Pitfalls for Testbed Evaluations of Internet Systems

David R. Choffnes
Northwestern University
drchoffnes@eecs.northwestern.edu

Fabian E. Bustamante
Northwestern University
fabianb@eecs.northwestern.edu

ABSTRACT

Today's open platforms for network measurement and distributed system research, which we collectively refer to as *testbeds* in this article, provide opportunities for controllable experimentation and evaluations of systems at the scale of hundreds or thousands of hosts. In this article, we identify several issues with extending results from such platforms to *Internet wide* perspectives. Specifically, we try to quantify the level of inaccuracy and incompleteness of testbed results when applied to the context of a large-scale peer-to-peer (P2P) system. Based on our results, we emphasize the importance of measurements in the appropriate environment when evaluating Internet-scale systems.

Categories and Subject Descriptors

C.2.5 [**Communication Networks**]: Local and Wide-Area Networks–Internet; C.4 **[Performance of Systems]:** Measurement techniques

General Terms

Experimentation, Performance, Measurements.

Keywords

Internet-scale systems, peer-to-peer, evaluation.

1. INTRODUCTION

Today's open platforms for network measurement and distributed system research, which we collectively refer to as *testbeds* in this article, provide opportunities for controllable experimentation and evaluation of systems at the scale of hundreds or thousands of hosts [8, 15, 20]. These testbeds have been successfully used to advise a variety of important applications addressing IP reachability, prefix hijacking and routing anomalies.

Recent studies, however, suggest that testbed results for Internet systems do not always extend to the targeted deployment. For example, Ledlie et al [11] and Agarwal et al. [1] show that network positioning systems perform much worse "in the wild" than in PlanetLab deployments.

In this article, we identify several limitations of the accuracy and completeness of testbed results when applied to the context of a large-scale peer-to-peer (P2P) system. To inform this study, we use a unique collection of traces gathered from a deployment containing hundreds of thousands of users located at the network edge.

We focus our analysis on the following three issues that affect the validity of any study extending testbed results to an Internet scale. First, we find that large and significant portions of the Internet topology used by P2P systems are invisible to research testbeds, limiting the effectiveness of testbed-based inference of Internet paths and relationships between autonomous systems (ASes). Next, we show that inferred properties of these topologies (latencies and throughput) are inaccurate. Finally, we discuss how these issues prevent accurate evaluations of performance for distributed systems running on these topologies.

Based on our results, we argue for a third stage for evaluating Internet-based systems – beyond emulation and overlay testbeds – that includes edge-based measurements. There is a number of approaches to achieve this goal, including the use of application-level and network-level traces from the edge of the network (e.g., via MLab[1] and Ono datasets). We also encourage the design and deployment of additional edge-based monitoring services, either built into existing distributed services or provided independently with the proper incentives for Internet-scale adoption.

2. EDGE SYSTEM TRACES

The Internet is growing in ways that make increasingly difficult to attain a global view of the network. Large swathes of the network cannot be probed directly from our research testbeds and a number of valuable measurement techniques have side effects that render them impractical.

For our study, we address this issue using network- and application-level traces from BitTorrent users running the Ono plugin [4]. In particular, our software passively records the volume of data transferred over each host's connection and performs active ping and traceroute measurements to a subset of these connections.

Our installed user base has grown to cover 204 countries, 53,000 routable prefixes and more than 7,000 ASes since December, 2007. The collection of traces used in this study was contributed by a subset of the total installed subscribers, consisting of tens of thousands of users online during the measurement period. While the amount of data collected per unit time varies according to the online user population, each day we record between 2.5 and 3.5 million traceroutes, tens of millions of latency measurements and more than 100 million per-connection transfer-rate samples. As part of this work, we are making this dataset available to researchers through our EdgeScope project.[2]

In the following sections we use this dataset to explore several key pitfalls of testbed-based evaluations for Internet scale systems. We begin by exploring the completeness of the view from such testbeds.

3. GENERALIZING NETWORK VIEWS

[1] http://www.measurementlab.net/

[2] http://www.aqualab.cs.northwestern.edu/projects/EdgeScope.html

Tier-1	Customer-Provider	Peering
3.14%	12.86%	40.99%

Table 1: Percent of links missing from public views, but found from edge systems, for major categories of AS relationships.

A number of studies explicitly or implicitly rely on network topologies for estimating Internet resiliency, inferring Internet paths and estimating cross-network traffic costs, to name a few. While several research efforts have successfully extracted detailed topologies from public vantage points (i.e., the *public view*), it is well known that these Internet views are incomplete [2, 3, 19].

