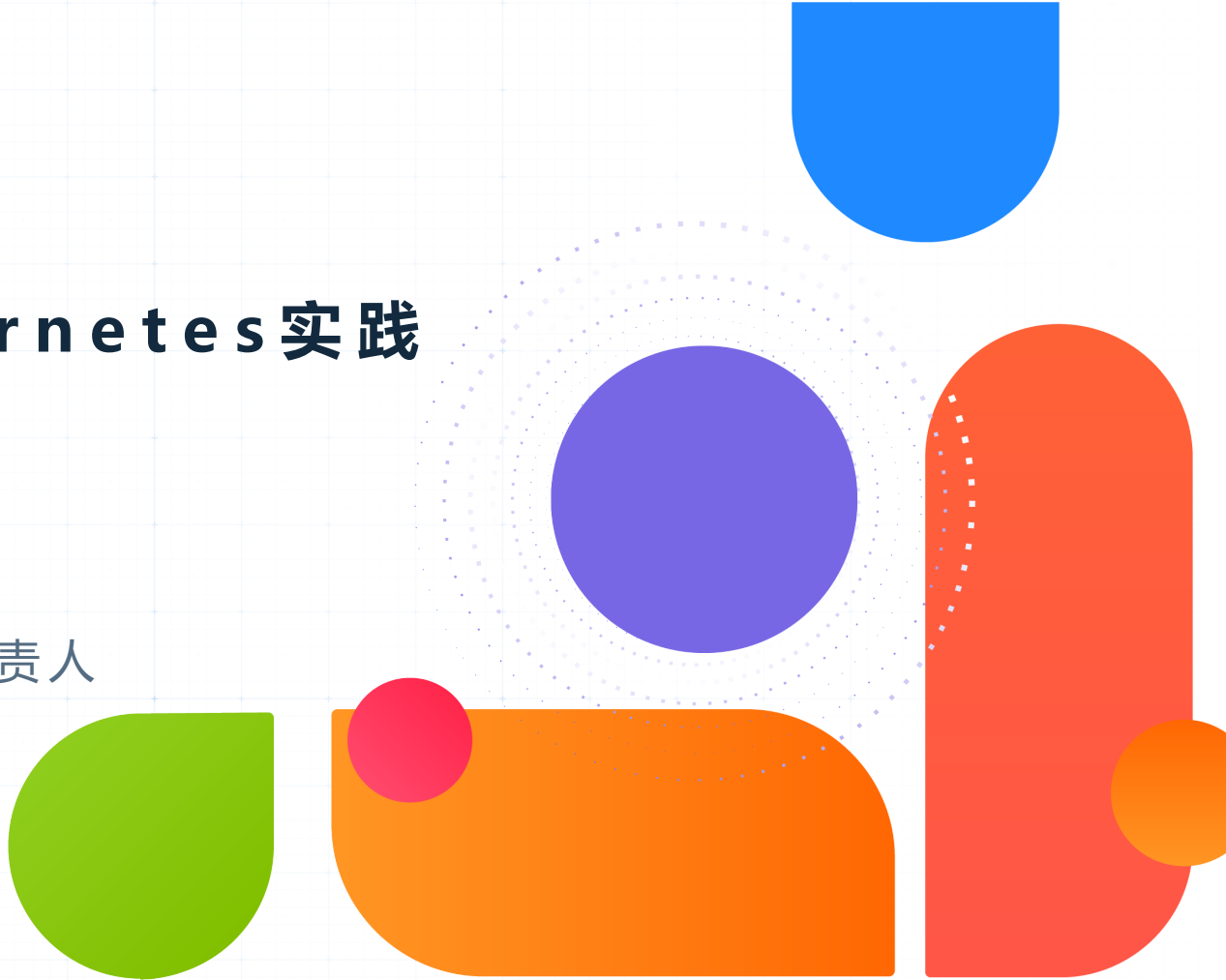


企业级 kubernetes 实践

- 安雪艳 -

优刻得KUN平台负责人



1. 基于Kubernetes的技术平台-KUN
2. KUN平台方案
3. Kubernetes中间组件管理方案
4. 优刻得云原生产品

KUN-支撑研发全生命周期管理和运行资源

23 地域

300+ 服务器

