

业务选择网关的连接模型研究

Service Selection Gateway Connecting Model Research

(西安电子科技大学)丁振国 张恒新
DING ZHENGUO ZHANG HENGXIN

摘要:业务选择网关是随着 Internet 和网络个性化服务的发展应运而生的。本文通过对基于请求内容分配机制的主要技术之一 TCP Handoff 技术的分析介绍,提出了一种不同于传统的业务选择网关请求连接和实现的模型。利用 TCP handoff 技术能将客户请求的响应结果直接返回客户端,消除了由业务选择网关节点而引起的网络通信瓶颈,从而可以显著提高整个服务系统的吞吐率,使整体性能得到极大的改善。

关键词:业务选择网关;TCP Handoff;Linux

中图分类号:TP391.05 文献标识码:A

Abstract:With the development of the Internet and the individuation network service, the Service Selection Gateway comes into being. By introducing and analysing TCP Handoff technology of content-aware scheduling, this paper presents a new Service Selection Gateway Connecting Model, in which a new technology such as the TCP Handoff is adopted. In this paper, by using the TCP Handoff technology, the Client requesting responses return directly to the Clients, the communication bottleneck of the front Service Selection Gateway can be resolved. Thus, the whole system performance can be improved remarkably.

key words:Service Selection Gateway,TCP Handoff,Linux

1 引言

随着 Internet 在世界范围内的迅速发展和网络宽带化进程的加快,人们对电信服务商提供服务的要求日趋多样性和个性化,这些服务包括:在线游戏、IP 电话、视频会议、视频点播、Web 浏览、电子商务和远程教育等等。业务选择网关 SSG(Service Selection Gateway)作为网络内容服务提供商的控制网关,通过为客户提供个人服务的门户 Web 页面的方式为客户提供个性化服务的菜单,从而直接满足客户的个性化需求。

西方国家对此研究起步比较早,其功能最早开始时是在 Cisco IOS 软件中为 Cisco 的 6400 平台引入的。目前,已经出现了一些提供个性化网络服务的产品,其中具有代表性的是 Cisco 公司的业务选择网关(SSG)。SSG 作为 Cisco IOS 软件的一个模块化特性集,可以为针对不同 IP 目的地址的客户交互策略提供 RADIUS 身份认证和记账功能,包括同时登录多项服务的能力等。在我国由于这方面的研究起步比较晚,所以现在还没有成熟的产品。一些较大的网络服务提供商 ISP(如上海网通等)都在使用 Cisco 和北电等国外较大设备提供商的产品。现在还没有为内容提供商设计的基于内容定制的个性化网络服务产品。

在基于页面选择的业务选择网关业务模型中,针对客户请求采用基于请求内容的调度方式,文献的研究表明可大大提高系统的整体性能。传统 TCP 网关的基于请求内容调度实现方式,其主要思想是运行一个基于应用层的调度程序,由它负责

在客户和后端服务器之间转发请求和响应。这种技术的缺陷是显而易见的:由于网关作为数据中转站,它的数据处理能力将是系统性能的瓶颈。特别是后端服务器将处理请求的结果数据均通过网关发送给用户时,而这些结果中往往包含大量的数据(例如:文件数据信息或语音、图形、图像、视频、动画等类型的多媒体信息)时,这势必占用网关系统中大量的处理能力和带宽,很容易造成网关过载而难以负载。目前有两种较新的实现方式采用的是 TCP 粘合(TCP Splicing)和 TCP 迁移(TCP Handoff)技术。TCP 粘合的实现思想与传统的基于请求内容调度的实现方法类似,区别只是在操作系统内核中实现。由于避免了数据报文在上下文切换时的拷贝,减少了核心空间到用户进程的通信开销,整体效率相比较传统方法是要好一些。

TCP Handoff(TCP 迁移)作为一项新型技术,它的提出并在 LVS(Linux Virtual Server)Linux 服务器集群系统中应用,收到了较好的效果。其实现基本思想是根据客户的请求内容,将客户与基于页面选择的业务选择网关的 TCP 连接对客户透明地迁移到后端服务器上。后端服务器的响应直接发给客户,不再经过基于页面选择的业务选择网关,既有效地减少了客户端的响应延迟,又大大降低了基于页面选择的业务选择网关的负载,从而有效地为服务提供商提高了整个系统的性能。

目前,国际上只有少数几个公司(Resonate, IBM)和大学实现了或正在研究 TCP Handoff 技术,国内的研究也已经开始。本文通过对基于请求内容分配机制的主要技术之一 TCP Handoff 技术的分析介绍,提出了一种不同于传统的业务选择网关请求连接分配和实现的模型。

