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 —

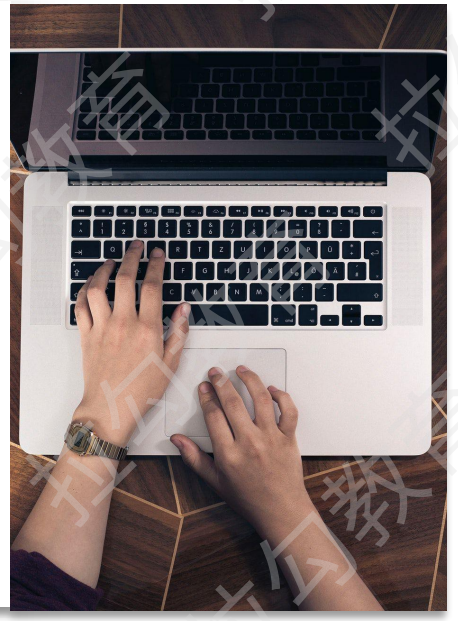
26 | 网络插件：Kubernetes 搞定网络原来可以如此简单？

前言

Kubernetes 可以跑在任何环境中

比如公有云、私有云、物理机、虚拟机、树莓派

任何基础设施（Infrastructure）对网络的需求都是最基本的



Kubernetes 网络模型基本原则

每一个 Pod 都拥有一个自己的独立的 IP 地址

且这些 Pod 之间可以不通过任何 NAT 互相连通

Kubernetes 网络模型

拉勾教育
互联网人实战大学

基于 IP-per-Pod 的方式

用户使用时不需要再额外考虑如何建立 Pod 之间的连接

