

智能时代的新运维

CNUTCon

全球运维技术大会

携程的容器化交付实践

大浪

携程研发经理

主办方 **Geekbang** **InfoQ**
极客邦科技

目录

- PaaS 概况
- PaaS 平台的容器交付
- 容器的镜像管理
- Mesos 迁移 K8s

PaaS 概況

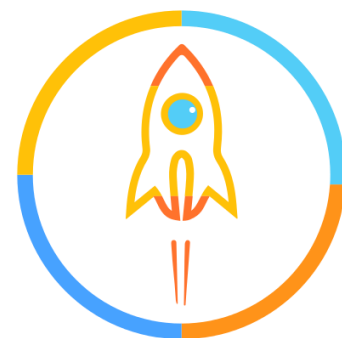
Overview



9000+ 应用
60% 容器

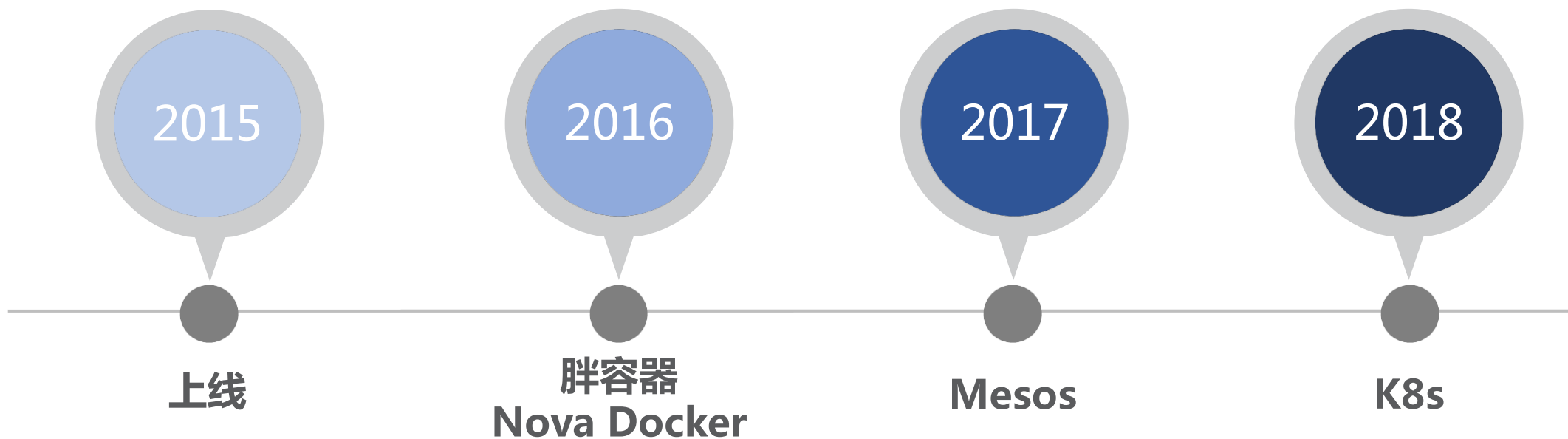


3000+ 研发
1500+ DAU

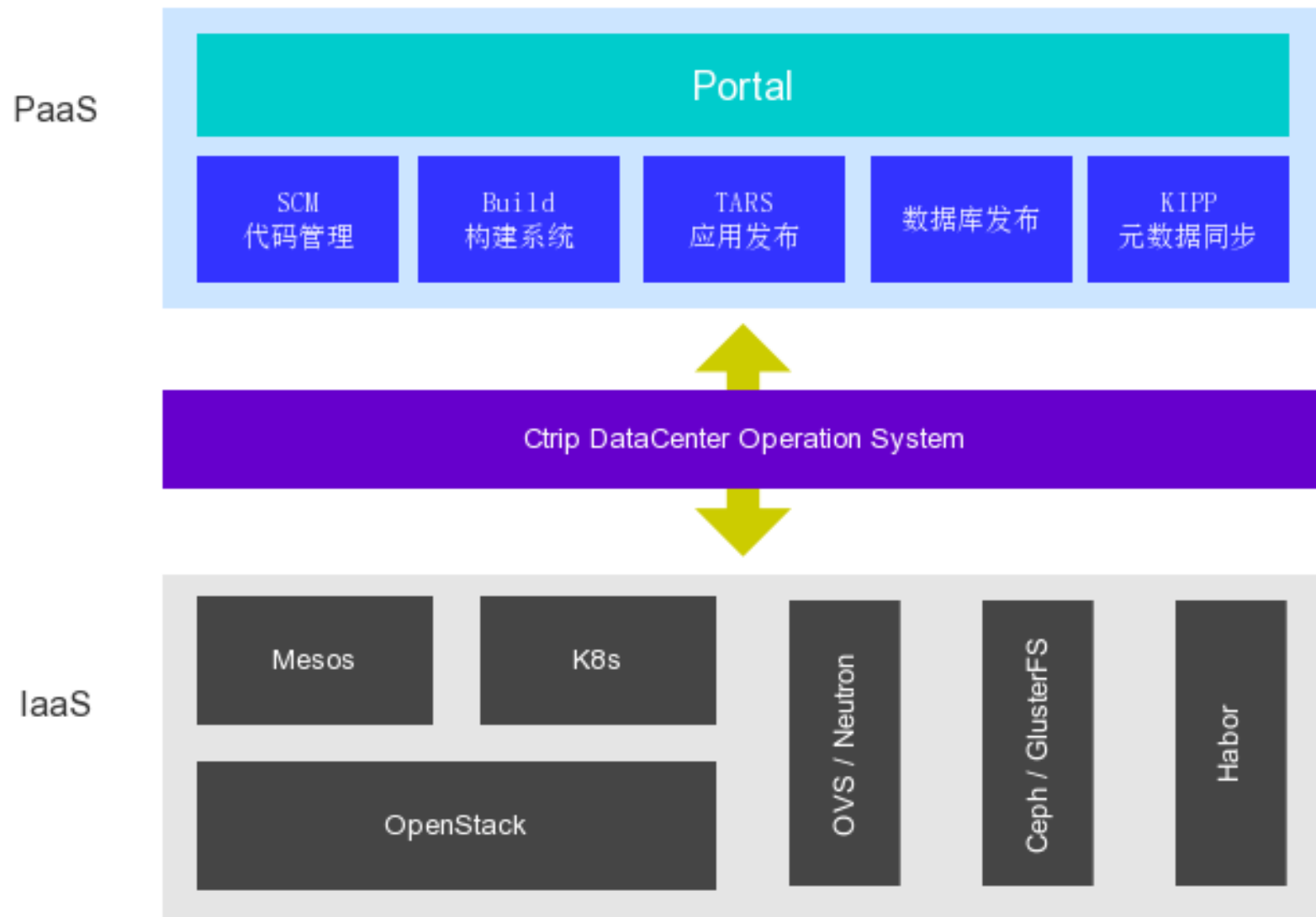


