From IPv4 Scarcity to IPv6 Abundance

European MUM – 2018

Berlin / Germany

Wardner Maia

Introduction

Wardner Maia

Electronic and Telecommunications Engineer;
Internet Service Provider since 1995;
Training Business since 2002;
Certified Mikrotik Trainer since 2007;
MD Brasil IT & Telecom CTO;
Member of the board of directors of LACNIC.

MD Brasil

ISP (radio and optical)
Distributor and training center

Previous Participations on European MUMs

1) Wireless Security (2008 – Krakow/PL)
2) Wireless Security for OLPC project (2009 – Prague/CZ)
3) Layer 2 Security (2010 – Wroclaw/PL)
4) Routing Security (2011 – Budapest/HU)
5) IPv6 Security (2012 - Warsaw/PL)
6) BGP Filtering (2013 – Zagreb/CR)
7) MPLS VPNs Security (2014 – Venice/IT)
8) Network Simulation (2015 – Prague/CZ)
9) DDoS – detection and mitigation (2016 – Ljubljana/SL)
10) IoT, IPv6 and new ISP challenges for Internet Security (2017 - Milan/IT)

http://mikrotikbrasil.com.br/artigos

md brasil
TI & TELECOM

x

Scarcity x Abundance

Everything that’s needed for future survival and progress is getting scarce or running out.

http://blog.strategiccoach.com/scarcity-vs-abundance/

Scarcity x Abundance

Everything that's needed for future survival and progress is getting scarce or running out.

IPv4 World!

http://blog.strategiccoach.com/scarcity-vs-abundance/

Scarcity x Abundance

Everything that's needed for future survival and progress is getting scarce or running out.

IPv4 World!

Scarcity leave us feeling overwhelmed, depressed and paralyzed;

http://blog.strategiccoach.com/scarcity-vs-abundance/

Scarcity x Abundance

Everything important is getting bigger and better as a result of capabilities that make things faster, easier, and cheaper.

http://blog.strategiccoach.com/scarcity-vs-abundance/

Scarcity x Abundance

Everything important is getting bigger and better as a result of capabilities that make things faster, easier, and cheaper.

IPv6
World!

http://blog.strategiccoach.com/scarcity-vs-abundance/

Scarcity x Abundance

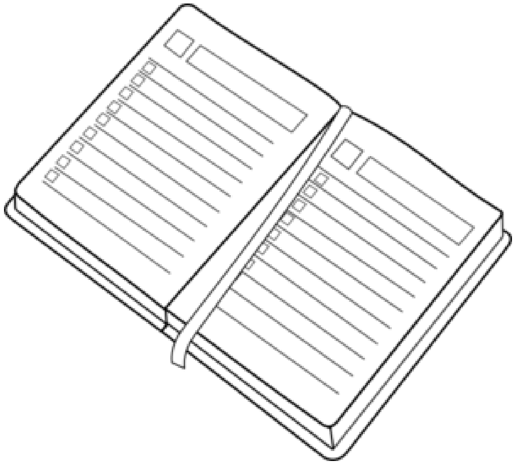
Everything important is getting bigger and better as a result of capabilities that make things faster, easier, and cheaper.

IPv6 World!

Abundance makes us feel excited motivated and ready to action;

http://blog.strategiccoach.com/scarcity-vs-abundance/

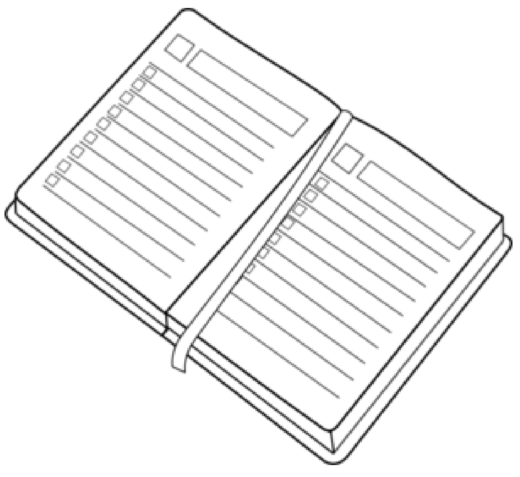
Agenda



Current status of IPv4 exhaustion and IPv6 adoption;

4’

Agenda

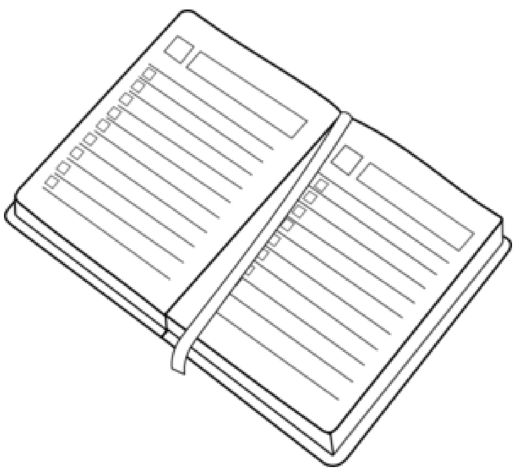


Current status of IPv4 exhaustion and IPv6 adoption;

Issues related to IPv4 scarcity and shared address solutions;

4′

Agenda



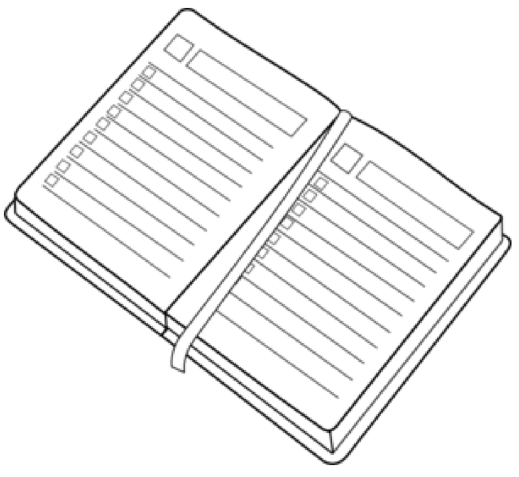
Current status of IPv4 exhaustion and IPv6 adoption;

Issues related to IPv4 scarcity and shared address solutions;

CGNAT implementation with low cost and good performance;

4′

Agenda



Current status of IPv4 exhaustion and IPv6 adoption;

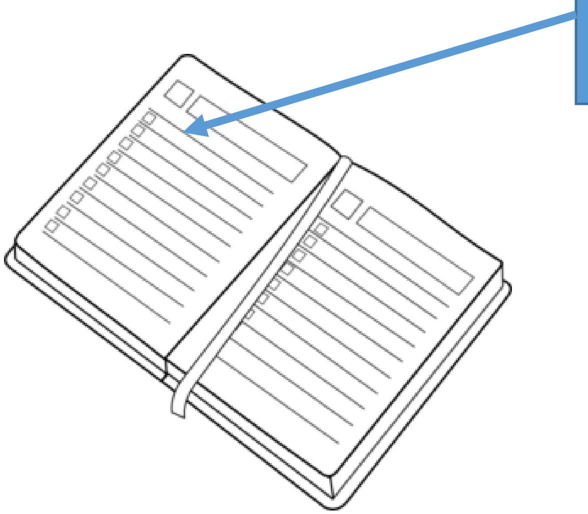
Issues related to IPv4 scarcity and shared address solutions;

CGNAT implementation with low cost and good performance;

Best practices for IPv6 deployment in an small/medium ISP access network;

4′

Agenda



Current status of IPv4 exhaustion and IPv6 adoption;

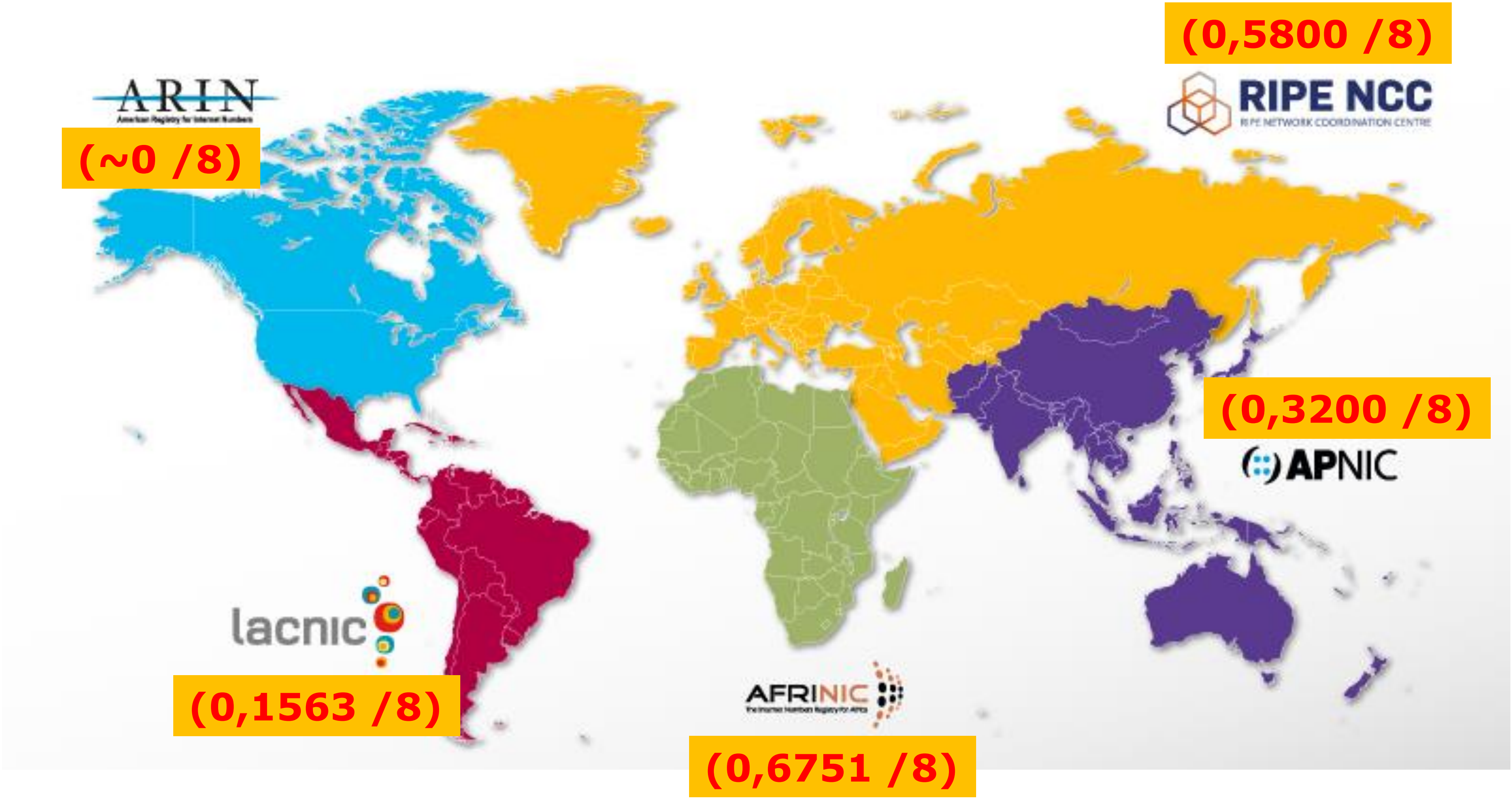
Issues related to IPv4 scarcity and shared address solutions;

CGNAT implementation with low cost and good performance;

Best practices for IPv6 deployment in an small/medium ISP access network;

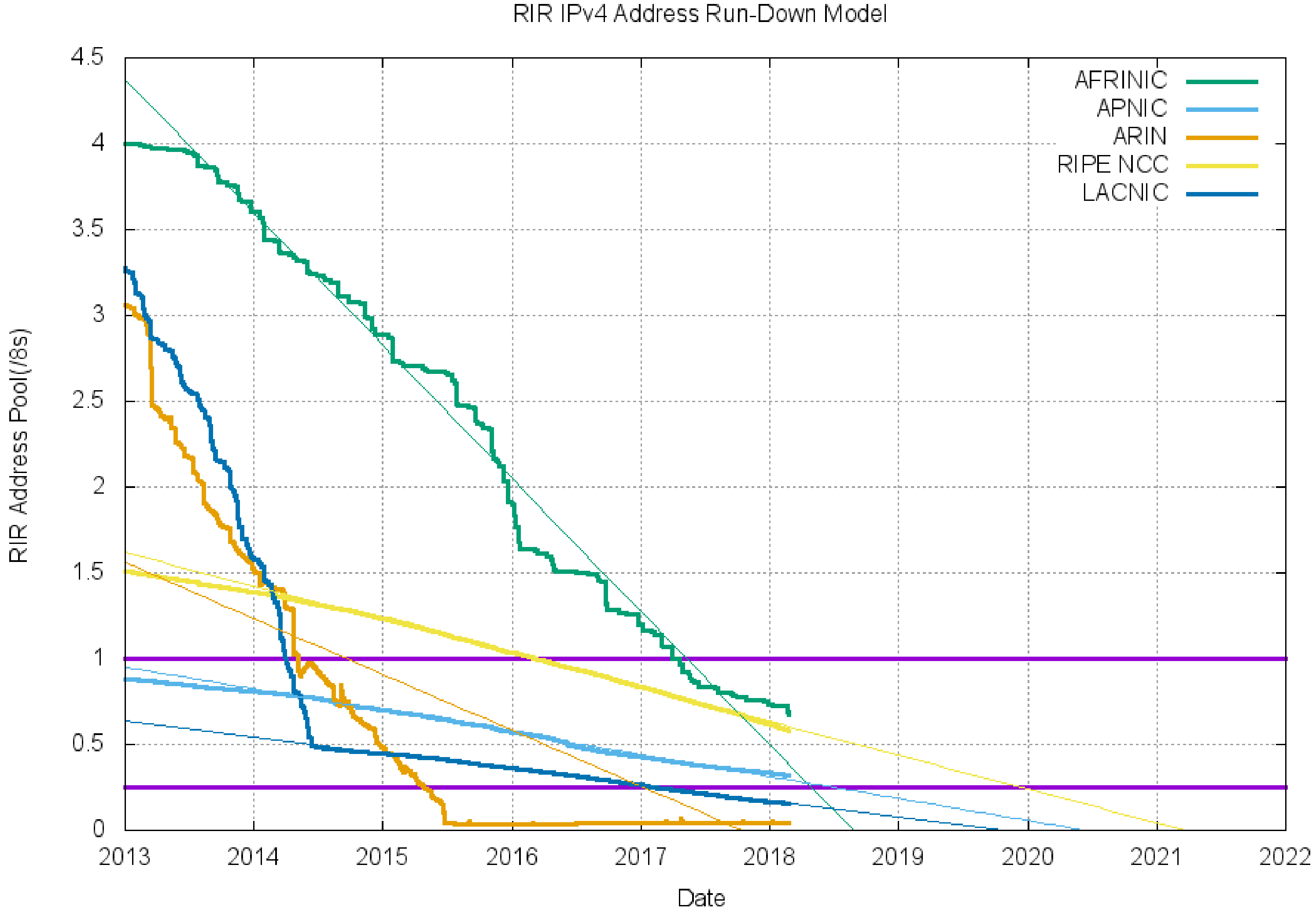
4′

IPv4 Around the World



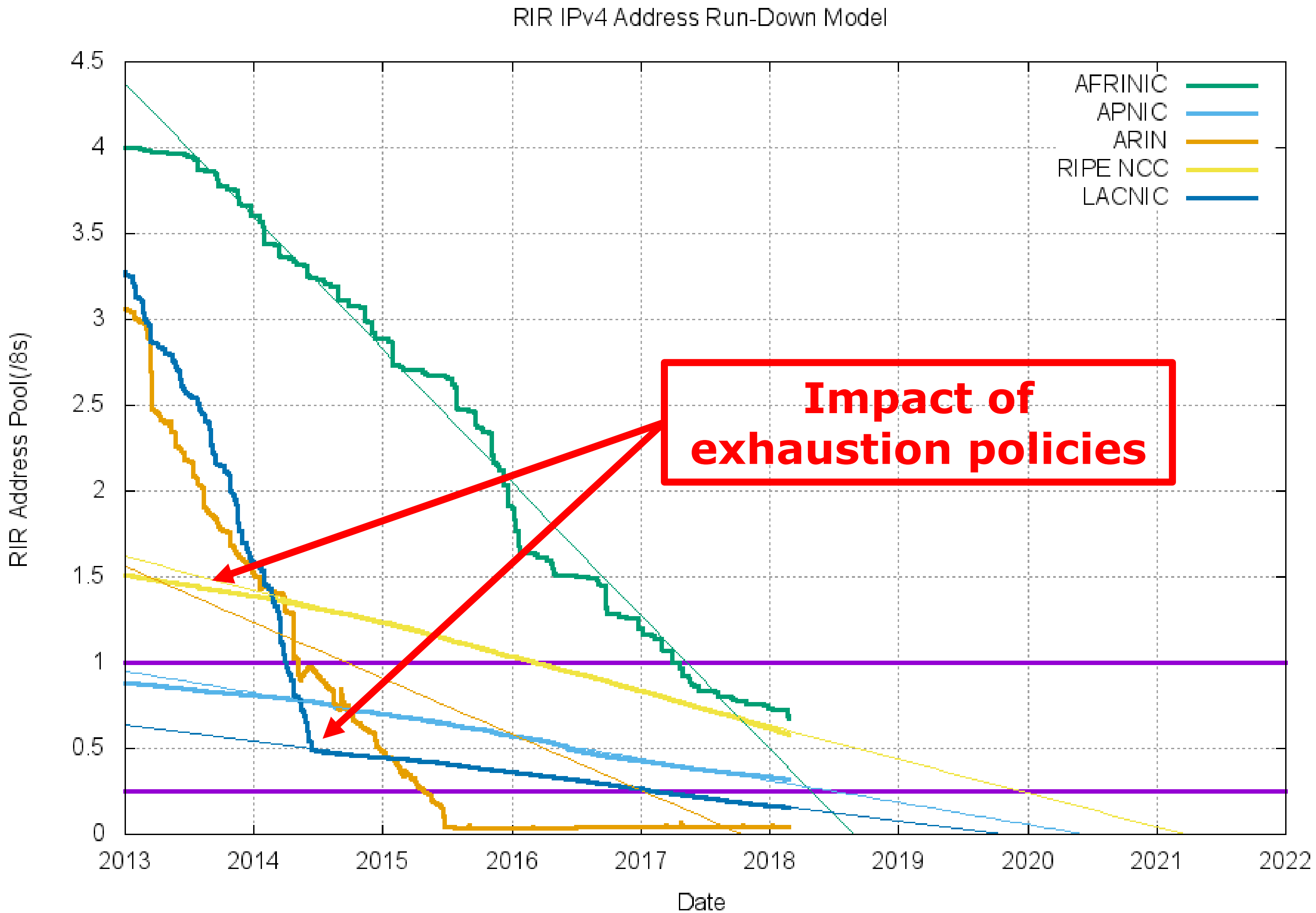
IPv4 status

https://ipv4.potaroo.net/



IPv4 status

https://ipv4.potaroo.net/



What About Legacy and Reserved Space?

