LACeS

An Open, Fast, Responsible, and Efficient **L**ongitudinal **A**nycast **Ce**nsus **S**ystem

IMC'25 - Madison, Wisconsin, USA

October 28, 2025

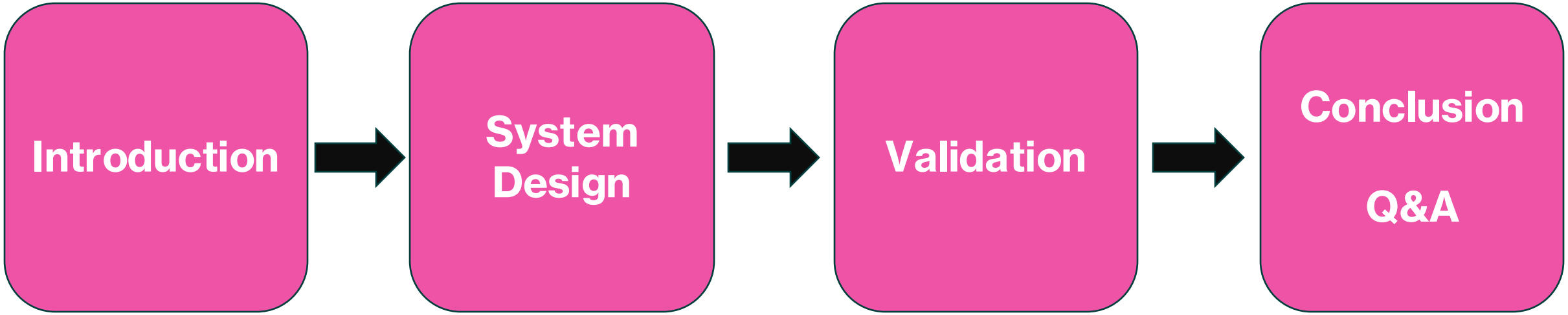
Remi Hendriks, Matthew Luckie, Mattijs Jonker,

Raffaele Sommese, Roland van Rijswijk-Deij

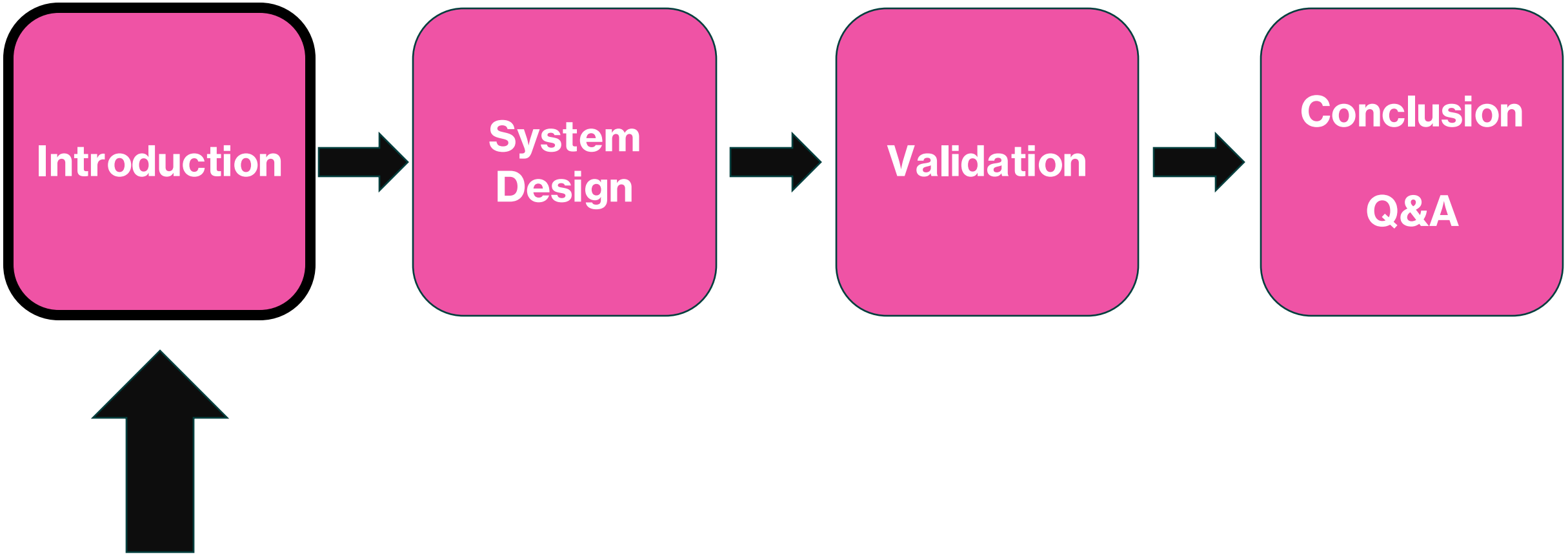
UNIVERSITY OF TWENTE.

caida

Structure

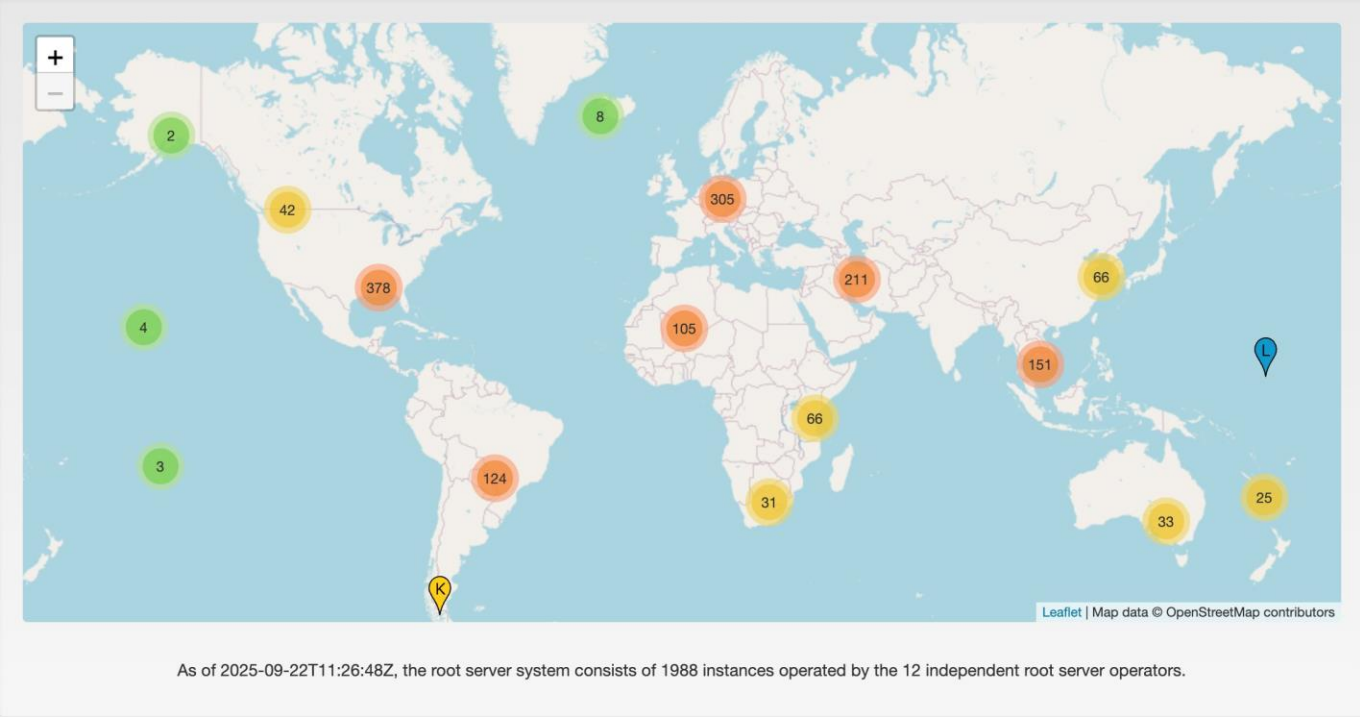


Structure



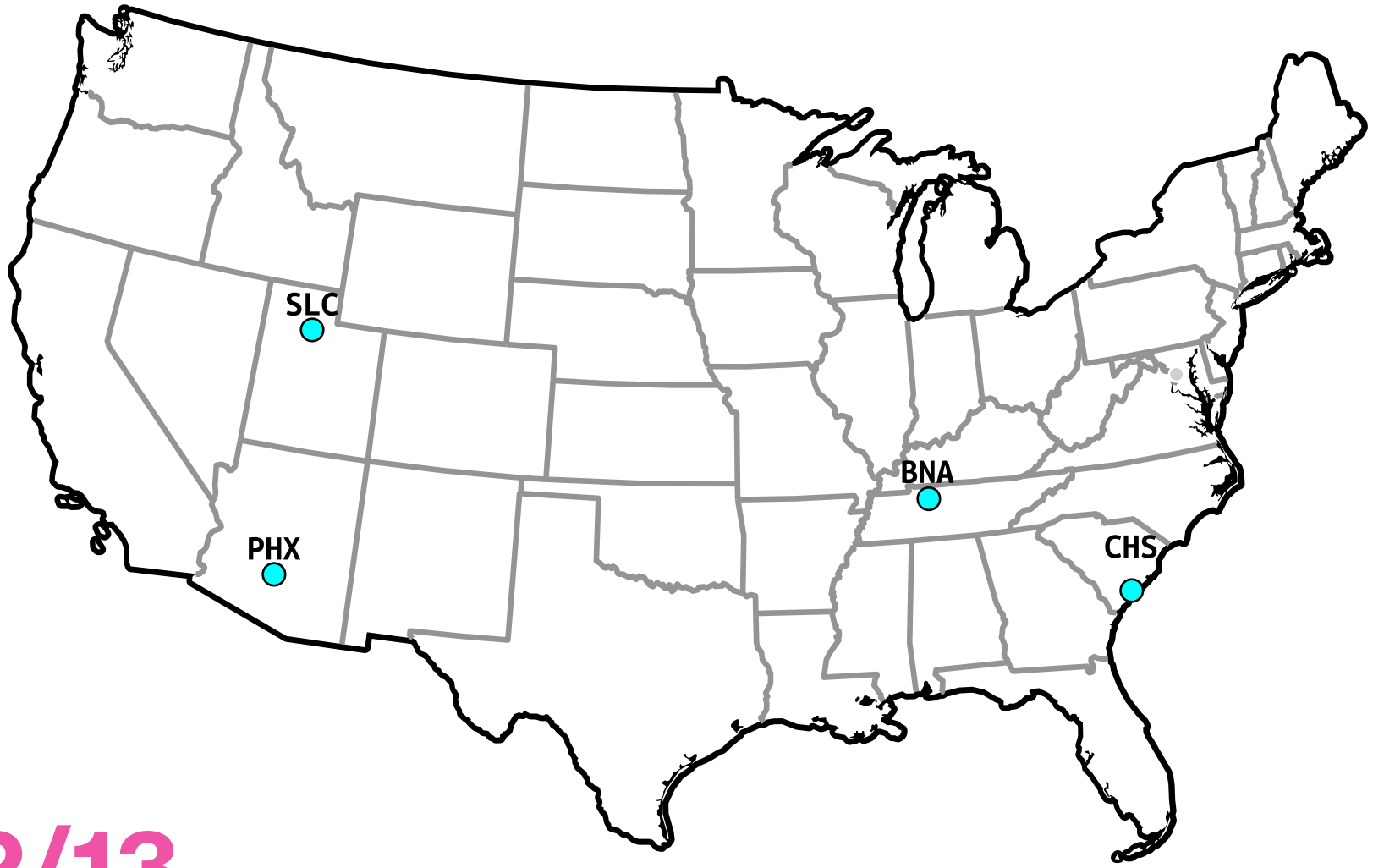
Introduction
Motivation

- Anycast: Making an IP address available in multiple locations (sites)
  - *How*? - Announcing an IP prefix at multiple locations using BGP
  - *Why?* - Distribute load, improve resilience, reduce latency
- Anycast is widely used
  - Critical Internet infrastructure (*e.g.*, DNS)
  - Content Delivery Networks
- **Problem**: anycast is opaque
  - Unclear if an address/service is anycasted
  - If yes, where is it located?
- **Solution**: Anycast census



Source: root-servers.org
"All 12 DNS root-letters are deployed using anycast; combined totaling 1,988 sites"

