Assessing Affinity Between Users and CDN Sites

Xun Fan[1,2] Ethan Katz-Bassett[2] John Heidemann[1,2]

1: USC/Information Sciences Institute 2: USC/Computer Science Dept.
xunfan@usc.edu ethan.kb@usc.edu johnh@isi.edu

Abstract. Large web services employ CDNs to improve user performance. CDNs improve performance by serving users from nearby Front-End (FE) Clusters. They also spread users across FE Clusters when one is overloaded or unavailable and others have unused capacity. Our paper is the first to study the dynamics of the user-to-FE Cluster mapping for Google and Akamai from a large range of client prefixes. We measure how 32,000 prefixes associate with FE Clusters in their CDNs every 15 minutes for more than a month. We study geographic and latency effects of mapping changes, showing that 50–70% of prefixes switch between FE Clusters that are very distant from each other (more than 1,000 km), and that these shifts sometimes (28–40% of the time) result in large latency shifts (100 ms or more). Most prefixes see large latencies only briefly, but a few (2–5%) see high latency much of the time. We also find that many prefixes are directed to several countries over the course of a month, complicating questions of jurisdiction.

1 Introduction

Large web services serve their content from multiple sites to reduce client latency, to spread load, and to provide redundancy against failure. These services use Content Distribution Networks (CDNs) that operate *Front-End (FE) Clusters*, each consisting of multiple servers in a specific location [7, 31]. The CDN dynamically directs users to specific FE Clusters at the granularity of network prefix which Google does and perhaps so do other CDNs. The CDN may direct a user to a FE Cluster using routing (anycast with BGP) or using DNS controlled by a mapping algorithm [3, 6, 14, 28].

Ideally user prefixes might map to the nearest FE Cluster to minimize network latency. In practice, user-FE Cluster mapping is often more involved—a FE Cluster may be temporarily down, a nearby FE Cluster may be overloaded, estimates of user location may be incorrect or out-of-date, or peering costs may influence FE Cluster choice, as reported by Facebook [16].

There are several reasons users, regulators, researchers, and CDN operators should care about the dynamics of a CDN's mapping from users to FE Clusters. Users care about performance, and we show that changes in FE Cluster can result in noticeable performance differences (§ 4). Regulators and some users may care about *where* their data goes, particularly when different political jurisdictions have different requirements for privacy. Countries have different policies about censorship [29], and requirements for law enforcement access to user data vary by jurisdiction. Recent concerns about surveillance prompted countries to suggest data should be kept domestically [8]. While prior studies enumerated and geolocated CDN networks [1, 2, 15], an understanding of dynamics helps interpret such mappings. In addition, a better understanding of user-FE Cluster mapping might help CDN operators understand better how other CDNs work.

The first contribution of this paper is to provide the first evaluation of how user prefixes associate with FE Clusters of CDNs from a large number of network prefixes. We regularly collect data for the Google and Akamai CDNs from a very broad range of vantage points for an extended period—we consider over 32k user prefixes, covering 180 countries and 5158 ASes, with data every 15 minutes for four weeks (§ 3). In addition, we use 192 PlanetLab nodes to measure network and application latency of the two CDNs over one week. We find that many user prefixes experience mapping changes frequently. About 20% of Google user prefixes and 70% of Akamai user prefixes see more than 60 mapping changes (twice everyday on average) in a month (§ 4.1).

Second, we show how changes in user/FE Cluster associations may affect user performance (§ 4). We find that, over one month, most prefixes (50–70%) are redirected from one FE Cluster to another that is very distant, and that sometimes (28–40%) these shifts result in large changes in latency. These shifts are usually brief, but a few users (2–5%) receive poor performance much of the time. We also identify several reasons for these changes, including load balancing and servers being temporarily taken out of production and later restored.

Finally, we look at the geographic footprint of which FE Clusters users employ (§ 4.5). We find that many prefixes are directed to several countries over the course of a month, complicating questions of jurisdiction.

2 Background: CDNs and DNS Redirection

CDNs deploy *front-ends* around the Internet. Front-ends (FEs) are servers that users connect to request web pages or services. For our purposes, we are interested in *FE Clusters*, each of which represents the FEs in a single physical and network location that provide the same services.

Some CDNs use DNS to direct users to front-ends. When a user performs a DNS lookup for CDN-hosted content, the CDN's DNS returns IP addresses of a front-end(s) to serve that user. In practice, CDNs generally perform the same redirection for all users in a given network prefix. We call this association between network prefix and front-end the CDN's *prefix-FE Cluster mapping*. Generally, CDNs strive to map prefixes to nearby FE Clusters to reduce network latency, but the mapping may also be influenced by load, maintenance, or other factors. This paper focuses on observing the results of CDN's prefix-FE Cluster mapping; we do not attempt to reverse engineer the CDN's specific algorithm.

When a prefix p is mapped to FE Cluster A at one time, then later mapped to FE Cluster B, we call this a *prefix-FE Cluster mapping change*. We call (A, B) the *switching pair*. Our goal is to understand these mapping changes—how often do they occur, how many users change, where did they go before and after.

