Towards Improving Outage Detection with Multiple Probing Protocols

Manasvini Sethuraman, Zachary S. Bischof, Alberto Dainotti
Georgia Tech

Georgia Tech

Monitoring the Internet for Outages

- Internet is critical infrastructure
- Enormous, distributed system
- Outages aren't uncommon
- Outages from Feb 26 to March 2
  - Source: IODA[1]
- 226 outages so far this year

[1]https://ioda.inetintel.cc.gatech.edu/

Georgia Tech

How are outages detected?

- Active probing via ICMP
  - Probe a subset of IP addresses in a /24 block via ICMP and record responses [SIGCOMM 13]
  - Infer if there is an outage based on some threshold criterion

ICMP Probing

Advantages

- Wide coverage
- Fast: each probing cycle lasts ~11 minutes

ICMP Probing

Advantages

- Wide coverage
- Fast: each probing cycle lasts ~11 minutes

Issues

- Some networks block ICMP
- /24 blocks with few responsive hosts require more than one probing cycle

Internet Wide Scans: What can we learn?

- Internet wide scans use different protocol handshakes to establish "liveness" of hosts
- Censys [CCS 2015] uses ZMap to scan the IPv4 address space using HTTP, SSH, DNS and other protocols for open ports and service discovery
- Combination of TCP + UDP + ICMP yields highest number of hosts [CCR 2018]

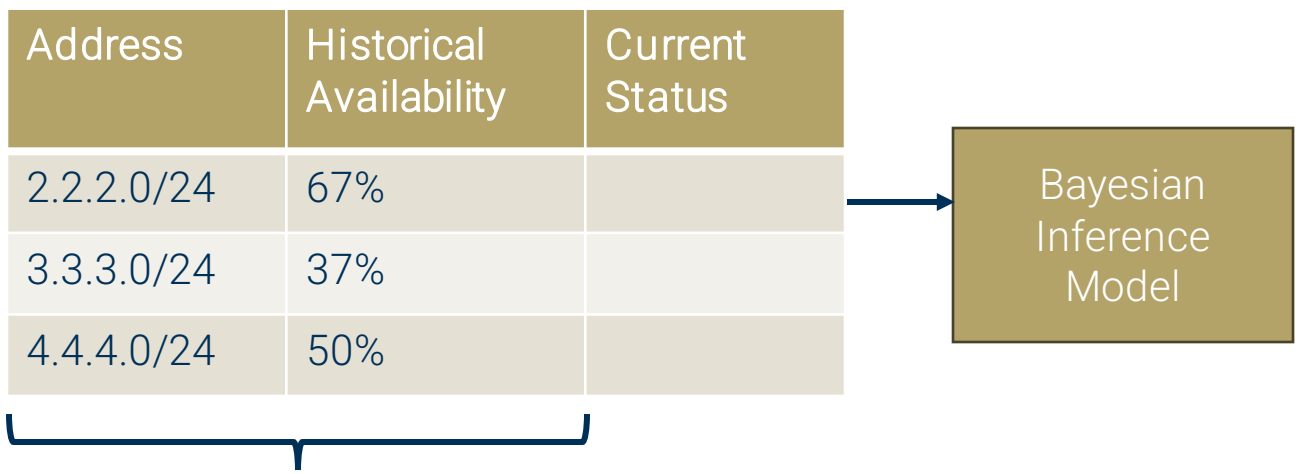
TCP/UDP probing in outage detection?

Can we quantify the improvement that adding TCP/UDP probing may bring in terms of block coverage in outage detection?

Active Probing in Practice

- Trinocular [SIGCOMM 2013] uses ICMP probes to estimate /24 block status (Up/down) probability using Bayesian inference
- Used in at least 2 operational systems at ISI and GaTech

Address	Historical Availability	Current Status
2.2.2.0/24	67%	
3.3.3.0/24	37%	
4.4.4.0/24	50%	