Background
MAnycast² *[1]*

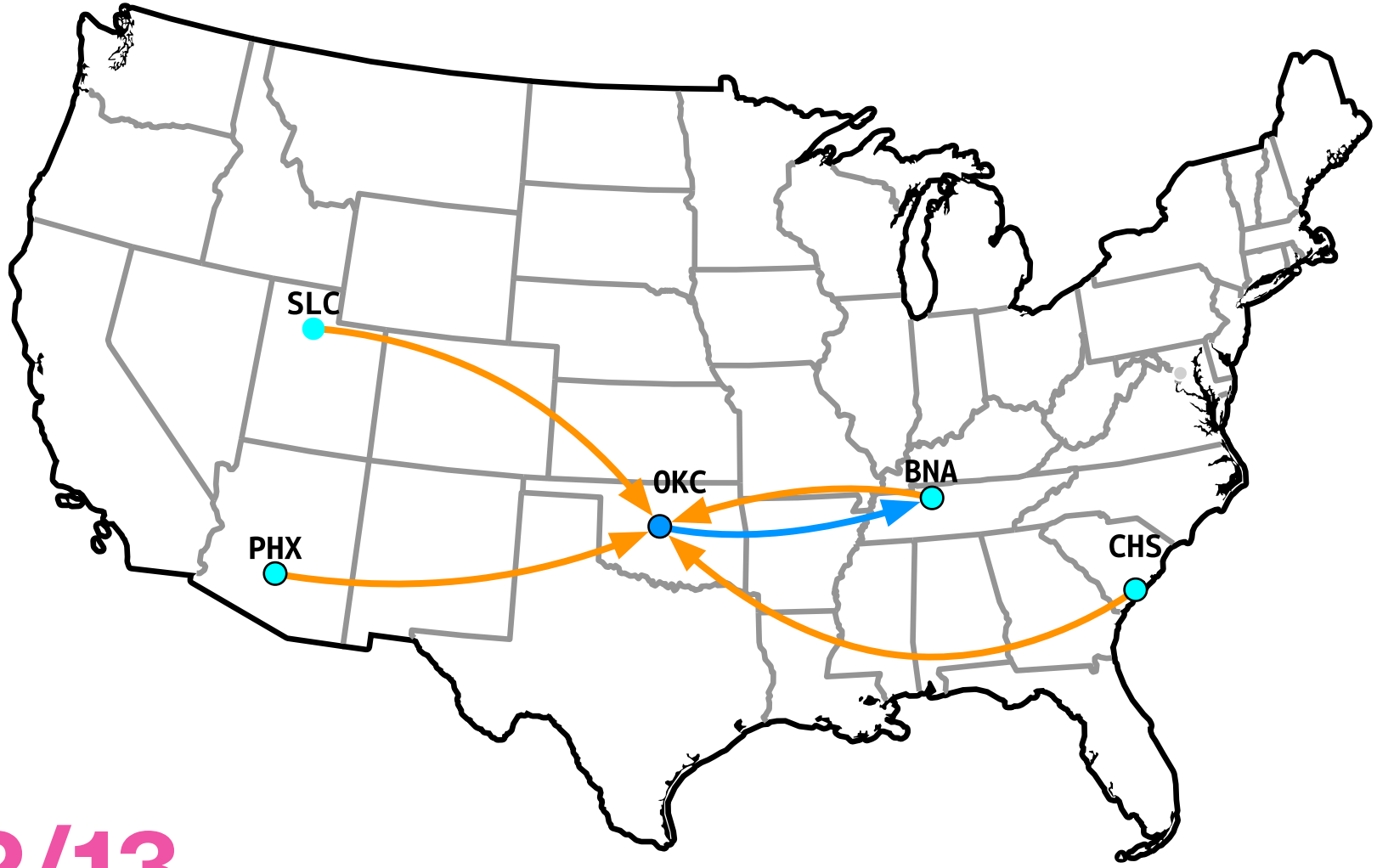


Using anycast (light blue)

to detect anycast.

[1] Sommese, Raffaele, et al. "MAnycast²: Using Anycast to Measure Anycast" IMC'20

Background
MAnycast² *[1]*



Using anycast (light blue)

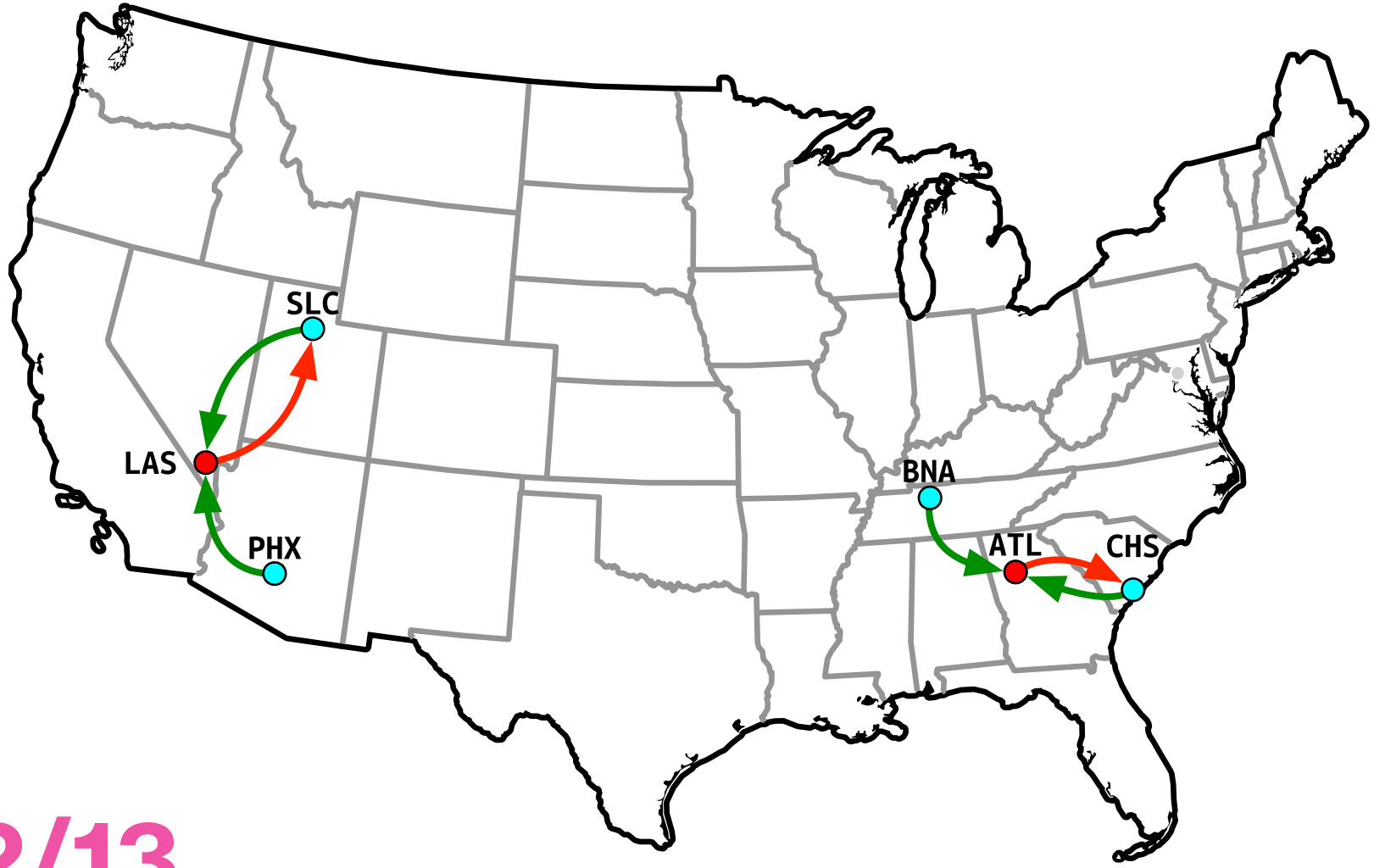
to detect anycast.

OKC (blue) received at

one Vantage Point (VP) -> unicast

[1] Sommese, Raffaele, et al. "MAnycast²: Using Anycast to Measure Anycast" IMC'20

Background
MAnycast² *[1]*



Using anycast (light blue)

to detect anycast.

OKC (blue) received at

one Vantage Point (VP) -> unicast

ATL, LAS (red) received at

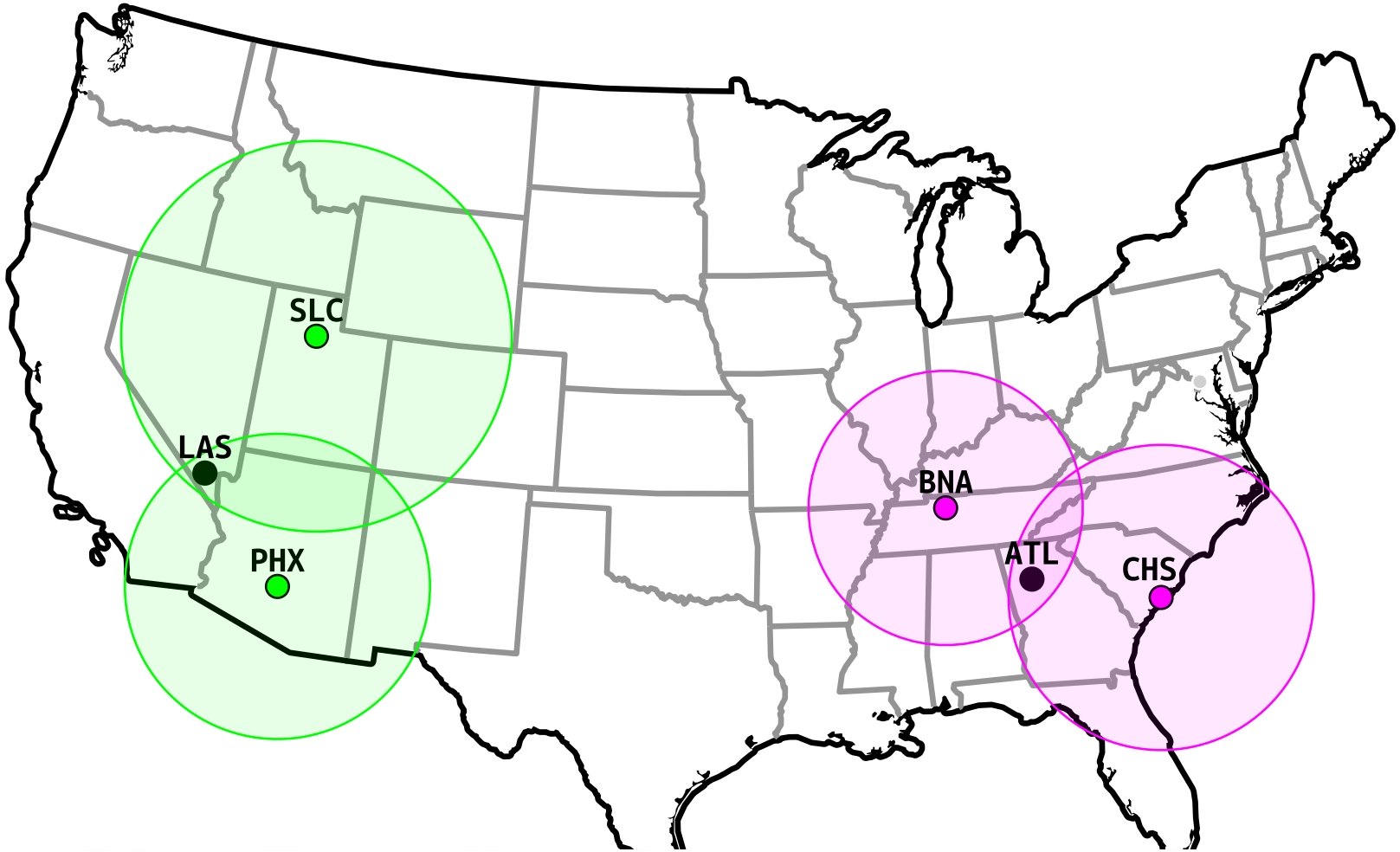
more than one VP -> anycast

[1] Sommese, Raffaele, et al. "MAnycast²: Using Anycast to Measure Anycast" IMC'20

Background
iGreedy *[2]*

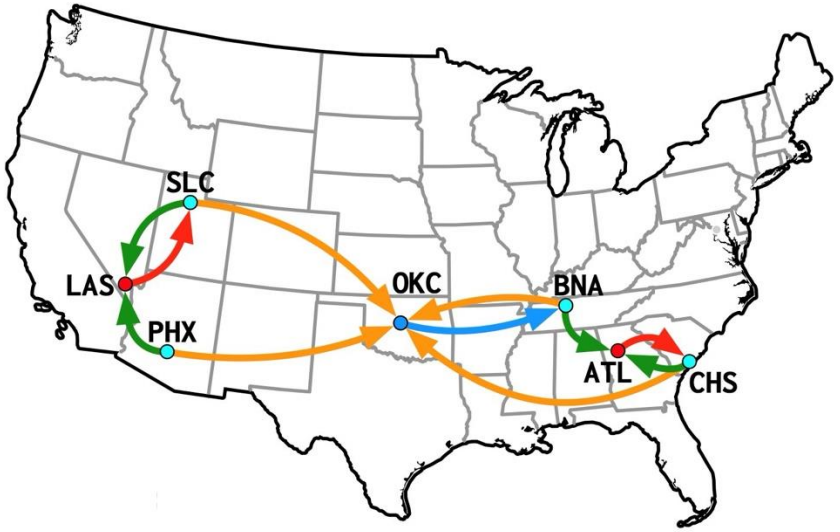
Using latency measurements

1) Detect anycast using speed-of-light violations
2) Enumerate (MIS)
3) Geolocate (largest city)



[2] Cicalese, Danilo, et al. "A Fistful of Pings: Accurate and Lightweight Anycast Enumeration and Geolocation" INFOCOM'15

