Reasons Dynamic Addresses Change

Ramakrishna Padmanabhan
University of Maryland
ramapad@cs.umd.edu

Amogh Dhamdhere
CAIDA/UCSD
amogh@caida.org

Emile Aben
RIPE NCC
emile.aben@ripe.net

kc claffy
CAIDA/UCSD
kc@caida.org

Neil Spring
University of Maryland
nspring@cs.umd.edu

ABSTRACT

Applications often use IP addresses as end host identifiers based on the assumption that IP addresses do not change frequently, even when dynamically assigned. The validity of this assumption depends upon the duration of time that an IP address continues to be assigned to the same end host, and this duration in turn, depends upon the various causes that can induce the currently assigned IP address to change. In this work, we identify different causes that can lead to an address change and analyze their effect in ISPs around the world using data gathered from 3,038 RIPE Atlas probes hosted across 929 ASes and 156 countries across all 12 months of 2015. Our observations reveal information about ISP practices, outages, and dynamic address prefixes. For example, we found 20 ISPs around the world that periodically reassign addresses after a fixed period, typically a multiple of 24 hours. We also found that address changes are correlated with network and power outages occurring at customer premises equipment (CPE) devices. Furthermore, almost half of the address changes we observed on the same CPE were to an entirely different BGP-routed prefix.

1. INTRODUCTION

Academia and industry often rely on a simplifying assumption that IP addresses uniquely identify end-hosts [1, 8, 9, 11, 15, 16, 18, 25, 34, 36, 37, 40–42, 47]. This assumption allows researchers to track end host behavior over time [16, 18, 35], or to count participating users in peer-to-peer systems [9, 34, 37]. Many organizations create blacklists of suspicious IP addresses based on previously observed malicious traffic associated with those addresses [8, 11, 40, 41].

We seek to verify the assumption that even dynamic IPv4 addresses are reasonably static over the time scales of these measurements or malicious behaviors. As a first step toward validating this assumption, we have analyzed dynamic address assignments from a large set of customer premises equipment (CPE) devices to understand more about the events and agents associated with dynamic address changes. Though several studies have investigated dynamic address churn rates [2, 7, 13, 17, 19, 21, 48], only Maier et al. have attempted to attribute dynamic address changes to their cause [19], for a single ISP in one urban area.

Anecdotal evidence is in conflict: some may report that their address changes often, others that their address changes extremely rarely [43–46]. In private conversation, ISP operators have claimed that they change dynamic addresses frequently, others appear to do so rarely. Despite the potential for dynamic address changes, the DHCP protocol tries to preserve address assignments even for expired leases (section 4.3.1 of RFC 2131 [10]). We attempt to reconcile these conflicting experiences so that researchers can estimate the likely duration of dynamic address assignments in different regions and networks.

In this paper, we introduce a new dataset for detecting dynamic address changes: the RIPE Atlas connection-logs dataset. RIPE Atlas is a measurement infrastructure consisting of thousands of probes deployed in various ISPs all over the world. The RIPE Atlas dataset is uniquely suited to studying the events associated with dynamic address change. It provides visibility into not only host IP address changes, but also events occurring at the host around the same time that likely triggered the change because a RIPE Atlas probe makes continuous additional measurements of the network and reports when it last booted. Further, the dataset allows the study of dynamic address prefixes from which a CPE device is allocated successive addresses since it includes

IMC 2016, November 14 - 16, 2016, Santa Monica, CA, USA

ISBN 978-1-4503-4526-2/16/11…$15.00

DOI: http://dx.doi.org/10.1145/2987443.2987461

every IP address that was assigned to the CPE device over time.

We analyze 95 ISPs in the dataset that have at least 5 dynamically assigned probes and find 20 ISPs we label as *periodic* because these ISPs renumber many of their customers after they have held their address for a specific duration. This time limit varies across ISPs. Of 2272 dynamically assigned probes in the dataset, 193 (8.5%) change addresses periodically with a period of 24 hours, and 123 (5.4%) do so with a period of one week. We take a first step toward understanding address changes that are coincident with outages, identifying the probability of address changes given power and network outages: about 15% of RIPE Atlas probes with dynamic addresses can expect an address change with each outage. For some ISPs, this probability increases with outage duration, while others renumber upon any outage—even outages that are so short that they are likely CPE device reboots or reconnects. Surprisingly, when addresses change, they often also change to a different BGP prefix as well, and are not constrained to the same /24. Private communication with a large European ISP corroborated our inferences regarding that ISP's renumbering patterns: the ISP periodically renumbers every 24 hours, renumbers upon outages of any duration, and renumbers across prefixes.

Much of this paper describes the process of excluding false or irrelevant data to arrive at precise estimates of address lifetime and the cause of outages. Multi-homed and dual-stack probes can alternate the addresses that they use without any dynamic reassignment. Firmware updates can cause RIPE Atlas probes to reboot to install the update as a *result* of an address change, when it loses a connection to the central controller, rather than a reboot being a cause. Probes may sometimes move physically from one network to another. Although it may be possible to design a cleaner measurement methodology from scratch, we see this work as an example of the challenge of repurposing existing measurement data sets toward a problem not originally considered. We also believe that the RIPE Atlas deployment, while biased toward technical users like most measurement infrastructure, and biased toward European deployment, is sufficiently large and long-established to yield novel and reliable observations about dynamic addressing in practice.

2. BACKGROUND AND MOTIVATION

An IP address can be used to uniquely identify the end-host it is assigned to until the end-host's address changes for some reason. The duration of time that a dynamic IP address continues to be assigned to the same CPE device depends upon various causes that can induce the assigned IP address to change. In this section, we present techniques used for assigning dynamic addresses and the events and agents involved in dynamic address changes.

2.1 Dynamic Host Configuration Protocol

Dynamic addresses are often assigned to devices using the Dynamic Host Configuration Protocol (DHCP) [10]. DHCP issues an IP address to a host for a lease duration configured by the ISP. The host will try to renew the lease before it expires, typically half-way into the lease. However, whether the same IP address is renewed, or a different one is assigned, depends upon ISP policy. We speculate that the typical behavior of ISPs using DHCP is to renew the lease of the currently assigned IP address, since one of the stated design goals in the DHCP specification is that a DHCP client should be assigned the same address in response to each request, whenever possible. Thus, we typically only expect an ISP using DHCP, to change the address of a CPE, if something happens to prevent the CPE from renewing its lease (like an outage).

2.2 Point-to-Point Protocol

In some networks, end-hosts connect to an ISP using point-to-point links. For these networks, the Point-to-Point Protocol (PPP) first configures and establishes the point-to-point link [39]. Next, a Network Control Protocol (NCP) like the Internet Protocol Control Protocol (IPCP) configures IP addresses [20]. The PPP specification notes that the link will remain configured for communication until the link is actively closed down through network administrator intervention or when an inactivity timer expires.

2.3 Potential dynamic address change causes

Next, we identify the reasons dynamic addresses assigned using the above techniques could change. We classify the following categories of address change:

- **Changes after outages** If the client is disconnected or loses power long enough to fail to renew a DHCP lease, its address may be assigned to another; when it returns, it may then get a new address. We call such changes *outage-caused address changes.*
- **Changes after reboot/reconnect** While we expect addresses assigned through traditional DHCP to change only when the outage duration is long enough to prevent lease renewal, addresses assigned through PPP can change upon outages of any duration. Any reboot or network reconnect event could cause the client to forget its prior address and request a new one, or the state associated with a connection may be lost. We call such address changes *reboot-caused address changes.*
- **Administrative address changes** A purpose of dynamic address assignment is to allow reconfiguration of the network; it is possible that a reconfiguration of the DHCP server will force a change to the subnet on which the client lies. We expect such reassignment to be rare.

- **Periodic address changes** We observe that some ISPs limit the session length associated with an address, causing a reassignment after a fixed duration, typically one day to one week depending on the ISP.

Intuitively, the address change is either caused by the ISP (administrative or periodic), or caused by the client (or an interruption in network service to the client) in a reboot or outage.

3. THE RIPE ATLAS DATASETS

Analyzing periodic and administrative address changes requires visibility of the dynamic addresses assigned to a sample of the ISP's customers and the ability to see these addresses change over time. Analyzing outage-caused and reboot-caused address changes requires knowledge of the events occurring on the end-host at the time of an address change. Prior studies of dynamic addressing have typically relied on incoming connections that have a unique client identifier, such as a user name, but changing addresses, and thus have no information about what caused a change or precisely when it occurred. The RIPE Atlas dataset is unique since it includes necessary information about both address changes and contemporaneous events at the host.

The RIPE NCC's Atlas project deploys small devices, called probes, that conduct measurements from globally distributed networks [27]. In this section, we first describe the connection logs dataset from RIPE Atlas that we use to detect IP address changes. We then describe the k-root ping and SOS-uptime datasets from RIPE Atlas that we use to learn about events occurring on end-hosts.

3.1 RIPE Atlas connection logs dataset

RIPE Atlas probes connect to the RIPE Atlas infrastructure through a single SSH session over TCP port 443 (typically used by HTTPS) [3]. RIPE Atlas servers record the establishment and termination of these connections in *connection logs.* Table 1 shows connection log entries for a RIPE Atlas probe in the dataset for the first five days in January 2015.

