Zippier ZMap

Internet-Wide Scanning at 10Gbps

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

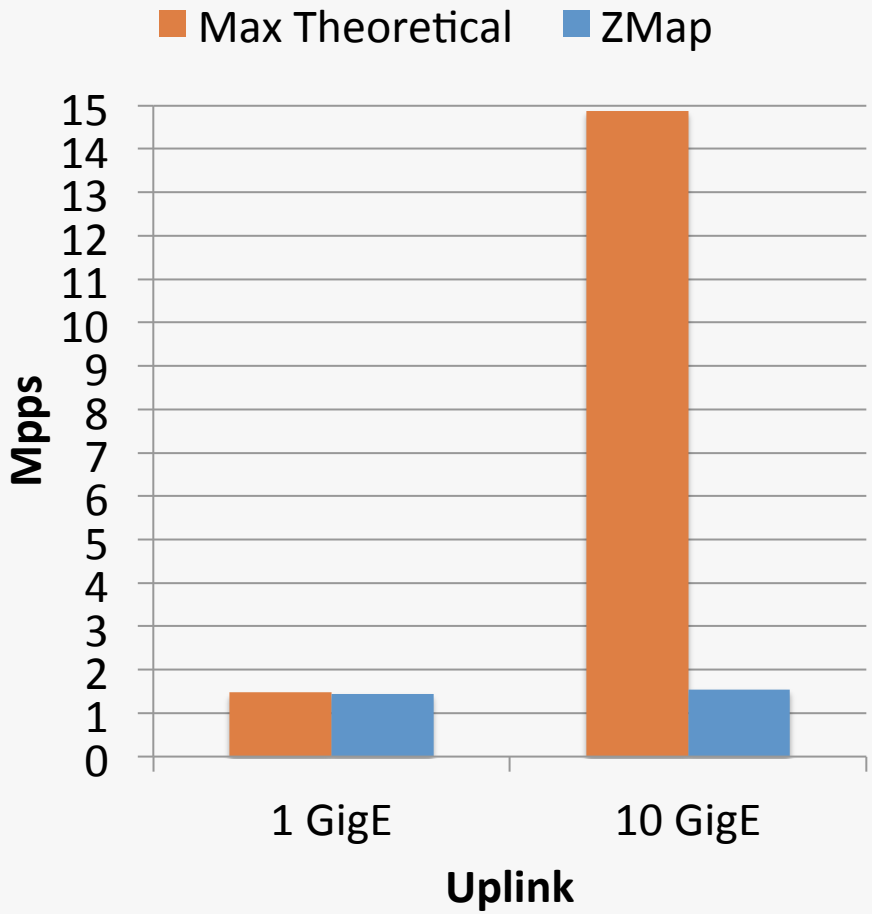
ZMap completes a single-port TCP SYN scan of all of IPv4 in forty-five minutes