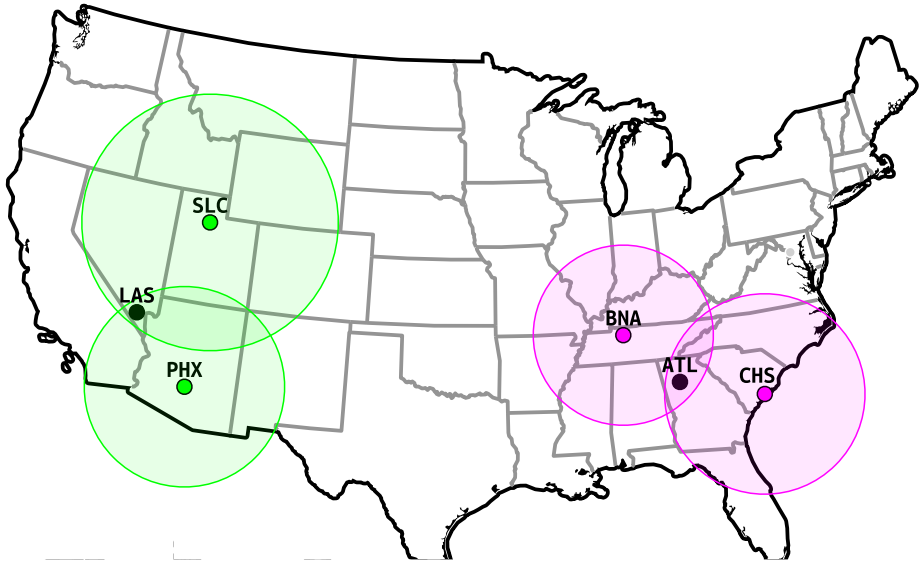
Background



MAnycast²

Pros: Lightweight detection

Cons: No geolocation & suffers from False Positives



iGreedy

Pros: Accurate & Geolocation

Cons: Requires large measurement platform

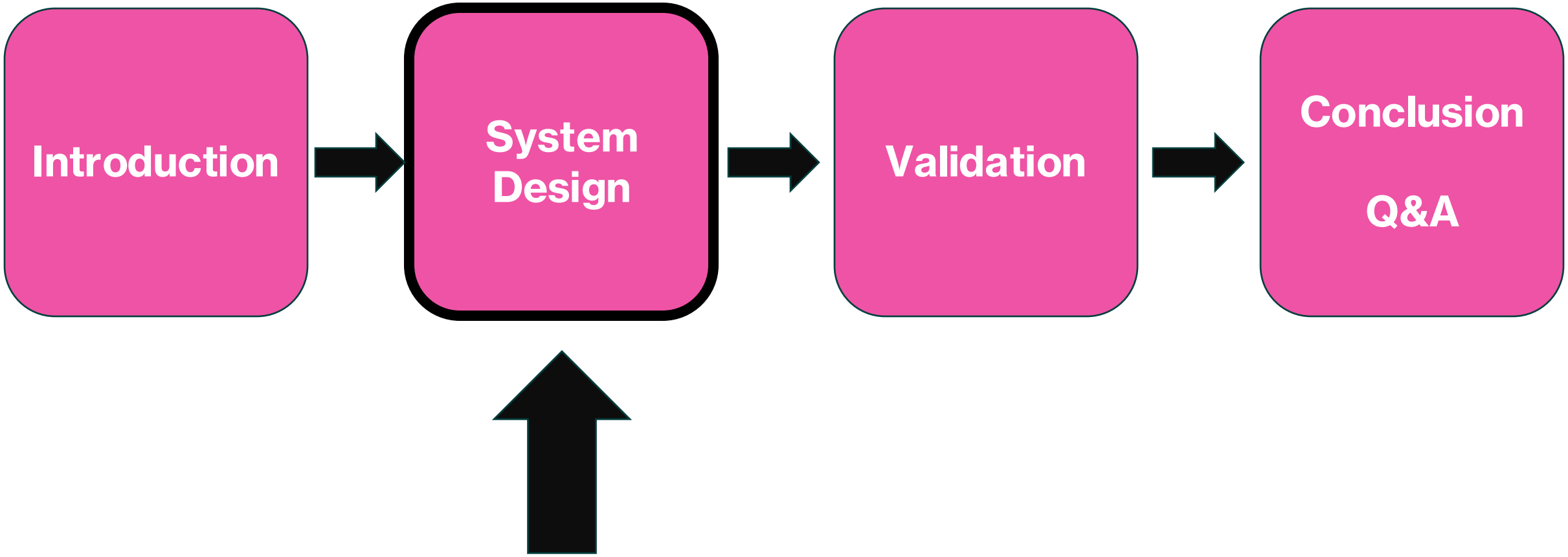
State-of-the-art

- Cicalese et al. released monthly iGreedy censuses [3] (Dec'15 – May'16) [discontinued]
- Commercial providers:
  - BGPTools provides a dataset of anycast prefixes (anycast-based like MAnycast[2])
  - IPInfo, MaxMind, IP2location, … have anycast flags (latency-based like iGreedy)
  - None provide enumeration and geolocation (only detection)

- We provide a non-commercial, publicly accessible, daily anycast census
  - Detection & enumeration & geolocation

[2] Cicalese, Danilo, et al. "A Longitudinal Study of IP Anycast" SIGCOMM'18

Structure

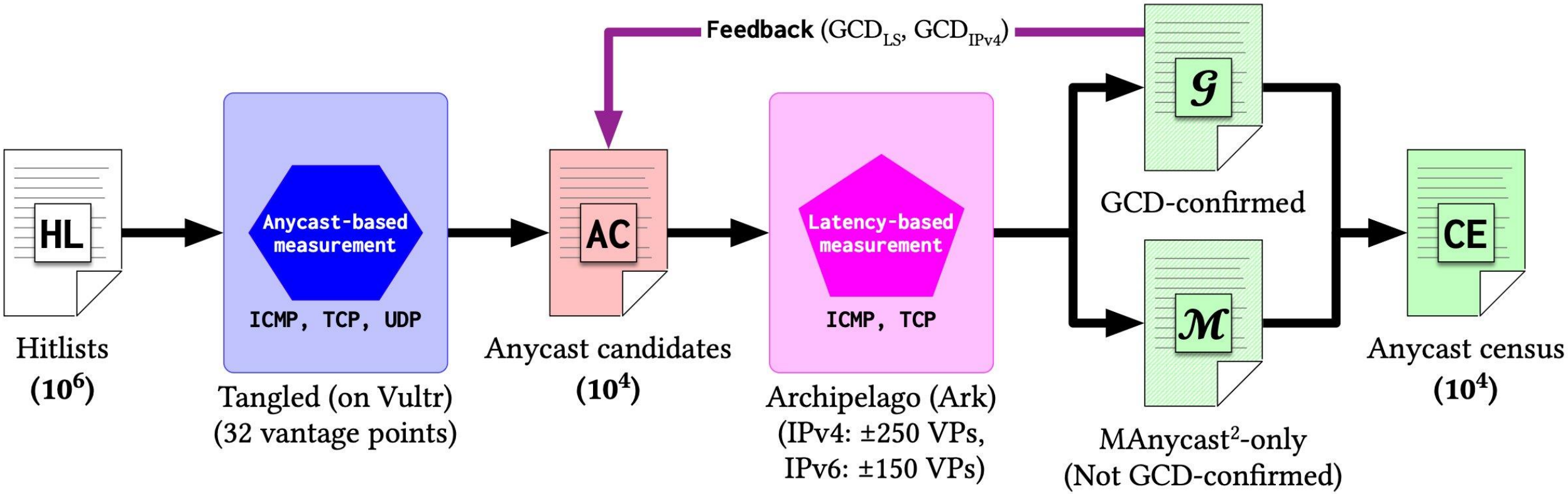


Requirements

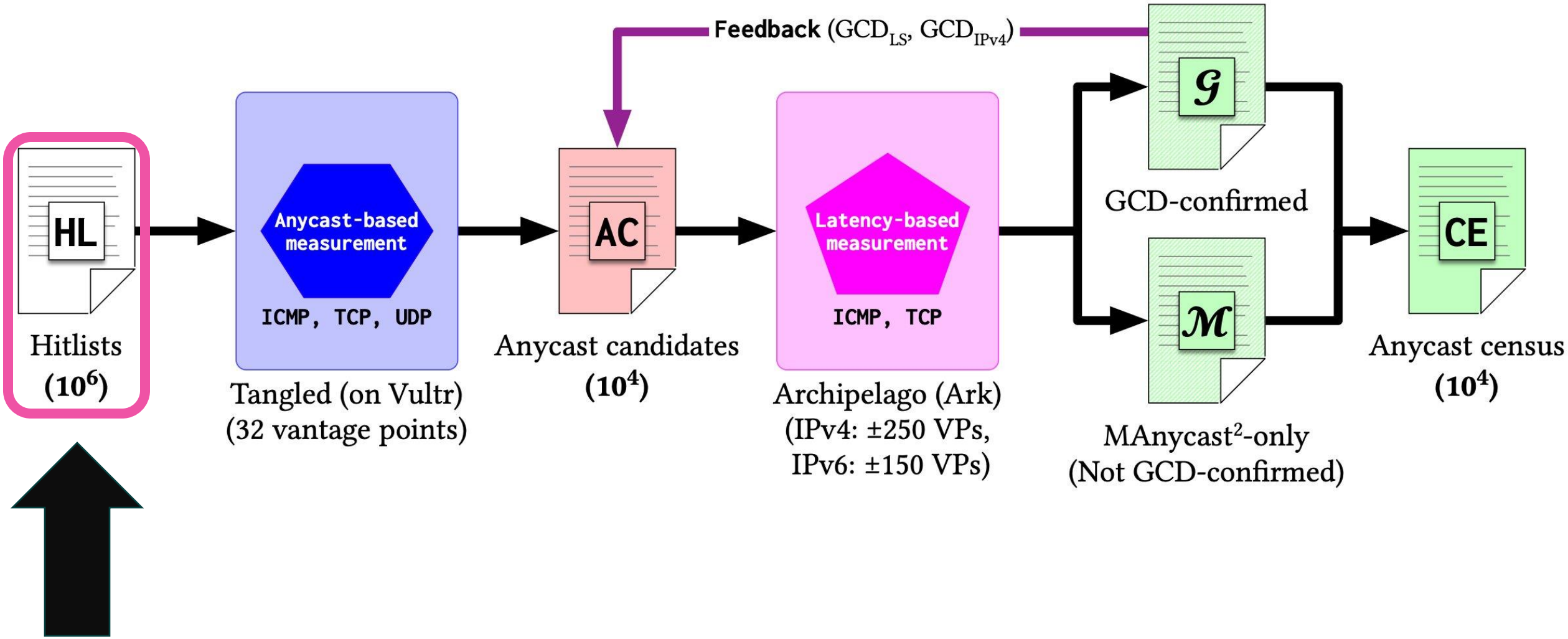
- **Accurate**. Minimize FPs and FNs
- **Precise**. Longitudinal consistency
- **Responsible**. Low impact on Internet and measurement infrastructure
- **Multi-Protocol**. Improve coverage with IPv6, TCP, DNS/UDP probing
- **Performant.** Allow for hitlist-based daily scanning
- **Reproducible**. Others can perform censuses using LACeS

[2] Cicalese, Danilo, et al. "A Longitudinal Study of IP Anycast" SIGCOMM'18

Census Pipeline

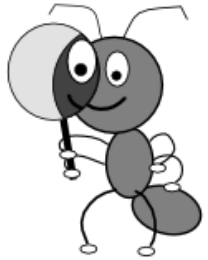


Census Pipeline



Methodology
Hitlists

- Targets
  - IPv4: USC/ISI ANT *[4]* (ICMP), zmap-based (TCP), OpenINTEL *[5]* infra:ns (DNS)
  - IPv6: TUM's public IPv6 hitlist *[6]* + OpenINTEL AAAA record addresses
- Probing strategy
  - One target per smallest propagated BGP prefix size (IPv4: /24, IPv6: /48) as representative
- Coverage:
  - 4.3M /24s & 5.6M /48s covering 7.3k ASes [hitlists updated regularly]

