FROM THREAT INTELLIGENCE TO RAP SHEETS (IN JUST A FEW STEPS)

Dragos Comaneci, Principal Software Engineer
Adrian Hada, Security Research Engineer

ABOUT

- Software Engineer, Ixia
- Loves to build, debug & understand distributed systems

- Security Research Engineer, Ixia
- Spends most of his time around malware, botnets and the like

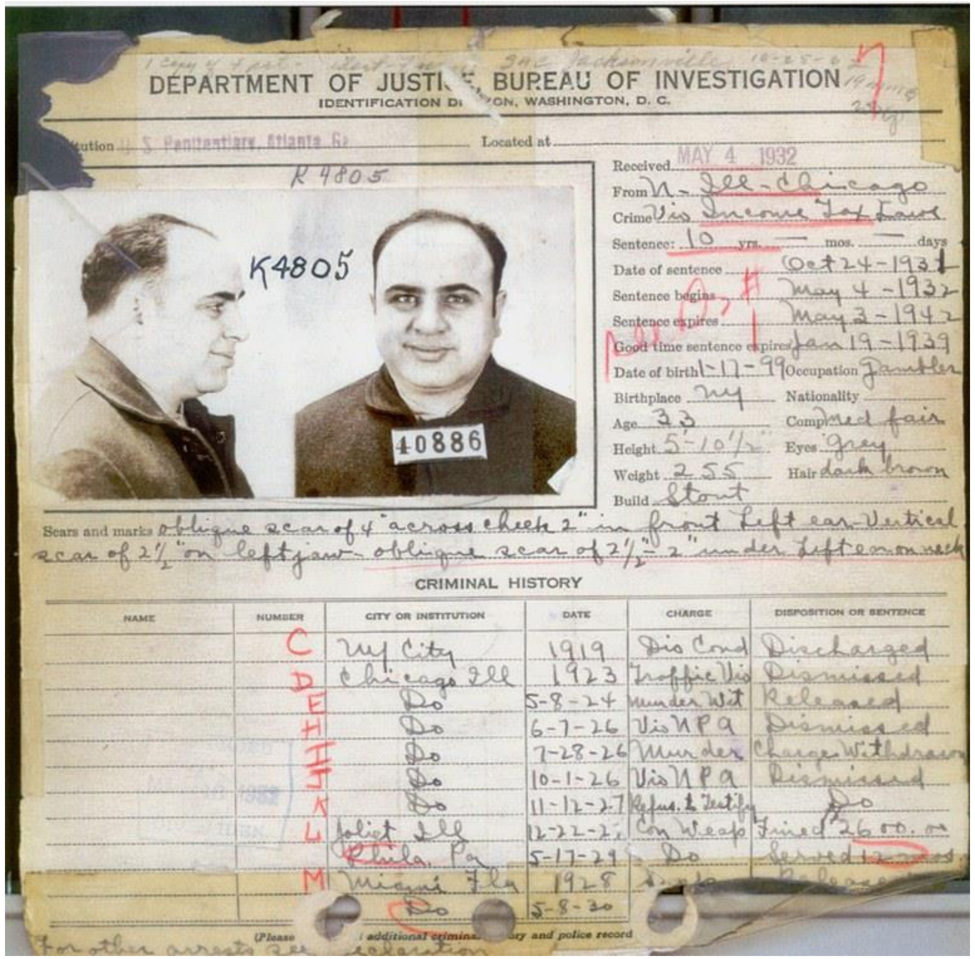
AGENDA

- What’s a Rap Sheet?
- Threat Identification
- Storing and Interpreting Data
- Stats

DISCLAIMER

- Not product placement
- Our perspective on developing a threat intelligence system

WHAT'S A RAP SHEET?



1 Copy of 4 pst - [illegible] Jacksonville 10-25-62

DEPARTMENT OF JUSTI[illegible], BUREAU OF INVESTIGATION
IDENTIFICATION DI[illegible]ON, WASHINGTON, D. C.

[illegible]tution U S Penitentiary, Atlanta Ga   Located at

K 4805

K4805

40886

Received MAY 4 1932
From N. Ill - Chicago
Crime Vio Income Tax Laws
Sentence: 10 yrs. — mos. — days
Date of sentence Oct 24 - 1931
Sentence begins May 4 - 1932
Sentence expires May 3 - 1942
Good time sentence expires Jan 19 - 1939
Date of birth 1-17-99   Occupation Gambler
Birthplace NY   Nationality
Age 33   Comp Med fair
Height 5'-10½"   Eyes Grey
Weight 255   Hair dark brown
Build Stout

Scars and marks oblique scar of 4" across cheek 2" in front left ear - Vertical scar of 2½" on left jaw - oblique scar of 2½" - 2" under left ear on neck

CRIMINAL HISTORY

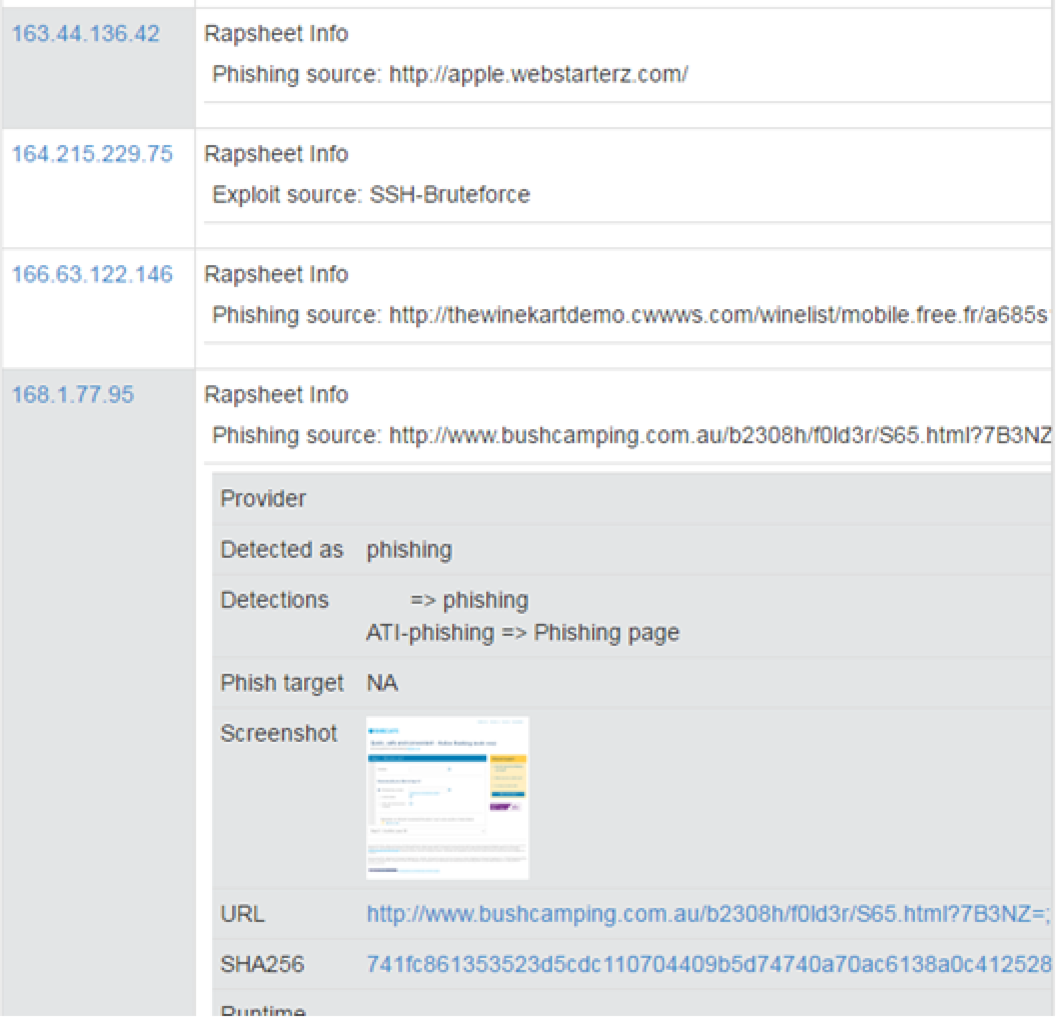
NAME	NUMBER	CITY OR INSTITUTION	DATE	CHARGE	DISPOSITION OR SENTENCE
	C	Ny City	1919	Dis Cond	Discharged
	D	Chicago Ill	1923	Traffic Vio	Dismissed
	E	Do	5-8-24	Murder Wit	Released
	H	Do	6-7-26	Vio N P A	Dismissed
	I	Do	7-28-26	Murder	Charge Withdrawn
	J	Do	10-1-26	Vio N P A	Dismissed
	K	Do	11-12-27	Refus. to Testify	Do
	L	Joliet Ill	12-22-2[illegible]	Con Weap	Fined 2600.00
		Phila. Pa	5-17-29	Do	Served 12 mos.
	M	Miami Fla	1928	[illegible]	Released
		Do	5-8-30		

