NAT Revelio: Detecting NAT444 in the ISP

Andra Lutu[1], Marcelo Bagnulo[2], Amogh Dhamdhere[3], and kc claffy[3]

[1] Simula Research Laboratory, Norway
[2] University Carlos III of Madrid, Spain
[3] CAIDA/UC San Diego, CA

Abstract. In this paper, we propose *NAT Revelio*, a novel test suite and methodology for detecting NAT deployments beyond the home gateway, also known as NAT444 (e.g., Carrier Grade NAT). Since NAT444 solutions may impair performance for some users, understanding the extent of NAT444 deployment in the Internet is of interest to policymakers, ISPs, and users. We perform an initial validation of the NAT Revelio test suite within a controlled NAT444 trial environment involving operational residential lines managed by a large operator in the UK. We leverage access to a unique SamKnows deployment in the UK and collect information about the existence of NAT444 solutions from 2,000 homes and 26 ISPs. To demonstrate the flexibility of NAT Revelio, we also deployed it in project BISmark, an open platform for home broadband internet research. We analyze the results and discuss our findings.

1 Introduction

The Internet Assigned Numbers Authority (IANA) officially announced the depletion of IPv4 addresses in February 2011. But many Internet services and applications still require IPv4, motivating the standardization and deployment of protocols that support more aggressive, i.e., multi-level, sharing of IPv4 addresses [11], e.g., NAT444 within access ISP networks. NAT444 involves two phases of address translation, from a private IPv4 address block in the subscriber's network, to another local IPv4 address block in the provider's network, and finally to globally routable IPv4 addresses. NAT444 technology adds significant operational complexity that can impede performance or even break applications [6, 8]. In particular, NAT444 removes the control that the residential user usually has to configure port forwarding over single-level NAT, e.g., for peer-to-peer gaming. NAT444 also limits the number of ports available per subscriber, threatening the availability of popular applications that use many ports, e.g., Google Maps [3]. Another complication of NAT444 is customer identification, since the subscriber no longer maps to a unique globally routable IP address. Finally, pervasive NAT444 deployment may slow down the transition to IPv6, promoting the likelihood of the Internet's fragmentation between the two protocols. With such potentially negative impacts of what seems a likely future scenario, it behooves policymakers, ISPs and Internet users to monitor the extent of NAT444 deployment in the Internet. But like many aspects of Internet structure, systematic measurement and monitoring of NAT444 deployment in the wide area is challenging.

We propose *NAT Revelio*, a novel test suite methodology for detecting NAT444 deployments within the ISP access network. In order to detect NAT444 cases, the Revelio

test suite aims to determine the location of the device translating to the globally routable public IP address that identifies the subscriber to the global Internet. If we find that the subscriber's home network is not hosting this device, we conclude that the ISP deploys NAT444. Our approach relies on detecting network configuration characteristics peculiar to NAT444 deployment in an access network. We design our solution to be highly versatile and not require prior knowledge of the setup that we are about to test. In particular, we target deployment of Revelio on large-scale measurement platforms deployed in subscriber homes, such as the SamKnows large scale measurements platform [14] and BISmark [16], an open platform for home broadband internet research.

2 Generic NAT444 Deployment Architecture

We design *NAT Revelio* [1] to detect a wide range of NAT444 solutions in various configurations in ISPs, without any prior knowledge on the environment we test. The *Revelio client* executes in a device deployed in the home network, such as a measurement device or a computer. The *Revelio client* performs six active tests against different elements, including one or more servers deployed in the public Internet. In the rest of this section we establish the terminology we use in this paper and give an overview of possible NAT444 deployment architectures. We use the latter to explain how we deploy NAT Revelio to detect NAT444 in the ISPs we test.

There are various NAT444 implementations. We describe next the NAT444 deployment architecture in the context of DSL access technology, although this maps cleanly to other access technologies, e.g., FTTx, cable. One type of NAT444 technology is *Carrier-Grade NATs* (CGN), also known as Large Scale NAT (LSN). DSL-based CGN devices are available in three configurations: (i) stand-alone, (ii) Broadband Remote Access Server (BRAS) insertion-card and (iii) Core Router (CR) insertion card. Also, NAT444 deployments can be distributed (at each BRAS) or centralized (at the CR). For simplicity of presentation, we describe a centralized deployment of stand-alone CGN directly connected to the CR in the ISP access network to explain our detection approach. Other NAT444 solutions are available [15].