Legacy and Underused Space

At the very beginning of the Internet a lot of big blocks have been assigned to institutions. Some of them never used the IP space

NETWORKWORLD FROM IDG

MICROSOFT SUBNET An independent Microsoft community View more

Home > Microsoft Subnet

THE MICROSOFT UPDATE

By Julie Bort, Network World | MAR 24, 2011 4:35 PM PT

Microsoft pays Nortel $7.5 million for IPv4 addresses

NETWORKWORLD

BUZZBLOG

By Paul McNamara, News Editor, Network World | APR 21, 2017 8:29 AM PT

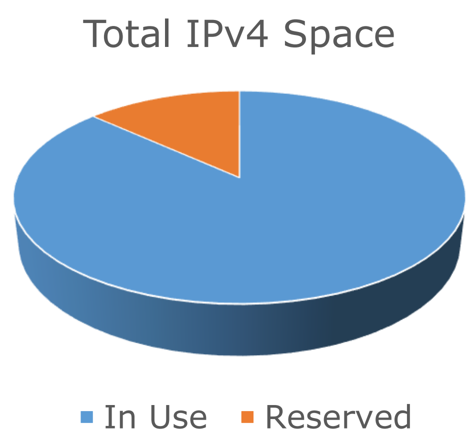
About

In addition to my editing duties, I have written Buzzblog since January, 2006. Feel free to e-mail me at buzz@nww.com.

MIT selling 8 million coveted IPv4 addresses; Amazon a buyer

Reserved IPv4 Space

Space in use = ~ 221 x /8



- 16 /8 (224.0.0/4) – Multicast
- 16 /8 (240.0.0.0/4) – Future use
- 1 /8 (0.0.0.0/8) – Local Identification
- 1 /8 (127.0.0.0/8) – Loopback
- 0.078 /8 (other small blocks) – RFC5735

Total reserved space by IETF ~ 35 /8

Recovering such space could be a solution?

The New Internet Scenario – IoT!

Gartner
WHY GARTNER ANALYSTS RESEARCH EVENTS CONSULTING ABOUT

Newsroom

Press Release

STAMFORD, Conn., November 10, 2015

Gartner Says 6.4 Billion Connected "Things" Will Be in Use in 2016, Up 30 Percent From 2015

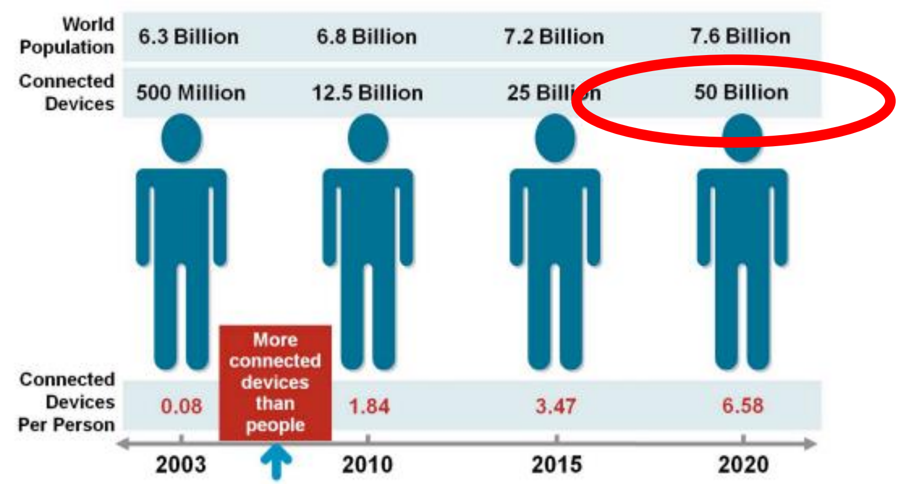
Table 1: Internet of Things Units Installed Base by Category (Millions of Units)

Category	2014	2015	2016	2020
Consumer	2,277	3,023	4,024	13,509
Business: Cross-Industry	632	815	1,092	4,408
Business: Vertical-Specific	898	1,065	[illegible]	[illegible]
Grand Total	**3,807**	**4,902**	**6,392**	**20,797**

IEEE SPECTRUM

Tech Talk | Telecom | Internet

Popular Internet of Things Forecast of 50 Billion Devices by 2020 Is Outdated



Source: Cisco IBSG, April 2011

The New Internet Scenario – IoT!

Gartner
WHY GARTNER ANALYSTS RESEARCH EVENTS CONSULTING ABOUT

Newsroom

Press Release

STAMFORD, Conn., November 10, 2015

Gartner Says 6.4 Billion Connected "Things" Will Be in Use in 2016, Up 30 Percent From 2015

Table 1: Internet of Things Units Installed P… …egory (Millions of Units)

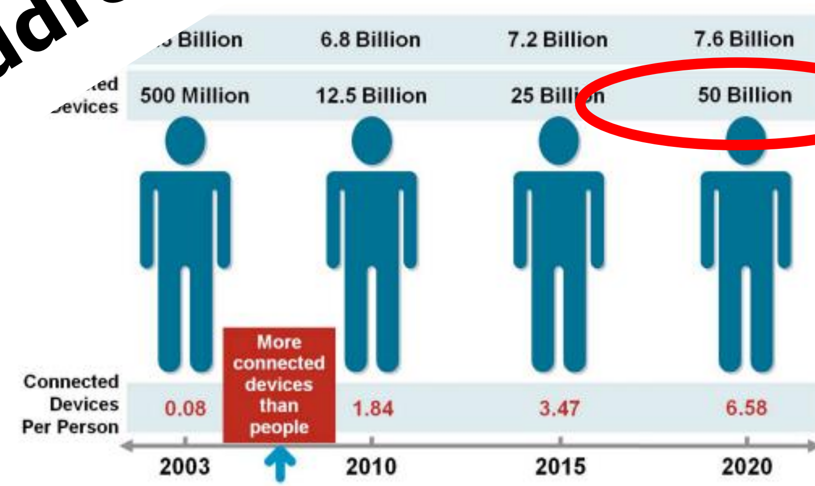
Category	2015	2016	2020
Consumer	3,023	4,024	13,509
Business: Cros…	815	1,092	4,408
Busi…	…,065	[illegible]	[illegible]
	4,902	6,392	20,797

IPv4 space is "only" 4,3 billion addresses

IEEE SPECTRUM

Tech Talk | Telecom | Internet

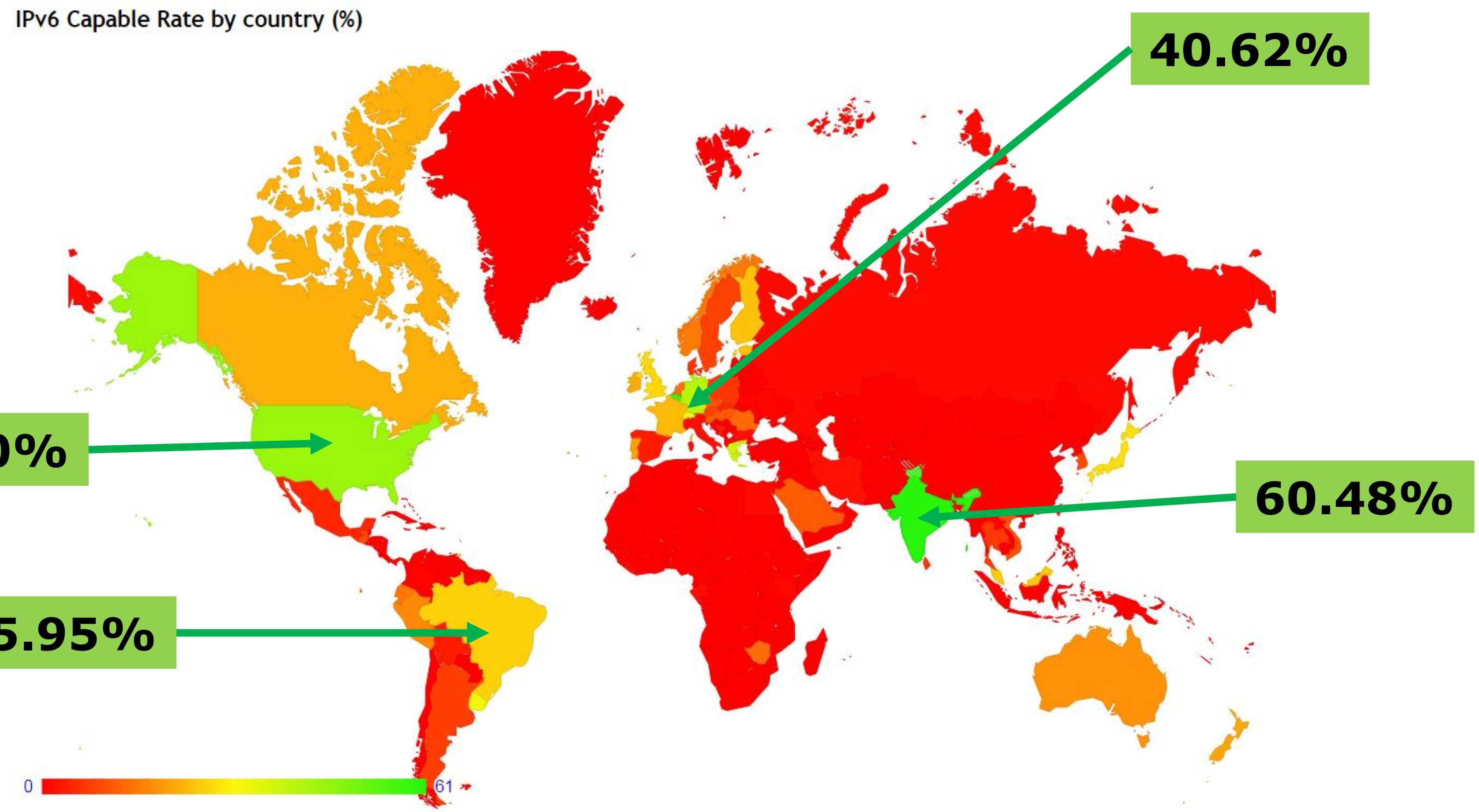
Popular Internet of Th… …recast of 50 Billion Devices by 202… …s Outdated



