



爱奇艺基于 Docker 的 App Engine 实践

杨成伟 <yangchengwei@qiyi.com>

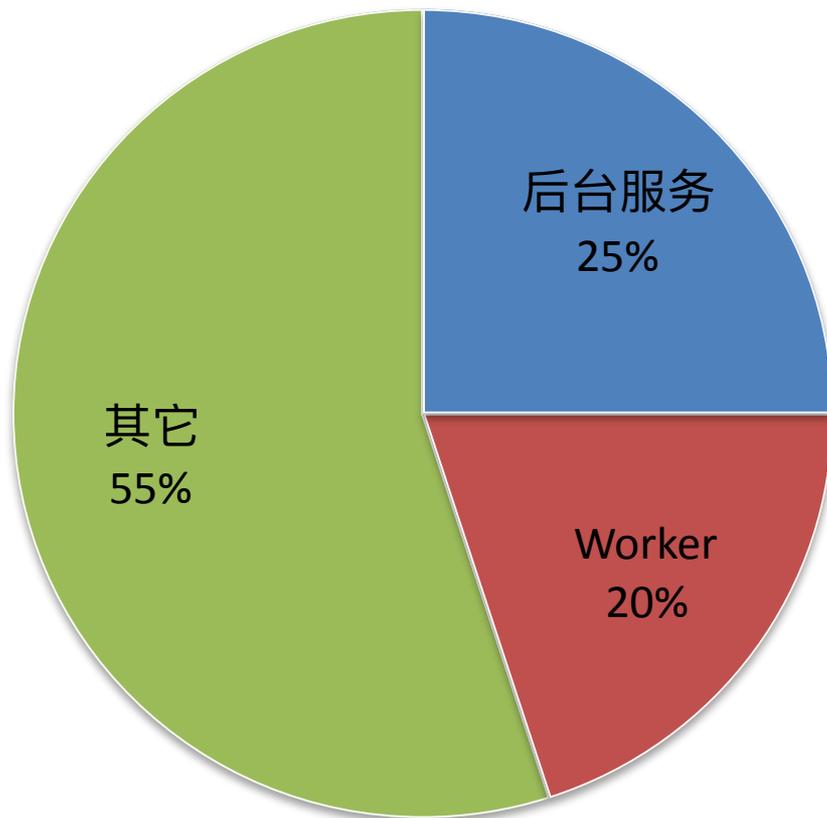
@container2016

内容提要

- 背景、出发点、目标
- iQIYI App Engine 设计与实现
- 一些经验（坑）
- 现状与展望

背景 — 业务上

虚拟机承载的业务



背景 — 技术上

业务接入	iQIYI App Engine	其它		
任务调度层	Long-Running (Marathon 0.5.0)	Time-Based Batch (Chronos)	Batch (Sisyphus)	Big Data (Storm)
任务运行时	Docker 1.0			Mesos builtin
计算资源层	Mesos 0.16.0 CentOS 6.4	Mesos 0.16.0 CentOS 6.4	Mesos 0.16.0 CentOS 6.4	