也不用考虑如何将容器的端口映射到主机端口

Kubernetes 网络模型

- Pod 内容器之间的网络通信
- Pod 之间的网络通信
- Pod 到 Service 之间的网络通信
- 集群外部与内部组件之间的网络通信

Kubernetes 网络模型

同一 Pod 内的网络通信

Kubernetes 会为每一个 Pod 创建独立的**网络命名空间（netns）**

Pod 内的容器共享同一个网络命名空间

Kubernetes 网络模型

拉勾教育
互联网人实战大学

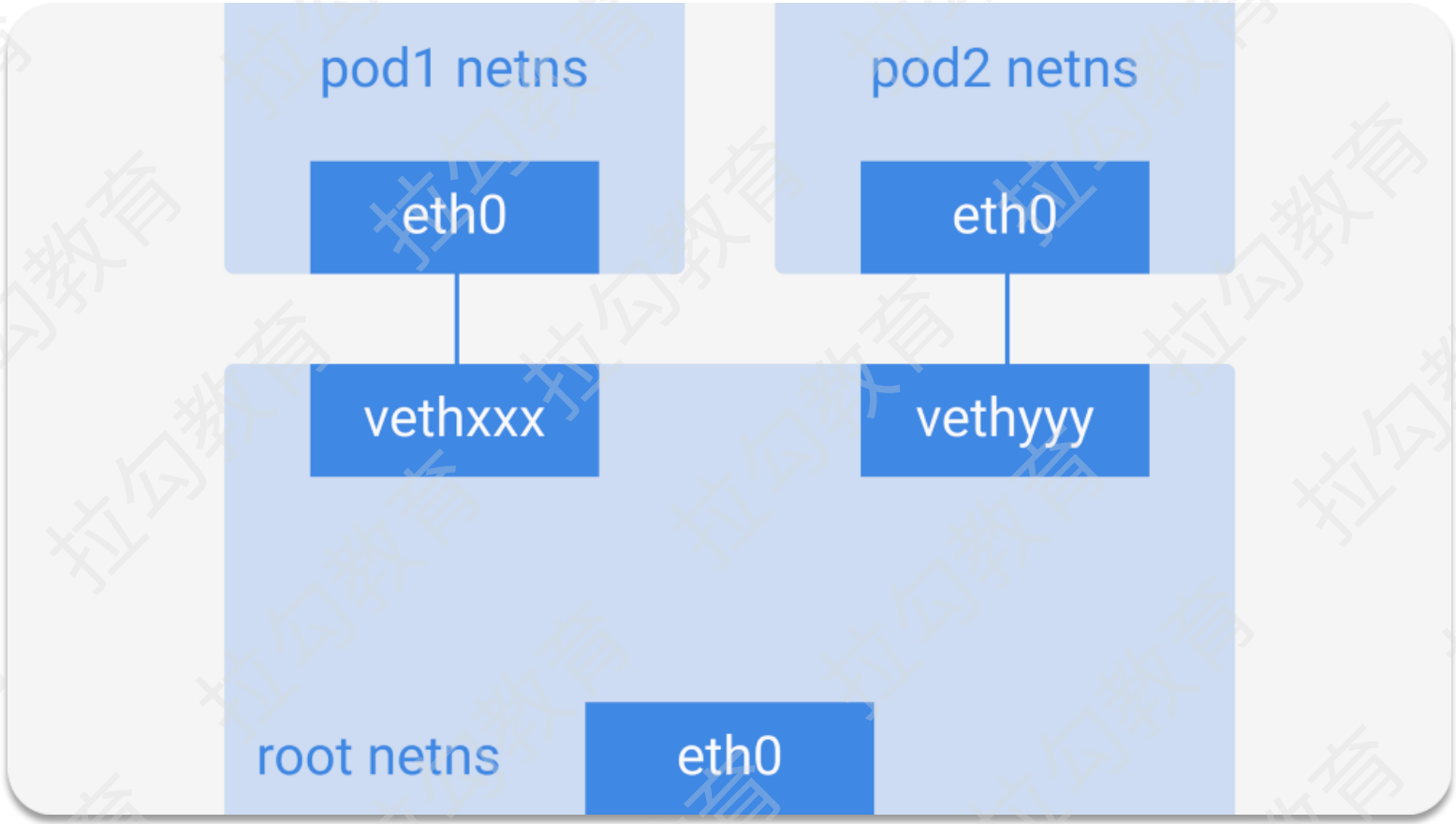
同主机通信

Pod 之间的网络通信

跨主机通信

Kubernetes 网络模型

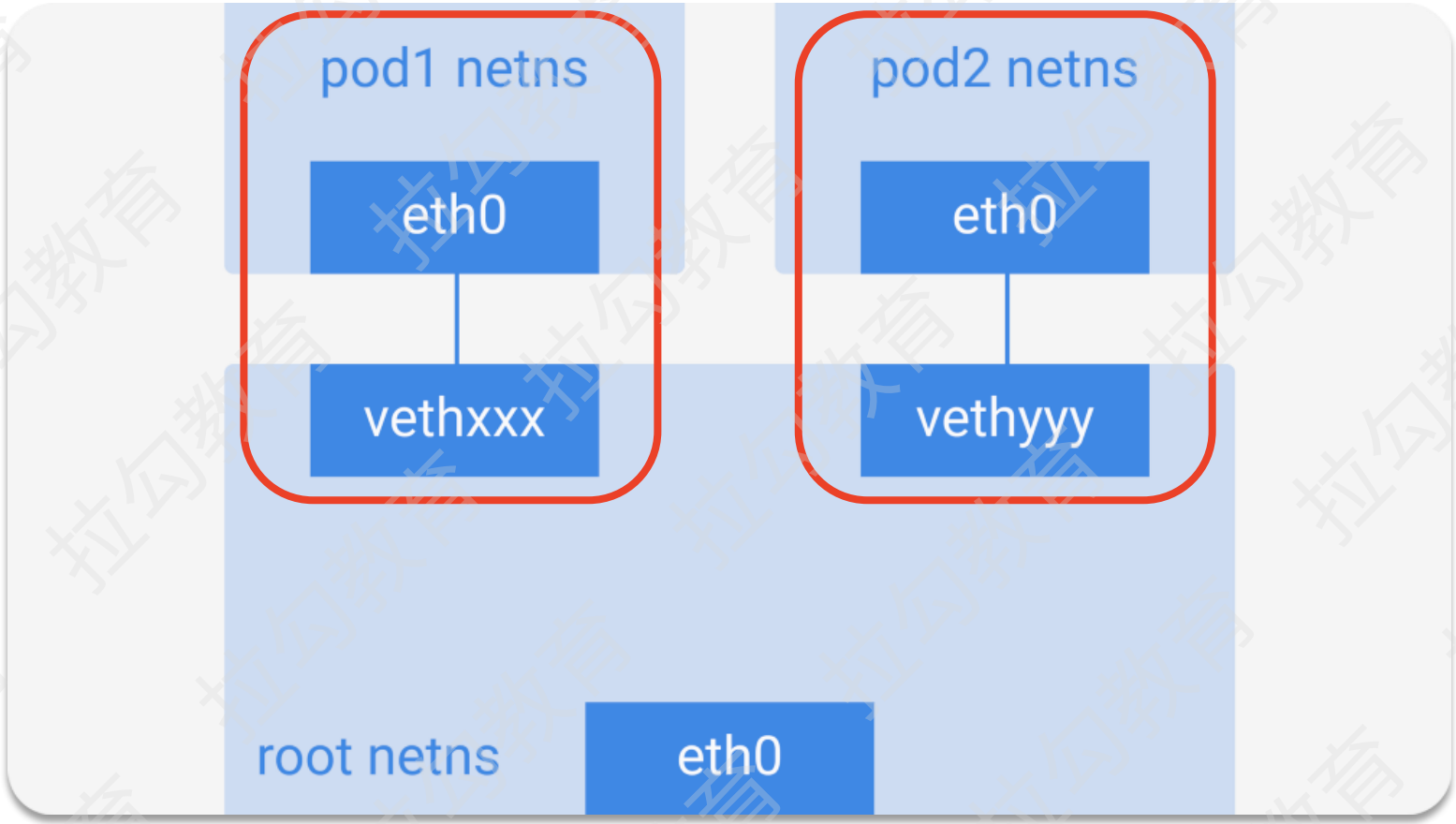
