Measurement and Analysis of Internet Interconnection and Congestion

David Clark
Steven Bauer
William Lehr
CSAIL/MIT
ddc,bauer,wlehr@mit.edu

kc claffy
Amogh Dhamdhere
Bradley Huffaker
Matthew Luckie
CAIDA
UC, San Diego
kc,amogh,brad,mjl@caida.org

August 15, 2014

Abstract

This paper examines the nature of congestion in the interior of the Internet: its location and its intensity. We present results of a new method for probing the Internet which allows us to detect the presence of congestion on specific links. We are particularly interested in the links that interconnect ISPs, including what are called peering links. There is little data on the extent of congestion across interconnection links, leaving users to wonder if peering links are properly engineered, or whether they are a major source of problems transferring data across the Internet.

Questions about congestion are an important part of the current debates about reasonable network management. Our results suggest that peering connections are properly provisioned in the majority of cases, and the locations of major congestion are associated with specific sources of high-volume flows, specifically from the Content Delivery Networks (CDNs) and dominant content providers such as Netflix and Google. We see variation in the congestion patterns among these different sources that suggests differences in business relationships among them. We argue that a major goal for the industry is to work out methods of cooperation among different ISPs and higher-level services (who may be competitors in the market) so as to jointly achieve cost-effective delivery of data with maximum quality of service and quality of experience. We discuss some of the problems that must be addressed in this coordination exercise, as well as some of the barriers.

1 Introduction

This paper reports early results from an ongoing project to detect and measure congestion in the core of the Internet. Congestion on the Internet, in particular at points of interconnection among ISPs, is an issue of recent popular interest due to the changing nature of Internet traffic over the last decade. In particular, the growth of streaming video and popular content distribution platforms, including social networks, has increased the volume and character of traffic. Much of today's Internet traffic originates from large content providers and their (own or partner) content delivery networks (CDNs). In 2013, Sandvine [20] reported that about half of all peak period downstream consumer traffic came from Netflix or Youtube. As CDNs are deployed, both CDN operators and ISPs must negotiate for suitable interconnection among parts of the Internet to carry this traffic, and the terms of these negotiations have sometimes been contentious, with the apparent consequence that traffic is flowing over links with insufficient capacity, leading to congestion.

The rise of large content providers and their CDNs is a significant shift in the industry structure of the Internet ecosystem, with implications for the balance of power and control among its players. Although disputes over interconnection[1] are not new [9, 18, 19], heated peering disputes between powerful players in the U.S. have increased in the last four years [1, 2, 4–6, 11, 12, 14, 22], raising questions about appropriate network management practices as well as concerns about intentional degradation of performance as a business strategy to obtain interconnection fees.

In this paper we report early results of an ongoing measurement study that attempts to locate and analyze episodes of persistent congestion on the Internet, specifically focused on interconnection links between ISPs, i.e., not links internal to a single ISP's infrastructure. Our motivation includes the press reports cited above and now increasing government interest in the possibility of congestion on interconnection links that is affecting the consumer experience. We report results from our initial development and deployment of the tool during the spring of 2014, a period that coincided with some high-visibility disputes involving Netflix and several large broadband access providers, including Comcast, as a result of which significant episodes of congestion manifested on paths carrying high volumes of video traffic. We saw peering links carrying video traffic congested for 18 hours a day. We also observed such congestion vanish as new interconnection links were put in place and activated. However, we did not find evidence of widespread persistent congestion in interconnection links between the U.S. access ISPs that we probed.

We emphasize the preliminary nature of this work, and its several limitations, outlined in more detail in Section 3 and our technical paper documenting this study [15]. Nonetheless, our data has led to productive conversations with network operators, edge service providers and regulators. Most of the congestion events we identified this year are related to specific circumstances that both parties (on either side of the congested links) have acknowledged, providing some external validation of our methods, trust in our preliminary conclusions, and confidence in our plans to extend the scope of this measurement and data analysis.

Sections 2 clarifies our terminology, and Section 3 explains our measurement methodology

[1]These disputes are often called "peering disputes", since the term to describe many of these interconnections is "peering".

and its limitations. Section 4 illustrates our methods using several case studies. Section 5 discusses why diagnosing congestion is important and challenging; we carefully frame and justify our interpretations of this data to avoid its possible misuse. Section 6 summarizes our conclusions and outlines future work we would like to both pursue and see others undertake.

2 Congestion: its definition and management

In this work, we take an operational view of congestion: Internet congestion occurs when a part of the Internet, e.g., a link or server, does not have the capacity to service all the demand presented to it. In particular, we are looking for persistent congestion that might signal a long-term mismatch between offered load and available capacity.

The duration of a congestion episode is an important factor in evaluating its potential causes and impact on users. Some instantaneous congestion on a given link is an inherent characteristic of Internet traffic. Specifically, the dominant Internet transport protocol (TCP) adjusts its sending rate based on congestion feedback from the network.[2] Occasional congestion episodes of moderate duration and frequency are also inevitable, since traffic demand is dynamic, and links in the core of the network are shared by many users. However, repeated and persistent congestion episodes of significant duration – several hours per day over many weeks or months – is not the result of any fundamental technology limitation; it is an indication that the operators of the congested links have not resolved who should take which steps to alleviate the congestion, either by increasing capacity or reducing traffic load on the congestion link(s).