3 Data Collection

We measure Google and Akamai using existing methodology. Our contribution is new long-term observations and analysis of dynamics. Our datasets (Ta-

name	where used	target	coverage (prefixes)	frequ-ency	start date (length)
Google-15min-EDNS	§ 4.1 § 4.2 § 4.5	Google	32,871	15 min.	2014/03/28 (30)
Akamai-Apple-15min-ODNS	§ 4.1 § 4.2 § 4.5	Akamai	29,535	15 min.	2014/03/28 (30)
Akamai-Huff-15min-ODNS	§ 4.1 § 4.2 § 4.5	Akamai	28,308	15 min.	2014/11/17 (30)
PlanetLab-DNS-TTL	§ 4.3	both	192	20 s/5 m	2014/04/23 (7)
Google-15min-early	§ 4.4	Google	32,324	15 min.	2013/12/13 (30)
Google-location-EDNS	§ 3.3	Google	10,057,110	1 day	2014/03/28 (30)
Akamai-Apple-location-ODNS	§ 3.3	Akamai	271,357	once	2014/04/14 (-)
Akamai-Huff-location-ODNS	§ 3.3	Akamai	185,370	once	2014/11/12 (-)
ODNS-2013	§ 3	-	271,357	once	2013/10/21 (-)

Table 1: Datasets collected as part of this work.

ble 1) provide daily observations for a month from 10M prefixes, and frequent (15-minute) observations for a 30k subset of prefixes.

3.1 Enumerating CDN Front-End Servers with DNS

We focus on the Google and Akamai CDNs because they are massively distributed, host popular services, and use DNS (not anycast) to map users to FE Clusters. Following prior work, we enumerate CDN infrastructure by issuing DNS queries for a service hosted by the CDN. For Google, we query for `www.google.com`. For Akamai, we query `www.apple.com` in Akamai-Apple-15min-ODNS dataset and `www.huffingtonpost.com` in Akamai-Huff-15min-ODNS dataset. They are both static websites hosted by Akamai. We query two websites for Akamai because our initial queries for `www.apple.com`, turned out to only cover a small set of Akamai's FE Clusters while `www.huffingtonpost.com` has larger coverage. We expect our results for the specific Google and Akamai services that we study to generalize to other services they each operate that also use DNS-based redirection. Since the fundamentals of replica selection are similar, they may also apply to application-level redirection such as in YouTube and Akamai's web caching, but we do not evaluate application-level services in this paper.

To better understand prefix-FE Cluster mapping we use three techniques. We get broad coverage with both *EDNS-client-subnet* and queries through open resolvers. We get more controlled, detailed measurements from PlanetLab.

Broad probing We probe Google with the DNS EDNS-client-subnet extension, following prior work [2,24]. This approach allows one to simulate queries from any location, but while Google supports it, Akamai added support only in mid-2014, which as part-way through our study [23]. Thus we do not use it with Akamai and instead probe Akamai with open DNS resolvers to make DNS queries from around the globe, again following prior work [9,15]. Open resolvers are often in people's homes, so we use them judiciously to measure Akamai. We choose a subset of global open resolvers that we collected in 2013 (*ODNS-2013*) as the source user prefixes. It contains 32,871 open resolver IPs, each from a unique /24 prefix, and covers 180 countries/regions and 5158 ASes. We use about 32k open resolvers so that our measurement settings can finish a query in 15 minutes. To identify this subset, we start with all open resolvers and take five complete enumerations of mappings for both CDNs over two months. We then discard

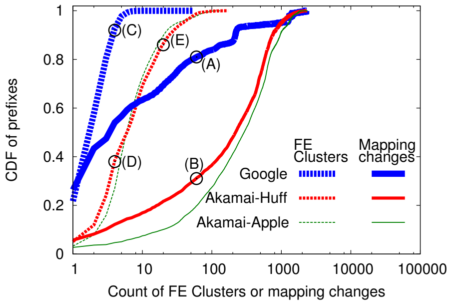


Fig. 1: Number of different FE Clusters and number of mapping changes that user prefixes seen in one month for Google and Akamai. Datasets: Google-15min-EDNS, Akamai-Apple-15min-ODNS and Akamai-Huff-15min-ODNS

	Google		**Akamai -Huff**	
Total IPs	24,150	100%	9,492	100%
Clustered	22,679	94%	8,843	93%
Un-clustered	1,471	6%	649	7%
Geolocated	22,101	92%	7,953	84%
Un-geolocated	2,049	8%	1,593	16%
Clustered and Geolocated	20,861	86%	7,953	84%
Total FE Clusters	983		1,195	

Table 2: Statistics on the number of IPs and FEs found for Google and Akamai. Datasets: Google-15min-EDNS and Akamai-Huff-15min-ODNS

those that do not respond in every trial, and finally we keep only those necessary to complete the IP-level enumeration that we saw in our five trials.

For Google, we issue DNS EDNS-client-subnet extension queries for the /24 prefixes[1] of the chosen open resolvers. Google hosts front-ends both on its backbone network and data centers (*on-net*) and in other ISPs around the world (*off-net*). We select prefixes to get broad coverage of FE Clusters, thus under-representing prefixes that are served directly from on-net FE Clusters. However, we believe our data is not drastically different from what we observe from all routable /24 prefixes, as the difference is moderate (70% of prefixes are mapped to on-net FE Clusters in our data and 88% of all routable /24 prefixes are mapped to on-net FE Clusters from Google-location-EDNS dataset). For Akamai, we probe directly to the chosen open resolvers. We probe both Google and Akamai every 15 minutes for all the 32,871 prefixes. We choose 15 minutes to limit load we impose on open resolvers.

Since open resolvers sometimes do not respond, we discard prefixes that miss more than 10% of their probes, leaving 29,535 and 28,308 prefixes in Akamai-Apple and Akamai-Huff.

