



CLOUD
NATIVE
CON
Europe 2017



KubeCon
A CNCF EVENT



Kubernetes Scheduling Features

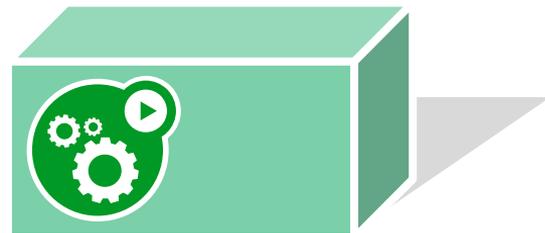
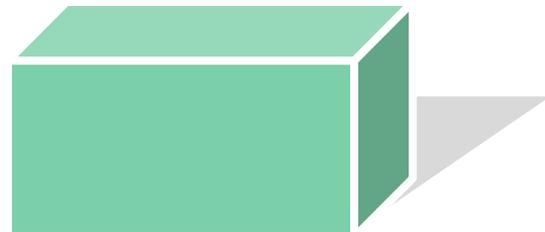
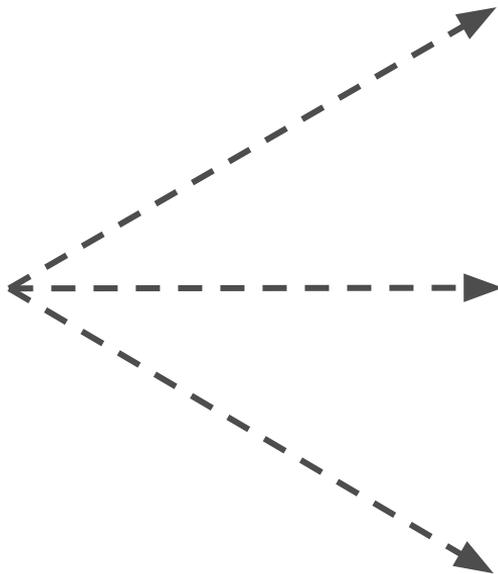
Or

How Can I Make the System Do What I Want?

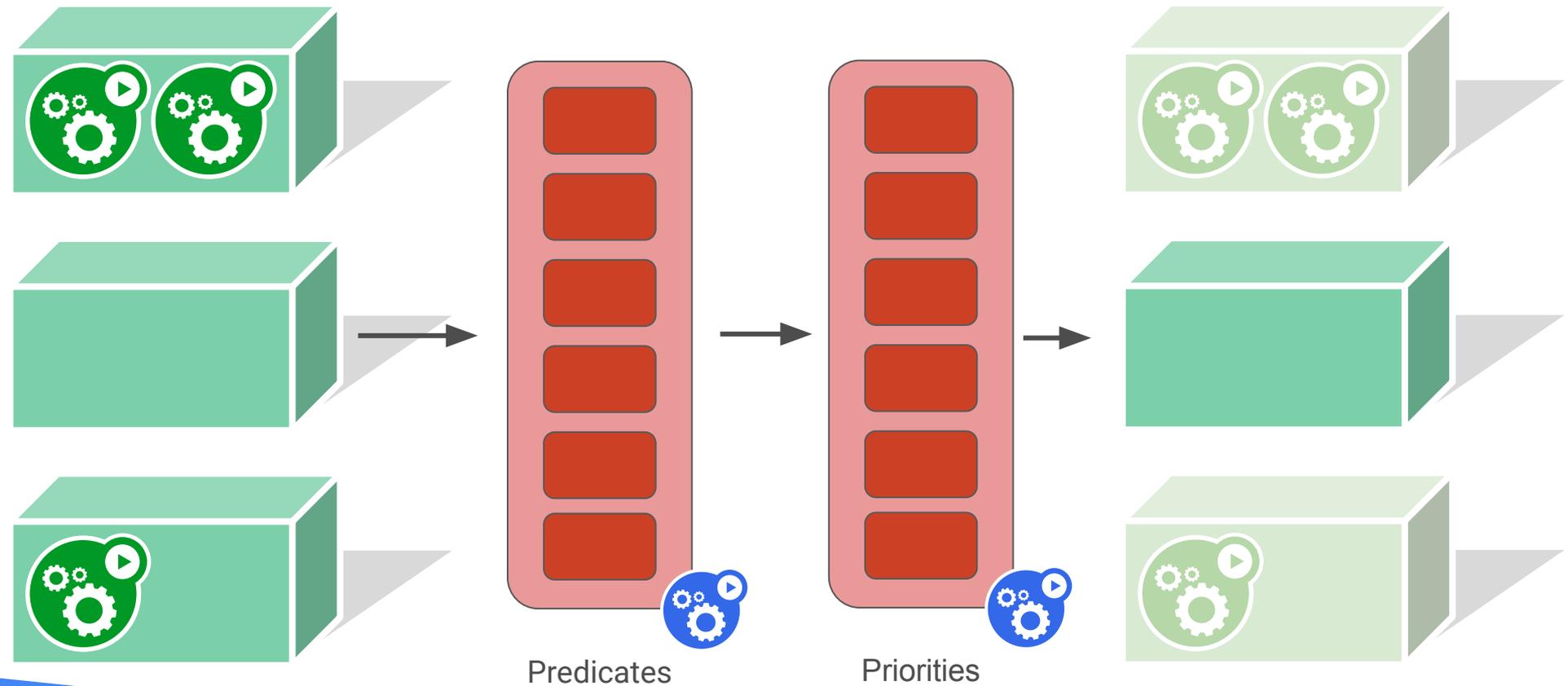
Marek Grabowski, Software Engineer, Google

Wojciech Tyczyński, Senior Software Engineer, Google

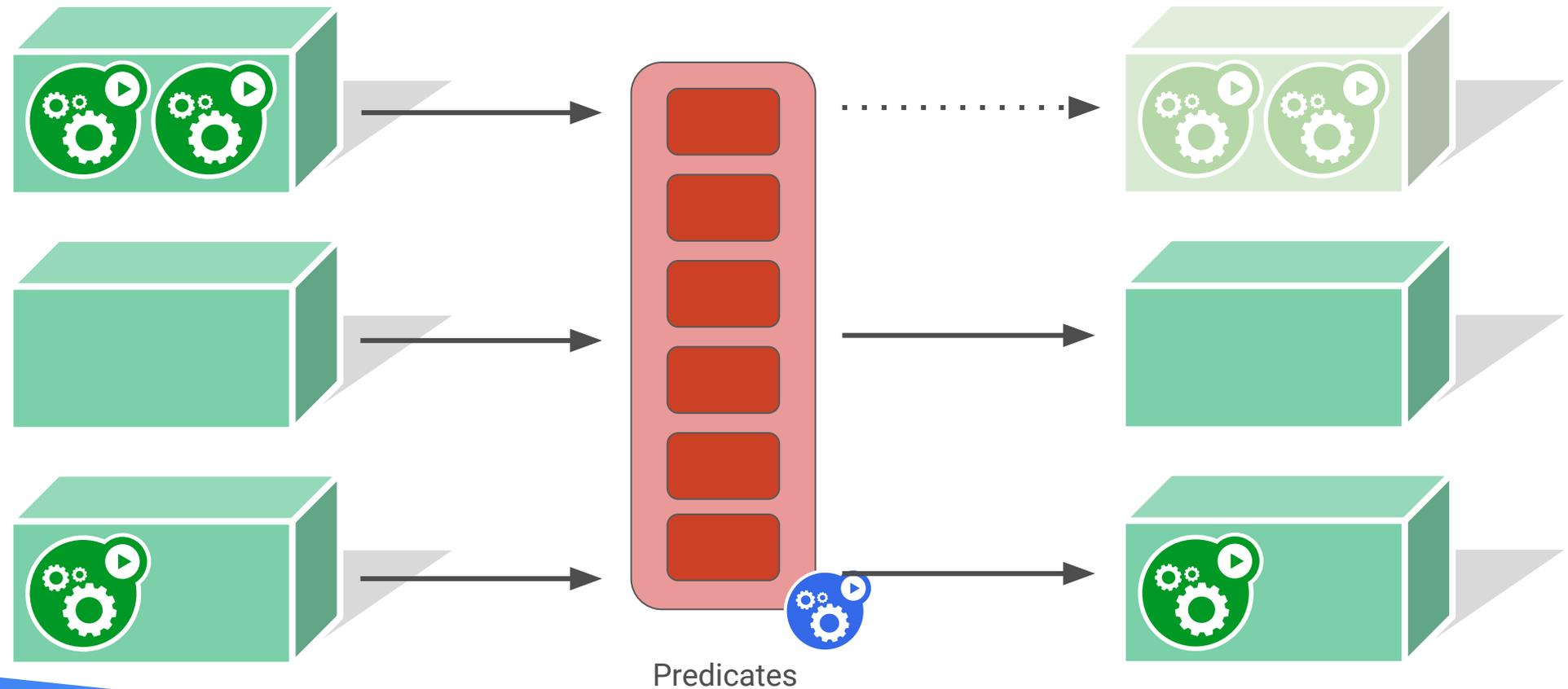
Scheduling



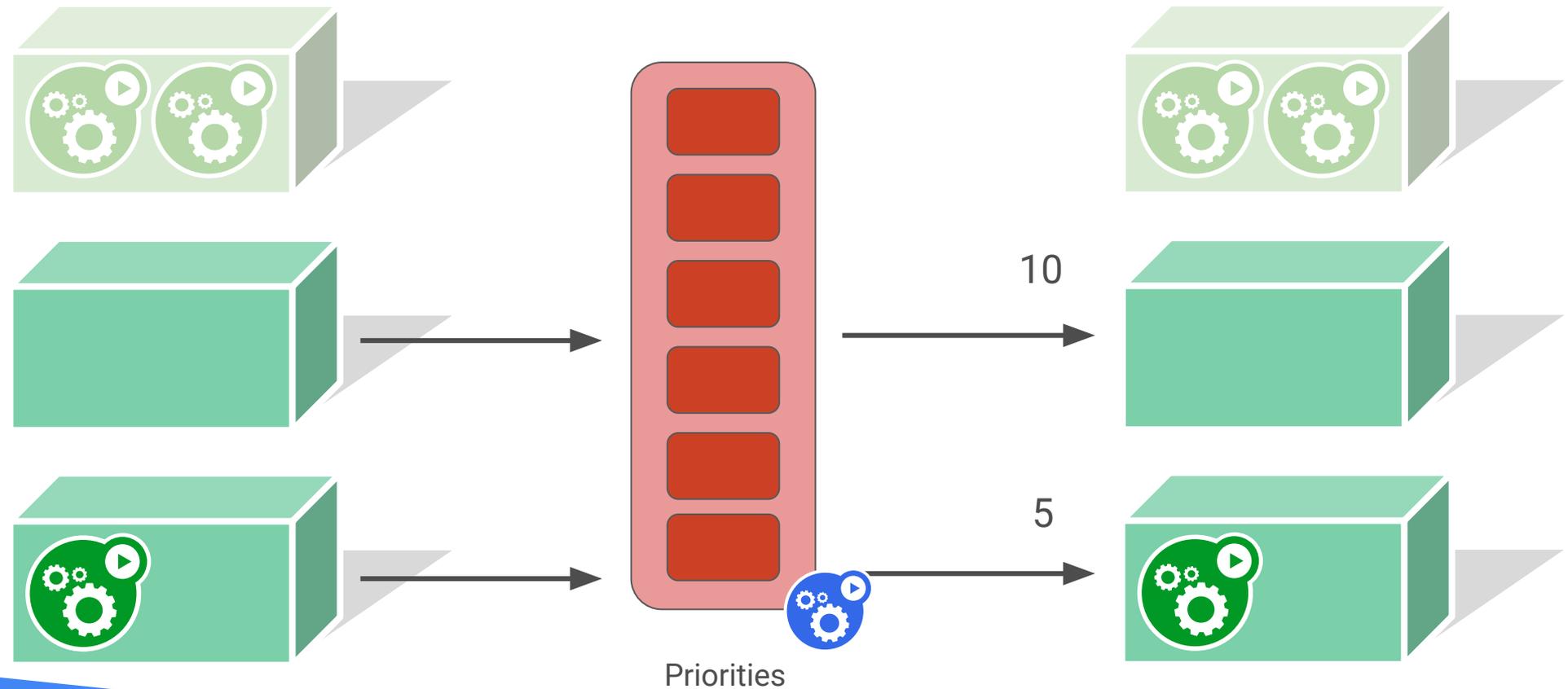
Scheduling



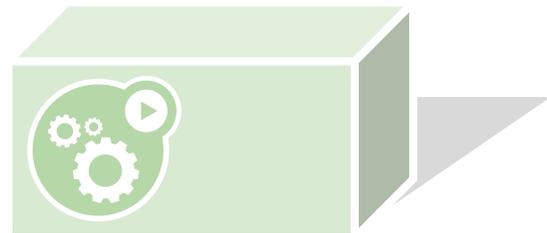
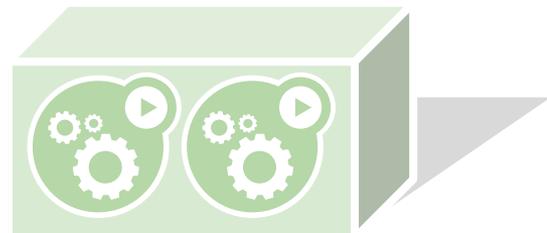
Scheduling



