

当Open-Falcon遇上Kubernetes

— 高泗俊 平安科技云平台事业部 —

个人简介

硕士毕业于复旦大学，非科班出生，喜欢开源和写作。关注DevOps和容器技术。

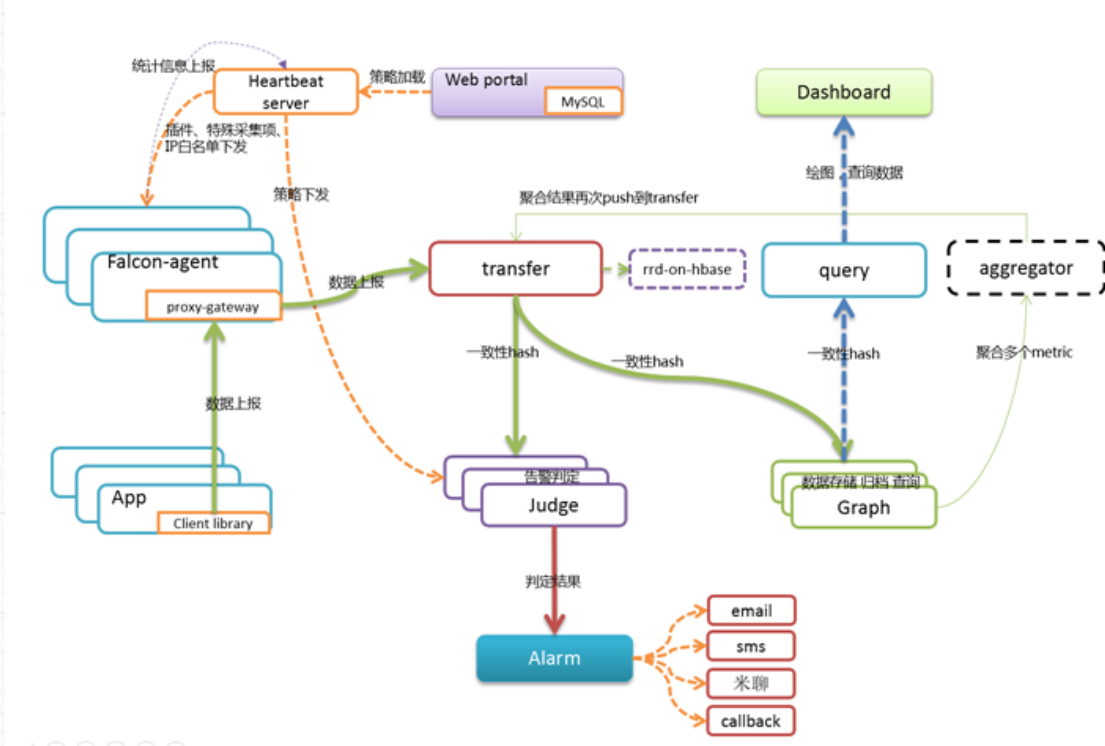
开源电子书：github.com/hackstoic/kubernetes_practice

