On Understanding the Internet Via Edge Measurement

May 14, 2015

Matt Sargent

Advisor: Mark Allman

CASE SCHOOL OF ENGINEERING
CASE WESTERN RESERVE UNIVERSITY

Introduction

- "Smart" edge vs. "Dumb" core
  - Logic for connections pushed to edges
  - Core networks properly route packets

- Core has gained functionality (slowly)
  - Edge responsible for rapid evolution

Introduction

- Empirical measurement keeps understanding of network properties up-to-date
- Measurement challenges mental models
  - E.g., packet reordering
  - E.g., session arrival times

Introduction

- Leverage empirical measurement to study edge-driven shifts
  - Available bandwidth
  - Transport protocols
  - Policy and security threats
- Presenting a subset of results

Available Bandwidth

Fiber-To-The-Home Traffic:

Characterization and Performance

Motivation

- Last mile bandwidth has leapfrogged past current content offerings
  - E.g., Google Fiber, municipal fiber
- What will users do with significantly higher capacity?
- Are protocols up to the task of utilizing significantly higher bandwidth?

Data

- Observe traffic in a Fiber-To-The-Home network, the Case Connection Zone (CCZ)
  - ~90 homes with bi-directional 1 Gbps
- Use Bro IDS to continuously collect data
- Collect packet traces one week per month

Result 1

Users behave similar to residential users with significantly less bandwidth

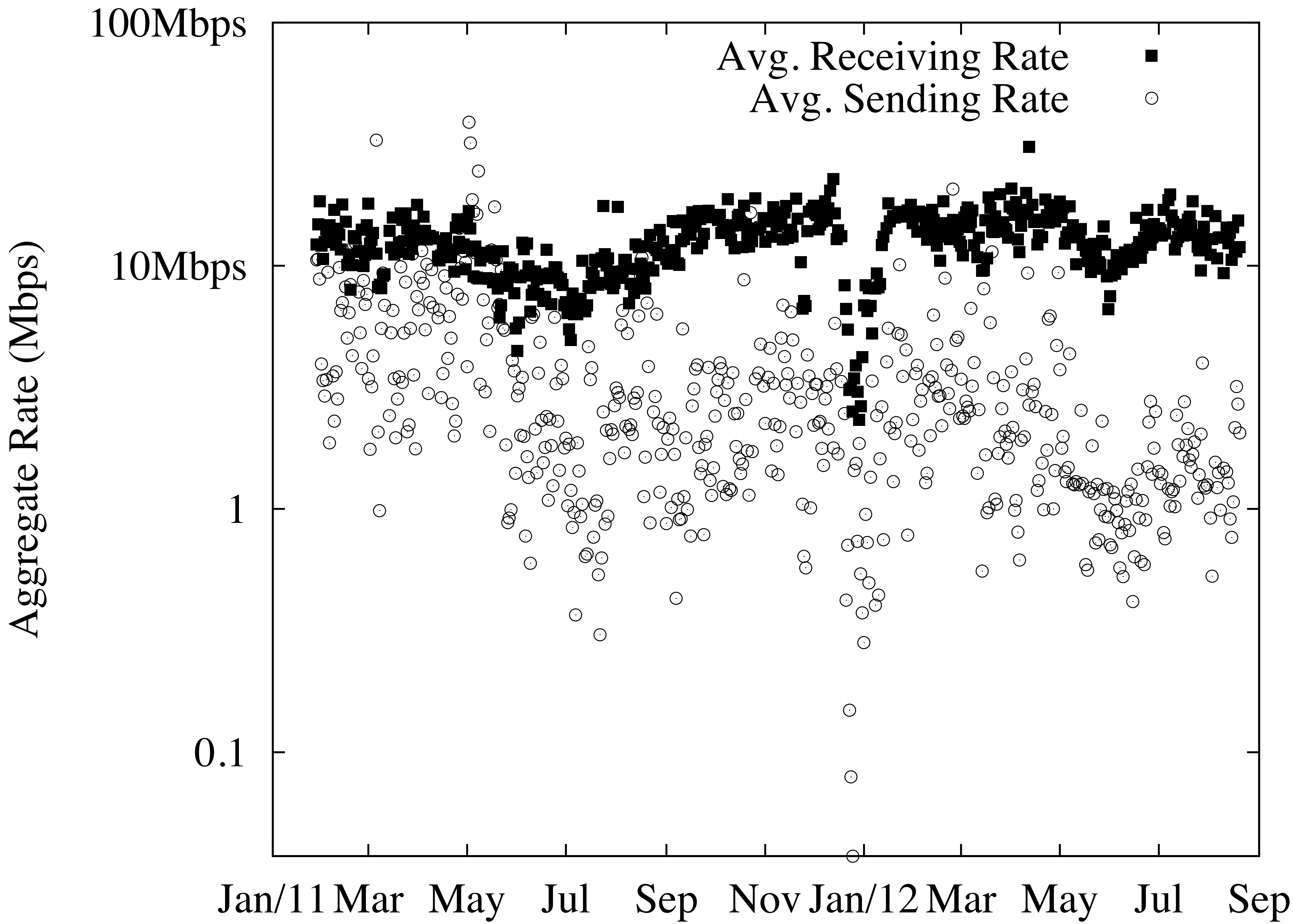
Result 1 - Traffic Mix

Service	Hosts	Conns.	Sent	Rcvd.
HTTP	90	321 M	1.1 TB	62 TB
Flash	89	444 K	6.0 GB	4.5 TB
BitTorrent	72	28 M	9.7 TB	3.4 TB
HTTPS	90	52 M	776 GB	1.9 TB
Steam	65	442 K	176 MB	819 GB
DNS	90	255 M	11.2 GB	63.7 GB
Other-39457	25	956 K	290 GB	45.3 GB
Other-1111	30	1.4 M	776 GB	40.1 GB
Other-31690	33	166 K	293 GB	23.6 GB
Minecraft	27	6.2 M	353 GB	7.7 GB
Unclassified	88	92.8 M	8.1 TB	5.0 TB
	98%	12%	38%	6%

Result 2

Even with essentially unlimited bandwidth, connection performance is low

Result 2 - Aggregate Sending Rates



Result 2 - Fast Sending

- For 99% of the time users send data under a rate of 0.5 Mbps
- For 99% of the time users receive data under a rate of 3.2 Mbps

- Each day, a user averages just over 1 minute of receiving at a rate of at least 10 Mbps

