

# Tips & Tricks to connect to a Layer2 IXP

MUM Chile 2019



# Agenda

- Introducción
- creando peers BGP4
- Filtros BGP (in & out)
- Usando Comunidades BGP
- Filtros de seguridad mínimos
- Tips para Layer2
- Q & A

# Quien les habla



- **Jaime Cruz M.** <jaime.cruz@hablaip.com>
  - Integrador de sistemas con más de 10 años de experiencia en Redes IP, telecomunicaciones, Sistemas Linux, Mikrotik, Asterisk y telefonía IP.



- Interconnection Strategy en IXP Pit Chile.

- **Introducción**
- creando peers BGP4
- Filtros BGP (in & out)
- Usando Comunidades BGP
- Filtros de seguridad mínimos
- Tips para Layer2
- Q & A

# Introducción

- Que es un IXP?

Un PIT o Punto de Intercambio de Tráfico (IXP Internet eXchange Point en inglés) es una instalación donde se intercambian los datos entre diferentes redes de diferentes proveedores y consumidores de internet.

Es decir, un PIT permite interconectar directamente (algunas de) las redes que conforman el internet.

# Introducción

- Que es un IXP?

- Esquema tradicional  
Con un proveedor de Internet.



# Introducción

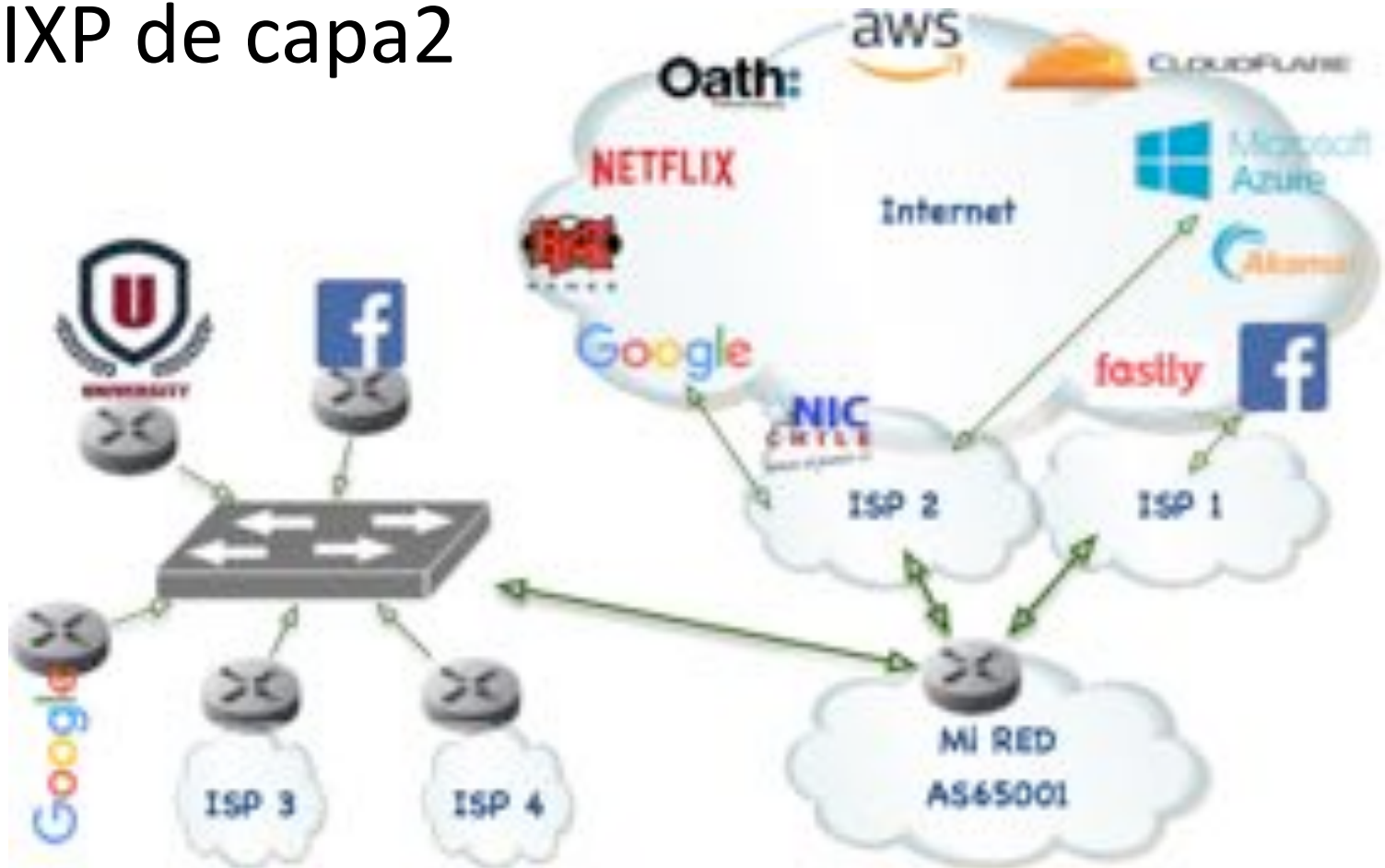
- Que es un IXP?

