# Web请求的时延刻画和降低时延的协作分配模型*)

傅建明 朱福喜
(武汉大学计算机学院 武汉430072)

## The Analysis of the Response Delay Times of Web Requests

FU Jian-Ming ZHU Fu-Xi
(Computer School of Wuhan University, Wuhan 430072)

**Abstract** With evolution of Internet, many cooperative information systems are developed and deployed. These systems may not meet the requirements of clients because of best-efforts in Internet sometimes, so a lot of techniques are provided to satisfy the Web-QOS(Quality Of Services)now. It is important to understand Web delay and to decrease the delay. In this paper, the Web delay is characterized from both network protocols and network architecture, and some basic techniques are discussed to shorten the delay, furthermore assessments of these techniques are given. Finally, a cooperative assigning model is presented to balance workloads and to decrease the delay on the basis of measuring performances of both networks and servers.

**Keywords** Web delay, DNS, HTTP, Workload balance, Cooperative assigning model

## 1 引言

随着互联网的发展,基于Web的协作系统越来越多,如基于Web的分布式设计系统,基于Web的网格计算,基于Web的电子商务,以及基于Web的工作流办公系统等。这些服务都是以Web为平台,任何客户只要连接上互联网,就可以享受这些系统提供的服务,且客户请求和响应服务都以Web页面存取。互联网的尽力而为的服务模式,可能导致系统不能满足客户的要求。客户的服务质量(QOS)除安全性外最重要的是Web请求的响应时延。另外,客户请求的媒体除文本外,音频和视频也越来越多,这些都会增加时延,仅用足够的网络带宽不能完全解决问题。目前,许多测量表明Web流量占整个互联网流量的70%到80%,成千上万的客户访问各种提供服务的Web页面。客户们都想获得更快更好的服务,希望Web请求的响应时间越小越好;与此同时,服务提供商(ISP)也迫切希望缩短Web请求的时延,提高服务器的容量,满足客户的质量服务。因此,一方面,理解和刻画Web请求的响应时延对客户和服务提供商都十分关键;另一方面,服务提供商往往升级硬件和软件环境以提升系统的服务能力,但该方法的性价比往往不尽人意,如何降低时延不是单一技术能解决的,是各方共同努力的方向。

## 2 相关工作

目前没有专门的文章系统地讨论Web请求的响应时延以及降低时延的方法。文[1]从测量角度讨论了互连网中的物理时延、网络时延和端点的处理时延,以及估算这些时延的技术。文[2]把Web响应时间分为地址解析时间、连接建立时间、页面传输时间和连接拆除时间,并分析了代理缓存对它们的影响。文[3]研究了Web请求中域名服务(DNS)时延对服务选择的作用。文[4]对各种负载平衡技术进行了分类和讨论,大体上把负载平衡技术分为四类:基于客户的方法、基于DNS的方法、基于分派的方法和基于服务器的方法,这些技术从某种程度上可以用来减少Web请求的响应时延。

## 3 基于协议的时延

基于Web的协作系统从应用层对应用提供服务,其实现的标准协议是超文本传输协议(HTTP),此协议承载在TCP上,TCP通过IP在互联网中实现信息的存储和转发。基于协议的时延是从协议角度考虑系统提供服务时产生的时延,下面分析这种时延的组成。

### 3.1 时延的组成

基于协议的时延主要从客户角度和协议处理的角度出发,考虑时延的构成。按照RFC的要求,一个客户发出Web的请求,首先地址解析器完成服务器URL的IP解析,这种解析由DNS完成,接着客户端与服务器建立TCP连接,然后传输客户所需的信息,传输完毕后断开TCP连接。因此,一个HTTP请求的完成包括以下步骤:① 解析器(通过DNS)解析URL的IP;② TCP连接的建立;③ 基于Web的服务信息传输;④ TCP连接的拆除。相应地,Web请求的响应时延$T_{delay}$为:$T_{delay}=T_{DNS}+T_{connection}+T_{content}$,其中,$T_{DNS}$表示IP地址的解析时间,$T_{connection}$表示连接建立和拆除时间,$T_{content}$表示页面传输时间。

