易宝集团
YEEPAY GROUP

容器云在易宝的落地

于涛

2016年8月

自我介绍

- 98年工作，2000加入互联网大潮，互联网运维探索和实践者
- 先后服务过大型国企、ISP、门户网站、网游、移动互联网等
- 现在在易宝支付任职系统架构师
- 做过运维方面的网络、系统、安全、架构设计等工作
- 从手动运维到容器时代，一路走来

易宝集团 YEEPAY GROUP

议程——技术流与方法论

- 从手工运维到自动运维
- 云时代与运维
- Kubernetes与云
- Kubernetes新特性
- 易宝容器云
- 在生产中的落地
- 后期工作及未来展望

异构环境的问题

- 问题
  - 硬件五花八门
  - 黑盒问题
  - 厂家主导：产品＋服务模式
  - 互联互通、匹配、协议等问题引发的扯皮
  - 各自为政的系统：网管系统、监控系统.....
  - 响应与故障解决引发的效率问题

- 解决之道
  - 从封闭到开放
  - 变被动为主动到驾驭

多些功劳，少些苦劳，少些疲劳

这几天微信群流传着一张神秘的图片，说台湾同胞也拜机房。

自动化工具

- 系统标准化环境：kvm、vmware、docker
- 环境、软件、配置管理自动化：cfengine、puppet、salts
- 产品部署、产品版本管理：Capistrano
- 数据迁移与变更：dbdeploy

监控：粒度、关联度、可视化

- 服务器：cpu、内存、磁盘、网卡、丢包、延迟
- 网络设备：cpu、session、ping、错误日志
- 网络质量：延时、丢包、流量、业务流量（端口）
- 应用监控：port、进程、业务流量、状态码、预期内容、依赖关系、数据文件、资源消耗（内存、cpu）、延时、容量指标
- 开源选型：zabbix（推荐）、nagios、cacti…
- 自建监控：通过js上报、idc互相监控
- 第三方监控：gomez、基调…
- 关联监控
- 数据可视化

ITIL与互联网运维

- 以ITIL为核心的运维：CMDB、变更管理、配置管理……xx管理
- ITIL与持续交付

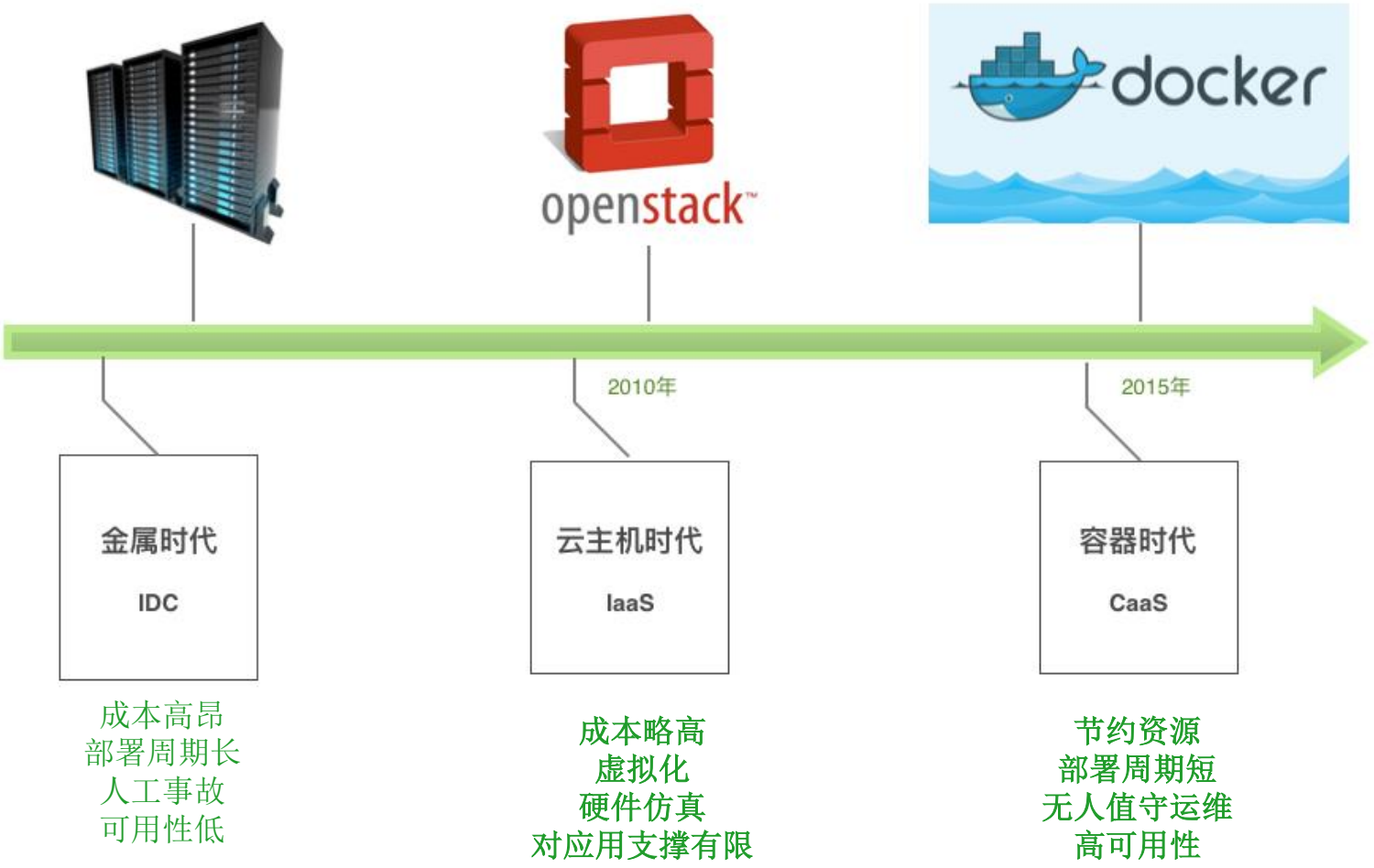
  “
  持续交付的关注点在于迭代和增量交付，以及跨功能职责角色之间的协作。ITIL从服务设计与服务运营的角度来考虑这些事情。但是，当谈到服务转化（尤其是开发、测试和部署）时，就有点被忽视了。
  ”

