
MesosCon 2016 - Alex Rukletsov

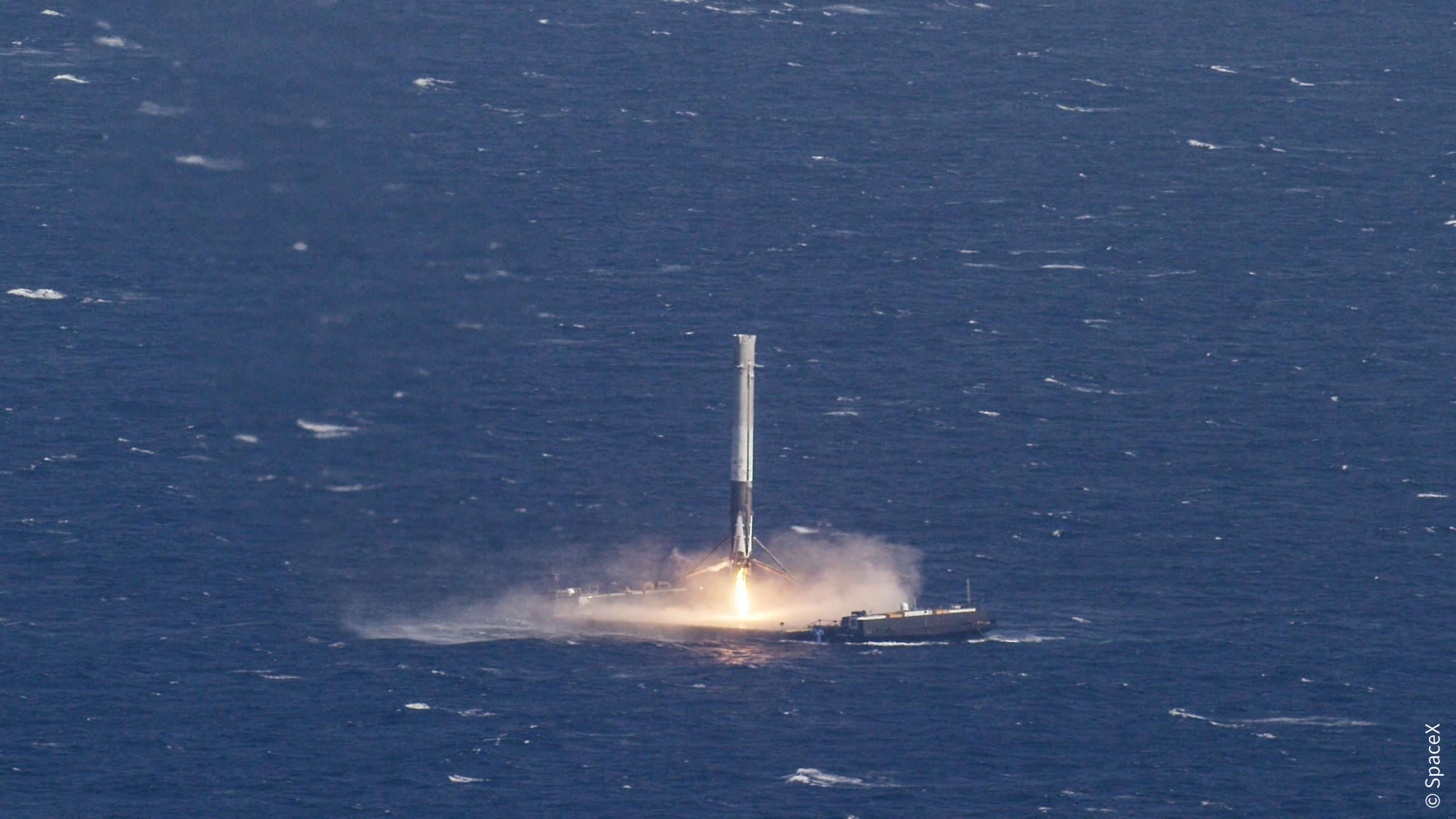
Leveraging Role Quotas to Guarantee and Limit Resource Allocations



Alex Rukletsov

Distributed Systems Engineer

Alex Rukletsov is an Apache committer and Mesos PMC member at Mesosphere. Prior to that Alex was segmenting medical images and investigating behaviour of human vessels in several German research institutes. His areas of interests include distributed systems, object recognition, probabilistic and heuristic algorithms.