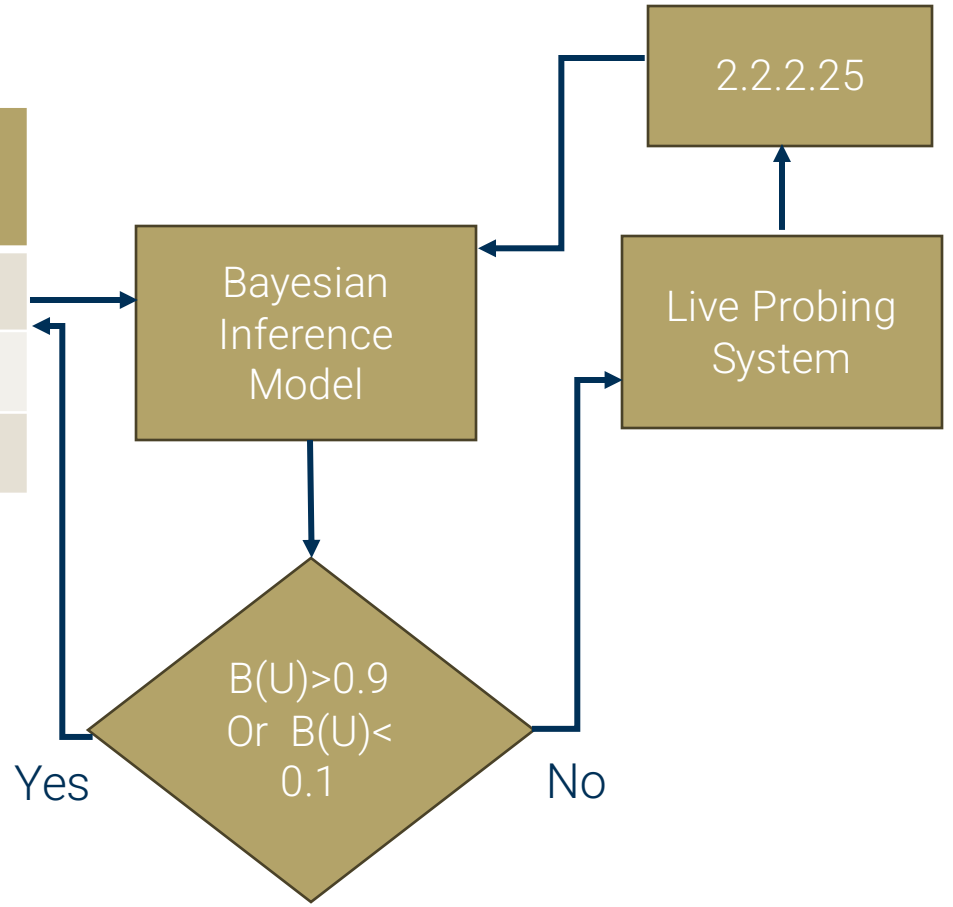


Georgia Tech

Active Probing in Practice

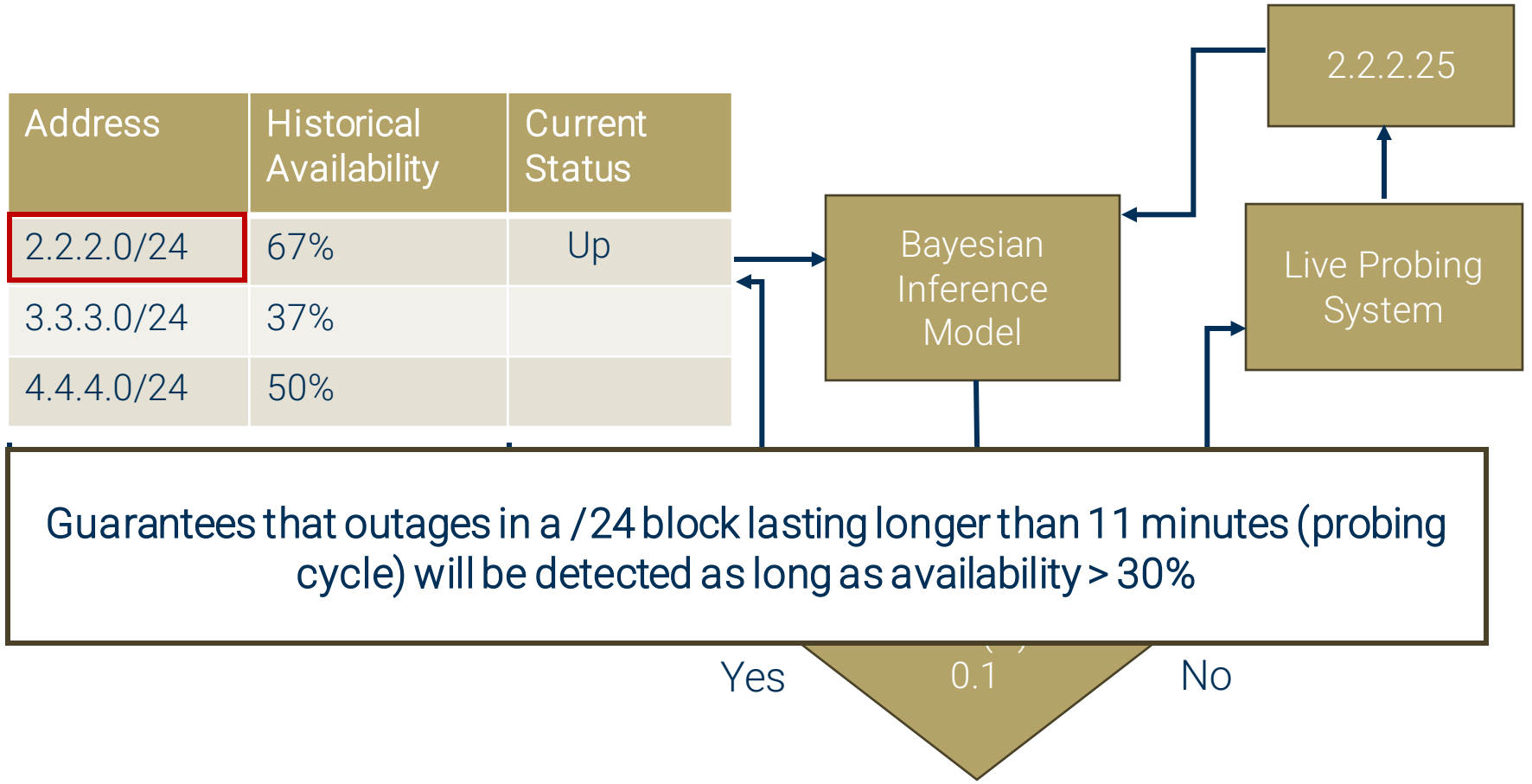
- Trinocular [SIGCOMM 2013] uses ICMP probes to estimate /24 block status (Up/down) probability using Bayesian inference
- Used in at least 2 operational systems at ISI and GaTech

Address	Historical Availability	Current Status
2.2.2.0/24	67%	Up
3.3.3.0/24	37%	
4.4.4.0/24	50%	



Active Probing in Practice

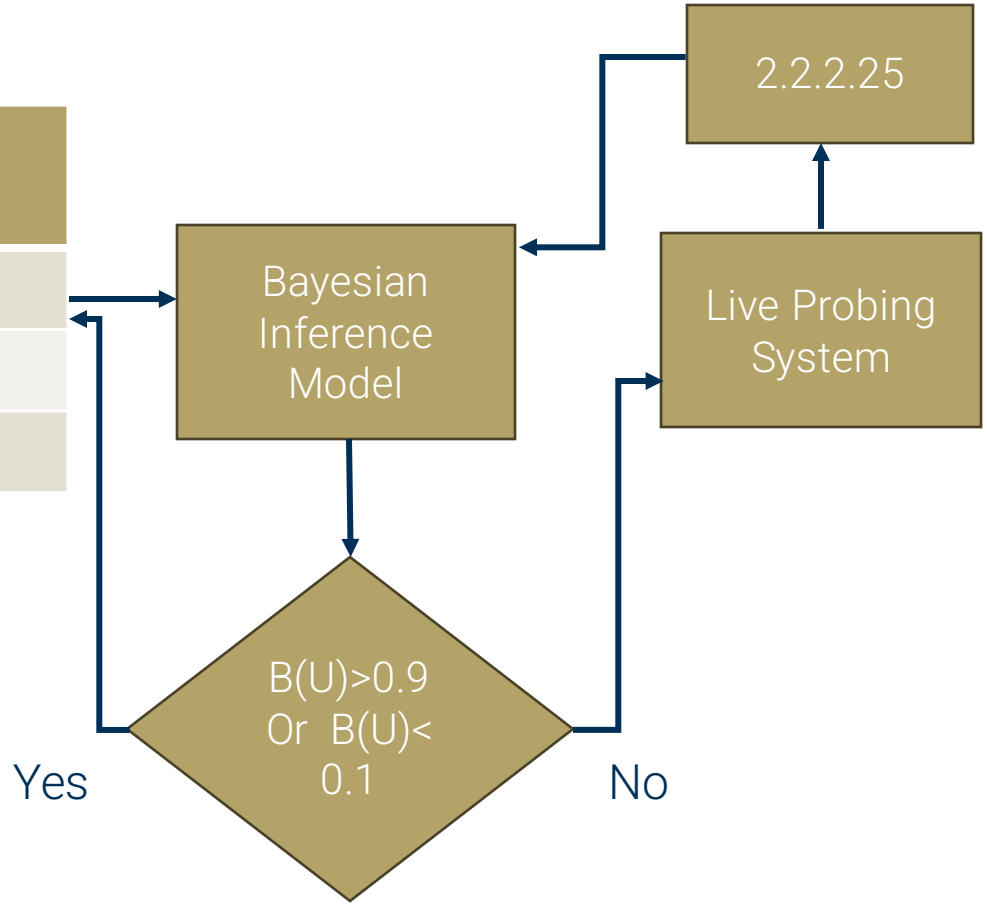
- Trinocular [SIGCOMM 2013] uses ICMP probes to estimate /24 block status (Up/down) probability using Bayesian inference
- Used in at least 2 operational systems at ISI and GaTech



Georgia Tech

Active Probing in Practice

Address	Historical Availability	Current Status
2.2.2.0/24	67%	
3.3.3.0/24	37%	
4.4.4.0/24	50%	