  ——《持续交付》P344

易宝集团
YEEPAY GROUP

终于，云时代来了

基础设施服务的演进


openstack™
docker
2010年
2015年
金属时代
IDC
云主机时代
IaaS
容器时代
CaaS
成本高昂
部署周期长
人工事故
可用性低
成本略高
虚拟化
硬件仿真
对应用支撑有限
节约资源
部署周期短
无人值守运维
高可用性

容器云对应用的支撑

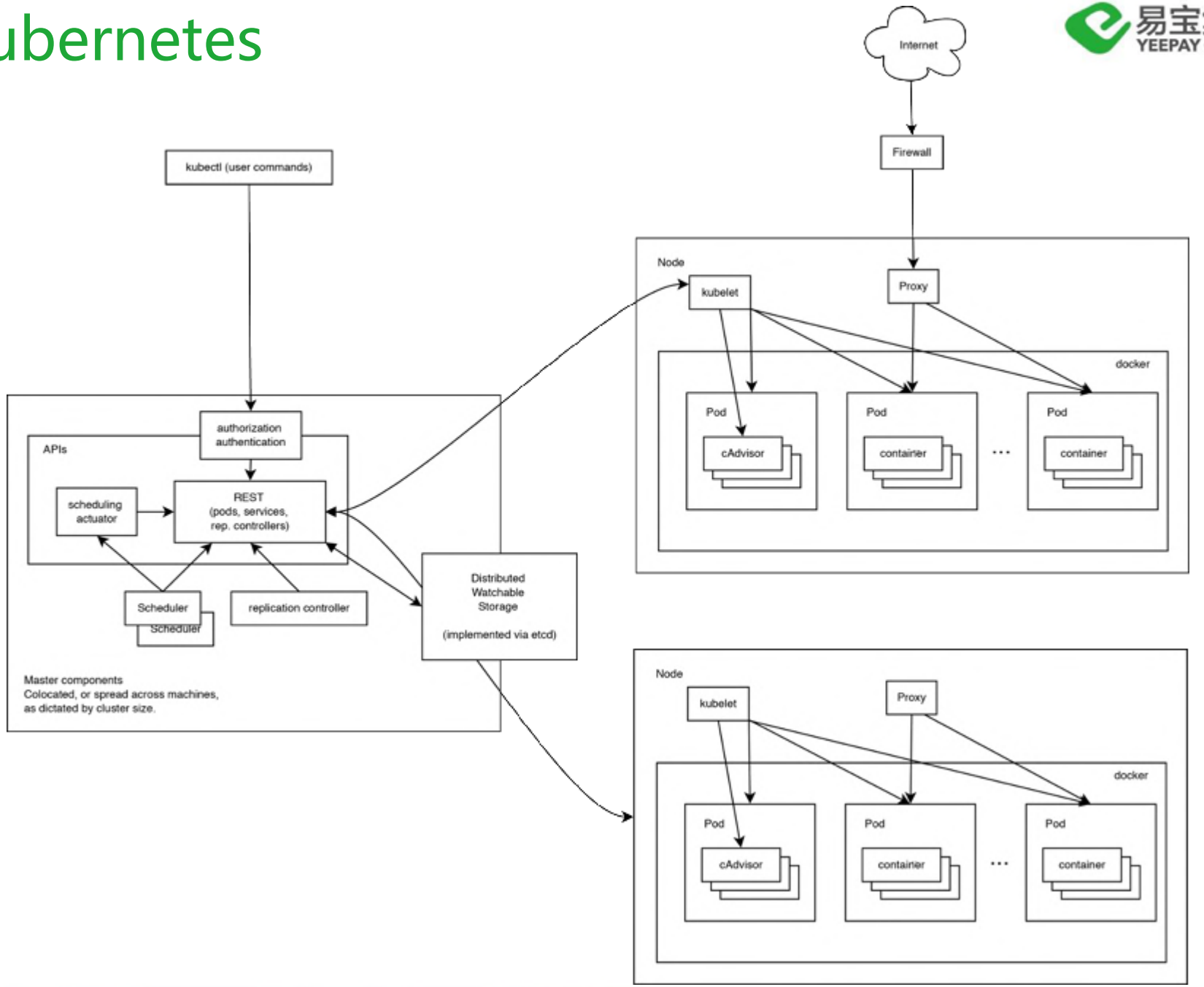
- 屏蔽底层，对应用透明
- 让基础设施如丝般顺滑
- 提供应用的全生命周期管理
- 随时、随地上线和回滚
- 弹性计算：无感知的按需扩容
- 持续集成、持续交付

