ARTICLE TYPE

A System for Profiling the IXPs in a Region and Monitoring their Growth: Spotlight at the Internet Frontier

Rodérick Fanou†1 | Víctor Sanchez-Agüero2,3 | Francisco Valera3 | Michuki Mwangi4 | Jane Coffin5

1Center for Applied Internet Data Analysis (CAIDA)/University of California, San Diego (UCSD), San Diego, USA
2IMDEA Networks Institute, Madrid, Spain
3Universidad Carlos III de Madrid (UC3M), Madrid, Spain
4Internet Society (ISOC), Nairobi, Kenya
5Internet Society (ISOC), Reston, USA

Abstract

This work aims at designing and implementing a system able to profile and help manage the set of IXPs in an Internet region. As part of the Internet Society’s strategy to help monitor and understand the evolution of IXPs in a particular region, a route-collector data analyzer tool was developed before being deployed and tested in AfriNIC. In fact, traffic localization efforts in the African peering ecosystem would be more sustained and their efficacy assessed if they were supported by a platform, which evaluates and reports in real-time about their impact on the Internet. We thus built the “African” Route-collectors Data Analyzer (ARDA), an open source web platform for analyzing publicly available routing information collected since 2005 by local route-collectors. ARDA evaluates pre-defined metrics that picture the status of the interconnection at local, national, and regional levels. It shows that a small proportion of AfriNIC ASes (roughly 17%) are peering in the region. Through them, 58% of all African networks are visible at one IXP or more. These have been static from April to September 2017, and even February 2018, underlining the need for increased efforts to improve local interconnectivity. We show how ARDA can help detect the impact of policies on the growth of local IXPs, or continually provide the community with up-to-date empirical data on the evolution of the IXP substrate. Given its features, this tool will be a helpful compass for stakeholders in the quest for better traffic localization and new interconnection opportunities in the targeted region.

KEYWORDS: IXPs, Growth, Monitoring, Interconnection.

1 | INTRODUCTION

Since Internet connectivity appears to be a lever of development in connected areas, there is an increasing interest from the Internet community in continually characterizing local interconnection in under-connected regions for efficiently helping them improve. Meanwhile, Internet Service Providers (ISPs) are more and more interested in acquiring updated details about the current situation to identify potential positioning opportunities in those geographical areas.

†This work has mostly been done while Rodérick Fanou was a PhD Student at IMDEA Networks Institute and Universidad Carlos III de Madrid (UC3M), Spain.
Although several works have been investigating the existence of Internet eXchange Points (IXPs), building tools to map IXPs within traceroutes paths, quantifying the impact of IXPs on the AS-level topology structure of the Internet, investigating interdomain congestion notably at IXPs, or exploring possible scenarios to enlarge the peering ecosystem within a region, the Internet community lacks a platform that enables an automated and detailed study of their growth or an inspection of the advantages of peering at each of those facilities. **Filling this gap is the primary objective of this paper.** In fact, such a tool is essential, as it will help ISPs willing to either invest in an Internet region or extend their footprint to efficiently detect where they could set up their next Points of Presence (PoPs), according to their interests. When answered well, this critical question becomes a pillar that empowers the connectivity in the said region. Further, regulators and Internet developmental institutions will also be able to keep an eye on the evolution of the ecosystem with respect to their expectations. Researchers may also use this tool to quickly assess the interconnectivity within an Internet region.

In this perspective, this work aims at designing and implementing a system able to profile the set of IXPs of a region. As part of the Internet Society’s (ISOC) strategy to allow IXPs to monitor and understand the evolution of those peering fabrics in a particular region, we developed a route-collector data analyzer tool. After that, we deployed and tested it in the AfriNIC region, which represents the Internet frontier due to its low Internet penetration. Moreover, this study is in line with the need for a longitudinal measurement and supervision of its evolving IXP infrastructure, mentioned in Fanou et al. In fact, the African peering ecosystem has been the subject of much attention over the last years with the goal of meeting the traffic localization challenge: the efforts of stakeholders in this direction have led to the setup of more IXPs. There are 42 IXPs in Africa (hosted in 32 countries), of which 20 have been set up since 2009 as of April 2018.

However, 21 years after the launch of the first local IXP, the monitoring and measurement infrastructure of the region still challenges the evaluation of the progress made on traffic localization. Assessing the impact of related activities – such as policy implementation and infrastructure developments – is quite challenging, considering that very few IXPs provide publicly accessible data on current traffic statistics or colocation data. As noticed in PeeringDB and PCH public datasets on IXP colocation are not up-to-date when it comes to IXPs in Africa because some IXP members either do not register as peers at existing infrastructures, or do not add their prefixes to their information in those datasets. Besides, locally useful data essential to support the growth of peering in the region is unavailable: this is particularly important in regions such as AfriNIC or LACNIC (Latin America and Caribbean Network Information Centre), where the hidden Internet topology complicates the analysis of the expansion possibilities. Further, other increasingly useful measurement resources (e.g., the RIPE Atlas network) still offer limited visibility in the African IXP substrate: despite intensive deployment efforts, only 17.5% of local networks host a RIPE Atlas probe as of April 2018.

The ISOC then decided to help offset the lack of progressive, visual, and near real-time information on the status of networks operating in a given Internet region, developing in a joint effort with the Universidad Carlos III de Madrid (UC3M) a methodology that enforces the collection, collation, and publication of useful data points, from an internal Vantage Point (VP). In fact, automating these tasks will ease the monitoring and reporting of the progress being made on interconnection and traffic exchange in the said region. We actively contributed to the definition of that methodology and designed the system automating it: the implementation of the said system for the AfriNIC region led to the African Route-collectors Data Analyzer (ARDA).

ARDA is an open-source tool – accessible through a web interface – that constantly collects raw routing data from route-collectors, which are deployed at African IXPs and capture a peering viewpoint of the Internet. The system then inspects this data from various angles to assess peering evolution in the region. It was built and tested from December 2015 during an overall period of 18 months and is freely available since April 18, 2017 at [https://arda.af-ix.net](https://arda.af-ix.net), i.e., hosted in the domain of the African IXP Association (Af-IX, [www.af-ix.net](http://www.af-ix.net)).

Such a compass is intended to: (i) provide network operators with supporting information for peering decisions, (ii) provide empirical data to support business investment decisions and opportunities in the region (*Internet business development*). Besides, this tool will (iii) inform development organizations and policy-makers on gaps and state of interconnection in the region (*Internet community*), and (iv) help researchers undertake interconnection studies or complement measurement studies that use other data sources, such as RIPE Atlas network, Ark measurement infrastructure, etc.

The contributions of this paper can be summarized as follows: first, we present the design and the implementation of the deployed system and show how it can help achieve the goals previously mentioned. We expect that the Internet research community, IXP and network operators, Internet developmental institutions, regulators, and stakeholders who have interests in ICT will make use of this tool to obtain in-depth analysis of the provided data and statistics in their own benefit. Next, we give striking examples of the different possibilities that ARDA offers in Section 4.3.2. We then present our definition of
the techniques automated by the route-collectors data analyzer as well as our design and implementation of the ARDA platform, including the key algorithms used to analyze data, analysis results from the BGP data, and use cases showing their value for the Internet ecosystem. As it will be seen, ARDA is built so that it can easily be applied to other regions in the future: in this direction, its code is also made public on GitHub at https://github.com/rodricfanou/African_Route-collectors_Data_Analyzer-ARDA.git. It is still a living project that the ISOC keeps on supporting, and some of its possible expansions will be commented on Section [1].

The methodology adopted to achieve our purposes constitutes the remainder of this paper. After defining the requirements of the route-collectors data analyzer in Section [2], we introduce its architecture in Section [3]. We then present the different steps of the data collection and storage process (Section [4.1]), followed by the data analysis (Section [4.2]): while doing so, we highlight our technical choices for the implementation of the ARDA platform from which arose some striking (visualization) results that underline its relevance for the Internet community (Section [4.3]). Next, we discuss the related work in Section [5]. Finally, we conclude and present our future directions for improving this application in Section [6].

2 | REQUIREMENTS OF THE ROUTE-COLLECTORS DATA ANALYZER

In this section, we better highlight the main visible outputs expected at the end of this work. They can be listed as follows:

1. **IXP growth and business potential**: the route-collectors data analyzer should constantly provide graphical views of the visible networks at each IXP, help IXPs market their features, and help end-users identify sub-regions that are connected to a particular IXP.

