RIPE 86

Rotterdam, Netherlands
22 – 26 May 2023

Rinse Kloek
May 2023

Deploying IPv6 and/or CGNAT ?

BACKGROUND OF THIS PRESENTATION

- ISP wants to be able to grow to > 1M end users. (Mainly consumer connections)"

- CFO : "This is the last /16 IPv4 we bought, please seek an alternative"

- IPv6 project + some transition technique was born

IPV6 + A TRANSITION TECHNIQUE

Transition technique	Pro's	Result	Why not choosen
MAP-T/MAP-E/Lw4o6	ISP side is stateless	❌	CPE won't support it
NAT64	No ipv4 addressing needed	❌	IPv6 rollout needs much more time…(legacy hardware)
NAT444 a.k.a CGNAT	Works over existing V4 net	✓	

- CFO: "Can we save on IPv4 costs ? => Deploy Carrier Grade Nat (CGNAT)

- Engineers: *But when do we finally do IPv6…… ??*


⚠ CAUTION

Legacy IP Only
This product does not support the current generation of Internet Protocol, IPv6.

THE BUSINESS CASE OF CGNAT

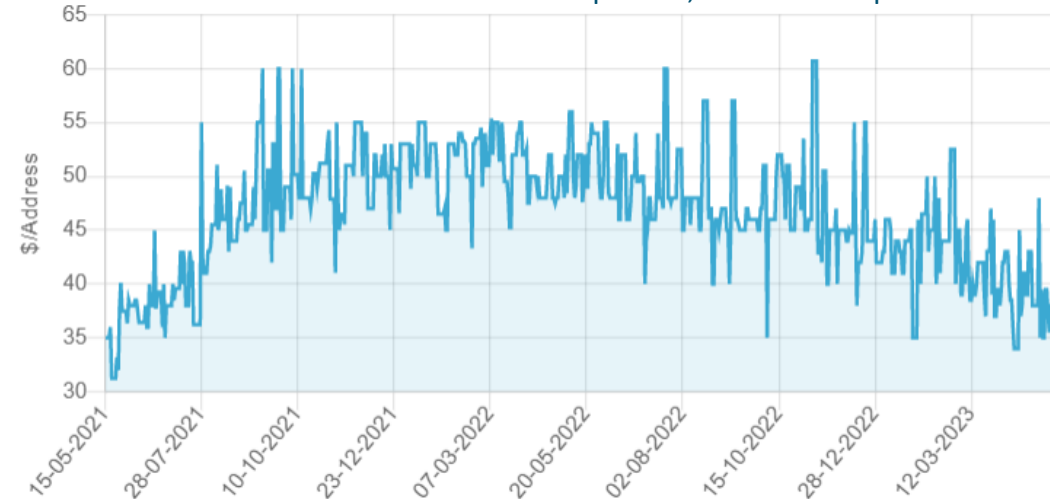■ CGNAT costs per user is significant lower than current IPv4 market price

■ With average traffic per user of ~5 mbit, CGNAT costs are 15 EUR per user.

