Has Internet Delay Gotten Better or Worse?

DK Lee, Kenjiro Cho†, Gianluca Iannaccone‡, and Sue Moon
KAIST, †IIJ Research Laboratory, ‡Intel Research Berkeley
{dklee, sbmoon}@an.kaist.ac.kr, †kjc@iijlab.net, ‡gianluca.iannaccone@intel.com

ABSTRACT

Delay is a key Internet performance metric and its stability, variation, and abrupt changes have been well studied. However, little could have been said about the Internet-wide delay distribution. In order to build a representative sample set for the Internet-wide delay distribution, one needs to draw data from a random selection of source hosts to destination hosts and there is no measurement system with access to every AS and subnet of the Internet.

In this work we propose to apply the *path-stitching* algorithm to archival measurement data and reconstruct the past history of Internet delay distribution. The two main advantages of path stitching are that data from existing measurement projects is sufficient to provide accurate estimates and it produces delay estimates between almost any two hosts in the Internet. As a first step towards the longitudinal study of the Internet-wide delay distribution, we examine how the Internet delay changes from 2004 to 2009. Our work is the first ever systematic approach to Internet delay distribution. We report the overall delay distribution has gotten worse from 2004 to 2009, while the delay distribution for the same set of host pairs remains almost identical or slightly improved.

Categories and Subject Descriptors

C.2.5 [**Local and Wide-Area Networks**]: Internet (e.g., TCP/IP)

Keywords

Internet measurement, Internet-wide delay distribution

1. INTRODUCTION

The Internet today is the most widely spread platform for information dissemination and plays a vital part in communication and collaboration of our modern lives. The network performance of the Internet is critical to all aspects of communications and online services. Many large-scale projects have been proposed and deployed to collect Internet-wide measurements data [1, 5, 6, 9, 10, 12, 14]. Longitudinal study about the evolution of the Internet AS topology [4,11] and Internet traffic [2] have revealed that the AS peering practice has switched from hierarchical to peer-to-peer and dominant traffic types have changed from web to peer-to-peer. But what do we know about the overall Internet delay performance? Internet delay is one of the key performance metrics, closely tied to application performance and user satisfaction. As a key end-to-end performance metric, stability, variation, and abrupt changes of delay as a path statistic have been well studied. However, little could have been said about the Internet-wide delay performance. In order to build a representative sample of the Internet-wide delay distribution, one needs data from a *random* selection of source hosts to destination hosts and there is no measurement system with access to every AS and subnet of the Internet. Only statistics from a selective partial set have been available [16].

In order to estimate the delay distribution of the Internet, it is essential to run point-to-point measurement between any source and destination pairs that are randomly drawn from every possible IP address. Instead of instrumenting end-hosts to collect measurements, we consider a different approach to estimate end-to-end delay. In our previous work we have proposed a structural path and round-trip delay estimation scheme called *path stitching* [8]. The main idea is to decompose existing end-to-end measurements by the AS and reconstruct the end-to-end path and delay. The two main advantages of path stitching is that data from existing measurement project is sufficient to provide estimates better than active measurement assisted estimation schemes and it can answer queries about most part of the Internet. We can apply path-stitching to any measurement, past or present, and reconstruct end-to-end path and delay. This unique capability together with random sampling of the Internet enables us to raise and address the following long-cherished and interesting questions.

- Has the Internet grown shorter in delay?
- What are the basic rules that govern the long-term dynamics of the Internet delay? How has it evolved? At what rate? When and why did the rate change?

As a first step towards the longitudinal study of the Internet-wide delay distribution, we investigate the feasibility of reconstructing the past history of Internet delay distribution with the existing measurements data. In this work, as a preliminary result, we examine how the Internet delay changes from 2004 to 2009. Our work is the first ever systematic approach to Internet delay distribution. We report that overall delay distribution has gotten worse from 2004 to 2009, while the delay distribution for the same set of host pairs remains almost identical or slightly improved. Our study of Internet delay distribution evolution does not focus on the individual mi-

*DK Lee and Sue Moon were supported by the IT R&D program of MKE/KEIT [KI001878, “CASFI : High-Precision Measurement and Analysis Research”].

CFI '10 June 16–18, 2010, Seoul, Korea

croscopic behavior, but is more of a macroscopic summary of the evolution trend, yet accounting for all the microscopic changes.

