Botnets

How (not) to count the Internet.

The Mirai Botnet Attacks

DDoS attack that disrupted internet was largest of its kind in history, experts say

Dyn, the victim of last week's denial of service attack, said it was orchestrated using a weapon called the Mirai botnet as the 'primary source of malicious attack'

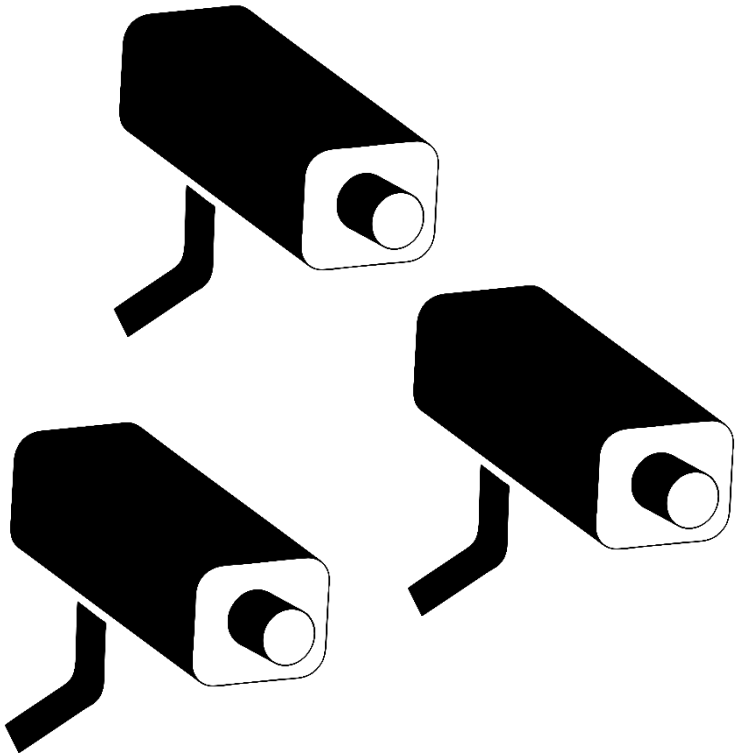
- Major cyber attack disrupts internet service across Europe and US

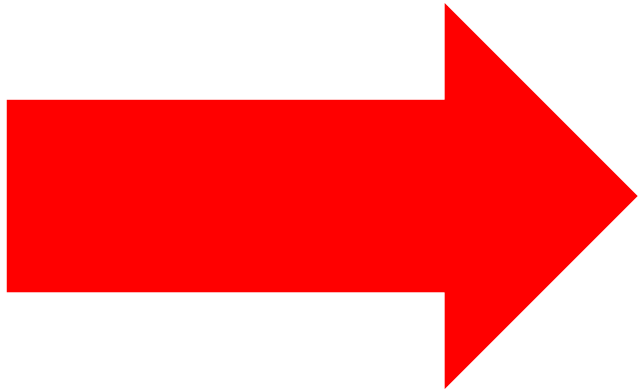
The Mirai Botnet Attacks

World's largest 1 Tbps DDoS Attack launched from 152,000 hacked Smart Devices

Tuesday, September 27, 2016 Swati Khandelwal

The Mirai Botnet Attacks





reddit

twitter

GitHub

airbnb

Denial of Service - DoS

Prevent legitimate users from accessing a service…

Distributed Denial of Service - DDoS

Prevent legitimate users from accessing a service…

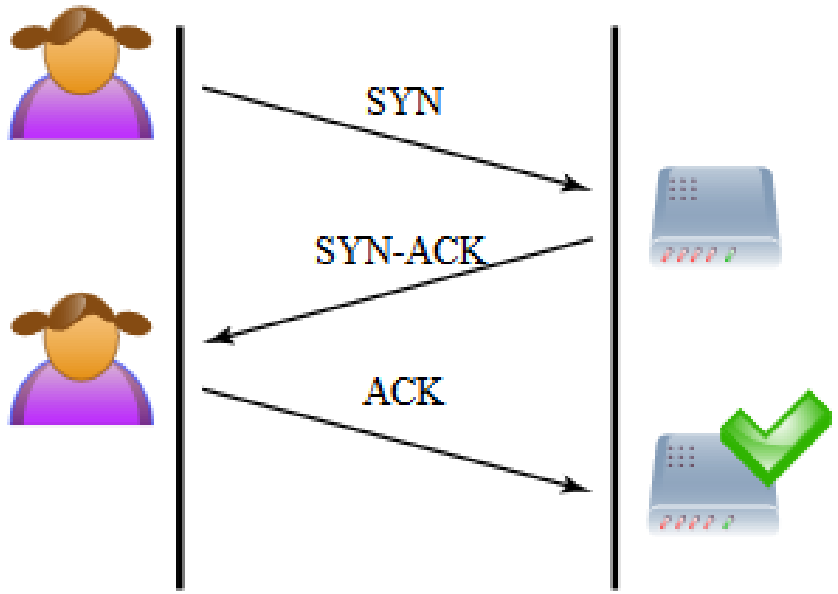
…using a distributed network, e.g. a Botnet.

Usually by sending a lot of packets.

