Insight Into Africa's Country-level Latencies

Josiah Chavula*, Amreesh Phokeer†
AFRINIC
11th Floor Standard Chartered Tower
Ebene, Mauritius
Email: *†{josiah.chavula,amreesh}@afrinic.net

Agustin Formoso
LACNIC
Rambla Republica de Mejico 6125
Montevideo, Uruguay
Email: agustin@lacnic.net

Nick Feamster
Princeton University
35 Olden Street
Princeton, New Jersey
Email: feamster@cs.princeton.edu

***Abstract*—This paper provides insight into the effects of cross-border infrastructure and logical interconnections in Africa on both intra-country and cross-border latency on end-to-end Internet paths, by comparing Internet performance measurements between different countries. We collected ICMP pings between countries using Speedchecker and applied a community detection algorithm to group countries based on round trip times (RTTs) between countries. We observed three main latency clusters: East and Southern Africa; North Africa; and West and Central Africa. An interesting observation is that these clusters largely correspond to countries that share the same official languages or past colonial history. The cluster in Eastern and Southern Africa is the most strongly clustered: these countries have the lowest inter-country latency values. We also found that some countries have a much higher intra-country latency than expected, pointing to the lack of local peering or physical infrastructure within the country itself. This finding underscores the importance of physical networking infrastructure deployment and inter-network relationships at a country and regional level.**

***Keywords*—Latency, Internet measurements, peering, topology, clustering**

I. Introduction

Africa has become the new frontier for Content Distribution Networks (CDNs) and other Over-the-Top (OTT) service providers, who are aiming to continuously expand their network infrastructure and customer base. By expanding their infrastructure on the continent, these service providers aim to offer better quality of service to users in the region.

Unlike many other regions where Internet consumption is dominated by real-time entertainment, however, a report from Sandvine shows that Africa is not such a great consumer of online video content, which only amounts to 23% of data on fixed lines and around 6% on mobile. This is far less than in Asia-Pacific (40% both on mobile and wireline) or in Latin America (about 40% wireline and 30% mobile) [1]. There are at least two reasons that could explain the low consumption of real-time content in Africa: (i) the high cost of Internet connectivity and (ii) bad Quality of Experience (QoE). Several recent Africa-centric Internet performance studies show that there are many areas of improvement for the African Internet. Notably, various studies on interdomain routing in Africa have shown that there is a lack of BGP peering between countries the African continent, which causes data that would otherwise stay on the continent to travel much further afield, ultimately affecting end-to-end latency [2, 3, 4, 5].

This study analyzes the quality of national and regional interconnections of African countries, as evident from end-to-end latencies. Using data obtained through active measurements conducted on the SpeedChecker platform [1], we analyze end-to-end latencies within each African country as well as between African countries, and thus characterize the quality of national interconnectivity for African countries. For this purpose, a matrix of latencies between 53 African countries were collected and applied to Louvain [6], a community detection algorithm. We used the Louvain algorithm to group countries based on inter-country round-trip-times (RTT), forming clusters that highlight the different levels of interconnectivity between countries on the continent. The latencies and clusters provide insight into both regional and cross-border infrastructure, such as regional and national Internet exchange points (IXPs); and the logical interconnections such as transit or peering agreements. Studies on the quality of interconnectivity within and between regional countries if vital for discovering and identifying positive strategies of the countries that have achieved better internal and regional interconnectivity. Towards this end, this study sheds additional light on countries that are not interconnecting so well. The results of this study should lead to better and more targeted interventions in these regions by Internet development organizations such as ISOC and AFRINIC.

II. Related Work

Several recent studies have highlighted routing and traffic engineering inefficiencies in Africa's Internet topology [2, 3, 4, 5]. Mostly, these studies have highlighted Internet performance problems that are attributed to lack of peering among Africa's ISPs, inefficient DNS configurations [7], a lack of local content caching servers, and a lack of cross-border cable systems and usage of satellite links. A recurring observation from these studies has been the general lack of local and regional peering among African ISPs, which has resulted in a significant fraction of Africa's Internet traffic being exchanged via intercontinental routes, often through Europe. These studies have also consistently shown that intra-continental end-to-end Internet latencies are comparably much higher in Africa than in most other continents.

