The Reputation of Networks – RIPE Region

Manish Karir, Kyle Creyts

(Merit Network Inc)

Outline

- Goal
- Background: IPv4 address allocation distribution in RIPE, commonly used blocklists
- Analysis
  - foreach(country, asn, bgp prefix)
    - SPAM Lists Distribution
    - Malware/Phishing Lists Distribution
    - Active Malicious Activity Lists
    - Highlight points of interest in data
- Network Reputation Discussion

The Problems

- How do you create incentives for the need to run a clean network
- How do you measure the relative security posture of a given network
- How do you balance the need to communicate with another network with the risks
- How can you estimate the likelihood of malicious activity from another network
- Can you assign a risk metric with a BGP path

➢ *You need to know about the historical and current reputation of networks*

Network Reputation

- Network reputation is an attempt to construct a metric or set of metrics that illustrate the collective reputation of all hosts in your administrative domain
- While infected hosts and botnets are a fact of life, how much of such activity represents an acceptable level of network pollution 1%? 10% of all hosts?
- Hosts that engage in malicious activity such as spam, phishing, malware, scanning in a network reduce the externally visible global network reputation of that network – it does not go un-noticed
- Reputation of hosts on your network has an impact on the usability of your network as portions might get blocked for various services
- It can be seen that not all networks are equal when it comes to network reputation. What policies, topology, connectivity, other factors make some networks better than others? How can we learn from them?

Reputation based security policies

- Network reputation is not just something other people know about you.
- You can use it to craft flexible local policies which better manage your risk profile
- We are creating an index of reputation for networks based on aggregation of many diverse sources of reputation data

Reputation-based Security Policies

- Some Interesting possibilities:
  - BGP : For each path compute the relative reputation over entire path or the lowest hop AS in any path and influence policy to avoid that path
  - SPAM Scoring – use reputation of source in scoring, but more interestingly – bypass other checks if reputation is > 95

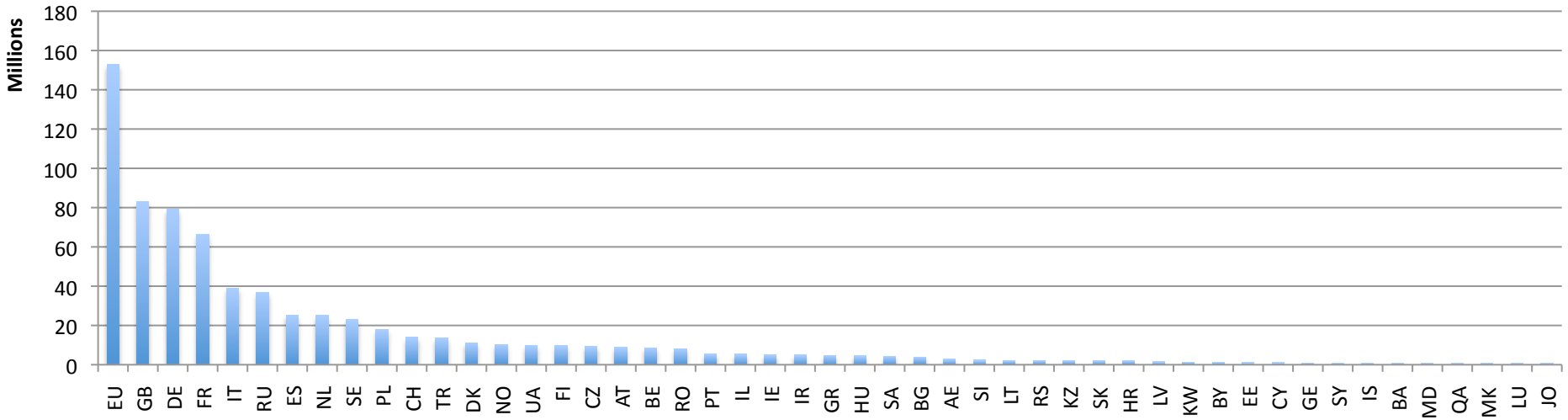
Other interesting possibilities

- Inbound Firewall – Allow all traffic to regular servers from sources with reputation > 10 but for reputation < 10 send traffic to alternate servers or services, require additional authentication etc
- Outbound traffic – disallow traffic to networks with poor reputation
- Making DPI viable/scalable for more people – normally route traffic for reputation > 30 but for poor reputation sources pass traffic through DPI for further inspection

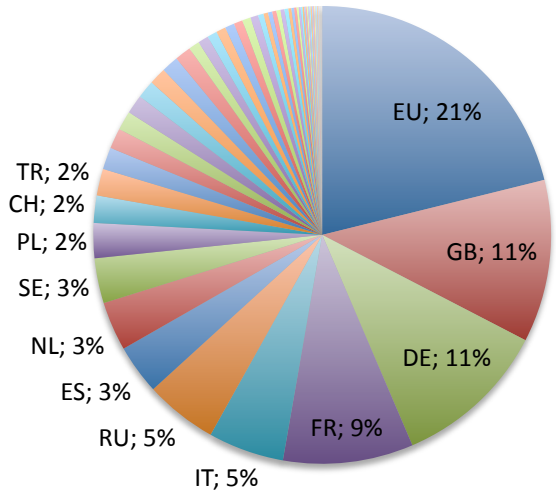
Common -Host- Reputation Block Lists (RBLs)

- RBLs are mostly lists of IP addresses of domains that have been observed to participate in suspicious behavior
- RBLs can be clustered by type of activity on which it is based:
  - SPAM Lists: SPAMHAUS(CBL), BRBL, SpamCop, wpbl, UCEPROTECT
  - Malware/Phishing hosts: SURBL (multi), phishtank, hpHosts
  - Active Attack Behavior: Darknet Scanner (merit), Dshield, ssh brute-force (fail2ban, denyhosts)
- Our goal is to analyze relative distribution of hosts on these lists to determine if there are some common traits that can broadly characterize the observed relative malicious activity originating from a country, ASN, and prefix

RIPE Address Space Distribution by Country



- Roughly 2.8M /24 blocks allocated ~ 733M IP addresses
- EU is 21% of allocations, GB, DE, FR, together account for another 30% of all allocations

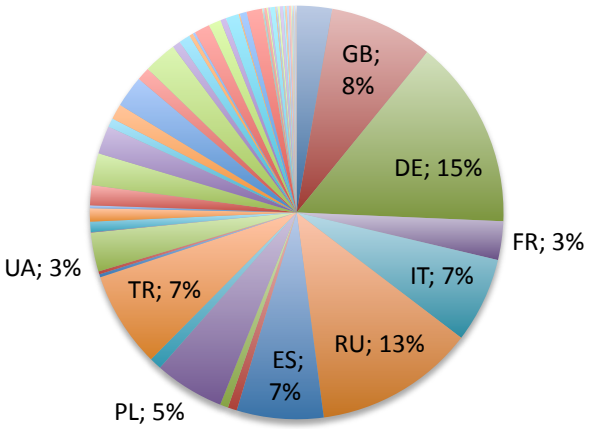


SPAM Lists Distribution Analysis

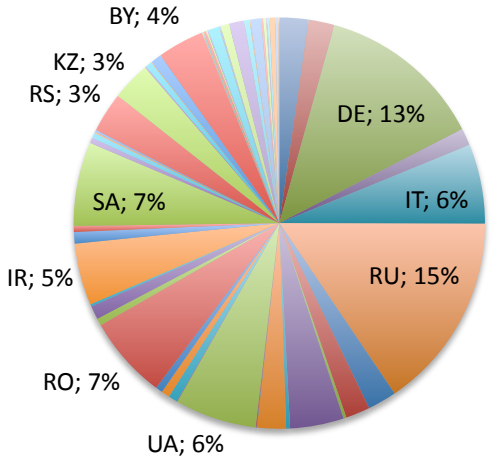
- Consider 3 largest/most popular SPAM Lists:
  - Barracuda BRBL
  - SPAMHAUS – CBL
  - SpamCop
  - Other SPAM data sources as well such as weighted private block list (wpbl), UCEPROTECT also analyzed but omitted here due to similarity
- Determine portions of those lists relevant to the RIPE region
- Determine relative distribution by country within RIPE region

