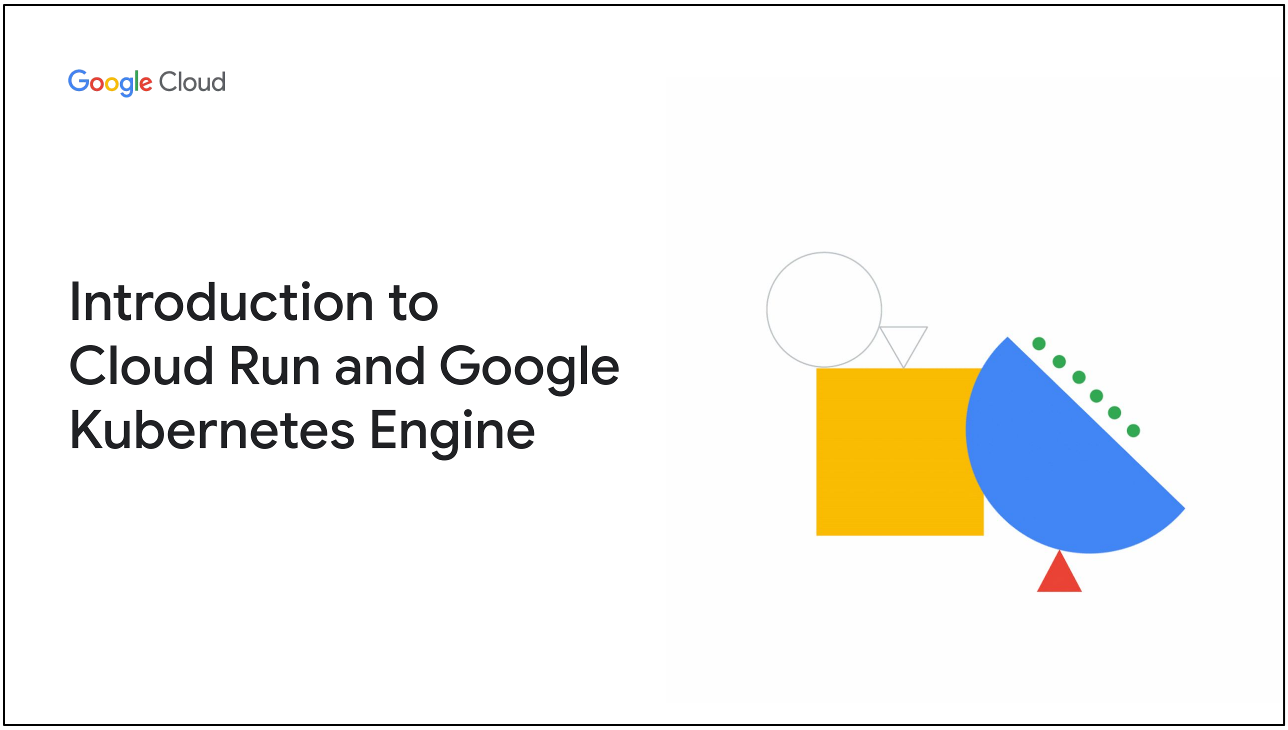
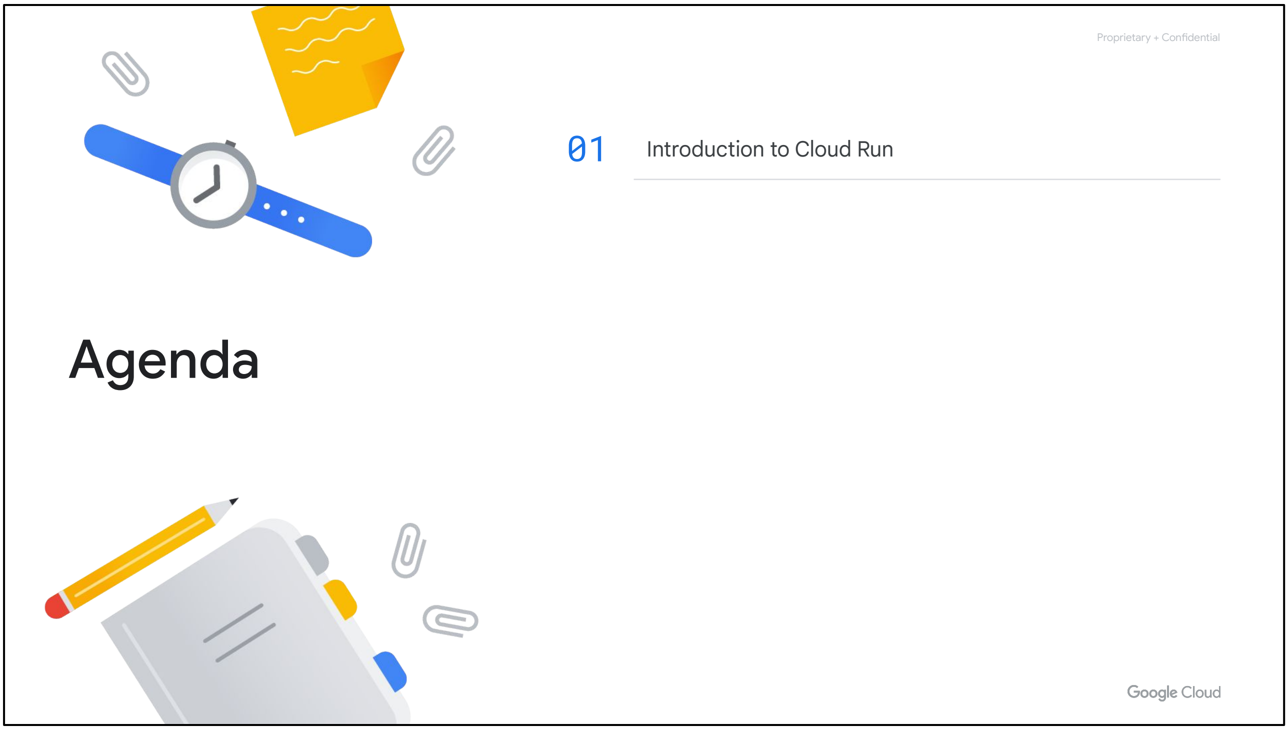


