DarkDNS: Revisiting the Value of Rapid Zone Update

Raffaele Sommese
University of Twente
r.sommese@utwente.nl

Gautam Akiwate
Stanford University
gakiwate@cs.stanford.edu

Antonia Affinito
University of Twente
a.affinito@utwente.nl

Moritz Muller
SIDN Labs and University of Twente
moritz.muller@sidn.nl

Mattijs Jonker
University of Twente
m.jonker@utwente.nl

KC Claffy
CAIDA
kc@caida.org

ABSTRACT

Malicious actors exploit the DNS namespace to launch spam campaigns, phishing attacks, malware, and other harmful activities. Combating these threats requires visibility into domain existence, ownership and nameservice activity that the DNS protocol does not itself provide. To facilitate visibility and security-related study of the expanding gTLD namespace, ICANN introduced the Centralized Zone Data Service (CZDS) that shares daily zone file snapshots of new gTLD zones. However, a remarkably high concentration of malicious activity is associated with domains that do not live long enough make it into these daily snapshots. Using public and private sources of newly observed domains, we discover that even with the best available data there is a considerable visibility gap in detecting short-lived domains. We find that the daily snapshots miss at least 1% of newly registered and short-lived domains, which are frequently registered with likely malicious intent. In reducing this critical visibility gap using public sources of data, we demonstrate how more timely access to TLD zone changes can provide valuable data to better prevent abuse. We hope that this work sparks a discussion in the community on how to effectively and safely revive the concept of sharing Rapid Zone Updates for security research. Finally, as a contribution of this work, we are releasing a public live feed of newly registered domains, with the aim of enabling further research in early abuse identification.

CCS CONCEPTS

• **Networks** → **Naming and addressing**; *Network measurement*;

1 INTRODUCTION

Malicious actors exploit (abuse) the DNS namespace to launch spam campaigns, phishing attacks, malware, and other harmful activities. In many dimensions the DNS ecosystem is more opaque than other

aspects of Internet transport. Unlike BGP, DNS is a pull protocol, so learning internal dynamics requires an entry point *i.e.,* a domain name. Without knowing this entry point, any abuse that lies behind a domain remains opaque to everyone except the targets.

This inherent opacity has motivated many attempts to improve visibility into the DNS ecosystem to facilitate research, analysis, and improved operational security. A prominent example is the ICANN Centralized Zone Data Service (CZDS) which coordinates collection and restricted sharing of daily zone file snapshots of participating TLDs. However, previous studies have found that the 24-hour zone files, while an unprecedented enabler of DNS-related security research, leave vital visibility gaps in studies of abuse and exploits [13, 14, 35]. Since so much malicious activity often begins soon after a domain is registered for that purpose, and the CZDS snapshots are updated only daily, commercial threat intelligence services [21] have tried to fill the gap by detecting domains soon after registration.

In this work we try to quantify, using public and private sources of data, the gaps in the current models of data sharing, both regulatory and market-driven. We focus on domains that are created and removed in under 24 hours, and thus do not appear in the CZDS zone files. The likely predominant reason for such *transient* domains is a fraudulent or malicious activity. For example, if registrars receive evidence of malicious use or flag potentially fraudulent payment methods after they process the registration, the registrar will remove the domain from the zone immediately *i.e.,* before the next CZDS zone snapshot. Unfortunately, abuse prevention currently relies solely on the registrars. Each registrar must independently identify the same signals to block the same threat actors, who continuously move across registrars to evade detection. This results in an endless cat-and-mouse game where attackers maintain the upper hand. Fine-grained visibility into TLD zone changes can enable verified third parties such as security researchers to build signals to prevent abuse and raise the cost for threat actors.

In using public and private sources of data to get fine grained visibility into TLD zone changes, we can reduce but not completely close the significant visibility gap. Using ground truth from a medium-size European ccTLD registry, we detected only one-third of these transient domains even with the best public available data. Given the measurable expansion of cybercrime, and the failure of current data sharing models to effectively combat it, security-conscious TLDs should consider offering a subscription service to a feed of changes to zones immediately as they occur. Verisign offered such a service for the .com TLD 15 years ago [25] but later terminated it due to its potential for abuse. We recognize these concerns of abuse but believe now is the time to develop a framework to

manage these concerns, just as ICANN is doing for access to even more sensitive types of registration data [26].

To that end, we make the following contributions:

(1) We develop, apply, and validate a method that uses CT logs to estimate a lower bound on the visibility gap in the daily CZDS snapshots, a blind spot for defenders that even commercial threat intelligence sources do not capture.
(2) We provide an open data feed of newly registered domains (including transient) we discover [33].
(3) We investigate the infrastructural landscape of transient domains and the possible reasons for their early removal.
(4) We discuss how to better support transparency while balancing privacy and commercial interests.

2 RELATED WORK

Defenders of networked infrastructure have long had incentives to improve detection of malicious domains early in their lifecycle. A 2020 study found that the average phishing campaign from start to the last victim took 21 hours [32]. A parallel study found phishing indicators that blocklists were unable to detect, concluding that better protection against phishing would require expansion of evidence-based reporting protocols [31].

