

# 一种基于 FPGA 设计的本地 DNS 服务器\*

李沛杰, 张兴明, 沈剑良

(国家数字交换系统工程技术研究中心, 郑州 450002)

**摘要:** 为了解决高负载情况下本地通用 DNS 服务器并发处理能力差的问题,通过分析 DNS 请求与响应的流处理特征,基于 FPGA 实现了一种硬件定制的大容量高并发 DNS 服务器。在 DNS 请求处理过程中,采用边解析边响应的并行处理方式,极大地提升了 DNS 服务器的本地解析能力。与通用 DNS 服务器相比,性能提升了 10 倍左右,能够适应高负载情况下的并发处理需求。

**关键词:** DNS 服务器; 域名解析; 流处理; 并发处理

**中图分类号:** TP302      **文献标志码:** A      **文章编号:** 1001-3695(2014)04-1102-03

**doi:**10.3969/j.issn.1001-3695.2014.04.035

## Design of local DNS server based on FPGA

LI Pei-jie, ZHANG Xing-ming, SHEN Jian-liang

(National Digital Switching System Engineering & Technology R&D Center, Zhengzhou 450002, China)

**Abstract:** The local general DNS server had poor concurrent processing capability in the condition of high load. This paper was aimed to improve the above-mentioned. By analyzing the streaming processing characteristics of the DNS request and response, it proposed a FPGA-based local DNS server with large capacity and high concurrency. It parsed while responded when processing the DNS requests. This DNS server could extremely promote the local analytical ability. Compared with the general DNS server, about 10 times performance was improved. This DNS server can meet the demands of concurrent processing when high load occurred.

**Key words:** DNS server; domain name analyze; streaming process; concurrent process

由于网络信息量的增大以及用户访问量的快速增长,简单提升服务器的性能并不能解决其对 Web 服务器的压力。域名系统 DNS(domain name system)以分布式数据库的形式用于 TCP/IP,主要实现服务器 IP 与访问域名之间的对应<sup>[1]</sup>。DNS 服务器基于某种负载均衡策略,能够将所有的请求分配给 Web 服务器集群中的所有服务器,其性能主要由响应速度决定。

对于本地 DNS 服务器,当接收到的 DNS 请求不在缓存中时,会将该请求转发给最近的根 DNS 服务器,并对根 DNS 服务器应答的域名和对应 IP 进行缓存,显然响应延时越短,其性能越好;而对于局域网内部的 Web 集群,本地 DNS 服务器只需记录该集群中的 Web 服务器 IP,当有 DNS 请求时,将这些 IP 依某种策略响应给用户。因此,当系统负载过高时,本地 DNS 服务器的并发处理能力将成为整个系统的瓶颈。DNS 的处理建立在数据包交换的基础上,因此数据包的处理是 DNS 服务器的基础。DNS 服务器检查到达的每一个 DNS 请求包,对包进行一系列的操作,然后将响应回送出去。由于 DNS 请求包及响应包均为短包(80 Byte 和 100 Byte),因此提升 DNS 服务器并发处理能力的途径是采用流水线结构和并行运算。

现场可编程门阵列(field programmable gate array, FPGA)具有高速、高可靠性、开发周期短的特点,且其硬件编程语言(hardware description language, HDL)具有并行处理和流水线结构的特点。MicroBlaze 是针对 Xilinx FPGA 器件而优化的功能强大的 32 bit 微处理器,具有 3/5 级流水线、高速缓存、异常

