Master Research Project

Measuring the deployment of DNSSEC over the Internet

Nicolas Canceill <Nicolas.Canceill@os3.nl>

Master *System and Network Engineering*
University of Amsterdam, The Netherlands

July 7, 2014

Abstract

This document is a Research Project report for the Master education in *System and Network Engineering* at the University of Amsterdam, The Netherlands. The study is mainly focused on measuring the deployment of DNSSEC over the Internet using the Atlas network.

Despite being a core technology of today's Internet, DNS is a notoriously insecure protocol. After a long development period, DNSSEC was finally deployed at the DNS root level, and the majority of the top-level domains are now signed. Nonetheless, DNS resolvers do not often support DNSSEC, let alone validation, and many nameservers still serve unsigned zones.

In order to properly evaluate the benefits of DNSSEC, which highly depend on its deployment among all DNS resolvers, the Internet community needs up to date information from all over the Internet. During this project, close to five thousand probes from the Atlas network were used to conduct measurements.

Measurements were run with the Atlas network, using a nameserver at NLnet Labs. The results provide new insight on the distribution of DNSSEC support among resolvers, and notably show that around 90% of resolvers are DNSSEC-aware, about 30% validate answers, and 65% provide the client with the cryptographic signatures when available. Moreover, some particular cases are outlined: the existence of insecure fallbacks in case of missing signatures, and a small issue with signatures of secure wildcard records.

Instructor Prof. C. de Laat
University of Amsterdam, The Netherlands

Supervisors B. Overeinder, W. Toorop
NLnet Labs, Amsterdam, The Netherlands

Acknowledgements

This work was conducted in Amsterdam, The Netherlands, with the assistance of NLnet Labs, and I am particularly grateful to Benno Overeinder and Willem Toorop for the welcome they showed me. They nicely spent a long time answering my many questions, helping me fix my many mistakes, and sharing with me a drop of their incredible DNS knowledge.

This project is part of the programme managed by Cees de Laat for the Master education in *System and Network Engineering*, and I am deeply thankful for his confidence and support. More generally, I would like to thank the entire staff of the Master, for guiding me through the final part of my engineering education.

Finally, I would like to thank my family, for supporting me all along my studies. This work would not have been possible without the love and care I received from my relatives and friends.

Nicolas Canceill

Contents

1 Introduction

We don't know.
Only the Machines know,
and they are going there and taking us with them.
Isaac Asimov (in *I, Robot*, 1950)

The Domain Name System, known as DNS, is one of the essential foundations of the Internet as we know it. Its main purpose is to translate domain names into IP addresses, so users can easily browse the Web, e-mails can be routed, and many other functionalities can be provided. For instance, someone wanting to access a webpage at `www.example.com` will have to resolve that name to an IP address in order to reach the corresponding webserver.

There are three classes of actors in DNS: the clients, the resolvers, and the nameservers. When a client needs DNS resource information, it queries a resolver, which will take care of gathering the information, by asking the one or many nameservers holding the requested resources. Nameservers are managed by domain owners, while resolvers are usually supplied to clients by their Internet Service Provider.

Unfortunately, DNS presents a major security vulnerability: information can be subject to forgery. Since the system is distributed, DNS packets often travel through the Internet, and anyone can very easily tamper with them. Middle boxes may modify records in transit, and non-authoritative nameservers may inject poisonous information. In the aforementioned example, this could allow an attacker to point the user to a different IP address, and potentially serve a malicious webpage.

The system was thus extended by DNSSEC in order to offer a secure naming system. It uses asymmetric cryptography to sign all records in the answers, and also to authenticate every delegation of authority. This takes advantage of the hierarchical nature of DNS in order to build a chain of trust from the root of the DNS tree.

The DNS root zone is now signed, but DNSSEC is still not widely deployed over the Internet. Even when a zone is properly authenticated and resources are correctly signed, resolvers must validate the answers for the client to receive trusted data. Otherwise, all benefits of DNSSEC are lost; one alternative would be to perform validation directly on the client, but this still requires the resolver to include the signatures in the answer.

1.1 History

The Domain Name System (DNS) was designed in 1983 [1] by Paul Mockapetris. Before that, the domain namespace was not structured, and domain names were merely host names [2]. At that time, DNS was a big step forward, so it mainly focused on functionality and was developed with little consideration for security.

DNS introduced a hierarchical tree structure for domain names, and a format for DNS resource information. It also defined the concept of authoritative zones, the roles of nameservers and resolvers, and the protocol allowing to query and serve the resource information. A year later, the first top-level domains were defined [3]. In 1987, DNS finally became an IETF standard [4, 5], constituting one of the core technologies of the Internet's infrastructure.

A decade later, DNS had become a critical component in Internet operations, and the absence of security mechanisms called for the first DNS Security Extensions (DNSSEC), published in 1997 [6] and revised in 1999 [7]. A final specification was published in 2005 [8], though more features have been added afterwards, for instance the NSEC3 specification.

DNSSEC builds a chain of trust starting from the DNS root, which is a trust anchor, by signing all subdomain delegations with specific resource records. It also defines signatures for resource records, in order to protect the integrity of the data. Development took a long time, and DNSSEC was not fully deployed at the DNS root level until July 2010. Subsequently, the DNS root and the majority of the top-level domains were signed with DNSSEC in the course of 2010 and 2011. As a result, the trust anchor of the root is now available to build the chain of trust.