2. **Interconnection development progress and gaps**: the route-collectors data analyzer should monitor local and regional interconnection growth, help identify IXPs that are facing potential challenges, as well as track local and regional policy and track regulatory impact on interconnection development.

3. **Technical support**: the route-collectors data analyzer is expected to report on the networks that are likely to have routing inefficiencies at each IXP.

We next determine the main aspects around which the designed system can be centered. The route-collectors data analyzer must depend on a reliable system, which locally collects and stores in a common format, the historical and current routing data previously fetched by passive measurements at the IXP. Pre-defined statistics, termed metrics in the rest of this paper, are then expected to be computed based on this collected data and presented under the following three views: (i) the IXP View whose metrics are per IXP (ii) the National View whose metrics involve the set of IXPs in the same country, and (iii) the Regional View for which the provided metrics cover all IXPs in the region. This naming was agreed with the African network and IXP operators during the Africa Peering and Interconnection Forum (AfPIF) 2016. Further, the designed route-collector data analyzer should have the ability to integrate private route-collectors deployed by local IXPs and the ability to be configured for other regions. Regarding its implementation in the AfriNIC region, it is essential to select a suitable location on the Internet, where the designed system could be hosted so as to be delivered with a high Quality of Service (QoS) to its potential users: IXP operators, Internet developmental institutions, current and potential peers (network operators, Content Providers (CPs)), etc.

3 | PROPOSED ARCHITECTURE OF THE ROUTE-COLLECTORS DATA ANALYZER

The architecture of the route-collectors data analyzer is composed of three modules (Figure [1]), which have been defined given the above-listed tasks. First, the data collection module is in charge of automatically identifying existing route-collectors, their type, and location in the studied region. It then ensures the concurrent download and parsing of BGP data from those sources to extract entries corresponding to those of our data structure. Not only this module collects historical BGP data in the background, but it also downloads hourly or daily the latest available routing data. Second, the data storage and metrics computation module ensures the storage of the key information from among those previously extracted and their usage to compute our metrics using data for the last month, the previous year, or the whole period of the dataset. The results are divided into weeks, months, and years respectively. Optimized algorithms, parallelism for fast computations are essential for delivering those results in real-time. This module thus contains numerous scripts (playing distinct functionalities) of which any set are concurrently launched by an
orchestrator to satisfy the need to update the values corresponding to each metric on time. Finally, the visualizations module generates and presents in the most appealing way dynamic graphs depicting the evolution of the previously computed metrics. Those charts are classified depending on the three views mentioned in Section 2. As one can notice, each module logically relies on the results obtained by the previous ones and on their good functioning. Our implementation of these modules in the ARDA platform and their functioning are detailed in Section 4.

4 | IMPLEMENTATION OF ARDA AND RESULTS

In this section, we present how we implemented ARDA following the steps data collection, data storage and metrics computation, and visualization. After that, we give some examples of empirical data that are easily accessible thanks to the platform and highlight the interesting insights they convey. We made the source code of ARDA publicly available on GitHub at https://github.com/rodricfanou/African_Route-collectors_Data_Analyzer-ARDA.git in March 2018.

4.1 | Data collection

The selection of the suitable Data Sources (DSes) is critical for successfully meeting the requirements while implementing ARDA’s data collection module. Poese et al.14 proved that using geolocation databases individually may lead to wrong inferences. For the geolocation of any new route-collector, we thus retained four DSes: OpenIPMap (OIM)32, Maxmind (MM)33,34, reverse DNS lookups outputs (RDNS), and Team Cymru (TC)35. Similarly to 3,21, these DSes are cross-checked. When all DSes having an entry do not return the same Country Code (CC) for the detected route-collector, ARDA does not suggest any location and lets its administrator manually add it. Based on the CC, the route-collector can then be tagged as deployed or not in Africa. Contrary to 3,21, we however do not add Regional Internet Registry (RIR) DSes among the DSes, or use ping measurements for tie-breaking in case of geolocations conflicts between the datasets: these are planned as future work. However, to depict the percentages of ASNs assigned to each country, which are seen at each IXP or at a set of IXPs on our maps, we use the RIRs DSes altogether as explained in Section 4.3.

To make ARDA give a broad view of the IXP substrate in the AfriNIC region, we chose to design it so that it combines data from existing RouteViews collectors with those from Packet Clearing House (PCH) and IXPs private route-collectors to perform its analysis. However, both the peering and transit links of Liquid Telecom and Network Platforms LTD (two Johannesburg Internet eXchange (JINX) members) are captured by JINX RouteViews, as shown in Figure 2a: this route-collector thus does not have a peering viewpoint. Consequently, we removed it from the set of route-collectors. For similar reasons, the RouteViews collector deployed at NAPAfrica (South Africa (ZA)) on September 7, 2017, cannot be considered either (cf. Figure 2b). Nevertheless, as JINX and NAPAfrica host PCH route-collectors with a peering viewpoint and both are connected to a more significant set of peers, not retaining their respective Routeviews collectors has little impact on the quality or the scope of the collected dataset.

We could not involve all the 42 African IXPs in this process, since not all of them host a PCH or RouteViews route-collectors for collecting routing data on a daily basis. Table 1 summarizes the type and number of route-collectors per IXP covered by our dataset and their corresponding country host. In that table, RCs means route-collectors, PCH stands for Packet Clearing House and RV stands for RouteViews. Table 1 also specifies the year of the launch of each IXP and the date of the deployment
FANOU ET AL

(a) Summary of the RIB entries of the JINX Routeviews collector. AS30844 (Liquid Telecom, United Kingdom (UK)) and AS37497 (Network Platforms LTD, South Africa (ZA)) advertise their full view of the Internet to the route-collector given the number of IP prefixes received from these ASes.

(b) Summary of the RIB entries of the NAPAfrica Routeviews collector. AS37468 (Angola Cables, Angola (AO)), AS37497 (Network Platforms LTD, ZA), AS37353 (Macrolan, ZA) and AS328145 (Lyca Digital, ZA) advertise their full view of the Internet to the route-collector given the number of IP prefixes received from those ASes.

FIGURE 2 Outputs of “sh ip bgp sum” run on JINX and NAPAfrica RouteViews collectors as of October 22, 2017, showing that they capture routing information received via both peering and transit links by some of their peers.

TABLE 1 List of the 24 African IXPs and the corresponding 41 route-collectors subject of this study.