Connection logs record each TCP connection made by the probe to a central controller and include the timestamp of the beginning and end of the connection (defined by the last receipt of data), the peer address of the connection that represents the publicly visible IP address used by the probe, and a unique identifier of the probe device. Probes are typically deployed behind the Customer Premise Equipment (CPE) of a user, so that the publicly visible IP address appearing in the connection logs belongs to that of the CPE. We term this address the "probe's address" or the "end-host address," since it is the useful, publicly visible address that the probe uses, even though the address may technically belong to the CPE and the probe has a different, private, RFC 1918 address.

ID	Start time	End time	IP Address	Dur
206	Dec 31 03:21:34	Jan 1 02:57:37	91.55.174.103	NA
206	Jan 1 03:22:16	Jan 1 17:34:11	91.55.169.37	14.2
206	Jan 1 18:00:54	Jan 1 18:42:31	91.55.132.252	0.7
206	Jan 1 19:06:46	Jan 2 02:19:16	91.55.155.115	7.2
206	Jan 2 02:41:55	Jan 3 02:18:00	91.55.141.95	23.6
206	Jan 3 02:43:14	Jan 4 02:16:59	91.55.165.167	23.6
206	Jan 4 02:40:58	Jan 5 02:15:45	91.55.163.252	23.6
206	Jan 5 02:38:39	Jan 6 02:14:48	91.55.141.63	NA

Table 1: Connection log sample for the first five days of 2015. We compute the address duration, shown in the last column in hours.

We find IP address changes by inspecting these connection logs. A new entry in a probe's connection log is created whenever an event occurs that causes the existing TCP connection to break. This connection will break when the probe's IP address changes, when a probe reboots, or when there is an outage. We can infer that the address changed between the end time of one connection and the start time of the next, if the addresses differ in consecutive entries. For example, in Table 1, there are seven address changes. Between changes, we can identify the duration that the probe held an address, shown in hours. In this example, each connection had a different address, so the address durations are equal to the connection duration, though this is not always the case. The duration of the first address is unknown because we do not know when that IP address was first assigned to the probe; the duration of the last address is also unknown.

The interval between connections, in the example of Table 1, typically 20–25 minutes, is information we also use in concert with other datasets described below to determine the type and duration of the event that led to a new connection. An active RIPE Atlas probe should report experiments back to the central controller about every three minutes [14]. We attribute this long delay between the end of one connection to the beginning of the next when there is an address change to waiting for TCP to exhaust its retransmission attempts (RFC 1122 Section 4.2.3.5) [4].

We obtained connection logs from January 1, 2015 to December 31, 2015 belonging to 10,977 active RIPE Atlas probes that had been connected to their central controllers for more than 30 days in 2015. We first found the list of active probes as of December 31, 2015, using the RIPE Atlas probe archive [31], and found 16,584 active probes. Next, we scraped each active probe's connection logs directly from the probe's webpage [30]. Subsequently, we found 10,977 probes who had been connected to their central controllers for an aggregate duration of more than 30 days in 2015.

3.2 Probe filtering

We omit from our analysis two sets of data: probes that are connected using a method where using different addresses does not indicate changes to the addresses that were assigned, for example, multihomed probes, as well as connection log entries that represent movement from one location or provider to another. Once we omit a probe for anomalous behavior in connection logs, we omit that probe from our analysis of the other RIPE Atlas datasets as well.

Category	Probes
Total Probes	10,977
Not Analyzable	
Never changed	3,073
Dual Stack	3,728
IPv6	237
Multihomed / Core / Data-center (tags)	174
Multihomed (alternating addresses)	511
Only address change from 193.0.0.78	216
Analyzable (geography)	**3,038**
Multiple ASes	766
Analyzable (AS-level)	**2,272**

Table 2: Of the 10,977 probes in the dataset, we are able to find address changes on 3,038 probes. 766 probes had addresses from multiple ASes; we discard address changes across ASes for these probes from our geographic analysis and filter these probes altogether in our AS-level analysis.

Table 2 provides an overview of the probes we omitted from the analysis.

IPv6 and dual-stacked probes

Probes that communicate, even occasionally, over IPv6 are not useful for understanding IPv4 address dynamics. We found 237 probes that made connections solely over IPv6 and 3,728 that used both IPv4 and IPv6. The 3,728 that connect over both protocols often alternated between address types, providing little information about the duration that the probe held any particular IPv4 address. Concretely, if a dual-stacked probe established one TCP connection to the central controller over IPv4 and the next TCP connection over IPv6, we cannot tell whether or when the IPv4 address changed while the IPv6 connection was active. We would need consecutive IPv4 connections from three different IPv4 addresses to determine how long the probe held the address in the middle of the sequence. In practice, a sequence of such IPv4 connections is rare for a dual-stack probe.

Multihomed and datacenter probes

We cannot use the connection logs dataset to observe address changes accurately on multihomed probes (probes that have more than one available IP address concurrently). For these probes, a connection from a new address could simply be a connection from the other address assigned to the CPE, much like a dual-stack probe. Probes at exchange points or in data centers are relatively few and seemed more likely to be problematic (by exhibiting multihomed behavior) than instructive (by representing address changes experienced by customers).

To filter multihomed probes, we first looked for hints in user-provided "tags" associated with a probe: 174 probes had at least one of the tags "multihomed," "datacentre," or "core." Tags are provided voluntarily and so probes may not be tagged with those labels even if they were in fact multihomed; thus, we looked for common features among the tagged probes which we could then use to omit probes with similar behavior. The most common feature we found was that connections from the tagged probes alternated between one fixed address and another potentially changing address; we found this feature on 36 of the 174 tagged probes. We found 511 other probes that matched this behavior and removed them from the dataset. We expect that it is far more likely that when a host returns to using a previously-used address, the host is choosing from among addresses it holds for a long time rather than that the ISP reassigned a previously held address to the host. We combine this behavioral, alternating-addresses, definition of multihomed with the tags to choose probes to omit from analysis.

3.3 Connection log entry filtering

We omit some entries in the connection log because of properties of either the address involved or because the detected address change was such that a probe reported an address from one autonomous system for one connection and an address from a different autonomous system for the next connection. Removing these connection log entries does not generally remove probes entirely from analysis.

Testing addresses

Some probes had their first address transition from the same IP address, 193.0.0.78. This address belongs to the RIPE NCC, and was used for testing before being shipped to volunteers. There were 427 such probes that started with this address; we remove this connection log entry. That left 216 additional probes with no further address changes in 2015, so we omitted those probes in Table 2.

Address changes across ASes

When attributing behavior to individual autonomous systems, we omit from analysis any probes where address changes indicated a change from the address space of one autonomous system to the address space of another. We used CAIDA's IP-to-AS dataset [6] to map each IP address to its autonomous system. CAIDA publishes the IP-to-AS dataset monthly; thus, we found the month in which a new IP address was assigned to a probe and used CAIDA's IP-to-AS dataset for that month to find the AS for that address. We found 766 probes with at least one address change spanning different autonomous systems. These ASes could be sibling ASes owned by the same ISP, but could also belong to different ISPs if the owner of the probes switched ISPs. For our geographic analysis (Section 4.2), we discarded the address changes spanning ASes for these probes, but retained the address changes within the same AS.

For our AS-level analysis of renumbering behavior (Section 4.3), we made the conservative choice of filtering these probes altogether.

Table 2 summarizes the dataset and the number of probes filtered. After the filtering process we had 2,272 probes analyzable for AS-level renumbering behavior, and 3,038 probes analyzable for geographic renumbering behavior. For each analyzable probe in Table 2, we found address changes along with the time of the address change and used them to find the duration for which addresses were assigned before changing.

3.4 k-root ping dataset

ID	Timestamp	N sent	N success	LTS
16893	Jan 27 09:01:42	3	3	86
16893	Jan 27 09:05:48	3	0	151
16893	Jan 27 09:09:45	3	0	388
16893	Jan 27 09:13:36	3	0	619
16893	Jan 27 09:17:49	3	0	872
16893	Jan 27 09:21:40	3	0	1103
16893	Jan 27 09:25:39	3	3	1342
16893	Jan 27 09:29:36	3	3	146

Table 3: Sample of k-root ping dataset for probe ID 16893 when a network outage occurred. We detect a network outage when pings to the k-root server are lost and when this ping loss is accompanied by increasing Last Time Synchronized (LTS) values. Here we detect a network outage beginning at Jan 27 09:05:48 and ending at Jan 27 09:21:40.

We detect network outages using two items from the built-in RIPE Atlas probe measurements. Every four minutes, each probe sends three pings to the k-root DNS server and logs the number of sent pings and the number of successful responses [29]. Table 3 shows a sample of this log. Probes report the results of these and other measurements via HTTP POST to the central controller once every four minutes. Along with the measurement data, the probe also reports the current *LTS* or “last time synchronised” value. This value indicates when the probe last synchronized its clock with that of the central controller. Typically, probes synchronize their clocks by NTP or upon receipt of the HTTP verify response from the controller [14], so in the absence of an outage, the reported LTS value should be less than four minutes (240 seconds).