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

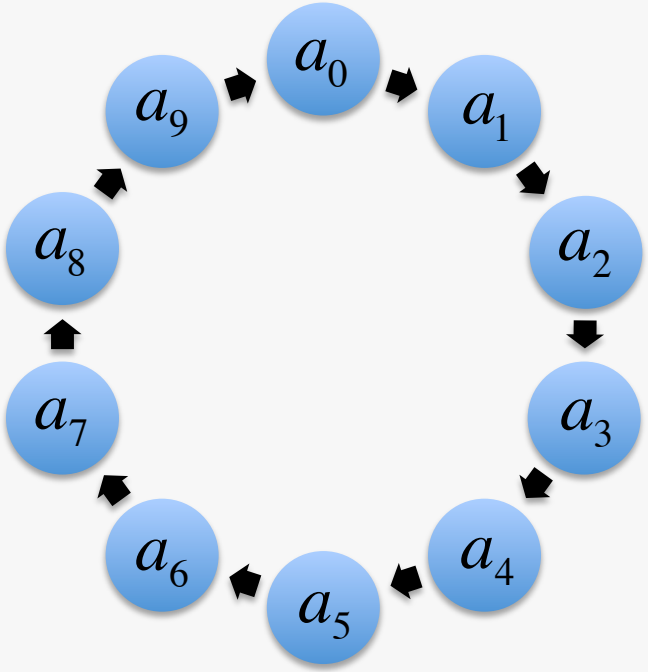
Efficient blacklisting and whitelisting

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

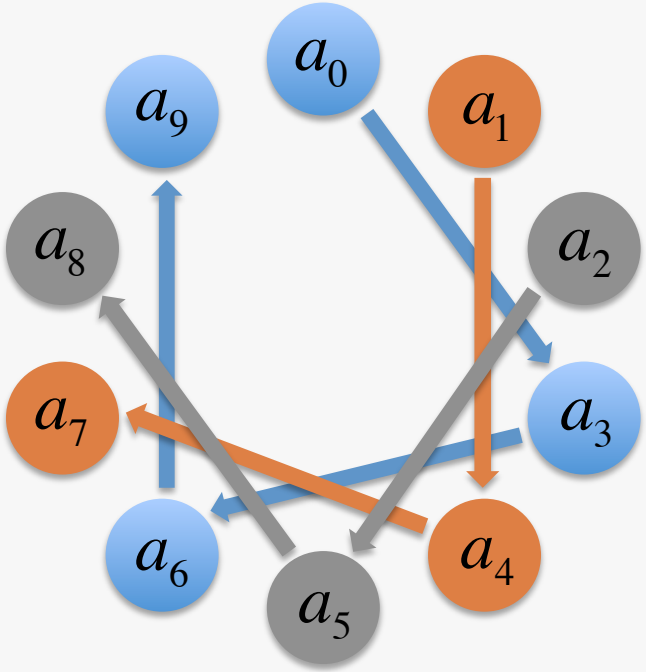


$$a_{i+1} = g \cdot a_i \bmod p$$

Address Generation

How do we address outgoing packets?

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



$$a_{i+1} = g \cdot a_i \bmod p$$

$$a_{i+n} = g^n \cdot a_i \bmod p$$

Shard the cycle into disjoint sets

Address Constraints

Good Internet citizenship demands honoring blacklist requests

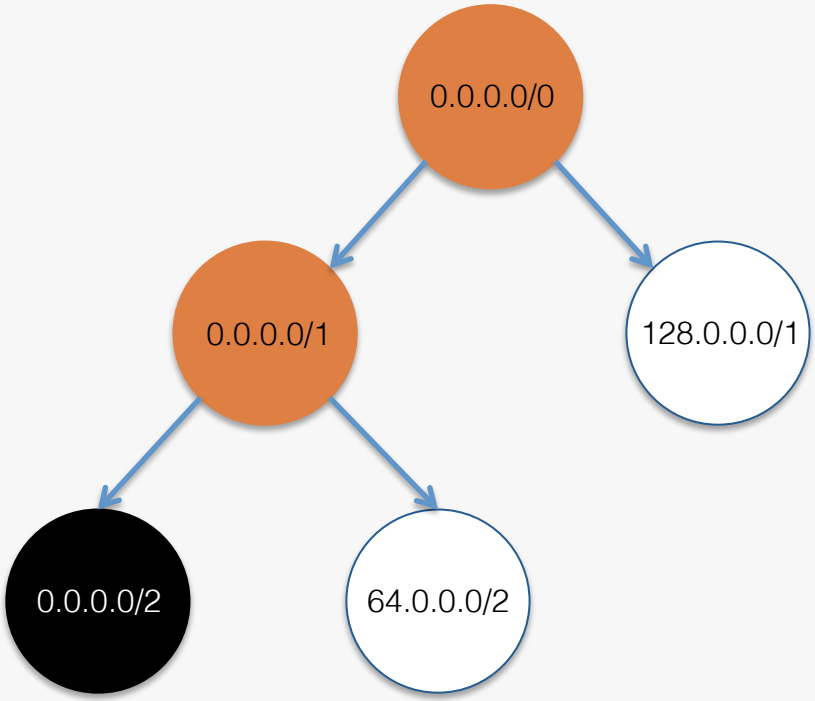
1100 entries from 208 organizations on our blacklist, 0.15% of IPv4 address space

