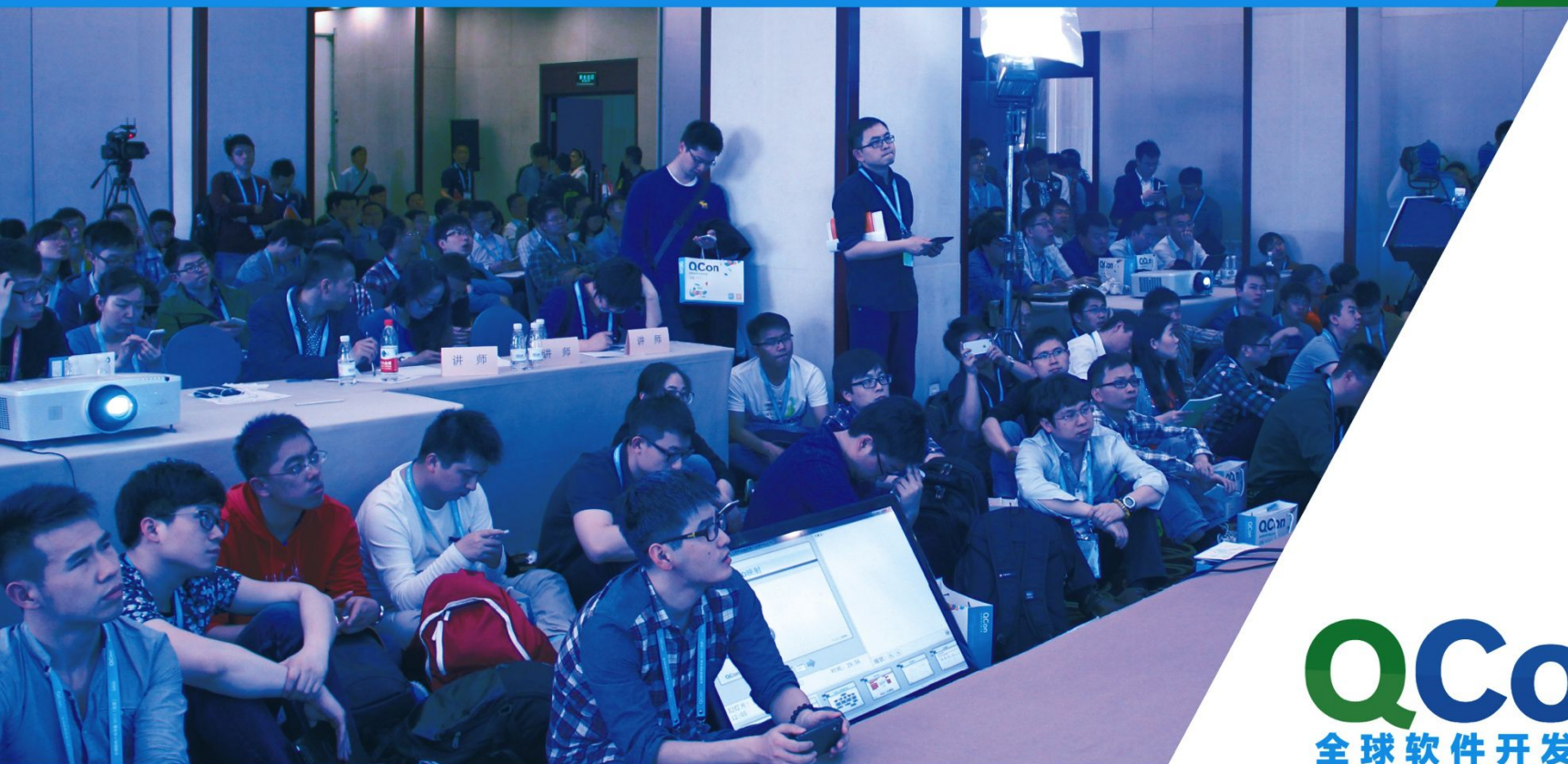


QCon全球软件开发大会

International Software Development Conference



QCon
全球软件开发大会

Geekbang>

极客邦科技

全球领先的技术人学习和交流平台

扫我，码上开启新世界



Geekbang>

InfoQ | EGO NETWORKS | StuQ

InfoQ

专注中高端技术人员
的社区媒体

EGO NETWORKS

EXTRA GEEKS' ORGANIZATION
高端技术人员
学习型社交网络

StuQ

实践驱动的IT职业
学习和服务平台



促进软件开发领域知识与创新的传播



实践第一 案例为主

时间：2015年12月18-19日 / 地点：北京·国际会议中心

欢迎您参加ArchSummit北京2015, 技术因你而不同



