A Longitudinal Study of Utilization at Internet Interconnection Points

Nick Feamster
Princeton University
http://interconnection.citp.princeton.edu/

Abstract

The increase in high-volume traffic flows due to applications such as video streaming draw new attention on utilization at the interconnections between the Internet's independently operated networks. This paper surveys the findings from nearly two years of Internet utilization data provided by seven participating ISPs—Bright House Networks, Comcast, Cox, Mediacom, Midco, Suddenlink, and Time Warner Cable—whose access networks represent about 50% of all U.S. broadband subscribers. The dataset spans 18 months and includes about 97% of the paid peering, settlement-free peering, and ISP-paid transit links of each of the participating ISPs. Analysis of the data—which comprises more than 1,000 link groups, representing the diverse and substitutable available routes—suggests that many interconnects have significant spare capacity, that this spare capacity exists both across ISPs in each region and in aggregate for any individual ISP, and that the aggregate utilization across interconnects is roughly 50% during peak periods.

1 Introduction

As traffic demands increase due to the rise of large asymmetric traffic flows such as video streaming, interconnection arrangements must evolve to meet these new demands. The nature, causes, and location of Internet congestion has spawned contentious debate over the past two years. End users have become increasingly invested in this topic as well, although they have sometimes conflated the issues of Internet congestion with other concerns about the prioritization of Internet traffic. Discussion about interconnection can and should be better informed by accurate, up-to-date information about where capacity bottlenecks exist. Prior to this study, data about traffic utilization at Internet interconnection points has been elusive, due to confidentiality and business constraints. This opacity has led users, policymakers, and researchers to rely on end-to-end probes to isolate congestion [2, 5, 12]; these techniques are often inaccurate.

As David Clark said, "An issue that has come up recently is whether interconnection links are congested. The parties who connect certainly know what's going on, but that data is generally not disclosed. The state of those links matters to a lot of people ... and there have been some misunderstandings around congestion and interconnection links" [4]. To help shed light on this important issue, a consortium of ISPs have shared data concerning the utilization of of interconnection links. This data yields some information concerning the utilization of network ports that face each network's "peers" (i.e., the networks that each ISP connects to directly). This data illuminates the utilization of each networks externally facing switch ports. Although this data cannot by itself tell the complete story about the location of congestion along end-to-end Internet paths, it is an important piece of the overall picture.

Each participating ISP provided information about its interconnection to neighboring networks (e.g., ISPs, content providers) in each region, as well as the capacity of each interconnect. The data that participating ISPs provide account for about 97% of links from all participating ISPs in any given month; the only links that are missing from the dataset are those where the measurement infrastructure has not yet been deployed. This information offers sufficient information to ascertain the capacity of each interconnect between an ISP and neighboring networks. Given this information, we can compare this provisioned capacity against traffic statistics for traffic that traverses each of these network ports and compare the measured utilization to the provisioned capacity to achieve an estimate of utilization.

Ideally, the information would reveal the utilization and capacity for each port, for every ISP—in such a scenario, comparing utilization to provisioned capacity would be straightforward. Business realities introduce complications: even the existence of an individual interconnection is sometimes considered proprietary, not to mention the business agreement surrounding that interconnection, as well as the capacity and utilization of the interconnection. The ISPs release traffic flow statistics that are aggregated across link groups in each region; these aggregates a high-level picture of capacity and utilization per region and ISP, as well as how this utilization varies over time.

The data has some limitations that make it inappropriate for answering certain questions about utilization. First, it is sampled, which makes it difficult to answer certain types of questions about flow size distributions, characteristics of small flows, and utilization by application. Second, to preserve proprietary information, the data is aggregated and anonymized, preventing conclusions about utilization at specific interconnection points. Throughout this paper, we are careful to highlight conclusions that we can and cannot make with the data that the ISPs have provided. Based on feedback from other experts, we have also iterated on the data that the ISPs have agreed to release, resulting in a careful balancing act between preserving proprietary information and revealing information about utilization at interconnection points that can inform ongoing debates.