However, the nameservers are only one side of the story. In order to ascertain the authenticity of answers, the resolvers must not only request the signatures and keys from the nameservers, but also validate the answers they receive. Since the amount of validating resolvers determines the real DNSSEC coverage experienced by Internet end-users, it is a highly interesting metric for administrators to evaluate the benefits of implementing DNSSEC at their own nameservers.

1.2 Background research

Since the incentive for nameserver administrators to implement DNSSEC relies on a wide support of DNSSEC among DNS resolvers, the Internet community has been monitoring DNSSEC deployment for a long time. Although research focused mostly on nameserver performance and packets size at first, there were early studies on DNSSEC resolvers [9, 10].

The effort on measuring DNSSEC deployment has been mainly led by the Internet Assigned Numbers Authority, and the five Regional Internet Registries. In 2013, scientists from APNIC presented an extensive study on various aspects of the deployment [11].

There are various ways to check a resolver for proper DNSSEC validation. For instance, Yingdi *et al.* developed a method to check for indicators of DNSSEC validation, using software to remove signatures and force validation resolvers to retry the query [12]. Parallelly, Lian *et al.* experimented with a method to generate queries from client browsers, by using advertising networks to load dummy 1-pixel images from controlled domains [13].

These methods all try to address the challenge of estimating end-users' experience without presence in the client networks. This project presents another approach, thanks to the Atlas network. Concerning the data for the measurements, inspiration was taken from [13] when defining zones with corrupted signatures — in order to check the protection offered by DNSSEC.

1.3 Atlas

Originally developed by RIPE Network Coordination Center (NCC) in late 2010, Atlas[1] is a worldwide network of probes, able to provide a deep understanding of the inner operations of the Internet, and probably the largest[2] Internet measurement network ever made.

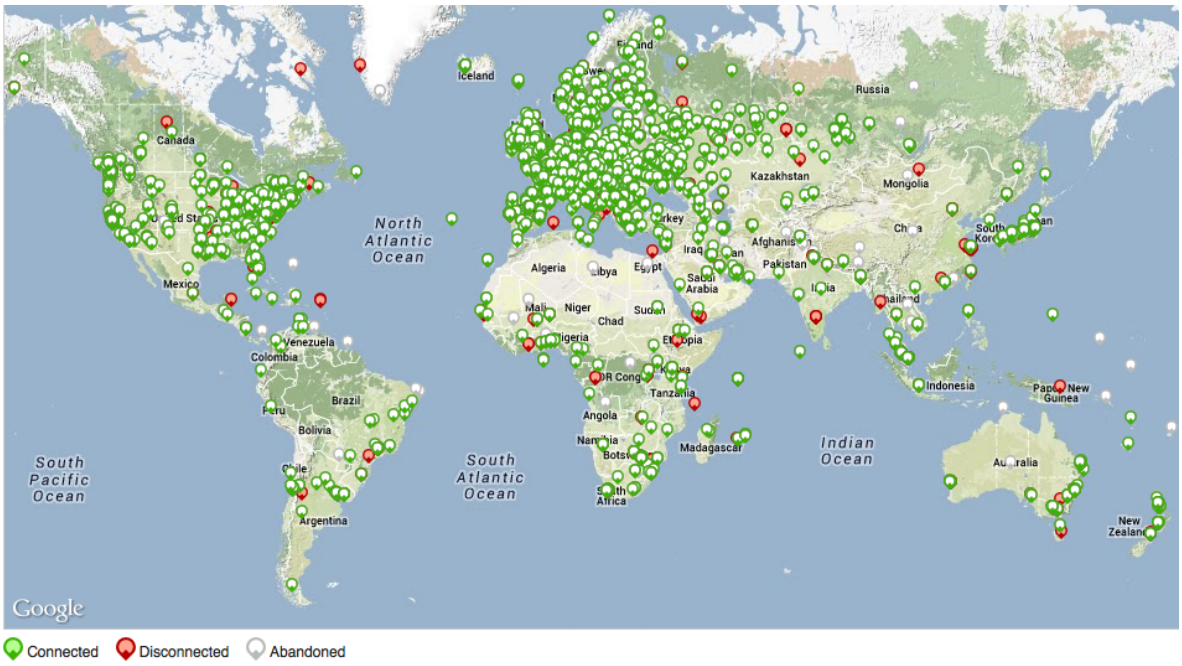

Figure 1.1: The global distribution of Atlas probes
Source: RIPE Labs

The probes are small, USB-powered hardware devices, attached to a host network. Although mostly located in Europe, they are all over the world (see Fig. 1.1), and they can be requested through the Altas API to run a wide variety of measurements on Internet connectivity and reachability.

Atlas is a community project, where people volunteer to host probes, and then receive credits that can be spent to run measurements. The measurement system itself is open-source, and the API is still under active development. It is a prototype service of the RIPE NCC, with no guaranteed quality of service. As shown on Fig. 1.2, there are around eight thousand

[1] See https://atlas.ripe.net.

[2] Excluding botnets used for research, like by Stone-Gross *et al* in *Your Botnet is My Botnet*, or for the Internet Census 2012.

