360搜索技术论坛长期交流群

360搜索

基于Mesos的360搜索私有云平台

谭博侃

01 360搜索面临的挑战

02 Mesos介绍

03 私有云平台架构

04 经验分享和未来展望

一.360搜索面临的挑战

海量数据

收录几千亿网页
索引几百亿网页

大规模集群

数万台机器, 多个IDC

时效性高

每天上万次服务数据变更
每周PB级别数据更新

稳定性高

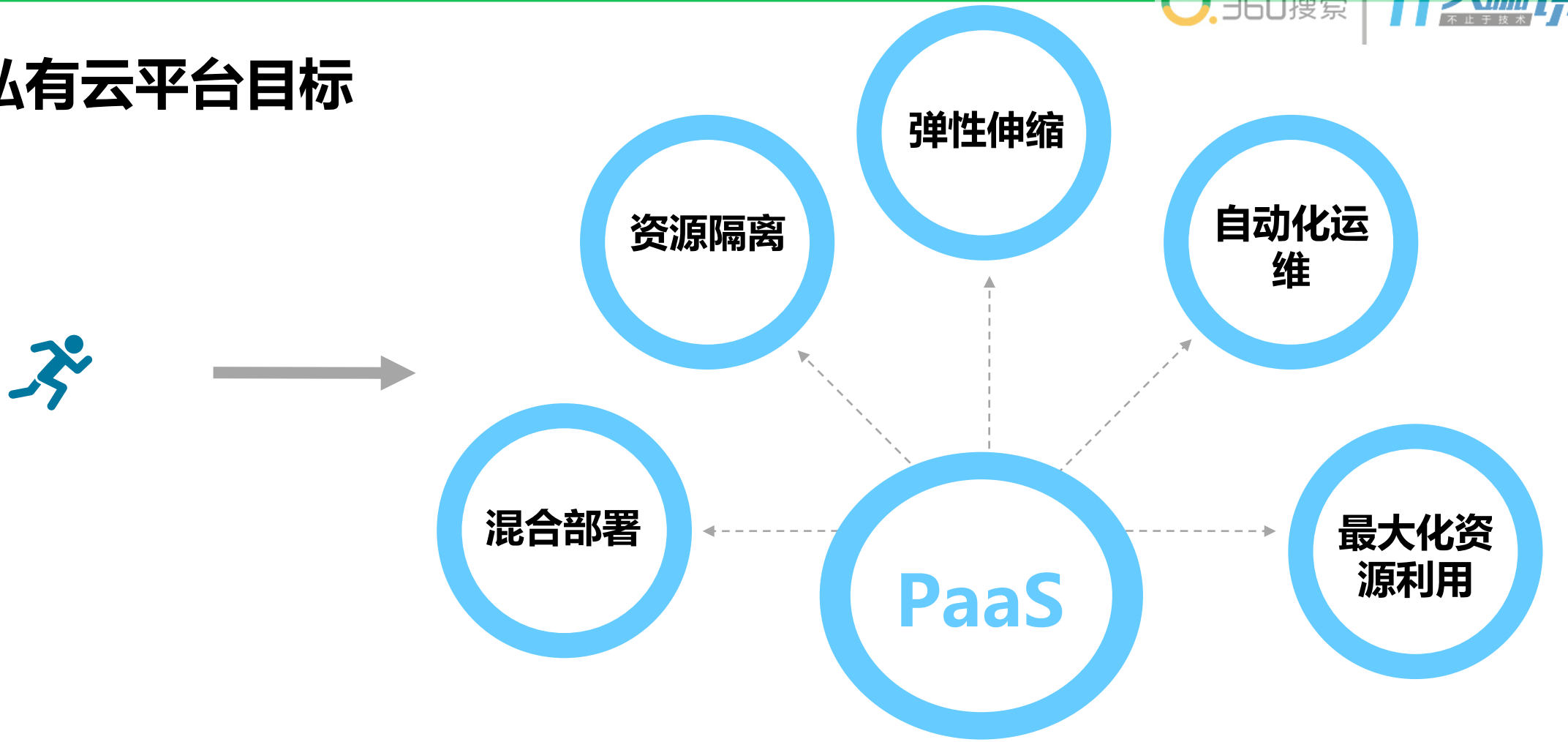
每秒数千次检索请求
可靠性要求 99.99%

传统运维存在的问题

- 服务器静态划分
- 资源利用不均衡
- 应用耦合多
- 无法弹性伸缩
- 管理成本高

一.360搜索面临的挑战

私有云平台目标



最终目的： 优化搜索引擎的延迟和吞吐，从而应用更复杂的相关性算法来提高搜索质量

一.360搜索面临的挑战

集群管理调度系统:

Google Borg

Apache Mesos

Google Kubernetes

01 360搜索面临的挑战

02 **Mesos介绍**

03 私有云平台架构

04 经验分享和未来展望

二.Mesos介绍- What’ s Mesos

Mesos: A kernel for data center applications

Program against your datacenter like it’s a single pool of resources

整个数据中心当成一个资源池来使用

二.Mesos介绍- Why Mesos

- **生产环境验证**
  - Twitter/Apple/Netflix
- **扩展性好**
  - twitter 3w+
  - 接口兼容性好
- **可定制性强**
  - Framework/Executor
  - Allocator/Container/Network/Storage

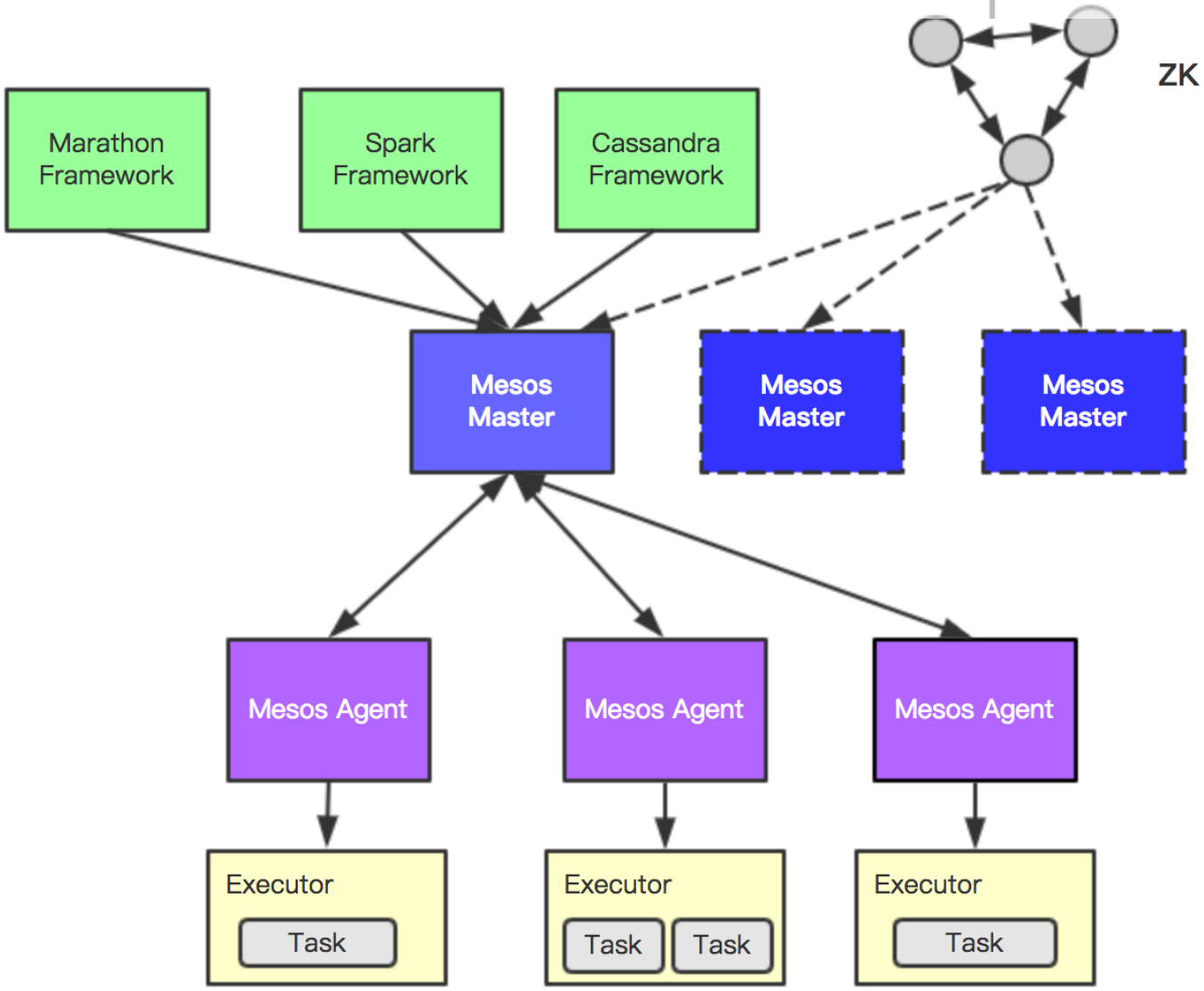
二.Mesos介绍-架构

Zookeeper:
实现Master高可用

Master:
接受Agent上报的资源
分配Offer给上层Framework

Framework:
基于Mesos API实现的调度器
根据调度需求决定是否接收Offer

Agent:
接受并执行Master的命令
通过Executor启动Task

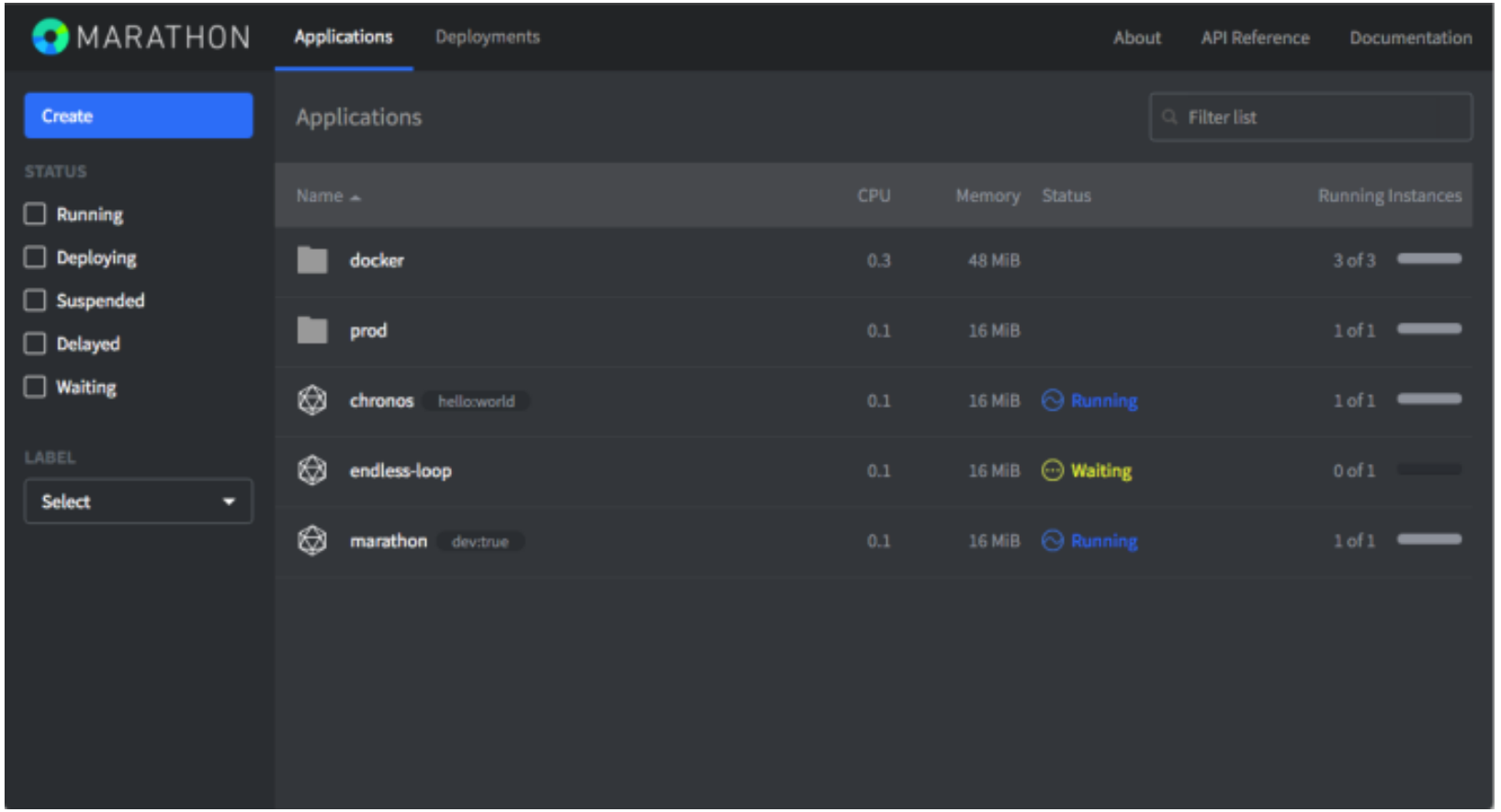


二.Mesos介绍-Framework

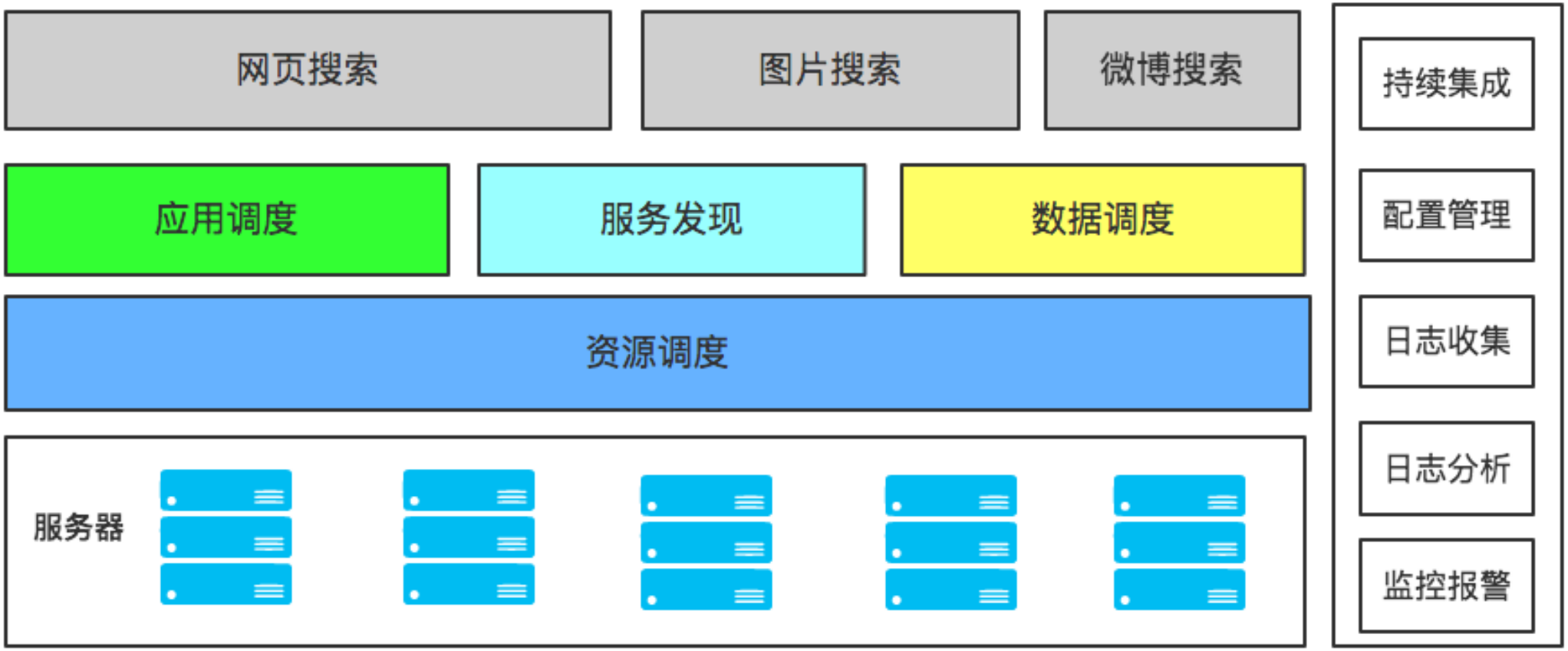
二.Mesos介绍-Marathon

Marathon特性

- 长服务
- 约束性条件
- 服务发现
- 健康监测
- Event Bus
- WebUI/RestAPI



三.私有云平台架构



三.私有云平台架构-技术选型

	已选方案	其他选择
自动化部署	Ansible	Puppet/SaltSlack
服务调度	Mesos + Marathon	K8S/Swarm
服务发现	Marathon-lb(haproxy)	Nginx/LVS/Zookeeper
配置管理	Consul + Consul-template	Zookeeper/Qconf
数据依赖	自研发数据调度DS	NFS/Ceph/S3/HDFS
日志收集/分析	Filebeat + ELK / QLog	Scribe / Flume
容器	Mesos / Docker	OCI/APPC
网络	Host	CNI
监控报警	InfluxDB/Sensu	Prometheus