In Figure 1, we illustrate this NAT444 architecture in DSL networks using the terminology of the IETF's Large-Scale Measurement of Broadband Performance working group (LMAP WG) reference path [4]. The path elements include:

- **Subscriber Device**: which initiates and terminates communications over the IP network. In the context of our measurement experiment this is the *measurement device* inside the subscriber's home network that executes the *Revelio client*.
- **Private Network**: a network of devices the subscriber operates in the home network, possibly using multiple layers of NAT, each operating different chunks of RFC1918 private address space.
- **Service Demarcation point**: where the ISP-managed service begins, usually the interface facing the public Internet on a residential gateway or modem.
- **Intra IP Access**: first point in the access network that uses a globally routable IP address.
- **Globally Routable Address Gateway (GRA GW)**: the point of interconnection between ISP's administrative domain and the rest of the Internet.

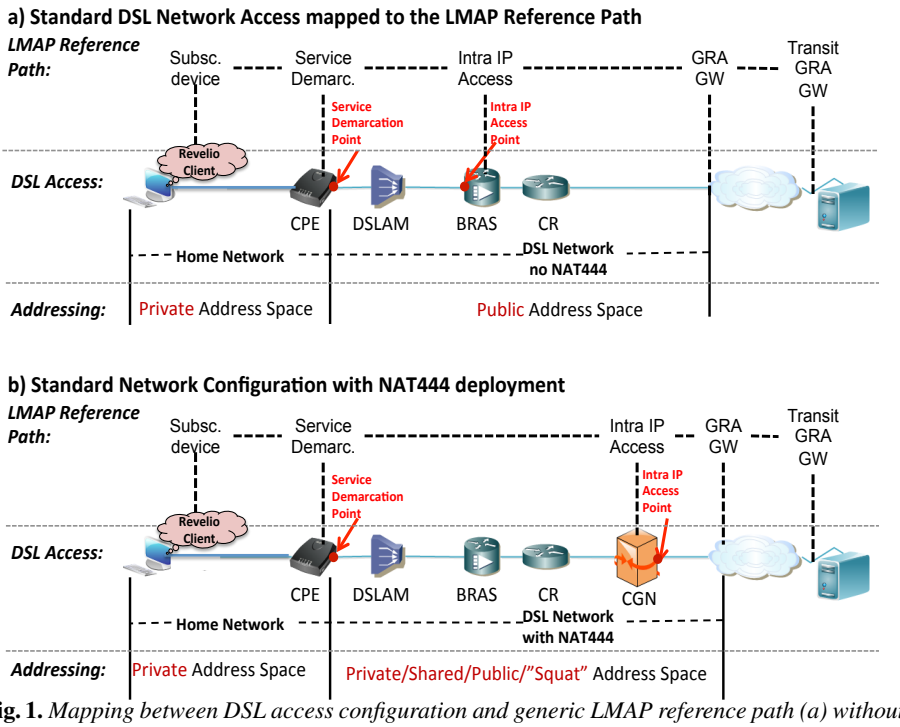

Fig. 1. *Mapping between DSL access configuration and generic LMAP reference path (a) without NAT444 and (b) with NAT444 (in this case, a stand-alone CGN) in the access network.*

Figure 1 illustrates the mapping between the LMAP reference path and a standard DSL network architecture, both (a) without NAT444 (but with traditional NAT), and (b) with NAT444 technology, using a stand-alone CGN device that connects to the CR. The customer premises equipment (CPE) usually performs the NAT function, translating private addresses in the home network to public addresses in the access network. The CPE is the Service Demarcation device; its Internet-facing interface is the Service Demarcation point. The BRAS is the Intra IP Access point – the first point *after the Service Demarcation point* that uses a globally routable IP address. The GRA corresponding to the subscriber maps to the IP address the ISP configures at the Service Demarcation point.

In the NAT444 configuration in Figure 1(b) the subscriber uses private addresses within the home network, prior to the Service Demarcation point. For the address space used between the Service Demarcation point and the Intra IP Access point, the access ISP can use private, shared [18], or public (legitimate or stolen/"squat") IPv4 addresses [3]. In this case, the Intra IP Access point maps to the NAT444 device (the stand-alone CGN), and the GRA is the IP address at the Intra IP Access point.

3 NAT Revelio Test Suite

This section describes the tests we use in the proposed test suite, and how we interpret them to infer the presence of NAT444 solutions in access ISPs.