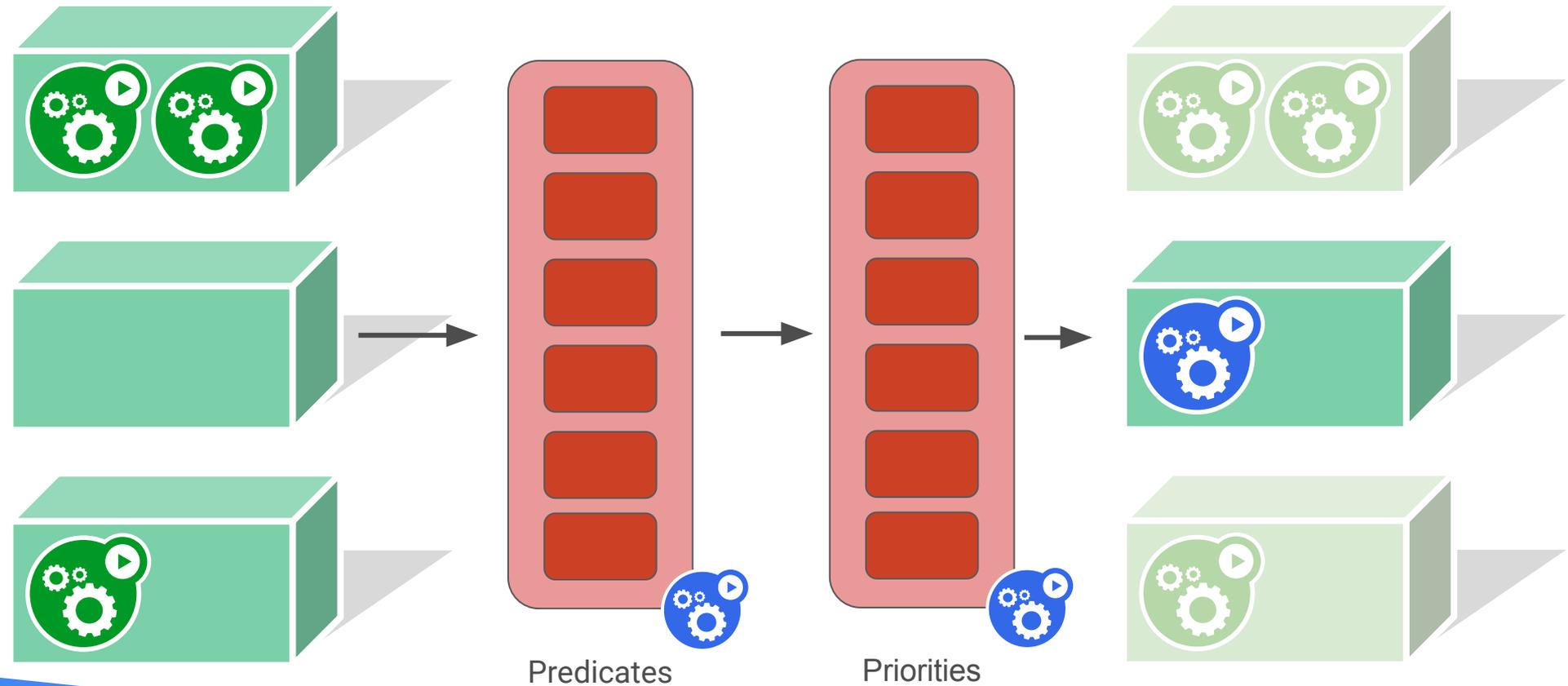
Scheduling



Scheduling



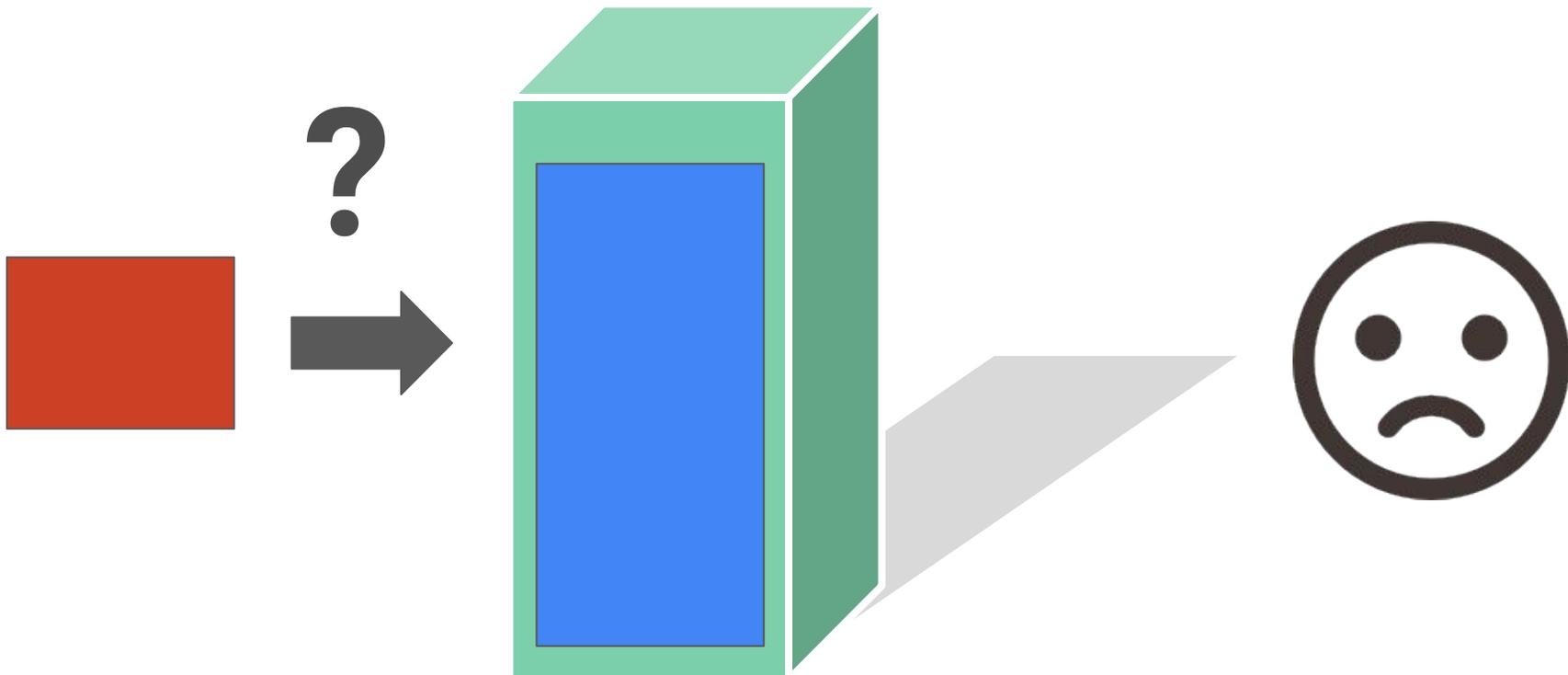
Scheduling



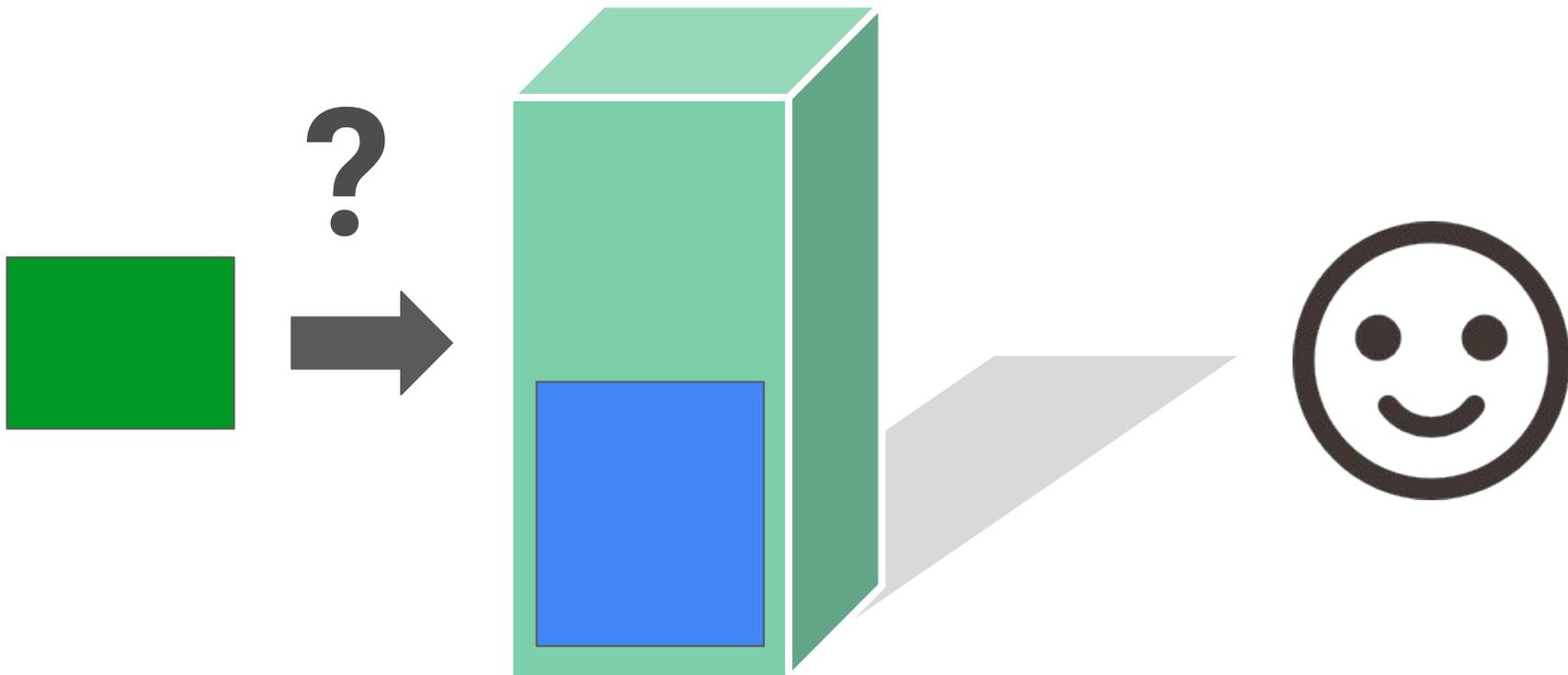


Predicates deep dive

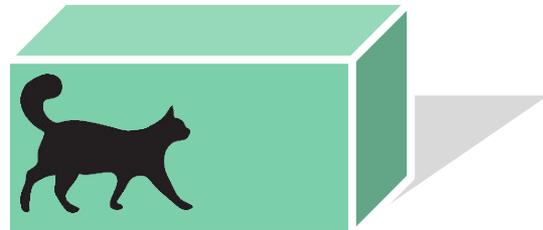
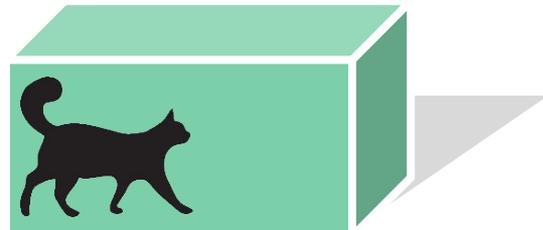
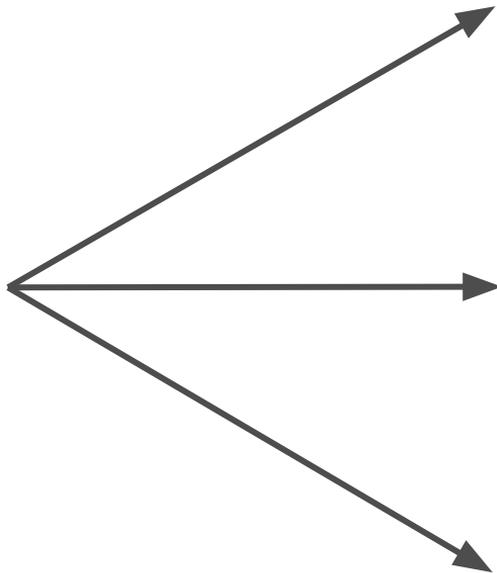
Prevent overcommit