6000+ 发布/天
2/3 容器

容器实践 Milestone

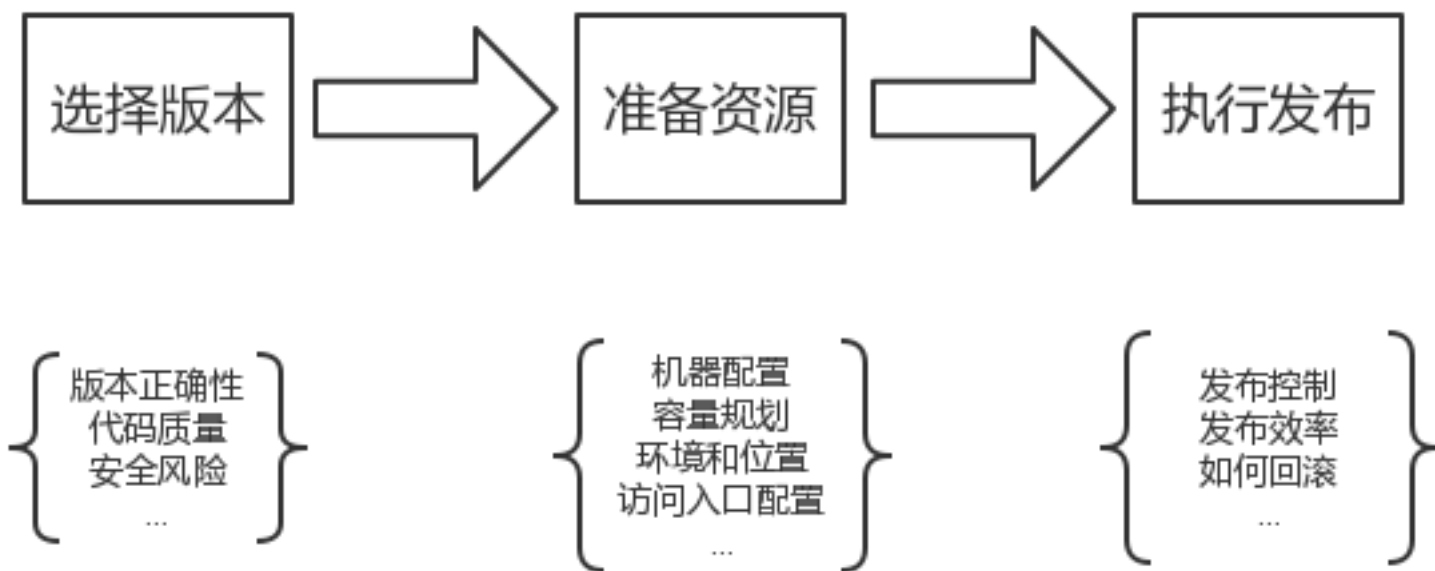
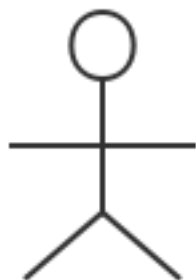


架构

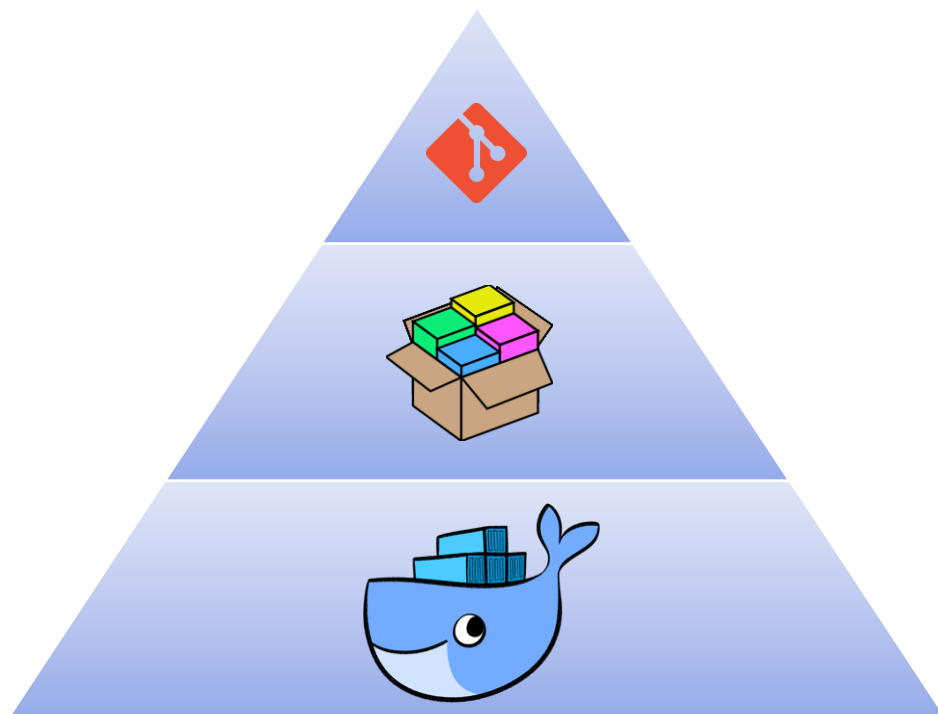


PaaS 平台的容器交付

交付



版本的演进



Git Tags

源码

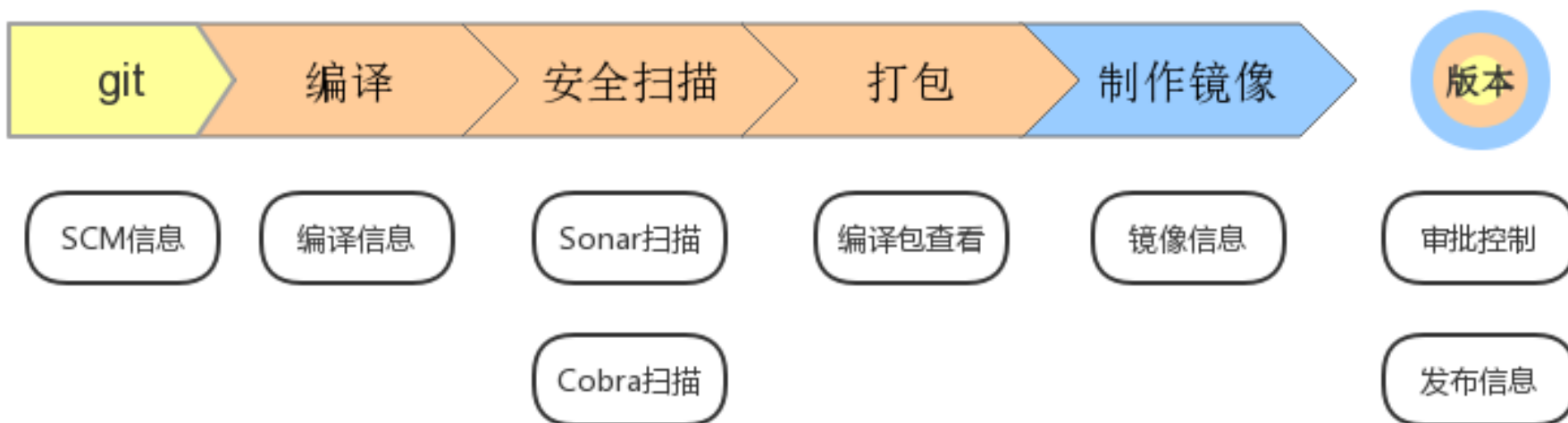
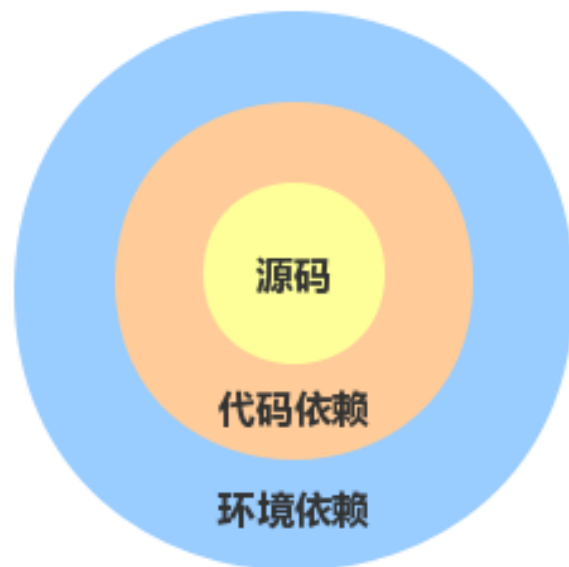
Packages