Barron *et al.* [16] found significant correlation between early domain name deletions and potentially malicious activities, and noted that often zone files do not reveal whether a domain is registered. Other studies have found that spammer domains generally had short lifetimes, and that blocklists generally did not detect them until they were actively used and reported rather than when initially registered [24]. Affinito *et al.* [15] found patterns of suspiciously short lifetimes domain names in some TLDs (*e.g.*, 80% under .xyz), many of which were reported as malicious. Their study revealed that domains registered under some TLDs (*e.g.*, .xyz, .icu) are taken down within a few days, while others (e.g., .top) remain in the zone longer before removal.

Commercial threat intelligence services have tried to narrow this visibility gap by monitoring DNS queries in traffic. In 2018 Domain Tools reported that most newly observed domains become inactive after 4 hours [34], presumably because they have been detected and flagged as malicious and no longer serve their purpose. Observing DNS queries from their CDN, Akamai identified 20.1% (13M) of all newly registered domain they detected as malicious. [12]. In 2022, Palo Alto Newtorks reported that 70% of newly registered domains detected from passive DNS data and zone files across 1530 TLDs were *malicious* or *suspicious* or *not safe for work* [19].

In the spike of COVID-related fraudulent online activity, a group of threat intelligence providers temporarily volunteered to share threat data in hopes of achieving better coverage. They found that for novel abuse types, of which there were many during COVID, the aggregated threat intelligence detected signals that registrars and registries did not [17].

A recent study is relevant to our goal of detection of newly observed domains without access to proprietary data. Last year, Sommese *et al.* [36]. demonstrated that CT transparency logs are a reliable and up-to-date source of domain names for addressing the visibility gaps in ccTLDs, most of which do not participate in CZDS. They used CT logs to independently reconstruct more than

half of the domains in several ccTLDs for which they had the full zone file as ground truth. Their work focused on spatial coverage whereas we use CT logs to analyze temporal gaps, i.e., how quickly we detect new and transient domains [36].

3 METHODOLOGY

Our approach to estimating the extent of newly registered domains not captured by daily zone snapshots relies on public data sources: available zone snapshots, certificate transparency logs [28], active DNS resolution measurements, and registration data (RDAP) [30]. Unlike passive DNS, our approach does not require the domain to be actively queried by users, but the the domain must have an issued certificate. Our pipeline has five steps: identify registered (pay-level) domains for which certificates are issued but are not present in the latest CZDS snapshots; collect RDAP data; monitor DNS changes of newly observed domains by performing active measurements; cross-validate inference with RDAP data; and identify transient (short-lived) domains. We feed the results of each measurement into Kafka topics and store them in Parquet format in our object storage for longitudinal analysis.

Step 1: Infer newly registered domains from Certificate Transparency logs. We start with the daily zone files extracted by our collector, which is populated with all latest zones snapshots available from ICANN CZDS. We use the open-source Certstream [18] to captures logs of newly issued certificates, [1] and extract domain names from the Common Name (CN) and the Subject Alternative Name (SAN) fields of these certificates. We discard domain names that already appeared in the latest daily zone snapshots. This step has three limitations. First, some CAs do not always perform Domain Validation step[2] [29], so domains with certificates fewer than 398 days old may no longer exist (§4). Second, zone file publication may be delayed by days, leading to inaccurate inference of domain existence. Lastly, we are able to detect domains only if a related certificate is issued. This step yields a stream of potentially newly registered domain names that we feed into our analysis pipeline. While, this stream contains only domains with an associated issued certificate, it represents a publicly accessible data feed for researchers, in contrast with commercial passive DNS feeds. We release this feed publicly at [33].

Step 2: Collect RDAP registration data. We verify how accurately we detect a newly registered domain using its RDAP-reported registration date and time. In addition, we collect registrar and registrant information, the latter of which is typically redacted to protect privacy. To collect RDAP data, we deploy a script in Azure cloud that retrieves newly registered domains from our Kafka topic and uses the `whoisit` Python library to perform RDAP queries for each domain. We leverage Azure functions to avoid aggressive rate-limiting by cycling measurements over different IPv4 addresses. To minimize overhead, we did not retry failed queries.

---

[1] We only consider PreCertificate entries, because they must always be published before the issuance of the final certificate [28]
[2] Per Section 4.2.1 of the CA Browser Forum baseline requirements [22], a CA may reuse cached validation information collected no more than 398 days prior to issuing the Certificate.

Step 3: Monitor changes to hosting and DNS providers for newly observed domains. A reactive measurement infrastructure powered by our Kafka streaming pipeline issues DNS (A, AAAA, and NS) queries every 10 minutes to these domains for the first 48 hours of their existence. Sixteen instances of our measurement worker, backed by Unbound (a caching resolver), execute these queries from 16 separate nodes. Our resolver is configured with a maximum 60-second cache to avoid cached A and AAAA records. For NS queries, the measurement workers send queries directly to the domain's TLD authoritative nameserver to more accurately infer domain removal from the zone, and to prevent misclassification of lame delegated or misconfigured domain names as deleted.

