

拉勾教育

— 互联网人实战大学 —

# 《Kubernetes 原理剖析与实战应用》

正范

— 拉勾教育出品 —

## 02 | 高屋建瓴：Kubernetes 的架构 为什么是这样的？

如何将 Kubernetes 应用到自己的项目中

首先就**需要了解 Kubernetes 的架构**

Google 使用 **Linux 容器** 时间超过 15 年  
期间共创建了三套容器调度管理系统  
分别是 **Borg**、**Omega** 和 **Kubernetes**



# Borg 的架构

拉勾教育

— 互联网人实战大学 —

**Borg 论文地址**

<https://research.google/pubs/pub43438/>



Borg 用 Cell 来定义一组机器资源

Google 内部一个中等规模的 Cell 可以管理

1 万台左右服务器

这些服务器的配置可以是异构的

# Borg 的架构

拉勾教育

— 互联网人实战大学 —



Cell

Borg 用 Cell 来定义一组机器资源

Google 内部一个中等规模的 Cell 可以管理

1 万台左右服务器

这些服务器的配置可以是异构的



Cluster

即集群

一个数据中心可以同时运行一个或者多个集群

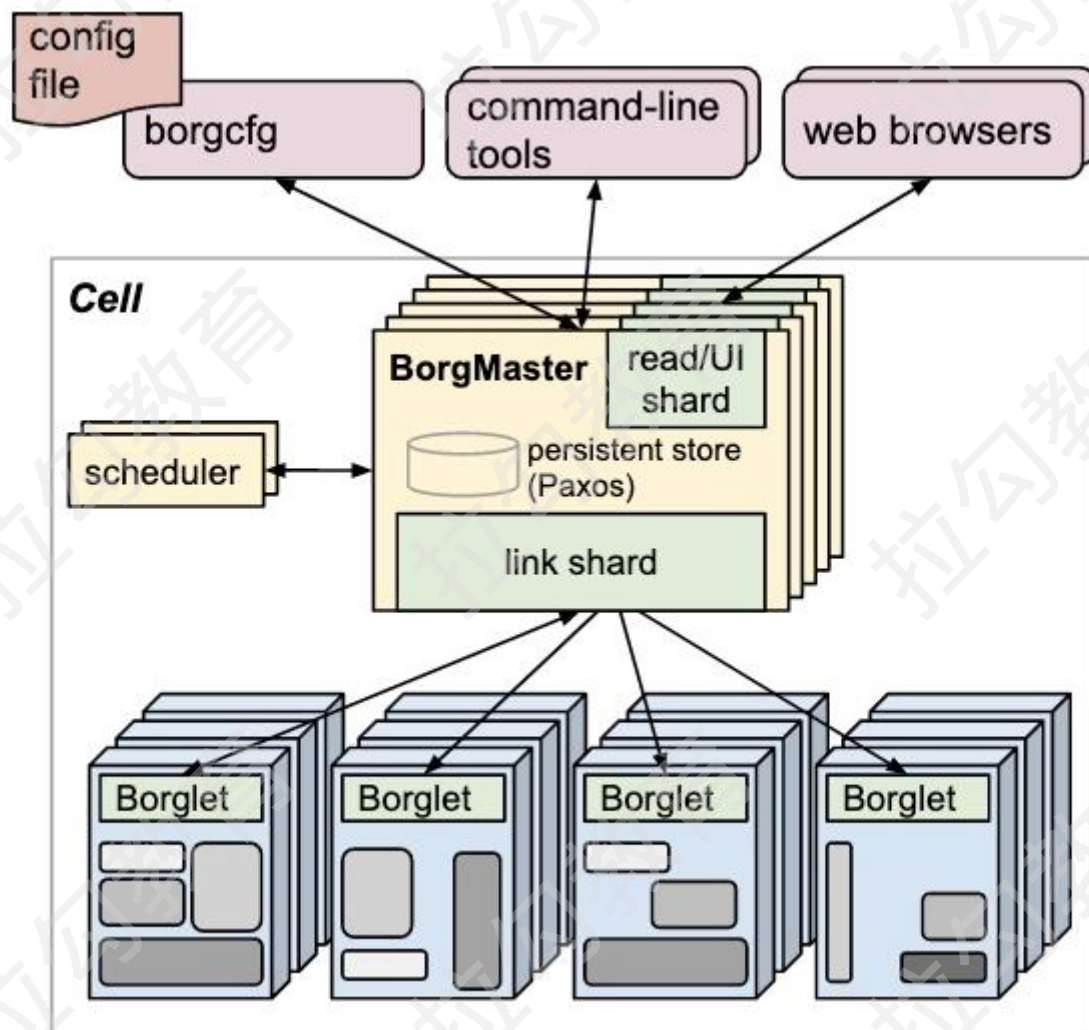
每个集群又可以有多多个 Cell

比如一个大 Cell 和多个小 Cell

# Borg 的架构

拉勾教育

— 互联网人实战大学 —

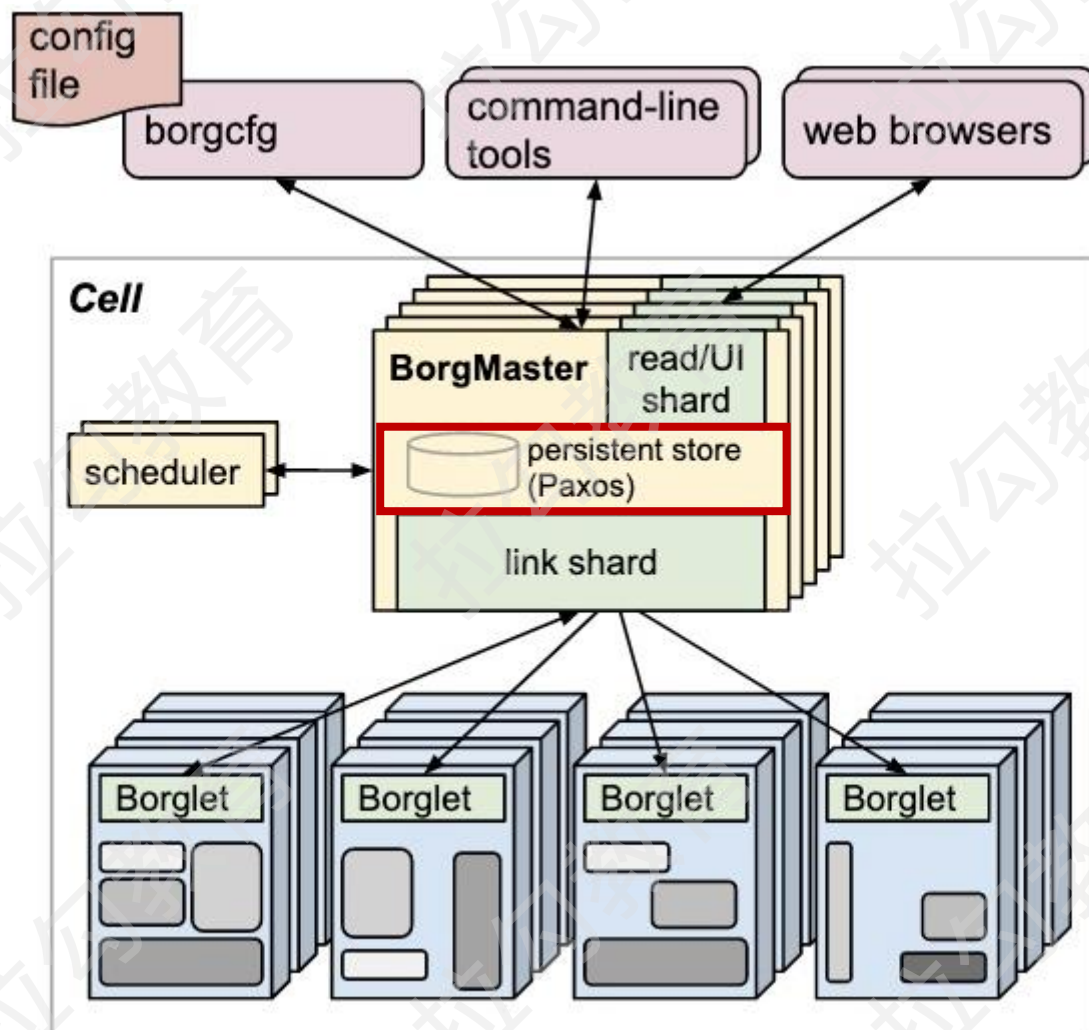




# Borg 的架构

拉勾教育