- Esquema tradicional  
Con /dos/ proveedores  
De Internet.



# Introducción

- Un IXP de capa2





# Introducción

- IXP/PIT's en Chile: (según Subtel)

- PIT Claro Chile
- PIT Entel
- PIT Level3
- PIT Intercity
- PIT NAP Chile
- PIT Orange
- PIT Telefónica Mundo
- PIT Chile
- PIT Tecnoera
- PIT Concepción

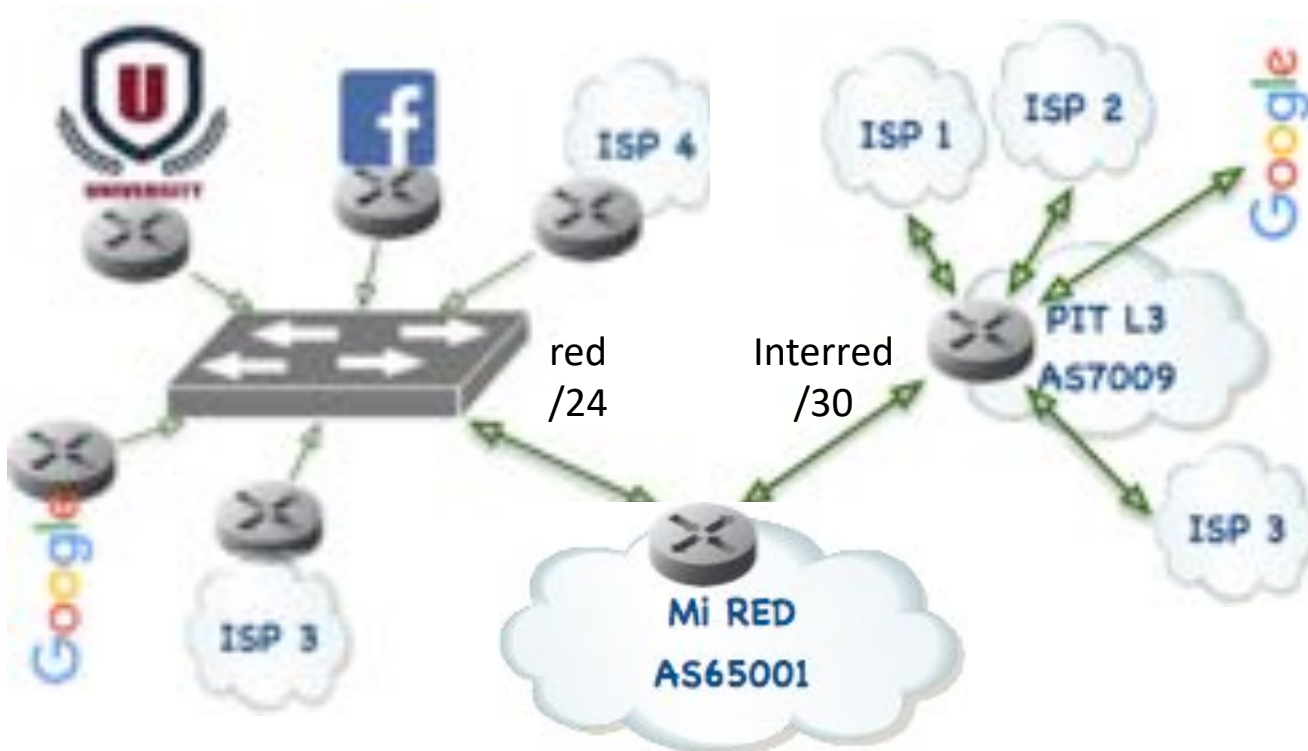
Valparaiso

(fuente: <https://www.subtel.gob.cl/normativa-tecnica-internet> )

# Introducción

- Diferencia entre IXP Layer3 v/s Layer2

Los PIT en capa2 ofrecen conectividad total y libre entre sus miembros.



# Introducción

- Requisitos para conectar a un IXP.
  - Contar con recursos IPv4/IPv6
    - Minimo /24 en IPv4.
  - Contar con un ASN (número autónomo)
  - Solicitarlos en LACNIC para Latinoamerica
    - [www.lacnic.net](http://www.lacnic.net)
  - Contar con equipamiento de red que soporte BGP4. ← Mikrotik ;-)

# Introducción

- Que es un ASN (Sistema Autonomo)?
  - Conjunto de redes administrado por una misma entidad, que a su vez tiene autonomía en internet.
  - Entidades de gobierno, financieras, universidades, proveedores de internet, etc.
  - Ej. SII(AS15208), Codelco(AS52226), BancoChile(AS22975), U.Chile(AS23140), GTD(AS14259), etc.

