



QCon 全球软件开发大会
INTERNATIONAL SOFTWARE
DEVELOPMENT CONFERENCE

BEIJING 2017

基于 Mesos 搭建 PaaS 平台你可能 需要修的路

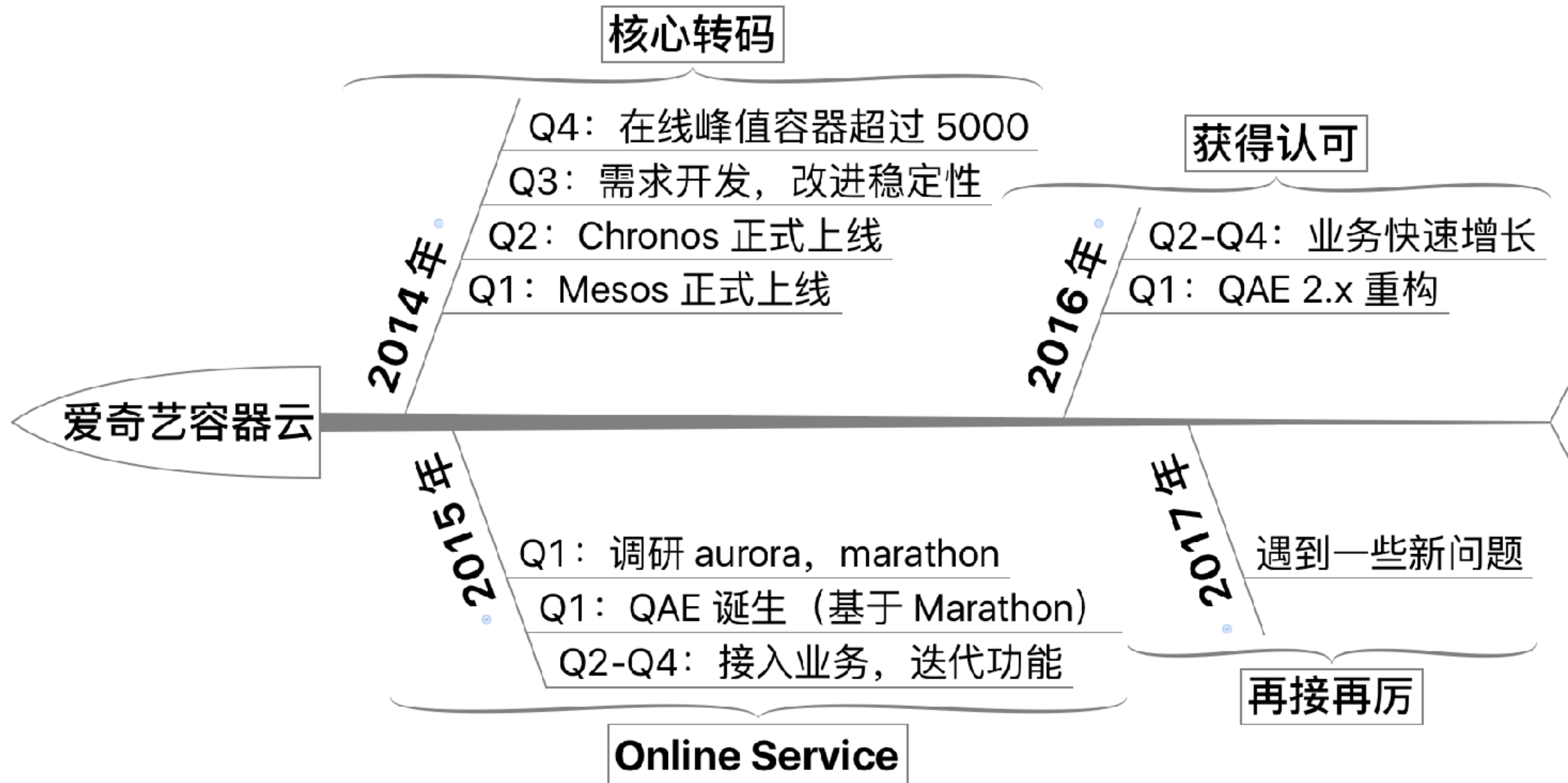
杨成伟@爱奇艺

内容提要

1. 背景介绍
2. QAE (iQIYI App Engine)
3. 延迟发布
4. 平滑升级
5. 灰度发布及 AB 测试
6. 日志
7. 总结及展望



容器云时间线



服务质量

业务对服务质量的要求越来越高：

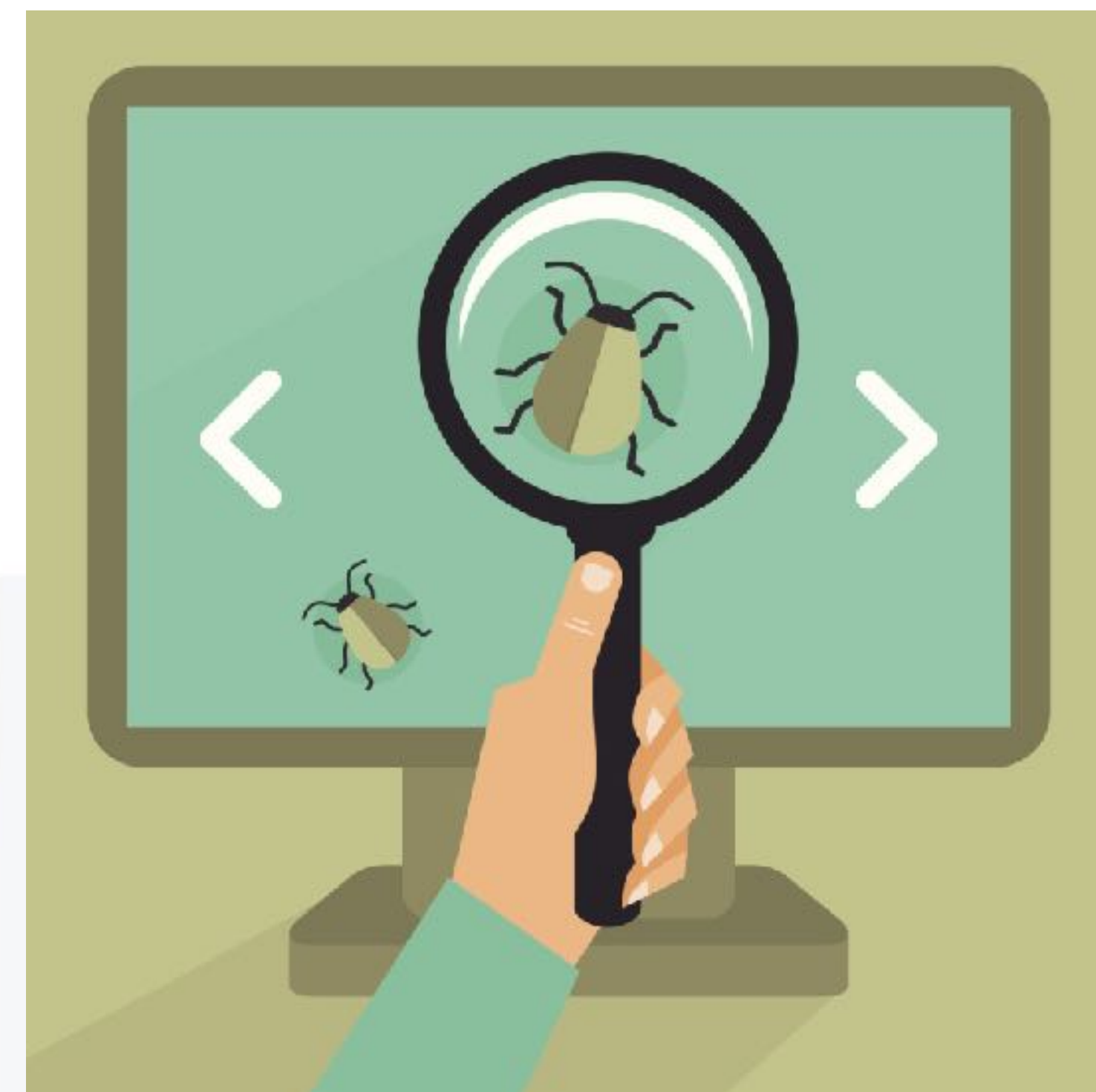
- 延迟发布，充分预热
- 流量缓释，避免容器压力
- 平滑升级，无流量丢失
- 灰度发布，发布过程可控
- AB 测试，能提供有效的对比手段
- 等等