In previous work, we explored a lower bound for missing topology information using AS-level paths and AS relationships inferred from traceroute data gathered from hundreds of thousands of users located at the edge of the network [3]. We briefly summarize some of our key finding regarding the portion of new AS links we found, then provide new results indicating the impact of these links when evaluating Internet scale systems.

Missing links. For this analysis of missing links, we used a dataset that includes probes from nearly 1 million source IP addresses to more than 84 million unique destination IP addresses, all of which represent active users of the BitTorrent P2P system. By comparison, the BitProbes study [10] used a few hundred sources from the PlanetLab testbed to measure P2P hosts comprising 500,000 destination IPs. Naturally, the number of vantage points available from edge systems in our dataset far outnumbers those from public views, particularly for lower tiers of the Internet hierarchy where most of the ASes reside.

This unique perspective allows us to identify links invisible to the public view; in total, we found 20% additional links missing from the public view. The vast majority of these links were located below Tier-1 ASes in the Internet hierarchy with, not surprisingly, their number increasing in lower tiers.

In addition to the locations of links in the Internet hierarchy, it is useful to understand what kinds of AS relationships are included in these missing links. Table 1 categorizes links into Tier-1, customer-provider or peering links and shows the missing links as a fraction of existing links in the public view, for each category. Note that there is a large number of additional peering links (44%) and a significant fraction of new customer-provider links (12%). Based on these results, it seems clear that when evaluating the interaction between network topologies and Internet systems at the edge – often located in lower Internet tiers – testbed-based topologies are less likely to include many relevant links and relationships.

Impact of missing paths. To better understand how this missing information affects studies of Internet-scale systems, we investigate the impact of missing links using three weeks of connection data from P2P users. In particular, we try to determine how much of these users' traffic volumes can be mapped to AS-level paths – an essential step for evaluating P2P traffic costs and locality.

We begin by determining the volume of data transferred over each connection for each host, then we map each connection to a source/destination AS pair using the Team Cymru service [23]. We use the set of paths from public views and P2P traceroutes [3] and, finally, for each host we determine the portion of its traffic volume that could not be mapped to *any* AS path in our dataset.[3]

Figure 1 uses a cumulative distribution function (CDF) to plot these unmapped traffic volumes using only BGP data (labeled *BGP*) and the entire dataset (labeled *All*). The figure shows that

[3]For simplicity, we assume that publicly announced BGP paths coincide with those that data actually traverses.

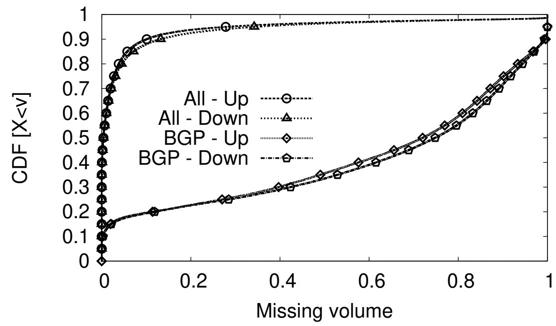


Figure 1: CDF of the portion of each host's traffic volume that could not be mapped to a path based on both public views and traceroutes between a subset of P2P users.

when using only BGP information the median volume of unaccounted traffic is nearly 75%. In fact, complete path information is available for only 7.4% of hosts and 7.3% of hosts use connections for which BGP data provides no path information. When using *All* path information, we cannot locate complete path information for 66% of hosts; fortunately, the median portion of traffic for which we cannot locate an AS path is only 0.4%. From the set of hosts in our dataset, 4% of them use connections for which we have path information for only half of their traffic volumes; less than 2% use connections for which we have no path information at all.

While our topology data adds 20% more links to the Internet graph, it allows us to map an order of magnitude more P2P traffic than using BGP alone. The difference is not due to a failure of existing path measurements, but rather the limitations of their coverage. As these results show, one must use caution when drawing conclusions from an Internet wide study from today's testbed environments.

We also note that there is a small portion of traffic (3.8% on average) that cannot be mapped even with the links we add to the public view – this occurs because our traceroute measurements are issued to a randomly selected subset of connected P2P users. While this is a non-negligible volume of unmapped traffic, we do not expect it to significantly impact our conclusions in the remainder of this article because our path measurements are performed at random and thus minimize any bias in missing paths.

4. GENERALIZING MEASUREMENTS

While the previous section showed that large portions of the network are invisible to current testbeds, in this section we illustrate how properties of topology links measured from testbed vantage points do not extend to those measured from the edge of the network. We begin by focusing on estimating distances between Internet hosts, which is essential to a variety of network performance optimizations including server selection, central leader election and connection biasing. We close the section by examining Internet-wide achieved throughput as measured by BitTorrent throughput from users at the edge of the network, which is essential for modeling and simulating system dynamics.