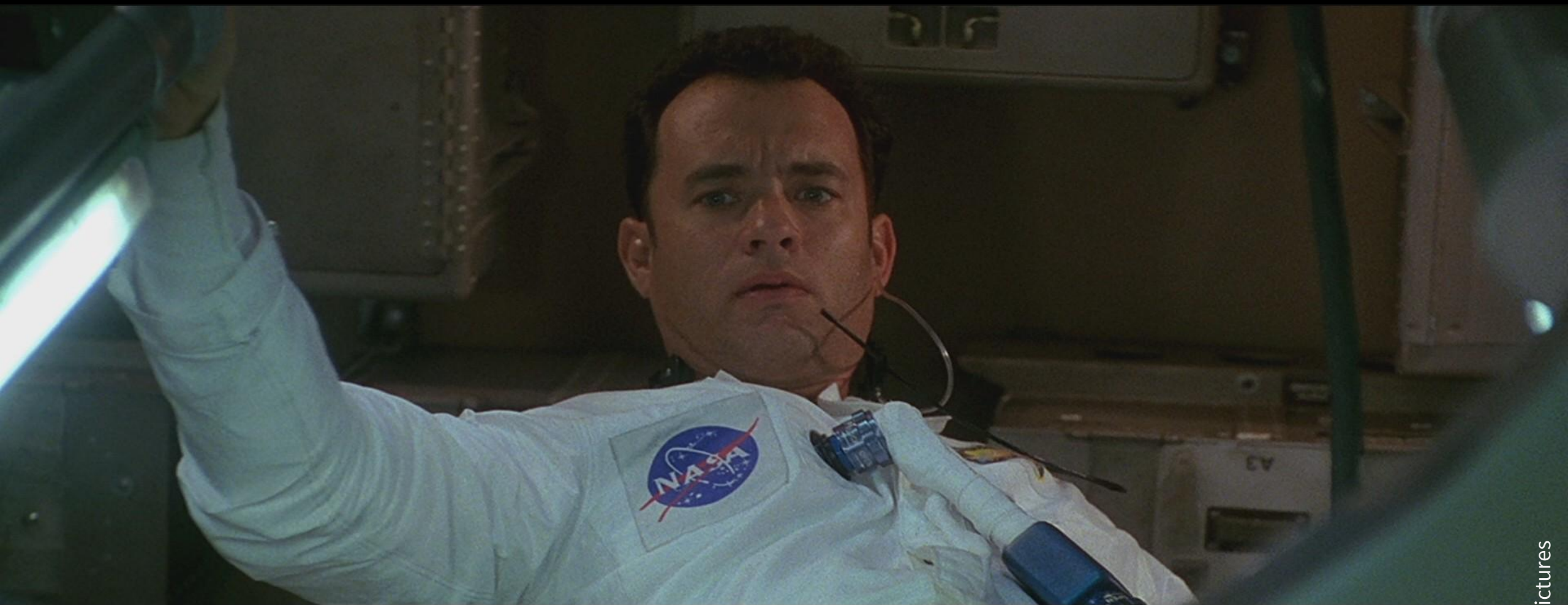
My goals (for this talk):

- share operational experience
- reflect on using quota
- talk about limitations
- explain the future (of quota)
- don't fall asleep



① Why quota ?





