AfriNREN Project Proposal

ROBERT PASSMORE, University of Cape Town
ROSYLN SANBY, University of Cape Town
CHANTAL YANG, University of Cape Town

HUSSEIN SULEMAN, University of Cape Town - Supervisor
JOSIAH CHAVULA, University of Cape Town – Co-supervisor

1. PROJECT TITLE

AfriNREN: Visualising Internet traffic and network structure of African NRENS

2. PROJECT DESCRIPTION

A National Research and Education Network (NREN) is a mesh of interconnected networks that support the needs of education and research communities in a country [Fryer 2012]. In order to facilitate better research and communication, NRENs aim to reduce latencies between educational institutions, promote bandwidth sharing and improve traffic engineering [Fryer, 2012].

Previous research on National Research and Education Networks (NRENs) in Africa has shown that 75% of traffic originating from and destined to African institutions traverse circuitous routes through Europe [Chavula et al. 2014]. This indicates that there is a high latency in traffic exchange in and between African NRENs [Chavula et al. 2014]. Latency is a measurement of the time it takes for a packet to move from its source to its destination and for the receiver to acquire the acknowledgement packet [Chavula et al. 2014]. Thus, latency has an effect on network performance. This is of particular importance to NRENs that facilitate cross-border research collaborations, where it is likely that large amounts of data need to be transmitted at high speeds [Fryer 2012].

That being said, little research has been conducted on quantifying the amount or nature of traffic in and between NRENs. By understanding the traffic patterns of NRENs and having a tool for visualising the physical (actual) and logical (apparent/abstract) topology of NRENs, better insight could be gained into why such high latencies occur. For this reason, the project will involve the collection of data for network topology discovery of African NRENs and the analysis of NetFlow Internet logs. Topology mapping will be done, making use of various forms of traceroute, a network diagnostic tool that determines end-to-end paths of packets and relays round trip time (RTT) at each hop. Internet NetFlow data logs will be provided by institutions participating in the study and be used to determine the type of traffic sent within and between NRENs (eg. http, p2p, etc.).

Once this is done, a Web application will be built that displays visualisations of the network topologies and internet traffic, making use of both the collected traceroute data and NetFlow Internet logs. Two types of visualisations will be implemented: a logical visualisation (non-geospatial) that presents an abstracted

representation of the physical network topology; and a geographical/geospatial visualisation, in which network structure is related to physical locations. Both of these will also visualise the collected NetFlow data of the traffic exchanged between locations in the network.

It is hoped that such a visualisation tool will aid in future research and help network engineers manage and discover why such high latencies occur in and between NRENs.

3. PROBLEM STATEMENT

The network managers and engineers of Africa's research and educational institutions and those of the UbuntuNet Alliance seek to improve the performance of information exchange between these institutions to facilitate the network-intensive collaborative work done between them [Fryer 2012; UbuntuNet Alliance 2014]. Currently, there is a lack of research into effective network data collection and visualisation tools in the context of the UbuntuNet and African NRENs. This research would be useful in that it helps develop tools to facilitate network design, planning and operation and, in the distinct context of NRENs, aids policy makers in the determination of whether the establishment of an NREN between two communicating institutions within the same country is viable [Karagiannis et al. 2005; Kim et al. 2007; Rossi and Valenti 2010].

3.1 Aims

With this in mind, the following aims have been established for this research project:

- Collect traceroute and NetFlow data from different vantage points within UbunutuNet
- Evaluate the type and amount of traffic exchanged between African NRENs
- Evaluate the quantity of traffic that either originates or terminates at NRENs but is destined for, or originates from, networks outside of an African NREN
- Identify the routes used for traffic exchanged between African NRENs
- Format the collected data to be used in a visualisation
- Create a visualisation tool to display the collected data and assist in analysis of the data and management of the network; aiding in the identification of patterns, gaps or anomalies.

3.2 Research Questions

The project seeks to answer the following research questions:

- Can one efficiently and reliably collect traceroute data for this purpose, reducing the number of measurements to perform and increasing the accuracy of these measurements?

- Can a dashboard of interactive non-geospatial visualisations of network topology and NetFlow data effectively communicate network structure, latencies and network traffic information about a network to the network managers?

