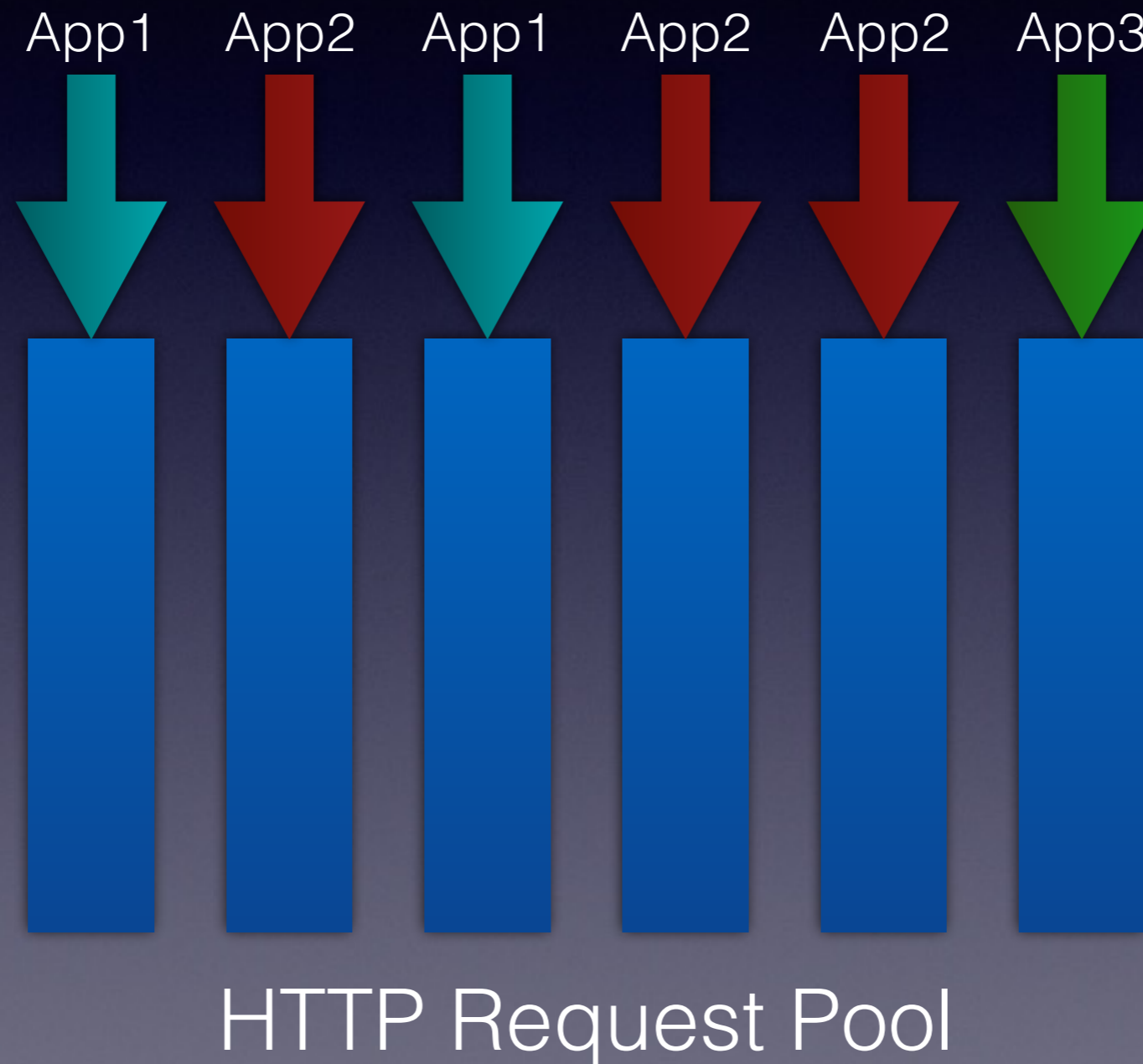


基于Kubernetes打造SAE容器云

丛磊

2016.1.24

目前SAE基于请求的架构



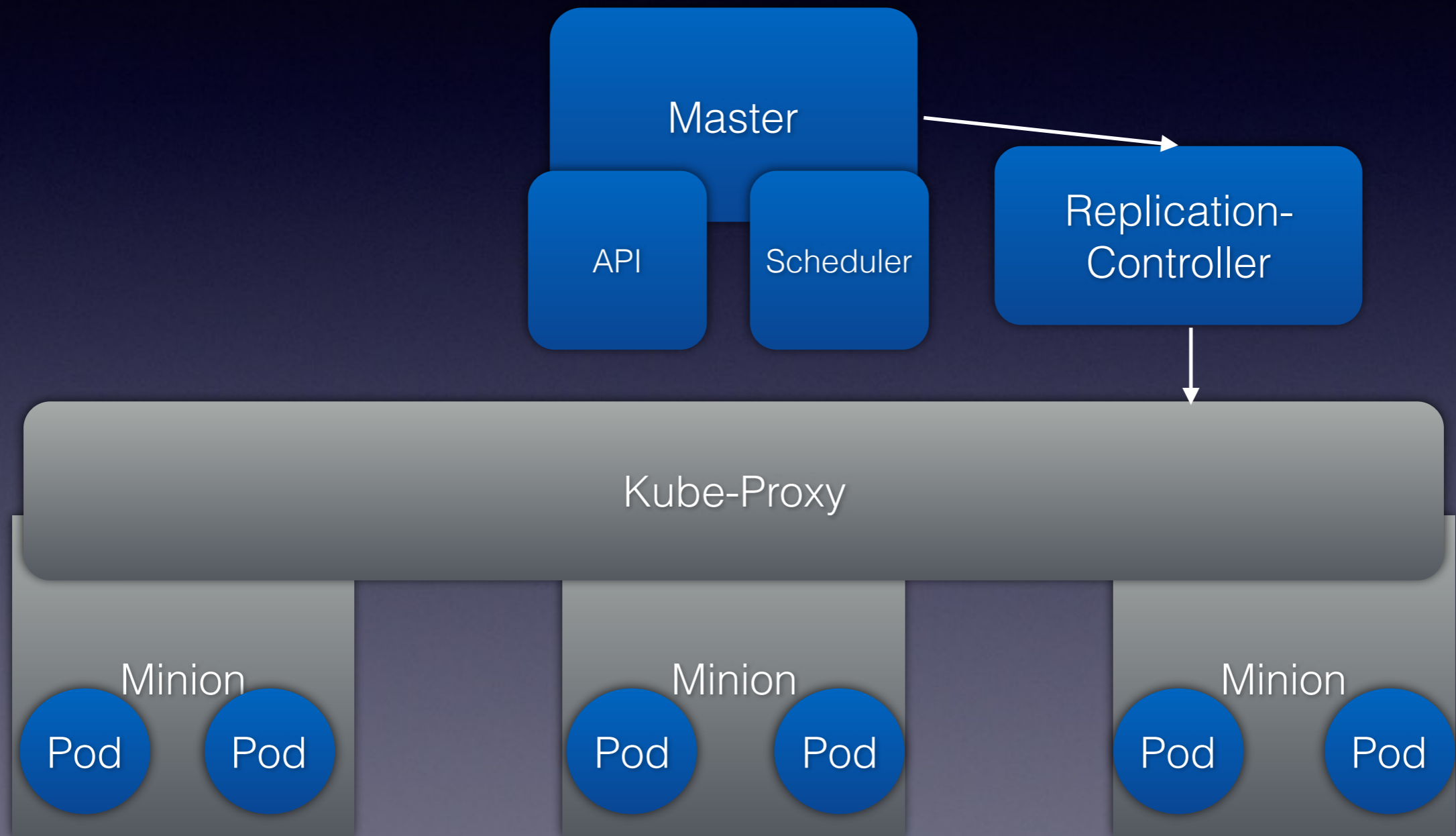
目前SAE基于请求的架构

- 优点
 - 进程内隔离，消耗资源最小
 - 无感扩容&缩容，用户无成本
 - Health&Redispatch，升级切换无成本
- 缺点
 - 无法提供独立的namespace
 - 无法Build&Ship&Run