源码 + 代码依赖

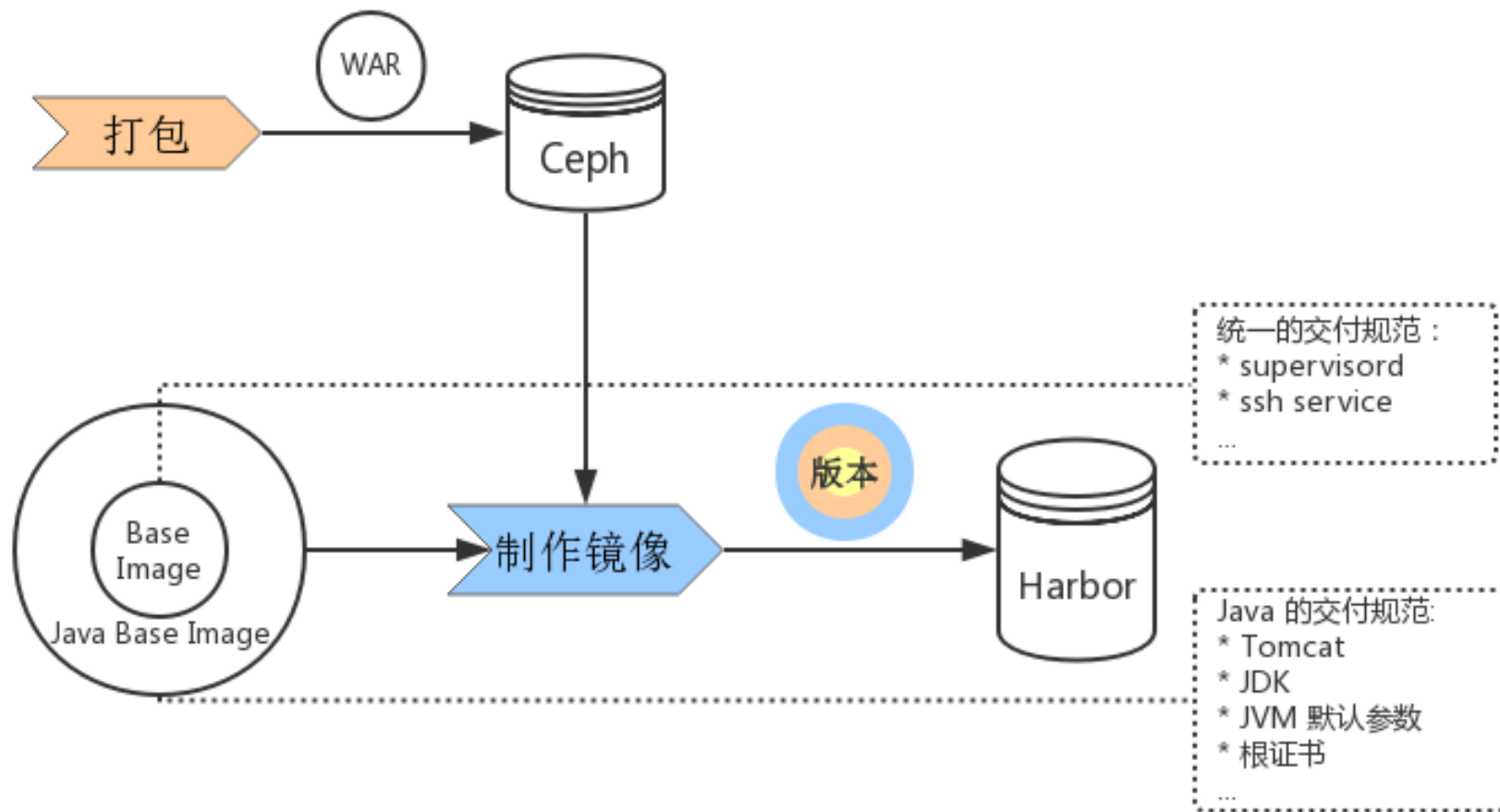
Images

源码 + 代码依赖 + 环境依赖

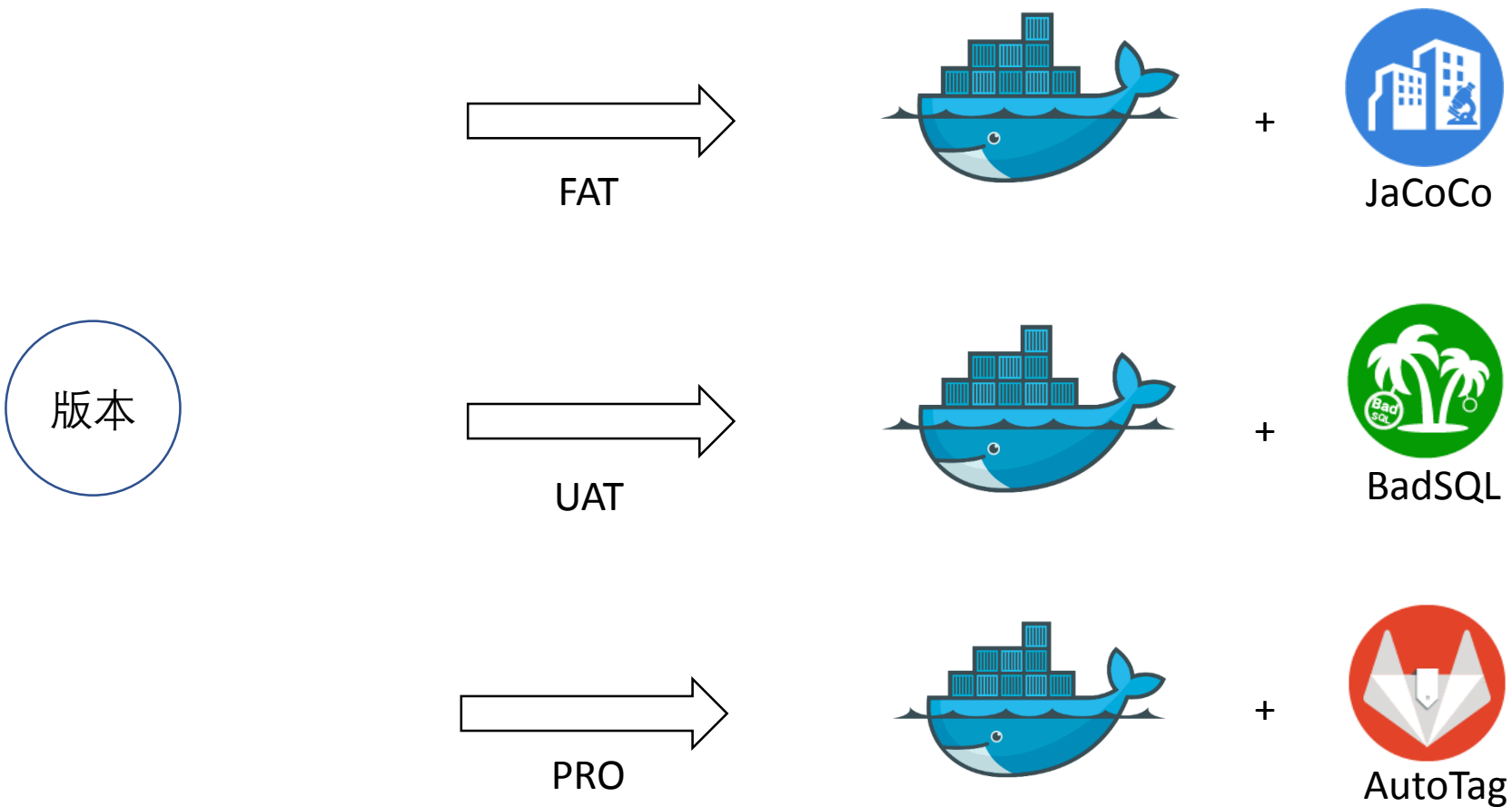
版本管理



制作镜像



分环境交付镜像



服务集市的解决方案

```
FROM scratch
```

```
LABEL name="CentOS Base Image" \  
  vendor="CentOS" \  
  license="GPLv2" \  
  build-date="20171128"
```

```
ADD centos-7-docker.tar.xz
```

