Panabit

无中生有

基于骨干网全量应用识别的威胁情报基础数据采集

北京派网软件有限公司

RSA总裁Amit Yoran

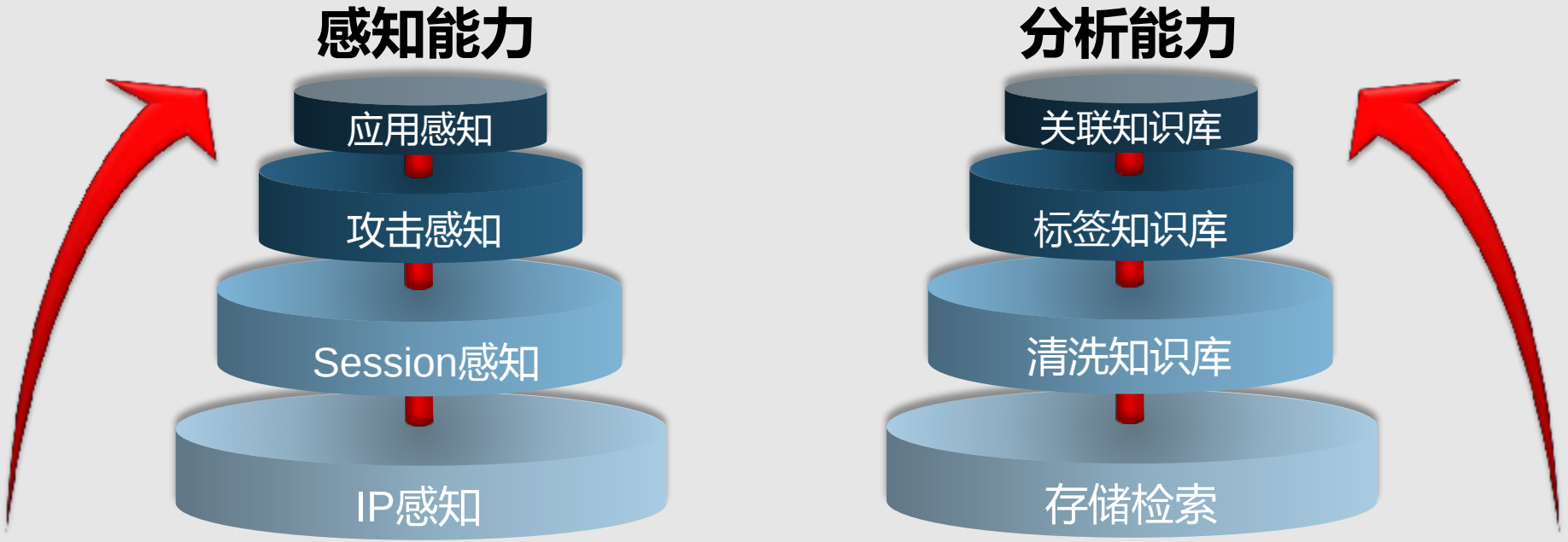
人们身处黑暗之中都会产生恐惧，因为当看不清周围的环境，却能听到声响或人影晃动时，人们将无法判断其中是否潜在着一些危险，这就像今天信息安全产业所面临的一个状态。

我们需要一张新的“地图”。这张“地图”一方面不依赖于预先保护机制；一方面强调普遍的可视性；一方面可以很好地进行身份认证和识别，掌握来自外部的威胁情报；一方面又能基于业务的重要级别，进行安全资源的优化部署。

威胁情报获取的难题？

威胁情报关键能力建设

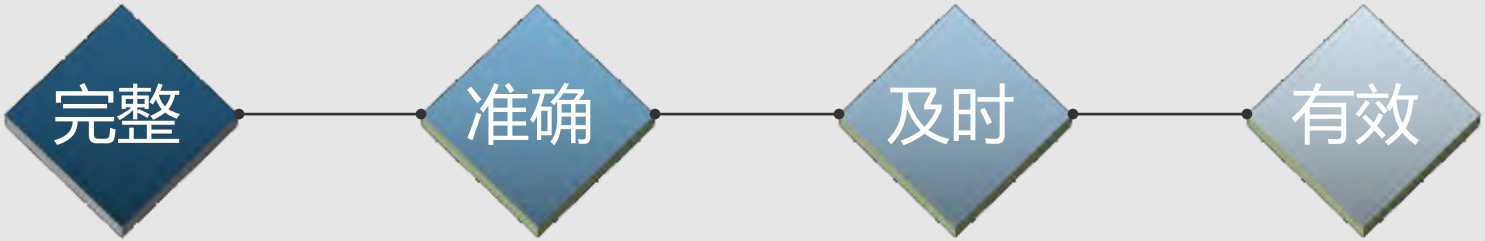
- 在互联网应用的推动下，从IP为实体的三层感知进化为应用为实体的七层感知，威胁情报理念发展的必然趋势。稳定的存储检索、准确的清洗、标签和关联知识库则是分析能力的基础，它们的表现直接决定了结果的有效性。



感知有多难

技术链条

没有互联网的飞速发展，就没有威胁情报，数据量猛增是互联网发展所带来的必然结果，准确及时的大数据分析是掌握网络运行状况和控制网络用户行为的有效手段。



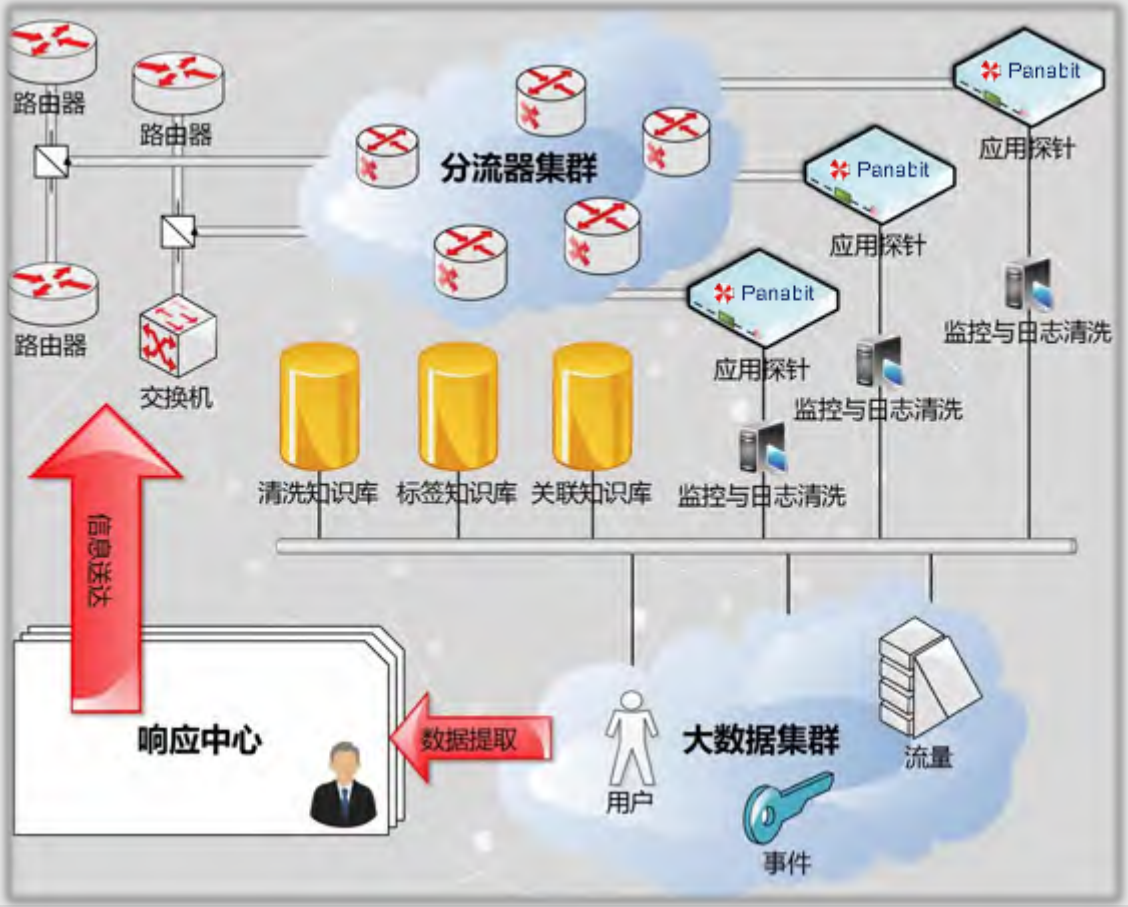
- ✓ 速度：超高处理性能适应互联网带宽发展趋势
- ✓ 广度：通达第七层，关键信息一个都不能少

