



# 后台数据处理架构演进

-- 与QingCloud云资源的实践

——泰捷软件 陈剑豪

# 概览

- 业务介绍
- 架构的演变
- LB
- 运维

# 现状

- 注册用户
  - 20m
- 活跃用户
  - 2.5m
  
- 压力
  - 服务器规模/配置
    - API服务器——50台/4核4G
    - DB服务器——20台/8核8G
    - 其它
  - API调用量：600k/min
  - 单台数据库压力：10k/min
  - 灾备
    - 1.5倍

# 在很久很久以前...

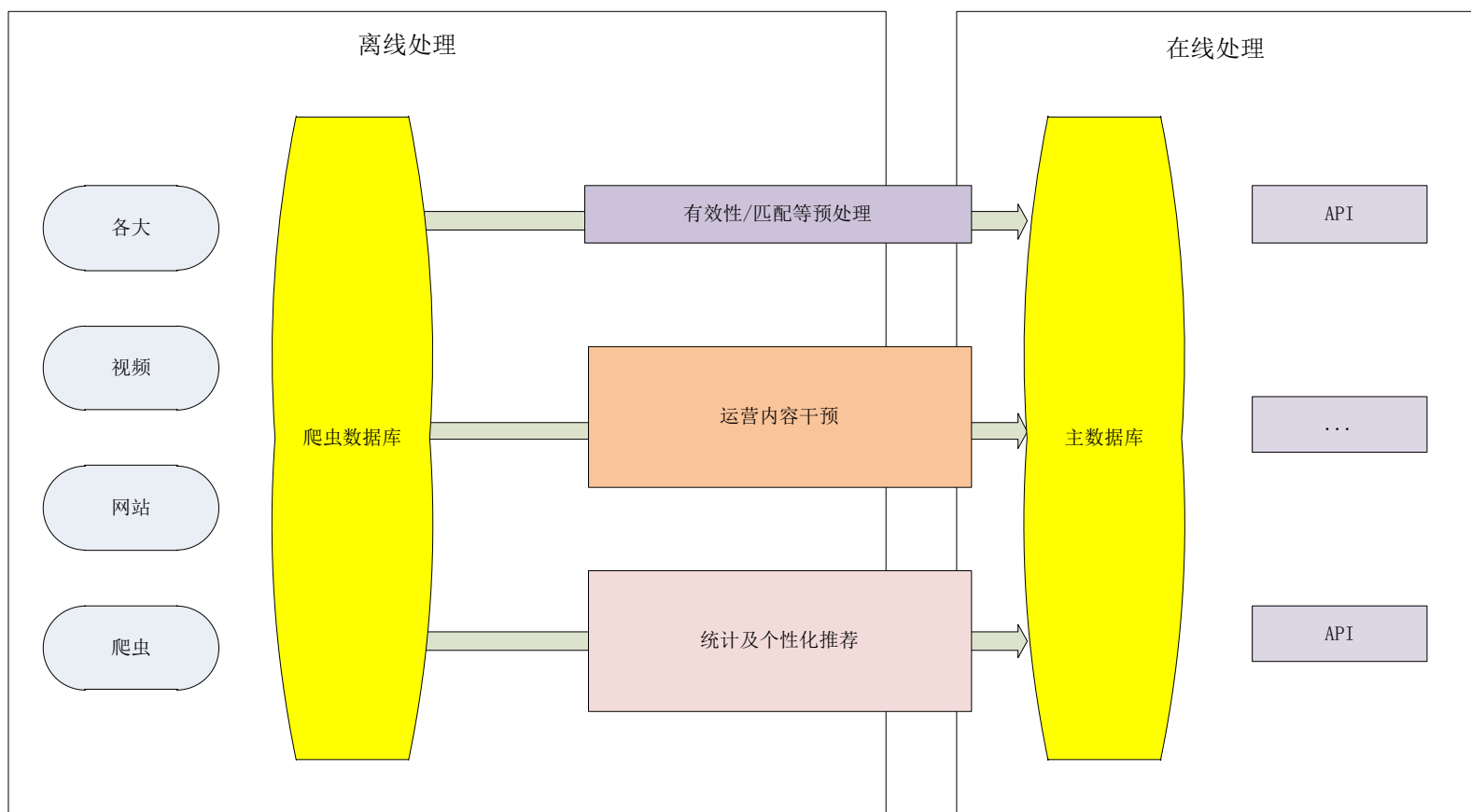
- 一台物理服务器
  - 一只爬虫
  - 一个后台管理+API
  - 一个数据库
    - 面向客户端
    - 面向爬虫
    - 面向后台管理



# 系统的拆分

- **水平拆分**
  - 预计算/实时计算，拆分成**离线/在线**处理
- **业务拆分**
  - 元数据
  - 自动更新数据
  - 内容运营数据
- 提高系统的稳定性
  - 数据库中间件/异步队列/匀速队列
- 实现“半自动化”的离线处理