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

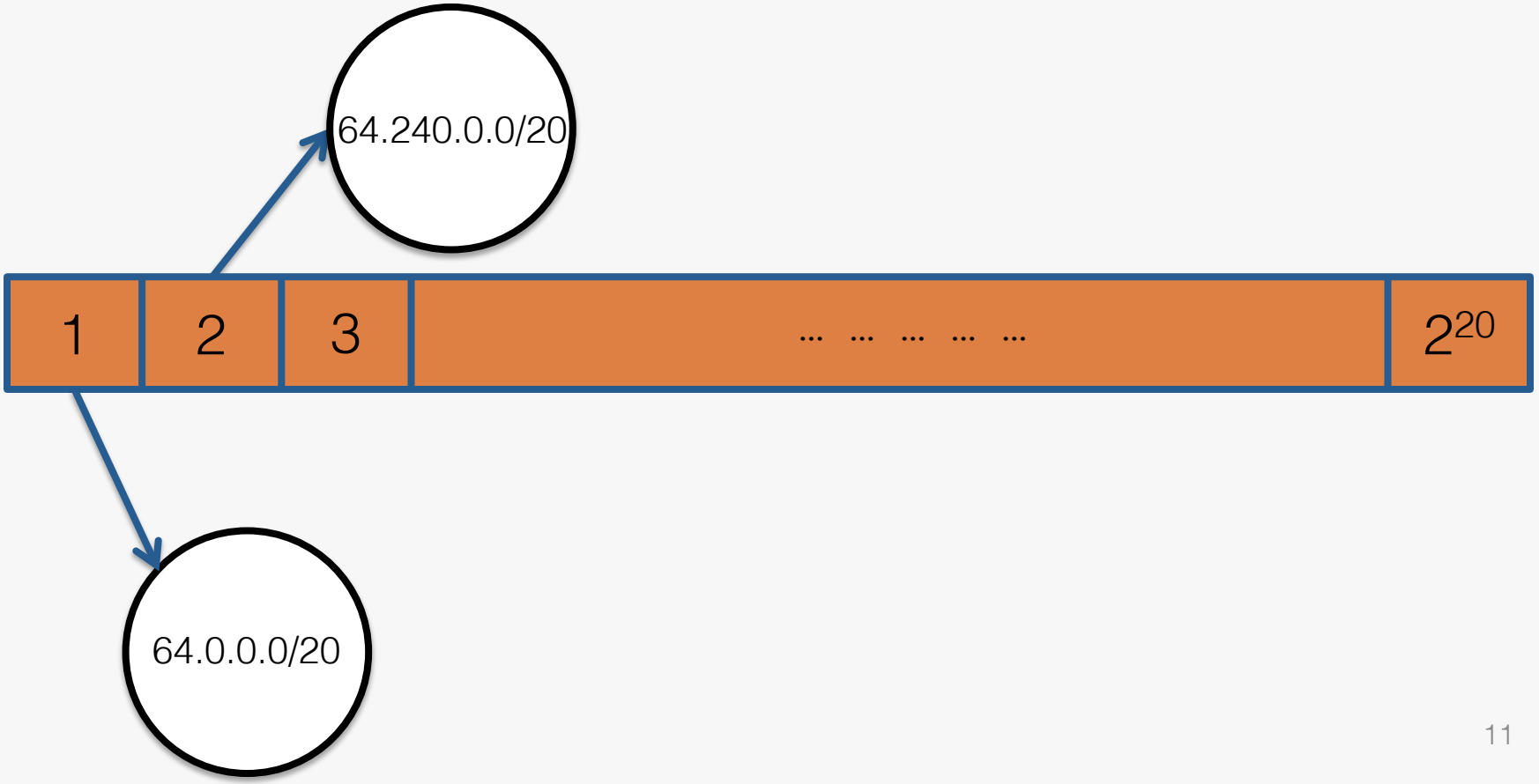


Use tree to determine number of allowed addresses *n*, and map indices *1…n* to addresses $a_1 \ldots a_n$

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?

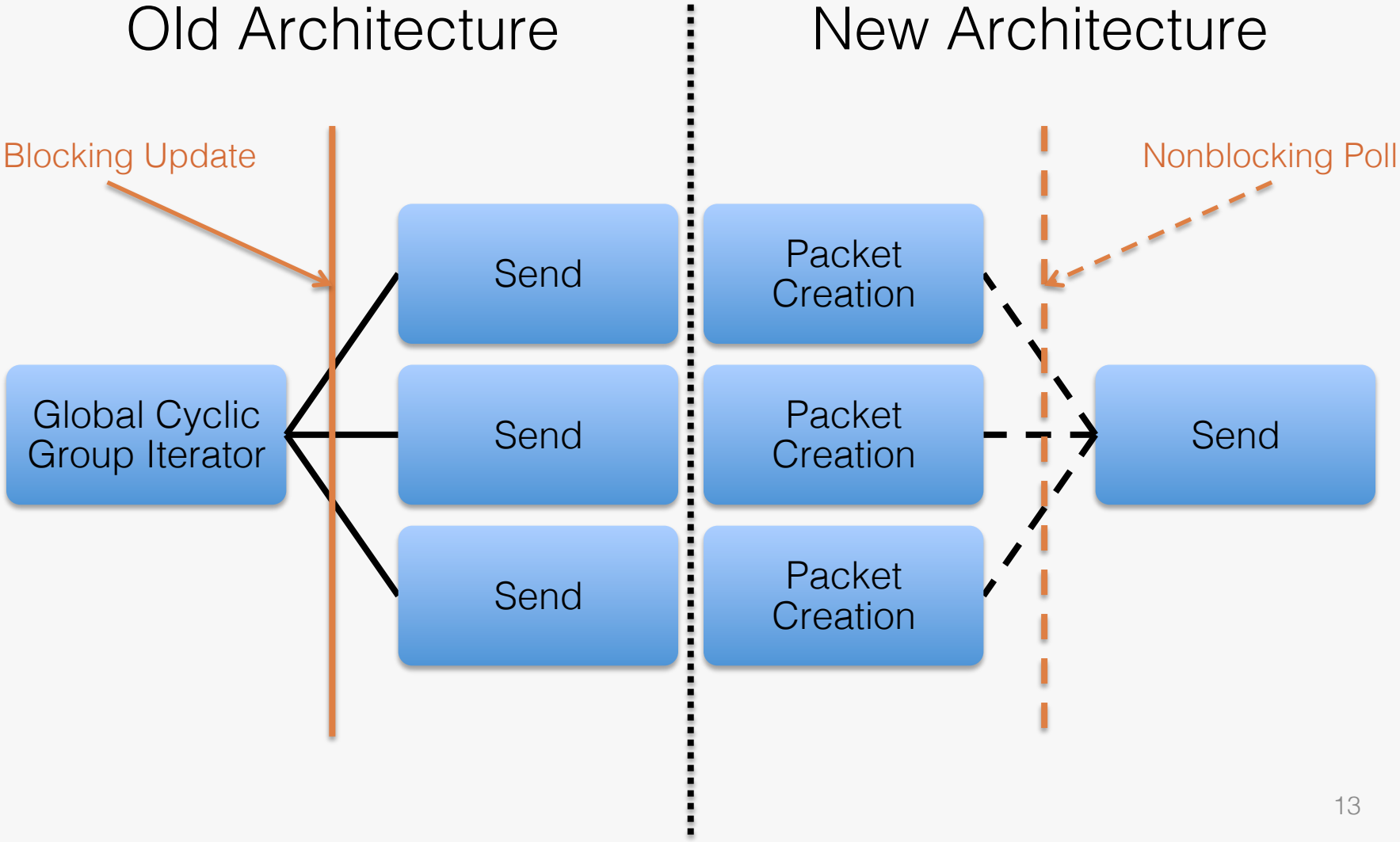
The Linux kernel is not capable of sending 64 byte packets at 10 GigE linespeed – 14.88 million packets per second