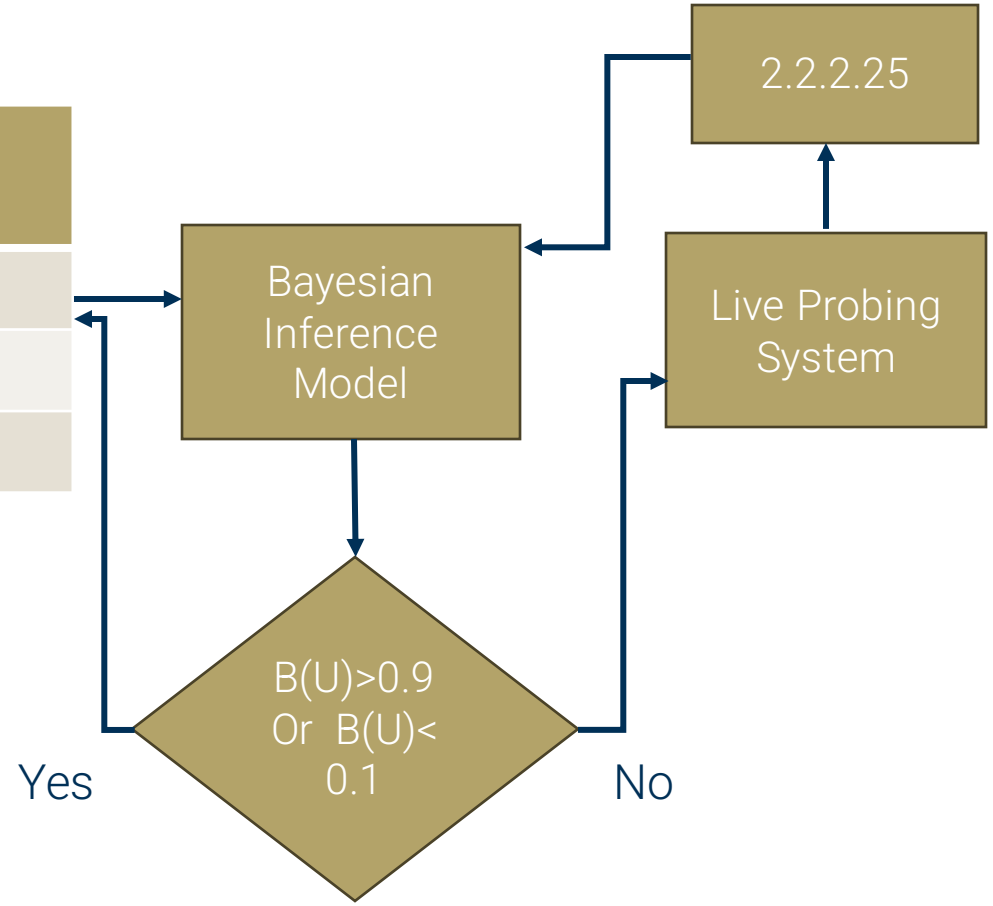


Georgia Tech

Active Probing in Practice

- Compare historical availability used to seed the inference model

Address	Historical Availability	Current Status
2.2.2.0/24	67%	
3.3.3.0/24	37%	
4.4.4.0/24	50%	



Georgia Tech

Quantifying improvement

- What metrics should we use?
  - Availability of a host
  - Reliability of a /24 block

Quantifying improvement: Host Availability

- Host availability
  - # of successful responses from a host / # of probes to the host

Quantifying improvement: Host Availability

- Host availability
  - # of successful responses from a host / # of probes to the host

Address	Survey 1	Survey 2	Survey 3	Average
2.2.2.1	0	1	1	67%
2.2.2.2	0	0	1	33%
2.2.2.25	0	1	0	33%

Quantifying improvement: Host Availability

- Host availability
  - # of successful responses from a host / # of probes to the host

Address	Survey 1	Survey 2	Survey 3	Average
2.2.2.1	0	1	1	67%
2.2.2.2	0	0	1	33%
2.2.2.25	0	1	0	33%

Quantifying improvement: Host Availability

- Host availability
  - # of successful responses from a host / # of probes to the host

Address	Survey 1	Survey 2	Survey 3	Average
2.2.2.1	0	1	1	67%
2.2.2.2	0	0	1	33%
2.2.2.25	0	1	0	33%

Quantifying improvement: Block Reliability

- Block availability
  - (avg availability of each address) / # of addresses

Address	Survey 1	Survey 2	Survey 3	Average
2.2.2.1	0	1	1	67%
2.2.2.2	0	0	1	33%
2.2.2.25	0	1	0	33%

Quantifying improvement: Block Reliability

- Block availability
  - (avg availability of each address) / # of addresses
- Availability for this block
  - (67+33+33) / 3
  - 44.4%

Address	Survey 1	Survey 2	Survey 3	Average
2.2.2.1	0	1	1	67%
2.2.2.2	0	0	1	33%
2.2.2.25	0	1	0	33%

Quantifying improvement: Block Reliability

- Block availability
  - (avg availability of each address) / # of addresses
- Availability for this block
  - (67+33+33) / 3
  - 44.4%

Address	Survey 1	Survey 2	Survey 3	Average
2.2.2.1	0	1	1	67%
2.2.2.2	0	0	1	33%
2.2.2.25	0	1	0	33%

- Trinocular guarantees outage (> 11 min) detection in blocks with availability > 30%
- A block is reliable if its availability is at least 30%

Datasets

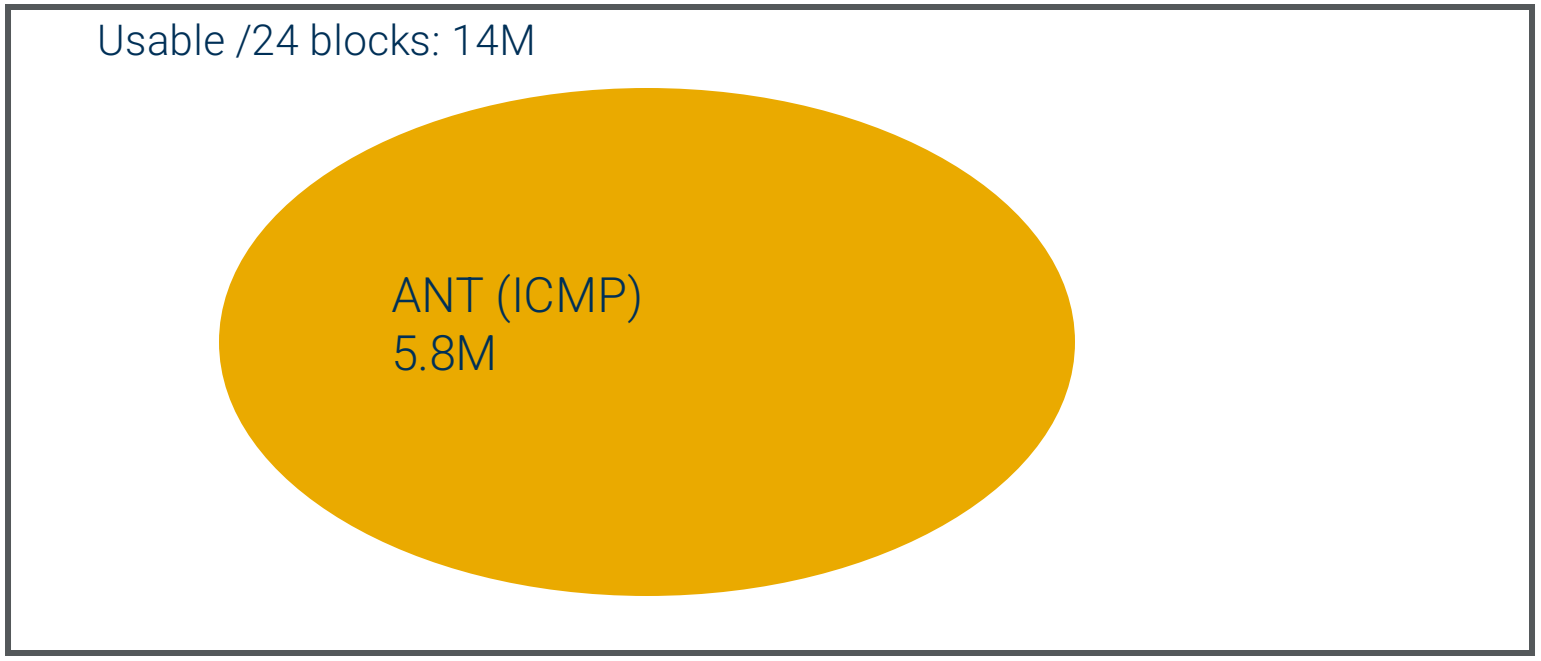
- ICMP
  - ISI ANT IP History Dataset
  - ICMP response from hosts
    - 0 = failure/error/no response, 1=success
  - Survey are ~2 months apart

