Best Of technozaure

💻 + 🚀 + ☁️ = 💸

# Introduction au FinOps avec Google Cloud

# Julien Landuré

Cloud Practice Leader

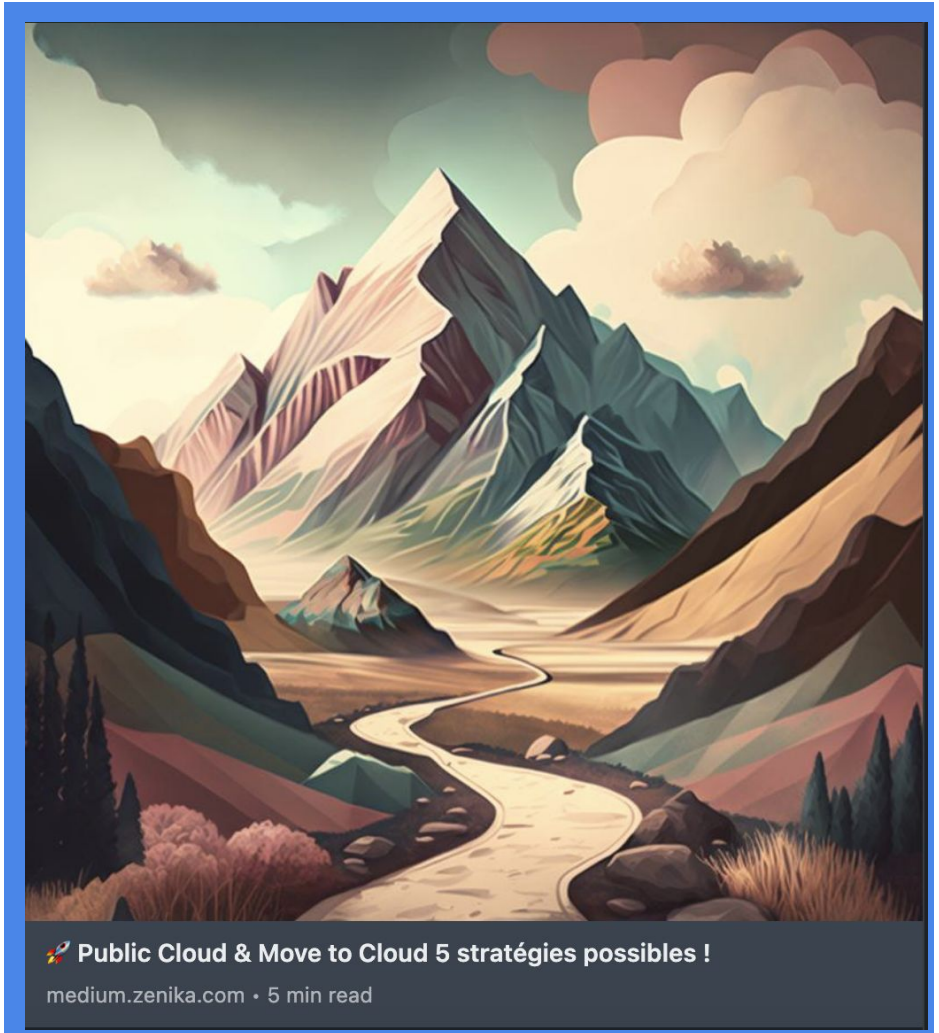
**GDG & DevFest Nantes Organizer**
**GDE Cloud**

jlandure.dev
@jlandure

# Back to the future

# 5 stratégies de déploiement

Rehost
Replatform
Refactor
Rebuild
Replace



# Compute Spectrum


IaaS PaaS SaaS

Servers
VM instances

Clusters
Cluster management

Serverless, autoscaling

Compute Engine

Kubernetes Engine

App Engine

Cloud Run

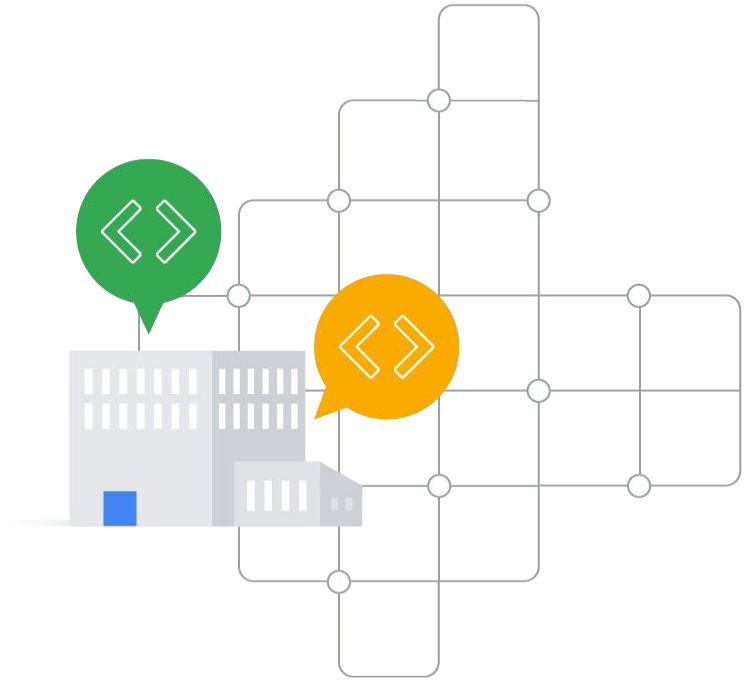
Cloud Functions

Google Cloud

# Payment models

## IaaS

Pay for what they allocate

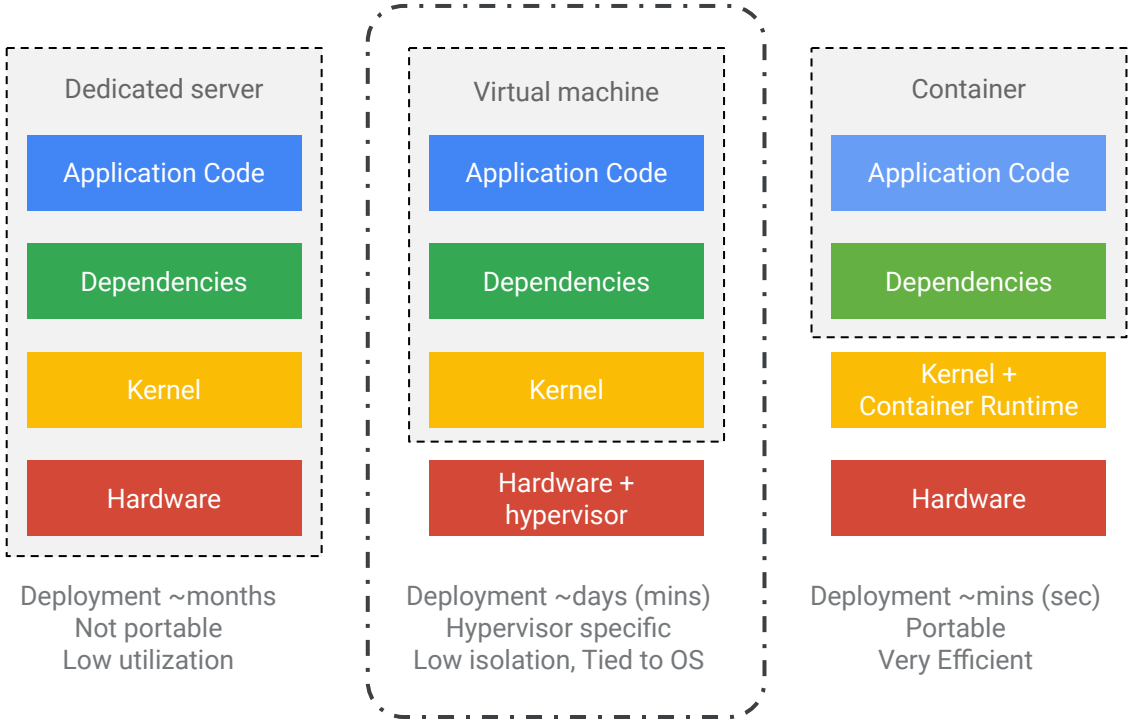


## PaaS

Pay for what they use

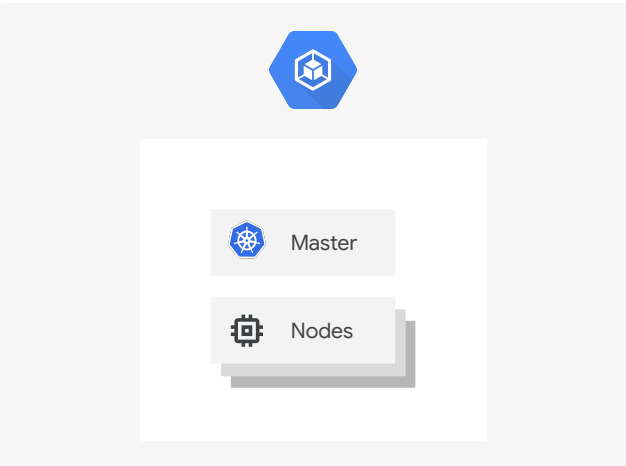
# Ce qui existe en standard…
# “Pay-as-you-go” model

