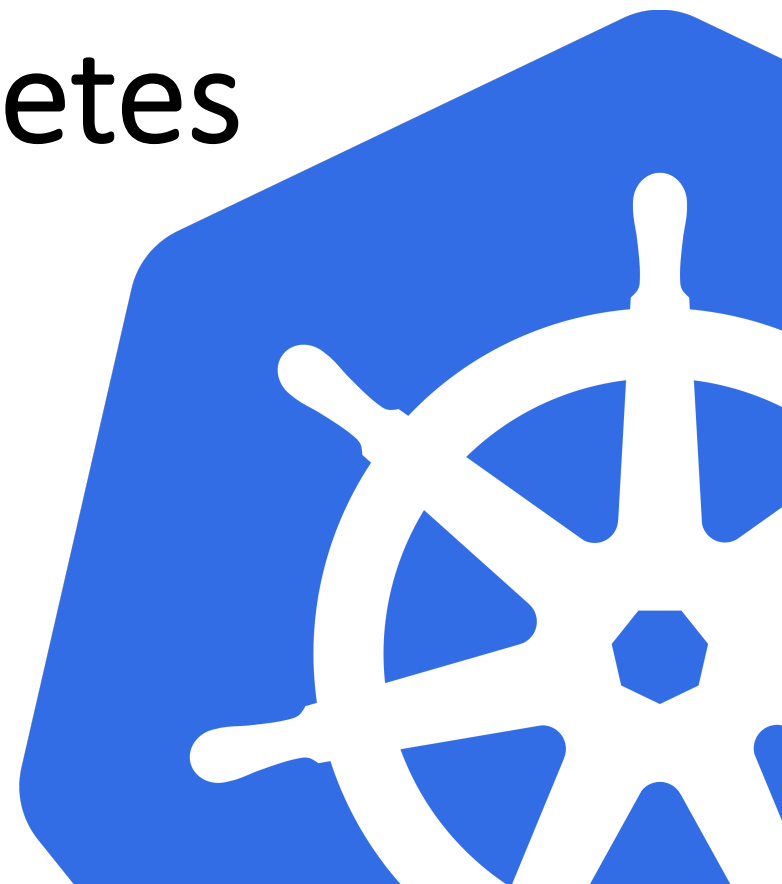


# Повышаем отказоустойчивость и снижаем расходы в Kubernetes

Павел Селиванов  
VK Cloud Solutions





# Павел Селиванов

Developer Advocate VK Cloud Solutions

- Работаю с Kubernetes с 2016 года
- Первый кластер версии 1.4
- Помогаю инженерам осваивать современные подходы работы с инфраструктурой



