IPv6 Hitlist Service: Lessons Learned From 10 Years of Operation

Oliver Gasser
IPinfo
oliver@ipinfo.io

Lion Steger
Technical University of Munich
stegerl@net.in.tum.de

Patrick Sattler
BENOCS
psattler@benocs.com

Johannes Zirngibl
Max Planck Institute for Informatics
johannes.zirngibl@mpi-inf.mpg.de

Abstract

After becoming an Internet Draft more than 30 years ago, IPv6 has seen an increase in deployment and use in the past years. As measurements in the IPv6 Internet require new approaches due to the vastly larger address space, hitlists have come along as one possible source for finding IPv6 targets. One of the most prominent hitlists is provided by the IPv6 Hitlist Service.

In this paper, we share insights from 10 years of operations of the IPv6 Hitlist Service: We show different evolutions of the service, highlighting important changes along the way. To better understand the representativeness of the hitlist, we perform a coverage analysis using real-world traffic data from a major central European ISP and Tier-1 network, finding that at least one address is known to the IPv6 Hitlist Service for 87.1 % of ASes and 56.5 % of /48 prefixes originating IPv6 traffic. We also share results from a conducted user survey and analyze users accessing the IPv6 Hitlist Service, finding diverse use cases and access patterns across time (e.g., one-off vs. continuous downloads) and available data (e.g., all vs. responsive addresses). Finally, we provide best practice recommendations when working with the hitlist and share lessons learned during its 10-year operation.

CCS Concepts

• **Networks** → **Network measurement**.

Keywords

IPv6, Hitlist, Internet scanning

1 Introduction

In March 2026, IPv6 user traffic to Google services for the first time surpassed IPv4 traffic [18]. This milestone shows that IPv6—more than 30 years after becoming an Internet Draft [9]—has finally surpassed this important threshold. An increase in IPv6 traffic is not unique to Google, but has been observed by other companies and organizations as well [2, 7]. The many architectural changes introduced by IPv6 compared to its predecessor also affect the Internet measurement community. One of these important changes is the address space, which in IPv6 is 10^{29} times larger compared to IPv4. This means that while in IPv4, brute-force scanning of the complete address space is possible within hours [11, 20], it is not feasible in IPv6. Therefore, IPv6 measurements usually use techniques such as hitlists [15, 17, 37, 40, 42, 51], Target Generation Algorithms (TGAs) [23, 31, 41, 46], or similar approaches to perform smart target selection.

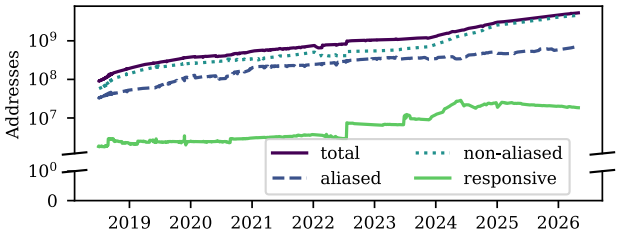

Figure 1: Temporal development of the IPv6 Hitlist Service.

One of the most prominent examples of this is the IPv6 Hitlist Service [25], which was first created back in 2016. The initial IPv6 hitlist TMA 2016 paper [17] focused on the analysis of potential hitlist sources and their differences and its service was run mostly internally, providing only partial access to IPv6 addresses upon request [16]. The follow-up IMC 2018 paper [15] had a public IPv6 hitlist [25], which could be directly used by other researchers, as one of its main contributions. The ongoing service was actively maintained and extended in the following years [42, 51]. As can be seen in Fig. 1, the number of addresses has seen a large increase over time: 60x total addresses and 10x responsive addresses. Moreover, since 2018, the IPv6 Hitlist Service has seen more than 150 registered users and an even larger number of people downloading the publicly available version. Furthermore, the hitlist has been instrumental in many IPv6 measurement research works and has been cited by more than 240 publications as of May 2026. The majority of these publications are from the Internet measurement community (22x ACM IMC, 13x PAM, 3x TMA), but also the broader networking community (10x IEEE INFOCOM, 7x IEEE/ACM TON, 6x ACM CoNEXT) and the security community (3x USENIX Security, 2x IEEE S&P, 2x IEEE EuroS&P) are making use of the hitlist.

In this paper, we share insights on 10 years of operations of the IPv6 Hitlist Service. Specifically, we make the following main contributions:

- **Changes Over Time:** We highlight different changes affecting the hitlist over time, such as additional data sources and changing characteristics of IPv6 networks (see Sec. 2).
- **Coverage:** We compare the coverage of the hitlist to real-world traffic data from a major central European ISP and Tier-1 network. At least one address is known to the IPv6 Hitlist Service for 87 % of ASes and 56 % of /48 prefixes originating IPv6 traffic (see Sec. 3).

- **User Survey:** We report on results from a conducted user survey which shows different use cases of the hitlist data and underlines its importance for IPv6 measurement research (see Sec. 4).
- **User Behavior:** We share insights on hitlist user behavior and show different access patterns (daily vs. one-off, all vs. responsive addresses) for un/registered users (see Sec. 5).
- **Best Practices & Lessons Learned:** We provide recommendations for best practices in hitlist usage and share lessons learned from building a successful measurement service (see Sec. 6).

2 Technical Development

In this section, we detail the architecture and evolution of the IPv6 Hitlist Service. This includes an analysis of the technical changes, modifications to the input sources and their respective impact on the data, and an automated analysis of trends in the data.

IPv6 Hitlist Service Pipeline: The first step in the IPv6 Hitlist Service pipeline is to combine different sources of addresses, which we further analyze in the following paragraphs, as well as the cumulative input from the last scans. These addresses are filtered for duplicates, resulting in 5.24 B addresses as of May 2026, which are published as the *input* dataset. After that, a blocklist filter is applied in order to enforce the opt-out mechanisms we implement for ethical scanning, followed by a filter for routable prefixes, during both of which only 81 M addresses are filtered out. An overview can be seen in Fig. 5 in Sec. A.

At this point, we apply our Aliased Prefix Detection (APD) method, where we sort all addresses into BGP prefixes and subnets of `/64` to `/120` in steps of 4 bits, considering any subnet which contains more than 100 addresses, except for `/64`, where all subnets are considered. For all of these subnets, we generate 16 random addresses, one for each possible value `0–f` of the first nibble after the subnet part of the address. We then scan all of these addresses with one ICMP and TCP/80 probe each and label the subnet as aliased if we receive a response for either probe for all 16 addresses in this or the previous two scans (see [15, 51] for more details). This step removes 678 M addresses. Finally, any address that was unresponsive for more than 30 days is filtered, which avoids repeatedly scanning inactive addresses. This step removes the majority (4.02 B) of addresses, leaving 458 M addresses to be scanned.