- Can a geospatial visualisation effectively and accurately communicate the network topology of African NRENs, allowing users to identify networks, where they connect and the routes this traffic traverses as well as convey network traffic information such as latencies and network traffic type?

4. PROCEDURES AND METHODS

Given the layered nature of the project, procedures and methods can be split into those that are used for the collection of data and those used for the development of the visualisations. The visualisation is further split into two parts: a geospatial visualisation and a non-geospatial visualisation.

4.1 Traceroute Data Collection

Traceroute data provides information about the routes taken by network traffic, Round Trip Time (RTT) and direction of flow of traffic [Motamedi et al. 2013]. This will be useful for the visualisation as it will allow us to map paths between nodes to show how traffic is exchanged between NRENs. It has been found that having many vantage points (points from which to launch traceroutes) in a diverse spread of locations is advantageous for unbiased results [Shavitt and Weinsberg 2009; Shavitt and Shir 2005].

In order to collect traceroute data, we will make use of the RIPE Atlas platform. RIPE Atlas is a network of probes that provides a real-time understanding of the Internet by measuring Internet connectivity and reachability. RIPE Atlas has the most vantage points in Africa as it has over 200 probes in Southern East Africa. These probes are small USB-powered hardware devices attached to an Ethernet port that conduct measurements (ping, traceroute, DNS and SSLcert) which are relayed to the RIPE Network Coordination Centre (NCC), a Regional Internet Registry. This data is then aggregated with data collected from other RIPE Atlas probes [RIPE Atlas n.d a]. As this project is researching traffic between and within NRENs, it is important for probes to be hosted within NRENs. RIPE Atlas

gives us access to many probes, not only in Africa, but also within NRENs in Southern East Africa.

There are three main steps to collecting the traceroute data:

1. Identify the probes from which we would like to launch network measurements, as well as destination IP addresses belonging to different universities in Southern East Africa.
   The identification of these probes will either be done manually by going through the map of Atlas probes showing the Atlas Network Coverage and choosing each probe to use as a vantage point, or by using the RIPE API to get probe information [RIPE Atlas n.d b]. As we are targeting the Universities and research centres in the UbuntuNet Alliance, we will identify probes and IP addresses in Southern East Africa that are situated in NRENs and/or universities.

   The results from NetFlow analysis will also be used as input for traceroute probes to determine which paths the traffic in the NetFlow data takes.

2. Launch user defined measurements using the measurement creation API. This involves writing scripts specifying destination IP addresses and ASes of vantage points as well as the measurement type (in our case traceroute) and protocols to use. The scripts are then run using the 'curl' command to push the scripts to the server.
   We will perform traceroute measurements to identify round trip time (RTT) and hops between network points in Africa. We will try three different protocols for these traceroutes - namely ICMP, UDP and TCP - to see if the use of different protocols affect the measurements gathered. We will also compare the use of normal traceroute versus Paris traceroute [RIPE Atlas n.d c].

   To set up and launch a network measurement with RIPE Atlas, one needs to use credits that are collected by hosting an online probe. UCT hosts a probe and has, as a result, collected credits which we will be able to use to conduct our experiments. Optimisations will be done to minimise the number of traceroutes required. These optimisations include graph analyses of traceroutes to minimise redundant traceroutes.

3. Obtain results using the RIPE Atlas REST API [RIPE Atlas n.d d]. The RIPE Atlas REST API is used to fetch the results of the measurements in a JSON format, which will be formatted to be available for use by the visualisation tools. This data will be stored in a database such as MongoDB or CouchDB.

Once these results are obtained, the AS numbers and IP addresses will need to be mapped to physical locations for use by the geospatial visualisation. This will be done using the MaxMind GeoIP Database.

4.2   Netflow Log Data Formatting

NetFlow logs are collected by a number of network managers within some African NRENs. Acquiring this data relies on their cooperation. For the purpose of this project, we have already acquired NetFlow logs of two institutions (the Kenyan NREN (KENET) and the Uganda Christian University (UCU)).

Once this data has been collected, it will then need to be processed and explored using the nfdump and nfSen tools. nfdump is a set of tools to process NetFlow data [Sourceforge.net 2014], while nfSen provides a graphical web-based summary of the NetFlow data [Sourceforge.net 2011].

After we have explored the data and decided how to visualise it, we will format the data for our use. This format will depend on the implementation of the visualisations but will be serialised in JSON format for ease of communication across applications.

