



GIAC

全球互联网架构大会

GLOBAL INTERNET ARCHITECTURE CONFERENCE

爱奇艺高并发网关的实践及优化之路

吴杰珂 爱奇艺 资深科学家



GIAC

全球互联网架构大会

GLOBAL INTERNET ARCHITECTURE CONFERENCE



关注msup
公众号获得
更多案例实践

GIAC 是中国互联网技术领域行业盛事，组委会从互联网架构最热门领域甄选前沿的有典型代表的技术创新及研发实践的架构案例，分享他们在本年度最值得总结、盘点的实践启示。

2018年11月 | 上海国际会议中心



高可用架构
改变互联网
的构建方式



Agenda

- 高并发&高可用遇到的问题
- 如何选型
- 4层优化架构
- 7层优化架构
- SSL优化
- 网络传输优化
- 接入层调度优化
- 后续展望



高并发&高可用遇到的问题

- 业务扩展、机器扩容需求大
- 突发流量，冗余度难以估计
- 百万并发，高流量支持
- 确保IDC的出入口没有单点
- 高并发下TLS/SSL的性能、延时问题
- 集群高可用
- DNS劫持问题



如何选型

• 4层负载均衡

- 为何传统LVS速度不够快
 - Kernel是高性能网络瓶颈
 - 资源共享：竞争与锁
 - 上下文切换开销
 - 完善却复杂的Kernel协议栈
- 使用Kernel bypass技术加速

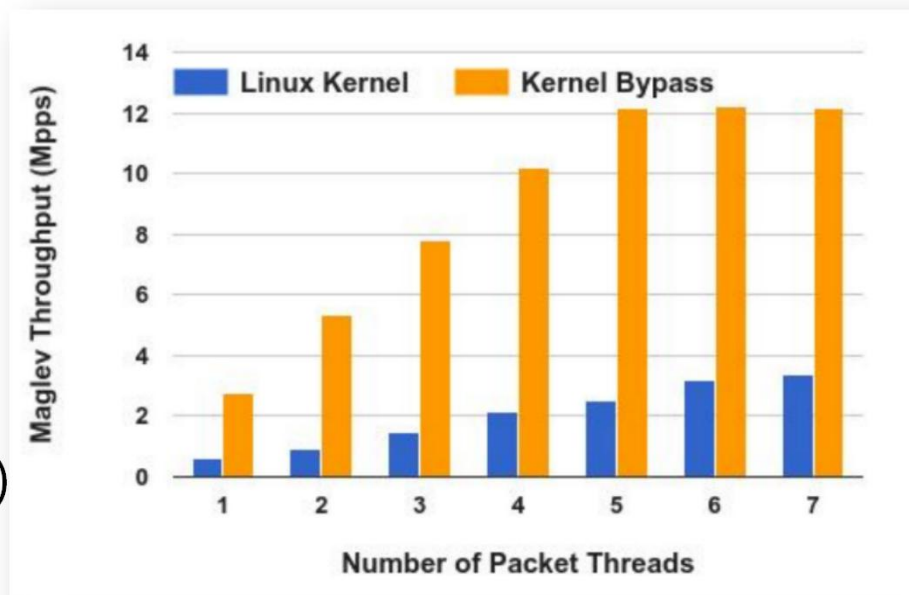
• 7层负载均衡

- 使用硬件加速卡（qat+nginx的稳定性，成本，可运维问题）
- 开源选型：ODP/OFP，mTCP，f-stack
- 分离出计算集群
- Session共享技术



4层优化架构

- Kernel By-pass技术 (DPDK)
- 优势
 - 低延时
 - 高性能
 - 开发、调试方便
- 劣势
 - 自研代码及周期
 - 功能特性 (缺少协议栈)
 - 配套配置工具 (ifconfig, ip, ...)