2. RECONSTRUCTING PAST HISTORY

In order to reconstruct past history of the Internet delay we need matching data and a methodology to combine them and produce end-to-end path and delay between arbitrary hosts. The core estimation methodology used in this work is path stitching [8]. We present a brief overview on how path stitching works and what types of data it uses in different steps.

Datasets

In this work, we rely on two types of the Internet's historical data: (1) end-to-end Internet forwarding path and delay measurements and (2) routing information.

These two types of measurement data have been available for over a decade: traceroute measurements collected CAIDA's Skitter and Ark projects [6, 7] and BGP routing table snapshots collected by RouteViews [1] and RIPE Routing Information Service (RIS) [12]. While they are among the largest data archives publicly available and hold constantly updated information about IP and AS-level topologies, those datasets obviously do not provide a complete map of the Internet (Ark traceroutes are generated by tens of systems in total). But they still provide a good starting point for our investigation into the representative delay distribution of the Internet.

From Ark, we use one round of traceroute outputs taken in June, 2004 and in June, 2009 (a total of approximately 50 million traceroute outputs.) A round of data in Skitter and Ark refers to a set of traceroute outputs to all routable /24 prefixes from the sources. From RouteViews and RIPE RIS, we use all available BGP table snapshots of the same period as our Ark data.

Path stitching

Path stitching is at the core of this work, enabling us to reconstruct end-to-end path and delay between any two arbitrary end hosts in the Internet. Figure 1 is a step-by-step illustration of how path stitching works. When a query for the path and delay from x to z arrives, path stitching produces delay estimate as follows. In Step 1 it maps the two IP addresses x and z to their AS numbers, X and Z, based on the routing information. In Step 2 it infers the AS-level path between the two ASes, X and Z. In Step 3 it stitches path segments along the inferred AS path, and finally returns an end-to-end delay estimate.

The two main source of input data to path stitching are hop-by-hop delay measurements and the BGP routing tables. The former is segmented by ASes and is transformed to a path segment repository. The latter is used in prefix-to-AS mapping, AS path inference, and routable /24 prefix compilation.

Path stitching does not always return a result. It fails when the inferred AS path between the source and destination IP addresses has an AS of which path segment does not exist in the path segment repository. It means that the Ark data we use failed to collect traceroute measurement about that specific AS. It also fails when end points of path segments from two adjoining ASes on the inferred path do not line up and cannot be stitched. In this case we employ approximation rules, such as using reverse path segments and clustering at /24 prefixes. On the other hand, path stitching may return multiple stitched paths for a given query. In such a case, path stitching applies preference rules to rank candidates and select the best one. Preferences are given to those segments with IP addresses that are close to the destination address, for the same destination prefix, and, lastly, to the most recent segment.

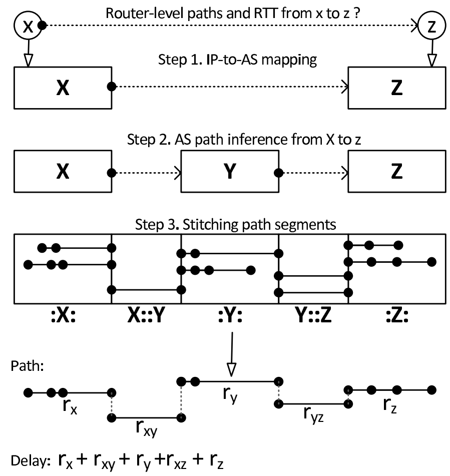


Figure 1: Path stitching algorithm, where :X: is a set of path segments in AS X, X::Y is a set of path segments between AS X and Y.

Path stitching reports less than 10 ms error for 75% of the cases when the query sources reside in the same ASes as the probing monitors are; and 50% when the query sources are not in the same ASes as any of the probing monitors. This performance is comparable to or slightly better than iPlane that has shown the best performance network delay estimation.

Host pair sample size

Having downloaded the archival traceroute data and routing information and armed with path stitching, we now design our sampling methodology for Internet delay distribution estimation. The complete delay distribution between every possible pairs of hosts on the Internet is impossible to obtain. Then, how many samples of host pairs are representative enough?

We regard the Internet as a finite set of pairs of communicating hosts. Instead of counting all possible pairs of individual host addresses, we assume that there are N unique pairs of /24 IP prefix blocks in the Internet. Because individual addresses in the same /24 prefix blocks are very likely to be assigned and managed by the same administrative entity, we expect that hosts in the same /24 block are likely to experience similar performance, such as network delays and packet losses. We choose a simple random sampling without replacement over N unique pairs.