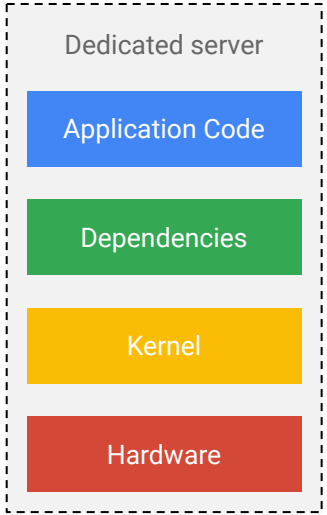
# VM World



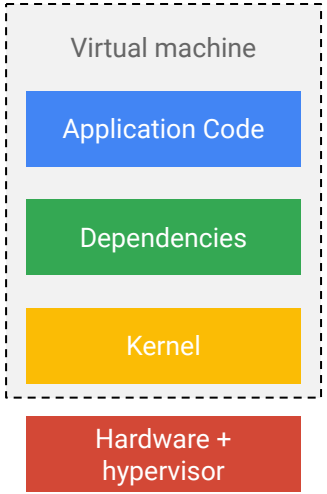

Google Cloud

# Container World (Kubernetes)

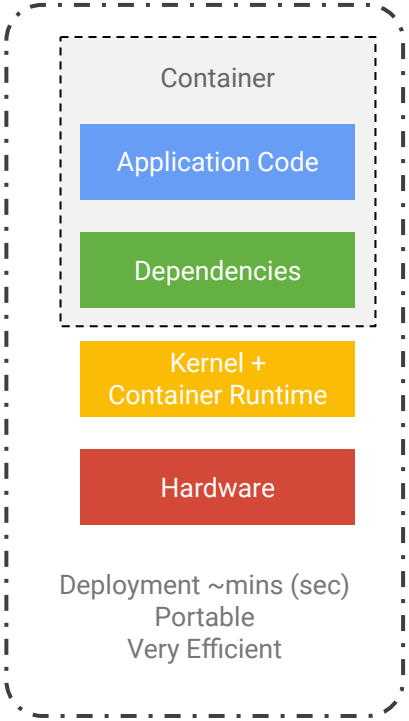



Deployment ~months
Not portable
Low utilization



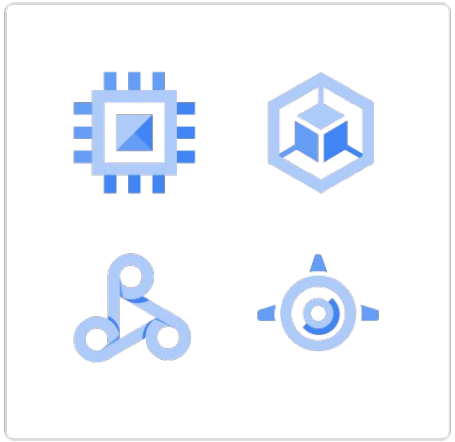

Deployment ~days (mins)
Hypervisor specific
Low isolation, Tied to OS




Deployment ~mins (sec)
Portable
Very Efficient

Google Cloud

# Google Compute products are billed per-second



# What is Compute Engine?

Per VM max:

- 416 vCPUs, 12 TB of RAM, GPUs, TPUs
- Independently configurable, including disk!

Start quickly, ~20 seconds to user code running

Pre-built images, or create your own

- Debian, CentOS, CoreOS, Ubuntu, RHEL, SUSE
- Windows Server Datacenter: 2008, 2012 R2, 2016 , 2019

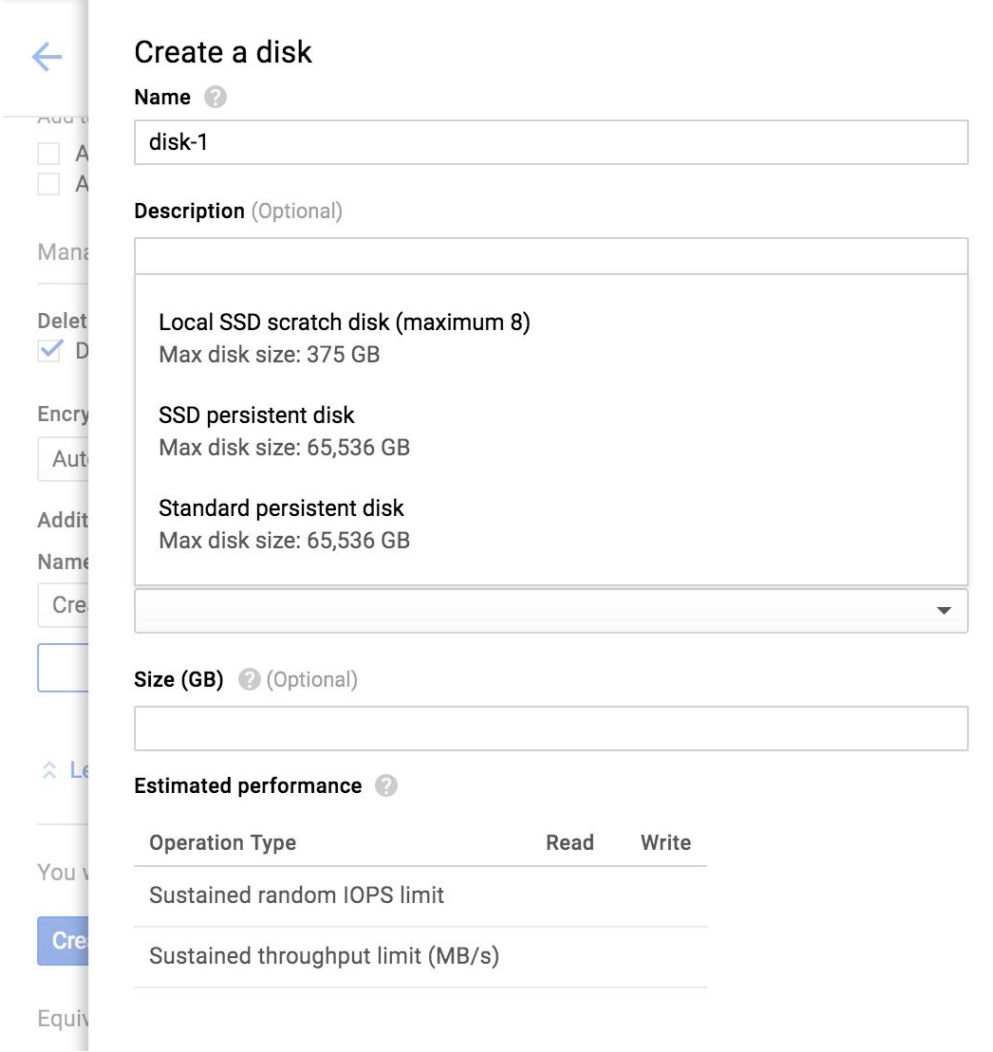
# Disk

**Persistent Disk**

- Network based
- Configure size (to 64 TB)
- Disk or SSD
- Independent of VM
  - Global Snapshots
  - Multi-VM RO mount
- Larger disks == more throughput

**Local SSD**

- Lowest latency
- 375 GB per disk (up to 8)
- Tied to VM lifecycle




Google Cloud

# GCE
# Machine Types

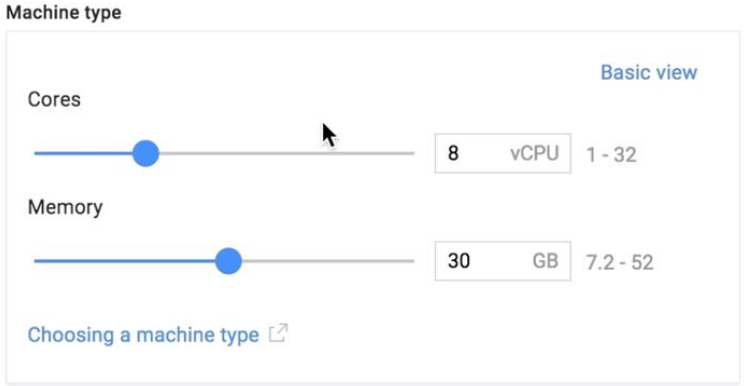
| Workload type | | | | | |
|---|---|---|---|---|---|
| General-purpose workloads | | | Optimized workloads | | |
| Cost-optimized | Balanced | Scale-out optimized | Memory-optimized | Compute-optimized | Accelerator-optimized |
| E2 | N2, N2D, N1 | Tau T2D | M2, M1 | C2, C2D | A2 |
| Day-to-day computing at a lower cost | Balanced price/performance across a wide range of VM shapes | Best performance/cost for scale-out workloads | Ultra high-memory workloads | Ultra high performance for compute-intensive workloads | Optimized for high performance computing workloads |
| • Web serving<br>• App serving<br>• Back office apps<br>• Small-medium databases<br>• Microservices<br>• Virtual desktops<br>• Development environments | • Web serving<br>• App serving<br>• Back office apps<br>• Medium-large databases<br>• Cache<br>• Media/streaming | • Scale-out workloads<br>• Web serving<br>• Containerized microservices<br>• Media transcoding<br>• Large-scale Java applications | • Medium-large in-memory databases such as SAP HANA<br>• In-memory databases and in-memory analytics<br>• Microsoft SQL Server and similar databases | • Compute-bound workloads<br>• High-performance web serving<br>• Gaming (AAA game servers)<br>• Ad serving<br>• High-performance computing (HPC)<br>• Media transcoding<br>• AI/ML | • CUDA-enabled ML training and inference<br>• HPC<br>• Massive parallelized computation |