Pod 之间的网络通信



同主机上 Pod 间的通信

Kubernetes 网络模型

Pod 之间的网络通信



同主机上 Pod 间的通信

Kubernetes 网络模型

拉勾教育
互联网人实战大学

Pod 之间的网络通信

1 L2（ARP跨节点）

2 L3（IP路由跨节点表）

3 Overlay 网络

4 弹性网卡

Kubernetes 网络模型

Pod 之间的网络通信

整个 Kubernetes 集群中，每个 Pod 的 IP 地址必须是唯一的，不能与其它节点上的 Pod IP 冲突

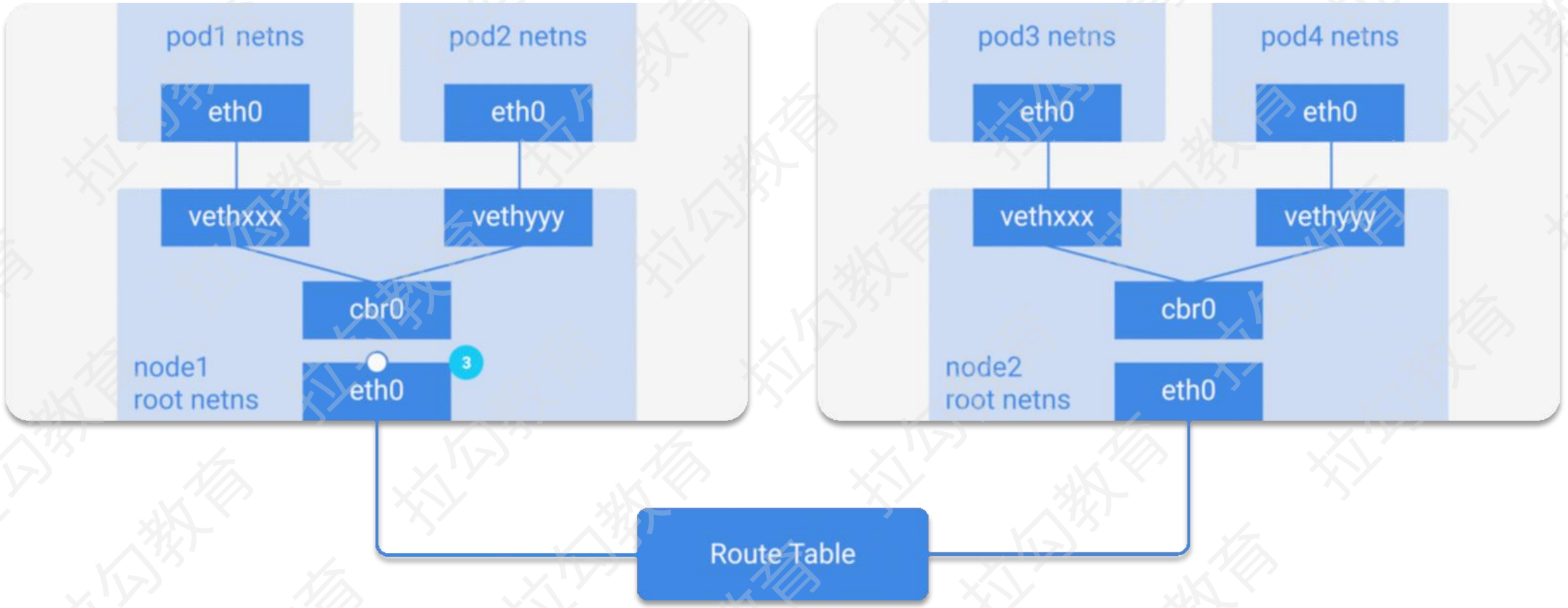
从 Pod 中发出的数据包不应该进行 NAT ，这样通信双方看到的 IP 地址就是对方 Pod 实际的地址