Network distances. There is a large body of research addressing the issue of how to measure, calculate and encode Internet distances in terms of round-trip latencies [7, 13, 21, 22]. Generally, these solutions rely on methods to predict latencies between arbitrary hosts without requiring the N^2 number of measurements that provide ground-truth information. Previous work has identified the following key properties that impact network positioning performance: the structure of the latency space, the impact of last-mile delays and the rate of triangle-inequality violations (TIVs) in the latency space. We now review recent results showing that these key properties are significantly different when measured exclusively from edge systems compared to those measured from testbed environments [5].

We base our results on 2 billion latency samples gathered from edge systems during June 10–25th, 2008. Unlike studies that use PlanetLab hosts to measure latencies or infer them based on latencies between DNS servers [9], this dataset consists exclusively of directly measured latencies between edge systems. It is also an order of magnitude larger than the set used by Agarwal et al. [1] to evaluate server selection in gaming systems.

Latencies. To begin, Fig. 2 compares the average latencies seen by hosts using the Ono plugin (labeled P2P) to those seen from three related projects: the RON testbed (MIT), PlanetLab (PL) and Ledlie et al.'s study (PL-to-P2P). The graph shows that latencies from edge systems are generally much larger than those from MIT King [7] and PlanetLab (PL). In fact, the median latency in our dataset is twice as large as reported by the study by Ledlie et al. [11], which used PlanetLab nodes to probe Vuze P2P users (PL-to-P2P).

To determine whether the latency distribution was affected by P2P traffic, we compared our results with those gathered only during periods when hosts were not transferring data. While we did find differences in latencies between peers actively transferring data and those in the complete dataset, the measured difference in median latencies was less than 10%.

Because a large portion of the difference in latencies remains unexplained, we investigate whether they are due to the fact that paths from PlanetLab to the rest of the Internet are significantly different from those between networks at the edge. For example, PlanetLab hosts are known to offer higher speed interconnections across countries and continents than what is generally available to ISPs hosting home users. Further, such testbed hosts are not subject to last-mile factors such as low-performance middleboxes and access technologies such as interleaving in DSL and contention/queuing for the shared medium in cable broadband.

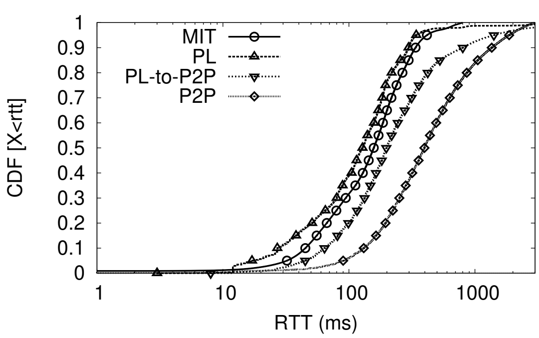


Figure 2: CDFs of latencies from different measurement platforms (semilog scale). Our measurement study exclusively between peers in Vuze (labeled P2P) exhibits double the median latency "in the wild" (labeled PL-to-P2P).

Last-mile effects. To analyze last-mile effects, we dividr the traceroute-based IP-level path between hosts into quartiles and determining the portion of the end-to-end latency contained in each quartile. If the latency were evenly distributed among IP hops along a path, each quartile would contain 25% of the end-to-end latency. In contrast, the first quartile (which is very likely to contain the entire first mile) accounts for disproportionately large fractions of the total end-to-end latency. For instance, when looking at the median values, the first quartile alone captures 80% of the end-to-end latency. The middle two quartiles, in contrast, each account for only 8%. [4]

Also note that the first quartile (and a significant fraction of the last quartile) has a large number of values close to and larger than 1. This demonstrates the variance in latencies along these first and last miles, where measurements to individual hops along the path can yield latencies that are close to or larger than the total end-to-end latency (as measured by probes to the last hop). In fact, more than 10% of the first quartile samples have a ratio greater than 1. While the performance impact of the last-mile is well known, the problem is particularly acute in typical network edge settings. However, most of today's network positioning systems either ignore or naively account for the severity of this issue. For instance, while Vivaldi uses "height" to account for (first- and) last-mile links [7], this analysis suggests that a single parameter is insufficient due to the large and variable latencies in a large-scale P2P environment.

Triangle-Inequality Violations. TIVs in the Internet delay space occur when the latency between hosts A and B is larger than the sum of the latency from A to C and C to B ($A \neq B \neq C$). This is caused by factors such as network topology and routing policies (see, for example, [12, 22]). Wang et al. [24] demonstrate that TIVs can significantly reduce the accuracy of network positioning systems.

