KCD
GROWING CLOUD NATIVE TOGETHER
El Salvador
SAN SALVADOR 2025

Kubernetes al límite: ¿cómo request & limits pueden proteger o degradar tu aplicación?

Apoyado Por

CLOUD NATIVE COMPUTING FOUNDATION

K8S restaurant

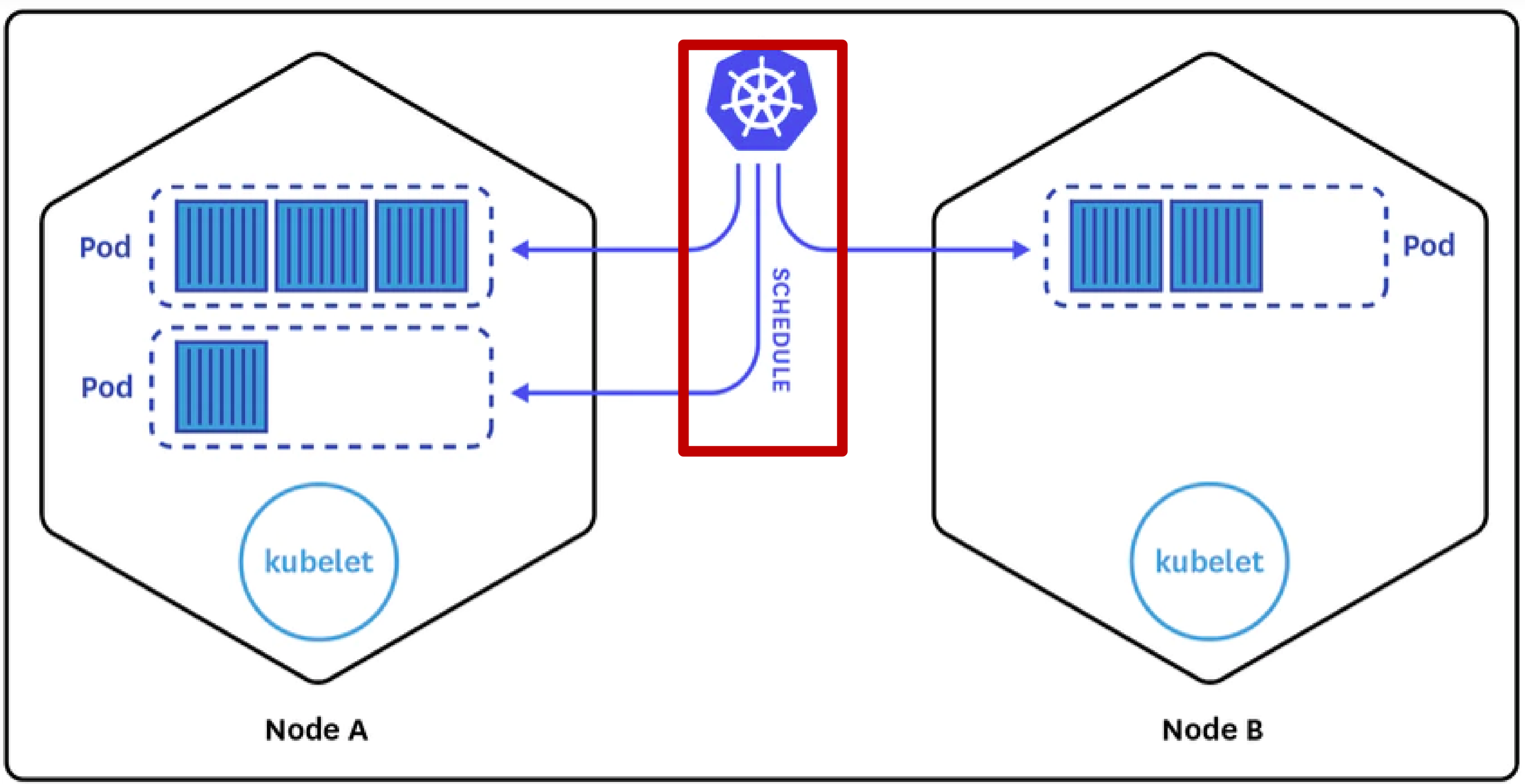
KCD
GROWING CLOUD NATIVE TOGETHER
El Salvador
SAN SALVADOR 2025