3.1 NAT Revelio Overview and Design Challenges

Building a test suite for large-scale deployment of NAT444 measurements must account for possible non-standard configurations. Specifically, we need to account for cases where the subscriber deploys several levels of NAT within the home network. In particular, false inferences of NAT444 deployment can occur when we assume that the Revelio Client is directly connected to the Service Demarcation device when, in fact, two in-home NAT devices are in the path between the Subscriber Device and the Service Demarcation point. A naive NAT444 detection test could falsely assume that the first NAT device is the Service Demarcation point, and falsely map the second in-home NAT device to an Intra IP Access point.

Thus, we design NAT Revelio to operate in two phases: (i) *Environment Characterization* and (ii) *NAT444 Detection*. In the first phase, Revelio aims to establish the location of the Revelio Client within the home network relative to the Service Demarcation point. In the second phase, Revelio tests for the presence of NAT444 solutions and interprets the measurement results using the environment information.

NAT Revelio performs active measurements from a device running the Revelio Client in the subscriber network (see Figure 1). This step attempts to ascertain where the IPv4 address translation to the subscriber GRA occurs: in the subscriber home network (CPE) or in the ISP access network (a NAT444 device).

Figure 2 depicts a flow diagram of our test methodology. When deploying Revelio, we perform all the measurements in the test suite and merge their results to make an inference regarding the existence of NAT444 in the ISP.

3.2 Environment Characterization Phase

In the *Environment Characterization* phase Revelio runs three tests to determine the position of the Service Demarcation point relative to the Subscriber Device running the *Revelio client*. This step avoids false positive inferences of NAT444 and ensure accurate results over a wide range of in-home configurations. Figure 2 encloses the environment characterization tests in green rectangles. We use the information we retrieve here to interpret the results of the tests we run in the subsequent NAT444 Detection phase. Additionally, this phase allows us to detect the IP addresses configured in the home network of the subscriber and the ones in the ISP access network.

1. Identify subscriber's GRA. First, we use the Session Traversal Utilities for NAT (STUN) [13] protocol to discover the *Globally Routable Address (GRA)* that corresponds to the subscriber. STUN is a standard client-server protocol that allows a user behind a NAT to learn its public mapped address. We program the *Revelio client* to behave as a STUN client that queries an external STUN server (we use stun.stunprotocol.org), which replies with the GRA of the subscriber. If the ISP does not deploy NAT444 (Figure 1(a)), this GRA corresponds to the address exposed at the Service Demarcation point. If the ISP deploys NAT444 using a topology similar to that of Figure 1(b), this GRA corresponds to the public IP address exposed at the Intra IP Access point along the reference path. This step corresponds to the very first block of the Revelio flowchart in Figure 2, labeled *STUN Binding Request*. The information we retrieve by performing this action is illustrated in the flowchart by the data block labeled *Subscriber GRA*.