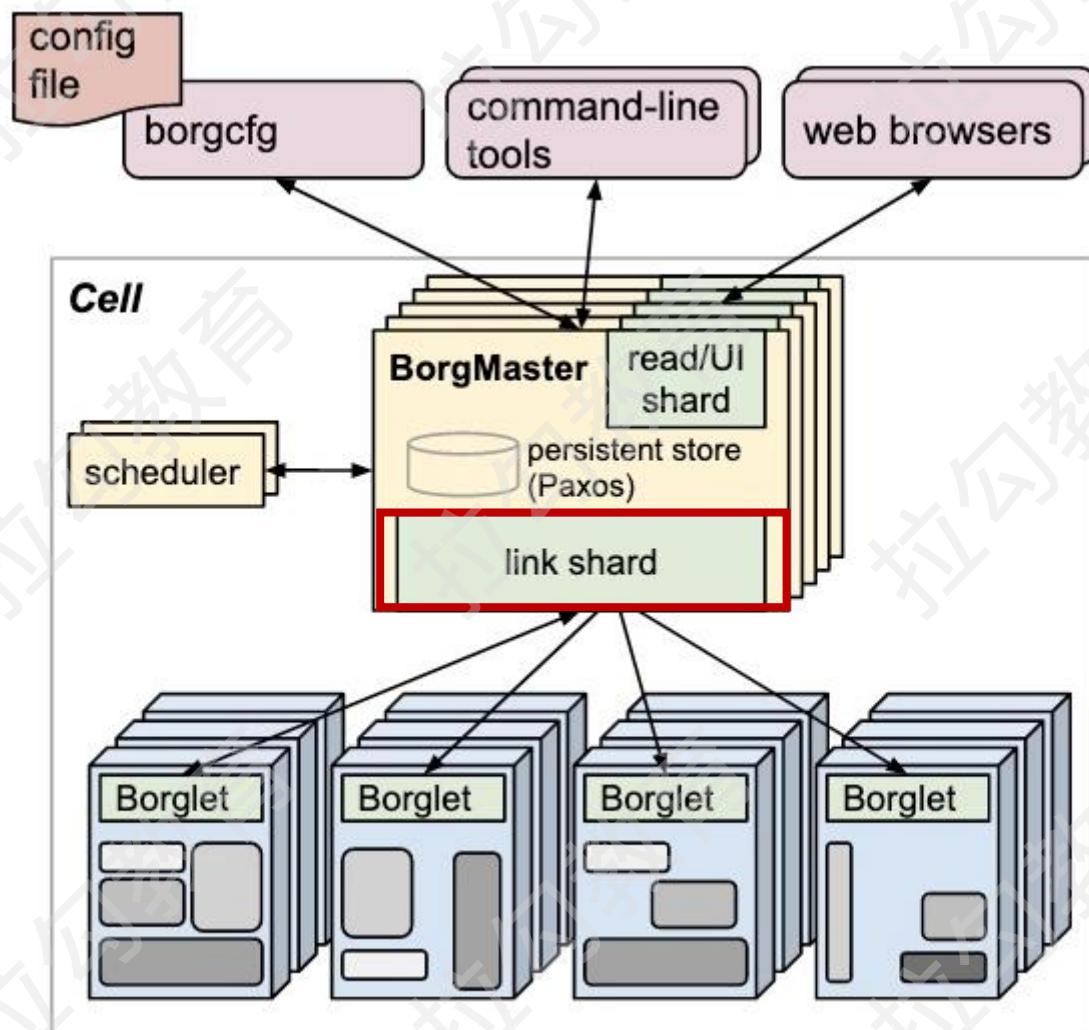
— 互联网人实战大学 —



# Borg 的架构

拉勾教育

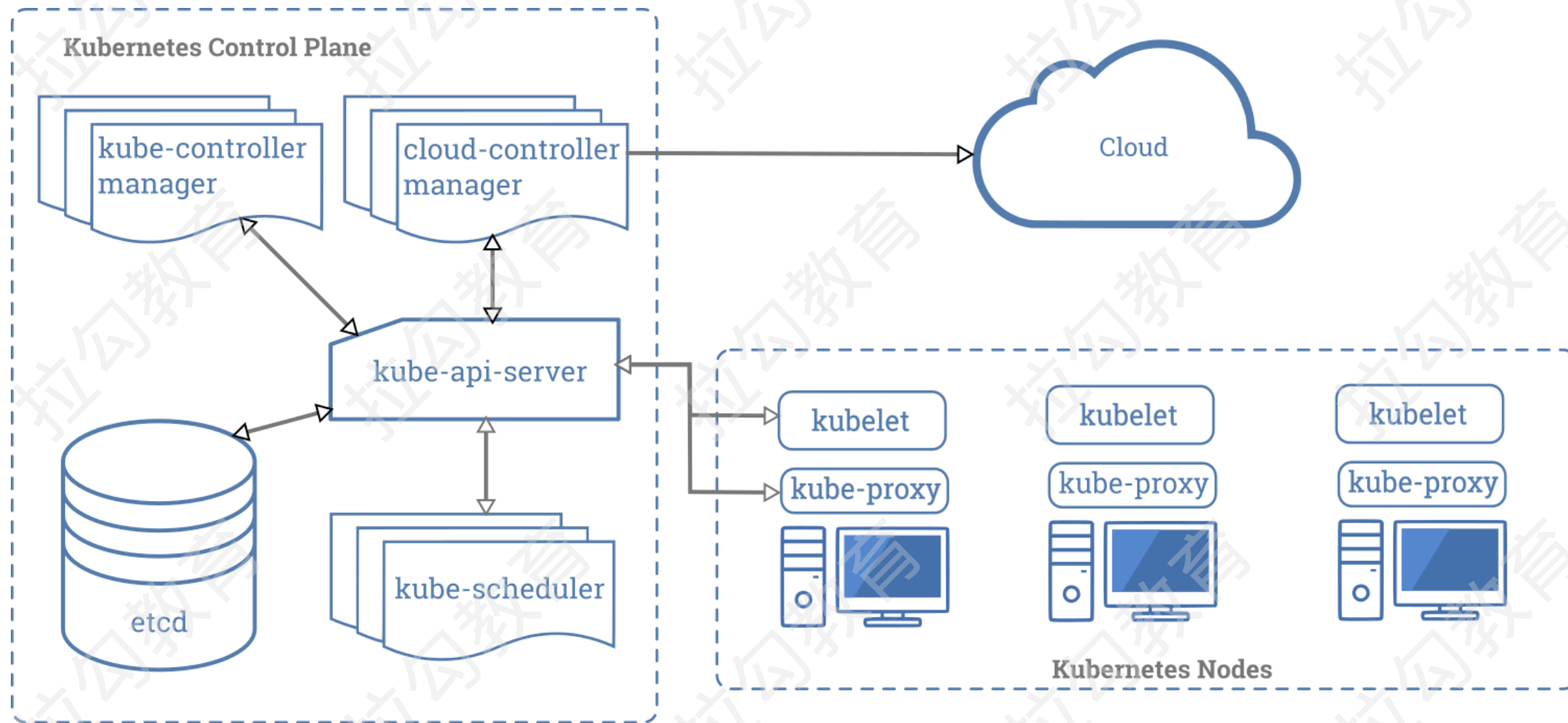
— 互联网人实战大学 —



# Kubernetes 的架构

拉勾教育

— 互联网人实战大学 —



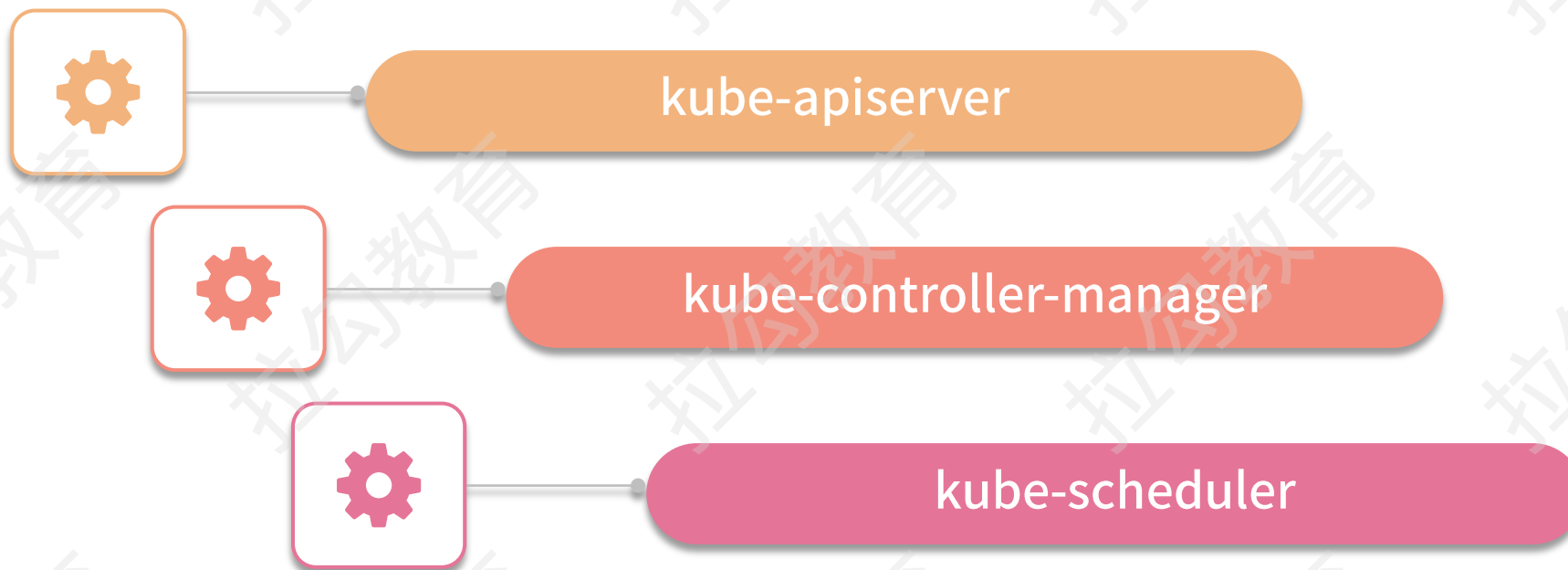
Paxos 和 Raft 的前世今生

<https://cloud.tencent.com/developer/article/1352070>

# Kubernetes 的组件

拉勾教育

— 互联网人实战大学 —



# Kubernetes 的组件

拉勾教育

— 互联网人实战大学 —



