Does It Spin?
On the Adoption and Use of QUIC's Spin Bit

Ike Kunze
kunze@comsys.rwth-aachen.de
RWTH Aachen University
Germany

Constantin Sander
sander@comsys.rwth-aachen.de
RWTH Aachen University
Germany

Klaus Wehrle
wehrle@comsys.rwth-aachen.de
RWTH Aachen University
Germany

ABSTRACT

Encrypted QUIC traffic complicates network management as traditional transport layer semantics can no longer be used for RTT or packet loss measurements. Addressing this challenge, QUIC includes an optional, carefully designed mechanism: the spin bit. While its capabilities have already been studied in test settings, its real-world usefulness and adoption are unknown. In this paper, we thus investigate the spin bit's deployment and utility on the web.

Analyzing our long-term measurements of more than 200 M domains, we find that the spin bit is enabled on ~10% of those with QUIC support and for ~ 50% / 60% of the underlying IPv4 / IPv6 hosts. The support is mainly driven by medium-sized cloud providers while most hyperscalers do not implement it. Assessing the utility of spin bit RTT measurements, the theoretical issue of reordering does not significantly manifest in our study and the spin bit provides accurate estimates for around 30.5 % of connections using the mechanism, but drastically overestimates the RTT for another 51.7 %. Overall, we conclude that the spin bit, even though an optional feature, indeed sees use in the wild and is able to provide reasonable RTT estimates for a solid share of QUIC connections, but requires solutions for making its measurements more robust.

CCS CONCEPTS

• **Networks → Network monitoring**; **Network measurement**.

KEYWORDS

QUIC; spin bit; Internet measurements

ACM Reference Format:
Ike Kunze, Constantin Sander, and Klaus Wehrle. 2023. Does It Spin? On the Adoption and Use of QUIC's Spin Bit. In *Proceedings of the 2023 ACM Internet Measurement Conference (IMC '23), October 24–26, 2023, Montreal, QC, Canada.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3618257.3624844

1 INTRODUCTION

Network operators require visibility into their networks to locate and fix faults and performance bottlenecks. Passive flow measurements are one commonly used source of information for this purpose: they track traffic already passing through the network without adding additional overhead to the production traffic. For example, TCP semantics can be leveraged to estimate flow round-trip times (RTTs) [29] or track reordering [32] and packet loss [5].

Such passive methods are, however, challenged by the rising share of encrypted transport protocols. QUIC, e.g., denies passive observers access to the protocol header fields, making the mentioned TCP methods neither applicable nor adaptable to QUIC. Instead, there is a push toward adding well-defined, explicitly measurable signals to protocols that can then be analyzed by passive observers [4, 9]. The QUIC spin bit is one such mechanism, enabling explicit RTT measurements on otherwise encrypted traffic [14].

While the addition of the spin bit is a promising sight for operators hoping to regain measurability of their networks, there is one significant hindrance: the spin bit is only an optional feature, i.e., QUIC implementations do not have to support it. Inspecting the publicly available QUIC stacks [2] already reveals that only few implement the spin bit. Furthermore, the spin bit has mainly been studied in test settings [6, 10]. Whether or not network operators can rely on the spin bit for accurate RTT estimation in the wild is thus an important open question and there is little publicly known information on these practical issues.

In this paper, we, thus, set out to investigate the adoption and accuracy of the spin bit in the wild. Performing long-term, large-scale measurements of more than 200 M domains, including many popular targets of toplists, we study the host support of the spin bit on the web, finding that it is active on ~10 % of domains that support QUIC and mostly driven by small- to medium-sized providers. Judging the utility of spin bit-based measurements by comparing them to QUIC's built-in RTT estimates, we observe that the spin bit generally overestimates the real RTT as expected, but still provides accurate measurements for around 30.5 % of connections using the mechanism. Overall, we contribute the following:

- We study the QUIC spin bit adoption in the web using repeated measurements of more than 200 M domains.
- We find that around 10 % of domains with QUIC support also support the spin bit with adoption mainly driven by small to medium cloud providers.
- Assessing its measurement accuracy, we observe that the spin bit mostly overestimates the real RTT and yields sensible estimates for a third of the studied cases.

2 QUIC AND ITS SPIN BIT

The spin bit allows passive network observers to determine flow RTTs and can be seen as an answer to Allman et al.'s call for adding explicit measurement mechanisms into protocols [4]. It addresses a need of network operators who, with the rising share of QUIC [24, 34], increasingly lose visibility into the performance

IMC '23, October 24–26, 2023, Montreal, QC, Canada

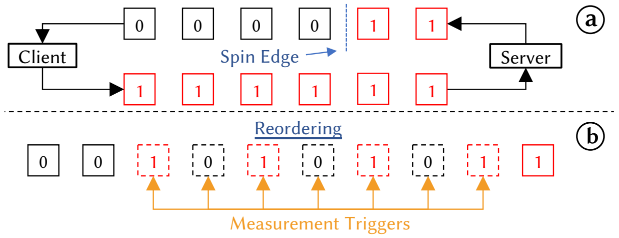

Figure 1: ⓐ: The client of a connection always *flips* the spin bit; the server *reflects* it. ⓑ: Reordering near the spin edges can cause faulty measurements.

of their networks as established measurements relying on implicit protocol semantics, e.g., exposed by TCP, can no longer be used.