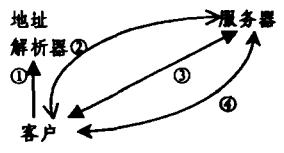


图1 Web请求的基本过程

### 3.2 DNS时延

互联网上的域名解析系统是全球的分布式数据库系统,负责所有URL名字到IP地址的解析和维护。文[5]表明,若

*)国家自然基金项目(90104005,66973034)。傅建明 博士,主要研究方向为高速信息网络与系统安全。

以包计，DNS流量大约为整个流量的8%，若以字节计，大约为4%。在Web请求的响应时延中，DNS的时延是不能忽略的。降低DNS流量，可以增加网络带宽，减少DNS时延。缓存技术通过缓存IP地址降低DNS时延，如图2所示。

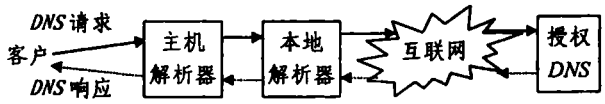


图2 地址解析过程

客户请求URL的某个页面时，由主机解析器解析，若解析失败，则由本地解析器解析。如果再解析失败，最后由授权DNS解析。如果所有的解析失败，则客户的请求无法提交。一般地，主机解析器驻留在客户的主机中，其TTL(Time To Live)由浏览器设定；本地解析器位于客户端的LAN中，其TTL由授权DNS设置。如果客户成功从授权DNS获得服务器的IP地址，则(IP，URL，TTL)三元组缓存在本地解析器和主机解析器中。这种IP地址的缓存可以由客户自己产生，也可以由其他客户产生[5]。如果多个客户请求同一个URL的地址，则一个成功的解析可以抑制其他相同的解析请求。文[3]表明，采用DNS缓存机制可以明显减少DNS时延，最多可以减少2个数量级。

缓存机制存在两个不足。首先是当提供服务的IP已变动或因超载导致此服务器的IP不可用，从而使得缓存的IP无效。此时客户发出URL的地址解析，客户得到一个无效的IP，用此IP建立TCP连接失败，最终导致客户重新解析URL的IP地址，其结果是反而增加了DNS时延。其次，缓存IP给服务器产生隐形加权负载，此时客户的请求没有经过服务器端授权DNS，使得基于DNS的负载平衡算法失效，引发服务器的负载不平衡。

### 3.3 连接时延

HTTP协议使用TCP进行传输，必须负责TCP连接的建立和拆除。如果页面有多个对象，则每一个对象使用一个连接，明显增加连接时延。对Web流量迹的分析表明[6]，一个网页平均有18个对象。近来的分析说明[7]，页面对象的数量不断增加，而对象的大小有递减趋势，从而连接时延所占的比例会越大。采用HTTP/1.1的永久连接，则传输一个页面的所有对象仅需一个连接，也就是说，只有一个连接的建立和拆除时间。文[7]分析结果显示有大约15%的网页通过永久连接进行传输。

文[8]从网络、CPU、磁盘系统等测量了HTTP/1.1和HTTP/1.0的性能。当服务延迟的瓶颈在网络时，使用永久连接优于HTTP/1.0；当瓶颈在CPU时，两种连接性能差异很小；当瓶颈在磁盘I/O时，使用永久连接的系统，其性能反而下降。

### 3.4 内容传输时延

内容传输时延是整个Web响应时延的主要组成部分。一般地，客户直接从远端的服务器获取请求页面。为了减少内容传输成本，客户端和服务端都可以使用缓存技术。客户端缓存内容可降低时延和服务成本。客户首先访问位于客户主机的主机缓存，如果命中失败，接着访问位于客户局域网的本地代理(如果存在)；如果再次失败，最后访问远端的Web服务器。缓存内容的TTL设置越长，则传输时延越短。这种内容缓存由客户本身或局域网内其他客户的访问产生。

