Get more from v4 with CGNAT and Proxy ARP

Eric Miller
Principal Network Architect

GiGstreem
Orlando, FL

GiGSTREEM

NANOG
ATLANTA
N87
2023
13–15 FEB

GiGstreem

- Service Provider specializing in multi-family communities and commercial MDU
- Somewhere between an integrator operator and a WISP
- Collectively steward a /18 IPv4 + /32 IPv6 for our products:
  - InstantOn – managed Calix routers and BYOD
  - Ubiquitous WiFi – WiFi everywhere using Enterprise APs
  - Commercial MDU – Enterprise and SoHo ethernet delivery

Who’s this talk for?

- Operators trying to maximize the efficiency of their IPv4 allocations

- Operators curious how another provider has implemented dual-stack ipv4-CGNAT and IPv6-global effectively

Who’s this talk for?

- It’s shameful, but we’ve all been there:
  - Saving IPs by using /24 on a shared L2 segment
  - Multiple customers sharing the same VLAN (private + isolated)
  - Sitting on an IPv6 assignment for 5 years with no deployment

Where can we get more IPv4?

🟥 ARIN	🟨 Default Route Providers	🟩 Trench coat guy over there by the bar has a /22 for sale
🟩	🟨	🟨

Where can we get more IPv6?

ARIN

ARIN

ARIN

ARIN

ARIN

ARIN

Solutions for IPv4 exhaustion

Utilizing RFC1918 space for customers

- Comfortable – plentiful - easy to issue and maintain
- Not prestigious – painful - conflict with BCP38 ACLs

10.0.0.0/8 **172.16.0.0/12** **192.168.0.0/16**

Solutions for IPv4 exhaustion

Utilizing RFC6598 space for customers

- Comfortable – plentiful - easy to issue and maintain
- Prestigious – less painful – easily draft or convert existing BCP38 ACLs
- Section 5 addresses specific application concerns, promoting a special space to dictate proper behavior.
- Allows for further granularity of IPv4 /24 usage within a latency zone
- Recommended setup for some managed routers like Calix

100.64.0.0/10

Solutions for IPv4 exhaustion

Deploy IPv6!

- Less Comfortable – very plentiful – learning curve for issuance and maintenance
- Highly Prestigious – who else do you know running a dual stack network?
- Resolves CDN problems relating to:
    - Addresses labeled as VPN
    - Better Geolocation glue because of a 1-1 association with your customers

Implementing CGNAT

You’re probably already meeting these main requirements:

- Have a collection of IPv4 that you’re trying to squeeze more from
- Using a Layer 3 gateway covering a small, definable geographical area
- Have something that you call “IPAM” to keep track of what’s assigned where

You may even already have an assignment of IPv6!

Implementing CGNAT

1 – IPAM

2 – Plan your Equipment

3 – VLAN Configuration

4 – Customer Interface Configuration

5 – NAT Concepts

6 – BCP38 ACL

Implementing CGNAT

Step 1 – IPAM

- You must know how to slice up your pie
- Every operator’s network is different, so you must customize!
- Take your time
- Then, take some more
- Use the right tool to hold your data (mothball that spreadsheet!)

Implementing CGNAT *(Step 1 – IPAM)*

Sample /24 market plan:

192.0.2.0/24:

- 192.0.2.0/29 – OSPF-0 - Datacenter Core Device Loopbacks (8)
- 192.0.2.8/29 – OSPF-1 – Web, Mail, DNS Servers, Anycast Service IPs (8)
- 192.0.2.16/28 – OSPF-2 - Metro E-Line Router Loopbacks (16)
- 192.0.2.32/28 – OSPF-3 - Wireless 10Gbps Lakeside Ring (16)
- 192.0.2.48/29 – OSPF-66 - Wireless 1.5Gbps Valley Ring (8)
- 192.0.2.64/27 – OSPF-10,20,xx - Wavelength P2P Router Loopbacks (32)
- 192.0.2.128/25 – Block of blocks for static customer assignments:
  - 192.0.2.144/30 – Metro Garden Apartments (4)

Implementing CGNAT *(Step 1 – IPAM)*

Sample /22 market plan:

198.18.0.0/22:

- 198.18.0.0/26 – OSPF-0 - Datacenter Core Device Loopbacks (64)
- 198.18.0.64/26 – Block of blocks for statics in Valley OSPF-5 (64)
- 198.18.0.128/25 – Block of blocks for statics in Lakeside OSPF-3 (128)
- 198.18.1.0/27 – OSPF-1 – Web, Mail, DNS Servers, Anycast Service IPs (32)
- 198.18.1.32/28 – OSPF-3 - Wireless 10Gbps Lakeside Ring (16)
- 198.18.2.48/29 – OSPF-5 - Wireless 1.5Gbps Valley Ring (8)
- 198.18.2.128/25 – Block of blocks for statics in Wavelength OSPF-x (128)
- 198.18.3.128/25 – OSPF-10,20,xx Wavelength P2P Router Loopbacks (128)

Implementing CGNAT

Step 2 – Plan your Equipment

- You have to consider the load that NAT places on CPU and Memory and spend $$ accordingly
- You can choose between a dedicated BNG package on x86 or use the software feature sets of a traditional router
- Syslog – connection logging might be required for CALEA or other jurisdictions and will generate LOTS of data. (I’m not a lawyer!)