- ✓ 内容无关：规整合并、垃圾过滤
- ✓ 内容相关：数据标签化、摘要常态化
- ✓ 存储优化

- ✓态势分析 ：运行状况、应用预警
- ✓经营分析：成本利润、资源引入、聚类客户评价、 潜在客户挖掘

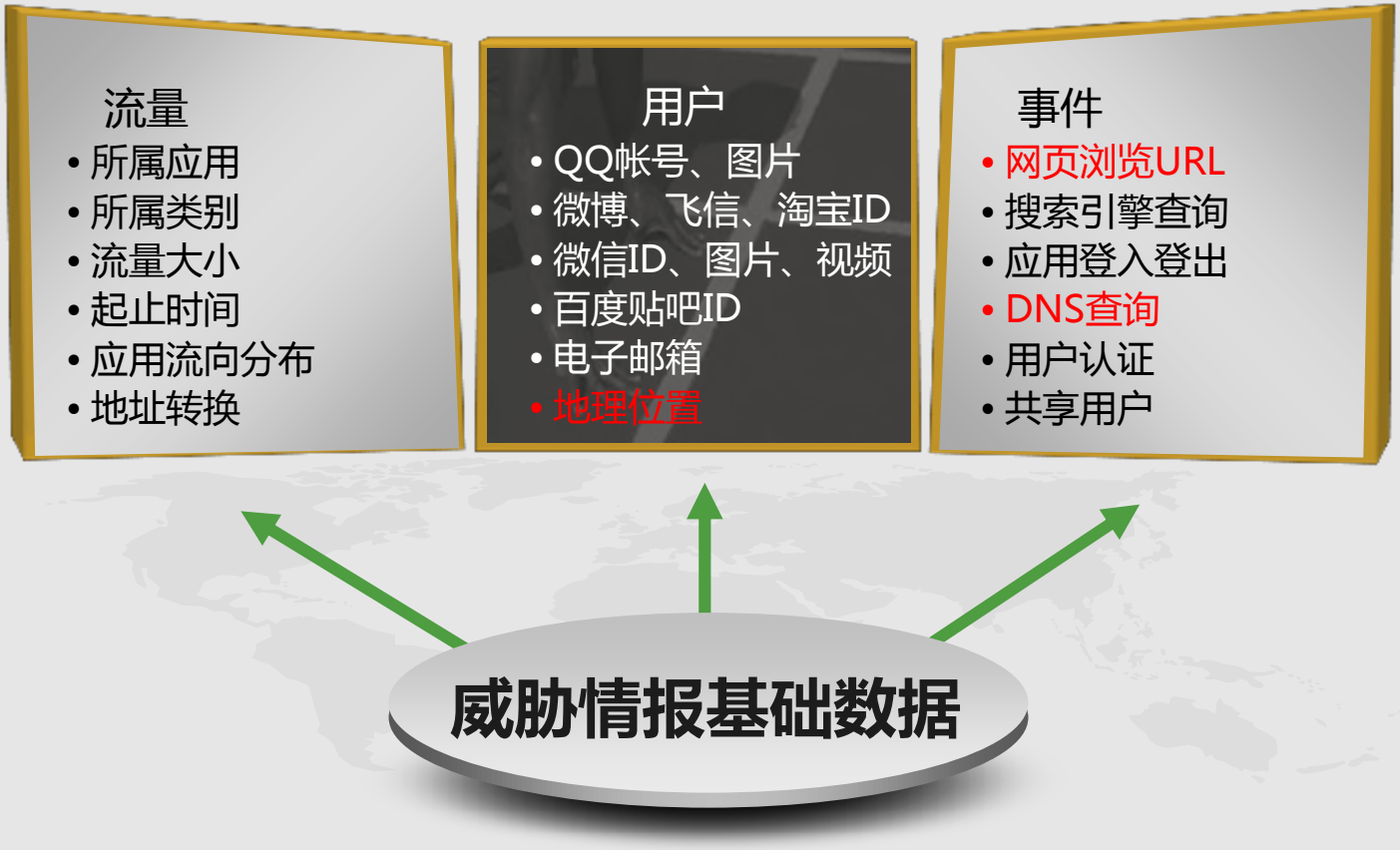
- ✓ 身份定位：实体客户与网络身份映射
- ✓ 到达手段：上门服务、电话、短信、Web信息、微信、QQ等

鸟揽威胁情报系统



- 分流子系统
  - 协议转换，同源同宿
- 探针子系统
  - 用户、事件和流量日志生成
  - 内容无关清洗
  - 数据初加工
- 知识库子系统
  - 内容相关清洗知识库
  - 标签知识库
  - 关联知识库
- 数据存储分析子系统
  - 海量存储、标签化、关联分析
- 响应子系统
  - 响应信息生成、过滤与送达

