Venafi

Kubernetes Confessions: Tales of Overspending and Redemption

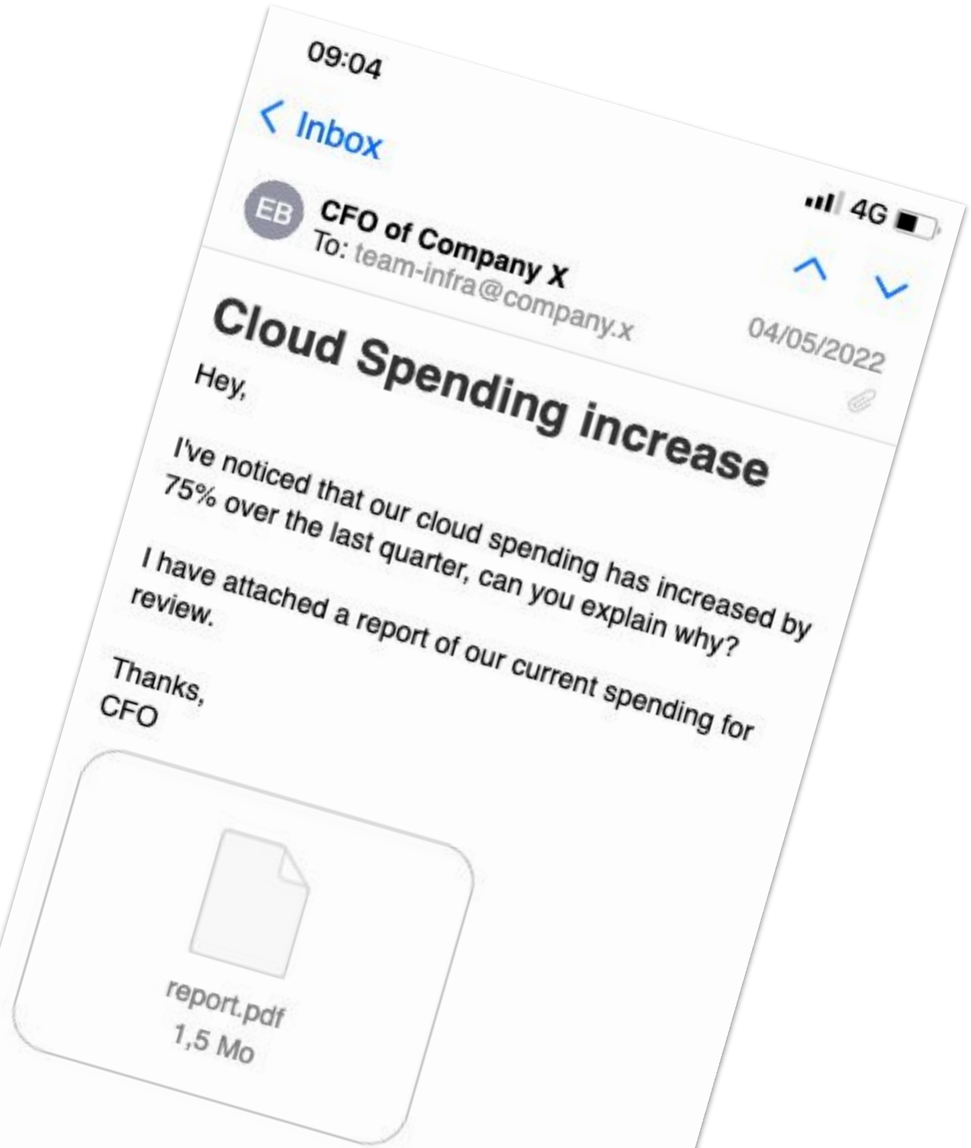
Becky Pauley
Solutions Engineer

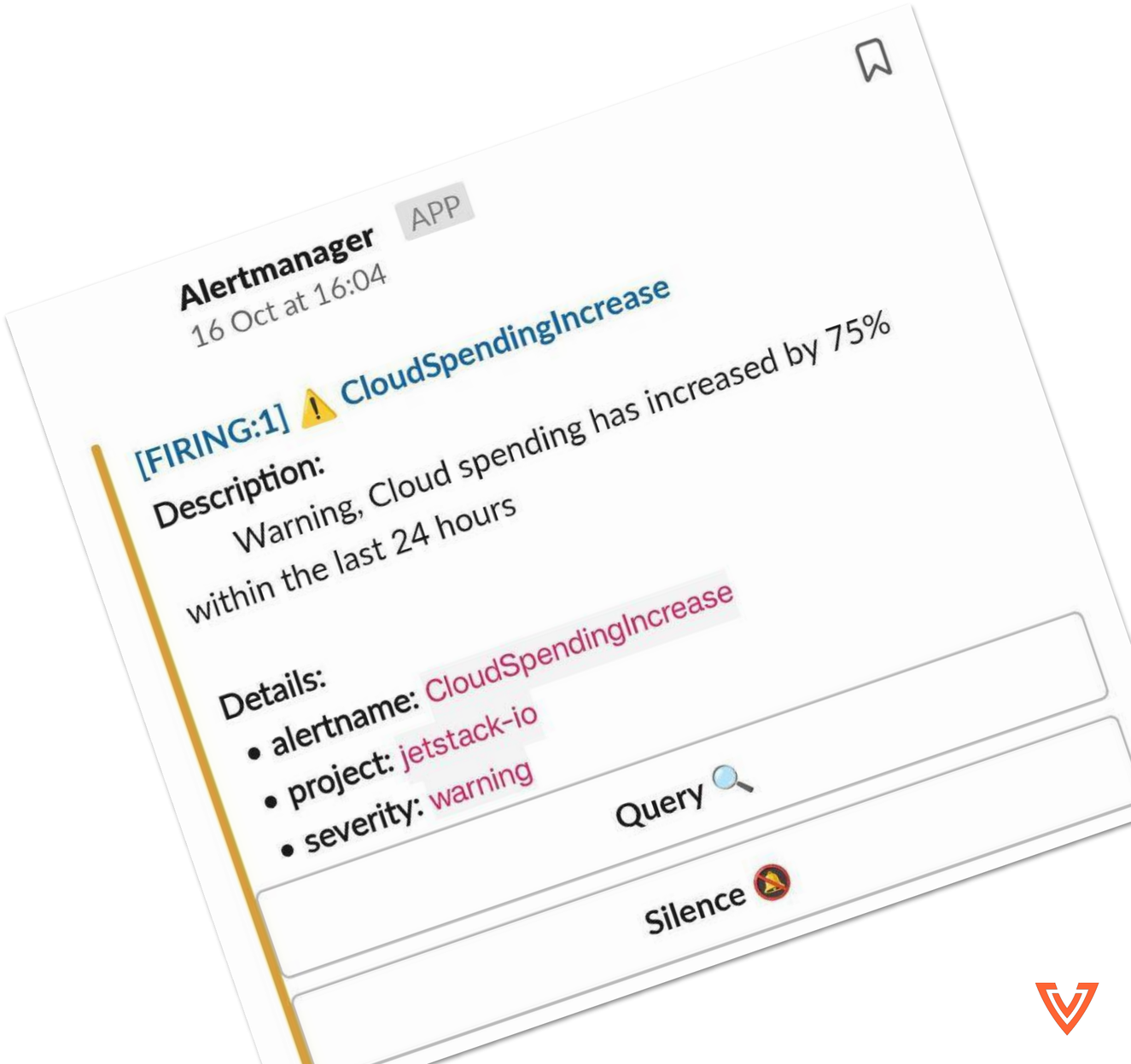
David Collom
Staff Solutions Engineer

A peek behind the scenes...

Image credit: HexandCube, Unsplash

Confession time





How to optimise your Kubernetes Cloud Spend

Optimization == Savings?

Value

"... the practice of getting the most business value for your cloud spend."

J.R. Storment, Mike Fuller: Cloud FinOps

Image credit: Tahlia Doyle, Unsplash

Reducing waste

Image credit: Jilbert Ebrahimi, Unsplash

Where does FinOps fit in?