We use a combination of the ping responses and the LTS value to infer a network outage, so that we have two (mostly) independent measurements that indicate that the probe's network has failed. We consider the network outage to start at the first measurement where all pings to the k-root server were lost, and to end at the last measurement where all pings were lost. If the LTS value did not grow, that would indicate that the probe was still able to communicate with the controller, and thus would not be an outage. Note that this interval underestimates the duration of a network outage by up to eight minutes.

3.5 SOS-uptime dataset

ID	Timestamp	Uptime counter value
206	Jan 1 03:15:18	262531
206	Jan 1 17:50:26	315038
206	Jan 1 17:50:55	19
206	Jan 1 17:53:59	203
206	Jan 1 18:59:44	4147

Table 4: Sample of SOS-uptime records from RIPE Atlas for January 1 2015 for probe ID 206. The third row shows that the uptime counter had reset 19 seconds before 17:50:55, allowing us to infer that the probe rebooted at 17:50:36.

The SOS-uptime dataset contains probe uptime counter values over time. The uptime counter on each probe is 64 bits long and counts the number of seconds since the probe booted. Probes report their uptime counter value to the central controller every time they make a new TCP connection to the controller.

We use the SOS-uptime dataset to determine when RIPE Atlas probes rebooted by finding when the uptime counter was reset. For example, consider the sample SOS-uptime records from the RIPE Atlas dataset for probe ID 206 shown in Table 4. The first entry at 03:15:18 on January 1st shows that the probe had been up for 262,531 seconds. Later that evening, the probe is shown to have been up for 315,038 seconds, but the next uptime counter value reports that the probe was up for only 19 seconds. We infer that a reboot occurred 19 seconds earlier, at 17:50:36.

After finding reboot times, we use the k-root ping dataset to measure how long each power outage lasted. When we detect a reboot, we use the difference in time between successive pings to the k-root server to estimate the power outage duration.

3.6 Associating inter-connection gaps with outage events

The next task is to synthesize these three datasets to identify outage events that occur between TCP connections to the central controller. The TCP connection to the central controller breaks when the IP address changes, when the probe reboots, when the CPE reboots, or when there is a power outage or significant network outage. For example, the reboot at 17:50:36 in Table 1 corresponds to rows 2 and 3 in Table 1 since the reboot time falls between the end of the connection log entry ending at 17:34:11 and the start of the connection log entry beginning at 18:00:54.

We use a priority ordering to assign outages to inter-connection gaps. If the k-root dataset indicated a network outage in the gap, we associate it with a network outage. If instead the SOS-uptime dataset indicates a reboot coincident with missing attempted k-root pings from the k-root dataset, we associate the gap with a power outage. If neither occurred, we mark the gap as a "no-outage" indicating that the reconnection was not associated with any outage.

4. PERIODIC ADDRESS CHANGES

ISPs can assign dynamic addresses for as long as they wish. In DHCP, long leases simplify administration, while short leases can be more efficient in reclaiming unused addresses. DHCP leases, however, are meant to be renewable by devices that are still active. In this section, we look at periodic address reassignment: instances where a device changes address periodically, despite actively using the address. Periodic reassignment is atypical for devices using DHCP since a device that is continuously renewing its lease should continue to keep its current address [10].

4.1 Metric to detect periodic address durations

If ISPs intentionally renumber after specific durations, we would expect those address durations to be prominent in a distribution of all address durations belonging to that ISP. We initially considered studying distributions of raw address durations, similar to the analyses by Maier et al. [19] and Moura et al. [21], but found that short address-durations were overrepresented. For example, in Table 1, inspecting the cumulative distribution of address durations would suggest that only half the addresses (3 of 6) were assigned for 24 hours. However, when trying to reason about the expected duration that an address will continue to be assigned to the CPE, we would like to know the fraction of total time that each duration accounted for. For example, in Table 1, the CPE was assigned 24 hour long addresses for roughly three-quarters of the total measured time. This latter notion is more useful to find whether an ISP is using periodic durations consistently, since the modes at intervals on the scale of days will be more visible.

To capture this notion we define a metric, the *total time fraction*. For a given probe and an address duration d, we define the total time fraction for d as the fraction of time spent by the probe in durations of length d. We compute the total time fraction for a given probe and a duration d by obtaining the total address time for the probe, and computing the fraction of the total address time that was accounted for by address durations of length d. For a probe p, if $n(d)$ is the number of times the probe had an address duration d and D is an array containing all address durations that were assigned to the probe, the total time fraction for the address duration d is given by:

$f_d^p = d \times n(d)/\Sigma(D)$

We use a similar procedure for computing the total time fraction considering all probes in an ISP, country, or continent. We believe that the total time fraction offers a better representation of the probability that an address was assigned for a certain duration than a simple inspection of the address durations.

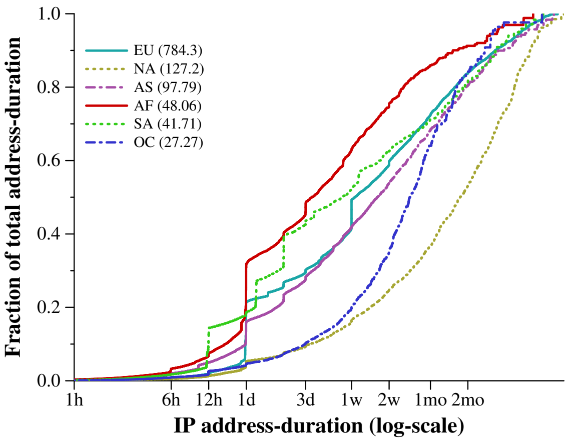

Figure 1: Cumulative distribution of total time fraction by continent. Modes (vertical segments in the CDF) indicate periodic renumbering. Addresses in North America are relatively long lived and free of periodic renumbering.

4.2 Periodic address changes by geography

We begin by inspecting how address durations vary across continents. We expected that address scarcity might affect address durations, leading to longer durations in North America and shorter durations in Asia. We use RIPE Atlas's probe database to find the country to which each probe belongs. Next, we aggregate the address durations of probes by their respective countries and subsequently, to their continents. Figure 1 shows the cumulative distribution of the total time fraction for each continent, i.e., the y-axis shows the fraction of total address duration accounted for by durations less than the x-axis value. The number in parentheses in the legend for each continent shows the total address duration for that continent in years ($\Sigma(D)$).

In Europe, Asia, Africa, and South America, address durations exhibit well-defined modes, mostly at intervals that are multiples of 24 hours. The most common mode is exactly at 24 hours: the total time fraction for European addresses at 24 hours is 0.16, African addresses is also 0.16, and Asian addresses is 0.07. One week address durations are also common in Europe, with the total time fraction at 1 week equaling 0.08. South American addresses exhibit multiple modes: their total time fraction is 0.11 at 12 hours, 0.07 at 28 hours, 0.09 at 48 hours, and 0.03 at 192 hours (8 days).

The curves for North America and Oceania do not have well-defined modes, suggesting that ISPs in these continents do not periodically change addresses. Further, North American probes typically retain their dy-

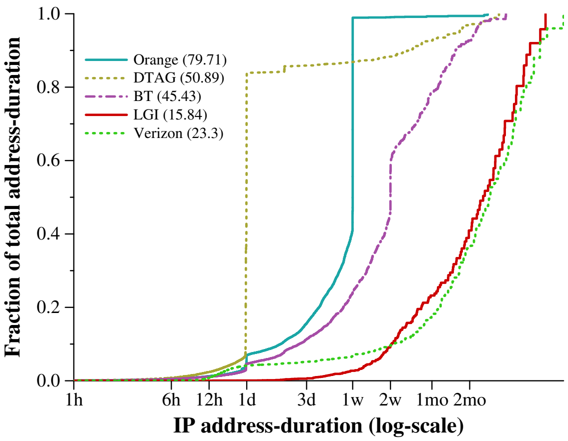

Figure 2: Cumulative distribution of total time fractions for ASes with most RIPE Atlas probes that yielded at least one address duration. Probes from Orange and DTAG spent more than half of their total duration in periodic durations of 1 week and 1 day respectively. BT also showed evidence of periodic renumbering with a mode at two weeks. On the other hand, LGI and Verizon have no modes at any durations, and spent most of their total time in durations that were weeks long.

namic addresses for much longer durations than other continents; North American addresses spent more than half of the total time in address durations longer than 50 days. This suggests that IP addresses can be used as end-host identifiers in North America for several weeks.

4.3 Periodic address changes by AS

We next considered whether the configuration decision to renumber periodically was uniform across an AS, or could reflect some other feature. For example, periodic renumbering could be a result of an unexpected cron job on the RIPE Atlas probe or a faulty DHCP client that could not renew. Periodic renumbering could be due to government regulations in countries, perhaps as a privacy measure. It could also simply reflect ISP policy, perhaps to hinder users from running web servers as anecdotal evidence suggests [12]. Investigating AS-level behavior can inform whether the periodic renumbering behavior is concentrated in some ASes and absent in others, shedding light on its potential cause.