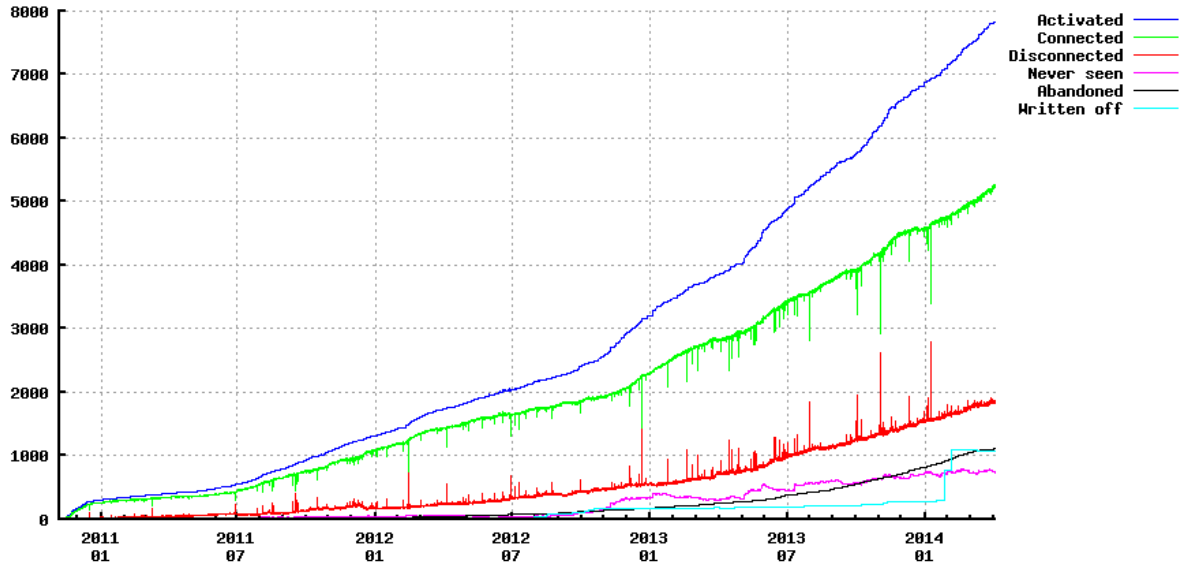

Figure 1.2: The growing number of Atlas probes
Source: RIPE Labs

probes currently deployed, and on average five thousand of them are active and ready to run measurements.

The methods mentioned in the previous section try to check if resolvers validate their answers, but they cannot measure the real effects on end-users' experience. Compared to those methods, using the Atlas network provides a vantage point: presence in the client network. The probes are directly connected to a host network, and receive the addresses of the client network's default resolvers through Dynamic Host Configuration Protocol (DHCP).

As a result, the probes can directly report the answers they receive. Since they use their host network's default resolvers, they are experiencing the same DNS conditions as any other device on the host network. Consequently, the results from the Atlas probes represent the end-users' experience.

1.4 Research scope

This project focused on the following research question:

> **What is the status of DNSSEC deployment over the Internet and how does it impact Internet users?**

The problematics can be thereafter divided into several sub-questions:

- Which DNS resolvers can be queried from clients?
- What methods can properly assess DNSSEC support?
- How does DNSSEC support influence user experience?

User experience is defined not only by potential vulnerability to DNS data forgery, but also by the other improvements that may be provided by DNSSEC. Some resolvers, known as "DNSSEC-aware", support certain DNSSEC features: for instance, they will set the `DO` bit in their requests to nameservers, they will include `RRSIG` records in their answers to the clients, and/or they will be able to query `DS` records.

More importantly, some resolvers will build the chain of trust and validate the answers they receive from nameservers, and will notify the clients that the data is authenticated with the `AD` bit. These validating resolvers enable clients to receive trusted answers, and thus can tremendously influence user experience.

The experiments were exclusively run with the Atlas network, and the scope was restricted to free, open-source software. The hardware and open-source software necessary to run a nameserver were provided by NLnet Labs.

The following sections present the research project itself. Firstly, the main challenges of the project are listed, and the methodology designed to address them is detailed. Then, the results are presented, along with a statistical analysis. Those results are discussed in a final section, with indicative answers to the research questions.

2 Methodology

In the scope of this project, there are two types of actors to investigate: the resolvers, and their clients. The former because it will determine the benefits of implementing DNSSEC, and the latter because the impact of DNSSEC depends on the protection provided to the clients.

Fortunately, the Atlas probes allow presence in the client network, and so valuable data was gathered about both actors. Moreover, the measures also give indications about the distribution of the clients among the resolvers.

The general setup for measurements is shown on Fig. 2.1: probes are used to query a zone under authority of a specific nameserver, which is controlled in order to capture all DNS packets flowing in or out.

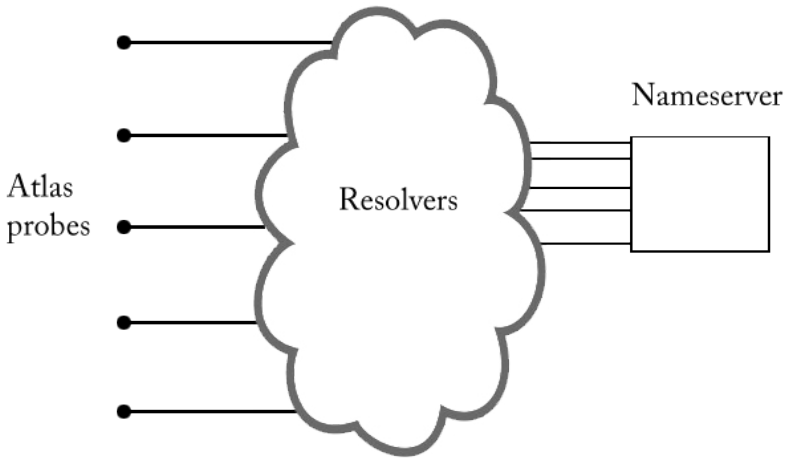

Figure 2.1: Overview of measurements setup

Since they use the client networks' configurations, the working assumption concerning the atlas probes is that they will experience the same DNS conditions as an end-user from the same local network. Consequently, they can be used to characterize the DNSSEC validation coverage. They can also help visualize the protection provided by DNSSEC, by comparing the answer rate from a secure zone to the rate from a corrupted zone. They may even report particular cases, for instance clients that successively query multiple resolvers.