## Que es un BGP4?

- Protocolo de enrutamiento de borde o /exterior/
  - Conecta diferentes Sistemas Autónomos
  - Utiliza el puerto tcp/179
- 
- Internet es un conjunto de sistemas autónomos que hablan BGP.

- Introducción
- **creando peers BGP4**
- Filtros BGP (in & out)
- Usando Comunidades BGP
- Filtros de seguridad mínimos
- Tips para Layer2
- Q & A

# Mikrotik en el borde

- Ejemplo de organización conectada a dos ISP's y un PIT.
  - RED a conectar: AS65001
  - ISP A: AS18474 nac. e inter.
  - ISP B: AS7094 inter.
  - PIT CL: AS61522

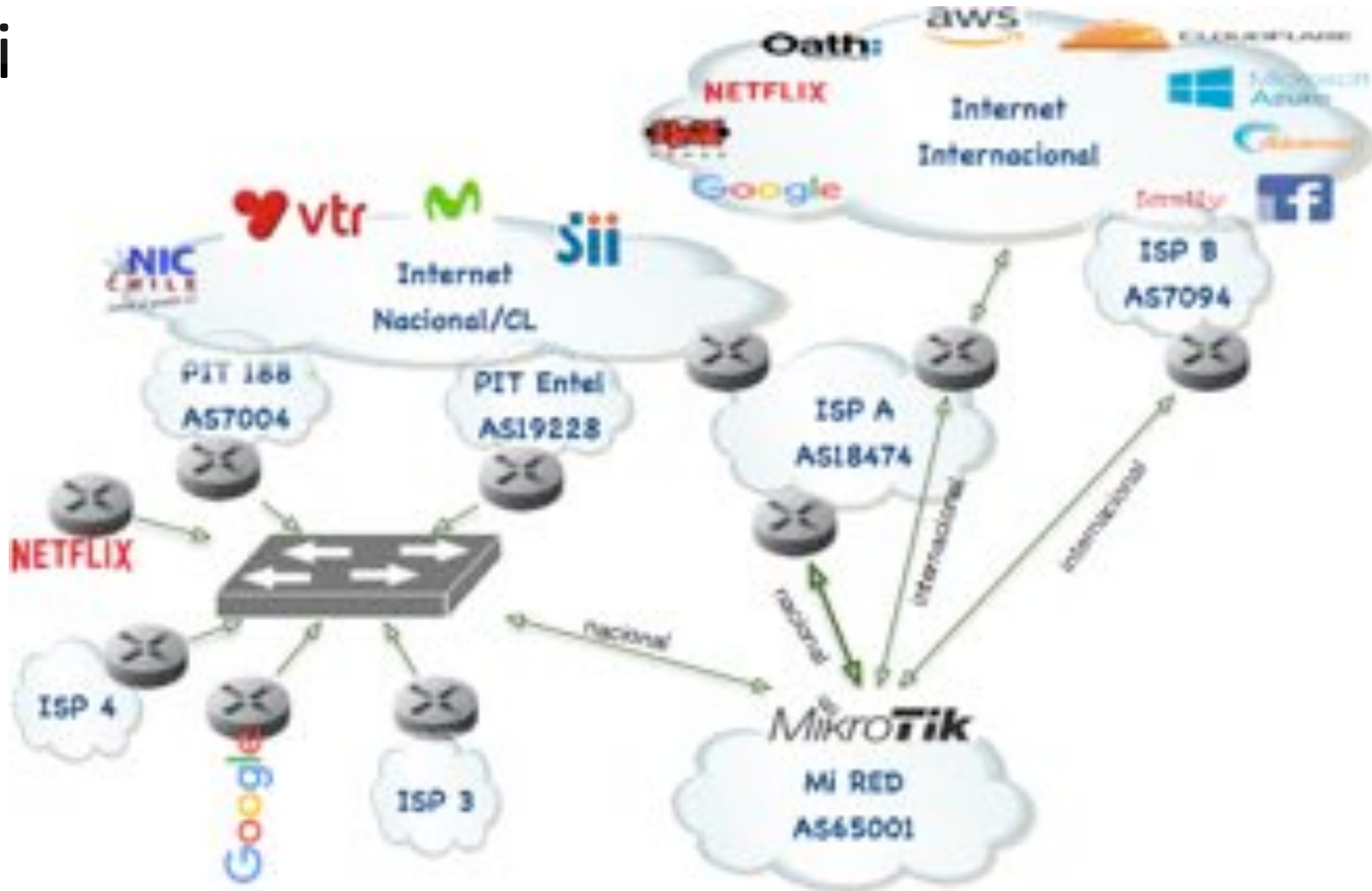


## Mikrotik en el borde





## Mi



# Creando peers BGP4

- Crear una “Instancia BGP”
- Routing->BGP->Instance-> +
  - Nombre descriptivo
  - ASN local es 65001
  - RouterID opciones
  - “redistribute XXXXX”

New BGP Instance

Name: bgpMiRed

AS: 65001

Router ID: [dropdown]