威胁情报源仅仅是URL和DNS？ No！

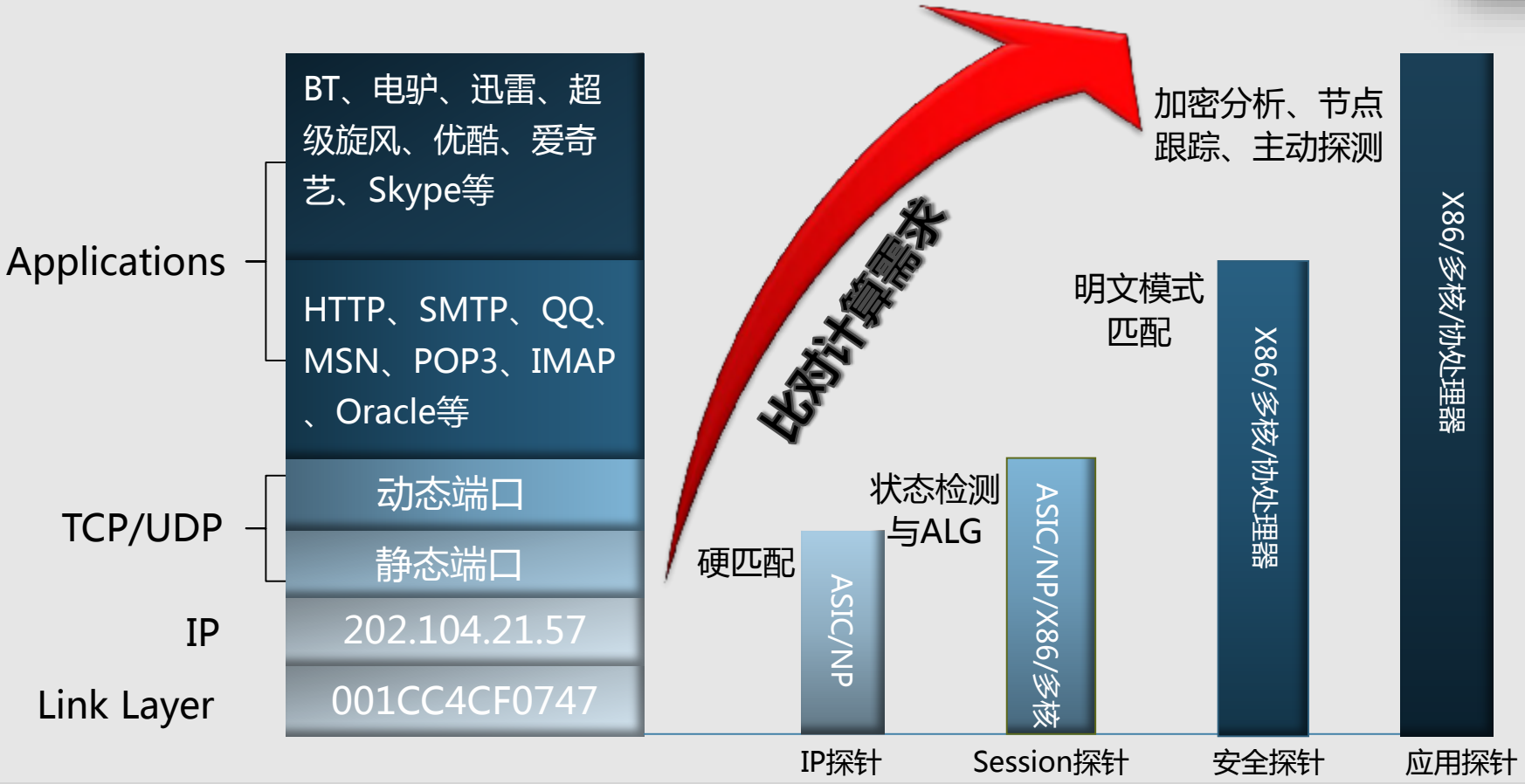


理想照进现实

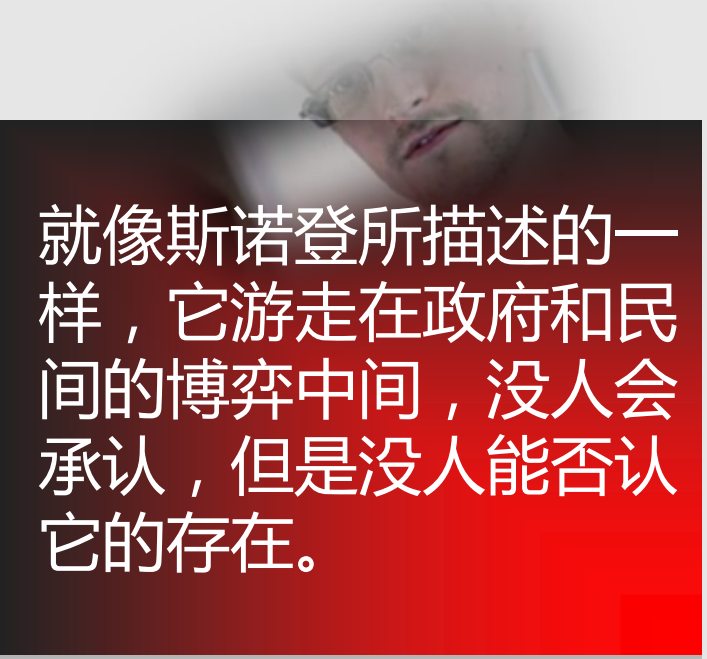
- 记录全世界？
  - 能力对等
- 数据究竟有多大？
  - 用户、流量、事件
  - 10GE到300G字节
- 数据收敛从哪里开始？
  - 分流子系统的局限性
  - 没有数据是多余的
  - 应用探针的全流量需求
- 需要多少专家才能走完整个链条？
  - 让专业的人做专业的事
  - 通信专家、应用识别专家、大数据专家、行业知识专家

探针技术演进

探针设备数据面的变化



你是谁？



就像斯诺登所描述的一样，它游走在政府和民间的博弈中间，没人会承认，但是没人能否认它的存在。

档案->219.216.110.96

TTL秒	在线时间(秒)	流出流量	流入流量	流出bps	流入bps
597	1229292	575.26G	847.27G	789.98K	25.28M
MAC地址		连接数		被拒绝的连接数	
00.e0.fc.46.2f.bc		2814		0/0	

序号	身份类型	身份信息	最近使用时间(s)
1	QQ号码	1355427908	2012-09-24
2	内部私有IP	192.168.0.91	2012-09-24
3	QQ号码	515348695	2012-09-24
4	POP3账号	kedadiqipi@163.com	2012-09-24
5	QQ号码	775315895	2012-09-24
6	QQ号码	411825361	2012-09-24