2.1 The QUIC Spin Bit

The spin bit is an optional, version-dependent feature of QUIC version 1 that uses one bit in short header packets [14]. In short, it modulates a square wave onto a QUIC flow with a wavelength equal to the flow's RTT measurable by observers.

Mechanism. The client starts the signal by transmitting packets with a value of 0. As visualized in Fig. 1 ⓐ, the server reflects the value it has received, setting the value on outgoing packets to the value seen on the latest incoming packet with the highest packet number. In contrast, the client spins the bit, i.e., it inverts the latest value. Observers can estimate the RTT by measuring the time between two consecutive spin bit flips ("spin edges").

Manageability considerations. As the spin bit needs support from both peers of a connection, either endpoint can unilaterally decide to disable the mechanism. RFC 9000 [14] further mandates that even end-points actively using the spin bit "MUST" disable it on at least one in every 16 connections to ensure that disabled spin bits appear regularly in Internet traffic. One major reason for this decision is that some endpoints may want to disable the spin bit to prevent RTT measurements, and this decision should not be cause for any mistreatment by the network [15]. Whether these guidelines are indeed followed in practice is unknown.

Disabling the spin bit. RFCs 9000 and 9312 [14, 15] recommend that end-points disable the bit via greasing, i.e., setting it randomly on a per-packet or per-connection ID basis. They can further set it arbitrarily, meaning that using the same value on all connections is also possible. Hence, despite the recommendations, it is unknown which form of disabling is commonly chosen on the web.

Accuracy considerations. Spin bit RTT measurements are vulnerable to different causes of inaccuracy. On the one hand, the spin bit always includes end-host delays, e.g., the time needed to process a request, which might cause inflated RTT estimates. On the other hand, the measurements can be disturbed by packet reordering as is visualized in Fig. 1 ⓑ. More specifically, reordering around spin edges can introduce ultra-short spin cycles and, hence, very short RTT estimates. While a solution addressing the first issue was originally proposed in the form of the Valid Edge Counter (VEC) [10], it has not found its way into the standard. Addressing the second issue, RFC 9312 [15] proposes the use of heuristics, but the usefulness and necessity of these heuristics has not yet been tested on a larger scale in academic work.

2.2 Related Work on QUIC and the Spin Bit

Studies by Rüth et al. [24] in 2018 and Zirngibl et al. [34] in 2021 show an increasing deployment of QUIC on the Internet and in the web. However, neither study investigates the use of the spin bit, which is the only optional feature of QUIC.

Academic work on the spin bit goes back to De Vaere et al.'s three-bit mechanism [10], which, in addition to the spin bit, included the VEC designed to provide more stable measurements. The authors study the capabilities of the spin bit subject to packet loss and reordering, and compare it to traditional measurements using the TCP timestamp option. While their results show that the spin bit is a useful mechanism comparable to TCP timestamps, the experiments are small-scale and do not consider the validity of spin bit measurements in larger deployment scenarios. Bulgarella et al. [6] confirm the general findings, but again focus on small-scale testbed experiments. Kunze et al. [18] show that the spin bit observer logic can be implemented on P4 hardware switches and that the heuristics of RFC 9312 can filter faulty RTT measurements to some extent.

In summary, QUIC deployment numbers are increasing [24, 34] and the spin bit has been shown to provide accurate latency measurements in small-scale and testbed experiments [6, 10]. However, its adoption in the wild, real-life usage patterns, e.g., regarding the guidelines defined in RFCs 9000 and 9312 [14, 15], and the practical measurement utility, are still unknown.

3 STUDY DESIGN

The adoption and usefulness of the spin bit in the wild are unknown. Hence, to shed light on this topic, we perform a large-scale Internet measurement campaign that focuses on the web.

Approach. We first identify suitable targets for our study via domain top- and zonelists (Sec. 3.1). We then conduct HTTP/3 requests via QUIC from our university network to these targets and our adapted webclient stores relevant connection information in the qlog [22] format (Sec. 3.2). Using these logs, we assess the spin bit adoption and its usefulness for RTT measurements (Sec. 3.3).

3.1 Determining Target Hosts

We assemble our target population using multiple datasets.

3.1.1 *Domain Toplists.*

Relying on different toplists, we first gather a target set that focuses on popular domains: (i) Alexa Top 1M[1] [13], (ii) Cisco Umbrella [31], (iii) Majestic Million [21], and (iv) the Tranco Research List [20]. We further deduplicate their entries.

3.1.2 *Domain Zonelists.*

For a broader assessment and to also include less popular domains, we enrich our target population with domains from zone files accessible via the ICANN Centralized Zone Data Service (CZDS). For our measurements in calendar week (CW) 20, 2023, we use 1140 individual zones, including various gTLDs such as .com, .net, and .org. Combined, the top- and zonelists amount to a total of around 219 M domains for CW 20, 2023.

[1] The Alexa list is deprecated since 05-2022; it stopped changing in 02-2023

3.2 Querying the Target Population

We perform HTTP/3 requests to retrieve the landing webpages of the target population.

