How Far Is Facebook from Me? Facebook Network Infrastructure Analysis

Reza Farahbakhsh, Angel Cuevas, Antonio M. Ortiz, Xiao Han, and Noel Crespi

ABSTRACT

Facebook is today the most popular social network with more than one billion subscribers worldwide. To provide good quality of service (e.g., low access delay) to their clients, FB relies on Akamai, which provides a worldwide content distribution network with a large number of edge servers that are much closer to FB subscribers. In this article we aim to depict a global picture of the current FB network infrastructure deployment taking into account both native FB servers and Akamai nodes. Toward this end, we have performed a measurement-based analysis during a period of two weeks using 463 PlanetLab nodes distributed across 41 countries. Based on the obtained data we compare the average access delay that nodes in different countries experience accessing both native FB servers and Akamai nodes. In addition, we obtain a wide view of the deployment of Akamai nodes serving FB users worldwide. Finally, we analyze the geographical coverage of those nodes, and demonstrate that in most of the cases Akamai nodes located in a particular country service not only local FB subscribers, but also FB users located in nearby countries.

INTRODUCTION

Facebook (FB) is the most popular online social network (OSN) with more than 1 billion subscribers all over the world. According to Alexa Ranking,[1] FB is the second most popular website in the world. A system of that dimension needs to be sustained by a robust and reliable architecture. Toward this end, FB owns and manages a number of centralized data centers located in the United States and Ireland [1]. However, those data centers are far from a large number of FB subscribers, who could incur very high delays to reach them. Access delay is a very sensitive parameter that impacts user experience and may have a very negative effect on online services if it is not bounded. Some illustrative examples of the actual relevance of delay reported in [2] are:

- 100 ms delay increment implies 1 percent sales loss for Amazon.
- An extra latency of 400 ms reduces Google search volume by 0.74 percent.
- 500 ms of delay decrements the revenue per user in Bing by 1.2 percent.

These numbers state that the lower the delay, the better the quality of experience of the users.

Therefore, to provide efficient service, a worldwide popular system like FB needs to rely on a distributed infrastructure that provides subscribers good quality of service (e.g., low access delay). To achieve this goal FB uses Akamai,[2] a content distribution network (CDN) with 170,000 servers deployed in 102 countries, which delivers between 15–30 percent of all web traffic.[3]

In this context, an intriguing question that motivates our research is how this complex infrastructure offers FB services[4] to FB subscribers, and whether all countries experience the same quality of service in terms of their delay in accessing those services. The goal of this article is to present a rigorous measurement study that allows us to construct the actual FB infrastructure (including Akamai servers) and see how it is being used to meet subscribers' demand.

To answer the previous question, it is essential to determine how the Akamai servers that offer FB services are distributed around the world, and to which Akamai locations FB subscribers are redirected when they access a particular service. Toward this end, we followed a systematic methodology that allows us to identify which Akamai servers are offering what FB services as well as geolocating them. This methodology is composed of four basic steps:

- Identify the URLs associated with FB services.
- Execute ping and traceroute commands from edge machines distributed worldwide to extract IP addresses associated with servers attending queries related to the discovered FB services.
- Geolocate those IPs and determine which ones are associated with native FB servers and which ones belong to Akamai servers.

Reza Farahabkhsh, Xio Han, and Noël Crespi are with Institut Mines-Telecom, Telecom Sud-Paris.

Ángel Cuevas is with Universidad Carlos III de Madrid and Institut Mines-Telecom, Telecom Sud-Paris.

Antonio M. Ortiz is with Montimage and Institut Mines-Telecom, Telecom Sud-Paris.

[1] http://www.alexa.com/topsites

[2] http://www.akamai.com

[3] Akamai Facts & Figures, 2014, www.akamai.com/html/about/facts_figures.html

[4] We refer to any activity that a regular FB subscriber can perform when she/he is connected to FB as an FB service, including visualizing pictures, watching videos, gaming, chatting, and so on.

 0163-6804/15/$25.00 © 2015 IEEE

- Determine which source nodes (in which locations) are assisted by which Akamai servers.

To apply this methodology we used 463 PlanetLab (PL) [3] nodes distributed across 41 countries all over the world, which sent ping and traceroute probes to 47 different FB URLs (grouped into 16 different service categories) six times a day for two weeks, from May 7 to May 21, 2013. Overall we collected almost 2 million delay samples from PL nodes to FB native servers and Akamai nodes.

Based on the results obtained from our measurements, we present a discussion that mainly covers two aspects:

- The quality of service (in terms of delay) experienced by subscribers depending on their location
- The picture of where Akamai nodes offering access to FB services are located and which geographical areas they cover (i.e., whether an Akamai node located in country A only receives queries from nodes located in that country or if it also serves nodes in other countries, and in such cases whether these are neighboring countries or not)

The results of our research serve as a solid benchmark to understand the performance offered by CDNs to large demanding clients with hundred of millions of subscribers distributed all over the world. Therefore, researchers aiming to improve CDN services could use the results presented in this article to validate their solutions with respect to the performance offered by the largest commercial CDN. In addition, it opens a door to the networking community to analyze what are the main sources of delay in order to propose solutions that minimize end users' access delay to services like FB. Finally, the simple but efficient methodology employed in the article can be replicated with other online sites and CDNs to perform comparative analysis to our work.

