

基于用户感知的 DNS 解析网络性能测量技术

杜跃进^{1,2,4} 张兆心^{1,3} 王克¹ 杨逍¹ 胡萍¹

(1. 哈尔滨工业大学计算机科学与技术学院, 哈尔滨, 150001;

2. 国家计算机网络应急技术处理协调中心, 北京, 100029;

3. 中国科学院计算技术研究所, 北京, 100190; 4. 北京邮电大学计算机学院, 北京, 100876)

摘要:通过对 DNS 解析数据流的分析,从终端用户角度出发,选取路由路径平均长度、路由往返时延 RTT、时延抖动、路径瓶颈带宽和丢包率作为 DNS 解析网络性能测量参数,并建立相应的测试方法。对全国各省市首选及备份的 60 个 DNS 递归解析服务器进行主动测量。测量结果表明:DNS 解析路由路径平均长度为 13.29 跳,平均往返时延为 104.33 ms,平均时延抖动为 2.403 ms,瓶颈带宽分别为 80.248 Mb/s(ICMP 方式)和 87.437 Mb/s(UDP 方式),丢包率为 6.077%。

关键词:网络安全;域名系统;用户感知;主动测量;性能参数

中图分类号:TP393

文献标志码:A

文章编号:1005-2615(2013)01-0110-06

Performance Measurement Technology of DNS Resolution Network Based on User Perception

Du Yuejin^{1,2,4}, Zhang Zhaoxin^{1,3}, Wang Ke¹, Yang Xiao¹, Hu Ping¹

(1. School of Computer Science and Technology, Harbin Institute of Technology, Harbin, 150001, China;

2. National Computer Network Emergency Response Technical Team/Coordination Center of China, Beijing, 100029, China;

3. Institute of Computing Technology, Chinese Academy of Sciences, Beijing, 100190, China;

4. School of Computer Science, Beijing University of Posts and Telecommunication, Beijing, 100876, China)

Abstract: Through analyzing on the resolution flow of domain name system(DNS), from the viewpoint of the end user, average length of routing path, round-trip time(RTT), delay jitter, bottleneck bandwidth, and packet loss rate are extracted as the performance measurement parameters of DNS resolution network, and the corresponding measurement methods are established. 60 DNS recursive servers of some provinces are measured. The measurement results show that, the average length of routing path of DNS is 13.29 hops, the average RTT is 104.33 ms, the average delay jitter is 2.403 ms, the average bottleneck bandwidth is 80.248 Mb/s(ICMP model) and 87.437 Mb/s(UDP model), and the packet loss rate is 6.077%.

Key words: network security; domain name system; user perception; active measurement; performance parameter

域名系统(Domain name system, DNS)^[1-4]是 Internet 的核心基础设施,是各种网络应用得以正

基金项目:国家高技术研究发展计划(“八六三”计划)重点(2007AA010503)资助项目;国家自然科学基金(61100189, 61003261)资助项目;山东省中青年科学家奖励基金(BS2011DX001)资助项目;威海市科技攻关(2010-3-96)资助项目;哈尔滨工业大学科研创新基金(HIT.NSRIF.2011119)资助项目。

收稿日期:2012-07-16; **修订日期:**2012-10-16

通信作者:张兆心,男,博士,副教授,1979 年出生, E-mail:heart@hit.edu.cn。

常运行的前提和保障。20 多年来, DNS 发展迅速, 截止到 2011 年第 1 季度, 全球大约有 2.098 亿个域名, 比去年增加 7.9%, 逼近 2.1 亿。同时, 根据研究机构的探测统计结果, 截至 2011 年 5 月 20 日, 监测到世界范围内域名服务器总量为 15 457 337 个, 其中权威域名服务器 2 852 319 个, 递归域名服务器 12 605 018 个。活跃域名服务器数量为 1 245 135 个, 其中权威域名服务器 561 250 个, 递归域名服务器 683 885 个。

