A View From the Edge: A Stub-AS Perspective of Traffic Localization and its Implications

Bahador Yeganeh
University of Oregon
byeganeh@cs.uoregon.edu

Reza Rejaie
University of Oregon
reza@cs.uoregon.edu

Walter Willinger
NIKSUN, Inc.
wwillinger@niksun.com

Abstract—**Serving user requests from near-by caches or servers has been a powerful technique for localizing Internet traffic with the intent of providing lower delay and higher throughput to end users while also lowering the cost for network operators. This basic concept has led to the deployment of different types of infrastructures of varying degrees of complexity that large CDNs, ISPs, and content providers operate to localize their user traffic. Prior measurement studies in this area have focused mainly on revealing these deployed infrastructures, reverse-engineering the techniques used by these companies to map end users to close-by caches or servers, or evaluating the performance benefits that "typical" end users experience from well-localized traffic.**

To our knowledge, there has been no empirical study that assesses the nature and implications of traffic localization as experienced by end users at an actual stub-AS. This paper reports on such a study for the stub-AS UOnet (AS3582), a Research & Education network operated by the University of Oregon. Based on a complete flow-level view of the delivered traffic from the Internet to UOnet, we characterize the stub-AS's *traffic footprint* (*i.e.* a detailed assessment of the locality of the delivered traffic by all major content providers), examine how effective individual content providers utilize their built-out infrastructures for localizing their delivered traffic to UOnet, and investigate the impact of traffic localization on perceived throughput by end users served by UOnet. Our empirical findings offer valuable insights into important practical aspects of content delivery to real-world stub-ASes such as UOnet.

I. Introduction

During the past two decades, various efforts among different Internet players such as large Internet service providers (ISP), commercial content distribution networks (CDN) and major content providers (CP) have focused on supporting the *localization* of Internet traffic. Improving traffic localization has been argued to ensure better user experience (in terms of shorter delays and higher throughput) and also results in less traffic traversing an ISP's backbone or the interconnections (*i.e.*, peering links) between the involved parties (*e.g.*, eyeball ASes, transit providers, CDNs, CPs). As a result, it typically lowers a network operator's cost and also improves the scalability of the deployed infrastructure in both the operator's own network and the Internet at large.

The main idea behind traffic localization is to satisfy a user request for a certain piece of content by re-directing the request to a cache or front-end server that is in close proximity to that user and can serve the desired piece of content. However, different commercial content distribution companies use different strategies and deploy different types of infrastructures to implement their business model for getting content closer to the end users. For example, while Akamai [1] operates and maintains a global infrastructure consisting of more then 200K servers located in more than 1.5K different ASes to bring the requested content by its customers closer to the edge of the network where this content is consumed, other CDNs such as Limelight or EdgeCast rely on existing infrastructure in the form of large IXPs to achieve this task [2]. Similar to Akamai but smaller in scale, major CPs such as Google and Netflix negotiate with third-party networks to deploy their own caches or servers that are then used to serve exclusively the CP's own content. In fact, traffic localization efforts in today's Internet continue as the large cloud providers (*e.g.*, Amazon, Microsoft) are in the process of boosting their presence at the edge of the network by deploying increasingly in the newly emerging 2nd-tier datacenters (*e.g.*, EdgeConneX [3]) that target the smaller- or medium-sized cities in the US instead of the major metropolitan areas.

These continued efforts by an increasing number of interested parties to implement ever more effective techniques and deploy increasingly more complex infrastructures to support traffic localization has motivated numerous studies on designing new methods and evaluating existing infrastructures to localize Internet traffic. While some of these studies [4]–[7] have focused on measurement-based assessments of different deployed CDNs to reveal their global [4], [5] or local [8], [9] infrastructure nodes, others have addressed the problems of reverse-engineering a CDN's strategy for mapping users to their close-by servers or examining whether or not the implemented re-direction techniques achieve the desired performance improvements for the targeted end users [6], [7], [9]. However, to our knowledge, none of the existing studies provides a detailed empirical assessment of the nature and impact of traffic localization as seen from the perspective of an actual stub-AS. In particular, the existing literature on the topic of traffic localization provides little or no information about the makeup of the content that the users of an actual stub-AS request on a daily basis, the proximity of servers that serve the content requested by these users (overall or per major CP), and the actual performance benefits that traffic localization entails for the consumers of this content (*i.e.*, end users inside the stub-AS).

In this paper, we fill this gap in the existing literature and report on a measurement study that provides a detailed assessment of different aspects of the content that arrives at an actual stub-AS as a result of the requests made by its end users. To this end, we consider multiple daily snapshots of unsampled Netflow data for all exchanged traffic between a stub-AS that represents a Research & Education network (*i.e.*, UOnet operated by the University of Oregon) and the Internet II. We show that some 20 CPs are responsible for most of the delivered traffic to UOnet and that for each of these 20 CPs, the CP-specific traffic is typically coming from only a small fraction of source IPs (Section III). Using RTT to measure the distance of these individual source IPs from UOnet, we present a characterization of this stub-AS' *traffic footprint*; that is, empirical findings about the locality properties of delivered traffic to UOnet, both in aggregate and at the level of individual CPs (Section IV). In particular, we examine how effective the individual CPs are in utilizing their infrastructure nodes to localize their delivered traffic to UOnet and discuss the role that *guest servers* (*i.e.*, front-end servers or caches that some of these CPs deploy in third-party networks) play in localizing traffic for this stub-AS (Section V). As part of this effort, we focus on Akamai and develop a technique that uses our data to identify all of Akamai's guest servers that delivered content to UOnet. We then examine different features of the content that arrived at UOnet from those guest servers as compared to the content that reached UOnet via servers located in Akamai's own AS. Finally, we investigate whether or not a CP's ability to localize its traffic has implications on end user-perceived performance, especially in terms of observed throughput (Section VI).

II. Data Collection for a Stub-AS: UOnet

The stub-AS that we consider for this study is the campus network of the University of Oregon (UO), called UOnet (ASN3582). UOnet serves more than 24K (international and domestic) students and 4.5K faculty/staff during the academic year. These users can access the Internet through UOnet using wireless (through 2000+ access points) or wired connections. Furthermore, more than 4,400 of the students reside on campus and can access the Internet through UOnet using their residential connections. UOnet has three upstream providers, Neronet (AS3701), Oregon Gigapop (AS4600) and the Oregon IX exchange. Given the types of offered connectivity and the large size and diversity of the UOnet user population, we consider the daily traffic that is delivered from the rest of the Internet to UOnet to be representative of the traffic that a stub-AS that is classified as a US Research & Education network is likely to experience.