Source: Cisco IBSG, April 2011

IPv6 adoption status

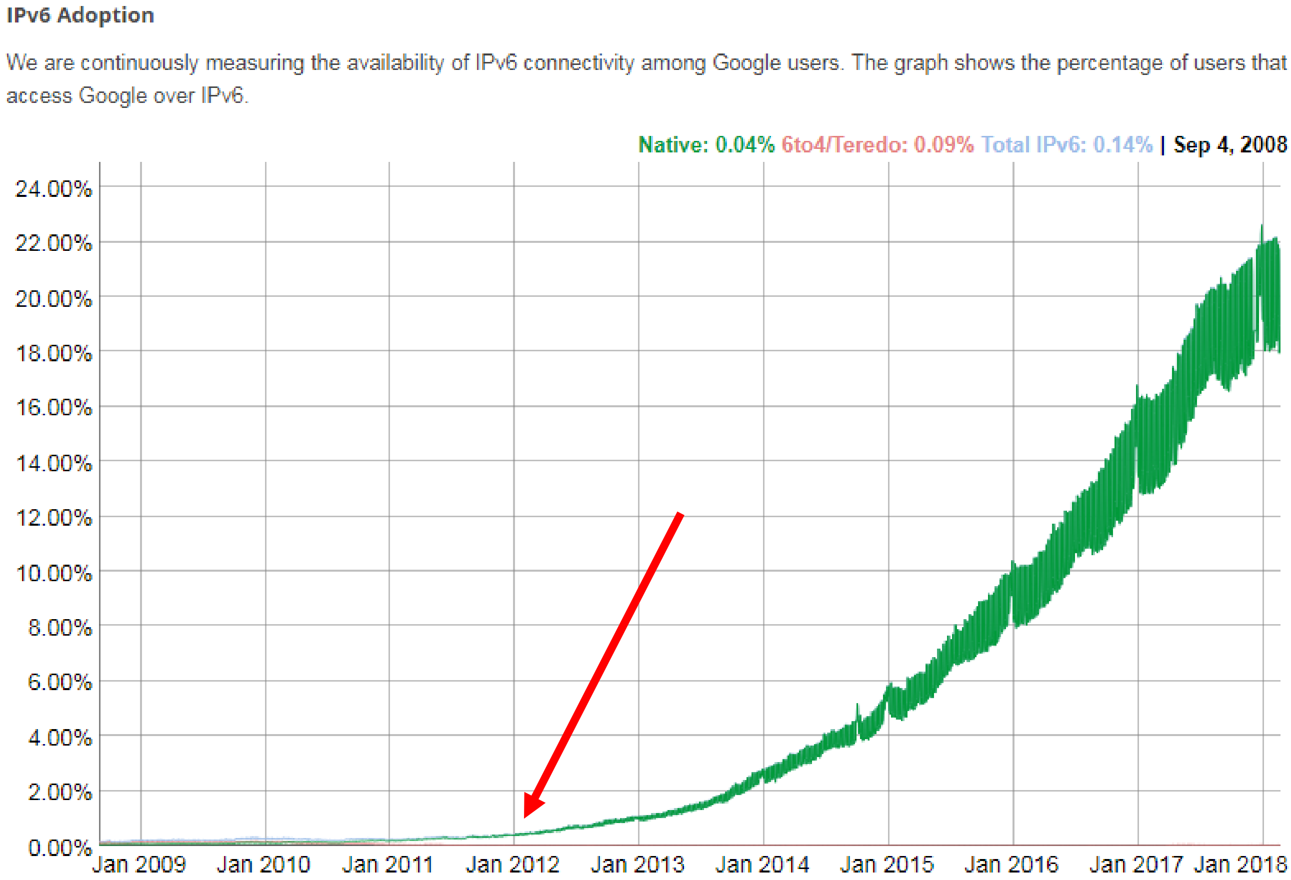
Ipv6 adoption as seen by APNIC



https://stats.labs.apnic.net/ipv6

statistics on February, 26 2018

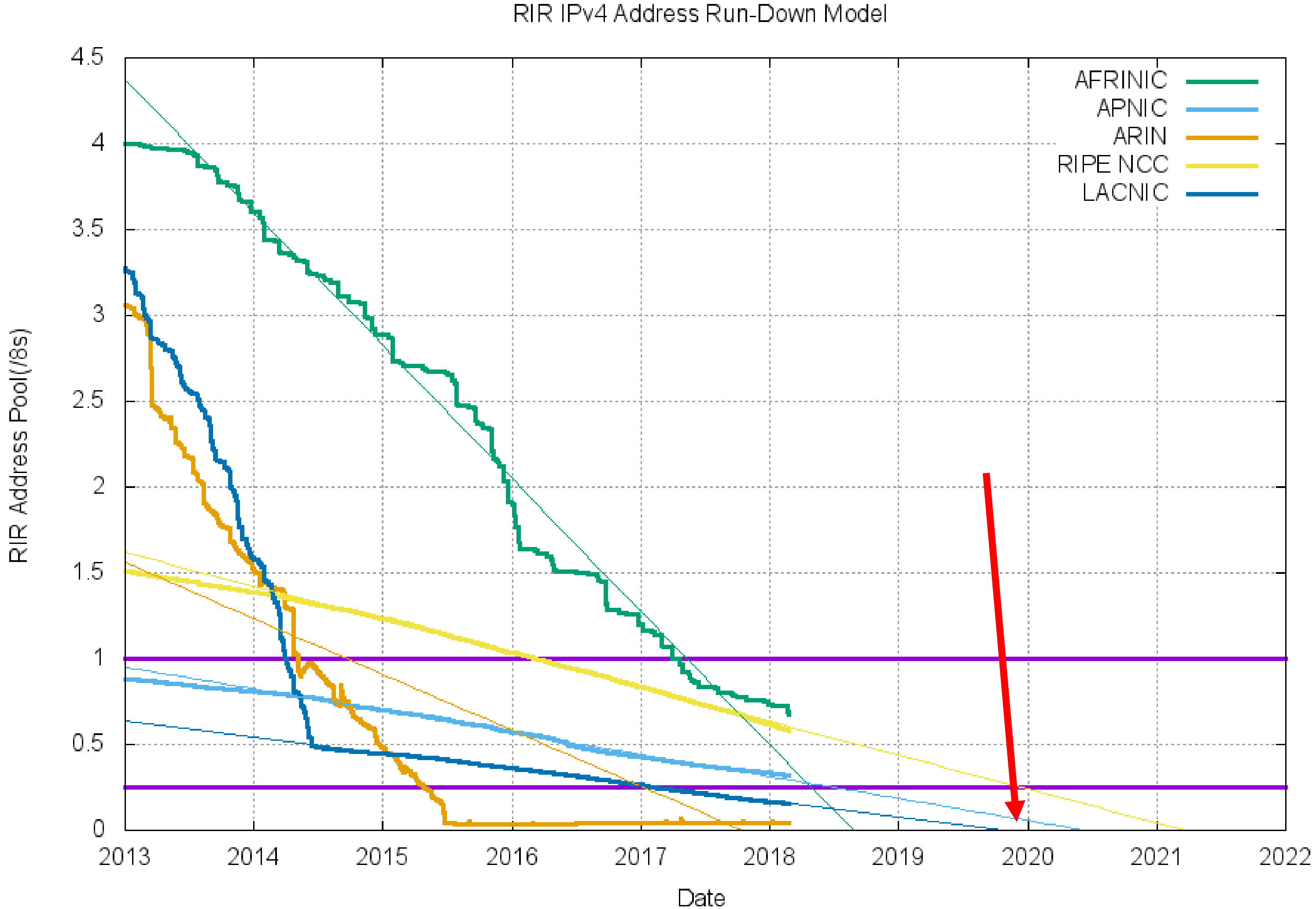
IPv6 adoption as seen by Google



https://www.google.com/intl/en/ipv6/statistics.html#tab=ipv6-adoption&tab=ipv6-adoption

IPv4 status

https://ipv4.potaroo.net/



IPv4 to IPv6 transition

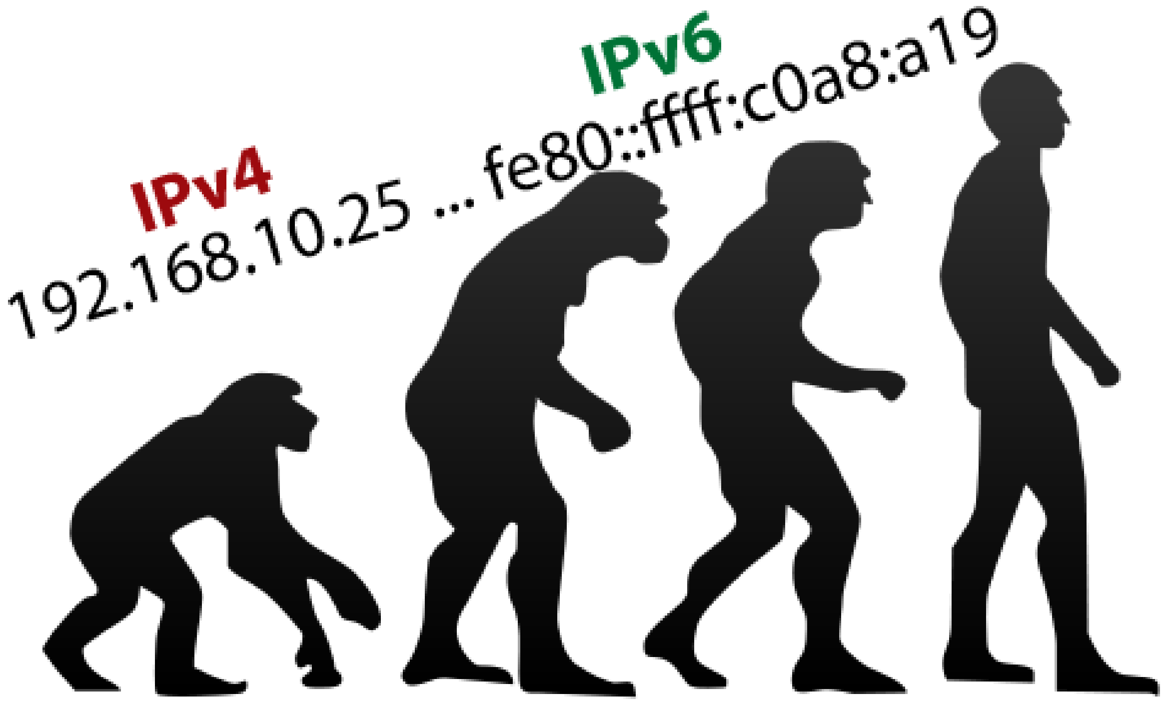
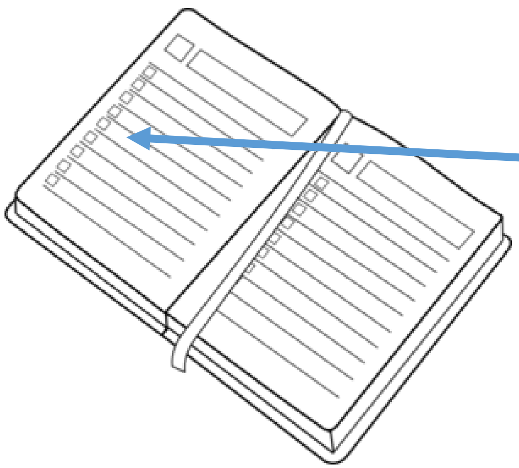


Image credit: thetelecomblog.com

Service providers and enterprises are faced with growing their networks using IPv6, while continuing to serve IPv4 customers.

Agenda ✓

Current status of IPv4 exhaustion and IPv6 adoption; ✓



Issues related to IPv4 scarcity and shared address solutions;

CGNAT implementation with low cost and good performance;

Best practices for IPv6 deployment in an small/medium ISP access network;

12’

Shared Address Solutions

→ Dual-Stack Lite [DS-Lite]

→ NAT64 [RFC6145] [RFC6146];

→ Address+Port (A+P) proposals [A+P] [PORT-RANGE]

→ Stateless Address Mapping [SAM]

→ Carrier Grade NAT (CGN) or Large Scale NAT (LSN) [LSN-REQS]

NAT, CGNAT (NAT444)

Issues and implementation

Issues with IP Address Sharing - Traceability

disruptive.asia

HOME | NEWS | VIEWS | FEATURES | OFF THE WIRE | VIDEOS | EVENTS | D

Here's why it's getting harder for law enforcement to find you via your IP address

March 10, 2017 · Don Sambandaraksa · Views · 0

Credit: VectorShots / Shutterstock.com

EUROPOL — ABOUT EUROPOL | ACTIVITIES & SERVICES | CRIME AREAS & TRENDS | PARTNERS & AGREEMENTS | CAREERS & PROCUREMENT

HOME › NEWSROOM › ARE YOU SHARING THE SAME IP ADDRESS AS A CRIMINAL? LAW ENFORCEMENT CALL FOR THE END OF CARRIER GRADE NAT (CGN) TO INCREASE ACCOUN...

ARE YOU SHARING THE SAME IP ADDRESS AS A CRIMINAL? LAW ENFORCEMENT CALL FOR THE END OF CARRIER GRADE NAT (CGN) TO INCREASE ACCOUNTABILITY ONLINE

17 October 2017

IEEE Xplore® Digital Library — Institutional Sign In

Browse | My Settings | Get Help | Subscribe

All | Enter keywords or short phrases (searches metadata only by default)

Browse Journals & Magazines > IEEE Security & Privacy > Volume: 15 Issue: 5

Availability of Required Data to Support Criminal Investigations Involving Large-Scale IP Address–Sharing Technologies

Sign In or Purchase to View Full Text | 87 Full Text Views

Technical Arguments

RFC 6302

(Logging Recommendations for Internet-Facing Servers)

Logging Recommendations

RFC 6302


[Docs] [txt|pdf] [draft-ietf-tsvwg-...] [Diff1] [Diff2] [Errata]

Internet Engineering Task Force (IETF)
Request for Comments: 6056
BCP: 156
Category: Best Current Practice
ISSN: 2070-1721

BEST CURRENT PRACTICE
Errata Exist
M. Larsen
Tieto
F. Gont
UTN/FRH
January 2011

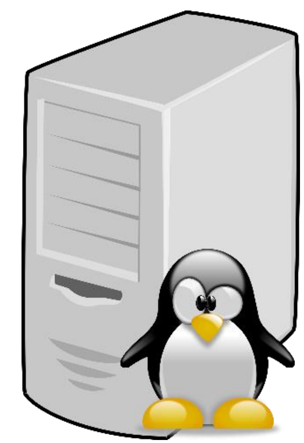
Recommendations for Transport-Protocol Port Randomization

Abstract

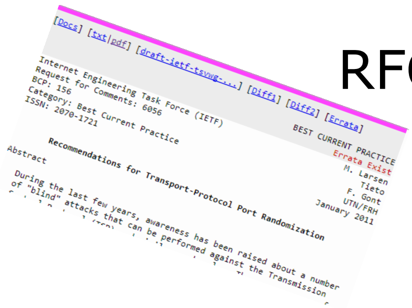
During the last few years, awareness has been raised about a number of "blind" attacks that can be performed against the Transmission

Server Considerations

*"In the wake of IPv4 exhaustion and deployment of IP address sharing techniques, this document recommends that Internet-facing servers **log port number** and accurate timestamps in addition to the incoming IP address."*



RFC 6302


[Docs] [txt|pdf] [draft-ietf-tsvwg-...] [Diff1] [Diff2] [Errata]

BEST CURRENT PRACTICE
Errata Exist

Internet Engineering Task Force (IETF) M. Larsen
Request for Comments: 6056 Tieto
BCP: 156 F. Gont
Category: Best Current Practice UTN/FRH
ISSN: 2070-1721 January 2011

Recommendations for Transport-Protocol Port Randomization

Abstract

During the last few years, awareness has been raised about a number of "blind" attacks that can be performed against the Transmission

RFC 6302

ISP Considerations

"ISP deploying IP address sharing techniques should also deploy a corresponding logging architecture to maintain records of the relation between a customer's identity and ***IP/port resources*** *utilized."*

How Big would be ISP logs for all customers?

Logging in an ISP

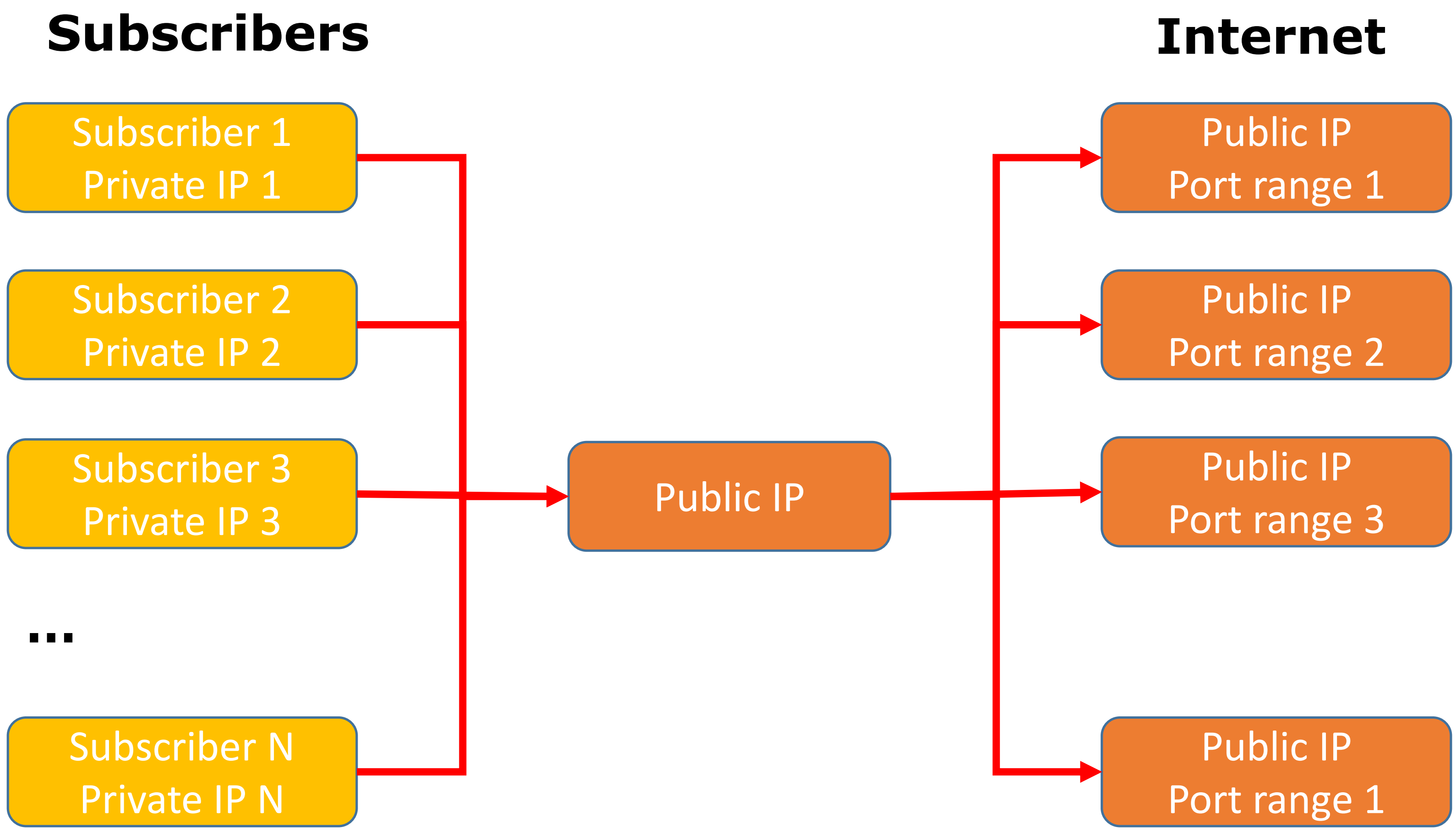
To log user’s port is painful and space consuming. The strategy is to divide available ports among customers in a fixed relation.