DNS 的网络性能影响海量互联网应用的服务质量。从用户的角度, DNS 是一个典型的端到端网络应用, 用户主要关注其准确性和时效性。其中准确性由 DNS 系统保证, 而时效性和用户与 DNS 服务器之间的网络状况有直接的关系。如何测量用户和 DNS 服务器之间的网络状况, 从而为合理地选择 DNS 服务器提供数据依据, 成为一个亟待解决的问题。凯斯西储大学的 Qian 等人^[5]使用一种新的 P2P 的 DNS 机制, 使 DNS 服务器更接近用户。Bell 实验室的 Deb 等人^[6]通过自动注册的方式来评估 DNS 服务器的响应时间, 从而解决 DNS 服务器选择问题。AT&T 实验室的 Morley 等人^[7]提出了一种精确高效的评价客户与本地 DNS 服务器的接近性方法。Fomenkov 等人^[8]从 DNS 互联网拓扑结构和宏观性能方面测量了根名称服务器, 结合 skitter 工具建立了一个评价最优根服务器安置和效率的框架系统。文献^[9]通过对域名服务器和客户端的实际位置和服务表现指标分析, 讨论了关于 DNS 服务器选择的有效性。香港大学的 Pan 等人^[10]针对内容分发网络 (CDN), 基于用户的感知体验, 研究了 DNS 服务器的选择方案。

上述研究主要从如何让 DNS 服务器更加靠近用户的角度出发, 提高了域名解析的效率。但域名解析的效率不仅与用户和 DNS 服务器之间的距离相关, 而且与用户和 DNS 之间的网络状况有直接相关, 更加接近并不代表效率更高。因此, 本文从网络服务质量角度出发, 选取与 DNS 解析相关的网络性能指标参数, 并建立相应的测试方法集。通过主动探测的方式获取各网络性能指标参数值, 从而为 DNS 服务器的选取提供数据依据。

1 DNS 网络性能指标选取

网络性能与终端性能及用户操作无关, 是网络本身特性的体现。网络性能代表了网络基础设施向上层应用提供的服务质量, 它是由一系列对于运

营商有意义的并可用于系统设计、配置、操作和维护的参数组成。对于 DNS 服务的性能, 用户主要关心的是端到端的性能。通过对 DNS 解析数据流的分析, 在此选取路由路径平均长度、路由往返时延 RTT、时延抖动、路径瓶颈带宽和丢包率作为性能参数。

1.1 路由路径平均长度

当针对的测量对象为客户机和 DNS 服务器这种端到端的网络性能时, 其数据包转发所经过的路径是非常重要的一项。理想状态下, 如果经由转发的路由器个数越少、路由规则越简单, 则传输速度会更快, 处理和响应时间也越及时, 同时相应的丢包可能性也会大大减小。

设客户机至 DNS 服务器端经过的路由器数为 m , 则单次测量时访问的路由路径长度 $L=m$, 按照 Poisson 分布的时间间隔探测路径长度的次数为 n , 则要测量的网络性能指标之一路由路径平均长度定义为

$$\bar{L} = \frac{\sum_{i=1}^n m}{n} \quad (1)$$

1.2 路由往返时延 RTT

数据分组在网络中传输的延时时间 (包括单向时延和往返时延), 即 IP 包穿越一个或多个网段所经历的时间。延迟由固定延迟和可变延迟两部分组成。固定延迟基本不变, 由传播延迟和传输延迟构成; 可变延迟由中间路由器处理延迟和排队等待延迟两部分构成。对于单向延迟测量要求时钟严格同步, 这在实际的测量中很难做到, 许多测量方案都采用往返延迟, 以避免时钟同步问题。

设在 T 时刻源结点向目的结点发出一个数据报文, 在 $T+dT$ 时刻源结点收到来自目的结点的应答报文的最后一位, 则往返时间 (Round-trip time, RTT) 为 dT 。