Step 4: Validate inference against RDAP data. We check that the RDAP registration timestamp is consistent (within 24 hours) with the observed certificate timestamp. The difference between these two timestamps provides a ground truth indication of how correctly and promptly we detect newly registered domain names. This step only works when our RDAP resolution succeeded ($\approx$ 97% of cases[3]).

Step 5: Identify transient (short-lived) domains, which we define as those registered and deleted in the time span (max 24 hours) between the two zone snapshots. Note that we do not include short-lifetime domains that the zone snapshots do manage to capture. In our data set, these transient domains are the subset of newly registered domains that do not appear in any zone file throughout our analysis window (1 Nov 2023 - 31 Jan 2024).

4 RESULTS

Recall, our methodology uses CT logs to identify *use of domains before* they show up in the CZDS zone snapshots. From 1 Nov 2023 to 31 Jan 2024, we found 6.8 million domain names (Table 1) that did not appear in the CZDS zone snapshot before we saw them in the CT logs. These 6.8*M* domain names are either transient domains or eventually show up in the next CZDS zone snapshots. Most of these domains (3.7M) were in `.com`, followed by `.xyz` (300K) and `.shop` (284K).

When comparing our list of detected newly registered domains to the diff between two daily CZDS zone snapshots, we found that our methodology identified 42% of the newly registered domains before they were published in the zone snapshots in our 3-month observation period. The remaining 58% did not have certificates issued immediately. Since our goal is to identify transient domains *i.e., ,* domains that do not show up in zone snapshots, this comparison is purely functional.

4.1 Detection Speed

Figure 1 plots the time difference between when we recorded the domain (*i.e.,* fetched it from CT Logs[4]) and the RDAP-reported timestamp. We detected half of the newly registered domains within the first 45 minutes of their existence, and $\approx$30% within the first 15 minutes. The distribution has a long tail toward 1 day (less than 2% have a difference greater than a day), which we believe derives from our misclassification of newly registered domains (*e.g.,* due

---

[3]We missed $\approx$ 3% due to RDAP rate limiting or other collection errors.
[4]We use the Certstream-reported timestamp because neither precertificates nor CT logs provide an insert timestamp
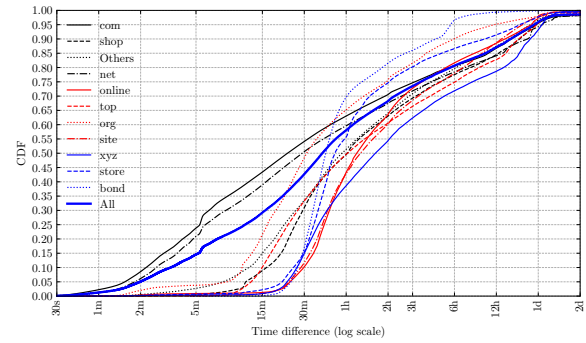
Figure 1: Difference in registration time per RDAP vs. CT logs. We detected 50% of domains within 45 minutes of their existence, and ≈30% within 15 min.

to incorrect extraction of the second-level domain (SLD) using the Public Suffix List) or zones published with several days of delay. The time differences across TLDs can be explained by the operational update times of their zones (*i.e.,* the domain needs to be active in the zone before a certificate can be issued). For example, the `.com` and `.net` registries update these TLD zones on average every 60 seconds while other gTLDs registries update their zones every 15-30 minutes. We validated this assumption by probing the zones of Figure 1 for SOA serial changes, and found consistent timestamps.

This result demonstrates that our methodology correctly and quickly identifies newly registered domains. The RDAP data helps to exclude corner cases and misclassified (as newly registered) domain names, and provides relevant metadata (*e.g.,* registrar identity, accurate creation time).

Furthermore, our reactive measurements allowed us to provide a lower-bound estimation of DNS infrastructural changes. In our data set, most (97.5% of) newly registered domains kept their initial nameserver infrastructure for the first 24 hours. A few (2.5%) changed NS infrastructure within the first 24 hours such that the daily zone snapshot differences could miss it, depending when in the 24-hour window the change occurred.

4.2 Transient Domains

While early detection of newly registered domains represents a significant advance detecting possible malicious behavior, those domains will generally show up in the zone snapshots in the days following their registration. However, a small percentage (1%) of newly registered domains never appeared in a zone snapshot due to their short lifespan falling between two snapshots. We call these domains *transient domains.* We identified these names by excluding from our list of newly registered domains those that appeared in our zone collection during the 3 month observation window (+/- 3 days to account for late zone snapshots publishing).

Our method inferred a lower bound of 68,042 transient domains (Table 2), i.e., ≈1% of CT-observed newly registered domain names. The TLD share of these two populations (transient and newly registered) differs (Table 1 vs. Table 2), perhaps related to TLD pricing oscillation and bulk malicious registration campaigns [1, 27].

Table 1: Top 10 TLDs ranked by count of newly registered domains (NRDs) for Nov 23 - Jan 24). The Zone NRD columns shows how many NRDs appeared in the daily zone file snapshots.

