

阿里业务网络服务质量分析

阿里巴巴阿里云事业部核心系统研发
朱友志/德泰
2014.08.13



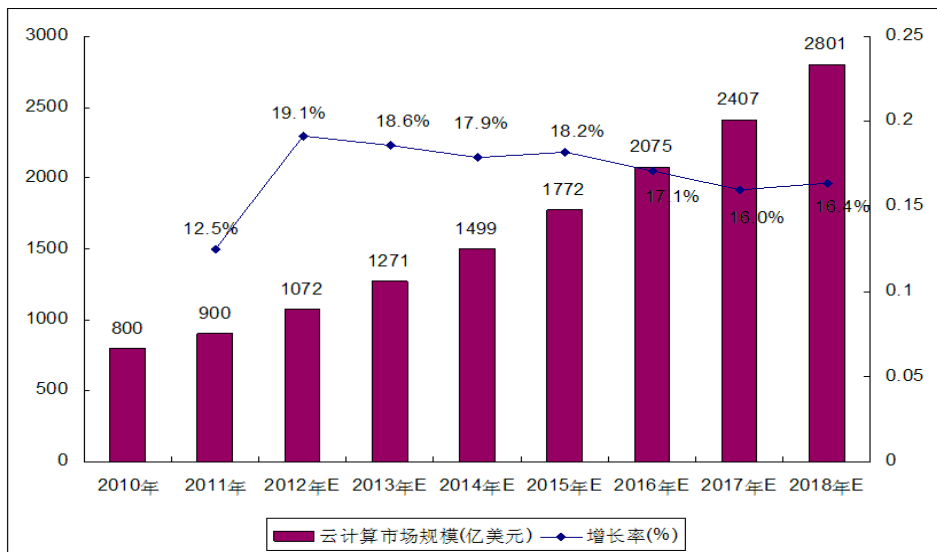
报告提纲

- 背景描述
- 应用案例
- 系统构成
- 设计要点
- 未来展望



背景描述

● 云端的精彩



2010-2018年全球云服务市场规模分析及预测
数据来源: 弘博报告网

2009-2015年中国网民及移动互联网用户规模图 (单位: 亿)
数据来源: 前瞻网

- 阿里云 当前: 电商云、金融云 计划: 集团云化(支付宝/淘宝)
- 无线 2013年双11, 手机支付单日成交额53.5亿, 手淘活跃用户1.27亿, 交易笔数占21%



背景描述

- 云端带来网络服务质量方面的挑战

- 云业务

在可用性、响应时间方面是否有提升？

- 无线互联网

客户端、接入网络制式、网络架构等更加复杂，如何提升无线应用网络服务质量？



背景描述 - 用户体验量化

图1 响应时间与用户体验(2-5-8s法则 (1*))

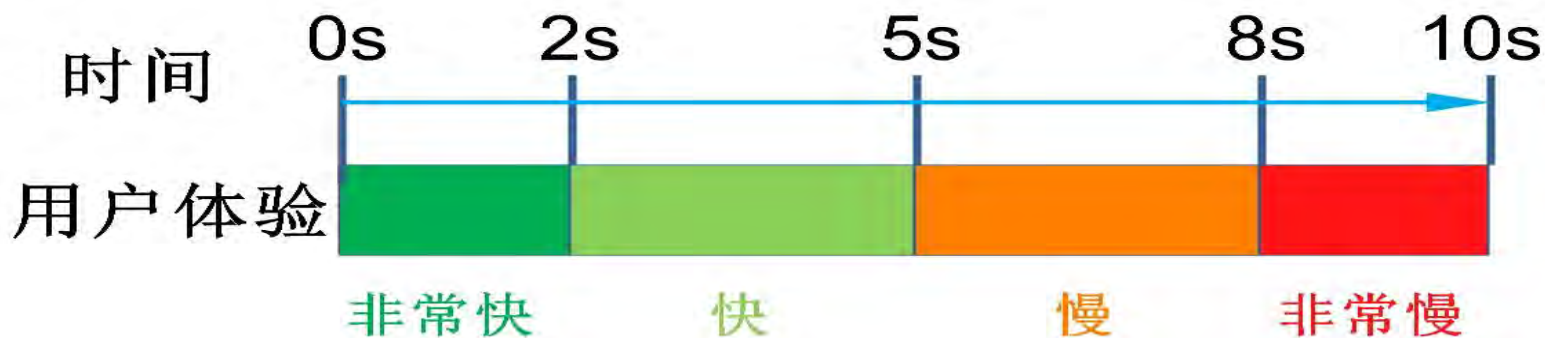


图2 用户体验的三个关键数字(2*)



背景描述 – 用户体验重要性

● 业务网络服务质量为什么重要？

- Amazon: 增加 100ms 延迟将导致收入下降1%
- Google: 400 ms 延迟将导致每用户搜索请求下降 0.59%
- Yahoo!: 400ms 延迟会导致流量下降 5-9%
- Bing: 2 秒的延迟将导致收入降低 4.3%/用户
- Shopzilla: 将页面载入时间从 7秒缩减到 2秒，转化率提升了 7-12%，页面请求增加 25%