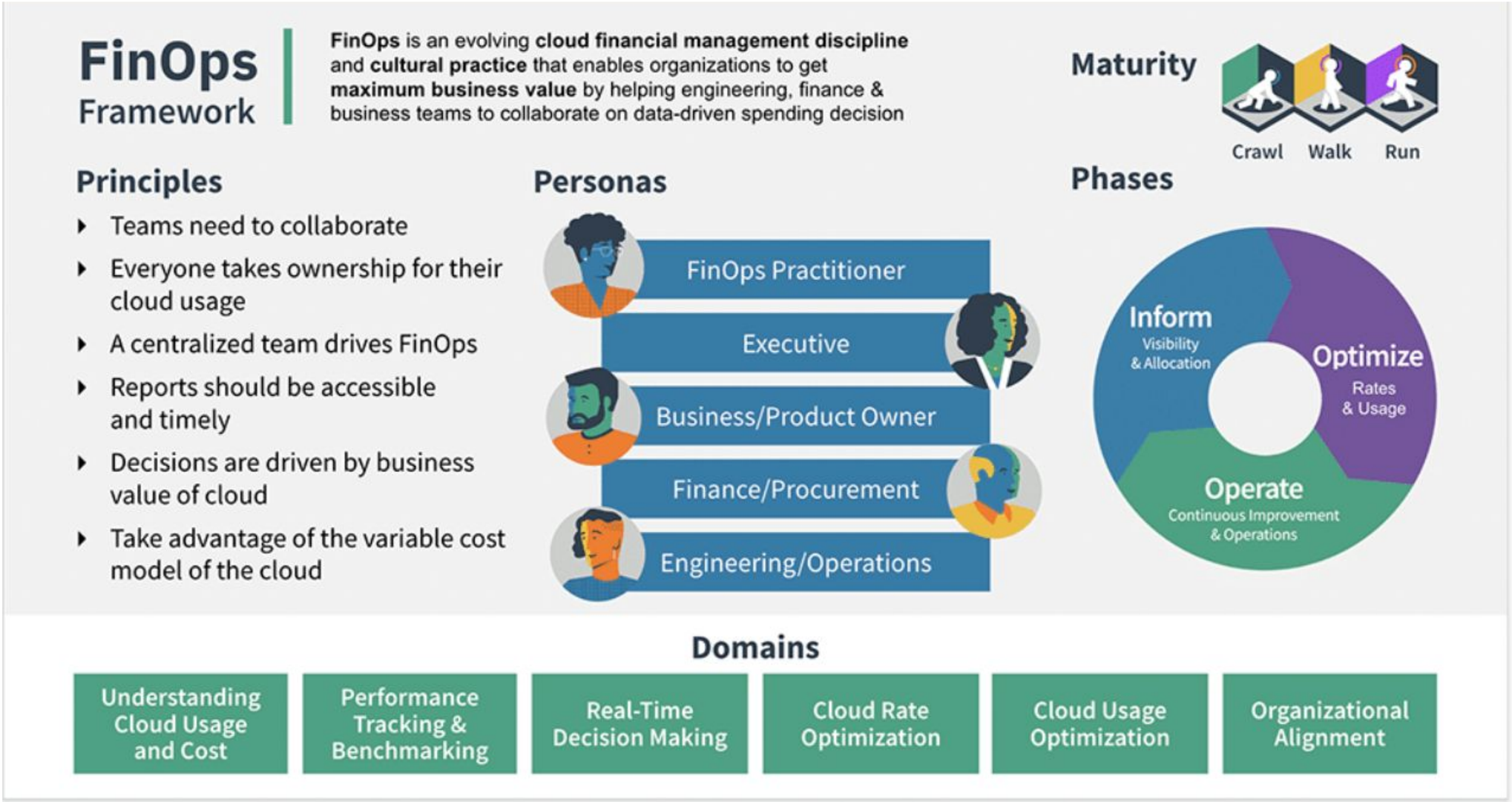


Image credit: What is FinOps? by FinOps Foundation

Where do we start?

~~Savings~~

“If you’re feeling intense pressure to save on cloud costs, you may be inclined to make drastic optimization decisions — don’t.

Making rash decisions without **visibility and context** into cloud spend can create unnecessary chaos. It may compromise productivity and contribute to a subpar customer experience.”

Google Cloud
Framing up FinOps: How to optimize your cloud costs on Google Cloud

Visibility

Image credit: Sarang Pande, Unsplash

Visibility

Cost allocation

Visibility

What do we want to show?

Visibility

David Collom’s bestie Barney

Visibility

David Collom’s bestie Barney

Visibility

Cost allocation

Efficiency/utilisation

Waste

Value

Something’s missing…

Confession:
The absent alert

You wake up on a Monday morning.
Before your morning Coffee, you see slack has lit up with notifications.
QA Reports that their test environment isn’t working.

You start your investigation and find nothing is working!
You continue to search for the Root Cause, and notice Production isn’t working either!

You then notice that this has been going on since Friday, and there are now 16k deployments and over 100 extra Nodes in the cluster....

How did nobody notice?

Alerting

Alerting

Alertmanager APP
16 Oct at 16:04

[FIRING:1] ⚠️ CloudSpendingIncrease

Description:

Warning, Cloud spending has increased by 75% within the last 24 hours

Details:

- **alertname:** CloudSpendingIncrease
- **project:** jetstack-io
- **severity:** warning

Query 🔍

Silence 🔕

Alertmanager APP 1:03 AM

[FIRING:1] ❓ HighCloudSpend

Alerts Firing:

Owner: @[redacted]

Account jetstack-[redacted] on gcp is predicted to exceed 310.72% of its 30 day limit

Predicted account spend of GBP310.72 is exceeding the limit of GBP100.00

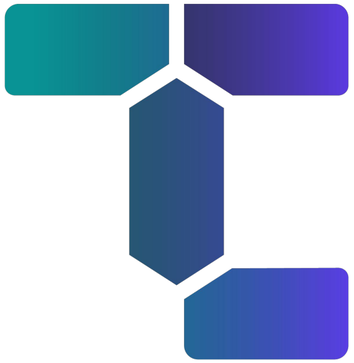
Details:

- **alertname:** HighCloudSpend
- **account:** [redacted]
- **cloud:** gcp
- **currency:** GBP
- **owner:** [redacted]
- **project:** [redacted]
- **type:** billing

Show less

Query 🔍 Silence 🔕

Tooling



Confession: Worker Node Sizing

“

During a routine review of the platform you happen to notice that the cluster has 100+ Nodes in Dev.

You review the number of applications within each environment, and discover you’re running one replica per node.

Right-sizing all the things

Image credit: Joshua Coleman, Unsplash

Requests

```
apiVersion: v1
kind: Pod
metadata:
  name: helm-controller-b78f6cf88-dhzlf
  namespace: flux-system
spec:
  containers:
  - name: manager
    image: ghcr.io/fluxcd/helm-controller:v0.28.1
    resources:
      requests:
        cpu: 250m
        memory: 100Mi
[..SNIP..]
```


Limits

```
apiVersion: v1
kind: Pod
metadata:
  name: helm-controller-b78f6cf88-dhzlf
  namespace: flux-system
spec:
  containers:
  - name: manager
    image: ghcr.io/fluxcd/helm-controller:v0.28.1
    resources:
      requests:
        cpu: 250m
        memory: 100Mi
      limits:
        cpu: 2500m
        memory: 1600Mi
[..SNIP..]
```

‘Optimizing’ too far

```
NAME          READY   STATUS      RESTARTS       AGE
memory-demo   0/1     OOMKilled   2 (22s ago)    26s
```

Confession:
The Diligent Time Sink

“

When I was newer to the industry, I received an email from my CTO asking me to reduce our monthly Kubernetes cloud spend.