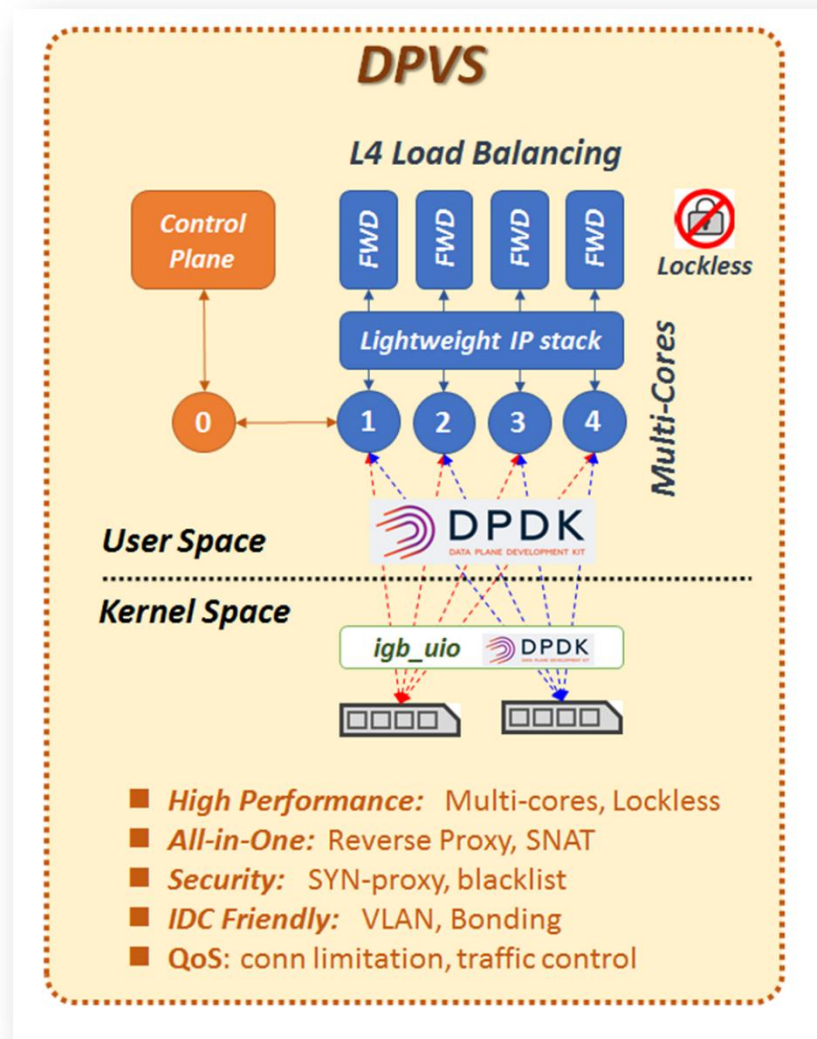


[图片出处](#) Google Maglev



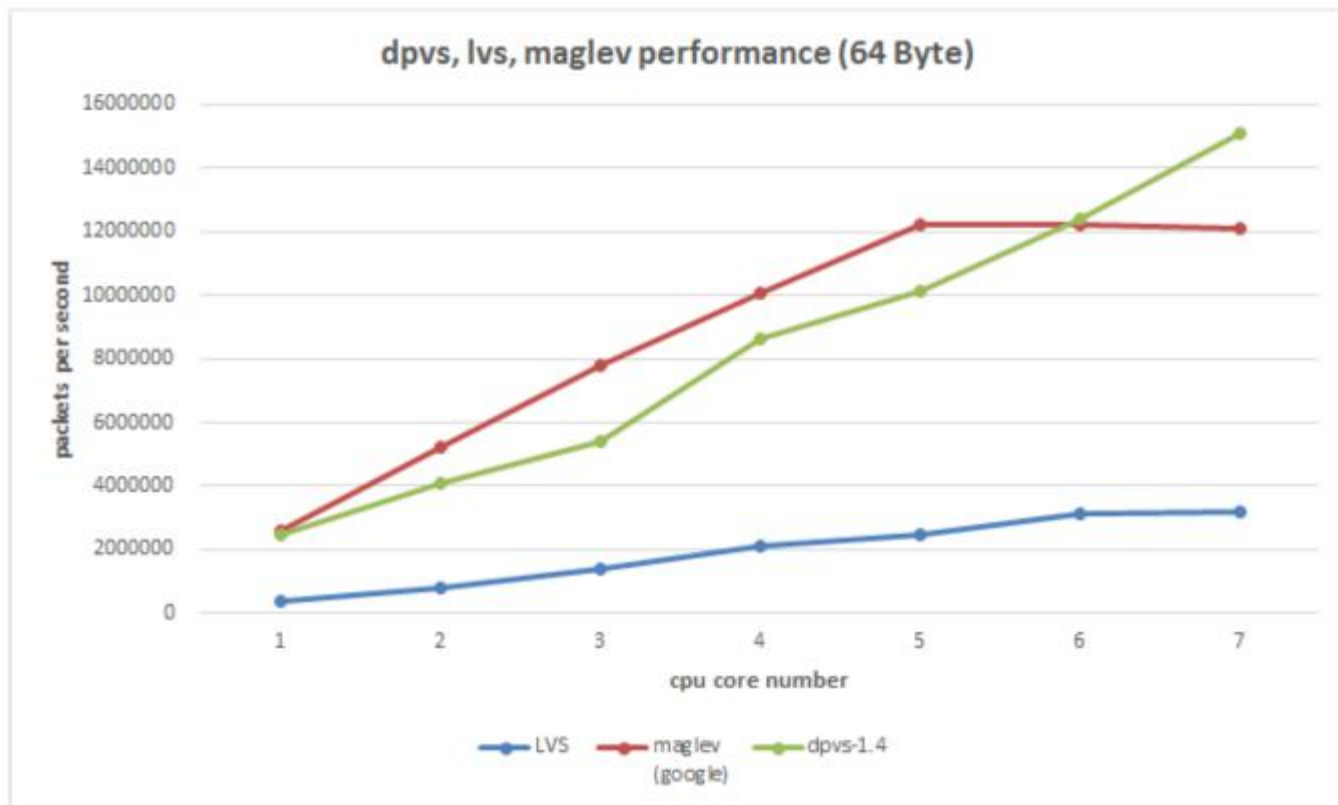
4层优化架构 (2nd)

- 用户态实现
- Master/Worker模型
- 避免上下文切换