(Please [illegible] additional crimina[illegible]ory and police record

For other arrests see [illegible]

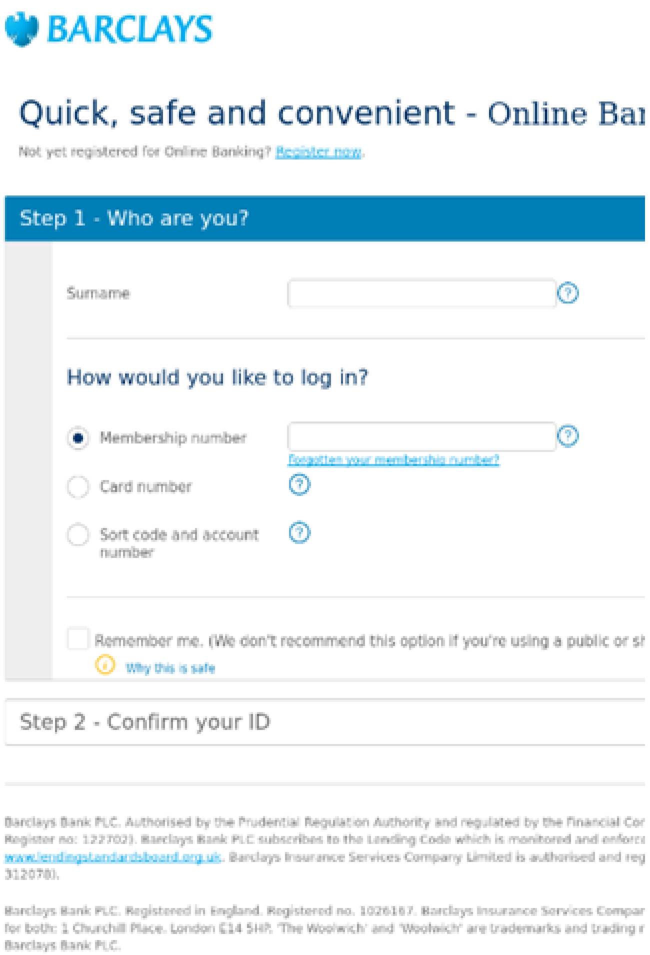
- "a list kept by the police of all the times a person has been arrested" (m-w.com)
- an official police document that lists the crimes that a particular person has committed (dictionary.cambridge.org)

WHAT'S A RAP SHEET?



163.44.136.42	Rapsheet Info Phishing source: http://apple.webstarterz.com/
164.215.229.75	Rapsheet Info Exploit source: SSH-Bruteforce
166.63.122.146	Rapsheet Info Phishing source: http://thewinekartdemo.cwwws.com/winelist/mobile.free.fr/a685s
168.1.77.95	Rapsheet Info Phishing source: http://www.bushcamping.com.au/b2308h/f0ld3r/S65.html?7B3NZ

Provider	
Detected as	phishing
Detections	=> phishing ATI-phishing => Phishing page
Phish target	NA
Screenshot	
URL	http://www.bushcamping.com.au/b2308h/f0ld3r/S65.html?7B3NZ=;
SHA256	741fc861353523d5cdc110704409b5d74740a70ac6138a0c412528
Runtime	



BARCLAYS

Quick, safe and convenient - Online Bai

Not yet registered for Online Banking? Register now.

Step 1 - Who are you?

Surname

How would you like to log in?

● Membership number

Forgotten your membership number?

○ Card number

○ Sort code and account number

□ Remember me. (We don't recommend this option if you're using a public or sh

Why this is safe

Step 2 - Confirm your ID

Barclays Bank PLC. Authorised by the Prudential Regulation Authority and regulated by the Financial Cor Register no: 122702). Barclays Bank PLC subscribes to the Lending Code which is monitored and enforc www.lendingstandardsboard.org.uk. Barclays Insurance Services Company Limited is authorised and reg 312078).

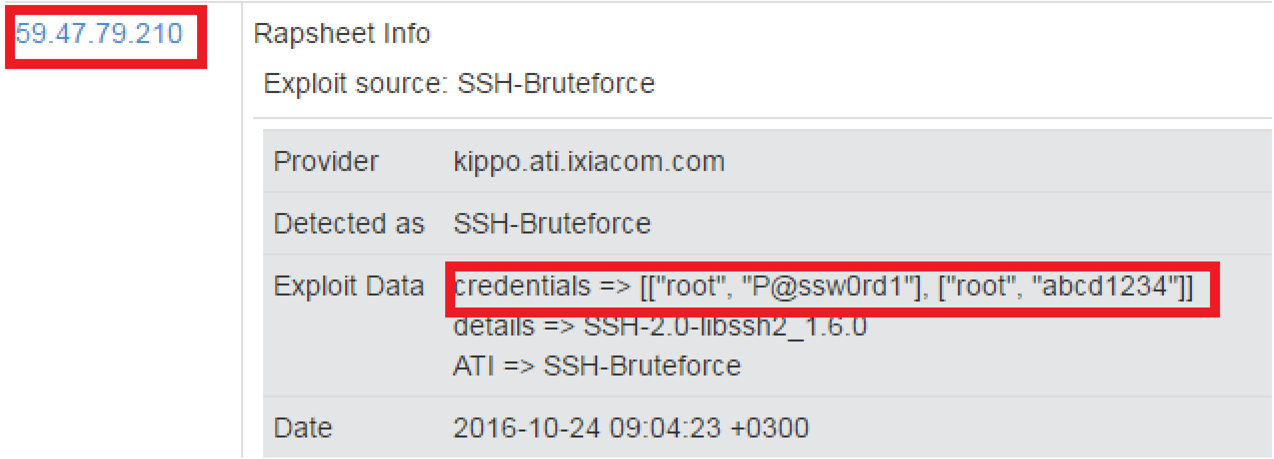
Barclays Bank PLC. Registered in England. Registered no. 1026167. Barclays Insurance Services Compar for both: 1 Churchill Place, London E14 5HP. 'The Woolwich' and 'Woolwich' are trademarks and trading r Barclays Bank PLC.

- Expanding upon the idea
- Track all malicious IPs on the Internet over time

WHAT'S A RAP SHEET?

- IP address or domain
- Proof of maliciousness
- **100%** certainty

- No moral judgement



59.47.79.210

Rapsheet Info

Exploit source: SSH-Bruteforce

Provider	kippo.ati.ixiacom.com
Detected as	SSH-Bruteforce

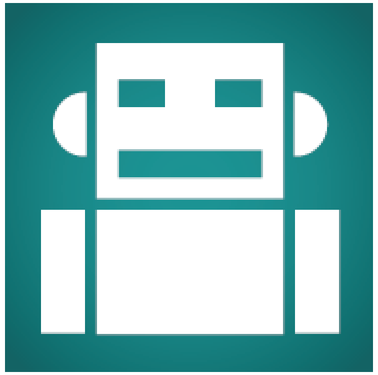
Exploit Data	credentials => [["root", "P@ssw0rd1"], ["root", "abcd1234"]]
	details => SSH-2.0-libssh2_1.6.0
	ATI => SSH-Bruteforce

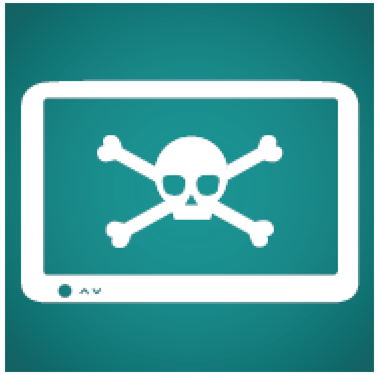
Date	2016-10-24 09:04:23 +0300

WHAT'S A RAP SHEET?