Determined to find the most cost-effective solution, I invested time in a detailed analysis of different node pools and costings, diligently answering questions - including the possibility of migrating to a more managed alternative.

In the end, the cost (in salary) of the time spent on the activity probably cancelled out the savings I made for the first year. No-one seemed to mind because the cloud bill went down - but I’d approach it very differently now.

Machines and clusters

Image credit: Frank McKenna, Unsplash

Image credit: Matt Duncan, Unsplash

Scaling down: Out of hours

- Scale down our Nginx Deployment
- Every Night at 5pm(18:00)
- Monday - Friday

```yaml
apiVersion: batch/v1
kind: CronJob
metadata:
  name: kubectl-scaler
spec:
  schedule: "0 18 * * 1-5"
  jobTemplate:
    spec:
      template:
        spec:
          serviceAccountName: kubectl-scaler
          containers:
          - name: kubectl-scaler
            image: bitnami/kubectl:1.25.11
            command: ["/bin/sh", "-c"]
            args:
            - kubectl scale deployment nginx --replicas=0
          restartPolicy: OnFailure
```

Sample Code used from: https://github.com/nodematiclabs/cron-k8s-deployment-scaling/blob/main/resources.yaml

Tooling

The real picture:

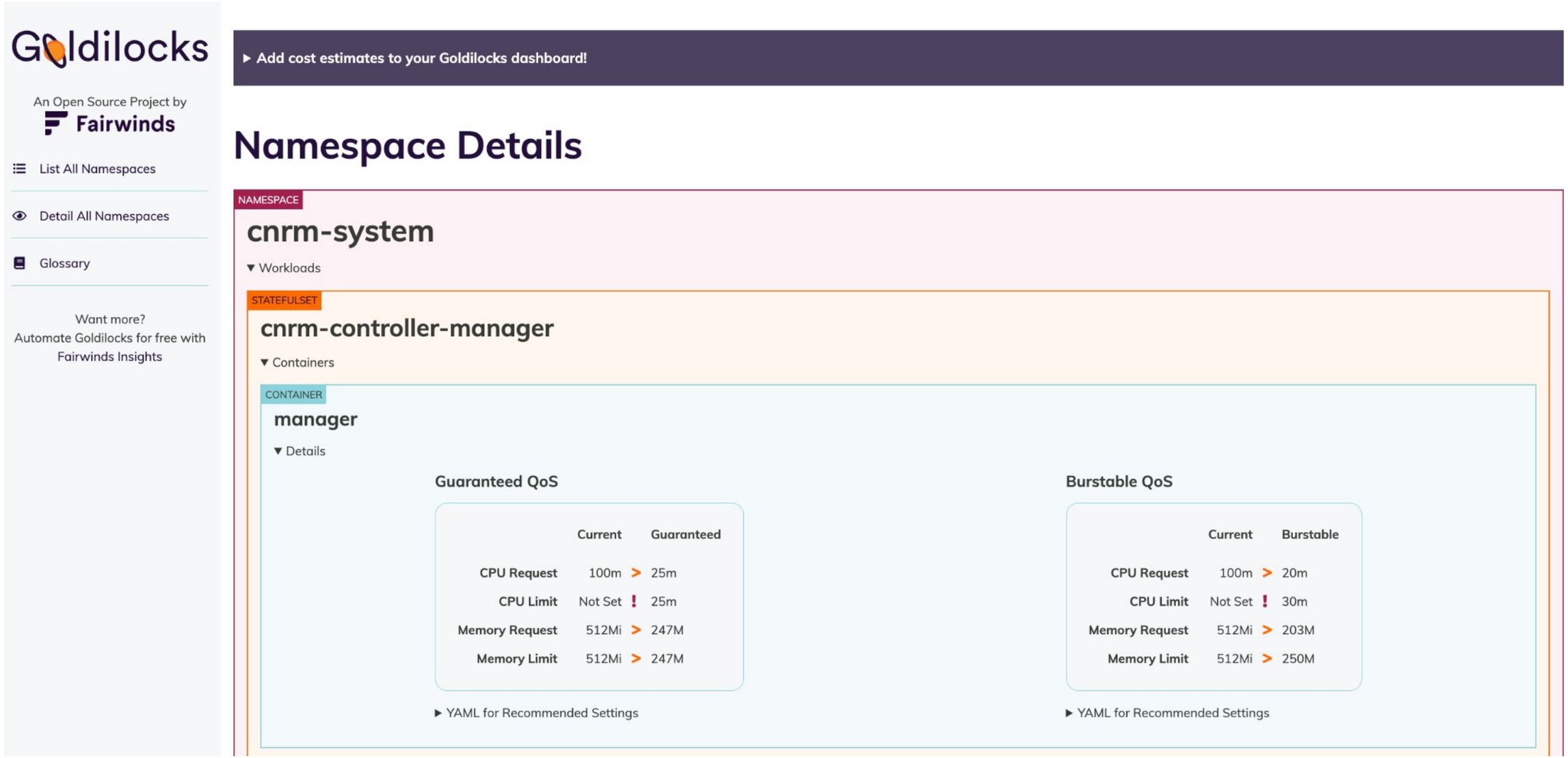
- Vertical Pod Autoscaler (VPA)
- Fairwinds Goldilocks
- Horizontal Pod Autoscaler (HPA)
- Keda https://keda.sh/
- Alpha feature from 1.27 re: resizing
- Application profiling

