On The State of Interdomain Routing in Africa

Roderick Fanou
IMDEA Networks Institute
Universidad Carlos III de Madrid (UC3M), Madrid, Spain
roderick.fanou@imdea.org
Thesis Tutors: Albert Banchs, UC3M and Pierre Francois, IMDEA Networks Institute

ABSTRACT

With IP networking booming on the African continent, initiatives to promote BGP peering in the region emerge, and changes in the transit behavior of ISPs serving Africa are expected. However, there is little knowledge on the IP transit topology currently forming the African Internet. Enhancing the RIPE Atlas measurement infrastructure, we shed light on both v4 and v6 topologies interconnecting ISPs based on the continent. We deployed 21 RIPE Atlas probes in 15 ASes located in 11 African countries, and performed measurements among a total of 214 probes scattered throughout the continent. Our results emphasize the dominance of ISPs based outside Africa for the provision of intra-continental paths, as well as the impact of this aspect on AS path length and end-to-end delay. They also reveal that South Africa is being adopted as a hub, thanks to its well connected infrastructure. We show a large variety of transit habits, depending on factors such as the language, the currency or the geographic location of the ISP. We finally highlight the emergence of new IXPs in Africa.

Keywords

RIPE Atlas, IP transit, African Internet, IXP

1. INTRODUCTION

Major investments in optical fiber infrastructure have been realised in Africa during the last decade. The excessive cost and low Quality of Service (QoS) bound with satellite communications, as well as the rapid expansion of mobile communications requiring transport infrastructures explain this evolution [1, 2, 3]. Despite such investments, there is still a considerable gap in the quality of Internet access in comparison with developed regions [4]. As a matter of fact, Internet access is perceived as of low quality, characterized by high latency and low bandwidth, while both the installation and usage costs paid by customers are high [5, 6].

In this context, the Internet Society, based on a preliminary study for Kenya and Nigeria, reported in [7] on the interest of setting up Internet eXchange Points (IXPs) to improve QoS on the continent and decrease costs for ISPs. Initiatives such as the AXIS (African Internet Exchange System) project or the CAB (Central Africa Backbone) have been established to respectively facilitate the creation of peering links through IXPs and improve the highly fragmented status of the physical infrastructure, through the deployment of intra-regional fiber [3, 8].

Obtaining relevant topological data for Africa is essential to understand its current state and observe its foreseen evolution [11]. However, little is known about the current state of the Internet topology in Africa despite extensive research performed on Internet topology discovery. African countries are indeed under-represented in existing Internet measurement projects. Recent work related to the African Internet such as [9, 6, 10] relied on a very limited set of vantage points, and had a different focus, as explained further in this paper.

In this thesis, we set as key milestone to obtain an interdomain map that is not biased towards the South African perspective, using probes scattered all over the continent. To this end, we met ISPs in Benin, Burkina Faso, Congo, Ghana, Ivory Coast, Mauritania, Morocco, Niger, Nigeria, Senegal, and Togo, in order to deploy RIPE Atlas probes within their networks. As a host can use any of the 6750 probes of the RIPE Atlas infrastructure [12], we could perform active measurements using v4 and v6 probes under our control and those already deployed in other countries. To obtain insights on the evolution of the peering ecosystem, our measurement campaign covers a period of 6 months, monitoring end-to-end paths every 3 hours on average.

We discover quite diverse ISPs transit habits. We observe that such habits are not only correlated with the location, but also with the official language, and the currency used in the country. Our results also illustrate that, with the exception of ISPs based in South Africa, the provision of intra-continental paths is dominated by ISPs based outside Africa. They however emphasize that South Africa is being adopted as a hub for intra-African communications. We further study the impact of such aspects on the end-to-end delay among the monitored ISPs, notably among networks based in the same country. Finally, we detect the emergence of new IXPs.

The remainder of this document is structured as follows. Section II discusses related work and summarizes previous results. Section III provides a description of the African interconnection landscape with a comparison to other continents. Section IV presents our measurement infrastructure and our methodology, while Section V exposes and analyzes our results. Section VI concludes the work and describes our plans for further research.

2. RELATED WORK

An extensive amount of research has been carried out on the discovery of the Internet topology, both at the router level and the AS level [13, 14, 15, 16, 17].

CAIDA has a long history running Internet measurement platforms. Its latest active measurement platform, Archipelago, has the goal of reducing the effort needed to develop and deploy sophisticated large-scale measurements [18]. Of the 94 Archipelago monitors deployed, only 5 were deployed in Africa. For this study, a larger base of deployed vantage points was needed, so Archipelago was not used.

In [19], Spring *et al.* analyzed with Rocketfuel a topology dataset issued from RouteViews BGP table dumps, combined with traceroutes performed by 750 vantage points targeting 10 ISPs in US. A key aspect of this work is the targeted analysis of a restricted set of ISPs instead of trying to map the whole Internet. We follow the same focused approach, targeting the African continent.

The PingER project [20] aims at measuring Internet end-to-end performance, and was notably used to quantify the digital divide. It involves 89 monitors and 1090 remote monitored nodes at 956 sites in 169 countries. But, among the 46 African countries involved, only Burkina Faso and South Africa host a monitoring site.

Some authors have focused on improving the knowledge of the African Internet [21, 22, 6] with restricted focuses such as the performance of broadband access in South Africa or the Nigerian telecom market. Gilmore *et al.* mapped, in [9], both the router level and AS level graphs of intra-African traffic. Traceroutes from South Africa towards all the IPs allocated by AFRINIC were performed for a week. To improve their results, they used the RIBs of routers in the South African Tertiary Education Network for link inference. They could extract from these BGP feeds links among ASes registered in AFRINIC and those towards their direct neighbouring ASes. As a consequence, they obtained one way paths from which they infered a tree, of which South Africa is the root. They acknowledged that the link density might look different if the traceroute probes were sent out from other countries in Africa. Moreover, they used the Maxmind database to perform both IP to AS and IP to location lookup. They did not analyze the content of the traceroute results or the BGP data, but focused instead on their visualization.

In contrast, we perform paris-traceroutes [23] to avoid the biases of classic traceroutes [24] results. We run these tests among an extensive amount of probes, covering a relevant number of countries in various African regions. For geolocation, we use and validate an algorithm which selects the most accurate router geolocation within 6 databases (DBs) and we adopt the Team Cymru database, more accurate for IP to AS lookup. We also provide insights on the evolution of the African Internet topology, instead of a snapshot.

Recently, Gupta *et al.* investigated Internet connectivity between Kenya, Tunisia, and South Africa [10], by performing traceroutes from access networks to sites hosting popular content. They noticed that 66.8% of paths from their vantage points towards Google cache servers located in Africa leave the continent. They generalized, from their results, the nature of intra-domain interconnectivity on the continent. Nevertheless, these are countries in which broadband access networks are more developed [6] in comparison with most African countries, so that, as acknowledged by the authors, the obtained dataset may not reflect the nature of the paths from other countries.

Our study instead presents discoveries of the Internet infrastructure based on measurements performed from access to access networks, as we aim at studying how Africans communicate with one another. We perform these measurements among a large variety of networks and for a long enough period of time to study the dynamics of the African Internet topology. The paths in our dataset are typically not seen in RouteViews, RIS and PCH datasets, as these do not host monitors in the studied regions.

By exhibiting newly established IXPs located in Africa, as well as the use of other ones, we show that ISPs do peer now in Africa, illustrating the first benefits of the initiatives promoting peering.

3. EVOLUTION OF THE AFRICAN TELECOM INFRASTRUCTURE

In the early 60's, the gateway license of most African countries and a monopoly on telecom markets was usually attributed by governments to incumbent operators which operate phone networks. Since the 90's, the markets have been liberalized everywhere, except in Djibouti and Ethiopia. Over the last years, many mobile operators and ISPs were therefore created. Besides, telecom operators invested in both domestic long haul and intercontinental optical fiber deployments [1, 3, 2, 25] to reduce their reliance on satellites links. Meanwhile, new backbone operators were formed through merging utility infrastructure with the incumbent (e.g. Tanzania), government owned backbones (e.g. Kenya), or public/private joint ventures (e.g Angola Cables). Most attempted to establish cross-border interconnection with neighbouring countries.

As a consequence, the physical transport infrastructure is now mainly based on submarine and terrestrial cables, but it is still fragmented. Central Africa and the Sahel are the main gaps on the map, isolating the other main areas of connectivity [3].