-- IResearch 《中国网络零售年志》（3*）



应用案例

- 如何评估服务质量

- 可用性 四层/七层
- 时间相关 rtt , rt , 首字节时间 , 服务器处理时间
- 丢包率 建立连接阶段、关闭连接阶段、请求、响应

- 云业务服务质量挑战

- 秒级可用性监控与报警

- 无线业务服务质量挑战

- 支持tcp长连业务
- 分析维度丰富：网络接入制式、客户端版本、设备、os版本等



应用案例 - SLB云业务服务质量监控

- 监控内容

- 系统指标：CPU/MEM/NET
- 网络可用性
- 时间
- 丢包率
- 日志

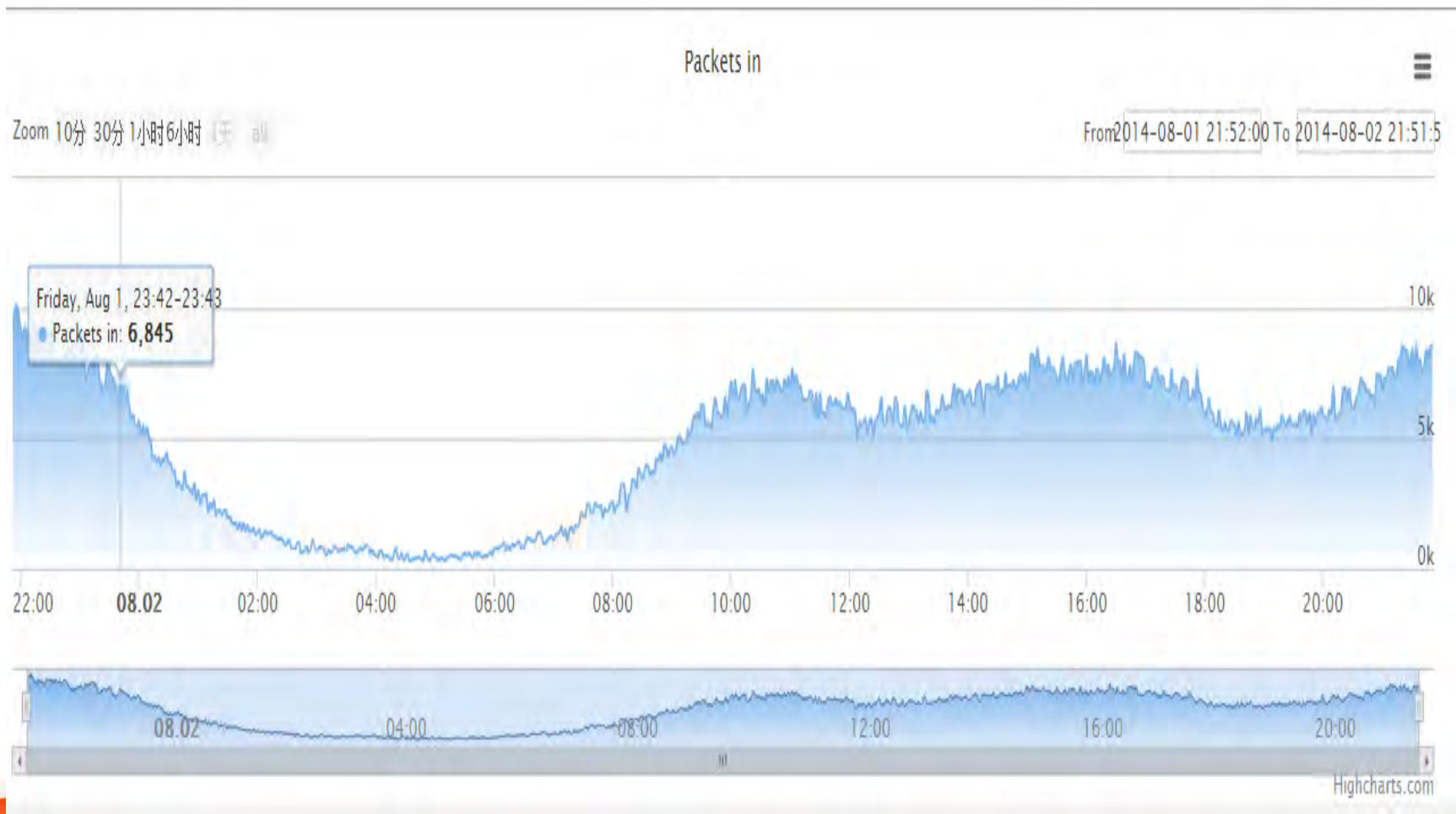
- 监控组件

- lvs/keepalived
- Tengine
- 管理系统



应用案例 - 云业务per-IP秒级流量与可用性监控

解决分钟级监控无法探测“闪断”的问题



应用案例 - 云业务CDN服务质量分析用途

- 常态化监控
- 优化
