多网安全隔离交换系统的设计与实现

丁烽祥, 张 怡, 王勇军
(国防科学技术大学计算机学院, 湖南 长沙 410073)

摘要: 以隔离网闸技术为基础, 设计并实现了一个多网安全隔离交换系统. 系统采用了多线程、ARP缓存、网络地址转换等技术, 使多个外网能够并行接入系统, 对内网进行安全访问; 采用了零拷贝技术, 并优化了规则匹配算法以提高处理性能. 经验证, 设计实现的多网安全隔离交换系统能够对内外网络进行物理隔离, 并对网络数据进行深度内容检查和安全控制, 且具有较高的网络处理性能.

关键词: 隔离网闸; 协议安全处理; 零拷贝; ARP缓存; 网络地址转换
中图分类号: TP 393 08 文献标识码: A 文章编号: 043820479(2007)S220092206

随着互联网的广泛应用, 网络安全问题越来越受关注. 各种专有网络连接到互联网时通常在接入处部署防火墙来保护内部网络免遭来自互联网的攻击. 防火墙能够通过包过滤、状态检测、应用代理等技术限制非法流量, 提供对内部网络的保护, 但是防火墙技术本身存在的问题使得它并不能完全保证内部网络的安全性. 防火墙是基于 TCP/IP协议进行信息交换和过滤的, 而 TCP/IP协议本身就存在漏洞; 它在可信网络与不可信网络之间提供了网络的直连通道, 不能完全保证通过数据的安全性. 针对防火墙存在的问题, 隔离网闸技术[1]应运而生. 隔离网闸的实现原理是将可信网络与不可信网络进行物理隔离, 通过专用通信设备和专用通信协议在不同安全级别网络之间进行数据交换, 彻底阻断网络间的直接 TCP/IP连接, 因此增加了网间数据交换的安全性, 杜绝了由于操作系统和网络协议自身漏洞带来的安全风险.

隔离网闸是架设在内网和外网之间的安全设备. 内网与外网为不同的安全域, 具有不同用户对象和数据敏感程度. 通常意义上, 内网即为可信网络, 外网即为不可信网络. 隔离网闸对内网提供安全防护. 虽然隔离网闸具有高安全性, 但由于单一的内外网接入, 使得灵活性较防火墙差.

在实际网络环境中经常存在多个外网同时接入内网的情况, 要求外网和内网之间能够进行安全访问, 而外网与外网之间相对独立, 且不允许任何通信. 例如, 在公安业务应用系统中, 作为数据中心的公安内网需要同时与移动警务系统、旅馆业治安管理系统系统、道路交通违章信息系统等类型的分散的多个公安外网进行安全通信, 且各公安外网间没有直接的业务关系. 在内网与每个外网间架设普通的隔离网闸设备, 会造成资源浪费且不易维护. 因此, 本文在隔离网闸技术的基础之上设计和实现了多网安全隔离交换系统, 其应用环境如图 1.

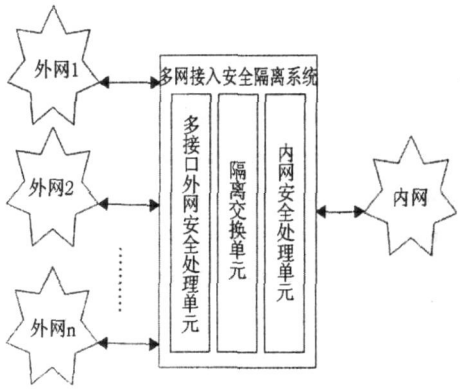

图 1 多网安全隔离交换系统的应用环境

多网安全隔离交换系统处于内网与多个外网之间, 其内网安全处理单元与内网相连并且是对应内网通往外网的网关, 外网安全处理单元与多个外网相连并且是对应多个外网通往内网的网关, 隔离交换单元隔离内外网, 并且作为内网处理单元与外网处理单元信息交互的桥梁. 外网相互之间独立, 即它们各自的网络拓扑结构独立, 地址空间独立, 所对应隔离交换系统上的网络接入接口独立, 且无法通过隔离交换系统进行相互通信. 但多个外网可以并行通过隔离交换系统对内网进行安全访问.

