My 1st RIPE Meeting

RACI thanks!

Grant Agreement n. 318627

Where are the anycasters?

Dario Rossi
Professor
dario.rossi@telecom-paristech.fr
http://www. telecom-paristech.fr/~drossi/anycast

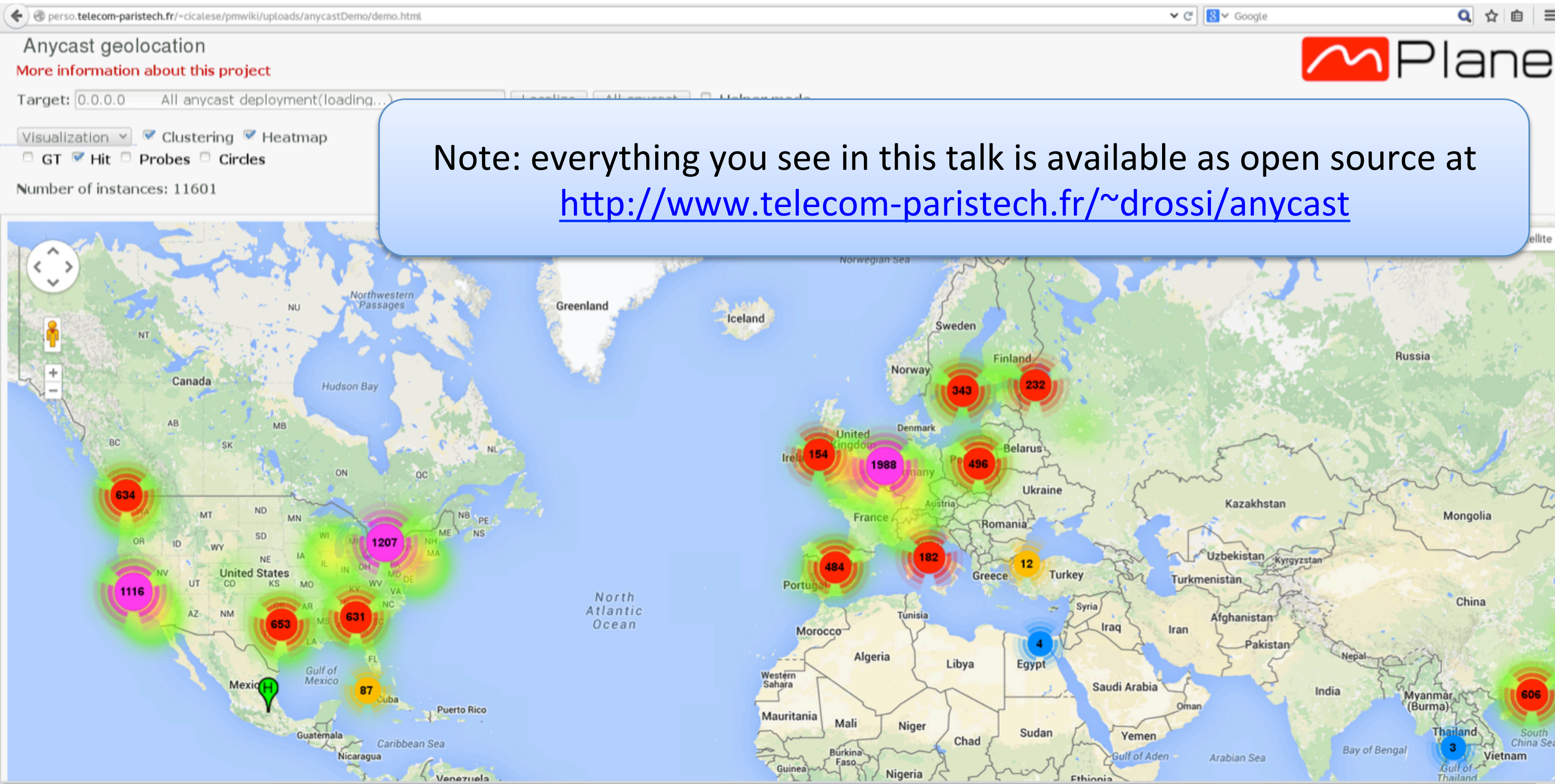
TELECOM ParisTech

Joint work with Danilo Cicalese, Diana Joumblatt, Jordan Auge and Timur Friedman

RIPE71 Plenary

Talk Teaser

A seminal work [4] at NANOG'59 investigates *who* are the IP anycasters. Our focus is instead on *where* they are



Note: everything you see in this talk is available as open source at
http://www.telecom-paristech.fr/~drossi/anycast

Demo shortlink: goo.gl/Ff8gdQ

Talk outline

Foreword: Quick overview, complementary technical talk at RIPE MAT WG if interested!

