

Vipimo

Frank Habicht

September 14, 2007

Abstract

It has become evident that Internet infrastructure in Africa is lacking inter-country connections. For multiple reasons (and excuses) international links are not installed - or too slowly. Small or unknown traffic volumes are one reason quoted. This is a project to establish measurements of Internet Traffic between different networks where direct connections do not exist.

1 Motivation — Aims and Background

At a meeting of African IXP operators in Abuja on April 30, 2007, Michuki Mwangi (and I) made the point that the existing IXPs in East Africa could and should be connected with (direct) terrestrial link. Transport is understood to be outside the scope of an IXP, and rightly so - the “P” stands for “point”. Transport is normally left to carrier networks, deriving profit from this operation. However for East Africa we are of the opinion that the main national IXPs are located relatively closely, and terrestrial infrastructure between them is feasible. Licensing appears to be an issue, as without this requirement any operator in one country could just operate in another as well, and also cross borders inside its network. Also, it appears infeasible for any single operator to establish a link to a distant IXP for only its own traffic; if we could argue for one link aggregating all traffic taking this route it would likely make this setup more feasible with current traffic patterns, as volumes remain relatively low, not requiring multiple links.

For physical networking infrastructure to be established across borders and used for internet traffic, a case is still to be made, including measurement of traffic volumes. The need for supporting data was duly pointed out by Mr. Randy Bush during the above mentioned meeting. With some limitations and exceptions these measurements can be done by individual ISPs, as this is common practice in other regions of the world. However, many African networks don't have their own identifiers (ASN), most don't receive full internet routing information (full BGP view), and for most - especially the smaller - players the development and setup cost for a measurement solution is often too high.

A project of AfrISPA with the name “Ungana” has not yet produced results and does not seem to be too actively pursued at the moment. I had been shown some notes about it. The project described here is based only on original developments, starting 07/07/07.

2 Definitions and Assumptions

2.1 Definitions

IXP route server a server or router at an IXP receiving routing info from networks connected at that IXP and having an open peering policy

Vipimo Route server a BGP speaking server, receiving network announcements from the IXP route servers. Similar to RouteViews¹.

local Vipimo server (Flow cruncher) Server used for this project at an IXP, receiving and evaluating internet traffic information from ISPs' routers.

2.2 Assumptions

There can be more than one IXP per country. In our context an IXP is of greater importance as it specifies one place where networks are currently interconnected

At each participating IXP some networks peer openly and, it is assumed, would also peer with other networks.

The measurement is presently restricted to IPv4 traffic only. This is caused by the Netflow data used and the Perl tools for evaluation.

3 Description

The logical setup of Internet Exchange Points can be quite different from one to the other. IP addresses at the Exchange itself can be public or private, they can be globally "seen"/routable or not. Route servers can exist. These can be mandatory. Other servers and services can exist. Some logical components of this project can be implemented in multiple ways in order to cater for varying setups.

The project relies on local components at each IXP to collect both network reachability information and traffic flow information. These components should be provided and administered by a similar or the same person(s) as the one(s) in charge of administering the IXP. Some parts of this project require to be handled by someone who is sufficiently neutral with regards to the competing ISP peers at an Exchange. In order to produce benefit, the project should be implemented together with a number of IXPs. These will be called participating IXPs.

The project needs as input data from each IXP a list of all IP networks peering there openly. Ideally it is intended to collect this information live at a route collector receiving BGP feeds from the IXPs (similar to RouteViews project). This is preferred as changes will be effective faster and less manual intervention for updates is required. Where this is impossible, an update by email to the central project contact should be made, and an update via authenticated web submission is planned. These should contain a list of CIDR blocks. The possibility of automated retrieval from a looking glass is in consideration.

Lists of IP blocks assigned by RIRs such as AfriNIC can be obtained freely, and could be sorted by country. These are here considered inaccurate, since

¹<http://www.routeviews.org/>

many networks we are aware of still use IP addresses from their upstream providers, that are registered in other countries.

Centrally, a unique number is assigned to each participating IXP.

The central server regularly compiles a list of these identifiers and the associated CIDR blocks. This list is published at a known web address for all participants to download. This information will be accessible by others, which is not posing a problem, as the data is considered to be public information. [opinions?] An alternative would be to disseminate these via BGP (turning the route collector into a route server). This would require a receiving BGP speaker on the Flow collector described later. It would also require a community or other attribute to provide information identifying the IXP at which the block mentioned is advertised.