收稿日期: 2007208216
基金项目: 国家自然科学基金(90604006)资助
作者简介: 丁烽祥, 男, 硕士研究生.
* 通讯作者: dingzhuol10@tom.com

系统中, 隔离交换单元采用专门设计的硬件隔离交换卡, 安全处理单元采用普通的 CPU 卡及专门设计的安全处理软件. 本文主要介绍软件部分的设计与实现.

1 软件体系结构

1.1 系统软件体系结构

多网安全隔离交换系统软件表现为一个分布式结构, 其软件结构如图 2

(1)管理审计软件: 接受管理员命令, 实现对系统的远程监控、配置管理、审计数据维护等操作.

(2)管理控制软件: 接受管理审计软件的管理控制信息, 实现对隔离交换单元和内 外网处理单元的配置、管理和控制.

(3)协议安全处理软件: 运行在安全处理单元上, 实现对网络数据的多层次协议解析和安全检查, 并将解析和检查结果的审计信息发送给管理审计软件.

系统的软件体系结构与隔离网闸的软件体系结构相类似, 最大的区别在于协议安全处理软件实现了对多个接入网络的处理, 这是多网安全隔离系统与一般隔离网闸的主要区别, 也是本文论述的重点.

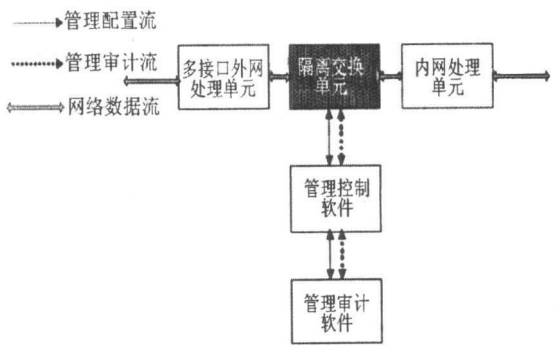

图 2 软件体系结构 (隔离交换单元为硬件)

1.2 协议安全处理软件

协议安全处理软件运行在内外网安全处理单元上, 分别对应一个内部和多个外部业务网络, 主要功能有:

1)使系统在网络拓扑环境中作为所接入网络通往另一不同安全等级网络的网关;

2)允许相互独立的多个外网并行对内网进行安全访问, 而不会产生冲突;

3)针对网络数据进行网络层访问控制和应用层深度处理, 即对应用内容进行还原, 完成各种访问控制、安全内容检查和安全审计等功能.

(1)多线程软件处理结构

协议安全处理软件采用了多线程并行处理方式, 实现对多个接口的数据隔离交换处理, 如图 3

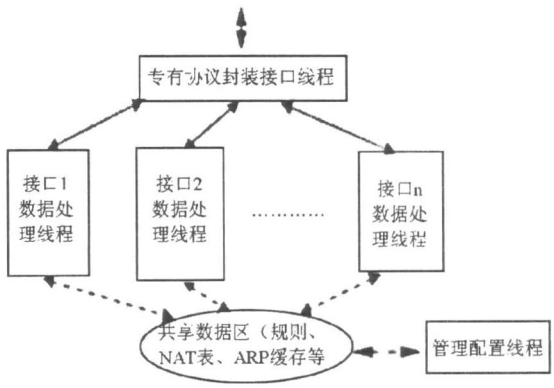

图 3 多线程软件处理结构

专有协议封装接口线程对应多个数据处理线程, 实现专有封装隔离通路的复用. 数据处理线程访问共享数据区中的有关数据, 包括访问控制规则、审计规则、NAT 表、ARP 缓存等, 对数据进行处理. 而这些数据是由管理配置接口线程通过接收管理控制软件发出的配置命令写入维护的.