得知道 Pod IP 和所在节点 Node IP 之间的映射关系

Kubernetes 网络模型

拉勾教育
互联网人实战大学

Pod 之间的网络通信

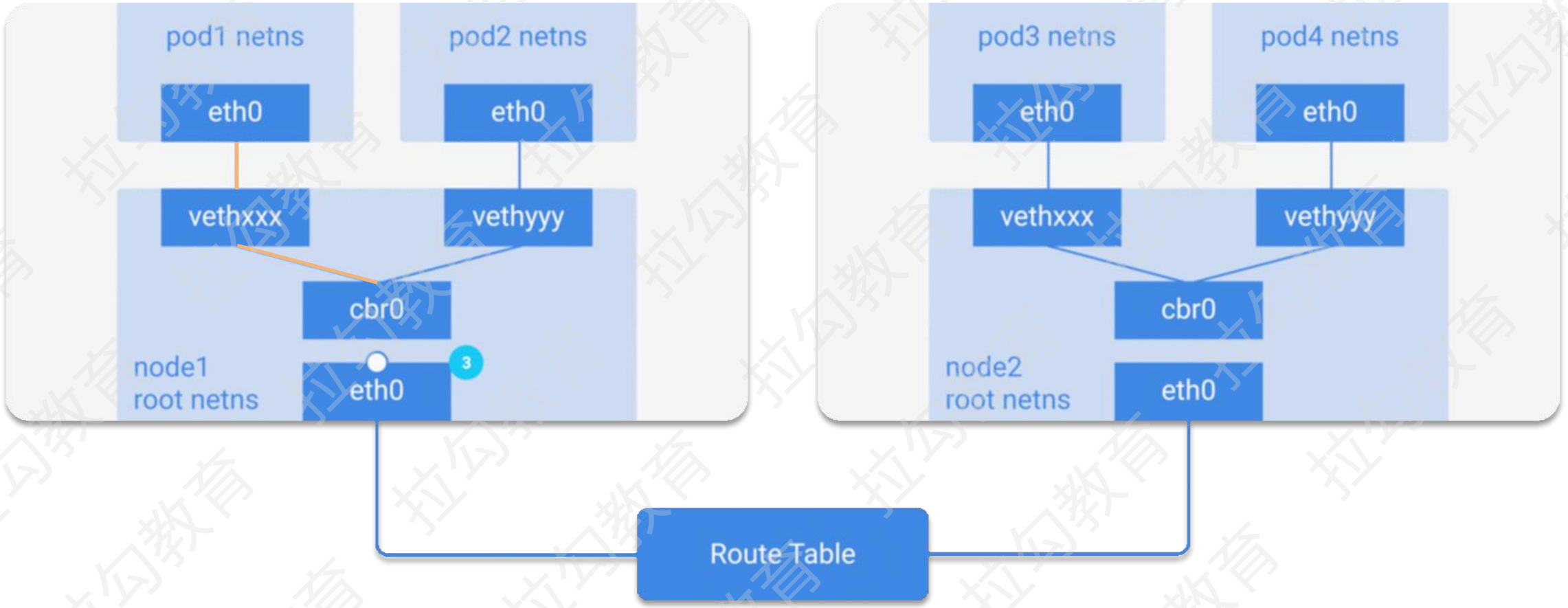


基于 L2 的跨节点 Pod 互相访问时的网络流量走向

Kubernetes 网络模型

拉勾教育
互联网人实战大学

Pod 之间的网络通信

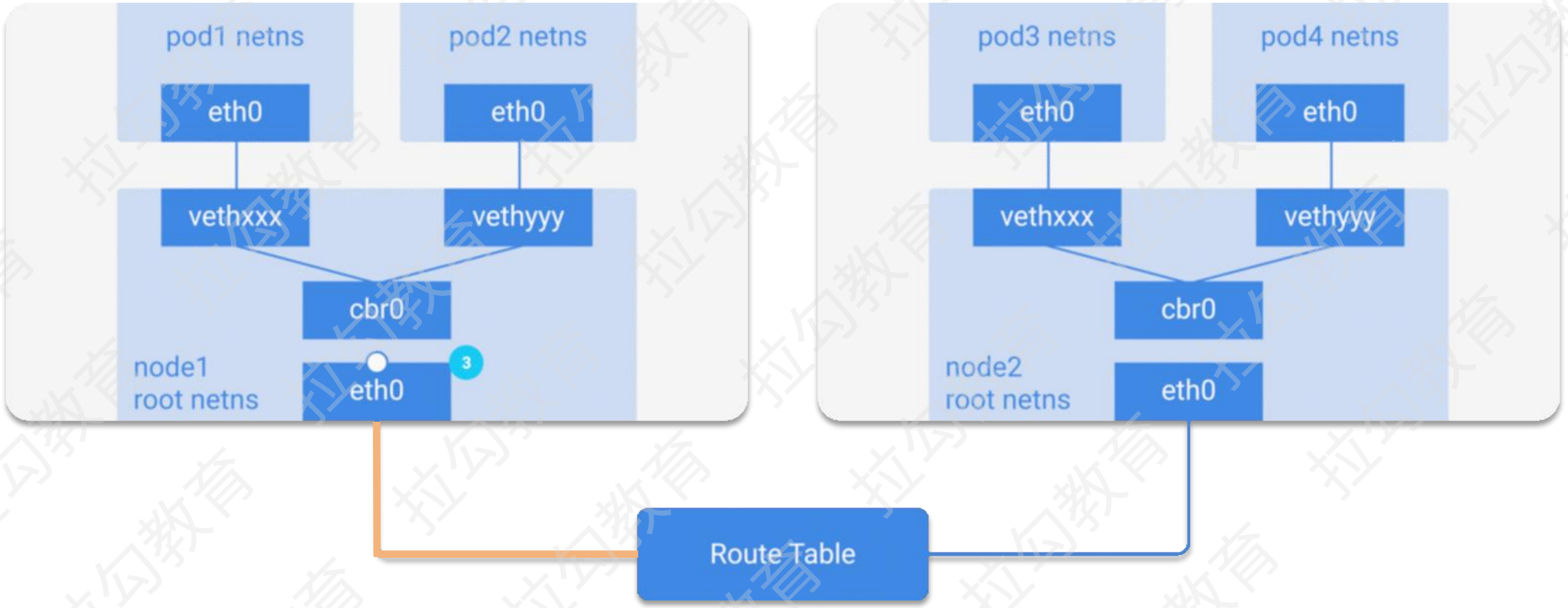


基于 L2 的跨节点 Pod 互相访问时的网络流量走向

Kubernetes 网络模型

拉勾教育
互联网人实战大学

Pod 之间的网络通信

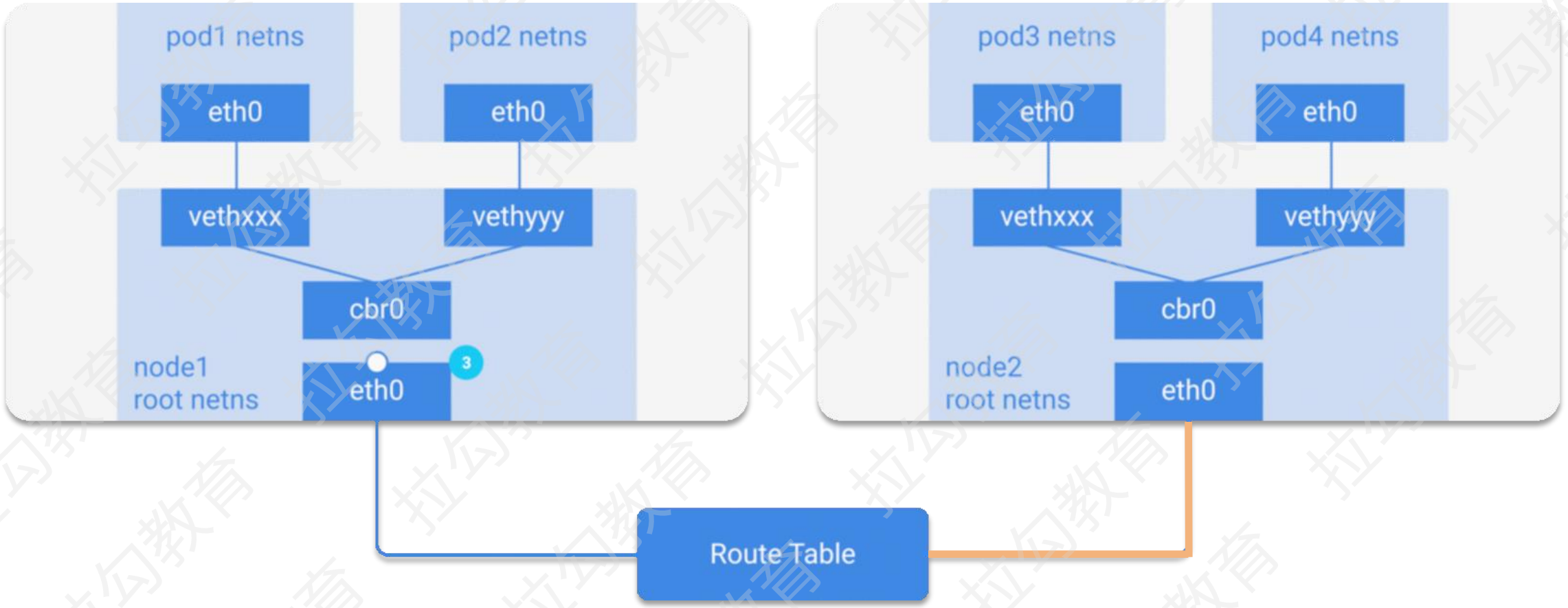


基于 L2 的跨节点 Pod 互相访问时的网络流量走向

Kubernetes 网络模型

拉勾教育
互联网人实战大学

Pod 之间的网络通信