- 关键数据无锁化
- 跨CPU无锁通信
- 支持FNAT/SNAT多种模式
- 支持VLAN/Bond/Tunnel等
- 安全：synproxy, 黑名单, 限流



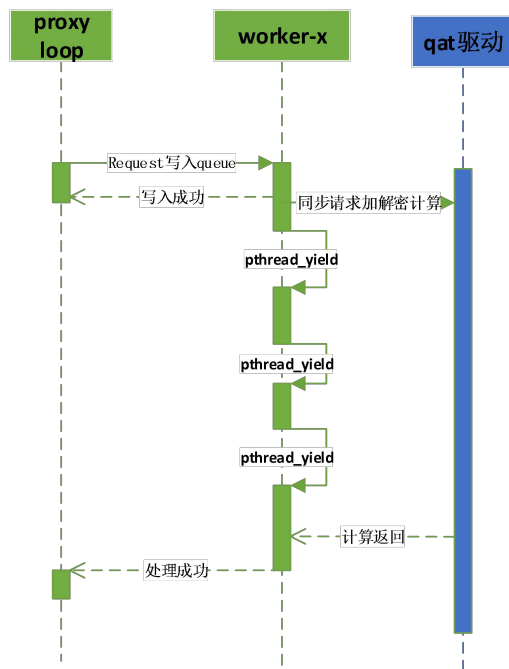
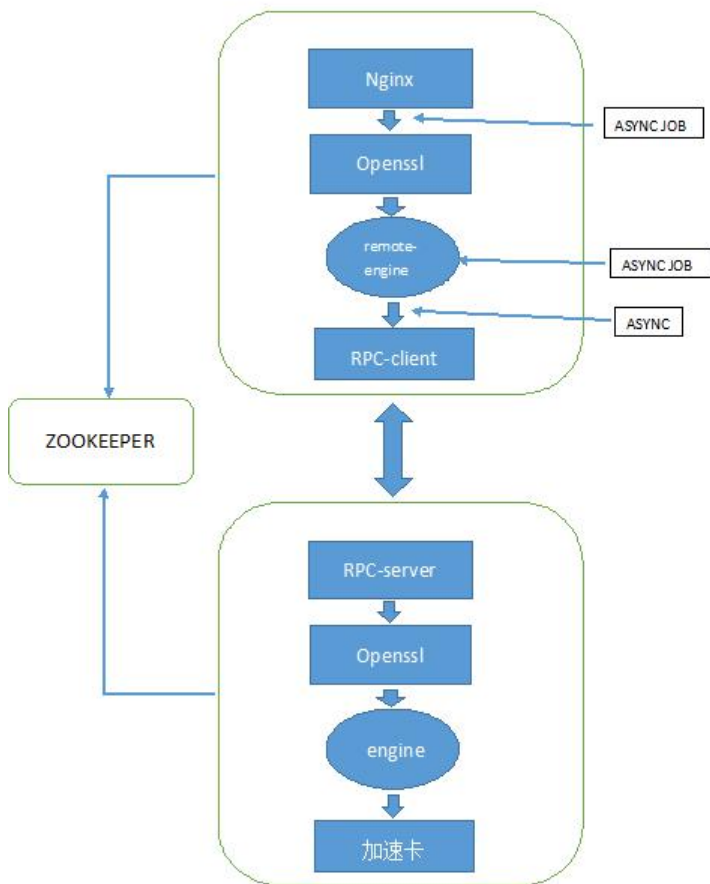