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?



Talk Roadmap

1. Performance Enhancements to ZMap
2. Evaluation of scanning at >1 Gbps
3. Applications and Conclusions

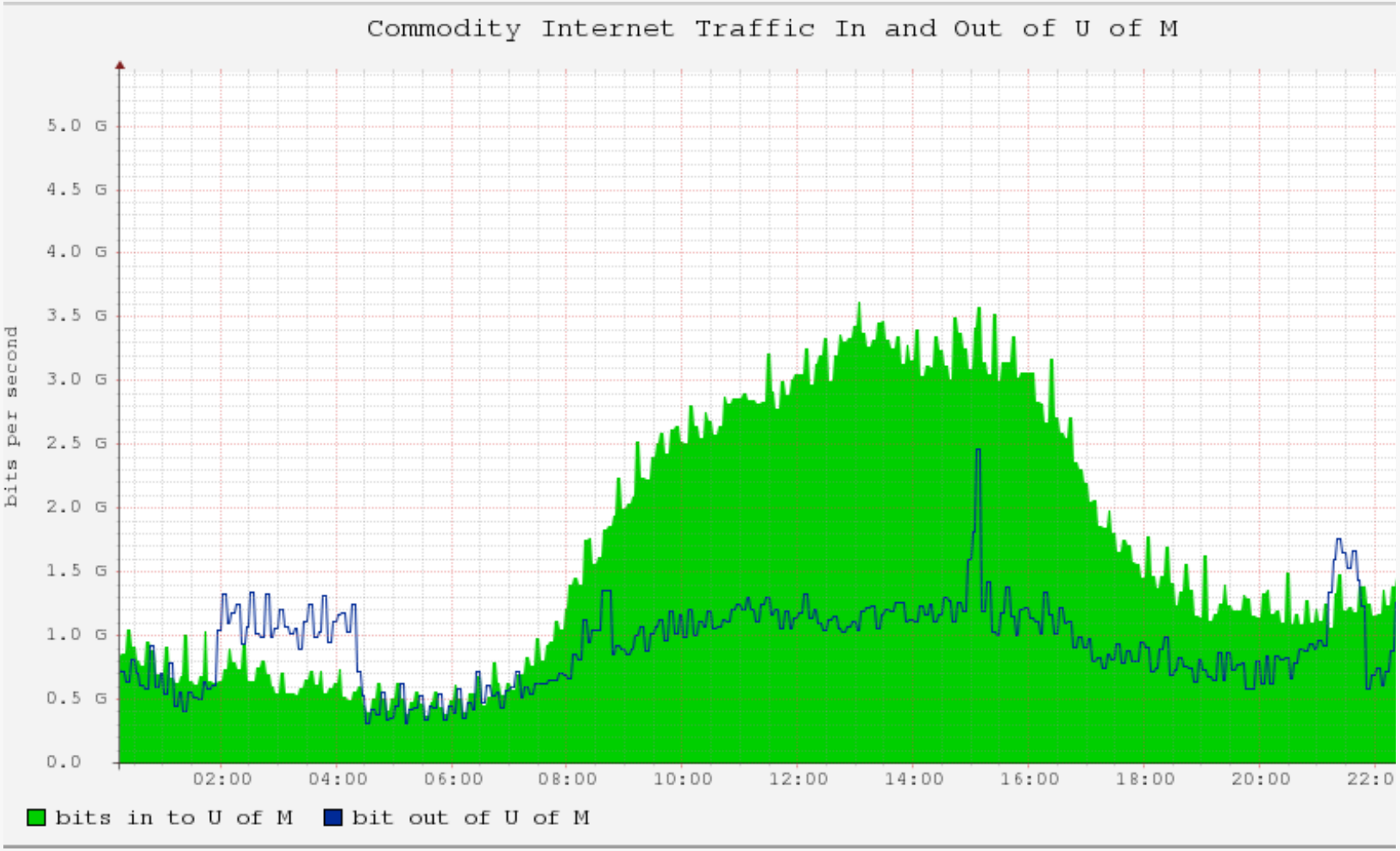
10 GigE is Fast

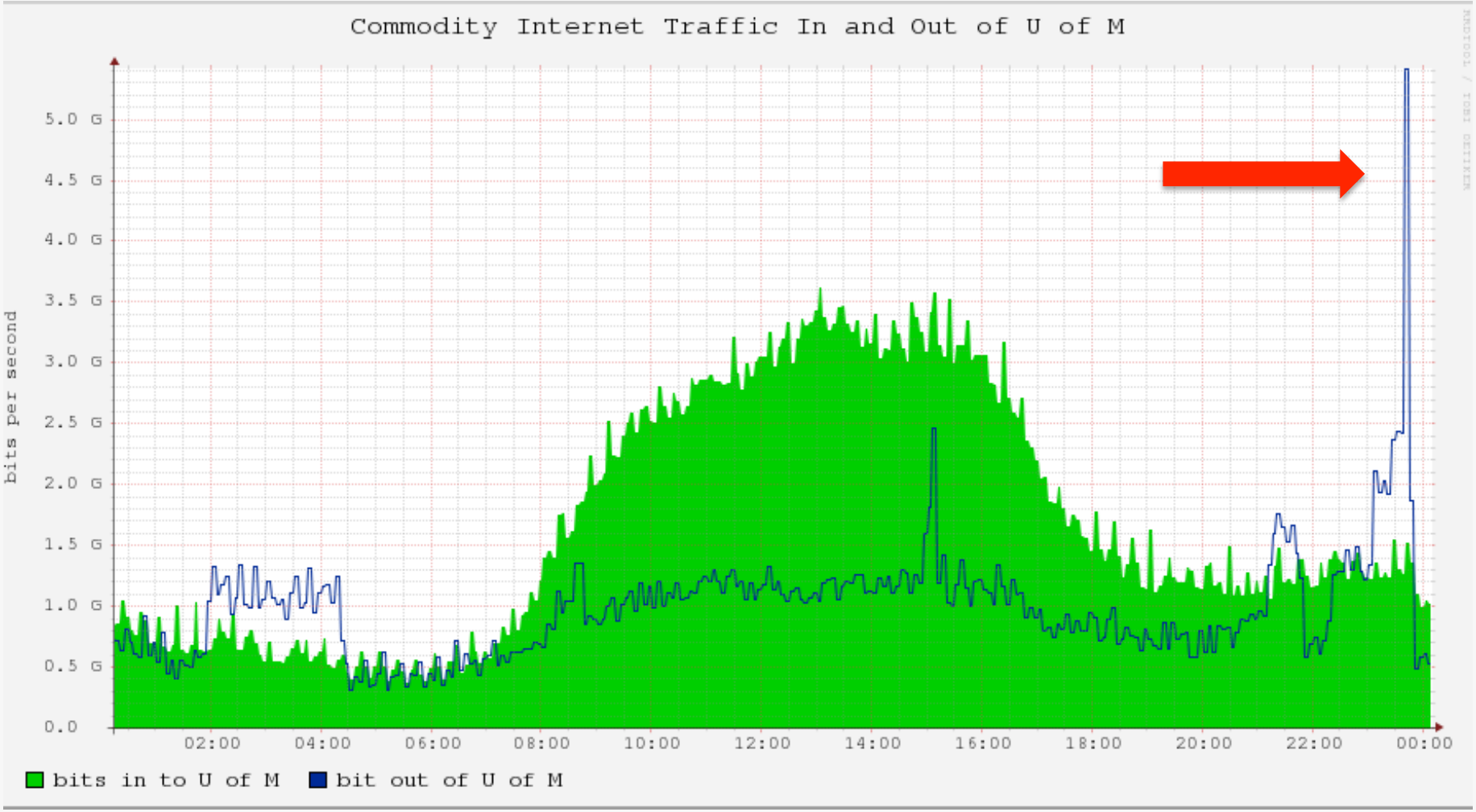
Your mileage may vary.

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.


Commodity Internet Traffic In and Out of U of M
bits per second
5.0 G
4.5 G
4.0 G
3.5 G
3.0 G
2.5 G
2.0 G
1.5 G
1.0 G
0.5 G
0.0
02:00
04:00
06:00
08:00
10:00
12:00
14:00
16:00
18:00
20:00
22:0
bits in to U of M
bit out of U of M


Commodity Internet Traffic In and Out of U of M
bits per second
5.0 G
4.5 G
4.0 G
3.5 G
3.0 G
2.5 G
2.0 G
1.5 G
1.0 G
0.5 G
0.0
02:00
04:00
06:00
08:00
10:00
12:00
14:00
16:00
18:00
20:00
22:00
00:00
bits in to U of M
bit out of U of M

Complete Scans

How fast can we complete full scans of the Internet?

