The Record Route Option is an Option!

Brian J Goodchild$^{\alpha,\beta,\gamma}$ Yi-Ching Chiu$^{\alpha}$ Rob Hansen$^{\delta}$ Haonan Lu$^{\alpha}$ Matt Calder$^{\alpha,\epsilon}$

Matthew Luckie$^{\zeta}$ Wyatt Lloyd$^{\alpha}$ David Choffnes$^{\delta}$ Ethan Katz-Bassett$^{\alpha,\gamma}$

$^{\alpha}$USC, $^{\beta}$Rutgers – Camden,$^{\gamma}$Columbia, $^{\delta}$Northeastern, $^{\epsilon}$Microsoft, $^{\zeta}$University of Waikato

Abstract

The IPv4 Record Route (RR) Option instructs routers to record their IP addresses in a packet. RR is subject to a nine hop limit and, traditionally, inconsistent support from routers. Recent changes in interdomain connectivity—the so-called "flattening Internet"—and new best practices for how routers should handle RR packets suggest that now is a good time to reassess the potential of the RR Option.

We quantify the current utility of RR by issuing RR measurements from PlanetLab and M-Lab to every advertised BGP prefix. We find that 75% of addresses that respond to ping without RR also respond to ping with RR, and 66% of these RR-responsive addresses are within the nine hop limit of at least one vantage point. These numbers suggest the RR Option is a useful measurement primitive on today's Internet.

CCS Concepts

• **Networks** → **Routing protocols**; *Topology analysis and generation*;

Keywords

Record Route, IP Options, Routing, Topology, Traceroute

ACM Reference format:

Brian Goodchild, Yi-Ching Chiu, Rob Hansen, Haonan Lu, Matt Calder, Matthew Luckie, Wyatt Lloyd, David Choffnes, Ethan Katz-Bassett. 2017. The Record Route Option is an Option!. In *Proceedings of IMC '17: Internet Measurement Conference , London, United Kingdom, November 1–3, 2017 (IMC '17),* 7 pages.

https://doi.org/10.1145/3131365.3131392

1 Introduction

Researchers and network operators need to understand Internet routing topology to troubleshoot problems. However, the protocols are not designed to reveal many details of Internet operations, the Internet is vast and complex, and it is administered by autonomous networks. Thus the set of tools for measuring topology is small, and many aspects of Internet routing remain opaque.

Given limited visibility, the IP Record Route (RR) Option offers promising attributes as a measurement primitive. IP Options are a standard part of the Internet Protocol and can be enabled on any packet. Like traceroute, RR reports IP addresses along an Internet path from a source to a destination, but it offers several advantages over traceroute. For example, RR can piece together hop-by-hop the reverse path back from a destination [11], which is invisible to traceroute and other traditional techniques; and it can uncover some hops that do not respond to traceroute probes [20, 21].

However, RR has a number of limitations. It can only measure nine hops, which may not suffice to measure a full route. Even routers within that limit may drop or rate limit RR packets, or forward them without recording an IP address, especially since IP Options are processed on the slow path by a router's resource-constrained route processor [10]. In fact, a 2005 technical report found that 46% of PlanetLab-to-PlanetLab paths dropped RR packets, leading the report to be titled, "IP Options are not an option" [8].

While it is true that IP Options are not a great option for supporting end-to-end IP extensibility—the use case investigated by that technical report—in this paper we revisit the suitability of the IP Record Route Option as a measurement primitive for today's Internet. Specifically, we contribute the following results:

- *Most pingable destinations respond to Record Route pings.* Of nearly 300,000 IP addresses—each in a different routable BGP prefix—that responded to ICMP Echo Requests (pings), 75% also responded to pings we sent with the RR Option. In total, we received RR responses from destinations in 40,545 Autonomous Systems (AS) (out of 49,100 for which we had destinations that responded to pings) (§3.2)
- *Most responsive destinations are within Record Route range of our vantage points.* Probing the destinations that respond to RR pings from M-Lab and PlanetLab, two-thirds are within the nine hop RR limit of at least one vantage point, and 60% are within the eight hop limit necessary to measure reverse paths from them to any host we control [11] (§3.3). Measurements from our vantage points from 2011 and 2016 show that the number of destinations within range is greater today than in the past (§3.4).
- *The flattening Internet suggests that cloud providers can employ Record Route to good effect.* Large cloud and content providers build out their backbones and peer broadly to bring their networks and services close to end-users [3]. Our measurements suggest

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

IMC '17, November 1–3, 2017, London, United Kingdom

© 2017 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

ACM ISBN 978-1-4503-5118-8/17/11.
https://doi.org/10.1145/3131365.3131392

that Google is within eight hops of 86% of destinations that respond to RR pings; thus the IP Option can potentially solve much of Google's need to uncover the paths from end-users to Google [13] (§3.6).
- *Careful experiment design can increase RR response rates by avoiding triggering rate limiting.* We reduce the impact of rate limiting and increase the response rate to RR probes by throttling the probing rate of specific vantage points and limiting the initial TTL of RR packets (§4.1, §4.2).

