KUBERNETES
COMMUNITY DAYS DALIAN

Kube-OVN：面向多租户网络安全的探索

灵雀云专家工程师

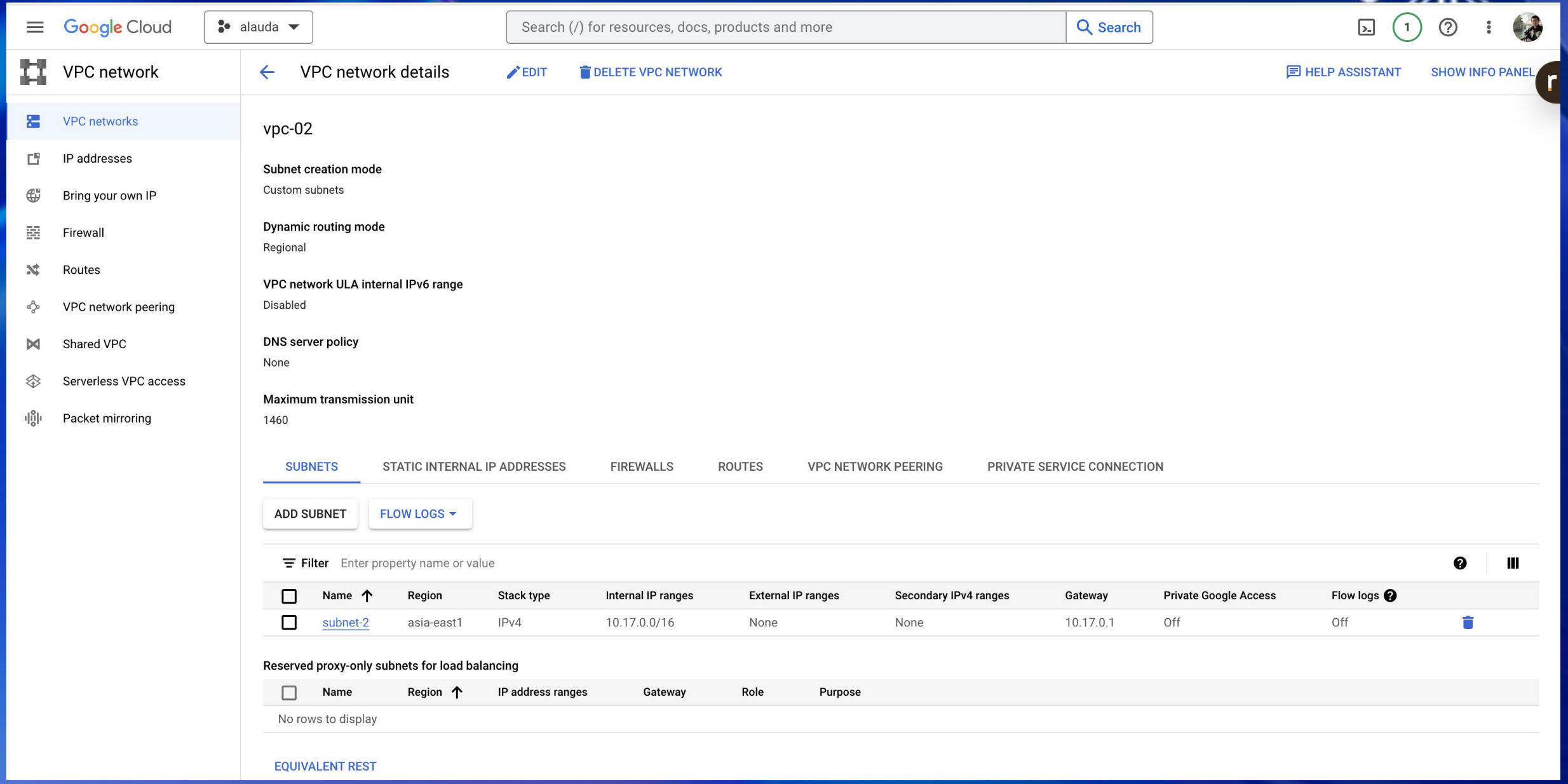
刘梦馨

Kube-OVN 是什么？

- 利用社区成熟的 OVS 作为网络底座
- 基于 Kubernetes 架构原生设计
- 结合企业侧多年实践

多租户网络是什么？