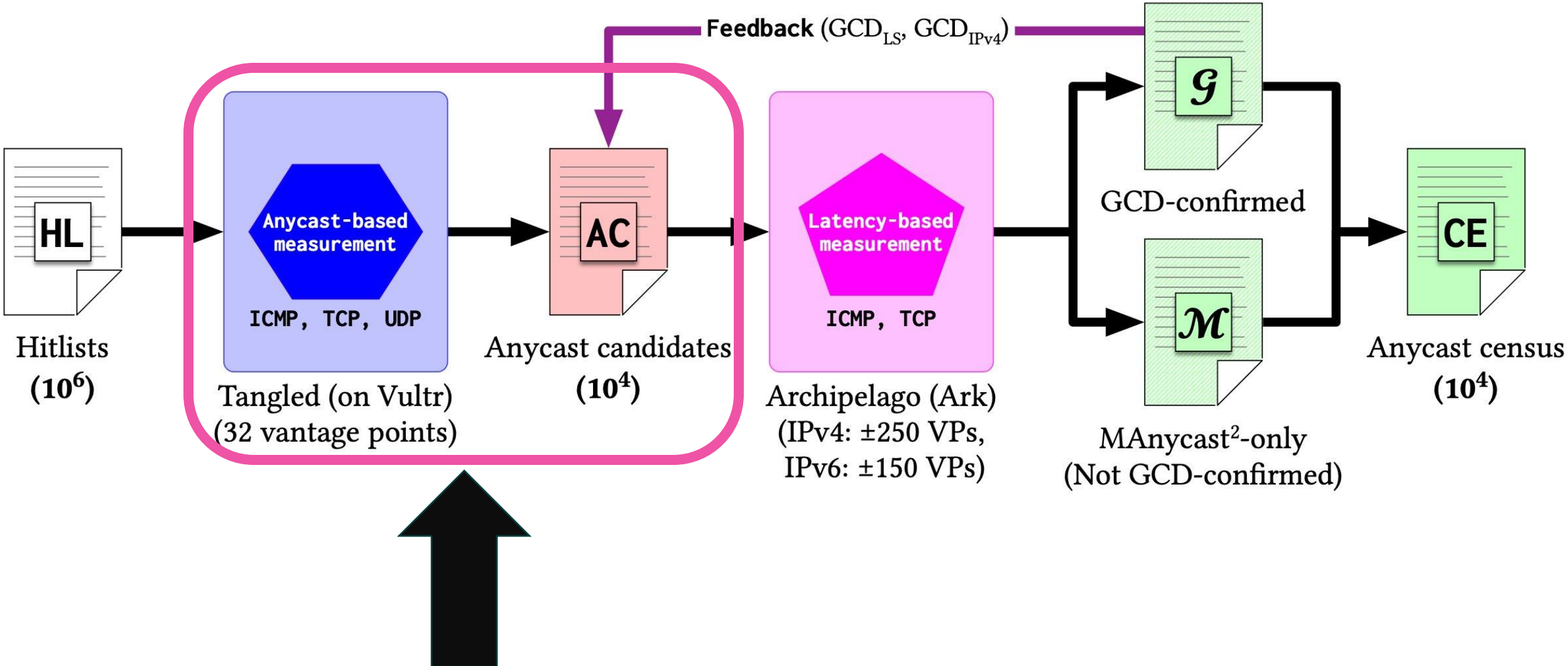


Technical University of Munich

Open INTEL

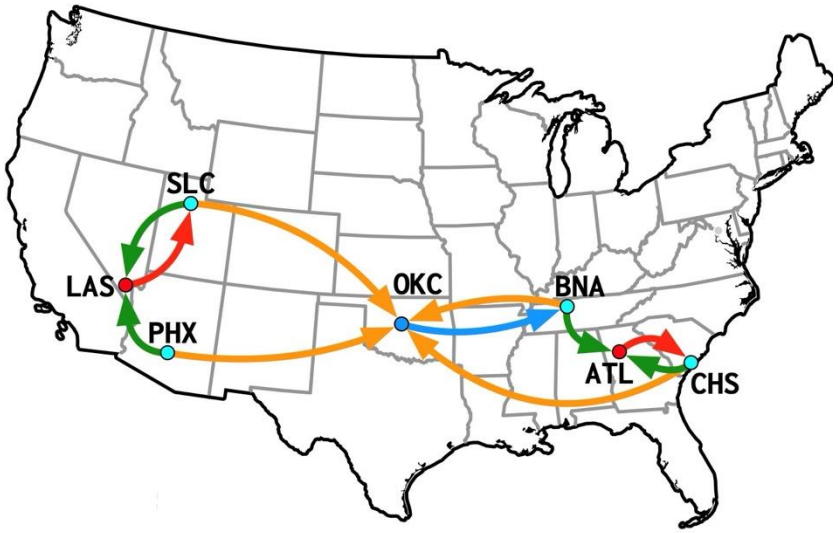
[4] USC/LANDER project: Internet Addresses Survey Dataset. isi.edu/ant/lander
[5] A High-Performance, Scalable infrastructure for Large-Scale Active DNS Measurements. van Rijswijk-Deij et al. IEEE Journal '16
[6] Clusters in the Expanse: Understanding and Unbiasing IPv6 hitlists. Gasser et al. IMC '18

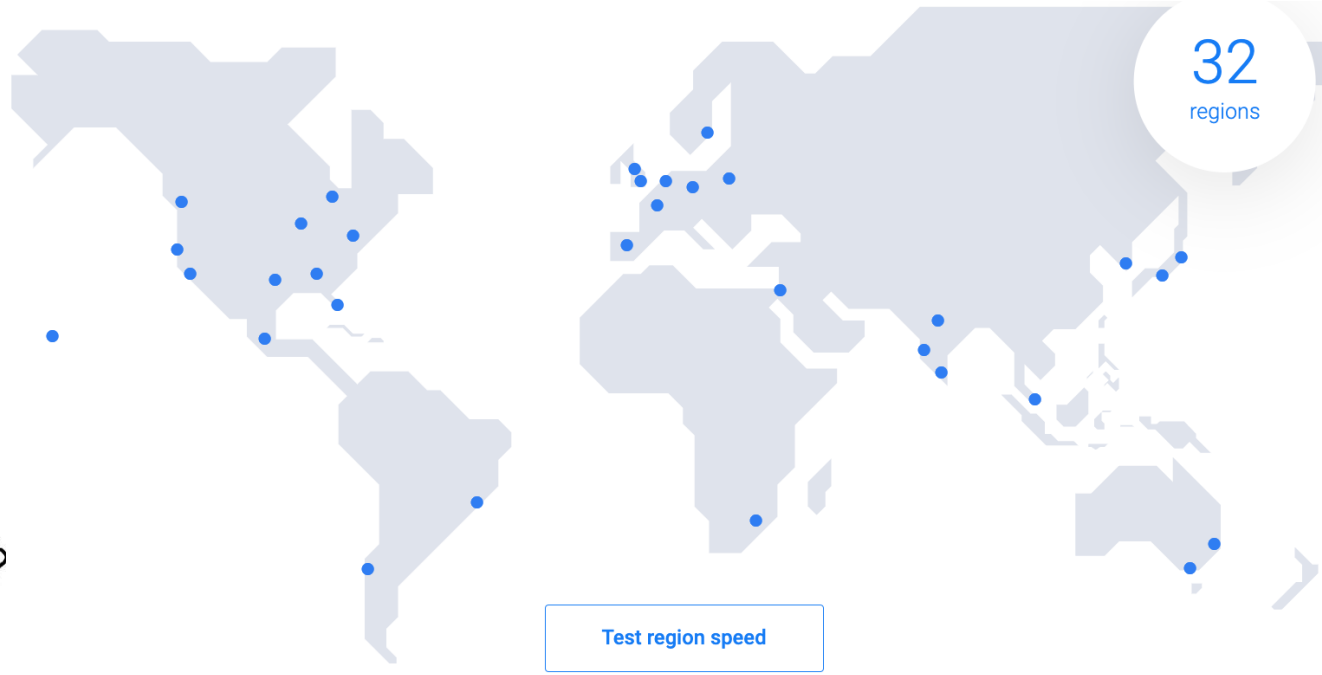
Census Pipeline



Methodology
Anycast-based

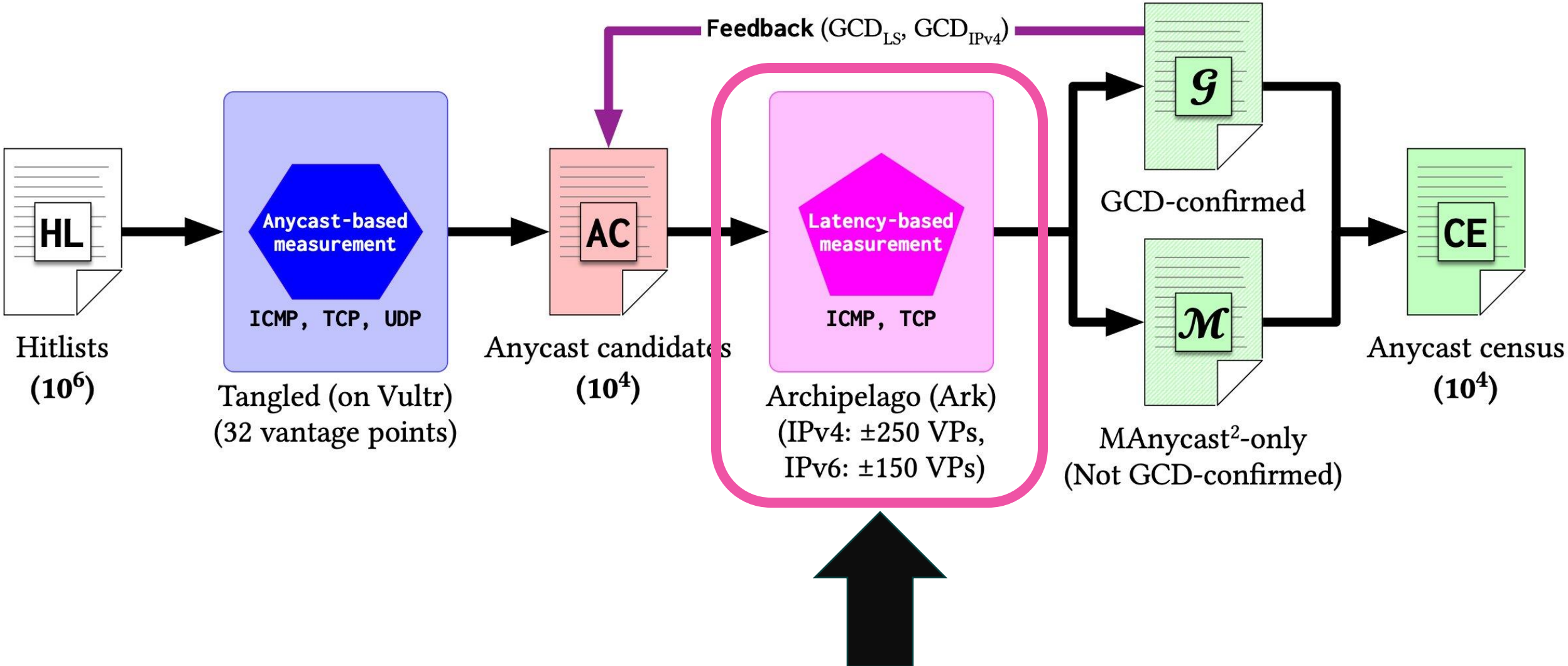
- Targets full hitlists (10^6)
- Similar methodology to MAnycast^2
- Performed using TANGLED (32 sites)
- **Output**: Anycast Candidates (10^4)
  - Anycast prefixes + MAnycast^2 False Positives





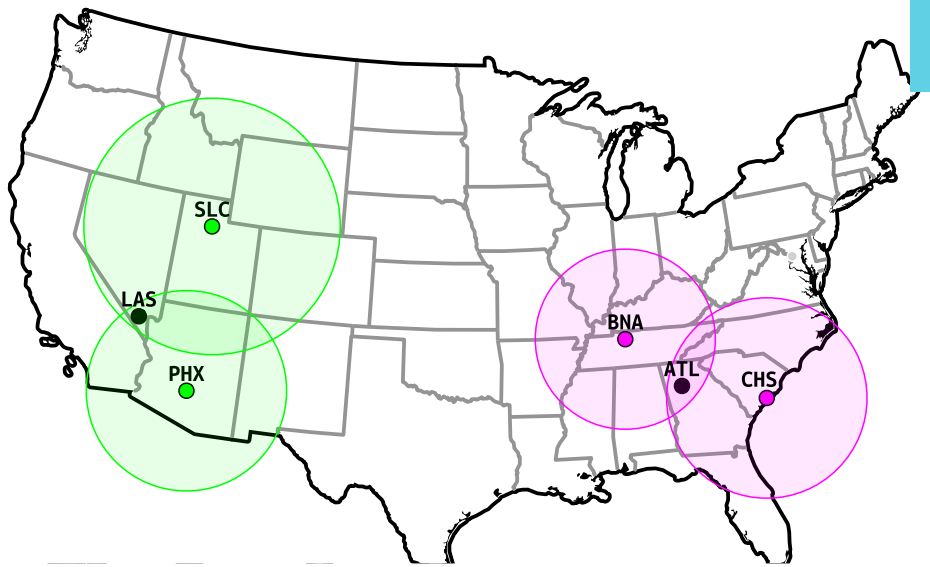
https://vultr.com/features/datacenter-regions/
TANGLED deployed on Vultr