Tooling

```
~$ k vpa-recommendation -A
NAME                    MODE   TARGET
cert-manager            Auto   cert-manager
helm-controller         Auto   helm-controller
kustomize-controller    Auto   kustomize-controller
source-controller       Auto   source-controller
ingress-nginx           Auto   ingress-nginx-controller
ttl-controller          Auto   ttl-controller
~$
```

```
~$ k  vpa-recommendation -A
NAME                                          MODE   TARGET                                % CPU DIFF   % MEMORY DIFF
cert-manager                                  Auto   cert-manager                          +0.00       +0.00
goldilocks-cert-manager                       Off    cert-manager                          +12400.00   +170.27
goldilocks-cert-manager-cainjector            Off    cert-manager-cainjector               -           -
goldilocks-cert-manager-webhook               Off    cert-manager-webhook                  -           -
goldilocks-cnrm-controller-manager            Off    cnrm-controller-manager               +270.37     +109.84
goldilocks-cnrm-deletiondefender              Off    cnrm-deletiondefender                 +900.00     +453.51
goldilocks-cnrm-resource-stats-recorder       Off    cnrm-resource-stats-recorder          +233.33     +77.78
goldilocks-cnrm-webhook-manager               Off    cnrm-webhook-manager                  +12400.00   +58.02
goldilocks-configconnector-operator           Off    configconnector-operator              +4900.00    +306.35
goldilocks-external-dns                       Off    external-dns                          -           -
goldilocks-external-secrets                   Off    external-secrets                      -           -
goldilocks-external-secrets-cert-controller   Off    external-secrets-cert-controller      -           -
goldilocks-external-secrets-webhook           Off    external-secrets-webhook              -           -
goldilocks-helm-controller                    Off    helm-controller                       +3025.00    +132.56
goldilocks-kustomize-controller               Off    kustomize-controller                  +1150.00    +1.01
goldilocks-source-controller                  Off    source-controller                     +3471.43    +0.65
helm-controller                               Auto   helm-controller                       +0.00       +0.00
kustomize-controller                          Auto   kustomize-controller                  +0.00       +1.01
source-controller                             Auto   source-controller                     +0.00       +0.65
goldilocks-alertmanager                       Off    alertmanager                          +300.00     +57.85
goldilocks-collector                          Off    collector                             +242.86     +234.95
goldilocks-gmp-operator                       Off    gmp-operator                          +166.67     +1.73
goldilocks-rule-evaluator                     Off    rule-evaluator                        +128.57     +27.16
goldilocks-goldilocks-controller              Off    goldilocks-controller                 +25.00      +1405.88
goldilocks-goldilocks-dashboard               Off    goldilocks-dashboard                  +25.00      +1063.64
goldilocks-ingress-nginx-controller           Off    ingress-nginx-controller              +2400.00    +345.22
ingress-nginx                                 Auto   ingress-nginx-controller              -           +0.00
goldilocks-jwtredirector                      Off    jwtredirector                         -           -
goldilocks-calico-node                        Off    calico-node                           +400.00     -
goldilocks-calico-node-vertical-autoscaler    Off    calico-node-vertical-autoscaler       -           -
goldilocks-calico-typha                       Off    calico-typha                          +6566.67    -
goldilocks-calico-typha-horizontal-autoscaler Off    calico-typha-horizontal-autoscaler    +400.00     -
goldilocks-calico-typha-vertical-autoscaler   Off    calico-typha-vertical-autoscaler      -           -
goldilocks-event-exporter-gke                 Off    event-exporter-gke                    -           -
goldilocks-fluentbit-gke                      Off    fluentbit-gke                         +426.32     +300.00
goldilocks-fluentbit-gke-256pd                Off    fluentbit-gke-256pd                   -           -
goldilocks-fluentbit-gke-max                  Off    fluentbit-gke-max                     -           -
goldilocks-gke-metadata-server                Off    gke-metadata-server                   +1328.57    +203.03
goldilocks-gke-metrics-agent                  Off    gke-metrics-agent                     +100.00     +122.22
```

Tooling



Confession: The Wild West Dev project

“

I spun up a GKE cluster to test something that minikube / kind (probably could, but) didn't really fit the bill for. I did my testing over a few days / weeks (obviously I didn't scale anything down when it wasn't in use), and then... pff!

