

Sampling TCP Data-Path Quality with TCP Data Probes

Rocky K. C. Chang, Edmond W. W. Chan, Waiting W. T. Fok, and Xiapu Luo

Department of Computing
The Hong Kong Polytechnic University
Kowloon, Hong Kong
{csrchang|cswwchan|cswtfok|csxluo}@comp.polyu.edu.hk

ABSTRACT

In this paper, we present preliminary results of measuring TCP data-path quality using a new measurement tool called OneProbe. Unlike the existing tools, OneProbe uses legitimate TCP data probes to profile TCP data-path quality by sampling round-trip delay, one-way loss rate, and one-way reordering rate at the same time. This paper presents a set of recent measurement studies on a set of web servers hosting the Beijing Olympic Games in China and Hong Kong.

1 Introduction

This paper considers active measurement methods for end-to-end paths without controlling the remote endpoint. We refer this class of methods to as *solo-active*. In contrast, *duet-active* methods require the control of both endpoints. Measuring an end-to-end path quality from only one endpoint is very challenging in the Internet landscape today. A major challenge is to induce suitable responses from arbitrary Internet nodes for measurement. The 28-year-old Ping, though still popular, has become less effective. A recent measurement study showed that using TCP SYN packets could reach 20%–30% more hosts than Ping [17]. Our own measurement results for the 37,000 web servers used for OneProbe validation tests also showed that only around 82.7% of them responded to Ping.

Furthermore, delay and loss measurements obtained from ICMP packets do not necessarily match with that for data packets, because they could be processed on different tracks. For example, to curb ICMP and TCP based DoS attacks, ICMP and TCP SYN/RST packets are subject to router’s rate-limiting. Other path intermediaries (such as, load balancers, accelerators, and firewalls) can further magnify the differences in the path quality experienced by these probe packets and data packets. Another source of inaccuracy comes from adversarial modification of the probe packets [5].

A third challenge is to diagnose path quality problem beyond round-trip measurement. Since Internet paths are known to be asymmetric, it is reasonable to expect that forward-

path quality and reverse-path quality are generally different. Although measuring one-way delay is very difficult without controlling the remote host [6], one-way measurement for other metrics is possible—sting [14] for packet loss and POINTER [11] for packet reordering. However, these tools may encounter problems similar to those by Ping and TCP control packets, because their methods of inducing response packets are based on “exception packet behavior”, such as sending a large number of reordered packets in [14].

This paper presents measurement results obtained by a new solo-active tool called OneProbe [10]. Unlike previous approaches, OneProbe measurement is based on normal TCP data probes. Therefore, OneProbe’s results can accurately reflect the path quality experienced by data packets (i.e., data-path quality). Besides, OneProbe measurement will not be easily tampered by path intermediaries and malicious activities on the path, because the probe and response packets cannot be easily distinguished from normal TCP data packets. Furthermore, OneProbe can sample round-trip delay, and one-way path metrics for packet loss and packet reordering at the same time using a single probe.

In the next section, we present measurement results for a set of Beijing Olympic Games web servers. The results reveal a number of interesting and useful path properties some of which cannot be obtained by the existing solo-active tools.

2 Beijing Olympic Games measurement

We deployed OneProbe to measure the data-path quality of a set of web servers hosting Beijing Olympic Games’ content. The measurement results presented here are drawn from the Olympic measurement (we also conducted measurement on PlanetLab nodes [10]). We chose the Olympic websites, first of all, for the event’s worldwide popularity and its timeliness for our study. Measuring the paths to these web servers under tight screening for security threats also presented an excellent testing ground for OneProbe.

2.1 The setup

We deployed OneProbe at a data center in Hong Kong (HK) to monitor data-path quality for nine Beijing Olympic Games web servers over two time periods. The first was 08 August 2008 03:30 UTC to 25 August 2008 19:19 UTC, which covers the entire Olympic Games period¹. The second was after the Olympic Games: 28 August 2008 16:06 UTC to 17 September 2008 16:56 UTC. We chose to deploy OneProbe at the

¹The opening ceremony started at 8 August 2008 12:08 UTC, and the closing ceremony ended at 24 August 2008 13:55 UTC.

data center (instead of the PlanetLab nodes) for its more stable environment (e.g., guaranteed bandwidth and dedicated hosts). Moreover, our collaborators at the data center could provide the “ground truth” about their network setup that might affect our measurement results. For example, we confirmed with them that our probes were not rate-limited in their network.