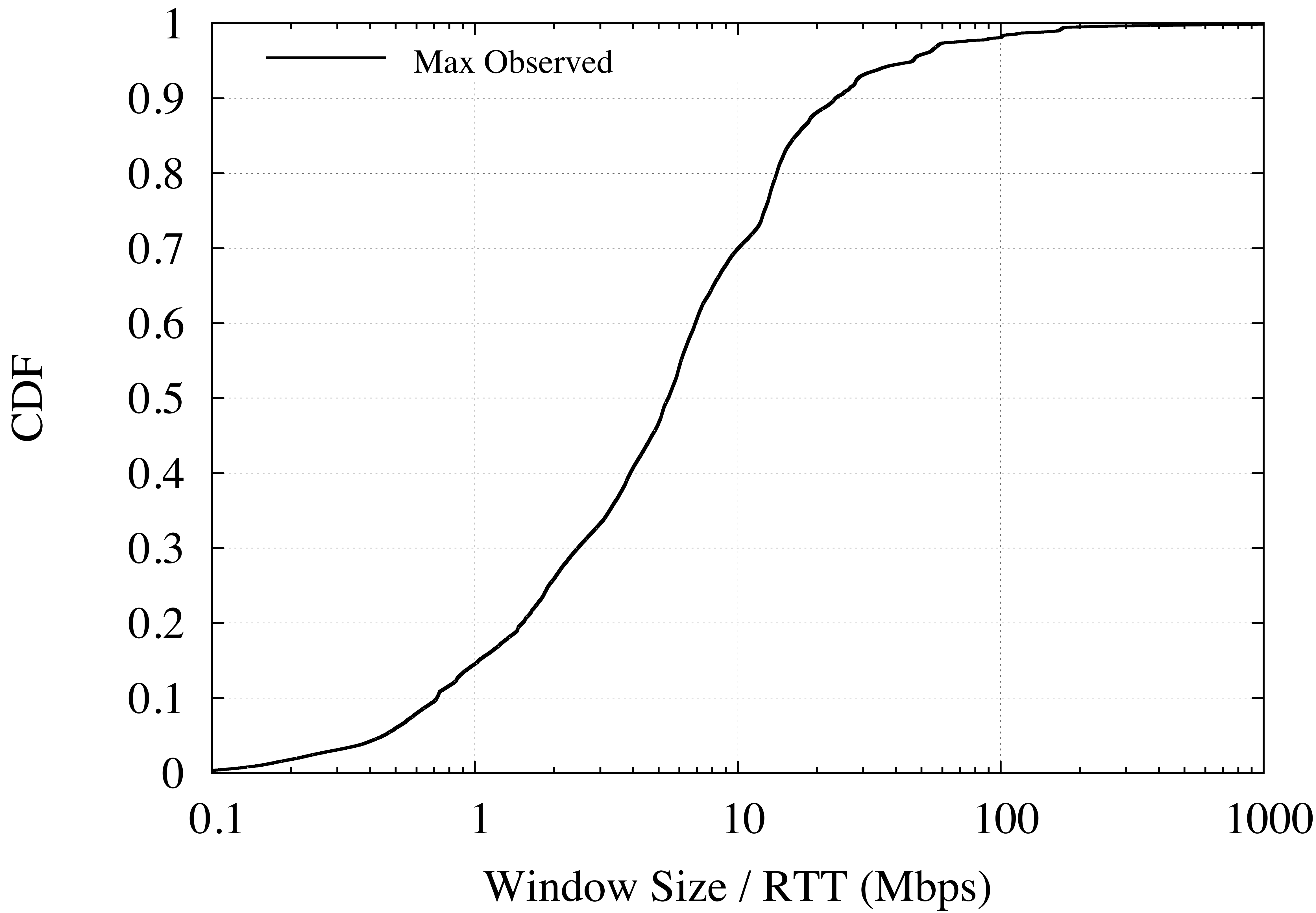
Result 3

TCP implementations limit connection performance

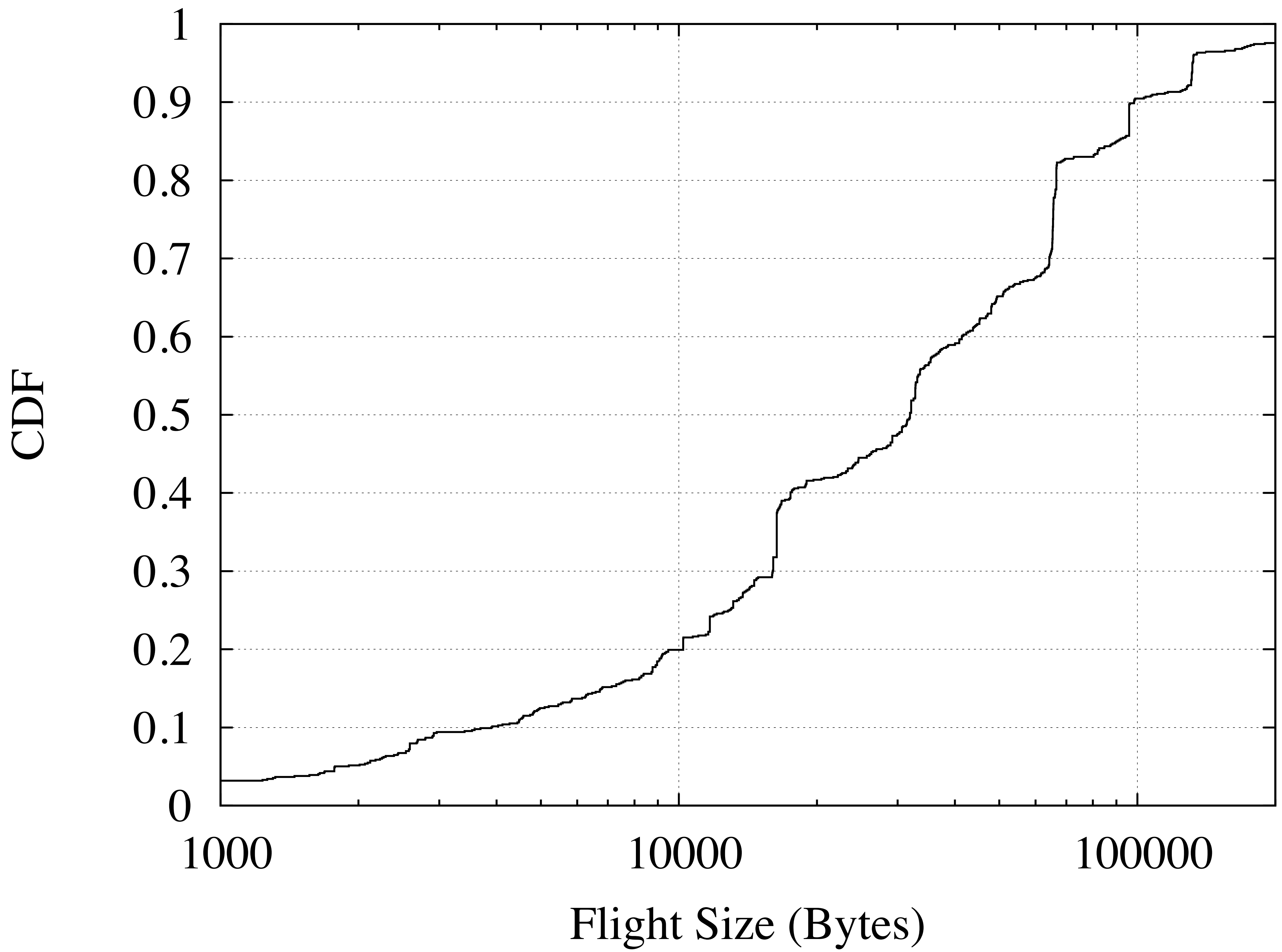
Maximum TCP Throughput

$$Throughput = \frac{WindowSize}{RTT}$$

Advertised window - Outgoing



Flight Size



1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
CDF
1000
10000
100000
Flight Size (Bytes)

Other Results

- Incoming bytes more evenly distributed across homes than outgoing bytes
- CDNs and streaming video make up bulk of incoming HTTP traffic
- HTTP and BitTorrent dominate fast incoming and outgoing transmission periods, respectively
- Based on loss rate, TCP theory suggests faster connection speeds are possible
- Etc.

Publications

- [SA14] Matt Sargent and Mark Allman. Performance Within A Fiber-To-The-Home Network. ACM Computer Communications Review, 44(3), July 2014.

- [SSDA12] Matt Sargent, Brian Stack, Tom Dooner, and Mark Allman. A First Look at 1 Gbps Fiber-To-The-Home Traffic. Technical Report 12-009, International Computer Science Institute, August 2012.

Transport Protocols

Revisiting TCP’s Initial Retransmission Timeout

Deriving Application Sending Patterns From the Transport Layer

Revisiting TCP's Initial Retransmission Timeout

CASE SCHOOL OF ENGINEERING
CASE WESTERN RESERVE UNIVERSITY

Motivation

- TCP requires a timeout to recover from certain types of loss
