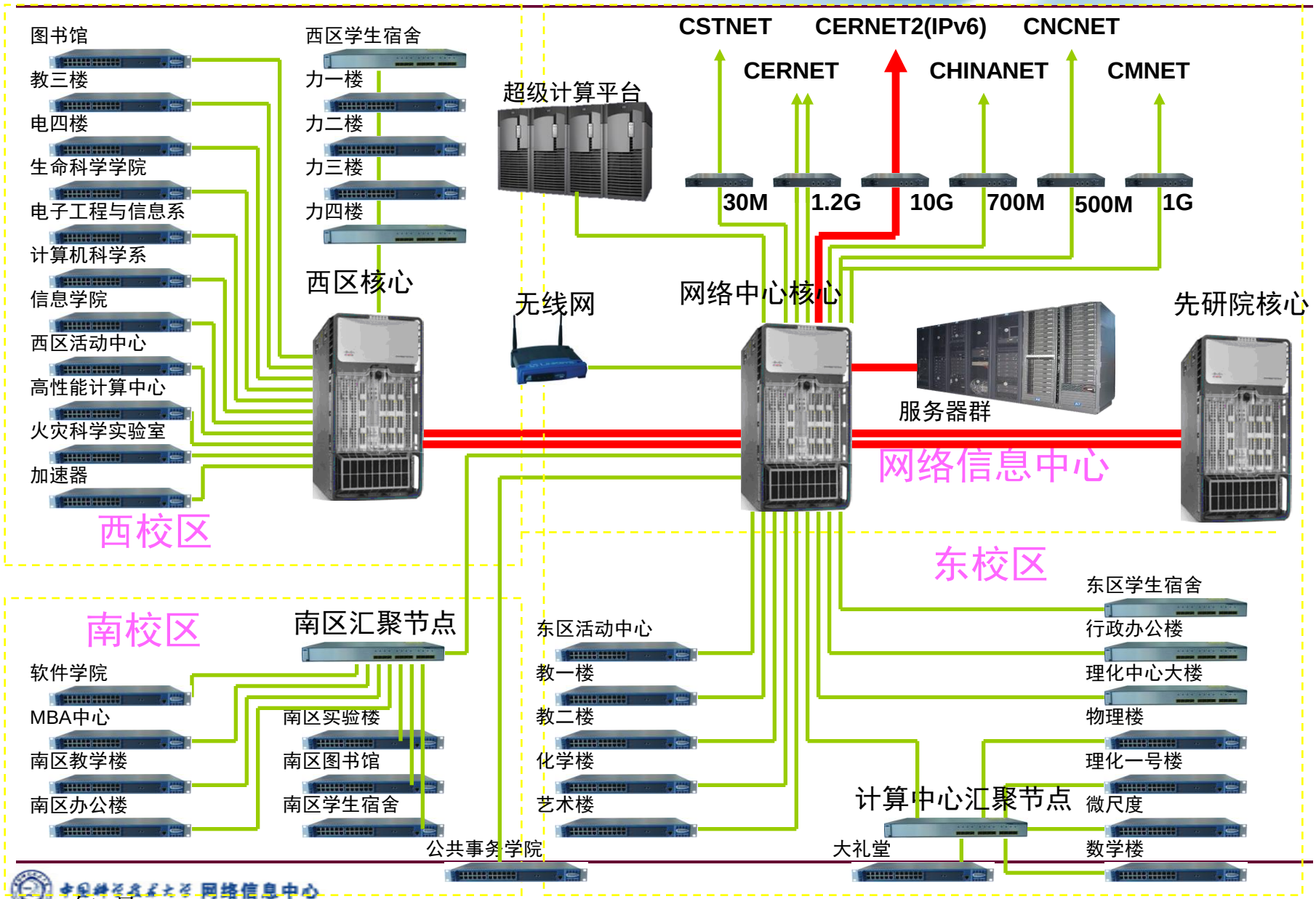
用户自主选择的校园网出口策略路由实现

张焕杰 james@ustc.edu.cn
中国科学技术大学 网络信息中心

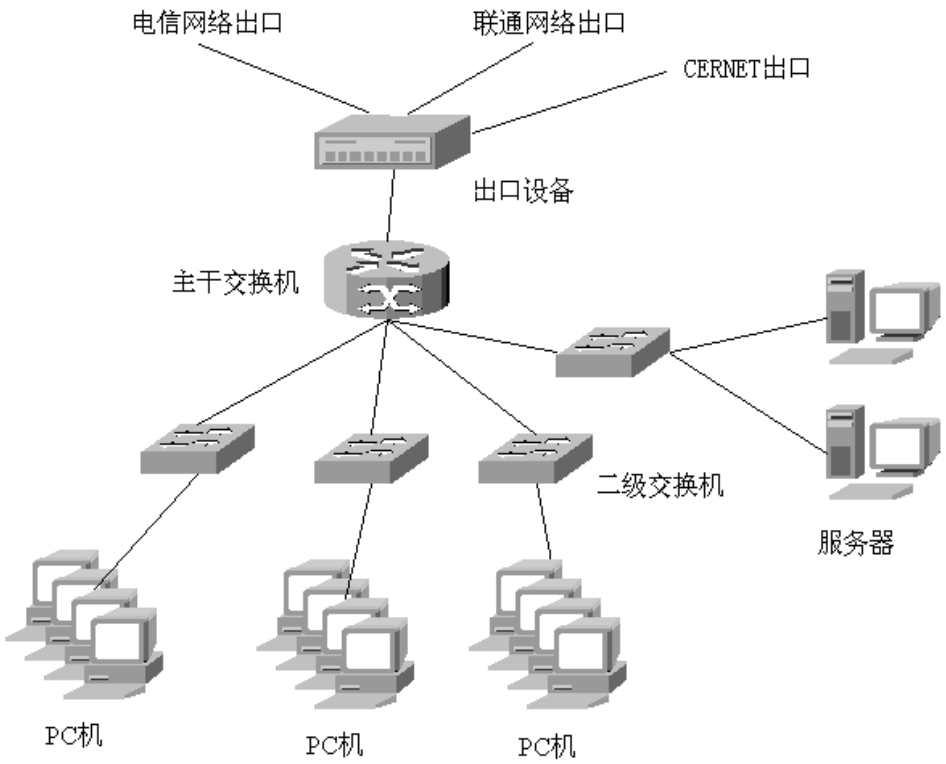
多出口缘由

- 校园网除连接教育网外，往往还连接电信、联通等网络
  - 国内的ISP仅仅在有限的几个点互连，导致跨ISP访问速度不理想：教育网、科技网、电信、联通、移动（铁通）
  - 不同ISP的计费策略不同
- 校园网出口多样和管理复杂
  - 很少有ISP像国外那样与客户提供BGP互连，而是仅仅提供一段各自的IP供客户使用
  - 从某ISP出口出去的数据包，要求源地址是该ISP分配的IP地址段，因此大部分情况下需要NAT和策略路由的应用

中国科学技术大学校园网络主干及出口示意图



典型的校园网出口示意图



- 出口设备需设置策略路由、地址转换等相关配置
- 由网络管理员设置，维护管理麻烦
- 校内用户无自由选择权，多样需求无法满足

网络通

文件(F) 编辑(E) 查看(V) 收藏(A) 工具(T) 帮助(H)