We randomize the addresses to evenly distribute scanning traffic over networks and conduct traceroute measurements for all addresses. We probe the addresses with different probe types, including ICMPv6 Echo Requests, TCP SYN packets on ports 80 and 443, and UDP packets with a QUIC initial packet for port 443 and a DNS query on port 53. As a last step we send multipath TCP probes to ports 80 and 443 [3] and an SNMPv3 probe on UDP port 161 [1]. All addresses responsive to at least one of these protocols (excluding traceroute) are published as *responsive addresses* without registration, and the exact scan results for the different probes are published in the registration-first section.

Technical Changes: Next, we summarize the most important technical changes made to the IPv6 Hitlist Service pipeline over time; a summary can be found in Tab. 1. The service ran without major changes until 2021, when first scans with Multipath TCP (MPTCP) probes were introduced, then UDP probes for port 161 with an SNMPv3 payload were added. As part of the 2022 paper [51], we identified many DNS responses as injections by the Great Firewall of China (GFW) and filtered them. This not only improved the quality of our scan results, but also greatly reduced the runtime of the subsequent scans as visible in Fig. 3. Since this filter still did not remove all injections from the GFW, we changed the queried domain in the DNS probe from `www.google.com`, which is blocked by the GFW, to a custom, non-blocked domain. Between November 2024 and April 2025, we implemented optimizations which decreased the scanning runtime, see Fig. 3. Other upward spikes in scanning runtime are due to outages and hardware migrations.

Table 1: Summary of the most important changes to the IPv6 Hitlist Service. Bold changes are visualized in subsequent plots.

Date	Summary
2021-03-22	Addition of MPTCP scans
2021-06-05	Addition of UDP/161 scans
2022-01-13	**Removal of injected DNS responses**
2023-03-24	Change of DNS payload to avoid injections
2024-03-16	**Removal of IPinfo sources as input**
2024-11-07	**Runtime optimizations**

Table 2: Input sources of the IPv6 Hitlist Service. Regular sources are included for each new scan. One shot sources are fixed sets of addresses being added only once or a few times. The first scan of the IPv6 Hitlist Service was conducted and published on July 01, 2018, using address sources marked as *First scan*. Some sources were excluded later for different reasons, e.g., data unavailability. Bold sources are highlighted in subsequent plots.

Type	Timespan	Source
Regular	First scan	DNS resolutions + AXFR
Regular	First scan – 2023-04-01	Bitnode addresses
Regular	First scan	Traceroutes from last scan
Regular	First scan	RIPE Atlas traceroutes
Regular	First scan	Open IP map
Regular	First scan, 2020-08-24 – 2022-05-31, 2023-02-24 – 2025-04-16	**Rapid7 DNS ANY**
One shot	2018-08-27	**rDNS data (Fiebig et al. [13])**
One shot	2019-02-19	**rDNS data**
One shot	2022-06-20	**TGA addresses (IMC)**
One shot	2023-06-16	**TGA addresses (TMA)**
Regular	2023-11-27 – 2025-01-22	**IPinfo 1**
Regular	2023-11-27 – 2024-03-16	**IPinfo 2**
One shot	2024-03-16	30-day-filtered addresses
Regular	2024-10-11	CAIDA ITDK traceroutes

Changes to Input Sources: The address sources used in the IPv6 Hitlist Service were subject to changes over time, as summarized in Tab. 2. While all sources added until November 2018 were already described in the 2018 publication by Gasser et al. [15], multiple new address sources were added afterwards. We summarize these changes in the following and analyze their impact. We mark changes with observable impact with dotted vertical lines in Fig. 2. Fiebig et al. [13] proposed a methodology to enumerate IPv6 addresses based on rDNS by exploiting the NXDOMAIN (denial of existence) semantics. Data from their work was added to the IPv6 Hitlist Service in 2018 and our own scans were conducted in 2019 following the same approach. From mid 2020 onward, we included the Rapid7

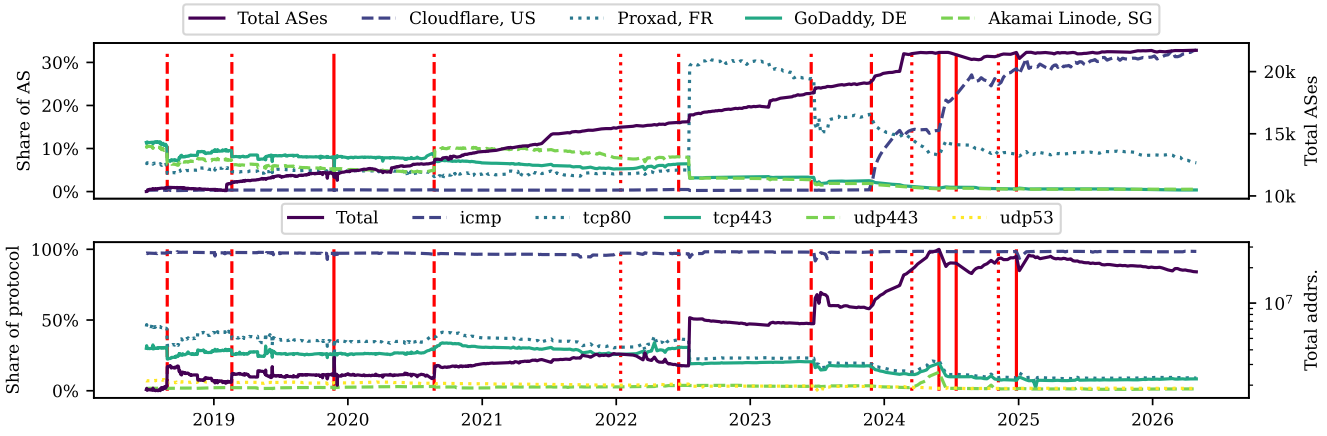

Figure 2: Development of the IPv6 Hitlist Service measured by different metrics. Vertical lines represent the addition of selected new sources (dashed), major changes to the pipeline (dotted), changes without a corresponding change to the service (solid).