TLD Category	Nov	Dec	Jan	Total	Zone NRD	Coverage NRD (%)
com	1 127 727	1 109 804	1 505 044	3 742 575	8 467 641	44.2%
xyz	114 582	87 051	107 740	309 373	649 010	47.7%
shop	76 626	99 660	107 675	283 961	775 253	36.6%
online	76 674	76 693	109 964	263 331	648 922	40.6%
bond	75 779	81 265	84 997	242 041	292 552	82.7%
top	82 746	74 134	83 837	240 717	532 363	45.2%
net	79 660	71 922	84 320	235 902	643 030	36.7%
org	53 377	53 767	76 400	183 544	481 870	38.1%
site	46 695	47 879	65 801	160 375	465 542	34.4%
store	42 931	38 699	50 279	131 909	326 383	40.4%
Others	328 570	333 000	380 551	1 042 121	3 009 575	34.6%
Total	2 105 367	2 073 874	2 656 608	6 835 849	16 292 141	42.0%

Table 2: Transient domain names observed

TLD Category	Nov	Dec	Jan	Total
com	9363	10 597	21 232	41 192
online	1800	2369	1990	6159
site	1578	1381	890	3849
net	702	866	1544	3112
org	595	602	1176	2373
shop	688	497	507	1692
xyz	321	316	624	1261
store	422	414	377	1213
top	213	161	276	650
fun	185	175	160	520
Others	1609	1958	2454	6021
Total	17 476	19 336	31 230	68 042

During our data collection, we noticed that RDAP failure rate of transient domains was noticeably higher (≈34%) than for newly observed domains (≈3%). We identified three major causes for this higher failure rate: (i) we detected too late, i.e., when we attempted to collect the RDAP data, the domain had already been removed; (ii) we were too early, *i.e.,* RDAP data was not yet in sync, and (iii) we detected domains that no longer existed but for which a certificate had been issued.

While we cannot investigate causes (i) and (ii) due to lack of RDAP data, we examined several domains in the third category, contacting the CERT teams of the CAs that issued certificates for these non-existent domains. GlobalSign, Sectigo, and Cloudflare confirmed that issuing a certificate for a non-existent domain (more precisely, issuing a certificate without validating the existence of the domain) is allowed under the condition that the CA possesses a previously obtained and valid Domain Validation (DV) token (see §3 and footnote 2). A necessary but not sufficient condition for this case is that the domain existed in the past. We conducted a comparison of the 34% of transient domains for which we failed to collect RDAP data with the CAIDA DZDB historical zone collection [23] and found that approximately 97% of those domains were registered

in the past. Another possible approach to filter out these domains is to individually check if a certificate was issued in the DV token validity period in the past. We leave the complexity of longitudinal investigation of CT Logs as future work.

Since domains with non-responding RDAP may not be transient (cause iii), we filtered them from subsequent analysis. Using the RDAP-reported registration timestamp, we also filtered domains misclassified as newly registered. This filtering yielded 42358 confirmed transient domains over the 3-month period. These transient domains represent only a portion of all the transient domains across the zones we analyzed. Because we do not have access to ground truth (*i.e.,* registries view), we cannot establish the total number of transient domains or the extent of our methodology's coverage. In subsection 4.4, we will compare our detection with a passive DNS source and the ground truth from a ccTLD.

4.2.1 Lifetimes of Transient Domains. To analyze how quickly transient domains disappear from their zone, we subtracted the last time the TLD nameserver provided a valid response for the NS query for that specific domain from the RDAP registration time of the domain. Per this method, over 50% of these domain names died within their first 6 hours of life (Figure 2).

4.2.2 Registrar and Hosting Landscape of Transient Domains. The collected RDAP data revealed that market leader GoDaddy topped the list of registrars holding transient domains, with 19% of such domains (Table 3). The registrar landscape is dominated by large registrars, with smaller registrars making up the long tail of the distribution (see Others in Table 3). These results suggest that transient domains are a widespread phenomenon across registrars.

We used our active resolution measurements to analyze the DNS hosting (Table 4) and web hosting (Table 5) infrastructure of transient domains. For DNS hosting, we investigated the most popular nameservers SLDs, while for Web hosting we looked at the ASNs of the A records of the domains. Half of these transient domains were using Cloudflare as DNS provider (*i.e.,* for their authoritative nameservers) and ≈35% of them used Cloudflare as a CDN provider. We identified Hostinger's parking DNS nameservers as the second most prominent DNS provider, as accounting for ≈8%
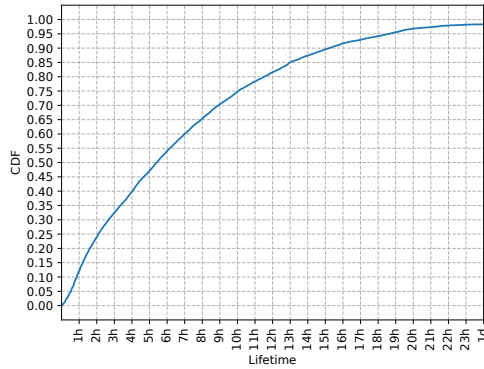
Figure 2: Lifetime of transient domain names.