自助排障

用户对自助排障的需求越来越多：

- 监控：系统监控、自定义监控
- 容器健康检查：基于应用指标均值？
- 在线排障：Web 控制台
- 历史容器
- 日志：实时日志、索引日志



内容提要

1. 背景介绍
2. QAE (iQIYI App Engine)
3. 延迟发布
4. 平滑升级
5. 灰度发布及 AB 测试
6. 日志
7. 总结及展望

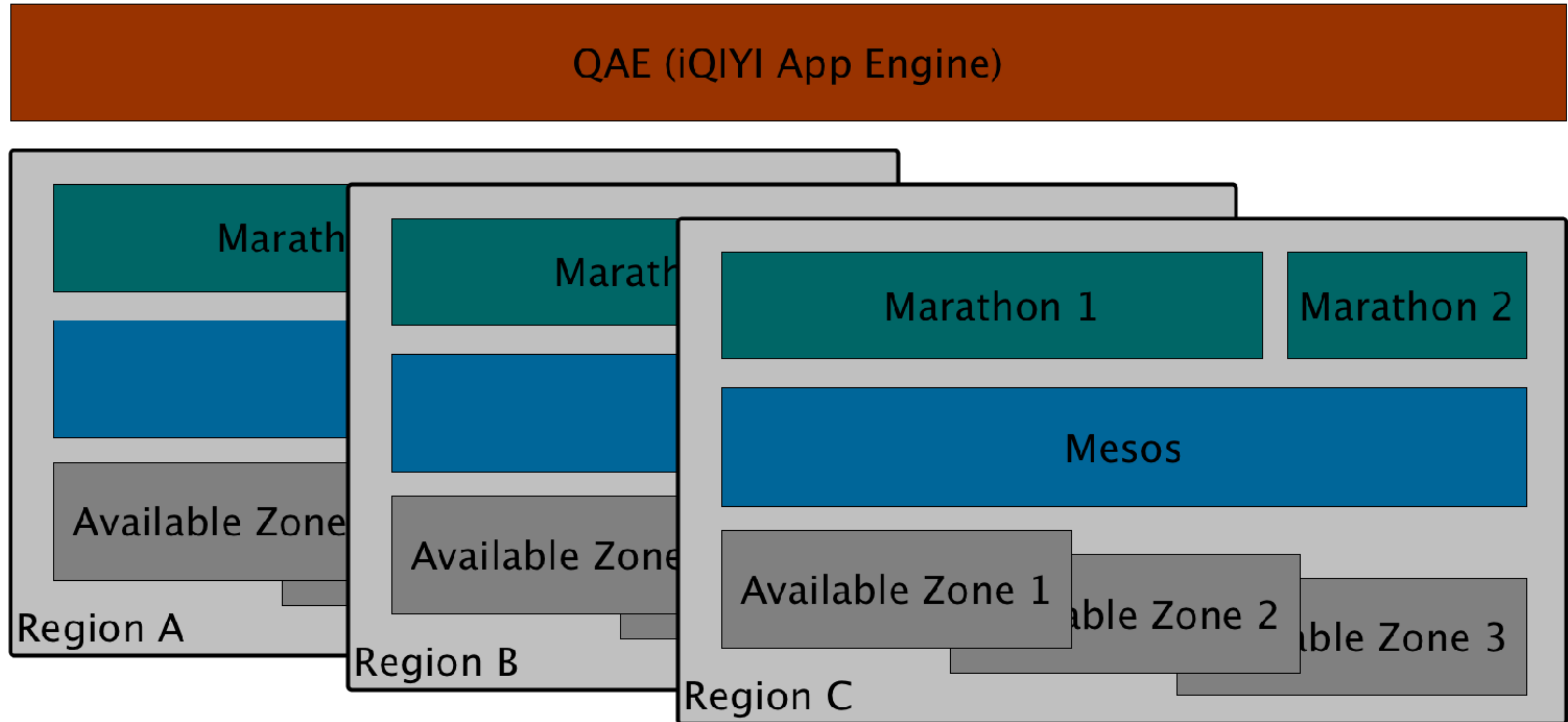


QAE 的基石——Marathon

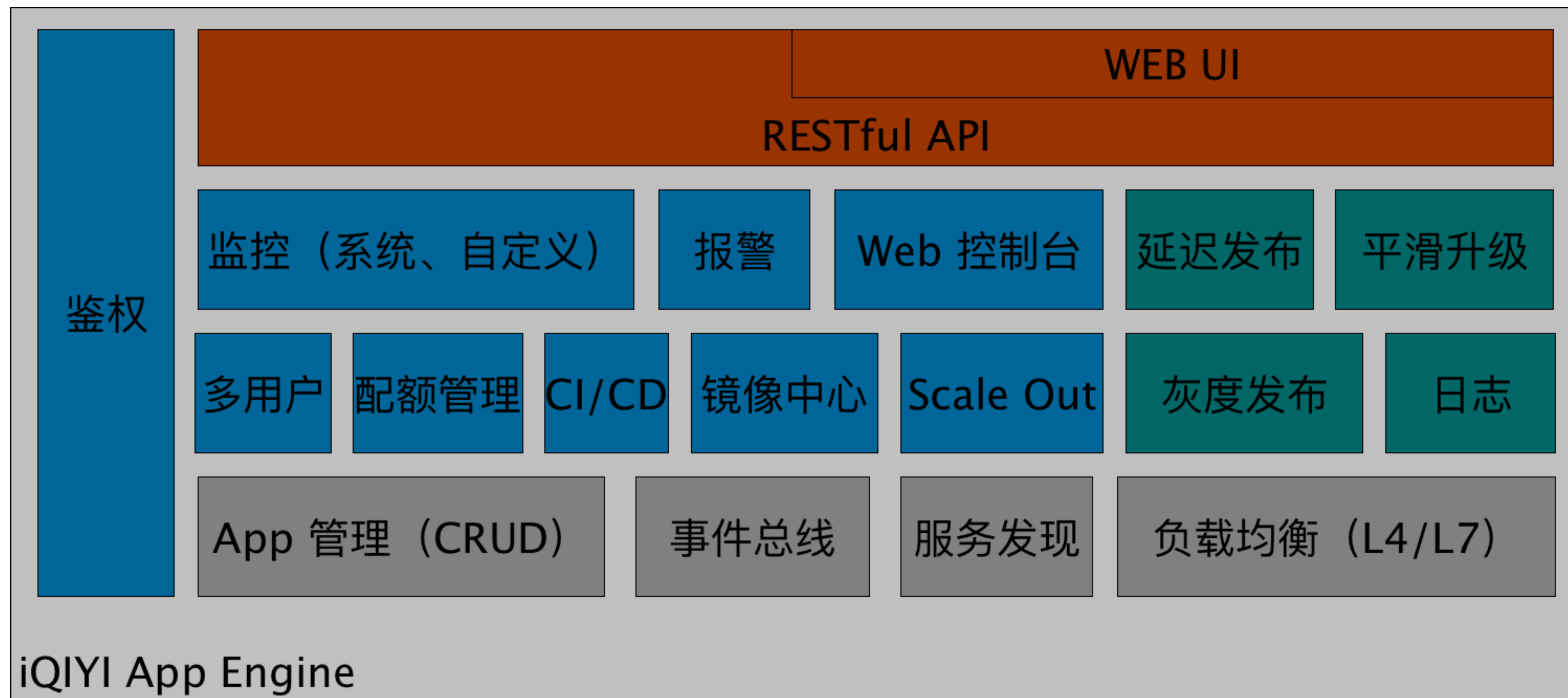
Marathon 主要特性：

- High Available, 高可用性
 - 容器分布限制
 - Web UI
 - RESTful API
 - Metrics
 - 服务发现和负载均衡
 - 健康检查
 - 事件总线
- *HA* -- run any number of Marathon schedulers, but only one gets elected as leader; if you access a non-leader, your request gets proxied to the current leader
 - *Constraints* - e.g., only one instance of an application per rack, node, etc.
 - *Service Discovery & Load Balancing* via HAProxy or the events API (see below).
 - *Health Checks*: check your application's health via HTTP or TCP checks.
 - *Event Subscription* lets you supply an HTTP endpoint to receive notifications, for example to integrate with an external load balancer.
 - *Marathon UI*
 - *JSON/REST API* for easy integration and scriptability
 - *Basic Auth* and *SSL*
 - *Metrics*: query them at `/metrics` in JSON format or push them to graphite/statsd/datadog.