# План

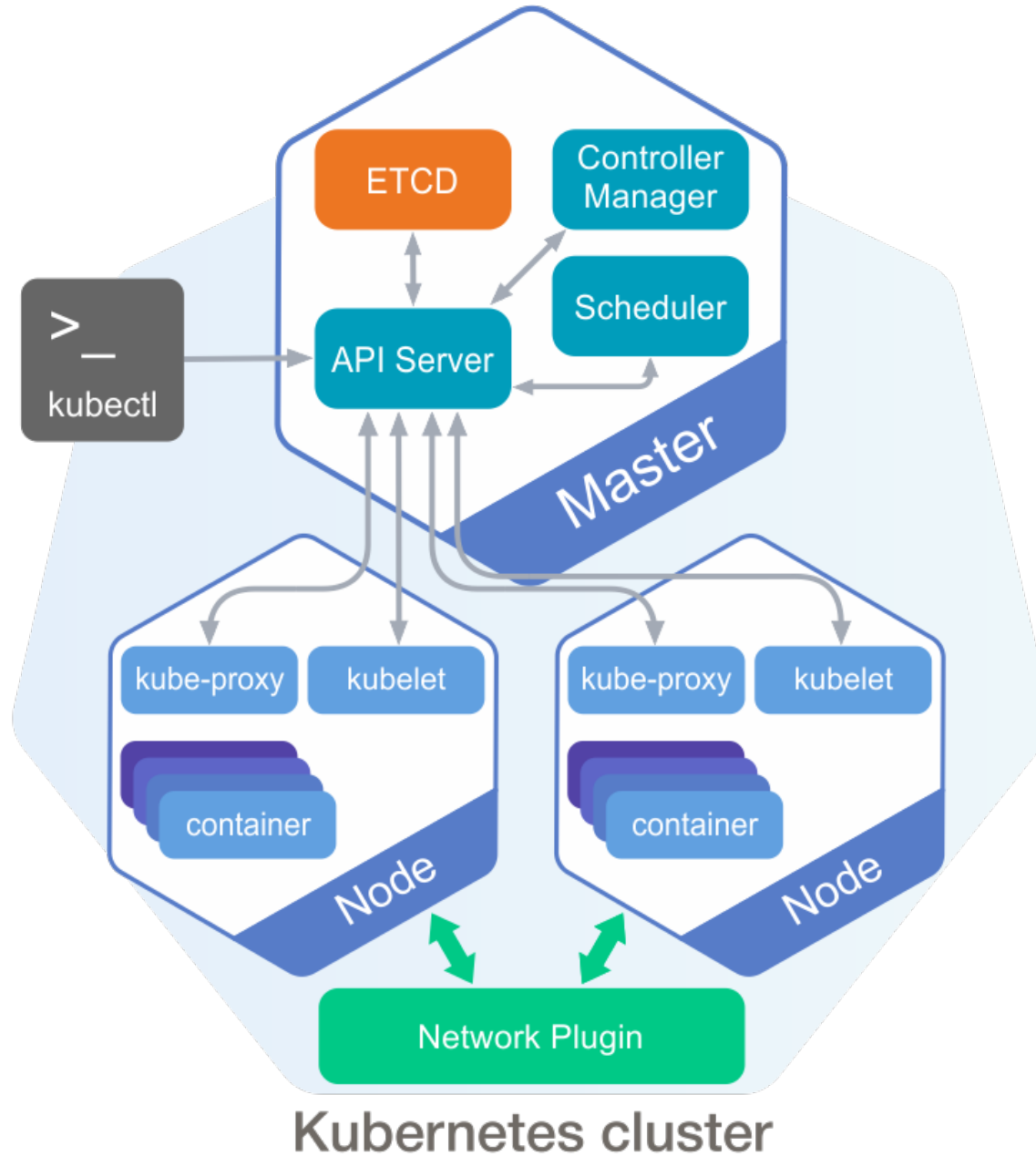
- Отказоустойчивость современных IT систем
- Kubernetes и зоны доступности
- Уменьшение затрат на кластер

# Избыточность

- Компоненты электронных схем **дают сбои**
- Дело не в производителе
- «**Надежные** машины из **ненадежных элементов**»?



