

BGP 与 Kube-router

BGP 简介

BGP-> Border Gateway Protocol

BGP-4(第四版本), RFC 4271

运行于 TCP 179 端口

能为多宿主的 ISP 提供更好的网络冗余

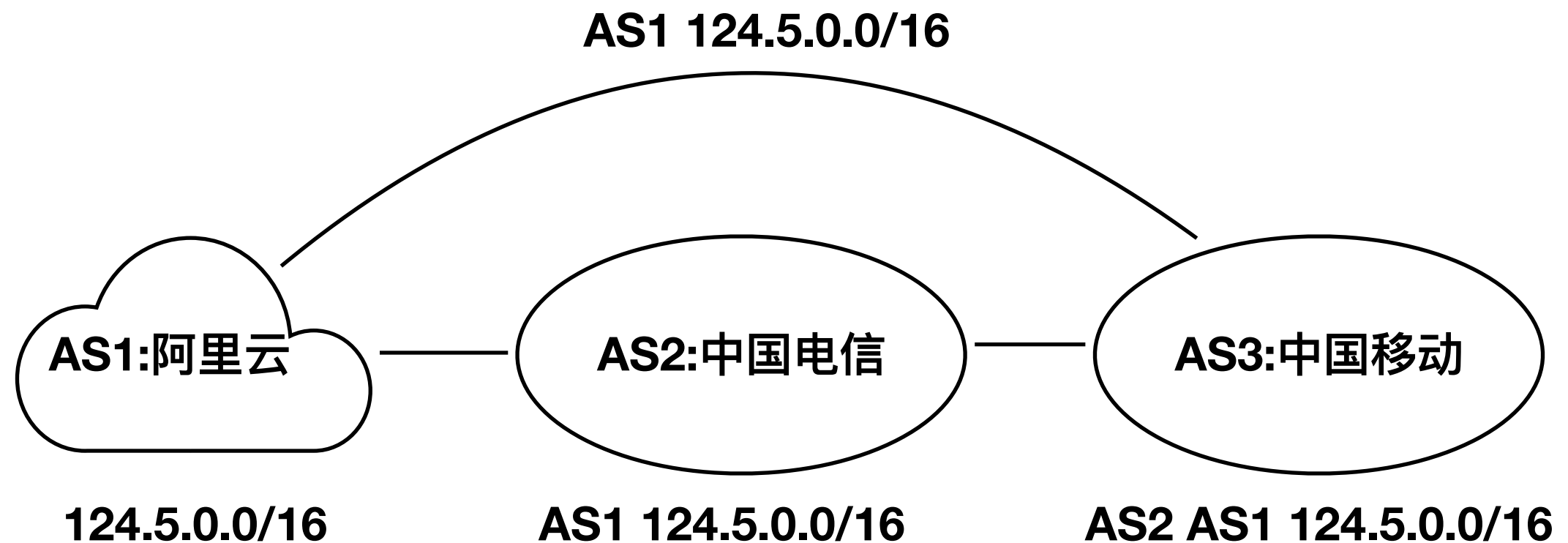
完全分散的路由系统

问题：浏览器发出的请求，如何通过中间的层层路由转发，最后到达指定的服务器？
分散的路由系统是如何实现的？

BGP

AS->autonomous system, 自治系统

AS 号码: 1-65411(Public), 64512-65535(private)