3.2.1 Tooling. We use an adapted version of zgrab2 [3], a golang-based banner grabber, as the general framework for our measurements. Underneath, we add *quic-go* [1], which supports QUICv1, and extend it for draft versions 27, 29, 32, and 34, and the spin bit. We prepend each domain with "www" for targeting websites, as is common practice [25, 33], and capture quic-go's qlog [22] information (which we extended with the spin bit state) for each established connection. Note that there can be more than a single connection for one domain, e.g., due to redirects; to limit the impact of our measurements, we only follow up to 3 redirects. Given our web-focused view, we further assume that the queried domains provide websites. Hence, we cannot assess the behavior of domains not responding to HTTP/3.

3.2.2 Vantage point and measurement schedule. We conduct our measurements from within the RWTH Aachen University network which connects to the well-connected German research network (DFN), e.g., peering at DE-CIX with many ASes. In general, we perform weekly measurements using IPv4. We further perform measurements using IPv6 in selected weeks with the same methodology to allow for comparisons between IPv4 and IPv6 while still reducing the load of our measurements. Our zonelist measurements start on Wednesdays and run through Fridays while the toplist measurements start on Fridays and finish by Saturdays.

3.3 Studying the Target Population

We assess the behavior of the target population by systematically analyzing the collected qlog data.

QUIC support. We first identify hosts supporting QUIC by checking whether connections could be established with remote endpoints. For this, we check whether the endpoints answer to QUIC packets and whether the actual initialization process finishes.

Spin bit support. For hosts with QUIC support, we focus on the received packets from the qlog and extract (i) the spin bit state, (ii) the QUIC packet number, and (iii) the corresponding timestamp. We identify connections that see spin bit flips, i.e., values of both 0 and 1, as candidates that might support the spin bit.

RTT assessment and grease filter. For the identified candidate connections, we use the spin bit mechanism to calculate the RTTs based on the received order of the packets. We then compare these RTTs to estimates provided by the QUIC stack which relies on richer information. In essence, the QUIC stack measures the time until a specific packet is acknowledged and additionally factors in processing delays as reported by the other host. Based on these two kinds of RTT values, we filter out connections that presumably grease the spin bit as soon as one spin bit RTT estimate is smaller than the minimum of all QUIC client RTT estimates.

Impact of reordering. Finally, we also assess the impact of reordering on the spin bit measurements. For this, we correct the ordering of the received packets based on the packet numbers and repeat the RTT calculations.

		Total	Resolved	QUIC	Spin
Toplists	#Domains	2.73 M	1.94 M	547.11 k	6.9 %
	#IPs		774.83 k	118.54 k	15.2 %
CZDS	#Domains	216.52 M	183.74 M	22.21 M	10.2 %
	#IPs		10.27 M	259.77 k	45.3 %
com/net/org	#Domains	183.05 M	158.89 M	18.42 M	11.1 %
	#IPs		9.20 M	242.88 k	46.4 %

Table 1: Overview of our IPv4 results for CW 20, 2023.

4 SPIN BIT USE IN THE WILD

We characterize the use of the spin bit from diverse perspectives, aiming to identify usage patterns, drivers of its deployment, and its utility in the wild. We base our assessment on long-term, large-scale IPv4 measurements (cf. Sec. 3). In this section, we focus on the general spin bit use before we analyze its RTT measurement utility in Sec. 5.

4.1 Dataset and Spin Bit Use Overview

We first give a broad overview on general statistics describing our measurement data and the use of the spin bit. Table 1 shows the results of our latest IPv4 measurements in CW 20, 2023 for the *Toplists*, all *CZDS* lists, and a focused view on the .com, .net, and .org zones (*com/net/org*). General observations identified for CW 20 are also representative for our longitudinal measurements.

Domain view. The respective first rows show the overall number of domains contained in the different domain lists (*Total*), the number of resolved domains (*Resolved*), the number of domains with at least one QUIC connection (*QUIC*), and the percentage of QUIC-enabled domains that are *potentially* spin bit enabled, i.e., see spin bit activity on at least one connection (*Spin*). Looking at the toplists, we can see some spin bit activity for 6.9 % of the corresponding QUIC connections. In comparison, there is a bit higher activity for CZDS (10.2 %) and, especially, com/net/org (11.1 %). Hence, there is small support for the spin bit, but it is not only focused on topdomains often provided by large hyperscalers.

IP view. Drilling deeper into this consideration, we next study the individual IPs that are contacted for each domain to assess the host support for QUIC and the spin bit. Here, we consider all IPs to which there is at least one connection with spin bit activity. Taking a broad look at the results, we can see that the 547 k QUIC-enabled toplist domains are served by ~119 k IPs, i.e., a ratio of 21.7 %, while the corresponding ratios are smaller for CZDS (1.2 %) and com/net/org (1.3 %). Surprisingly, almost 50% of the IPs serving CZDS and com/net/org show spin bit activity while only 15.2 % of the toplist IPs show support. This observation supports our initial conjecture that the main drivers of the spin bit might not be hyperscalers, but instead smaller providers.

Takeaway. *About 10% of CZDS domains and 50% of IPs serving these domains show spin bit activity while there is significantly less support in the toplists of less than 20% of IPs.*

↑	Total #	AS Organization	Spin #	Spin %	Spin ↑
1	11.48 M	Cloudflare	0	0 %	0
2	6.16 M	Google	6.87 k	0.1 %	54
3	1.55 M	Hostinger	802.59 k	51.9 %	1
4	326.23 k	Fastly	0	0 %	0
5	219.25 k	OVH SAS	132.40 k	60.4 %	2
6	218.21 k	A2 Hosting	129.58 k	59.4 %	3
7	173.50 k	SingleHop	102.53 k	59.1 %	4
8	148.71 k	Server Central	100.52 k	67.6 %	5
	2.52 M	<other>	1.34 M	53.3 %	

Table 2: QUIC connections and spin bit activity resolved for the corresponding AS organization for com/net/org and the IPv4 results of CW 20, 2023.