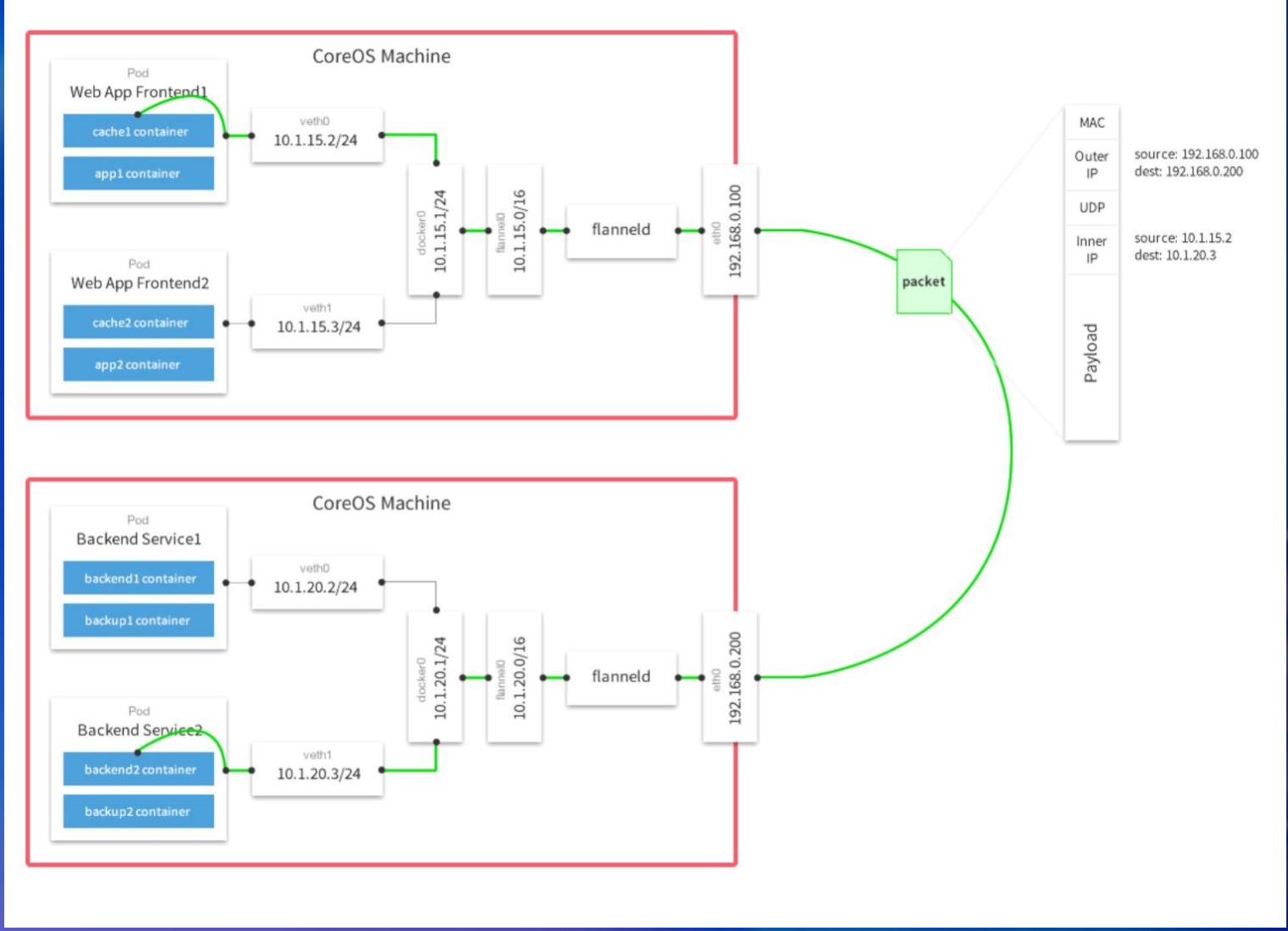
- 经典网络
  - 公有云早期网络形态
  - 传统网络架构，基于物理设备和拓扑结构
  - 所有租户2层互通，通过防火墙进行隔离
  - 难以扩展，管理和维护困难
  - 大部分公有云都已经淘汰该网络模型
- 多租户网络
  - 每个租户提供独立的虚拟网络
  - 每个租户拥有独立的网络拓扑、IP地址空间、安全策略等
  - 可以实现资源隔离、安全性和灵活性