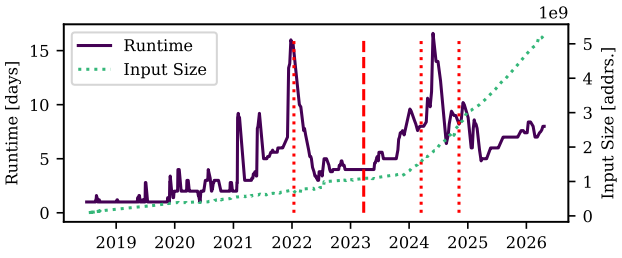

Figure 3: Temporal development of the scan runtime of the IPv6 Hitlist Service compared to the total input size. Dashed vertical lines represent the addition of selected new sources, dotted vertical lines represent major changes to the pipeline.

DNS ANY dataset regularly, excluding it when our research access was not renewed, leading to an increase in total responsive addresses and a slightly increased share of addresses from AS63949 (Akamai Linode). In 2022, following the publication by Zirngibl et al. [51], addresses generated by different Target Generation Algorithms (TGAs) were included, resulting in a 173 % increase in responsive addresses, a smaller increase in covered ASes and a shift in AS and protocol distribution. The newly included addresses were mainly responsive to ICMP probes, with 27 % from one AS (AS12322), which shifted the protocol and AS distributions. In 2023, TGA outputs from another publication [42] were included, inducing another increase in responsive addresses and covered ASes with a smaller share of the top ASes. From November 2023 onward, two datasets from IPinfo [24] were regularly included. This introduced a strong increase in IP addresses, covered addresses, but also scanning runtime, see Fig. 3 (dashed vertical line). The latter was mainly caused by the increase in `/64` prefixes included in the cumulative input, each of which are scanned by our APD mechanism. This source also introduced a lot of addresses from AS13335 (Cloudflare), which shifted the share of represented ASes in the IPv6 Hitlist Service. The large number of addresses and other side effects led to the removal of these sources in 2024 and 2025 respectively, in turn leading to a (delayed) decrease in scan runtime. Lastly, the source also led to an increase in total addresses in the IPv6 Hitlist Service, while the number of addresses in aliased prefixes remained relatively stable, indicating the introduction of many non-aliased addresses. An analysis of the APD results for prefixes from Cloudflare shows that we often receive fewer than 16 responses when evaluating subnets of aliased prefixes. This is an anomaly [15, Section 5.1], which might be caused by ICMPv6 rate limiting. Erdemir et al. [12] suggest adapting the response threshold from 16 to values as low as 10 to minimize this effect. This shows that continuously adding sources adds value, but ongoing maintenance is required.

Automated Analysis of Changes: Lastly, we conduct an automated trend analysis of the data. We start by comparing all scan dates that exhibit an absolute change in responsiveness of more than 120 % compared to the next scan date. We exclude dates which are already explained by changes in our sources or pipeline and analyze the remaining changes leaving us with the solid vertical lines in Fig. 2. In late 2019, we observe a drop in responsiveness throughout all ASes and protocols, which was caused by errors in our pipeline following a hardware fault. In mid 2024, we observe two strong decreases in responsiveness. The IPinfo source introduced a constant number of active addresses, which however changed a lot between scan dates. Addresses usually became unresponsive after few scans, being replaced by other responsive addresses, leading to a lot of address churn. The removal of the IPinfo source led to the short-lived addresses not being replaced anymore, which in turn led to an overall decrease in responsiveness with substantial drops when bigger providers rotated addresses or prefixes.

3 IPv6 Hitlist Coverage

While the IPv6 Hitlist Service collects addresses from different sources, it is difficult to quantify its coverage of the IPv6 ecosystem. To shed light on this question, we collaborate with a major central European ISP and Tier-1 provider and compare addresses of the IPv6 Hitlist Service to visible traffic. We use flow data collected at

Table 3: Coverage of ASes and /48 prefixes in the IPv6 Hitlist Service (Apr. 30, 2026) compared to toplist resolutions in three months of flow data (containing at least one address or aliased prefix). Traffic indicates the share of IPv6 traffic these prefixes originate during the observation period.

		ASes		/48		Traffic [B]
		#	%	#	%	%
IPv6 Hitlist Service	Input	21073	87.1	3899910	56.5	97.2
	Responsive	18030	74.8	607863	8.8	54.7
	↪ ICMP	17863	74.1	593607	8.6	53.3
	↪ TCP/80	10098	41.9	91068	1.3	52.4
	↪ TCP/443	9623	39.9	91551	1.3	52.7
	↪ UDP/53	7060	29.3	36554	0.5	11.7
	↪ UDP/443	3358	13.9	12350	0.2	32.1
	Aliased	677	2.8	32178	0.5	51.4
	Aliased + Responsive	18169	75.4	637647	9.2	95.3
Toplists	1-day	5432	22.5	24906	0.4	86.3
	7-day	5826	24.2	27718	0.4	88.5
	30-day	6233	25.9	30537	0.4	88.8

the edge of their network, aggregated per hour and /48 prefix of the source address. The flow data is sampled and does not contain any information about flow targets. Nevertheless, it allows us to provide an overview of the coverage of the IPv6 Hitlist Service compared to real network traffic.

The dataset contains flows collected over three months (February – April 2026). It contains IPv6 flows from 24.1 k source ASes and 6.9 M /48 prefixes, with a daily median of 13 k observed ASes and 1.3 M /48 prefixes. The total number of prefixes is influenced by a few ASes, *i.e.,* 10 ASes already account for 54.7 % and 100 ASes account for 86.0 % of the /48 prefixes, respectively. A majority of these prefixes are only visible for 1–2 days, potentially indicating rotating prefix assignments in ISP networks. 58.9 k further ASes are visible in the flow data.

We compare this flow data to addresses present in the IPv6 Hitlist Service results from April 30, 2026 in three dimensions (ASes, prefixes, traffic covered) in Tab. 3. The cumulative input of the IPv6 Hitlist Service contains addresses originated by 21.1 k ASes (87.1 %) and 3.9 M of prefixes (56.5 %) visible in flow data. Furthermore, the prefixes with at least one address known to the IPv6 Hitlist Service account for 97.2 % of the traffic volume captured in the available flow data.

Comparing only responsive addresses to our flow dataset shows that there is at least one responsive address in 18 k (74.8 %) ASes, mapped to only 607.9 k /48 prefixes (8.8 %). Still these prefixes account for 54.7 % of traffic captured by the flow data. The coverage of protocol-specific responsive addresses follows an expected trend: while addresses responsive to ICMP cover the most ASes and prefixes, addresses responsive to TCP/80 and TCP/443 cover prefixes resulting in slightly more traffic. UDP/443 (the default port for HTTP/3 over QUIC) covers the fewest ASes and prefixes while still covering 32.1 % of traffic. As shown by Zirngibl et al. [49, 50], QUIC is mostly used by large CDNs.