4.3   Visualisation

Visualisations will be designed to be shown in a Web browser, making use of the collected traceroute data and NetFlow Internet logs. This will allow users to access and interact with the visualisation easily.

The following tools can be used to create a Web application that displays and allows interaction with the visualisation:

- **Web Development:**
  - Python 2.7 or Python 3 [Python.org 2015]
  - SQLAlchemy ORM toolkit [Sqlalchemy.org 2015]
  - PostgreSQL [Postgresql.org 2015]
  - Flask Web framework [Flask.pocoo.org 2015]
  -
- **Visualisation Development:**
  - D3.js [GitHub 2015]
  - Processing and Processing.js programming languages [Processingjs.org 2015]
  - amcharts.js [amCharts 2015]

We will use a User-Centred Design (UCD) approach when designing the geospatial and non-geospatial visualisations. User-Centred Design describes a process where

users are involved in every step of the Design Process Cycle (Design, Prototype, Evaluate) [Abras et al. 2004]. This will involve obtaining initial Visual Queries (an information need addressed by a visualisation) from potential users of the visualisation tool during the Design Stage. Once this has been done, different visualisation prototypes (Prototype Stage) for both the geospatial and non-geospatial will be developed and then presented to user for evaluation (Evaluation Stage). Upon receiving feedback, the design will be edited incorporating this feedback before being presented to users once again. Iteration will conclude once users are able to answer a list of predefined visual queries with ease using the visualisations we have developed.

The effectiveness of these visualisations will be evaluated using feature-based evaluation frameworks, design guidelines as well as metrics and techniques outlined in relevant published literature.

5. ETHICAL AND LEGAL ISSUES

As effectiveness of the visualisation will depend on input from users, consent from users during user testing will be required.

This project also relies on the use of data gathered by institutions for primarily internal use. The KENET and UCU institutions have granted us the use of the data for our research purposes. It is important to formalise the agreement of use with the original holder of the data and to insure that we abide by the terms of use of said agreement. This might manifest in distributing our research tool - along with the data we used - to other parties.

6. RELATED WORK

Research was conducted by Gupta et al. [2014] and Chavula et al. [2014] that found that Internet traffic in Africa follows circuitous routes through Europe. Studies focused on traffic exchanged between and within NRENs specifically, found that about 75% of traffic sent from Africa and destined for Africa follow circuitous routes through Europe and that these circuitous routes result in higher latencies [Chavula et al. 2014].

It was identified that the use of (some form of) traceroute would be appropriate for topology discovery at both an Autonomous System (AS) and PoP level despite its challenges and limitations which have been well-documented [Mao et al., 2003; Donnet et al., 2007; Motamedi et al, 2014]. Additionally, while there have been studies conducted on mapping Internet topology within Africa [Gilmore et al. 2007], studies in relation to the topology of NRENs on the continent have not been found.

Many platforms have been developed to measure the topology of the Internet, but many of them do not have enough vantage points in Africa. Since the focus of this research project is on NRENs in Africa, multiple vantage points within Africa are necessary for an unbiased dataset [Shavitt and Weinsberg. 2009, Shavitt and Shir. 2005]. It would appear that measurement tools such as RIPE Atlas would be ideal

to measure the connectivity between and within African NRENs, as it appears to have more vantage points inside African NRENs, and also because it allows the use of many probes from many vantage points in different locations [RIPE Atlas n.d b].

An analysis of literature on information visualisations found that the principles of Schneiderman's [1996] Visual Information Seeking Mantra are an important guideline for interface design. In terms of network visualisations specifically, use of graph representations of nodes and links is common, even though Fowler et al. [2014] presents a unique representation of AS level topology that differs from this. In Withall et al. [2007], literature on numerous network visualisation applications are presented and although the paper classifies these into geographic visualisations (nodes are represented in respect to physical locations), abstract topological visualisations (nodes are independent of physical locations) and plot-based visualisations (focus on a single network point with respect to time), the effectiveness of these visualisations is not critically evaluated on the basis of any design principles.

7. ANTICIPATED OUTCOMES

7.1 System

We expect to develop a functional Web application that allows a user to interactively explore visualisations of the network structure and network traffic exchanged between research institutions and NREN members of the UbuntuNet Alliance, based on the collected data (traceroute data and Netflow Internet Logs).