1. Implement CGNAT and migrate users
2. Sell IPv4 blocks on the market
3. Do it ASAP as V4 prices are decreasing
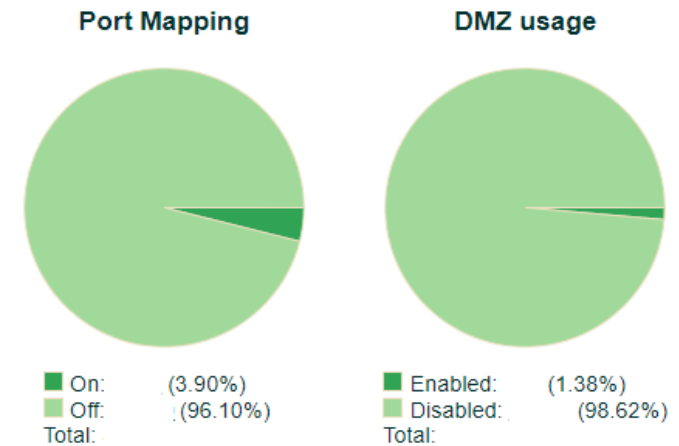
CGNAT cost per user
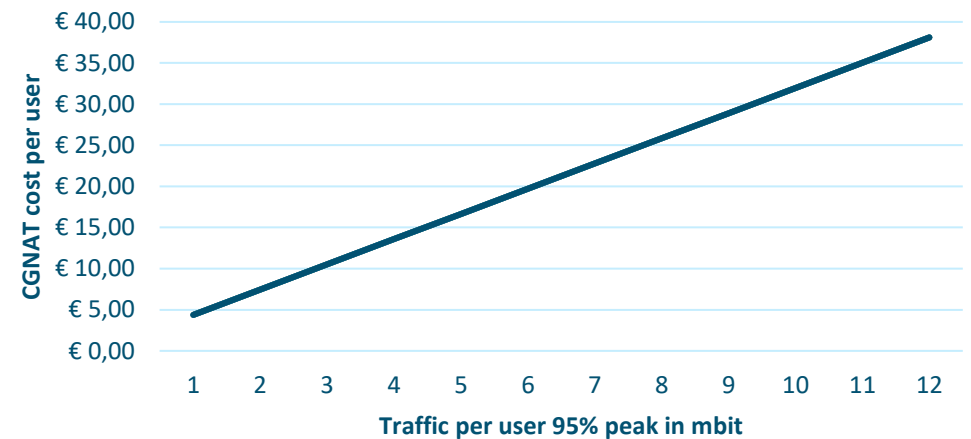
[1] Calculations based on 2N CGNAT setup RFO's, details on request

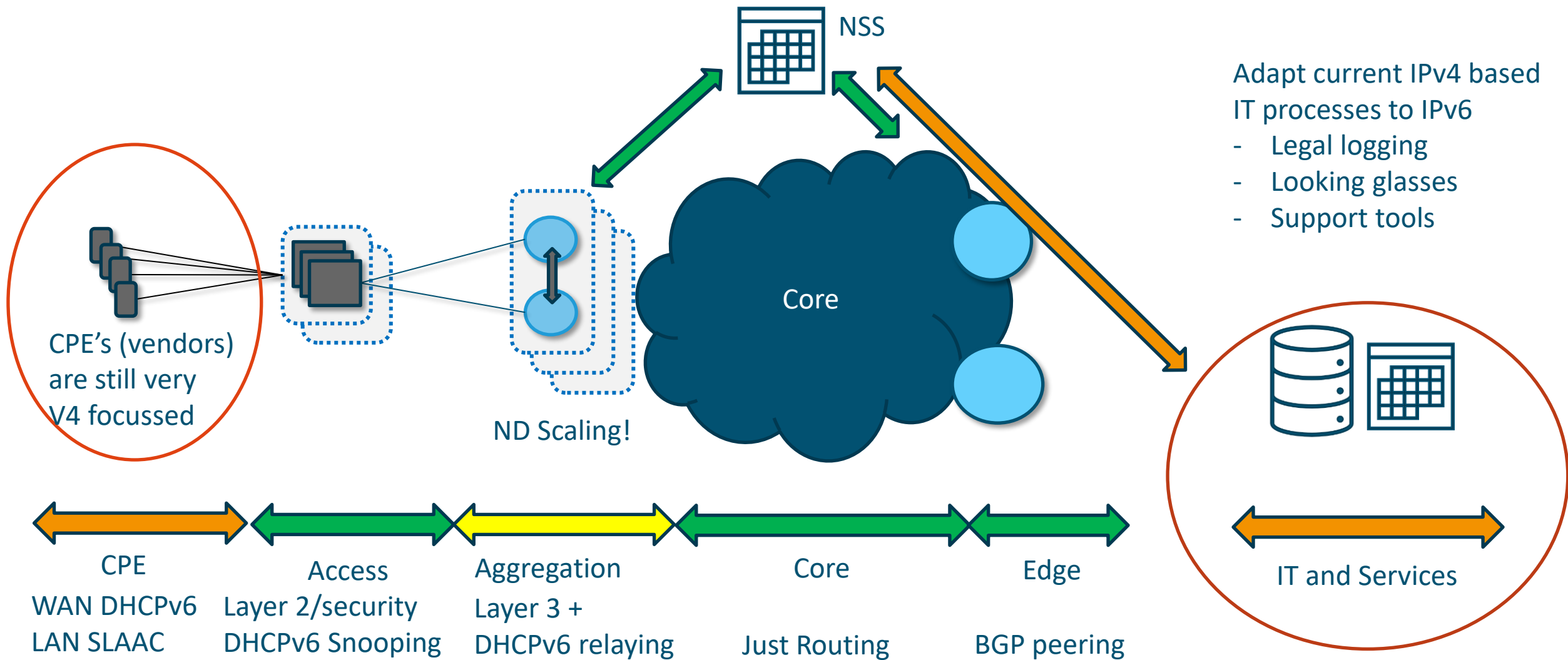IPv4 Prior sales graph from https://auctions.ipv4.global/prior-sales