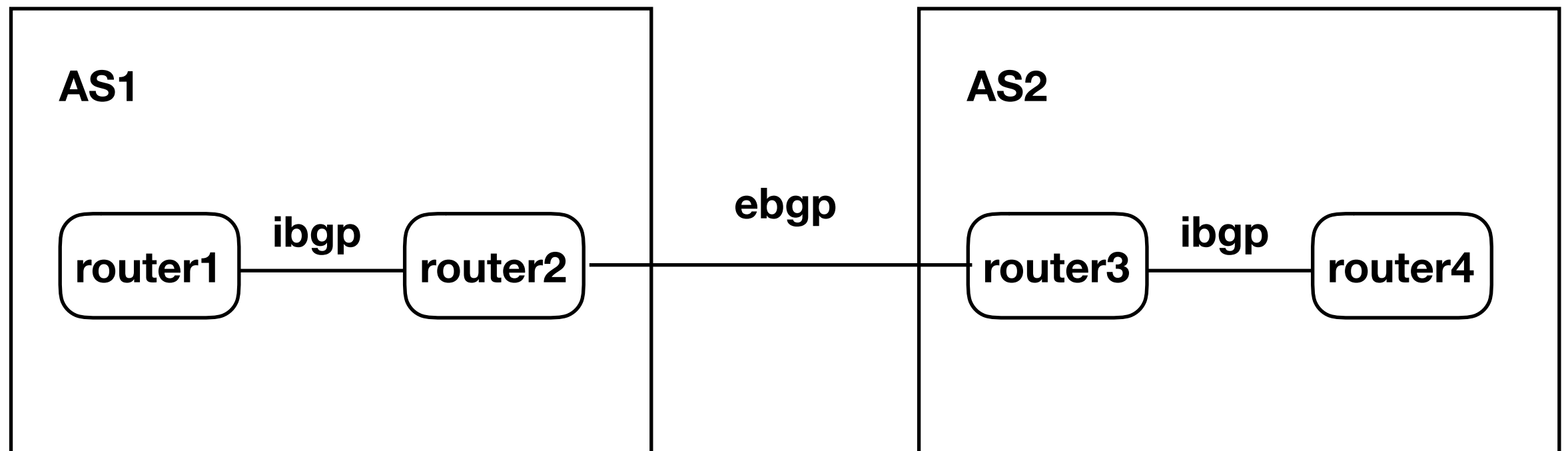


BGP

EBGP: external BGP, 用于 AS 之间的路由信息交换

IBGP: internal BGP, 用于 AS 内部的路由信息交换