用户的需求

- 面对代码 vs 面向容器
- 定义一切
- run anywhere
- 无感扩容/缩容

为什么选择Kubernetes



为什么选择Kubernetes

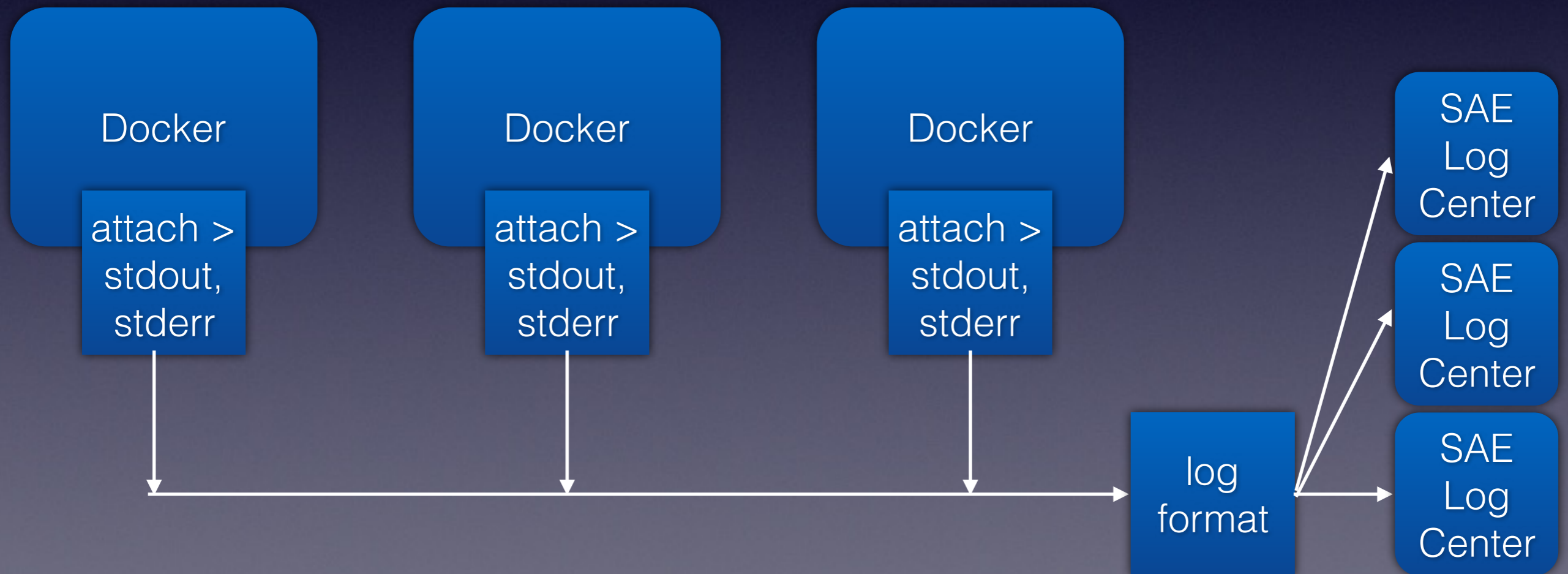
- Pod
- Replication
- Go
- Easy for CentOS6

