doi:10.3969/j.issn.1000-436x.2013.z2.004

用户自主选择的校园网出口策略路由实现

张焕杰，夏玉良
(中国科学技术大学 网络信息中心，安徽 合肥 230026)

摘 要：校园网有多个出口时，出口路由设备必须要对校内 IP 发出的数据分组进行策略路由处理，才能保证正常的通信。传统由网络管理员设置的源地址策略路由灵活性不够，无法满足用户多样变化的需求。因此提出由用户根据使用需要，自主选择所用 IP 地址的策略路由，在网络出口设备上对校园网发出的数据包进行正确的路由处理，实现更灵活和方便的校园出口策略路由。通过 10 年校园网的实际使用，证明该方式能满足用户的各种需求，运行稳定可靠。

关键词：多出口；策略路由；自主选择
中图分类号：TP393.18 **文献标识码**：A **文章编号**：1000-436X(2013)Z2-0014-03

User demanded policy routing for multi-homing campus network

ZHANG Huan-jie, XIA Yu-liang
(Network and Information Center, University of Science and Technology of China, Hefei 230026, China)

Abstract: Edge router must use policy routing for multi-homing campus network. The flexibility of source address based policy routing defined by network administrator is not sufficient to meet the changing needs of diverse users. One scheme using Linux system was proposed to do user demanded policy routing, which was more flexible and convince than the traditional administrator defined source address based policy routing. Through 10 years of actual use in campus network, It proves that this method can meet the various needs of users, operating stabely and reliablely.

Key words: multi-homing; policy routing; user demanded

1 引言

为了提高校园网对外的灵活性和通信速度，不少学校除连接中国教育科研网(CERNET)外，往往还连接有中国电信、中国联通等运营商的网络。

一个典型的连接 3 个网络出口的校园网结构如图 1 所示。图中的出口设备分别连接 3 个网络出口和内部主干交换设备。

本文提出采用 Linux 系统作为出口设备，在 IP rule 提供策略路由的功能上进行优化，实现用户自主选择出口策略路由的功能。

2 校园出口的策略路由需求

在如图 1 所示连接方式下，处理校内计算机对外的访问比较简单，并不需要设置策略路由。校内的计算机多使用由 CERNET 分配的 IP 地址。通过在出口设备上增加静态路由，把中国电信/中国联通网络地址段的下一跳地址设置为中国电信/中国联通的出口网关，使得校内计算机在访问中国电信/中国联通地址段时，能利用中国电信/中国联通公网出口，并由出口处的 NAT 设备转换成中国电信/中国联通提供的 IP 地址对外通信。

而校外发起对校内服务器的通信访问，服务器发送的数据分组在网络出口处的路由处理则相对复杂，不能简单地按照目的地址路由，通常的做法是由网络管理员设置基于源地址的策略路由[1,2]。某台服务器，如 DNS 服务器，希望外部用户从 CERNET 出口访问，需在出口设备上设置源地址策略路由，使得该服务器发出的数据分组永远从 CERNET 出口发出，任何时候都不会选择其他出口。