区别: ebgp 转发时会修改 next-hop, ibgp 获取到的路由不被转发给其他 ibgp



BGP报文

Open: 用于建立 BGP 连接

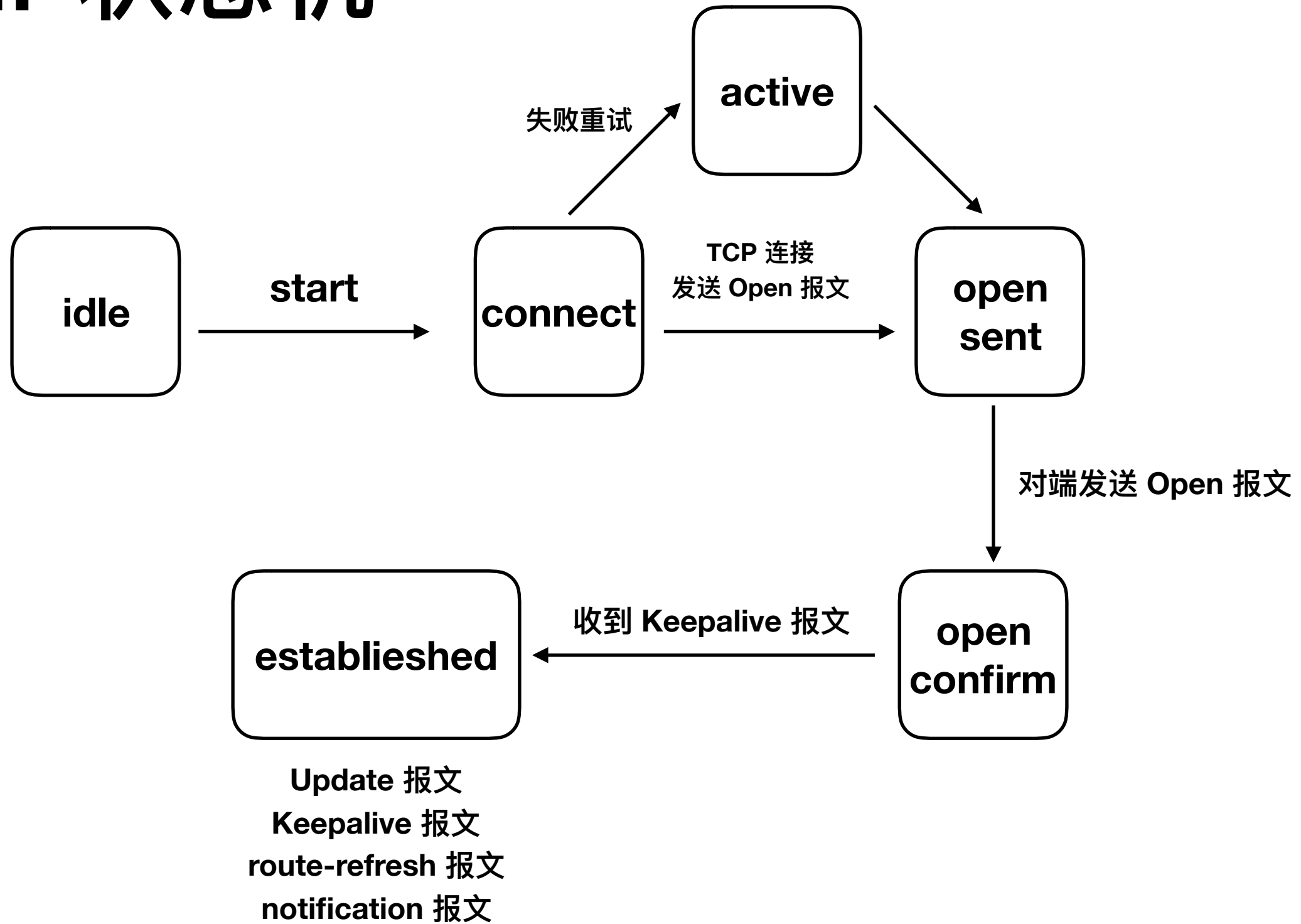
Update: 用于交换路由信息