To conduct our analysis, we rely on un-sampled Netflow (v5) data that is captured at the different campus border routers. As a result, our Netflow data contains all of the flows between UOnet users and the Internet. The Netflow dataset contains a separate record for each incoming (and outgoing) flow from (to) an IP address outside of UOnet, and each record includes the following flow attributes: *(i)* source and destination IP addresses, *(ii)* source and destination port numbers, *(iii)* start and end timestamps, *(iv)* IP protocol, *(v)* number of packets, and *(vi)* number of bytes. We leverage Routeviews data to map all the external IPs to their corresponding Autonomous Systems (ASes) and use this information to map individual flows to particular providers (based on their AS number) and then determine the number of incoming (and outgoing) flows (and corresponding bytes) associated with each provider. In our analysis, we only consider the incoming flows since we are primarily interested in delivered content and services from major content providers to UOnet users. An incoming flow refers to a flow with the source IP outside and destination IP inside UOnet. We select 10 daily (24 hour) snapshots of Netflow data that consist of Tuesday and Wednesday from five consecutive weeks when the university was in session, starting with the week of Oct 3rd and ending with the week of of Oct 31st in 2016. Table I summarizes the main features of the selected snapshots, namely their date, the number of incoming flows and associated bytes, and the number of unique external ASes and unique external IPs that exchanged traffic with UOnet during the given snapshot. In each daily snapshot, wireless connections are responsible for roughly 62% (25%) of delivered bytes (flows) and residential users contributed to about 17% (10%) of incoming bytes (flows).

TABLE I
Main features of the selected daily snapshots of our UOnet Netflow data.

Snapshot	Flows (M)	TBytes	ASes (K)	IPs (M)
10/04/16	196	8.7	39	3.3
10/05/16	193	8.5	37	3.0
10/11/16	199	9.0	41	4.1
10/12/16	198	9.1	41	4.7
10/18/16	202	8.8	40	3.7
10/19/16	200	9.1	38	3.3
10/25/16	205	8.7	37	2.9
10/26/16	209	9.1	40	4.1
11/01/16	212	8.6	39	3.5
11/02/16	210	8.7	40	4.3

III. Identifying Major Content Providers

Our main objective is to leverage the UOnet dataset to provide an empirical assessment of traffic locality for delivered flows to UOnet and examine its implications for the end users served by UOnet. Here by "locality" we refer to a notion of network distance between the servers in the larger Internet that provide the content/service requested from within UOnet. Since the level of locality of delivered traffic by each CP depends on both the relative network distance of its infrastructure and its strategy for utilizing this infrastructure, we conduct our analysis at the granularity of individual CPs and focus only on those that are responsible for the bulk of delivered content to UOnet. Moreover, because the number of unique source IPs that send traffic to UOnet on a daily basis is prohibitively large, we identify and focus only on those IPs that are responsible for a significant fraction of the delivered traffic.

Inferring Top CPs: Figure 1 (left y-axis) shows the histogram of delivered traffic (in TB) to UOnet by those CPs that have the largest contributions in the 10/04/16 snapshot. It also shows

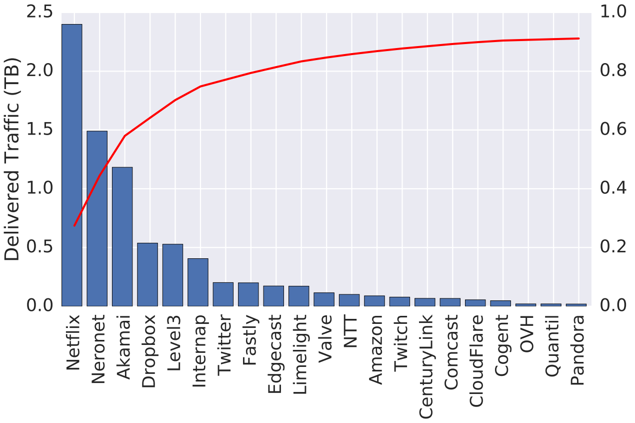

Fig. 1. The volume of delivered traffic from individual top CPs to UOnet along with the CDF of aggregate fraction of traffic by top 21 CPs in the 10/04/16 snapshot.

(right y-axis) the CDF of the fraction of aggregate traffic that is delivered by the top-k CPs in this snapshot. The figure is in full agreement with earlier studies such as [10], [11] and clearly illustrates the extreme skewness of this distribution – the top 21 CPs (out of some 39K ASes) are responsible for 90% of all the delivered daily traffic to UOnet.

To examine the stability of these top CPs across our 10 daily snapshots, along the x-axis of Figure 2, we list any CP that is among the top CPs (with 90% aggregate contributions in delivered traffic) in at least one daily snapshot (the ordering is in terms of mean rank, from small to large for CPs with same prevalence). This figure shows the number of daily snapshots in which a CP has been among the top CPs (*i.e.* CP's prevalence, left y-axis) along with the summary distribution (*i.e.*, box plot) of each of the CPs rankings among the top CPs across different snapshots (rank distribution, right y-axis). We observe that the same 21 CPs consistently appear among the top CPs. These 21 CPs are among the well-recognized players of today's Internet and include major CPs (*e.g.* Netflix, Twitter), widely-used CDNs (*e.g.* Akamai, LimeLight and EdgeCast), and large providers that offer hosting, Internet access, and cloud services (*e.g.* Comcast, Level3, CenturyLink, Amazon). In the following, we only focus on these 21 CPs (called *target CPs*) that are consistently among the top CPs in all of our snapshots. These target CPs are also listed in Figure 1 and collectively contribute about 90% of the incoming daily bytes in each of our snapshots.