1.3 时延抖动

数据报文时延抖动是指由于在网络上变动的排队时延等原因造成的数据报文之间的时间间隔不统一, 它代表了数据报文传输过程的平稳度。数据报文时延抖动通过测量数据报文之间的时间间隔来获得。

对于某一个链路, 通过对时延做时间上的扩展, 抽样时间间隔设定为满足 Poisson 分布, 可得到一段时间内的时延序列 $d_i (i=1, 2, \dots, n)$ 。定义 $\text{Avg}(d_i)$ 为该时延序列的平均值。则时延抖动定义为

$$D = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (d_i - \text{Avg}(d_i))^2} \quad (2)$$

1.4 路径瓶颈带宽

路径的容量是链路或路径在单位时间内能够支持的最大传输量。路径的容量不因链路中其他流量的存在而改变,是路径的一个不经常改变的属性,只有在底层硬件设施发生变化时(如路由器或链路的速度发生了变化)才会改变。一条路径的容量是由其中容量最小的链路决定的。如图1所示,该链路路径的瓶颈带宽为50 Mb/s。

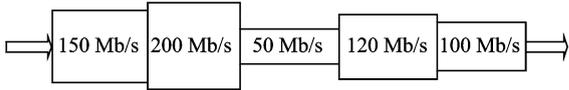


图1 瓶颈带宽示例

1.5 丢包率

由于链路或者路径的路由不稳定性和传输排队、拥塞等,会造成数据报的丢包问题,丢包率定义为链路或路径在一段时间内丢失包占传输包总数的比例。一个链路或路径在时段 n 的丢包率为: $l_n = L_n / C_n$ 。其中, C_n 为在时段 n 内进入网络的报文数, L_n 为时段 n 内丢失的报文数。对于一个由 H 跳组成的端到端的路径,如果其 i 跳路径在时段 n 内的丢包率为 $l_{n,i}$,则整条路径的丢包率为

$$l_n = 1 - \prod_{i=1}^H (1 - l_{n,i}) \quad (3)$$

2 测量方案设计

为评价和分析一种分布式网络应用的服务性能,必须要求有效地、准确地获取性能特征数据,网络测量技术^[1]是获取性能特征准确而有效的手段。

2.1 测量方法与测量抽样

端到端性能指标的测量可基于已有协议进行设计和实现。对于路由往返时延、路由路径长度、时延抖动、丢包率指标可通过在测量结点主动发送探测报文,并对接收到的报文数量、时间、顺序等进行分析获得端到端性能参数。

(1)测量方法 现有的测量方法主要包括基于 ICMP,UDP 和 TCP 的探测方法。基于 ICMP 的探测方法,由于路由器、防火墙、服务器等设备对服务质量的支持,配置过滤 ICMP 的探测报文,使得探测结果的真实性存在偏差;基于 UDP 高端口的探测方法,同样存在被防火墙等设备过滤的问题;基于 TCP 协议的探测方法基本可穿透大多数防火墙。因此在测量方法中,综合采用基于 ICMP,

UDP 和 TCP 的方法实现测量。

(2)测量抽样 由于网络的性能处于不断变化中,这种按照一定间隔进行的测量实际上是对网络性能的抽样,尤其是对于丢包等小概率事件。因此,为了准确地刻画网络性能的变化情况,在主动测量中需要对探测报文的探测时间间隔进行设计以捕获到网络性能的变化。如何设计抽样间隔才能更加接近真实性能,一直是网络性能测量关注的问题。Poisson 分布是 IPPM 推荐的测量间隔分布,该分布对于估计均值等原始性能特征是无偏的。Poisson 抽样是一种随机增量抽样,对于被抽样过程,该抽样采集到该过程任意时刻值的概率相等。Poisson 抽样得到的每个状态的比率和该状态在总体时间中所占的比率是一致的,也称为具有 PASTA 特性。Poisson 抽样的随机性可有效地避免与网络中存在的周期性事件同步。因此,针对各个网络性能指标的主动测量,均采用 Poisson 分布的时间间隔测量和抽样。