Nowadays, operators are actively following the recommendations to improve the level of intra-regional connectivity. The number of IXPs in Africa has considerably increased, from 8 in 2008 [26] to 27 in 2014 [27, 28]. However, this number represents only 5% of the 435 IXPs globally established by February 2014. For comparison, there are 46 IXPs operating in Latin America and the Caribbean, 88 in North America (US and Canada) and over 130 in Europe [29].

Despite this positive evolution, African IXPs are still not well spread on the continent. Only about a third of African countries host an IXP [29]. Moreover, the IXPs are relatively small, compared to those on other continents, with an average of 14 member ISPs. The largest IXPs are in South Africa.

Our objective is to monitor the evolution of this interconnection landscape, at the IP level, with the ultimate goal of providing ground for the definition of potential improvements.

4. METHODOLOGY

In this section, we describe the approach that was followed to identify ISPs playing a key role in the IP transit between any couple of Autonomous Systems (ASes) hosting a probe, and analyze the impact of the characteristics of the observed inter-domain paths on end-to-end delay. We also explain how loops in the infered AS paths were dealt with, as well as the approach followed to discover new peering links and IXPs.

4.1 Measurement Infrastructure

Multiple challenges influenced our choice for the measurement infrastructure. First, whereas network operators are frequently reticent to the intrusion of foreign devices, for legitimate security and privacy reasons, we have to find a relevant number of hosting locations for the measurement devices. Second, the physical devices have to be robust, as power outages and surges occur very often in the studied countries. Third, the devices cannot be expensive, as we have no guarantees that all our collaborators will keep them online. Finally, we wanted to deploy devices that are useful for others, notably network operators and researchers, for the study of African networking.

These requirements lead us to chose the RIPE Atlas measurement platform, which consists of over 6700 devices deployed worldwide in a wide variety of locations [12]. Any individual wanting to host a RIPE Atlas probe can do so. The probes are free, secure and easy to install. They consume low bandwidth, do not need maintenance, and are robust against frequent power outages [30, 31, 32]. These properties of the RIPE Atlas project make it well suited to satisfy our objectives. However, as of June 2013, Africa only contained a handful of active devices, with almost no deployment in West Africa.

4.2 Data Collection

We deployed, mostly in West Africa, 21 RIPE Atlas probes in 15 ISPs networks, starting in July 2013. These devices are hosted by either ISPs, universities, or home networks. We make sure that none of them are behind a wireless access link to reduce the impact of last mile latency on our results.

RIPE Atlas ambassadors and institutions such as AFRINIC and ISOC simultaneously deployed a considerable amount of probes in Southern and Eastern regions as well. Thanks to these efforts, up to 282 active probes are now hosted in 129 ASes and spread across 42 countries all over Africa [12].

We used paris-traceroute [23] for all our measurements to discover as many paths among probes as possible, and not suffer from inconsistent results caused by load balancing, as happens with classic Jacobson's traceroute [24]. We forced the probes to perform 16 independent traceroutes using different paris_id values, every time a path is being measured. Note that we used the UDP variant of traceroute, to avoid the potential bias caused by differentiated traffic handling of ICMP packets.

We conducted 3 measurement campaigns. During the 1st one, we performed paris-traceroutes from each active probe in Africa to all the others with a period of roughly 3 hours, starting on November 30, 2013 and ending on April 06, 2014. It results in 675,421 traceroute outputs. Since v6 probes were only deployed in Southern Africa, we carried out similar measurements among both v4 and v6 probes in the region, starting on June 01, 2013 and ending on August 01, 2014, to compare v4 and v6 routing infrastructures. It results in 408,383 v4 and 21,744 v6 traceroute outputs. To highlight the launch of the Serekunda IXP of Gambia, we finally performed hourly, during the 2nd week of August 2014, 3,161 publicly available mesh paris-traceroutes among RIPE atlas probes in Gambia [33].

CC	Country	ASes	Used	% ASes	% IPs
AO	**Angola**	36907, **17400**	2	6.1%	4.8%
BJ	*Benin*∗	37090, 28683, 37292	15	37.5%	73.2%
BF	*Burkina Faso*∗	25543, 8513, 37073	4	28.6%	64.9%
BW	**Botswana**	*14988*, **37678**	3	11.1%	81.7%
CI	*Ivory Coast*∗	36974, 29571	3	16.7%	68.8%
CM	Cameroon	16637, 15964	2	7.7%	32.9%
ET	Ethiopia	24757	2	50%	33.3%
GA	Gabon	16058	1	11.1%	81.2%
GH	*Ghana*∗	30988, 29614, 37140	2	6%	19.5%
GM	*Gambia*	37309, 37524, 327719, 37323, 25250	5	71.4%	80.8%
KE	Kenya	12556, 37061, 15399	4	3.9%	5.5%
LS	**Lesotho**	**37057**	1	10%	68.5%
MA	Morrocco∗	30983, 6713	2	25%	61.6%
MG	**Madagascar**	37054, 37608	3	25%	48.8%
MR	Mauritania∗	8657	1	33.3%	24.6%
MU	**Mauritius**	37006, 37100, 23889, 30844, 327681, 3215	10	12.5%	80.5%
MZ	**Mozambique**	30619, 42235, **31960**, 6939∗	4	37.5%	8.9%
NA	**Namibia**	36996, **33763**	4	13.3%	31.1%
NG	*Nigeria*∗	30988, 30984	3	1.5%	0.9%
NE	*Niger*∗	37205, 37385	4	28.6%	33.3%
RW	Rwanda	37228, 37006	2	12.5%	66.5%
SC	**Seychelles**	36867, *36958*, 36902, **37343**	20	50%	34.7%
SD	Sudan	37197	1	14.3%	4.1%
SN	*Senegal*∗	8346, 37196	4	66.7%	76.8%
SZ	**Swaziland**	**19711**		16.7%	68.6%
TG	*Togo*∗	30982	1	50%	41.4%
TN	Tunisia	2609	2	10%	27%
TZ	**Tanzania**	37045, 36909, *37084*, **37182**, **33765**	4	10.4%	24.1%
UG	Uganda	37063	2	2.9%	12.3%
ZA	**South-Africa**	32653, *10474*, **36877**, 37542, *2018*, **37172** **12258**, **6968**, 33762, 37497, *37520*, *3741* **11845**, **37618**, 37403, 36937, 37457, *6083* **37358**, 5713, 16637, 22355, *37105*, 18931∗ 37251, 29975, 37253	100	7.8%	40.2%
ZM	**Zambia**	*37043*, **37154**, **30844**	2	18.8%	5.9%
ZW	**Zimbabwe**	***30844***	1	6.2%	3.2%

Table 1: *Involved Probes and ASes*

Table 1 summarizes the geographical and networking spread of the probes used in our study. Before filtering, our raw data involves 214 probes hosted in 90 ASes operating in 32 African countries (Table 1). We add the symbol $*$ to ASes hosting only v6 probes and put in italics those hosting both v4 and v6 probes. ASes that participate to only the second campaign are in bold. Besides, we put Southern African countries in bold while Western African ones are in italics. We also add the symbol $*$ to countries in which operate ASes hosting our deployed probes. In the rest of this paper, we use ISO 2-letter country codes (CC) (Table 1).

We compute the percentage of ASes allocated by AFRINIC [34] covered per country as well as the percentage of IPv4 space of each country covered by ASes hosting the probes in order to give an overview of the granularity of our results. Note that, for instance, in SN, TN, ZM and ZW, we consider ASes 8346 (Sonatel), 2609 (Tunisia Backbone) and 30844 (Liquid Telecom) of local operators while computing these percentages, although these ASes have been allocated by other Regional Internet Registries (RIRs).

4.3 Dataset Treatment

4.3.1 IP to Country Code Mapping and Validation

Geolocation is said to be of poor quality, especially for IPs located in Africa [35]. Although there are many public datasets, none of them is fully accurate. To geographically locate the 8328 v4 and 465 v6 public IPs found in the traceroute data, we compare 6 public databases (DBs) and select the most accurate entry, in a context where these DBs can be inconsistent. We then validate our results by a 2nd methodology based on delay measurements.

Because publicly available information on geolocation of Internet infrastructure is not of high enough quality for our purposes, we turn to using data from OpenIPMap (*OIM*), which is a project aimed at obtaining city-level accuracy of Internet infrastructure, by crowdsourcing this information from network operators and other interested parties. Currently the *OIM* project has 14 contributors, mostly operators.