Table 3: Top 10 Transient Domains Registrars Distribution.

Registrar	Domains	%
GoDaddy	8213	19.39%
Hostinger	6418	15.2%
NameCheap	4195	9.9%
Squarespace	2820	6.7%
Public Domain Registry	2625	6.2%
IONOS	2352	5.6%
Metaregistrar	1866	4.4%
NameSilo	1853	4.4%
Network Solutions, LLC	1670	3.9%
Tucows	1304	3.1%
Others	9042	21.3%
Total	42358	-

of the transient domains, followed by a long tail of popular DNS hosting services including NS1, Squarespace, and GoDaddy. These findings suggest that those domains are generally hosted on similar infrastructure of long-lived domains. In contrast, the large presence of parked domains could indicate domains that have already been removed or caught before misuse.

Table 4: Top 5 DNS Hosting Domains (NS records) of Transient Domains

Name	NS Record SLD	Domains	%
Cloudflare	cloudflare.com	20981	49.5%
Hostinger	dns-parking.com	3682	8.7%
NS1	nsone.net	2938	6,9%
Squarespace	squarespacedns.com	2908	6.9%
GoDaddy	domaincontrol.com	2315	5.5%
Others	-	9534	22.5%
Total	-	42358	-

Table 5: Top 5 Web Hosting (A records) of Transient Domains

Name	ASN	Domains	%
Cloudflare	13335	15322	36.2%
Hostinger	47583	5930	14.0%
Amazon	16509	3198	7.6%
Squarespace	53831	2257	5.3%
Namecheap	22612	1650	3.9%
Others	-	14001	33.1%
Total	-	42358	-

4.3 Reasons for Early Removals

In private conversations with two top registrars in Table 3, they confirmed that transient domains are mostly likely malicious. The registrars did not comment on exact numbers but anecdotally acknowledged that with few exceptions, reasons for early removal include abuse, account suspensions, or credit card fraud. Some legitimate cases are domain tasting and right of cancellation, but they are exceptionally rare. To investigate indicators of maliciousness of transient domain names, we analyzed how often blocklists manage to flag them. We considered both early-removed newly registered domain names and 42358 transient domains (Table 3). With *early-removed*, we refer to newly registered domains that were removed before the end of our analysis period. Unlike transient domains, early-removed domains appear in the zone file snapshots but still have a shorter lifespan than the 1-year typical duration of a registered domain due to the likely forced removal by the registrar [15]. We observed 555491 domains that were deleted before 1 February 2024 (10% of newly registered domains we detected).

We examined ten blocklists daily from 1 November 2023, to 29 April 2024, extending beyond our observation period to capture late insertions: DBL [4], Phishtank [8], Phishingarmy [7], Cybercrime-tracker [2], Tolouse DDoS, Crypto, Malware [10], Digitalside [3], Openphish [5], Vxvault [11], Ponmocup [6], Quidsup [9].

Newly registered domains. Of 555491 newly registered domains, at least one blocklist in our set identified 6.6% (37188) as malicious. Of these, 92% (34233) were active when we detected them, 3% (1072) were already included in the blocklists before their registration date, and 5% (1882) appeared on blocklists after deletion.

Transient domains. Of the 42358 transient domains, at least one blocklist flagged 5% (2123) as malicious. Of these, 5% (105) were flagged on their registration date, 1% (12) were already on the blocklists before their registration date, and 94% (2006) appeared after domain deletion.

Summary. Blocklists identify and label as malicious a tiny percentage of domains before their registration dates, which may indicate a re-registration. Additionally, a significant percentage (94%) of transient malicious domains appeared on the blocklists post-deletion, indicating that blocklists do not promptly or in some cases ever detect short-lived malicious domains [17].

4.4 Visibility Gap in Domain Detection

Commercial threat intelligence services [21] using passive DNS have tried to fill the gap by detecting domains soon after registration, and are generally the best available data source for identifying transient domains. To understand the differences between identified transient domains using public and private data sources, we collected and compared one day (May 9, 2024) worth of newly registered domains provided by DomainTools's SIE Newly Observed Domain (NOD) feed [21] with our feed. To simplify the comparison and avoid accounting for domains detected after several days, we compared only domains that were registered on May 9, 2024, according to the RDAP data from both our feed and the SIE feed. We considered only gTLD domains because they were available in both data feeds.

We first compared newly registered domains, finding that the SIE NOD feed detected ≈5% more domains than our method. However, the overlap between the two data sources is ≈60%, indicating that each method detects a unique subset of domains. Looking at transient domains identified by the two sources, we find the overlap drops further. A total of 855 transient domains were identified by either one of the two sources. Of these, only 33% were detected by both, with the SIE NOD feed identifying 10% more transient than our method. While the SIE NOD detects a larger number, the disjoint nature of the intersecting set of domains suggests the need to combine them to narrow the visibility gap.

