WHOIS sunset? A primer in Registration Data Access Protocol (RDAP) performance

Carlos H. Gañán
Internet Corporation for Assigned Names and Numbers (ICANN)
carlos.ganan@icann.org

***Abstract*—The access to registration data regarding domain names and IP addresses was originally intended to enable timely resolution of any technical problems involving domain configuration and operation. WHOIS has been the de-facto communication protocol for more than 35 years to query registration data. Unfortunately, this protocol presents several limitations such as lack of standardized format and inability to authenticate users. To overcome these limitations, a new Registration Data Access Protocol (RDAP) was designed and deployed by registrars and registries. While some of these services have been running for more than 10 years, their performance has not been documented before. In this paper we performed large-scale distributed measurements to evaluate and quantify the response time of these services. We executed more than 10 million RDAP queries against 533 RDAP services. Our results show that 95% of all RDAP queries were responded within 3 seconds; but there were significant differences depending on the RDAP operator. Nevertheless, the median response time for RDAP queries was 3 times slower than WHOIS queries. We further investigate the impact on the response time of several factors such as location, IP address type, response size and HTTPS transaction timings.**

I. Introduction

Internet identifiers –namely IP addresses and domain names– are essential for the functioning of nearly all protocols, e.g., routing protocols such as BGP or application layer protocols such as HTTP. These identifiers are managed by a group of interdependent organizations that assign and allocate identifiers. From the 5 regional Internet registries (RIR) that manage Internet number identifiers to thousands of domain name registries and registrars that manage the reservation of Internet domain names.

To identify who is responsible for a given identifier at any point in time, in the early 1980s the WHOIS protocol [1] was designed to provide access to registration information for Internet identifiers such as IP addresses and domain names. Even though the original design of WHOIS presented several limitations such as lack of standardized response format or inability to authenticate users, it became the main mechanism for external parties to automatically obtain metadata about Internet identifiers. Registration information is useful not only for network administrators to fix system problems but also to combat threats such as spam. Thus, timely and accurate registration data is important and provides powerful information for mitigating potential threats.

In 2015, a new protocol was proposed by the IETF [2] to standardize registration data access while supporting internationalized entries and client authorization. Known as the Registration Data Access Protocol (RDAP), it enables access to current domain names, IP address, and AS number registration data and was envisioned to replace the WHOIS protocol. RDAP was designed to overcome the limitations of the WHOIS protocol while providing security by design. Thus, RDAP operations are structurally different from WHOIS. Contrary to WHOIS which is a text-based protocol that uses a specialized protocol and port [3], RDAP uses a RESTful interface over HTTP with standardized responses specified in JSON [2]. Moreover, at the core of the RDAP design lies a "bootstrapping" service that allows finding the authoritative server of the relevant registration data, avoiding limiting the query to a single specific registry operator or registrar. This is different from the WHOIS protocol, where the information is not linked across the different systems.

Despite all these functionalities and the fact that some RDAP services have been running for more than 10 years, WHOIS is still the most used protocol to access registration data. For instance, in 2020, the RIR for the Asia-Pacific reported an average rate of queries per second 4 times higher for WHOIS than RDAP [4]. Previous research only examined whether RDAP service operators comply with ICANN's policies [5] and deployment issues [6]. However, the performance of both protocols when it comes to retrieve registration data has never been compared before.

This paper presents the first quantitative study on the performance of RDAP services as deployed by registries and ICANN-accredited registrars. We measure the response time of RDAP services by deploying a large-scale measurement infrastructure and executing more than 10 million RDAP domain queries. The major contributions of this paper are:

- We design a measurement methodology to evaluate the performance of the RDAP services provided by RIRs, top-level domain (TLD) registries, and ICANN-accredited registrars;
- We perform large-scale response time measurements from ten vantage points to 533 RDAP services. These represented all publicly known RDAP services at the time of the study;
- We quantify the impact of four factors (location, IP address type, response size and HTTPS transaction timings) on the RDAP response time;
- We quantify response time differences between WHOIS and RDAP queries.

978-3-903176-40-9

II. Related work

Registration data is used for a wide variety of purposes not only by researchers but also by commercial entities (e.g., registrars), law enforcement authorities and the general public. For instance, it is used to identify the operator of a website or the owner of an IP address. Researchers have used registration data to gain insights on cybercriminal activities such as bulletproof hosting [7]–[9] and botnet operations [10]–[12]. Law enforcement agencies regularly depend on registration data to identify criminals [13], [14]. Trade-maker owners also use registration data to avoid possible conflict and infringements [15].

Unfortunately, registration data is also prone to misuse [16]. For instance, domain speculators (aka 'domainers') scrape registration data to acquired lucrative domains and then monetize their portfolios (e.g., via domain tasting or kiting) while they waited for offers to purchase their domain names [17]. Criminals also scrape contact details of registrants to then spam them via email, phone and/or postal letter [16].

Multiple studies have documented issues with registration data such as inaccuracy [18] or rate limiting [16]. Nevertheless, even though WHOIS has been running for more than 35 years, to the best of our knowledge, no study has documented its performance in terms of response time. In the rest of this section, we discuss the evolution of different protocols to access registration data.