This study reflects our current understanding of capacity and utilization at interconnection points, and how it has evolved over time, from October 2015 to the present (this paper reports findings through March 2017). The dialog surrounding interconnection is ongoing, and, accordingly, we are continuing to collect and analyze data. As a resource to interested parties—and to promote further academic research in this field, we have periodically updated the findings and data from this project on the project website [8] since its launch in March 2016.

The rest of this paper is organized as follows. Section 2 describes related work and analysis techniques. Section 3 describes the measurement techniques and data, as well as the effects of various phenomena such as sampling on the accuracy of the collected data. Section 4 describes the findings from a preliminary analysis of the data collected as part of the project. Section 5 concludes with a summary and suggestions for possible next steps.

2 Related Work

In this section, we briefly outline other attempts to measure both end-to-end performance of Internet paths and infer congestion along these paths (and at interconnects) using these datasets. All of these techniques and approaches involve inference based on measurements from end hosts, as opposed to direct measurements of utilization at the interconnect. They also do not report longitudinal measurements, as we do in this study.

2.1 Measurements from Network Endpoints

Measurements from end-hosts. The network diagnostic tool (NDT) [2] performs throughput tests. One version of the tool runs as a Java applet from a web browser. Measurement Lab runs a version of the Java applet from its website that measures throughput to the collection of deployed Measurement Lab servers around the world [12], using geolocation to map the client to a nearby NDT server for the purposes of the throughput test (the accuracy of a TCP throughput test depends on measuring throughput to a nearby server, since TCP throughput is inversely proportional to round-trip latency). NDT also forms the basis of well-known measurement efforts, such as the Internet Health Test. Unfortunately, MLab's NDT test setup is known to be inaccurate due to its use of only a single thread to measure TCP throughput, which our previous work shows can significantly underestimate the throughput of the link [14]. Additionally, NDT provides no mechanism for locating a throughput bottleneck along an end-to-end path.

Measurements from routers. The Broadband Internet Service Benchmark (BISmark) [1, 14] project runs custom throughput, latency, and packet loss measurements from home routers that run OpenWrt. The project has been collecting performance data from home networks since 2011; at its peak, the project was collecting data from about 400 home networks in more than 30 countries, and about 70 home routers are actively deployed. The FCCs Measuring Broadband America project [11] produces periodic reports using similar measurements from a larger deployment. Their reports are less frequent (typically once per year), as opposed to BISmark's "real time" visualizations of throughput, latency, and packet loss. In both cases, the performance measurements end-to-end.

2.2 Measuring Interconnect Performance

Because the above tools can only measure from end-host vantage points, they do not provide direct information about utilization or congestion at interconnection points. Because congestion manifests as an increase in latency, the measurement techniques that we have discussed above can often detect congestion along an Internet path. Yet, detecting congestion at a particular interconnection point is difficult to do with these types of measurements. We discuss various other methods to indirectly infer or directly measure congestion at an interconnection point in this section.

2.2.1 Indirect: Tomography & Round-Trip Latency

A general measurement approach sometimes referred to as *network tomography* attempts to use a combination of performance measurements along different end-to-end Internet paths to infer specific links where congestion or failures may be occurring [5]. The intuition is quite simple: Given near-simultaneous measurements of two end-to-end Internet paths that may share one or more links, if one end-to-end path experiences symptoms of congestion (i.e., an increase in latency) whereas a second end-to-end path does not, we can infer that the congestion must be occurring on the portion of the second path that is not common with the first path. Unfortunately, it is difficult to obtain a comprehensive set of vantage points in practice because most end-to-end paths will share more than one interconnection point or link in common. For example, in an M-Lab report released in 2014 [13], many of the end-to-end paths between NDT vantage points and the M-Lab servers could (and likely do) share multiple end-to-end links along the path—not only the interconnection point (where the report implies congestion is taking place) but also other links along the path (e.g., links within transit providers). The second scenario is a distinct possibility that previous reports have outlined in detail [6]. Additionally, providers can (and have) gamed these active measurement techniques by prioritizing probe traffic [9].

