

携程容器云优化与实践

王潇俊

QCon

全球软件开发大会

10月17-19日 上海·宝华万豪酒店



扫码锁定席位

九折即将结束

团购还享更多优惠，折扣有效期至9月17日

扫描右方二维码即可查看大会信息及购票



如果在使用过程中遇到任何问题，可联系大会主办方，欢迎咨询！

微信：qcon-0410

电话：010-84782011

ArchSummit

全球架构师峰会 2017



扫码锁定席位

12月8-9日 北京·国际会议中心

七折即将截止立省2040元

使用限时优惠码AS200，

以目前最优惠价格报名ArchSummit

仅限前20名用户，优惠码有效期至9月19日，

扫描右方二维码即可使用



如果在使用过程中遇到任何问题，可联系大会主办方，欢迎咨询！

微信：aschina666

电话：15201647919

极客搜索

全站干货，一键触达，只为技术

s.geekbang.org



扫描二维码立即体验

有没有一种搜索方式，能整合 InfoQ 中文站、极客邦科技旗下12大微信公众号矩阵的全部资源？

极客搜索，这款针对极客邦科技全站内容资源的轻量级搜索引擎，做到了！

扫描上方二维码，极客搜索！

这里只有 技术领导者

EGO会员第二季招募季正式开启



—— E小欧 ——

报名时间：9月1日-9月15日
扫描添加E小欧，
邀您进入EGO会员预报名群

立即报名



关于我

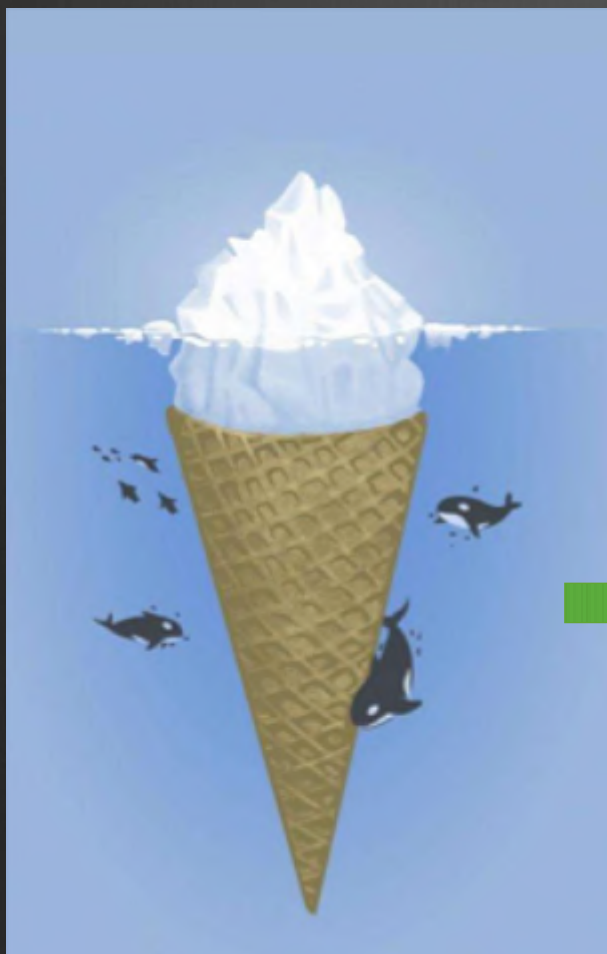
- 携程系统研发部
- 正在经历容器化的过程
- 多年专注于微服务，持续交付，容器，DevOps等领域

