Mercury: Revealing Hidden Interconnections Between Access ISPs and Content Providers

Manuel Palacin, Alex Bikfalvi, and Miquel Oliver

Networking Technology and Strategies (NeTS) Research Group,
Universitat Pompeu Fabra (UPF)
{manuel.palacin,alex.bikfalvi,miquel.oliver}@upf.edu

Abstract. Knowing the detailed topology of the Internet at the Autonomous System (AS) level is extremely valuable for both researchers and industry when making network policies. Although there are many measurement projects and databases that provide this information, such as ARK, RETRO, ONO and PeeringDB, they only offer a partial view for analyzing end-to-end Internet routing paths and they do not focus on the hidden direct interconnections between Access ISPs and Content Providers. In order to address these shortcomings, we present Mercury, a web platform focused on the AS-level interconnection between content providers and content consumers. Mercury enables users to visualize the AS topology, aggregating data from traceroute measurements of participants and AS information from several databases. The advantage of Mercury is that it discovers how operators connect to other organizations and how content providers organize their server's infrastructure (CDN) to reach their target audience. To this end, Mercury identifies Internet Exchange Points (IXPs) and AS relationships along an Internet path and presents this information via a web site and a built-in API. We evaluate its potential probing a dataset of 100 popular web URLs from the major Spanish ISPs and we successfully identify many direct interconnections that were hidden for other methodologies.

1 Introduction

Discovering the hidden interconnections between access Internet Service Providers (ISPs) and Content Providers (CPs) is a challenge for engineers due to the lack of data available to the general public and due to the several weaknesses of the existing tools and methodologies to extract it. From the perspective of the ISP business relationships, these direct interconnections are critical because of the increasing demand for multimedia content requiring an optimal quality of experience (QoE) for the end users. These interconnections are elusive because they are usually based on peering agreements where the BGP announcements are only visible between peers and their customers. We also focus on another type of hidden interconnection between access ISPs and CPs, where the CP servers infrastructure is placed within the ISP.

Several efforts have been made to discover the Internet topology at the Autonomous System (AS) level. Historically, two approaches have been used: the analysis of BGP paths and the *traceroute* traces. The BGP approach discovers the AS topology using a set of distributed monitors that sniff the BGP messages. This methodology uses the AS paths included in the BGP messages to infer the interconnections between the different ASes. However, BGP announcements describe the control plane which does not necessary correspond to the real path of the Internet traffic. Also, it is not effective for discovering the hidden interconnection and requires much more monitors to obtain a global topology. In contrast, the traceroute approach has the advantage that can discover the real path of the Internet traffic between two end points. As a disadvantage, the traceroute tool works at the IP level and is affected by the multi-path diversity of the IP protocol and requires an IP-to-AS mapping to infer the traversed ASes.

Toward this end we introduce *Mercury*[1], a platform for discovering the AS-level interconnections between content providers and content consumers. Mercury enables users to visualize the AS topology of access ISPs when they connect to other organizations and to identify how CPs organize their server infrastructure. To perform this, Mercury combines the two approaches presented previously. On one hand, a desktop client collects traceroute measurements from project participants. We have extended a version of traceroute, called Paris traceroute [4], which attempts to mitigate the multi-path problem for routers that implement per-flow load balancing. We rely on our own BGP datasets, collected from several data sources, to perform the IP-to-AS translation. The client uploads the measurements to the Mercury platform which presents all this information and statistics using a web site.

By leveraging this data, Mercury discovers the AS paths and interconnection relationships, while putting a special emphasis on the detection of interconnections between CPs and access ISPs. Mercury offers information about the number of AS hops to reach selected CPs, the type of AS relationships and the existence of IXPs in the path. Furthermore, Mercury allows researchers to consult aggregated statistics of different AS paths from multiple geographical locations to the same destination. This analysis is particularly useful for the identification of the server infrastructure used by Content Distribution Networks (CDNs). In addition, the platform provides a built-in API for expert users. In the following, we present several of the motivating scenarios.

Identification of the hidden interconnections: Mercury discovers many direct interconnections between CPs and access ISPs as the traceroute measurements are mostly done from commercial access ISPs with destination the most popular web sites. This information could be very useful to demonstrate that CPs are getting closer to the access ISPs.

Identify the architecture of CDNs: We identify the interconnection for CDNs, by revealing different paths to the same web resource from many points of origin. This helps researchers to analyze and classify existing CDN deployments, and to propose novel caching techniques. At the same time, CPs can evaluate the optimal strategies for reaching their target customers.