4.3.1 Is periodic renumbering prevalent across all ISPs?

We first investigate the ASes with the largest deployment of RIPE Atlas probes where we detected at least two instances of address changes. Recall that we only obtain an address duration when the address began and ended during the interval we studied, so that a minimum of two address changes are necessary for a probe to yield an address duration. Figure 2 shows the cumulative distribution of total time fractions for the five autonomous systems with the most probes that yielded address durations. In this figure, Orange, an ISP from France, appears to change addresses after a duration of 168 hours (1 week): 55% of its total address duration was a week long. The German ISP, Deutsche Telekom AG (DTAG) reassigns addresses after 24 hours: 76% of the total address duration lies in that mode. British Telecom (BT) has a mode at 336 hours (2 weeks) with 13% of its total duration being in 2 week intervals. We study these ASes further in Section 4.4.

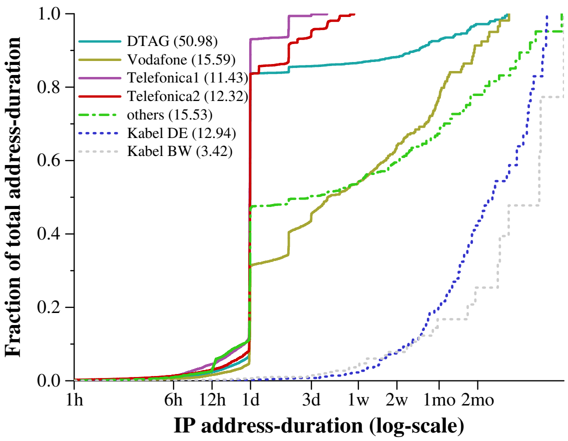

Figure 3: Cumulative distribution of total time fractions for ASes in Germany. Many German ISPs appear to change addresses every 24 hours. However, some ISPs have more stable addresses.

The other two ISPs do not exhibit any evidence of periodic renumbering. Liberty Global, an ISP to which probes spread across Europe belong, does not appear to change addresses periodically and neither does Verizon (US). Among these ASes, Verizon has the longest address durations.

Since periodic renumbering behavior is widespread in some ISPs and non-existent in others, we conclude that the cause of periodic renumbering is likely ISP policy.

4.3.2 Is periodic renumbering geographically correlated?

Next, we investigate how the periodic renumbering behavior of ISPs correlates with the country in which they operate. Germany has more than a hundred RIPE Atlas probes deployed across several ISPs, thus we study their address durations in Figure 3 for ISPs with probes that contributed at least 3 years of total time. Many ISPs in Germany change addresses every 24 hours: 77% of the duration in DTAG (AS 3320), 76% in Telefonica1 (AS 6805), 74% in Telefonica2 (AS 13184), and 29% in Vodafone (AS 3209), is 24 hours. We observe that the ’other’ ISPs also have a mode at 24 hours, suggesting that German ISPs are particularly likely to renumber every 24 hours. However, this behavior is not universal: Kabel Deutschland (AS 31334) and Kabel BW (AS29562) do not exhibit a mode at 24 hours; instead, more than 90% of their total address duration was spent in durations longer than two weeks.

These results suggest that periodic renumbering behavior can exhibit some geographic correlation, but is likely largely caused by ISP policy.

Private communication with a large European ISP confirmed that the ISP renumbers every 24 hours, since the ISP considers this scheme to be more 'privacy secure' although there is no government regulation that forces this feature. The ISP also reported that it uses PPPoE instead of DHCP for its DSL lines (which accounted for the vast majority of its customers). Since periodic behavior would be atypical of DHCP but consistent with PPP techniques for address assignment, we speculate that periodic renumbering is a property of ISPs that use PPP.

4.4 ISPs that renumber periodically

In this section, we look specifically at ISPs that renumber periodically to infer the period over which they renumber, the fraction of the ISPs' probes which periodically renumber, how reliably the renumbering occurs at the end of the period, and whether the renumbering is synchronized across probes. We classify a probe as "periodic" when its total time fraction at some duration d exceeds 0.25. We set the threshold to 0.25 because we expect a probe whose address is reassigned periodically to sometimes have a shorter duration, say, due to a reboot, and sometimes have a longer duration, say, by receiving the same address again.

We consider autonomous systems having at least five probes with an address change of which at least three probes are periodic, and provide an overview of their renumbering period and behavior in Table 5. The periodic duration d is shown in hours; 24 hour durations are typical. Renumbering in this table is primarily a feature of central Europe, with some in Russia, Kazakhstan, Mauritius, and South America. We describe the rest of the columns in the next subsections.

4.4.1 *What fraction of probes is periodic?*

Even for ISPs such as Orange and DTAG which have total time fraction at period d in excess of 0.5, not all address durations equal d; some durations are shorter and others longer as seen in Figure 2. One possible explanation is that only a few probes in these ISPs were periodically renumbered while others were not. Alternately, periodic probes sometimes have address durations not equal to d. We find that it is usually a combination of both factors that lead to non-periodic durations in these ISPs, although the extent to which each is responsible varies by ISP.

In Table 5, the N column shows the number of probes with at least one address change in the dataset. The next column, $f_d^p > 0.25$, shows the number of periodic probes—those having a time fraction of more than 0.25 at duration d. In some ISPs, only a small fraction of probes are periodically renumbered. For example, only a fifth of the probes in BT were periodic with a 2-week period, partially explaining why the total time fraction at 2-weeks for BT in Figure 2 is only 0.13.

The subsequent columns, $f_d^p > 0.5$ and $f_d^p > 0.75$ show what percentage of the periodic probes are persistently so, where the total time fraction at duration d is more than half or three quarters. We show percentages rather than raw counts in these columns to simplify the comparison, given that these providers have different sizes. A high percentage indicates that most of the periodic probes (with $f_d^p > 0.25$), are strongly so ($f_d^p > 0.75$). A low percentage indicates that probes may either be reassigned early (due to outages) or late (due to inconsistent reassignment). We can see that only 15% of the periodic probes in BT had $f_d^p > 0.5$ and none had $f_d^p > 0.75$, providing further explanation for why the total time fraction at 2-weeks for BT is low.

Other ISPs have a much larger fraction of their probes that are periodic: more than 80% of probes in Orange, DTAG, Telefonica Germany, A1 Telekom, Hrvatski, ISKON, ANTEL, Global Village Telecom, Mauritius Telekom, Orange Polska, and Digi Tavkozlesi are periodic. For each of these ISPs, more than 75% of probes are persistently periodic, having $f_d^p > 0.5$. For DTAG, Telefonica, A1 Telekom, Hrvatski, ANTEL, and Orange Polska, more than 75% of probes have $f_d^p > 0.75$. Notable is Orange Polska, which has four of its ten probes periodic at 24 hours, and five more probes periodic at 22 hours, but 100% of them have a time fraction at their respective durations greater than 0.75.

Probes in these ISPs typically have address durations capped at d. Address durations can sometimes be shorter—potentially due to outages or reboot/reconnect events as we show in Section 5—but can occasionally be larger than d as well. We study these next.

4.4.2 *Why are some address durations longer than the period?*

Some address durations exceed the typical period, d, for an ISP. We would like to determine whether this is a behavior limited to a few probes in the ISP (potentially caused by unusually designed CPE devices), or if the longer-than-typical durations are spread across probes.

How many periodic probes have an address duration longer than d? We expected that no address duration for such probes would exceed the periodic duration d. That is, if the ISP was renumbering a probe on a schedule, then some additional renumbering would be possible due to other reasons, but the probe would never keep its address longer than d. It turns out that this expectation is not the case. The column $MAX \leq d$ shows the percentage of the periodic probes that had their maximum address duration less than d (to capture only those durations that clearly exceeded d, we adjusted d to be d + 5% for this column). Across all periodic probes, 94% of those that appear to be on a

AS	**ASN**	**Country**	d	**N**	$f_d^p > 0.25$	$f_d^p > 0.5$	$f_d^p > 0.75$	$MAX \leq d$	**Harmonic**
All			24	2272	193	88.6%	68.4%	43.5%	89.6%
All			168	2272	123	74.0%	13.8%	94.3%	98.4%
Orange	3215	France	168	122	111	77%	14%	98%	99%
DTAG	3320	Germany	24	63	51	96%	86%	78%	98%
Telefonica DE 2	6805	Germany	24	17	15	93%	80%	27%	93%
Telefonica DE 1	13184	Germany	24	14	14	93%	86%	21%	100%
PJSC Rostelecom	8997	Russia	24	22	13	100%	69%	23%	100%
BT	2856	U.K.	337	67	13	15%	0%	38%	62%
Proximus	5432	Belgium	36	41	12	83%	8%	0%	83%
A1 Telekom	8447	Austria	24	12	11	100%	91%	73%	100%
Vodafone GmbH	3209	Germany	24	21	9	78%	11%	0%	89%
Hrvatski	5391	Croatia	24	7	7	100%	100%	43%	86%
ISKON	13046	Croatia	24	6	6	83%	33%	0%	100%
ANTEL	6057	Uruguay	12	6	6	100%	100%	33%	100%
Global Village Telecom	18881	Brazil	48	6	6	100%	67%	0%	17%
Mauritius Telecom	23889	Mauritius	24	6	5	100%	20%	20%	100%
JSC Kazakhtelecom	9198	Kazakhstan	24	15	5	80%	80%	60%	80%
Orange Polska	5617	Poland	22	10	5	100%	100%	60%	80%
VIPnet	31012	Croatia	92	7	4	75%	0%	75%	75%
Proximus	5432	Belgium	24	41	4	50%	25%	0%	75%
Digi Tavkozlesi	20845	Hungary	168	4	4	100%	25%	100%	100%
Orange Polska	5617	Poland	24	10	4	100%	100%	50%	100%
Free SAS	12322	France	24	12	3	100%	67%	0%	67%
SONATEL-AS	8346	Europe	24	7	3	33%	33%	33%	33%
Net by Net	12714	Russia	47	7	3	100%	100%	67%	100%

