Working with Containers

Memi Lavi
www.memilavi.com

Working with Containers

- Containers have become extremely popular in recent years
- Allow isolation, host density, single code base
- App Engine Flexible Environment works with containers

Working with Containers

- GCP supports two services for running containers:

Cloud Run

Google Kubernetes Engine

- First let's learn more about containers

Containers

- Traditional deployment:
  - Code was copied and built on the production server
  - Problems were found on the servers that weren't found in the dev machines

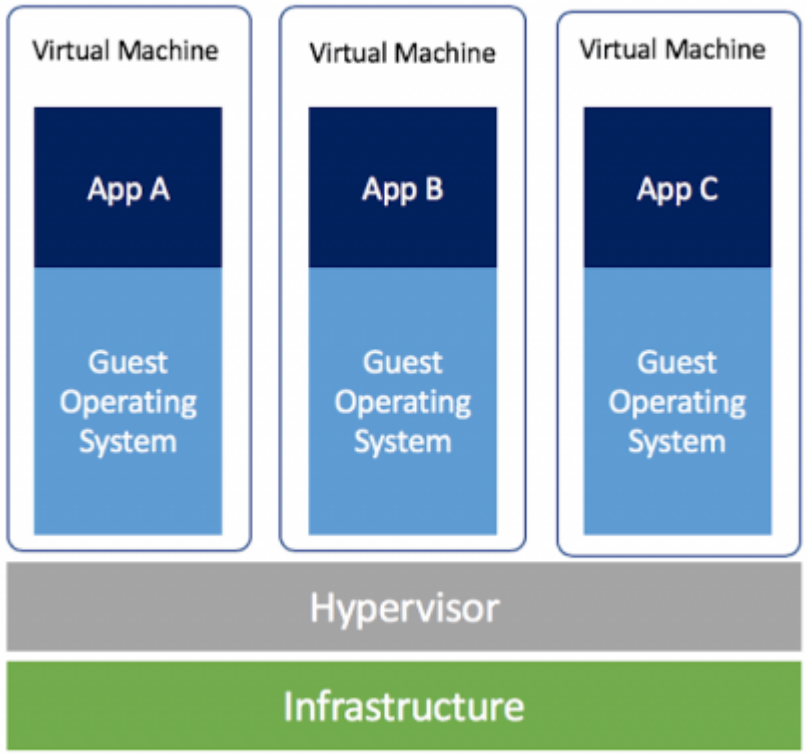
Containers to the Rescue!

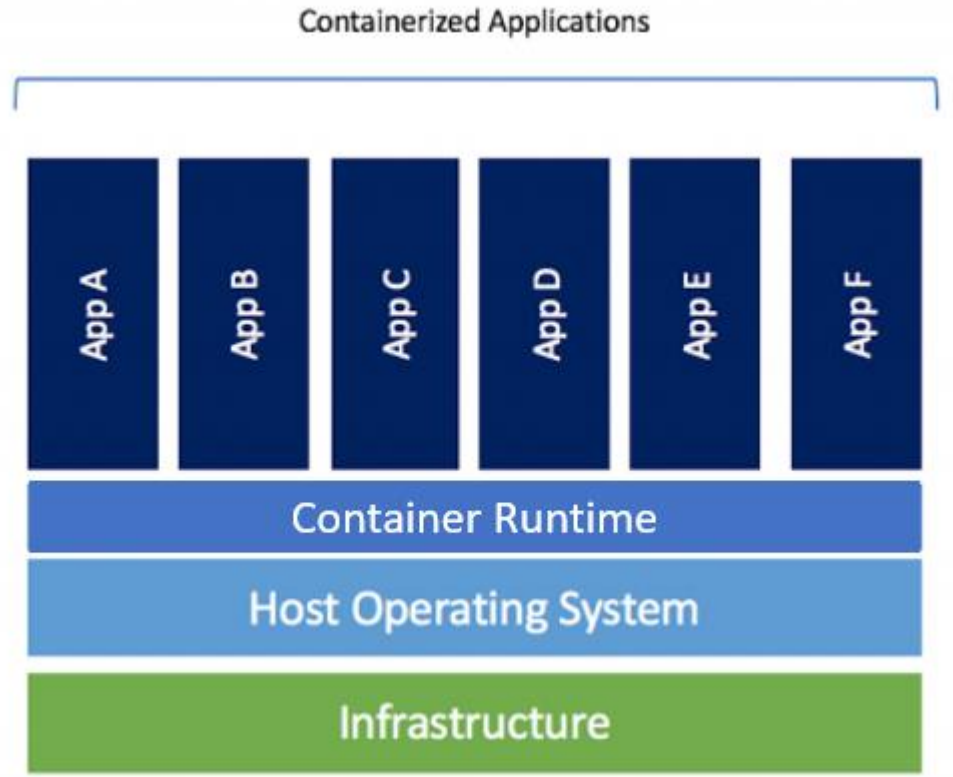
http://www.developermemes.com/2013/12/23/defect-production-works-machine/

Containers

- Thin packaging model
- Packages software, its dependencies, and configuration files
- Can be copied between machines
- Uses the underlying operating system

Container vs VM





Why Containers?