Table 2 shows the total number of front-end IP addresses we find using broad probing. In total, we find 24,150 Google front-end IPs. For Akamai, we find 685 front-end IPs hosting `www.apple.com` (the *Akamai-Apple* dataset, omitted from the table for space) and 9,492 Akamai front-end IPs hosting `www.huffingtonpost.com` in 30 days (*Akamai-Huff*, shown in the table). We will see later that there are also many more FE Clusters hosting `www.huffingtonpost.com` than `www.apple.com`, and we believe this difference comes from the different SLAs used by the two sites. Compared to published reports of the sizes of the Google [2] and

[1] We always use /24 prefixes and so just write *prefix* from here.

Akamai [19] CDNs, we know that our coverage is incomplete, but we believe we cover a good part of Google's CDN (about 70% of prior results [2]). Akamai runs tens of thousands of servers; our methodology tracks only the part of that infrastructure used by our targets. We focus on specific clients hosted by Akamai so we can study user-prefix dynamics for thousands of user prefixes without creating excessive measurement traffic. We observe about three times more IPs in Google's clusters compared to Akamai's. Our methodology of sampling specific URLs means that we do not fully enumerate clusters, and load-balancing and other factors mean IP addresses do not necessarily indicate cluster size, so we focus on clusters rather than IP addresses.

Performance probing In order to also study the effects of mapping changes on user-experienced performance, we use PlanetLab to collect ping times to the front-ends and application-level page fetches, as described in § 4.3.

We also issue frequent DNS queries from PlanetLab. Following prior work [25], we probe on DNS TTL intervals (the quickest an end user might experience changes) to capture prefix-FE Cluster mapping changes. (TTL for Google DNS is 5 minutes and Akamai is 20 seconds.)

We collect our *PlanetLab-DNS-TLL* dataset using probing at these rates for 7 days. We use 192 PlanetLab nodes, each in a distinct /24 prefix.

3.2 FE Cluster Identification

Since we are interested in mapping changes between FE Clusters, not IP addresses, we use our previous technique to group IP addresses into FE Clusters based on similarity of round-trip times from PlanetLab [2]. Table 2 shows our clustering results. We find 983 FE Clusters for Google from 22,679 replying IP addresses. We were unable to cluster 1,471 Google IPs because they do not respond to the pings we need for clustering. For Akamai, we find 1,195 Akamai FE Clusters from 9,492 IP addresses in Akamai-Huff dataset, (336 Akamai FE Clusters from 650 IP addresses in Akamai-Apple, not in the table), with 649 IPs we could not cluster. We have no way of identifying, clustering, or geolocating IP addresses that do not reply to measurements, so we must discard them.

3.3 Front-End Geolocation

We geolocate FE Clusters in our datasets using our previous CCG technique (Client-Centric-Geolocation) [2]. CCG geolocates FE Clusters by averaging the locations of the prefixes they serve after aggressively removing prefixes clearly distant from the FE. From that earlier work, we have daily measurements of Google since 2013. We use one month of that data (dataset: *Google-location-EDNS*), selecting the period and subset of prefixes to match our prefix-FE Cluster mapping datasets.

We use an alternate source of data for geolocation since Akamai did not support EDNS-client-subnet queries when our measurements began (§ 3.1). We collect data from open resolvers and apply the CCG algorithm to it ourselves. We use the whole set of open resolvers (ODNS-2013) we collected in 2013 as clients for CCG. The set of open resolver contain 600,000 open resolver IP addresses from 271,357 distinct /24 prefixes, covering 217 countries/regions and 11,793

ASes. Since it covers a fraction of the 10 million total routable /24 prefixes, we validate the use of CCG with open resolvers and find that it provides similar accuracy to CCG with all routable /24 IP prefixes. Our geolocation is accurate, with 90% of IP addresses having distance error within 500km [10].

CCG does not provide locations for 8% of Google IP addresses and about 16% of Akamai IPs (Table 2). Typically, CCG fails for FE Clusters that see an insufficient number of clients, so these servers may be relatively unimportant.

4 Dynamics of User Redirection

4.1 Are user prefixes mapped to different FE Clusters?

We first examine how many mapping changes and how many FE Clusters each user prefix observes over one month. Figure 1 shows the cumulative distribution. We see that 20% and 70% of prefixes observe more than 60 mapping changes ((A) and (B) in Figure 1) in a month (average 2 a day) for Google and Akamai respectively, suggesting mapping changes are common for many prefixes. (The number of changes we report here is much smaller than prior work [25] because we report the changes between clusters, not just IP addresses.) In addition, we see that most user prefixes have fairly stable mappings for Google, with 92% of them being mapped to at most 4 FE Clusters ((C) in Figure 1). Akamai user prefixes seem to experience more variation, with only around 40% being mapped to 4 FE Clusters or fewer and 14% being mapped to 20 or more FE Clusters ((D) and (E) in Figure 1). This analysis shows that *mappings changes are common*, with some users changing frequently and most occasionally.

4.2 Distances of Mapping Changes

We next examine the distance between the FE Clusters that users switch between. We expect that a user would see little latency change when switched between nearby FE Clusters, while mapping changes between very distant FE Clusters are more likely to lead to large latency change. Unless the client is equidistant between the old and new FE Clusters, a large change in FE distance suggests a non-optimal choice of a FE.

We measure distance between the switching pair of a prefix-FE Cluster mapping change. We randomly choose an observation time t, then find the switching pair of the next mapping change (A, B) for each prefix after time t. We then plot the CDF of distance between A and B over all prefixes. We see nearly identical distributions after three trials and so report one case as representative.