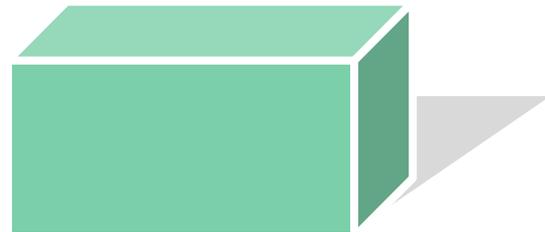
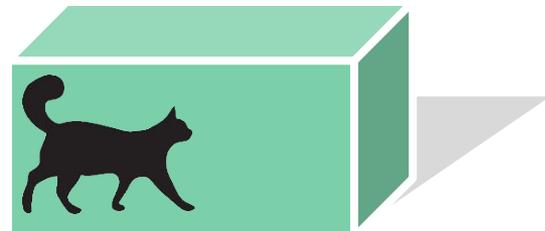
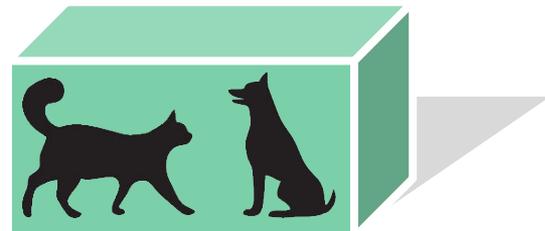
Prevent overcommit



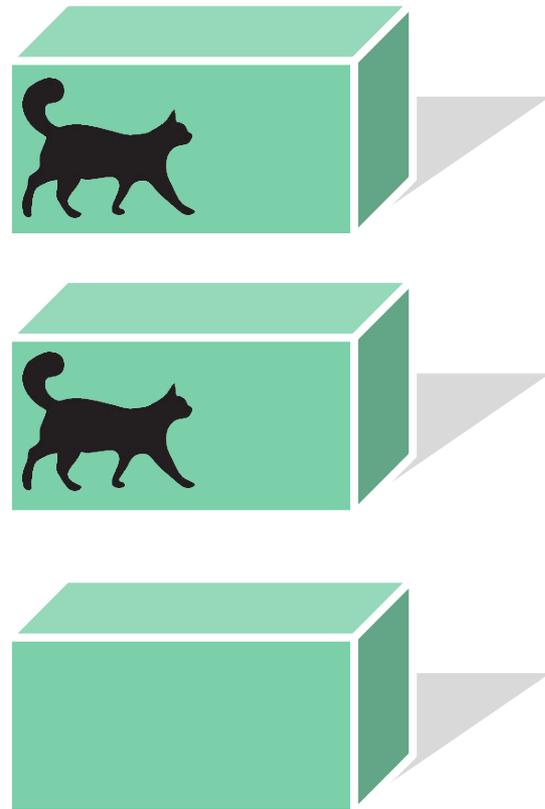
Prevent co-scheduling (pod anti-affinity)



Prevent co-scheduling (pod anti-affinity)

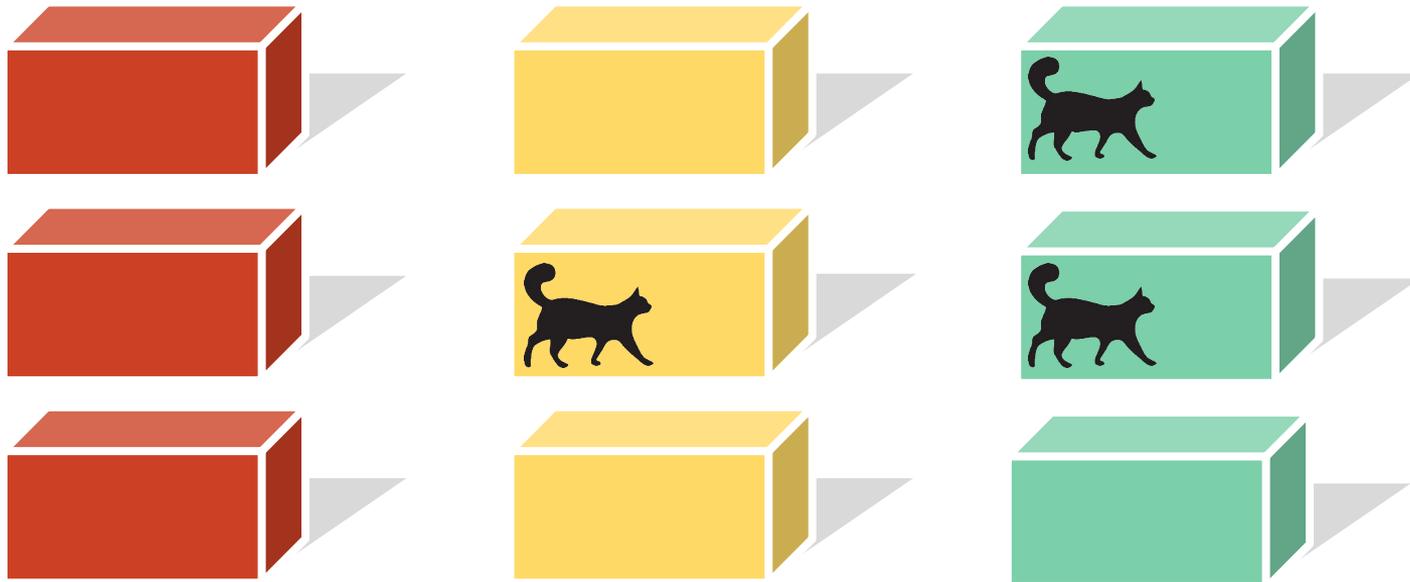


Prevent co-scheduling (pod anti-affinity)

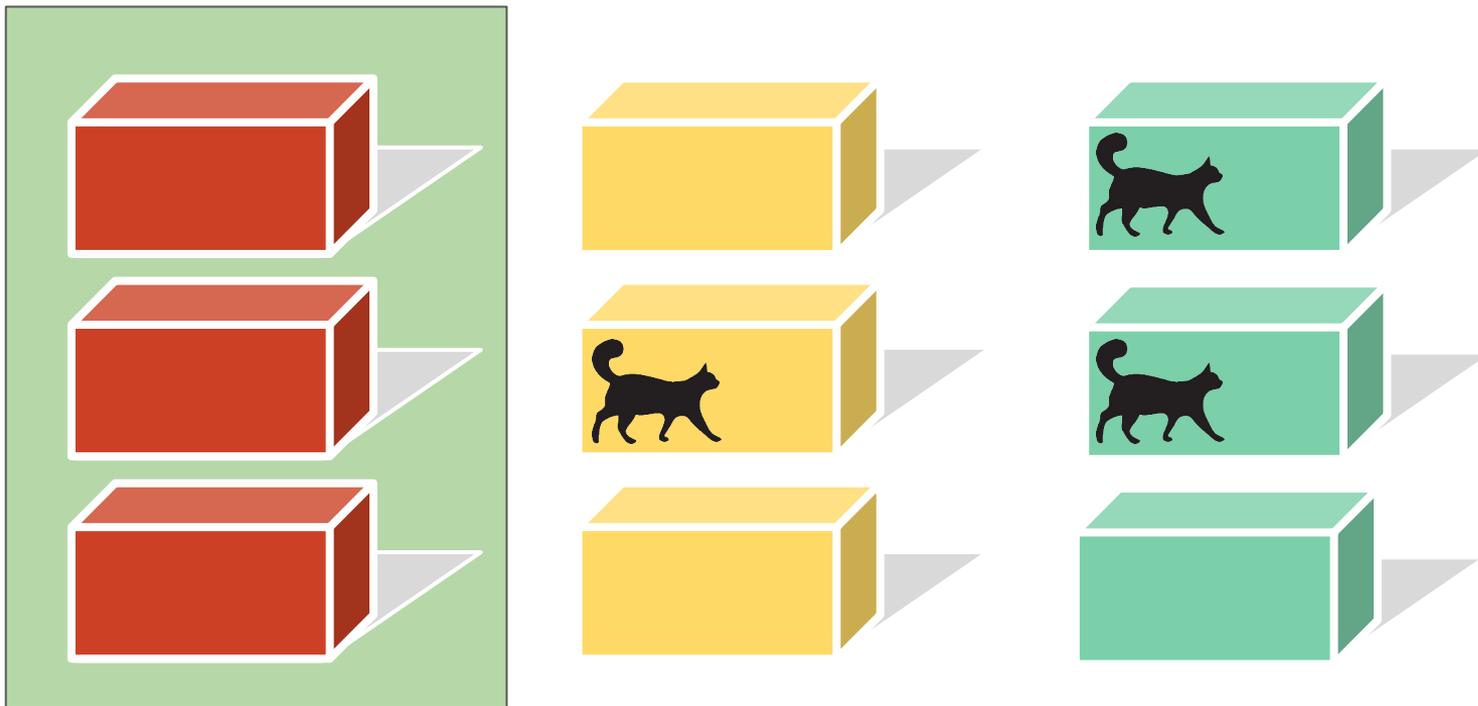


```
PodAntiAffinity: {  
  TopologyKey: "hostname",  
  LabelSelector: "type:cat"  
}
```

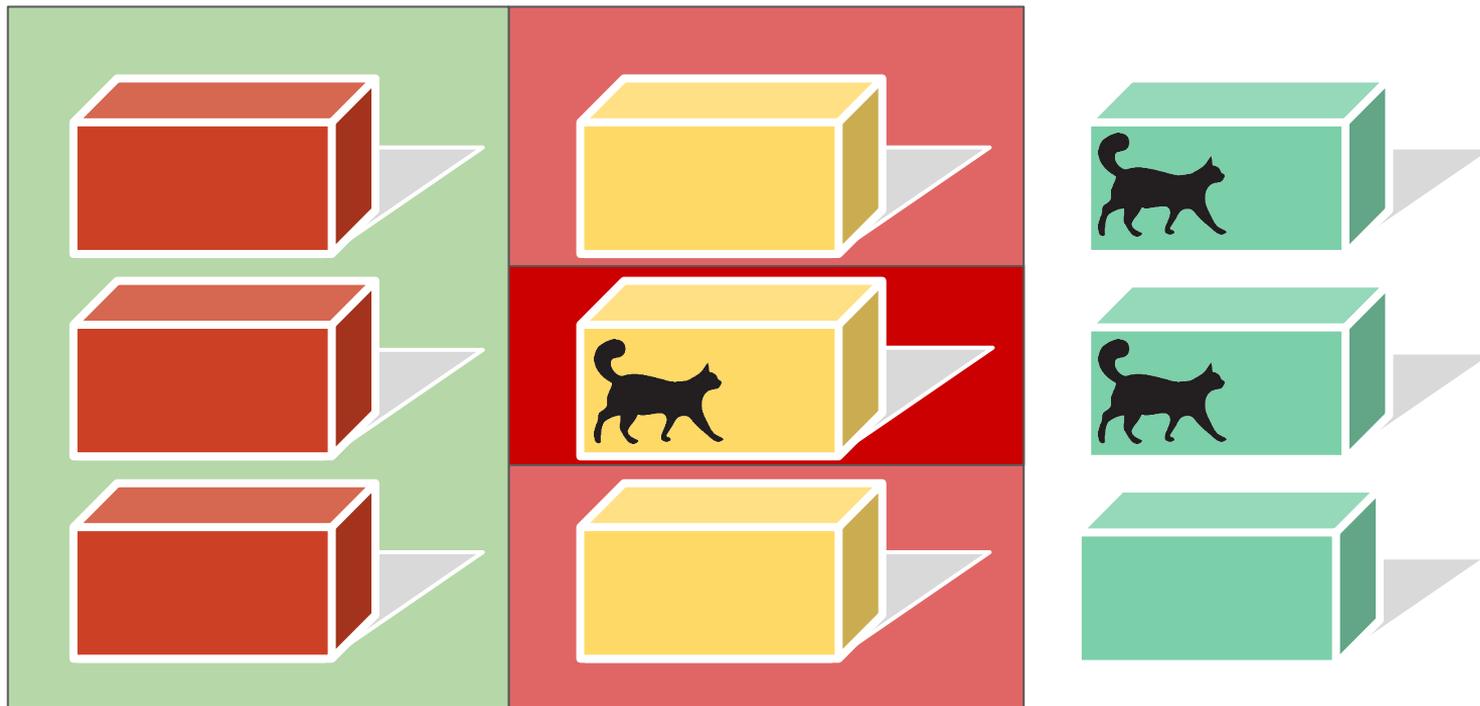
Prevent co-scheduling (pod anti-affinity)