Google Cloud
alauda
Search (/) for resources, docs, products and more
Search
VPC network
VPC networks
IP addresses
Bring your own IP
Firewall
Routes
VPC network peering
Shared VPC
Serverless VPC access
Packet mirroring
VPC network details
EDIT
DELETE VPC NETWORK
HELP ASSISTANT
SHOW INFO PANEL
vpc-02
Subnet creation mode
Custom subnets
Dynamic routing mode
Regional
VPC network ULA internal IPv6 range
Disabled
DNS server policy
None
Maximum transmission unit
1460
SUBNETS
STATIC INTERNAL IP ADDRESSES
FIREWALLS
ROUTES
VPC NETWORK PEERING
PRIVATE SERVICE CONNECTION
ADD SUBNET
FLOW LOGS
Filter
Enter property name or value
Name
Region
Stack type
Internal IP ranges
External IP ranges
Secondary IPv4 ranges
Gateway
Private Google Access
Flow logs
subnet-2
asia-east1
IPv4
10.17.0.0/16
None
None
10.17.0.1
Off
Off
Reserved proxy-only subnets for load balancing
Name
Region
IP address ranges
Gateway
Role
Purpose
No rows to display
EQUIVALENT REST

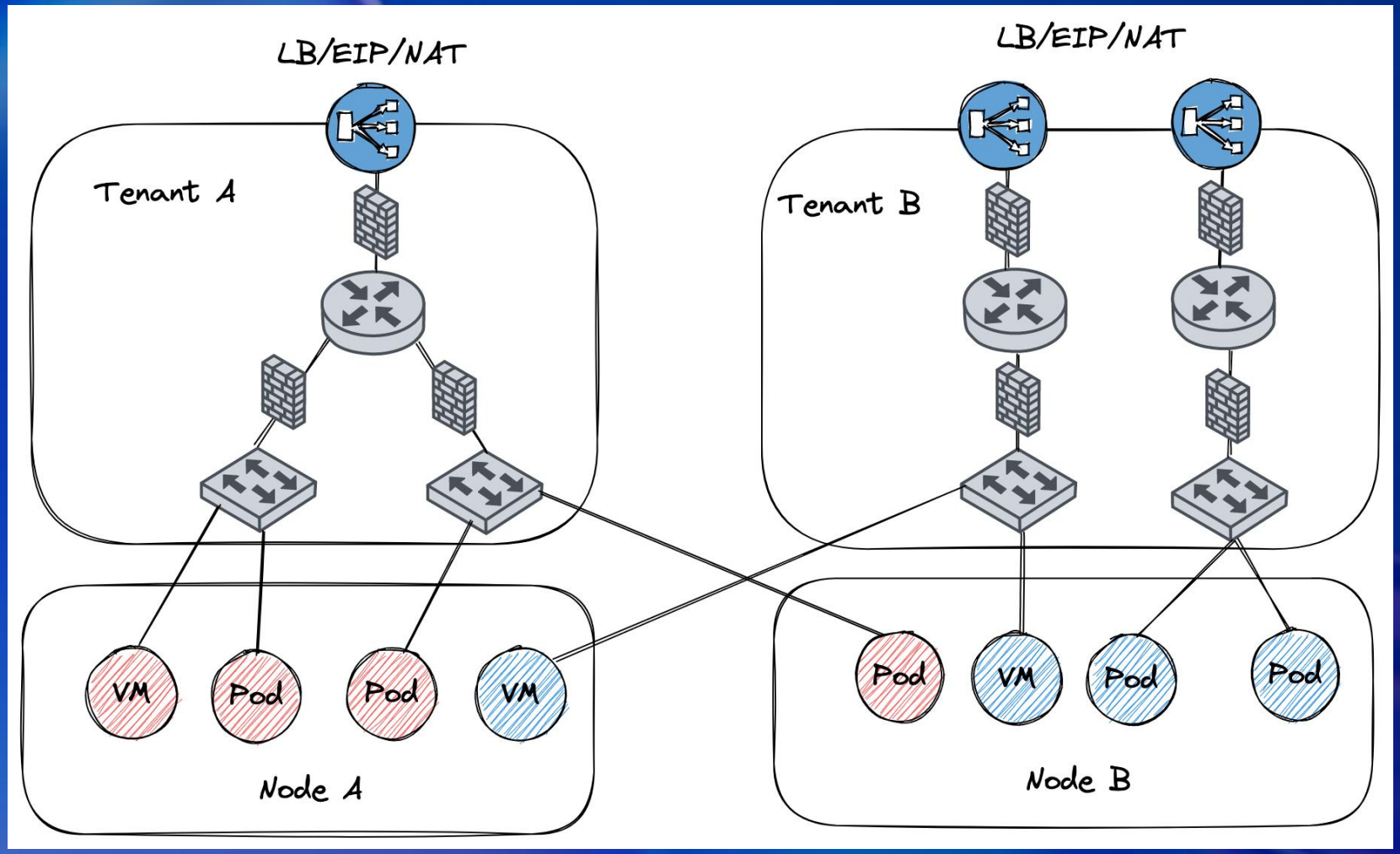
Kubernetes 为什么需要多租户网络？

- 面向应用的 PaaS 平台并没有强烈的多租户网络需求
  - 基础的容器网络互通
  - 使用kube-proxy提供的服务发现能力
  - NetworkPolicy做网络隔离和安全
  - Flannel,Calico,Cilium ...
- 用户开始使用 Kubernetes 管理基础设施
  - 使用Kubernetes + Kubevirt 替代原先的 IaaS
  - 同时管理容器、虚机、裸金属资源
  - 构建公有云或集团云对外提供服务
  - 多租户网络是公有云服务的基本能力和使用标准