(2)功能模块组成

协议安全处理软件的软件模块划分及模块与模块间的作用关系如图 4

软件主要的功能模块有: 高速数据包捕获和发送模块、ARP 处理模块、NAT 模块、TCP/IP 协议还原重组模块、应用协议解析及安全控制模块、专有隔离协议模块和管理控制模块等.

高速数据包捕获和发送模块实现高速从网络上截获数据包并将处理完的数据包发送到网络上.

ARP 处理模块主要负责获取接入网络中各节点的硬件地址和提供系统对应接入网络的接口的硬件地址.

NAT(网络地址转换)模块根据内外网网络地址规划, 通过管理控制软件的 NAT 配置, 从管理控制模块中获取地址转换信息, 对来自内外网的数据包进行网络地址转换, 从而完成内外网的互通.

TCP/IP 协议还原重组模块根据数据报的网络层和传输层头部信息进行合法性检查、还原和重组. 对于来自本协议安全处理单元接入网络的报文提取出其中的五元组信息 (源 IP、目的 IP、源端口、目的端口、协议号, 其中源端口和目的端口只针对 TCP/UDP 协议), 通过管理控制模块查找五元组规则, 进行访问控制. 类似防火墙的包过滤技术, 这属于底层的安全控制.

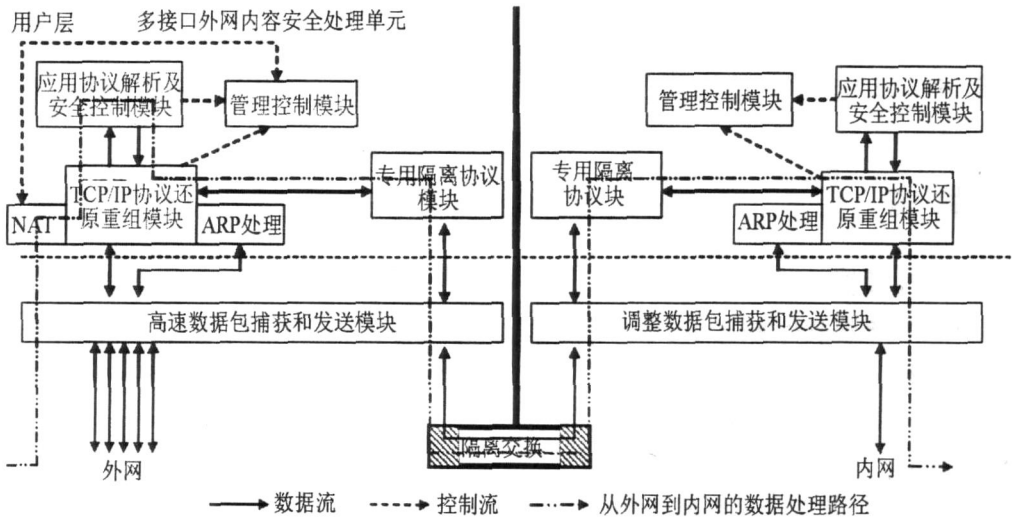

图 4 软件的模块组成

应用协议解析及安全控制模块根据传输层的端口号判断应用层协议类型, 调用相应的应用协议解析模块. 解析模块根据应用协议规范, 对应用层数据进行解析、还原、状态跟踪, 并进行格式化和内容检查, 通过管理控制模块查找应用层访问控制规则, 实现应用层的安全控制.

专用隔离协议模块完成对内外网间交换数据的专用协议封装. IP 数据要经过了协议头剥离, 封装成内部的专用格式, 通过高速数据报发送模块发送至隔离交换单元. 这样避免了内外网之间 IP 报文的直接交换, 解决了标准 TCP/IP 协议漏洞引起的安全隐患.

管理控制模块是连接管理控制软件和协议安全处理软件的桥梁, 它接收来自管理控制软件的管理配置信息, 并将协议安全处理软件产生的审计日志、统计值等信息发送给管理控制软件. 管理控制模块通过安全策略管理、审计日志管理等模块加以支持, 间接影响数据处理路径.