When location information is not available in *OIM*, we perform Reverse DNS (RDNS) lookups for the IP addresses in our traceroutes. As many network operators embed location information in hostnames, for instance airport codes or abbreviated city names, we deduce location from these hints. For example, 'xe-3-2-1.**was14**.ipv4.gtt.net.' corresponds to an AS3257 (TINET, DE) router located in Washington (US), 'if-4-1-2.core2.**COV-Cochin**.as6453.net', to an AS6453 (TATA, US) router located in Cochin (IN) while 'be2321.ccr22.**ams03**.atlas.cogentco.com.', to an AS174 (Cogent, US) router located in Amsterdam (NL).

We then select, like [9], the v4 and v6 versions of Maxmind GeoIP2City DB (*MM*) [36]. However, *MM* is more oriented to end-users geolocation and less accurate for routers. We therefore choose the Team Cymru (*TC*) DB [37] whose data is obtained directly from the RIRs. But, performing the IP to country lookup with *TC* could also introduce biases in the geolocation since an AS could span many countries.

To differentiate IP addresses in Africa from those on other continents, we use AFRINIC database (*AF*) gathering AFRINIC allocations and assignment files recorded from 2005 to 2014 [34]. Although Whois DB (*Whois*), obtained from registration of IP ranges to RIRs, is considered as not up-to-date by previous work [4, 35, 13], we retain it as well as a last resort solution when no other DBs contains an entry for a given IP.

We compare in Table 2 the entries of the selected DBs. *TC*, *Whois* and *MM* locate most IPs. Any entry of these DBs which is not a valid CC is ignored ('EU', 'AP', 'ZZ', 'A1', 'A2', etc). This rule leads us to ignore, among others, all the available entries for 133 IPs which belong to IP ranges of either AS702 (UUNET, US), AS1273 (CW, GB), AS12995 (TELIANET, SE), or AS6774 (BELGACOM, BE). Compared to the other DBs, *OIM*, *AF* and *RDNS* have few percentages of entries, but never contain non-valid CCs.

	IPv4 entries			IPv6 entries		
DB	**Cov.**	**Inv. CC**	**Trust**	**Cov.**	**Inv. CC**	**Trust**
OIM	26%	0%	96.8 %	30.1%	0%	94.9%
RDNS	56.7%	0%	93%	46.7%	0%	78.5%
MM	83.9%	16%	72.7%	99.1%	0%	71%
TC	86.7%	13.3%	69.4%	99.1%	0.9%	79.7%
AF	36.2%	0%	93%	56.7%	0%	84.4%
Whois	85.6%	11.8%	66.1%	43.2%	0%	67.5%

Table 2: *DBs entries comparison*

In a first step, we rank the databases according to our intuition of trustworthiness. Since *OIM* is specifically aimed at accurate router geolocation, we consider its data as most trusted, if available. We find otherwise the geolocation of the remaining IPs among other DBs classified in the order *RDNS*, *MM*, *TC*, *AF*, *Whois* (Algorithm 1). It allows us to tie break *country conflicts*, cases in which CCs provided by DBs are different. We locate 98.4% v4 and 100 % v6 IPs.

To validate our geolocation results, we perform twice a set of ping (16,203 v4 and 776 v6) measurements [33] to IPs for which at least 1 valid entry is available among all DBs. We ping each IP from up to 10 random RIPE Atlas probes hosted in each possible country given by the DBs. 81.2% of v4 IPs (resp. 92.2% of v6 IPs) answer to our pings. We compute for each IP the minimum delay recorded per possible country. To deduce the location from this latency-based method, we assume that the probe is in the country from which the minimum delay is the lowest one.

We then compare the results of both methods. The percentages of locations successfully found by each DBs, while executing our algorithm, are listed, for both v4 and v6, in the "Trust" columns of Table 2. It appears that our algorithm finds in total 95.1% of those obtained by the validation methodology. We also notice that the minimum delay from a retained country, in cases of *country conflicts*, is often close to those in which all available entries in DBs are identical and correspond to that country. We could therefore update 66 v4 (resp. 7 v6) entries and insert 6,161 v4 (resp. 325 v6) entries in *OIM* among which we denote 2,982 v4 (resp. 200 v6) IPs located in Africa. By adding this data to *OIM*,

Algorithm 1: IP to Country Mapping Algorithm

Data: *OIM, TC, MM, Whois, AF*, and *RDNS* (any DB entry which is not a Country Code is ignored)
Input: IP
Output: Country Code *CC* of IP
/* *OIM*: 2167 v4 && 140 v6 IPs */
if *CC of IP in OIM* **then** Return *CC* ;
/* *RDNS*: 3303 v4 && 105 v6 IPs */
else if *CC of IP in RDNS* **then** Return *CC* ;
/* *MM*: 2634 v4 && 220 v6 IPs */
else if *CC of IP in MM* **then** Return *CC* ;
/* *TC*: 47 v4 && 0 v6 IPs */
else if *CC of IP in TC* **then** Return *CC* ;
/* *CC*: 0 v4 && 0 v6 IPs */
else if *CC of IP in AF* **then** Return *CC* ;
/* *Whois*: 46 v4 && 0 v6 IPs */
else if *CC of IP in Whois* **then** Return *CC* ;
/* *CC not found*: 133 v4 && 0 v6 IPs */
else Return *CC not found* ;

we significantly improve the publicly available geolocation information of African Internet resources, allowing further studies to built upon this. With the geolocation data, we finally compute the country path corresponding to the IP path of each traceroute output, as explained in [35].

However, we acknowledge that our validation method is not 100% precise. In case we have only probe(s) in a country A with high-latency first hops, where the bordering country B has probes with low-latency first hops, we could select country B, even when in reality the IP we look up is in country A.

4.3.2 IP to AS Lookup and Raw Data Sanity Check

Let us consider the following output as the result of a traceroute performed from ip_s towards ip_d.

$$P_{trace}(ip_s, ip_d) = ip_s, ip_2, ip_3, ip_4, ip_5, ip_6, ip_7, ip_8, ip_9, ip_d$$

We map with *TC*, public IP addresses of each traceroute output into ASes, as explained below.

$$Mapping_{ip \to AS}(P_{trace}(ip_s, ip_d)) = \underbrace{ip_s, ip_2}_{AS_s}, \underbrace{ip_3, ip_4, ip_5, ip_6}_{AS_1}, \underbrace{ip_7, ip_8}_{AS_2}, \underbrace{ip_9, ip_d}_{AS_d}$$

$$Mapping_{ip \to AS}(P_{trace}(ip_s, ip_d)) = Path(AS_s, AS_d) = AS_s, AS_1, AS_2, AS_d, \text{where } AS_s \neq AS_1 \neq AS_2 \neq AS_d$$

As for the raw data cleaning approach, we separate all the obtained AS paths for which source and destination ASes do not correspond to the known ASes hosting the probes and we retain the others. We then make a second check, keeping in a data structure the AS adjacencies of the retained ones. If we find, in a traceroute that does not reach its target, an immediate upstream of the source or the destination AS, learned from the retained traceroutes, we still consider the said traceroute result. But we use it only for AS adjacencies and paths inference

Second, we keep track of either all the IPs for which there is no entry in *TC*, or IP hops at which the router do not reply [13] when these issues only occur within the AS paths i.e. between source and destination ASes. We refer to them, in the paper, as respectively *unresolved* and *unknown* ASes. In fact, the hop may belong to an AS different from the neighbouring ones. Note that if an AS path B is identical to another one A, except 1 *unknown* or *unresolved* AS, we assume the missing AS in B is the one corresponding to the AS with the same position in A.

After that, we compress AS paths by suppressing inexistent adjacencies. If a path contains, for instance, twice the same AS separated by *unresolved* or *unknown* ASes, we replace this sequence by the AS itself. Besides, an AS sequence of x consecutive times the same AS is replaced by that AS. Actually, we infer an edge from AS1 to AS2 only when in the AS path, AS1 and AS2 are known and consecutive. At this stage, all the paths with a length lower than 2 are discarded as well. We then look for load-balancing cases.