A. Protocols to access registration data

WHOIS is the initial and still prevailing protocol to access registration data. It was specified in 1982 (see RFC-812 [1]); but had from several major modifications along the way. In 1985, RFC-945 [19] required the use of TCP (dropping NCP) and introduced the concept of a registrar. In 1994, referral WHOIS (aka rWHOIS) [20] was defined to take into account that registration data was no longer stored in a single database and it was necessary to perform a series of lookups to reach the final maintainer. Again, one year later, another extension (WHOIS++ [21]) was specified aiming at increasing the scalability of the service by structuring registration data. Unfortunately, due to the lack of extensible markup languages, the deployment of WHOIS++ did not succeed. The final update of WHOIS protocol was in 2004, with RFC-3912 [3], reaffirming the technical specification of WHOIS as TCP Port 43 transaction-based query-response service that did not require additional client software.

In parallel to WHOIS extensions, other protocols were proposed. The Shared WHOIS Project (SWIP) [22] aimed at ensuring the various WHOIS databases were consistent and complete, since records are located in many places. The introduction of internationalized domain names (IDNs) also showed that WHOIS did not allow querying consistently for domain names that were not in ASCII characters. To handle IDNs and other concerns, the Cross Registry Internet Service Protocol (CRISP) working group [23] designed the Internet Registry Information Service (IRIS) protocol [24].

Early 2010s, the IETF stated that IRIS was not a successful replacement of WHOIS mainly due to its complexity. With two RIRs already serving registration data via RESTful web services, the IETF created a working group to develop the Web Extensible Internet Registration Data Service (WEIRDS). WEIRDS was eventually renamed to the Registration Data Access Protocol (RDAP) [2].

B. Registration Data Access Protocol (RDAP)

RDAP stands in many respects as an enhanced version of WHOIS. RDAP was designed to overcome the weaknesses of the previous protocol and has focused mainly on the security, structuring and internationalization aspects when developing the new consultation protocol. As a replacement to WHOIS it stands out for the following novelties: (i) structured question and answer semantics, including standardized error messages; (ii) secure access to requested contact details (for example, via HTTPS); (iii) expandability; (iv) bootstrapping mechanism to find the appropriate authoritative DNS server; (v) standardized transmission of requests; (vi) web-based and in compliance with REST; and (vii) simple translation of output data and possibility of providing differentiated access to contact data. Six RFC documents [2], [25]–[29] cover the specification of all the elements that the RDAP protocol consists of.

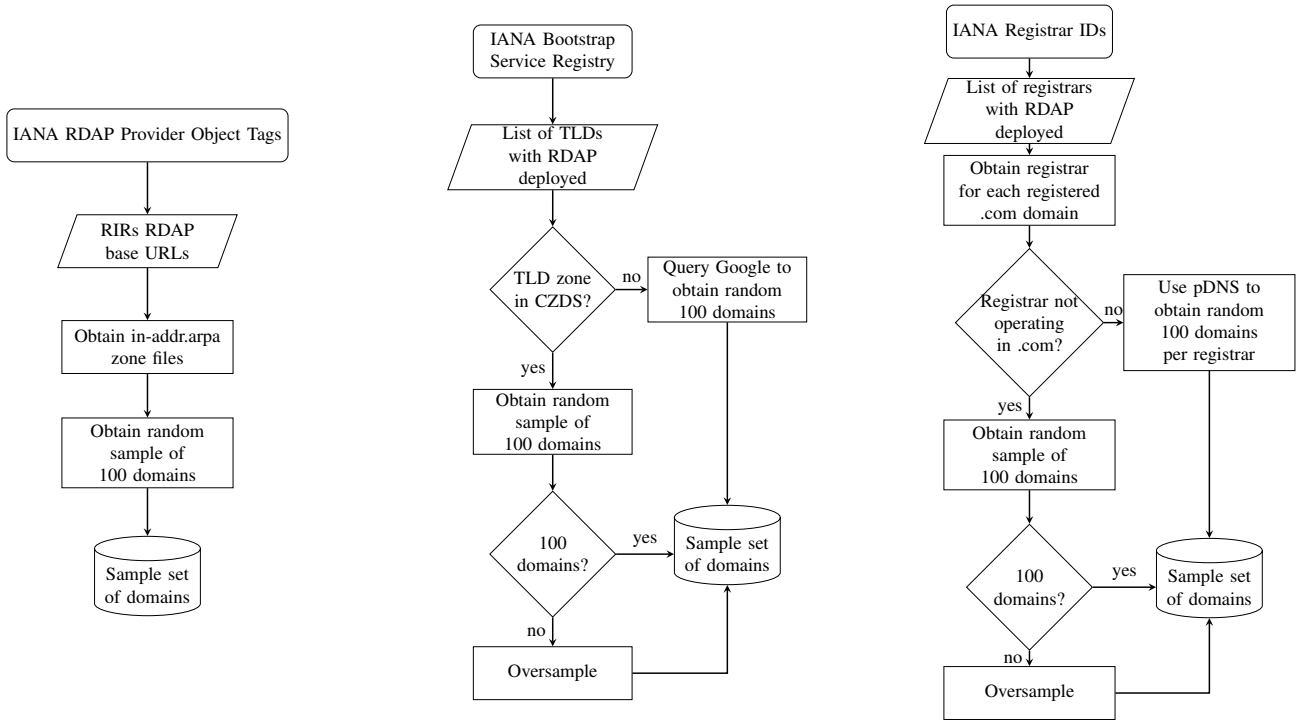
III. Methodology

We deployed ten vantage points distributed across the globe to estimate different metrics related to the performance of RDAP services. All measurements were conducted from outside the network where the RDAP services were running. Thus, metrics related to the internal usage of resources such as CPU or memory usage were not quantified. The method consisted of two steps: (i) creating a random sample of domain queries, and (ii) executing the queries while recording the response and response time.

