

网易云 & K8S

网易云如何基于K8S提供容器服务

陈定斌

资深解决方案架构师

主办方：



协办方：





主题 1：网易容器服务的历史

- ◆ 2015年初，容器云容器服务启动调研（k8s 1.0都没有发布，但是我们已经在调研k8s 0.6版本）
- ◆ 2015年10月左右，基于k8s 1.0的基于网易基础设施研发的容器服务v1.0灰度发布
- ◆ 2016年5月，网易云启动了微服务版本的容器平台研发，基于k8s 1.3版本，容器全面支持计算、存储、网络和服务发现，并与2016年10月正式发布容器服务v2.0版本
- ◆ 2017年上半年，产品启动新版计算（一期），在k8s 1.6基础上，重构原有代码和架构，以一种对k8s侵入更少的方式以方便后续快速升级k8s版本。
- ◆ 2017年9月份左右上线基于k8s 1.3的vpc网络容器和裸机容器公有云上线
- ◆ 2017年11月启动基于k8s 1.8的k8s社区一致性和同时支持容器“fargate模式”和k8s api的版本的最新容器版本研发
- ◆ 2018年4~5月份 建德私有云上线基于k8s 1.9（1.8直接线下升级到1.9）的版本



主题 2：网易K8S使用历程

基于k8s 1.0的基
于网易基础设施
研发的容器服务

K8S 1.0

基于k8s 1.3的vpc网
络容器和裸机容器公
有云上线

K8S 1.3

因为年中vpc
的适配和裸
机容器开发
没有上线

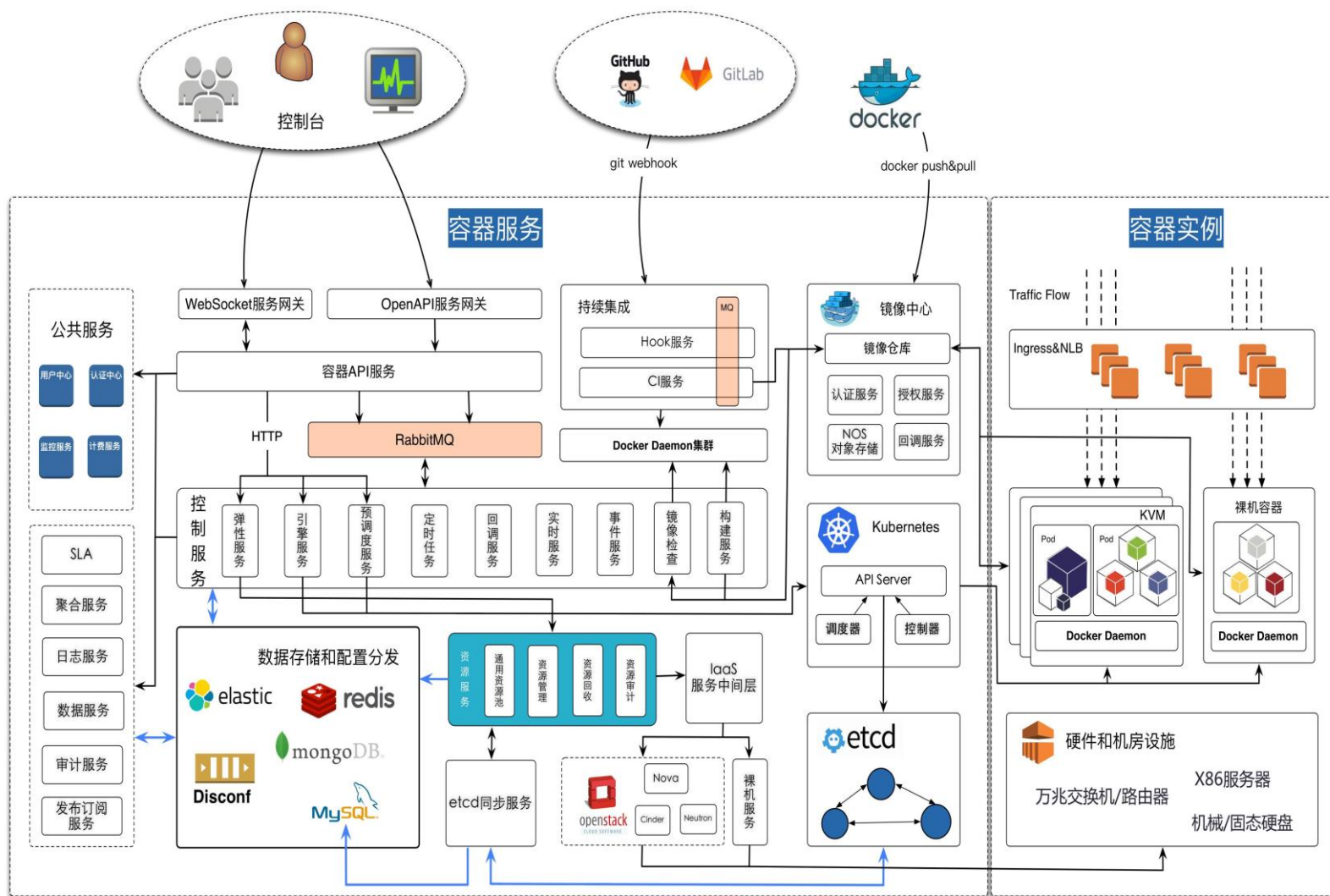
K8S 1.6

建德私有云
上线基于k8s
1.9版本容器
服务

K8S 1.9



网云容器服务



网易云-----容器网络



网易云对Kubernetes的定制

- 容器服务和Kubernetes服务

两种使用模式同时支持

- 深度整合网易云基础资源

虚拟主机 / 高性能物理机

经典网络 / VPC网络

多种类型块存储

负载均衡

- 公有云全面改造

多租户隔离

安全加固

动态资源管理

- 性能优化/功能增强

单集群规模

主流程耗时全链路优化

高级有状态服务

数据自动迁移

垂直扩容

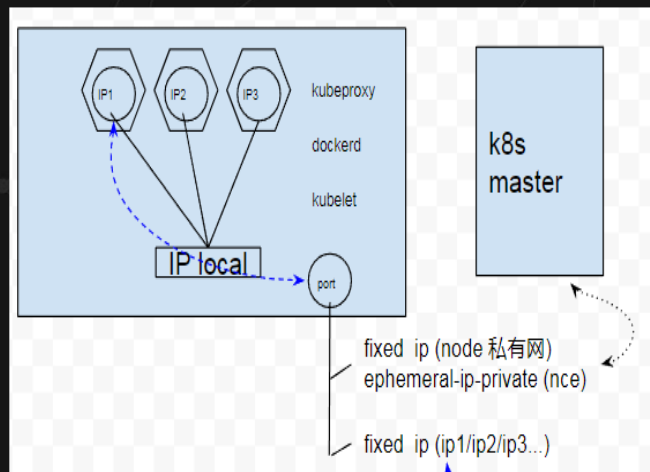
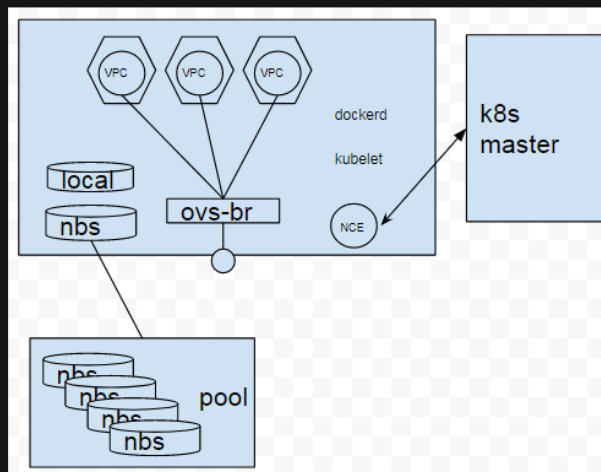
...

