# Quantifying Cloud Misbehavior


**Rajat Tandon**
*rajattan@usc.edu*

PI: Dr. Jelena Mirkovic
*sunshine@isi.edu*

Pithayuth Charnsethikul
*charnset@usc.edu*

*University of Southern California Information Sciences Institute*


*https://steel.isi.edu/Projects/Cloud_Misbehavior/*


Published in 2020 IEEE International Conference on Cloud Networking


USC

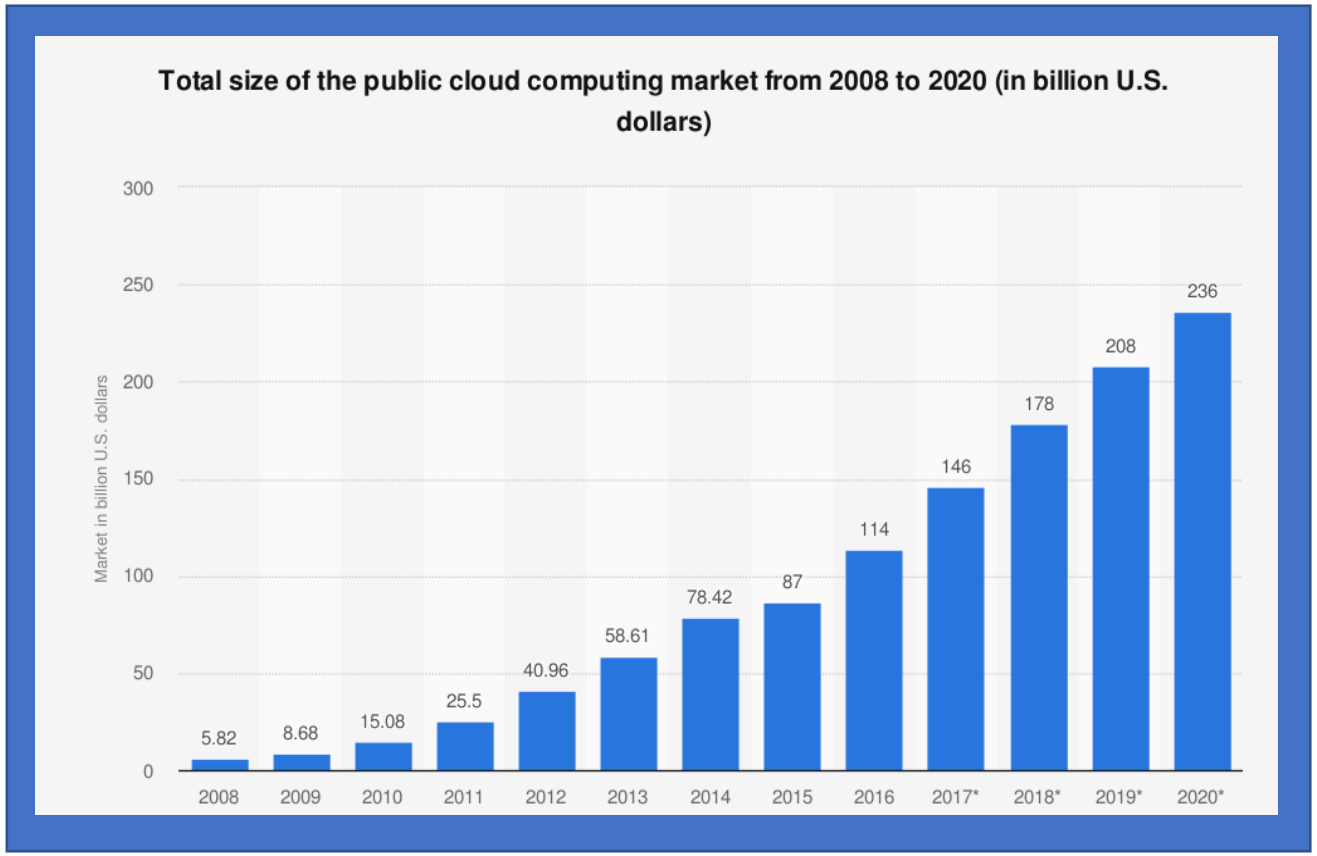
ISI Information Sciences Institute




Funded by the Department of Homeland Security, USA - BAA HSHQDC-14-R-B0005, and the Government of UK - D15PC00184

# Background

- Clouds have gained popularity over the years. They provide:
  - High storage capacities
  - High computing power
  - Reduced hardware costs
  - On-demand availability