2014 年 Q3

出发点



目标



目标 一会变



目标 — 会变



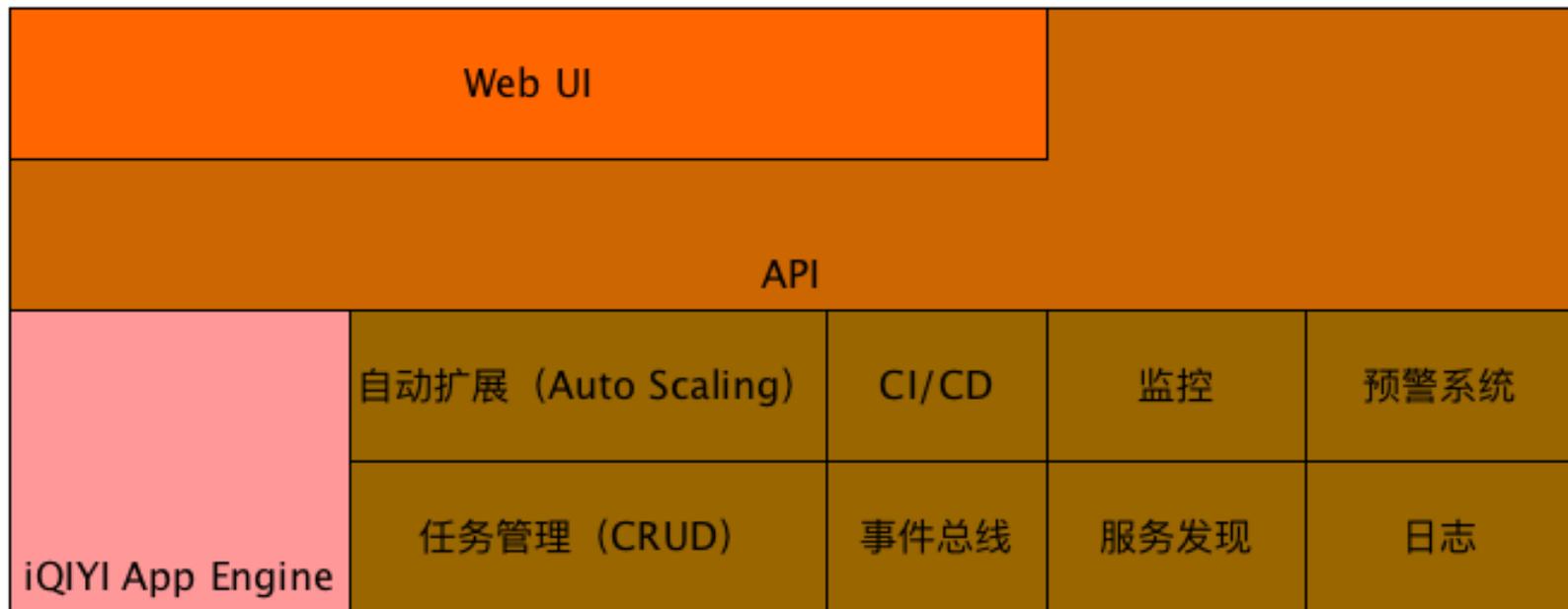
用户受益

- 资源到位快