We obtained a set of web servers’ IP addresses by crawling the domain `beijing2008.cn` using Larbin [3]. The APNIC and ARIN whois databases showed that the addresses belonged to three groups of providers: ChinaCache, China Network Communications Group Corp. (CNCGroup), and Akamai and Quest. ChinaCache, being a CDN, hosted the Olympic Games sites in various parts of China. The CNCGroup hosted the origin web servers in Beijing. Akamai’s and Quest’s servers were located in HK. We selected three servers from each set based on their RTTs and refer them to as CC-1/2/3 (ChinaCache), BJ-1/2/3 (Beijing), and QT (Quest) and AK-1/2 (Akamai).

We employed three hosts H-1/2/3 to measure the three sets of servers. Each host measured a server from each set, as shown in the first column of Table 1. In this way, the measurement results for a particular set would not be biased by a single machine. Besides, we measured the same set of paths using Ping [12], PPing [8], and Httping [9] for comparison and consistency checks. We included PPing, because the data center uses it to monitor path performance. PPing uses TCP SYN/SYN-ACK packets to measure round-trip delay and loss rate. Including Httping—it uses both SYN/SYN-ACK packets and HTTP HEAD request messages to measure round-trip delay and loss rate—is also useful, because it uses the HTTP/TCP channel for measurement. However, unlike OneProbe, Httping opens a new TCP connection for each observation.

Table 1: A summary of the nine paths’ characteristics obtained by TCPTraceroute.

Probers/ servers	No. hops	Forward path: Region/Network (no. hops)
H-1	CC-1	13 HK (5)→CNCGroup Backbone (4)→Hebei Province Network (4)
	BJ-1	16 HK (5)→AP-TELEGLOBE (2)→CNCGroup Backbone (4)→Beijing Province Network (4)
	QT	5 All in HK
H-2	CC-2	19 HK (5)→CNCGroup Network (6)→CNCGroup Backbone (3)→Shandong Province Network (5)
	BJ-2	16 Almost the same as BJ-1’s
	AK-1	5 First four hops same as QT’s
H-3	CC-3	16 HK (5)→Korea (2)→CNCGroup Backbone (4)→Henan Province Network (5)
	BJ-3	16 Almost the same as BJ-1’s
	AK-2	5 First four hops same as QT’s

The tools sampled each path continuously for one minute, then idled for the next four minutes. After that, the cycle repeated again. An one-minute interval is sufficiently long for sampling loss rate [16] but is not intrusively long to the server. Within each one-minute interval, the sampling process was periodic. OneProbe’s probe rate was two per second, and it opened new TCP connections for every new interval and closed them at the end of it. To reduce the experiment complexity, we fixed both probe and response packet sizes to 1500 bytes. For other tools, we limited their

probe rates to one per second in order to prevent routers and firewalls from blocking or rate-limiting their traffic.

Moreover, to facilitate accurate comparisons, we synchronized the tools’ starting times for each interval. We took several measures to avoid cross-induced or self-induced packet losses. Most important, the probe rates were chosen to restrict the aggregated rate of the probe and response packets to 217 Kbps for each host. OneProbe also includes a self-diagnosis module to detect self-induced losses. We further analyzed the measurement data to confirm negligible self-induced losses.

We also performed Traceroute and TCPTraceroute every five minutes to obtain their forward paths which are summarized in Table 1. The paths were generally stable, although there were some flip-flops between two IP addresses for some hops. The CC paths were most diverse. Despite that they all went through the CNCGroup Backbone, they were diverged to different province networks where the servers were located. The CC-3 path even detoured to Korea. The BJ paths, on the other hand, were almost the same until reaching the servers. The AK/QT paths also shared the same path except for the last hop; however, they were different from the BJ paths in their short RTTs. For the reverse paths, the IP hop counts stayed almost unchanged.

3 Measurement results

This section reports the following path metrics for OneProbe’s one-minute measurement interval. Wherever appropriate, we will aggregate the results obtained from a number of intervals.