Another approach is to use traceroutes to discover an end-to-end path and subsequently send latency probes to either side of an interconnect. Although this approach is more direct than network tomography, the approach entails significant shortcomings, which are outlined in detail in previous work [10]. Among the limitations are the difficulty in accurately identifying interconnection points along and end-to-end traceroute, as well as the fact that increases in latency might be occurring along reverse paths, as opposed to the forward path that the probes are attempting to measure.

2.2.2 Direct Measurement: Packet Capture and IPFIX

Packet traces capture a record of the packets that traverses a particular interface. A capture may include the complete packets, the first bytes of each packet, or simply the “headers” or metadata for each packet. Packet capture provides complete timing information about the arrival of individual packets and the header information on individual packets (including the TCP window size), and can as such be used to compute or infer properties of traffic flows including jitter, packet loss, and instantaneous throughput. These methods are typically not practical at large, high-throughput interconnection points; they tend to be costly to deploy, and they produce more data than can be reasonably backhauled to a data-center for post hoc analysis. This project is the first to report interconnect utilization based on IPFIX measurements. An earlier version of this work was presented at the *Technology Policy Research Conference* [7], which does not publish archival proceedings. This paper presents an update to that technical report, augmented with an additional year of longitudinal data.

3 Measurement and Data

We now describe the data that each ISP provides concerning the utilization of each network port, and how this data is sampled and aggregated.

3.1 Traffic Flow Statistics and Utilization

The most common version of flow statistics is the IPFIX protocol (Cisco’s version is “NetFlow”) [3]. A IPFIX record contains metadata about the flow, including the number of bytes transferred, the number of packets in the flow, the start and end times for the flow, and the network interface associated with the flow. Accordingly, the statistics in a flow record can give useful information about the average utilization over a period of time in terms of either bytes or packets. For example, if the flow record has a duration of ten seconds and reports that 1 gigabyte of traffic was transferred during that ten seconds, then the average utilization over that ten-second period would be 800 megabits per second.

IPFIX records are often based on a sample of the packets in any particular flow. Most of the ISPs report traffic flow statistics based on a sampling rate of 1/1,000, meaning that statistics are based on a sampling of every thousandth packet, on average; all of the ISPs who are contributing data implement a sampling rate of at least 1/8,000. Some of the ISPs in the study use deterministic sampling and others use random sampling; given that the goal is to estimate capacity on links where much of the traffic flows that contribute to congestion are large, long-running video streams—which have fairly large packet and byte counts—neither the sampling rate nor the mode of sampling should affect the accuracy in estimating utilization. We empirically verified that sampled IPFIX records are sufficient for aggregate capacity utilization. We compared the SNMP byte counters to sampled IPFIX records

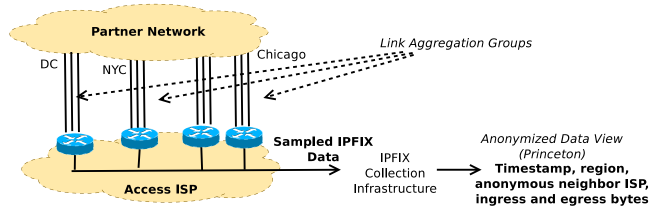


Figure 1: *Data collection infrastructure and approach.*

with a 1/8,000 sampling rate across 250 interconnect links for one of the largest participating ISPs in the study for a single day. The mean and median of the ratio between both metrics were both around 0.98, with a standard deviation of 0.095.