 - 调度策略优化
 - 协议栈参数调整(init_cwnd)
 - 连接复用率调整
 - 拥塞控制算法改进



应用案例 - 云业务CDN http请求rt、连接复用率分析

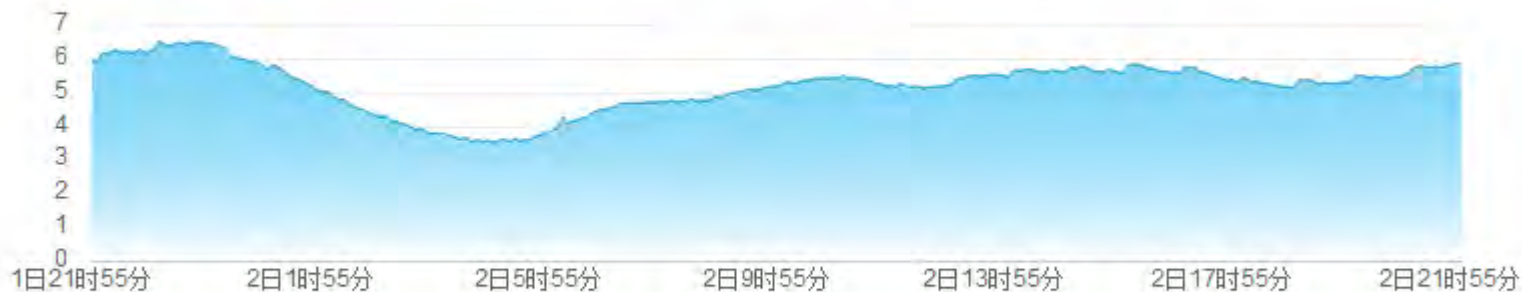
服务器端网络服务质量测量，为系统优化和精细化运维提供基础

响应时间



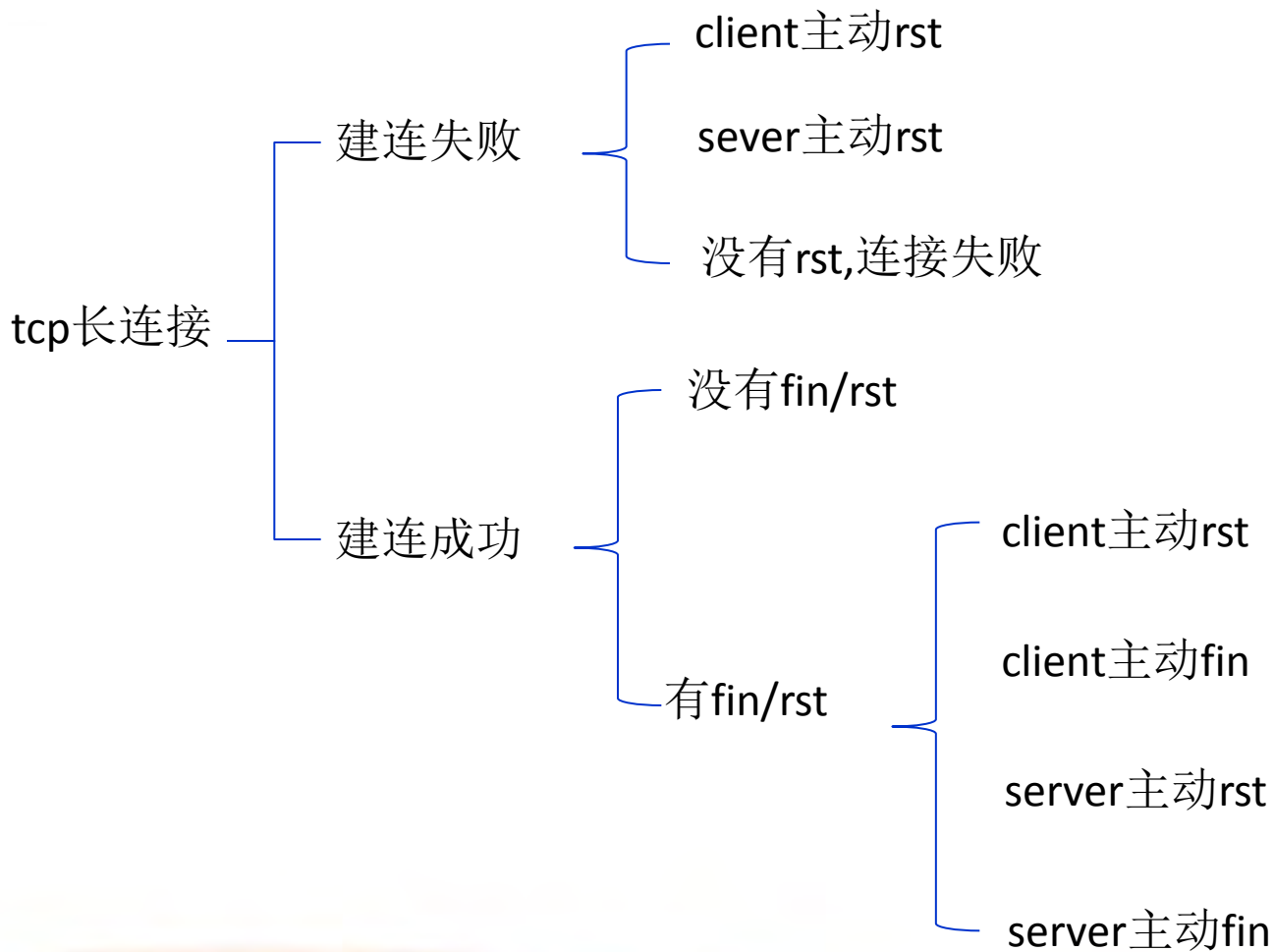
—●— 响应时间

TCP连接复用率



—●— HTTP连接复用率

应用案例 - 无线长连服务断连原因分析



应用案例 - 无线访问特征与优化

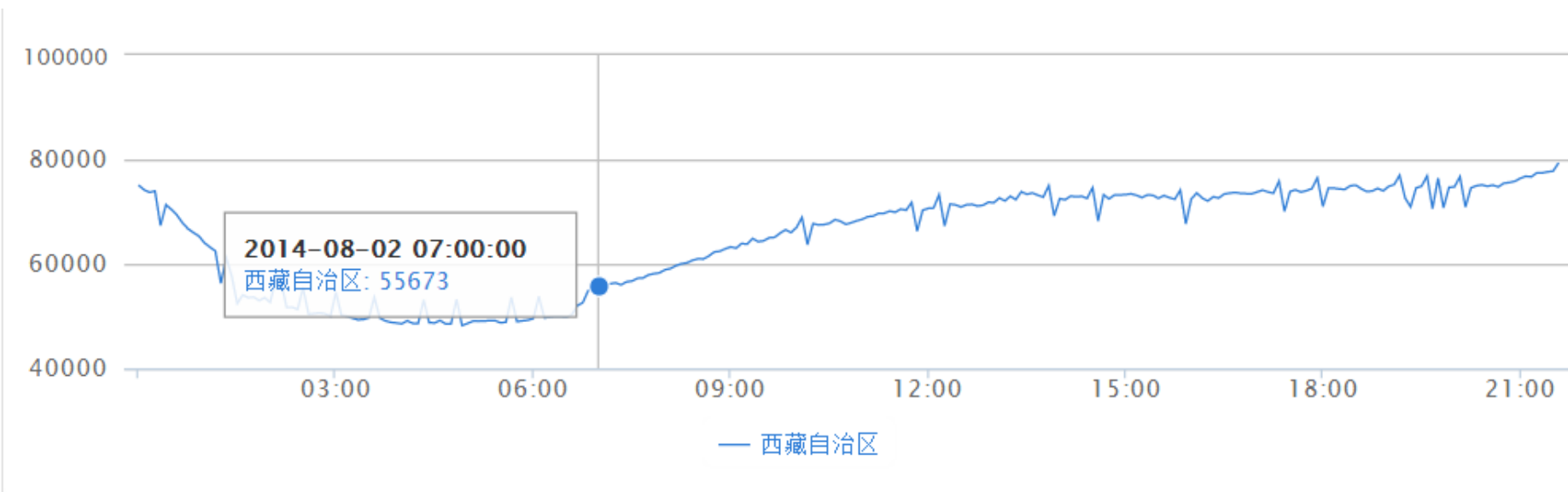
- MSS分布与IP分片统计
 - MSS协商
- 建连丢包影响
 - 服务器端回复的synack丢失率高
- 协议栈参数优化
 - (初始)连接超时
 - 重试次数 `/proc/sys/net/ipv4/tcp_retries2`