知乎专栏：云时代的运维开发（二维码）



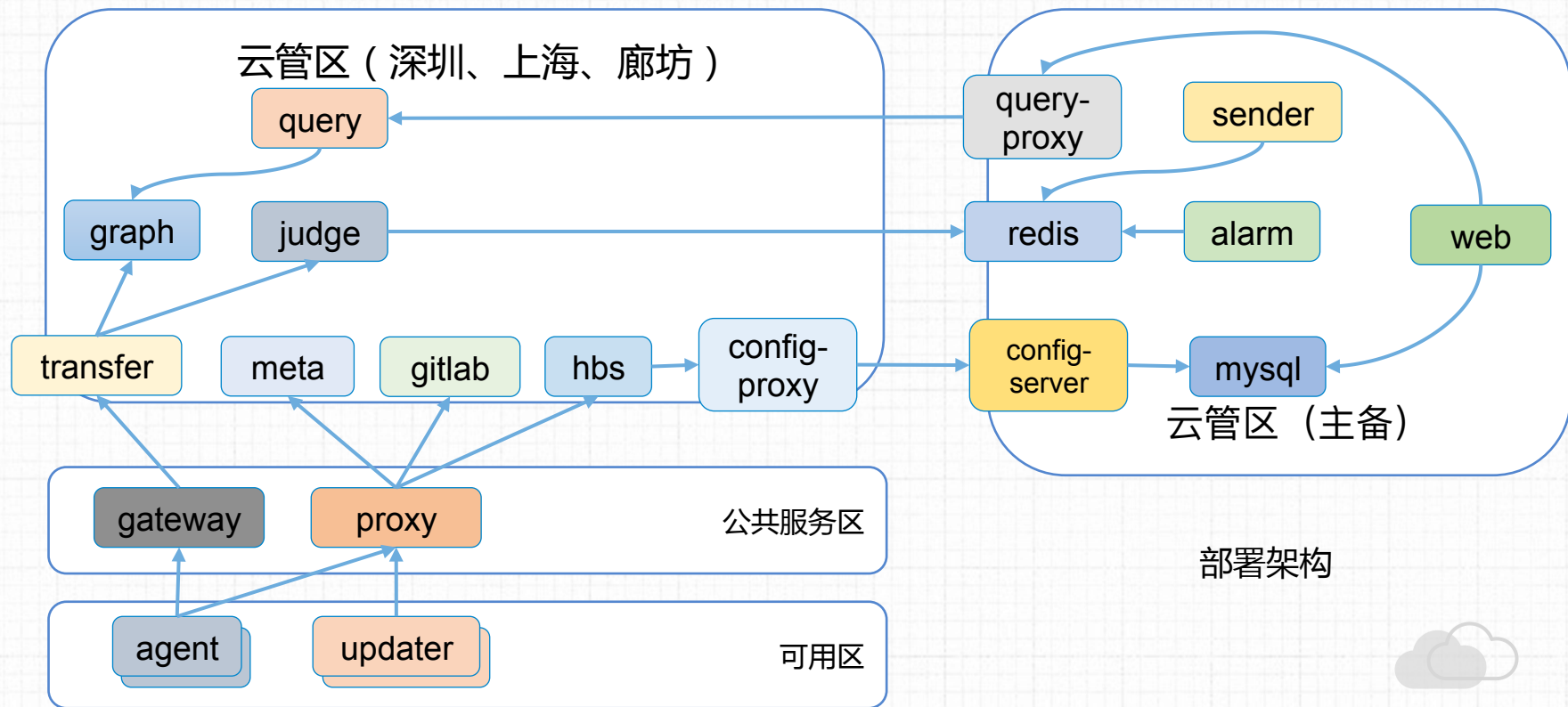
What is Open-Falcon?

基本介绍:小米开源的互联网企业级监控系统, 水平扩展能力, 性能均优于Zabbix。



平安云监控系统 Argus

基本介绍: 结合平安的场景, 在开源的open-falcon v0.1版本基础上二次开发的监控系统



Kubernetes + Argus?