CAAS

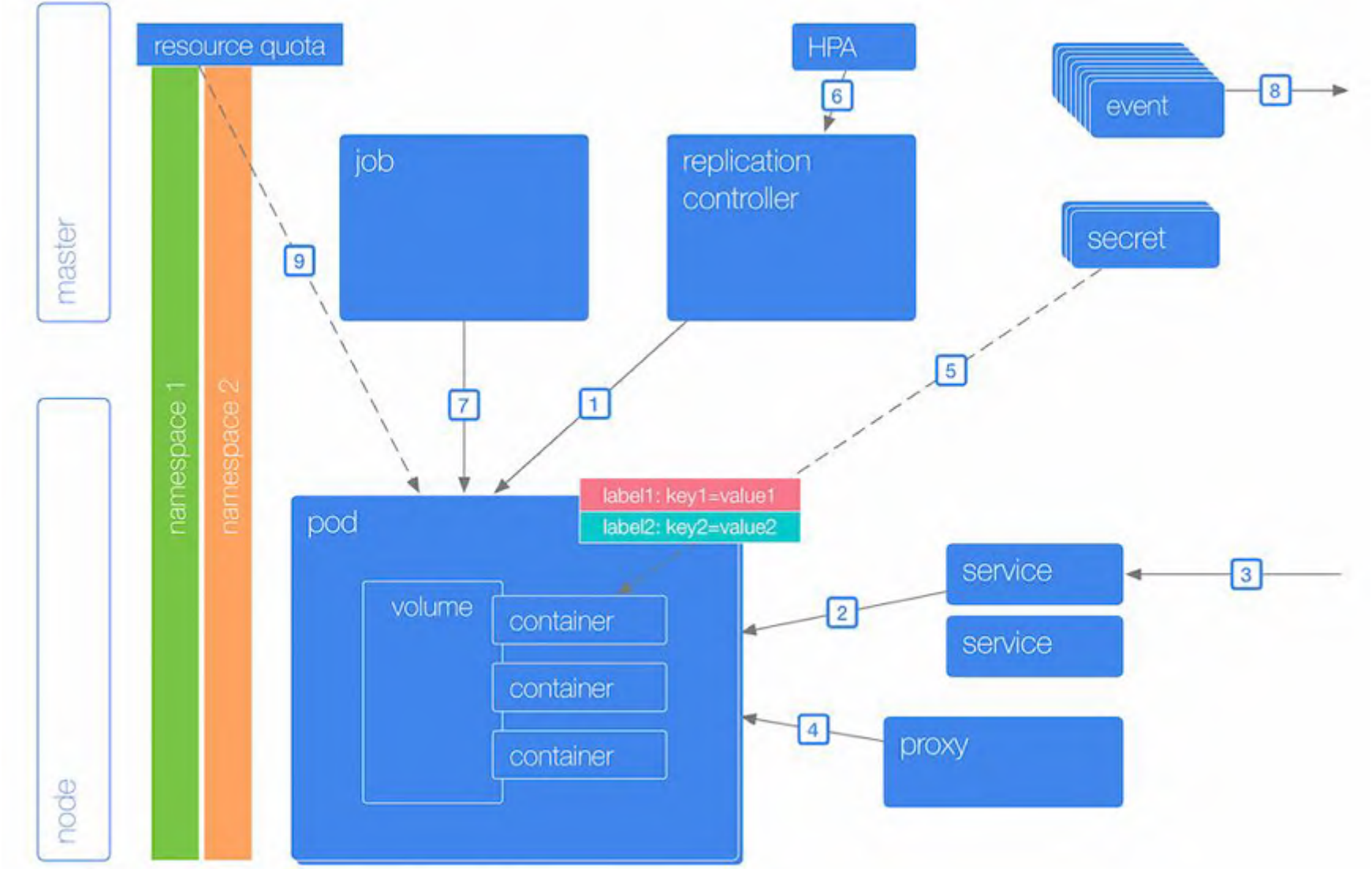
主流容器平台

- Mesos
- Kubernetes
- Docker/Swarm

Kubernetes


Internet
Firewall
kubectl (user commands)
Node
kubelet
Proxy
docker
Pod
cAdvisor
Pod
container
Pod
container
authorization
authentication
APIs
scheduling
actuator
REST
(pods, services,
rep. controllers)
Scheduler
Scheduler
replication controller
Distributed
Watchable
Storage
(implemented via etcd)
Master components
Colocated, or spread across machines,
as dictated by cluster size.
Node
kubelet
Proxy
docker
Pod
cAdvisor
Pod
container
Pod
container

Kubernetes


master
node
resource quota
namespace 1
namespace 2
job
HPA
6
replication controller
event
8
secret
9
7
1
5
label1: key1=value1
label2: key2=value2
pod
volume
container
container
container
service
service
3
2
4
proxy

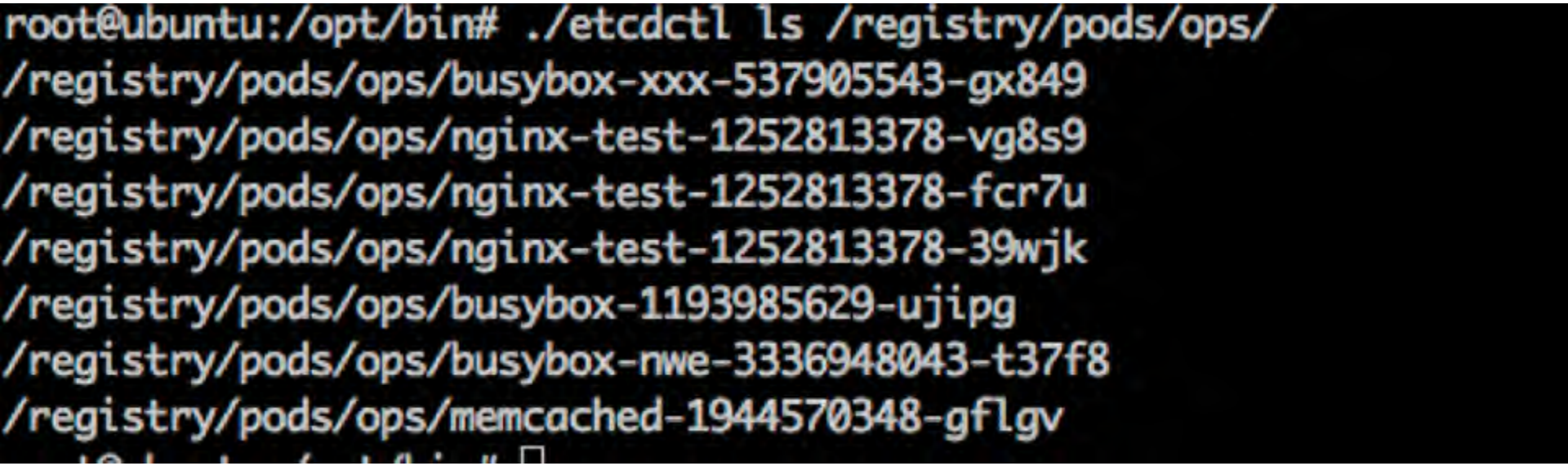
ETCD主机信息 - - minion

```
root@ubuntu:/opt/bin# ./etcdctl list /registry/minions/
No help topic for 'list'
root@ubuntu:/opt/bin# ./etcdctl ls /registry/minions/
/registry/minions/172.21.1.24
/registry/minions/172.21.1.25
/registry/minions/172.21.1.26
/registry/minions/172.21.1.27
/registry/minions/172.21.1.11
/registry/minions/172.21.1.21
/registry/minions/172.21.1.22
/registry/minions/172.21.1.23
```

网络配置 - - flannel

```
root@ubuntu:/opt/bin# ./etcdctl ls /coreos.com/network/subnets/
/coreos.com/network/subnets/10.0.102.0-24
/coreos.com/network/subnets/10.0.17.0-24
/coreos.com/network/subnets/10.0.47.0-24
/coreos.com/network/subnets/10.0.84.0-24
/coreos.com/network/subnets/10.0.46.0-24
/coreos.com/network/subnets/10.0.62.0-24
/coreos.com/network/subnets/10.0.99.0-24
/coreos.com/network/subnets/10.0.4.0-24
```