# Kubernetes





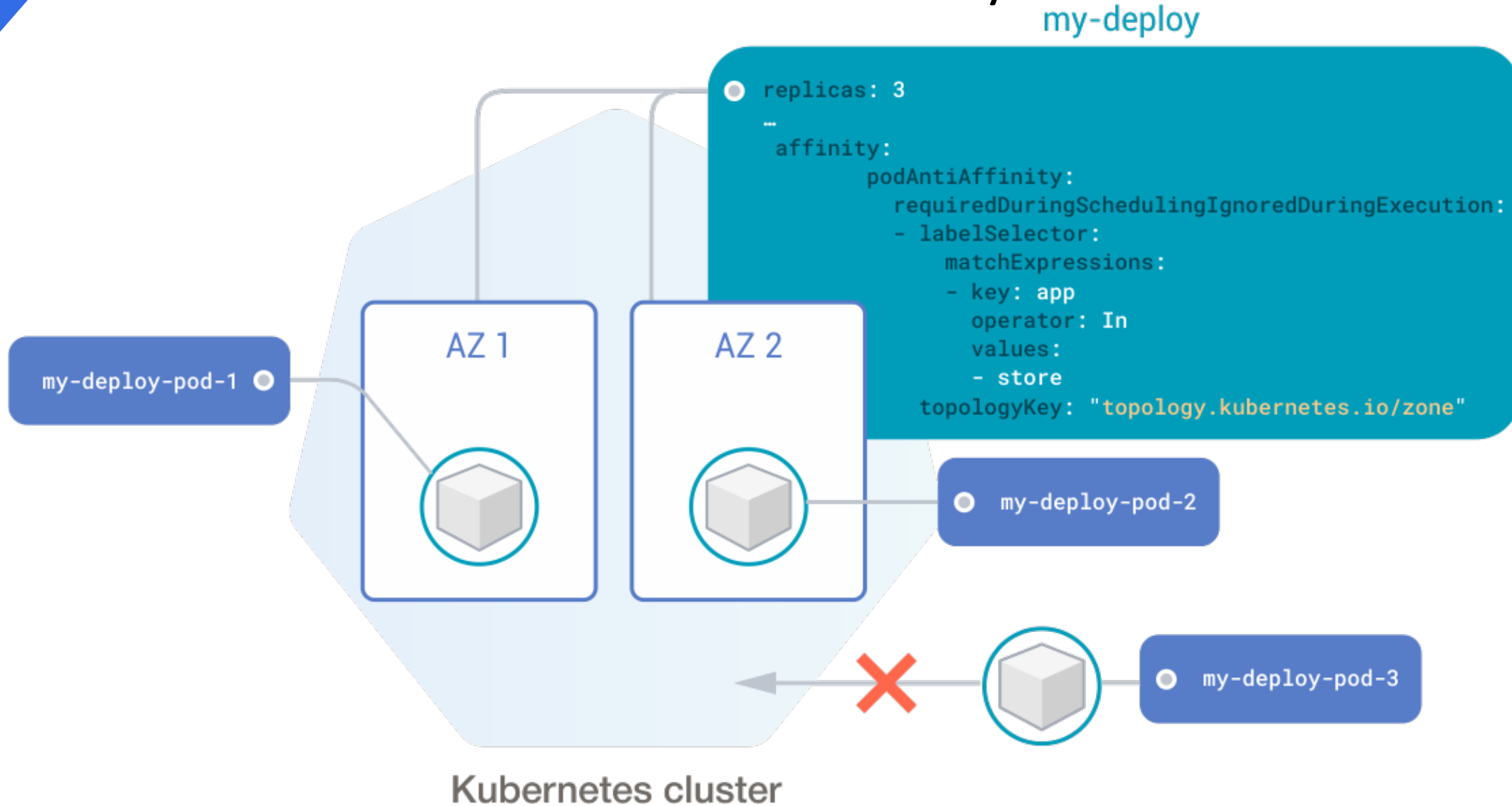
Облачная инфраструктура  
не обладает высоким уровнем надежности  
в классическом понимании

# Тогда зачем нам облака?

- **Один** сбой на **3 часа хуже** чем **3** сбоя по **5 минут**
- Инструменты для управления инфраструктурой
- Инструменты для создания отказоустойчивой инфраструктуры



# Kubernetes Pod AntiAffinity





# Pod Topology Spread Constraints

<https://kubernetes.io/docs/concepts/workloads/pods/pod-topology-spread-constraints/>

```
kind: Pod
apiVersion: v1
metadata:
  name: myapp
  labels:
    app: myapp
spec:
  topologySpreadConstraints:
  - maxSkew: 1
    topologyKey: topology.kubernetes.io/zone
    whenUnsatisfiable: DoNotSchedule
    labelSelector:
      matchLabels:
        app: myapp
```

Разница в количестве  
подов по зонам

# Pod Topology Spread Constraints

<https://kubernetes.io/docs/concepts/workloads/pods/pod-topology-spread-constraints/>

```
kind: Pod
apiVersion: v1
metadata:
  name: myapp
  labels:
    app: myapp
spec:
  topologySpreadConstraints:
  - maxSkew: 1
    topologyKey: topology.kubernetes.io/zone
    whenUnsatisfiable: DoNotSchedule
    labelSelector:
      matchLabels:
        app: myapp
```

Разница в количестве  
подов по зонам

Лейбл на нодах  
по которому делать  
распределение

# Pod Topology Spread Constraints

<https://kubernetes.io/docs/concepts/workloads/pods/pod-topology-spread-constraints/>

```
kind: Pod
apiVersion: v1
metadata:
  name: myapp
  labels:
    app: myapp
spec:
  topologySpreadConstraints:
  - maxSkew: 1
    topologyKey: topology.kubernetes.io/zone
    whenUnsatisfiable: DoNotSchedule
    labelSelector:
      matchLabels:
        app: myapp
```