4层优化架构





7层优化架构（SSL计算分离）



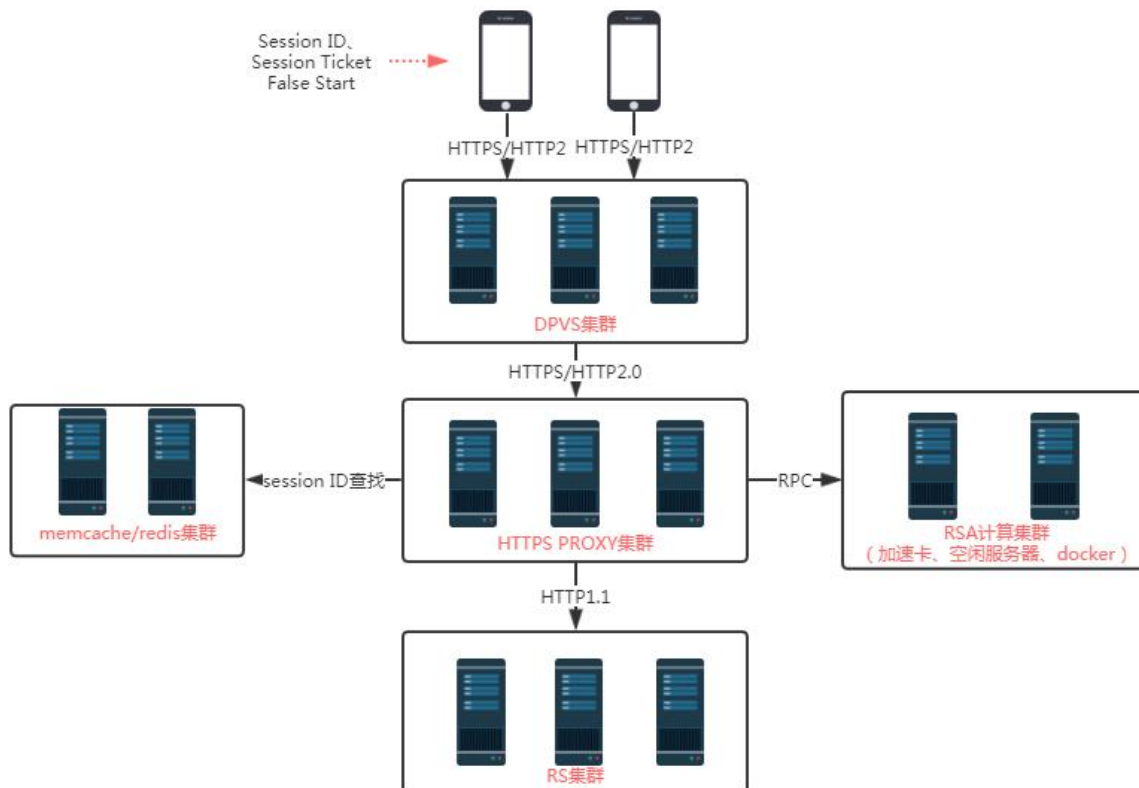
- HTTPS计算分离
– 优化非对称加密

	qat 1张	qat 4张
同步	QPS 4.02w CPU 98%	QPS 8.3w CPU 99%
异步	QPS 4.05w CPU 11.6%	QPS 16.2w CPU 27.3%