Predictability	The same package is deployed from the dev machine to the test to production
Performance	Container goes up in seconds vs minutes in VM
Density	One server can run thousands of containers vs dozens of VMs

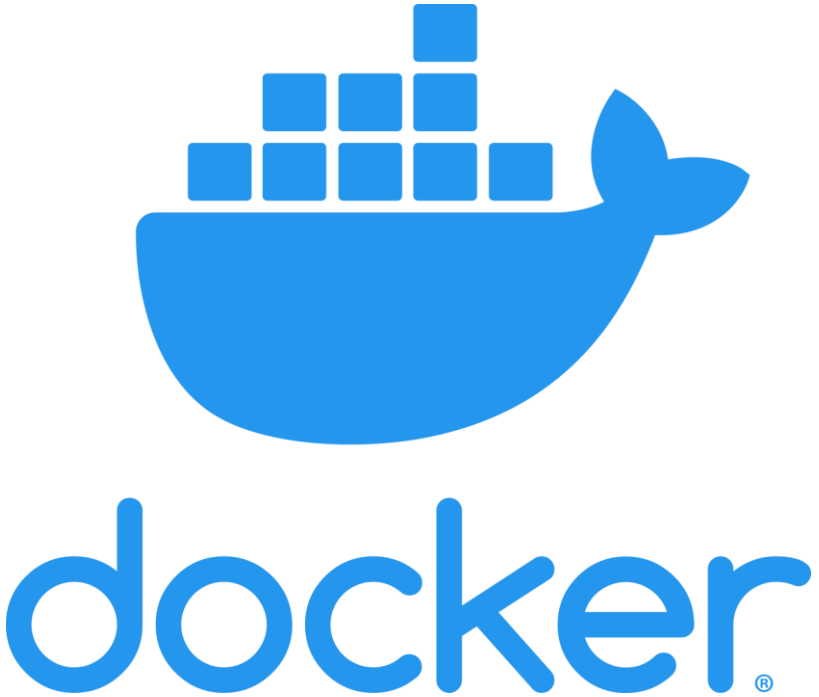
Why Not Containers?

Isolation

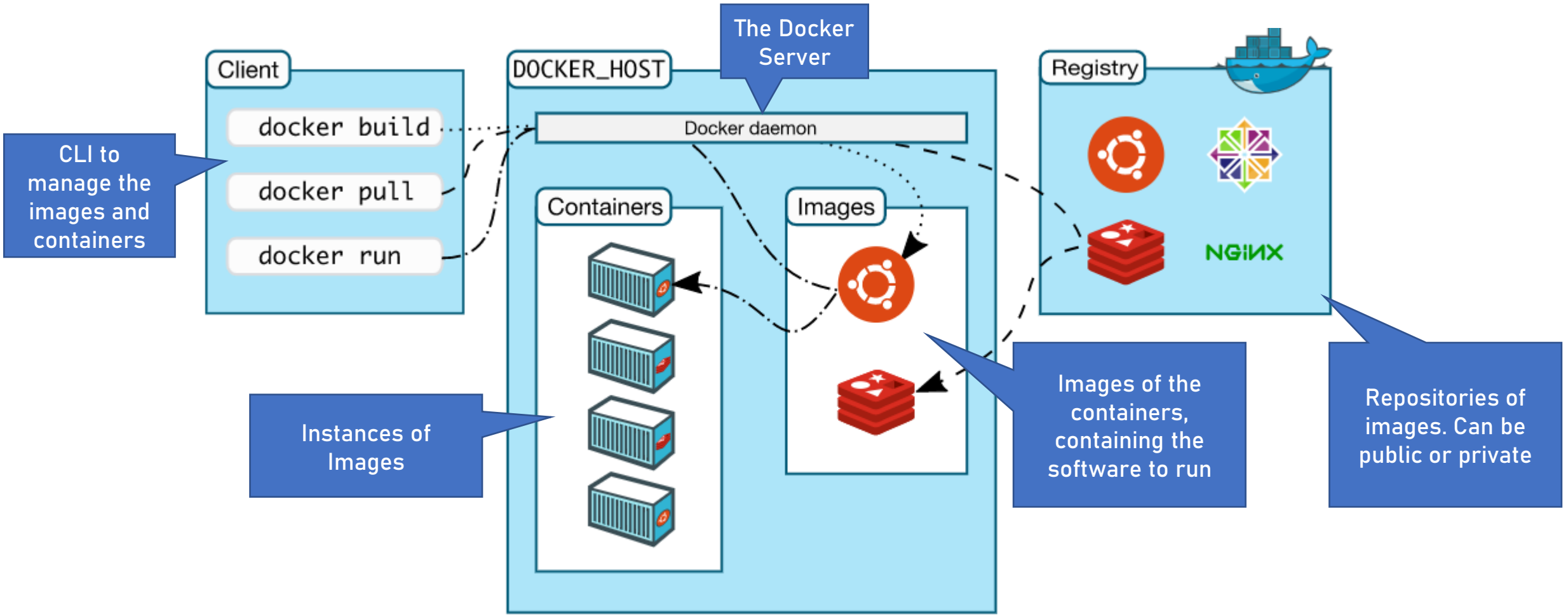
Containers share the same OS, so isolation is lighter than VM

