



PRÉSENTÉ PAR



DENIS GERMAIN
BLOG.ZWINDLER.FR



QUENTIN JOLY
UNE-TASSE-DE.CAFE



DEVFEST NANTES - 2025

LIMITS, REQUESTS, QOS, PRIORITYCLASSES

**ON BALAIE CE QUE VOUS PENSIEZ SAVOIR
SUR LE SCHEDULING DANS KUBERNETES**

LUCCA

DENIS GERMAIN / QUENTIN JOLY

DEVFEST NANTES - 2025

LRQPQBCQVPSSLSDK

LUCCA

DENIS GERMAIN / QUENTIN JOLY

RAPPELS RAPIDES





**“ON VA METTRE DES
LIMITES PARTOUT !”**

PARTIE 1 : LIMITS ET REQUESTS

- **LIMITS : C'EST LA CONTRAINTE DE RESSOURCES QUE VOTRE POD NE PEUT PAS DÉPASSER.**
- **REQUESTS : C'EST CE QUE VOTRE POD DEMANDE AU SCHEDULER POUR ÊTRE PLACÉ SUR UN NŒUD.**

LES TYPES DE RESSOURCES

COMPRESSIBLES (CPU)

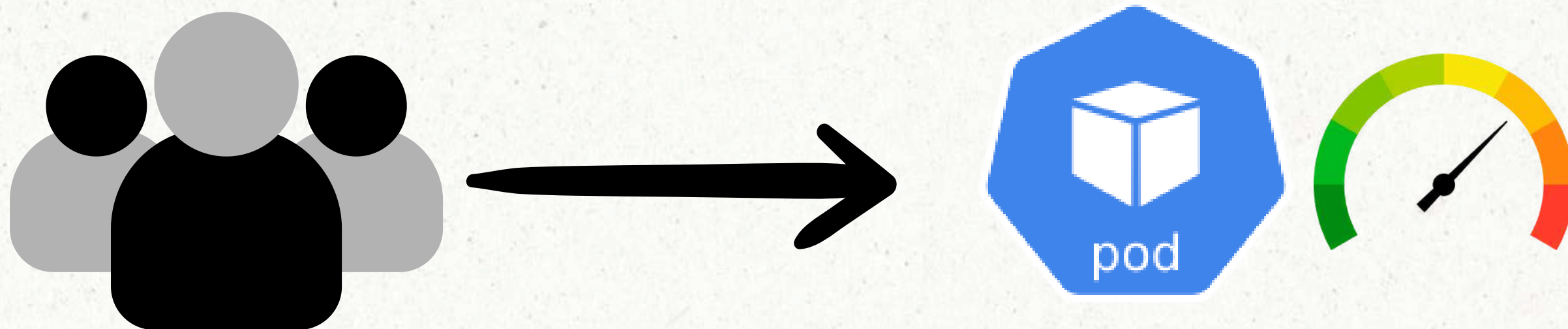
- ON PEUT PARTAGER "À L'INFINI" LES CYCLES CPU
- ÇA SERA JUSTE PLUS LENT 🐌

INCOMPRESSIBLES (RAM, DISQUE)