2 Motivation and Goal

Despite its drawbacks, Record Route has advantages over traceroute. First, while traceroute measures one hop per packet, RR allocates space in the header for nine IP addresses so that each router can record an address as long as space remains. Because RR can record nine routers traversed by a single packet in that packet, RR can avoid artifacts that traceroute can introduce upon encountering a load balancer [2]. Second, traceroute only measures the forward path from source to destination, whereas RR data in an ICMP Echo Request (ping) packet can be copied by the destination into its ICMP Echo Reply packet, and any empty slots can be filled on the reverse path from destination back to source. This mechanism forms the basis of our reverse traceroute system that can measure the reverse path back to any local host from a destination within range of at least one local host [11]. Third, RR can capture some hops that are invisible to traceroute, such as routers that do not decrement TTL or routers inside tunnels with certain configurations [6, 22]. To be clear: RR is *not a replacement* for traceroute, rather it can complement traceroute. The two tools can be used in combination to augment our understanding of network topology, a topic both explored in previous work [17, 20] and open to future study.

The conventional wisdom overemphasizes the drawbacks. Based on informal conversations with colleagues, we believe that the *title* of the 2005 technical report—"IP Options are not an option"—led to the *unintended interpretation* that IP Options were not a good option *for measurement,* even though the contents of the report suggest the potential of RR for measurement [8]. The goal of the report was to assess whether IP Options could be used to provide IP extensibility, and the high fraction of paths that dropped Options packets means that they are not a general vehicle for supporting Internet-wide end-to-end functionality. However, the study also found that, for 91% of the paths that dropped them, the drops occurred at the source or destination AS.

Reinterpreting this result in a measurement context, a host that can send RR packets without being filtered locally can likely reach most destinations that support the Option. Further, even if the packets are filtered somewhere along the path between the source and destination, they can potentially provide useful partial path measurements.

The drawbacks of Record Route may be diminishing as the Internet evolves. Given this potential as a measurement tool, we believe that recent trends motivate the need to reevaluate support for RR on the Internet. Increased peering means that parts of the Internet such as clouds and colocation facilities may be more richly interconnected [1, 3, 14]. Vantage points in or near these locations may be able to reach many destinations within the nine hop limit of Record Route. Large content and cloud providers have especially short paths [3] and could use RR to measure the paths from end-users to their networks [11], which is necessary to improve performance [13].

Goal: *Given the demonstrated advantages of Record Route, we reassess the coverage of Record Route support on today's Internet. Since some trends suggest that coverage may have improved over time, we hope to show that Record Route is a widely-supported, useful measurement primitive; present approaches to mitigate some of its limitations; and thereby revise the conventional wisdom.*

3 Results: Does The Internet Support Record Route?

Our measurements (explained in §3.1) reveal that the answer is often "yes." The majority of destinations we probed respond to RR (§3.2), a large fraction of these destinations are within RR's 9 hop limit of at least one of our vantage points (§3.3), and, over time, this fraction has increased (§3.4). We find no evidence of ASes that systematically forward packets without recording IP addresses (§3.5). Finally, the interconnectivity of large cloud providers suggests that they could achieve even better coverage with RR than our vantage points (§3.6).

3.1 Dataset, Methodology, and Terminology

The primary dataset we used for investigating the usefulness of Record Route is composed of results from two measurement studies.[1] Each study sent probes to the same destination set that included 1 IP address in each advertised BGP prefix collected by RouteViews on September 24, 2016[19]. For each prefix, the set includes the address that was most responsive to previous ping probes [7]. We conducted the first study during September 24–25, 2016 and sent one ping with the Record Route Option (henceforth: `ping-RR`) from 141 vantage points (VPs). The VPs included one randomly chosen machine at each operational PlanetLab (55) and M-Lab (86) site. We conducted the second study in early October 2016 and sent three pings (without any IP Options enabled) to each destination from one machine at USC. Both studies used `scamper` [15] to send probes and sent 20 probes per second per machine. Section 4.1 explores the impact of probing rate.