DevOps

- 打破混乱之墙，而非塑造全能之神
- 专业分工仍旧存在
- 水下是基础，水上是服务



水下部分



- 容器云的主体架构

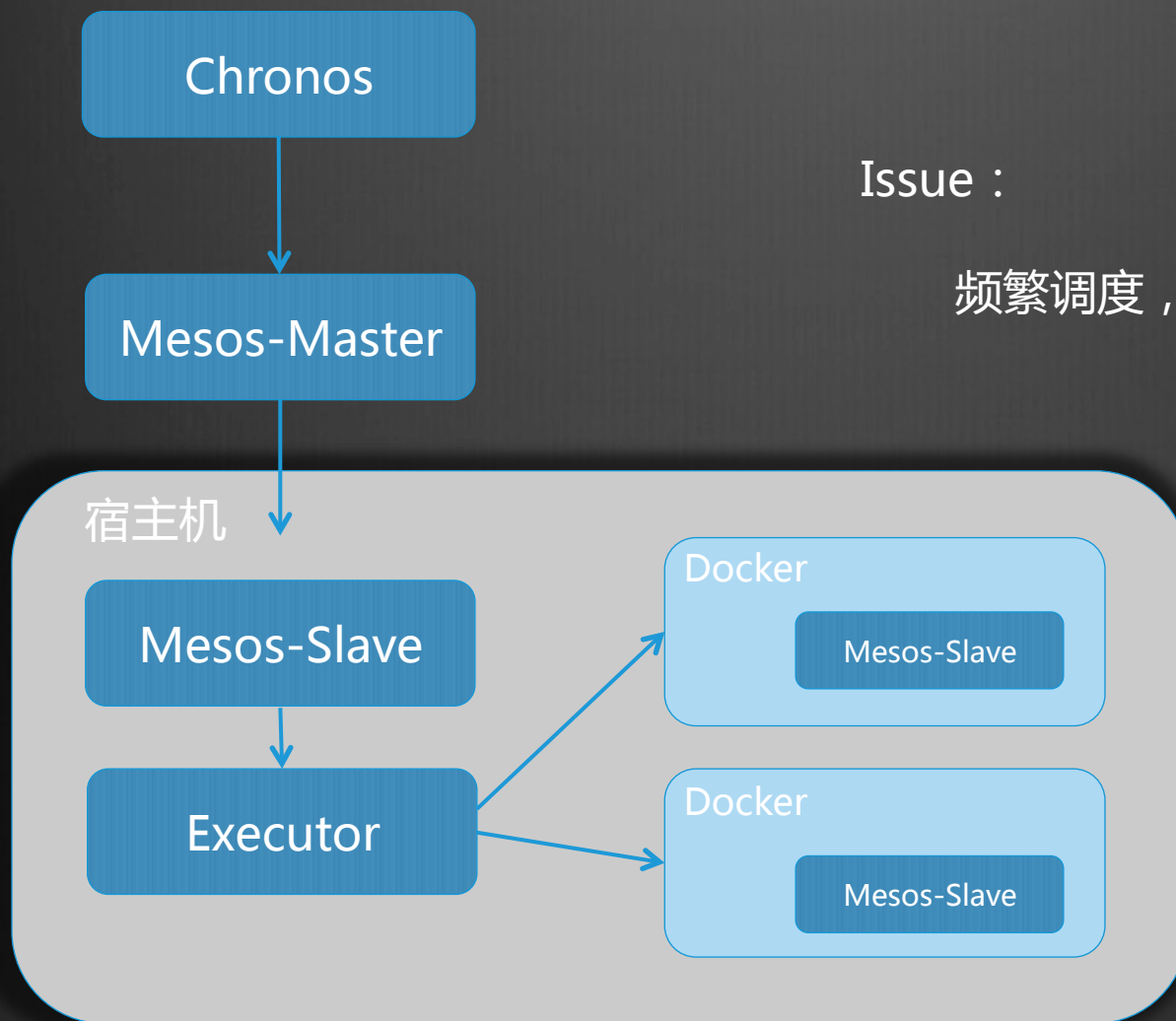
主要阶段

类VM，通过
Openstack进
行管理