We further map aliased prefixes identified by the IPv6 Hitlist Service to visible /48 prefixes in the flow data. Aliased prefixes are only identified in 677 ASes (2.8 %) with visible traffic and 32.2 k prefixes (0.5 %). Nevertheless, these prefixes still cover 51.4 % of traffic. This aligns with findings by Zirngibl et al. [51]: they showed

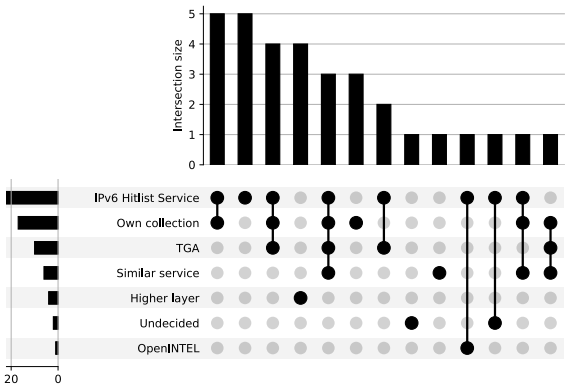

Figure 4: UpSet plot showing use of other address sources.

that aliased prefixes are often seen in large CDNs and are most likely fully responsive prefixes used for load-balancing, and are not aliases of a single host. Responsive and aliased prefix coverage can overlap due to the aggregation on /48 of flow traffic. In combination, aliased prefixes and responsive addresses cover 95.3 % of the traffic volume. This reinforces their suggestion to include addresses (e.g., at least one) of aliased prefixes for IPv6 scans.

To show the usefulness of cumulative address collection over time, we compare the IPv6 Hitlist Service coverage to the coverage of DNS resolutions of domain toplists, another technique commonly used for scans. Tab. 3 shows the coverage of toplist-based lists of IPv6 addresses. We use DNS resolutions of Google Crux [19], Cloudflare Radar [8], Majestic [34] and Cisco Umbrella [6] from 1, 7, and 30 days (daily resolution of up-to-date lists respectively). We do not test the responsiveness of addresses but evaluate the coverage based on IPv6 addresses from AAAA records. Toplist addresses are mapped to 30.5 k /48 prefixes that still are the source of 88.8 % of the traffic in bytes. Many domains found in toplists are hosted by major CDNs or hypergiants accounting for large traffic shares. However, compared to the IPv6 Hitlist Service, fewer /48 prefixes and especially ASes are covered. While toplists account for the top ASes, the long tail is easily missed by solely relying on DNS resolution of toplists.

In general, the evaluation shows good coverage of the IPv6 Hitlist Service compared to traffic visible at a major central European ISP and Tier-1 provider. At least one address is known to the IPv6 Hitlist Service for 87.1 % of ASes and 56.5 % of /48 prefixes originating IPv6 traffic over a 3-month period and cover prefixes originating 97.2 % of the IPv6 traffic.

4 User Survey

To better understand the current use of the IPv6 Hitlist Service, we conduct a survey (see Sec. B for details) and advertise it via networking mailing lists, Internet-measurement-related Slack and Mattermost channels, and on LinkedIn. We also reach out to 147 registered IPv6 Hitlist Service users via their registration emails. In total, we receive 33 responses to our survey in a 30-day period (February 18, 2026 – March 20, 2026).

IPv6 Measurements: We first ask survey participants, whether they are performing IPv6 measurements. 21 answered that they are

currently performing IPv6 measurements, 8 did so in the past, and 3 plan to do so in the future. Only one responded negatively to this introductory question.

Hitlist Use: We continue the remainder of the survey with the 32 responses which answered positively to the IPv6 measurement question. Next, we inquire whether the respondents are using the IPv6 Hitlist Service for their measurements. 22 responded that they are using the IPv6 Hitlist Service, 9 responded that they are not using it, and one responded that they may be using it in the future. Interestingly, we find that respondents who had performed IPv6 measurements in the past are more likely to make use of the IPv6 Hitlist Service (75 %) compared to respondents currently running IPv6 measurements (66 %). When asked about reasons of why they are not using the IPv6 Hitlist Service, they indicated that they were not aware of it, or it did not align with their needs.

IPv6 Address Sources: Next, we analyze the use of other IPv6 address sources. 17 respondents use their own IPv6 address collection, 10 indicate the use of TGAs, and 6 responses indicate the use of a similar service as the IPv6 Hitlist Service. Four respondents indicate that they do not need dedicated IPv6 address targets as they are running their IPv6 measurements on a higher layer (e.g., domains, URLs). Other respondents use OpenINTEL data [45], are not using any additional address sources or still need to decide. When looking at the combinations as shown in Fig. 4, we find that (1) using the hitlist or (2) using their own collection in combination with the hitlist are the two most common choices among respondents. Furthermore, four respondents each are using their own collection with TGAs and the hitlist or do not need any dedicated IPv6 addresses as their measurements are run on a higher layer. Some respondents also provide additional details on their IPv6 address target collection, stating that they collect addresses from various sources, *e.g.,* TLS certificates, toplists, TLD domain lists, /56 gateway addresses from LEO providers, running an NTP pool server, and traceroutes.

Hitlist Usage Patterns: To better grasp how the IPv6 Hitlist Service is being used, we ask about usage patterns in the survey. When it comes to duration, we find that 20 % use the hitlist one-off, 15 % use it for 1 month and another 15 % for 2–6 months, 10 % for 7–12 months, and 40 % for more than 1 year. Next, we ask all users about their frequency of use: 5 % use the service daily, 15 % weekly, 25 % monthly, 45 % less frequently, and 10 % as a one-off. As the hitlist is available to users either publicly or through a registration-first service, we ask respondents how many are registered for the IPv6 Hitlist Service. 16 respondents indicate that they are registered for the service, whereas 4 are not registered.

Hitlist Used Data: Next, we want to shed light on the IPv6 Hitlist Service data used by researchers. The service provides a publicly available list of responsive addresses and lists of aliased/non-aliased prefixes, which can be downloaded by everyone without registering an account (publicly available data). Additionally, after registration researchers from academic institutions get access to additional data, *i.e.,* all input addresses (including non-responsive addresses), responsive addresses by different TCP, UDP, and ICMPv6 protocols, and access to historical data (registration-first data). From the respondents, the majority (11) stated that they are using the publicly available as well as the registration-first data. Moreover, 5 respondents use only publicly available data, whereas the remaining 4 respondents use only registration-first data.