- ON NE PEUT PAS PARTAGER UN OCTET EN RAM
- SI ON DÉPASSE, ÇA CASSE 💥

EN PRATIQUE

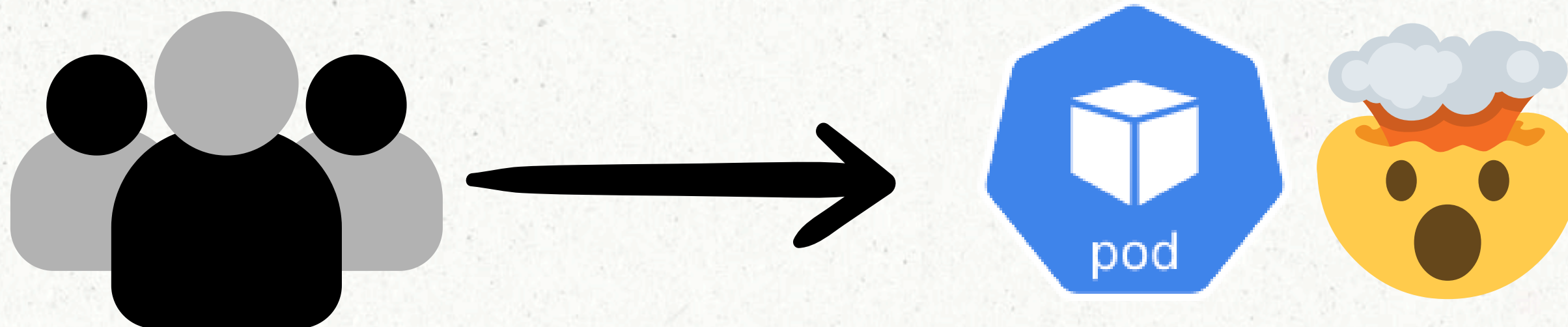
CPU THROTTLING



DEMO 1

METTRE DES LIMITS CPU ?

FBI OU PAS ?



CPU LIMITS

IL SE PASSE QUOI SOUS LE CAPÔT ?

(YAML)

...

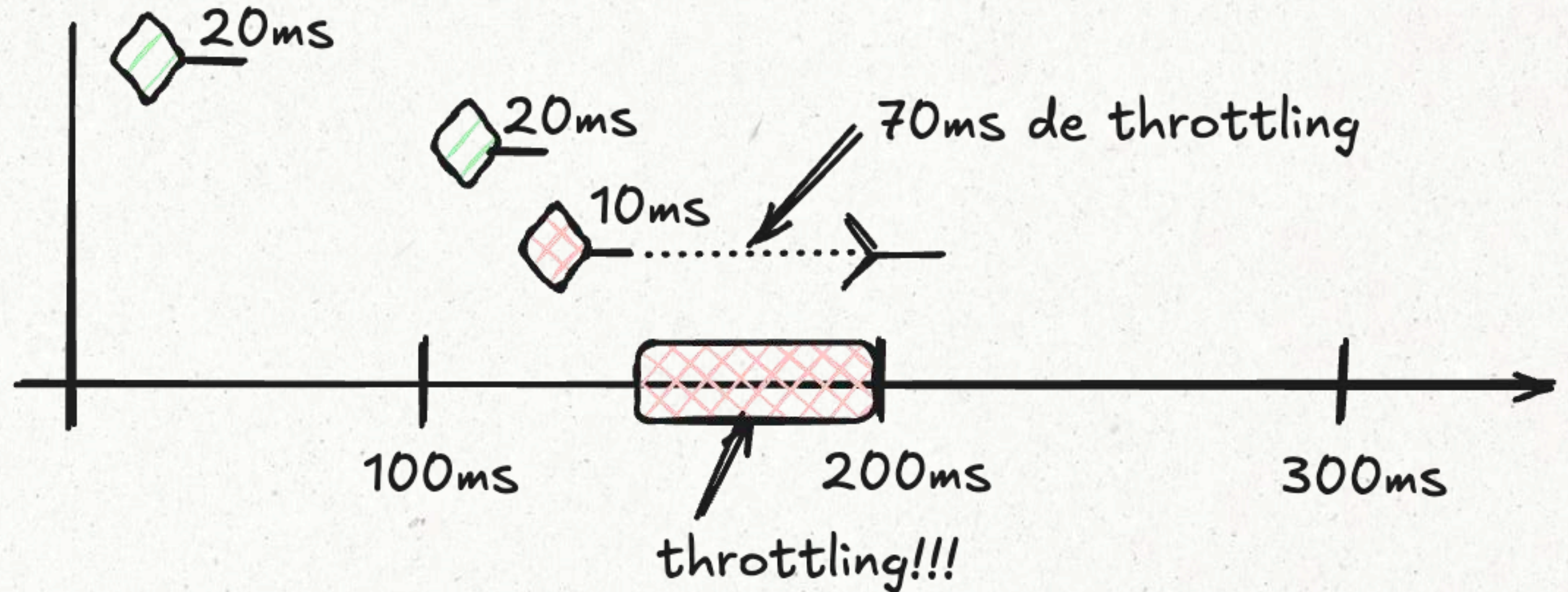
resources:

limits:

cpu: 300m

=

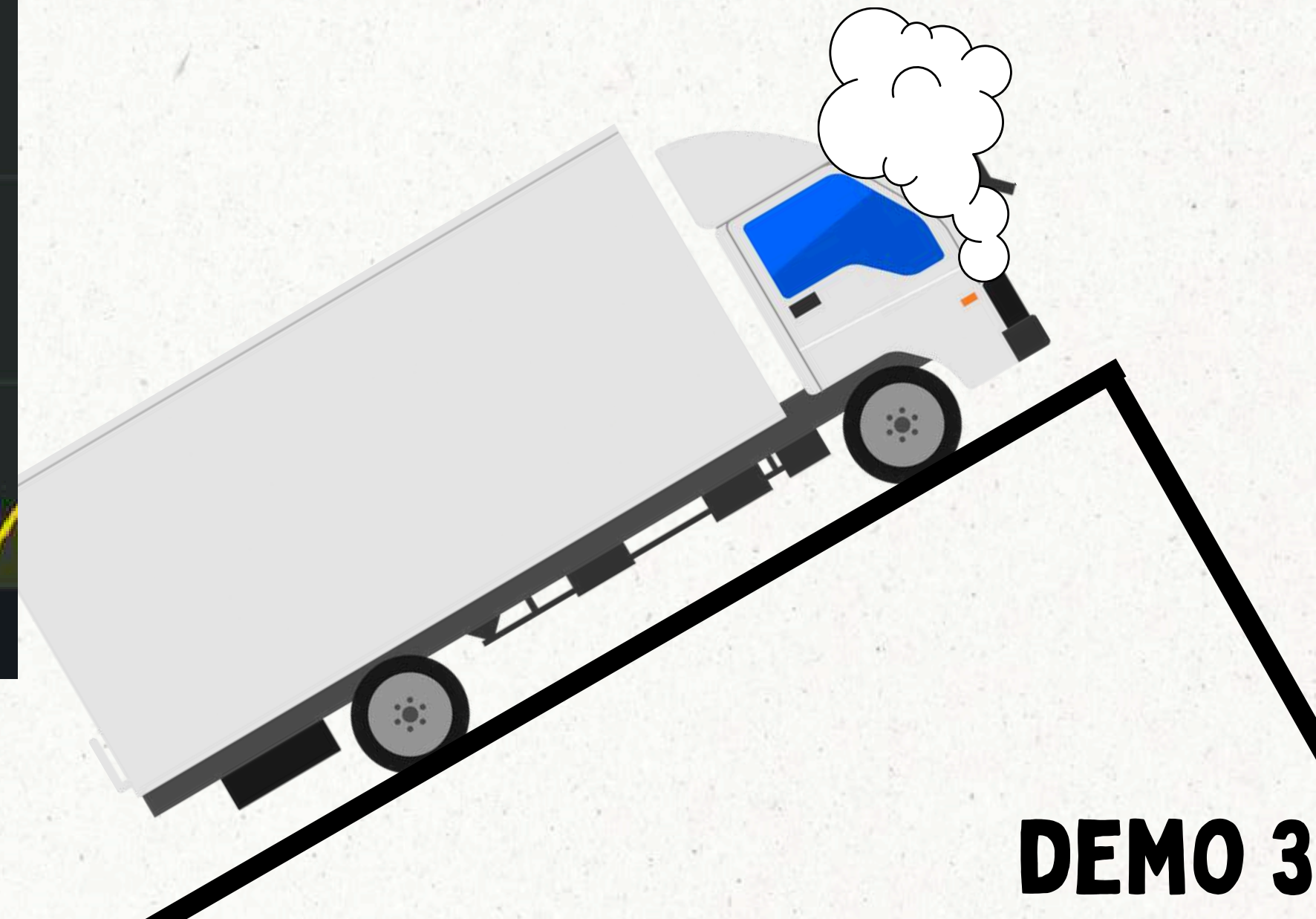
30ms de CPU
toutes les 100 ms



DÉPASSER LE QUOTA D'UNE RESSOURCE INCOMPRESSIBLE ?



OOMKILLED
OUT OF MEMORY



DEMO 3



**"POURQUOI
S'EMBETER A
SPECIFIER DES
REQUESTS ? LE
SCHEDULER N'A QU'A
SE DEBROUILLER..."**

PARTIE 2 : REQUESTS & SCHEDULING

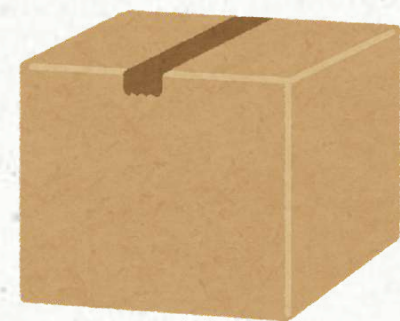
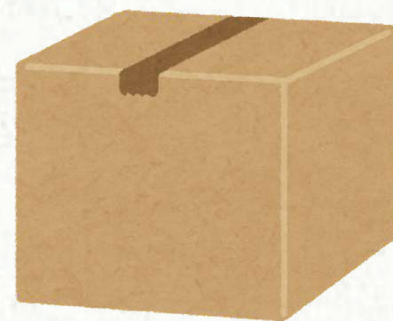
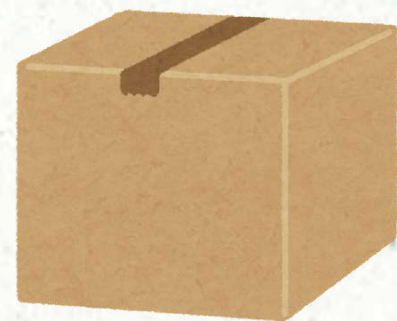
200KG