应用案例 - 无线长连服务质量分析

● 无线架构、软件优化基础 -- 服务质量细粒度分析

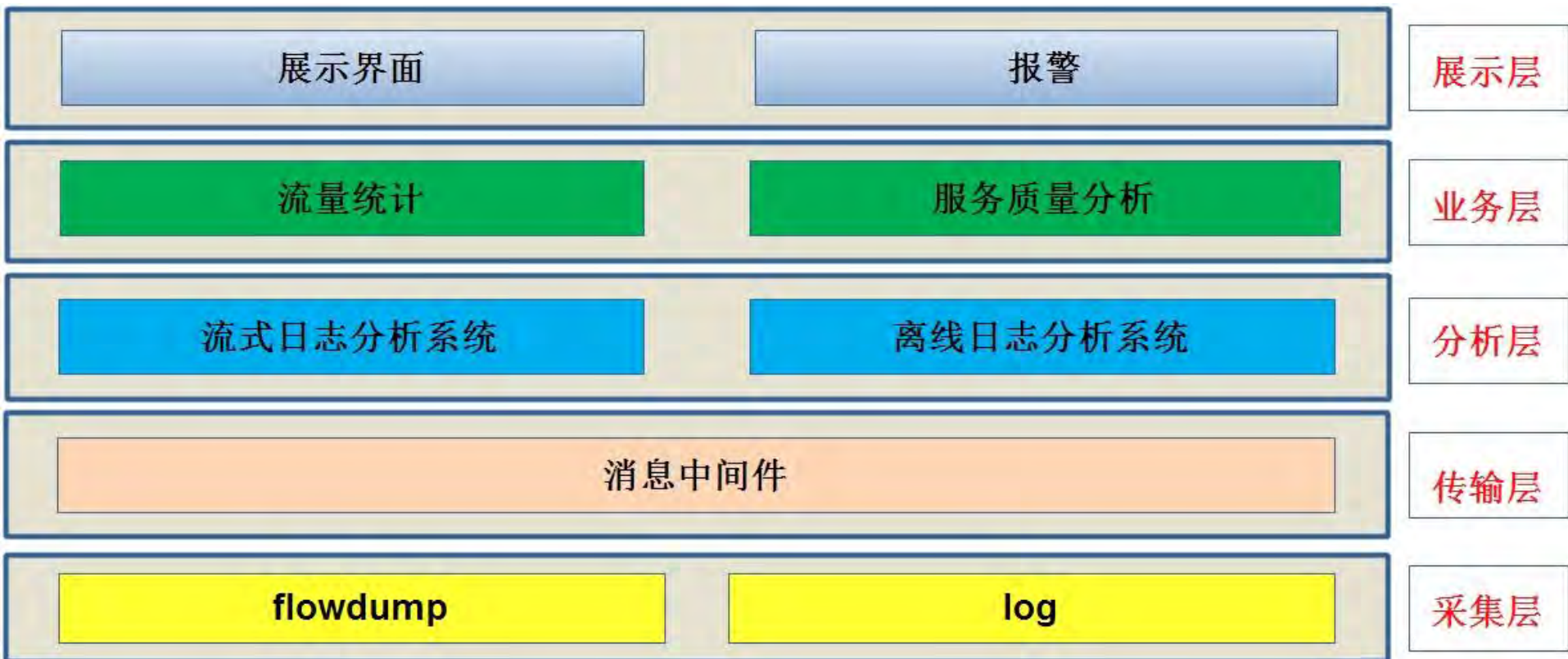
- 分析维度: 域名、地域、ISP、网络接入制式、客户端版本



西藏自治区长连数变化曲线



系统构成



设计要点 - 采集方式选型

| | 客户端采集方案 | 日志采集方式 | 页面埋点方式 | Flowdump服务器端抓包采集方式 |
|-------------|---------|--------|--------|--------------------|
| 是否能采集页面性能数据 | 是 | 否 | 是 | 否 |
| 数据准确性 | 高 | 低 | 高 | 中 |
| 是否真实用户请求 | 否 | 是 | 否 | 是 |
| 是否用户侧时间 | 是 | 否 | 是 | 是 |
| 能否采集丢包率 | 是 | 否 | 否 | 是 |
| 采样率 | 低 | 高 | 低 | 高 |
| 成本 | 高 | 低 | 低 | 低 |
| 可运维性 | 低 | 高 | 中 | 高 |