Examples

- Malware is dirty
- Exploit kits are dirty
- Bots exploiting vulnerabilities are dirty
- Phishing pages are dirty
- Spam is "clean"
- pr0n is "clean"





THREAT IDENTIFICATION (FINDING THE BADDIES)

THREAT IDENTIFICATION

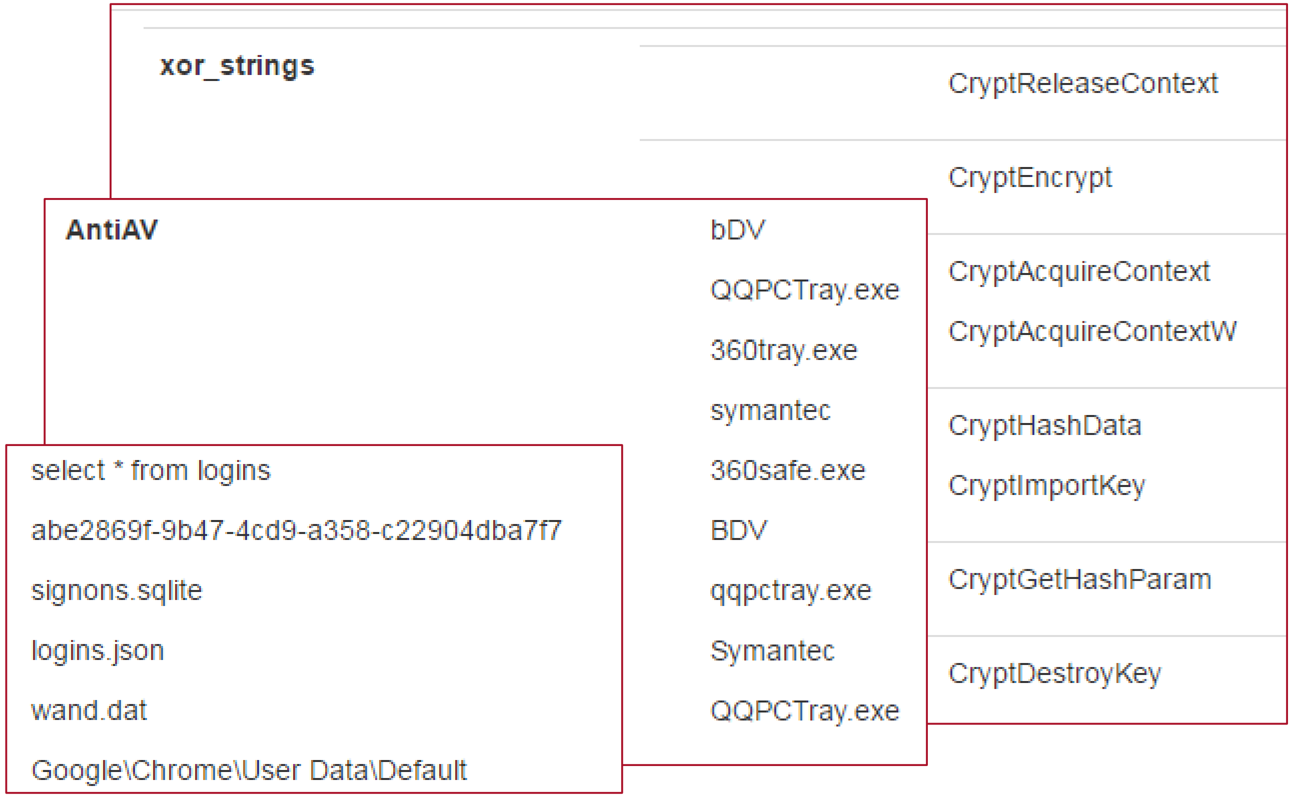
Virus Scanning

- Battery of AV products
- Threat intelligence feeds
- Detection threshold
    - Lower chance of False Positives

THREAT IDENTIFICATION

Static Analysis

- No execution
- Interesting properties/artefacts
  - Imported/mentioned functions
  - Sections
  - Entropy
  - Certificates
  - Particular strings
    - Children’s game using Mimikatz
    - Educational app dumping passwords
- Other intel on binary



xor_strings

CryptReleaseContext

CryptEncrypt

CryptAcquireContext
CryptAcquireContextW

CryptHashData
CryptImportKey

CryptGetHashParam

CryptDestroyKey

AntiAV

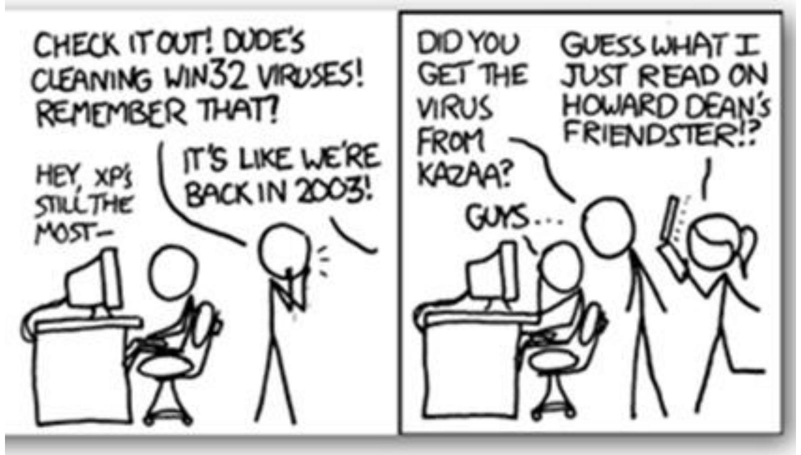
bDV
QQPCTray.exe
360tray.exe
symantec
360safe.exe
BDV
qqpctray.exe
Symantec
QQPCTray.exe

select * from logins
abe2869f-9b47-4cd9-a358-c22904dba7f7
signons.sqlite
logins.json
wand.dat
Google\Chrome\User Data\Default

THREAT IDENTIFICATION

Dynamic Analysis

- Cuckoo Sandbox
  - Great project!
  - VM/sandbox hardening is a must
    - Cuckoo does some of this work for you
  - VMs are easy to revert and reuse
  - (alternatively) Execute on hardware – slow cleanup


CHECK IT OUT! DUDE'S CLEANING WIN32 VIRUSES! REMEMBER THAT?
HEY, XP'S STILL THE MOST–
IT'S LIKE WE'RE BACK IN 2003!
DID YOU GET THE VIRUS FROM KAZAA?
GUESS WHAT I JUST READ ON HOWARD DEAN'S FRIENDSTER!?
GUYS...

https://xkcd.com/694/

- Analyze behavior on execution
- All sorts of honey
  - Applications, documents, credentials
- Grab more Intel – dropped files, contacted IPs, URLs, etc

THREAT IDENTIFICATION

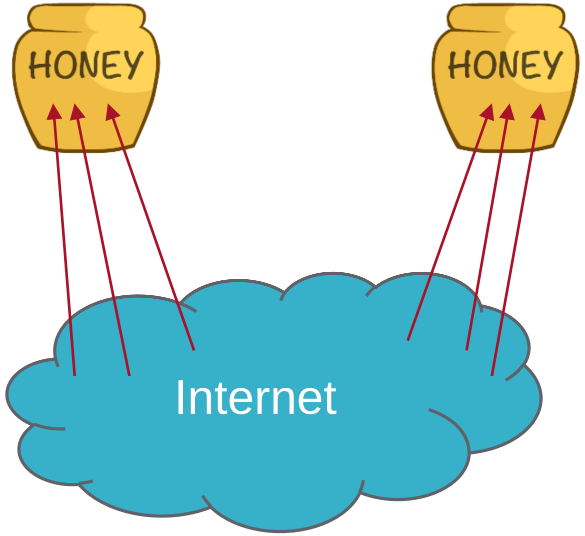
Dynamic Analysis

- Responsibility!
- Try not to:
    - Spam others
        - Gathering spam is also useful
    - DoS others
    - Brute force others
    - Infect others
- Sometimes more easier said than done!