Next, we identify and separate from the remaining traceroutes, those whose AS paths contain loops. We do not consider those containing IXPs and sibling ASes, i.e. ASes belonging to the same organisation [38, 39, 28]. Indeed, an AS sequence containing such ASes can make us identify the paths as with an AS-loop. For the 1st campaign, we find 4,648 cases: among others, AS3356 (Level3, US) causes 32% while AS37282 (MainOne, NG) and AS37054 (Data Telecom Service, MG) respectively 15% and 11%. For the 2nd campaign, we find 3,284 v4 cases: among others, AS3741 (Internet Solutions, ZA) causes 20%, AS12258 (MWEB, ZA) 12% while AS16637 (MTN SA, ZA) and Data Telecom Service respectively 9% and 8%. We also find 55 v6 cases: the only ASes causing loops are AS37100 (Seacom, MU) and AS10474 (MWEB, ZA) with respectively 60% and 40%. In the last campaign, we do not find any loop.

By the end of this raw data cleaning method, we reject for the 1st campaign only 82,292 v4 traceroutes, while for the 2nd one, we discard 11,147 v4 traceroutes and 2,149 v6 traceroutes. For the last campaign, we discard 413 outputs. The dataset resulting from this filtering process comprises paris-traceroutes outputs from 181 probes located in 30 African countries and hosted in 90 ASes. The v4 measurements involve 181 probes hosted in 88 ASes operating in 30 countries, while the v6 ones, 19 probes in 13 ASes hosted in 5 countries. Furthermore, our dataset contains 1,862 pairs of v4 ASes and 116 pairs of v6 ones.

Finally, we infer both the delay at the last IP hop and the delay from the source to the destination AS, as the difference between the one at the egress point of the AS source and the one at the ingress point of the AS destination.

5. RESULTS

In this section, we first analyze the biases and the completeness of our dataset. We then investigate the dynamics of the observed paths. Next, we show that, mostly in West Africa and contrarily to Southern Africa, a large number of intermediate ASes are found on paths between geographically collocated ISPs. We highlight the dominance of ISPs based outside Africa to provide inter-domain connectivity among studied ASes, except those in South Africa. We for example discover that most of the paths from, towards and within the CFA monetary region, go through Orange France Telecom. We then illustrate the impact of the intercontinental aspect of paths on the RTTs among African ISPs. We emphasize that South Africa, thanks to its well connected infrastructure, is being adopted as a hub for a very relevant proportion of v4 and v6 intra-African communications. We finally detect new peerings and IXP infrastructures on the continent.

5.1 Result accuracy

5.1.1 Caveats and Limitations

We acknowledge all the probes were not deployed at the beginning of our 1st measurement campaign. The constant evolution of the RIPE Atlas infrastructure on the continent leads us to daily add new probes to the set of probes that we use. Besides, our dataset involves 7.2% of the ASes allocated by AFRINIC. ASes hosting our probes partially cover IP ranges allocated per country (Table 1). Finally, we should acknowledge the shortcomings of IP to AS mapping. For instance, 40.6% (resp. 35.9%) of the unique v4 (resp. v6) AS paths contain at least one either *unknown* or *unresolved* AS, as defined in the previous section.

5.1.2 Dataset Completeness

We compare our dataset to the set of paths (announced or accepted by the 90 studied ASes) available in RouteViews, RIPE RIS, PCH datasets [39] collected by route servers from 2013 to 2014. We only consider the 2,529 v4 (resp. 91 v6) paths containing no *unknown* or *unresolved* ASes. Our dataset is more precise as it contains end-to-end African paths contrarily to these public datasets: most of the routes collectors are hosted outside the continent. Among the 960 v4 (resp. 63 v6) AS adjacencies we infer from the discovered paths, 733 v4 (resp. 35 v6) are not visible in these public datasets. Quite intuitively, 2,519 v4 (resp. 79 v6) paths are not visible in RouteViews, RIS, and PCH either. Most of the AS paths found in both datasets actually belong to the set of those originated by AS3741 (Internet Solutions, ZA): this ISP hosts a RouteViews collector. Besides, most of the AS adjacencies found in both datasets are those among ASes based outside the continent.

5.2 Path dynamics

We discuss in this section the dynamics of AS paths used to serve communications among AS pairs. To this end, we identify the unique AS paths for any two ASes and compute for each of them the frequency of its usage to serve the said AS pair. This value is the ratio between the number of times the unique path is observed for the AS pair and the number of usable traceroutes among these ASes. In the rest of this paper, we refer to the path with the highest frequency of usage from an AS to another, as the preferred path.

Figure 1 depicts, in ascending order of values, the frequency of the preferred path per AS pair. Only 4% of the v4 AS pairs have used their preferred paths with a frequency lower than 50% whereas no v6 preferred path has a frequency of usage lower than 50%. Meanwhile, about 72.6% (resp. 82%) of the v4 (resp. v6) preferred paths have been used with a frequency higher than 90%.

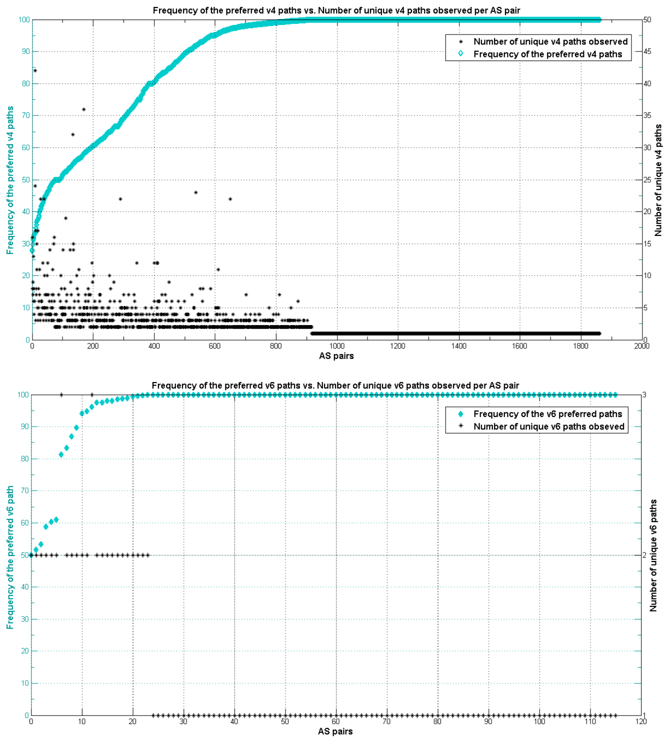

Figure 1: *Frequency of the preferred v4 and v6 paths vs. Number of unique paths observed per AS pair in the dataset*

Let us consider the 9.2% of the v4 AS pairs for which we observed more than 4 AS paths. We identify among them AS37090 (Isocel Telecom, BJ) and AS25543 (Onatel, BF) as being either the source or the destination in respectively 17% and 16.4 % of these pairs of ASes. 32% of the said AS pairs contain at least one of them. We notice the highest numbers of unique paths (42, 32, 24, 23, 22, etc) for pairs of ASes containing AS37090 as shown by Figure 2.

We validated these results by visiting Isocel Telecom, discovering that this ISP is constantly making traffic engineering in order to offer the best possible QoS to its customers. However, BGP route failures, flap damping, etc, could also be listed as the reasons of such changes. Note that most of

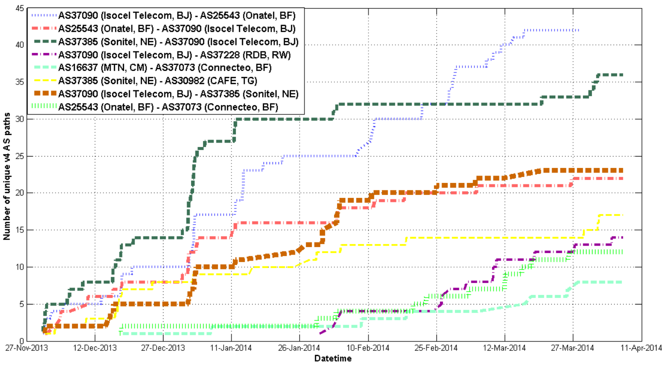

Figure 2: *Number of newly discovered v4 paths per hour for some AS pairs with a high AS path dynamics*

the curves of Figure 2 have their largest slopes between December 12, 2014 and January 11, 2014. At the time of this writing, we are still investigating the reasons of this bump.

5.3 AS path length distribution

We study the distribution of the length of v4 AS paths among pairs of ASes operating within Africa, within West Africa, as well as the one of v4 and v6 paths within Southern Africa. We also carry out a specific analysis for pairs of ASes located within the same country (Figure 3). To highlight the particularities of such distributions, we remove *unknown* and *unresolved* ASes from paths used for plotting the graphs of Figure 3, except those corresponding to paths length distribution within countries. Thus, the AS paths in those cases could be even longer than what is presented.