Google Cloud

Introduction to Cloud Run and Google Kubernetes Engine

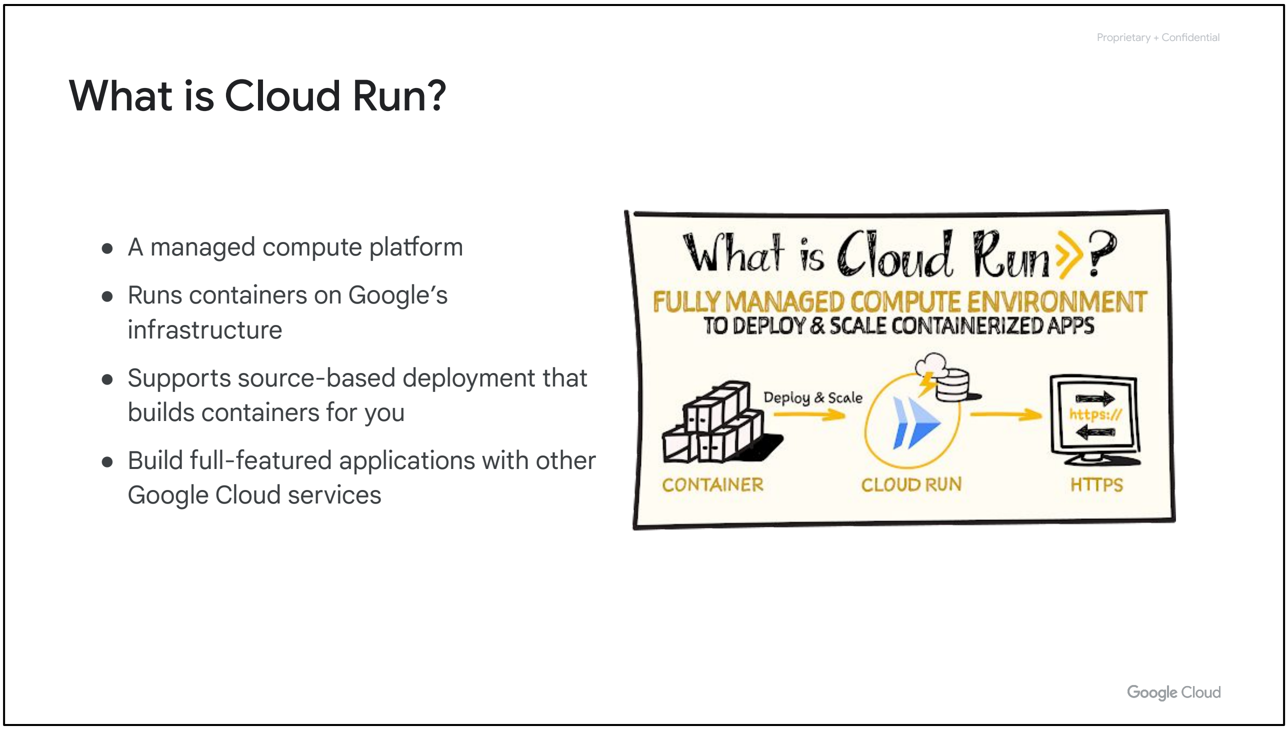
In this module, we introduce Cloud Run and Google Kubernetes Engine (GKE) for application development.