Google Cloud

# Custom Machine feature

## Config your own custom machine, with average 19% savings

- Flexible configurations
- Create a machine type with 1 vCPU and up to 96 vCPUs
- Extended Memory enables higher Mem:vCPU ratios
- Sizing recommendation




```
// CREATE INSTANCE WITH 6 vCPUs and 5 GB MEMORY
gcloud compute instances create my-vm --custom-cpu 6
--custom-memory 5
```

Google Cloud

# Preemptible / Spot VMs

**Short lived, low cost virtual machines.**

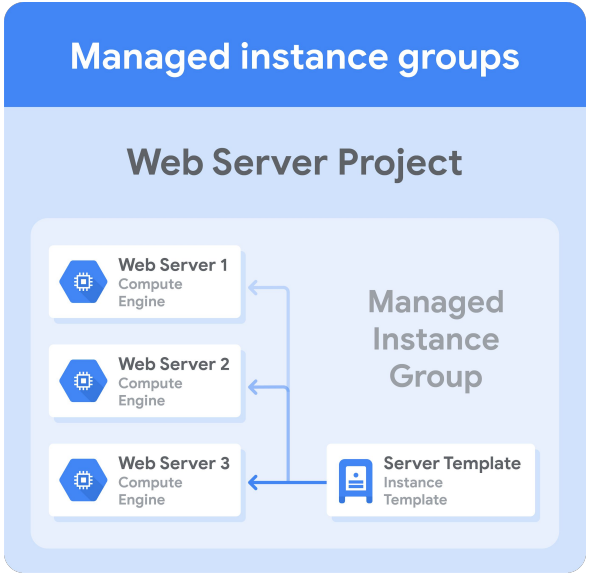
- Can run up to 24 hours
- Might be preempted at any time with 30 seconds notification
- Much lower price (up to 80% cheaper) than regular instances
- Offers the same machine types and options as regular VM instances
- Best for running non-urgent fault-tolerant jobs or short-lived tasks, like genomics, media transcoding, etc.


Google Cloud

# Managed instance groups (MIGs)

**A collection of identical VM instances that you control as a single entity, using an instance template.**

- Instance group can be resized
- Manager ensures all instances are RUNNING
- Typically used with **autoscaler**
- Can be single zone or regional
- Supports **autohealing**

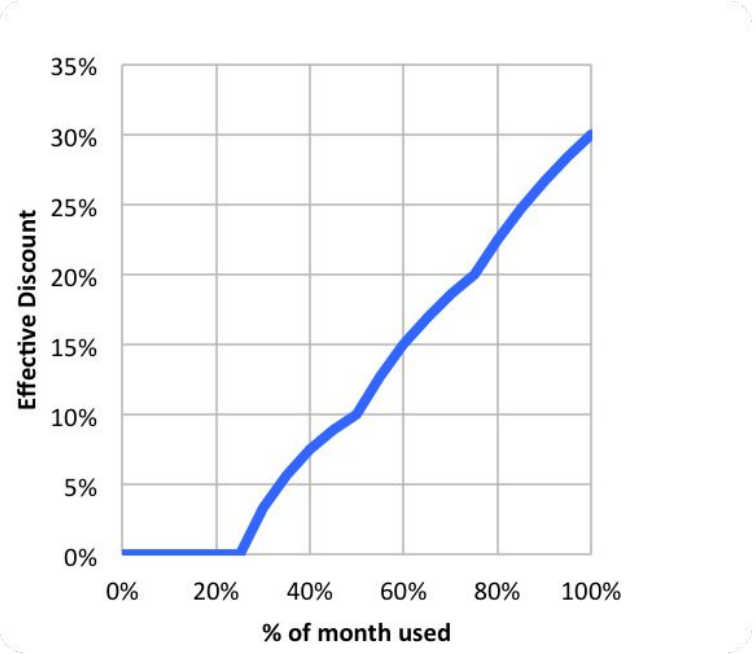



Google Cloud

# Sustained use discounts

**Up to 30%** net discount for instances that run the entire month.

| Usage Level (% of month) | % at which incremental is charged |
|---|---|
| 0% - 25% | 100% of base rate |
| 25% - 50% | 80% of base rate |
| 50% - 75% | 60% of base rate |
| 75% - 100% | 40% of base rate |