网易云-----容器网络

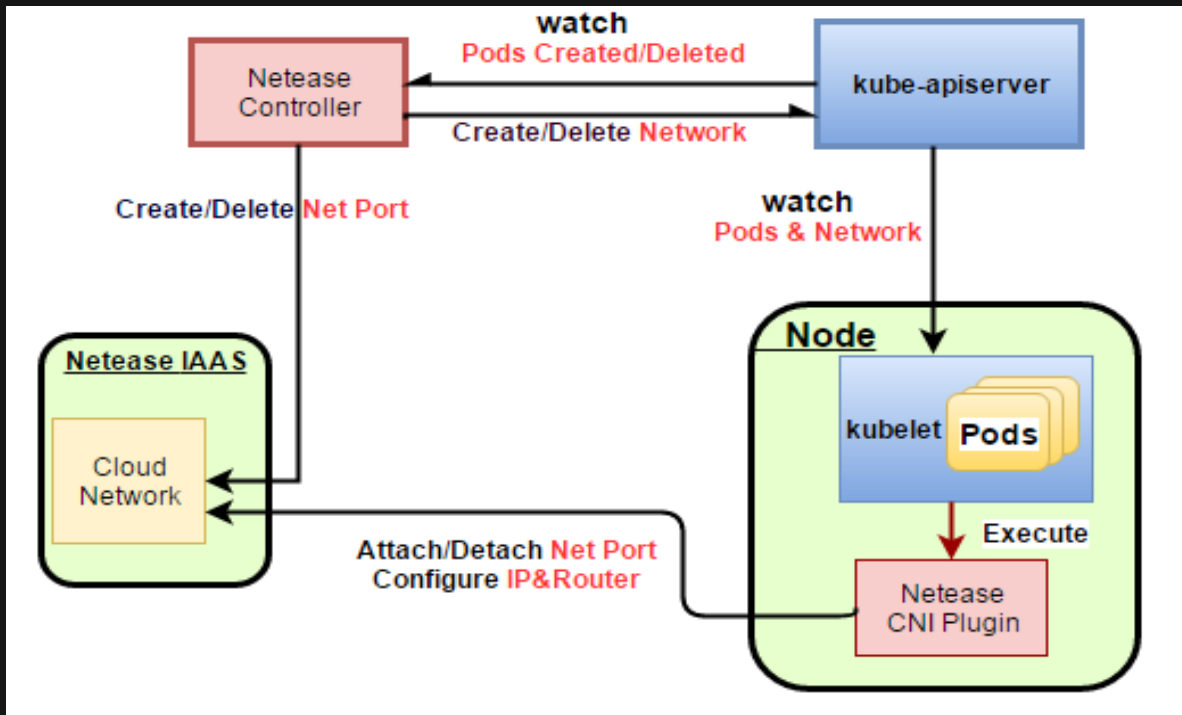


VPC网络是公有云平台的基础设施。为用户构建一个逻辑隔离的网络空间，自定义IP 地址范围网段、路由表，并通过安全组实现安全防护。

- 网络隔离
- 灵活
- 安全



网易云-----容器网络



- 三层扁平网络：容器私有网，容器公网，node管理网
- ip 管理：私有网 IPAM 不冲突。有状态希望永远不变，所以 pod 必须与网络 ip 绑定
- cni 插件：网桥管理，net ns 设置路由表。如果是私有网配置网桥等，外网还需要挂载、卸载设置路由等。