- The retransmission timeout (RTO) adjusts as a connection progresses
  - Adjustments based on round trip times

Motivation

- Initial RTO value should reflect a "reasonable" timeout

- RFC 2988 specifies initial RTO of 3 seconds
  - But RTTs are typically under 1 second

- What impact would lowering the initial RTO from 3 seconds to 1 second have on network traffic?

Data

Name	Dates	Packets	Connections	Clients	Servers
LBL-1	Oct/05–Mar/06	292M	242K	228	74K
LBL-2	Nov/09–Feb/10	1.1B	1.2M	1,047	38K
ICSI-1	Sep/11–18/07	137M	2.1M	193	486K
ICSI-2	Sep/11–18/08	163M	1.9M	177	277K
ICSI-3	Sep/14–21/09	334M	3.1M	170	253K
ICSI-4	Sep/11–18/10	298M	5M	183	189K
Dartmouth	Jan/4–21/04	1B	4M	3,782	132K
SIGCOMM	Aug/17–21/08	11.6M	133K	152	29K
Total	Jan/2004–Sep/2010	3.3B	17.7M	5.9K	1.4M

Result 1

Up to 2% of connections retransmit their SYN in each dataset

Result 2

- Fewer than 0.1% of connections have RTTs greater than 1 second (1.1% at Dartmouth)
  - Send a spurious SYN
  - Congestion window will collapse

Result 3

- 10% performance improvement:
  - ranges from 43% (LBL-1) to 87%(ICSI-4)
- 50% performance improvement:
  - 17% (ICSI-1 / SIGCOMM) to 73% (ICSI-4).

Publications

- [PACS11] Vern Paxson, Mark Allman, Jerry Chu, and Matt Sargent. Computing TCP's Retransmission Timer, June 2011. RFC 6298.