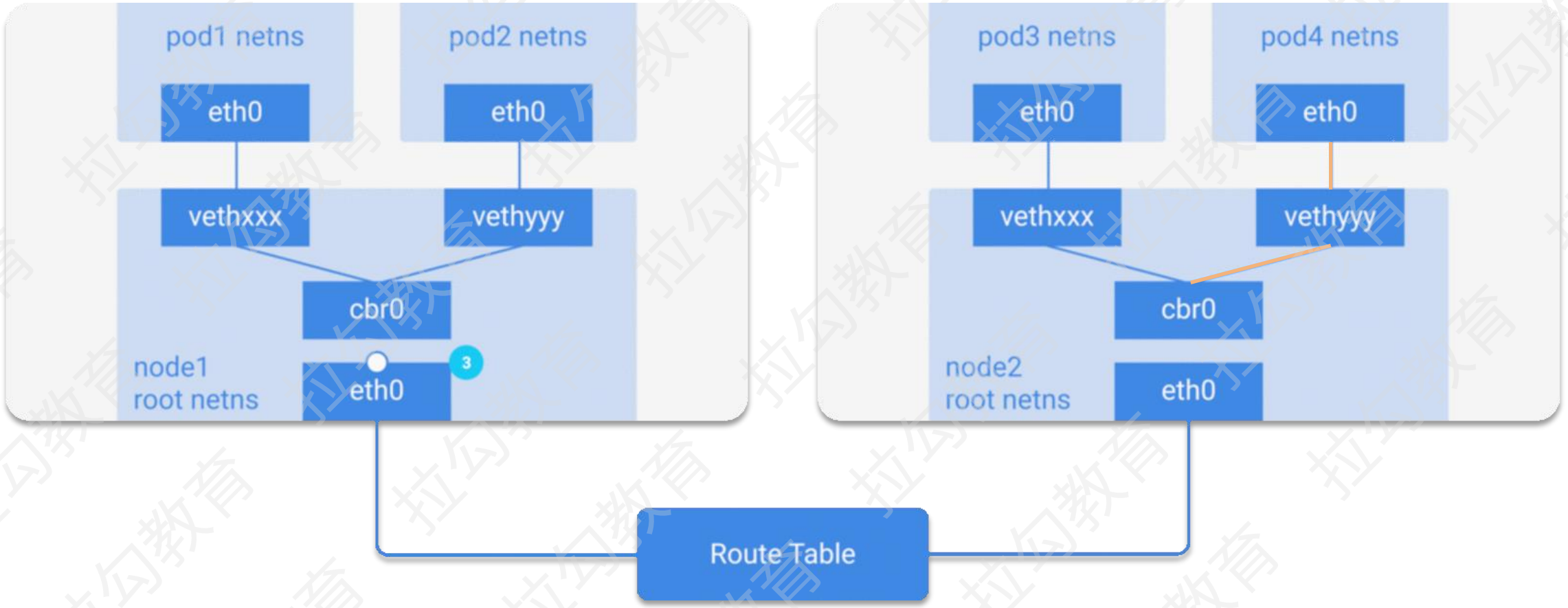


基于 L2 的跨节点 Pod 互相访问时的网络流量走向

Kubernetes 网络模型

拉勾教育
互联网人实战大学

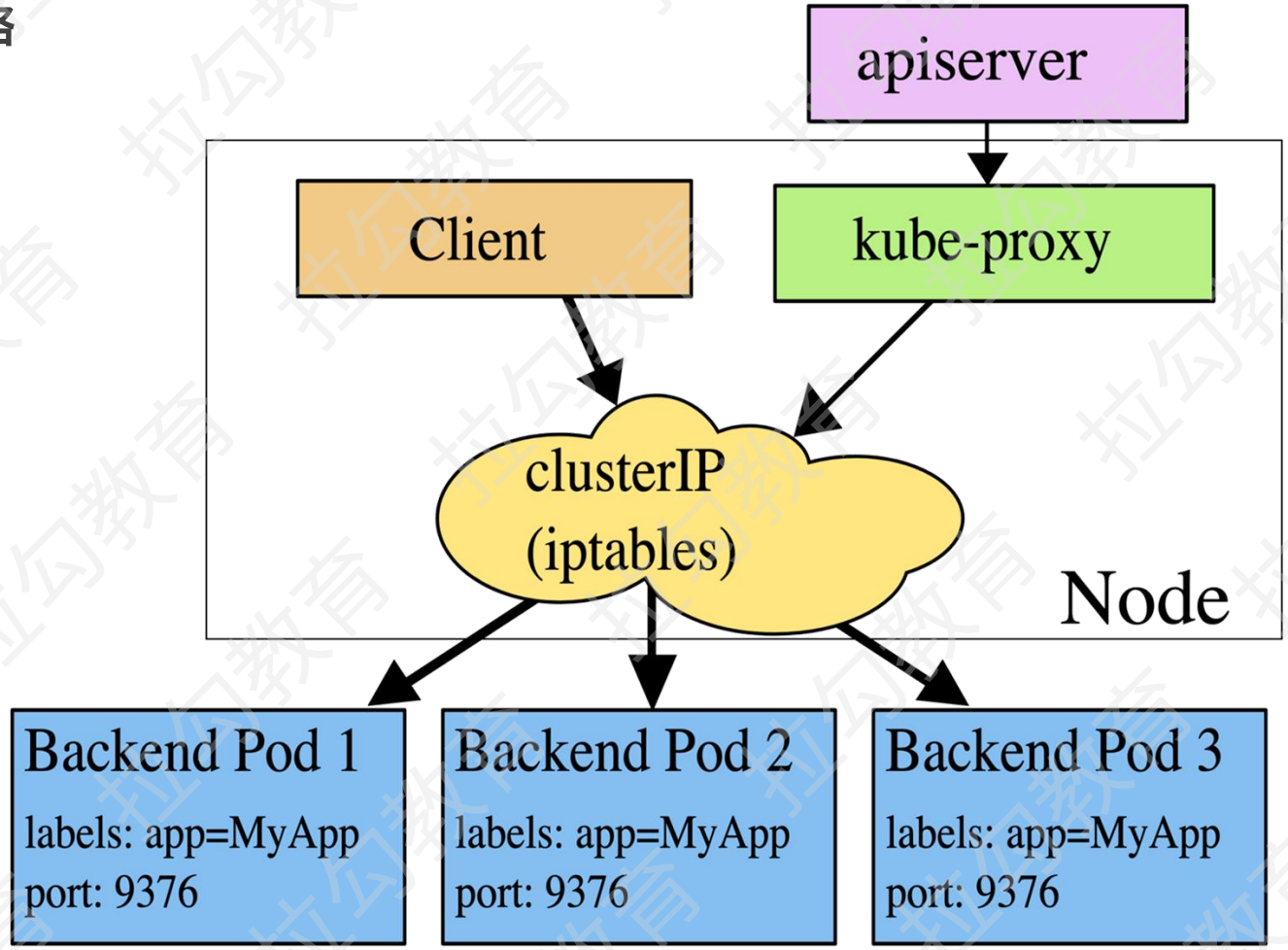
Pod 之间的网络通信



基于 L2 的跨节点 Pod 互相访问时的网络流量走向

Kubernetes 网络模型

Pod到 Service 的网络



iptables 模式

Kubernetes 网络模型

Pod到 Service 的网络

1. 对已经删除的 Service 进行清理

   删除不需要的 iptables 规则

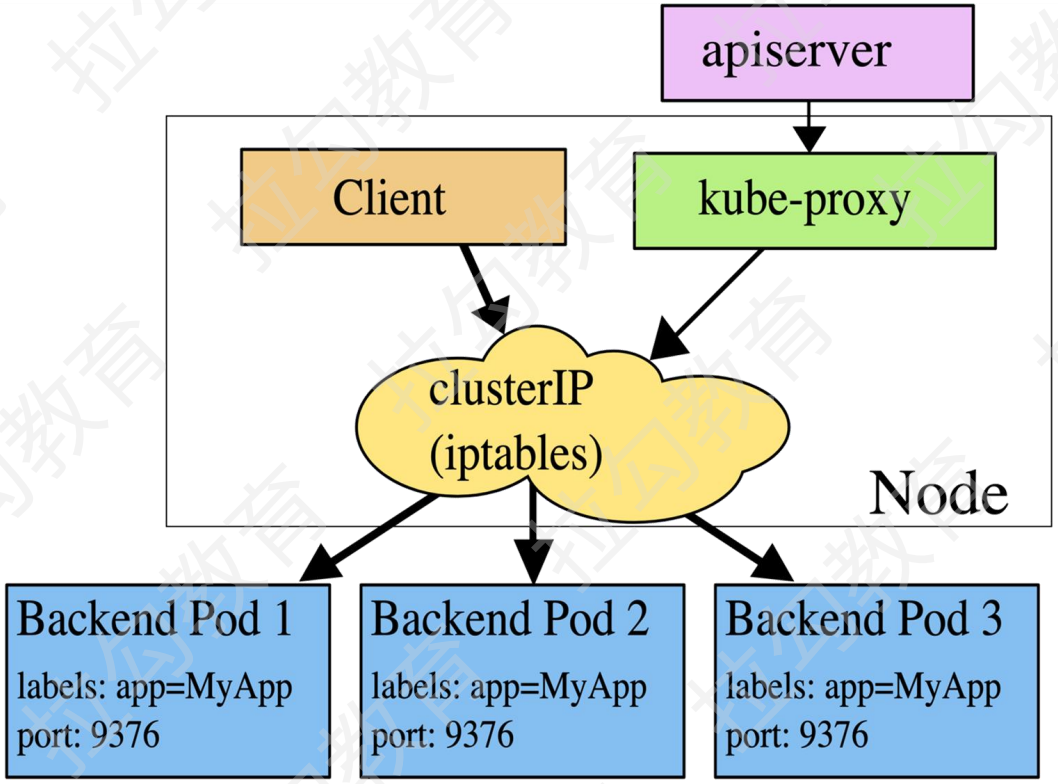
2. 如果一个新的 Service 没设置 ClusterIP

   则直接跳过，不做任何处理