15211 Services

18659 Pods

管理区资源利用率显著提升，产品开发测试上线节奏加快

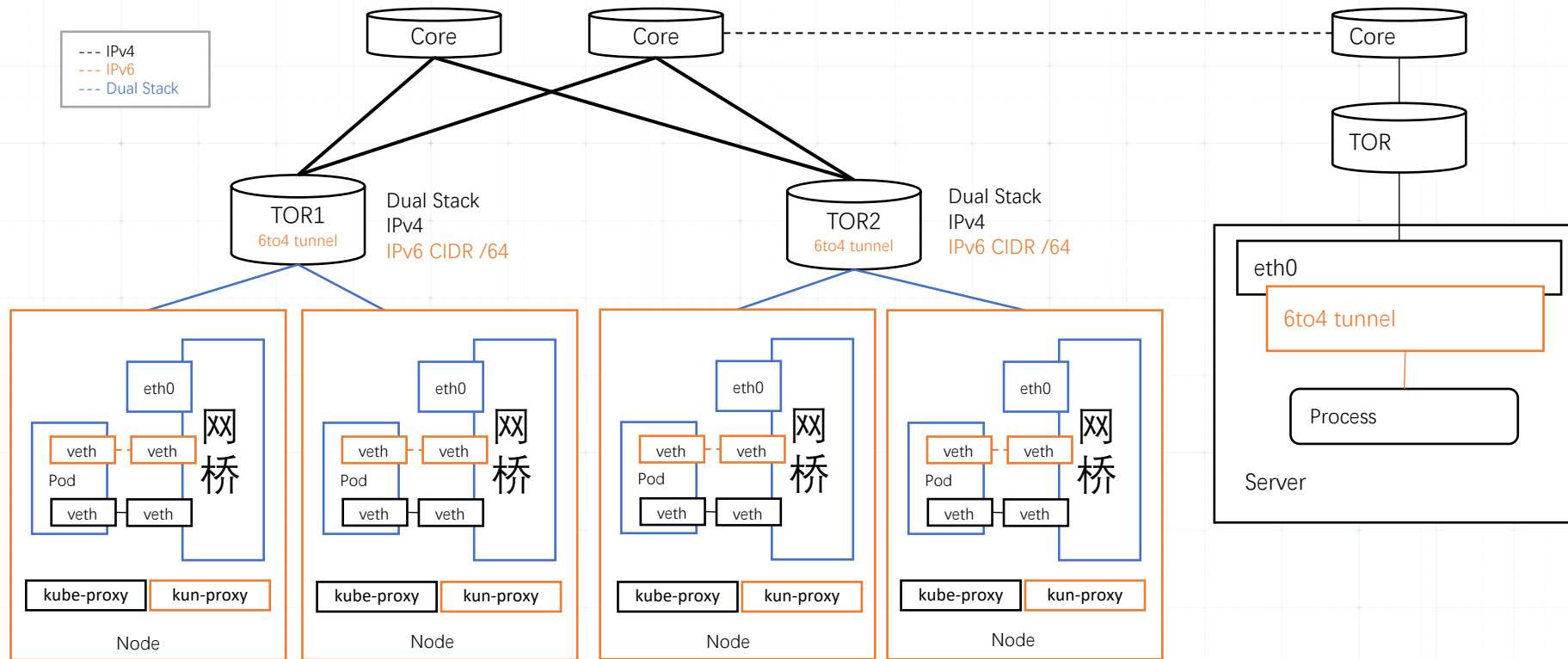
KUN技术平台相关工作

- 自研双栈
- IP冷却
- 自研存储组件
- CI/CD开发
- 组件operator开发
- 高可用监控方案
- 带缓存日志系统
-

网络需求

