

#### AUSNOG 2019 CGNAT, IPv6, International POPs and JEFF



# **Recap from AusNOG 2018**



# At AusNOG 2018 we had a number of challenges to solve.

- Rapidly running out of IPv4 address space
- Predicting that we would see total aggregate network traffic hit 380G by June 2019.
- Had only 55 10G protected backhaul links operational with the rest being 1G unprotected.
- Had 10 POIs on stop sell due to capacity.
- Were planning to add a further 100,000 services in the 12 months to August 2019.
- Needed to add better transit options, Telstra Global wasn't cutting it anymore.

#### **Recap from AusNOG 2018**





<sup>-</sup> Twitch 2 Max: 856 Mbps Current: 784 Mbps - Megaport Vic VDC01 Max: 383 Mbps Current: 268 Mbps - Netflix 1 Max: 1.8 Gbps Current: 1.0 Gbps - Netflix 2 Max: 1.9 Gbps Current: 1.0 Gbps - Edgecast Max: 2.9 Gbps Current: 1.6 Gbps

## Network August 2018





INT - Interactive, VDC - Vocus DCs, YDCH YDCE - Your DC, OTW - Over the Wire, SY1 - Equinix, P1 B1 S1 - NextDC, GS - Global Switch

## Fast forward to today



# Today we are connecting more than 10,500 new nbn services a month.

Consistent and solid growth has seen us connect over 105,000 nbn services in the last 12 months. Churn is at 0.8% down from 1.2% last year.

Today we have over 158,000 active nbn services on the network and in the peak they are consuming over 330G of traffic.

We have 375G of CVC provisioned giving each user an average CVC allocation of 2.37 Mbits

12 months ago, we were provisioning 1.9 Mbits average per user – a 20% increase

#### **Telstra & Caches**





## **Domestic Peering**





#### International





## Network today (JEFF)





Cisco Powered Network using the latest NCS55A1 and ASR9K platforms.

100% fully protected core with path and carrier diversity.

International connectivity and points of presence in the USA and Singapore.

Direct peering links with all major content providers

Connected directly to all 121 nbn<sup>™</sup> points of interconnect (POIs).

## State level design





## **Intercapital Design**





## CGNAT



# For us CGNAT became the necessary evil that saved the IPv4 exhaustion day



- We selected F5 as our equipment vendor and we use the BIG-IP i15800 in VIC, NSW & QLD and the BIG-IP i11800 in SA & WA. We have two boxes in each state.
- The i15800 have 100G uplinks, the i11800 have 40G uplinks.
- Overall we have 1,320G of redundant CGNAT capacity.

# **Our CGNAT setup**



CGNAT has been one of the easiest and quickest equipment deployments we've done.

- We use the 100.64.0.0/10 range broken down into /17s or /16s per state for the inside range
- Each state has a separate VRF for the internal side, WA is VRF666, this keeps the 100.64.0.0 out of our main route table
- We leak some key internal IPs like mail and voice servers into the VRF so they are not cgnatted
- We leak IPv6 into the VRF so that this traffic doesn't have to go through the CGNAT boxes
- They are setup in an active/standby mode

## **Data flow path**





#### In a failure





# Things we've learnt



Logging can consume considerable space if you don't choose the right port allocation method.

- Initially we used DNAT (Deterministic NAT)
- We allowed ~2000 ports per customer, so a /22 of ips (1024 x (65536 1024) gives 66 million ports. Divide by a /17 (32768 ips for the internal /17) gives 2016 ports per customer. This was ok for most people, but some used more, eg bittorrent or trojans.
- For these customers using more ports, they overflowed into a separate /24 which needed to be logged separately
- This consumed around 10 GBytes a day in logging space for the overflow into the /24

# Things we've learnt



F5 suggested we move to PBA (Port Block Allocation) and our logging dropped to almost nothing.