This information will be used by the local Flow collectors at an IXP. It will also receive Netflow data from ISPs' transit interfaces - where international links terminate. The Netflow data will be evaluated by a lookup in the netblocks tables to determine whether it involves a connection to addresses present at another IXP, and which IXP if so. For each sample interval the traffic volume (in bytes and packets) between the local ISP and any remote IXP gets added up and stored - for both incoming and outgoing traffic. Data will be stored in RRD, and made available through a protected web interface to each ISP. Also it is desirable to aggregate traffic data from all local ISPs to show aggregate traffic of this location with the other aggregation centers (IXPs).

4 Components

4.1 Route server - collecting networks info

It is preferred and assumed that an IXP operates a route server (or route collector). This can automate the collection of information regarding networks available at this IXP. Networks announced to the route collector are then for the purposes of this project assumed to have an open peering policy. This means they would also peer with other networks at the same IXP.

It is intended to collect this information for several IXPs at a central server operated for the project. Preferably it would receive a BGP feed similar to how the Routeviews project of the University of Oregon operates. This information is published as a downloadable file, described below. It could also be distributed via BGP.

For this project, a BGP route collector is running at address rs.eaix.net or 41.222.63.195 . Routing tables are also published as a data file and frequently updated.

4.2 Local flow collectors

At a local IXP one server will have to do some data collection and processing tasks. The server does not necessarily need to be dedicated to this job. Likely an existing one can be used for this project in addition to other functions it performs.

A daemon program running on that server will receive the Netflow information from the local ISPs' routers. store the information temporarily and add

information showing which ISP it came from. Following that another program, that has access to the Routing tables obtained from the central route server, will evaluate the stored netflow data and classify for each if it is communication to or from a network at another IXP.

For each statistics interval of 5 minutes, the amount of traffic any ISP is exchanging with networks at each of the other IXPs is recorded, as bytes incoming and outgoing as well as packets incoming and outgoing. An aggregate for all IXPs at the specific IXP is also calculated.

Programming techniques that scale for a growing number of IXPs sending flows as well as many IXPs and prefixes to look up should be used.

4.3 IXPs international gateways

Routers with external (upstream) connections are required to send Netflow information to the local flow collectors. If Netflow information is already processed, it is also possible to re-export that necessary data from the processing server instead.

This action does transfer internal data of the ISP to an outside entity, and IXPs should consider the risk/benefit of this action.

Specific implementation depends on vendor and model of the equipment and software, one example is below.

4.3.1 cisco code

```
ip flow-cache timeout inactive 120
ip flow-cache timeout active 5
!
interface upstream-external-link
  ip route-cache flow
!
ip flow-export source FastEthernet0/0  ! an internal interface
ip flow-export version 5
ip flow-export destination <your-existing-flow-collector> <port>
ip flow-export destination <ixp-Vipimo-flow-collector> <port>
```

4.4 Central website

A website should be created to inform about the project, reference the local installations and resources and publish collected public data.

4.5 Timing considerations

Flows are only sent to the flow collector some time after the actual packets pass the interfaces. Thus for flows recorded at the Collector, packets may have actually moved at a previous time. This limitation is inherent in the concept of flow collection. For our project however, real time reporting is not important, as the overall volume of traffic is of greater significance.

5 Management

If the project is adopted by some IXPs as intended, it is expected that scope, implementations, regulations, results will be governed by this document.

6 Funding

It is expected that each participating IXP can source a flow collector server independently. Work time can hopefully be done provided on voluntary basis by IXP operators.

7 Security and Confidentiality

As ISPs are submitting internal traffic data to the flow collector server, the confidentiality of this data should be maintained. It is expected that the Flow collector server is controlled by the same trusted team as the neutral switching fabric of the IXP. The data pertaining to transit traffic of any specific ISP should only be accessible by this ISP, it should be password protected.

Processes used for this project should be publicly documented. Sensitive data should be protected

8 Information policy

all processes / procedures / code published on website

9 Code

9.1 Data structures/files

The data files described below should be stored in a subdirectory ./data .

