发的方法,进一步引导学生使用计算机进行相关专业的学习,并创造条件使学生能够利用计算机处理毕业设计、毕业论文,甚至研制、设计、开发计算机应用系统(软件).其次,计算机基础课教师也要转变观念,加大计算机科学与技术、专业知识及应用的研究,加强同专业教师的合作以及指导的力度,促进计算机基础教学向计算机应用的深层次发展.

编辑:李志敏

面向宽带网络的通用信息侦听模型**

云晓红　张艳萍
(牡丹江市职工大学　牡丹江　157000)(牡丹江师范学院　牡丹江　157012)

摘　要　探讨了面向宽带网络的通用信息侦听模型.在该模型中,利用可扩展集群系统的并行计算机能力处理宽带网络所产生的海量数据;为实现适合各种网络信息捕获,提出了物理层捕包策略;为均衡各节点机的负载,建立了一个基于 TCP/IP 协议四元组的数据划分模型.

关键字　网络安全;宽带网络;侦听

1　引言

Internet 的飞速发展一方面直接促进了全球经济的爆发性增长,另一方面也带来了严重的信处安全危机.如何解决 Internet 网络安全问题已成为目前 Internet 技术领域中重要的研究课题.各国把研究网络安全问题,特别是把基于 Internet 的网络安全问题提到了国家安全的战略高度.

信息侦听是诸如入侵检测、网络安全管理等网络安全研究领域的核心技术之一[1].然而,传统的网络侦听技术具有局限性.主要表现在:①传统的网络侦听技术大多是针对广播式网络,其基本原理是利用广播式网络所具有的广播特性,在网络的数据链路层捕获所有流经监测路口的数据信息.从广义上讲,有两种类型的网络传输技术:广播式网络和点对点网络[2].对于点对点网络,这一侦听技术将不再适用.②在网络带宽较低的情况下,网络侦听主要由单节点计算机完成.然而随着宽带网络的迅速普及,单一计算机的计算能力不足以处理宽带网络所产生的海量数据.

为此,突破传统的信息侦听框架,建立新的信息侦听体系结构是实现宽带网络侦听的关键所在.

2　并行侦听模型

为测试单节点计算机处理网络数据的能力,采用普通 PC 服务器进行了实验.实验中,用两台同方服务器根据真实统计数据产生相应比例大小的随机地址的包,由第三台同方服务器接收.被测服务器配置为双 PⅢ600 CPU,1G 内存,SCSI 硬盘,Inteleepro100 NIC(SpeedoⅡ系列芯片)网卡,操作系统为 Linux RedHat 6.2,用 libpcap 捕包之后,将包地址和端口存入内存中的二叉树结构中.

表 1　实验记录

捕包网卡数目	发送速率 (Mbps)	收到包数 (k)	发送包数 (k)	收到比率 (%)
1	36.8*2=73.6	591	946	62.4
1	31.7*2=63.4	774	946	81.8
2	36.8*2=73.6	620	946	65.5
2	31.7*2=63.4	763	946	80.6
2	26.3*2=52.6	946	946	100

结果表明,捕包很占CPU时间以及系统 I/O 中断,以至于在高速率多处理一些的时候就会丢包.而且在实验中发现只做很简单的处理,只能达到 50 Mbps 不丢包,由于实际应用程序要做的处理比它要均匀,所以如果没有很大改善,不容易超过这个数字.

显然,对于宽带网络,采用单节点计算机无法满足数据处理的需要.为此,在模型中,采用可扩展集群的体系结构对网络数据进行侦听.这样,对于带宽小于 50 M 的网络,安置单节点 PC 服务器,而对于大于 50 M 的网络,则根据实际情况,由 N 节点集群系统进行并行侦听.

* 收稿日期:2002-11-18
** 国防"九五"预研项目(编号:15.7.2)

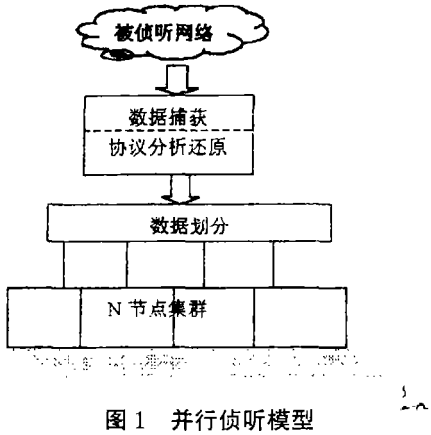


图 1 并行侦听模型

为实现通用高效的并行侦听,本模型必须实现如下两个关键技术:

(1)数据捕获.由于网络的多样性,因此需要采用通用技术对各种网络的信息进行捕获;

(2)数据划分.由于消除数据依赖并保证负载平衡是提高并行计算性能的关键所在[3],因此需要对捕获的数据进行有效的数据划分.