SPAM Lists Distribution by Country

barracuda

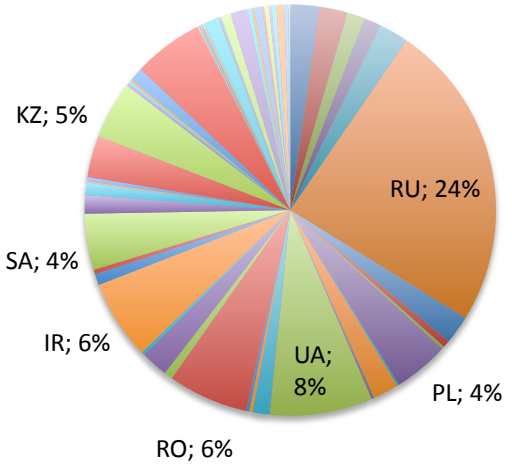


cbl



List	Total IPs	RIPE IPs
Barracuda	128M	65M(17%)
SPAMHAUS CBL	8.1M	2.6M(12%)
SpamCop	325K	66K(8%)

spamcop



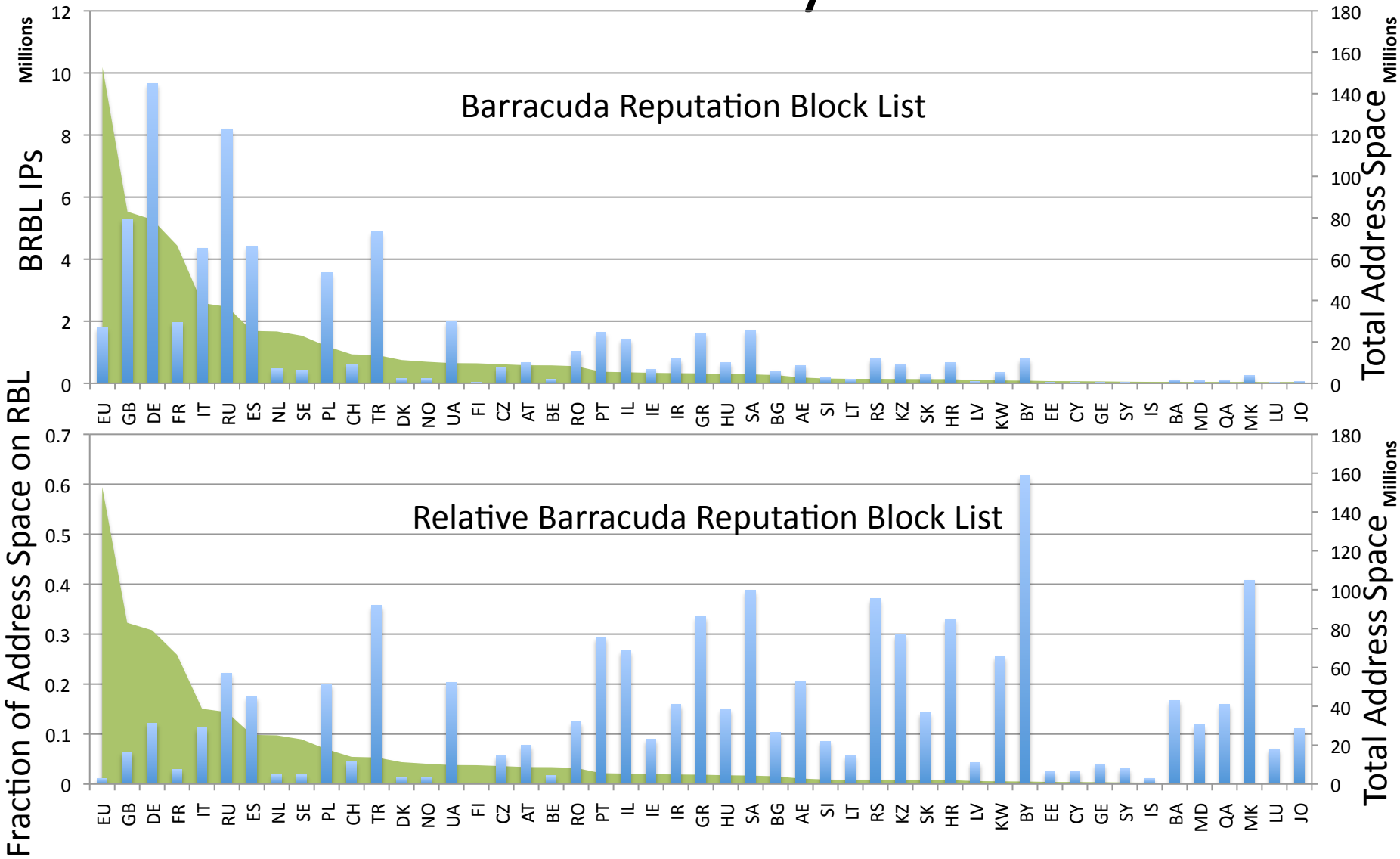
SPAM List Relative Distribution

- In general: countries with larger allocations have more entries in block lists
- Expected – if you assume infection rates are a steady fact of life.
- This assumes that on average, a constant % of any given IP address range will be on a block list

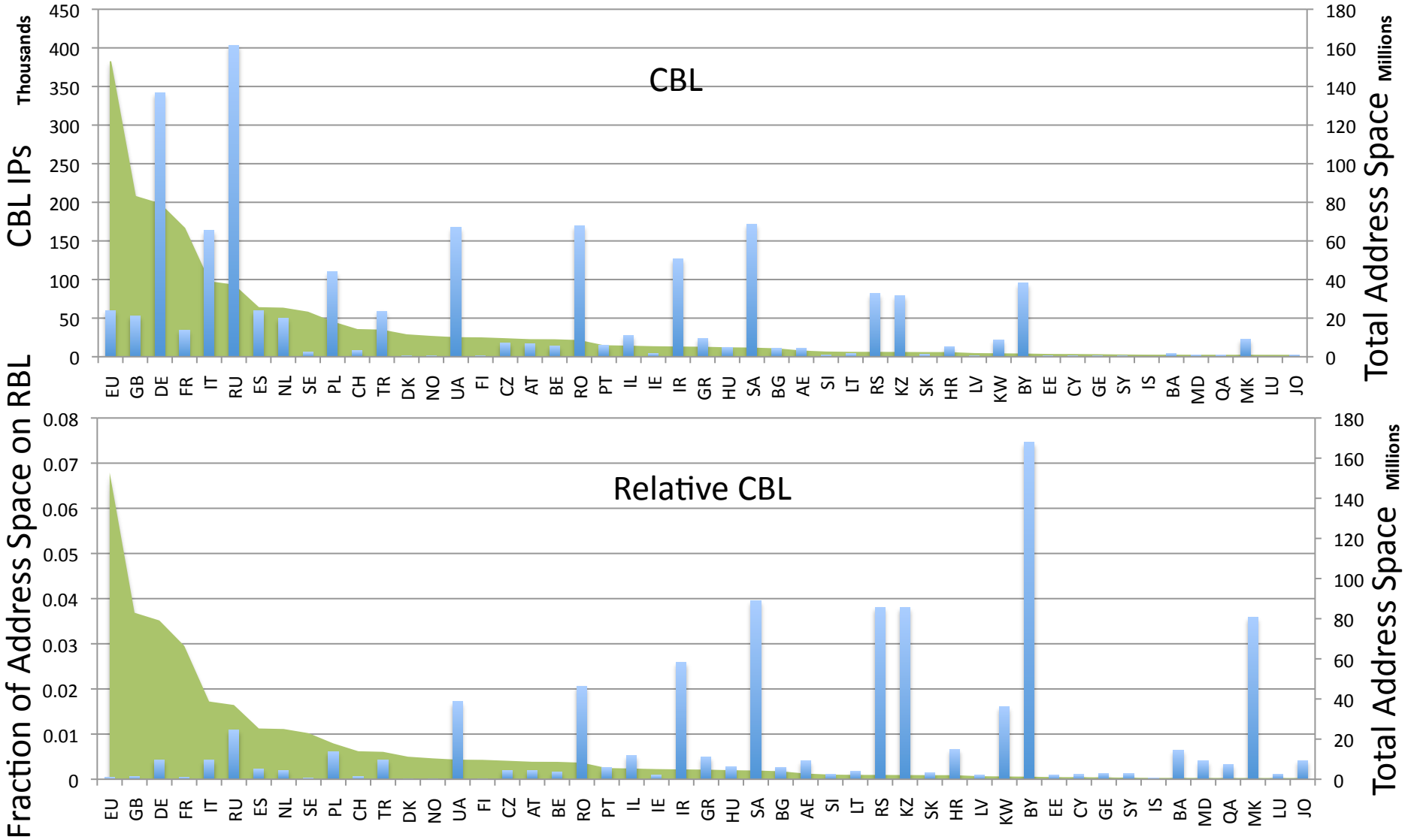
Is this true?