为什么要改进Kubernetes

- 不足之处：
 - 无感扩容
 - 监控
- 不适合SAE之处：
 - Kube-Proxy&VIP
 - Etcd

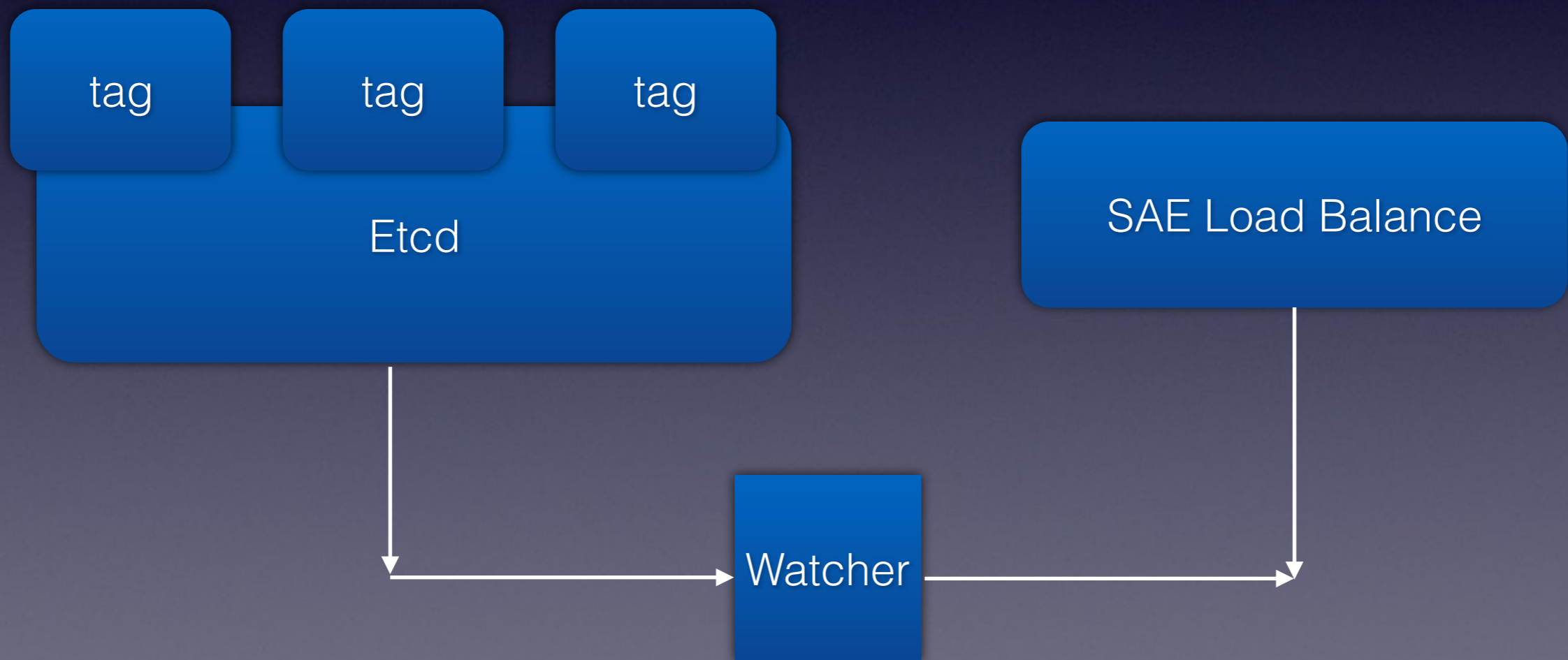