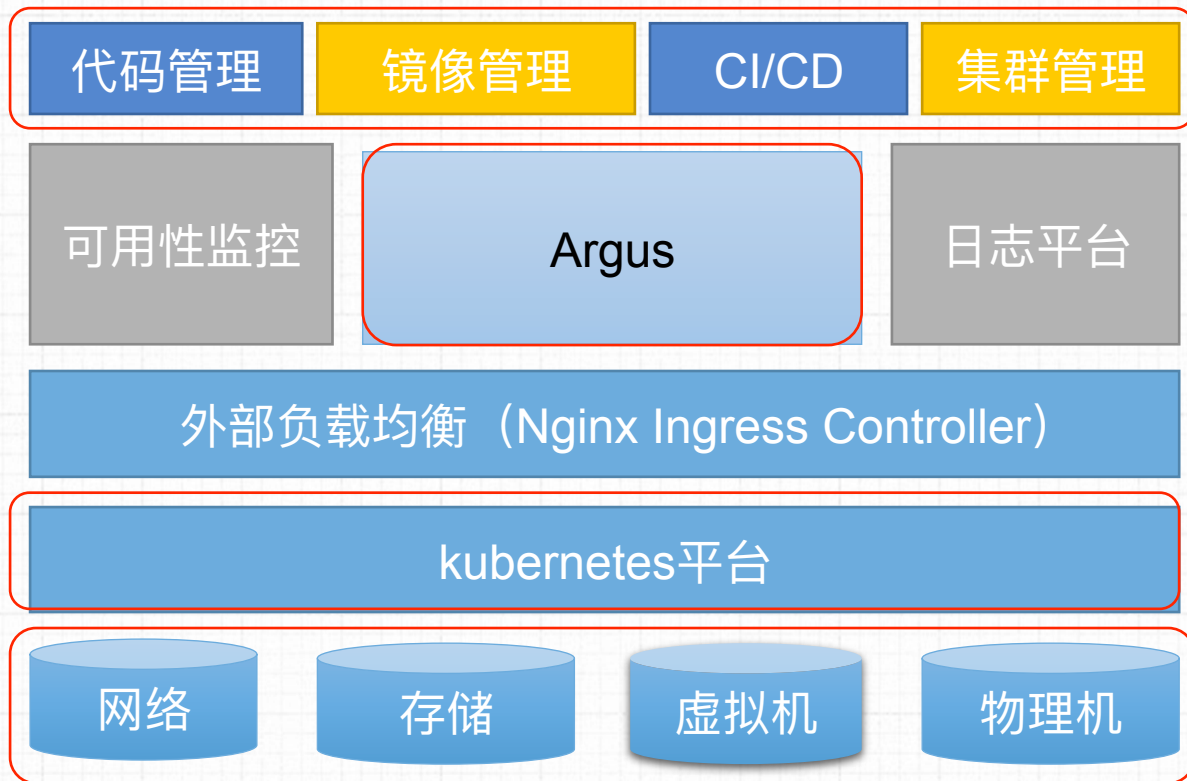
Argus运维痛点

- 集中管理问题： 30+组件， 6套集群， 300+机器
- 各个组件的高可用问题
- 后期的升级，回滚，启停，监控等运维工作
- 主机和IP资源利用率问题

kubernetes特性

- 对应用容器进行集中的调度管理
- 通过复制控制器维持应用的实例数
- Deployment提供方便的实例伸缩，回滚功能，水平扩展等
- 方便启停和管控应用容器
- 方便限制应用使用的资源
- 通过资源隔离提升资源利用率

架构体系



盖房子和搭系统



集群规划(1)