Google Cloud

# Ce qui existe en standard…
# “Pay-as-you-use” model

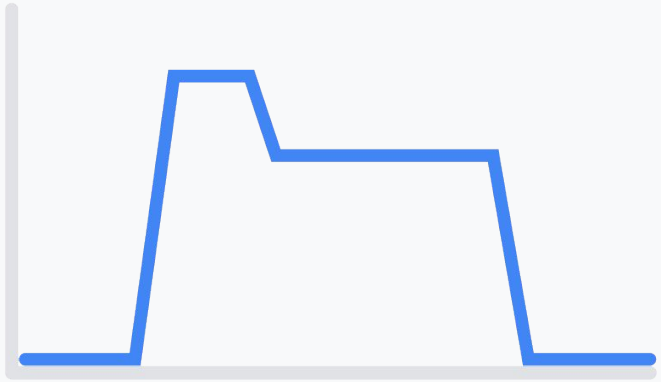
# Serverless

Deploy and scale applications fast and securely in a fully managed environment

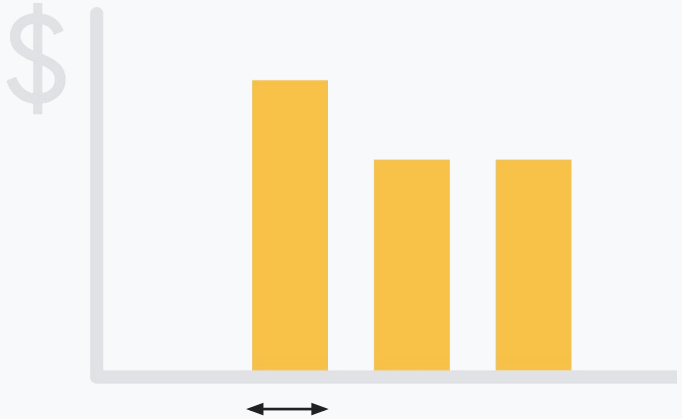
Auto-scaling

No Infra Management

# Pay per use

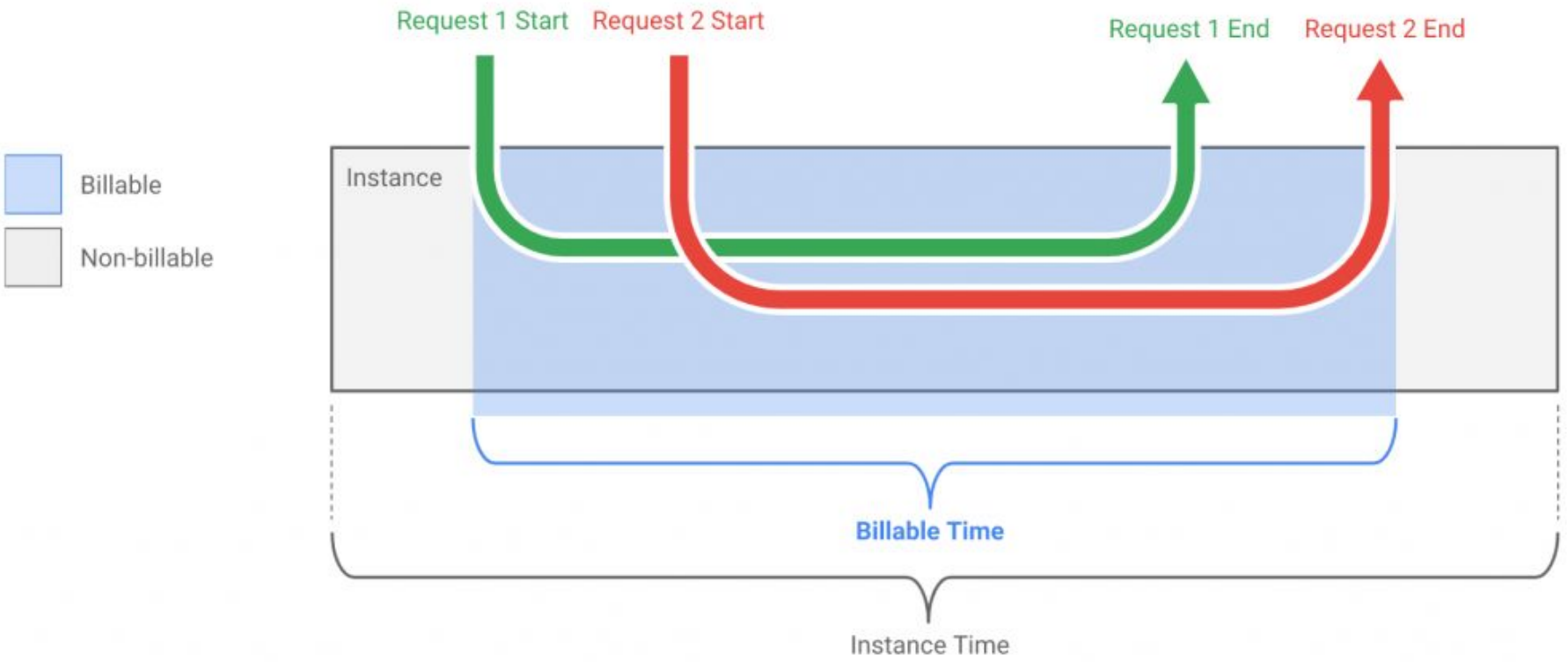


CPU / Memory / Requests

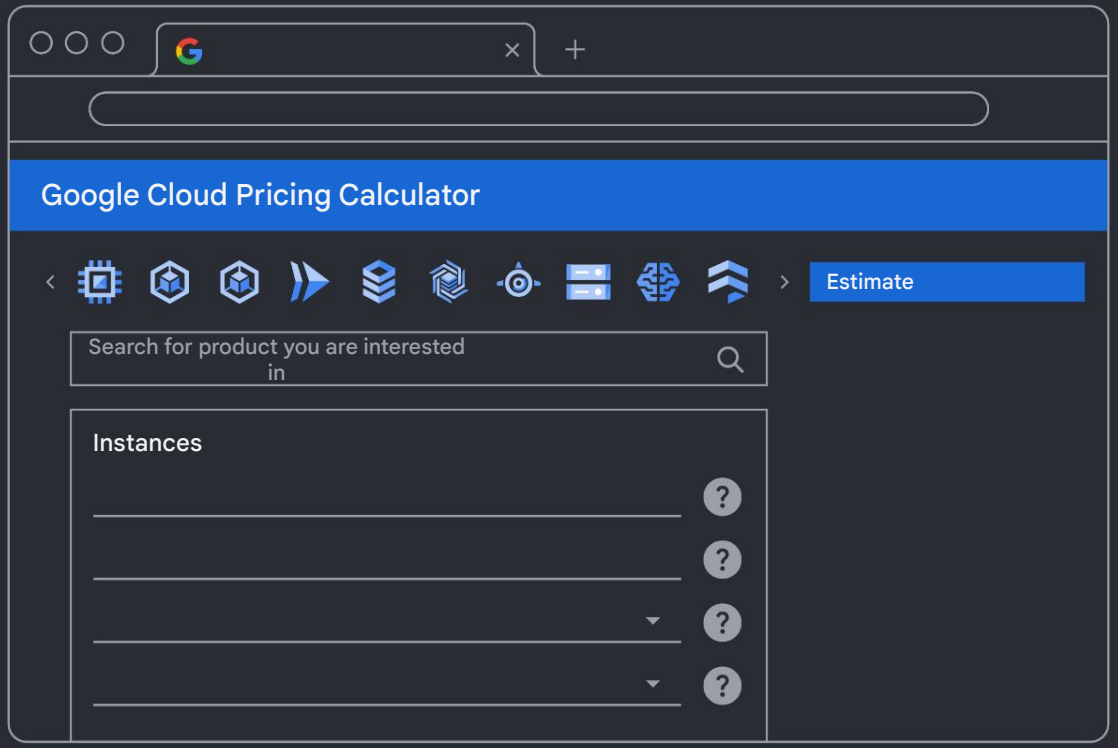



100ms

# Cloud Run – Billable container instance time



# Ce qui existe en standard…
# Pour mesurer et comprendre




cloud.google.com/products/calculator

# Billing tools help to budget and monitor usage

Budgets

Alerts

Reports

Quotas

# Ce qui existe en standard...
# Free tiers

# Free tier

### Compute Engine
Machines virtuelles évolutives et hautes performances

1 instance e2-micro par mois

### Cloud Storage
Des performances exceptionnelles, une fiabilité exemplaire et des tarifs compétitifs pour tous vos besoins de stockage

5 Go-mois de stockage standard

### BigQuery
Entrepôt de données d'analyse entièrement géré, à l'échelle du pétaoctet

1 To de requêtes par mois

### Google Kubernetes Engine
Orchestration de conteneurs en un clic via des clusters Kubernetes, gérés par Google

Un cluster Autopilot ou zonal par mois

### App Engine
Plate-forme pour créer des applications Web et des backends mobiles adaptables

28 instances-heure par jour

### Cloud Run
Un environnement entièrement géré permettant d'exécuter des conteneurs sans état

2 millions de requêtes par mois

### Cloud Build
Des compilations rapides, cohérentes et fiables sur Google Cloud

120 compilations-minute par jour

### Suite Operations (anciennement Stackdriver)
Surveillance, journalisation et diagnostics pour les applications sur Google Cloud

Attributions mensuelles pour la journalisation et la surveillance

### Firestore
Base de données de documents NoSQL qui simplifie le stockage, la synchronisation et l'interrogation des données d'applications

1 Go d'espace de stockage

### Pub/Sub
Un service de messagerie et de diffusion des données fiable et en temps réel, à l'échelle mondiale

10 Go de messages par mois

### Cloud Functions
Un environnement sans serveur pour créer et connecter des services cloud avec du code

2 millions d'appels par mois

### Vision AI
Détection de thèmes, reconnaissance optique de caractères, détection de visages et bien plus encore

1 000 unités par mois

https://cloud.google.com/free

# FinOps

https://www.finops.org

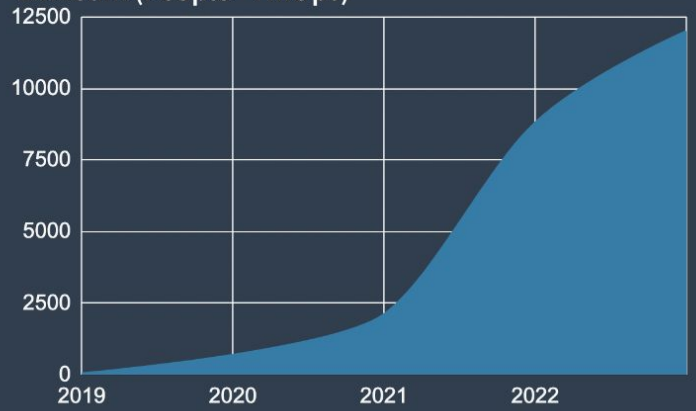
# Trends

# The FinOps industry has hit an inflection point this year...

*(we just launched jobs.finops.org too)*

## Social proof of those offering or seeking skills

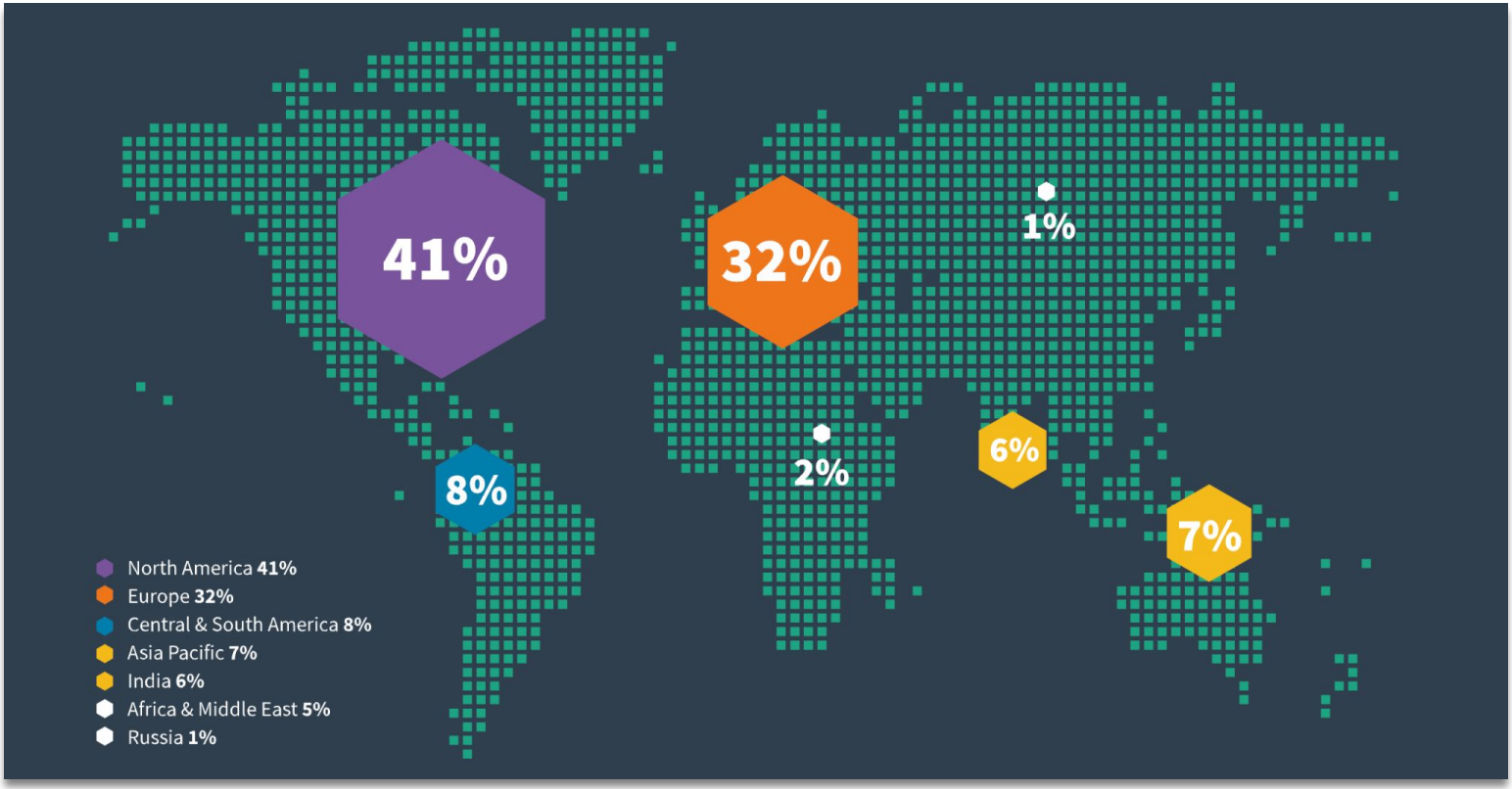
Linkedin (People - FinOps)