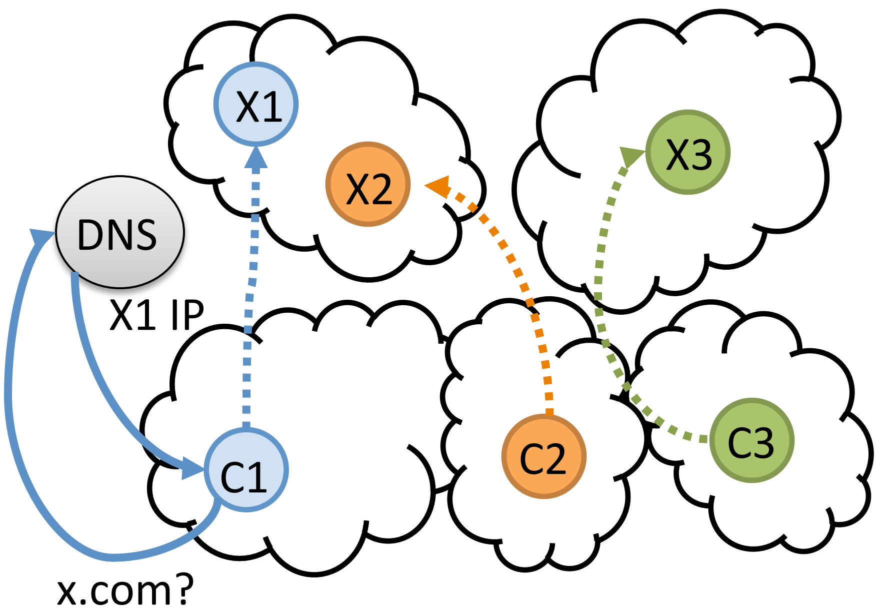
- Background & related work
- What we have (just finished!)
  - ✓ Latency-based anycast geolocation technique [1]
  - ✓ IPv4 anycast censuses [2]
  - ✓ Demo, source code, ground truth and more [3]
- What we miss (just started!)
  - Study infrastructure evolution & usage
  - Application to BGP hijack detection
- At last, New Yarely Conclusions And Summary of Talk

Notation:
RIPE Atlas credits
cost per action

[1] D. Cicalese et al. A Fistful of Pings: Accurate and Lightweight Anycast Enumeration and Geolocation, IEEE INFOCOM, Apr 2015.
[2] D. Cicalese et al. Characterizing IPv4 Anycast Adoption and Deployment, ACM CoNEXT, Dec 2015.
[3] http://www.telecom-paristech.fr/~drossi/anycast.

Application-layer anycast

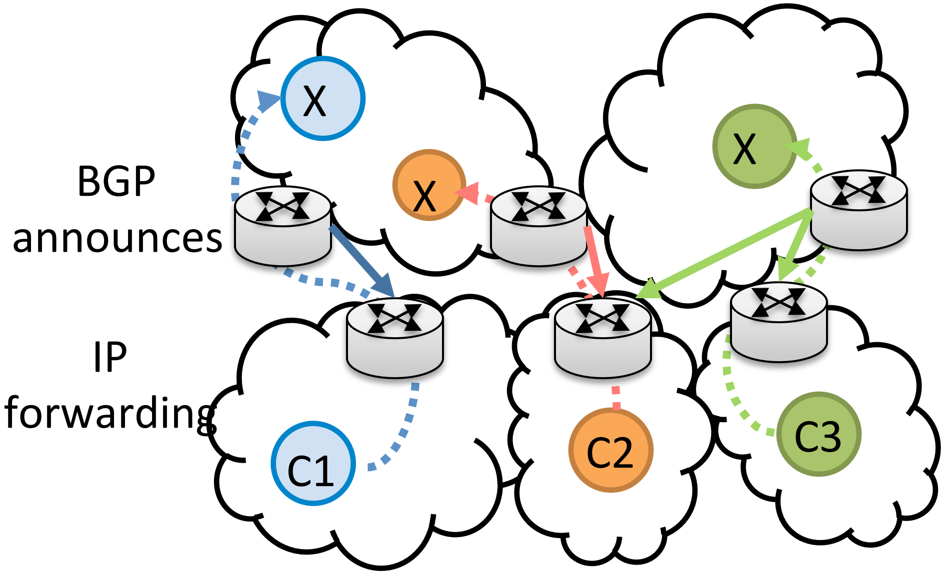
- How it works
  - Relies on IP unicast
  - Server selection via DNS redirection, URL rewriting



- Pros
  - Ability to specify selection criteria
  - Fine-grained control over server load
  - Maintains connection affinity
  - Fast failover
- Cons
  - Ad hoc, per service configuration
  - Overhead to collect selection metrics
  - E.g. List of servers up, latency, load

IP-layer anycast

- How it works
  - Shared IP address for replicate servers
  - BGP Prefix advertised from multiple points of origin



- Pros
  - Natively supported by IP
  - Transparent to upper layers
  - Visibility of servers (Global vs Local)
- Cons
  - Lack of fine-grained control
  - Destination determined by IP routing metrics
  - Prone to prefix hijacking
  - No guarantees of server affinity (e.g., connection-oriented services)