4.2 Drivers of the Spin Bit

Following our intuition that smaller providers might be at the forefront of spin bit support, we next aim to identify the main drivers of the spin bit. For this, we inspect (i) the organizational affiliations of the supporting IPs, and (ii) the webservers serving these connections for com/net/org.

Organizational affiliations. Analyzing the affiliation of all connections, we first map each IP to its corresponding ASN using BGP data of RIPE's RIS archive [23] and then lookup the corresponding organizations using CAIDA's as2org dataset [7].

AS results. Table 2 shows the overall number of connections (*Total #*) and the corresponding overall rank (↑) for the top 5 organizations in terms of connection share. We further add the number and percentage of connections with some spin bit activity (*Spin #*/*Spin %*) and the top 5 organizations in terms of absolute spin bit support as indicated by their rank (Spin ↑). As can be seen, the two organizations with the most connections, being responsible for more than 75 % of QUIC connections observable from our vantage point, show no (Cloudflare) or very little support (Google, rank 54) for the spin bit. Instead, medium-sized hosters that also have a relatively high share of QUIC connections represent most of the spin bit support, each having spin activity on more than 50 % of their connections. Hostinger has the largest absolute share with 802.59 k connections with spin bit support. Interestingly, 53.3 % of the remaining 2.52 M connections that do not belong to the top ASes show spin bit support, indicating that there is a broad base of support.

Confirming our previous intuition, we can conclude that the spin bit is mainly in use by smaller organizations. This stands in contrast to general trends for QUIC for which Zirngibl et al. [34] find hyperscalers to be the main drivers.

Webserver support. Given the broad spin bit support among many ASes, we next dig deeper into which servers support the spin bit. For this, we look for HTTP server information, focusing on those connections where we can unambiguously match this information with the qlog information.

Webserver results. By far the most connections reach *LiteSpeed* webservers, making up more than 80 % of all connections (not shown), while another 7 % are served by *imunify360-webshield*, which we also suspect to build upon *LiteSpeed*. Hence, we conclude that the overwhelming share of spin bit support is provided by LiteSpeed deployments, i.e., likely a single stack.

Takeaway. *Large drivers of QUIC, such as Cloudflare, and Google, do not support the spin bit. Instead, smaller hosters, such as Hostinger, carry the bulk of spin bit support, although there is also large support among smaller ASes. Most of this support can be traced back to LiteSpeed webservers.*

	All Zero	All One	Spin	Grease
Toplists	507.97 k (92.8 %)	859 (0.2 %)	37.77 k	58 (0.0 %)
CZDS	19.85 M (89.4 %)	62.38 k (0.3 %)	2.26 M	5.31 k (0.0 %)
com/net/org	16.28 M (88.4 %)	53.72 k (0.3 %)	2.05 M	4.65 k (0.0 %)

Table 3: Spin behavior of all QUIC domains for our measurements in CW 20, 2023. Spin column corresponds to Table 1.

4.3 Spin Bit Configuration

Having established the sources of spin bit support, we next dig deeper into how the spin bit is configured in practice. RFCs 9000 and 9312 [14, 15] recommend that end hosts should disable the spin bit via greasing, i.e., setting it to random values. Greasing can be done on a per-packet or per-connection basis while using a fixed value is also possible. We next assess which practices are used by analyzing how the spin bit is set in the wild.

How is the spin bit disabled? Table 3 shows how the different domains set the spin bit. As can be seen, most domains that do not use the spin bit use a value of zero (*All Zero*) while only few exclusively send a value of one (*All One*). Additionally, we only filter a small number of connections with our simplistic grease filter, indicating that RTT underestimations do not occur frequently.

Is the RFC "MUST" followed? Another critical aspect of RFC 9000 is that the spin bit "MUST" be disabled on one in 16 connections (or one in eight as per RFC 9312) to prevent ossification. We assess this configuration aspect by inspecting the behavior of domains over a longer period leveraging our longitudinal measurement data. In particular, we first select n measurement days spread across our *measurement campaign* (CW 15, 2022 – CW 20, 2023) and gather all domains that had the spin bit enabled on *any* of these days. We then select the domains to which we could establish a connection in *every* week and count in how many weeks they showed spin bit activity. We choose $n = 12$ to include as many weeks as possible while still having a large number of considered domains.

Across all domain lists, we could identify a total of 4.37 M domains that have spun at any time during the 12 selected weeks, out of which only 2.26 M domains had working connections in every week. Fig. 2 shows a relative histogram of the share of domains that had the spin bit enabled in the corresponding number of weeks as well as shares computed based on the RFC descriptions usinng probability theory. As can be seen, slightly less than 20 % of domains spin the spin bit in all weeks while the shares for the other groups are around 5 to 10 % each. Additionally, the domains spin less than would be allowed by the two interpretations of the specification, indicating that the guidelines on regularly disabling the spin bit are indeed followed. We discuss an alternative methodology in Sec. 6.

