

# DockerOnYarn调度系统实践

Sogou-申贤强



## 关于我们

- I. 来自搜狗大数据平台部
- II. 基于Apache Hadoop生态，建设搜狗海量数据存储和计算平台
- III. 提供稳定高效的数据分析系统，为搜狗各类型大数据应用，提供一站式数据处理服务
- IV. 每天数十亿的数据增量，数以万计的数据计算流程，使数据的价值得到充分利用
- V. 最前沿技术落地及推进开源技术的发展

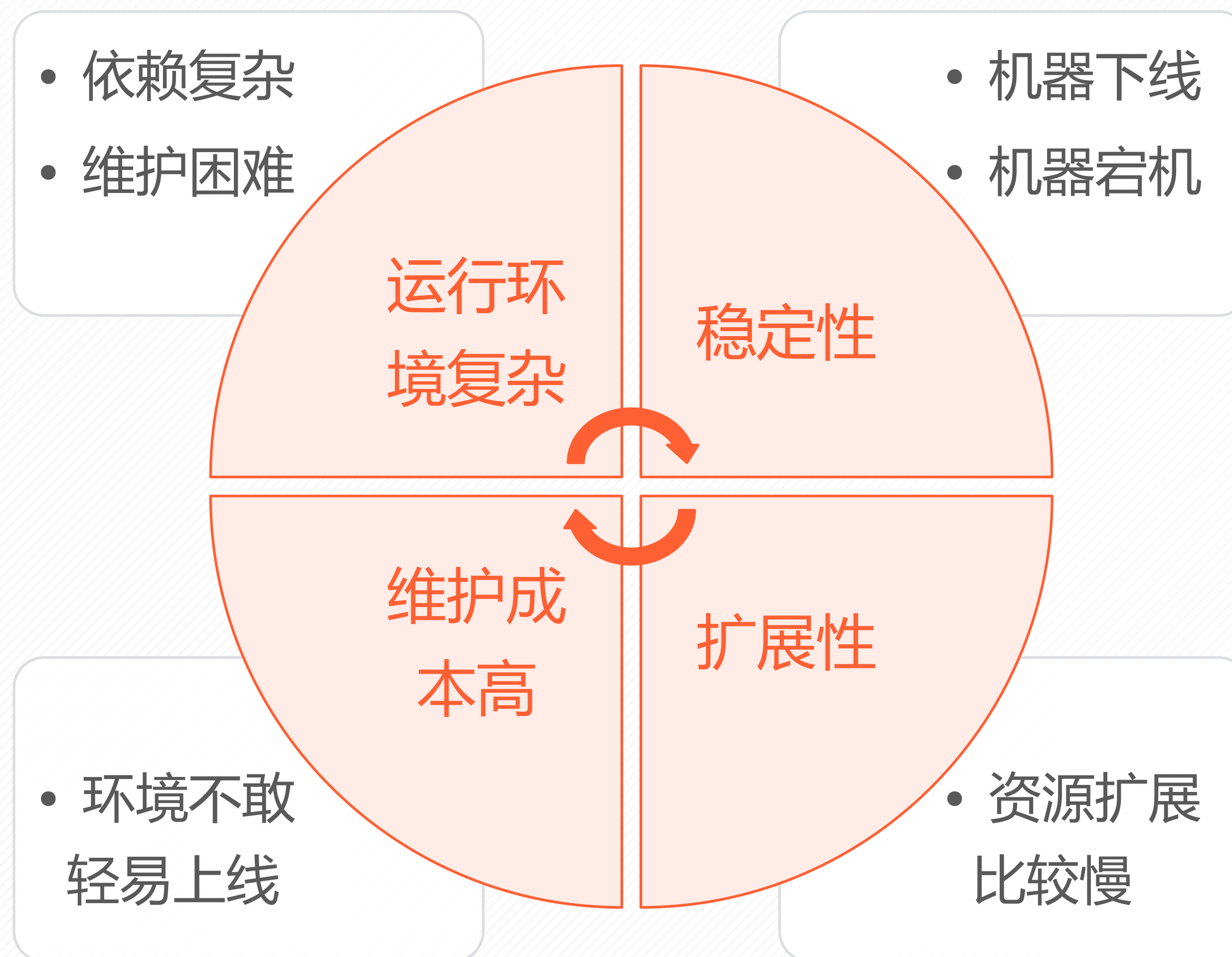
# Agenda

## I. 背景

## II. DockerOnYarn调度实践

- a) Docker在sogou的应用
- b) DockerOnYarn框架
- c) Docker基础服务优化
- d) 任务提交与容灾
- e) CIPS-Sogou实践