The traffic in this dataset covers interconnection points for access ISPs that account for about 50% of the broadband subscribers in the United States. Figure 1 shows how data is collected from each interconnection point between an access ISP and neighboring partner network. Each participating access ISP may connect to a partner network in multiple geographic regions. The access ISP collects IPFIX data at each interface that interconnects with a neighboring partner network. The traffic statistics that each ISP reports are based on IPFIX records that are exported at least as frequently as every 60 seconds and subsequently aggregated across a link group; to protect the confidentiality of information pertaining to usage on specific interconnects, the data is aggregated into a single link group per geographic region. (Section 3.2 describes this approach in more detail, and how it affects the conclusions we can draw.) The statistics represent an aggregate that is computed based on the sum of peak five-minute intervals in each hour, for each {neighbor network, circuit group} pair.

The dataset contains about 97% of links from all participating ISPs in any given month; the only links that are missing from the data set are links that are not configured in the measurement system. In this paper, we analyze data over a timespan of 18 months, from October 2015 through March 2017. All interconnections between participating ISPs and neighboring partner networks are private (i.e., none of the interconnections in this study involve public IXP switch fabrics). Each row in the dataset that we receive includes timestamp (representing a five-minute interval), region (representing an aggregated link group), anonymized partner network access ISP, total ingress bytes, total egress bytes, and total link capacity. Because flows do not begin and end on discrete five-minute intervals, each five-minute timestamp represents the sum of utilization of active traffic flows that were active during that interval. Suppose that, at a given time, a set of flows are active. Then, the total ingress bytes for that five-minute interval for a single flow would be the average bitrate for that flow over its total duration, multiplied by the amount of time that the flow was active during the given five minute interval. The total utilization for the link aggregation group is the sum of all such statistics, for any flows that were active during that five-minute interval.

3.2 Aggregation and Load Balancing

When measuring the contribution of a traffic flow to a link's utilization, it is also important to ensure that flows are not double counted. An ISP's ports may be configured as a link aggregation group (we are aware of this configuration for at least one ISP in the study). In this ISP's case, the router balances outbound traffic flows across the links; a single flow always goes across a single link. The allocation of outbound traffic flows to links is based on a hashing algorithm on the router; given enough traffic flows, this type of load balancing typically works well enough to balance load evenly across the available links in any given aggregation group. It is extremely rare for any ISP to have multiple LAGs in a region to a given partner network.

Although we do not know the load balancing practices of neighbor networks for traffic that is inbound to an ISP, we assume that these networks also use typical load balancing practices for their outbound traffic (and, hence, we can assume a relatively uniform load balance of inbound traffic flows for a link aggregation group). In networks where there exist only a small number of flows, such as in commercial VPNs, it is possible that utilization across links might become unbalanced, but on links carrying consumer traffic, such as those in this study, it would be highly unusual for traffic to be unevenly balanced across links, due to the nature of the router hashing function, which is designed to randomly assign these flows to available links.

3.3 Limitations

In the private dataset, statistics aggregated across link groups; thus, any fluctuations that occur on only a single link may not be reflected in the aggregate statistics. drawing conclusions based on average utilization per link within a link group is reasonable. Additionally, short-term periods of high utilization across an entire link group may not be evident in the data, because utilization is reported on five-minute averages.

Due to aggregation requirements in the public dataset, it is possible to assess the overall utilization in some region across all ISPs and partner networks, but not for any individual interconnection point in a region. Similarly, it is possible to see the aggregate utilization for any of the participating ISPs, but not for a specific region or neighboring AS. As a result, the aggregates make it difficult to drill down into the utilization between any pair of networks, either as a whole or for any particular region. As a result, it is not possible to conclude that no interconnection links experience high utilization. Because the public data shows utilization across each ISP, we can conclude that each ISP has spare capacity—although we cannot conclude that it has spare capacity in each region or on any individual port.