We classify a destination as *ping-responsive* if we received a response to at least one of the three normal pings. We classify a destination as *RR-responsive* if at least one VP received a response to its `ping-RR`. A RR-responsive destination responds to a `ping-RR` with an ICMP Echo Reply that copies the Record Route option with any recorded IP addresses into the header of its response. We define a destination as *RR-reachable* from a given VP if a `ping-RR` sent from the VP arrives at the destination with empty slots available in the RR header. Sometimes, we refer to a destination as RR-reachable from our set of VPs, meaning that it is RR-reachable from at least one VP. For our analyses, we test if a destination is RR-reachable by observing if the destination IP address appears in the RR response header. This

---

[1] All datasets and tools used in this study have been made publicly available at: https://www.measurementlab.net/publications/#the-record-route-option-is-an-option

		Total	Transit/Access	Enterprise	Content	Unknown
By IP	All Probed	**510,305 (100%)**	388,959 (100%)	61,204 (100%)	44,295 (100%)	15,847 (100%)
	Ping Responsive	**394,644 (77%)**	296,011 (76%)	51,579 (84%)	37,299 (84%)	9,755 (62%)
	RR-Responsive	**296,734 (58%)**	225,000 (58%)	34,917 (57%)	28,786 (65%)	8,031 (51%)
By AS	All Probed	51,920 (100%)	19,888 (100%)	24,920 (100%)	2,250 (100%)	4,862 (100%)
	Ping Responsive	49,100 (95%)	19,282 (97%)	23,454 (94%)	2,198 (98%)	4,166 (86%)
	RR-Responsive	40,545 (78%)	17,250 (87%)	17,876 (72%)	1,960 (87%)	3,459 (71%)

Table 1: *Response rates for pings with/without RR, both total and by AS type. Top shows all probed IP addresses. Bottom counts ASes with at least one IP address. Of 394,644 ping-responsive IP addresses, 296,734 (75%) also respond to RR. Of 49,100 ASes with at least one ping-responsive destination, 40,545 (82%) also contain at least one RR-responsive destination.*

test allows for some false negatives, which we explore further at the end of Section 3.3.

3.2 Do Destinations Respond to RR?

The responsiveness results of our measurement study are shown in Table 1. The bolded entries show the number of destination that were probed, the number that were ping-responsive, and the number that were RR-responsive. Of all destinations, 77% are ping-responsive and 58% are RR-responsive, meaning that 75% of destinations that respond to ping also respond to `ping-RR`.

Because different AS types may have different policies and because different sets of researchers are interested in the behavior of different AS types, Table 1 reports responsiveness results by destination AS type from CAIDA [23]. We find that there is not a substantial difference across AS types and that the high rate of response holds across all types: the ratio of $\frac{\text{RR-responsive}}{\text{ping-responsive}}$ addresses for each type is over 0.67.

We initially suspected that some ASes would implement AS-wide policies to filter RR. To investigate, we group the destinations by AS and classify an AS as ping-responsive or RR-responsive if at least one address in the AS responded. The bottom three rows of Table 1 show that 95% of ASes had at least one ping-responsive destination, and 78% had at least one RR-responsive destination, meaning that 82% of the ASes that are ping-responsive are also RR-responsive. This result demonstrates that, while some AS-wide filtering occurs, most ASes are RR-responsive and thus do not filter at this granularity.

Our classification of a destination as RR-responsive required a `ping-RR` response to one or more VPs. We investigated the distribution of the number of VPs that received a response from each RR-responsive destination. Roughly 80% of destinations that responded to at least one VP responded to over 90. This result is consistent with the finding from previous work that filtering of Options packets mainly occurred in a small number of edge ASes [8].

3.3 Are Destinations Within the 9 Hop Limit?

We now investigate the utility of the RR Option according to its ability to measure forward and reverse paths. Our primary reachability finding is that 66% of RR-responsive destinations are RR-reachable, i.e., they are within nine hops of their closest VP, and that VP can measure its complete forward path to the destination. We further find that *at least one vantage point was within 8 hops of nearly 60% of RR-responsive destinations, offering the potential to measure the reverse path from them to any other vantage point [11].*

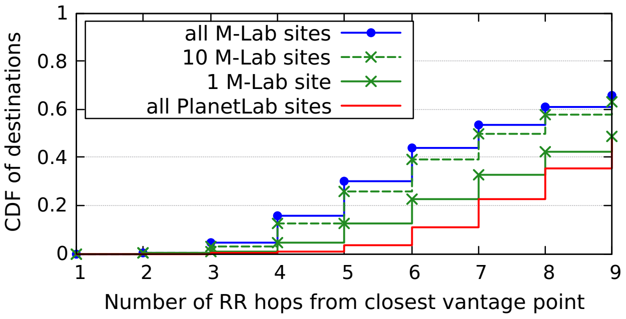

Figure 1: *RR hops from closest vantage point (in various sets) to RR-responsive destinations. M-Lab vantage points are closer than PlanetLab, and 10 M-Lab sites can provide most of the benefit.*