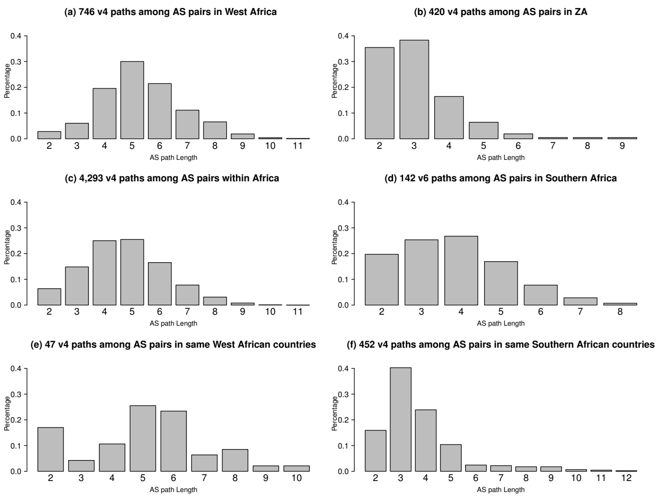

Figure 3: *Path length distributions*

On Figure 3(c), we show the AS path length distribution across the paths used by the 1,862 v4 AS pairs of the dataset. Since ASes in West Africa are based in geographically collocated countries, one could guess that paths would be shorter. But, there is only a few paths of length 2. Moreover, we find a low proportion of national paths going through at least 3 intermediate ASes. They connect, for example, AS37073 (Connecteo) with AS25543 (Onatel) in BF, AS37385 (Sonitel) with AS37205 (Atlantique Telecom) in NE.

As depicted in Figure 3(a), we however discover unusually long paths of 5 ASes on average in West African communications. In contrast, one can notice, when comparing Figure 3(f) to Figure 3(b), that Southern Africa contains more direct links and short paths, most of them within ZA.

As for the paths among ASes operating in the same African country, 40% have a length of 2. These peering or transit links are for example observed in BJ, SC, TZ, and ZA. Figure 3(b) and 3(f) proves that most of them are in Southern Africa and precisely in ZA. This confirms that focusing solely on mesurements from ZA does not provide a representative sample of Internet paths characteristics for the rest of the continent. The v6 path lengths (Figure 3(d)) are all lower than the maximum of 12 traversed by paths among ASes operating in Southern Africa (Figure 3(f)). The mode of the distribution is 4 and the percentage of paths of direct links is non negligible (20%).

5.4 AS centrality

In this section, we study the importance of the transit role of ASes found in the set of paths. We first characterize the 164 ASes in the dataset according to their geographical scope in the Internet. To provide insights for the different African regions, we classify ASes based on their region of operation. Category *WAf* includes ASes based in West Africa, category *SAf*, those in Southern Africa, category *EAf*, those in East Africa, while category *RAf* includes ASes operating in Africa but not based in any of the previous regions. Category *Int* (Intercontinental) includes all ASes based outside the continent. Note that we find 61 *Int* Ases.

To quantify the role of transit played by each ISP covered by this study, we define the concept of AS-centrality of an AS as the percentage of paths for which that AS is within an AS path. Mindly that the AS-centrality of an AS is only accounted for the paths in which that AS is neither the source nor the destination, as detailed in Equation 1. It is different from the well-know Betweeness Centrality metric since we only consider paths among studied couples of ASes, and not those towards ASes within the path, while computing the numerator and the denominator of Equation 1.

$$C(n)=\frac{\sum\limits_{\substack{s,d\\ s\neq d\neq n}} \delta_{(s,d)}(n)}{\#Paths(s,d)}, \text{ where } \delta_{(s,d)}(n)=\begin{cases} 1 \text{ if } n\in Path(s,d) \\ 0 \text{ otherwise} \end{cases} \tag{1}$$

We now precise our centrality analysis by defining the concept of "joint AS-centrality", which captures the centrality of tuples of ASes present together on the observed AS paths, as defined by Equation 2.

$$J_C(n_1,n_2,\ldots,n_k) = \sum_{\substack{s,d \\ \{s,d\} \not\subset \{n_1,n_2,\ldots,n_k\}}} \frac{\delta_{(s,d)}(n_1,n_2,\ldots,n_k)}{\#Paths(s,d)} \quad (2)$$

$$\text{where } \delta_{(s,d)}(n_1,n_2,\ldots,n_k) = \begin{cases} 1 \text{ if } n_1,n_2,\ldots,n_k \in Path(s,d) \\ 0 \text{ otherwise} \end{cases}$$

Figure 4 depicts the centrality of each AS, in the whole set of paths (blue curve), among West African networks (orange curve), and among Southern African networks (black curve).

We sort the ASes according to their centrality on the whole set of paths and represent them with different markers given the category they belong to. Note that we only plot ASes which play a non negligible role of transit in Africa, i.e. their AS-centrality is greater than the threshold 0.5%, leaving 37 ASes out.

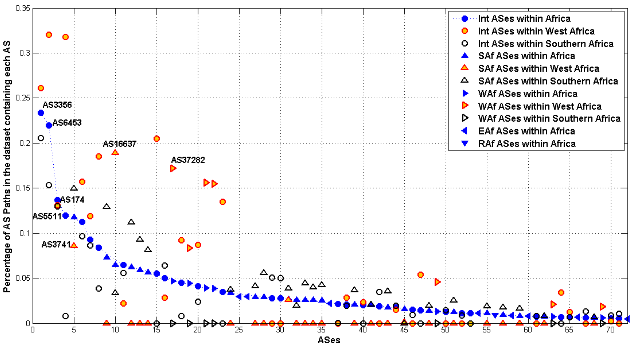

Figure 4: *AS-centrality. ASes are sorted according to their AS centrality within the African interdomain topology (blue curve).*

As indicated by the orange curve, the 4 most central ASes in our view of the v4 African interdomain topology are all intercontinental ones: AS3356 (Level3, US) with 23.4% of the 4,293 AS paths, AS6453 (TATA, US) with 22%, AS174 (Cogent, US), 13.6%, and AS5511 (Orange, FR), 12%. 65.2% of the AS pairs were served using at least one of these 4 ASes. It also appears that AS3741 (Internet Solutions, ZA) has a centrality of 11.6%.

Within West Africa, 57.1% of the ASes playing a role in IP transit are *Int* ASes. Orange seems to be a dominating ISP in the region, with an AS-centrality of 37.8%. TATA and Level3 respectively own 32% and 26.1% of the West African market share. 80.2% of the paths within West Africa are through at least one of these 3 *Int* ASes. We notice that a relevant percentage of paths (18.9%) connecting ASes in the region transit via AS16637 (MTN SA, ZA). The most central local AS is AS37282 (MainOne, NG), whose business objectives include providing pan-African connectivity to West African ISPs [40]. We find it on 17.2% of the West African paths. MainOne is also the most present local AS for the provision of paths among West Africa and the rest of the continent. However, we still notice the dominance of *Int* ASes in that category, with one of TATA, Level3, and Orange found on 76.42% of the paths.

The reliance on *Int* transit providers to serve communications is considerably lower in the Southern African region. Although the top 2 ASes remain the same as in Africa and Western Africa (Level3 with 20.5% and TATA with 15.3%), the AS-centrality of local ASes increases: Internet Solutions with 15%, AS5713 (SAIX-NET, ZA) with 12.9%, AS10474 (MWEB, ZA) with 11.2%. Still, Cogent is also present on 12% of the paths. We finally notice that the reliance on ISPs based on other regions and particularly West and East Africa in Southern Africa is insignificant.

In the v6 routing infrastructure, some ASes which are not involved in the v4 routing have a high AS-centrality. The top 4 ASes are mostly *SAf* ones: AS6939 (Hurricane Electric, US) with 23.9%, AS2018 (TENET, ZA) with 22.5%, AS30844 (Liquid Telecom, GB) with 21.1%. We denote that 19.7% of the v6 paths traverse the IXP AS5459 (LINX-AS, GB) while 14.7%, the IXP AS1200 (AMSIX, NL). Level3 and TATA are present on only 9.8 % and 9.1% of the paths. This could mean that operators deploying v6 routing tend to also give importance to peering.