后退 搜索 收藏夹

地址(D) http://wlt.ustc.edu.cn/cgi-bin/ip?cmd=disp 转到 链接

用户james拥有的权限 刷新显示 使用中任何问题请联系网络信息中心用户服务部

序号	权限	状态	到期日	其他
1	VPN接入权限	开通		
2	拨号接入权限	开通		
3	完全访问权限	开通		最大连接数10

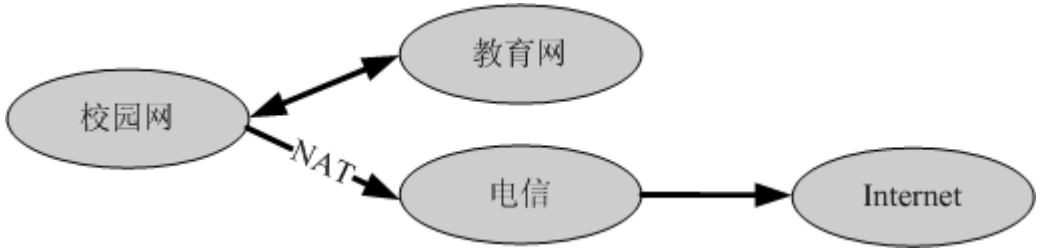
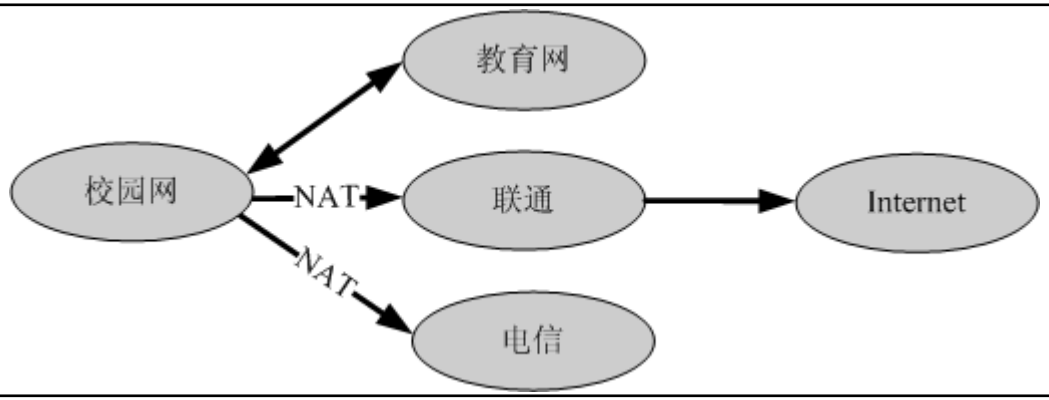
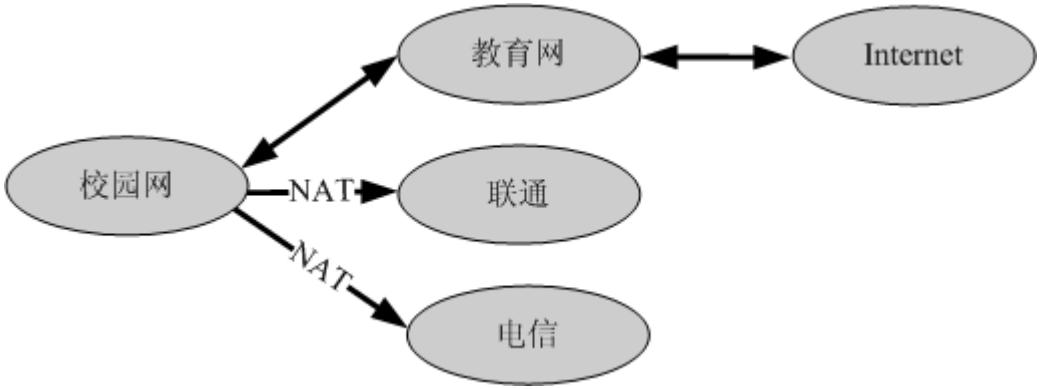
您当前使用IP地址202.38.64.143的状态如下:
MAC地址:00055D61BA54
出口: 网通网出口3
权限: 国内

出口选择	使用时限
○ 1 教育网出口(国内) ◉ 2 电信网出口(国际,到教育网走教育网) ○ 3 网通网出口(国际,到教育网走教育网) ○ 4 电信网出口2(国际,到教育网免费地址走教育网) ○ 5 网通网出口2(国际,到教育网免费地址走教育网) ○ 6 电信网出口3(国际,文献出口) ○ 7 网通网出口3(国际,到教育网走教育网,到电信走电信) 注: 选择出口2、3无法使用的某些电子资源, 使用出口4、5、6也许可以正常使用 几种出口的解释	○ 动态 ○ 1小时 ◉ 4小时 ○ 11小时
确认选择 恢复缺省	

