LSU

College of

Engineering

School of Electrical
Engineering & Computer Science

College of
Engineering
School of Electrical Engineering
& Computer Science

MudHunter: Internet-Scale DNS Cache Snooping for Cyber Threat Intelligence

Bassel Succar[1]
bsucca1@lsu.edu

Joseph Khoury[1]
Joseph.khoury@lsu.edu

Antonia Affinito[2]
a.affinito@utwente.nl

Elias Bou-Harb[1]
ebouharb@lsu.edu

- [1]Louisiana State University, LA, USA
- [2]University of Twente, Enschede, NL

Introduction

Effective defense begins with visibility — and visibility requires continuous measurement.

Motivation

- *DNS is the Internet's backbone - processes trillions of lookups daily.*
- *Over 60% of traffic flows through a few public resolvers (Google, Cloudflare, Quad9, OpenDNS).*
- *Resolvers resolve both benign and malicious activity:*
  - *Command-and-control (C2) operations*
  - *Phishing sites*
  - *DNS tunneling and data exfiltration*

Google Public DNS

OpenDNS

quad9

CLOUDFLARE®

Motivation

- *DNS is the Internet’s backbone - processes trillions of lookups daily.*

🧠 ***Insight: Public DNS resolvers are a gold mine for Cyber Threat Intelligence.***

- *Command-and-control (C2) operations*
- *Phishing sites*
- *DNS tunneling and data exfiltration*

Google Public DNS

CLOUDFLARE

Limitations of Current approaches

- *Passive DNS: valuable, but incomplete and delayed (multi-hour lag).*
- *Sinkholes: offer direct visibility but works for domains you can register, take over, or that someone else (e.g., an ISP) agrees to redirect.*
- *DNS logging or telemetry: rarely available at Internet scale due to privacy and jurisdictional constraints.*
- *Result: no scalable, privacy-preserving way to estimate domain activity across resolvers.*

💡 *DNS Cache Snooping - A Promising Alternative* 👀

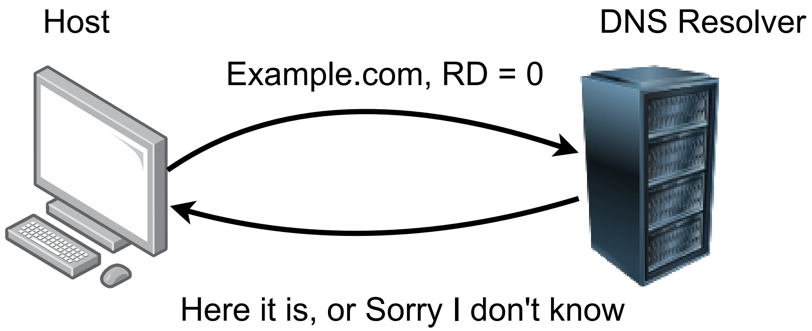
- *Uses non-recursive queries (RD=0) to check if a domain is cached → evidence of recent lookups.*
- *Requires no cooperation from resolvers or domain owners.*
- *Privacy-preserving: no user data, only aggregate cache state.*
- *Enables real-time, global-scale estimation of domain activity across networks.*

What is Cache Snooping ?

- *DNS uses caching to reduce latency and offload authoritative servers. Each record includes a Time-to-Live (TTL) , the time (in seconds) a resolver can reuse the cached answer.*

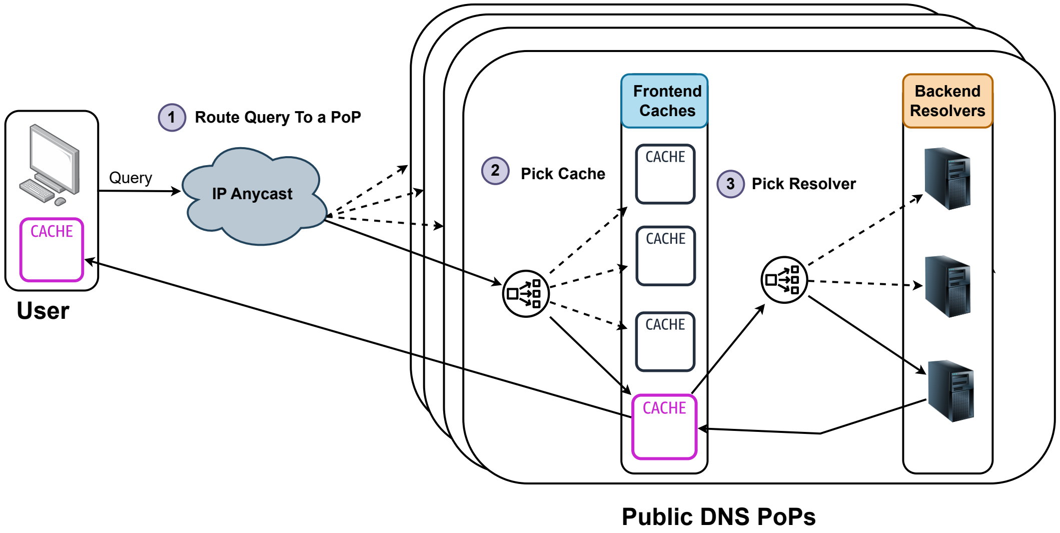
- *During this TTL window, the resolver serves the cached result instead of re-querying upstream servers.*