Deriving Application Sending Patterns From the Transport Layer

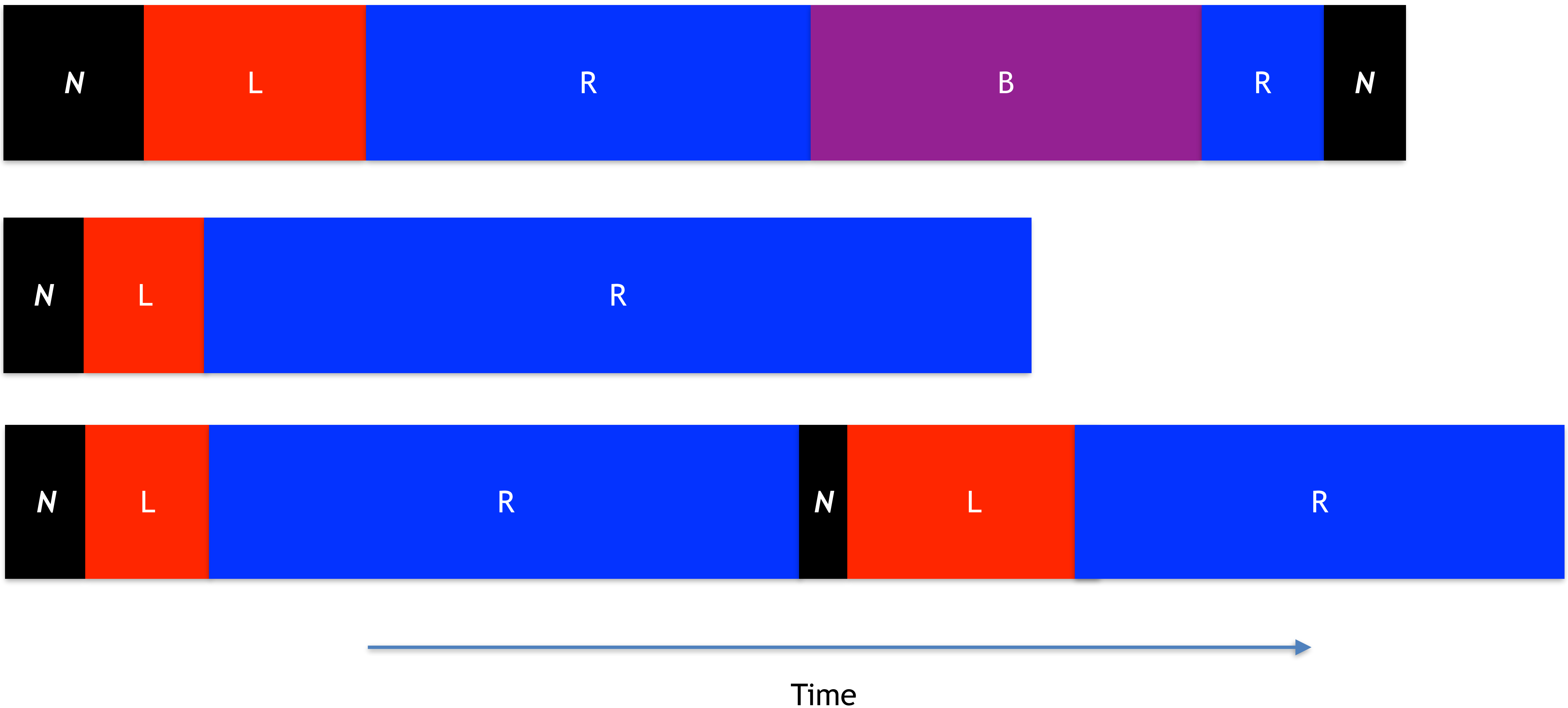
Motivation

- Applications are responsible for handing data to TCP
  - TCP is tuned for bulk transfers
  - No longer strictly bulk transfer
- Can we understand application sending patterns by studying the transport layer?

Methodology

- Collect packet traces from the CCZ and the International Computer Science Institute
- Split connections into sending periods
  - ***L***ocal
  - ***R***emote
  - ***B***oth
  - ***N***one

Methodology



Result 1

How often does silence appear in connections?

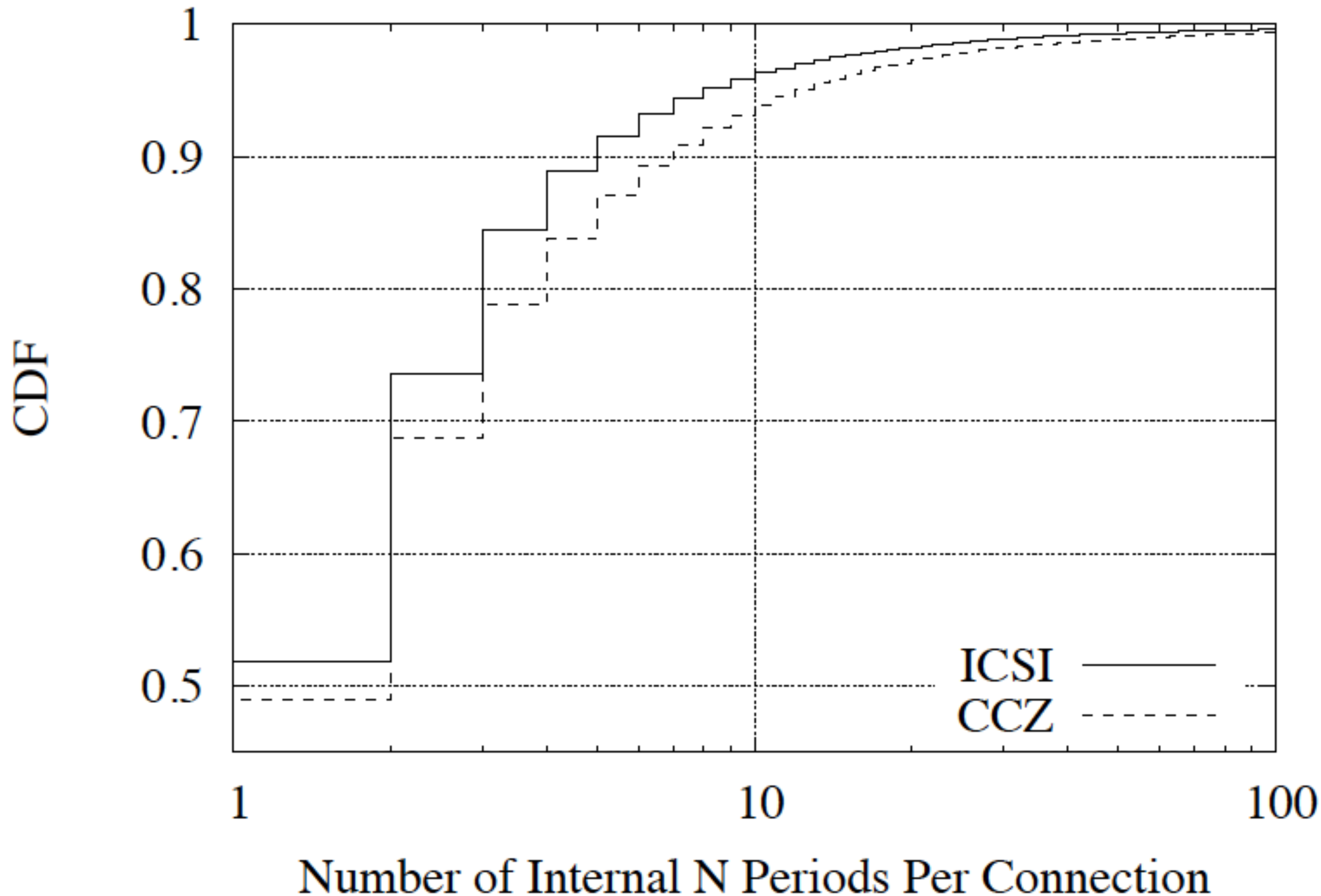
Result 1

Location	**CCZ**	**ICSI**
No N	31%	51.2%
Internal-only	14.4%	18.3%
Trailing-only	32.3%	20.7%
Internal & Trailing	22.3%	9.8%