A. Creating sample of queries

To test the response time of each RDAP service, it is necessary to create a sample list of queries. While some RDAP services allow to query for different types of resources (mainly IP address, autonomous system numbers, and/or domains), this study focused on domain queries as this is the object type that is common to all RDAP services including those provided by RIRs, registries, and registrars. To capture potential differences among the responses of the same RDAP service, 100 different queries were created. This threshold was chosen heuristically to make the measurement process viable in machines with limited resources.

We followed three processes to create RDAP queries depending on the type of RDAP operator, i.e., RIRs, TLD registries, and ICANN-accredited registrars. The flowcharts in Figure 1 represent the different steps that were carried out to obtain a random set of domain names. These domain names were, in turn, used to create sample RDAP queries. We created a new set of queries every day to account for potential domains being unregistered/expired during the measurement period.



(a) RIRs (b) TLD Registries (c) ICANN-accredited Registrars

Fig. 1: Process to obtain a random sample of domains to create RDAP queries

1) RDAP queries for RIRs: RIRs RDAP base URLs are listed by IANA as part of the method that can be used to identify the authoritative server for processing queries [30]. With the base URL and domain names included in the `in-addr.arpa` zone files, we created a random set of 100 RDAP queries per RIR. In particular, we carried out the following steps (see Figure 1a):

(i) Obtain the list of regional RIR RDAP base URLs from IANA's bootstrap list[1].
(ii) For each RIR, randomly select 100 domains from the `in-addr.arpa` zone files.
(iii) Create domain queries by appending the sample set of domains to each RDAP URL from step (i) following RFC 7482 [26].

2) RDAP queries for TLD registries: As defined in RFC 7484 [28], RDAP requires of a bootstrap service registry for domain names. This bootstrap service includes the RDAP base URLs of the services ran by TLD registries. With this information and leveraging the zone files publicly accessible through ICANN's Centralized Zone Data Service (CZDS), we constructed 100 random sample queries for each one of the registries with an RDAP service. To create this sample list of domains we perform the following steps (see Figure 1b):

(i) Obtain the TLDs for which IANA has an RDAP base URL listed in the bootstrap service[1].
(ii) For each TLD:
- If the zone file is available, get a random set of 100 domains (if the zones do not contain 100 domains, then oversample);
- If the zone file is not available, get a random set of 100 unique domains by extracting second level domains from the results of a web search (e.g., Google query: "* site:*TLD*"). If the results of the search do not contain 100 different domains, then oversample.

(iii) Create domain queries by appending the sample set of domains to each RDAP URL as stated in RFC-7482 [26].

3) RDAP queries for ICANN-accredited registrars: To create a sample set of queries for the registrars, first we needed to identify which registrars are providing an RDAP service. We used IANA's list of registrars[2] which specifies the RDAP Base URL for each registrar. To create a sample of domains per registrar, we leveraged `.com` zone files as provided by the registry. In particular we carried out the following steps to create RDAP sample queries for registrars (see Figure 1c):

(i) Obtain the list of base URLs for the RDAP servers operated by registrars from IANA's registrar identifier list.
(ii) For each domain in the `.com` zone file, obtain the corresponding registrar via `WHOIS`.
(iii) Randomly select 100 domains for each registrar that had an RDAP server mapped in step (i) by using the registrar information of step (ii):
- If there are not 100 `.com` domains, then oversample;

[1] https://data.iana.org/rdap/

[2] https://www.iana.org/assignments/registrar-ids/registrar-ids.xhtml

- If no `.com` domains for a particular registrar are found, then use Spamhaus pDNS API [31] to retrieve domains for the missing registrar.

(iv) Create domain queries by appending the sample set of domains to each RDAP URL from step (i) following RFC 7482 [26].

B. Executing RDAP queries

To measure the response time of the different RDAP services, ten different virtual machines (VM) were provisioned to conduct the measurements. Each virtual machine was located in a different autonomous system spread across 6 different continents to account for potential measurement biases due to routing. Figure 2 shows the location of these VMs.

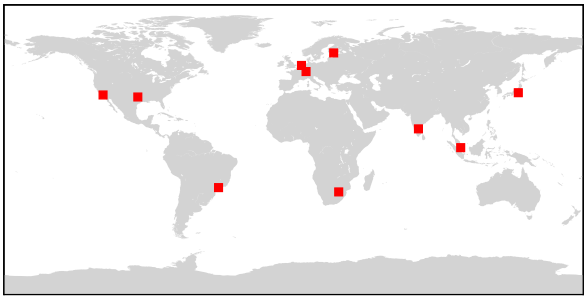


Fig. 2: Geolocation of the measurement VMs

The queries ran periodically (i.e., every 5 minutes) and were executed at the same time from the different nodes. Each RDAP service was only queried once every 5 minutes by each node. The RDAP services were grouped based on the IP address and origin AS of their domains. This allowed us to group RDAP services that share the same infrastructure.

The measurements aimed at quantifying response time. Response time is one of the most important metrics to track the performance of a REST API. For the purpose of this study, response time is defined as the time elapsed since an RDAP query is executed until the response is received minus the time to resolve the domain name. TLS1.2 was forced in all queries.