收稿日期：2013-09-04

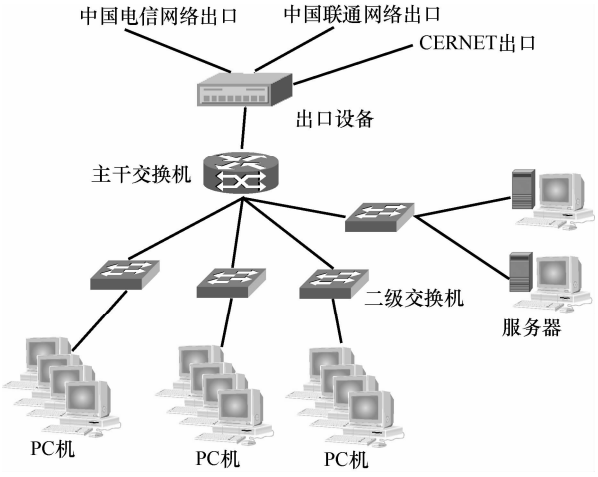


图 1　3 个网络出口校园网拓扑

这种由网络管理员设置的源地址策略路由有 2 个弊端。

1) 校内服务器较多时，网络管理员的维护工作量增大，无法应对用户的频繁更改需求。

2) 校内用户只能接受网络管理员的设置，对网络出口的选择没有自主权。

当校园出口较多并且用户的需求很分散时，网络管理员很难设置用户满意的出口路由策略，用户的意见较大。

3　用户自主选择的策略路由

为提高校园网出口策略路由的灵活性，中国科学技术大学采用用户自主选择路由方式，具体做法是：Linux 服务器作为出口设备，网络管理员设定若干基于目的 IP 地址的路由表，由校内用户决定所用 IP 对外通信时使用其中的哪一个路由表。用户通过 Web 界面，可以随时按自己的意愿修改路由表的选择。

Linux 服务器支持 1~255 共 255 个路由表，系统默认使用其中的 253、254、255 这 3 个路由表，管理员可以设置剩余的 252 个路由表。

中国科学技术大学出口设置了 10 个路由表，根据需要可以随时扩充，典型的路由表内容如表 1 所示。

用户选择路由表 100，其 IP 发出所有对外的数据分组交由教育网关，即完全使用教育网线路。用户选择其他路由表，根据目的地址的不同，发出的数据分组会交给不同的设备处理，从而使用不同的出口线路（使用中国电信/中国联通等线路时，需要经过 NAT 设备进行地址转换处理）。

表 1　　校园出口路由

编号	目的地址	下一跳
100	默认	教育网关 IP
101	教育网	教育网关 IP
	默认	电信网出口 NAT IP
103	国内免费 IP	教育网关 IP
	默认	电信网出口 NAT IP
105	教育网	教育网关 IP
	中国电信 IP	电信网出口 NAT IP
	默认	联通网出口 NAT IP
108	中国电信 IP	电信网出口 NAT IP
	中国联通 IP	联通网出口 NAT IP
	默认	教育网关 IP

用户可以根据需求随意修改路由表的选择。如需外网用户访问他的计算机时，可以选择路由表 100。正常上网希望使用 CERNET 线路访问国际图书文献资源时，则选择路由表 108 会比较合适。通过将选择权交给用户，能提高用户满意度，减少用户与管理员的矛盾。

4　Linux 下大规模策略路由的实现

Linux 系统中，使用 IPrule 命令可以维护策略表[3]，实现不同源 IP 使用不同路由表的功能。系统的默认实现，在查找这些策略时使用简单的顺序查找方式，因此一旦设置较多的策略条目，查找的效率会很低，并不适合在校园出口使用。

考虑到校园网内部 IP 地址分为若干块，利用一个数组存放一个 IP 地址块的路由表信息。以中国科学技术大学共有 7 段 IP 地址为例，其主要数据结构如表 2 所示。

表 2　　校内 IP 策略路由表数据结构

地址	掩码	路由表数组指针
114.214.160.0	255.255.224.0	长度 8 192 的数组
114.214.192.0	255.255.192.0	长度 16 384 的数组
202.38.64.0	255.255.224.0	长度 8 192 的数组
210.45.64.0	255.255.240.0	长度 4 096 的数组
210.45.112.0	255.255.240.0	长度 4 096 的数组
211.86.144.0	255.255.240.0	长度 4 096 的数组
222.195.64.0	255.255.224.0	长度 8 192 的数组

对于114.214.160.0/255.255.224.0地址段，一共有8 192个IP地址，长度8 192的数组中，每个单元存放一个IP地址选择的路由表编号。如数组中的第二个单元存放的是100，则说明114.214.160.1选择的是100号路由表；第3个单元存放的是108，则说明114.214.160.2选择的是108号路由表。利用以上数据结构，查询次数仅仅与校内IP地址段数有关，而与设置的策略多少无关。对于中国科学技术大学环境，最多查询7次即可以得到对应IP选择的路由表编号。