ArchSummit北京二维码



【北京站】

2016年04月21日-23日



关注InfoQ官方信息
及时获取QCon演讲视频信息

选型指南 CDN系统中DNS设计与开发

SpeedyCloud

李孟

2015年10月14日

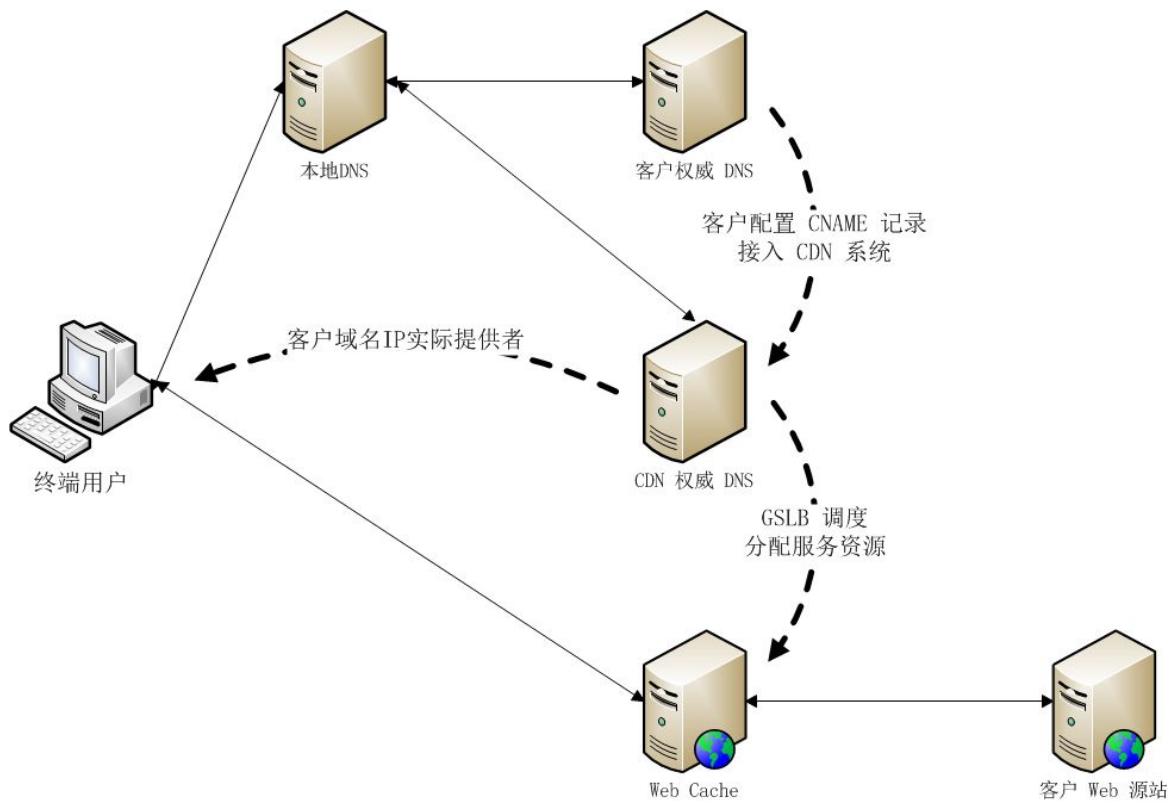
目录

- 智能DNS解析概述
- 剖析域名解析过程
- 性能
- 智能DNS与CDN流量调度的关系

智能解析基础

- 一般业务形式
- 典型CDN DNS支持协议标准
- 典型权威DNS通信特点
- 系统消耗特征

典型业务形式



典型CDN DNS支持协议标准

- rfc 1035 : 基础
- rfc 2671 : EDNS0
- rfc 3596 : AAAA
- ECS支持: draft-vandergaast-edns-client-subnet-04