+ 防火墙日志 + 虚拟身份信息库 + IP位置信息库

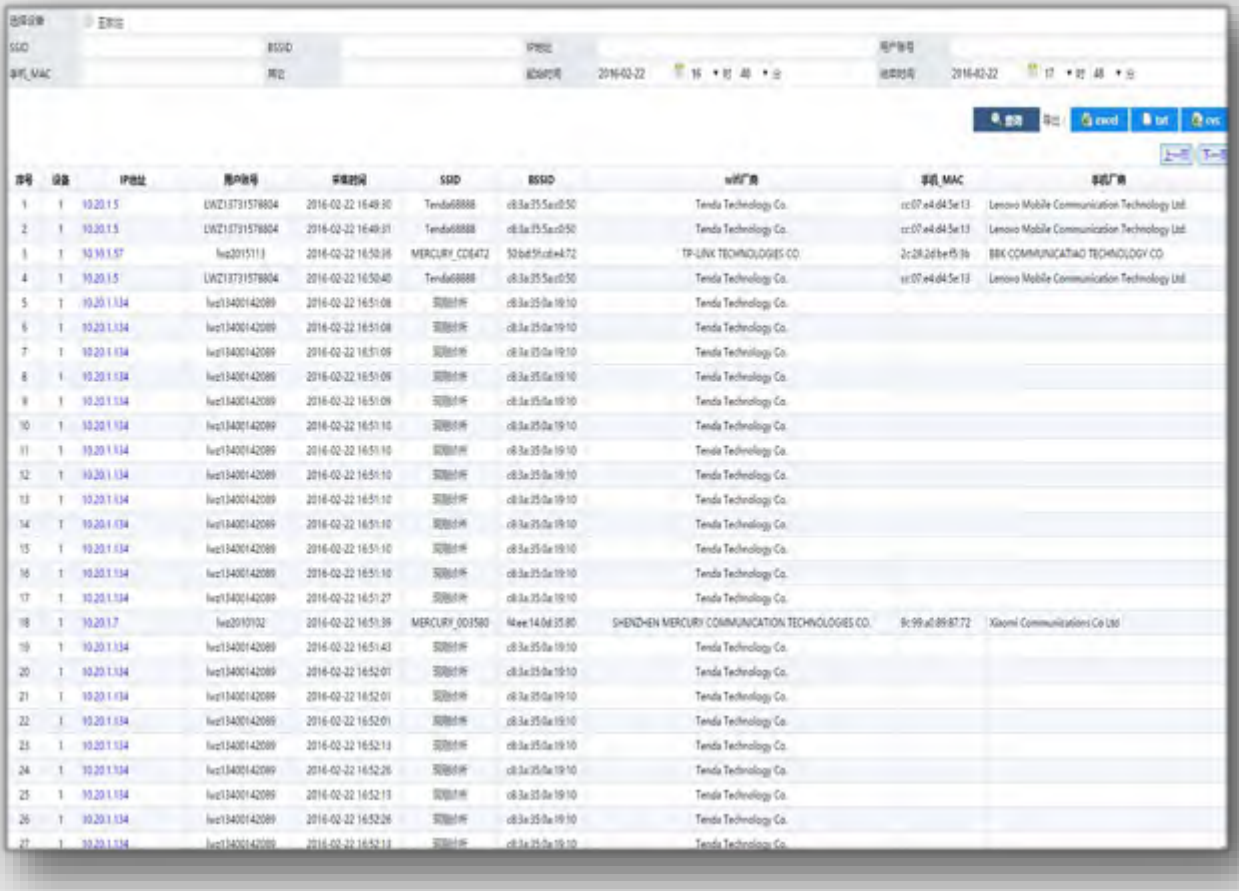
= （张三在某酒店某房间登录了某种应用）

- 为什么是Panabit？
  - 部署位置位于通信主干，无需改变连接拓扑
  - 有足够性能处理关键要素审计，不需要增加设备
  - 对应用协议分析透彻，分析协议可以涵盖邮件、即时通信、社交应用和游戏等
  - 附加移动属性：微博、微信、贴吧等

从哪里来？

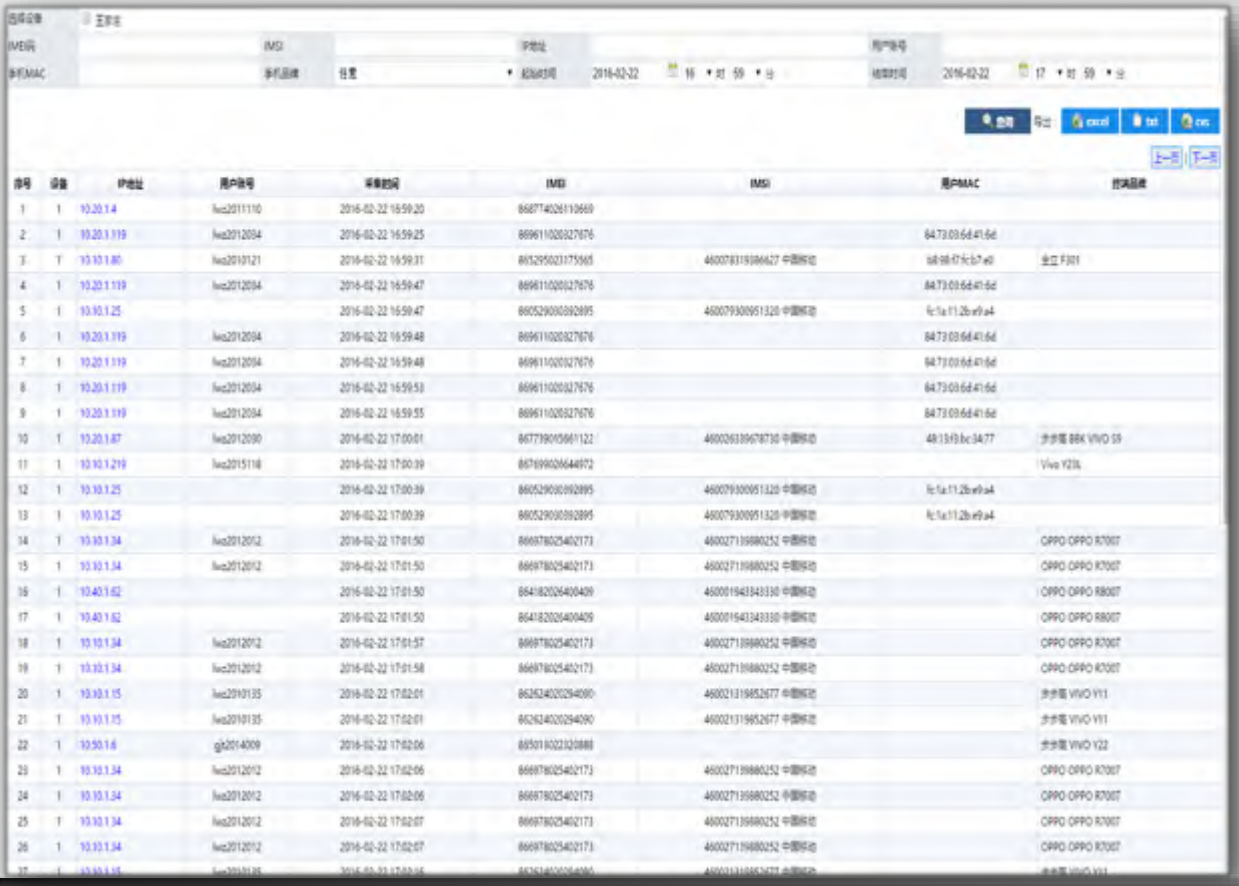
- 实时精确锁定固网IP地址地理位置
  - 精确度达到街道楼宇级别
  - 不依赖手机移动网提供数据
  - 可定位内网IP地理位置
  - 噪声来源：异地访问，设备切换
  - 依赖大数据统计规律
- 核心用途
  - 用户地理热力分析
  - 应用地理热力分析
  - 虚拟身份定位
  - 无线环境定位
  - 带宽资产核查

无线环境信息



- 无线环境
  - IP和时间为索引
  - 无线AP的SSID
  - 无线AP的MAC地址
  - 客户机MAC地址
  - 无线环境探查

移动设备信息



- 移动设备信息
  - IP和时间为索引
  - IMEI：唯一标识UE硬件
  - IMSI：唯一标识用户ID
  - 手机MAC
  - 大数据分析主要索引，例如他网手机定位

到哪里去？

- 应用去往何处？
  - 所属应用：根据应用协议判定流量所属于的具体应用种类，比如：迅雷、微信、爱奇艺、魔兽世界等。
  - 所属类别：归类具体应用所属的应用类别，比如：视频、即时通信、P2P下载等，目的是为了批量处理。
  - 流量大小：流量所属session总数据量
  - 起止时间：流量所属session开始和结束时间
  - 应用流向分布：根据源地址和目的地址判别session在运营商之间或者地理版图上分布。
- 与其他业务系统的配合
  - 应用镜像：根据应用复制到探针其他物理接口
  - 应用路由：根据应用选择下一条路由