2 关键技术的实现

系统既需要实现多网安全隔离交换的功能, 同时还需要尽可能的提高网络数据处理性能, 因此, 协议安全处理软件采用了以下的关键技术.

2 1 高速数据包捕获和发送

高速数据包捕获和发送技术的基本原理是: 在外网处理单元与外网连接的接口上, 内网处理单元与内网连接的接口上, 采用零拷贝技术进行以太网数据包的捕获和发送.

零拷贝[2]的基本思想是: 数据包从网络设备到用户程序空间或从用户程序空间到网络设备的传递过程中, 减少数据拷贝次数, 减少系统调用, 实现 CPU 的零参与. 零拷贝用到的最主要的技术是 DMA 数据传输技术和内存区域映射技术. 实现中, 在用户层分配一个适当大小的用户缓冲区 (UserBuffer), 存放接收和发送的数据包. UserBuffer被分成多个大小为 2 kb的块, 每块存放一个数据包. 在内核区创建一个 Dum[2], 作为用户程序与网络接口进行交互的中介, 并成为上层应用操作网络的句柄. Dum为用户缓冲区的每个块建立一个描述符, 并将每个描述符挂入 Dum不同的队列里. Dum共建立了发送队列、接收队列、空闲队列 3个环形队列. 待发送的数据包的描述符挂入发送队列, 接收到的数据包的描述符挂入接收队列, 没有数据的块的描述符挂入空闲队列. 同时, Dum还维护一张虚地址))) 物理地址对应表 (VPAT). 用户缓冲区有多少虚地址页, 这张表就有多少项, 每一项存放该虚地址页对应的物理页面的首地址. Dum本身在内核中的缓冲区空间被映射到用户区, 和 UserBuffer一起成为一块由内核和用户共享的区域. 这块区域包括 UserBuffer, 3个环形队列和地址对应表. 这样网卡和用户都可以自由地访问这块共享区域的数据. 修改网卡的驱动程序, 将网卡设备的私有域填上 Dum设备的信息, 网卡进行直接存储器存取 (DMA)传输时, 直接根据 Dum中的描述符找到真正的位于用户区的缓冲区, 然后直接将数据包从网卡传到用户区; 同理, 用户程序根据 Dum中的描述符在共享缓冲区中填入将要发送的数据包, 并通过 DMA 传送给网卡进行发送.

采用零拷贝技术的主要原因是: 第一, 增大系统在网络中的吞吐量, 减少时延, 提高系统的网络数据处理性能; 第二, 避免了操作系统内核协议栈的参与, 减少了由于系统漏洞和协议漏洞带来的危害.

2.2 ARP缓存

在以太网环境中, 多网安全隔离交换系统要作为所接入网络通往另一不同安全等级网络的网关, 就必须先支持 ARP(地址解析协议). 以太网数据链路层的寻址机制依靠 48 bits的 MAC地址 (硬件地址), 这是使用数据链路的任何网络层都必须遵从的. 当一台主机把以太网数据帧发送到位于同一局域网上的另一台主机时, 是根据 48 bits的以太网数据链路层地址来确定目的接口的. ARP为 IP地址到对应的硬件地址之间提供动态映射[5].

由于网络环境中有多个外网, 系统对多个外网进行编号, 使每个外网对于系统来说都有唯一的网号.

ARP缓存的目的是为了在系统转发报文时查找目的 MAC地址. 系统的 ARP缓存以哈希表的方式进行维护, 以便快速查找目的 MAC地址. 通过网号 (只针对外网安全处理单元)和 IP地址计算出哈希键值, 用来申请、释放、查找哈希表项. 哈希表项的内容就是所对应的 MAC地址. 同时, 每个表项维护一个定时器. 在固定的一段时间 (定时器设定的时间)后, 定时器超时, 表项被释放.