As we define congestion, the consequences are a queue of packets that form at the input to the congested link, and dropped packets if the queue is full.[3] This operational definition of congestion is not the only possible one. A network operator might conclude that a link is becoming congested when the observed peak load on the link has reached 70 or 80% of capacity, and start planning an upgrade. An economist would not look at technical features of congestion (queuing delay or dropped packets), but would define congestion in terms of a condition that imposed a negative externality on the users of the resource. An extensive discussion of the character of congestion and its different definitions can be found in [3].

Our measurement scheme can detect persistent congestion, but does not permit us to diagnose its cause, determine the extent to which it is harming the user experience, or identify how to best alleviate it. Disputes over interconnection often involve claims that one party is causing the congestion. A more constructive way to understand the management of congestion is to recognize that there are many ways to deal with congestion, and different actors have different means at their disposal. CDNs may have many copies of a given piece of content, and by picking among different copies they can control the source (and thus the path) of their data coming into an ISP, so they can potentially increase (or alleviate) loading and congestion on different points of interconnection. By investing in strategic storage of

[2] A sender doing a bulk data transfer increases its sending rate until it detects evidence of congestion, e.g., a dropped packet or an explicit control message. When it detects one of these signals, it reduces its sending rate, and then starts increasing it again, repeatedly probing for the maximum rate that triggers congestion.

[3] Since a dropped packet is the normal indication of congestion on the Internet, some congestion management schemes (called Active Queue Management (AQM)) may drop a packet before the queue is full.

Table 1: *Current vantage points running TSLP methods*

Host	AS number	Location
bed-us	7922	Comcast, Concord, MA, US
bed2-us	701	Verizon FiOS, Groton, MA, US
bed3-us	6079	RCN, Lexington, MA, US
ith-us	7843	Time Warner, Ithaca, NY, US
lex-us	7843	Time Warner, Lexington, KY, US
lpi-se	8473	Bahnhof, Linkoping, SE
mry-us	7922	Comcast, Carmel, CA, US
msy-us	7843	Cox, New Orleans LA, US
osl2-no	2119	Canal Digital, Oslo, NO
san2-us	22773	Cox, San Diego, CA, US
san4-us	7843	Time Warner, San Diego, CA, US
stn-uk	5400	British Telecom, Ipswich, UK
nce-fr	12322	Free, Nice, France

more copies, they increase the options for source selection. They can also adjust the coding of content. Decisions about the capacity of interconnection links are joint decisions by both parties – neither side can unilaterally upgrade capacity.[4] CDNs and ISPs can also negotiate to position the CDN's content inside the access ISP's network, which reduces traffic crossing the interconnection link. Thus, the best response to congestion is not necessarily to upgrade the congested link. However, given the range of options to manage congestion, strategic behavior by a CDN or ISP is possible, and may create externalities for other infrastructure operators, and inevitable regulatory interest in the resulting shifts in bargaining, money flows, and potential market power related to interconnection.

3 Measurement infrastructure and methodology

Our measurement infrastructure consists of Archipelago [7] measurement probes (vantage points, or VPs) in residential customer networks of the major U.S. broadband access providers, including Comcast, Time Warner, Verizon and Cox. We have also deployed VPs at several ISPs outside the U.S., including BT (UK) and Free (France). Table 1 lists our current measurement vantage points, a subset of the larger Archipelago measurement infrastructure. These probes continually send and receive packets, measure the Round Trip Time (RTT) between each Vantage Point (VP) and specific destination Internet addresses.[5] The nugget of our technical method – called time-sequence latency probing (TSLP) – is to measure in a particular way to capture a sequence of RTT values to the near and far ends of known interconnection links that connect routers of two different ISPs.

[4]It is also true that either party to a dispute can throttle, or in extreme cases, disconnect the link, which is referred to as *de-peering*. Such de-peering events have generally prompted public attention to the fact that bargaining negotiations have broken down, which in some cases has fueled calls for regulatory intervention.

[5]The probes are small computers, currently Raspberry Pis, deployed in the home of a customer of an ISP. The probe typically connects to the customer's home router, through which it sends packets to the Internet.

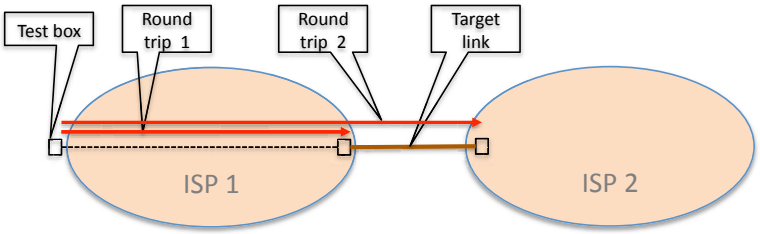

Figure 1: *Finding evidence of congestion by measuring round trip time to near and far side of a specific link, in this case connecting two ISPs.*

Our TSLP method takes advantage of two facts. First, traffic demand on Internet links typically exhibit a strong diurnal pattern, with peak demand during the evening, and low demand in the early morning. Congested links are thus normally congested for only part of a day, corresponding to times of heavier load. The second fact we exploit is router queuing behavior under load. Specifically, when a link is congested, a queue of packets forms at the input to that link. Measured round trip times are a function of the queue lengths of the routers on the forward and reverse paths: as queue lengths increase, so does round trip time. When round trip times increase to the far router but not to the near router, we can infer that a queue between these two routers induced the delay. If we see a diurnal variation in the round trip delay to the far end but not the near end of the link, we can assume that there is congestion associated with this link. Figure 1 illustrates the probing setup.