Kube-OVN多租户网络架构



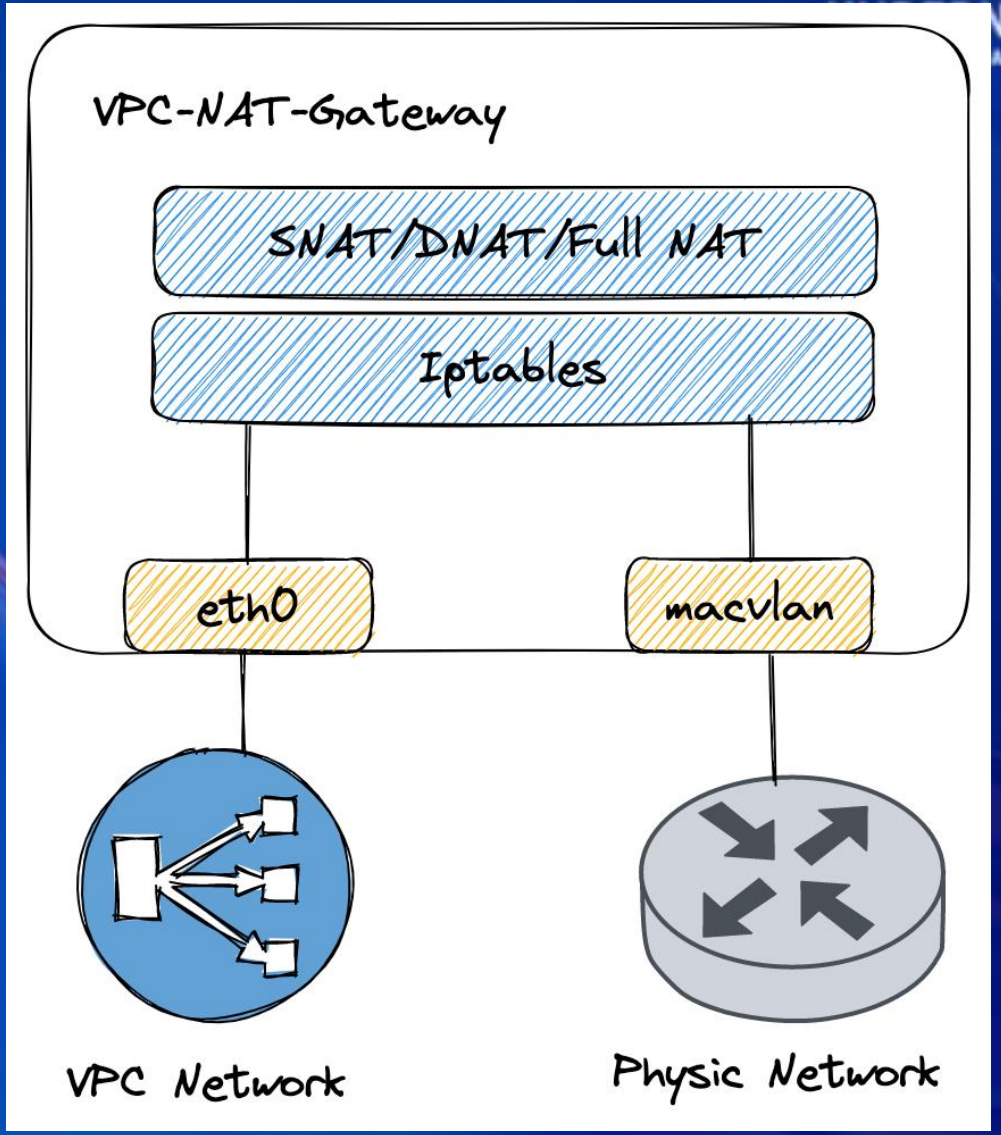
Kube-OVN多租户网络架构



Kube-OVN多租户网关设计

- 多租户VPC内的Pod如何访问外部网络？

- 每个VPC独立的出口
- 双网卡Pod，一个接入虚拟网络，一个接入物理网络
- 设置VPC内路由，访问外网流量路由到 vpc-nat-gateway
- 通过Iptables进行SNAT/DNAT/Full Nat 等操作



多租户安全设计

- 租户间网络逻辑上隔离
  - 可使用相同的网段、IP
  - 租户独立出网网关，出网流量隔离
- 安全组设计
  - NetworkPolicy 缺少优先级，黑名单等功能组合较为困难
  - 公有云 Security Group 接口更符合用户使用习惯
  - 通过CRD映射 Security Group 和 OVN 内的 ACL

```yaml
apiVersion: kubeovn.io/v1
kind: SecurityGroup
metadata:
  name: sg1
spec:
  allowSameGroupTraffic: true
  egressRules:
    - ipVersion: ipv4
      policy: allow
      priority: 1
      protocol: all
      remoteAddress: 0.0.0.0/0
      remoteType: address
  ingressRules:
    - ipVersion: ipv4
      policy: allow
      priority: 10
      protocol: icmp
      remoteAddress: 0.0.0.0/0
      remoteType: address
    - ipVersion: ipv4
      policy: allow
      priority: 10
      protocol: tcp
      portRangeMin: 3306
      portRangeMax: 3306
      remoteAddress: 192.168.2.100
      remoteType: address
    - ipVersion: ipv4
      policy: allow
      priority: 10
      protocol: all
      remoteSecurityGroup: sg2
      remoteType: securityGroup
```

用户自定义负载均衡

- Servie ClusterIP 在多租户下的问题
  - ClusterIP 实现了集群内负载均衡的功能
  - 整个集群共享一个 Service CIDR
  - 用户无法自定义负载均衡的地址

- 使用新的 CRD 来实现多租户下的负载均衡
  - 创建 Headless Service，分配 VIP 给 Service
  - 复用Endpoint机制
  - 调用 OVN L2 LB 能力实现内部负载均衡

多租户的DNS

- 现有CoreDNS 在多租户下的问题
  - CoreDNS部署在系统VPC下，和用户VPC默认无法联通
  - CoreDNS默认记录了所有Service的解析，无法做到租户隔离