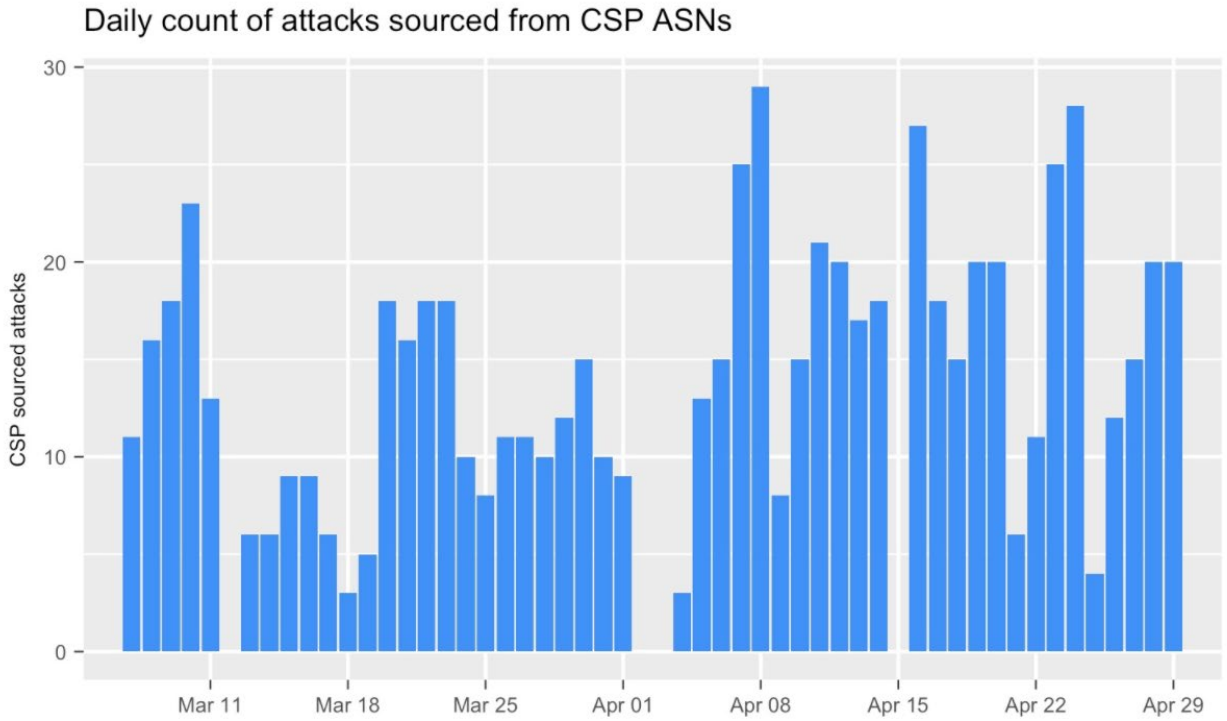



https://www.statista.com/statistics/510350/worldwide-public-cloud-computing/

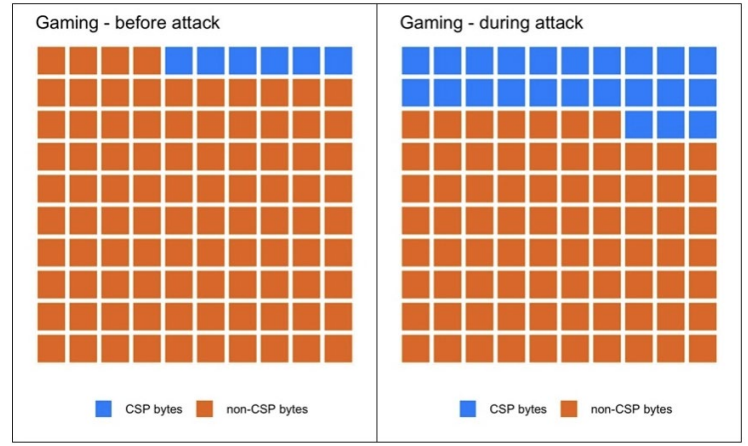
# Background

- Cloud users often gain superuser access to cloud machines
- But superuser access without the support  of experienced system administrators, can create fertile ground for accidental or intentional misuse
- Attackers can rent cloud machines or hijack them from cloud users
- They leverage them to generate unwanted traffic
- Bulletproof hosting permits malicious traffic generation

# Motivation




https://blogs.akamai.com/2019/05/do-ddos-attacks-originate-from-cloud-service-providers.html




Home › Network Security

**Cybercriminals Abuse Amazon Cloud to Host Linux DDoS Trojans**

By Eduard Kovacs on July 28, 2014

RSS

Cloud services provided by Amazon and other companies are being abused by profit-driven cybercriminals to host DDoS bots, Kaspersky Lab reported on Friday.

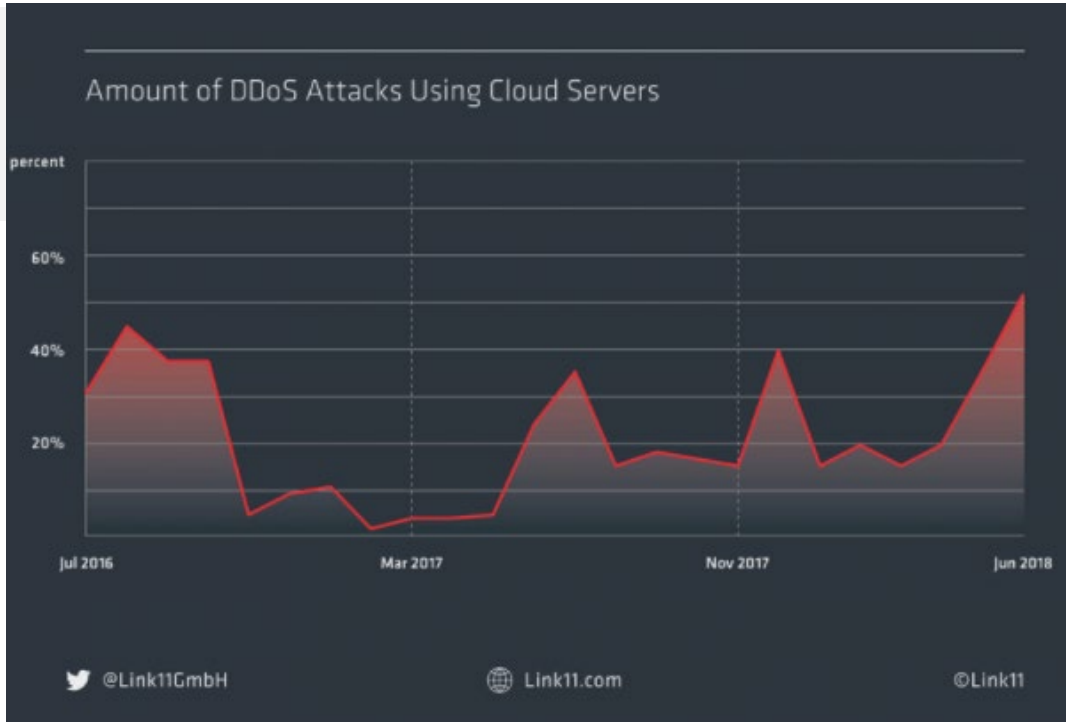
Earlier this month, researchers published an **analysis** of a sophisticated Linux Trojan (Backdoor.Linux.Mayday.f) capable of conducting DNS amplification DDoS attacks. After further investigation, Kaspersky identified two new variants of this threat, which the security firm detects as Backdoor.Linux.Mayday.g.

https://www.securityweek.com/cybercriminals-abuse-amazon-cloud-host-linux-ddos-trojans

# Motivation

Cloud services from leading providers including AWS, Microsoft Azure and Alibaba used in 25% of all DDoS attacks in Europe from July 2017 to June 2018

https://www.link11.com/en/blog/threat-landscape/public-cloud-services-increasingly-exploited-to-supercharge-ddos-attacks-new-link11-research/




threatpost Cloud Security / Malware / Vulnerabilities / Waterfall Security Spotlight