Our focus

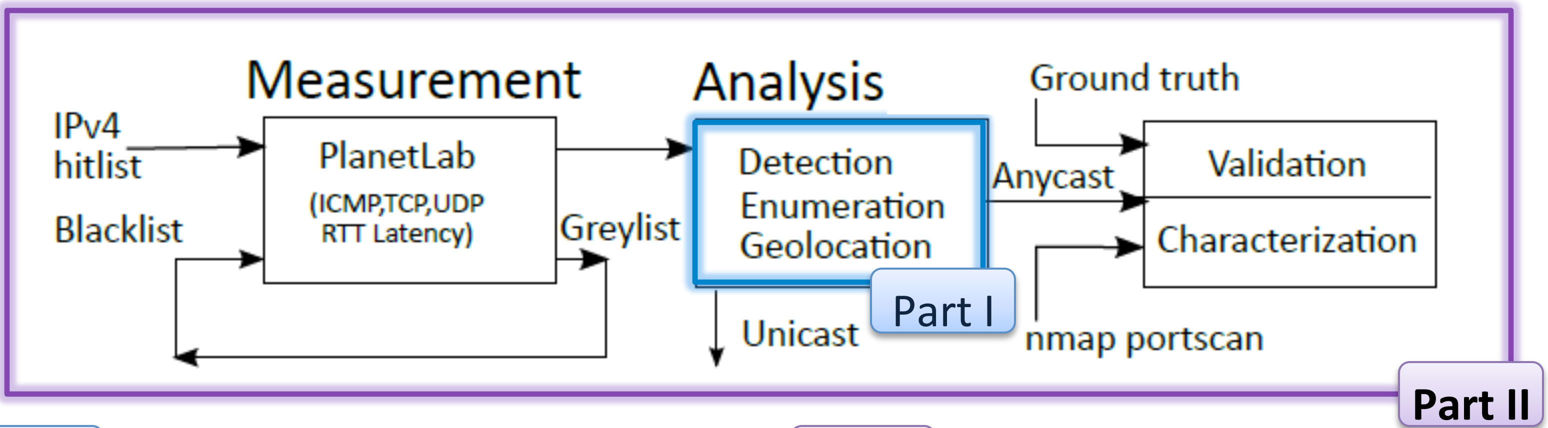
Related work

	[4]	[5]	Part I	[6]	[7]	[8]	[9]	[10]	[11]	[12]	Part II
Platform(#VPs)	Renesys monitors	PL (238), Netalyzr (62k), rDNS (300k)	*PL (300), RIPE (6k)*	End-hosts $O(100)$	PL (300)	DNSMC (77)	PL (129)	rDNS (20K)	C,F,K root	Renesys monitors	*PL (300)*
Technique	BGP vs. traceroute	DNS CHAOS +traceroute	*Latency probes*	DNS CHAOS	DNS CHAOS	DNS CHAOS +BGP	DNS CHAOS	DNS queries	pcap	BGP	*Latency probes*
Targets	IPv4 prefixes	F root, TLDs, AS112	*F,I,K,L root EdgeCast CloufFlare CDNs*	C,F-K,M root	B,F,K root, TLDs	C,F-K,M root	C,F-K,M root, AS112	F,J root, AS11		1 CacheFly prefix	*all IPv4*
Detect	✓	✓	✓								✓
Enumerate		✓	✓								✓
Geolocalize			✓								✓
Proximity					✓		✓	✓	✓		
Affinity				✓	✓		✓	✓	✓	✓	
Availability					✓	✓		✓		✓	
Loadbalance								✓			

Main differentiators

- **Part I**: first lightweight and protocol-agnostic technique able to detect, enumerate and geolocate anycast instances
- **Part II**: first large-scale (several IPv4 censuses) geolocation of anycast instances

Workflow



Part I PlanetLab and RIPE Atlas

- **Lightweight**: O(1) pings per target per vantage point
- **DNS**: use past RIPE Atlas measurements (0 credits!)
- **CDN**: issue new ICMP measurements

  500 credits per target (good enough coverage)
  180K credits (all nodes)
- **PlanetLab**: use different protocols (ICMP, DNS/UDP, DNS/TCP, HTTP/TCP)

Part II PlanetLab only

- **Heavyweight**: apply technique of Part-I to over 5,000,000 targets, per census
- Estimate of RIPE Atlas credits, 3 Billions per census (same footprint as PlanetLab ~300)
- Solution: combine PlanetLab and RIPE Atlas (more next!)

Anycast geolocation

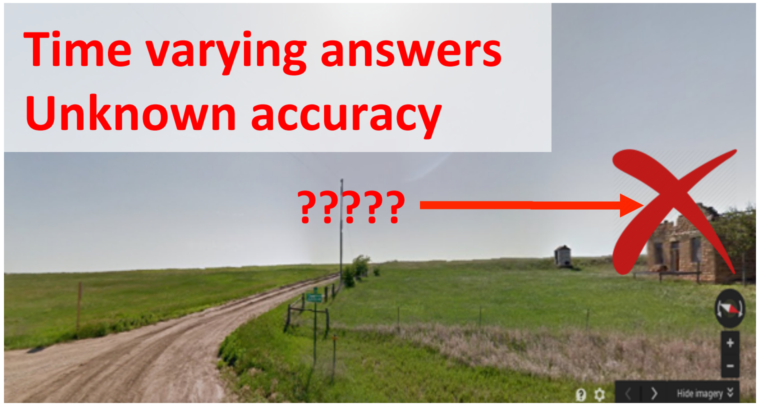
- Problem statement: where are the anycast instances?
  - E.g., Google 8.8.8.8 or CloudFlare, or EdgeCast or root servers, etc.