2.2 路径瓶颈带宽测量原理与设计

端到端的网络带宽测量通常采用主动测量的方式进行,通过向网络中注入数据报文并对测试结果进行计算。带宽测量的方法通常是观察链路的属性对数据报文时延的影响。链路带宽在两个方面影响数据报文的传输时延:排队时延和传输时延。数据报文的排队时延是由链路的容量及它前面的数据报文队列大小决定。数据报文的传输时延是由链路的容量和数据报文的大小决定。

针对网络路径中的瓶颈(窄链路)带宽,采用基于探测报文间隔模型的方法进行主动测量。通过向测量路径中发送一对探测报文,然后分析这些探测报文进入网络时的时间间隔与离开网络时的时间间隔之间的关系计算路径的带宽,它既可以探测路径容量,也可以用来探测路径的可用带宽。报文对测量方法通过发送两个相同大小的背靠背数据报文,当数据报文沿着传输路径传输时,在容量最小的链路,将会由于排队时延而导致报文对之间的时间间隔增大,这个时延将会在后续传输中被保持,在接收端可以测得两个报文的时延差,据此可以测得容量最小的链路的容量。

对于一个有 H 跳的路径,假设路径上没有任何背景流量,第 i 跳链路的容量为 C_i , Δ_i 为长度为 L 的 2 个数据报文的最后一位到达第 i 跳时的时间间隔,即为时延差量。由于 2 个报文的大小相等,因此它们会遭遇到同样的传输时延、传播时延、路由器处理时延。如果报文对进入链路的时延间

隔小于链路传输第 1 个报文的时延,第 2 个报文将会等待第一个报文传输完成后才能开始传输,因此导致排队延迟,报文对之间的时延差量增大。相反,如果报文对进入链路的时延间隔大于链路传输到第 1 个报文的时延,报文对之间的时延间隔将保持不变。因此数据报文离开第 i 跳时的时间间隔为

$$\Delta_{i+1} = \max(\Delta_i, \frac{L}{C_i}) \quad (4)$$

报文对之间的时延差量会在后续的传输中被保持或增大,因此报文对离开探测路径时的时延差量为

$$\Delta_{H+1} = \max_{i=0, \dots, H} (\Delta_0, \frac{L}{C_i}) = \max(\Delta_0, \frac{L}{\min_{i=0, \dots, H} C_i}) \quad (5)$$

如果发送探测报文的速度 Δ_0 小于探测报文在窄链路上的时延间隔,即

$$\Delta_0 < \frac{L}{\min_{i=0, \dots, H} C_i} = \frac{L}{C} = \Delta_{H+1} \quad (6)$$

式中: C 为路径的容量。从式(6)中可以得到路径的容量为

$$C = \frac{L}{\Delta_{H+1}} \quad (7)$$

因此,通过报文对之间时延差量的测量,可推断出窄链路的容量,也就是网络路径的容量。为了让探测报文能够在窄链路上背对背地排队,发送探测报文的速度 Δ_0 必须小于探测报文在窄链路上的时延间隔,即 $\Delta_0 < L/C$ 。

为了提高探测效率和准确性,在实际的探测中,结合单个 ICMP 报文的时延和报文对的时延差量进行路径容量估算。如果测量报文对中的任意一个报文都没有在路径上遭遇排队延迟,那么这两个报文的时延之和一定是最小的,而且在一个路径中,虽然两个报文同时没有遭到排队时延的概率较小,但在多次测量的情况下仍然是存在的^[12]。通过这样的报文对估算得到的路径容量具有较大的准确性。

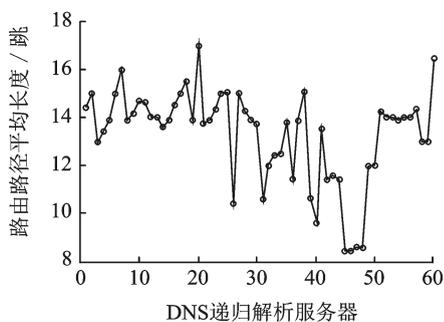
3 实验及结果分析

分别选取全国各个省市的首选、备份 60 个 DNS 服务器作为探测实验对象,涉及到电信、网通、教育网三大 ISP。以 Alex top 500 域名为测试数据集,进行测试,测试结果如下。