We performed a TIV analysis on our dataset and found that over 13% of the triangles had TIVs (affecting over 99.5% of the source/destination pairs). Lumezanu et al. [12] study the dynamics of TIVs and demonstrate that using the minimum RTTs, as done in this study, is likely to underestimate the rate of TIVs. Thus our results can be considered a lower bound for TIVs in a large-scale P2P environment.

Compared to TIV rates reported in an analysis of datasets from Tang and Crovella [22], TIV rates in the P2P environment we studied are between 100% and 400% higher, and the number of source/destination pairs experiencing TIVs in our dataset (nearly 100%) is significantly greater than the 83% reported by Ledlie et al. [11]. These patterns for TIVs and their severity hints at the challenges in accounting for TIVs in coordinate systems. In Section 5 we show the impact of these violations on their accuracy.

Bandwidth capacities. Bandwidth capacities are an important factor in the design of distributed systems, from making encoding decisions in video streaming to informing peer selection in P2P systems. While there are many proposals for estimating capacities, these techniques are not amenable to widespread studies due to limitations on measurement traffic volumes and the need for compliant endpoints. Further, previous work has cast doubts on their accuracy [16].

Perhaps more important than raw capacities, the maximum transfer rates that hosts actually can achieve are essential for modeling the dynamics of any large-scale data-sharing system such as P2P file sharing (e.g., when modeling BitTorrent using only tracker information [6]). Of course, these achieved transfer rates can be affected by ISP interference (e.g., traffic shaping), limited available bandwidth in the P2P system and user-specified limits on the maximum throughput consumed by a P2P application. We now investigate how these rates compare with bandwidth capacities estimated by tools run from testbed environments.

[4]The first and last quartiles are not symmetric because large numbers of hosts reside behind middleboxes that block traceroute probes.

After removing samples containing user-specified limits on transfer rates, we find the maximum upstream and downstream transfer rates seen by each host during a three-week period in April, 2009. We base our analysis on transfer-rate samples taken every 30 seconds; during the measurement period, 90% of hosts were online for more than 15 minutes and median session times were on the order of hours. In addition to quantifying achieved transfer rates, our results represent a lower bound for each host's bandwidth capacity.

Figure 3 depicts a CDF of maximum upstream and downstream throughput seen for each host in our study. First, we note the lack of step-like functions in the CDFs, which would occur if BitTorrent were, as commonly believed, most often saturating the full bandwidth capacity. Thus, while BitTorrent attempts to saturate each user's downstream bandwidth capacity, in practice it does not always do so.

We also find that the median upstream rate is 54 KB/s while the median for downstream rates is 102 KB/s. While these values indicate the effects of asymmetric bandwidth allocation, typically such allocations offer order-of-magnitude larger downstream rates. The smaller difference in these two values seems to indicate that the transfer rates achieved by P2P systems are thus limited by the peers' upstream capacities.

It is important to note that these CDFs do not imply that the ratio of upstream to downstream capacities is greater than 0.5 for most hosts – some of the above samples contain only upstream transfer rates. For those hosts where we can measure both upstream and downstream throughputs, we find that the median ratio is 0.32 and the 90th percentile ratio is 0.77. This is in line with the asymmetric bandwidth allocations typical of DSL and cable Internet technologies being used by the majority of our vantage points.

We now compare these lower-bound estimates of capacities with those measured from PlanetLab in 2006 as reported by Isdal et al. [10]. One of their findings is that 70% of hosts have an upload capacity between 350Kbps and 1Mbps. Studies from the ITU[5] indicate that bandwidth capacities have increased by about 35% per year in the subsequent three years. Thus, even if BitTorrent consumed only a fraction of a host's bandwidth capacity, we would expect achieved transfer rates to be equal to or larger than the measured capacities from the BitProbes study.

Based on results from P2P users, however, we surprisingly find that only 45% of hosts in our study achieve such transfer rates. In fact, 40% of hosts in our study achieve *less than* 350 Kbps maximum upstream rates. This suggests that even if the testbed-based bandwidth capacity measurements were accurate, they are insufficient for predicting *achieved* transfer rates in a P2P system.

For completeness, Fig. 3 shows that *downstream* rates closely track upstream rates until after the 30th percentile, where downstream rates significantly exceed upstream ones. We cannot compare these values with Isdal et al. because their approach does not extend to downstream rates at the edge of the network.

Finally, we analyze the maximum throughput achieved by hosts grouped by country in Fig. 4. We find that hosts in Germany, Romania and Sweden achieve the highest transfer rates while those in India, the Philippines and Brazil achieve the lowest. This is in line with results from independent bandwidth tests from Speedtest.net, indicating that maximum transfer rates measured from P2P users, when grouped by location, are in fact predictive of the bandwidth capacity *rankings*.