Service category	#URLs	Service provider
Access website	2	Facebook and Akamai
Authentication	4	Facebook and Akamai
Blog site	1	Facebook
Chat	2	Facebook
Developer site	1	Facebook
Error	1	Facebook
Friend finder	1	Akamai
Friend site	1	Facebook
Game applications	3	Facebook
Group site	1	Facebook
Multiple services	4	Facebook and Akamai
News feed	4	Facebook
Photo upload	1	Facebook
Photo view	19	Facebook and Akamai
Post site	1	Facebook
Video view	1	Akamai

Table 1. Facebook service categories, number of URLs for each service, and the service provider (Facebook and/or Akamai).

FB owns and manages a number of centralized data centers located in the United States and Ireland [1]. However, those data centers are far from a large number of FB subscribers, who could incur very high delays to reach them. Access delay is a very sensitive parameter that impacts user experience and may have a very negative effect.

Methodology

The goal of this article is twofold: to analyze the user experience in accessing FB services from different countries in terms of latency, and to describe a geographical picture for the location of those servers (with a special focus on Akamai nodes) offering FB services, and, linked to that location, whether they only cover a local region or also serve users located in different countries. Toward this end, we have employed a simple yet meaningful methodology that could be replicated to evaluate the performance in terms of the access delay a CDN offers to a particular website. Next, we define in detail the steps followed in our methodology.

Step 1. Identify URLs Associated with the Service Offered by the Website (i.e., FB): We asked several Facebook subscribers to perform a number of activities in FB such as login to the site, access their profiles, access photos and videos, and access friends' content. In parallel, we used a network protocol analyzer tool [4] that collected all the traffic associated with each of the described actions. After a simple filtering of the network traces we could map each FB action to one (or more) URLs that could refer to either an FB native server (e.g., profile.facebook.com) or an Akamai server (e.g. photos-a.ak.fbcdn.net). We identified 47 URLs that correspond to 16 different FB services. To be sure that the URLs were not location-dependent we repeated this exercise on several machines at different geographical locations leading to the same results. Table 1 shows the 16 identified FB service categories included in this study as well as the information on which service provider, Akamai and/or FB, is in charge of replying to the queries for these services.

Step 2. Script to Measure Access Delay and Network Path to the URLs: We implemented a simple script, following a standard discovery method [5], that executes ping and traceroute operations from the machine where it is executed to all 47 identified URLs. The ping measures the latency from the source node to the queried server, which served us to evaluate the performance in terms of access delay. The traceroute reports the intermediate hops between the

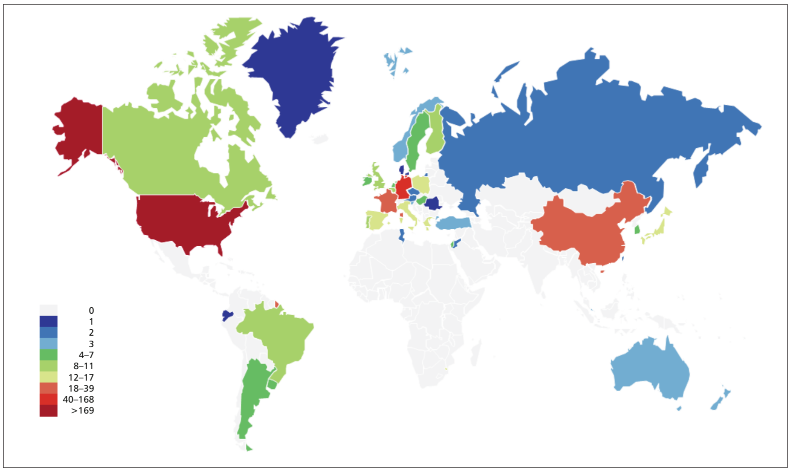

Figure 1. Presence and distribution of the 463 PlanetLab nodes (PL_node) per country.

Country	Acr.	#PL_node	Country	Acr.	#PL_node
United States	US	169	Argentina	AR	4
Germany	DE	40	Hungary	HU	4
China	CN	19	Korea, Rep.	KR	4
France	FR	18	Netherlands	NL	4
Italy	IT	16	Australia	AT	3
Poland	PL	16	New Zealand	NZ	3
Spain	ES	16	Norway	NO	3
Greece	GR	12	Singapore	SG	3
Japan	JP	12	Slovenia	SI	3
Switzerland	SZ	12	Turkey	TR	3
Canada	CA	11	Austria	AT	2
United Kingdom	UK	11	Czech Rep.	CZ	2
Belgium	BE	9	Jordan	JO	2
Brazil	BR	8	Puerto Rico	PR	2
Finland	FI	8	Russia	RU	2
Portugal	PT	8	Taiwan	TW	2
Israel	IL	6	Tunisia	TN	2
Sweden	SE	6	Denmark	DK	1
Hong Kong	HK	5	Ecuador	EC	1
Ireland	IE	5	Romania	RO	1
Uruguay	UY	5			

Table 2. Distribution of the 463 PlanetLab nodes (PL_node) per country.

source node and the server, and the delay to each hop (in case the intermediate router accepts ICMP traffic). The traceroute results may serve to dig into the particular reasons why a particular source node-server path is incurring unexpected delays and try to identify the elements in the paths leading to that situation. However, that individualized analysis goes beyond the scope of this article and would require an article itself.

Step 3. Create a Distributed Infrastructure to Obtain Comprehensive Results from Different Geographic Locations: The goal of this research required measuring access delay to the servers serving the 47 URLs from a large number of source machines distributed all over the world. For this purpose we relied on PL [3]. In particular, we distributed the script described in step 2 across 463 PL nodes located in 41 different countries (Fig. 1) as shown in Table 2. In addition, in order to have a large enough and robust dataset that avoids eventual network effects which could corrupt the average delay results, we ran the script six times a day (every four hours at the same time across all machines) in each PL node during a period of two weeks from May 7 to May 21, 2013. Our dataset contains more than 2 million ping and traceroute probes.