Of the 15 respondents using the registration-first data, the majority use multiple types of data: 4 use responsive addresses, all addresses, and aliased data; 3 use responsive and aliased data; 3 use only responsive data; 2 use all addresses and aliased prefixes; another 2 use all addresses and responsive addresses; and a single respondent was using only the list of all addresses. This underlines the heterogeneity of use cases for researchers using the service.

Measurement Types, Tools, and Aliasing: To better understand what types of IPv6 measurements researchers using the IPv6 Hitlist Service perform, we next ask questions about targeted device types, used measurement tools, and handling of aliased prefixes. Looking at the responses we find two equally important types of devices. 12 responses indicate mainly targeting servers (HTTP, DNS, etc.) and routers in their measurements. Clients on the other hand are targeted by 9 respondents. A few single respondents provided custom answers, highlighting that they are targeting exclusively DNS servers, a single address per /48 prefix, or all IPv6 hosts. Note that respondents could provide multiple answers.

In terms of tooling, the majority of respondents use the port-scanning tool ZMapv6 [52] or the application-layer scanning tool ZGrab2 [48], with 13 and 9 respondents, respectively. Moreover, traceroute is used by 6 respondents, scamper [32] by 5 respondents, yarrp6 [4] by 4 respondents, and Nmap [33] by 3 respondents. Finally, DoE-Hunter [10], ZDNS [27], MAnycastR [21], and XMap [30] each saw a single mention among survey participants. Note that respondents could provide multiple answers.

When asked about how survey respondents handle aliased prefixes in their measurements, the majority (8) rely on the provided lists by the IPv6 Hitlist Service, while 4 respondents use those lists in addition to running their own alias prefix detection measurements. 6 respondents do not filter out aliased prefixes at all.

Impact: Next, we try to assess the academic impact of the IPv6 Hitlist Service in terms of publications. We ask survey participants whether their research using the IPv6 Hitlist Service resulted in a research paper. 14 respondents indicated that this research resulted in research output in the form of a publication: for 9 respondents this was the case for multiple publications, while for 5 respondents this resulted in a single publication. For the remaining 6 respondents this has not resulted in a publication, of which 1 has a paper under submission and another 2 are planning a submission at the time of the survey. We also ask survey participants about their perceived importance of the IPv6 Hitlist Service to their own research, where 9 responded that it is very important, and they could not do their research without it; 11 participants responded that it is important as parts of their research rely on the IPv6 Hitlist Service. No participant responded that the hitlist was not important. When asked about whether the IPv6 Hitlist Service should be continued, 19 out of 20 respondents answered with yes, 1 answered with neutral.

These responses—in addition to the 240 citations for the IMC 2018 IPv6 Hitlist paper [15] as of May 2026—further underline the importance and impact that the IPv6 Hitlist Service has on IPv6 measurement research.

Other Feedback: Finally, we ask survey participants for additional feedback to better understand IPv6 Hitlist Service user needs. We receive several custom responses, with ideas such as providing an API to retrieve the data, categorizing IP addresses by responsiveness over time or likelihood of being responsive in scans, and providing subsets of the hitlist for router addresses, servers, CDNs, etc. We also receive several thank-you notes in this free-form feedback.

5 IPv6 Hitlist Use

Besides self-reported user behavior based on the survey, we also evaluate access patterns to the published data. Therefore, we collect access logs of the webserver used to provide the data for 120 days. We delete the logs after this study. During this period, the IPv6 Hitlist Service receives more than 91 k successful HTTP requests from 288 ASes (this might include crawlers and bots). Next, we break those requests down into registered users and access to the openly available data.

Over the period of data collection 17 users and 2 measurement services [1, 3] access results available via the registration-first service. Registered users access different file types indicating different use cases and potential follow-up scans. 4 users access all three types of files, 4 users access the output and aliased files while 1 user accesses the input and aliased files. The remaining 10 users focus on a single file type (3× input, 7× output). Thus, the response files are the most commonly accessed files (15×) followed by the aliased prefixes (9×) and input (8×). As part of the IPv6 Hitlist Service, continuous scans test responsiveness over different protocols (see Sec. 2). For those responsive IP addresses for different protocols we see different access patterns, *e.g.,* all protocols or only UDP/53 indicating interest in DNS specifically. Besides different interest in available results, access patterns over time differ. 2 users access the data on one day only, 11 users on 2 to 9 days over multiple weeks. 4 users and the 2 measurement services access the data once per week. One user downloaded the complete historical aliased prefix data going back all the way to 2018 within 3 days. This shows that the availability of historic data can be useful for different purposes. Furthermore, the available data granularity is actively used and indicates that the IPv6 Hitlist Service data is used in a focused and careful manner, *e.g.,* to reduce the scale of follow-up scans.

Next, we analyze access patterns of the publicly available IPv6 Hitlist Service data. As there is no user or registration information available for this data, we evaluate access based on an AS level in combination with user-agent information. Throughout the 120-day period, we see 653 distinct AS-user-agent pairs accessing the publicly available data. We discard individual IP addresses before conducting any analysis and do not store logs outside the scope of this study. While this does not necessarily allow us to attribute individual Internet measurement or research entities, *e.g.,* if multiple users access the IPv6 Hitlist Service from a cloud provider with the same user-agent, we are still able to analyze access patterns to the provided data. Similar to registered users, access to publicly available data shows different patterns. A majority of entities access the data on 5 days or fewer (609, 93 %). In contrast, 8 entities access the data daily throughout the 120-day period, *e.g.,* Shadowserver [39], Leitwert [29], or nxthdr [35], most likely seeding regular follow-up IPv6 scans. Besides registered academic users, measurement organizations frequently access the hitlist data over time.

6 Best Practices & Lessons Learned

Response Types: The IPv6 Hitlist Service focuses on addresses with services resulting in positive responses, *i.e.,* sending an ICMP Echo Reply, a TCP SYN/ACK or a valid answer via UDP. Thus, the lists of responsive addresses contain few routers that mostly send ICMP error messages or respond to specific services only, and few clients. Furthermore, due to scalability reasons, the IPv6 Hitlist Service focuses on specific ports and protocols (DNS and web). While the list of responsive addresses to ICMP probes contains hosts potentially hosting different services, follow-up scans for services should consider an initial scan with the cumulative input.

Cumulative Input: The input of the IPv6 Hitlist Service is a combination of all addresses collected by the service since 2018 independent of whether they have been responsive to the tested protocols once or not. While larger than the list of responsive addresses by multiple orders of magnitude, it is a valuable source for different goals. For example, it contains router addresses extracted from traceroutes conducted by RIPE Atlas [36], CAIDA Ark [5] and hitlist-specific yarrp scans. Even if they do not return positive responses during follow-up scans, the input conserves these historical lists and can be used for future scans. Furthermore, it might reveal changing address patterns or prefix rotation behavior of different networks. TGAs could improve their evaluation of address patterns in networks by considering not only currently responsive addresses but also historical patterns.