If a link is so busy that for part of the day the queue is always close to full, a time series of round trip time measurements to the far router will appear as a square wave, with the minimum round trip time during the low state reflecting probes that did not experience delay, and the minimum round trip time during the high state reflecting probes consistently encountering a queue close to full. Queue lengths are finite, reflected by the top of the square wave. Figure 2 shows such a round trip time pattern on a peering link between Comcast and Cogent; the minimum round trip time measured every five minutes to the Cogent router increased from 20ms to 70ms for 14-18 hours per day.

The height of the elevated period is not an indication of the *degree* of congestion, but rather the size of a queue in the router serving the interdomain link. The width of the elevated period indicates its duration, i.e., how much of the day the link was congested.

While our method interprets increased delay as a signal of congestion, it is important to understand that the actual consequence of congestion is not just increased delay. If during a peak period a link has less than the capacity necessary to carry the offered load, then all the flows through that link will have their rate reduced proportionally. It is this reduction in rate as well as the increase in delay that may lead to impairment of the user experience. For a bulk data transfer where no user is waiting for the result, the consequence of a reduced rate may be minimal. For a user looking at a web page, the result may be a less responsive and more sluggish experience. For streaming video, the result might be rebuffering episodes, or inability to watch the movie at a high resolution (or at all). at its highest resolution. For a teleconference, both the queue (latency and jitter) and the reduced throughput may impair the call. Given that these links are in the core of the Internet, for example an interconnection link, all sorts of traffic will be mingled, each affected in their own way by

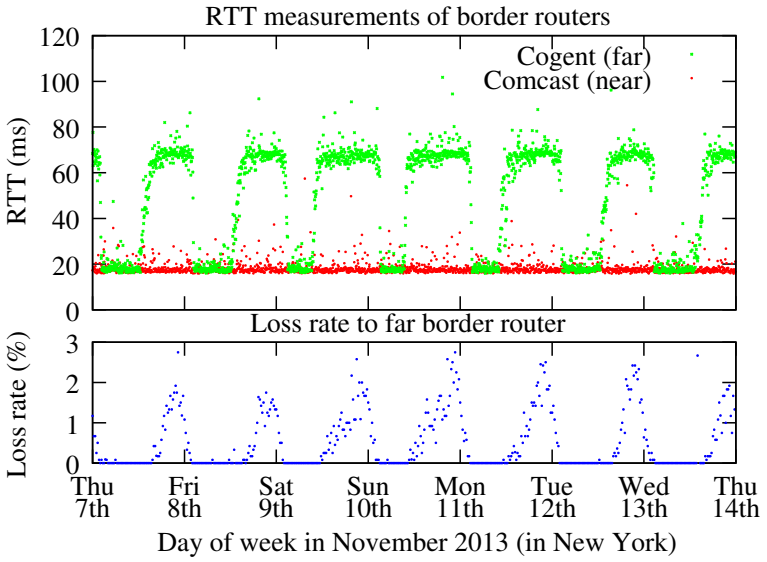

Figure 2: *Congestion on an interdomain link between Comcast and Cogent, measured from a VP within Comcast. The round trip time to the Cogent (far) router increases from 20ms to 70ms while the round trip time to the Comcast (near) router is stable at 20ms. The square wave indicates the queue is always close to full when the round trip time increases to 70ms. The loss rate from the Cogent router increases after this level shift occurs.*

the common experience of flowing through this congested link.

From one vantage point, we can find every interconnection link that connects to that serving region of the access ISP. A large access ISP like Comcast or Time Warner might have as many as 50 peers, and in many cases, they would interconnect at multiple cities, presumably both for resilience and efficient forwarding. So for any of our probes, we might measure hundreds of links. For a large ISP with a national footprint, different serving areas may be configured to use different interconnection links, since large ISPs may interconnect at many points across their respective networks. Therefore, we can say with some confidence that we have found all the points of interconnection that relate to a user in one serving area, but we do not have the density of probes at the present time to be sure that we have found every link leaving a large ISP. So at the present time an access ISP might have a congested link that we fail to detect. However, we do not believe that this possibility prevents us from drawing useful insights from our data.

We can also infer some evidence of congestion using RTT time-series measurement data collected in the background (not the custom TSLP measurement process) by other Archipegalo monitors, although this measurement process slowly probes the entire IPv4 space, so so only a subset of links of the VP network are sampled frequently enough to reveal diurnal fluctuations Figure 4 (discussed in Section 4.2) uses this background measurement data to show a longitudinal view of an interesting link, and Figures 5 (b) and (c) (discussed in Section 4.4) use these background measurements because we did not have a Level3 vantage point running the TSLP process at the time of the study.