2 业务选择网关(SSG)的业务模型

以 Cisco 公司的业务选择网关为原型进行讨论。SSG 一般可以提供两种实现或业务模型:第二层业务选择或第三层业务

丁振国:教授

基金项目:国家 863 基金项目,“基于国产 LINUX 的网络教育教学质量国家评估系统与网络教育机构教学质量自我评价系统”(2004AA1Z2520)

选择模型。在第二层业务选择与第三层业务选择之间的主要区别就是：在第二层业务选择中，用户在发起一条点到点协议(Point-to-Point Protocol, PPP) 会话之前就已经确定了他的最终目标，并且在该 PPP 会话处于活跃的整个过程中，最终目标是保持相同；在第三层业务选择中，用户在进入了 NAP 或 NSP 的“带有围墙的花园(walled garden)”(本地网络)后，它可以有多种业务(最终目标)供其选择，然后用户可以选择他需要连接的具体业务，而不必断开 PPP 会话(在第二层业务选择中就需要断开原来的 PPP 会话)。

第三层业务选择功能使得 SSG 对于任何从事零售业务的 SP 而言,成为一种必备的特征,并于接入方式无关:不论是 xDSL、电缆、还是无线方式。第三层业务选择通常也被称为基于 Web 页面的业务选择(Web-Based Service Selection)。

本文主要讨论 Linux 操作系统下，基于页面业务选择使用一个 HTTP 服务器或 Web 服务器的业务选择网关。它也被称之为业务选择控制板(Service Selection Dashboard, SSD)。通过基于页面的业务选择方式，服务提供商可以向用户提供一组基于页面的菜单，菜单内列出了用户能够选择的业务。用户在唯一的用户名和密码登录之后，就可以看到这个菜单。在用户登录后，他可以选择任何已经订购的业务，并访问之。如图 1-1 给出了基于页面选择的业务选择网关的典型应用。

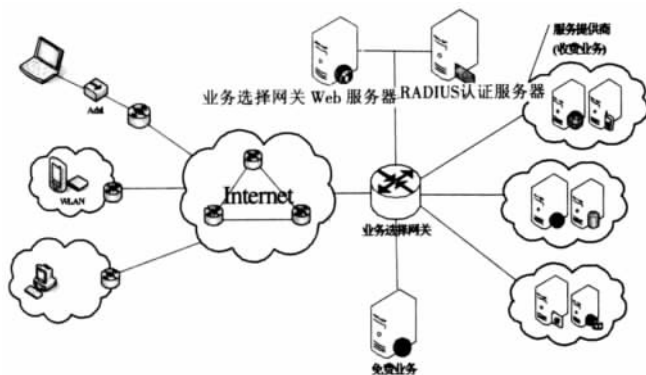


图 1 基于页面选择的业务选择网关的典型应用

针对基于页面选择的业务选择网关的典型应用,可以看出客户对服务的请求是通过业务选择控制面板即 Web 服务器的 HTTP 请求完成的。HTTP 协议是基于连接的协议,即客户请求时必须通过业务选择网关与后端服务器建立连接才能获得 HTTP 请求的内容。在基于请求内容的会话调度中采用 TCP Handoff 技术,调度的依据不仅有客户的 IP 地址,而且还考虑客户的请求内容,从而可以充分利用访问服务资源存放在不同的后端服务器上的特点,大大提高整个访问的性能而且,这样有利于扩充服务资源。

3 TCP Handoff 技术分析

TCP Handoff(TCP 迁移)技术,其主要思想是根据客户请求的内容,将客户与基于页面选择的业务选择网关的 TCP 连接对客户透明地迁移到后端的内容服务器上。后端服务器的响应直接发给客户,不再经过业务选择网关。

从目前国际上 TCP Handoff 的实现技术来看, 均需要修改操作系统协议栈, 有的甚至要重写服务程序, 并且源代码和协议内容均不公开。然而, 在基于 Linux 内核 2.4 操作系统的业务选择网关和后端服务器模型之下, 利用 Linux 内核 2.4 中新开发的数据包过滤、数据包处理、NAT 等功能框架 Netfilter 技术,

可以在操作系统协议栈处理过程中截获报文,并对通过其进行相应处理来实现 TCP Handoff 技术。

3.1 技术实现思想

利用 Linux 操作系统对 Netfilter 功能框架的支持，在前端的业务选择网关和后端内容服务器上，分别挂接 IP 报文遍历的 IP_LOCAL_IN 链和 IP_LOCAL_OUT 链，用于截取、改写 TCP 报文。在前端业务选择网关上，主要是记录连接、迁移信息，分析 HTTP 请求内容，转发报文给后端服务器。在后端服务器上，则主要是改写序列号，丢弃某些报文，通过这些方法来实现 TCP Handoff。