Result 2

Most connections have only a few internal silent periods

Result 2

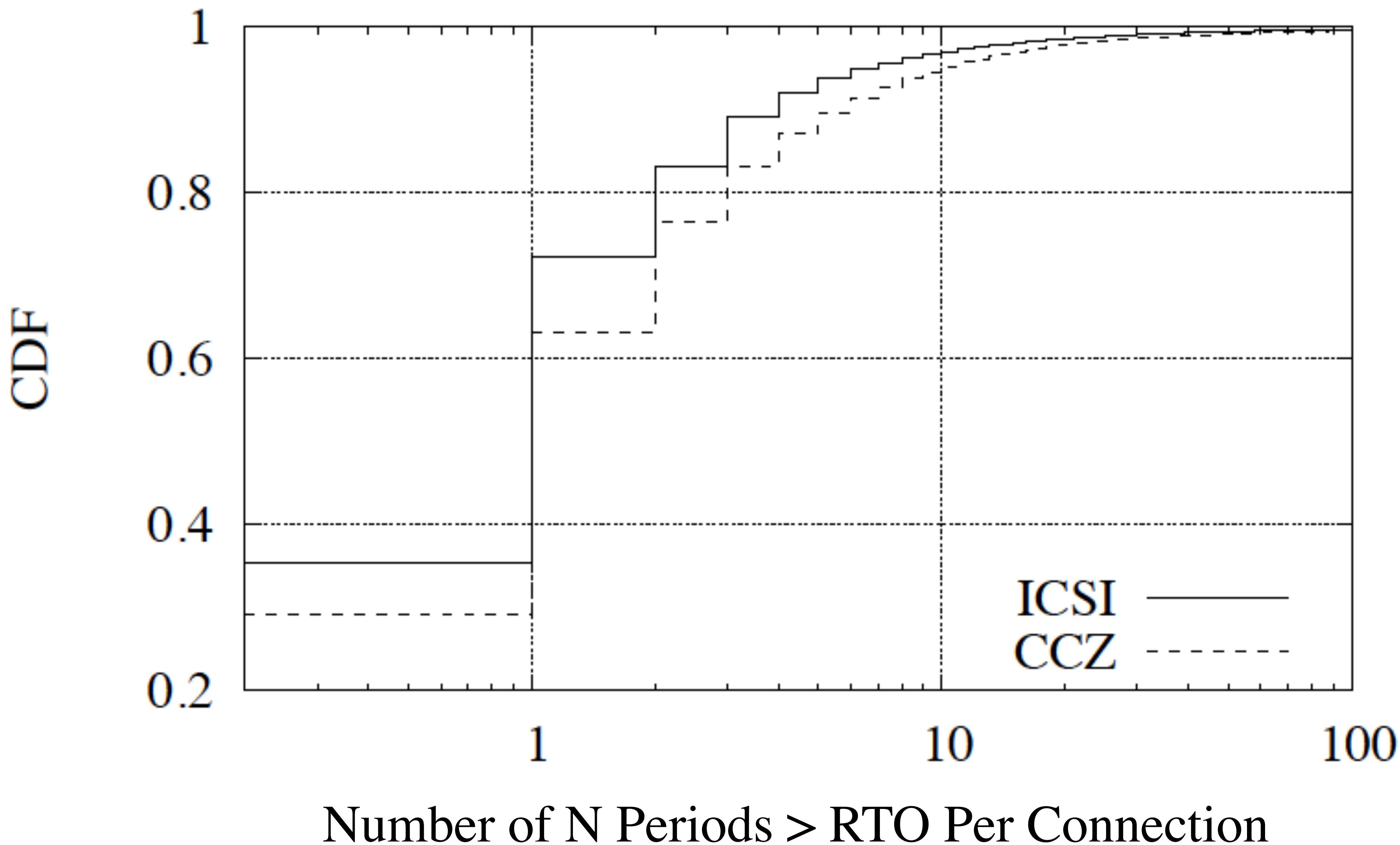


CDF
1
0.9
0.8
0.7
0.6
0.5
1
10
100
Number of Internal N Periods Per Connection
ICSI
CCZ

Result 3

Silent periods are long enough to negatively affect TCP performance

Result 3


CDF
1
0.8
0.6
0.4
0.2
1
10
100
Number of N Periods > RTO Per Connection
ICSI
CCZ

Other Results

- Trailing silences highlight persistent behavior in TCP connections
- Focus on silent period characteristics for specific applications
- Around 1/3 of connections with silent periods spend at least 90% of their duration in silence
- Etc.

Publications

- [SBA 14] Matt Sargent, Ethan Blanton, and Mark Allman. Modern Application Layer Transmission Patterns from a Transport Perspective. In Passive and Active Measurement Conference, March 2014.

Policy and Security Threats

Inferring Filtering via Passive Observation

Understanding IGMP *Neighbors2* Response Behavior

Inferring Filtering via Passive Observation

Motivation

- Traffic filtering is used by edge networks
  - No idea how wide spread specific filtering is
  - Previous efforts require active measurements

- Can we come up with a passive method to infer policy filters?

Data

- Collect packet traces at 5 /8 darknets
  - 2.5% of IPv4 address space
  - Receive packets from 4.1M /24s

Methodology

- Use *traffic markers* to infer filtering policy
  - Types of traffic that we can expect to observe from many network locations
  - Initial focus is on Conficker traffic

Result 1

We expect Conficker on 1.6M out of 4.1M /24s

Result 2

We judge 55% of /24s that contain Conficker infectees

Result 2

Expect Conficker?	Observe Conficker?	>=5* known infectees?	Judgement	Total
F	F	-	None	
F	T	-	Rare	<1%
T	T	-	No Filter	27%
T	F	T	Filtering	28%
T	F	F	None	45%

* Threshold developed in dissertation

Result 3

Aggregating up to routed prefix enables us to judge 699M IP addresses (28% of routable addresses)

Limitations

- Traffic markers are imperfect

- Finding a traffic marker is difficult
  - Most types of scanning traffic arrive at the darknet from < 1% of /24s

Other Results

- Additional details on Conficker behavior
- Validation of our methodology against Netalyzr "ground truth"
- More detailed breakdown of judgements for routed prefixes
- Evidence of multiple policies in place for routed prefixes, especially for large prefix sizes
- Etc.

Publications

- [SCAB15] Matt Sargent, Jakub Czyz, Mark Allman, and Michael Bailey. On The Power and Limitations of Detecting Network Filtering via Passive Observation. In Passive and Active Measurement Conference, March 2015.

Understanding IGMP *Neighbors2* Response Behavior

Introduction

- Internet Group Management Protocol (IGMP)
  - Multicast group membership management

- Distance Vector Multicast Routing Protocol (DVMRP)
  - Enables routers to exchange multicast routing information