5.5 Technico-economic insights on routing trends

As a further objective, we try to discover trends in the transit habits of African ISPs, depending on the official language (French vs. English), and the geographical location (coastal vs. inland) of the source and destination ASes. We also highlight differences given the currency used in countries they are operating in (XAF - Central African CFA franc/ XOF - West African CFA franc or not): with 14 countries, XAF-XOF is, similarly to EURO, one of the biggest formal monetary unions, making the region interesting from a cooperation perspective. We also distinguished, for the specific case of paths among West African ASes, the business history of the ISPs (historical state-owned vs. challenger). In this study, we consider the AS-centrality of TATA, Level3, Orange, as well as their joint AS-centrality.

As shown in Figure 5, some paths used for communications among French speaking, inland, and XOF/XAF countries contain all 3 ASes: they typically correspond to paths among collocated West African countries from Sonitel (NE) to Connecteo (BF) or AS30982 (CAFE, TG), from AS28683 (Benin Telecom, BJ) to Connecteo, and from AS37292 (OTI, BJ) to CAFE or Connecteo.

Many paths are also served by a combination of two of the considered *Int* ASes. French speaking countries mostly rely on Orange/France Telecom, which serves 17% of the West African AS pairs, without TATA nor Level3. Another 14.5% of AS pairs are served by Orange, jointly with TATA or Level3. Orange almost completely disappears from our African internetworking map when it comes to communications among English speaking countries. This phenomenon is even stronger when we analyze paths within based on monetary regions or not: 36.7% of paths among ISPs in XOF/XAF transit via Orange while 24.9%, for communications from and towards them.

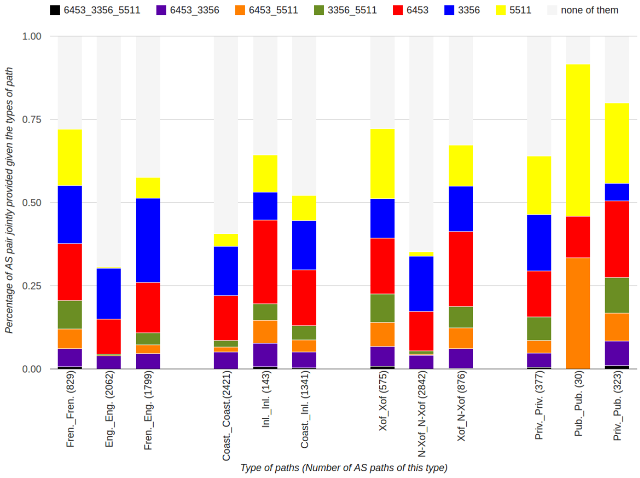

Figure 5: *Joint centrality of AS3356 (Level3, US), AS6453 (TATA, US) and AS5511 (Orange, FR) as a function of the type of path*

ISPs	Ownership	intra West-African towards ISP	intra West-African from ISP	intra African towards ISP	intra African from ISP
ISP 1	X ⋙ 50%	83.3%	26.66%	75%	26.66%
ISP 2	X ⋙ 50%	13.33%	75%	20%	91.66%
ISP 3	X ≫ 25%	53.33%	80%	63.6%	100%
ISP 4	N/A	50%	93%	50%	83.33%
ISP 5	N/A	29.9%	23.6%	50%	25%

Table 3: *Percentage of paths via Orange for both intra-West African and intra-African communications of partially owned ISPs*

We compute, for incumbent operators in West Africa partially owned by Orange, the percentage of paths that transit through its AS, in our view of the African interdomain topology (Table 3). This case highlights the impact of the business relationships among ISPs on their transit habits in the region.

To summarize, African French speaking countries are mostly served via Orange/France-Telecom while English-speaking countries mostly via TATA, with a notable presence of Level3 in paths among both types of countries. African inland AS pairs rely much more on TATA (54.17%) than on Level3 (29.16%), and not on Orange. Such differences may be explained by the scarcity of Internet transit offerings in inland countries.

As for communications among West African ISPs, we observe radically different transit approaches between historical ASes and challenger ones. Historical state-owned ASes of West Africa mostly rely on Orange, while the connectivity of challenger ASes is mostly established through Level3 and TATA.

Such statistics emphasize the widespread phenomenon of "intercontinental tromboning" that African ISPs face in their intra-continental communications. They also emphasize that there is no such thing as an "African Internet", since transit habits vary radically from region to region, and from ISP to ISP.

5.6 Country centrality

In this section, we answer to the following question: What are the countries lying on a relevant proportion of v4 and v6 paths among African ISPs? To this end, we consider all the country paths corresponding to the 128,312 v4 (1,207 v6) unique IP paths in the dataset and compute the country centrality, as defined in [35].

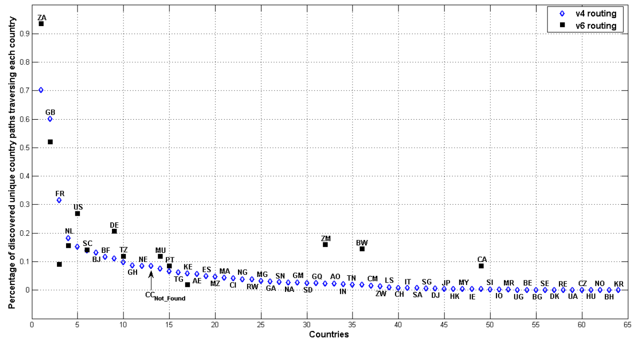

Figure 6: *Country centrality*

Figure 6 shows the countries involved in intra-African communications, and the amount of IP paths that pass through each of them. We notice that ZA appears on 70.3% of v4 (93.5% of v6) paths. Indeed, most of the country paths among local ISPs stay in ZA and traverse JINX, DINX and/or CINX, IXPs where these ISPs peer. However, we find some pairs of ASes for which country paths corresponding to v4 IP paths stay in ZA or connect 2 countries while those of v6 ones, pass through US, NL, GB or DE. We can list, for instance, pairs of ASes AS2018 (TENET, ZA) and Internet Solutions, AS6083 (POSIX, ZA) and Internet Solutions in ZA, TENET and AS37105 (Neology, ZA), or Internet Solutions and AS14988 (BFN, BW).

Since some ISPs based in ZA are present in diverse African regions (c.f. Figure 4 for the case of West Africa), a relevant proportion of paths from a region to another traverse IXPs in ZA, used as a hub by these continental ISPs. For instance, such a property was observed on paths from AS30982 (CAFE, TG) to AS36974 (AFNET, CI), from AS37043 (PNL, ZM) to AS37154 (ZAMTEL, NA), and AS36958 (CWS, SC) to AS3057 (Robtex, LS).

Besides, we find many links between some ISPs based in ZA and those in countries close to it, such as SC, NA, BW, LS, MU, MZ. Similarly to ZA, country paths corresponding to communications among some ISPs of TZ, BJ, NA, CI and SC stay in those countries.

Figure 6 illustrates that 60.1% of v4 (52.1% of v6) paths traverse GB, 31.6% of v4 (9% of v6) paths, FR, and 15.3% of v4 (26.8% of v6) paths pass across US. In fact, the corresponding IP paths traverse a set of *Int* ASes and/or IXPs located in these countries, since the source and the destination ASes are either non-connected or mono-homed and route the packets to their transit providers. We denote that on most of the said paths, specifically those from and towards the Southern and Western regions, ZA also lie on the paths.

To summarize, ZA appears as a hub traversed by a relevant number of packets destined to African countries, especially those close to it or in which South African ISPs are having presence. Countries such as GB, FR, NL, US, DE, etc, lie on a considerable amount of paths since the physical infrastructure in Africa is still fragmented, and non-connected ISPs route packets for intra-African communications overseas.

5.7 End-to-end delay analysis

In this section, we analyze the end-to-end round trip time among our probes, highlighting the impact of path characteristics on experienced delay. We then focus on a specific subset of our data, to pinpoint the impact of the establishment of a new IXP on the end-to-end delay among probes whose paths were affected by the new deployment.

5.7.1 Minimum RTT distribution

Figure 7 shows, for 4,082 v6 and v4 AS paths, the distribution of the minimum RTTs. We first identify per AS path, the IP path with the minimum RTT in the dataset as well as its corresponding country path. We then group AS paths into two categories depending on countries lying on the said country path. Continental AS paths (from 1 to 1,073) are those which stay within Africa, whereas intercontinental ones (from 1,074 to 4,082) do not.