- What happens when we look at block list entries relative to allocation sizes
- We should look at both the large and the small ends of allocation spectrum to remove allocation size from the equation
- What do we expect to see?

Relative SPAM List Distribution by Country



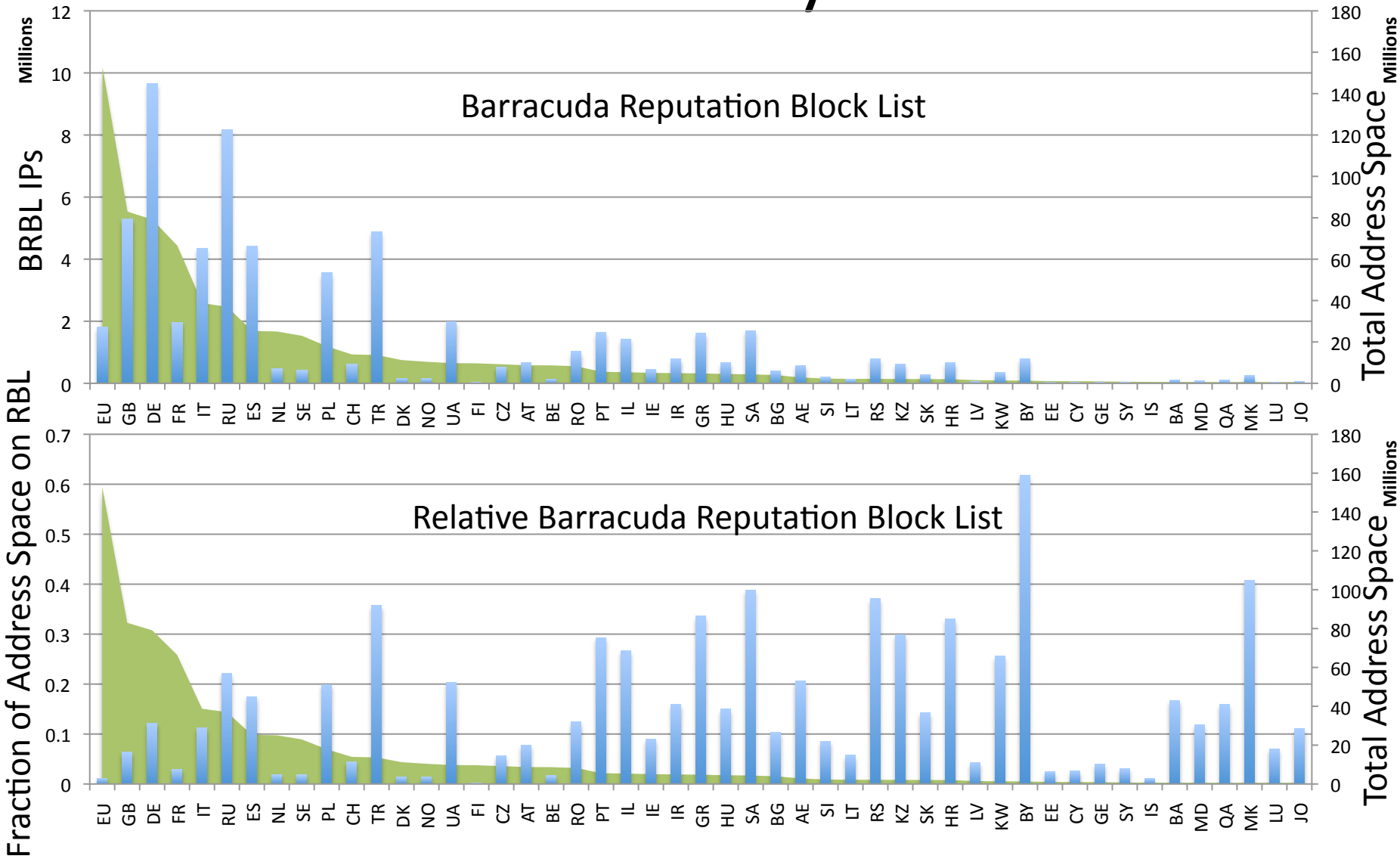
Relative SPAM List Distribution by Country



SPAM List Discussion

- All networks are not created equal when it comes to entries on a SPAM list
- Interesting things to notice:
  - Almost 65% of Belarus is on BRBL
  - Almost 40% of Saudi Arabia is on BRBL
  - Almost 35% of Turkey is on BRBL
  - Only 10% of Germany but that is a lot of IPs
  - More than half of the countries have greater than 10% of their IP addresses on BRBL
  - Given the allocation sizes Netherlands, Sweden, Denmark and Norway have unusually low listing rates on BRBL
  - Smaller percentages of listed IPs on other lists but the relative trends between countries seem to be the same
- What accounts for these regional variations? Local policy? Connectivity? Network topology?

Relative SPAM List Distribution by Country

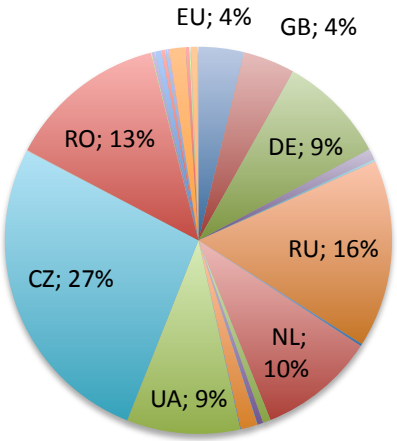


Malware/Phishing Lists Distribution Analysis

- Consider 3 common malware/phishing Lists:
  - SURBL-multi
  - hpHosts
  - phishtank
  - Other popular data sources such as malwaredomains and malwaredomainsList are included in the SURBL-multi dataset
- Use same methodology as SPAM analysis

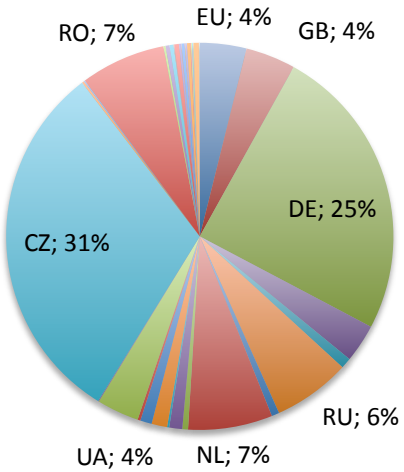
Malware/Phishing Lists by Country

surbl

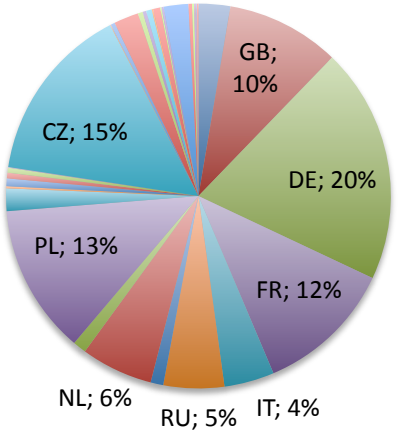


List	Total IPs	RIPE IPs
SURBL	360K	107K
Hphosts	185K	71K
Phishtank	4700	1700

hphosts



phishtank

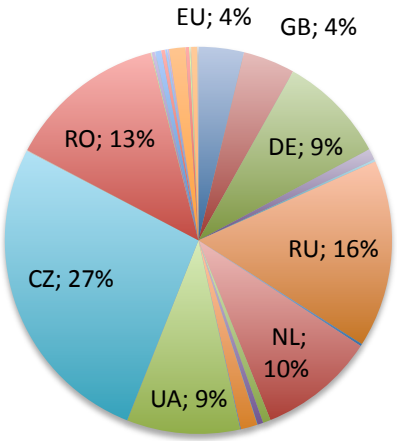


Malware/Phishing Discussion

- Czech Republic relatively higher percentage of Malware/Phishing listed domains ~ 30% of all RIPE region domains
- Poland and France have a unusually high percentage of IPs listed as hosting phishing sites.
- Aside from Russia there appears to be little in common with SPAM blocklists