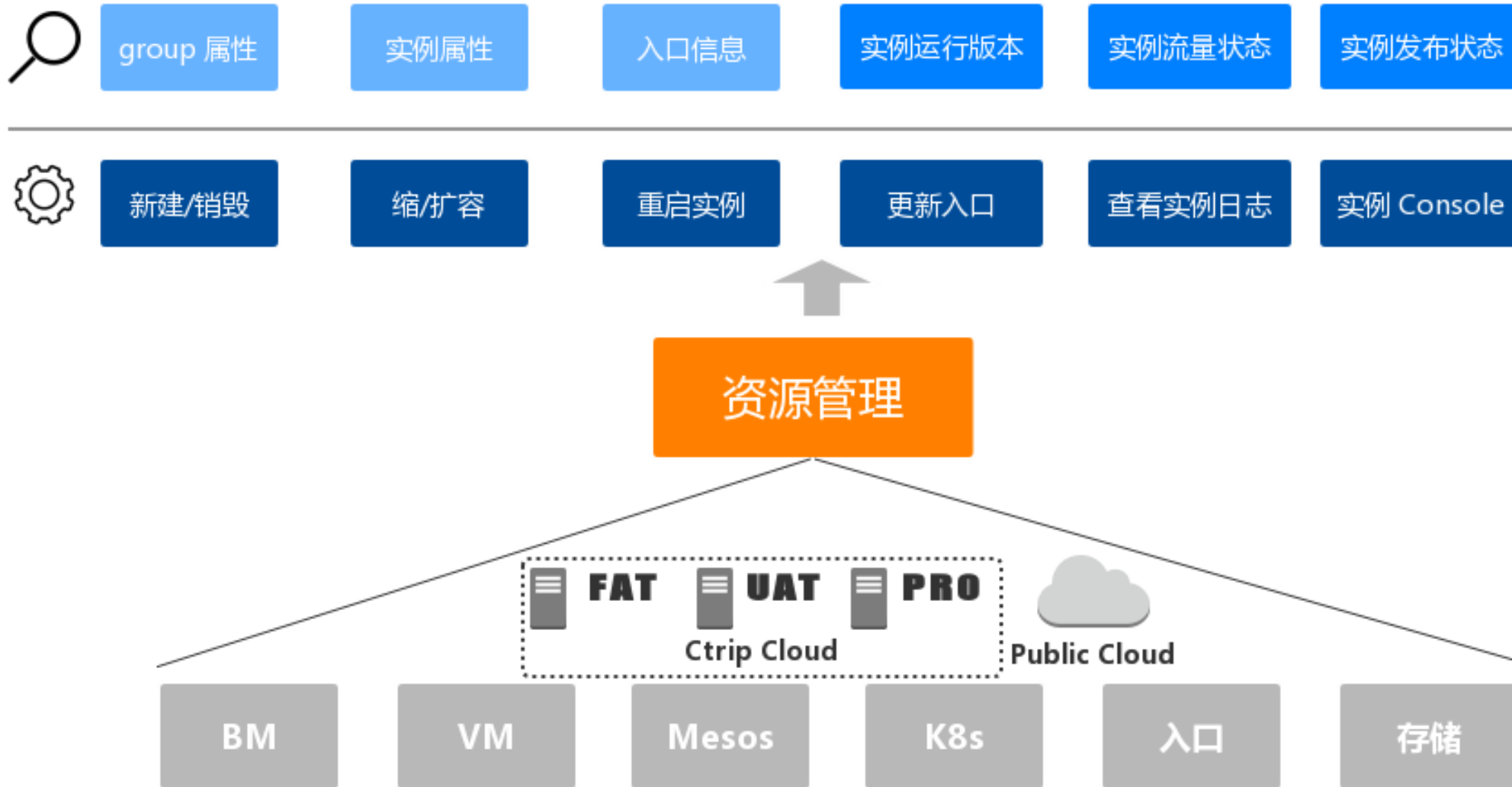
```
CMD ["/bin/bash"]
```

集市服务数量: 12
开通服务数量: 5000+



```
...  
RUN copy jacoco.jar  
&& bash  
jacocoagent.sh
```

资源管理



全环境 全类型 一站式的管理平台

日志集成

- 日志类型

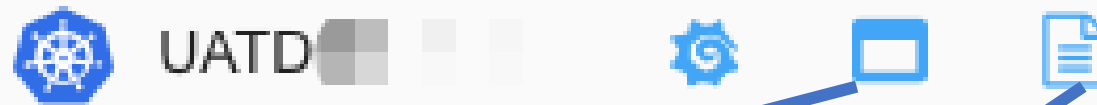
- tomcat catalina gc access.log ...
- stdout stderr

- 日志收集

- mount 宿主机 + filebeat

- 日志清理

- logrotate
- xfs project quota



Web Console

```
Web Console 10.5.170.18
sys-fordocker-test-service [用于docker service测试]
testgroup@0727_100006124
10.5.170.18
UATD@101693
Command + c
Command + v
ip@ip:~$
ip@ip:~$ cd /opt/logs/100006124
ip@ip:~/logs$ ls
ip@ip:~/logs$ cat catalina.out
ip@ip:~/logs$ cat gc-20180727134826.log
ip@ip:~/logs$ cat host-manager.log
ip@ip:~/logs$ cat localhost.log
ip@ip:~/logs$ cat manager.log
ip@ip:~/logs$
```

Log Viewer

日志文件

选择日志文件

/opt/logs/100006124

- applog
- catalina.log
- catalina.out
- gc-20180727134826.log
- host-manager.log
- localhost.log
- manager.log

查看日志

catalina.log 下载文件

关键字查找

Jul 27, 2018 1:48:27 PM org.apache.catalina.startup.VersionLoggerListener log
INFO: Server version: Apache Tomcat/7.0.73
Jul 27, 2018 1:48:27 PM org.apache.catalina.startup.VersionLoggerListener log
INFO: Server built: Nov 7 2016 21:27:23 UTC
Jul 27, 2018 1:48:27 PM org.apache.catalina.startup.VersionLoggerListener log
INFO: Server number: 7.0.73.0
Jul 27, 2018 1:48:27 PM org.apache.catalina.startup.VersionLoggerListener log
INFO: OS Name: Linux
Jul 27, 2018 1:48:27 PM org.apache.catalina.startup.VersionLoggerListener log
INFO: OS Version: 4.10.13-1.el7.elrepo.x86_64
Jul 27, 2018 1:48:27 PM org.apache.catalina.startup.VersionLoggerListener log
INFO: Architecture: amd64
Jul 27, 2018 1:48:27 PM org.apache.catalina.startup.VersionLoggerListener log
INFO: Java Home: /usr/java/jdk1.8.0_144/jre
Jul 27, 2018 1:48:27 PM org.apache.catalina.startup.VersionLoggerListener log
INFO: JVM Version: 1.8.0_144-b01
Jul 27, 2018 1:48:27 PM org.apache.catalina.startup.VersionLoggerListener log
INFO: JVM Vendor: Oracle Corporation
Jul 27, 2018 1:48:27 PM org.apache.catalina.startup.VersionLoggerListener log
INFO: CATALINA_BASE: /opt/tomcat
Jul 27, 2018 1:48:27 PM org.apache.catalina.startup.VersionLoggerListener log