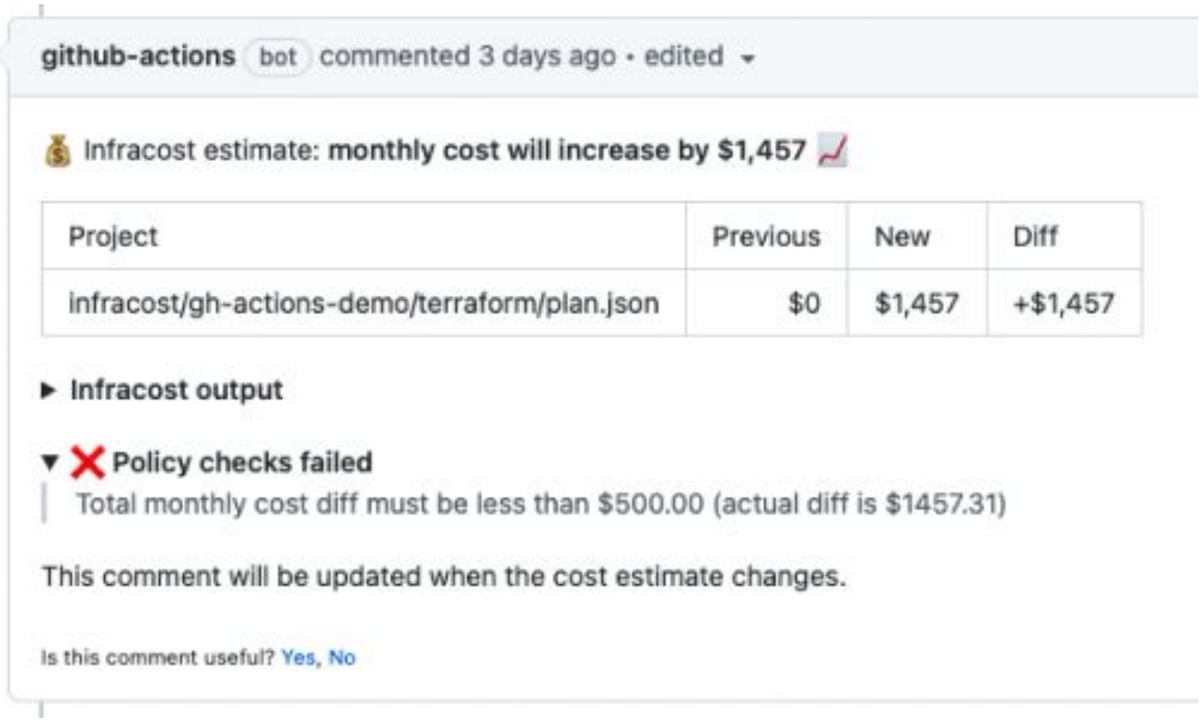
My brain moved on from that matter entirely. *4 MONTHS*, and  £573 ($700+) later, I finally discovered that the cluster (and a bunch of detached persistent disks) was still provisioned... lying there idle.

The best part? My org still to this day hasn't asked a single question about it. #workperks

Guard rails

Image credit: Josh Applegate, Unsplash

Guard rails



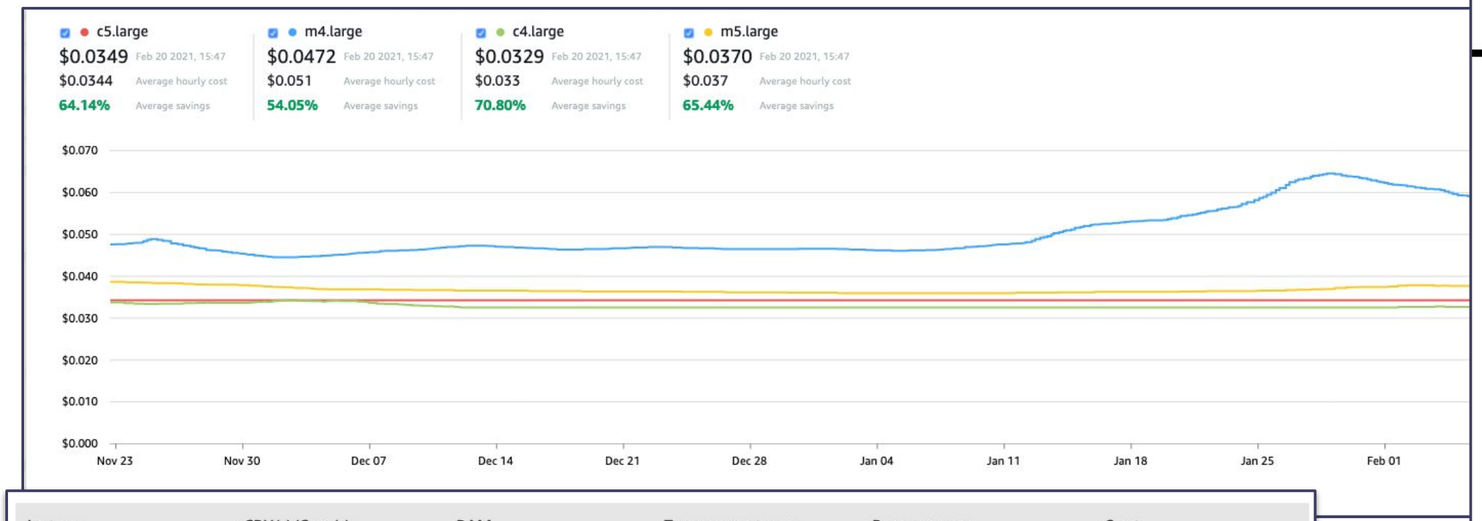

```
package infracost
deny[out] {
  maxDiff = 5000.0

  msg := sprintf(
    "Total monthly cost diff must be less than $%.2f
(actual diff is $%.2f)",
    [maxDiff, to_number(input.diffTotalMonthlyCost)],
  )
 `failed` property.
  out := {
    "msg": msg,
    "failed": to_number(input.diffTotalMonthlyCost) >=
maxDiff
  }
}
```