Docker

- The most popular container environment
- De-facto standard for containers
- Released in 2013



Docker Architecture



https://docs.docker.com/get-started/overview/

dockerfile

- Contains instructions for building custom images

```
1  WORKDIR /opt/node_app
2  COPY package.json package-lock.json* ./
3  RUN npm install --no-optional && npm cache clean --force
4  ENV PATH /opt/node_app/node_modules/.bin:$PATH
5  WORKDIR /opt/node_app/app
6  COPY . .
```

https://www.docker.com/blog/keep-nodejs-rockin-in-docker/

Artifact Registry

- GCP registry for Docker images and software packages
- Reliable and fast
- Easy to maintain
- Regional and multi-regional
- Can store artifacts from Cloud Build

Artifact Registry Supported Formats

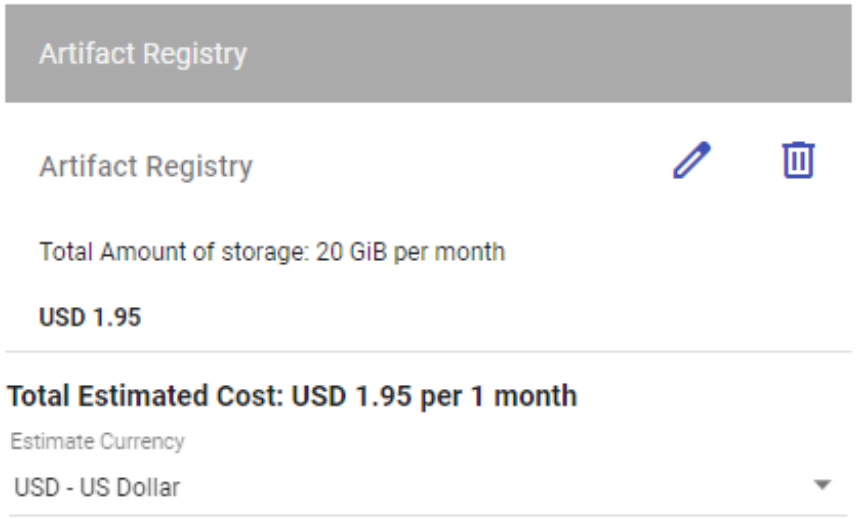
- Docker images
- Language packages: `Java, Node.js, Python, Go`
- OS Packages: `Debian, RPM`

Artifact Registry Integration

- Integrates with:
  - Compute Engine
  - App Engine Flexible Environment
  - Cloud Run
  - Google Kubernetes Engine (GKE)

Artifact Registry Pricing

- Price per storage
- First 0.5GB is free
- After that - $0.10/GB/month



Cloud Run

- Managed compute platform for running containers
- Pulls images from various registries
  - Best integration with Artifact Registry
- Auto scaling
- Traffic management between revisions
- SLA 99.95%

Cloud Run

- Can be public or private
  - Public: Accessible from the internet
  - Private: Accessible only from internal resources

Cloud Run

- Supports Services and Jobs

Services

- Respond to web requests
- Used for web apps, web APIs, etc.

Jobs

- Run code that performs a work
- Quit when the work is done
- Do not listen to outside requests
- Can be scheduled
- Used for data processing, migrations, etc.

Cloud Run CPU Allocation

- Two types of CPU allocation:

Only during requests

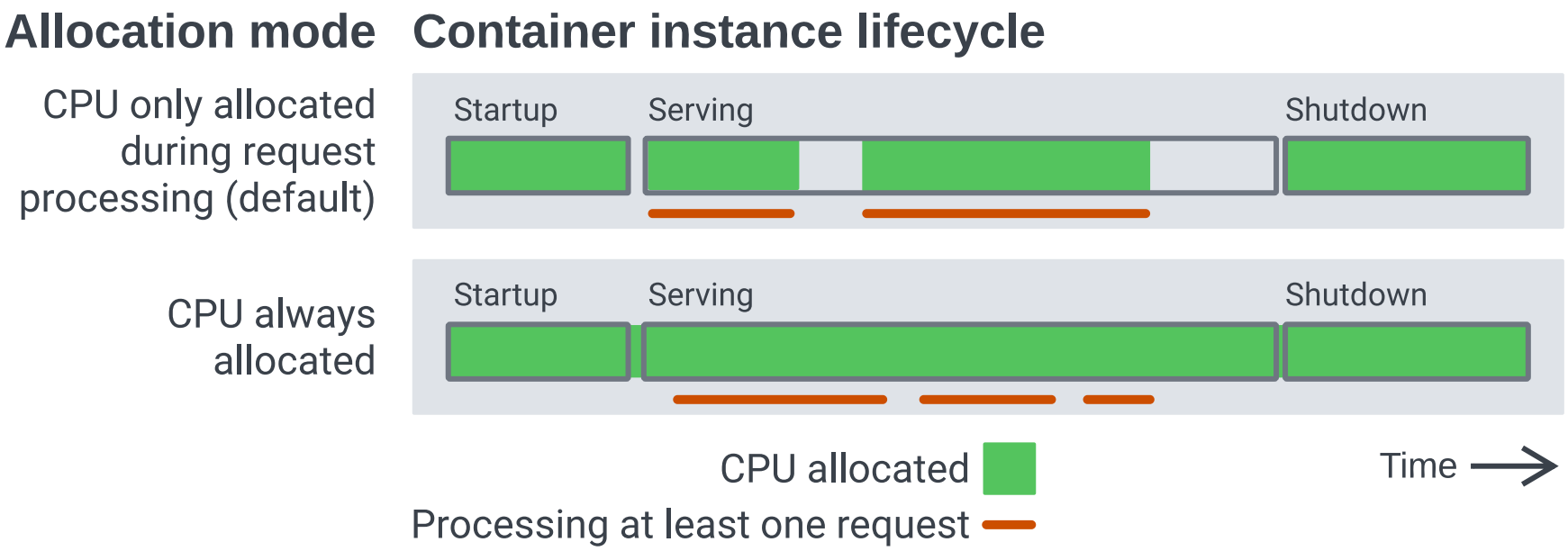
- CPU is allocated to the instance only when requests are incoming

Always allocated

- CPU is always allocated even when there are no incoming requests

Cloud Run CPU Allocation

- Two types of CPU allocation:

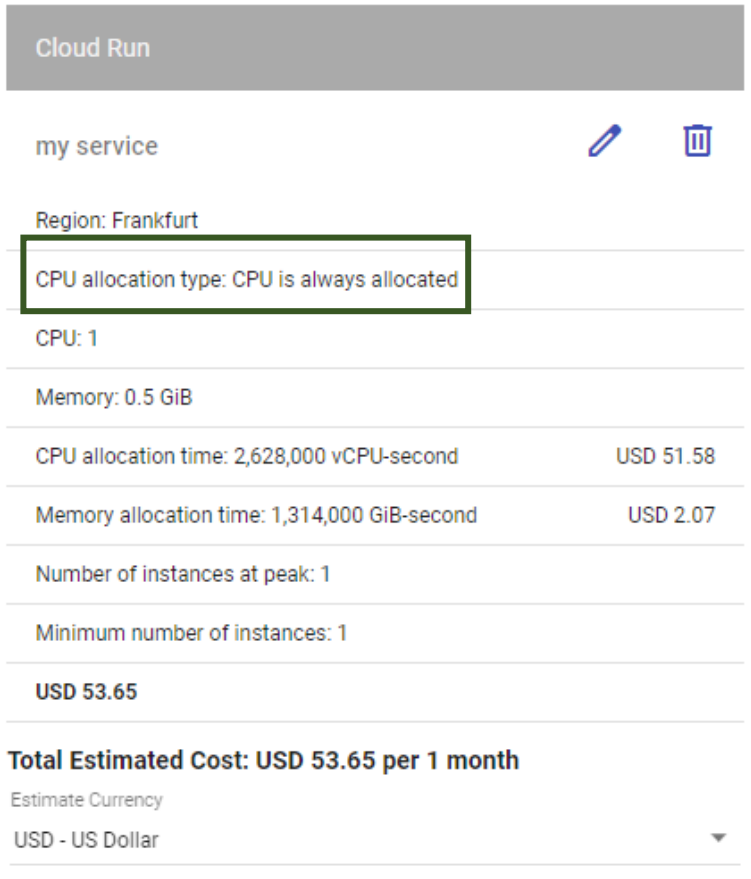


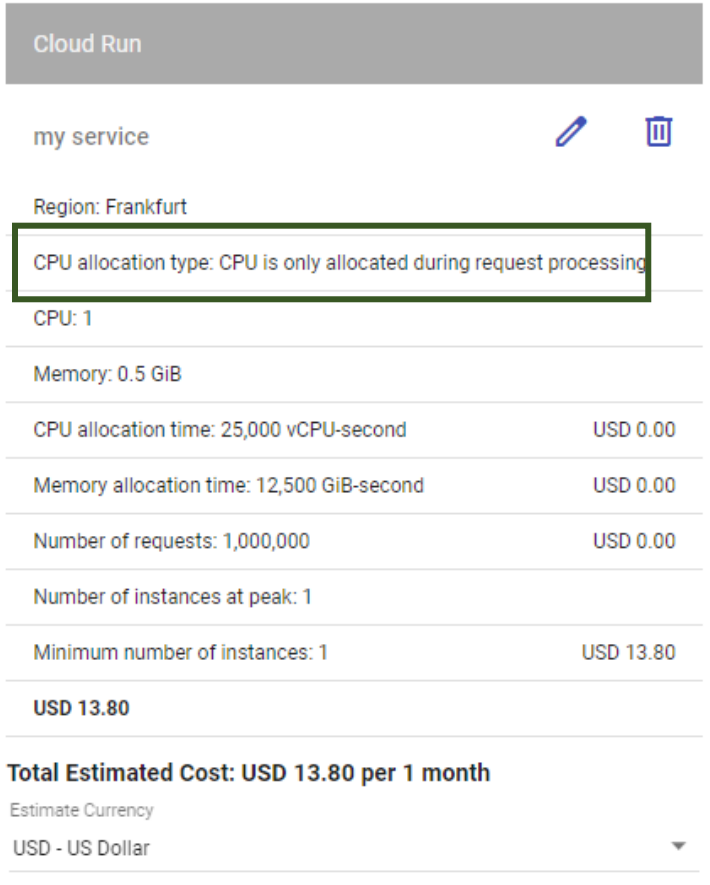
Cloud Run CPU Allocation

- Affects:
  - Background processing (requires always allocated)
  - Pricing

Cloud Run Pricing

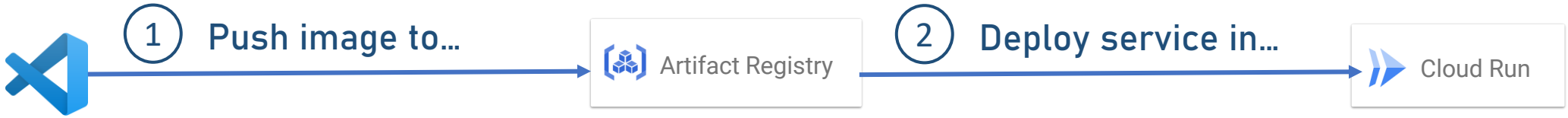
- Depends on the CPU Allocation configuration





Deploying from Source

- So far we deployed to Cloud Run in two steps:



Deploying from Source

- With Deploy from Source we can do that in a single step:

Deploying from Source

- Behind the scenes:
  - A new repo will be created in Artifact Registry
    - Named `cloud-run-source-deploy`
  - Image is pushed to the repo
  - Cloud Run pulls from the repo and runs the image