[1] http://www.speedchecker.xyz/

Gilmore et al. [5] performed a logical mapping of Africa's Internet topology, highlighting both the router level and AS level paths followed by intra-Africa traffic. Their analysis was based on traceroute data obtained from measurements conducted from a single vantage point in South Africa towards all AFRINIC allocated IP addresses. The resulting logical topology, which contained one-way paths from South Africa to the rest of Africa, showed that most of the routes traversed the United Kingdom, Scandinavia and the USA. Similarly, Chavula et al. [3] used the CAIDA Archipelago platform to conduct logical topology mapping for Africa's national research and education networks. Their study showed that over 75% of Africas inter-university traffic followed intercontinental routes. The study further showed that intercontinental paths were characterised by latencies that were more than double those of intra-continental routes.

Fanou et al. [2] assessed the African interdomain routing topology by performing traceroute measurements from 214 RIPE Atlas probes located in at least 90 ASes located in 32 African countries. Their results also showed a lack of direct interconnection among African ISPs. Accordingly, most inter-ISP paths in Africa, with the exception of those in South Africa, often relied on international transit providers. The work also discovered that many ASes that are geographically co-located in the same countries had much longer inter-AS paths than expected: The average end-to-end RTTs for continental paths were between 50 and 150 ms, whereas for intercontinental routes, the average RTTs were around 200 ms. Most of the latencies between 100 ms and 400 ms (95%) were through Europe, whereas the latencies above 750 ms were for paths that went through satellite links.

Finally, Fanou et al. [8] showed that the sub-optimal Internet performance of many Internet services in Africa largely arises from significant inter-AS delays, which also result in local ISPs not sharing their cache capacity. The authors showed that the observed poor Internet performance also at least partially results from the sub-optimal DNS configurations used by some ISPs on the continent, which sometimes counteract the attempts of providers to optimize interconnectivity and content delivery. In the same vein, a study by Gupta et al. [4] showed that around 66% of Africa-based Google cache content consumed by end users in South Africa, Kenya, and Tunisia was being served through intercontinental links. The analysis from this study also showed that Africa's ISPs do peer to each other much more through European IXPs such as London and Amsterdam, than they do at national or regional IXPs.

While past efforts focused on Africa's end-to-end latencies and routing [2, 3, 4, 5], this paper provides new insights by focusing on country-level latencies. The paper evaluates in-country and cross-border latencies, highlighting the quality of peering within each country, as well as the quality of interconnectivity between neighboring African countries. By applying a clustering algorithm to the latency data, we are able to observe the extent to which ISPs are peering at regional level. An evaluation of the country level latencies reveals the impact of national and regional Internet infrastructure, including fiber optic cables and IXPs.

III. Data Collection

Collecting data samples that properly represent regional connectivity is challenging, as it requires multiple vantage points located in a diverse set of networks, as well as performing measurements to many target networks. The approach employed in this study is similar to the used by Formoso et al. [9] to study the in-country and inter-country connectivity in Latin America and Caribbean region.

A. Measurement Platforms: RIPE Atlas and Speedchecker

An extensive list of Internet measurement platforms is available, with each platform having its own advantages and disadvantages. In this study, one of the most important features was the distribution of measurement vantage points (measurement probes) throughout the continent. RIPE Atlas and Speedchecker provided the most extensive distribution of vantage points in Africa, with Speedchecker appearing to have the highest spread of nearly 850 probes covering most subregions of Africa. RIPE Atlas was found to have around 250 active probes during the measurement period. The experiment covered a total of 322 networks across 53 African countries (Figures 1 and 2). On average, 3300 experiments were conducted per country.

TABLE I
The four country clusters resulting from inter-country latency analysis.