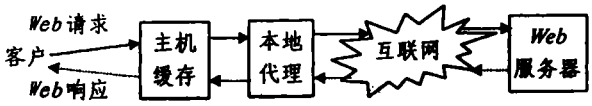


图3 Web页面的存取

服务器端既可缓存一些动态生成的页面或静态页面，也可预取一些页面。当客户访问页面时，可以减少动态页面的生成时间或磁盘读取时间，从而降低内容的传输时延。

页面缓存的不足是缓存的页面可能过期，导致客户访问不真实的信息。对一些时间敏感的数据，如天气预报、股票信息以及一些重要的商品价格信息等，尽量减少缓存的TTL时间，按照客户的需求针对不同的页面内容设置不同的TTL。

## 4 基于结构的时延

基于Web的协作系统在其网络拓扑结构上进行协作和服务，可以安装在局域网上，也可以安装在企业的虚拟网上或在广域网上。客户既可通过局域网访问服务，也可通过拨号上网访问服务。从网络拓扑看，Web请求的时延由客户端和服务端处理时延，以及网络中网元设备的处理时延和链路的传输时延，这种时延称为基于结构的时延。

### 4.1 时延的组成

客户通过局域网的路由器可以接入提供服务的网络，或由调制解调器(猫)拨号连接ISP，通过ISP获得所需的Web服务。一个Web请求从客户端发出，中间经过许多网元设备和物理线路，最终到达服务器端，其响应的路径正好相反，如图4所示。

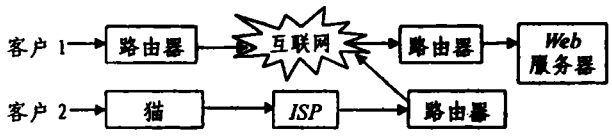


图4 Web访问的基本网络结构

根据Web访问的逻辑网络结构，Web请求的响应时延可以分割为以下几个部分[1]：① 客户端的处理时延；② 物理时延，由物理线路、路由器、交换机、调制解调器产生；③ 网络时延，数据包在路由器中排队和阻塞时产生的；④ 服务器端的处理时延。

### 4.2 客户端的处理时延

客户使用的主机，由于其CPU和内存等资源的不同，导致系统的处理能力不同。为了减少客户端的处理时延，一方面可以减少主机上的作业，提升主机的处理能力；另一方面，优化和升级主机的硬件或软件，从根本上增强系统的处理能力。前者限制了系统提供的服务；后者增加了客户的成本。客户端的处理时延还与TCP协议的具体实现有关。

### 4.3 物理时延和网络时延

物理时延和网络时延涉及的因素很多，主要是从客户端到服务器端的链路上的网元个数、物理链路的长度以及网络带宽。增加网络带宽，降低网络负载，提升网元的处理能力，可以减少网络中数据包的排队时间和阻塞时间，缩减网络时延；减少网元个数(或者跳数)，缩短物理链路长度，可以降低物理时延。在网络环境不变时，采用就近服务的技术，把服务器从

客户的远端复制到客户的近端，从而减少链路长度和网元个数。这种技术主要结合 DNS 实现，根据客户的 IP、域名、往返时间(RTT)以及其他启发式信息[9,10]，DNS 决定客户访问哪一个服务器。这种技术也可以在客户端实现，客户按照某种策略(跳数，RTT，频率)来分配 Web 请求。

就近技术可以较好地解决网络时延和物理时延。但如果服务器放置不当，或 Web 请求分配不当，则会导致 Web 请求响应时延的增加，引发服务器的负载不平衡。

### 4.4 服务器端的处理时延

服务器承载的服务较多，其 CPU 和内存等资源的有限使得并发服务的客户数有限。为了减少时延，可以关闭某些服务，优化和升级硬件或软件，增加服务器的吞吐量，如目前流行的应用服务器或功能服务器。这种时延与 TCP/HTTP 协议的实现紧密相关。除此之外，服务器簇技术可提升服务的处理能力，并获得较高的性能价格比。文[7]指出，其流量迹中35%的流量来源于服务器簇(Server-farm)。服务器簇技术的关键问题解决各服务器之间的负载平衡(特别是异构服务器簇)。文[4]对此问题进行了深入的探讨。