Agenda

01 Introduction to Cloud Run

Let’s first discuss what is Cloud Run.



What is Cloud Run?

- A managed compute platform
- Runs containers on Google’s infrastructure
- Supports source-based deployment that builds containers for you
- Build full-featured applications with other Google Cloud services

What is Cloud Run?
FULLY MANAGED COMPUTE ENVIRONMENT
TO DEPLOY & SCALE CONTAINERIZED APPS
Deploy & Scale
https://
CONTAINER
CLOUD RUN
HTTPS

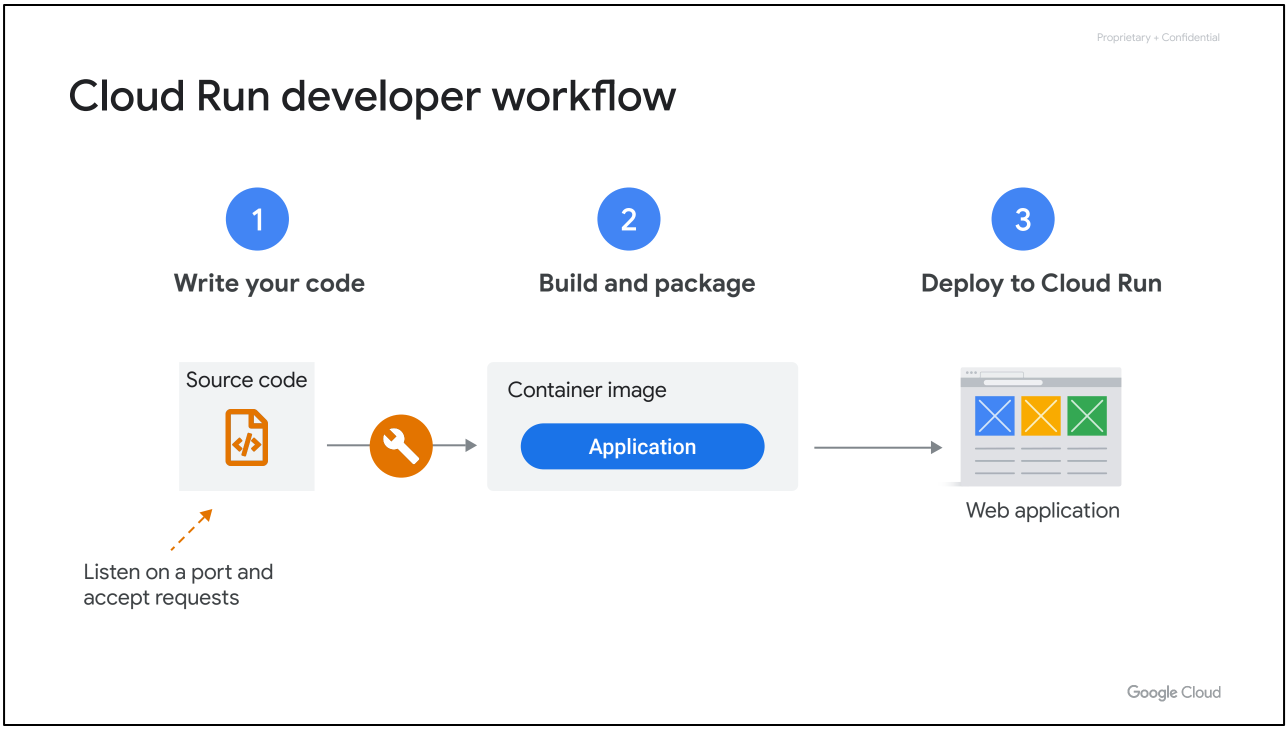
Google Cloud

Cloud Run is a fully managed compute platform that lets you deploy and run containers directly on top of Google's scalable infrastructure.

If you can build a container image of your application code written in any language, you can deploy the application on Cloud Run.

You can use the source-based deployment option that builds the container for you, when developing your application in Go, Node.js, Python, Java, .NET Core, or Ruby.

Cloud Run works well with other services on Google Cloud, so you can build full-featured applications without spending too much time operating, configuring, and scaling your Cloud Run service.



