An Empirical Evaluation of Wide-Area Internet Bottlenecks

Aditya Akella Srinivasan Seshan
Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213
aditya@cs.cmu.edu, srini+@cs.cmu.edu

Anees Shaikh
Network Services and Software
IBM T. J. Watson Research Center
Hawthorne, NY 10532
aashaikh@watson.ibm.com

ABSTRACT

Performance limitations in the current Internet are thought to lie at the edges of the network – *i.e* last mile connectivity to users, or access links of stub ASes. As these links are upgraded, however, it is important to consider where new bottlenecks and hotspots are likely to arise. Through an extensive measurement study, we discover, classify and characterize non-access bottleneck links in terms of their location, latency and available capacity. We find that nearly half of the paths explored have a non-access bottleneck with available capacity less than 50 Mbps. The bottlenecks identified are roughly equally split between intra-ISP links and links between ISPs. Also, we find that low-latency links, both intra-ISP and peering, have a significant likelihood of constraining available bandwidth. These results have implications on issues such as the choice of access providers and route optimization.

Categories and Subject Descriptors

C.2 [**Computer Systems Organization**]: Computer-Communication Networks; C.2.5 [**Computer-Communication Networks**]: Local and Wide-Area Networks

General Terms

Measurement, Performance

1. INTRODUCTION

A common belief about the Internet is that poor network performance arises primarily from constraints at the edges of the network. As access technology evolves, enterprises and end-users, given enough resources, can increase the capacity of their Internet connections by upgrading their access links. The positive impact on overall performance may be insignificant, however, if other parts of the network subsequently become new performance bottlenecks. In this study, we consider the likely location and characteristics of future bottleneck links in the Internet. Such information could prove very useful in the context of choosing intermediate hops in overlay routing services or inter-domain traffic engineering, and also to customers considering their connectivity options.

Our aim is to study the characteristics of links within or between carrier ISPs that could *potentially* constrain the bandwidth available to long-lived TCP flows, called *non-access* bottleneck links. Using a large set of network measurements, we discover and classify such links according to their location in the Internet hierarchy and their estimated available capacity. We make two key contributions: 1) a methodology for measuring bottlenecks links (Section 2) and 2) a classification of non-access bottleneck links in terms of their location, available bandwidth and latency (Section 3).

This research was sponsored by DARPA under contract F30602-99-1-0518. Additional support was provided by IBM. Views and conclusions contained in this document are those of the authors and should not be interpreted as respresenting the official policies, either expressed or implied, of DARPA, IBM or the U.S. government.

SIGMETRICS'03, June 10–14, 2003, San Diego, California, USA.
ACM 1-58113-66401/03/0006.

We find that nearly half of the paths measured have a non-access bottleneck link with available capacity less than 50 Mbps. Moreover, the percentage of observed paths with bottlenecks grows as we consider paths to lower-tier destinations. Surprisingly, the bottlenecks identified are roughly equally split between intra-ISP links and peering links between ISPs. Also, we find that low-latency links, both within and between ISPs have a significant probability of constraining available bandwidth. Finally, of the bottlenecked paths through public exchanges, the constrained link appeared at the exchange point itself in nearly half the cases.

Our observations provide key insights into the location and nature of performance bottlenecks in the Internet, and in some cases, address common impressions about constraints in the network. We hope that our work could prove instrumental in improving the performance of future network protocols and services in terms of which bottlenecks to avoid (and how to avoid them).