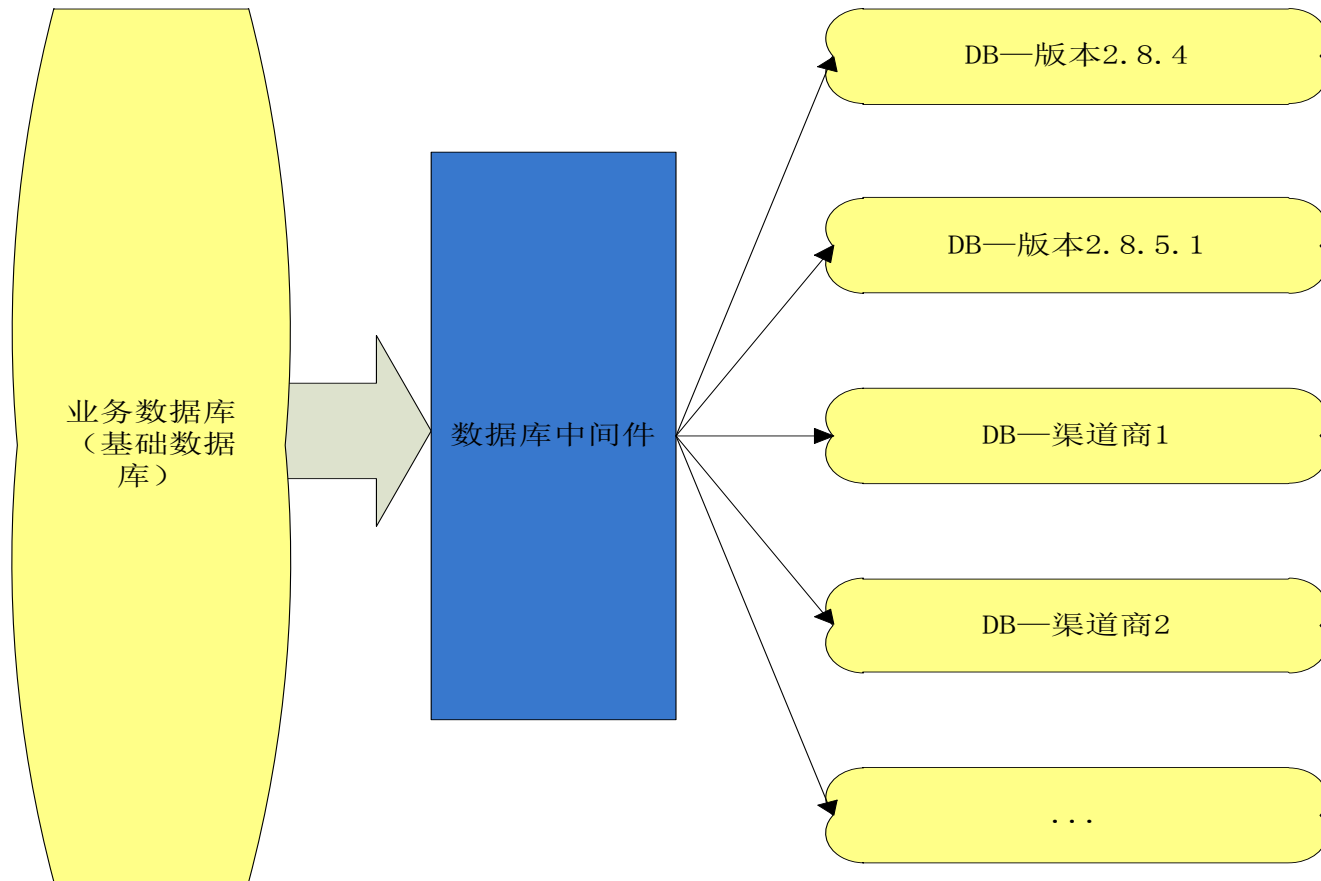
# Stage 1——针对离线处理的架构改进



# 数据库的拆分

- 剥离离线/在线处理
- 数据的多版本变体支持？
  - 版本、设备、渠道商.....