Datasets

- ICMP
  - ISI ANT IP History Dataset
  - ICMP response from hosts
    - 0 = failure/error/no response, 1=success
  - Survey are ~2 months apart
- TCP/UDP
  - Censys Internet wide scan data
  - Protocol handshakes for HTTP, DNS, SSH, SMTP, NTP, FTP
  - Weekly surveys

Datasets

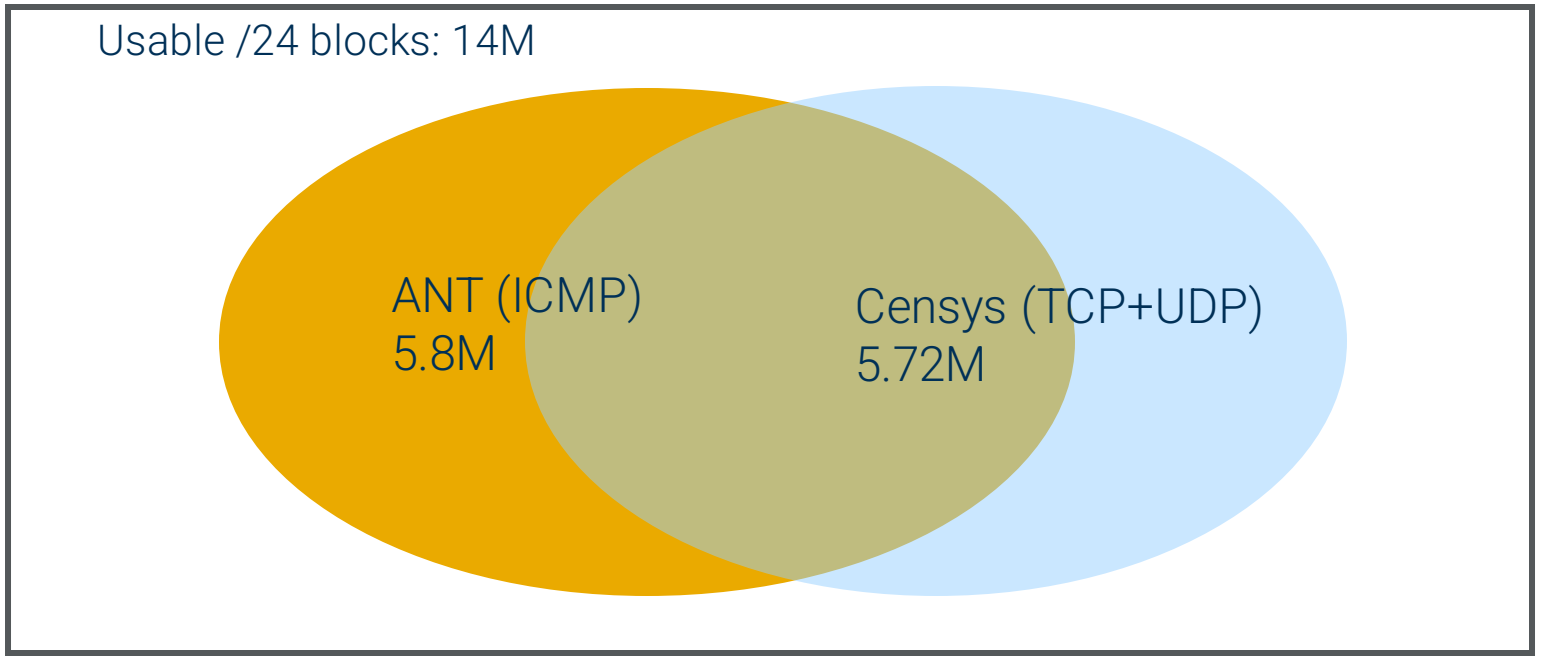
- ICMP
  - ISI ANT IP History Dataset
  - ICMP response from hosts
    - 0 = failure/error/no response, 1=success
  - Survey are ~2 months apart
- TCP/UDP
  - Censys Internet wide scan data
  - Protocol handshakes for HTTP, DNS, SSH, SMTP, NTP, FTP
  - Weekly surveys



Georgia Tech

Datasets

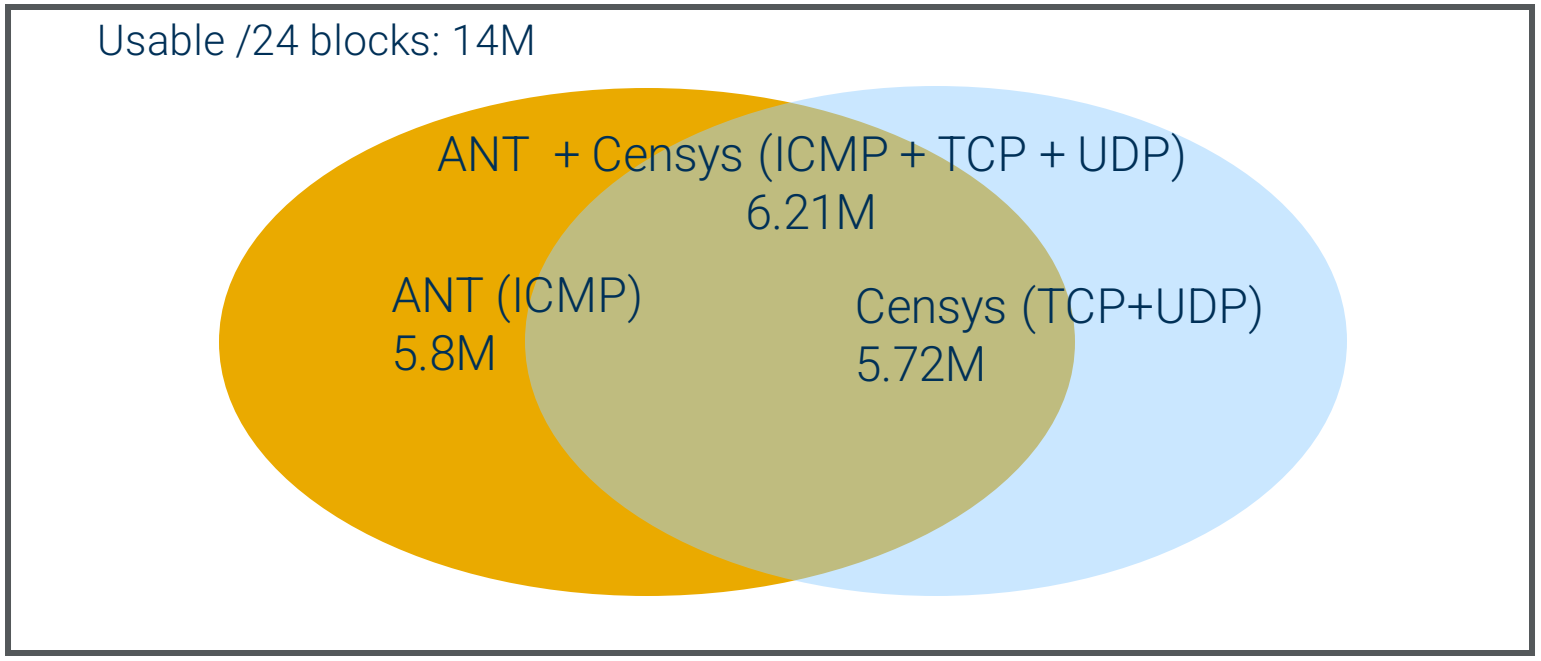
- ICMP
  - ISI ANT IP History Dataset
  - ICMP response from hosts
    - 0 = failure/error/no response, 1=success
  - Survey are ~2 months apart
- TCP/UDP
  - Censys Internet wide scan data
  - Protocol handshakes for HTTP, DNS, SSH, SMTP, NTP, FTP
  - Weekly surveys