Apoyado Por CLOUD NATIVE COMPUTING FOUNDATION

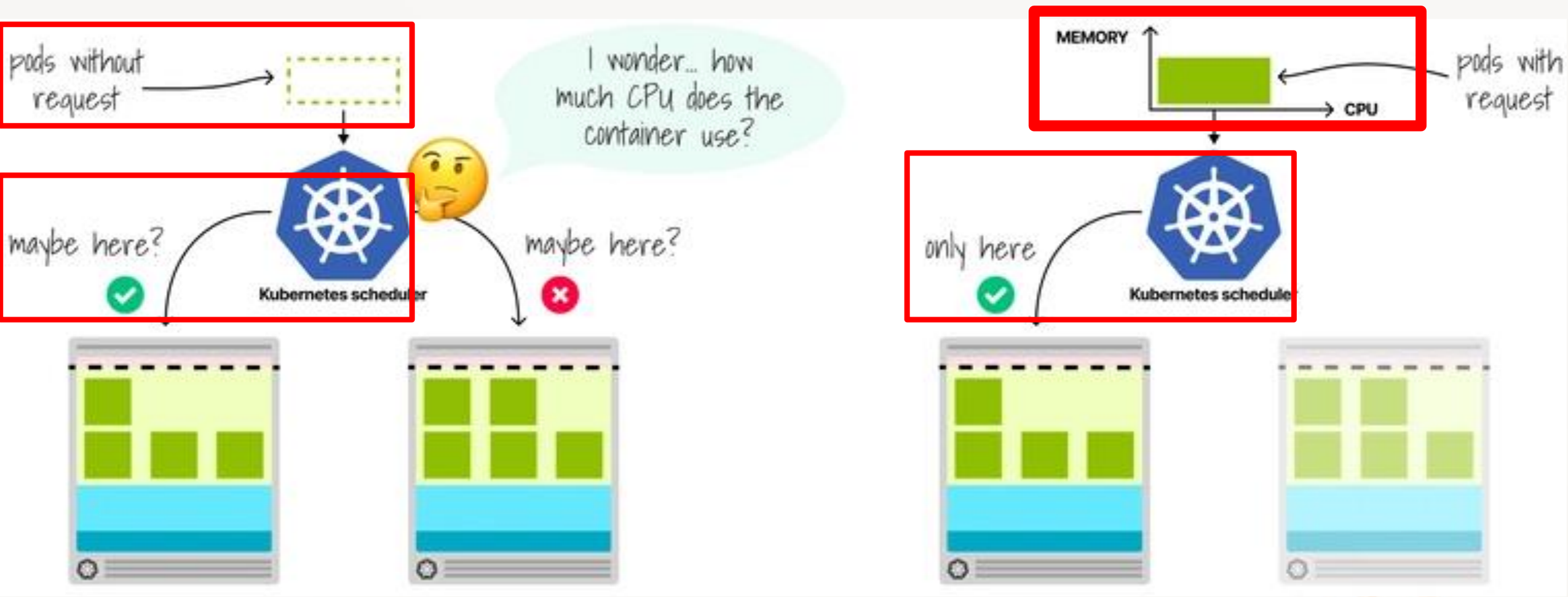
Adalberto García

BYTE

www.linkedin.com/in/adalberto-garcia/



Fuente: https://blog.nashtechglobal.com/schedule-pod-over-nodes/



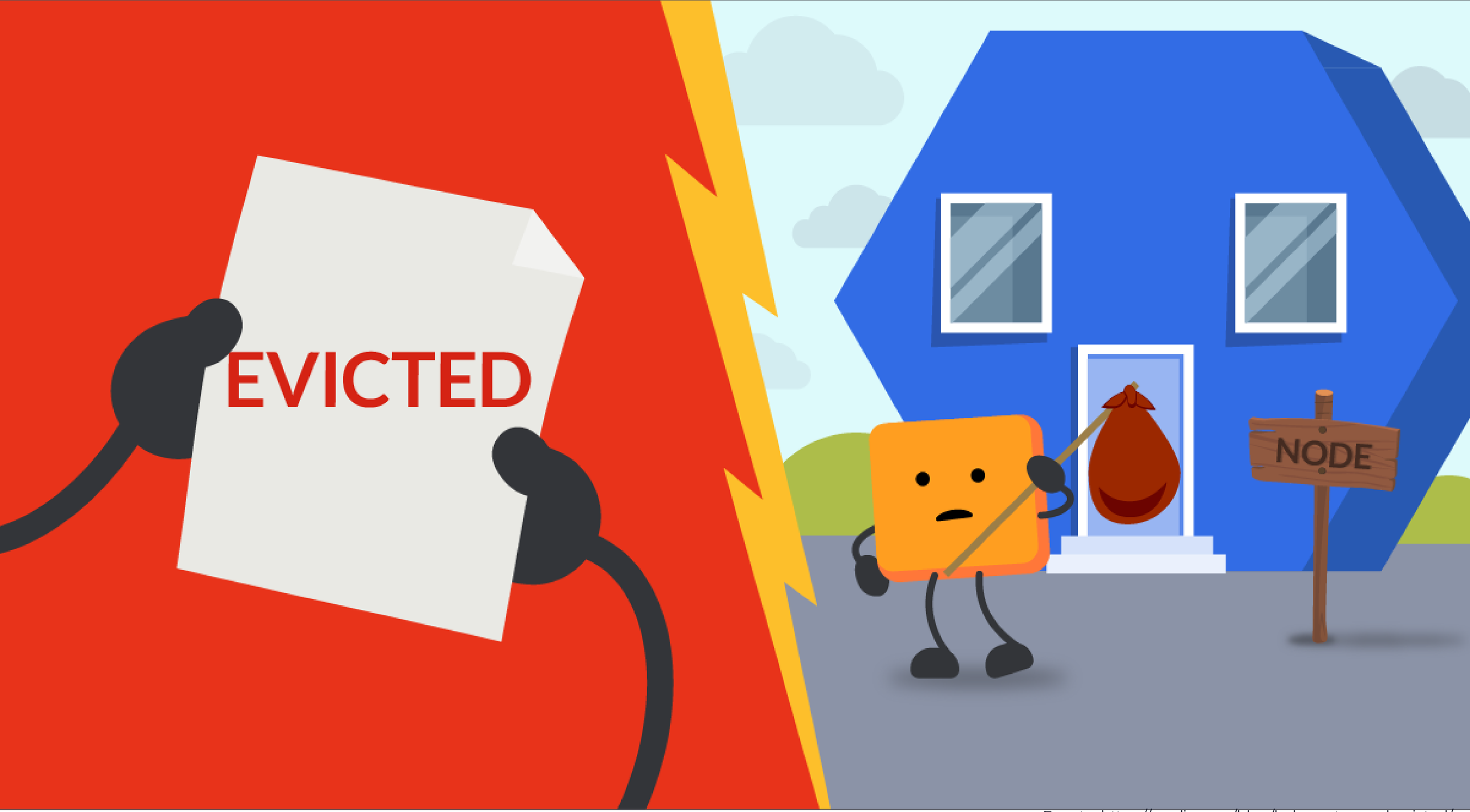
Fuente: https://dev.to/danielepolencic/cpu-requests-and-limits-in-kubernetes-10mh

```yaml
apiVersion: apps/v1
kind: Deployment
metadata:
  name: app1-nginx-deployment
  labels:
    app: app1-nginx
spec:
  replicas: 1
  selector:
    matchLabels:
      app: app1-nginx
  template:
    metadata:
      labels:
        app: app1-nginx
    spec:
      containers:
        - name: app1-nginx
          image: stacksimplify/kube-nginxapp1:1.0.0
          imagePullPolicy: Always
          ports:
            - containerPort: 80
          # Requests & Limits for usermgmt-webapp Container
          resources:
            requests:
              cpu: "100m"
              memory: "128Mi"
            limits:
              cpu: "200m"
              memory: "256Mi"
```

Fuente: https://stacksimplify.com/course-images/azure-kubernetes-service-resources-requests-limits-2.png

request & limits

¿ porqué algo tan importante es opcional y no obliga a definirlo de manera mandatoria?



Fuente: https://sysdig.com/blog/kubernetes-pod-evicted/

Pod Priority

Quality of Service Classes (QoS)