Разница в количестве  
подов по зонам

Лейбл на нодах  
по которому делать  
распределение

По какому лейблу  
распределять поды

# Pod Topology Spread Constraints

<https://kubernetes.io/docs/concepts/workloads/pods/pod-topology-spread-constraints/>

```
kind: Pod
apiVersion: v1
metadata:
  name: myapp
  labels:
    app: myapp
spec:
  topologySpreadConstraints:
  - maxSkew: 1
    topologyKey: topology.kubernetes.io/zone
    whenUnsatisfiable: DoNotSchedule
    labelSelector:
      matchLabels:
        app: myapp
```

Разница в количестве подов по зонам

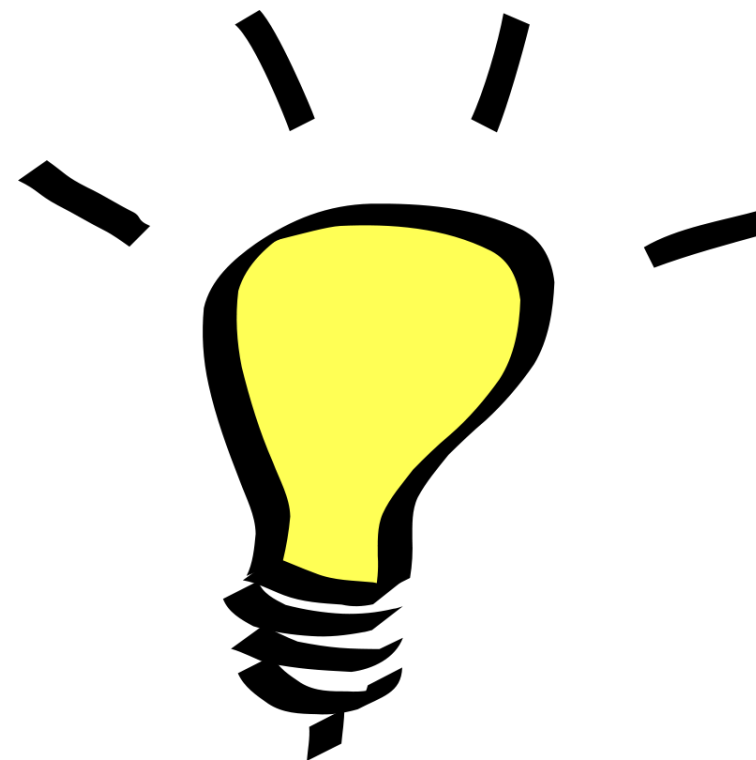
Что делать если нельзя удовлетворить настройки распределения