处理、调试逻辑等根据自主性能需求设计的性能特性。

针对 DNS 请求与响应的流程以及 FPGA 及 HDL 的开发特点,本文基于 FPGA 实现了一种支持大容量高并发的本地 DNS 服务器。

## 1 相关研究

传统的 DNS 服务器对于客户端的 DNS 响应采用即时查表响应的方式,但是即使在 ISP(internet service provider)网络中的根节点 DNS 服务器中,当访问负载增大时:a)由于处理不及造成的 DNS 暂存将增大 DNS 的响应延迟;b)当请求过大时,极易造成 DNS 服务器的分布式拒绝服务。因此 DNS 服务器并发处理能力的高低关系到 Web 服务器集群系统的性能及安全。对于 DNS 服务器性能的研究,文献[2]研究了缓存 DNS 请求的有效性;文献[3]采用自回归模型来提升 DNS 服务器的处理速度;文献[4]采用非权威的 DNS 服务器对 DNS 请求进行权威响应,从而减少 DNS 响应时长,提升处理能力。上述文献均是在低负载情况下从域名更新的角度改善 DNS 的响应速度,但均没有研究 DNS 并发处理能力对响应速度的影响。基于 Linux 系统配置的通用 DNS 服务器<sup>[5,6]</sup>均在操作系统上进行配置,基于 DHT 表进行本地解析,对于每个 DNS 请求,在判断出其请求的域名之后均需在 DHT 表中进行查询匹配,当本地负载增大时,由此造成的响应延迟将严重影响 DNS 的并发处理性能;文献[7]对突发流量下本地 DNS 服务器解析性能的

**收稿日期:** 2013-07-13; **修回日期:** 2013-08-18      **基金项目:** 国家“863”计划资助项目(2009AA012201);上海市科委重大科技攻关项目(08dz501600)

**作者简介:** 李沛杰(1990-),男,山西襄汾人,硕士研究生,主要研究方向为 Web 服务安全、高性能计算(lpj0724@163.com);张兴明(1963-),男,教授,主要研究方向为宽带信息网络、SoC 设计等;沈剑良(1982-),男,讲师,主要研究方向为嵌入式系统与 SoC、可重构计算等。

测试显示出通用 DNS 服务器性能在高负载情况下的瓶颈。以上几种方式都是采用软件配置的方式在通用服务器实现本地 DNS 服务器,由于采用通用的处理器实现方式,因而 DNS 响应延时及其性能提升空间不大。

本文针对新概念高效能计算机体系结构及系统研究开发设计的一种基于 FPGA 实现的本地 DNS 服务器,依据数据包处理的流特性,并克服通用 DNS 服务器高负载情况下处理不及时的问题,采用硬件定制的方式实现了一种基于流处理的具有本地解析能力的大容量高并发 DNS 服务器。

## 2 大容量高并发本地 DNS 服务器的 FPGA 模型

本文针对 DNS 请求与响应的流程,着重研究 DNS 服务器接收和响应的流水线处理,基于 FPGA 实现了一种支持大容量高并发的本地 DNS 服务器。其实现框图如图 1 所示。

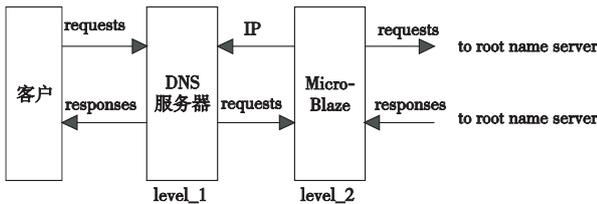


图1 DNS服务器实现框图

在图 1 所示的实现框图中, MicroBlaze 对外自留接口, 用于转发 DNS 服务器不存在的域名给根域名服务器和接收来自根域名服务器的解析信息。对于这种情况, 本文不作过多讨论, 而着重研究缓存中存在请求域名时, 本地 DNS 服务器对高并发 DNS 请求接收和响应的流水线处理。

要想实现大容量、高并发的 DNS 处理, 必须对各个功能模块进行并行流水处理。由于 DNS 的处理建立在数据包处理的基础上, 因此针对数据包的处理流程划分及构建功能模块显得十分重要。鉴于此, 本文将整个 DNS 服务器分为三部分, 即 DNS 请求接收、DNS 查找以及 DNS 组包响应。其中 DNS 查找采用文献[4]所述方式实现, DNS 请求接收及组包响应则采用 FPGA 实现, 重点则用于后者的研究。图 2 所示为本地 DNS 服务器的处理框图。