THE BUSINESS CASE OF CGNAT PART 2

- ▪ Support calls can kill the case, be aware for the "5%"
  - Exclude users that do DMZ/Port forwarding (Give them public V4)
  - Easy opt-out users from CGNAT via portal/app
  - Test, test ,test….

- ▪ Lowering CGNAT traffic, improves business case
  - Bypass CGNAT for internal services (mail/dns)
  - Directly route Google/Netflix traffic from Local Caches.
    (Thanks Google and Netflix for killing the V6 business case)
  - **Implement IPv6 will decrease CGNAT traffic,**
    Finally a business case for IPv6!

Port Mapping

DMZ usage

- On: (3.90%)
- Off: (96.10%)
- Total:

- Enabled: (1.38%)
- Disabled: (98.62%)
- Total:

CGNAT stats/user	@Peak / user
UDP sessions	24
TCP sessions	29
Total Sessions	**53**
Conn rate	**0.91/s**
Traffic in Mbps	**2.0 Mbit**

IPV6 IMPLEMENTATION AND CHALLENGES

NSS

Core

Adapt current IPv4 based
IT processes to IPv6
- Legal logging
- Looking glasses
- Support tools

CPE's (vendors)
are still very
V4 focussed

ND Scaling!

IT and Services

CPE	Access	Aggregation	Core	Edge
WAN DHCPv6	Layer 2/security	Layer 3 +		
LAN SLAAC	DHCPv6 Snooping	DHCPv6 relaying	Just Routing	BGP peering

IPV6 IMPLEMENTATION

- RIPE690 (BCOP about prefix assignment) very good starting point!
  - Thank you RIPE for having this document, it made my life easier!

- DHCPv6 (IA_NA /128 on WAN link + IA_PD Prefix Delegation /56
  - Backwards compatible for clients without Prefix Delegation

- Persistant IPv6 Prefix + WAN address static based on DHCPv6 Option 37
  - ISC's KEA DHCP + Flex_id plugin makes it very easy to choose where to bind lease to.
  - Option 37 is filled with unique connection id (User can swap CPE and keep same prefix)

- Dropping DHCPv6 Release prevents prefix 'release'.
  - We don't want host reservation as it will need provisioning of users.

```
kea-dhcp.conf
  "client-classes": [
        { "name": "DROP", "test": "pkt6.msgtype==8" }
  ],
```

IPV6 IMPLEMENTATION – NETWORK CHALLANGES

- **Aggregation Router resources**

Seperate Global /128 prefix on WAN link **takes 3 HW entries extra**

- **After discussion with CPE vendor, they made it possible to do DHCPv6 Prefix without DHCPv6 WAN.**

|	IPv4	PD + NA	PD
Triple play + MGMT VLAN*	4	2	1
HW entries per address**	2	3	3
Total entries	**8**	**6**	**3**

*   Only Ipv6 in the Internet VLAN
** IPv6 uses 3 HW entires, IPv4 only 2 HW entries