Commercial databases

Mountain View, CA (IP2Location)
New York, NY (Geobytes)
United States (Maxmind)

Distributed measurement

Tools using distributed measurement aren't better !



Anycast geolocation

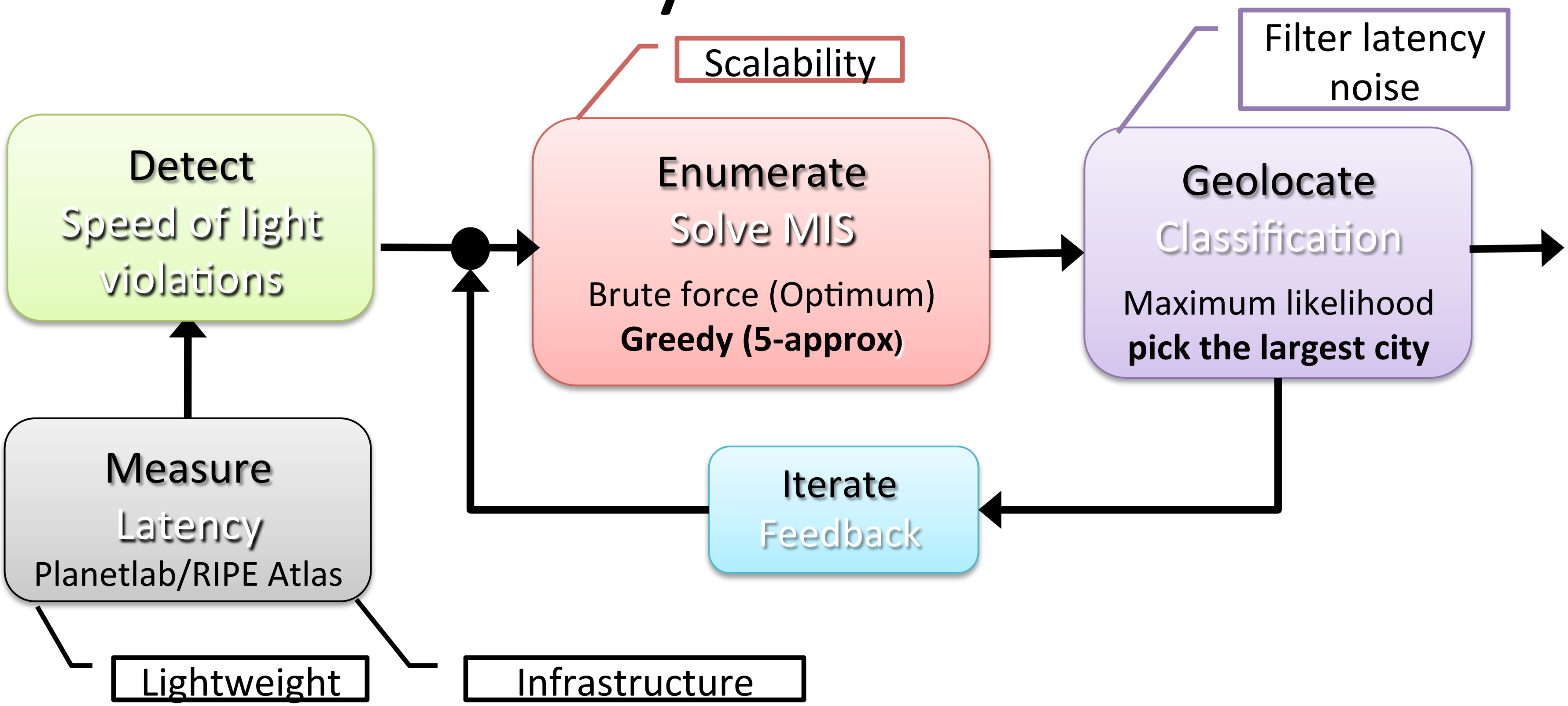
- Problem statement: where are the anycast instances?
  - E.g., Google 8.8.8.8 or CloudFlare, or EdgeCast or root servers, etc.
- Solution
  - Leverage inconsistency of latency measurement
  - This was used in NANOG'59 to detect **who** are the anycasters
  - We raise this to the next level and geolocate **where** they are
- Properties
  - **Lightweight**: few pings
  - **Protocol agnostic**: ICMP probes
  - **Accurate** against ground truth
  - **Fast**: greedy, but as good as costly optimum solution

Distributed measurement

Tools using distributed measurement aren't better !

But they could!