THREAT IDENTIFICATION

Honeypots

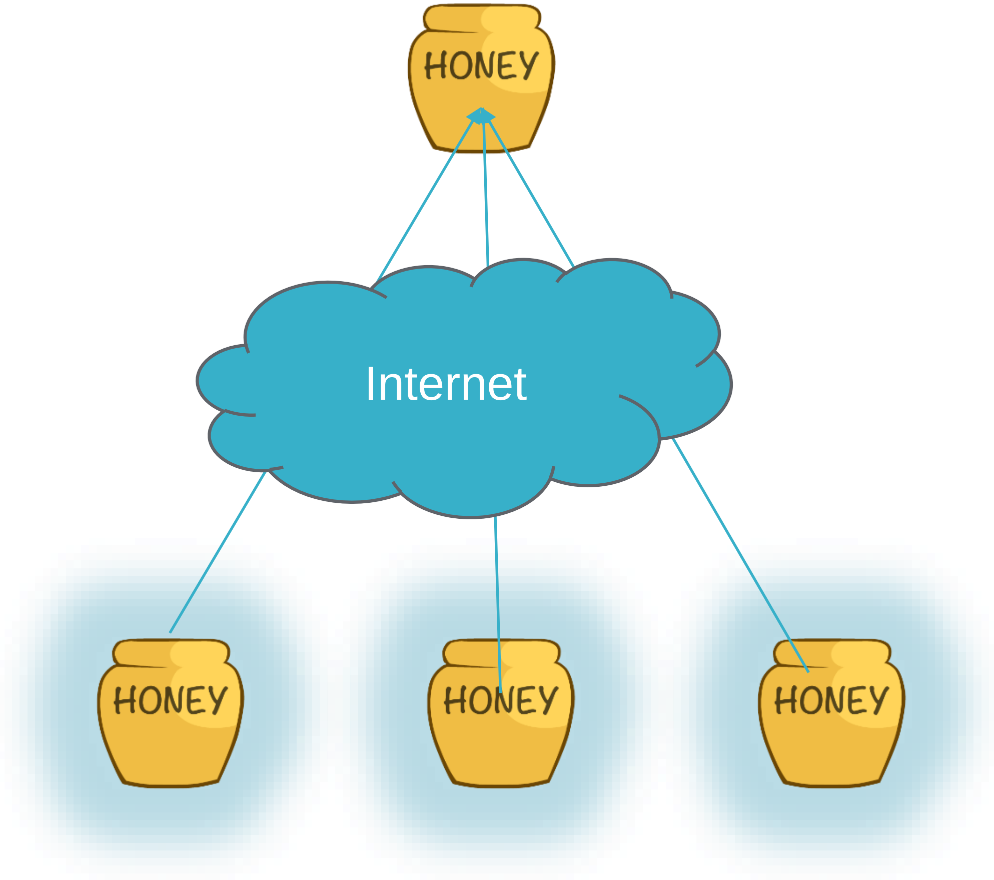
- Multiple honeypots
    - Dionaea, Kippo/Cowrie, Glastopf and others
- Globetrotting
    - Different continents
    - Different countries
    - Different provider sizes
- Hard to administer!



THREAT IDENTIFICATION

Honeypots

- Learning from the enemy
- Honeypot proxies
  - Forward to real honeypot
  - Easy to deploy/redeploy
  - No dependency/OS issues
- Honeypot blacklisting
- One jar for many flies
- Dump to central repository



Honey courtesy of Jeff Geerling and http://cliparts.co/

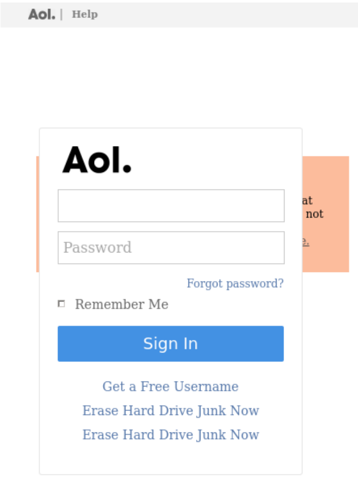
THREAT IDENTIFICATION

Honeypots

- False positives
  - Scanners – malicious or not?
  - Indexing bots – real or fake?
- Only identify attacks
  - Signatures for attacks

THREAT IDENTIFICATION

Phishing Detection



Aol. | Help

Aol.

Password

Forgot password?

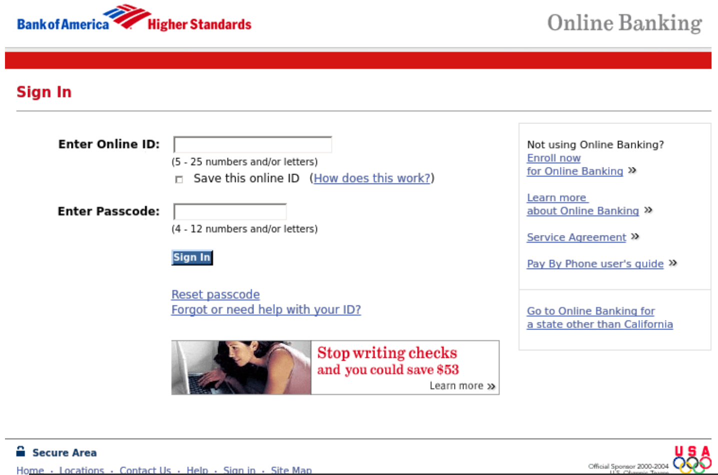
□ Remember Me

Sign In

Get a Free Username

Erase Hard Drive Junk Now

Erase Hard Drive Junk Now



Bank of America Higher Standards

Online Banking

Sign In

Enter Online ID: (5 - 25 numbers and/or letters)

□ Save this online ID (How does this work?)

Enter Passcode: (4 - 12 numbers and/or letters)

Sign In

Reset passcode

Forgot or need help with your ID?

Stop writing checks and you could save $53

Learn more »

Not using Online Banking?

Enroll now for Online Banking »

Learn more about Online Banking »

Service Agreement »

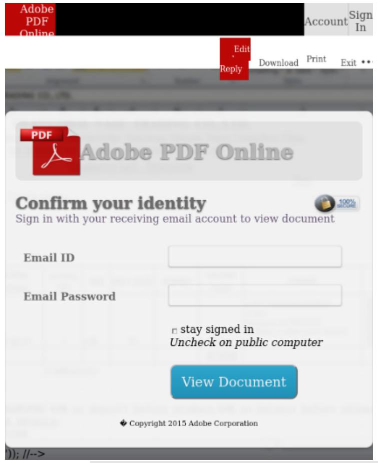
Pay By Phone user's guide »

Go to Online Banking for a state other than California

Secure Area

Home · Locations · Contact Us · Help · Sign in · Site Map

Official Sponsor 2000-2004



Adobe PDF Online

Account

Sign In

Edit Reply

Download

Print

Exit

Adobe PDF Online

Confirm your identity

Sign in with your receiving email account to view document

Email ID

Email Password

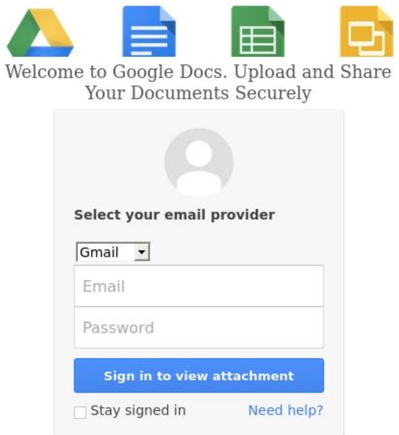
□ stay signed in

Uncheck on public computer

View Document

� Copyright 2015 Adobe Corporation

')); //-->



Welcome to Google Docs. Upload and Share Your Documents Securely

Select your email provider

Gmail

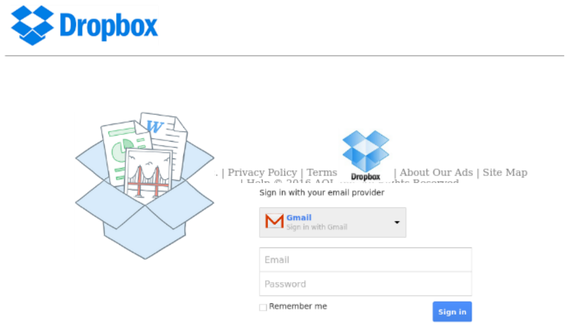
Email

Password

Sign in to view attachment

□ Stay signed in

Need help?