DHCP IA_PD + DHCP IA_NA (2 ND entries, 1 prefix route)	DHCP_PD over link-local (1 ND entries, 1 prefix route)
#show ipv6 neighbors | i 6057.47c8.10ec 2001:4c3b:XXXX::1:13 33 6057.47c8.10ec REACH BV10500 fe80::6257:47ff:fec8:10ec - 6057.47c8.10ec REACH BV10500	#show ipv6 neighbors | i 6057.47c8.10ec fe80::6257:47ff:fec8:10ec - 6057.47c8.10ec REACH BV10500

#show route ipv6 subscriber
A   2001:4c3b:XXXX:300::/56
    [1/0] via fe80::6257:47ff:fec8:10ec, 00:01:55, BVI10500
	#show route ipv6 subscriber
	A 2001:4c3b:2000:300::/56
	    [1/0] via fe80::6257:47ff:fec8:10ec, 00:30:43, BVI10500

BUGS AND CAVEATS, STORIES BY THE CAMPFIRE

- **After upgrade of firmware the CPE started sending DHCPv6 releases after reboot**

  - That's the opposite to what RFC9096 suggests.

- **DHCPv6 relaying deamon crashed after ~30 days on Aggregation router**

- **Firewall in CPE did not firewall the ports on link-local address**

- **Samsung Smart TV did not work on CGNAT for most customers** (DNS interception feature rate limtted DNS traffic, Solution: disable DNS interception)