典型权威DNS通信特点

- 典型报文特征
 - UDP为主
 - 单包请求应答
 - 小数据包
 - 五元组不重复
- 业务特征
 - 延迟敏感
 - 分布式部署

系统消耗特征

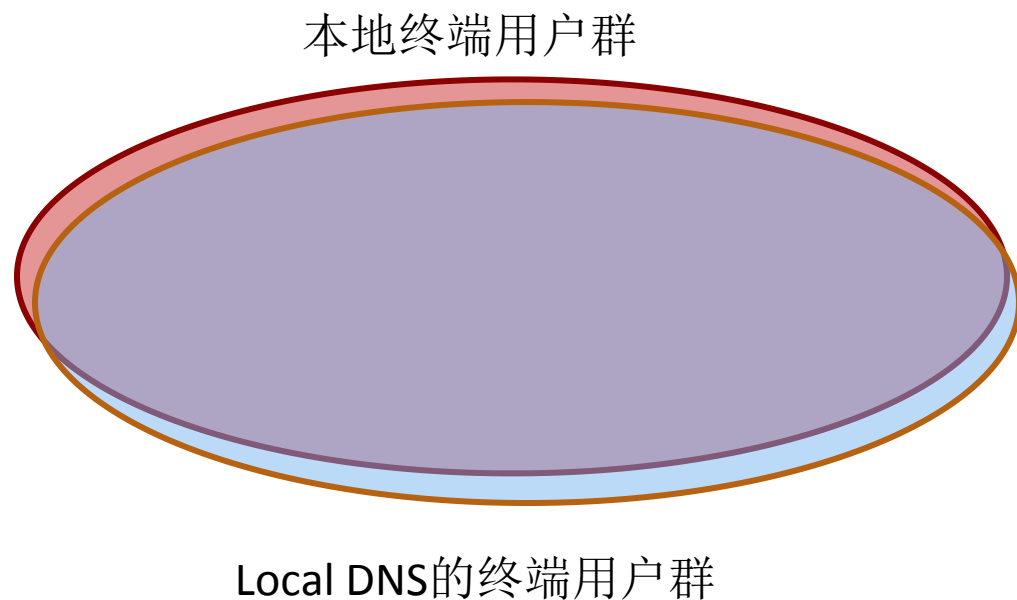
- 网络I/O密集型
 - 小包高频
 - 一次性网络I/O
- CPU密集型
 - 小包高频
 - 一次性网络I/O
 - 域名为字符串查询和处理为主。
 - 额外进行CDN策略计算