选择 动态时, 本浏览器窗口关闭时自动恢复缺省

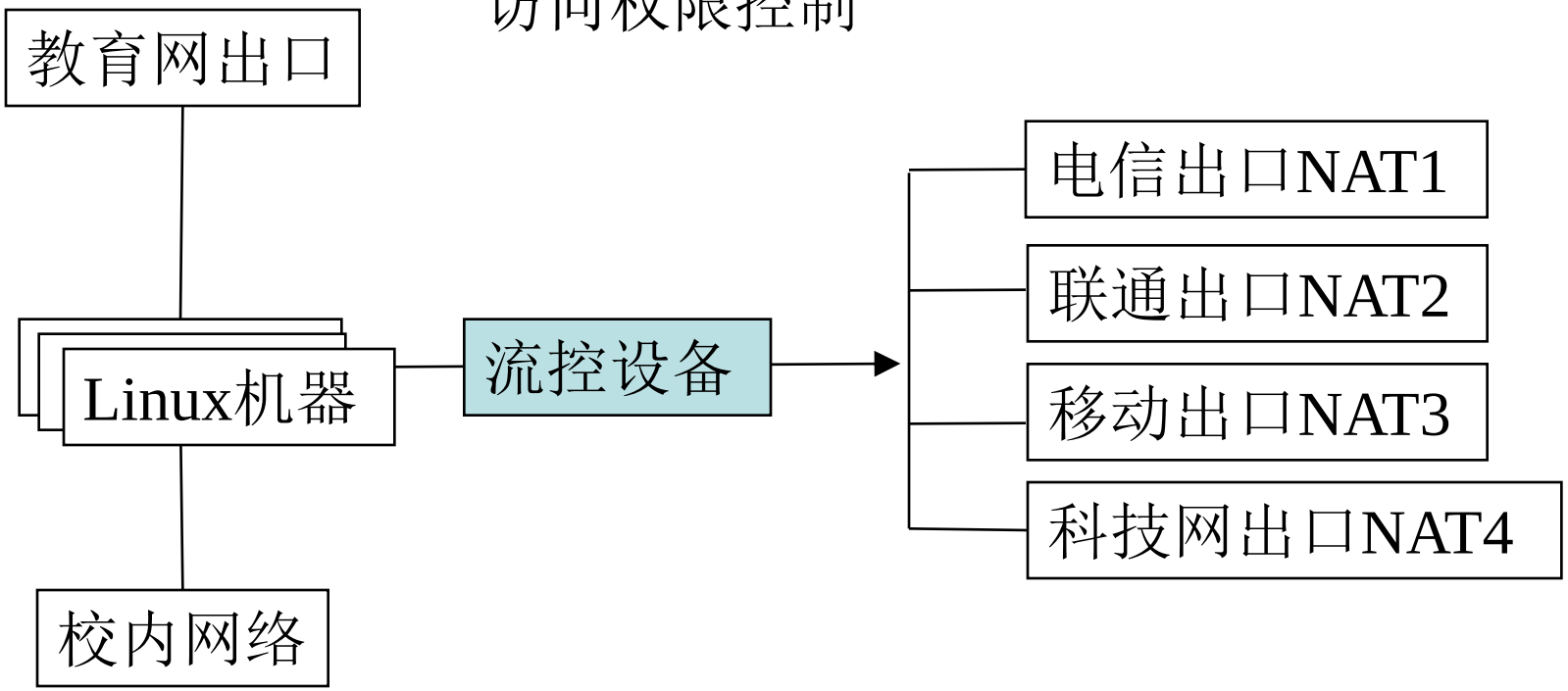
用户自己选择出口的界面

完毕 Internet

教育网出口		走教育网
电信网出口		教育网走教育网，其他NAT走电信
联通网出口3		教育网走教育网，电信地址走电信，其他联通
教育网出口2		电信地址经走电信，联通地址走联通，其他经过教育网