Georgia Tech

Datasets

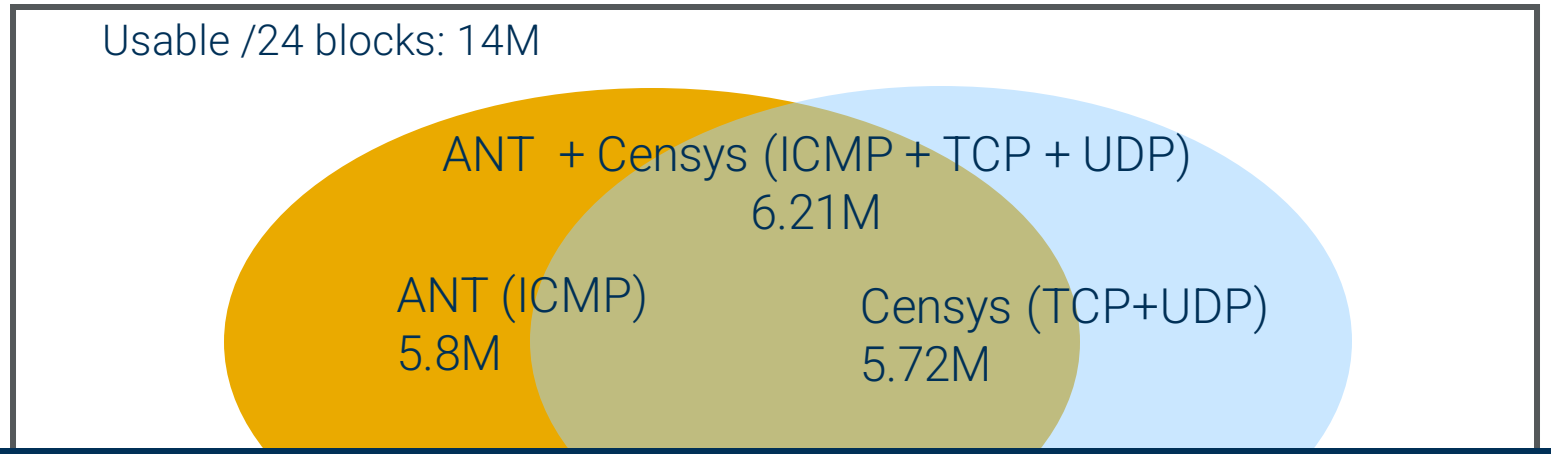
- ICMP
  - ISI ANT IP History Dataset
  - ICMP response from hosts
    - 0 = failure/error/no response, 1=success
  - Survey are ~2 months apart
- TCP/UDP
  - Censys Internet wide scan data
  - Protocol handshakes for HTTP, DNS, SSH, SMTP, NTP, FTP
  - Weekly surveys



Georgia Tech

Datasets

- ICMP
  - ISI ANT IP History Dataset
  - ICMP response from hosts
    - 0 = failure/error/no response, 1=success
  - Survey are ~2 months apart
- TCP/UDP
  - Censys Internet wide scan data
  - Protocol handshakes for HTTP, DNS, SSH, SMTP, NTP, FTP
  - Weekly surveys



- 9 snapshots for each dataset spanning 2 years (Nov 2020 to Dec 2022)
- Combined 820M hosts, 6.21M /24 blocks

Findings: At the host level

- Consider hosts that appear in both datasets across the 9 snapshots
  - On average, how many snapshots does a host respond in?
  - Does the addition of TCP/UDP probes cause an increase in availability?

Metric	Only ICMP	ICMP + TCP + UDP
Average response count	5.24	6.02
Hosts appearing in all snapshots	22%	28%

Findings: At the host level

- Consider hosts that appear in both datasets across the 9 snapshots
  - On average, how many snapshots does a host respond in?
  - Does the addition of TCP/UDP probes cause an increase in availability?

Metric	Only ICMP	ICMP + TCP + UDP
Average response count	5.24	6.02
Hosts appearing in all snapshots	22%	28%

