The GeoNet Network

From 1 pps to 10,000 hits/second

Colin Dyer (GeoNet)
Ewen McNeill (Naos Ltd)

Presented at NZNOG 2013
http://naos.co.nz/talks/geonet-network/

About GeoNet

- http://www.geonet.org.nz/
- Non-profit
- Collaboration of EQC and GNS Science
- Additional funding from LINZ
- Collects seismic data (earthquakes, volcanic activity, tsunami)
- And GPS data (surveying, etc)
- From wider New Zealand region

Network characteristics (1/3)

- Low data rates
  - By modern standards anyway (1 **M**bps is fast!)
- Near-realtime requirements
  - Data immediately useful if received in the first minute
  - After that “save for future research”
- Remoteness
  - Long drive, often 4WD, sometimes helicopter
  - Affects power, network design, operations

Network characteristics (2/3)

Sometimes what you are observing...
comes to visit you

http://www.flickr.com/photos/geonet_nz/

http://info.geonet.org.nz/pages/viewpage.action?pageId=2655654

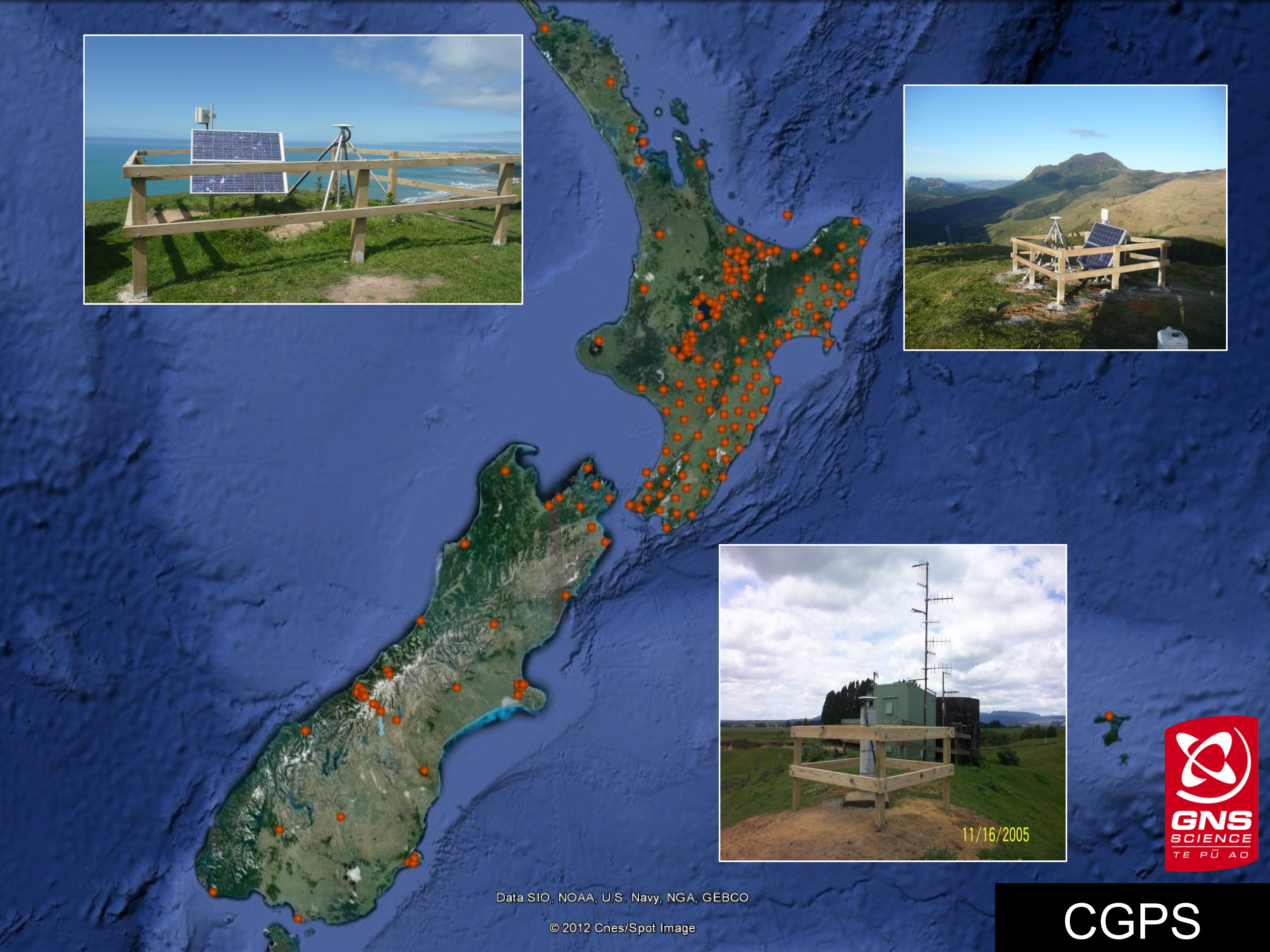
Network characteristics (3/3)