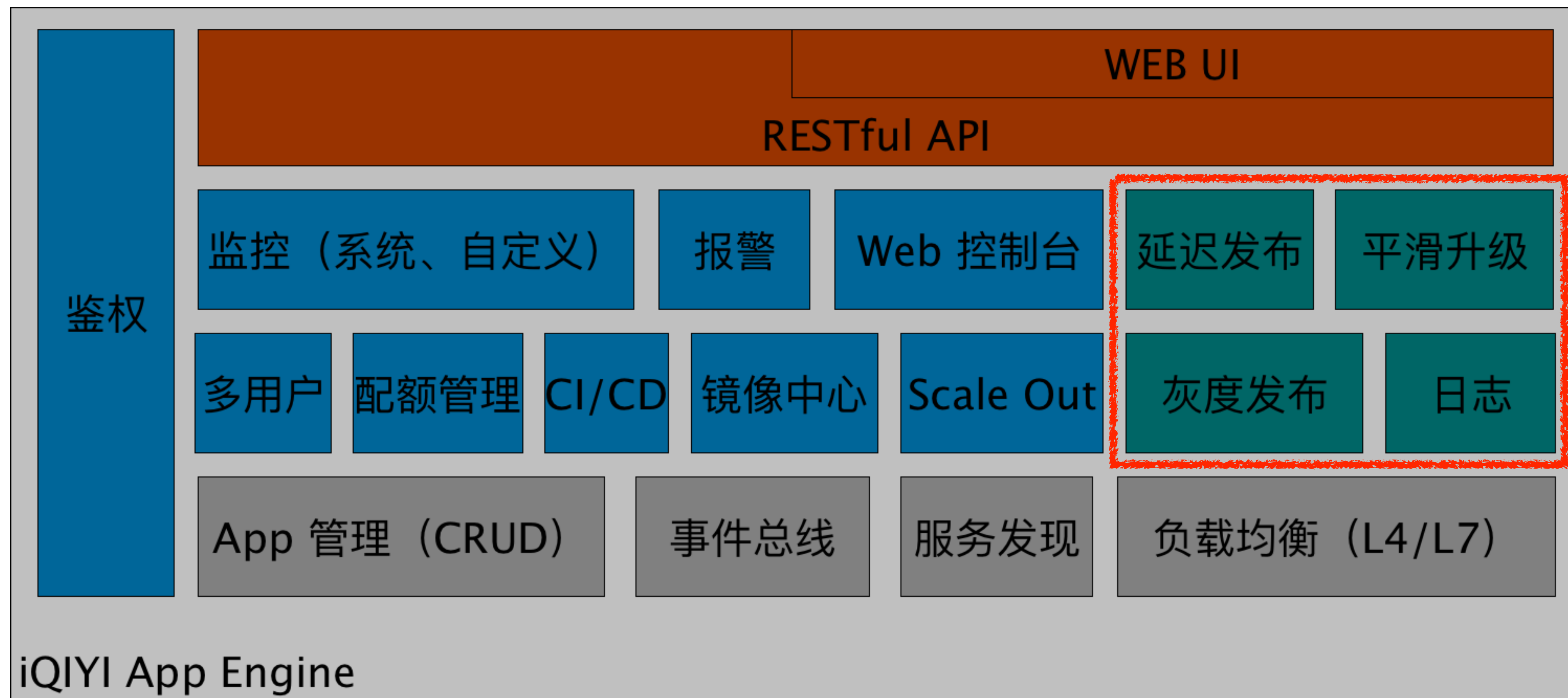
QAE 和 Marathon



QAE 的主要特性



QAE 的主要特性

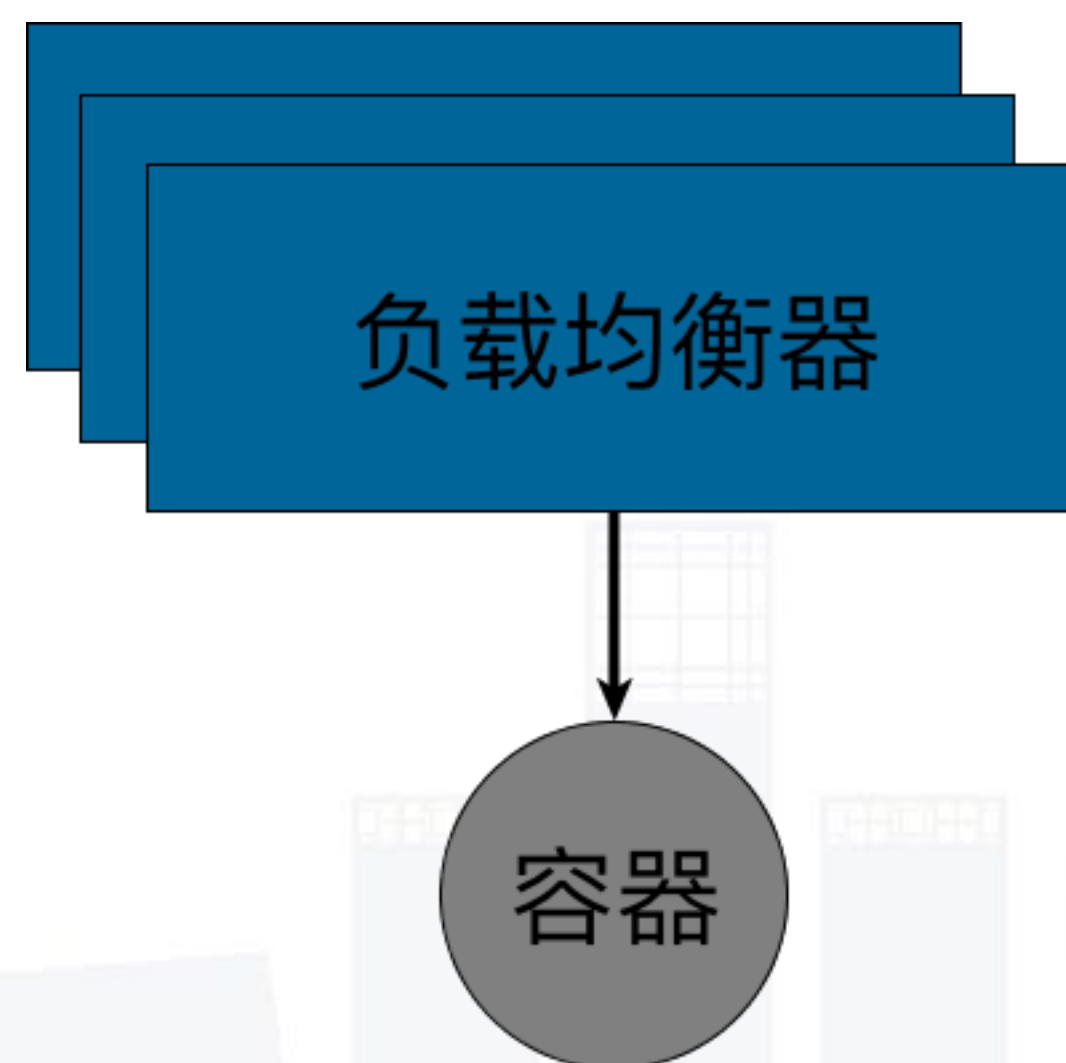


内容提要

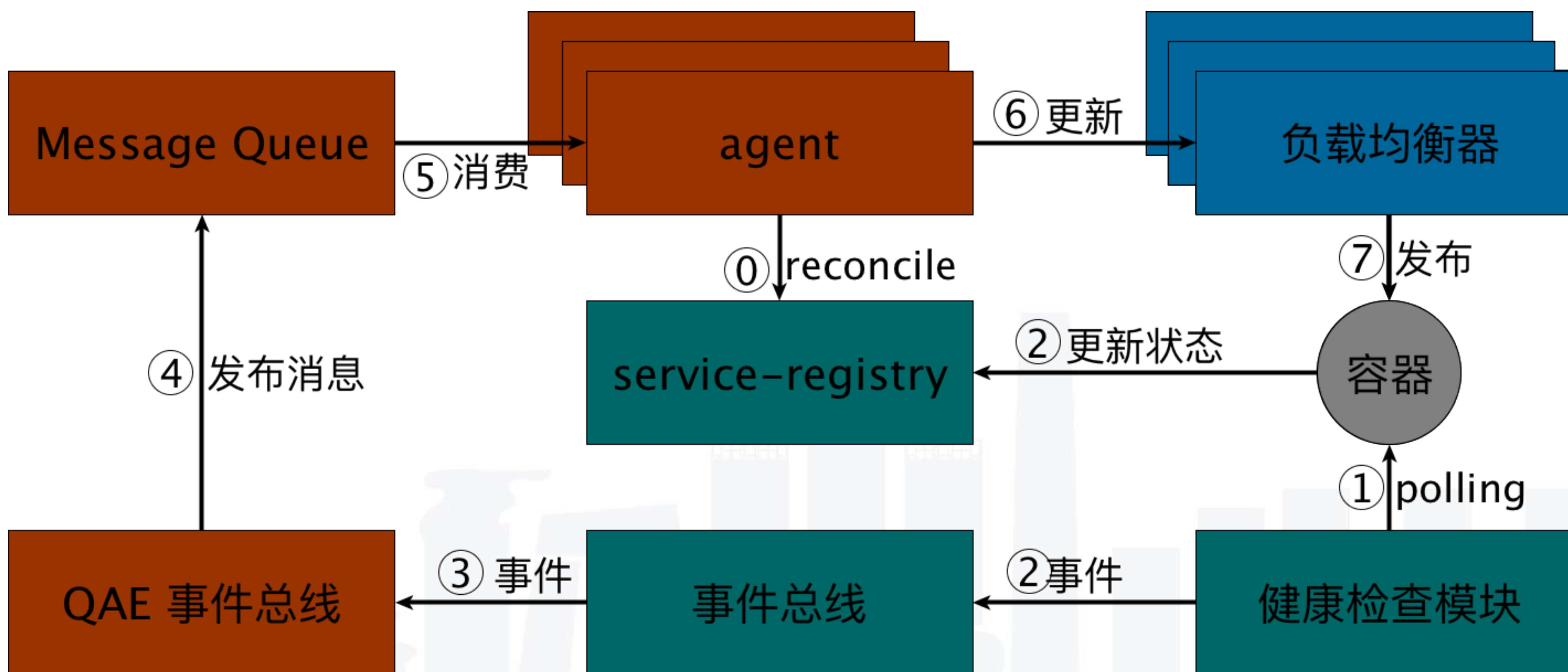
1. 背景介绍
2. QAE (iQIYI App Engine)
3. 延迟发布
4. 平滑升级
5. 灰度发布及 AB 测试
6. 日志
7. 总结及展望



服务发现与负载均衡



服务发现与负载均衡



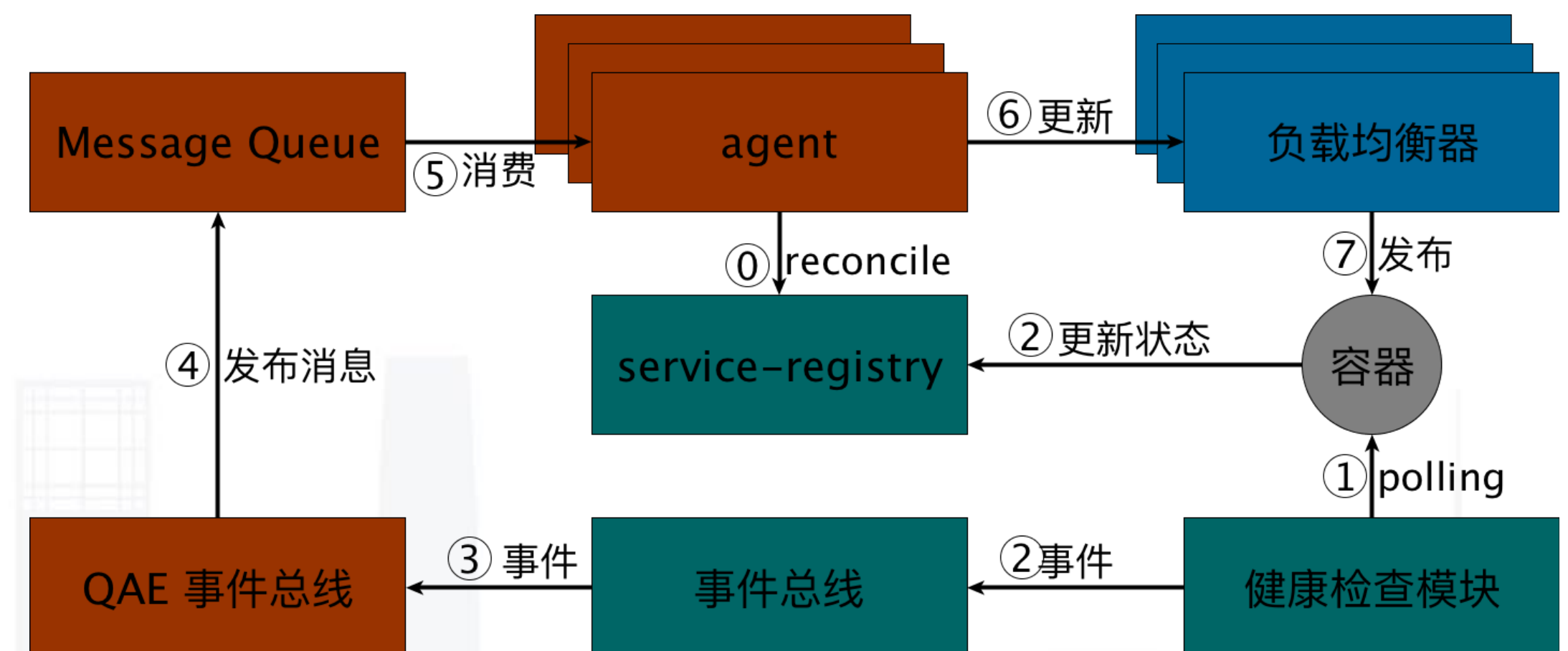
容器发布

Marathon:

- 能够通过 mesos 及时获得容器的状态更新