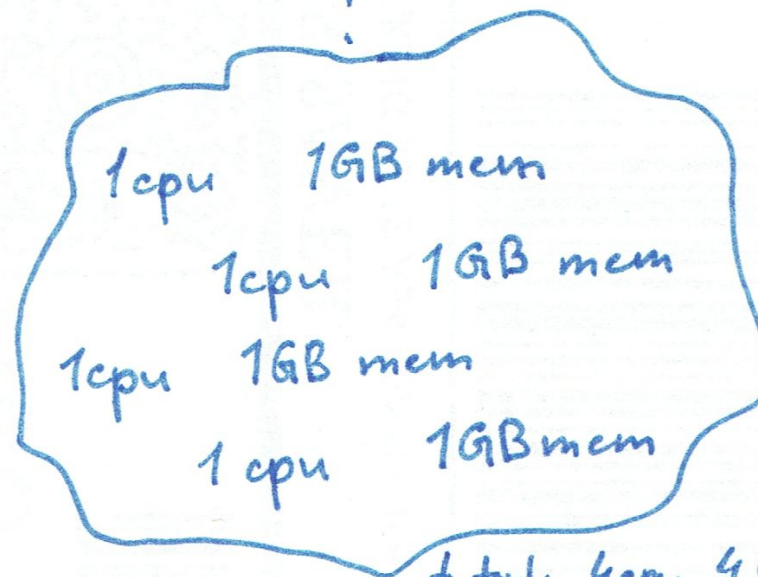
"Houston, we have a problem." Jim Lovell

Frameworks:

① Production Web-App

② Spark analytics

Mesos master



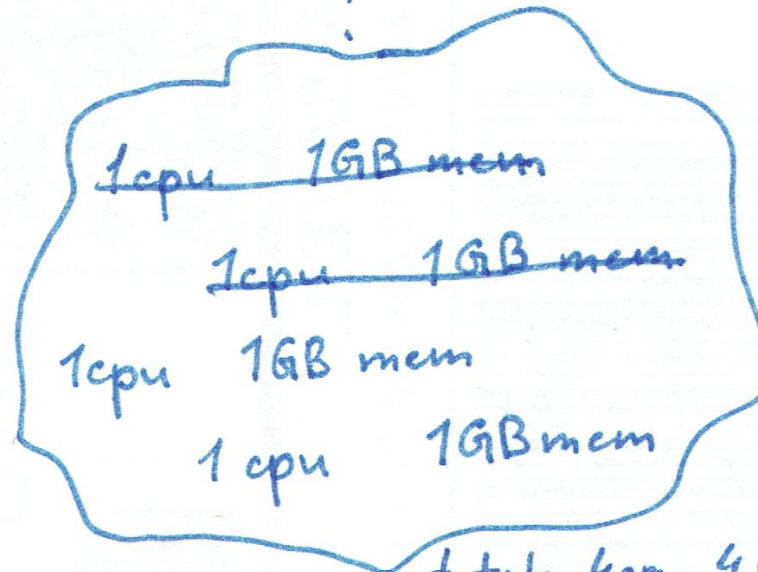
elapsed: 0s

Frameworks:

① Production Web-App
1cpu 1GB mem

② Spark analytics
1cpu 1GB mem

Mesos master



total: 4cpu, 4GB mem

elapsed: 1s

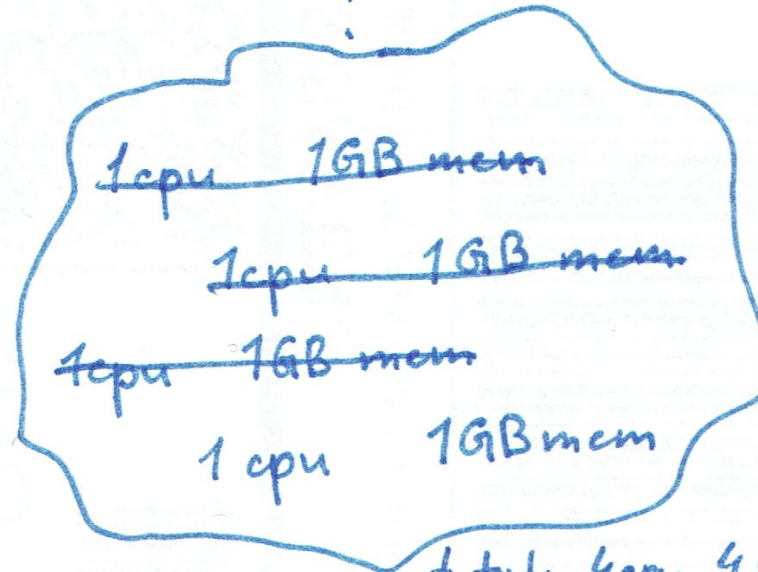
Frameworks:

① Production Web-App
1cpu 1GB mem

decline offer : 5s

② Spark analytics
1cpu 1GB mem
1cpu 1GB mem

Mesos master



total: 4cpu, 4GB mem

elapsed: 1s

Frameworks:

① Production Web-App

1cpu 1GB mem

decline offer : 5s

② Spark analytics

1cpu 1GB mem

1cpu 1GB mem

1cpu 1GB mem

Mesos master

~~1cpu 1GB mem~~

~~1cpu 1GB mem~~

~~1cpu 1GB mem~~

~~1cpu 1GB mem~~

total: 4cpu, 4GB mem

elapsed: 2s

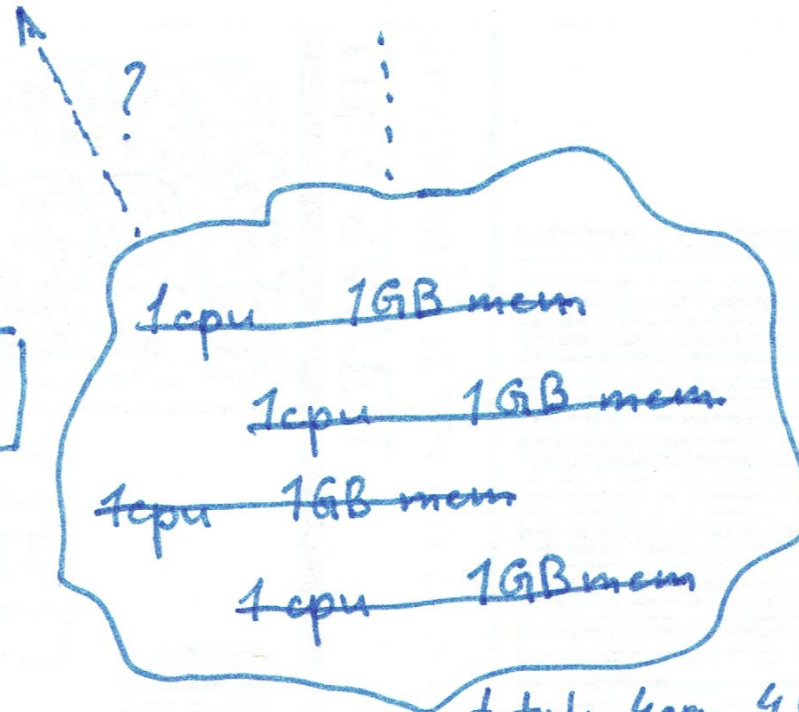
**10 seconds after:
Web-App needs to scale. Now.**

Frameworks:

① Production Web-App
1cpu 1GB mem

② Spark analytics
1cpu 1GB mem
1cpu 1GB mem
1cpu 1GB mem

Mesos master



total: 4cpu, 4GB mem

elapsed: 12s

Frameworks:

① Production Web-App : (

1cpu 1GB mem

② Spark analytics

1cpu 1GB mem

1cpu 1GB mem

1cpu 1GB mem

Mesos master

1cpu 1GB mem

1cpu 1GB mem

1cpu 1GB mem

1cpu 1GB mem

total: 4cpu, 4GB mem

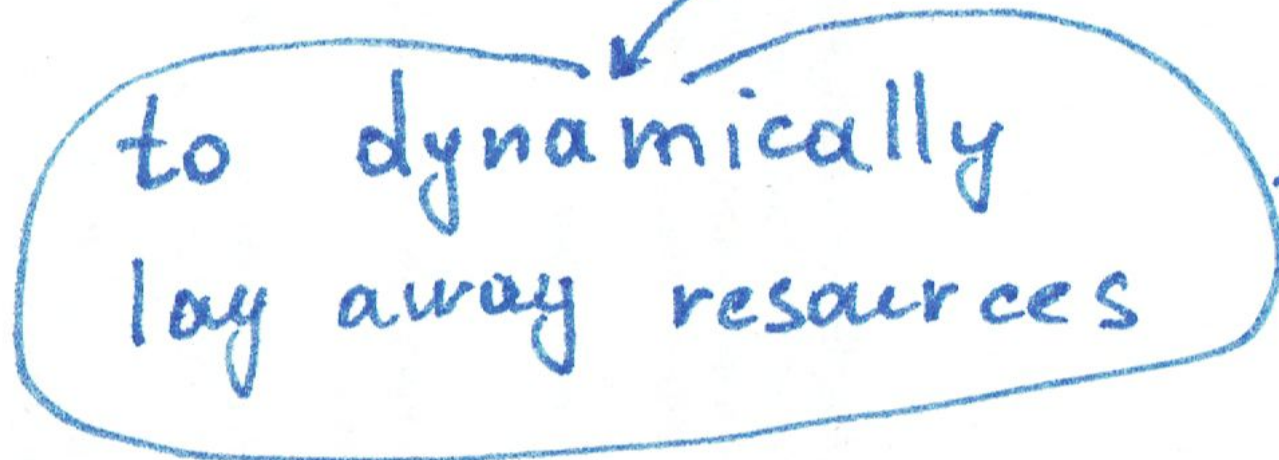
Quota is...

