ZMap
Fast Internet-Wide Scanning and its Security Applications

Zakir Durumeric   Eric Wustrow   J. Alex Halderman

University of Michigan
Internet-Wide Network Studies

Previous research has shown promise of Internet-wide surveys

Mining Ps and Qs: Widespread weak keys in network devices (2012)

EFF SSL Observatory: A glimpse at the CA ecosystem (2010)

Census and Survey of the Visible Internet (2008)
Internet-Wide Network Studies

Previous research has shown promise of Internet-wide surveys

Mining Ps and Qs: Widespread weak keys in network devices (2012)

25 hours across 25 Amazon EC2 Instances (625 CPU-hours)

EFF SSL Observatory: A glimpse at the CA ecosystem (2010)

3 months on 3 Linux desktop machines (6500 CPU-hours)

Census and Survey of the Visible Internet (2008)

3 months to complete ICMP census (2200 CPU-hours)
ZMap: Fast Internet-Wide Scanning and its Security Applications
What if...?

What if Internet surveys didn’t require heroic effort?

What if we could scan the HTTPS ecosystem every day?

What if we wrote a whole-Internet scanner from scratch?
Introducing ZMap

an open-source tool that can port scan the entire IPv4 address space from just one machine in under 45 minutes with 98% coverage.

With Zmap, an Internet-wide TCP SYN scan on port 443 is as easy as:

```
$ zmap -p 443 -o results.txt
34,132,693 listening hosts
(took 44m12s)
```

97% of gigabit Ethernet linespeed
Talk Roadmap

1. **Philosophy and Architecture of ZMap**
2. Characterizing ZMap's Performance
3. Applications of High Speed Scanning
4. Scanning and Good Internet Citizenship
# ZMap Architecture

## Existing Network Scanners

- Reduce state by scanning in batches
  - Time lost due to blocking
  - Results lost due to timeouts

- Track individual hosts and retransmit
  - Most hosts will not respond

- Avoid flooding through timing
  - Time lost waiting

- Utilize existing OS network stack
  - Not optimized for immense number of connections

## ZMap

- Eliminate local per-connection state
  - Fully asynchronous components
  - No blocking except for network

- Shotgun Scanning Approach
  - Always send $n$ probes per host

- Scan widely dispersed targets
  - Send as fast as network allows

- Probe-optimized Network Stack
  - Bypass inefficiencies by generating Ethernet frames

---

*ZMap: Fast Internet-Wide Scanning and its Security Applications*
Addressing Probes

How do we randomly scan addresses without excessive state?

1. Scan hosts according to random permutation
2. Iterate over multiplicative group of integers modulo $p$

Negligible State

1. Primitive Root
2. Current Location
3. First Address
Validating Responses

How do we validate responses without local per-target state?

Encode secrets into mutable fields of probe packets that will have recognizable effect on responses
Validating Responses

How do we validate responses without local per-target state?

Encode secrets into mutable fields of probe packets that will have recognizable effect on responses.

---

**Ethernet**
- Receiver MAC address
- Sender MAC address
- Length
- Data

---

**IP**
- V
- IHL
- ... (Variable)
- Sender IP address
- Receiver IP address
- Data

---

**TCP**
- Sender port
- Receiver port
- Sequence number
- Ack. number
- ... (Variable)
- Data
Validating Responses

How do we validate responses without local per-target state?

Encode secrets into mutable fields of probe packets that will have recognizable effect on responses.

**Diagram:**

- **Ethernet**:
  - receiver MAC address
  - sender MAC address
  - length
  - data

- **IP**:
  - V
  - IHL
  - ... sender IP address
  - receiver IP address
  - data

- **TCP**:
  - sender port
  - receiver port
  - sequence number
  - ack. number
  - ... data
Packet Transmission and Receipt

How do we make processing probes easy and fast?

1. **ZMap framework** handles the hard work

2. **Probe modules** fill in packet details, interpret responses

3. **Output modules** allow follow-up or further processing
Talk Roadmap

1. Philosophy and Architecture of ZMap
2. Characterizing ZMap's Performance
3. Applications of High Speed Scanning
4. Scanning and Good Internet Citizenship
Scan Rate

How fast is too fast?

No correlation between hit-rate and scan-rate.

Slower scanning does not reveal additional hosts.
Coverage

Is one probe packet sufficient?

We expect an eventual plateau in responsive hosts, regardless of additional probes.