Furthermore, the response time is broken down into 5 different times related to HTTPS transactions:

1) *Connect time*: the time it took from the start until the connect to the remote host was completed. This measures the TCP three-way handshake from the client's perspective. It ends just after the client sends the `ACK`, i.e., it does not include the time taken for that `ACK` to reach the server. It should be close to the round-trip time (RTT) to the server.
2) *Appconnect time*: the time it took from the end of the connection time until the TLS handshake was completed. This captures the TLS setup, in the case of TLS 1.2 around two RTTs. The client is then ready to send the RDAP `HTTP GET` request.
3) *Pretransfer time*: the time it took from the end of the TLS handshake until the response transfer is just about to begin (0 RTTs).
4) *Start-transfer time*: the time it took from the end of the pretransfer until the first byte is received. This captures the response generation on the remote RDAP server.
5) *Transfer time*: the time it took from the first byte until the last byte is received, i.e., time until the client has sent the `FIN` connection tear down.

Note that some services redirect the queries to another RDAP URL. In these cases, the "response time" metric will also account for these redirections by calculating the total time from query to response including any redirection.

IV. RDAP RESPONSE TIME

After measuring RDAP response time from February 23, 2021 until March 09, 2021, a total of 10,598,236 valid RDAP responses were collected. We observed that, on average, RIRs provide the fastest RDAP service followed by gTLDs. On the other hand, the response time of the RDAP services provided by registrars is the highest. Notwithstanding these differences, 95% of all queries were answered in less than 4 seconds. Table I evidences the presence of outliers in the measurements. In the case of the RDAP response time provided by registrars, the standard deviation is twice higher than the mean value. Looking at the maximum values it is also worth noting that some queries took more than 5 minutes to get a response.

	Response time (sec)						
	mean	std	min	50%	95%	99%	max
RIR	0.50	0.74	0.00	0.27	1.46	2.17	90.72
Registry	0.88	1.52	0.04	0.75	1.99	3.07	373.15
Registrar	1.22	2.93	0.01	0.83	3.06	9.19	300.84

TABLE I: Summary statistics: response time

Figure 3 shows the cumulative density function (CDF) of the response time for all the queries per RDAP operator type. This figure corroborates the existence of extreme outliers in the case of RDAP services provided by registrars and registries which in a handful of cases took several minutes to receive a response. As can be seen in the zoomed inset in this figure, 88% of all the queries were answered within a second.

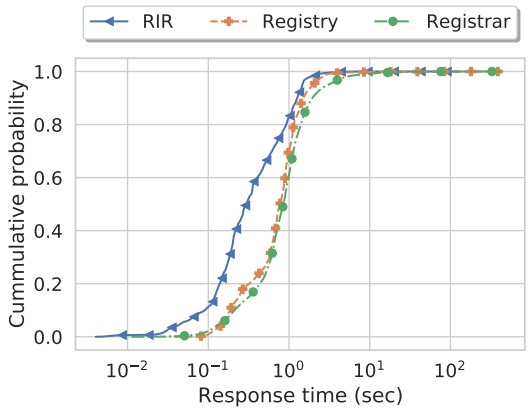


Fig. 3: CDF of the RDAP response time (sec) for all the queries per operator type

A. RIRs' RDAP performance

Over the measurement period, we executed more than a 250,000 RDAP queries against the RIRs' RDAP service. RIRs have been providing RDAP services for longer than registrars or registries. In fact, ARIN deployed an RDAP pilot in October 2009 that was fully operation in July 2010. This might explain why their RDAP response time was the lowest on average. Nevertheless, not all the RIRs have deployed the same type of infrastructure to host their RDAP services.

Table II shows the main descriptive statistics of the response time for the RIRs' RDAP service. LACNIC's RDAP service has the slowest response time on average and APNIC's RDAP service is the fastest. In fact, we could divide the RIRs into 2 groups according to their RDAP performance. The RDAP services from AFRINIC, ARIN and LACNIC responded more than twice slower than the RDAP services provided by APNIC and RIPE. It is also worth noting that there were some outliers in the measurements. For instance, some queries took more than a minute to return the corresponding RDAP response. Again, these were extreme outliers as 99% of all queries finished in less than 5 seconds.

RIR	mean	std	min	50%	95%	99%	max
AFRINIC	0.66	0.99	0.02	0.35	1.49	1.72	90.72
APNIC	0.24	0.59	0.02	0.18	0.55	0.77	76.54
ARIN	0.62	0.64	0.03	0.43	1.68	2.37	64.87
RIPE	0.28	0.33	0.00	0.14	0.96	1.04	10.33
LACNIC	1.85	1.62	0.08	1.61	3.36	4.82	16.74

TABLE II: Response time(sec) statistics per RIR

We again compare the performance of the individual RDAP service operators. Figure 4 shows the cumulative distribution of the average response time for each RDAP service. On average RIRs' RDAP service was faster than 90% of the services provided by registries and registrars, i.e., only a handful of registries and registrars provided RDAP responses as fast as RIRs.