# 数据库中间件

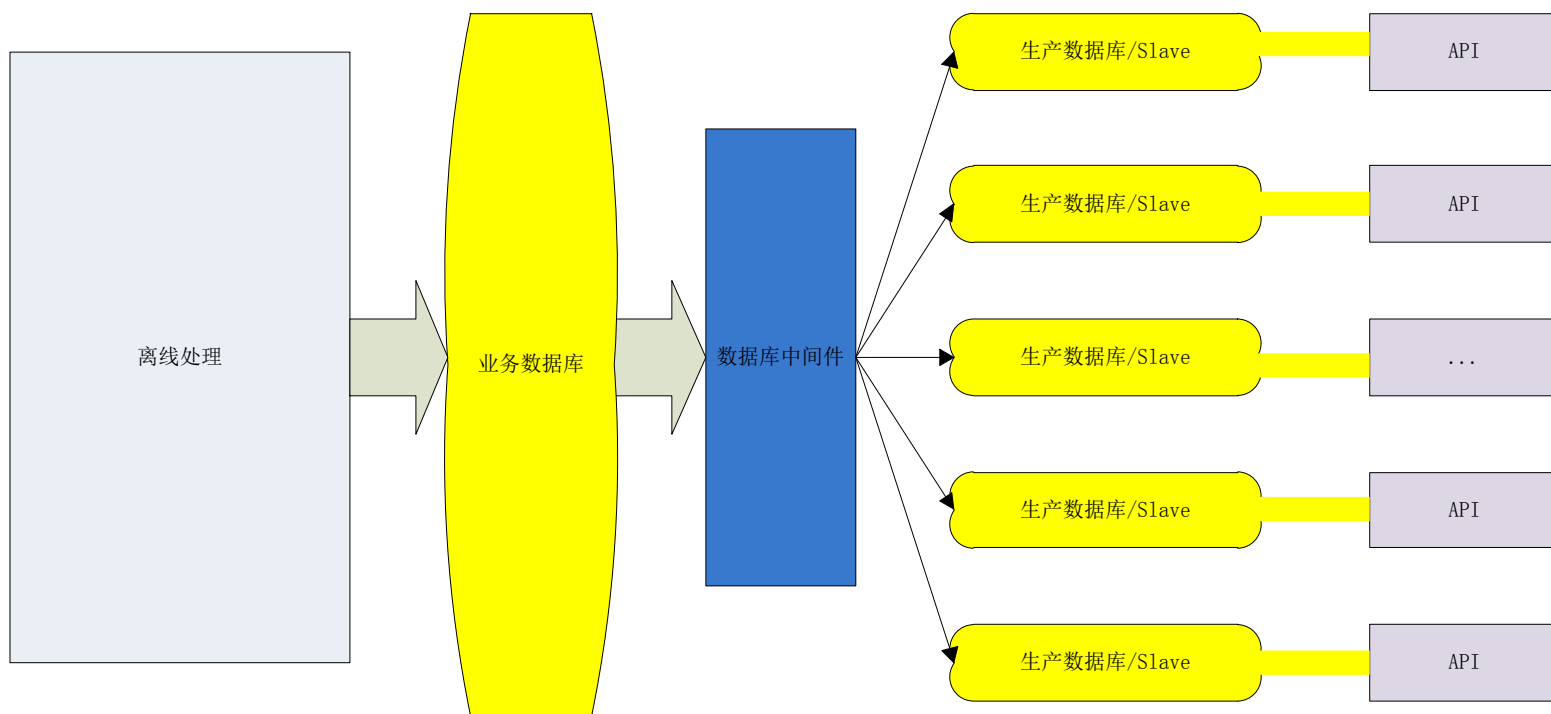




# 数据库中间件

- 统一数据库入口
- 数据多变体的支持
- 突破MongoDB 集群限制，部署跨机房数据中心

# Stage 2——在线业务的优化



# API应用架构

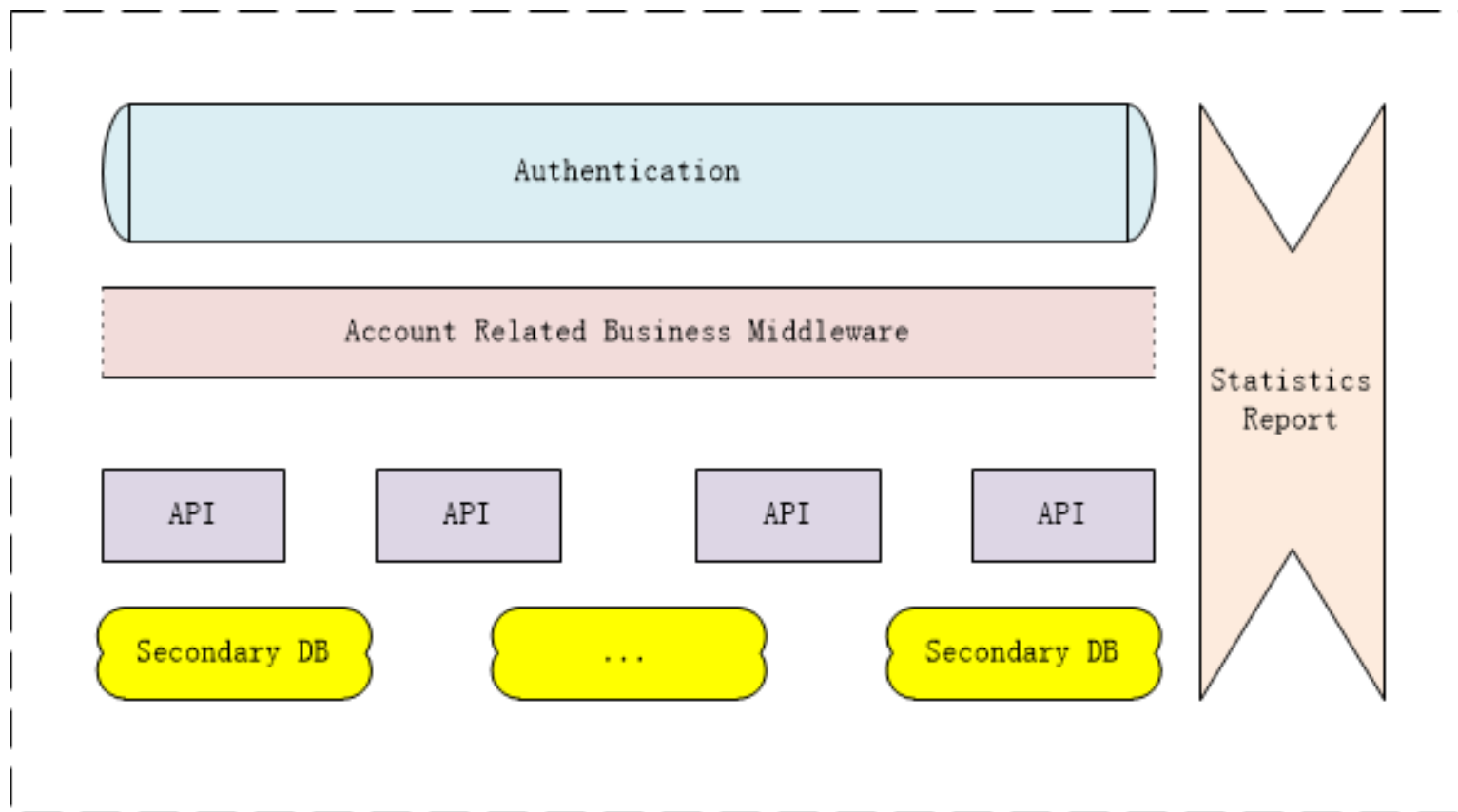
- “深井式”应用部署
  - 一个API
  - 一个DB
  - 一个CACHE
  
- 那么，问题就来了：
  - 10台API服务器 === 10台DB Slave?