- (Median RTT) We have adopted the suggestion of using median for combining statistics for a path [13]. There are two different RTT metrics. The first (just called RTT) reports observations for probes and response packets that do not experience loss or reordering. The second is loss-pair RTT: the RTT observed by a loss pair for which only the second packet is lost.
- (Forward/reverse-path loss rate) This derived metric is given by the first-packet loss rate that is the percentage of the first probe/response packet losses. We do not consider the second-packet losses, because they could be biased by the first packets [4].
- (Forward/reverse-path reordering rate) This derived metric is given by the percentage of probes/response packets that are reordered when arriving at the other endpoint. Therefore, this metric considers only the observations that do not involve any packet loss.

We computed the RTT and round-trip loss rate obtained from Ping/PPing/Httping similarly as for OneProbe.

Comparing the tools’ RTTs is rather straightforward, but loss comparison requires careful considerations. First, other tools can only measure round-trip loss rates. Therefore, we computed a round-trip loss observation for OneProbe by combining the two one-way observations: an observation reported a round-trip loss if there was a loss on *either* path. Second, OneProbe’s sampling rate was double of other tools’ rates. To obtain a fair comparison, we filtered every other observations made by OneProbe.

3.1 The first set

The measurement experiments were conducted smoothly for all the tools. However, unlike Ping, PPing, and Httping, OneProbe might not always receive a full set of 120 ob-

servations during an one-minute interval. Since Pping and Httpping use TCP SYN/SYN-ACK packets for measurement, a failure of establishing a TCP connection will be counted as a loss observation. However, the same failure event will prevent OneProbe from making any observation from the data channel. Besides, all TCP connections employed in an OneProbe measurement session could be blocked, especially when there are losses. Our results showed that 98.95% of the measurement intervals contained a full set of observations, and over 99.90% contained at least 99 observations which were still enough for computing the statistics.

Table 2 summarizes the overall statistics of the data-path quality obtained by OneProbe for the first set of Olympic measurement. In the following, we report some of the findings, and more results will be available from [1]. We use FW, RV, and RT to refer to forward path, reverse path, and round trip, respectively.

3.1.1 Diurnal RTT and loss patterns

OneProbe measurement captured diurnal RTT and round-trip loss patterns for the CC and BJ paths, as shown in Figure 1. In particular, the weekend RTT patterns were distinct from the weekdays' by their longer periods of peak RTTs (i.e., more time spent on Internet). It is also important to observe that most of the packet losses occurred during the peak RTT periods. The losses were more prevalent on weekends than weekdays. Another interesting observation is that the RTT of all the six paths experienced a sudden increase of 17 ms at 23 August 20:37 UTC, and the increase was maintained thereafter. Since the change took place at the same time for all the CC and BJ paths (but not for the QT/AK paths), the change was due to an unknown event in the CNCGroup Backbone which was the only network the six paths traversed. The TCPtracert results, however, did not reveal any change in the paths.

As expected from the TCPtracert results, the BJ paths shared very similar quality. This observation also confirms that the three hosts' measurement results were consistent. On the other hand, the CC-1 and CC-2 path quality also shared some similarity, but the CC-3 path quality was very different from the other two. In particular, the CC-3 path's RTT and loss rate are the highest in the set, which was possibly due to its detour to Korea, instead of using the CNCGroup networks as the other two did. Compared with the BJ paths, the CC paths' RTTs also saw much higher variations.

3.1.2 Discrepancy between Ping and OneProbe RTTs

While Ping RTT and OneProbe RTT captured similar trends, we observed from the CC-3 path that they could be very different. Figure 2 depicts that their RTTs consistently differed by around 100 ms during the peaks, but they were similar during the valleys. We did not observe similar behavior from the other two CC paths. Therefore, we suspect that the route via Korea was responsible for this behavior. Interestingly enough, due to an event unseen to the IP layer, their RTTs "converged" at 12 Aug. 2008 16:39 UTC². We will also comment on the change in forward-path loss rates below.

To probe further, we present scatter plots in Figure 3 to com-

²Since the local time was 1 hour 39 minutes into the midnight, it is plausible that a network configuration change took place at that time.