- Kubernetes 集群内部的 Pod 和 Service 需要和集群外部互通
- 需要访问外网
- IPv4 地址不够用

网络-Pod



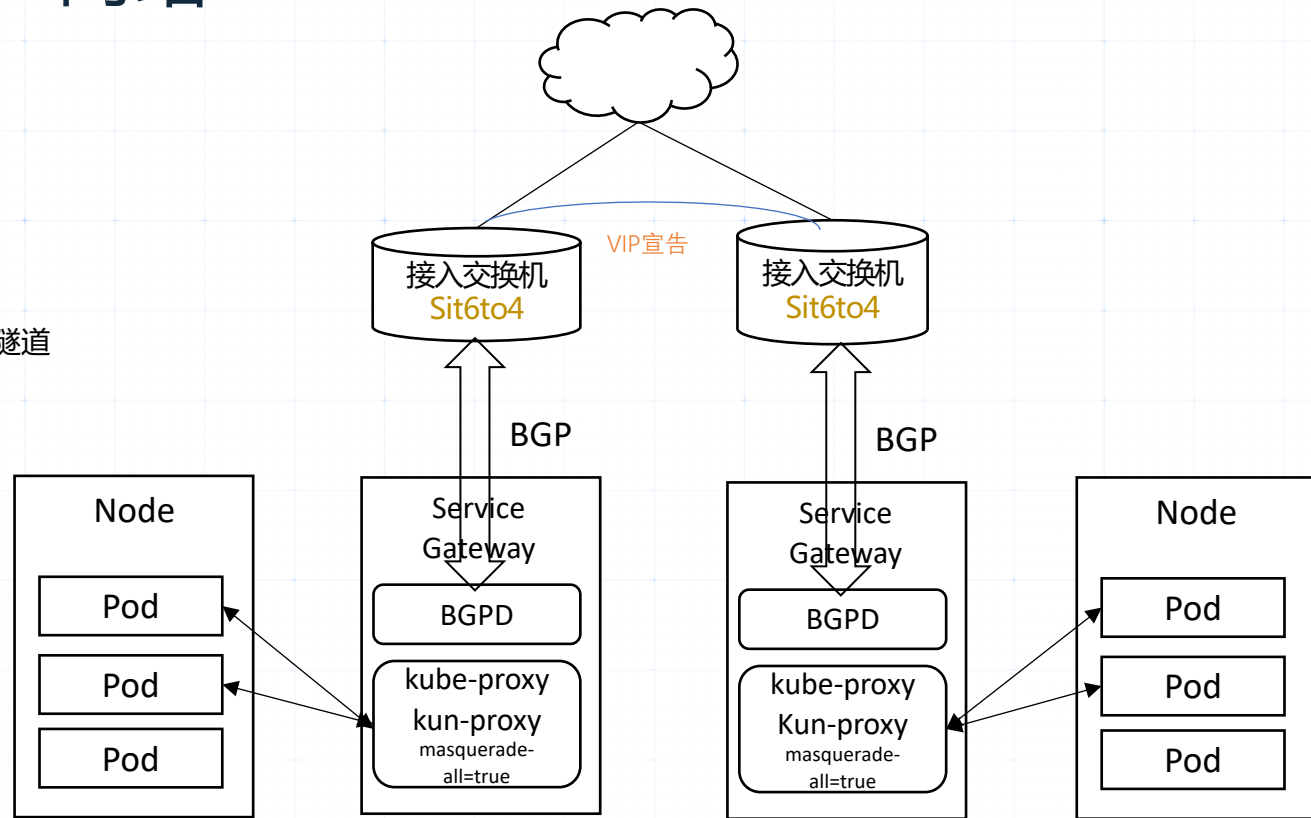
网络-Service

方案1.0版本

- 多台网关节点，高可用
- 交换机VIP宣告，配置sit6to4隧道
- IPv4 网关上BGP全网广播

特点

- 三层转发