LokiBot Trojan Spotted Hitching a Ride Inside .PNG Files

**Hackers Abuse Google Cloud Platform to Attack D-Link Routers**

https://threatpost.com/hackers-abuse-google-cloud-platform-to-attack-d-link-routers/143492/

Home » Forensics » Azerbaijan » Azerbaijan and the fineproxy DIY DDOS service (Region40 / QualityNetwork)

AZERBAIJAN AND THE FINEPROXY DIY DDOS SERVICE (REGION40 / QUALITYNETWORK)

https://www.qurium.org/alerts/azerbaijan/azerbaijan-and-the-region40-ddos-service/

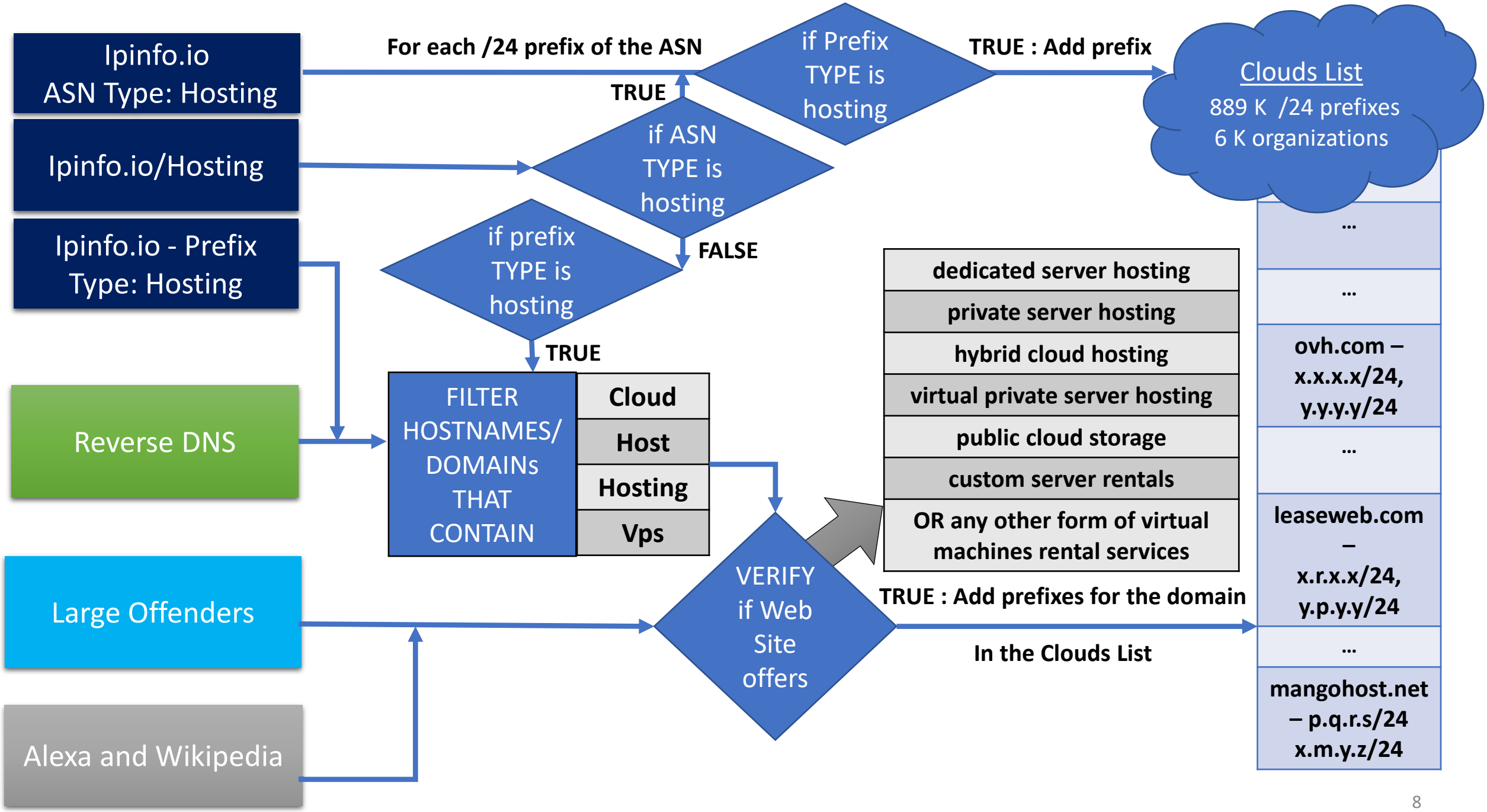
# Goal

- ***Quantify cloud misbehavior***
  (we analyze 13 diverse datasets, containing different variants of unwanted traffic)
- ***Identify clouds,*** that most often and most aggressively generate unwanted traffic

Note: We define as a “cloud”, an organization that offers servers for rent, and allows users to install custom software on these servers.

# Methodology

- Identifying Clouds
  - Identify all /24 prefixes in the Internet that offer server hosting services
  - (Because RIRs require a minimum assignment of /24 prefix for any end user organization)
  - We focus on /24 prefixes and not entire organizations
  - (Because large organizations may dedicate only a portion of their address space to cloud services)


Ipinfo.io
ASN Type: Hosting
Ipinfo.io/Hosting
Ipinfo.io - Prefix
Type: Hosting
Reverse DNS
Large Offenders
Alexa and Wikipedia
For each /24 prefix of the ASN
if Prefix TYPE is hosting
TRUE
TRUE : Add prefix
Clouds List
889 K /24 prefixes
6 K organizations
if ASN TYPE is hosting
FALSE
if prefix TYPE is hosting
TRUE
FILTER HOSTNAMES/ DOMAINs THAT CONTAIN
Cloud
Host
Hosting
Vps
VERIFY if Web Site offers
dedicated server hosting
private server hosting
hybrid cloud hosting
virtual private server hosting
public cloud storage
custom server rentals
OR any other form of virtual machines rental services
TRUE : Add prefixes for the domain
In the Clouds List
...
...
ovh.com – x.x.x.x/24, y.y.y.y/24
...
leaseweb.com – x.r.x.x/24, y.p.y.y/24
...
mangohost.net – p.q.r.s/24 x.m.y.z/24