In summary, our study of achieved throughput in BitTorrent

[5] http://www.itu.int/publ/D-IND/en

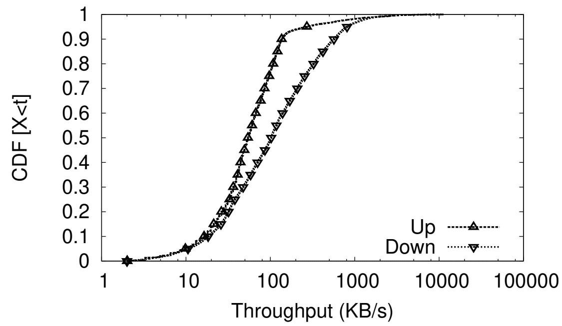


Figure 3: CDF of transfer rates for all users, where the median is only 50 to 100 KB/s. This suggests that the BitTorrent system is dominated by mid-to-low-capacity hosts.

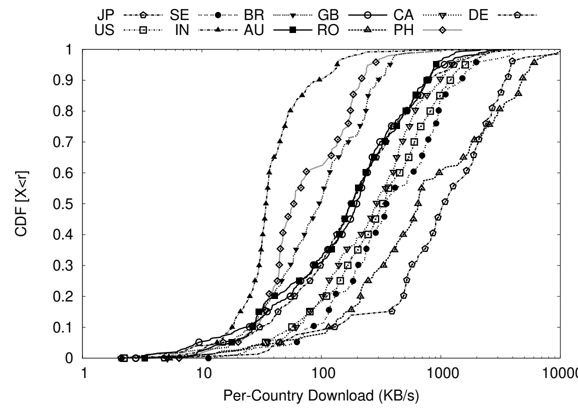


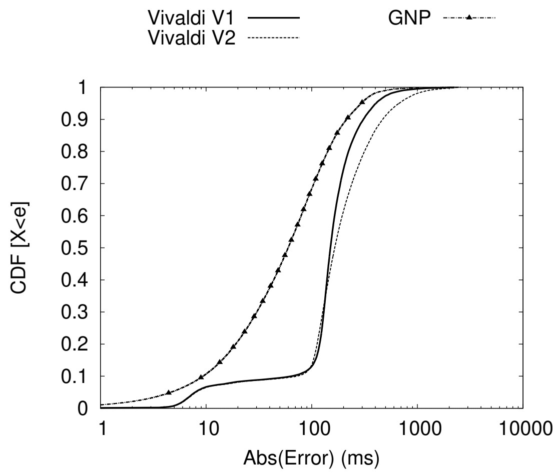
Figure 4: Per-country throughput CDFs, showing that Germany, Romania and Sweden have the highest average capacities while India, the Philippines and Brazil have the lowest.

indicates that bandwidth capacity measurements from testbeds are likely to overestimate achieved throughput at the edge, while this throughput tends to correlate with bandwidth capacity rankings across countries. Since achieved throughput significantly impacts the network flows generated by many distributed systems (e.g., P2P file sharing), using bandwidth capacities can lead to misleading conclusions about system dynamics. For example, we show in the next section that our empirically derived flows produce significantly different results in terms of P2P traffic costs when compared to estimates in previous work.

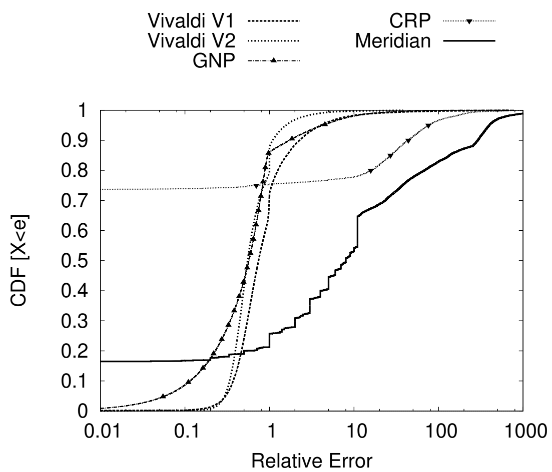
5. INFERRING SYSTEM PERFORMANCE

While measurements from the edge of the network help us better understand network topologies, delay behavior and bandwidth capacity distributions, they also are essential to designing, evaluating and optimizing distributed systems that run in this environment. We now show how more accurate views of the edge of the network affect system performance when compared to evaluations conducted from testbed environments.

Network positioning. We begin with network positioning systems and determine how the latency space measured in the previous section affects accuracy for a variety of positioning systems including GNP [13], Vivaldi [7], Meridian [25] and CRP [21].



(a) Absolute error.



(b) Relative error.

Figure 5: Absolute value of errors between estimated and measured latencies, in milliseconds (right), and absolute value of *relative* errors between estimated and measured latencies (left).