Лейбл на нодах по которому делать распределение

По какому лейблу распределять поды

# Как это можно использовать

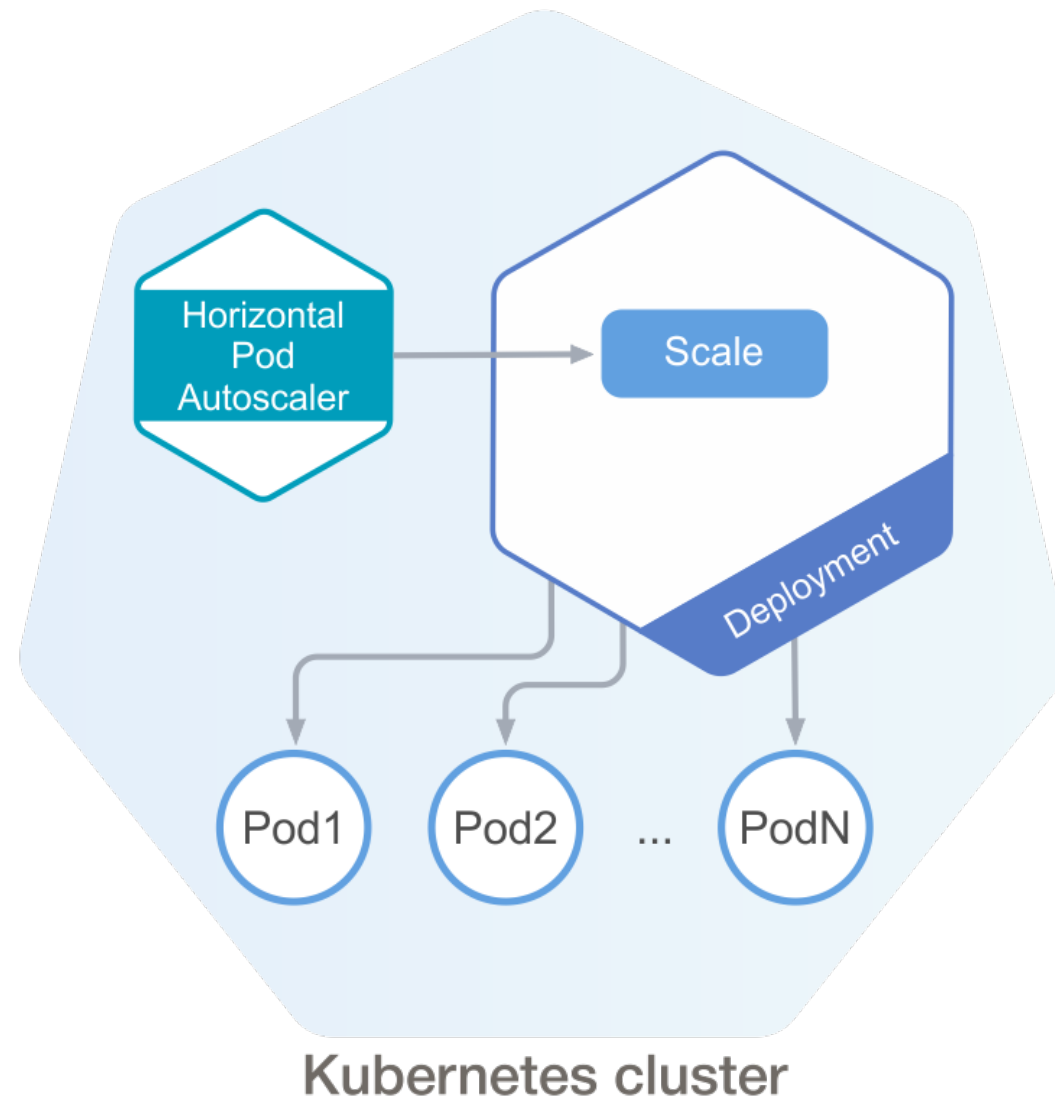
- Распределение подов по **зонам доступности**
- Распределение подов по **стойкам**
- Экономия денег
  - Запуск подов на **спотах**
  - **Гарантия** запуска части подов на обычных хостах





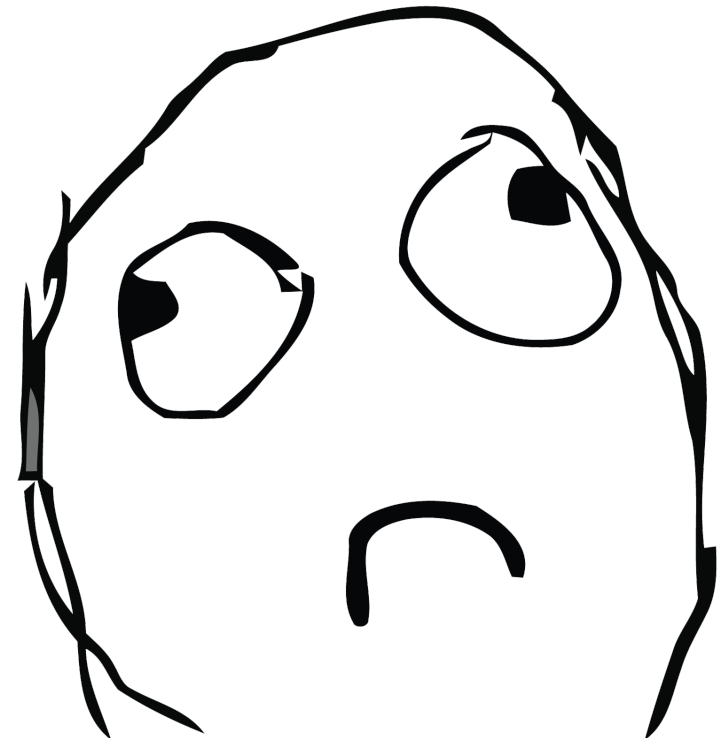
Еще одна проблема –  
облака это очень дорого