Step 4. Source Nodes, FB Servers, and Akamai Servers Geolocation: Until this step we have a large dataset in which each ping probe is associated with a source IP address (i.e., PL node), destination IP address (i.e., FB or Akamai server), and delay. However, in order to perform the study described in the introduction we have to geolocate each IP address so that for each ping entry in our dataset we also know location of source node and location of destination node. To geolocate each source node, FB server and Akamai server we used the Maxmind database[5] to bind each IP address to its respective location. The location included country and city (if available).

We would like to note that the final dataset employed in our research is publicly available for the research community.[6]

End Users' Access Delay to Facebook Services

In this section we aim to understand the performance level experienced by end users in terms of the latency in accessing FB services located in either native FB or Akamai servers. Table 3a shows the detail of the average access delay (and its standard deviation) per country to access FB services in servers located in FB facilities, and Table 3b shows the same parameters in relation to Akamai servers. In addition, Fig. 2 shows the average access delay to access FB services in servers located at FB facilities (Fig. 2a)

[5] http://www.maxmind.com/

[6] http://www.it.uc3m.es/acrumin/papers/FB_Arch_project.rar

(a) Facebook				(b) Akamai			
Country	Avg.Delay (ms) ± STD	Country	Avg.Delay (ms) ± STD	Country	Avg.Delay (ms) ± STD	Country	Avg.Delay (ms) ± STD
(1)				(1)			
Singapore	193.66 ± 59.41	Brazil	169.78 ± 60.93	China	174.59 ± 213.30	Argentina	124.98 ± 79.67
Romania	190.07 ± 50.55	Israel	167.14 ± 90.85	Uruguay	157.40 ± 78.98		
China	187.14 ± 227.29	Australia	164.11 ± 43.22	(2)			
Uruguay	179.96 ± 65.08	Argentina	155.38 ± 67.49	New Zealand	95.98 ± 83.41	Hong Kong	71.41 ± 80.92
Portugal	177.91 ± 69.02	New Zealand	152.02 ± 38.00	Korea, Rep.	90.14 ± 90.75	Jordan	68.72 ± 38.89
Slovenia	169.86 ± 48.50			Australia	87.03 ± 89.32	Tunisia	63.05 ± 27.39
(2)				Ecuador	79.62 ± 55.81	Israel	54.64 ± 78.04
Denmark	140.93 ± 38.52	Ecuador	106.69 ± 36.66	Brazil	78.22 ± 68.44		
Finland	137.12 ± 61.17	Tunisia	104.47 ± 50.99	(3)			
France	133.12 ± 61.04	Norway	104.01 ± 62.21	Portugal	49.43 ± 16.24	Canada	22.59 ± 38.57
Korea, Rep.	128.84 ± 76.56	Italy	102.57 ± 75.37	Singapore	45.32 ± 73.15	Finland	22.54 ± 17.42
Japan	126.96 ± 64.96	Taiwan	101.71 ± 85.34	Puerto Rico	41.86 ± 41.65	Slovenia	**18.70 ± 17.11**
Sweden	114.28 ± 56.11	Spain	100.94 ± 73.13	Turkey	39.14 ± 45.65	U.S.	**15.90 ± 25.02**
Jordan	109.95 ± 61.85	Hong Kong	100.58 ± 84.43	Taiwan	35.74 ± 57.75	Italy	**15.06 ± 12.83**
Puerto Rico	108.42 ± 36.14	Hungary	100.05 ± 76.77	Greece	33.78 ± 24.88	Germany	**10.94 ± 8.58**
(3)				Japan	30.42 ± 42.95	U.K.	**10.80 ± 11.74**
Poland	99.69 ± 58.80	Russia	77.49 ± 52.54	Spain	27.25 ± 19.00	Belgium	10.68 ± 27.21
Greece	92.70 ± 69.36	Netherlands	59.52 ± 54.77	Russia	26.38 ± 20.41	Sweden	**10.20 ± 10.98**
UK	90.46 ± 50.67	Austria	53.75 ± 50.77	Romania	26.36 ± 17.35	Hungary	**8.80 ± 7.36**
Switzerland	88.40 ± 66.13	Turkey	51.37 ± 64.37	Ireland	24.35 ± 41.62	Switzerland	**8.56 ± 12.02**
Germany	84.47 ± 61.80			France	24.34 ± 46.36	Netherlands	**7.77 ± 13.00**
(4)				Norway	23.33 ± 15.62	Denmark	7.09 ± 6.02
Czech Rep.	48.36 ± 51.95	Canada	38.51 ± 46.15	Poland	23.15 ± 10.31	Austria	6.84 ± 5.76
Ireland	45.88 ± 50.55	US	36.81 ± 34.72				
Belgium	42.70 ± 56.02						

Table 3. Average delay (milliseconds) ± standard deviation to access FB services from different countries for services located in a) FB servers; b) Akamai servers.

and at Akamai facilities (Fig. 2b). Overall, FB users need 113 ms in average to access native FB servers, but only 43 ms to reach Akamai nodes providing FB access. This means that accessing FB services in Akamai nodes reduces the delay 2.5× the delay. Next, we provide a detailed analysis of the access delay performance per country.

Access Delay to Native Facebook Servers

Based on the results of Table 3, we have defined four groups in terms of their access delay to FB servers, which are illustrated in different color range in Fig. 2 as well.