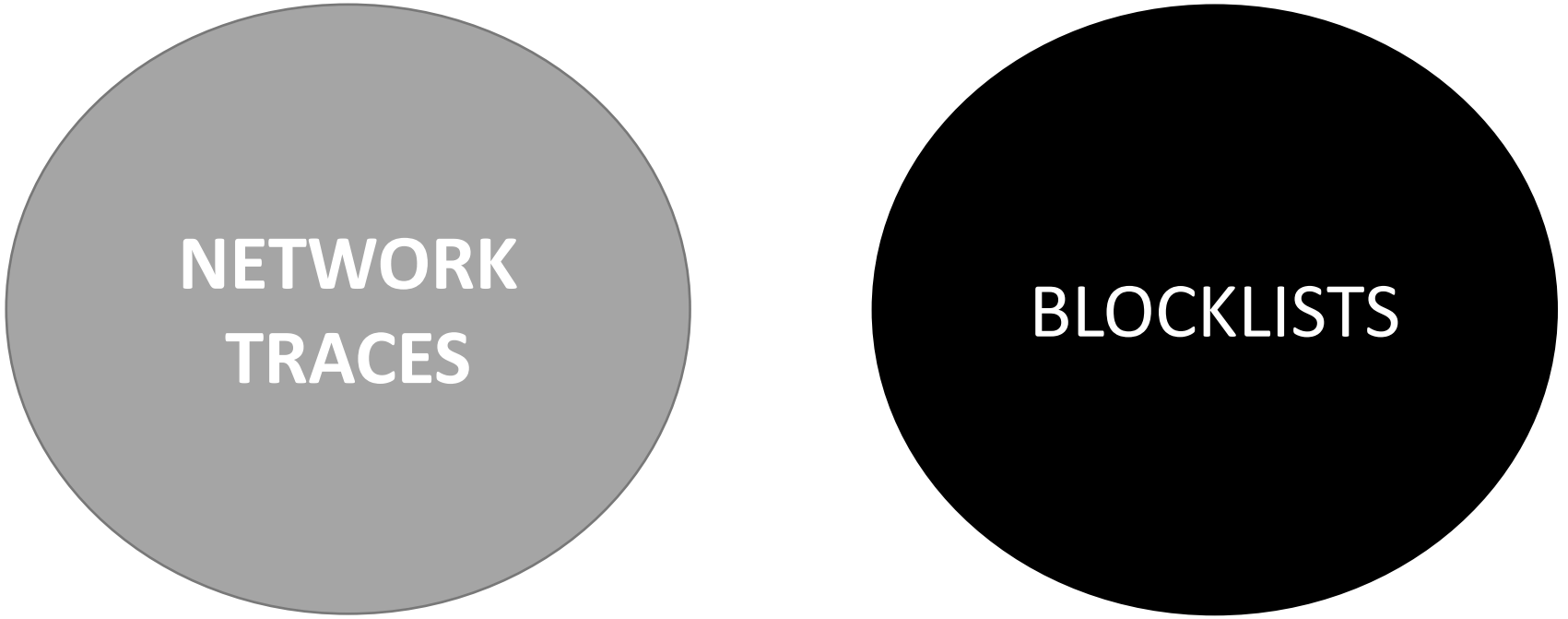
# Limitations

- No ground truth that we could use to evaluate accuracy of our cloud identification.
- Manual verification may in some cases be inaccurate.
- We may miss some clouds, because our candidate identification process misses them.

Our list of cloud prefixes is available as open source, and we hope that other researchers can help improve its accuracy.

*https://steel.isi.edu/Projects/Cloud_Misbehavior/*

# Datasets



# NETWORK TRACES

| Dataset | Source | Format | Description | Start Date | End Date | Size |
|---|---|---|---|---|---|---|
| CAIDA | [33] | PCAP files | Real-Time Network Telescope Data | 12-Mar-20 | 28-Apr-20 | 80 GBs hourly compressed files (a few billion packets) |
| Merit | [19] | PCAP files | Real-Time Network Telescope Data | 11-Mar-20 | 19-Mar-20 | 2 GBs hourly compressed files (fewer than 0.1 billion packets) |
| RONX | Anon. | Netflow files | Live ISP data (all 5-minutes long Netflow records) | 24-Feb-20 | 26-Apr-20 | 1.15 TB |

# BLOCKLISTS

| Dataset | Source | Format | Description | Start Date | End Date | Size |
|---|---|---|---|---|---|---|
| Scamalytics | [35] | IP addresses | Top 100 monthly IPs with the maximum fraud in online dating | 1-Mar-20 | 30-Apr-20 | 204 IP addresses |
| udger.com | [15] | IP addresses | Source IP addresses associated with different attacks | 4-May-20 | 5-May-20 | 3,030 IP addresses |
| Cybercrime Tracker | [10] | IP addresses | IPs that spread malwares | 1-Jan-19 | 12-May-20 | 3,471 IP addresses |
| Google Safebrowsing | [14] | URLs | Malicious URLs List from maltiverse.com | 8-May-20 | 16-May-20 | 7,886 URLs (that belong to non-bogon IPv4 address range) |
| COVID-19 Hostnames | [8] | Hostnames / URLs | Malicious hostnames/URLs that contain the word "COVID-19"/"corona" and are associated with generating different variants of unwanted traffic | 1-Jan-20 | 16-May-20 | 9,874 malicious hostnames |
| COVID-19 Phishing | [7] | URLs | Phishing URLs related to COVID-19 content | 13-Mar-20 | 16-May-20 | 239 phishing URLs |
| Openphish | [20] | ASNs and ASNs Domains | Top 10 ASNs associated with phishing | 15-May-20 | 22-May-20 | 54 snippets of top 10 ASNs |
| BGP Ranking | [3] | ASNs | Maintains top 100 malicious ASNs | 21-Aug-20 | 21-Aug-20 | 101 ASNs |
| BLAG (2018) | [47], [38] | IP addresses | Publicly available blacklisted IPs list collected daily using 157 popular blocklists | 1-Jan-18 | 31-Dec-18 | 0.5 billion IP addresses (14.5 million unique IPs) |
| BLAG (2019) | [47], [38] | IP addresses | Publicly available blacklisted IPs list collected daily using 157 popular blocklists | 1-Jan-19 | 31-Jan-20 | 5 billion IP addresses |
| F5 Labs: Attack Traffic | [18] | IP addresses / ASNs | Top 50 malicious ASNs and IPs ranked by number of attacks for more than 14 million attacks globally | 1-Aug-19 | 31-Oct-19 | 50 ASNs and 50 IPs |