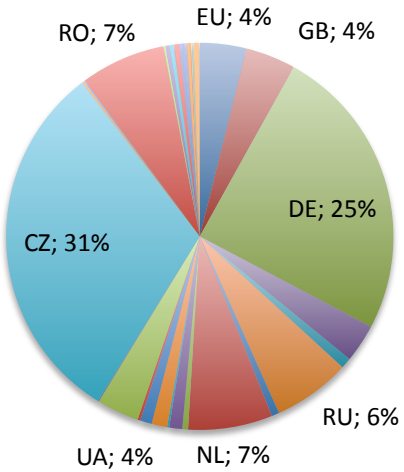
Malware/Phishing Lists by Country

surbl

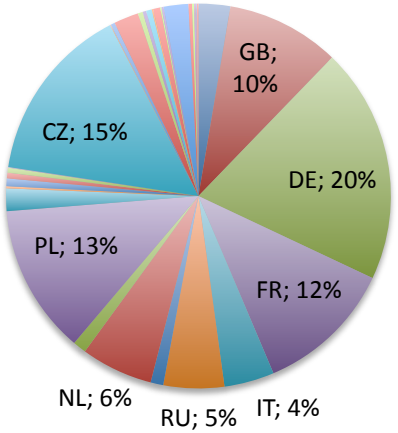


List	Total IPs	RIPE IPs
SURBL	360K	107K
Hphosts	185K	71K
Phishtank	4700	1700

hphosts

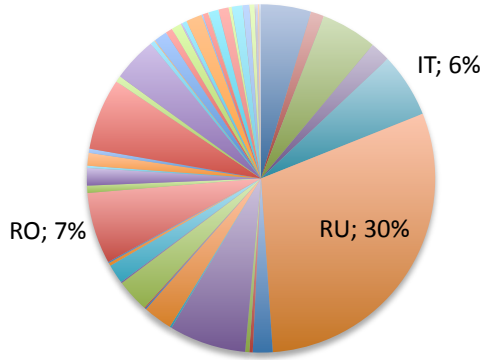


phishtank

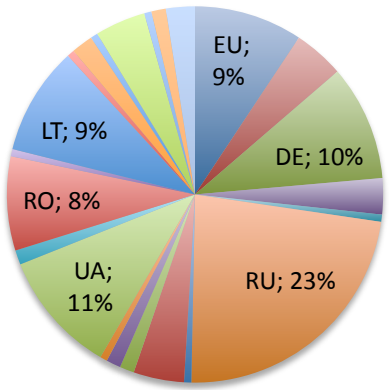


Active Malicious Activity by Country

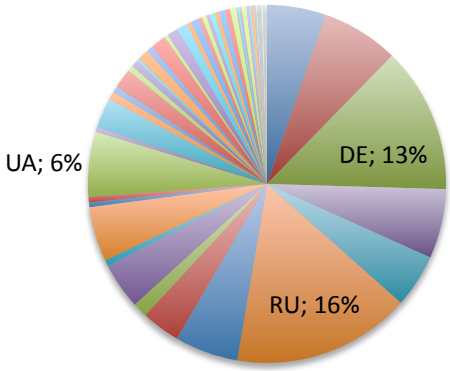
Darknet Scanning



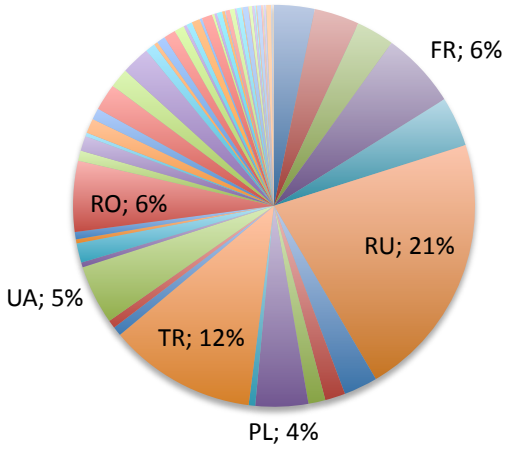
zeus



ssh bruteforce



dshield



Active Malicious Activity Discussion

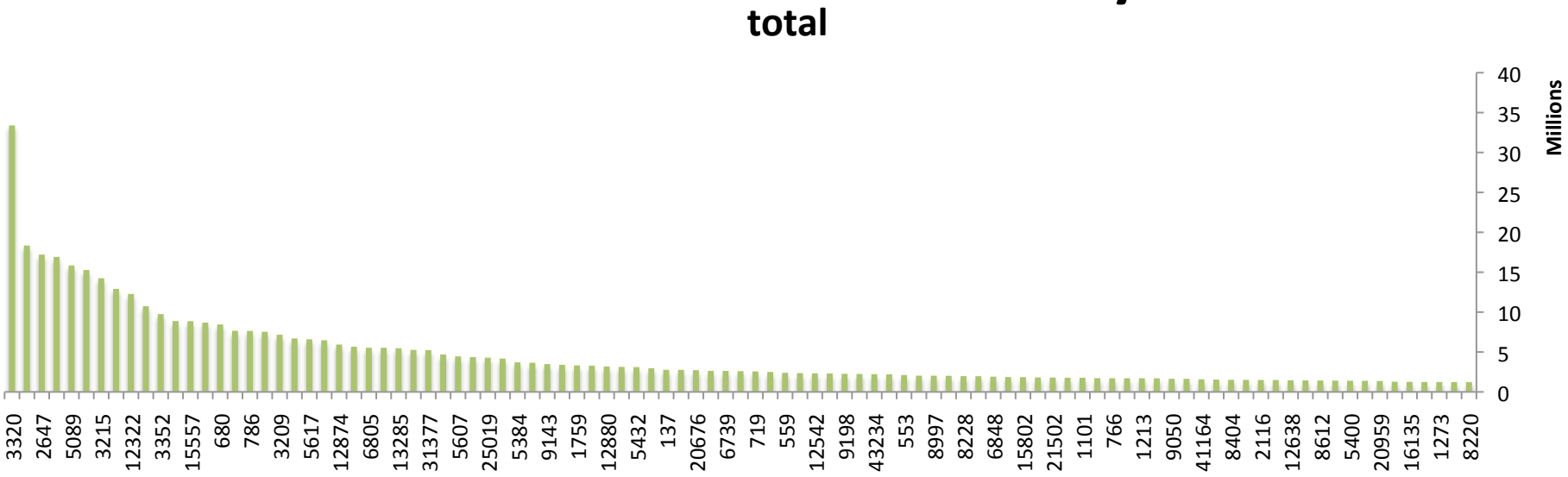
List	Total IPs	RIPE IPs
ssh brute-force	68K	22K
Dshield	754K	314K
Darknet Scanning	156K	83K
Zeus	215	161

- Russia accounts for ~ 30% of darknet scanning activity from RIPE region
- Ukraine, Lithuania, Romania togther account for 30% of Zeus C&C in RIPE region, Russia is an additional 23%
- Unsually lower IP listings from France

Boundaries

- What are effective boundaries
- What other boundaries can we use
- National borders are not good enough for making truly effective policy decisions
- Certainly not good enough for making service decisions