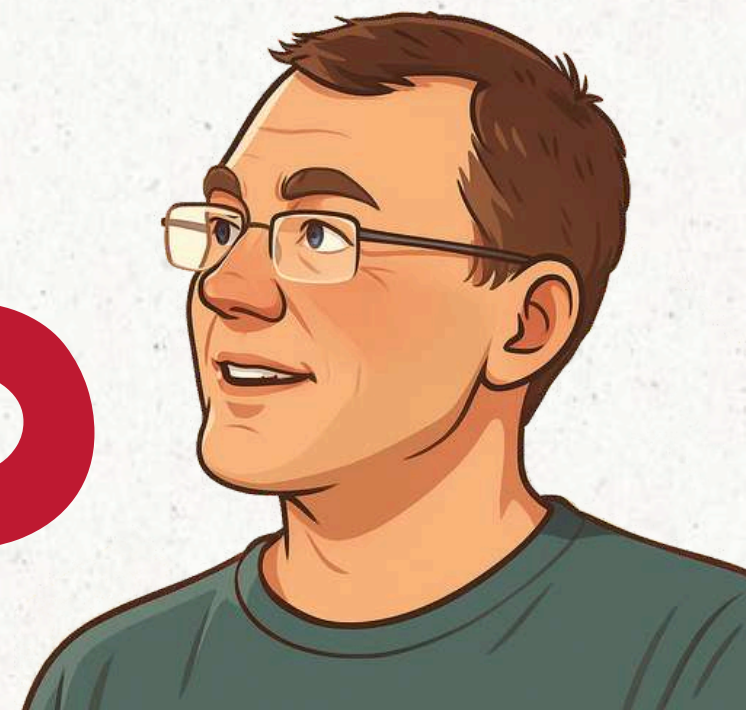
100KG



50KG



COMMENT SAVOIR QUEL CAMION PEUT ACHEMINER TELLE CARGAISON ?
ON NE PEUT PAS SI ON A PAS D'INDICATION