镜像发布，使
用出Chronos
跑Job

镜像发布，自
研 Mesos
Framework

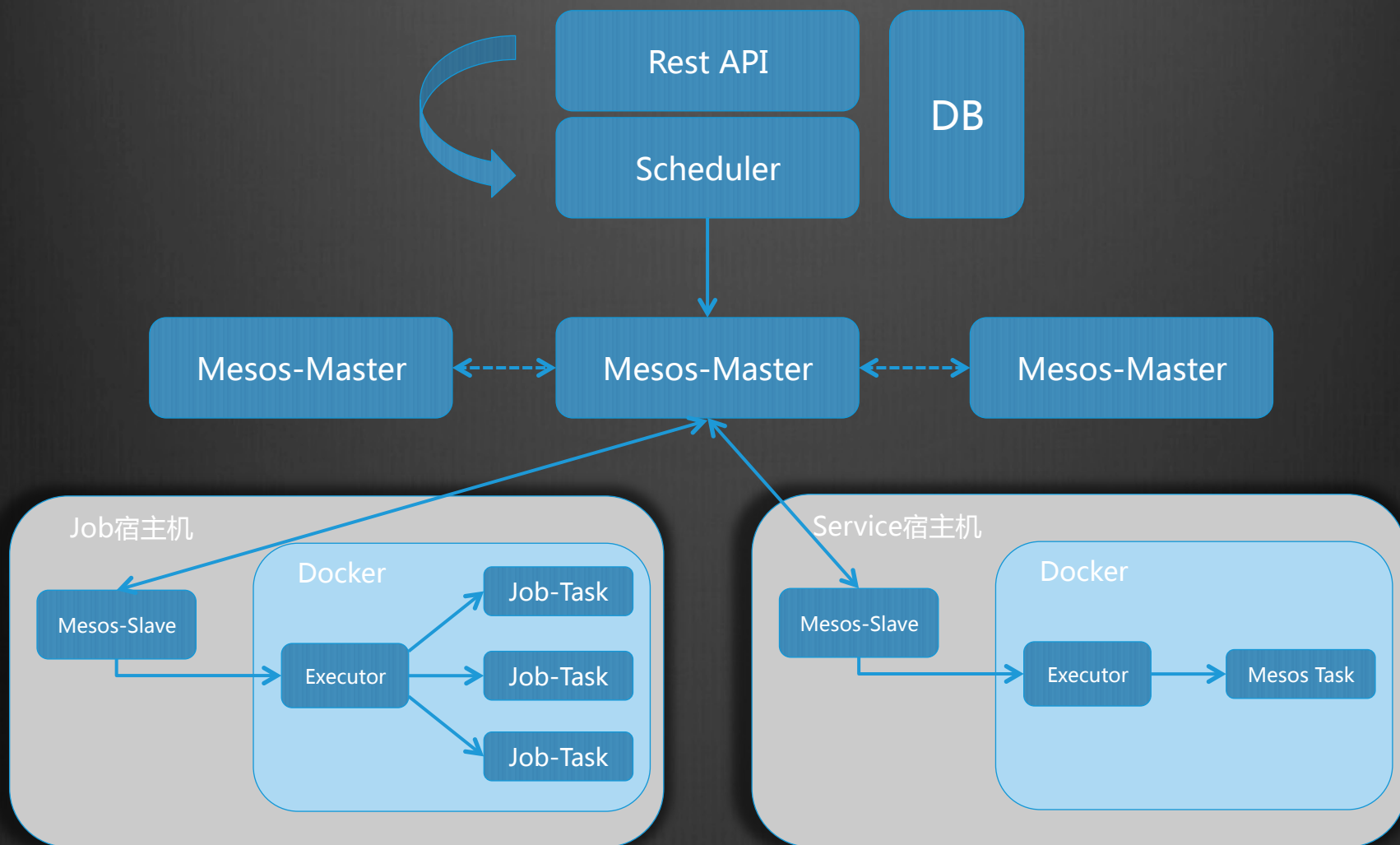
二代架构



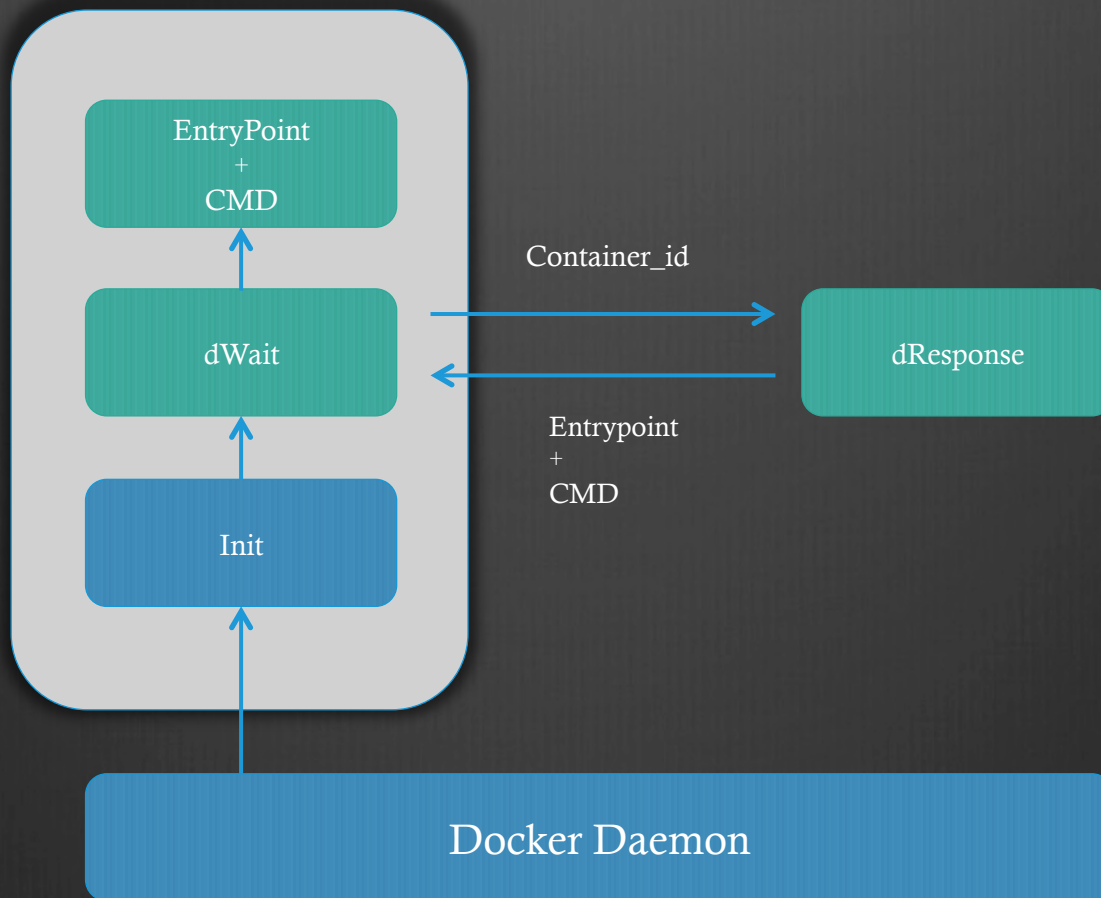
Issue :