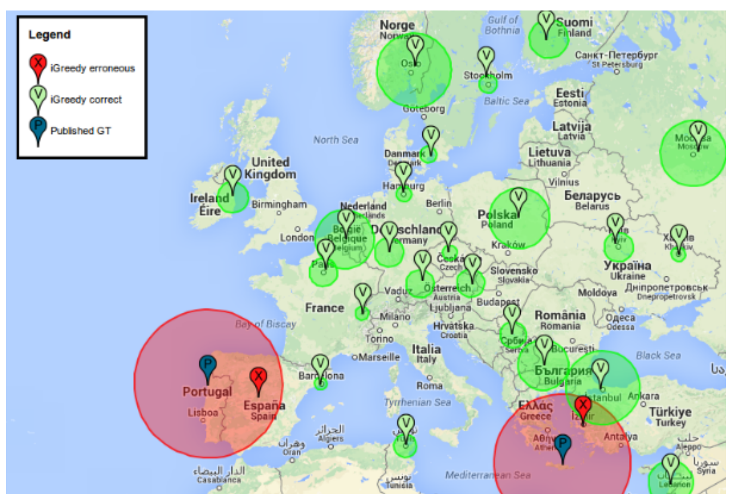
iGreedy overview



Measure | Detect | Enumerate | Geolocate | Iterate

iGreedy performance

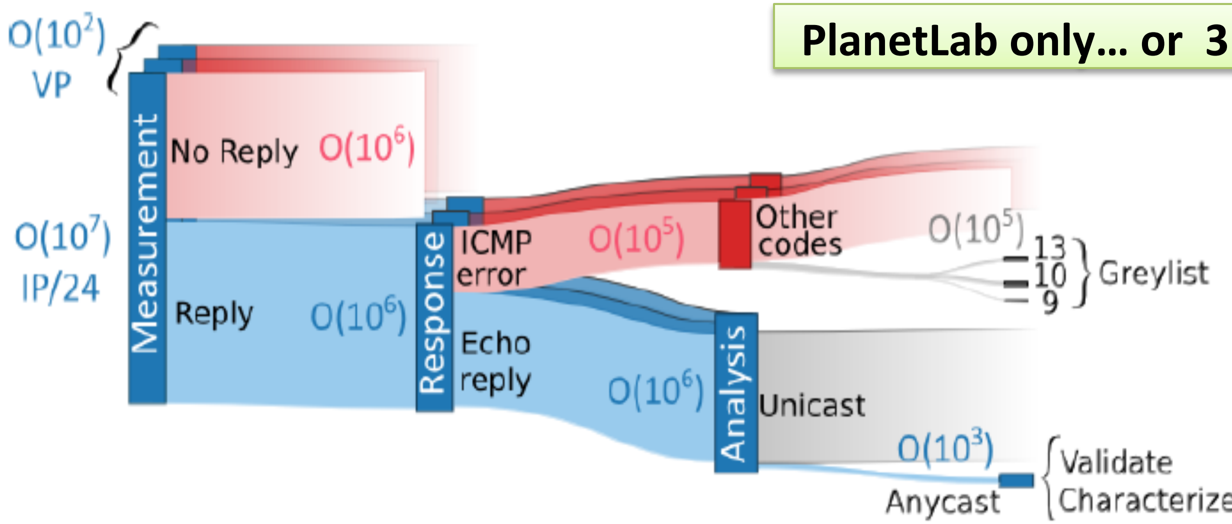
- At a glance
  - **Accurate enumeration** over 75% recall
  - **Precise geolocation** over 75% true positives
  - **Protocol agnostic** DSN and CDN, etc.
  - **Lightweight** 100x less probes than previous work
  - **Fast** O(100ms) greedy vs O(1000sec) for brute force
  - **Ready** Open source code [3] to measure, analyze and map!



IPv4 anycast censuses

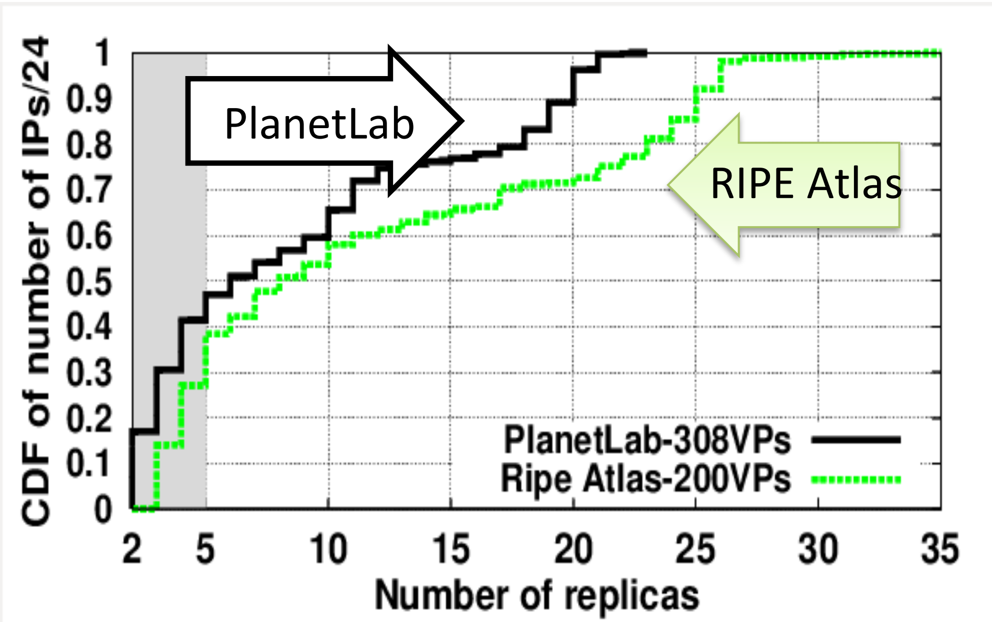
- Significant re-engineering, typical workflow:

PlanetLab only… or 3Billion credits!

