DPDK
DATA PLANE DEVELOPMENT KIT

THE LINUX FOUNDATION

# 百万用户规模级vBRAS实践

# SCALING VBRAS TO MILLION-USER NETWORK

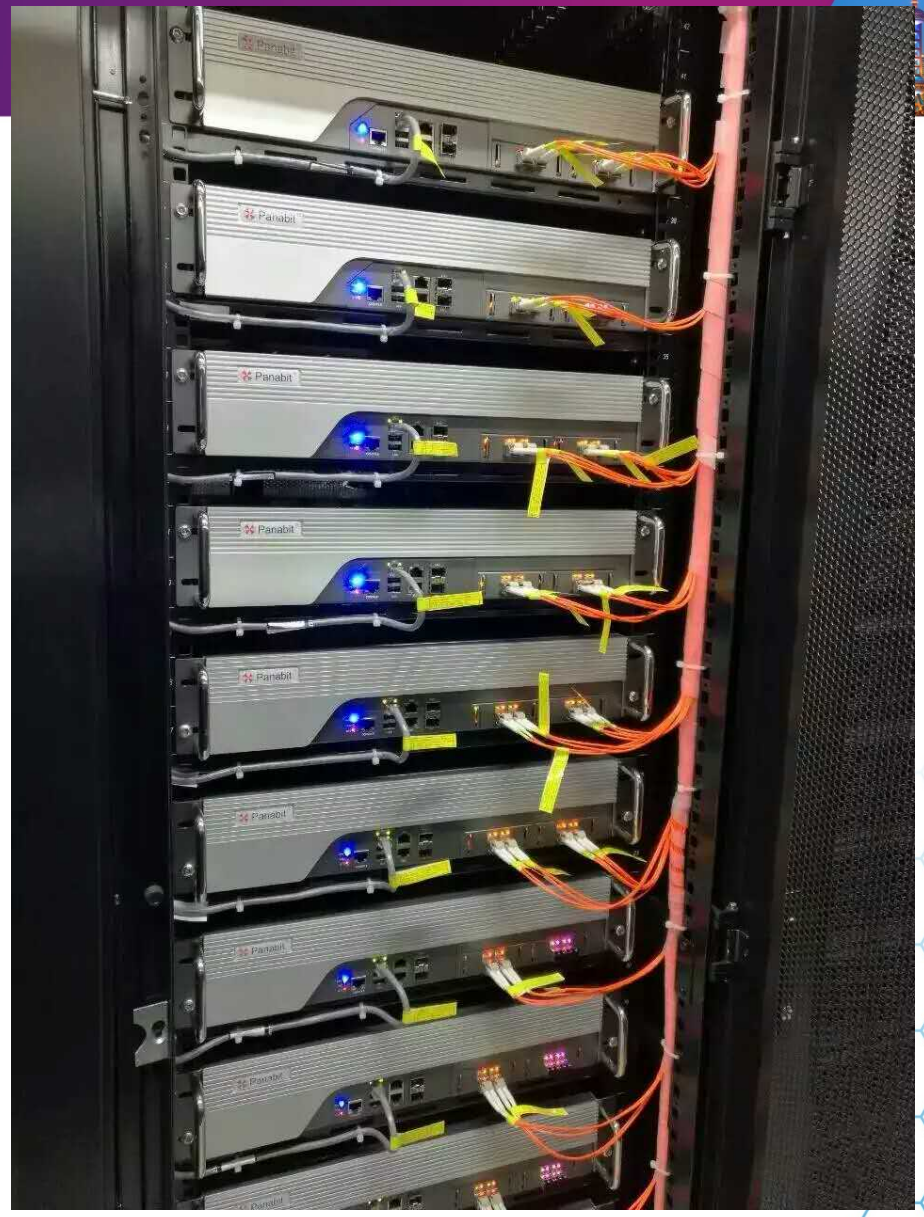
北京派网软件有限公司　孙朝晖

主办方：intel

参与方：腾讯云　ZTE　美团云　Panabit　太一星晨　UnitedStack有云　云杉网络 Yunshan Networks

协办方：SDNLAB 专注网络创新技术　　视频支持方：IT大咖说