域名分布异常

域名 www.sina.com.cn 起始时间 2016-05-08 结束时间 2016-05-09 查询

编号	区域	URL访问量	操作
1	北京	14413	域名详细
2	重庆	8317	域名详细
3	广东	2408	域名详细

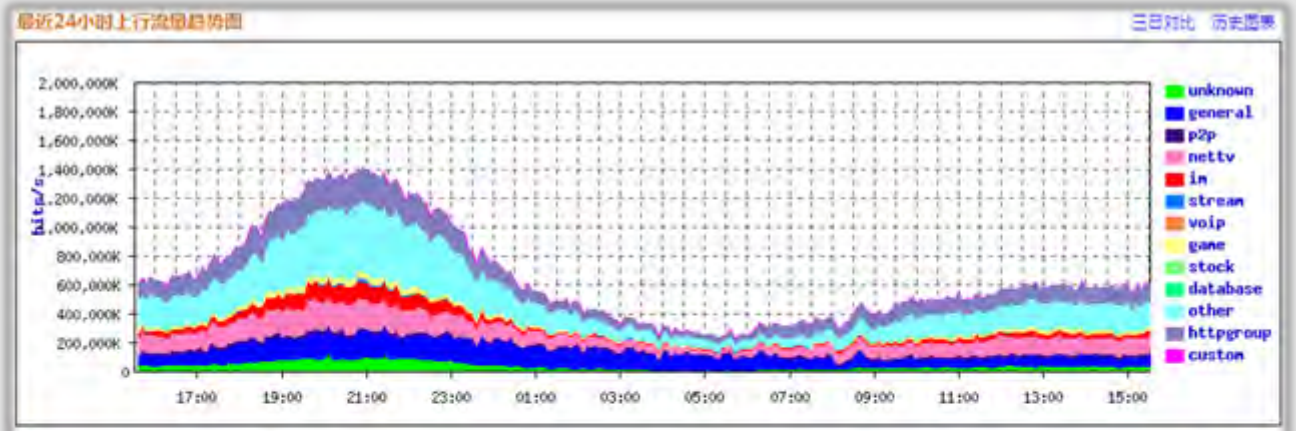
微软威胁情报中心总经理 John Lambert

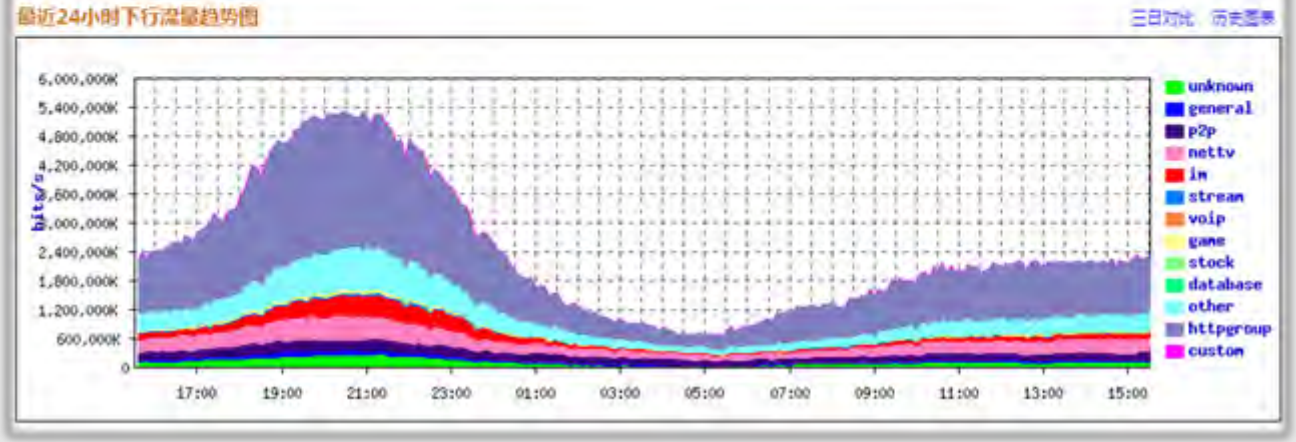
What is the most important network security spend: Sensor appliances? SIEM? Threat intelligence feeds? It's your analyst team.

最重要的网络安全开支是什么？传感器设备？安全信息和事件管理？威胁情报来源？

都不是，最重要的是你的分析师团队。

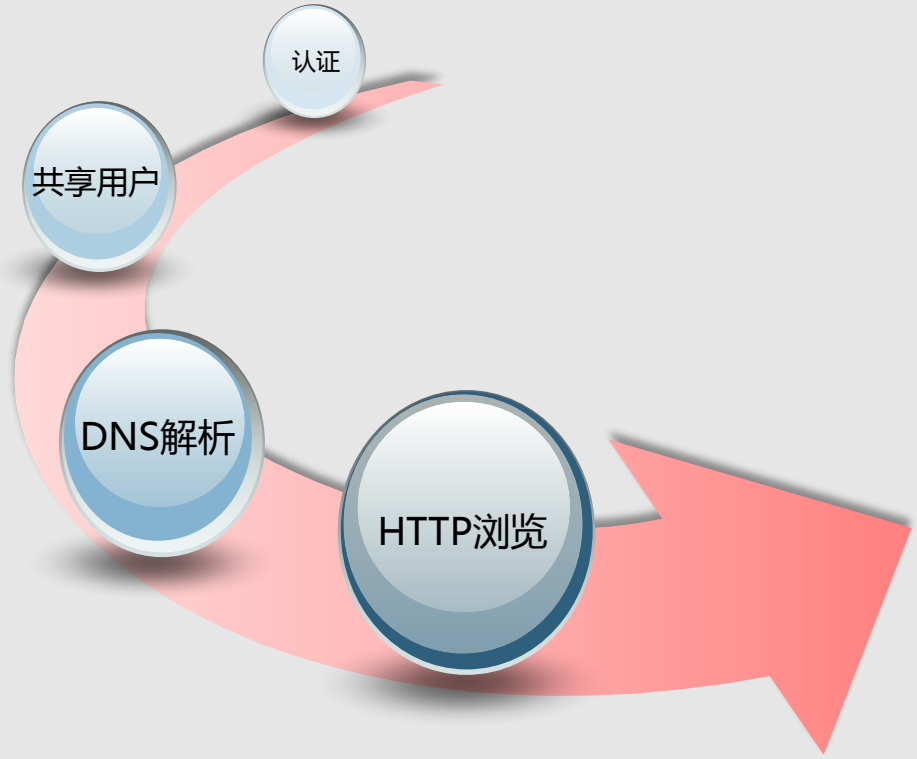
全业务分析





- 全业务流量
  - 14大类
  - 近1000种应用
  - 如实反映协议完整性，从三层覆盖到七层
- 可视化分析
  - 分布
  - 极值
  - 趋势
  - 相似性
- 流量与事件的事件序列分析是主要未知威胁情报来源

时间序列分析



- HTTP浏览
  - HOST排名：发现异常，例如：病毒
  - URL分类库：判定行为，分布
  - 按IP段的URL排名：区域互联网活跃度，数据中心发展情况
- DNS解析
  - DNS解析排名：发现异常，例如：DNS Flooding，C&C
  - 时间段解析情况：应用热度
- 认证
  - 针对AAA系统的攻击与滥用
- 共享用户事件
  - 实际用户数监测：NAT后面有乾坤

应用画像

用户画像

30.20.169.109档案

已运行1天22小时32分23秒 (admin)