DDoS Attack Examples

- Syn flood attack.
- Reflector attack.
- …

Syn flood attack



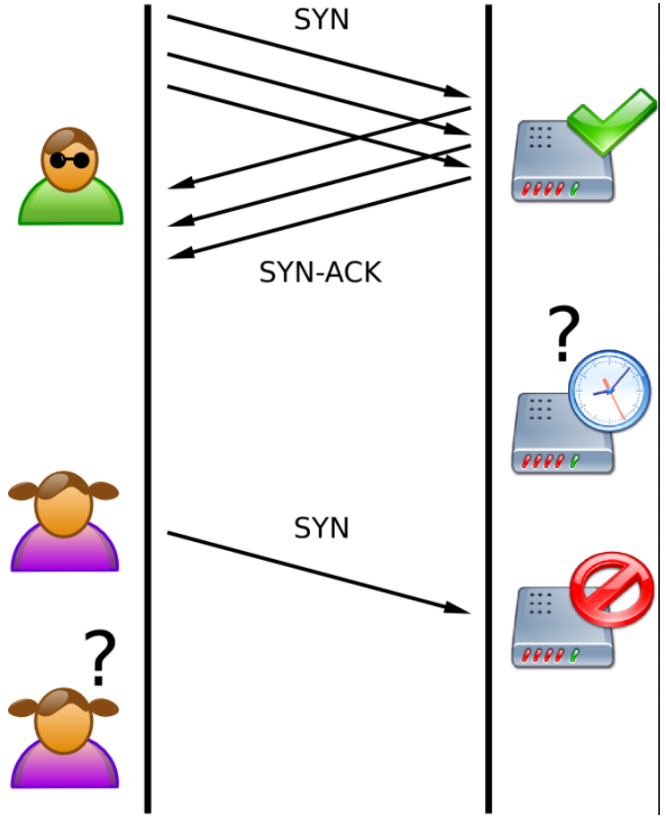
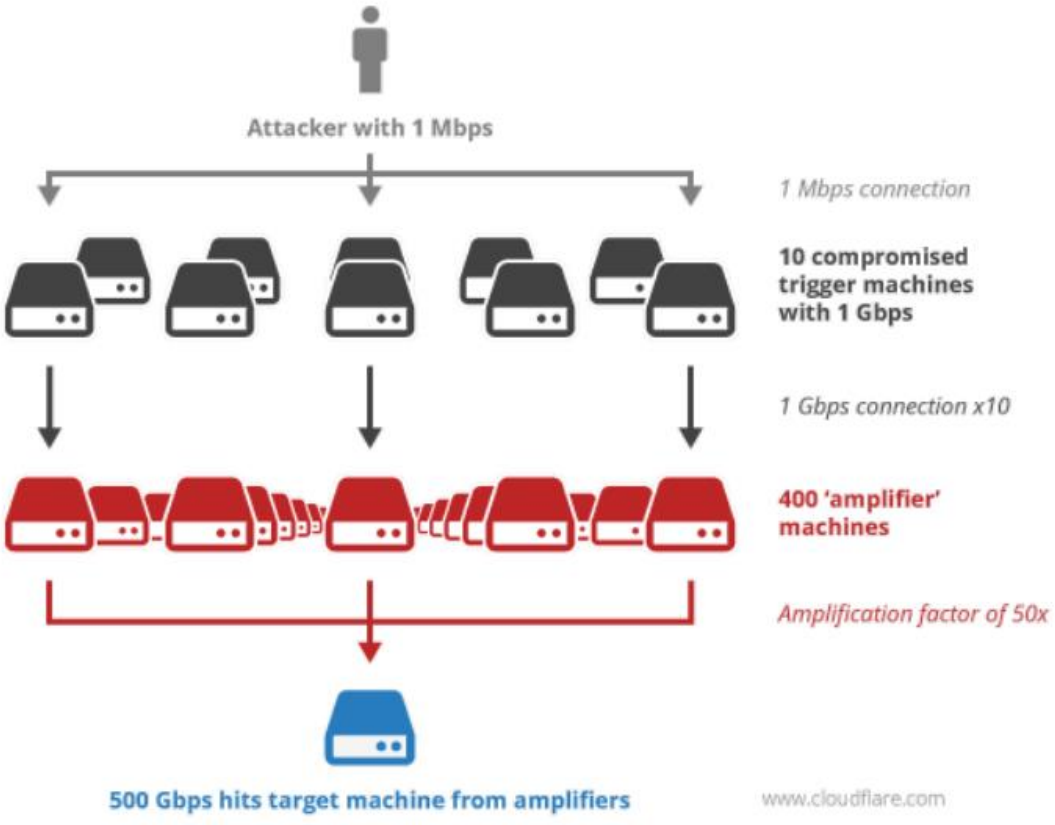
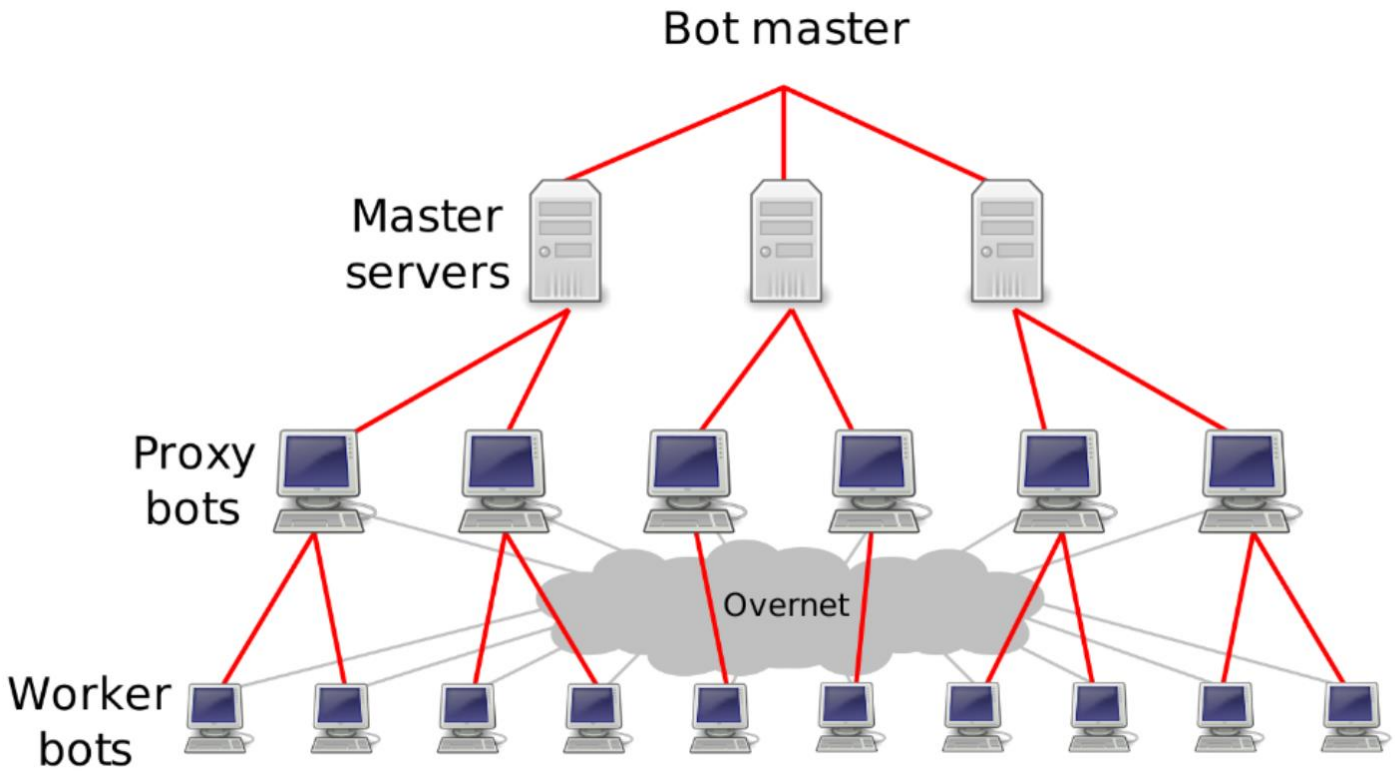


Image Source: Wikipedia

Reflection attack



Botnet Topology



Source: Spamalytics: An Empirical Analysis of Spam Marketing Conversion, Kanich et al.