- Availability
  - **Really** want to keep collecting data through/after natural disaster

GeoNet

Seismic

CGPS


Data SIO, NOAA, U.S. Navy, NGA, GEBCO
© 2012 Cnes/Spot Image
11/16/2005
GNS SCIENCE
TE PŪ AO

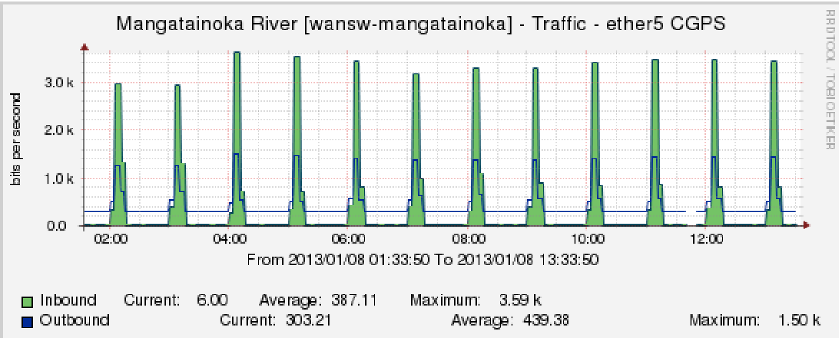
Comms

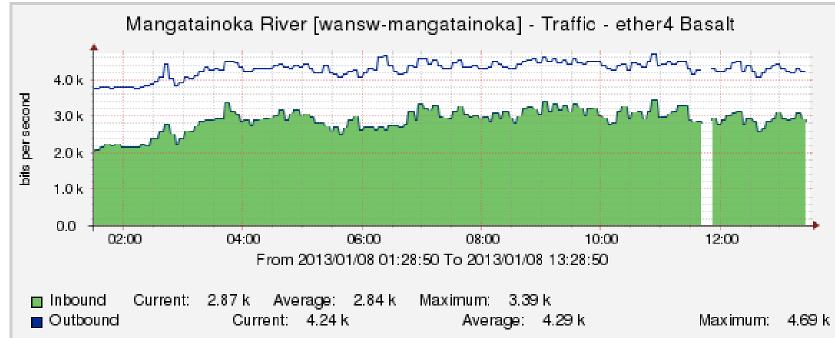
Optus VSAT
Telecom One Office DSL
Telecom Remote Office
BayCity VSAT
Gisborne.Net.NZ Limited
Vodafone Corporate Connect
© 2012 Cnes/Spot Image
Data SIO, NOAA, U.S. Navy, NGA, GEBCO

The GeoNet WAN

- Science requirements determine locations
- Network: whatever will get data back
  - GeoNet and third party radio links (licensed + WiFi)
  - 2 cellular providers
  - 2 VSAT (satellite) providers
  - DSL
  - Ethernet (over copper/fibre)
  - Still some dialup modems!