- PBA gives blocks of 512 ports per customer to use, but if they need more then it will assign another 512 ports
- This method logs the start time a group of ports are allocated and the stop time when they are taken back again
- If a customer's connection stays up, it can be weeks between a log entry for starting a port allocation to ending a port allocation
- This reduced our logging from 10 GBytes per day to about 100 MBytes per week
- Most customers have only needed one block of 512 ports

# Log reduction





Total E		Total Entries: 2
✓ ■ Overall	555.93	0.05
Common/Sydney-PBA	7.35	0
	745.10	0.07

# **Opt-Out**



# We see 10% of customers opting out of CGNAT.

Opt-Out Reason	
Blacklisted in a game or console	40.5%
Requires port forwarding	33.8%
Runs servers	22.3%
Blacklisted on a website	2.6%
External VoIP not working	0.5%
Seeing too many captchas	0.4%

We allow customers to freely opt-out of CGNAT and business customers are not put onto CGNAT

- Game consoles that use peer to peer gaming are the primary problem that our helpdesk has and these customers are opted-out almost immediately.
- There is a "stigma" around CGNAT and more tech savvy customers will opt-out even if they don't experience any issues.
- Customers need to call our support to opt-out, we don't provide an online mechanism to reduce unnecessary opt-out.

## **Emergency services**



We all need to help educate emergency services and support services like lifeline to capture port information with the IP address.

- We receive daily requests for customer information where port number information is not captured along with the IP Address
- An education campaign is needed focused on sites like Lifeline and Beyond Blue who pass on life threatening IP address leads to police but not port information
- This is not just affecting us, but would be affecting the mobile carriers as well
- Services like Cloudflare also need to setup their systems to capture port information
- We want to help, but need the port information to be able to identify the customer at risk

# **Time is still running out**



Even with CGNAT, we will still exhaust our current IPv4 space in September 2021.

- We currently have a 200% oversubscription rate for our IP address utilisation
- With a 10% opt-out rate, and business customers getting static IP ranges we are consuming 70 IPv4 addresses per day based on our 480 customer a day sign up rate.
- Will need to purchase more address space in mid 2020 to allow time to clean up that space and try to get filters updated to reflect its now in Australia.
- IPv6 is the answer, but .....

# IPv6 - if only it just worked



#### It's 2019 and the number of poor implementations of IPv6 in consumer CPE is just staggering.

Residential CPE vendors need to get their act together and make sure their IPv6 stacks actually work. If we were to universally turn on IPv6 today, our helpdesk would light up light a Christmas tree!

- Many lockup and need to be rebooted daily to restore the connection to working
- Some receive their IPv6 lease but then don't hand it out to their internal LAN reliably
- Some are not correctly implementing the standards and randomly lose their IPv6 lease, or don't get it in the first place
- Our own CPE vendor only fixed their stack 3 months ago!

# Major sites still IPv4 only



#### News, banking, weather, still IPv4 only.

news.com.au heraldsun.com.au nine.com.au abc.net.au sbs.com.au anz.com commbank.com.au nab.com.au westpac.com.au bom.gov.au

# What we're doing



# IPv6 is now available to all our nbn customers via an opt-in system.

Over the last 12 months we've worked with Cisco to resolve the issues we were seeing in the ASR9901 BNG firmware – this was to do with being able to mark the DHCP packets to TC4 so they could be carried by nbn rather than being dropped at the POI.

- We've had our residential CPE vendor (Netcomm) fix their firmware issues.
- Customers can opt-in to IPv6
- All services we provide except voice and Fetch now support IPv6.

## **International POPs**





We never thought we'd get to build an international network, but it's been the key to delivering lower latency services to our customers.

- Even when you're using a door to door freight agent like Fedex, expect issues with customs (Singapore was easy, USA was hard get a CBP number first).
- At Coresite, don't expect PDUs to be included. We've used managed APC PDUs that can remote switch.
- Have the DC throw in a blended IP service for OOB.
- Allow a good amount of time for the install.
- If using Vocus for any links, make them test end to end (ie the wavelengths and any local DF they use)

#### International Network







# **Questions?**

Download this presentation at:

aubb.me/ausnog





aussiebroadband.com.au