As for the continental paths, very low RTTs mostly correspond to paths among pairs of ASes based in the same country, or those passing through collocated regional ISPs, for example in SC, ZA, TZ, BJ, CI, etc. Slightly longer RTTs (50-150ms) are seen among AS pairs directly peering from geographically distant countries. For instance, a path from a KE ISP to a ZA ISP, only served by African transit ISPs, shows a minimum RTT of 80ms. A striking result comes from the presence of very long RTTs in paths that are categorized as continental ones. These v4 paths are typically those between Eastern African ISPs and Western African ISPs, which are served by ZA transit ISPs. The longest RTTs are recorded on paths from TZ to ZA via AS37100 (SEACOM, MU), from ZA to TZ via KE, or from SAIX-NET to TENET in ZA. All the continental v6 paths traverse ZA.

88.8% of the pairs of ASes based in West Africa communicate through intercontinental paths. 84.8% of the pairs of ASes based in Africa, but not in ZA, use intercontinental paths as well. Only 3.77% of the paths among ZA ISPs are intercontinental ones; ISPs in ZA tends to rely less on *Int* ASes than the rest of Africa. We did not expect intercontinental paths to have a low RTT (i.e. $\prec$100ms). In fact, not only these AS paths contains *Int* ASes but also the IPs have been located in either GB, NL, FR or US. However, it seems that packets following these paths are not leaving the continent. Investigating the router level traces of these paths, we realized that these *Int* ISPs are actually routing this traffic within an African-based infrastructure.

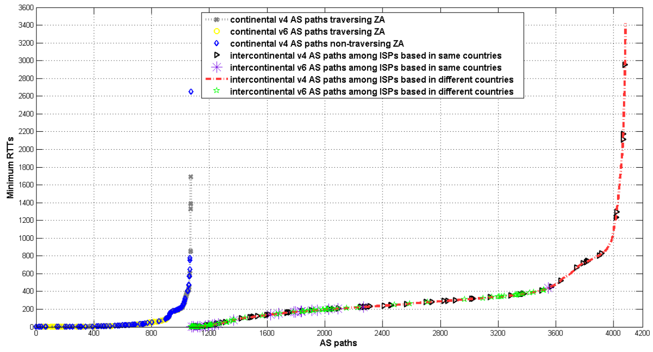

Figure 7: *Minimum RTTs distribution per AS pairs. We collect, sort and plot the minimum RTT per AS path depending on the category of path (continental paths from 1 to 1,073, and intercontinental ones, from 1,074 to 4,082).*

Most of the measured RTTs in this category however reflect intercontinental transit of continental traffic with a delay around 200ms. A first gap is found between such measurements and another group of RTTs scattered around 750ms. These paths are typically from and towards ISPs that are mostly served by Satellite service providers, routing traffic through other continents. For example, a path in this group is from Connecteo in BF to AFNET in CI, passing through SkyVision, Level3 (via New York), Level3/GlobalCrossing (via London), and MTN SA. The paths measured with an RTT of around 1200ms were individually checked at the router level. One is from Connecteo in BF to Sonitel in NE, also going through the US and Europe, but arriving in NE via another satellite, provided by IntelSat.

We notice that many AS pairs, based in neighbouring West African countries, were exchanging packets through intercontinental AS paths. We also highlight, for both v6 and v4, RTTs corresponding to intercontinental AS paths among ASes operating within the same country. They are observed for example in BJ, CM, MA, MZ, MU, etc.

5.7.2 Emergence of new IXPs

We present in this section some known peering links and IXPs used by the ISPs as well as new IXPs found during our measurement campaigns, with an emphasis on the eXchange Point AS327719 (SIXP, GM). To this end, we compare AS adjacencies in our paths to those among peers in PCH and Peeringdb datasets [39, 28]. As mentioned above, we notice in Southern Africa, the use of IXPs and peering links in ZA and TZ by studied ISPs operating in those countries. We detect peering links, for instance, between AS36877 (IWAY) and AS36996 (Telecom Namibia) in NA and between AS29571(CI Telecom) and AFNET in CI.

We find new IXPs in SC, BJ and GM. To the IXP in SC are connected 4 peers: CWS, AS36902 (Intelvision), AS37343 (Telecom Seychelles) and AS36867 (Kokonet Ltd). At the IXP in BJ, we find 3 peers: Benin Telecom, Isocel Telecom and AS37292 (OTI). At SIXP in GM, to which an AS has already been attributed by AFRINIC, we also notice the presence of 3 peers (Figure 8): AS37309 (QCell), AS37323 (Netpage) and AS25250 (GAMTEL).

Figure 8 depicts RTTs among peers ASes and those from these peers to SIXP. The RTTs are very low and around 1.5ms among QCell, NetPage and SIXP. There is a direct link (path of length 2) in both directions between QCell and SIXP. Moreover, GAMTEL and SIXP appear within the path from SIXP to QCell. This shows an inaccuracy of the IP to AS mapping (the IXP address space), but hints to the fact that GAMTEL is also a peering partner. Indeed, to peer at SIXP, GAMTEL and QCell use the IXP address space, while the BGP peering is either direct or via a route-server. The IXP typically holds the AS that announces the IXP address space, which causes its appearance in between the peers at that IXP, when using IP to AS mapping with Team Cymru (or data from RIPE RIS). Meanwhile, from GAMTEL to SIXP, our measurements show a direct link. Besides, paths from GAMTEL to QCell all contain SIXP. These prove the success of SIXP, the new IXP of GM.

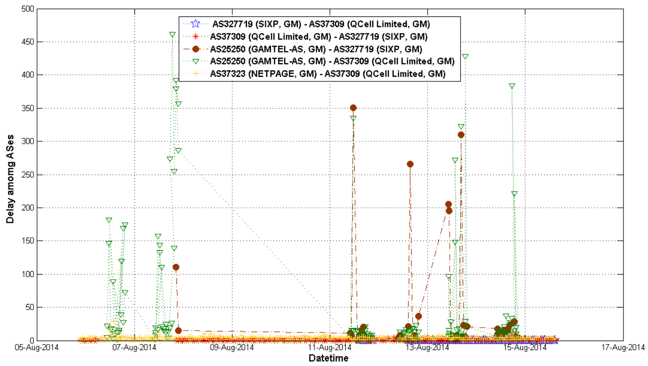

Figure 8: *Delays among peers at SIXP (GM)*

However, RTTs between GAMTEL and QCell and those between GAMTEL and SIXP fluctuate between low (0.96ms) and high values (460ms). After comparison with RTTs between Netpage and QCell (0.037ms - 18.93ms), we deduce that the link from GAMTEL to SIXP is responsible of such delays. Furthermore, we notice that, although it peers at the IXP, the paths from Netpage to SIXP transit via Orange, Cogent, AS8657 (PT Comunicacoes, PT), and QCell. In fact, Netpage filters packets destined to SIXP address space and not those destined to its peers, while QCell does not and advertises SIXP address space to Internet. As a consequence, any packet going from Netpage to SIXP transits via their upstreams and via QCell, with a long delay. These shortcomings could be solved by the peers.

6. CONCLUSION AND FUTURE WORK

In this paper, we present initial steps towards understanding the global African interdomain routing topology. To this end, we enhanced the RIPE Atlas infrastructure, by deploying new probes in 15 ISPs based in 11 African countries which were not hosting any. Overall, we collected and analyzed traceroutes data from 214 RIPE Atlas probes located in 90 ASes, covering 32 African countries.

We then propose and successfully validate an algorithm for router geolocation. Based on this data, we study the country centrality, as proposed in [35], for both v4 and v6 African routing infrastructures. The authors of [35] indeed explain that the lack of knowledge on IP transit in sparsely connected countries biases their results on country centrality for the concerned regions. We hope that our work can help towards resolving this issue.

Moreover, we find a diversity of transit operators playing a role in the provision of both v4 and v6 African interdomain paths, but we also highlight the dominant reliance on intercontinental ISPs for the establishment of continental connectivity. This leads to long AS paths and RTTs, sometimes among ISPs in the same country. Our results show a lack of interconnection among African ISPs in v4 (South Africa being an exception) confirming the interest of initiatives to promote peering on the continent. We also show that local paths and peering are, in contrast, more often observed on v6. This result is however biased by the fact that the v6 deployments have been made where the v4 topology was also properly meshed. That is, there is no v6 deployments where the v4 topology is sparse and relies on intercontinental transit.

We notice striking differences in transit habits of operators, notably depending on the official language of the country and the monetary region, specifically in West Africa. These illustrate how critical it is to have quantity and diversity in the vantage points used in our measurement campaign.