- *A non-recursive query (RD = 0) only returns answers already in cache or an empty response if not cached.*



A deeper look into Public DNS Caching architecture

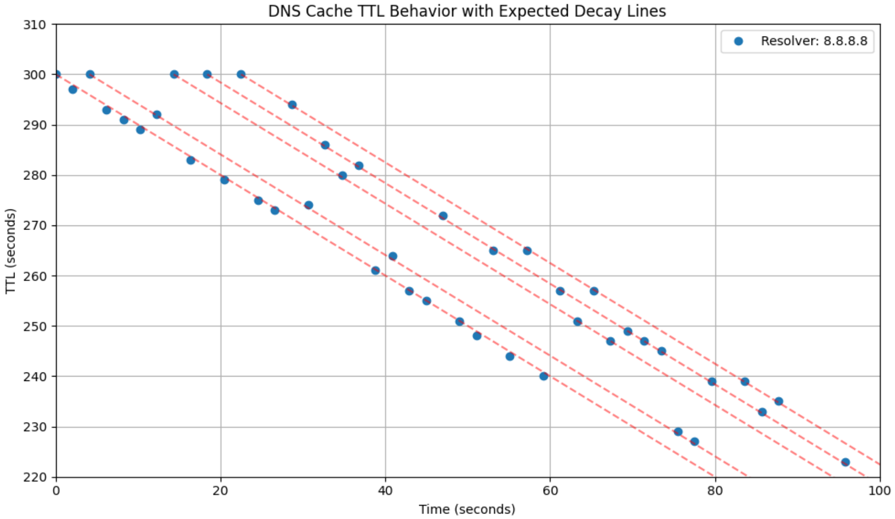
- *Public DNS Points of Presence (PoP) are distributed to handle the big load of traffic they receive.*
- *Each PoP has 2 layers: front-end caches and back-end resolvers, with load balancers in between.*
- *Due to local caching, one user can fill one cache for the duration of the TTL of the domain.*



Domain Activity Estimation Through Cache Snooping

AHA

By sending non-recursive (RD=0) queries repeatedly to the same PoP and observing TTL Values, we can estimate how many independent caches hold a domain giving a lower bound on how many users queried it.



Building on TruffleHunter

TruffleHunter (IMC 2020) first proved that DNS cache snooping could estimate global domain activity, but required manual, per-node deployments that limited real-world use.