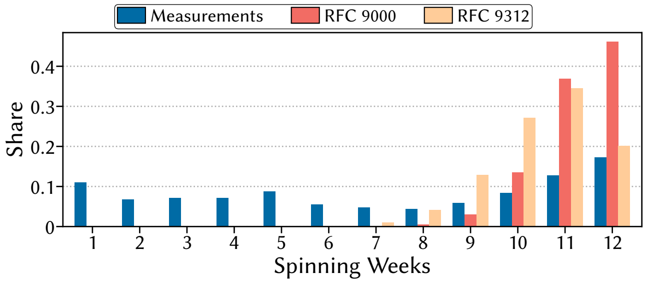

Figure 2: Histogram showing the number of weeks in which the selected domains had the spin bit enabled. RFC values are computed using probability theory and reflect theoretical behavior when the spin bit is disabled in one in eight (RFC 9312) / 16 (RFC 9000) connnections.

Takeaway. *Most hosts seem to opt for zeroing the spin bit. Additionally, greasing the spin bit on a per-packet basis, indicated by very short RTT estimates, only occurs seldomly as indicated by small grease filter numbers. Finally, the long-term behavior of the spin-enabled domains shows that the RFC mandate seems to be generally followed.*

4.4 IPv6 and the Spin Bit

In addition to our IPv4 measurements, we have also conducted IPv6 measurements using the same methodology in some weeks of our measurement campaign to assess the state of the spin bit in the IPv6 space. Table 4 shows the IPv6 results for CW 20, 2023. While the overall number of domains with QUIC support is lower, we can see drastically more IPs supporting QUIC for CZDS and com/net/org, each also having a larger share of connections with spin bit support of more than 60 %. In contrast, there is much lower spin bit support for the toplists of only 2.3 % of domains and 8.1 % of hosts. Similar to the IPv4 results, LiteSpeed servers are responsible for around 80 % of all connections and Hostinger represents a large share of the spinning connections with 91.5 %. Overall, the IPv6 results draw a two-fold picture: on the one hand, more modern IPv6 deployments seem to coincide with a higher spin bit support for CZDS and com/net/org. On the other hand, the toplists, typically a driver for innovation, show a much worse support than with IPv4.

Takeaway. *The spin bit support is broader for IPv6 compared to IPv4, yet worse when only focusing on the toplists. General trends regarding the main drivers of the spin bit are similar to IPv4.*

5 RTT MEASUREMENT ACCURACY

While we have found that flows indeed use the spin bit on the web, the utility of the corresponding RTT estimates is still unclear. In particular, De Vaere et al. [10] have identified different sources of inaccuracy when analyzing the performance of the spin bit in a testbed and a small set of real-world measurements; the impact of these issues on a larger scale and in practice is, however, unknown. Hence, we now leverage our large-scale measurements to shed light on the practical accuracy of spin bit RTT estimates.

		Total	Resolved	QUIC	Spin
Toplists	#Domains	2.73 M	569.52 k	368.33 k	2.3 %
	#IPs		166.13 k	94.53 k	8.3 %
CZDS	#Domains	216.52 M	21.47 M	9.10 M	8.2 %
	#IPs		2.12 M	1.18 M	62.6 %
com/net/org	#Domains	183.05 M	17.03 M	6.63 M	10.2 %
	#IPs		1.85 M	1.04 M	63.6 %

Table 4: Overview of our IPv6 results for CW 20, 2023.

5.1 Assessment Methodology

We base our assessment on the raw data of *all* IPv4-based QUIC connections with spin bit activity throughout our *entire measurement campaign* (cf. Sec. 4.3), amounting to ~86 M connections in total. For each connection, we compare the mean of the spin bit estimates (*spin*) to the mean of the QUIC stack estimates (*QUIC*) using the following two methods.

1) Absolute measurement accuracy. We assess the absolute accuracy of the spin bit by computing the difference between the two means: $abs = spin - QUIC$.

2) Relative measurement accuracy. We assess the relative accuracy of the spin bit by computing the ratio of the means. We always divide by the smaller of the two and multiply the ratio by -1 if $spin < QUIC$. Negative values indicate an underestimation of the real RTT while positive values represent an overestimation.

Terminology. We assess the performance of spinning connections (*Spin*) and connections filtered by our grease filter (*Grease*). We also distinguish results where we use the *received* packet order (*R*), potentially including reordering, and with packets *sorted* according to their packet numbers (*S*) to judge the impact of reordering.

5.2 Spin Bit Accuracy Results

Fig. 3 shows a relative histogram of the mean differences for *Spin* and *Grease*, with (*S*) and without (*R*) correcting the packet order, while Fig. 4 shows the corresponding ratio.

General usability. 97.7 % of the Spin (R) results overestimate the real RTT while there is only a small share of measurements underestimating it. We can further observe that 28.8 % of the connections have an absolute difference of ≤ 25 ms, yet 41.3 % of the connections overestimate the real RTT by more than 200 ms.

To assess the impact of these large absolute differences, we study the corresponding ratios in Fig. 4. As can be seen, the spin bit estimates are very close (less than 25% difference) to the real RTT for 30.5 % of the spinning connections and 36.0 % are within a factor of 2. However, a non-negligible portion of results (51.7 %) significantly overestimates the real RTT by more than a factor of 3. Thus, while many estimates have an acceptable accuracy, the general tendency of the spin bit overestimating the real RTT becomes apparent.