Findings: At the host level

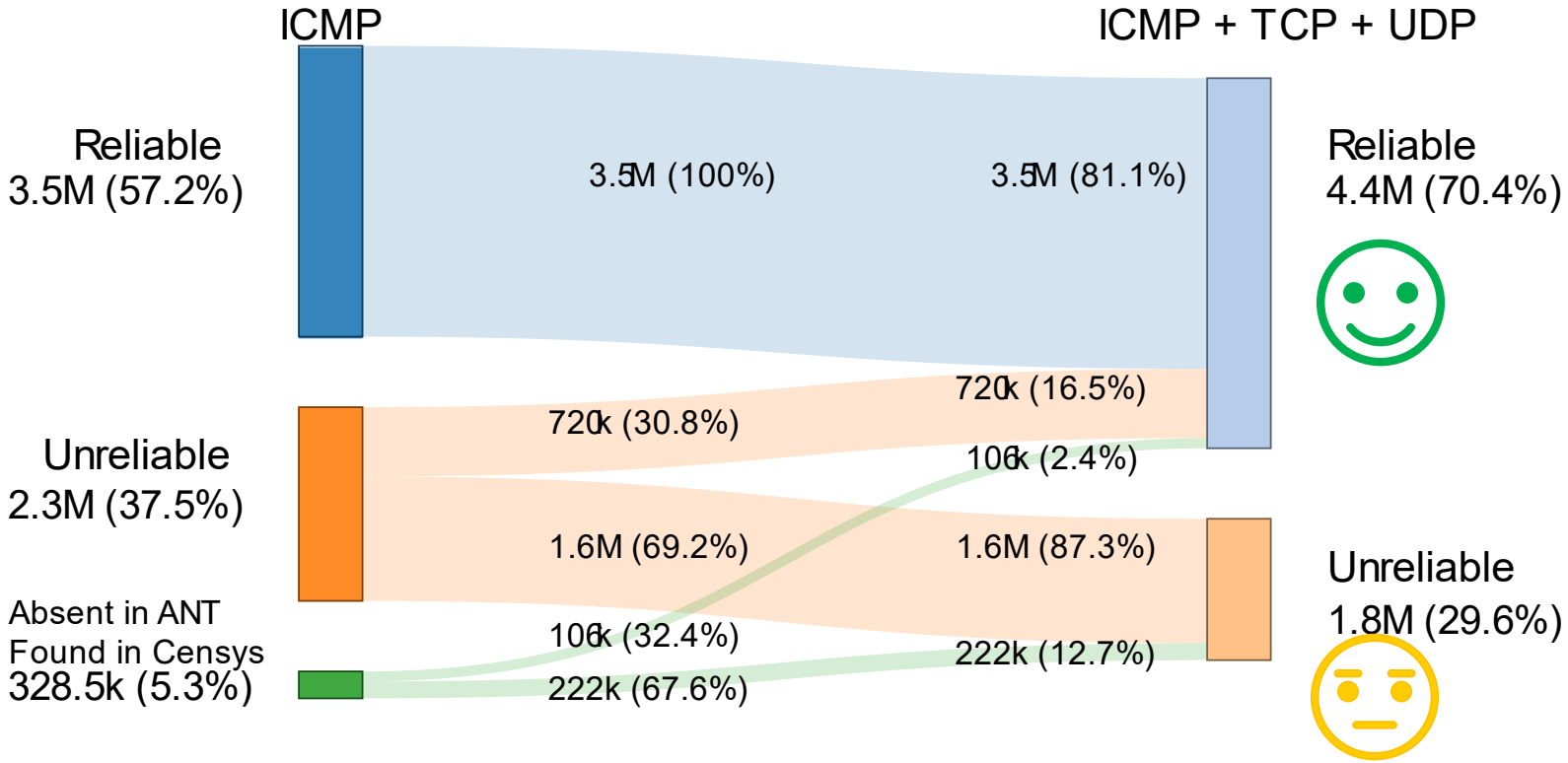
- Consider hosts that appear in both datasets across the 9 snapshots
  - On average, how many snapshots does a host respond in?
  - Does the addition of TCP/UDP probes cause an increase in availability?

Metric	Only ICMP	ICMP + TCP + UDP
Average response count	5.24	6.02
Hosts appearing in all snapshots	22%	28%

The availability of a host on average is higher when we use ICMP + TCP + UDP probing

Findings: At the /24 block level

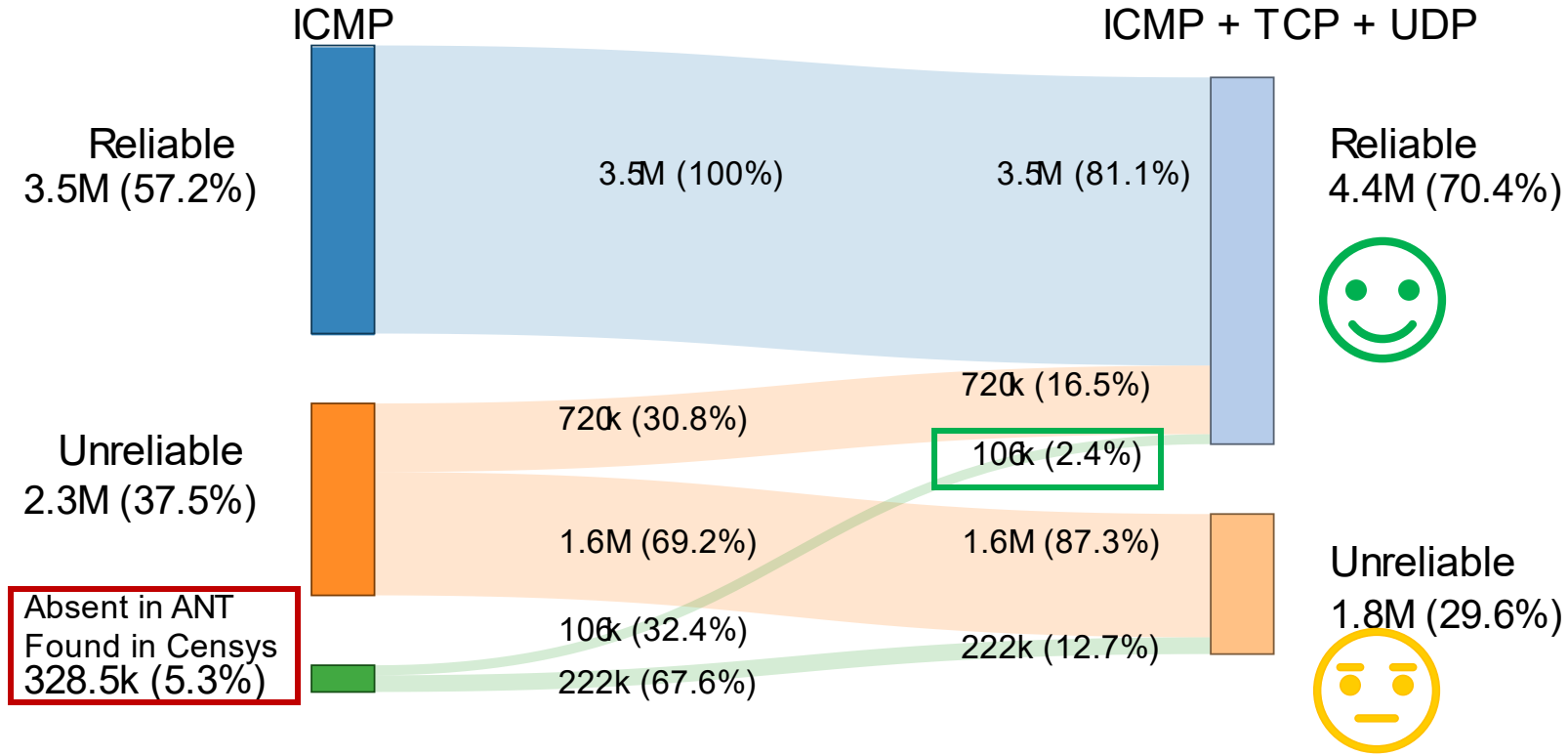
- Reliable /24 blocks (Availability >= 30%)