To validate our visibility gap against a source of ground truth, we use the perspective of a mid-sized ccTLD registry (.nl) on transient domains. From November 2023 until January 2024, the registry observed 714 domain names that were deleted in less than 24 hours in their registration system. Of those domains, 334 were registered and deleted such that they were never captured in zone file snapshots. Applying our methodology to this ccTLD, we found only 99 transient domains, or 29.6% of the 334 ccTLD-identified transient domain names.[5] This result shows that while we can catch a fraction of these transient domains, researchers still have a huge blindspot regarding intra-day events.

5 DISCUSSION AND CONCLUSION

It is well-established that domain-related cybercrimes generally run their course or affect the most victims within 24 hours of attack onset [20, 24, 32, 34]. The phenomenon of transient domains indicates some measure of success that registrars detect and take down so many short-lived domains in their early stages, before they can do damage. However, in this current model, each registrar has to independently re-learn the same signals as threat actors move across different registrars to evade detection. This state of affairs not only inflicts additional cost on registrars but also increases the time to remediate abuse. Even well-intended registrars and registries cannot always investigate reports of abuse in a timely manner, and the process for doing so is not well-defined [20].

In the meantime, transient domains, in which malicious activity tends to dominate, have largely been invisible to the research community. At the same time, research has established that the existing DNS-related threat intelligence community, *e.g.*, blocklist companies, are less effective against new and emerging threats [17], suggesting the need for new approaches.

Learning from history. Verisign long ago supported *rapid zone updates*, enabling updates to the .com and .net zone every 5 minutes. The data included domain names, nameservers, IP address additions, deletions and modifications, to "promote security and stability by providing a useful tool to online security companies, ISPs, search engines, financial services companies, and other stakeholders" [25], [B]. This service illuminated the activity of bad actors, including phishing, fraud and identity theft [24]. Concerns about abuse of the data by spammers led to Verisign's termination of this service. In particular, attackers exploited a brief time window where the zone updates were publicly available, but before domain owners had set up protection of their domain.

Resurrecting RZU. Given the inherent advantages that attackers have over defenders, the ineffectiveness of existing uncoordinated countermeasures, and the limited obligations (or ability) of registrars to mitigate harm,[6] we see a need to expand transparency and accountability mechanisms.

Registrars and registries who want to establish themselves as serious about security could resurrect the capability to subscribe to rapid zone updates along with a framework to safeguard against abuses, and to learn from history how to mitigate the risk of abuse. While risks of abuse of access to non-public DNS data has created tension and controversy for decades, the CZDS program is testament to the ability to manage these risks. More recently, ICANN has coordinated a model for requesting access to other non-public registration data in a more consistent and standardized format [26]. We advocate for the development of a similar model for data sharing among approved and trusted parties, including operators, law enforcement, and the research community. Such access to rapid zone updates would enable a source of labeled data that would allow application of machine learning techniques not only to security research on systemic harms, but also to anti-abuse efforts of registries, registrars, and law enforcement agencies.

Future Directions. To demonstrate the visibility gap of daily zone snapshots, we proposed a methodology to identify newly registered domains based on CT logs and active measurements, which we released as a public stream [33]. Our results shed light on previously undetected short-lived domains, often related to malicious activities, which we named transient domains. Preliminary analysis and contact with registrars suggest these domains are likely malicious. In the future we plan to expand our measurements beyond DNS infrastructure records, including mail extensions (*e.g.,* SPF, MX, *etc.*), subdomains, and web-crawling. These measurements, combined with early detection using our methodolgy, can support machine learning-based approaches to proactively identify malicious domains before they do harm. Proactive registries can leverage such use cases to justify and sustain a responsible RZU service.

---

[5]We were unable to compare the SIE NOD feed to the registry data because we had access only to one day's worth of data. However, based on our previous results we argue that combining the two data sources would still leave a large visibility gap.

[6]Interisle reports that ICANN's recently revised contract language does not attribute any duty of care to registrars or registries, and does not address vulnerabilities such as exact-match brand registration and suspicious bulk registration [20].