Inferring Top IPs per Target CP: To assess the locality of the traffic delivered to UOnet from each target CP, we consider the source IP addresses for all of the incoming flows in each daily snapshot. While for some target CPs, the number of unique source IP addresses is as high as a few tens of thousands, the distribution of delivered traffic across these IPs exhibits again a high degree of skewness; *i.e.* for each target CP, only a small fraction of source IPs (called *top IPs*) is responsible for 90% of delivered traffic. Figure 3 shows the summary distribution (in the form of box plots) of the number of top IPs across different snapshots along with the cumulative number of unique top IPs (blue line) and all IPs (red line) across all of our 10 snapshots. The log-scale on the y-axis shows that the number of top IPs is often significantly smaller than the number of all IP addresses (as a result of the skewed distribution of delivered content by different IPs per target CP). A small gap between the total number of top IPs and their distribution across different snapshots illustrates that for many of the target CPs, the top IPs do not vary widely across different snapshots. In our analysis of traffic locality below, we only consider the collection of all top IPs associated with each of the target CP across different snapshots. Focusing on these roughly 50K IPs allows us to capture a rather complete view of delivered traffic to UOnet without considering the millions of observed source IPs.

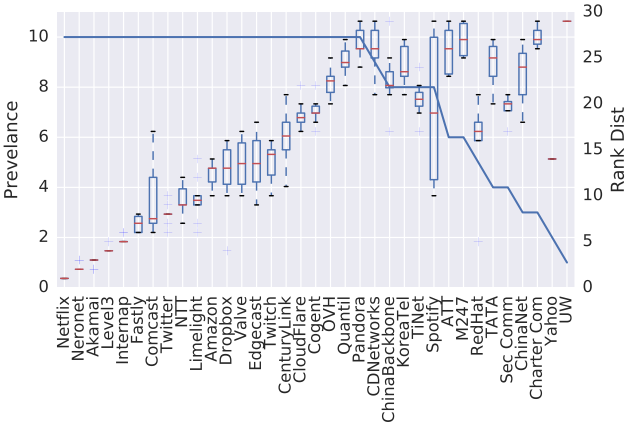

Fig. 2. The prevalence and distribution of rank for any CP that has appeared among the top CPs in at least one daily snapshot.

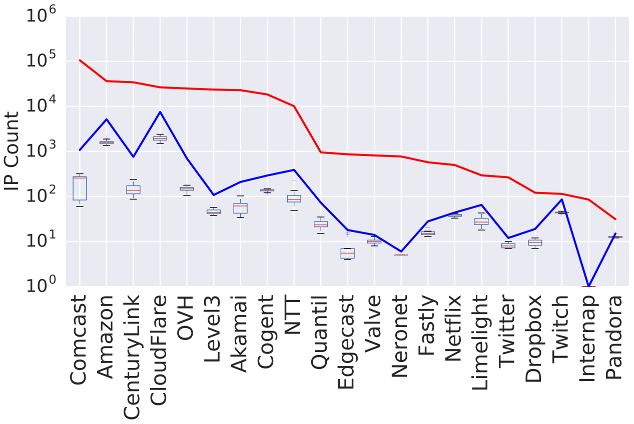

Fig. 3. Distribution of the number of top IPs across different snapshots in addition to total number of unique top IP addresses (blue line) and the total number of unique IPs across all snapshots (red line) for each target CP.

Measuring the Distance of Top IPs: Using the approximately 50K top IPs for all 21 target CPs, we conducted a measurement campaign (on 11/10/16) that consisted of launching 10 rounds of traceroutes[1] from UOnet to all of these 50K top IPs to infer their minimum RTT.

Note that the value of RTT for each top IP accounts for

[1] We use all three types of traceroute probes(TCP, UDP, ICMP) and spread them throughout the day to reach most IPs and reliable capture minimum RTT

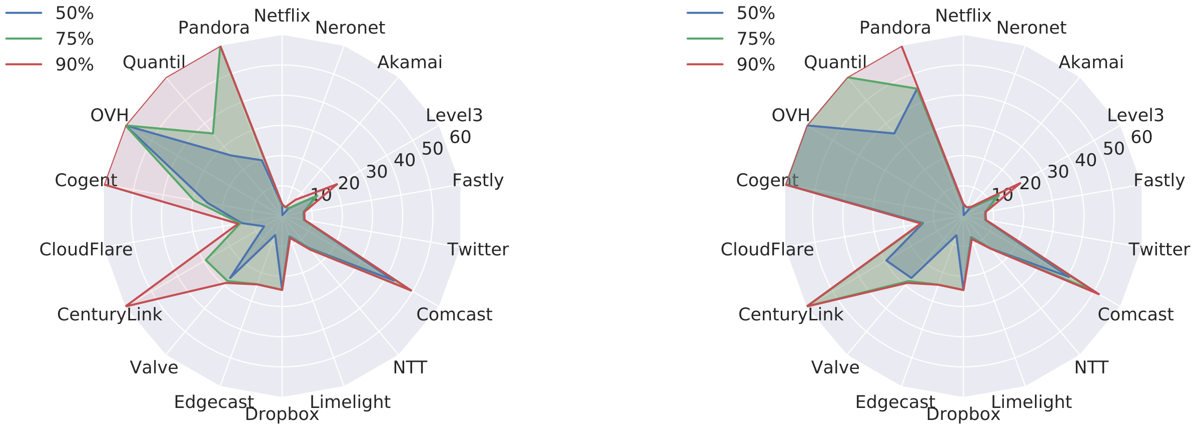

Fig. 4. Radar plots showing the aggregate view of locality based on RTT of delivered traffic in terms of bytes (left plot) and flows (right plot) to UOnet in a daily snapshot (10/04/2016).

possible path asymmetry between the launching location and the target IP and is therefor largely insensitive to the direction of the traceroute probe (*i.e.* from UOnet to a top IP vs. from a top IP to UOnet). Our traceroute probes successfully reached 81% of the targeted IP addresses. We exclude three target CPs (*i.e.*, Internap, Amazon and Twitch) from our analysis because their servers did not respond to more than 90% of our traceroute probes. All other target CPs responded to more than 90% of our probes.