- 无需ingress控制器



痛点：TOR 需要特殊配置支持 IPv6 隧道，扩容、运维成本较大

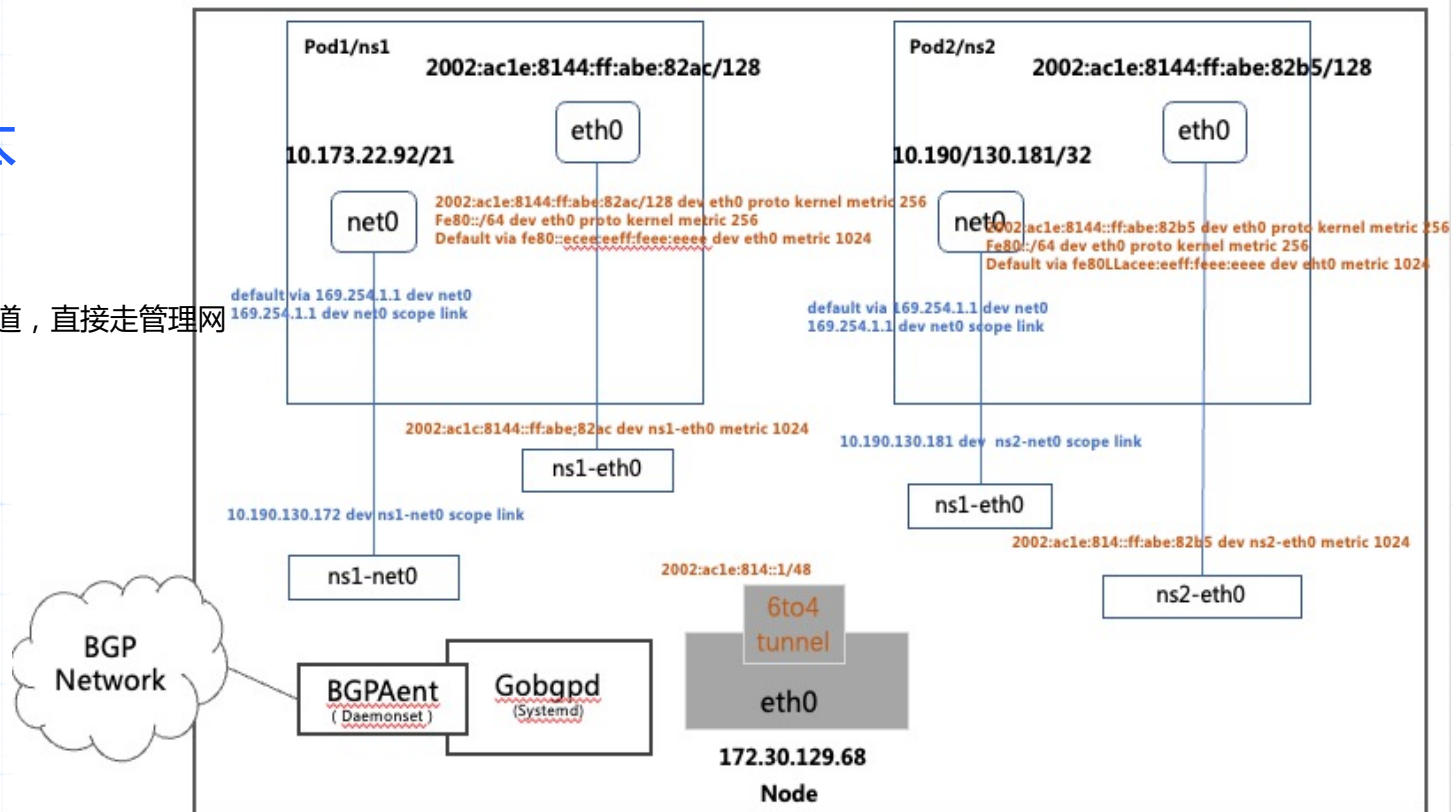
网络-Pod

方案2.0版本

- Pod到node路由模式
- IPv4 通过BGP全网广播
- IPv6 通过节点sit6to4隧道，直接走管理网

特点

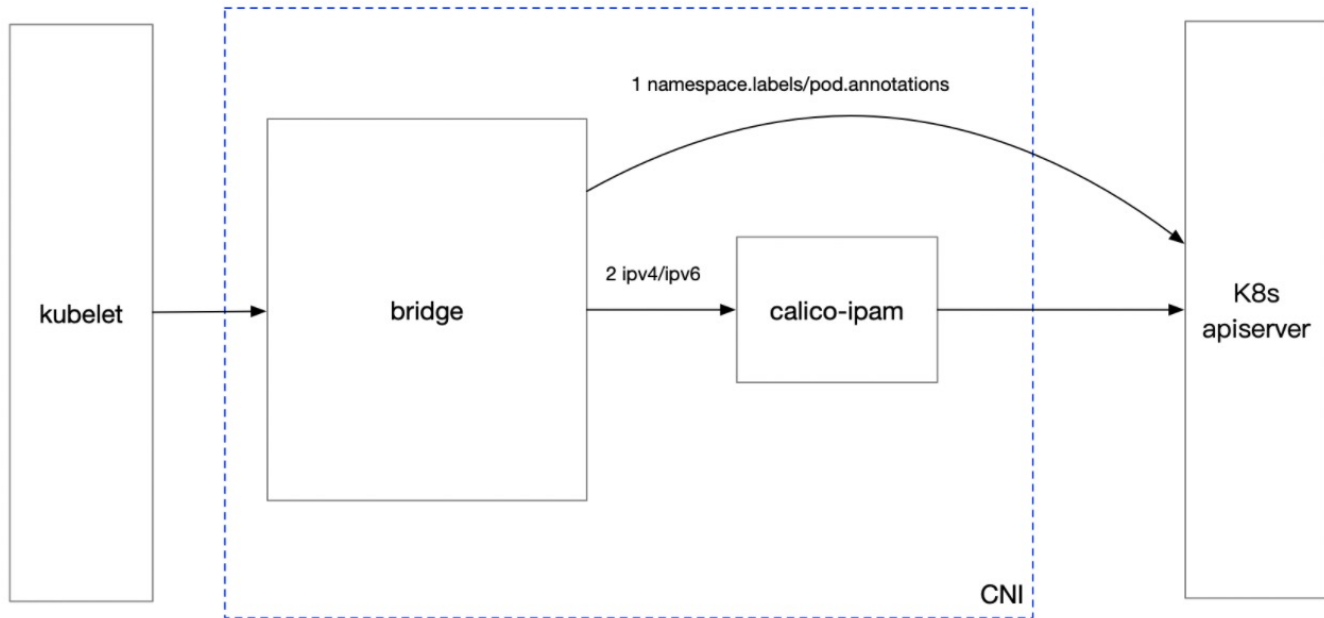