Concerning the resolvers, their support of DNSSEC must be determined. Indeed, a non-validating resolver may still have some understanding of DNSSEC, which could be helpful to the client. Depending on the matching between probes and resolvers, the level of DNSSEC support of the resolvers may be determined.

The nameserver is configured with several zones, each presenting a different security profile. Decoding and indexing the DNS packets captured allows to cross-reference them with the measurements results from Atlas.

2.1 Challenges

Despite having presence in the client networks, the first issue with the measurements is to match the probes with the queries intercepted at the nameserver. Many probes are actually using resolvers with local IP addresses, and some others are using resolvers with anycast addresses: as a result, the request will be coming to the nameserver from a different IP address.

In order to circumvent that problems, it is possible to instruct the probes to prepend their unique identifier as an extra label in the query. For instance, probe `12345` can query `12345.example.com` instead of `example.com`. Coupled with a wildcard record, this allowed to match queries with the probes. Unfortunately, it is not a full solution, for instance it cannot cover the `DS` queries sent to the upper level: thus it is not possible to match measurements results with the packet captures.

Another issue is that probes may receive multiple resolvers via DHCP, which can lead to ambiguous results, and makes it harder to map probes to their resolvers. For instance, when a probe tries multiple resolvers but only one responds: if the probe does not see the public IP address of the resolver, then it is undetermined which resolver answered.

Moreover, those resolvers may use forwarding, which can also impact measurements. For example, some forwarders may not always follow DNS specifications, which can lead them to omit the `AD` bit or the EDNS pseudo-record.

Finally, there is always the probability of a few resolvers being misconfigured or out of order. This will be presented by the types of errors received, and their meaning will be discussed. However, this should only concern a small fraction of the results.

The Atlas network itself has its limits. First of all, even though the probes are spread worldwide, most of them are in Europe (see Fig. 1.1 p. 7). This introduces a bias in the measurements, which can be partially balanced by the geographical distribution of Internet users, as shown in Table 2.1. This issue is discussed in the final section.

Country	Probes	Country	Internet users (in 2012)
United States	853	China	568,192,066
Germany	819	United States	254,295,536
Russia	724	India	151,598,994
United Kingdom	605	Japan	100,684,474
Netherlands	457	Brazil	99,357,737
France	397	Russia	75,926,004
Ukraine	364	Germany	68,296,919
Belgium	184	Nigeria	55,930,391
Italy	166	United Kingdom	54,861,245
Czech Republic	161	France	54,473,474

Table 2.1: Top ten countries by probes and by Internet users
Source for probes: RIPE Labs
Source for Internet users: Wikipedia[1]

[1]See http://en.wikipedia.org/wiki/List_of_countries_by_number_of_Internet_users, and the orig-

Besides, the Atlas network is volatile: probes can go offline or come online at any time, and about two thirds of them are active simultaneously (see Fig. 1.2 p. 8). This is an issue because each measurement takes time to run; it is probable that some probes from one measurements will not be available for the next one.

Additionally, there is a hard limit on the number of probes that can be queried simultaneously for a measurement, so it is impossible to request all active probes for a single measurement. As a result, running enough measurements to get all active probes takes about 100 minutes. Consequently, the specific probes used for a certain set of measurements are not necessarily the same ones used for the next set, even when the total amount of probes is the same.

2.2 Process

The general process of a measurement is the following:

1. List all active probes
2. Start packet capture at the nameserver
3. Launch measurement on Atlas probes
4. Wait for measurement results
5. Stop packet capture
6. Repeat steps 2-5 until all active probes have been used

Fives different zones were used, all of which were subdomains of a securely delegated upper level. The `secure` zone was securely delegated from the upper level and fully signed. The `insecure` zone was neither securely delegated nor signed — a regular DNS zone, without any DNSSEC. The three remaining zones were securely delegated but were incorrectly signed on purpose:

- The label counts included in the `RRSIG` records of the `badlabel` zone were increased by 1
- One character from each `RRSIG` record in the `badrrsigs` zone was arbitrarily modified
- All `RRSIG` records were removed from the `norrsigs` zone

The nameserver itself is running the `nsd` software[2], and the asymmetric key pairs and signatures are managed using the `ldns` library[3] — `ldns` was even patched and recompiled to produce bad signatures for the badly signed zones. Packet captures are conducted directly at the nameserver, using the Wireshark network analyzer[4].

The measurements are run by calling the Atlas API through a Python[5] 2 library written by NLnet Labs engineers. Each measurement has a unique identifier that can be used to retrieve

inal source from the International Telecommunications Union [14].

[2]See `http://www.nlnetlabs.nl/projects/nsd`.

[3]See `http://www.nlnetlabs.nl/projects/ldns`.

[4]See `http://www.wireshark.org`.

[5]See `https://www.python.org`.

the results. All identifiers of measurements used for this study are presented in App. B.

A set of Python 2 scripts[6] was specifically written for this project in order to automate the measurement process, and compute statistics on the results. These scripts are also used to cross-reference the measurement results with the packet captures, and to run statistics on the packets, using the `dpkt` library[7].

On each zone, all measurements were run on `TXT` and `A` records, because they are the most widely supported types. It was confirmed that there is no difference between statistics on each of the two record types. Additionally, measurements were run on the `secure` zone with the `DS` type, because they can detect DNSSEC-awareness.