7.2 Key Features

This Web application will use a database for access to the information required to render the visualisations.

The key features of the visualisation will be developed iteratively with the users but the initial features to be presented with the user have been detailed as follows.

The user can select between viewing the network topology and the network traffic data visualisations or both. The user can also view the data in a geospatial format - where the network structure is projected onto a geographical map - or in a no format - where the network is represented as an abstract graph of nodes and links on a 2D plane.

In each visualisation format (geospatial and non-geospatial), the user will be able to interact with the visualisation by zooming, panning and filtering the data with pre-defined filters. The filters will allow the user to choose the level of network topology they wish to view (AS- or PoP-level). The filters will also allow the user

to highlight a pair of nodes in the visualisation and the routes of links and nodes that connect the pair.

Another feature to be implemented is to visualise the data from different viewpoints. Since the traceroute data is not symmetric (it may travel along different paths depending on which direction the data is flowing in), visualisations would differ based on chosen vantage points.

7.3 Expected Impact of Project

The expectation is that, upon completion of the project, users of the system (network managers/monitors/policy makers) will have the tools needed to effectively explore the structure and traffic information of the UbuntuNet Alliance Regional Research and Educational Network - helping users of the system to deduce where latencies are high and what the effect of latencies and circuitous routes are on the traffic flow in Sub-Saharan Africa.

It is expected that the geospatial visualisation of the network will be more effective at communicating the topology of the network while the non-geospatial visualisation will be most effective at communicating the NetFlow traffic data of the network.

7.4 Key Success Factors

The intended user can successfully answer the primary visual queries that represent their functional requirements of our system. Although these visual queries and features will only be properly determined after further collaboration with a group of users, the following initial examples have been developed:

- Which institutions/networks exchange large quantities of data for research and educational purposes?
- What is the route taken by network traffic between institution X and institution Y?
- How much of the intra-African network traffic between network X and Y travel along intercontinental circuitous routes?

Another success factor is whether the data collected is of a high enough accuracy to be relied upon for the user's requirements. This accuracy will be deduced by comparing the results collected from performing traceroutes with different protocols (ICMP, UDP and TCP) as well as by conducting different types of traceroutes - namely ordinary and Paris traceroutes. The accuracy will also be measured based on comparisons with other tools and published results to verify the data. Furthermore,
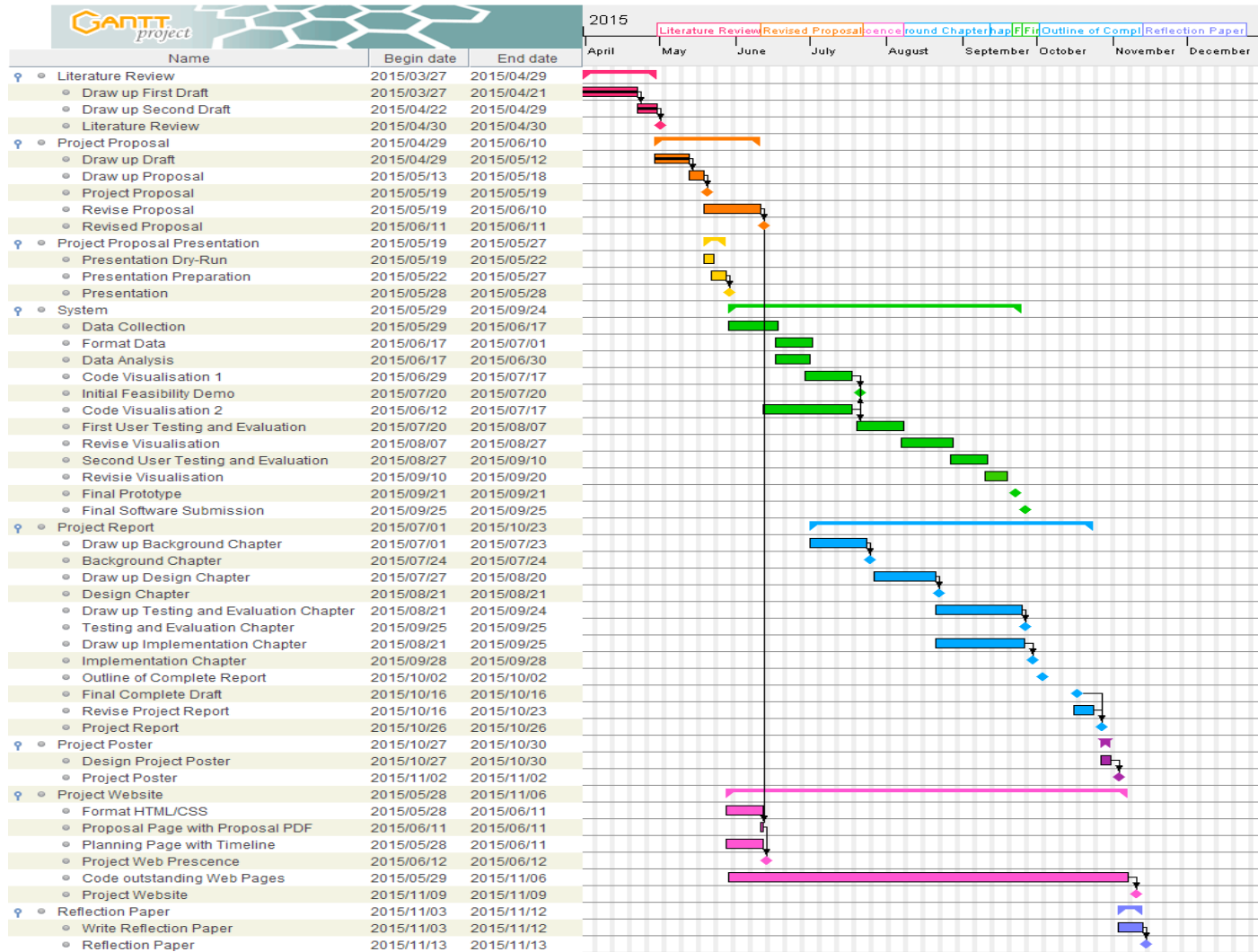
8. PROJECT PLAN