Topology Attributes

- Command latency
- Resilience
- Bot awareness
- Planning

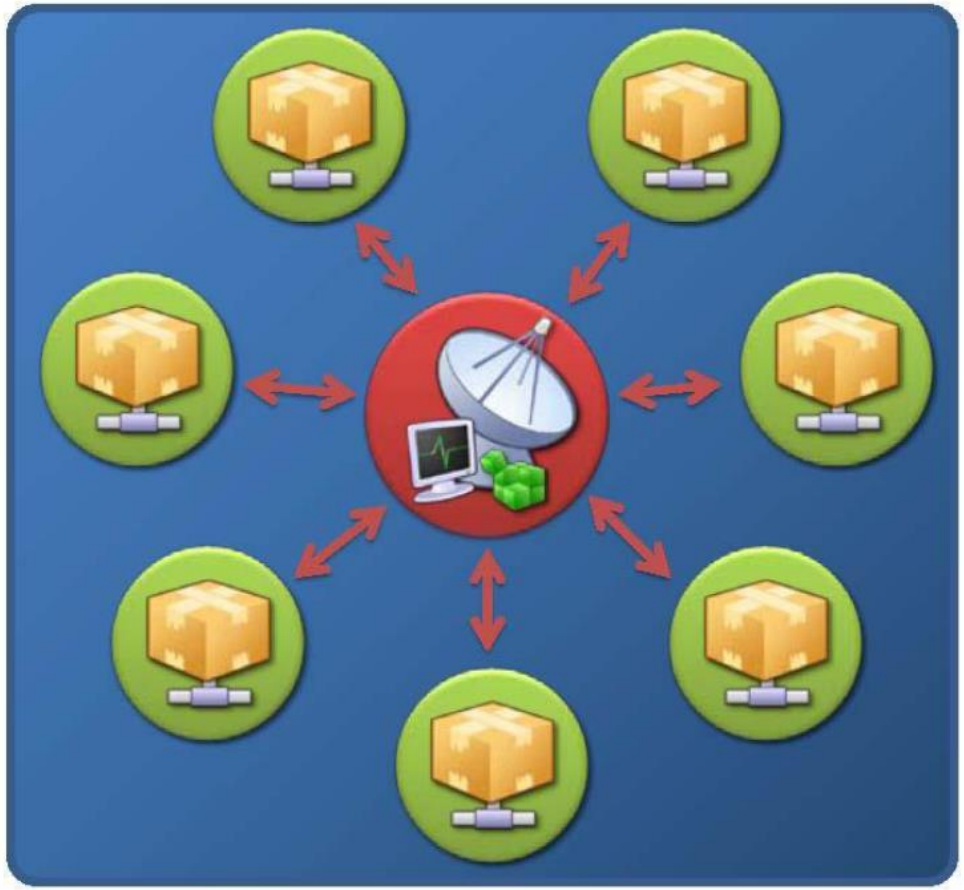
Star

Command latency (+)

Resilience (-)

Bot awareness(-)

Planning (+)



Source: G. Ollmann - Damballa, http://bit.ly/BotCom

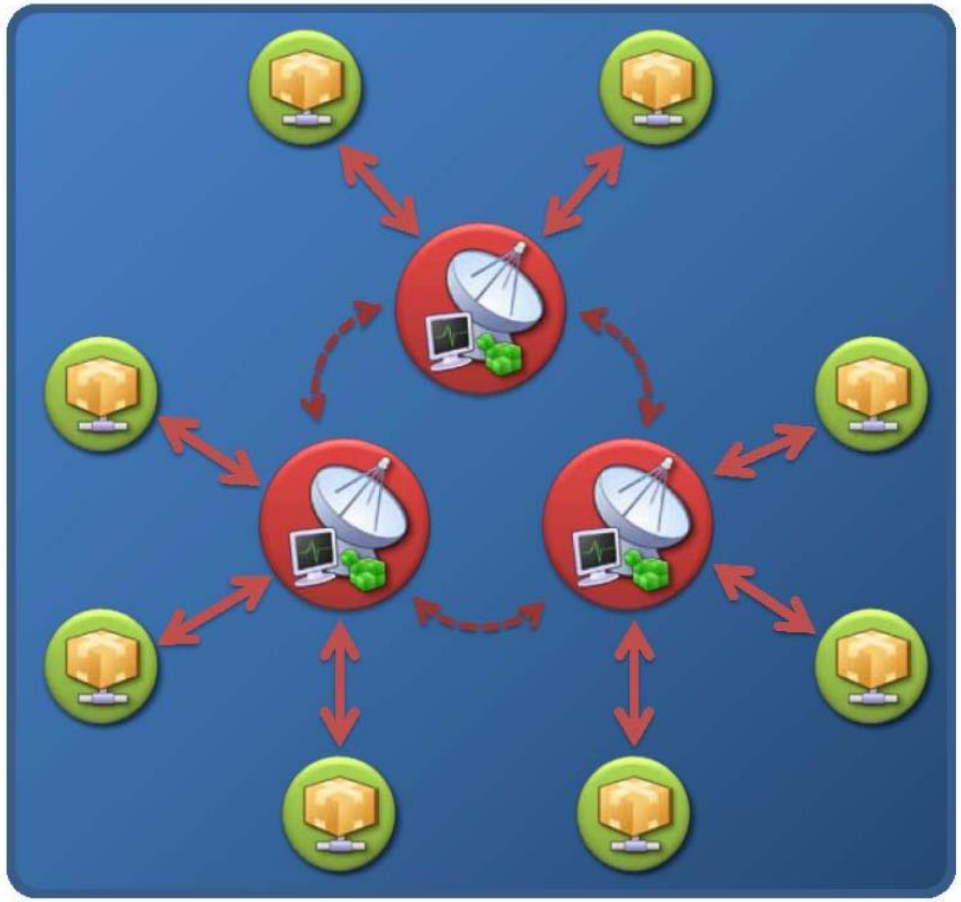
Multi Server

Command latency (+)

Resilience (+)

Bot awareness(-)

Planning (-)



Source: G. Ollmann - Damballa, http://bit.ly/BotCom

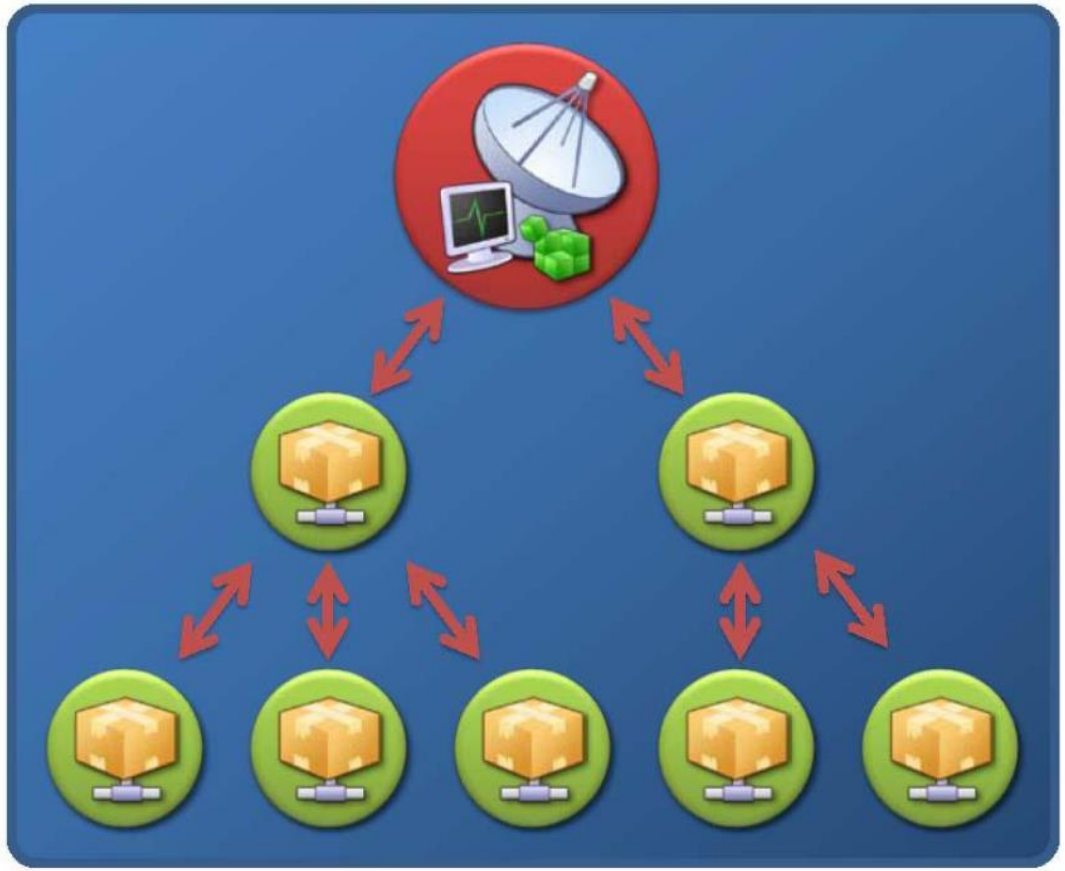
Hierarchical

Command latency (-)