剖析域名解析过程

- 终端与Local DNS交互特征
- Local DNS与 CDN DNS交互特征
- Local DNS与 DNS集群的交互特征

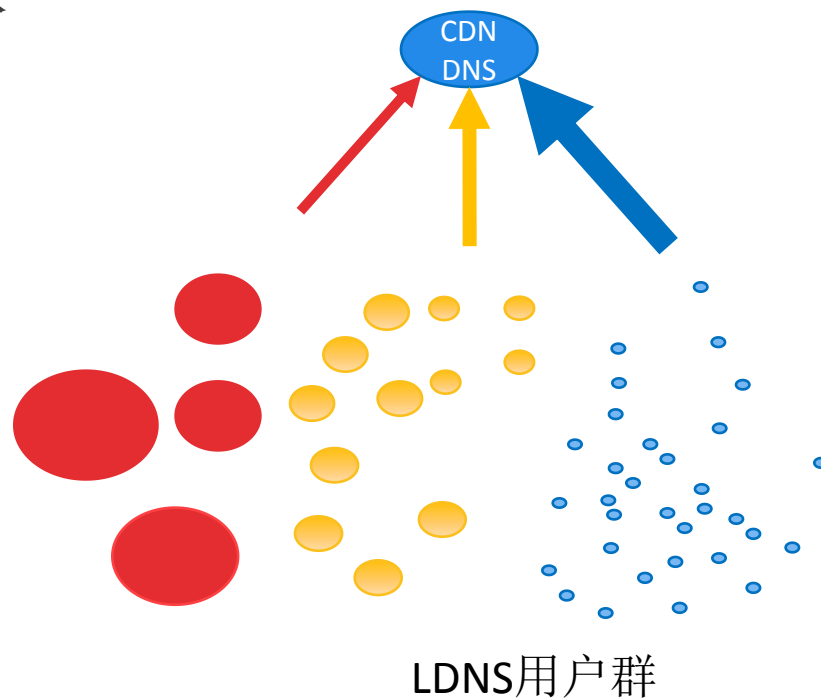
终端与Local DNS交互特征

- 大部分终端使用是当地 Local DNS



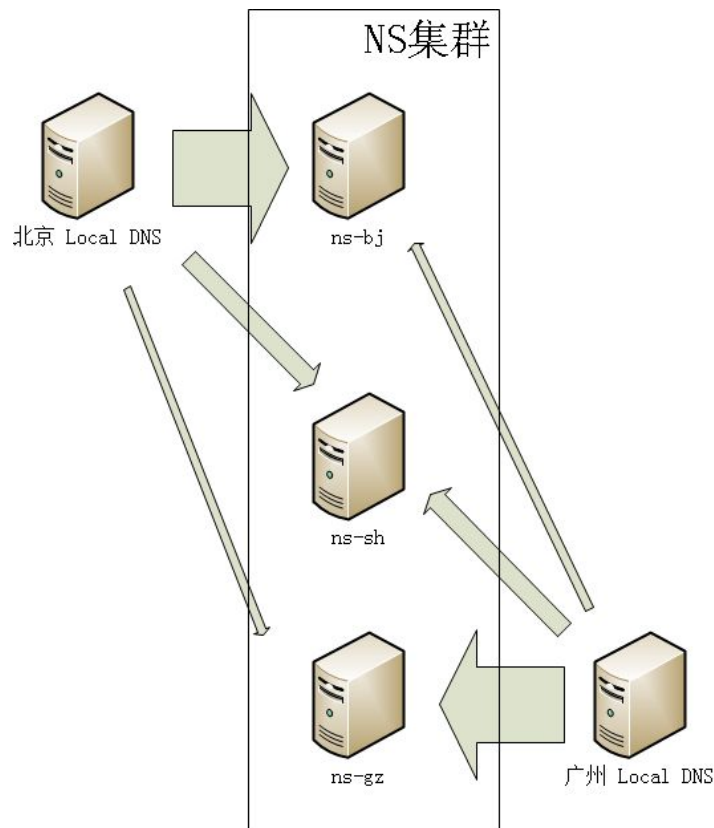
Local DNS与智能DNS交互特征

- Local DNS的用户规模差异极为悬殊
- 从CDN DNS视角
 - Local DNS的cache稀释了热点
- 流量调度视角
 - 每次DNS访问涵盖用户群大小不一