Redistribute Connected

Redistribute Static

Redistribute RIP

Redistribute OSPF

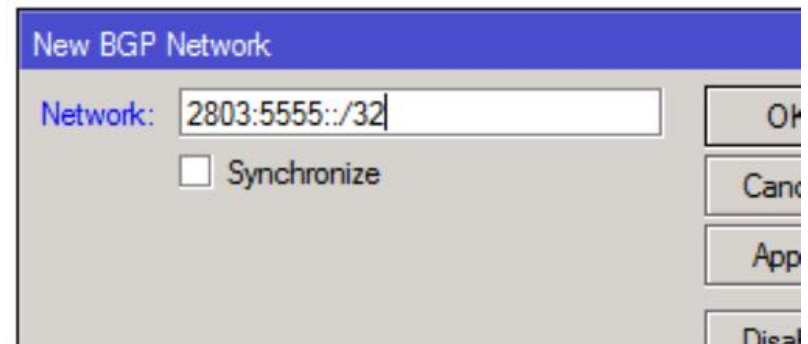
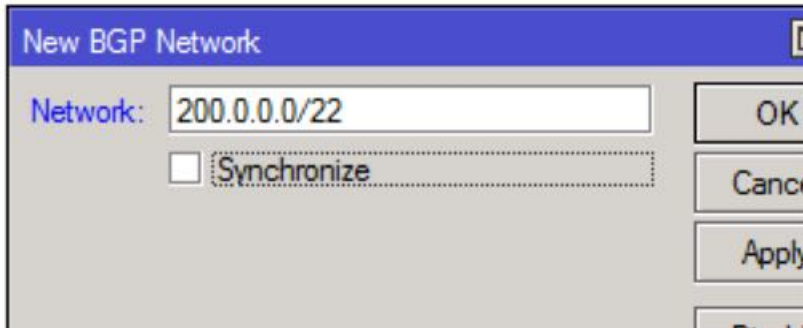
Redistribute Other BGP

Out Filter: [dropdown]

Confederation: [dropdown]

# Creando peers BGP4

- Definir la redes a anunciar
- Routing->BGP->Networks-> +
  - Requisito contar con IPv4 y/o IPv6



# Creando peers BGP4

- Crear un Peer BGP
- Routing->BGP->Peers -> +
  - Requisitos: ip y asn remoto (10.1.1.1/30 AS7094)

The screenshot shows the 'New BGP Peer' configuration window with the 'General' tab selected. The following fields are highlighted with red boxes:

- Name: peer\_ISP-A
- Instance: bgpMiRed
- Remote Address: 10.1.1.1
- Remote AS: 7094

Other visible fields include Remote Port, TCP MD5 Key, Nexthop Choice (set to default), and checkboxes for Multihop and Route Reflect.

The screenshot shows the 'New BGP Peer' configuration window with the 'Advanced' tab selected. The following fields are highlighted with red boxes:

- Address Families:  ip,  ipv6,  l2vpn,  vpn4
- Update Source: none
- Cisco VPLS NLRI Length Format: auto bits

# Creando peers BGP4

- Ejemplo de visualización de Peers BGP

Nombre	Remotos IP	ASN		Uptime	Prefix Count	Estado
 Pit_PCH-42	200.23.206.220	42	200.23.206.220	8d 11:36:53	127	established
 Pit_Microsoft2	200.23.206.211	8075	207.46.32.82	23d 06:58:36	212	established
 Pit_Microsoft1	200.23.206.210	8075	207.46.32.81	23d 06:58:33	212	established
 Pit_V6_Chile1	2801:14:9000::1	61522	200.23.206.1	88d 05:11:48	235	established
 Pit_V6_Chile2	2801:14:9000::2	61522	200.23.206.2	88d 05:11:41	235	established
 Pit_Google	200.23.206.225	15169	74.125.251.16	23d 06:57:47	579	established
 PitChile1	200.23.206.1	61522	200.23.206.1	18d 22:14:07	5180	established

# Creando peers BGP4

- Opciones “Peer BGP”
  - Opcion clave MD5
  - Max prefix limit
  - Max Pref. Restart T.
  - In & Out Filter

The screenshot shows a configuration window for a BGP peer. The 'General' tab is selected. The following fields are visible and highlighted with red boxes:

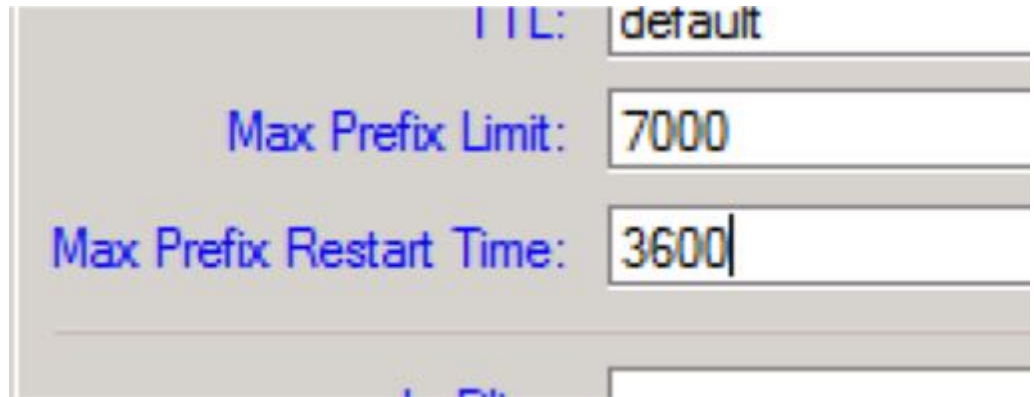
- TCP MD5 Key:** A text input field.
- Max Prefix Limit:** A text input field.
- Max Prefix Restart Time:** A text input field.
- In Filter:** A text input field.
- Out Filter:** A text input field.

Other visible fields include:

- Name:** peer\_ISP-A
- Instance:** bgpMiRed
- Remote Address:** 10.1.1.1
- Remote Port:** (empty)
- Remote AS:** 7094
- Nexthop Choice:** default
- Multihop
- Route Reflect
- Hold Time:** 180
- Keepalive Time:** (empty)
- TTL:** default
- AllowAS In:** (empty)

# Creando peers BGP4

- Opciones “Peer BGP”
  - Max prefix limit ~ por sobre un 20%
  - Max Pref. Restart T. (en segundos)  
0 -> (infinito)



The image shows a configuration window for a BGP peer. It contains several input fields. The 'Max Prefix Limit' field is set to 7000. The 'Max Prefix Restart Time' field is set to 3600. Other fields like 'TTL' are set to 'default'.

TTL:	default
Max Prefix Limit:	7000
Max Prefix Restart Time:	3600

(evitar recibir “escapes” de rutas)

# Filtros BGP: IN & OUT

- Diferencias filtro IN v/s OUT
  - Son diferentes a filtros de firewall
  - OUT: Que prefijos publico
  - IN: Que prefijos acepto/filtro





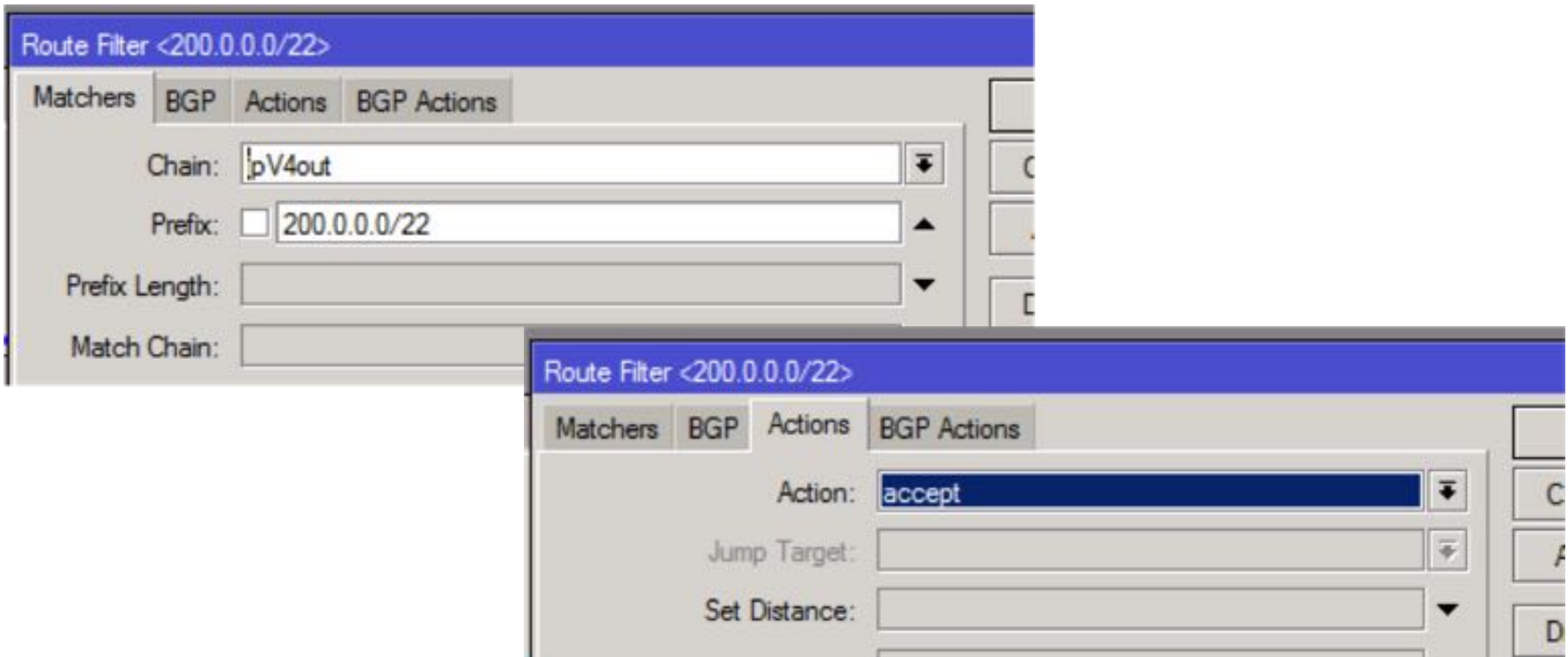
# Filtros BGP: IN & OUT

- Filtros OUT:
  - Solo publicar mis prefijos!!!
  - Publicar lo mas sumariado posible