- 更加灵活
- 资源不受限



网络-按需分配

设置关键字，按需分配：

```
cni.networking.kun/ipv4Enabled: false
```



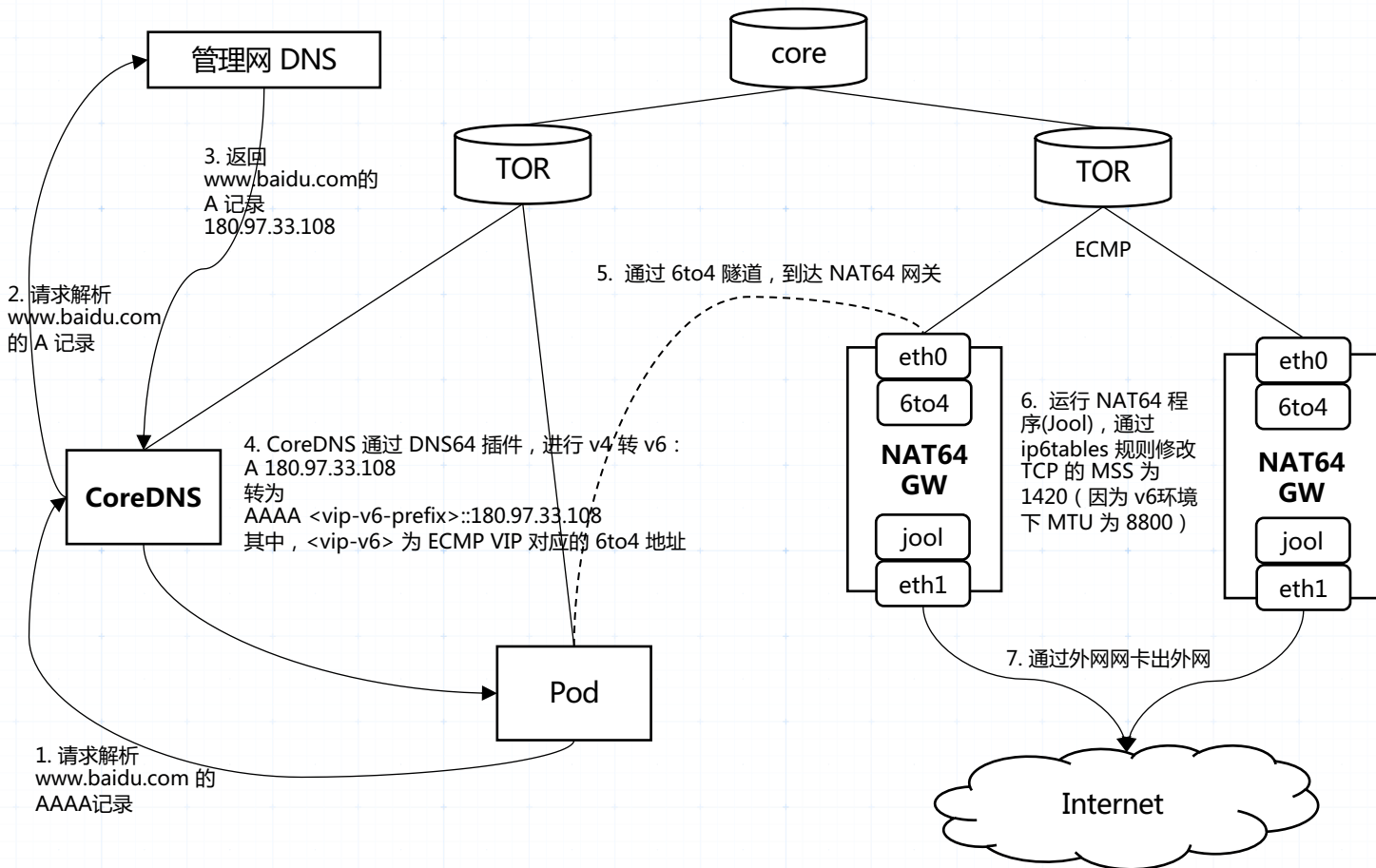
网络-外网

方案

- IPv6协议栈
- DNS+nat64
- ECMP (BGP)