7层优化架构（Session共享）

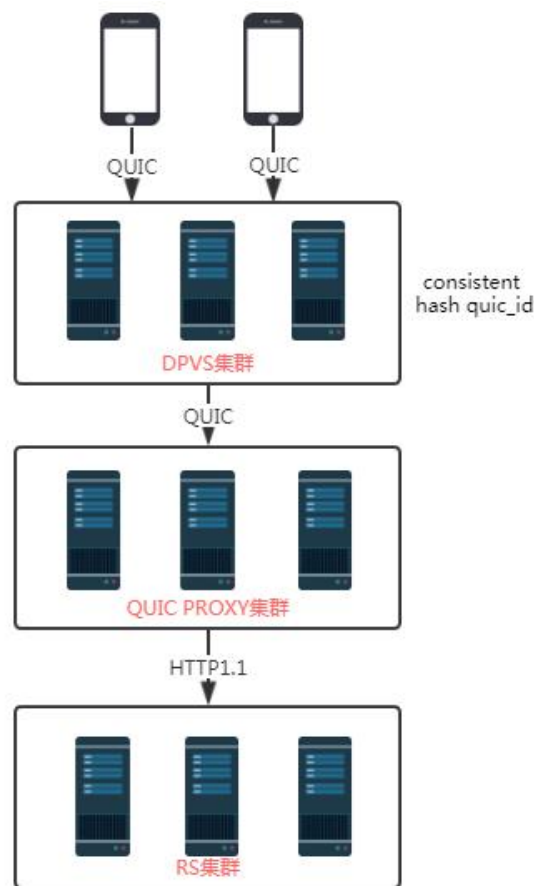
- Session ID共享
- Session Ticket机制
- False Start
- 多级cache
 - 内存
 - memcache/Redis