We derive the sample size n of host pairs in order to guarantee a certain level of accuracy in the delay distribution estimation. As we expect the delay distribution not to follow a normal distribution– from our empirical data we observe that delay distributions are heavy tailed–, median is a better metric than mean.

Given n samples of round-trip delays, $y_1, y_2, \ldots, y_n$, we estimate the median of the population (q_m) using the order statistics ($y_{[1]} \leq y_{[2]} \leq \ldots \leq y_{[n]}$). Then, the estimator of the median is defined by $\hat{q}_m = y_{[\lceil n/2 \rceil]}$. The distribution of $\hat{q}_m$ around the true value (q_m) approaches a normal distribution asymptotically as the sample size n grows (see chapter 2.3.3 of [13]). The estimator is also known to be unbiased ($E[\hat{q}_m] = q_m$) and consistent ($\hat{q}_m \rightarrow q_m$ as $n \rightarrow \infty$) [3]. Then, the $100 \times (1 - \alpha)$% confidence interval of median estimator is given by

$$\hat{q}_m \pm z_\alpha \frac{\sqrt{0.5(1-0.5)}}{Pr[Y = q_m] \cdot \sqrt{n}} \quad (1)$$

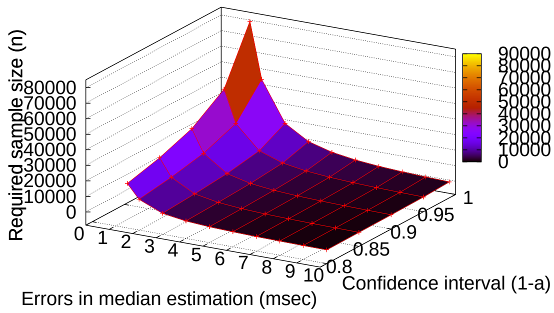


Figure 2: Sample size when $Pr[Y = q_m] = 0.003$

Except for the $Pr[Y = q_m]$, all variables in (1) are easily derived from the samples. Because we do not make any assumption on the population distribution, we do not know $Pr[q_m]$ in advance. We choose to approximate $Pr[q_m]$ from our empirical delay distribution from $100,000$ sample pairs. We have observed that the empirical observation of $Pr[q_m]$ is normally distributed, and the value converges as we increase the number of observations.

In Figure 2, we illustrate how many samples are required for the estimate to fall within the confidence interval. In this figure, we use $Pr[q_m] = 0.003$ that we have observed in 2009. We see that the sample size of $n = 50,000 \sim 60,000$ shows very small errors (about 1 msec) for a very tight confidence interval for $\alpha = 0.99$.

In this work, we choose the number of sample size $n = 100,000$ to maximize the accuracy of estimation. In the next Section, we will see that the average success rate of path stitching with random host pairs are about 65%. That is, when we try $100,000$ random host pairs, we successfully estimate path and delays for about $65,000$ host pairs, and it still provides very small errors (about 1 msec) for the 95% confidence interval.

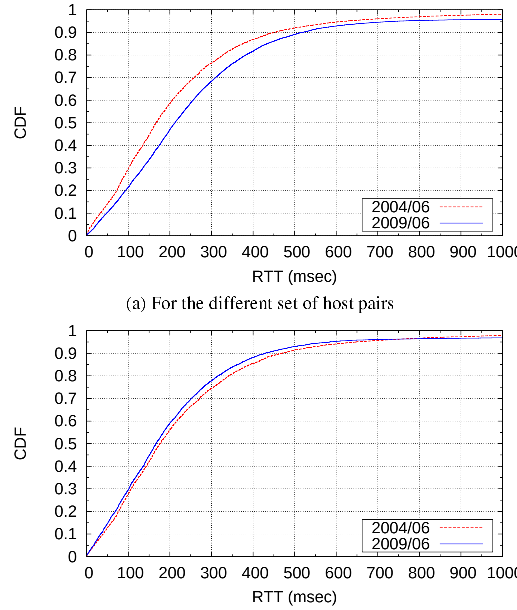
3. PRELIMINARY RESULTS

To get a sense of the feasibility about analyzing the Internet delay history, we take a quick look at the delay distributions in 2004/06 and in 2009/06. In this section, as a preliminary result for the work, we examine the observed differences between two distributions, and provide possible explanations.