Cluster	mean RTT	Regional mean RTT	Strength	Category
0	256	265	6	weak
1	187	256	69	strong
2	179	287	108	strong
3	209	278	69	strong

Based on the intra-cluster and inter-cluster latencies, it is possible to quantify then strength of the relationships between the different clusters. Figure 3 expands on the data shown in Table I and quantifies each of the relationships obtained from the clustering analysis. Figures 4 and 5 show examples of a strong cluster and a weak clusters.

B. Measurements

We collected ICMP pings from Africa-based Speedchecker probes to Speedtest servers also located in African countries. Three times a day, 20 SpeedChecker probes were randomly selected from each country and used as vantage points. Each set of vantage points was configured to run measurements to one Speedtest server randomly selected from each country. Each measurement consisted of ten consecutive pings that were sent at one-second intervals. We performed these measurements over two months (56 days, ending on 30th

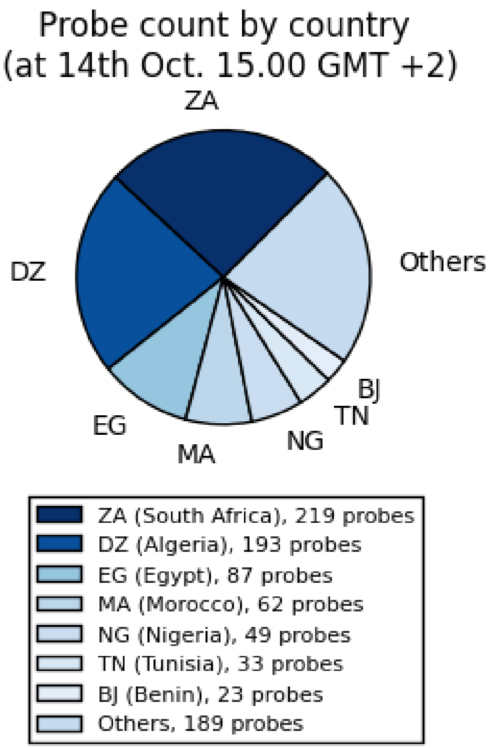


Fig. 1. Number of Speedchecker probes per country; South Africa, Algeria, Egypt and Morocco had the highest number of probes per country

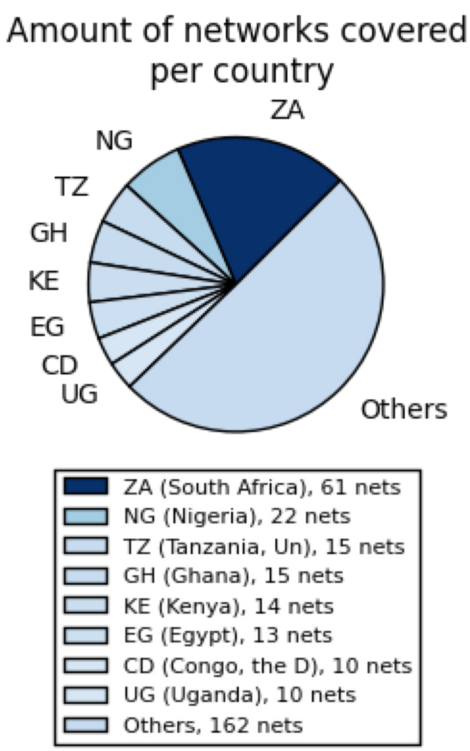


Fig. 2. Number of networks (ASNs) hosting Speedchecker probes per country; South Africa has the highest number of networks covered

September, 2016), generating a dataset of around 121,000 measurements from 53 African countries.

C. Data Post-processing

An initial exploration of the RTT dataset revealed that the RTTs between countries are often asymmetric: Measurements originated in country A and targeting country B, are not necessarily similar to those originating in B and targeting country A. This asymmetry is an important aspect to consider in the analysis since the clustering algorithm we use is based on undirected graphs. The weight from the edge connecting nodes A and B is based on an average calculated from the samples between countries A and B: specifically, we compute the average of A→B and B→A measurements. If A → B and B→A delays are too different (too asymmetric), then their averages would be unrepresentative. For this reason, we did not consider the highly asymmetrical paths in our analysis. Up to 5% of the dataset had an asymmetry greater than 100 ms. Figure 9 shows the distribution of RTT differences and shows how the RTT differences decrease linearly up to the 100 ms, beyond which there is no clear pattern.