Scan Rate	Duration	Normalized Hit Rate
1.44 Mpps (~1 Gbps)	42:08	1.00
3.00 Mpps	20:47	0.99
4.00 Mpps	15:38	0.97
14.23 Mpps (~10 Gbps)	**4:29**	**0.63**

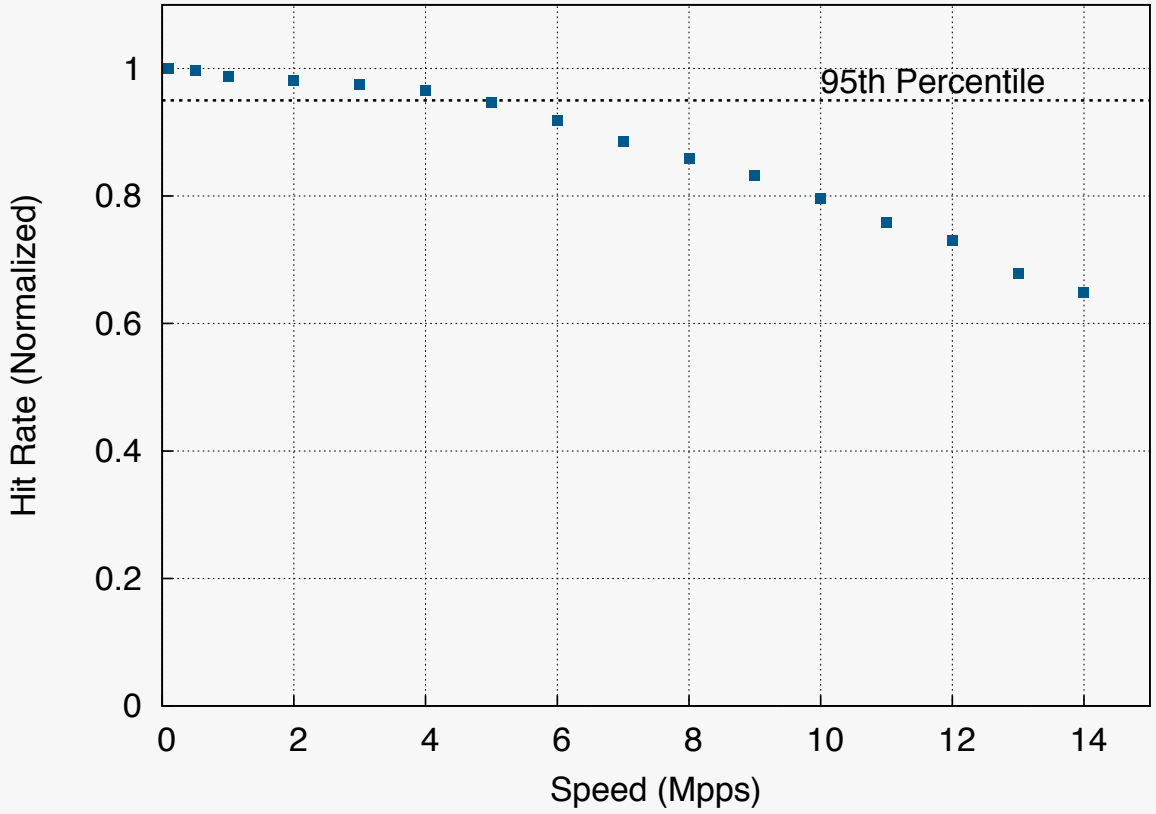
95% 10 GigE linespeed

37% Drop

Complete scans of port 443 with our enhancements and blacklist

Hit Rate vs. Scan Rate

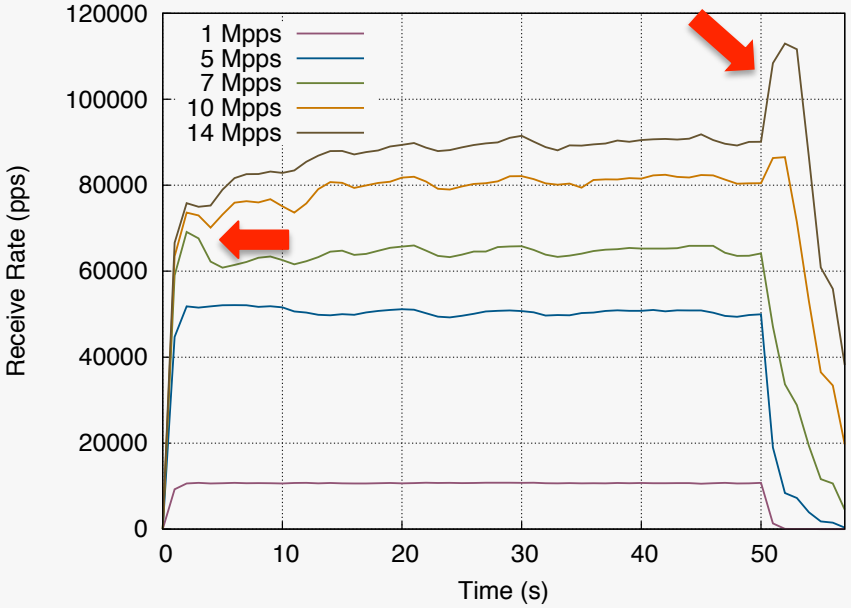
When does fast become too fast?