↕

. . .


Everyone loves a discount!
SALE

Spot Instances



Instance	vCPU(s)/Core(s)	RAM	Temporary storage	Pay as you go	Spot ▲
D4pls v5	4	8 GiB	N/A	$124.10/month	$12.41/month 90% savings
D4ps v5	4	16 GiB	N/A	$130.67/month	$13.07/month 90% savings
D4plds v5	4	8 GiB	150 GiB	$140.16/month	$14.02/month 90% savings
A4 v2	4	8 GiB	40 GiB	$139.43/month	$14.35/month 90% savings
D4as v5	4	16 GiB	N/A	$147.46/month	$14.75/month 90% savings
D4pds v5	4	16 GiB	150 GiB	$154.76/month	$15.48/month 90% savings
F4s v2	4	8 GiB	32 GiB	$154.76/month	$15.48/month 90% savings
D4a v4	4	16 GiB	100 GiB	$163.52/month	$16.35/month 90% savings

VM	Region	Total Savings Over On-Demand
C2	us-west1	90.90%
	us-east1	90.90%
	us-cental1	90.90%
N1	asia-east1	78.80%
	asia-northeast1	78.80%
	asia-northeast3	84.00%
	asia-south1	83.30%
	asia-southeast1	83.30%
	asia-southeast2	86.50%
	australia-southeast1	84.50%
	europe-west1	84.00%
	europe-west3	84.00%
	europe-west4	83.00%
	northamerica-northeast2	83.00%
	us-west3	83.00%
N2	asia-east1	81.00%
	asia-south1	83.90%
	asia-southeast1	81.00%
	europe-west1	78.90%
	europe-west3	81.00%
	us-west4	81.00%
	europe-west8	80.80%
	europe-west6	80.80%

Images for illustration purposes - Please see your cloud provider for more accurate details at the time.

Commitments

Spend-based CUDs	Resource-based CUDs
Spend-based commitments for Compute Engine are purchased and measured in terms of the dollars per hour of equivalent on-demand spend.	**Resource-based** commitments for Compute Engine are purchased and measured in terms of the underlying vCPU, memory, GPU, and local SSD resources.
Spend-based commitments are purchased from your Cloud Billing account; they apply to eligible usage in any projects paid for by that Cloud Billing account. ★ **Note**: After you purchase a spend-based commitment, you cannot change the Cloud Billing account associated with the commitment.	**Resource-based** commitments are purchased in the context of an individual project, rather than that of a Cloud Billing account. You can enable discount sharing so that the Compute Engine committed use discounts are shared across all projects that are paid for by the same Cloud Billing account. You can change the Cloud Billing account that pays for the project where you purchased the resource-based commitments. Learn about changing the Cloud Billing account for projects.