Address Distribution by ASN

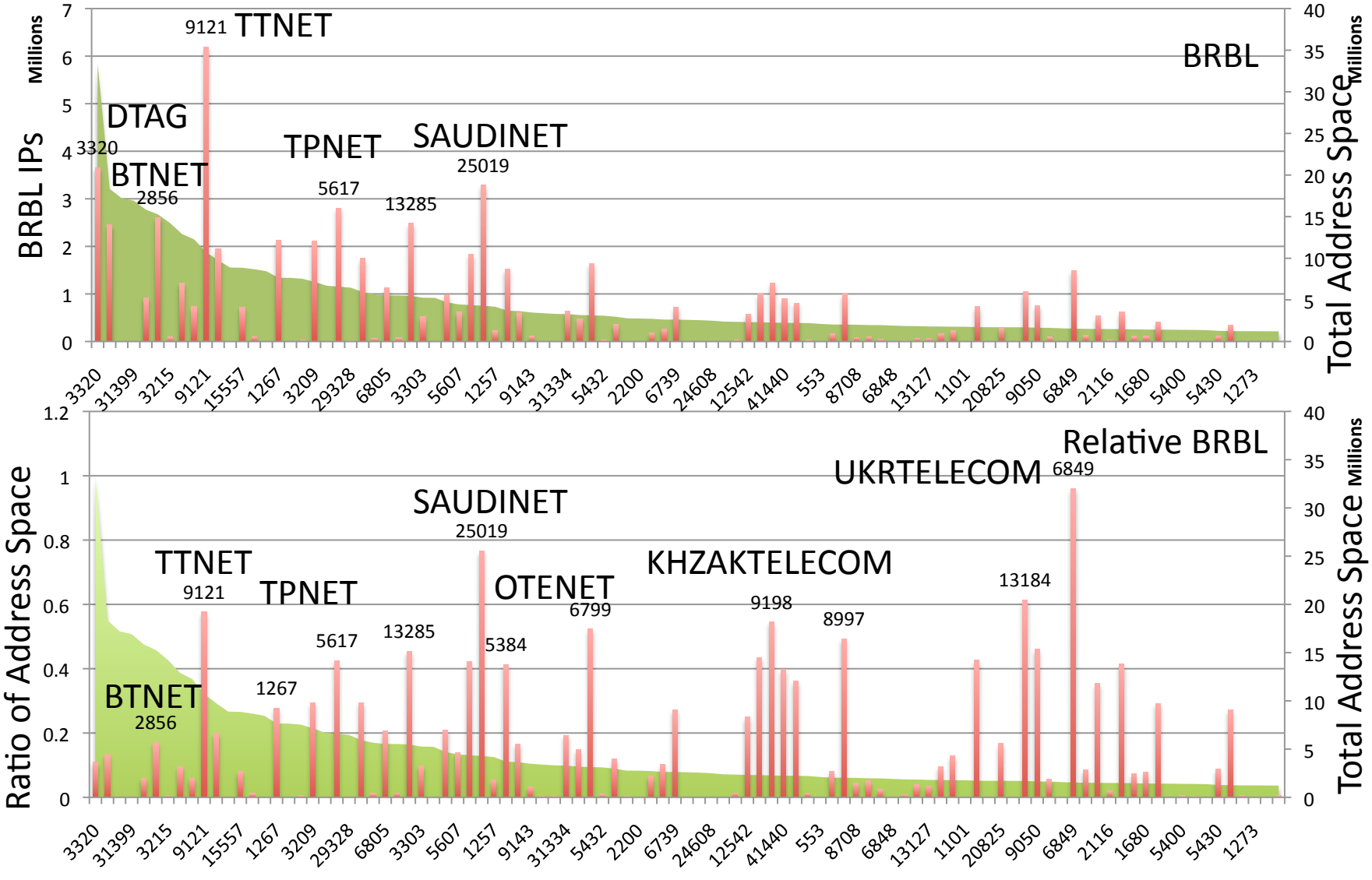


- Roughly 16.5K ASNs in use in RIPE region
- They account for roughly 88.1K of prefixes in the BGP routing table (total 360K entries)
- A total of 733.6M IPs
- We focus on the largest 100 ASNs
- Total number of IPs announced by these ASNs varies from 34M to 1.2M

Top 10 ASNs by Size

ASN	Name	IP Addresses
3320	DTAG Deutsche Telekom AG	33M (4.5%)
3269	ASN-IBSNAZ Telecom Italia S.p.a.	18M (2.4%)
31399	DAIMLER-AS Daimler Autonomous System	17M (2.3%)
5089	NTL NTL Group Limited	17M (2.3%)
2856	BT-UK-AS BTnet UK Regional network	16M (2.2%)
3215	AS3215 France Telecom - Orange	15M (2.0%)
6830	UPC UPC Broadband	14M (1.9%)
12322	PROXAD Free SAS	13M (1.7%)
9121	TTNET Turk Telekomunikasyon Anonim Sirketi	12M (1.6%)
3352	TELEFONICA-DATA-ESPANA TELEFONICA DE ESPANA	10M (1.4%)

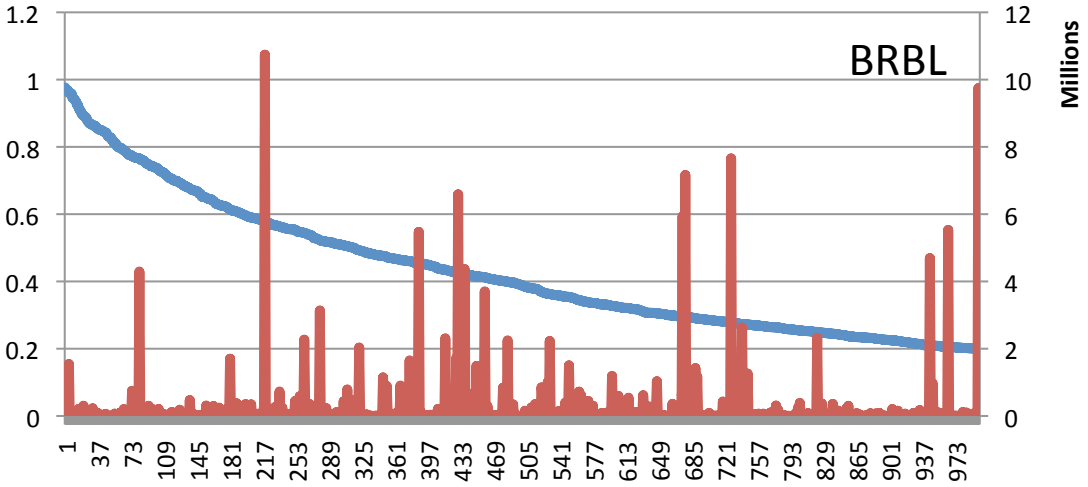
SPAM List IP Distribution by ASN

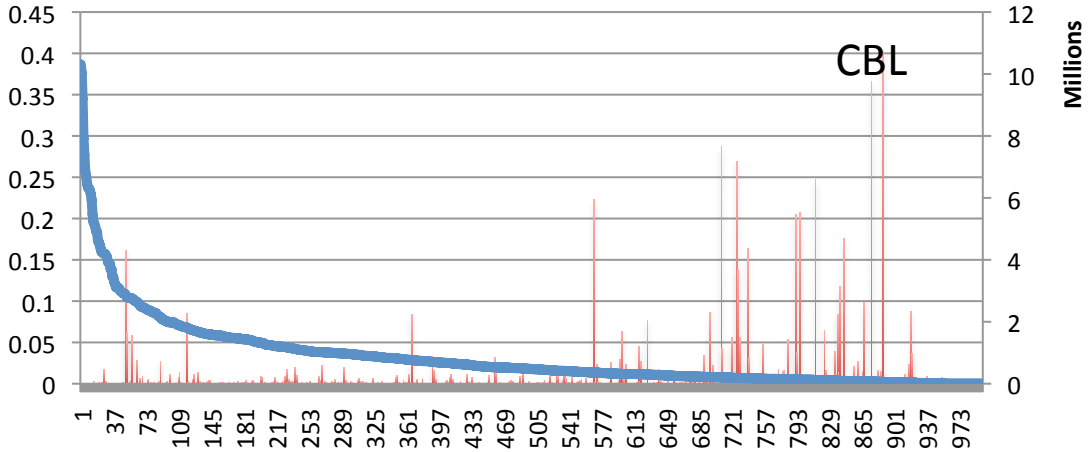


SPAM List IP Address Distribution by ASN Discussion

- Top 10 network AS9121 TTNET accounts for 6M IPs in BRBL which is almost 60% of its total
- AS 2647 SITA which has 17M IPs has negligible number of BRBL and CBL entries similar trend for AS3215 – France Telecom
- AS6849 UKRTELECOM is almost entirely on BRBL
- 15 of the largest 100 ASNs have more than 40% of their address space on the BRBL