网络传输优化（QUIC）

- QUIC Proxy
- 高可用（负载均衡）
- 获取真实Client IP（UOA）
- 连接一致性保证（CID）



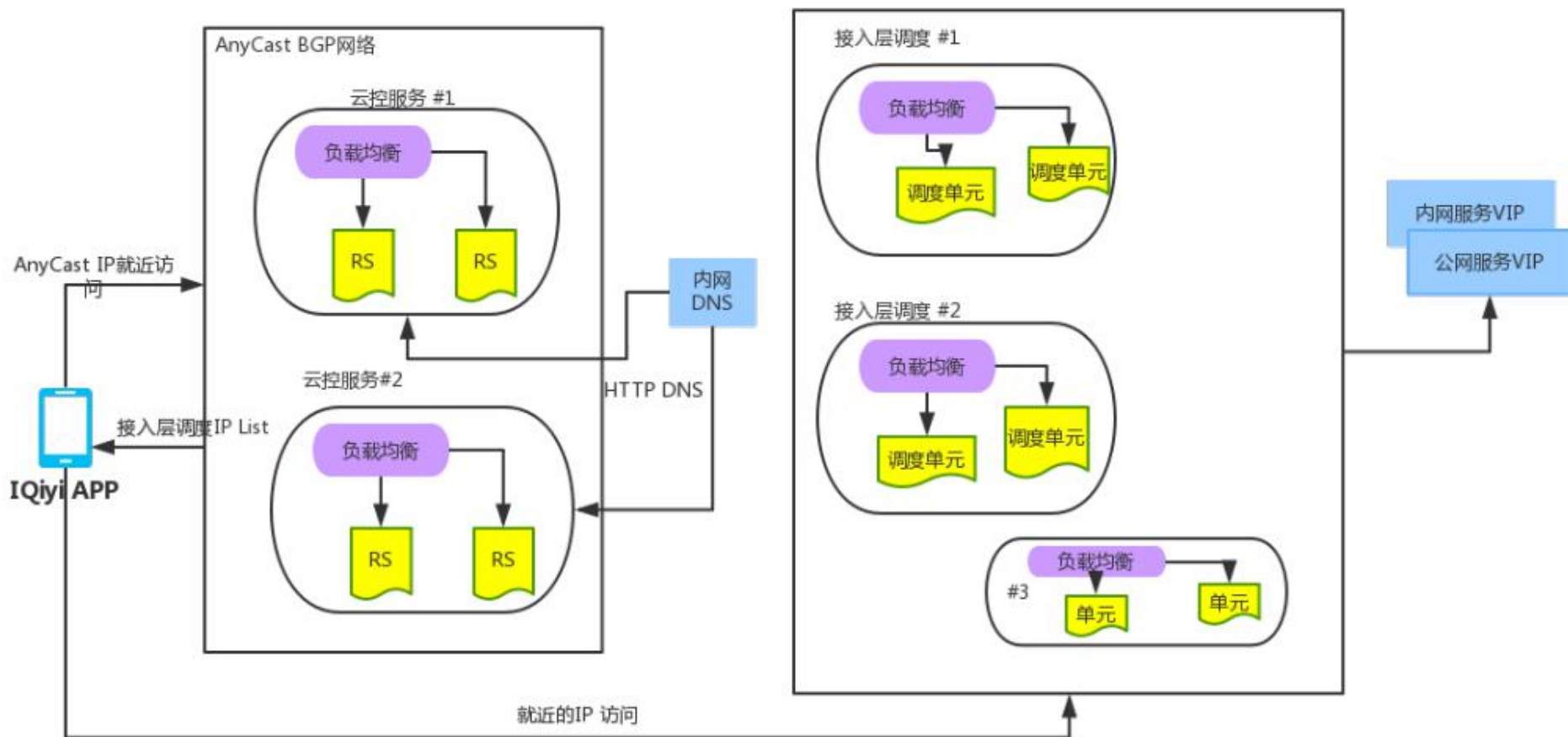


接入层调度详细问题

- 用户的dns解析错误或dns发生故障所导致的问题。
- 用户依据dns解析结果访问业务接口时，业务接口所在的数据中心网络发生故障或业务接口服务出现异常。
- 用户https请求被劫持后重试问题。



接入层调度





接入层调度

- 接入层调度-云控IP
 - 使用BGP Anycast技术，多地区部署了接入集群，与国内多个运营商进行了对等互联。
 - 云控IP根据大数据分析计算出最优的若干接入IP并返回给客户端。
- 接入层调度-接入IP
 - 客户端使用接入IP直连到接入层调度单元，跳过用户dns。
 - 客户端在使用接入IP时，需要对若干接入IP进行连通性测试，优先使用连通性测试通过且延迟最小的接入IP。
 - 客户端一旦与接入IP建连，接入层调度单元将会找到最优的RS。



接入层调度

• 接入层调度-调度单元

- 调度单元负责进行域名解析并把外网域名转为内网域名。
- 调度单元对已申请使用接入层调度的业务域名进行实时健康检测。
- 调度单元根据不同地区业务RS健康状态、往返延迟等因素进行定向调度。
- 如果客户端使用接入层调度提供的sdk对请求进行加密，调度单元将负责解密，并把原始请求发送给业务RS。
- 客户端可以把任何不符合预期的服务器返回结果投递给调度单元，调度单元进行记录用户ip、异常域名与异常响应结果。



后期展望

- 4层负载均衡

- Session共享
- 日志支持
- 精确的流量隔离，监控
- 25G/40G大网卡及性能优化

- 7层负载均衡

- QUIC Proxy与Nginx集成
- Fstack + 计算分离

- 接入层调度

- 更强大的BGP Anycast 云控IP?
- 通过使用机器学习，对接入IP对用户的可用性连通性进行预测，达到更优的推荐效果。
- 调度单元后端长连接

- 4层负载均衡：

<https://github.com/iqiyi/dpvs>

GIAC

全球互联网架构大会

GLOBAL INTERNET ARCHITECTURE CONFERENCE



关注公众号获得
更多案例实践