(If this talk is helpful, invite me back next time for a talk on what I do for syslog)

Implementing CGNAT

Step 3 – VLAN Configuration

Move to a template configuration

- Scalability amongst the junior engineers is a requirement for success
- Templates allow for the repetition needed to fully learn by understanding
- Allows for the parallel customer that’s still “working” when the other is not

Implementing CGNAT *(Step 3 – VLANs)*

Site Template for networks southbound of Gateway

1 – Reserved

2-255 – Reserved for Property Management NNI if Common Area WiFi is deployed

300 – IDF Switch Management

1001-3999 – Customer VLANs on IDF Switches

- IDF #1 – 1101-1124
- IDF #2 – 1201-1224
- …
- IDF #29 – 3901-3924

4000-4094 – Reserved

Implementing CGNAT

Step 4 – Customer interface Configuration

Interface VLAN:

- IPv4 address
- IPv6 global eui-64
- IPv6 dhcp-pd support
- BCP38 ACL

Implementing CGNAT *(Step 4 – Customer Interface examples - Juniper JUNOS)*

```
interfaces {
        et-0/0/0 {
                unit 999 {
                        description v999_Customer_1;
                        vlan-id 999;
                        family inet {
                                mtu 1500;
                                address 100.76.99.1/27;
                        }
                        family inet6 {
                                mtu 1500;
                                address 2001:db8:2800:f001::/64 {
                                        eui-64;
}}}}}
```

Implementing CGNAT *(Step 4 – Customer Interface examples - Cisco IOS XE)*

```
interface GigabitEthernet0/0/1.999
 description v999_Customer_1
 encapsulation dot1Q 999
 ip address 100.76.99.1/27
 ip nat inside
 ipv6 address 2001:db8:2800:f001::/64
 ipv6 enable
 ipv6 nd prefix 2001:db8:2800:f001::/64
 ipv6 nd other-config-flag
 ipv6 dhcp server dhcp-v6-server
!
```

Implementing CGNAT *(Step 4 – Customer Interface example - Mikrotik RouterOS)*

```
/interface vlan
add comment="v999_Customer_1" interface=sfp-sfpplus1 name=vlan999 vlan-id=999
/ip address
add address=100.76.99.1/27 interface=vlan999 network=100.76.99.0
/ipv6 address
add address= 2001:db8:2800:f001::/64 interface=vlan999 eui-64=yes advertise=yes
/ipv6 nd
set [ find default=yes ] other-configuration=yes
```

Implementing CGNAT

Step 5 – NAT Concepts

- Source NAT Customer->World behind shared IPv4
- Large sites possibly need multiple Public IPs in a pool
- UPNP creates destination NAT World->Customer
- No static Destination NATs allowed, upsell customer into static IP
    - 1-1 NAT for Customer's device w/ source NAT all hosts behind static IP
    - Use /32 route static IP to VLAN – relies on Proxy ARP (save some popcorn)

Implementing CGNAT

Step 6 – BCP38 ACL

- Don’t allow RFC1918 and other traffic where it doesn’t belong.
- Do we hairpin NAT RFC6598 or allow it to route inside a CGNAT region?

IPv4 Efficiency

What if you are not made of IPv4 addresses?

(I don’t have any to sell you, but I’m sure somebody here does)

One VLAN per customer pairs very well with Private IP addressing but not for public IP (typical /30 used means 64 customers per /24)