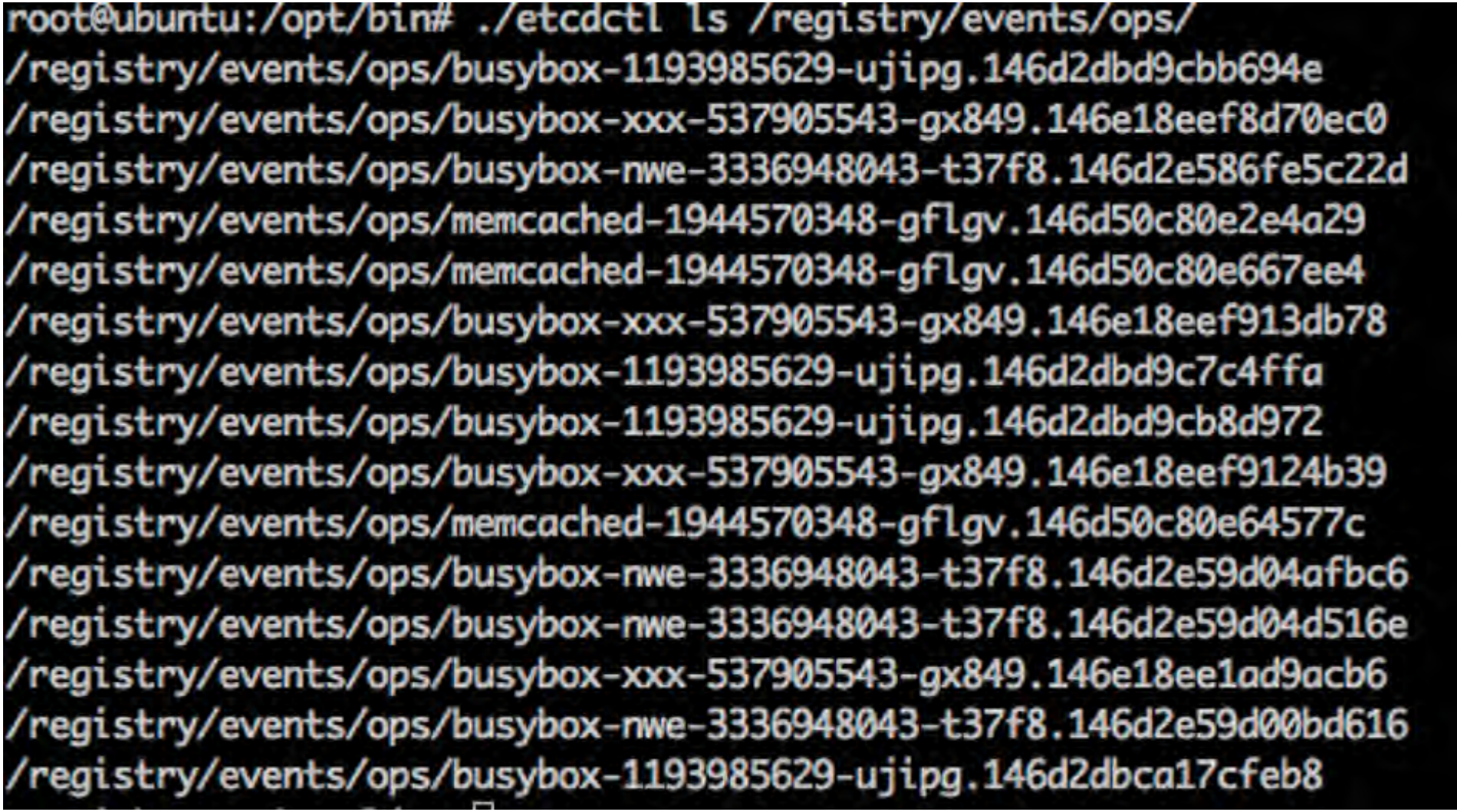
应用信息 - - pod


root@ubuntu:/opt/bin# ./etcdctl ls /registry/pods/ops/
/registry/pods/ops/busybox-xxx-537905543-gx849
/registry/pods/ops/nginx-test-1252813378-vg8s9
/registry/pods/ops/nginx-test-1252813378-fcr7u
/registry/pods/ops/nginx-test-1252813378-39wjk
/registry/pods/ops/busybox-1193985629-ujipg
/registry/pods/ops/busybox-nwe-3336948043-t37f8
/registry/pods/ops/memcached-1944570348-gflgv

服务信息 - - Service

```
root@ubuntu:/opt/bin# ./etcdctl ls /registry/services/specs/default/
/registry/services/specs/default/mongo-svc
/registry/services/specs/default/mongo-svc-b
/registry/services/specs/default/proxy
/registry/services/specs/default/redis-slave
/registry/services/specs/default/registry
/registry/services/specs/default/ui
/registry/services/specs/default/fyy-service
/registry/services/specs/default/fyy1-service
/registry/services/specs/default/mysql
/registry/services/specs/default/redis-master
/registry/services/specs/default/kubernetes
/registry/services/specs/default/mongo-svc-c
/registry/services/specs/default/mongo-svc-a
```