The existence and nature of interconnection is considered proprietary, as are the decisions about where any particular ISP has a point of presence and where any ISP tends to route different types of traffic. These details reflect both business strategy (e.g., provisioning), business relationships, the source and destination of traffic demands, and decisions

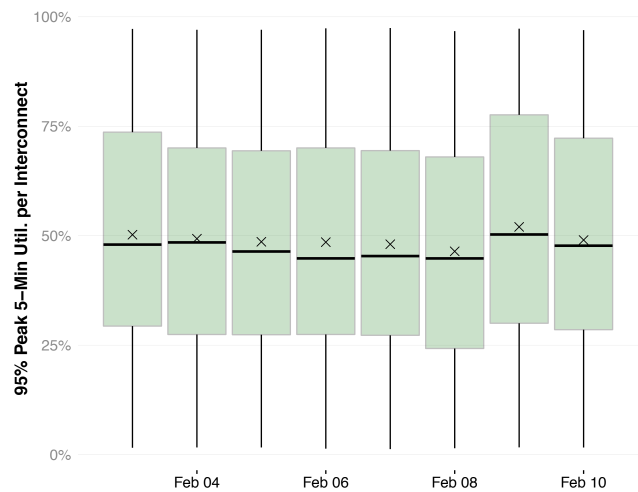


Figure 2: *Utilization of each interconnect group over one week in February 2016, normalized by capacity of the interconnects.*

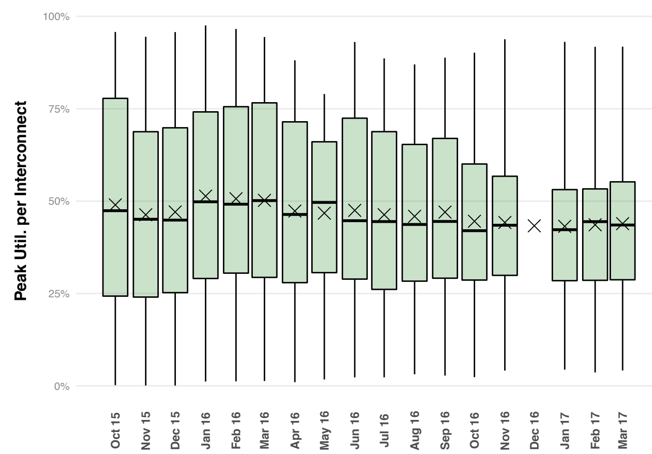


Figure 3: *Per-month utilization of all participating interconnects.*

about network management and operations. The restrictions on our ability to disclose data to the public result not from a specific agreement with the ISPs but rather from the *mutual non-disclosure agreements between the ISPs and their content providers*, which are intended to protect both parties.

4 Utilization at Interconnection Points

This section presents analysis of the utilization measurements from the interconnect groups from the participating ISPs, both overall and by region. From October 2015 through March 2017, aggregate interconnect capacity has been roughly 50% utilized at peak, and capacity has grown consistently by about 3% each month on average over the period of our study.

4.1 Aggregate Utilization

Figure 2 shows the interconnect utilization over time, for a one-week period in February 2016 across all regions. Each

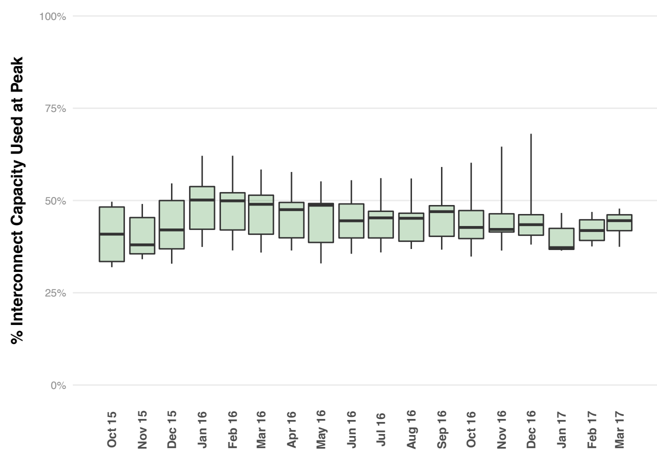


Figure 4: *Distribution of 95th percentile peak ingress utilization across all ISPs, with all ISPs equally weighted.*

data point in the timeseries shows a box plot illustrating the distribution of utilization across interconnect points for all five-minute intervals. The median utilization across interconnects is consistently below 50%, even at peak times, and many of the links have significantly less utilization. Less than 4% of the link aggregation groups exceed 95% utilization in any five-minute interval, and the vast majority of the link aggregation groups see much less utilization, even at peak times. In the next section, we explore these trends for individual regions.