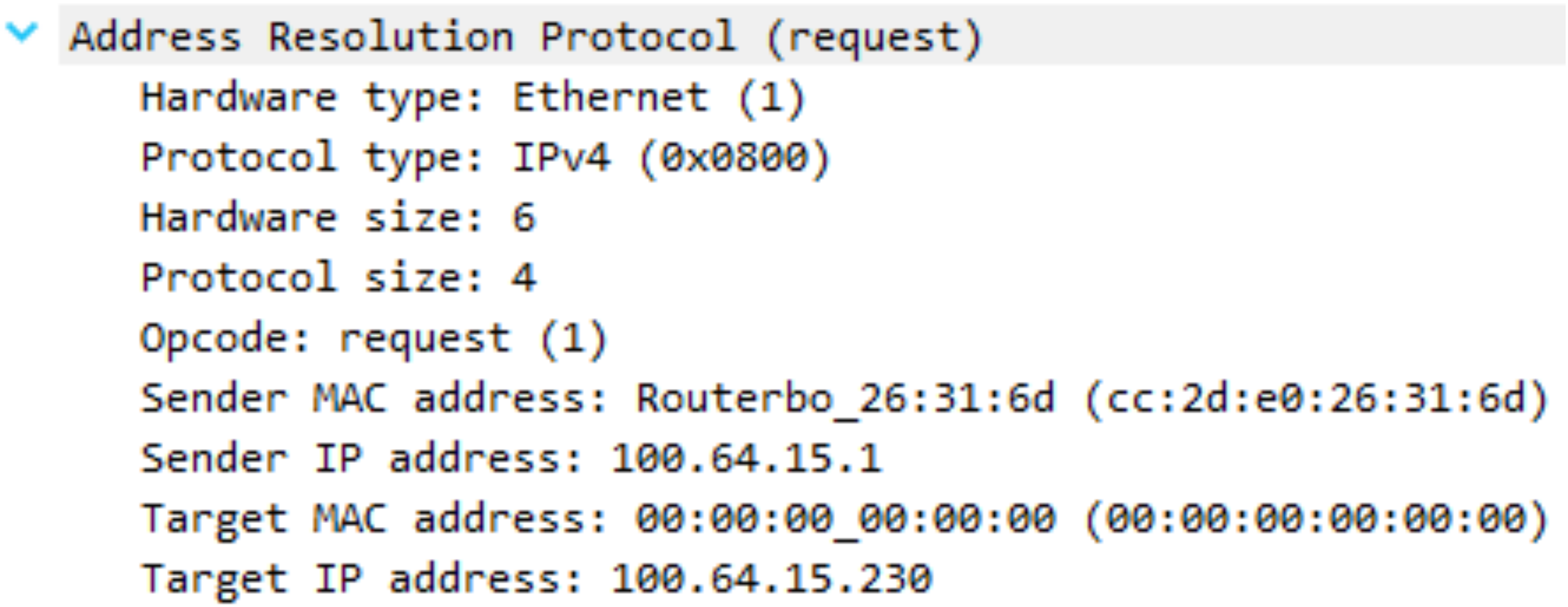
- /31, if compatible
- PPPoE is retired
- Have you tried teaching a junior engineer Private VLANs?

IPv4 Efficiency

- Proxy ARP is back!

  (Does anybody even remember how ARP works?)

- No Subnet Mask or CIDR in an ARP message



```
Address Resolution Protocol (request)
    Hardware type: Ethernet (1)
    Protocol type: IPv4 (0x0800)
    Hardware size: 6
    Protocol size: 4
    Opcode: request (1)
    Sender MAC address: Routerbo_26:31:6d (cc:2d:e0:26:31:6d)
    Sender IP address: 100.64.15.1
    Target MAC address: 00:00:00_00:00:00 (00:00:00:00:00:00)
    Target IP address: 100.64.15.230
```

IPv4 Efficiency

- Interface Address option depends on Vendor:
  - Static address (ie. 192.0.2.1/32)
  - IP unnumbered (ie. “ip unnumbered loopback0”)
- Use /32 route to interface as the downlink route
- Proxy ARP is used to simulate CPE’s “Default Gateway” for uplink traffic
- Dual Stack v4/v6 allows v6 configuration to remain the unchanged

IPv4 Efficiency

Proxy Arp for Gateway IPs allows maximum usage of any block size

/24	-	Gateway	.1	Assignable	.2 - .254
/22	-	Gateway	.1	Assignable	.0.2 - .3.254

IPv4 Efficiency *(Step 4 – Proxy ARP examples – Juniper JUNOS – v6 removed)*

```
interfaces {
  et-0/0/0 {
    unit 2 {
      description "IP Unnumbered Root";
      vlan-id 2;
      family inet {
        address 192.0.2.1/32;
    }}}
    unit 999 {
       proxy-arp unrestricted;
       vlan-id 999;
       family inet {
         mtu 1500;
         unnumbered-address et-0/0/0.2;
}}}}

routing-options {
  static {
    route 192.0.2.13/32 {
      qualified-next-hop et-0/0/0.999;
    }
  }
}
```

IPv4 Efficiency *(Step 4 – Proxy ARP examples – Cisco IOS XE – v6 removed)*

```
interface GigabitEthernet0/0/1.4
 description IP unnumbered root
 encapsulation dot1Q 4
 ip address 192.0.2.1 255.255.255.0
!
interface GigabitEthernet0/0/1.999
 description v999_Customer_1
 encapsulation dot1Q 999
 ip unnumbered GigabitEthernet0/0/1.4
!
ip route 192.0.2.66 255.255.255.255 GigabitEthernet0/0/1.99
```

IPv4 Efficiency *(Step 4 – Proxy ARP examples – Mikrotik RouterOS – v6 removed)*

```
/interface vlan
add arp=proxy-arp comment="v999_Customer_1" interface=sfp-sfpplus1 name=vlan999 vlan-id=999

/ip address
add address=192.0.2.1/32 interface=vlan999 network=192.0.2.1

/ip route
add distance=1 dst-address=192.0.2.57/32 gateway=vlan999
```

Questions