Cloud Run developer workflow

1 Write your code

2 Build and package

3 Deploy to Cloud Run

Source code

Container image

Application

Web application

Listen on a port and accept requests

Google Cloud

The Cloud Run developer workflow is a three-step process:

1. First, you write your application using your favorite programming language. This application should start a server that listens for web requests.
2. Second, you build and package your application into a container image.
3. Finally, you deploy the container image to Cloud Run.

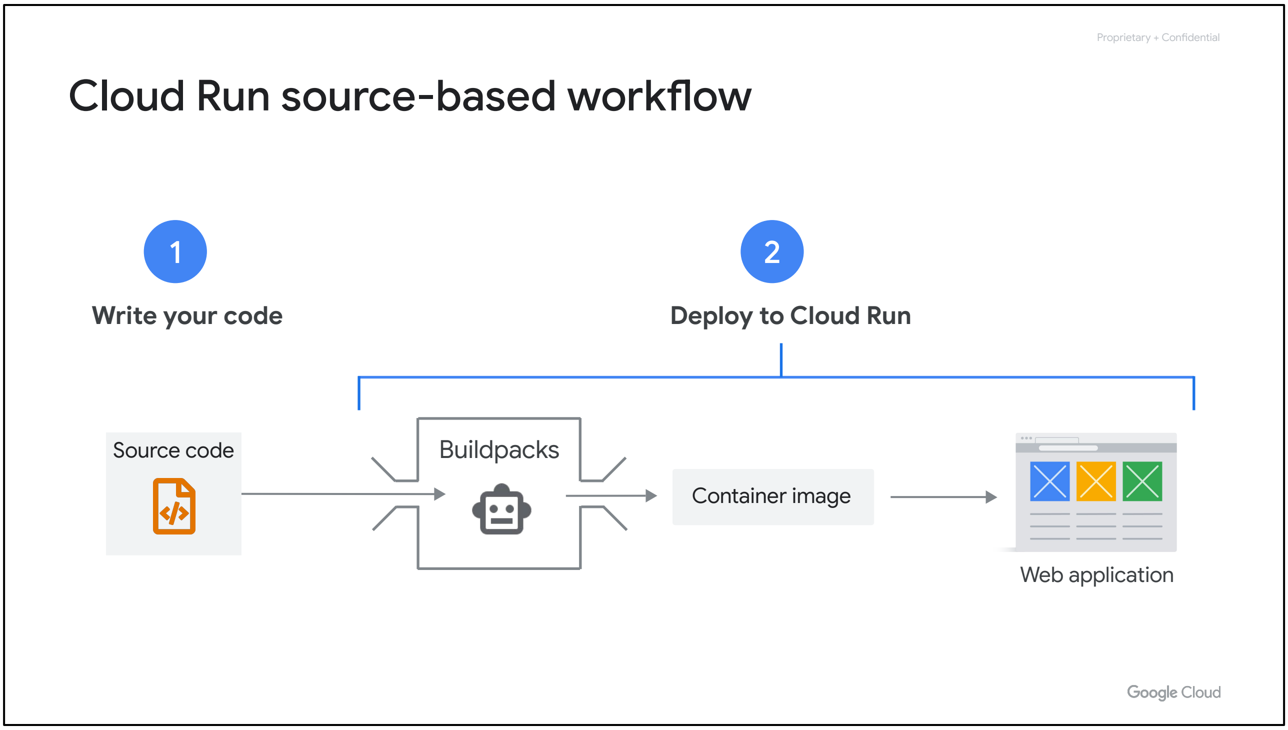
You can deploy your container to Cloud Run using the Google Cloud console, the gcloud CLI, or declaratively from a YAML configuration file.

Cloud Run also supports the deployment of services using Terraform, a popular open source tool that is used to create and manage infrastructure as code.

Once you've deployed your container, you get a unique HTTPS URL.

Cloud Run then starts your container on demand to handle requests, and ensures that *all* incoming *requests* are handled by dynamically adding and removing containers.

Cloud Run is serverless. That means that you, as a developer, can focus on building your application, and not on building and maintaining the infrastructure that powers your application.