8.1 Risks

Table1: Risk Matrix for AfriNREN Project

Risk	Probability	Impact	Mitigation strategy
Loss of Team Member (eg. dropping out)	Low to medium. As we are already almost half way through the year, this risk is unlikely	High. This would result in the scope needing to be adjusted. As the tasks identified are separable and data is already available, the scope of the project would change either to only collecting data, only visualising data geographically or only visualising data logically. The visualisations would also only make use of the NetFlow data if the collection of data aspect is dropped.	Keep each team members section of work independent of the others'. In the case of dropping out, the remaining team members can continue work without drastic setbacks in the project.

Additionally, insist that all work done is in repository that is accessible by all team members. The work that is contained within this repository should be well documented to allow recovery from team member loss. |
| Missing deadlines | Low. All team members have fewer modules scheduled for the third block, so more time will be available to work on this project from July. However, two members of the team will be going away for two and four weeks respectively during the June/July holiday. This will increase time-pressure and difficulty in meeting deadlines at the beginning of the third block. | Medium to high. The impact of this would be quality of work as handins might be rushed. Also, by missing one deadline, this could cause an avalanche effect where all subsequent deadlines are also missed. | We will begin work before the holiday period and keep each other accountable for work needing to be done. |
| Loss or corruption of data | Low to medium | High. As the visualisations are based both on the data that we already have and on the data we will be collecting, loss or corruption of this data would result in visualisations with minimal data or data that is not realistic. | Backup data regularly. |

Scope creep	Medium. Most tasks are clearly defined, however one can get carried away with design features.	High. If we do too much, the quality of the project could be affected.	Set a realistic scope and make sure all members of team adhere to it.
Load shedding	High. With load shedding being implemented more frequently, this risk is very likely to happen.	Medium to high. If there is load shedding at campus and laptops are not charged, no work will be able to be done on campus. If any team members experience load shedding while working from home, members have the opportunity to work from another team member's home as team members live near enough to each other for quick travel, yet all are in different load shedding zones.	Save work regularly to avoid losing work when load shedding strikes. Make sure laptops are charged. Keep track of load shedding schedules and plan to have a backup location available at all times that isn't being load-shed. The de facto location being Rosy Sanby's house which has yet to experience loadshedding.
Illness of team member	Medium. Winter time does lead to more illnesses.	High. As the project team consists of three people, if one person were to fall ill, this would have a significant impact on the project.	Ensure that work is separable enough so that one section is not entirely reliant on another. This way, each member can work on their own section of the project without depending on other team members.

