UNIVERSITY OF WAIKATO SUMMER RESEARCH SCHOLARSHIP FINAL REPORT TO RESEARCH COMMITTEE

PROJECT TITLE:	**Mapping the New Zealand Internet**
STUDENT:	**Christopher Lorier**
SUPERVISOR:	**Richard Nelson**
DEPARTMENT:	**Computer Science**
START DATE OF PROJECT:	**14/11/2011**
COMPLETION DATE OF PROJECT:	**10/02/2012**
DATE OF REPORT:	**16/03/20**

Project achievement:

This project was to produce maps of the New Zealand Internet, and software so that the maps can be reproduced with new data at regular intervals. These maps can be as educational tools for conveying the structure and nature of the Internet within New Zealand, or for investigation of changes over time, including the impact of UFB and RBI or the deployment of IPv6.

Procedure and Tools:
The graphs are produced using Gephi. Gephi is an open-source graphing programme that allows you to edit nodes, run layout algorithms, recolour nodes, and group many nodes into a single node based on attributes of the nodes.

Scamper is software to conduct Internet measurement tasks to large numbers of IPv4 and IPv6 addresses. Scamper finds traceroute data for a set of addresses created with APNIC whois data as well as alias data for all addresses in the traces.

The addresses are extracted from the traces and submitted to Team Cymru's IP to ASN mapping service for ASN data.

The addresses are then processed to eliminate foreign nodes. This was done by checking the ASN data for any hop that had a response time of greater than 3500ms. If the AS name didnt meet a list of exceptions identifying nodes as being local, it was removed.

Then the traces were read in to a programme to convert them into a GEXF file. It also reads in the alias data, and selects aliases for a given router arbitrarily from the network that most aliases of that router belong to based on the AS data of the aliases.

A GEXF file is a type of XML graph file used by Gephi. In gephi the graphs were edited to include AS information that was missing from the Team Cymru data and to remove any remaining foreign nodes. Then they were recoloured based on their ASN data. The layout was produced using the Force Atlas 2 algorithm, with dissuade hubs used briefly to clarify the centre of the graph and some editing by hand. At this point the graph was exported as the device graph.

In Gephi the nodes were then grouped according to their AS names, and exported as a GEXF file (with the option "visible only" selected). This file was edited to remove edgeweights and to rescale the node sizes, this was necessary as Gephi's sizing of the combined nodes caused the graphs to be unreadable. It was then reopened in Gephi and labels were added and the layout was modified with the Label Adjust algorithm. This was then exported as the network graph.

Software produced in this project:

sc_warts2graph
sc_warts2graph is written in c++ and takes warts files (the files produced by scamper) and produces graph files as well as useful text files.
It reads ASN data, scamper trace data and Alias data stores it into maps, and then exports it as a text file. It was written in C++ to allow it to read warts files directly, as well as use the "scamper_addr" structure to store address data, allowing it to potentially use IPv6 addresses. Though at the moment it doesn't allow that, as there is no IPv6 data.
There is a minor memory leak in this programme when the Alias data contains addresses not found in the trace data. There may also be another memory leak that I was unable to track down because I couldn't isolate it from the first leak. As the programme stores all the data until it is ready to print, after which it closes, these are not significant concerns.

It can make GEXF; DOT; and LGL edgefiles, vertex colour files, and node colour files. The options for each of these are -G for gexf, -D for dot, -L for LGL edgefile, -V for LGL vertex colour files, -E for LGL edge colour files.
It will colour things according to an ASN file, in the format you get from team-cymru when you submit a bulk ip to ASN query. The option to include an ASN file is -A. It will also output the AS number and the name if the output is a gexf file. It will assume there is an ASN file if you are outputting a vertex colour file or an edge colour file.

To submit a bulk query to Team Cymru you need a file containing all the IPs in the traces, listed one per line, with "begin" as the first line and "end" as the last. To produce this, you can use the -I option, or this is the default. You need to use GNU netcat to submit this (not nc), instructions can be found on the Team Cymru website.

The option to include alias data is -a. Alias data needs to be in a file containing two ips each line, separated by a space, where both ips are an alias of the same router. If you have ASN and alias data, it will select the alias for a given node by picking an arbitrary reannz alias if it has one, or if not, arbitrarily from any of the ASes the node has equal most aliases from. If you have alias data and not ASN data, it picks an alias entirely arbitrarily. You can print the alias data with every alias of a router in a line with the -l option.