The number of Online user supported by Panabit vBRAS has succeeded 5000,000



# vBRAS

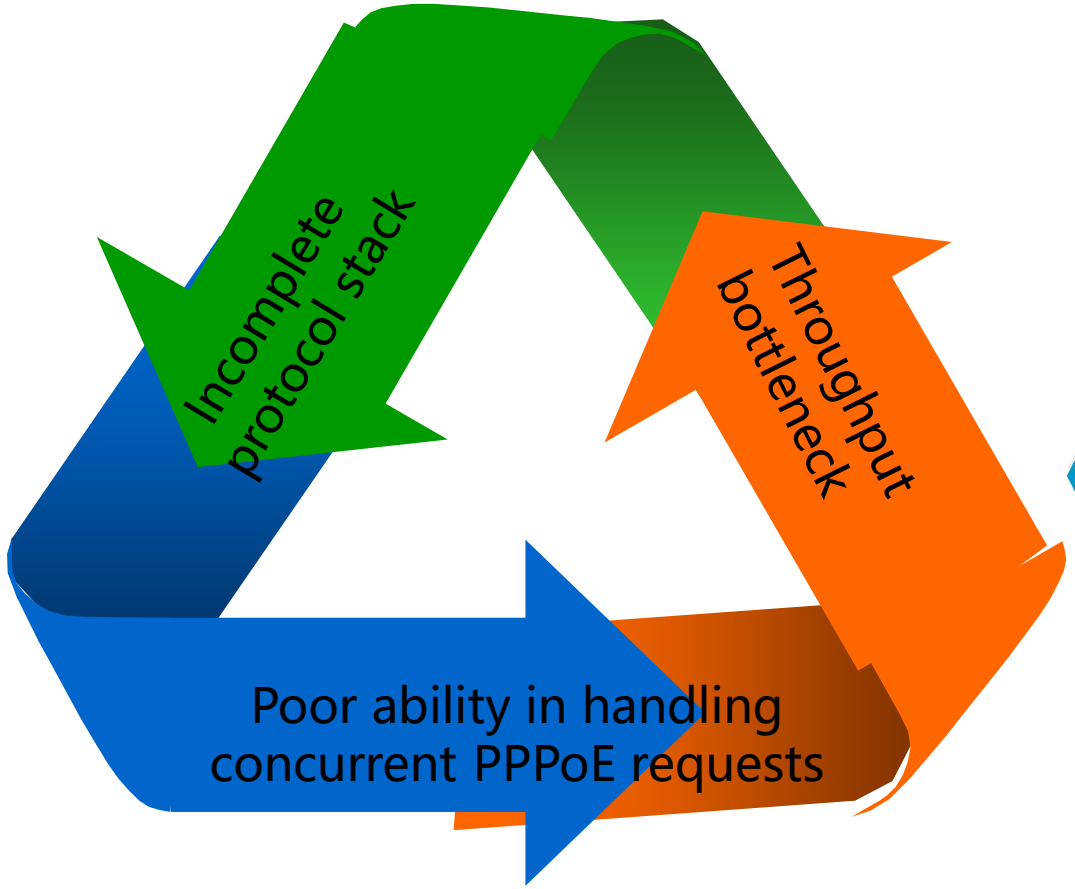
Migrating to virtualization

Industry consensus

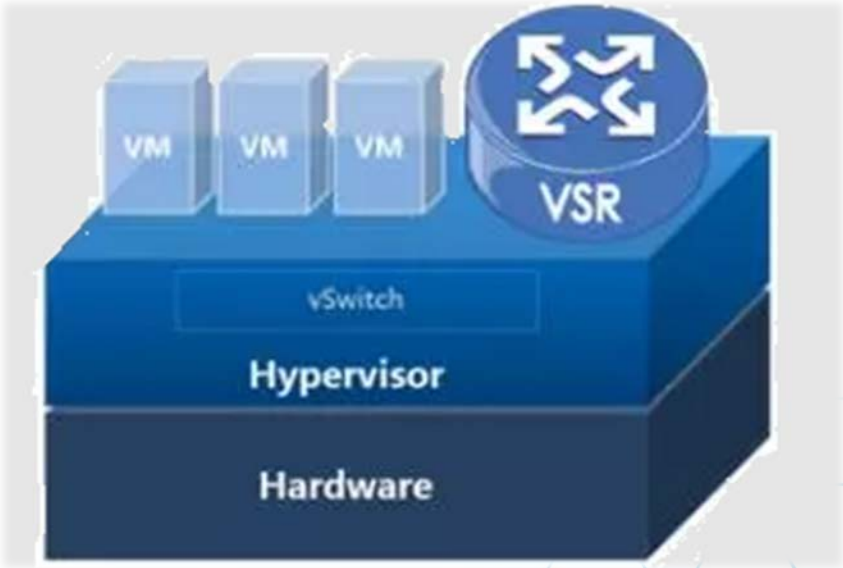
- Tightly coupoled hard&software
- Poor scability
- Low reuse density of resources
- Poor elasticity & flexibility