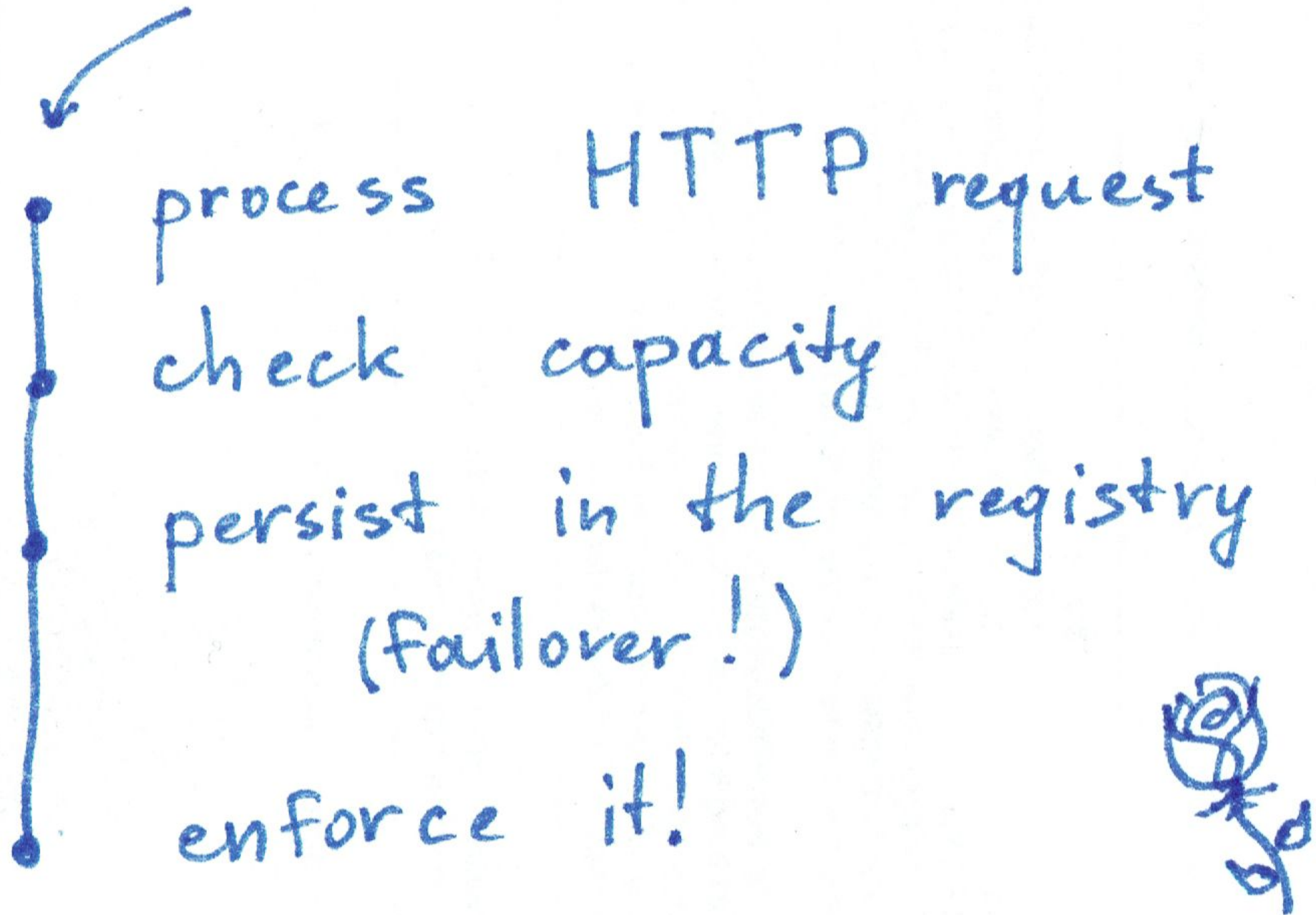
- resource guarantee per role
- RESTlike API for operators

Quota is...