是整个 Kubernetes 集群的“灵魂”，是信息的汇聚中枢

提供了所有内部和外部的 API 请求操作的唯一入口

同时也负责整个集群的认证、授权、访问控制、服务发现等等能力

# Kubernetes 的组件

拉勾教育

— 互联网人实战大学 —



负责维护整个 Kubernetes 集群的状态，比如多副本创建、滚动更新等

# Kubernetes 的组件

拉勾教育

— 互联网人实战大学 —



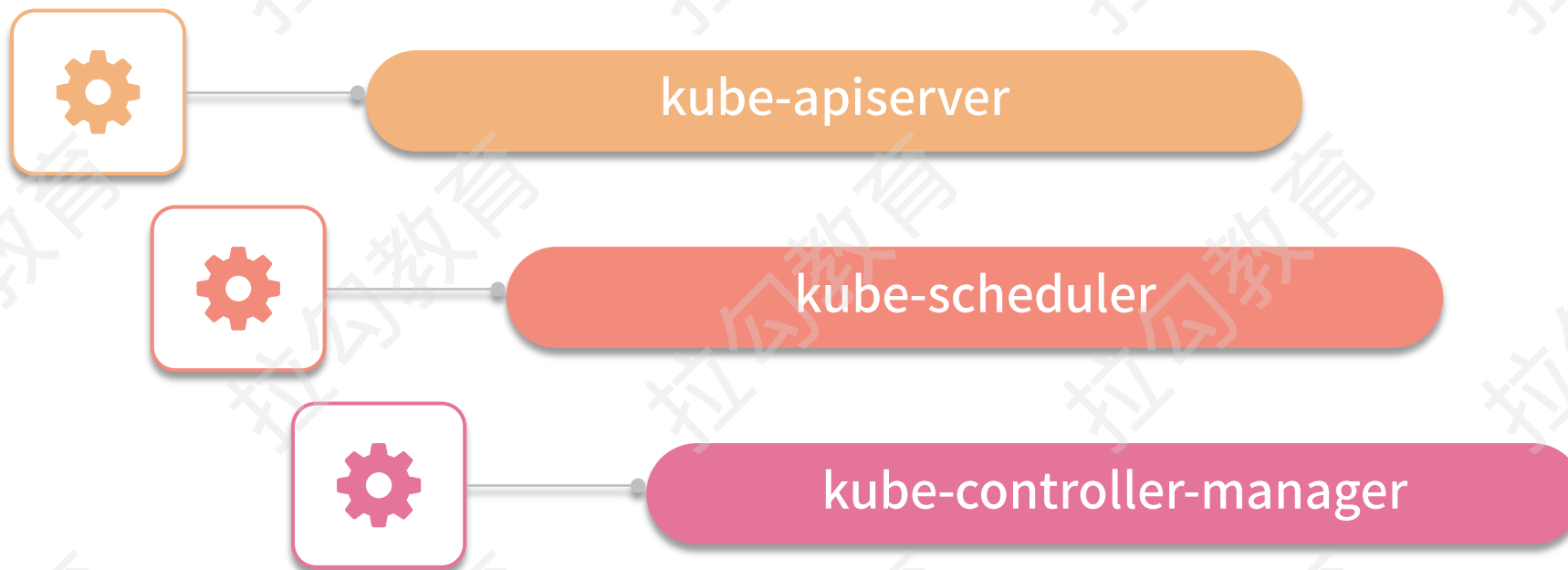
是监听未调度的 Pod，按照预定的调度策略绑定到满足条件的节点上



# Kubernetes 的组件

拉勾教育

— 互联网人实战大学 —



# Kubernetes 的组件

拉勾教育

— 互联网人实战大学 —