# DockerOnYarn调度实践

## Docker在Sogou的应用



### 线上业务

- 企业搜索核心模块，包括搜索相关核心模块，以及前端系统

### 线下业务

- 调研流程，各类线下实验平台

### Clotho

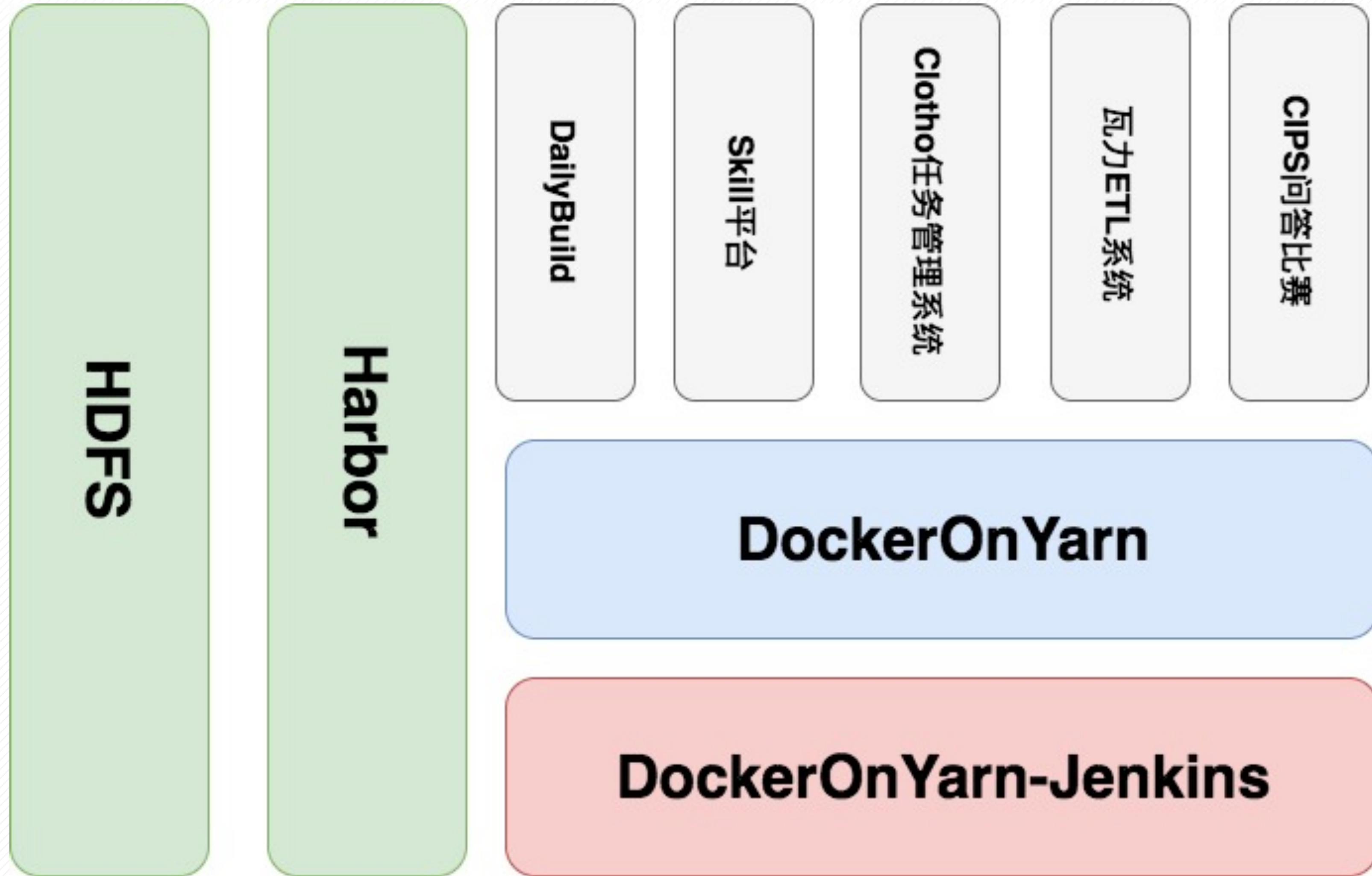
- Hadoop离线任务管理，主要涉及相关性rank Link分析，意图识别，反作弊，数据统计等
- Long Running Service，如ELK，grafana，query Sensu监控，报警系统，ganglia等各类服务