# Автоскейлинг в Kubernetes



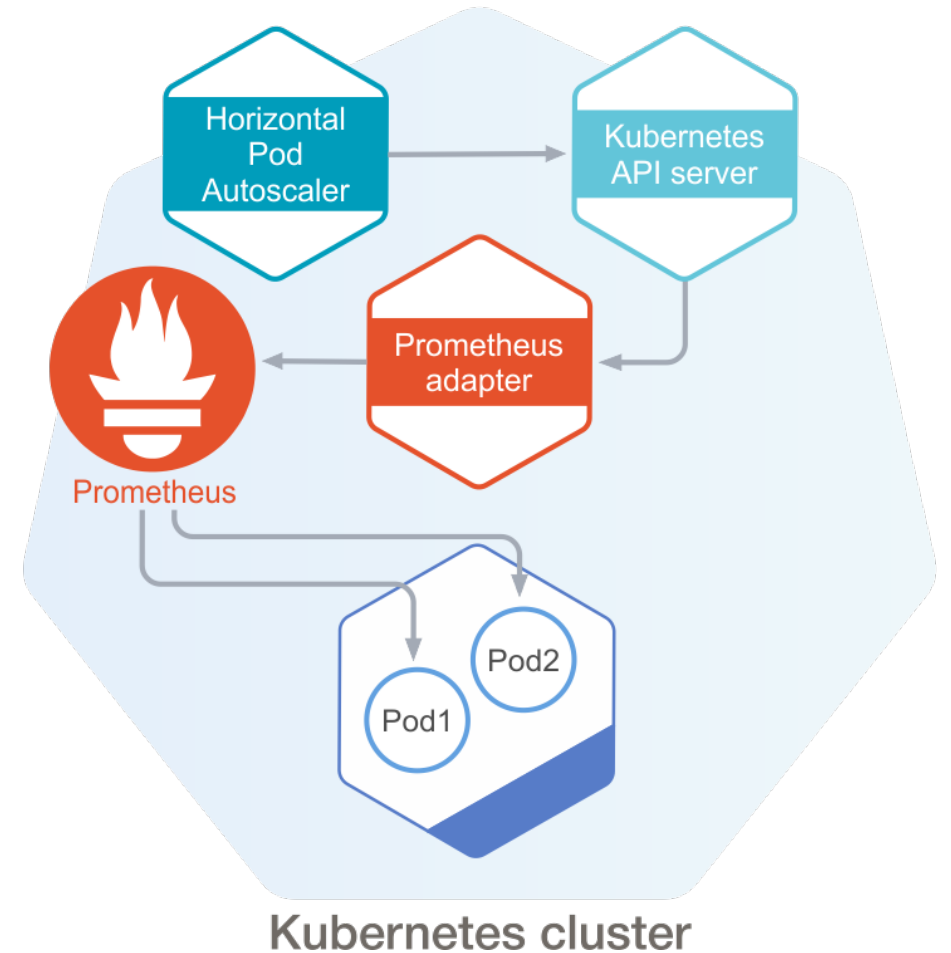
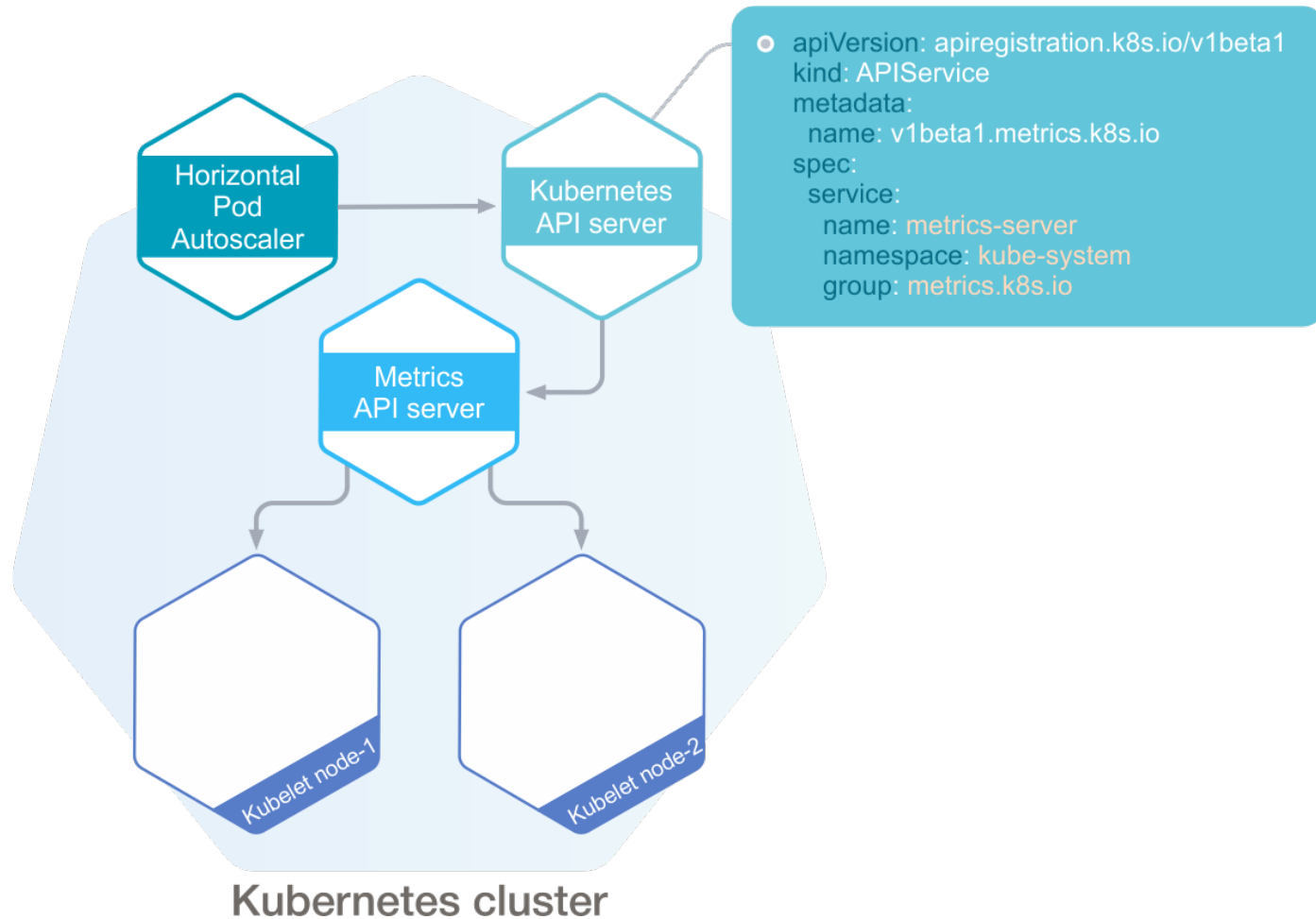
# Автоскейлинг в Kubernetes