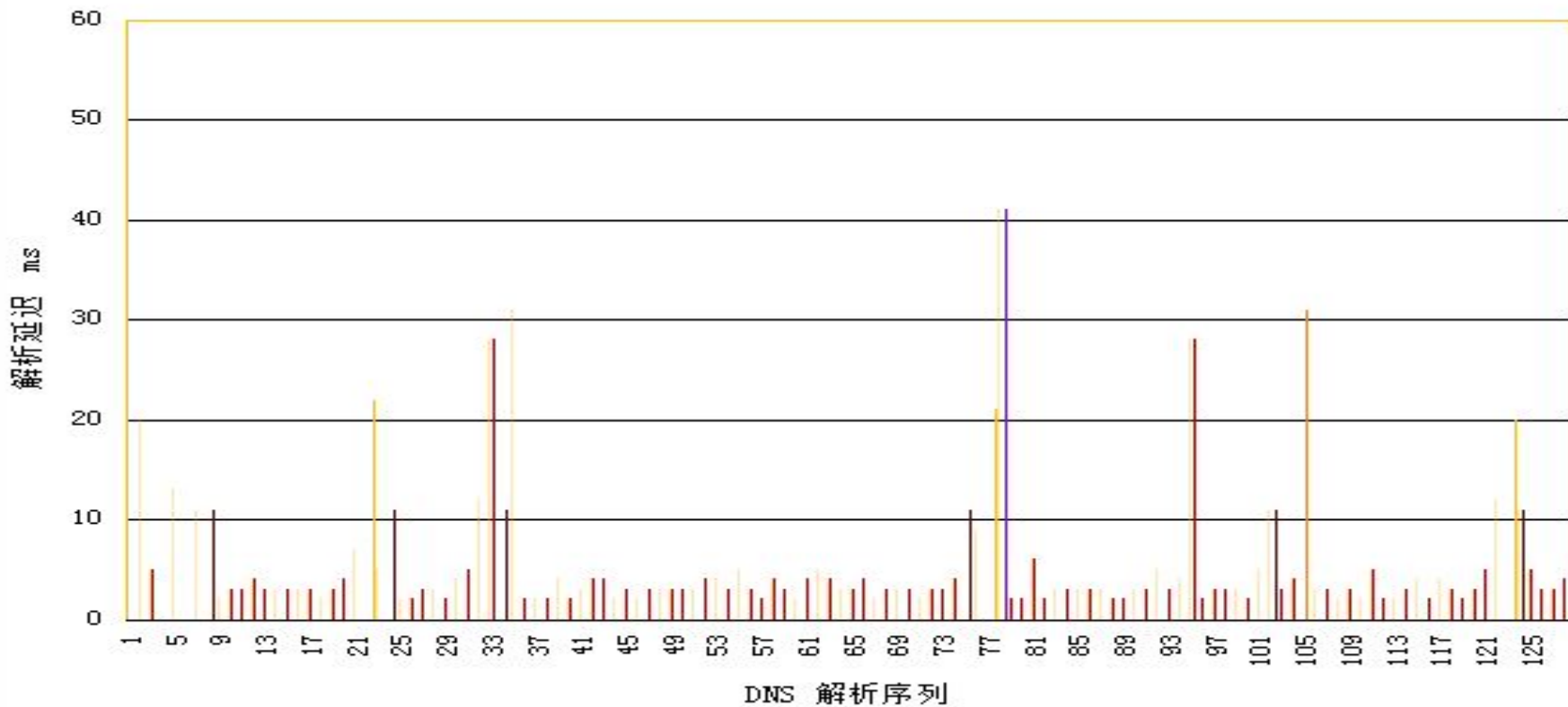


Local DNS与DNS集群的交互特征

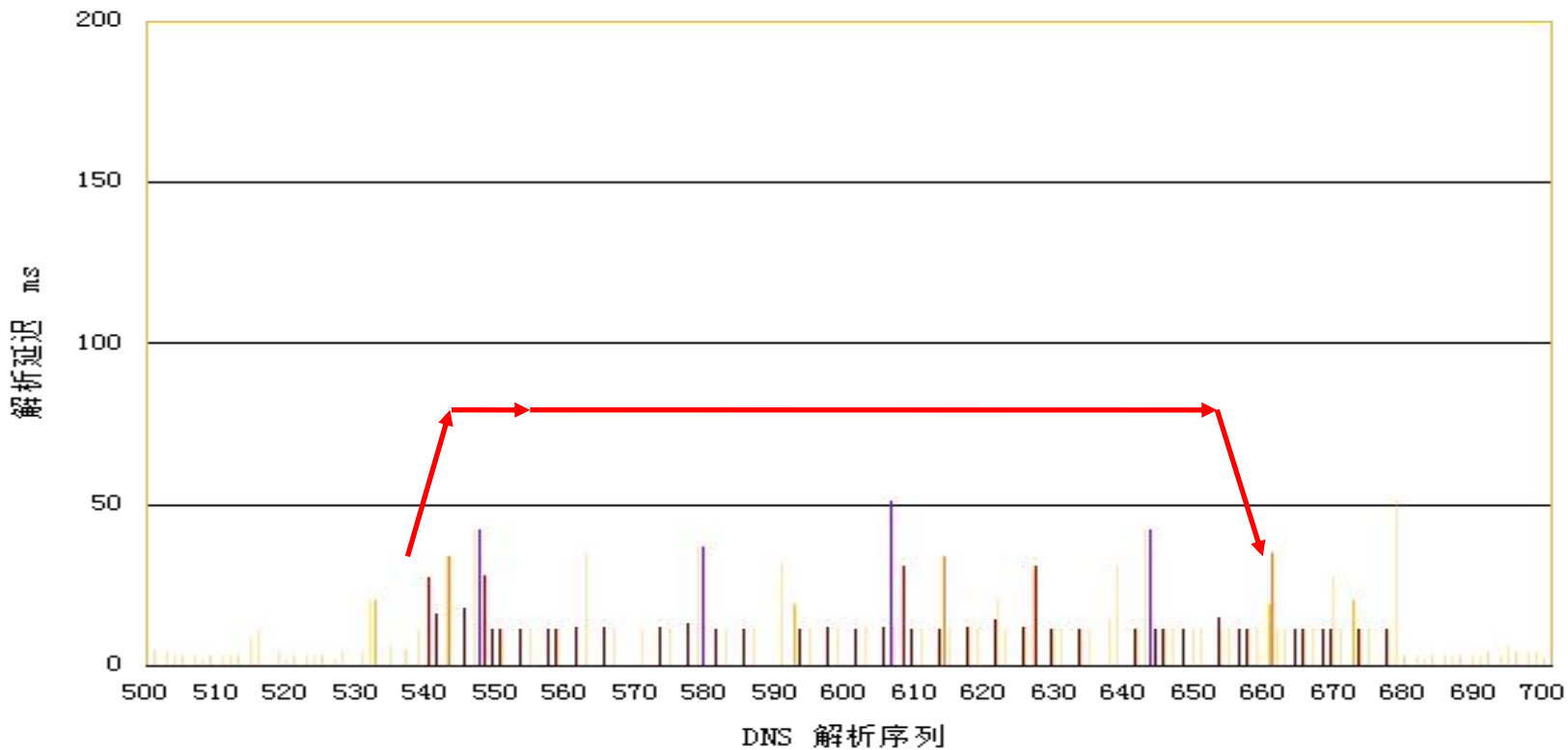
- Local DNS择优：**解析时延敏感**
 - 高概率选择近NS（RTT短）
 - 低概率选择远NS（RTT长）
- 流量调度视角
 - 每个NS获得的访问不对等



集群择优实例



高延迟情形下惩罚实例



性能要求

- 高品质DNS设备的性能特征
- 评估性能设计指标
- 性能测试
- 网络I/O选型
- 新近出现的DNS数据存储

高品质DNS设备的性能特征

- 与ICMP echo服务（ping）特征相仿

性能指标评估

性能量级

C40M - 40Gb

C10M - 10Gb

C1M - 1Gb

C500K-C100 - 100Mb

C10K - 10Mb

C1K - 1Mb

C100K

域名解析量

DDoS

小型DDoS

小型DDoS

大型网站峰值解析

活跃的网站

单机性能

Knot、NSD、BIND、PowerDNS

BIND+DLZ+NOSQL

热点稀释对性能影响

- 热度稀释原因
 - 上游Local DNS是cache型服务，有效期内不会再次访问
 - NS集群分散了访问加大了访问周期
- 热度稀释效应
 - 线上性能比测试性能低一大截
 - 软硬件各级cache miss率偏高，
 - 云环境下，因为自身数据活跃性的劣势，软硬件各级cache miss率偏高。
- 需要注意
 - 评估中需要额外考虑，留出足够富余
 - 对于存储和算法等的选型中，随机访问的性能稳定性需要额外考量
 - 回放或clone在线流量进行DNS测试。

性能测试

- queryperf
- DNSPerf
- tcpreplay 回放
- tcpcopy
- 性能测试仪器

性能评价误区 1

- CDN DNS的并发性能是多少？
 - 作为权威DNS，业务特点是迅速完成一次性的应答。
 - 成熟的权威DNS可以快速处理超量访问，避免堆积并发。
- **QPS是评价CDN DNS性能的基本指标**