The first group: Refers to all those countries with an access delay longer than 150 ms

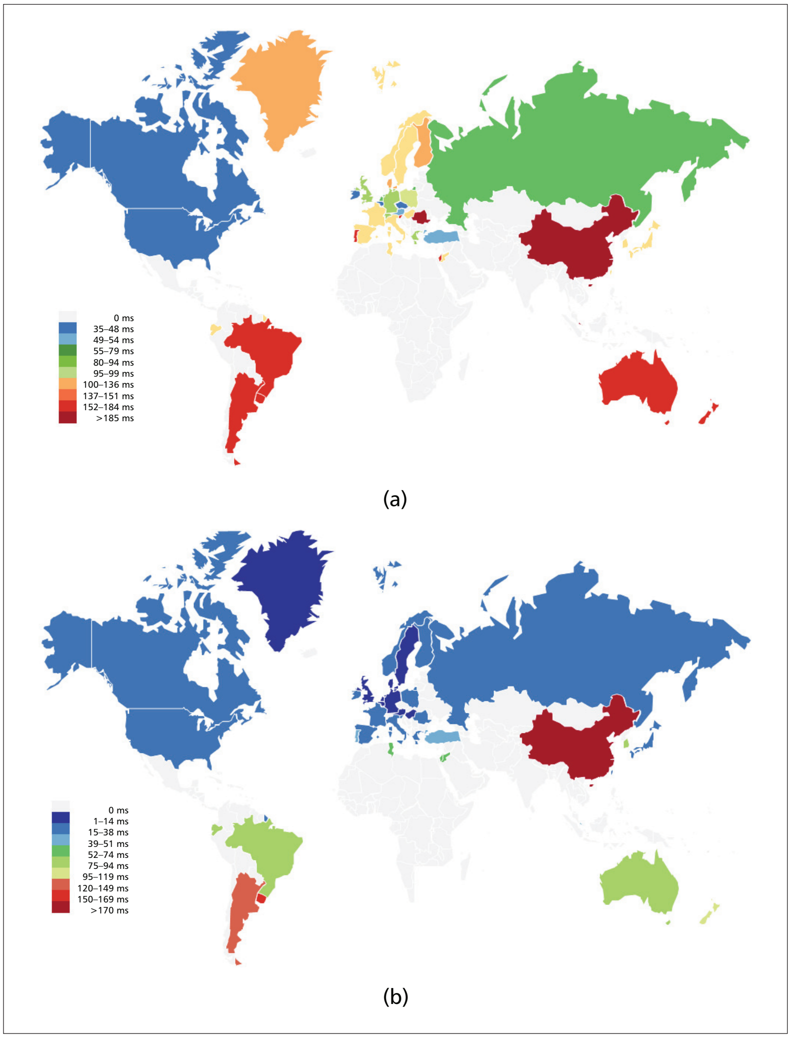

Figure 2. Average delay (miliseconds) to access FB services from different countries for services: a) located in FB servers; b) located in Akamai servers.

(red group in Fig. 2a). This group is formed by countries that are quite far from the United States (e.g., Australia, New Zealand), South American countries, and three countries we did not expect to find in this group (Portugal, Slovenia, and Israel) since their surrounding neighbors show a considerably lower delay.

The second group: Formed of those countries whose delay ranges between 100 and 150 ms (orange group). This group includes Northern European countries, Asian countries with deep penetration of high-speed access connections (e.g., Japan, South Korea, Hong Kong), countries from Central America, and Mediterranean countries including some important European ones such as France, Italy, and Spain.

The third group: Includes those countries with a delay greater than 50 ms but less than 100 ms (green group). This group is mainly formed by countries located in Central Europe plus Greece, Turkey, and the United Kingdom.

The last group: Contains those countries with access delay under 50 ms (blue group). This includes the two countries hosting native FB servers, the United States and Ireland [1], and Canada due to its proximity and good connectivity with the United States. Surprisingly, this group also includes Belgium and the Czech Republic, which intuitively would have fit better in the third group.

Access Delay to Akamai Servers

In the case of Akamai nodes we just define three groups for our discussion.

The first group: Formed by three countries that experience an average delay longer than 100 ms (red group in Fig. 2b). These countries are China, Argentina, and Uruguay. This happens because an important portion of the FB queries from these countries are redirected to remote Akamai nodes, which could be located, for instance, in the United States.

The Second Group: Consists of countries with an average access delay ranging between 50 and 100 ms (green group). This include far eastern countries like Australia, New Zealand, South Korea, and Hong Kong; two countries in South America, Brazil and Ecuador; and three countries from North Africa and the Middle East: Jordan, Tunisia, and Israel. As seen in the next section, the first six countries count on their own Akamai nodes, but a relevant portion of their demand is attended to by foreign Akamai servers. In addition, Jordan and Tunisia do not host any Akamai nodes, but are served by Akamai nodes located in Europe, which is relatively close. It is surprising that Australia (as a developed country) experiences quite bad performance in accessing FB services through Akamai nodes. To have better insight, we leveraged the FB ads planner[7] to retrieve the potential reach for ads in each country. We have found that Australia has a potential reach of 13 million FB users, while some of the countries in the third group, like Greece and Slovenia, which present 50 and 70 ms less average access delay, have a potential reach of 4.4 and 0.7 million, respectively. Another surprising case in this group is Brazil, a huge country with a population of more than 200 million and potential reach of an audience of 86 million for FB ads, and shows an average Akamai access delay around 78 ms.

The third group: Includes the countries with access delay below 50 ms (blue group). This group mainly includes developed countries from Europe, Asia (i.e., Japan and Singapore), and North America (United States and Canada). This is a good estimation of a short list of important countries for FB, where FB is interested in offering a better quality of service through Akamai nodes.