Keepalive: 保持 BGP 连接

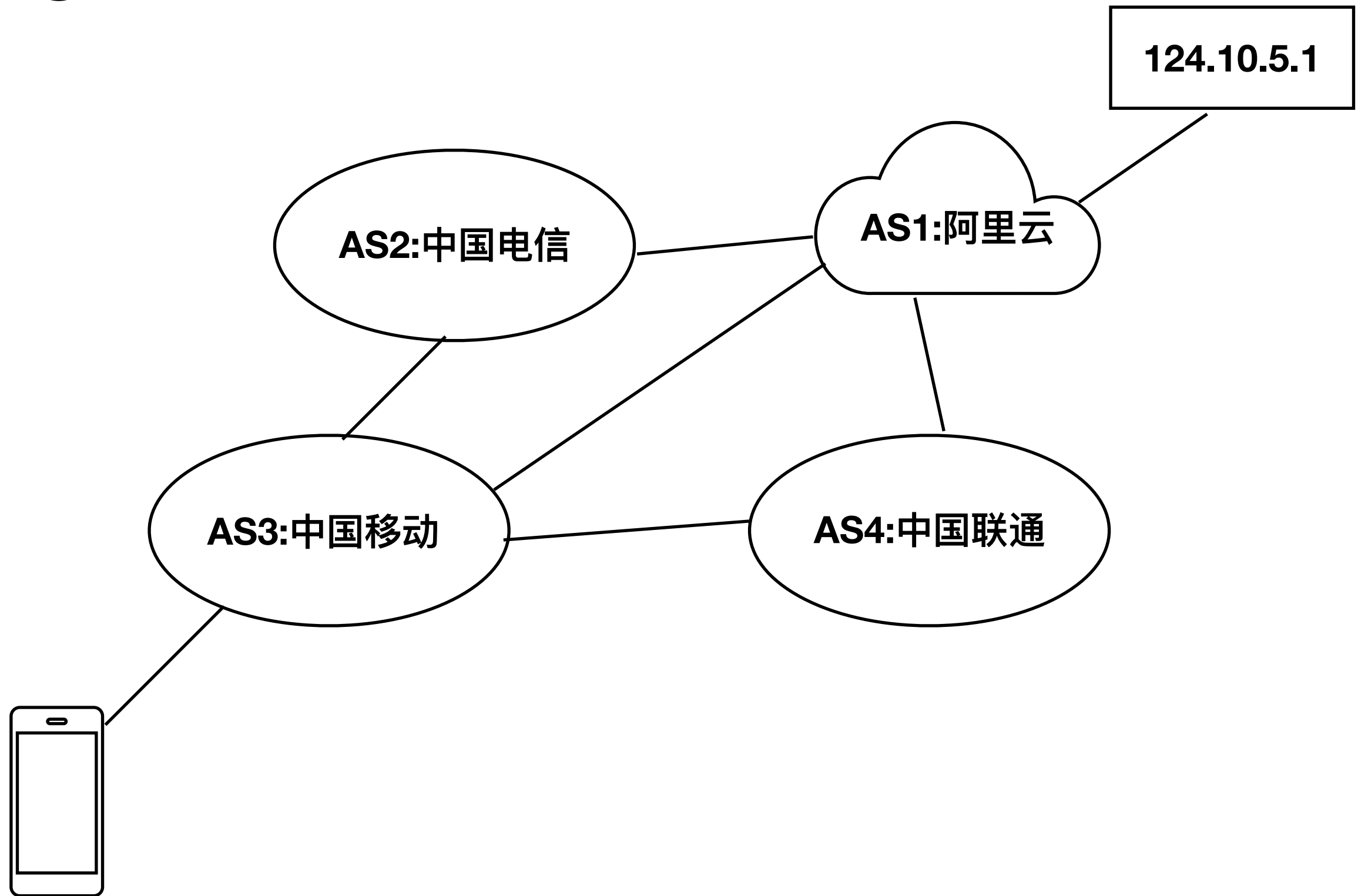
Route-refresh: 改变策略后请求对端重发路由信息

Notifycation: 错误时中断连接