## 5 基于测量的协作分配模型

网络性能测量一直是研究的热点。测量结果反映网络的运行状况，并指导网络优化，提升网络性能。网络测量从单纯的测量，过渡到服务于各种应用系统，以反映系统的状况和优化系统的性能。基于测量的协作分配模型，以测量网络和服务器的负载为基础，根据协作策略，合理地分配客户的请求，旨在平衡负载和减少 Web 请求的响应时延。该模型的结构见图5所示。

Web 请求的响应时延既涉及到 HTTP 协议/DNS 协议以及底层网络协议的实现，也涉及到从客户到服务器的链路上的所有资源(客户端主机，服务器主机，路由器，交换机等)。另外，Web 的响应时延与提供服务的服务器负载紧密相关，重负载时其时延长，轻负载时其时延短。服务器超载时拒绝客户服务，此时整个系统应通过包重写技术或重定向技术把 Web 请求分配到其他服务器上。

### 5.1 模型的构件

该模型的测量构件为 agent 和 prober。agent 负责服务器工作负载的测量，测量单位时间内的请求数(requests)和系统的空闲资源(resources)，测量结果用二元组S(requests，resources)表示，S 表示系统的状态。prober 测量网络的吞吐量(throughtput)和网络的空闲资源(resources)，测量结果为二元组 N(throughtput，resources)，N 表示网络的状态。该模型的请求分配构件为 switch 和 proxy。switch 管理服务器端的

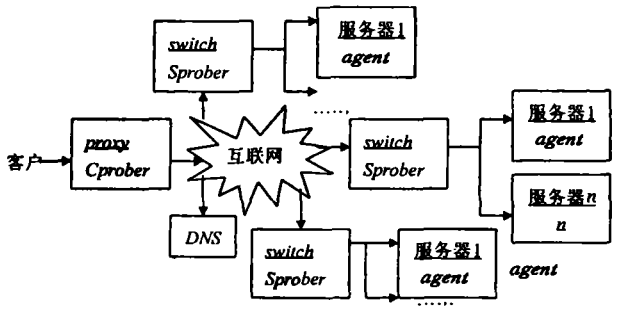


图5 Web 请求分配的协作分配模型

请求分配，而 proxy 在客户端附近，负责分配客户的请求。

### 5.2 构件的布置与通信

构件的布置按客户和系统的需要配置。一般，在客户端的局域网中放置一个 prober(记为 Cprober)和一个 proxy，Cprober 负责测量和计算客户的时延，监视链路的变化。proxy 负责页面的本地缓存，当缓存失败后，根据 Cprober 的测量结果把 Web 请求分配到适当的服务器。

各服务器簇按照地域或客户的访问量放置在适当的地方，所有服务器上的服务都是同步的和可用的。每一个服务器簇都有一个网络探测器 prober(记为 Sprober)和一个请求分配的 switch。Sprober 收集网络带宽的利用率以及簇内所有服务器的利用率，switch 负责此簇的所有请求分配。簇内每一个服务器都有一个性能监测代理 agent，agent 负责测量服务器的性能。一个服务器簇内的 Sprober 和簇内的所有 agent 构成一个组播组，agent 广播其本身服务器的状态，Sprober 和 agent 接受组内成员广播的信息。Sprober 把自己测得的信息和簇内利用率最低的服务器地址传递给 DNS。当 DNS 收到客户的地址解析请求时，DNS 根据已有的信息和某种策略对 URL 进行地址解析，实现客户的就近访问，以及平衡服务器和链路的负载。

agent 广播其本身服务器的状态可以是时间驱动，也可以是事件驱动。该模型采用事件驱动，减少组播次数。设置服务器状态的两个域值 $S_{min}$ 和 $S_{max}$。$S_{min}$ 表示服务器的空闲状态，此时服务器可以接受更多的请求；$S_{max}$ 表示服务器的忙碌状态，此时服务器不能接受新的请求。定义两个事件：

(1) Agent_Event_idle：当系统的状态 $S<S_{min}$ 时，agent 广播其服务器的状态(S，idle)；

(2) Agent_Event_busy：当系统的状态 $S>S_{max}$ 时，agent 广播其服务器的状态(S，busy)。

其中，idle 和 busy 为状态的标记。

SProber 发布网络状态可以是时间驱动，也可以事件驱动。同理，该模型采用事件驱动。设置网络状态的两个域值 $N_{min}$ 和 $N_{max}$。$N_{min}$ 表示网络的空闲状态，此时网络可以接受更多的请求；$N_{max}$ 表示网络的忙碌状态，此时网络不能接受新的请求。定义两个事件：