Comparing reachability across sets of vantage points. Using *both* PlanetLab and M-Lab requires operational overhead from researchers. When designing systems or conducting studies where reachability from any VP is a key metric (as is the case when measuring reverse paths [11]), it is helpful to know the usefulness of using both platforms. Figure 1 shows the distribution of the distance from the RR-responsive destinations to the closest VP from various subsets of VPs. To avoid clutter, we omit the line corresponding to the full set of M-Lab and PlanetLab VPs; it is within 1% of the *all M-Lab sites* line at all points, indicating that PlanetLab provides little added benefit over M-Lab for this use case. While M-Lab VPs are within range of 99% of the full set of RR-reachable destinations, PlanetLab VPs are only within range of 72%. We suspect a disparity in site placement is responsible—M-Lab VPs are in centrally-located transit networks and colocation facilities, while most PlanetLab VPs are hosted in university networks. Strategically choosing vantage points from other measurement platforms, such as RIPE Atlas [18], could further improve coverage into networks out of range of M-Lab. However, Atlas currently does not allow measurements with IP Options, and their strict rate limits could complicate the process of finding VPs in range of particular destinations [5].

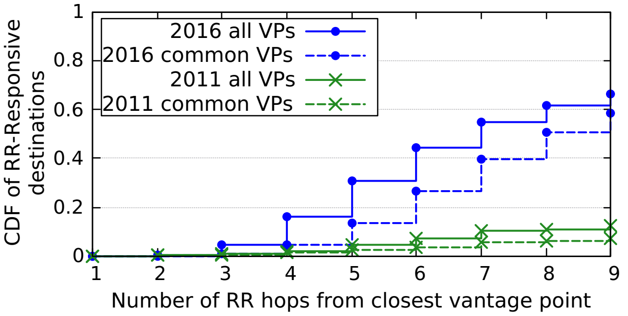

Figure 2: *RR hops from closest M-Lab or PlanetLab vantage point to RR-responsive destinations, in 2011 versus 2016. In 2016, vantage points tend to reach a higher fraction of RR-responsive destinations within 9 hops.*

In systems where the key metric is distance from the closest VP, it can be tempting to issue batches of probes from all VPs. Such exhaustive probing techniques introduce large numbers of RR packets into a network, which may trigger abuse reports or rate limiters (discussed in Section 4.1). In search of more prudent probing methodologies, we investigated the trade-off in reachability that comes with restricting ourselves to small sets of VPs. By greedily selecting M-Lab sites, we find that we can reach 73% of RR-reachable destinations with just one site (NYC), 82% with two sites (adding LA), 86% with three (Denver), 91% with five (Miami and Milan), and 95% with 10. We added lines in Figure 1 comparing some of these subsets. Not only are 95% of RR-reachable destinations reachable with only 10 VPs, but the majority of destinations are reachable at the same distance they would be if we were to use the full set. By carefully choosing small subsets of VPs, we can greatly reduce probing load while maintaining a high probability of finding a VP of minimal distance to any RR-reachable destination.

Uncovering Additional Reachability. As noted in Section 3.1, our analysis determined if destinations were RR-reachable by checking if the destination IP address appears in the RR response header. This allows for some false negatives. In this section, we introduce two situations in which destinations may have been falsely classified, then describe and briefly evaluate two tests to account for them.

In the first situation, a destination device could record a different IP address than the one we target [20], a so-called alias. To be able to identify some of these cases, we performed alias resolution on the 1,079,779 IP addresses that were RR-responsive destinations and/or appeared in the RR headers of our measurements. Using MIDAR [12] we uncovered 48,937 alias sets composed of 205,017 unique IP addresses. With this data we inferred that 5,637 of the destinations recorded an alias in the RR header, but never recorded the destination, thus they are RR-reachable.

In the second situation, probes may reach the destination with RR slots remaining, but the response RR header does not contain any IP address from the destination because the destination does not honor RR, a case mentioned in previous work [20]. To detect such instances, we first composed a set of destinations containing all IP addresses determined to be RR-responsive but not RR-reachable from any VP. Next, we used `scamper` [15] to send pings to high-numbered UDP ports with the RR Option enabled (henceforth `ping-RR`$_{\texttt{udp}}$), with the intent to trigger "port unreachable" error responses. Typically, when devices generate error messages, they will quote the offending packet in the response, including the contents of the offending IP header [16]. By viewing the RR hops recorded in the quoted packet header, we can determine if the offending packet arrived at the destination with RR slots available, and thus that the destination must not honor RR. This methodology allows us to reclassify an additional 4,358 destinations as RR-reachable.

Using these two techniques, we were able to reclassify a total of 9,995 destinations as RR-reachable, meaning we could potentially measure their entire forward and reverse paths. We note that the design of more creative techniques to extract useful information from RR-based measurements is an open area of research [17], and the utility of the RR Option increases with each new application.