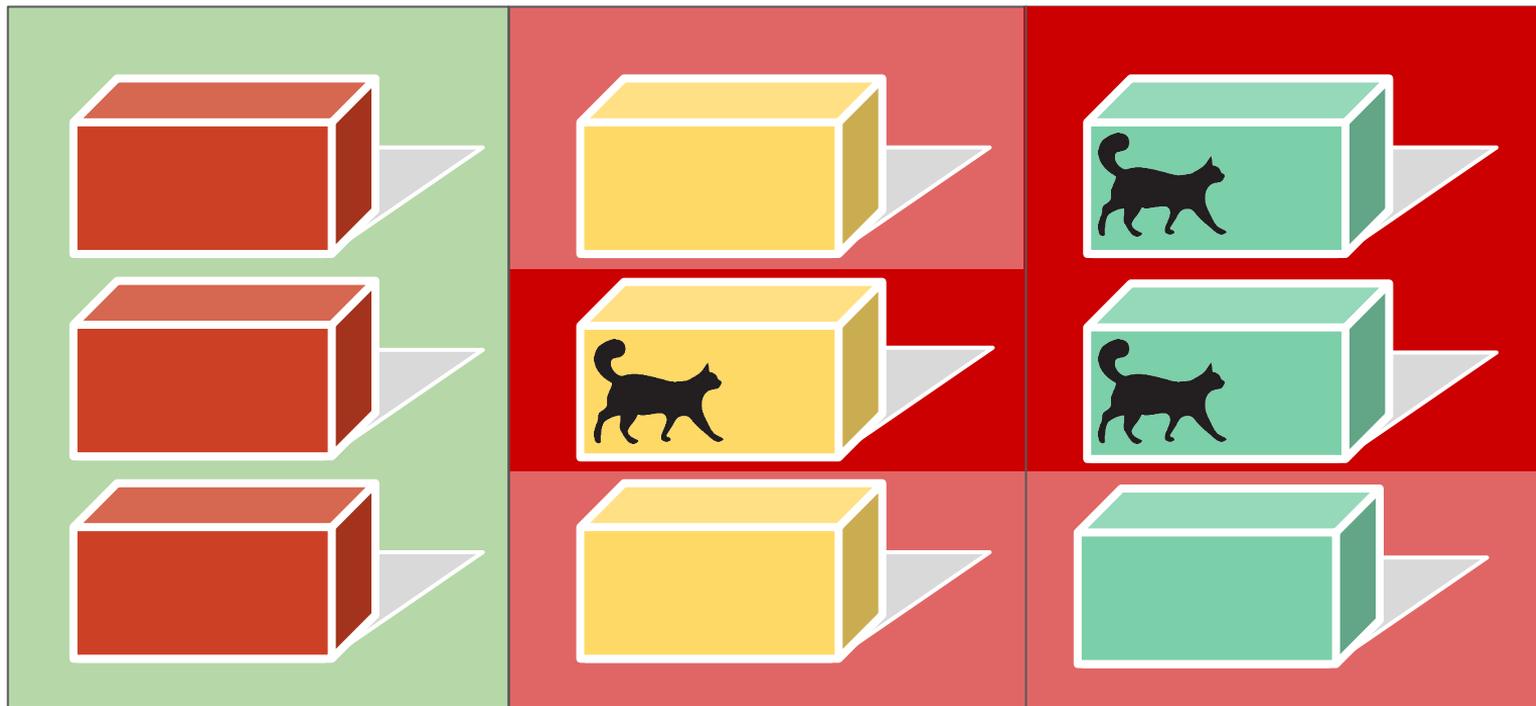
Prevent co-scheduling (pod anti-affinity)



Prevent co-scheduling (pod anti-affinity)



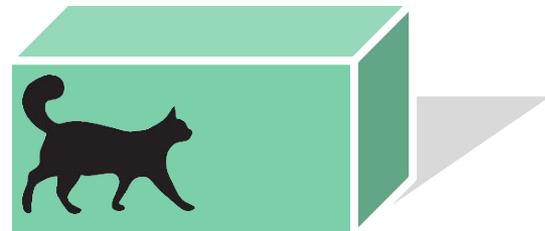
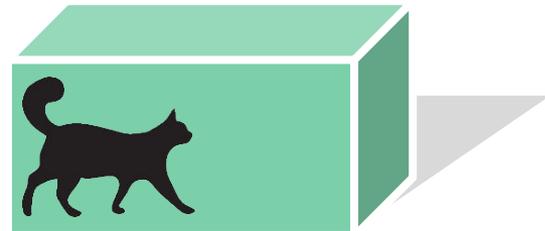
Prevent co-scheduling (pod anti-affinity)



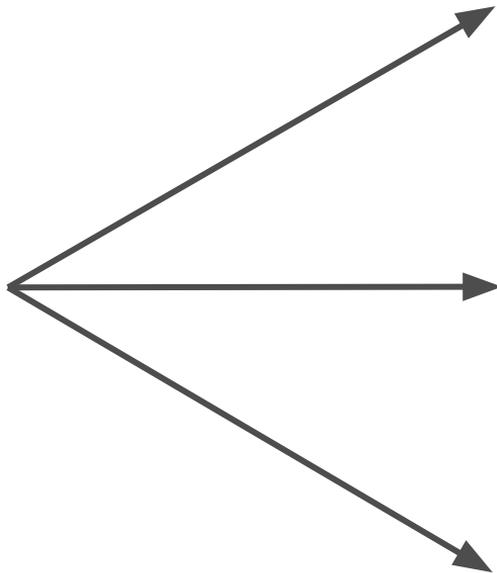
Prevent co-scheduling (pod anti-affinity)



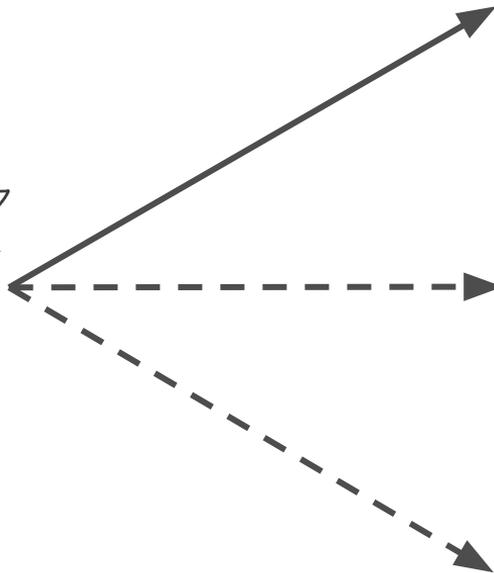
```
PodAntiAffinity: {  
  TopologyKey: "hostname",  
  LabelSelector: "type:cat"  
}
```



Force co-scheduling (pod affinity)

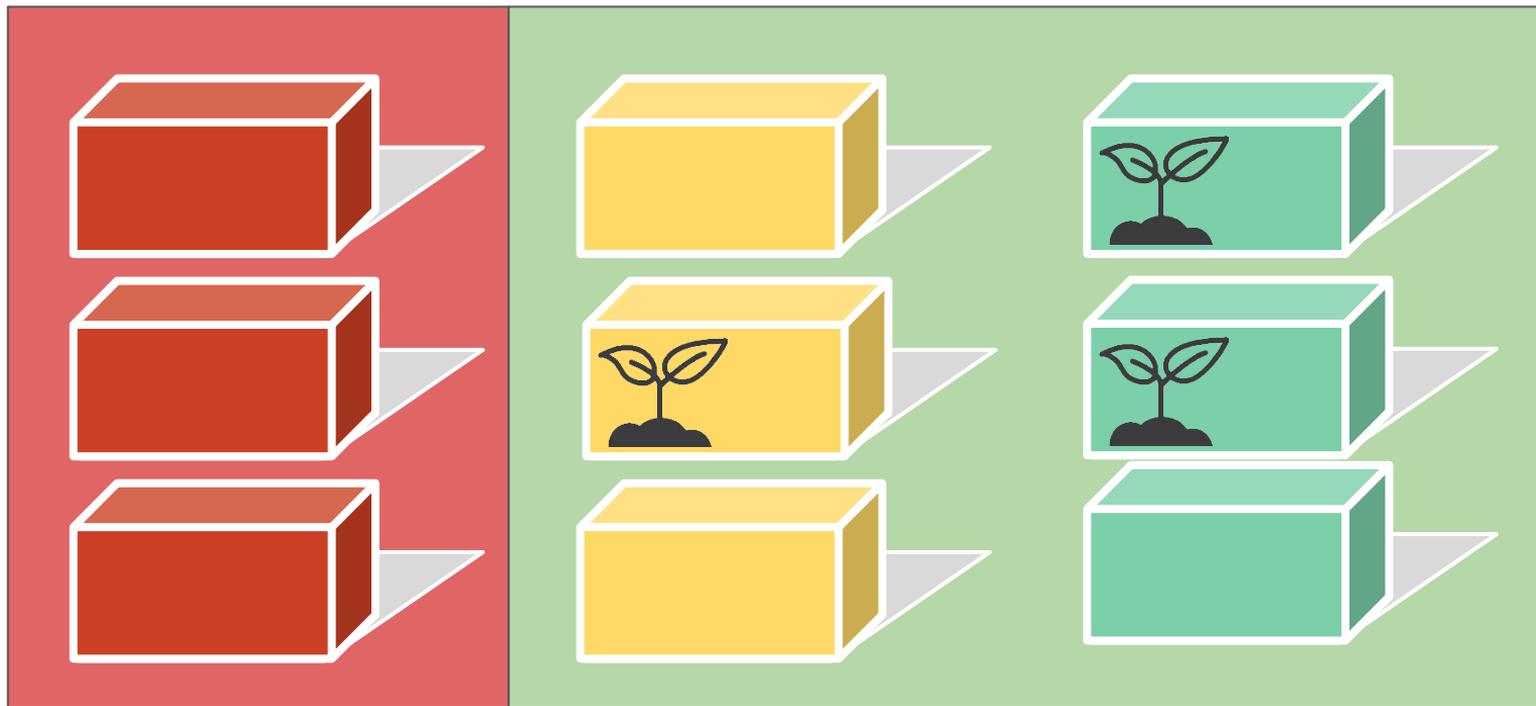


Force co-scheduling (pod affinity)

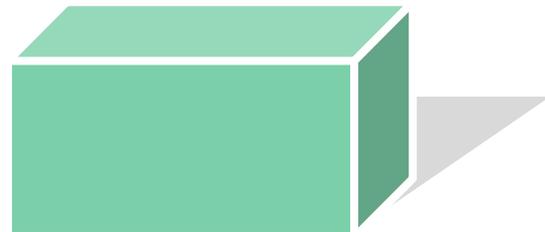


```
PodAffinity: {  
  TopologyKey: "hostname",  
  LabelSelector: "type:sapling"  
}
```