频繁调度，消耗高

自研 Framework



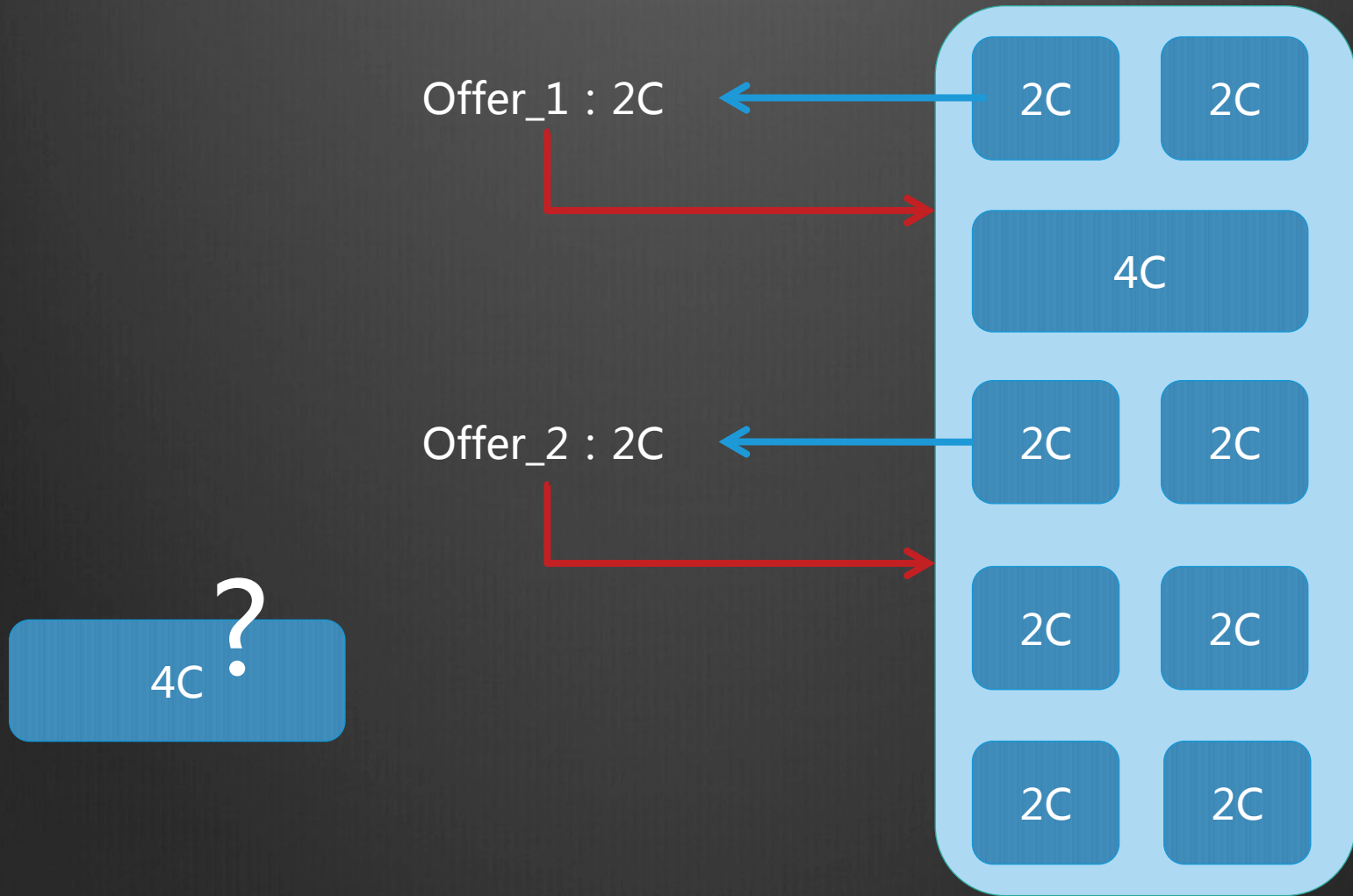
网络处理



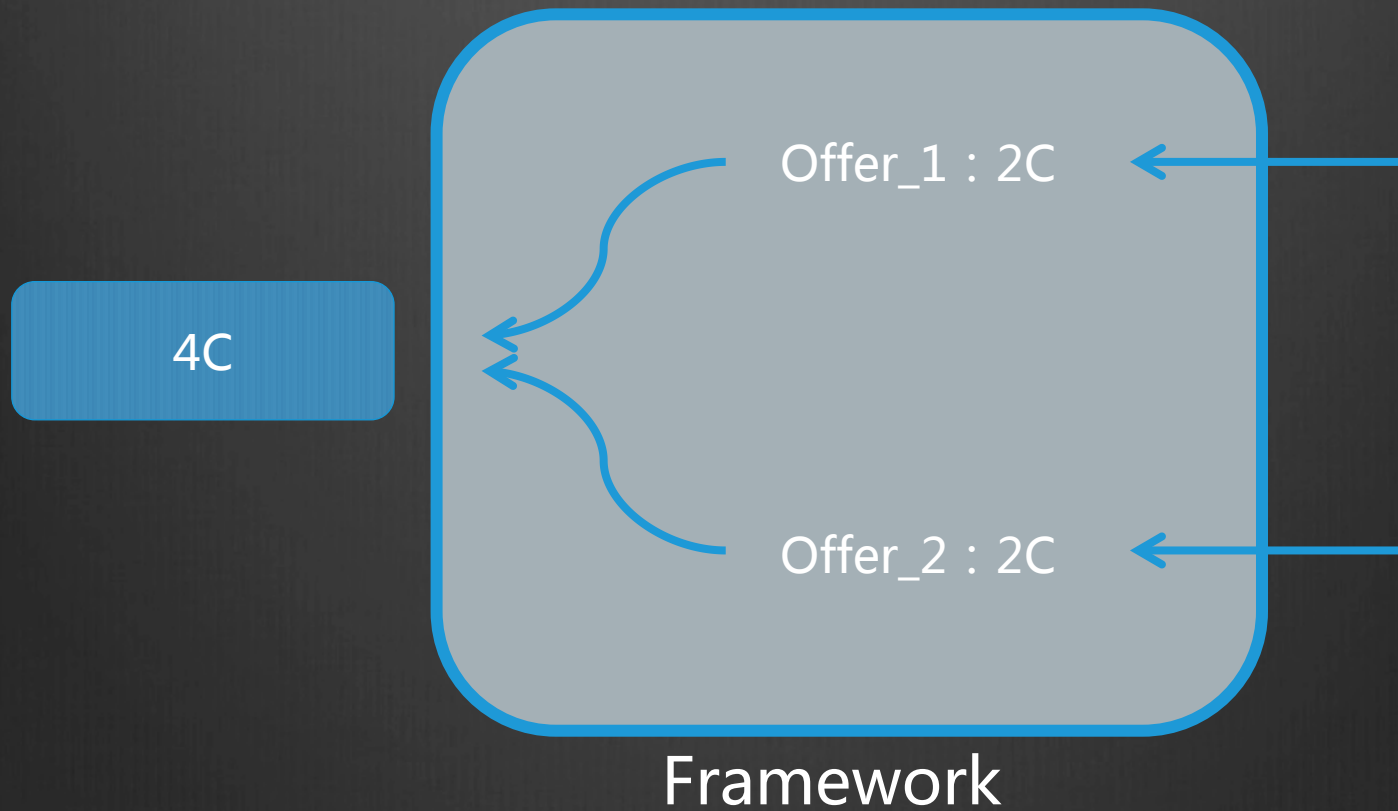
Stick

1. Stateful
2. 日志连续性
3. ARP

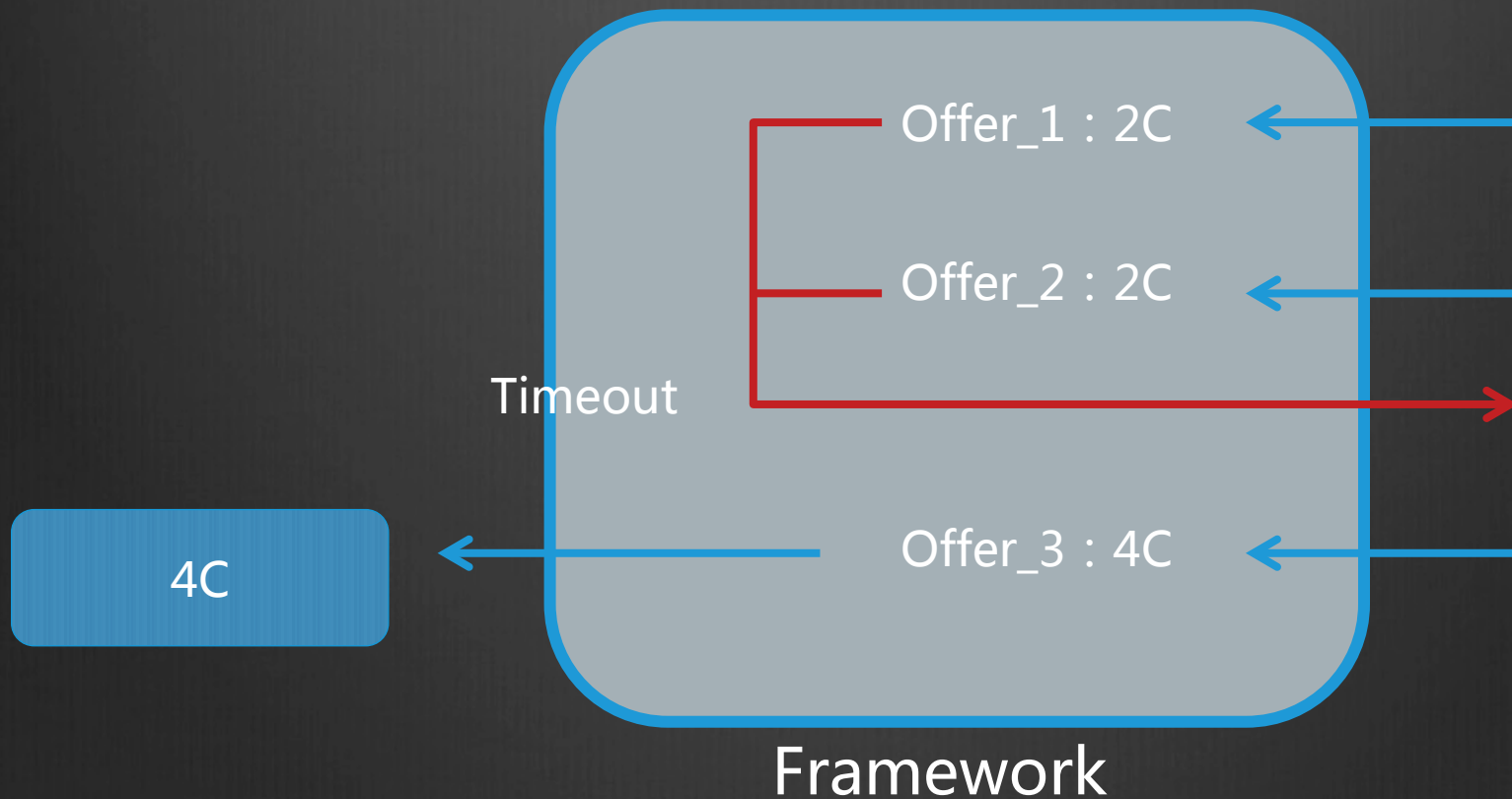
Offer 碎片



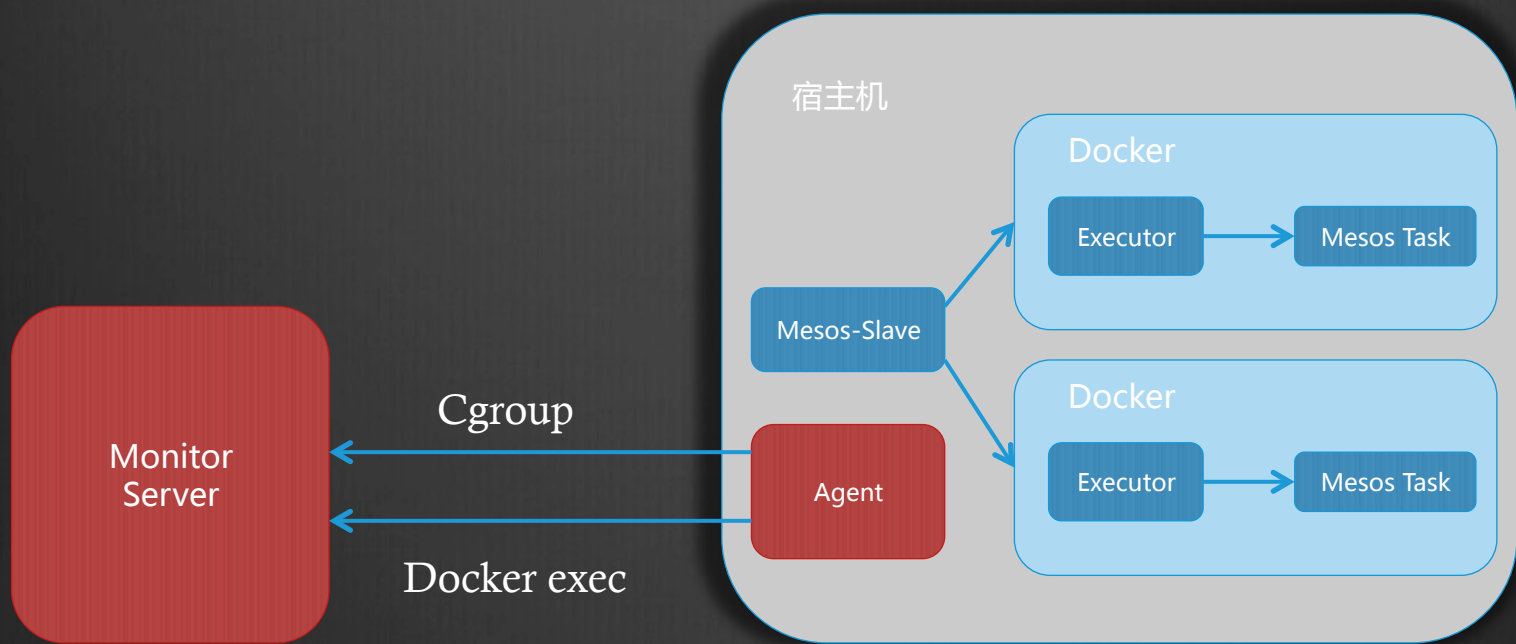
Solution 1



Solution 2



监控



自研的动机

- 轻量化，专注需求
- 兼容性，适配原有中间件
- 程序员的天性，改不如重写😊

混乱的水上部分

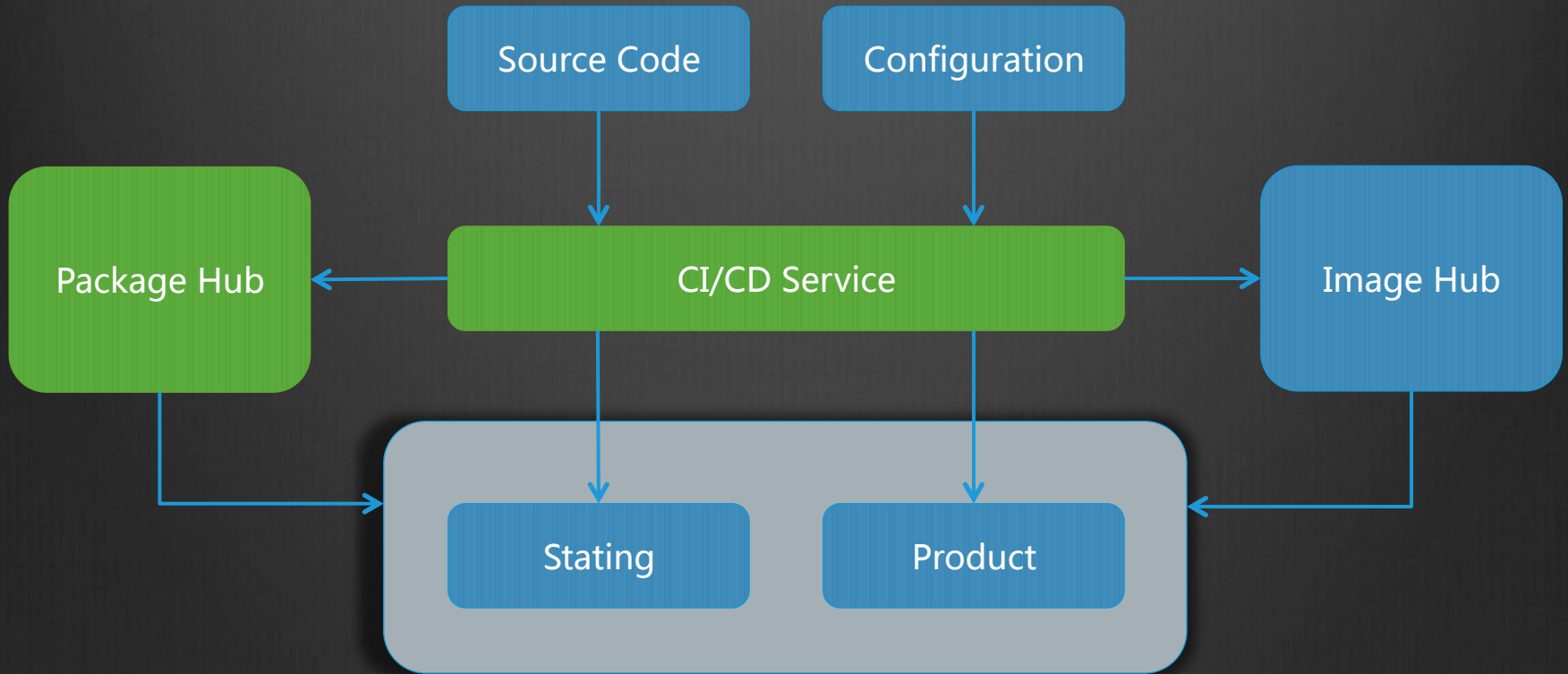


- Delivery
- Runtime
- Engineering

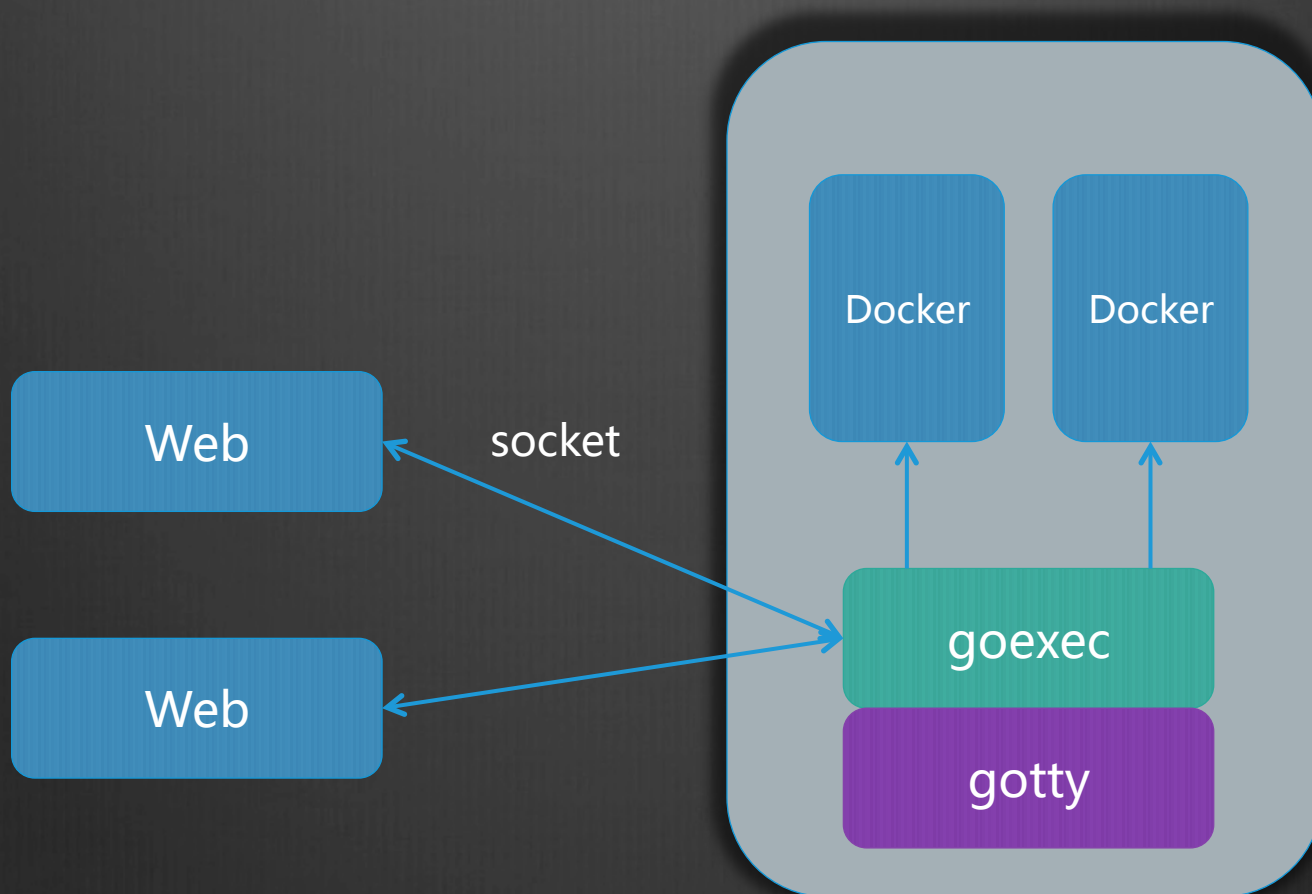