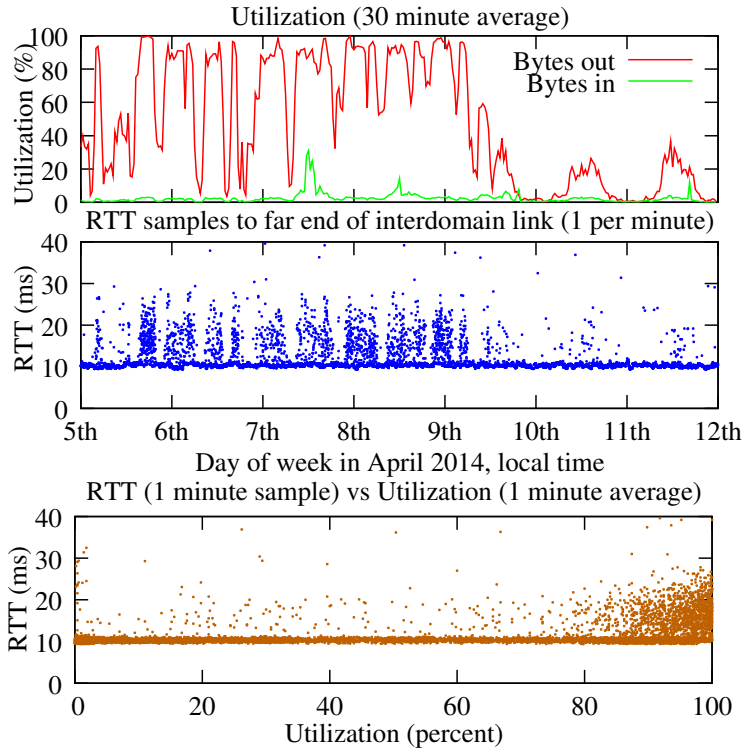

Figure 3: *Comparing measured round trip times with utilization on a 50Mbps customer link operated by a R&E network. The bottom scatterplot suggests that as the 1-minute average load exceeds ≈85%, increasing numbers of probes to the far side of the link encounter a queue, suggesting congestion at that moment.*

3.1 Validation of our inferences of congestion

As a means to validate our method, we used the case of Comcast and Cogent, illustrated in Figure 2, and we also probed every second to observe packet loss across this link; we only observed losses in periods where we also observed increased round trip times. Increased RTTs and packet losses are the two primary indicators of Internet congestion; our data suggests some correlation between them during the periods we observe.

Lending further confidence to this method, Figure 3 plots a week of traffic across a research and education network link known to be congested, where we had access to traffic data that showed that link's utilization, and round trip time measurements made using the TSLP method. The 30-minute average utilization on the link (top graph) correlates with periods when some probes experienced increased round trip times to the far end of the interdomain link (middle graph). The bottom graph shows that most round trip time measurements above 10ms occurred when the average utilization was above 85%, which given the stochastic nature of Internet traffic, represents a high level of capacity utilization.

To validate our results more generally, we asked the ISPs involved to validate our inferences of congestion for their links that exhibited this diurnal RTT behavior. Although they are generally blocked by NDAs from sharing traffic data, informal feedback from content, transit, and access network operators has given us confidence in our observations.

3.2 Benefits and limitations of the TSLP approach

Several strengths and weaknesses of the TSLP method bear mention. First, we emphasize that our probing does not reveal anything not already known to the network operators involved; they generally have internal data that reveals the level of congestion on interconnection links. But operators are typically unwilling or unable to reveal interconnection traffic or performance information, much of which has been covered by NDAs for many years.

The primary strength of our method is that it allows an independent, third-party determination of where congestion exists, using a lightweight (low cost) measurement tool. Other proposed methods, e.g., [16], try to instantaneously detect congestion along a path, whereas our method must gather data over time to find evidence of the diurnal fluctuation. However, this method require a server at the far end of the path being measured, which means probing all paths exiting a network such as Comcast would require perhaps a hundred servers. Furthermore, this method only identifies that a path from a client to one of the supported servers is congested, not which link. Our TSLP method does not require a server, so it can probe any link from one of our probes, so log as no intermediate links are also congested (in which case the two signals would interfere). These points highlight the fact that no single tool is best for all circumstances, and the availability and concurrent use of multiple tools will likely offer additional insights.

A congested high capacity link might affect many more users than a smaller capacity link. Because we have no access to traffic data, and we cannot measure the raw capacity of any link, we cannot generally draw conclusions about the fraction of traffic entering these access ISPs that may have been affected by observed congestion. This lack of visibility limits our ability to interpret our measurements in terms of harm to consumers or content/application providers, or to recommend an appropriate response (further discussed in Section 6).

While TSLP is a simple and surprisingly effective method for inferring congestion, there are many challenges to applying it effectively: accurately finding and identifying all interdomain links exiting the network being probed (i.e., hosting a vantage point); proving the response to our probe returns over the targeted interdomain link; determining the direction of congestion; robustness to router (ICMP) queuing behavior; adapting to path dynamics; and automating processing to scale to thousands of paths. For a discussion of the technical issues related to this measurement approach, see the discussion in [15].

Finally, our measurement method can fail to detect congestion in ISPs who implement network management practices that affect our probes. For example, if ISPs segregate traffic into classes, and implement a queuing discipline to handle them separately, e.g., Weighted Fair Queueing (WFQ), then the ISP can control the degree of congestion that each class experiences (a type of fast/slow-lane scheme). One class of traffic may see an uncongested path, while another may see congestion and a full queue. If our probe traffic happens to be in the uncongested class, we will not see the congestion that the other class observes. We believe that we have seen examples of this sort of traffic management in some of our measurements.

4 Case studies

In this section we present five brief case studies illustrating the most interesting aspects of the empirical data we have gathered and analyzed.