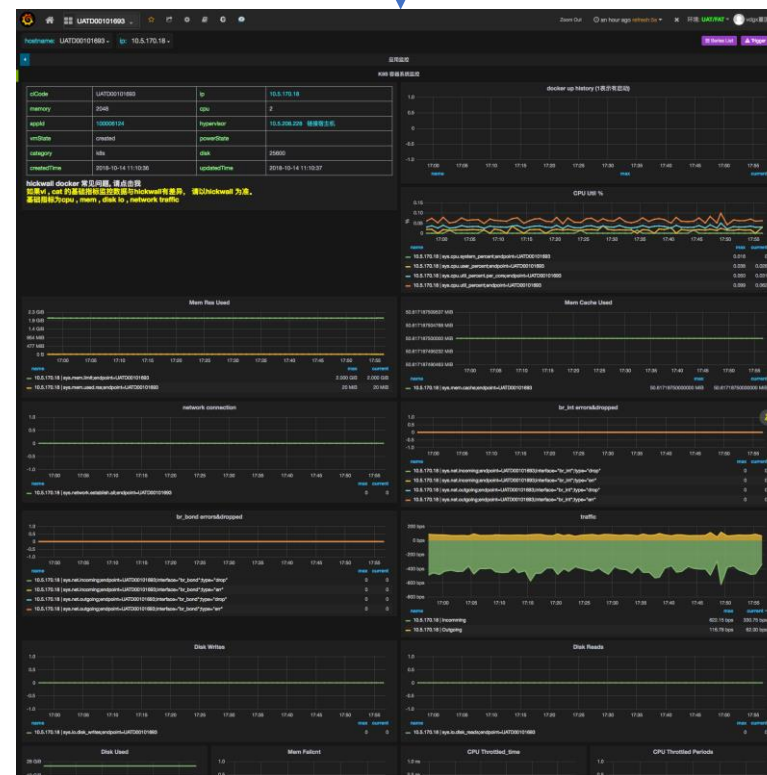
< 返回

监控集成

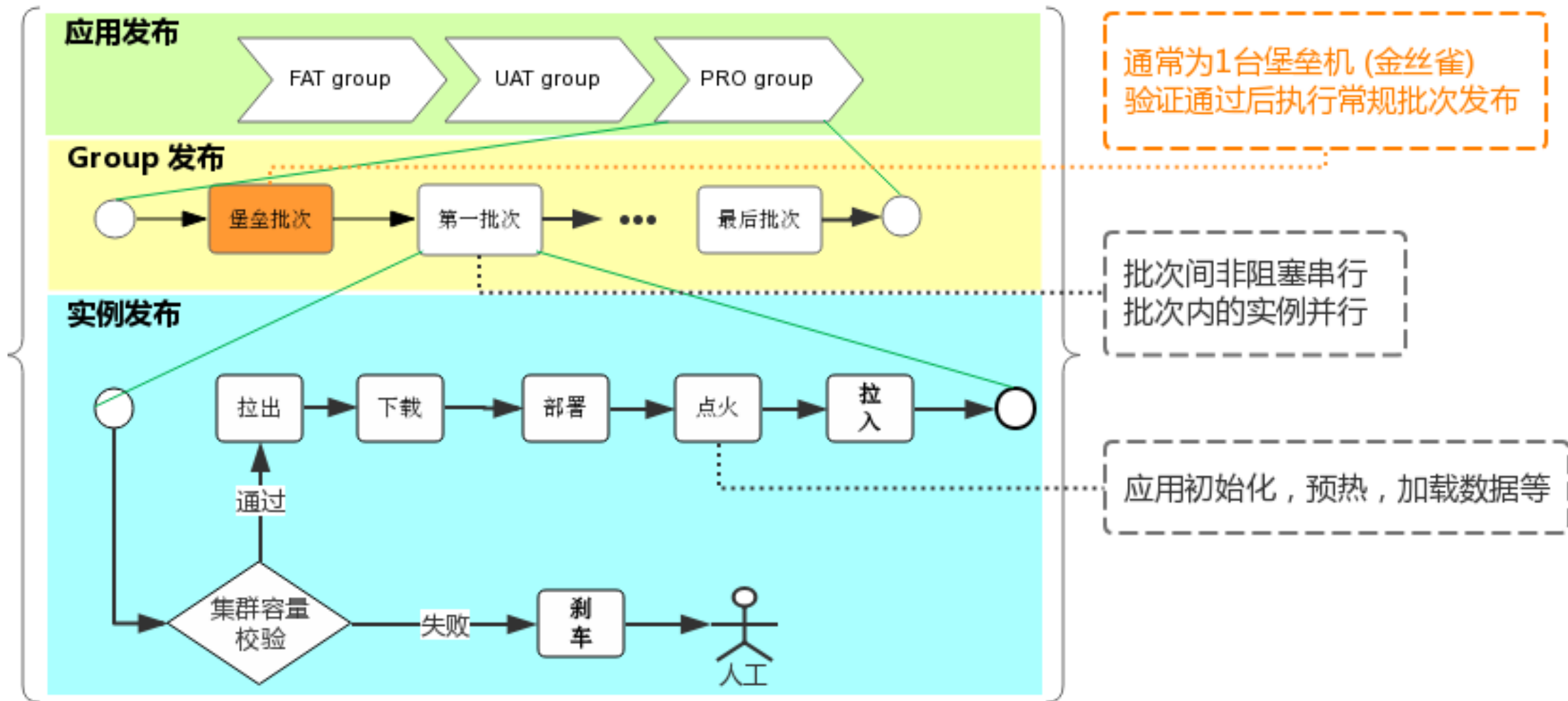
- Hickwall (携程的监控平台)
 - Hickwall server -> kubelet API
- kubelet API
 - cadvisor: /stats/xxx, 环境变量
- OOM
 - docker: cgroup notification