IV. Results

A. Inter-Country Latency Clusters

To obtain insight into the quality of country-level interconnectivity, we clustered the data using the Louvain community detection algorithm [10]. The algorithm classifies countries into closely-related communities, referred as clusters based on country-to-country RTTs. Clusters in this sense are therefore groups of countries that share similar inter-country latencies within neighbouring groups as determined by the Louvain community detection algorithm. The Louvain algorithm is based on the concept of modularity optimization of a partition of the network. The higher the modularity, the denser the connections are within a cluster, and the less dense the connections are between nodes of different clusters. The Louvain has been tested in multiple studies that have confirmed the efficiency of the algorithm on large networks, especially in the study of Online Social Networks (OSNs) [11, 6].

We applied the Louvain algorithm to the complete dataset of African inter-country latencies so as to define the community of countries based on the reported latencies. The latency dataset had to be transformed into a graph of regional latencies, with the countries as nodes and latencies as the edges. The dataset was firstly transformed into a matrix M, where M_{ij} would represent the median RTT from country i to country j. A graph can be extracted from the matrix where the edge weights would correspond to the average of M_{ij} and M_{ji}. M_{ii} would represent in-country latency for country i, and these M_{ii} values were not included in the clustering computation, so that nodes could not have self referencing links. Stripping out self-references was done to reinforce the concept that country clusters implicitly refer to relationships between different countries.

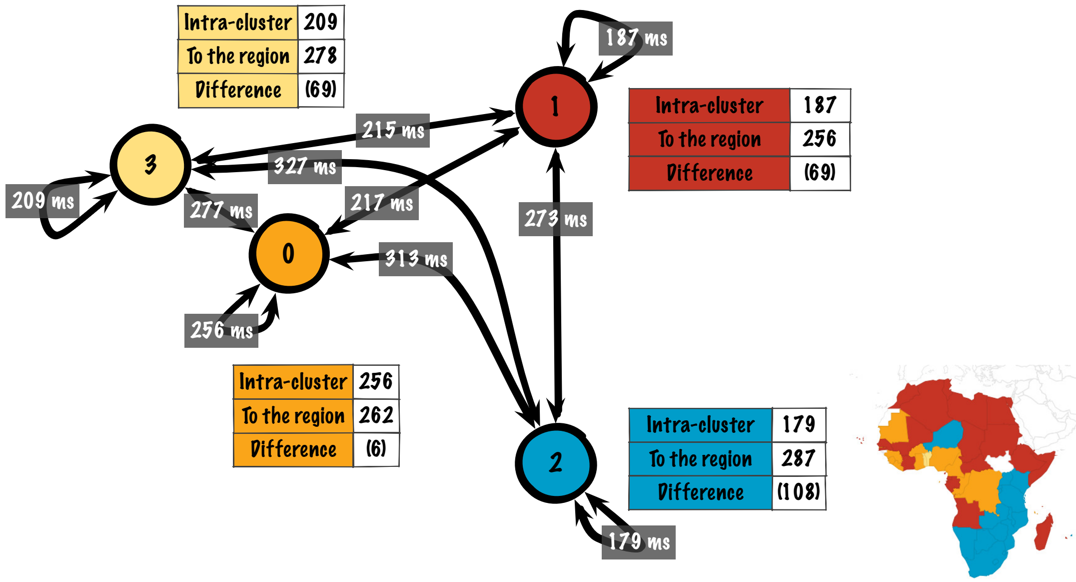


Fig. 3. Graph representation of clusters, with all clusters having significant latencies of 250 ms between them, and relatively lower in-cluster latencies that are below 200 ms. Cluster 0 (West Africa) distinctively has higher in-cluster latencies