ARP处理的流程比较简单, 主要是针对 ARP请求报文和 ARP应答报文. 若从接口获得的报文类型是 ARP(无论是请求还是应答), 则提取出它的源 IP和源 MAC, 更新 ARP缓存; 若此 ARP报文是对本接口 IP询问 MAC地址的请求报文, 则响应 ARP应答; 若系统准备将来自一个网络的报文转发至另一个网络时, 根据报文目的 IP查找 ARP缓存没有发现对应的 MAC地址, 则发送 ARP请求报文.

2.3 网络地址转换

由于多个外网的地址空间的划分是独立的, 且是任意的, 就有可能引起 IP地址冲突, 从而造成外网与内网的通信混乱. 为此, 引入网络地址转换 (NAT)技术. 只有外网协议安全处理单元上存在 NAT模块, 它根据内外网网络地址规划, 通过管理控制软件的 NAT配置而产生地址转换表, 对来自外网的数据包进行源地址转换 (SNAT), 对来自内网的数据报进行目的地址转换 (DNAT), 从而完成内外网的正常互通.

NAT可以分为静态地址转换、动态地址转换和复用动态地址转换[6]. 在系统中采用了静态地址转换和动态地址转换两种方式. 为了方便起见, 本文称转换前的 IP地址为原始地址, 称转换后的 IP地址为可用地址. 静态地址转换是将原始地址与可用地址进行一对一的转换, 这种转换关系由管理控制软件的 NAT配置指定, 在系统运行过程中不会动态改变, 适用于对服务器的 IP地址进行转换. 动态地址转换也是将原始地址与可用地址进行一对一的转换, 但它是从管理控制软件 NAT配置的地址池中动态的选择一个未被使用的可用地址对原始地址进行转换, 这种转换关系在系统运行的过程中会动态改变, 适用于对客户端的 IP地址进行转换. 地址池是一个链表结构, 其中每个结点用来维护一个未被使用的可用地址. 当地址池中的某个可用地址被选择时, 与其对应的结点将从链表中被删除, 即可用地址从地址池中移出.

NAT地址转换表采用的方式和 ARP缓存相同, 都是使用哈希表的方式. 根据转换内容可以分为源地址转换表和目的地址转换表 (每个一对一的转换关系在两个表中都有一个对应项). 源地址转换表是通过网号和源 IP地址 (原始地址)计算哈希键值, 哈希表项内容是可用地址; 目的地址转换表是通过目的 IP地址 (可用地址)计算哈希键值, 哈希表项内容是原始地址和网号. 对于以动态地址转换方式生成的表项, 面向连接和非连接协议的处理方式不同. 面向连接的协议, 在连接结束时表项被释放, 可用地址返还给地址池; 面向非连接协议, 启动相应的定时器, 在固定的一段时间 (定时器设定的时间)表项未被使用后, 定时器超时, 表项被释放, 可用地址返还给地址池.

2.4 规则匹配的优化

TCP/IP协议还原重组模块提取出数据包中的五元组信息, 通过管理控制模块查找和匹配规则, 进行五元组访问控制. 应用协议解析及安全控制模块提取出数据包中的应用层信息, 通过管理控制模块查找和匹配规则, 进行应用层访问控制. 为了缩短当前网络状况下数据包匹配规则的时间, 提高系统的网络数据处理性能, 采用统计分析的方法对访问控制规则进行优化: 统计访问控制规则在最近一段时期的使用频率, 对规则的相对次序进行动态调整, 使得使用最频繁的规则位于规则列表的最前面.

动态调整规则的次序以适应当前的网络流量状况, 首先要合理的调整采样统计时间. 采样统计时间的选取与当前网络流量的大小有着直接的关系: 当网络流量比较大时, 由于能在较短的时间内统计足够的数据, 得到网络流量的发展趋势, 为了避免滞后性就应该缩短统计时间, 使规则次序尽快适应当前网络流量; 当网络流量比较小时, 应该延长统计时间, 从而统计足够多的数据包, 较为准确地预测当前网络流量的发展趋势. 在实现中, 设置一个定时器, 定时器的超时值为统计时间. 系统设定初始的统计时间为 T_{init}(也是最小的统计时间), 最大的统计时间为 T_{top}, 统计时间计算因

子 F_{time}=系统最大处理流量/当前网络流量. 当定时器超时, 统计时间被重新计算, 计算表达式为 $T=T_{init}*F_{time}$, 若 T 大于 T_{top}则 T 被赋值为 T_{top}, 将统计时间 T 赋值给定时器的超时值, 重新启动定时器.