Cloud Run source-based workflow

1 Write your code

2 Deploy to Cloud Run

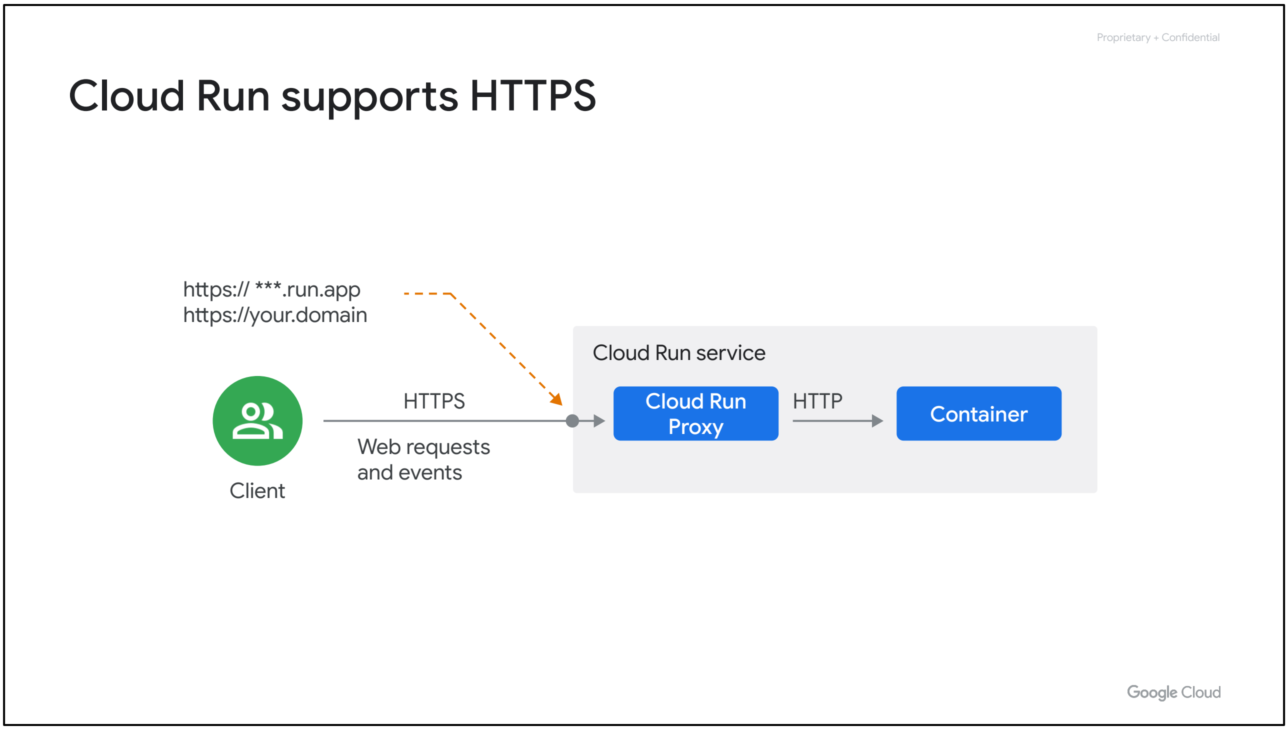
Source code → Buildpacks → Container image → Web application

For some use cases, a container-based workflow is great, because it gives you a great amount of transparency and flexibility.
If you build the container image yourself, you decide exactly which files are packaged in your container image.

However, building an application is hard enough already, and adding containerization adds additional work and responsibilities.
If you're just looking for a way to turn source code into an HTTPS endpoint by creating and deploying a containerized application in a secure, well-configured, and consistent manner, you can use Cloud Run.

With Cloud Run, you can use a container-based workflow, and a source-based workflow.

If you use the source-based approach, you deploy your source code, instead of a container image. Using Buildpacks, Cloud Run then builds your source, and packages the application along with its dependencies into a container image for you.