# 此外

- + Account Resource Manage
- + **API Authentication**

- 那么，问题又来了：
  - API单机处理能力
    - 3k

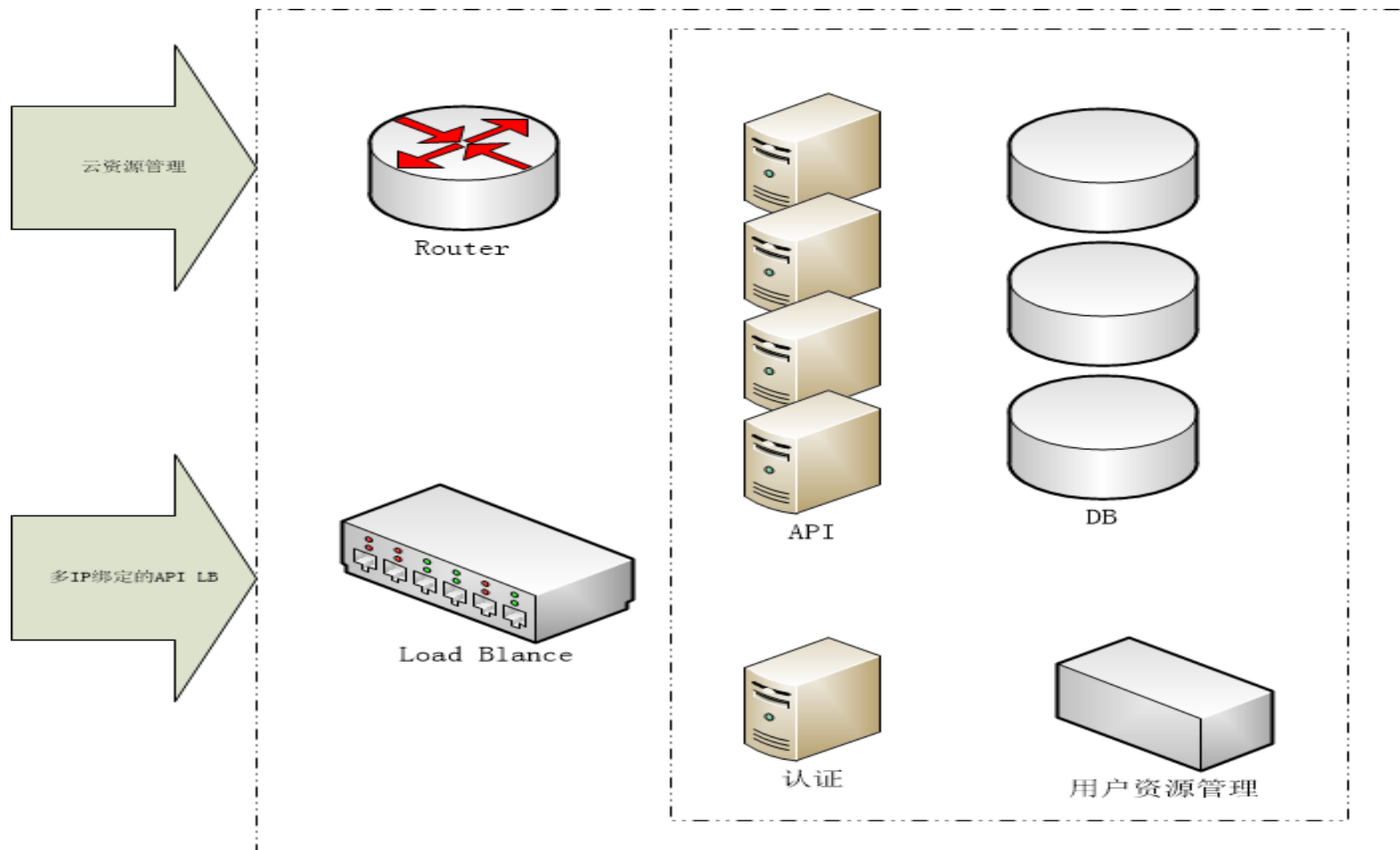