3. 获取该 Service 的所有端口定义列表

   并逐个读取 Endpoints 里各个示例的 IP 地址

   生成或更新对应的 iptables 规则



iptables 模式

Kubernetes 网络模型

外部和服务间通信

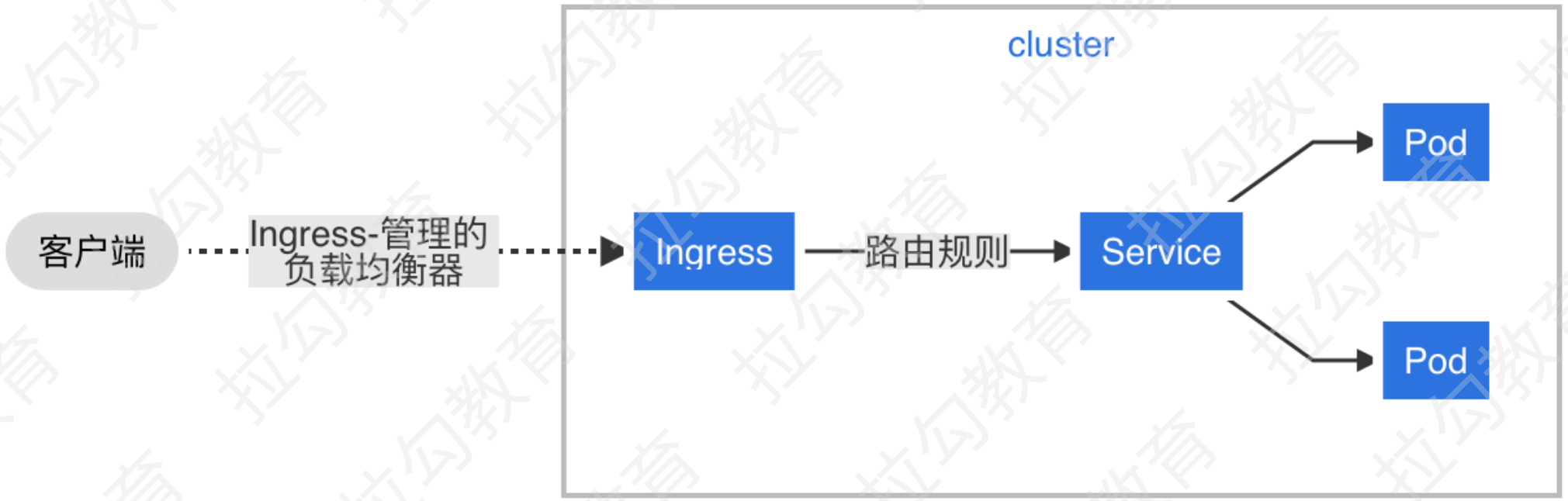
Pod 和 Service 都是 Kubernetes 中的概念

Ingress 可以将集群内服务的 HTTP 和 HTTPS 暴露出来，以方便从集群外部访问

Kubernetes 网络模型

拉勾教育
互联网人实战大学

外部和服务间通信

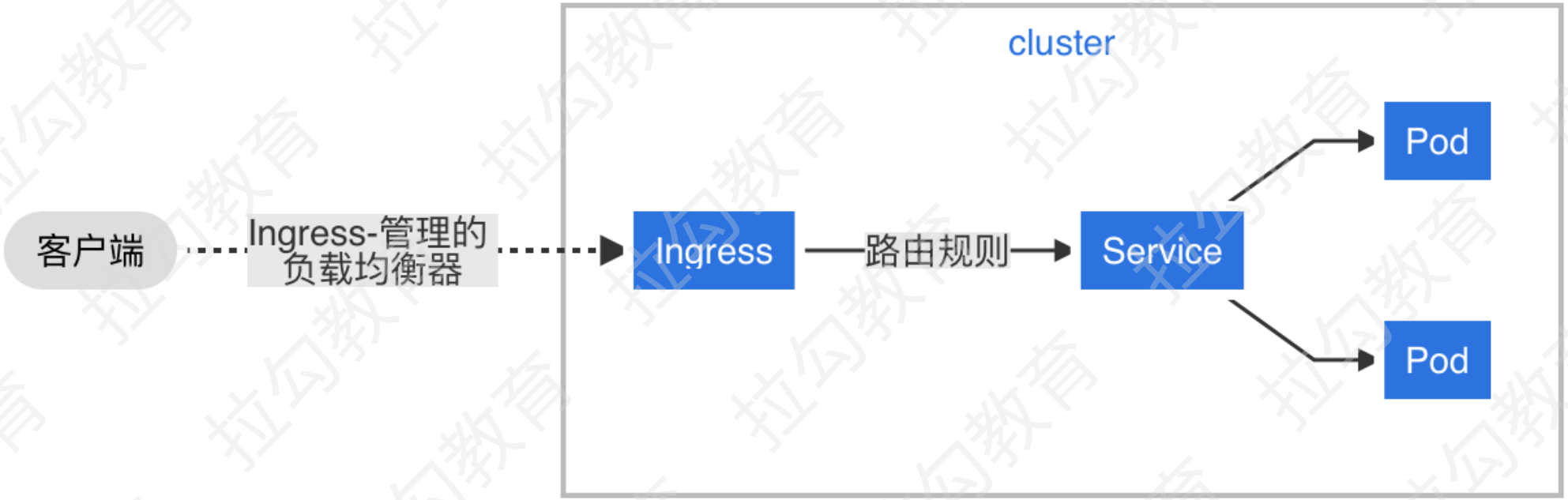


将所有流量都发送到同一 Service 的 Ingress

Kubernetes 网络模型

拉勾教育
互联网人实战大学

外部和服务间通信



将所有流量都发送到同一 Service 的 Ingress