特性

- ECMP高可用
- 科学上网



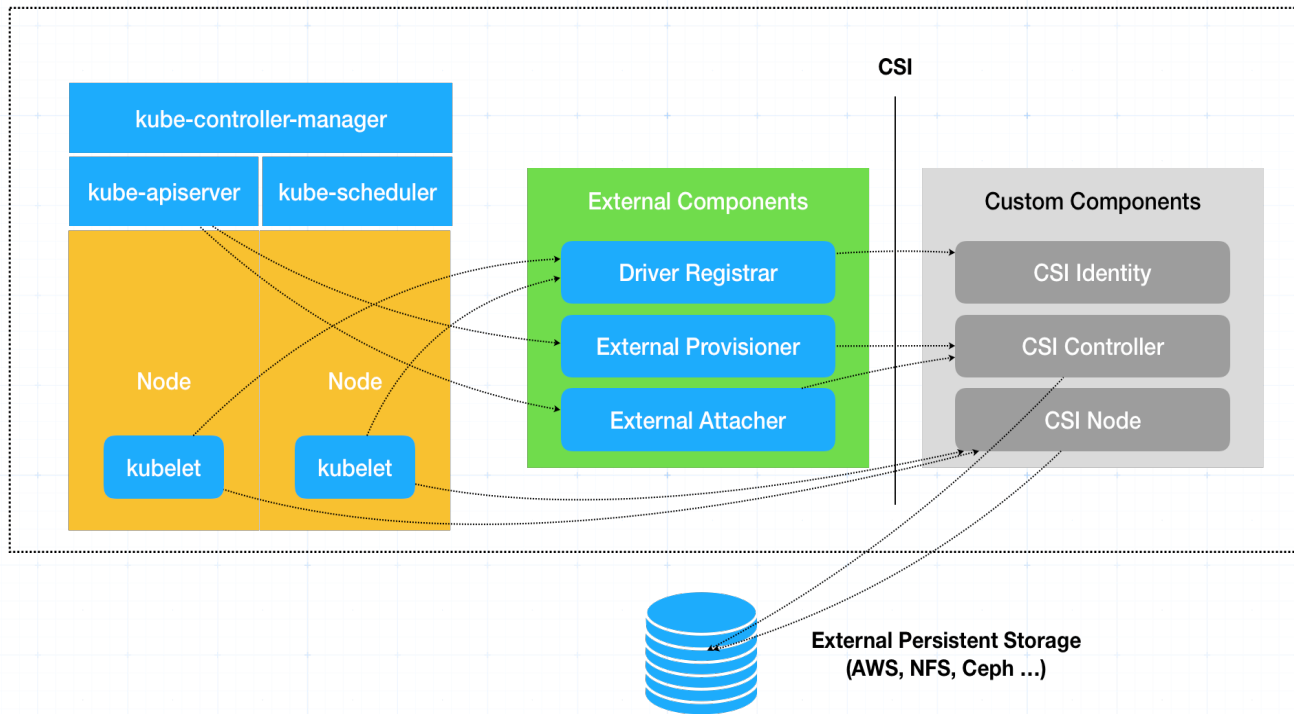
网络改造

- IPAM 可能会出现 IP 短时间内重复利用的问题，我们做了修改，增加了 IP “冷却时间”
- 每个 Pod 可创建 2 对 Veth Pair，分别用于 IPv6 和 IPv4，并且做了映射，IPv6 地址的最后 32 位即为 IPv4 地址
- 使用 gobgpd 进行 BGP 通信，开发batman用于管理BGP路由
- 增加了 Pod Annotation 开关控制是否需要 IPv4
- 我们开发了 kun-proxy (基于 IPVS)、kunservice (CRD) 以及 kunservice operator，和 kube-proxy 以及 service一起工作，分别支持 IPv4 和 IPv6。同时实现 AZ 亲和性。
- 开发了 CoreDNS 插件提供 6to4、NAT64、科学上网的功能

存储

CSI (Container Storage Interface)