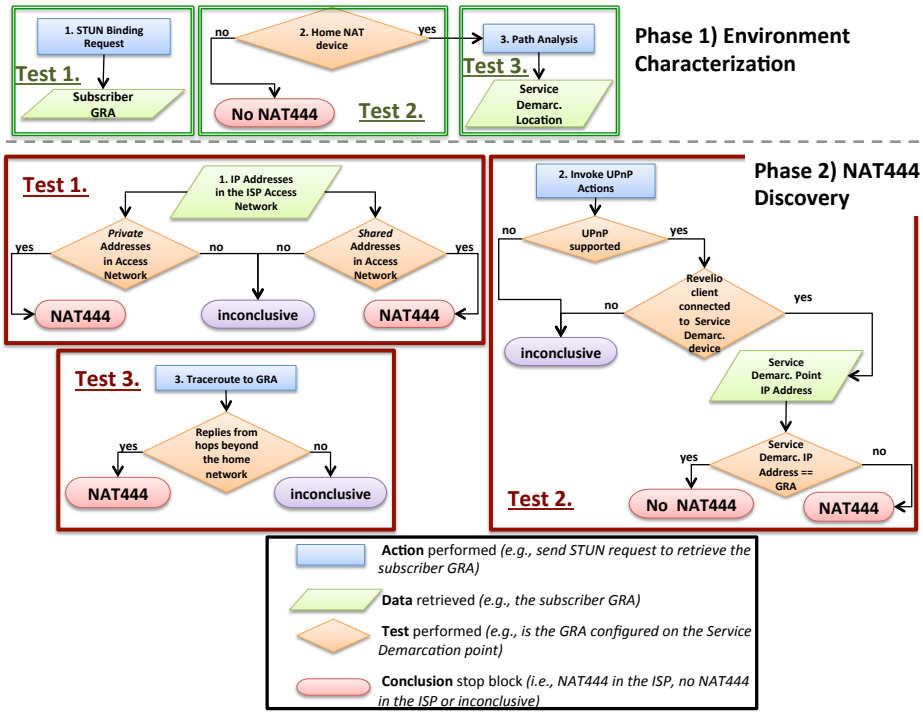

Fig. 2. *The NAT Revelio test suite flowchart: measurement actions (sending/receiving packets) are in blue rectangles; measurement data is in green parallelograms; tests on retrieved data are in orange rhombuses. Inferences of NAT444 are in red stop blocks. We use the data we collect in phase 1 of environment characterization for all subsequent NAT detection test we run in phase 2.*

2. Discover home NAT device. Second, we establish whether the Service Demarcation device performs NAT. Specifically, we verify that the local IP address of the Subscriber Device running the *Revelio client* in the home network is in private address space [12]. If the IP address of the Subscriber Device is a public address, we conclude that the client is not behind a NAT (and, implicitly, not a NAT444 device either). We further confirm this scenario when comparing the local IP address to the GRA. If these two match, then there is no NAT device along the path. We represent this step of the Environment Discovery phase in the NAT Revelio flowchart with the test block labeled *Home NAT device*. Depending on the results of the test, we include in the flowchart a stop block with *No NAT444* (i.e., a negative result), or we move on to the next step in this phase.

3. Locate Service Demarcation point [Path Analysis]. If the CPE performs NAT, we test to identify the location of the access link (i.e. the link between the Service Demarcation point and the first hop in the access network of the ISP) relative to the *Revelio client*. We heuristically identify the access link by assuming it is the first link on the outbound path with a transmission latency at least an order of magnitude higher than its neighboring links [17].

To quantify per-link latency we use a technique similar to *pathchar* [7]. Namely, we estimate per-link delay parameters by taking the minimum values of repeated Round Trip Time (RTT) measurements with different UDP packet sizes along a path, and assuming negligible queuing and processing delays (similar to [7]). To minimize the impact of these measurements on the subscriber's network, we gather the data by running traceroute hourly over a period of two days, using 21 different packet sizes varying from 120 bytes to 1,400 bytes, and using as a destination a high-availability IP address in Level 3's network. Limiting the number of packets to 21 per test allows us to complete one run of the NAT Revelio measurements in 30 seconds. Running Revelio once per hour for 2 days results in 48 RTT samples per TTL per packet size. We analyze these values to estimate per-link propagation delay, and infer that the first link with a ten times latency increase relative to its neighboring links is the access link. We use the pathchar result (labeled *Service Demarc. Location* in the flowchart) in the tests we perform in the second phase of NAT Revelio.

3.3 NAT444 Discovery Phase

This phase seeks to identify the location of the device performing NAT to the GRA mapped to the subscriber, namely before or after the Service Demarcation point. Figure 1(b) depicts the scenario with NAT444 (CGN) deployed in the DSL access ISP network, after the BRAS and the Core Router. When the ISP deploys NAT444, the location of the Intra IP Access point changes compared with the case where the ISP does not use NAT444 (Figure 1(a)). In Figure 2, we depict enclosed in red rectangles the three tests we run for NAT444 detection. We perform all three tests and interpret the set of results we obtain together with the information we collect in the *Environment Characterization* phase to make an inference regarding NAT444 deployment in the ISP we measure. To increase the robustness of the test suite to non-standard architectures, e.g., when the ISP does not deploy NAT444, but configures private addresses in its access network, we assign a different confidence level to each test. One strength of Revelio lies in being able to compare the results of multiple tests for the same subscriber. To control against false positives, when test results conflict, we give priority to the negative result, concluding there is no NAT444 deployment in the ISP.

1. Identify private/shared addresses in the ISP access network. The first method in the NAT444 Discovery phase detects the use of private or shared IP addresses in the access network, between the Service Demarcation point and the Intra IP Access point. Figure 1(b) depicts an ISP using special address domains (i.e., private or shared address space) in its access network when a NAT444 solution is in place. We characterize the path obtained by traceroutes in Phase (1), step 3, including inferring the position of the Service Demarcation point. We then check if private or shared addresses are configured along the path toward the public Internet target which is a router inside Level3, and if so, determine their location relative to the Service Demarcation point. This discovery helps us to establish if the private/shared addresses we identify are configured in the ISP access network. The information allows us to correctly distinguish cases of multiple levels of NAT in the home network, which can otherwise be easily confused with NAT444 deployment. The flowchart (Figure 2) represents this step by including the data

block labeled *IP Addresses in the Access Network* (which gets as input the location of the Service Demarcation point relative to the *Revelio client*) and the two following tests: *Private/Shared Addresses in Access Network*.