2. MEASUREMENT METHODOLOGY

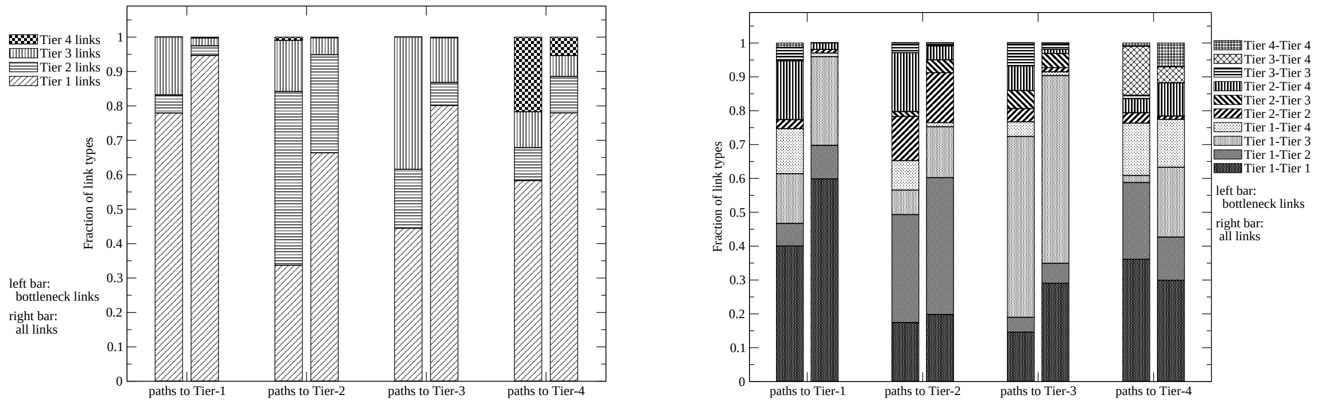
Our methodology addresses four key issues:

Choosing sources: Since we aim to characterize the bottlenecks faced by well-connected end-points, we choose sources (PlanetLab [2] nodes) that they have no bottlenecks in their own access networks, are geographically dispersed, and do not introduce biases due to connectivity to a few upstream ASes.

Choosing destinations: The network paths we measure must be representative of typical Internet paths. Therefore, our destinations consist of routers within various ISPs belonging to each of the four tiers of the Internet hierarchy [3]. We also include paths through public exchange points, which are commonly considered significant bandwidth bottlenecks, by picking destination routers within small tier-4 ISPs attached to popular public exchanges like MAE.

Measurement tools: To identify bottlenecks and report the available bandwidth and latency we developed a tool, *BFind*, that uses techniques motivated by TCP's bandwidth probing behavior and operates in a single-ended mode without requiring superuser access (unlike most bandwidth measurement tools). BFind initiates a variable rate UDP stream to the destination and regularly monitors the delay on the path. If the delay on any link on the path during a given monitoring interval is close to the base link latency, BFind increases the sending rate. Otherwise, the rate is kept steady over successive intervals as long as the link delay remains high. If the delay remain high for reasobably many successive intervals, BFind identifies the corresponding link as the bottleneck, outputs the current send rate (available bandwidth) and the latency of the link. If no such link is found within 180 seconds BFind quits without discovering a bottleneck. BFind limits its rate such that bottlenecks with an available capacity > 50 Mbps cannot be discovered.

Classification metrics: For the bottlenecks links discovered by BFind we identify if the link was within an ISP or between ISPs.



(a) intra-ISP links

(b) Peering links

Figure 1: Graphs (a) and (b) show the average composition of different paths in terms of intra-ISP and peering links, respectively.

We further classify the links according to the tier of the ISP(s), and comment on the observed available bandwidth on different types of bottlneck links.

3. RESULTS

Location of bottlenecks: Nearly half the paths we measured had a bandwidth bottleneck. Figure 1 summarizes the characteristics of each type of path, i.e., to destinations in tiers 1–4. Figure 1(a) shows the breakdown of intra-ISP links observed, averaged across all paths observed for type of path. For each path type, the left bar shows the average fraction of various types of intra-ISP *bottleneck* links in the path, and the right bar shows the *overall* average composition of the paths. Figure 1(b) similarly illustrates the breakdown of peering links appearing as bottlenecks, and overall in the paths. For example, in Figure 1(b), the graph shows that on paths to tier-1 destinations, tier-1 to tier-1 peering links are the bottlenecks in about 40% of the cases on average, and these same links make up about 60% of the average path to tier-1 destinations.

