Anycast Discovery: Daily mapping the Anycast landscape for enhanced Internet resilience

AIMS WORKSHOP 2024

RAFFAELE SOMMESE – REMI HENDRIKS

UNIVERSITY OF TWENTE

Our Team @ DACS

Remi Hendriks
PhD Candidate
University of Twente

Raffaele Sommese
Assistant Professor
University of Twente

Mattijs Jonker
Assistant Professor
University of Twente

**Roland Van Rijswijk**
Full Professor
University of Twente

Why Anycast Census?

▶ Anycast is one of the most effectively distribution and **resilience** technique used in the world.

▶ The working principle of Anycast relying on BGP makes it **opaque** to the rest of the Internet.

▶ Mapping the evolution of the adoption of Anycast at scale is fundamental for the analysis of the development of the global Internet.

A «new» approach to measure Anycast: MAnycast2

- Developed in an IMC2020 submission.

- We leverage the concept of using anycast to measure anycast:

  - Pinging a unicast address from an anycast network results in packet to be routed always to a SINGLE node, regardless of the source anycast site.

  - But pinging an anycast address from an anycast network results in packet routed to different MULTIPLE nodes, depending on the source site.

  - We leverage this behaviour to identify anycast networks.

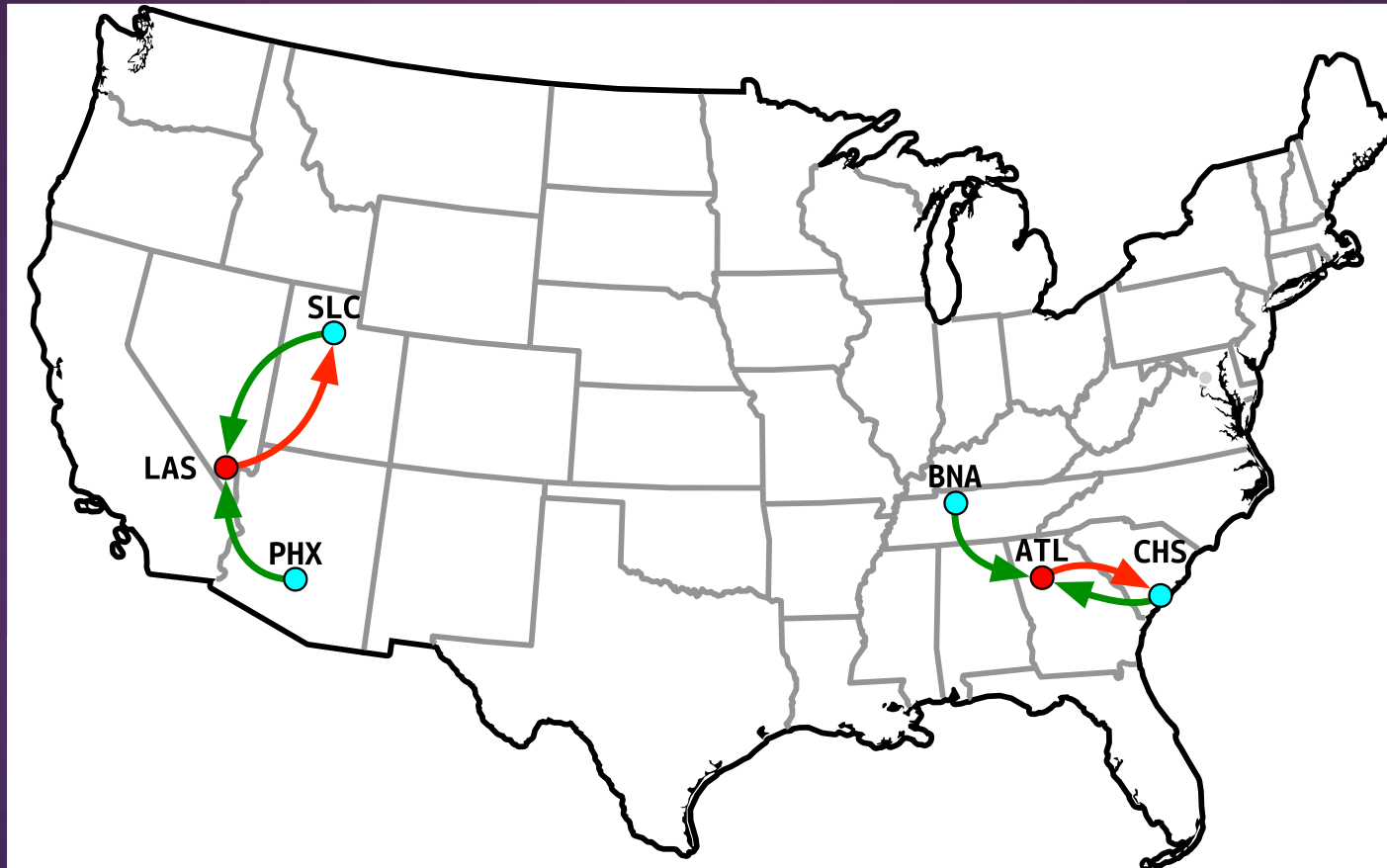# Anycast-based measuring Set-up
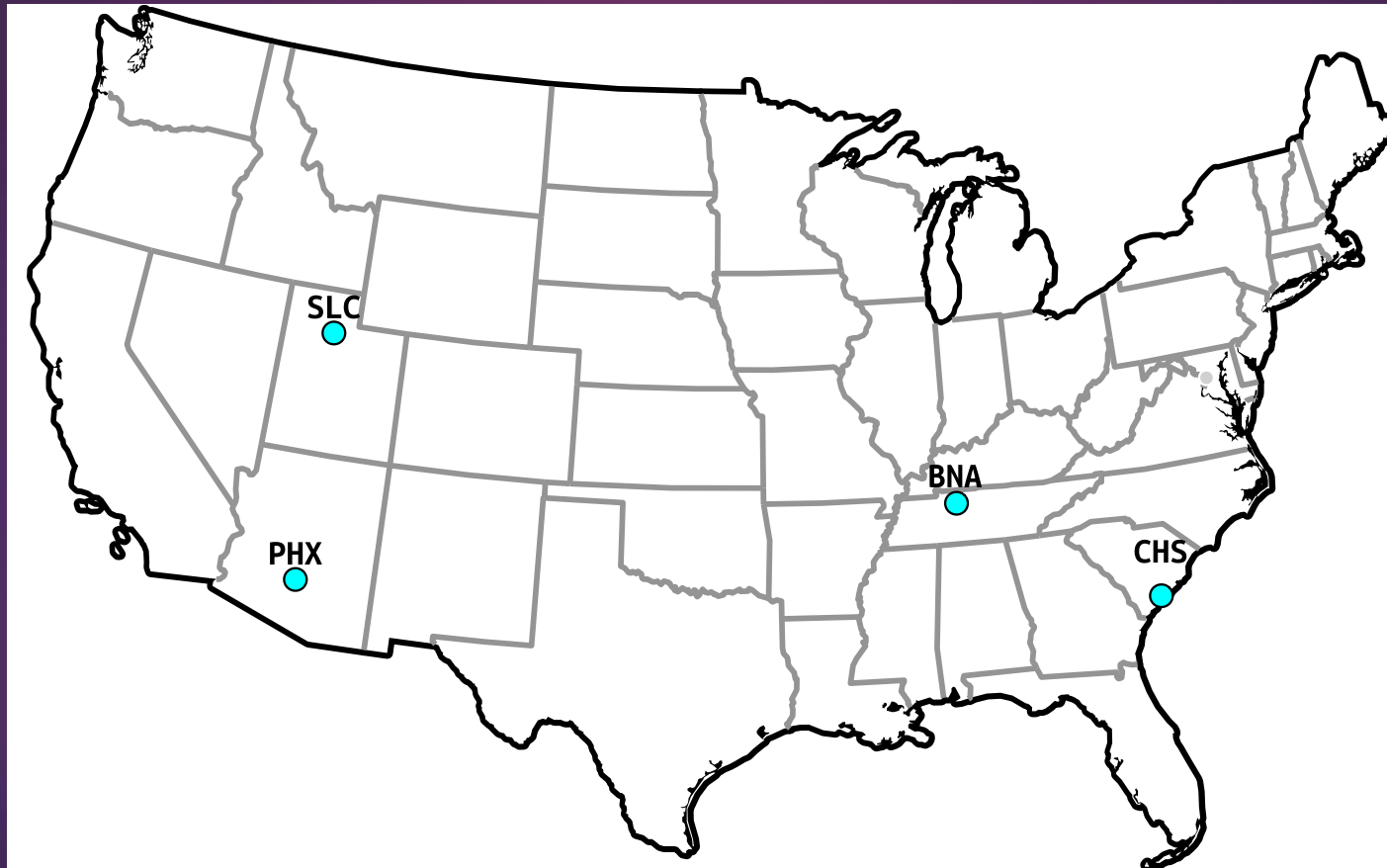
Anycast-based measuring Unicast
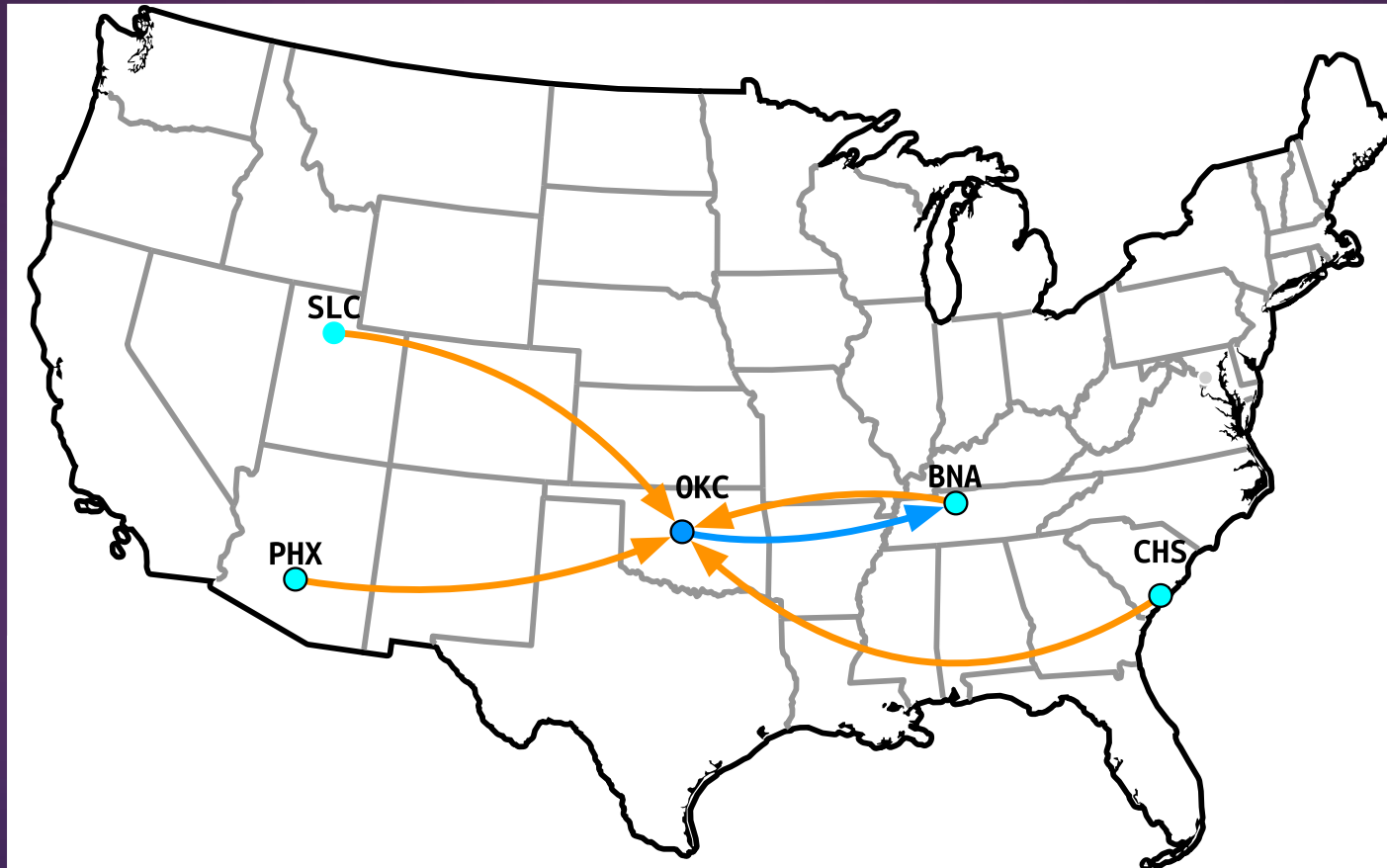
Anycast-based measuring Anycast