Furthermore, it is interesting to note that Akamai offers the best delay performance (i.e., below 10 ms) to small countries roughly located in Central Europe (Hungary, Switzerland, Netherlands, Denmark, Austria, and Czech Republic). This happens because these are very small countries (in size) that experience a very small delay due to the short distance to a large number of Akamai nodes located in Central Europe.

[7] https://www.facebook.com/ads

Akamai Nodes Distribution to Provide Access to Facebook Services

This section provides a global picture of the deployment of Akamai nodes to serve FB services worldwide.

Local vs. External Access

Figure 3 shows which portion of the queries for each country (i.e., ICMP echo) is managed by Akamai servers hosted in the same country than the source node(s) and which portion is served by Akamai nodes in foreign countries.

There are only two countries showing a higher portion of local access to Akamai servers compared to external access which are US and Singapore with 90 percent and 62 percent of the queries going to local Akamai servers. The case of Singapore might be unexpected, but as we will show later, Singapore has a high number of Akamai nodes (i.e., IPs). Close to Singapore performance, we find the case of Taiwan in which half of the queries are dealt with local servers and half by foreign servers.

We found that there are a limited number of countries that use local Akamai nodes to serve between 30 percent-40 percent of their queries. These are:

- The largest European countries by size (i.e., Germany, France and Spain) all of which have a large number of Akamai servers;
- Australia, another large country with high number of FB subscribers, that are located far from native FB servers and thus FB is motivated to use Akamai CDN to offer a good performance to Australian subscribers;
- Three European countries, Switzerland, Sweden and Romania, each particularly distributed geographically in the center, north and east of Europe respectively.

The Akamai infrastructures in Switzerland and Sweden bring them to have access delays to Akamai nodes in the order of 10ms. Finally, Romania has just six Akamai servers that service 35 percent of the queries generated in Romanian nodes.

Next, we found a large number of countries keeping between 7 percent and 30 percent of FB queries were responded locally, while most of them were serviced by foreign Akamai servers. Each of these countries have more or less Akamai nodes that allow keeping part of the queries locally, but their delay is mostly affected by how far those Akamai nodes are located from the major part of their queries.

Finally, there were 10 countries for which we could not identify any local Akamai server. Among these are five European countries (Belgium, Denmark, Hungary, Portugal and Slovenia), each with a population under 10M and close to countries with a significant deployment of Akamai nodes running FB services. The fact that these countries are experiencing a very good service by accessing Akamai nodes in nearby countries explains the low presence of Akamai servers.

This group of countries without Akamai servers also includes Turkey, which we found similar delay to some European countries like Greece or Portugal because the three PL nodes used for our experiments are located in the western part of Turkey (i.e., Istanbul and Izmir). Next, we discuss the case of Uruguay, a small South-American country surrounded by Argentina and Brazil that already contains some Akamai servers. Interestingly, the results in Table 4 show that the five PL nodes placed in Uruguay access Akamai servers located in Brazil as well as servers in Mexico and US that are far away, but never go to Argentina. Two small countries, Tunisia and Jordan, both of them are served by Akamai nodes located (mainly) in Europe. Finally, we find China which is currently blocking FB, and thus it does not make sense to deploy Akamai nodes to serve FB subscribers and they are served by Akamai nodes all over the world.

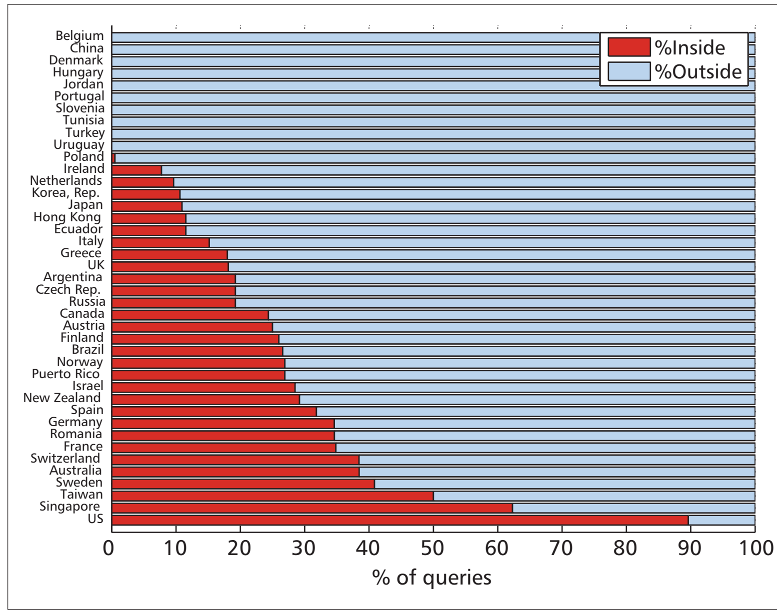

Figure 3. Portion of FB queries from each country served by local (% inside) and foreign (% outside) Akamai nodes.

Country Coverage by Akamai Servers

Table 4 shows for each country which is hosting Akamai nodes, the overall number of IPs linked to Akamai nodes located in that country (column #IP), and the list of countries hosting nodes that access those IPs[8] (column Served To(#IP)). For each source-querying country we represent the overall number of IPs (between brackets) accessed in the destination country hosting Akamai nodes.

We found 35 countries that host Akamai nodes to provide FB access to the 41 countries represented by PL nodes. Among them, at the top of Table 4, we find 13 countries where Akamai nodes only serve local users. In the middle of the table we list four countries: Azerbaijan, Malaysia, Mexico and Panama, whose Akamai nodes only serve foreign countries. In fact, this behavior responds to the fact that we did not have any PL node located in those countries. Otherwise, we would very likely have observed that these Akamai nodes also serve local users.