The clustering algorithm returned the African countries grouped into four different clusters based on inter-country latencies. Clusters are considered strong when they comprise countries that together have distinctive inter-country latencies, as shown in Figure 3. The criterion to classify clusters as strong or weak is therefore based on the difference between the clusters mean latency (i.e., the average of inter-country latencies within the cluster, (C)), and the average latency from cluster to the rest of the region R. The strength of a cluster is taken as the difference between C and R, i.e., U_{DIFF} = R - C. This means that strong clusters have low RTTs between their own members but are collectively further away from the rest of Africa. On the other hand, weak clusters do not have very distinct latencies between themselves and the rest of the continents. Figure 3 and Table I present the results of the clustering. The strongest cluster appears to be the one comprising countries in Southern and Eastern Africa (Cluster 2). This cluster, which includes the regional hubs South Africa and Kenya, demonstrates generally lower latencies between its members, but much higher latencies to the rest of the continent. This cluster is thus considered to be distant from the rest of Africa. This clustering would be expected, considering that both South Africa and Kenya act as major regional hubs for the region, with a high number of undersea fiber optic cables passing through these two countries' coastlines. The two countries also host some of Africa's most highly utilized Internet exchange points: the Johannesburg Internet Exchange point (JINX) and Cape Town Internet Exchange point (CINX) in South Africa; and the Kenya Internet exchange point (KIXP) in Kenya. The cable maps also indicate that a number of terrestrial fiber cables from the Southern and East Africa region terminate into South Africa.

The Western African cluster (Cluster 0) includes the major economy in West Africa, Nigeria, and some of its neighbors, including Ghana and Cameroon. This cluster appears to be the weakest of the clusters, having high in-cluster latencies that are almost similar to the latencies experienced by the cluster's members to the rest of Africa. Apart from being a weak cluster, the West African cluster is also "geographically fuzzy" in the sense that neighboring countries arbitrarily belong to different clusters. For example, countries such as Senegal and Cote d'Ivoire, which would have been expected to be in the Western cluster, are instead clustered with the North African cluster. This absence of distinctively lower latencies between countries in Western Africa block suggest a lack of network level integration, either due to lack of cross-border physical connectivity, or lack of regional peering between service providers in the region. Such lack of cross-border network integration does result in regional Internet traffic being circuitously routed through distant exchange points, and this could explain the same level of high latencies between neighboring countries in the region, and to the rest of the continent.

The other prominent cluster is the North Africa cluster (Cluster 1). This cluster includes the countries in the Horn of Africa, along the Red Sea, Mediterranean Sea, all the way around the North Atlantic coast of Africa. Again, this appears to be a natural clustering, considering that major international undersea fiber optic cables along North Africa pass through these countries as depicted by the Telegeography Submarine cable map[2]. However, a major anomaly with this cluster is that it includes other countries that are geographically distant