Pros and cons

▶ Pros:
  - o **Low probing-cost** (suitable for Internet scale measuring)
 - o **Low FP rate** (rarely misclassifies anycast as unicast)

▶ Cons:
  - o **Considerable FP rate** (falsely classifying unicast as anycast)
  - o **No geolocation** of sites (only detection & enumeration)

GCD measuring (iGreedy)

- GCD (**Great Circle Distance**)
- Latency-based measuring using
  speed-of-light violations
- Pros:
  - Low FP/FN rate (**highly accurate**)
  - **Geolocation possible**
- Cons:
  - Requires large measuring platform
  - High probing cost

  (unsuitable for Internet scale)
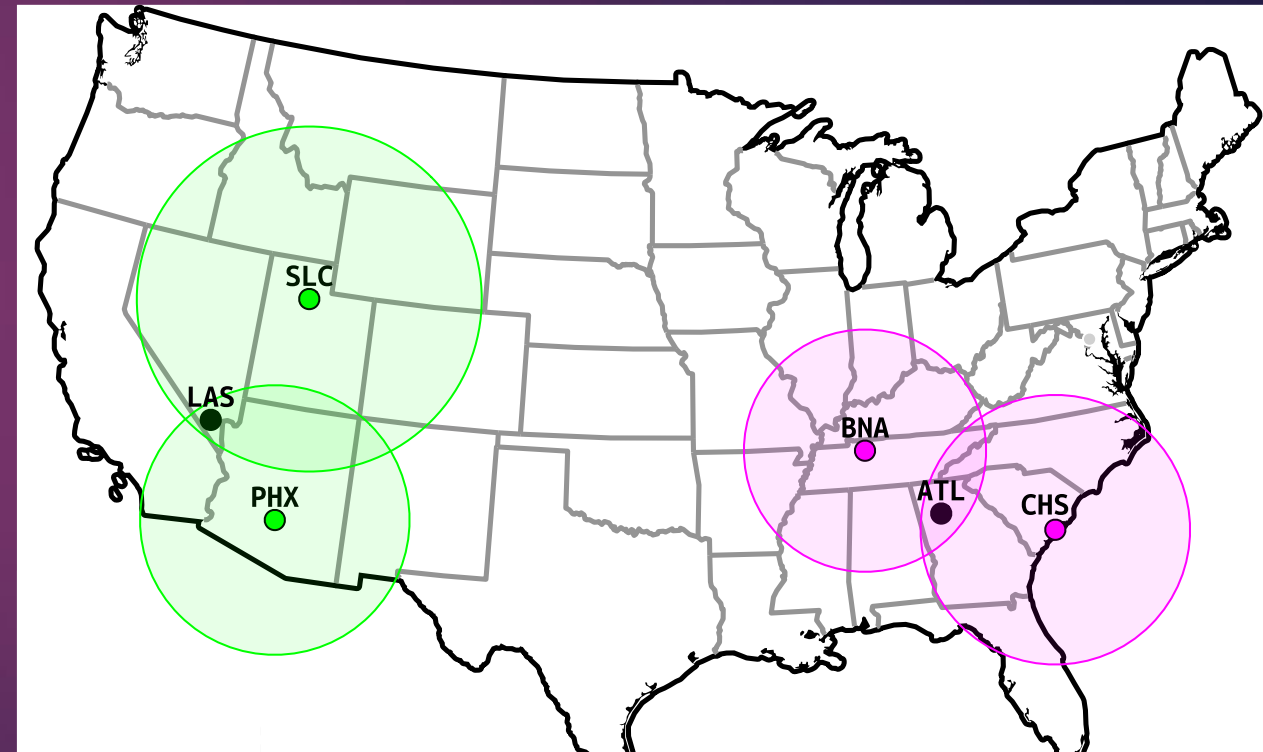
Combining the two