Note that we assign different confidence levels to these two tests. When we observe shared address space in the ISP, beyond the Service Demarcation point, we are *highly confident* of the presence of a NAT444 solution, given that these addresses are specifically for use in NAT444 deployment. However, when we observe RFC1918 private addresses beyond the Service Demarcation point, we give a *low confidence level* to our results, because the ISP might use private address space for its internal infrastructure without deploying NAT444. Moreover, in the case where NAT Revelio does *not* detect any private or shared addresses past the Service Demarcation point, the test suite cannot discard the possibility of a NAT444 deployment in the ISP. This case can occur when the ISP configures public addresses (legitimate or stolen "squat" address space) in the access network as part of a NAT444 deployment.[4]

2. Invoke UPnP actions. NAT Revelio runs a series of tests that aim to infer the hop count between the Service Demarcation device and the device performing the final translation to the subscriber GRA. To check if the Service Demarcation device is the device translating to the subscriber GRA, we verify whether the address configured on the Service Demarcation point matches the subscriber's GRA.

If the *Revelio client* directly connects to the Service Demarcation device (Figure 1), we leverage the Universal Plug and Play (UPnP) IGP protocol [2] if supported by the CPE. The *Revelio client* sends a UPnP client control message to the CPE that retrieves the IP Address of the WAN interface of the CPE, which, in this case, maps to the Service Demarcation point. In the case of a match, we infer that the ISP **does not** use NAT444. A mismatch between these addresses means that the ISP **does indeed** deploy NAT444. We give a *high level of confidence* to this result.

Otherwise, if the Subscriber Device running the *Revelio client* does not connect to the Service Demarcation device, we find ourselves in a non-standard configuration, where multiple NAT devices are present within the home network. In this case, we cannot draw any conclusion regarding the presence of NAT444 in the ISP from this test, since the UPnP test retrieves the IP address of the innermost CPE device within the home network, and not the IP address at the Service Demarcation point. The NAT Revelio flowchart includes this set of tests, following the *yes* branches both for the *UPnP Supported* and the *Revelio client connected to Service Demarcation device* tests, in the NAT444 Discovery phase in Figure 2.

3. Traceroute to the subscriber GRA. We also run traceroute from the *Revelio client* to the subscriber GRA to measure the hop count between them. Without NAT444, the GRA is at the Service Demarcation point (Figure 1(a)), and all traceroute-responding hops are inside the home network. With NAT444 (Figure 1(b)), the GRA is at the Intra IP Access point, which is past the Service Demarcation point. If we already know the

[4] A common configuration is to assign private or shared address space only to the interface of the Service Demarcation point attached to the ISP network, while other elements of the ISP network use public addresses.

location of the Service Demarcation point relative to the *Revelio client* (from the first phase), a UDP traceroute to the GRA distinguishes these two cases.

We assign to the *Traceroute to GRA* test a *high confidence level*, since it relies on no CPE-specific capabilities, nor on the assumption that the ISP configures private or shared IP addresses in the access network. Nonetheless, this test still may fail to determine the presence of NAT444 in the ISP, for example when the ISP actively blocks ICMP packets triggered by the traceroute. Thus, NAT Revelio cannot conclusively determine the presence of a NAT444 solution in the ISP. Figure 2 illustrates this possibility in the NAT Revelio flowchart with the purple *inconclusive* stop block.

4 Validation and Large-Scale *Revelio* Measurement Campaigns

4.1 Revelio Validation in Controlled Environment

With the help of a large UK ISP operator, we tested NAT Revelio on a controlled set of subscribers included in a trial deployment of a CGN implementation of NAT444 within the ISP network. The trial environment consisted of operational DSL residential lines connected behind a stand-alone CGN NAT444 implementation. We ran the *Revelio client* on 6 Subscriber Devices, 2 of which were behind the NAT444 device. We found that *NAT Revelio* accurately detected the deployment configuration of all 6 devices. We explain details of the test results below.

After running the **Environment Discovery** (Section 3.2), we learned that all six Subscriber Devices running the Revelio Client connected directly to the Service Demarcation device within the home network.