事件信息 - - Events


root@ubuntu:/opt/bin# ./etcdctl ls /registry/events/ops/
/registry/events/ops/busybox-1193985629-ujipg.146d2dbd9cbb694e
/registry/events/ops/busybox-xxx-537905543-gx849.146e18eef8d70ec0
/registry/events/ops/busybox-nwe-3336948043-t37f8.146d2e586fe5c22d
/registry/events/ops/memcached-1944570348-gflgv.146d50c80e2e4a29
/registry/events/ops/memcached-1944570348-gflgv.146d50c80e667ee4
/registry/events/ops/busybox-xxx-537905543-gx849.146e18eef913db78
/registry/events/ops/busybox-1193985629-ujipg.146d2dbd9c7c4ffa
/registry/events/ops/busybox-1193985629-ujipg.146d2dbd9cb8d972
/registry/events/ops/busybox-xxx-537905543-gx849.146e18eef9124b39
/registry/events/ops/memcached-1944570348-gflgv.146d50c80e64577c
/registry/events/ops/busybox-nwe-3336948043-t37f8.146d2e59d04afbc6
/registry/events/ops/busybox-nwe-3336948043-t37f8.146d2e59d04d516e
/registry/events/ops/busybox-xxx-537905543-gx849.146e18ee1ad9acb6
/registry/events/ops/busybox-nwe-3336948043-t37f8.146d2e59d00bd616
/registry/events/ops/busybox-1193985629-ujipg.146d2dbca17cfeb8

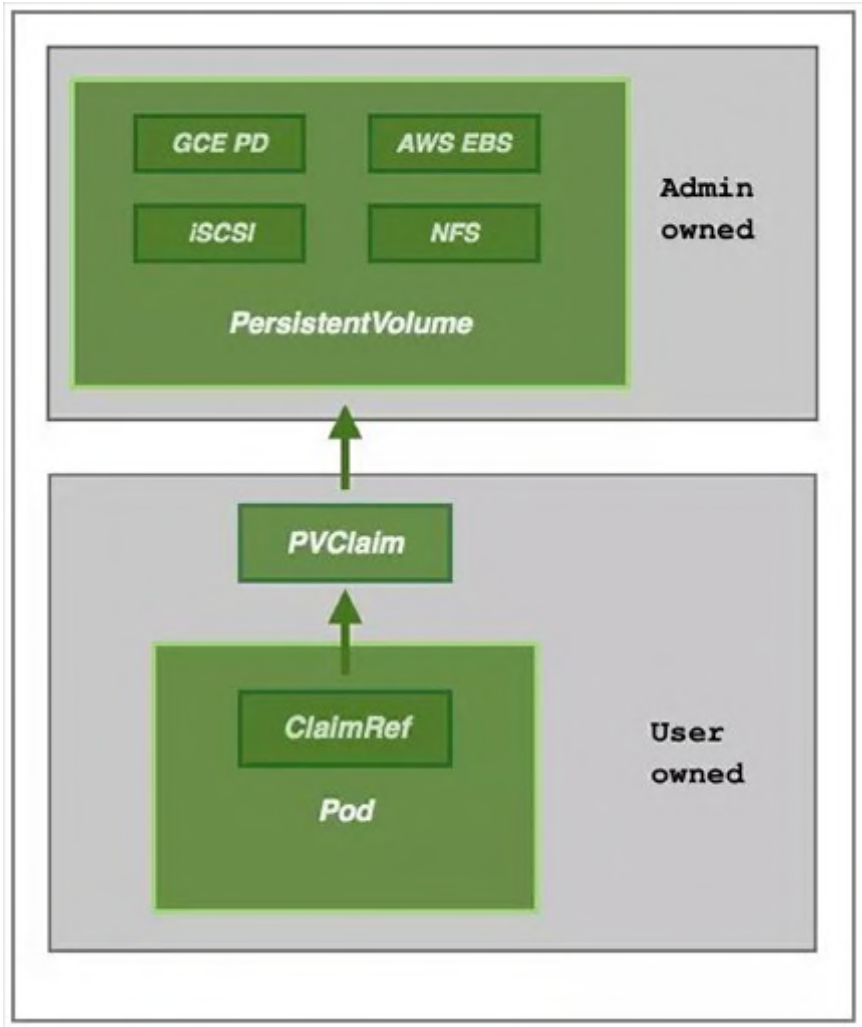
Kubernetes1.4—kubeadmin

易宝集团 YEEPAY GROUP

- 两条命令创建集群
- kubeadm init创建master
- kubeadm join 把node并入集群
- 支持apt yum源

Kubernetes1.4—动态PVC

- 对计算节点pod，屏蔽底层存储实现
- 让存储像计算资源一样抽象成：容量、iops、吞吐量
- 存储资源管理，动态匹配
- 支持分类定义、访问模式定义、回收策略等
- 未来发展空间很大