(1) Prober_Event_idle：当网络的状态 $N<N_{min}$ 时，Sprober 发送其网络的状态(N，idle)；

(2) Prober_Event_busy：当网络的状态 $N>N_{max}$ 时，Sprober 发送其网络的状态(S，busy)。

### 5.3 分配策略

switch 的分配策略依赖于服务器信息。switch 每收到一个 agent 的消息则更新 available-List，把状态为 idle 的服务器加入到 available_List 中。switch 收到一个 Web 请求，从 available_list 中使用轮循算法[3]把该请求分配给相应的服务器。

DNS 的分配策略取决于网络信息。DNS 每收到一个 Sprober 的消息则更新 available-List，把状态 idle 的网络节点加入到 available_List 中。DNS 收到一个服务器地址的解析请求后，根据客户的 IP 地址决定客户的地理位置，根据就近原则和 available_list 选择合适的 IP 进行解析。

**结论** Web 请求的时延可能成为基于 Web 的协作信息系统的瓶颈，分析产生时延原因，寻找解决此问题的办法。理解和刻画时延是分析时延的关键，本文从协议和结构两个方

(下转第138页)

到达一个管理域时，接受广播报文后，提取相应的参数并且进行存储。

注册：

通常情况下，当 MN 确定要进行一次切换时，MN 首先将自己计算出来的切换频率(在当前层次上的切换次数/发生这些切换的时间)与 $P_{max}$ 和 $P_{min}$ 进行比较，如果大于则向高一层的 FA 注册，如果小于，则向低一层的 FA 注册，否则向同一层的 FA 注册。

FA 收到注册请求后，还要进行 MN 速率的判断，此时只判断 MN 的速率是否大于 $V_{max}$，是则将注册请求转发给更高一级的 FA，否则响应注册请求 。

在较高级的 FA 中，对应于每个注册在其上的 MN 均设置一个最低速率计时器，只要 MN 的速率降低到 $V_{min}$ 以下，计时器开启，如果 MN 的速率回升到 $V_{min}$ 之上，则计时器归零。当计时器达到设定值时，可以根据当前 FA 的系统容量占用情况决定是否把 MN 的注册转移到更低一级的 FA 上去。

路由更改：

如果切换发生在同一个 FA 中的不同基站之间，只需要在该 FA 的路由表中更改指向的基站地址即可。

如果切换发生在不同的 FA 之间(包括同一层次的和不同层次的 FA)，那么在新的 FA 到交叉节点之间的所有节点上均要增加路由记录，在旧的 FA 到交叉节点之间的所有节点均要删除原来的路由记录，在交叉节点上更改路由记录。

**结论** 在上面的描述中，我们介绍了关于在多层小区拓扑上实现移动 IP 的移动性管理的一种方案。由于该方法的引入，可以把高速移动的 MN 或者切换频繁的 MN 注册到宏 FA，从而明显地降低 MN 的切换次数，从而减少大量的网络控制报文的流量，与此同时，对于低速的或者办公室环境下的 MN，则注册到微 FA 或者微微 FA 上，在保证较低切换次数的基础上，也尽量地增加系统可接入的 MN 的数量，同时层次化的 FA 拓扑结构，也保证了在切换时的路由更改非常简单易行。

在本方法中，引入了两个参数对移动用户的类型进行分类，速率参数显然是必须的，而切换频率参数则有待在以后的研究过程中进行详细深入的论证后决定是否可以舍弃或者使用替代的参数。

本文提出了有关移动 IP 移动性管理策略方面的改进思路，在以后的研究中，还需要从以下几个方面进行更深入的工作：1. 对于实现该方案所需要的控制报文格式的改进；2. 该方案在减少切换次数、减少报文流量等方面的定量分析；3. 移动通信网络中的不同切换方式对移动性管理性能的影响和优化；4. 与此相关的网络仿真工作。