改进Kubernetes

- 日志系统

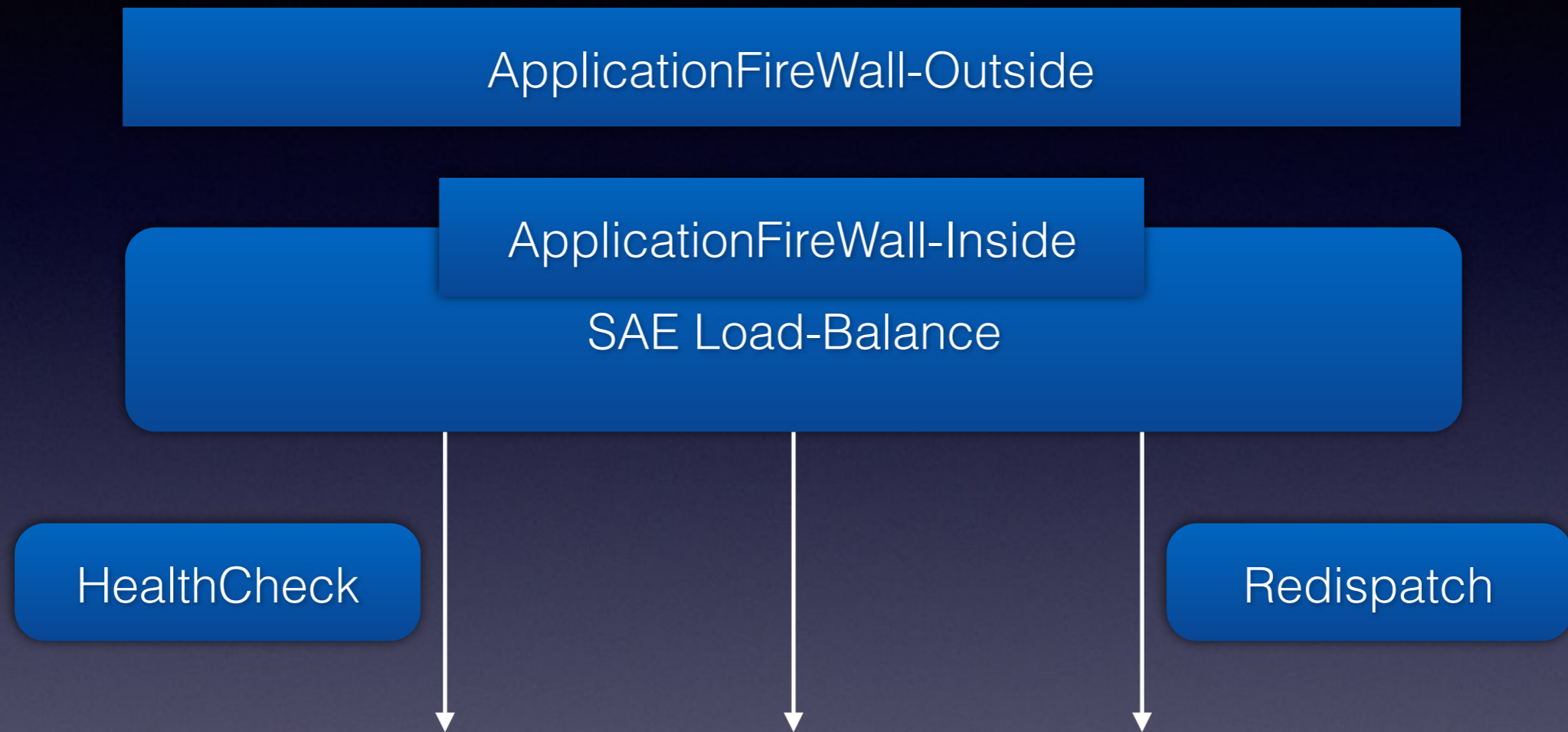


改进Kubernetes

- 接入SAE Load Balance



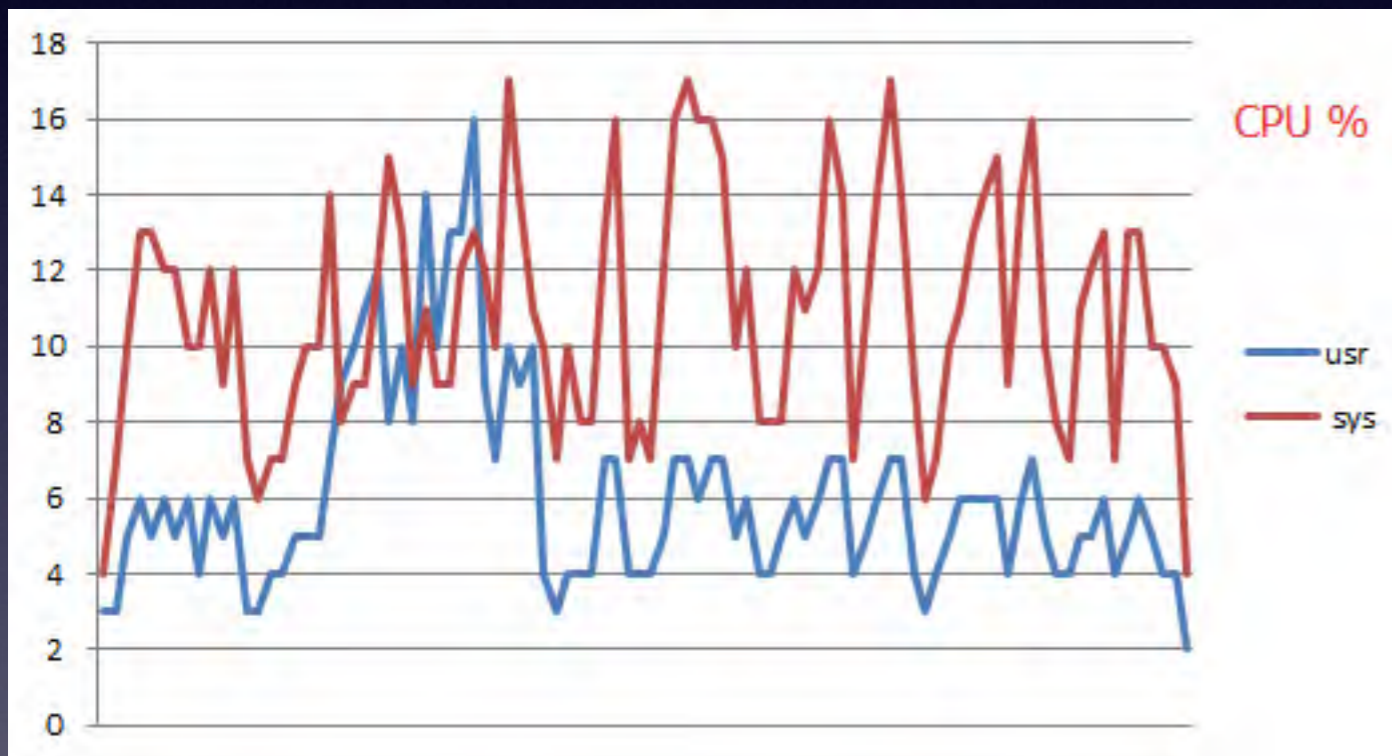