Aliased Prefixes: Aliased prefixes are an important aspect to consider when running IPv6 measurements. Often, they are completely filtered or ignored. However, both extremes impact results. On the one hand, ignoring aliased prefixes can easily skew address-based results or render scans infeasible. Our provided list of aliased prefixes is based on the addresses known to the IPv6 Hitlist Service. Research adding other sources or using provided addresses, *e.g.,* for TGAs, needs to check for additional aliased prefixes to correctly interpret results. On the other hand, removing aliased prefixes completely removes important parts of the Internet. Aliased prefixes correlate with /48 prefixes originating 51.4 % of traffic as seen in flow data (see Sec. 3). As discussed by Zirngibl et al. [51], aliased prefixes are not necessarily aliases used by a single host but often are fully responsive due to load-balancing artifacts from large CDNs. Addresses from these prefixes should be sampled and used for follow-up scans to allow a representative view of the Internet.

Providing a Measurement Service: Over the past 10 years, we have learned important lessons on running a measurement service. The initial version of the IPv6 Hitlist Collection, which ran from 2016 to 2018, only provided data to researchers upon request. When relaunching the service as the IPv6 Hitlist Service in 2018, we changed this model to a publicly available subset of the data, with raw data available after registration. This flexible model allows everyone (researchers, network operators, even companies) to quickly download and work with the data, while it still allows researchers access to raw data if needed. Furthermore, we can use registration

emails to reach out to users of the service to announce new measurement data, published papers, or even conduct surveys, as we have done for this paper. In general, we encourage researchers not only to publish data as a one-off dump for reproducibility purposes, but rather to consider running their measurements as a service to provide continuously updated data to fellow researchers. Even though the work needed to maintain a measurement service is non-negligible, as things constantly change on the Internet, this changing nature also means that the results from long-running measurements often contain interesting insights which remain hidden when only performing one-off measurements. Recently, we see a positive trend of more and more Internet measurement researchers adopting the model of running a measurement service accompanying their published paper [1, 3, 14, 22, 26, 28, 38, 44].

Acknowledgments

This work was partially funded by the Horizon Europe programme under projects GreenDIGIT (101131207) and SLICES-IP (101287692), by the German Research Foundation (DFG) under project SLICES-SUSTAINABILITY (566292327) and by the German Federal Ministry of Research, Technology and Space (BMFTR) under projects ALL#HANDS (02K25A026) and 6G-life (16KIS2414).

References

[1] Taha Albakour, Oliver Gasser, Robert Beverly, and Georgios Smaragdakis. 2021. Third Time's Not a Charm: Exploiting SNMPv3 for Router Fingerprinting. In *Proc. ACM Int. Measurement Conference (IMC)*. doi:10.1145/3487552.3487848

[2] APNIC LABS. [n. d.]. *IPv6 Capable Rate by country (%)*. https://stats.labs.apnic.net/ipv6/

[3] Florian Aschenbrenner, Tanya Shreedhar, Oliver Gasser, Nitinder Mohan, and Jörg Ott. 2021. From Single Lane to Highways: Analyzing the Adoption of Multipath TCP in the Internet. In *Proceedings of the IFIP Networking Conference*. doi:10.23919/IFIPNetworking52078.2021.9472785

[4] Robert Beverly. 2016. Yarrp'ing the Internet: Randomized High-Speed Active Topology Discovery. In *Proc. ACM Int. Measurement Conference (IMC)*. doi:10.1145/2987443.2987479

[5] CAIDA. 2025. *Archipelago (Ark) Measurement Infrastructure*. Retrieved 2025-10-26 from https://www.caida.org/projects/ark/

[6] Cisco. 2022. *Umbrella Top 1M List*. https://umbrella.cisco.com/blog/cisco-umbrella-1-million

[7] Cloudflare. [n. d.]. *Adoption & Usage*. https://radar.cloudflare.com/adoption-and-usage

[8] Cloudflare. 2023. *Cloudflare Radar*. https://radar.cloudflare.com/

[9] S. Deering and R. Hinden. 1995. *Internet Protocol, Version 6 (IPv6) Specification*. RFC 1883. IETF. https://www.rfc-editor.org/rfc/rfc1883.txt

[10] DoE-Hunter authors. 2026. DoE-Hunter on GitHub. https://github.com/steffsas/doe-hunter.

[11] Zakir Durumeric, Eric Wustrow, and J Alex Halderman. 2013. ZMap: Fast Internet-wide scanning and its security applications. In *22nd USENIX Security Symposium*.

[12] Mert Erdemir, Frank Li, and Paul Pearce. 2025. *Understanding IPv6 Aliases and Detection Methods*. doi:10.1007/978-3-031-85960-1_3

[13] Tobias Fiebig, Kevin Borgolte, Shuang Hao, Christopher Kruegel, and Giovanni Vigna. 2017. Something from Nothing (There): Collecting Global IPv6 Datasets from DNS. In *Proc. Passive and Active Measurement (PAM)*. doi:10.1007_978-3-319-54328-4_3

[14] Romain Fontugne, Malte Tashiro, Raffaele Sommese, Mattijs Jonker, Zachary S. Bischof, and Emile Aben. 2024. The Wisdom of the Measurement Crowd: Building the Internet Yellow Pages a Knowledge Graph for the Internet. In *Proc. ACM Int. Measurement Conference (IMC)*. doi:10.1145/3646547.3688444

[15] Oliver Gasser, Quirin Scheitle, Pawel Foremski, Qasim Lone, Maciej Korczyński, Stephen D Strowes, Luuk Hendriks, and Georg Carle. 2018. Clusters in the Expanse: Understanding and Unbiasing IPv6 Hitlists. In *Proc. ACM Int. Measurement Conference (IMC)*. doi:10.1145/3278532.3278564

[16] Oliver Gasser, Quirin Scheitle, Sebastian Gebhard, and Georg Carle. [n. d.]. *IPv6 Hitlist Collection*. https://www.net.in.tum.de/projects/gino/ipv6-hitlist.html

[17] Oliver Gasser, Quirin Scheitle, Sebastian Gebhard, and Georg Carle. 2016. Scanning the IPv6 Internet: Towards a Comprehensive Hitlist. In *Proc. Network Traffic Measurement and Analysis Conference (TMA)*.