Dropbox

. | Privacy Policy | Terms | About Our Ads | Site Map

| Help © 2016 AOL ... Rights Reserved

Sign in with your email provider

Gmail

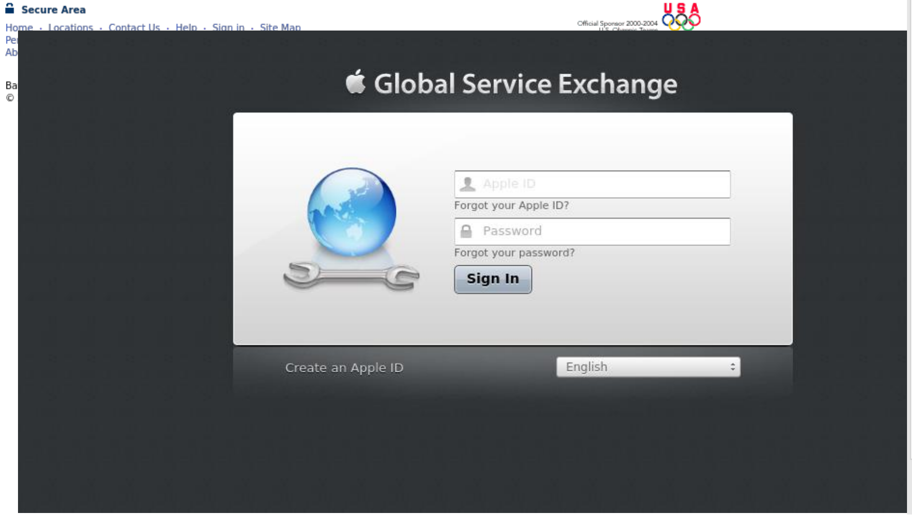
Sign in with Gmail

Email

Password

□ Remember me

Sign in



Global Service Exchange

Apple ID

Forgot your Apple ID?

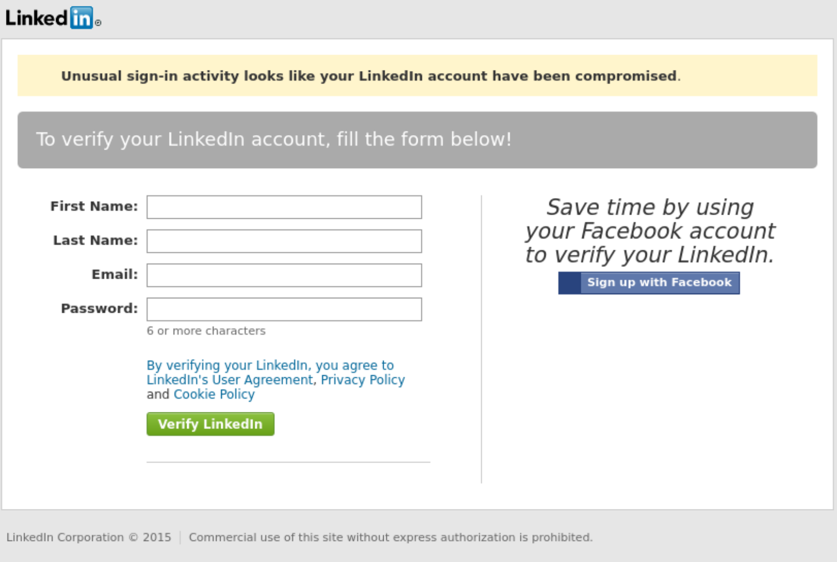
Password

Forgot your password?

Sign In

Create an Apple ID

English



LinkedIn

Unusual sign-in activity looks like your LinkedIn account have been compromised.

To verify your LinkedIn account, fill the form below!

First Name:

Last Name:

Email:

Password:

6 or more characters

By verifying your LinkedIn, you agree to LinkedIn's User Agreement, Privacy Policy and Cookie Policy

Verify LinkedIn

Save time by using your Facebook account to verify your LinkedIn.

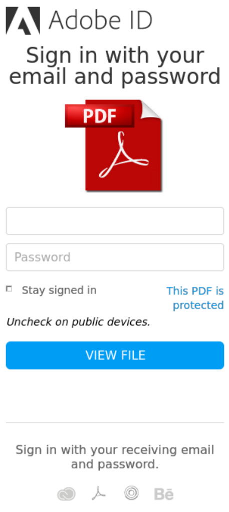
Sign up with Facebook

LinkedIn Corporation © 2015 | Commercial use of this site without express authorization is prohibited.

THREAT IDENTIFICATION

Phishing Detection

- Static “signatures” for larger targets
- Plenty of challenges
  - Signature development
    - No false positives
    - Originals will always match
      - Phish of phish of phish…
  - Redo periodically
  - Limited detection
  - Easy to bypass
    - Must look Facebook-y or Google-y, not exact clone



Adobe ID

Sign in with your email and password

Password

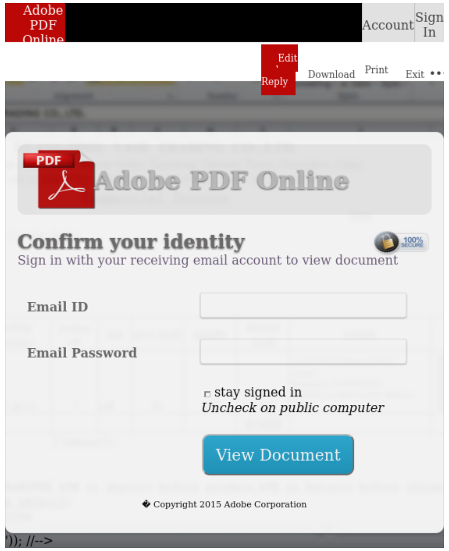
□ Stay signed in

This PDF is protected

Uncheck on public devices.

VIEW FILE

Sign in with your receiving email and password.



Adobe PDF Online

Confirm your identity

Sign in with your receiving email account to view document

Email ID

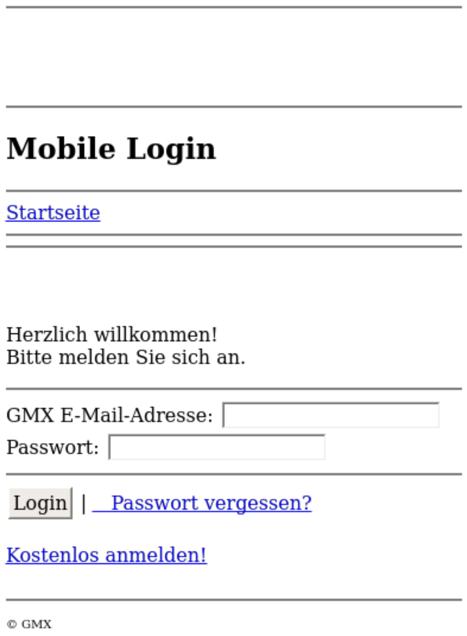
Email Password

□ stay signed in
Uncheck on public computer

View Document

� Copyright 2015 Adobe Corporation

')); //-->



Mobile Login

Startseite

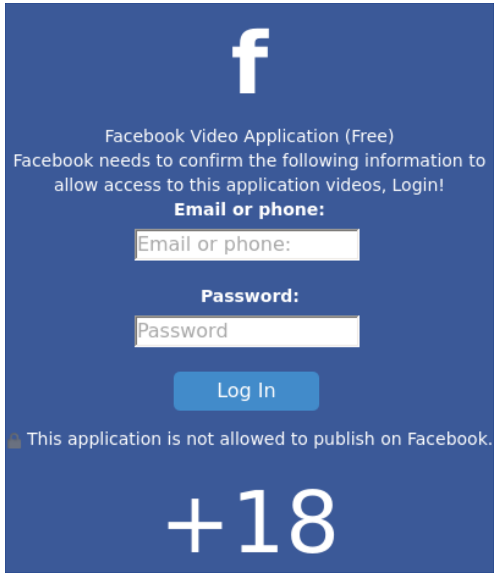
Herzlich willkommen!
Bitte melden Sie sich an.

GMX E-Mail-Adresse:
Passwort:

Login | Passwort vergessen?

Kostenlos anmelden!

© GMX



Facebook Video Application (Free)
Facebook needs to confirm the following information to allow access to this application videos, Login!

Email or phone:

Email or phone:

Password:

Password

Log In

This application is not allowed to publish on Facebook.