ASN IP Blocklisting Distribution

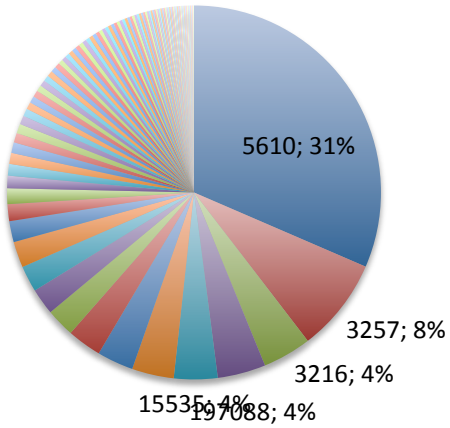




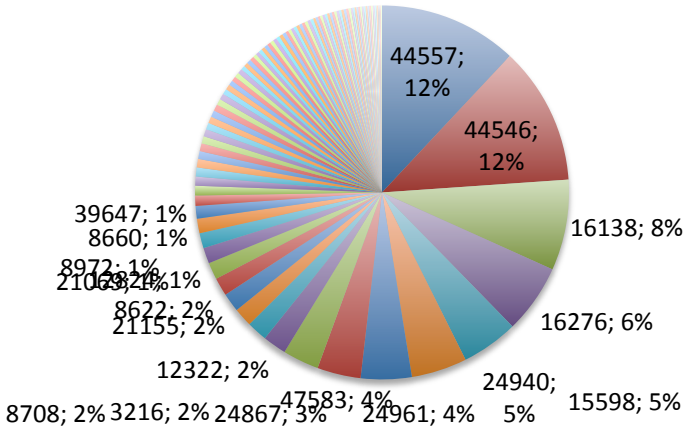
- Top 1000 ASNs with largest percentage of their networks on SPAM blocklists
- Almost 500 ASNs have atleast 40% of their Ips on BRBL
- Almost 200 ASNs have at least 5% of their IPs on CBL

Malware/Phishing Domains Distribution by ASN

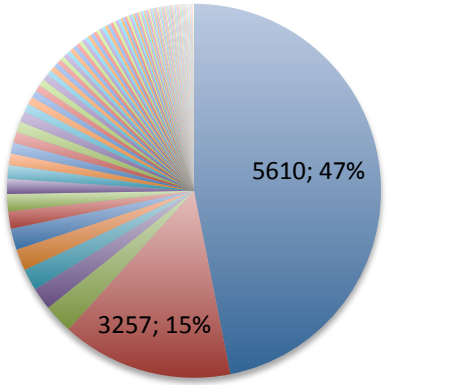
surbl



phishtank



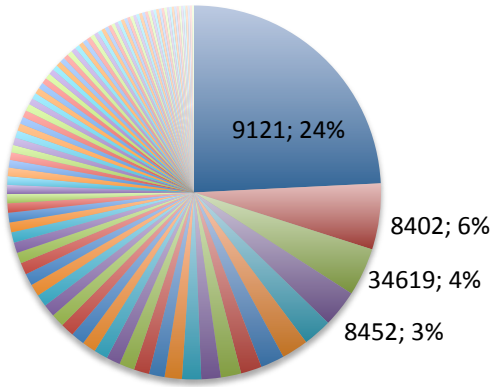
hphosts



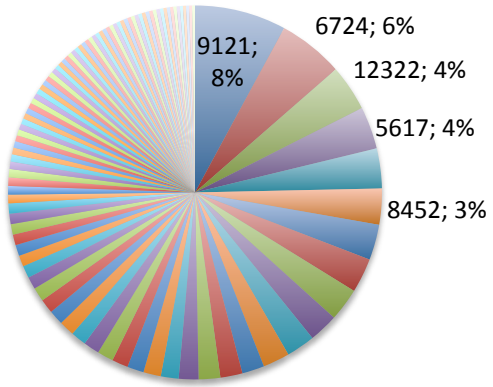
- AS5610 - Telefonica O2, Czech Republic represents 30% of SURBL RIPE region entries and 47% of hphosts entries
- AS 3257 - TINET-BACKBONE Tinet is the next highest contributor
- AS 44557 (Dragonara) and AS4546 (AlfaTelecom)- together represent 25% of the RIPE region phishtank listings

Active Malicious Activity by ASN

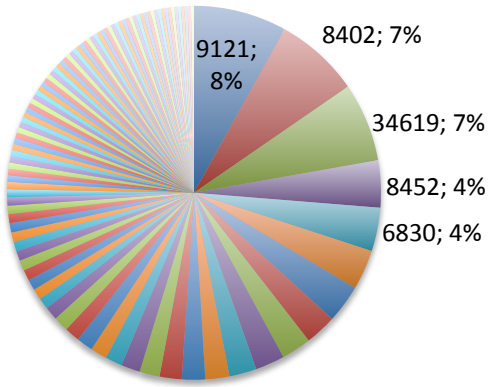
Darknet Scanners



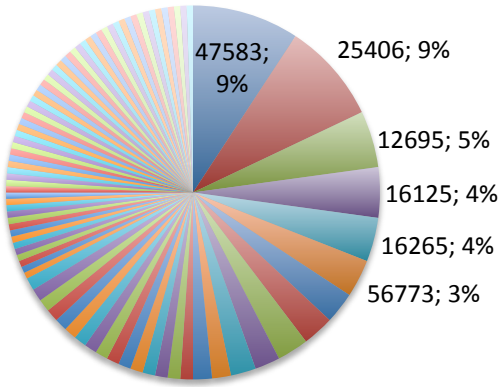
denyhosts



dshield



zeus



Active Malicious Activity Discussion

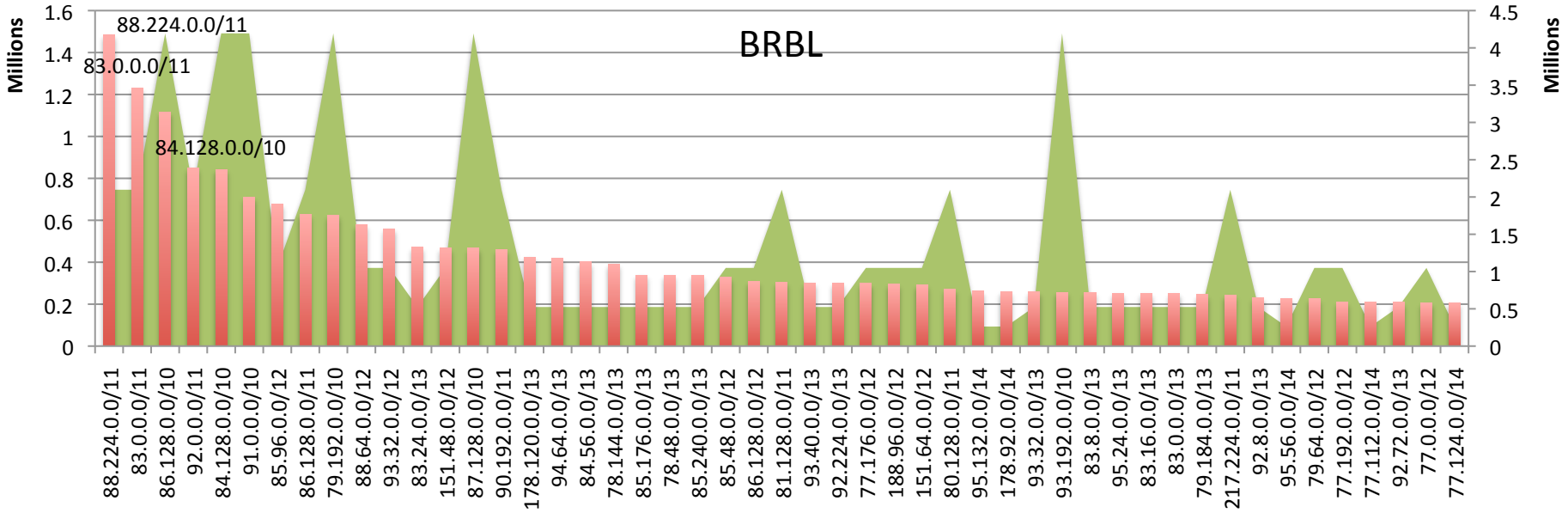
List	Total IPs	RIPE IPs
ssh brute-force	65K	22K
Dshield	754K	314K
Darknet Scanning	158K	83K
Zeus	215	161

- AS9121 - TTNET Turk Telekomunikasyon accounts for almost 25% of darknet scanning IPs from RIPE region
- AS9121 – TTNET Turk Telekomunikasyon is also almost 10% of IPs on ssh-brute-force activity lists as well as dshield. Unsually lower IP listings from France
- Zeus list IPs too few for meaningful results but more than half of all reported C&C IPs are in RIPE region.

Are ASNs the most useful boundary?