[1] A work-in-progress version is available at `http://mercury.upf.edu/mercury`.

Modeling the AS topology and the taxonomy of AS interconnections: Mercury aggregates data of Internet paths, facilitating the generation of AS graphs. The interconnection degree increases with the addition of new measurements. Such a dataset can serve for the evaluation of interconnection models of commercial networks.

Network economics: Mercury provides information about the AS relationships. Network operators can use this to optimize their traffic routing policies based on the observation of other AS topologies. Thanks to this data, an ISP can discover peering partners and avoid using transit connections.

Although Mercury was implemented to collect data from participants around the world, this paper only focuses on the Spanish interconnection market to facilitate the evaluation of the system. Our objective is to answer the following research questions: (i) how common are the direct interconnections between CPs and access ISPs and (ii) what are the most common content delivery solutions for CPs.

2 Related Work

Improving the accuracy of the Internet AS topology through measurements has been a major objective for researchers for many years. The Route-Views project provided one of the first datasets expressing the AS topology of the Internet, based on BGP measurements [13]. Despite their pioneering work, researchers realized that measuring the Internet topology from few vantage points leads to partial results [7]. To address this, Mao et al. developed an AS-level traceroute tool, creating a database of IP-to-AS mappings based on the observation of both BGP announcements in combination with IP traceroute measurements [11]. However, the authors noticed the difficulty of detecting IXPs and sibling relationships, as well as mapping mismatches due to measurements in a limited geographic region.

Recognizing the benefit of measuring the Internet from the edge, Shavitt et al. propose DIMES, a measurement infrastructure using a large number of clients [17]. Although the project is open and its data is freely available, the information does not include the relationship between ASes. In parallel, Dimitropoulos et al. focused on this issue. Their work started initially as an effort to model and generate synthetic but realistic AS topologies [9]. Subsequently, they attempted a classification of the Internet ASes using data collected from the Internet Routing Registries and RouteViews. Their data in combination with the active measurements of the Ark project sponsored by CAIDA has contributed to the improvement in the knowledge of the AS Relationships[5].

In parallel, other projects such as PeeringDB, PCH and EuroIX were developed to facilitate the AS interconnection. They put a special emphasis on peering interconnections, such that participant ASes can register their IP prefixes. Based on this data, there have been many new contributions augmenting the AS topology with relationship information [14]. He et al. merge both data from these databases with their own traceroute tool, called RETRO, to identify IXP participants [10]. While they collect data using public traceroute servers, Chen et al. increase the number of traceroute sources by developing a measurement add-on, called ONO, for a popular BitTorrent client [8]. In parallel, Augustin et al. provided a new approach for detecting IXPs and inspecting their AS participants based on various databases and traceroute measurements [3]. They detected 223 out of 278 IXPs and demonstrated that most of the remaining IXPs are invisible to tracerouting.

Almost all of these works point to discover the Internet topology in a global way, not focusing in any specific player. Ager et al. [1] proposed a new methodology for detecting web content infrastructure based on the evaluation of BGP snapshots and DNS queries to the most popular and long-tail web sites. Their results revealed what ASes have the content and where are they located. Our work, although uses a similar methodology to these previous efforts, differs in that we focus on identifying a specific type of interconnection, the CP to access ISP interconnection.

3 Objectives and Methodology

Mercury is a measurement platform dedicated to discovering the interconnection between content providers and content consumers. Fig. 1 illustrates the AS information that one can discover using Mercury. On the consumer's side, Mercury illustrates how access ISPs reach the CPs. This result should reveal whether access ISPs use different interconnections strategies depending on their size. On the provider's side, Mercury facilitates the identification of content distribution architectures, depending on the determined geographical, IP or AS destinations.

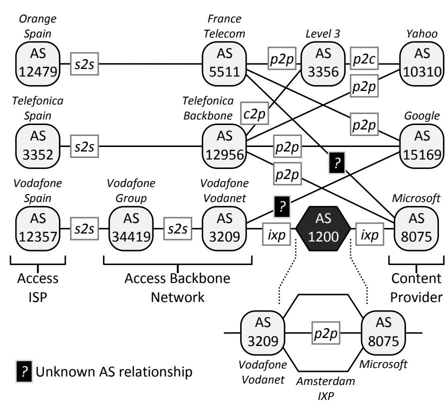

Fig. 1. Example of AS interconnections.