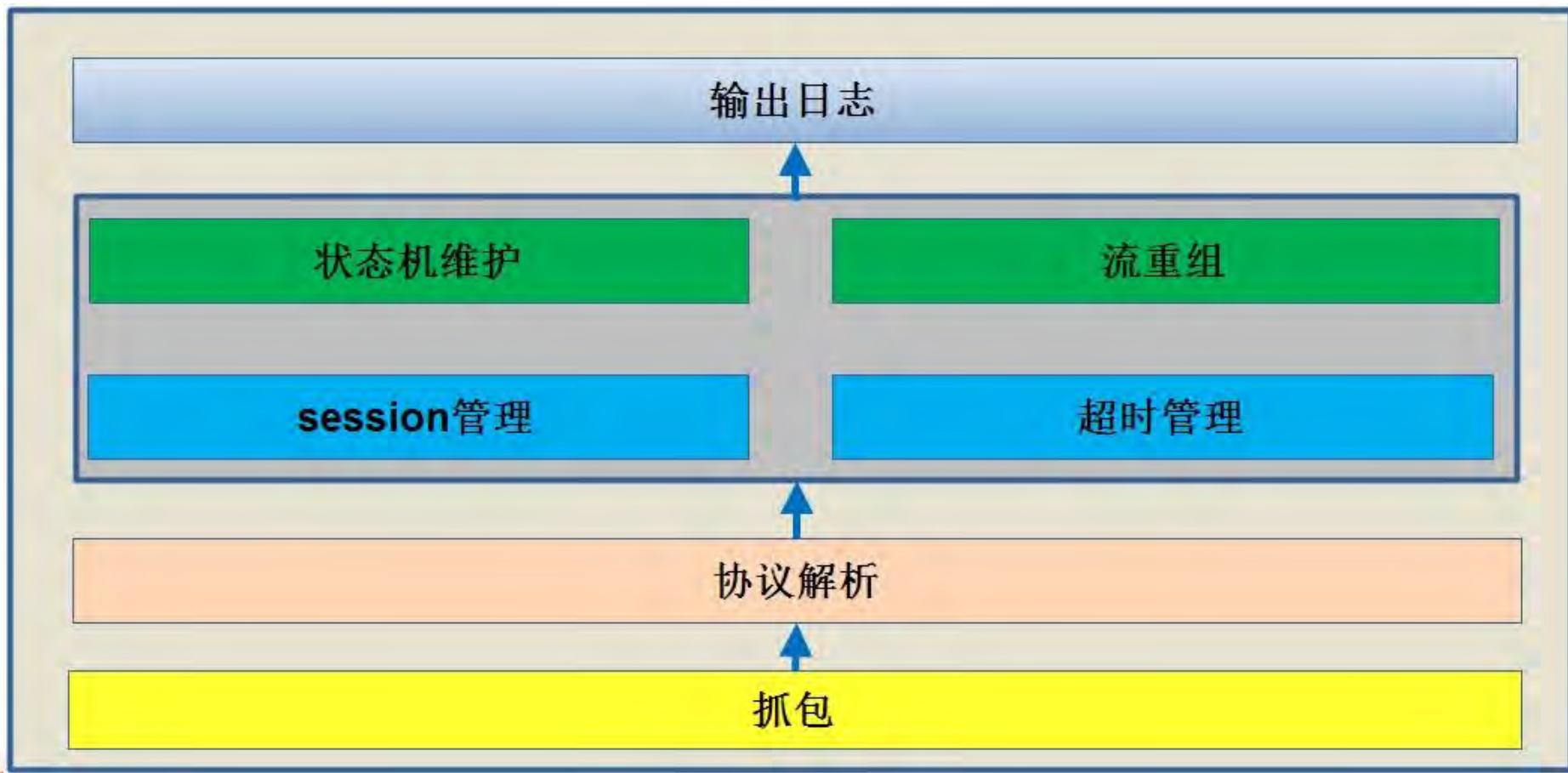
设计要点 - 抓包库选型

| 采集软件版本 | 性能 | 通用性 | 备注 |
|------------|----|-----|-----------|
| LIBPCAP | 低 | 高 | |
| 内核模块方式 | 高 | 高 | 七层信息分析困难 |
| PF_RING(*) | 中 | 中 | |
| DPDK | 高 | 低 | 只能独占服务器使用 |



设计要点 - 采集软件设计

Flowdump整体流程



设计要点 - 日志格式设计

| 序号 | 名称 | 样例 | 备注 |
|----|--------------------|----------------------------|-----------------|
| 1 | type | HTTP | 日志类型 |
| 2 | date | 08-05-01 | 日期 |
| 3 | time | 16:59:46 | 时间 |
| 4 | sip.port | 220.181.37.42.54520 | 源ip和port |
| 5 | dip.port | 220.181.37.156.80 | 目的ip和port |
| 6 | packets_in | 10 | 收到的数据包个数（被监测ip） |
| 7 | packets_out | 10 | 发送的数据包个数（被监测ip） |
| 8 | bytes_in | 640 | 收到的数据字节数（被监测ip） |
| 9 | bytes_out | 1300 | 发送的数据字节数（被检测ip） |



设计要点 - 日志格式设计

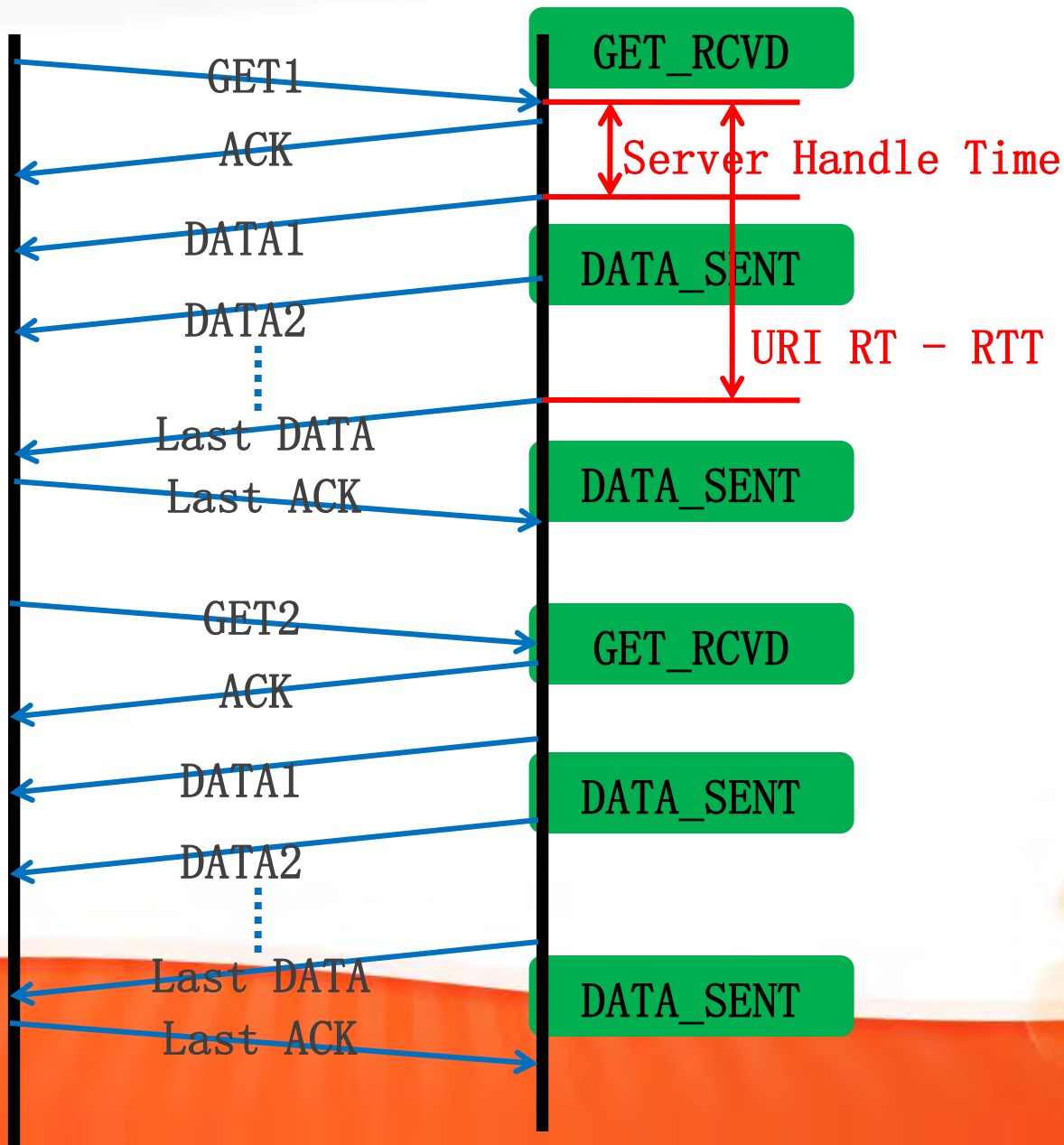
| | | | |
|----|-------------------------|---------------------|-----------------------------|
| 11 | http net body bytes | 640 | http 七层响应体字节数 |
| 12 | http offset | 1 | 同一个 tcp 流中第几个 http |
| 13 | http request | GET/POST | HTTP 请求类型 |
| 14 | Host | img01.taobaocdn.com | http 请求的 host 信息 |
| 15 | http URI | /1.jpg | URI 信息 |
| 16 | http response | 200/403/..... | HTTP 应答 code (当没有响应时返回 N/A) |
| 17 | http total data packets | 506 | HTTP 数据包的总数 (含 GET 请求) |
| 18 | http loss data packets | 4 | HTTP 请求和响应丢失的数据包数(含 GET 请求) |
| 19 | http out loss packets | 2 | HTTP 响应的丢包数 |
| 20 | http init wnd | 3 | HTTP 响应在收到第一个 ack 前发送的包数 |
| 21 | server handle time | 4 | 服务器收到 GET 请求到开始发送数据时间(ms) |
| 22 | rt | 5 | HTTP 请求的响应时间(ms) |