RFC 6269 (Issues with IP Address Sharing)

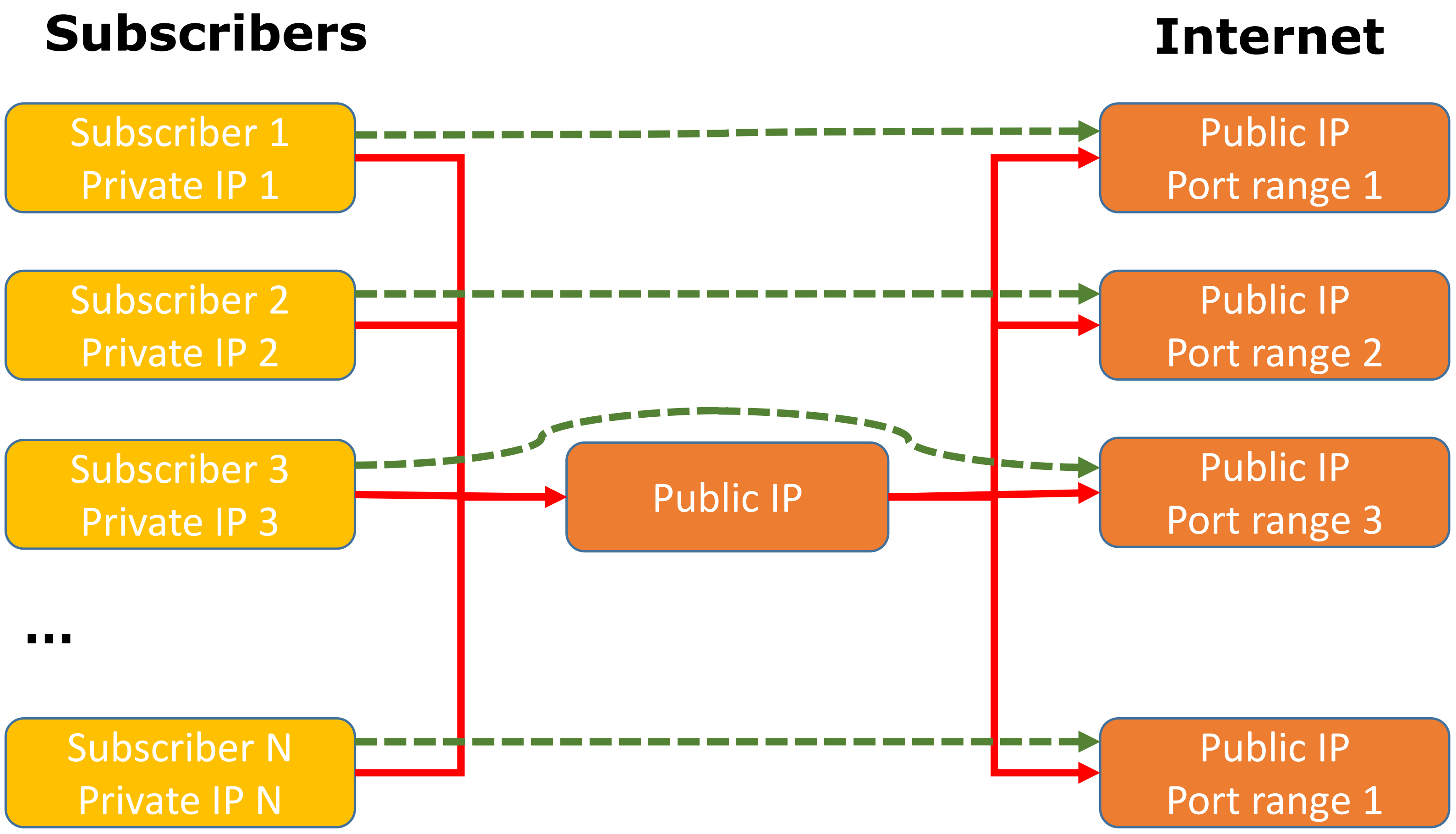
“Address sharing solutions may mitigate these issues to some extent by ***pre-allocating groups of ports.*** *Then only the allocation of the group needs to be recorded, and not the creation of every session binding within that group.”*

This way logs are not necessary, but only a table with pre-allocated group or ports per user.

ISP Strategy



ISP Strategy



Which Ports?

Which Ports?

IANA Ports

- **Well-Known Ports:** from 0 through 1023;
- **Registered Ports:** from 1024 through 49151;
- **Dynamic and/or Private Ports:** from 49152 through 65535.

md brasil TI&TELECOM

Which Ports?

Can we use the **“registered ports”**?

Although the term **“registered ports”** could lead to an understanding that such ports have some kind of restriction, RFC 4787 makes it clear that the use of port space (1024-65535) is safe:

"mapping a source port to a source port that is already registered ***is unlikely to have any bad effects****".*

(RFC 4787)

In total, we have 65535 – 1024 = **64511** ports for CGNAT

How many ports per user?

How many ports per user?

It is not a good idea to reserve a small amount of ports per user.

There are some implications related to the number of available port per user.

→ TCP time wait

→ Port randomization

Number of Ports
TCP time wait

After a TCP connection has been concluded it enters in TIME-WAIT state (typically 4 minutes);

The purpose is avoid duplicate connections in case of overlap of new and old TCP sequence numbers;

TIME-WAIT delay gives enough time to connections to die before reopening them.

→ This implies in a **bigger reservation** of available ports for users than the **real number** needed for connections

Number of ports
Security Considerations

There are several types of “Blind” attacks against TCP and similar Protocols

Possible Consequences: throughput reduction, broken connections and/or data corruption

Attacks rely on attackers ability to know (or guess) the five-tuple (Protocol, Source Address, Destination Address, Source Port, Destination Port)

How many Ports per User?

Depending on the type of connection we can have different needs. E.g. for mobile phones clients, few ports could be enough. For fixed fiber broadband, a bigger number should be allocated;

The bigger number of ports we can reserve for a single user, less probability of future problems.

The question to be answered is:

How many times our IPv4 space should be multiplied to attend our future necessities?

IPv4 Planning in Times of Scarce Resources...

Hypothetical Scenario

Current situation:

- Small ISP starting a business in a region with 200K inhabitants;
- Potential market for fixed broadband customers – 50K houses;

Growth forecast for the next 3 years

- 50% of market share in fixed bandwidth (25K customers)

Currently ISP is near 1K customers and only 1 IPv4 /22 block (1024 IPs) ☹

Hypothetical Scenario

Sharing ratio:

To attend this scenario, ISP will have to do CGNAT and the "sharing ratio" would be 1:25

Number of ports per customer:

Considering 64511 ports, the number of ports will be:

- 64511 / 25 ~= 2580 ports per user

subscriber 1: 1.024 - 3.604

subscriber 2: 3.605 - 6.185

…

subscriber 25k: 65.285 – 65.535

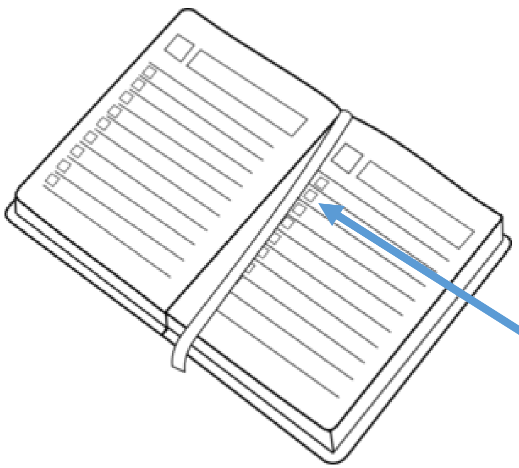
How to deploy?

How to deploy CGNAT

You can buy an expensive “dedicated” box:

Or you can follow the RFCs, and use RouterOS saving money for implementing IPv6 ☺

Agenda



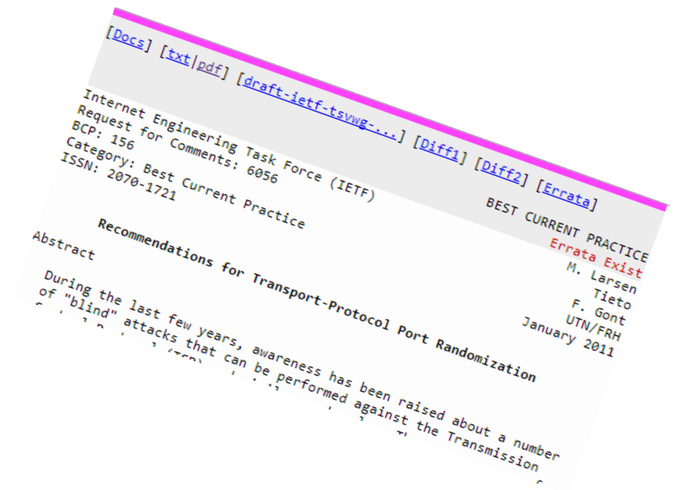
- Current status of IPv4 exhaustion and IPv6 adoption;
- Issues related to IPv4 scarcity and shared address solutions;
- **CGNAT implementation with low cost and good performance;**
- Best practices for IPv6 deployment in an small/medium ISP access network;

21’

How to deploy CGNAT

RFC 6598

The reserved space for CGNAT or NAT444, according to RFC 6598 is 100.64.0.0/10


[Docs] [txt|pdf] [draft-ietf-tsvwg-...] [Diff1] [Diff2] [Errata]

Internet Engineering Task Force (IETF)
Request for Comments: 6056
BCP: 156
Category: Best Current Practice
ISSN: 2070-1721

BEST CURRENT PRACTICE
Errata Exist
M. Larsen
Tieto
F. Gont
UTN/FRH
January 2011

Recommendations for Transport-Protocol Port Randomization

Abstract

During the last few years, awareness has been raised about a number of "blind" attacks that can be performed against the Transmission

First approach, NAT for each IP address:

https://wiki.mikrotik.com/wiki/Manual:IP/Firewall/NAT

*In fact the documentation address block is 198.51.100.0/24

CGNAT script

```
:global sqrt do={
  :for i from=0 to=$1 do={
    :if (i * i > $1) do={ :return ($i - 1) }
  }
}

:global addNatRules do={
  /ip firewall nat add chain=srcnat action=jump jump-target=xxx \
    src-address="$($srcStart)-$($srcStart + $count - 1)"

  :local x [$sqrt $count]
  :local y $x
  :if ($x * $x = $count) do={ :set y ($x + 1) }
  :for i from=0 to=$x do={
    /ip firewall nat add chain=xxx action=jump jump-target="xxx-$($i)" \
     src-address="$($srcStart + ($x * $i))-$($srcStart + ($x * ($i + 1) - 1))"
  }
```

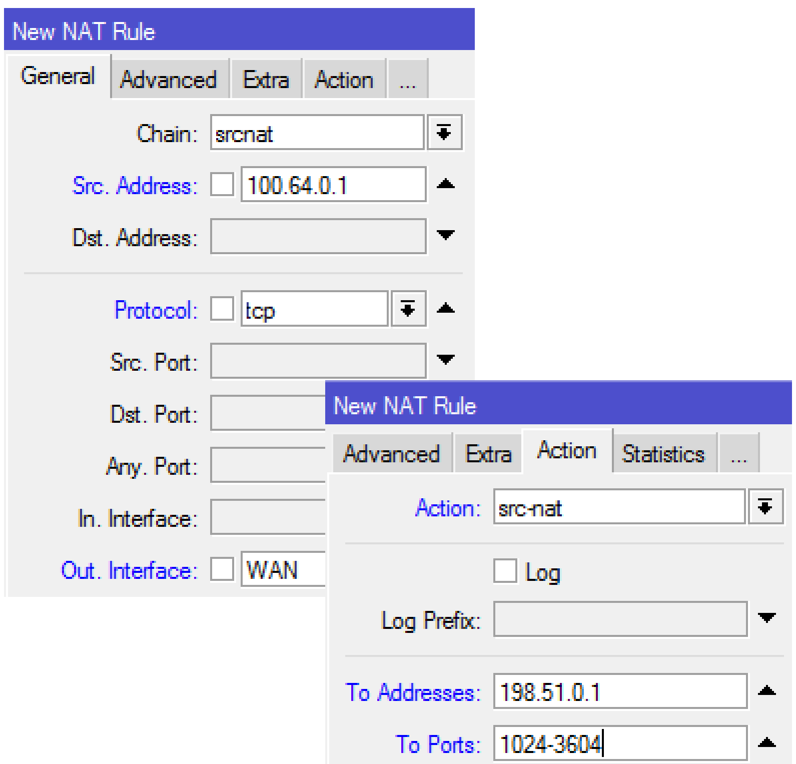
CGNAT script

```
:for i from=0 to=($count - 1) do={
  :local prange "$($portStart + ($i * $portsPerAddr))-$($portStart + (($i + 1) * $portsPerAddr) - 1)"
  /ip firewall nat add chain="xxx-$($i / $x)" action=src-nat protocol=tcp src-address=($srcStart + $i) \
   to-address=$toAddr to-ports=$prange
  /ip firewall nat add chain="xxx-$($i / $x)" action=src-nat protocol=udp src-address=($srcStart + $i) \
   to-address=$toAddr to-ports=$prange
 }
}
```

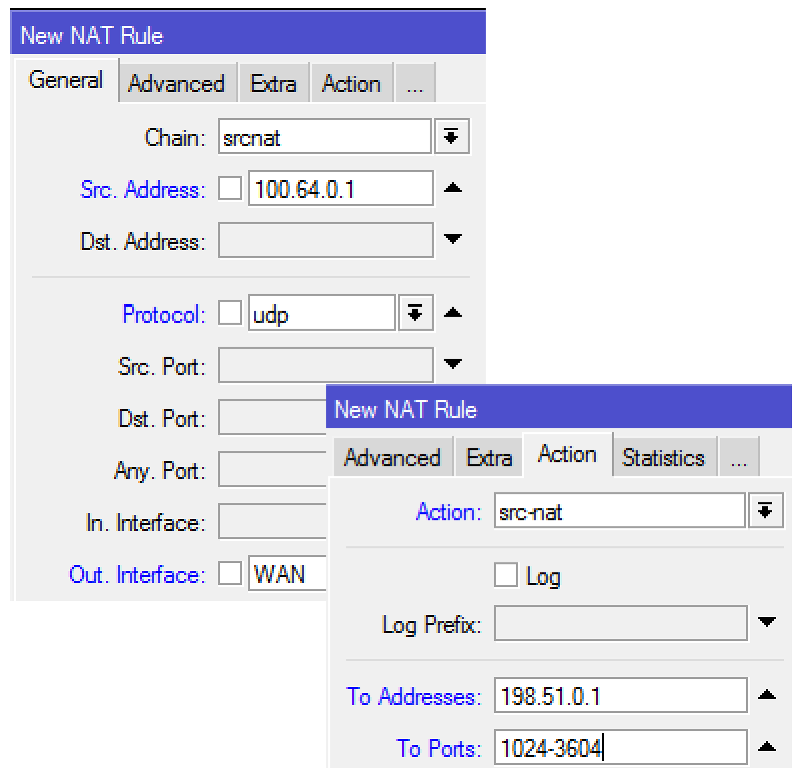
CGNAT

Basically we'll create 3 NAT rules per IP address:

TCP Protocol

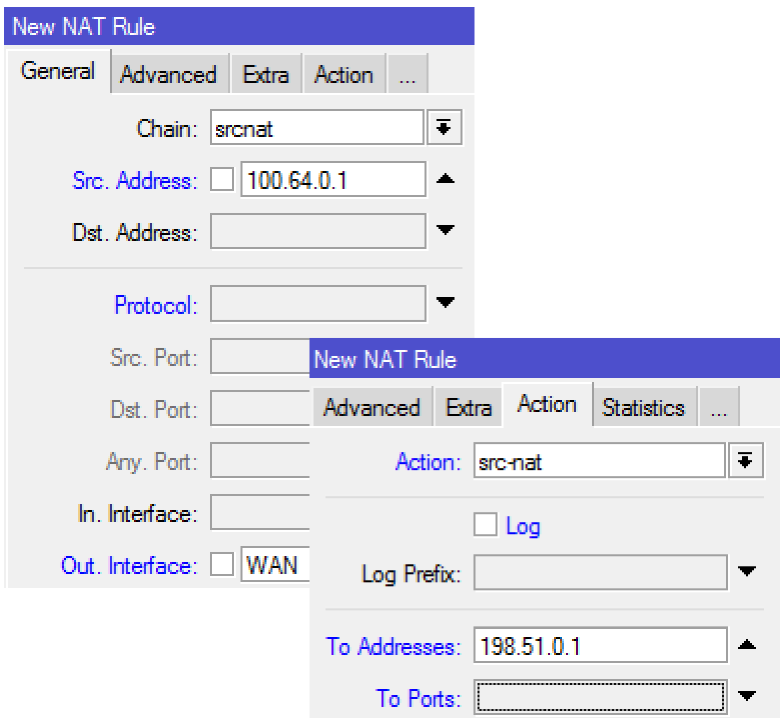


UDP Protocol



CGNAT

Other Protocols (non port-oriented)



For a sharing ratio of 25:1, total

3 x 25 = 75K rules!

Deploying CGNAT

With this approach we'll have 3 x 25 x 100 = 75k rules!