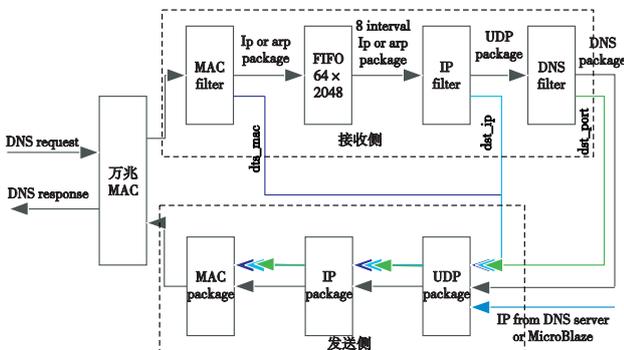


图2 DNS服务器的FPGA实现框图

### 2.1 MicroBlaze (MB) 的处理机制

嵌入式的 MicroBlaze 实现两种功能:a)统计本地 Web 服务器 IP 信息,缓存在 level-2 中的数据库中,并以某种策略将这些 IP 交付 DNS 服务器;b)用于实现文献[4]中的方法,解决 DNS 的请求延时问题。如图 1 所示,当用户请求的域名存储在 level-1 中 DNS 服务器的数据库中时, DNS 服务器直接对客户请求作出响应。若本地局域网内部 Web 服务器出现变动时, MicroBlaze 更新 level-2 中的数据库信息,并告知 level-1 中的 DNS 服务器更新其内部数据库;若用户所请求的域名及其对

应的 IP 地址不在 level-1 中时, MicroBlaze 转发 DNS 服务器交付的 DNS 请求给最近的上级根 DNS 名字服务器 (name server), 等待 name server 查询并接收其应答的 IP。

### 2.2 DNS 服务器的设计

#### 2.2.1 流水线设计思想

从图 2 中所示的 DNS 请求相应处理过程中可以看出, DNS 服务器主要由两大部分组成, 即请求侧和响应侧, 请求侧用于处理 DNS 服务器接收的 DNS 请求, 其主要工作为过滤解析 DNS 请求包; 响应侧主要用于对解析好的 DNS 请求作出应答, 其主要工作为组合 DNS 响应包。DNS 请求与响应的流处理特征决定了各个功能模块只有均做到并行处理才能保证大容量高并发 DNS 请求的实时处理。

通用 DNS 服务器对用户 DNS 请求有如下处理:a)对 DNS 请求提供缓存队列, 在一个包处理完成之后再行另一个包的处理;b)在解析 DNS 请求后, 对于 DNS 响应用到的信息作暂存处理, 组合 DNS 响应包时查找组合。

这样当 DNS 请求增大时, 极易造成缓存队列的溢出及流处理的中断。本文通过改善通用 DNS 服务器的处理机制, 采用 FPGA 流水线处理的特性, 关键对 DNS 请求及响应的数据流处理过程进行如下改进:a)对于 DNS 请求不作请求缓存队列, 这样只要保证硬件处理的流水性, 便能及时处理用户请求;b)对于解析的 DNS 源信息不作寄存, 直接作为 DNS 响应的目的信息, 随响应包即时处理;c)对于流水处理过程中耗时最多的模块进行边处理边接收或者发送的处理机制, 以此最大程度地节省处理周期, 提升处理速度。