- Из коробки работает только на основе нагрузки CPU
- Решает только задачу скейлинга подов





# Horizontal Pod Autoscaler v2



# Full Metal Jacket автоскейлинг

- HPA v2
- Cluster Autoscaler
- Автоскейлинг CI/CD
- Динамические стенды разработки



# Давайте прикинем цифры

- 10 микросервисов (mSvc)
- Каждый под = 500m/512Mi/200 RPS
- Входящий RPS днём = 1000 RPS / ночью = 50 RPS
- Ночь = 8 часов
- Нода = 7 руб / час



# Давайте прикинем цифры

- Днём  $1000 \text{ RPS} / 200 \text{ RPS} * 10\text{mSvc} = 50$  подов
  - По ресурсам 25 CPU / 25 Gi RAM = 5 нод
- Ночью  $50 \text{ RPS} / 200 \text{ RPS} * 10\text{mSvc} = 20$  подов (избыточность сохраняем)
  - По ресурсам 10 CPU / 10 Gi RAM = 2 ноды
- $8 \text{ часов} * 3 \text{ ноды} * 7 \text{ руб} / \text{ час} * 30 \text{ дней} = 5040 \text{ руб} / \text{ мес}$



## Еще накинём

- Девы / стейджи
- CI/CD раннеры

---

До 10 тыр / мес при общих затратах 30 тыр / мес





**30 % ЭКОНОМИЯ**



Облачная инфраструктура требует  
облачных подходов

# Контакты

 p.selivanov@vk.team

 [linkedin.com/in/pjamm](https://www.linkedin.com/in/pjamm)