同时运行的Docker Container达到万级以上，每天完成几十万离线Job，托管了上千类型的Long Running Service



# DockerOnYarn调度实践

## Docker在Sogou的应用



# DockerOnYarn调度实践

## DockerOnYarn框架

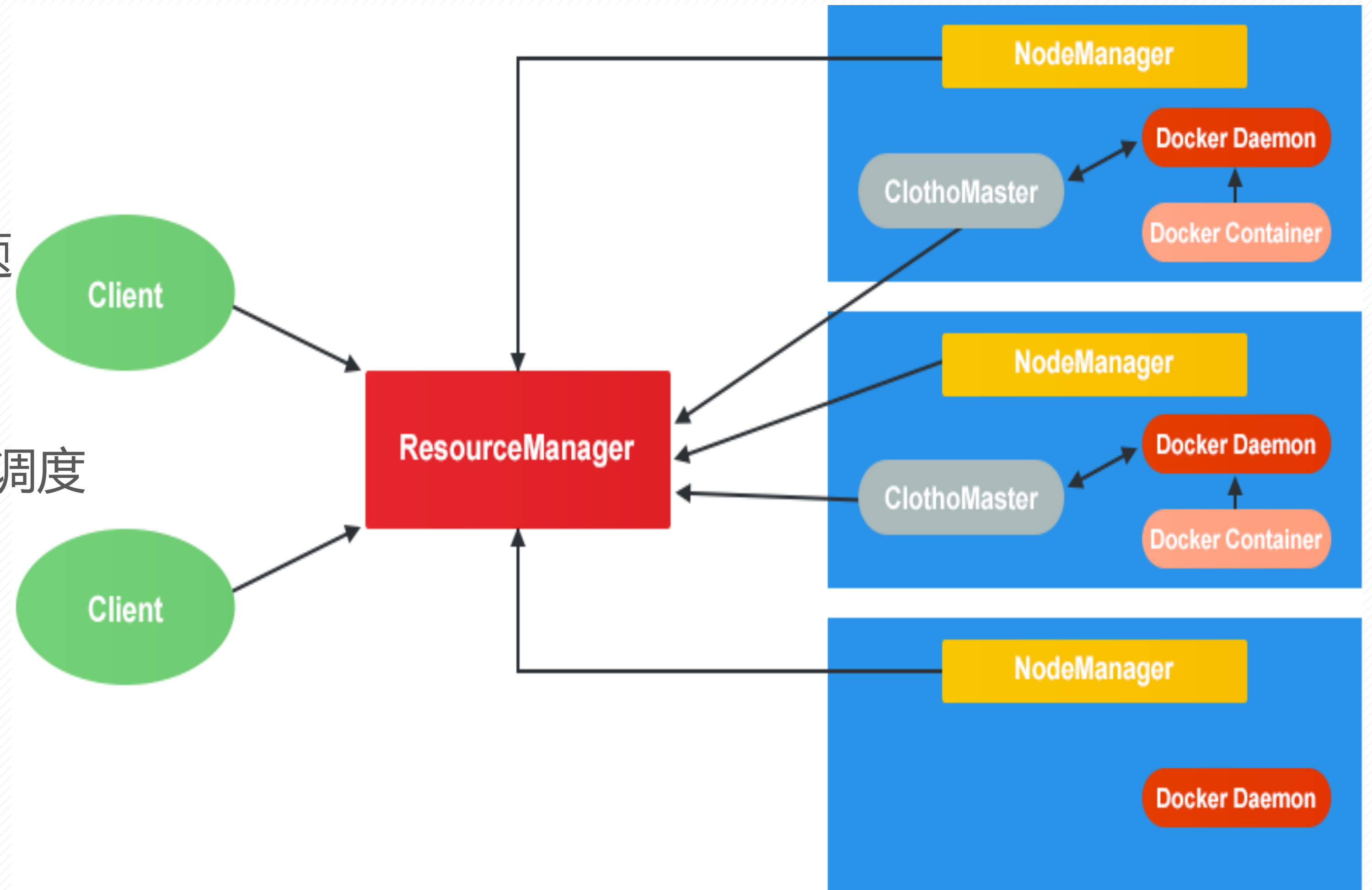
### 1. Docker 1.13.1

解决随机端口冲突

解决docker daemon僵死问题

commit exceed max depth

### 2. 指定机器/Label/Rack/Group调度

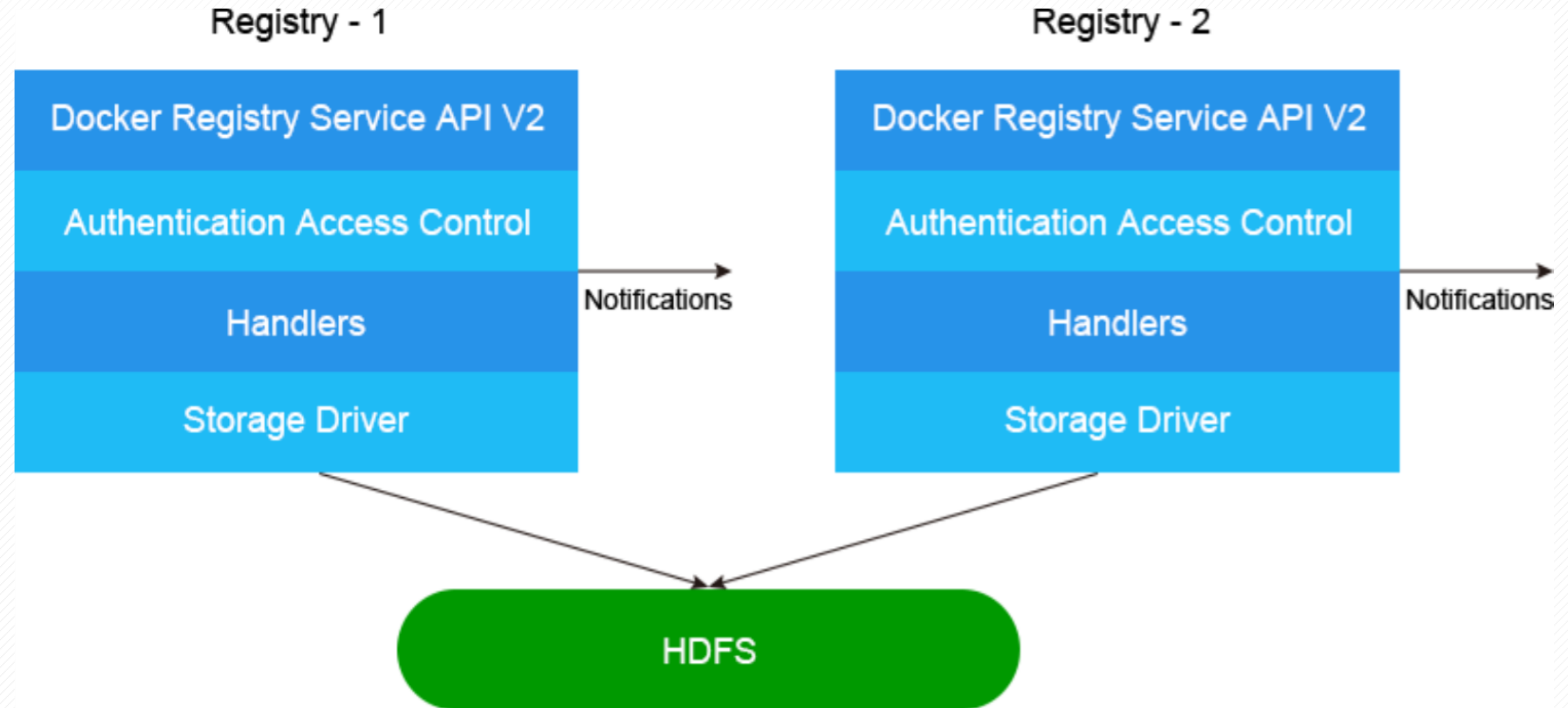




# DockerOnYarn调度实践

## Docker基础服务优化

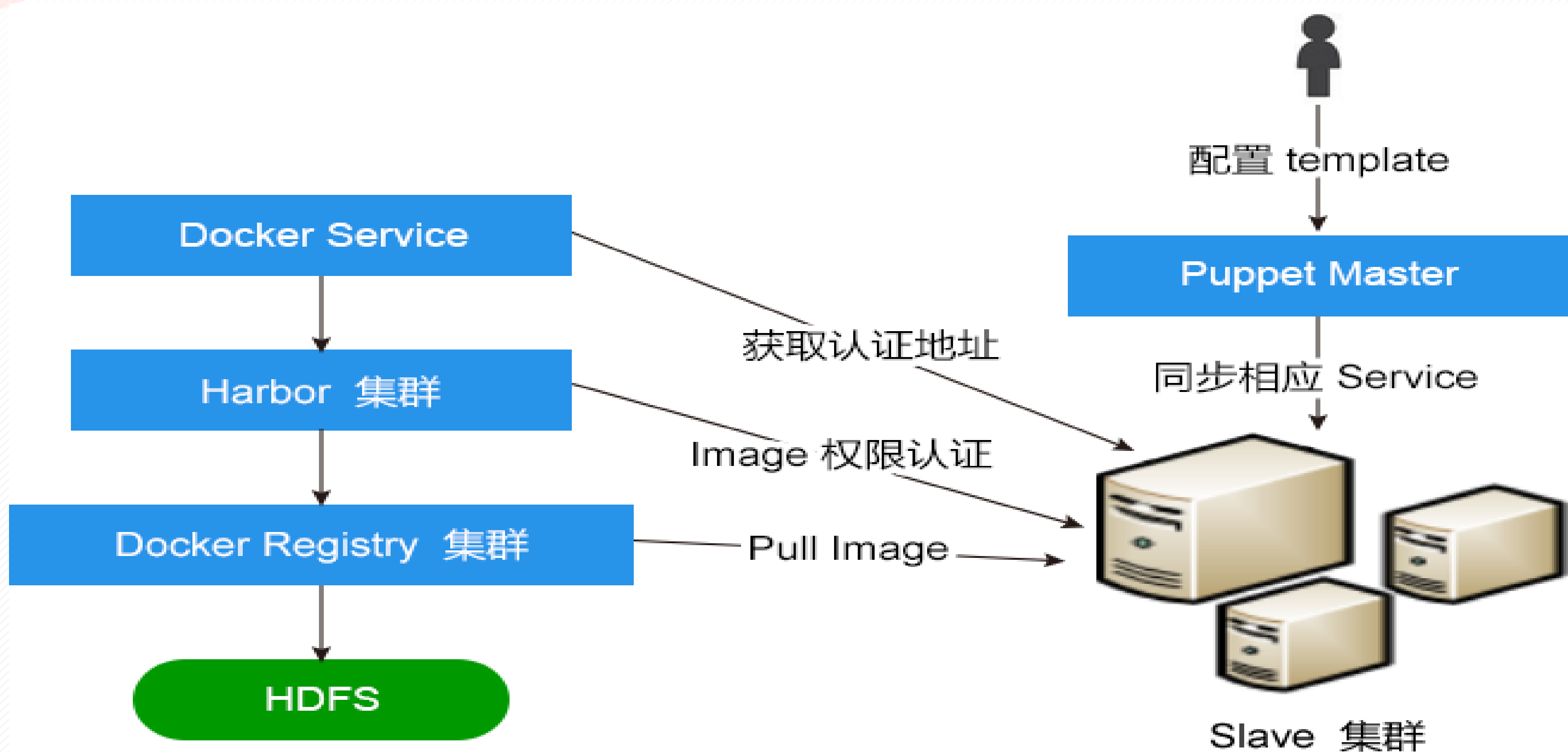