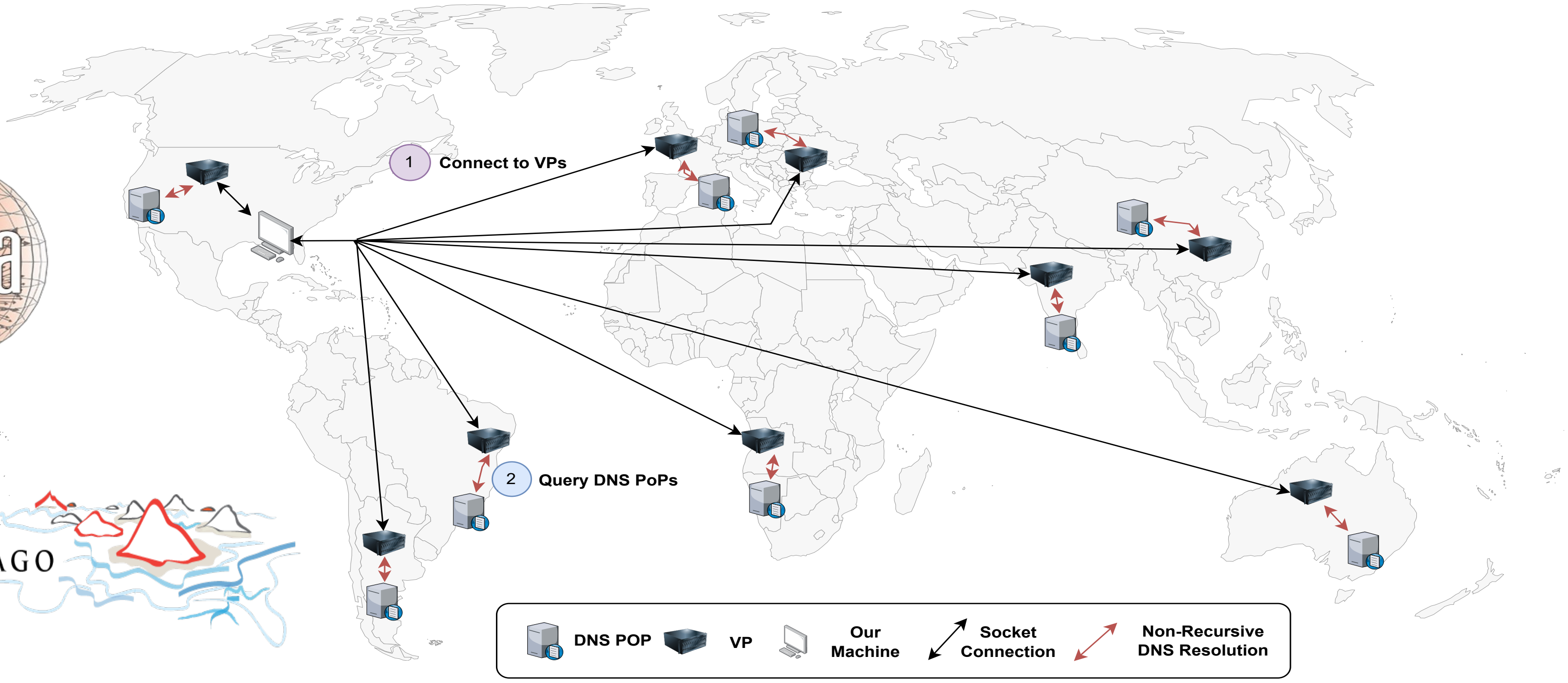
MudHunter removes this barrier automating Internet-scale measurements and turning cache data into real-time, geographically resolved threat intelligence.

Link to TruffleHunter Paper: https://dl.acm.org/doi/10.1145/3419394.3423640

Methodology

caida

ARCHIPELAGO



Measurement Phases

PoP Discovery & mapping

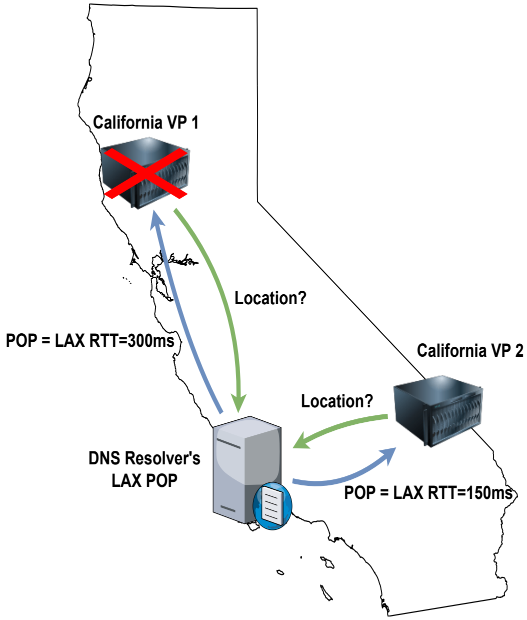
- *Different resolvers expose PoP info differently:*
- *Google → o-o.myaddr.l.google.com*
- *Cloudflare / Quad9 → CHAOS TXT id.server*
- *OpenDNS → TXT debug.opendns.com*
- *Aggregated into a global VP-to-PoP map (using IATA airport codes).*
- *Refreshed periodically to track routing changes.*