TTL(秒)	在线时间(秒)	流出流量	流入流量	流出bps	流入bps
591	4464	3.20M	9.63M	10.95K	54.74K

MAC地址	连接数	虚拟身份	共享用户	移动终端
6c:9c:ed:33:00:28	21	0	0/0	无

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)
DNS	udp	30.20.169.109:1059--180.153.162.150:53		1476	0/0	0/2071479
DNS	udp	30.20.169.109:1065--180.153.162.150:53		1476	0/0	0/2065804
DNS	udp	30.20.169.109:4278--119.167.195.3:53		1809	0/0	927582/0
DNS	udp	30.20.169.109:3087--59.46.2.109:53		940	0/0	0/172056
DNS	udp	30.20.169.109:3086--59.46.2.109:53		940	0/0	0/147288
WWW	tcp	30.20.169.109:1066--115.230.124.15:80		1233	0/0	0/94953
传奇系列	tcp	30.20.169.109:1781--117.27.240.155:8001		1601	0/0	0/5502
未知应用	tcp	30.20.169.109:4071--106.120.160.86:80		1816	0/0	235/0
未知应用	udp	30.20.169.109:4137--98.126.77.138:53		156	0/0	143/0
未知应用	tcp	30.20.169.109:1623--111.206.79.230:80		1755	0/0	134/0
SYN_ACK	tcp	30.20.169.109:1968--101.71.24.29:46539		5	0/0	0/0
SYN_ACK	tcp	30.20.169.109:2189--219.129.237.189:8885		9	0/0	0/0
传奇系列	tcp	30.20.169.109:1516--110.80.140.140:895		8	0/0	0/0
传奇系列	tcp	30.20.169.109:1503--120.41.32.238:7301		8	0/0	0/0
传奇系列	tcp	30.20.[illegible]:3308--120.41.32.238:7301		8	0/0	0/0
SYN_ACK	tcp	30.2[illegible]		1	0/0	0/0
SYN_ACK	tcp	30.[illegible]		1	0/0	0/0
传奇系列	tcp	30.20.[illegible]		3	0/0	0/0
SYN_ACK	tcp	30.20.169.109:1348--101.71.24.29:46539		5	0/0	0/0
传奇系列	tcp	30.20.169.109:1564--113.107.181.67:7107		5	0/0	0/0
传奇系列	tcp	30.20.169.109:1535--110.80.140.140:47410		6	0/0	0/0

重度游戏客户

111.136.57.205档案

已运行162天17小时6分0秒 [JLCTT2]

TTL(秒)	在线时间(秒)	流出流量	流入流量	流出bps	流入bps
599	55481	1.66G	56.69G	306.50K	13.00M

MAC地址	连接数	虚拟身份	共享用户	移动终端
00:18:82:35:3e:3c	1575	1	0/0	无

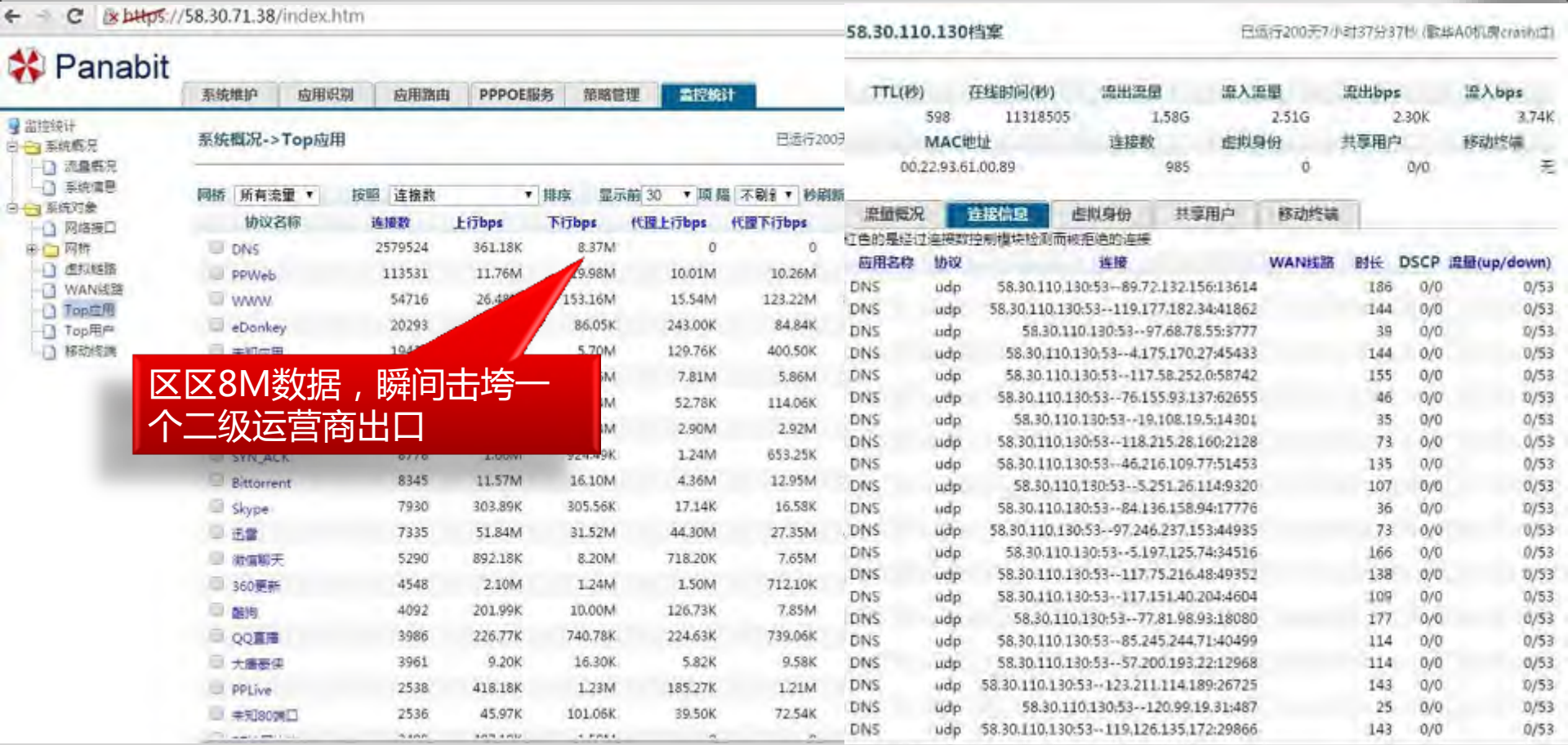
流量概况 | 连接信息 | 虚拟身份 | 共享用户 | 移动终端

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)
仿IE下载	tcp	111.136.57.205:60445--113.142.21.76:443		1918	0/0	535/666549331
仿IE下载	tcp	111.136.57.205:63394--113.142.13.200:443		997	0/0	535/177235011
仿IE下载	tcp	111.136.57.205:63492--113.142.13.180:443		963	0/0	1072/174987935
仿IE下载	tcp	111.136.57.205:65398--123.138.162.97:443		138	0/0	487/68025556
QQ聊天	udp	111.136.57.205:52946--111.161.88.90:8000		39775	0/0	157092/903396
QQ聊天	udp	111.136.57.205:29929--113.142.10.71:8000		39664	0/0	125952/34506
超级旋风	udp	111.136.57.205:29909--58.39.230.117:29909		4782	0/0	127452/0
迅雷	udp	111.136.57.205:14759--112.65.1.104:15473		2677	0/0	28528/0
迅雷	udp	111.136.57.205:14759-[illegible]655		2476	0/0	26387/0
迅雷	udp	[illegible]59-[illegible]306		2476	0/0	25279/0
迅雷	udp	111.136.57.205:14759--123.139.176.174:28418		2350	0/0	25255/0
迅雷	udp	111.136.57.205:14759-		2470	0/0	23499/0