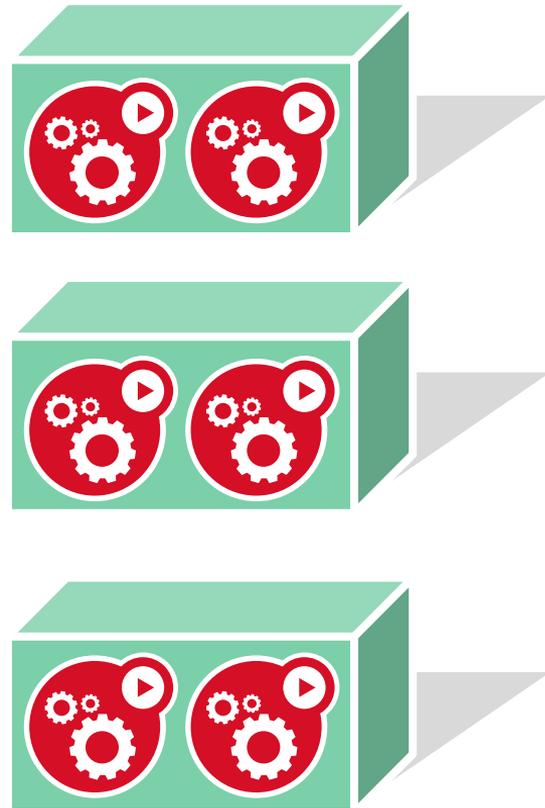
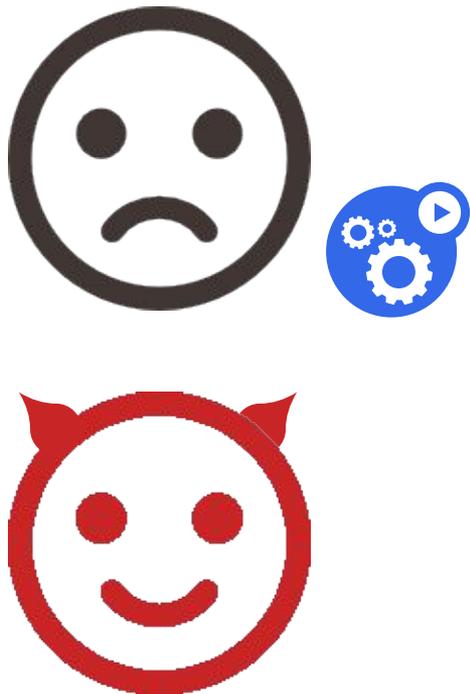
Force co-scheduling (pod affinity)



Dedicated machines (Taints)

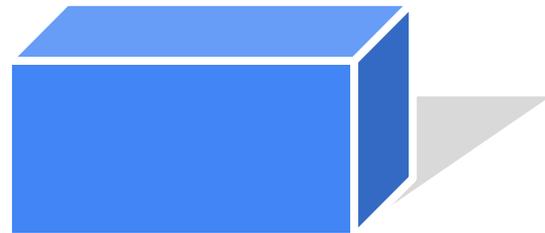


Dedicated machines (Taints)



Dedicated machines (Taints)

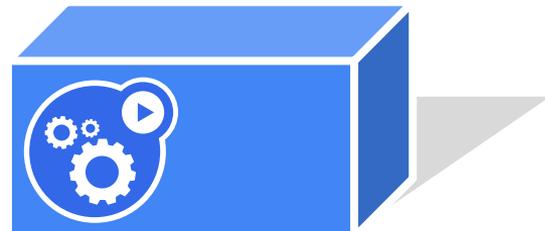
```
Taint: {  
  TaintEffect: "NoSchedule",  
  Key: "color",  
  Value: "blue"  
}
```



```
Toleration: {  
  Key: "color",  
  Value: "blue",  
  Operator: "Equal",  
  TaintEffect: "NoSchedule"  
}
```

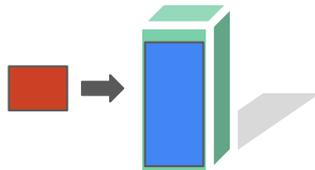


Dedicated machines (Taints)



Predicate summary

- Prevent overcommit



- Prevent co-scheduling



- Force co-scheduling



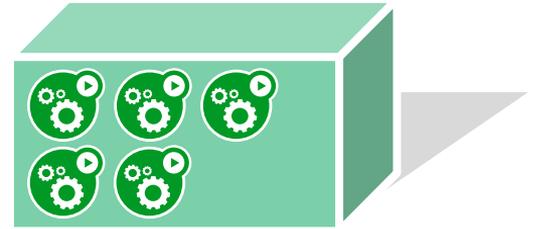
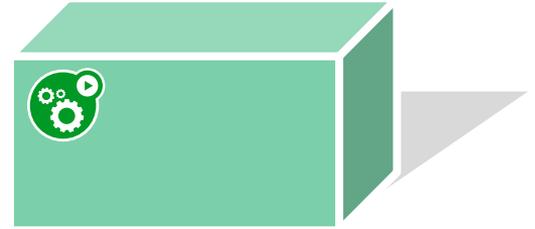
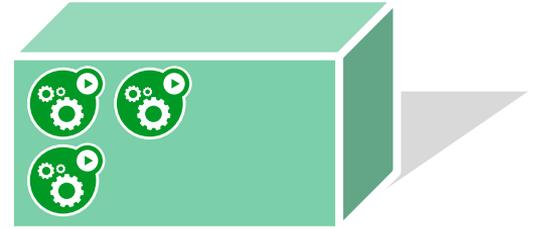
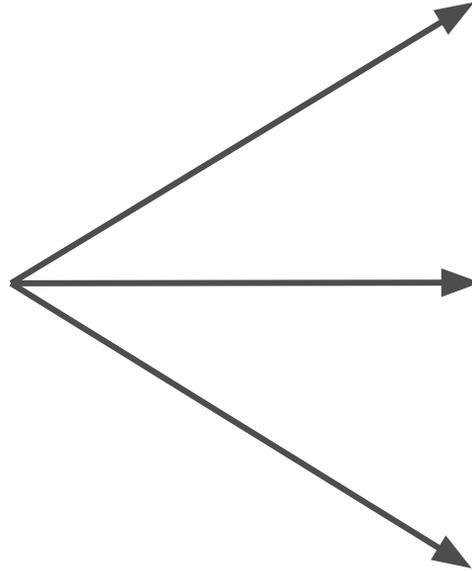
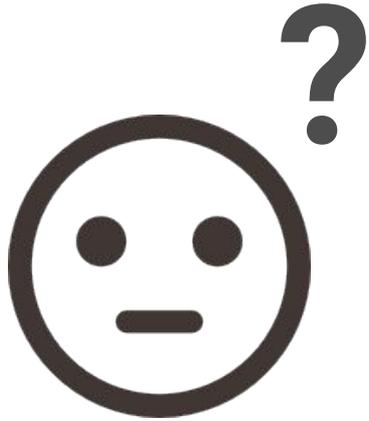
- Dedicated Nodes



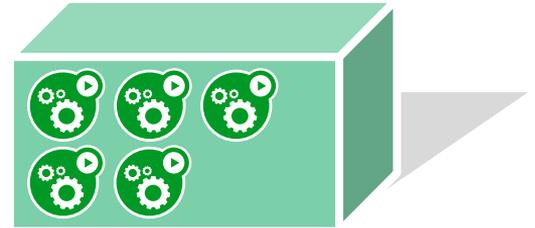
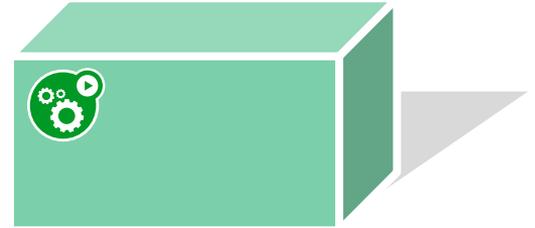
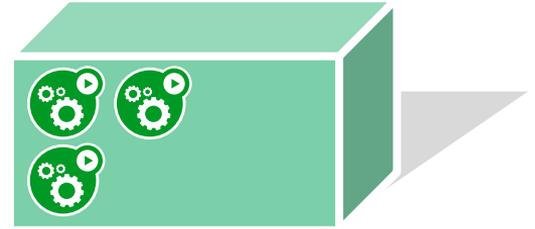
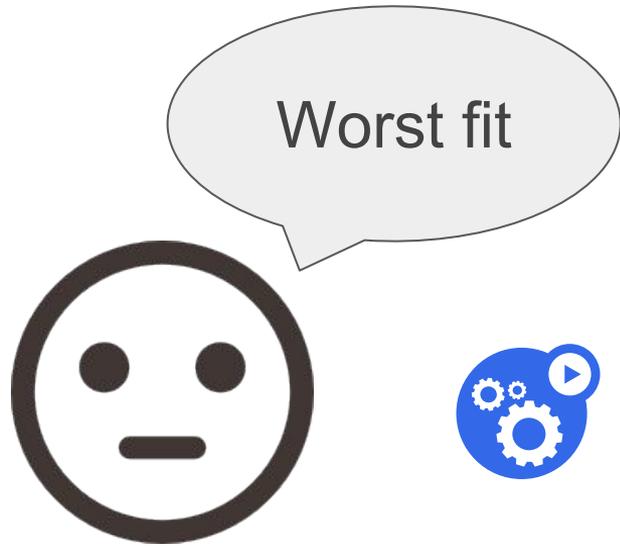
Priorities deep dive



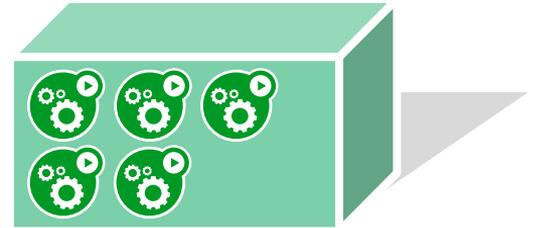
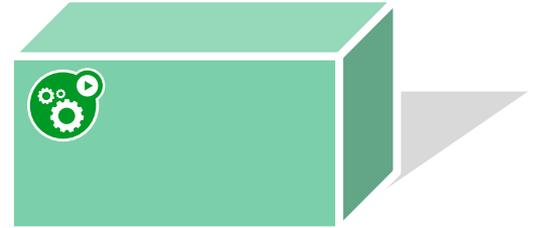
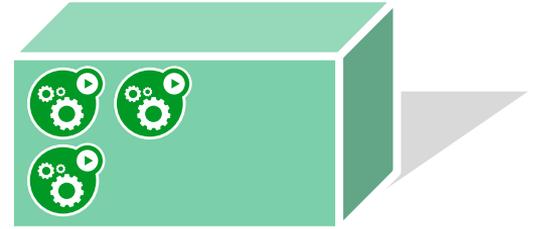
Best fit vs worst fit



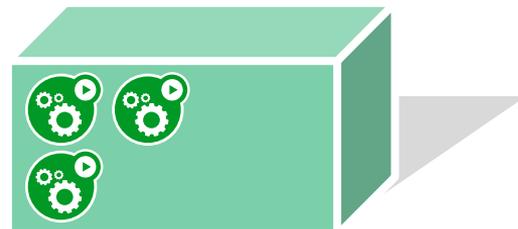
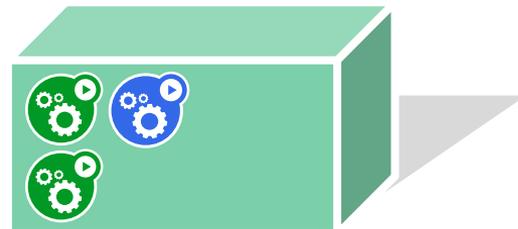
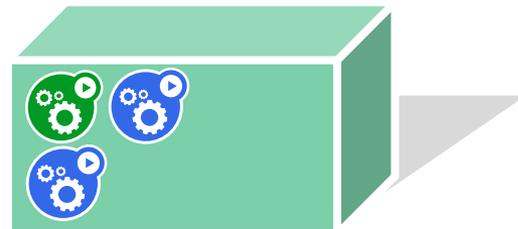
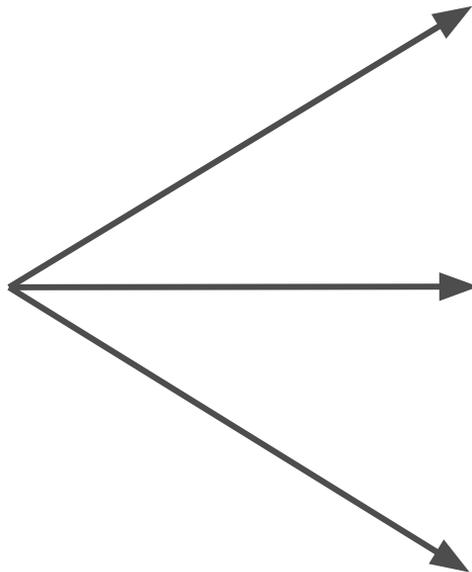
Best fit vs worst fit