Introduction

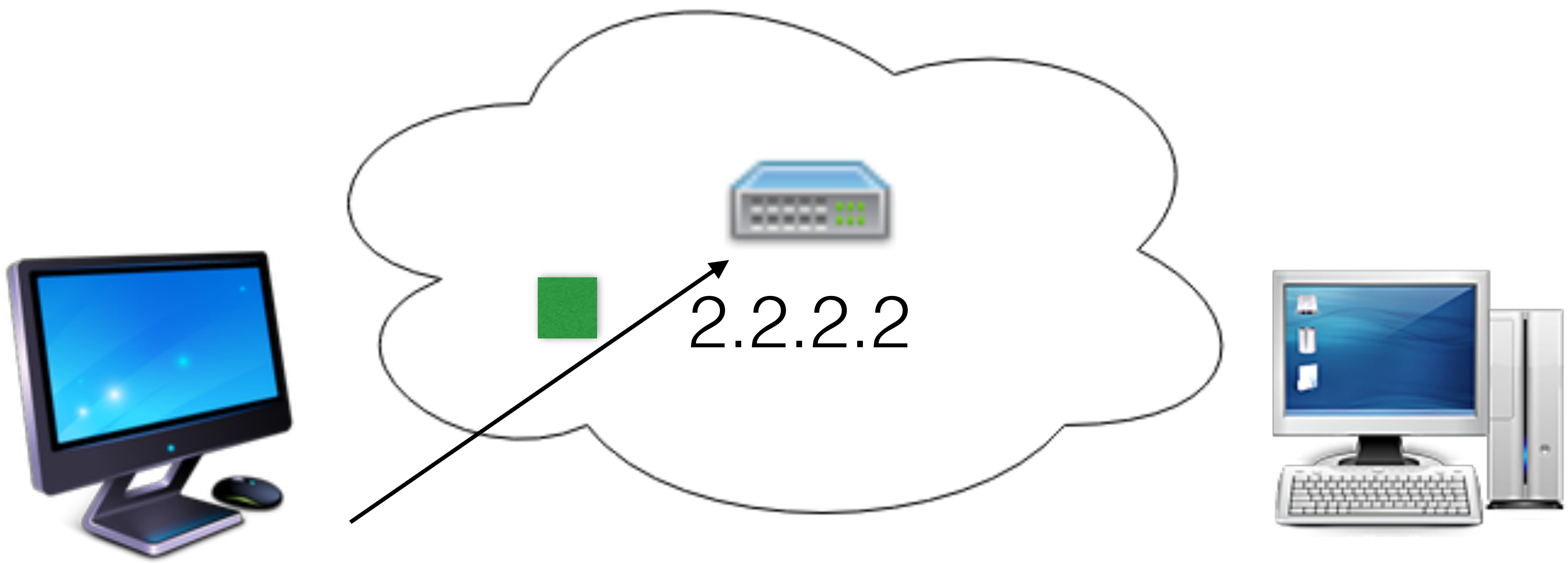
- *AskNeighbors2* packets explicitly request routing information from a router

- *Neighbors2* packets contain multicast neighbor information

Introduction

- *AskNeighbors2* packets have been used to study network topology
  - *MERLIN*
  - *mrinfo*
- Connectionless exchange of information creates a potential attack vector

Reflection



Attacker

1.1.1.1

Victim

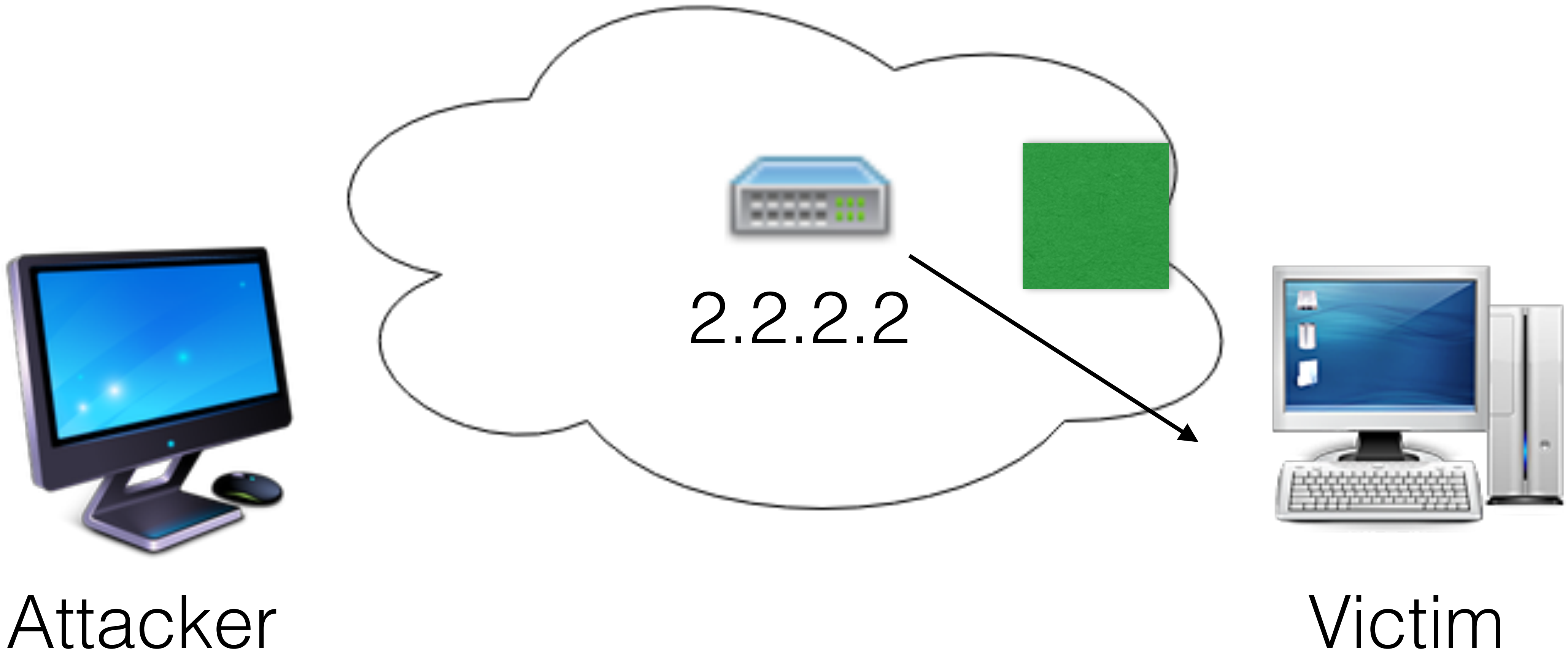
3.3.3.3

AskNeighbors2

Source IP: 3.3.3.3

Destination IP: 2.2.2.2

Amplification



Neighbors2
Source IP: 2.2.2.2
Destination IP: 3.3.3.3

Methodology

- Write custom probing module for *ZMap*
- Scan IPv4 address space with *AskNeighbors2* requests
- Capture *Neighbors2* responses
  - Re-probe responding hosts 10, 20, and 30 days after the initial scan