## Traditional BRAS

# Pain point of large operator's vBRAS



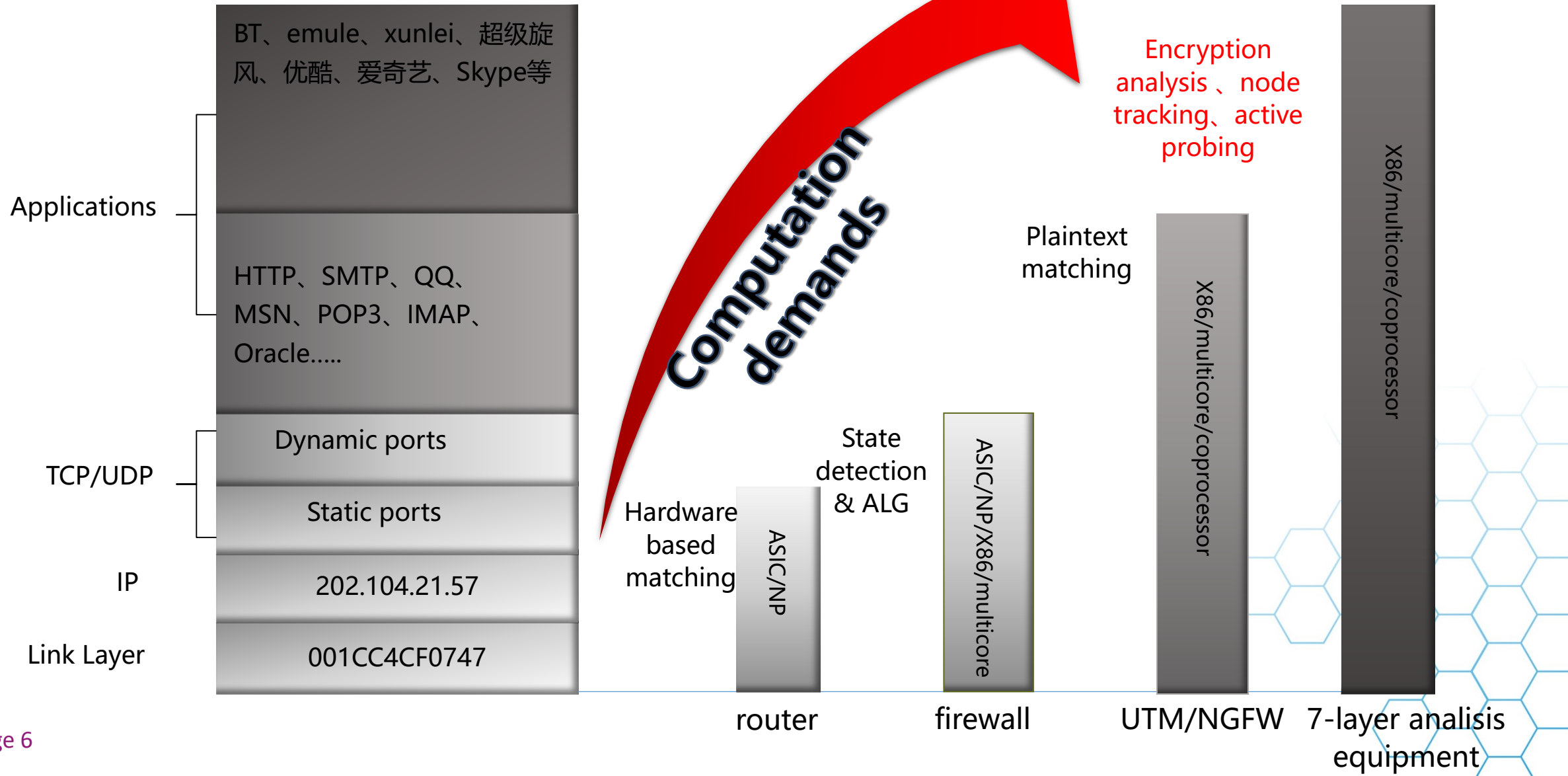

3 pain point




vBRAS structure of large operators

# what's the key to problem

- **Technology view**
  - throughput
  - PPPoE session maintainance ability
  - Account compatibility
  - virtualization
  - Added value and profit
- **Ecnomic view**
  - enormous value of the stock market
  - decoupling of control and forwarding


BT、emule、xunlei、超级旋风、优酷、爱奇艺、Skype等
Applications
HTTP、SMTP、QQ、MSN、POP3、IMAP、Oracle.....
Dynamic ports
TCP/UDP
Static ports
IP
202.104.21.57
Link Layer
001CC4CF0747
Computation demands
Hardware based matching
ASIC/NP
State detection & ALG
ASIC/NP/X86/multicore
Plaintext matching
X86/multicore/coprocessor
Encryption analysis 、node tracking、active probing
X86/multicore/coprocessor
router
firewall
UTM/NGFW
7-layer analisis equipment

## 《给力吧，x86》专题连载九：英特尔Sandy Bridge平台网络通信性能测试分析

作者 老韩 | 2011-12-14 | 类型 研发动态, 科技普及, 网络安全, 行业动感, 通讯产品 | 57条用户评论 »

## 《给力吧，x86》专题连载八：英特尔5520平台网络通信性能测试分析（下）

作者 老韩 | 2011-12-12 | 类型 研发动态, 科技普及, 网络安全, 行业动感, 通讯产品 | 26条用户评论 »