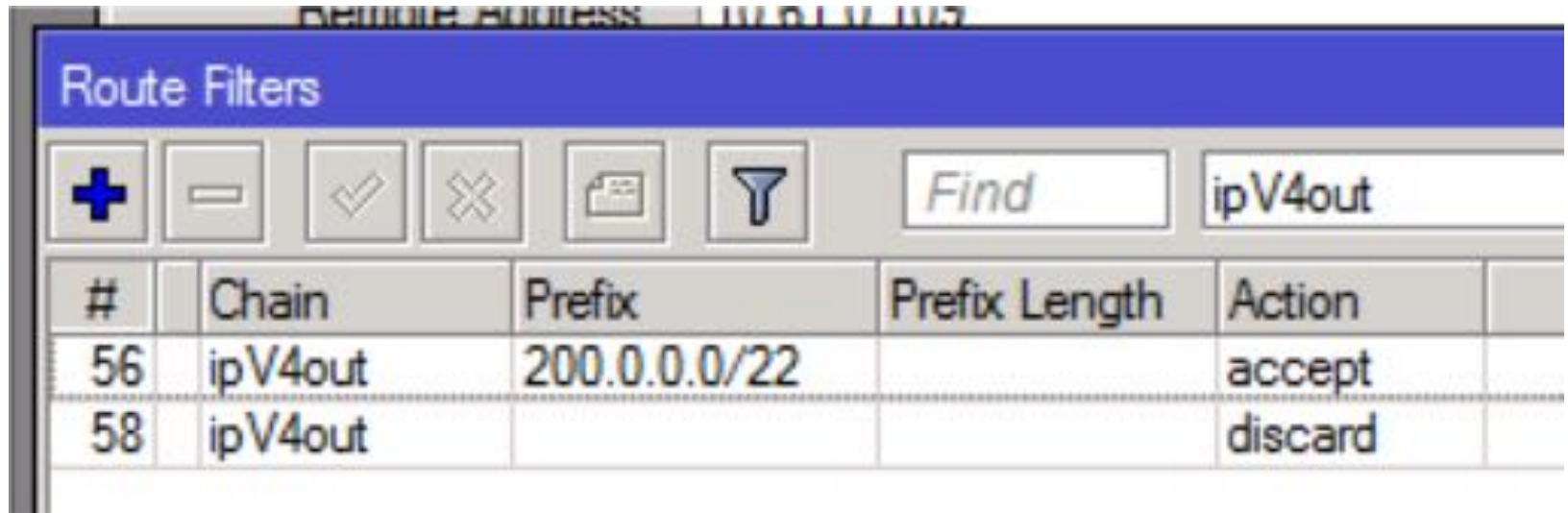
# Filtros BGP: IN & OUT

- Filtros OUT: (solo mis prefijos)
- Routing->Filters -> +



# Filtros BGP: IN & OUT

- Filtros OUT: (solo mis prefijos)
- Routing->Filters -> +



The screenshot shows the 'Route Filters' configuration window in Mikrotik WinBox. The window title is 'Route Filters'. Below the title bar, there are several control buttons: a plus sign (+), a minus sign (-), a checkmark (✓), a cross (✗), a folder icon, and a funnel icon. To the right of these buttons is a 'Find' search box containing the text 'ipV4out'. Below the controls is a table with the following data:

#	Chain	Prefix	Prefix Length	Action
56	ipV4out	200.0.0.0/22		accept
58	ipV4out			discard

# Filtros BGP: IN & OUT

- Filtros OUT: Aplicado al peer BGP.
- Routing->Filters -> +

General | Advanced | Status

Name: peer\_ISP-A

Instance: bgpMiRed

Remote Address: 10.1.1.1

Remote Port:

Remote AS: 7094

TCP MD5 Key:

Nexthop Choice: default

Multihop

Route Reflect

Hold Time: 180

Keepalive Time:

TTL: default

Max Prefix Limit:

Max Prefix Restart Time:

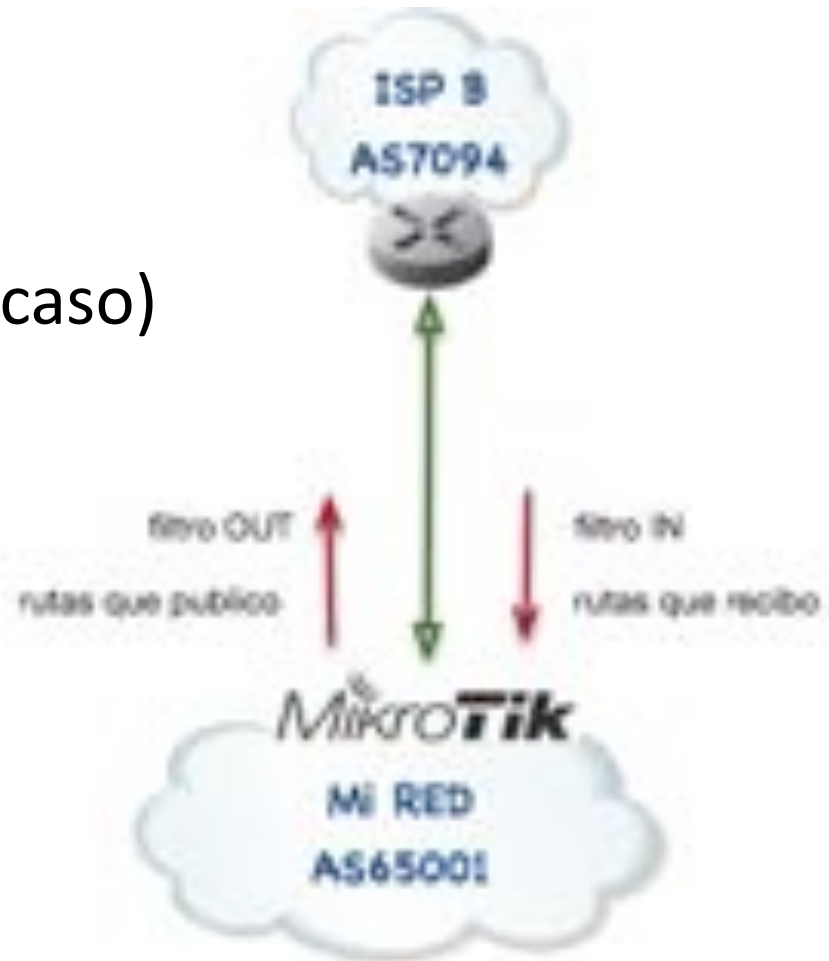
In Filter: |

Out Filter: ipV4out

AllowAS In:

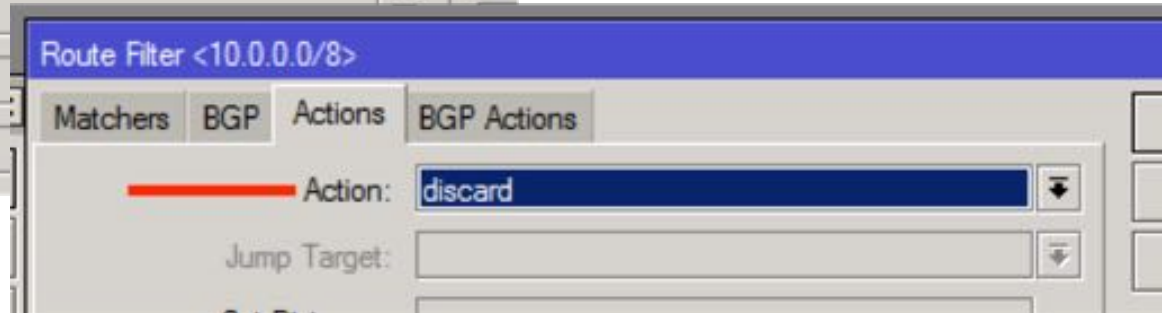
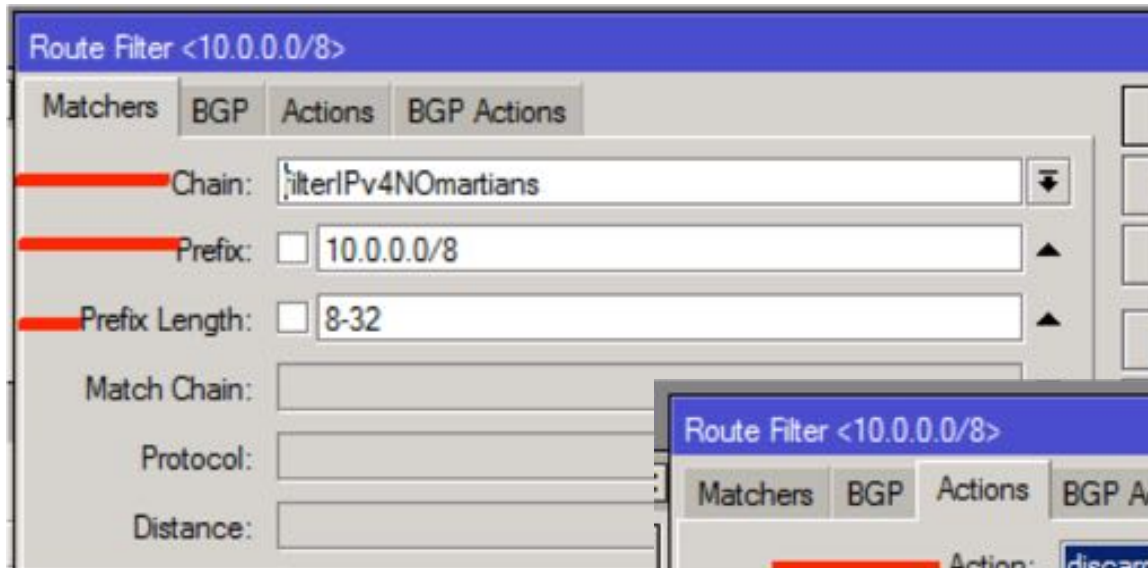
# Filtros BGP: IN & OUT