- 稳定性高：99.99%
- 并发度高
- 扩展性好





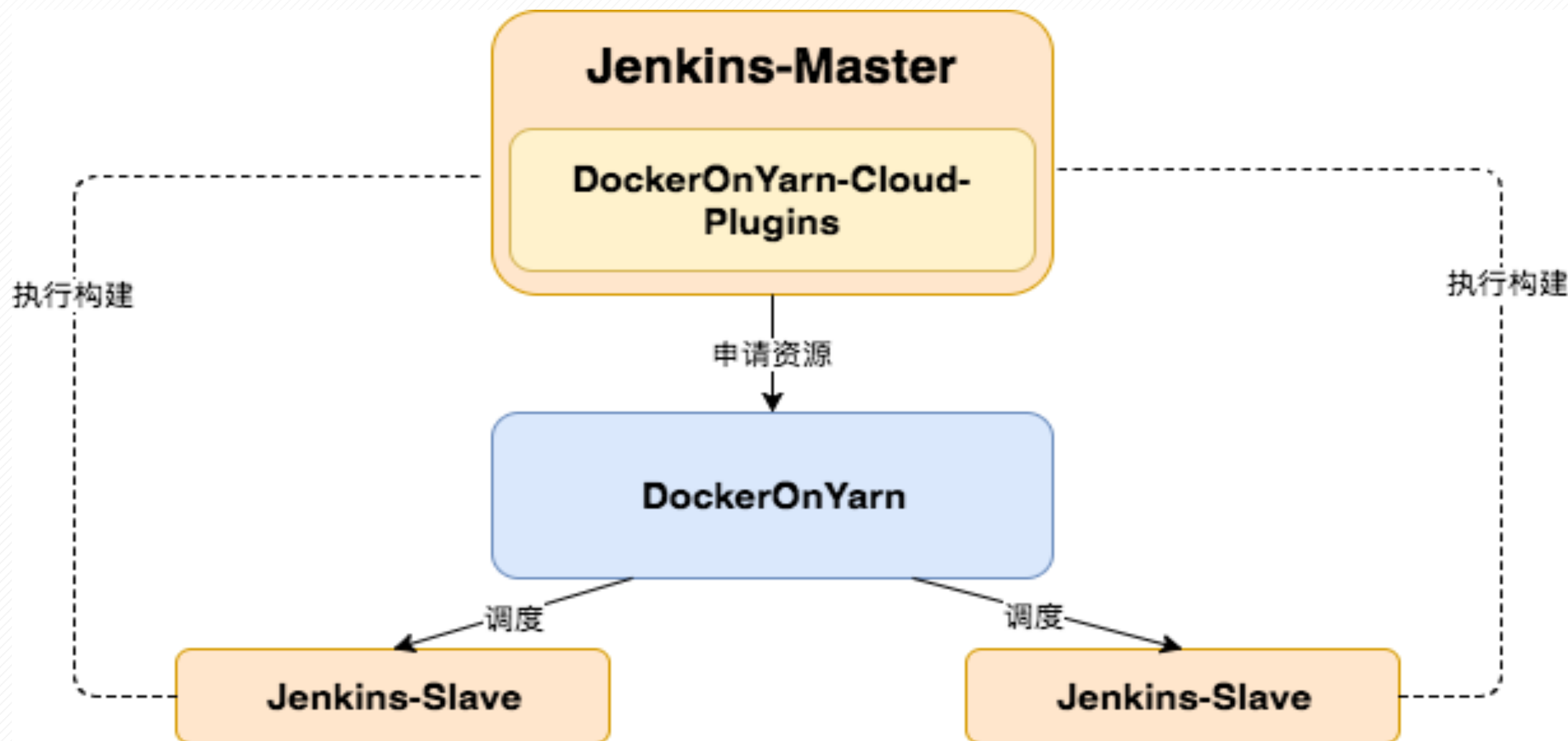
# DockerOnYarn调度实践

## Docker基础服务优化



# DockerOnYarn调度实践

## Docker基础服务优化





# DockerOnYarn调度实践

## Docker基础服务优化

☒ Restrict where this project can be run

Label Expression

[Label zhanghongda@sogou-inc.com:123](#) is serviced by no nodes and 1 cloud. Permissions or other restrictions provided by plugins may prevent this job from running on those nodes.

Docker Image Build Step

Dockerfile

Config Build Steps

Base Image

Build Steps

Add File in Image

File To Add

File Path In Image

Execute Command

Command

Install Yum Packages

Packages

Packages List Splits by ','

Set Environment Variable

Key

Value

Set Work Directory

Work Directory

保存

应用

# DockerOnYarn调度实践

## 任务提交与容灾

### 服务SPEC配置：

- AppName：唯一标识
- Image：实际运行的Image信息
- Tag：标识服务状态
- Port：发布的端口
- Instances：启动的Service数目