Georgia Tech

Findings: At the /24 block level

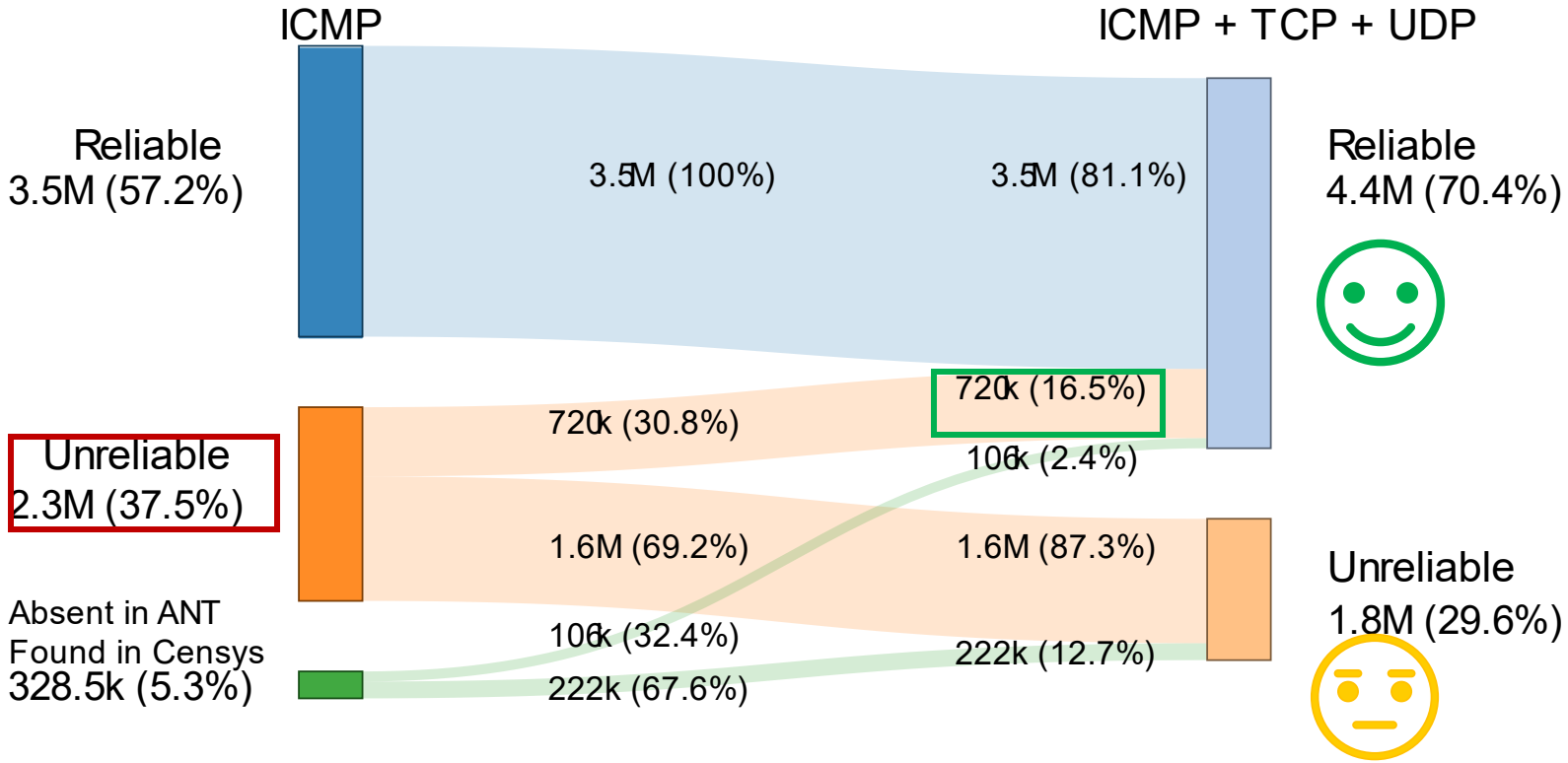
- Reliable /24 blocks (Availability >= 30%)
- TCP/UDP probes can discover previously unseen blocks



Georgia Tech

Findings: At the /24 block level

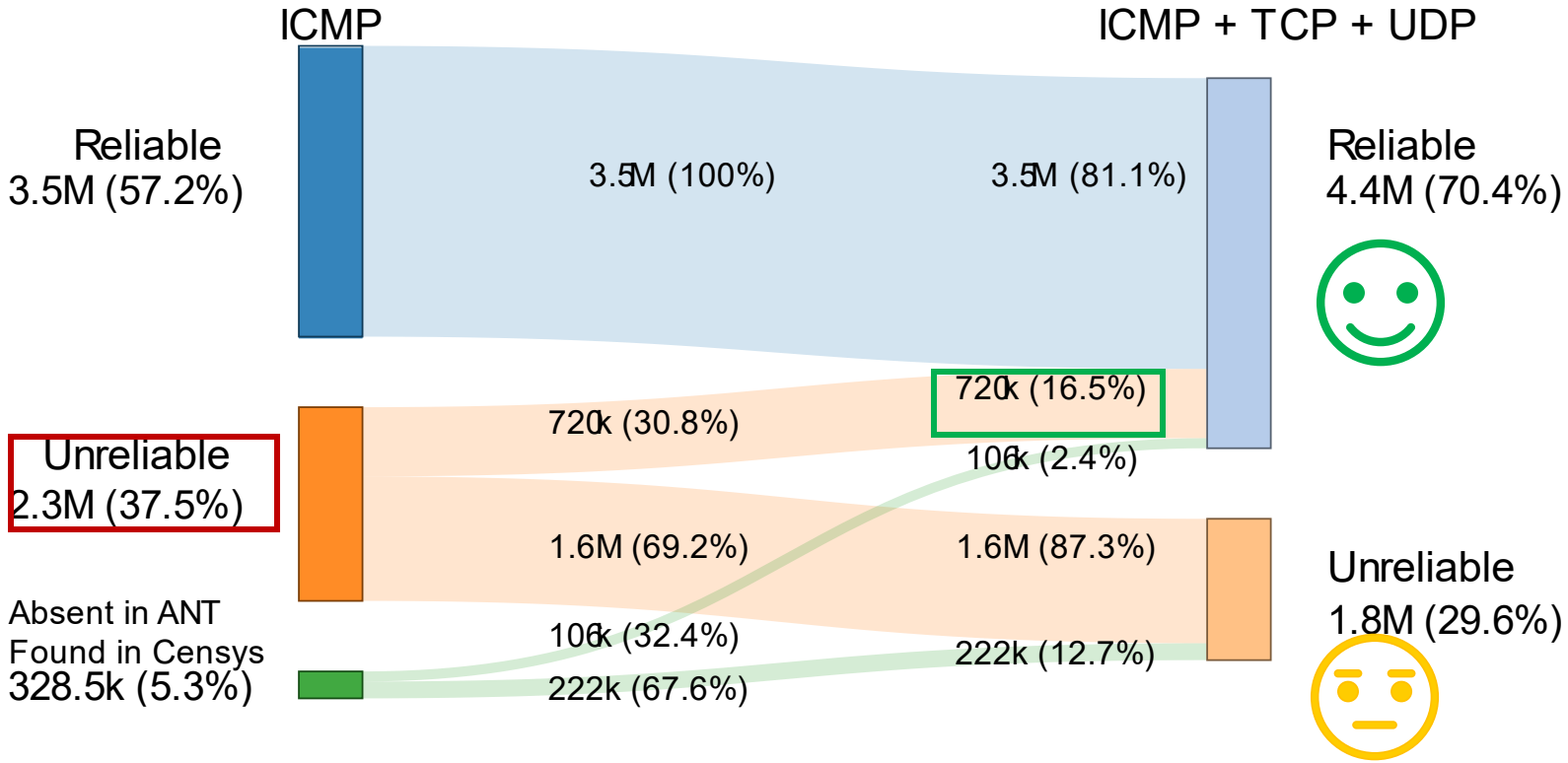
- Reliable /24 blocks (Availability >= 30%)
- TCP/UDP probes can discover previously unseen blocks
- Improve availability of existing /24 blocks