改进Kubernetes



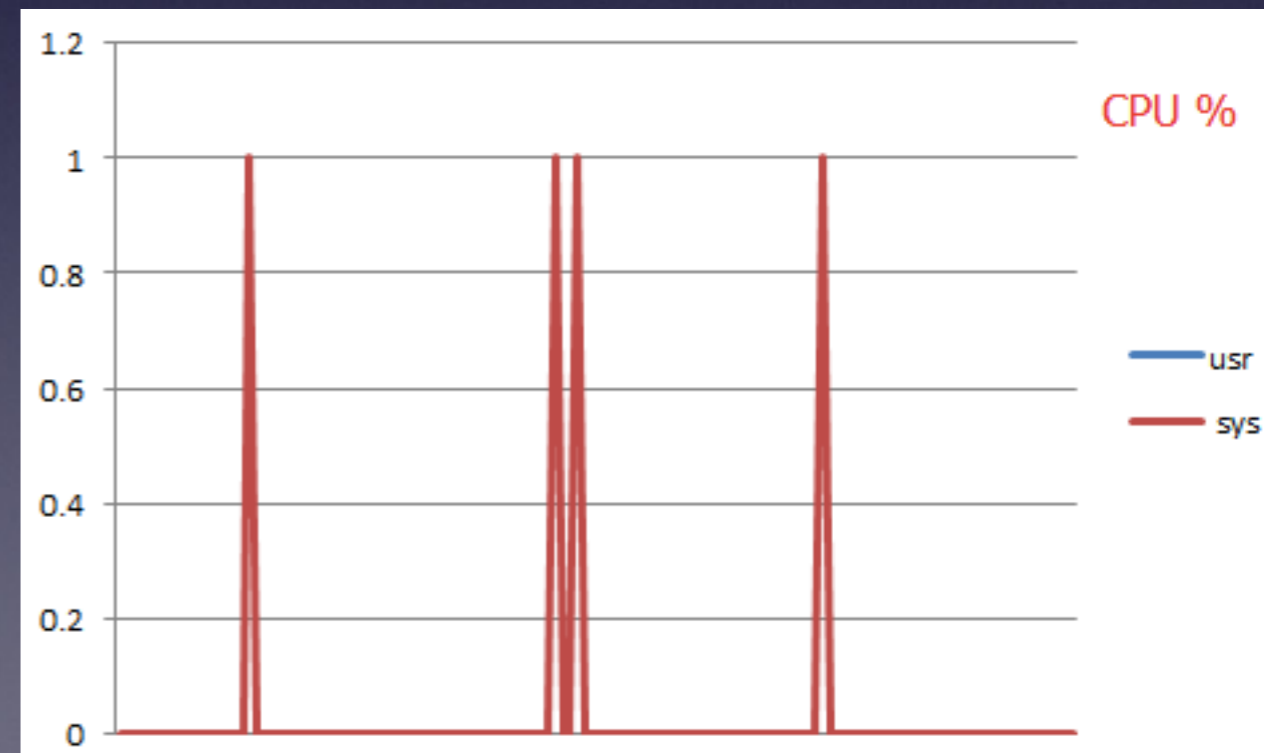
- Redispatch
- 403 vs 444
- Drop vs Reject
- Netlink-Queue