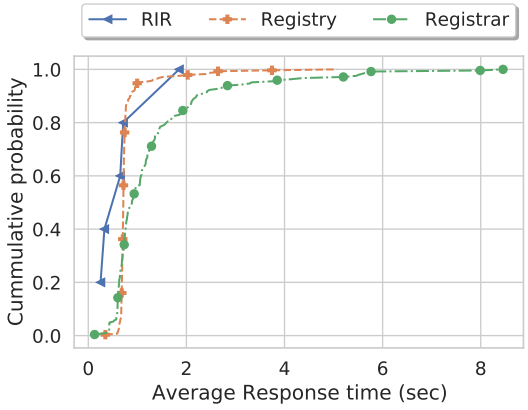


Fig. 4: CDF of the average RDAP response time(sec) for each RDAP service operator.

B. Registries' RDAP performance

We queried a total of 287 RDAP services operated by TLD registries which resulted in more than 2 million RDAP responses. We observed significant differences between the different services. In particular, the response time was higher for RDAP services providing country-level registration data (ccTLDs) than services providing gTLD registration data. As can be seen from Table III, the average response time for ccTLDs was 1.51 seconds while for gTLDs it did not reach one second. The response times for RDAP services related to sponsored TLDs (sTLD) and restricted generic TLDs (grTLD) did not differ significantly from those related to gTLDs.

Type	mean	std	min	50%	95%	99%	max
ccTLD	1.51	2.56	0.10	1.17	2.94	7.33	331.79
gTLD	0.76	1.00	0.04	0.71	1.58	2.38	373.15
grTLD	0.68	0.34	0.16	0.70	1.19	1.22	8.18
sTLD	0.67	0.70	0.09	0.66	1.32	1.91	65.70

TABLE III: Response time (sec) statistics per registry type

When comparing the average response time among the different operators, we observe that 98% of the different RDAP services provided by registries responded within 2 seconds, and only 6 registry's RDAP services took more than 2 seconds in average to respond (see Figure 4).

C. Registrars' RDAP performance

Even though there are 4,136 registrars offering RDAP services, only 242 of these services were unique, i.e., different RDAP base URLs. Querying these 242 every 5 minutes over the measurement period resulted on more than 8 million valid RDAP responses. The RDAP service performance of these 242 different registrars was diverse and presented significant differences. As can be seen in Figure 4, 80% of the registrar's RDAP services replied within 2 seconds. However, 34 registrars have deployed RDAP services that took between 2 and 8 seconds on average to respond.

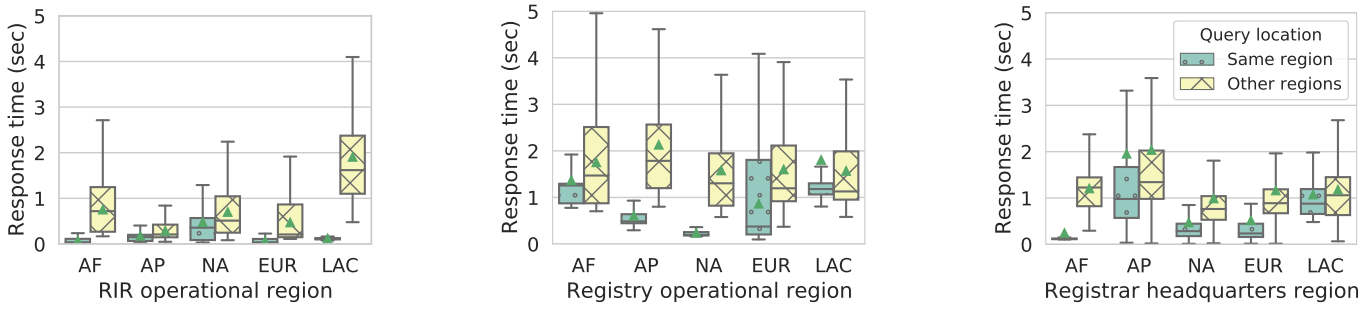
We further classified the registrars depending on whether they belong to a family of registrars (i.e., they operate under multiple IANA registrar identifiers) or they operate individually. 158 of the registrars providing an RDAP service belonged to a family, while 84 operate individually. Table IV shows that, on average, registrar's belonging to a family provided RDAP responses a few deciseconds slower than individual registrars.

Registrar	mean	std	min	50%	95%	99%	max
Family	1.24	3.06	0.01	0.85	3.21	8.56	298.66
Individual	1.18	2.56	0.01	0.78	2.75	10.33	300.84

TABLE IV: Registrar's response time(sec) statistics

V. Factors influencing RDAP query response time

As with any other REST API service, RDAP service performance can be influenced by different factors. In this section, we analyze the impact of a series of factors, namely source location of the query, RDAP response size, type of IP address



(a) RIRs (b) ccTLD registries (c) ICANN-accredited Registrars

Fig. 5: Response time distribution depending on the location of query source and the operational region of the RDAP service operator. (AF = Africa; AP = Asia/Australia/Pacific; EUR = Europe; LAC = Latin America/ Caribbean; NA = North America)

over which the query was performed, and HTTPS transaction timings.

A. Source query location

We measure the response time for each query executed from the 10 vantage points (VP) deployed across the globe. Significant differences were observed depending on the location from which the queries were run. Table V shows the main descriptive statistics for the response time per VP. The response time was significantly higher for the VPs located in Asia and South America.