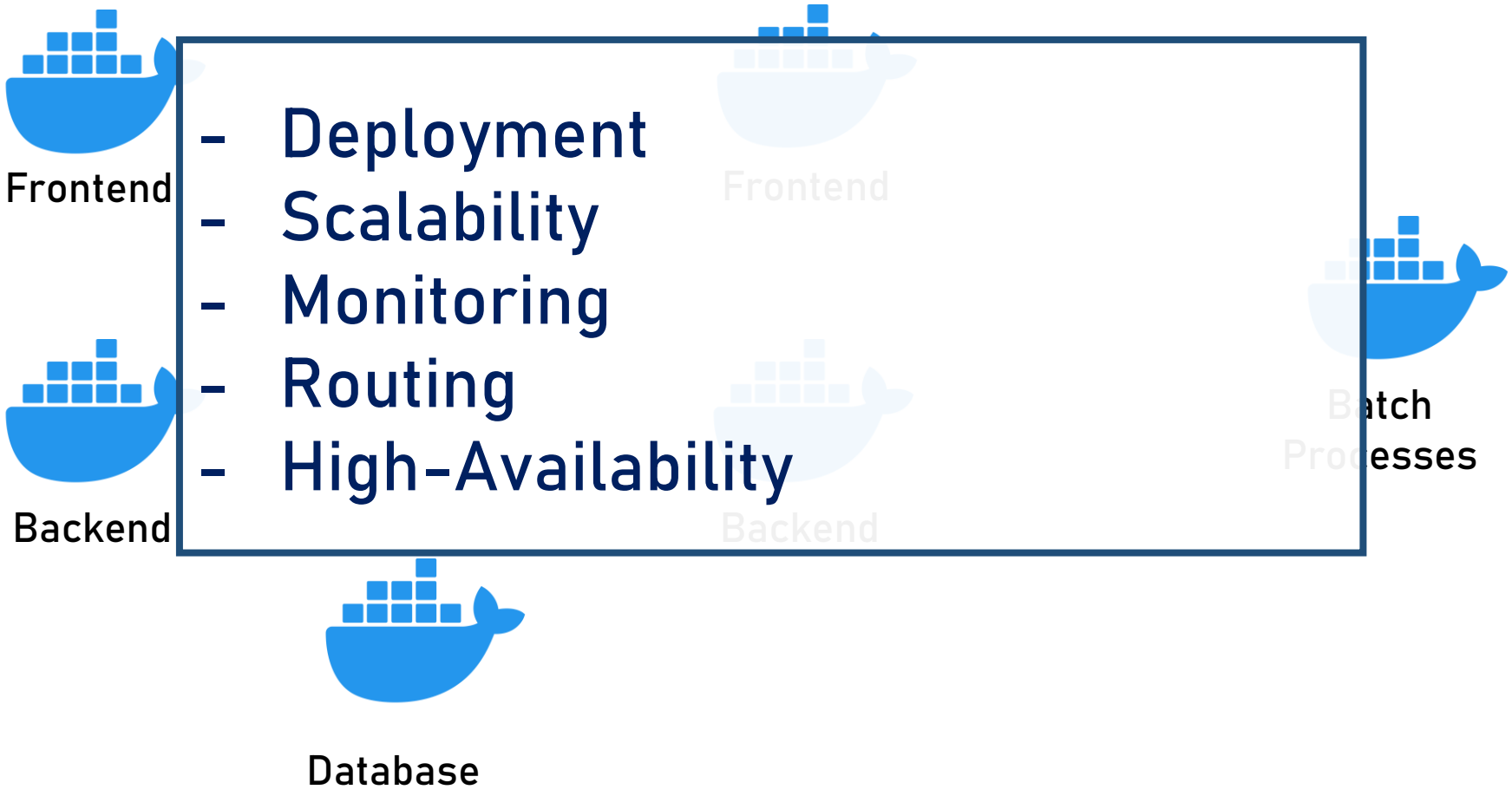
Cloud Run Jobs

- Cloud Run can be used to run also jobs
- A unit of work that does something until work is done
- Can be scheduled

Containers Management

- Containers are a great deployment mechanism
- Gain popularity
- What happens when there are too many of them?