上述数据结构存放在Linux kernel中，使用proc文件系统来完成应用程序与kernel的通信过程，以便修改或读取其中的选择信息。

系统启动时，注册控制文件/proc/iprule/control，对该文件写特定的命令可以修改数据结构中的内容。支持的写命令共2条，如表3所示。

表3 /proc/iprule/control写命令

写命令	含义
A 202 38 64 0 255 255 240 0	增加一个校内IP地址段
C 100 202 38 64 1	设定202.38.64.1选择路由表100

增加一个校内IP地址段命令执行后，会在/proc/iprule下添加2个文件，供程序获取IP的路由表选择信息。上述例子命令执行后在/proc/iprule下自动添加2个文件：202.38.64.0_255.255.224.0和B202.38.64.0_255.255.224.0。读取文本文件202.38.64.0_255.255.224.0可以得到每个IP的路由选择信息，内容如下：

100 202.38.64.1
100 202.38.64.2
100 202.38.64.3
100 202.38.64.4
100 202.38.64.5

B202.38.64.0_255.255.224.0是长度为8 192的二进制文件，其中，每个字节是对应路由表数组的内容，更加方便程序的处理。

5 用户自主选择路由界面

用户通过Web界面来修改所用IP地址的路由选择，该界面同时实现Web出口认证功能。

采用C语言开发CGI程序实现界面功能，用户的认证信息如用户名、密码等信息存放在MySQL数据库中。CGI程序确认用户身份后，按照用户权限和选择设定相应的路由信息，实现用户自主选择路由的功能。其工作截图如图2所示。

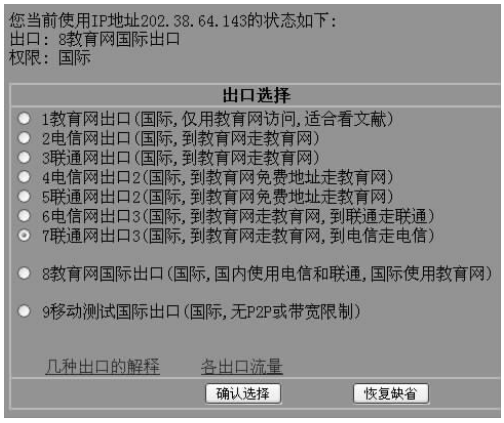


图2 用户自主选择路由界面

6 系统优化与扩展

中国科学技术大学2003年2月17日开始使用上述用户自主选择的校园出口策略路由系统，至今已运行超过10年。早期的kernel程序基于Linux kernel 2.4开发，2010年7月移植到Linux kernel 2.6。10年的使用过程中系统运行稳定，即使偶尔碰到异常DoS流量有过高的PPS(packet per second)导致系统工作慢，一旦异常流量消失，系统立即恢复正常。

由于系统中不存放和使用UDP/TCP的连接信息，因此上述系统非常容易平行扩展。只要在出口处并行放置若干台Linux服务器，其中一台运行Web界面为主设备，其他服务器与主设备同步/proc/iprule目录下文件，并在出口路由器上增加等代价路由到这些服务器，既可以实现平行扩展。

根据实际测试，单台服务器依据CPU性能的不同，可以支持1G-4G BPS的带宽。目前使用3台服务器，网络正常时，处理约3G BPS的带宽，一台10年前的服务器CPU利用率在40%左右，其余两台服务器CPU利用率均小于5%。

7 结束语

在多出口校园网上使用Linux系统作为出口设备，由用户自主控制所用IP地址的策略路由选择，可以实现更灵活的校园网出口策略路由功能，在不影响网络性能的前提下，满足校内用户对网络出口的多样性需求，提高用户的满意度。

（下转第22页）

路。在此基础之上，为了避免或者减少 AS 路径环路，本文还为 ISP 的 BGP 路由管理提出了建议。

参考文献：

[1] MOY J. RFC 2328: OSPF version 2[EB/OL]. http://www. ietf.org/rfc/rfc2328

[2] ORAN D, RFC 1142: OSI IS-IS intra-domain routing protocol[EB/OL]. http://tools.ietf.org/rfc/rfc1195.txt.