<table>
<thead>
<tr>
<th>CC</th>
<th>Country</th>
<th>IXP</th>
<th>Year of IXP launch</th>
<th>Type (#) RCs</th>
<th>1st date of RC deployment</th>
<th>Gap period (in years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BJ</td>
<td>Benin</td>
<td>BENIN-IX</td>
<td>2013</td>
<td>PCH (1)</td>
<td>29/07/2015</td>
<td>2</td>
</tr>
<tr>
<td>BW</td>
<td>Botswana</td>
<td>BINX</td>
<td>2005</td>
<td>PCH (1)</td>
<td>08/07/2016</td>
<td>11</td>
</tr>
<tr>
<td>EG</td>
<td>Egypt</td>
<td>CAIX</td>
<td>2002</td>
<td>PCH (2)</td>
<td>17/10/2011</td>
<td>9</td>
</tr>
<tr>
<td>GM</td>
<td>Gambia</td>
<td>SIXP</td>
<td>2014</td>
<td>PCH (1)</td>
<td>20/02/2015</td>
<td>1</td>
</tr>
<tr>
<td>KE</td>
<td>Kenya</td>
<td>KIXP</td>
<td>2002</td>
<td>RV (1)</td>
<td>07/10/2005</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PCH (3)</td>
<td>06/08/2010</td>
<td>8</td>
</tr>
<tr>
<td>LR</td>
<td>Liberia</td>
<td>LIBERIA-IX</td>
<td>2015</td>
<td>PCH (1)</td>
<td>13/01/2016</td>
<td>1</td>
</tr>
<tr>
<td>MG</td>
<td>Madagascar</td>
<td>MGIX</td>
<td>2016</td>
<td>PCH (1)</td>
<td>15/03/2016</td>
<td>0</td>
</tr>
<tr>
<td>MW</td>
<td>Malawi</td>
<td>MIX</td>
<td>2008</td>
<td>PCH (2)</td>
<td>11/07/2013</td>
<td>5</td>
</tr>
<tr>
<td>MU</td>
<td>Mauritius</td>
<td>MIXP</td>
<td>2008</td>
<td>PCH (1)</td>
<td>25/05/2015</td>
<td>7</td>
</tr>
<tr>
<td>MZ</td>
<td>Mozambique</td>
<td>MOZIX</td>
<td>2002</td>
<td>PCH (2)</td>
<td>21/07/2010</td>
<td>8</td>
</tr>
<tr>
<td>NA</td>
<td>Namibia</td>
<td>WHK-IX</td>
<td>2014</td>
<td>PCH (1)</td>
<td>17/06/2015</td>
<td>1</td>
</tr>
<tr>
<td>NG</td>
<td>Nigeria</td>
<td>IXPN</td>
<td>2007</td>
<td>PCH (2)</td>
<td>30/01/2015</td>
<td>8</td>
</tr>
<tr>
<td>RW</td>
<td>Rwanda</td>
<td>RINEX</td>
<td>2004</td>
<td>PCH (1)</td>
<td>11/05/2015</td>
<td>11</td>
</tr>
<tr>
<td>SD</td>
<td>Sudan</td>
<td>StaP</td>
<td>2011</td>
<td>PCH (2)</td>
<td>10/12/2014</td>
<td>3</td>
</tr>
<tr>
<td>ZA</td>
<td>South Africa</td>
<td>JINX</td>
<td>1996</td>
<td>PCH (3)</td>
<td>19/07/2005</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DINX</td>
<td>2012</td>
<td>PCH (2)</td>
<td>21/02/2014</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CINX</td>
<td>1997</td>
<td>PCH (2)</td>
<td>21/07/2010</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NAPAfricaCT</td>
<td>2012</td>
<td>PCH (3)</td>
<td>18/04/2013</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NAPAfricaDB</td>
<td>2012</td>
<td>PCH (1)</td>
<td>22/09/2015</td>
<td>3</td>
</tr>
<tr>
<td>TZ</td>
<td>Tanzania</td>
<td>AIXP</td>
<td>2006</td>
<td>PCH (1)</td>
<td>15/06/2015</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TIX</td>
<td>2004</td>
<td>PCH (1)</td>
<td>06/06/2015</td>
<td>11</td>
</tr>
<tr>
<td>TN</td>
<td>Tunisia</td>
<td>TUNIXP</td>
<td>2011</td>
<td>PCH (3)</td>
<td>09/12/2014</td>
<td>3</td>
</tr>
<tr>
<td>UG</td>
<td>Uganda</td>
<td>UIXP</td>
<td>2001</td>
<td>PCH (1)</td>
<td>13/06/2016</td>
<td>15</td>
</tr>
</tbody>
</table>

Total 24 IXPs From 1996 PCH (39) From 2010 RV (1) From 2005 0 — 17

of the first route-collector of each type; using these dates, we computed the gap period needed to better point out the dataset limitations. IXPs private route-collectors have not yet been included in our DSes. In total, ARDA involves data from all (41) route-collectors of the region, which could be taken into account for this work. These are deployed at 24 IXPs in 18 African countries located in four African sub-regions out of five.

Further, we used RIRs datasets [36, 37, 38, 39, 40, 41] to extract information related to Autonomous System Numbers (ASNs) and prefixes assignments. Finally, we selected DSes from APNIC’s routing table analysis [42] (from which the Weekly Routing Table Report [43] is extracted) for any comparison between routing information at the IXPs and those appearing on the Internet.
4.2 Data storage and metrics computation

In this section, we enumerate and define some metrics evaluated by ARDA before detailing the algorithms used for their computations, showing how they fit into the three main aspects listed in Section 2.

4.2.1 Key concepts

Let us consider as showcase the Routing Information Base (RIB) entries below, extracted on March 1st, 2018 from the Route-Views route-collector located at LINX (London, United Kingdom). This routing data is known to be in the Multi-Threaded Routing Information Export format (RFC6396) as detailed in Section 2.

```
TABLE_DUMP2|1519862400|B|195.66.236.29|5413|1.0.4.0/22|5413|6939|4826|38803|56203|IGP|195.66.236.29|0|0|NAG|
TABLE_DUMP2|1519862400|B|195.66.224.29|5413|1.0.4.0/22|5413|6939|4826|38803|56203|IGP|195.66.224.29|0|0|NAG|
TABLE_DUMP2|1519862400|B|195.66.226.66|8426|1.0.4.0/22|8426|6939|4826|38803|56203|IGP|195.66.224.21|0|0|NAG|
```

From the left to the right of each RIB entry above, we can list the type of dump (TABLE_DUMP_V2/IPV4_UNICAST), the timestamp (i.e., 1519862400, which can easily be converted into datetime – 03/01/18 00:00:00), the Peer IP address or IP address, which provides update for the RIB entry (e.g., 195.66.236.29 for the first line), its corresponding ASN (e.g., AS5413 for the two first lines), the visible prefix (e.g., 1.0.4.0/22 for all three lines), the AS path, the origin (e.g., IGP for all three lines), the next hop (e.g., 195.66.224.21 for the last line), as well as the sequence number (a simple incremental counter for each RIB entry), etc.

- The number of prefixes visible at a given IXP or number of prefixes seen at that IXP: represents the number of prefixes advertised/reachable by networks peering at that IXP. Considering the example above, we would only have one prefix advertised at the IXP hosting the route-collector, i.e., 1.0.4.0/22.

- A bogon prefix is a prefix that is not routable on the Internet as it belongs to an RFC 1918 or reserved address space.

- The origin ASN of an AS path: is the rightmost ASN of that AS path or the last ASN from the left of that path. As an example, if we consider the AS path “5413 6939 4826 38803 56203” of the first RIB entry, the origin ASN is AS56203.

- The peering ASN of an AS path: is the leftmost ASN of that AS path or the first ASN from the left of that path. If we consider the AS path “5413 6939 4826 38803 56203” of the first RIB entry, the origin ASN is AS5413. AS5413 is peering at the IXP and thus, provides updates for the RIB entry.

- The number of origin ASNs seen at an IXP: represents the number of distinct ASNs that have originated prefixes visible at the IXP. The three RIB entries above would only account for one origin ASN, which is AS56203.

- The number of peering ASNs seen at an IXP: represents the number of distinct ASNs, which are peering with the route-collector located at the IXP and which update at least one RIB entry. When considering the three RIB entries above, we would count two peering ASNs: AS5413 and AS8428.

Let us now consider the following lines extracted from the AfriNIC delegated file on March 1st, 2018, listing some records of type ASN performed by that RIR. It is worth noting that prefixes allocations are stored in a similar format.

```
afrinic[ZA][asn][2905][1][19930910][allocated][F367678F]
afrinic[DZ][asn][3208][1][20090928][allocated][F366B63E]
afrinic[ZZ][asn][5536][1][available]
afrinic[ZZ][asn][8770][1][reserved]
afrinic[MA][asn][6713][1][20070920][allocated][F36FBD4CD]
afrinic[EG][asn][6879][1][19961216][allocated][F363C475]
```

From the left to the right of these RIR DS entries, we can list the RIR (AfriNIC), the country to which the ASN has been allocated (e.g., South Africa (ZA) for the first line), the type of Internet resource allocated i.e., ASN in all cases, the ASN itself, the count of ASN from this start value, the date at which this allocation/assignment was made in the format YYYYMMDD, the status or type of record (i.e., available, allocated, assigned, or reserved), and an opaque-id.
• The number of ASNs assigned to a country: represents in this paper the number of distinct ASNs assigned or allocated to the said country by its corresponding RIR. If we were to consider the example above, ZA, DZ, MA, and EG would have one assigned ASN each. Note, “ZZ” is a non-valid country code.

• The number of prefixes assigned to a country: represents the number of distinct IPv4/IPv6 prefixes assigned or allocated to that country by its corresponding RIR.

• A prefix $Pref_B$ overlaps another prefix $Pref_C$ when a portion of the IP addresses contained in $Pref_B$ is also contained in $Pref_C$. In other words $Pref_B \cap Pref_C \neq \varnothing$.

4.2.2 IXP growth and business potential

To evaluate the growth of each involved IXP, the number of visible prefixes, origin ASNs, and peering ASNs are quantified per week, month, and year.

The number of visible prefixes at an IXP represents the number of distinct prefixes seen at all its route-collectors. While computing it, bogon prefixes are separated from those routable on the Internet to help identify IXPs at which peers announce more bogon prefixes. Similarly, the distinct origin/peering ASNs visible in the routing data collected at each IXP are listed. While the origin AS (whose identifier is the last ASN from the left) of a given AS path is the network, which originates the prefix, the peering AS (first ASN from the left) is that connected to the IXP route-server, as detailed in Section 4.2.1.