In a future study, we plan to measure the connectivity between African ISPs and the rest of the world. We would also like to establish a cost model for IP transit in Africa, essential to exhibit ground truth data and incentives for peerings among ISPs. Besides, we'll analyse traffic flows exchanged by ISPs.

Given the current state of transit, RTTs among African ISPs tend to be worse than among an African and a European ISP. It is clear that establishing a local CDN cache in one of these networks only makes sense for the direct customers of that network, and not for the rest of the region. We hope to observe, in our further works, the success of projects aimed at fostering IXP establishments in the area, so that popular content could be served by CDN caches deployed in the region, improving the QoS for the local users while reducing the transit cost of local operators.

7. ACKNOWLEDGMENT

We are grateful to whoever deploy or host a RIPE Atlas probe for helping building such a powerful measurement infrastructure for the Internet community. We also wish to thank Vesna Manojlovic, Emile Aben, Andra Lutu, Nishal Goburdhan, Michuki Mwangi and all those who gave us contacts, advice or credits during our measurement campaign.

8. REFERENCES

[1] Submarine Telecoms Forum, Inc., "Submarine Telecoms Industry Report 2013," Tech. Rep. 74, Terabit Consulting, March 2013.

[2] S. Song, "African Undersea Cables." `https://manypossibilities.net/african-undersea-cables/`, March 2014.

[3] S. Song, "Mapping Terrestrial Fibre Optic Cable Projects in Africa." `https://afterfibre.net/`, March 2014.

[4] V. Pejovic, D. Johnson, M. Zheleva, E. Belding, L. Parks, and G. Van Stam, "The bandwidth Divide: Obstacles to Efficient Broadband Adoption in Rural Sub-Saharan Africa," in *International Journal of Communication 6*, no. 24, pp. 2467–2491, 2012.

[5] R. Les Cottrell, "How Bad Is Africa's Internet?." `https://spectrum.ieee.org/telecom/internet/how-bad-is-africas-internet`, January 2013.

[6] M. Chetty, S. Sundaresan, S. Muckaden, N. Feamster, and E. Calandro, "Measuring Broadband Performance in South Africa," in *Proceedings of the 4th Annual Symposium on Computing for Development*, ACM DEV-4 '13, (New York, NY, USA), ACM, 2013.

[7] M. Kende and C. Hurpy, "Assessment of the impact of Internet Exchange Points (IXPs) - empirical study of Kenya and Nigeria," Tech. Rep. 59, ISOC, 2012.

[8] V. Foster, "Africa Infrastructure Country Diagnostic. Overhauling the Engine of Growth: Infrastructure in Africa," Tech. Rep. 26, The World Bank, September 2008.

[9] J. Gilmore, N. Huysamen, P. Cronje, M. de Klerk, and A. Krzesinski, "Mapping the African internet," in *Proceedings Southern African Telecommunication Networks and Applications Conference (SATNAC), Mauritius*, September 2007.

[10] A. Gupta, M. Calder, N. Feamster, M. Chetty, E. Calandro, and E. Katz-Bassett, "Peering at the Internet's Frontier: A First Look at ISP Interconnectivity in Africa," in *PAM*, p. 10, 2014.

[11] SOFRECO Led Consortium, "Etude sur le Programme de Développement des Infrastructures en Afrique (PIDA) : RAPPORT PHASE III – SECTEUR TIC," Tech. Rep. 156, The New Partnership for Africa's Development (NEPAD), Union Africaine (UA), Banque africaine de développement (BAD), 2011.

[12] RIPE NCC, "Global RIPE Atlas Network Coverage." `https://atlas.ripe.net/results/maps/network-coverage/`, April 2014.

[13] H. Haddadi, M. Rio, and A. Moore, "NETWORK TOPOLOGIES: INFERENCE, MODELING, AND GENERATION," in *IEEE Communications Surveys and Tutorials*, pp. 48–69, IEEE, 2nd Quarter 2008.

[14] L. Gao, "On Inferring Autonomous System Relationships in the Internet," in *IEEE/ACM Transactions on Networking, vol. 9, no. 6*, December 2001.

[15] L. Subramanian, S. Agarwal, J. Rexford, and R. H. Katz, "Characterizing the Internet Hierarchy from Multiple Vantage Points," in *Proceedings of IEEE INFOCOM*, p. 12, 2002.

[16] Z. M. Mao, J. Rexford, J. Wang, and R. Katz, "Towards an Accurate AS-level Traceroute tool," in *Proceedings of ACM SIGCOMM*, pp. 365–378, 2003.

[17] Z. M. Mao, L. Qiu, J. Wang, and Y. Zhang, "On AS-level path inference," in *ACM SIGMETRICS*, no. 11, pp. 339–349, 2005.

[18] Y. Hyun, "Archipelago Infrastructure." `https://www.caida.org/projects/ark`.

[19] N. Spring, R. Mahajan, D. Wetherall, and T. Anderson, "Measuring ISP Topologies with Rocketfuel," in *IEEE/ACM Transactions on Networking*, no. 13, pp. 931–943, 2004.

[20] PingER, "PingER Project." `https://www-iepm.slac.stanford.edu/pinger/`, 2013.

[21] A. Wills and G. Daniels, "Nigeria Telecommunications Market: A Snap Shot View. Africa Analysis," April 2003.

[22] M. Zennaro, E. Canessa, R. Sreenivasan, A. Rehmatullah, and R. Cottrell, "Scientific Measure of Africa's Connectivity," No. 10, pp. 55–64, he Massachusetts Institute of Technology Information Technologies and International Development, 2006.

[23] B. Augustin, X. Cuvellier, B. Orgogozo, F. Viger, T. Friedman, M. Latapy, C. Magnien, and R. Teixera, "Avoiding traceroute anomalies with Paris traceroute," in *IMC '06 Proceedings of the 6th ACM SIGCOMM conference on Internet measurement*, pp. 153–158, October 2006.

[24] J. Van, "Traceroute." `ftp://ftp.ee.lbl.gov/traceroute.tar.gz`, February 1989.

[25] G. Mahlknecht, "Greg's Cable Map." `http://www.cablemap.info/`, December 2013.

[26] M. Michuki, "Overview of African IXPs." `https://www.isoc.org/isoc/conferences/inet/08/docs/inet2008_mwangi.pdf`, June 2008.

[27] Packet Clearing House (PCH), "Internet Exchange Point Growth." `https://prefix.pch.net/applications/ixpdir/summary/growth/`, 2014 April.

[28] Peeringdb, "IXP repository." `https://www.peeringdb.com/private/exchange_list.php`, 2004–2013.

[29] H. Galperin, "Connectivity in Latin America and the Caribbean: The role of Internet Exchange Points," Tech. Rep. 20, Internet Society, November 2013.

[30] R. Bush, "RIPE Atlas probes." `https://archive.psg.com/130614.AtlasProbes.pdf`, June 2013.

[31] S. Bortzmeyer, "RIPE Atlas probes and the User-Defined Measurements." `https://ripe67.ripe.net/presentations/153-`

ripe-atlas-udm-api-1.pdf, October 2013.

[32] N. Melnikov, "Hands-on: RIPE Atlas." https://cnds.eecs.jacob-university.de/users/nmelnikov/aims2013-ripe-atlas.html, AIMS 2013.

[33] Authors, "Our Public Measurements." https://atlas.ripe.net/api/v1/measurement/1717241-to-1717244-and-1740800-to-1760809-and-1760855-to-1765509/result/, September 2014.

[34] AFRINIC, "AFRINIC database." ftp://ftp.afrinic.net/, 2014.

[35] J. Karlin, S. Forrest, and J. Rexford, "Nation-State Routing: Censorship, Wiretapping, and BGP," *CoRR*, vol. abs/0903.3218, no. http://arxiv.org/abs/0903.3218, 2009.

[36] MaxMind, "GeoIP accuracy." https://www.maxmind.com/en/city_accuracy, 2014.

[37] Team Cymru, "Team Cymru Services." https://www.team-cymru.com/.

[38] T. C. A. for Internet Data Analysis (CAIDA), "AS relationships." https://www.caida.org/data/active/as-relationships.

[39] PCH, "PCH IXP directory." https://prefix.pch.net/images/applications/ixpdir/ip_asn_mapping.txt, 2014.

[40] MainOne Cable System, "Mainone cable." https://www.mainonecable.com/network, 2014.