Vantage Point Filtering

- *Multiple VPs may route to the same PoP → redundant measurements.*
- *For each PoP, select one VP with the lowest RTT (fastest path).*
- *Avoids double-counting TTL lines and unnecessary probe traffic.*

Parallel Cache Probing

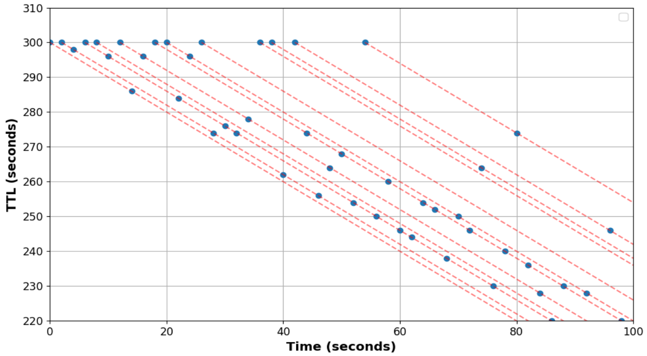
- *Selected VPs send 50 non-recursive (RD=0) queries to four resolvers: Google, Cloudflare, Quad9, OpenDNS.*
- *Each query round records TTL values → aggregated by {resolver, PoP}.*
- *Insert 2s delay between rounds to respect rate limits.*



Cache Filling Experiment

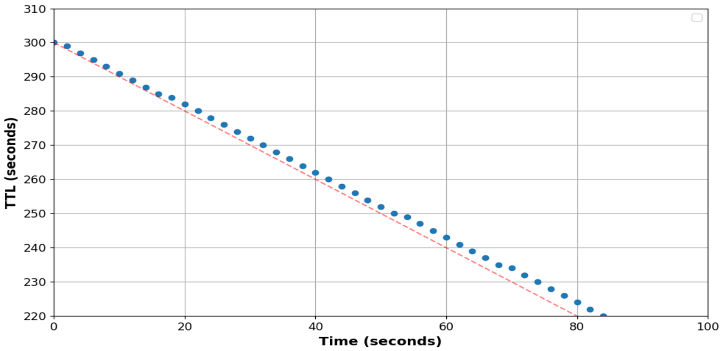
- *Registered our own domain with TTL = 300s to observe resolver behavior in isolation.*
- *Sent recursive (RD=1) queries every 2 s from each vantage point × 50 rounds.*
- *TTL = 300 → authoritative; TTL < 300 → cached response.*
- *Tracked TTL decay to visualize cache filling patterns per resolver.*

- *GPDNS, Quad9 and OpenDNS operates using independent caching.*
- *Cloudflare Operates using unified Caching.*