For the two subscribers connected to the ISP behind a NAT444 solution, all tests in the **NAT444 Discovery** (Section 3.3) successfully indicated the presence of NAT444 within the access network. First, after retrieving the CPE's WAN IP address which corresponds to the Service Demarcation point address (as per the test we describe in Section 3.3.1), we identified it as shared address space, which is a clear symptom of NAT444 deployment. Second, we confirmed that the subscriber GRA did not match the Service Demarcation point address (as per the test we describe in Section 3.3.2), reinforcing evidence of NAT444 deployment. Third, when verifying how far from the Service Demarcation device the translation to GRA occurred (as per the test we describe in Section 3.3.3), we measured 6 hops between the Subscriber Devices and the device translating to the GRA. Only the first of these hops belonged to the home network, leaving 5 hops between the Service Demarcation device and the device performing translation to the GRA.

NAT Revelio successfully inferred that the other 4 Whiteboxes were not behind a NAT444 solution after *Invoking UPnP Actions* (Section 3.3.2) and concluding that the IP addresses at the Service Demarcation point matched the GRA of the subscriber.

To illustrate Revelio's robustness to non-standard configurations, we also tested our NAT444 detection approach on 24 residential DSL lines operated by a large Italian ISP that does not employ NAT444 solutions in its DSL network. However, in its access network configuration, the ISP does use private IP address space for its infrastructure. This is a non-standard configuration that can wrongly mimic the presence of a NAT444

solution in the ISP. Due to the fact that we consider multiple tests to detect NAT444 in the ISP, we were able to discard such cases on the basis of conflicting results. We found that the first test in NAT444 Discovery (Section 3.3.1) indicated the existence of a NAT444 solution in the ISP based on the detection of RFC1918 address space beyond the Service Demarcation point. Since the operator disabled UPnP on its home routers, we could not invoke any UPnP actions (Section 3.3.2). However, *traceroute to the subscriber GRA* (Section 3.3.3) showed that the GRA is, in fact, at the Service Demarcation point. As we mention in Section 3.3, when we have conflicting results from Revelio tests, we give priority to the negative test to avoid false negatives. Thus, we accurately concluded that the Italian ISP does not have any NAT444 deployment.

4.2 Large-Scale Measurement Campaigns

After the above validation exercise, we experimented with *NAT Revelio* on two different large-scale measurement platforms (SamKnows' UK deployment and BISmark), targeting multiple ISPs and potential NAT444 solutions.

SamKnows Deployment. We deployed the Revelio Client on a set of SamKnows Whiteboxes within home networks in the UK. A SamKnows Whitebox is a custom hardware device that residential users host voluntarily. We ran *NAT Revelio* from 2,000 Whiteboxes that allowed us to test 26 different ISPs for NAT444 solutions. We had no previous knowledge of the configuration of these ISPs. We collected results of tests of two different Revelio deployments that we performed 5 months apart, in June 2014 and October 2014. Although they did not cover the same subscribers, both campaigns yielded similar results, indicating that the NAT444 deployment did not expand during the five-month period.

The results of June 2014 campaign revealed that out of the approximately 2,000 residential lines we tested, we inferred that 10 different end-users connected behind a NAT444 solution. The 10 users were spread across 5 different ISPs. Thus, the proportion of end-users we inferred were behind a NAT444 solution was 0.5% of all the residential lines we tested. We were able to validate these findings with the operators for only for one case.[5] The operator in question validated our inferences for the lines we found to be deployed behind a NAT444 solution.

Analyzing the results from the June 2014 campaign, we inferred that a total of 90% of tested end-users were not connected through a NAT444 solution (no NAT444). The **Environment Characterization** phase of *NAT Revelio* helped us discard 60% of the cases of in-home cascaded NATs that would have otherwise emerged as false positives.

In the **NAT444 Discovery** phase, the *Invoking UPnP Actions* test (Section 3.3.2) successfully ran on 82% of the SamKnows Whiteboxes, further identifying 81.2% of the tested customers as **not** configured to use a NAT444 solution. In the other 18% of the cases, UPnP was not supported by the home gateway, so we could not run this test. Additionally, the *Traceroute to the GRA* (Section 3.3.3) independently classified approximately 50% of the end-users we tested as not behind a NAT444 deployment. In

[5] Attempting to validate our findings, we have contacted all the 5 ISPs, but we have yet to receive a reply from 4 of them.

9.5% of observed cases, we could not draw a conclusion because all tests included in the NAT444 Discovery phase gave **inconclusive** results.