The `DS` record type is a peculiar case in DNS. It was introduced by DNSSEC: when a domain securely delegates authority, the delegated domain creates a key to sign its records; in order to build the chain of trust, the delegating domain needs to authenticate that key. In the end, a domain `deleg.example.com` is not authoritative for its own `DS` record: the upper-level domain `example.com` is. Consequently, the `DS` record is considered out-of-zone data by resolvers that are not aware of RFC 4033 [8].

As a result, if a resolver returns an answer for a `DS` record, it may mean that it has at least a basic support of DNSSEC. Indeed, when a probe sends a request for the `DS` type, if the resolvers queries the upper-level nameserver for the `DS` record, it is DNSSEC-aware. However, there is also the possibility that the `DS` record was cached during the initial lookup, because the queries always had the `DO` bit set: hierarchically, the upper-level nameserver `example.com` will be queried before the `target.example.com`, which may have the side-effect of caching the `DS` record of `target.example.com`, even if the resolver is not DNSSEC-aware.

As a final step of the process, it was attempted to completely map the probes to their resolvers. However, due to the issues mentioned earlier, this task revealed impossible to automate: there is often no match between the resolver's address seen by the probe, and the address seen by the nameserver. This happens when the probes access the resolvers either through a local IP address such as `192.168.0.254` or through an anycast address like `8.8.8.8`, and prevents an easy matching. As a result, only the mostly used resolvers were identified.

[6] See https://github.com/ncanceill/atlas-dnssec.
[7] See https://code.google.com/p/dpkt.

3 Results

3.1 Resolvers

The most interesting characteristic of resolvers is their understanding of DNSSEC: are they aware of DNSSEC specificities? and are they able to validate DNSSEC answers?

The first indicator of DNSSEC support is the `DO` (DNSSEC OK) bit: it is set by the probe in the query, and it indicates that a secure answer is requested from the resolver. The `DO` bit is set in the `OPT` pseudo-record as defined in EDNS [15], which is sent in the Additional sections. If the resolvers set the `DO` bit in their own queries to the nameserver, it means they indicate their own support of DNSSEC.

Additionally, it is interesting to check if resolvers include `RRSIG` records in their answer. Not only does it indicate that they are DNSSEC-aware, it means the client receives the signatures. Then, any application could be used to perform validation on the client side.

The results in Table 3.1 are from `TXT` queries on a regular record in the `secure` zone. The probes set the `DO` bit, then packets are intercepted at the nameserver in order to see wether the resolver also set the `DO` bit when they query the nameserver.

Probes	**Resolvers**	**Setting `DO` bit**	**`RRSIG`s**
4673	5139	4534 [88.23%]	3448 [67.09%]

Table 3.1: Results of `TXT` queries intercepted on `secure` zone

According to the specifications, all DNSSEC-aware resolvers should set the `DO` bit when they query the nameserver, so this should provide a good estimation. There are more resolvers than probes because many probes have multiple resolvers, and query them successively until a response arrives.

As mentioned in the previous section, support of `DS` requests is also a good indicator of DNSSEC-awareness, because `DS` records may be considered out-of-zone data by DNSSEC-unaware resolvers. Table 3.2 presents the results from `DS` queries on the `secure` zone. Packets were intercepted since the upper-level zone is served by the same nameserver.

Probes	**Answer**	**AD bit**	**`RRSIG`s**	**No `RRSIG`s**	**`FORMERR`**
5602	5323 [95.01%]	1557 [27.79%]	2176 [38.84%]	1590 [28.38%]	268 [4.78%]

Table 3.2: Results of `TXT` queries intercepted on `secure` zone

The percentage in the **Answers** column of Table 3.2 only counts the replies containing an Answer section.

It was also possible to observe the effect of resolvers re-trying requests on corrupted zoned. As shown in Table 3.3, the corrupted zone generates close to six times more queries.

Zone	Resolvers	Requests/probe
`secure`	4586	2.20
`badrrsigs`	49048	11.89

Table 3.3: Results of `TXT` measurements on `secure` and corrupted zones

3.2 Users

3.2.1 Distribution of resolvers

The first step in understanding end-users' experience with DNS is to try and pair DNS resolvers with their clients. This is presented in Fig. 3.1: the graph shows the amount of resolvers that have a specific amount of probes.

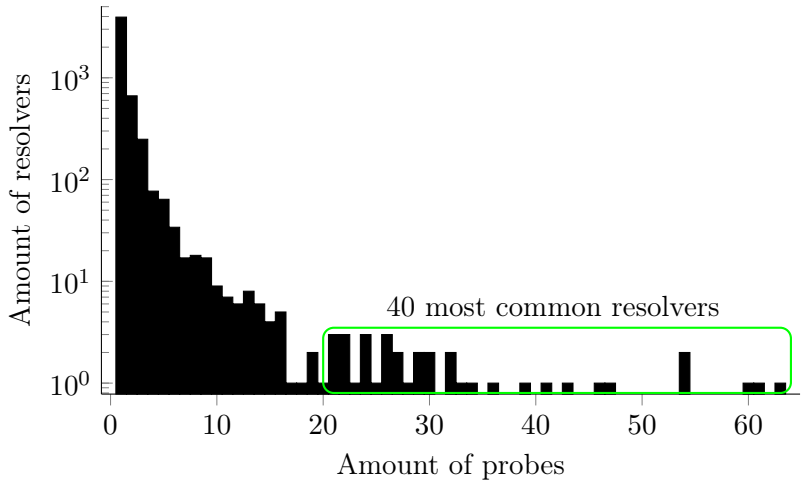

Figure 3.1: Distribution of resolvers among probes

This clearly show that most resolvers only impact between 1 and 5 probes, and that just about 40 resolvers impact more than 20 probes each. Out of these 40 most common resolvers, 38 came from IP addresses registered at Google, and 2 from addresses at OVH. The addresses are listed in App. A.