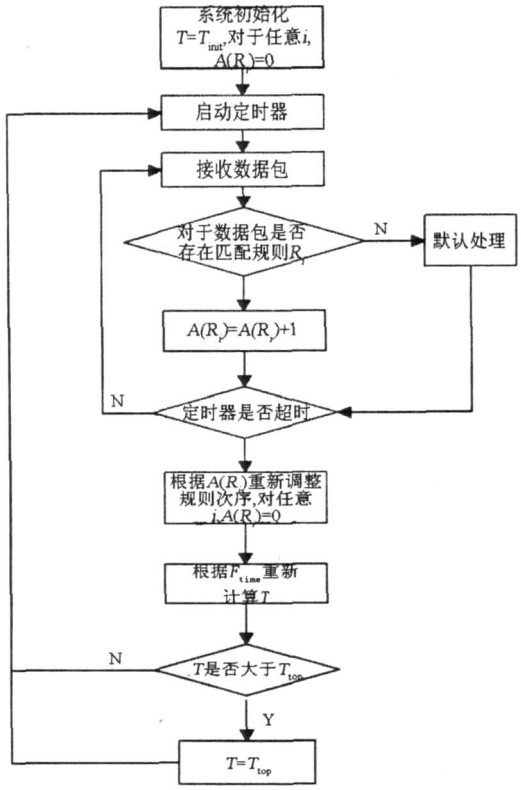

图 5 规则匹配优化算法流程图

R_i 表示规则列表第 i条规则, $A(R_i)$表示规则 R_i 统计使用频率的计数器. 在定时器超时以前 (即在统计时间内), 若一个数据包匹配规则 R_i, 则 $A(R_i)$计数加 1; 当定时器超时, 根据规则列表中各规则的使用频率, 从高到低重新对规则进行排序 (当然, 相互关联的规则的相对位置不能改变, 以免影响系统的安全性能), 然后清空 $A(R_i)$, 重新计数.

规则匹配优化算法的流程如图 5.

3 功能测试与性能测试

3.1 功能测试

为了验证多网安全隔离交换系统的主要功能, 设计了如图 6的实验环境.

3个外网各有一台客户机, 每台客户机的配置相同, 操作系统为 Windows XP, 256 MB 内存, CPU 2.0 GHz, 硬盘 40 GB, IP 地址为 10.129.212.5, 掩码为 255.255.255.0, 网关 IP地址为 10.129.212.1. 内网服务器的配置: 操作系统为 Fedora Core 5, 1 GB内存, CPU 2.8 GHz, 硬盘 120 GB, IP地址为 192.168.0.5, 掩码为 255.255.255.0, 网关 IP地址为 192.168.0.1. 三个外网的客户机需要通过 FTP同时从内网服务器上下载文件, 则在隔离交换系统上通过管理控制软件进行以下设置: 配置各外网接口的 IP 地址为 10.129.212.1(对应外网的网关地址), 内网接口 IP 地址为 192.168.0.1(对应内网的网关地址); 配置 NAT, 将外网 1、2、3的 10.129.212.5这个 IP地址分别转换为 1.190.11.1, 1.190.11.2, 1.190.11.3; 配置五元组规则和 FTP 应用层规则 (包括用户名、FTP 命令、文件名、文件大小等). 实验结果如表 1.