多出口网络示意图

Linux机器完成认证、策略路由、
访问权限控制



Linux机器利用ip route设置多个路由表
优化的程序将用户IP与路由表关联

Ip route 路由表示例

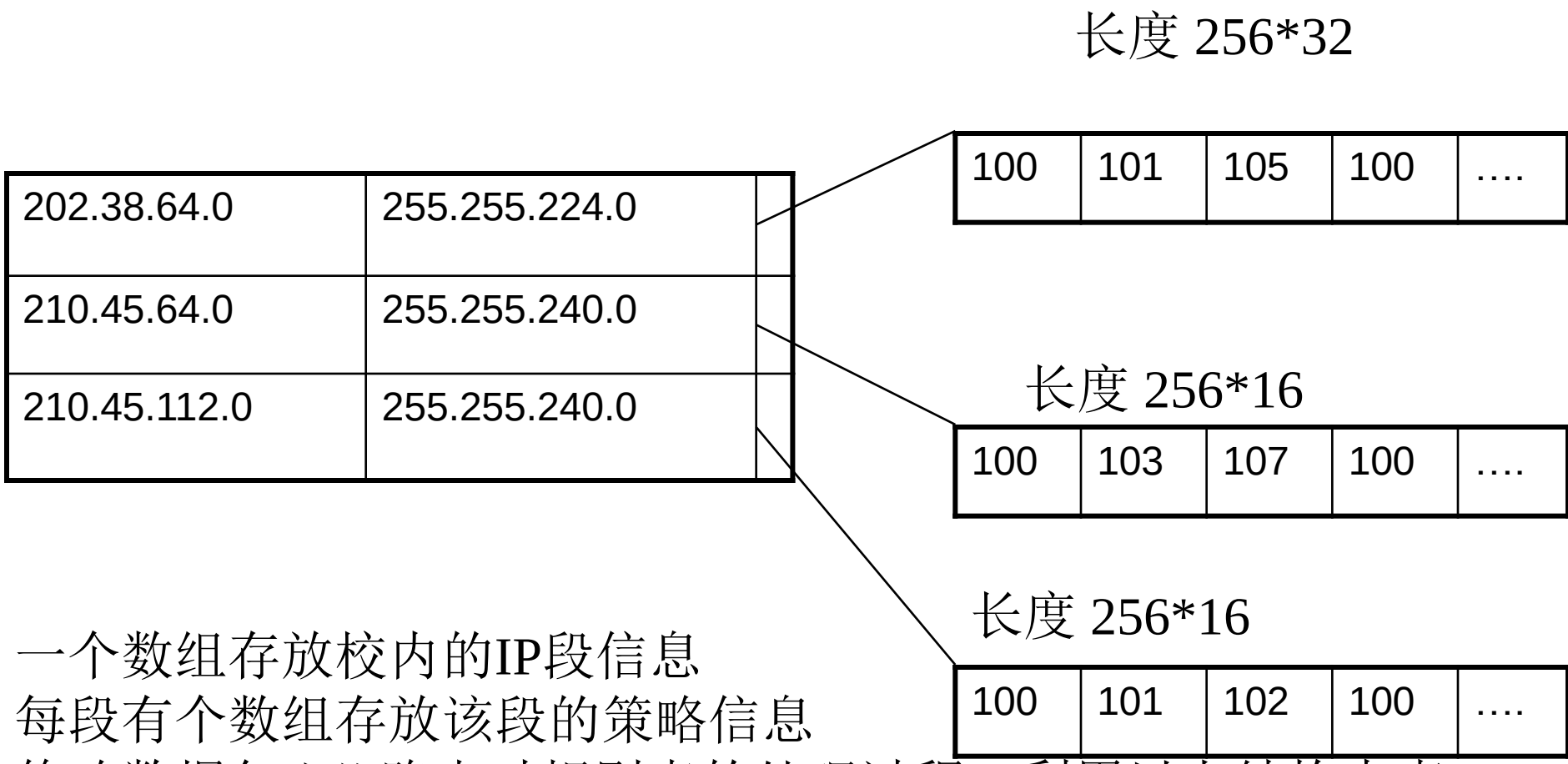
```
[root@gateway root]# ip route show table 100
default via 210.45.224.254 dev eth1.300

[root@gateway root]# ip route show table 101
default via 202.38.96.200 dev eth1.3

[root@gateway root]# ip route show table 102
default via 202.38.96.201 dev eth1.3

[root@gateway root]# ip route show table 103
118.123.232.0/22 via 210.45.224.254 dev eth1.300
202.4.252.0/22 via 210.45.224.254 dev eth1.300
122.10.160.0/22 via 210.45.224.254 dev eth1.300
123.151.172.0/22 via 210.45.224.254 dev eth1.300
…..
default via 202.38.96.200 dev eth1.3
```