Usually when the programme runs, it ignores all unresponsive hops in the traces. If you want to include the unresponsive hops, then use the option -u. It will ignore consecutive unresponsive hops (IE if there is a line of three unresponsive hops in between two nodes, this will be represented as one unresponsive hop), it will ignore any unresponsive hop that is at the end of a trace, and it will assume any unresponsive hop between any two given nodes is the same node.

The file inputs to the programme go in the order: ASN file, alias file, warts files.

preprocess

preprocess is written in c++ and takes warts files and the ASN file received from Team Cymru and will delete any trace that ends in a hop that takes over 3500ms to respond unless its hostname is listed as an exception. The hostname information is taken from the ASN file and if it is missing in that file it will perform a lookup. It outputs information about the looks-up it performs as well as a warts file combining all the warts files inputs into one. The inputs go in the order ASN file, warts files. The output file for the warts file is hardcoded as "foo.warts".

sizefix.py

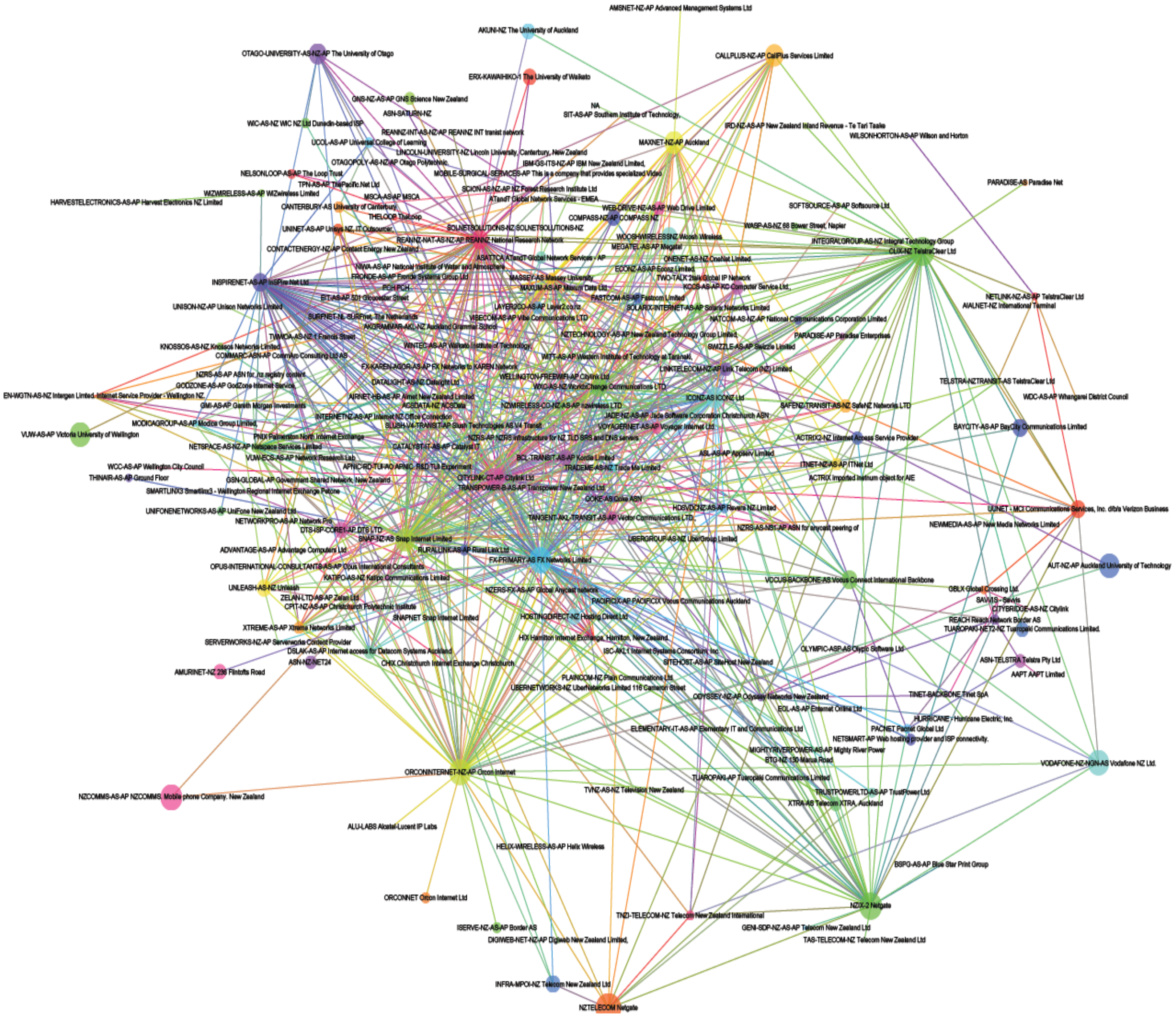
sizefix.py is written in python and edits the GEXF file outputted by gephi and resizes nodes and removes edgeweights. This is needed because the scale used by gephi when you group nodes is ridiculous, with Telecom being enormous and most nodes being so small they cannot be seen.