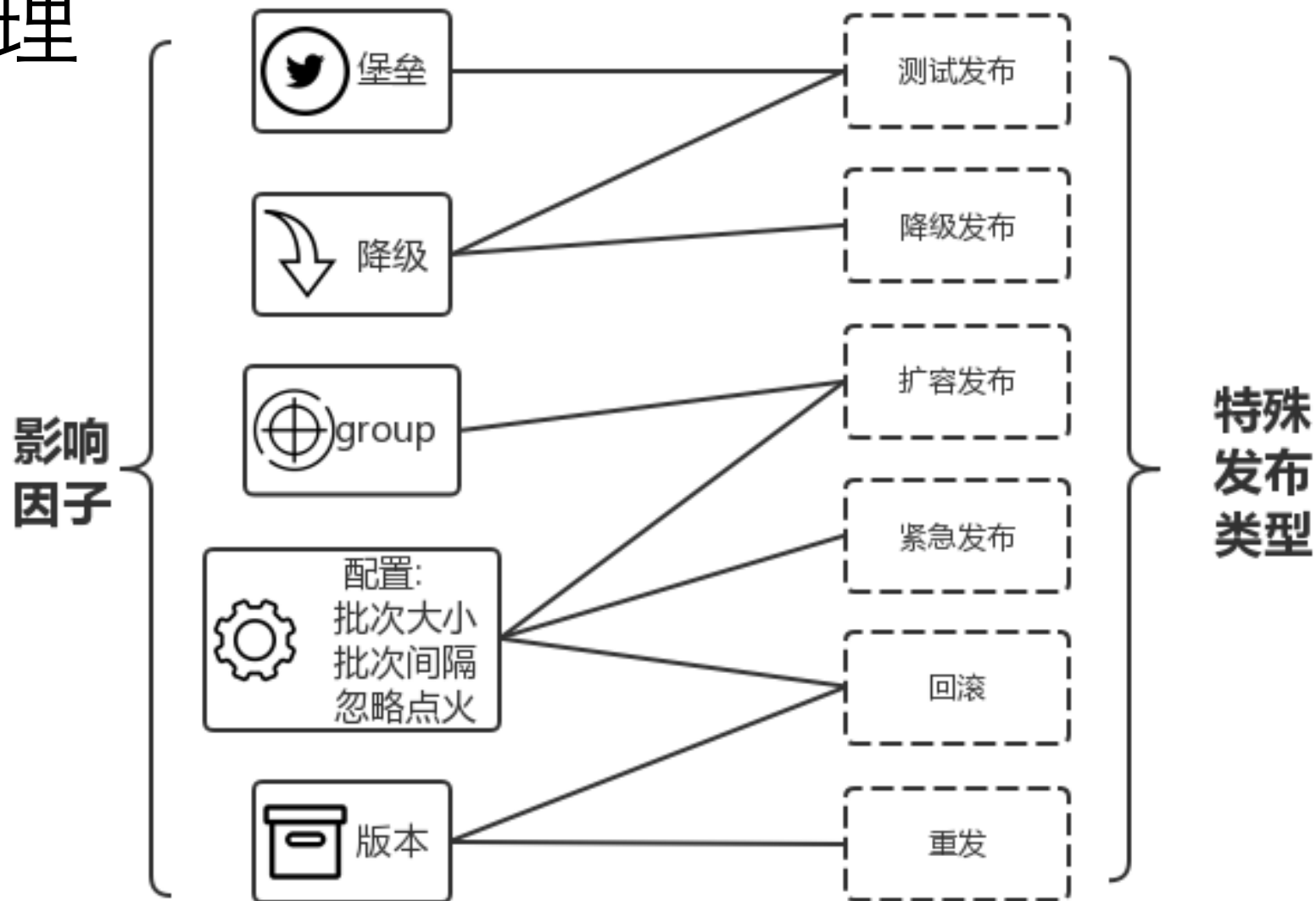
UATD



发布模式 分环境的带堡垒批次滚动发布



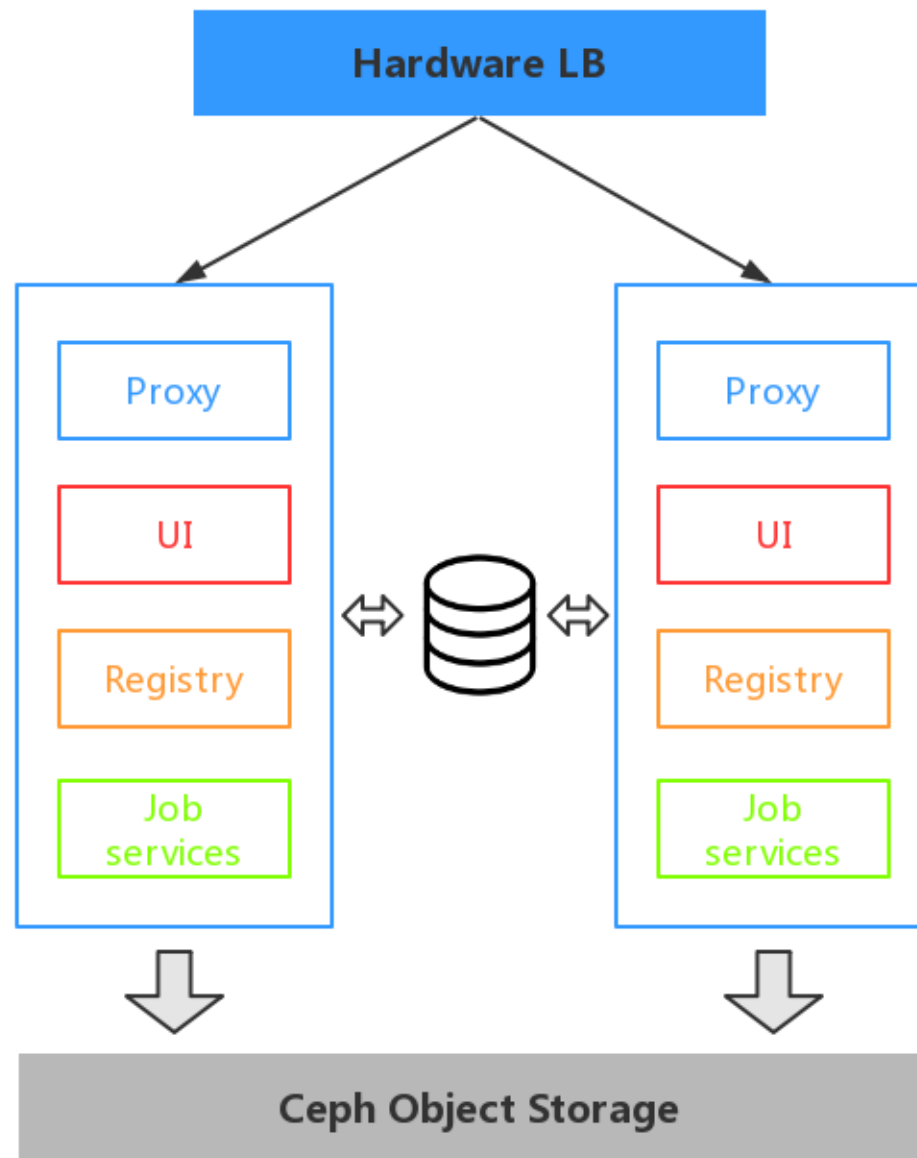
发布管理



容器的镜像管理

Harbor 单集群架构

- HA Solution
- One region One cluster

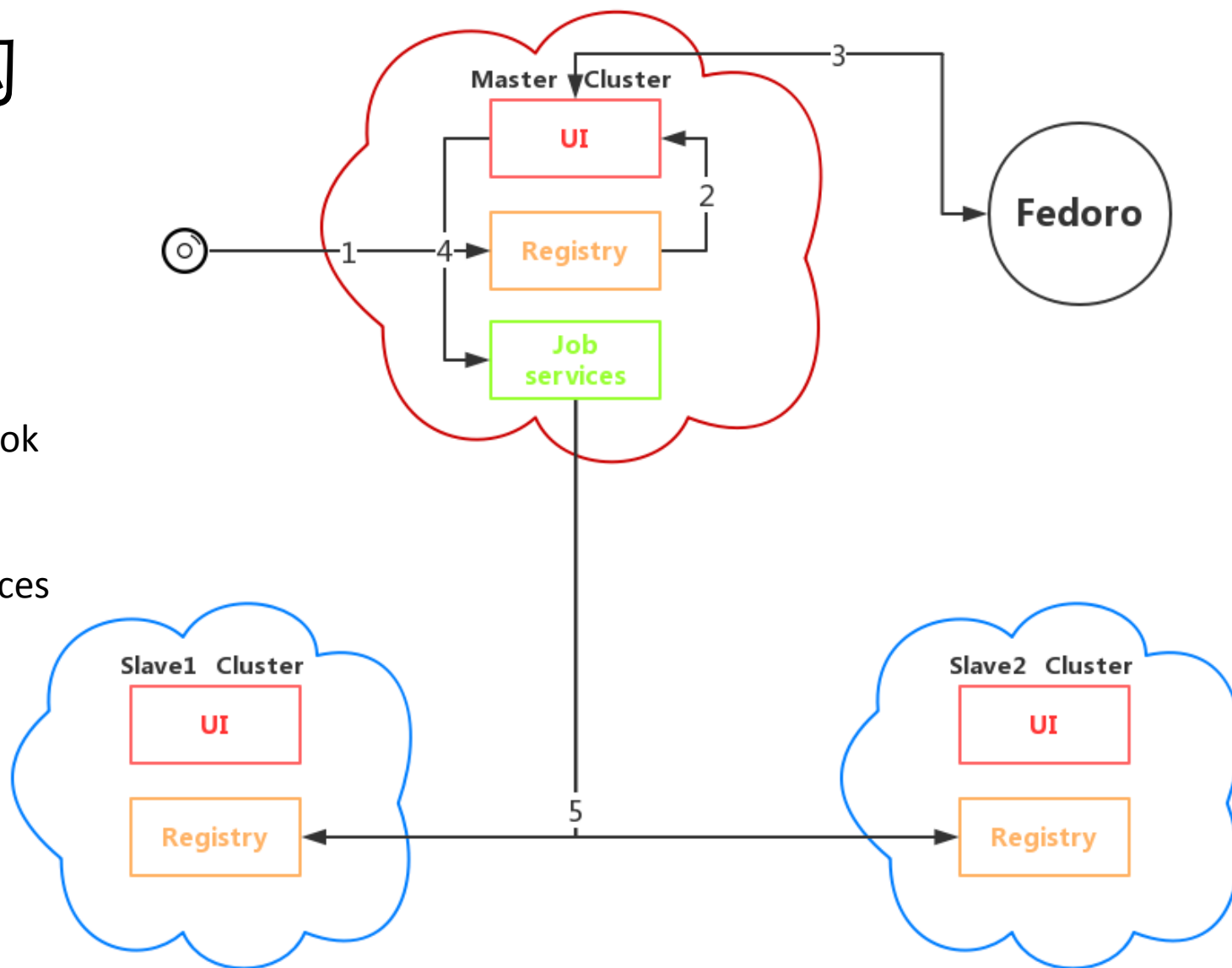


Harbor 联邦架构

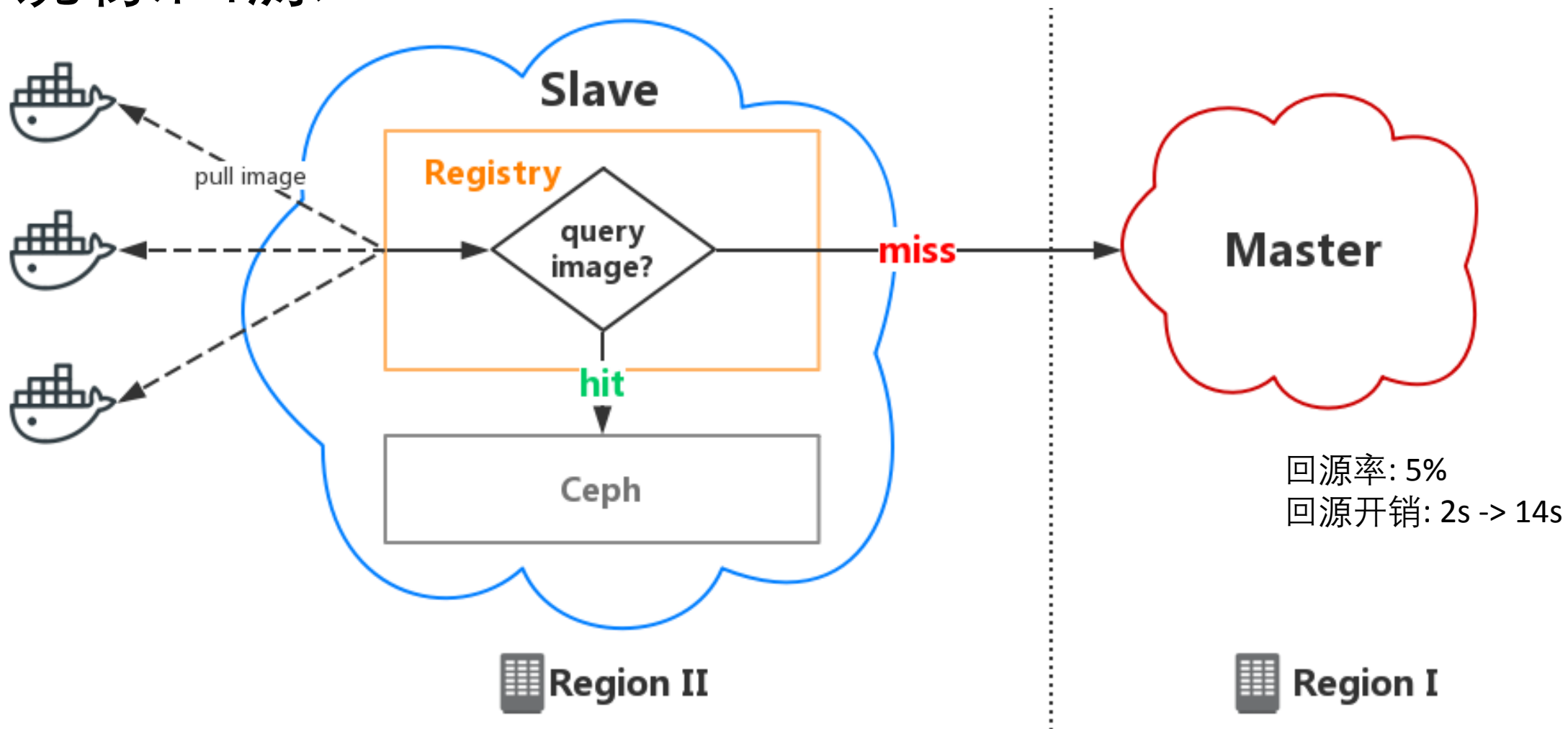
Cross-IDC Image Sync

1. Push image to harbor
2. Populate event to UI via Webhook