内部的数据结构



一个数组存放校内的IP段信息
每段有个数组存放该段的策略信息
修改数据包查找路由时规则表的处理过程，利用以上结构来查

Ip rule 策略表示例

[root@gateway root]#ip rule

0: from all lookup local

20: from all lookup main

60: from all lookup 200

1000: from 255.255.255.255 lookup 199

32700: from all lookup 100

32766: from all lookup main

32767: from all lookup 253

修改ip rule处理程序，from 255.255.255.255转入处理流程

与kernel的通信

- 应用程序和kernel采用/proc文件系统通信，文件/proc/iprule/control为控制文件
- 写A 202 38 64 0 255 255 224 0到control,增加一段内部地址
  - 增加地址后/proc/iprule下会增加2个文件
    - 202.38.64.0_255.255.224.0和
    - B202.38.64.0_255.255.224.0
    - 读这两个文件可以得到某个IP的策略信息
    - 前面一个是ASCII的，方便管理员操作
    - 后面一个是binary的，偏移量0的字节就是202.38.64.0的策略，方便程序处理
- 写 C XXX 202 38 64 1到control修改策略，xxx为路由表
  - 如echo “C 101 202 38 64 51” > control
  - 设定202.38.64.51使用路由表101

校内选择不同路由表的IP数实例

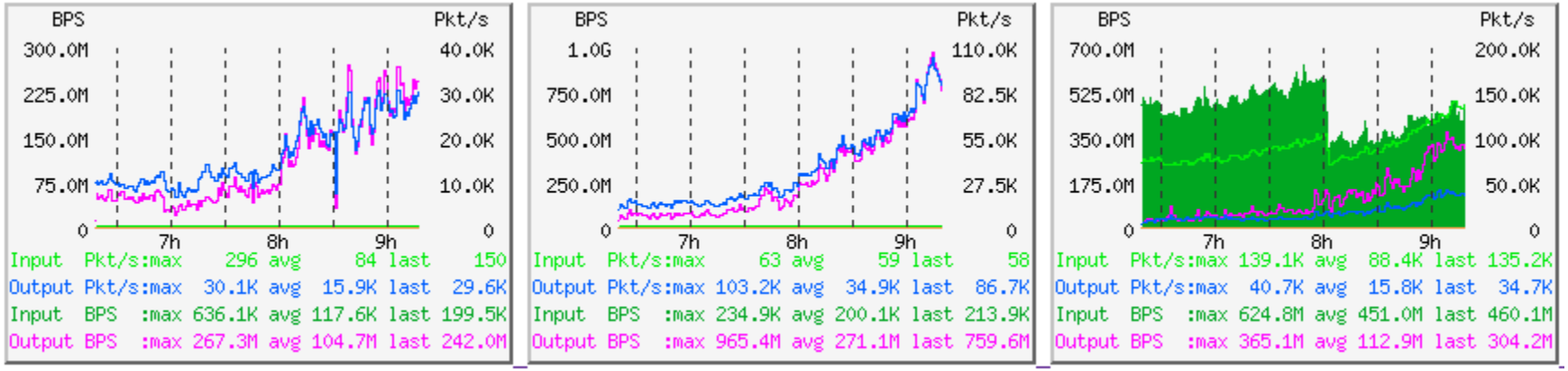
```
[root@gateway iprule]# cd /proc/iprule/
[root@gateway iprule]# ls
114.214.160.0_255.255.224.0 114.214.192.0_255.255.192.0
202.38.64.0_255.255.224.0 control
210.45.64.0_255.255.240.0 210.45.112.0_255.255.240.0
211.86.144.0_255.255.240.0 222.195.64.0_255.255.224.0
B114.214.160.0_255.255.224.0  B114.214.192.0_255.255.192.0
B202.38.64.0_255.255.224.0 B210.45.64.0_255.255.240.0
B210.45.112.0_255.255.240.0 B211.86.144.0_255.255.240.0
B222.195.64.0_255.255.224.0
[root@gateway iprule]# cut -f1 -d' ' [0-9]* | sort | uniq -c
     25 100
    706 101
    669 102
    162 103
     47 104
   2806 105
   3140 106
   2225 107
    158 108
```

系统优化与扩展

- 2003年2月17日开始使用，至今已运行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。
- 由于系统中不存放和使用UDP/TCP的连接信息，容易平行扩展。
  - 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。
- NAT设备对内的数据包，可以绕过Linux服务器直接送到校内，减少1/3的流量。

出口流量和服务器CPU利用率

- 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。
- 目前我校使用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
  - 其余两台近年购置的服务器CPU利用率均<5%



科大网络通出口

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。