Future Work:
The programmes need to be updated if they are going to be used with IPv6 data.

Research results:

The first graph shows connectivity between devices connected to the Internet inside New Zealand. The colours are based on ASNs.

The second shows connectivity between networks based on ASN and doubles as the index for the first. The colours are the same as in the device graph, the positions of the nodes are based on the positions of nodes within that network on the device graph, and the sizes of the nodes are based on the number of nodes within that network on the device graph. The scale is intended to be larger than shown here, in order for the labels to be more readable.



AMSNET-NZ-AP Advanced Management Systems Ltd
AKUNI-NZ The University of Auckland
OTAGO-UNIVERSITY-AS-NZ-AP The University of Otago
CALLPLUS-NZ-AP CallPlus Services Limited
ERX-KAWAIHIKO-1 The University of Waikato
GNS-NZ-AS-AP GNS Science New Zealand
NA
ASN-SATURN-NZ
SIT-AS-AP Southern Institute of Technology,
WIC-AS-NZ WIC NZ Ltd Dunedin-based ISP
IRD-NZ-AS-AP New Zealand Inland Revenue - Te Tari Taake
REANNZ-INT-AS-NZ-AP REANNZ INT tranist network
WILSONHORTON-AS-AP Wilson and Horton
UCOL-AS-AP Universal College of Learning
MAXNET-NZ-AP Auckland
LINCOLN-UNIVERSITY-NZ Lincoln University, Canterbury, New Zealand
OTAGOPOLY-AS-NZ-AP Otago Polytechnic.
IBM-GS-ITS-NZ-AP IBM New Zealand Limited,
NELSONLOOP-AS-AP The Loop Trust
MOBILE-SURGICAL-SERVICES-AP This is a company that provides specialized Video
PARADISE-AS Paradise Net
TPN-AS-AP ThePacific.Net Ltd
SCION-AS-NZ-AP NZ Forest Research Institute Ltd
WIZWIRELESS-AS-AP WiZwireless Limited
MSCA-AS-AP MSCA
ATandT Global Network Services - EMEA
HARVESTELECTRONICS-AS-AP Harvest Electronics NZ Limited
CANTERBURY-AS University of Canterbury
SOFTSOURCE-AS-AP Softsource Ltd
WEB-DRIVE-NZ-AS-AP Web Drive Limited
THELOOP TheLoop
COMPASS-NZ-AP COMPASS NZ
WASP-AS-NZ 68 Bower Street, Napier
SOLNETSOLUTIONS-NZ SOLNETSOLUTIONS-NZ
UNINET-AS-AP Unisys NZ, IT Outsourcer,
WOOSHWIRELESSNZ Woosh Wireless
REANNZ-NAT-AS-NZ-AP REANNZ National Research Network
INTEGRALGROUP-AS-NZ Integral Technology Group
MEGATEL-AS-AP Megatel
CONTACTENERGY-NZ-AP Contact Energy New Zealand
CLIX-NZ TelstraClear Ltd
ASATTCA ATandT Global Network Services - AP
ONENET-AS-NZ OneNet Limited
NIWA-AS-AP National Institute of Water and Atmosphere
ECONZ-AS-AP Econz Limited.
FRONDE-AS-AP Fronde Systems Group Ltd
MASSEY-AS Massey University
TWO-TALK 2talk Global IP Network
INSPIRENET-AS-AP InSPire Net Ltd
PCH PCH
MAXUM-AS-AP Maxum Data Ltd
KCCS-AS-AP KC Computer Service Ltd.,