GPDNS,Quad9,OpenDNS

Cloudflare



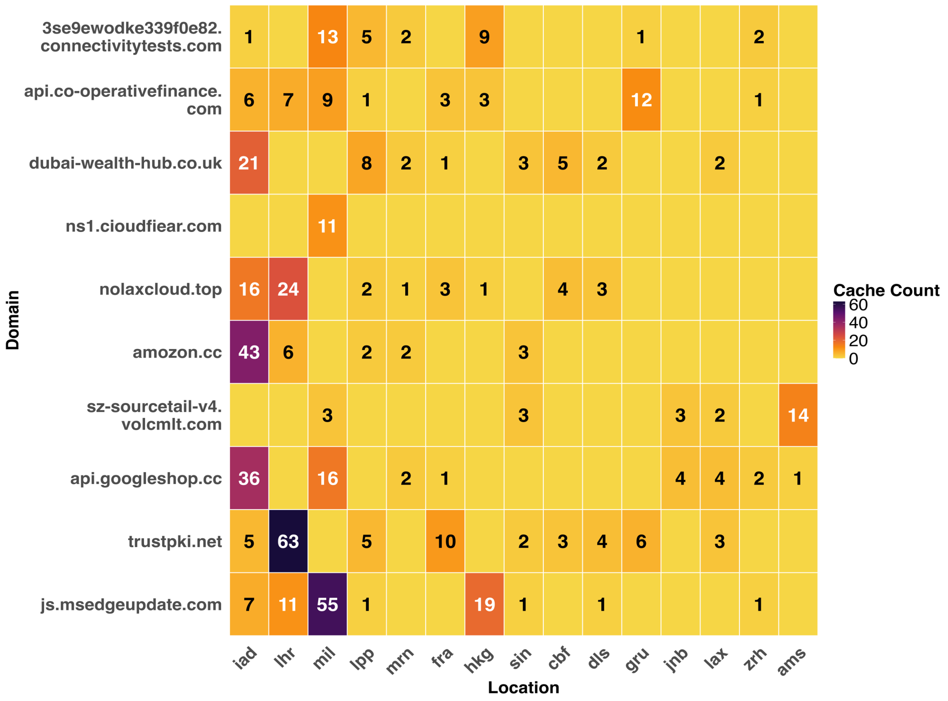
Mapping Botnet C2 Infrastructure

- *Botnets depend on DNS to locate and control C2 servers.*

- *MudHunter probed 1,247 verified C2 domains (Apr 3-10 2025) every 6 hours via the 4 major resolvers.*

- *Used cache-based heatmaps to reveal regional C2 activity.*

- *Found concentrated hotspots, not uniform global spread → evidence of regionally targeted botnet operations.*

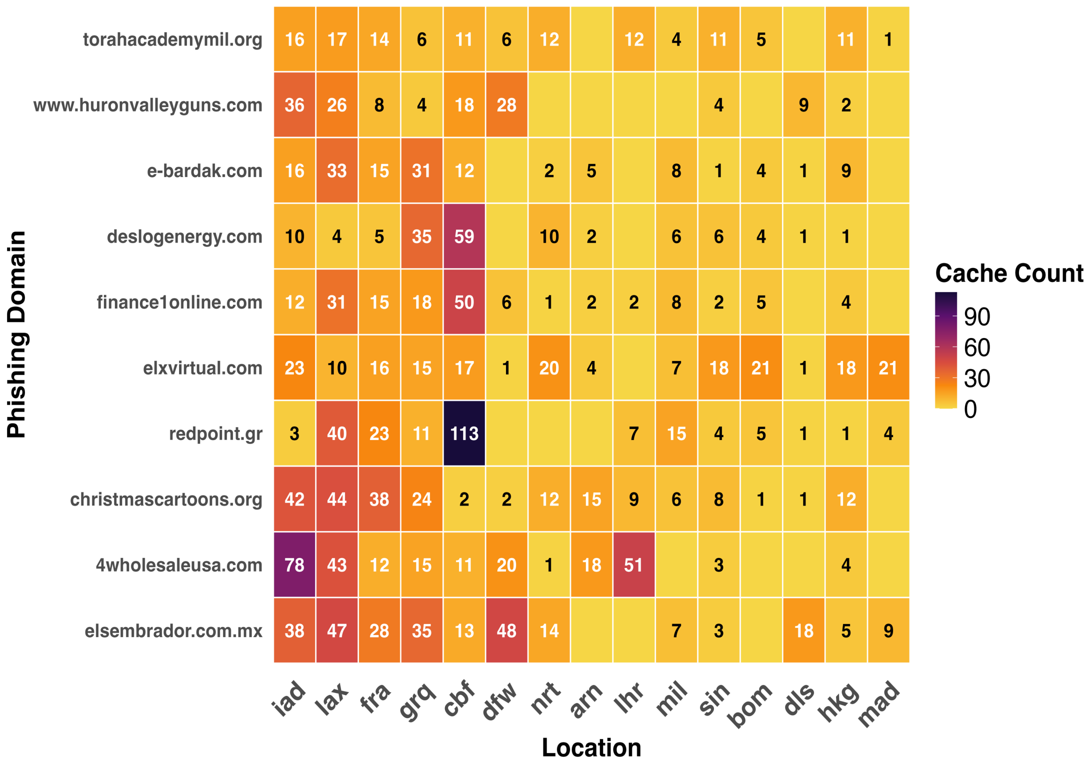


Domain	iad	lhr	mil	lpp	mrn	fra	hkg	sin	cbf	dls	gru	jnb	lax	zrh	ams
3se9ewodke339f0e82.connectivitytests.com	1		13	5	2		9				1			2	
api.co-operativefinance.com	6	7	9	1		3	3				12			1	
dubai-wealth-hub.co.uk	21			8	2	1		3	5	2			2		
ns1.cioudfiear.com			11												
nolaxcloud.top	16	24		2	1	3	1		4	3					
amozon.cc	43	6		2	2			3							
sz-sourcetail-v4.volcmlt.com			3					3				3	2		14
api.googleshop.cc	36		16		2	1						4	4	2	1
trustpki.net	5	63		5		10		2	3	4	6		3		
js.msedgeupdate.com	7	11	55	1			19	1		1				1	