改进Kubernetes

- LoadBalance抗CC攻击压力对比



Nginx 444



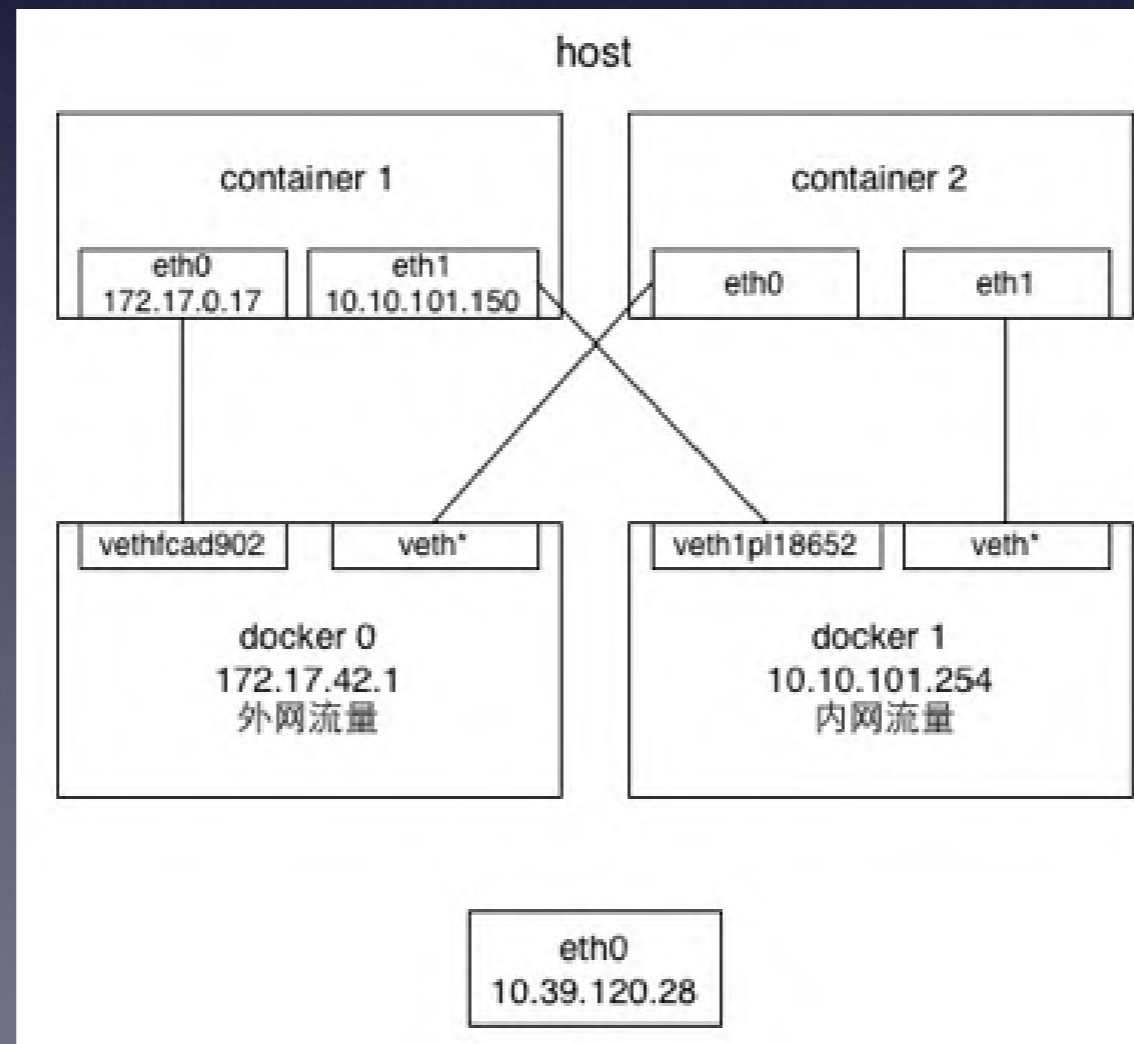
Iptables + Netlink Queue

改进Kubernetes

- PaaS SDN和IaaS SDN的区别
- 网络隔离
 - NAT
 - Bridge（更主流）
- 我们选择NAT
 - NAT提速

改进Kubernetes

- Simple Docker Network
- 内外网流量分开



改进Kubernetes

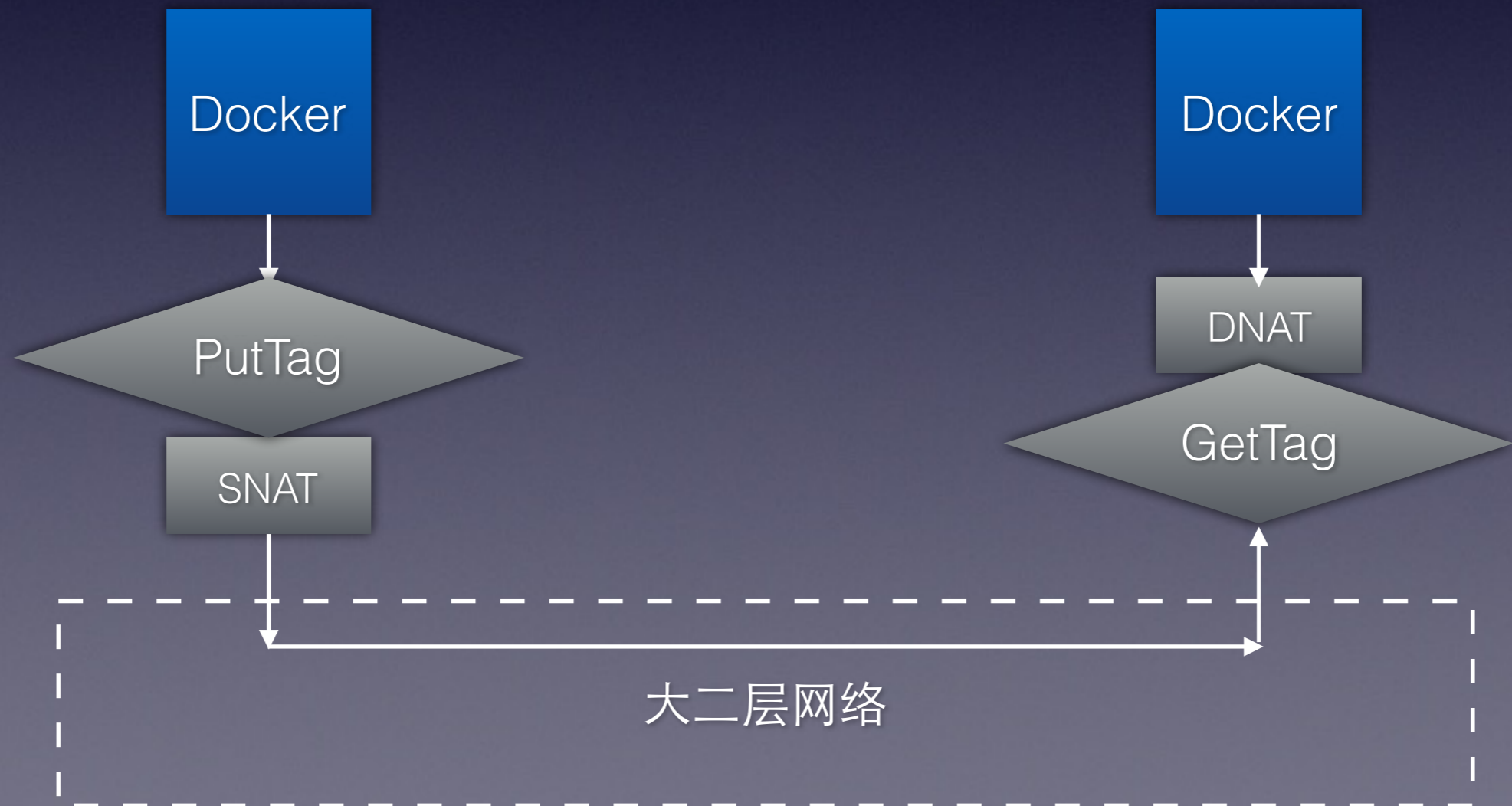
- Simple Docker Network