# 独立的服务单元



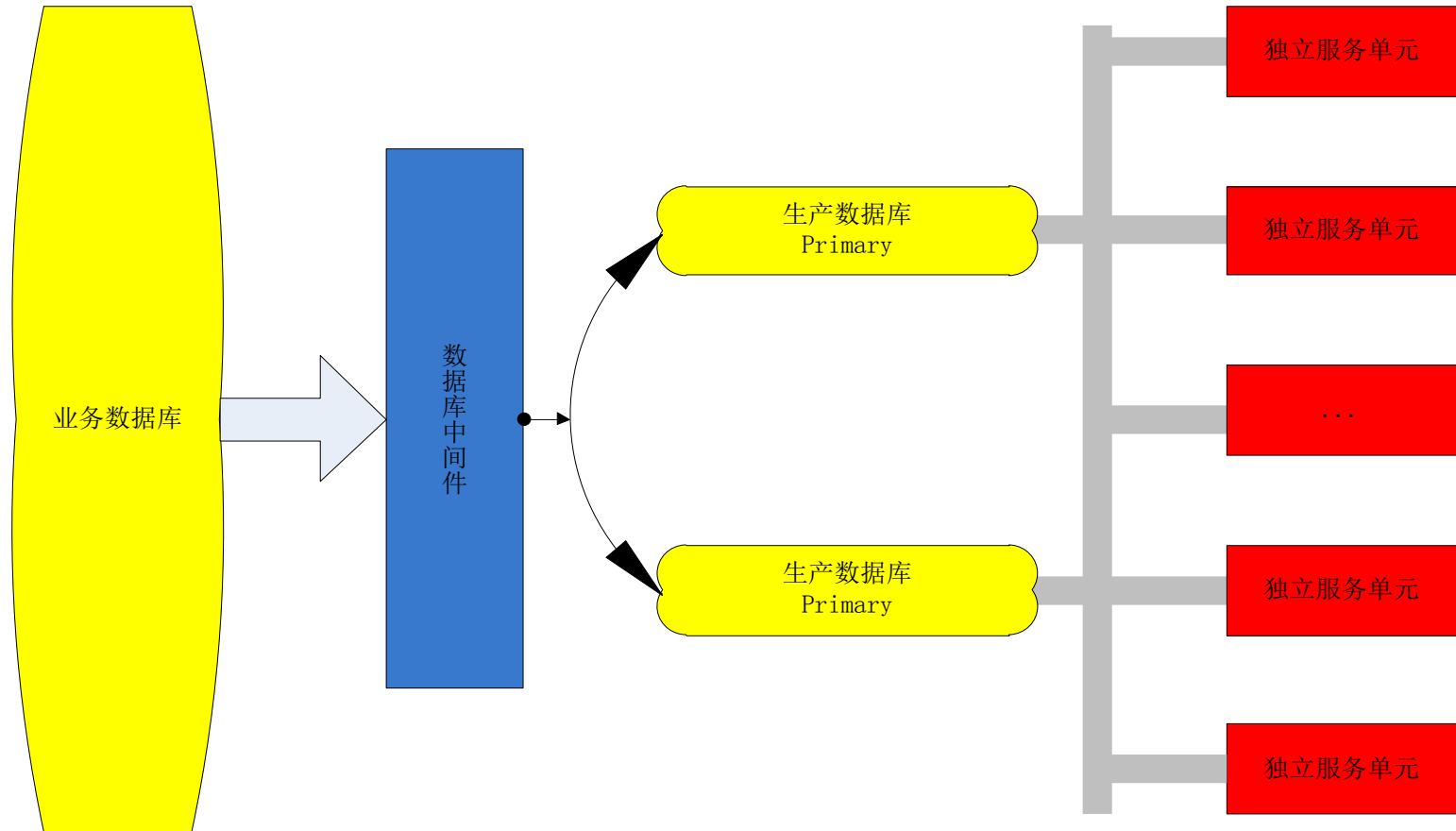
# 独立的服务单元

- 特点：
  - 独立
  - 支持动态水平扩展
  - 局域网访问
- 使业务的相关性更为清晰
- 提高单机服务能力
  - 加倍：3k → 6k
- 单元处理能力
  - 12k \* 10