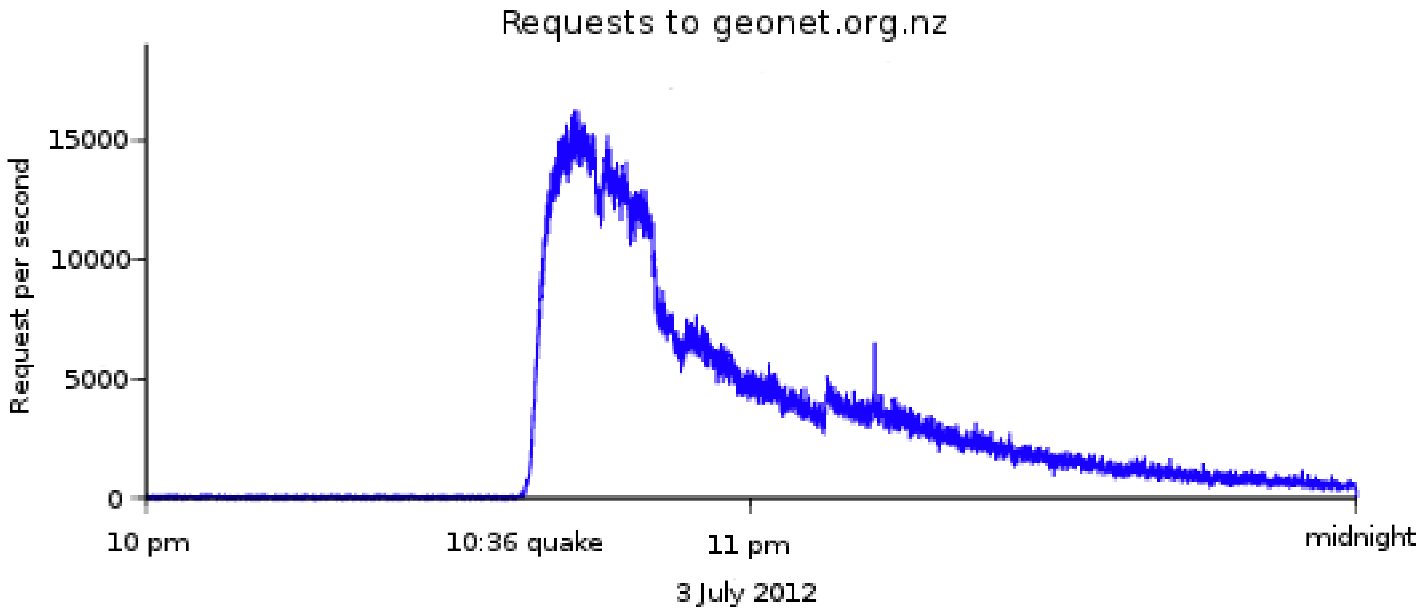
Network – instrument view





- Units are **bits** per second
  - Top of graphs are about 4 **k**bps
- CGPS – Continuous GPS station (left)
- Basalt is a seismic instrument (right)
- Regular data is compressed, sent in burst
- More data when “something happens”

Network – public view



- Units are web cluster hits per second
    - Peak is 16,257 hits/second (baseline at 0)

http://geonet-dev.blogspot.co.nz/2012/09/geonet-web-hosting-16257-requests-per.html
(Quake: http://www.geonet.org.nz/quakes/region/taranaki/2012p498491)

Data collection network

- Instrument data is small, distributed
- Multiple measurements/packet, delta compressed
- Sent to one or both of core sites
  - Fast links: to both sites; slow links: to one site
- Replicated to other site, sent for analysis
- Raw data streamable over the Internet
  - Near real-time
- Archived for long term research

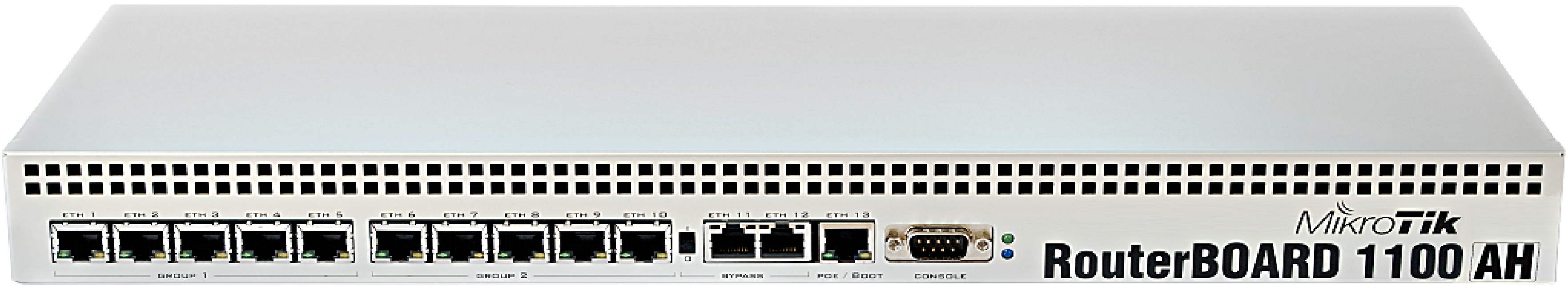
Network pre 2012

- Radio networks bridged together
  - In some areas spanning over 100km
- Other networks (static) routed by provider
- Each provider net differently numbered
- Data collection hosts multi-homed
  - Static routes on hosts for each provider network
- Minimal link redundancy outside core
  - Link failover was a manual operation
- Earthquakes located by hand, on VMS