The October 2014 deployment covered fewer subscribers (approximately 1,500 SK Whiteboxes) than the one in June 2014 (approximately 2,000 SK Whiteboxes). We found that 4 ISPs deployed NAT444 solutions. The results we obtained for 3 of the 5 ISPs were consistent with the results we inferred of the June 2014 campaign. We detected one additional ISP for which the Subscriber Device (Whitebox) connected directly to the Service Demarcation Device, but for which the Service Demarcation point address was a private (Section 3.3.2). We give high confidence to this result.[6]

BISmark Deployment. Between 7-9 February 2015, we deployed NAT Revelio on 37 OpenWRT routers that are part of the BISmark measurement platform. Our BISmark experiment involved fewer vantage points than our SamKnows UK experiment, but they had much wider geographical distribution. We deployed the Revelio client in Subscriber Devices hosted in 24 different ISPs active in 13 countries distributed across the five Regional Internet Registries (RIRs). Using the Revelio test suite, we inferred the presence of NAT444 in three different ISPs: Vodafone for DSL customers in Italy, Embratel in Brasil and Comcast in the US. In all three cases, we inferred a NAT444 solution by establishing the presence of RFC1918 private addresses in the ISP access network (Section 3.3.1). The traceroute to the GRA (Section 3.3.3) gave inconclusive results in all three cases. Also, in the case of Embratel and Comcast, the Revelio client could not invoke UPnP actions (Section 3.3.2). Since an ISP may use RFC1918 addresses in the access network without deploying a NAT444 solution, we give low confidence to the latter two results, and mark them as potential false positives. In the case of the Subscriber Device connected to Vodafone Italia, the Revelio client could invoke UPnP actions and verify the presence of the NAT444 solution in the ISP. We give high confidence level to this result, where two Revelio tests detected NAT444 deployment.

5 Related Work

In recent years, detection of middleboxes, and characterization and assessment of their impact on the Internet, has become a topic of interest. In particular, researchers have studied how to identify the presence of middleboxes on the Internet path, including NAT444 solutions. A recent study proposed *NATAnalyzer* [10], an algorithm capable of discovering previously unknown cascaded NAT configurations. NATAnalyzer requires control of the client and server sides of the test, whereas *NAT Revelio* is a client-side discovery mechanism. NATAnalyzer determines the position of the NAT devices using repetitive traceroutes. First, the test establishes address mappings in NAT devices on the path by running a traceroute from the end-user side to the server. NATAnalyzer then relies on fixed NAT state timers to sequentially ensure that the per-hop mappings expire, while maintaining the rest of the mappings by sending traffic from the external server towards the client (a NAT configuration that represents a security risk and is not

[6] Attempting to validate this result, we found that several subscribers reported on the ISP's online customer support forum that they had identified the presence of the CGN by detecting the presence of shared address space in the ISP.

recommended). The algorithm does not account for timers that may differ for multiple NAT configurations across various networks. Revelio does not rely on any features of NAT devices, treating them as black boxes along the path.

The Netalyzr [9] tool, initially meant as a networking debugging tool, is continuously running a survey of the health of the Internet's edge by detecting anomalous configurations. This survey includes detection of NAT solutions. Unlike Revelio, Netalyzr is not specifically tailored to detecting NAT444 solutions, and might not be robust to non-standard configurations inside home networks.

Tracebox [5] is an extension to the widely used traceroute tool that detects various types of middlebox interference over an Internet path. The solution is prone to open issues affecting traceroute. Though this can also potentially impact Revelio, our test-suite also includes other tests which we can fallback on.

6 Conclusions and Future Work

Despite concerns about its performance impact, NAT444 is part of the technology landscape during this ongoing phase of transition from IPv4 to IPv6. In this paper, we proposed *NAT Revelio*, a novel methodology and test suite aimed at accurately detecting NAT444 deployments by running active tests from the home network. We validate the accuracy of our approach by evaluating the status of a control set of 6 residential lines tested in a NAT444 deployment trial within the network of a large UK operator. We tested the robustness of the test suite to a non-standard configuration by evaluating the status of 24 DSL residential lines connected to a large Italian ISP that does not deploy NAT444, but uses private addresses in its access network.

The large scale *NAT Revelio* distribution across the UK showed that NAT444 solutions are still in early stages of deployment in the UK. However, our results infer that operators are at least testing these solutions to potentially move them in production. Using the BISmark platform, we tested 24 additional ISPs active in 13 countries distributed across the five Regional Internet Registries (RIRs). We inferred the presence of NAT444 in three different ISPs and proved our solution to be highly versatile.