PARTIE 2 : REQUESTS & SCHEDULING

200KG



100KG



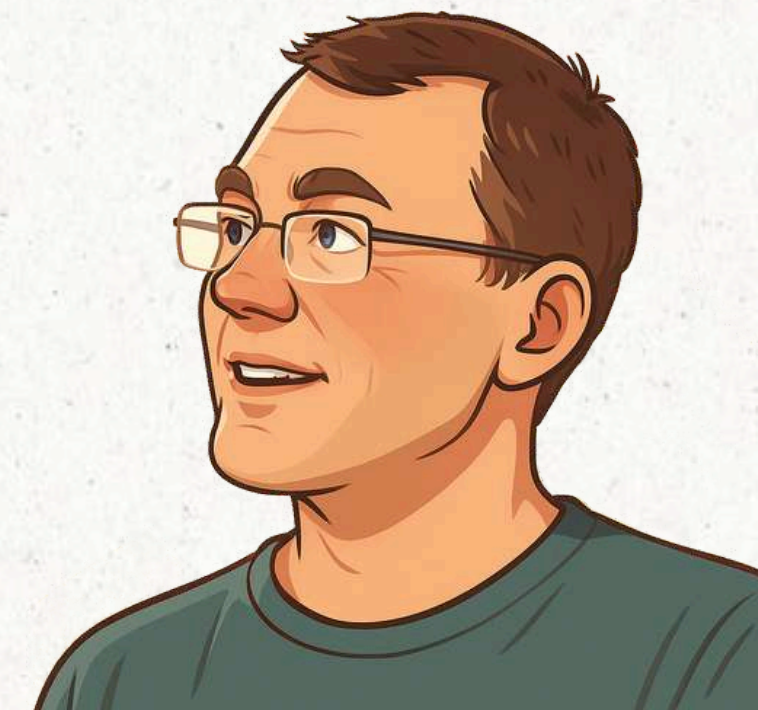
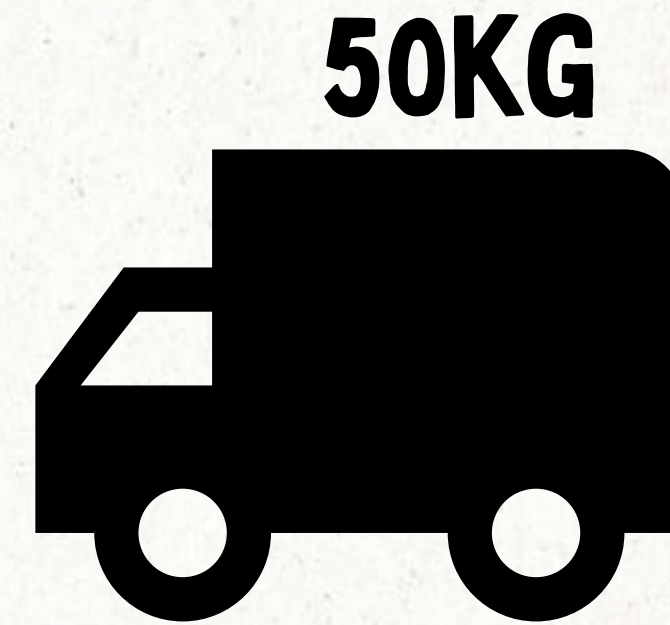
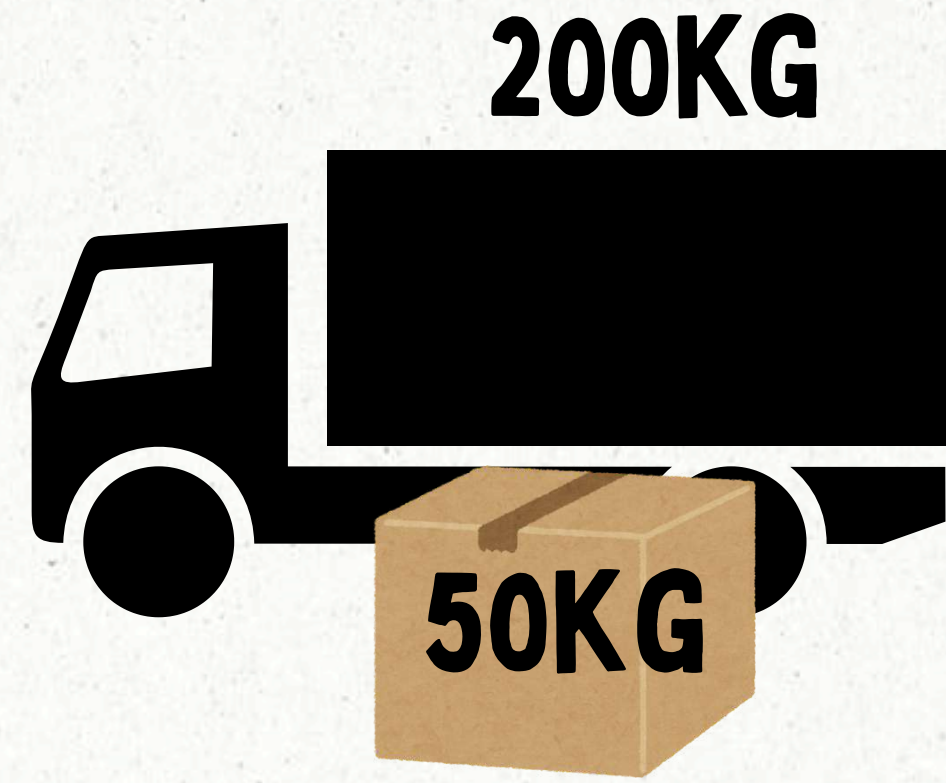
50KG