- 灵活、可配置的健康检查机制

QAE:

- 可配置的最小发布延时
- **只发布**健康的容器
- 流量缓释



可配置的最小发布延时

Marathon gracePeriodSeconds:

- 在这段时间内，忽略健康检查失败
- 但是容器一旦健康，立即发布
- 依赖健康检查准确性（扔给用户）

QAE gracePeriodSeconds:

- 容器启动后，**至少**等待这么长时间
- 然后检查容器健康状态，健康则发布，失败则重启



只发布健康的容器

Marathon:

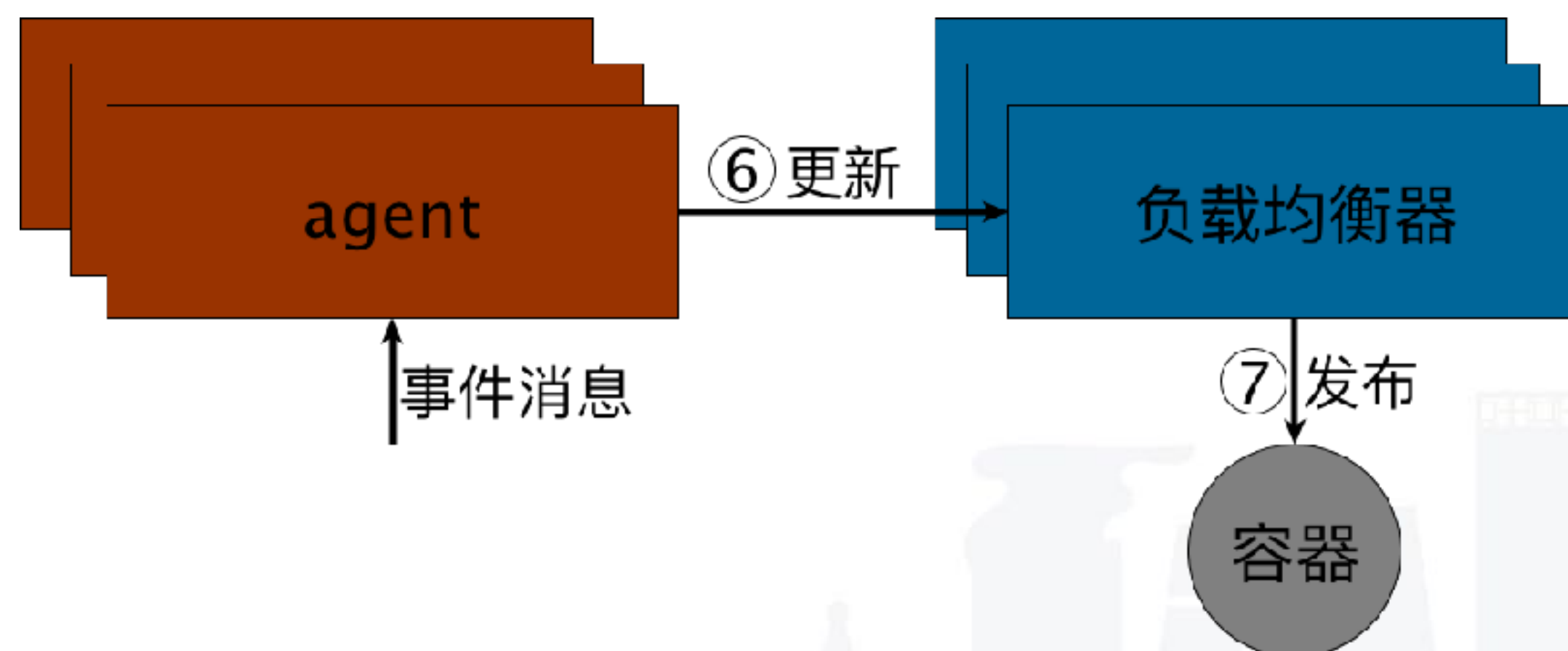
- 虽然，能标记容器的健康状态
- 但是，没有持久化容器的健康状态
- 切换 Leader, **BANG!**