Our objective is transform Mercury into a comprehensive public dataset of end-to-end Internet paths. In addition to accessing the data, users can contribute, sending their own traceroute measurements using a desktop client. Although we evaluated the option of implementing a web-browser plugin, we discarded this option due to the inability of accessing to certain restricted functions of the operating system such as opening raw sockets. Therefore we finally implemented a stand-alone application inspired by the DIMES and ONO projects.

Mercury features a user-friendly web environment to navigate through the available datasets, query for certain data entries and to visualize interactive plots. Expert users can alternatively use the provided API to query for stored information. The API is open and based on the REST-WS protocol.

3.1 Measurement Methodology

As illustrated in Fig. 2, Mercury collects topology data from two main sources: (i) traceroutes from end-points clients across the Internet, and (ii) Internet topology information from several trusted databases. The measurement methodology is the following.

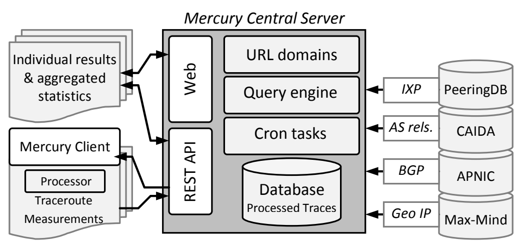

Fig. 2. The Mercury architecture.

Content Destination Selection: we selected the top 100 web site destinations from Alexa Top Sites [2] for 18 different countries. Once, we have the top 100 destinations for each country, we parse each web site in order to extract all the URL resources. With this step we determine all web servers that contribute content to the rendering of an individual web page, a step of particular importance when embedded content is hosted by a CDN. Finally, we obtain between 700 and 800 unique URLs per country that we store in the Mercury main server and the clients download their URL list according to their country.
Traceroute Collection: Mercury platform receives IP traceroute paths from the desktop clients and stores them into a database. Our client extends and implements Paris *traceroute* [4] in order to prevent multi-path anomalies across routers that perform load balancing. Because the correctness of traceroute data is paramount to the inference of real AS relationships, we resort to multiple tests per destination, at least 50 in most cases, which helps us to filter correct traceroutes. Using a desktop client, as opposed to a specially deployed traceroute servers in vantage points, guarantees a greater topological and geographical diversity [8]. Processed traceroute information is uploaded via the API to the Mercury platform. Clients probe a set of web destinations provided according to the geographical region. The rationale for this approach rests on the belief that access ISPs and CPs have incentives to directly connect when the corresponding content is popular in their country.
AS Resolution: Upon executed a traceroute measurement, the client translates the IP addresses to the equivalent AS number requesting mapping data to the

Mercury platform. A scheduled task in the Mercury platform collects every 24 hours IP prefixes from the APNIC/RouteViews from BGP announcements [16] and updates its own database with the new IP-to-AS mappings.
Geolocation: In parallel to the AS resolution, the Mercury client obtains, when it is available, the geographical location of IP addresses. The Mercury platform has a scheduled task that requests the IP-to-geo dataset from the MaxMind service [12]. Although we offer this information, we have to note that sometimes these mappings are not precise, specially those related to IP address that belong to large CPs or CDNs.
IXP Identification: We search the IP addresses matching an IXP from the PeeringDB database [15]. Fig. 1 illustrates the principle of IXP interconnection, where the IP addresses within the IXP subnet and revealed by the traceroute, actually belong to the connecting ASes.
AS Relationships: Finally, we examine the interconnection relationships between ASes using data from CAIDA [5]. Thanks to this dataset, we identify the peering (*p2p*), customer-to-provider (*c2p* or *p2c*) and sibling (*s2s*) relationships (see Fig. 1). In addition to this, we provide an extra dataset with (*s2s*) relationships extracted from analyzing the ISPs owners.

3.2 Implementation Internals

Mercury is formed by two main software instances: (i) the *desktop clients*, geographically distributed around the world and used by the participants, and (ii) the *Mercury Central Server* or *MCS*. The MCS is the responsible of formatting and storing the information from the external datasets and the responsible of storing and publishing the processed measurements sent by the clients. The client performs the traceroute probes and processes the results for obtaining the IP-to-AS translation, the IXP detection and the AS relationships of the end-to-end path.