The evolution of those numbers highlights how popular is a local IXP compared to others and how fast it has been growing. It also helps identify IXPs with the highest/stable number of peers or reachable networks in the region/each sub-region, as well as those, which are not functional for a while and the corresponding malfunction period.

With routing data covering the last four weeks, the percentage of prefixes (assigned to each country in the world), which are seen at any local IXP, is then computed. Towards this end, the set of prefixes allocated by an RIR to its country members is fetched from each RIR database. ARDA then verifies if any of the prefixes visible at an IXP overlaps any such allocated prefix. The percentage of prefixes assigned to a given country that are visible at the considered IXP, therefore, represents the ratio of the number of prefixes seen at the IXP that overlapped those assigned to the country to the total number of prefixes assigned to that country.

Such statistics will give IXP members and prospects an accurate knowledge of which countries/regions they are (they will be able) to reach while (after) peering at any IXP in Africa. These are intended to help prospects compare those IXPs by their ability to allow them to reach countries of their interests. The results are presented in Section 4.2.3.

Next, ARDA compares the percentage of IPv4 to that of IPv6 blocks assigned to the country hosting a given IXP, which are seen or not at that IXP. To achieve this, all IPv4 and IPv6 blocks allocated to the country host of the IXP are identified. ARDA then checks if any prefix seen at the considered IXP overlaps any such blocks. The ratio of the number of visible prefixes at an IXP found to overlap the assigned IPv4/IPv6 blocks to the total number of assigned IPv4/IPv6 blocks is then computed.

4.2.3 Interconnection development progress and gaps

The metrics listed in Section 4.2.2 are evaluated at both the national and regional levels to monitor interconnection development growth and gaps. While the national level gathers data from all IXPs in a given African country, the regional level presents data from all IXPs located on the continent.

Technical support

ARDA also reports on networks, which are likely to have routing inefficiencies at each IXP. First, the number of prefixes of various length announced at each peering point over time is quantified. Second, the behavior of IXP members on aggregation and de-aggregation when announcing their respective prefixes is compared to that at their upstream. To inspect this, all assigned prefixes are fetched and individually cross-checked with the set of prefixes visible on the Internet available at 42. We adopted three looking glasses APNIC’s router in Washington, US, APNIC’s router at DIX-IE, Japan, and Bhutan Telecom’s router at LINX, London. This cross-checking enables the identification of allocated blocks, which match those announced on the Internet. The length of the latter prefixes is then contrasted with the length of those visible at each IXP. The goal of this comparison is to single out prefixes whose announcement at the public peering fabric are shorter, match exactly (best practice), or are longer. Performing such an analysis aims at raising awareness amongst IXP members that are not applying the best practice.
4.3  |  Visualizations and Results

Before presenting some showcases of its functionalities, underlining their usefulness, and revealing striking results that demonstrate how ARDA can help profile the African IXP substrate in real-time, we specify technical details related to its implementation.

4.3.1 | Technical choices

For local content to be hosted locally and be as close as possible to most potential users, the server destined to host ARDA was planned to be deployed within the infrastructure of an African IXP. The JINX infrastructure (in ZA) was selected, given the stability it has acquired as the oldest IXP (Table 1) in the region and since several networks are connected to it.

The hardware destined to host the web server was then selected for a high availability service (64 GB of RAM, two Intel Xeon 2.4 GHz processors, redundant power supplies, 18 TB of disk space composed of hot-swappable hard drives, etc.). These choices were also made considering the number of concurrent clients (expected to reach thousands of people), the loads of answering their requests, and the computation tasks that the server will have to support. Next, a Linux – Apache – MySQL – PHP (LAMP) server was built. We only included open source technologies so that anybody can interact with the scripts without expenses, once the code is released.

End-users that will interact with ARDA were classified into two categories: the common users and the administrator. The common user can be an IXP member/operator, an ISP engineer, a decision-making institution, a member of the Internet community, or a researcher. The administrator is responsible for ARDA maintenance and management.

To avoid long waits while end-users are accessing the results, pre-computing the numerical values of each metric was preferred to computing them upon requests. The corresponding set of python scripts, therefore, uses the needed raw data to frequently compute the metrics listed in Section 4.2, so as to deliver up-to-date information to the visualization module. Those outputs are then directed to text files, which are re-used by the PHP and Javascript scripts to display the graphs. Another measure taken to achieve this goal was to physically separate the computation from the visualization module (Figure 3). Three virtual machines (VMs) are thus hosted on the server. The first one termed Pulse is destined to the computations. The second one termed Front-end plays the role of the web server. The last one hosts a monitoring system that supervises the three VMs and the host machine. All of them host the OS Ubuntu 14.04.3-server.

Every 30 mins, the outputs of metrics computations are transferred from Pulse to Front-end, under text files formats, some of which can be downloaded by the end-users upon requests. The adopted technical architecture (Figure 3) allows Front-end, and thus, ARDA to still be functional with end-users accessing the results of the last computations, even if Pulse were to experience a failure. It will also enable caching (Front-end at diverse locations) in the near future.

Further, a MySQL database for hosting the raw routing information, a database for hosting the RIRs assignment data, and another one for user-related information (Figure 3) were built. The former was indexed for more efficiency in the data storage and their provision to our scripts. The main information composing its data structure are the type of the route-collector, the route-collector name, the AS path, the origin ASN, the network, etc. Details related to route-collectors are stored in the same table. Any IXP at which a new route-collector is later deployed has its information automatically added in that table and is included in the next series of computations.

To avoid overloading Pulse, the maximum number of computation scripts running simultaneously was set to 8, given their individual workload. In addition, the historical or real-time data downloaders are always running in the background. An orchestrator was then designed to play the role of tasks scheduler i.e., it identifies per view, every four hours, the script whose end of execution date is the oldest and relaunches it when the maximum number of scripts is not exceeded (Figure 5). By doing so, it ensures that every 15 days, most of ARDA’s analysis results are updated at least once.

Regarding data collection and storage, IPv4 and IPv6 RouteViews real-time data are hourly fetched using BGPStream since June 2016. This operation, combined with the data parsing and storage, usually ends within the first 15 mins of each hour. Meanwhile, IPv4 and IPv6 snapshot are daily fetched from PCH website: it is unfortunately the only way to get this information as of this writing, since there is no API to access this data. On average 8 min are needed per day to download and store the data from the 41 PCH route-collectors of Table 1, while pausing in between any two of them for a random period. Further, we defined the format of the outputs of private route-collector data sources as being the same as those of PCH so that a similar treatment can be applied to both inputs.
It is worth mentioning that IPv4 and IPv6 historical data were downloaded for the period 2005 to end of May 2016. As of April 2017, all PCH and RouteViews historical data were fully downloaded, parsed, and stored. Since then, this dataset has been fetched on a daily basis. The size of the database as of May 2018 is 426.1 GB, and it increases at a rate of roughly 0.6 GB per week, when storing only daily snapshots.

Nevertheless, some issues arose during the implementation of ARDA. As an example, the PCH website was constantly evolving forcing us to often rewrite our downloaders. Moreover, PCH route-collectors are not publicly associated with an IXP. Upon request, we were provided by PCH with this information. Managing the simultaneous run of computations scripts to keep the displayed results always up-to-date was also challenging.

4.3.2 Showcases of the relevance of ARDA

In this section we present as examples two of the different views that ARDA offers (Section 2): the Regional view and the IXP view. We also highlight some interesting results and insights learned from the global picture of the peering ecosystem offered by ARDA before deep diving into individual IXPs to understand how they contribute to the evolution of the African IXP substrate as a whole. To begin with, the user can download the detailed list of values obtained for each metric, used to plot the graphs displayed in those views.

4.3.2.1 Regional View

Our results show that, as of April 15, 2017, only 17.2% of the ASNs assigned by AfriNIC are directly peering at any African IXP involved in this study. Such a percentage highlights the need for increased efforts to improve local interconnectivity in the AfriNIC region. Meanwhile, 58.2% of assigned ASNs are visible as origin ASNs at any such IXP. As of September 18, 2017, these fractions are 17.2% and 57.6% respectively. When considering the same fractions in February 2018, these are respectively 17.4% and 59%. The remaining 41% of ASNs correspond to African networks or content providers that are currently transiting local traffic. Considerably reducing this percentage should constitute a concern for the Internet community. Overall, our findings show that the level of regional interconnection has remained static notably over five months, despite the growing and highly dynamic peering ecosystem in specific countries, such as Kenya (KE) and South Africa (ZA).
other words, when considering, for instance, the global picture of the peering ecosystem in the region over the period April-September 2017, there is almost no perceptible evolution of its size, or perceptible evolution regarding local interconnectivity. However, we show in Section 4.3.2.2 that when inspecting individual IXPs, notably IXPs within KE and ZA, the evolution is significant in the meantime and there is a noticeable boom, which may attract more network operators and content providers. It also means that IXPs in those two countries and their respective models can be singled out as positive examples for the remaining African IXPs.