Interestingly enough, most of those requests coming from Google DNS resolvers originated from probes that did not have Google Public DNS[1] as their default resolver. The most likely explanation for this is that Internet Service Providers redirect to Google resolvers, probably through the clients' local home routers (acting as local resolvers), or a set of forwarders.

[1]See https://developers.google.com/speed/public-dns.

3.2.2 DNSSEC protection coverage

From the end-users' point of view, the main question is wether they are protected against DNS attacks. This is determined by the difference between the amount of answers accepted for a query on a secure zone, and for a query from a zone with security issues.

This difference, due to resolvers not accepting the answers that present security issues, can be observed in Table 3.4, which summarizes the results of measurements of `TXT` queries (with the `DO` bit set) on the corresponding wildcard records.

In case the probe did not receive an Answer section in the response, the response code is taken into consideration. Indeed, non-acceptance of a badly signed record should lead the resolver to return a `SERVFAIL` code.

Zone	**Probes**	**No Answer**	`SERVFAIL`	`FORMERR`	Parse Error
`secure`	5457	297 [5.44%]	12 [0.22%]	263 [4.82%]	100 [1.83%]
`badlabel`	5366	1735 [32.33%]	1410 [26.28%]	302 [5.63%]	81 [1.51%]
`badrrsigs`	5427	1739 [32.04%]	1417 [26.11%]	299 [5.51%]	67 [1.23%]
`norrsigs`	5491	1737 [31.63%]	1416 [25.79%]	306 [5.57%]	20 [0.36%]

Table 3.4: Results of `TXT` measurements on `secure` and bad zones when Answer section is not present

It was confirmed that the `FORMERR` (format error) code was seen because the queries had the `DO` bit set: it is not in the DNS packet header, but in the `OPT` pseudo-section introduced by EDNS0 [16]. Resolvers seem to set it to indicate that they do not support EDNS0. The "Parse Error" column does not indicate a DNS return code: it simply counts packets that were received malformed and could not be parsed.

Since those are wildcard records, a single signature is not enough: in order to validate the answer, the client needs an `NSEC` record to be sure that the name queried did not exist — thus proving that the wildcard is indeed the correct match. This difference is outlined in the two penultimate columns of Table 3.5.

Zone	**Probes**	**Answers**	**AD bit**	`RRSIGs+NSEC`	`RRSIGs` **only**	**Just Answer**
`secure`	5457	5160 [94.55%]	1472 [26.97%]	1109 [20.32%]	967 [17.72%]	1612 [20.54%]
`badlabel`	5366	3631 [67.66%]	0 [0.00%]	1014 [18.90%]	1004 [18.71%]	1613 [30.06%]
`badrrsig`	5427	3688 [67.95%]	0 [0.00%]	1017 [18.74%]	1034 [19.05%]	1636 [30.15%]
`norrsigs`	5491	3754 [68.37%]	0 [0.00%]	0 [0.00%]	0 [0.00%]	3754 [68.37%]

Table 3.5: Results of `TXT` measurements on `secure` and bad zones when Answer section is present

It is also very important to confirm that signing a zone does not harm users' experience. This is determined by the difference between the amount of answers from the secure zone and the amount from the insecure zone: indeed, 93.71% of the probes got an answer for the insecure zone, to compare to the 94.55% in Table 3.5.

4 Conclusions

Combining a packet-capturing nameserver with the power of the Atlas network allowed to run DNSSEC measurements with novel methodologies, and to estimate the impact on end-users' experience with high accuracy. Among a sample of more than five thousand resolvers, around 90% were found to be DNSSEC-aware, and more than 27% were even validating their answers and rejecting corrupted signatures.

4.1 Results discussion

The most important findings are the proportion of DNSSEC-aware and validating resolvers among the sample available through the Altas network. The observed 28% of validating resolvers is a significative improvement from the mere 3% found in [13]. The observed 88% of DNSSEC-aware resolvers (sending the `DO` bit) is close to the 80% found in [11], but that study only found around 8.3% of validating resolvers.

4.1.1 Distribution of resolvers

It should be taken into account that the global distribution of Atlas probes does not necessarily reflect the real distribution of Internet users. As presented in Table 2.1 p. 11, some countries like India and Brazil are clearly under-represented in the Atlas network. Paralelly, early DNSSEC adopters (United States, Western and Northern European countries) are very well represented, and so the results may reflect a deployment that is higher than the global average.

Moreover, Atlas is simply a project, not an official RIPE service. Only the members of the project can volunteer to set up probes, and their host networks may not always be representative of all Internet users. For instance, they might be more likely to implement DNSSEC on a local resolver, instead of using a standard solution from an Internet Service Provider, which could bias the measurements.

The distribution of resolvers among the probes (see Fig. 3.1 p. 15) is an expected result: they are usually served by DHCP by the Internet Service Provider, and it is assumed that most of them are local forwarders, therefore only covering a few probes. Besides, the "big players" like Google have anycast IP addresses, which are seen as multiple distinct IP addresses by the nameserver.

