基于 Netfilter 的流量检测与控制系统

蒋 华[1],王汝传[1,2],李致远[1]
(1. 南京邮电大学 计算机学院,江苏 南京 210003;
2. 南京邮电大学 计算机研究所,江苏 南京 210003)

摘 要:P2P 流量已成为互联网流量的重要组成部分,由于其占用大量带宽,严重影响了其他网络业务的运行,因此如何有效地检测和控制网络流量已成为目前面临的一个重要难题。基于 Linux 的 Netfilter 防火墙和连接跟踪机制,结合应用层协议识别工具 L7-filter 和流量控制器 TC 设计并实现了一个简单而高效的流量控制系统。该系统首先对数据包进行应用层协议识别并分类,然后对有限的网络带宽进行合理分配,实现流量控制。实验证明此系统能有效地检测和控制网络流量。

关键词:流量检测与控制;防火墙;第七层过滤;连接跟踪
中图分类号:TP31 **文献标识码**:A **文章编号**:1673-629X(2012)01-0090-03

Traffic Identification and Control System Based on Netfilter

JIANG Hua[1], WANG Ru-chuan[1,2], LI Zhi-yuan[1]
(1. College of Computer, Nanjing University of Posts and Telecommunications, Nanjing 210003, China;
2. Institute of Computer Technology, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract: P2P traffic has taken great portions in the network traffic. It is a serious influence to the rest of network service that peer-to-peer flows occupy the network bandwidth seriously, therefore how to effectively identify and control network traffic has been a very important problem. A simple and efficient traffic control system was designed and realized based on Linux Netfilter firewall and connecting tracking strategy. It combined the L7-filter which identifies packets based on application layer data and TC which is a traffic control tool. The system can identify and classify packets based on application layer data and allocate limited network bandwidth properly to realize traffic control. The experiment proves that the system can identify and control network traffic effectively.

Key words: traffic identification and control; netfilter; L7-filter; connecting tracking

0 引 言

随着因特网的飞速发展,各种网络应用日趋复杂,网络流量不断增长并且呈现多样化。一些非关键业务,特别是 P2P 业务[1],消耗大量网络带宽,严重影响了 HTTP 等关键业务的传输质量,如何高效、准确地识别和控制网络流量显得尤为重要[2-4]。传统的一些基于端口的协议识别已经不能满足需求,如今越来越多的 P2P 软件采用动态端口。文中基于 Linux 的防火墙 Netfilter,使用 iptables 的补丁插件 L7-filter(Layer7 filter)工具对数据包进行应用层协议识别并分类,再利用 Linux 的流量控制器 TC(Traffic Control)对每类数据包进行流量控制,设计并实现了一个简单而高效的应用层协议流量控制系统。

收稿日期:2011-05-29;**修回日期**:2011-09-08
基金项目:国家自然科学基金(60973139,61003039,61003236);江苏省科技支撑计划(工业)项目(BE2010197,BE2010198);江苏省高校自然科学基础研究项目(10KJB520013,10KJB520014);高校科研成果产业化推进工程项目(JH10-14);江苏省六大高峰人才项目(2008118);教育部高等学校博士学科点专项科研基金(20103223120007)

作者简介:蒋 华(1986-),男,江苏苏州人,硕士研究生,研究方向为计算机软件、计算机通信、对等网络、流量控制技术等;王汝传,教授,博士生导师,研究方向为计算机软件、计算机通信、信息安全、无线传感器网络、移动 Agent 技术等。

1 相关工作

1.1 Netfilter/iptables 框架

Netfilter[5,6]是 Linux 2.4 内核中实现的第三代防火墙框架,它由一系列基于协议栈的钩子组成,内核模块可以对每种协议的一个或多个钩子进行注册挂载,实现对数据包的过滤、修改等功能。在 IPv4 协议中定义了 5 个钩子函数(hook),如图 1 所示。

IP 数据包从左边进入系统,NF_IP_PRE_ROUTING 和 NF_IP_POST_ROUTING 分别是数据包进入和离开系统的钩子函数;NF_IP_LOCAL_IN 和 NF_IP_LOCAL_OUT 分别是数据包进入和离开用户空间本地进程的钩子函数;NF_IP_FORWARD 是转发数据包的