Proprietary + Confidential

Cloud Run supports HTTPS

https:// ***.run.app
https://your.domain

Client

HTTPS

Web requests
and events

Cloud Run service

Cloud Run Proxy

HTTP

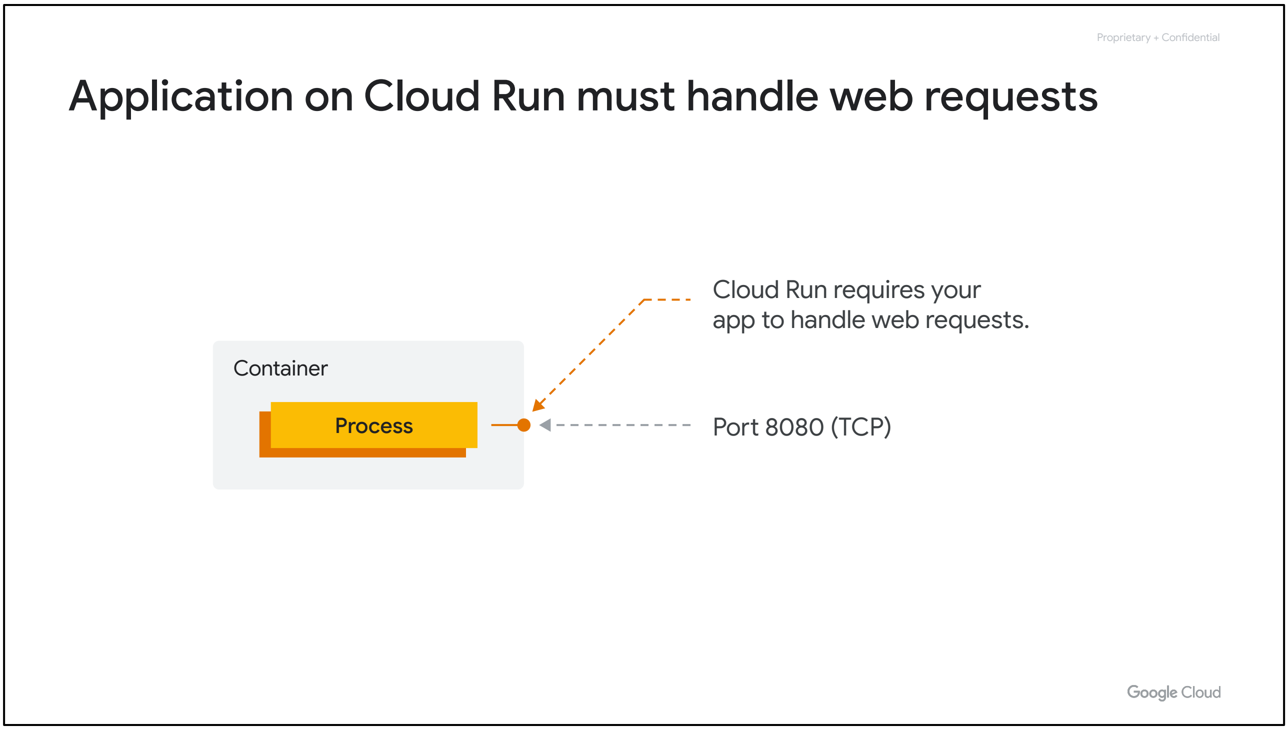
Container

Google Cloud

Cloud Run supports secure HTTPS requests to your application. On Cloud Run, your application can either run continuously as a service or as a job. Cloud Run services respond to web requests, or events, while jobs perform work and quit when that work is completed.

Cloud Run:

- Provisions a valid TLS certificate, and other configuration to support HTTPS requests.
- Handles incoming requests, decrypts, and forwards them to your application.



Application on Cloud Run must handle web requests

Container

Process

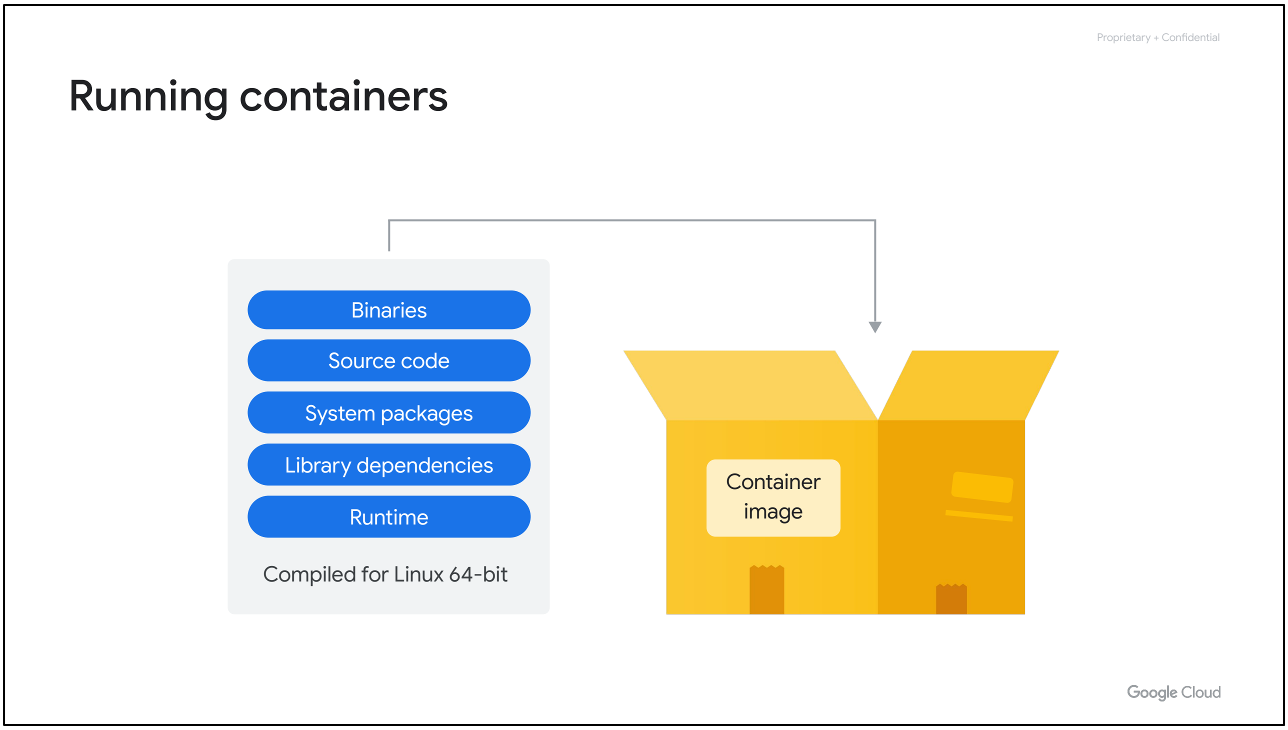
Cloud Run requires your app to handle web requests.