[8] For simplicity during the discussion we use the number of IPs as the number of servers/nodes, even though we are aware that it is feasible that the same physical server could hold more than one IP (multiple network cards, virtualization, etc.)

Country	#IP	Served To (#IP)
Argentina	9	Argentina (9)
Canada	28	Canada (22)
Ecuador	3	Ecuador (3)
Greece	7	Greece (7)
HongKong	6	HongKong (6)
Israel	18	Israel (18)
Korea	7	Korea (7)
Poland	2	Poland (2)
Puerto Rico	8	Puerto Rico (8)
Romania	6	Romania (6)
Russia	7	Russia (7)
Spain	35	Spain (35)
Taiwan	9	Taiwan (9)
Azerbaijan	1	China (1)
Malaysia	21	HongKong (3), NewZealand (16), Singapore (2)
Mexico	4	Uruguay (4)
Panama	4	Canada (4)
Australia	15	Australia (10), Japan (4), Taiwan (1)
Austria	49	Austria 9), Greece (24), Hungary (26), Israel (2), Poland (37), Slovenia (27)
Brazil	26	Brazil (22), Uruguay (16)
Czech	11	Czech (6), Poland (7), Russia (4)
Finland	24	Finland (19), Norway (4), Russia (3), Sweden (12)
France	176	Belgium (20), Finland (4), France (60), Germany (4), Greece (1), Hungary (1), Ireland (7), Israel (3), Jordan (20), Poland (10), Singapore (3), Spain (41), Switzerland (5), Tunisia (4), Turkey (1), United Kingdom (48)
Germany	473	Australia (3), Austria (11), Belgium (30), China (6), Czech (13), Denmark (4), Finland (19), France (25), Germany (184), Greece (43), Hungary (11), Ireland (9), Israel (20), Italy (20), Jordan (5), Netherlands (12), Norway (5), Poland (31), Portugal (24), Romania (7), Russia (12), Slovenia (2), Spain (47), Sweden (11), Switzerland (86), Tunisia (9), Turkey (16), UnitedKingdom (14), United States (4)
Ireland	6	China (1), Ireland (5)
Italy	49	China (1), Greece (4), Hungary (2), Israel (1), Italy (20), Jordan (14), Switzerland (1), Tunisia (6), Turkey (1), United States (2)
Japan	36	China (17), Hong Kong (4), Japan (16), Korea (5)
Netherlands	39	Belgium (3), China (1), France (4), Ireland (14), Netherlands (6), Tunisia (6), UnitedKingdom (1), United States (3)
NewZealand	11	China (1), NewZealand (10)
Norway	8	Finland (2), Norway (5), Sweden (2)
Singapore	110	Argentina (3), Brazil (26), China (2), Ecuador (4), HongKong (13), Japan (1), Korea (3), New Zealand (1), Puerto Rico (15), Singapore (26), Taiwan (1), United States (30), Uruguay (12)
Sweden	77	Denmark (1), Finland (31), Ireland (7), Norway (6), Poland (7), Russia (10), Sweden (24), United Kingdom (19)
Switzerland	49	Australia (5), Poland (9), Sweden (8), Switzerland (33)
United Kingdom	246	Belgium (21), Denmark (4), France (19), Germany (22), Greece (34), Hungary (7), Ireland (17), Israel (34), Italy (2), Netherlands (23), Norway (13), Poland (29), Portugal (21), Romania (1), Spain (21), Sweden (5), Switzerland (11), Tunisia (9), Turkey (15), UnitedKingdom (45), UnitedStates (4)
UnitedStates	2505	Argentina (67), Australia (27), Austria (19), Belgium (39), Brazil (48), Canada (148), China (177), Czech (17), Denmark (13), Ecuador (17), Finland (16), France (69), Germany (117), Greece (32), HongKong (67), Hungary (26), Ireland (16), Israel (11), Japan (96), Jordan (1), Korea (66), Netherlands (18), NewZealand (21), Norway (17), Poland (47), Portugal (79), Puerto Rico (17), Romania (7), Singapore (21), Slovenia (14), Spain (24), Sweden (22), Switzerland (9), Taiwan (19), Tunisia (6), Turkey (13), UnitedKingdom (32), UnitedStates (1668), Uruguay (52)

Table 4. The first column shows the list of countries hosting Akamai nodes offering access to FB services. The second column shows the number of identified Akamai-related IPs in each country. The third column shows the list of countries including nodes querying Akamai IPs in the country referred to in the first column. The number between parentheses reflects the number of IPs accessed in the reference (first column) country.

Finally, at the bottom of the table, we find a major part of the countries (18 in total) with Akamai nodes that process queries from both local and foreign PL nodes. Next, we discuss the most interesting aspects for this group.

First, we observe that large countries with a relatively heavy weight in the geopolitical environment such as the United States, United Kingdom, France, Germany, and Italy have a high number of Akamai nodes (i.e., associated IPs) that serve a large number of countries. The four European countries mainly serve nodes from all over Europe, at a minor level nearby non-European countries like Israel, Jordan, Tunisia, and Turkey, and on a very small scale the United States and China. We also found a similar pattern in The Netherlands, although it has a lower Akamai presence. Furthermore, we discovered more Akamai nodes in the United States than in the rest of the countries together. These servers process queries from users located all over the world. This clearly has an impact on the delay for those countries that access Akamai nodes in the United States for a large portion of their queries, despite being far from the United States (e.g., Uruguay, Argentina, China, and Korea).