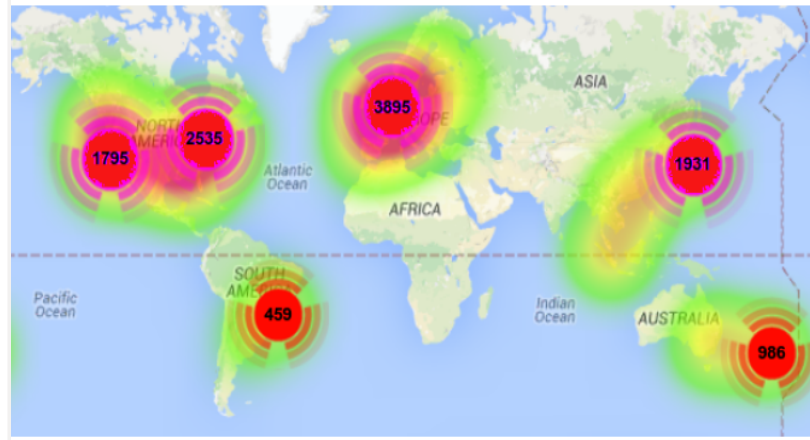


- $O(10^7)$ targets x $O(10^2)$ active probes
- $O(10^3)$ targets/sensor/second , 1 census = few hours
- $O(10^7)$ runs of iGreedy later….
- $O(10^3)$ targets are anycast – *proverbial needle in the IPv4 haystack*

Censuses at a glance



	IP/24	ASes	Cities	Countries	Replicas
All	1,696	346	77	38	13,802
$\geq$ 5 Replicas	897	100	71	36	11,598
∩ CAIDA-100	19	8	30	18	138
∩ Alexa-100k	242	15	45	29	4,038



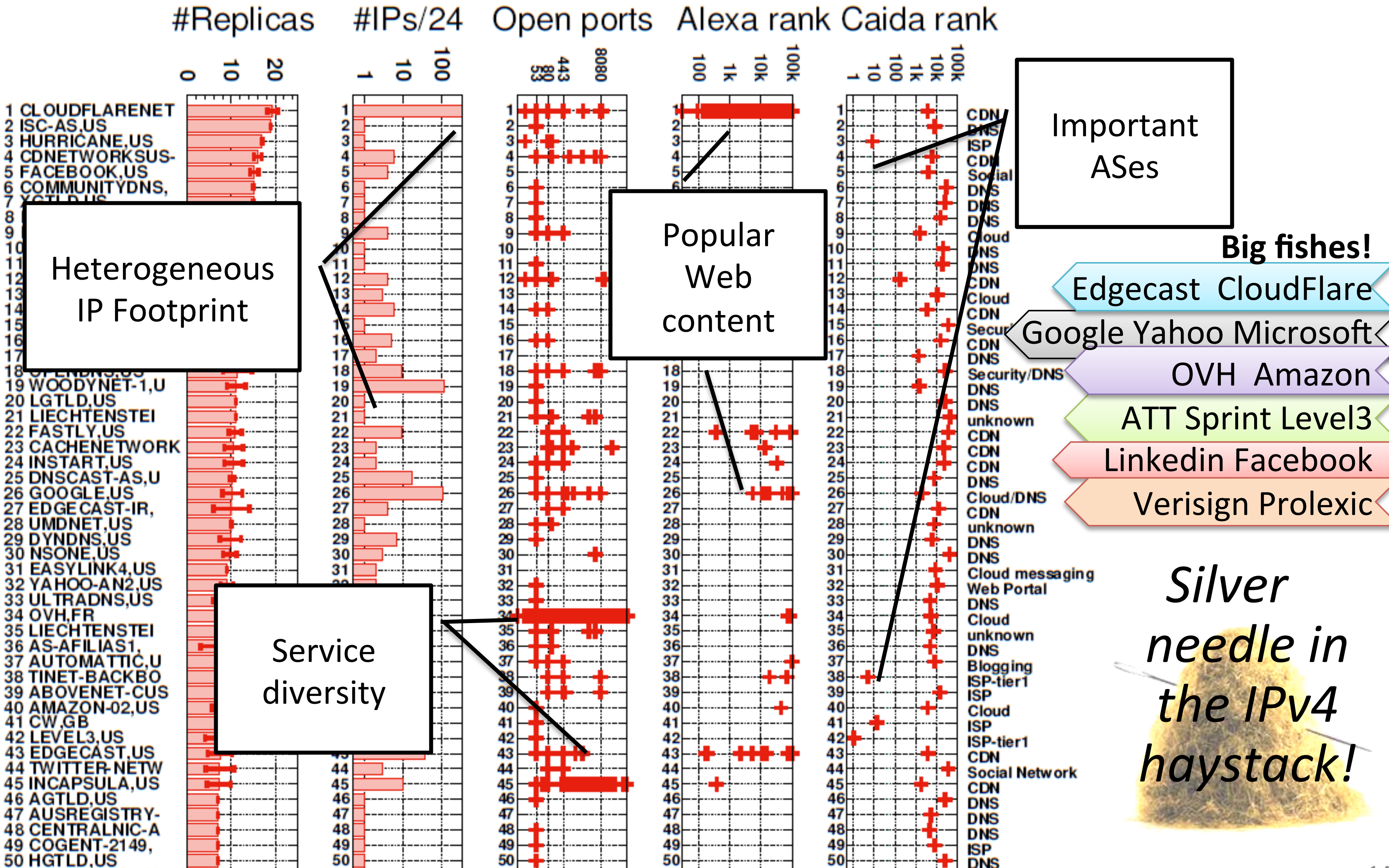
- Combine PlanetLab and RIPE Atlas
  - Ameliorate coverage