4.1 Netflix traffic to Comcast via Cogent

Figure 2, discussed above, illustrated the congestion we observed on a link between Comcast and Cogent. In the summer of 2013, Netflix began to use Cogent as a transit provider to serve traffic to Comcast (and other access providers). These measurements, taken from a probe located in Comcast's residential broadband network in the Boston serving area, shows the interconnection point between Cogent and Comcast serving Boston for a week in November 2013. The data indicates that the link was congested between 14 and 18 hours a day during this week. This congestion signal is stronger on weekends, including Monday which in this case happened to be a holiday. While our data does not itself prove that the observed congestion was harmful, it seems reasonable to conclude that congestion of this persistence is likely to have adversely affected many flows (and thus users).

4.2 The rise and fall of congestion on Comcast links

Our vantage point from a Comcast residential business customer's home network in Monterey, CA continuously performed measurements across all inter-AS (peering and transit) links Comcast uses in this serving area, including interdomain links with Cogent, Level3, and Tata that were known to have carried Netflix traffic during our period of study. At the time, Tata was the transit provider for Comcast. Figure 4 shows the estimated duration of congestion on these links (in hours per day) from February 2013 to April 2014. Both the Cogent and Level3 links grow from being congested 4-6 hours per day in February 2013 to a maximum of 18 hours (Cogent) and 14 hours (Level3) in February 2014. From mid-May to mid-July, the congestion signal on the Level3 link is replaced with a congestion signal on the Tata link, suggesting a significant volume of traffic shifted from the Level3 to the Tata link. In late February 2014, Netflix and Comcast agreed to peer directly, following which the congestion signal on the Cogent and Level3 links disappeared.

After the February 2014 agreement between Comcast and Netflix, our measurement probes from this Comcast VP started traversing direct peering links toward Netflix prefixes. For most interconnections to Netflix, there was no level shift in round trip time values that indicated a queue persistently full during part of the day. However, the peering link between Comcast and Netflix in San Jose, CA still appeared congested for 2-3 hours per day in April 2014 (figure 5a). We asked Netflix about this congestion and learned that they had not completely deployed the required peering capacity with Comcast at San Jose.

This case study illustrates the extent to which congestion manifested during the period of negotiation between Comcast and Netflix, and also reveals that congestion can more or less instantly shift (in a day or so) from one path to another. Again, while we do not argue that our data shows the cause of congestion, these rapid shifts strongly suggest that the correct response to growing congestion is not always to add more capacity. On the contrary,

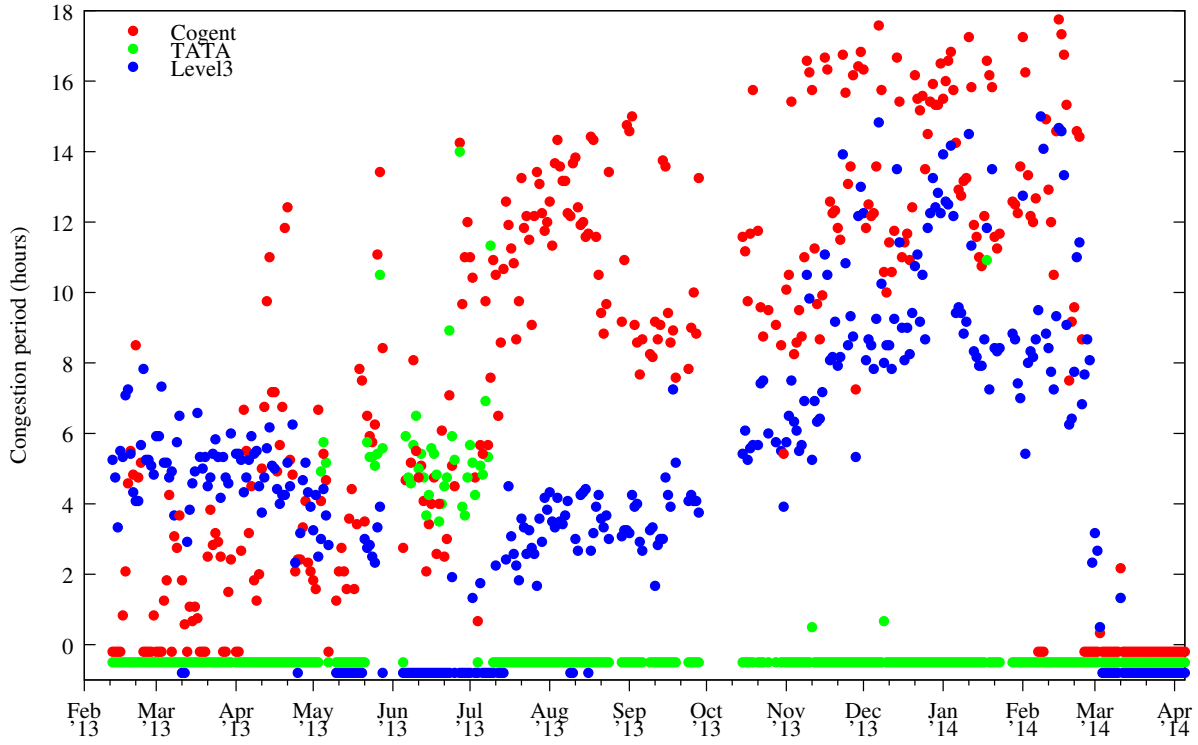

Figure 4: *Estimated congestion duration for links connecting three major networks to Comcast. By February 2014, the Cogent and Level3 links were congested up to 18 and 14 hours per day, respectively. After Netflix and Comcast signed a peering agreement in February 2014, congestion on those links disappeared.*

adding capacity to a link might be a poor investment if a content provider can shift a huge fraction of the traffic from that link to another link overnight.