IT'S ALWAYS DNS

EVEN WHEN ITS NOT DNS
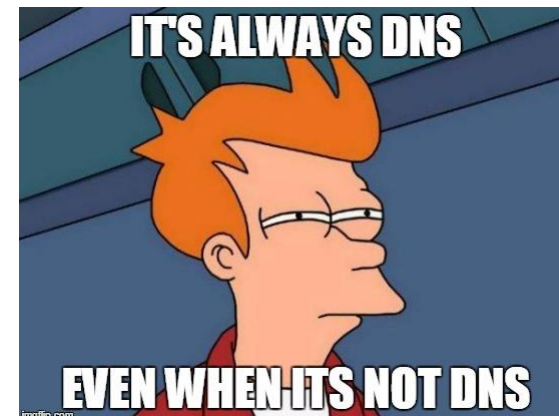
CONCLUSIONS

- There is a **(small) business case** for IPv6 if you have deployed CGNAT

- Give customers the option to easily **opt-out** from CGNAT

- In case of migration to CGNAT, exclude customers that have **Port forwarding/DMZ**
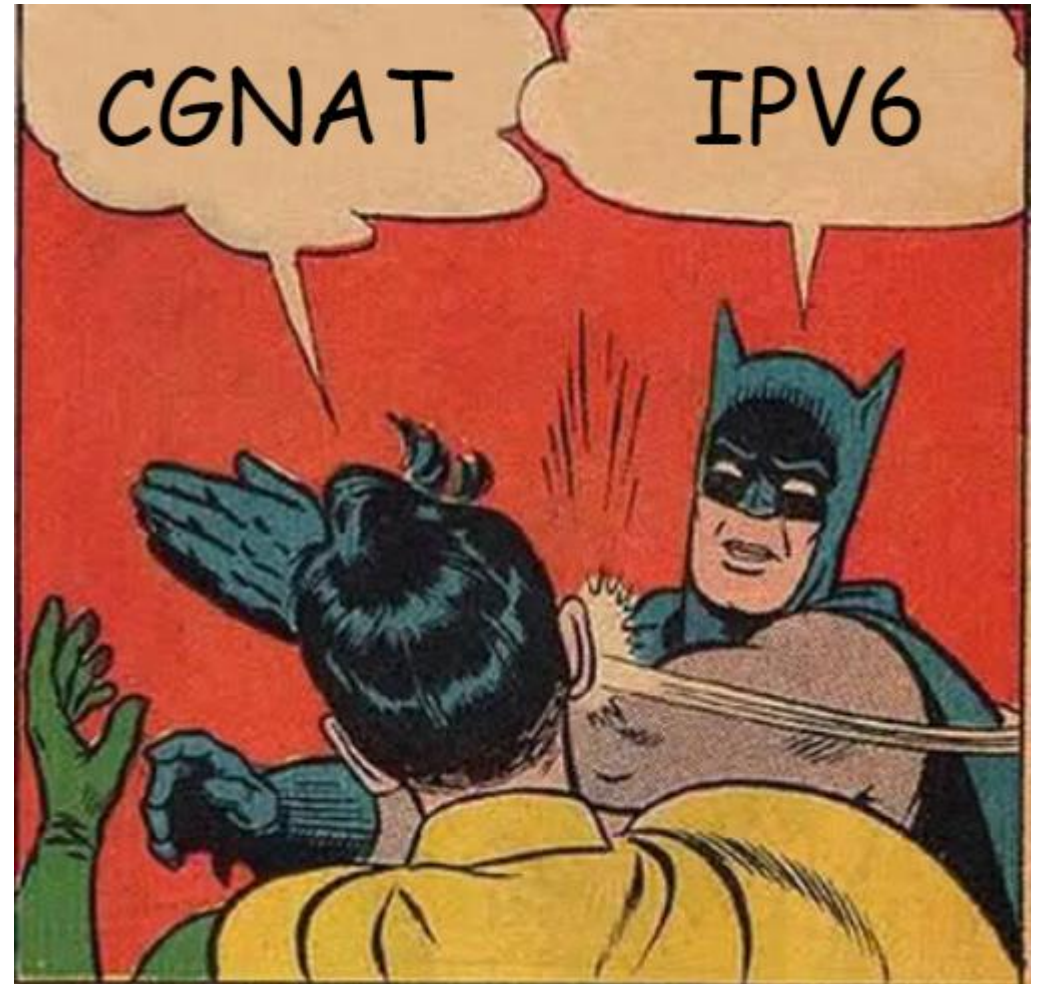
- IPv6 implementation, challenges are at the edges (CPE and IT processes)

- **Persistant prefixes** are strongly recommanded

- Be aware of the **ND/ARP table growth** when enabling IPv6

- CTO Quote:"***We don't just enable IPv6 because of the business case, but also because our customers want it***"

QUESTIONS ?

rinse@kindes.nl
rinse @ IRCNET

APPENDIX I

References used in project and presentation :

[RIPE 690] IPv6 prefix assignment for end-users, https://www.ripe.net/publications/docs/ripe-690

[GOOGLE] Google Interconnect help Carrier Grade NAT, https://support.google.com/interconnect/answer/7658745?hl=en

[NETFLIX] Netflix Open Connnect CGNAT, https://openconnect.zendesk.com/hc/en-us/articles/360036171812-Are-OCAs-compatible-with-CGNAT-Carrier-Grade-NAT-

[KEA] KEA DHCPv6 Design Considerations, https://kb.isc.org/docs/kea-dhcpv6-design-considerations

[KEA] Inifitife leases, https://gitlab.isc.org/isc-projects/kea/-/issues/897

[RFC 6598] IANA-Reserved IPv4 Prefix for Shared Address Space, https://www.rfc-editor.org/rfc/rfc6598

[RFC 7857] Updates to Network Address Translation (NAT) Behavioral Requirements, https://www.rfc-editor.org/rfc/rfc7857

[RFC 7768] Port Management to Reduce Logging in Large-Scale NATs, https://www.rfc-editor.org/rfc/rfc7768

[RFC 7599] Mapping of Address and Port using Translation (MAP-T), https://www.rfc-editor.org/rfc/rfc7599

[RFC 7597] Mapping of Address and Port using Translation (MAP-E), https://www.rfc-editor.org/rfc/rfc7597

[RFC 9096] Improving the Reaction of Customer Edge Routers to IPv6 Renumbering Events, https://www.rfc-editor.org/rfc/rfc9096

[IPv4.global] IPv4 Prior Sales, https://auctions.ipv4.global/prior-sales