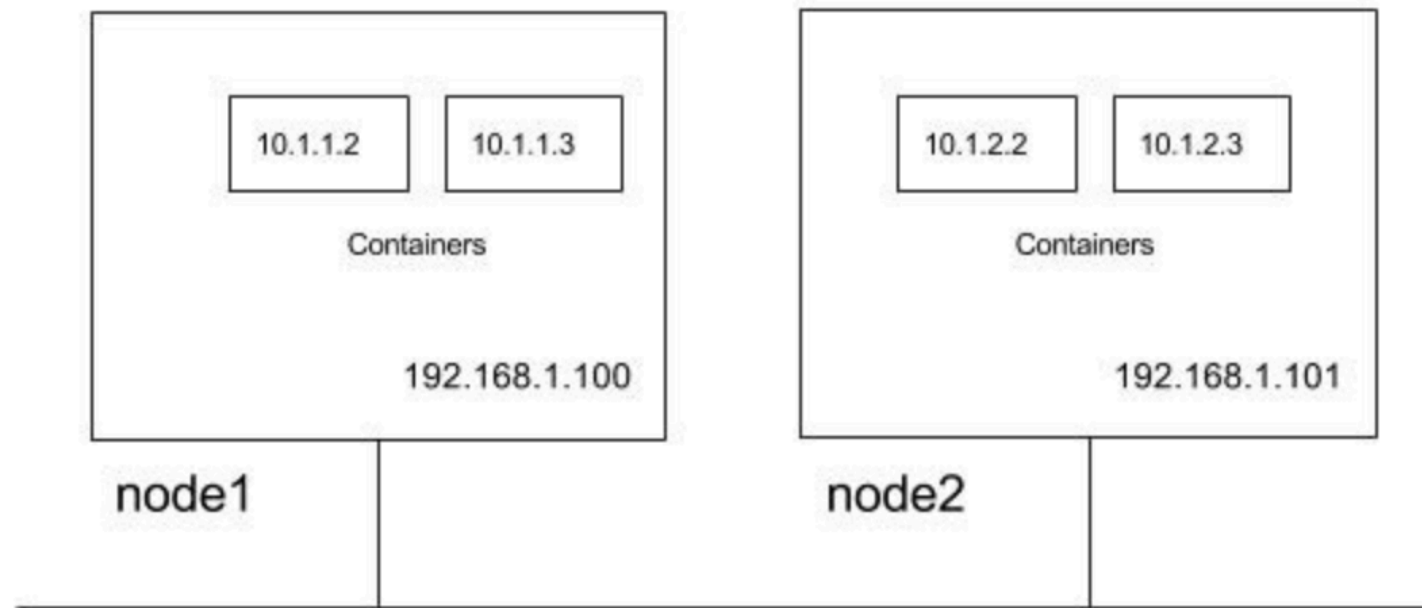
BGP状态机



BGP



K8s 网络

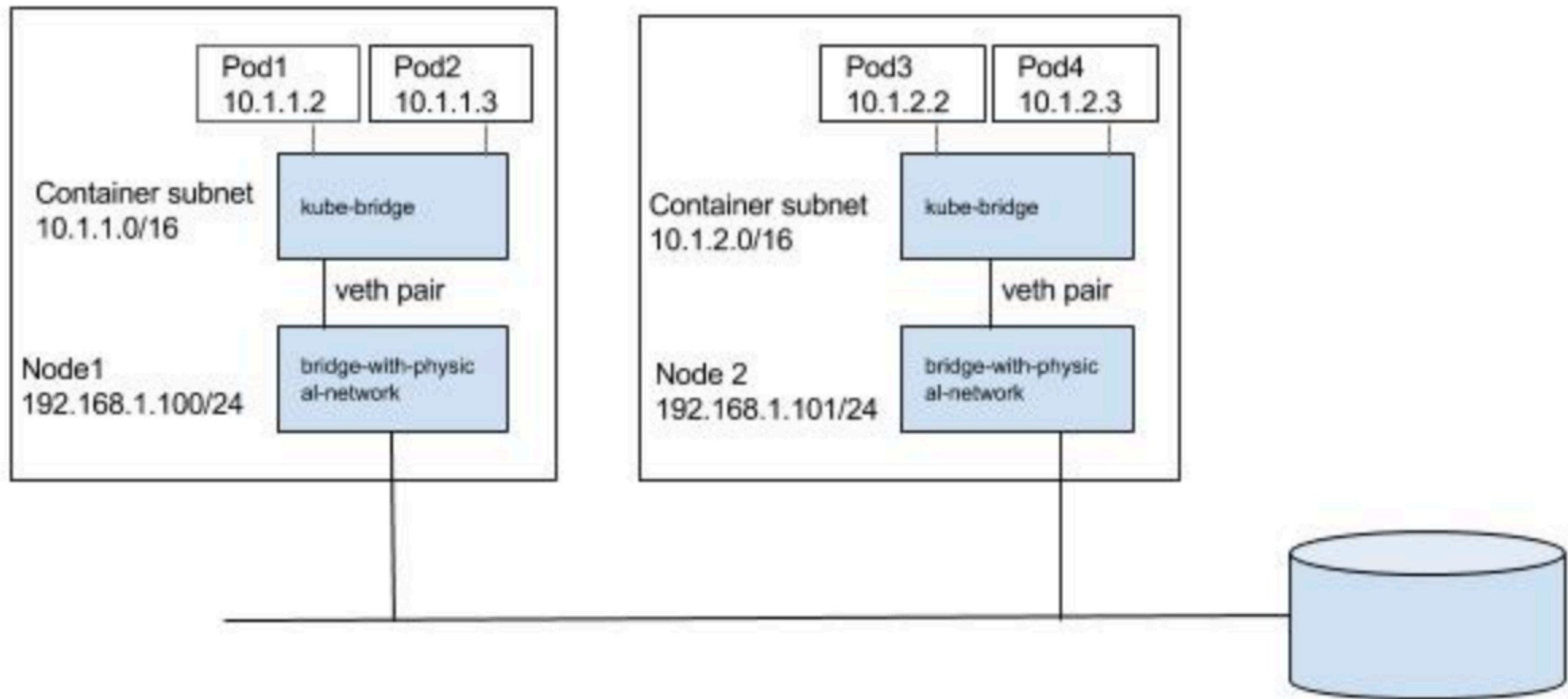


k8s 要求网络能够实现如下特性，且不使用 NAT：