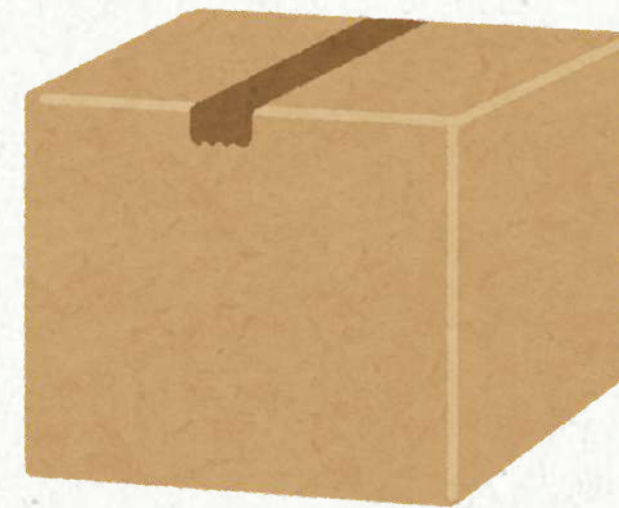
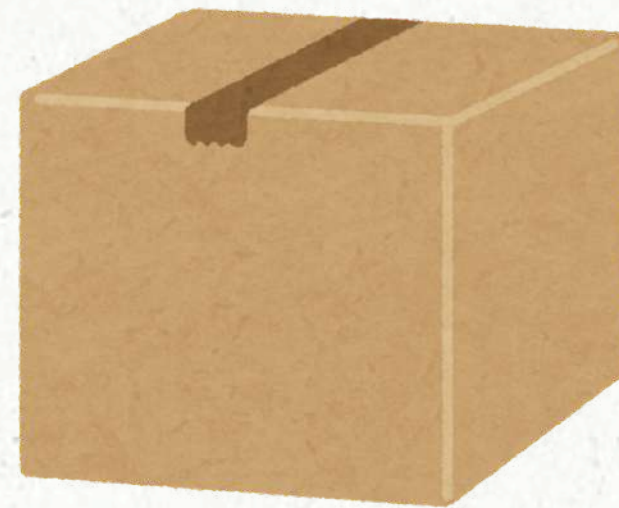
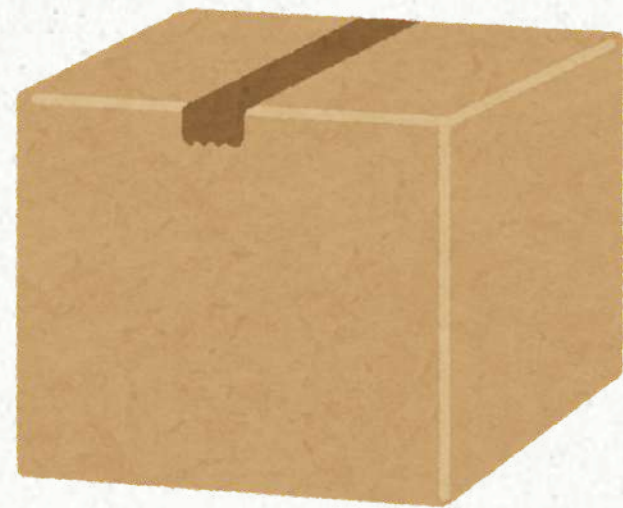
SI ON CONNAIT LE POIDS, ON PEUT FACILEMENT METTRE
LES CARTONS DANS LES BONS CAMIONS