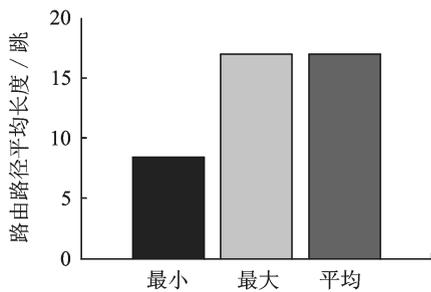
(1) 路由路径平均长度

对每个 DNS 服务器进行主动探测,并对探测

结果进行统计,获得每个目标服务器的路由平均路径长度分布如图 2 所示。从图 2(a)中可以看出,用户到 DNS 递归解析服务器之间的路由路径长度变化比较大,但 80% 的路由路径长度集中在 [11.29, 15.29] 的范围内;从图 2(b)中可以看出,路由路径长度在 [8.4, 17] 范围内,其中整体访问路径平均长度为 13.29 跳。



(a) 路由路径平均长度分布



(b) 路由路径平均长度范围

图 2 路由路径平均长度探测结果

(2) 路由往返时延 RTT

通过对路径探测中往返时间进行计算,可得用户与 DNS 服务器间的往返时延。各 DNS 服务器的访问时延分布如图 3 所示。从图 3(a)中可以看出,路由往返时延随 DNS 服务器的不同有较大变化,但 83.33% 以上的路由往返时延低于 160 ms;从图 3(b)中可以看出,路由往返时延分布在 [9.54, 300.77] 区间中,其中整体平均往返时延为 104.33 ms。

(3) 时延抖动

按照 Poisson 分布的探测时间间隔进行分时段段的独立探测,得到不同 DNS 服务器的访问时延抖动,各 DNS 服务器的访问时延抖动分布如图 4 所示。从图 4(a)中可以看出,时延抖动随 DNS 服务器的不同变化不大,93.33% 以上时延抖动小于 4 ms,说明网络状况比较稳定;从图 4(b)中可以看出,时延抖动分布在 [0.606, 15.237] 区间中,其中

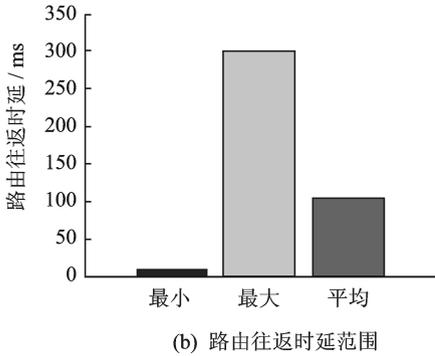
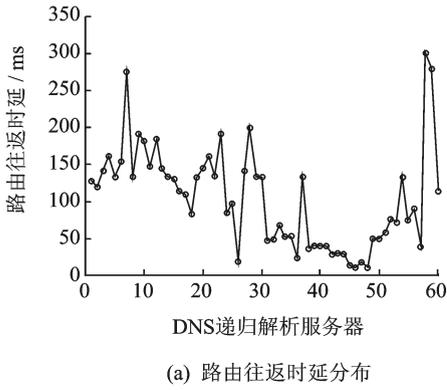


