# Update on Domain Abuse Activity Reporting tool.

**John Crain**
**15 March 2018**

ICANN 61
COMMUNITY FORUM

**SAN JUAN**
10–15 March 2018

# The Domain Abuse Activity Reporting system

## What is the Domain Abuse Activity Reporting system?

- A system for reporting on domain name registration and abuse data across TLD registries and registrars

## How does DAAR differ from other reporting systems?

- Studies all gTLD registries and registrars for which we can collect zone and registration data
- Employs a large set of reputation feeds (e.g., blocklists)
- Accommodates historical studies
- Studies multiple threats: phishing, botnet, malware, spam
- Takes a scientific approach: transparent, reproducible

# DAAR & the Open Data Initiative

- Goal of Open Data Initiative is to facilitate access to data that ICANN organization or community creates or curates
- DAAR system uses data from public, open, and commercial sources
  - DNS zone data
  - WHOIS data
  - Open source or commercial reputation blocklist (RBL) data
    - Certain data feeds require a license or subscription
- In cases where licensing permits, DAAR data or reports will be published and included in the Open Data Initiative

# Project Goals

- DAAR data can be used to
  - Report on threat activity at TLD or registrar level
  - Study histories of security threats or domain registration activity
  - Help operators understand or consider how to manage their reputations, their anti-abuse programs, or terms of service
  - Study malicious registration behaviors
  - Assist operational security communities

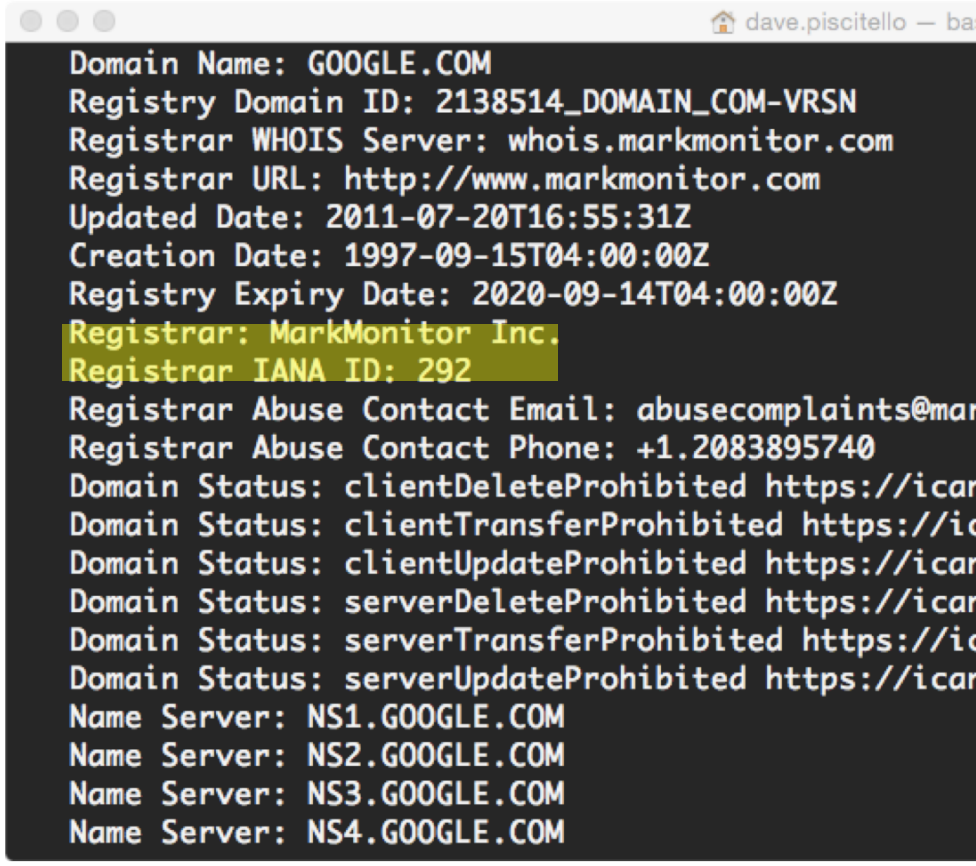
*The purpose of DAAR is to provide data to support community, academic, or sponsored research and analysis for informed policy consideration*

# DAAR Uses TLD Zone Data

- Collects all gTLD zones for gTLD registry analytics
- DAAR uses publicly available methods to collect zone data
  - Centralized Zone Data Service, zone transfer)
- DAAR only uses domain names that appear(ed) in zones
- Currently, system collects zones from ~1240 gTLDs
  - Approximately 195 million domains

# DAAR Uses Whois

- DAAR uses published registration data (Whois)
  - Uses only registration data necessary to associate resolving domain names in zone files with sponsoring registrars
- Reliable, accurate registrar reporting depends on Whois
  - Collecting registration records for millions of domains is a big challenge