容器编排引擎和存储
系统间一套标准的存
储接口定义



存储

块

Ceph on Rook 或 Ustore
云硬盘 UDisk

文 件

UFS(NFS)
自研基于IPVS的NFS高
可用客户端

对 象

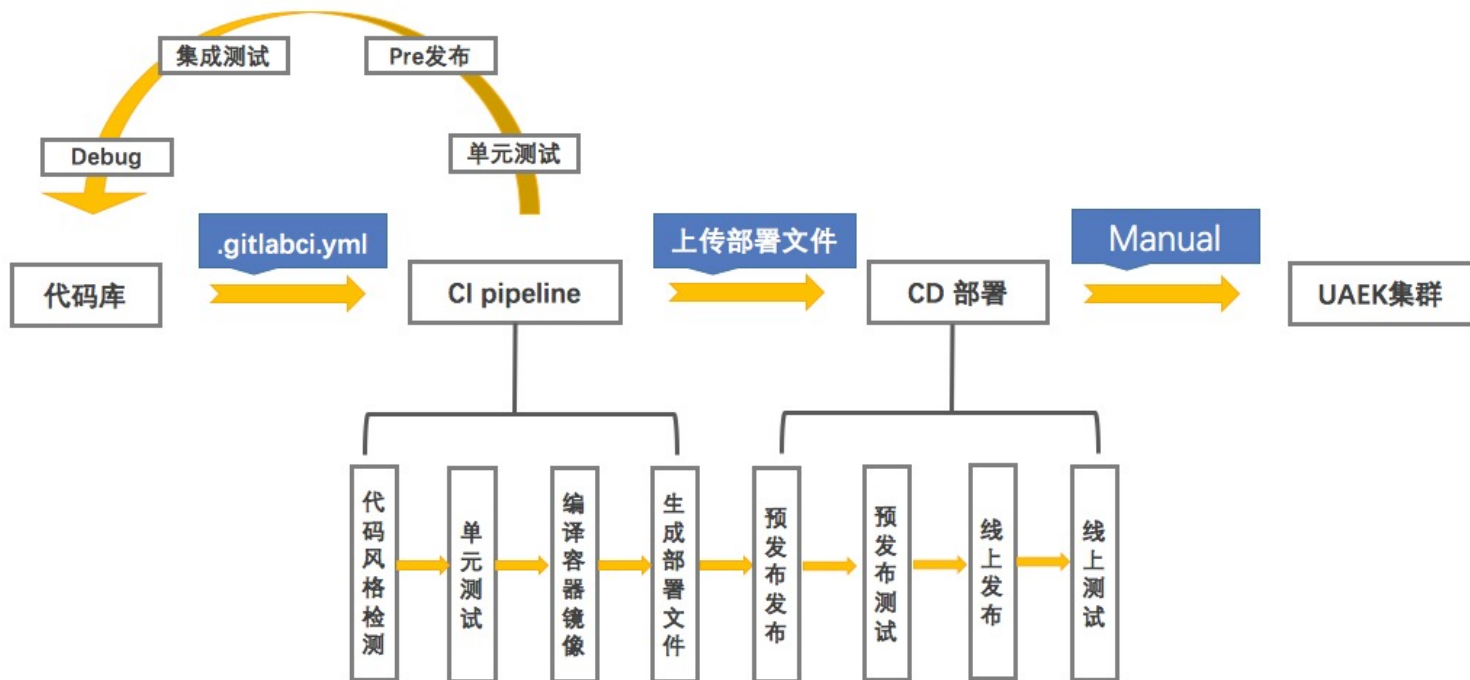
US3(s3fs mount)

CI/CD

Gitlab Runner

+

Kubernetes



CI/CD

- 改造 Gitlab Runner :
 - 接入 IAM 进行鉴权认证 :
 - 虚拟用户
 - 增加缓存机制
- 使用 Kaniko 进行 Docker 镜像构建 + 缓存机制
- 自研CD系统 , 或者 Argo CD