4.3 Google and Free

Inspired by customer reports of poor performance of Youtube videos on Free's network [1], we used our vantage point in Free's network to measure its direct links to Google. Figure 5b shows our measurements to the near and far end of one of these links. This link appears congested for 4-6 hours at peak time on weekdays, and more than 12 hours on the weekends (March 22nd and 23rd). This case shows that not all of the business issues that can lead to congestion are due to Netflix. Notably, in this case we also found congestion on other links out of Free.[6]

4.4 Level3 and AT&T, Verizon

In May 2014, Level3 published an article on their persistently congested links with five large broadband consumer networks in the U.S. and one in Europe [21]. They published a graph

[6]Because we observed anomalous evidence in his case, i.e., two links from Free to different ASes with the same congestion pattern, we need additional measurements or validation to confirm these inferences.

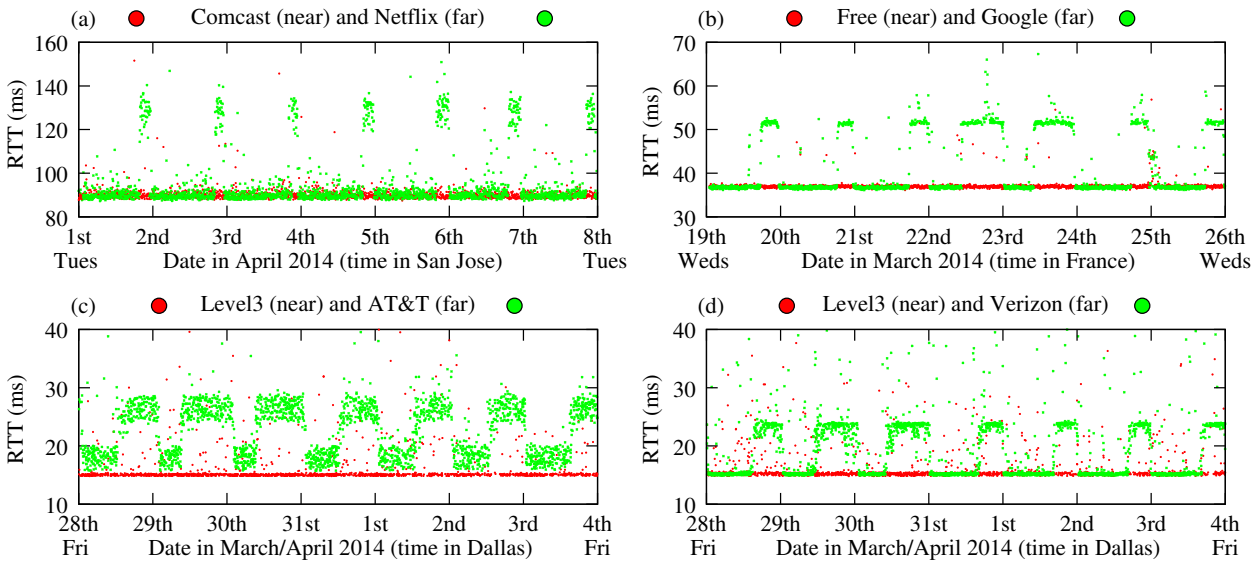

Figure 5: *Case studies of four interdomain connections. All graphs are plotted in the local time zone of the respective links.*

of their interconnect with an unnamed peer in Dallas, Texas. From our probe in Level3's network, we were able to observe congestion on Level3's interconnections with both AT&T and Verizon (we presume two of the broadband consumer networks whose interconnection links Level3 depicted) in Dallas (figures 5c and 5d). These links appeared congested for 6 hours during weekdays and up to 16 hours during weekends in March/April 2014, and both are consistent with the anonymized graph published by Level3. While both time series are similar for this selected week, we observe differences in the period of congestion on March 31st: AT&T was congested longer. This example illustrates the power of our technique to provide independent evidence of congestion on specific interconnection links, identify the parties involved when the parties themselves cannot do so because NDA agreements, and inform the public debate.

4.5 General observations about our measurements

Most of the thousands of interdomain links we have measured so far show little evidence of persistent congestion. We see a few congested paths to peers in many networks we probe, but on a small fraction of paths, and the congestion is usually brief. Since Comcast has been in the news recently, we looked specifically at the data from Comcast. Looking at data gathered in April 2014, Comcast has more than 50 peers, and for some of them multiple paths. We find moderate congestion (mostly between 4%-7% of the time, with one path at 19%) on links going directly to Netflix. We see one peer with two links where both links seem congested (13% and 37% of the time) and four other peers where one of several links shows congestion. The other networks we have probed (Table 1) show similar results. The networks have between 50 and 100 peers, many with multiple paths, and typically show congestion, usually of short duration, on a few paths (less than 10). We also observed congested links connecting some business customers of the access networks we probed, which is reasonable

since end-users balance their willingness to tolerate peak congestion against the costs of a higher bandwidth service. We also saw congestion events on some ISP-internal trans-oceanic links, which are especially expensive to upgrade.

The fact that persistent congestion is not pervasive is good news, but the lack of public data on interdomain congestion amplifies the value of a lightweight technique that can locate interdomain congestion using a tool at the edge of the network.