01 360搜索面临的挑战

02 Mesos介绍

03 私有云平台架构

04 **经验分享和未来展望**

私有云平台三点经验分享

服务发现

静态主机和端口变成动态
上游需要能感知下游服务

数据调度

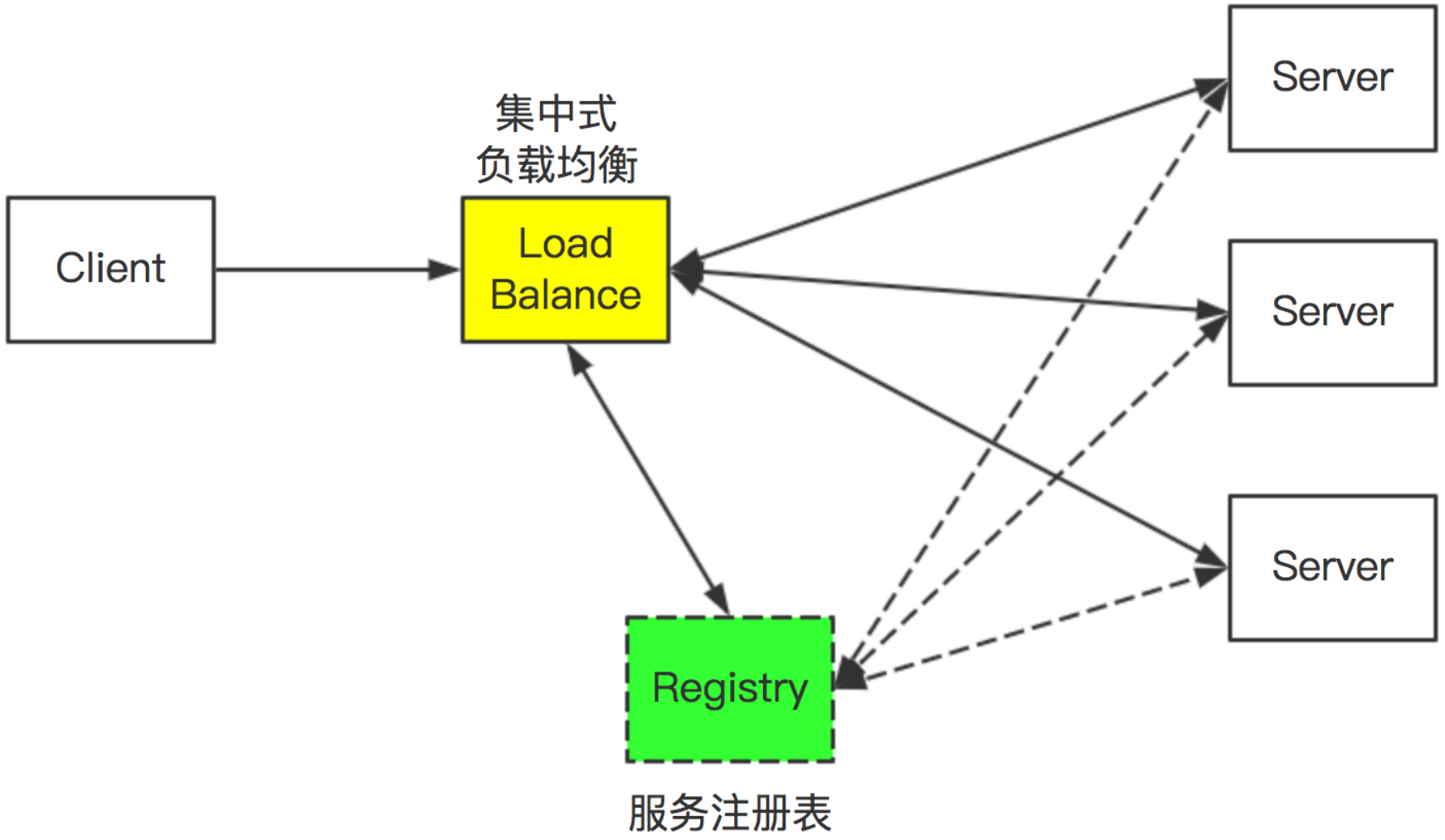
有状态业务迁移云平台难
自研发的DS数据调度系统
解决有状态服务的迁移问题

高可靠保证

开发者担心迁移后稳定性
如何保证服务可靠性

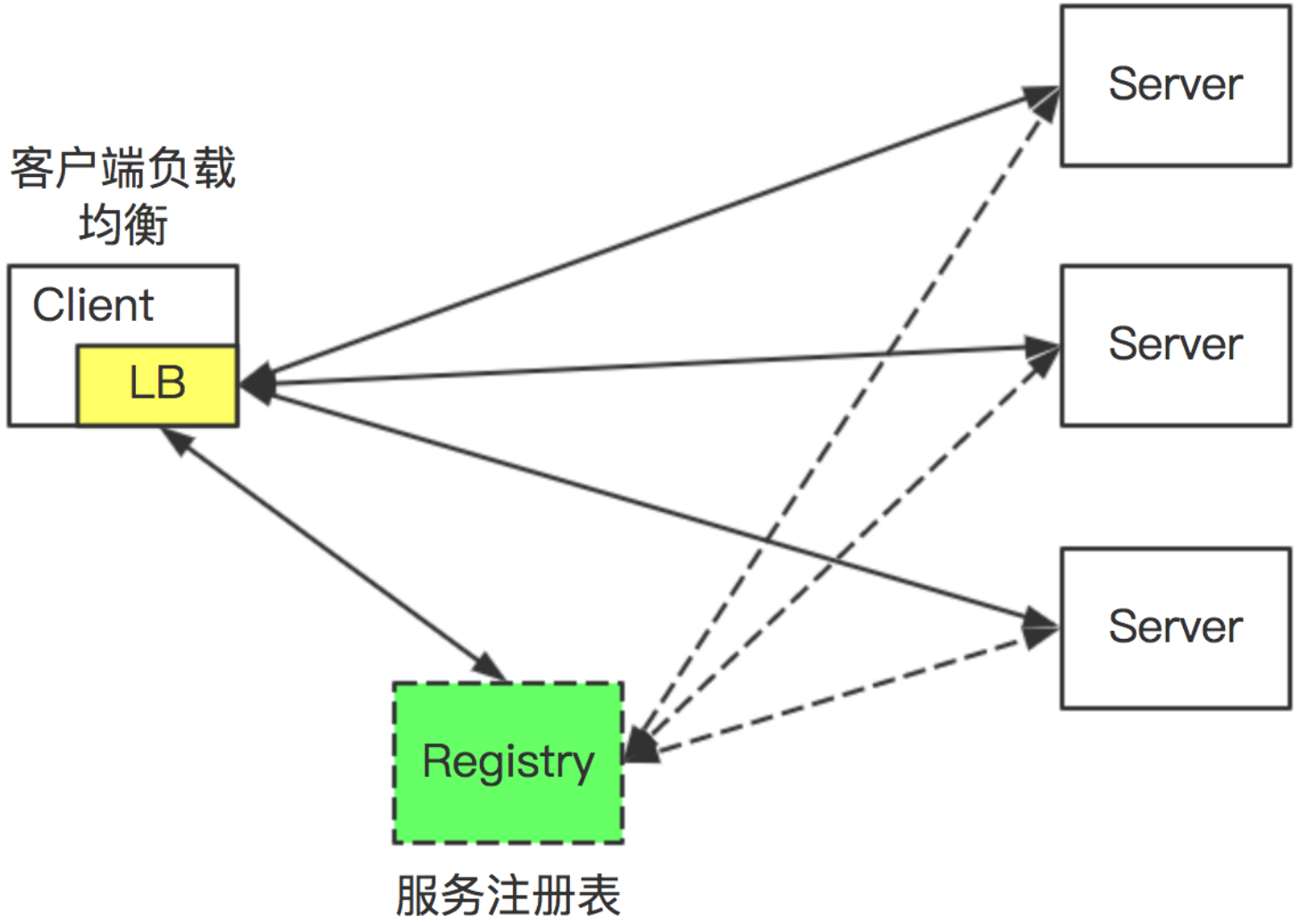