Figure 2 shows the CDF of the distance between the switching pair for all prefixes over one randomly chosen observation times for Google and Akamai. While some prefixes switch between FE Clusters that are near each other (about 26–33% are within 100 km), many prefixes change between FE Clusters that are far apart. More than 50% Google changes and 30% of Akamai changes move between switching pairs more than 1000 km apart.

Long-distance remapping: akamai When measured at a random time We see that many prefixes change between FE Clusters that are distant from each other. We next consider this question for *every* time over a month. Figure 3 plots the distribution of the maximum distance of switching pairs seen by every prefix

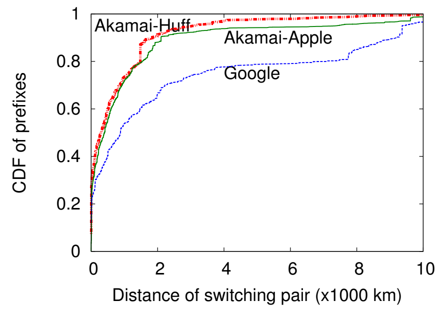


Fig. 2: CDF of distance of switching pairs over all prefixes after a random observation time t.

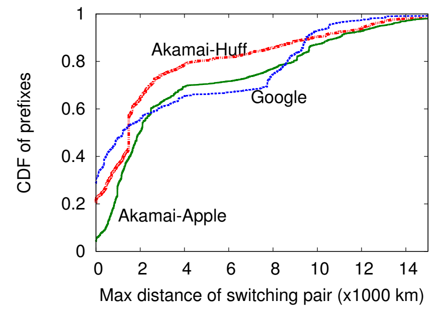


Fig. 3: CDF of maximum distance of switching pair seen in one month over all prefixes.

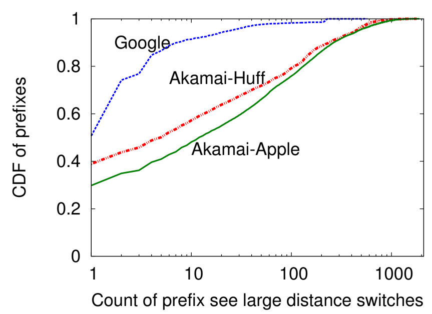


Fig. 4: CDF of the number of times in a month each prefix sees mapping changes of large distances (more than 1000km).

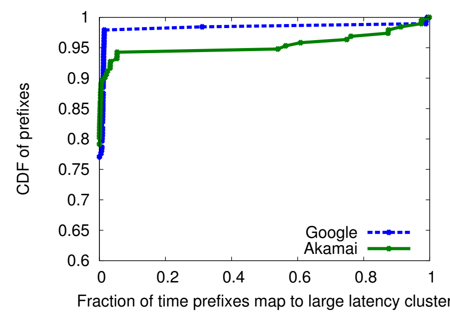


Fig. 5: CDF of fraction of time user prefixes spend on a FE Cluster with large latency (where page-fetch time is 100 ms worse than in the prior/next mapping).

in one month. Many prefixes experience long-distance changes. For example, 50% of prefixes switch between Google FE Clusters that are at least 1000 km apart, and 60-70% experience such a switch for Akamai servers. Figure 4 shows the distribution of the number of times prefixes experience large distance switching pairs. We see that a few Google prefixes (9%) and many Akamai prefixes (40-50%) move large distances (1000 km) more than 10 times in a single month, suggesting it's not rare for these long distance re-mappings to happen. In § 4.4 we explore reasons why these changes may occur.

4.3 Effects of Mapping Changes on Users

To understand how changes to prefix-FE Cluster mappings affect users, we consider when mapping changes affects (or does not affect) user latency.

Large Distance Leads to Larger Latency While § 4.2 showed that users are sometimes mapped to FE Clusters in very different places, it does not directly measure performance. While a prefix equidistant between two FE Clusters may see similar performance from both, in most cases we expect that a prefix that is redirected to a very different place will see different user-visible performance.

Here we study measurements taken from 192 prefixes hosting PlanetLab sites since evaluating user performance requires measurements taken from inside each prefix. Although these sites are only a small subset, we verified that they generally are representative of our measurements with 32,871 prefixes [10].

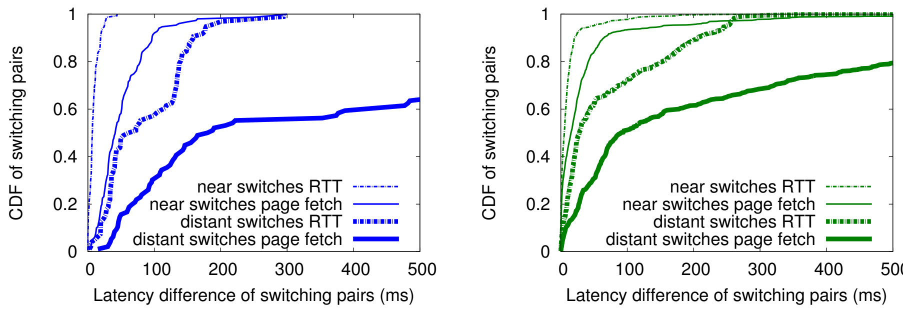
We assess user performance by measuring network latency and application performance. We measure *network* and *application* latencies every DNS TTL, and also immediately after we observe a prefix p has changed its mapping from FE Cluster A to B (prior work measured latency [25,27], but not around mapping changes). We measure network latency with ICMP echo request (ping), observing $RTT_{p,A}$ and $RTT_{p,B}$. We measure application latency by fetching a web page to observe $PFT_{p,A}$ and $PFT_{p,B}$. To avoid noise in individual observations, each observation uses two pings and one page fetch, and analysis uses the second smallest of the 10 most recent observations. For Google we fetch a 75 kB web page corresponding to a search for "USA" (`http://www.google.com/search?q=USA`). For Akamai we fetch the 9.5 kB home page of Apple (`http://www.apple.com`). We then evaluate the absolute value of the difference of these metrics: $RTT^{\delta}_{p,A,B} = |RTT_{p,A} - RTT_{p,B}|$ and $PFT^{\delta}_{p,A,B} = |PFT_{p,A} - PFT_{p,B}|$. We use absolute value to judge overall changes, since data shows that at steady state, mapping changes generally alternate between nearer to further FE Clusters.