# DAAR Uses Many Threat Data Sets

- DAAR counts “unique” abuse domains
  - A domain that appears on *any* RBL reporting to DAAR is included in the counts *once*
- DAAR uses multiple domain or URL abuse data sets to
  - Generate daily counts of domains associated with phishing, malware hosting, botnet C&C, and spam
  - Calculate daily total and cumulative abuse domains
  - Calculate newly added abuse domains (a monthly count), and cumulative abuse domains (365 day count)
  - Create histograms, charts, days in the life views

*DAAR reflects how entities external to ICANN community see the domain ecosystem*

# Current Reputation Data Sets

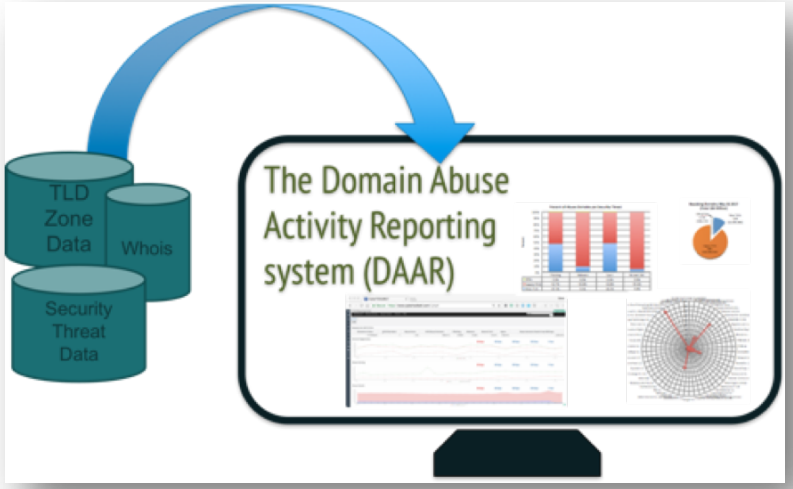
- SURBL lists (domains only)
- Spamhaus Domain Block List
- Anti-Phishing Working Group
- Malware Patrol (Composite list)
- Phishtank
- Ransomware Tracker
- Feodotracker

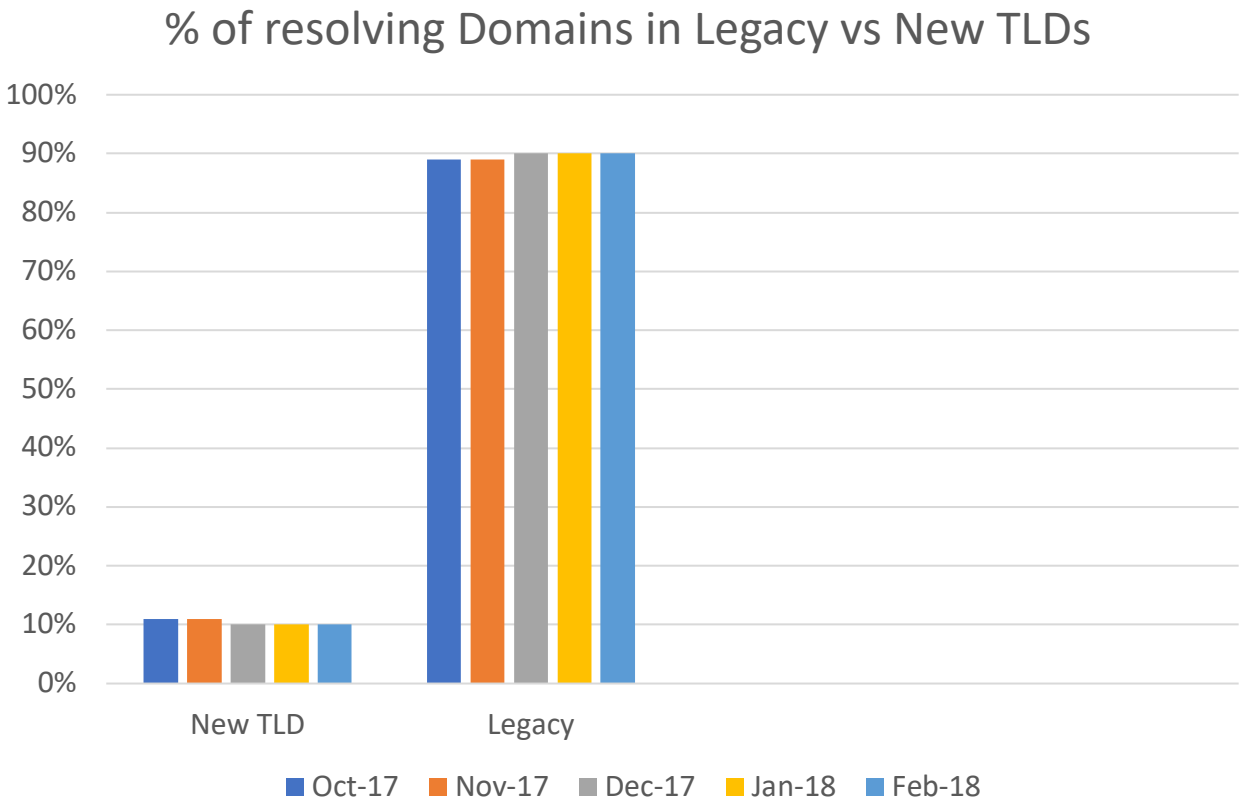
SpamAssassin: malware URLs list
Carbon Black Malicious Domains
Postfix MTA
Squid Web proxy blocklist
Symantec Email Security for SMTP
Symantec Web Security
Firekeeper
DansGuardian
ClamAV Virus blocklist
Mozilla Firefox Adblock
Smoothwall
MailWasher