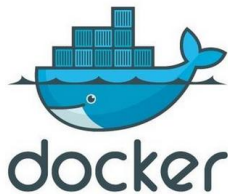
➤ 组件或系统版本



1.6.0



3.1.5



1.12.
3



flannel

0.7.0



CentOS

7.2

➤ 机器配置



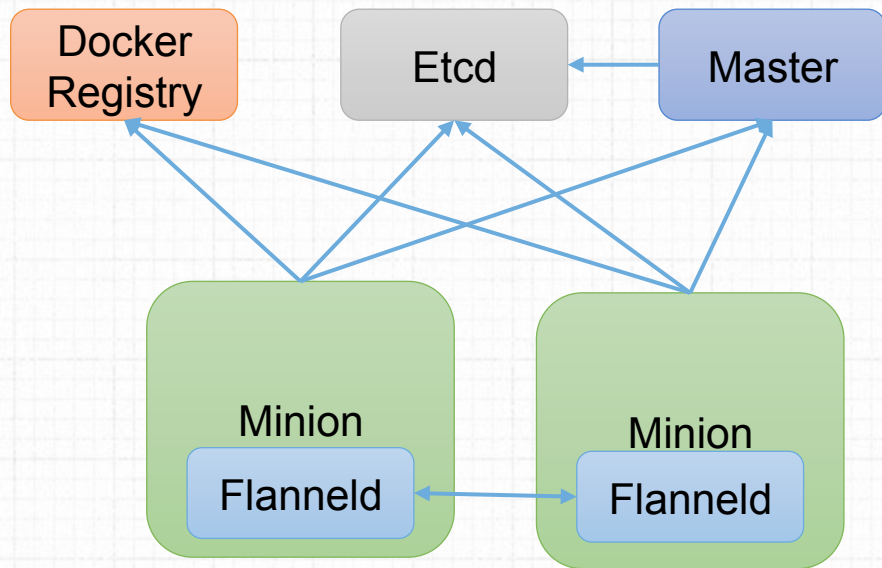
集群规划(2)

➤ 端口与防火墙策略

- Minion → Master
- Master → Etcd
- Minion → Docker registry
- Flanneld → Flanneld
- Flanneld → Etcd

➤ 网络规划

- Flannel的网段 192.168.0.0/16
- Flannel子网的掩码 26 (单机支持 $2^{(32-26)} - 2 = 62$ 个pod)
- Service 网段 10.10.1.0/22
- Node port 的端口范围 1025 ~ 32767



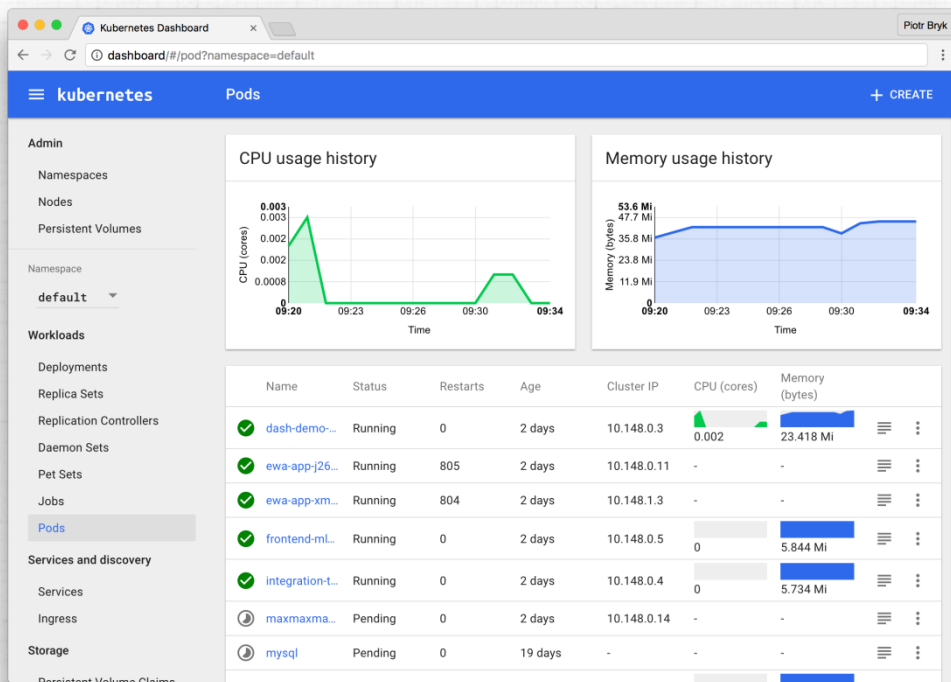
网络&存储

网络方案	性能	上手难度	租户隔离性
Flannel udp	较差	简单	不支持
Flannel vxlan	适中	简单	不支持
Calico	好	较复杂	支持
Weave	差	简单	不支持