9.1.1 IXP's NLRI

```
#IXP
193.220.232.0/23
194.9.64.0/23
194.9.82.0/23
195.202.64.0/19
```

9.1.2 compiled IXPs' NLRI

```
#IXP_id CIDR
1 216.6.26.0/24
...
1 216.104.206.0/23
2 41.220.128.0/20
...
3 217.199.153.252/30
```

```
4 41.220.0.0/20
4 41.210.128.0/18
```

9.2 Route Server

The route server maintains a number of data files containing CIDR blocks for particular IXPs. These can be updated either manually, via snapshot from a BGP feed, or querying a public route server.

... for querying remote routeservers, Perl module p5-Net-Telnet is installed on the route server.

these individual files will be validated and CIDR blocks will be optimised by

```
perl -w -m'Net::CIDR' -e 'while (<>) {@cidr_list=Net::CIDR::cidradd($_, @cidr_list);}print
```

p5-Net-CIDR cd /usr/ports/net-mgmt/p5-Net-CIDR/ sudo make install clean
and then prepended with the IXP number

9.3 Setup files - flow collector

9.3.1 flow collector startup

in /etc/rc.local:

9.3.2 flow collector config file

```
pidfile "/var/run/flowd.pid"
logfile "/var/log/flowd/flowd"
listen on 127.0.0.1:12345
listen on [::1]:12345
listen on <interface_ip>:9996
store TAG
store SRCDST_ADDR
store PACKETS
store OCTETS
discard all
discard quick src 192.168.0.0/16
discard quick src 172.16.0.0/12
discard quick src 10.0.0.0/8
discard quick dst 192.168.0.0/16
discard quick dst 172.16.0.0/12
discard quick dst 10.0.0.0/8
#ISP1
accept tag 1 agent <flow-exporter1-ip>
#ISP2
accept tag 2 agent <flow-exporter2-ip>
```

9.3.3 crontab - statistics

Crontab entry in /etc/crontab:

```
*/5 * * * * root /usr/local/bin/run_vipimo_stats.sh
```

Shell script `run_vipimo_stats.sh` to be run:

```
#!/bin/sh
mv /var/log/flowd/flowd /var/log/flowd/flowd-`date +%F-%H-%M-vipimo`
kill -HUP|USR1 `cat /var/run/flowd.pid`

# if fast & easy ie we have bgpfeed from RS to local bgpd
#/usr/local/vipimo/bin/update_nlri.sh

#run_perl_cruncher
for flowfile in /var/log/flowd-*
do
  /usr/local/bin/flowd-reader $flowfile | /usr/local/vipimo/bin/vipimo.pl
  if $? .....
  rm or mv $flowfile
done
update nlri.txt and flowd.conf
# log?
```

9.4 FlowCruncher program

- init arrays clear comments: `<nlri.txt grep -v ' *#' - parse stdin (or file in ARG)`

9.5 Installation instructions

9.5.1 OpenBSD

9.5.2 FreeBSD

```
flowd:
cd /usr/ports/net-mgmt/flowd && make install clean
or pkg_add -r flowd
  p5-Net-CIDR
cd /usr/ports/net-mgmt/p5-Net-CIDR/
sudo make install clean
```

```
Patricia (lib):
cd /usr/ports/net/p5-Net-Patricia && make install clean
or pkg_add -r p5-Net-Patricia
  pkg_add -r rrdtool
  mkdir /var/log/flowd
```

9.5.3 RRD files

install flowd from ports config `/etc/flowd.conf` make flowd start automatically on server startup it will logflow info into a binary file periodically (every 5 minutes), a script should be run that will: - move the current log file - send a SIGUSR1 to the flowd process (`kill -USR1 `cat /var/run/flowd.pid``) - evaluate the binary flow log file by: `flowd-reader |logfile; ...` — special program

evaluating and updating RRDs - then move the logfiles to archive or delete them periodically (daily?) a file

10 Glossary

CIDR Classless Interdomain Routing - RFC 4632

CIDR block ip network block like 196.223.0.0/16

IXP Internet Exchange Point

BGP Border Gateway Protocol, defined in RFC 4271 with extensions in RFC 4760

ASN Autonomous System Number

route server BGP speaking router or server collecting and distributing routing information

open peering policy to interconnect with any other willing network at no cost

NetFlow format of internet traffic measurement information sent from routers to servers for evaluation

11 FAQ

how to join email Frank