Next, we observe that Akamai nodes in Northern European countries (Norway, Finland, and Sweden) mainly respond to the demands of users located within those northern countries. A third observation is that Ireland and New Zealand should actually be located at the top of the table since they mostly attend to local FB demand, along with a few queries from China. Fourth, Akamai nodes located in small Central European countries such as Austria, Czech Republic, and Switzerland service FB demand mainly from local and nearby countries' users. We can find a similar pattern for Japan and Brazil, and additionally Australia, where the nodes mostly deal with internal demand for FB services but also receive some queries from nodes located in Japan and Taiwan. Finally, Singapore (the fourth country in terms of number of Akamai IPs) presents rarer results. On one hand, Akamai nodes in Singapore exhibit an expected behavior by serving users located in Asia. On the other hand, we discovered a very strange pattern in which Akamai nodes in Singapore attend quite a few nodes located all over America (including North and South America).

In summary, we can conclude that FB subscribers' queries are usually attended by Akamai nodes located either locally or in some nearby country. This provides a bounded access delay leading to the result presented above that indicates a delay 2.5× lower when an FB query is resolved by an Akamai node instead of a native FB server. However, we can still find some odd cases where source nodes access Akamai nodes located far away, which has a harmful impact on their access delay to FB services.

Related Work

We found a number of works related to our article that can be classified into two different categories: CDN infrastructure analysis and Facebook services analysis.

CDN Infrastructure Analysis

There are some prior studies that analyzed different aspects of large CDNs like Akamai [6, 7] or the CDN used by Google to serve youtube videos [8]. In the latter study the authors aim at understanding from where videos are served, and how effective is this distribution. One of the main conclusions of this study is that round-trip time (RTT) is used to select the preferred data center to serve the video. The studies on Akamai CDN go from a general overview [6] to a more detailed analysis of Akamai's system components and architecture [7] in which the authors probe an Akamai network from 140 PlanetLab nodes during two months and characterize some aspects of Akamai architecture deployment such as server diversity, redirection dynamics, and latency. Finally, we found a study [9] in which the authors examined how CDNs are used to host and serve FB content from a network perspective. This work relies on a dataset including one month of HTTP traces collected in mid-2013 from the third generation (3G) mobile network of a large European ISP.

Facebook Services Analysis

There are also some research works that carried out different performance analyses on Facebook services. The authors in [1] look at the established connections when FB users login to the system. In particular, they identify different sections in the FB wall page of a user, and analyze how the information filling those sections is retrieved. An earlier work from 2010 [10] identified some performance degradation (delay, packet losses, etc.) for some users accessing FB from outside the United States. Finally, we have found another interesting study [11] which states that photo viewing is the most critical service for FB, and presents a detailed description on how FB photos are distributed to CDN Akamai servers. However, it does not perform a geographical analysis to understand how different regions of the world are being served as we do in our article.

Lessons Learned and Recommendations

In this section we present the most important lessons extracted from our work and provide some recommendations that could improve the current delay experienced by users in some relevant countries.

1: Our study confirms the good work Akamai does for a large-scale web service such as Facebook. Our results show that FB is reducing delay 2.5× by using the Akamai nodes. This latency reduction is of great importance for Facebook or any other Internet service given the monetary implications associated the delay experience by end users [2].

2: At the time of our study, Akamai nodes were mostly responsible for serving heavy content mainly associated with photos and videos shared on Facebook. In contrast, Facebook native servers were in charge of processes like registration and authentication.

We can conclude that FB subscriber queries are usually attended by Akamai nodes located either locally or in some nearby country. This provides a bounded access delay leading to the result that indicates a delay that is 2.5x lower when a FB query is resolved by an Akamai node instead of a native FB server.

At the time of our study, Akamai nodes were mostly responsible of serving heavy content associated mainly to photos and videos shared in Facebook. In contrast, Facebook native servers were in charge of processes like the registration and authentication.

3: Akamai is very efficient (< 50 ms delay) in serving Facebook content in Europe and North America, which is explained by two factors:
- Akamai is very well positioned there with a huge number of servers.
- A major part of the revenues obtained by FB from advertisement happens in Europe and North America; thus, it is very important to offer good quality of service to the subscribers in those locations.

4: There is some room for improving the current Facebook infrastructure in some countries like Australia and Brazil. These two countries have 13 and 86 million subscribers, respectively, according to the data reported by the FB ads planner, and experience much higher access delay (87 and 78 ms, respectively) than other countries with much lower numbers of subscribers like Slovenia. Therefore, we believe that Facebook should find a solution to improve the experience of Australian and Brazilian users by further exploiting Akamai nodes in those countries.

Conclusions

This study presents a comprehensive measurement-based analysis of the FB network infrastructure with special emphasis on depicting how Akamai nodes replying to FB queries from subscribers are distributed throughout the world. In this context, we have analyzed the average access delay FB subscribers experience to access FB services delivered from native FB servers as well as Akamai servers depending on the country in which they are located. Moreover, we have thoroughly discussed the coverage offered by those Akamai nodes serving FB services.

Acknowledgments

This works is partially supported by the European Celtic-Plus project CONVINcE and eCOUSIN (EUFP7-318398). This work is also funded by the Ministerio de Economía y Competitividad of SPAIN through the project BigDatAAM (FIS2013-47532-C3-3-P).