8.2 Timeline

Figure 1: Gantt Chart for AfriNREN Project

Name	Begin date	End date
Literature Review	2015/03/27	2015/04/29
Draw up First Draft	2015/03/27	2015/04/21
Draw up Second Draft	2015/04/22	2015/04/29
Literature Review	2015/04/30	2015/04/30
Project Proposal	2015/04/29	2015/06/10
Draw up Draft	2015/04/29	2015/05/12
Draw up Proposal	2015/05/13	2015/05/18
Project Proposal	2015/05/19	2015/05/19
Revise Proposal	2015/05/19	2015/06/10
Revised Proposal	2015/06/11	2015/06/11
Project Proposal Presentation	2015/05/19	2015/05/27
Presentation Dry-Run	2015/05/19	2015/05/22
Presentation Preparation	2015/05/22	2015/05/27
Presentation	2015/05/28	2015/05/28
System	2015/05/29	2015/09/24
Data Collection	2015/05/29	2015/06/17
Format Data	2015/06/17	2015/07/01
Data Analysis	2015/06/17	2015/06/30
Code Visualisation 1	2015/06/29	2015/07/17
Initial Feasibility Demo	2015/07/20	2015/07/20
Code Visualisation 2	2015/06/12	2015/07/17
First User Testing and Evaluation	2015/07/20	2015/08/07
Revise Visualisation	2015/08/07	2015/08/27
Second User Testing and Evaluation	2015/08/27	2015/09/10
Revisie Visualisation	2015/09/10	2015/09/20
Final Prototype	2015/09/21	2015/09/21
Final Software Submission	2015/09/25	2015/09/25
Project Report	2015/07/01	2015/10/23
Draw up Background Chapter	2015/07/01	2015/07/23
Background Chapter	2015/07/24	2015/07/24
Draw up Design Chapter	2015/07/27	2015/08/20
Design Chapter	2015/08/21	2015/08/21
Draw up Testing and Evaluation Chapter	2015/08/21	2015/09/24
Testing and Evaluation Chapter	2015/09/25	2015/09/25
Draw up Implementation Chapter	2015/08/21	2015/09/24
Implementation Chapter	2015/09/28	2015/09/28
Outline of Complete Report	2015/10/02	2015/10/02
Final Complete Draft	2015/10/16	2015/10/16
Revise Project Report	2015/10/16	2015/10/23
Project Report	2015/10/26	2015/10/26
Project Poster	2015/10/27	2015/10/30
Design Project Poster	2015/10/27	2015/10/30
Project Poster	2015/11/02	2015/11/02
Project Website	2015/05/28	2015/11/06
Format HTML/CSS	2015/05/28	2015/06/11
Proposal Page with Proposal PDF	2015/06/11	2015/06/11
Planning Page with Timeline	2015/05/28	2015/06/11
Project Web Prescence	2015/06/12	2015/06/12
Code outstanding Web Pages	2015/05/29	2015/11/06
Project Website	2015/11/09	2015/11/09
Reflection Paper	2015/11/03	2015/11/12
Write Reflection Paper	2015/11/03	2015/11/12
Reflection Paper	2015/11/13	2015/11/13

8.3.1   Human

Appropriate users with sufficient knowledge of networks will need to be identified for participation in user testing of the effectiveness of visualisations.

8.3.2   Technological

Software

The software used for the collection and preparation of the data into different formats is free and open-source and therefore easily acquired (eg. RIPE Atlas platform, MaxMind GeoIP Database and lookup tool, Python) [RIPE Atlas n.d a; MaxMind GeoIP 2015; Python 2015]. Having said that, it is possible that a commercial version of the MaxMind GeoIP Databases (eg. GeoIP2 City vs GeoIP2Lite City) may need to be purchased in order to obtain more accurate results for the geolocation (coordinates) of IP addresses [MaxMind GeoIP 2015].

Furthermore, the software used to develop and host the Web application as well as the visualisations is also free and open-source.

Hardware
Development Hardware:
No special hardware is required to develop the proposed project system. Personal laptops and the computers available in the UCT Computer Science Honours lab can be utilised for development purposes.