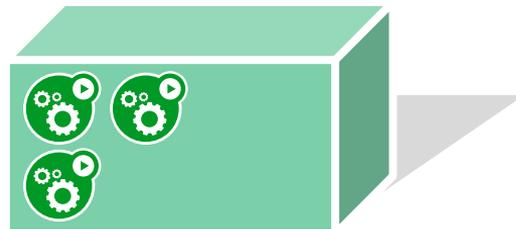
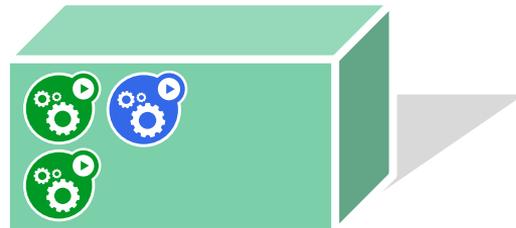
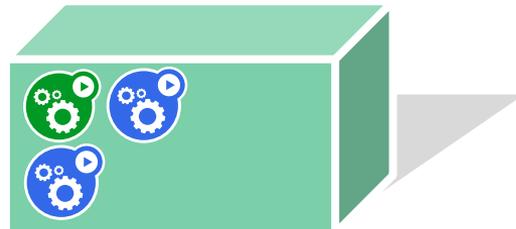
Best fit vs worst fit



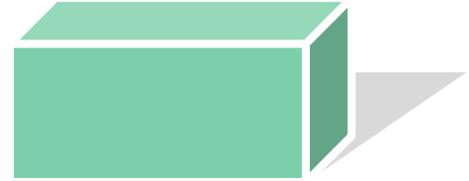
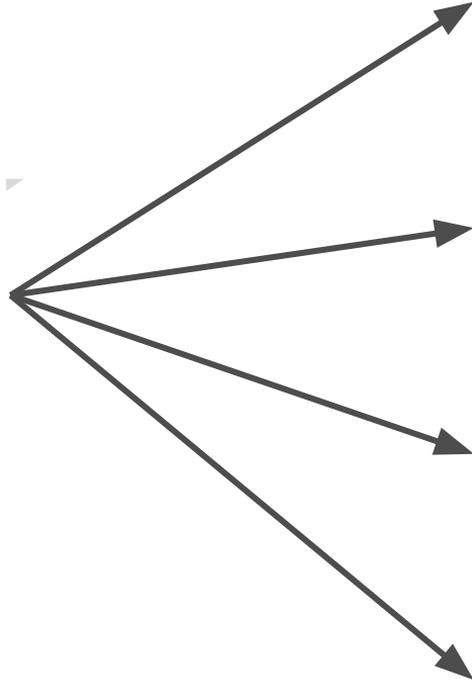
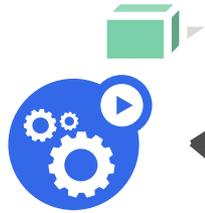
Selector Spreading



Selector Spreading



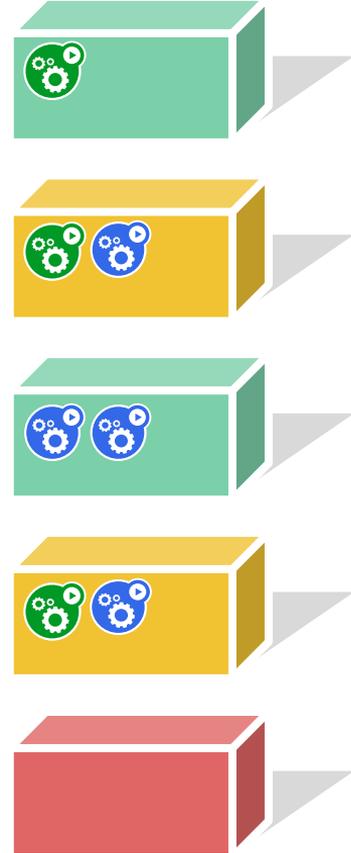
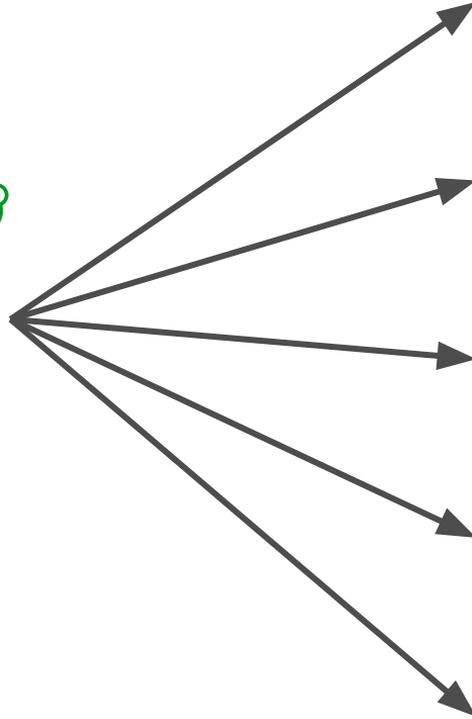
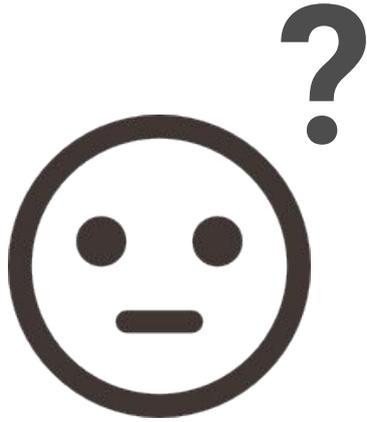
Node Affinity



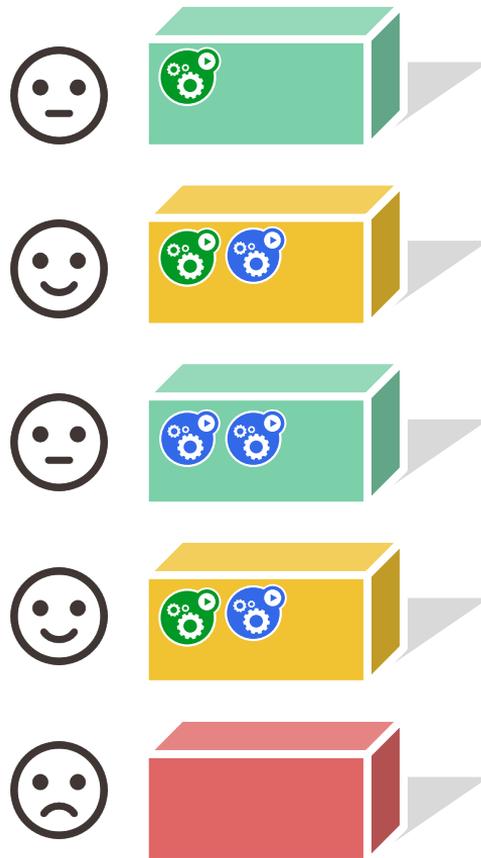
Node Affinity



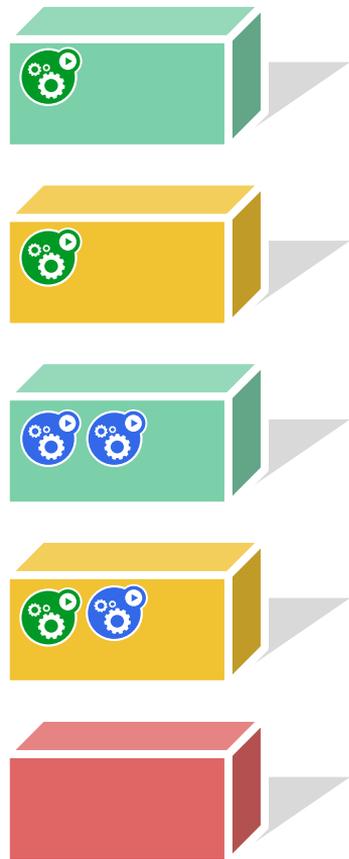
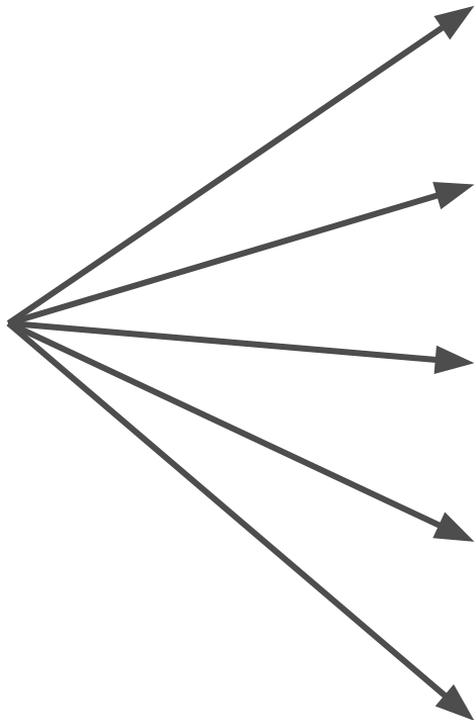
Pod Affinity