# Misbehavior Metrics

$mal^{d}_{scans}t$ Number of scans a /24 prefix sends during the time t in the the network trace dataset d

$mal^{d}_{bl}$ Number of times a /24 prefix appears in the the blocklist dataset d

$$malorg^{d}_{score} = \frac{prefs_d \cdot r_d}{prefs_{tot}} \quad ;$$

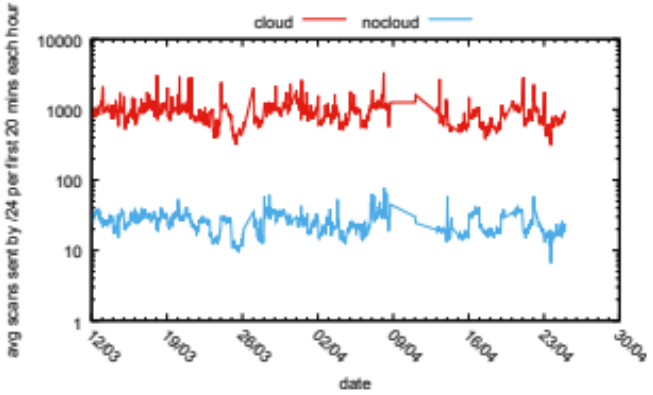
$prefs_{tot}$ is the number of /24 cloud prefixes owned by a given organization

$prefs_d$ is the number of /24 prefixes from an organization that appear in the dataset d

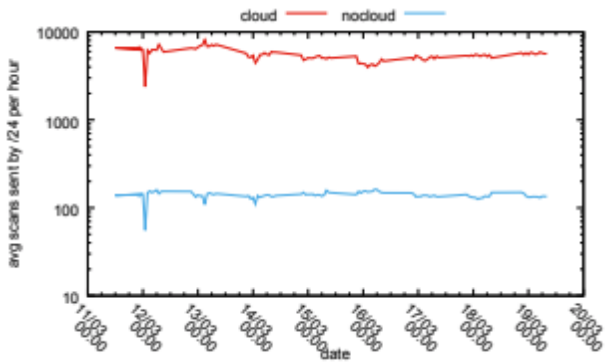
$r_d$ is the fraction of contribution of this organization to the total scans or entries in the dataset d

# Findings : NETWORK TRACES

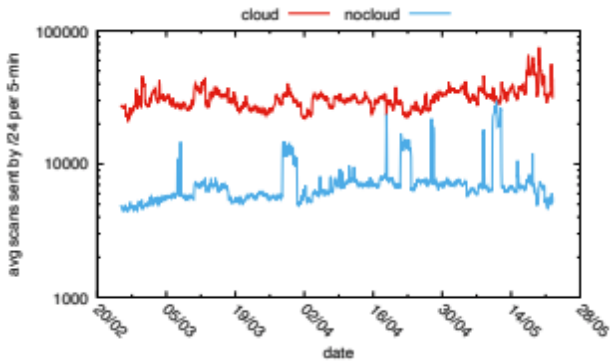
- Cloud prefixes are more aggressive than non-cloud prefixes across the network traces datasets.
- **The average $mal_{scans}$ for a cloud prefix is 20 - 100 times higher than the average $mal_{scans}$ for a non-cloud prefix.**