[18] Google. [n. d.]. *Statistics: IPv6 Adoption*. https://www.google.com/intl/en/ipv6/statistics.html

[19] Google. 2026. *Chrome User Experience Report*. https://developer.chrome.com/docs/crux

[20] Robert Graham. [n. d.]. *Masscan*. https://github.com/robertdavidgraham/masscan

[21] Remi Hendriks, Matthew Luckie, Mattijs Jonker, Raffaele Sommese, and Roland van Rijswijk-Deij. 2025. LACeS: An Open, Fast, Responsible and Efficient Longitudinal Anycast Census System. In *Proc. ACM Int. Measurement Conference (IMC)*. doi:10.1145/3730567.3764484

[22] Remi Hendriks, Matthew Luckie, Mattijs Jonker, Raffaele Sommese, and Roland van Rijswijk-Deij. 2025. LACeS: An Open, Fast, Responsible and Efficient Longitudinal Anycast Census System. In *Proceedings of the 2025 ACM Internet Measurement Conference*. 445–461.

[23] Bingnan Hou, Zhiping Cai, Kui Wu, Tao Yang, and Tongqing Zhou. 2023. 6Scan: A High-Efficiency Dynamic Internet-Wide IPv6 Scanner With Regional Encoding. *IEEE/ACM Transactions on Networking* (2023). doi:10.1109/tnet.2023.3233953

[24] IPinfo. [n. d.]. *The trusted source for IP address data*. https://ipinfo.io/

[25] IPv6 Hitlist Authors. [n. d.]. *IPv6 Hitlist Website*. https://ipv6hitlist.github.io/

[26] Katherine Izhikevich, Ben Du, Sumanth Rao, Alisha Ukani, and Liz Izhikevich. 2024. Trust, But Verify, Operator-Reported Geolocation. *arXiv preprint arXiv:2409.19109* (2024).

[27] Liz Izhikevich, Gautam Akiwate, Briana Berger, Spencer Drakontaidis, Anna Ascheman, Paul Pearce, David Adrian, and Zakir Durumeric. 2022. ZDNS: A Fast DNS Toolkit for Internet Measurement. In *Proc. ACM Int. Measurement Conference (IMC)*. doi:10.1145/3517745.3561434

[28] Maynard Koch, Raphael Hiesgen, Marcin Nawrocki, Thomas C. Schmidt, and Matthias Wählisch. 2025. Scanning the IPv6 Internet Using Subnet-Router Anycast Probing. *Proc. ACM Netw.* (2025). https://doi.org/10.1145/3768997

[29] Leitwert. [n. d.]. *Tailored Internet Intelligence*. https://www.leitwert.net/

[30] Xiang Li, Zixuan Xie, Lu Sun, Yuqi Qiu, Zuyao Xu, and Zheli Liu. 2026. XMap: Fast Internet-wide IPv4 and IPv6 Network Scanner. *arXiv preprint arXiv:2602.09333* (2026).

[31] Zhizhu Liu, Yinqiao Xiong, Xin Liu, Wei Xie, and Peidong Zhu. 2019. 6Tree: Efficient dynamic discovery of active addresses in the IPv6 address space. *Computer Networks* (2019). doi:10.1016/j.comnet.2019.03.010

[32] Matthew Luckie. 2010. Scamper: a Scalable and Extensible Packet Prober for Active Measurement of the Internet. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*. doi:10.1145/1879141.1879171

[33] Gordon Fyodor Lyon. 2009. *Nmap network scanning: The official Nmap project guide to network discovery and security scanning*. Insecure.

[34] Majestic. 2022. *The Majestic Million*. https://majestic.com/reports/majestic-million/

[35] nxthdr. [n. d.]. *Internet Research Platform*. https://nxthdr.dev/

[36] RIPE. 2025. *RIPE Atlas*. Retrieved 2025-10-26 from https://atlas.ripe.net/

[37] Erik Rye and Dave Levin. 2023. IPv6 Hitlists at Scale: Be Careful What You Wish For. In *Proceedings of the ACM SIGCOMM 2023 Conference*. 904–916.

[38] Patrick Sattler, Johannes Zirngibl, Mattijs Jonker, Oliver Gasser, Georg Carle, and Ralph Holz. 2023. Packed to the Brim: Investigating the Impact of Highly Responsive Prefixes on Internet-wide Measurement Campaigns. *Proc. ACM Netw.* (2023). doi:10.1145/3629146

[39] Shadowserver. [n. d.]. *The Shadowserver Foundation*. https://www.shadowserver.org/

[40] Guanglei Song, Lin He, Zhiliang Wang, Jiahai Yang, Tao Jin, Jieling Liu, and Guo Li. 2020. Towards the Construction of Global IPv6 Hitlist and Efficient Probing of IPv6 Address Space. In *2020 IEEE/ACM 28th International Symposium on Quality of Service (IWQoS)*. doi:10.1109/IWQoS49365.2020.9212980

[41] Guanglei Song, Jiahai Yang, Lin He, Zhiliang Wang, Guo Li, Chenxin Duan, Yaozhong Liu, and Zhongxiang Sun. 2022. AddrMiner: A Comprehensive Global Active IPv6 Address Discovery System. In *2022 USENIX Annual Technical Conference (USENIX ATC 22)*.

[42] Lion Steger, Liming Kuang, Johannes Zirngibl, Georg Carle, and Oliver Gasser. 2023. Target Acquired? Evaluating Target Generation Algorithms for IPv6. In *Proc. Network Traffic Measurement and Analysis Conference (TMA)*. doi:10.23919/tma58422.2023.10199073

[43] Nik Sultana, Hyunsuk Bang, Elena Yulaeva, Ricky KP Mok, Kc Claffy, and Richard Mortier. 2025. A Survey on Packet Filtering. *ACM SIGCOMM Computer Communication Review* (2025). doi:10.1145/3711992.3711994

[44] Kedar Thiagarajan, Esteban Carisimo, and Fabián E. Bustamante. 2025. The Aleph: Decoding Geographic Information from DNS PTR Records Using Large Language Models. *Proc. ACM Netw.* (2025). doi:10.1145/3709374

[45] Roland van Rijswijk-Deij, Mattijs Jonker, Anna Sperotto, and Aiko Pras. 2016. A High-Performance, Scalable Infrastructure for Large-Scale Active DNS Measurements. *IEEE Journal on Selected Areas in Communications* (2016). doi:10.1109/JSAC.2016.2558918

[46] Grant Williams, Mert Erdemir, Amanda Hsu, Shraddha Bhat, Abhishek Bhaskar, Frank Li, and Paul Pearce. 2024. 6Sense: Internet-Wide IPv6 Scanning and its Security Applications. In *33rd USENIX Security Symposium*.