官方文档：https://kubernetes.io/zh/docs/concepts/services-networking/ingress/

CNI（Container Network Interface）

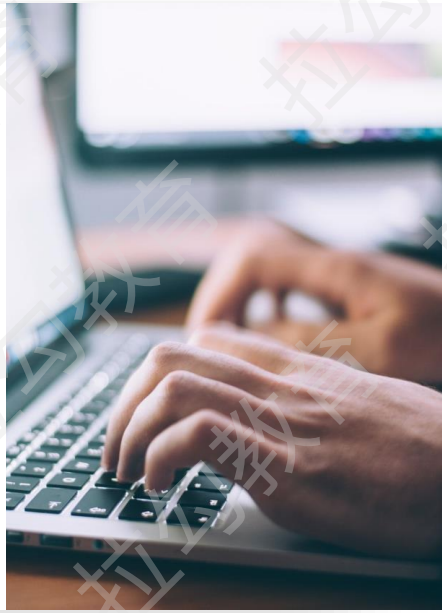
拉勾教育
互联网人实战大学

CNI

定义一套标准的、通用的接口

创建容器时分配网络资源，删除容器时释放网络资源

框架开放，可支持不同的网络模式，容易实现



CNI（Container Network Interface）

- 解析配置信息
- 执行具体的网络配置 ADD 或 DEL
- 对于 ADD 操作还需输出结果

CNI 插件的执行过程

CNI（Container Network Interface）

GitHub 项目

https://github.com/eranyanay/cni-from-scratch

插件列表

https://kubernetes.io/zh/docs/concepts/cluster-administration/networking/#%E5%A6%82%E4%BD%95%E5%AE%9E%E7%8E%B0-kubernetes-%E7%9A%84%E7%BD%91%E7%BB%9C%E6%A8%A1%E5%9E%8B

写在最后

IP-per-Pod

基于 IP-per-Pod 的基本网络原则

Kubernetes 设计出 Pod - Deployment - Service 这样经典的 3 层服务访问机制

极大地方便开发者在 Kubernetes 部署自己的服务

Next：《27 | 网络插件：Kubernetes 搞定网络原来可以如此简单？》

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

L / A / G / O / U