钩子函数。数据包根据路由规则依次通过各个钩子函数,并检查各个点的过滤规则,并返回 NF_DROP 或 NF_ACCEPT,分别表示丢弃或接受该数据包。

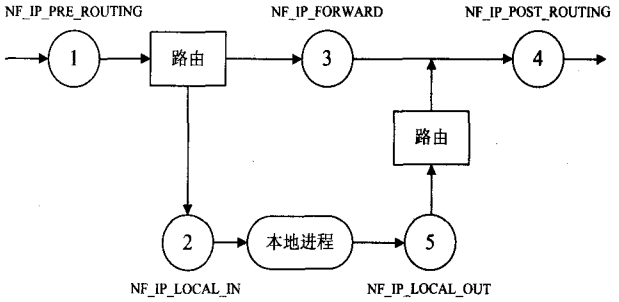


图 1　IPv4 中 Netfilter 的 5 个钩子函数

iptables 是 Netfilter 的用户配置工具,用来设置各个钩子点的过滤规则。

1.2　连接跟踪

在对流量进行识别分类时,以连接为单位比以数据包为单位进行识别分类的效率高很多。当模块识别了某连接的前几个数据包属于某协议类时,可以直接把后续数据包也都标记为此协议类。Netfilter 防火墙框架包含了连接跟踪[7]模块 nf_conntrack,此模块通过一个双向哈希链表记录所有连接的状态,从而实现跟踪管理各个连接。链表中记录了数据包的源地址、目的地址、源端口、目的端口及传输层协议这 5 个元组(tuple)生成的哈希值,以此来唯一标识网络中一个连接,如图 2 所示。

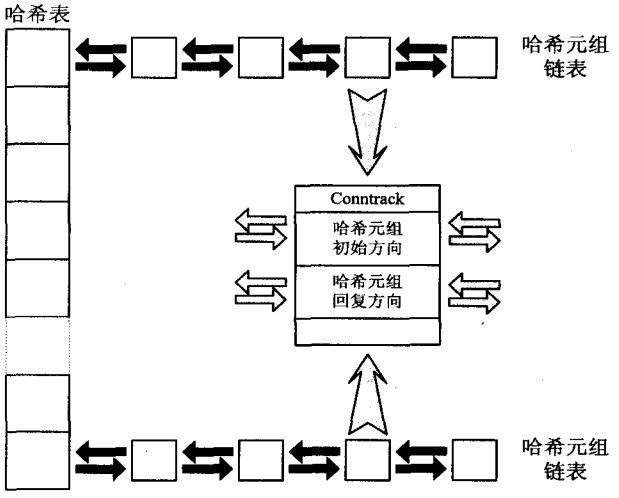


图 2　连接跟踪哈希表

1.3　协议识别 L7-filter 及流量控制 TC

L7-filter[8,9]是一个 iptables 的增强型补丁插件,L7-filter 作为 Linux 上的数据包识别分类器,和其他大多数分类器不同,它并不只是查看数据包的端口号、IP 地址等简单属性,而是使用正则表达式来匹配应用层数据,从而得知当前数据包属于何种协议并打上标记。因此,L7-filter 能够按不同应用层协议对数据包分类,特别是能对 P2P 协议高效识别。目前 L7-filter 不仅默认能识别实际网络中的上百种协议,还可以通过添加匹配模式文件方便地进行扩展。

Linux 在发送数据包中加入了流量控制模块,它通过命令行用户接口 TC 来控制。该工具可以让用户自由配置流量控制框架的 3 个要素:排队规则(qdisc)、类(class)及过滤器(filter)。TC 将流经网络接口的数据包先放入队列中,然后过滤器把数据包放入不同分类中,最后通过控制每个分队列数据包发送的速率限制每个分类的带宽。

2　流量检测与控制系统的设计与实现

本系统分为协议识别分类和流量控制两个模块,首先协议识别分类模块使用 L7-filter 识别数据包的应用层协议[10],然后用 iptables 的 CLASSIFY 目标对识别出的不同协议数据包进行分类,最后流量控制模块用 TC 对各类流量分配不同的带宽和优先级,从而实现流量控制。