Port 8080 (TCP)

A process running in a container has access to its own private, virtual network stack, so your application can always open a port and listen for incoming connections.

Cloud Run expects your container to listen on port number 8080 to handle web requests. The port number 8080 is a configurable default, so if this port is unavailable to your application, you can change the application’s configuration to use a different port.

You don’t need to provide an HTTPS server, Google’s infrastructure handles that for you.



Proprietary + Confidential

Running containers

Binaries

Source code

System packages

Library dependencies

Runtime

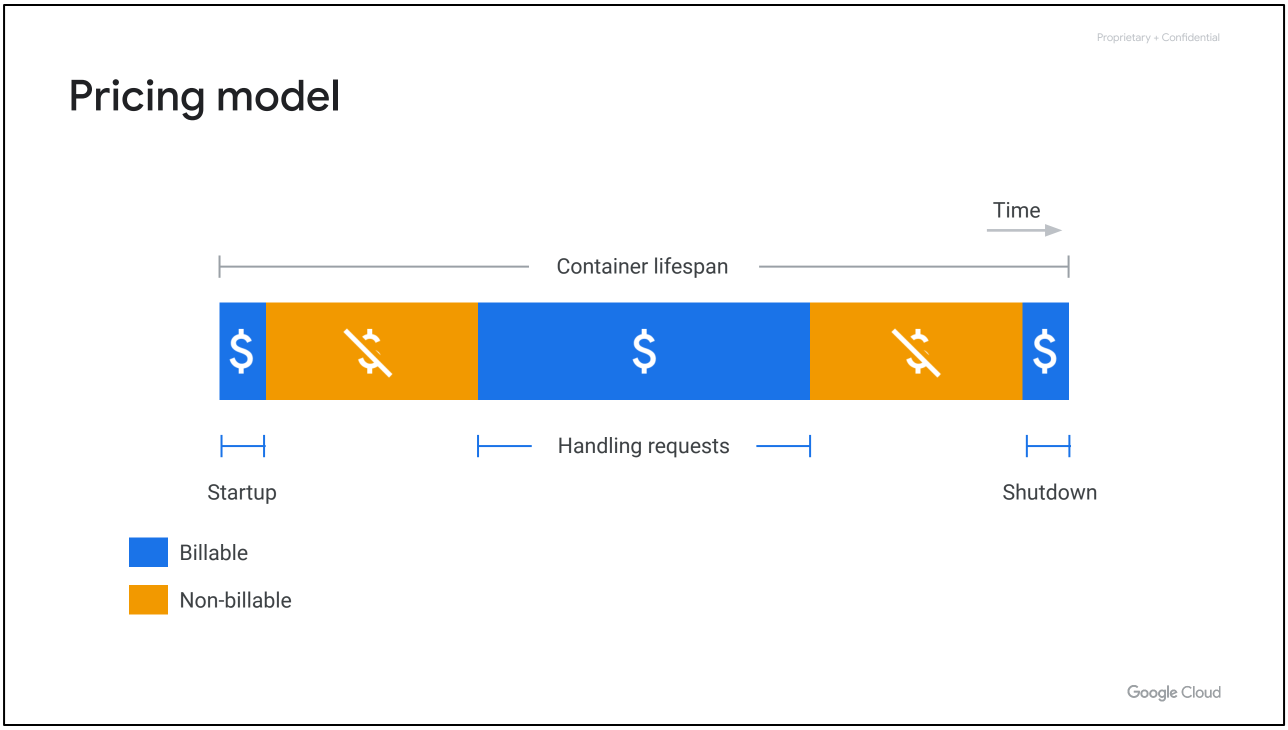
Compiled for Linux 64-bit

Container image

Google Cloud

One major advantage of Cloud Run is that it runs containers.

