



# 饿了么容器化的落地

黄跃龙

饿了么 核心基础设施部

资深工程师



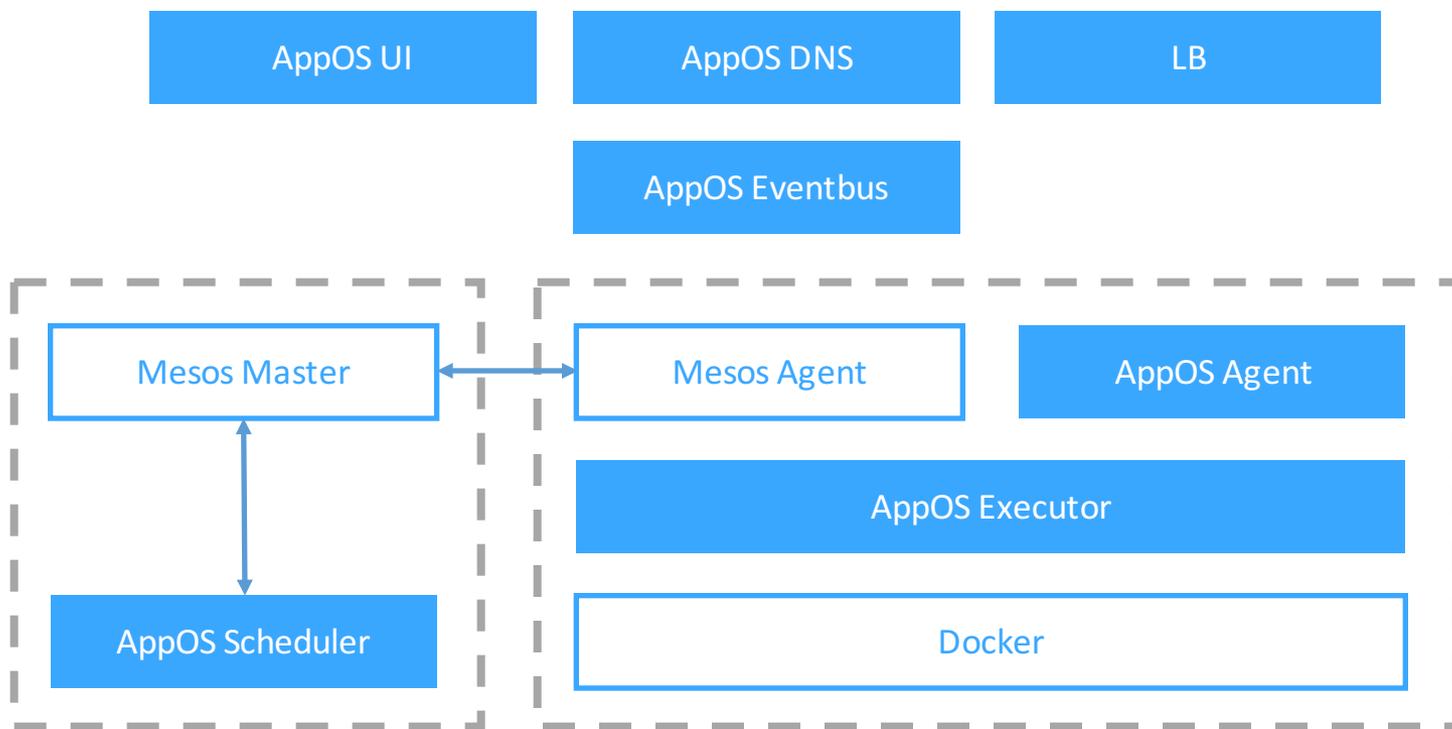
- 调度与编排
- 存储与网络
- 服务发现
- 镜像
- 宿主机管理

## 怎样才算落地？

---

- ✓ 能接入业务关键路径上的组件
- ✓ 现有应用能简单地接入和迁移
- ✓ 应用配套设施接入后不打折扣

# AppOS 简要架构



# 调度与编排

从单机跨出的第一步

User Interface

Scheduler

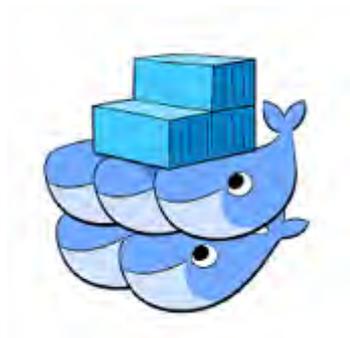
Container Runtime

OS

封装与业务无关的调度细节

封装集群管理的细节

封装进程资源隔离的细节



官方支持  
轻量级



成熟的大规模调度  
兼容其他类型任务



功能丰富  
社区非常活跃

没有特殊需求选kubernetes就好

**选哪个都没关系**

# 存储与网络

服务可用的基本条件

容器化真正的技术难点



# 存储

---

■ Storage Driver    Devicemapper

■ Volume Driver    Local

## 如何考虑

- 基础设施限制
- 性能
- 有没有排障的能力

# 持久化

---



理想

可靠的高性能分布式存储

现实

NFS, Mesos Persistent Volume

妥协

只接入无状态应用

# 网络

---

■ 公有云 Flannel

■ 私有云 MacVlan

## 如何考虑

- 基础设施限制
- 性能
- 有没有排障的能力

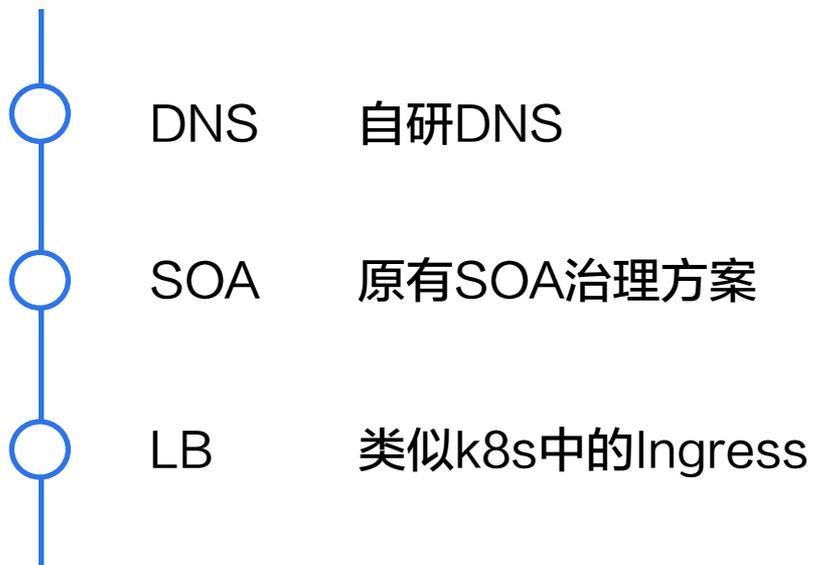
# 服务发现

向外提供服务