For each prefix, we evaluate all mapping changes over the entire measurement period, giving a set of observations of many $RTT^{\delta}_{p,A,B}$ and $PFT^{\delta}_{p,A,B}$. Since changes are generally symmetric, we merge the (A, B) and (B, A) directions and take the median value of all observations to get $RTT^{m\delta}_{p,A,B}$ and $PFT^{m\delta}_{p,A,B}$. Finally, to understand if large distance switches affect performance, we divide observations into *distant switches*, where A and B are 1000 km apart or more, and *near switches* where they are less than 1000 km. We then plot the CDF of $RTT^{m\delta}$ and $PFT^{m\delta}$ for each group.

Figure 6 shows results for Google and Akamai. We first see that the switches between distant FE Clusters (the wider, right-most lines) show much greater performance changes than switches between nearby ones (the thinner, left lines). For Google, near switches show smaller performance changes ($RTT^{m\delta} < 50$ ms and $PFT^{m\delta} < 150$ ms), while for distant switches group, more than 40% have changes more than twice that ($RTT^{m\delta} > 100$ ms and $PFT^{m\delta} > 400$ ms). The results of Akamai are similar, with only 2% of near switches showing $RTT^{m\delta} >$ 100 ms, while the number is 28% for distant switches.

To summarize, prefixes that switch between FE Clusters that are far apart tend to also observe large network and page-fetch latency changes.

How Long Do Prefixes Stay On Non-Optimal FE Clusters? Fortunately, we next show that switches that increase user latency are usually brief for most prefixes. We analyze our PlanetLab data to see what fraction of time user prefixes spend in a mapping that has large latency (for this subset of data). We focus on *distant* switching pairs, those with distance larger than 1000 km, and of these, those with *long* differences in page-fetch times ($PFT^{m\delta} > 100$ ms). The



(a) Google

(b) Akamai

Fig. 6: Prefix-FE Cluster latency changes after a mapping change, measured by RTT (dashes) and page fetch time (solid). Left line are near switches, right line are distant switches. (Dataset: PlanetLab-DNS-TTL)

resulting subset are all prefixes with large distance switches that raise application latency. Finally, we look at how long each prefix remained at the larger-latency FE Cluster, computing the fraction of observations the prefix spent there.

Figure 5 shows the CDF of fraction of time user prefixes spend on FE Clusters with large latency (where page-fetch time is 100 ms worse than in the prior mapping). Most of these FE Clusters are used only briefly (97% of Google and 93% of Akamai prefixes spend less than 5% of their time at FE Clusters with high application latency). But the tail is long, with 2% of Google and 5% of Akamai prefixes spending more than 60% of time on distant FE Clusters and seeing higher application latencies, even though lower-latency FE Clusters exist.

4.4 Reasons for Mapping Changes

We have shown that mapping changes are common. We next evaluate *why* they occur. Although we cannot categorize every change, we see three general reasons: FE Clusters drain and restore (that is, temporarily shut down), load balancing, user-to-FE Clusters mapping reconfiguration. We cannot completely separate these categories without inside knowledge of each CDN. However, our external observations provide some evidence of each.

FE Clusters Drain and Restoration CDN sometimes *drain* some of their FE Clusters, assigning no user prefixes to them, in order to, for example, perform maintenance or troubleshoot problems. For example, Facebook recently drained an entire datacenter as part of an infrastructure stress test [30]. As an example drain event, Figure 7 shows the number of active FE Clusters in Google over our Google-15min-EDNS dataset. We see a large drop around April 23rd (from 900 to 60 FE Clusters). Examination of the clusters before and after the drop shows that Google stopped directing clients to all FE Clusters not in Google ASes (the off-net FE Clusters). They restored broader service, then shut off-net FE Clusters again on April 28th.

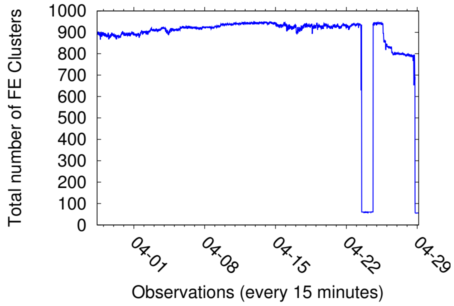


Fig. 7: Total number of Google FE Clusters seen from all prefixes at each observation.

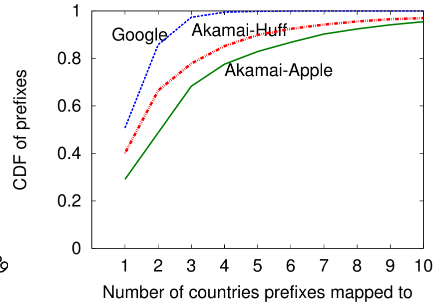


Fig. 8: CDF of the number of different countries to which prefixes are mapped.