Posts in #jobs in the Slack 2021-2022


## 7000+ Community Members

# Identified FinOps practitioners



# Definition

L'approche **FinOps**, contraction des termes de finance et d'opérations, vise au monitoring et à l'optimisation des coûts en matière de cloud computing. Il est question, dans le même esprit que la stratégie DevOps, d'encourager la coopération et l'assimilation des moyens IT par les métiers. De ce fait, les pratiques FinOps ajoutent une dimension financière dans la gestion des coûts du Cloud.

# DevOps

❌ a standard
❌ a certification
❌ a goal

✅ a culture
✅ a philosophy
✅ a journey

# 📌 Key principles in FinOps

1. Teams need to collaborate
2. Everyone takes ownership for their cloud usage
3. A centralized team drives FinOps
4. Reports should be accessible and timely
5. Decisions are driven by business value of cloud
6. Take advantage of the variable cost model of the cloud.

# Cloud FinOps is an evolving cloud financial management discipline and cultural practice that:

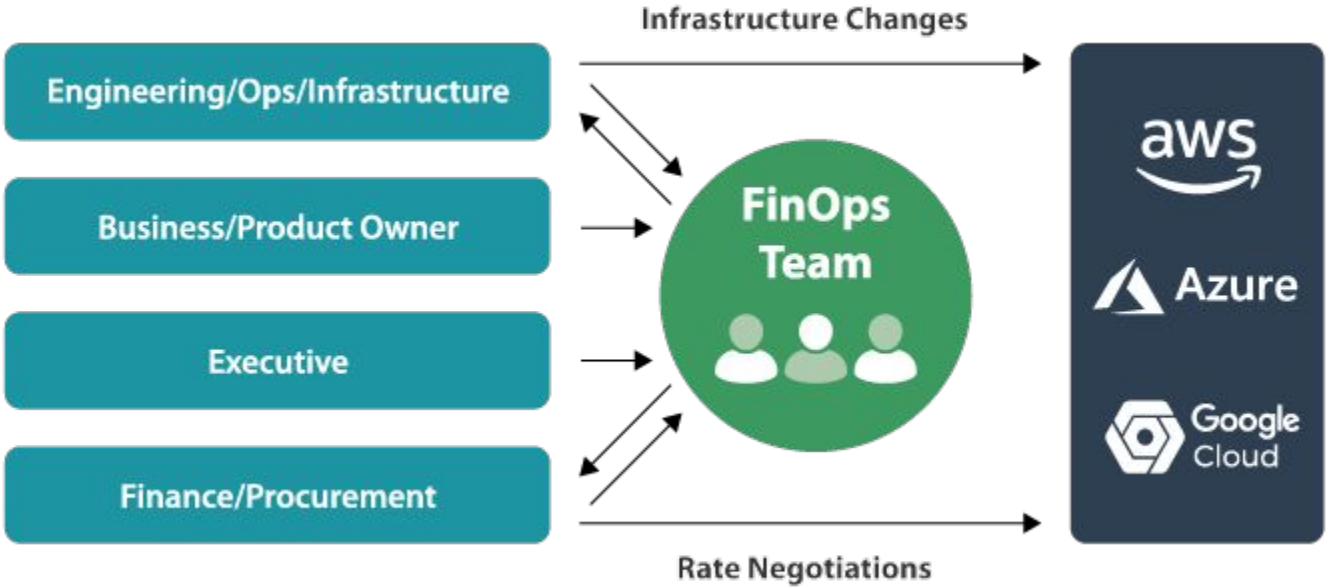
enables organizations to get **maximum business** value

by **helping** engineering, finance and business teams

to **collaborate** on data-driven spending **decisions**

## FinOps.org

# FinOps Foundation



# Ecosystem

FinOps
Foundation

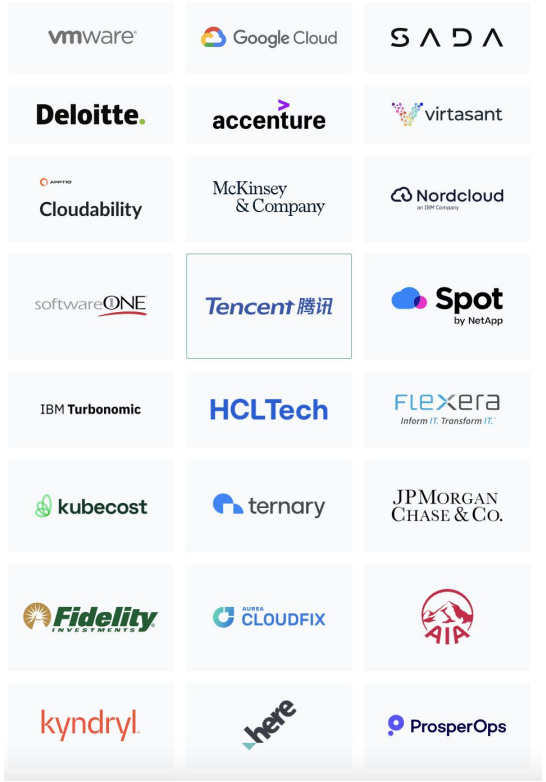
CLOUD COMPUTING > PERFORMANCE MANAGEMENT

# FinOps Foundation Brings New Focus to Cloud Costs

The FinOps Foundation is joining the Linux Foundation to accelerate its efforts to bring the costs of the cloud in line with financial accounting within companies.