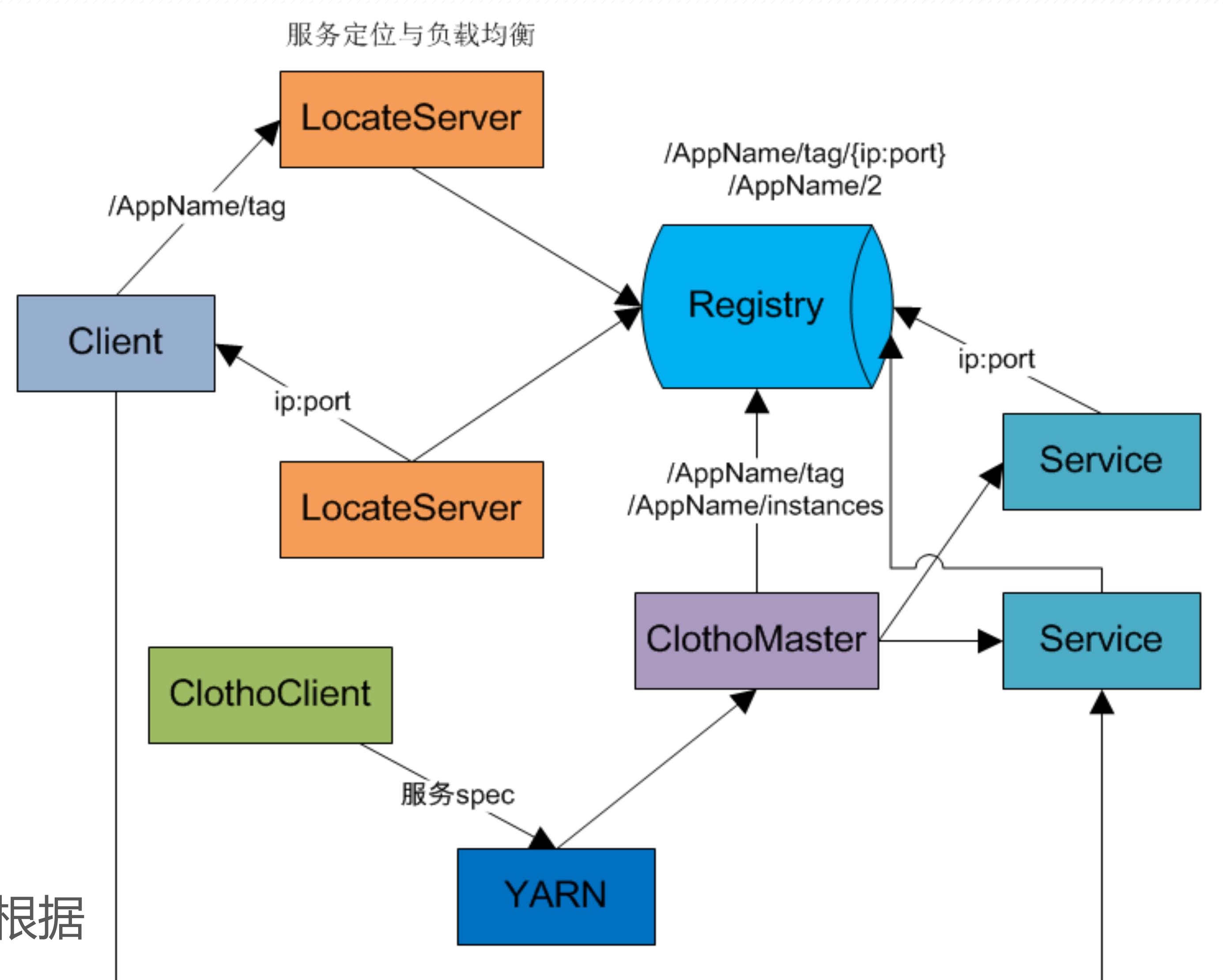
### 负载均衡：

多实例情况下简单随机，含地域特性

### 弹性伸缩：

直接修改Registry Service里