PARTIE 2 : REQUESTS & SCHEDULING



SCHEDULING DES WORKLOADS EN PRATIQUE





**"ON VA METTRE DES
TOUTES PETITES
REQUESTS POUR
ÊTRE SUR QUE NOS
PODS SOIENT
SCHEDULES"**

LEÇON SUR LES CPU SHARES

IL SE PASSE QUOI SOUS LE CAPÔT ?

REQUESTS.CPU

```
[...]  
containers:  
- name: container1  
  resources:  
    requests:  
      memory: "256Mi"  
      cpu: "250m"
```

CPU.SHARES

```
Pod 100m: 4  
Pod 500m: 20
```





**"QU'EST CE QUI SE
PASSE SI ON
CHARGE LA MULE ?"**

PARTIE 3 : QOS

Limits:

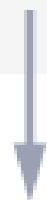
cpu: 1

memory: 500Mi

Requests:

cpu: 1

memory: 500Mi



Guaranteed

Limits:

cpu: 500m

memory: 500Mi

Requests:

cpu: 300m

memory: 400Mi



Burstable

<No Limits or Requests>



BestEffort

ASSURER LA LIVRAISON : L'ÉVICTION





**“ET SI MON CLUSTER
EST PLEIN DE TRUCS
PAS PRIORITAIRES ?”**

PARTIE 4 : PRIORITY CLASSES

COMMENT REMPLIR UN CAMION DÉJÀ PLEIN ?



CE COLIS EST PLUS IMPORTANT !



TAKEAWAYS

**PAS DE LIMITS CPU* (*SAUF EXCEPTION)
REQUESTS / LIMITS RÉALISTES PARTOUT**

VS

QOS	SURVIE EN CAS DE SURCHARGE INATTENDUE
PRIORITY CLASSES	ASSURER LE DÉPLOIEMENT D'APPS CRITIQUES



ET VIENDEZ VOIR LE STAND DES COPAINS

MERCI



↑ Scannez moi !!

SAUREZ-VOUS TROUVER LA DÉMO MANQUANTE?