- **Perform anycast-based census**
  - Input: Internet wide hitlist (USC/ISI IPv4, TUM + OpenINTEL IPv6) (10^6)
  - Output: set of "*anycast targets*" (includes TPs and FPs)
- **Perform GCD-based measurement**
  - Input: "*anycast targets*" (10^4)
 - *Output: Anycast prefixes + enumeration + locations individual sites*

Realizing daily census

▶ Developed anycast-based measuring tool

  ▢ Scalable

  ▢ Efficient

  ▢ Robust

  ▢ Responsible

  ▢ Increased precision (reduced number of FPs)

  ▢ Increased coverage (IPv4, IPv6, TCP, DNS, ICMP, ..)

▶ Deployed using Vultr (32 locations)

Ark/Scamper

- GCD measurements with Ark VPs
  - ~ 140 for ICMPv4
- Implemented using Scamper tool

- Able to enumerate/geolocate majority of sites
  - For small deployments
- Able to enumerate/geolocate up to 50 anycast sites
 - For large deployments

Pipeline
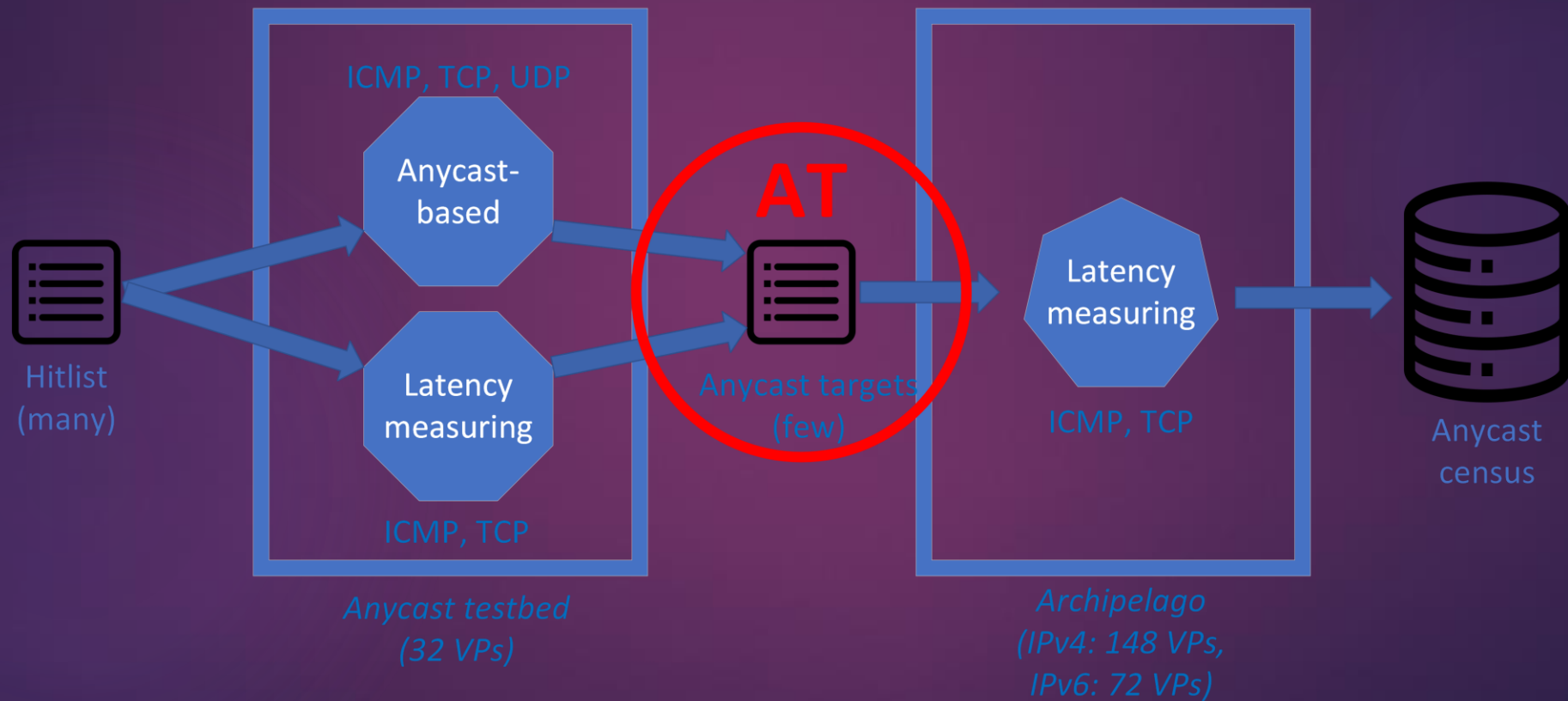
Some numbers

- ~12.3k anycasted /24s (783 ASes)
- ~4.9k anycasted /48s (364 ASes)
- 249 ASes found to anycast both IPv4 and IPv6

AS	Organization	IPv4	IPv6
396982	Google Cloud	3,345	3
13335	Cloudflare	3,131	162
16509	Amazon	1,235	86
54113	Fastly	438	56
15169	Google	282	6
209242	Cloudflare Spectrum	234	2,836
19551	Incapsula	2	292
12041	Afilias	222	207
44273	GoDaddy	31	122

Table 6: Largest ASes originating anycast prefixes for IPv4 and IPv6.