REFERENCES

[1] 2023. Phishing Landscape 2023: A Study of the Scope and Distribution of Phishing. https://www.cybercrimeinfocenter.org/phishing-landscape-2023.
[2] 2024. Cybercrime-tracker. (2024). https://cybercrime-tracker.net/.
[3] 2024. DigitalSide Threat-Intel Repository. (2024). https://osint.digitalside.it/.
[4] 2024. Domain Blocklist (DBL). (2024). https://www.spamhaus.org/blocklists/domain-blocklist/.
[5] 2024. Openphish. (2024). https://openphish.com/.
[6] 2024. Oracle. DynDNS.org Malware Feeds. (2024). http://security-research.dyndns.org/pub/malware-feeds/.
[7] 2024. Phishingarmy. (2024). https://phishing.army/.
[8] 2024. PhishTank. (2024). https://phishtank.org/.
[9] 2024. QuidsUp - NoTrack - Block List. (2024). https://quidsup.net/notrack/blocklist.php.
[10] 2024. Tolouse. (2024). https://dsi.ut-capitole.fr/blacklists/.
[11] 2024. Vxvault. (2024). http://vxvault.net/ViriList.php.
[12] Akamai. 2024. Flagging 13 Million Malicious Domains in 1 Month with Newly Observed Domains. (2024). www.akamai.com/blog/security-research/newly-observed-domains-discovered-13-million-malicious-domains.
[13] G Akiwate, M Jonker, R Sommese, I Foster, G Voelker, S Savage, and k claffy. 2020. Unresolved Issues: Prevalence, Persistence, and Perils of Lame Delegations. In *ACM Internet Measurement Conference (IMC)*. https://doi.org/10.1145/3419394.3423623
[14] G Akiwate, S Savage, G Voelker, and k claffy. 2021. Risky BIZness: Risks Derived from Registrar Name Management. In *ACM Internet Measurement Conference (IMC)*. https://doi.org/10.1145/3487552.3487816
[15] Antonia Affinito and Raffaele Sommese and Gautam Akiwate and Stefan Savage and Kimberly Claffy and Geoffrey M. Voelker and Alessio Botta and Mattijs Jonker. 2022. Domain Name Lifetimes: Baseline and Threats. In *Traffic and Measurement Analysis*.
[16] Timothy Barron, Najmeh Miramirkhani, and Nick Nikiforakis. 2019. Now You See It, Now You Don't: A Large-scale Analysis of Early Domain Deletions. In *22nd International Symposium on Research in Attacks, Intrusions and Defenses (RAID 2019)*.
[17] Xander Bouwman, Victor Le Pochat, Pawel Foremski, Tom Van Goethem, Carlos H. Ganan, Giovane C. M. Moura, Samaneh Tajalizadehkhoob, Wouter Joosen, and Michel van Eeten. 2022. Helping hands: Measuring the impact of a large threat intelligence sharing community. In *31st USENIX Security Symposium (USENIX Security 22)*.
[18] Calidog. 2024. CertStream. (2024). https://certstream.calidog.io/.
[19] Zhanhao Chen and Jun Javier Wang. 2022. Newly Registered Domains: Malicious Abuse by Bad Actors. (2022). unit42.paloaltonetworks.com/newly-registered-domains-malicious-abuse-by-bad-actors/.
[20] Dave Piscatello and Colin Strutt. 2023. *Cybercrime Supply Chain 2023: Measurements and Assessments of Cyber Attack Resources and Where Criminals Acquire Them.* Technical Report. Interisle. https://interisle.net/s/CybercrimeSupplyChain2023.pdf.
[21] Domain Tools. 2024. Security Information Exchange Newly Observed Domains. (2024). https://www.domaintools.com/resources/user-guides/newly-observed-domains-nod/.
[22] CA/Browser Forum. 2024. Baseline Requirements for the Issuance and Management of Publicly-Trusted TLS Server Certificates - Version 2.0.2. (2024). https://cabforum.org/uploads/CA-Browser-Forum-TLS-BRs-v2.0.2.pdf.
[23] I Foster and R Koga. 2024. DZDB. https://catalog.caida.org/dataset/dzdb. (2024).
[24] Shuang Hao, Matthew Thomas, Vern Paxson, Nick Feamster, Christian Kreibich, Chris Grier, and Scott Hollenbeck. 2013. Understanding the domain registration behavior of spammers. In *Proceedings of the 2013 Conference on Internet Measurement Conference.* https://doi.org/10.1145/2504730.2504753
[25] ICANN. 2007. VeriSign Application for Registry Service: âĂIJRapid Zone UpdatesâĂİ. (2007). https://www.icann.org/en/system/files/files/memo-dns-update-service.pdf.
[26] ICANN. 2023. Registration Data Request Service (RDRS). (2023). https://www.icann.org/en/blogs/details/icann-launches-rdrs-releases-first-metrics-report-17-01-2024-en.
[27] Maciej Korczynski, Maarten Wullink, Samaneh Tajalizadehkhoob, Giovane C. M. Moura, Arman Noroozian, Drew Bagley, and Cristian Hesselman. 2018. Cybercrime After the Sunrise: A Statistical Analysis of DNS Abuse in New gTLDs. In *Proceedings of the 2018 on Asia Conference on Computer and Communications Security.* https://doi.org/10.1145/3196494.3196548
[28] B. Laurie, A. Langley, and E. Kasper. 2013. *Certificate Transparency.* RFC 6962. IETF. http://tools.ietf.org/rfc/rfc6962.txt
[29] Zane Ma, Aaron Faulkenberry, Thomas Papastergiou, Zakir Durumeric, Michael D. Bailey, Angelos D. Keromytis, Fabian Monrose, and Manos Antonakakis. 2023. Stale TLS Certificates: Investigating Precarious Third-Party Access to Valid TLS Keys. In *Proceedings of the 2023 ACM on Internet Measurement Conference.* https://doi.org/10.1145/3618257.3624802