3 物理层捕包模型

本文提出一个面向有线传输的通用物理层捕包模型.该模型对于广播式网络和点对点网络同样适用.其原理是根据网络物理层的协议特性,将传输线路上的物理信号旁路下来.然后采用与所侦听网络对等的协议栈,将其从物理层还原成有意义的网络数据包,如 IP 包.由于当前的有线传输介质主要是光纤、电缆、双绞线等光或电传输介质.因此,本模型主要利用光、电信号的物理特性,对其进行捕获.对于采用光传输的网络,利用光信号的反射和透射原理[4],用分光器代替进行传输的光纤(如图 2),这样所有传输的光信号除正常传输外同时将被反射到分光口中,从而实现光信号的物理层捕获.

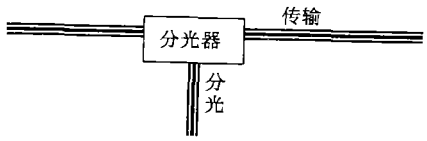


图 2 光信号物理层捕包模型

对于采用电传输的网络,由于数学信号是通过高低电平表示 0、1 代码,因此通过采用电路并联的方法,可将传输的电信号同时旁路下来(如图 3).

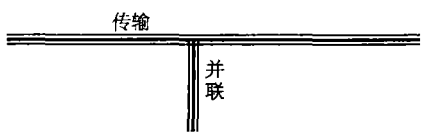


图 3 电信号物理层捕包模型

当前的数据传输大多采用不同的线路分别对信号收发,如光纤传输中采用两路不同的纤芯分别收发信号,而在 100 BaseTX 网络中,分别用双绞线的 1,2 芯发送信号,3,6 芯接收信号.这样将数据传输中的双向信号完整地捕获下来,需要分别利用两个接收端口捕获收发两个方向信号.同时为了不干扰网络的正常传输,侦听设备上配置发送端口.

在正常的网络传输中,接收设备对于目的地址非自身的数据拒收.因此,捕包设备必须设置为全接收模式,以便能够将所有的数据全部捕获下来.

4 数据划分策略

为使集群系统的并行侦听性能达到最优,一个理想的数据划分策略应该能够保证如下两方面:①数据被均匀地划分到各个节点机上,以保证节点机间的负载均衡.②任意一个 TCP 连接的双向数据都被分流到同一节点机上,以保证各节点机间无数据依赖.

在 TCP/IP 协议中,四元组(源 IP、目的 IP、源端口、目的端口)唯一地确定了一个连接[5].在我们的数据划分策略中,通过将该四元组的各项异或获取该连接的特征值,然后将该特征值对节点机数取模,所得的结果即为对该连接 进行处理的节点机号.

(源 IP ⊕ 目的 IP ⊕源端口 ⊕目的端口)MOD N

其中,N 为节点机数.

上面的策略在没有 IP 分片的情况下,效率会很高.但一旦出现 IP 分片,由于必须对其进行重组后,才能获得整个 IP 包的连接端口号,因此会影响数据划分的效率.但所幸的是随着网络通信技术的发展,在各种网络上一般都已经支持传输较大的数据包.因此在 INTERNET 上传输的 IP 分片很少,TCP 协议的 IP 分片几乎没有.下面的实验数据是在哈工大校园网出入口上所获取的数据.

(1)每 10 秒 IP 分片包总数为 25 个,每 10 秒 IP 分片包总长度 18 176 bytes;

(2)每秒内 IP 分片平均包数 3 个,每秒内 IP 分片平均流量 18 540 bps;

(3)每秒内 IP(TCP)分片平均包数 0 个,每秒内 IP(TCP)分片平均流量 0 个;

(4)每 10 秒 IP 包总数 236 576 个,每 10 秒 IP 包总长度 66 359 415 bytes;

(5)每秒内 IP 平均包数 23 657 个,每秒内 IP 平均流量 53 087 560 bps.