The outcome of our measurement campaign is the list of top IPs along with their min RTT and the percentage of delivered traffic (in terms of bytes and flows) for each target CP. With the help of this information, we can now assess the locality properties of the content that is delivered from each target CP to UOnet. Note that in theory, any distance measure could be used for this purpose. However, in practice, neither AS distance (*i.e.*, number of AS hops), nor hop-count distance (*i.e.*, number of traceroute hops), nor geographic distance are reliable metrics. While the first two ignore the commonly encountered asymmetry of IP-level routes in today's Internet [12], the last metric suffers from known inaccuracies in commercial databases such as IP2Location [13] and Maxmind [14] that are commonly used for IP geolocation. In this paper, we choose the RTT distance (*i.e.*, measured by min RTT value) as our metric-of-choice for assessing the locality of delivered traffic since it is the most reliable distance measure and also the most relevant in terms of user-perceived delay.

IV. Traffic Locality for Content Providers

Overall View of Traffic Locality: We use radar plots to present an overall view of the locality of aggregate delivered traffic from our target CPs to UOnet based on RTT distance. Radar plots are well suited for displaying multi-variable data where individual variables are shown as a sequence of equiangular spokes, called radii. We use each spoke to represent the locality of traffic for a given target CP by showing the RTT values for 50th, 75th and 90th percentiles of delivered traffic (in bytes or flows). In essence, the spoke corresponding to a particular target CP shows what percentage of the traffic that this CP delivers to UOnet originates from within 10, 20,..., or 60ms distance from our stub-AS. Figure 4 shows two such radar plots for a single daily snapshot (10/04/16). In these plots, the target CPs are placed around the plot in a clock-wise order (starting from 12 o' clock) based on their relative contributions in delivered bytes (as shown in Figure 1), and the distances (in terms of min RTT ranges) are marked on the 45-degree spoke. The left and right plots in Figure 4 show the RTT distance for 50, 75 and 90th percentile of delivered bytes and flows for each CP, respectively. By connecting the same percentile points on the spokes associated with the different target CPs, we obtain a closed contour where the sources for 50, 75 or 90% of the delivered content form our target CPs to UOnet are located. We refer to this collection of contours as the *traffic footprint* of UOnet. While more centrally-situated contours indicate a high degree of overall traffic locality for the considered stub-AS, contours that are close to the radar plot's boundary for some spokes suggest poor localization properties for some CPs.

The radar plots in Figure 4 show that while there are variations in traffic locality for different target CPs, 90% of the delivered traffic for the top 13 CPs are delivered from within a 60ms RTT distance from UOnet and for 9 of them from within 20ms RTT. Moreover, considering the case of Cogent, while 50% of bytes from Cogent are delivered from an RTT distance of 20ms, 50% of the flows are delivered from a distance of 60ms. Such an observed higher level of traffic locality with respect to bytes compared to flows suggests that a significant fraction of the corresponding target CP's (in this case, Cogent) large or "elephant" flows are delivered from servers that are in closer proximity to UOnet than those that serve the target CP's smaller flows. Collectively, these findings indicate that for our stub-AS, the overall level of traffic locality for delivered bytes and flows is high but varies among the different target CPs. These observations are by and large testimony to the success of past and ongoing efforts by the different involved parties to bring content closer to the edge of the network where it is requested and consumed. As such, the results are not surprising, but to our knowledge, they provide the first quantitative assessment of the per-CP traffic footprint (based on RTT distance) of a stub-AS.

Variations in Traffic Locality: After providing an overall view of the locality of the delivered traffic to UOnet for a single snapshot, we next turn our attention to how traffic locality of a CP (with respect to UOnet) varies over time. To simplify our analysis, we consider all flows of each target CP and bin them based on their RTTs using a bin size of 2ms. The flows in each bin are considered as a single group with an RTT value given by the mid-bin RTT value. We construct the histogram of percentages of delivered bytes from each group of flows in each bin and define the notion of *Normalized Weighted Locality* for delivered traffic from a provider CP in snapshot s as:

$$NWL(s,CP) = \sum_{i\epsilon RTTBins(CP)} \frac{FracBytes(i) * RTT(i)}{minRTT(CP))}$$

NWL(s,CP) is simply the sum of the fraction of delivered traffic from each RTT bin ($FracBytes(i)$) that is weighted by its RTT and then normalized by the lowest RTT among all bin ($minRTT(CP)$) for a CP across all snapshots. NWL is an aggregate measure that illustrates how effectively a CP localizes its delivered traffic over its own infrastructure. A NWL value of 1 implies that all of the traffic is delivered from the closest servers while larger values indicate more contribution from servers that are further from UOnet.

The top plot in Figure 5 presents the summary distribution of $NWL(s,CP)$ across different daily snapshots for each CP. The bottom plot in Figure 5 depicts min RTT for each CP. These two plots together show how local the closest server of a CP is and how effective each CP is in utilizing its infrastructure. The plots also demonstrate the following points about the locality of traffic. For one, for many target CPs (*e.g.* Netflix, Comcast, Valve), the NWL values exhibits small or no variations across different snapshots. Such a behavior suggests that the pattern of delivery from different servers is stable across different snapshots. In contrast, for CPs with varying NWL values, the contribution of various servers (*i.e.* the pattern of content delivery from various CP servers) changes over time. Second, the value of NWL is less than 2 (and often very close to 1) for many CPs. This in turn indicates that these CPs effectively localize their delivered traffic to UOnet over their infrastructure. The value of NWL for other CPs is larger and often exhibit larger variation due to their inability to effectively utilize their nodes to localize delivered traffic to UOnet.

V. Traffic From Guest Servers

To improve the locality properties of their delivered content and services to end users, some content providers expand their infrastructure by deploying some of their servers in other networks. We refer to such servers as *guest servers* and to the third-party networks hosting them as *host networks or host ASes*. For example, Akamai is known to operate some 200K such servers in over 1.5K different host networks, with the servers using IP addresses that belong to the host networks [7], [15].

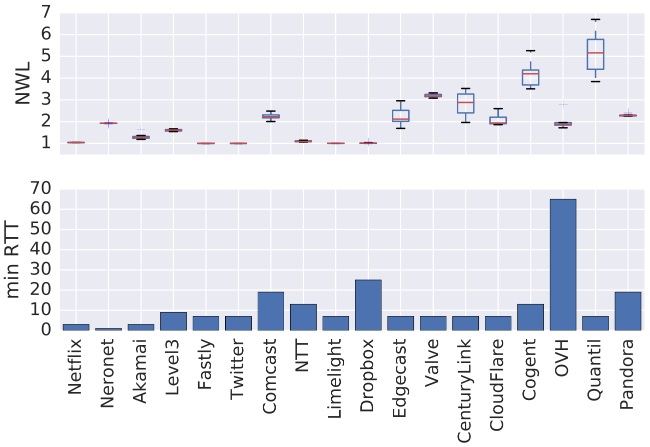

Fig. 5. Two measures of traffic locality, from top to bottom, Summary distribution of NWL and the RTT of the closest servers per CP (or minRTT).