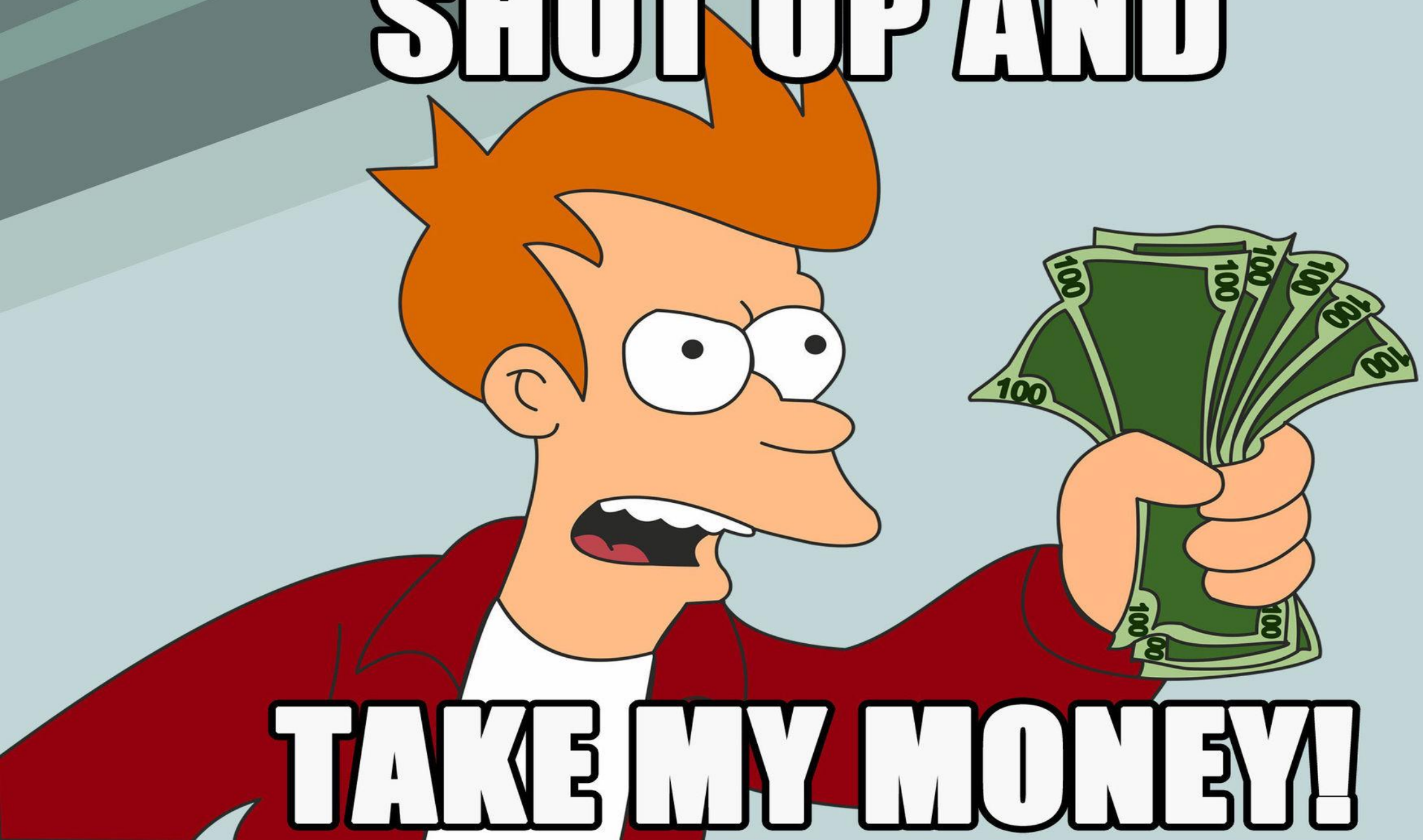
- resource guarantee per role
- RESTlike API for operators