The Vivaldi and CRP systems are implemented in our measurement platform, so their values represent true, "in the wild" performance. For evaluating GNP performance, we use the authors' simulation implementation. The results are based on three runs of the simulation, each using a randomly chosen set of 15 landmarks, 464 targets and an 8-dimensional coordinate space. We also simulate Meridian using settings proportional to those in the original evaluation, with 379 randomly selected Meridian nodes, 100 target nodes, 16 nodes per ring and 9 rings per node. Our results are based on four simulation runs, each of which performs 25,000 latency queries.

We begin our analysis by evaluating the accuracy of GNP and of the Vuze Vivaldi implementations in terms of errors in predicted latency. Meridian and CRP are omitted here because they do not provide quantitative latency predictions. Figure 5(a) presents the cumulative distribution function (CDF) of errors on a semilog scale, where each point represents the absolute value of the *average* error from one measurement host. We find that GNP has lower measurement error (median is 59.8 ms) than the original Vivaldi implementation (labeled V1, median error is $\approx$ 150 ms), partially due to GNP's use of fixed, dedicated landmarks. Somewhat surprisingly, Ledlie et al.'s Vivaldi implementation (labeled V2) has slightly larger errors in latency (median error is $\approx$ 165 ms) than GNP and V1; however, we show in the next paragraph that its relative error is in fact smaller.

Relative error, the difference between the expected and measured latency, is a better measure of accuracy for network positioning systems. To compute relative errors, we first calculate the absolute value of the relative error between Vivaldi's estimated latency and the ping latency for each sample, then find the average of these errors for each client running our software. Fig. 5(b) plots a CDF of these values; each point represents the average relative error for a particular client. For Vivaldi V1, the median relative error for each node is approximately 74%, whereas the same for V2 is 55% – both significantly higher than the 26% median relative error reported in studies based on PlanetLab nodes [11]. Interestingly, the median error for Vivaldi V2 is approximately the same as for GNP, indicating that decentralized coordinates do not significantly hurt relative performance. Finally, because Meridian and CRP do not predict distances, Fig. 5(b) plots the relative error for the closest peers they found. Meridian finds the closest peer approximately 20% of the time while CRP can locate the closest peer more than 70% of the time.

In summary, network positioning systems that rely on predicting latencies perform worse than those using direct measurement and relative positions, as expected from the latency-space analysis in Section 4.

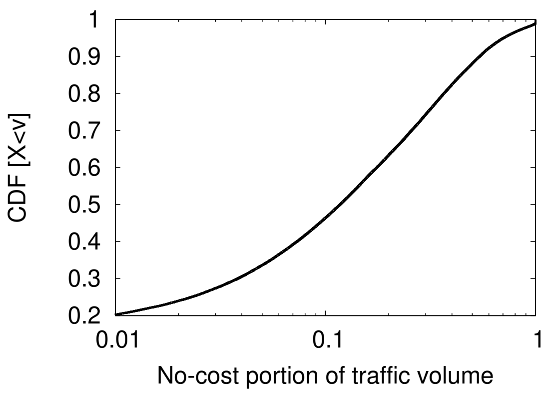
ISPs costs/revenue from P2P file sharing. Large traffic volumes generated by P2P file-sharing systems have generated a great deal of publicity as network providers attempt to reduce their costs by blocking, shaping or otherwise interfering with P2P connections. Given the popularity of these systems, a number of research efforts have investigated this issue by designing systems to reduce cross-network traffic [4, 26] and evaluate the potential for P2P traffic locality [14].

Most previous work in this area relies on limited deployments and/or simulation results to estimate network costs of P2P systems. We now show how measurements from a large-scale, live deployment – combined with more complete AS topology information – provides a different view of the costs incurred by these systems.

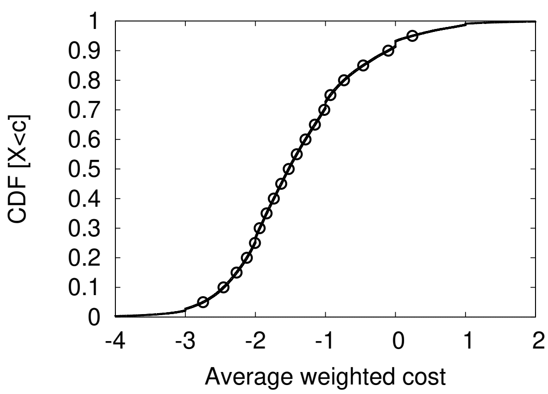
A number of studies estimate the costs of P2P traffic as proportional to the number of AS hops along paths to different hosts. In this context, traffic is considered *no-cost* (also referred to as *local*) if it stays entirely in the same AS. We now refine this metric to include all paths for which no hop contains a customer-provider relationship (or vice-versa); i.e., our definition of no-cost includes traffic that remains in the origin AS or traverses peering and sibling links. This is based on the assumption that traffic costs borne by ISPs occur at customer/provider edges.