We present two examples to illustrate the deployment of guest servers. First, our close examination of delivered traffic from Neronet which is one of UOnet's upstream providers revealed that all of its flows are delivered from a small number of IPs (see Figure 3) associated with Google servers, *i.e.* Google caches [5] that are deployed in Neronet. This implies that all of Google's traffic for UOnet is delivered from Neronet-based Google caches and explains why Google is not among our target CPs. Second, Netflix is known to deliver its content to end users through its own caches (called Open Connect Appliances [16]) that are either deployed within different host networks or placed at critical IXPs [4]. When examined the DNS names for all the source IPs of our target CPs, we observed a number of source IPs that are within another network and their DNS name follow the `*.pdx001.ix.nflxvideo.net` format. This is a known Neflix convention for DNS names and clearly indicates that these guest servers are located at an IXP in Portland, Oregon [4].

A. Detecting Guest Servers

Given the special nature of content delivery to UOnet from Google (via Neronet) and Netflix (via a close-by IXP), we focus on Akamai to examine how its use of guest servers impact the locality of delivered traffic to UOnet. However, since our basic methodology that relies on a commonly-used IP-to-AS mapping technique cannot identify Akamai's guest servers and simply associates them with their host network, we present in the following a new methodology for identifying Akamai's guest servers that deliver content to UOnet.

Our proposed method leverages Akamai-specific information and proceeds in two steps. The first step consists of identifying the URLs for a few small, static and popular objects that are likely to be cached at many Akamai servers. Then, in a second step, we probe the observed source IP addresses at other target CPs with properly-formed HTTP request for the identified objects. Any third-party server that provides the requested objects is considered an Akamai guest server. More precisely, we first identify a few Akamai customer websites and interact with them to identify small, static and popular objects (*i.e.*, “reference objects”). Since JavaScript or CSS files are less likely to be modified compared to other types of objects and thus are more likely to be cached by Akamai servers, we used in our experiments two JavaScript objects and a logo from Akamai client web sites (*e.g.* Apple, census.gov, NBA). Since an Akamai server is responsible for hosting content from multiple domain names, the web server needs a way to distinguish requests that are redirected from clients of different customer websites. This differentiation is achieved with the help of the HOST field of the HTTP header. Specifically, when constructing a HTTP request to probe an IP address, we set the HOST field to the original domain name of the reference object (*e.g.* apple.com, census.gov, nba.com). Next, for each reference object, we send a separate HTTP request to each of the 50K top source IP addresses in our datasets (see Section 3). If we receive the HTTP OK/200 status code in response to our request and the first 100 bytes of the provided object match the requested reference object [2], we consider the server to be an Akamai guest server and identify its AS as host AS. We repeat our request using other reference objects if the HTTP request fails or times-out. If all of our requests time-out or receive a HTTP error code, we mark the IP address as a non-Akamai IP address.

To evaluate our proposed methodology, we consider all the 601 servers in our dataset whose IP addresses are mapped to Akamai (based on IP to AS mapping) and send our HTTP requests to all of them. Since all Akamai servers are expected to behave similarly, the success rate of our technique in identifying these Akamai servers demonstrates its accuracy. Indeed, we find that 585 (97%) of these servers properly respond to our request and are thus identified as Akamai servers. The remaining 3% either do not respond or respond with various HTTP error codes. When examining these 16 failed servers more closely, we discovered that 11 of them were running a mail server and would terminate a connection to their web server regardless of the requested content. This suggests that these Akamai servers perform functions other than serving web content.

Using our proposed technique, we probed all 50K top source IP addresses associated with our 21 target CPs in all of our snapshots. When performing this experiment (on 11/20/16), we discovered between 143-295 Akamai guest servers in 3-7 host ASes across the different snapshots. In total, there were 658 unique guest servers from 7 unique host ASes, namely NTT, CenturyLink, OVH, Cogent, Comcast, Dropbox and Amazon. Moreover, these identified Akamai guest servers deliver between 121-259 GBytes to UOnet in their corresponding daily snapshots which is between 9-20% of the aggregate daily traffic delivered from Akamai to UOnet. These results imply that the 34-103 Akamai-owned servers in each snapshot deliver on average 12 times more content to UOnet than Akamai’s 143-295 guest servers. Moreover, we observed that the bulk of delivered bytes from Akamai’s guest servers to UOnet (*i.e.*, 98%) is associated with guest servers that are deployed in two CPs, namely NTT (76.1%) and CenturyLink (21.9%).

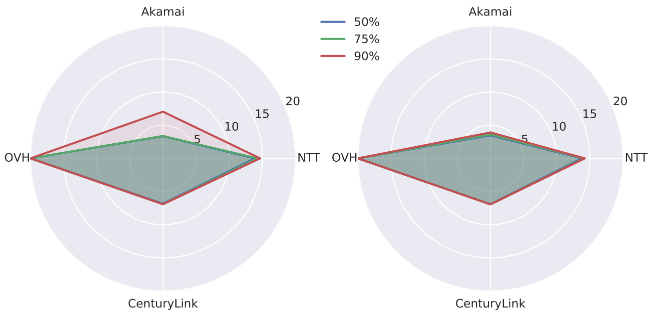

Fig. 6. Locality (based on RTT in ms) of delivered traffic (bytes, left plot; flows, right plot) for Akamai-owned servers as well as Akamai guest servers residing within three target ASes for snapshot 2016-10-04.