- 部署快（上线、升级；持续集成、发布）
- 扩容快
- 自动故障转移，免运维
- 完善的监控预警

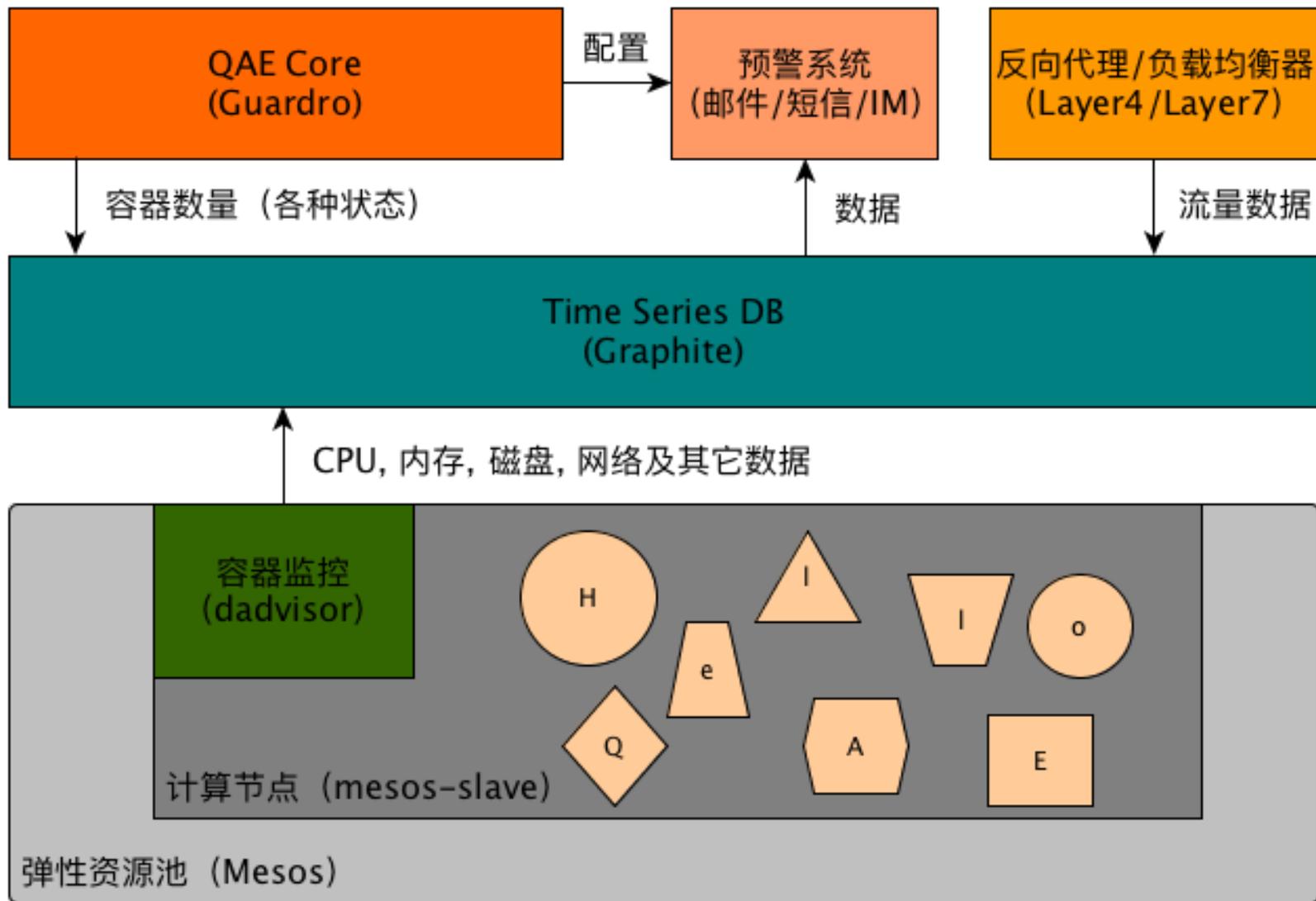
内容提要

- 背景、出发点、目标