CAIDA



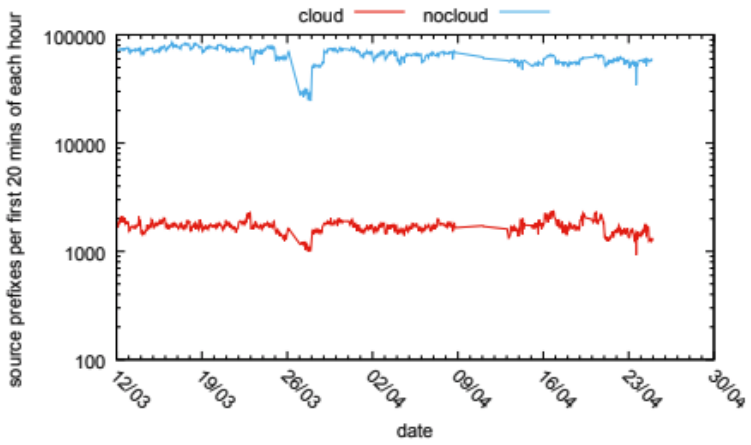

MERIT



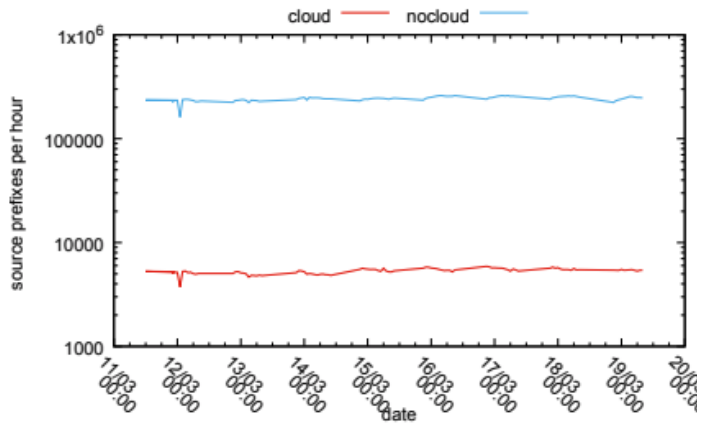

RNOX

# Findings : NETWORK TRACES

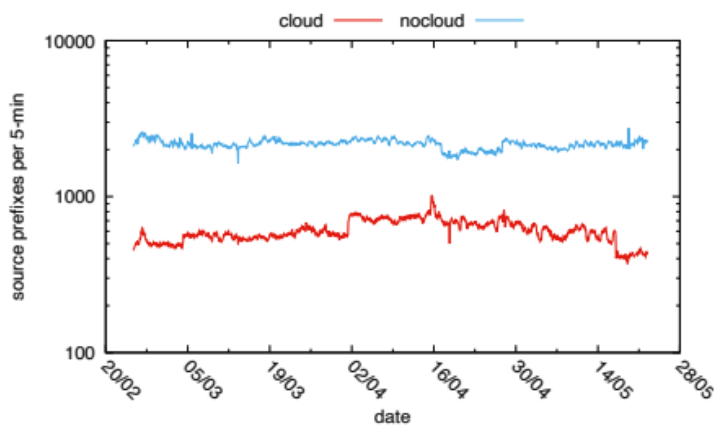
- **The number of /24 prefixes is 30 – 60 times higher for non-clouds than for clouds.**




CAIDA



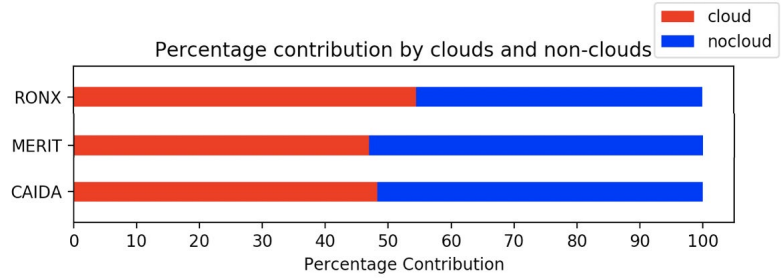

MERIT

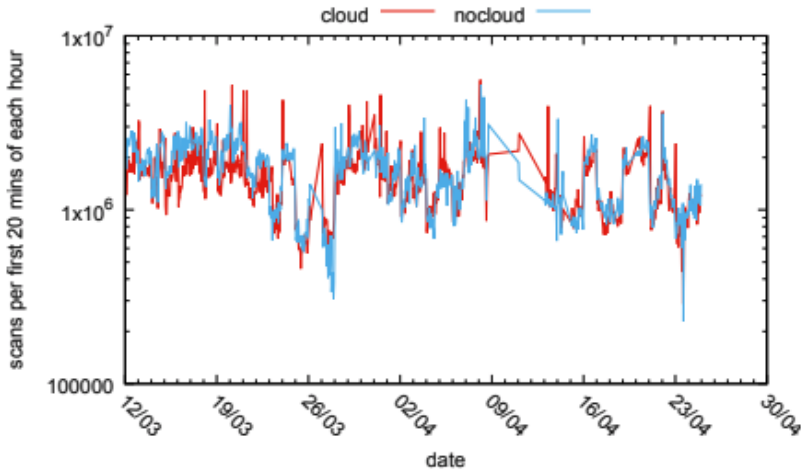



RNOX

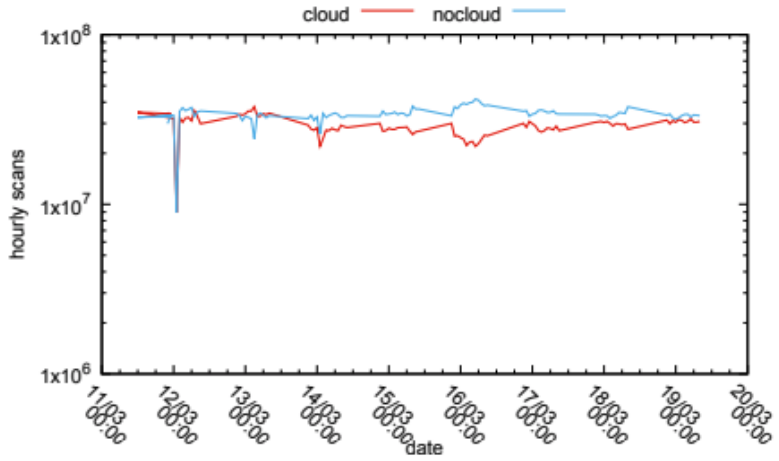
# Findings : NETWORK TRACES

- **Cloud** prefixes make only **2-3% of prefixes** in the network trace datasets, while **Non-Cloud** prefixes make **97–98% of prefixes**.
- **Total $mal_{scans}$ from clouds is similar to total $mal_{scans}$ from non-clouds**

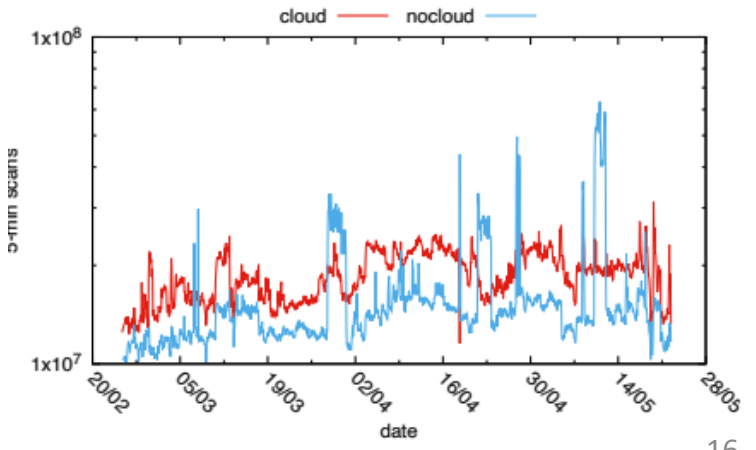



CAIDA




MERIT




RNOX

# Findings : NETWORK TRACES

- **Digital pickpocketers rely on clouds for possible monetary gains.**
- Port 8545 contributes to 1.7%, 1.3% and 3.7% of cloud scans in CAIDA, Merit and RONX datasets, respectively
- Conversely it only contributes to 0.3%, 0.4% and 1% of non-cloud scans respectively.
- **Around 25 clouds and 200 non-clouds, generate 90% of the malicious traffic.**