We checked if these drains biased our previous observations (§ 4.2 and § 4.3). To do so, we re-examined the distance user prefixes switched with and without these days where all off-net FE Clusters drained. We confirmed that overall changes are small, meaning regular changes in mapping dominate our results.

Load Balancing We observe two patterns of behavior that we believe are due to load-balancing of user prefixes across multiple FE Clusters. First, we sometimes see some prefixes (about 10% for Google and 30% for Akamai) switch between two FE Clusters quite frequently (on average every hour). We sample 10 prefixes from each of these groups, and for each prefix, both FE Clusters they switched between are close to each other (within 200 km). This behavior may indicate that the CDN is spreading the load between FE Clusters at two different PoPs. Second, we see that a few Google FE Clusters (about 10 of 900) display diurnal patterns (as seen in spectral analysis [21]), suggesting some load balancing due to changes in diurnal traffic patterns.

Reconfiguration of User-to-FE Clusters Mapping Both Google and Akamai strive to optimize performance for users by associating prefixes with nearby FE Clusters [7, 18]. Long-term shifts in routing, user population, and FE Cluster deployments may shift this mapping as the CDN re-optimizes. In early data (Google-15min-early dataset), we saw that Google would occasionally shift one-third of user prefixes at the same time [31]. These bulk shifts have diminished in recent observations of Google and never appeared in Akamai, but both CDNs currently have a few percent of user prefixes that have stable mappings for weeks.

Changes also happen at short timescales—Facebook reconfigures their mapping over the course of a day due to changes in observed client latency [16]. We know that Google and Akamai also have short-term changes, but we do not know if they are responses to changes in user latency or responses to changes due to their CDN infrastructure, such as load balancing.

Unknown We also observe some mapping changes that are not explained by the above reasons. For example, we see Google sometimes map prefixes to very distant Google FE Clusters (across continents) for a single observation.

source country	Google %	non-domestic 1st	2nd	3rd	Akamai-Huff %	non-domestic 1st	2nd	3rd
us (United States)	11%	be (4%)	nl (4%)	de (3%)	98%	ca (38%)	gb (27%)	fr (27%)
kr (S. Korea)	97%	jp (58%)	us (19%)	cn (18%)	99%	tw (99%)	jp (6%)	nl (3%)
ru (Russia)	99%	us (35%)	be (6%)	nl (5%)	96%	se (74%)	no (43%)	de (40)
jp (Japan)	55%	us (30%)	nl (9%)	be (7%)	100%	cn (92%)	us (67%)	vn (9%)
br (Brazil)	48%	nl (18%)	be (17%)	us (14%)	83%	us (78%)	cl (53%)	ar (35%)
tw (Taiwan)	45%	us (24%)	be (9%)	nl (9%)	99%	cn (74%)	us (72%)	vn (48%)
cn (China)	51%	us (27%)	nl (11%)	be (11%)	99%	jp (93%)	us (89%)	gb (67%)
it (Italy)	60%	us (40%)	de (19%)	fr (5%)	–	–	–	–
gb (U. Kingdom)	54%	us (40%)	nl (19%)	be (8%)	–	–	–	–
au (Australia)	52%	us (24%)	nl (18%)	be (11%)	–	–	–	–
hk (Hong Kong)	–	–	–	–	90%	cn (88%)	jp (25%)	vn (12%)
tr (Turkey)	–	–	–	–	91%	it (82%)	se (46%)	de (23%)
fr (France)	–	–	–	–	99%	pl (69%)	gb (57%)	es (56%)

Table 3: Top 10 source countries (with ISO country codes) and their percentage of prefixes that had been mapped to FE Clusters in other countries, and the to three non-domestic countries serving them. Datasets: Google-15min-EDNS and Akamai-Huff-15min-ODNS

4.5 Geographic Footprint Seen by User Prefixes

Prefix-FE Cluster mapping changes across long distances, suggesting that users may see FE Clusters in different countries.[2] For some users, traffic leaving a given country may raise concerns about privacy or legal jurisdiction. We next show that some prefixes in many countries are often mapped abroad.

First, we assess how many countries each prefix is mapped to over the course of a month in Figure 8. We see that *more than half* of prefixes are mapped to different countries over time (50% for Google, and 60–70% for Akamai). It is common for a user to be served from multiple countries. We caution that this result reflects two biases in our data: first, our prefix selection under-representing prefixes that are served directly from the provider, as described in § 3.1. Second, because of cluster drain (§ 4.4), we expect many prefixes to shift from off-net FE Clusters, present in many countries, to on-net FE Clusters that operate in only a few countries.

We next consider from where prefixes are served. For each service we select the 10 countries that originate the most user prefixes, then identify from where they are served. (We exclude prefixes that are never served domestically on the assumption that they have no local option or that our geolocation is wrong.) For each country we consider two questions: what portion of prefixes leave the country? Where does their traffic go?

Table 3 shows the results for Google and Akamai. (The top countries differ because the CDNs are different.) For each country, the first column shows how many of that country's prefixes that are sometimes mapped outside its borders. The following three columns show which other countries most often provide service. For Akamai, we show only Akamai-Huff data here for space; we show Akamai-Apple data in [10] and summarize any differences here.

We see that all prefixes but U.S.-ones have many non-domestic mappings—around 50% of user prefixes for Google and more than 90% for Akamai. We see

[2] We use the term country generically, sometimes considering smaller or larger regions.

that Google often serves from the U.S., Belgium and Netherlands, perhaps those countreis have good connectivity and host Google datacenters [22]. For Akamai, we see that U.S. FE Clusters serve prefixes from other countries, perhaps because of good U.S. connectivity. Akamai-Huff selection (and also Akamai-Apple) shows a stronger geographic locality than Google, with French and Turkish prefixes remaining in Europe and Hong Kong prefixes in Asia. Surprisingly, most Chinese prefixes are sent abroad in both Akamai datasets.