Georgia Tech

Findings: At the /24 block level

- Reliable /24 blocks (Availability >= 30%)
- TCP/UDP probes can discover previously unseen blocks
- Improve availability of existing /24 blocks

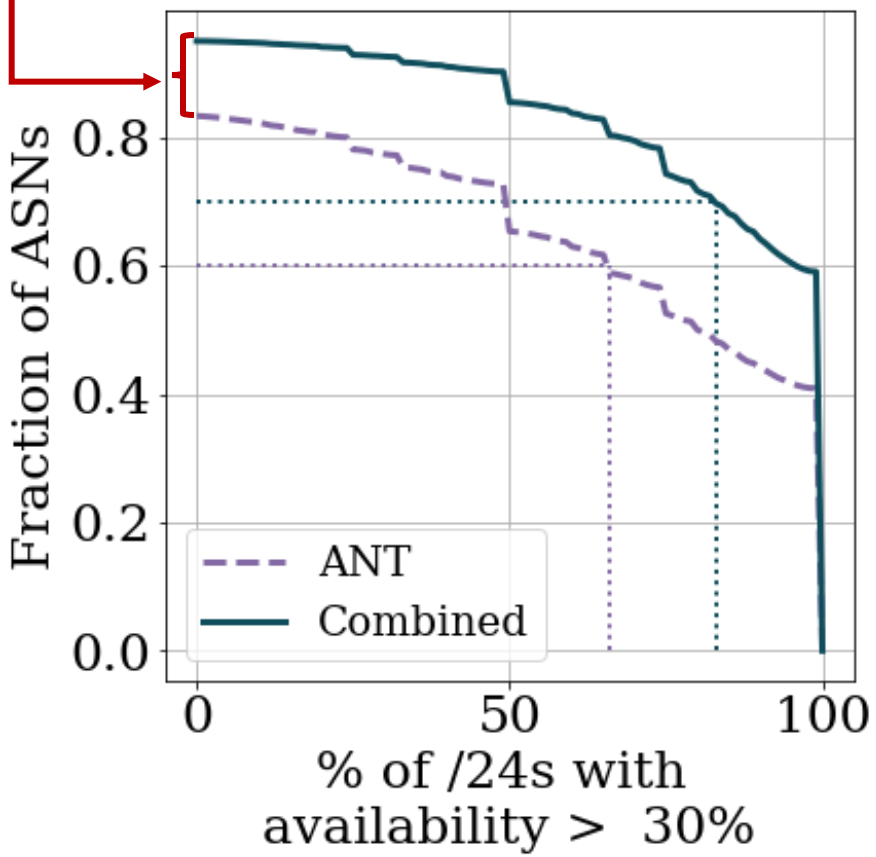


Adding TCP/UDP probes can reduce the number of unreliable /24 blocks by ~826k

Georgia Tech

Findings: At the AS level

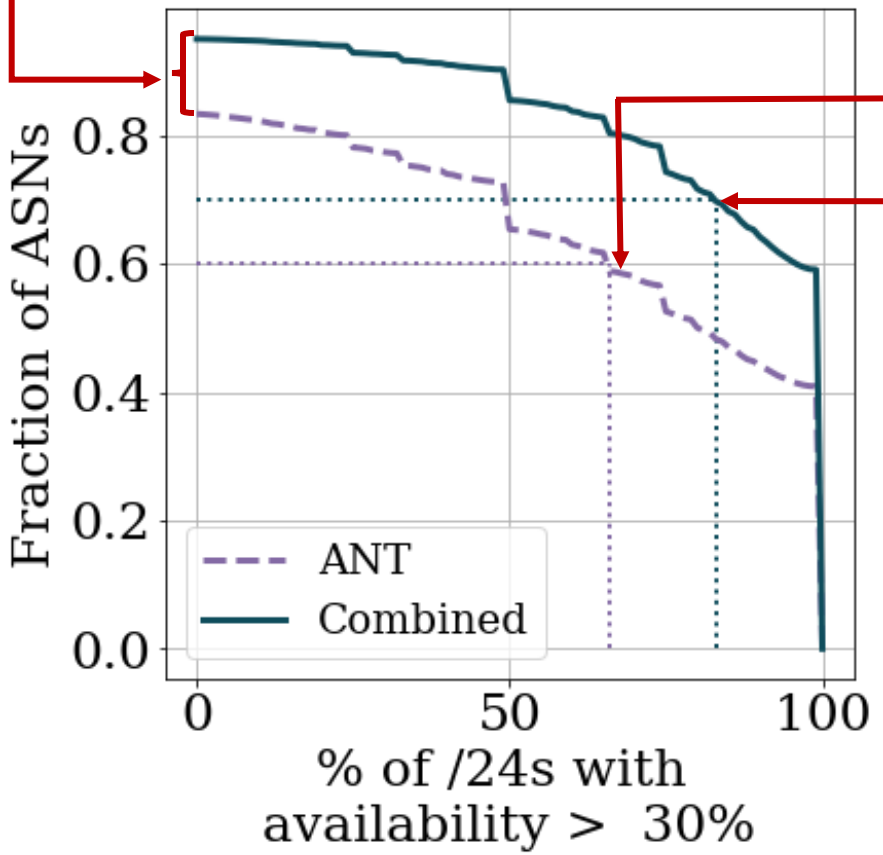
- Increase number of ASes covered by 1.8K



Georgia Tech

Findings: At the AS level

- Increase number of ASes covered by 1.8K
- Increase % of reliable blocks within an AS on avg from 66% to 83%



1.8K new ASes are discoverable when TCP/UDP probing is incorporated

Georgia Tech

Summary

- The availability of a host on average is higher when ICMP + TCP + UDP probing is used

Summary

- The availability of a host on average is higher when ICMP + TCP + UDP probing is used
- Reduce the number of unreliable /24 blocks by **~826k** (5% of all /24 blocks)

Summary

- The availability of a host on average is higher when ICMP + TCP + UDP probing is used
- Reduce the number of unreliable /24 blocks by **~826k** (5% of all /24 blocks)
- **1.8K new ASes** (~2% of assigned AS numbers) are discoverable

Limitations and Future work

- May incur overhead from additional probing

Summary

- The availability of a host on average is higher when ICMP + TCP + UDP probing is used
- Reduce the number of unreliable /24 blocks by **~826k** (5% of all /24 blocks)
- **1.8K new ASes** (~2% of assigned AS numbers) are discoverable

https://github.com/InetIntel/ioda-censys-isi

Limitations and Future work

- May incur overhead from additional probing
- Don't consider the impact of IP address churn

Summary

- The availability of a host on average is higher when ICMP + TCP + UDP probing is used
- Reduce the number of unreliable /24 blocks by **~826k** (5% of all /24 blocks)
- **1.8K new ASes** (~2% of assigned AS numbers) are discoverable

https://github.com/InetIntel/ioda-censys-isi

Questions?

- May incur overhead from additional probing
- Don't consider the impact of IP address churn
- We now know the potential for improvement in outage detection
  - To what degree can we see improvements when we implement TCP/UDP probing?

Summary

- The availability of a host on average is higher when ICMP + TCP + UDP probing is used
- Reduce the number of unreliable /24 blocks by **~826k** (5% of all /24 blocks)
- **1.8K new ASes** (~2% of assigned AS numbers) are discoverable

https://github.com/InetIntel/ioda-censys-isi

Thank You

Georgia Tech