和K8S结合使用

K8s内部的负载均衡

- 直接使用Service名称访问
- Server->VIP->pod IP
- 自动适配service扩展

无状态服务，建议使用service

- 域名服务+Service
- Service 原理



网易云----负载均衡

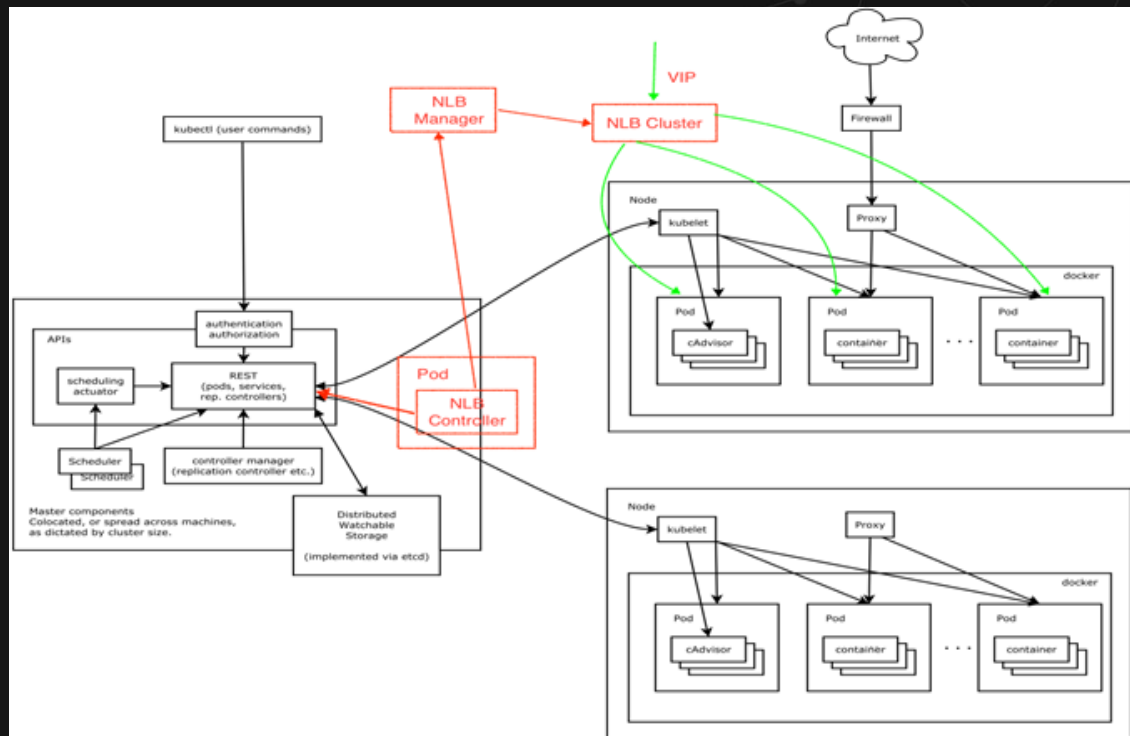


外部负载均衡

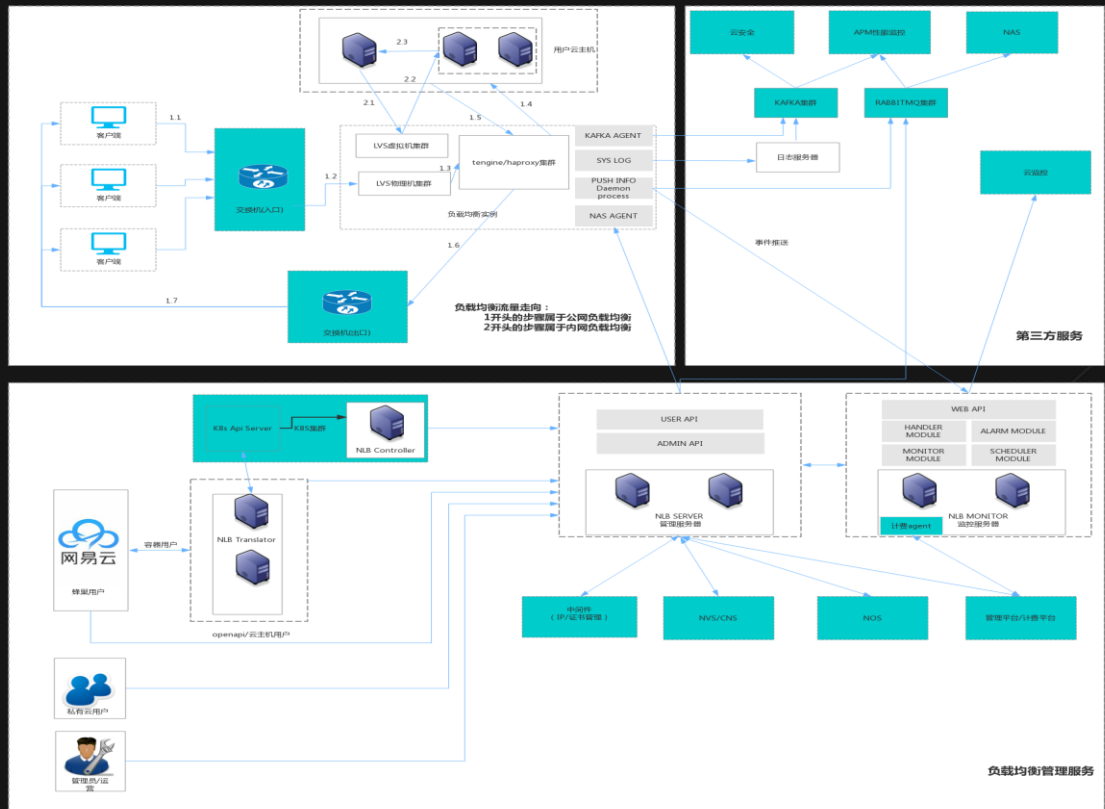
K8S 官方的几种选择

网易云提供的

- Ingress基础之上的扩展
- 负载均衡直接接入service
- 支持K8s 命令行



网易云----负载均衡



负载均衡使用上需要注意容器与云主机两种类型的使用方式。

- 软件实现的负载均衡
- 基于LVS+haproxy提供四、七层负载均衡
- 暂时不支持后端同时存在云主机和POD
- 基于DPDK实现高性能负载均衡转发能力

K8S升级——1.0~1.3



微服务版本的容器服务从整体产品形态上就有很大的变动，包括上下层的技术架构、表结构、数据、接口等都会没法兼容。

从整体来说，迁移从上到下涉及的模块主要包括：

- 前端及web服务
- nce后端服务
- k8s编排服务
- docker容器
- nlb外网负载均衡
- 日志服务
- 性能监控

K8S升级——1.0~1.3



K8S升级:

k8s容器编排层的升级迁移主要包括k8s各组件二进制程序和k8s管理的元数据两部分:

程序升级

- 涉及到master和node
- 程序二进制文件升级
- 初始化脚本升级

元数据迁移

- etcd数据库中已有数据的升级，包括namespace、pod、node、deployment等几乎所有资源的数据升级订正。
- service/endpoint/ingress资源要补齐
- namespace 对应一个网易云用户（tenant ID），现在一个用户可以有多个namespace，namespace名字命名格式都要修改



1. 容器(Pod)级故障

应对措施:

- 容器退出自动重启
- 探测异常自动重启

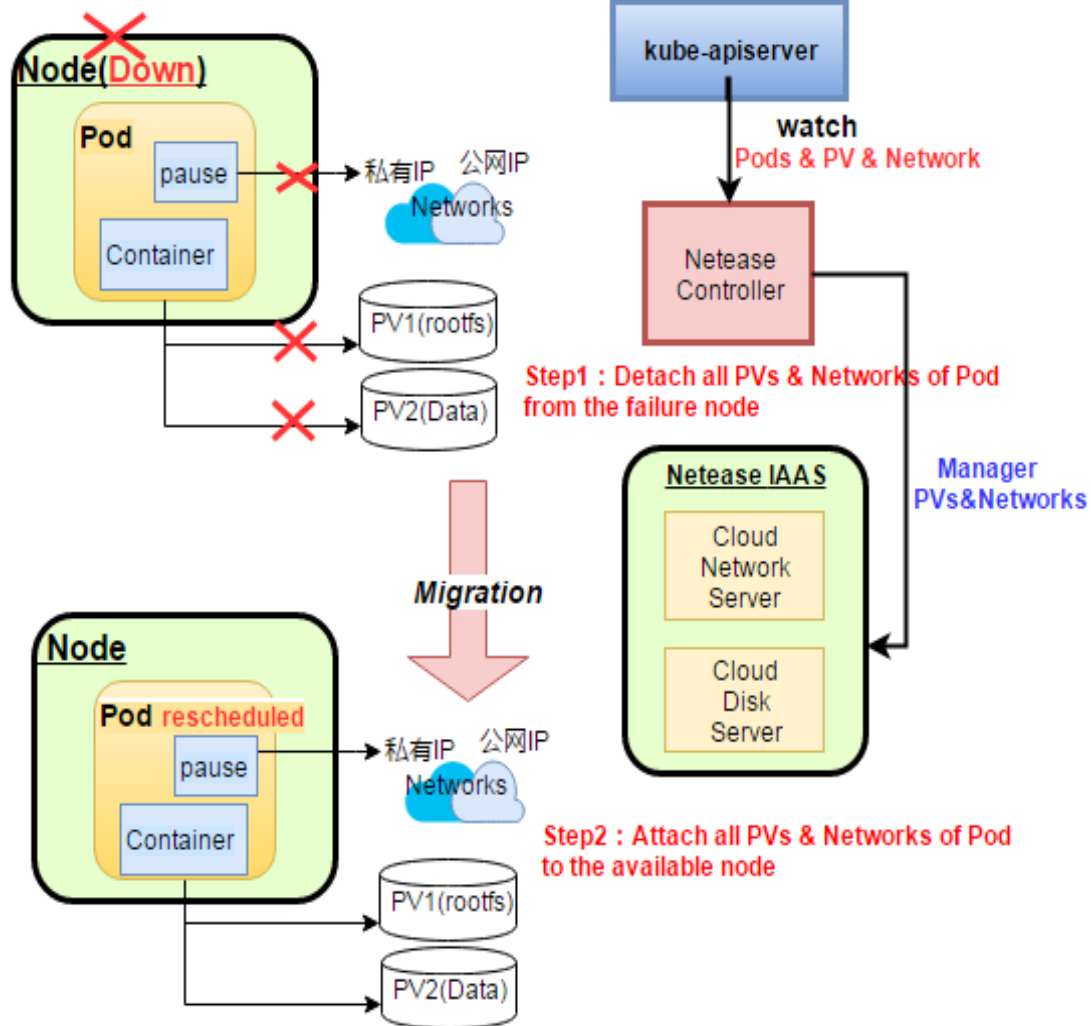
2. 节点(Node)级故障

应对措施:

- Node心跳定时探测
- 离线后容器自动重建
- 有状态云盘自动迁移



网易云1.0有 状态容器故障 迁移



谢 谢



主办方：



caicloud 才云



 TensorFlow

协办方：



网易云