## 《给力吧，x86》专题连载七：英特尔5520平台网络通信性能测试分析（上）

作者 老韩 | 2011-12-12 | 类型 研发动态, 科技普及, 网络安全, 行业动感, 通讯产品 | 1条用户评论 »

## 《给力吧，x86》专题连载六：网络通信硬件平台巡览•D525篇

作者 老韩 | 2011-12-11 | 类型 研发动态, 科技普及, 网络安全, 行业动感, 通讯产品 | 8条用户评论 »

## 《给力吧，x86》专题连载五：网络通信硬件平台巡览•G41篇

作者 老韩 | 2011-12-10 | 类型 研发动态, 科技普及, 网络安全, 行业动感, 通讯产品 | 21条用户评论 »

## 《给力吧，x86》专题连载四：网络通信平台评估软件NCPBench应用分析

作者 老韩 | 2011-12-07 | 类型 研发动态, 科技普及, 网络安全, 行业动感, 通讯产品 | 15条用户评论 »

## 《给力吧，x86》专题连载三：x86平台网络应用效能实测

作者 老韩 | 2011-12-07 | 类型 研发动态, 科技普及, 网络安全, 行业动感, 通讯产品 | 11条用户评论 »

## 《给力吧，x86》专题连载二：x86平台在网络领域的发展应用分析

作者 老韩 | 2011-12-07 | 类型 研发动态, 科技普及, 网络安全, 行业动感, 通讯产品 | 23条用户评论 »

## 《给力吧，x86》专题连载一：x86平台在网络产品中的应用回顾

作者 老韩 | 2011-12-05 | 类型 研发动态, 科技普及, 网络安全, 行业动感, 通讯产品 | 21条用户评论 »

# PANAOS-Standing On The Shoulder Of Intel Giants

Problems of home-made gateway products
Relying on modifying the kernel of general purpose operating system such as Linux/FreeBSD
Poor stability, performance and scalability
Can't satisfy high level demand through low level repeatation

advantages of PANAOS

- Data plane oriented development of OS key elements, such as driver, memory management etc.
- Complete decoupling of data plain and control plain, adopting independent IP protocol stack and driver, guarantee extremely high performance.
- Dual OS backup system, guarantee high stability.
- Built-in functions of routing, NAT, load balance, application recognition and control, providing integrated solution.
- Provide App virtualization engine for third party applications, support built-in third party module.

# How fast could Intel run?

- wrong cognitions of gateway performance
  - Forwarding performance ≠ application performance
  - Use 256 byte packet in benchmark throughput testing
  - Performance bottleneck：CPU、network card、driver、application software