Resilience (+)

Bot awareness(+)

Planning (+)



Source: G. Ollmann - Damballa, http://bit.ly/BotCom

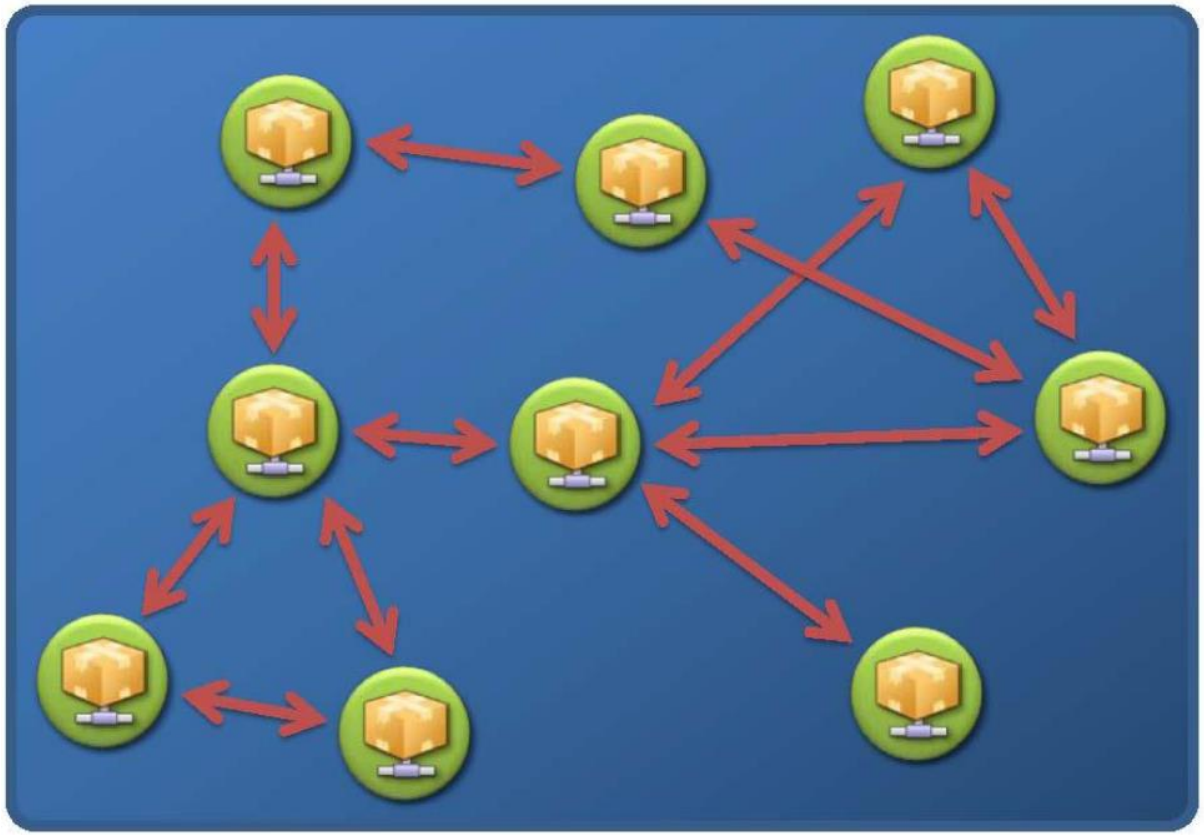
Random P2P

Command latency (-)

Resilience (+)

Bot awareness(-)

Planning (+)



Source: G. Ollmann - Damballa, http://bit.ly/BotCom

Botnet Communication

- No communication
- Public channels
- Private channels
- Hybrid

How Do Bots Find The Master?

- IP flux
- Domain flux

IP Flux

Fully qualified domain name. *e.g. mypc.atl.damballa.com*

Constantly change IP address of this domain.

Single flux vs. double flux

Domain Flux

Inverse of IP flux

Domain Generation Algorithms (DGAs)

DGA Example - TorPig

- Three fixed domains to be used if all else fails.
- Daily/weekly domain name (dd/wd)
- Every 20 minutes bot attempts to connect (in order) to:

    wd.com, wd.net, wd.biz

    dd.com, dd.net, dd.biz

    the three fixed domains

Source: http://www.cs.ucsb.edu/~kemm/courses/cs177/torpig.pdf

Newly Infected Device – What Now?

Persist, avoid detection

Social attacks

Eventually aggressive attacks

Rent it to someone else

Paper 1: Botnets As A Service

Analysed traffic across the globe.

Labeled IPs to known botnets.

Source: Characterizing Botnets-as-a-Service, Chang et al.

Characteristics

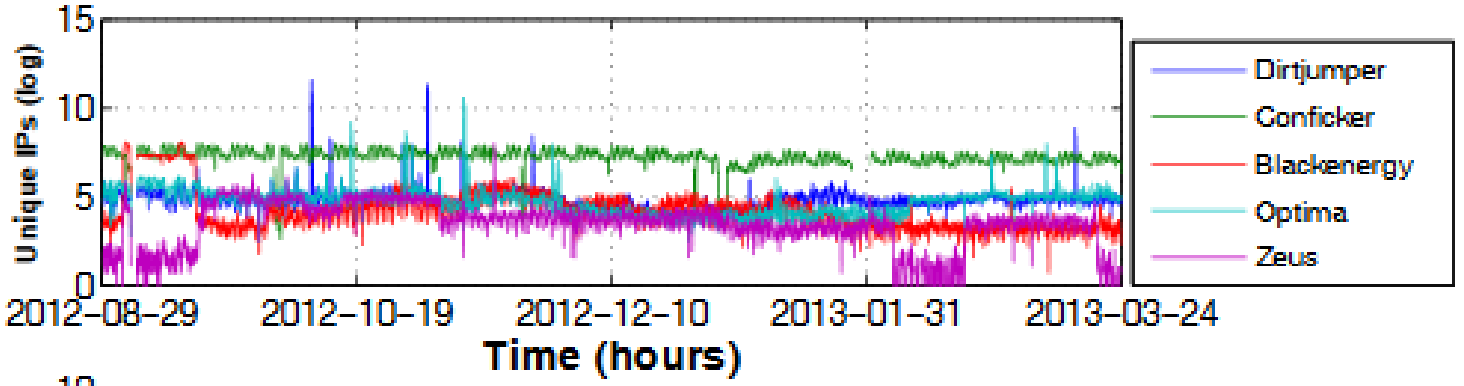
Size

Stability

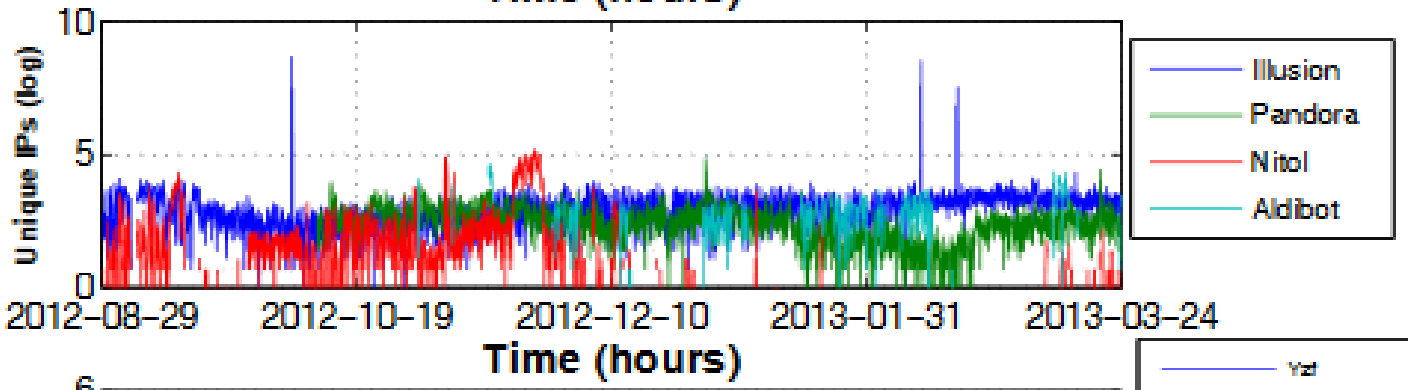
Elasticity

Source: Characterizing Botnets-as-a-Service, Chang et al.