# 服务发现

---



# DNS

---



与调度器打通，动态更新频繁

可以用来存储元数据

负载均衡的一种选项

# 反向代理

---



与调度器打通，动态更新

向外部提供服务时保证网络隔离

负载均衡

# 镜像

Docker 生态的灵魂



# 镜像

---



Registry

自研Registry

镜像分发

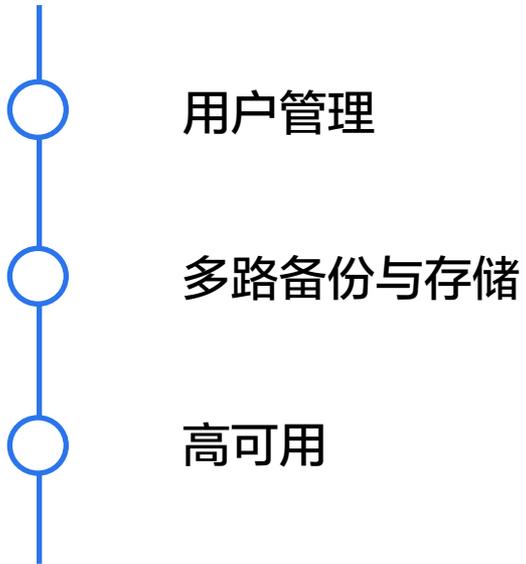
预热以加速部署

基础镜像

方便应用接入

# Registry

---



# 基础镜像

---



1GB!

开箱即用 我们也很绝望啊

定制的启动脚本

标准化的容器应用

能当虚拟机用

方便调试与接入

# 宿主机管理



# 宿主机管理

---



快速的标准化部署

稳定性与可扩展性的基础

弹性扩缩容

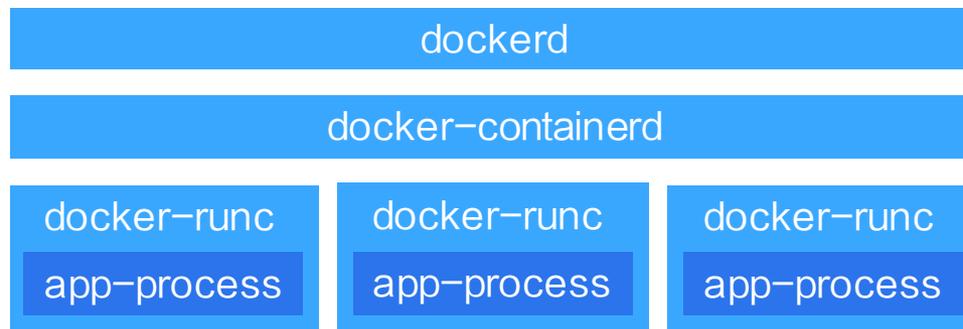
逻辑扩缩容与物理扩缩容

Docker Daemon

# Docker Daemon

---

- 
- 为应用多增加了一层依赖
  - 不是100%可靠
  - 挂掉之后容器怎么办



- 开启 `live-restore`
- `daemon`挂掉后，通过`runc`或直接操作进程来控制容器
- 减少对`daemon`功能的依赖

# 开发的需求

---



远程登录	通过exec实现
日志	log driver, logging plugin, mesos sandbox
监控	自己实现指标采集

## 其他

---

- 无论选择何种容器调度方案，仍需大量自主研发
- 60%的问题不是技术问题，而是治理问题
- 关注容器相关的标准如OCI, CNI, CSI
- 面向用户的系统一定要有界面



Q&A