容器运行时主要负责容器的镜像管理以及容器创建及运行

# Kubernetes 的组件

拉勾教育

— 互联网人实战大学 —



容器运行时主要负责容器的镜像管理以及容器创建及运行



负责维护 Pod 的生命周期，比如创建和删除 Pod 对应的容器

# Kubernetes 的组件

拉勾教育

— 互联网人实战大学 —



容器运行时主要负责容器的镜像管理以及容器创建及运行



负责维护 Pod 的生命周期，比如创建和删除 Pod 对应的容器



主要负责 Kubernetes 内部的服务通信  
在主机上维护网络规则并提供转发及负载均衡能力

# Kubernetes 的组件

拉勾教育

— 互联网人实战大学 —



CoreDNS

负责为整个集群提供 DNS 服务



Dashboard

提供 GUI 可视化界面



Ingress Controller

为服务提供外网接入能力



Fluentd + elasticsearch

为集群提供日志采集、存储与查询等能力

# Master 和 Node 的交互方式

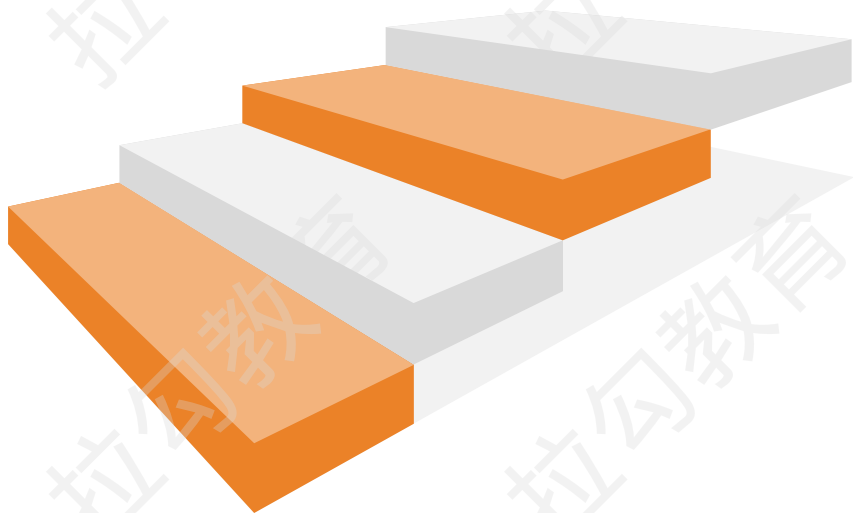
拉勾教育

— 互联网人实战大学 —

Kubernetes 中所有的状态都是采用**上报**的方式实现的

API Server 不会主动跟 Kubelet 建立请求链接

所有的容器状态汇报都是由 Kubelet 主动向 API Server 发起的



# Master 和 Node 的交互方式

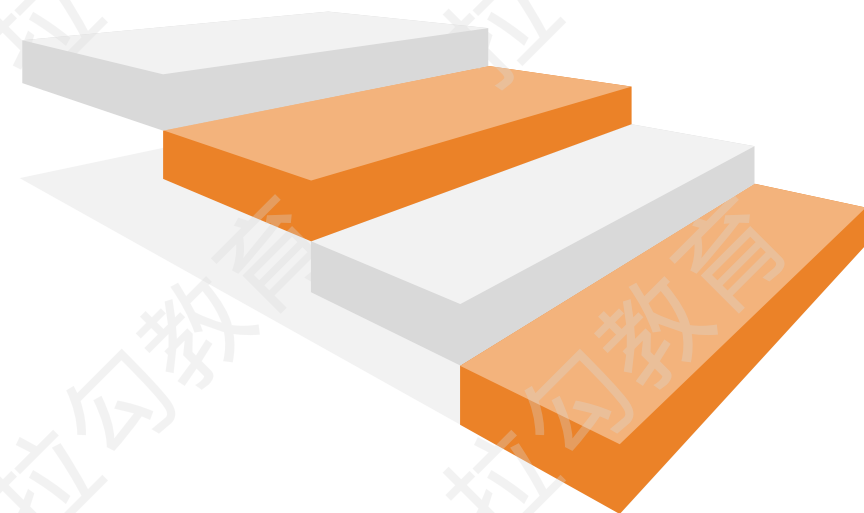
拉勾教育

— 互联网人实战大学 —

新增的 Node 被 APIServer 纳管后

Kubelet 进程就会定时向 APIServer 汇报 “心跳”，即汇报自身的状态

包括自身健康状态、负载数据统计等



# Master 和 Node 的交互方式

拉勾教育

— 互联网人实战大学 —

当一段时间内心跳包没有更新

那此时 kube-controller-manager 就会将其标记为 **NodeLost (失联)**



## Kubernetes 系统在设计时很注重容错性和可扩展性

通过 backoff retry、多副本、滚动升级等机制

增强集群的容错性，提高 Kubernetes 系统的稳定性

Next: 《03 | 集群搭建：手把手教你玩转 Kubernetes 集群搭建》

# 拉勾教育

— 互联网人实战大学 —



关注拉勾「教育公众号」  
获取更多课程信息