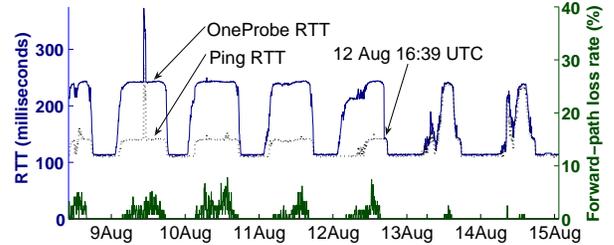


Figure 2: Different RTTs measured by OneProbe and Ping for the CC-3 path in the first five days.

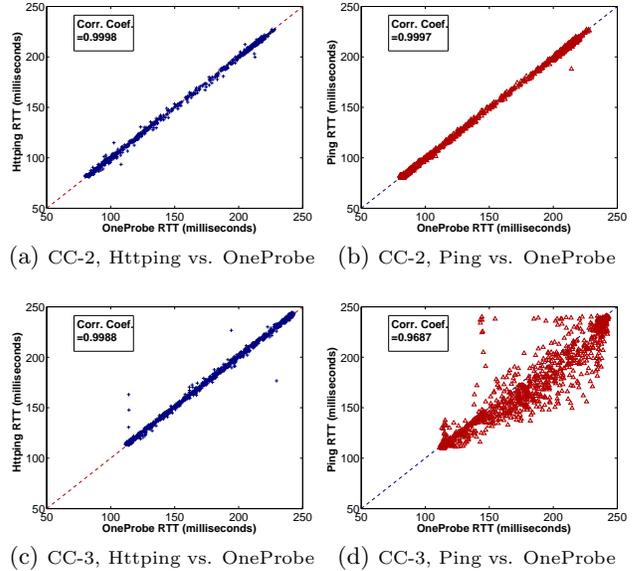


Figure 3: Comparing the RTT measurement for Httpping, Ping, and OneProbe during 14–21 August 2008.

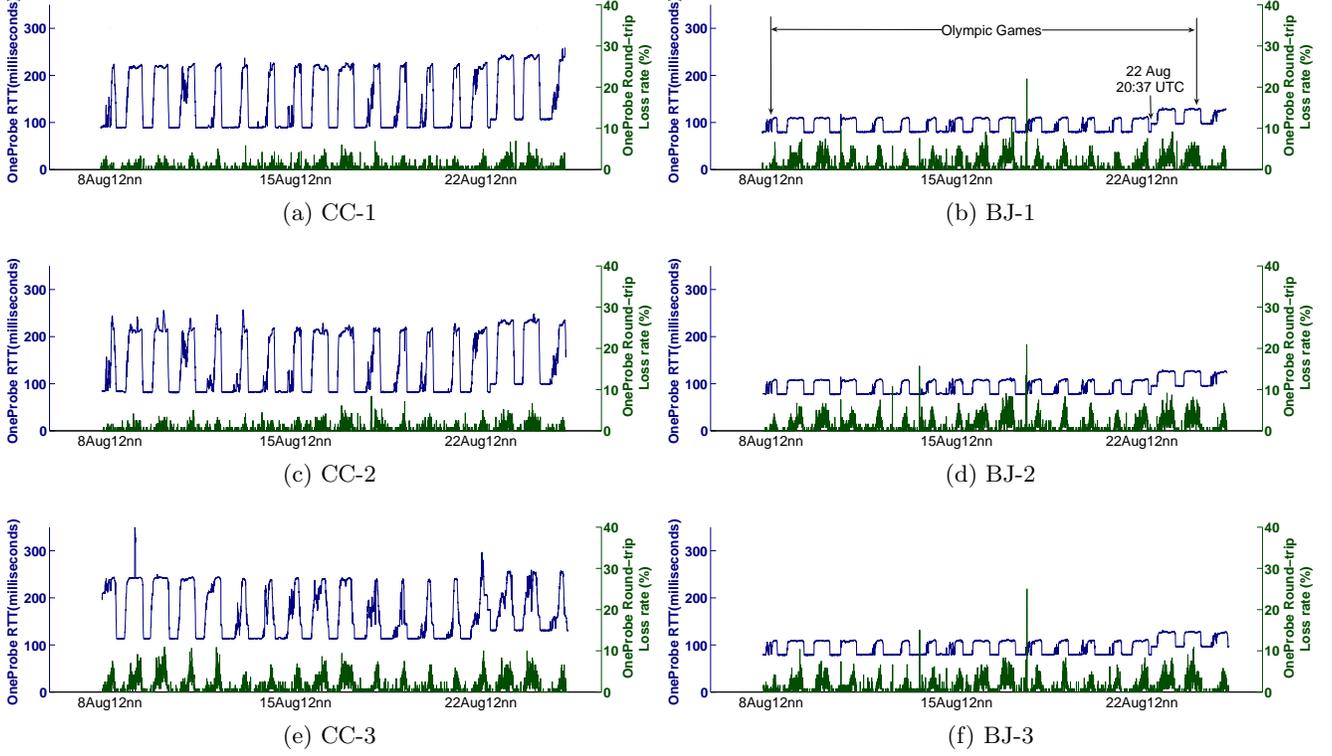
pare OneProbe's RTT with other tools' RTTs based on one week of data starting on 14 August 2008. Even after Ping RTT and OneProbe RTT converged on 12 August 2008, the figure shows that their RTTs still disagreed when the RTTs ramped up and down between the valleys and the peaks (i.e., 140–240ms). This shows that ICMP packets and TCP data packets could experience very different RTTs during a traffic "surge-up" or "surge-down". However, Ping RTT and OneProbe RTT matched very well for the CC-1/2 paths. Figure 3 also shows that Httpping and OneProbe RTTs were in agreement, because both were conducted in TCP data channels. The slight variations were due to the HTTP request's nonconstant packet size and the small TCP SYN-ACK packet size. Results for other paths using Httpping and Pping were generally in agreement with OneProbe measurement.