存储方案	性能	数据可用性	是否分布式	运维难度
Ceph rbd	好	高	是	复杂
Nas	一般	中	是	适中
本地盘	好	低	否	简单

部署与管理

- Ansible playbook (3天 → 30分钟)
- Kubernetes Dashboard & Kubectl



Argus组件容器化



Dockerfile



Deployment



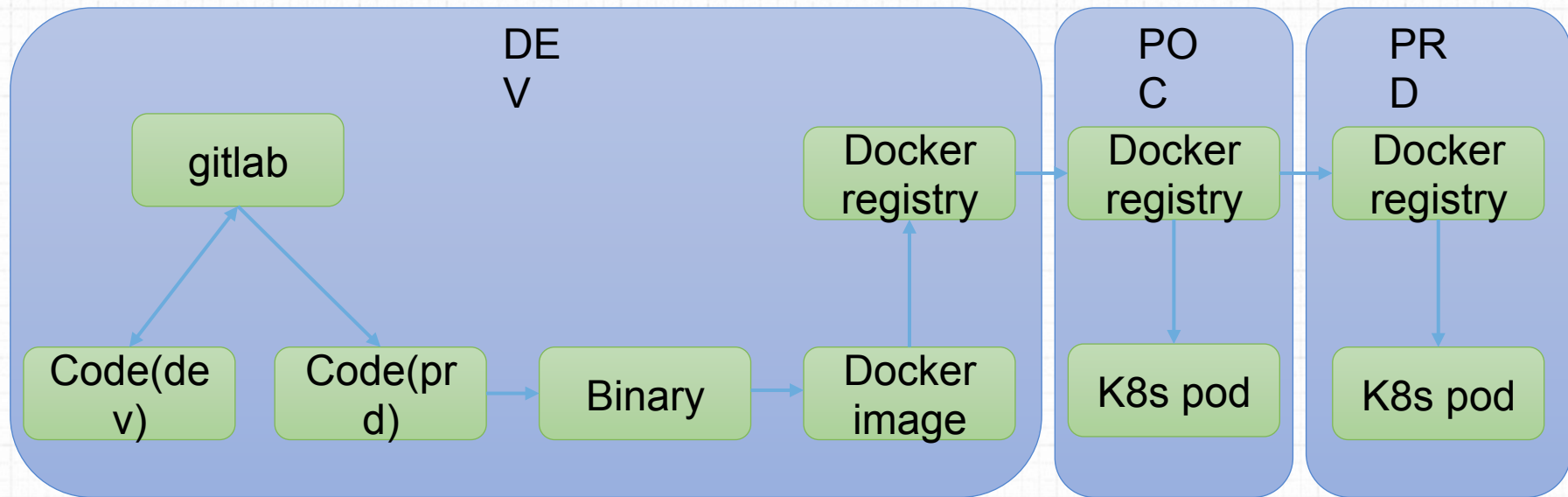
Service



Configmap



开发，发布，部署



监控与日志

- 监控： Argus + Heapster + Anteye
 - 性能监控（1）： 集群主机的监控
 - 性能监控（2）： 容器指标监控
 - 可用性监控（1）： k8s master和Etcd进程监控
 - 可用性监控（2）： Argus核心组件的监控

- 日志： Filebeat + ELK



踩过的坑

- 在flannel 网络中 pod 1 ping 不通 pod2 的几个原因：

- iptables forward 链 默认规则为drop
- 没有打开ip_forward
- 两个主机flannel的侦听端口不一样
- 两个主机没有开通vxlan的防火墙策略
- flannel上报到etcd的是一个外部不可访问的ip



To-Do List

- 应用自身优化
 - 配置热加载
 - 应用健康检查
 - 资源限制和限额
- 集群看板（集群巡检， 集群概览， 系统容灾切换等）
- 使用Init Container 初始化数据库， 并检查数据库是否Ready
- API Gateway（对外接口的权限验证， 限流， 审计等）
- Meta， gitlab等非核心有状态模块使用Ceph rbd替代本地盘存储
- 整合CI/CD流程



应用场景拓展



Thanks

| 平安云 • 金融云 |