Country	Region	Response time (sec)						
		mean	std	min	50%	95%	99%	max
BR	LAC	1.30	2.40	0.02	0.97	2.69	6.31	298.49
FI	EUR	0.90	1.90	0.03	0.51	2.40	6.53	298.66
FR	EUR	0.88	3.23	0.01	0.43	2.36	7.07	331.79
IN	AP	1.31	2.88	0.03	0.90	2.71	8.54	288.28
JP	AP	1.30	2.35	0.02	1.07	2.78	6.15	281.43
NL	EUR	0.85	2.56	0.00	0.42	2.43	7.02	373.15
ZA	AF	1.38	2.02	0.02	0.99	2.91	7.77	293.85
SG	AP	1.39	3.35	0.01	0.98	2.75	10.64	325.12
US-WC	NA	0.97	2.53	0.01	0.69	2.25	5.56	300.84
US-EC	NA	0.87	1.84	0.03	0.59	2.23	5.51	276.25

TABLE V: Response time statistics per vantage point

To further understand why some VPs experienced higher response time than others, we grouped the measurements according to the location of the source of the query and the operational region of the RDAP service operator. For instance, for the European RIR, we consider the queries launched from the VPs in Netherlands, France and Finland to be from the "same region", and any other query from the remaining seven VPs to be from "other regions". Note that for gTLD registries their operational region is not restricted to a particular territory, thus we exclude them from this analysis.

Figure 5 shows the boxplots for the response time of each operator depending on these regions. Just by comparing the size of the different boxes, a clear pattern can be observed independently of the RDAP operator, i.e., queries executed from the same region as the RDAP operator are responded faster on average that queries executed from a different region. Even though this pattern is present in all regions, in some areas this difference can lead to experience two or three times higher response times. For instance, operators in Africa can provide fives times faster responses for queries that are executed from Africa than when is executed from another region. However, there is one exception to this pattern. APNIC's RDAP responded almost at the same speed independently of the location where the query originated. This is mainly due to the fact that this RIR uses Google Cloud Platform to provide their RDAP service.

B. RDAP response size

We also calculated the size in bytes of each RDAP-level JSON response. The average response size of all the performed queries was of around 6 kB with a standard deviation of 4.4 kB. This evidences a great diversity of response sizes across the different RDAP services. Table VI breaks down the response size statistics per RIR, TLD registries, and registrars. As can be seen, registrars' RDAP services provided the shortest responses on average while some registrars provided a response of more than 122 kB.

	Response size (bytes)					
	mean	std	min	50%	95%	max
RIR	6,334	5,686	1,118	3,038	17,646	28,351
Registry	7,893	3,228	1,930	8,893	11,928	61,202
Registrar	4,446	2,632	1,285	4,167	8,352	122,536

TABLE VI: RDAP response size statistics

Finally, we investigated a potential relationship between the average response time and average response size. Figure 6 plots these variables against each for each one of each on. While the responses from ARIN's RDAP service were larger on average, the average response time was still faster than the RDAP service of AFRINIC. Similarly, the registries' and registrars' RDAP response size also vary significantly per operator, being sometimes two order of magnitude different. A Pearson product-moment correlation coefficient was computed to assess the relationship between the response size and response time. No strong correlation was found ($\rho = 0.00001$, $p - value < 0.00005$).

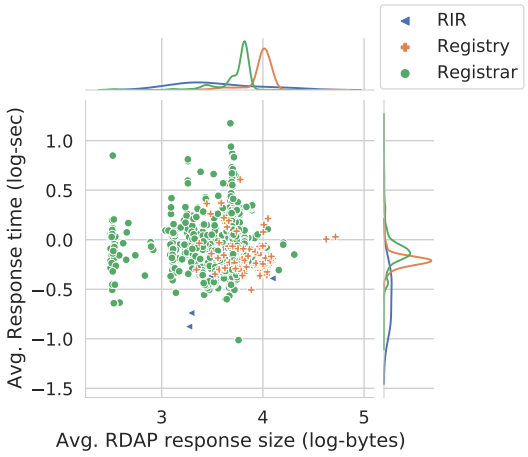


Fig. 6: Avg. response time vs RDAP response size per operator

C. IP address type

A subset of the vantage points nodes allowed to perform RDAP queries using IPv6 source addresses. From these, the same queries were performed using IPv4 and IPv6 source addresses. Figure 7 shows that queries performed over IPv6 were slightly faster, but no statistically significant differences are found between the queries that used IPv4 compared to IPv6. The median response time for queries over IPv4 was 0.87 seconds while the same queries executed over IPv6 were responded slightly faster with a median response time of 0.67 seconds. However, IPv4 queries suffered from more variance (std=1.02 sec) than IPv6 queries (std=0.68 sec).

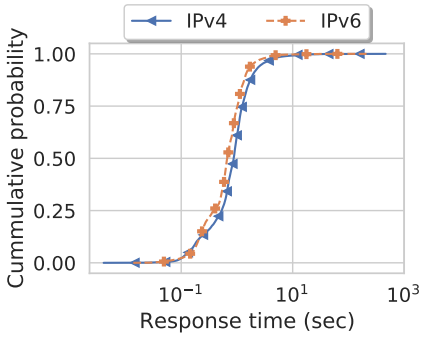


Fig. 7: CDF response time per IP address type

D. HTTPS transaction times

All the queries in this study were performed using HTTPS, thus the total response time is also affected by the TLS handshake. To understand to what extent the HTTPS protocol affects RDAP performance, the total response time is broken down into the 5 metrics as defined in section III.

Table VII shows the mean value of the different timings. Between 320 to 430 milliseconds were spent on average per RDAP query on the TLS handshake, and between 226 to 500 milliseconds were spent on transferring the response data.