3. Retrieve slave info from Fedoro
4. Commit syncing job to Job services
5. Sync image to slaves

Success Rate: 99.7%
Total Image: 3500 / day
SLA: < 20s

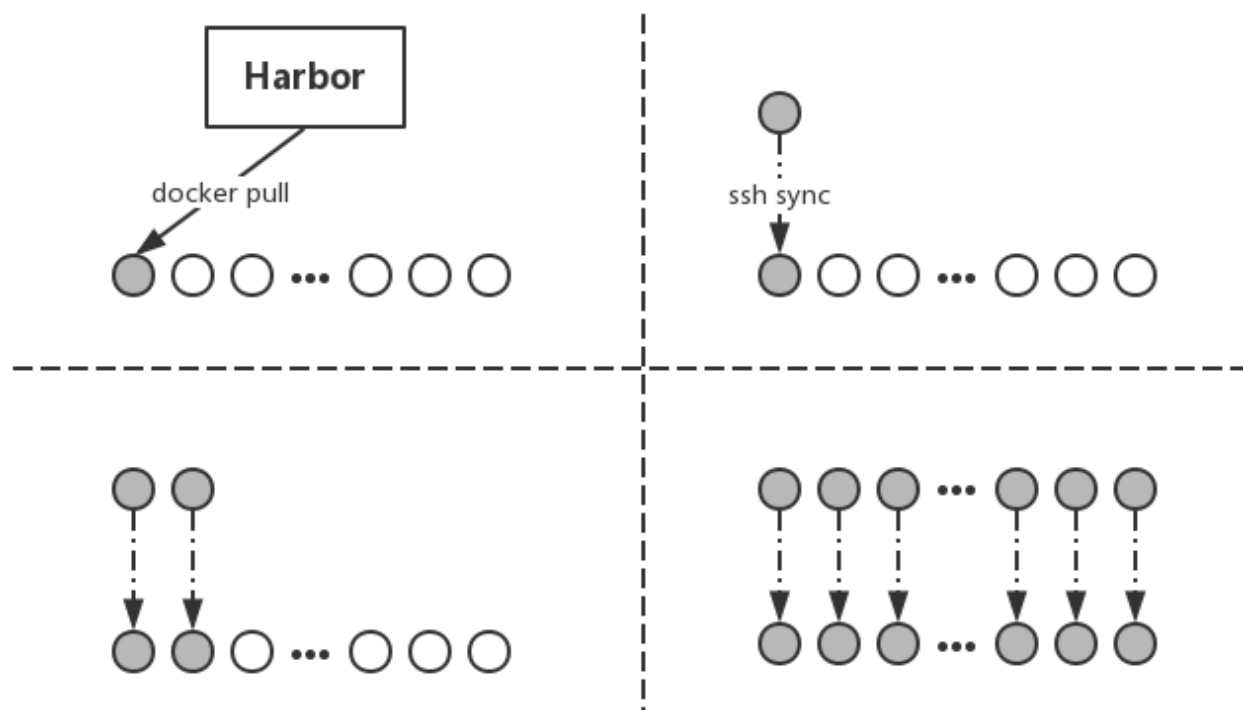


镜像回源



基础镜像分发

- 去中心化的分发策略



- 灰度发布
 - 分环境
 - 指定应用比例的灰度
 - 白名单

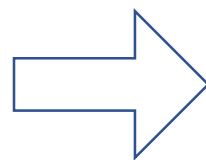
Mesos 迁移 K8s

为什么要迁移

	已解决	待解决
应用交付	无状态应用的交付	有状态应用的交付
资源弹性	手动扩缩容	自动弹性伸缩

为什么要迁移 cont.