Both Google and Akamai often map prefixes outside their originating country. Countries that have expressed privacy concerns, such as Brazil [8], or regions with strict privacy laws, such as the European Union, may find traffic leaving their legal jurisdiction weakens their ability to implement some policies. For example, Brazil's exact set of foreign countries varies depending on CDN or service, but in all cases their prefixes are served outside Brazil. In other cases, prefixes in some countries find services in others that have strict limits on domestic handling of some topics. Examples include South Korea and Japan receiving service from China (with limits on Chinese politics), and in Akamai-Apple data where Brazil served from Germany (with limits on Nazi politics). While such issues may not be a concern for Apple or Huffingtonpost's home page, it may be for other services using these CDNs.

5 Related Work

Prior work compared the performance of CDN-selected front-end servers and other servers of the same CDN [17, 20, 25, 27]. Su *et al.* use Akamai's choice of server location to influence their selection to leverage Akamai's network measurements [25]. Triukose *et al.* compare the page download performance difference between Akamai selected server with 80 other randomly selected Akamai servers to study if CDNs enhance performance [27]. Krishnamurthy *et al.* study CDN DNS load balancing performance by using two dozen clients to detect DNS load balancing every 30 minutes and performing file download when observing CDN server changes [17]. Otto *et al.* compare HTTP latency between CDN servers returned by different DNS servers to measure the impact of using remote DNS on CDN performance [20]. Our work differs from this prior work by exploring how CDNs *change* their prefix-FE Cluster mappings over time, and how these changes affect network and application latency for users.

The Ono system uses large set of clients (120,000) to study affinity between users and CDN servers [5]. They use this information to help peer selection in peer-to-peer networks to reduce cross-ISP traffic. Our work also uses a large set of client prefixes to assess user-to-CDN affinity, but we focus on understanding the properties of prefix-FE Cluster mapping changes and their potential impact on both users and previous CDN studies.

Huang *et al.* studied the cache dynamics from users to Facebook Edge Caches as viewed from within Facebook [16]. Facebook optimizes to balance latency, server load, and peering cost, sometimes directed users to caches that are not physically nearest. Our paper complement theirs by looking from the user side.

Torres *et al.* studied mechanism and policy of user to content server mapping of Youtube using video flow data collect from 5 distinct locations over a week [26].

They Geolocate Youtube datacenters using CBG and find that non-negligible fraction of traffic are provided by *non-preferred* datacenter. They find that the reasons of non-preferred datacenter access include load balancing, DNS server variations, limited availability of rarely accessed videos and alleviating hot-spot due to popular videos. Our work differs from theirs by focusing on the effects of user to FE Cluster mapping changes on users, while they focus on understanding the mapping dynamics themselves. We also have a broader coverage on user prefixes and CDN FE Clusters while theirs is deeper from a few vantage points.

Cases *et al.* [4] and Finamore *et al.* [13] each study associations between web services, hosting organizations, content-server IPs, and service provisioning. They use min-RTT estimates to cluster IPs to datacenters. They use measurements from one ISP and observe user/datacenter switches suggesting load balancing. We also cluster IPs to datacenters, but with many vantage points [2]. Both their work and ours identifies load balancing and mapping changes, but they apply their work to provisioning while we study its effects on end-users.

Fiadino *et al.* use a month of HTTP flow data collected from a major European ISP to study the traffic anomaly caused by cache selection dynamics and the impacts on both ISP and users [11, 12]. They found Facebook traffic anomaly by identify large amount of flow shift from Akamai to other hosting organization of Facebook. They report the anomaly may increase the transit cost of the users' ISP. They also found Youtube traffic anomaly that shift traffic to different set of /24 subnets of Youtube and found that the shift affect user experienced throughput. Our work differs from them in following ways. First, the methodologies are quite different. They detect synchronized mapping changes for particular web services by watching for large shifts in flow volumes, while we directly measure target FE Clusters with EDNS-client-subnet and direct DNS queries. Their approach is ideal for studying a single ISP when traffic is available, but the second difference is that our approach allows us to provide much broader coverage. We examine 32k user prefixes from hundreds of countries and ASes, while their study focuses only on users of a single ISP. Last, we study how often users traffic changes countries.

6 Conclusions

This work provides the first evaluation of the dynamics of CDN redirection of user's network prefixes to Front-End Clusters from a large range of prefixes. We gather new data about Google and Akamai, and we find that some prefixes switch between FE Clusters that are long distances apart, often seeing large changes in latency and application-level performance. While most of prefixes only stay shortly on FE Clusters that have large application level latency, a few percent of prefixes are mapped to those FE Clusters much of the time. We also find that many user prefixes are directed to multiple countries in a month, complicating questions of jurisdiction.

Acknowledgments and Data Availability: Our data is publicly available at http://www.isi.edu/ant/traces/mapping_cdns/. This work was identified by the USC IRB (IIR00001412, March 2013) as non-human subject research. We thank Matt Calder for his assistance with CCG.

This research is partially sponsored by the Department of Homeland Security (DHS) Science and Technology Directorate, HSARPA, Cyber Security Division, BAA 11-01-RIKA and Air Force Research Laboratory, Information Directorate under agreement number FA8750-12-2-0344, NSF CNS-1351100, and via SPAWAR Systems Center Pacific under Contract No. N66001-13-C-3001. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views contained herein are those of the authors and do not necessarily represent those of DHS or the U.S. Government.