We then inspect the distribution of ASNs at all IXPs in Africa displayed by Figure 4. We find that in 2005, most (78%) ASNs peering at all IXPs in the region were AfriNIC ASNs against few RIPE and ARIN ASNs. From 2006 to 2008, most (84% – 42.3%) ASNs visible at IXPs in the region were ARIN ASNs. Next, the distribution of ASNs seen at all African IXPs from 2009 to 2012 is similar to that of 2005 (roughly 66% of AfriNIC ASNs), although the amount of RIPE ASNs has slightly increased. We then notice that from 2013 up to 2018, period characterized by the set up and launch of new IXPs in the context of the AXIS project, there is a constant bump of the amount of reserved ASNs, as well as that of ASNs from RIPE, APNIC, and ARIN regions present at all African IXPs. 

This highlights the triggered interests of ASNs from other regions to take part of the peering ecosystem in Africa, and thus the positive outcomes of the AXIS project.

![Figure 4](image.png)

**FIGURE 4** Evolution of the distribution of ASNs at all IXPs in Africa from 2005 to 2018

Last, but not least, the Internet community may wonder (with the rising of concerns about the penetration of IPv6 due to the exhaustion of IPv4 blocks49) which percentage of IPv4 and IPv6 blocks assigned by AfriNIC are seen at any IXP in the region. We learn from ARDA that as of September 2017, in total 51.9 % IPv4 blocks assigned to any African country are seen at one or more local IXPs, whereas only 20.4 % of IPv6 blocks are seen. When considering the same fractions in May 2018, these are respectively 53.5 % and 22.2 %. As a takeaway message more efforts need to be done to drastically increase both the IPv6 adoption and its usage at local IXPs.

### 4.3.2.2. IXP View

Table 2 summarizes the values obtained as of April 15, 2017 and September 18, 2017 for the metrics presented in Section 4.2. ARDA provides a lower boundary of how many networks are peering at each African IXP and identifies those networks. Table 2 indeed shows that the African IXP having the highest number of members connected to its route-collectors is NAPAfrica Cape Town (124 members in April and 144 in September 2017) located in South Africa (ZA). The smallest number of members is 2, registered for SIxP (Sudan (SD)) and AIXP (Tanzania (TZ)) regardless of the month. On average, 21 members are peering at the studied IXPs as of April 2017; this number has increased to 24 in September 2017. Further, at 78.3 % IXPs, notably NAPAfrica, JINX (ZA), TIX (TZ), etc., almost all IXP members are connected to the deployed route-collector: it does not imply that each member peers with everyone at the IXP, however. For instance, one can notice that the number of peering ASNs at KIXP (Kenya, KE) found by ARDA in April 2017 (30) is close to that on KIXP website50 (32). Nevertheless, this is not the case of IXPN (Nigeria, NG) for which the number of detected peering ASNs (6 in April 2017) is really low compared to the 36 members listed on IXPN website51. Peers at IXPs in similar cases (whose names are not followed by a * in Table 2) need to remedy this situation. We remark that this was later corrected by the peers at IXPN in June 2017; consequently, the number of peering ASNs at the said IXP is 37 (identical to the ground truth) as of September 18, 2017.
TABLE 2  IXP View: Overview of some metrics evaluated by ARDA per African IXP in the dataset as of April 15, 2017 and September 18, 2017. IXPs at which almost all members are peering with the route-collectors are followed by a *. N/A stands for no data in the route-collector for the considered period.

<table>
<thead>
<tr>
<th>IXPs (CC host) involved in the dataset</th>
<th>#visible peering (origin) ASNs at the IXP</th>
<th>#visible local (external) origin ASNs at the IXP</th>
<th>#visible prefixes at the IXP</th>
<th>%IPv4 (%IPv6) blocks assigned to the country</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>15/04/2017 18/09/2017</td>
<td>15/04/2017 18/09/2017</td>
<td>15/04/2017 18/09/2017</td>
<td>15/04/2017 18/09/2017</td>
</tr>
<tr>
<td>Benin-IX* (BJ)</td>
<td>5 (8) 6 (9)</td>
<td>3 (5) 6 (5)</td>
<td>176 186</td>
<td>45.8 % (0 %) 48.3 % (0 %)</td>
</tr>
<tr>
<td>BINX (BW)</td>
<td>6 (20) 6 (22)</td>
<td>10 (10) 13 (9)</td>
<td>210 212</td>
<td>64.9 % (0 %) 64.1 % (0 %)</td>
</tr>
<tr>
<td>CAIX* (EG)</td>
<td>3 (67) 2 (65)</td>
<td>47 (20) 49 (16)</td>
<td>3,363 3,078</td>
<td>73.7 % (25 %) 72.8 % (21.4 %)</td>
</tr>
<tr>
<td>SIXP (GM)</td>
<td>6 (9) 6 (9)</td>
<td>7 (3) 6 (3)</td>
<td>66 68</td>
<td>60 % (0 %) 60 % (21.4 %)</td>
</tr>
<tr>
<td>KIXP* (KE)</td>
<td>30 (413) 29 (6,756)</td>
<td>48 (365) 66 (6,690)</td>
<td>3,888 50,126</td>
<td>70 % (38.2 %) 68.5 % (25.4 %)</td>
</tr>
<tr>
<td>LIBERIA-IX* (LR)</td>
<td>4 (8) 4 (9)</td>
<td>4 (4) 6 (3)</td>
<td>88 94</td>
<td>50 % (0 %) 57.2 % (0 %)</td>
</tr>
<tr>
<td>MGIX* (MG)</td>
<td>5 (8) 6 (9)</td>
<td>2 (6) 3 (6)</td>
<td>183 576</td>
<td>50 % (0 %) 72.7 % (0 %)</td>
</tr>
<tr>
<td>MIX* (MW)</td>
<td>N/A (N/A)</td>
<td>N/A (N/A)</td>
<td>N/A N/A</td>
<td>N/A (N/A) N/A (N/A)</td>
</tr>
<tr>
<td>MIXP* (MU)</td>
<td>9 (12)</td>
<td>7 (5) N/A (N/A)</td>
<td>204 N/A</td>
<td>24 % (14.3 %) N/A (N/A)</td>
</tr>
<tr>
<td>MOZIX* (MZ)</td>
<td>12 (23)</td>
<td>12 (11) 14 (11)</td>
<td>339 861</td>
<td>62.5 % (0 %) 65.4 % (0 %)</td>
</tr>
<tr>
<td>WHK-IX (NA)</td>
<td>4 (8) 4 (8)</td>
<td>5 (3) 5 (3)</td>
<td>97 105</td>
<td>53.1 % (0 %) 53.1 % (0 %)</td>
</tr>
<tr>
<td>IXPN* (NG)</td>
<td>6 (109)</td>
<td>77 (32) 120 (68)</td>
<td>1,503 2,264</td>
<td>49.8 % (0 %) 63.5 % (0 %)</td>
</tr>
<tr>
<td>RINEX* (RW)</td>
<td>12 (85)</td>
<td>8 (77) 11 (82)</td>
<td>660 804</td>
<td>77.3 % (0 %) 75 % (10 %)</td>
</tr>
<tr>
<td>SixoP* (SD)</td>
<td>5 (8) 6 (9)</td>
<td>7 (1) 6 (2)</td>
<td>675 692</td>
<td>70.4 % (0 %) 63.3 % (0 %)</td>
</tr>
<tr>
<td>JINX* (ZA)</td>
<td>63 (22,659)</td>
<td>172 (22,487) 294 (24,769)</td>
<td>140,967 162,936</td>
<td>56 % (49.3 %) 61.1 % (47.9 %)</td>
</tr>
<tr>
<td>DINX* (ZA)</td>
<td>15 (165)</td>
<td>58 (107) 166 (146)</td>
<td>1,263 2,462</td>
<td>14.4 % (7.6 %) 40.2 % (17.1 %)</td>
</tr>
<tr>
<td>CINX* (ZA)</td>
<td>19 (464)</td>
<td>148 (316) 211 (240)</td>
<td>4,629 4,685</td>
<td>53.2 % (22.2 %) 51.2 % (20 %)</td>
</tr>
<tr>
<td>NAPAfricaCT* (ZA)</td>
<td>124 (18,022)</td>
<td>171 (17,851) 258 (28,208)</td>
<td>160,418 212,885</td>
<td>46.4 % (29.9 %) 48.6 % (47.5 %)</td>
</tr>
<tr>
<td>NAPAfricaDB* (ZA)</td>
<td>44 (401)</td>
<td>124 (277) 197 (248)</td>
<td>3,669 3,703</td>
<td>28.4 % (11.1 %) 29.9 % (15.4 %)</td>
</tr>
<tr>
<td>AIXP* (TZ)</td>
<td>2 (42)</td>
<td>17 (25) 19 (23)</td>
<td>352 348</td>
<td>32.2 % (42.3 %) 28.2 % (34.1 %)</td>
</tr>
<tr>
<td>TIX* (TZ)</td>
<td>28 (169)</td>
<td>39 (130) 56 (127)</td>
<td>1,324 1,496</td>
<td>78.3 % (50 %) 80 % (43.9 %)</td>
</tr>
<tr>
<td>TUNIXP* (TN)</td>
<td>4 (29)</td>
<td>9 (15) 13 (16)</td>
<td>1,250 1,290</td>
<td>98 % (14.3 %) 98 % (10 %)</td>
</tr>
<tr>
<td>UIXP* (UG)</td>
<td>24 (238)</td>
<td>17 (221) 24 (256)</td>
<td>2,287 2,495</td>
<td>72.1 % (18.8 %) 74.2 % (14.3 %)</td>
</tr>
</tbody>
</table>
Further, we depict on Table 3, the evolution of the African IXPs in terms of peering ASNs. It shows how the highest growth in terms of IXP members from the date of the deployment of the first route-collector are noticed at KIXP (Kenya, KE), NAPAfricaCT (South Africa, ZA), and IXPN (Nigeria, NG). Some IXPs have also declined and currently have no members or have their route-collectors not functional: we can list for example MIXP (Mauritius, MU) and SixP (Sudan, SD). These highlight the need to further invest in projects aiming at deploying fiber networks across countries and improving interconnectivity notably in West (WAf), North (NAf), and Central Africa (CAf) or to implement IXP interconnection scenarios as suggested by for increasing the number of IXPs members, essential to better save on transit costs.