Table 5: Autonomous systems that had at least three probes with a total time fraction for duration d (in hours) greater than 0.25. $f_d^p > 0.25$ shows the number of probes that had a total time fraction at d greater than 0.25; $f_d^p > 0.50$ and $f_d^p > 0.75$ show the percentage of those probes that had fractions greater than 0.5 and 0.75 for the same duration. $MAX \leq d$ shows the percentage of probes whose maximum duration was no greater than d. "Harmonic" represents the percentage of probes that, if not renumbered after d, are renumbered after some multiple of d hours. The ASes are sorted in decreasing order of $f_d^p > 0.25$.

one-week renumbering schedule did not have an address duration longer than one week; only 44% of those that appeared to be on a one-day renumbering schedule had all durations limited by twenty-four hours.

This fraction seemed surprisingly low. Why would so many probes show daily renumbering, even reporting a total time fraction of 0.75, when the probe might also keep its address longer? We considered two possible explanations that would have the same symptoms: that a periodic renumbering was skipped or that the same address was (perhaps by random chance) assigned again. In these cases, rather than see an address change after 24 hours, we might see one at 48 or even 72 hours. We term such address changes "Harmonics", and consider what fraction of the time all address changes are at or before d (as expected), or occur at a multiple of d. The percentage of probes that match this loosened definition (a superset of those in $MAX \leq d$) appears in the last column of Table 5. Most periodic probes from all ISPs except Global Village Telecom and SONATEL-AS have maximum durations of this kind.

4.4.3 Are changes synchronized?

We imagine two broad strategies for daily renumbering: either leaving each customer on an independent, free-running clock that resets after 24 hours, or synchronizing all address changes to an off-peak time when few would be interrupted. Both seem reasonable strategies: independent clocks seem simple to implement, synchronized address changes seem more likely to shuffle addresses since many addresses are made available during the synchronized interval. However, if one were to blacklist addresses for misbehavior, knowing which strategy is in use would help to choose for how long to keep the blacklist entry. We expect that plotting the time of day at which addresses change for each ISP will expose whether the renumbering is synchronized.

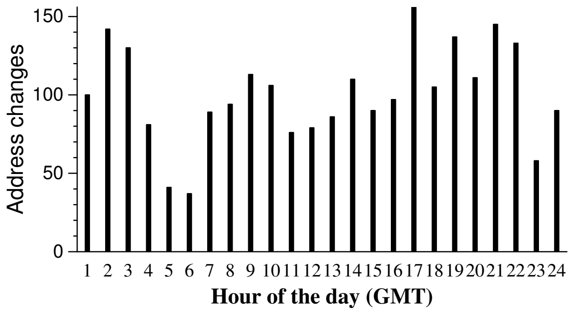

Figure 4: Periodic address changes in Orange appear more evenly distributed among the hours of the day.

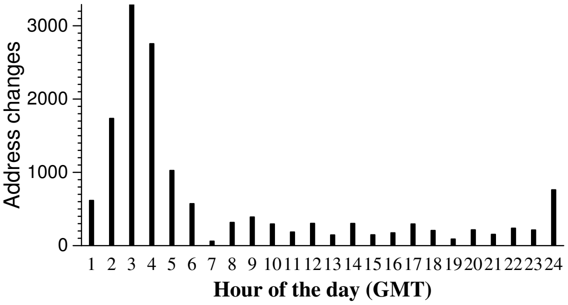

Figure 5: Periodic address changes are more likely in some hours for Deutsche Telekom.

For Orange and DTAG, the two ISPs with the most periodic probes, we choose the hour of the day in which every address duration that had duration d ended and show these in Figure 4 and Figure 5. For Orange, periodic address changes are not concentrated during any specific hours of the day. However, DTAG assigns periodic durations more often during some hours of the day. In private correspondence with a large European ISP, we learned that many CPE devices come with an option to choose the time at which they should disconnect and reconnect to receive a new address, as a privacy feature. Figure 5 supports this deployment scenario, observing almost three quarters of all periodic address changes between hours 24 to 6 (in GMT). However, some CPEs do not have this feature because a quarter of the periodic address changes happen at other hours of the day.

5. OUTAGE-CAUSED ADDRESS CHANGES

In Section 4.4, we saw that even probes from ISPs that renumber periodically often have durations shorter than the typical period. In this section, we study another potential cause of address change: outages occurring at the CPE (customer premises equipment), due to loss of power or network connectivity. Here, we quantify how frequently and for which probes an outage event at the CPE device appears to cause the reassignment of its IP address. If an outage event occurs at approximately the same time as an address change, we assume that the outage caused the address change. If an outage event occurs distant in time from an address change, then we assume that the outage did not cause an address change.

There are three versions of RIPE Atlas probes: v1,v2, and v3. More than 75% of probes are v3, although the distribution of versions within individual ISPs varies. We find network outage events on all versions of probes since network outages are by definition caused when a probe was up and reporting measurements. However, finding power outage events is confounded by the presence of potential false positives and negatives. We address these in detail next and describe our approach for filtering falsely inferred power outages.

5.1 Filtering falsely inferred power outages

The SOS-uptime data (Section 3.5) allows us to determine when the *probe rebooted.* Ideally, however, we would like to know *when the CPE rebooted.* Fortunately, probe reboots are often representative of CPE reboots due to a combination of how the RIPE NCC suggests that probes be installed [28] and expected fate sharing of co-located devices powered together, as we describe next.

The RIPE Atlas probe gets power from USB; because of this design, the probe can be powered by the USB port on the CPE and will be power-cycled whenever the CPE reboots. When the probe is plugged into the CPE, or both together are power-cycled, a probe reboot indicates that the CPE also rebooted. These represent the typical cases that are useful for the analysis of power outage related address changes. The potential error scenarios are as follows. When the CPE alone is rebooted but the probe is not, we would not observe a power outage, leading to a false negative. When the probe alone is rebooted but the CPE is not, we would detect a power outage, leading to a false positive. Although we expect probe reboots to be rare, a specific scenario in which they occur is when the probe receives a firmware upgrade. We discuss how to remove probe reboots due to firmware upgrades below in Section 5.2.

Older probe hardware (v1,v2) can also confound our inference of power outages, because these probes may reboot when they create new TCP connections, since they are vulnerable to memory fragmentation [33]. Address changes create new TCP connections and could induce such reboots, so for our power outage analysis we discard data from these older probes.

5.2 Removing reboots caused by firmware updates

The RIPE Atlas servers push firmware updates to probes simultaneously. When a probe's TCP connection to the central controller breaks, the probe will reboot and install the firmware update. Our goal is to filter reboots that were associated with a firmware update, since these reboots occur *as a result of* a dropped connection rather than as a cause. Figure 6 shows the number of unique probes that rebooted on each day of 2015. We observe five periods during the year when probes experienced more than twice as many reboots as the median for at least two consecutive days.

For each of these periods, we found the first day corresponding to the spike, and identify that day as when the firmware update was distributed. Some dates (April 14, July 6, October 5), agree precisely with documented RIPE Atlas firmware and UI updates [32]. Other dates are close—we observe March 23 instead of March 28, and January 25 instead of January 14—but nevertheless show the same spike in reboots. We then discard the first reboot for each probe that occurred after the firmware update.

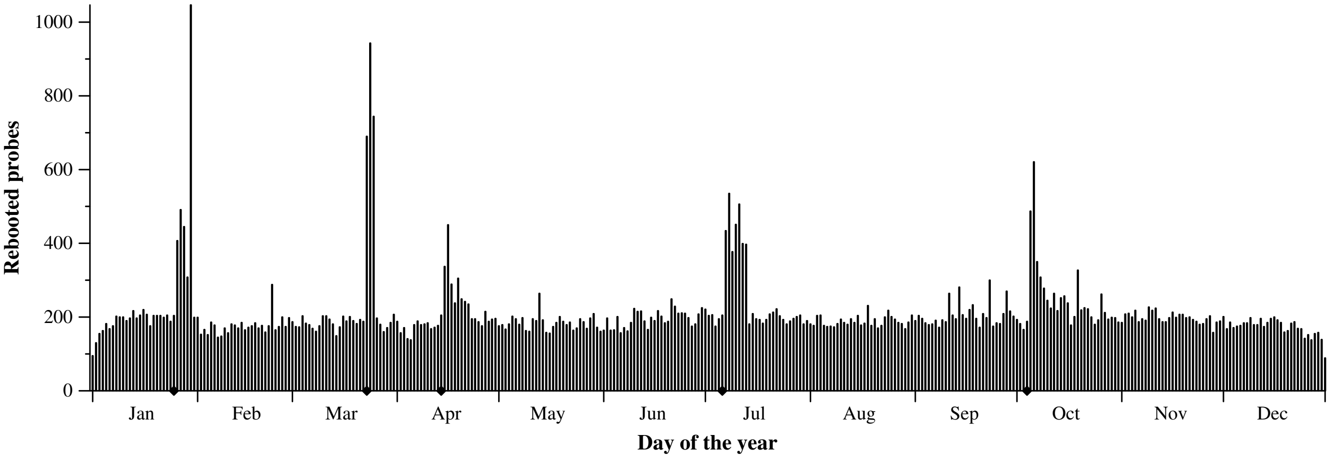

Figure 6: Number of unique probes that rebooted on each day of the year. Days with exceptionally many reboots follow the distribution of firmware updates. We indicate days where updates seem to have been distributed with diamonds along the x-axis.