Fortunately RouterOS provides another features to make things better: "netmap"

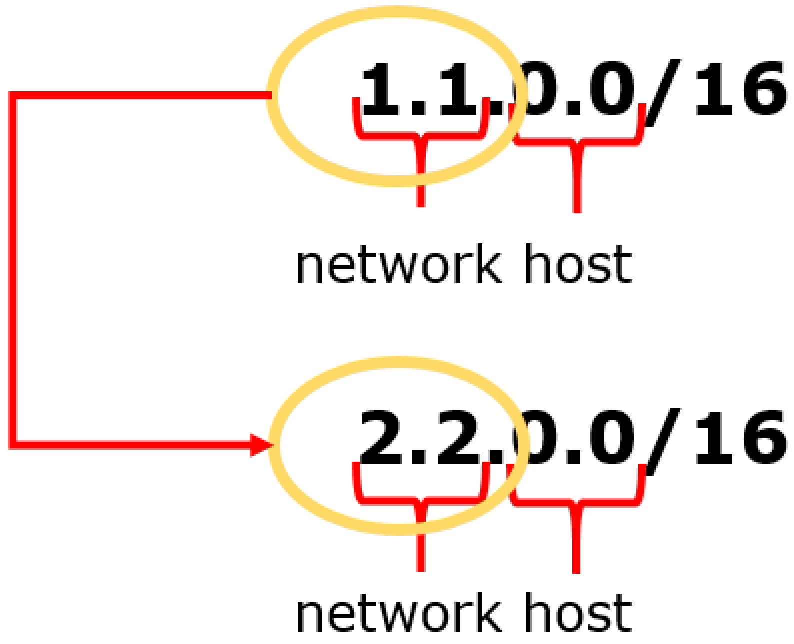
Netmap was initially implemented on Linux iptables in the packet "patch-o-matic" and ported to RouterOS.

MD Brasil started to deploy CGNAT with netmap in its own network and for some customers with good results.

How Netmap works

How netmap works

Netmap is an implementation of source or destination NAT where only the network part of an IP is natted. The host part remains as is. E.g. netmap network 1.1.0.0/16 into 2.2.0.0/16

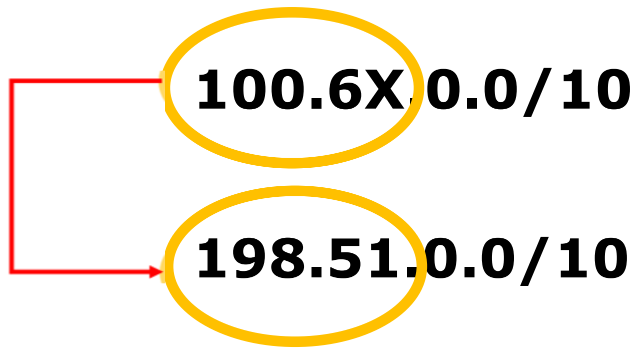


An IP address **1.1.X.Y** will be translated to **2.2.X.Y**

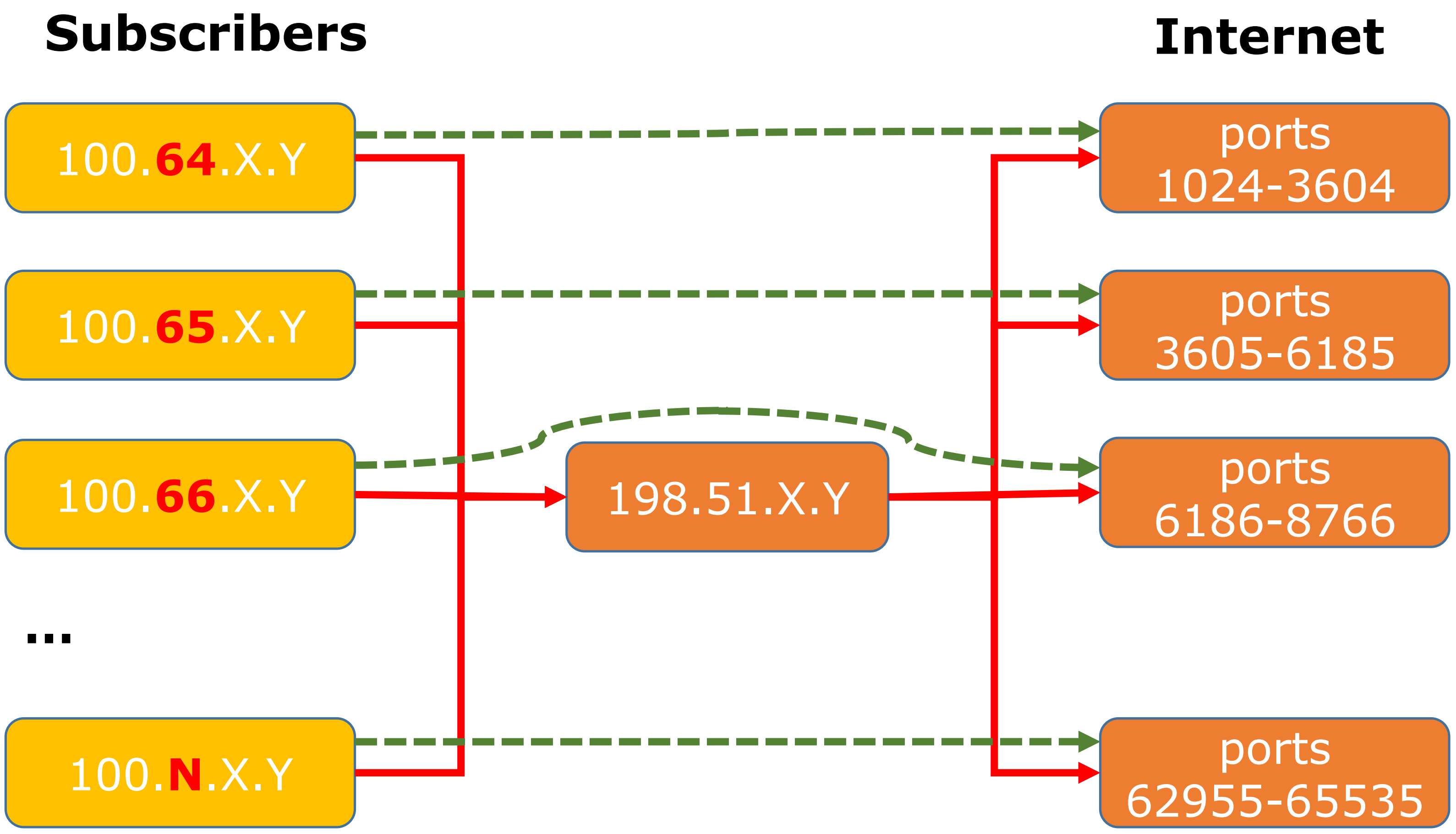
Using Netmap

Netmap makes NAT 1:1, where the host part is kept as is and only the network part is changed.

Our strategy will be to map the network 100.64 into our public IP address

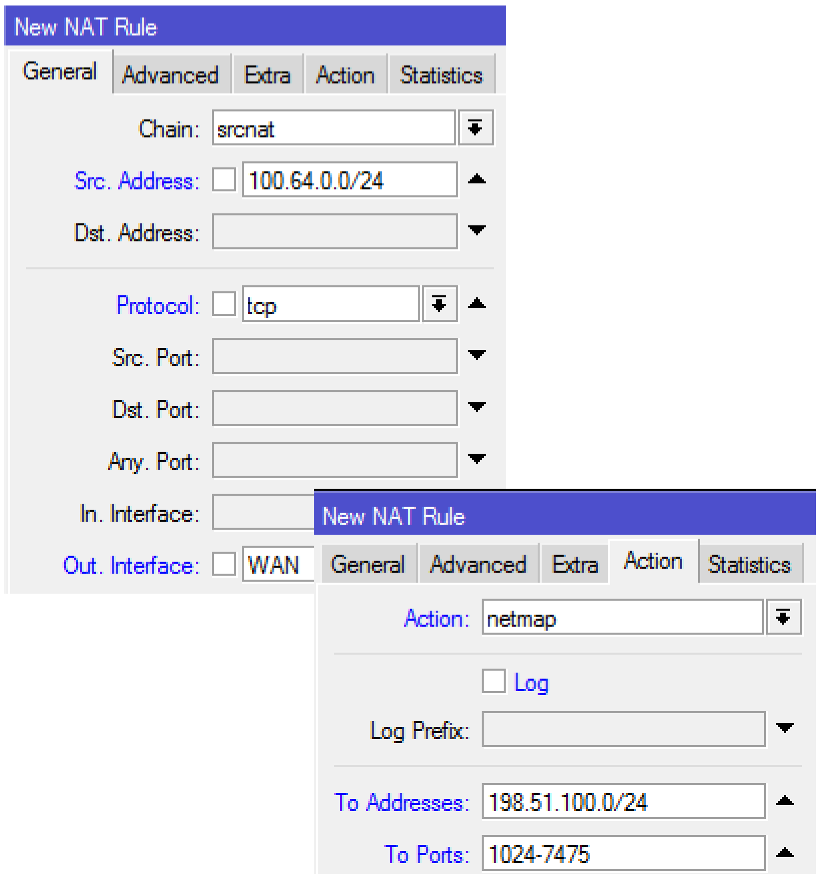


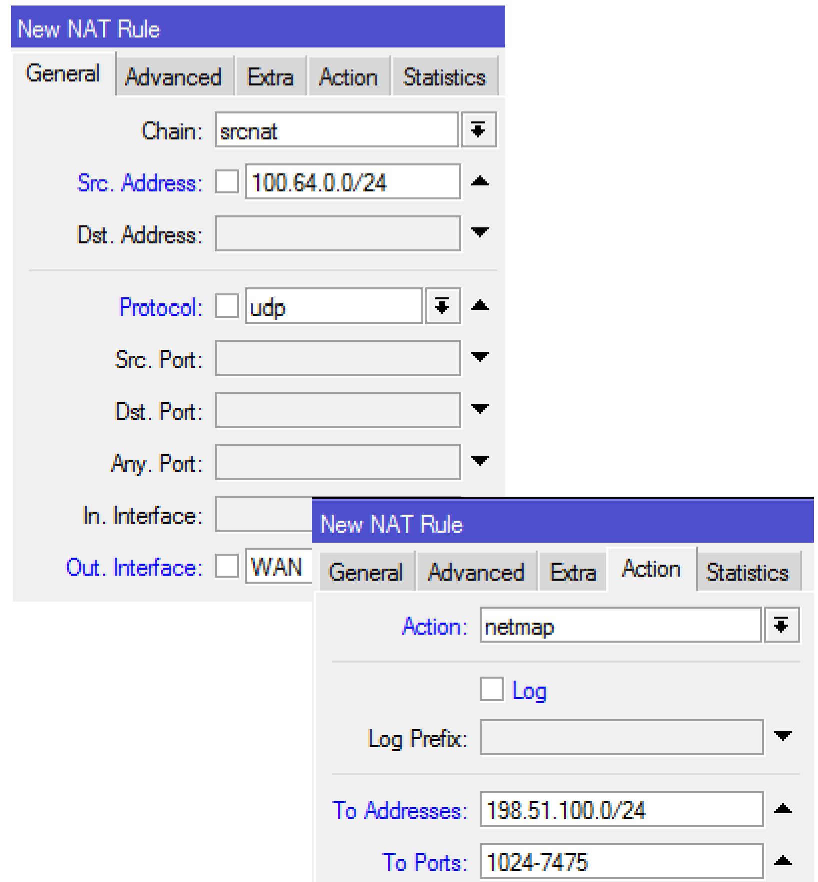
Using Netmap



Using Netmap

Typical netmap rules:





Using Netmap

The quantity of the rules will depend only on **sharing ratio.** If we have a sharing ratio of 1:N, we'll need

→ N rules for TCP

→ N rules for UDP

→ 1 rule for non port-oriented protocols

Regardless of the network size we'll have always

2N+1 rules!

Our hypothetical ISP will have a 51 rules, instead of 75K

RouterOS Implementation example

Netmap Example

Using a sharing ratio of 1:7

We'll have:

- 7 rules for TCP
- 7 rules for UDP and
- 1 rule for other protocols

- Total 15 rules

In this example we'll do a very simple schema to make more intuitive the distribution.

CGNAT Planning

Public prefix allocated to the LIR

- 198.51.0.0/22* (198.51.0.0 – 198.51.3.255)

Port division:

1) 1024 – 9999 (range 0);
2) 10000 – 19999 (range 1);
3) 20000 – 29999 (range 2);
4) 30000 – 39999 (range 3);
5) 40000 – 49999 (range 4);
6) 50000 – 59999 (range 5);
7) 60000 – 65535 (range 6);

* The prefix used in this presentation in fact is not a reserved range for documentation (RFC 5737).

CGNAT Planning

The proposition is to have the following distribution:

IP address allocated to a single user: 100.10**X.Y.Z**
where:

X = Port range (0 to 6)

Y = Third octet of the shared public IP address

Z = Fourth octet os the shared public IP address

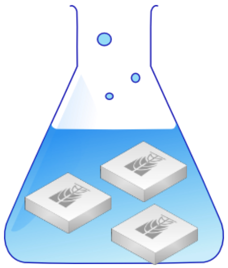
CGNAT Planning

Suppose we have one POP with the prefix 198.51.2.64/27

For instance, the IP 198.51.2.70 will be shared among 7 subscribers:

Subscriber	CGNAT IP	Public IP	Port Range
subscriber 1	100.**100**.2.70	198.51.2.70	**1024-9999**
subscriber 2	100.**101**.2.70	198.51.2.70	**10000-19999**
subscriber 3	100.**102**.2.70	198.51.2.70	**20000-29999**
subscriber 4	100.**103**.2.70	198.51.2.70	**30000-39999**
subscriber 5	100.**104**.2.70	198.51.2.70	**40000-49999**
subscriber 6	100.**105**.2.70	198.51.2.70	**50000-59999**
subscriber 7	100.**106**.2.70	198.51.2.70	**60000-69999**

CGNAT



Based on the schema, we can easily identify the subscriber who is behind the pair IP/port:

- 198.51.2.145, port 4045 → **100.100.2.145**
- 198.51.0.27, port 50045 → **100.105.0.27**
- 198.51.3.66, port 13016 → **100.101.3.66**

Netmap Rules

TCP Protocol

```
/ip firewall nat
add action=netmap chain=srcnat out-interface=wlan1 protocol=tcp src-address=100.100.X.Y/27 to-addresses=198.51.X.Y/27 to-ports=1024-9999
add action=netmap chain=srcnat out-interface=wlan1 protocol=tcp src-address=100.101.X.Y/27 to-addresses=198.51.X.Y/27 to-ports=10000-19999
add action=netmap chain=srcnat out-interface=wlan1 protocol=tcp src-address=100.102.X.Y/27 to-addresses=198.51.X.Y/27 to-ports=20000-29999
add action=netmap chain=srcnat out-interface=wlan1 protocol=tcp src-address=100.103.X.Y/27 to-addresses=198.51.X.Y/27 to-ports=30000-39999
add action=netmap chain=srcnat out-interface=wlan1 protocol=tcp src-address=100.104.X.Y/27 to-addresses=198.51.X.Y/27 to-ports=40000-49999
add action=netmap chain=srcnat out-interface=wlan1 protocol=tcp src-address=100.105.X.Y/27 to-addresses=198.51.X.Y/27 to-ports=50000-59999
add action=netmap chain=srcnat out-interface=wlan1 protocol=tcp src-address=100.106.X.Y/27 to-addresses=198.51.X.Y/27 to-ports=60000-65535
```

Netmap Rules

UDP Protocol

```
/ip firewall nat
add action=netmap chain=srcnat out-interface=wlan1 protocol=udp src-address=100.100.X.Y/27 to-addresses=198.51.X.Y/27 to-ports=1024-9999
add action=netmap chain=srcnat out-interface=wlan1 protocol=udp src-address=100.101.X.Y/27 to-addresses=198.51.X.Y/27 to-ports=10000-19999
add action=netmap chain=srcnat out-interface=wlan1 protocol=udp src-address=100.102.X.Y/27 to-addresses=198.51.X.Y/27 to-ports=20000-29999
add action=netmap chain=srcnat out-interface=wlan1 protocol=udp src-address=100.103.X.Y/27 to-addresses=198.51.X.Y/27 to-ports=30000-39999
add action=netmap chain=srcnat out-interface=wlan1 protocol=udp src-address=100.104.X.Y/27 to-addresses=198.51.X.Y/27 to-ports=40000-49999
add action=netmap chain=srcnat out-interface=wlan1 protocol=udp src-address=100.105.X.Y/27 to-addresses=198.51.X.Y/27 to-ports=50000-59999
add action=netmap chain=srcnat out-interface=wlan1 protocol=udp src-address=100.106.X.Y/27 to-addresses=198.51.X.Y/27 to-ports=60000-65535
```

Netmap Rules

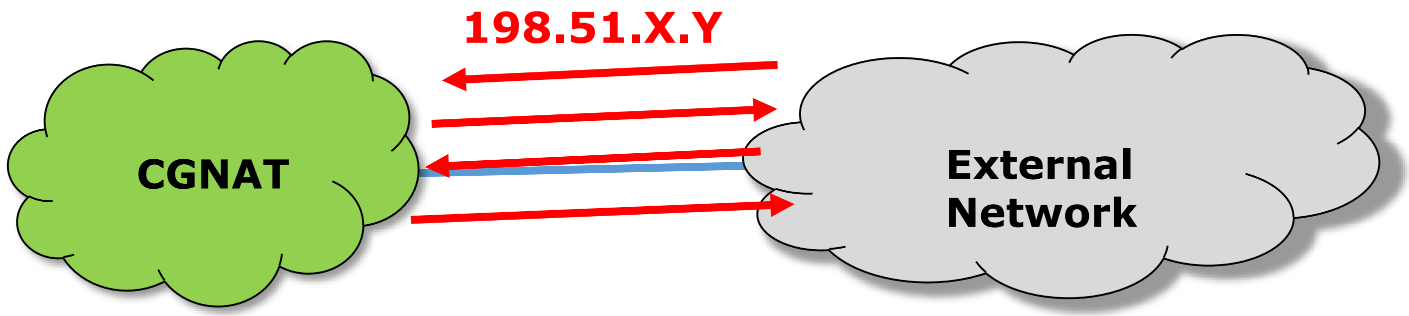
Non port oriented traffic

/ip firewall nat

add action=masquerade src-address=100.100.X.Y/27 chain=srcnat out-interface=wlan1

Netmap and "static loops"

With this implementation, any packet originated from outside the network and destined to one Public IP used by the CGNAT, won't have routes, leading to a static loop.