资源提供者



服务提供者

- 只提供容器实例，无编排
- 发布系统需要负责网络实例的状态管理
- 用户关注实例、IP

- 提供更高层次的抽象
- 屏蔽 IaaS 层的资源管理
- 用户关注服务、策略

K8s 的选型



- Deployment 无状态应用
- StatefulSet 有状态应用
- CRD Operator



- 全部固定 IP
- 应用有状态/无状态边界模糊
- 保持一致的发布体验

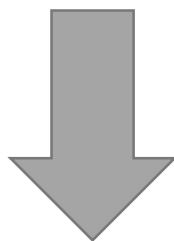
Stateless on StatefulSet

迁移方案

存量Mesos应用迁移

单cluster容量瓶颈

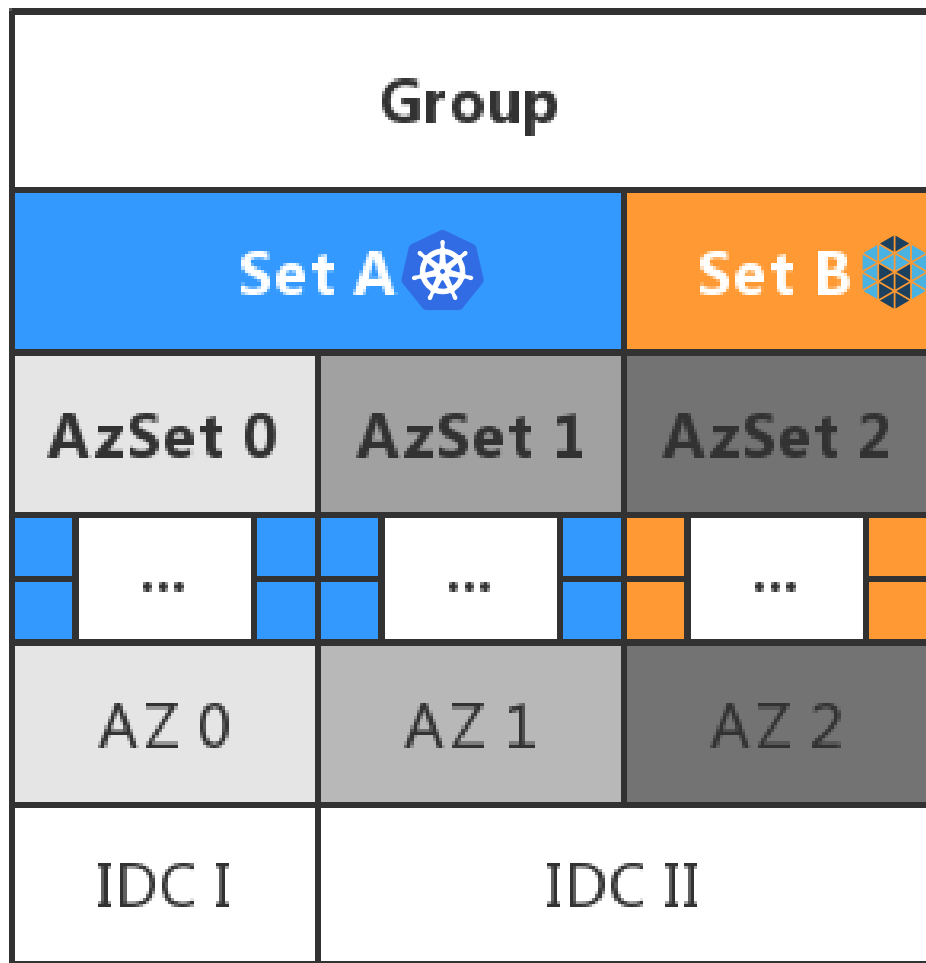
K8s可用性风险



兼容 Mesos K8s

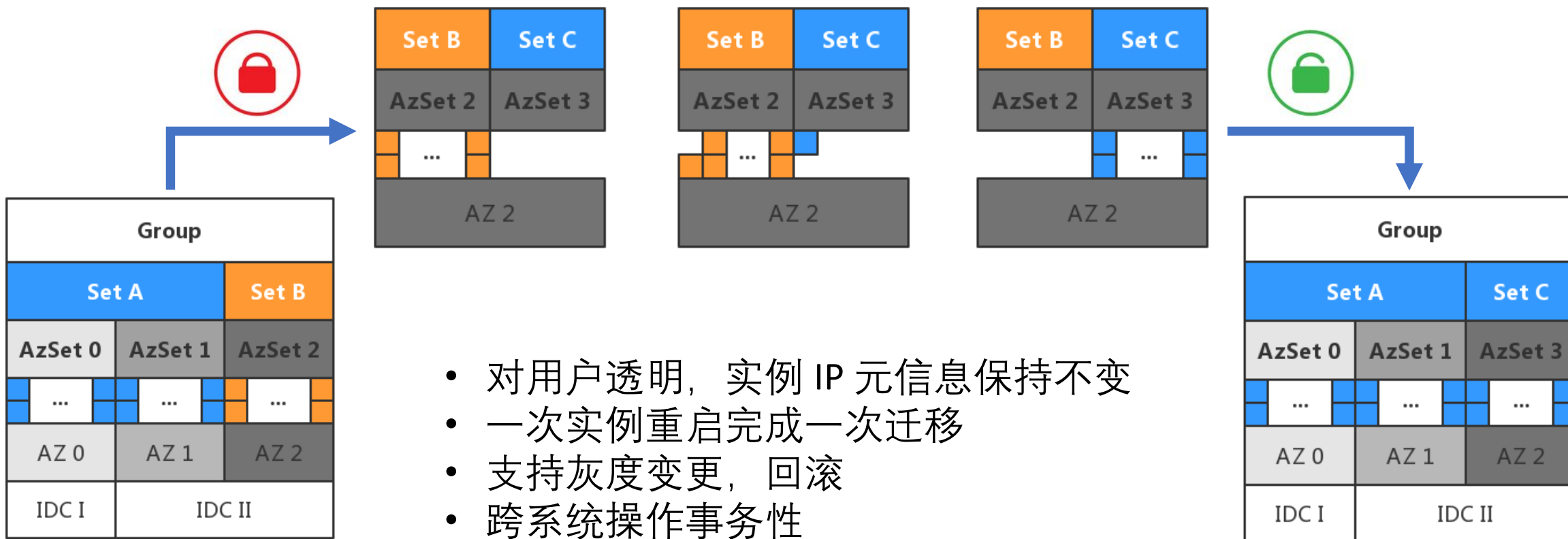
单 group 跨多 cluster

K8s-App 架构



- 属于同个应用的一组实例，发布的最小单位
- 一个 Set 对应一个 Spec 版本
- 一个 AzSet 对应一个 StatefulSet 同时在 K8s 和 Mesos 做了实现
- 根据 Set 上记录的 Spec 信息创建的实例
- AZ 对应网络的边界
- 跨 IDC 部署

迁移过程

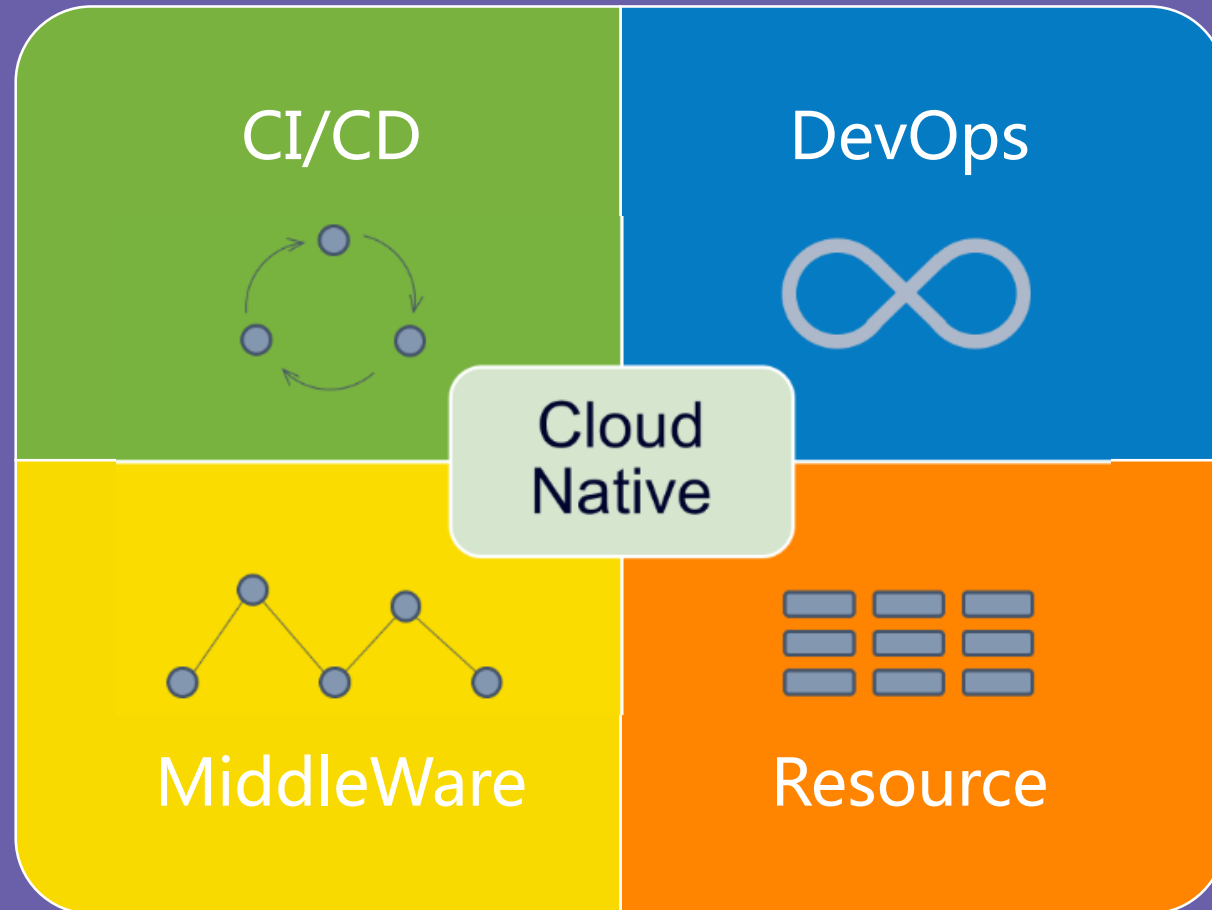


实施策略

- 同AZ的Mesos到K8S迁移按物理机迁移
 - 使用 stackstorm 自动化 pipeline
-

- 一台宿主机迁移 < 10min
- 正在迁移生产最后一个Mesos集群

What's Next



谢谢