# Does DAAR Identify All Abuse?

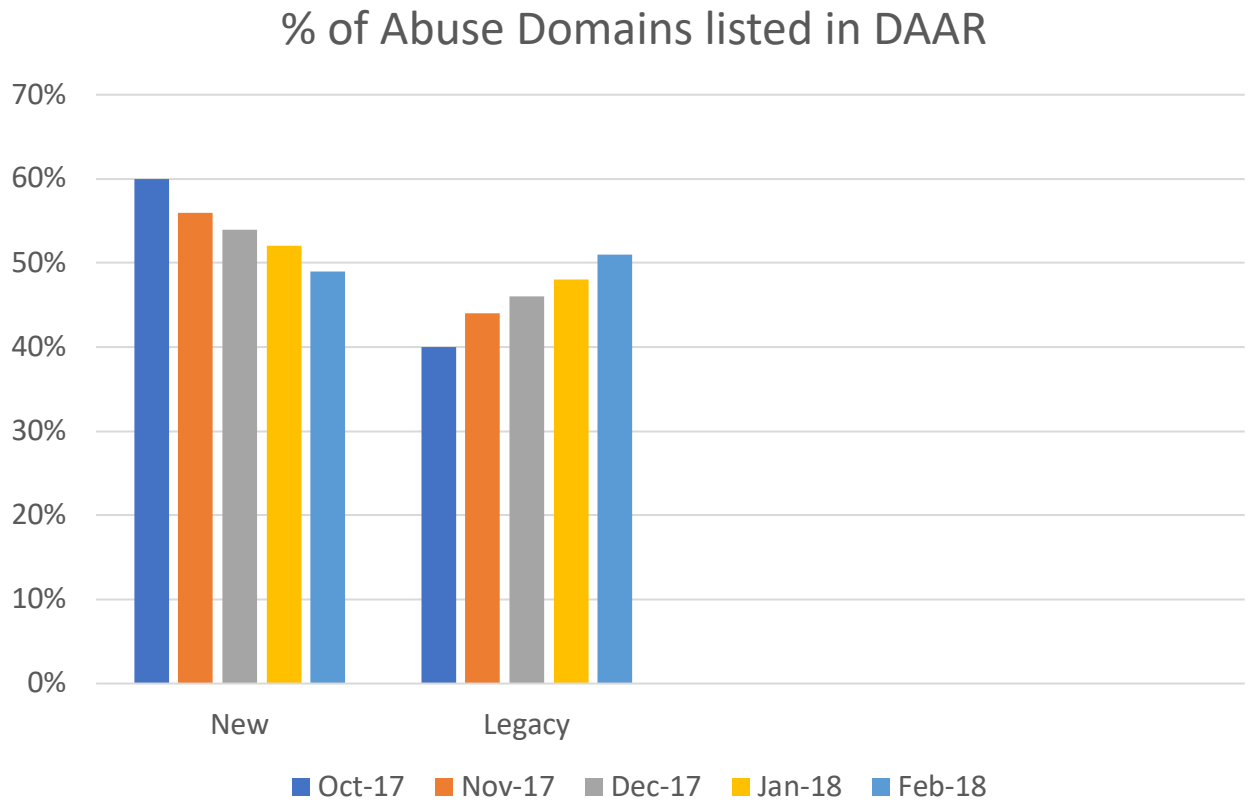
- No reputation provider can see all the abuse
  - Each is catching only some (what they see)
- Providers look for different types of abuse, use different methods or infrastructures
- Some lists are big and some are small.
  - The smaller the list, the less percent of overlap it might have with a larger list