集中负载均衡

- LVS
- HAProxy
- Nginx



四.经验分享-服务发现

智能客户端



Mesos周边解决方案

- Mesos-DNS
- Marathon-lb(集中负载均衡)
- Zookeeper + SDK(智能客户端)
- Traefik
- Linkerd

四.经验分享-数据调度

无状态的服务　　**有状态的服务**

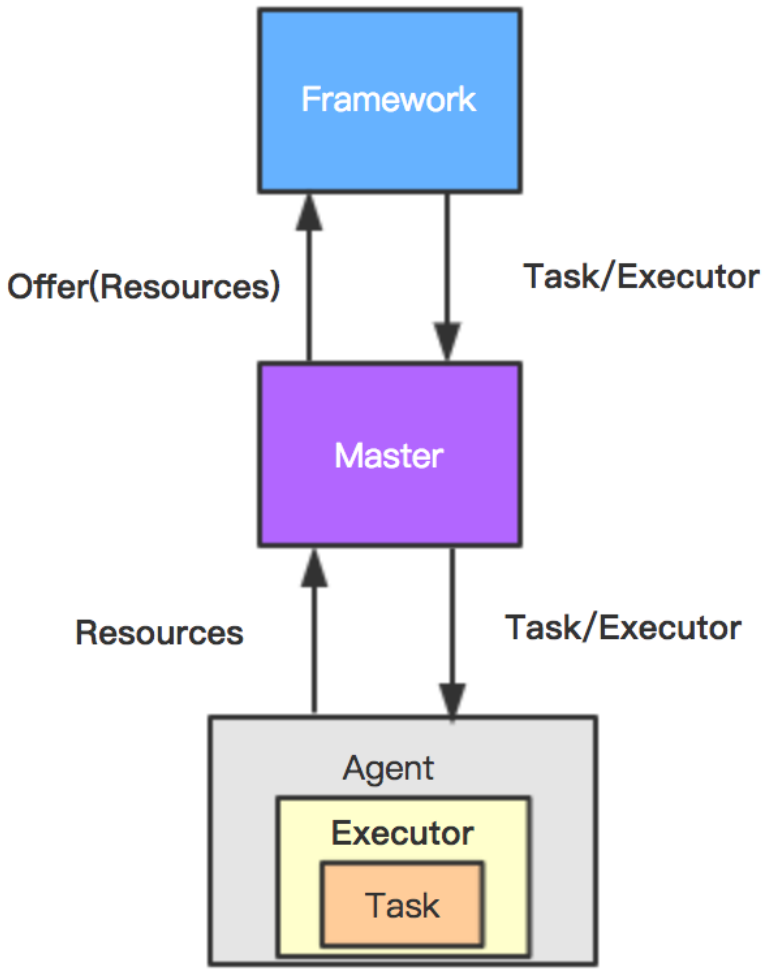
四.经验分享-数据调度