Tracking Banking Phishing Domains

- *Phishing domains are short-lived (often hours) and use homographs, subdomain abuse, and typosquatting to evade detection.*
- *MudHunter probed 892 verified banking phishing domains (Apr 3-10 2025) every 6 hours via the 4 major resolvers.*

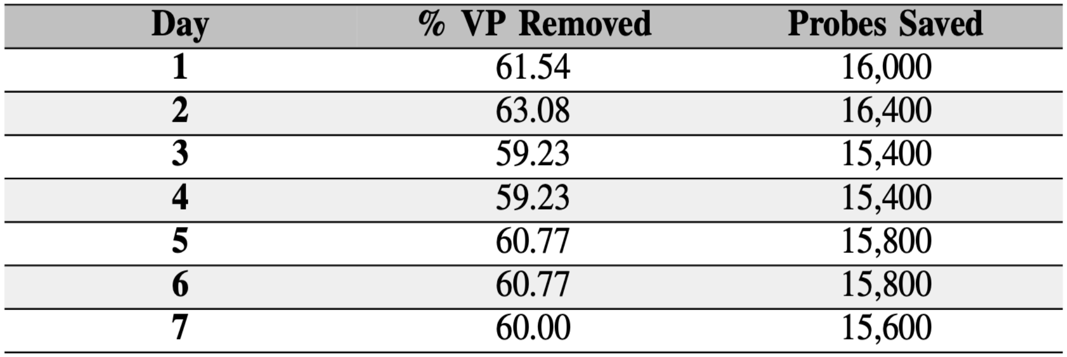
- *Activity shows localized, short-term campaigns, not global spread, consistent with targeted phishing operations.*



Phishing Domain	iad	lax	fra	grq	cbf	dfw	nrt	arn	lhr	mil	sin	bom	dls	hkg	mad
torahacademymil.org	16	17	14	6	11	6	12		12	4	11	5		11	1
www.huronvalleyguns.com	36	26	8	4	18	28					4		9	2	
e-bardak.com	16	33	15	31	12		2	5		8	1	4	1	9	
deslogenergy.com	10	4	5	35	59		10	2		6	6	4	1	1	
finance1online.com	12	31	15	18	50	6	1	2	2	8	2	5		4	
elxvirtual.com	23	10	16	15	17	1	20	4		7	18	21	1	18	21
redpoint.gr	3	40	23	11	113				7	15	4	5	1	1	4
christmascartoons.org	42	44	38	24	2	2	12	15	9	6	8	1	1	12	
4wholesaleusa.com	78	43	12	15	11	20	1	18	51		3			4	
elsembrador.com.mx	38	47	28	35	13	48	14			7	3		18	5	9

Vantage point filtering Effectiveness

- *≈ 60.8 % (± 1.4 %) of VPs filtered daily → ~79 / 130 VPs removed*
- *Each removed VP would have sent 200 probes (k = 50 × 4 resolvers)*
- *Optimization saves ≈ 15.7 K DNS probes per domain (range 15.4 K–16.4 K)*



Day	% VP Removed	Probes Saved
1	61.54	16,000
2	63.08	16,400
3	59.23	15,400
4	59.23	15,400
5	60.77	15,800
6	60.77	15,800
7	60.00	15,600

Limitations Of Cache Snooping

- *Load Balancers non-determinism: Probes hit different FE/BE → false misses, undercount.*
- *Resolver policy variance for RD=0: REFUSED / SERVFAIL → blind spots.*
- *Geo sparsity: Limited VP coverage.*
- *TTL churn & eviction: Races between user hits and probes skew results.*
- *VPNs/proxies can skew geo inference: cache hits may reflect shared VPN/proxy infrastructure rather than unique local users, concentrating activity at certain PoPs and blurring true location.*

Conclusion

- *Our experiments confirmed that cache snooping remains viable across today's major public resolvers, offering renewed visibility into domain activity.*

- *By coordinating 130 vantage points worldwide, it transforms cache snooping from a research trick into a reproducible measurement system.*

- *Through this lens, resolver caches become signals, quietly reflecting where malicious infrastructure is active without touching user data.*

- *Ultimately, MudHunter lowers the barrier for global active DNS intelligence, empowering defenders and researchers to measure, not guess, where threats emerge.*

Thank you