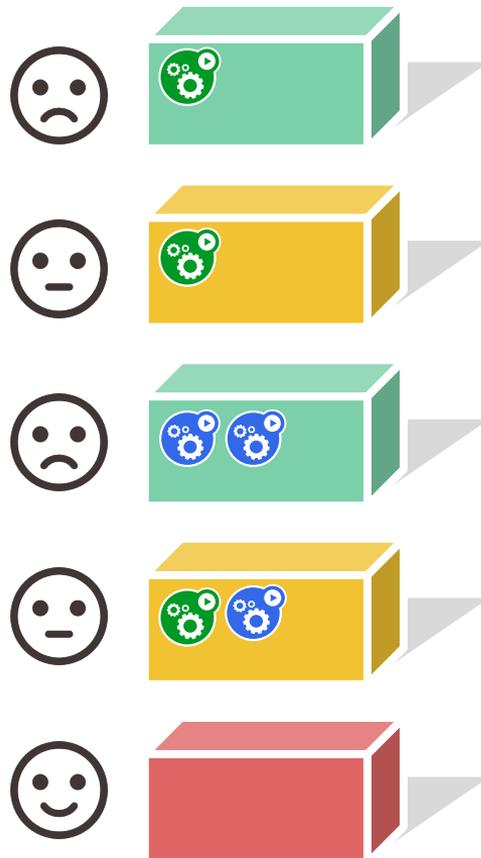
Pod Affinity



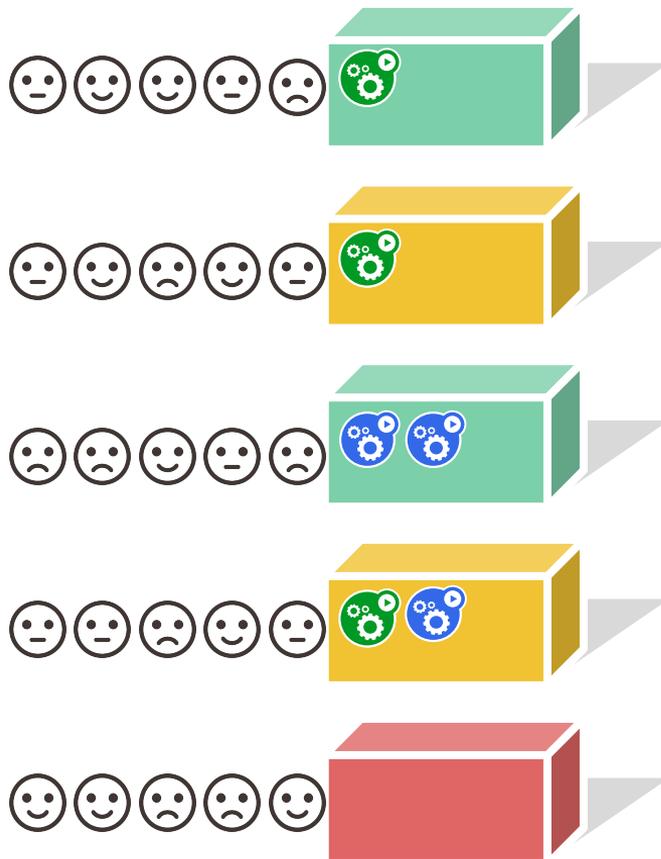
Pod Anti-Affinity



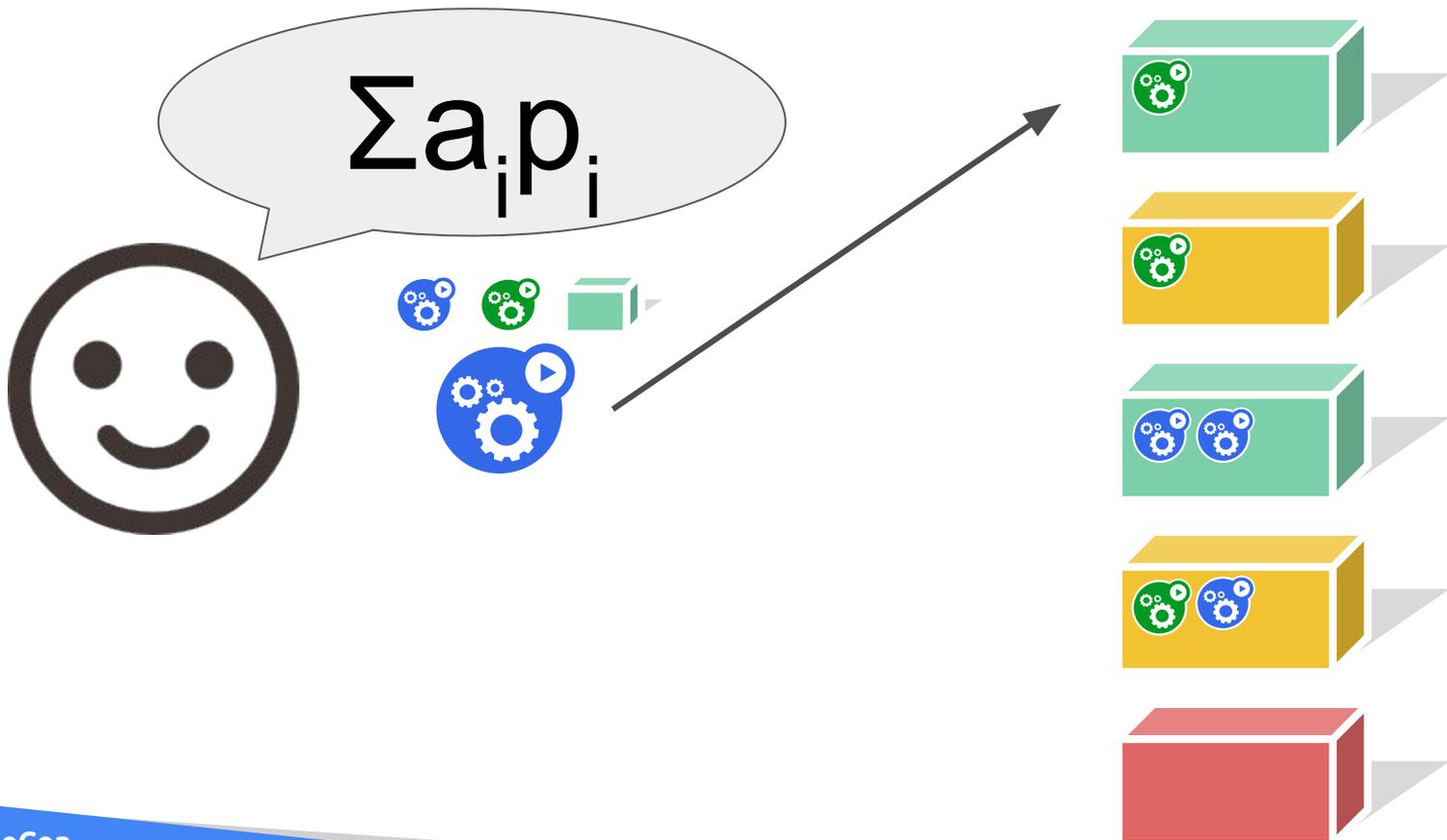
Pod Anti-Affinity



How to combine scores?



Linear combination



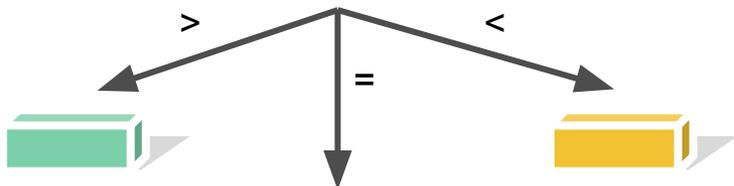
Dangers of depending on priorities



Decision tree

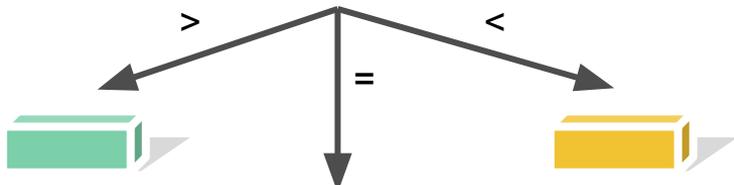


Node Affinity score ? Node Affinity score



$1000 * (\text{Node Affinity score}) +$

Spreading score ? Spreading score



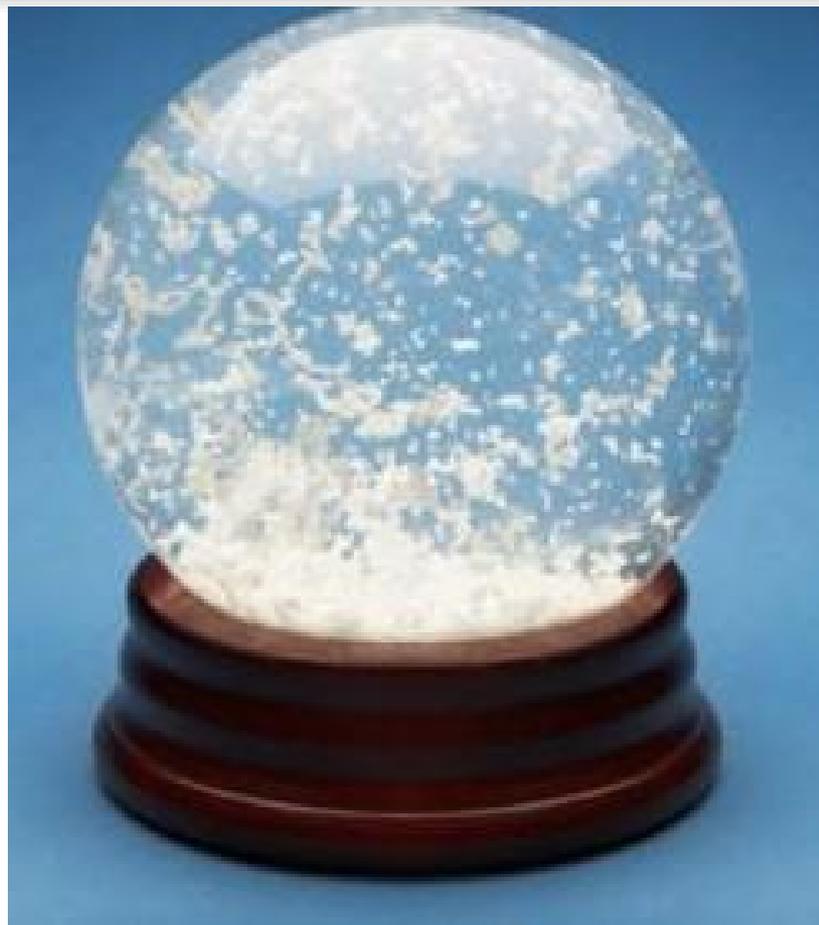
$100 * (\text{Spreading score}) +$

...

...

Priorities are powerful

... but require deep understanding



Feature state as of 1.6

- [beta in 1.6] Multi-scheduler support
- [beta in 1.6] Node affinity
- [beta in 1.6] Pod affinity/anti-affinity
- [beta in 1.6] Taints & tolerations

Future plans

- Priorities & preemptions
- Rescheduler
- Resource estimation

SIG Scheduling

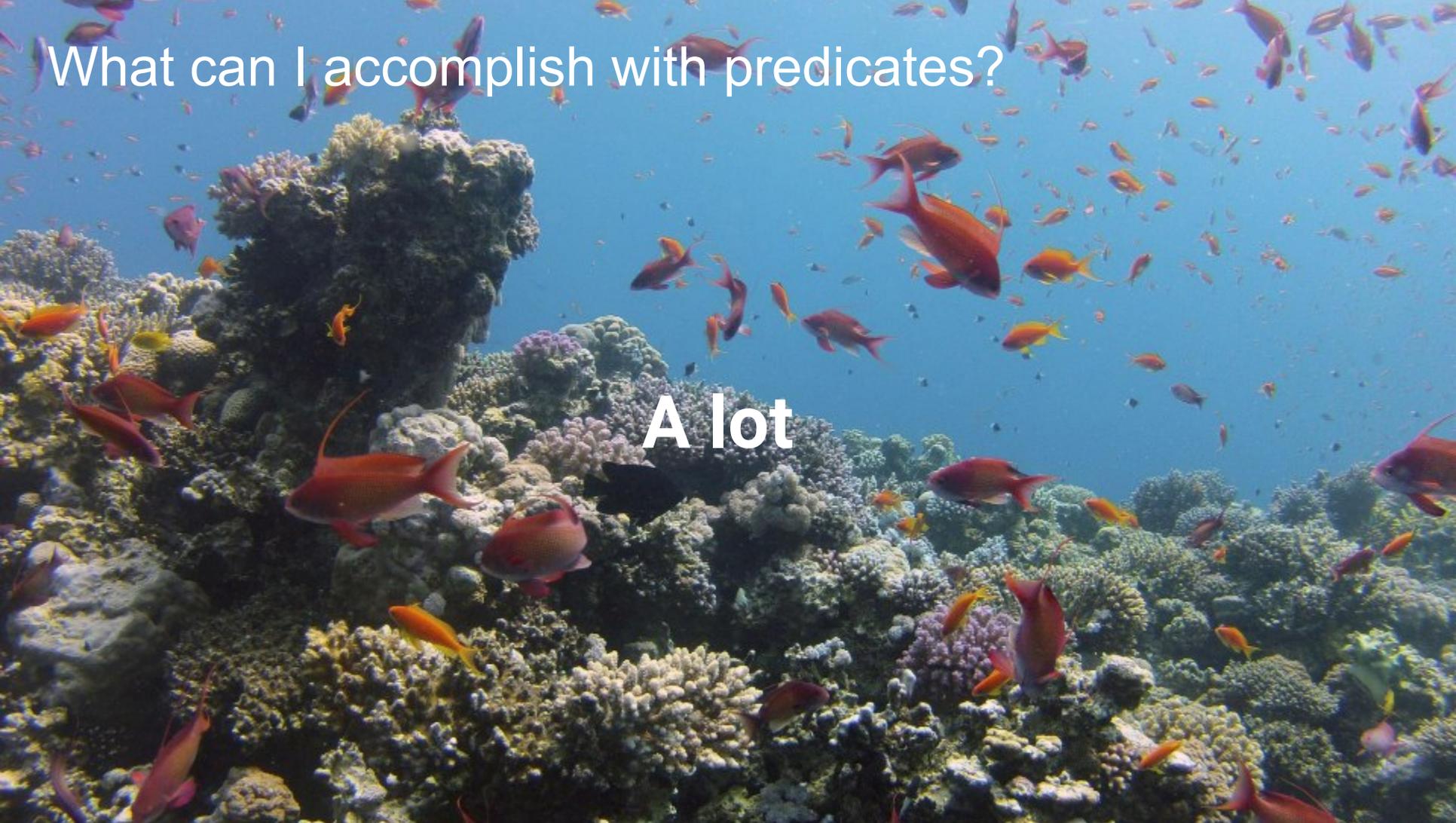
- Meeting every other Monday 2pm PST
- #sig-scheduling Slack channel
- kubernetes-sig-scheduling@ mailing group