Delivery

谁开发，谁运行

交付过程



Web Console



Tomcat

需求：

- Tomcat 的一切行为可追溯

容器的通常行为：

- Tomcat 为主进程，启动失败，则实例销毁

矛盾：

- Tomcat 启动失败，是否等同于实例启动失败？

镜像内进程

UID	PID	PPID	C	STIME	TTY	TIME	CMD
root	1	0	0	06:43	?	00:00:00	/cexecutor/cexecutor -logtostderr -enable-dwait
root	22	1	0	06:44	?	00:00:07	/usr/bin/python /usr/bin/supervisord -c /etc/supervisord.conf
root	41	22	0	06:44	?	00:00:00	python /usr/local/sbin/listener_hook tomcat source /opt/tomcat/
root	42	22	0	06:44	?	00:00:00	python /usr/local/sbin/post_start_hook tomcat curl -s http://gi
root	43	22	0	06:44	?	00:00:00	python /usr/local/sbin/listener_hook tomcat source /opt/tomcat/
root	44	22	0	06:44	?	00:00:00	/usr/sbin/sshd -D
Deploy	45	22	0	06:44	?	00:00:32	/usr/bin/java -Djava.util.logging.config.file=/opt/tomcat/conf/
root	1225	0	0	14:31	?	00:00:00	hook

- Supervisord 负责注册自定义的 hook
- 并且负责维护 Tomcat 进程

IP

- 预分配 IP
- 发布时，不改变实例的 IP
- Job是个特例 (long running)