Limits:
cpu: 1
memory: 500Mi
Requests:
cpu: 1
memory: 500Mi

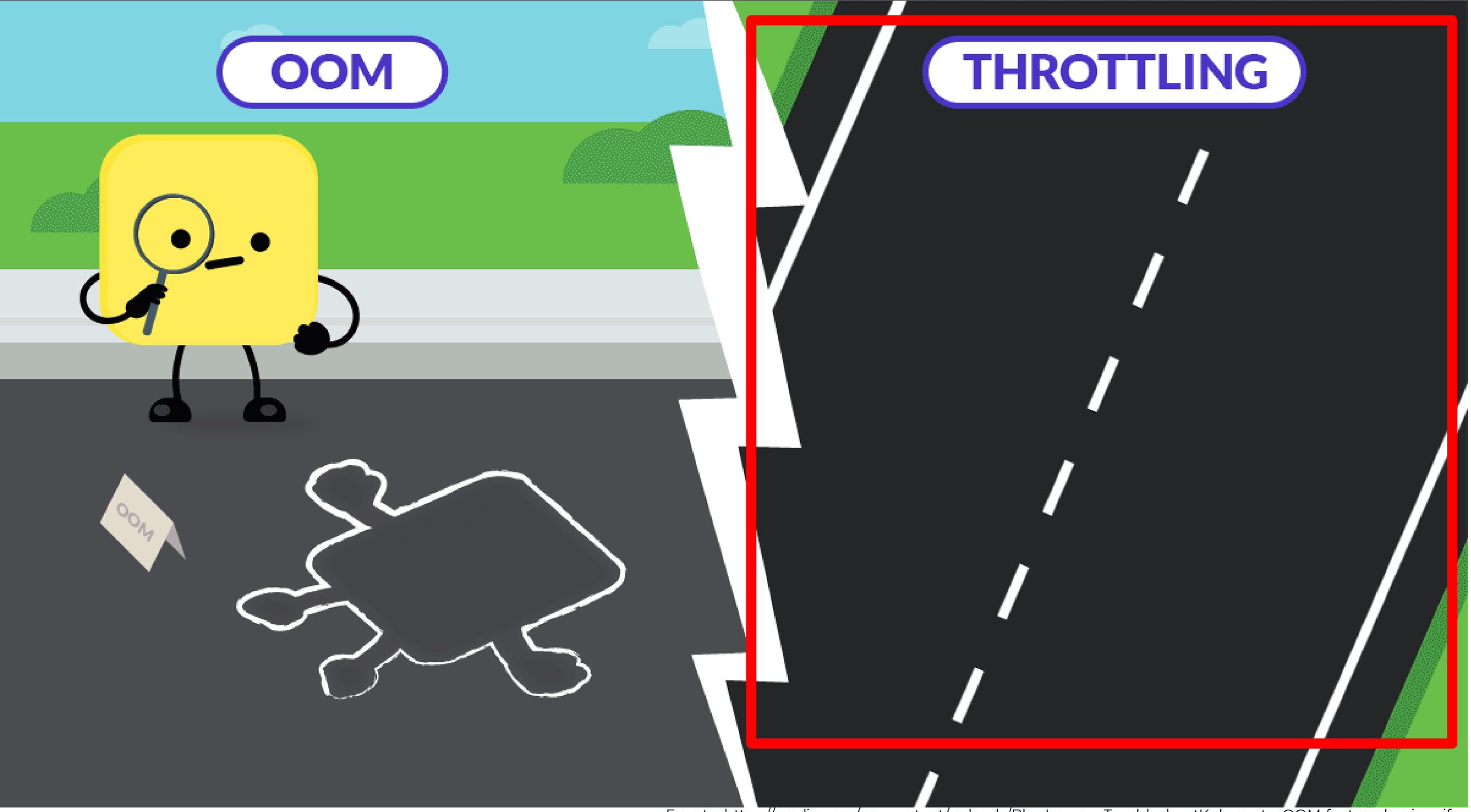
Guaranteed

Limits:
cpu: 500m
memory: 500Mi
Requests:
cpu: 300m
memory: 400Mi

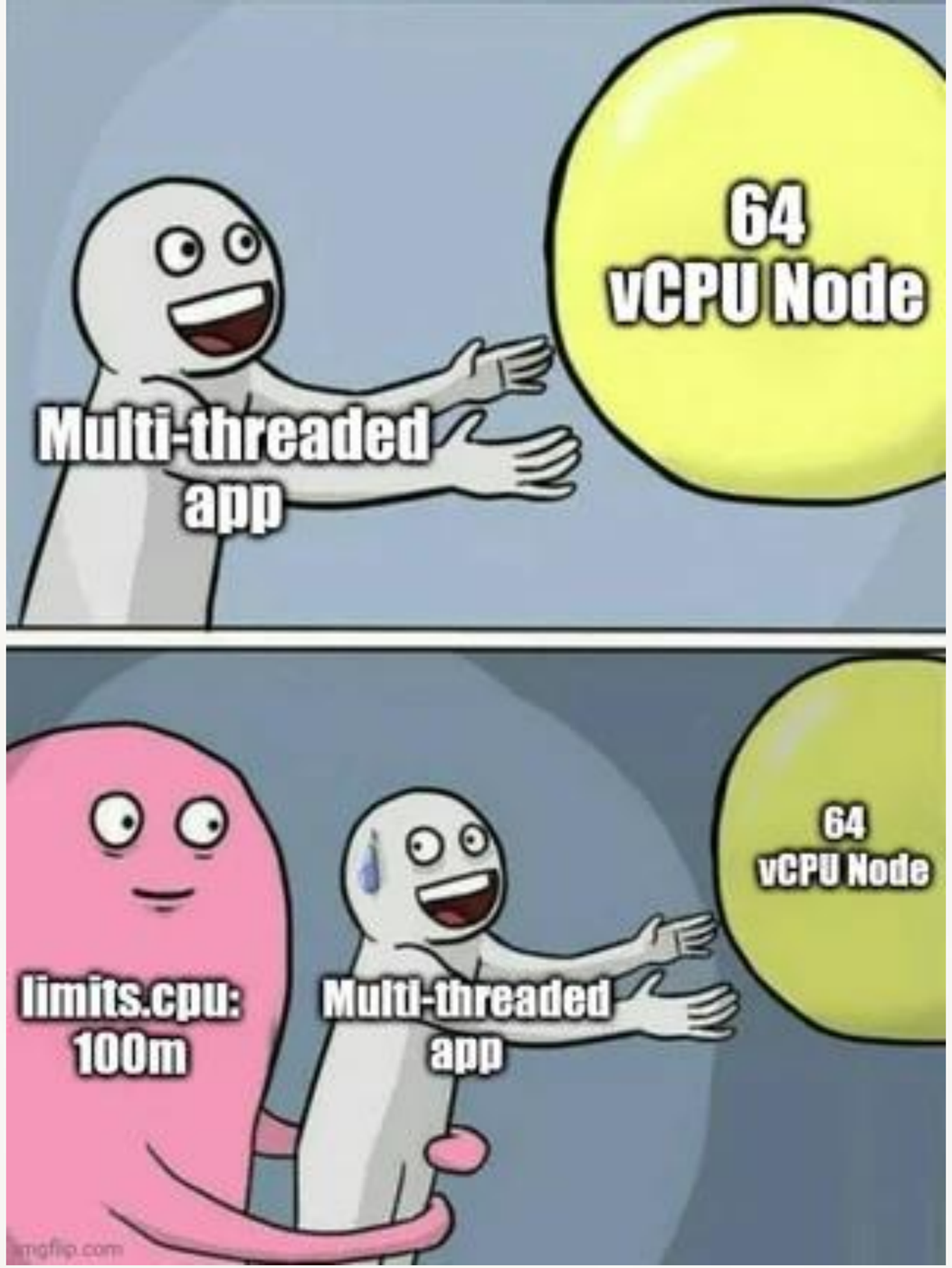
Burstable

<No Limits or Requests>

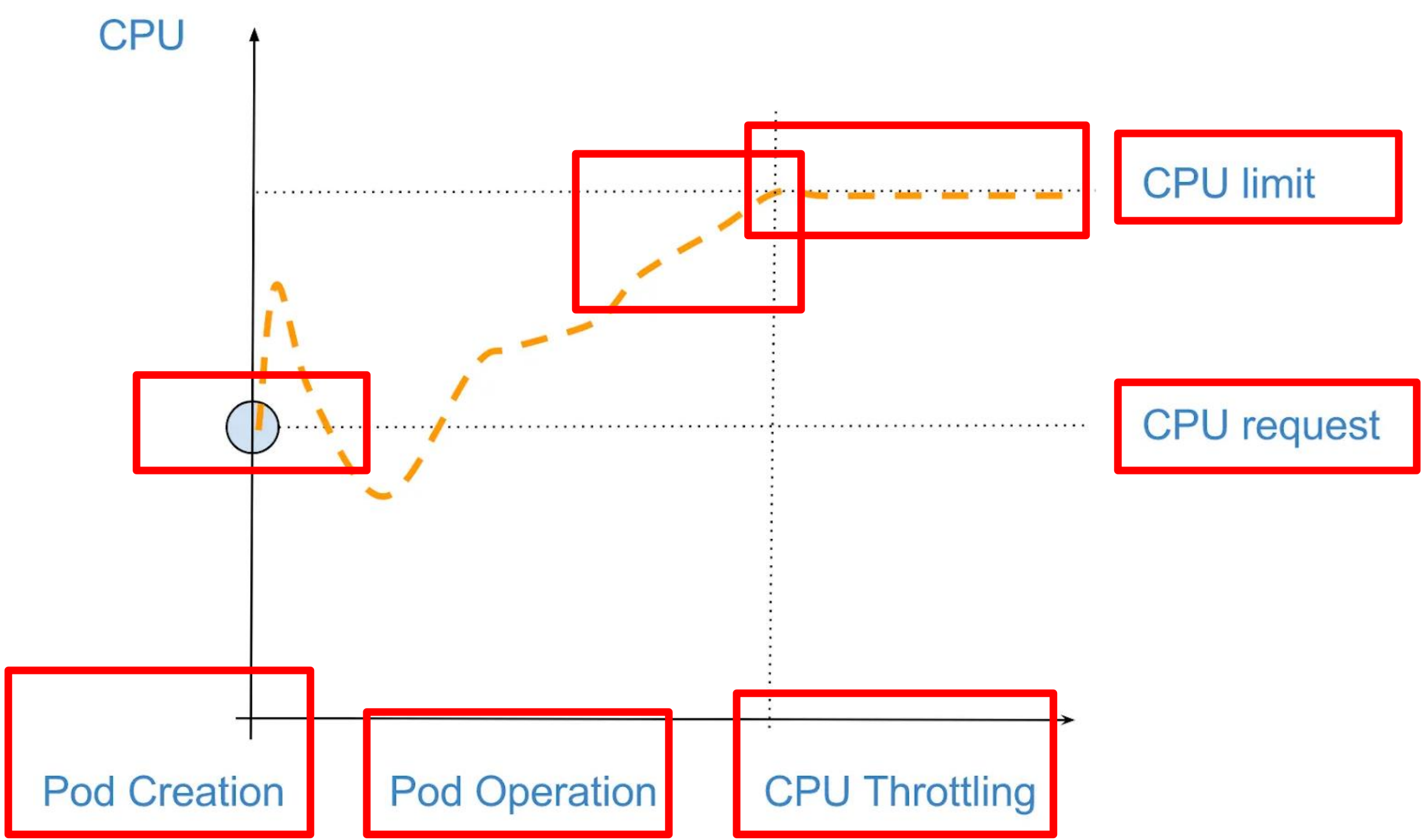
BestEffort



Fuente: https://sysdig.com/wp-content/uploads/BlogImages-TroubleshootKubernetesOOM-featured-anim.gif



Fuente: https://www.reddit.com/r/kubernetes/comments/qdcex8/humor_cpu_shares/



Fuente: https://faun.pub/how-to-use-priorities-in-kubernetes-e1bb1b722b6a

DEBATE

¿Deberíamos establecer límites de CPU o dejar que los contenedores tomen lo que necesiten?

Request & Limits

10 consejos prácticos para no sufrir con tus aplicaciones...

1

Siempre define CPU request: garantiza planificación adecuada

2

Usa Pod Priority: protege tus cargas críticas..

3

Evita CPU limit salvo que tengas razones técnicas claras para usarlos.

4

```
kind: LimitRange
```

```
limits:
  cpu:
  "500m"
  memory:
  "1Gi"
```

Fuente: generado con chatgtp..com

5

```
kind: ResourceQuota
```

```
hard:
  cpu: "4"
  memory: "8Gi"
```

Fuente: generado con chatgtp..com

6

No es solo monitorear: entiende el impacto del throttling en tu aplicación (Grafana / Prometheus)