Kubernetes1.4—安全

易宝集团 YEEPAY GROUP

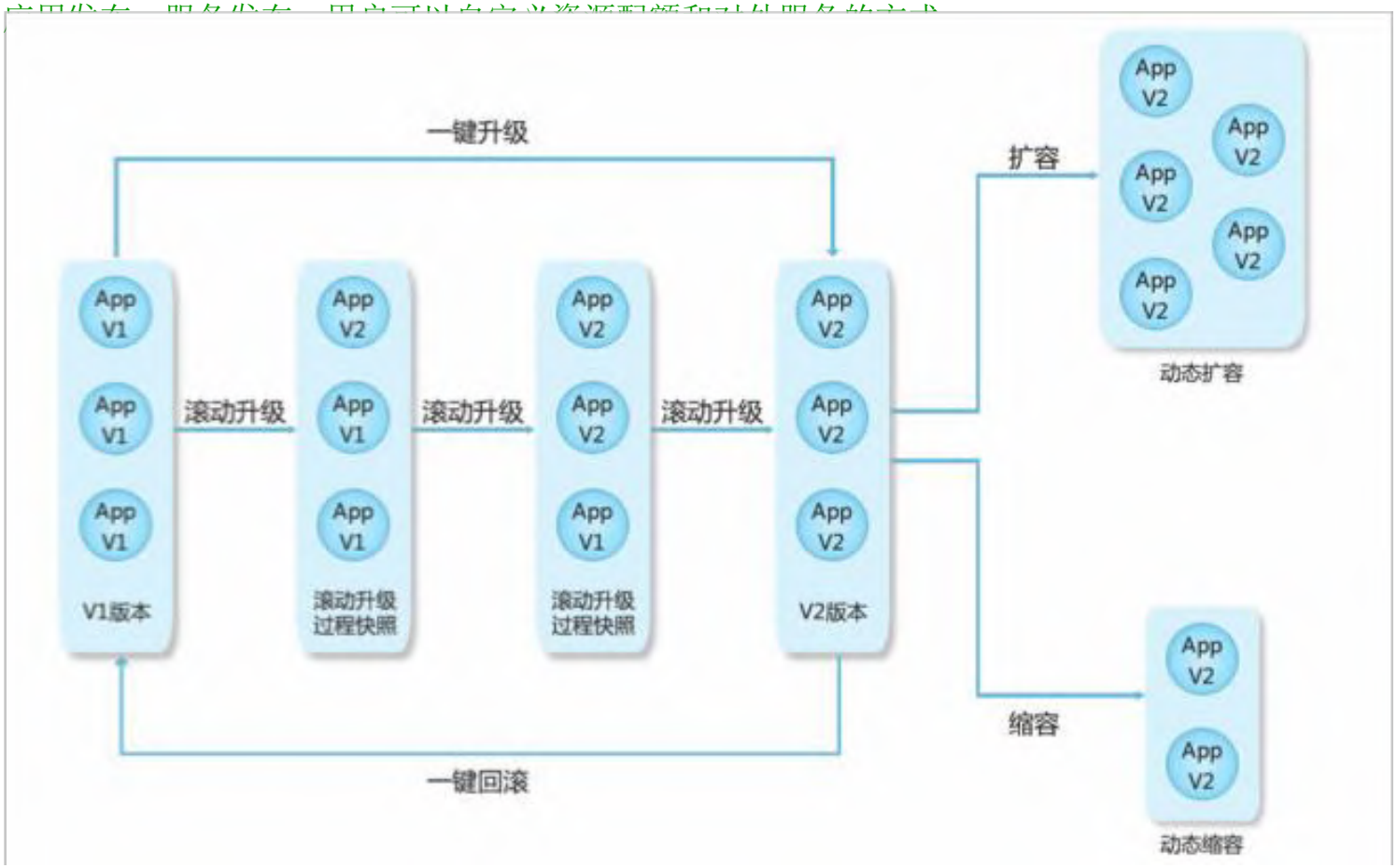
- Pod安全策略，根据用户策略、组策略定义安全上下文
- AppArmor支持，使部署安全，提供审计和监控功能