在数据划分策略中,网络流量的动态性使得各台节点机分配的流量在某一瞬间未必能够理想的均摊,但从统计学的角度,在各个时间段内每个节点机的负载将是均衡的.

5 结论

通过采用并行处理技术、物理层捕包技术和

基于TCP/IP协议四元组的数据划分策略建立了对于宽带网络的通用侦听模型,该模型在网络安全的研究中具有重要意义.

参考文献

1 胡昌振,李贵涛等.在向21世纪网络安全与防护[M].北京:北京希望电子出版社,1999.
2 Andrew S. Tanenbaum 著.计算机网络[M].熊桂喜,王小虎译.北京:清华小学出版社,2000.
3 云晓春等.Sigma系统中数据依赖关系分析的完善与增强[J].计算机研究与发展,1998,(3)
4 吕海宝,冯勤群等.强度型光纤传感检测中的强度补偿技术[J].激光技术,1999,(2)
5 W. Richard Stevens 著.TCP/IP详解[M].范建华,胥光辉等译.北京:机械工业出版社,2000.

编辑:李志敏

考号生成器的设计与实现*

赵晓霞　王长龙　陈纯锴

(①③牡丹江师范学院计算机系　牡丹江　157012)(②黑龙江林业职业技术学院　牡丹江　157012)

摘　要　利用vb 60数据库支持电子表格,可完成对电子表格中字段的查找,记录的添加功能,实现考号的自动生成.

关键字　数据源;生成;译码;Visual Basic 6.0

随着时代的进步,社会的发展,计算机已深入到人们生活的各个领域,特别是在教育教学中,计算机教学的重要性得到了普遍的认可,但在教学管理中结合实际的应用却不多.例如在考试中对学生考号的安排,许多教学管理人员对计算机的应用仅仅是将最终结果由计算机输出,即只利用字表处理软件将手工排列的序号输出对应的文档,没有充分利用计算机的潜力.教学管理人员没有从繁琐的工作中解脱出来.因此开发出实用的考号生成软件已成为必需,以实现教学管理的自动化.

1　传统的编排考号方法的弊端

传统的编排考号是完全采用手工的方法,由工作人员将年级、系别、考场号、座位号按一定顺序组合在一起后分配给某位学生,如果用机器批卷,还要将其姓名翻译为相应代码,序号加姓名代码近20位数字,这些工作均有教学管理人员手工完成,不仅费时费力,且准确率不高,易出错.

2　采用考号生成器进行考号生成的可能性

由于人们对计算机知识的地位和作用的认识日益清晰,加之计算机技术的飞速发展,计算机性能价格比的快速提高,许多学校在教学管理方面加大了资金投入力度,一般的办公室都配备了计算机,并且在教学管理人员的实际操作能力培训等方面加大了力度,使每位教学管理人员都能掌握计算机的基本操作技能.这些都为本软件的普遍应用提供了可靠保证.

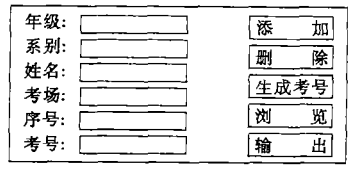
3　考号生成器的基本原理与实现

3.1 开发软件的选择

本软件主要完成对已有的数据源进行统计分析,将多条记录不同的字段连接生成新记录的一个字段,该数据源是一个,我们选用面向界面、具有高强数据库处理能力的Visual Basic 6.0进行开发,以实现对数据库进行读写操作的要求.

3.2　软件的基本原理与实现

本软件是由一个窗体构成的工程,窗体(见附图),用以完成学生年级、系别、姓名、考场、序号的录入(即数据源相应字段的录入);相应信息录完后可完成考号的自动生成.考号的输出由数据源(电子表格)直接输出.



4　考号生成器实现的关键技术

本软件实现的关键技术是汉字的译码.在考号编排中涉及到将学生姓名翻译为相应编码的过程.这个问题的完成需要有大量汉字译码及根据译码建立的一个汉字及其译码的数据库,该库的建立工作量较大,需要大量的信息录入.有了

* 收稿日期:2002-09-16