The MCS obtains the IP-to-AS mappings from the BGP monitors through the Routing Report project [16]. We download daily the BGP report that contains the AS origin of the running IP prefixes. This information is structured and stored in a database and combined with IXP mappings from PeeringDB. In addition, we obtain the relationship type of each interconnected AS pair from the CAIDA AS relationships table [5] and the geolocation of the IPs from MaxMind [12]. Finally the central server has the list of URLs to be examined for each country. All these datasets are stored in the MCS and the clients can download them using the REST-API.

The Mercury client is a Desktop application written in C# that automatically downloads the list of URLs and executes the traceroute probes. Before tracerouting, the client executes a *nslookup* query for each URL in order to obtain the corresponding IP address. Sometimes it obtains multiple IPs for a same URL. This denotes the existence of load-balancing which is a technique of CDN architectures. Then, the client execute the traceroute measurements for the set of IPs of each URL.

The traceroute is performed using a modified version of Paris traceroute [4] to minimize the effect of multiple paths during IP routing. That is, using Paris traceroute we modify the IP packets for generating traffic flows that follow the same paths across routers that implement per-packet load balancing. In our implementation, for each destination IP address we generate 5 different flows with 5 attempts per flow, for both ICMP and UDP traffic, respectively. Therefore, we generate 50 different traceroutes for each destination. The client translates each IP address from each IP hop to an AS number using the MCS-API. We are mostly interested in the AS level of a traffic path. This means that we are mostly concerned about detecting and correcting inconsistencies at edge routers between different ASes. This step is crucial because it aggregates flows and detects, corrects (if possible) and discards multi-path behaviours that lead us to obtain loops and missing ASes. To solve these anomalies we use an algorithm that detects AS hop inconsistencies and corrects them based on the analysis of the previous and next AS hops (see Fig. 3). We also assume a hot potato policy at the AS level. Therefore, the combination of Paris traceroute with this algorithm minimizes the number of incorrect traces.

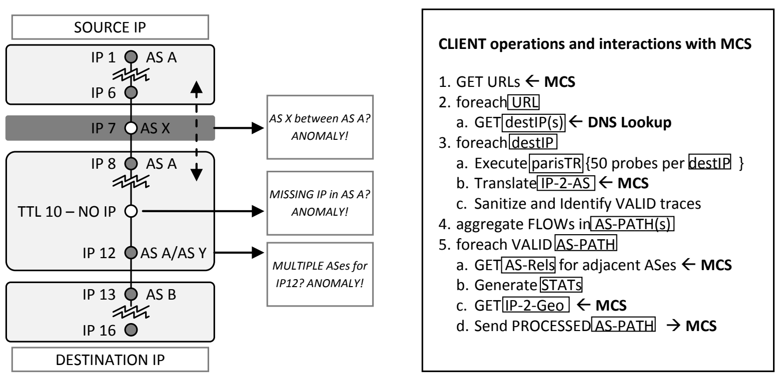

Fig. 3. Processing AS paths.

Once the traceroutes are executed, the client adds extra information from the external datasets of the MCS. In this step it includes to the traces the AS relationships between adjacent ASes, the geolocation of the end-points and it generates an statistical summary. The client is the responsible of processing all the information releasing the MCS for doing only searching and publication tasks. Finally, when the client ends processing, all the data is sent to the MCS.

All the stored measurements are publicly available using the Web interface and the REST API. Researchers can obtain aggregated statistics for a set of paths. For example one can aggregate paths filtering by destination URL, AS number or geographic location. These statistics include average, mean and standard deviation about the number of AS hops and the type of the AS relationships but also other network indicators like the number of destinations IPs and ASes that host a destination URL, e.g. *www.20minutos.es* is in average at 0.88 AS hops, points to two different IPs (193.148.34.26 and 89.140.253.190) and is hosted in two different ASes (AS3324 and AS6739).

4 Case Studies

We conducted an experiment using a set of traceroute measurements from various end-points located in the major Spanish access ISPs. We select Spain because is a typical post-monopoly market and we expect that this experiment will obtain similar results in the larger European countries. This experiment does not require too many participants as we consider that most of the users of a certain access ISP will be routed, at the AS-level, using the same policy. Hence, at least one participant in each one of the five major Spanish access ISPs will be enough to draw conclusions on their interconnections. The objective of this experiment is twofold. Firstly, we test the detection and aggregation of interconnection relationships for a set of given URL destinations, showing the types of connectivity for a given end-point. Secondly, we leverage the known relationships between end-points and the Internet AS graph to show how we can discover the architecture of complex distributed systems, such as CDNs.