We extract all /24 routable IP prefixes from the BGP routing table snapshots in 2004 and in 2009. Total number of announced /24 prefixes in 2004 and in 2009 are $5,170,229$ and $10,071,994$, respectively. A random sample of 100,000 host pairs have been drawn from those all routable /24 prefixes, and delay estimates are produced by path stitching. We have observed that path stitching successfully estimated paths and delays for the 67% in 2004 pairs and 65% in 2009 pairs.

In Figure 3(a) we plot the CDF of round-trip delay distributions for 2004/06 and 2009/06. We show that overall delay distribution got worse in 2009 than 2004. The median delays are 166.0 msec and 213.0 msec, respectively. Where does the difference of 50 msec median delay come from? Internet has been expanding in terms of hosts, ASes, countries, or geographic region. IP address usage must have expanded from 2004 to 2009. Some of /24 prefixes of those hosts with large delays in 2009 may not existed in 2004. Similarly, ASes of those hosts in 2009 may not existed in 2004.

Then what if we choose the same set of host pairs for 2004 and for 2009. We pick $100,000$ host pairs in 2004/06 again and use the same set of host pairs for 2009/06. We have observed that $41,905$ pairs always responded. We plot the result in Figure 3(b). We observe the opposite results; delay distributions for the same set of sample host pairs got slightly better in 2009 than 2004. In this case, the median delays are 173.0 msec and 163.4 msec, respectively.



(a) For the different set of host pairs

(b) For the same set of host pairs

Figure 3: Delay Distributions, 2004/06 and 2009/06

Where does the 10 msec improvement in median delay from 2004 to 2009 come from? One possible explanation for the 10 msec improvement is at the transmission rate upgrade. Back-of-the-envelope calculation of transmission time of a $1,500$ bytes packet over 1 Mbps is 12 msec and over 10 Mbps is 1.2 msec. Can we say that the major transmission technology has evolved from 1 Mbps to 10 Mbps in the past six years? We do not have concrete evidences for the claim.

Regional growth of the Internet

Why has the Internet delay gotten worse from 2004 to 2009 for independent samples of host pairs? We find a possible reason for this global trend from the newly appeared ASes or prefixes. The prefixes and ASes of those hosts with large delays in 2009 may not exist in 2004, and they can cause the overall Internet performance to degrade.

In Figure 4 we plot the geographic regional distribution of host pairs in 2004 and 2009. In the figure, we can see that the fraction of host pairs in North America (NA-NA in the figure) decreased significantly from 40% to 20%. Interestingly, the fractions of all other regional pairs increased (except for the North America - Oceania pair that has remained constant).

The effect of the change in the geographical distribution of sample hosts becomes clear with Figure 5(a). In the figure, we plot the CDF of delays for the North America - North America pairs. The delay distributions for those pairs in 2004 and 2009 are almost identical. It does not mean that the delay distributions for the other regional pairs have gotten worse. For example, in Figure 5(b), the delay performance for the Africa - Europe pairs for most part improved significantly from 2004 to 2009. But 10% of Africa - Europe pairs experience delays than 1 sec in 2009. Their median delay is still much larger than that of intra North America pairs.

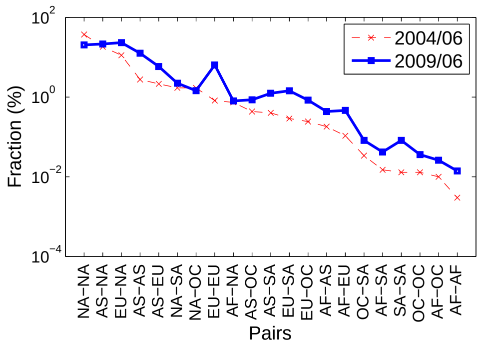
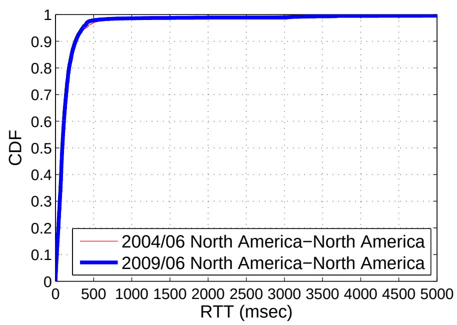
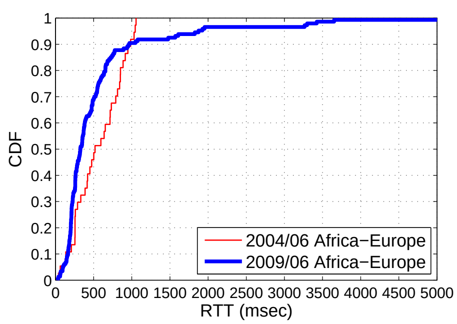