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

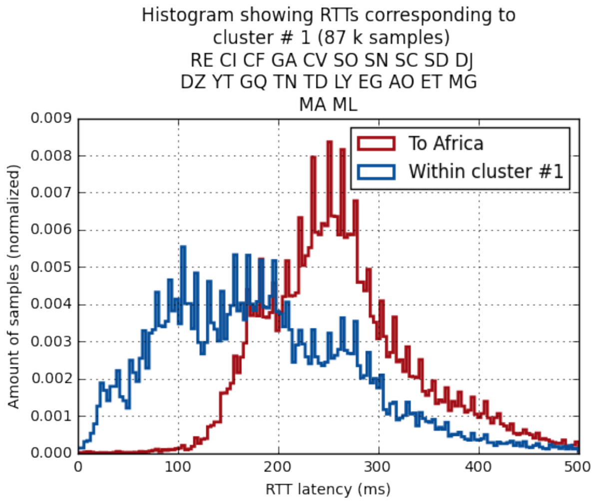


Fig. 4. Example of a strong cluster, where U_{DIFF} = 69

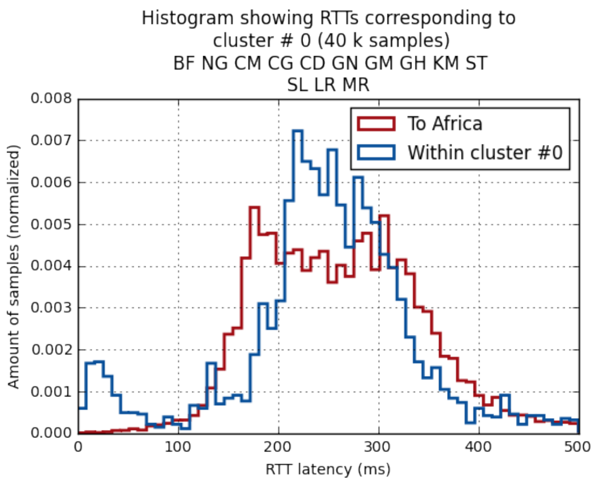
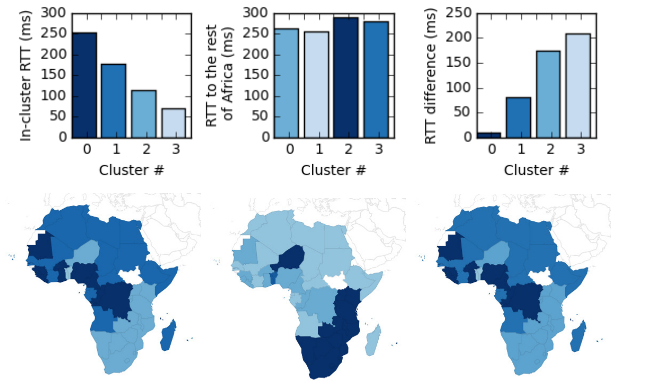


Fig. 5. Example of a weak cluster, where U_{DIFF} = 6

from the main cluster. These distant countries include Angola and Gabon in south western coast of Africa, Senegal, and Cote d'Ivoire in west Africa, as well as Madagascar and Reunion in the Indian Ocean. The inclusion of these distant countries in the North African cluster can be attributed to the weak nature of the cluster (i.e., due to high latencies between cluster members). Indeed, apart the international fibre optic cable along the countries' shores, there appears to be less cross-border fibre optical cables between the North African countries.

Another way to visualize the latency clusters is through a matrix heat map. In Figure 7, the rows and columns are arranged such that countries from the same cluster are grouped together. The resulting heat map is a visualization that provides information on both the range of latencies and clustering proximity of the whole countries. Noticeable aspects of heat map are:

- The cluster formed by Benin and Togo on the lower right corner



(a) Latency between cluster members (in-cluster RTTs)

(b) Latency from the different clusters to the rest of Africa

(c) Difference between the measurements from (a) and (b)

Fig. 6. Cluster Latencies

- The South East cluster starting with Burundi and Botswana, and ending with South Africa and Zimbabwe
- Tunisia stands out as having lowest average latencies to the rest of Africa, while Madagascar appear to have the highest average latencies to the rest of the continent.

B. Intra-Country Latencies

We observed that the clusters exhibit different levels of intra-country latencies. Figure 8 indicates the distribution of mean intra-country latencies among countries. Members of the Western Africa cluster appear to have higher in-country latencies compared to the other clusters. Almost all members of this cluster registered in-country latencies of over 150 ms, with Sierra Leone and Democratic Republic of Congo having the highest in-country latencies of over 350 ms. Again, this could be indicative of the lack of country-level network integration, due to either lack of physical interconnections, or lack of peering among network operators.

Countries in all the other three clusters registered in-country latencies averaging less than 100 ms. In the North African cluster, Tchad had distinctively high in-country latencies of about 400 ms. In the Southern and East African cluster, the highest in-country latencies were recorded in Uganda, which with an average in-country latency of 200 ms. This could be indicative of lack of fiber optic cables in these countries, as this level of latencies indicates high usage of satellite communication.