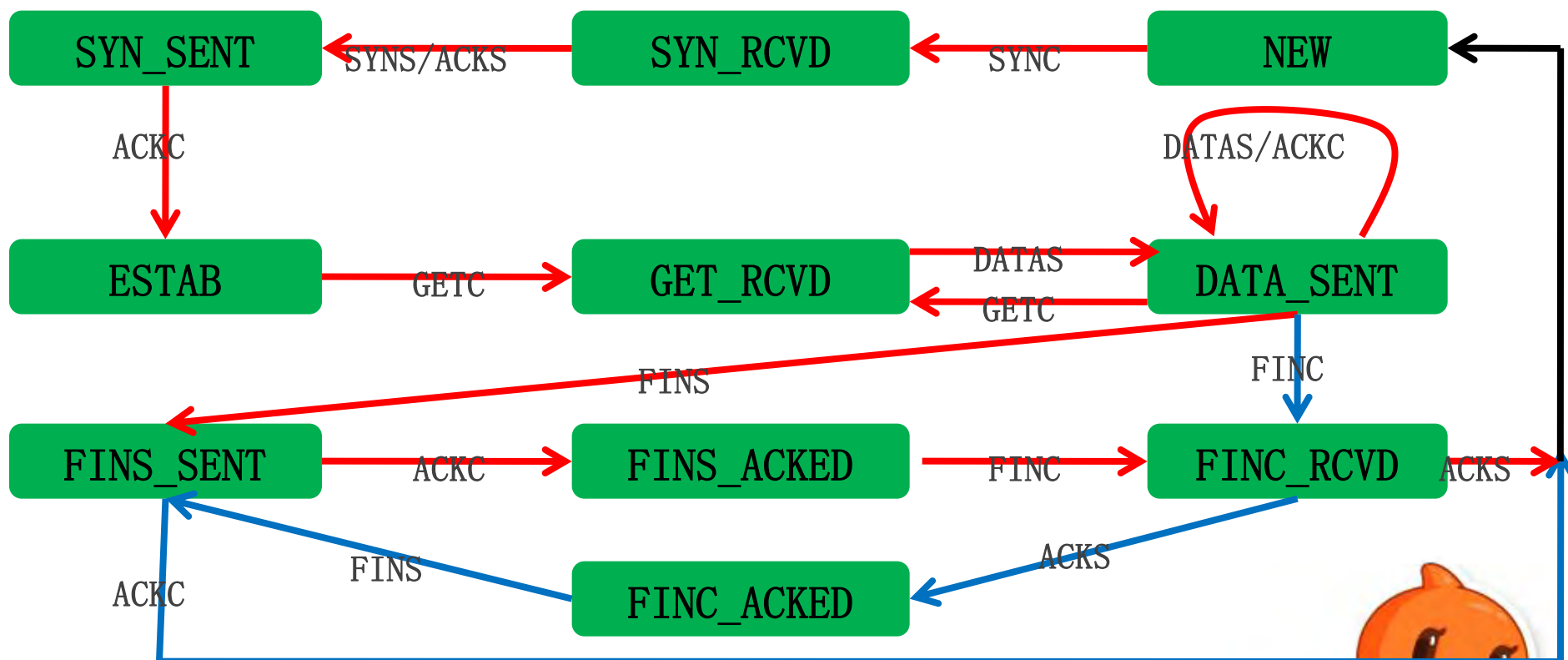
设计要点 - 服务质量参数计算原理

Client

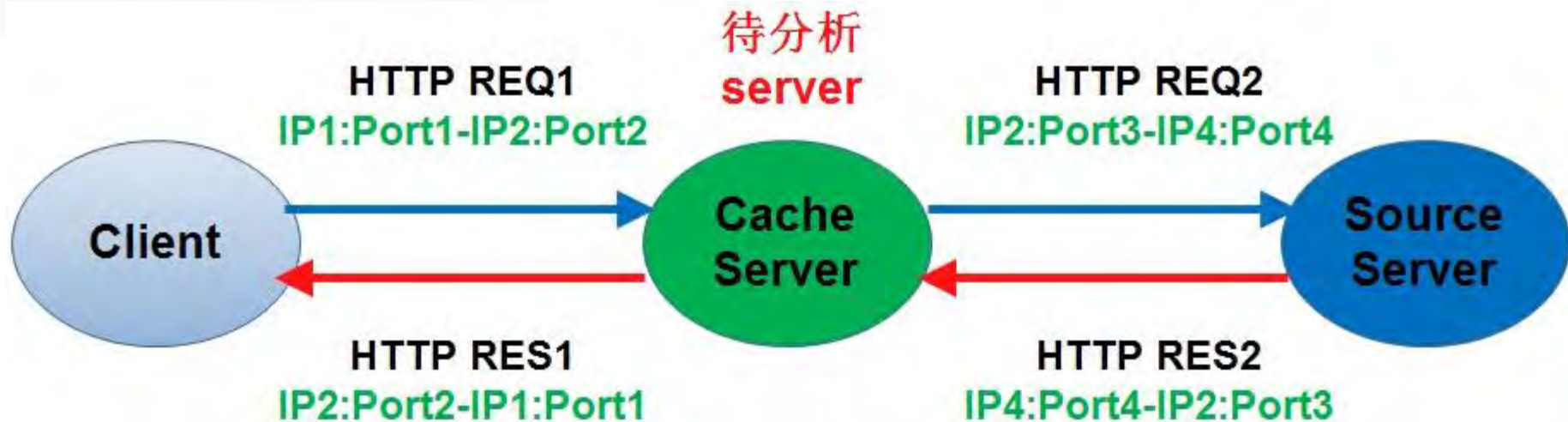
Server



设计要点 - 服务质量参数计算实现



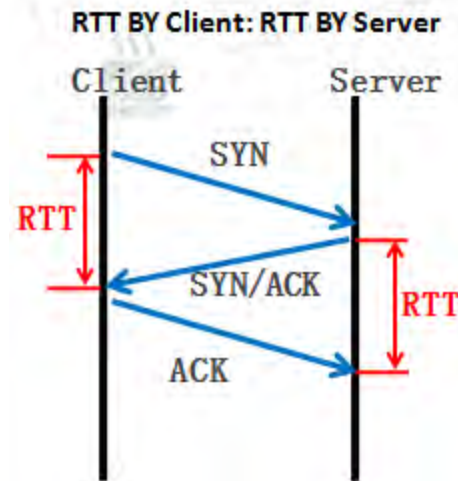
设计要点 - 服务器角色、包方向判断



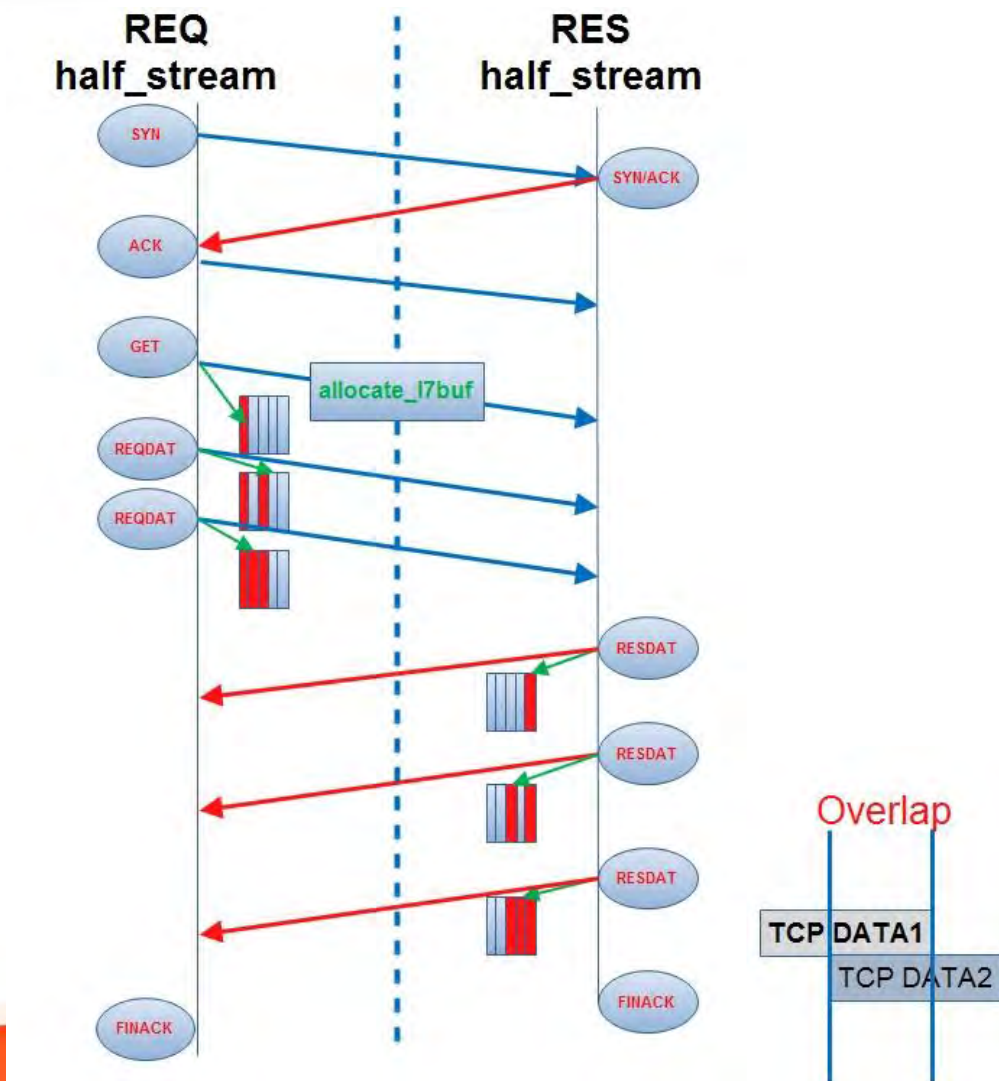
Q1: 请求和响应的四元组不同，如何归一化？

Q2: 待分析服务器角色不同，对服务质量参数有何影响？

Q3: 对于待分析机器，如何判断其在通信中的角色？