2012 redesign

- Put routing on routers
  - Both at core sites, and regional aggregation points
  - Add dynamic routing (OSPF)
- Abstract from backhaul addresses
- Add link redundancy
  - At aggregation points
  - And remote/inaccessible sites
- Data support for automatic quake location
  - Linux, compute intensive: “in the cloud”

Core routing



- Mikrotik routers
  - Mikrotik already used in network (several radio links)
  - Inexpensive, useful feature set
  - RB1000AH has 13 GigE
- 5/9/10 interface models used nearer edge
  - Can be powered from 12V

About those GigE links...

- Mikrotik RB2011LS-IN
  - 1 SFP, 5 GigE Copper, 5 10/100 Copper
- Supports Gigabit SFPs – **only**
  - So have some 1 **G**bps multimode in-building links
  - With under 1 **M**bps of traffic on them...
  - LevelOne SFP-3001 (1000Base-SX) seem okay

~~Avoiding~~ Reducing SPoF

- Too expensive to eliminate **all** SPoF
- But availability is important
- Two core sites, different set of natural disaster risks
- Key functionality replicated at two sites
  - Backhaul, routing, server processing, storage
- Data replicated to second site ASAP
- Key data streamed to overseas partners

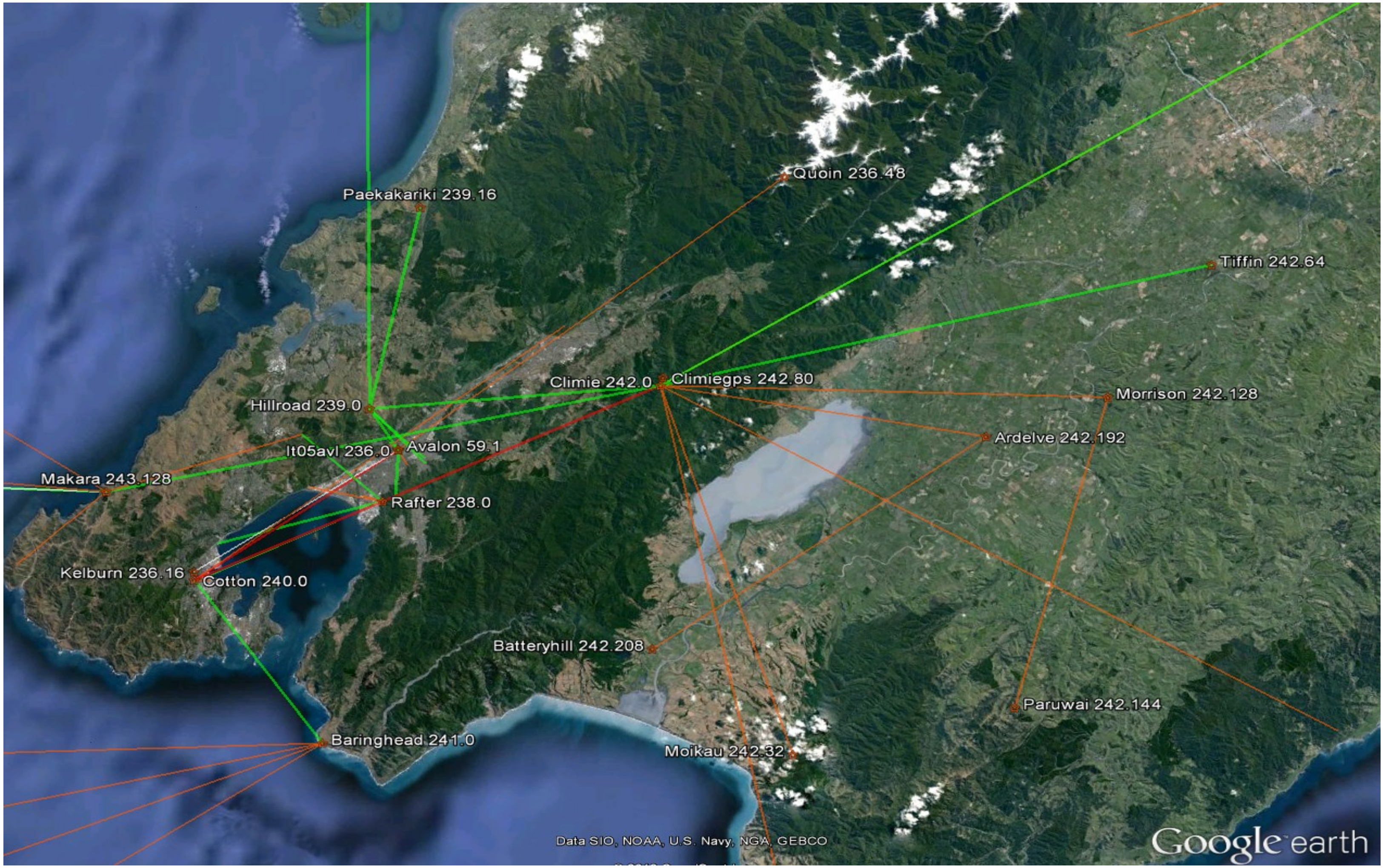
Reducing network SPoF