<table>
<thead>
<tr>
<th>Country</th>
<th>IXP</th>
<th>Year of RC deployment</th>
<th># of peering ASNs during RC deployment</th>
<th># of peering ASNs in 02/2018</th>
<th>Growth in terms of IXP members</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benin</td>
<td>BENIN-IX</td>
<td>2015</td>
<td>3</td>
<td>6</td>
<td>200%</td>
</tr>
<tr>
<td>Botswana</td>
<td>BINX</td>
<td>2016</td>
<td>6</td>
<td>6</td>
<td>0%</td>
</tr>
<tr>
<td>Egypt</td>
<td>CAIX</td>
<td>2011</td>
<td>2</td>
<td>2</td>
<td>0%</td>
</tr>
<tr>
<td>Gambia</td>
<td>SIXP</td>
<td>2015</td>
<td>15</td>
<td>24</td>
<td>160%</td>
</tr>
<tr>
<td>Kenya</td>
<td>KIXP</td>
<td>2005</td>
<td>1</td>
<td>31</td>
<td>3100%</td>
</tr>
<tr>
<td>MSA-IX</td>
<td>2017</td>
<td>7</td>
<td>7</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>Liberia</td>
<td>LIBERIA-IX</td>
<td>2016</td>
<td>4</td>
<td>3</td>
<td>-25%</td>
</tr>
<tr>
<td>Madagascar</td>
<td>MGIX</td>
<td>2016</td>
<td>6</td>
<td>8</td>
<td>133.3%</td>
</tr>
<tr>
<td>Malawi</td>
<td>MIX</td>
<td>2013</td>
<td>8</td>
<td>0</td>
<td>-100%</td>
</tr>
<tr>
<td>Mauritius</td>
<td>MIXP</td>
<td>2015</td>
<td>8</td>
<td>8</td>
<td>0%</td>
</tr>
<tr>
<td>Mozambique</td>
<td>MOZIX</td>
<td>2010</td>
<td>9</td>
<td>15</td>
<td>166.7%</td>
</tr>
<tr>
<td>Namibia</td>
<td>WHK-IX</td>
<td>2015</td>
<td>3</td>
<td>4</td>
<td>133.3%</td>
</tr>
<tr>
<td>Nigeria</td>
<td>IXPN</td>
<td>2015</td>
<td>4</td>
<td>38</td>
<td>950%</td>
</tr>
<tr>
<td>Rwanda</td>
<td>RINEX</td>
<td>2015</td>
<td>12</td>
<td>14</td>
<td>116.7%</td>
</tr>
<tr>
<td>Sudan</td>
<td>SixP</td>
<td>2014</td>
<td>2</td>
<td>0</td>
<td>-100%</td>
</tr>
<tr>
<td>South Africa</td>
<td>JINX</td>
<td>2013</td>
<td>12</td>
<td>77</td>
<td>641.7%</td>
</tr>
<tr>
<td>Tunisia</td>
<td>TUNIXP</td>
<td>2014</td>
<td>2</td>
<td>0</td>
<td>-100%</td>
</tr>
<tr>
<td>Uganda</td>
<td>UXIP</td>
<td>2016</td>
<td>24</td>
<td>23</td>
<td>-4.2%</td>
</tr>
</tbody>
</table>

In the perspective of depicting the growth of local peering fabrics overtime, ARDA also gives an insight into the origin ASNs seen at each IXP. As of April 2017, while for the category “peering ASNs,” JINX is the runner-up IXP with 63 ASNs, it appears as the top local IXP for the category “origin ASNs” with 22,659 ASNs (Table 2). This number corresponds to roughly the 2/5 of the total number of networks composing the Internet during that period (57,015 ASNs according to CAIDA’s inferred AS relationships). Five months later, the highest number of origin ASNs (28,466) is seen at NAPAfrica (Table 2), which thus becomes the top local IXP in terms of peers (144). The highest amount of visible prefixes is also registered at that IXP in both April and September 2017 (with 160,418 and 212,885 prefixes, respectively). These results confirm our aforementioned observations regarding the African peering ecosystem as a whole vs. that at individual IXPs (see Section 4.3.2.1).

Another functionality offered by ARDA is the ability to match ASNs visible at an IXP with reachable countries worldwide. Note that all IXPs selected in the following examples can be considered as mature Internet markets in the region, given their launch date. Let us split into 8 categories, the set of origin ASNs visible at KIXP (launched in 2002) and JINX (1996) as examples: local AfriNIC ASNs, which gather ASNs assigned to networks operating/licensed in the country hosting the IXP by AfriNIC; external AfriNIC ASNs i.e., ASNs assigned by AfriNIC to networks operating/licensed in African countries different from the country hosting the IXP; private ASNs; reserved ASNs; RIPE NCC ASNs; ARIN ASNs; LACNIC ASNs; and APNIC ASNs. A detailed analysis of the proportions of ASNs corresponding to these categories in April and September 2018 will shed light on how those IXPs have been growing over that period.

Figures 5b (left) and 5a (left) show that the percentage of KIXP-visible ASNs that belong to the category external AfriNIC ASNs (64.9 %) is higher than that seen at JINX in April 2017 (1.7 %). It is due to the considerable amount of origin ASNs from other regions visible at the latter IXP compared to that of KIXP. Five months later, as the number of external origin ASNs visible at KIXP has increased from 365 to 6,690 (Table 2), one can notice by comparing Figures 5a (left) and 5a (right) that the fraction of external AfriNIC ASNs at KIXP has dropped from 64.9 % to 2.8 %. Meanwhile, the percentage of KIXP-visible ASNs that
belong to the category ARIN ASNs has drastically increased from roughly 5% to 42.5% of all origin ASNs seen at that IXP. The dynamics of the African IXP ecosystem are also noticeable at the above-listed IXPs when considering the evolution of the fractions corresponding to other origin ASNs categories seen at each of those IXP. As an example, the fraction of KIXP-visible ASNs that belong to the category RIPE ASNs has increased from 6.5% in April 2017 to up to 18.5% in September 2017. In the meantime, the percentage of KIXP-visible prefixes belonging to the local AfriNIC ASNs category (13.6%), higher to that seen at JINX (1%) as of April 2017, has drastically decreased to 0.2% (Figures S5 and Figures S5).