Containers Management



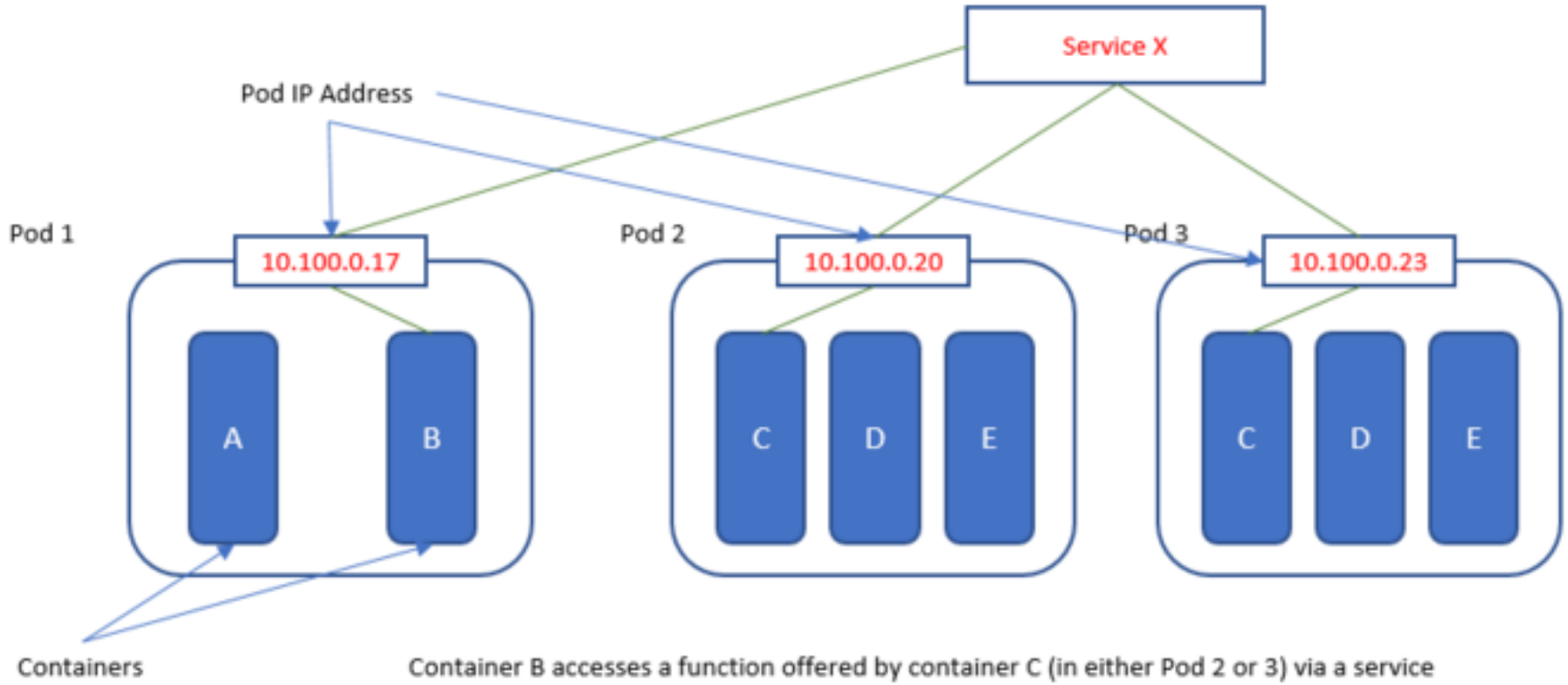
Kubernetes

- The most popular container management platform
- De-facto standard for container management
- Released by Google in 2014

Kubernetes

- Provides all aspects of management:
  - Routing
  - Scaling
  - High-Availability
  - Automated Deployment
  - Configuration Management
  - And more…

Kubernetes Architecture



https://en.wikipedia.org/wiki/Kubernetes

GKE

- Google Kubernetes Engine
- Fully managed Kubernetes environment in the cloud
- Manages the various Kubernetes components:
  - Cluster
  - Nodes
  - Pods
  - Etc.

GKE

- Major capabilities:
  - Autoscaling
  - Platform management (upgrades, maintenance etc.)
  - Security
  - Auto-repair
  - Logging and monitoring
  - SLA: 99%

GKE Editions and Modes

- GKE has two editions:

Standard

- All basic capabilities
- Runs on a single cluster
- Autoscaling
- Logging and monitoring

Enterprise

- All the Standard capabilities
- Manages multi-cluster deployments
- Team management
- Policy management
- Built-in Service Mesh
- Works also on AWS and Azure
- Built on Anthos

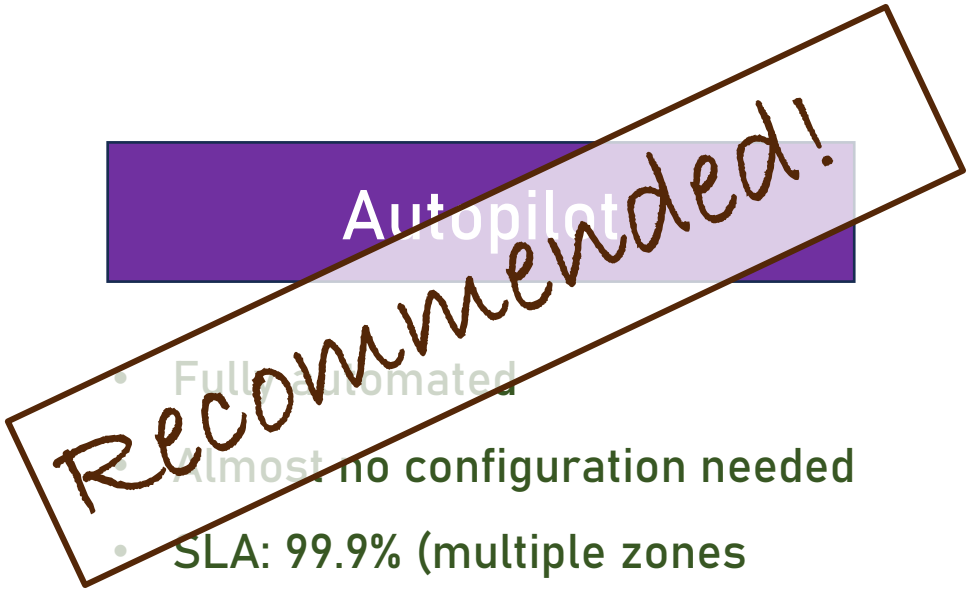
GKE Editions and Modes

- GKE runs on two modes:

Standard

- Manual configuration of:
  - Cluster
  - Nodes
  - Security
  - Scheduling
  - Scaling
  - And more…