5 Economic and policy considerations of congestion

Our data reveals only the presence and duration of certain congestion episodes; it does not answer questions about their causes or implications. In this section we speculate about history, causes, and implications of congestion and peering disputes, with the goal of shaping a policy discussion around these issues. We stress that our data does not directly lead to these sorts of conclusions, but in conjunction with other measurements, and analysis of industry structure and practice, our data can inform this discussion.

As noted earlier, congestion is a normal occurrence in well-managed systems of all types, and a key design component of the Internet. However, congestion may induce service quality degradation and efficiency losses (e.g., under-utilization of other resources due to a congested bottleneck), and may signal incipient or on-going market failures that may require regulatory intervention. The objective of policymakers is thus not to avoid all types of congestion, but to identify and avoid harmful congestion. This challenge requires reconciling total welfare (which may be harmed) and private welfare (where some gain while others lose), a process that may involve different perspectives on congestion and strategies for handling it, e.g, traffic discrimination.

We also noted above the recent heated debates over who should be responsible for mitigation of congestion on interconnection links. For example, in the Netflix/Comcast dispute, there have been dueling claims that either Netflix or Comcast was responsible for the congestion and should pay for upgrading the interconnection point. The acrimony of the debate is increased by mutual accusations of intentional performance degradation, and widespread concern that broadband access providers, including Comcast, hold market power over their customers, who must incur significant costs to switch providers, if there is even a comparable alternative. In the language of regulation, a broadband access provider is a terminating monopoly–the only path to the customers of that ISP is through the facilities of that ISP. This reality suggests that an important question to debate is the extent to which an ISP has an obligation to provide an unimpaired path for others to reach their customers.

Economists usually view congestion as a negative externality (cost) that is imposed on other users of a shared resource by an individual's use. In the case of Internet infrastructure, expanding capacity requires investment that is significantly sunk, fixed or shared, which creates a classic free-rider problem. That is, once the capacity is deployed, the incremental cost of using the capacity is typically small, and users have an incentive to avoid paying for more than the incremental cost. A common economic mechanism to address this problem is the use of capacity contracts, which interconnection agreements are, and they have increased in type and complexity (e.g., partial transit, paid peering, etc.) to accommodate the changing nature of Internet traffic and topology over the last two decades [8, 13]. In well-functioning

markets, expansion in contracting options may be a healthy sign of competition, innovation, and customization to adapt to changing market circumstances. However, the increased complexity, and in particular the opacity, of interconnection agreements might obscure market problems, including potential abuses of market power.

5.1 Content distribution options

As we discussed earlier, content providers or CDNs have additional degrees of freedom to control congestion, by hosting content (e.g. web pages, video, streaming music) at locations close to its consumers, and strategically selecting from which source to transmit the content, in order to affect congestion, improve performance, or reduce the cost of the transmission. Effective use of this strategy requires a sufficient number of uncongested paths to reach a CDN's customers, at a reasonable price. Of course, for business reasons, a content provider might also choose a lower cost path instead of a less congested path to its destination. In abstract terms, when a content provider directly connects to an access ISP (peering, whether revenue-neutral or paid), the two parties have asymmetric options for congestion control. The two parties jointly decide which paths to provision, and then the content provider picks which of these paths to use. Both decisions can either create or mitigate congestion. These degrees of freedom suggest that if regulators are contemplating intervention to deal with observed congestion, they should first try to facilitate cooperation among existing players to resolve the congestion and deliver content efficiently.

5.2 An historical perspective

Despite the current media attention to disputes around content delivery, congestion due to unresolved business issues is not a new phenomenon. Allowing persistent congestion to occur is a well-known strategy [17] that has been used by multiple market participants, including by ISPs on either side of bilateral interconnection negotiations to induce the other ISP to alter an interconnection arrangement. Sometimes the desired outcome is to replace a transit relationship with a peering agreement (revenue-neutral or paid), or to expand capacity for a transit agreement. In other cases, service providers may opt for rate-limiting or otherwise allow peak-period congestion in order to minimize interconnection costs, i.e., to reduce payments for paid peering or transit services.

As Sean Donelan recounted on the NANOG mailing list in 2001 [10], one of the early peering battles in the 1990s involved congestion on a link into one of the first exchange points (the Commercial Internet eXchange, or CIX) that allowed commercial networks to freely exchange traffic:

> The second peering battle was one of packet loss. Sprint tried to make things as painful as possible by never upgrading its connection to the CIX router above a T1. So even though other providers were exchanging traffic at 34Mbps to 45Mbps, Sprint kept their quality of service limited to 1.5Mbps at the CIX.

This was a deliberate strategy by Sprint to try to force other networks to become Sprint customers. If one had traffic that they wanted to exchange with another Sprint customer, exchanging it at the CIX router over Sprint's peering link could be a painful experience.

6 Conclusion and future directions

We have provided a preliminary report of a CAIDA/MIT collaborative measurement research project in its early stages. The project is a multi-year undertaking that has an ultimate goal of producing a "heat map" of congestion, in a form that is useful both to technologists and policymakers. To begin, we developed and tested a simple method – the Time-Series Latency Probing (TSLP) method – to identify congestion on links deep within the network, and used it to study several congestion episodes on links between ISPs (interdomain links) that correlate with reports of contested business negotiations between them. This method cannot detect all types of Internet congestion, but it captures evidence of a type of persistent congestion that generally degrades the user experience, i.e., it suggests a potential consumer harm that merits empirical investigation and consideration of mitigation strategies.