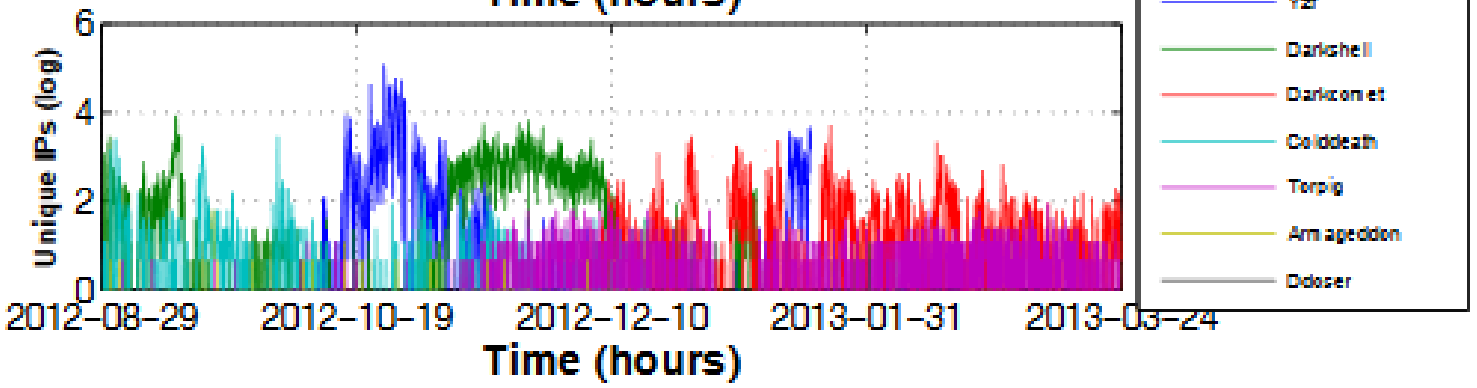
Large



Medium



Small



Source: Characterizing Botnets-as-a-Service, Chang et al.

Botnet Trends

Large botnets have dynamic stability.

Large botnets tend to be more elastic.

Botnets collaborate.

Botnet Trends

Large botnets have dynamic stability.

Large botnets tend to be more elastic.

Botnets collaborate?

What Does This Remind You Off?

- Service based
- Scalable
- Elastic
- Metered
- Redundant
- Highly available

“Cloud Computing”

Are Botnets Always Bad?

Paper 2: Internet Census 2012

Count number of used IPs.

Used a botnet for scanning.

Published anonymously in March 2013

Carna Botnet

“Carna was the roman goddess for the protection of inner organs and health and was later confused with the goddess of doorsteps and hinges. This name seems like a good choice for a bot that runs mostly on embedded routers.”

The Beginning

Discovered vulnerable devices when playing around with nmap scripting engine.

Scanned on port 23.

Small binary loaded into vulnerable devices.

in one night ~30 thousand devices infected.

Implementation – Be Nice!

- Don’t change passwords.
- No permanent changes.
- Limited scanning speed to ~10 IPs/s.
- Added a Readme file.

Found Devices

Routers, set-top boxes	~25%
IPSec routers, BGP routers, industrial control systems, door security systems,…	~75%

Targets

- Only Routers and set-top boxes
- Approx. 420k devices infected
- Some Bots act as middle nodes (proxies)

Tools

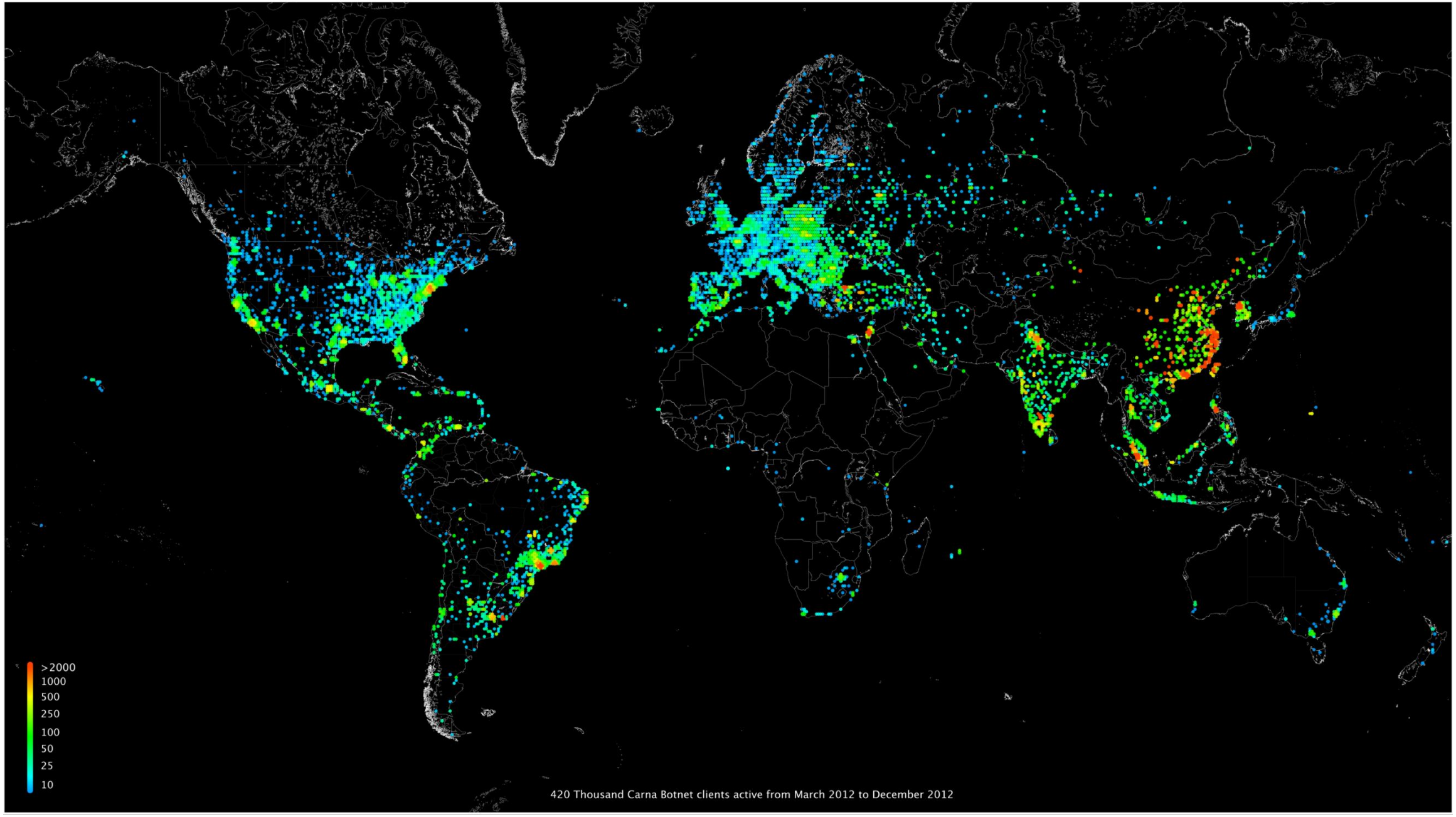
- Binary written in C
- Web Interface written in PHP
- Python scripts
- Apache Hadoop with PIG to handle data