3.1.3 Highly asymmetric loss rates

OneProbe measurement results revealed that the loss rates were highly asymmetric for all nine paths. Figure 4 shows three of them. We skip the plots for the AK-1/2 and BJ-2/3 paths, because they are very similar to the QT and BJ-1 paths, respectively. The figures clearly show that the

Table 2: The overall data-path quality measured by OneProbe for the first set of Olympic measurement.

Servers	Median RTT	Packet loss rate	Reordering rate
CC-1	Diurnal, peak \approx 225ms, valley \approx 90ms	Weak diurnal, peak \approx 5%, RV-loss rate $>$ FW loss rate	None
CC-2	Diurnal, peak \approx 225ms, valley \approx 90ms	Weak diurnal, peak \approx 5%, average FW-loss rate \approx 0%	None
CC-3	Diurnal, peak \approx 250ms, valley \approx 120ms	Diurnal, peak \approx 8%, persistent RV-loss rate, change in FW loss rate	None
BJ-1	Diurnal, peak \approx 110ms, valley \approx 80ms	Diurnal, peak \approx 6%, occasional high FW loss rate	None
BJ-2	Diurnal, peak \approx 110ms, valley \approx 80ms	Diurnal, peak \approx 6%, occasional high FW loss rate	None
BJ-3	Diurnal, peak \approx 110ms, valley \approx 80ms	Diurnal, peak \approx 6%, occasional high FW loss rate	Some high FW rate
QT	Average \approx 4ms	RV loss rate \approx 0%, FW loss rate \approx 5% but became 0%	None
AK-1	Average \approx 4ms	RV loss rate \approx 0%, FW loss rate \approx 5% but became 0%	None
AK-2	Average \approx 4ms	RV loss rate \approx 0%, FW loss rate \approx 5% but became 0%	None

**Figure 1: Time series of OneProbe’s RTT and round-trip loss rates for the CC and BJ paths.**

round-trip loss rates were mostly dominated by reverse-path losses. Although there were some high forward-path loss rates observed occasionally for the BJ paths, they were not persistent enough to affect the path quality. The plots for the CC-1/2 paths (not shown here) are also similar in that they experienced persistent losses on their forward paths, but their rates were low compared with the reverse-path loss rates.

The CC-3 and QT/AK paths presented perhaps the most interesting loss results. The loss rates on the CC-3 forward and reverse paths were quite balanced for the first five days. However, the forward-path loss rate diminished significantly at 12 August 2008 16:00 UTC; as a result, the reverse-path losses dominated the loss rates thereafter. Similar to the previous case, the change occurred at midnight on the local time, it is therefore likely that the change was also caused by a change in network configuration. For the QT/AK paths, their forward-path loss rates were almost zero; the reverse-path losses were persistently high until 14 August 2008 06:50 UTC at which the loss rate dropped to zero. We will come

back shortly to correlate the loss pattern changes with RTT changes.