# Findings : NETWORK TRACES

| CLOUDS | | | | NON-CLOUDS | | | |
|---|---|---|---|---|---|---|---|
| Organizations | Average Ranking | # /24 Prefixes Seen | # /24 Prefixes Owned | Organizations | Average Ranking | # /24 Prefixes Seen | # /24 Prefixes Owned |
| selectel.ru | 1.33 | 47.3 | 1077 | ssnet.bg | 1.00 | 1.0 | 1 |
| ipvolume.net | 2.00 | 9.7 | 476 | megavantage.cn | 3.33 | 1.0 | 16 |
| perhost.net | 2.67 | 1.7 | 2 | chinaunicom.cn | 4.00 | 99841.7 | 285544 |
| inter-host.net | 3.33 | 1.7 | 2 | chinatelecom.com.cn | 5.33 | 107922.0 | 417135 |
| novogara.com | 5.00 | 1.7 | 7 | censys.io | 6.33 | 2.0 | 2 |
| digitalocean.com | 5.33 | 1275.7 | 8702 | chinamobile.com | 6.67 | 23311.7 | 149026 |
| rm-injinering.ru | 8.33 | 3.7 | 7 | networkdedicated.com | 7.33 | 1.7 | 440 |
| ovh.net | 8.67 | 1305.3 | 13744 | dm-auto.eu | 9.33 | 1.0 | 1 |
| colocrossing.com | 9.00 | 147.7 | 3086 | vitox.eu | 9.33 | 10.7 | 17 |
| reliablesite.net | 9.67 | 17.3 | 244 | wenzhouglasses.com | 11.00 | 3.0 | 210 |

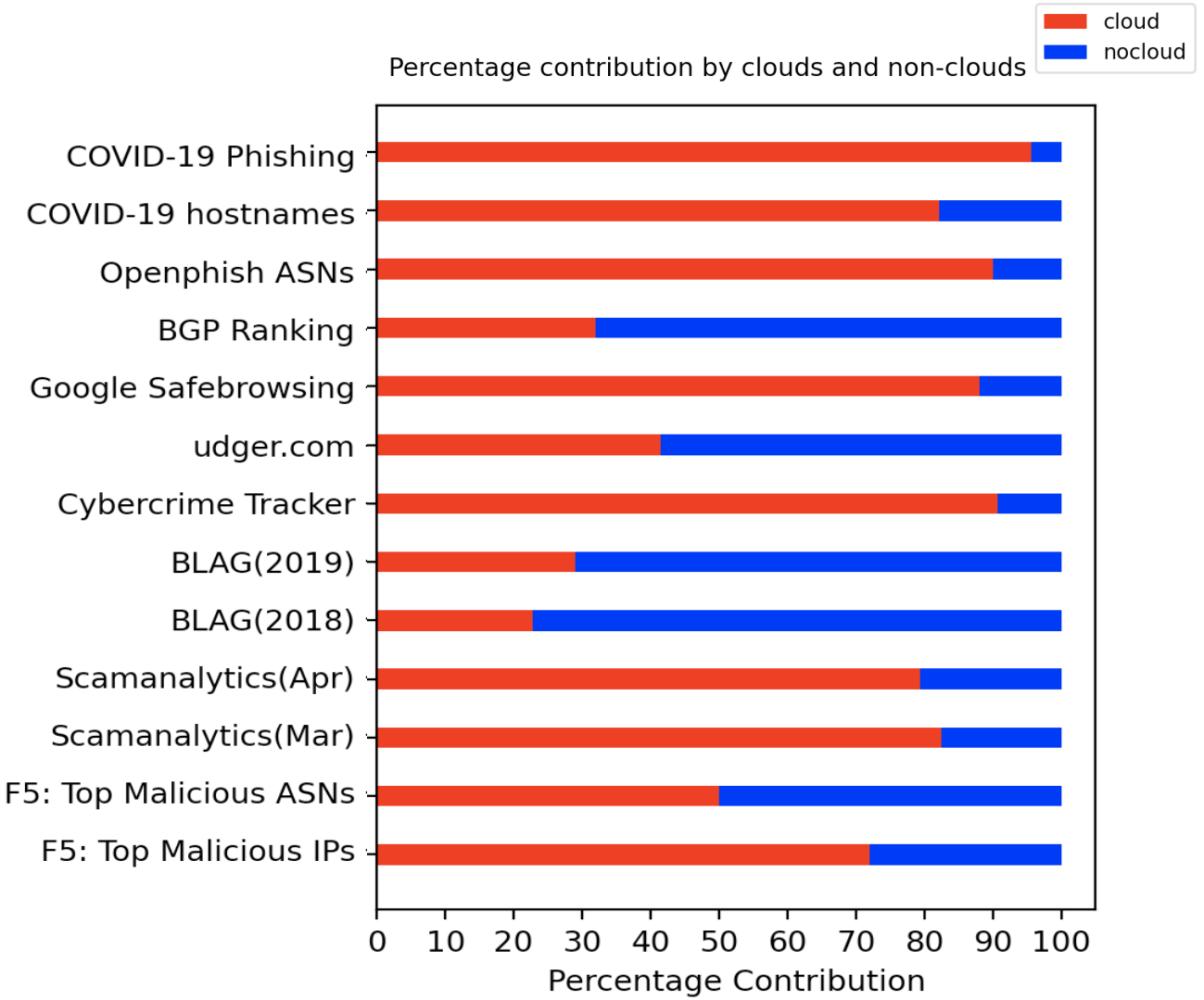
(a) Ordered by organization wise total $mal_{scans}$

| CLOUDS | | | | NON-CLOUDS | | | |
|---|---|---|---|---|---|---|---|
| Organization | $malorg_{score}$ | # /24 Prefixes Seen | # /24 Prefixes Owned | Organization | $malorg_{score}$ | # /24 Prefixes Seen | # /24 Prefixes Owned |
| perhost.net | 5.19 | 1.7 | 2 | ssnet.bg | 9.18 | 1.0 | 1 |
| inter-host.net | 4.24 | 1.7 | 2 | chinaunicom.cn | 0.97 | 99841.7 | 285544 |
| rm-injinering.ru | 0.99 | 3.7 | 7 | chinatelecom.com.cn | 0.96 | 107922.3 | 417135 |
| novogara.com | 0.92 | 1.7 | 7 | dm-auto.eu | 0.68 | 1.0 | 1 |
| digitalocean.com | 0.57 | 1275.7 | 8702 | vitox.eu | 0.44 | 10.7 | 17 |
| ipvolume.net | 0.51 | 20.3 | 476 | megavantage.cn | 0.23 | 1.0 | 16 |
| selectel.ru | 0.31 | 47.3 | 1077 | chinamobile.com | 0.23 | 23311.7 | 149026 |
| ovh.net | 0.22 | 1305.3 | 13744 | vnpt.vn | 0.20 | 14901.7 | 28918 |
| linode.com | 0.16 | 395.7 | 1914 | viettel.com.vn | 0.18 | 9782.3 | 24012 |
| nforce.com | 0.13 | 60.3 | 389 | telkom.co.id | 0.16 | 8897.3 | 16185 |