Scanning Methods

- ICMP ping
- Reverse DNS
- Nmap SYN scans
- Nmap service probes
- Traceroute

Scanning Methods

- ICMP ping
  - 52 billion probes
- Reverse DNS
  - 10.5 billion stored records
- Nmap SYN scans
  - 2.8 billion records for ~660 million IPs with 71 billion ports tested
- Nmap service probes
  - 4000 billion probes sent, ~175 billion answered
- Traceroute
  - 68 million records


>2000
1000
500
250
100
50
25
10
420 Thousand Carna Botnet clients active from March 2012 to December 2012

Surprises

Suspicious Binary in /tmp folder

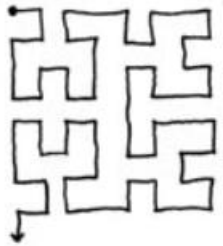
AIDRA: Classic Botnet with IRC CnC Server

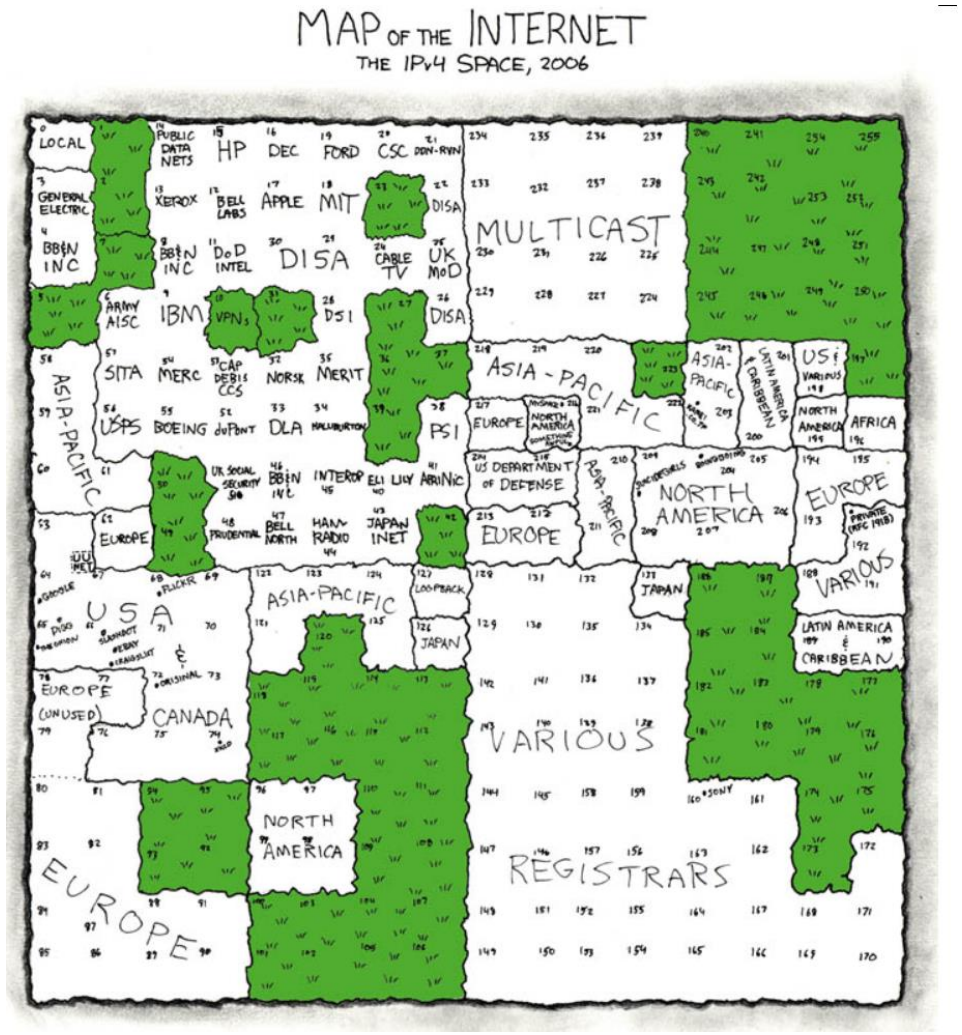
Over 250KB size

Less bots than Carna

Analysis

0 1 14 15 16 19 →
3 2 13 12 17 18
4 7 8 11
5 6 9 10

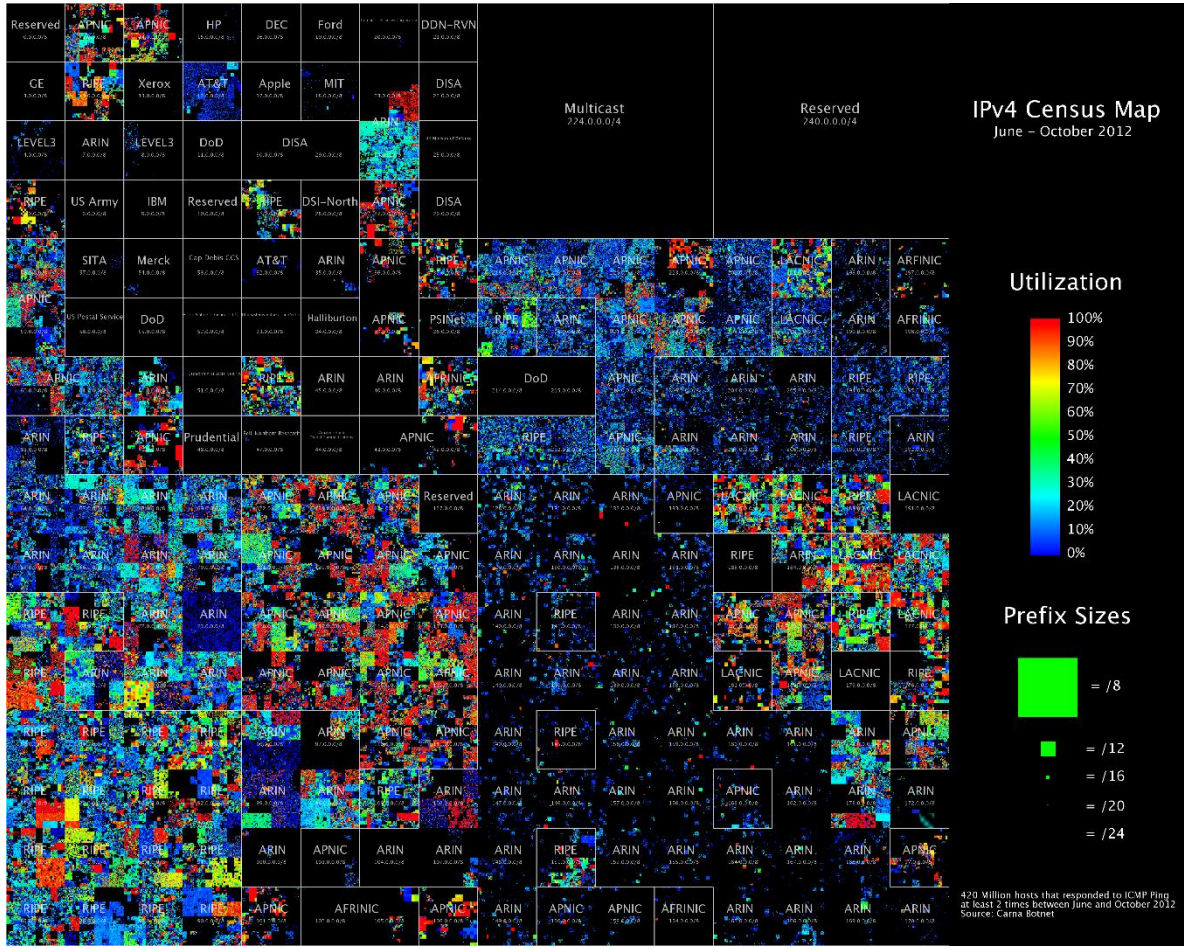


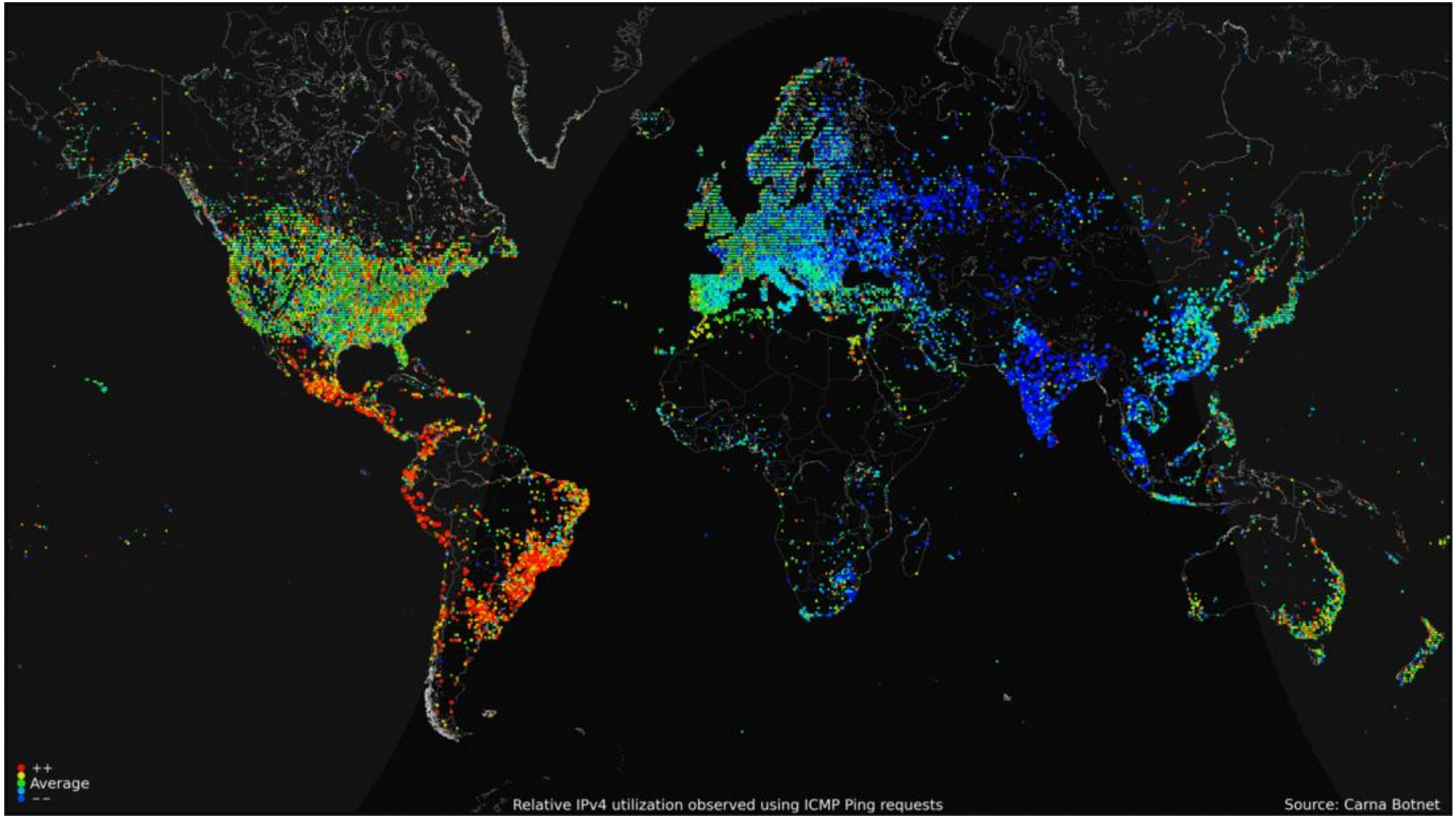


Source: xkcd.com/195

Analysis


MAP OF THE INTERNET
THE IPv4 SPACE, 2006
LOCAL
PUBLIC DATA NETS
HP
DEC
FORD
CSC
DDN-RVN
GENERAL ELECTRIC
XEROX
BELL LABS
APPLE
MIT
DISA
BBN INC
CABLE TV
UK MOD
ARMY AISC
IBM
VPNs
DSI
SITA
MERC
CAP DEBIS CCS
NORSK
MERIT
USPS
BOEING
duPONT
DLA
HALLIBURTON
ASIA-PACIFIC
MULTICAST
EUROPE
NORTH AMERICA
US DEPARTMENT OF DEFENSE
PSI
UK SOCIAL SECURITY
INTEROP
ELI LILLY
ARINIC
PRUDENTIAL
BELL NORTH
HAM RADIO
JAPAN INET
USA
CANADA
EUROPE (UNUSED)
LOOPBACK
JAPAN
VARIOUS
REGISTRARS
SONY
AFRICA
LATIN AMERICA
CARIBBEAN


IPv4 Census Map
June – October 2012
Multicast
224.0.0.0/4
Reserved
240.0.0.0/4
Utilization
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Prefix Sizes
= /8
= /12
= /16
= /20
= /24