We also compared OneProbe’s round-trip loss rates with that obtained by Httping, Pping, and Ping. To provide a clearer comparison, the results are based on 60-minute data aggregations (i.e., combining 12 one-minute intervals). The metric used is the ratio of Httping/Pping/Ping’s loss rate to OneProbe’s loss rate obtained during the same 60-minute data aggregation, provided that OneProbe’s loss rate is greater than 0. A ratio of less than one therefore means that OneProbe’s loss rate was higher than the other tool. Figure 5 plots the cumulative distribution function (CDF) for this ratio. Unlike RTT, OneProbe’s and other tools’ loss rates generally did not agree. The degrees of discrepancy were also different for the three different sets of paths and for the three tools. For the QT/AK paths, OneProbe measured a higher loss rate almost all the time. The only case that OneProbe gave a lower loss rate in a majority of the 60-minute intervals is Httping for the CC-1/2 paths in Figure 5(c).

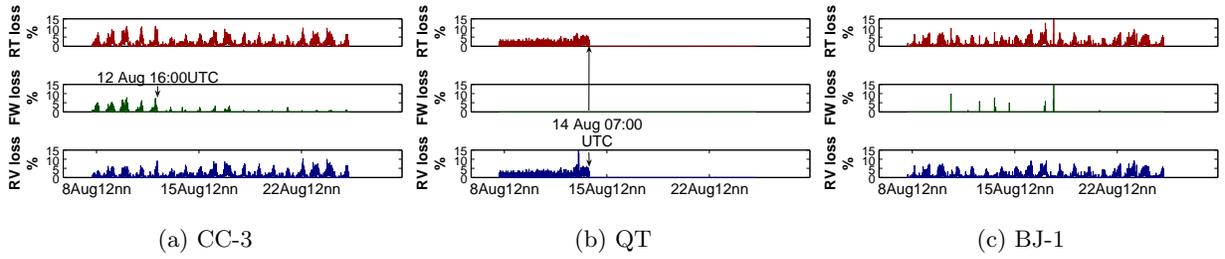


Figure 4: Time series of round-trip, forward-path, and reverse-path loss rates measured by OneProbe.

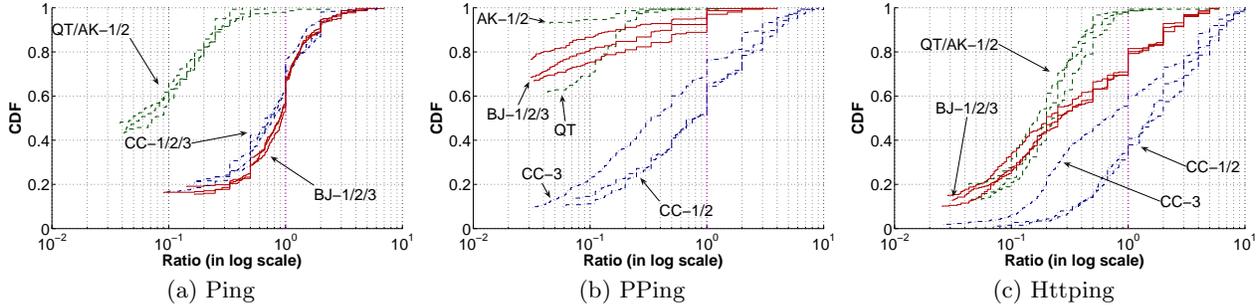


Figure 5: CDF of the ratio of Httping/PPing/Ping’s loss rate and OneProbe’s loss rate with 60-minute data aggregations.

3.1.4 Abrupt changes in data-path quality

OneProbe uncovered three suspicious configuration changes unseen to TCPTraceroute and Traceroute. In a chronological order, the first occurred only to the CC-3 path (Figure 2), the second to all QT/AK paths (Figure 4(b)), and the third to all CC and BJ paths (Figure 1). OneProbe detected them by either abrupt changes in a single path metric experienced by multiple paths or abrupt changes in multiple path metrics.