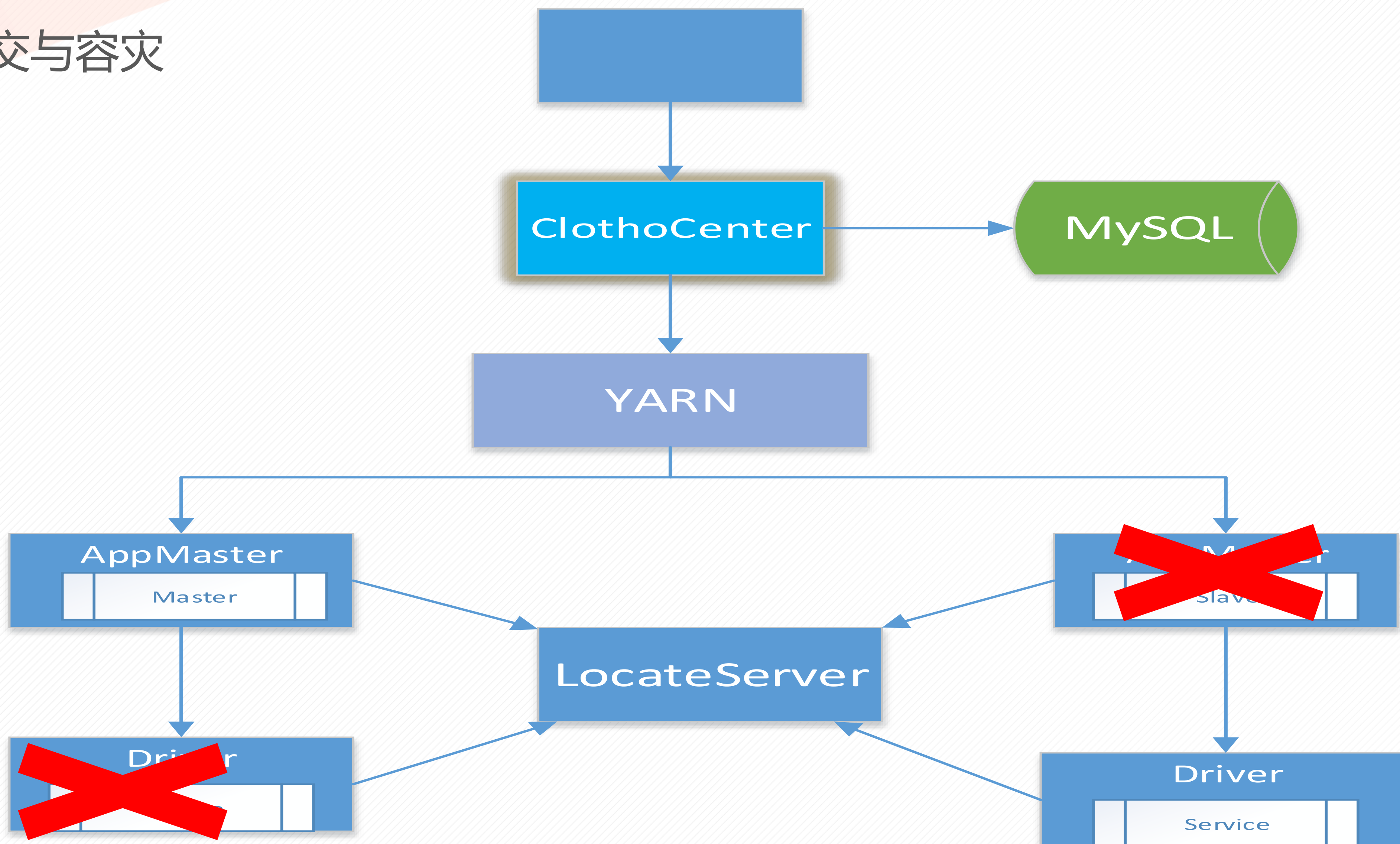
/AppName/instances，ClothoMaster根据instances动态调整申请的资源