Comparing the pie charts of KIXP and JINX – Figures S5 and S5 to those of CAIX (Egypt, launched in 2002), TIX (Tanzania, 2004) – Figures S5 and S5, and RINEX (Rwanda, 2004) hinted the existence of some policy issues at CAIX. In fact, no external AfriNIC ASNs are visible as origin ASNs at CAIX (noticed from August 2011 to September 2017) contrary to the other IXPs, although KIXP, TIX, RINEX, and CAIX were launched during approximately the same period. After discussing with the CAIX operator, our hypothesis was confirmed. We were informed that CAIX does not allow any network operator not licensed in Egypt to peer at the IXP.

As of February 2018, existing publicly available empirical data do not allow to perform a more comprehensive study of the regulations of all 54 African countries. It is not the purpose of this study neither. However, we further investigated the possible motivations of this particular regulation on the CAIX in Egypt (EG). We learned that allowing only licensed networks in Egypt to peer at the local Internet eXchange Point is a law that was legislated at a time that meant well (the 2011 Egyptian revolution) for strategical reasons. It traduces the will of the government to have a police state, a control over the IXP. This is quite understandable, given the interests of the government during that period. It then remained out of convince. In fact, the process of getting licensed is not easy in that the network operator needs to accept to be controlled. For example, one of the conditions is that the government may ask the network to shutdown, which most foreign networks do not accept. However, as EG has won the African Union bid to build a “Regional Internet eXchange Point”, the regulator is more aware of the issue and they promised to do their best to fix it. ARDA further shows how this policy sadly limits the scope of CAIX (cf. Figure 6b(left)), and why the country will gain by removing it. It indeed prevents networks licensed in EG and peering at CAIX from reaching other African networks and even ASes worldwide with a low latency while saving on transit costs. Further, no content provider caches are available at the IXP, again forcing the peers to reach popular content after tromboning through expensive transit links. All these translate into poor QoS for their end-users, as quantified in the following measurements studies.

In addition, ARDA matches origin ASNs visible at each IXP to the countries they have been assigned to by their respective RIRs, and colors those countries depending on the percentage of allocated ASNs seen at the IXP. Such a feature could be of strategic importance when helping the Internet community to understand the reach of networks connected to a given IXP. As an example, Figure 6a highlights the results obtained in the case of KIXP (413 visible origin ASNs in April 2017 and 6,756 visible origin ASNs in September 2017). Figure 6a(left) shows that 50% African countries had no ASN seen at the IXP in April 2017. Further, no ASN allocated to a country in North Africa (NAf) was directly peering or seen at the IXP: this may be due to the closeness of NAf to larger IXPs in Europe. The top five countries whose origin ASNs were visible at KIXP are ZA (69 ASNs), KE (48), the nearby countries TZ (35), Uganda (UG, 20), and finally Brazil (BR, 19). They represent respectively 20.8%, 60%, 59.3%, 62.5%, and 0.4% of the ASNs assigned to the said countries. Few networks assigned to European and North American countries were also visible, which may be of particular interest to progressive out of region networks looking to expand into East Africa (EAf). Figure 6a(right) shows that in September 2017, the reach of the networks peering at KIXP has improved significantly due to the increase in the number of origin ASNs seen at the IXP (Table 2). We can also specify that most of the new ASNs seen at KIXP are those allocated to the US: they correspond to 31.2% of the ASNs visible at the IXP and 34.1% of the number of ASNs assigned to the US. Contrary to April 2017, ASNs allocated to countries in NAf are seen as well; only three African countries still have no allocated ASN seen at KIXP (Chad, Eritrea, and Ethiopia).

Figure 6b(left), which highlights the reach of networks peering at CAIX (EG) contrasts with Figure 6b(right), that of IXPN (NG). 70.5% of networks seen at CAIX are ASNs assigned to EG: no ASNs assigned to another country in the NAf region is seen at CAIX, consequence of the policy adopted by that IXP, which we discussed above. By contrast, networks assigned to countries in West Africa (WAf) and to ZA in Southern Africa (SAf) are seen at IXPN: this does not prevent 61.9% of the ASNs allocated to NG to be visible at IXPN; instead, it enables regional interconnection, which is beneficial to both stakeholders and end-users in terms of QoS and AS paths lengths and is thus essential to the development of the local interconnectivity in the AfriNIC region. This shows why any network operator, even the ones non-licensed in the countries hosting an IXP, need to be encouraged to peer at those IXPs.

Finally, Figure 6c compares the reaches of networks peering at UIXP (UG) and TIX (TZ), to show how well countries in the EAf and SAf regions are interconnected. After its comparison to Figure 6c, one can deduce that CAIX and IXPN need...
FIGURE 5 Percentage of ASNs assigned by each RIR visible as origin ASNs at JINX (South Africa, launched in 1996) and KIXP (Kenya, 2002), CAIX (Egypt, 2002), and TIX (Tanzania, 2004) as of April 15, 2017 and September 18, 2017.
to increase their marketing toward networks operating in other African sub-regions and different continents to expand their reach to such parts of the world. Such an observation is identical when considering most IXPs in WAf, CAf, and NAf. In case this strategy is well implemented, the reach of those IXPs will be similar to those of the largest local IXPs NAPAfrica and JINX (ZA), depicted in Figure 6d.

4.3.2.3. Impact of ARDA on the Internet community

• ARDA adoption

A week after its launch, ARDA counted 389 users connecting from 155 ASes and located in 56 countries worldwide. Table 4 gives more details about the number of distinct IP addresses, which connected to it from April to February 2018, their corresponding ASes and CCs. Notably during network operators’ meetings African Internet Summit (AIS) 2017 and Africa Peering and Interconnection Forum (AfPIF) 2017, several IP addresses from the local ASes offering connectivity and from the countries hosts were frequently connecting to the platform: these explain the peaks in the number of users in May-June and August 2017, respectively. We note that the highest bump in the amount of distinct IPs, ASes, and countries (from which ARDA is accessed) is in December 2017. We found the top five ASes for that month to be AS8075 (Microsoft), AS15169 (Google), AS3352 (Telefonica), AS14061 (DigitalOcean), and AS39572 (AdvanceHosters-AS). Over the whole period, we find Microsoft (US), Google (US), and AdvanceHosters-AS (NL) (which all provide content) to often belong to the top five ASNs from which the platform is accessed. Meanwhile, countries from which most end-users IPs are geolocated are United States, Spain, France, China, Russia, and Brazil. All these traduce the interest generated by ARDA worldwide. Specifically in Africa, we interestingly discover that there is more access to the platform from English-speaking countries than from French-speaking countries: this could be corrected in the future by giving end-users the option of selecting between both languages. However, more checks need to be done on the corresponding IPs and their activities on the platform for separating bots from real end-users.

<table>
<thead>
<tr>
<th>Month/Year</th>
<th># distinct end-users' IPs accessing ARDA</th>
<th># distinct ASes hosting end-user's IPs</th>
<th># distinct CCs in which end-user's IPs are geolocated</th>
</tr>
</thead>
<tbody>
<tr>
<td>04/2017</td>
<td>377</td>
<td>131</td>
<td>33</td>
</tr>
<tr>
<td>05/2017</td>
<td>458</td>
<td>168</td>
<td>50</td>
</tr>
<tr>
<td>06/2017</td>
<td>478</td>
<td>146</td>
<td>48</td>
</tr>
<tr>
<td>07/2017</td>
<td>423</td>
<td>131</td>
<td>48</td>
</tr>
<tr>
<td>08/2017</td>
<td>484</td>
<td>157</td>
<td>66</td>
</tr>
<tr>
<td>09/2017</td>
<td>489</td>
<td>123</td>
<td>45</td>
</tr>
<tr>
<td>10/2017</td>
<td>446</td>
<td>120</td>
<td>42</td>
</tr>
<tr>
<td>11/2017</td>
<td>411</td>
<td>122</td>
<td>43</td>
</tr>
<tr>
<td>12/2017</td>
<td>702</td>
<td>202</td>
<td>65</td>
</tr>
<tr>
<td>01/2018</td>
<td>481</td>
<td>121</td>
<td>42</td>
</tr>
<tr>
<td>02/2018</td>
<td>569</td>
<td>106</td>
<td>39</td>
</tr>
</tbody>
</table>

• The practical value of this study

The main value of this study is to have succeeded in implementing a platform entitled ARDA, which publicly provides Internet developmental institutions, network operators, content providers, IXPs associations, or stakeholders with empirical data to support policy implementation or removal (e.g., case of the CAIX in Egypt), to guide/justify (additional) investments in a country or a sub-region for encouraging more local interconnectivity. Having actual data that highlight the evolution or not of Internet eXchanges in a region from the setup of the Regional Internet Registry (in this case 2005) and in real-time is an asset that may support long term projects aiming at pushing for efforts to improve Internet connectivity in the studied region (see Table 3).