to dynamically
lay away resources

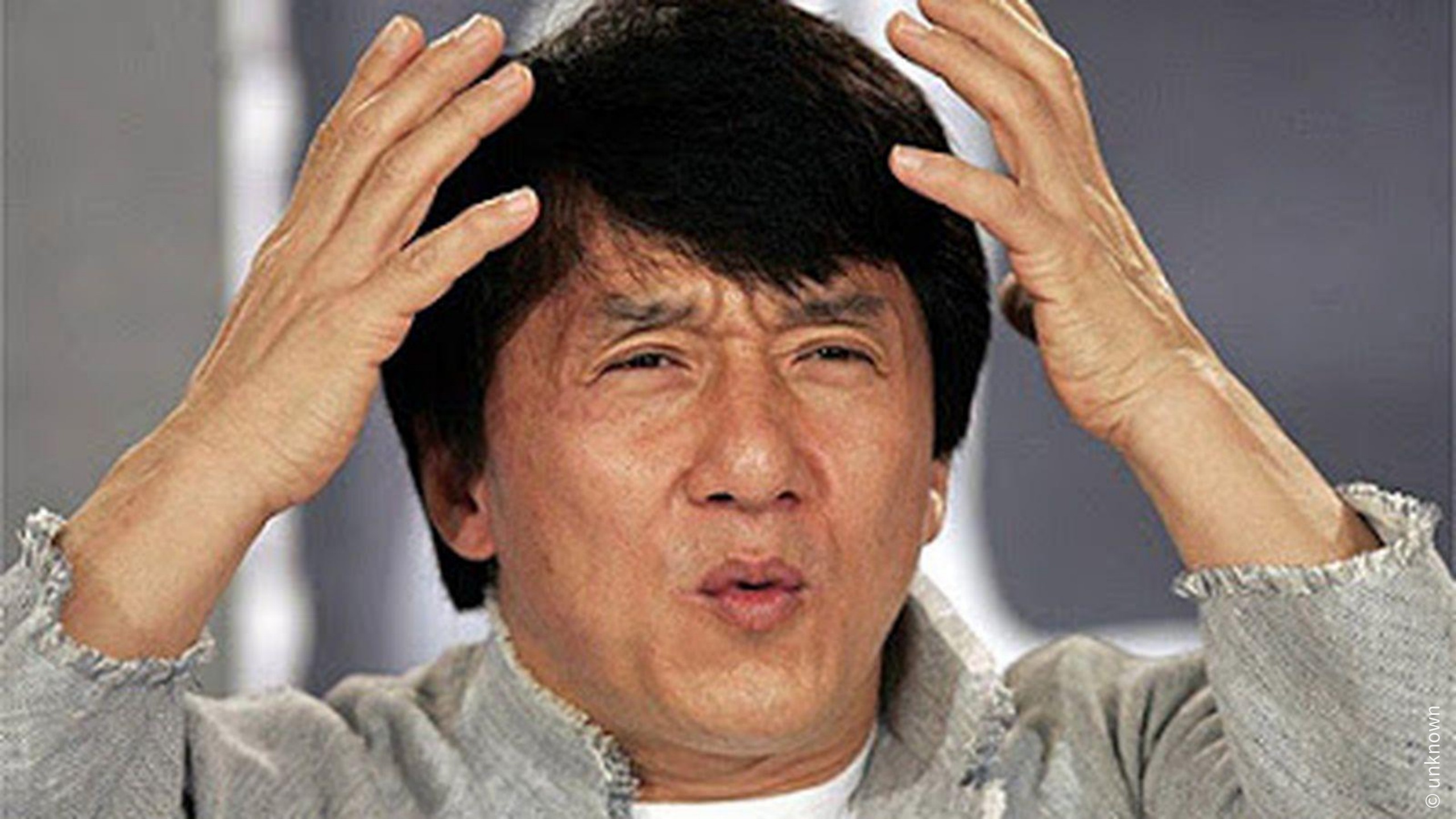




SHUT UP AND



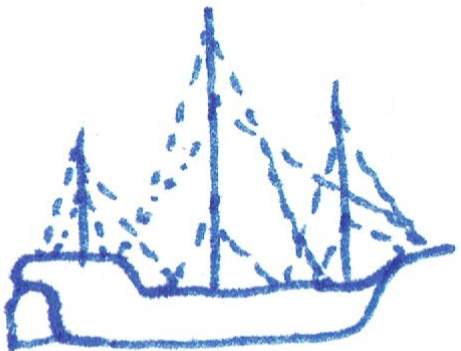
TAKE MY MONEY!



Limitations (some):

- laid away resources are not offered
- limit only, not <guarantee, limit>
- only scalar resources
- no atomic update
- quota is gameable
- no granularity

② Lessons learned



Limitations (some):

- (a) - laid away resources are not offered
- (b) - limit only, not <guarantee, limit>
 - only scalar resources
 - no atomic update
 - quota is gameable
 - no granularity

preemption
↓
oversubscription for quota
↓
(a)

Original idea

Resources beyond guarantee up to fair share.

Protect a production framework

vs.

Current behavior^{u/r}

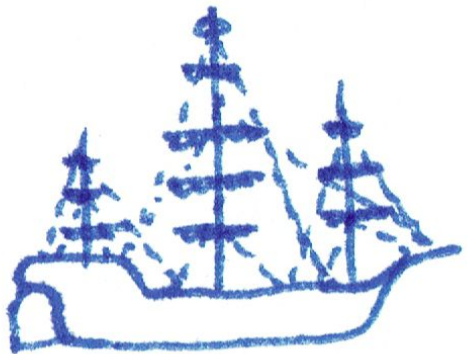
No resources beyond quota; guarantee is also limit

Limit a greedy framework

What's in revocable-by-default world?



③ Common questions (and pitfalls)



1) Quota is per role.

→ careful with multiple frameworks in the same role.

Roles can be implicit!

2) "remove" + "set"

≠