	Average HTTPs transaction times (msec)				
	connect	appconnect	pretransfer	starttransfer	transfer
RIR	69.54	181.98	0.09	188.12	57.48
Registry	162.82	375.10	0.74	301.92	35.39
Registrar	211.46	395.66	0.08	605.34	8.41

TABLE VII: HTTPS transaction timings

As can be seen in Figure 8, on average the TLS handshake takes around 40% of the total response time. Most of the time (around 60%) is spent in starting the connection and transferring the RDAP response only accounts for less than 40% of the total response time on average. However, these proportions differ between RDAP operators. For instance, queries against RIRs' RDAP services spent more time transferring the response than connecting to the server which is partly due to the considerably large size of the RDAP responses provided by the ARIN.

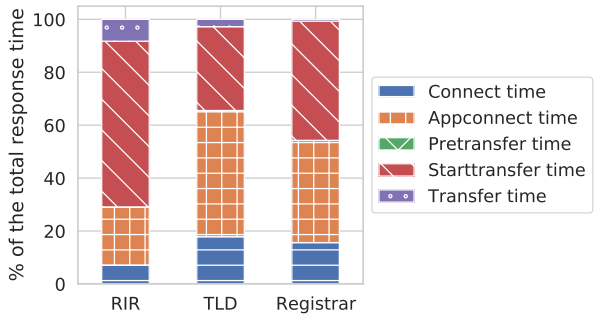


Fig. 8: Breakdown of RDAP response time into HTTPS transaction timings

VI. WHOIS vs RDAP response time

While the creation of RDAP was not meant to make access to registration data faster, due to the number of round-trip times involved in an RDAP query it is expected to be slower than WHOIS. In this section, the latency of WHOIS vs RDAP is compared.

For the sample population of domains that was used in the previous section, we executed WHOIS queries over port 43 from the same 10 VPs in parallel to the RDAP queries. We measured the response time of these queries.

	Response time (sec)						
	mean	std	min	50%	95%	99%	max
RDAP	1.37	2.78	0.00	0.97	3.38	9.58	373.17
WHOIS	1.19	5.87	0.00	0.28	2.17	33.61	846.31

TABLE VIII: Response time statistics: RDAP vs WHOIS

Table VIII shows the main statistics of the response time for both WHOIS and RDAP queries. There are 2 salient patterns: (i) WHOIS response time was on median 3 times faster than RDAP; and (ii) WHOIS response time had larger variance with a larger presence of outliers than RDAP queries. Figure 9 shows the distribution of the response time for all the

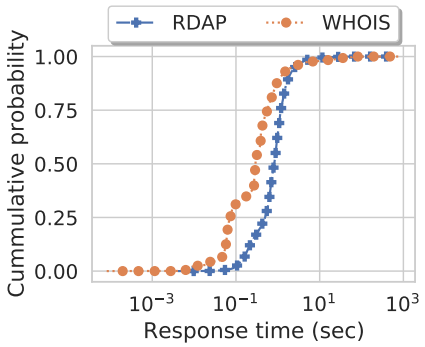


Fig. 9: CDF of the response time for WHOIS vs RDAP queries

queries performed with RDAP and for the same type of query performed via WHOIS. As can be seen, 50% of the WHOIS queries were answered within the first 300 milliseconds, while only 15% of the total number of RDAP queries were responded within the same time frame.

VII. Conclusions

The current deployment of RDAP services is diverse and, while 95% of all performed domain queries were answered within 3 seconds, some queries can take up to several minutes to be answered. Registrars' RDAP services are currently the slowest on average while RIRs' RDAP services are the fastest.

Several factors influenced the response time. The source location of the query had a significant impact on the response time. On average, queries originating from Europe and North-America received faster responses than those from Asia or Africa. Queries executed over IPv6 had slightly lower response times than those over IPv4. Response size did not seem to have a significant impact on the response time apart from the RDAP service operated by ARIN whose responses were one order of magnitude larger than the other RIRs.

When comparing RDAP to `WHOIS`, our results show that the median RDAP response time is 3 times larger with RDAP than with `WHOIS`. This can be partly attributed to the use of HTTPS as on average 40% of the RDAP response time is spent on establishing a secure connection.

References

[1] K. Harrenstien and V. White, "NICNAME/WHOIS," RFC 812, RFC Editor, Fremont, CA, USA, Mar. 1982, obsoleted by RFCs 954, 3912.

[2] A. Newton, B. Ellacott, and N. Kong, "HTTP Usage in the Registration Data Access Protocol (RDAP)," RFC 7480 (Proposed Standard), RFC Editor, Fremont, CA, USA, Mar. 2015.

[3] L. Daigle, "WHOIS Protocol Specification," RFC 3912 (Draft Standard), RFC Editor, Fremont, CA, USA, Sep. 2004.

[4] Asia Pacific Network Information Centre, "APNIC 2020 annual report," 2020. [Online]. Available: https://www.apnic.net/wp-content/uploads/2021/03/APNIC-2020-Annual-Report.pdf

[5] G. Aaron, "Domain Name Registration Data at the Crossroads," Interisle Consulting Group, LLC, Tech. Rep., 2020.

[6] M. Blanchet, "RDAP Deployment Findings and Update," RFC Editor, Fremont, CA, USA, Dec. 2019.