B. *Relative Locality of Guest Servers*

Deploying guest servers in various host ASes enables a CP to either improve the locality of its traffic or provide better load balancing among its servers. To examine these two objectives, we compare the level of locality of traffic delivered from Akamai-owned servers vs Akamai’s guest servers. The radar plots in Figure 6 illustrate the locality (based on RTT) of delivered content from Akamai-owned servers shown at 12 o’clock (labeled as Akamai) as well as from Akamai’s guest servers in all three host networks in the snapshot from 10/04/16. The guest servers are grouped by their host ASes and ordered based on their aggregate contribution in delivered bytes (for Akamai flows) in a clock-wise order. We observe that traffic delivered from Akamai-owned servers exhibits a higher locality – 75% (90%) of the bytes (flows) are delivered from servers that are 4ms (8ms) RTT away. The Akamai traffic from CenturyLink, NTT and OVH is delivered from servers that are at RTT distance of 8, 15 and 20ms, respectively. While these guest servers serve content from further away than the Akamai-owned servers, they are all relatively close to UOnet which suggests that they are not intended to offer higher level of locality for delivered content to UOnet users.

VI. Implications of Traffic Locality

Improving end user-perceived performance (*i.e.* decreasing delay and/or increasing throughput) is one of the main motivations for major CPs to bring their front-end servers closer to the edge of the network. In the following, we examine whether such performance improvements are indeed experienced by the

[2] The second condition is necessary since some servers provide a positive response to any HTTP requests.

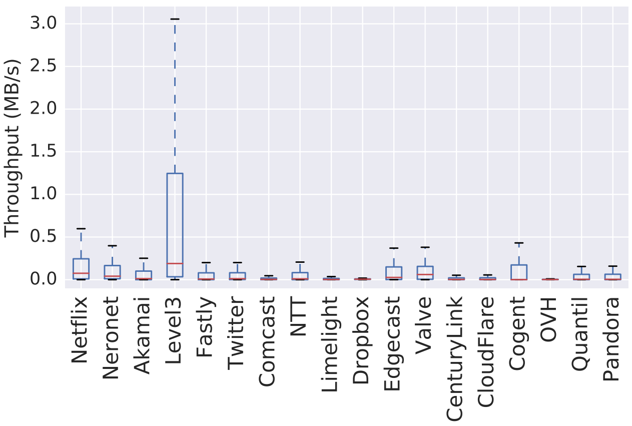

Fig. 7. Summary distribution of average throughput for delivered flows from individual target CPs towards UOnet users across all of our snapshots.

end users served by UOnet and to what extent for a given CP the observed performance is correlated with that CP's traffic locality.

We already showed in Figure 4 that the measured min RTT values for a majority of CPs (with some exceptions such as OVH, Quantil, Cogent) are consistently low (<20ms) across all flows. The average throughput of each flow can be easily estimated by dividing the total number of delivered bytes by its duration [3]. To get an overall sense of the observed average throughput, Figure 7 shows the summary distributions of the measured throughput across delivered flows by each target CP. We observe that 90% of the flows for all target CPs (except Level3) experience low throughput ($<$ 0.5MB/s, and in most cases even $<$ 0.25MB/s). This raises the question why these very localized flows do not achieve higher throughput.

In general, reliably identifying the main factors that limit the throughput of individual flows is challenging [17]. The cause could be any combination of factors that include

- *Content Bottleneck*: the flow does not have sufficient amount of content to "fill the pipe";
- *Receiver Bottleneck*: the receiver's access link (*i.e.* client type) or flow control is the limiting factor;
- *Network Bottleneck*: the fair share of network bandwidth is limited due to cross traffic (and resulting loss rate);
- *Server Rate Limit*: a CP's server may limit its transmission rate implicitly due to its limited capacity or explicitly as a results of the bandwidth requirements of the content (*e.g.* Netflix videos do not require more than 0.6 MB/s for a Full-HD stream [18]).

Rather than inferring the various factors that affect individual flows, our goal is to identify the primary factor from the above list that limits the maximum achievable throughput by individual CPs. To this end, we only consider 3-4% (or 510-570K) of all flows for each target CP that their size exceeds 1 MB and refer to them as "elephant" flows.[4] These elephant flows have typically several 100s of packets and are thus able to fully utilize available bandwidth in the absence of other limiting factors (*i.e.* content bottleneck does not occur). More than 0.5 million elephant flows for individual CPs are delivered to end users in UOnet that have diverse connection types (wireless, residential, wired). Therefor, receiver bottleneck should not be the limiting factor for the maximum achievable throughput by individual CPs. This in turn suggests that either the network or the server are responsible for limiting the achievable throughput.

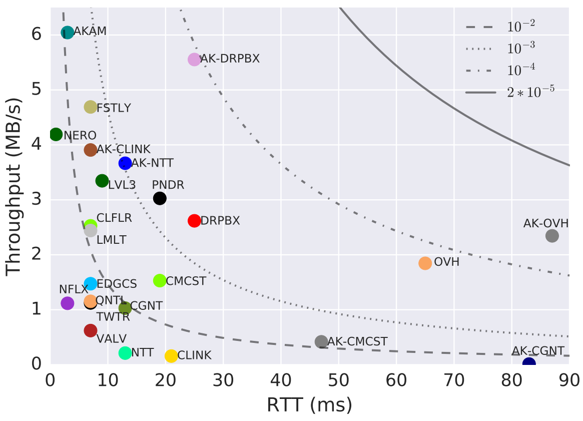

Fig. 8. Maximum Achievable Throughput (MAT) vs MinRTT for all CPs. The curves show the change in the estimated TCP throughput as a function of RTT for different loss rates.

To estimate the *Maximum Achievable Throughput (MAT)* for each CP, we group all elephant flows associated with that CP based on their RTT into 2ms bins and select the 95% throughput value (*i.e.* median of the top 10%) in the bin as its MAT with its mid-bin RTT value as the corresponding RTT. Since a majority (96%) of these flows are associated with TCP connection and thus are congestion controlled, we can examine the key factors responsible for limiting throughput. Figure 8 shows a scatter plot where each labelled dot represents a target CP with its y-value denoting its MAT and its x-value denoting the associated RTT. We also group all Akamai flows from its guest servers at each host AS_x, determine their separate MAT and exclude them from AS_x's own flows to avoid double-counting them. For example, Akamai flows that are delivered from OVH are marked as AK-OVH. To properly compare the measured MAT values across different RTTs, we also plot an estimated TCP throughput as a function of RTT for three different loss rates that we obtain by applying the commonly-used equation [20]: $T < \frac{MSS}{RTT} * \frac{1}{\sqrt{L}}$. In this equation, MSS denotes the Maximum Segment Size which we set to 1460; L represents the loss rate. We consider three different loss rate values, namely 10^{-2}, 10^{-3}, 10^{-4}, to cover a wide range of "realistic" values.