Impact of reordering. Comparing the results of Spin (R) and Spin (S), we can observe that there is almost no perceivable difference between the two. Going into more detail, we find differing results for only 0.28 % of all connections with 98.7 % of these differences having an absolute impact of less than 1 ms, slightly improving the measurement accuracy in 93.1 % of the cases. Hence, while

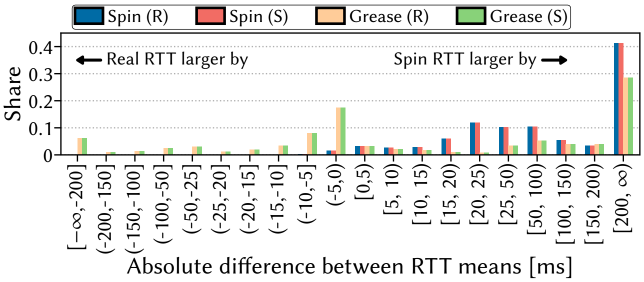

Figure 3: Histogram of abs. difference between means of spin bit and QUIC estimate for all connections.

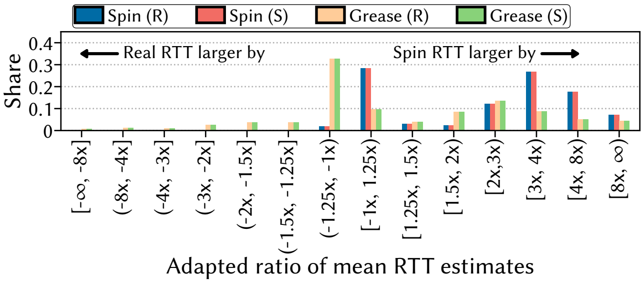

Figure 4: Histogram of mapped ratio of the means of spin bit and QUIC estimate for all connections.

reordering does occur, it does not seem to have a large-scale impact on the measurement accuracy, at least not for our vantage point, indicating that it might be more of a theoretical issue.

Grease filtering. Inspecting the results of Grease (R/S), we suspect that our grease filter might create false positives. While we do notice a larger share (46.0 %) of connections that underestimate the real RTT, there are also many connections significantly overestimating it, although not as many as for Spin (R/S). Furthermore, the relative impact of the underestimations seems to be rather small as 62.5 % of the connections are within a factor of 2. We conjecture that the underestimations are not caused by greasing and suspect reordering to be one culprit, but only find 7.9 % of the connections to be affected. Thus, there seem to be yet unknown effects warranting further analysis which we enable by sharing our underlying spin bit measurement data set (cf. Appendix B).

Takeaway. *The spin bit allows accurate RTT estimates for many connections. An even larger share sees drastic overestimations. While reordering has no pronounced impact on our results, unknown effects cause unexpected underestimation warranting further inspection.*

6 DISCUSSION

Our measurement campaign shows that the spin bit is indeed used in the wild but that it is mainly driven by smaller providers. We have further shed light on RFC compliance regarding disabling the spin bit and we find that many spin bit measurements are rather inaccurate. In this section, we discuss possible reasons for and implications of our findings as well as limitations of our work.

Reasons for disabling the spin bit. One deliberate reason for not using the spin bit are privacy concerns. In particular, using the spin bit exposes the RTT which can theoretically be used for locating users. However, Trammell and Kühlewind [30] have shown that these concerns are negligible. A more pragmatic reason for not using the spin bit could be that QUIC stacks are still in the process of completing the mandated functionality. Given that the spin bit does not provide an immediate benefit for the stack providers, implementing the optional spin bit might just not be a priority.

Measurement accuracy. By design, spin bit RTT estimates include end-host delays which are likely reasons for the identified large absolute differences (cf. Sec. 5.2). These delays are most prominent at connection starts, which our approach focuses on, while measurements tend to stabilize over longer durations [19]. Hence, apart from assessing the usefulness of the spin bit for practical applications, such as network tomography [8], it might be worthwhile to study the spin bit accuracy for longer connections.

Limitations. Due its large footprint and other measurements conducted from our vantage point, our measurement campaign only performs weekly one-shot measurements without the ability to run additional measurements. As a consequence, we use the results of consecutive weeks for assessing RFC 9000 and 9312 compliance. Hence, our results might be affected by longer-term deployment changes as we note when studying the use of ECN in QUIC based on the same dataset [28]. To study RFC compliance more accurately, one could first identify domains with an enabled spin bit in a large-scale measurement and then follow up with multiple measurements of a smaller target set, e.g., querying them n=16 times to more accurately reflect RFC descriptions.

7 CONCLUSION

The spin bit promises to provide network operators with visibility into the latency behavior of their networks in times of rising shares of QUIC traffic. In this paper, we present the results of our measurement campaign investigating the practical use of the spin bit on more than 200 M domains, including many popular targets.

We find that around 10 % of domains with QUIC support also use the spin bit, fueled by ~50 (IPv4) to ~60 % (IPv6) support of the underlying hosts. The support is mainly driven by small to medium cloud providers, each using the spin bit on more than 50 % of their connections. For connections with spin bit support, we find a high measurement accuracy of less than 25 % difference to the baseline for around 30 % of connections while 50 % of connections experience an overestimation by a factor of more than 3.