3.2 TCP handoff 工作流程

基于页面选择的业务选择网关的典型应用,可以看出客户通过页面对服务的请求是通过业务选择控制面板即 Web 服务器的 HTTP 请求完成的。HTTP 协议是基于连接的协议,即客户请求时必须通过业务选择网关与后端服务器建立连接才能获得 HTTP 请求的内容。因此,在基于请求内容调度下采用 TCP Handoff 技术系统的工作流程如图 1-2 所示。

首先,客户端和业务选择网关使用 TCP/IP 协议,通过正常的 3 次握手过程建立起一条 TCP 连接,如图中前 3 个报文序列。接着,客户方发出请求内容,该报文在在基于页面选择的业务选择网关中分析并改写,然后转发给到后端的服务器上。在后端的服务器上,连接被重建,响应报文直接发送给客户方。该连接上的后续报文由基于页面选择的业务选择网关在 IP 层直接转发给后端的服务器,服务器发出的响应报文则直接发送给客户方。最后,报文发送结束,终止转接。

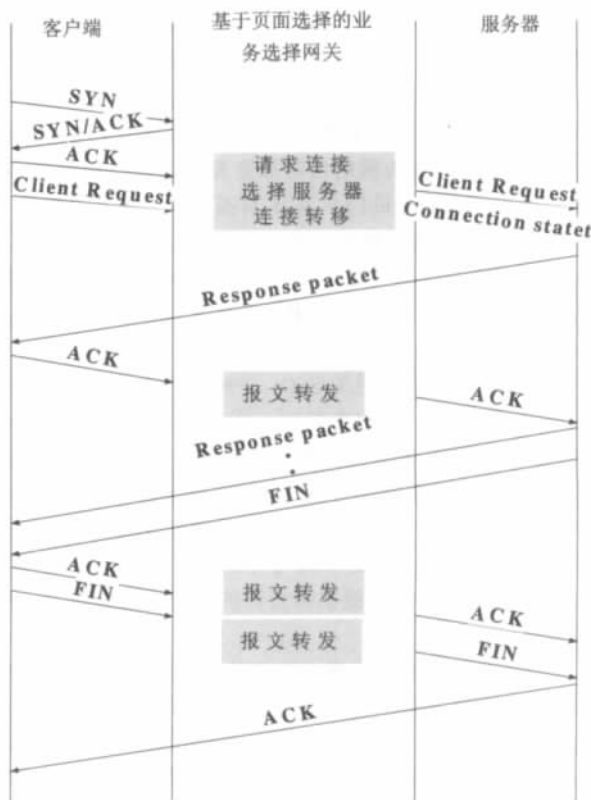


图 1-2 TCP Handoff 技术系统的工作流程

4 小结

随着 Internet 的飞速发展, 网络服务进一步细分并向个性

化服务的方向发展,加之服务提供商对网络增值服务利益的不断追求,必将带动业务选择网关的新发展。将有更多的新技术应用于业务选择网关中研究和实现中,本文介绍的 TCPHandoff 技术具有技术新颖,实现效果好,对客户响应透明等优点。但是, TCP Handoff 技术的实现涉及到修改操作系统的内核,所以较为复杂,还有待继续改进和研究。

本文作者创新点:将 TCP Handoff(TCP 迁移)这一新型技术,引入业务选择网关连接模型中,将客户与基于页面选择的业务选择网关的 TCP 连接对客户透明地迁移到后端服务器上。使后端服务器的响应直接发给客户,不再经过基于页面选择的业务选择网关,从而有效地为服务提供商提高了整个系统的性能。

参考文献

[1]梁润秋,丁振国. 基于 SSG 构建个性化宽带服务平台[J]. 天津通信技术. 2003, 12. NO.4. PP.26~28.

[2]Sanjeev Mervana,Chris Le[M],刘东昱,李逢天[译]. 设计与实现基于 DSL 的接入方案. 人民邮电出版社. 2002, 09. PP.138~157.