**Atom D525**

Two-way throughput 1G
1,000,000 concurrent connections

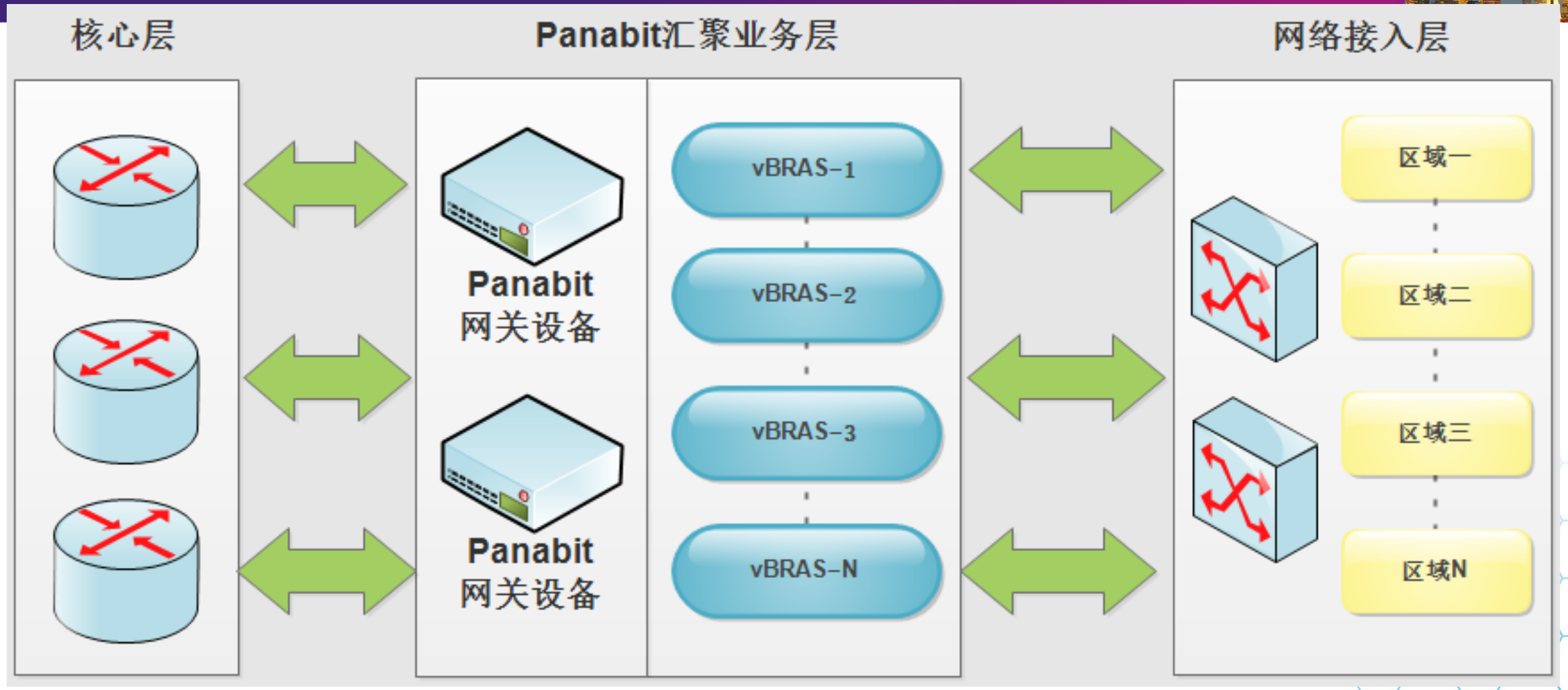
**G41 Q8400**

Two-way throughput 6G
4,000,000 concurrent connections

**C206 I7 2600**

Two-way throughput 40G
10,000,000 concurrent connections

- Performance of panabit
  - Two-way throughput: 80Gbps
  - maximum concurrent connections:15,000,000
  - Maximum concurrent IP address: 600,000
  - Session establishment speed≥650,000/s
  - Application forwarding delay<0.1ms




**Support more than 500 PPPoE virtual server !**

**PPPOE认证->服务列表**

| 名称 | 网卡 | 服务 | 网关地址 | DNS | VLAN | MTU | 认证方式 | 地址池 | RADIUS | 在线用户/最大 | 流入速率 | 流出速率 | 备注 | 添加>> |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PPPOE-SRV01 | ix0 | | 1.1.1.1 | 210.21.196.6,221.5.88.88 | 100 | 1492 | 先本地后RADIUS | pool | 默认服务 | 3362/0 | 87.65M | 1143.84M | | |
| PPPOE-SRV02 | ix0 | | 2.2.2.2 | 210.21.196.6,221.5.88.88 | 200 | 1492 | 先本地后RADIUS | pool | 默认服务 | 3569/0 | 105.54M | 1337.93M | | |
| PPPOE-SRV03 | ix0 | | 3.3.3.3 | 210.21.196.6,221.5.88.88 | 300 | 1492 | 先本地后RADIUS | pool | 默认服务 | 2771/0 | 78.91M | 1020.31M | | |
| PPPOE-SRV04 | ix0 | | 4.4.4.4 | 210.21.196.6,221.5.88.88 | 400 | 1492 | 先本地后RADIUS | pool | 默认服务 | 2813/0 | 73.68M | 922.60M | | |