- Core routers are paired at each site
  - VRRP at edge, OSPF towards core network
- Most backhaul providers into two core sites
- Building out Auckland POP – more options
- In-filling with alternative backhaul paths
  - With additional (meshed) radio links where possible
  - Cellular or second provider alternative link
- Aiming to push routing out towards edge
  - Reduce number of hops in isolated spurs

Lower impact of failures

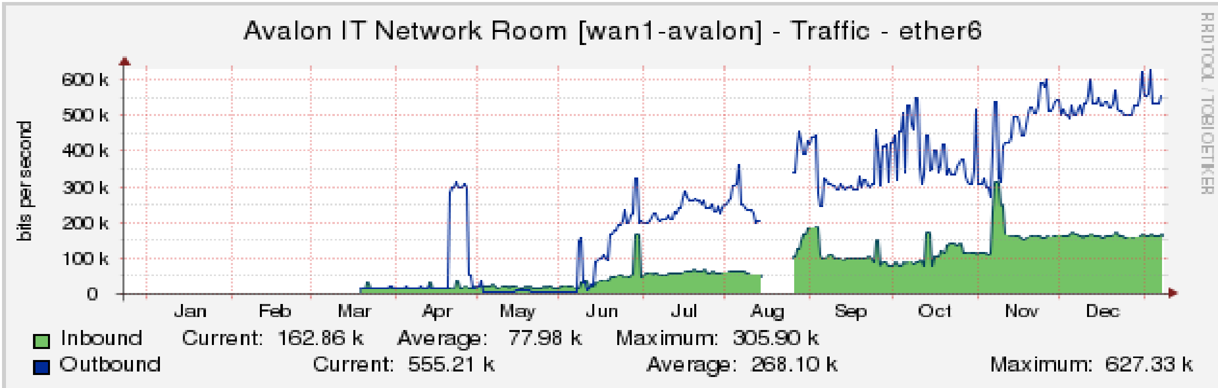
- Increase instrument density
  - Observe same event from multiple locations
  - Useful science possible with M of N locations working
- “Checkerboard” layout for backhaul
  - Try to use different backhaul for nearby sets of sites
- Many instruments can buffer data
  - Download cached data when reconnect
  - Useful for later comprehensive analysis
- Design sites with extended autonomy
  - eg, several days of battery life, storage, etc

Current data WAN state (1/2)



Wellington region radio network converted

Current data WAN state (2/2)



- Core wellington router doing ~ 600 **k**bps
- Selected other radio networks converted
- Links with one cellular provider completed
- Other 3G/DSL/radio changes underway
- Satellite sites are work in progress