QAE:

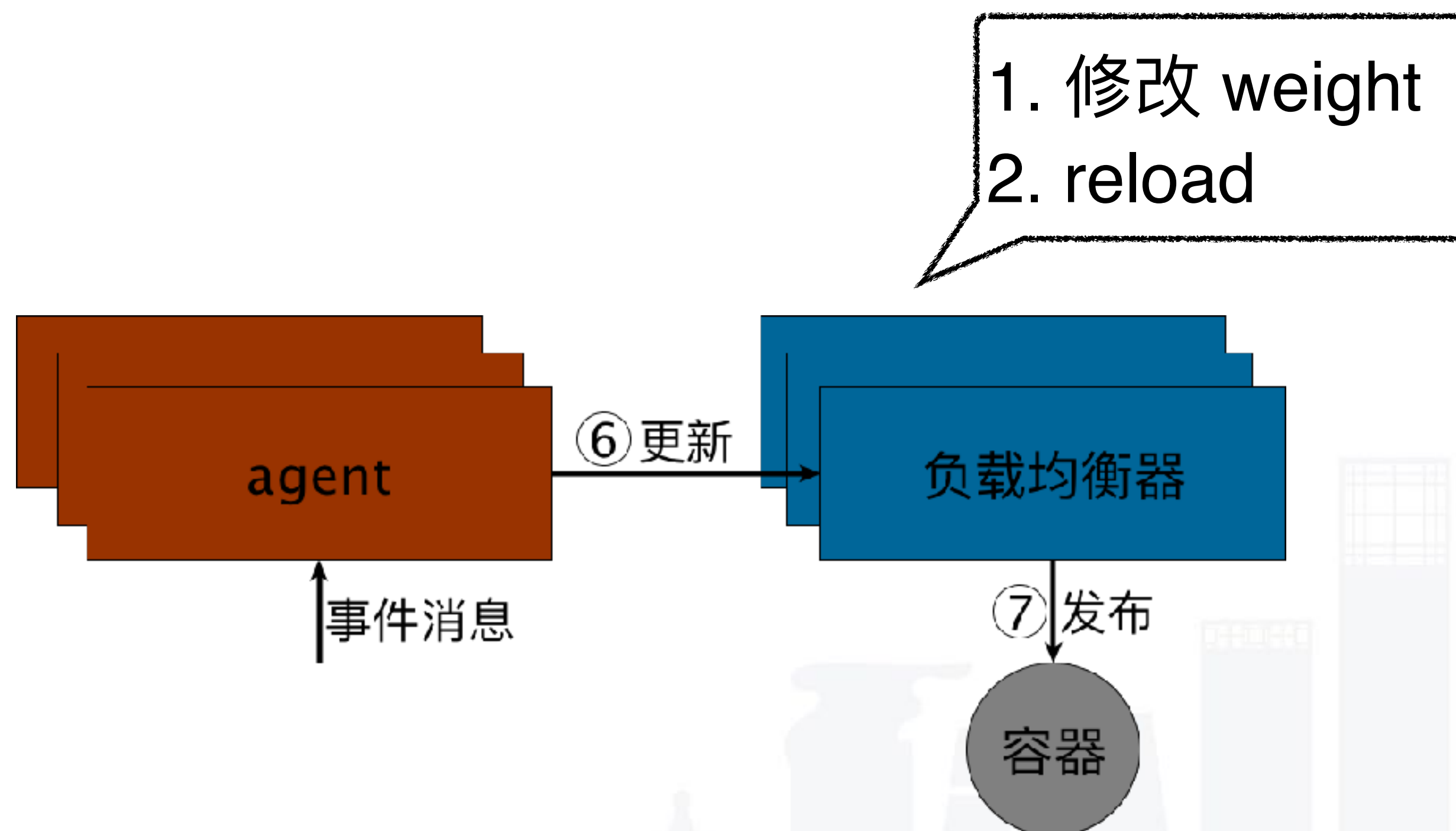
- 持久化容器的最近一次健康检查结果



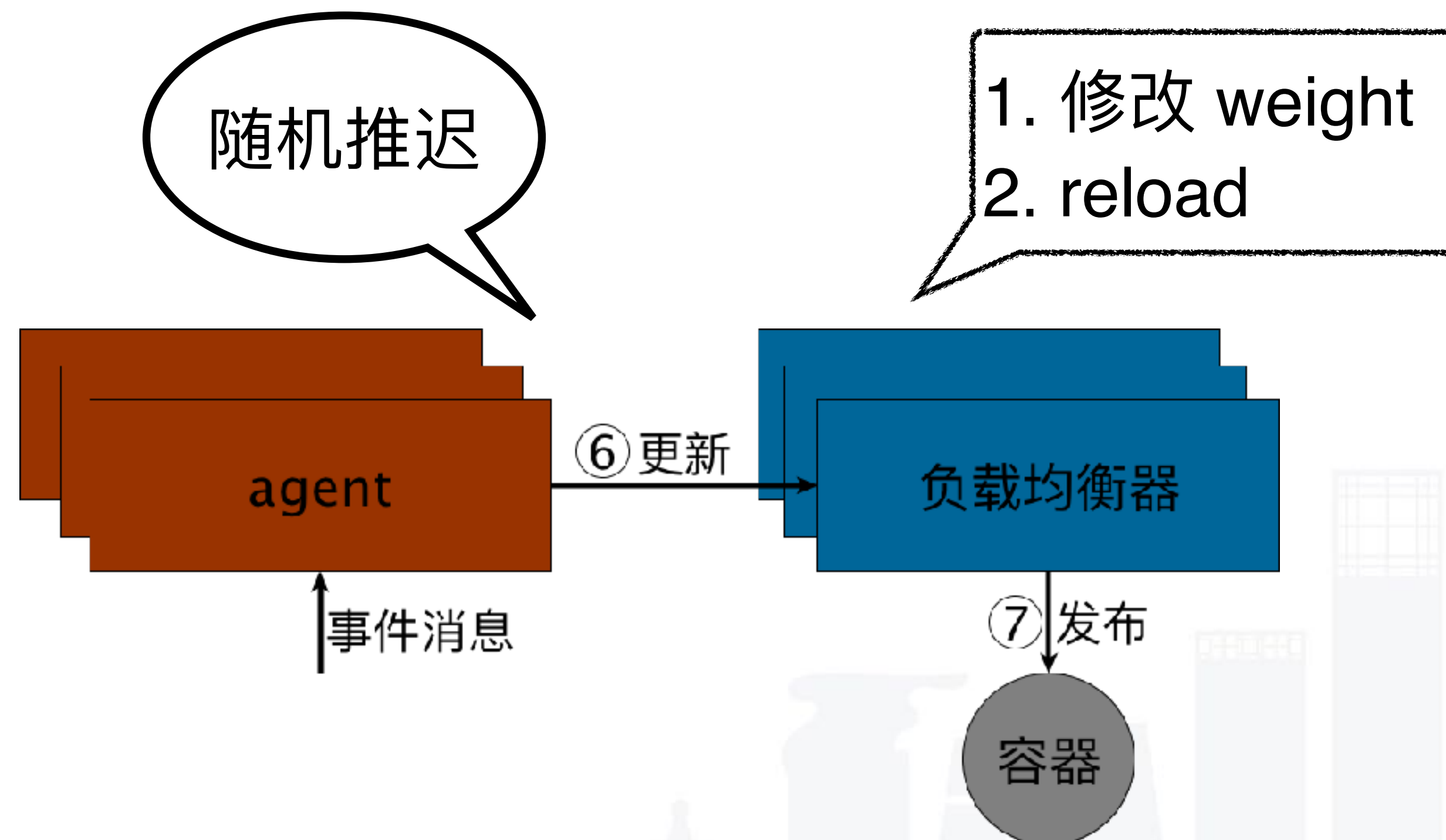
流量缓释



流量缓释



流量缓释



内容提要

1. 背景介绍
2. QAE (iQIYI App Engine)
3. 延迟发布
4. 平滑升级
5. 灰度发布及 AB 测试
6. 日志
7. 总结及展望



Marathon——滚动升级



升级过程：

1. 根据配置和资源启动一批新容器
2. 新容器启动或健康后，停止旧容器
3. 重复上述过程至所有容器升级完成

不足：

- 新容器一旦健康，立即发布
- 杀掉旧容器前不能先将其下线