Toward this end, we selected a set of 100 web destinations from the Alexa Top 100 list in Spain as CPs. From these 100 sites we extract the embedded URLs from each site giving us more than 700 URLs. We parse the embedded URLs in order to identify media content hosted in external CDNs. On the side of content consumers, we have 5 volunteers from Barcelona from each of the 5 major Spanish access ISPs: Telefonica, Orange, ONO, Jazztel and Vodafone (see Fig. 4).

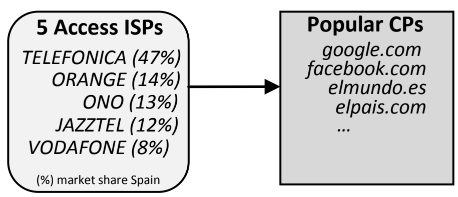

Fig. 4. Major access ISPs probing popular CPs

Each participant uses the client for tracerouting the set of destinations, and uploading the results to the MCS. We require clients located at commercial access ISPs in order to see how these ISPs interconnect with the CPs. Upon receiving the measurement data, Mercury distinguishes between *completed* and *inconsistent* traceroutes. Then it stores the measurements with a flag that identifies the different inconsistences of a path.

4.1 Revealing Hidden Interconnections

We use Mercury to discover direct interconnections between CPs and access ISPs without intermediaries and CPs that place their servers within the access ISP network. Table 1 shows the identification of direct interconnections for a subset of popular web sites that include global and local CPs. It compares the existence of direct interconnection relationships, either physically direct or across a sibling AS of the same organization, with the AS relationships from the the CAIDA dataset [5]. In these results, we use checkmarks to emphasize the matches between Mercury and CAIDA, crossings to indicate relationships not found in CAIDA, and dashes represents the non-existance of direct interconnections.

Table 1. Identification of direct interconnections.

	Google	Facebook	Yahoo	Twitter	Amazon	MSN	Wikipedia
Telefonica	Sibling ✗	Sibling ✓	Sibling ✓	No –	Sibling ✓	Sibling ✓	Sibling ✓
Orange	Sibling ✗	Sibling ✓	No –	No –	No –	Sibling ✗	Sibling ✓
ONO	Direct ✗	No –	No –	No –	No –	No –	No –
Jazztel	Direct ✗	Direct ✗	Direct ✗	IXP ✗	IXP ✗	IXP ✗	No –
Vodafone	Sibling ✓	Sibling ✓	Sibling ✓	IXP* ✓	IXP* ✓	Sibling ✗	Sibling ✓

Note: IXP* is a relationship where a sibling AS is connected to an IXP

Mercury identifies more links than CAIDA dataset when we focus on detecting direct interconnections. Although Mercury sometimes is not capable to identify the AS relationship type of a direct interconnection, it at least detects it, making possible to focus on this link in future studies in order to detect the relationship type. However, we conjecture that most of the direct interconnections are based on peering or paid-peering relationships according to the peering policies of both access ISPs and CPs. One can observe that access ISPs and large CPs find more attractive this formula than using an intermediary AS (see Google and MSN with Jazztel). They find the direct interconnection mutually beneficial, i.e. the content provider can be closer to the users and can offer a better QoE while the access ISP obtains an economic compensation from the paid-peering agreement.

Mercury shows us that Google has direct connections with all the major spanish ISPs. We also observe that not all CPs have direct interconnections to all access ISPs. This could be for different reasons: there are some CPs that only allocate the cacheable content using direct interconnections and there are also some CPs that have only agreed direct interconnections with some certain access ISPs and require an intermediary to reach the rest of access ISPs. This can be seen in Fig. 5, where one of the major Spanish press groups (ElMundo.es) maintain its own AS, but they contract a third party for delivering their cached content.

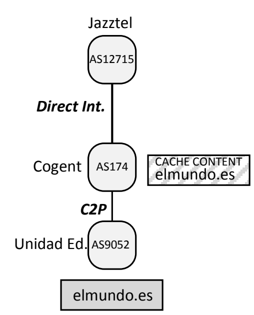

Fig. 5. Multiple content delivery strategies for web sites.