    1.8M credits for the top-100
  - Validate small deployments

    150K credits for 2-instances deployments

More on this at MAT WG

Top-50 IPv4 anycast deployments



IPv4 anycast software census

- Nmap census

nmap portscan statistics

IPs/32	ASes	Ports (with SSL)	Well-known	Software
812	81	10,499 (185)	457	30

  - Stealthy scan, all ports, 1 IP per each anycast /24
  - Not only DNS! Lots of CDN/Web (++), but also e.g., Gmail(??) or >10,000 open ports on OVH (!!)

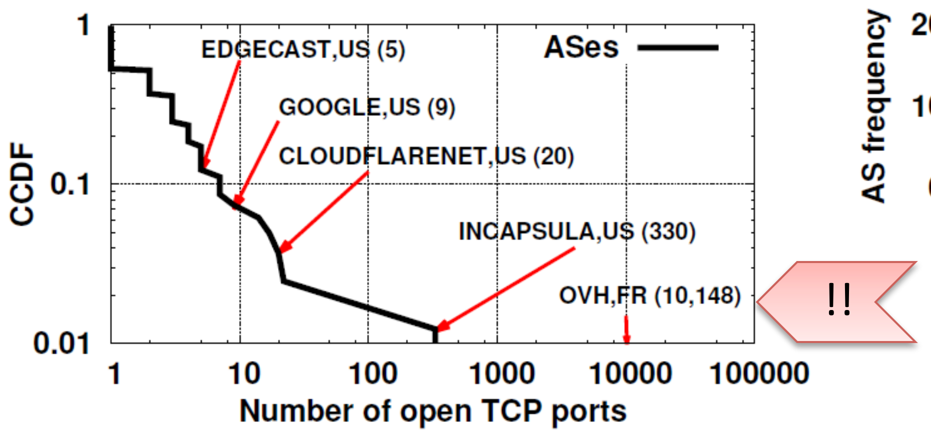


Figure 9: Complementary CDF of the number of open TCP ports per AS.

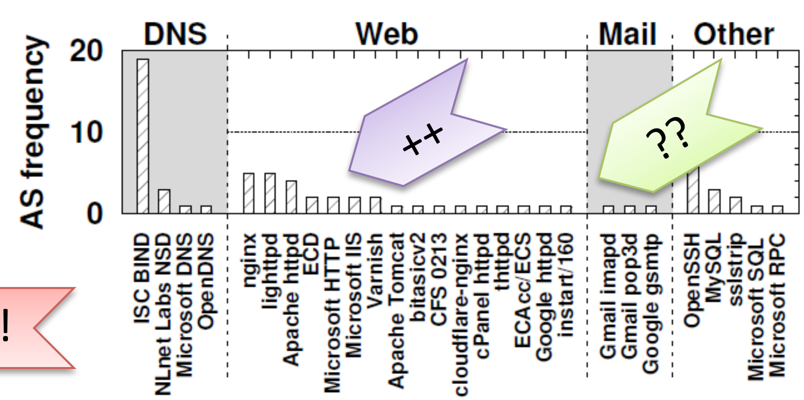
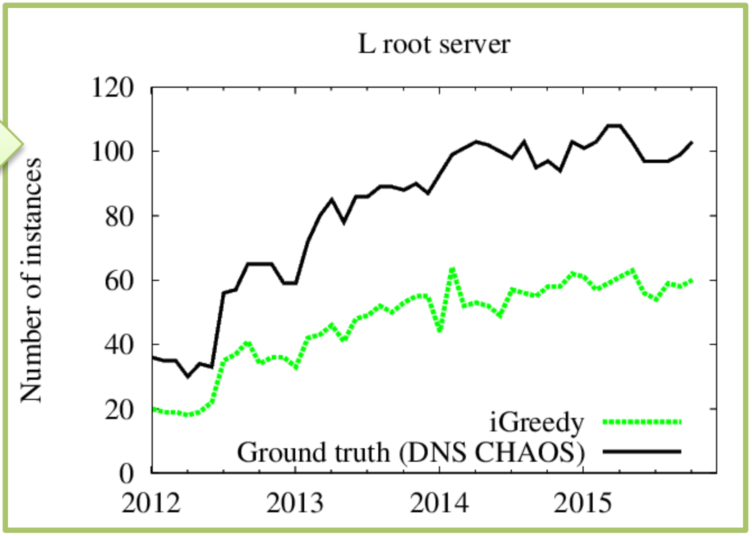
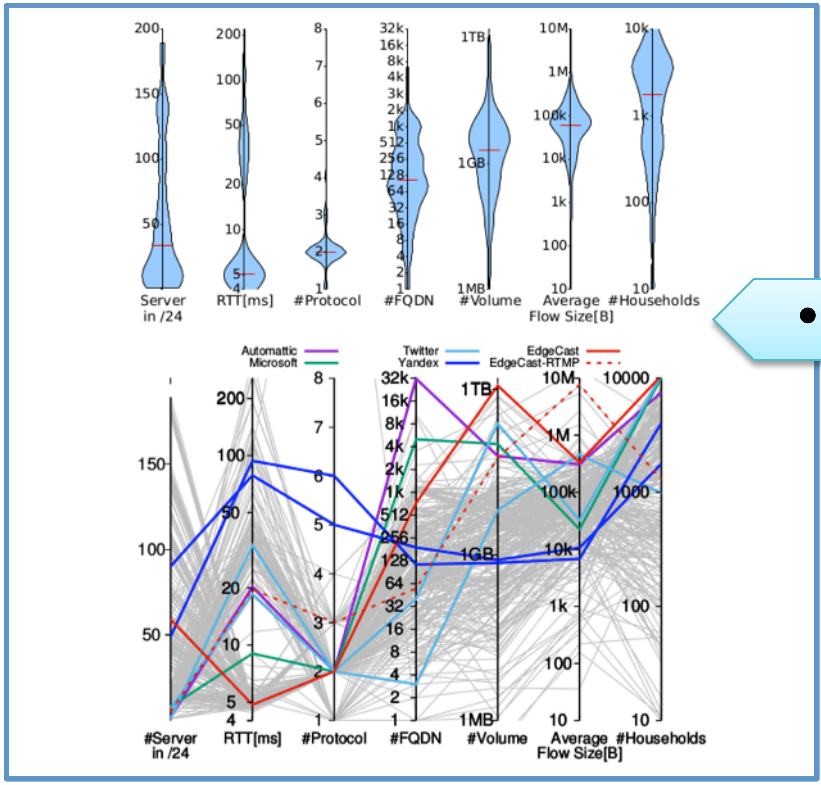