2.1　协议识别分类模块的设计与实现

协议识别分类模块使用 L7-filter 对流经 Netfilter 防火墙的数据包进行基于连接跟踪的应用层协议检测,并用 iptables 的 CLASSIFY 目标对协议分类,按服务流量的重要性把流量分成 6 类:第一类为 SSH、Telnet、DNS 和带有 SYN、ACK 标记的数据包,第二类为 HTTP 协议的数据包,第三类为 SMTP 和 POP3 的邮件协议,第四类为 FTP 协议,第五类为电驴、BT、迅雷等 P2P 协议,其他的默认为第六类。

例如,把 edonkey 协议的数据包分到 1:14 类的命令为:

```
iptables -t mangle -A POSTROUTING -m layer7 --l7proto edonkey -j CLASSIFY --set-class 1:14
```

2.2　流量控制模块的设计与实现

每块网卡都有一个出口根排队规则,每个排队规则都指定一个句柄,句柄由一个主号码和一个次号码组成。本系统中 TC 排队规则使用 HTB[11,12](Hierarchical Token Bucket)即分类的令牌桶过滤器,HTB 能对一个固定速率的链路分成多种不同的用途使用,为每种用途做出带宽承诺并实现定量的带宽借用。HTB 还可以设定带宽分配的优先级。每个流量类有以下参数:AR(assured rate):保证的最小带宽值,TC 中用 rate 表示;CR(ceil rate):可获得的最大带宽值,TC 中用 ceil 表示;P(priority):优先权,该值为 0 时优先权最大,数值越大优先权越小,TC 中用 prio 表示;Q(Quantum):控制带宽借用的参数,通常取系统默认值。则对于流量类 c 的实际带宽值 R_c 的表达式为:

$$R_c = \min(CR_c, AR_c + B_c) \quad (1)$$

其中 B_c 为借用带宽比,其表达式为:

$$B_c=\begin{cases}\dfrac{Q_cR_p}{\sum\limits_{i\in D(p)}(Q_i\mid P_i=P_c)},\min(P_i\mid i\in D(p))\geqslant P_c\\ 0,\text{其他}\end{cases}\tag{2}$$

其中p为c的父类,$D(p)$为所有需要从p类借用带宽的子类。由以上表达式可知,流量类的带宽至少为AR,而优先级高的类优先获得剩余带宽。

根据应用需求,首先对不同流量类创建一个HTB树,见图3。

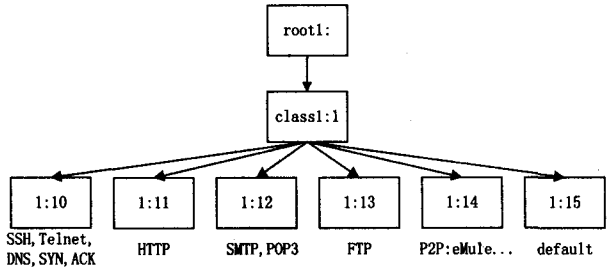


图3　流量控制模块的HTB树

然后利用TC对不同协议流量类设定流量控制参数,相关命令为:

tc qdisc add dev eth0 root handle 1: htb default 15

此命令为网络接口eth0绑定一个HTB队列,并且指定一个名称为handle的句柄,用1:htb标识它下面的子类,没有被分类的流量被分配到类1:15。

tc class add dev eth0 parent 1: classid 1:1 htb rate2048kbit ceil 2048kbit

此行命令为刚才建立的队列建一个主干类,带宽为2048kbit,最大速率为2048kbit。

tc class add dev eth0 parent 1:1 classid1:10 htb rate 600kbit ceil 2048kbit prio 0

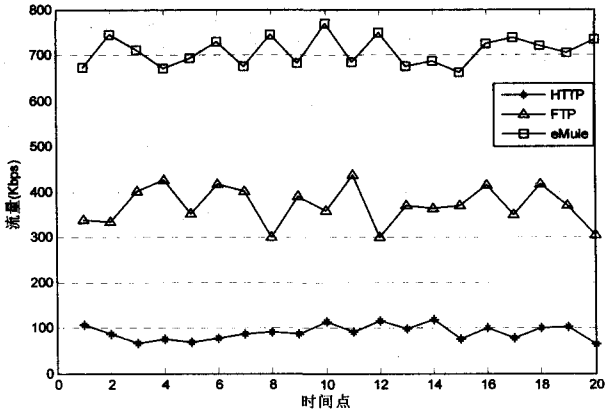
此命令建一个类1:10,带宽为600kbit,最大速率为2048kbit,优先级为0。这是一个最高优先级的类,这个类中的数据包拥有最低延迟并最先取得空闲带宽。把SSH、Telnet、DNS和带有SYN、ACK标记的数据包这些要求低延迟的服务归属到该类中。对其他5类数据流设定相应带宽和优先级(见表1)。结合前面用iptables对数据包进行分类的相关命令就可以实现对不同应用层协议的流量进行分类和控制。