- 植入Tenant ID

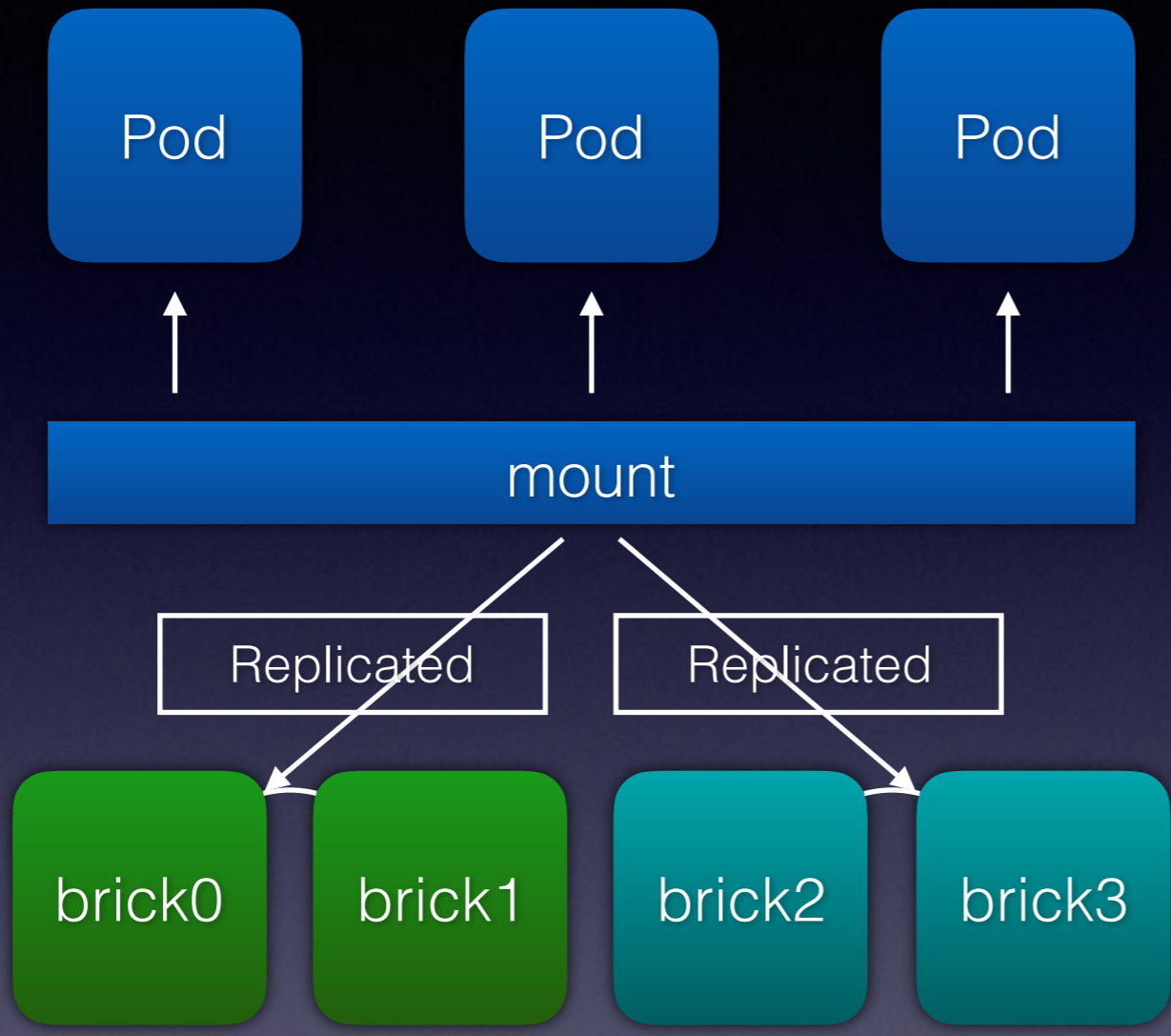
```
case IPOPT_CIPSO:
    if ((!skb && !ns_capable(net->user_ns, CAP_NET_RAW)) || opt->cipso) {
        pp_ptr = optptr;
        goto error;
    }
    opt->cipso = optptr - iph;
    if (cipso_v4_validate(skb, &optptr)) {
        pp_ptr = optptr;
        goto error;
    }
    break;
case IPOPT_SEC:
case IPOPT_SID:
default:
    if (!skb && !ns_capable(net->user_ns, CAP_NET_RAW)) {
        pp_ptr = optptr;
        goto error;
    }
    break;
```

net/ipv4/ip_options.c

改进Kubernetes

- Simple Docker Network





SAE容器云

- 功能：
 - 镜像仓库
 - BuildPkg
 - 无感扩容
 - 共享存储
- 正式发布！





Q & A

丛磊
2015.11.6