- 解决方案
  - 租户内自建 DNS 服务
  - 默认 VPC DNS 通过负载均衡对外暴露
  - 租户内DNS网络和默认VPC打通

实际应用案例

基于Kube-OVN的网络方案

多租户容器网络

- 提供基于VPC的多租户网络，具备更好的可扩展性以及更强的隔离性和安全性。
- 可应用于公有云、边缘计算、金融业务等对隔离性和安全敏感的场景

跨集群容器网络

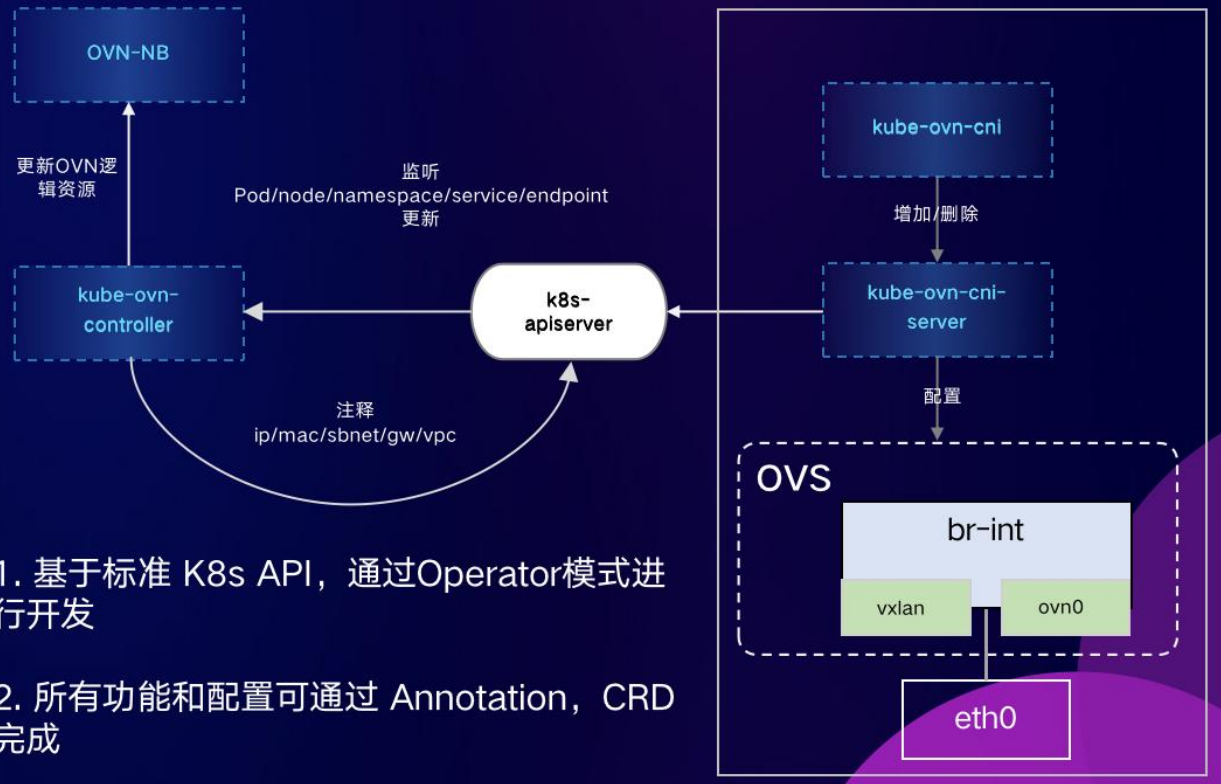
- 基于隧道的多集群网络互通方案，可提供跨集群的容器IP直接互访
- 多集群API网关的功能，支持对多集群流量进行更为灵活的调度和控制

硬件加速

- 通过硬件能力加速容器网络的处理能力，极大提升了容器网络的性能
- 提供了OVS-DPDK的集成，DPDK应用可以获得高性能网络处理的能力

智能运维

- 自带自动化运维工具，并提供了100+监控项
- 全面监控主机、容器、服务之间网络质量以及稳定性指标，保证企业网络的稳定运行。



1. 基于标准 K8s API，通过Operator模式进行开发

2. 所有功能和配置可通过 Annotation，CRD 完成

未来的发展

- 智能网卡管理裸金属资源网络
- 云上云下环境打通
- 更多云网络资源引入VPN，QoS，计费
- 更大规模集群支撑

KUBERNETES
COMMUNITY DAYS DALIAN


KUBERNETES
COMMUNITY DAYS DALIAN
Q&A