Census Pipeline



Methodology
Latency-based

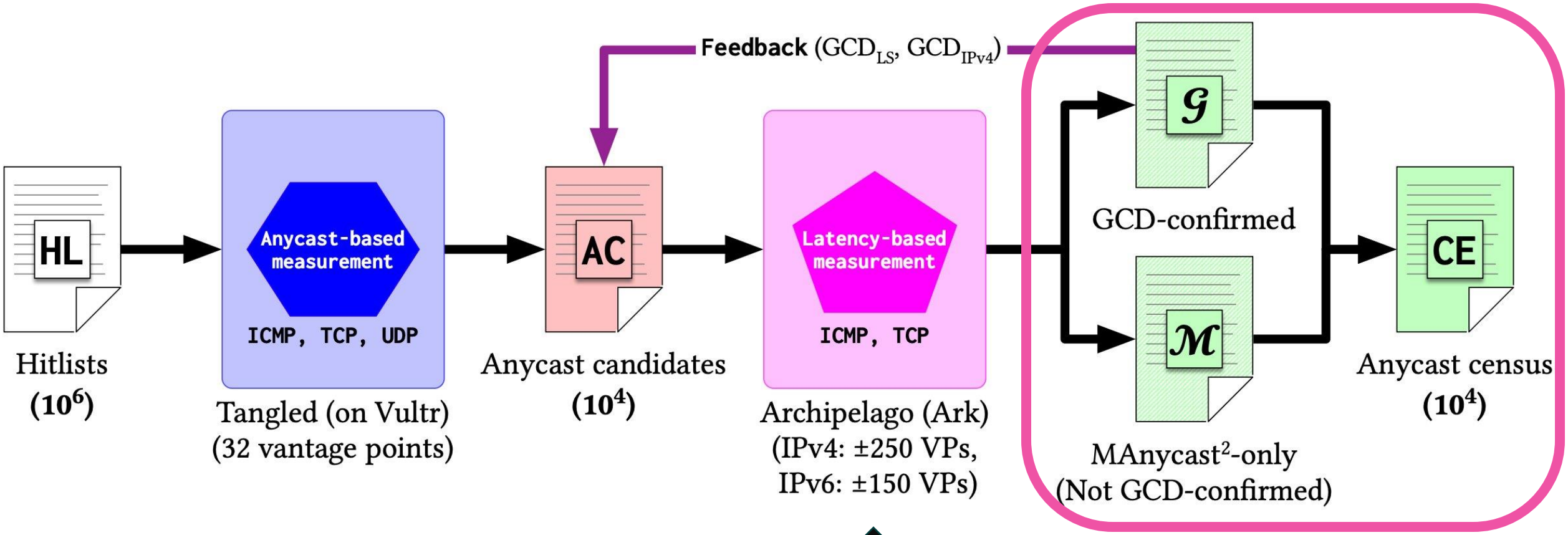
- Targets Anycast Candidates (10^4)
- Great-Circle-Distance latency measurements
- Performed using CAIDA's Ark (250+ VPs)
- **Output**: GCD-confirmed prefixes (10^4)
  - Including geolocation





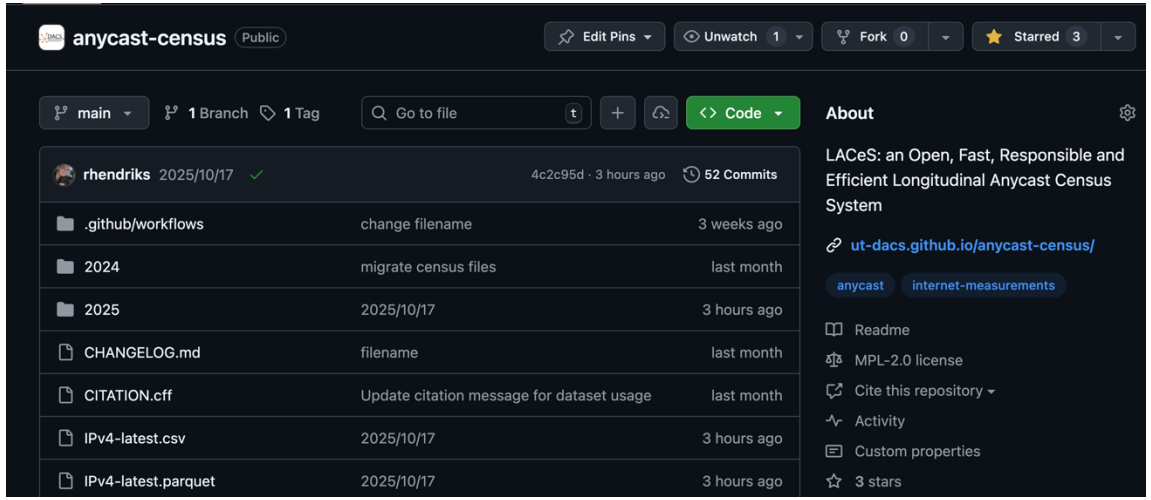
https://www.caida.org/projects/ark/locations/
Ark Vantage Point locations

Census Pipeline



Methodology
Output

- Results pushed daily to GitHub
  - github.com/ut-dacs/Anycast-Census/
- User-friendly dashboard
  - manycast.net



Census Date: Tue Mar 25 2025

IP: 1.1.1.1

Prefix: 1.1.1.0/24

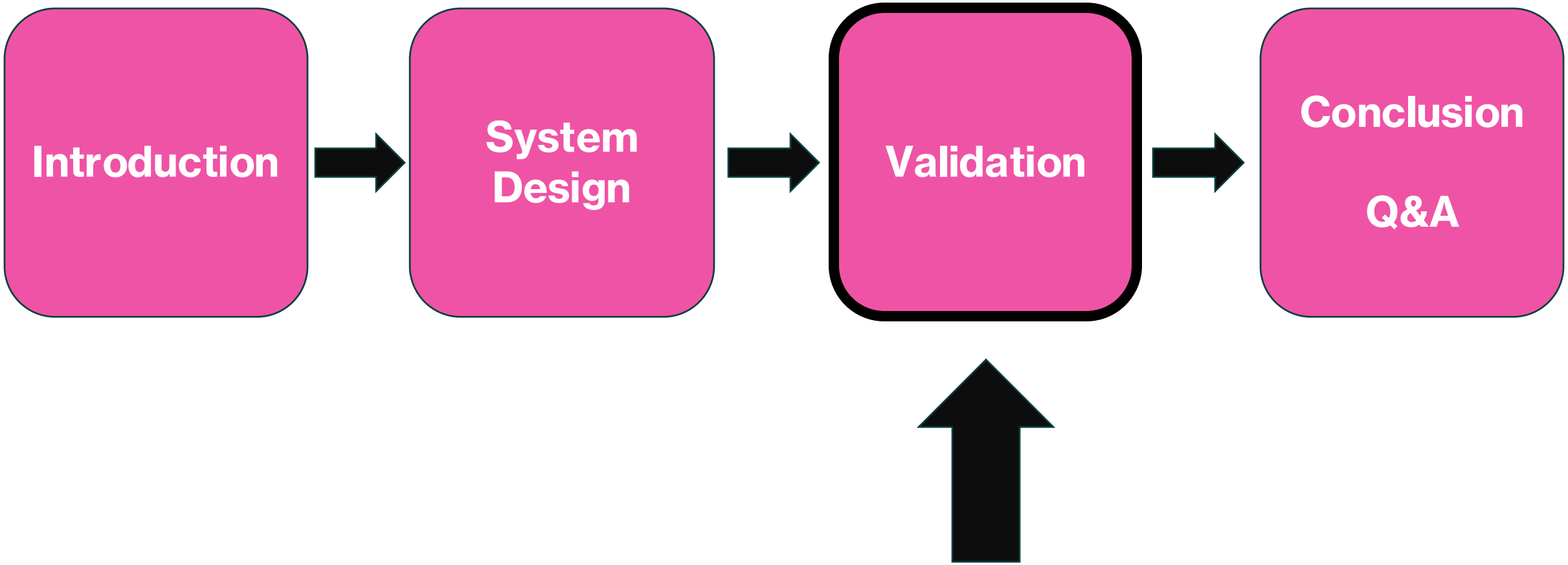
Type: Anycast

DETECTOR	ANYCAST	INSTANCES
Anycast-based ICMPv4	Yes	27
Anycast-based TCPv4	Yes	26
Anycast-based UDPv4	Yes	24
GCD-based ICMPv4	Yes	60
GCD-based TCPv4	Yes	19

Geolocation Data

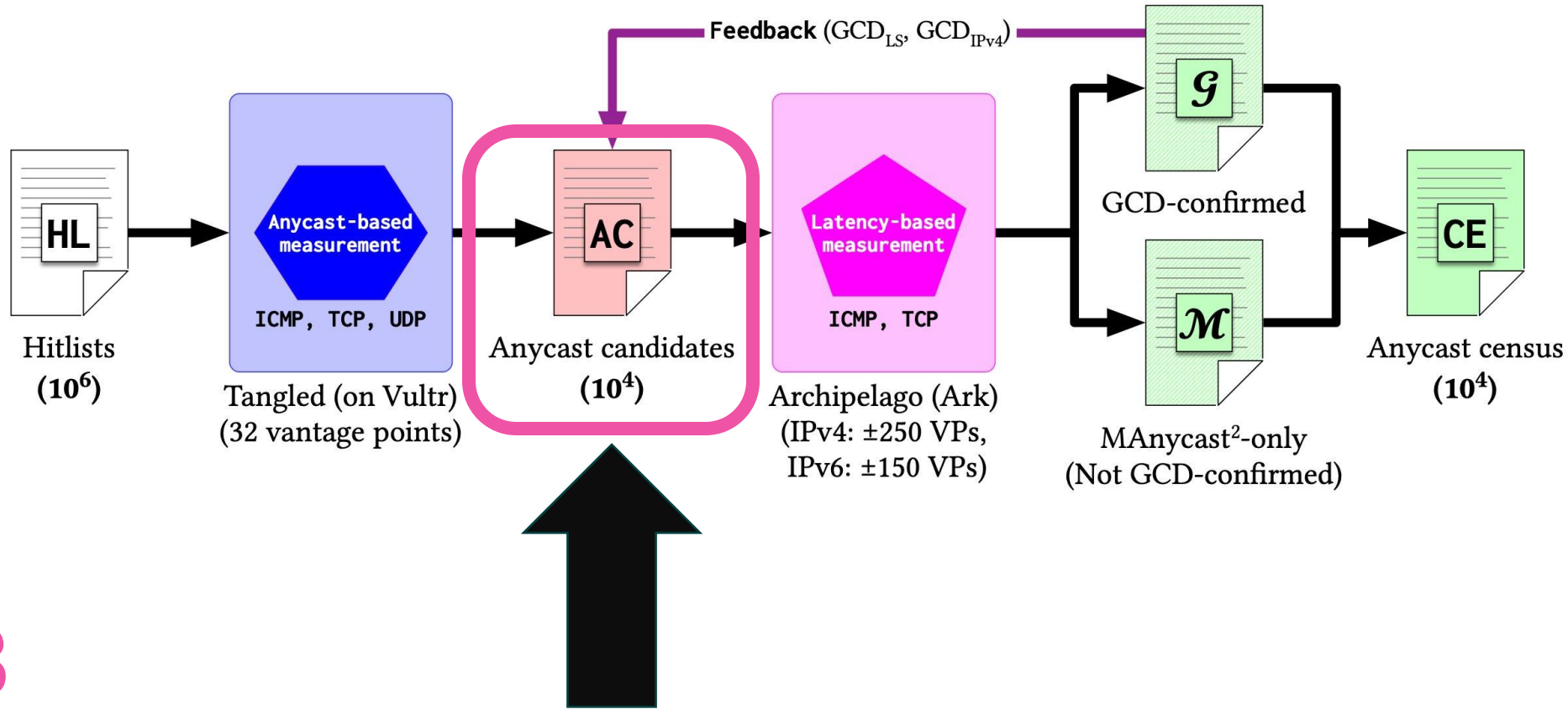


Structure



Validation
Anycast-based FNs

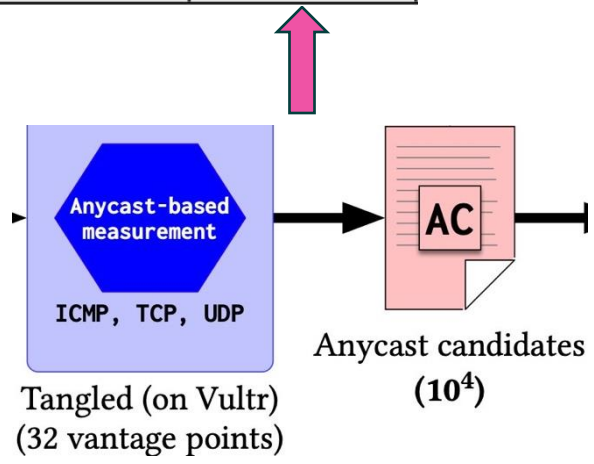
- How many prefixes are missed by the anycast-based method?