"update"

3) Setting quota may lead to offer rescinding in your cluster.

4) Unused quota'ed resources are idle. They are currently not offered to other frameworks.

5) After quota is stored, next available resources are laid away.

6) If there are not enough resources to satisfy quota, it is either :

- rejected, or
- 'force' d

7) Quota

- quantity (somewhere)
- on free resources
- only for operators

Dynamic reservations

- identity (specific node)
- also on offered
- also for frameworks

vs

behaviour in presence of failures

AOL

MONTH

OCT

DAY

28

YEAR

2005

AM

PM

HOUR

04

MIN

29

DESTINATION TIME

MONTH

JUL

DAY

02

YEAR

2002

AM

PM

HOUR

08

MIN

49

PRESENT TIME

MONTH

OCT

DAY

28

YEAR

1985

AM

PM

HOUR

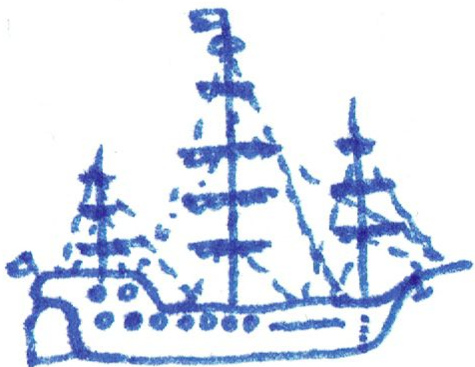
08

MIN

20

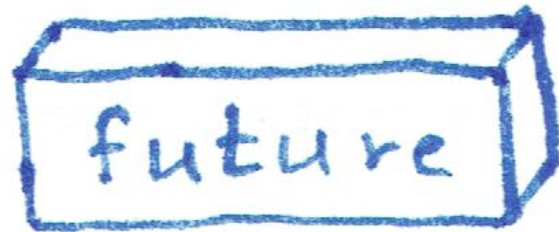
LAST TIME DEPARTED

④ Future work



The Future:

- atomic update
- hierarchical roles support
- oversubscription for quota





VISIT OUR BOOTH

Learn more by visiting dcos.io and mesosphere.com

THANK YOU!