Scan Coverage

1 Packet: 97.9%
2 Packets: 98.8%
3 Packets: 99.4%
## Comparison with Nmap

Averages for scanning 1 million random hosts

<table>
<thead>
<tr>
<th></th>
<th>Normalized Coverage</th>
<th>Duration (mm:ss)</th>
<th>Est. Internet Wide Scan</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nmap (1 probe)</strong></td>
<td>81.4%</td>
<td>24:12</td>
<td>62.5 days</td>
</tr>
<tr>
<td><strong>Nmap (2 probes)</strong></td>
<td>97.8%</td>
<td>45:03</td>
<td>116.3 days</td>
</tr>
<tr>
<td><strong>ZMap (1 probe)</strong></td>
<td>98.7%</td>
<td>00:10</td>
<td>1:09:35</td>
</tr>
<tr>
<td><strong>ZMap (2 probes)</strong></td>
<td>100.0%</td>
<td>00:11</td>
<td>2:12:35</td>
</tr>
</tbody>
</table>

ZMap is capable of scanning more than 1300 times faster than the most aggressive Nmap default configuration ("insane")

Surprisingly, ZMap also finds more results than Nmap
 Probe Response Times

Why does ZMap find more hosts than Nmap?

Response Times

- 250 ms: < 85%
- 500 ms: 98.2%
- 1.0 s: 99.0%
- 8.2 s: 99.9%

Statelessness leads to both higher performance and increased coverage.
Talk Roadmap

1. Philosophy and Architecture of ZMap
2. Characterizing ZMap's Performance
3. Applications of High Speed Scanning
4. Scanning and Good Internet Citizenship
Visibility into Distributed Systems

Gaining near real-time perspective into the CA ecosystem

ZMap enables us to scan the public HTTPS Ecosystem every day

Completed 110 scans of the HTTPS ecosystem in the last year

We collected more than **42 million unique certificates** of which **6.9 million were browser trusted**. Identified 2 sets of misissued CA certificates.
Tracking Protocol Adoption

Examining the growth in global HTTPS adoption

June 2012–May 2013

- 10% increase in HTTPS servers.
- 23% increase in use on Alexa Top-1M sites.
- 11% increase in browser-trusted certificates.

ZMap: Fast Internet-Wide Scanning and its Security Applications
Enumerating Vulnerable Hosts

Discovering UPnP Vulnerabilities En Masse

HD Moore disclosed vulnerabilities in several common UPnP frameworks in January 2013.

Under 6 hours to code and run UPnP discovery scan. Custom probe module, 150 SLOC.

We found that 3.34 M of 15.7 M devices were vulnerable.

Compromise possible with a single UDP packet!
Uncovering Hidden Services

Enumerating Unadvertised Tor Bridges

Scanning has potential to uncover unadvertised services

We perform a Tor handshake with public IPv4 addresses on port 9001 and 443

We identified 86% of live allocated bridges with a single scan

Tor has developed obfsproxy that listens on random ports to count this type of attack
Further ZMap Potential

Further Potential Applications

Detect Service Disruptions
Track Adoption of Defenses
Study Criminal Behavior

Other Security Implications

Anonymous Communication
Track users between IP leases

Snapshot of HTTPS outages caused by Hurricane Sandy
Talk Roadmap

1. Philosophy and Architecture of ZMap
2. Characterizing ZMap's Performance
3. Applications of High Speed Scanning
4. Scanning and Good Internet Citizenship
Ethics of Active Scanning

Considerations

Impossible to request permission from all owners
No IP-level equivalent to robots exclusion standard
Administrators may believe that they are under attack

Reducing Scan Impact

Scan in random order to avoid overwhelming networks
Signal benign nature over HTTP and w/ DNS hostnames
Honor all requests to be excluded from future scans
User Responses

Responses from 145 users

Blacklisted 91 entities
(3.7 M total addresses)

15 hostile responses

2 cases of retaliatory traffic

<table>
<thead>
<tr>
<th>Entity Type</th>
<th>Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Business</td>
<td>41</td>
</tr>
<tr>
<td>Home User</td>
<td>38</td>
</tr>
<tr>
<td>Corporation</td>
<td>17</td>
</tr>
<tr>
<td>Academic Institution</td>
<td>22</td>
</tr>
<tr>
<td>Government</td>
<td>15</td>
</tr>
<tr>
<td>ISP</td>
<td>2</td>
</tr>
<tr>
<td>Unknown</td>
<td>10</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>145</strong></td>
</tr>
</tbody>
</table>
Future Work

10gigE Network Surveys
TLS Server Name Indication
Scanning Exclusion Standards
IPv6 Scanning Methodology?

Use ZMap to do great research!
Public Release

Releasing ZMap as a fully documented open source project

Downloaded it now from https://zmap.io

Scanning the Internet really is as simple as:

$ zmap -p 443 -o results.txt

Be sure you have adequate bandwidth and be a good Internet neighbor!
Conclusion

Living in a unique period
IPv4 can be quickly, exhaustively scanned
IPv6 has not yet been widely deployed

ZMap lowers barriers of entry for Internet-wide surveys
Now possible to scan the entire IPv4 address space from one host in under 45 minutes with 98% coverage

Explored potential applications

Ultimately we hope ZMap enables future research