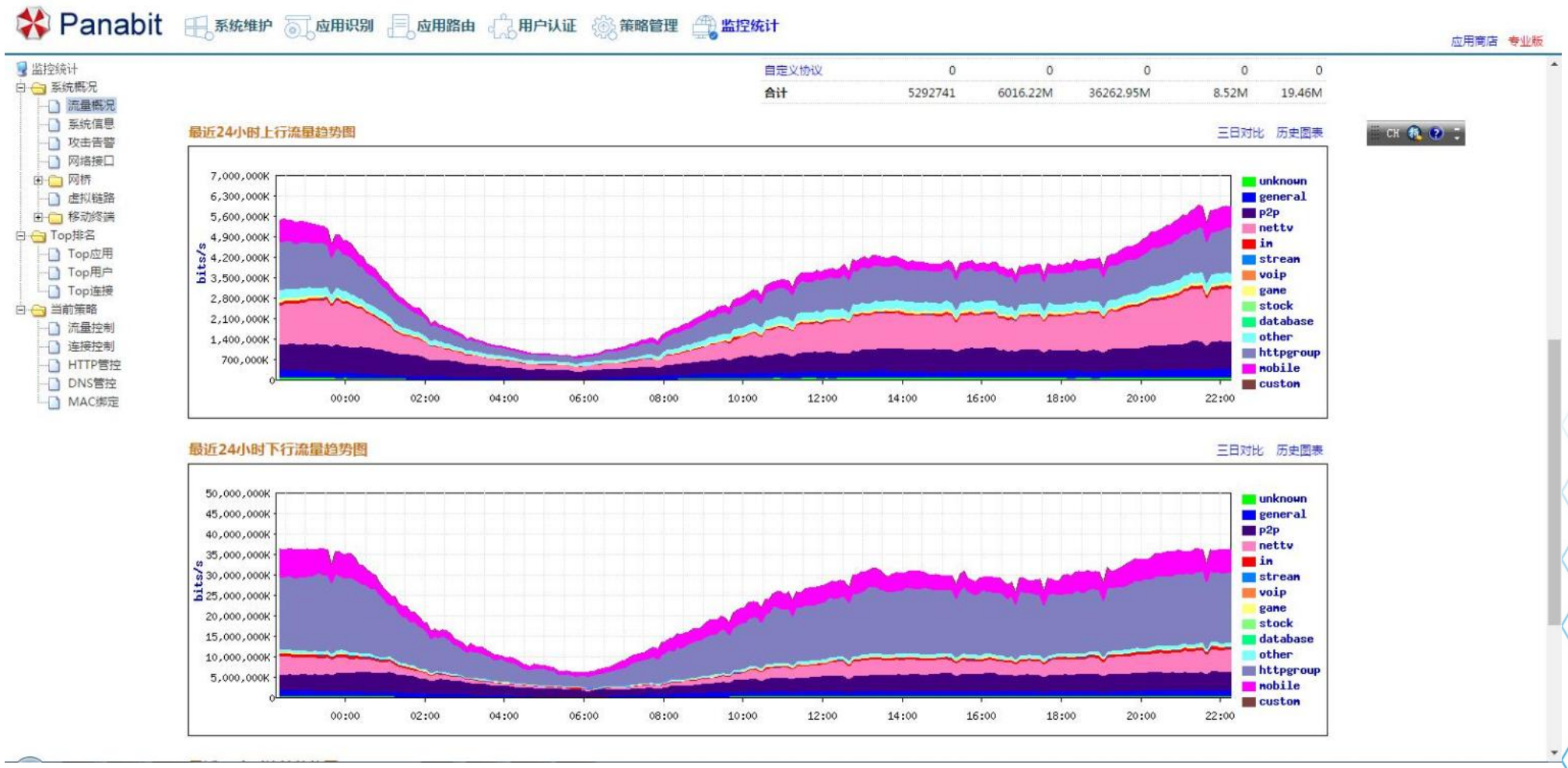
**PPPOE认证->在线用户->本地用户**

地址池 所有账号 服务器 任意服务器 VLAN 0 认证方式 所有方式 其他条件 查询 绑定MAC 强制下线

| 序号 | MAC地址 | IP地址 | 账号 | VLAN | MTU | 认证方式 | 限速(上/下) | 地址池 | 服务器 | 在线时间 | 操作 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 12411 | 28-2c-b2-c3-40-03 | 11.11.248.202 | 15889722799 | 340/0 | 1480 | RADIUS | 800/10100 | pool | PPPOE-SRV01 | 9天6时35分44秒 | 查看 |
| 12412 | 08-10-75-f1-c0-cb | 11.11.252.141 | 13267334563 | 354/0 | 1492 | RADIUS | 350/4099 | pool | PPPOE-SRV03 | 9天6时3分11秒 | 查看 |
| 12413 | 24-69-68-68-7a-49 | 11.11.252.92 | 15014005765 | 53/0 | 1480 | RADIUS | 420/10000 | pool | PPPOE-SRV04 | 9天6时4分57秒 | 查看 |
| 12414 | 14-75-90-c9-5b-11 | 11.11.244.122 | 13713210307 | 354/0 | 1480 | RADIUS | 400/6500 | pool | PPPOE-SRV01 | 9天7时19分34秒 | 查看 |
| 12415 | f4-6a-92-10-5a-51 | 11.11.244.82 | 13691797855 | 358/0 | 1480 | RADIUS | 480/6300 | pool | PPPOE-SRV02 | 9天7时21分13秒 | 查看 |
| 12416 | 48-5b-39-94-53-33 | 11.11.241.247 | 13530561361 | 53/0 | 1492 | RADIUS | 450/6100 | pool | PPPOE-SRV03 | 9天7时49分28秒 | 查看 |
| 12417 | 08-10-79-8d-a1-9d | 11.11.245.115 | 18274078321 | 358/0 | 1492 | RADIUS | 350/4099 | pool | PPPOE-SRV04 | 9天7时8分0秒 | 查看 |
| 12418 | f0-b4-29-20-48-74 | 11.11.241.42 | 18750457533 | 354/0 | 1480 | RADIUS | 350/4099 | pool | PPPOE-SRV01 | 9天8时1分6秒 | 查看 |
| 12419 | a4-56-02-d6-ea-f8 | 11.11.238.117 | 13714717206 | 354/0 | 1480 | RADIUS | 350/4099 | pool | PPPOE-SRV04 | 9天8时55分21秒 | 查看 |
| 12420 | 08-10-79-83-44-06 | 11.11.240.199 | 13713925734 | 358/0 | 1492 | RADIUS | 350/4099 | pool | PPPOE-SRV04 | 9天8时7分53秒 | 查看 |
| 12421 | 08-10-78-aa-96-e9 | 11.11.236.151 | 18944708139 | 354/0 | 1492 | RADIUS | 350/4099 | pool | PPPOE-SRV02 | 9天9时36分17秒 | 查看 |
| 12422 | 08-10-78-64-1a-ba | 11.11.236.54 | 13360505919 | 30/0 | 1492 | RADIUS | 350/2055 | pool | PPPOE-SRV03 | 9天9时44分45秒 | 查看 |