Validation
Anycast-based FNs

- How many prefixes are missed by the anycast-based method?
- Comparing
  - Anycast Candidates (anycast-based)

Protocol	AC
ICMPv4	30,609
ICMPv6	11,166

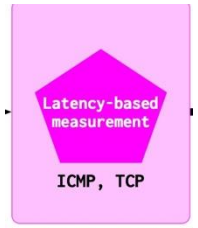


Validation
Anycast-based FNs

- How many prefixes are missed by the anycast-based method?
- Comparing
  - Anycast Candidates (anycast-based)
  - Large-scale GCD measurement (full hitlist)

Protocol	AC	GCD_{LS}
ICMPv4	30,609	13,514
ICMPv6	11,166	11,286





Archipelago (Ark)
(IPv4: ±250 VPs,
IPv6: ±150 VPs)

Validation
Anycast-based FNs

- How many prefixes are missed by the anycast-based method?
- Comparing
  - Anycast Candidates (anycast-based)
  - Large-scale GCD measurement (full hitlist)

Protocol	AC	GCD_{LS}	$AC \cap GCD_{LS}$
ICMPv4	30,609	13,514	12,698 (94.0%)
ICMPv6	11,166	11,286	10,617 (94.1%)



- Main takeaways
  - 94% coverage of prefixes detectable with GCD

Validation
Anycast-based FNs

- How many prefixes are missed by the anycast-based method?
- Comparing
  - Anycast Candidates (anycast-based)
  - Large-scale GCD measurement (full hitlist)

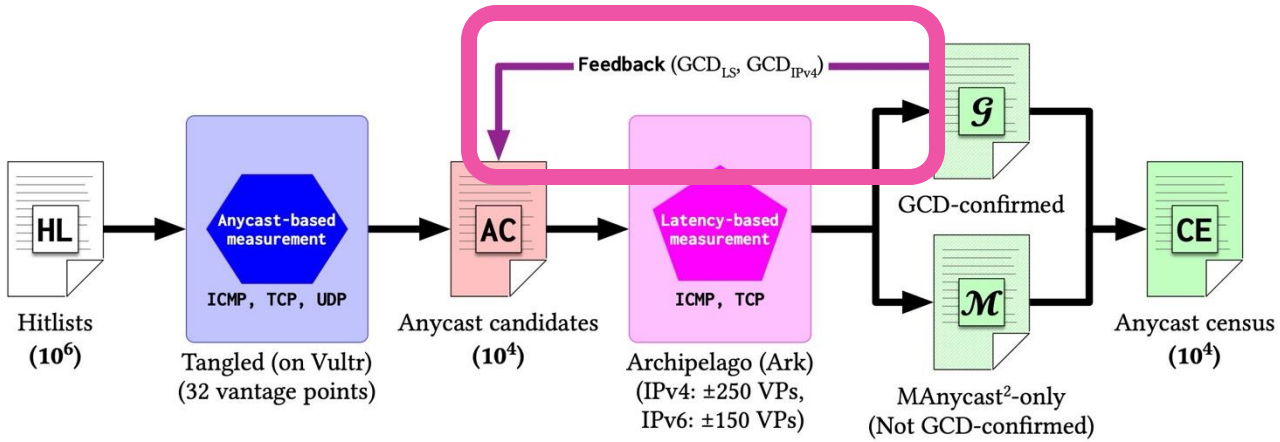
Protocol	AC	GCD_{LS}	$AC \cap GCD_{LS}$	FNs (FNR%)
ICMPv4	30,609	13,514	12,698 (94.0%)	816 (6.0%)
ICMPv6	11,166	11,286	10,617 (94.1%)	669 (5.9%)



- Main takeaways
  - 94% coverage of prefixes detectable with GCD
  - 6% FN rate

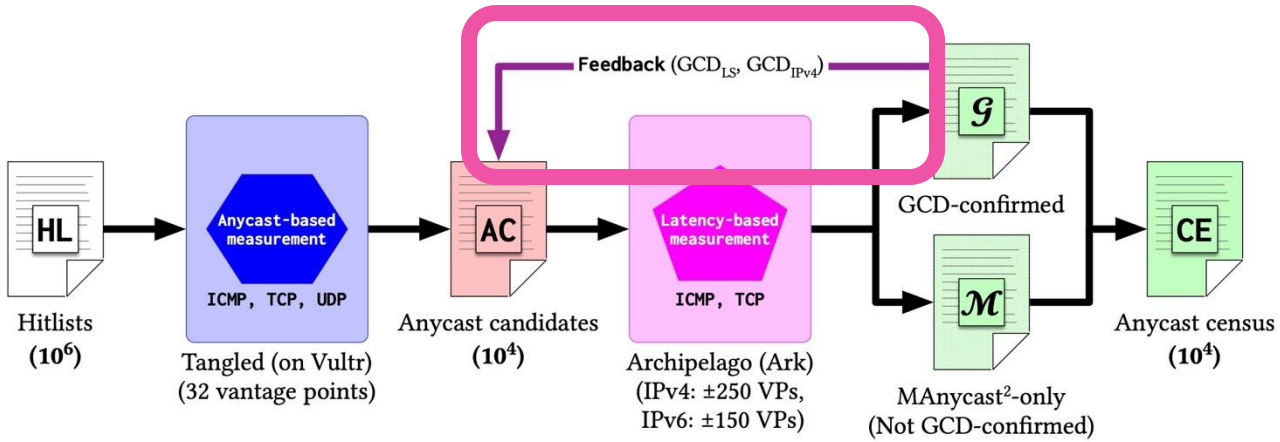
Validation
What about the FNs?

- FNs covered using feedback loop
  - Prefixes always fed to daily latency-based measurements



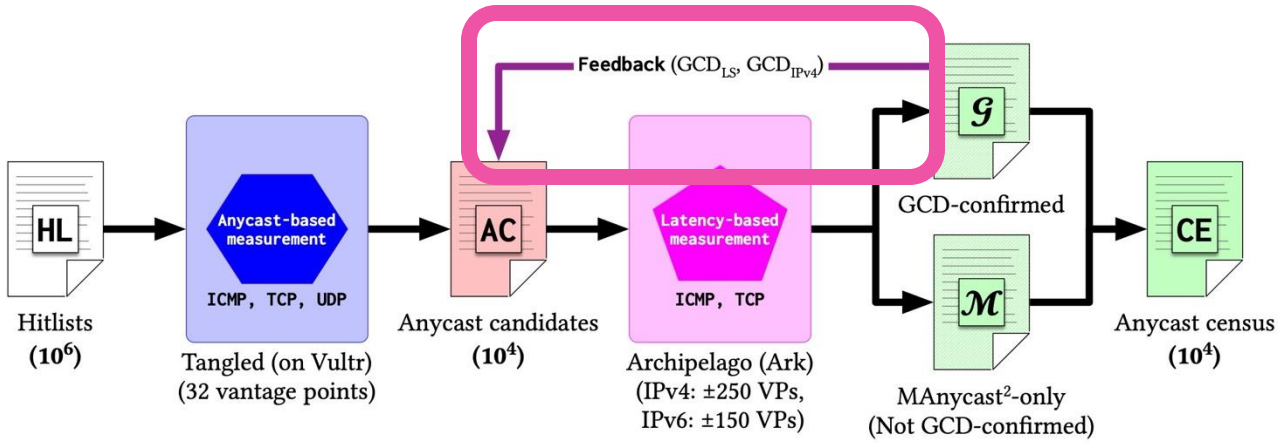
Validation
What about the FNs?

- FNs covered using feedback loop
  - Prefixes always fed to daily latency-based measurements
- GCD(LS)
  - Full hitlist latency-based measurements
  - Coverage of anycast-based FNs



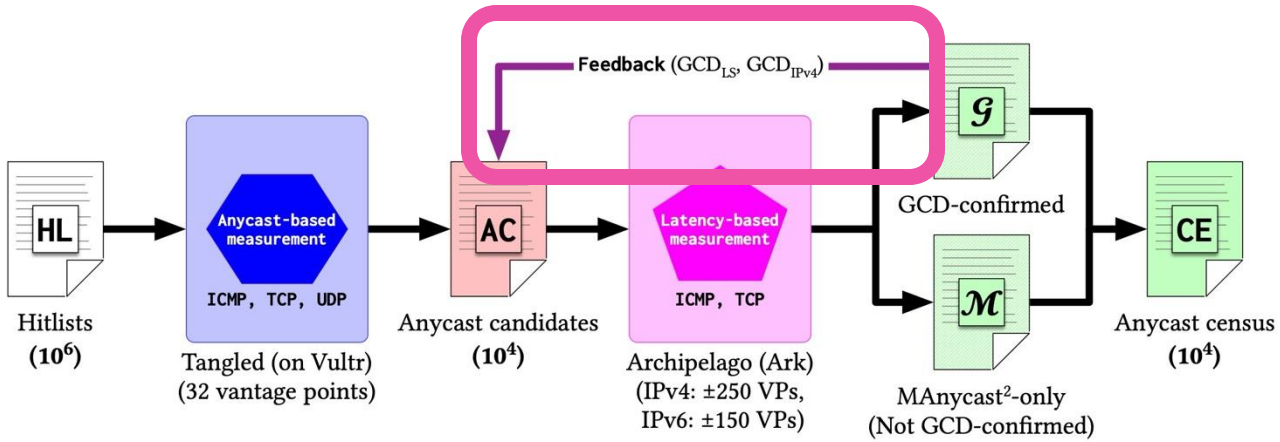
Validation
What about the FNs?

- FNs covered using feedback loop
  - Prefixes always fed to daily latency-based measurements
- GCD(LS)
  - Full hitlist latency-based measurements
  - Coverage of anycast-based FNs
- GCD(IPv4)
  - Full IPv4 measurement (with fewer VPs)
  - Coverage for prefixes missed by hitlists



Validation
What about the FNs?

- FNs covered using feedback loop
  - Prefixes always fed to daily latency-based measurements
- GCD(LS)
  - Full hitlist latency-based measurements
  - Coverage of anycast-based FNs
- GCD(IPv4)
  - Full IPv4 measurement (with fewer VPs)
  - Coverage for prefixes missed by hitlists
- Both repeated bi-annually
  - Too burdensome for daily measurements



Validation
Operator Ground Truth

- Most anycast prefixes originate from large CDNs
  - We contacted operators and used public operator datasets for validation
  - 100% IPv4 accuracy for dark green

AS	Organization	IPv4 (/24)		IPv6 (/48)	
396982	Google Cloud	3,627	(1st)	5	
13335	Cloudflare	3,133	(2nd)	284	(3rd)
16509	Amazon	1,286	(3rd)	120	
54113	Fastly	435	(4th)	65	
209242	Cloudflare Spectrum	289	(5th)	3,338	(1st)
19551	Incapsula (Imperva)	2		352	(2nd)
12041	Afilias	221		222	(4th)
44273	GoDaddy	32		122	(5th)

Table 5: Largest ASes by number of anycast /24-prefixes.

Validation
Operator Ground Truth

- Most anycast prefixes originate from large CDNs
  - We contacted operators and used public operator datasets for validation
  - 100% IPv4 accuracy for dark green
  - Near 100% (some FPs) for orange

AS	Organization	IPv4 (/24)		IPv6 (/48)	
396982	Google Cloud	3,627	(1st)	5	
13335	Cloudflare	3,133	(2nd)	284	(3rd)
16509	Amazon	1,286	(3rd)	120	
54113	Fastly	435	(4th)	65	
209242	Cloudflare Spectrum	289	(5th)	3,338	(1st)
19551	Incapsula (Imperva)	2		352	(2nd)
12041	Afilias	221		222	(4th)
44273	GoDaddy	32		122	(5th)

Table 5: Largest ASes by number of anycast /24-prefixes.

Validation
Operator Ground Truth

- Most anycast prefixes originate from large CDNs
  - We contacted operators and used public operator datasets for validation
  - 100% IPv4 accuracy for dark green
  - Near 100% (some FPs) for orange
  - IPv6 hitlist limitations light green

AS	Organization	IPv4 (/24)		IPv6 (/48)	
396982	Google Cloud	3,627	(1st)	5	
13335	Cloudflare	3,133	(2nd)	284	(3rd)
16509	Amazon	1,286	(3rd)	120	
54113	Fastly	435	(4th)	65	
209242	Cloudflare Spectrum	289	(5th)	3,338	(1st)
19551	Incapsula (Imperva)	2		352	(2nd)
12041	Afilias	221		222	(4th)
44273	GoDaddy	32		122	(5th)

Table 5: Largest ASes by number of anycast /24-prefixes.

Validation
Operator Ground Truth

- Most anycast prefixes originate from large CDNs
  - We contacted operators and used public operator datasets for validation
  - 100% IPv4 accuracy for dark green
  - Near 100% (some FPs) for orange
  - IPv6 hitlist limitations light green
- Reached out to ccTLDs
  - More interesting as some deploy regional anycast
  - Main takeaways:
    - Correct detection for vast majority
    - Geolocation closely match reality
    - ^ struggles with sites in near-proximity (*e.g.*, Prague & Vienna)
    - Longitudinal LACeS data capture confirmed infrastructure expansions

AS	Organization	IPv4 (/24)		IPv6 (/48)	
396982	Google Cloud	3,627	(1st)	5	
13335	Cloudflare	3,133	(2nd)	284	(3rd)
16509	Amazon	1,286	(3rd)	120	
54113	Fastly	435	(4th)	65	
209242	Cloudflare Spectrum	289	(5th)	3,338	(1st)
19551	Incapsula (Imperva)	2		352	(2nd)
12041	Afilias	221		222	(4th)
44273	GoDaddy	32		122	(5th)

Table 5: Largest ASes by number of anycast /24-prefixes.