420 Million hosts that responded to ICMP Ping at least 2 times between June and October 2012
Source: Carna Botnet


++
Average
--
Relative IPv4 utilization observed using ICMP Ping requests
Source: Carna Botnet

Authors Comments

“nobody would connect that to the Internet, really nobody”

there are at least 1000 people who did.

“that shouldn't be on the Internet but will probably be found a few times”

it's there a few hundred thousand times. Like half a million printers, or a Million Webcams,…”

Authors Comments

“nobody would connect that to the Internet, really nobody”

there are at least 1000 people who did.

“that shouldn't be on the Internet but will probably be found a few times”

it's there a few hundred thousand times. Like half a million printers, or a Million Webcams, or devices that have root as a root password.”

So How Big Is The Internet?

- 420 Million pingable IPs.
- 36 Million that had one or more ports open.
- 141 Million IPs firewalled.
- 729 Million more IPs just had reverse DNS records.

Total ~1.3 billion IPs in use.

Conclusion

“…to our knowledge, the largest and most comprehensive IPv4 census ever.”

No, it’s not. Bigger Census done in 2004, 2009,…

“We hope other researchers will find the data we have collected useful”

Difficult to say.

Problems With This Work

- Hard to verify
- Technically illegal

Krenc et al.

“CAIDA has confirmed that the scanning took place”

Reverse DNS: separate, external dataset from Nov 2012

95.2% exact matches

ICMP Dataset

- 2 claims: complete scans within 24h and scans over six weeks

  “from June 2012 to October 2012” no data from June/July!