3.4 Has Reachability Changed Over Time?

We hypothesized that two trends may mean that, collectively, today's available VPs are in range of more RR-responsive destinations than in the past. First, M-Lab now includes machines in a greater number of locations. Second, M-Lab sites tend to be hosted in colocation facilities, where an increase in Internet peering over time has the potential to create shorter paths to more destinations [14, 3].

We compare our 2016 data, which uses 55 Planetlab sites and 86 M-Lab sites to measure 296,734 RR-responsive destinations, to 2011 measurements from 294 Planetlab sites and 14 M-Lab sites to 3,506,984 RR-responsive destinations [11] to. Both sets of destinations were chosen from a contemporaneous list of historically ping-responsive addresses [7], but the 2011 destinations also included IP addresses harvested from `ping-RR` probes [11], guaranteeing that they are RR-reachable from at least one vantage point. We include all destinations in our results (not just IP addresses or prefixes that were common across both dates) because the large gap between the measurements means that even common addresses may have been repurposed or relocated. The *all VPs* lines in Figure 2 show an increase in the fraction of RR-responsive destinations that were RR-reachable from 0.12 in 2011 to 0.66 in 2016. The *common VPs* lines in the same figure indicates a similar increase even if we just consider the 34 PlanetLab and 11 M-Lab sites that were used in both years, suggesting that changes to the set of available VPs cannot alone account for this difference, and that individual VPs are "closer" to more destinations than they were in the past.

3.5 Do ASes Refuse to Stamp Packets?

One potential drawback when using RR is missing hops due to routers not *honoring* RR, since recent recommendations for best practices suggest that routers forward packets without recording their IP addresses [9]. Given these recommendations, we wondered whether some ASes had globally configured their routers to forward packets without stamping them. Here we compare RR paths to traceroute paths at coarse granularity to uncover evidence of such ASes. We use this analysis as a proxy to estimate the accuracy of RR at the level of AS hops.

One approach to discovering when IP hops have been missed due to routers not honoring RR is to align IP paths recorded using RR to corresponding traceroutes. However, previous attempts have shown this to be difficult [20]. Since we were interested in AS-wide behavior, we took a different approach. To test if any ASes systematically refuse to honor RR, we issued traceroutes from each M-Lab vantage point to that VP's RR-reachable destinations (choosing 10,000 randomly for VPs with more than that). In total, 130,000 distinct destinations were considered. By restricting our comparison only to RR-reachable destinations, we were able to avoid the problem of determining which fraction of the traceroute path was present in RR (a subset of the path alignment problem).

We compared the AS paths derived from these traceroutes with those from the corresponding `ping-RR`s. If an AS consistently appears in traceroutes, but not in RR, we would have evidence suggesting a global configuration of routers within that AS to not honor RR. Of 7,185 ASes extracted from these measurements, only two appeared in traceroute but never RR; 143 were usually seen in both, but not always; and the vast majority, 7,040 were always present in the RR path if they were in the corresponding traceroute. This evidence suggests that operators are not adopting AS-wide policy to forward RR packets without stamping them. Furthermore, given the high probability that an AS appears in both traceroute and RR, we are confident that RR is accurate at the granularity of AS hops.

3.6 Could RR Be Useful to Cloud Providers?

Large content and cloud providers are expanding their infrastructures to bring content closer to users [3]. As their paths shorten, nine hops will represent a larger fraction of the round-trip path, and `ping-RR` could become an effective way for providers to measure paths back from their users to their networks (where they lack visibility [13]).

Because we do not have the ability to issue `ping-RR`s from a cloud provider,[2] we instead estimate the number of destinations potentially RR-reachable from three cloud providers using traceroutes we issued in August 2015 to destinations around the world. First, we examine traces we issued in May 2017 from M-Lab VPs to RR-reachable destinations to give a rough estimate of the distribution of traceroute path lengths to RR-reachable destinations (a topic further explored in Section 4.1). Next, for each cloud provider, we calculate a distribution of traceroute path lengths to compare to the M-Lab traceroute length distributions, to roughly calibrate whether the distributions suggest that many are RR-reachable. To do so, we first select all destinations from the 2015 traceroutes that are in the same /24 prefix as a destination that is RR-responsive in 2017, since destinations in a prefix generally share similar paths from a vantage point. By equating destinations in the same /24, we can expand the set of destinations for which we have measurements in both M-Lab `ping-RR`s, and cloud-issued traceroutes. Based on our experience issuing measurements while working at large cloud providers, we assume it is feasible to tunnel the packet to the edge of the cloud provider's AS without using any RR hops, and so we count the length of the traceroute starting at the first hop outside of the cloud provider's AS.