- iQIYI App Engine 设计与实现
- 一些经验（坑）
- 现状与展望

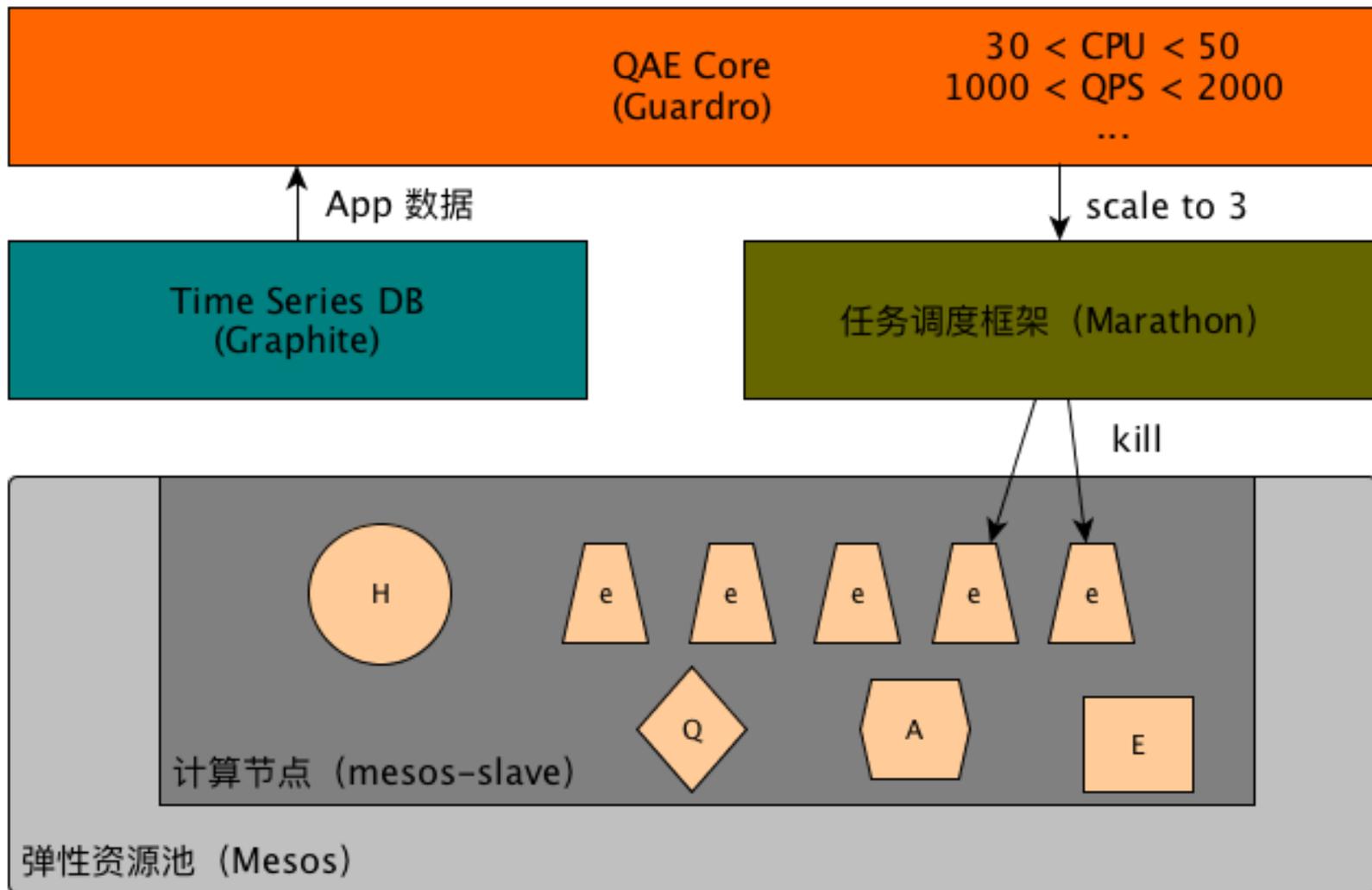
iQIYI App Engine (QAE)