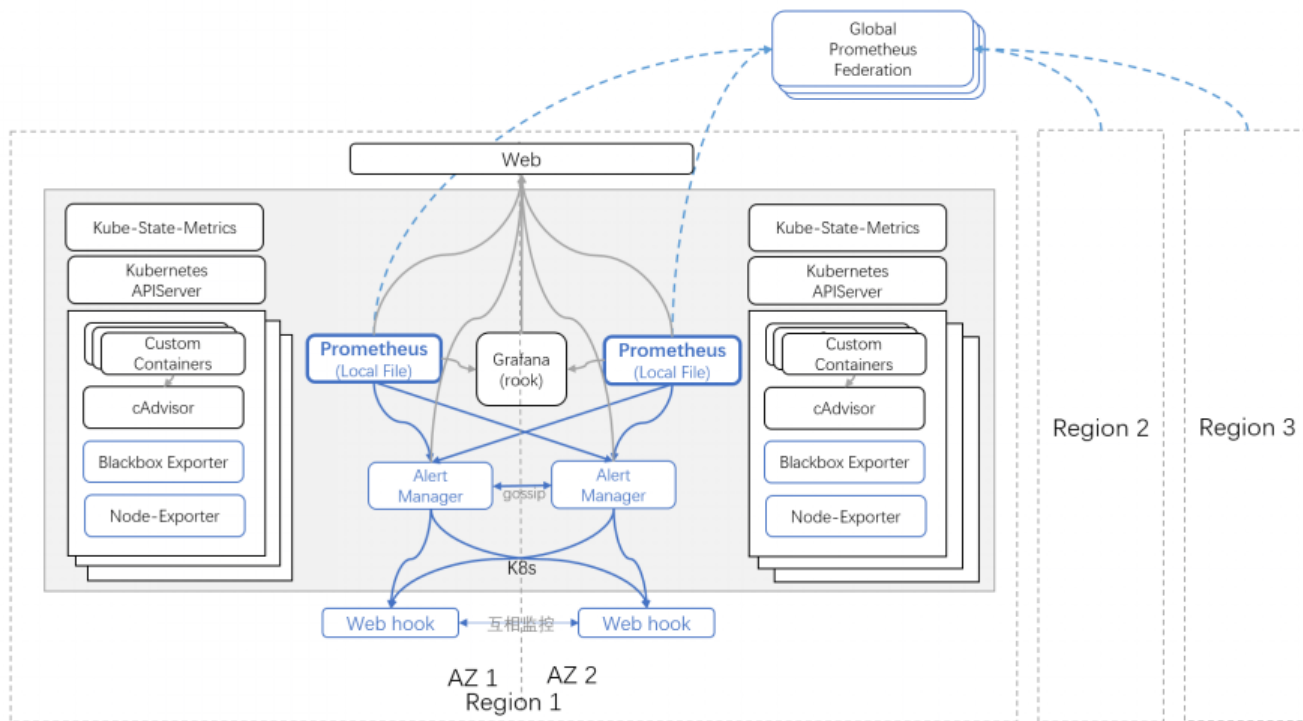
监控

公共监控高可用

- 单集群：多实例 + 本地存储
- 多集群thanos + grafana展示
- 自研monitor+manager：探活监控组件 + 自杀告警
- 接入消息助手：支持noc + 邮件 + 微信告警

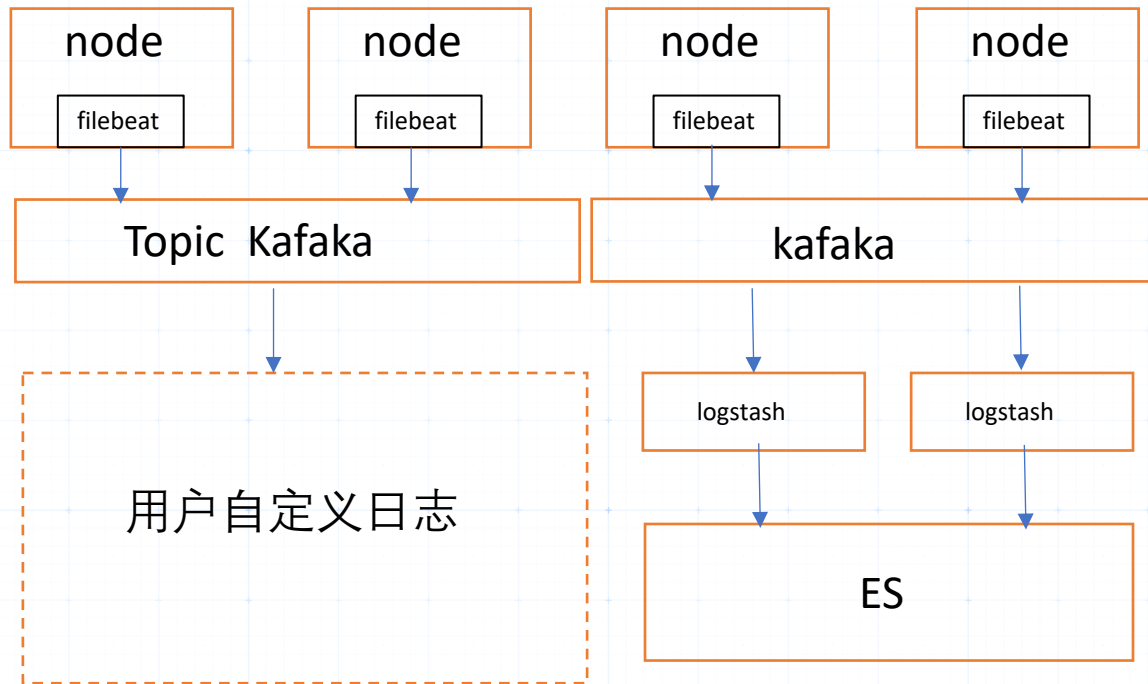
业务定制监控

- 自研Export
- 主动探测业务层级



日志

- Filebeat + Kafka + Logstash + ES
- 每个 Node 部署 Filebeat
- 两套 Kafka
 - 公共 Kafka
 - 以 Namespace 分 Topic Kafka, 用于给用户消费, 自己解析日志



公告组件

Helm Chart + Operator

自研

Redis Sentinel Cluster Operator

Redis Cluster Operator

Grafana Operator

灰度发布 Operator

KUN Service Operator

BGP 相关 Operator

开源

Prometheus Operator

Etcd Operator

Argo Workflow, Argo Events

NATS Operator

总结

网络

BGP+VIP , 云原生方案 (Calico , Flannet)

存储

ROOK , 私有云存储Ustore

CI/CD

Gitlab Runner + Argo CD

日志

ELK

监控

Prometheus + Thanos + Grafana

公告组件

Helm + Operator



UK8S
公有云



Cube
Serverless



优钛云
私有云

THANKS!



添加小优，加你入群