- 升级过程不可控，不支持灰度发布

Marathon——滚动升级



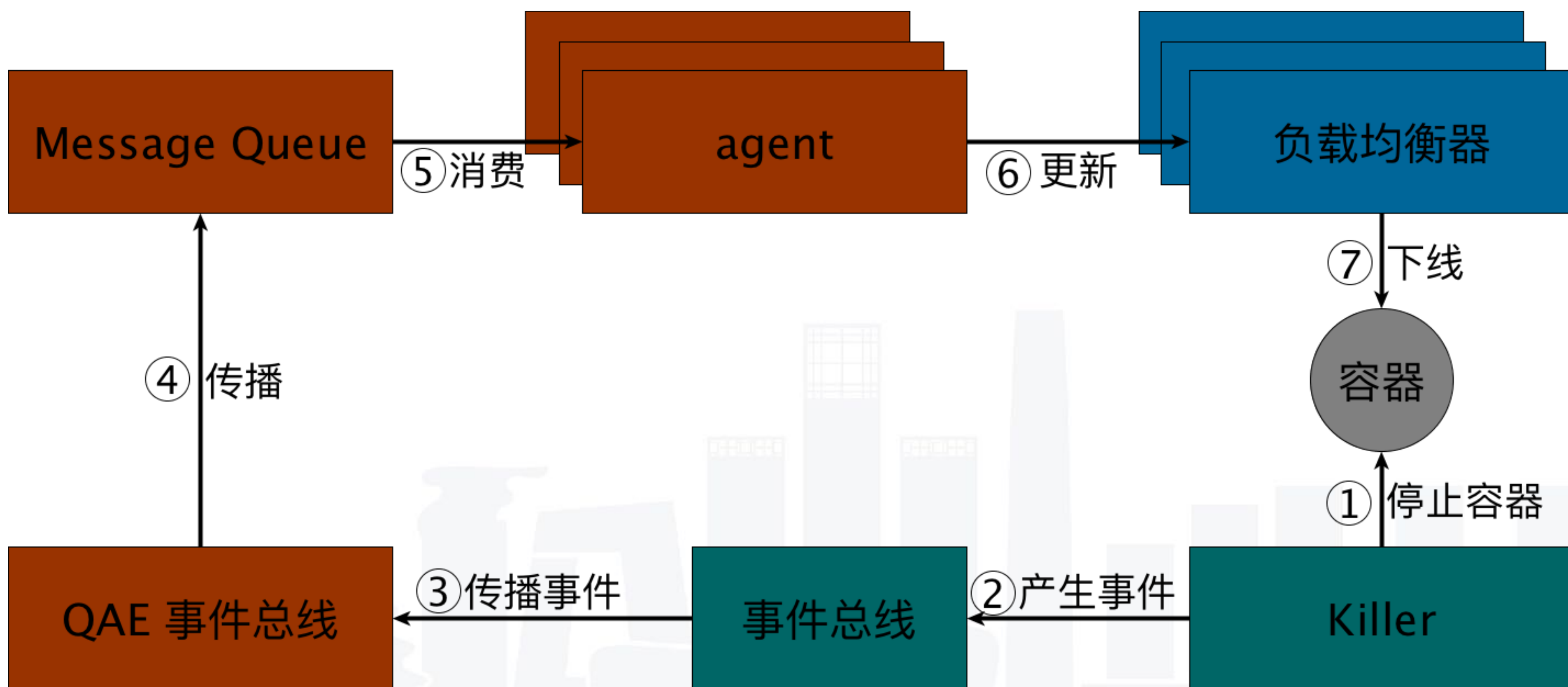
升级过程：

1. 根据配置和资源启动一批新容器
2. 新容器启动或健康后，停止旧容器
3. 重复上述过程至所有容器升级完成

不足：

- 新容器一旦健康，立即发布
- 杀掉旧容器前不能先将其下线
- 升级过程不可控，不支持灰度发布

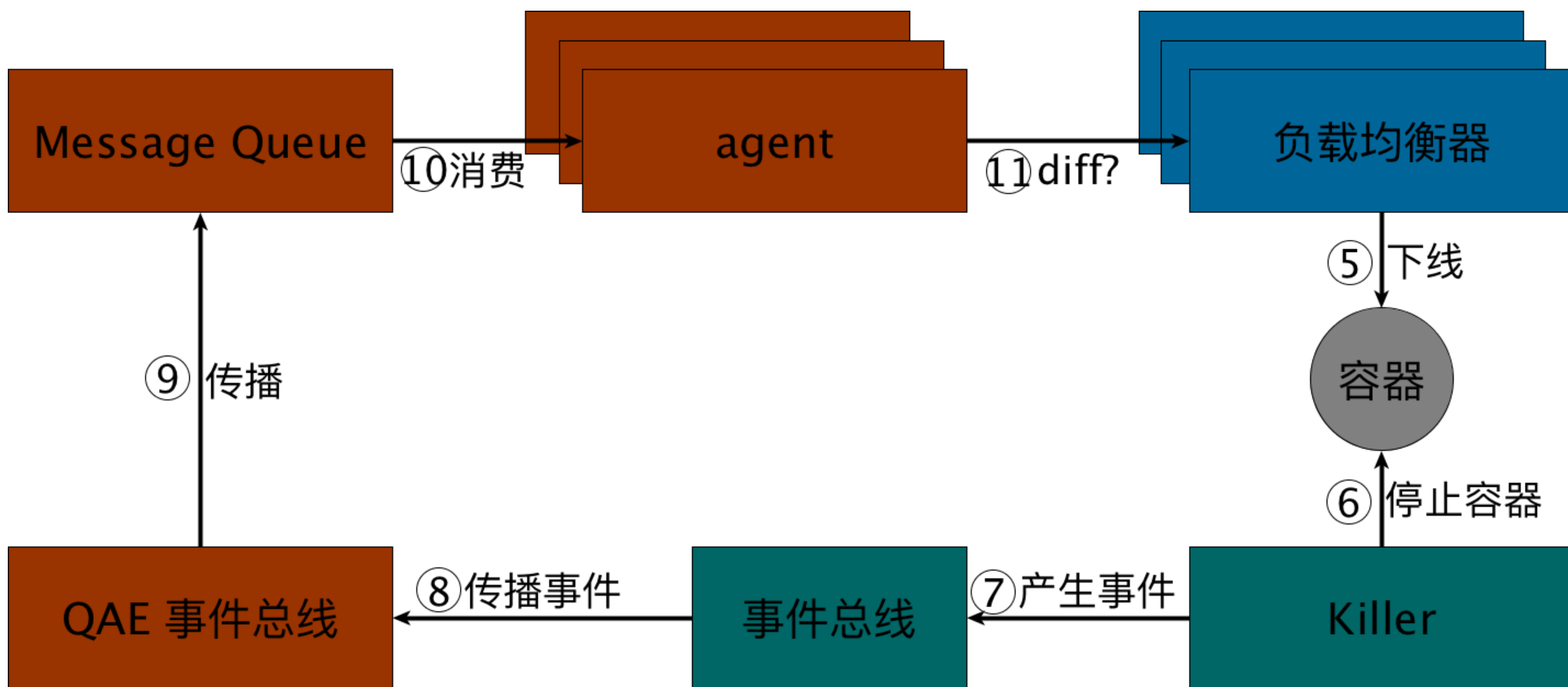
先斩后奏



先奏后斩



先奏后斩



先奏后斩

能够处理所有可预知的容器下线：

- 平滑升级、灰度发布及 AB 测试
- 手动停止、重启**指定**容器
- 手动/自动 Scale out (**随机**容器)