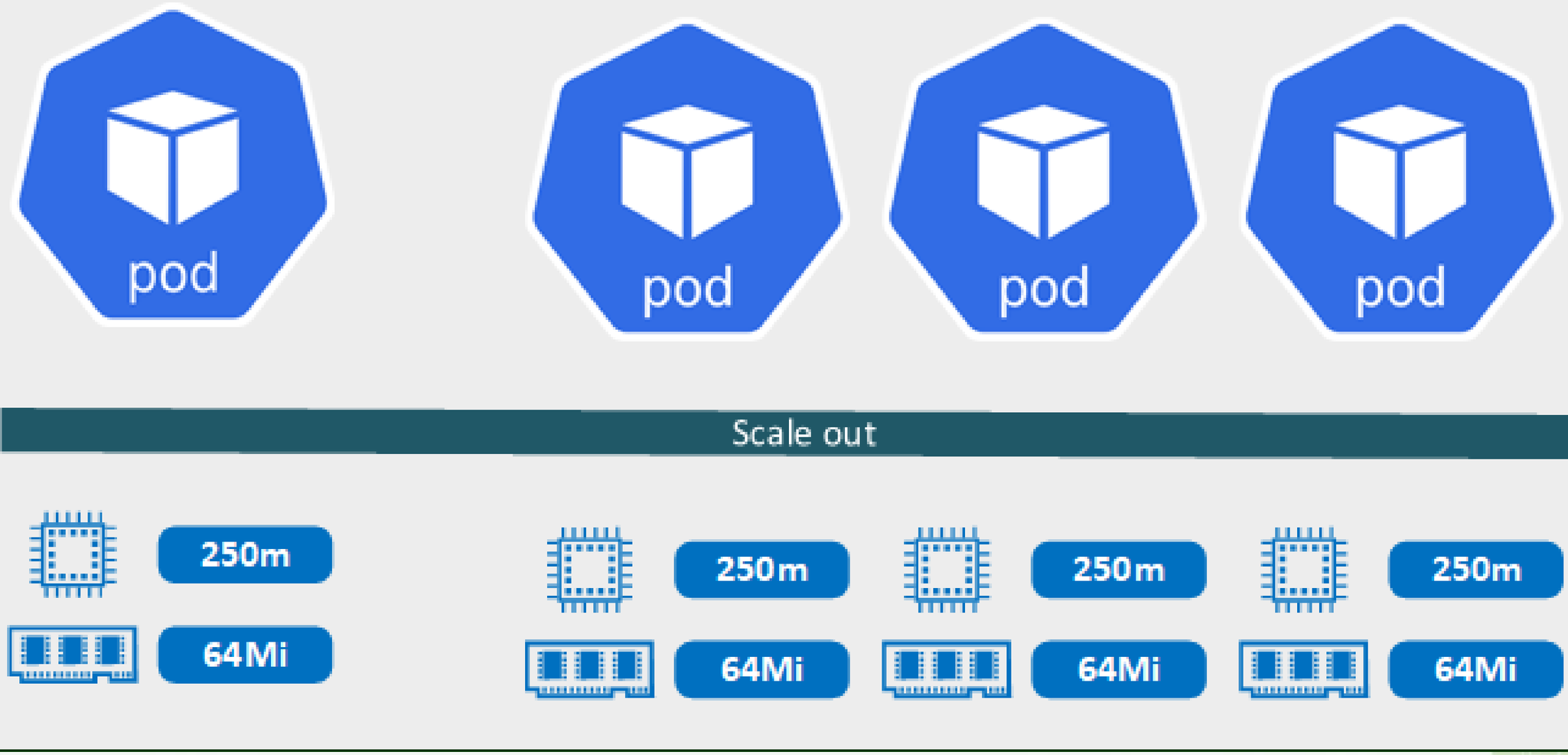
7

Usa limits (al inicio)...

Hasta que sepas cuándo no usarlos (Del límite fijo al escalado inteligente).

8

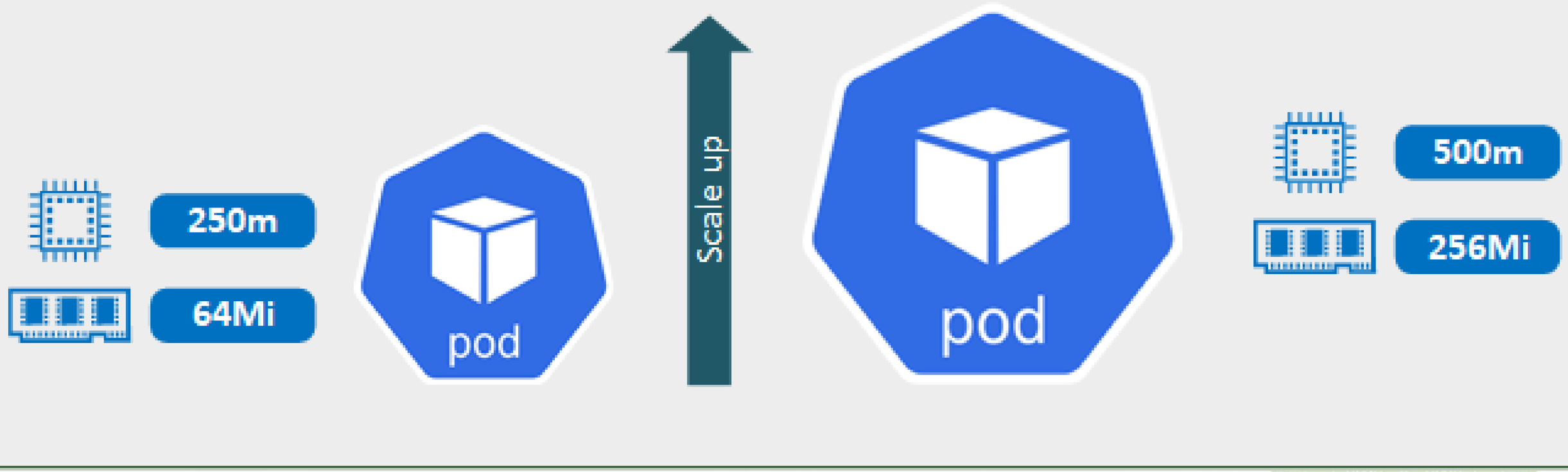
Horizontal Pod Autoscaler (HPA)