5.3 Most outages result in an address change for some ASes

We found network and power outage events and associated them with inter-connection gaps as described in Section 3. If the connection log entries on either side of the inter-connection gap used different addresses, we infer that the event caused an address change and call the address change an *Address change with network outage*, *Address change with power outage*, and *Address change with no-outage*, depending upon the event.

For each individual probe, we consider the conditional probability of an address change given a detected outage. $P(ac|nw)$ represents the conditional probability that an address change occurred given a network outage and $P(ac|pw)$ represents the same for a power outage. We estimate this probability using the fraction of outages occurring contemporaneously with an address change (out of the total number of outages). We show the distribution of these probabilities by probe to estimate whether the group of probes (by geography or ISP) is dominated by those that always or seldom change addresses on an outage.

We find that the likelihood of address change upon an outage event differs across ASes. Figure 7 shows the CDF of $P(ac|nw)$ for the five ASes that host the most probes with at least one address change and at least three network outage events. We find that probes in ASes that periodically renumber—Orange, DTAG, and BT—have high $P(ac|nw)$ compared to probes from ASes that do not periodically renumber, LGI and Verizon. Around half of the probes in both Orange and DTAG had $P(ac|nw)$ equal to 1: every network outage was accompanied by an address change!

Figure 8 shows $P(ac|pw)$ for these ASes. Recall that we discarded probes with versions 1 and 2 due to their potential to reboot as a result of an address change, thus we have fewer samples. The AS-level behavior for power outages is similar to network outages. DTAG and Orange tend to renumber frequently upon power outages; half of the probes in Orange and 40% of the probes in DTAG have $P(ac|pw)$ equal to 1. Verizon and LGI do not renumber frequently upon power outages; only about half of their probes had an address change even once upon an outage. Since the likelihood of an address change upon an outage can also depend upon the duration of the outage, we investigate the distribution of outage durations and the likelihood of address changes for different outage durations in Section 5.4.

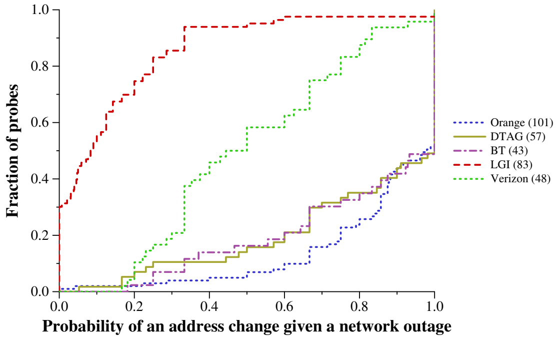

Figure 7: Distribution of $P(ac|nw)$ per probe for the ASes with the most probes that had at least one address change. Probes in DTAG, Orange, and BT, are far more likely to change addresses upon a network outage than probes in Verizon and LGI.

Since the ASes in Figure 7 and Figure 8 exhibit such disparate behavior, we considered if some ASes are particularly likely to renumber upon outages. To investi-

AS	ASN	Country	N	$P(ac\|nw) > 0.8$	$P(ac\|nw) = 1$	$P(ac\|pw) > 0.8$	$P(ac\|pw) = 1$
All			1113	29.1%	16.9%	28.3%	14.6%
Orange	3215	France	84	79%	54%	77%	50%
Telecom Italia	3269	Italy	28	71%	50%	57%	21%
BT	2856	U.K.	22	64%	55%	50%	14%
Proximus	5432	Belgium	20	70%	45%	60%	30%
DTAG	3320	Germany	19	58%	47%	47%	42%
Vodafone GmbH	3209	Germany	12	83%	75%	58%	42%
Wind Telecomunicazioni	1267	Italy	12	67%	42%	83%	42%
SFR	15557	France	16	38%	25%	50%	6%
ISKON	13046	Croatia	6	100%	50%	83%	67%
PJSC Rostelecom	8997	Russia	7	71%	29%	57%	14%

Table 6: Probes likely to change addresses upon network outages are also likely to change addresses upon power outages. The table shows autonomous systems with at least five probes whose conditional probability of address change upon network outage was greater than 0.8. The N column shows the number of probes with at least three network outages and at least three power outages. $P(ac|nw) > 0.8$ and $P(ac|nw) = 1$ show the percentage of N for which the conditional probability of address change upon network outage was greater than 0.8 and equal to 1 respectively, and $P(ac|pw) > 0.8$, $P(ac|pw) = 1$ show the same for power outages.

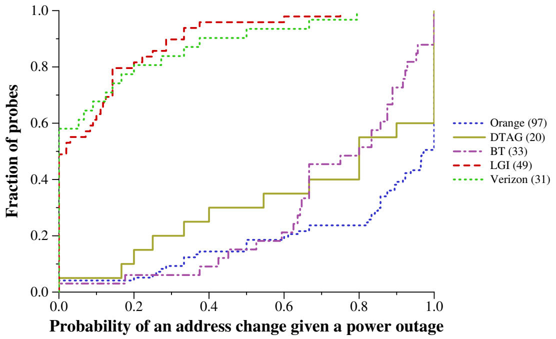

Figure 8: Distribution of $P(ac|pw)$ per probe for probes running version 3. As with network outages, probes in DTAG and Orange are more likely to change addresses upon power outage than probes in Verizon and LGI.

gate this, we found the set of probes with at least three network and power outages. We then found probes with $P(ac|nw)$ of 0.8 or more and show ASes with 5 or more such probes in Table 6.

First, we observe strong geographic correlation; all these ISPs are in Europe. Second, we observe that $P(ac|pw)$ is also high; $P(ac|nw) > 0.8$ and $P(ac|pw) > 0.8$ are similar for all these ISPs (although $P(ac|pw) = 1$ tends to be lower because our power outage detection technique is more prone to false positives). This suggests that both types of outages are likely to cause address changes. Third, we find that 7 of the 10 ISPs also appeared in Table 5. Maier et al. [19] studied the logs from an urban area of a major European ISP that used Radius to assign addresses: neither CPE nor Radius servers remember addresses. The behavior of these ISPs that nearly always renumber is consistent with the behavior of the large DSL provider in that study. Private communication with a large European ISP whose probes consistently had an address change upon outage confirmed that they use PPPoE and Radius to assign addresses for their DSL lines. We expect that this property can be used as evidence in inferring a device's link type.

5.4 Is there a relationship between outage duration and address changes?

Dynamic addresses assigned using DHCP should typically retain their addresses as long as they continue to renew their lease half-way into the lease duration as the standard recommends [10]. However, an outage could prevent them from renewing their lease. Depending upon the address churn at the time, the address they had previously been assigned may be reassigned to another device. In this way, an outage longer than half a lease duration could potentially cause an address change.

To investigate this, we analyzed the conditional probability of an address change given the occurrence of network or power outages of different durations for probes from LGI (AS 6830) and Orange (AS 3215) in Figure 9. For network outages, we considered outages from all versions of probes while for power outages, we only considered outages from probes running v3. We chose these ISPs due to their difference in address change behavior upon the occurrence of outages as seen in Figure 7 and Figure 8.

The behavior upon outages for the two ISPs is strikingly different. LGI's behavior appears consistent with what we would expect for dynamic addresses assigned using DHCP: fewer than 3% of outages of up to an hour resulted in an address change. More than 25% of outage durations that lasted at least twelve hours resulted in an address change. This behavior is consistent with a DHCP lease duration on the order of a few hours. Not every outage longer than twelve hours resulted in an address change, consistent with DHCP behavior when a

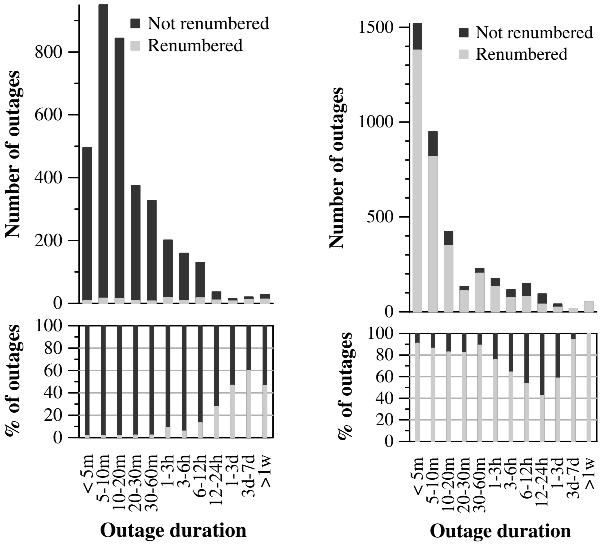

Figure 9: The likelihood of an address change (renumbering) given network or power outages of different durations in LGI (left) and Orange (right). The top graph is a histogram; the complete bar represents the number of outages observed across all probes in that AS. The lightly-shaded bar extends for those outages that also saw an address change. The lower graph shows the same data as a percentage. Although relatively few outages lasted longer than a day, the majority of these were coincident with an address change in both ISPs. However, Orange (right) changed addresses even on the shortest outages.

client returns after an expired lease and the previously assigned address is still available.