性能评价误区 2

- 使用queryperf、DNSPerf评价在线CDN DNS在线性能？
 - CDN DNS的Local DNS变化维度很难覆盖到
 - 五元组固定，不能正确体现出底层消耗
 - queryperf、DNSPerf配合被测DNS设备的应答节奏进行固定强度压力测试
- 互联网访问节奏通常服从泊松分布，但会波动。
 - 但DNS也存在人为相关性的访问加剧波动。
 - 如周期性探测，周期性的网页爬取。

网络I/O选型

性能量级

Q40M - 40Gb

Q10M - 10Gb

Q1M - 1Gb

Q100K - 100Mb

Q10K - 10Mb

Q1K - 1Mb

网络I/O

DPDK

PF_RING, Netmap

BPS, netfilter

Socket

新近出现的DNS数据存储

- Open LDAP LMDB
 - Lightning Memory-Mapped Database
- PowerDNS + LMDB
 - 400kqps
- Knot DNS
 - 700kqps

负载均衡选型

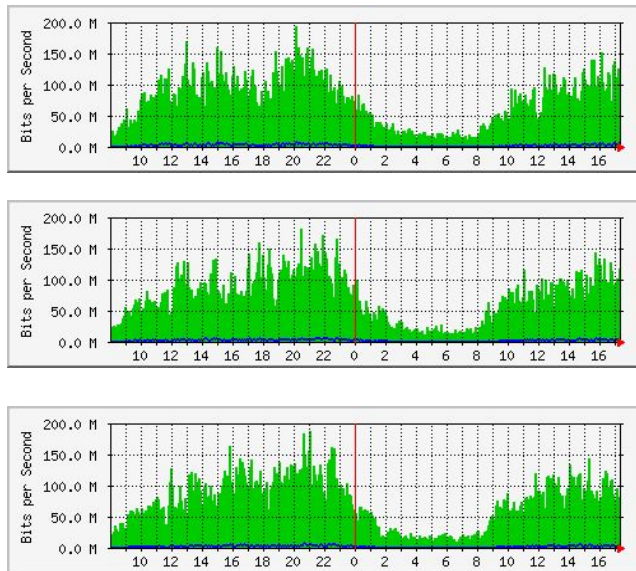
- 交换机/路由器 + 节点内的Anycast
 - 如果条件允许首选方式
- 硬负载均衡
- 软负载均衡

智能DNS与CDN流量调度

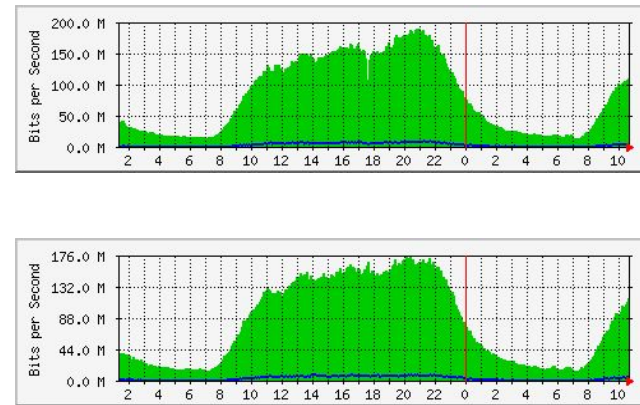
- 实例分析
- 影响DNS调度流量表达的因素
- 如何评价DNS流量调度效果
- 可采用调度算法