- 1、pod 间可以通信**
- 2、node 可与任意 pod 通信，反之亦然**
- 3、一个 pod 看到的自己的 ip，和其他 pod 看到的相同**

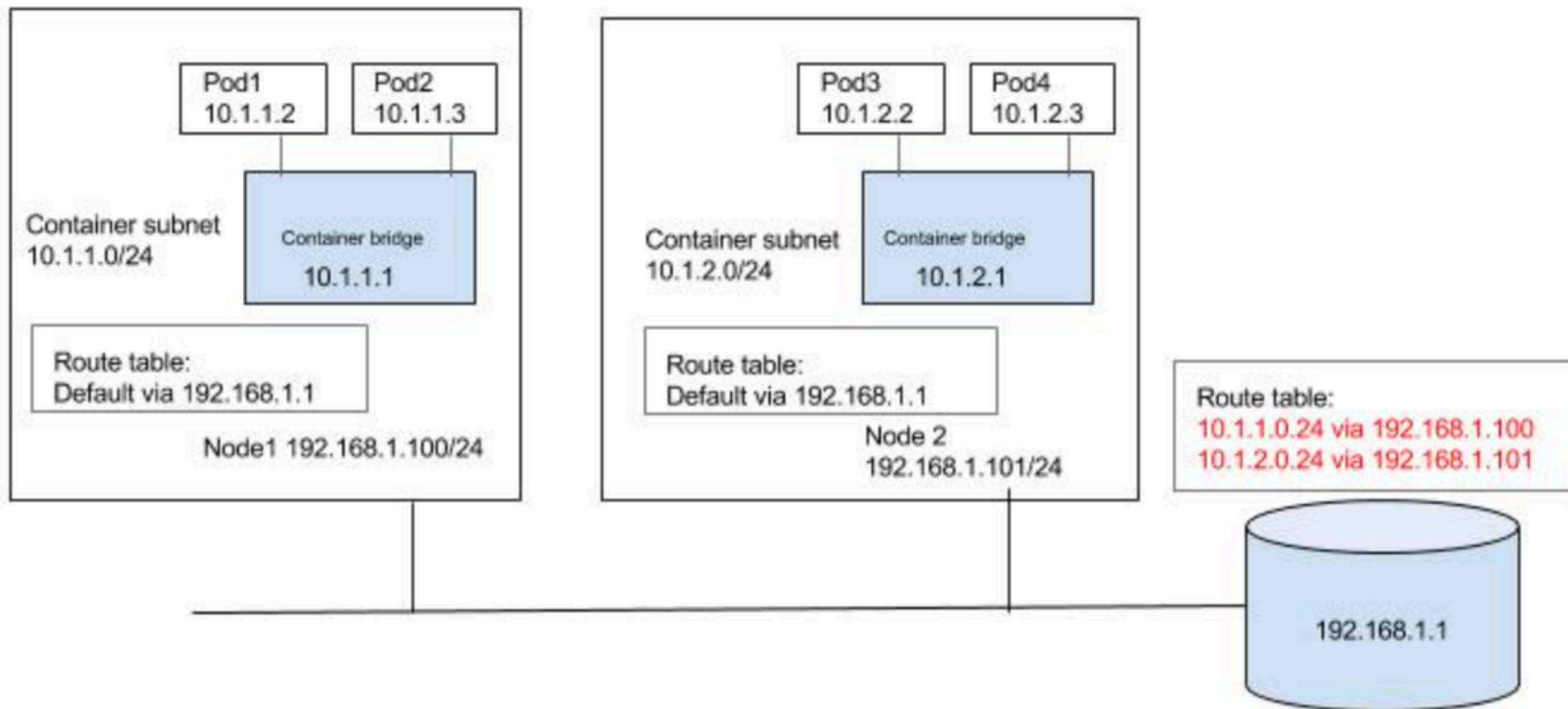
为什么 pod 需要观察到自己的 ip