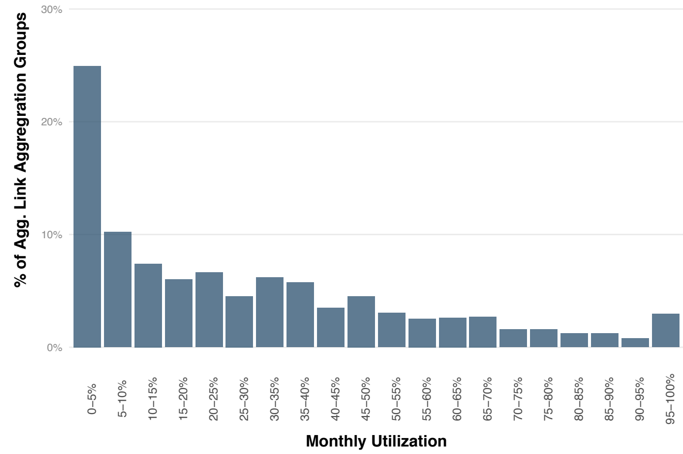
Figure 3 shows the distribution of interconnect capacity over all five-minute intervals across link aggregation groups for each month, for all aggregation groups. The box plot shows the inter-quartile ranges, the horizontal line shows the median utilization, and the whiskers show the 5th and 95th percentiles. Due to the aggregation that we must perform on the data, we cannot expose whether there are no links in the aggregation group running at full utilization. It is evident, however, is that there *exist* links in the aggregation group with sufficient spare capacity.

4.2 Utilization by ISPs and Links

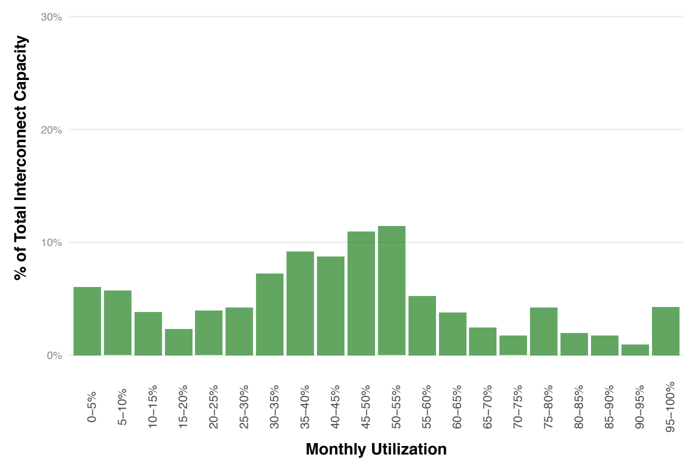
Figure 4 shows the distribution of 95th percentile peak ingress utilization across all ISPs, normalized by capacity. The median ISP in the group of seven ISPs experienced a 95th percentile peak ingress utilization that was less typically around 50% of the available capacity. This plot shows that each ISP has significant spare capacity across its set of links and regions. This figure does *not* indicate whether a particular ISP is experiencing congestion in a particular region, to a particular partner network, or across a set of links.

Unfortunately, we cannot show utilization for specific links or neighbor networks, because the existence of a particular business relationship or even the existence of a specific link in a region may reveal proprietary information. We can, however, explore the utilization across the aggregate of all links, which also shows the existence spare capacity. Specifically, we can show how the characterization of peak utilization across *all* links, weighted both by links and by overall capacity, as shown in Figure 5. Figure 5a shows the distribution of 95th percentile peak monthly utilization across all links, for all participating ISPs. This figure shows that more than 25% of all links are significantly underutilized, and that less than 10% of all links experience a 95th percentile peak utilization that exceeds 90%.