For Orange, we found that even very short outages resulted in address changes. 91% of outages that lasted less than five minutes resulted in an address change, and for every outage duration longer than five minutes and shorter than three hours, more than 75% occurred with an address change. For outages between three hours to three days long, the percentage of address changes was closer to 50%, suggesting the presence of some CPE devices that do not renumber upon every outage. However, as the outage duration increases beyond 3 days, almost every outage results in an address change.

Private communication with a large European ISP confirmed that this behavior is expected for PPPoE based DSL lines in that ISP: any reboot/reconnect event will result in the assignment of a new address from the ISP's dynamic address pool. Since outages of such short durations can result in an address change, a simple reboot of the CPE (resulting in a power outage), or unplugging and replugging the network cable (resulting in a network outage), can change the dynamic address assigned to the end-user. That end-users can change their dynamically assigned address has implications for researchers and operators who use IP addresses to identify end-hosts, particularly when IP addresses are being used to blacklist malicious actors.

6. DOES A USER'S DYNAMIC ADDRESS PREFIX CHANGE?

It is tempting to expect that a new address, when reassigned, will typically be drawn from nearby addresses, say, from the same enclosing /24 prefix. If such an assumption were true, it would allow blacklisting of the enclosing prefix of a malicious host, if it were thought that the malicious host could cause its address to change via reboot or by waiting a day. However, we find that such locality of addresses is rare and address changes typically span prefixes.

We examined whether the dynamic address assigment also varies the enclosing prefix, defined three ways. For each instance of address change that we observed, we found the BGP prefix of the previous address and the new address using CAIDA's IP-to-AS dataset [6], as described in Section 3. We also extracted the /16 and /8 prefixes from the previous and new addresses. We then compared how often the prefix of the previous address differed from the prefix of the new address. Table 7 presents the results for the overall AS-level dataset with 2,272 probes and for the ten ASes with the most probes that had at least one address change.

ISPs varied prefixes even for consecutive addresses assigned to the same customer; nearly half of the 166,644 total address changes we observed also changed BGP prefixes. Unlike periodicity and renumbering upon outages, assigning addresses out of different prefixes appears to be a common behavior for ISPs. For the ten ASes in Table 7, Verizon and DTAG had the lowest percentage of address changes across prefixes, but even for these ASes, almost a quarter of all address changes were across /16s and a fifth of all address changes were across /8s. Thus, it is not just the dynamic addresses that change; their prefixes change too. When a malicious actor receives a new address, even blacklisting the entire enclosing /8 prefix of the old address would fail to prevent access for a third of the address changes we observed.

7. RELATED WORK

Previous work studied the performance of DHCP in small campus networks [5,17] and settings where smartphone usage is widespread [23] and developed techniques to reduce network address utilization and DHCP broadcast traffic. The goal of those studies was to improve the performance of DHCP by tuning configuration.

Conceptually, so long as there is some uniquely identifying feature that remains constant across a host's address change, it is possible to track IP address changes over time for that host. Several studies have used this broad method [2, 7, 13, 17, 19, 21, 48]. UDmap [48] studied dynamic address properties using Hotmail user login traces where the user's login serves as the identifying

AS	ASN	Country	Diff BGP		Diff /16		Diff /8	
	All		81,571	48.9%	79,430	47.7%	55,835	33.5%
Orange	3215	France	7,016	68%	6,961	67%	5,513	53%
LGI	6830	many	171	56%	168	55%	136	45%
BT	2856	U.K.	1,736	44%	2,685	68%	1,735	44%
DTAG	3320	Germany	4,706	24%	5,391	28%	4,610	24%
Verizon	701	U.S.	241	23%	241	23%	209	20%
Comcast	7922	U.S.	76	37%	74	36%	63	31%
Proximus	5432	Belgium	2,152	49%	2,331	53%	1,983	45%
Telecom Italia	3269	Italy	4,281	85%	4,412	88%	2,374	47%
Ziggo	9143	Netherlands	18	35%	22	43%	16	31%
Virgin Media	5089	U.K.	46	84%	49	89%	39	71%

Table 7: Number of address changes across prefixes. Diff BGP shows the number of address changes where the previous address and the next address belonged to different BGP prefixes. Diff /16 shows the number of address changes where the previous address and the next address belonged to different /16 prefixes and Diff /8 shows the number of address changes where the previous address and the next address belonged to different /8 prefixes. The % column shows the percentage of total address changes for that autonomous system.

feature. Casado et al. [7] tracked clients using HTTP cookies when clients access a CDN. Other studies [13,19] used continuous responsiveness of an address itself as the identifying feature, assuming that an address that responds continuously belongs to the same user and that when an address stops responding to pings, it has been reassigned.

While we share the same goal as these studies, our approach diverges in that we are interested in the events associated with an address change. Previous studies lacked access to end-host information that could reveal the cause of an address change. One exception, Maier et al. [19], used access to the Radius server of a European DSL provider from one urban area to identify why DSL sessions terminated, and noted that the DSL provider often limited Radius session length to 24 hours in that area. We extend this result to several ISPs in countries from Europe, Asia, and South America, and identify other typical session length limits. Argon et al. [2] used periodic measurements from end-hosts in the DIMES infrastructure [38]. DIMES software installed on an end-user computer is different from RIPE Atlas hardware probes primarily in that it reports back only every 30-60 minutes (as opposed to RIPE Atlas's 3 minutes), the agent can be installed on laptops that move (as opposed to RIPE Atlas probes that could move, but do not), the hosts running DIMES are often powered down (resulting in limited uptime), and DIMES hosts appear to have static IP addresses more often (they reported 60% had only one address). Nevertheless, Argon et al. observed that some small ISPs exhibited address alternation with a 24 hour periodicity. In IPv6, the RFC for privacy extensions for stateless address autoconfiguration recommends that IPv6 addresses be changed every 24 hours [22] and empirical results by Plonka and Berger found that more than 90% of client IPv6 addresses were ephemeral [24]. We showed that 24 hour defaults are not uncommon in IPv4 as well.

These studies relied on relatively uncontrolled observations of the address assigned to a device or user, both in terms of whether the devices are active, whether the users connect using multiple devices, and how frequently samples are provided. As a consequence, the dynamic IP address churn rates reported by these studies vary. While UDmap reported that over 30% of IP addresses have inter-user durations of 1–3 days [48], Heidemann et al. reported that 90% of IP addresses were occupied for less than a day [13]. Maier et al. [19] reported that a major European ISP had per-user median durations of just 20 minutes during their study in 2009 (we did not observe this duration in 2015). We believe that the perspective of a device using the dynamically assigned network is necessary for understanding the reasons behind the address change and for getting precise information about the duration that any address is held. Further, since RIPE Atlas probes provide continuous, longitudinal measurements enabling the inference of successive addresses assigned to a CPE device, we perform the first analysis of dynamic prefixes from which devices are assigned successive addresses.

8. CONCLUSIONS

Among other motivations, the scarcity of IPv4 addresses has led ISPs to dynamically assign IP addresses to customer devices as needed, rather than assigning static addresses to each device. Although dynamic addressing can offer tremendous efficiency gains in IP address usage, it creates challenges for those who legitimately need to track individual host behavior, e.g., operators, IP address blacklist maintainers, law enforcment, researchers. Understanding the reasons dynamic addresses change is the first step towards predicting how long one can reasonably expect the same address to identify an end-host.

To study the conditions in which addresses change, we used an existing set of logs from a vast measurement infratructure—RIPE Atlas with 3,038 globally distributed probes that saw address changes in 2015—to infer and analyze patterns of address changes and factors that induce them. We found several factors in play. Dynamic address durations vary by geography, with addresses from North American ISPs persisting for weeks and addresses from many German ISPs assigned for a day. Dynamic addresses change as a result of network and power outages in most ISPs. In some ISPs, an outage of any duration results in an address change, while in others, the likelihood of address change increases with outage duration.