Probing hardware:
The traceroute tests need to be run from hardware probes located around the continent within different research institutions. Access to the network of probes is gained by spending special credits given to owners of the probes. We have access to enough credits to conduct our planned data collection.

Data
NetFlow data captured within different African universities is required. We are in contact with these universities and has already delivered an initial dataset from KENET, the Kenyan NREN, and Uganda Christian University (UCU).

Traceroute data needed to understand the network topologies is collected using the RIPE Atlas platform. We will conduct user defined measurements using the Atlas API and then retrieve the results using the RIPE API [RIPE Atlas n.d d].

8.3.3 Financial

As previously mentioned, a commercial version of MaxMind GeoIP Databases may need to be purchased in order to obtain more accurate results. Several pricing options exist for this including the purchasing of a site licence in which a downloadable database can be obtained for US$370 (R4399.17); a web service of 50 000 lookups for US$20 (R237.79) or databases for individual continents (eg. Africa and Europe) starting at US$200 (R2377, 93). This will likely be funded by the University of Cape Town's Computer Science Department.

8.4 Deliverables

The following table presents scheduled deliverables and milestones for the project. Hard deadlines are indicated in bold font while non-bold items are soft deadlines as suggested in the Project Milestones PDF.

Table 2: Deliverables for AfriNREN Project

Deliverable	Due Date
Literature Review	**28 April 2015**
Project Proposal (including project plan)	**19 May 2015**
Presentation of Project Proposal	**26 & 27 May 2015**
Revised Proposal Finalized	11 June 2015
Project Web Presence: proposal, timeline/plan	12 June 2015
Initial Feasibility Demonstration	**20-24 July 2015**
Background/Theory Chapter	24 July 2015
Design Chapter	21 August 2015
Final prototype/experiment/performance test and write up	21 September 2015
Implementation and Testing Chapters (coding complete)	25 September 2015
Outline of complete report	2 October 2015
Final complete draft of report (Decide weighting for marking)	**16 October 2015**
Project report final submission	**26 October 2015**
Poster	**2 November 2015**
Website	**9 November 2015**
Reflection Paper	**13 November 2015**
Open Afternoon/Evening	18 November 2015

All Members
Web interface in which visualisation is embedded.

Rosy
Collection of traceroute data using user defined measurements using the RIPE Atlas platform and formatting of data. Verify and evaluate the results obtained from measurements.

Chantal
The creation of the geospatial graph visualisation of traceroute and netflow data. User testing of the geographical graph visualisation to evaluate its effectiveness.

Rob
The creation of the non-geospatial visualisation dashboard displaying traceroute and netflow data. User Centred Design of the non-geospatial graph visualisation and evaluation of its effectiveness using appropriate methods.

REFERENCES

Abras, C.,Maloney-Krichmar, D. & Preece, J. 2004. User-centered design. *Bainbridge, W. Encycl. Human-Computer Interact. Thousand Oaks Sage Publ.* 37, 4 (2004), 445–56. DOI:http://dx.doi.org/10.1.1.94.381

amCharts, (2015). *JavaScript Charts and Maps | amCharts.* Retrieved May 18, 2015 from amCharts: http://www.amcharts.com/

Becker, R. A., Eick, S. G., & Wilks, A. R. 1995 Visualizing network data. *Visualization and Computer Graphics, IEEE Transactions on*, 1(1), 16-28.

Chavula, J., Feamster, N., Bagula, A., & Suleman, H. 2014. Quantifying the Effects of Circuitous Routes on the Latency of Intra-Africa Internet Traffic: A Study of Research and Education Networks. in Proceedings of the AFRICOMM conference, (Uganda, 2014), AFRICOMM,1-10.

Donnet, B., & Friedman, T. 2007. Internet topology discovery: a survey. Communications Surveys & Tutorials,*Communications Surveys & Tutorials*, IEEE9, 56-69.

Flask.pocoo.org, (2015). *Welcome | Flask (A Python Microframework),* Retrieved May 18, 2015 from Flask: http://flask.pocoo.org/

Fowler, J. J., Johnson, T., Simonetto, P., Schneider, M., Acedo, C., Kobourov, S., Lazos, L. 2014. IMap: visualizing network activity over internet maps. in *Proceedings of the Eleventh Workshop on Visualization for Cyber Security*, ACM, 80-87.