OneProbe detected the change for the CC-3 path by measuring an abrupt drop in the forward-path loss rate. The convergence of Ping’s and OneProbe’s RTTs occurring at the same time further supports this conjecture. Since this change applied to one path, it was likely due to a network covered by only the CC-3 path, such as the Korea network. For the QT/AK paths, OneProbe detected the change by the disappearance of the reverse-path loss and a drop in the RTT by 0.5 ms (not shown here). Because the change was not observed by the CC and BJ paths, the change should be originated from a network in HK. Finally, OneProbe detected the change for all CC and BJ paths by a rise in RTT for all six paths.

3.1.5 Loss-pair RTT

Recall that RTTs in Figure 1 were based on the observations that did not detect loss or reordering events. However, a loss pair only “samples” the RTT distribution. Figure 6 shows the RTTs observed by a loss pair for the CC-3 path. The loss-pair RTTs are superimposed with the corresponding RTT time series in Figure 1 to identify which parts the loss pair sampled the most. To be consistent with the first-packet RTT measurement, we only plot the forward/reverse-path loss pairs with the second packet lost. Figure 6 shows that almost all the loss-pair RTTs on the

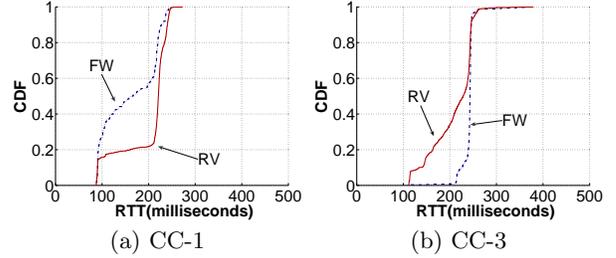


Figure 7: CDF of loss-pair RTTs measured by OneProbe for the CC-1/3 paths.

forward path were clustered at the RTT peaks. That is, the first packet experienced a largest RTT, and the second packet was dropped by the router. Therefore, the results suggest that the packets were dropped in a drop-tail router. However, the loss-pair RTTs on the reverse path behaved very differently. While many loss pairs saw the largest RTT, there were also many others seeing other values of RTT, including the lowest RTT. Figure 7 further quantifies the differences by comparing their CDFs of the loss-pair RTTs. Moreover, we had observed different loss-pair patterns for other paths. The loss-pair RTTs for the CC-1 reverse path, for example, clustered at both low RTT and high RTT, as shown in Figure 7(a).

3.2 The second set

We did not observe significant differences between this set of results and the first one. In this set, the RTTs were slightly higher, but the loss rates were slightly lower. The loss rates for the CC and BJ paths were still highly asymmetric with the domination of the reverse-path losses. Moreover, the

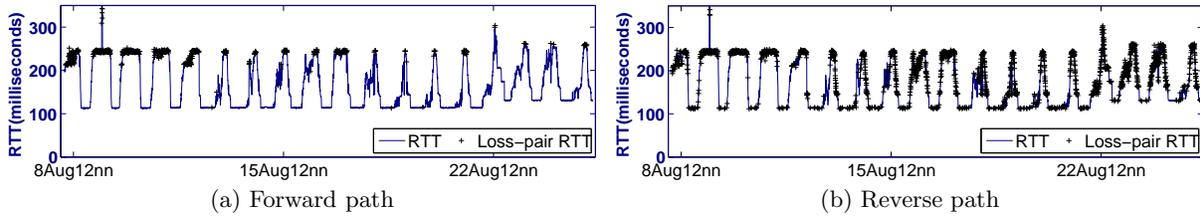


Figure 6: The RTTs observed by forward/reverse-path loss pairs for the CC-3 path.

three QT/AK paths' loss rates were almost zero, similar to the latter part of the results in the first set. The most obvious change was a significant drop in the CC-3 path's RTT on the 10th day of the measurement period.

4 Discussion and future work

We have selected three issues that require further discussion and even new measurement experiments.

Source of the diurnal patterns Although OneProbe measurement results revealed clear diurnal patterns for RTT and loss for the CC and BJ paths, they could not help identify its source. We in fact ran another set of experiments for the same set of web servers from the data center, in parallel to the Olympic measurement reported in this paper. However, the inbound paths for this unreported set were different (the traffic returned via a different upstream ISP). This set of unreported results, to our surprise, did not reveal diurnal RTT patterns, except for the CC-3 path. This observation therefore suggests that the diurnal patterns were not due to web surfing synchronization for the Games. Instead, the upstream ISP for the inbound traffic was responsible for the diurnal patterns.