# Visualizing DAAR Data

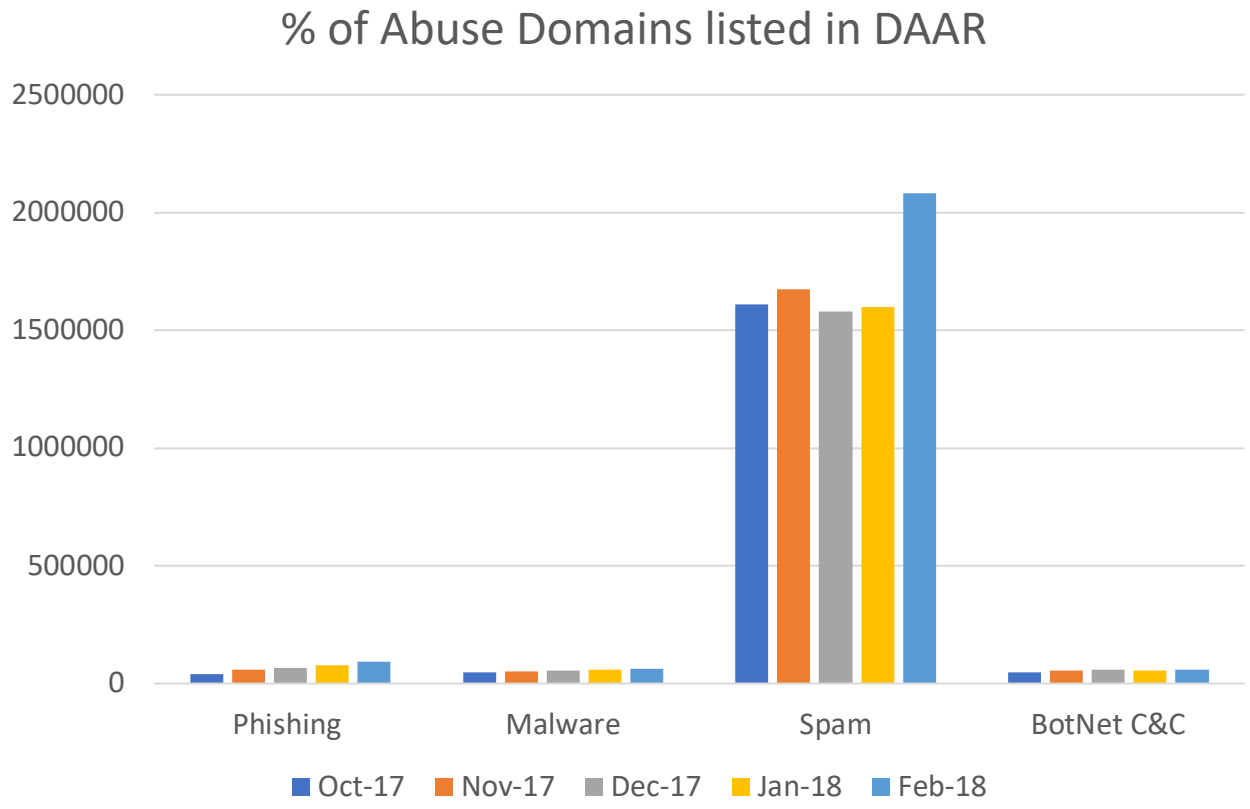



% of resolving Domains in Legacy vs New TLDs
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
New TLD
Legacy
Oct-17
Nov-17
Dec-17
Jan-18
Feb-18