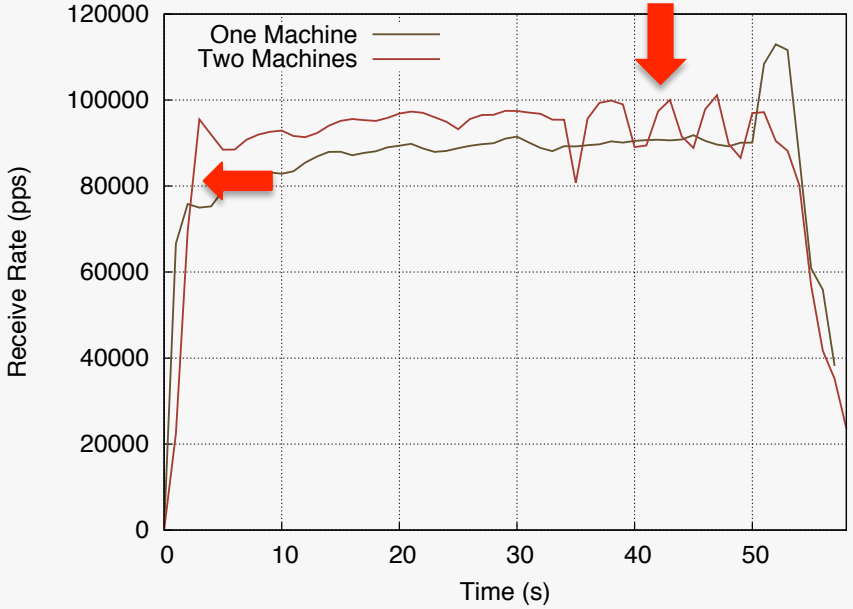
50 second long scans of random samples of IPv4 address space on port 443

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?