For instance, a ping originated on the Internet and destined to IP 198.51.X.Y will arrive in the concentrator, and sent back to the last router, which will sent to the concentrator again, etc, etc.

Netmap and "static loops"

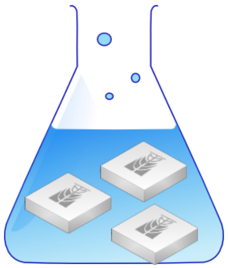
Possible solutions:

- Destination NAT rule pointing to a public IP address or to a blackhole IP configured on the concentrator;

- All public IP addresses configured on a loopback interface on the concentrator.

- Simply a blackhole route for the entire public network on the concentrator.

What about Network Topology?



The rules can be applied without any change in the current topology.

For distributed authentication:

→ Netmap rules on the concentrators

For centralized authentication:

→ Netmap rules on the central concentrator

Port Forwarding with CGNAT

For port forwarding, besides the normal dst-nat rule on the CPE, it is necessary another rule at concentrator level.

A previous set of rules could help support desk

For instance, both users **100.64.X.Y e 100.65.X.Y** want forwarding to port 80

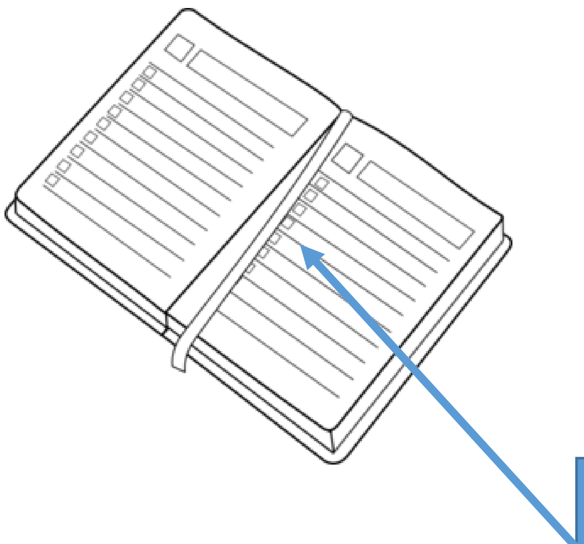
IP/external port	CGNAT IP	Internal IP
198.51.X.Y:80**64**	100.**64**.X.Y:80	192.168.1.180
198.51.X.Y:80**65**	100.**65**.X.Y:80	192.168.1.180

inform to subscribers

Pre configured on concentrator

Pre configured on CPE
Inform to subscribers

Agenda



Current status of IPv4 exhaustion and IPv6 adoption;

Issues related to IPv4 scarcity and shared address solutions;

CGNAT implementation with low cost and good performance;

Best practices for IPv6 deployment in an small/medium ISP access network;

37’

IPv4 x IPv6

Basic comparison:

	IPv4	IPv6
Address Size	32 bit	128 bit
Address Format	192.168.1.1	2001:db8:1:2:3:4:5:6
Possible Combinations	2^32	2^128

4,294,967,296

340,282,366,920,938,463,463,374,607,431,768,211,456

Planning IPv6 distribution

Planning IPv6 Deployment

2	0	0	1	:	0	d	b	8	:	0	1	2	3	:	4	5	6	7	:
04	08	12	16		20	24	28	32		36	40	44	48		52	56	60	64	

- Minimum allocation for ISPs in all RIRs is a /32

- Longer prefix allowed in BGP is a /48

- Minimum allocation that allow SLAAC to work is a /64

How Many Prefixes per Subscriber?

2	0	0	1	:	0	d	b	8	:	0	1	2	3	:	4	5	6	7	:
04	08	12	16		20	24	28	32		36	40	44	48		52	56	60	64	

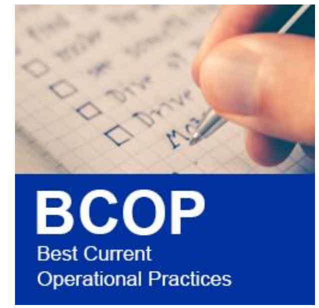
- Currently recommended allocation size (RFC6177):
  - /56 for residential customers
  - /48 for business customers

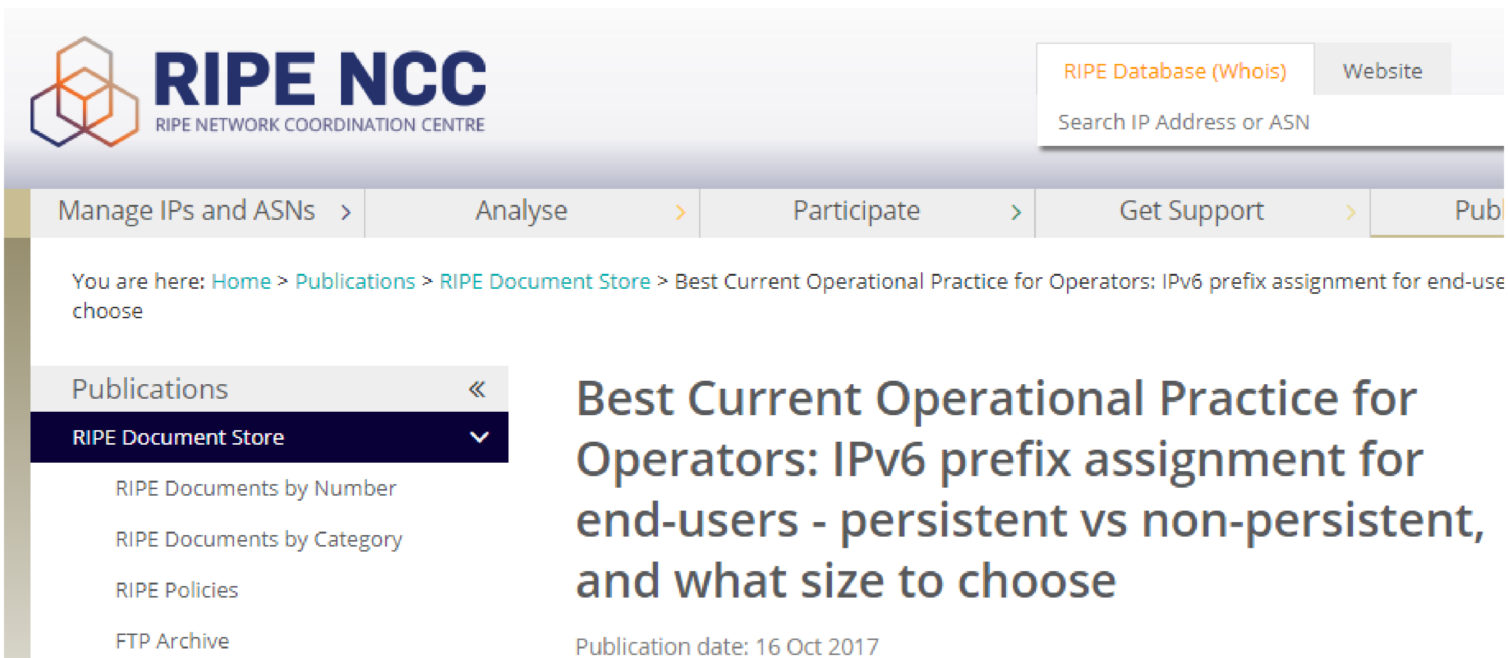
/56 → 256 /64 subnets

/48 → 65536 /64 subnets

Planning IPv6 Distribution

RIPE BCOP (Doc 690 – October 2017) recommends minimum allocation of /48 even for domestic users!





md brasil TI&TELECOM

/48 for everybody?

Would be RIPE trying to exhaust IPv6?

☺

/48 for everybody?

In fact RFC 3177 (2001), already recommended the use of /48 for residential and enterprise subscribers:

RFC3177

"....

In particular, we recommend:

- Home network subscribers, connecting through on-demand or always-on connections should receive a /48.

- Small and large enterprises should receive a /48.

- Very large subscribers could receive a /47 or slightly shorter prefix, or multiple /48's.

..."

/48 for everybody?

Ten years later RFC 6177 revised the understanding, however without much reasons for that.

RFC6177

"....

*While the /48 recommendation does simplify address space management for end sites, it has also been **widely criticized as being wasteful***

...

*While it seems likely that the size of a typical home network will grow over the next few decades, **it is hard to argue that home sites will make use of 65K subnets** within the foreseeable future.*

..."

/48 for everybody?

RIPE operators resumed the discussion and published the document RIPE690

"*...*

4.2.1. /48 for everybody

This is probably the most practical way to assign IPv6 prefixes to end customer CPE devices. In this case everyone has a /48 prefix and advanced end-users are less likely to make mistakes when addressing their networks and devices, resulting in much less call-centre time to sort out problems. It also has the advantage of sharing the same prefix size as ULAs and some transition mechanisms, so this facilitates a direct mapping of existing customer addressing plans to the delegated prefix.

…"

/48 for everybody?

BCOP 690 doesn't "condemn" RFC-6177, but points that such differentiation has much more commercial reasons than technical ones.

"*...*

4.2.2. /48 for business customers and /56 for residential customers

Some operators decide to give a /48 prefix to their business customers and a /56 to their residential customers. ***This rationale is understood to be mainly coming from sales and marketing departments where they wish to create some distinction in services between different types of customer.*** *This method can be considered as pragmatic, future-proof and has nearly no downsides, the same as the "/48 for everyone" approach.*

..."

/48 for everybody?

And suggests for the ones who want to provide a /56 for residential subscribers (for any reason) that reserve a /48 and use the first /56 for the customer.

"...

An alternative is to reserve a /48 for residential customers, but actually assign them just the first /56. If subsequently required, they can then be upgraded to the required prefix size without the need to renumber, or the spare prefixes can be used for new customers if it is not possible to obtain a new allocation from your RIR (which should not happen according to current IPv6 policies)

..."

md brasil TI&TELECOM

Is my block enough big to give /48 for all?

/48 for everybody?

Questions:

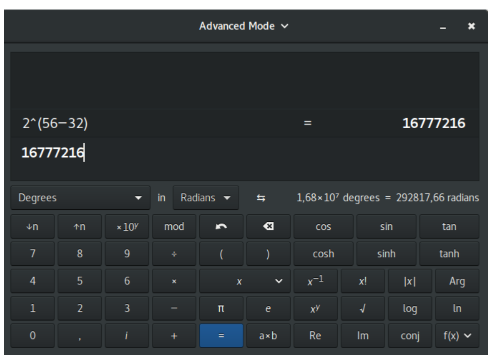
1) If we have a /32, how many customers could we provide service giving a /56?

2) And what about giving /48 for everybody?

/48 for everybody?

Questions:

1) If we have a /32, how many customers could we provide service giving a /56?



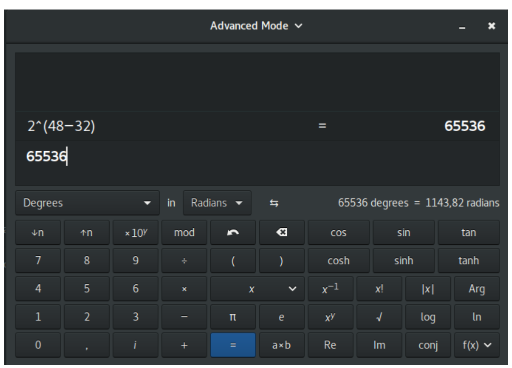
2 ^ (56 – 32) = 16.777.216

A little bit more then 16 million subscribers ☺

/48 for everybody?

Questions:

1) And what about providing /48 for everybody?



2 ^ (48 – 32) = 65.536

Happy with 65K subscribers? ☺ ☺ ☺ ☺

Another good reason in favor of /48

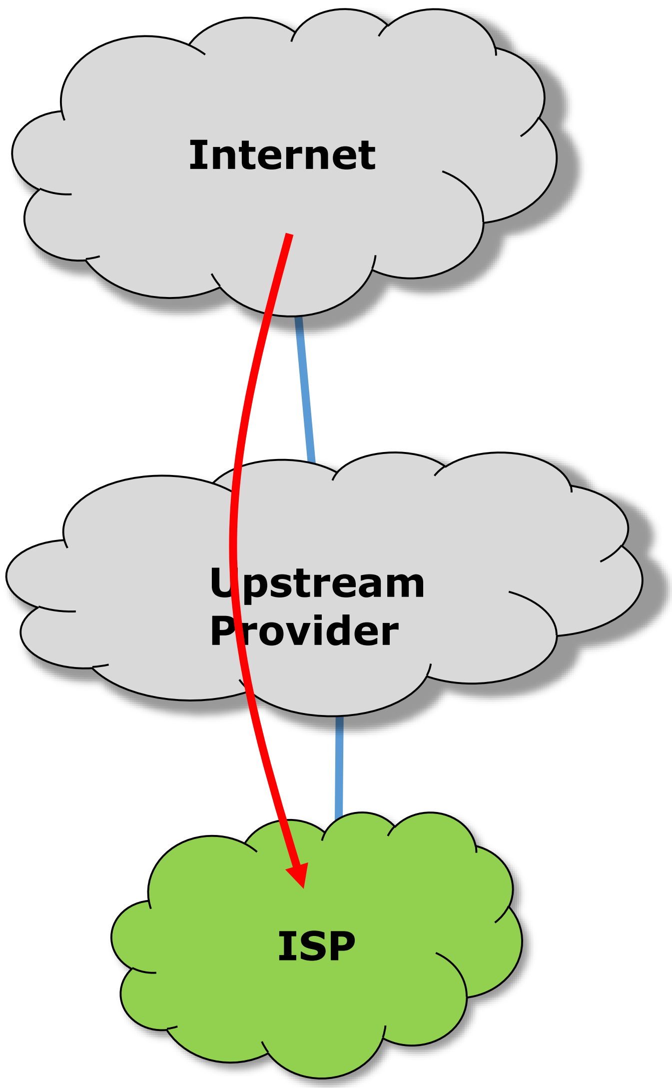
Another good reason that is not mentioned in BCOP 690 is security related.