+18

THREAT IDENTIFICATION

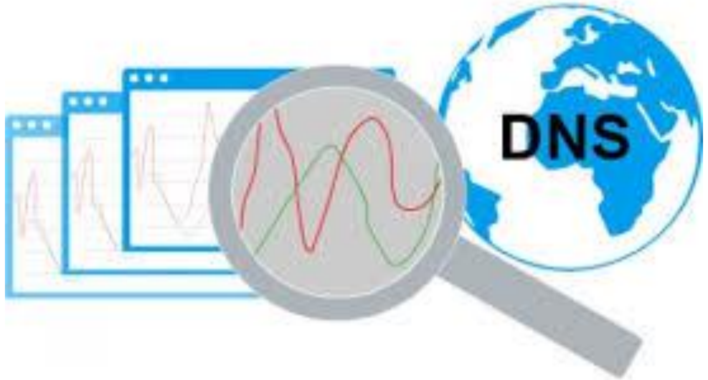
Phishing Detection

- Generic, machine learning-based approach
- Processes the HTML code
- Tries to classify correctly
- False positives likely

THREAT IDENTIFICATION

Passive DNS

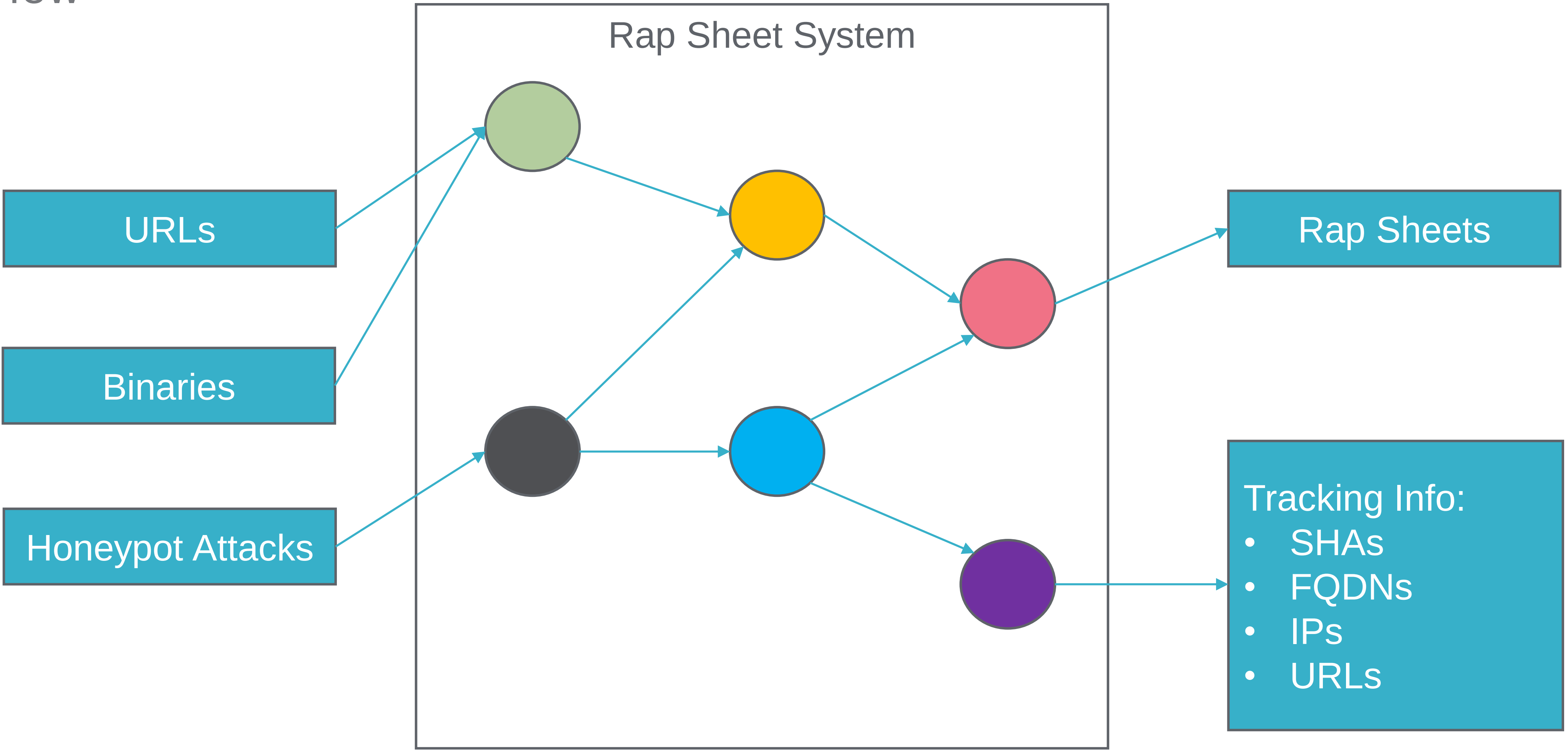
- Hostnames, domains and IP addresses
- Valuable information
  - Important infrastructure services (whitelisting)
  - Reoccurrences
  - Mapping threat actors



HOW DO WE MAKE SENSE OF ALL THE DATA WE'RE COLLECTING?

STORING AND INTERPRETING DATA

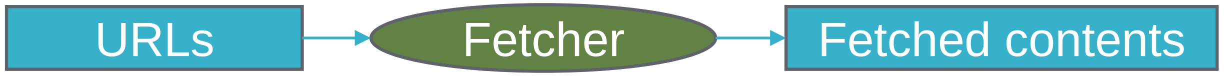
Overview



STORING AND INTERPRETING DATA

Some facts

- Real-time system
  - As soon as we have enough information to build a Rap Sheet, we build & publish it
- Dataflow model
  - Each node receives some input and produces some output. E.g..

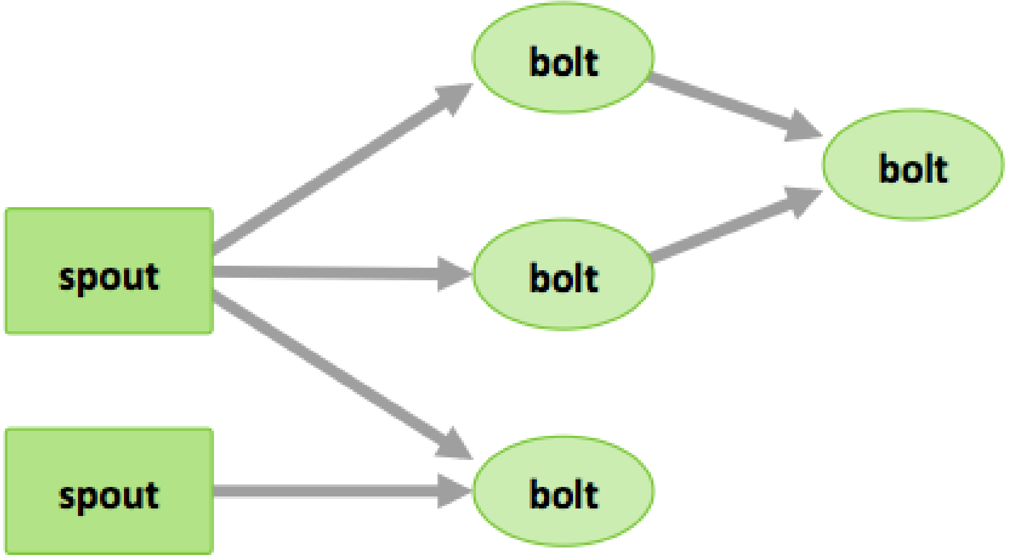


  - Nodes are connected to one another in a Topology
  - Nodes may interact with other external services (Databases, Storage, Sandbox execution, etc.)
  - Special nodes
    - Only produce output (e.g. scanning threat intelligence feeds and extracting URLs for processing)
    - Only receive input (e.g. storing the final information into the database)

STORING AND INTERPRETING DATA

Why this model ?

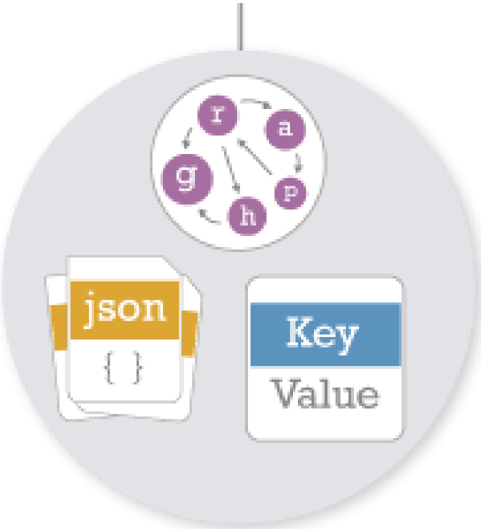
- Each node does one small thing (microservices anyone ? ☺)
    - Easy to develop & test
    - Easy to reuse
    - Easy to reason about
- Nodes can be combined in different ways in a topology
- Nodes can be scaled individually
- New nodes can be easily integrated
- Each node can be updated individually



