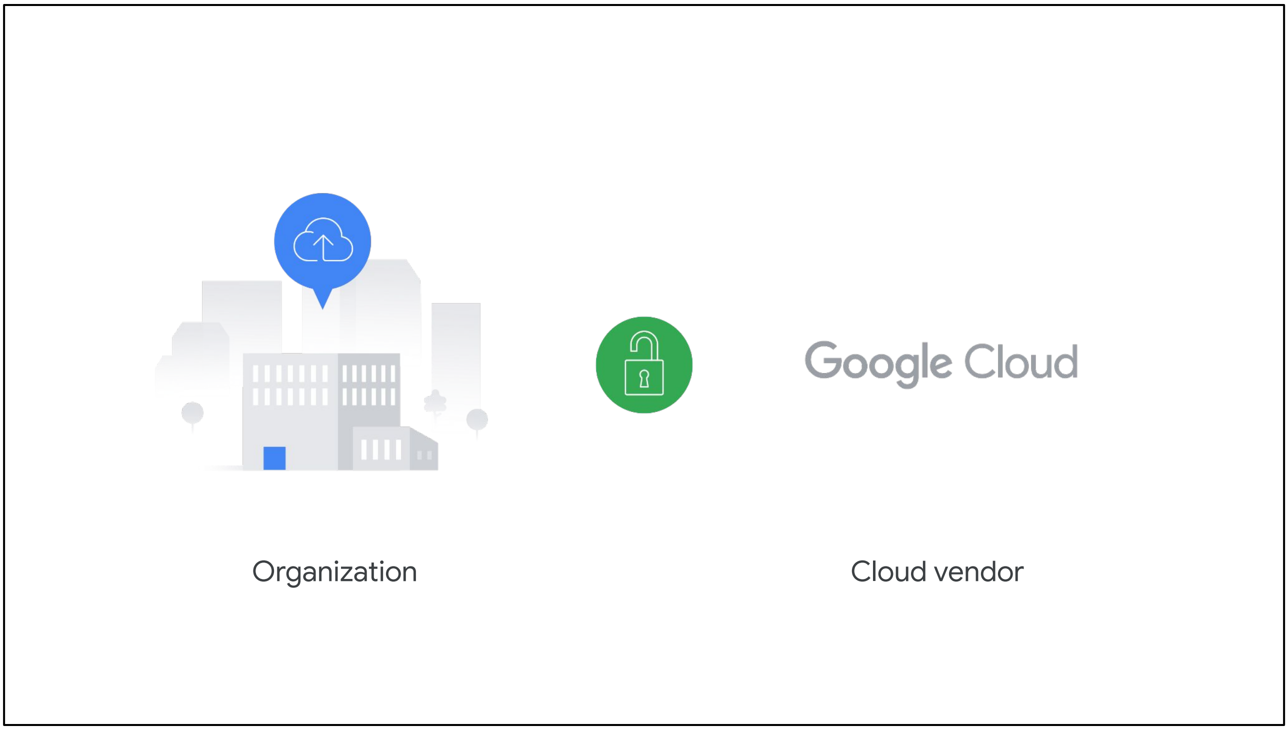
However, if, for whatever reason, a customer decides that Google



Organization

Cloud vendor

is no longer the best provider for their needs, we provide them with the ability



Organization

Cloud vendor

to run their applications elsewhere.