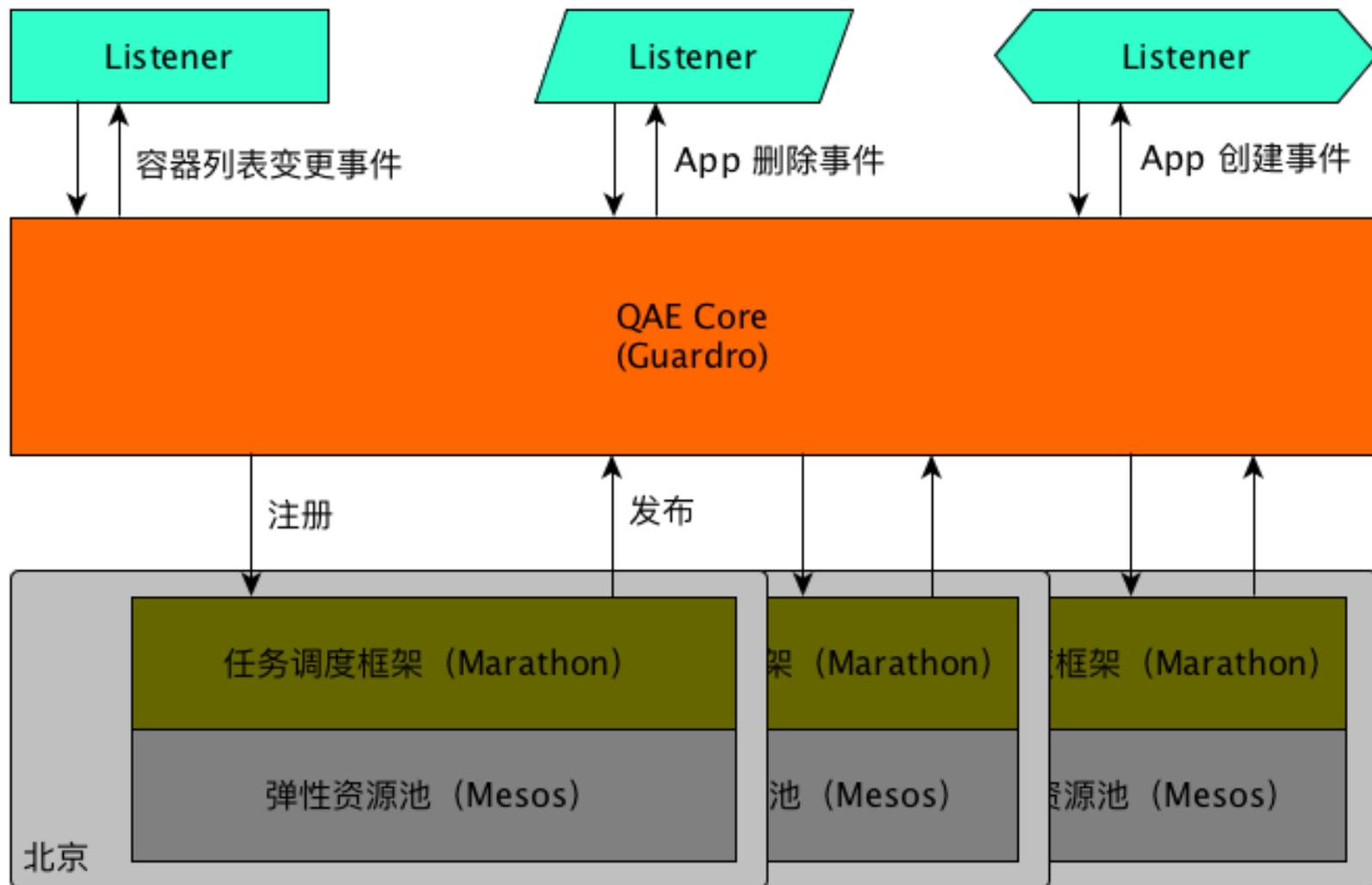
QAE — 监控和预警