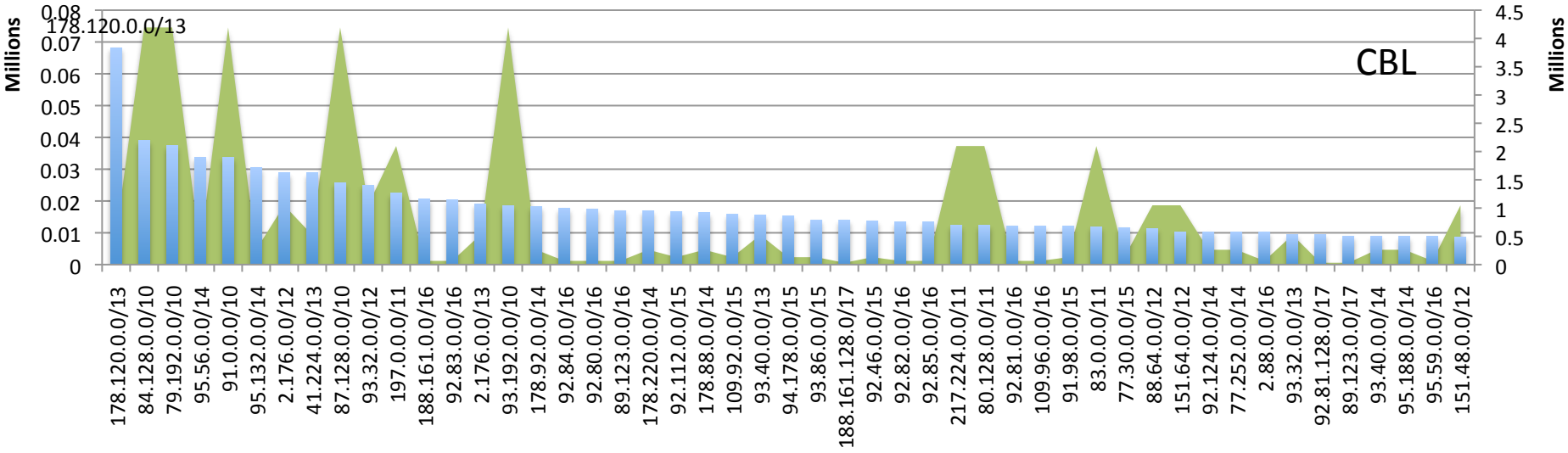
- ASN does not necessarily indicate administrative domain
- How can we more effectively identify administrative domains
- We used prefixes observed from our view
- Quality of reputation data, and proper identification of administrative domains should evolve from additional views

BGP Prefix SPAM List IP Distribution



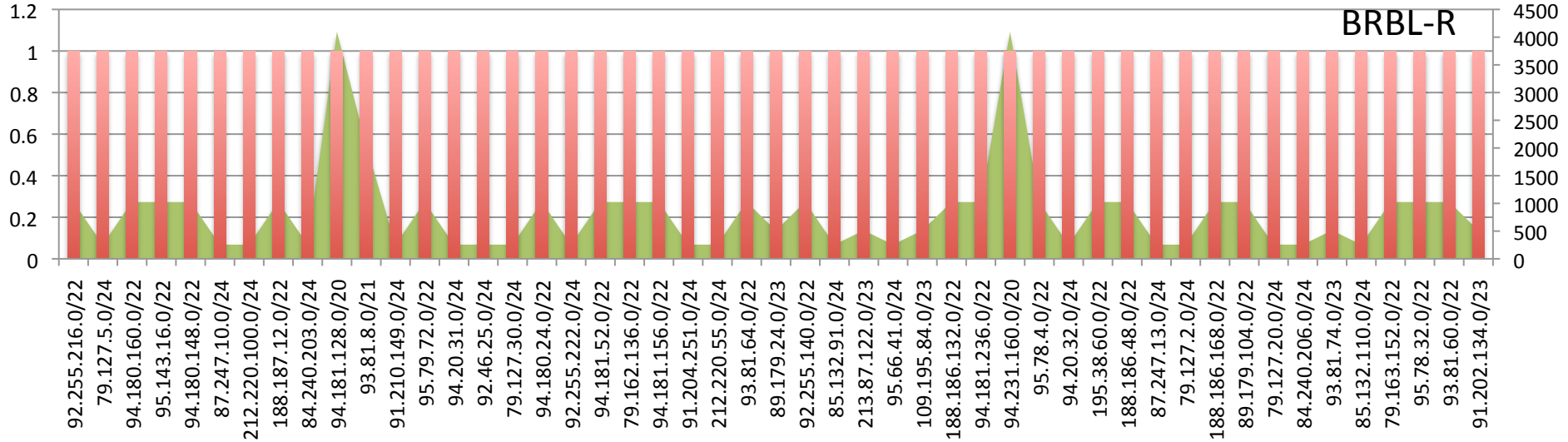
- BGP RIPE region prefixes 88350 out of total routing table of ~360K
- No surprise that large prefixes have large numbers of IPs in BRBL
- BUT – still a surprise that 15 prefixes have over 500K IPs in the BRBL
- 88.224.0.0/11 – Turk Telcom has 1.4M IPs out of an allocation of 2M on BRBL
- 83.0.0.0/11 - Telekomunikacja Polska S.A has 1.2M Ips out of 2M on BRBL
- All 50 prefixes shown above have atleast 200K Ips on BRBL or atleast 780 /24 blocks

BGP Prefix SPAM List IP Distribution



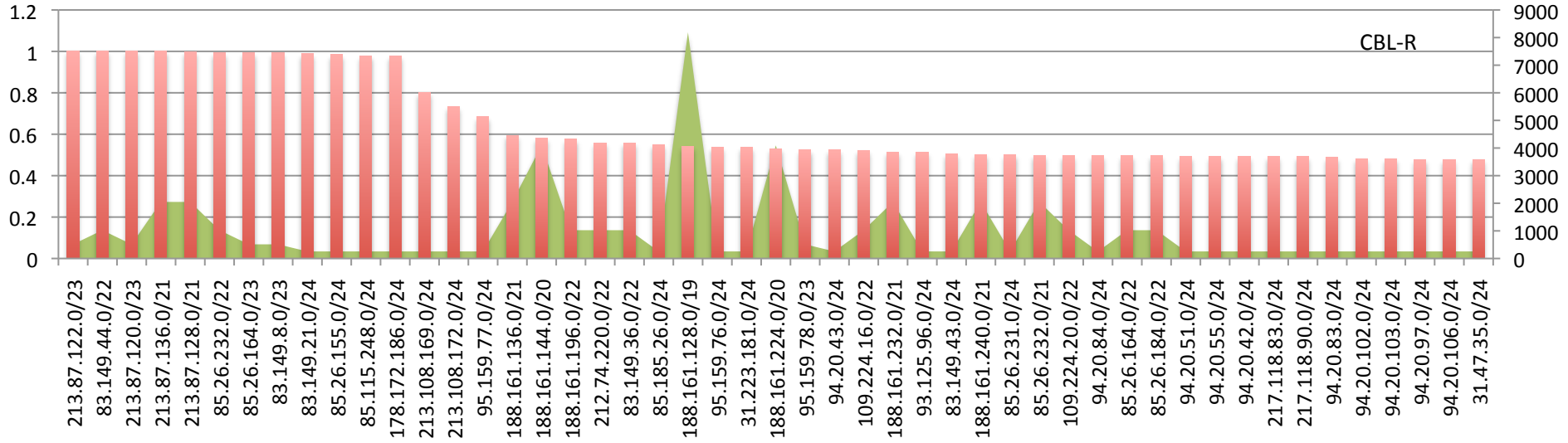
- Even for CBL all 50 of the prefixes shown above have atleast 7.5K IPs listed
- 178.120.0.0/13 – BELTELECOM has almost 70K IPs listed in the CBL
- 84.128.0.0/10 - Deutsche Telekom AG has roughly 35K IPs on CBL and 80K IPs in the BRBL