Figure 6(a) presents a CDF of the portion of each P2P user's traffic that is no-cost. Our results from 130,000 source IPs and 12 million destination IPs indicates that the vast majority of hosts naturally generate at least some no-cost traffic. This result contradicts those from Piatek et al. [14], who use inferred testbed-based results and a single deployed vantage point to question the effectiveness of reducing ISP costs in P2P systems. In fact, we find that the majority of traffic volumes are no-cost for a significant fraction (12%) of hosts.

It is important to note that these results may be biased by the fact that the measured hosts are using Ono to preferentially use no-cost peer connections. Due to the nature of Ono, we cannot control for cases where higher-cost paths can be used. As a result, the portion of no-cost traffic we measure may be larger than for



(a) Portion of “no-cost” traffic generated per host.



(b) Internet-wide costs incurred by BitTorrent traffic.

Figure 6: CDF of portion of “no-cost” traffic generated per host (right), and estimated Internet-wide costs incurred by BitTorrent traffic (left). The vast majority of hosts generate at least some no-cost traffic while the majority of traffic volumes are no-cost for 12% of hosts. Further, our results show that P2P traffic has a net effect of generating significant revenue for provider ISPs.

hosts not using the plugin. On the other hand, of the total number of connections, those that are biased are relatively small and this reduces the potential impact on our results.

Finally, Fig. 6(b) plots the average cost per byte for each user, based on the net costs of P2P traffic according to the traffic volumes per path and AS relationships along each path. Specifically, the cost of a path is the sum of the cost of each AS hop, where a hop between customer and provider is assigned a cost of 1, provider to customer a cost of -1 and zero otherwise (sibling and peer AS hops). While our cost model is simple, it allows us to perform the first analysis of traffic costs based on the common practice of charging for transit based (at least in part) on traffic volumes crossing customer/provider links. We then determine, for each host, the portion of all traffic volume generated by each of its connections and multiply this by the cost of the path. Each point in Fig. 6(b) represents the sum of these values for each host.

As the figure shows, the vast majority of hosts generate flows with a net effect of generating revenue (i.e., negative costs) for ISPs. While this result is in agreement with commonly held notions that P2P traffic has generated revenue for ISPs (particularly those in tier-1), we believe that we are the first to attempt to quantify this effect. We leave a study of which ISPs are benefiting from this (and by how much) as part of our future work.

6. WHERE DO WE GO FROM HERE?

While this article focused on limitations for testbeds, current edge measurement platforms are not without their own restrictions. Our own view from a large deployment on edge systems still does not provide complete network coverage, nor does it allow controlled experimentation. These are challenges that affect *all* approaches for edge measurement.

Further, our platform does not provide arbitrary measurement (by design), and our view of the network is restricted to what is offered by passive measurements from BitTorrent and limited active probing, only while users run our software. These limitations are not inherent to edge-based measurement in general and, given the need for edge-based views of the network, we believe that an important area of future work is addressing them in the design of new distributed research platforms.

While new work in this area can draw from the lessons we learned in our own deployment (e.g., dealing with data collection at scale and convincing users to install software), there is a large number of opportunities for new research in issues such as tradeoffs between security, privacy and experimental control, and how to manage and mine the data collected from such platforms.

Others in our community are also exploring opportunities to incorporate edge-based measurements when evaluating Internet wide systems. In addition to our EdgeScope project, related research efforts such as DipZoom [17], ShaperProbe[6], C’MON[7], HMN [18] and Glasnost[8] are collecting views of the edge of the network. We hope that public releases of their datasets will enable new evaluations of network properties and system performance in representative environments.

7. CONCLUSION

This article discussed potential issues with extending results from limited platforms to Internet wide perspectives. In particular, we showed that testbed-based views of Internet paths are surprisingly incomplete, the properties of these paths do not extend to the edge of the network and these inaccuracies have a significant impact on inferred system-wide performance for services running at the edge. These results make a strong case for research in new evaluation strategies for Internet-scale systems, both through edge-systems traces (such as those available via our EdgeScope project) and new evaluation platforms.

8. ACKNOWLEDGMENTS

We would like to thank the anonymous reviewers for their helpful comments. This work was supported by the National Science Foundation under CAREER Award CNS 0644062 and Award CNS 0917233.

9. REFERENCES

[1] AGARWAL, S., AND LORCH, J. R. Matchmaking for online games and other latency-sensitive P2P systems. In *Proc. of ACM SIGCOMM* (2009).