DDoS mitigation techniques for volumetric attacks can be improved in the cases of:

RTBH – Remote Triggered Blackhole

Mitigation done by a Scrubbing Center

Remote Triggered Blackhole in IPv4

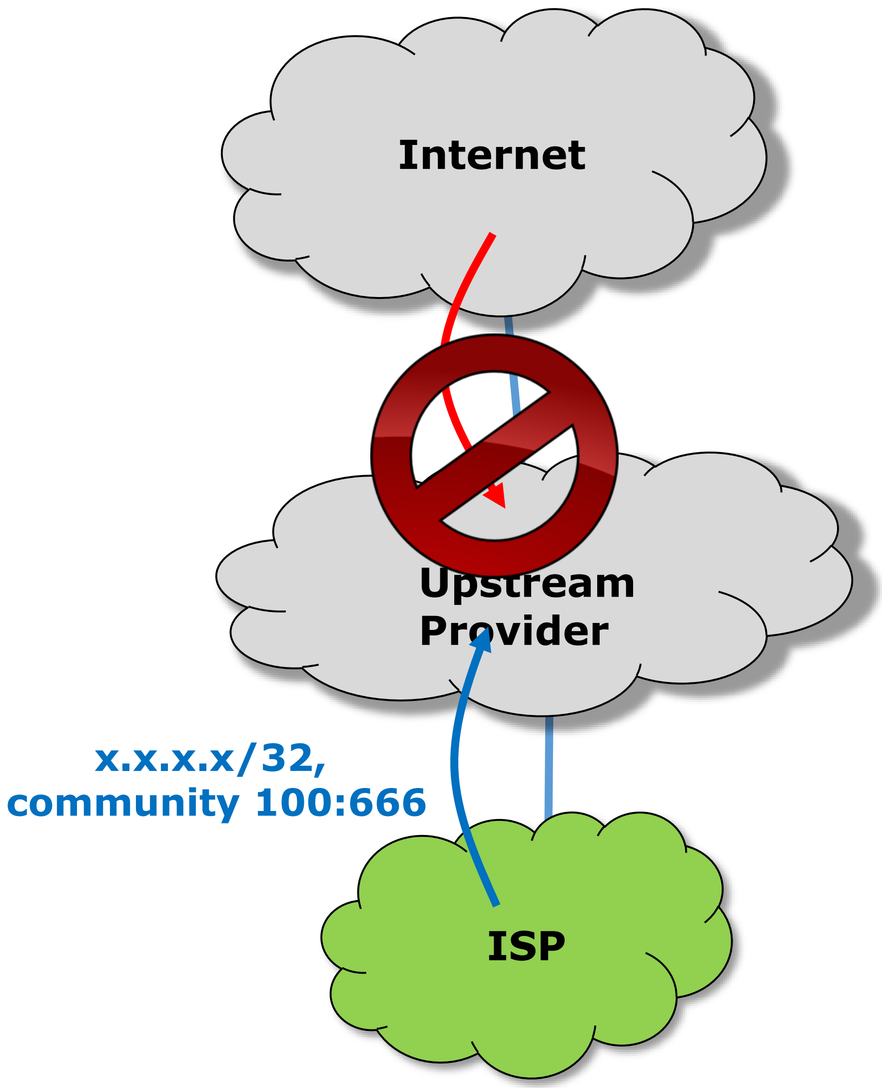


ISP is suffering a DDoS attack targeting some IPv4 /32;

Upstream provider (e.g. AS 100) provides a policy that black-hole any /32 announcement with a specific community (e.g. 100:666);

http://mum.mikrotik.com/presentations/EU16/presentation_2960_1456752556.pdf

Remote Triggered Blackhole in IPv4

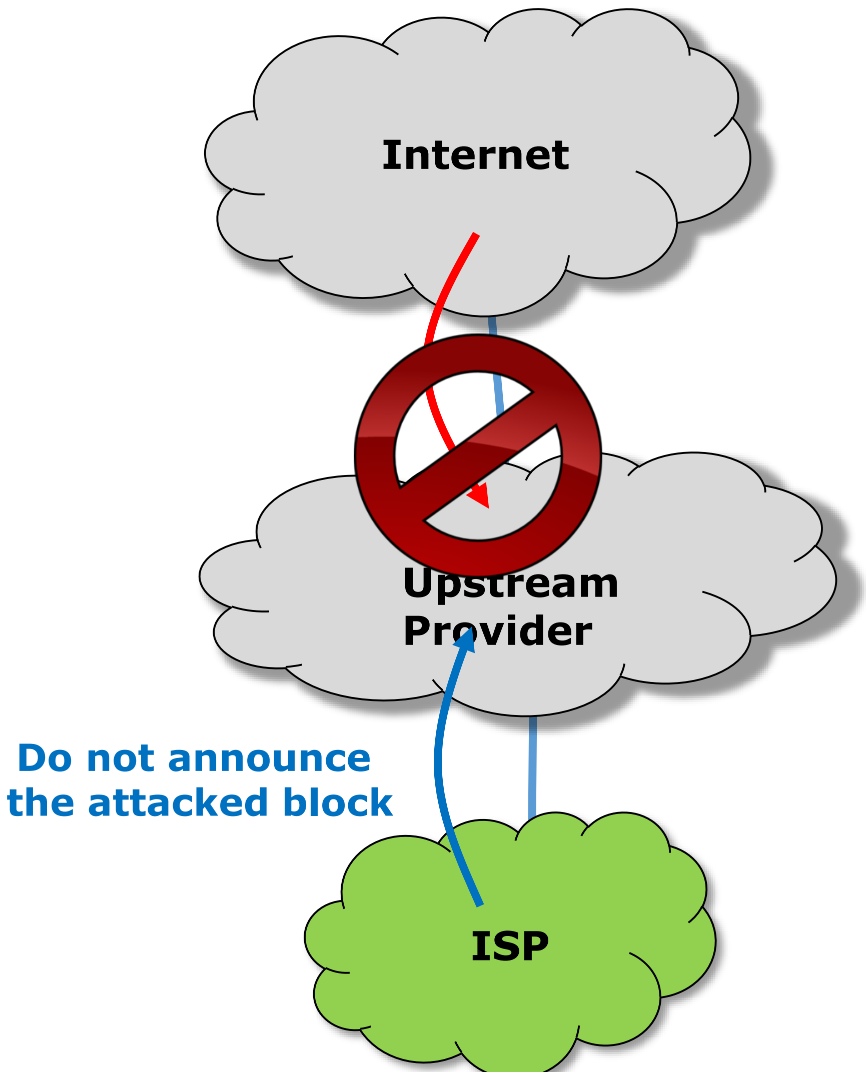


ISP announces to the Upstream provider the /32 with the community;

Upstream provider put the /32 in blackhole;

Communication with /32 is lost and channel overflow stops;

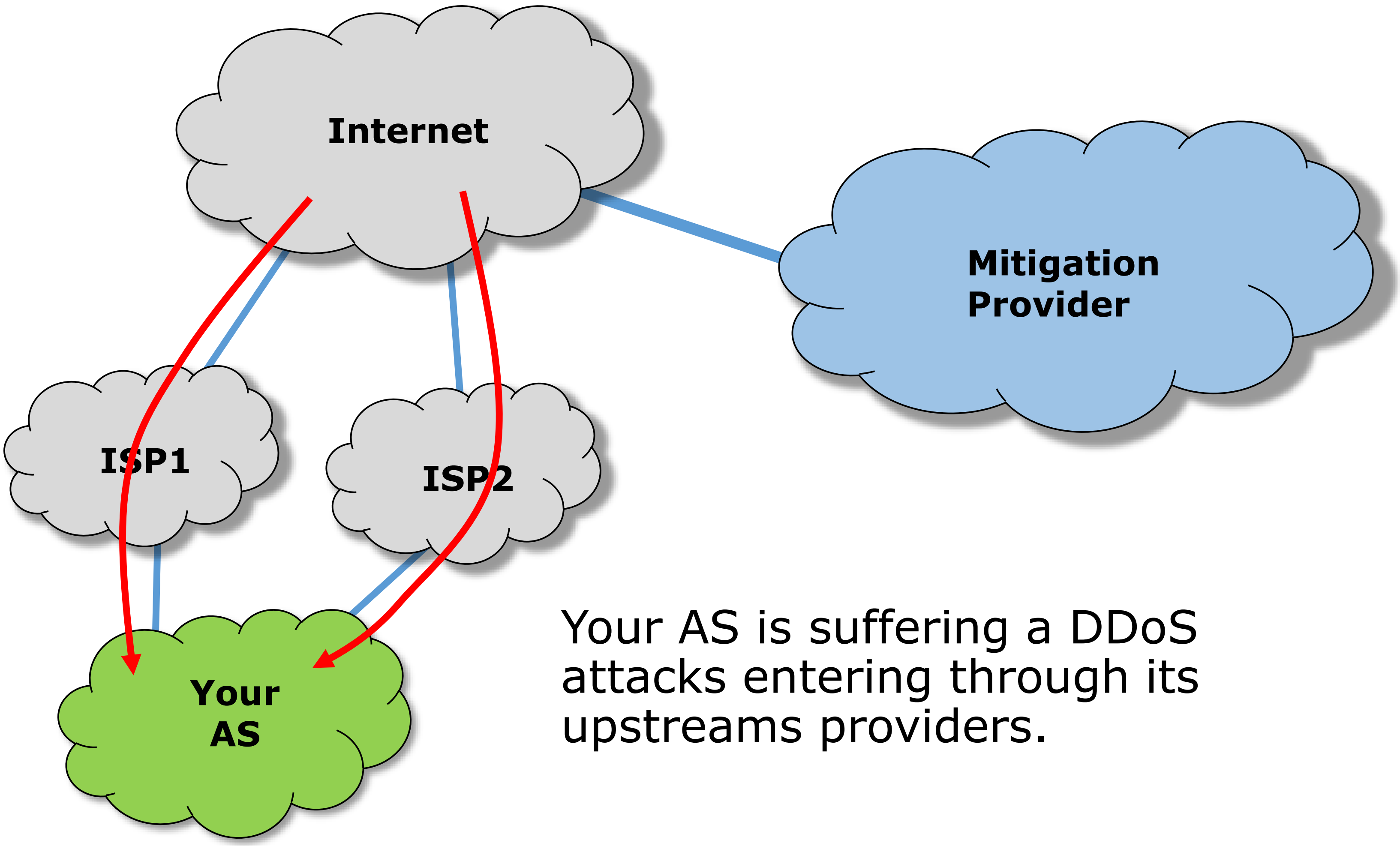
Remote Triggered Blackhole in IPv6



The same could be done with IPv6, the difference is that when it comes to IPv4 your upstream provider MUST have a blackhole policy and you depend on him;

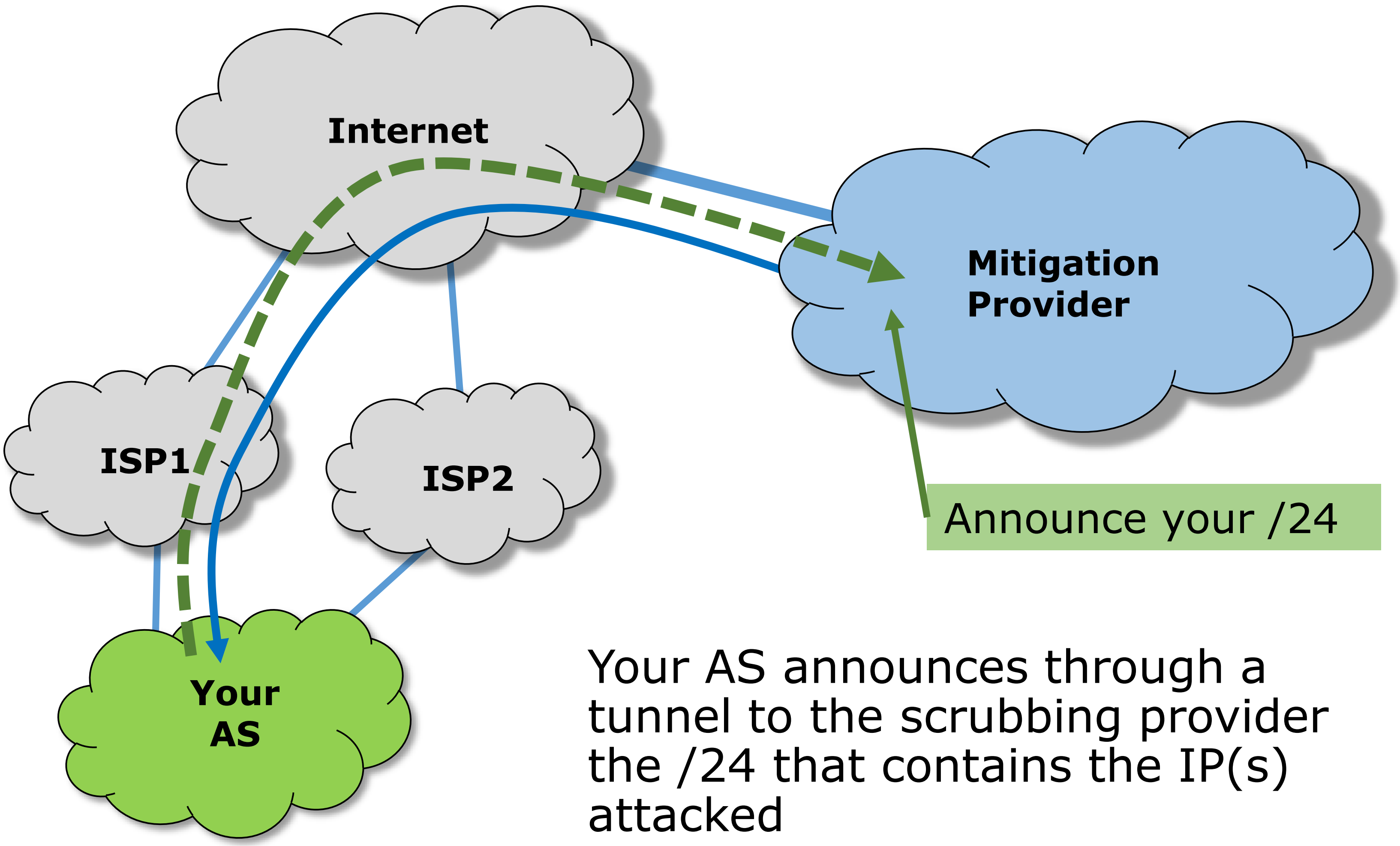
If you distribute IPv6 /48, you split your announcements and simply **do not announce the attacked block** and the attack will stop regardless of your upstream provider!

Mitigation On the Cloud on IPv4



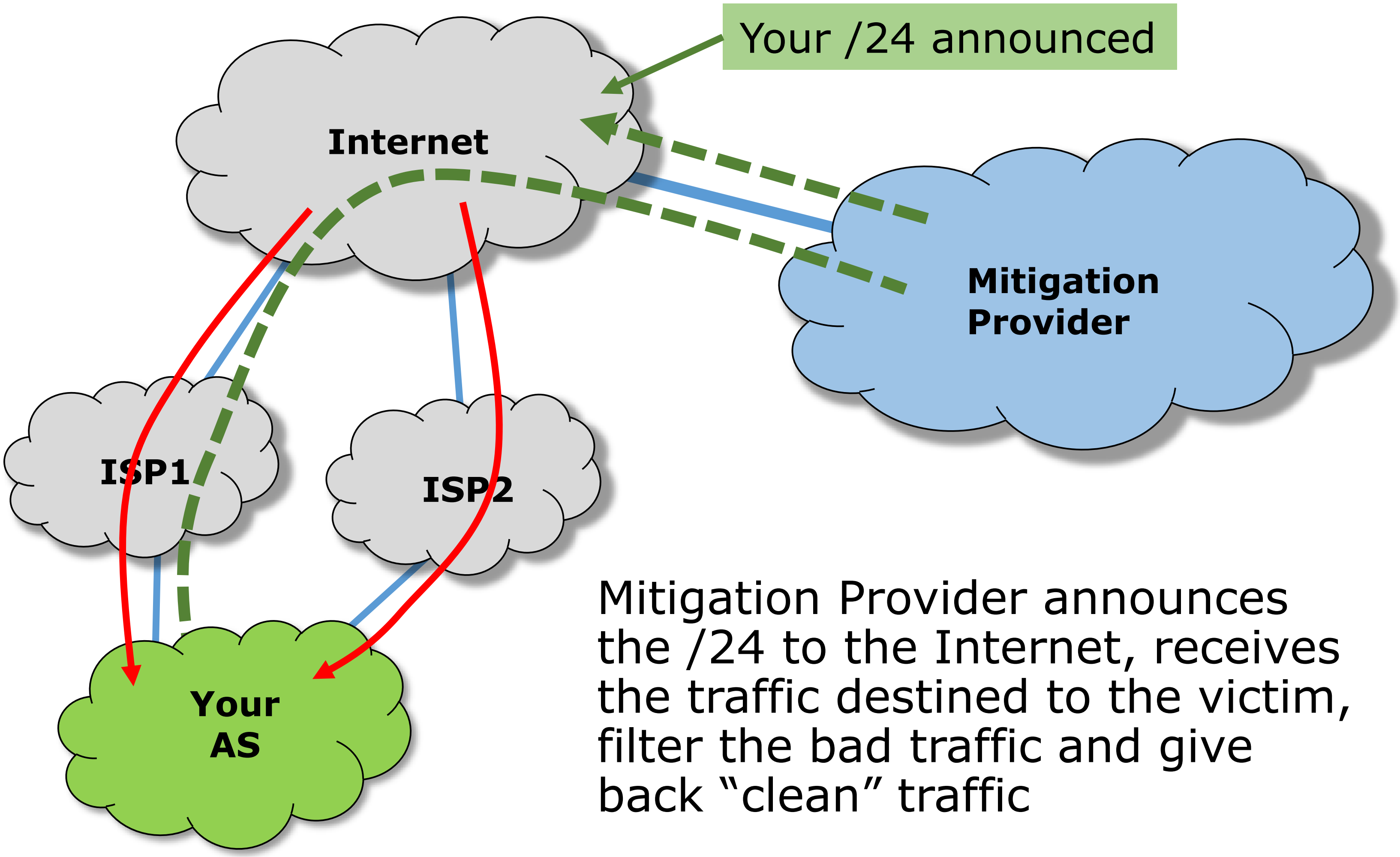
Your AS is suffering a DDoS attacks entering through its upstreams providers.