V. Conclusion and Future Work

This study on latency measurement in Africa has shown that countries within the same cluster share similar network characteristics. Lower in-country latencies are observed for the Southern and Eastern African countries, which also correlates with the more dense fiber cable distribution in the area[3]. In contrast, the Northern and Central blocks have

[3]https://afterfibre.nsrc.org/

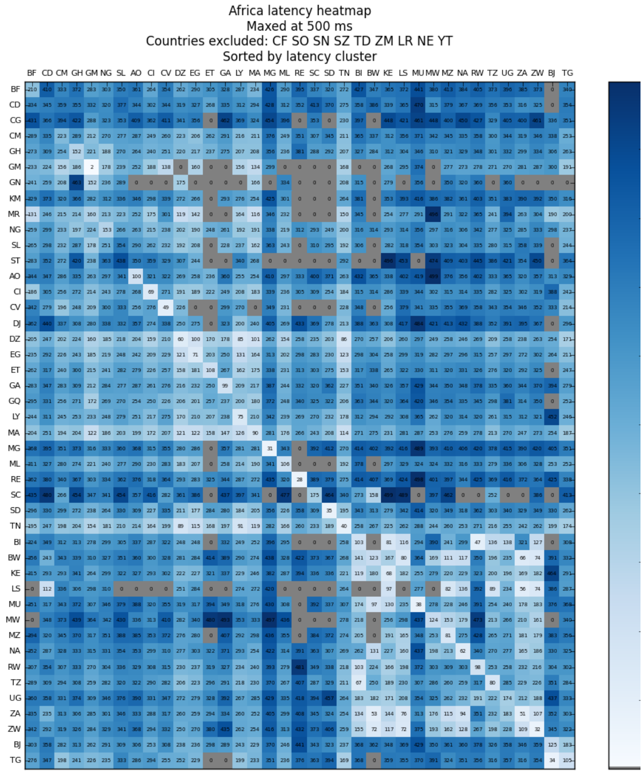


Fig. 7. Africa latency matrix, reordered by the clustering algorithm. 0 indicates no meaningful samples could be gathered between that pair of countries.

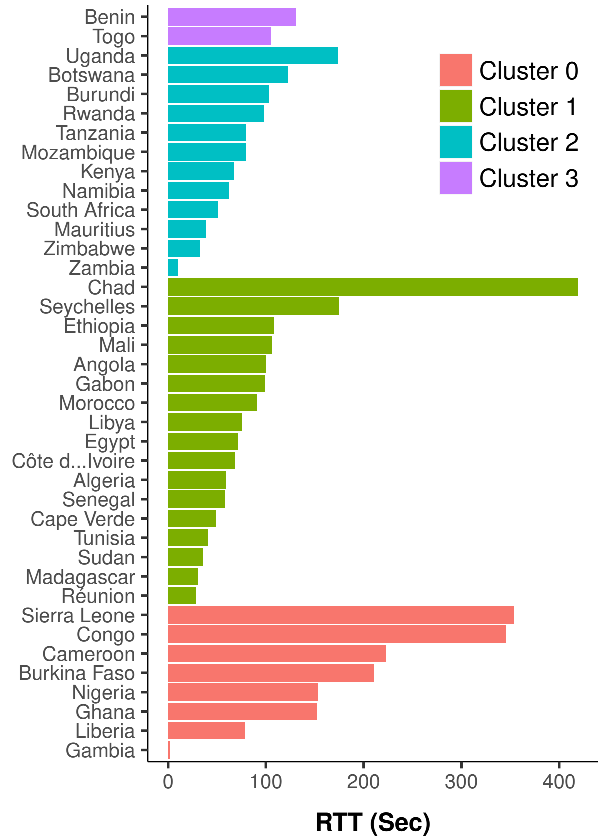


Fig. 8. Distribution of Intra-country latencies, grouped by clusters

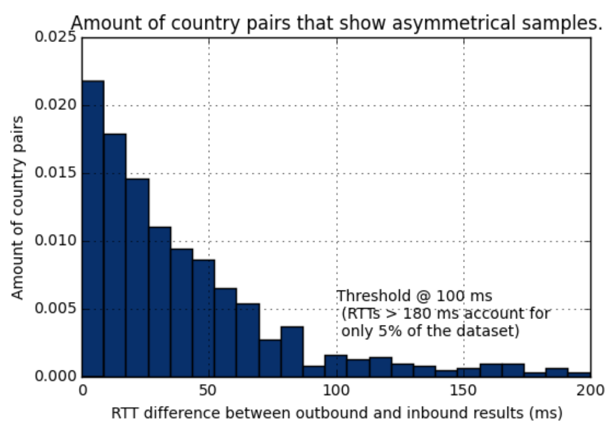


Fig. 9. Distribution of countries pairs that show asymmetrical samples. The data suggests a threshold at 100 ms, where the linear tendency is broken and marks the beginning of a long random tail

lesser fiber optic cable density and also experience higher in-country latencies.