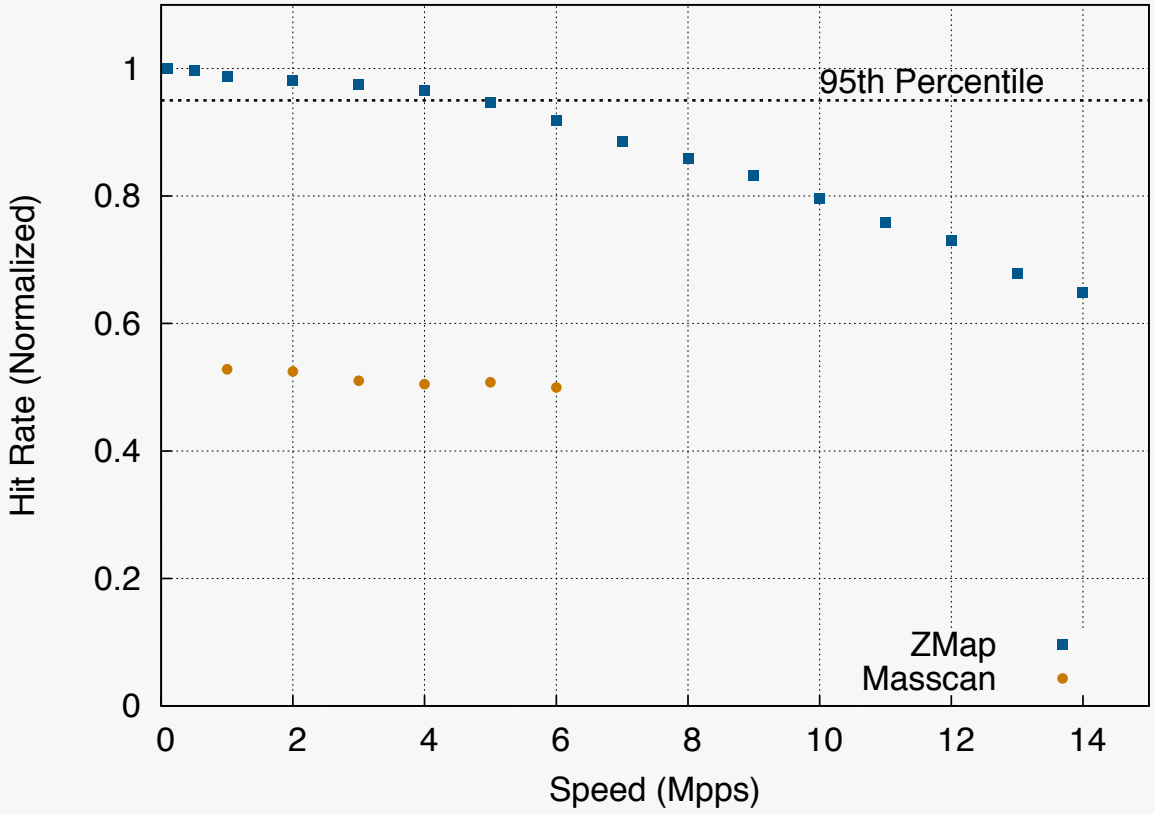
8-25 Mpps using dual 10 GigE ports

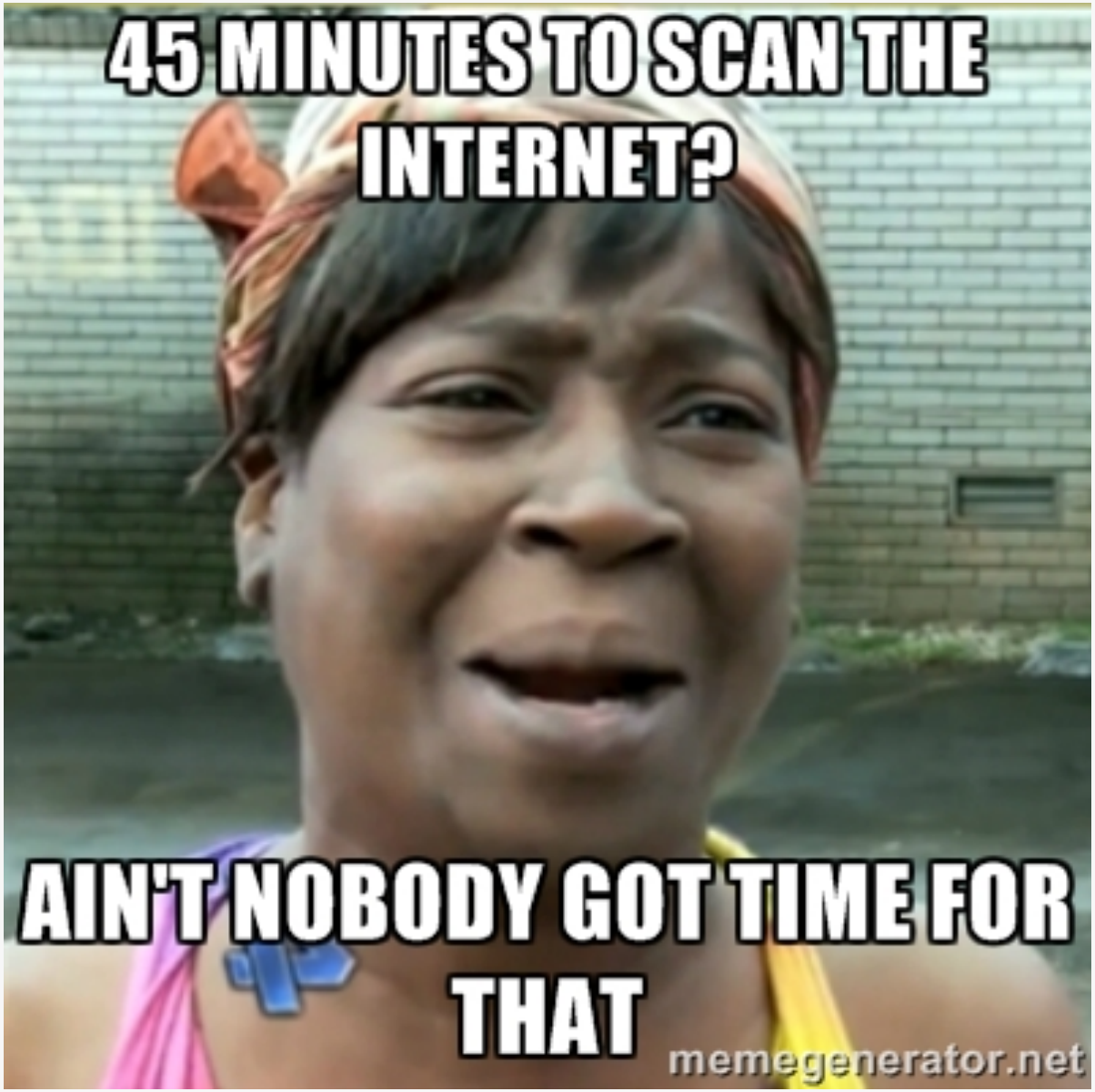
Did not have facilities to perform live network tests faster than 100,000 pps

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?




45 MINUTES TO SCAN THE INTERNET?
AIN'T NOBODY GOT TIME FOR THAT
memegenerator.net