重度下载客户

应用画像与用户画像交叉引用



极值实例-网站排名分析

域名 显示 50 项

按 不排序 排序

起始时间 2014 年 12 月 1 日 查询

结束时间 2014 年 12 月 2 日

序号	域名	访问数	上行流量	下行流量
1	mmsns.qpic.cn	71946	6.29G	88.04G
2	market.huobi.com			
3	wa.gtimg.com			
4	www.22mt.la			
5	116.211.111.250			
6	www.sqmai.cn	15069	2.46M	1.36M
7	mobads-logs.baidu.com	43508	496.7M	42.68M
8	short.weixin.qq.com	31830	12.64G	6.09G
9	sango.qzone.qzoneapp.com	48676	67.57M	270.16M
10	bbs.eduwest.com	48154	2.19M	2.89M
11	tieba.baidu.com	105111	1.01G	2.18G
12	appstore.duokanbox.com	56520	13.3M	5.44M
13	monitor.uu.qq.com	79537	4.23G	783.22M
14	3g.music.qq.com	90704	33.6M	6.17M
15	news.sina.com.cn	116733	70.57M	109.62M
16	mvads.kugou.com	60686	82.89M	1.13M
17	api.chat.xiaomi.net	56079	537.13M	140.06M
18	msg.api.xiaoenai.com	21114	1.05M	158.65K
19	qzonestyle.gtimg.cn	107903	1.35G	5.07G
20	dl28.yunpan.360.cn	41210	48.08M	10.26G
21	182.118.24.33	51465	2.8M	11M
22	caoimg.com	66842	8.55M	335.6M

批量推动微软升级，
产生的直接效果

区域流量 -> 宁夏

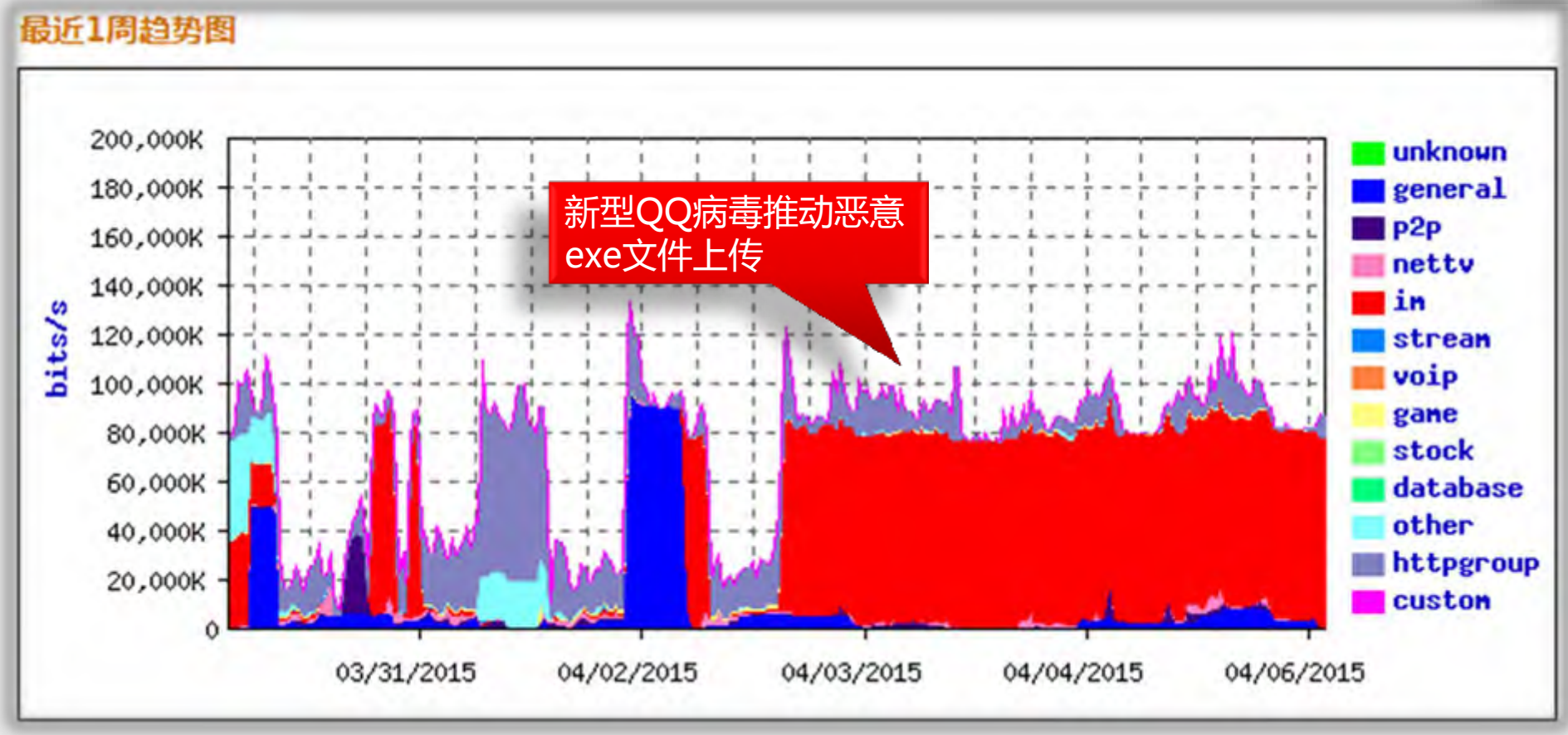
序号	域名	访问量	上行流量(Byte)	下行流量(Byte)	总流量(Byte)
1	vi6.6.cn	65524	1.06M	0	1.06M
2	www.download.windowsupdate.com	65015	46.44M	683.7M	730.14M
3	www.hx2car.com	64584	5.48M	586.94K	6.05M
4	portal.maxthonimg.com	64503	331.74K	2.69M	3.02M
5	vr5.6rooms.com	63842	1.09M	0	1.09M
6	assets.dwstatic.com	62475	28.37M	83.52M	111.89M
7	jebe.ximg.cn	61106	2.57M	21.3M	23.87M
8	www.doniv.net	60133	604.57K	644.68K	1.22M
9	sp.vip.com	59956	2.39M	38.39M	40.78M
10	cc.acgvideo.com	58486	5.22M	15.52G	15.52G
11	vi0.6.cn	55572	1.04M	0	1.04M
12	cu005.www.duba.net	55323	53.9M	2.77G	2.82G
13	vi0.6rooms.com	54561	1.24M	0	1.24M
14	vj6.6rooms.com	54492	901.38K	0	901.38K
15	snapshot.file.yy.com	54310	519.85K	652.42K	1.14M
16	screenshot.dwstatic.com	53736	9.34M	67.15M	76.49M
17	share.vip.com	53029	10.96M	3.91M	14.87M
18	image.yy.com	52973	143.42M	1.43G	1.57G
19	p0.pstatp.com	52609	1.49M	10.46M	11.95M
20	static.m.yy.com	52484	101.57M	475.66M	577.22M
21	i0.6.cn	52081	1.05M	0	1.05M
22	vj0.6rooms.com	51763	5.26M	29.5M	34.76M
23	ax.init.itunes.apple.com	51005	6.79M	29.22M	36.01M
24	s1.doyo.cn	50874	8.87M	31.66M	40.54M
25	img2.a0bi.com	48414	1.71M	7.98M	9.7M
26	uploadmp3.6rooms.com	48375	973.59K	90.87M	91.82M
27	s.cimg.163.com	47824	77.76M	490.86M	568.62M
28	123.94.26.134	47568	5.61M	1.94G	1.95G