Autopilot



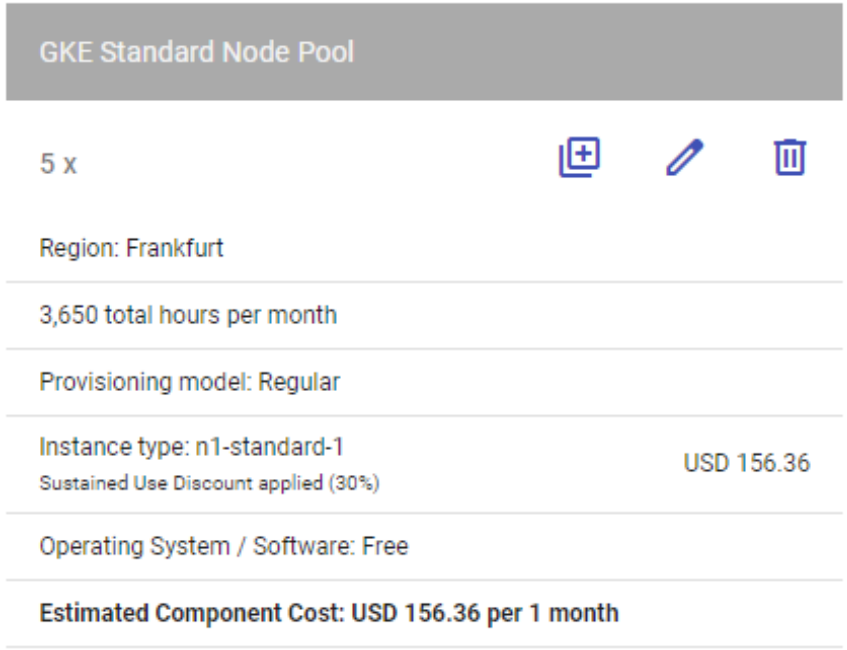
- Fully automated
- Almost no configuration needed
- SLA: 99.9% (multiple zones deployment)

GKE Pricing

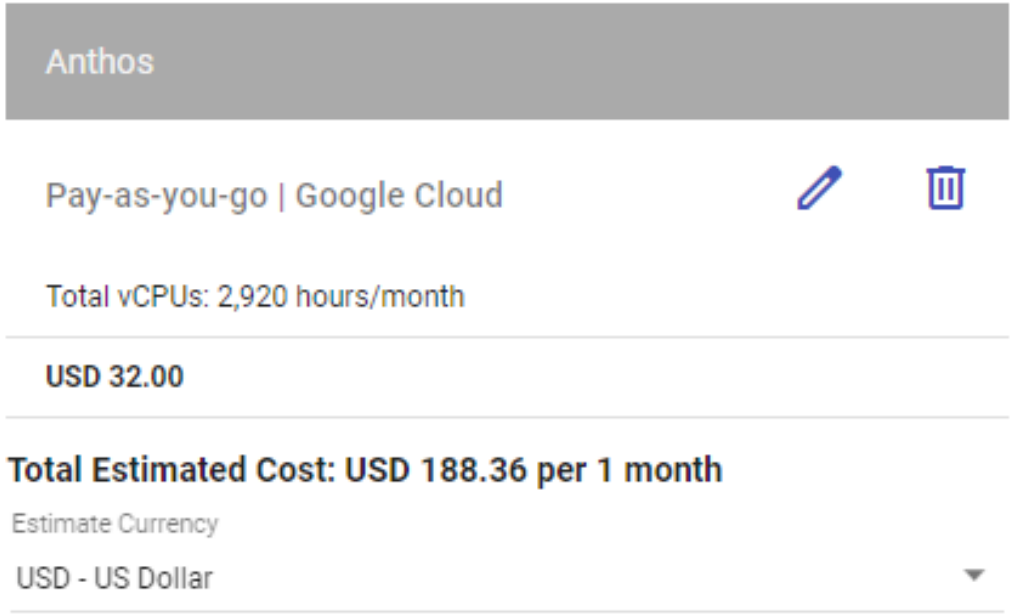
- Depends on the mode and edition:

Standard	Price of the underlying compute engine instances
Enterprise	Anthos license
Autopilot	Pay for resources actually used (vCPU, memory, disk)

GKE Pricing


GKE Standard Node Pool
5 x
Region: Frankfurt
3,650 total hours per month
Provisioning model: Regular
Instance type: n1-standard-1
Sustained Use Discount applied (30%)
USD 156.36
Operating System / Software: Free
Estimated Component Cost: USD 156.36 per 1 month

GKE Pricing


Anthos
Pay-as-you-go | Google Cloud
Total vCPUs: 2,920 hours/month
USD 32.00
Total Estimated Cost: USD 188.36 per 1 month
Estimate Currency
USD - US Dollar

GKE Pricing

General-purpose (default) Pods

Iowa (us-central1)

Item	Regular price
GKE Autopilot vCPU Price (vCPU)	$32.485
GKE Autopilot Pod Memory Price (GB)	$3.593425
GKE Autopilot Ephemeral Storage Price (GB)	$0.040004

Prices per month. Actual billing is per second.

ReadIt!

Architecture: ReadIt Cloud System