[3]张毅,董云艳 基于 Linux 的 TCP/IP 协议栈安全性研究[J]微计算机信息 2005 23

[4]Linux 服务器集群系统 LVS 简介 [EB/OL]http://tech.ccidnet.com/art/737/20050609/264501_3.html

[5]Aron M, Sanders D, Druschel P et al. Scalable Content-aware Request Distribution in Cluster-based Network Servers [C]. Proceedings of the USENIX 2000 Annual Technical Conference. 2000-06

[6]Pai V S, Aron M, Banga G, et al. Locality-aware Request Distribution in Cluster-based Network Servers[J]. ACM PLAN Notices, 1998,33(11): 205-216

作者简介:丁振国, (1959-), 男, (汉族), 陕西三原人, 教授, 西安电子科技大学网络教育学院副院长, 国家信息安全研究中心关键技术实验室总工程师, 中国工程技术信息网副总工程师, 主要研究方向为计算机网络与信息处理; 张恒新(1977-), 男, (汉族), 新疆奇台人, 助工, 在读研究生, 主要研究方向为计算机网络与信息处理。

Biography:Ding Zhenguo, (1959-), male, (the Han Nationality), the Shaanxi Sanyuan person, professor, Xidian University Network Education Institute vice-president, country information security research center key technologies laboratory chief engineer, Chinese project technology information network vice-chief engineer, mainly studies the direction for the computer network and the information processing; Zhang Hengxin (1977-), male, (the Han Nationality), the Xinjiang Qitai person, the assistant engineer, Xidian University Computer College master the graduate student, mainly studies the direction for the computer network and the information processing

(710071 西安 西安电子科技大学网络教育学院)丁振国

(710071 西安 西安电子科技大学计算机学院)张恒新

(School of Network Education Xidian Univ, Xi 'an 710071) Ding ZhenGuo

(School of Computer Science and Technology Xidian Univ., Xi 'an 710071)Zhang HengXin

通讯地址:(710071 西安 西安电子科技大学 396 信箱)张恒新

(收稿日期:2007.5.03)(修稿日期:2007.6.05)

(上接第 180 页)

单节点吞吐量只能达到最大值的 70%左右;而平均响应时间上升不超过原请求服务时间的 2 倍。第三, 系统在处理慢速请求源时, 每个节点的平均吞吐量下降不明显, 而平均响应时间上升剧烈。单节点吞吐量能达到最大值的 90%以上;而平均响应时间上升超过原请求服务时间的 9 倍以上。另外, 我们还对该算法进行了仿真实验, 实验结果与理论分析的结果大致相同。

5 结论

本文研究了机群系统下有状态应用的请求调度策略。通过理论分析和实验仿真, 指出了使用轮循算法系统性能的主要取决于请求到达速率与请求服务速率之比, 而受系统中并发会话状态数和应用服务器数的影响较小。研究表明, 该法在快速与慢速请求用户下, 会有不同的性能表现, 因此, 在实用过程中应根据服务要求来确定是否能使用该算法。另外, 本文提出的抽象模型对有状态应用中其他调度策略的性能分析也有借鉴意义。

本文作者创新点:通过理论分析和实验仿真, 指出了使用轮循算法下影响系统性能的主要因素, 并给出了系统在不同情况下的性能表现。利用本实验结果, 可以对项目中系统的性能进行预估计, 从而可以极大地减少设计和实现的成本。另外, 在理论上, 本文提出的抽象模型对有状态应用中其他调度策略的性能分析也有借鉴意义。

参考文献

[1]D. Jacobs. Distributed computing with BEAWebLogic server. In Proc. Conference on Innovative Data Systems Research, Asilomar, CA, 2003.

[2]Sun. iplanet server. <http://www.sun.com/>

[3]Microsoft. Internet information services (iis). <http://www.microsoft.com/iis>

[4]徐雪霖. Web 数据库访问技术探析.[J]微计算机信息 2004, No.2: 110-112.

[5]马征, 李建华. 基于多代理技术的分布式 Web 日志挖掘系统.[J]微计算机信息 2004, No.2:113-114,83.

[6]Debra VanderMeer,Helen Thomas, et.al. ReDAL: Request Distribution for the Application Layer. Proceedings of the 25th IEEE International Conference on Distributed Computing Systems, 2005.

作者简介:陈研(1977-), 男(汉), 上海交通大学, 博士生, 主要从事并行系统设计和分析的工作, 多次以访问学生的身份参与爱立信公司的项目合作, 并已在多个国际会议和期刊发表论文; 徐国治, 男(汉), 上海交通大学, 博导, 研究方向:并行系统设计和分析等。

Biography:Chen Yan, male, Shanghai Jiaotong University, Ph.D student, research on parallel system design and analysis. He has worked as visiting student in Ericsson for times, and has published multiple papers in various international conferences and journals;Xu Guozhi, male, Shanghai Jiaotong University, professor, research on parallel system design and analysis, etc.

(200240 上海 上海交通大学电子工程系)陈研 徐国治

(Shanghai Jiaotong University, Shanghai, 200240)Chen Yan Xu Guo Zhi

通讯地址:(200240 上海 上海交通大学电子工程系)陈研

(收稿日期:2007.6.03)(修稿日期:2007.7.05)