# Data Set: All gTLDs having at least 1 reported abuse domain



# Data Set: All gTLDs having at least 1 reported abuse domain

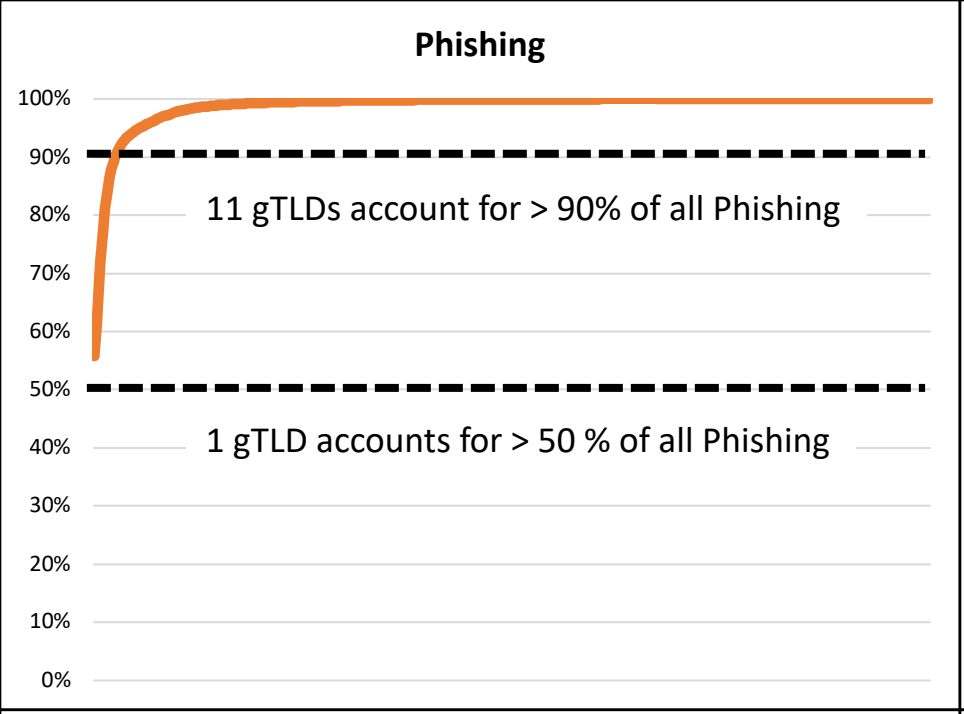


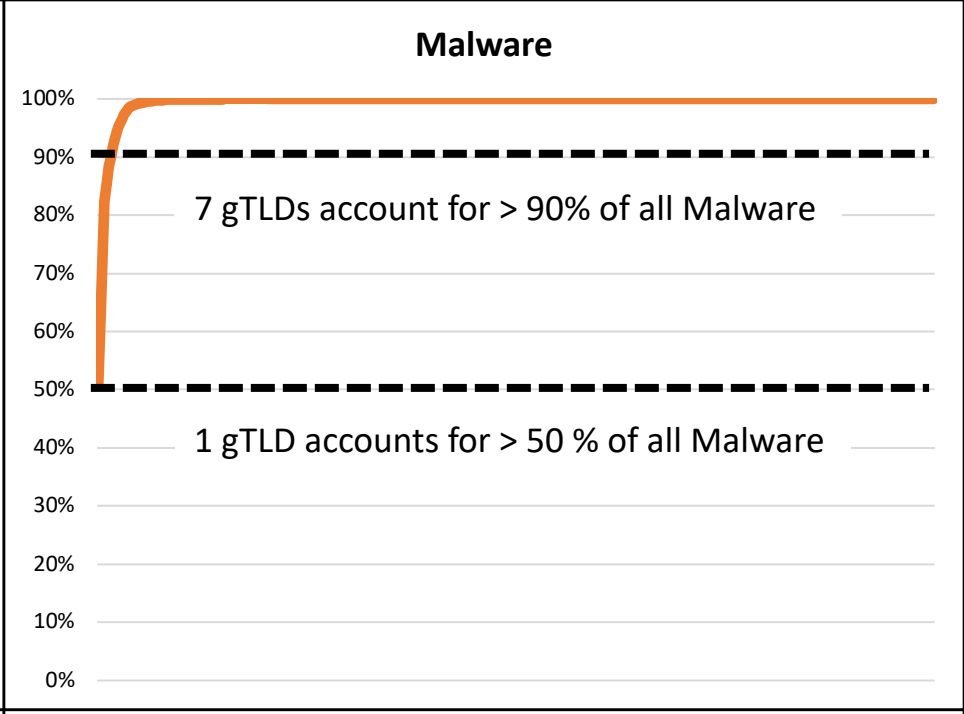
# M2.*: Number of Abused Domain per 10,000 Registrations