References

1. B. Ager and et al. Web content cartography. In *ACM IMC*, 2011.
2. M. Calder, X. Fan, Z. Hu, E. Katz-Bassett, J. Heidemann, and R. Govindan. Mapping the Expansion of Google's Serving Infrastructure. In *IMC*, Oct. 2013.
3. R. L. Carter and M. E. Crovella. Server selection using dynamic path characterization in wide-area networks. In *IEEE INFOCOM*, Apr. 1997.
4. P. Casas, P. Fiadino, and A. Bar. Ip mining: Extracting knowledge from the dynamics of the internet addressing space. In *ITC*, 2013.
5. D. Choffnes and F. E. Bustamante. Taming the torrent: A practical approach to reducing cross-ISP traffic in peer-to-peer systems. In *ACM SIGCOMM*, 2008.
6. M. E. Crovella and R. L. Carter. Dynamic server selection in the Internet. In *IEEE HPCS*, Aug. 1995.
7. J. Dilley, B. Maggs, J. Parikh, H. Prokop, R. Sitaraman, and B. Weihl. Globally distributed content delivery. *Internet Computing, IEEE*, 6(5):50–58, 2002.
8. A. Edgerton. NSA Spying Allegations Put Google on Hot Seat in Brazil. http://www.businessweek.com/news/2013-10-28/nsa-spying-allegations-put-google-on-hot-seat-corporate-brazil, 2013.
9. X. Fan, J. Heidemann, and R. Govindan. Evaluating anycast in the domain name system. In *IEEE INFOCOM*, 2013.
10. X. Fan, E. Katz-Bassett, and J. Heidemann. Assessing affinity between users and CDN sites (extended). http://www.isi.edu/~xunfan/affinity_tech_report.pdf.
11. P. Fiadino, A. D'Alconzo, A. Bar, A. Finamore, and P. Casas. On the detection of network traffic anomalies in content delivery network services. In *ITC*, 2014.
12. P. Fiadino, A. D'Alconzo, and P. Casas. Characterizing web services provisioning via cdns: The case of Facebook. In *TRAC*, 2014.
13. A. Finamore, V. Gehlen, M. Mellia, M. Munafò, and S. Nicolini. The need for an intelligent measurement plane: The example of time-variant cdn policies. *IEEE NETWORKS*, 2012.
14. J. D. Guyton and M. F. Schwartz. Locating nearby copies of replicated internet servers. In *ACM SIGCOMM*, pages 288–298, Aug. 1995.
15. C. Huang, A. Wang, J. Li, and K. W. Ross. Measuring and evaluating large-scale CDNs. Technical Report MSR-TR-2008-106, Microsoft Research, Oct. 2008.
16. Q. Huang, K. Birman, R. van Renesse, W. Lloyd, S. Kumar, and H. C. Li. An analysis of Facebook photo caching. In *ACM SOSP*, 2013.
17. B. Krishnamurthy, C. Wills, and Y. Zhang. On the use and performance of content distribution networks. In *ACM IMW*, pages 169–182, 2001.
18. R. Krishnan and et al. Moving beyond end-to-end path information to optimize CDN performance. In *ACM IMC*, 2009.
19. M. Mao and et al. Peer-assisted content distribution in akamai netsession. In *ACM IMC*, pages 31–42, 2013.
20. J. S. Otto and et al. Content delivery and the natural evolution of dns: remote dns trends, performance issues and alternative solutions. In *ACM IMC*, 2012.
21. L. Quan, J. Heidemann, and Y. Pradkin. When the Internet sleeps: Correlating diurnal networks with external factors. In *ACM IMC*, 2014.
22. F. Robinson. Google Sets Big Belgian Investment. http://blogs.wsj.com/brussels/2013/04/10/google-sets-big-belgian-investment/, Apr. 2013.
23. Stacey Higginbotham. Akamai signs deal with opendns to make the web faster. http://gigaom.com/2014/06/03/akamai-signs-deal-with-opendns-to-make-the-web-faster/.
24. F. Streibelt, J. Böttger, N. Chatzis, G. Smaragdakis, and A. Feldmann. Exploring EDNS-client-subnet adopters in your free time. In *ACM IMC*, 2013.
25. A.-J. Su, D. R. Choffnes, A. Kuzmanovic, and F. E. Bustamante. Drafting behind Akamai (Travelocity-based detouring). In *ACM SIGCOMM*, 2006.
26. R. Torres, A. Finamore, J. R. Kim, M. Mellia, M. M. Munafo, and S. Rao. Dissecting video server selection strategies in the Youtube CDN. In *ICDCS*, 2011.
27. S. Triukose, Z. Wen, and M. Rabinovich. Measuring a commercial content delivery network. In *ACM WWW*, pages 467–476, 2011.
28. P. Wendell, J. W. Jiang, M. J. Freedman, and J. Rexford. DONAR: Decentralized server selection for cloud services. In *ACM SIGCOMM*, Aug. 2010.
29. Wikipedia. Internet censorship by country. http://en.wikipedia.org/wiki/Internet_censorship_by_country.
30. Yevgeniy Sverdlik. Facebook turned off entire data center to test resiliency. http://www.datacenterknowledge.com/archives/2014/09/15/facebook-turned-off-entire-data-center-to-test-resiliency/.
31. Y. Zhu, B. Helsley, J. Rexford, A. Siganporia, and S. Srinivasan. LatLong: Diagnosing wide-area latency changes for CDNs. *IEEE TNSM*, 9(1), Sept. 2012.