表1　各类流量HTB参数设置

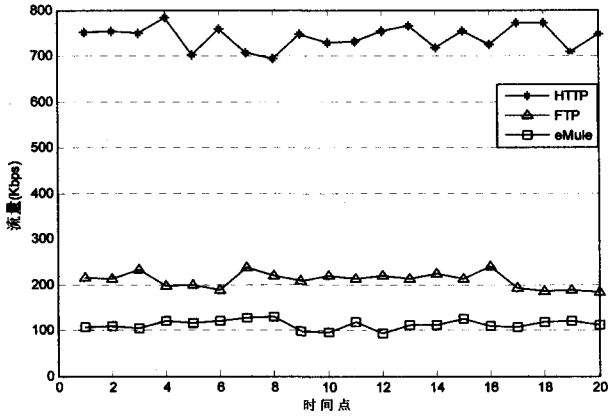
流量类	1:11	1:12	1:13	1:14	1:15
rate	600	400	200	100	100
ceil	2048	2048	600	300	300
prio	1	2	3	4	5

2.3　系统测试

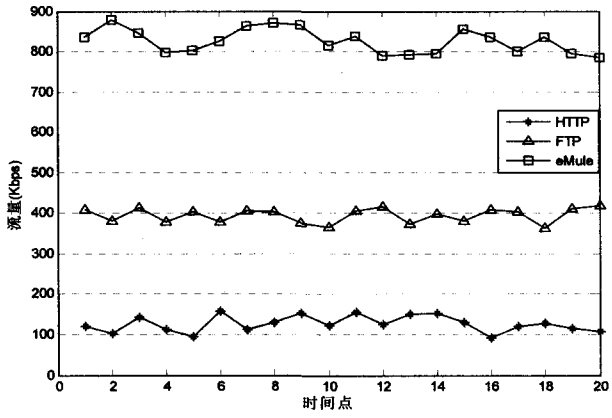
在实验室的Linux网关服务器上安装以上设计的流量控制系统,通过一条2Mbps的链路连接到公网。在内网客户端上运行Web浏览器、FTP和eMule客户



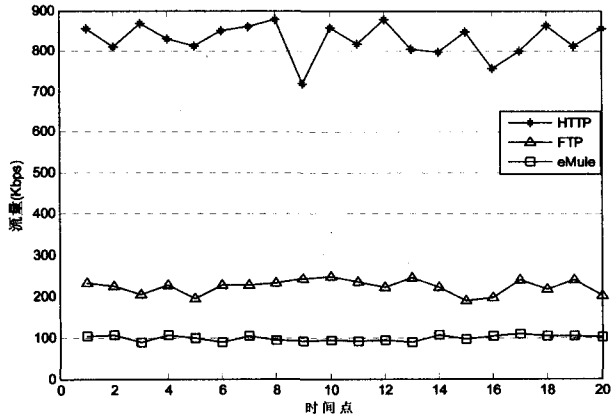
(a) 在10:00系统运行前



(b) 在10:00系统运行后



(c) 在20:00系统运行前



(d) 在20:00系统运行后

图4　流量控制系统运行前后的流量变化

端程序,数据采集时间段分别为 9 点至 11 点和 19 点至 21 点,每个小时取 20 个测量点,分别在 10 点和 20 点开始运行流量控制系统,得到系统运行前后一个小时的流量变化如图 4 所示。

由实验可知,FTP 和 P2P 的流量有效控制在 200kbps 和 100kbps 左右,与系统 HTB 设置一致。实验证明,此流量控制系统可以准确识别各种应用层协议并有效地进行流量控制。

3　结束语

文中基于 Linux 强大的 Netfilter 防火墙和流量控制器 TC,设计并实现了一个应用层协议流量控制系统。L7-filter 工具能够对数据包进行高效而准确的应用层协议识别与分类,TC 能对每类数据包进行灵活的流量控制。该系统安全、稳定、简单、高效并且价格低廉,对中小型网络来说是一个很好的解决方案。

参考文献:

[1] 吴国庆. 对等网络技术研究[J]. 计算机技术与发展,2008,18(7):100-104.