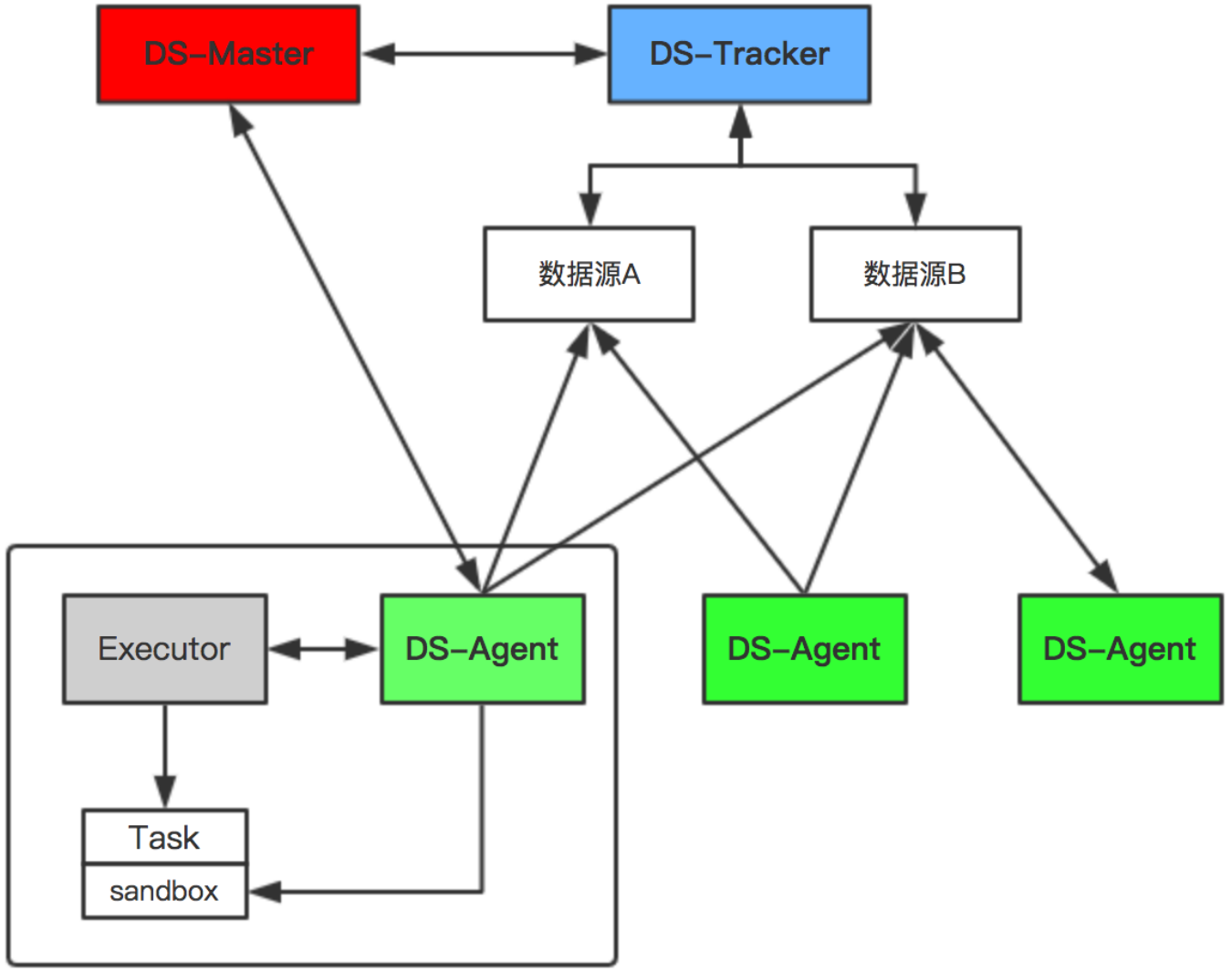
有状态服务:

- 约束性条件固定节点 (维护成本高)
- Mesos动态预留和持久化卷 （ 数据无法迁移 ）
- 分布式文件存储(网络中断，性能差)
- Framework和应用可以进行数据备份和数据迁移(类似Cassandra, 改造成本高)

搜索服务的状态特点

四.经验分享-数据调度 DS架构





四.经验分享-高可靠保证

架构高可用:

Mesos Master 无状态,通过ZK保证的高可用
Mesos 的Agent支持Recovery
Marathon通过ZK选主和持久化

服务高可用:

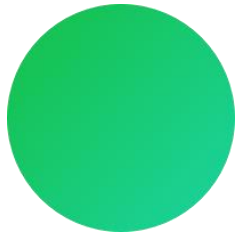
Marathon支持健康检查和滚动升级策略
Marathon + Mesos + UCS框架支持优雅退出机制
Marathon-lb支持灰度发布

索引服务调度到Mesos私有云平台

定制Framework和DS系统结合，支持数据亲和性调度

DS系统支持HDFS，P2P，跨机房的数据调度

更精准的资源调度，保障业务稳定的前提下最大化利用集群资源



QA