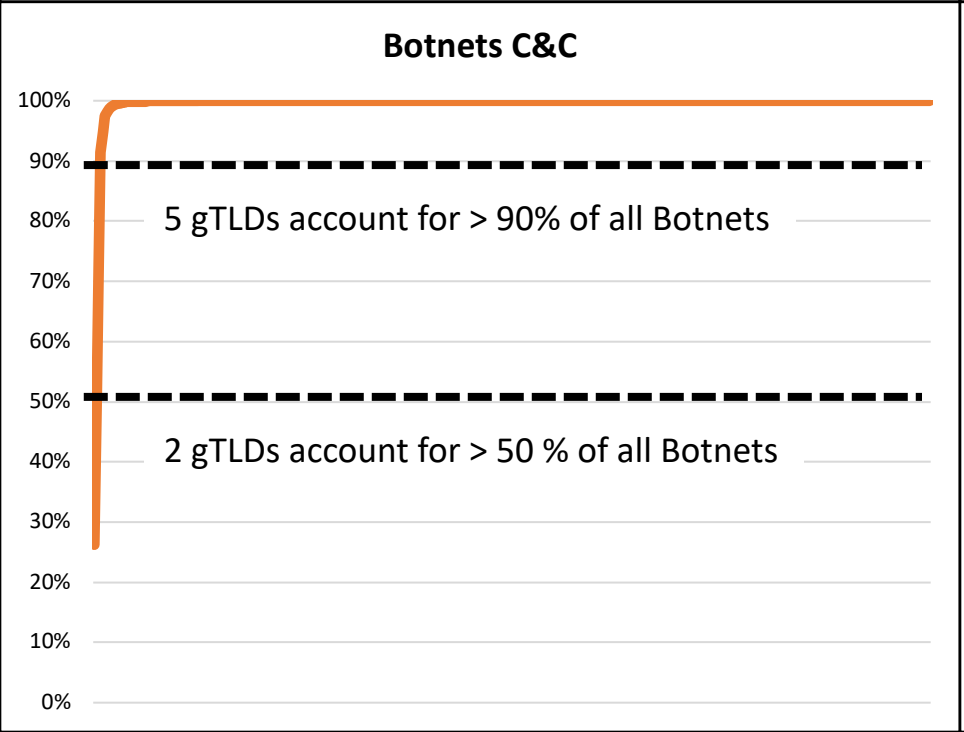
Data from 01/31/2018

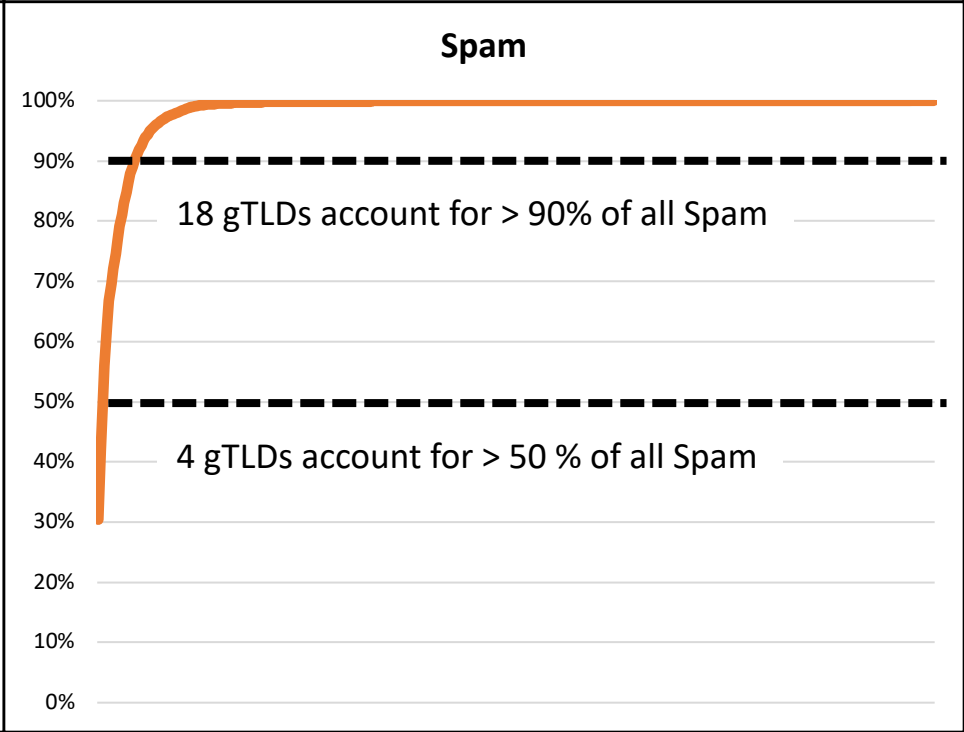
| M2 metric name | Global Average |
|---|---|
| M2.1 = number of Phishing Domains per 10000 registered domain names | 4.28 |
| M2.2 = number of Malware Domains per 10,000 registered domain names | 3.28 |
| M2.3 = number of Botnet C&C Domains per 10,000 registered domain names | 2.89 |
| M2.4 = number of Spam Domains per 10,000 registered domain names | 86.73 |

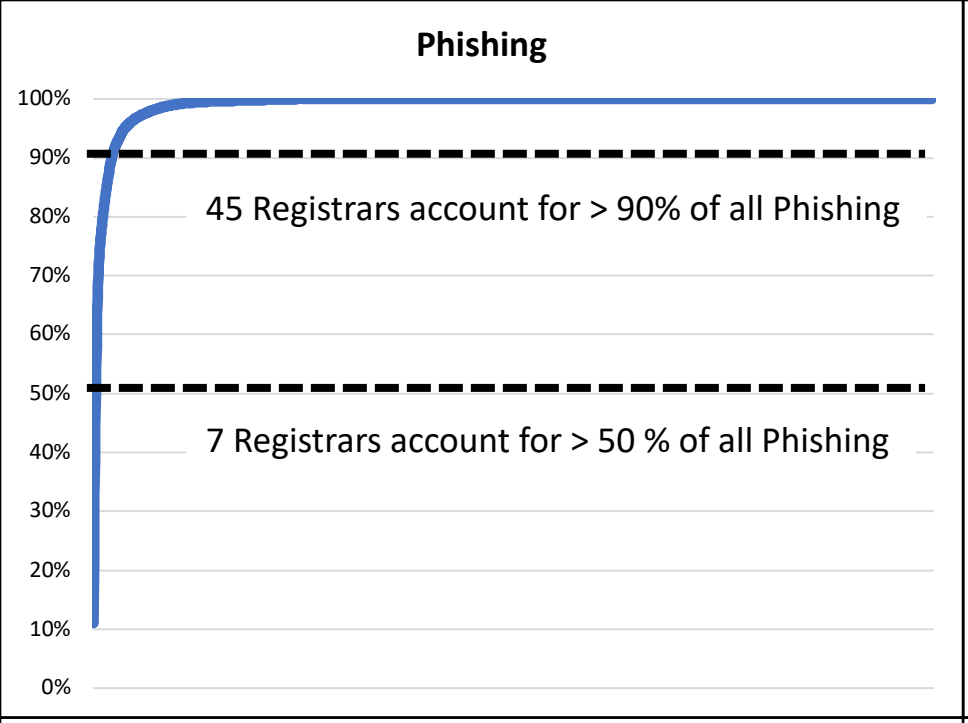
Total number of gTLDs: 1143, Total number of registrars: 1952