易宝容器云

- 基于K8s API
- Go web framework, IRIS

易宝容器云

- 应用全生命周期管理：一键扩容/自动扩容，一键回滚，滚动升级
- 应用发布、服务发布、用户可以自定义资源配额和对外服务的方式



易宝容器云

数据中心可视化：能够以拓扑关系图的方式展示每个应用的健康状态以及与调度到哪*台物理机上*

易宝容器云

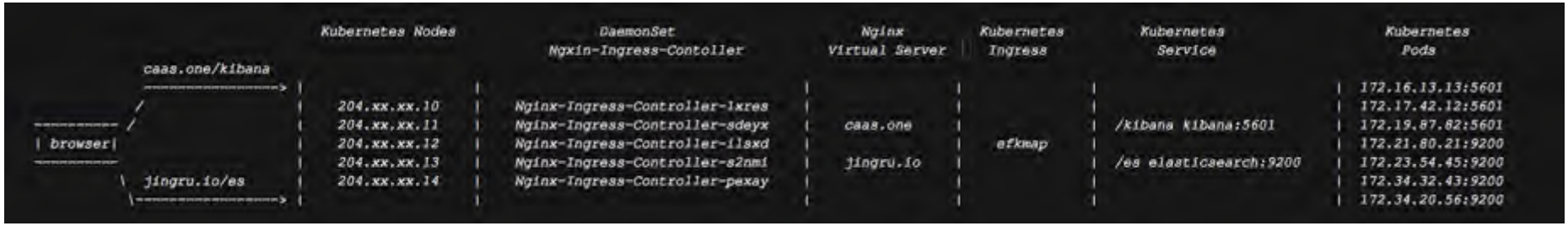
- 丰富的Dashboard功能，能过多维度，多粒度观察容器云集群和容器化应用的状态和关系
- 计算资源以组织为为单位供给配额，提供完整的从配额的购买、支付、消费一整套完整方案
- 考虑开源

容器云在生产中的落地

- 容器集群管理
  - 弹性负载均衡器：资源化、无感知对外发布
  - 内部服务发现：DNS，集群内，集群外
  - 内部服务自动负载均衡
  - 回滚到任意历史版本
  - 自动扩容、缩容（CPU使用率，冷却时间）

基础设施是代码 - - infrastructure as code

- 主机，计算资源——>pod
- 负载均衡——>Service
- Internet接入——>Ingress
- ……
- 部署、配置管理的便捷性、可重复性



不可变基础设施

易宝集团 YEEPAY GROUP

- 痛点：维护基础设施的动态和差异（系统、组件、人）
- 基础设施不可修改，以只读状态呈现
- 只能以创建新实例的方式更新
- 类似“阅后即焚”
- 部署、配置管理的便捷性、可重复性

面临的困难及解决

易宝集团 YEEPAY GROUP

- 初期解放运维，实现基础设施的自动化
- 边缘突破原则
- 无状态应用
  －session落到后端
  －本地文件放到对象存储
- 日志流式输出
- 使用configmap实现配置管理
- CI/CD集成

后期工作

易宝集团 YEEPAY GROUP

- 推动应用上云
- 细化容器管理平台
- 从磨合到融合
- 实践Devops
- 沉淀－实践－沉淀－实践，到形成文化

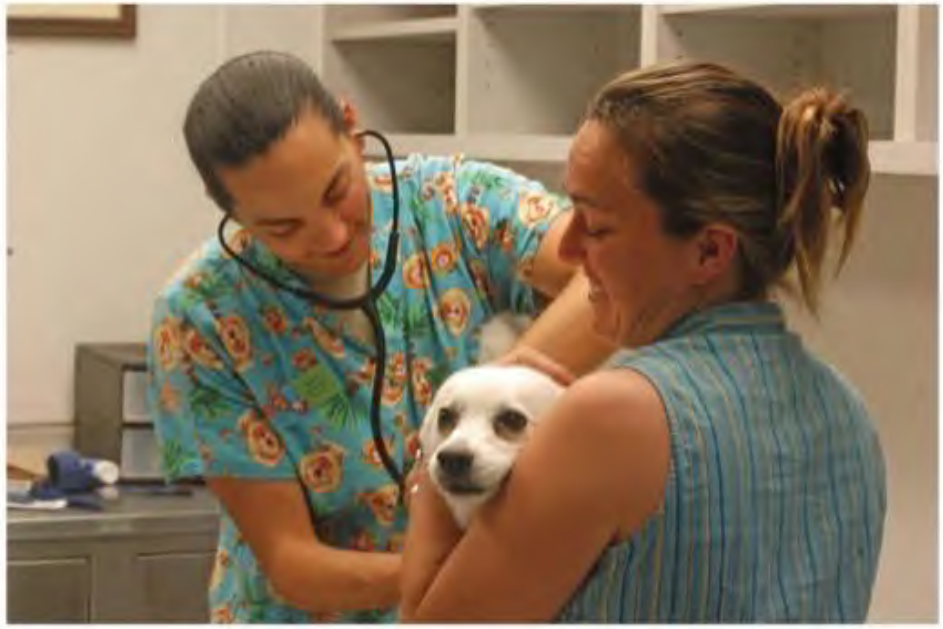
未来展望

易宝集团 YEEPAY GROUP

- 计算密度加强、服务器高配、数量减少
- 底层透明化
- 应用技术栈百花齐放，语言之争声音变小
- 应用微服务化，运维机械化. .

未来展望

- 程序从宠物到牲畜
- 监控数据以metric输出
- 运维从细致到粗放
- 推荐的运维工具：start/stop/restart, kill -9
- 重启大法好...



欢迎关注《容器时代》

易宝集团
YEEPAY GROUP

Thank You!

谢谢！