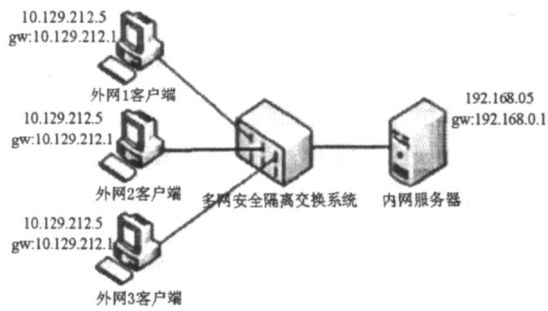

图 6 功能测试实验环境

实验结果验证了: 多个外网能够并行接入系统, 对内网进行安全访问; 系统对网络数据进行深度内容检查和访问控制.

3.2 性能测试

使用 Spirent Adtech AX/4000 Broadband Test System测试仪对系统进行性能测试. 隔离交换系统的

表 1 功能测试结果

外网号	客户端原始 IP地址	NAT转换后 IP地址	i的动作	R的动作	FTP文件传输	平均下载速率
1	10.129.212.5	1.190.11.1	允许	允许	成功	2.5Mbit/s
2	10.129.212.5	1.190.11.2	允许	允许	成功	2.6Mbit/s
3	10.129.212.5	1.190.11.3	允许	禁止	失败	无

注: 表中 r表示五元组规则, R表示应用层规则.

配置与功能测试的配置基本相同, 还需要添加规则, 允许测试仪构造的 UDP报文通过. 系统内外网安全处理单元的物理网卡都采用千兆网卡. 测试仪配置一个发送端, 连接系统的内网接口; 配置 3个接收端, 分别连接图 6中 3个外网接口. 测试仪在发送端构造三个 256字节的 UDP报文, 目的 IP分别为 1. 190. 11. 1, 1. 190. 11. 2, 1. 190. 11. 3, 且以相同的速率发送这 3个报文. 测试结果: 系统的最大吞吐量为 450 Mbit/s, 平均时延为 823 Ls 测试结果表明系统具有较高的网络数据处理性能.

4 结束语

多网安全隔离交换系统在隔离网闸技术的基础上, 采用多线程、零拷贝、NAT等技术实现了多个外网并行接入隔离交换系统与内网进行安全通信, 目前该系统已在实际环境中试用. 经实验和试用证明, 该系统具有较高安全性能和数据处理性能.

参考文献:

[1] 万平国. 网络隔离与网闸 [M]. 北京: 机械工业出版社, 2004.

[2] 王佰玲, 方滨兴, 云晓春. 零拷贝报文捕获平台的研究与实现 [J]. 计算机学报, 2005, 28(1): 46- 52.

[3] Kurmann A, Rauch F, Stricker T. Speculative defragmentation2leading gigabit Ethernet to True zero2copy communication [J]. Cluster Computing 2001, 4(1): 7- 18.

[4] Rubini A, Corbet J. LINUX device drivers [M]. Sebastopol: O. Reilly& Associates, 2002.

[5] W. Richard Stevens. TCP/IP详解 卷 1: 协议 [M]. 范建华, 胥光辉, 张涛, 译. 北京: 机械工业出版社, 2000.

[6] Egevang K., Francis P. The IP network address translator (NAT) [R]. RFC1631, 1994.

The Design and Implementation of Multi2network Security Isolation2Exchange System

DING Feng2xiang　ZHANG Yi　WANG Yong2jun

(School of Computer, National University of Defense Technology, Changsha 410073, China)

Abstract　Based on the technology of isolation network2gap, we design and implement a multi2network security isolation2exchange system. The system adopts the technologies such as multi2thread, ARP buffer and NAT. So it can be connected with more than one outside networks simultaneously, and permit the outside networks to access the inside network safely at the same time. Moreover, the system adopts zero2copy and optimizes the rule2match algorithm to improve process performance. It can be proved that the multi2network security isolation2exchange system not only implement physical isolation between the inside network and the outside networks, in2depth detect and security control the datagram on networks, but also attain high network process performance.

Key words　isolation network2gap; protocol security process; zero2copy; ARP buffer; NAT