Examining Figure 8, we notice that the relative location

[3] Note that we may have fragmented flows for this analysis. This means that long flows will be divided into 5min intervals. However, 5min is sufficiently long to estimate average throughput of individual flows.

[4] Selecting the 1 MB threshold for flow size strikes a balance between having sufficiently large flows [19] and obtaining a large set of flows for each CP.

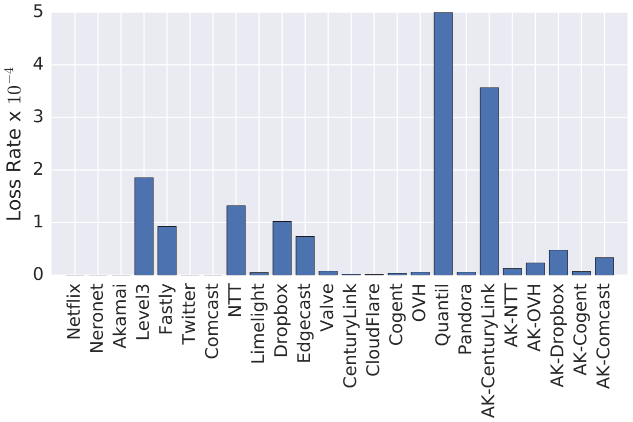

Fig. 9. Average loss rate of closest servers per target CP measured over 24 hours using ping probes with 1 second intervals. For each CP we choose at most 10 of the closest IP addresses.

of each labelled dot with respect to the TCP throughput lines reveals the average "virtual" loss rate across all elephant flows of a CP if bandwidth bottleneck were the main limiting factor. The figure shows that this virtual loss rate for many CPs is at or above 10^{-3}. However, in practice, average loss rates higher than 10^{-3} over such short RTTs (<20ms) are very unlikely in our setting (*e.g.*, UOnet is well provisioned and most incoming flows traverse the paths with similar or identical tail ends). To test this hypothesis, we directly measure the loss rate between UOnet and the closest servers for each CP using 170K ping probes per CP.[5] Figure 9 depicts the average loss rate for each target CP and shows that the measured average loss rate for all of the target CPs is at least an order of magnitude lower than the virtual loss rate for each CP. This confirms that all of the measured MAT values must be rate-limited by the server, either explicitly (due to the bandwidth requirements of the content) or implicitly (due to server overload).

Figure 8 also shows that the measured MAT values for Akamai guest servers are often much larger than those for the servers owned by the host AS. For example, the MAT value for AK-CLINK (AS-DRPBX or AK-NTT) is much higher than the MAT for CLINK (DRPBX, or NTT). Furthermore, the measured MAT value for all the flows from Akamai's guest servers is lower than its counterpart for all flows from Akamai-owned servers.

To summarize, there are two main take-aways from our examination of the performance implications of traffic locality. On the one hand, traffic locality is key to achieving the generally and uniformly very small measured delays for traffic delivered to UOnet. On the other hand, our results show that a majority of flows for all target CPs are associated with small files and thus do not reach a high throughput. Furthermore, the throughput for most of the larger flows are not limited by the network but rather by the front-end servers. In other words, high throughput delivery of content at the edge is either not relevant (for small objects) or not required by applications.

VII. Related Work

There exists an extensive literature on various aspects of content delivery. Past studies in this area can be broadly divided into two groups. The first group consists of studies that focus on uncovering the delivery strategy [5], [7], [8] or discovering the CP-specific infrastructures used for delivering content across the globe (*e.g.* Netflix [4], [6] or Akamai [15]). The latter are measurement studies that either rely on the availability of multiple vantage points across the globe [6]–[8], [15] or leverage a single vantage point by exploiting an intrinsic feature of the infrastructure of interest [4], [5].

The second group of studies is mainly concerned with assessing or measuring the performance of a specific content delivery infrastructure [21], [22] or of a specific type of traffic [9], [10], [23], [24]. These studies often examine various performance-related metrics to evaluate the efficiency of a given platform or to assess user-perceived quality-of-experience. For example, Jiang *et al.* [23] present a clustering algorithm to identify common features of subpar video sessions while Bermudez *et al.* [21] and Ager *et al.* [10] are concerned with the geo-disparity of content that is delivered to various regions, Akhtar *et al.* explore the latency of various photo CDNs from multiple vantage points. The work by Gehlen *et al.* [9] is most closely aligned with our work and utilizes passively collected packet level traces (with payload) from three ISPs to perform deep packet inspection algorithms to uncover the corresponding application (*e.g.* P2P, Web). The authors characterize the delivery strategy for web traffic, present the distribution of measured RTT values and throughput across *all* flows and then argue that the observed bimodal distribution for throughput across Akamai flows is due to the CDN's caching strategy. Cordero *et al.* [25] also employ an edge centric view of a network footprint. However, they mainly focus on the overlapping paths among incoming flows from different sources and do not offer any characteristics of traffic on a per CP basis.

Our work is also related to prior studies that examine the effects of new content delivery techniques on the AS-level hierarchy of the Internet topology. While some of this work is concerned with global implications of content delivery on the nature and volume of inter-domain traffic in the Internet [26]–[28], other work illustrates the emergence and growing importance of strategies whereby content providers deploy their own servers in third-party networks [4], [11]. In contrast,in this paper we are concerned with the implications that these strategies have on user-perceived performance at a stub-AS.

VIII. Summary

Our work contributes to the existing literature on content delivery by providing a unique view of different aspects of content delivery as experienced by the end users served by a stub-AS (*i.e.*, a Research & Education network). To this end, we examine the complete flow-level view of traffic delivered to this stub-AS from all major content providers and characterize this stub-AS' *traffic footprint* (*i.e.* a detailed assessment of the locality properties of the delivered traffic).

[5] Note that ping measures loss in both directions of a connection.