Sean Michael Kerner | Jun 30, 2020

# 💰 25 Premier Members & 51 general members

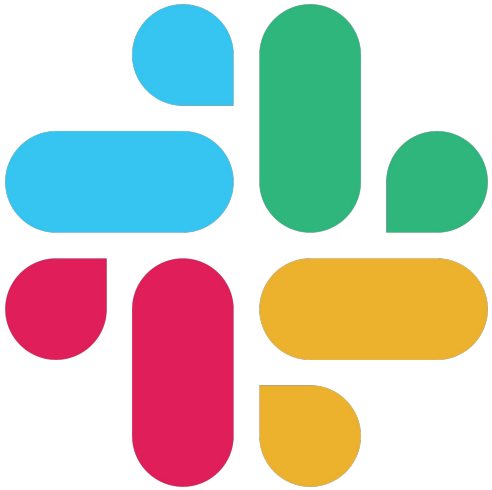


# 💪 Upskilling

# 👤 FinOps Personas

# 🎓 Training & Certification

# 💚 Community

# Events

## Upcoming FinOps Events

Suggest an event to be added


### Community FinOps Event – Feb 22, Washington, D.C

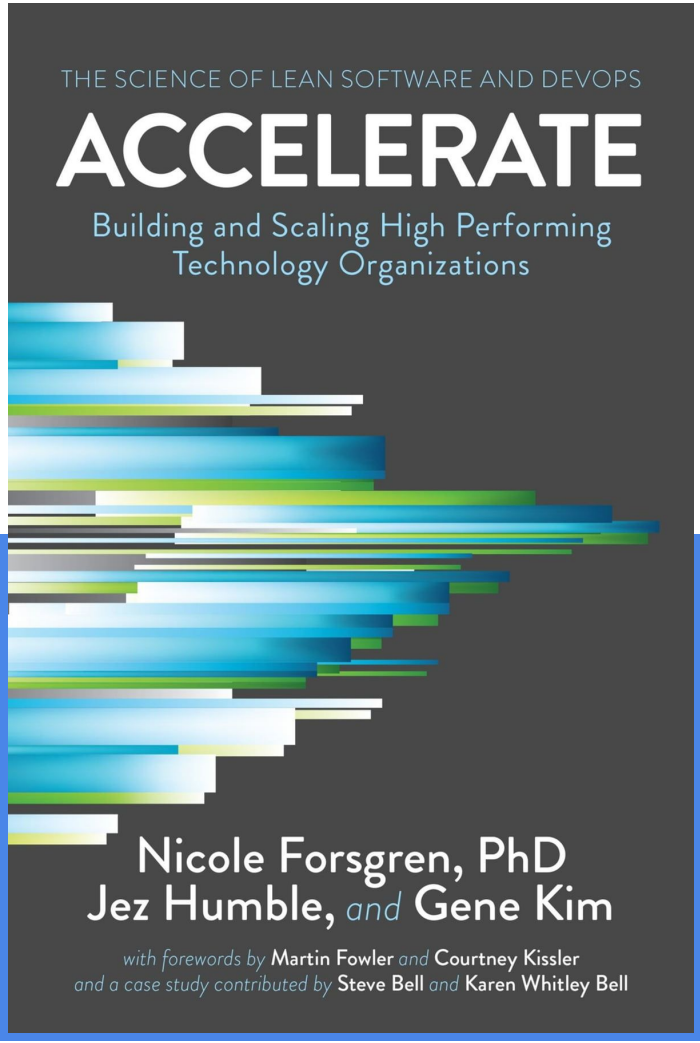

### FinOps Community Day Europe – March 21, Amsterdam


### FinOps X – June 28-29, San Diego, California

FinOps X is the official event

# Capabilities

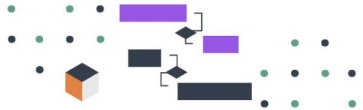
Accelerate / State of DevOps Report



# Domain & Capabilities

Understanding Cloud Usage and Cost

Performance Tracking & Benchmarking



Real-Time Decision Making

Cloud Rate Optimization

Cloud Usage Optimization

Organizational Alignment

# Tooling

kubecost

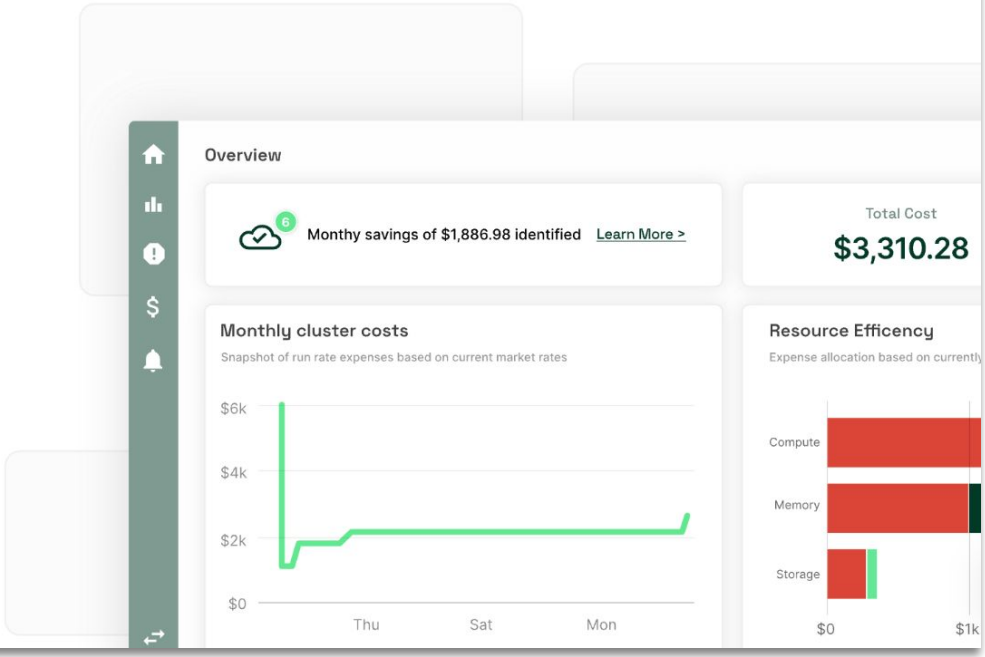
Features Pricing Resources About Contact Sales Get Started

Join us Feb 2nd to learn about monitoring your K8s costs in different environments

# Monitor & reduce Kubernetes spend

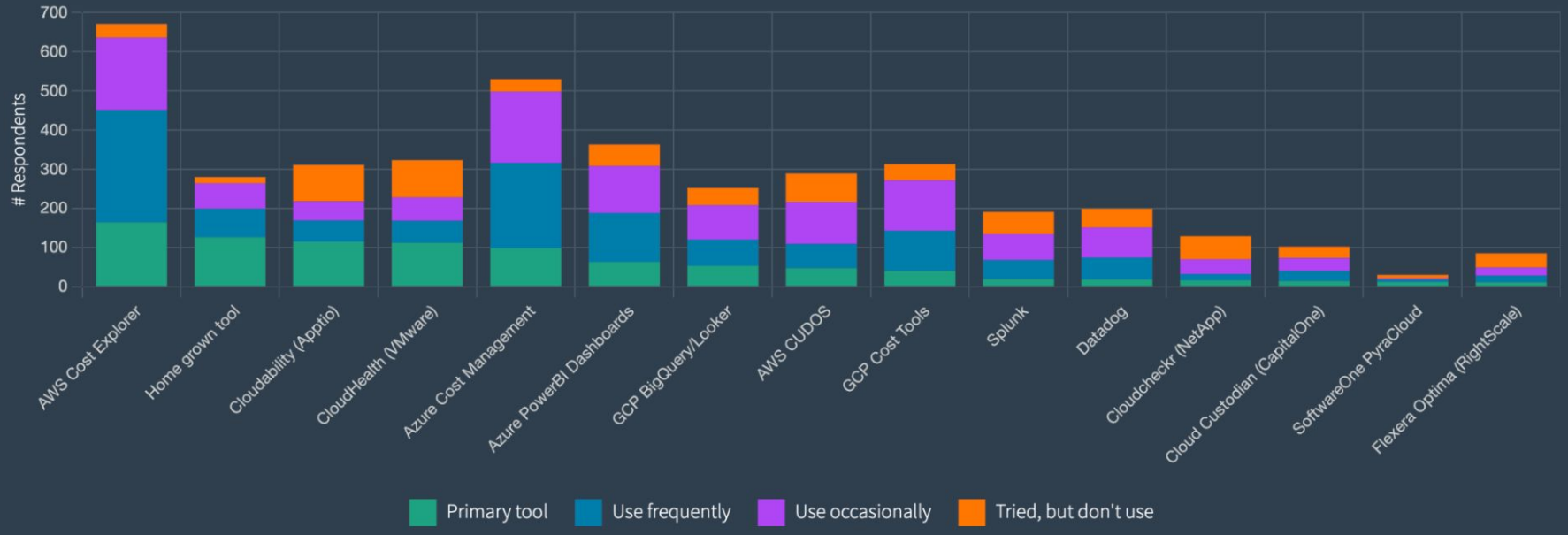
Kubecost provides real-time cost visibility and insights for teams using Kubernetes, helping you continuously reduce your cloud costs.

Watch a Demo Get Started



# Mix of Tooling Used to Support FinOps

Having the right FinOps tools available can smooth the most tumultuous parts of a journey (especially as cloud usage, cost, and complexity increase). Here are the top tools our respondents report they use daily to manage and forecast cloud costs. n=838, multiple choice



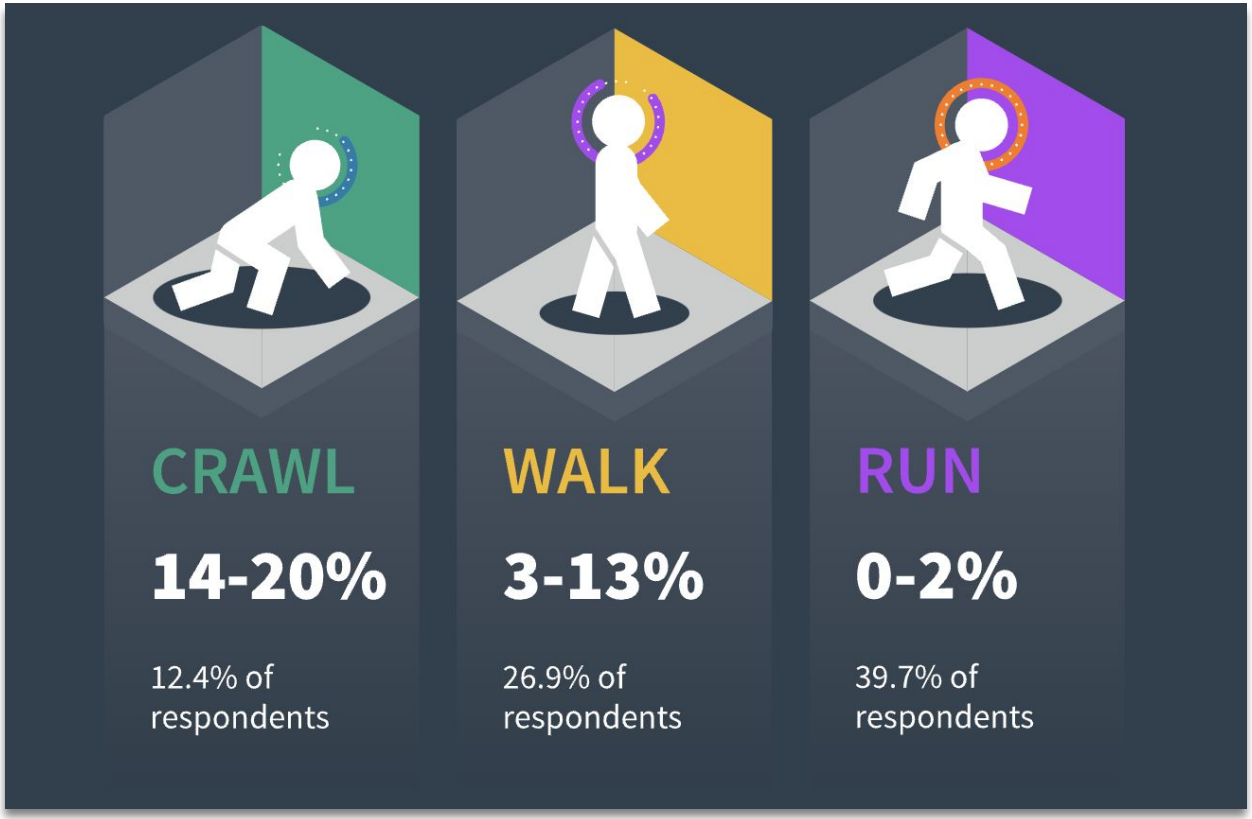
# 💡 Examples of measurement

| BUSINESS (Sales based on the way products are sold) | FINANCIAL (Service component value indicator) | TECHNICAL (Sum of FinOps tactics and engineering choices at infra value layer) |
|---|---|---|
| Cost per weekly active user | Cost per core hour | Cost variance reductions |
| Cost per million transactions | Cost per GB of RAM | Spend risk reductions |
| Cost per million in revenue | | |