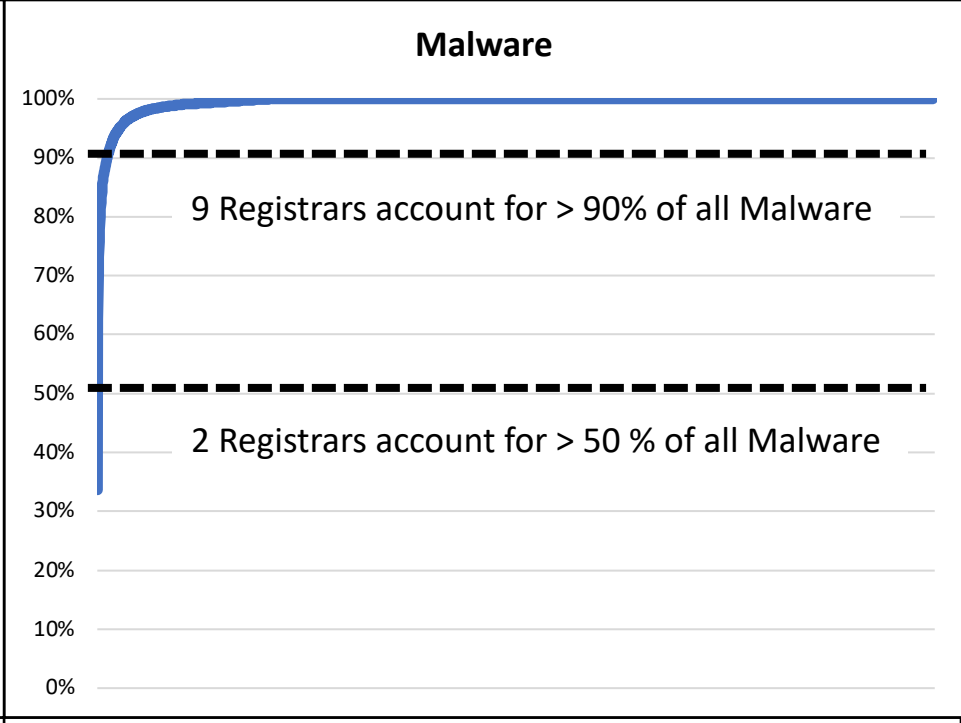

Phishing
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
11 gTLDs account for > 90% of all Phishing
1 gTLD accounts for > 50 % of all Phishing

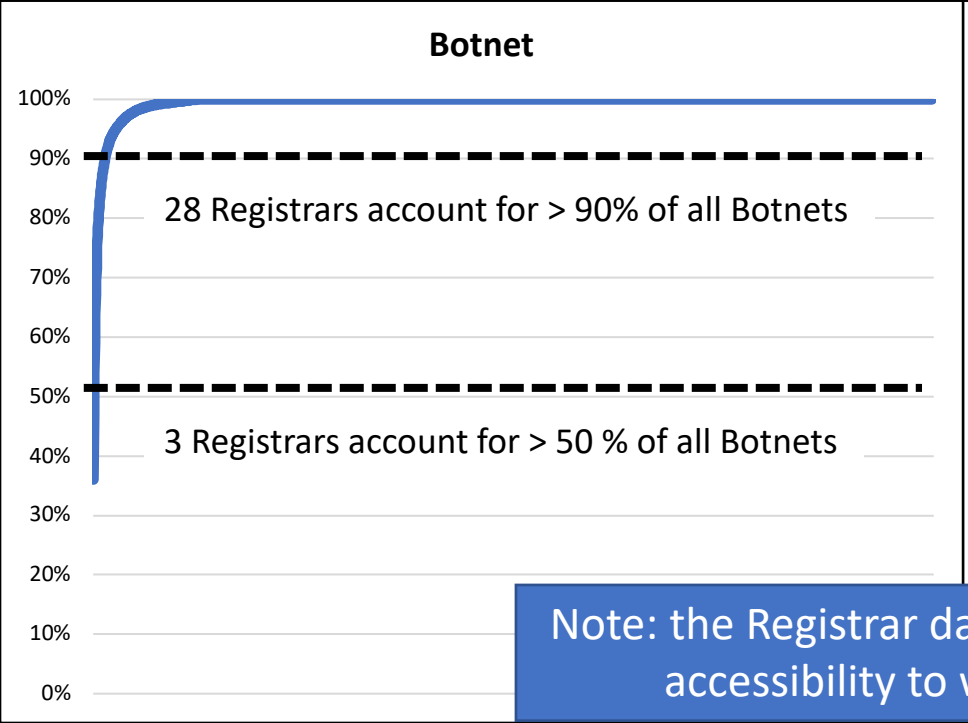

Malware
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
7 gTLDs account for > 90% of all Malware
1 gTLD accounts for > 50 % of all Malware