NETLINK-NZ-AS-AP TelstraClear Ltd
EIT-AS-AP 501 Gloucester Street
FASTCOM-AS-AP Fastcom Limited
AIALNET-NZ International Terminal
LAYER2CO-AS-AP Layer2.co.nz
UNISON-NZ-AP Unison Networks Limited
SOLARIX-INTERNET-AS-AP Solarix Networks Limited
SURFNET-NL SURFnet, The Netherlands
VIBECOM-AS-AP Vibe Communications LTD
NATCOM-AS-NZ-AP National Communications Corporation Limited
AKGRAMMAR-AKL-NZ Auckland Grammar School
NZTECHNOLOGY-AS-AP New Zealand Technology Group Limited,
PARADISE-AP Paradise Enterprises
TWMOA-AS-NZ 1 Francis Street
KNOSSOS-AS-NZ Knossos Networks Limited
WINTEC-AS-AP Waikato Institute of Technology,
SWIZZLE-AS-AP Swizzle Limited
COMMARC-ASN-AP CommArc Consulting Ltd AS
WITT-AS-AP Western Institute of Technology at Taranaki,
FX-KAREN-AGOR-AS-AP FX Networks to KAREN Network
LINKTELECOM-NZ-AP Link Telecom (NZ) Limited
NZRS-AS-AP ASN for .nz registry content
WELLINGTON-FREEWIFI-AP Citylink Ltd
TELSTRA-NZTRANSIT-AS TelstraClear Ltd
DATALIGHT-AS-NZ Datalight Ltd
GODZONE-AS-AP GodZone Internet Service,
WXC-AS-NZ WorldxChange Communications LTD
EN-WGTN-AS-NZ Intergen Limited. Internet Service Provider - Wellington NZ.
AIRNET-HB-AS-AP Airnet New Zealand Limited
ICONZ-AS ICONZ Ltd
WDC-AS-AP Whangarei District Council
GMI-AS-AP Gareth Morgan Investments
ACSDATA-NZ ACSData
NZWIRELESS-CO-NZ-AS-AP nzwireless LTD
SAFENZ-TRANSIT-AS-NZ SafeNZ Networks LTD
JADE-NZ-AS-AP Jade Software Corporation Christchurch ASN
INTERNETNZ-AS-AP Internet NZ Office Connection
MODICAGROUP-AS-AP Modica Group Limited,
SLUSH-V4-TRANSIT-AP Slush Technologies AS V4 Transit
VOYAGERNET-AS-AP Voyager Internet Ltd
BAYCITY-AS-AP BayCity Communications Limited
VUW-AS-AP Victoria University of Wellington
PNIX Palmerston North Internet Exchange
NZRS-AP NZRS infrastructure for NZ TLD SRS and DNS servers
ACTRIX2-NZ Internet Access Service Provider
NETSPACE-AS-NZ-AP Netspace Services Limited
CATALYST-IT-AS-AP Catalyst IT
ASL-AS-AP Appserv Limited
VUW-ECS-AS-AP Network Research Lab
BCL-TRANSIT-AS-AP Kordia Limited
ITNET-NZ-AS-AP ITNet Ltd
WCC-AS-AP Wellington City Council
APNIC-RD-TUI-AO APNIC R&D TUI Experiment
TRADEME-AS-NZ Trade Me Limited
CITYLINK-CT-AP Citylink Ltd
THINAIR-AS-AP Ground Floor
ACTRIX Imported Inetnum object for AIE
GSN-GLOBAL-AP Government Shared Network, New Zealand
TRANSPOWER-B-AS-AP Transpower New Zealand Ltd
SMARTLINX3 Smartlinx3 - Wellington Regional Internet Exchange Petone
COKE-AS Coke ASN
HDSVDCNZ-AS-AP Revera NZ Limited
UUNET - MCI Communications Services, Inc. d/b/a Verizon Business
UNIFONENETWORKS-AS-AP UniFone New Zealand Ltd
TANGENT-AKL-TRANSIT-AS-AP Vector Communications LTD.,