STORING AND INTERPRETING DATA

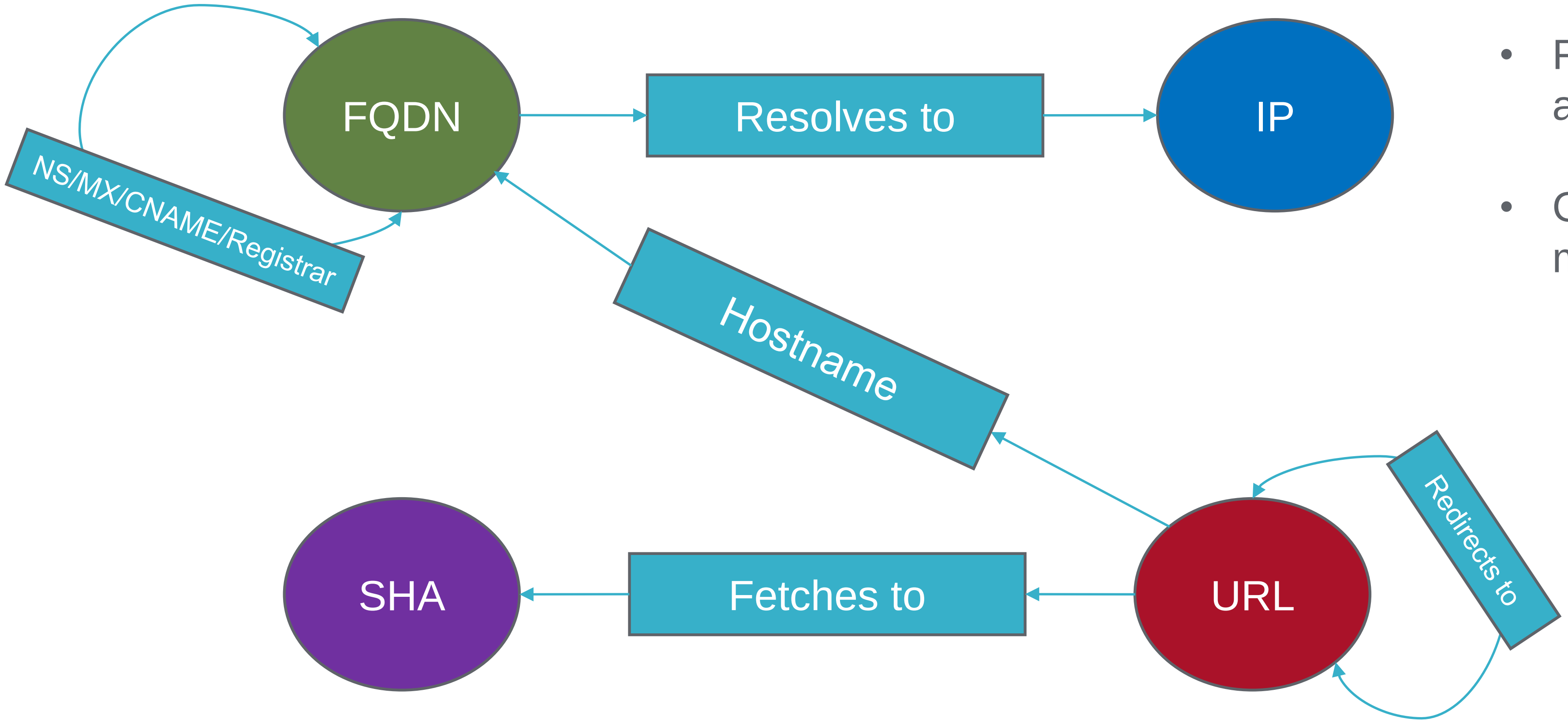
Where do we store data ?

- Blob data
    - Fetched URL contents, packet captures, dropped files, sandbox analysis results, screenshots, etc.
    - Cloud storage (currently AWS S3)
- Structured data
    - NoSQL multi-model DB called ArangoDB
        - Key/value
        - Document
        - Graph
    - Schemaless
        - JSON objects
        - Easy to add new fields & information
    - AQL, transactions, indexes, joins



STORING AND INTERPRETING DATA

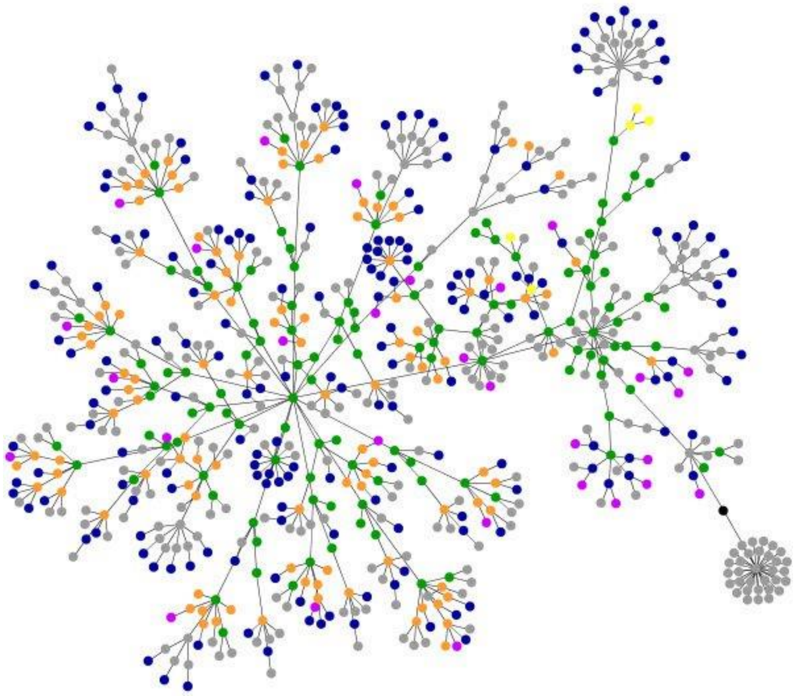
How do we store things?



- Four main entities with associated information
- Graph links between them with more information

STORING AND INTERPRETING DATA

Why do we store it this way?



- Graph naturally maps to the underlying problem domain

- We can do interesting queries like:
  - Finding all IPs that have served a certain malicious SHA
  - Finding redirector domains (& URL shortening services)
  - Finding other domains sharing the same NS/MX servers
    (as well the the usual PassiveDNS type queries)
  - Finding IPs & Domains which served SHAs which were dropped during dynamic analysis by SHAs coming from a specific IP.

- Flexibility:
  - Can easily tack on new information & entities to the Graph

STORING AND INTERPRETING DATA

Scaling it up

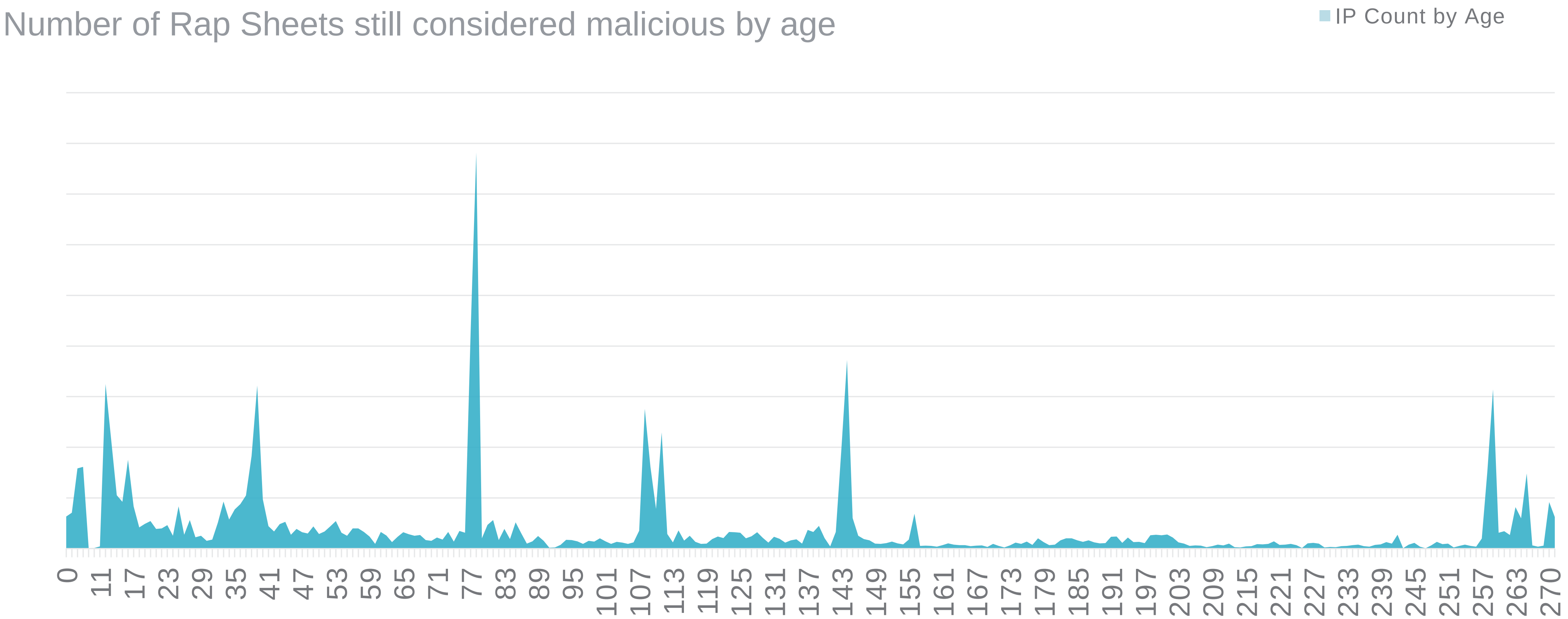
- Tech
  - Datacenter OS & Apache Mesos
    - Simple deployment for lots of distributed services (Redis, ELK, RabbitMQ, ArangoDB, Storm, etc.)
  - Containerize all the things and then run them using Marathon
  - Apache Storm