# 独立服务单元的资源分布图



# 当前在线架构

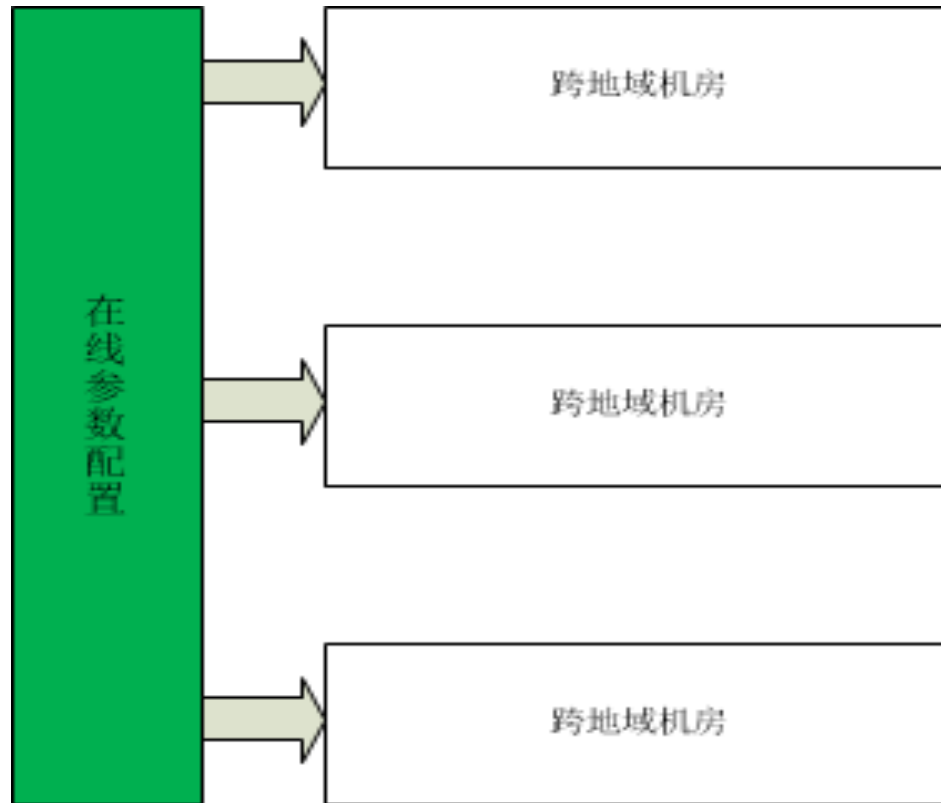




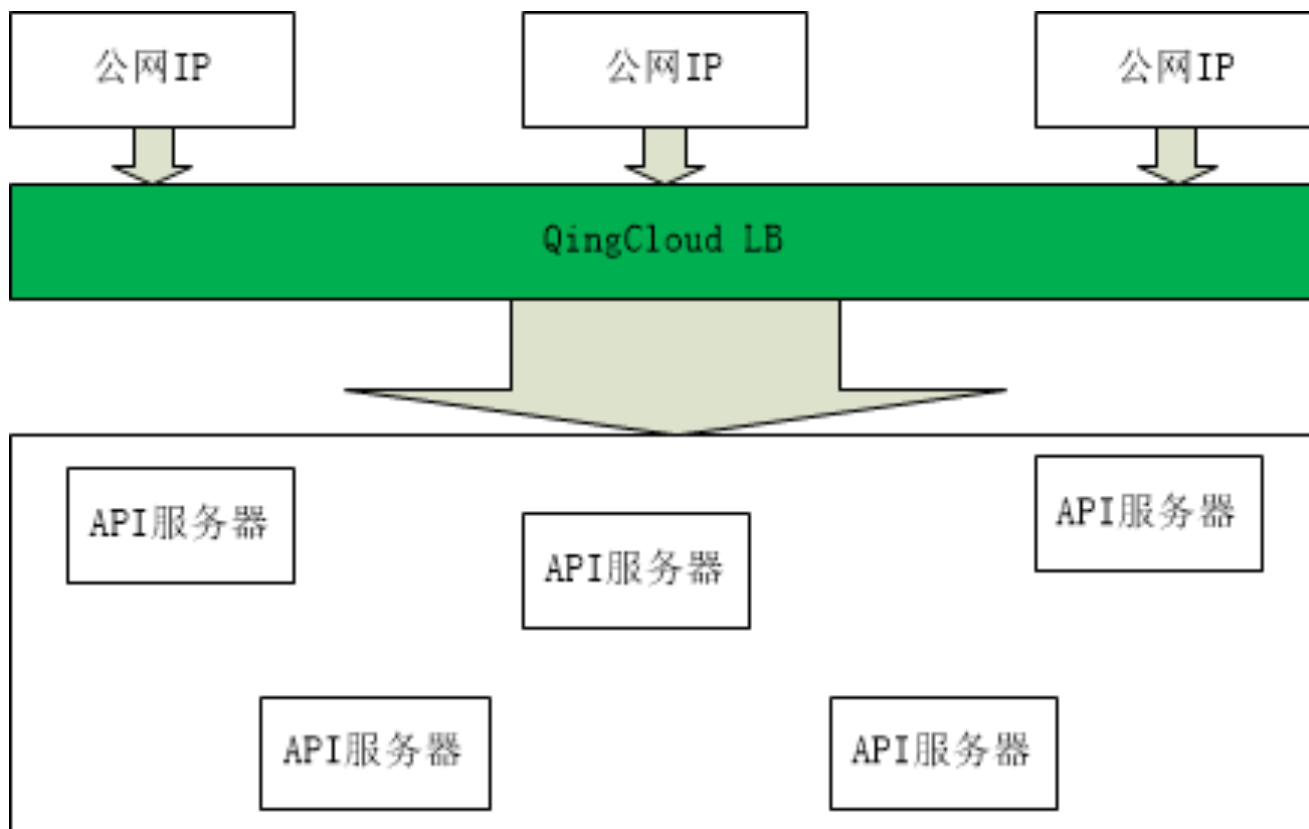
# 内部负载均衡系统

- 均衡
  - 在线参数均衡
  - API服务器集群内的均衡
  - 数据库负载均衡
- 过载
  - 过载保护