L2 层实现

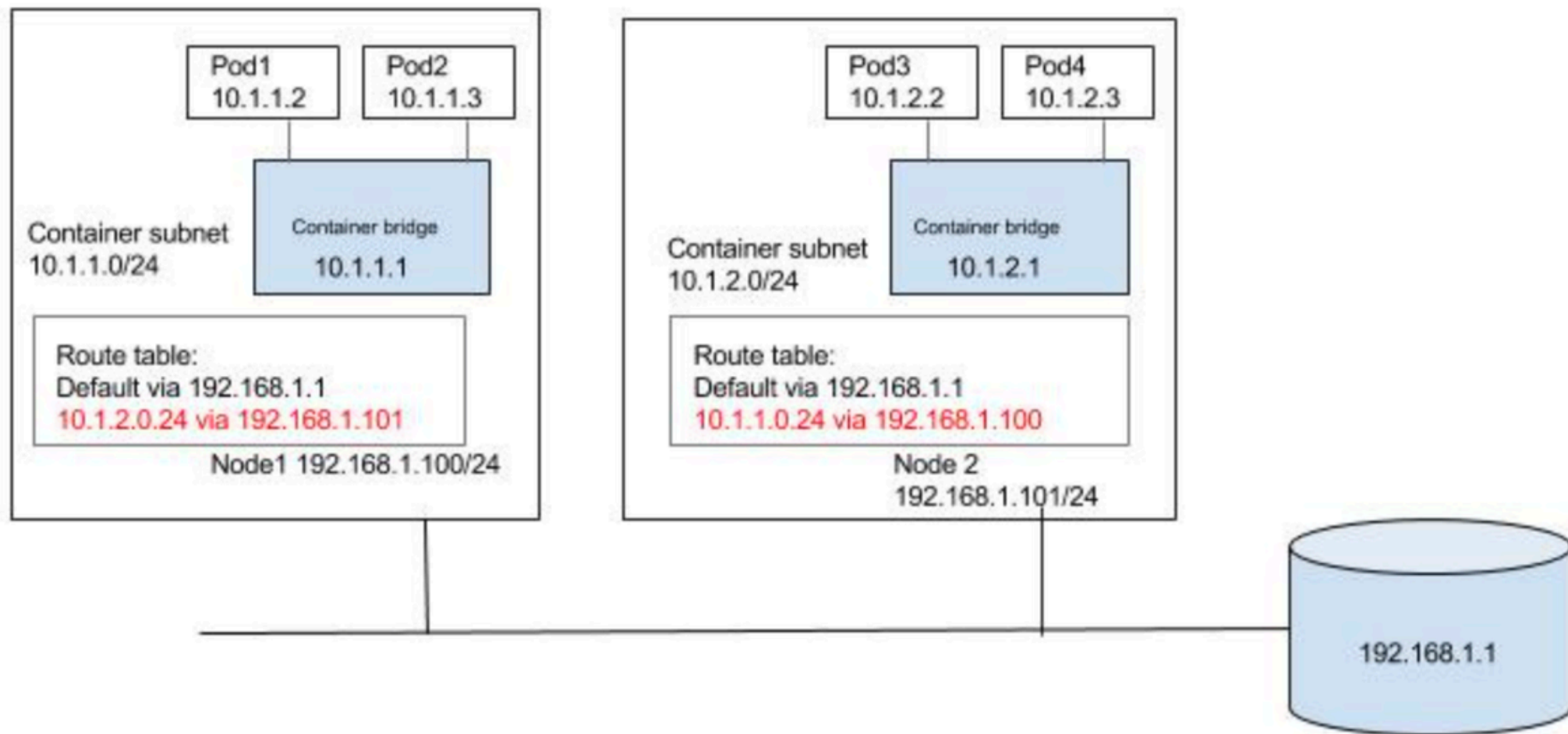


需要完善稿子

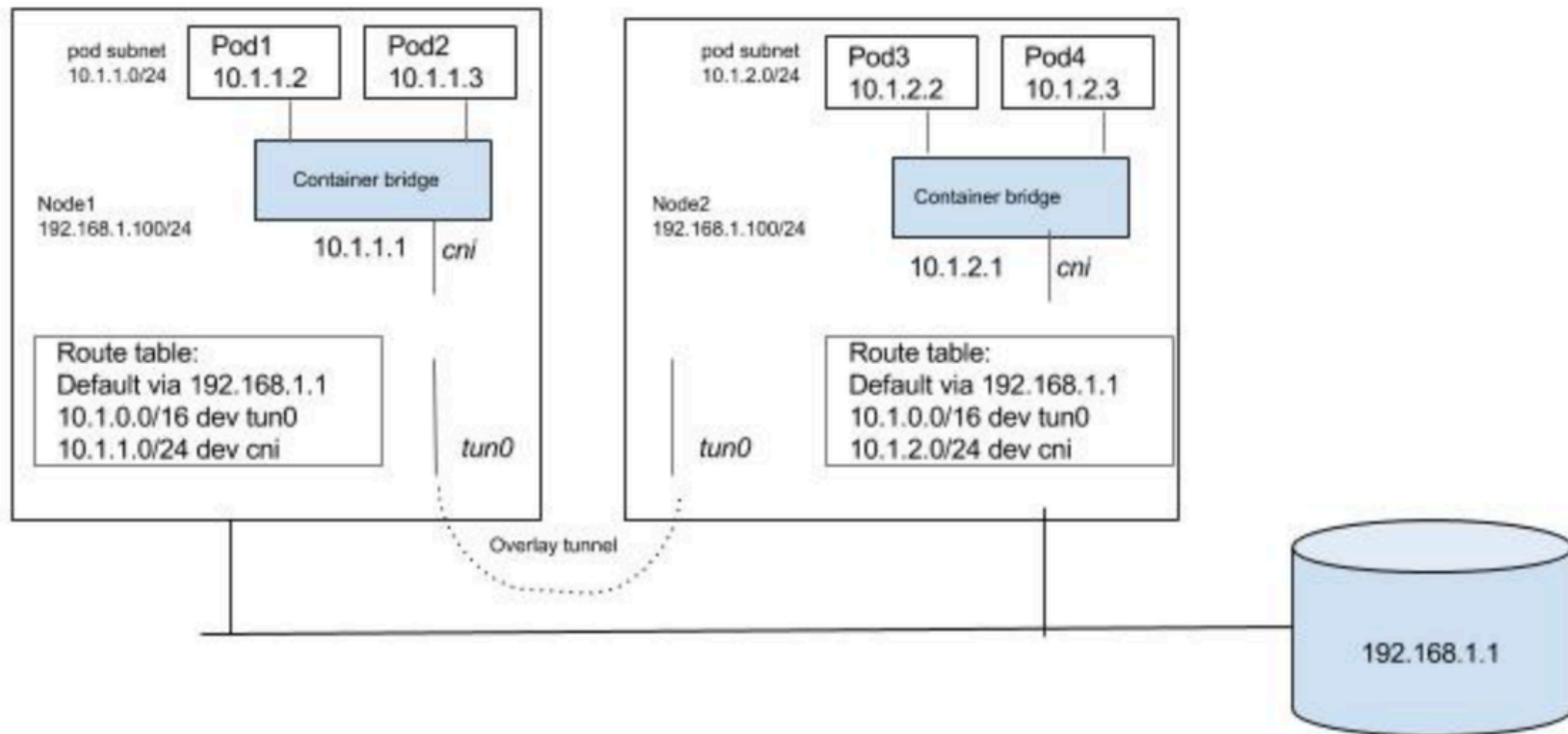
L3 层实现



L3 层实现