## 参 考 文 献


1 Perkins C. IP Mobility Support. Network Working Group, Oct. 1996. RFC 2002

2 Caceres R, Padmanabhan V N. Fast and Scalable Handoffs for Wireless Internetworks. In: Proc. of ACM MobiCom ' 96, Nov. 1996

3 Chen W, Lin E, Wei H. Dynamic location control for mobile nodes: [Technical Report 97-CSE-10]. Department of Computer Science and Engineering, Southern Methodist University, 1997

4 Yavatkar R, Pendarakis D, Guerin R. A Framework for Policy-based Admission Control. RFC2753; Jan. 2000

5 Andr'as G. Valk'o. Cellular IP: A New Approach to Internet Host Mobility. Computer Communication Review, 1999, 29(1): 50～65

6 Campbell A T, et al. Design, Implementation, and Evaluation of Cellular IP. IEEE Personal Communications, 2000, 7(4): 42～49

7 Subir D, Archan M, Prathima A. TeleMIP: Telecommunication-Enhanced Mobile IP Architecture for Fast Intradomain Mobility [J]. IEEE Personal Communications, 2000, 7(4): 50～58

8 Ramjee R, LaPorta T, Thuel S. IP micro-mobility support using HAWAII. http://www.bell-labs.com/user/ramjee/papers/draft-ietf-mobileip-hawaii-01.txt

9 Ramjee R, et al. HAWAII: a domain-based approach for supporting mobility in wide-area wireless networks. In: Proc. of the 7 th Intl. Conf. on Network Protocols (ICNP) 1999. 283～292

10 Chih-Lin I, Greenstein L J, Gitlin R D. A Micrecell/Macrocell Celluar Architecture for Low-and High-Mobility Wireless Users. IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, 1993, 11(6): 885～890

11 Hu L, Rappaport S S, Personal communication systems using multiple hierarchical cellular overlays. IEEE Journal on Selected Areas in Communications, 1995, 13(2): 406～415

12 Coombs R, Steele R. Introducing microcells into macrocellular networks: A case study. IEEE Transactions on Communications, 1999, 47(4): 568～576


(上接第131页)

面探讨了时延的产生和组成。在实际 Web 应用系统中，Web 流量的实时监测结果为客户选择服务提供商和服务提供商优化网络服务提供充分的支持。本文提出的基于测量的协作分配模型，从负载均衡的角度解决 Web 请求的时延问题，该模型首先从 DNS 进行 Web 请求的预分配，然后 switch 进行再分配，同时考虑了客户端的请求分配。将来重点研究根据客户的注册信息实现 Web-QOS 的分配调度算法。

## 参 考 文 献


1 Martin H S, McGregor A, Cleary J G. Analysis of Internet Delay Times, PAM'2000

2 Liu B, Abdulla G, Johnson T, Fox E A. Web Response Time and Proxy Caching. WebNet98, Orlando, Nov. 1998

3 Shaikh A, Tewari R, Agrawal M. On the Effectiveness of DNS-based Server Selection . In: Proc. of IEEE INFOCOM, 2001

4 Cardellini V, Colajanni M, Yu P S. Dynamic Load Balancing on Web-Server Systems. IEEE Computer, 1999, 3(3): 28～39

5 Danzig, Obraczka K, Kumar A. An Analysis of Wide-Area Name Server Traffic. ACM Comp. Commun. Review (SIGCOMM'92)

6 Mikhail Mikhailov Craig E. Wills Embedded Objects in Web Pages. WPI-CS-TR-00-05

7 Smith F D, Campos F H, Jeffay K, Ott D. What TCP/IP Protocol Headers Can Tell Us About the Web. SIGMETRICS/Performance'2001

8 Youn J. A Performance Evaluation of Hyper Text Transfer Protocols in Wide Area Network . http://cs-people.bu.edu/jisook/firstlayer/Wan_nistnet.ppt

9 Venkata N, Padmanabhan, Subramanian L. An Investigation of Geographic Mapping Techniques for Internet Hosts. SIGCOMM'2001

10 Govindan R, Tangmunarunkit H. Heuristics for Internet Map Discovery. IEEE INFOCOM 2000