#### 2.2.2 具体问题处理

对于上述流水线设计思想, 其 FPGA 的具体实现过程中分别对其改进进行实现, 其中具有如下关键问题需要解决:

a)由于 DNS 请求包与响应包均为短包, 接收端的瓶颈在于 IP 头的校验, 发送端的瓶颈在于 UDP 包的检验和计算, 在万兆网络中 (156 MHz 时钟, 64 bit 数据), 前者需要 4 个周期, 若 DNS 请求包间隔为 5 个时钟周期, 则后者需要 21 个周期, 因此假如按照原有的计算校验和之后再传输数据, 势必需要将 DNS 请求包以及需要作校验和的 DNS 响应包进行缓存。

b)DNS filter 模块一方面需要对 IP 过滤好的 UDP 包进行 DNS 解析, 另一方面需要对 DNS 请求进行响应, 而接收侧的 DNS 请求的源信息 (源 MAC、源 IP、源 PORT), 作为 DNS 响应的目的信息至关重要, 但是若暗处阿通的方式将这些信息进行缓存, 将同样会造成流水处理的瓶颈。

c)用户访问的突发性以及不均匀性会造成不同用户 DNS 请求的间隔很短 (最短 1 周期), 而对万兆 MAC 接收后的数据包的处理很难在极短的时间间隔内进行完整数据包的处理, 这样会造成 DNS 请求包的丢弃。

针对以上情况, 本文在 DNS 设计中有如下处理:

a)在 UDP package 模块中只加入一个短 FIFO, 用于暂存计算好的校验和, 而在最终发送数据包时, 计数到发送校验和时, 将其读出, 并插入, 从而保障数据包的流水线处理。其仿真结果如图 3 所示。图 3 所示最左标签处为计算好的校验并写入到短 FIFO 中, 中间标签所指的 UDP\_IPPSO\_nochecksum\_data\_i[66:0]处以及之前并不会插入计算好的校验和, 最右标签所指的 UDP\_IPPSO\_checksum\_data\_o[66:0]处及之后便插入了计算好的校验和。

b)接收侧接收到的源信息不作缓存, 直接将其作为目的

信息,随 DNS 响应包一同流水处理。其仿真结果如图 4 所示。

如图 4 所示,左 2 标签处从 DNS 请求包中解析到的源信息作为 DNS 响应包的目的地信息(dst\_mac\_addr[47:0]、dst\_ip\_addr[31:0]、dst\_port\_num[15:0])在左 2 标签后的各个标签处一直随包处理。

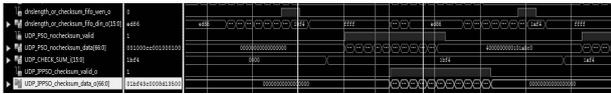


图3 检验和的处理



图4 源信息的处理

c)由图 5 所示仿真图可知,当 DNS filter 模块的输入包间隔在两个周期以下时,将会造成数据包的丢失,因此在接收侧加一个浅 FIFO 缓存 DNS 请求包,并将相邻两个 DNS 包的包间隔增加 8 个周期,这样既能保证 DNS 响应的处理,又不影响整体的流水处理机制。其仿真结果如图 6 所示。

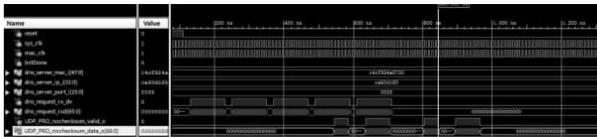


图5 DNS请求包丢失



图6 插入包间隔处理

如图 6 所示,DNS 请求包的包间隔为一个时钟周期,可以看出当解析出 DNS 请求包时,其包间隔被拉大到 3 个时钟周期以上,这样便能够保证所有的 DNS 包均被正确处理。

经过上述处理,基于 FPGA 的 DNS 服务器检查到达的每一个 DNS 请求包,对包进行一系列的流水操作,然后将处理结果发送出去。这样便充分利用了 FPGA 的并行处理特性,实现了真正意义的流水线处理。如图 1、2 所示,其工作流程如下:

- a)用户向基于 FPGA 的 DNS 服务器发送 DNS 请求。
- b)DNS 服务器在接收到请求之后,经过 MAC 层的解析处理,将其目的 MAC 传送至发送侧的 UDP 组包模块,同时将请求暂存在端 FIFO 中(FIFO 中至多暂存两个 DNS 包)。
- c)MAC 解析之后的请求包之间插入 8 个时钟周期的间隔。
- d)依次解析 DNS 请求的源 IP 和源 PORT 信息,并依次送至发送侧的 UDP 组包模块,作为发送侧的目的信息,随包流水处理。
- e)根据 DNS 请求组建 DNS 包,其中不包括域名解析对应的 IP 地址。
- f)Level-1 数据库根据请求域名查找对应的 IP,若有,则直接插入 DNS 响应包,到 i);否则将 DNS 请求发送给 MicroBlaze,到 g)。
- g)Level-2 的 MicroBlaze 收集 Web 服务器的 IP 信息,并告知 DNS 服务器,DNS 服务器将其插入 DNS 响应包,到 h),同时更新其 Level-1 的数据库。对于不存在的域名,将该请求转发

给上层根 name 服务器,直至查找到对应 IP,或者查找不到则结束;

h)对于根 name 服务器查找到的 IP 回送给 MicroBlaze, MicroBlaze 再转发到 DNS 服务器,同时更新 level-2 及 level-1 的数据库。

i)组建好的 DNS 响应包,送至 MAC 发送给用户。

### 3 仿真与验证

#### 3.1 验证平台

本文选择 Xilinx 公司提供的 Virtex-6 系列开发版,型号为 XC6VLX550T,输入时钟为 200 MHz,由开发版上自带晶振产生,FPGA 内部逻辑采用 156.25 MHz,以满足万兆 MAC 口的要求。开发工具为 ISE 13.4,用于 DNS 服务器设计,XPS 13 用于嵌入式 MicroBlaze 设计,仿真软件采用 Xilinx 自带的仿真软件 ISim。开发使用的计算机为 Intel Core,主频 3.40 GHz,内存为 8 GB。

模拟用户采用 IXIA 公司的 Array APV8600 测试仪,它是一种适用于多重服务、应用程序交付平台和网络安全设备的可扩展测试解决方案,具有强大的流处理能力,其应用交付性能够得到数倍提升<sup>[8]</sup>,且其采用每个端口独立多 CPU 内核的形式,能够产生瞬时多数据流,能够很好地模拟大容量、高并发的用户请求。

#### 3.2 验证结果

验证过程主要通过对比通用 DNS 服务器(IBM 3650 上配置 BIND 9.0)与本文设计的 DNS 服务器性能,考虑对比性,本文设定额定每秒 DNS 请求上限为 300 万,对其成功率、吞吐量及延时进行对比,测试时间为 132 s。每组图按照先本文 DNS 服务器后通用 DNS 服务器的顺序放置。对比结果如图 7~9 所示。

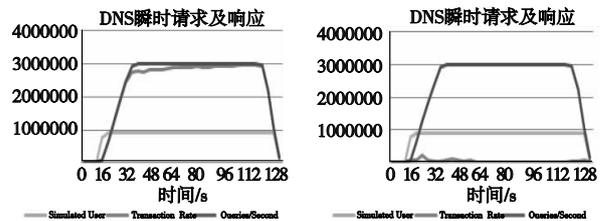


图7 DNS请求及响应瞬时统计

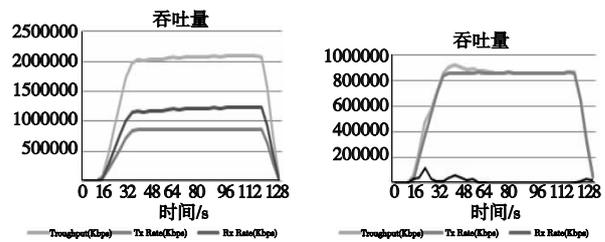


图8 DNS请求及响应吞吐量

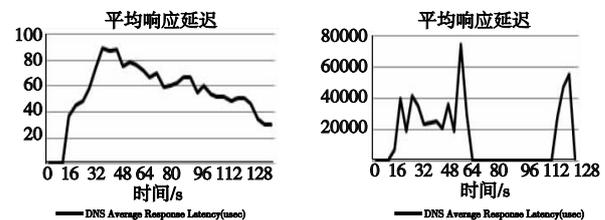


图9 DNS平均响应延迟

由测试结果可以看出,在目标额定每秒 (下转第 1108 页)