(b) Ordered by organization wise $malorg_{score}$

***$rank_{avg}$ of top clouds and non-clouds***

# Findings : BLOCKLISTS



# Findings : BLOCKLISTS

- **Clouds contribute anywhere from 22% to 96%, in spite of being only 5.4% of the routable address space.**
- On the average, the total $mal_{bl}$ per cloud is 1.82 times higher than the total $mal_{bl}$ per non-cloud.
- Clouds are almost twice as likely to engage in misbehavior that lands them on a blocklist than non-clouds
- Clouds play a vital role in spreading malware and supporting phishing URLs
- Clouds play vital role in web-application attacks.
- 61% of the global attacks from the F5 Labs Research dataset originated from clouds, i.e., 8.6 million out of the 14 million attacks.

# Findings : BLOCKLISTS

**CLOUDS**

| Organizations | Average Ranking | # /24 Prefixes Seen | # /24 Prefixes Owned |
|---|---|---|---|
| ovh.net | 5.0 | 1809.2 | 13744.0 |
| digitalocean.com | 5.3 | 1408.3 | 8702.0 |
| amazon.com | 7.7 | 53.5 | 190643.0 |
| godaddy.com | 8.0 | 364.9 | 3834.0 |
| cloudflare.com | 8.3 | 5.7 | 6098.0 |
| hetzner.de | 11.3 | 777.0 | 7454.0 |
| unifiedlayer.com | 12.7 | 396.4 | 6101.0 |
| quadranet.com | 13.3 | 219.2 | 1950.0 |
| google.com | 18.7 | 82.1 | 52084.0 |
| namecheap.com | 19.0 | 17.2 | 101.0 |

**NON-CLOUDS**

| Organizations | Average Ranking | # /24 Prefixes Seen | # /24 Prefixes Owned |
|---|---|---|---|
| chinatelecom.com.cn | 5.0 | 10257.7 | 417135 |
| vnpt.vn | 6.0 | 3333.4 | 28918 |
| rostelecom.ru | 6.3 | 4036.5 | 42457 |
| airtel.com | 6.8 | 2683.3 | 35105 |
| ptcl.net.pk | 8.0 | 1541.4 | 14509 |
| chinaunicom.cn | 8.0 | 6548.5 | 285544 |
| ertelecom.ru | 8.7 | 818.3 | 9858 |
| viettel.com.vn | 12.5 | 2517.2 | 24012 |
| chinamobile.com | 13.7 | 1725.6 | 149026 |
| fpt.com.vn | 17.6 | 687.5 | 5462 |

**(a) Ordered by organization wise total $mal_{bl}$**

**CLOUDS**

| Organization | $malorg_{score}$ | # /24 Prefixes Seen | # /24 Prefixes Owned |
|---|---|---|---|
| digitalocean.com | 0.27 | 1564.8 | 8702 |
| lanset.com | 0.22 | 34.1 | 92 |
| ovh.net | 0.20 | 1809.2 | 13744 |
| namecheap.com | 0.19 | 33.9 | 101 |
| unifiedlayer.com | 0.18 | 580.5 | 6101 |
| hostmaze.com | 0.14 | 3.1 | 7 |
| godaddy.com | 0.13 | 538.6 | 3834 |
| colocrossing.com | 0.10 | 147.7 | 3086 |
| quadranet.com | 0.10 | 278.3 | 1950 |
| datashack.net | 0.10 | 60.3 | 214 |

**NON-CLOUDS**

| Organization | $malorg_{score}$ | # /24 Prefixes Seen | # /24 Prefixes Owned |
|---|---|---|---|
| airtel.com | 0.22 | 13416.5 | 35105 |
| ptcl.net.pk | 0.23 | 4953.8 | 14509 |
| rostelecom.ru | 0.21 | 13242.2 | 42457 |
| vnpt.vn | 0.21 | 6943.8 | 28918 |
| chinatelecom.com.cn | 0.21 | 33649.3 | 417135 |
| ertelecom.ru | 0.19 | 2629.5 | 9858 |
| viettel.com.vn | 0.16 | 7191.4 | 24012 |
| tot.co.th | 0.15 | 2653.6 | 5127 |
| fregat.net | 0.12 | 407.5 | 450 |
| ufanet.ru | 0.11 | 628.5 | 966 |

**(b) Ordered by organization wise $malorg_{score}$**

***$rank_{avg}$ of top clouds and non-clouds***

# Conclusion

- Cloud machines can be misused either because of:
  - The negligence in adopting the security practices by their users
  - Or because they explicitly permit malicious traffic generation
- In all our 13 datasets clouds are much more aggressive than non-clouds.
- Clouds generate 20 – 100 times more scans per /24 prefix
- Clouds are twice more likely to appear on a blocklist
- Both clouds and non-clouds misbehave in heavy-tailed manner.
- Top 25 clouds account for 90% of the unwanted scans from clouds
- Top 10 clouds contribute more than 20% of cloud addresses that appear on blocklists.
- Thus, if efforts are focused on securing these clouds, Internet attacks can be greatly reduced.

*Contact us:*

*rajattan@usc.edu*

*sunshine@isi.edu*

*charnset@usc.edu*


We are grateful to **CAIDA**, **Merit Network (Dr. Michalis Kallitsis)**, **RONX** and **IPinfo.io** for allowing us to use their data in our research.

*https://steel.isi.edu/Projects/Cloud_Misbehavior/*