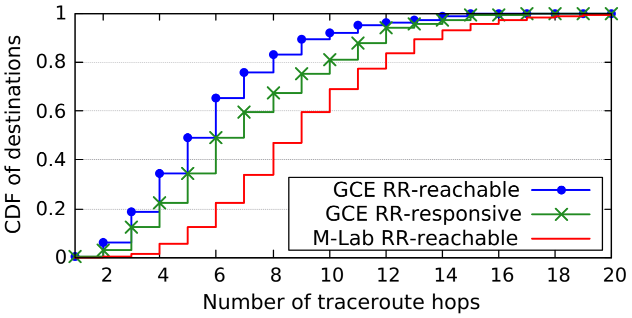

Figure 3: *Hop count from GCE and M-Lab to RR-reachable and RR-responsive destinations. GCE is closer to RR-responsive destinations than M-Lab is to RR-reachable ones, meaning it is likely in range to measure reverse hops.*

Figure 3 focuses on one cloud provider, Google Compute Engine (GCE). It shows the path length distributions of two sets of destinations—41,000 RR-reachable destinations, for which we have both M-Lab and GCE traceroutes, and 263,000 RR-responsive (but not RR-reachable) destinations, for which we have only GCE traceroute path lengths. In total, the set of RR-reachable destinations appear to be significantly closer in terms of traceroute hops to GCE than they are to the set of M-Lab vantage points. Given that these destinations are known to be within nine RR hops of M-Lab, we would expect a great many (if not all) of them to be within nine RR hops of GCE as well. Additionally, 49% of RR-responsive destinations are within 5 traceroute hops of GCE—a shorter distance than that between nearly 80% of M-Lab vantage points and their known RR-reachable destinations, meaning that a large fraction of these may be close enough for GCE to measure reverse paths. While GCE had the shortest path distribution, we also found Amazon EC2 to be within 8 RR hops of 40% of the same RR-responsive destinations and IBM Softlayer within 8 hops of 45%. These results suggest that, should cloud providers adopt RR, they would likely make good vantage points. Specifically in the case of Google, RR could help solve some of its need to uncover the paths from end-users to Google [13].

4 Mitigating Rate Limiting

Router configurations sometimes limit the allowable rate of Options packets to ten per second [4], which can severely hamper probing efforts, particularly when probing large sets of destinations from multiple vantage points. We briefly quantify the impact of rate limiting (§4.1), then consider an approach to avoid destination-proximate rate limiting by sending probes with limited TTL values (§4.2).

4.1 Finding Evidence of Rate Limiting

In Section 3, we probed at 20 packets per second (pps) in a loose attempt to limit the impact of rate limiting. To see whether we could have probed faster without impacting our results, we first randomly selected a set of 100,000 destinations previously deemed RR-responsive. Next, we probed these destinations from all vantage points at 100pps and 10pps. As with our original study, each VP

---

[2] GCE strips options headers from `ping-RR`s issued from VMs within the network, while Amazon EC2 and IBM Softlayer (now Bluemix) both outright filter `ping-RR`s.

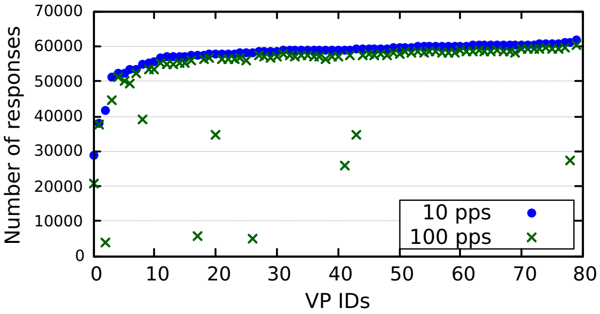

Figure 4: *Number of RR responses received by 79 vantage points at two different probing rates. While most VPs receive similar numbers of responses at both rates, some receive vastly fewer responses at 100pps due to source-proximate rate limiters.*

probed the destination set in random order. Figure 4 depicts the number of responses each VP received at each rate, excluding 56 VPs that received fewer than 1000 responses at either rate. While most vantage points receive only slightly fewer responses at 100pps, a few experience a drastic drop. For eight nodes (five Planetlab and three M-Lab), the response dropped by more than 25%. Our results suggest that rate limits (up to 100pps) have a limited impact on response rates for most VPs in our study, meaning it is not a significant factor in limiting the utility of RR measurements. Further, VPs with lower rate limits are easy to detect and can be configured to use lower VP-specific probing rates to achieve high response rates.