Figure 10: Breakdown of software running on anycast replicas.

(1/2) Infrastructure evolution & usage

- Time evolution of single deployment
  - Orthogonal to spatial dimension of census
  - Example from historic RIPE Atlas data





- Anycast usage across deployments
  - Orthogonal to active measurement
  - Use passive measurement at some ISP point of presence

(2/2) BGP hijack detection

Credits: renesys

- IP anycast
  - Syntactically equivalent to a BGP hijack in the BGP lingo
  - Only difference: router authorized to advertise the prefix or not

- iGreedy for BGP hijack detection

Reactive scan on BGP announces

- Analyse BGP feed and issue iGreedy on suspicion

Problem

- BGP Hijacks are of short duration
- Control plane information may arrive late at some monitor

Proactive Internet-wide scan

- Scan all /24 prefixes every minute

Problem

- Over 100x faster than current speed
- More challenge, more fun!

Google Faculty Research Awards

Conclusions

- iGreedy novel technique to investigate and especially geolocate anycast deployment
  - ✓ Practical lightweight, fast and protocol agnostic
  - ✓ Ready open-source software to issue, analyze and display RIPE Atlas measurement (using your credits!)
  - ✓ Useful Web interface to (significant subset of) census results already available

Interested ? Drop an email dario.rossi@enst.fr ! (but cc danilo.cicalese@enst.fr to get a timely reply)

References

This talk:

[1] D. Cicalese , D. Joumblatt, D. Rossi, J, Auge, M.O Buob, T. Friedman. **A Fistful of Pings: Accurate and Lightweight Anycast Enumeration and Geolocation** , IEEE INFOCOM, 2015

[2] D. Cicalese , J. Auge, D. Joumblatt, T. Friedman, D. Rossi, **Characterizing IPv4 Anycast Adoption and Deployment** , ACM CoNEXT, Dec 2015

[3] http://www.telecom-paristech.fr/~drossi/anycast

Related:

[4] D. Madory, C. Cook, and K. Miao, "Who are the anycasters," Nanog, 2013.

[5] X. Fan, J. S. Heidemann, and R. Govindan, "Evaluating anycast in the domain name system." in *Proc. IEEE INFOCOM*, 2013.

[6] P. Boothe and R. Bush, "DNS Anycast Stability: Some Early Results," CAIDA, 2005.

[7] S. Sarat, V. Pappas, and A. Terzis, "On the use of anycast in DNS," in *Proc. ICCCN*, 2006.

[8] D. Karrenberg, "Anycast and bgp stability: A closer look at dnsmon data," Nanog, 2005.

[9] H. Ballani and P. Francis, "Towards a global IP anycast service," in *Proc. ACM SIGCOMM*, 2005.

[10] H. Ballani, P. Francis, and S. Ratnasamy, "A measurement-based deployment proposal for ip anycast." in *Proc. ACM IMC*, 2006.

[11] Z. Liu, B. Huffaker, M. Fomenkov, N. Brownlee, and K. C. Claffy, "Two days in the life of the DNS anycast root servers." in *Proc. of PAM*, 2007.

[12] M. Levine, B. Lyon, and T. Underwood, "Operational experience with TCP and Anycast," Nanog, 2006.

?? || //