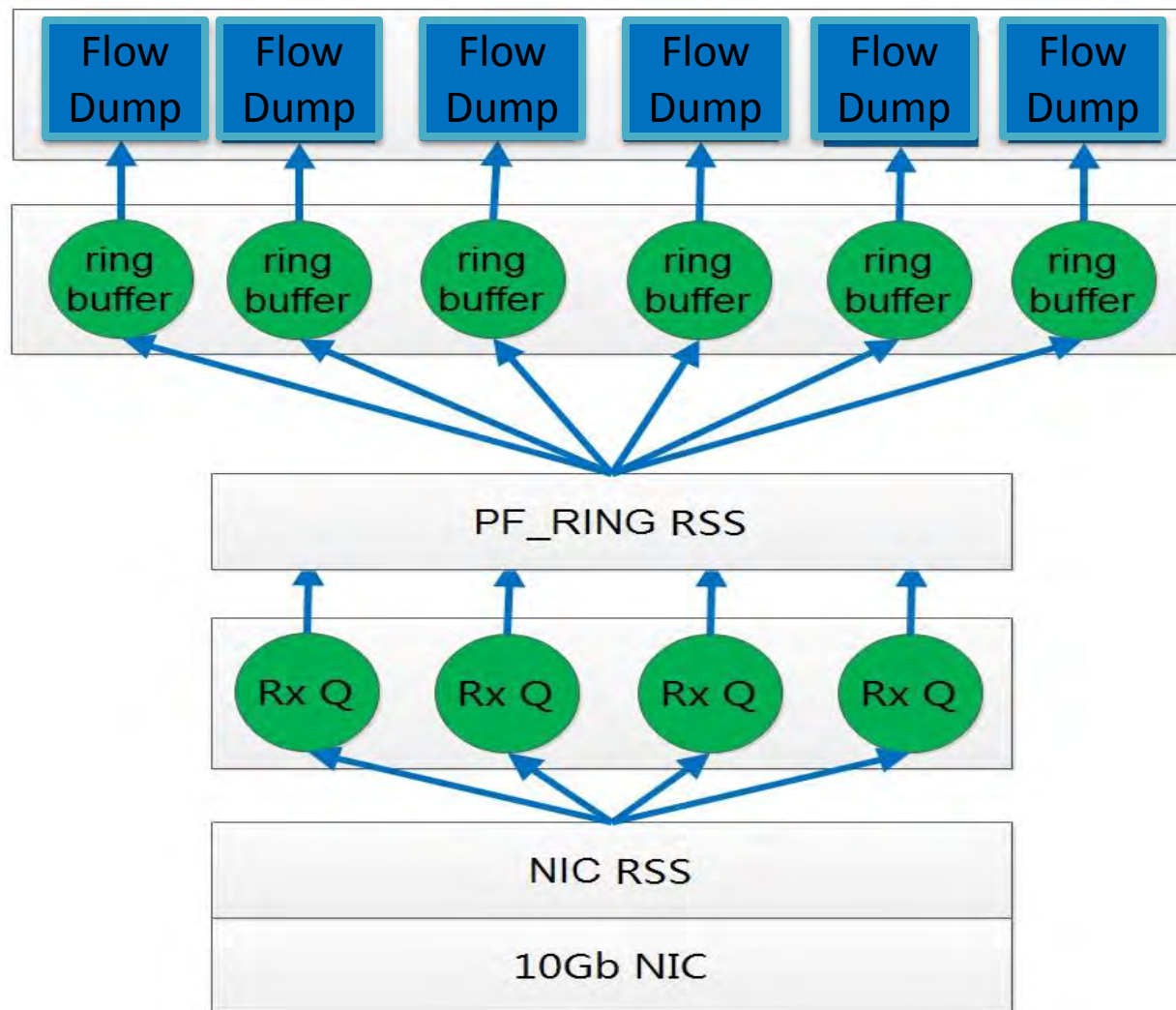
设计要点 - 功能 - tcp流重组



- Q1: 攻击情况如何处理?
- Q2: 报重传、交叉重叠等异常如何处理?



设计要点 - 性能 - 架构优化

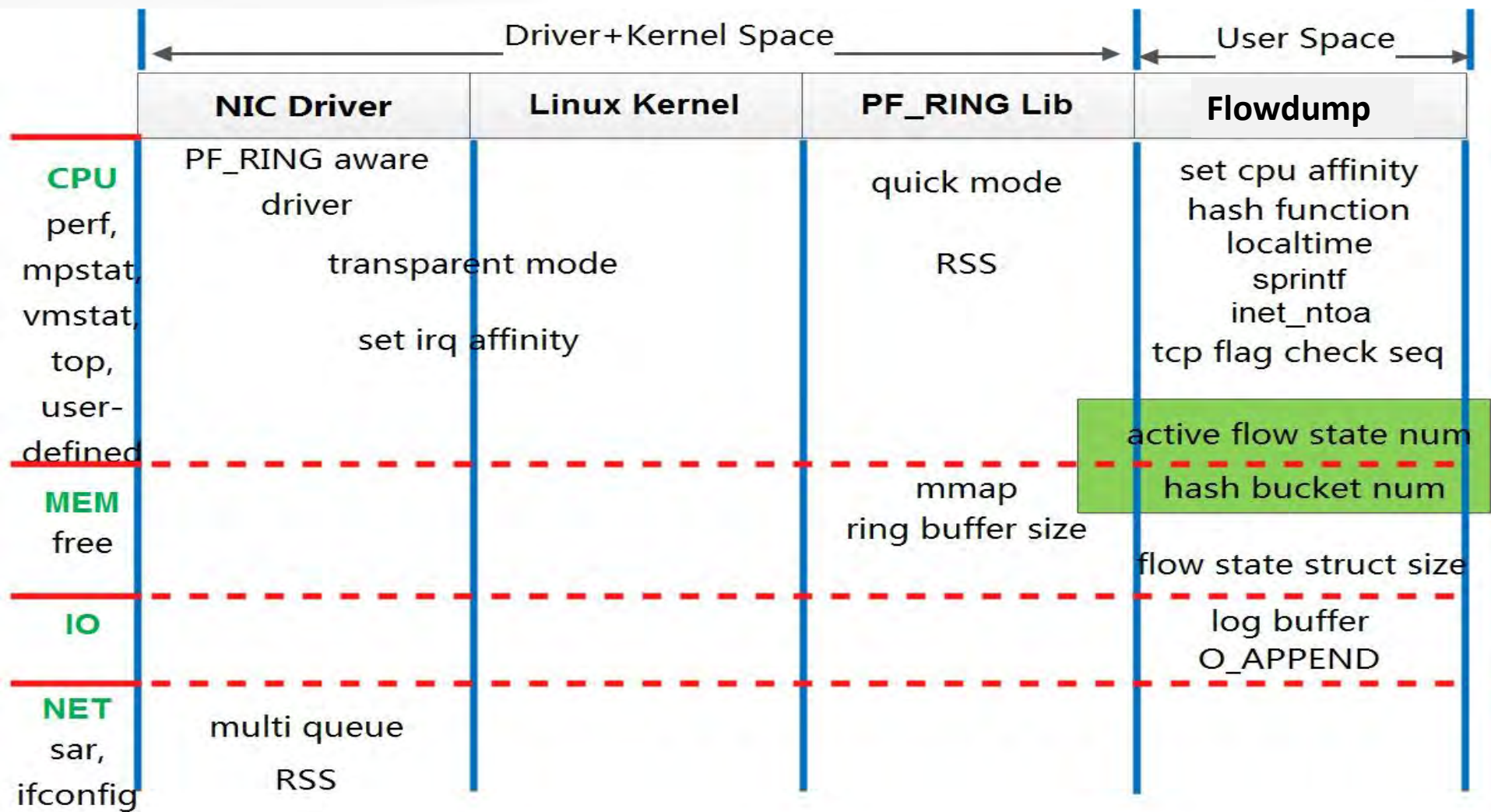


进程亲和性绑定

IRQ亲和性绑定



设计要点 - 性能 - 优化策略



未来展望

- 客户端、服务器端统一的服务质量分析平台
- 无线新协议的网络服务质量分析支持
- 性能提升



参考文献

- 1、 <http://www.websiteoptimization.com/speed/1/1-3.html>
- 2、 http://en.wikipedia.org/wiki/Ideal_Web_response_time#Speed_issues
- 3、 中国网络零售年志 2010~2011



PLA小组：

吴佳明/普空 pukong.wjm@alibaba-inc.com 新浪微博：吴佳明_普空

朱友志/德泰 detai@alibaba-inc.com 新浪微博：淘德泰

熊桥/诗然 xiongqiao.xq@alibaba-inc.com 新浪微博：天天2872390505

王中波/金庸 zhongbo.wzb@alibaba-inc.com 新浪微博：wzb56

Q&A