Validation
Longitudinal Dynamics

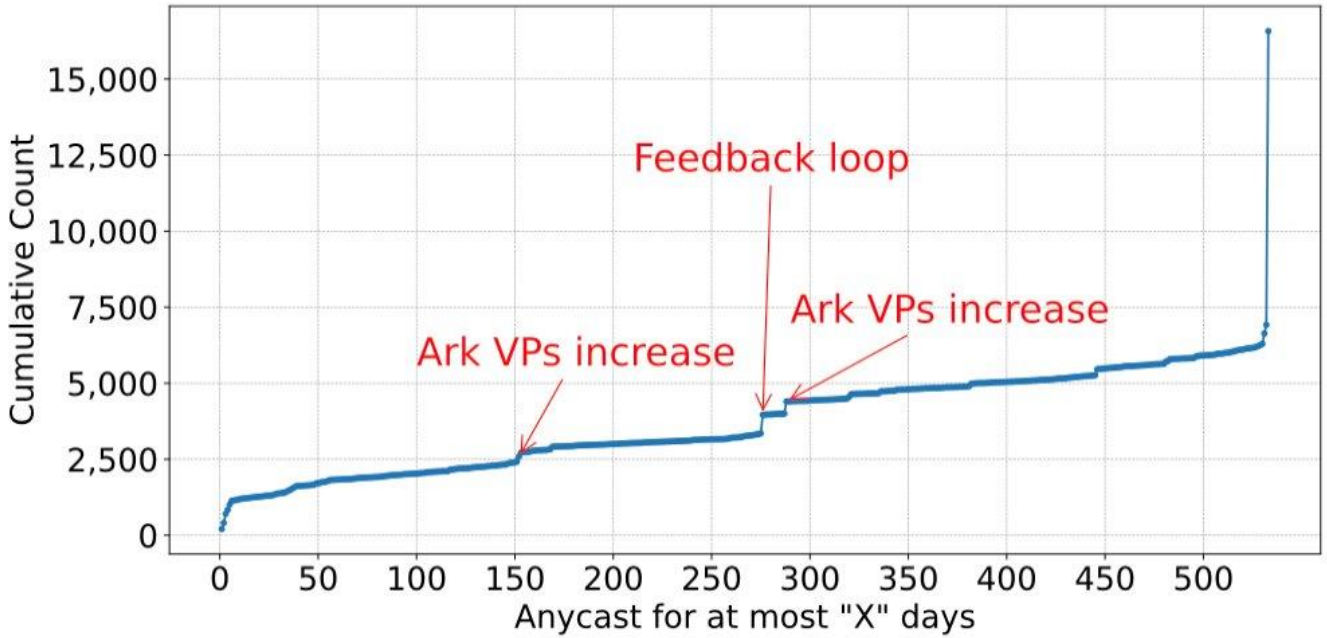


Figure 10: Cumulative count of anycast /24-prefixes detected to be anycast for at most "X" days.

Validation
Longitudinal Dynamics

- Main takeaways:
  - 9.7k /24s always visible
    - Majority from ground truth confirmed CDNs

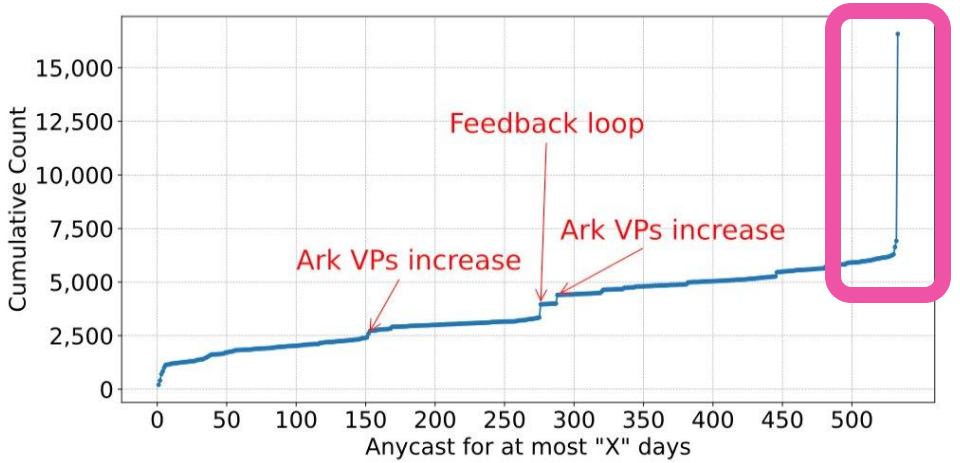


Figure 10: Cumulative count of anycast /24-prefixes detected to be anycast for at most "X" days.

Validation
Longitudinal Dynamics

- Main takeaways:
  - 9.7k /24s always visible
    - Majority from ground truth confirmed CDNs
  - 3.0k /24s visible for 20-80% of period
    - Changes in LACeS coverage
    - Prefixes that became anycast
    - Dynamic address utilization

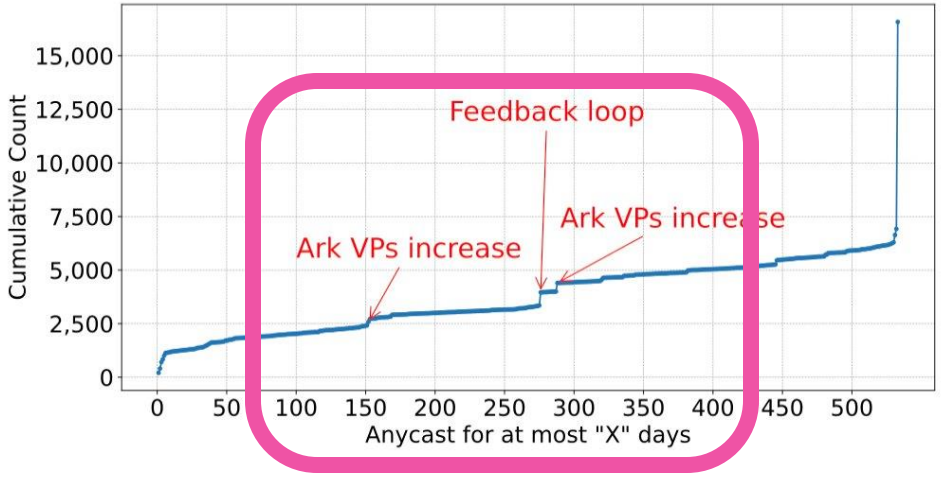


Figure 10: Cumulative count of anycast /24-prefixes detected to be anycast for at most "X" days.

Validation
Longitudinal Dynamics

- Main takeaways:
  - 9.7k /24s always visible
    - Majority from ground truth confirmed CDNs
  - 3.0k /24s visible for 20-80% of period
    - Changes in LACeS coverage
    - Prefixes that became anycast
    - Dynamic address utilization
  - 1.1k /24s visible up to 10 days
    - 572 belonging to Imperva (DDoS protection service?)

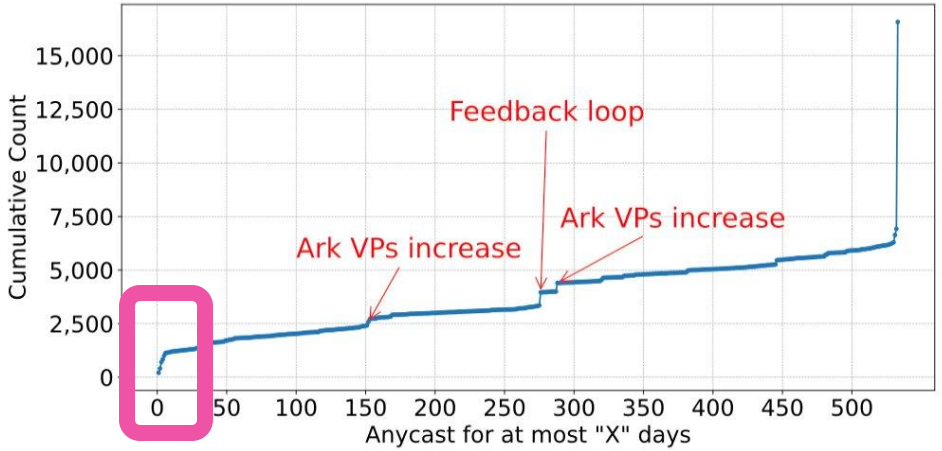


Figure 10: Cumulative count of anycast /24-prefixes detected to be anycast for at most "X" days.

Validation
Longitudinal Dynamics

- Main takeaways:
  - 9.7k /24s always visible
    - Majority from ground truth confirmed CDNs
  - 3.0k /24s visible for 20-80% of period
    - Changes in LACeS coverage
    - Prefixes that became anycast
    - Dynamic address utilization
  - 1.1k /24s visible up to 10 days
    - 572 belonging to Imperva (DDoS protection service?)
  - 191 anycast for 1 day
    - Suspected data-plane/BGP misconfigurations/hijacking
    - Seen at two location (actual and suspected hijacking location)

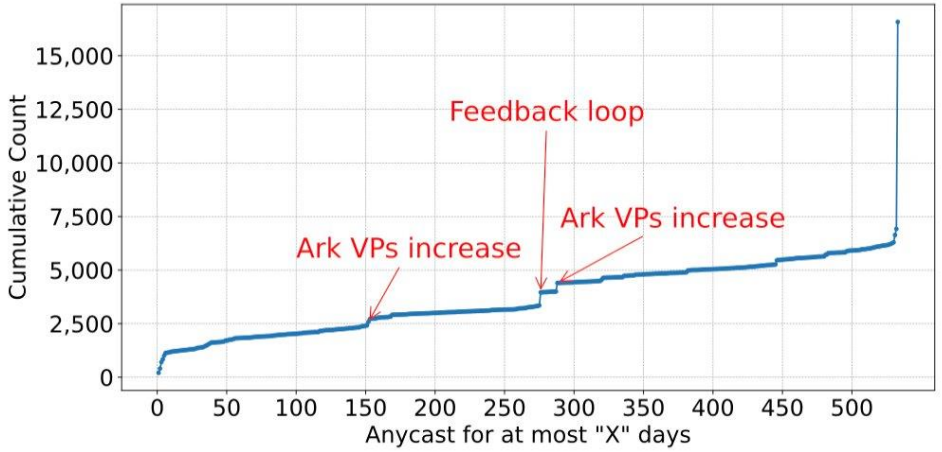
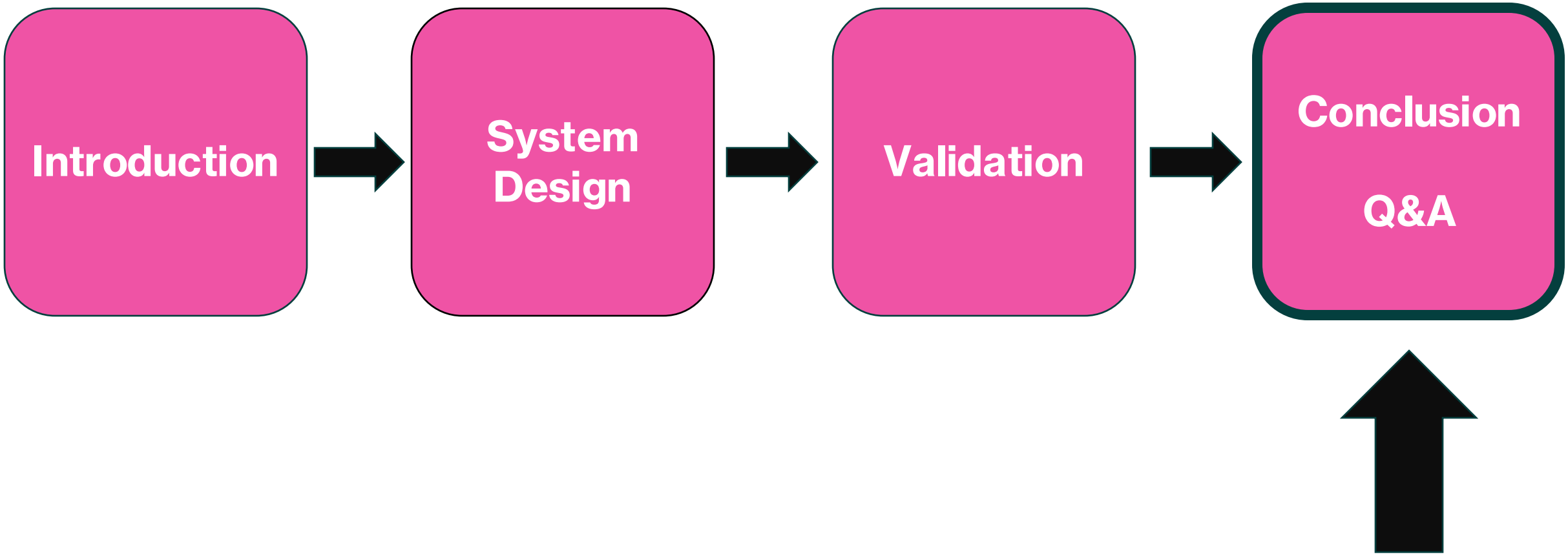


Figure 10: Cumulative count of anycast /24-prefixes detected to be anycast for at most "X" days.