# Adopt

# ⚡ What is the difference between forecasted spend and actual spend?

# 🎁 Percentage of commitment-based discounts that are estimated to go unused each month




## How do you promote FinOps adoption in the organization?

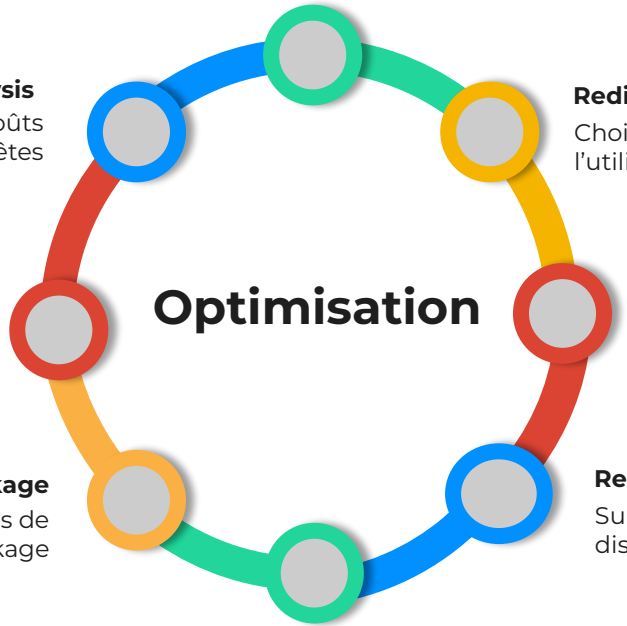
Crawl-stage respondents are in a learning state and rely on organic adoption, whereas run-stage respondents push executive directives, grass-roots initiatives, gamifications, and reward systems.

n=692

VIEW 'OTHER' ANSWERS →

# Comment l’appliquer au monde Google Cloud ?

# Des leviers disponibles




**Engagement de ressources**
Engagement en matière de consommation de services & campagne globale de réduction

**Redimensionnement**
Choisir les bons types d'instances de l'utilisation réelle du système

**Planification**
Allumer / éteindre les ressources en fonction de leur plage d'utilisation

**Ressources orphelines / inutilisées**
Supprimer les ressources inutilisées (instances, disk, snapshots, IPs, etc.)

**Ressources préemptibles**
Utiliser la surcapacité de l'infrastructure Google

**Solutions de stockage**
Utiliser les différentes classes de stockage, designer le stockage

**Architecture**
Revoir l'architecture pour utiliser les services les plus efficaces

**BigQuery Analysis**
Optimisation des coûts des requêtes

# Oui d’accord et ?

# Exemple 1 : Firebase

**Flavien Normand** for Zenika
Posted on Jan 18, 2021 • Updated on Jan 25, 2021 • Originally published at blog.zenika.com

❤️ 1

# Optimisations Firebase, moins cher et plus performant

#firebase #finops

# Exemple 2 : GKE

https://brennerm.github.io/posts/wasting-money-with-kubernetes.html

# FinOps Kubernetes : 10 indicateurs à surveiller pour rationaliser les coûts

**Benoit 🦊☸ COUETIL**
Ingénieur DevOps/Delivery

11 articles ✓ Following

# Exemple 3 : CUD

# 💡 Committed Use Discounts (CUDs)

## Resource-based CUDs

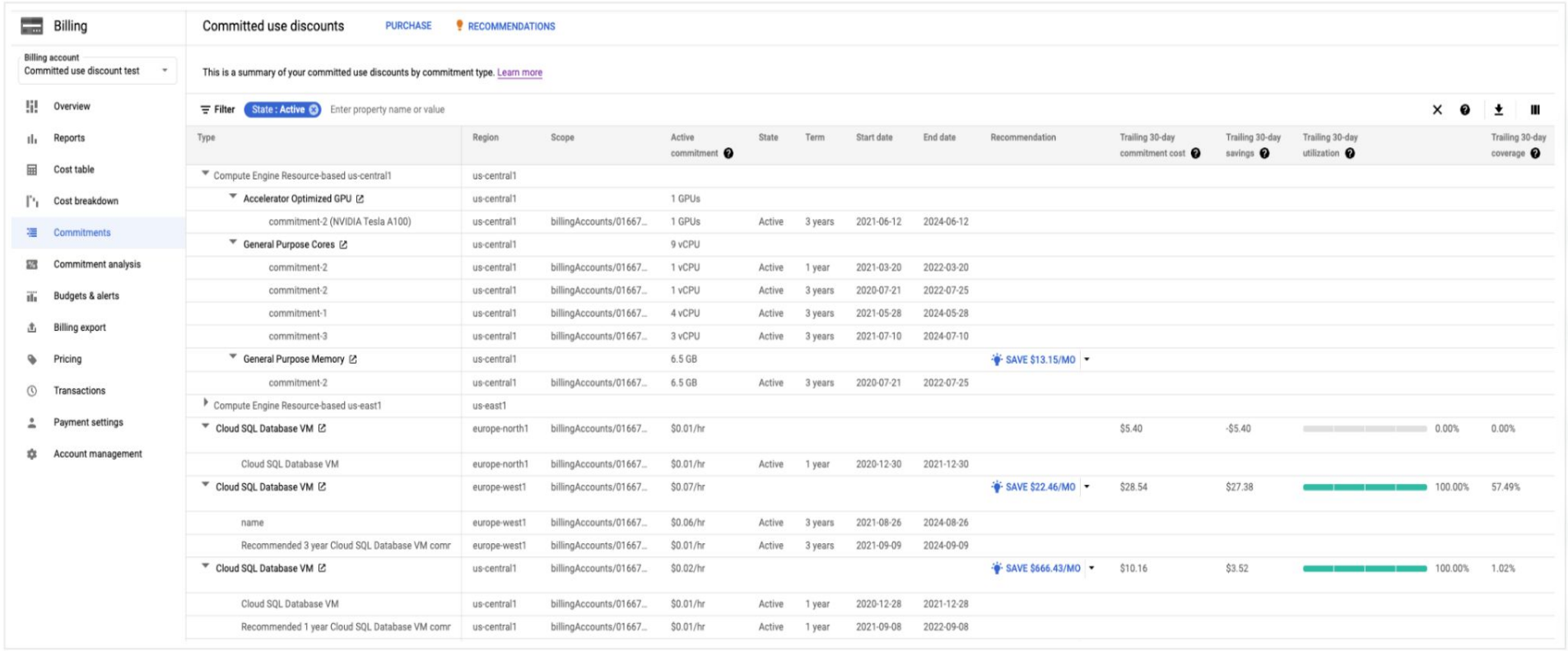
*Ex: 50 vCPU pour N2D dans us-central1*

Services pris en charge : Compute Engine, Dataproc, Google Kubernetes Engine (Standard)

## Spend-based CUDs

Exemple: $50/hour spend in Cloud SQL (PgSQL) dans us-central1

Services pris en charge : Cloud SQL, Cloud Spanner, Cloud Run, Google Cloud VMware Engine, Google Kubernetes Engine (Mode Autopilot