In Figure 5a all links are weighted equally, which does not reveal whether there exists significant excess *capacity*, only whether there exist links that have spare capacity. Exploring utilization where the set of links is weighted by their capacity



(a) *Weighted by links.*



(b) *Weighted by capacity.*

Figure 5: *The fraction of interconnect capacity, weighted by the number of links and the amount of total capacity, respectively, whose 95th percentile utilization in a month experienced a particular utilization level. The figure shows statistics for March 2017.*

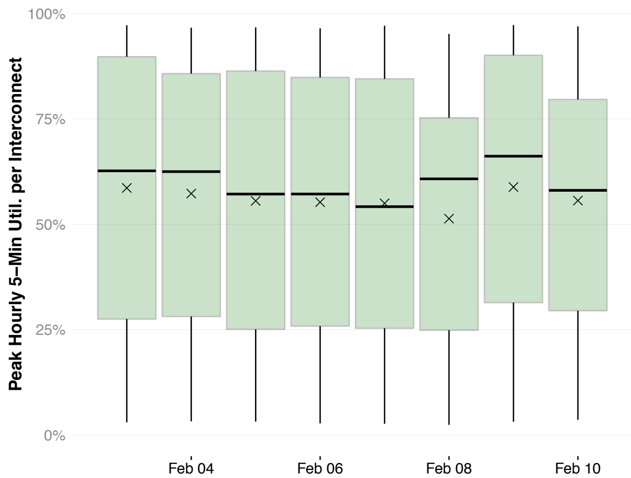
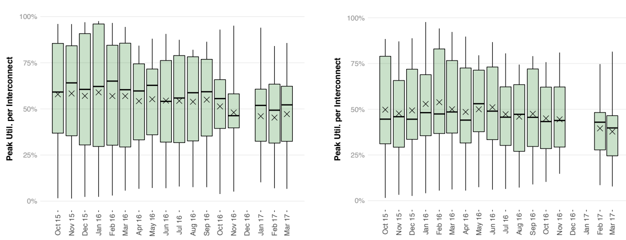


Figure 6: *Utilization of each interconnect group over one week in February 2016 across interconnects in Chicago, IL, normalized by capacity of the interconnects.*



(a) *Chicago.* **(b)** *Dallas.*

Figure 7: *Per-month utilization of participating interconnects in two example regions.*

reveals more information. Figure 5b shows the same distribution, where links are weighted by overall capacity. The figure shows that links that account for about 10% of overall interconnect capacity experienced a 95th percentile peak utilization that exceeded 95%. Most of the capacity experienced significantly less utilization.

Together, these plots present a picture of the existence of spare utilization across many of the interconnects that also account for much of the capacity at interconnects. Certain answers remain obscured, such as whether a particular partner network is experiencing persistent congestion, or whether particular types of connections (e.g., paid peering) are experiencing more or less congestion. Yet, the figures above do reveal a general picture of (1) all ISPs having spare capacity in aggregate across interconnects; (2) most interconnect capacity in aggregate showing spare capacity at peak. Both of these conclusions reveal significantly more than we have known to date; as this project matures and we receive further feedback, we hope to make additional views of the data available that also respect the private and proprietary information of each ISP.

4.3 Utilization by Region

We also explored how utilization evolves over time in individual regions, to determine whether utilization patterns at interconnects in specific regions agreed with the overall general trends that we observed in Figure 2. Figure 6 shows how utilization evolves over time across interconnects in Chicago; the trends in this specific region are similar to the overall trends. The trends are similar in other cities with busy interconnects; interconnects in Atlanta show similar distributions.