4.2 Revealing Interconnections Inside the Access ISP: The CDN

When seeking evidence of whether a content provider uses a CDN solution, Mercury provides a number of statistical indicators that are adequate to this task: (i) the number of AS hops to reach a server, (ii) the list of destination countries, and the list of (iii) IPs and (iv) ASes for a same URL. The number of AS hops to reach a CDN destination is a weak parameter of CDN existence, but it reveals the location of the content server relative to the user (when zero, the server is insider the user ISP). The list of destination countries is a better indicator of CDNs when the geolocation service is accurate, giving us the distribution of the server infrastructure. However, in practice, we must be careful because global CPs do not publish the location of their servers. For example Google servers are geolocated only in the USA, based on the registration of their AS. Currently, a lot of research effort is invested into improving the geolocation of IP addresses [6,1]. The list of destination IPs and ASes are the strongest indicators of a CDN. They show that a web site is deployed along multiple distributed servers, confirming the existence of some type of load balancing or caching technique, which are intrinsic to the use of CDN solutions. In addition, we can use them to determine the taxonomy of CDN strategies. As illustrated in Fig. 6, we can distinguish between two main types: (i) CDNs that host their cache servers inside the Tier 3 ISPs (Akamai and Google strategy) and (ii) CDNs that locate their cache servers in vantage points (VPs), near to the access operator (Cogent or Level3 strategies).

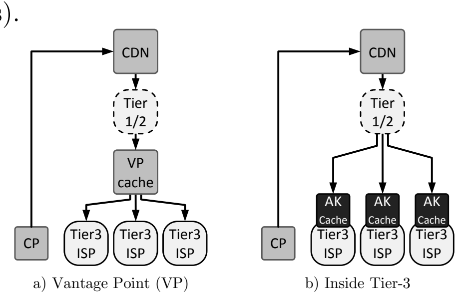

Fig. 6. Types of CDN architectures.

Table 2 summarizes the Mercury data for several CDN destinations. Our results indicates that most of the web sites we analyzed resort to some type of CDN. We observe that many global CPs like Facebook or Microsoft use Akamai, which deploys servers in both access ISPs and Tier 1 carriers, in addition to their own VPs. Furthermore it is also interesting that Elpais, the second major press group in Spain, also uses Akamai. The number of servers deployed by Akamai in other ASes stands well above the other CDNs, something observed by previous research publications [18]. Google uses a similar strategy and has direct interconnections with most of the access ISPs and has servers inside some access ISPs (Fig. 7 shows that Google has cache servers pointing to Jazztel IP addresses).

The remaining CPs use VPs to host their cache servers, based on different strategies. LinkedIn uses a multi-vendor CDN solution formed by Level3, Lime-

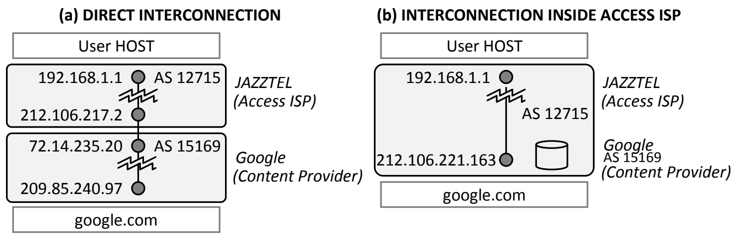

Fig. 7. Google places cache servers in vantage points and within the access ISP network

Light and others. This case is quite interesting because LinkedIn diversifies the spending in content delivery services. Similarly, the Spanish press groups El-Mundo and 20Minutos use a similar multi-vendor strategy. For Amazon CDN, we highlight the large number of servers and that it resells CDN services to other web sites like Instagram.

Table 2. CDN strategies for different content providers.

CP	Google	Facebook	Yahoo	Amazon	MSN	Instagram
CDN	Google	Akamai	Yahoo	Amazon	Akamai	Amazon
# servers	43	5	4	85	4	12
Inside ISP	✓	✓	–	–	✓	–
VP	✓	✓	✓	✓	✓	✓
Multi-vendor	–	–	–	–	–	–
CP	Elmundo	Elpais	LinkedIn	20minutos	Wikipedia	
CDN	Cogent Interroute	Akamai	Limelight Level3 ...	Level3 Fujitsu ONO	Wikipedia	
# servers	41	12	4	4	1	
Inside ISP	–	✓	–	–	–	
VP	–	✓	–	–	✓	
Multi-vendor	✓	–	✓	✓	–	