The advantages of this method are its conceptual, implementation, and deployment simplicity. In contrast to experiments that test access speeds, which require VPs at many access points, we can measure interdomain links from a given serving area of an ISP with one VP. Most importantly, the TSLP method does not require a server on the other side of the link being probed. Approaches that depend on a server may reveal evidence of congestion from one measurement, but require either instrumentation in the right two places, and/or complex correlation and tomography to localize the point of congestion. Since the vast majority of links do not exhibit persistent congestion, being able to localize congestion deep in the network from a single endpoint has benefits that justify further attention by the research community to resolve the challenges we have described.

Researchers, industry participants, and policy makers are actively engaged in trying to identify what data and metrics/methods for reporting would best serve market efficiency and the public's right to know about Internet performance, without being prone to misinterpretation and false inferences. Better tools and measurements to characterize congestion events will inform the debate, and support market-based light-handed regulation, by providing publicly available methods to identify harms. Because different actors have different options to deal with congested interconnection links, many of which require coordination by both parties (e.g., actions by a CDN and an access ISP with which it peers), regulators should first try to reduce barriers and provide incentives to collective action to facilitate the efficient delivery of high-volume content, before considering mandating specific link upgrades.

The tool we have developed is inherently an ex post enforcement tool. It can detect evidence of likely harmful congestion that may need remedying; but its failure to detect congestion does not imply there is no harm. However, such a tool has ex ante incentive benefits in that it will deter agents from inducing (or ignoring) congestion due to increased likelihood of being detected and penalized. There is also a risk of an arms race in detection tools and strategies for manipulating congestion.

This research highlights the multidisciplinary challenges of developing and maintaining appropriate measurement infrastructure for the continuously evolving Internet. Changing traffic patterns, user behavior, and and revenue opportunities are prompting service providers across the value chain to modify their business models, with important implications for network management practices. Ensuring that the public debate over Internet traffic management is informed with the right data, appropriately interpreted, will promote market efficiency and offers our best hope for avoiding misguided regulatory interventions.

References

[1] R. Andrews and S. Higginbotham. YouTube sucks on French ISP Free, and French regulators want to know why. *GigaOm*, 2013.

[2] Backdoor Santa. Some Truth About Comcast - WikiLeaks Style, 2010. `http://www.merit.edu/mail.archives/nanog/msg15911.html`.

[3] S. Bauer, D. Clark, and W. Lehr. The evolution of internet congestion. In *Proceedings 2009 The 37th Research Conference On Communication, Information and Internet Policy (TPRC*, 2009.

[4] J. Brodkin. Why YouTube buffers: The secret deals that make-and break-online video. *Ars Technica*, July 2013.

[5] S. Buckley. Cogent and Orange France fight over interconnection issues, 2011.

[6] S. Buckley. France Telecom and Google entangled in peering fight. *Fierce Telecom*, 2013.

[7] Center for Applied Internet Data Analysis. Archipelago Measurement Infrastructure. `http://caida.org/projects/ark`.

[8] Clark, David D. and Lehr, William and Bauer, Steven. Interconnection in the Internet: The Policy Challenge. In *TPRC*, 2011. `http://ssrn.com/abstract=1992641`.

[9] S. Cowley. ISP spat blacks out Net connections. *InfoWorld*, 2005.

[10] S. Donelan. Email to nanog mailing list, Jan. 2001. `https://www.nanog.org/mailinglist/mailarchives/old_archive/2001-01/msg00353.html`.

[11] J. Engebretson. Level 3/Comcast dispute revives eyeball vs. content debate, Nov. 2010.

[12] J. Engebretson. Behind the Level 3-Comcast peering settlement, July 2013. `http://www.telecompetitor.com/behind-the-level-3-comcast-peering-settlement/`.

[13] P. Faratin, D. Clark, S. Bauer, W. Lehr, P. Gilmore, and A. Berger. The growing complexity of Internet interconnection. *Communications and Strategies*, (72):51–71, 2008. `http://ssrn.com/abstract=1374285`.

[14] Jon Brodkin. Time Warner, net neutrality foes cry foul over Netflix Super HD demands, 2013.

[15] M. Luckie, A. Dhamdhere, D. Clark, B. Huffaker, and kc caffy. Challenges in inferring internet interdomain congestion. In *Internet Measurement Conference (IMC)*, 2014. accepted for presentation.

[16] M. Mathis. Model Based Bulk Performance Metrics. `http://datatracker.ietf.org/doc/draft-ietf-ippm-model-based-metrics/`, Oct 2013.

[17] W. Norton. *The 2014 Internet Peering Playbook: Connecting to the Core of the Internet*. DrPeering Press, 2014.

[18] H. Online. France Telecom severs all network links to competitor Cogent, 2005. http://morse.colorado.edu/~epperson/courses/routing-protocols/handouts/cogent-ft.html.

[19] M. Ricknas. Sprint-Cogent dispute puts small rip in fabric of Internet. *PCWorld*, Oct. 2008.

[20] Sandvine. Global Internet Phenomenon Report, 1H13, 2013.

[21] M. Taylor. Observations of an Internet Middleman, May 2014. http://blog.level3.com/global-connectivity/observations-internet-middleman/.

[22] Verizon. Unbalanced peering, and the real story behind the Verizon/Cogent dispute, June 2013. http://publicpolicy.verizon.com/blog/.