Washington, New York, Dallas, and Los Angeles exhibit similar trends, although utilization exceeded 90% less frequently than in Chicago and Atlanta, the two busiest regions. Figure 7 shows the distribution of interconnect capacity across link aggregation groups over all five-minute intervals in each month, for from late 2015 through early 2016, for two example regions. Figure 7a shows this distribution for a busier interconnect region (Chicago); Figure 7b shows the same distribution for Dallas. Interconnections in Dallas tend to have lower median utilizations across link groups, although the highest loaded link groups at peak time also follow similar trends as those that we observed in Chicago.

5 Conclusion

Public discourse about interconnection requires better visibility into utilization at interconnection points between ISPs and content providers. Unfortunately, existing measurement methods from the edge using active probes are inconclusive—they cannot accurately pinpoint congestion at interconnection, and they cannot disambiguate congestion that occurs on a forward path from congestion that occurs on a reverse path. Stronger conclusions require more direct measurements of utilization at the interconnection points.

This paper presents an initial step in that direction, based on data from interconnection points from seven major Internet service providers. Our analysis suggests that capacity continues to be provisioned to meet growing demand and certain interconnection points have spare capacity even though specific links may be experiencing high utilization.

Perhaps one of the most important unanswered questions about interconnect utilization concerns how metrics such as interconnect utilization affect user experience for different applications. Although answering will ultimately require more holistic data sources than simply interconnection utilization. Yet, information about the utilization of interconnection can shed light on an important aspect of this problem. For example, an important open question concerns how the utilization of a link ultimately affects a customer's quality of experience for a given application, such as video streaming. Conventional wisdom says that once utilization exceeds 70%, an ISP should provision additional capacity. Yet, if certain links could be operated at higher utilizations without adversely affecting customer experience, ISPs could save significant operational costs. Our ongoing work includes assessing the correlation between these network-level traffic statistics and the corresponding quality of experience for different types of applications.

References

[1] BISMark: Broadband Internet Service Benchmark. http://projectbismark.net/.

[2] R. Carlson. Network Diagnostic Tool. http://e2epi.internet2.edu/ndt/.

[3] B. Claise, B. Trammell, and P. Aitken. *Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information*. Internet Engineering Task Force, Sept. 2013. https://datatracker.ietf.org/doc/rfc7011/.

[4] D. Clark. Keynote: Internet measurement. https://youtu.be/2GE8qShOoJ4?t=6m53s, Feb. 2016. Comments on interconnection congestion start at 6:53.

[5] M. J. Coates and R. D. Nowak. Network tomography for internal delay estimation. In *Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP'01). 2001 IEEE International Conference on*, volume 6, pages 3409–3412. IEEE, 2001.

[6] N. Feamster. Where is Internet Congestion Occurring?, Apr. 2015. https://freedom-to-tinker.com/blog/feamster/where-is-internet-congestion-occurring/.

[7] N. Feamster. Revealing Utilization at Internet Interconnection Points. In *Technology Policy Research Conference*, Washington, DC, Sept. 2016. Technical Report. https://arxiv.org/abs/1603.03656.

[8] Interconnection Measurement Project. http://interconnection.citp.princeton.edu/.

[9] H. Kilmer. M-Labs data and Cogent DSCP markings, Nov. 2014. https://goo.gl/9m8Kk5.

[10] M. Luckie, A. Dhamdhere, D. Clark, B. Huffaker, et al. Challenges in Inferring Internet Interdomain Congestion. In *Internet Measurement Conference*, pages 15–22. ACM, 2014.

[11] Measuring Broadband America. https://www.fcc.gov/general/measuring-broadband-america. Federal Communications Commission.

[12] Measurement Lab. http://measurementlab.net.

[13] ISP Interconnection and its Impact on Consumer Internet Performance. Technical report, Measurement Lab, Oct. 2014. http://www.measurementlab.net/static/observatory/M-Lab_Interconnection_Study_US.pdf.

[14] S. Sundaresan, W. de Donato, N. Feamster, R. Teixeira, S. Crawford, and A. Pescapè. Broadband internet performance: A view from the gateway. Toronto, Ontario, Canada, Aug. 2011.