Runtime

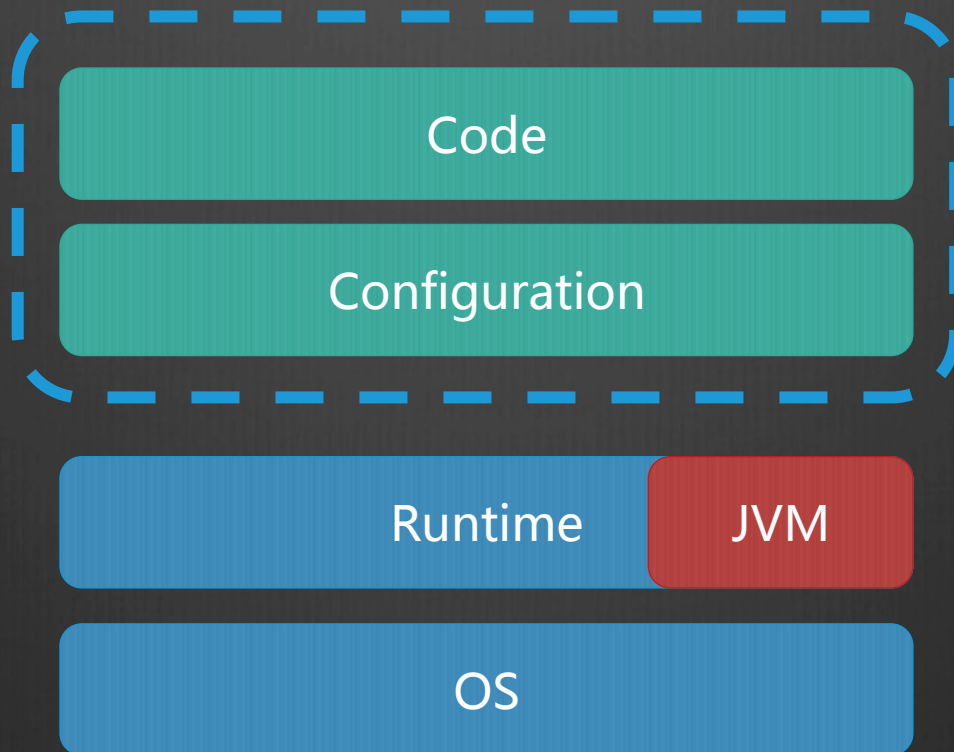
JVM 优化的锅

OOM

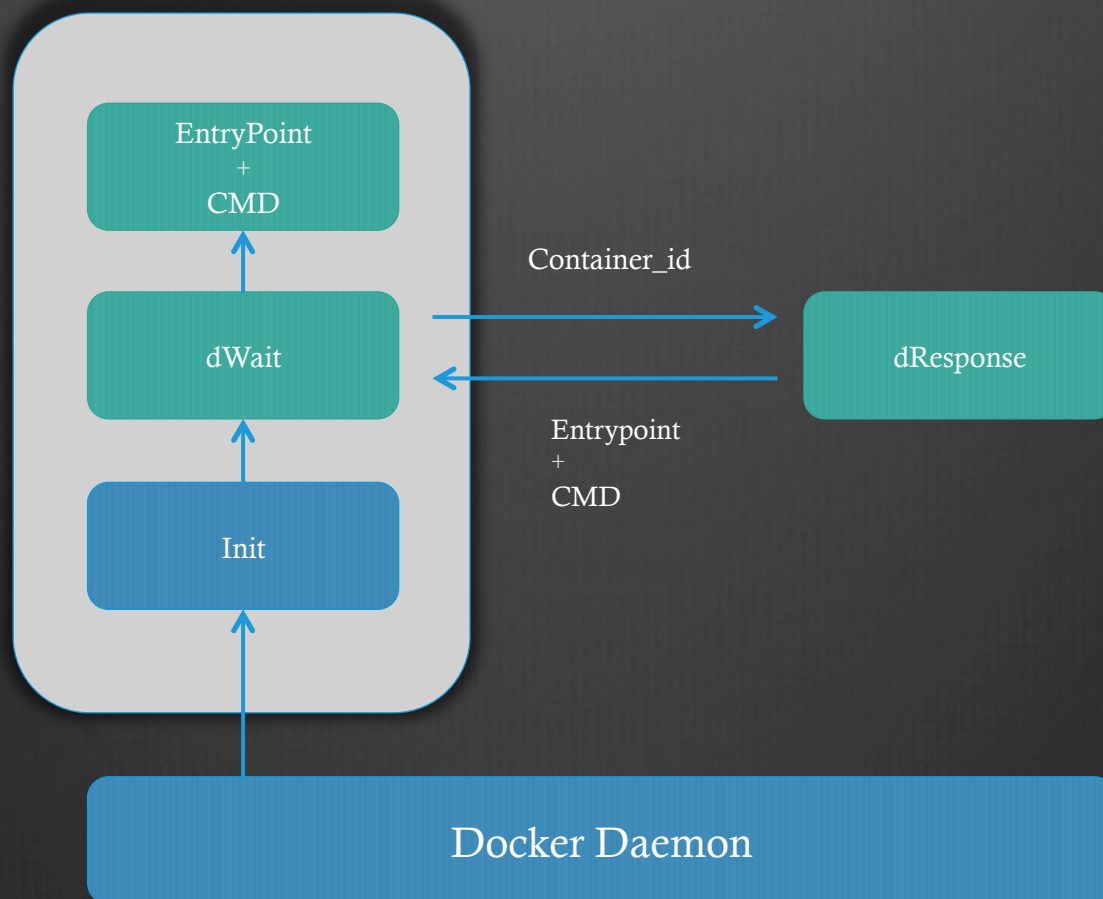
- 采用cpu quota , 无法获取准确的cpu数量
- Java 组件根据 cpu 数量确定 thread数量
- 造成堆外内存殆尽 , OOM

镜像分层

问题：谁来负责 JVM 的配置



实例启动过程



推荐配置

- Flavor : 4G
 - $X_{mx} = X_{ms} = \text{Flavor} * 80\%$
 - $X_{ss} = 256K$
- Flavor : 32G ?
- Flavor : 2G ?
- 堆外最小800 , 最大2G , 符合这个规则之内 , 以20%计