# DockerOnYarn调度实践

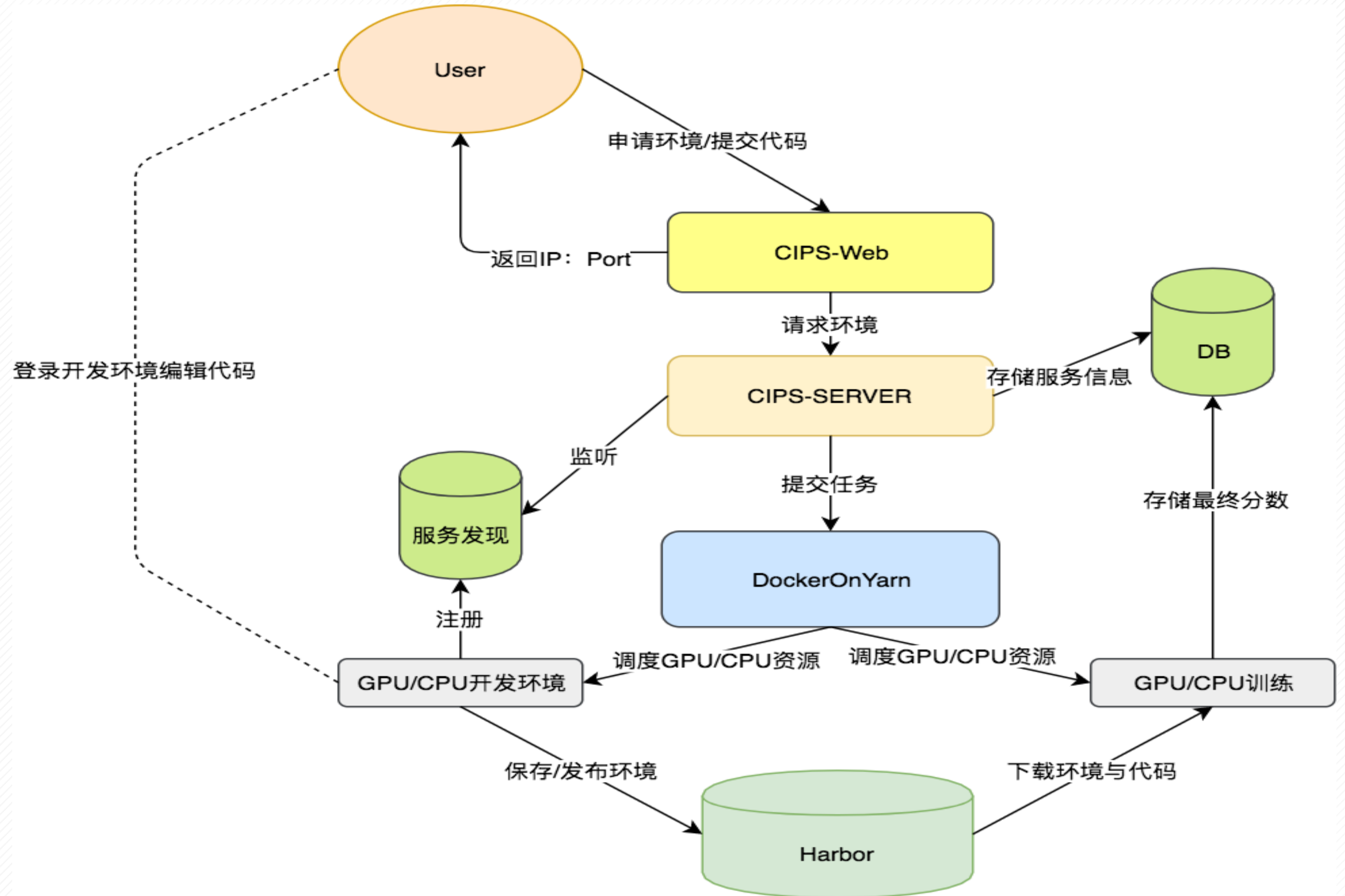
## 任务提交与容灾



# DockerOnYarn调度实践

## CIPS-Sogou实践

1. Docker隔离GPU资源
2. 实现GPU的板卡调度
3. 指定GPU/CPU组的随机调度与指定机器调度
4. 解决exceed max depth
5. 使用HDFS解决共享数据





# TODO

- I. 支持复杂DAG
- II. 支持更好的资源隔离策略
- III. 与Hadoop集群结合，支持离线和在线业务

Thank You