Further, this work plays the role of research-enabler on the African IXP substrate. Any researcher, regardless of her/his location in the world, can now easily download and access pre-computed and publicly available data on each IXP in the African region and use it for research purposes. We show in our results some examples of the various information that could be extracted from the metrics computed by ARDA as well as how they may be interpreted. In addition, the
Percentage of ASNs assigned to each country by its RIR visible at KIXP as origin ASNs as of April 15, 2017 (left) and September 18, 2017 (right).

Percentage of ASNs assigned to each country by its RIR visible as origin ASNs at CAIX, EG (left) and IXPN, NG as of September 18, 2017.

Percentage of ASNs assigned to each country by its RIR visible as origin ASNs at UIXP, UG (left) and TIX, TZ (right) as of September 18, 2017.

Percentage of ASNs assigned to each country by its RIR visible as origin ASNs at NAPAfrica, ZA (left) and JINX, ZA (right) as of September 18, 2017.

**FIGURE 6** Percentage of ASNs assigned to each country (worldwide by its corresponding RIR), which is visible as origin ASNs at selected African IXPs.
source code is publicly available for the replication of this system in other Internet regions, or the improvement of the ARDA platform by its open source community under construction.

5 | ROUTING DATA ANALYSIS

An extensive amount of research has been carried out on RouteViews data such as \[18,57,58,59,42,60\]. The Weekly Routing Table Report \[46,47\] presents a summary of the results obtained by the computations of \[48\] which we use for comparisons purposes in Section \[44\]. Recently, CAIDA has provided the Internet community with its open source software framework BGPStream, which facilitates live and historical BGP data analysis \[49,50\]. In fact, the University of Oregon RouteViews project \[41\] and RIPE RIS \[51\] are the most popular projects operating route-collectors and continuously updating their information. RouteViews manages a passive raw routing data collection system, which stores, under MRT format \[52\], the BGP routes exchanged among the peers at the IXPs at which it is deployed. Its data have been collected on a daily basis since 2004 and are publicly accessible. IXP participants, which peer with RouteViews may agree or not to exchange their full routing tables, thereby providing respectively either a global viewpoint or a peering viewpoint, seen from their respective IXPs. In this paper, the term peering viewpoint refers to the set of AS paths received by a route-collector (deployed at an IXP) to which IXP members solely announce their networks and those of their customers (but neither those of their peers nor those of their transit providers). As of September 2017, there are in total 19 RouteViews collectors in the five Internet regions. In the meantime, 21 RIPE RIS route-collectors, all deployed at IXPs in Europe, aim at achieving the same purposes. Similarly, PCH \[61\] adopted an open peering policy thanks to which it peers with all IXP members that are willing to do so. Contrary to RouteViews collectors, PCH boxes always offer a peering viewpoint, as their peers only exchange the routes of their customers, rather than their full routing tables. Since 2003, PCH has been peering worldwide at 139 IXPs covering 68 countries. The collected data is also made public at \[62\].

Unlike in other Internet regions, only three RouteViews collectors are located in Africa (at KIXP in KE, JINX, and recently NAPAfrica in ZA) as of September 2017. In contrast, PCH route-collectors are deployed at 23 (63 \%) IXPs, including at KIXP and JINX. These are hosted in 18 (33.3 \%) African countries (Section \[4,1\]). Furthermore, some local IXPs deployed their private route-collectors or route-servers with which each member is suggested to peer. These infrastructures enable the collection of exchanged routes locally and facilitate peering setup for newcomers, as noticed by \[63\]. Contrary to RouteViews and PCH datasets, these data are not publicly accessible.

Despite the existence of these facilities, there is a lack of studies on historical routing data collected at African IXPs. In fact, analyzing such data may give a glimpse of how ASes have been behaving at those IXPs, the evolution of those facilities over time, their richness regarding reachable ASes or prefixes, etc. In the context of overall efforts \[22\] to localize traffic, this study is critical for decision-making stakeholders, and the results can also incentivize new ISPs or content providers to join the existing IXPs of their choice, given their interests. These motivate this work that aims at designing and implementing a tool, which can evaluate in real-time key metrics that could help IXPs market their features, report on routing inefficiencies, and make everyone witness the interconnection growth and gaps, etc. Such a goal is inline with previous research such as \[41,51\], which aimed to help improve networks performance or topologies using visualizations tools that relied on collected BGP data.

Finally, it is worth noting that the Isolario measurement project \[22\] also aims at helping improve the knowledge of the Internet AS-level ecosystem by offering a real-time analysis of the interdomain routing from different points of view in return for ISPs’ full routing tables. However, as of this writing, only three African ISPs are involved of which one (Workonline Communications) is known to peer at local IXPs. Further, the design of Isolario contrasts with that of ARDA in that it does not shed light uniquely on the evolution of the IXP substrate in a given region, which is instead more valuable for answering the afore-mentioned questions related to the growth of the peering ecosystem on a regional level (more details in Section \[1\]).

6 | CONCLUSION AND FUTURE WORK

IXPs are known to have the potentiality to drive the growth of the Internet ecosystem in a region through the increase of traffic in local markets \[22,38,39,50,61,62\]. As part of the Internet Society (ISOC) strategy to allow the Internet community to monitor and understand the evolution of IXPs in a particular region, a route-collector data analyzer tool has been designed, and afterward it has been implemented, deployed, and tested in AfriNIC. We have thus obtained the “African” Route-collectors Data Analyzer (ARDA), an open source web platform for analyzing publicly available routing information collected since 2005 by all PCH and RouteViews collectors with a peering viewpoint. In this paper, we present its design and the process of its
implementation. We also shed light on some showcases to highlight its usefulness for the Internet community, and give an idea of the analysis that can be made on its computed metrics.

ARDA provides metrics, which picture the status of the interconnection at local, national, and regional levels. Upon provision of their BGP feeds to a route-collector, local IXPs participants are automatically taken into account. Thanks to this platform, we could understand the evolution of the peering ecosystem in the African region. To allow replication of the platform for other regions and allow the Internet community to openly contribute to it, we make the code source available on GitHub at [https://github.com/rodricfanou/African_Route-collectors_Data_Analyzer-ARDA.git](https://github.com/rodricfanou/African_Route-collectors_Data_Analyzer-ARDA.git). ARDA’s metrics computation results show that a small proportion of the ASNs assigned by AfriNIC (17%) are peering in that region. Through them, roughly 58% of all African networks are visible at one IXP or more. Further, we have noticed that these values have been static from April to September 2017, and are roughly the same in February 2018. Next, we have shown how ARDA can help detect the impact of policies on the growth of local IXPs or automatically provide up-to-date information on the evolution of each IXP in an Internet region and on that of the peering ecosystem as a whole. We believe that this tool will be a helpful compass in the quest for a better traffic localization or new interconnection opportunities in the studied region, since it maintains in real-time, detailed and updated information on its IXP substrate.

As we can obtain routing information solely from the members that accepted to peer with the studied route-collectors, we plan to include IXPs private route-collectors data in ARDA’s dataset in the near future. This will help increase the number of African IXPs covered by the system as well as improve the accuracy of the regional view. Further, integrating route-collectors from other regions (e.g., LACNIC) or deploying ARDA in those regions are possible plans for future work. This will help compare IXPs in both AfriNIC and LACNIC regions on the basis of the evaluated metrics to find out joint solutions for their issues related to Internet connectivity.

7 ACKNOWLEDGMENTS

This work was partially funded by the Internet Society (ISOC). Support to this work was also provided by IMDEA Networks Institute, the National Science Foundation (NSF) CNS-1414177, and NSF OAC-1724853. We are grateful to Nishal Goburdhan and Dibya Khatiwada for their technical support as well as to The African IXP Association (Af-IX), Packet Clearing House (PCH), and Hisham Ibrahim for their cooperation. We wish to thank anonymous reviewers for their invaluable comments, which contributed to significantly improve this manuscript.

References


16. Clark, David and Bauer, Steven and Lehr, William and Claffy, KC and Dhamdhere, Amogh and Huffaker, Bradley and Luckie, Matthew. Measurement and analysis of Internet interconnection and congestion. 2014;.