DNS调度优化实例对比

优化前



优化后

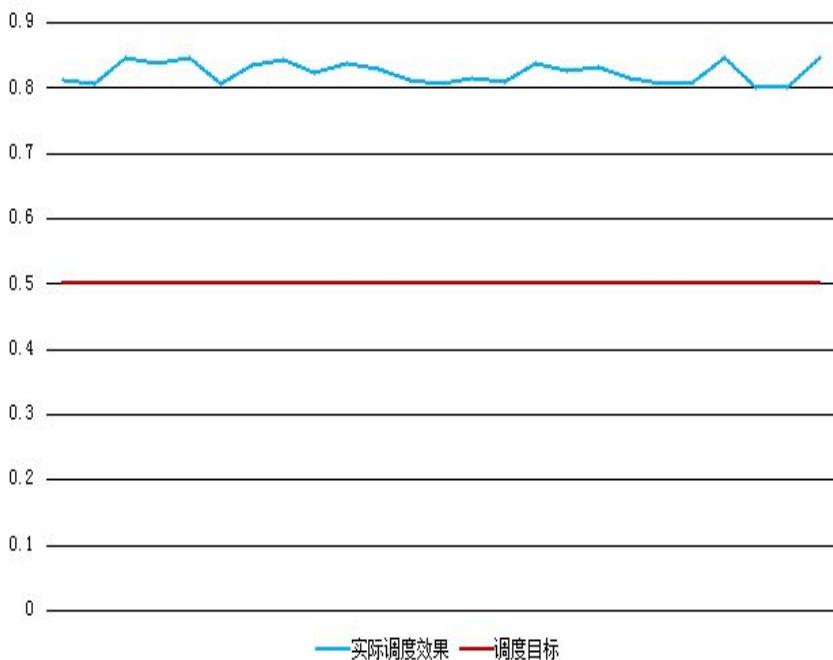


影响智能DNS调度效果的因素

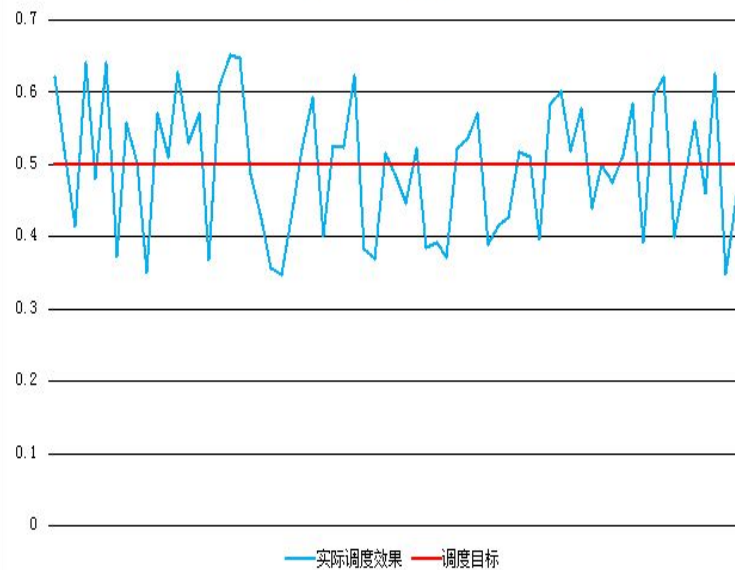
- 部分终端改变了自身的Local DNS
- Local DNS Cache 影响调度生效和失效
- Local DNS涵盖的用户群大小不一
- Local DNS择优行为影响数据对等

DNS智能调度的准度与精度

准度差，精度高



精度差，准度高



智能DNS的精度和准度特征的影响

- 精度特征
 - 影响调度系统动态调节难度
 - 影响设备带宽利用效果
- 准度特征
 - 影响调度系统预测规划难度

DNS调度依据

- DNS的统计计数数据作为调度依据
 - 不推荐直接使用
- 按照固有属性进行流量调度
 - Local DNS IP 地域属性
 - Local DNS IP 网络属性
- 按照无状态属性进行流量调度
 - Hash特征
 - 随机

常见调度表达方式

- 默认流量等分
- 按照比例
- IP分类
 - 静态的地域分类
 - 动态的网络链路
 - ...

流量调度常见形式特征

| 调度方法 | 准度 | 精度 |
|--------------------|----|-----|
| 地区切分 | 很差 | 好 |
| 默认等分 | 很好 | 很好 |
| 随机比例 | 准确 | 不稳定 |
| 按照Local DNS IP切分流量 | 差 | 好 |

智能DNS与其他调度方式结合

- 智能DNS粗分，其他调度方式细分
 - Http 302
 - 集群
 - 设备限连接，限流。
- 节点内混用各取所长
 - 高品质，小流量业务
 - 高精度智能DNS调度
 - 低要求，大流量业务
 - 高准度智能DNS调度，跨节点调度
 - 辅助设备限连接，限流。

THANKS

Brought by **InfoQ**

International Software Development Conference