David Adrian, Zakir Durumeric, Gulshan Singh, J. Alex Halderman

University of Michigan

WOOT ’14
San Diego, CA
One Year Ago…

We released ZMap

ZMap is an Internet-wide port scanner capable of scanning at *97% the maximum theoretical speed* of gigabit Ethernet.

ZMap completes a single-port TCP SYN scan of all of IPv4 in forty-five minutes.
Networks are Faster
Our own got 10x faster!

1 GigE ~ 1.48 million packets per second

10 GigE ~ 14.88 million packets per second

Why not full 10 GigE?
Zippier ZMap

A series of performance enhancements to ZMap, enabling scanning at 95% 10 GigE linespeed, completing a single-port TCP scan in under five minutes.
Talk Roadmap

1. Optimizations to ZMap

2. Evaluation of scanning at >1 Gbps

3. Applications and Conclusions
Performance Enhancements

What do we need to optimize?

Parallelize address generation

Efficient blacklisting and whitelisting

Very low overhead sends (~200 cycle budget)
Address Generation

How do we address outgoing packets?

Multithreaded iteration over a cyclic group of integers modulo $p$ requires a lock

\[ a_{i+1} = g \cdot a_i \mod p \]
Address Generation

How do we address outgoing packets?

Multithreaded iteration over a cyclic group of integers modulo \( p \) requires a lock

Shard the cycle into disjoint sets

\[
a_{i+1} = g \cdot a_i \mod p
\]

\[
a_{i+n} = g^n \cdot a_i \mod p
\]
Address Constraints

Good Internet citizenship demands honoring blacklist requests

1100 entries from 208 organizations on our blacklist, 0.15% of IPv4 address space

Use blacklist to exclude IANA-reserved addresses, 14% of IPv4 address space
Optimized Address Constraints

Model IPv4 as a binary tree populated with blacklist
Paint leaf nodes as whitelisted or blacklisted

Use tree to determine number of allowed addresses $n$, and map indices $1 \ldots n$ to addresses $a_1 \ldots a_n$
Optimized Address Constraints

Can we avoid the tree lookup?

Move the whitelisted /20 blocks out of the tree and into an array to bypass tree lookup.
Zero-Copy NIC Access
How can we send packets at line rate?

The Linux kernel is not capable of sending 64 byte packets at 10 GigE linespeed – 14.88 million packets per second

Use the PF_RING ZC library for direct NIC “zero-copy” access to reach linespeed

Bypass the kernel to reach 10 GigE linespeed
Zero-Copy NIC Access
How do we combine sharding with PF_RING?

Old Architecture

New Architecture

Global Cyclic Group Iterator

Blocking Update

Send

Send

Send

Packet Creation

Packet Creation

Packet Creation

Nonblocking Poll

Send
Talk Roadmap

1. Performance Enhancements to ZMap

2. Evaluation of scanning at >1 Gbps

3. Applications and Conclusions
10 GigE is Fast
Your mileage may vary.

This is as much a stress-test of the University of Michigan’s network as it is a study of ZMap

Building uplink is an aggregated 2x10 gigabit fiber channel

Performance may vary on other networks.
Complete Scans
How fast can we complete full scans of the Internet?

<table>
<thead>
<tr>
<th>Scan Rate</th>
<th>Duration</th>
<th>Normalized Hit Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.44 Mpps (~1 Gbps)</td>
<td>42:08</td>
<td>1.00</td>
</tr>
<tr>
<td>3.00 Mpps</td>
<td>20:47</td>
<td>0.99</td>
</tr>
<tr>
<td>4.00 Mpps</td>
<td>15:38</td>
<td>0.97</td>
</tr>
<tr>
<td><strong>14.23 Mpps (~10 Gbps)</strong></td>
<td><strong>4:29</strong></td>
<td><strong>0.63</strong></td>
</tr>
</tbody>
</table>

95% 10 GigE linespeed

Complete scans of port 443 with our enhancements and blacklist

37% Drop
Hit Rate vs. Scan Rate
When does fast become too fast?

50 second long scans of random samples of IPv4 address space on port 443
Receive Rate
Where are the packets going?

SYN ACK receive rate for 50s sample scans

Split send and receive between two machines

Packets get dropped on the network
Talk Roadmap

1. Performance Enhancements to ZMap

2. Evaluation of scanning at >1 Gbps

3. Applications and Conclusions
Applications

What can we gain from 10 GigE scanning?

Decrease the moving camera effect during Internet-wide scans

Faster multi-packet scanning-related applications

Large scale vulnerability detection and exploitation
Conclusion

As faster network infrastructure becomes available, scanning at 10 Gbps will enable powerful new applications for attackers and defenders alike.
Zippier ZMap

https://zmap.io

https://github.com/zmap

@davidcadrian

David Adrian, Zakir Durumeric, Gulshan Singh, J. Alex Halderman

zippier-team@umich.edu

University of Michigan
Backup Slides
Masscan
How are we different?

8-25 Mpps using dual 10 GigE ports

Did not have facilities to perform live network tests faster than 100,000 pps

Masscan peaked at 6.4 Mpps on our machines in a single-port configuration
Hit Rate vs. Scan Rate
When does fast become too fast?

![Graph showing Hit Rate vs. Speed (Mpps)]
45 MINUTES TO SCAN THE INTERNET?

AIN'T NOBODY GOT TIME FOR THAT