不能处理不可预知的容器下线：

- failover

内容提要

1. 背景介绍
2. QAE (iQIYI App Engine)
3. 延迟发布
4. 平滑升级
5. 灰度发布及 AB 测试
6. 日志
7. 总结及展望



灰度发布及 AB 测试

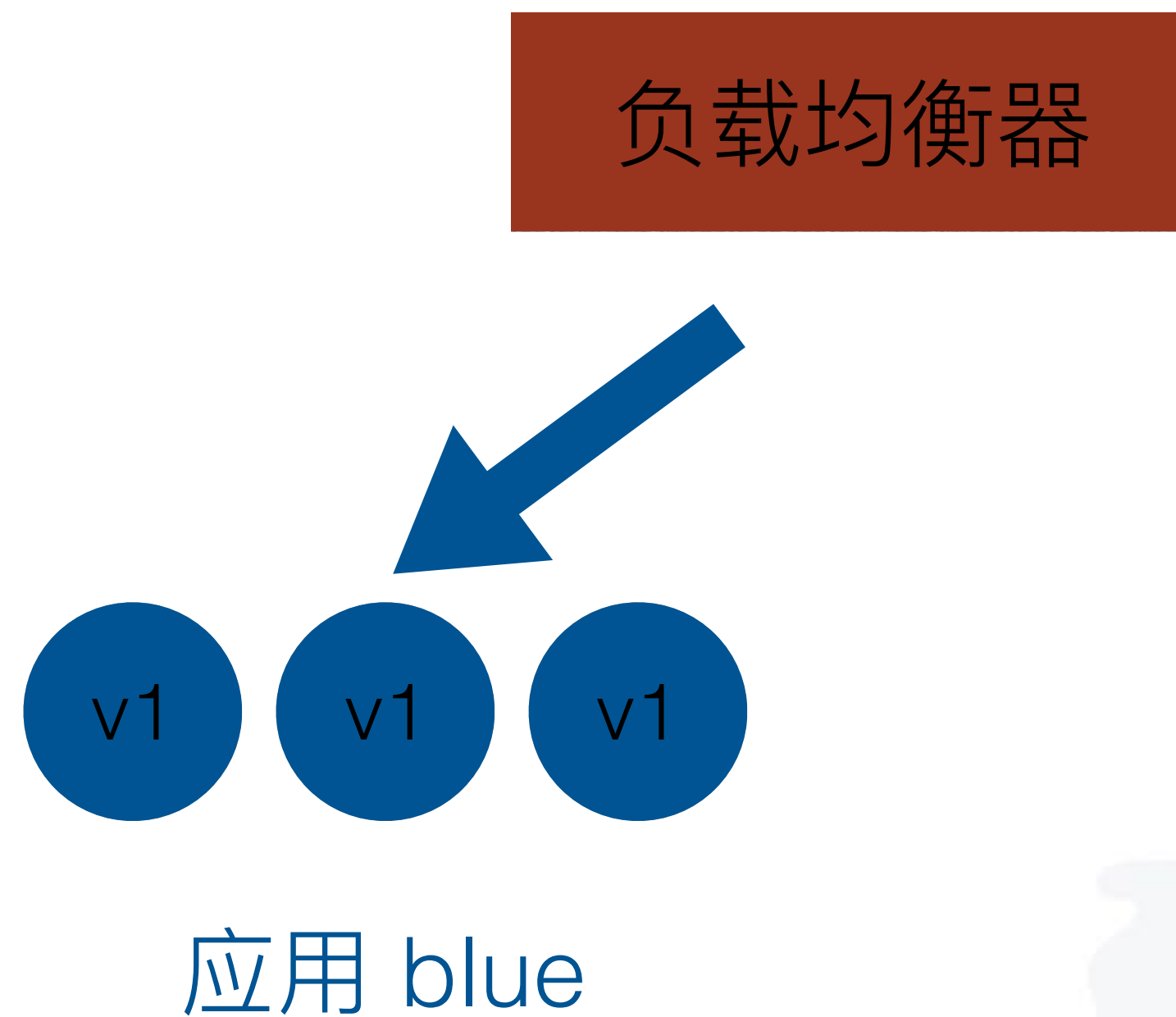
我们有什么？



我们缺什么？



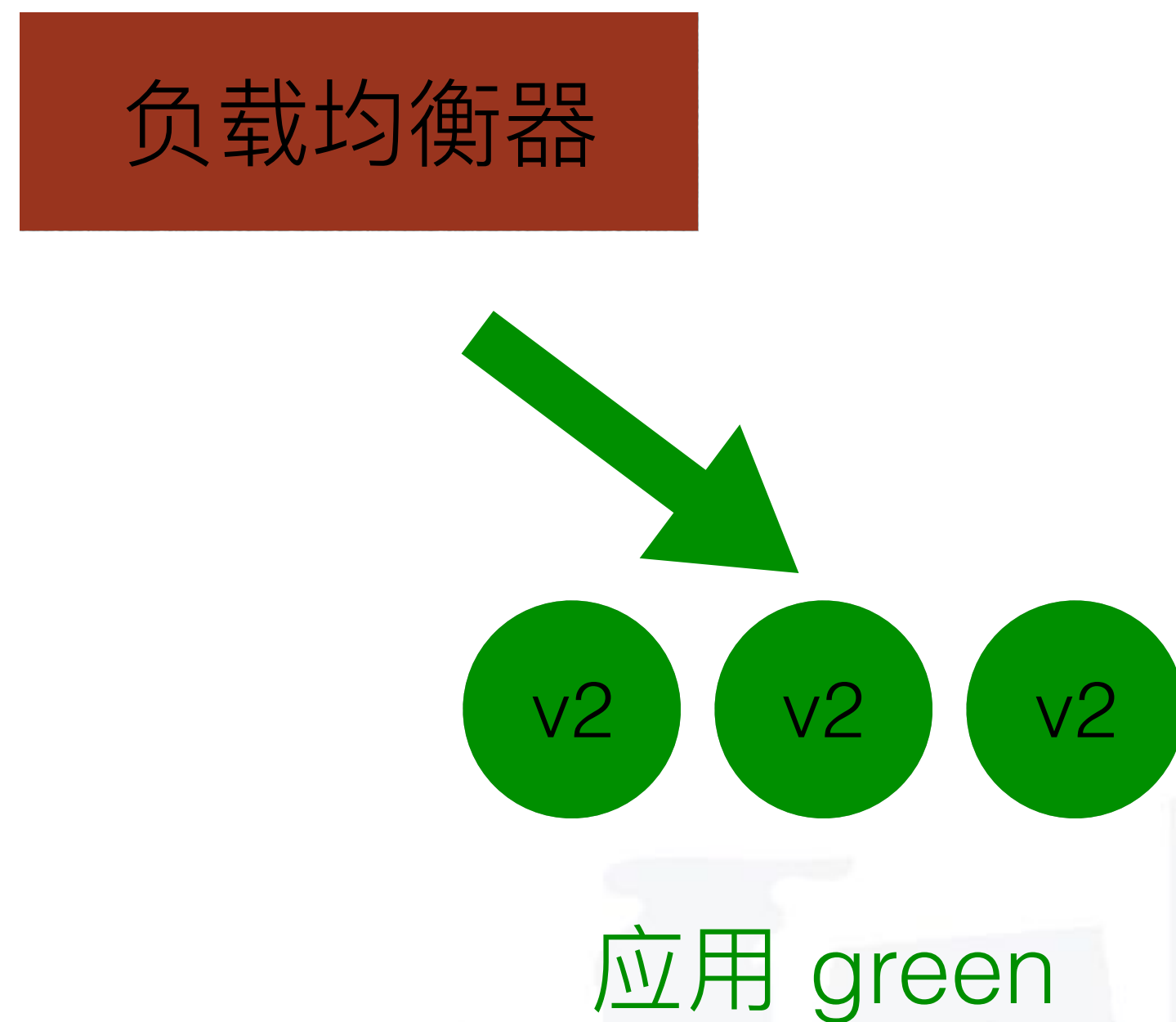
Blue-Green Deployment



看着很眼熟?

- 这不是 marathon rolling upgrade
- 需要**两个** marathon 应用来实现
- 分别控制应用 blue 和 green 的容器增减来实现

Blue-Green Deployment



看着很眼熟?

- 这不是 marathon rolling upgrade
- 需要**两个** marathon 应用来实现
- 分别控制应用 blue 和 green 的容器增减来实现

灰度发布及 AB 测试

直接修改 marathon?

- 社区很早就有类似需求[\[jira-4396\]](#)
- 现有代码架构实现起来很复杂

QAE 的实现

- 基于 blue-green deployment
- 保持向后兼容 (backward compatibility)
- 对用户暴露一个应用，**两个版本**
- 独立的监控 (AB 测试)
- 独立的容器列表 (kill, restart)

内容提要

1. 背景介绍
2. QAE (iQIYI App Engine)
3. 延迟发布
4. 平滑升级
5. 灰度发布及 AB 测试
6. 日志
7. 总结及展望



stdout & stderr



I build open source

- 日志散落在 /data/app/log 目录下
- foo.log, bar.log, deadbeef.log
- 为什么要改成 stdout, stderr?
- 你帮我改?