Two of our findings seem at odds with the address assignment practice specified in the DHCP standard [10], which states as a design goal that an address assigned to a client should persist as long as the client continues to renew its lease, even across client and DHCP server reboots. First, our dataset included periodically reassigned addresses for many European and nearby Asian probes; these reassignments potentially terminated sessions that were active during the time of renumbering. Second, many address changes seem to result from reboots and reconnect events. These observations—that some addresses change daily and may change under the control of a user—have implications both for researchers who might use addresses as a means of counting or tracking individual users and for operators that might blacklist addresses for misbehavior. We provide a list of ISPs that renumbered periodically and their renumbering parameters in Table 5; the maximum duration these ISPs are likely to assign an address to a CPE can be estimated accurately with high probability. We also provide a list of ISPs that renumber consistently upon reboot and reconnect events in Table 6; malicious users from these ISPs can evade blacklists by simply rebooting their device.

In starting this work, we anticipated that the rich dataset provided by RIPE Atlas would enable us to infer the configured duration of DHCP leases. It turns out that address reassignment was substantially more complex than we expected, with periodic address reassignment of even connected, functioning equipment being a common practice. The technical goal of efficient address assignment appears dominated by non-technical goals of ISP policy, e.g., privacy and preventing server operation. In particular, we believe that the address durations we measured are distinct from lease durations.

In this work, we found only one instance of administrative renumbering—reassignment of addresses en masse from one prefix to another—although this may be a limitation of our data. Recent research reports that there is continuous churn in the IPv4 address space: the set of addresses observed at a large CDN on one day differs from the set of address observed on the next day by 8% on average [26]. In future work, we plan to analyze how much of the observed churn in the address space can be attributed to administrative renumbering. Investigating dynamic addressing patterns in IPv6 and comparison with IPv4 is also future work.

Acknowledgments

We thank the anonymous reviewers and our shepherd, Paul Barford, for their helpful comments. We would like to express our sincere appreciation to the RIPE Atlas team for their efforts in deploying and maintaining thousands of probes over the world. The diversity of measurements from their probes along with their ease of access was invaluable to this work. We would especially like to thank Philip Homburg for his timely and thorough feedback about various measurements being conducted from the probes.

This research was supported by NSF grants CNS-1526635 and CNS-1528148.

9. REFERENCES

[1] Dennis Andriesse, Christian Rossow, and Herbert Bos. Reliable recon in adversarial peer-to-peer botnets. In *IMC*, 2015.

[2] Oded Argon, Anat Bremler-Barr, Osnat Mokryn, Dvir Schirman, Yuval Shavitt, and Udi Weinsberg. On the dynamics of IP address allocation and availability of end-hosts. *arXiv preprint arXiv:1011.2324*, 2010.

[3] About RIPE Atlas: FAQ: How does the probe connect to the Internet? https://atlas.ripe.net/about/faq/.

[4] R. Braden, Editor. Requirements for internet hosts – communication layers. IETF RFC-1122, October 1989.

[5] Vladimir Brik, Jesse Stroik, and Suman Banerjee. Debugging DHCP performance. In *IMC*, 2004.

[6] CAIDA. Routeviews prefix to as mappings dataset (pfx2as) for IPv4 and IPv6. https://www.caida.org/data/routing/routeviews-prefix2as.xml.

[7] Martin Casado and Michael J. Freedman. Peering through the shroud: The effect of edge opacity on IP-based client identification. In *NSDI*, 2007.

[8] The CBL. http://www.abuseat.org/.

[9] Jacky C. Chu, Kevin S. Labonte, and Brian N. Levine. Availability and locality measurements of peer-to-peer file systems. In *ITCom: Scalability and Traffic Control in IP Networks*, 2002.

[10] Ralph Droms. Dynamic Host Configuration Protocol. IETF RFC-2131, March 1997.

[11] Fail2ban. http://www.fail2ban.org/.

[12] Zwangstrennung (Forced IP address change). https://de.wikipedia.org/wiki/Zwangstrennung.

[13] John Heidemann, Yuri Pradkin, Ramesh Govindan, Christos Papadopoulos, Genevieve Bartlett, and Joseph Bannister. Census and Survey of the Visible Internet. In *IMC*, 2008.

[14] Philip Homburg. NTP measurements with RIPE Atlas. https://labs.ripe.net/Members/philip_homburg/ntp-measurements-with-ripe-atlas, February 2015.

[15] Jaeyeon Jung and Emil Sit. An empirical study of spam traffic and the use of DNS black lists. In *IMC*, 2004.

[16] Andrew J. Kaizer and Minaxi Gupta. Open resolvers: Understanding the origins of anomalous open DNS resolvers. In *PAM*, 2015.

[17] Manas Khadilkar, Nick Feamster, Matt Sanders, and Russ Clark. Usage-based DHCP lease time optimization. In *IMC*, 2007.

[18] Marc Kührer, Thomas Hupperich, Jonas Bushart, Christian Rossow, and Thorsten Holz. Going wild: Large-scale classification of open DNS resolvers. In *IMC*, 2015.

[19] Gregor Maier, Anja Feldmann, Vern Paxson, and Mark Allman. On dominant characteristics of residential broadband internet traffic. In *IMC*, 2009.

[20] Glenn McGregor. The PPP Internet Protocol Control Protocol (IPCP). IETF RFC-1332, May 1992.

[21] Giovane CM Moura, Carlos Ganán, Qasim Lone, Payam Poursaied, Hadi Asghari, and Michel van Eeten. How dynamic is the isps address space? towards internet-wide dhcp churn estimation. *IFIP*, 2015.

[22] Thomas Narten, Richard Draves, and Suresh Krishnan. Privacy extensions for stateless address autoconfiguration in ipv6. IETF RFC-4941, September 2007.

[23] Ioannis Papapanagiotou, Erich M. Nahum, and Vasileios Pappas. Configuring DHCP leases in the smartphone era. In *IMC*, 2012.

[24] David Plonka and Arthur Berger. Temporal and spatial classification of active IPv6 addresses. In *IMC*, 2015.

[25] Moheeb Abu Rajab, Jay Zarfoss, Fabian Monrose, and Andreas Terzis. My botnet is bigger than yours (maybe, better than yours): why size estimates remain challenging. In *Proceedings of the 1st USENIX Workshop on Hot Topics in Understanding Botnets, Cambridge, USA*, 2007.

[26] Philipp Richter, Georgios Smaragdakis, David Plonka, and Arthur Berger. Beyond Counting: New Perspectives on the Active IPv4 Address Space. In *IMC*, 2016.

[27] RIPE NCC. Atlas. http://atlas.ripe.net.

[28] RIPE NCC. Become a ripe atlas probe host. https://atlas.ripe.net/get-involved/become-a-host/.

[29] RIPE NCC. Built-in measurements. https://atlas.ripe.net/docs/built-in/.

[30] RIPE NCC. RIPE atlas connection logs url format. https://atlas.ripe.net/probes/<prb_id> /connection-history/<yyyy>/<mm>/.

[31] RIPE NCC. RIPE atlas probe archive. https://atlas.ripe.net/api/v1/probe-archive/.

[32] RIPE NCC. Technical updates. https://atlas.ripe.net/resources/announcements/.

[33] RIPE NCC Staff. RIPE Atlas: A global internet measurement network. *Internet Protocol Journal*, 18(3), September 2015.

[34] Stefan Saroiu, P. Krishna Gummadi, and Steven D Gribble. Measurement study of peer-to-peer file sharing systems. In *MMCN*, 2002.

[35] Aaron Schulman and Neil Spring. Pingin' in the rain. In *IMC*, 2011.

[36] Vyas Sekar, Yinglian Xie, Michael K. Reiter, and Hui Zhang. A multi-resolution approach for worm detection and containment. In *DSN*, 2006.

[37] Subhabrata Sen and Jia Wang. Analyzing peer-to-peer traffic across large networks. *IEEE/ACM Transactions on Networking (ToN)*, 12(2):219–232, 2004.

[38] Yuval Shavitt and Eran Shir. Dimes: Let the internet measure itself. *ACM CCR*, 35(5):71–74, 2005.

[39] William Simpson. The Point-to-Point Protocol. IETF RFC-1661, July 1994.

[40] Sorbs (spam and open-relay blocking system). www.sorbs.net/.

[41] The spamhaus project. http://www.spamhaus.org/.

[42] Brett Stone-Gross, Marco Cova, Lorenzo Cavallaro, Bob Gilbert, Martin Szydlowski, Richard Kemmerer, Christopher Kruegel, and Giovanni Vigna. Your botnet is my botnet: analysis of a botnet takeover. In *Proceedings of the 16th ACM conference on Computer and communications security*, 2009.

[43] Tech stuff - static and dynamic IP addresses. http://www.zytrax.com/isp/faqs/static.htm.

[44] Why does my external IP change every day? http://www.makeuseof.com/answers/why-does-my-external-ip-change-every-day/, June 2013.

[45] Why does your IP address change now and then? http://whatismyipaddress.com/keeps-changing.

[46] Why do ISPs change your IP address? http://www.howtogeek.com/163747/why-do-isps-change-your-ip-address/.

[47] Yinglian Xie, Vyas Sekar, David Maltz, Michael K. Reiter, Hui Zhang, et al. Worm origin identification using random moonwalks. In *Proc. of the IEEE Symposium on Security and Privacy*, 2005.

[48] Yinglian Xie, Fang Yu, Kannan Achan, Eliot Gillum, Moises Goldszmidt, and Ted Wobber. How dynamic are IP addresses? In *ACM SIGCOMM*, 2007.