[47] Fan Gabriella Xue and Matthew Caesar. 2025. Towards Immersive Cloud-Based IoT Education. *ACM SIGCOMM Computer Communication Review* (2025). doi:10.

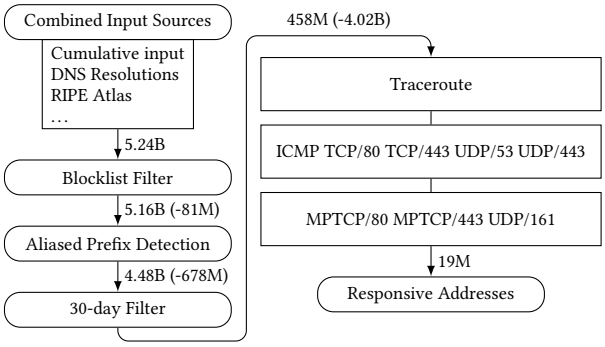

Figure 5: Pipeline of the IPv6 Hitlist Service.

1145/3711992.3711995
[48] ZGrab2 authors. 2026. ZGrab2 on GitHub. https://github.com/zmap/zgrab2.
[49] Johannes Zirngibl, Philippe Buschmann, Patrick Sattler, Benedikt Jaeger, Juliane Aulbach, and Georg Carle. 2021. It's over 9000: Analyzing early QUIC Deployments with the Standardization on the Horizon. In *Proc. ACM Int. Measurement Conference (IMC)*. doi:10.1145/3487552.3487826
[50] Johannes Zirngibl, Florian Gebauer, Patrick Sattler, Markus Sosnowski, and Georg Carle. 2024. QUIC Hunter: Finding QUIC Deployments and Identifying Server Libraries Across the Internet. In *Proc. Passive and Active Measurement (PAM)*. doi:10.1007/978-3-031-56252-5_13
[51] Johannes Zirngibl, Lion Steger, Patrick Sattler, Oliver Gasser, and Georg Carle. 2022. Rusty Clusters? Dusting an IPv6 Research Foundation. In *Proc. ACM Int. Measurement Conference (IMC)*. doi:10.1145/3517745.3561440
[52] ZMapv6 authors. 2025. ZMapv6 on GitHub. https://github.com/tumi8/zmap.

A Hitlist Pipeline

Fig. 5 provides an overview of the IPv6 Hitlist Service pipeline collecting input sources, applying filters and conducting scans. More details about the initial design can be found in [15, 51] and the current status and changes are discussed in Sec. 2.

B Survey Details

Similar to related work in Internet measurement [43, 47], we conduct a survey on IPv6 measurements and the use of the IPv6 Hitlist Service. We run the survey for a 30-day period (February 18, 2026 – March 20, 2026) and receive 33 responses in total. Below is the list of questions we asked in our survey. Note that for some questions we had several variants (*e.g.,* different tense), depending on previous questions; responses which are not present for all of these variants are denoted in parentheses. Questions which allow multiple responses are prepended with "Multi:", optional questions are prepended with "Optional:". Responses where participants could enter custom text are denoted as "*Other*" or "*Text*".

B.1 Are you performing IPv6 measurements?

- Yes, currently
- Yes, I did so in the past
- Yes, I plan to do so in the future
- No

B.2 Do/did you/do you plan to use the IPv6 Hitlist Service (ipv6hitlist.github.io) for your IPv6 measurements?

- Yes
- (Maybe)
- No

B.3 Multi: Which other source(s) for IPv6 addresses do/did you/do you plan to use?

- Similar service providing a list of IPv6 target lists
- I have my own collection of IPv6 targets
- I'm using a seed set and then generating new addresses with a target generation algorithm (TGA)
- I don't need dedicated IPv6 targets as I'm running my IPv6 measurements on a higher layer (e.g., domains, URLs)
- (I still need to decide)
- None
- *Other*

B.4 Optional: Optionally provide more details on your exact IPv6 address source(s) and how you use/used/ plan to use them.

- *Text*

B.5 Optional: Optionally provide details on why you do not/did not/do not plan to use the IPv6 Hitlist Service for your IPv6 measurements.

- *Text*

B.6 Over which total duration do/did you download data from the IPv6 Hitlist Service?

- One-off
- 1 month
- 2–6 months
- 7–12 months
- More than 12 months

B.7 How frequently do/did you download new data from the IPv6 Hitlist Service?

- Daily
- Weekly
- Monthly
- Less frequently than monthly
- One-off

B.8 Did you register to use the IPv6 Hitlist Service data?

- Yes
- No

B.9 Multi: Do/did you download publicly available data or data which is only available after registration for the IPv6 Hitlist Service?

- Publicly available data
- Data available after registration

B.10 Multi: Which data do/did you download from the IPv6 Hitlist Service?

- All collected addresses (also called "input")
- Responsive addresses (also called "output")
- Aliased prefixes
- Non-aliased prefixes

B.11 Multi: How do you filter aliased/highly-responsive prefixes from your IPv6 measurements?

- Using the list of aliased/non-aliased prefixes from the IPv6 Hitlist Service
- Running alias prefix detection measurements myself
- I do not filter them out

B.12 Multi: Which tool(s) do/did you use to run your IPv6 measurements?

- ZMapv6
- ZGrab2
- yarrp6
- nmap
- traceroute
- scamper
- *Other*

B.13 Which types of devices are you targeting in your IPv6 measurements?

- Routers
- Servers (HTTP, DNS, etc.)
- Clients
- *Other*

B.14 Did the measurements where you used the IPv6 Hitlist Service result in a peer-reviewed publication?

- Yes, more than one
- Yes, one
- Not yet, it is currently under submission
- No, but I'm planning a submission
- No, no submission is planned at this point

B.15 How important is/was data from the IPv6 Hitlist Service for your research?

- Very important, couldn't do it without it
- Important, parts of the research rely on it
- Not that important, I could do the research without the data

B.16 Should we continue to take efforts to keep the IPv6 Hitlist Service running?

- Yes, the service is important for IPv6 measurement research
- Neutral, it doesn't hurt to have the data available
- No, the service is not needed anymore today

B.17 Would you be willing to share IPv6 addresses/measurement data with the IPv6 Hitlist Service to further improve the service?

- Yes
- Maybe
- No

B.18 Optional: Any other feedback, suggestions, improvements you want to share about the IPv6 Hitlist Service?

- *Text*