Other honourable mentions

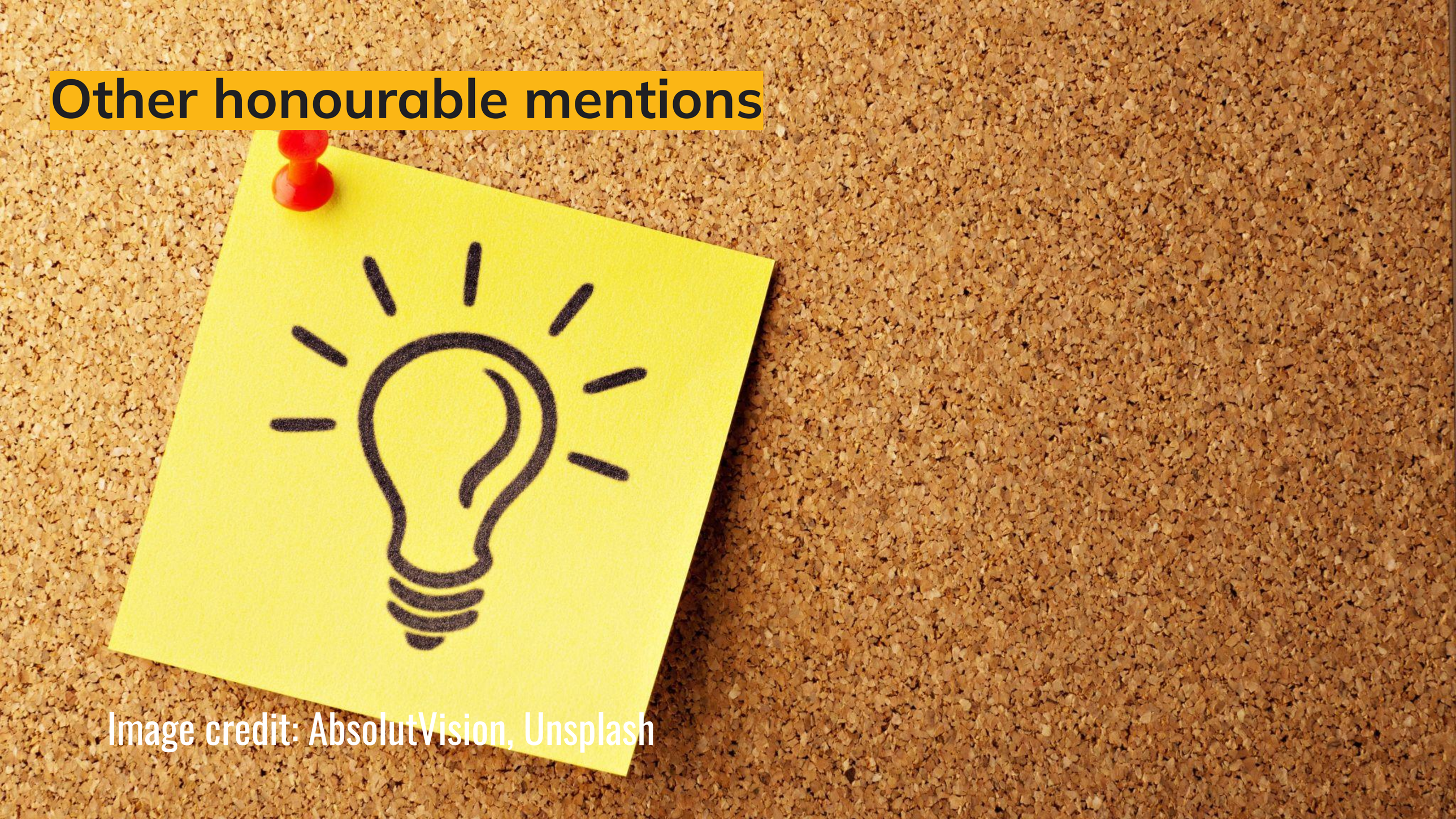
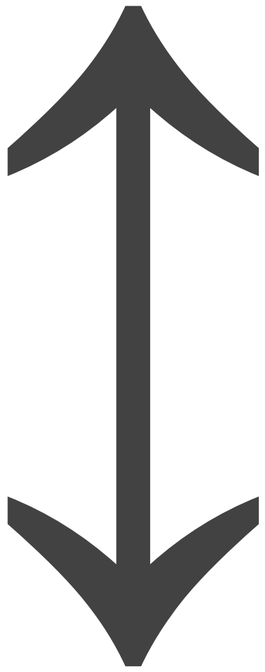
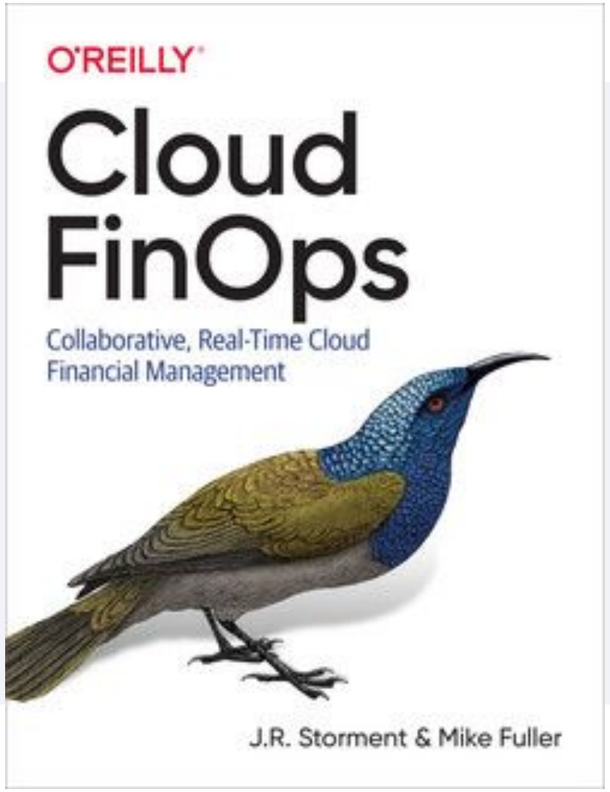


Image credit: AbsolutVision, Unsplash



Where do I look next?

Google Cloud: State of Kubernetes Cost Optimisation



What is Fleet Ops? - Venafi Jetstack Consult

Cloud native ecosystem knowledge and expertise

Venafi Jetstack Consult provides strategic consulting and advisory services for companies that are embracing cloud native.

CLOUD NATIVE COMPUTING FOUNDATION

https://venafi.com/jetstack-consult/

Venafi

Thank you!

Please submit Feedback at the following QR Code

Becky Pauly
Solutions Engineer
becky.pauly@venafi.com

David Collom
Staff Solutions Engineer
david.collom@venafi.com

Questions?