4.2 Choosing Low-Impact TTLs

A `ping-RR` accrues no additional value once the nine slots are full. However, the packet's IP Option will still have to be processed by the route processor of every remaining router on the forward and reverse path, leaving it susceptible to rate limiting or filtering, and incurring wasted processing on routers' slow paths [10].In the rate-limiting study in Section 4.1, we randomized the order of destinations at each VP to avoid triggering rate limiters closer to the destinations. However, there may be times when it is necessary to probe sets of destinations that are similarly located. In these cases, we propose to limit the TTL of the initial `ping-RR`, such that it is highly likely to expire around the time the RR slots are exhausted. The expiring packet will trigger a TTL Time Exceeded error message which will *not* have the RR Option enabled, but the original RR Option will be in the header of the packet quoted inside the error message, allowing us to read it at the source. While at first glance the right answer might seem to be a TTL of nine, there are routers that support RR but do not decrement the TTL or do not send TTL expired errors (i.e., anonymous routers[21]), and there are routers that decrement TTLs but do not stamp RR [20]. Too low a TTL leaves RR slots unused; too high risks rate limiting or filtering.

To study the trade-off, we issued a round of `ping-RR`s to an equal number of RR-reachable and non-RR-reachable, RR-responsive destinations per vantage point (i.e., each vantage point probed a set near it and a set far from it), with randomly assigned TTLs between 3 to 23 or the standard default TTL (64). In Figure 5, we show response rate separately for RR-reachable and non-RR-reachable destinations, grouped by initial TTL. For the former, we want to set a TTL such that most `ping-RR`s reach the destinations. For the latter, we want to minimize the number of hops that the `ping-RR`s traverse after filling their nine slots.

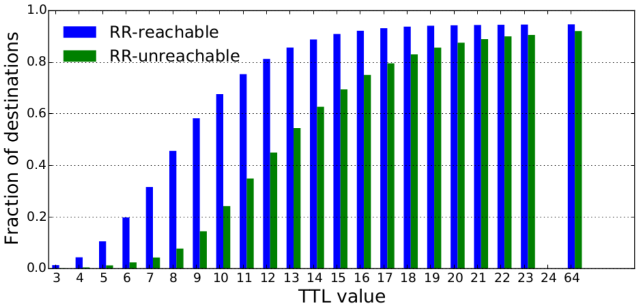

Figure 5: *Responsive rate for RR-reachable and RR-unreachable destinations with different TTL values. TTLs in the range of 10 to 12 offer a good tradeoff in allowing probes to distant destinations to expire while still reaching most RR-reachable destinations.*

For TTL smaller than 8, less than half of RR-reachable destinations respond, which means most of the time we would fail to retrieve complete path information due to a premature timeout. When TTL equals to 10, roughly 70% of the previously reachable destinations are responsive, while only 25% of the previously unreachable destinations respond. Above 12, we receive responses from the majority of non-RR-reachable destinations, which means we lose the benefit of expiring ineffective measurements. Though the exact tradeoff between efficiency and coverage depends on the application, the graph shows that setting TTLs between 10 and 12 could substantially reduce the impact of RR probes on routers while still reaching most RR-reachable destinations. This result implies that one can effectively TTL-limit `ping-RR` probes to mitigate their adverse impact on routers.

5 Conclusion

This paper addresses the question: "Is the IP Record Route Option useful for conducting Internet path measurements?" We found that, contrary to conventional wisdom, the answer is yes. Our measurements show that:

- The majority of ping-responsive IP addresses and ASes respond to RR probes, and their responses can be recorded from most of our PlanetLab and Measurement Lab vantage points.
- A large fraction of IP addresses that respond to RR are reachable within 9 hops of at least one of our vantage points, and a majority of those are within the 8 hops needed to measure reverse paths.
- Large cloud providers, like Google, Amazon, and IBM are sufficiently close to most destinations to record their forward and reverse paths, which they can use to diagnose and improve performance for clients.
- There is no evidence that, of ASes that do not filter RR packets, any systematically refuse to stamp them.

This level of Record Route support may not reflect conscious choices of network operators worldwide, but may be due to the fact that RR, and IP options more generally, have seen little use. Should there be a wide-scale increase in RR traffic, it is possible that some operators might configure routers within their networks to filter or refuse to stamp packets with RR enabled, leading to decreases in responsiveness and reachability. For this reason, we suggest exercising prudence when adopting RR for use in measurement systems and studies. However, there is reason for optimism. Our Reverse Traceroute system, which was well-received at the time of publication in 2010 [11], has been using RR for daily operation and related studies consistently over the past nine years. Our measurements suggest that support has not dwindled in the face of our traffic. Moving forward, we hope that the need for effective topological measurement tools outweighs competing concerns, so that researchers may continue to discover new and better uses of the Record Route Option.

Acknowledgments

We would like to thank our shepherd, Brian Trammell, for his feedback and guidance in the revision process, as well as the anonymous reviewers. We thank Rajiv Gandhi, whose tireless dedication to his students made this work possible. The authors are supported through NSF awards CCF-1433220, CCF-1218620, CNS-1405871, and CNS-1406042.