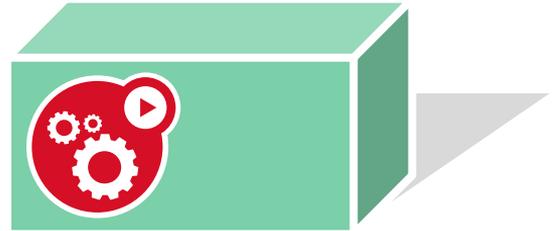
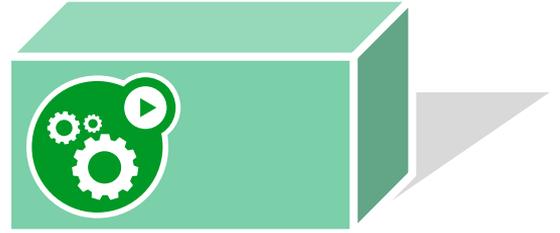


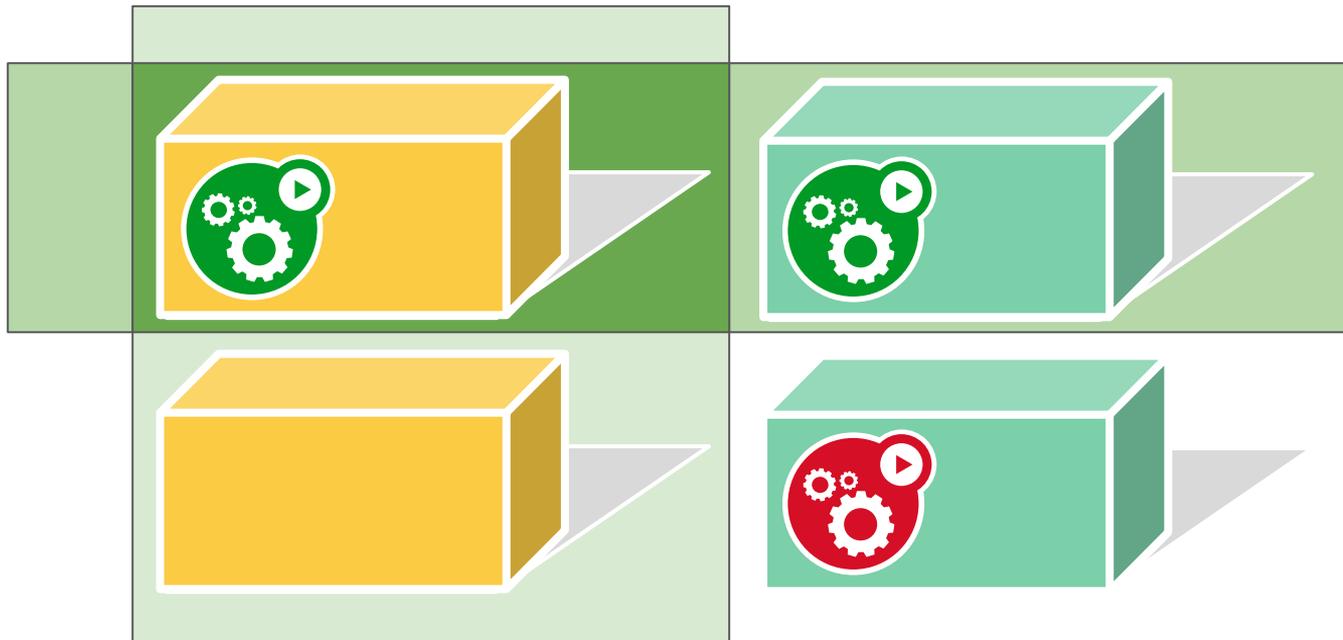
Backup slides

What can I accomplish with predicates?

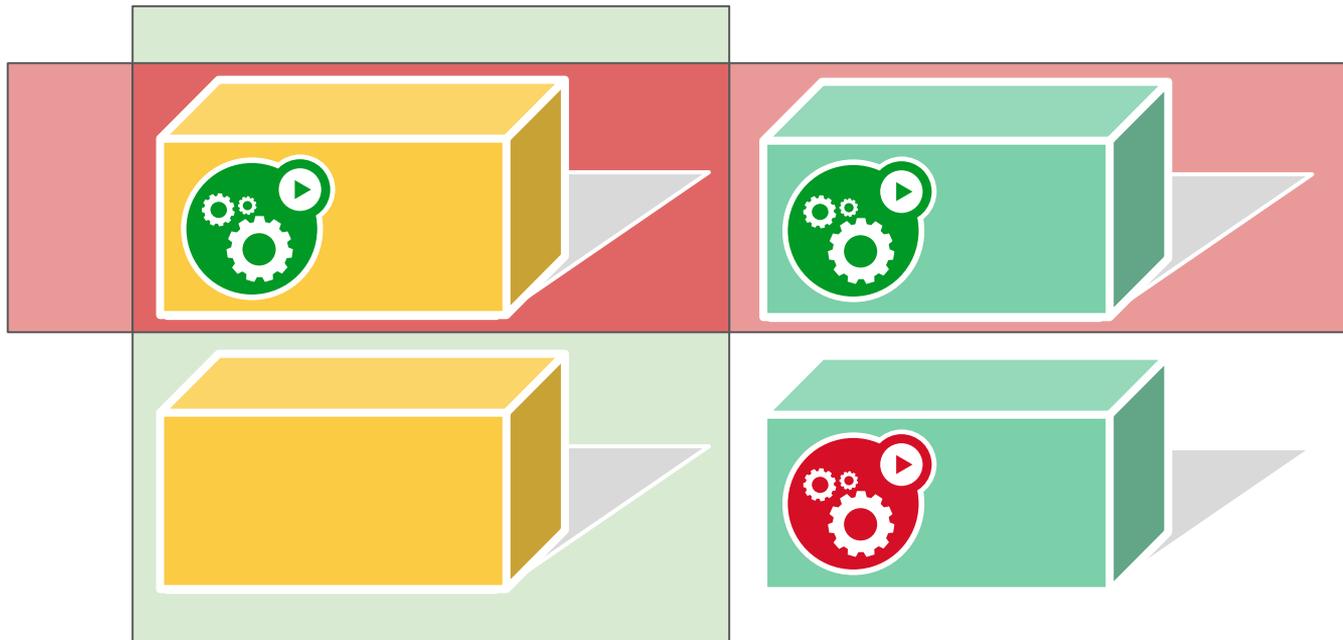
A lot

A vibrant underwater scene featuring a diverse coral reef. The reef is covered in various types of coral, including branching and table corals, in shades of green, yellow, and purple. Numerous colorful fish, including many orange and red ones, are swimming throughout the water. The background is a clear, bright blue. A black arrow points from the text 'A lot' towards the coral reef.

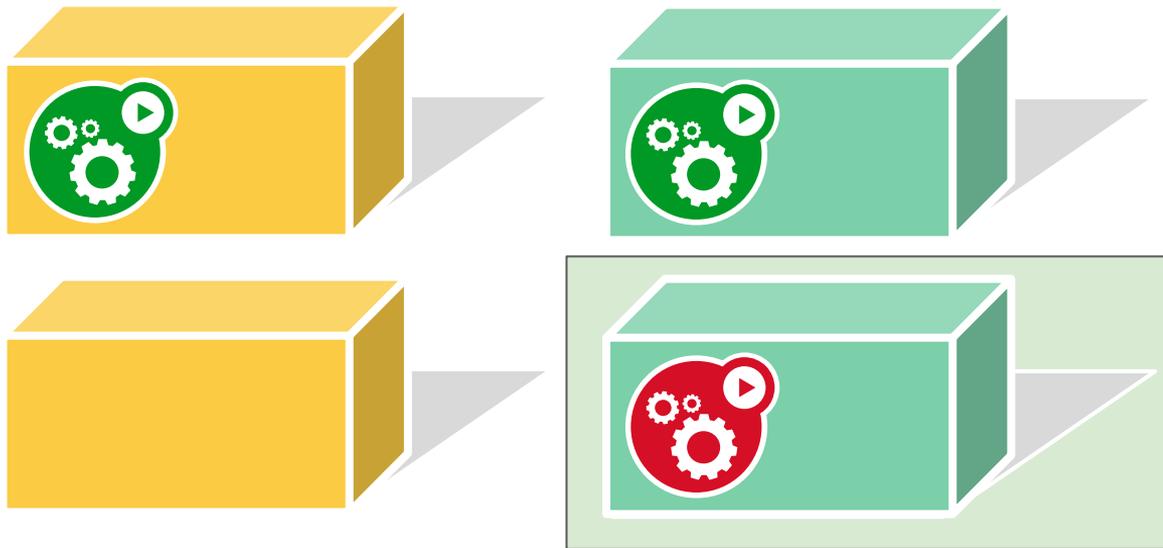




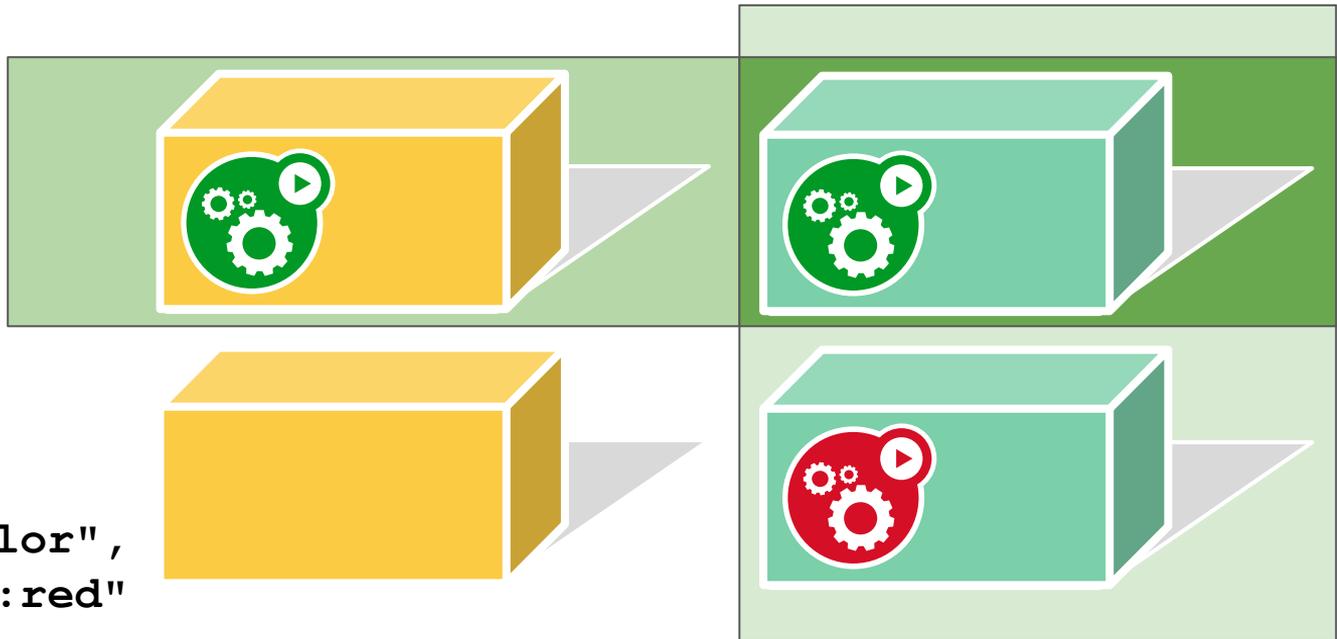
```
{  
  PodAffinity: {  
    TopologyKey: "hostname",  
    Selector: "color:green"  
  },  
  NodeSelector: "color:yellow"  
}
```



```
{  
  PodAntiAffinity: {  
    TopologyKey: "hostname",  
    Selector: "color:green"  
  },  
  NodeSelector: "color:yellow"  
}
```



```
{  
  PodAffinity: {  
    TopologyKey: "hostname",  
    Selector: "color:red"  
  },  
}
```



```
{  
  PodAffinity: {  
    TopologyKey: "color",  
    Selector: "color:red"  
  },  
  PodAffinity: {  
    TopologyKey: "hostname",  
    Selector: "color:green"  
  },  
}
```