For future work, we will expand testing to other regions, where NAT444 solutions are more popular. In particular, we will deploy NAT Revelio in the SamKnows FCC Measuring Broadband America testbed in the US. We also plan to tackle the limitations of the proposed methodology, namely by designing other detection algorithms in the case when assumed CPE capabilities are not implemented or networks actively block ICMP packets.

Acknowledgments

This work has been partially funded by the European Community's Seventh Framework Program (FP7/2007-2013) grant no. 317647 (Leone). This work was supported by the U.S. NSF grants CNS-1513283 and CNS-1528148 and CNS-1111449. We would like to thank Sam Crawford and Andrea Soppera for their feedback and numerous discussions while designing NAT Revelio, as well as the support for the large-scale deployments of

Revelio on the SamKnows UK panel. We also thank Guilherme Martins for his support during the BISmark deployment and Dario Ercole for his help validating NAT Revelio.

References

1. List of spells in Harry Potter. http://en.wikipedia.org/wiki/List_of_spells_in_Harry_Potter, accessed: 2015-10-04
2. UPnP Forum. Universal Plug and Play (UPnP) Internet Gateway Device (IGD) V 2.0. http://upnp.org/specs/gw/igd2/ (December 2010), accessed: 2014-06-15
3. Aitken, B.: MC/159 Report on the Implications of Carrier Grade Network Address Translators. Final Report for Ofcom (2013)
4. Bagnulo, M., Burbridge, T., Crawford, S., Eardley, P., Morton, A.: A Reference Path and Measurement Points for Large-Scale Measurement of Broadband Performance. RFC 7398 (February 2015)
5. Detal, G., Hesmans, B., Bonaventure, O., Vanaubel, Y., Donnet, B.: Revealing middlebox interference with tracebox. In: Proceedings of the 2013 conference on Internet measurement conference. pp. 1–8. ACM (2013)
6. Donley, C., Howard, L., Kuarsingh, V., Berg, J., Doshi, J.: Assessing the Impact of Carrier-Grade NAT on Network Applications. RFC 7021 (September 2013)
7. Downey, A.B.: Using pathchar to estimate internet link characteristics. In: Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication. SIGCOMM '99 (1999)
8. Ford, M., Boucadair, M., Durand, A., Levis, P., Roberts, P.: Issues with IP Address Sharing. RFC 6269 (June 2011)
9. Kreibich, C., Weaver, N., Nechaev, B., Paxson, V.: Netalyzr: illuminating the edge network. In: Proceedings of the 10th ACM SIGCOMM conference on Internet measurement. pp. 246–259. ACM (2010)
10. Müller, A., Wohlfart, F., Carle, G.: Analysis and Topology-based Traversal of Cascaded Large Scale NATs. In: Proceedings of the 2013 Workshop on Hot Topics in Middleboxes and Network Function Virtualization (2013)
11. Perreault, S., Yamagata, I., Miyakawa, S., Nakagawa, A., Ashida, H.: Common Requirements for Carrier-Grade NATs (CGNs). RFC 6888 (April 2013)
12. Rekhter, Y., Moskowitz, B., Karrenberg, D., de Groot, G., Lear, E.: Address Allocation for Private Internets. RFC 1918 (February 1996)
13. Rosenberg, J., Mahy, R., Matthews, P., Wing, D.: Session Traversal Utilities for NAT (STUN). RFC (October 2008)
14. SamKnows™: Methodology and technical information relating to the SamKnows™ testing platform - SQ301-002-EN (2012)
15. Skoberne, N., Maennel, O., Phillips, I., Bush, R., Zorz, J., Ciglaric, M.: IPv4 Address Sharing Mechanism Classification and Tradeoff Analysis. Networking, IEEE/ACM Transactions on 22(2), 391–404 (April 2014)
16. Sundaresan, S., Burnett, S., Feamster, N., De Donato, W.: Bismark: a testbed for deploying measurements and applications in broadband access networks. In: 2014 USENIX Conference on USENIX Annual Technical Conference (USENIX ATC 14). pp. 383–394 (2014)
17. Sundaresan, S., De Donato, W., Feamster, N., Teixeira, R., Crawford, S., Pescapè, A.: Broadband internet performance: a view from the gateway. In: ACM SIGCOMM computer communication review. vol. 41, pp. 134–145. ACM (2011)
18. Weil, J., Kuarsingh, V., Donley, C., Liljenstolpe, C., Azinger, M.: IANA-Reserved IPv4 Prefix for Shared Address Space. RFC 6598 (April 2012)