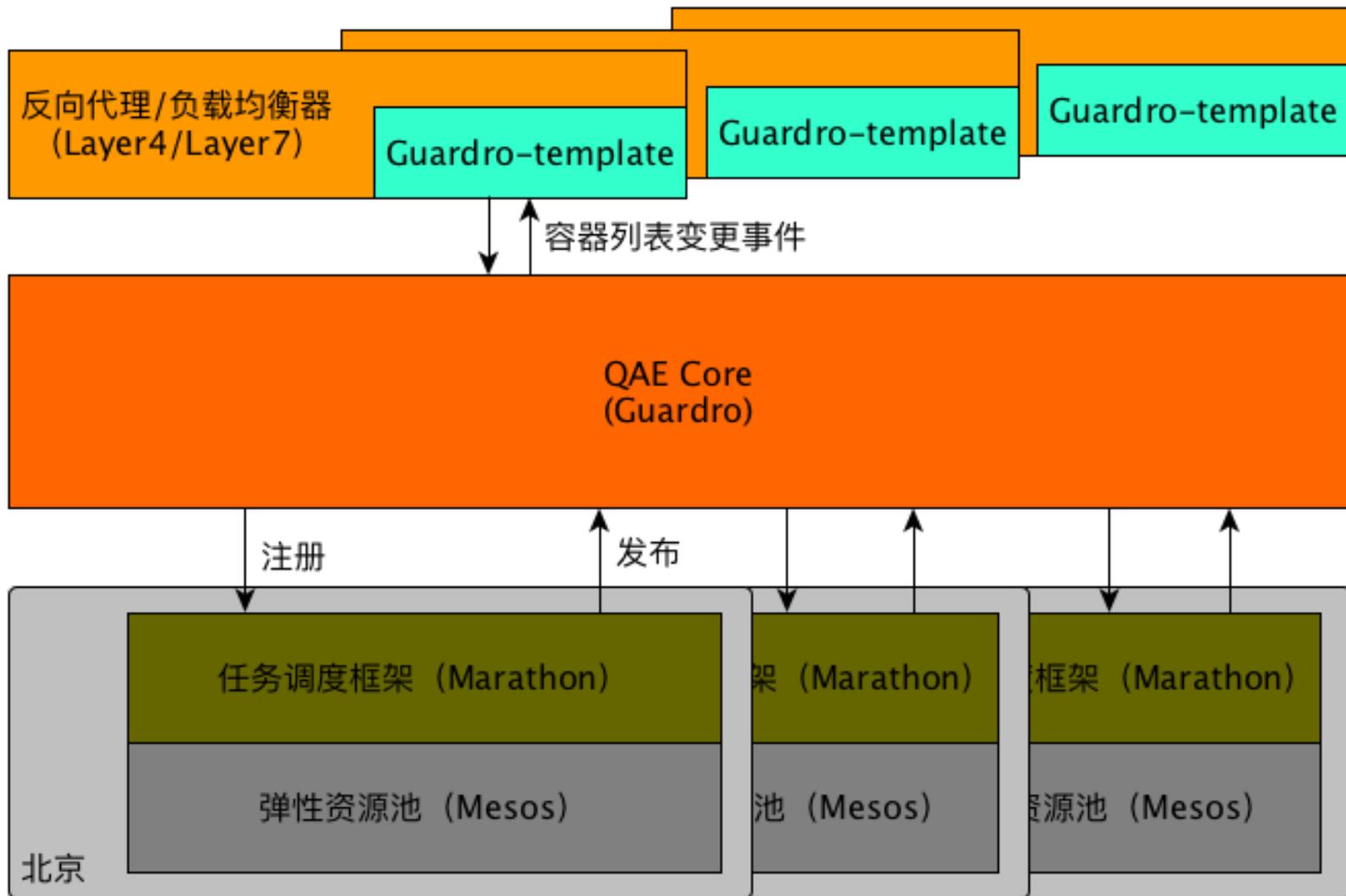
QAE — 自动扩展 (Auto Scaling)