图3 路由往返时延探测结果

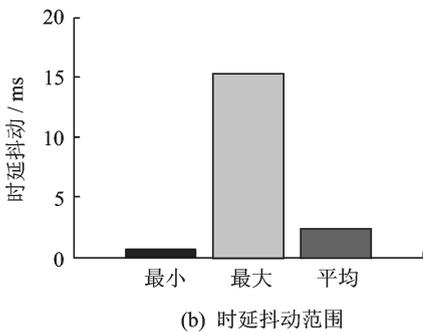
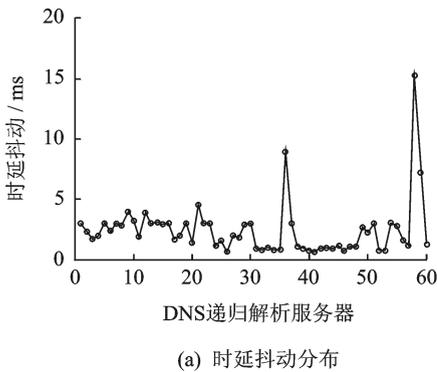


图4 时延抖动探测结果

平均时延抖动为 2.403 ms。

(4) 路径瓶颈带宽

针对 DNS 服务器的访问路径,采用 Poisson 分布的时间间隔探测瓶颈带宽,采用基于 ICMP 和

UDP 两种报文对每个服务器进行独立探测,每种方式探测 10 次,并对探测结果取平均值,探测结果如图 5 所示。

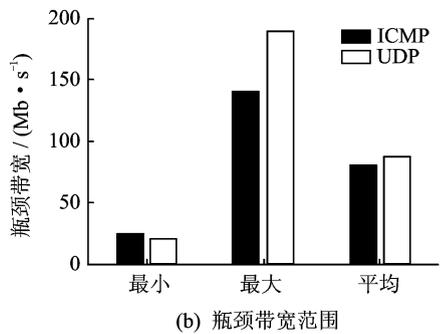
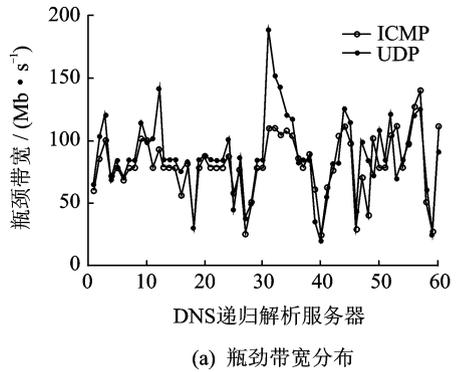
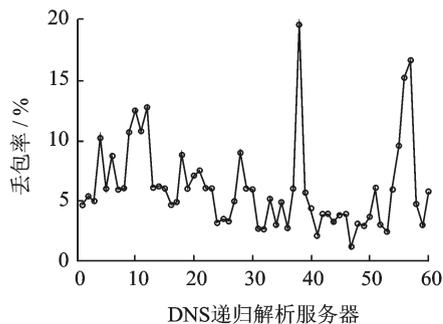


图5 瓶颈带宽探测结果

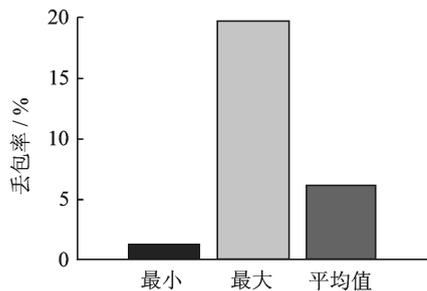
从图 5(a)中可以看出采用两种方式探测到的瓶颈带宽数据结果基本吻合,以 ICMP 和 UDP 两种方式探测的瓶颈带宽分别有 80%和 65%集中在 [50,110]区间中;从图 5(b)中可以看出,以 ICMP 和 UDP 两种方式探测的瓶颈带宽分别在 [24.08, 140.4]和 [19.48,188.98]区间中,平均瓶颈带宽分别为 80.248 Mb/s 和 87.437 Mb/s,相差约 8.22%。联合路由往返时延及时延抖动,说明现在网络的带宽基本满足 DNS 解析的需求。