[2] BUSH, R., MAENNEL, O., ROUGHAN, M., AND UHLIG, S. Internet optometry: assessing the broken glasses in internet reachability. In *Proc. of IMC* (2009).

[6] http://www.cc.gatech.edu/ partha/diffprobe/shaperprobe.html

[7] http://cmon.grenouille.com/

[8] http://broadband.mpi-sws.org/transparency/bttest.php

[3] CHEN, K., CHOFFNES, D., POTHARAJU, R., CHEN, Y., BUSTAMANTE, F., AND ZHAO, Y. Where the sidewalk ends: Extending the Internet AS graph using traceroutes from P2P users. In *Proc. of ACM CoNEXT* (2009).

[4] CHOFFNES, D. R., AND BUSTAMANTE, F. E. Taming the torrent: A practical approach to reducing cross-ISP traffic in peer-to-peer systems. In *Proc. of ACM SIGCOMM* (2008).

[5] CHOFFNES, D. R., SANCHEZ, M. A., AND BUSTAMANTE, F. E. Network positioning from the edge: An empirical study of the effectiveness of network positioning in p2p systems. In *Proc. of IEEE INFOCOM Mini Conference* (2010).

[6] CUEVAS, R., LAOUTARIS, N., YANG, X., SIGANOS, G., AND RODRIGUEZ, P. Deep diving into bittorrent locality. http://arxiv.org/abs/0907.3874.

[7] DABEK, COX, KAASHOEK, AND MORRIS, R. Vivaldi: A decentralized network coordinate system. In *Proc. of ACM SIGCOMM* (2004).

[8] EMULAB. Emulab. http://www.emulab.net/.

[9] GUMMADI, K. P., SAROIU, S., AND GRIBBLE, S. D. King: Estimating latency between arbitrary Internet end hosts. In *Proc. ACM IMW* (Nov. 2002).

[10] ISDAL, T., PIATEK, M., KRISHNAMURTHY, A., AND ANDERSON, T. Leveraging BitTorrent for end host measurements. In *Proc. of PAM* (2007).

[11] LEDLIE, J., GARDNER, P., AND SELTZER, M. Network coordinates in the wild. In *Proc. of USENIX NSDI* (Apr. 2007).

[12] LUMEZANU, C., BADEN, R., SPRING, N., AND BHATTACHARJEE, B. Triangle inequality variations in the internet. In *Proc. of IMC* (2009).

[13] NG, T., AND ZHANG, H. Predicting Internet network distace with coordinates-based approaches. In *Proc. of IEEE INFOCOM* (2002).

[14] PIATEK, M., MADHYASTHA, H. V., JOHN, J. P., KRISHNAMURTHY, A., AND ANDERSON, T. Pitfalls for ISP-friendly P2P design. In *Proc. of HotNets* (2009).

[15] PLANETLAB. Planetlab. http://www.planet-lab.org/.

[16] PRASAD, R. S., MURRAY, M., DOVROLIS, C., AND CLAFFY, K. Bandwidth estimation: Metrics, measurement techniques, and tools. *IEEE Network 17* (2003), 27–35.

[17] RABINOVICH, M., TRIUKOSE, S., WEN, Z., AND WANG, L. Dipzoom: The internet measurements marketplace. In *Proc. of IEEE INFOCOM* (Apr. 2006).

[18] RITACCO, A., WILLS, C., AND CLAYPOOL, M. How's my network? a java approach to home network measurement. In *Proc. ICCCN* (2009).

[19] ROUGHAN, M., TUKE, S. J., AND MAENNEL, O. Bigfoot, sasquatch, the yeti and other missing links: what we don't know about the as graph. In *Proc. of IMC* (2008).

[20] ROUTEVIEWS. http://www.routeviews.org/.

[21] SU, A.-J., CHOFFNES, D., BUSTAMANTE, F. E., AND KUZMANOVIC, A. Relative network positioning via CDN redirections. In *Proc. of ICDCS* (2008).

[22] TANG, L., AND CROVELLA, M. Virtual landmarks for the internet. In *Proc. of IMC* (2003).

[23] TEAM CYMRU. The Team Cymru IP to ASN lookup page. http://www.cymru.com/BGP/asnlookup.html.

[24] WANG, G., ZHANG, B., AND NG, T. S. E. Towards network triangle inequality violation aware distributed systems. In *Proc. of IMC* (2007).

[25] WONG, B., SLIVKINS, A., AND SIRER, E. Meridian: A lightweight network location service without virtual coordinates. In *Proc. of ACM SIGCOMM* (2005).

[26] XIE, H., YANG, R., KRISHNAMURTHY, A., LIU, Y., AND SILBERSCHATZ, A. P4P: Provider portal for (P2P) applications. In *Proc. of ACM SIGCOMM* (2008).