- virtualize 500 vBRAS services in one interface
- every vBRAS service support independent account and IP address configuration
- support more than 32k/U concurrent PPPoE users
- 40G/U throughput
- support QinQ and PPPoE agent
- Active content push and DPI big data analysis capability
- PPPoE service load balance
- complete QoS/NAT/routing protocol stack

# support 40G traffic in one CPU



# vBRAS vs BRAS: software agility

应用协议->迅雷档案　　已运行206天7小时1分

| 连接数 | 节点数 | 流量(上行/下行) | 速率(上行/下行) | 代理速率(上行/下行) |
|---|---|---|---|---|
| 62329 | 34524 | 171597.74G/602672.18G | 73.42M/249.63M | 0/0 |

协议配置 | 趋势图表 | 在线用户 | 连接排名 | 有关策略

按照 总速率 排序 显示前 20 项 每隔 不刷新 秒刷新 提交

| IP地址 | 账号 | 迅雷流入速率 | 迅雷流出速率 | 迅雷总速率 |
|---|---|---|---|---|
| 10.156.101.244 | | 28.49M | 311.30K | 28.80M |
| 10.15.217.38 | | 15.61M | 814.54K | 16.42M |
| 10.156.98.150 | | 9.39M | 531.48K | 9.92M |
| 10.15.23.86 | | 6.99M | 419.97K | 7.41M |
| 10.172.0.16 | | 6.48M | 878.14K | 7.36M |
| 10.173.0.18 | | 6.54M | 466.05K | 7.01M |
| 10.155.174.206 | | | 344.18K | 6.49M |
| 10.173.0.14 | | | 2.46M | M |
| 10.150.146.159 | | | | |
| 10.152.2.90 | | | | |
| 10.152.48.60 | | | | |
| 10.0.42.125 | | 0 | | |
| 10.0.42.126 | | 0 | | |
| 10.154.211.251 | | 4.70M | 336.33K | .04M |
| 10.155.198.84 | | 4.47M | 251.90K | 4.72M |
| 10.154.31.132 | | 4.43M | 240.18K | 4.67M |
| 10.173.0.16 | | 4.52M | 127.96K | 4.65M |
| 10.156.252.152 | | 4.31M | 253.34K | 4.57M |
| 10.157.253.220 | | 1.60M | 2.70M | 4.30M |
| 10.149.48.67 | | 3.87M | 200.97K | 4.07M |

IP

application

10.157.163.13档案　　已运行206天6小时58分26秒 [

| 账号信息 | MAC地址 | 累计流出 | 累计流入 | 流出bps | 流入bps |
|---|---|---|---|---|---|
| | 00-14-1b-37-c8-00 | 346.79M | 599.20M | 503.18K | 447.27K |
| 速率限制(入/出,kbps) | 在线时长(秒) | 连接数 | 虚拟身份 | 共享用户 | 移动终端 |
| 0/0 | 0/14:36:00 | 158 | 1/1 | 0/0 | 0 |