Our main observations are as follows: (1) Lower-tier intra-ISP links seem to be bottlenecks more frequently than we would expect based on the appearance of these links in the paths. For example, tier-2 links make up 50% of the bottlenecks we found to tier-2 destinations, but account for only about 29% of the links in these paths. (2) Peering links between tier-1 ISPs are bottlenecks much less frequently than their overall prevalence in the paths indicates. (3) Peering links to or from tier-2, tier-3 or tier-4 ISPs are bottlenecks more frequently than expected. For example, compare the proportion of tier-2 to tier-4 peering bottlenecks with the proportion of these links in the corresponding overall path length (e.g., 17% vs. 2% for paths to tier-1, and 17% vs. 4% for paths to tier-2). (5) The bottlenecks are equally split between inter- and intra-ISP links despite a smaller overall number of inter-ISP links. Thus, if there is a bottleneck link on a path, it is equally likely to be either in the interior of an ISP or between ISPs. But, as there are fewer peering links, each peering link has a higher likelihood of being a bottleneck. (6) The fraction of paths with bottlenecks grows as we consider paths to lower-tier destinations. About 32.5% of the paths to tier-1 destinations have bottlenecks. For paths to tiers 2, 3, and 4, the percentages are 36%, 50%, and 54%, respectively.

Available Bandwidth of Bottleneck Links: Our analysis of the available bandwidth of the bottleneck links identified by BFind shows the following: (1) Tier-1 and tier-3 ISPs have a distinct advantage in terms of bottleneck bandwidth over tier-2 ISPs. (2) Links in tier-4 ISPs exhibit the most limited available bandwidth distribution as expected. (3) For peering bottlenecks, Tier-1 to tier-1 peering links are the least constrained. (4) Tier-2 and tier-3 links exhibit very similar bandwidth characteristics in their peering links to tier-1 networks. (4) Peering links between tier-2 and tier-3 are not significantly different than tier-2 to tier-2 links. (5) Bottleneck peering links involving networks low in the hierarchy (e.g., tier-4) provide significantly less available capacity, as expected.

Bottlenecks at Exchange Points: Of the 466 measured paths through exchange points, 170 (36.5%) had a bottleneck link and about 15% (*i.e.* 70 paths) had bottlenecks at the exchange point. This is counter to the expectation that many exchange point bottlenecks would be identified on such paths. Notice that the probability that the bottleneck is at the exchange is about 41% ($= 70/170$) implying that if there is a bottleneck on a path through a public exchange point, it is very likely to be at the exchange itself.

A more detailed description of these results, along with an analysis of bottleneck link latency for different types of links, is available in [1]

4. DISCUSSION

Although our findings generally support conventional wisdom about Internet bottlenecks, there are a few unexpected observations.

There is a clear performance advantage to using a tier-1 provider. Small regional providers, e.g., tier-4 ASes in our study, provide the worst performance. At the same time, the tier-2 and tier-3 carriers perform similarly, implying that if a stub network desires reasonably wide connectivity, then choosing a tier-3 provider might be a beneficial choice, both economically and in terms of performance.

Since $> 50\%$ of the paths we probed seem to have an available capacity > 40 Mbps, we hypothesize that large portions of the network are potentially under-utilized. This confirms what many large ISPs report about the utilization of their networks. But the fact that this holds even for providers of smaller size (e.g., tier-3) as well as for most peering links and even links at NAPs, seems surprising. Also since we did not see as many peering bottlenecks as conventional wisdom suggests, this may imply that either the peering links are in fact quite well provisioned, or that a smaller portion of the entire Internet traffic traverses these links than expected.

Our results imply that buying bandwidth from two different tier-1 ISPs (e.g., to reduce peering-point crossings) may not be much better from a performance perspective than buying twice as much bandwidth from a single tier-1 ISP. It may also be more economical to buy from one ISP. Also, a shorter route to a destination that passed through a tier-1-tier-1 peering link might be better than a longer route within a single lower-tier provider.

5. REFERENCES

[1] A. Akella, S. Seshan, and A. Shaikh. An empirical evaluation of wide-area Internet bottlenecks. Technical report, IBM TJ Watson Research Center, 2003.

[2] PlanetLab. `http://www.planet-lab.org`, 2002.

[3] L. Subramanian, S. Agarwal, J. Rexford, and R. H. Katz. Characterizing the Internet hierarchy from multiple vantage points. In *Proceedings of IEEE INFOCOM*, June 2002.