NETWORKPRO-AS-AP Network Pro
NEWMEDIA-AS-AP New Media Networks Limited
NZRS-AS-NS1-AP ASN for anycast peering of
DTS-ISP-CORE1-AP DTS LTD
SNAP-NZ-AS Snap Internet Limited
UBERGROUP-AS-NZ UberGroup Limited
ADVANTAGE-AS-AP Advantage Computers Ltd
RURALLINK-AS-AP Rural Link Ltd
FX-PRIMARY-AS FX Networks Limited
AUT-NZ-AP Auckland University of Technology
OPUS-INTERNATIONAL-CONSULTANTS-AS-AP Opus International Consultants
KATIPO-AS-NZ Katipo Communications Limited
VOCUS-BACKBONE-AS Vocus Connect International Backbone
UNLEASH-AS-NZ Unleash
GBLX Global Crossing Ltd.
NZERS-FX-AS-AP Global Anycast network
ZELAN-LTD-AS-AP Zelan Ltd
SAVVIS - Savvis
PACIFICIX-AP PACIFICIX Vocus Communications Auckland
CPIT-NZ-AS-AP Christchurch Polytechnic Institute
CITYBRIDGE-AS-NZ Citylink
SNAPNET Snap Internet Limited
HOSTINGDIRECT-NZ Hosting Direct Ltd
REACH Reach Network Border AS
TUAROPAKI-NET2-NZ Tuaropaki Communications Limited.
XTREME-AS-AP Xtreme Networks Limited
HIX Hamilton Internet Exchange, Hamilton, New Zealand.
SERVERWORKS-NZ-AP Serverworks Content Provider
OLYMPIC-ASP-AS Olypic Software Ltd
DSLAK-AS-AP Internet access for Datacom Systems Auckland
ISC-AKL1 Internet Systems Consortium, Inc.
ASN-TELSTRA Telstra Pty Ltd
ASN-NZ-NET24
SITEHOST-AS-AP SiteHost New Zealand
CHIX Christchurch Internet Exchange Christchurch
AMURINET-NZ 236 Flintofts Road
AAPT AAPT Limited
PLAINCOM-NZ Plain Communications Ltd.
UBERNETWORKS-NZ UberNetworks Limited 116 Cameron Street
TINET-BACKBONE Tinet SpA
ODYSSEY-NZ-AP Odyssey Networks New Zealand
EOL-AS-AP Enternet Online Ltd
HURRICANE - Hurricane Electric, Inc.
PACNET Pacnet Global Ltd
ELEMENTARY-IT-AS-AP Elementary IT and Communications Ltd
NETSMART-AP Web hosting provider and ISP connectivity.
MIGHTYRIVERPOWER-AS-AP Mighty River Power
BTG-NZ 130 Manua Road
VODAFONE-NZ-NGN-AS Vodafone NZ Ltd.
ORCONINTERNET-NZ-AP Orcon Internet
TUAROPAKI-AP Tuaropaki Communications Limited
TVNZ-AS-NZ Television New Zealand
TRUSTPOWERLTD-AS-AP TrustPower Ltd
NZCOMMS-AS-AP NZCOMMS. Mobile phone Company. New Zealand
XTRA-AS Telecom XTRA, Auckland
ALU-LABS Alcatel-Lucent IP Labs
HELIX-WIRELESS-AS-AP Helix Wireless
BSPG-AS-AP Blue Star Print Group
ORCONNET Orcon Internet Ltd
NZIX-2 Netgate
TNZI-TELECOM-NZ Telecom New Zealand International
ISERVE-NZ-AS-AP Border AS
GENI-SDP-NZ-AS-AP Telecom New Zealand Ltd
DIGIWEB-NET-NZ-AP Digiweb New Zealand Limited,
TAS-TELECOM-NZ Telecom New Zealand Ltd
INFRA-MPOI-NZ Telecom New Zealand Ltd
NZTELECOM Netgate

Signed: ……………………………………………… (Student) **Date: ……………………………..**	**Signed: ………………………………………… (Supervisor)** **Date: ……………………………..**