Figure 4: Geographic regional distribution of host pairs (AS: Asia, AF: Africa, EU: Europe, OC: Oceania, NA: North America, SA: South America



(a) North America - North America pairs



(b) Africa - Europe pairs

Figure 5: Delay distributions for specific regional pairs

Even though we do not have exact information about the access technologies of newly emerging prefixes or ASes, the expansion of the Internet out of North America to far wider regions of the globe is a likely cause behind the mean delay increase from 2004 to 2009.

4. SUMMARY AND FUTURE WORK

In this work, we present the methodology for the Internet delay history analysis with the existing measurements and our path-stitching algorithm. We demonstrate that our approach is feasible and gives insights into the overall Internet delay distributions for the past as well as the present one.

Future work will focus on rigorous statistical analysis about the sources of error in our approach. As well as the sampling errors from the restricted number of sample size, the effect of non-respondents (about 35% of sampled host pairs in our work) and the measurements error from the path stitching should be carefully considered together.

We will also incorporate additional datasets from NLANR [10], RocketFuel [15], and iPlane projects. We will see the trend from 1999 to 2009, and match it with the Internet-wide upgrades, such as new undersea technology developments or DSL/cable deployment. This would allow us to better understand the perspectives on Internet performance growth. We expect significant change delay during the first half of the decade when the Internet experienced the unprecedented growth and the delay distributions from 1999 to 2004 would offer us a very insightful perspective to the evolution of Internet performance.

5. REFERENCES

[1] Advanced network technology center and University of Oregon. The RouteViews project. http://www.routeviews.org.

[2] P. Borgnat, G. Dewaele, K. Fukuda, P. Abry, and K. Cho. Seven years and one day: Sketching the evolution of Internet traffic. In *IEEE INFOCOM*, Rio de Janeiro, Brazil, April 2009.

[3] B.-Y. Choi, S. Moon, R. Cruz, Z.-L. Zhang, and C. Diot. Quantile sampling for practical delay monitoring in Internet backbone networks. *Computer Networks*, 51:2701–2716, 2007.

[4] A. Dhamdhere and C. Dovrolis. Ten years in the evolution of the Internet ecosystem. In *ACM SIGCOMM IMC*, October 2008.

[5] J. Heidemann, Y. Pradkin, R. Govindan, C. Papadopoulos, G. Bartlett, and J. Bannister. Census and survey of the visible Internet. In *ACM SIGCOMM IMC*, October 2008.

[6] Y. Hyun, B. Huffaker, D. Andersen, E. Aben, C. Shannon, M. Luckie, and kc claffy. The CAIDA IPv4 Routed /24 Topology Dataset. http://www.caida.org/data/active/ipv4_routed_24_topology_dataset.xml.

[7] kc claffy, T. Monk, and D. McRobb. Internet tomography. *Nature, Web Matters*, January 1999.

[8] D. Lee, K. Jang, C. Lee, G. Iannaccone, and S. Moon. Path stitching: Internet-wide path and delay estimation from existing measurements. In *IEEE INFOCOM mini-conference*, March 2010.

[9] H. V. Madhyastha et al. iPlane: An information plane for distributed services. In *USENIX OSDI*, November 2006.

[10] NLANR. Active measurement project. http://watt.nlanr.net.

[11] R. Oliveira, B. Zhang, and L. Zhang. Observing the evolution of Internet AS topology. In *ACM SIGCOMM*, Kyoto, Japan, August 2007.

[12] RIPE NCC. The Routing Information Service. http://www.ripe.net/ris.

[13] R. J. Serfling. *Approximation Theorems of Mathematical Statistics*. Wiley, 1980.

[14] Y. Shavitt et al. DIMES: Let the Internet measure itself. *ACM SIGCOMM CCR*, 35(5):71–74, 2005.

[15] N. Spring et al. Measuring ISP topologies with Rocketfuel. In *IEEE/ACM Transactions on Networking*, 2004.

[16] B. Zhang, T. S. E. Ng, A. Nandi, R. Riedi, P. Druschel, and G. Wang. Measurement-based analysis, modeling, and synthesis of the Internet delay space. In *ACM SIGCOMM IMC*, October 2006.