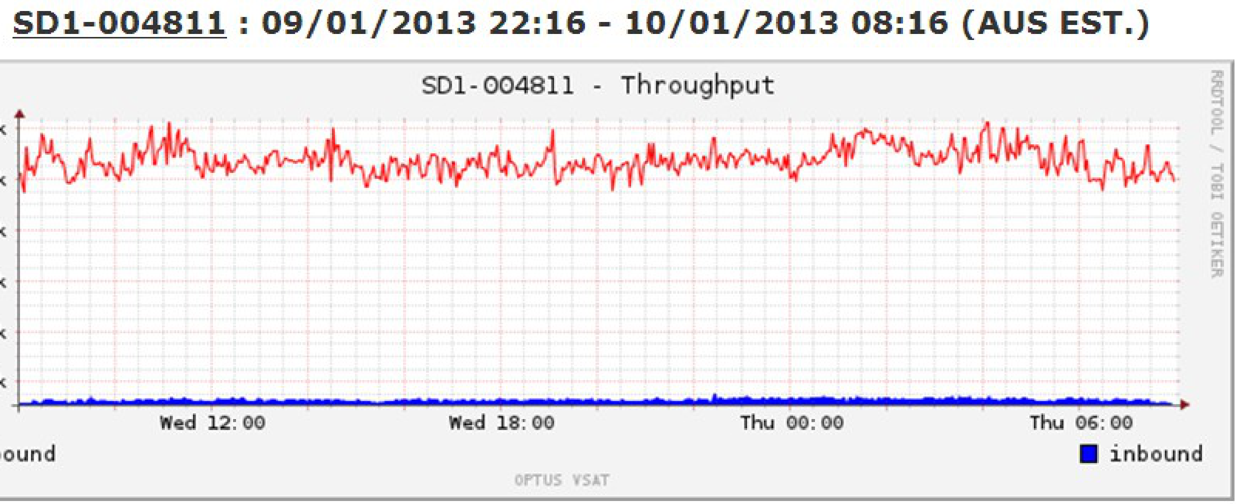
Provider abstraction – GRE

- Tunnel over provider networks
  - At least on “fast” links (not satellite!)
- Cisco-style GRE tunnel
  - Very simple, stateless
  - (Relatively) low overhead
  - Widely supported, including:
    - RouterOS 5.xx (/ip gre add …)
    - Hongdian cellular modems
- Static routing to leaf sites (mostly cellular)
- OSPF through tunnels to aggregation sites

Provider abstraction – Satellite

- Two providers, three backhaul paths
  - Terrestrial and double-hop satellite for one provider
  - Into each of two core sites
- Use policy routing to pick appropriate path
- Avoid tunnel overheads (mostly)
- Solution: NAT at both ends as abstraction
  - No on-wire packet overhead
  - But CPU overhead
  - … and sanity overhead!

Satellite – data volumes (1/2)



- Graph peak about 120 **k**bps
  - In to one of the core sites (outbound from router)
- Terrestrial backhaul:128 **k**bps Frame Relay
  - Two links, from Australia, installed 10 years ago
  - Prohibitively expensive to expand (as Frame Relay)

Satellite – data volumes (2/2)

- Load balancing on core links manual task
  - Load on some satellite sites can be issue (32 **k**bps)
  - "Catch up" transfer on reconnect can fill links
- There are also "double hop" satellite links
  - Two satellite hop has about 2s RTT
- Need to trade off "low latency" (1s RTT!)
- And "bulk data" that can handle latency

WAN IP design

- RFC1918 blocks: 10/8 and 192.168/16
  - Providers mostly using 172.16/12 for linknets
- "Region" number ("R") allocated
  - Usually one per aggregation site
- "Runnets": 192.168.R.z/28
  - Has all the science equipment
- Linknets: 10.R.x.y/28
  - Network equipment management in linknet ranges
  - Backhaul link from site "R.**z**" will be 10.R.**z**.y/28

GeoNet website

- Anycast cluster
  - APE, WIX, PNIX, CHIX, plus two hosts with PCH
  - Two IPs: www, static
- Handles spikes 0-10,000+ hits/second
- Was CMS rendered to static files + Apache
- Now Varnish cache
- Key 2012 goal: publish quake auto-locates
  - Now live: auto-updated as more data arrives
  - Final human confirmation follows later

How you can help

- Rural colocation/backhaul
  - Most data requirements well under 1Mbps
  - “Best effort” okay, especially for redundant links
- WiFi frequency coordination
- Peer at APE/WIX/PNIX/CHIX/with FX
  - Users in NZ should be accessing nearby server!
- Thoughts on GeoDNS anyone?
- Enter “Felt” reports – very helpful context
  - http://www.geonet.org.nz/quakes/felt

Questions?

- Colin Dyer <C.Dyer@gns.cri.nz>
- Ewen McNeill <ewen@naos.co.nz>

http://www.geonet.org.nz/

http://naos.co.nz/talks/geonet-network/