扰动。本文针对自适应预失真系统的精度问题,提出先对预失真器离线训练,使其逼近理想的放大器逆特性,再将训练好的预失真器放入自适应预失真系统中。其中离线训练过程是基于 BP 神经网络对功放逆向建模实现的,神经网络建模速度快、精度高。自适应部分采用 LMS 算法调整预失真器的系数。最后通过将本文所采用的方案与经典的直接—非直接学习结构在 MATLAB 软件中仿真比较,仿真结果验证了本文所提方案的可行性和高效性。该方案为神经网络在功放预失真系统中的应用奠定了理论基础。

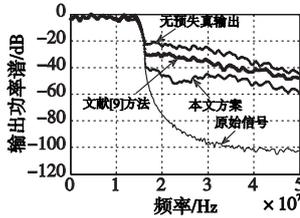


图7 自适应预失真真频谱对比图

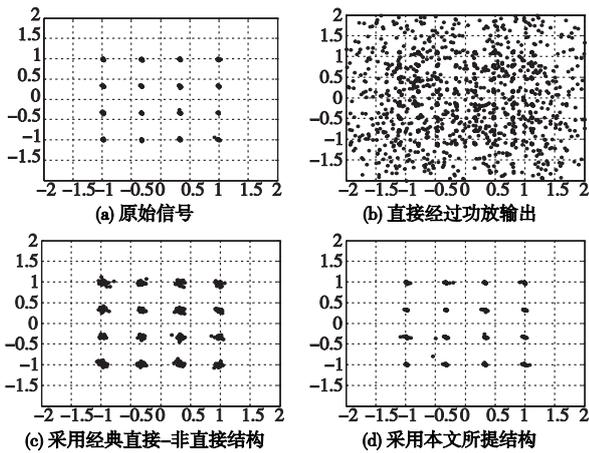


图8 自适应预失真真星座对比图

参考文献:

[1] XU Yong, TONG Xin-hai. A new pre-distorter based on neural network

(上接第 1104 页)300 万 DNS 请求时,如图 7 所示,本文所设计的 DNS 服务器 DNS 请求能够比较好的跟踪,其成功率为 96.73%,而通用 DNS 服务器最大只能达到 30 万/s,成功率仅为 1.30%,实际上,通用的 DNS 服务器能够支持每秒 30 万请求时成功率在 90% 以上;由于 DNS 响应包长为 100 Byte 左右,初步计算其最大吞吐量为  $300 \times 10^4 \times 100 \times 8 \text{ bps} = 2.4 \text{ Gbps}$ ,如图 8 所示,本文设计的 DNS 服务器吞吐量为 2.07 Gbps,基本达到理论最大吞吐率,而通用 DNS 服务器不到 1 Gbps 吞吐量;对于图 9 所示的响应延迟中,通用 DNS 服务器的平均响应延迟在 40 ms 左右,本文设计的 DNS 服务器平均响应延迟为 50  $\mu\text{s}$  左右。

4 结束语

采用 FPGA 设计,结合嵌入式软核 MicroBlaze 实现的大容量高并发 DNS 服务器在处理流程上采用流水线处理的方式,能够达到 DNS 请求的并行处理,实现大容量高并发 DNS 请求的实时处理,极大地提高了 DNS 的响应速度,缩短了其响应延迟,相比于通用的 DNS 服务器,其性能有了很大的提升。

参考文献:

[1] MOCKAPETRIS P. RFC 1034, Domain names: concepts and facilities

for the nonlinear HPA in satellite communications[C]//Proc of International Conference on Communications and Intelligence Information Security. 2010:168-171.