Initial Scan

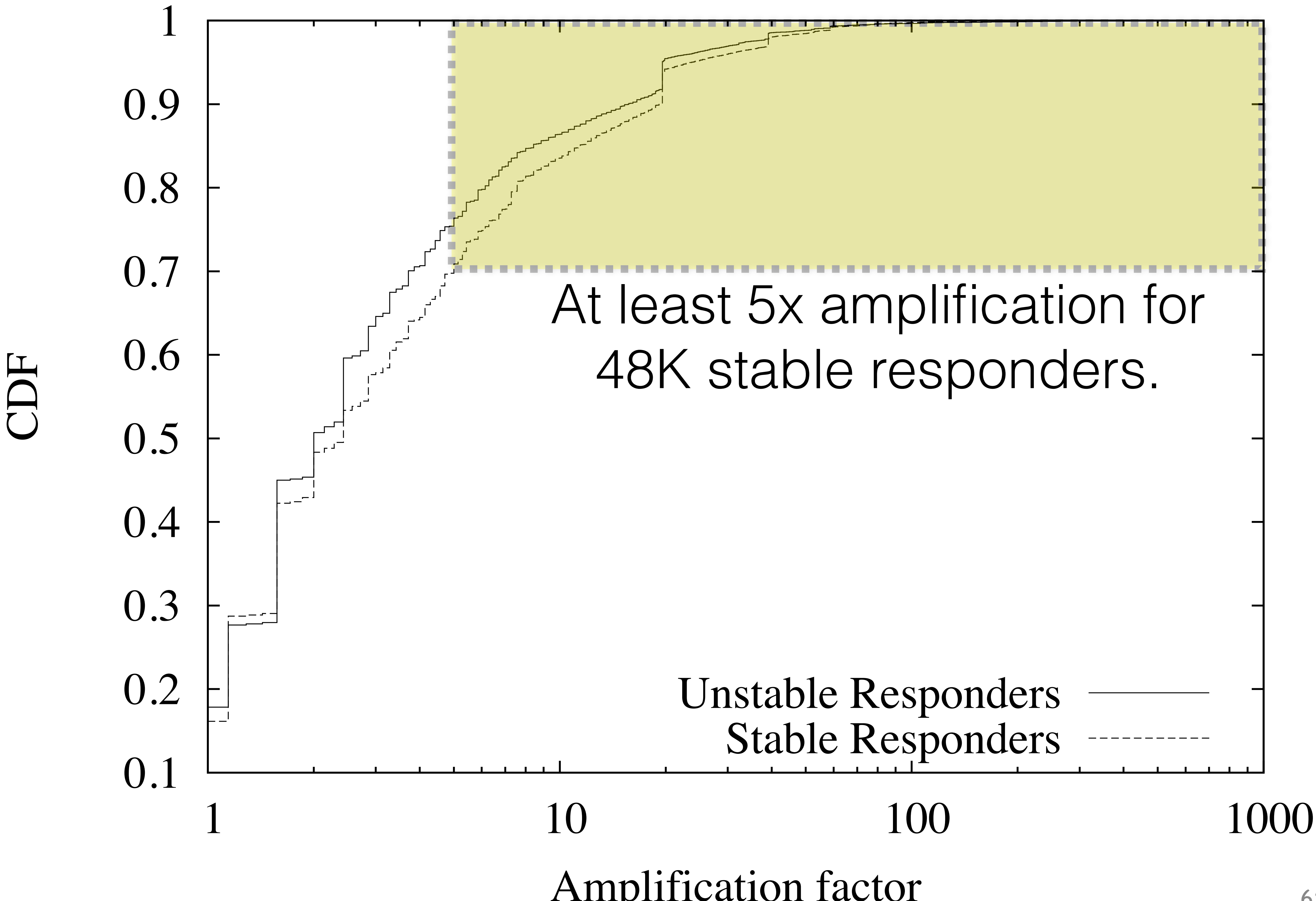
Start Date	End Date	Outgoing Pkts	Incoming Pkts.	Responding IPs
2015/01/12	2015/01/18	4.2B	263M	305K

Re-probes

- 262K (86%) out of 305K hosts respond in at least one of three re-probes

- 161K (52.8%) hosts respond to all three re-probes
  - Call these hosts “stable responders”

Amplification



Denial of service attack

- Hit list of 48K stable responders with at least 5X amplification

- Send each stable responder 53 packets per second

Denial of service attack

- This strategy produces 1.27 GB of data forwarded to the victim each second
  - Rate of 10.2 Gbps

- Requires 570 Mbps in total from the attacker

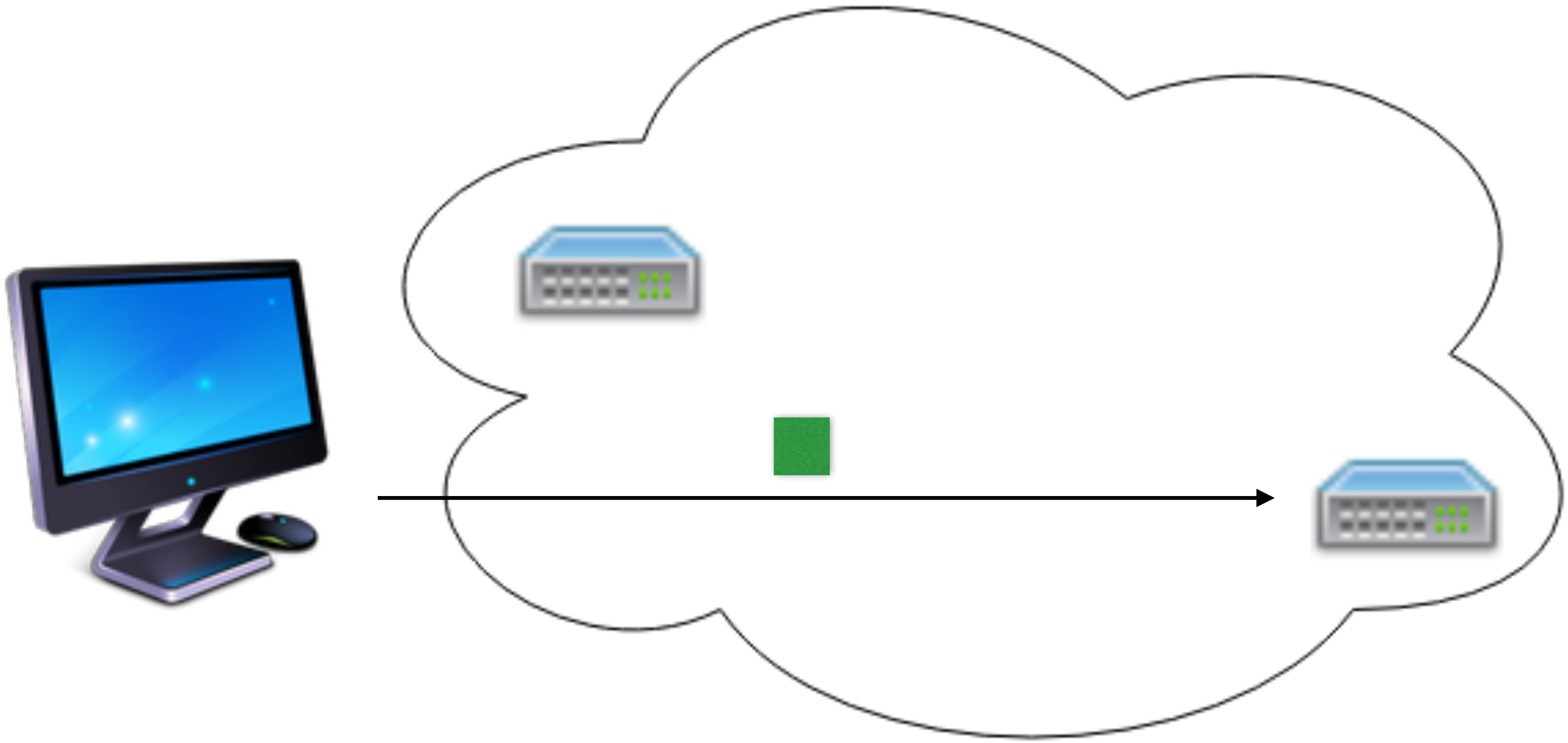
Pulse attack

- Similar to denial of service attack
- Rather than a sustained attack, direct a large burst of traffic to a victim
  - Repeat burst every few seconds
  - Disrupt congestion control with temporary congestion at the victim's network