- report states 52 billion probes, the data set only contains 49.5 billion probes

- Almost no metadata available

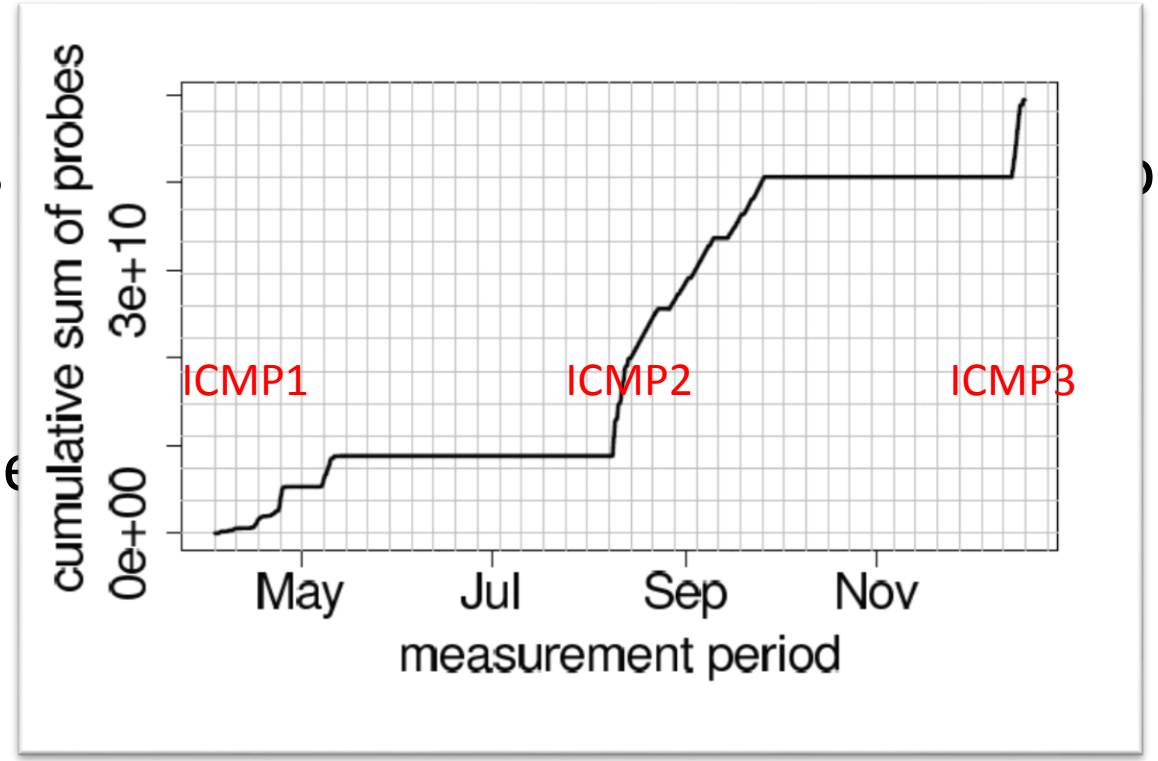
ICMP Dataset

- 2 claims: complete scans within 24h and scans over six weeks

  “from June 2012 to October 2012” no data from June/July!

- report states ntains 49.5 billion probes

- Almost no me



Finding The Scans

Analysing the probed IPs:

At most one complete scan possible .

Estimated between 1 and 12 “complete” scans.

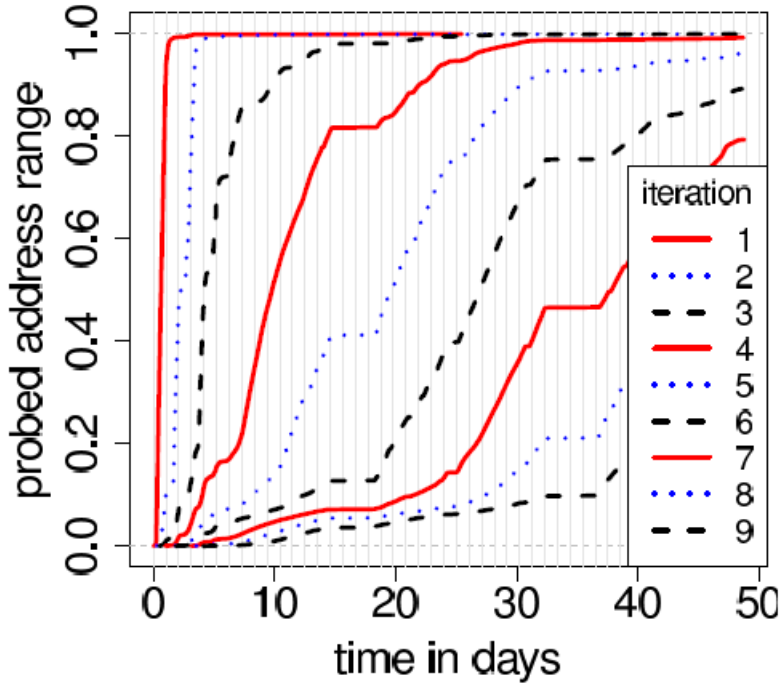


Figure 5: Overlapping iterations: First nine iterations over the probed address range in $icmp_2$.

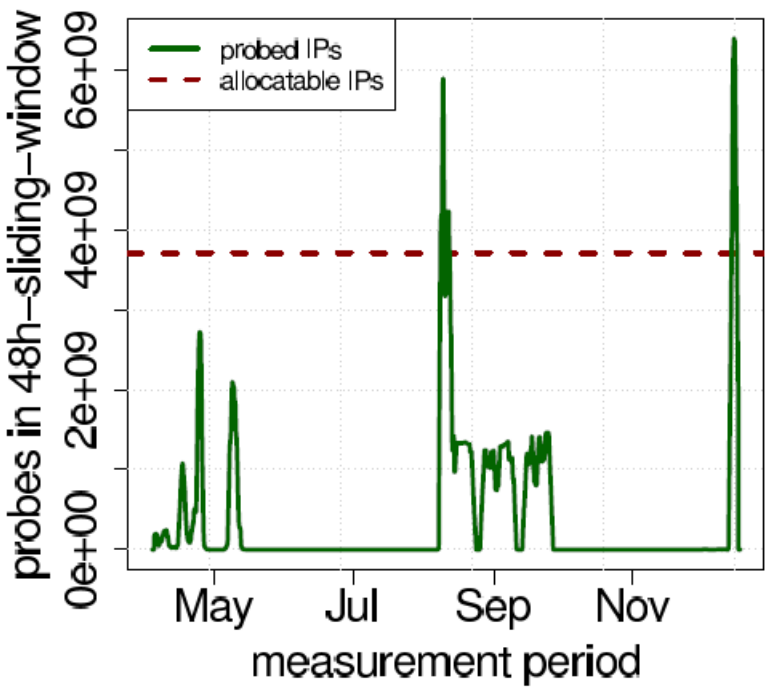


Figure 7: Finding *fast scans*: Sum of probes in 48h-sliding window over entire measurement period.

Real Size Of The Internet?

Not all reverse DNS entries are actually used.

Mixing incoherent measurement periods.

Number of IPs “in use” not necessarily equal to the size of the Internet.

IPs who do not respond to probes not necessarily “unused”.

Open Questions

Botnets are powerful, but illegal.

Can they still be used for good?

Using Botnets (even for research purposes) is unethical.

Should the data be used?

Questions?

What do you think?

Botnets are powerful, but illegal.

Can they still be used for good?

Using Botnets (even for research purposes) is unethical.

Should the data be used?