[2] 吴　敏,王汝传. 基于主机的 P2P 流量检测与控制方案[J]. 计算机技术与发展,2009,19(10):26-29.

[3] 蒋海明,张剑英,王青青,等. P2P 流量检测与分析[J]. 计算机技术与发展,2008,18(7):74-76.

[4] 杨　勇,王雪晶,陈良臣. QoS 在 IP 中的研究与应用[J]. 计算机技术与发展,2007,17(5):33-36.

[5] The netfilter. org project[EB/OL]. 2010. http://netfilter.org/.

[6] 徐苏磊,梁　伟. 基于 Netfilter/Iptables 内核扩展的 P2P 流量管理[J]. 计算机技术与发展,2010,20(6):101-104.

[7] Ayuso P N. Netfilter's Connection Tracking System[J]. LOGIN:The USENIX Magazine,2006,32(3):34-39.

[8] Application Layer Packet Classifier for Linux[EB/OL]. 2009. http://l7-filter. sourceforge. net/.

[9] Othman M,Kermanian M N. Detecting and preventing peer-to-peer connections by Linux iptables[C]//International Symposium on Information Technology. Kuala Lumpur, Malaysia:[s. n.],2008:1-6.

[10] 张五生,郑灵翔. 基于 Linux 的流量控制系统研究[J]. 厦门大学学报(自然科学版),2010,49(1):38-42.

[11] Hubert B. Linux advanced routing & traffic control HOWTO[EB/OL]. 2004-03-31. http://lartc. org/.

[12] Devera M. Hierachical token bucket theory[EB/OL]. 2002-05-05. http://luxik. cdi. cz/ ~ devik/qos/htb/manual/theory. htm.

(上接第 89 页)

3　结束语

在 3DS MAX 中建立基本实体模型,并在 VC++和 DirectX 开发环境中实现事件流程的仿真控制,开发周期短、效率高,可以取得较好的仿真效果。根据以上方法生成的某型导弹虚拟战场演示系统(见图 2 ~ 5),场面逼真,交互性好,沉浸感强,为相关导弹部队加快战斗力的生成提供了新的平台。为进一步完善该系统,仍有许多工作要做,比如对具体地图地形数据及对应纹理的快速获取技术、道路数据采集及作战要素部署简化技术、导弹战场环境完整作战过程展现技术等,这些都是以后要重点研究的内容。

参考文献:

[1] 黄安祥. 空战虚拟战场设计[M]. 北京:国防工业出版社,2007.

[2] 胡　令,关正西. 导弹武器发射环境视景仿真研究[J]. 信息化纵横,2009(5):66-68.

[3] 郭齐胜,董志明. 战场环境仿真[M]. 北京:国防工业出版社,2005.

[4] 许建中. 虚拟战场环境的三维地形构建技术研究[D]. 上海:上海交通大学,2007.

[5] 王克伟,杨　帆. 3ds max8 完全自学手册[M]. 北京:中国林业出版社,2006.

[6] 左小清. 公路三维模型建立与数据组织[J]. 武汉大学学报:信息科学版,2004,29(2):179-183.

[7] Thomas C M,Featherstone V J E. Validation of vincenty's formulas for the geodesic using a new fourth-or-der extension of kiviojas formula[J]. Journal of Surveying Engineering,2005,131(1):20-26.

[8] 朱　庆. 道路网络模型研究综述[J]. 中国学术期刊文摘,2007(20):4-4.

[9] 张　毅. 弹道导弹弹道学[M]. 长沙:国防科技大学出版社,1999.

[10] Chiou Y C,Kou C Y. Geometric approach to three dimensional missile guidance problem[J]. Journal of Guidance, Control and Dynamics,1998,21(2):199-205.

[11] 刘　月,刘文杰,刘　敏. 突发威胁下的航迹规划问题研究[J]. 飞行力学,2009,27(5):89-92.

[12] Dorigo M,Gambardella L M,Middendorfm M,et al. Guest editorial:special section on ant colony optimization[J]. IEEE Transactions on Evolutionary Computation,2002,6(4):317-319.

[13] 管　华,王双亭,王　净. 飞行物体虚拟环境仿真系统的研究[J]. 测绘学院学报,2003,20(1):29-31.