Mitigation On the Cloud in IPv4



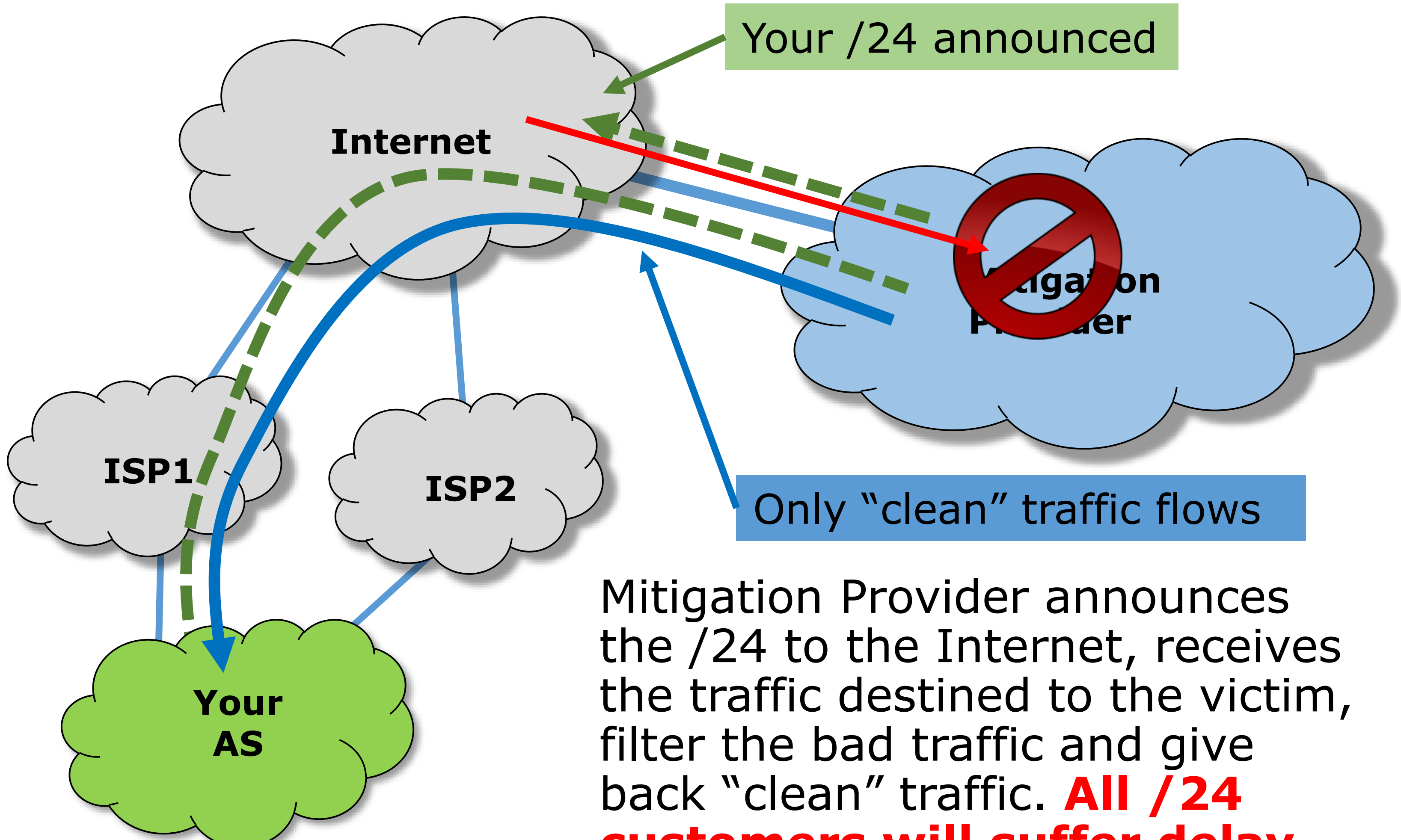
Your AS announces through a tunnel to the scrubbing provider the /24 that contains the IP(s) attacked

Mitigation On the Cloud in IPv4



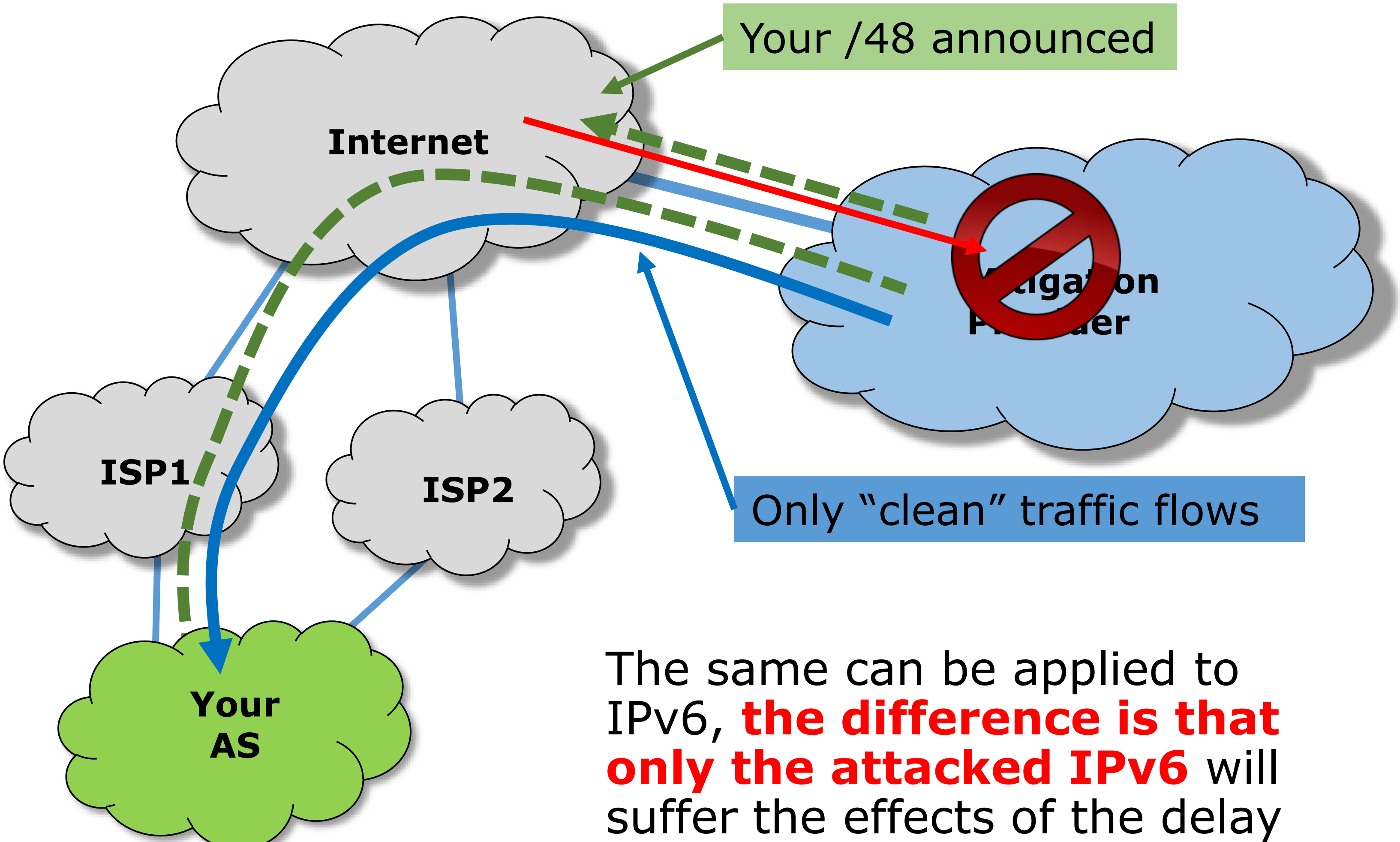
Mitigation Provider announces the /24 to the Internet, receives the traffic destined to the victim, filter the bad traffic and give back "clean" traffic

Mitigation On the Cloud in IPv4



Mitigation Provider announces the /24 to the Internet, receives the traffic destined to the victim, filter the bad traffic and give back “clean” traffic. **All /24 customers will suffer delay due to the tunnel.**

Mitigation On the Cloud in IPv6



The same can be applied to IPv6, **the difference is that only the attacked IPv6** will suffer the effects of the delay

Back to the Planning

IPv6 Anatomy

4 bits

2001:0db8:0123:4567:89AB:CDEF:0123:4567

8 bits

8 x 16 = 128 bits

md brasil TI&TELECOM

Back to the Planning

2 0 0 1 : 0 d b 8 : 0 1 2 3 : 4 5 6 7 :

04 08 12 16 20 24 28 32 36 40 44 48 52 56 60 64

Considering the "anatomy" of IPv6, it is interesting to plan the distribution with hops of 4 bits.

This will turn your distribution "cleaner", easy to understand and could avoid future configuration mistakes.

Back to the planning

e.g. POPs

Each City - 16 Pops /40

2001:db8:**01**00/40
2001:db8:**02**00/40
2001:db8:**03**00/40
2001:db8:**04**00/40
.....
2001:db8:**0F**00/40

16 /36 (e.g. Cities)

2001:db8:**01**00/40
2001:db8:**02**00/40
2001:db8:**01**00/40
2001:db8:**02**00/40
2001:db8:**01**00/40
2001:db8:**02**00/40

/32
(ISP Block)

2001:db8::/32

Back to the Planning

The allocation of /48 for everybody will turn your addressing plan very easy and understandable.

2001:db8:**0**000/36
2001:db8:**1**000/36
2001:db8:**2**000/36
2001:db8:**3**000/36
.....
2001:db8:**F**000/36

→

2001:db8:**01**00/40
2001:db8:**02**00/40
2001:db8:**03**00/40
2001:db8:**04**00/40
.....
2001:db8:**0F**00/40

→

2001:db8:**0100**/48
2001:db8:**0201**/48
2001:db8:**0302**/48
2001:db8:**0403**/48
.....
2001:db8:**0FFF**/48

256 /48 per POP

or

4096 /48 per City

md brasil TI&TELECOM

Back to the planning Infrastructure Addresses

Consider to reserve a part of the last /36 for infrastructure and divide in parts that that **may** be announced and **must** be announced.

/32

(ISP Block)

2001:db8::/32 →

- 2001:db8:**0**000/36
- 2001:db8:**1**000/36
- 2001:db8:**2**000/36
- 2001:db8:**3**000/36
- .....
- 2001:db8:**F**000/36

must be announced

2001:db8:**FE**00::/40

may be announced

2001:db8:**FF**00::/40

Continuing BCOP 690

Fixed or Dynamic Addresses?

Fixed or Dynamic Addresses

Some problems related to dynamic addresses:

→ Logs / accounting for tracking;

→ Issues related to internal services

→ Problems related to power outages

BCOP 690 recommendation:

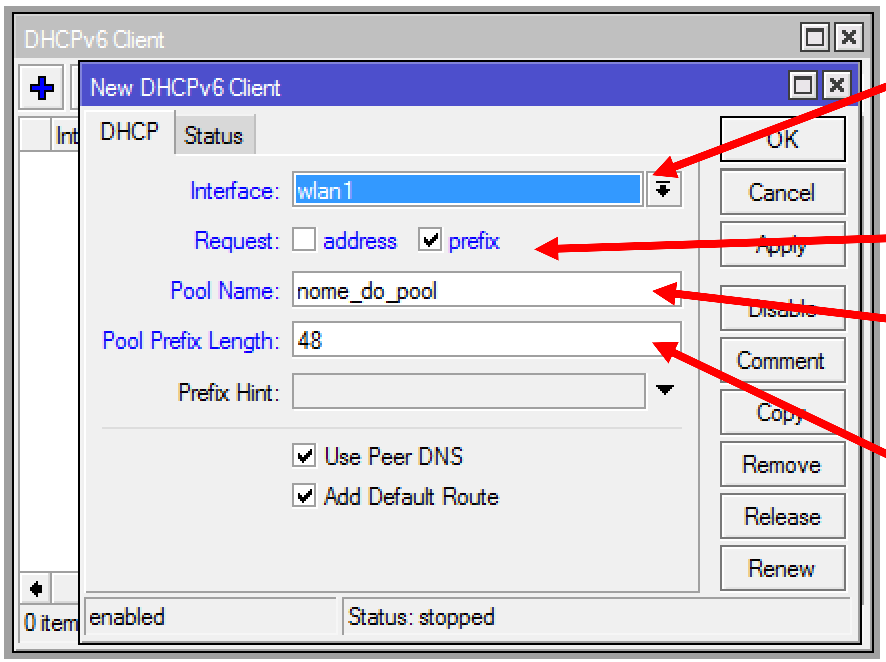
→ To use fixed (permanent) addresses

→ If for some reason (commercial for instance) you don't want to go this way, at least configure a big lifetime for the connection;

PPP and DHCPv6 support in RouterOS

DHCPv6 Client

DHCPv6 PD client get prefixes from a DHCPv6 PD server, and can subdivide among the subscribers, inserting a route to the DHCPv6 Server.



Interface where this client will run

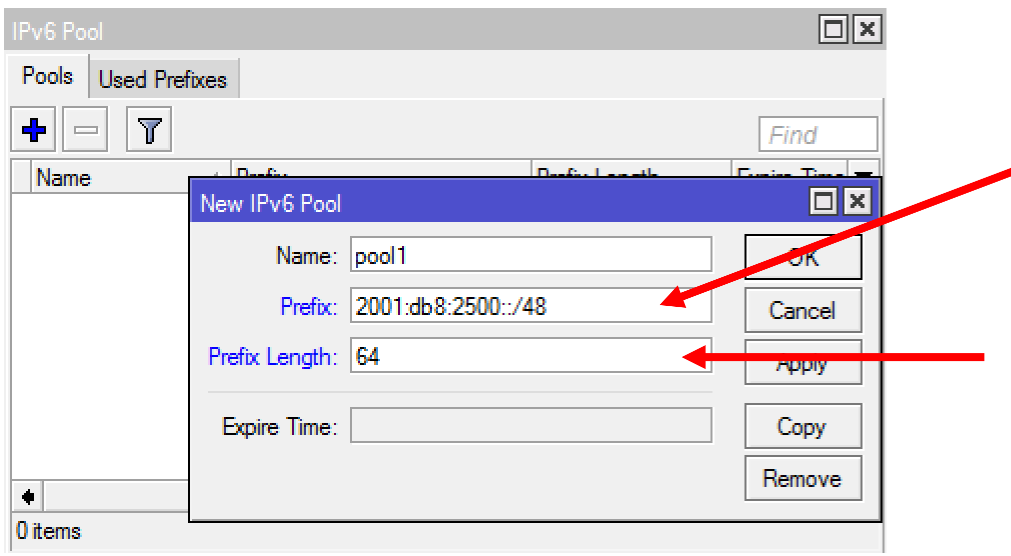
To request a prefix

Internal Pool that will be created

Prefix Length

IPv6 Pool

IPv6 Pool defines an address range for future use that will be available for SLAAC, DHCPv6 and PPP servers



Prefix allocated to the router

Bitmask for splitting the prefix. E.g. for SLAAC

PPP Support

Dynamic Allocation (not recommended)

Configuring a pool in the concentrator and distributing the prefixes via PPP;

For traceability, you must use another technique like a script “on-logon” and “on-logout” and send the logs to a Remote Syslog;

PPP Support

Fixed delegation via RADIUS

RouterOS offers support to PD (Prefix Delegation), however it is not possible to directly get the prefixes from a RADIUS server using the attribute "Delegated-IPv6-Prefix";

This attribute is not supported and there is a thread in Mikrotik Forum about this:

https://forum.mikrotik.com/viewtopic.php?t=89443

Although, as we'll see it is possible to circumvent this limitation.

PPP Support

Circumventing Fixed Allocation via RADIUS

RouterOS supports the attribute “Mikrotik-Delegated-IPv6-Pool” (string)

This string can be associated to a specific user in RADIUS

The Concentrator has to have a pool with the same name of the string.

The pool will be delegated to that specific user.

Conclusions

Conclusions

Some affirmatives can sound obvious, but it’s important to have in mind:

→ CGNAT techniques with low cost and good performance can be very helpful in the current scenario of scarcity. However CGNAT is not sustainable in the long term. Remember that ports are finite too!

→ IPv6 is totally different of IPv4. We should not use the same concepts and paradigms. IPv6 is Abundance, IPv4 is Scarcity.

→ For the ones that didn’t start IPv6 don’t wait the time limit, because there will be no time limit. Time limit has gone…

References

References

BCOP 690

Best Current Operational Practice for Operators: IPv6 prefix assignment for end-users - persistent vs non-persistent, and what size to choose

https://www.ripe.net/publications/docs/ripe-690

BCOP 631

IPv6 Troubleshooting for Residential ISP Helpdesks

https://www.ripe.net/publications/docs/ripe-631

Vielen Dank!