This means you can develop your applications in any programming language and run them on Cloud Run, as long as they can be compiled to a 64-bit Linux binary, and packaged in a container image.



Pricing model

Time

Container lifespan

Handling requests

Startup

Shutdown

🟦 Billable

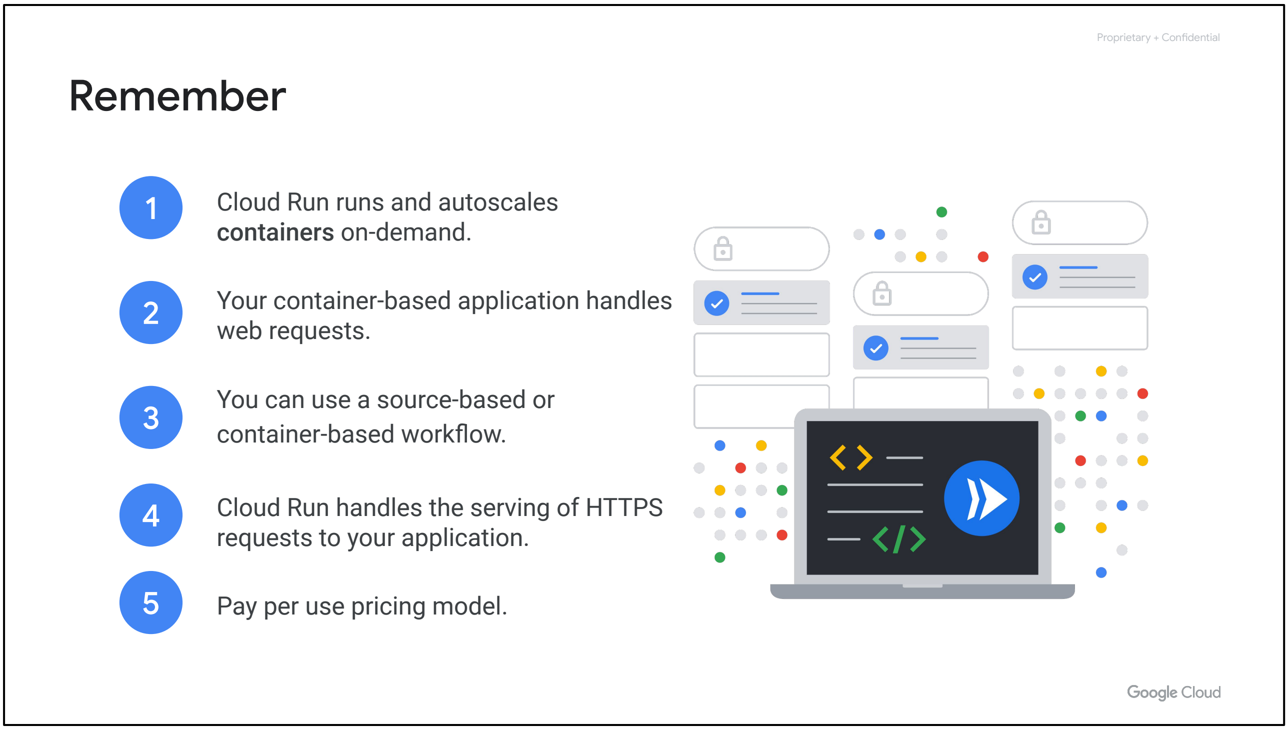
🟧 Non-billable

The Cloud Run pricing model is unique; as you only pay for the system resources that are used while a container is handling requests, and when it's starting or shutting down.

Cloud Run also supports a pricing model which charges you for the entire container lifecycle, with CPU always allocated to container instances even when there are no requests to your application. This model may be more economical for most steady state workloads.

The price of container time increases based on the number of vCPUs and memory allocated for the container.

For more information on the Cloud Run pricing model, view the documentation.



Remember

1. Cloud Run runs and autoscales **containers** on-demand.
2. Your container-based application handles web requests.
3. You can use a source-based or container-based workflow.
4. Cloud Run handles the serving of HTTPS requests to your application.
5. Pay per use pricing model.

Google Cloud

In summary:

- Cloud Run is a managed serverless product on Google Cloud that runs and autoscales containers on-demand.
- You can deploy any containerized application that handles web requests.
- You can employ a source-based or container-based workflow.
- Cloud Run handles HTTPS requests to your application.
- With the Cloud Run pricing model, you only pay for what you use.