Concluding, we find a rather high adoption of the spin bit given the low support in QUIC stacks. Our RTT utility assessment further suggests that the spin bit provides sensible capabilities for network operators for a solid share of connections, but that estimates have to be used with caution and can benefit from further research, e.g., studying the usefulness of filtering techniques described in RFC 9312 [15]. Overall, it remains to be seen if implementation support by hypergiants will follow naturally with growing maturity of the QUIC stacks or if it requires strong incentives, e.g., in the form of immediate benefits enabled by the spin bit.

ACKNOWLEDGMENTS

This work has been funded by the German Research Foundation DFG under Grant No. WE 2935/20-1 (LEGATO). We thank the anonymous reviewers and our shepherd Vaibhav Bajpai for their valuable feedback. We further thank Leo Blöcher for maintaining our quic-go / zgrab2 changes and the network operators at RWTH Aachen University, especially Jens Hektor and Bernd Kohler.

REFERENCES

[1] 2023. Quic-Go: A QUIC Implementation in Pure Go. https://github.com/quic-go/quic-go

[2] 2023. QUIC Implementations Overview. https://github.com/quicwg/base-drafts/wiki/Implementations

[3] 2023. Zmap/Zgrab2: Fast Go Application Scanner. https://github.com/zmap/zgrab2

[4] Mark Allman, Robert Beverly, and Brian Trammell. 2017. Principles for Measurability in Protocol Design. *ACM SIGCOMM Computer Communication Review* 47, 2 (2017), 2–12. https://doi.org/10.1145/3089262.3089264

[5] Péter Benkö and Andras Veres. 2002. A Passive Method for Estimating End-to-End TCP Packet Loss. In *Proceedings of the 2002 IEEE Global Telecommunications Conference (GLOBECOM)*. https://doi.org/10.1109/GLOCOM.2002.1189102

[6] Fabio Bulgarella, Mauro Cociglio, Giuseppe Fioccola, Guido Marchetto, and Riccardo Sisto. 2019. Performance Measurements of QUIC Communications. In *Proceedings of the 2019 ACM/IRTF Applied Networking Research Workshop (ANRW)*. https://doi.org/10.1145/3340301.3341127

[7] CAIDA. 2023. The CAIDA UCSD AS to Organization Mapping Dataset, 2022-2023. https://www.caida.org/catalog/datasets/as-organizations/

[8] Mark Coates, Alfred O. Hero, Robert Nowak, and Bin Yu. 2002. Internet Tomography. *IEEE Signal Processing Magazine* 19, 3 (2002), 47–65. https://doi.org/10.1109/79.998081

[9] Mauro Cociglio, Alexandre Ferrieux, Giuseppe Fioccola, Igor Lubashev, Fabio Bulgarella, Massimo Nilo, Isabelle Hamchaoui, and Riccardo Sisto. 2023. *Explicit Host-to-Network Flow Measurements Techniques*. Internet-Draft. IETF. https://datatracker.ietf.org/doc/draft-ietf-ippm-explicit-flow-measurements

[10] Piet De Vaere, Tobias Bühler, Mirja Kühlewind, and Brian Trammell. 2018. Three Bits Suffice: Explicit Support for Passive Measurement of Internet Latency in QUIC and TCP. In *Proceedings of the 2018 ACM Internet Measurement Conference (IMC)*. https://doi.org/10.1145/3278532.3278535

[11] David Dittrich and Erin Kenneally. 2012. *The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research.* Technical Report. U.S. Department of Homeland Security. https://doi.org/10.2139/ssrn.2445102

[12] Zakir Durumeric, Eric Wustrow, and J. Alex Halderman. 2013. ZMap: Fast Internet-Wide Scanning and Its Security Applications. In *Proceedings of the 2013 USENIX Security Symposium*. https://www.usenix.org/conference/usenixsecurity13/technical-sessions/paper/durumeric

[13] Alexa Internet Inc. 2021. Alexa. https://web.archive.org/web/20220101025437/https://www.alexa.com/topsites

[14] Jana Iyengar and Martin Thomson. 2021. *QUIC: A UDP-Based Multiplexed and Secure Transport.* RFC 9000. IETF. https://doi.org/10.17487/RFC9000

[15] Mirja Kühlewind and Brian Trammell. 2022. *Manageability of the QUIC Transport Protocol.* RFC 9312. IETF. https://doi.org/10.17487/RFC9312

[16] Ike Kunze, Constantin Sander, and Klaus Wehrle. 2023. Does It Spin? On the Adoption and Use of QUIC's Spin Bit – Raw Spin Bit Values. https://doi.org/10.5281/zenodo.8305843

[17] Ike Kunze, Constantin Sander, and Klaus Wehrle. 2023. Does It Spin? On the Adoption and Use of QUIC's Spin Bit – Raw Zgrab2 Toplist Results. https://doi.org/10.5281/zenodo.8354418

[18] Ike Kunze, Constantin Sander, Klaus Wehrle, and Jan Rüth. 2021. Tracking the QUIC Spin Bit on Tofino. In *Proceedings of the 2021 Workshop on the Evolution, Performance and Interoperability of QUIC (EPIQ)*. https://doi.org/10.1145/3488660.3493804