We also study the impact that this traffic footprint has on the performance experienced by its the end users and report on two main takeaways. First, this stub-AS' traffic locality is uniformly high across the main CPs; *i.e.*, the traffic that these CPs deliver to this stub-AS experiences in general only very small delays. Second, the throughput of the delivered traffic remains far below the maximum achievable throughout and is not limited by the network but rather by the front-end servers.

Lastly, to complement the effort described in this paper, assessing the locality properties of the traffic that constitutes the (long) tail of the distribution in Figure 1 and is typically delivered from source IP addresses that are rarely seen in our data or are responsible for only minuscule portions of the traffic delivered to UOnet looms as an interesting open problem and is part of future work.

IX. Acknowledgment

This material is based upon work supported by the National Science Foundation (NSF) Award NeTS-1320977. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

References

[1] Akamai, "Akamai Technologies Facts & Figures," https://www.akamai.com/us/en/about/facts-figures.jsp.

[2] Limelight, "Private global content delivery network," https://www.limelight.com/network/.

[3] EdgeConneX, "Space, power and connectivity," http://www.edgeconnex.com/company/about/.

[4] T. Böttger, F. Cuadrado, G. Tyson, I. Castro, and S. Uhlig, "Open connect everywhere: A glimpse at the internet ecosystem through the lens of the netflix cdn," *arXiv preprint arXiv:1606.05519*, 2016.

[5] M. Calder, X. Fan, Z. Hu, E. Katz-Bassett, J. Heidemann, and R. Govindan, "Mapping the expansion of google's serving infrastructure," in *Internet Measurement Conference*. ACM, 2013, pp. 313–326.

[6] V. K. Adhikari, Y. Guo, F. Hao, M. Varvello, V. Hilt, M. Steiner, and Z.-l. Zhang, "Unreeling Netflix: Understanding and Improving Multi-CDN Movie Delivery," in *INFOCOM*. IEEE, 2012, pp. 1620–1628.

[7] X. Fan, E. Katz-Bassett, and J. Heidemann, "Assessing affinity between users and cdn sites," in *Traffic Monitoring and Analysis*. Springer, 2015, pp. 95–110.

[8] R. Torres, A. Finamore, J. R. Kim, M. Mellia, M. M. Munafò, and S. Rao, "Dissecting video server selection strategies in the YouTube CDN," in *International Conference on Distributed Computing Systems*. IEEE, 2011, pp. 248–257.

[9] V. Gehlen, A. Finamore, M. Mellia, and M. M. Munafò, "Uncovering the big players of the web," in *Lecture Notes in Computer Science*. Springer, 2012, pp. 15–28.

[10] B. Ager, W. Mühlbauer, G. Smaragdakis, and S. Uhlig, "Web content cartography," in *Internet Measurement Conference*. ACM, 2011, pp. 585–600.

[11] N. Chatzis, G. Smaragdakis, J. Böttger, T. Krenc, and A. Feldmann, "On the benefits of using a large ixp as an internet vantage point," in *Internet Measurement Conference*. ACM, 2013, pp. 333–346.

[12] M. A. Sánchez, J. S. Otto, Z. S. Bischof, D. R. Choffnes, F. E. Bustamante, B. Krishnamurthy, and W. Willinger, "Dasu: Pushing experiments to the internet's edge." in *Networked Systems Design and Implementation*. USENIX, 2013, pp. 487–499.

[13] "IP Address Geolocation to Identify Website Visitor's Geographical Location," http://www.ip2location.com/.

[14] MaxMind, "Geoip: Industry leading ip intelligence," https://www.maxmind.com/.

[15] S. Triukose, Z. Wen, and M. Rabinovich, "Measuring a commercial content delivery network," in *World Wide Web*. ACM, 2011, pp. 467–476.

[16] Netflix, "Open connect appliance overview," https://openconnect.netflix.com/en/appliances-overview/.

[17] S. Sundaresan, N. Feamster, and R. Teixeira, "Measuring the performance of user traffic in home wireless networks," in *Passive and Active Measurement*. ACM, 2015, pp. 305–317.

[18] Netflix, "Internet connection speed requirements," https://help.netflix.com/en/node/306.

[19] I. Sodagar, "The mpeg-dash standard for multimedia streaming over the internet," in *IEEE MultiMedia*, 2011, pp. 62–67.

[20] M. Mathis, J. Semke, J. Mahdavi, and T. Ott, "The macroscopic behavior of the tcp congestion avoidance algorithm," in *SIGCOMM*. ACM, 1997, pp. 67–82.

[21] I. Bermudez, S. Traverso, M. Mellia, and M. Munafo, "Exploring the cloud from passive measurements: The amazon aws case," in *INFOCOM*. IEEE, 2013, pp. 230–234.

[22] P. Casas, A. D'Alconzo, P. Fiadino, A. Bär, A. Finamore, and T. Zseby, "When youtube does not work—analysis of qoe-relevant degradation in google cdn traffic," in *Transactions on Network and Service Management*. IEEE, 2014, pp. 441–457.

[23] J. Jiang, V. Sekar, I. Stoica, and H. Zhang, "Shedding light on the structure of internet video quality problems in the wild," in *CoNEXT*. ACM, 2013, pp. 357–368.

[24] Z. Akhtar, A. Hussain, E. Katz-Bassett, and R. Govindan, "Dbit: Assessing statistically significant differences in cdn performance," *Computer Networks*, vol. 107, pp. 94–103, 2016.

[25] J. A. Cordero and O. Bonaventure, "Understanding the topological properties of internet traffic: a view from the edge," in *Networking Conference, IFIP*. IEEE, 2014, pp. 1–9.

[26] C. Labovitz, S. Iekel-Johnson, D. McPherson, J. Oberheide, and F. Jahanian, "Internet inter-domain traffic," in *SIGCOMM*. ACM, 2010, pp. 75–86.

[27] Y.-C. Chiu, B. Schlinker, A. B. Radhakrishnan, E. Katz-Bassett, and R. Govindan, "Are we one hop away from a better internet?" in *Internet Measurement Conference*. ACM, 2015, pp. 523–529.

[28] A. H. Rasti, R. Rejaie, and W. Willinger, "Characterizing the global impact of p2p overlays on the as-level underlay," in *International Conference on Passive and Active Network Measurement*. Springer, 2010, pp. 1–10.