  - 恶意攻击

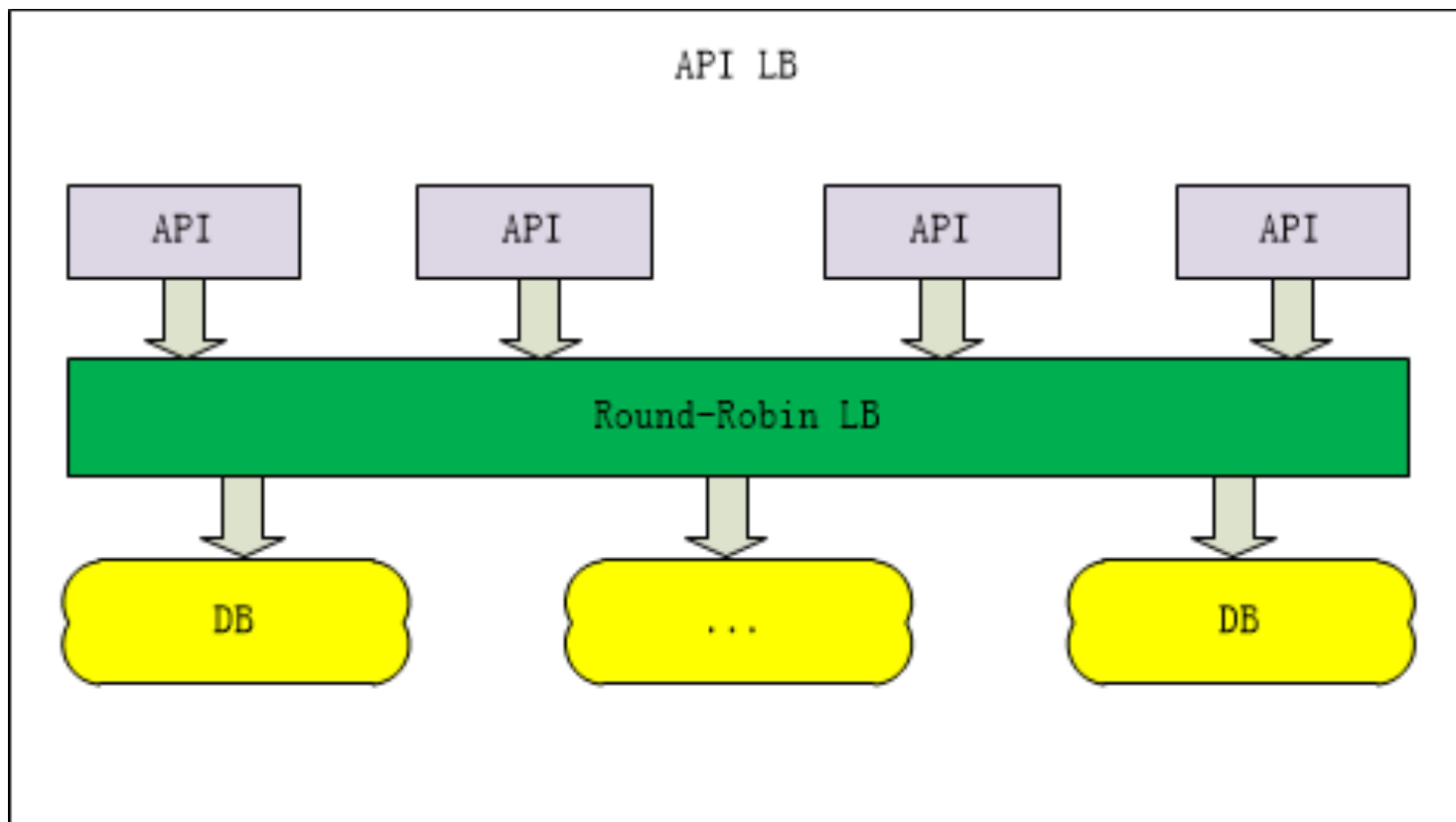
# 在线参数配置



# API服务器集群内的均衡



# 独立服务单元内的数据库均衡

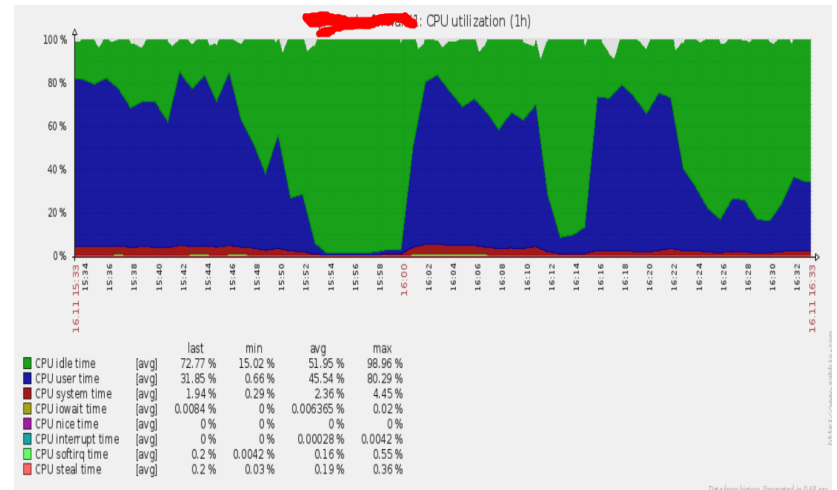
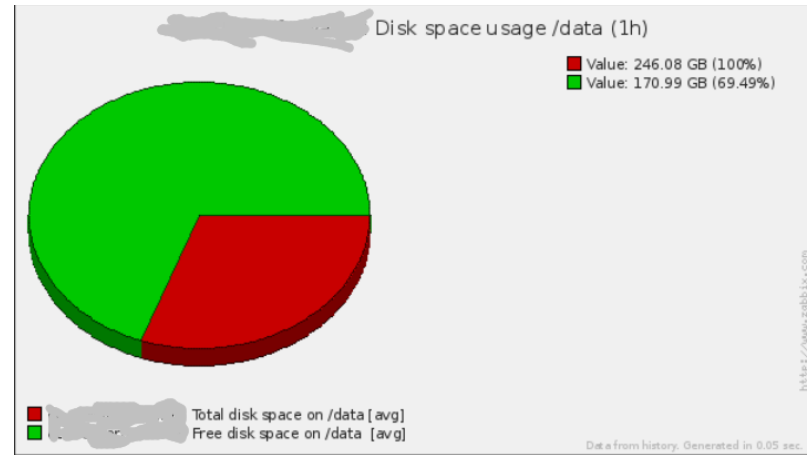