[19] Ike Kunze, Klaus Wehrle, and Jan Rüth. 2021. L, Q, R, and T - Which Spin Bit Cousin Is Here to Stay?. In *Proceedings of the 2021 ACM/IRTF Applied Networking Research Workshop (ANRW)*. https://doi.org/10.1145/3472305.3472319

[20] Victor Le Pochat, Tom Van Goethem, Samaneh Tajalizadehkhoob, Maciej Korczynski, and Joosen Wouter. 2019. TRANCO: A Research-Oriented Top Sites Ranking Hardened Against Manipulation. In *Proceedings of the 2019 Network and Distributed Systems Security Symposium (NDSS)*. https://doi.org/10.14722/ndss.2019.23386

[21] Majestic. 2023. The Majestic Million. https://majestic.com/reports/majestic-million

[22] Robin Marx, Luca Niccolini, Marten Seemann, and Lucas Pardue. 2023. *Main Logging Schema for Qlog.* Internet-Draft. IETF. https://datatracker.ietf.org/doc/draft-ietf-quic-qlog-main-schema work in progress.

[23] RIPE. 2023. RIPE - Route Collection Raw Data / Routing Information Service (RIS). https://data.ris.ripe.net/rrc00/

[24] Jan Rüth, Ingmar Poese, Christoph Dietzel, and Oliver Hohlfeld. 2018. A First Look at QUIC in the Wild. In *Proceedings of the 2018 International Conference on Passive and Active Network Measurement (PAM)*. https://doi.org/10.1007/978-3-319-76481-8_19

[25] Jan Rüth, Torsten Zimmermann, Konrad Wolsing, and Oliver Hohlfeld. 2018. Digging into Browser-based Crypto Mining. In *Proceedings of the ACM Internet Measurement Conference 2018 (IMC)*. https://doi.org/10.1145/3278532.3278539

[26] Constantin Sander. 2023. Quic-Go. https://github.com/COMSYS/quic-go

[27] Constantin Sander. 2023. ZGrab 2.0. https://github.com/COMSYS/quic-zgrab2

[28] Constantin Sander, Ike Kunze, Leo Blöcher, Mike Kosek, and Klaus Wehrle. 2023. ECN with QUIC: Challenges in the Wild. In *Proceedings of the 2023 ACM Internet Measurement Conference (IMC)*.

[29] Satadal Sengupta, Hyojoon Kim, and Jennifer Rexford. 2022. Continuous In-Network Round-Trip Time Monitoring. In *Proceedings of the 2022 ACM SIGCOMM Conference*. https://doi.org/10.1145/3544216.3544222

[30] Brian Trammell and Mirja Kühlewind. 2018. Revisiting the Privacy Implications of Two-Way Internet Latency Data. In *Proceedings of the 2018 International Conference on Passive and Active Network Measurement (PAM)*. https://doi.org/10.1007/978-3-319-76481-8_6

[31] Cisco Umbrella. 2016. Umbrella Popularity List. https://umbrella-static.s3-us-west-1.amazonaws.com/index.html

[32] Yufei Zheng, Huacheng Yu, and Jennifer Rexford. 2023. Detecting TCP Packet Reordering in the Data Plane. *arXiv.2301.00058* (2023). https://doi.org/10.48550/arXiv.2301.00058

[33] Torsten Zimmermann, Jan Rüth, Benedikt Wolters, and Oliver Hohlfeld. 2017. How HTTP/2 Pushes the Web: An Empirical Study of HTTP/2 Server Push. In *Proceedings of the 2017 IFIP Networking Conference*. https://doi.org/10.23919/IFIPNetworking.2017.8264830

[34] Johannes Zirngibl, Philippe Buschmann, Patrick Sattler, Benedikt Jaeger, Juliane Aulbach, and Georg Carle. 2021. It's Over 9000: Analyzing Early QUIC Deployments with the Standardization on the Horizon. In *Proceedings of the 2021 ACM Internet Measurement Conference (IMC)*. https://doi.org/10.1145/3487552.3487826

A ETHICS.

We follow widely accepted ethical measures and measurement guidelines to minimize the impact of our study [11, 12].

Information. We inform about our research by (1) adding reverse DNS entries to our measurement IPs showing our research context, (2) embedding our projectname as hint in every HTTP request, and (3) hosting a dedicated website on our measurement IPs. The website explains our research and how measurement targets can opt out of our study: via a single email or by blocking our measurement subnet. We quickly handle all (abuse) emails.

Load reduction. We conduct our measurements from within our network using a dedicated IP of our measurement subnet. We spread each measurement over multiple days to reduce the peak load on networks, end-hosts, and shared infrastructure, such as DNS, e.g., only processing 1000 domains per second. Furthermore, we maximize the utility of our measurements as distinct partial results are reused for different research projects (e.g., [28]), i.e., we only perform one set of measurements instead of several individual ones, significantly reducing the caused load.

B ARTIFACTS

We release our changes to zgrab [27], quic-go [26] and our raw measurement data of the toplists [17], stripping unused information to limit the file size. To enable future work on the spin bit, e.g., regarding RTT filtering mechanisms (cf. Sec. 5.2), we also add the extracted raw spin bit information for all domains as detailed in Sec. 3.3 together with qlog baseline information and corresponding analysis scripts [16].