[3] MALKIN G. RFC 2453: RIP version 2[EB/OL]. http://www.ietf.org/rfc/rfc2453.txt.

[4] EKHTER Y, LI T, RFC 1771: border gateway protocol 4[EB/OL]. http://www.ietf.org/rfc/rfc1771.txt.

[5] PEI D, WANG L, MASSEY D, *et al*. A study of packet delivery performance during routing convergence[A]. Proceedings of IEEE International Conference on Dependable Systems and Networks[C]. San Francisco, USA, 2003.

[6] DEERING S, HINDEN R. RFC 2460: Internet protocol, version 6 (IPv6) Specification[EB/OL]. http://tools.ietf.org/ html/ rfc2460.

[7] PEI D, ZHAO X, MASSEY D, *et al*. A study of BGP path vector route looping behavior[A]. Proceedings Internet Conference on Distributed Computing Systems[C]. Tokyo, Japan, 2004. 720-729.

[8] MAHAJAN R, WETHERALL D, ANDERSON T. Understanding BGP misconfiguration[A]. Proceedings of the 2002 SIGCOMM[C]. Pittsburgh, PA, 2002.

[9] SHI X, XIANG Y, WANG Z, *et al*. Detecting prefix hijackings in the Internet with argus[A]. Proceedings of the 2012 ACM Conference on Internet Measurement Conference[C]. Boston, USA, 2012.

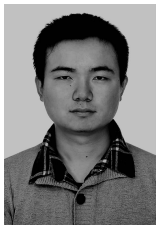
[10] University of oregon route views project[EB/OL]. http://www. route-views.org/

[11] ZHAO X, PEI D, WANG L, *et al*. An analysis of BGP multiple origin AS (MOAS) conflicts[A]. Proceedings of the 1st ACM SIGCOMM Internet Measurement Workshop[C]. Burlingame, USA, 2001.

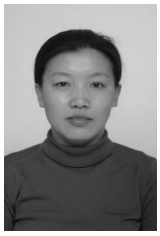
[12] CHIN K. On the characteristics of BGP multiple origin AS conflicts[A]. Telecommunication Networks and Applications Conference[C]. Christchurch, New Zealand, 2007.

[13] IPv6 global unicast address assignments[EB/OL]. http://www.iana.org/assignments/IPv6-unicast-address-assignments/IPv6-unicast-address-assignments.xhtml.

作者简介：



张圣林（1989-），男，山东滨州人，清华大学博士生，主要研究方向为下一代互联网体系结构。



刘莹（1973-），女，甘肃天水人，博士，清华大学副研究员，主要研究方向为下一代互联网体系结构、多播路由算法研究、多播路由协议设计、高性能路由器体系结构。

（上接第 16 页）

参考文献：

[1] 赵叶红等. NAT 环境下基于连接跟踪信息的策略路由[J]. 计算机应用，2006,26(7):1549-1551.

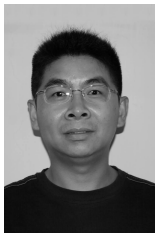
ZHAO Y H, *et al*. Policy routing based on connection tracking information in NAT[J]. Journal of Computer Applications, 2006,26(7): 1549-1551.

[2] 张焕杰等. 基于 Linux 系统的校园网多出口策略路由实现[J]. 通信学报, 2006,27(z1):130-133.

ZHANG H J, *et al*. Policy routing on Linux for multi-homing campus network[J]. Journal on Communications,2006,27(z1):130-133.

[3] Linux advanced routing & traffic control HOWTO [EB/OL]. http://lartc.org/howto/index.html, 2012.

作者简介：



张焕杰（1975-），男，安徽淮北人，中国科学技术大学高级工程师，主要研究方向为网络安全、网络管理等。

夏玉良（1975-），男，安徽岳西人，中国科学技术大学工程师，主要研究方向为软件开发、数据库等。