Pulse attack

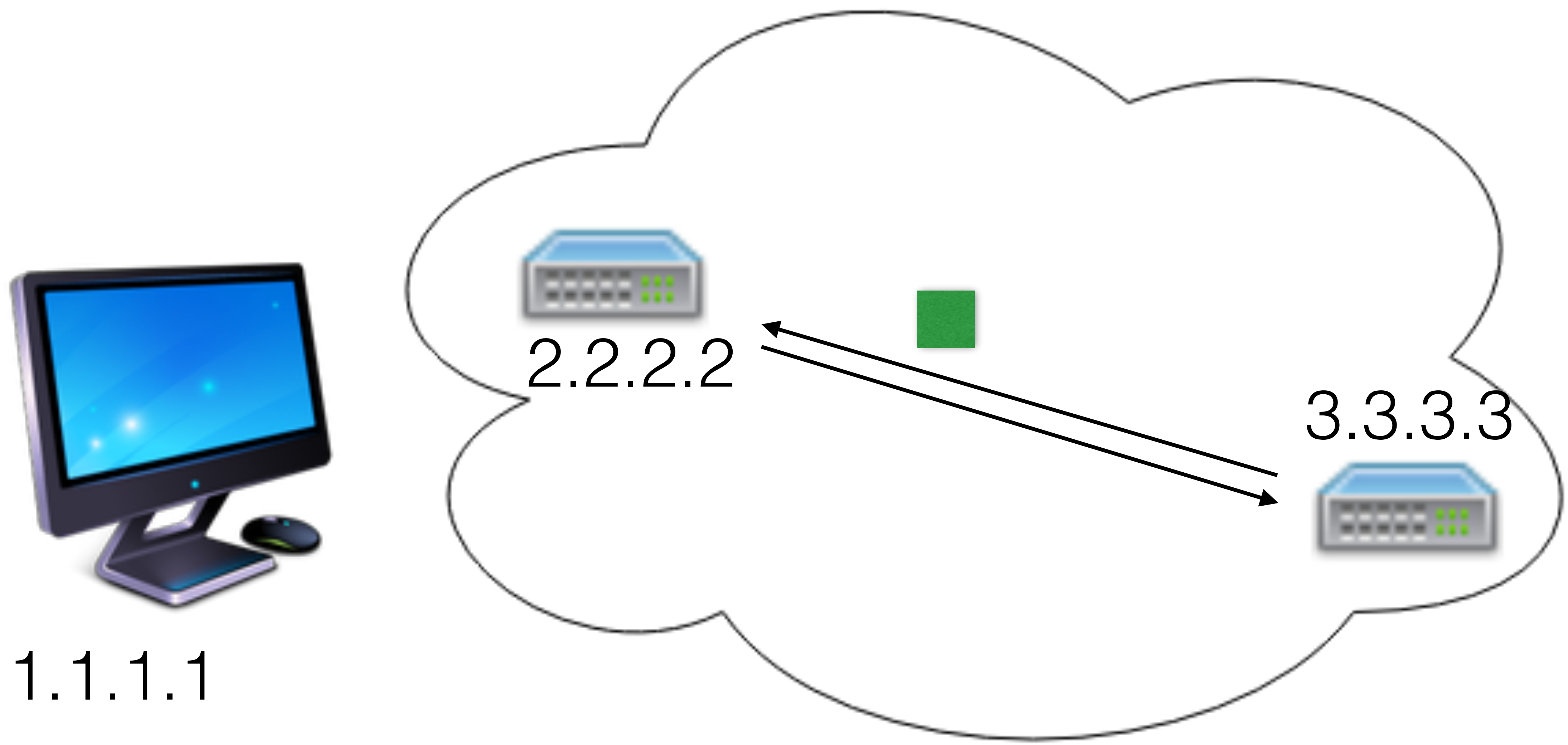
- Send a single packet to each of the 48K stable responders with at least 5x amplification
  - Generates at least 192 Mbps worth of traffic sent to the victim
  - Requires 10.7 Mbps from the attacker

Loop Attack



AskNeighbors2

Loop Attack



Other Results

- Unstable responders
- Packet amplification
- Anomalous responses
- Responder locality
- Etc.

Conclusion

Applications

- CCZ application sending patterns suggest prevalence of distinct transactions
  - Types of applications used on CCZ largely mirror other residential networks
  - Suggests non-bulk demand is pervasive

- May need to introduce additional mechanisms to improve TCP performance further

TCP Performance

- Behavior of TCP is defined by both the underlying specification and implementation
  - TCP implementations are outpaced by last mile bandwidth
  - TCP specification is outpaced by lower RTTs

Questions?

On Understanding the Internet Via Edge Measurement

May 14, 2015

Matt Sargent

Advisor: Mark Allman

Publications

- [SCAB15] Matt Sargent, Jakub Czyz, Mark Allman, and Michael Bailey. On The Power and Limitations of Detecting Network Filtering via Passive Observation. In Passive and Active Measurement Conference, March 2015.
- [SA14] Matt Sargent and Mark Allman. Performance Within A Fiber-To-The-Home Network. ACM Computer Communications Review, 44(3), July 2014.
- [SBA 14] Matt Sargent, Ethan Blanton, and Mark Allman. Modern Application Layer Transmission Patterns from a Transport Perspective. In Passive and Active Measurement Conference, March 2014.
- [SSDA12] Matt Sargent, Brian Stack, Tom Dooner, and Mark Allman. A First Look at 1 Gbps Fiber-To-The-Home Traffic. Technical Report 12-009, International Computer Science Institute, August 2012.
- [PACS11] Vern Paxson, Mark Allman, Jerry Chu, and Matt Sargent. Computing TCP's Retransmission Timer, June 2011. RFC 6298.