Customized

推荐配置



Override



强制配置

Engineering

Dockerfile 要受控

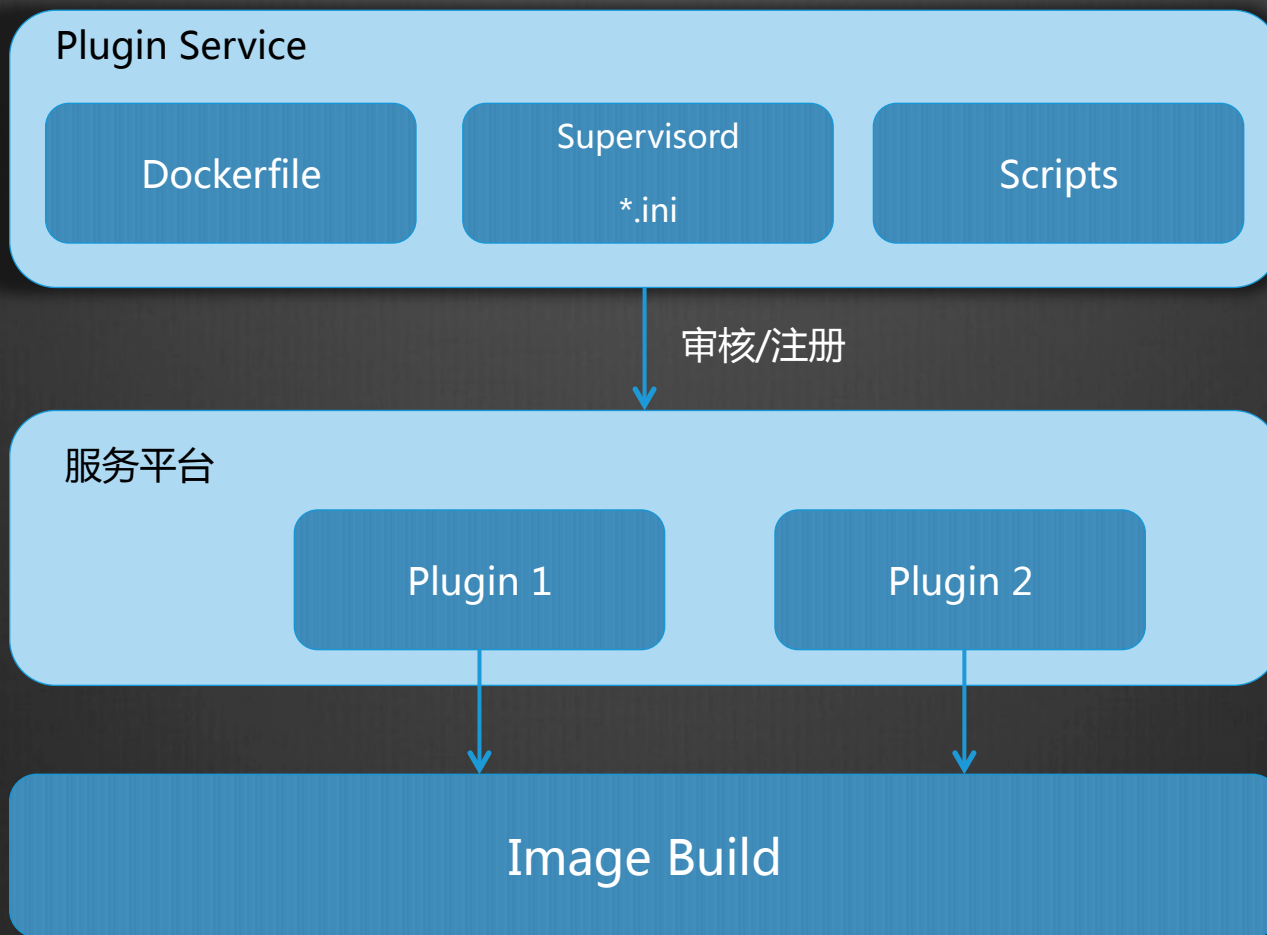
可爱的Dockerfile

- 不同的测试环境一定有不同的需求
- 通过 Dockerfile 定制 image
- 简单易懂，上手快

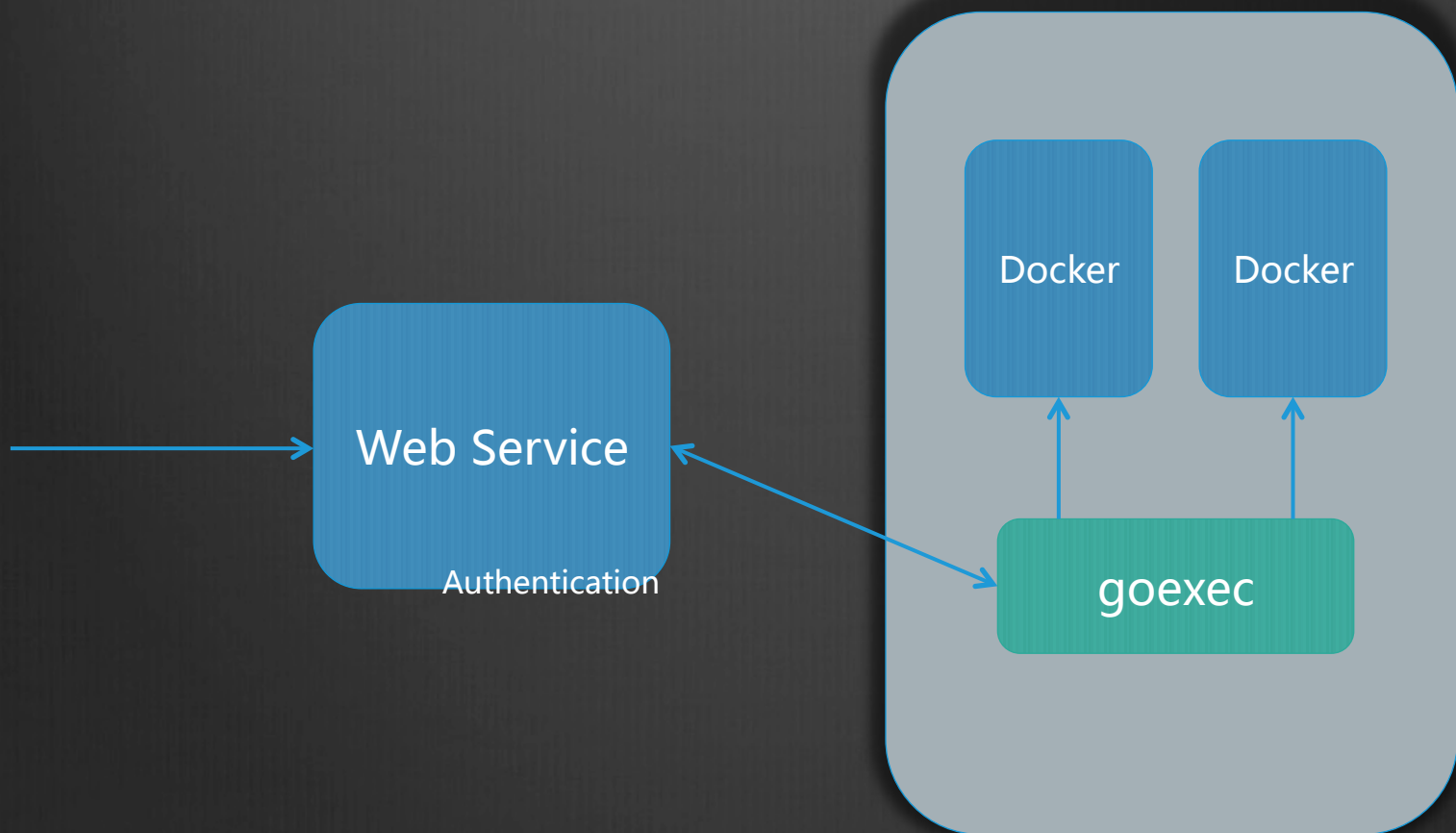
可恨的Dockerfile

- 无法执行条件运算
- 不支持继承
- 如何维护Dockerfile
- 就是个后门，环境标准化被完全破坏

Plugin 服务



Exec Scripts



Sample

测试环境，通过 Jacoco 手机代码覆盖率数据

Dockerfile :

1. 安装 Jacocoagent.jar
2. 完成环境配置

Post_start_hook.ini :

1. Tomcat启动后，通知 Jacoco Service

Scripts :

1. 重启 Tomcat

服务注册

创建服务

服务信息 *

名称 *

请填写服务名称

0 / 15

描述 *

请填写服务描述

0 / 50

图标

点击上传 png、jpg 或 jpeg 格式的图片，宽高 60*60，大小在 128KB 以内

环境配置 *

☐ FAT ☐ FWS ☐ LPT ☐ UAT ☐ PROD

下一步

创建服务

服务配置 - Plugin (可选)

Plugin 是 Docker 的插件，通过配置各阶段的自定义 Dockerfile 可以实现一些特定的功能。

自定义 Dockerfile

点击上传后缀名是 .sh 的文件

脚本名称 *

请填写脚本名称

脚本路径 *

请填写脚本路径

+

< 返回上一步

跳过Plugin配置，直接提交

首页 > 服务集市

创新工场

输入服务名称/描述搜索

创建服务



MySQL
关系型数据库



Redis
非关系型数据库



Mockserver
安装Mockserver



去哪儿Jacoco
安装Qunar Jacoco



酒店Jacoco
安装Hotel Jacoco

服务开通

soa-service-test

基本信息 安全信息 WEBHOOKS CHECKLIST SERVICE

Service 列表

开通 Service

名称	描述	操作
 Mockserver	安装Mockserver	
 去哪儿Jacoco	安装Qunar Jacoco	
 酒店Jacoco	安装Hotel Jacoco	开通

Page: 1 Rows per page: 20 1 - 3 of 3 < >

最后

- DevOps 观念的改变
- 水上和水下同样重要
- 适合自己的才是最好的

Thanks