Loss measurement and comparison Unlike RTT, accurate loss measurement is much harder to obtain [15]. We employed a periodic sampling method mainly for facilitating the process of verifying and analyzing the results. Comparing loss measurement across different tools is even more difficult. There are at least a couple of factors responsible for the measurement discrepancies. First, OneProbe sent more packets due to a higher rate and two-packet probes. Second, Httping and Pping might end their measurement beyond the one-minute interval when encountering losses, but OneProbe uses a probe scheduler to comply with the sampling pattern.

Data-path quality for UDP It is not clear whether the data-path quality is similar for UDP data. Measuring UDP data-path quality from a single endpoint is generally difficult, because most hosts and routers do not respond to UDP packets [7]. We are therefore planning to conduct duet-active experiments in the PlanetLab platform. The One-way Ping (OWAMP) [2], which reports loss, reordering, and delay statistics, will be a useful tool for this purpose.

5 Conclusions

In this paper, we presented preliminary measurement results obtained by OneProbe, a new solo-active system for measuring data-path quality. Unlike Ping and other methods based on TCP control packets, OneProbe measures the path quality experienced by TCP data packets. Using data probes also has the advantage that they are less susceptible to various path intermediaries and malicious activities.

The smooth monitoring of the nine Olympic servers for almost one month has successfully demonstrated the nonintrusiveness of OneProbe measurement. Moreover, the one-way measurement capability offers a number of advantages, such as revealing highly asymmetric loss patterns, detecting changes in network configurations, and obtaining loss-pair RTTs.

6 Acknowledgments

This work is partially supported by a grant (ref. no. ITS/152/08) from the Innovation Technology Fund in Hong Kong.

7 References

- [1] OneProbe website. <http://www.oneprobe.org>.
- [2] OWAMP. <http://e2epi.internet2.edu/owamp/>.
- [3] S. Ailleret. Larbin: Multi-purpose web crawler. <http://larbin.sourceforge.net/>.
- [4] J. Bolot. End-to-end packet delay and loss behavior in the Internet. In *Proc. ACM SIGCOMM*, 1993.
- [5] S. Goldberg, D. Xiao, E. Tromer, B. Barak, and J. Rexford. Path-quality monitoring in the presence of adversaries. In *Proc. ACM SIGMETRICS*, 2008.
- [6] O. Gurewitz, I. Cidon, and M. Sidi. One-way delay estimation using network-wide measurements. *IEEE/ACM Trans. Networking*, 14(6), 2006.
- [7] A. Haeberlen, M. Dischinger, K. Gummadi, and S. Saroiu. Monarch: A tool to emulate transport protocol flows over the Internet at large. In *Proc. ACM/USENIX IMC*, 2006.
- [8] M. Hedlund. Pping. <http://www.blazen.com/pping/>.
- [9] F. Heusden. <http://www.vanheusden.com/httping/>.
- [10] X. Luo, E. Chan, and R. Chang. Design and implementation of TCP data probes for reliable and metric-rich network path monitoring. In *USENIX Annual Technical Conference*, 2009.
- [11] X. Luo and R. Chang. Novel approaches to end-to-end packet reordering measurement. In *Proc. ACM/USENIX IMC*, 2005.
- [12] J. Postel. Internet control message protocol. RFC 792, IETF, September 1981.
- [13] M. Roughan and O. Spatscheck. What does the mean mean? In *Proc. Intl. Teletraffic Congress*, 2003.
- [14] S. Savage. Sting: a TCP-based network measurement tool. In *Proc. USENIX Symp. Internet Tech. and Sys.*, 1999.
- [15] J. Sommers, P. Barford, N. Duffield, and A. Ron. Improving accuracy in end-to-end packet loss measurement. In *Proc. ACM SIGCOMM*, 2005.
- [16] H. Song, L. Qiu, and Y. Zhang. NetQuest: A flexible framework for large-scale network measurement. In *Proc. ACM SIGMETRICS*, 2006.
- [17] L. Wenwei, Z. Dafang, Y. Jinmin, and X. Gaogang. On evaluating the differences of TCP and ICMP in network measurement. *COMCOM*, Jan. 2007.