4.1.2 DNSSEC support

All results tend to indicate that DNSSEC deployment is subject to gradation: there are multiple levels of DNSSEC support. There are at least two classes of DNSSEC-aware servers:

significantly more resolvers were observed accepting DS records than resolvers sending the DO bit to the nameserver — it could also be that some DNSSEC-unaware resolvers still send the DS records because they were previously cached while walking the DNS tree. There is also an unexplained difference between the resolvers setting the AD bit on valid answers and the ones returning SERVFAIL on corrupted signatures: protection is not always coincidental with validation.

There is an additional interest in DNSSEC-awareness: it allows resolvers to forward RRSIG records back to the client. This is very interesting, because it could allow the client to perform the signature verification by itself, at the application level. Moreover, such a solution could solve the "last mile problem": all the security of DNSSEC relies on the client's default resolver, which is often silently set via DHCP. In the measurements on the secure zone, about 67% of the resolvers sent the signatures in the answer (see Table 3.1 p. 14), which could allow many clients to validate answers themselves.

Comparing between these findings and the results reported in [13, 11], an important difference appears in the rate of validating resolvers. The first explanation is the possible bias introduced in this study by the distribution of Atlas probes, as discussed above. Another possible factor is that the methods in [13, 11] could also be biased: advertising networks may arbitrarily distribute the ads among users, and the resulting iframes and Javascript calls could be blocked in the browser, so their data may not reflect the average situation of Internet users either.

4.1.3 Particular cases

Among the resolvers able to send signatures back to the client, some seem not to understand the particular case of wildcard records: besides a signature authenticating the wildcard record itself, the client needs an authenticated NSEC record in order to prove that the record queried did not exist. Due to this issue, client-side validation is only possible on 47% of probes, compared to 67% for regular records.

Additionally, there seems to be a specific level of DNSSEC protection: Table 3.4 shows that the norrsigs zone generates about 0.5 percentage points less SERVFAIL than the other corrupted zones. A possible explanation would be that the corresponding resolvers have a safe mode (they fail when the RRSIG record is incorrect), but they will fallback to unsafe mode in case the RRSIG record is completely missing.

Finally, despite a novel methodology allowing presence both at the nameserver and in the client network, there is still an unknown part: this study only covers the extremities of the resolving chain. Forwarding DNS can be harmful, and it is not impossible that some resolvers affect the packet on the way — removing Additional or Answer sections, unsetting the AD bit...

4.2 Contributions

This project focused on providing the Internet community with a better understanding of the state and the impact of DNSSEC deployment. Now that the results have been presented and discussed, answers to the research questions may be proposed.

- **Which DNS resolvers can be queried from clients?**
  As for Internet users, the probes can have many different configurations: they may use a local resolver, a forwarder, or a remote revolver. The clients and resolvers may have multiple resolvers themselves, and try them successively until one answers.

- **What methods can properly assess DNSSEC support?**
  There are multiple levels of DNSSEC support. Using queries on `DS` records, checking for the `DO` bit at the nameserver, and looking for signatures in the answers, can estimate DNSSEC-awareness.

- **How does DNSSEC support influence user experience?**
  DNSSEC validation can bring authentication to the insecure world of DNS, while DNSSEC protection prevents DNS hijacking and other malicious attempts to disrupt Internet operations. Even DNSSEC-awareness is useful, since it ensures that the exceptions of DNSSEC are supported, and can even enable application-level validation at the client.
  Threats to user experience:
  It was found that DNSSEC wildcard records are badly handled in 20% of cases, thus preventing client-side validation. Moreover, the case of an insecure fallback when signatures are missing, discovered in [13], was confirmed with similar results.

What is the status of DNSSEC deployment over the Internet and how does it impact Internet users?
DNSSEC deployment is progressing over the Internet, especially concerning validation and protection. More and more users are benefiting from the tree of trust that allows validation. More importantly, users are increasingly protected against corrupted or spoofed signatures. Nevertheless, there are still some issues with DNSSEC validation that may negatively impact Internet users.

4.3 Further research

DNS is a complex protocol, and DNSSEC is even less simple. As a result, there are many parameters to tweak, many hypotheses to consider. This project studied the benefits of implementing DNSSEC, and there needs to be studies on the costs too.

This project did not focus on packet size and Maximum Transmission Unit, but these are also very important metrics. Particularly, the effect of UDP fragmentation and fallback to TCP should be examined.

The time dimension was out of the scope of the study, but measuring the delays in answers could reveal precious indications. DNSSEC validation necessarily takes more time than basic DNS. Moreover, this could help detect timeouts and re-tries.

The results prove the existence of a vast disparity of levels of DNSSEC support: aware of `DS` records, sending signatures, verifying signatures, validating and setting the `AD` bit. A strict method of assessing well-defined levels of DNSSEC support would be very interesting to push this study further.

It would also be worthy to have a full understanding of the differences between the various

studies on DNSSEC deployment, as mentioned in Section 4.1.2. In the end, the best future for this research would be an even larger probe network. It is suspected that measurements may be biased because the Atlas probes are not an accurate representation of all Internet users. Future studies will need methods to ensure a proper distribution of clients all over the Internet.

A Most common resolvers

The following addresses are the 40 most common resolvers from Fig. 3.1 p. 15, sorted from the most common to the least common.