日志需求

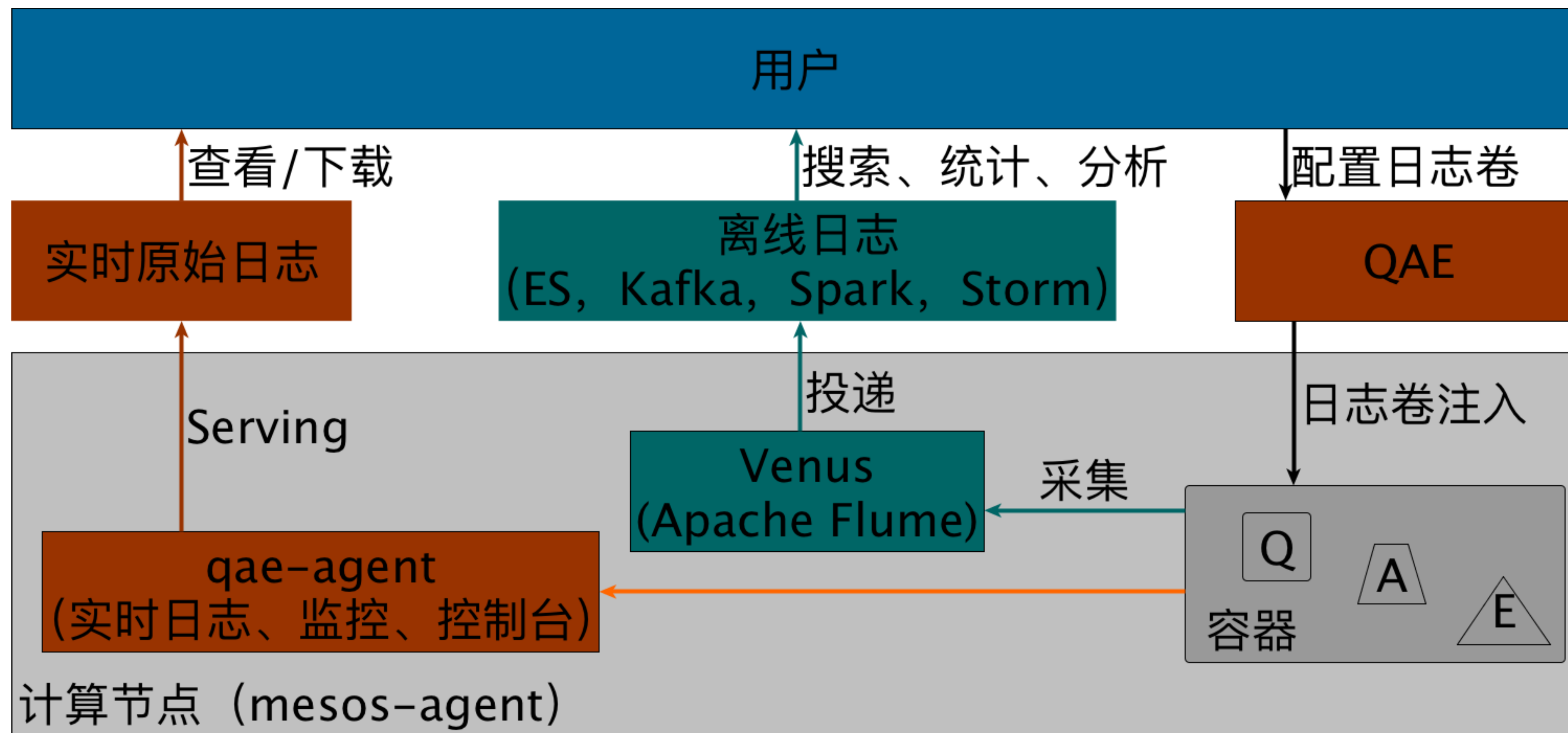
强需求：

- **实时**日志，业务故障了，火烧眉毛
- **原始**日志，怀疑日志搬运工，是不是搬错了
- 可查看，下载，方便快速定位，一眼看不出来再用脚本查找

弱需求：

- 聚合
- 索引、搜索：标签、全文
- 统计、分析
- 呈图、报表

QAE 的日志架构



卷注入

主机路径前缀

容器 ID (唯一)

卷类型

/data/qae/injected-volumes/**app-name**/**container-uuid**/**log** : **/qae/log**

应用名 (唯一)

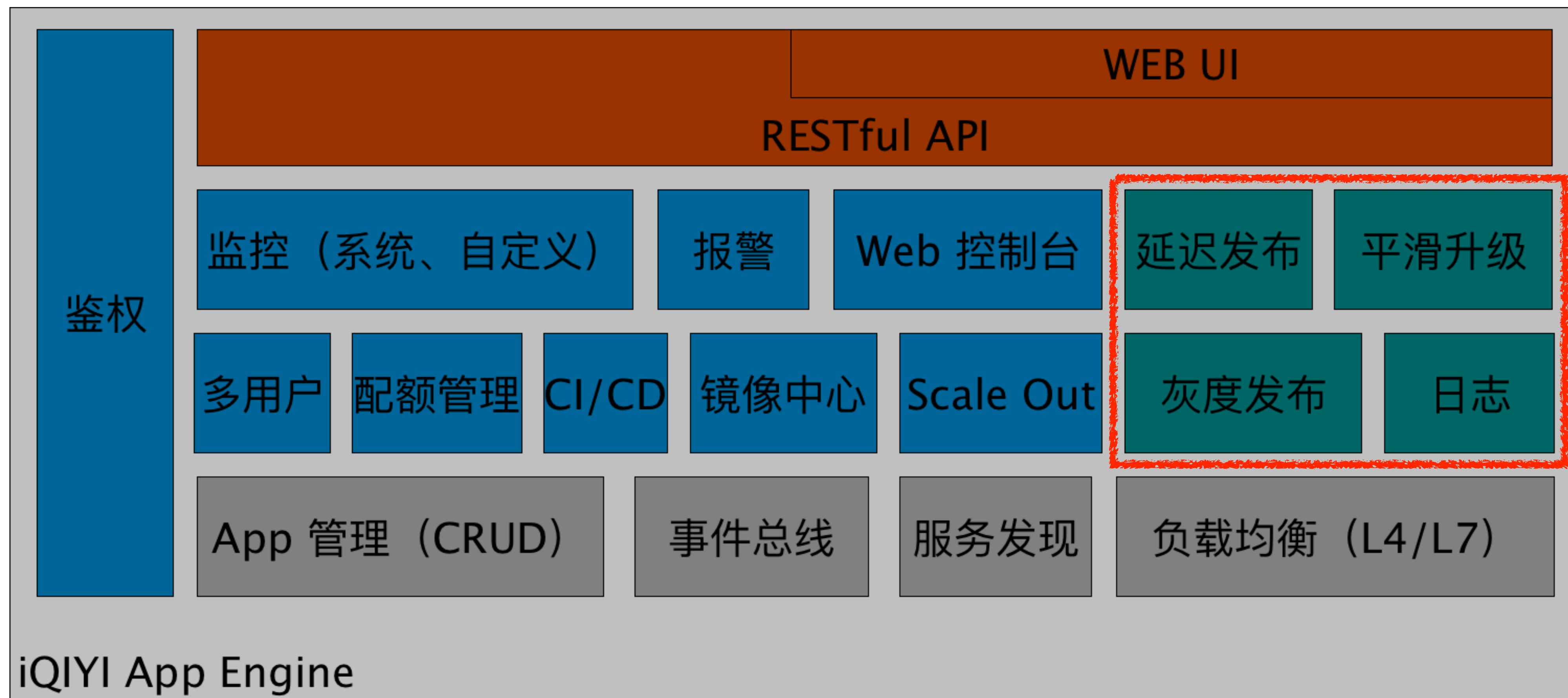
容器内路径 (可配置)

内容提要

1. 背景介绍
2. QAE (iQIYI App Engine)
3. 延迟发布
4. 平滑升级
5. 灰度发布及 AB 测试
6. 日志
7. 总结及展望



总结



展望

服务中心

分布式健康检查

API Gateway

oversubscription

隔离性

cpuset

混合部署

Mesos Unified Containerizer

存储

虚拟化

网络



关注QCon微信公众号，
获得更多干货！

Thanks!



主办方 **Geekbang** > **InfoQ**
极客邦科技