Relative Amounts of IP addresses in SPAM lists



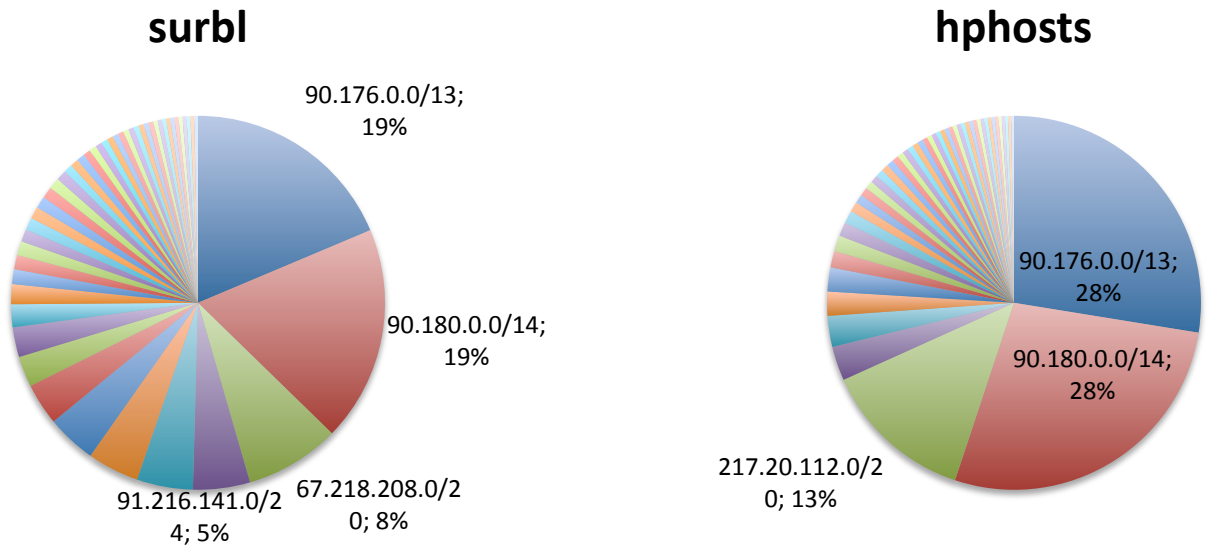
- 253 prefixes are completely included
- Over 3500 prefixes out of all RIPE region prefixes have over 85% of their IP address block listed in the BRBL

Relative Amounts of IP Address in SPAM Lists



- 12 prefixes mostly /24 - /23 have over 90% of their IPs listed in CBL
- All 50 of the prefixes shown above have at least 50% of their IPs listed in the CBL

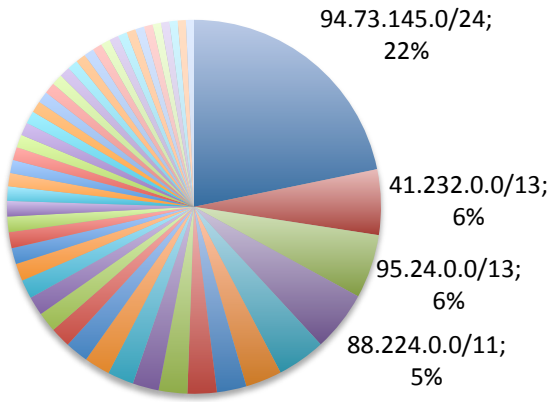
Malware/Phishing Hosting IP Address Distribution



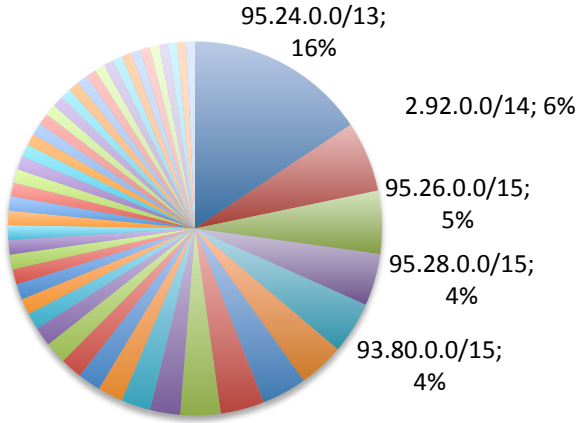
- Relative percentages of IPs for the top 50 prefixes for each data type are shown above
- 90.176.0.0/13 and 90.180.0.0/14 - Telefonica O2 Czech Republic appear on both lists. Together 40% of SURBL entries and 56% of hphosts entries

Active Malicious Activity List IP Distribution

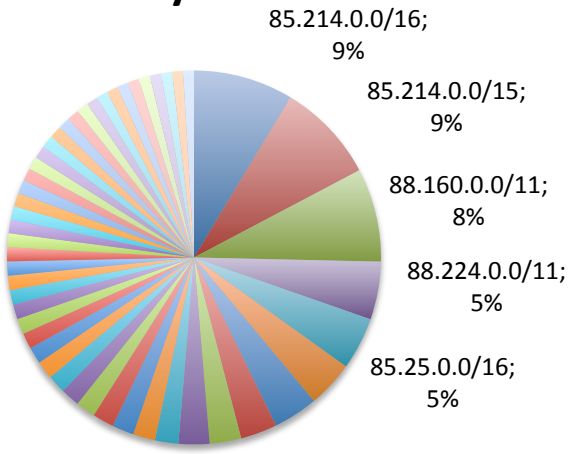
dshield



Darknet Scanners



denyhosts



- Relative percentages of IPs in the top 50 prefixes are shown above
- 95.24.0.0/13, 2.92.0.0/14, 93.80.0.0/15 and 95.26.0.0/15 - CORBINA TELECOM accounts for 31% of all scanning IPs in the top 50 prefixes in RIPE region
- 94.73.145.0/24 - Cizgi Telekom is almost 22% of the activity from top 50 prefixes in RIPE region
- 85.214.0.0/16 and 85.214.0.0/15 - Strato AG represent 18% of ssh brute-force activity
- 88.160.0.0/11 ProXad network – and 88.224.0.0/11 – Turk Telecom account for another 13%.

Global Regional Reputation Comparisons – SPAM RBLs

SPAM RBL List	Total IPs on RBL	ARIN IPs (1.65B)	LACNIC IPs (116M)	RIPE IPs (733M)	APNIC IPs (848M)
Barracuda	128M	8.8M (RBL: 6.8%)	22.7M (RBL: 17%)	65M (RBL: 51%)	32M (RBL: 25%)
SPAMHAUS CBL	8.1M	122K (RBL: 1.5%)	1M (RBL: 12%)	2.6M (RBL: 32%)	3.2M (RBL: 39%)
SpamCop	325K	3.2K (RBL: 1%)	28K (RBL: 8%)	66K (RBL: 20%)	125K (RBL: 38%)

ARIN Region has unusually low rates of member ship on SPAM lists,
RIPE region is comparatively high

Global Regional Reputation Comparisons – Malware RBLs

Malware RBL List	Total IPs on RBL	ARIN (Total IPs: 1.65B)	LACNIC (Total IPs: 116M)	RIPE (Total IPs: 733M)	APNIC (Total IPs: 848M)
SURBL	360K	194K (RBL: 54%)	3K (RBL: <1%)	107K (RBL: 30%)	51K (RBL: 14%)
Hphosts	185K	94K (RBL: 51%)	2K (RBL: <2%)	71K (RBL: 38%)	17K (RBL: 9%)
Phishtank	4700	2627 (RBL: 56%)	124 (RBL: < 3%)	1700 (RBL: 36%)	216 (RBL: 4%)

ARIN Region has unusually high rates of membership on malware lists,
RIPE region is also high, LACNIC and APNIC regions comparatively lower

Global Regional Comparisons – Active Malicious Activity RBLs

Active Malicious RBL List	Total IPs on RBL	ARIN (Total IPs: 1.65B)	LACNIC (Total IPs: 116M)	RIPE (Total IPs: 733M)	APNIC (Total IPs: 848M)
ssh brute-force	68K	11K (RBL: 16%)	11.6K (RBL: 17%)	22K (RBL: 32%)	20K (RBL: 29%)
Dshield	754K	128K (RBL: 17%)	61K (RBL: 8%)	314K (RBL: 42%)	224K (RBL: 29%)
Darknet Scanning	156K	7.8K (RBL: 5%)	28K (RBL: 17%)	83K (RBL: 53%)	36K (RBL: 23%)
Zeus	215	35 (RBL: 16%)	1 (RBL: 0%)	161 (RBL: 75%)	17 (RBL: 8%)

RIPE region has comparatively higher rates of membership on active malicious activity lists

Conclusions and Future Work

- Our goal is to develop a comprehensive global network reputation system that computes, for each prefix you observe in the BGP routing table, a reputation metric.
- Variations can allow arbitrary network boundaries not simply BGP boundaries but that is the starting point
- Data from common sources such as RBLs are a starting point for bootstrapping the reputation system, however in order to be successful the system must have data from many many vantage points
- Different networks have different views of reputations of other networks
- The more vantage points you have the closer to “true reputation you will get”
- The system must allow all networks to participate and contribute reputation information regarding all other networks while being resistant to collusion and false reporting
- Current project at Merit Network Inc is building such a system and an effort will soon be made to recruit participant networks on various mailing lists
- If you would like to participate please send email to: mkarir@merit.edu
- How reputable is your network?