- Principles
  - Split up your components
  - Split up your database
  - Common and battle hardened infrastructure components
  - Moving code to data instead of the other way around

STATS

STATS

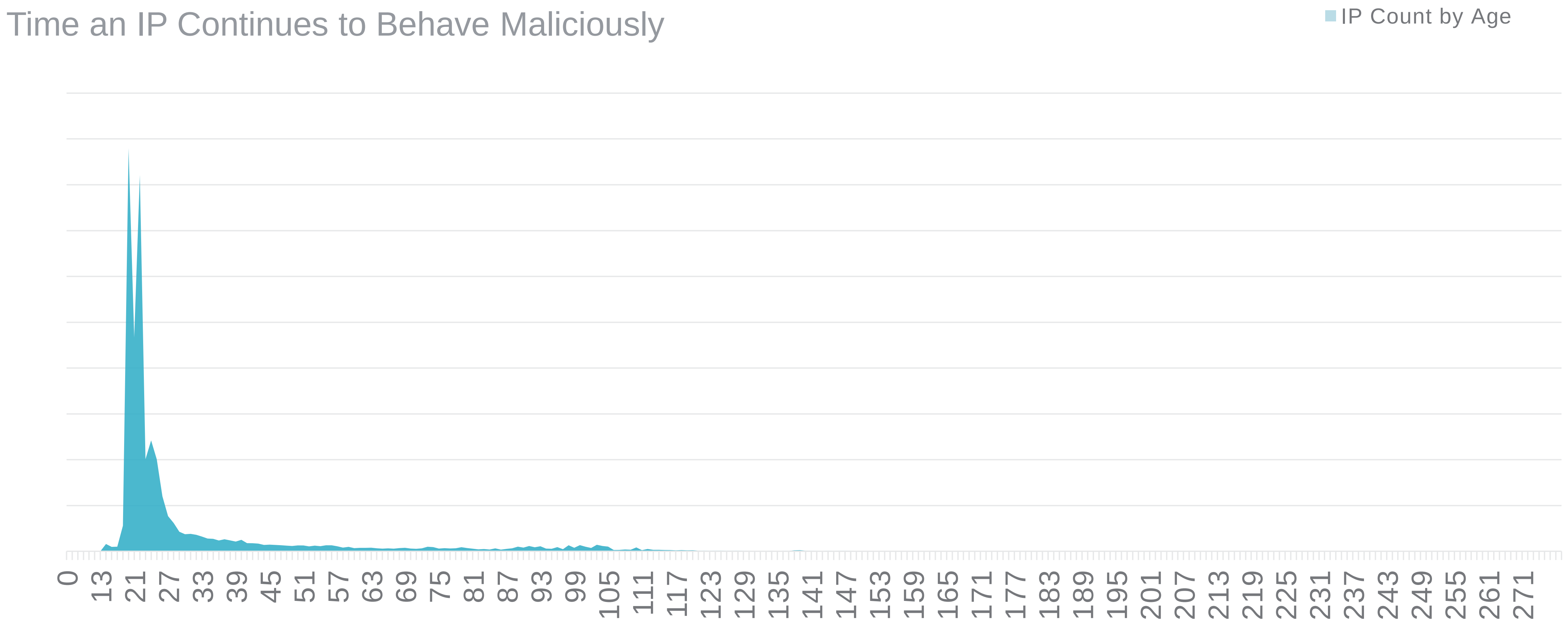
Active Rap Sheet Age



Number of Rap Sheets still considered malicious by age

STATS

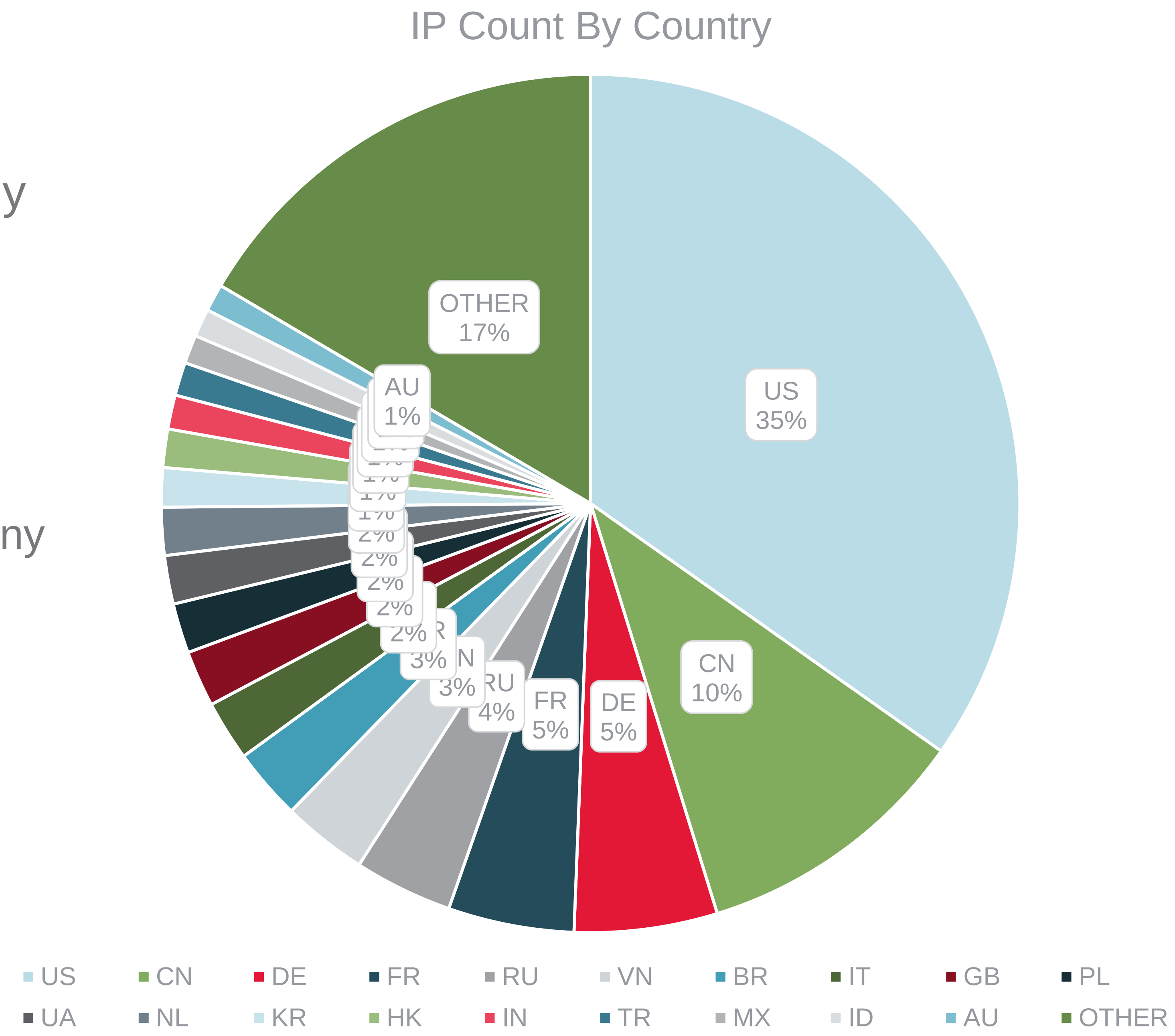
Delisted Rap Sheet Age



STATS

Malicious IP Addresses by Country

- Top 20 – 85%
- “Usual suspects” – US, China, Germany
- Surprising – Vietnam, Indonesia



|

ixia