5 Conclusions

In this paper we introduced Mercury, an Internet measurement platform that aggregates traceroute measurements from multiple locations and analyzes the AS interconnection relationships along a network path. Mercury stands-out over other solutions because it discovers the end-to-end network path at the AS-level, while including information about the AS relationships, detecting IXPs and adding geolocation. We evaluate Mercury for a set of web sites using clients located at major Spanish access ISPs. Our results reveal the existence of many direct AS interconnections between access ISPs and content providers that are hidden for other methodologies. This suggests that some access ISPs and CPs find this interconnection strategy more attractive over using an intermediary and it confirms the trend that CPs are increasingly closer to the end users. In addition, our results emphasize the interconnection degree of the Spanish market, which is relatively high due to the small number of AS hops between access ISPs and CPs. Finally, Mercury provides indicators for discovering the architecture of CDNs and we successfully identify different content delivery strategies used by many web sites. Mercury detects Akamai and Google servers inside some Spanish ISPs which demonstrates the interest of these companies in offering a high quality content delivery.

Acknowledgments: This work was partially supported by the Spanish government, through the project CISNETS (TEC2012-32354).

References

1. B. Ager, W. Mühlbauer, G. Smaragdakis, and S. Uhlig. Web content cartography. In *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*, IMC '11, pages 585–600, New York, NY, USA, 2011. ACM.
2. Alexa. Alexa Top Sites. `http://www.alexa.com/topsites`.
3. B. Augustin, B. Krishnamurthy, and W. Willinger. IXPs: mapped? In *Proceedings of the 9th ACM SIGCOMM Internet Measurement Conference (IMC)*, 2009.
4. Augustin, B. and Cuvellier, X. and Orgogozo, B. and Viger, F. and Friedman, T. and Latapy, M. and Magnien, C. and Teixeira, R. Paris traceroute. `http://www.paris-traceroute.net/`.
5. CAIDA. The CAIDA AS Relationships Dataset. `http://www.caida.org/data/active/as-relationships/`, July 2012.
6. M. Calder, X. Fan, Z. Hu, E. Katz-Bassett, J. Heidemann, and R. Govindan. Mapping the expansion of google's serving infrastructure. In *Proceedings of the 2013 Conference on Internet Measurement Conference*, IMC, 2013.
7. H. Chang, R. Govindan, S. Jamin, S. J. Shenker, and W. Willinger. Towards Capturing Representative AS-level Internet Topologies. In *Proceedings of ACM SIGMETRICS*, pages 280–281. ACM, 2002.
8. K. Chen, D. R. Choffnes, R. Potharaju, Y. Chen, F. E. Bustamante, D. Pei, and Y. Zhao. Where the Sidewalk Ends: Extending the Internet AS Graph Using Traceroutes From P2P Users. In *Proceedings of the 5th International Conference on Emerging Networking Experiments and Technologies*, 2009.
9. X. Dimitropoulos and G. Riley. Modeling Autonomous-System Relationships. In *Proceedings of the 20th Workshop on Principles of Advanced and Distributed Simulation (PADS)*, pages 143–149. IEEE Computer Society, 2006.
10. Y. He, G. Siganos, M. Faloutsos, and S. Krishnamurthy. A systematic framework for unearthing the missing links: Measurements and impact. In *Proceedings of 4th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, pages 187–200. USENIX, 2007.
11. Z. M. Mao, J. Rexford, J. Wang, and R. H. Katz. Towards an Accurate AS-level Traceroute Tool. In *Proceedings of ACM SIGCOMM*, 2003.
12. MaxMind. GeoLite Databases. `http://dev.maxmind.com/geoip/legacy/geolite`.
13. D. Meyer. University of Oregon Route Views. `http://www.routeviews.org/`.
14. R. V. Oliveira, B. Zhang, and L. Zhang. Observing the Evolution of Internet AS Topology. *ACM SIGCOMM Computer Communication Review*, 2007.
15. PeeringDB. PeeringDB. `https://www.peeringdb.com/`.
16. Philip Smith, Cisco Systems. BGP Routing Table. `http://thyme.apnic.net/`.
17. Y. Shavitt and E. Shir. DIMES: Let the Internet Measure Itself. *ACM SIGCOMM Computer Communication Review*, 35(5):71–74, 2005.
18. A.-J. Su, D. R. Choffnes, A. Kuzmanovic, and F. E. Bustamante. Drafting Behind Akamai: Inferring Network Conditions Based on CDN Redirections. *IEEE/ACM Transactions on Networking (TON)*, 17(6):1752–1765, 2009.