# 开发可能会结束，但运维永远不会

- Zabbix
- QingCloud 云资源监控项
- Public监控模块


# 资源使用率监控

- 磁盘
- CPU负载
- 内存
- 网卡流量



# QingCloud Console 监控




## 规则

  添加  应用修改  删除

| <input type="checkbox"/> | 监控项         | 条件 | 阈值         | 连续周期数 |
|--------------------------|-------------|----|------------|-------|
| <input type="checkbox"/> | 响应延迟时间      | >  | 500 毫秒     | 1     |
| <input type="checkbox"/> | 监听器 5xx 响应数 | >  | 25000 次/分钟 | 1     |

\* 提示：可通过在各个资源上点击“右键”来进行常用操作，以及“双击”来修改基本属性。

## 告警行为

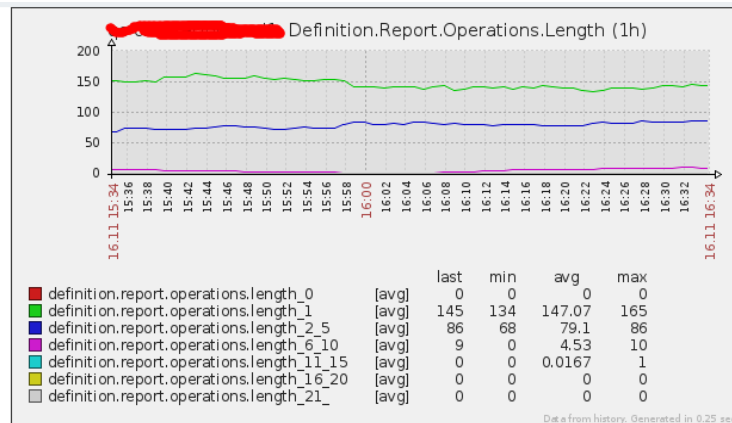
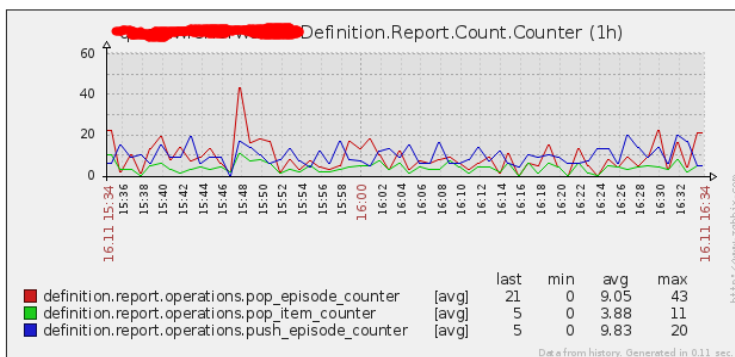
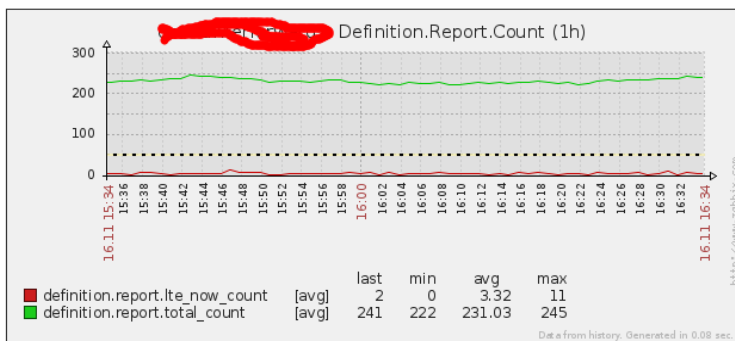
  添加  删除

| <input type="checkbox"/> | 触发状态 | 触发行为 | 通知列表  | 创建时间                |
|--------------------------|------|------|-------|---------------------|
| <input type="checkbox"/> | 告警   | 发送通知 | 邢稳    | 2014-09-26 17:09:18 |
| <input type="checkbox"/> | 告警   | 发送通知 | 陈剑豪   | 2014-08-25 12:13:00 |
| <input type="checkbox"/> | 告警   | 发送通知 | 李龙    | 2014-08-25 12:07:11 |
| <input type="checkbox"/> | 告警   | 发送通知 | costa | 2014-08-25 12:05:26 |

\* 提示：可通过在各个资源上点击“右键”来进行常用操作，以及“双击”来修改基本属性。

# 服务监控

- 自定义监控项





# 总结

- 硬件升级
- 清晰的架构划分
  - 功能拆分
    - 离线/在线
    - 业务逻辑
  - 逻辑拆分
    - 独立服务单元
    - 数据库
- 均衡/过载
- 缓存
- 异步消息队列
- 水平扩展
- 前后历经近一年左右，实现了完整的“半自动”运营的数据处理系统

# QingCloud 云资源的支持

- 灵活的SDN网络
- 快速响应的资源调度
- 精确到秒的计费
- 便捷的资源管理
- 稳定的资源

# QingCloud的工单系统

提交于: 2014-10-31 23:44:40

我们观察到10-31 22:51~22:54左右, lb-g5dhhx3s有异常。流量下降至0, 持续2分钟左右, 能查下吗?

附件



稍等, 我们看下。

Osier Yang

2014-10-31 23:45:29

辛苦了

2014-10-31 23:52:34

hi,

我们看了下监控没有发现有异常。不过这个问题确实不太正常, 我们会继续观察。想问下, 出问题的时候你们有做什么操作吗? 另外这个问题是LB自己恢复的, 还是你们手动更新了之后恢复的?

spencer

2014-11-01 01:41:46

# Q&A

- [cipher.chen2012@gmail.com](mailto:cipher.chen2012@gmail.com)
- <http://blog.cipherc.com>