[7] A. Noroozian, J. Koenders, E. Van Veldhuizen, C. H. Ganan, S. Alrwais, D. McCoy, and M. Van Eeten, "Platforms in everything: analyzing ground-truth data on the anatomy and economics of bullet-proof hosting," in *28th USENIX Security Symposium*, 2019, pp. 1341–1356.

[8] S. Tajalizadehkhoob, R. Böhme, C. Gañán, M. Korczyński, and M. Eeten, "Rotten apples or bad harvest? what we are measuring when we are measuring abuse," *ACM Transactions on Internet Technology (TOIT)*, vol. 18, no. 4, pp. 1–25, 2018.

[9] S. Tajalizadehkhoob, C. Gañán, A. Noroozian, and M. v. Eeten, "The role of hosting providers in fighting C2C infrastructure of malware," in *Proceedings of the 2017 ACM on AsiaCCS*, 2017, pp. 575–586.

[10] S. Tajalizadehkhoob, M. Korczyński, A. Noroozian, C. Ganán, and M. van Eeten, "Apples, oranges and hosting providers: Heterogeneity and security in the hosting market," in *IEEE/IFIP Network Operations and Management Symposium*. IEEE, 2016, pp. 289–297.

[11] M. H. Jhaveri, O. Cetin, C. Gañán, T. Moore, and M. V. Eeten, "Abuse reporting and the fight against cybercrime," *ACM Computing Surveys (CSUR)*, vol. 49, no. 4, pp. 1–27, 2017.

[12] O. Çetin, C. Ganán, L. Altena, T. Kasama, D. Inoue, K. Tamiya, Y. Tie, K. Yoshioka, and M. van Eeten, "Cleaning up the internet of evil things: Real-world evidence on isp and consumer efforts to remove mirai." in *NDSS*, 2019.

[13] M. Edwards, A. Rashid, and P. Rayson, "A systematic survey of online data mining technology intended for law enforcement," *ACM Computing Surveys (CSUR)*, vol. 48, no. 1, pp. 1–54, 2015.

[14] C. Gañán, U. Akyazi, and E. Tsvetkova, "Beneath the radar: Exploring the economics of business fraud via underground markets," in *APWG Symposium on Electronic Crime Research (eCrime)*, 2020.

[15] J. D. Lipton, "Beyond Cybersquatting: Taking Domain Name Disputes Past Trademark Policy," *Wake Forest L. Rev.*, vol. 40, p. 1361, 2005.

[16] N. Leontiadis and N. Christin, "Empirically measuring WHOIS misuse," in *Symposium on Computer Security*. Springer, 2014, pp. 19–36.

[17] M. Milam, "The rise of brandjacking against major brands," *Computer Fraud & Security*, vol. 2008, no. 10, pp. 10–13, 2008.

[18] United States Congress House, Committee on the Judiciary, *The US Government's Role in Ensuring Public Access to Accurate WHOIS Data*. US Government Printing Office, 2003.

[19] J. Postel, "DoD statement on the NRC report," RFC 945, RFC Editor, Fremont, CA, USA, May 1985, obsoleted by RFC 1039.

[20] S. Williamson and M. Kosters, "Referral Whois Protocol (RWhois)," RFC 1714 (Informational), RFC Editor, Fremont, CA, USA, Nov. 1994, obsoleted by RFC 2167.

[21] J. Gargano and K. Weiss, "Whois and Network Information Lookup Service, Whois++," RFC 1834 (Informational), RFC Editor, Fremont, CA, USA, Aug. 1995.

[22] C. Weider and R. Wright, "A Survey of Advanced Usages of X.500," RFC 1491 (Informational), RFC Editor, Fremont, CA, USA, Jul. 1993.

[23] J. White, "High-level framework for network-based resource sharing," RFC 707, RFC Editor, Fremont, CA, USA, Dec. 1975.

[24] A. Newton and M. Sanz, "IRIS: A Domain Registry (dreg) Type for the Internet Registry Information Service (IRIS)," RFC 3982 (Proposed Standard), RFC Editor, Fremont, CA, USA, Jan. 2005.

[25] S. Hollenbeck and N. Kong, "Security Services for the Registration Data Access Protocol (RDAP)," RFC 7481 (Proposed Standard), RFC Editor, Fremont, CA, USA, Mar. 2015.

[26] A. Newton and S. Hollenbeck, "Registration Data Access Protocol (RDAP) Query Format," RFC 7482 (Proposed Standard), RFC Editor, Fremont, CA, USA, Mar. 2015.

[27] ——, "JSON Responses for the Registration Data Access Protocol (RDAP)," RFC 7483 (Proposed Standard), RFC Editor, Fremont, CA, USA, Mar. 2015.

[28] M. Blanchet, "Finding the Authoritative Registration Data (RDAP) Service," RFC 7484 (Proposed Standard), RFC Editor, Fremont, CA, USA, Mar. 2015, updated by RFC 8521.

[29] L. Zhou, N. Kong, S. Shen, S. Sheng, and A. Servin, "Inventory and Analysis of WHOIS Registration Objects," RFC 7485 (Informational), RFC Editor, Fremont, CA, USA, Mar. 2015.

[30] S. Hollenbeck and A. Newton, "Registration Data Access Protocol (RDAP) Object Tagging," RFC 8521 (Best Current Practice), RFC Editor, Fremont, CA, USA, Nov. 2018.

[31] Spamhaus Technology, "Spamhaus' Passive DNS," 2021. [Online]. Available: https://pdns.spamhaustech.com/