Zmap: Fast Internet-Wide Scanning and its Security Applications
Paper Overview

- Presents a framework and an implementation (ZMap) to scan the entire IPv4 space
- Evaluates the implementation by comparing against Nmap.
  - Speed
  - Accuracy
  - Coverage
- Demonstrates security application
- Recommends good internet citizenship rules for using Zmap
Paper Flavor

- Hybrid between a system paper and measurement paper
- Contribution is a framework that enables TIMELY internet-wide scanning
- The evaluation of the tool is a measurement to demonstrate utility
- Highlights important ethical and legal expectations for operating high-speed scanning

“With Great Power, Comes Great Responsibility”
- Uncle Ben
Purpose and Significance

▪ How does network scanning help us empirically study internet security?

▪ Network scanners were introduced over 25 years ago, What is so special about ZMap?
  • Timely scanning 1300x faster
  • More scans (real-time)
  • Finer granularity
  • User friendly/modular
Purpose and Significance

- How does ZMap improve over prior scanners?
- What is some disadvantages in comparison with traditional scanner?
  - Less granular
  - Single port scanning
  - No retry (fire and forget)
  - More congestion due to speed
  - Async coordination is complicated
ZMap Architecture

- Three important components
  - Configuration
  - Packet generation
  - Result validation and aggregation

- Why does ZMap implement random IP selection?
  - Avoid overwhelming target
  - Avoiding being blocked
  - Representative sampling

- Would you adapt a different strategy?
  - Cloud Providers?
  - Residential?
  - Small amount of information shared between scanner instances
  - Spreading out load for network blocks
ZMap Architecture

- Support of IP address exclusion, why do you want to exclude network blocks?
  - Reserved/Non-Internet routable IPs
  - Users requests
  - Exclusion of sensitive networks

- How would you add additional enforcement to ensure exclusion?
  - Routing rules/firewall rules
ZMap Architecture

- What are some challenges you anticipate in implementing packet generation and result validation for Zmap?
  - Async coordinate
  - Multiple probe

- How do you think the authors chose an 8 sec wait time for probing?
  - Perhaps based on the timeout of each probe
The probe module uses TCP scanning, how does that differ from the TCP 3-hand-shake?

How much overhead do you think a 3-hand-shake will add to the current implementation?

Integrity check uses UMAC for each host, why do you think the authors chose UMAC over MAC or other hash algorithms?
- UMAC is efficiently calculated with 1 CPU cycle

UDP is application specific, whereas TCP is generic and provides stateful connection
Evaluation (Empirical Measurement)

▪ What is the metric hit rate vs scan rate?
  • Range of IP with end-host – measures the fastest scan rate 1.4M pps

▪ What other metric would you consider to evaluate efficacy of ZMap scanner?
Evaluation (Empirical Measurement)

- Is the coverage evaluation convincing? What problems do you see with the current results?
- Ground truth is difficult to obtain, do you find the current estimated ground truth robust for the purpose of this study?
Evaluation (Empirical Measurement)

- Authors discuss diurnal effect. How would you adjust your experiment to accommodate for the 3% variation in the results?
- Is 3% significant? Would you consider it problematic for your study? Would it affect the coverage results?
Evaluation (Empirical Measurement)

- Is the NMap vs ZMap comparison fair? What additional evaluation would you carry out?
- Is the SSL/TLS coverage comparison fair? What bias do you see?

<table>
<thead>
<tr>
<th>Scan Type</th>
<th>Coverage (normalized)</th>
<th>Duration (mm:ss)</th>
<th>Est. Time for Internet-wide Scan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nmap, max 2 probes (default)</td>
<td>0.978</td>
<td>45:03</td>
<td>116.3 days</td>
</tr>
<tr>
<td>Nmap, 1 probe</td>
<td>0.814</td>
<td>24:12</td>
<td>62.5 days</td>
</tr>
<tr>
<td>ZMap, 2 probes</td>
<td>1.000</td>
<td>00:11</td>
<td>2:12:35</td>
</tr>
<tr>
<td>ZMap, 1 probe (default)</td>
<td>0.987</td>
<td>00:10</td>
<td>1:09:45</td>
</tr>
</tbody>
</table>
Sec. Applications and Future Research

- How does ZMap enable security researchers to empirically document security problems in practice?
- What security applications stood out to you and you found to be interesting?
- How would you scan for UDP?
Sec. Applications and Future Research

- What security problem would you use ZMap for?
- How would you approach IPv6 with Zmap?
Sec. Applications and Future Research

- What recent work have you come across that uses ZMap to study a security problem?
- How would you approach IPv6 with Zmap?
Good Citizenship

- Why is it important to be a good citizen on the Internet?
- Legally, there are no laws against port scanning. If a net block owner requests exclusion but the net block is important to your study, how would you approach it?
- Would you amend any of the recommendations for good Internet citizenship?

1. Coordinate closely with local network admins to reduce risks and handle inquiries.
2. Verify that scans will not overwhelm the local network or upstream provider.
3. Signal the benign nature of the scans in web pages and DNS entries of the source addresses.
4. Clearly explain the purpose and scope of the scans in all communications.
5. Provide a simple means of opting out, and honor requests promptly.
6. Conduct scans no larger or more frequent than is necessary for research objectives.
7. Spread scan traffic over time or source addresses when feasible.
Good Citizenship

- One of ZMap’s security application is to find vulnerabilities, but Scanners can produce a very high false positive rate. How would validate the vulnerabilities?
- How would you design your measurement so that it does not affect production systems?
- How certain can you be your experiment design is safe?
Related Work

- Internet Census 2012 (Carna botnet)
- Masscan (2013)
  - [https://github.com/robertdavidgraham/masscan](https://github.com/robertdavidgraham/masscan)
- Nessus
  - [https://www.tenable.com/products/nessus](https://www.tenable.com/products/nessus)
- Censys
  - [https://censys.io](https://censys.io)
- Shodan
  - [https://www.shodan.io](https://www.shodan.io)
- ZoomEye
  - [https://www.zoomeye.org/about](https://www.zoomeye.org/about)
- Binary Edge
  - [https://www.binaryedge.io](https://www.binaryedge.io)