极值实例- App安全的那些梗

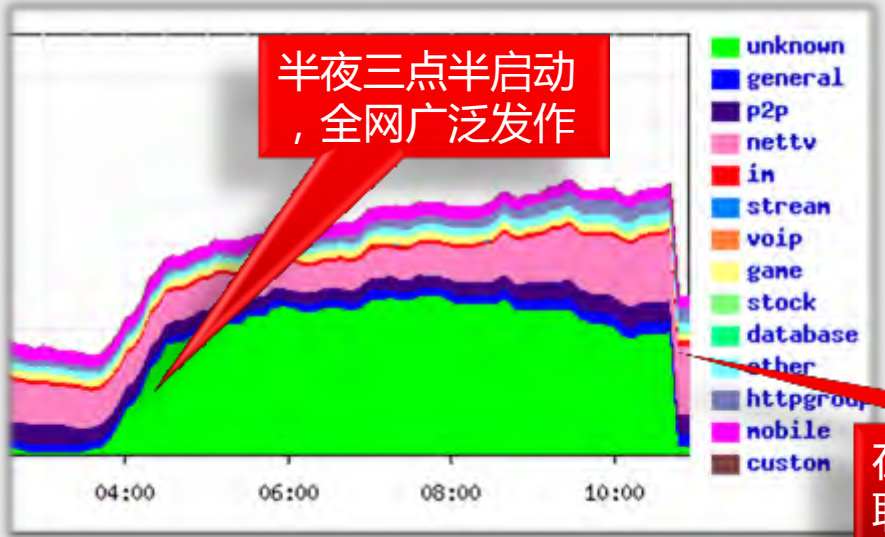
序号	域名	访问次数	上行流量	下行流量	总流量	操作
1	monitor.uu.qq.com	1881584	5.52G	3.51G	9.03G	
2	loc.map.baidu.com	1766395	10.82G	5.36G	16.18G	
3	mmsns.qpic.cn	1759599				
4	configsvr.msf.3g.qq.com	1701188				
5	inews.gtimg.com	1695213				
6	sdk.open.phone.igexin.com	1619056				
7	adash.m.taobao.com	1610455				
8	dc.51y5.net	1607358				
9	click.hd.sohu.com.cn	1585586				
10	irs01.com	1581510				
11	rqd.uu.qq.com	1575842				
12	ma.3g.qq.com	1571284				
13	mobads-logs.baidu.com	1556379				
14	i.gtimg.cn	1542240				
15	msg.71.am	1470466				
16	mbdlog.iqiyi.com	1429592				
17	dlied1.qq.com	1383214				
18	omgmta.qq.com	1377206				
19	wx.qlogo.cn	1374550				
20	android.bugly.qq.com	1348349				


```
"st" = "m"
"uhid" = "a00000000000000000000000000000001"
"dcType" = "005021"
"sign" = "D57843D62EF602EC0A0F3036EFB5FDD7"
"pid" = "00500101"
"msg" = "{"cts":"1467819787260","aid":"42a188baff4dce4"}"
"capBssid" = "14:e6:e4:d6:ba:9a"
"mac" = "00:9a:cd:ad:17:ea"
"mapSP" = "a"
"verName" = "4.1.28"
"userToken" = ""
"verCode" = "3058"
"appId" = "A0008"
"netModel" = "w"
"capSsid" = "yang1"
"chanId" = "guanwangn160704"
"ts" = "1467819788123"
"lati" = "25.312966"
"longi" = "110.306319"
"imei" = "A00000592EC4CE"
"dhid" = "f074429347df42c88581bd16a172bfe3"
"lang" = "cn"
"origChanId" = "guanwangn160704"
```

趋势实例-QQ空间病毒

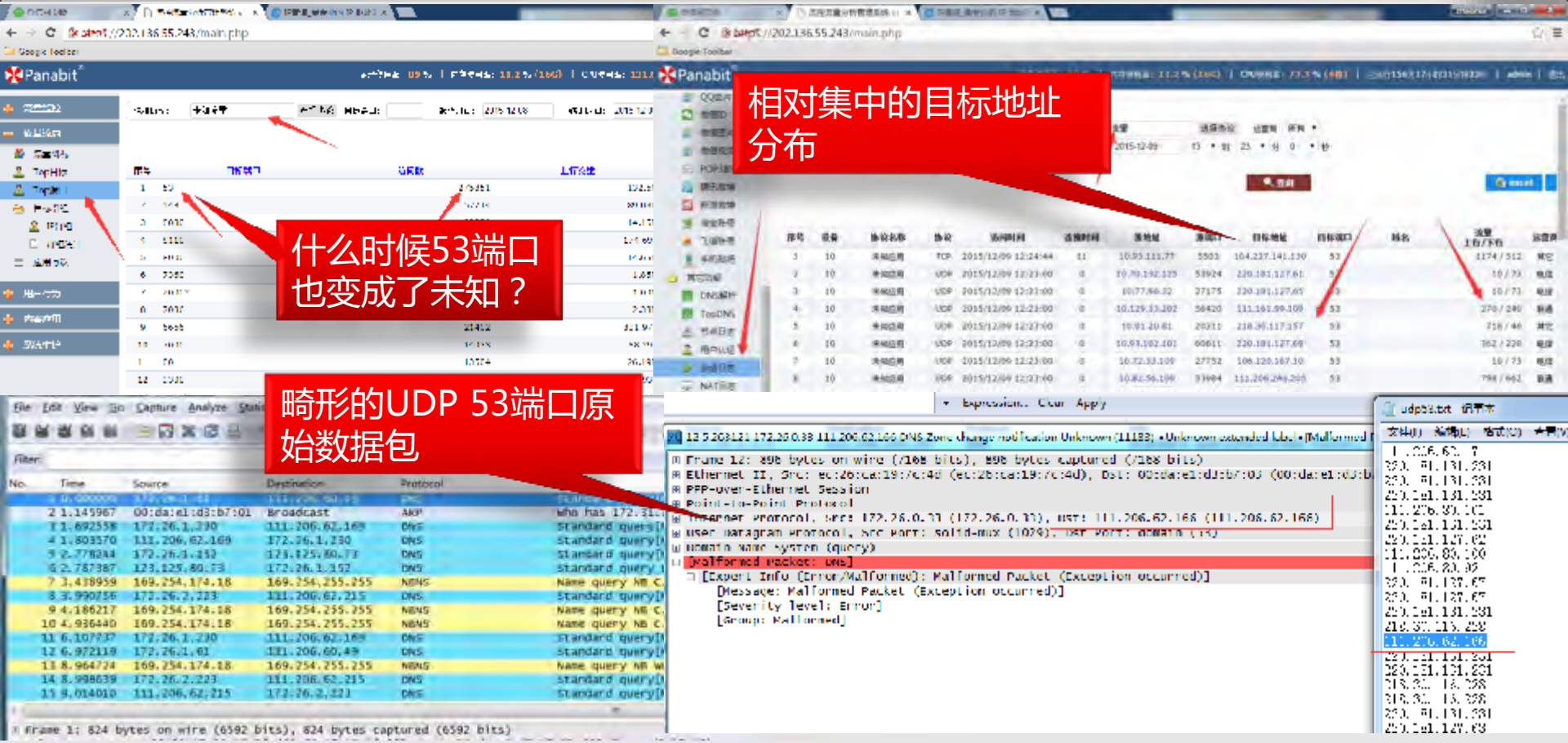


分布实例-磊科蠕虫事件



在Panabit协助下，中国互联网11:00启动全网封堵

相似性实例-真假DNS



全程开放策略

探针可以根据数据类型，以标准格式输出日志信息到PanaLog或者任何第三方接收平台。

Panabit运行的PANAOS操作系统可以加载第三方APP，在保证性能和稳定性前提下扩展探针功能。

PanaLog预制大数据接口，可在监控清洗基础上，将需要的日志输出到后端大数据平台。

Panabit