References

[1] Bernhard Ager, Nikolaos Chatzis, Anja Feldmann, Nadi Sarrar, Steve Uhlig, and Walter Willinger. "Anatomy of a large European IXP". In: *Proc. of SIGCOMM*. 2012.

[2] Brice Augustin, Xavier Cuvellier, Benjamin Orgogozo, Fabien Viger, Timur Friedman, Matthieu Latapy, Clémence Magnien, and Renata Teixeira. "Avoiding traceroute anomalies with Paris traceroute". In: *Proc. of IMC*. 2006.

[3] Yi-Ching Chiu, Brandon Schlinker, Abhishek Balaji Radhakrishnan, Ethan Katz-Bassett, and Ramesh Govindan. "Are We One Hop Away from a Better Internet?" In: *Proc. of IMC*. 2015.

[4] *Cisco Router Setting Recomendations*. http://www.cisco.com/c/en/us/about/security-center/copp-best-practices.html.

[5] Ítalo Cunha, Pietro Marchetta, Matt Calder, Yi-Ching Chiu, Brandon Schlinker, Bruno VA Machado, Antonio Pescapè, Vasileios Giotsas, Harsha V Madhyastha, and Ethan Katz-Bassett. "Sibyl: A Practical Internet Route Oracle." In: *NSDI*. 2016, pp. 325–344.

[6] Benoit Donnet, Matthew Luckie, Pascal Mérindol, and Jean-Jacques Pansiot. "Revealing MPLS tunnels obscured from traceroute". In: *ACM SIGCOMM CCR* 42.2 (2012).

[7] Xun Fan and John Heidemann. "Selecting Representative IP Addresses for Internet Topology Studies". In: *Proc. of IMC*. 2010.

[8] Rodrigo Fonseca, George Porter, R Katz, Scott Shenker, and Ion Stoica. *IP options are not an option*. Tech. rep. UC Berkeley, 2005.

[9] F. Gont, R. Atkinson, and C. Pignataro. *Recommendations on Filtering of IPv4 Packets Containing IPv4 Options*. BCP 186. RFC Editor, 2014.

[10] Ramesh Govindan and Vern Paxson. "Estimating router ICMP generation delays". In: *Passive & Active Measurement (PAM)*. 2002.

[11] Ethan Katz-Bassett, Harsha V Madhyastha, Vijay Kumar Adhikari, Colin Scott, Justine Sherry, Peter Van Wesep, Thomas E Anderson, and Arvind Krishnamurthy. "Reverse traceroute". In: *Proc. of NSDI*. 2010.

[12] Ken Keys, Young Hyun, Matthew Luckie, and Kim Claffy. "Internet-scale IPv4 alias resolution with MIDAR". In: *IEEE/ACM TON* (2013).

[13] Rupa Krishnan, Harsha V Madhyastha, Sridhar Srinivasan, Sushant Jain, Arvind Krishnamurthy, Thomas Anderson, and Jie Gao. "Moving beyond end-to-end path information to optimize CDN performance". In: *Proc. of IMC*. 2009.

[14] Craig Labovitz, Scott Iekel-Johnson, Danny McPherson, Jon Oberheide, and Farnam Jahanian. "Internet Inter-domain Traffic". In: *Proc. of SIGCOMM*. 2010.

[15] Matthew Luckie. "Scamper: A Scalable and Extensible Packet Prober for Active Measurement of the Internet". In: *Proc. of IMC*. 2010.

[16] David Malone and Matthew Luckie. "Analysis of ICMP quotations". In: *International Conference on Passive and Active Network Measurement*. Springer. 2007, pp. 228–232.

[17] Pietro Marchetta, Valerio Persico, Giuseppe Aceto, Alessio Botta, and Antonio Pescape. "Measuring Networks Using IP Options". In: *IEEE Network* (2017).

[18] *RIPE Atlas*. https://atlas.ripe.net/.

[19] Oregon RouteViews. "University of Oregon RouteViews project". In: *Eugene, OR.[Online]. Available: http://routeviews.org/bgpdata/2016.09/RIBS/rib.20160924.1200.bz2* ().

[20] Rob Sherwood, Adam Bender, and Neil Spring. "Discarte: a disjunctive internet cartographer". In: *Proc. of SIGCOMM* (2008).

[21] Rob Sherwood and Neil Spring. "Touring the Internet in a TCP sidecar". In: *Proc. of IMC*. 2006.

[22] Joel Sommers, Paul Barford, and Brian Eriksson. "On the prevalence and characteristics of MPLS deployments in the open Internet". In: *Proc. of IMC*. 2011.

[23] *The CAIDA AS Classification Dataset*. www.caida.org/data/as-classification/.