Fryer, T., Research and Education Networks around the World and their Use. in *Segunda Conferencia de Directores de Tecnología de Información*, (Lima, 2012), Red Clara, 1-13.

Gilmore, J. S., Huysamen, N. F., & Krzesinski, A. E. 2007. Mapping the african internet. in *Proceedings Southern African Telecommunication Networks and Applications Conference (SATNAC), (Sept 2007), Mauritius*, South African National Research Foundation.

GitHub, (2015). *mbostock/d3*. Retrieved May 18, 2015 from D3: https://github.com/mbostock/d3/wiki

Gupta, A., Calder, M., Feamster, N., Chetty, M., Calandro, E., & Katz-Bassett, E. 2014. Peering at the Internet's frontier: A first look at ISP interconnectivity in Africa in *Passive and Active Measurement*, (Los Angeles, 2014), Springer International Publishing, 204-213

Luckie, M., Hyun, Y., & Huffaker, B. 2008. Traceroute probe method and forward IP path inference. in *Proceedings of the 8th ACM SIGCOMM conference on Internet measurement*, (Vouliagmeni, 2008), ACM, 311-324 .

Mao, Z. M., Rexford, J., Wang, J., & Katz, R. H., 2003. Towards an accurate AS-level traceroute tool. in *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, (Karlsruhe, 2003), ACM, 365-378.

Maxmind.com, (2015). *MaxMind - GeoIP2 City.*, Retrieved 18 May, 2015 from MaxMind - GeoIP2 City: https://www.maxmind.com/en/geoip2-city.

Motamedi, R., Rejaie, R., Willinger, W., A Survey of Techniques for Internet Topology Discovery*, IEEE Communications Surveys & Tutorials*, 1-25.

Postgresql.org, (2015). *PostgreSQL: About*, Retrieved May 18, 2015 from PostgreSQL: http://www.postgresql.org/about/

Processingjs.org, (2015). *Processing.js*. Retrieved May 18, 2015 from Processingjs: http://processingjs.org/

Processing.org, (2015). *Processing.org*. Retrieved May 18, 2015 from Processing.org: https://processing.org/

Python.org, (2015). *Welcome to Python.org*. Retrieved 18 May, 2015 from Python.org: https://www.python.org/.

a   RIPE Atlas, (n.d), *What is RIPE Atlas?,* Retrieved May 7, 2015 from RIPE Atlas: https://atlas.ripe.net/about/

b   RIPE Atlas, (n.d), *Network Extent, Coverage, and Statistics*, Retrieved May 7, 2015 from RIPE Atlas: https://atlas.ripe.net/results/maps/network-coverage/

c   RIPE Atlas, (n.d), *Creating Measurements with the RIPE Atlas Restful API*, Retrieved May 7, 2015 from RIPE Atlas: https://atlas.ripe.net/docs/measurement-creation-api/

d   RIPE Atlas, (n.d), *REST API*, Retrieved May 7, 2015 from RIPE Atlas: https://atlas.ripe.net/docs/rest/

Shavitt, Y. & Shir, E. 2005. DIMES: Let the Internet measure itself. ACM SIGCOMM Computer

Communication Review 35, 71-74.

Shavitt, Y., & Weinsberg, U. 2009. Quantifying the importance of vantage points distribution in internet topology measurements. In INFOCOM 2009, IEEE. 792-800.

Sourceforge.net, (2014), *NFDUMP*, Retrieved May 15, 2015 from Sourceforge.net: http://nfdump.sourceforge.net/

Sourceforge.net, (2011), *NfSen - NetFlow Sensor*, Retrieved May 15, 2015 from Sourceforge.net: http://nfsen.sourceforge.net/

Sqlalchemy.org, (2015). *SQLAlchemy - The Database Toolkit for Python*, Retrieved May 18, 2015 from SQLAlchemy: http://www.sqlalchemy.org/

UbuntuNet Alliance, (2014), *What is UbuntuNet Alliance?*, Retrieved April 25, 2015 from UbuntuNet: http://www.ubuntunet.net/sites/default/files/What_Is_UbuntuNet_v1_20141030.pdf

Withall M., Phillips, I. and Parish, D., 2007. Network visualization : a review. IET Communications, 1 (3), pp. 365-372 [doi:10.1049/ietcom:20060169]