Overlay 实现

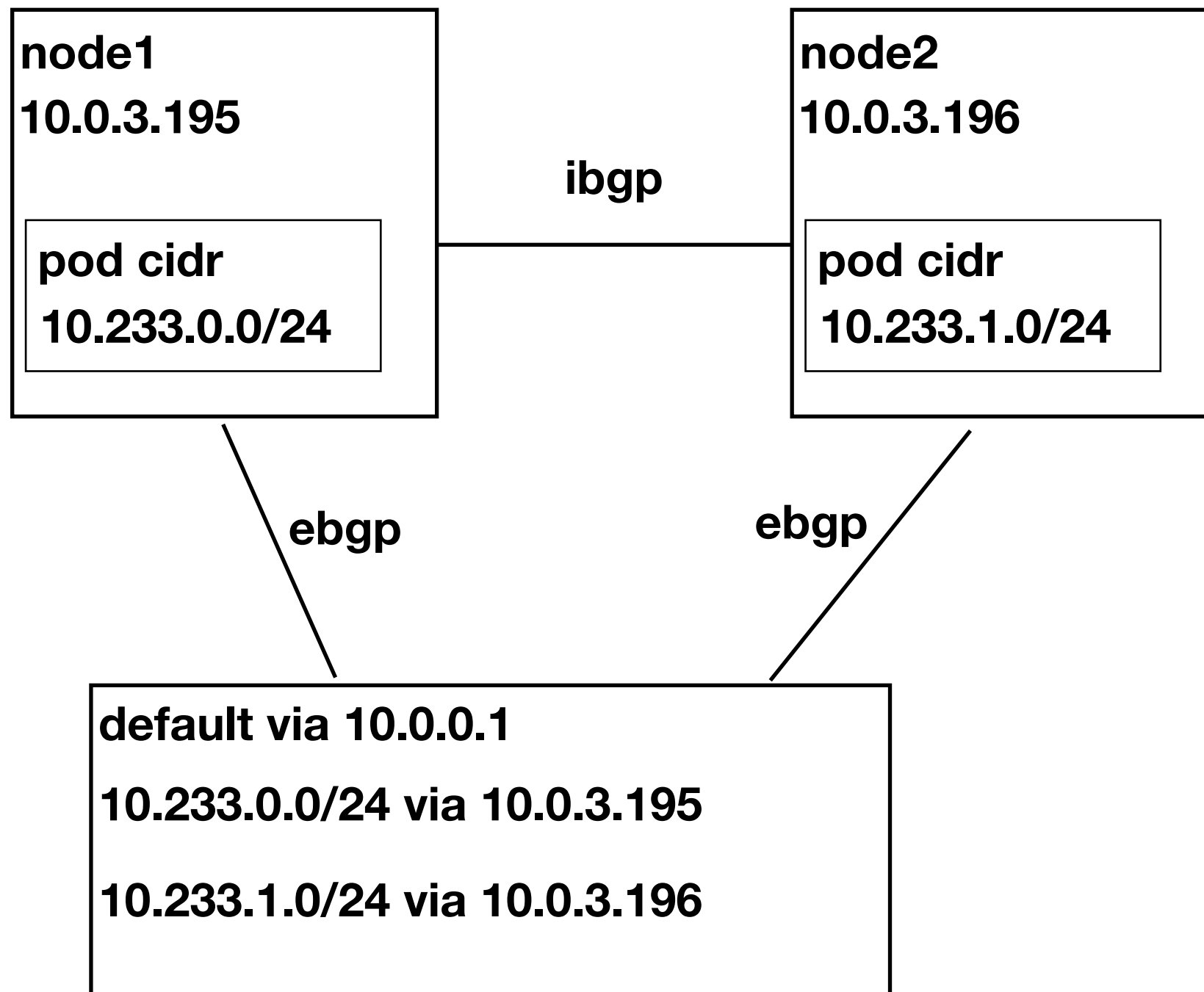


需要完善稿子

Kube-router

一个 k8s 网络解决方案(实现)

使用 BGP 协议与 GoBGP 实现了 pod 间网络



QA