[30] A. Newton and S. Hollenbeck. 2015. *Registration Data Access Protocol (RDAP) Query Format.* RFC 7482. IETF. http://tools.ietf.org/rfc/rfc7482.txt
[31] Adam Oest, Yeganeh Safaei, Penghui Zhang, Brad Wardman, Kevin Tyers, Yan Shoshitaishvili, and Adam Doupé. 2020. PhishTime: Continuous Longitudinal Measurement of the Effectiveness of Anti-phishing Blacklists. In *29th USENIX Security Symposium*.
[32] Adam Oest, Penghui Zhang, Brad Wardman, Eric Nunes, Jakub Burgis, Ali Zand, Kurt Thomas, Adam Doupé, and Gail-Joon Ahn. 2020. Sunrise to Sunset: Analyzing the End-to-end Life Cycle and Effectiveness of Phishing Attacks at Scale. In *29th USENIX Security Symposium)*.
[33] OpenINTEL. 2024. Zonestream. https://zonestream.openintel.nl/. (2024).
[34] Pawel Foremski and Paul Vixie. 2018. The Modality of Mortality in Domain Names. (2018). https://www.virusbulletin.com/uploads/pdf/magazine/2018/VB2018-Vixie.pdf.
[35] Security and Stability Advisory Committee. 2024. *SSAC Report on Registrar Nameserver Management.* Technical Report. ICANN. https://itp.cdn.icann.org/en/files/security-and-stability-advisory-committee-ssac-reports/sac-125-09-05-2024-en.pdf.
[36] Raffaele Sommese, Roland van Rijswijk-Deij, and Mattijs Jonker. 2024. This Is a Local Domain: On Amassing Country-Code Top-Level Domains from Public Data. *ACM SIGCOMM Computer Communication Review* 54, 2 (2024), 2–9.

A ETHICS

We acknowledge that making a public feed of newly registered domain names available may expose personal information (*e.g.,* personal name used as a domain name). However, we argue that our method only consolidates information already publicly available in the CT Logs ecosystem. Therefore, we decided not to remove any domain names from the published stream.

Our ethics review board does not require a specific ethics approval for internet measurements conducted according to community best practices. In measuring the DNS infrastructure of newly registered domains, we followed these best practices by aiming to strike a balance between coverage and slower probing rate. As a result, we decided to scan those domains every 10 minutes, with only three infrastructural query type (A, AAAA, NS), and without enumerating possible sub-labels. Furthermore, we ensured that the address space of our measurement infrastructure had pointers to direct interested users to a page explaining the project's goals and providing a contact for opting out from the measurements.

Regarding RDAP measurements, we queried approximately 76,000 domains per day. We limited our measurements to a maximum of one domain per second to accommodate possible bursts. This rate is generally acceptable and falls well below most registry rate limits (*e.g.,* 7,200 queries per hour for CentralNic) for RDAP. We distributed measurements across four workers with different IPs,

to avoid rare cases of registry RDAP servers enforcing extremely aggressive rate limiting and permanently blocking IPs. To prevent overburdening the infrastructure, we purposefully avoided repeating RDAP measurements in case of failure. We monitored the failure rate to ensure that we did not overload the targeted servers.

B APPENDIX: VERISIGN'S REQUEST TO PROVIDE RAPID ZONE UPDATES

The following is a quotation from Verisign's application to ICANN to provide rapid zone updates as a service [25]:

In September 2004, VeriSign implemented rapid zone updates, enabling updates to the .COM and .NET zone every 3 minutes (prior to this VeriSign propagated updates to the .COM and .NET zones every 12 hours). Although VeriSign updates these zones every 3 minutes, VeriSign publishes that updated data twice a day. This data includes domain names, nameservers, IP address additions, deletions and modifications. The proposed service would enable registrars and others (i.e., anyone who wishes) who currently obtain zone file access in the .COM and .NET TLDs twice daily, to receive updated zone information every five minutes.

....

VeriSign states that the service would be used by recipients to build brand protection and fraud detection services for their customers, and promote security and stability by providing a useful tool to online security companies, ISPs, search engines, financial services companies, and other stakeholders.

- The service would shed light on the activity of those engaged in domain tasting and expose bad actors. "This [service] elevates bad actors into the light of day." "All it does is out people who are probably bad guys."
- The service does not seem susceptible to gaming; "it prevents gaming."
- The service will increase choice for registrars: they can continue to obtain information in twelve-hour increments, purchase the service directly from VeriSign, or purchase the service from value-added providers.
- The service would provide "more granularity" regarding tasting activities but not impact the practice significantly.
- Phishers and others currently alter name servers and conduct fraudulent activity during the time between the twice daily publication of zone file updates. That is, they make nameserver changes to conduct fraudulent activity immediately after one publication and make the change back prior to the next publication – there is a twelve-hour window to conduct this activity.
- Intellectual property owners, brand protection managers and law enforcement would be able to improve their ability to search the .COM and .NET zones for "typo-squatters". "The service would advantage law enforcement to thwart phishing." (One opinion was that the additional twelve-hour extra notice period might not be a significant advantage.)
- After consideration and attempts to, "poke holes in it," there was no apparent ability to use the service for "bad" purposes.

Comments from other sectors of the community have indicated that VeriSign's proposal would assist those who use zone file information provided by VeriSign to address such problems as phishing, fraud and identity theft.

External counsel and ICANN staff agreed that there are no apparent competition issues that require that this matter be forwarded to relevant competition authorities at this time.