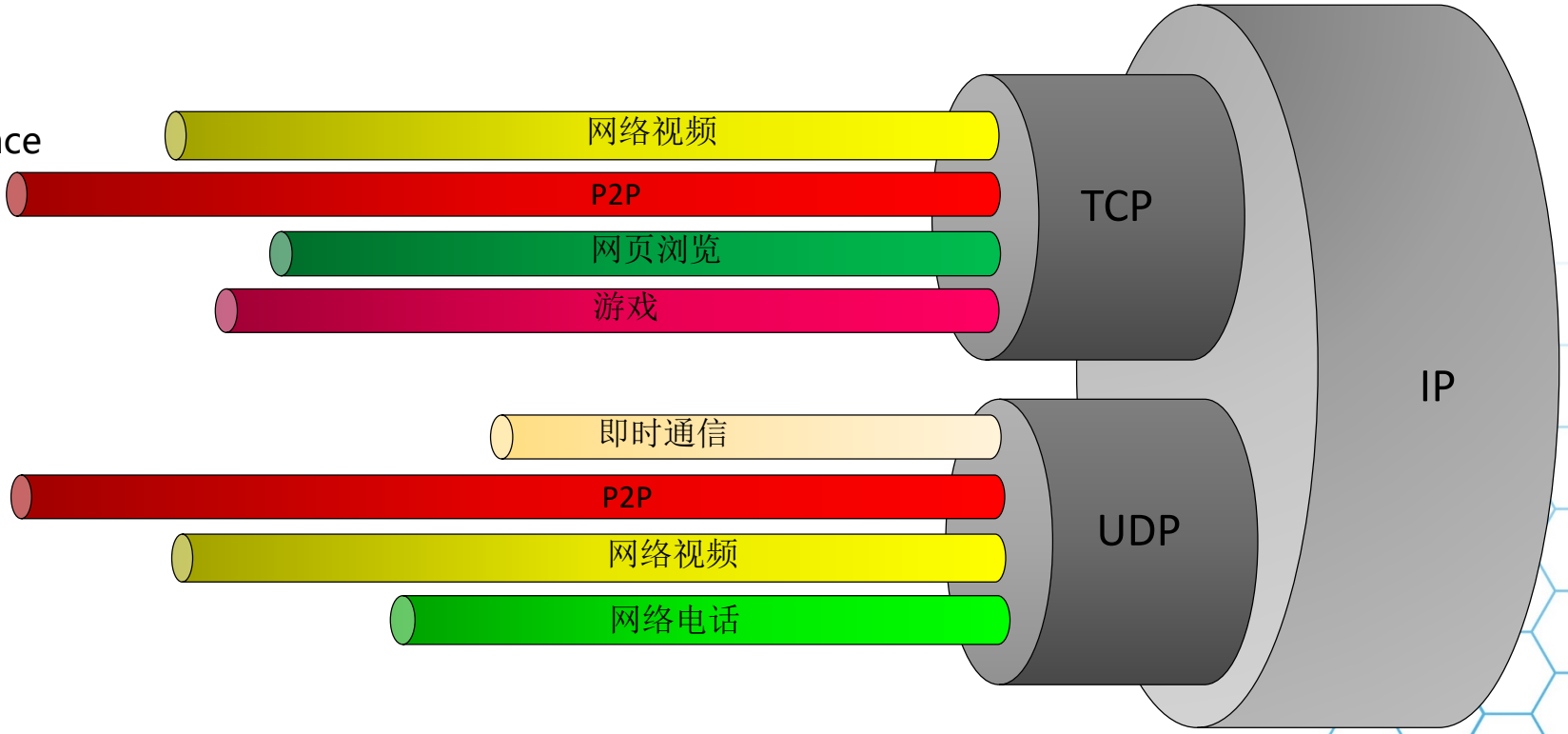
流量概况 | 连接信息 | 虚拟身份 | 共享用户 | 移动终端

红色的是经过连接数控制模块检测而被拒绝的连接

| 应用名称 | 协议 | 连接 | WAN线路 | 时长 | DSCP | 流量(up/down) |
|---|---|---|---|---|---|---|
| 地下城与勇士 | tcp | 10.157.163.13:53819--183.61.235.60:7106 | | 603 | 0/0 | 32122683/26235907 |
| 未知80端口 | tcp | 10.157.163.13:55900--14.17.57.249:80 | | 717 | 0/0 | 953158/173983 |
| 地下城与勇士 | tcp | 10.157.163.13:53802--183.61.160.179:10024 | | 637 | 0/0 | 139231/558938 |
| QQ聊天 | udp | 10.157.163.13:4015--183.60.56.33:8000 | | 652 | 0/0 | 42448/140449 |
| QQ幻想世界 | tcp | 10.157.163.13:50033--101.226.124.52:8108 | | 7439 | 0/0 | 29994/85317 |
| 歪歪语音 | tcp | 10.157.163.13:53114--61.130.29.136:8080 | wan3144 | 756 | 0/0 | 48707/46649 |
| 英雄联盟游戏 | udp | 10.157.163.13:60993--183.61.228.99:8081 | | 734 | 0/0 | 91912/0 |
| Teredo | udp | 10.157.163.13:52846--94.245.121.253:3544 | | 7683 | 0/0 | 15372/18966 |
| 51炫舞 | tcp | 10.157.163.13:53953--111.221.29.254:443 | | 54 | 0/0 | 28375/4556 |
| 穿越火线 | udp | 10.157.163.13:56029--183.61.228.99:8080 | | 716 | 0/0 | 24300/0 |
| 其它HTTPS | tcp | 10.157.163.13:49779--111.221.29.96:443 | | 7608 | 0/0 | 3751/6375 |
| 歪歪语音 | tcp | 10.157.163.13:45013--222.73.62.124:78 | wan3142 | 754 | 0/0 | 4546/3121 |
| PPWeb | tcp | 10.157.163.13:38383--116.31.72.59:443 | | 50101 | 0/0 | 1221/1220 |
| 优酷 | tcp | 10.157.163.13:36298--14.17.112.234:80 | | 20 | 0/0 | 1987/20 |

# vBRAS infrastructure -DPI

- Signature based DPI
  - Most common & effective
  - Evolve to DFI
- Node tracking
  - improving the accuracy and performance of inspection
  - Active probing
  - Probing the far end status of
  - Encrypted protocol
- Protocol multi-state machine
  - From plaintext dual-state machine to multi-state machine
  - Packet length and frequncy
  - Leaking message

# Q&A