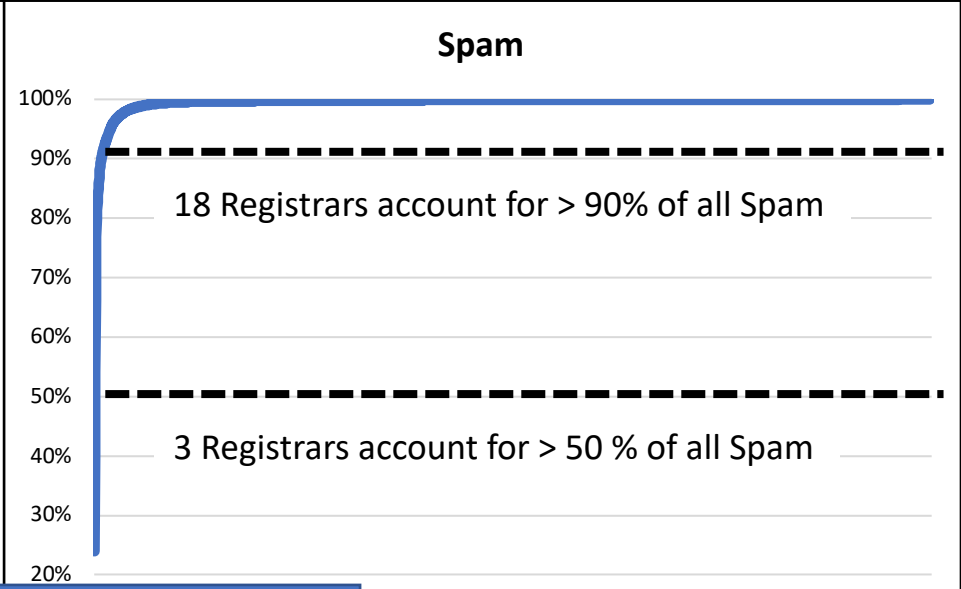

Botnets C&C
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
5 gTLDs account for > 90% of all Botnets
2 gTLDs account for > 50 % of all Botnets


Spam
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
18 gTLDs account for > 90% of all Spam
4 gTLDs account for > 50 % of all Spam




Note: the Registrar data is effected by accessibility to whois data

# M2.*: Concentration of Abuse

| Abuse | gTLD50 | Registrar50 | gTLD90 | Registrar90 |
|---|---|---|---|---|
| **Phishing** | 1 | 7 | 11 | 45 |
| **Malware** | 1 | 2 | 7 | 9 |
| **Botnet** | 2 | 3 | 5 | 28 |
| **Spam** | 4 | 3 | 18 | 18 |

Table shows the number of TLDs/Registrars to account for > 50%/90% of all abuse of the specified type.

Total number of gTLDs: 1143, Total number of registrars: 1952*

(*) We removed two parking registrars from those statistics

# M2.*: Number of Abused Domain per 10,000 Registrations

Data from 01/31/2018

| M2 metric name | Score |
|---|---|
| M2.1 = Average number of Phishing Domains per 10000 registered domain names | 4.28 |
| Highest | **190** |
| # of TLDs >10*M2.1 | 10 |
| | |

Total number of gTLDs: 1143, Total number of registrars: 1952

# M2.*: Number of Abused Domain per 10,000 Registrations

Data from 01/31/2018

| M2 metric name | Score |
|---|---|
| M2.2 = Average number of Malware Domains per 10,000 registered domain names | 3.28 |
| Highest | **417** |
| # of TLDs >10*M2.2 | **1** |
| | |

Total number of gTLDs: 1143, Total number of registrars: 1952

# M2.*: Number of Abused Domain per 10,000 Registrations

Data from 01/31/2018

| M2 metric name | Score |
|---|---|
| M2.3 = Average number of Botnet C&C Domains per 10,000 registered domain names | 2.89 |
| Highest | **71.46** |
| # of TLDs >10*M2.3 | **1** |
| | |

Total number of gTLDs: 1143, Total number of registrars: 1952

# M2.*: Number of Abused Domain per 10,000 Registrations

Data from 01/31/2018

| M2 metric name | Score |
|---|---|
| M2.4 = Average number of Spam Domains per 10,000 registered domain names | 86.73 |
| Highest | **4112** |
| # of TLDs >10*M2.4 | **15** |
| | |

Total number of gTLDs: 1143, Total number of registrars: 1952

- The top scorers in the four metrics are four different registries.

- We have contracted independent reviewers and expect to post the results in the coming weeks.

# Engage with ICANN

## Thank You and Questions

Visit us at **icann.org**
Email: email


@icann


facebook.com/icannorg


youtube.com/icannnews


flickr.com/icann


linkedin/company/icann


slideshare/icannpresentations


soundcloud/icann