(5) 丢包率

丢包率的测量是在时延抖动探测、瓶颈带宽探测的过程中进行统计。分别针对目标 DNS 服务器探测过程中的丢包进行统计计算,得到丢包率的测量结果,如图 6 所示。从图 6(a)中可以看出,丢包率的变化不大,其中 86.67%的丢包率低于 10%,但这样的丢包率已经非常大,说明当前的网络状况一般。究其原因,丢包率高于 10%的主要集中一些路由路径长度和路由往返时延相对较大链路上;从图 6(b)中可以看出,用户访问 DNS 服务器的丢包率在 [1.22, 19.67] 区间中,平均丢包率为 6.077%。



(a) 丢包率分布



(b) 丢包率范围

图 6 丢包率探测结果

4 结束语

本文通过对大量 DNS 解析数据流的深入分析,从终端用户角度出发,选取路由路径平均长度、路由往返时延 RTT、时延抖动、路径瓶颈带宽和丢包率作为 DNS 解析网络性能测量参数,并针对各参数建立相应的测试方法。以 Alex Top 500 域名为例,对全国各省市的首选、备份 60 个 DNS 服务器进行测量。测量结果表明:DNS 解析路由路径平均长度为 13.29 跳,平均往返时延为 104.33 ms,平均时延抖动为 2.403 ms,平均瓶颈带宽分别为 80.248 Mb/s (ICMP 方式)和 87.437 Mb/s (UDP 方式),丢包率为 6.077%。这些测量值说明了当前 DNS 系统的网络状况,解析路由路径较长的 DNS 服务器往往丢包率比较高或延迟比较大。故选取较近的 DNS 服务器,解析所需的时间也较短,建议用户优先选取。

参考文献:

[1] Mockapetris P. Domain names: concepts and facilities[S]. IETF RFC1034, 1987.
 [2] Mockapetris P. Domain names: implementation and

specification[S]. IETF RFC1035, 1987.

- [3] Delbende I, Rossi M, Daube O. DNS of flows with helical symmetry[C]// Theoretical and Computational Fluid Dynamics. 2012,26(1/4): 141-160.
 [4] Berger A, Natale E. Assessing the real-world dynamics of DNS[C]// Lecture Notes in Computer Science. Berlin:Springer,2012:1-14.
 [5] Qian H, W, Rabinovich M, Al-Qudah Z. Bringing local DNS servers close to their clients[C]//2011 IEEE Global Telecommunications Conference. Houston: Case Western Reserve Univ, 2011: 1-6.
 [6] Deb S, srinivasan A, sreenivasa K. An improved DNS server selection algorithm for faster lookups [C]//3rd International Conference on Communication System Software and Middleware. Bangalore: Bell Labs, 2008: 288-295.
 [7] ZhuoqingMorley Mao, Charles D Cranor, Fred Douglass, et al. A precise and efficient evaluation of the proximity between web clients and their local DNS servers[C]//Proceedings of the General Track, 2002 Usenix Annual Technical Conference. San Francisco: Business Wire, 2012: 229-242.
 [8] Fomenkov M, Claffy K, Huffaker B, et al. Macroscopic Internet topology and performance measurements from the DNS root name servers[C]//Proceedings of the LISA 2001 15th Systems Administration Conference. San Diego, California: Usenix sage, 2001:231-240.
 [9] Shaikh A, Tewari R, Agrawal M. On the effectiveness of DNS-based server selection[C]//IEEE INFOCOM. Washington: IEEE Computer Society Press, 2001: 1801-1810.
 [10] Pan J, Hou Y, T, Li B. An overview DNS-based server selection in content distribution networks[J]. Computer Networks, 2003, 43(6): 695-711.
 [11] 杨家海,吴建平,安常青. 互连网络测量理论与应用[M]. 北京:人民邮电出版社, 2009.
 Yang Jiahai, Wu Jianping, An Changqing. Internet measurement theory and applications[M]. Beijing: Press of People's Posts and Telecommunication, 2009.
 [12] Kapoor R, Chen L, Lao L, etc al. CapProbe: a simple and accurate capacity estimation technique[C]// Proceedings of SIGCOMM. New York: Association for Computing Machinery, 2004: 67-78.