[2] IBNKAHLA M. Neural network predistortion techniques for digital satellite communications [C]//Proc of IEEE ICASSP. 2004: 3506-3509.

[3] 黄圣,胡泽鑫,龚文斌.一种具有普适性的功放线性化预失真方案[J].计算机应用研究,2012,29(5):1854-1856,1859.

[4] 刘影,南敬昌,梁立明.功放的数字基带预失真系统研究与仿真[J].计算机与仿真,2009,26(12):303-306.

[5] ANDREA D A N,LOTTICI V,REGGIANNINI R. Efficient digital predistortion in radio relay links with nonlinear power amplifiers[J]. IEEP Proceedings Communications,2000,147(3):175-179.

[6] GUO Zhi-ying, NAN Jing-chang, LI Jiu-chao. Research for adaptive digital predistortion based on BP-LMS [C]//Proc of International Conference on Computational Problem Solving. 2010:131-135.

[7] 沈英杰,刘郁林,胡中豫.自适应数字预失真方法在功放线性化中的应用[J].重庆邮电学院学报,2005,17(6):687-690.

[8] 崔华,赵祥模,艾渤.记忆功放的 BP 神经网络分离预失真方法[J].西安电子科技大学学报,2010,37(3):565-569.

[9] 詹鹏,秦开宇,蔡顺燕.新的射频功放预失真线性化方法[J].电子科技大学学报,2011,40(5):676-681.

[10] ZAYANI R,BOUALLEGUE R,ROVIRAS D. Adaptive predistortions based on neural networks associated with levenberg-marquardt algorithm for satellite down links[J]. EURASIP Journal on Wireless Communications and Networking,2008,2008(2):1-15.

[11] 刘坤.非线性系统的神经网络逆模型控制[J].南京工程学院学报,2004,2(3):40-45.

[12] RAWAT M,RAWAT K,GHANNOUCHI F M. Adaptive digital predistortion of wireless power amplifiers transmitters using dynamic real-valued focused time-delay line neural networks[J]. IEEE Trans on Microwave Theory and Techniques,2010,58(1):95-104.

[13] 孙欣,丁香乾,石硕.BP 神经网络在柴油机涡轮增压系统故障诊断中的应用[J].计算机应用研究,2006,23(6):202-204.

[14] 王亚静,刘福才,张艳欣,等.基于 BP 网络逆模型的轧辊偏心自适应逆控制[J].化工自动化及仪表,2010,37(6):9-12.

[15] 张浩,叶梧,冯穗力,等.一种新型的基于自适应逆控制的预失真器设计[J].信号处理,2006,22(4):486-491.

[S]. 1987.

[2] JUNG J, SIT E, BALAKRISHNAN H, et al. DNS performance and the effectiveness of caching[J]. IEEE/ACM Trans on Networking, 2002,10(5):589-603.

[3] DEB S, SRINIVASAN A, PAVAN S K. An improved DNS server selection algorithm for faster lookups[C]//Proc of the 3rd International Conference on Communication Systems Software and Middleware and Workshops. 2008:288-295.

[4] PARWES M R, AKBAR M, HAIDER S, et al. DNS propagation delay: an effective and robust solution using authoritative response from non-authoritative server [C]//Proc of the 2nd IEEE International Conference on Information Management and Engineering. 2010:150-153.

[5] RAMASUBRAMANIAN V, SIRER E G. The design and implementation of a next generation name service for the internet[J]. ACM SIGCOMM,2004,34(4):331-342.

[6] 刘晨光,秦华.基于混合式 P2P 的域名系统中超节点选择策略[J].微计算机信息,2010,26(3):107-108.

[7] 郭亚杰,李华,敖腾河,等. DNS 服务器解析性能测试方案设计[J].广西大学学报:自然科学版,2011,36(1):55-60.

[8] 董培欣. Array APV8600 应用交付控制器产品评测[EB/OL]. http://www.cnw.com.cn/test-report/hm2011/20110708\_228729.shtml.