74.125.189.21	2a00:1450:4001:c02::153	74.125.17.145
74.125.189.20	74.125.18.210	74.125.18.83
74.125.189.16	2a00:1450:4001:c02::156	2a00:1450:4001:c02::151
74.125.189.19	173.194.98.147	2a00:1450:4001:c02::154
74.125.189.23	173.194.98.144	188.165.197.144
74.125.189.17	74.125.18.215	74.125.17.149
74.125.189.22	2a00:1450:4001:c02::152	74.125.18.209
74.125.189.18	74.125.18.213	74.125.17.146
173.194.98.146	74.125.18.211	2a00:1450:4001:c02::157
173.194.98.150	2a00:1450:4001:c02::155	74.125.18.81
74.125.18.82	173.194.98.149	74.125.17.144
173.194.98.148	74.125.18.208	74.125.18.84
173.194.98.145	173.194.98.151	74.125.181.80
74.125.18.80	91.121.161.184	

B Measurements identifiers

Measurements results can be freely downloaded from Atlas, given the measurement identifiers. For the sake of reproducibility, all measurements used for this study have their identifiers listed below.

1578280	1596404	1596641	1602195	1603637	1605008	1606438	1608023
1578284	1596405	1596642	1602196	1604812	1605009	1606439	1608024
1578500	1596406	1596644	1603555	1604952	1605011	1606440	1608025
1578883	1596408	1596646	1603566	1604953	1605012	1606441	1608184
1582162	1596409	1596648	1603567	1604954	1605014	1606442	1608195
1582567	1596411	1596649	1603569	1604956	1605015	1606443	1608206
1589840	1596412	1596650	1603570	1604957	1605016	1606447	1608218
1589851	1596413	1596652	1603571	1604958	1605017	1606679	1608229
1589853	1596414	1596653	1603574	1604960	1605020	1606686	1608240
1589867	1596417	1596654	1603575	1604961	1605022	1606687	1608241
1589868	1596418	1596655	1603576	1604962	1605037	1606688	1608244
1589869	1596419	1596657	1603577	1604964	1605038	1606689	1608247
1589870	1596593	1596662	1603579	1604984	1605039	1606690	1608249
1591162	1596595	1596667	1603580	1604985	1605040	1606691	1608250
1591170	1596608	1596668	1603581	1604986	1605042	1606693	1614235
1591173	1596611	1596669	1603585	1604987	1605043	1606819	1614236
1591178	1596613	1602168	1603626	1604988	1605044	1606969	1614237
1591184	1596614	1602171	1603627	1604990	1605047	1607131	1614238
1596393	1596623	1602174	1603628	1604995	1605048	1608014	1614240
1596394	1596627	1602175	1603629	1604996	1605049	1608015	1614241
1596395	1596628	1602176	1603630	1604999	1605051	1608016	1614245
1596396	1596630	1602177	1603632	1605000	1606430	1608017	1614247
1596398	1596632	1602178	1603633	1605001	1606431	1608018	1614248
1596400	1596634	1602180	1603634	1605003	1606433	1608020	1614249
1596402	1596637	1602187	1603635	1605006	1606434	1608021	1614251
1596403	1596640	1602192	1603636	1605007	1606435	1608022	

List of Figures

List of Tables

List of Acronyms

DHCP Dynamic Host Configuration Protocol

NCC Network Coordination Center

Bibliography

[1] P. V. Mockapetris. *RFC 882: Domain names: Concepts and facilities.* Obsoleted by RFC1034, RFC1035. Updated by RFC973. 1983.

[2] L. P. Deutsch. *RFC 606: Host Names On-line.* 1973.

[3] J. Postel and J. Reynolds. *RFC 920: Domain Requirements.* 1984.

[4] P. V. Mockapetris. *RFC 1034: Domain names — Concepts and facilities.* 1987.

[5] P. V. Mockapetris. *RFC 1035: Domain names — Implementation and specification.* 1987.

[6] D. Eastlake and C. Kaufman. *RFC 2065: Domain Name System Security Extensions.* Obsoleted by RFC2535. 1997.

[7] D. Eastlake. *RFC 2535: Domain Name System Security Extensions.* Obsoleted by RFC4033-4035. Updated by RFC2931, RFC3007, RFC3008, RFC3090, RFC3226, RFC3445, RFC3597, RFC3655, RFC3658, RFC3755, RFC3757, RFC3845. 1999.

[8] R. Arends et al. *RFC 4033: DNS Security Introduction and Requirements.* Updated by RFC6014, RFC6840. 2005.

[9] Jelte Jansen. *Measuring the effects of DNSSEC deployment on query load.* Tech. rep. NLnet Labs, 2006.

[10] Daniel Migault, Cédric Girard, and Maryline Laurent. "A Performance view on DNSSEC migration". In: *6th International Conference on Network and Service Management.* Niagara Falls, Canada: IEEE, 2010, pp. 469–474.

[11] Geoff Huston and George Michaelson. "Measuring DNSSEC Use". In: *IEPG Meeting 87.* Berlin, Germany: IETF, 2013.

[12] Yingdi Yu et al. "Check-Repeat: A new method of measuring DNSSEC validating resolvers". In: *32nd International Conference on Computer Communications.* IEEE, 2013, pp. 3147–3152.

[13] Wilson Lian et al. "Measuring the Practical Impact of DNSSEC Deployment". In: *22nd USENIX Security Symposium.* Washington, D.C., USA: USENIX Association, 2013, pp. 573–587.

[14] *Percentage of Individuals using the Internet 2000-2012.* International Telecommunications Union (Geneva). 2013.

[15] J. Damas, M. Graff, and P. Vixie. *RFC 6891: Extension Mechanisms for DNS.* 2013.

[16] P. Vixie. *RFC 2671: Extension Mechanisms for DNS (EDNS0).* Obsoleted by RFC6891. 1999.