References

[1] W. Wongyai and L. Charoenwatana, "Examining the Network Traffic of Facebook Homepage Retrieval: An End User Perspective," *Proc. Int'l. Joint Conf. Computer Science and Software Engineering*, 2012.
[2] A. Singla *et al.*, "The Internet at the Speed of Light," *Proc. 13th ACM Wksp. Hot Topics in Networks*, 2014.
[3] B. Chun *et al.*, "Planetlab: An Overlay Testbed for Broad-Coverage Services," *ACM SIGCOMM Computer Commun. Review*, vol. 33, no. 3, 2003.
[4] A. Orebaugh, G. Ramirez, and J. Beale, *Wireshark & Ethereal Network Protocol Analyzer Toolkit*, Syngress, 2006.
[5] S. K. R Siamwalla and R Sharma, "Discovering Internet Topology," *Proc. IEEE INFOCOM*, 1999.
[6] E. Nygren, R. K. Sitaraman, and J. Sun, "The Akamai Network: A Platform for High-Performance Internet Applications," *SIGOPS Op. Sys. Rev.*, 2010.
[7] A.-J. Su *et al.*, "Drafting Behind Akamai: Inferring Network Conditions Based on CDN Redirections," *IEEE/ACM Trans. Net.*, vol. 17, no. 6, 2009.
[8] R. Torres *et al.*, "Dissecting Video Server Selection Strategies in the Youtube CDN," *Proc. 31st Int'l. Conf. Distributed Computing Systems*, 2011.
[9] P. Fiadino, A. D'Alconzo, and P. Casas, "Characterizing Web Services Provisioning via CDNS: The Case of Facebook," *Proc. Wireless Communications and Mobile Computing Conf.*, 2014.
[10] M. P. Wittie *et al.*, "Exploiting Locality of Interest in Online Social Networks," *Proc. Int'l Conf. Emerging Networking EXperiments and Technologies*, 2010.
[11] D. Beaver *et al.*, "Finding a Needle in a Haystack: Facebook's Photo Storage," *Proc. 9th USENIX Conf. Operating Systems Design and Implementation*, 2010.

Biographies

Reza Farahabkhsh [M] (reza.farahbakhsh@it-sudparis.eu) received his B.S. degree in computer engineering from Qazvin Azad University in 2006, his M.S. degree in 2008 from the University of Isfahan, Iran, in 2008, and his Ph.D from Institut-Mines Telecom, Telecom SudParis (CNRS Lab UMR5157) jointly with Paris VI (UPMC) in May 2015. He is now a postdoctoral researcher at Institut Mines-Telecom, Telecom SudParis since May 2015. His research interests are online social networks, P2P networks, Internet measurements, as well as Mobile IPv6 and IMS. He is a member of ACM.

Ángel Cuevas (acrumin@it.uc3m.es) received his B.Sc. in telecomunication engineering, his M.Sc. in telematics engineering, and his Ph.D. in telematics engineering from Universidad Carlos III de Madrid in 2006, 2007, and 2011, respectively. He is currently a tenure-track visiting professor at the Department of Telematic Engineering at Unviersidad Carlos III de Madrid. Prior to that he was a postdoctoral researcher at Institut Mines-Telecom, Telecom SudParis from March 2011 until January 2013. His research interests focus on online social networks, P2P networks, wireless sensor networks, and Internet measurements. He is a co-author of more than 30 papers in prestigirous international journals and conferences such as IEEE/ACM Transactions on Networking, *Elsevier Computer Networks*, IEEE Network, *IEEE Communications Magazine*, ACM CONEXT, ACM MSWiM, IEEE ISCC, and IEEE ICC. He was co-recipient of the Best Paper Award at ACM MSWiM 2010.

Antonio M. Ortiz received his Master's degree in advanced computer technologies (2008) and his Ph.D. in computer science (2011) from the University of Castilla-La Mancha, Spain. In 2012, he worked as a post-doctoral researcher at the Letterkenny Institute of Technology, Ireland, and at the Institut Mines-Telecom, Telecom SudParis, France during 2013 and 2014. He is currently a research engineer and project manager at Montimage, France. He actively participates in a number of European research projects and has contributed to the ITU-T and oneM2M standardization bodies. His topics of interest are focused on advanced monitoring and testing solutions applied to diverse aspects of network communications and cloud-based technologies, as well as on the Internet of Things, social networks, and artificial intelligence.

Xio Han received her co-joint Ph.D degree from Telecom SudParis and Université Pierre et Marie Curie. She studied at the Automation School of Northwestern Polytechnical University, China, and received her B.Sc. and M.Sc. there in 2008 and 2011, respectively. Her research interests include social networks analysis and modeling, social-based applications, and social P2P networks.

Noël Crespi [M'07, SM'08] holds Master's degrees from the Universities of Orsay (Paris 11) and Kent (United Kingdom), a Diplome d'Ingénieur from Telecom ParisTech, and a Ph.D and an Habilitation from Paris VI University (Paris-Sorbonne). From 1993 he worked at CLIP, Bouygues Telecom and then at Orange Labs in 1995. He took leading roles in the creation of new services with the successful conception and launch of Orange's prepaid service and in standardization. In 1999, he joined Nortel Networks as telephony program manager, architecting core network products for the EMEA region. He joined Institut Mines-Telecom in 2002 and is currently a professor and program director, leading the Service Architecture Lab. He coordinates the standardization activities for Institut Mines-Telecom at ITU-T, ETSI, and 3GPP. He is also an adjunct professor at KAIST, an affiliate professor at Concordia University, and on the four-person Scientific Advisory Board of FTW, Austria. He is the scientific director of the French-Korean laboratory ILLUMINE. His current research interests are in service architectures, services webification, social networks, and Internet of Things/Services. http://noelcrespi.wp.tem-tsp.eu/