More details in the paper…

- Comparison with external datasets
- More in-depth longitudinal analysis
- Many more validating experiments
- /24-prefix scanning motivation (+ partial anycast)
- Additional ground-truth
- And much more…

Structure



Conclusion

- We built LACeS, a daily anycast census system
  - Extensive validation shows high accuracy, coverage, and precision
  - Confirmed over 60% of detected anycast using ground-truth
- We release daily LACeS censuses (17+ months to date)
  - Available at github.com/ut-dacs/anycast-census
  - We intend on providing the census for the foreseeable future
- Full source code for LACeS publicly available (see paper)

Q&A

Appendix slides

Comparison with external datasets

- BGPTools
  - Classifies entire RIB prefixes as anycast
  - Uses methodology similar to anycast-based
  - Our results show they suffer from MAnycast2 FPs
  - And from FNs by classifying entire RIB prefixes
- IPInfo
  - Has anycast flag in their dataset
  - Uses methodology similar to latency-based
  - We find large agreement between results
- Both only provide detection (no geolocation/enumeration)

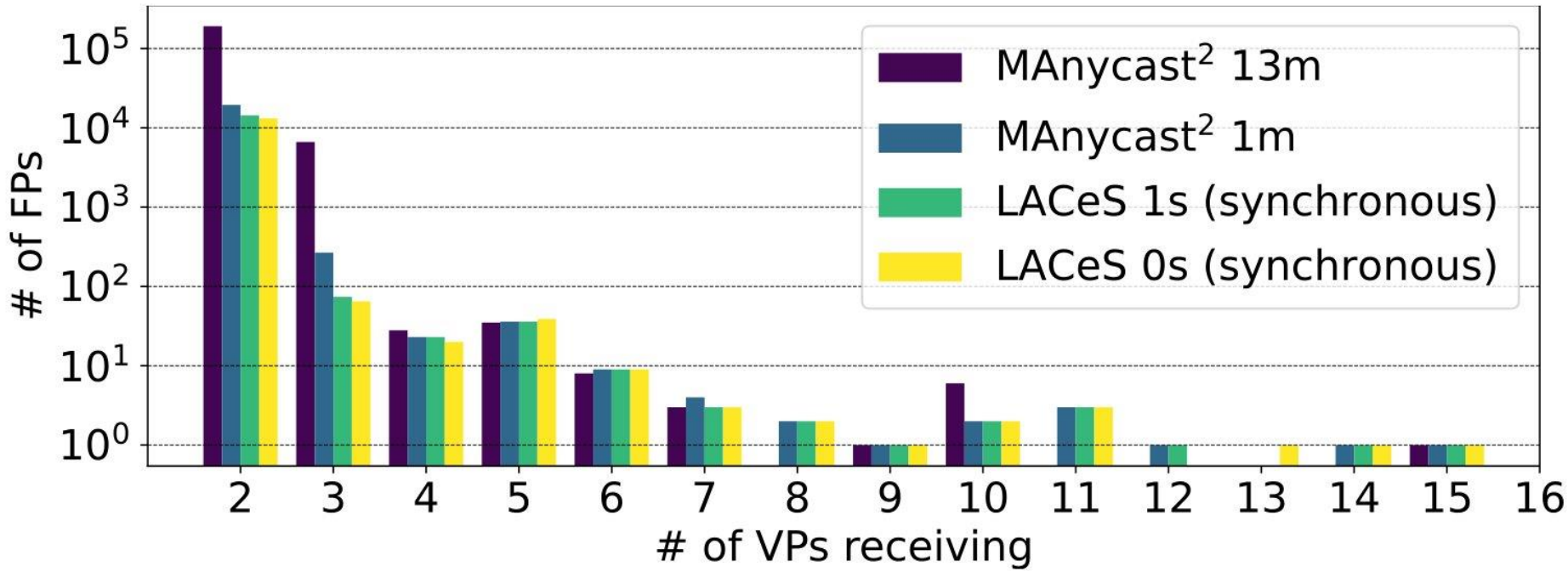




Not GCD-confirmed

# of sites receiving	Candidate anycast	$\mathscr{G}$ (GCD)	$\mathscr{M}$ (¬GCD)	Overlap (in %)
2	17,996	726	17,270	4.03%
3	597	366	231	61.31%
4	363	309	54	85.12%
5	253	215	38	84.98%
6-10	1,103	954	149	86.49%
11-15	854	713	141	83.49%
16-20	4,608	4,608	0	100.0%
21-25	1,415	1,404	11	99.22%
26-32	3,420	3,420	0	100.0%
Total	30,609	12,715	17,894	41.54%

Synchronous probing



Longitudinal Statistics

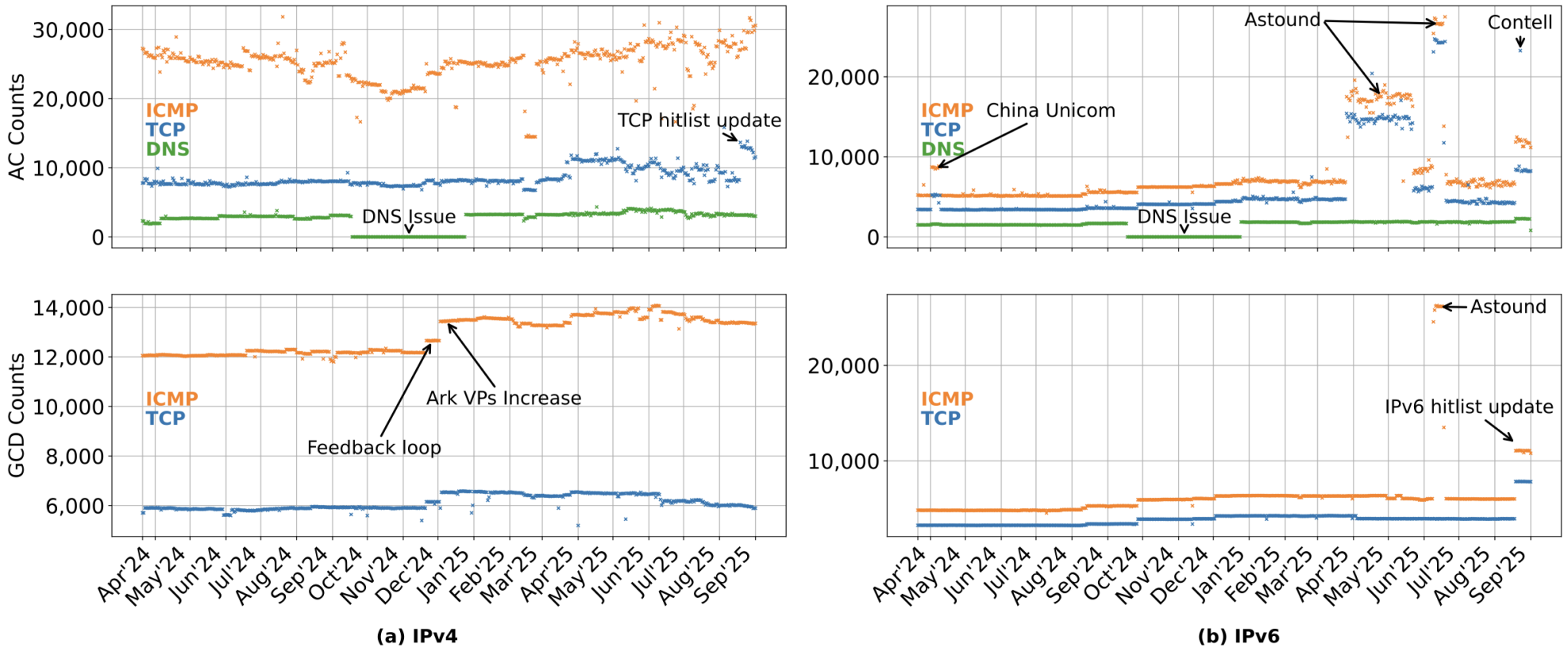


Figure 9: LACeS detection statistics by method and protocol over time for both IPv4 and IPv6.

Protocol coverage

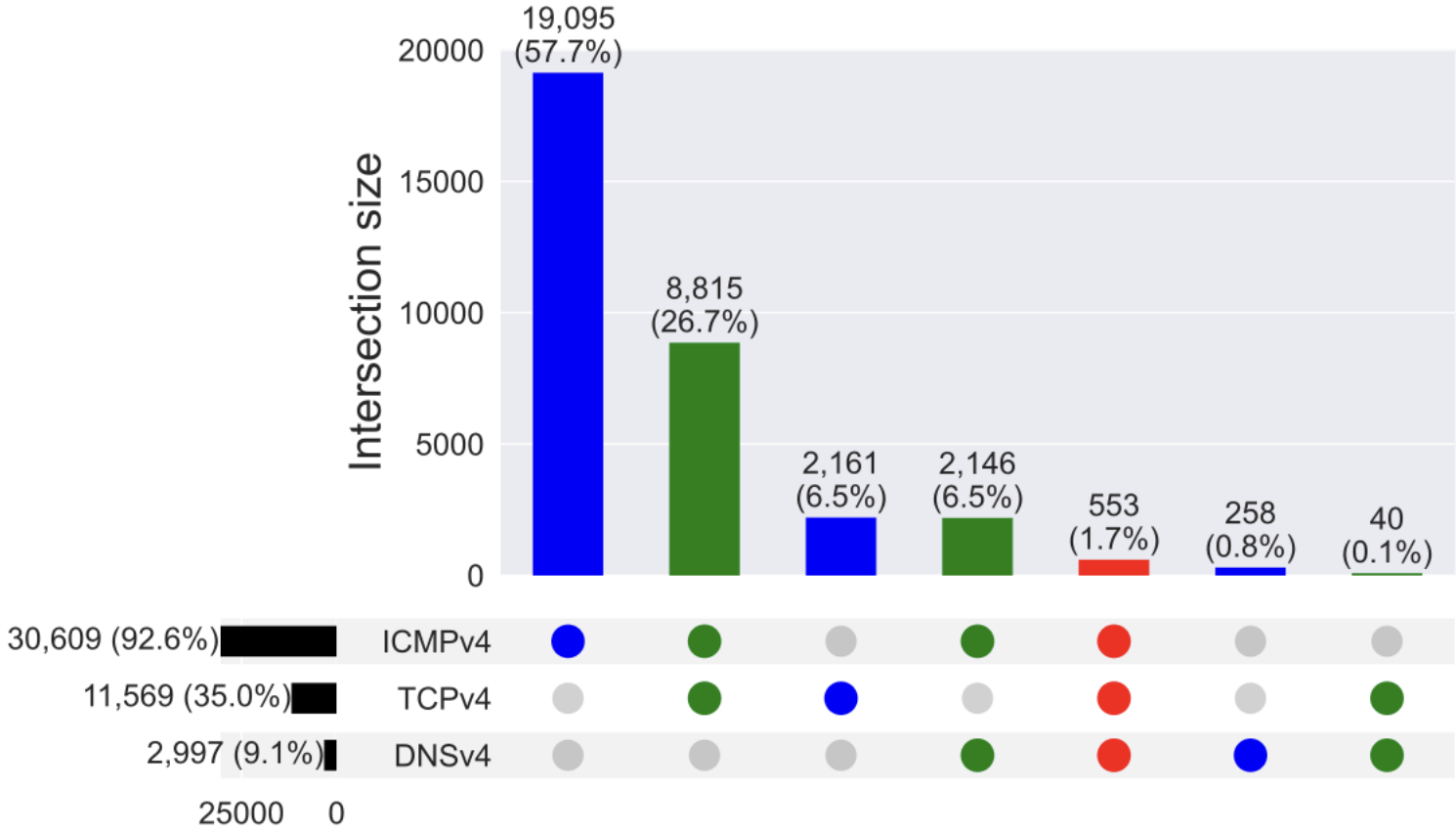


Figure 13: LACeS detection of anycast candidates for ICMPv4, TCPv4, and DNSv4.

Enumeration counts

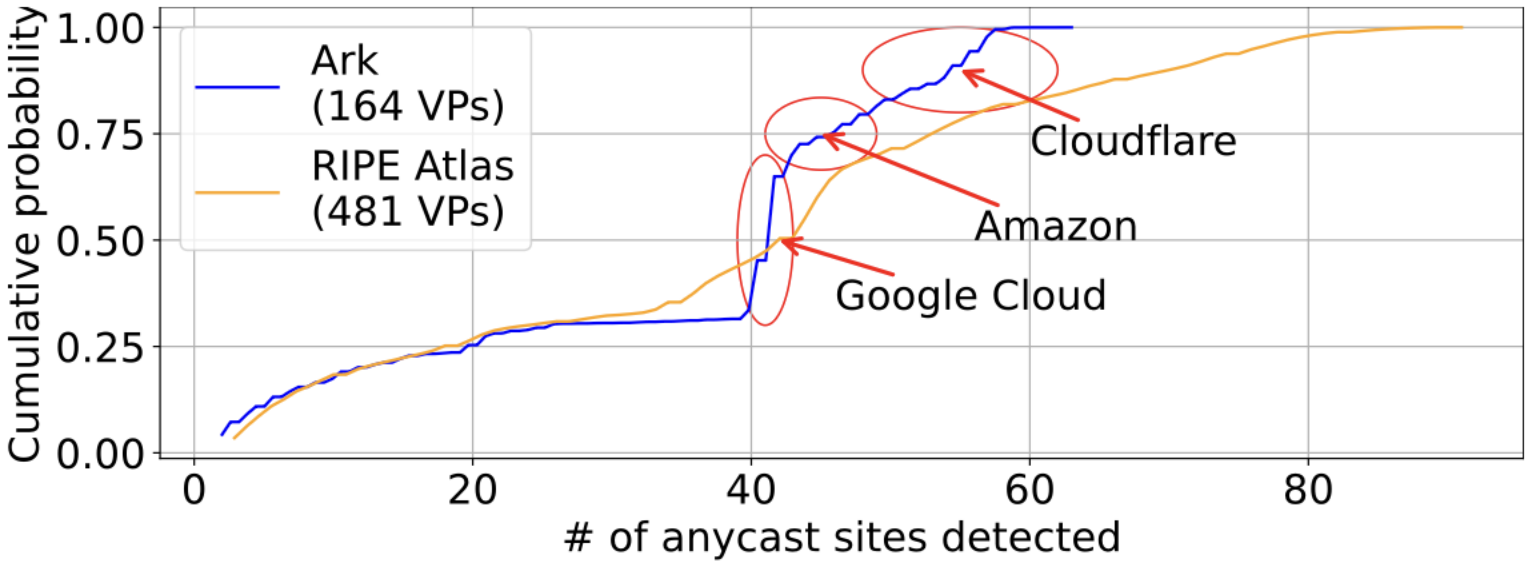


Figure 6: Number of sites detected per /24-prefix.