Fuente: https://www.kreyman.de/index.php/others/linux-kubernetes/249-was-ist-kubernetes-autoscaling

9

Vertical Pod Autoscaler (VPA)



Fuente: https://www.kreyman.de/index.php/others/linux-kubernetes/249-was-ist-kubernetes-autoscaling

10 Kubernetes Resource Recommender - krr

Scan result (99.96 points)

Number	Cluster	Namespace	Name	Pods	Type	Container	CPU Requests	CPU Limits	Memory Requests	Memory Limits
1.	gke_robusta-development_us-east5-a_arik…	default	crashpod	3	Deployment	crashpod	none -> ?	none -> none	none -> 10M	none -> 10M
2.	gke_robusta-development_us-east5-a_arik…	default	hamster	1	Deployment	hamster	100m -> 171m	300m -> none	50Mi -> 10M	none -> 10M
3.	gke_robusta-development_us-east5-a_arik…	default	ng.inx-deployment	3	Deployment	nginx	none -> 5m	none -> none	none -> 10M	none -> 10M
4.	gke_robusta-development_us-east5-a_arik…	default	nginx-deployment	3	Deployment	nginx	none -> 5m	none -> none	none -> 10M	none -> 10M
5.	gke_robusta-development_us-east5-a_arik…	kubewatch	ng.inx-deployment	1	Deployment	nginx	none -> 5m	none -> none	none -> 10M	none -> 10M
6.	gke_robusta-development_us-east5-a_arik…	robusta	inline-crashpod	1	Deployment	crashpod	none -> ?	none -> none	none -> 10M	none -> 10M
7.	gke_robusta-development_us-east5-a_arik…	robusta	robusta-forwarder	1	Deployment	kubewatch	10m -> 8m	none -> none	512Mi -> 37M	512Mi -> 37M
8. 9. 10.	gke_robusta-development_us-east5-a_arik…	robusta	robusta-grafana	1	Deployment	grafana-sc-dashboard grafana-sc-datasources grafana	none -> 9m none -> 5m none -> 5m	none -> none none -> none none -> none	none -> 97M none -> 93M none -> 79M	none -> 97M none -> 93M none -> 79M
11.	gke_robusta-development_us-east5-a_arik…	robusta	robusta-kube-prometheus-st-operator	1	Deployment	kube-prometheus-stack	100m -> 5m	none -> none	none -> 30M	none -> 30M
12.	gke_robusta-development_us-east5-a_arik…	robusta	robusta-kube-state-metrics	1	Deployment	kube-state-metrics	10m -> 5m	none -> none	none -> 19M	none -> 19M
13.	gke_robusta-development_us-east5-a_arik…	robusta	robusta-runner	1	Deployment	runner	250m -> 105m	none -> none	1Gi -> 918M	1Gi -> 918M
14. 15.	gke_robusta-development_us-east5-a_arik…	robusta	alertmanager-robusta-kube-prometheus-st-…	1	StatefulSet	alertmanager config-reloader	50m -> 5m 200m -> 5m	none -> none none -> none	200Mi -> 36M 50Mi -> 10M	none -> 36M 50Mi -> 10M
16. 17.	gke_robusta-development_us-east5-a_arik…	robusta	prometheus-robusta-kube-prometheus-st-pr…	1	StatefulSet	prometheus config-reloader	50m -> 201m 200m -> 5m	none -> none none -> none	none -> 1060M 50Mi -> 14M	none -> 1060M 50Mi -> 14M
18.	gke_robusta-development_us-east5-a_arik…	robusta	robusta-prometheus-node-exporter	3	DaemonSet	node-exporter	50m -> 5m	none -> none	none -> 16M	none -> 16M

https://github.com/robusta-dev/krr

One more thing {steve jobs}

KCD
GROWING CLOUD NATIVE TOGETHER
El Salvador
SAN SALVADOR 2025

Kubernetes v1.33: In-Place Pod Resize Graduated to Beta

Fuentes: https://kubernetes.io/blog/2025/05/16/kubernetes-v1-33-in-place-pod-resize-beta/
https://www.linkedin.com/feed/update/urn:li:activity:7335198614337576960/

Ten Presente

Estabilidad sin rendimiento es ineficaz.

Rendimiento sin estabilidad es insostenible.

Ten Presente

Kubernetes nos da el equilibrio!

KCD
GROWING CLOUD NATIVE TOGETHER
El Salvador
SAN SALVADOR 2025

Apoyado Por CLOUD NATIVE COMPUTING FOUNDATION

/adalberto-garcia

@fhcloudnative

/adalj8d/kubernetes-resources-demo/

09/oct/25