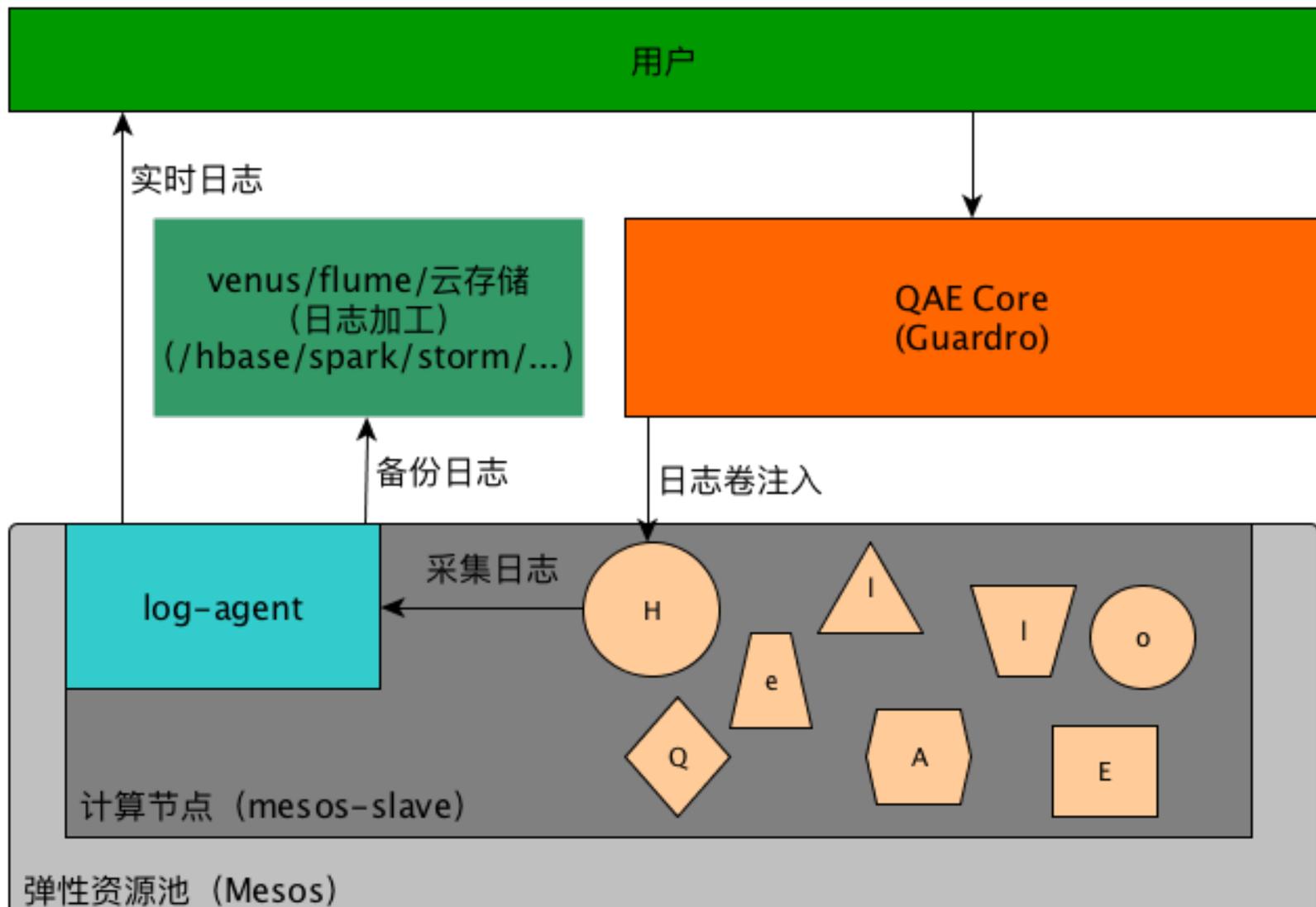
QAE — 事件总线



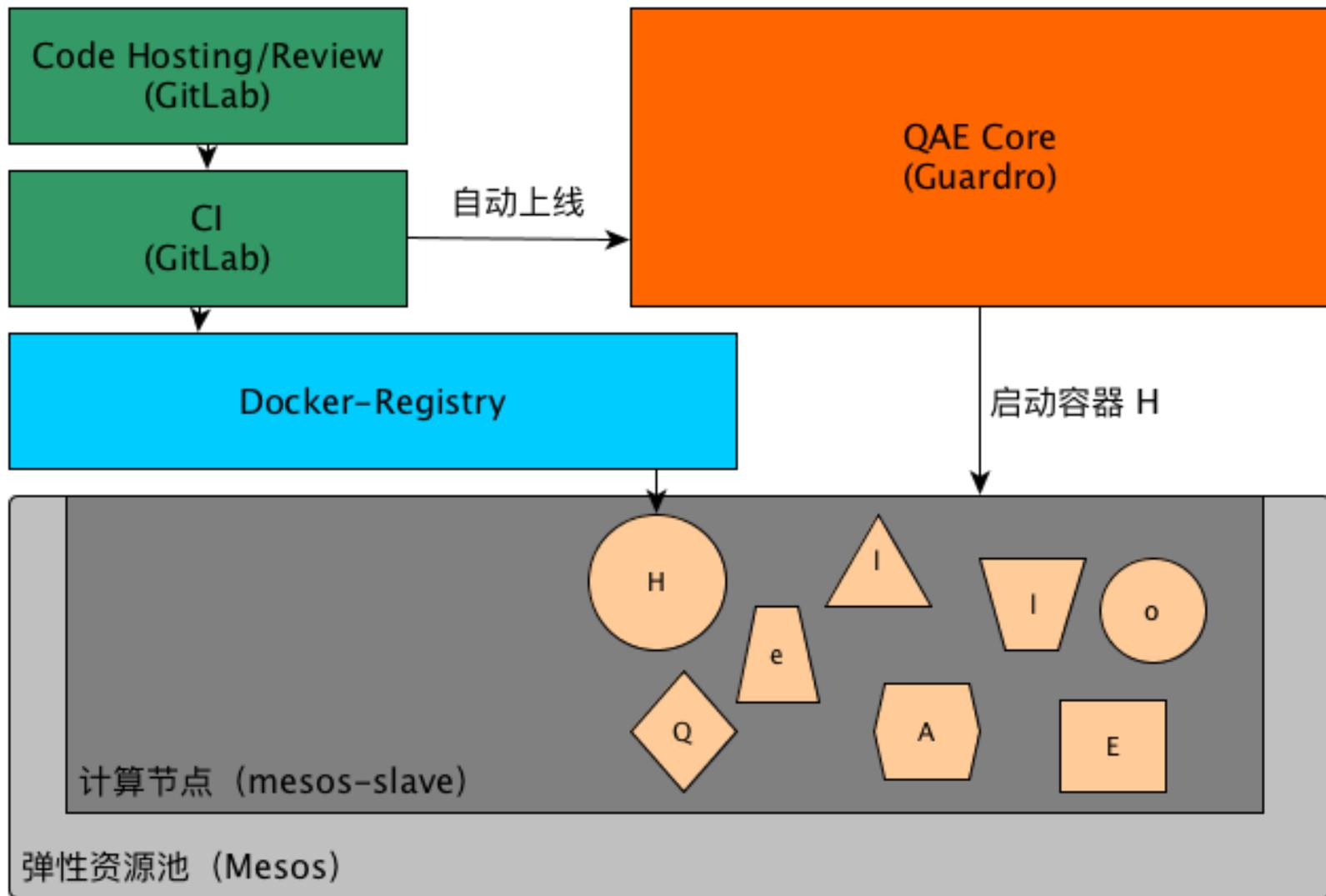
QAE — 服务发现



QAE 日志



QAE — CI/CD



内容提要

- 背景、出发点、目标
- iQIYI App Engine 设计与实现
- 一些经验（坑）
- 现状与展望

容器监控 — Zabbix

- 优势
 - 已有 Zabbix 基础服务
 - Auto Discovery 能发现容器启动/消失事件
- 劣势
 - 需要在容器内部安装 zabbix, overhead 太大
 - 不是很稳定
 - 数据统计比较复杂

容器监控 — cAdvisor

- 优势
 - 背靠 Google 且开源
 - 监控数据详尽
- 劣势
 - 集群化开发成本较高，Kubernetes heapster 实现
 - influxDB 而非 graphite (当时有基础服务)

容器监控 — dAdvisor

- Docker Advisor
- 采集：CPU, 内存, 磁盘, 网络
- 集成现有 Graphite
- 开发量不大

Time Series Database — Graphite

- 优势

- 老牌（ Since 1999 ），已有服务，经过验证
- 简单、内置 aggregator ， Render URL API

- 劣势

- 项目及社区已经几乎停滞