Billing
Billing account
Committed use discount test
Overview
Reports
Cost table
Cost breakdown
Commitments
Commitment analysis
Budgets & alerts
Billing export
Pricing
Transactions
Payment settings
Account management
Committed use discounts
PURCHASE
RECOMMENDATIONS
This is a summary of your committed use discounts by commitment type. Learn more
Filter
State : Active
Enter property name or value
Type
Region
Scope
Active commitment
State
Term
Start date
End date
Recommendation
Trailing 30-day commitment cost
Trailing 30-day savings
Trailing 30-day utilization
Trailing 30-day coverage
Compute Engine Resource-based us-central1 us-central1
Accelerator Optimized GPU us-central1 1 GPUs
commitment-2 (NVIDIA Tesla A100) us-central1 billingAccounts/01667... 1 GPUs Active 3 years 2021-06-12 2024-06-12
General Purpose Cores us-central1 9 vCPU
commitment-2 us-central1 billingAccounts/01667... 1 vCPU Active 1 year 2021-03-20 2022-03-20
commitment-2 us-central1 billingAccounts/01667... 1 vCPU Active 3 years 2020-07-21 2022-07-25
commitment-1 us-central1 billingAccounts/01667... 4 vCPU Active 3 years 2021-05-28 2024-05-28
commitment-3 us-central1 billingAccounts/01667... 3 vCPU Active 3 years 2021-07-10 2024-07-10
General Purpose Memory us-central1 6.5 GB SAVE $13.15/MO
commitment-2 us-central1 billingAccounts/01667... 6.5 GB Active 3 years 2020-07-21 2022-07-25
Compute Engine Resource-based us-east1 us-east1
Cloud SQL Database VM europe-north1 billingAccounts/01667... $0.01/hr $5.40 -$5.40 0.00% 0.00%
Cloud SQL Database VM europe-north1 billingAccounts/01667... $0.01/hr Active 1 year 2020-12-30 2021-12-30
Cloud SQL Database VM europe-west1 billingAccounts/01667... $0.07/hr SAVE $22.46/MO $28.54 $27.38 100.00% 57.49%
name europe-west1 billingAccounts/01667... $0.06/hr Active 3 years 2021-08-26 2024-08-26
Recommended 3 year Cloud SQL Database VM comr europe-west1 billingAccounts/01667... $0.01/hr Active 3 years 2021-09-09 2024-09-09
Cloud SQL Database VM us-central1 billingAccounts/01667... $0.02/hr SAVE $666.43/MO $10.16 $3.52 100.00% 1.02%
Cloud SQL Database VM us-central1 billingAccounts/01667... $0.01/hr Active 1 year 2020-12-28 2021-12-28
Recommended 1 year Cloud SQL Database VM comr us-central1 billingAccounts/01667... $0.01/hr Active 1 year 2021-09-08 2022-09-08

Cost Recommendations

EXPORT HISTORY SENI

## Recommended Compute Engine Commitments

You could be saving money with Compute Engine Commitments. Turn on discount sharing to see additional savings from all applicable cost recommendations. Learn more

**Cost savings**

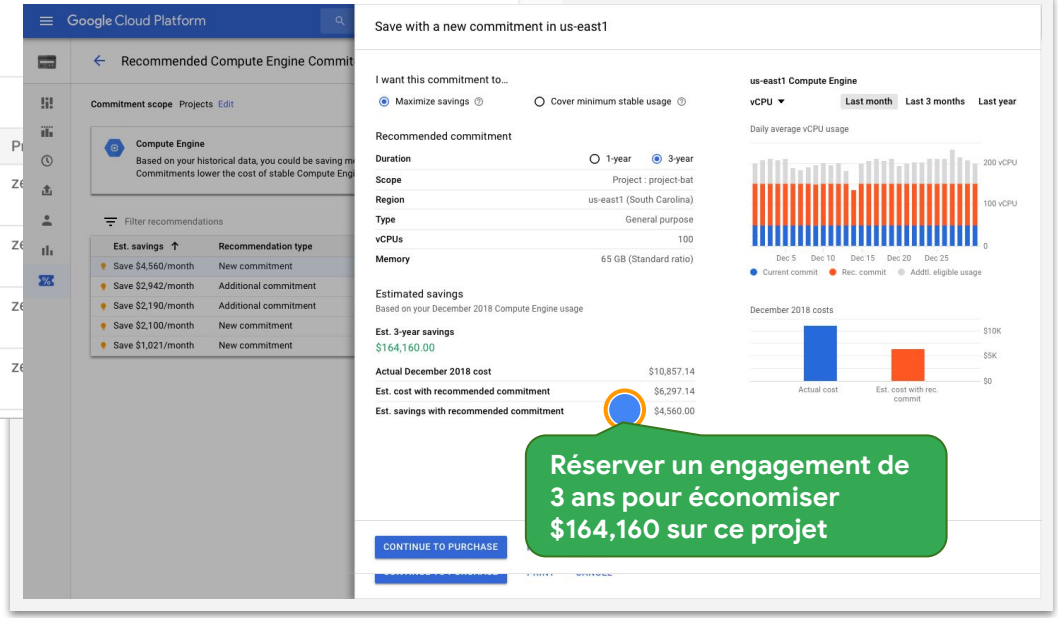
$79.30/month

## Recommendations

DISMISS

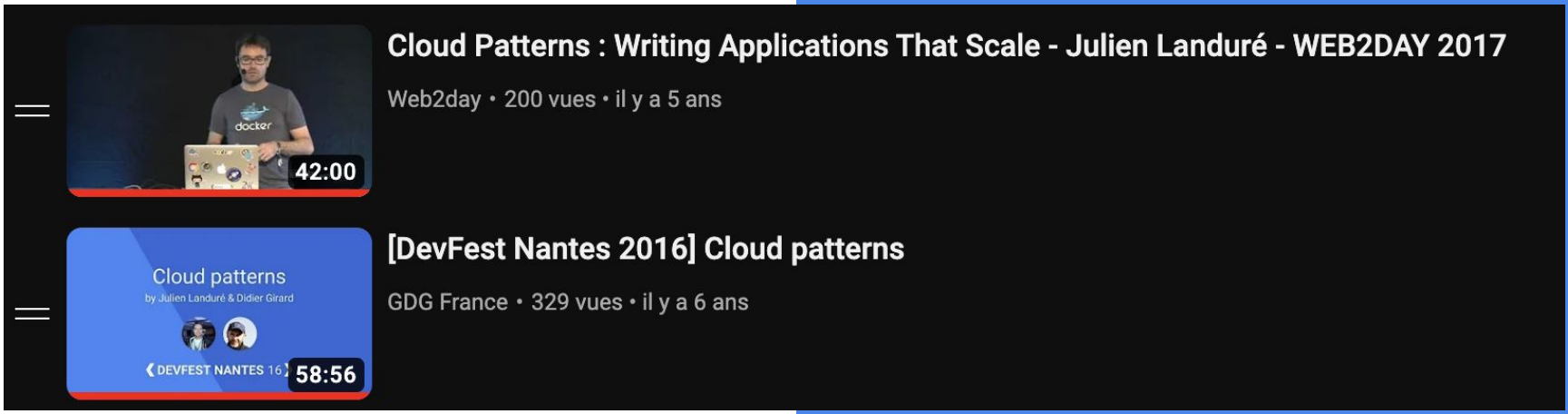
Filter Enter property name or value

| | Est. savings | Recommendation type | Region | Pr |
|---|---|---|---|---|
| ○ | Save $43.95/month | New commitment | europe-west3 | ze |
| ○ | Save $22.08/month | New commitment | europe-west3 | ze |
| ○ | Save $8.64/month | New commitment | us-west1 | ze |
| ○ | Save $4.63/month | New commitment | us-west1 | ze |



# Exemple 4 : Cloud Native Dev

# Développer en respectant les Cloud Patterns



# Conclusion

# Les points à retenir

- 👍 Un sujet tendance avec de vrais concepts
- 🔐 Démarche essentielle comme le SRE ou la Securité
- 💪 Une culture “FinOps” à diffuser
  - Autoscaling de 0 à N – par env (hors-prod / prod)
  - Respect de 12-Factors côté dev
  - Profiter au maximum du pricing model du Serverless
  - Choisir le bon outil pour faire la bonne chose
  - Observer, mesurer et maîtriser son architecture
  - Savoir quand se “commit” et sur quelle durée

# Next

# #GreenIT

# https://greensoftware.foundation

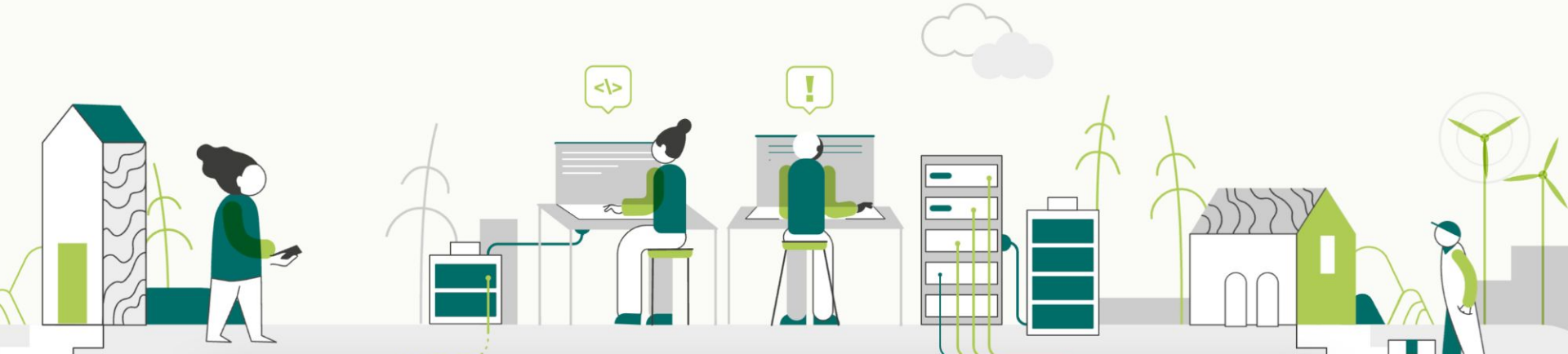
Green Software Foundation

About Working Groups Projects Resources Articles

JOIN US

We are building a trusted ecosystem of people, standards, tooling and best practices for

## GREEN SOFTWARE



# 🙌 Merci !

Introduction au FinOps avec Google Cloud ☁️


jlandure.dev
@jlandure