Some numbers

Figure 6: MAnycastR detection of anycast candidates for ICMPv4, TCPv4, and UDPv4.
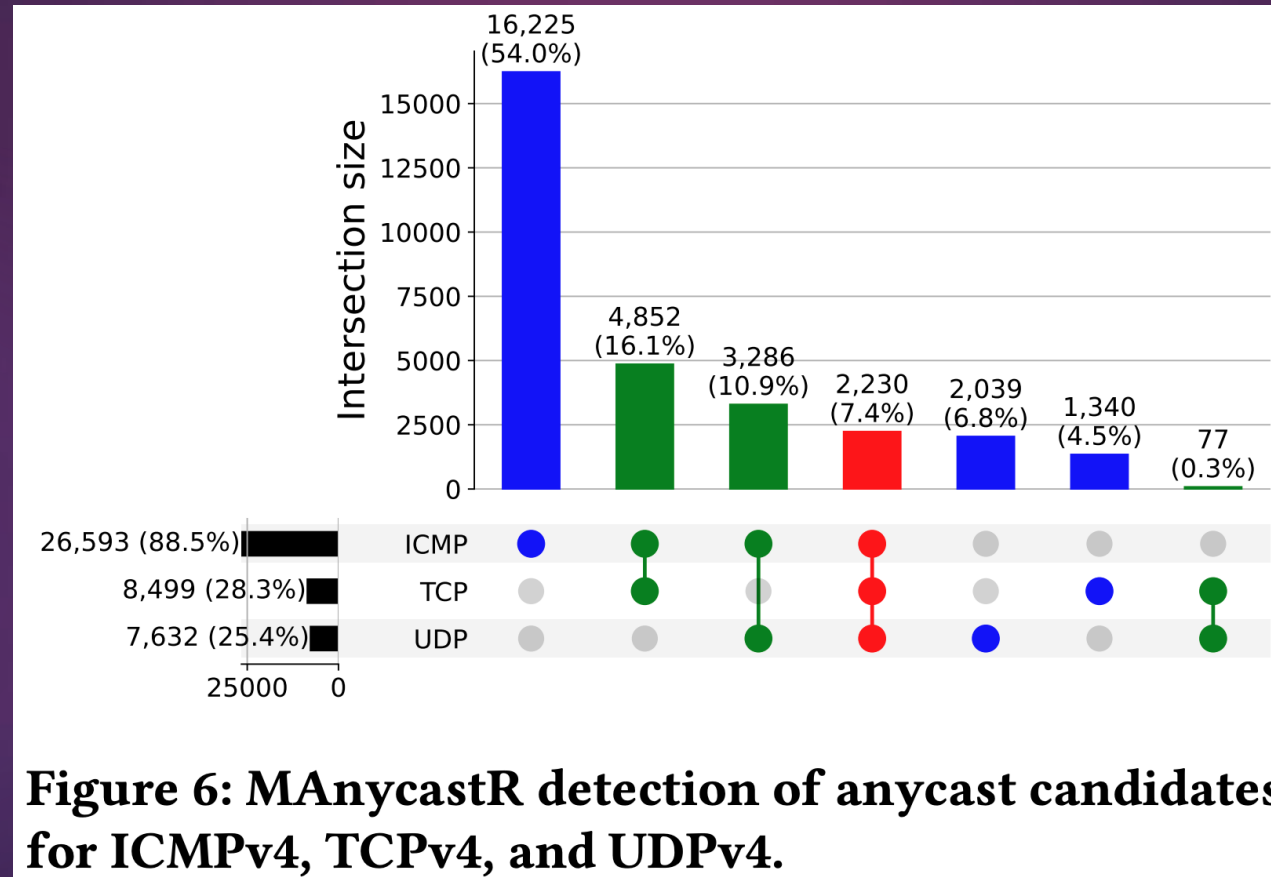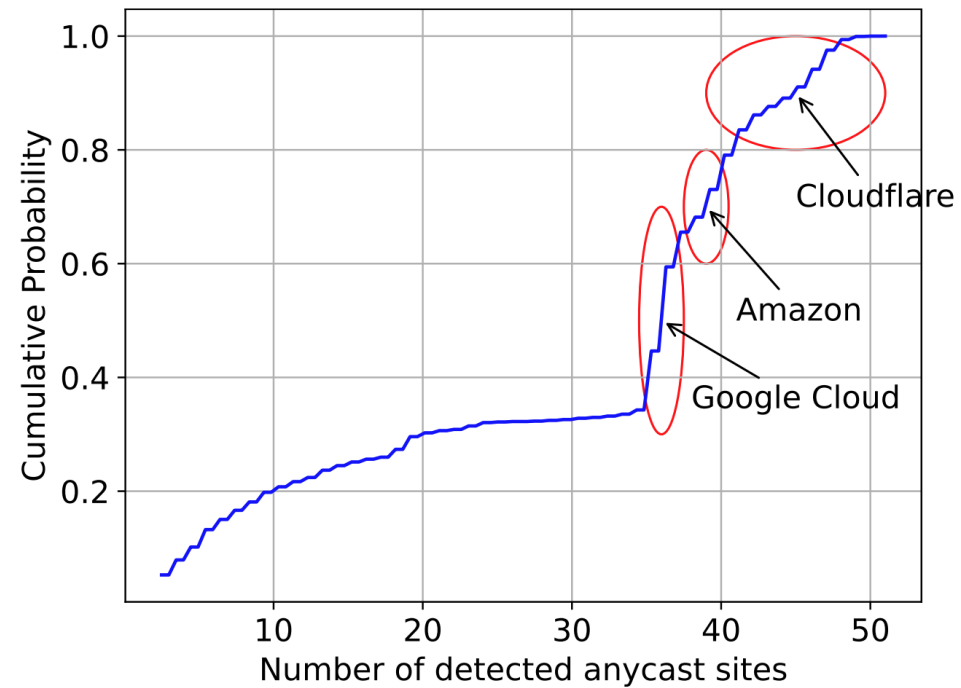
Some numbers

Figure 9: Number of anycast sites detected per /24 prefix with GCD using latency data from Ark VPs.

Conclusion

▶ Created **responsible**, **scalable**, **accurate** anycast measuring tool

▶ Uses **Ark** for latency measurements toward anycast prefixes

▶ A **daily** census of anycast

▶ We hope other researchers will use this data towards assessment of the resilience and development of Internet

Future

- Deploy Scamper on Vultr VMs
  - Perform GCD measurements using both Ark and Vultr
- Refining/improving pipeline
  - E.g., canary outage detection

- Public release of daily census (currently 4 month of census data)
- Public release of anycast-based measuring tool
- Allowing researchers to perform live measurements (web-interface/API)

- Longitudinal analysis of anycast

Questions?

- remi.hendriks@utwente.nl
- r.sommese@utwente.nl