While this paper has given some insight into the state of connectivity and logical proximity between African countries, there as still many other questions that need to be explored. For example, the relationship between infrastructure and the observed latencies needs further analysis. The expectation would be that countries that have direct physical interconnectivity should have lower latencies. However, the availability of physical infrastructure is not enough to ensure low latency: the logical interconnection between ISPs, either for transit (provider-customer relationship) or peering, is an important factor in keeping in-country latencies low. Although latency information does provide some insight into regional connectivity, and it can pinpoint both predictable behavior as well as anomalies, it has no way to demonstrate causation for the higher latencies. Future work will therefore focus on analyzing the correlating between physical and logical topologies with latency. In order to establish such relationships between infrastructure and latencies, there is need for analysis of the submarine or terrestrial cables datasets in relation to the observed latencies. In order to obtain meaningful insight about why latency patterns or anomalies are seen, the latency dataset must be analyzed in parallel with other datasets, such as physical cables and peering relations, which could contribute to a better understanding about the state regional networking.

References

[1] "Sandvine global internet phenomena report," 2016, accessed: 01-Jun-2016. [Online]. Available: https://www.sandvine.com/downloads/general/global-internet-phenomena/2011/2h-2011-global-internet-phenomena-report.pdf

[2] R. Fanou, P. Francois, and E. Aben, *On the Diversity of Interdomain Routing in Africa*. Cham: Springer International Publishing, 2015, pp. 41–54. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-15509-8{_}4

[3] J. Chavula, N. Feamster, A. Bagula, and H. Suleman, *Quantifying the Effects of Circuitous Routes on the Latency of Intra-Africa Internet Traffic: A Study of Research and Education Networks*. Cham: Springer International Publishing, 2015, pp. 64–73. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-16886-9{_}7

[4] 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," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8362 LNCS, pp. 204–213, 2014.

[5] J. Gilmore, N. Huysamen, and A. Krzesinski, "Mapping the african internet," in *Proceedings Southern African Telecommunication Networks and Applications Conference (SATNAC), Mauritius*. Citeseer, 2007.

[6] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, p. P10008, 2008. [Online]. Available: http://stacks.iop.org/1742-5468/2008/i=10/a=P10008

[7] A. Phokeer, A. Aina, and D. L. Johnson, "DNS Lame delegations : A case-study of public reverse DNS records in the African Region," in *Proc. AFRICOMM 2016, Ouagadougou, Burkina Faso, 6-7 Dec 2016*, 2016.

[8] R. Fanou, G. Tyson, P. Francois, and A. Sathiaseelan, "Pushing the frontier: Exploring the african web ecosystem," in *Proceedings of the 25th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 2016, pp. 435–445.

[9] A. Formoso and P. Casas, "Looking for Network Latency Clusters in the LAC Region," in *Proceedings of the 2016 Workshop on Fostering Latin-American Research in Data Communication Networks*, ser. LANCOMM '16. New York, NY, USA: ACM, 2016, pp. 10–12. [Online]. Available: http://doi.acm.org/2940116.2940130

[10] V. Blondel, J. Guillaume, R. Lambiotte, and E. Lefebvre, "The louvain method for community detection in large networks," *J of Statistical Mechanics: Theory and Experiment*, vol. 10, p. P10008, 2011.

[11] M. Cordeiro, R. P. Sarmento, and J. Gama, "Dynamic community detection in evolving networks using locality modularity optimization," *Social Network Analysis and Mining*, vol. 6, no. 1, p. 15, 2016. [Online]. Available: http://dx.doi.org/10.1007/s13278-016-0325-1