- 集群化配置复杂，需要详细规划
- 缺乏多用户认证，授权机制，黑白名单太单薄
- 数据删除缺少 API

服务发现 (Service Discovery)

- 负载均衡/反向代理方案
 - HAProxy : Marathon 支持 , 用户少
 - Nginx : 大量用户
- 服务发现
 - Flask 推送 , SSH , 不可靠 , 不安全
 - Ansible 推送 , SSH , 不可靠 , 不安全
 - event bus + guardro-template

日志

- ELK (Elasticsearch + Logstash + Kibana)
 - 延时较高 (分钟级)
 - 只支持格式化后的日志 (for machine)
- 用户诉求
 - 低延时 (秒级) , 准实时
 - 裸日志 (for real person)
 - 备份日志 (Apache Flume)

内容提要

- 背景、出发点、目标
- iQIYI App Engine 设计与实现
- 一些经验（坑）
- 现状与展望

现状

- 处于活跃开发状态
- 接入了几十个业务
- 主要业务类型：服务型，RPC，Worker
- 服务型业务 QPS 大于 10 万

展望

- Web Console
- 内置典型服务（如：Tomcat）数据统计
- Quota 和计费
- Docker 持久化卷和跨主机通信
- 支持更多类型服务：Cache, DB, MQ 等
- 支持短任务



悦 享 品 质