- Filtros IN:
  - Filtrar redes Marcianas!
  - Filtrar la red local
  - Filtrar ruta default (según caso)
  - Filtrar prefijos > /24



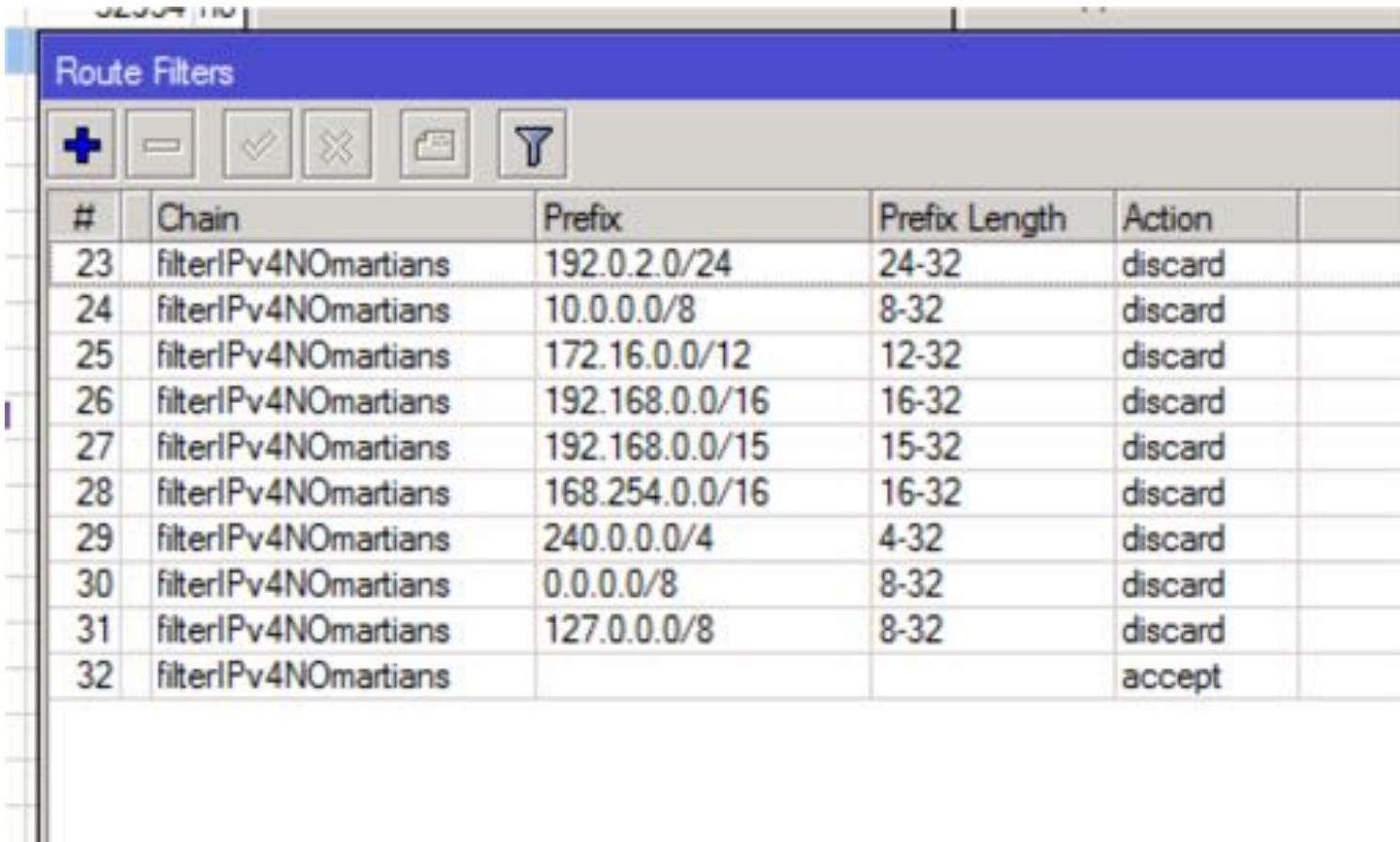
# Filtros BGP: IN & OUT

- Filtros IN:
- Routing->Filters -> +



# Filtros BGP: IN & OUT

- Filtros IN: (Martian Routes)

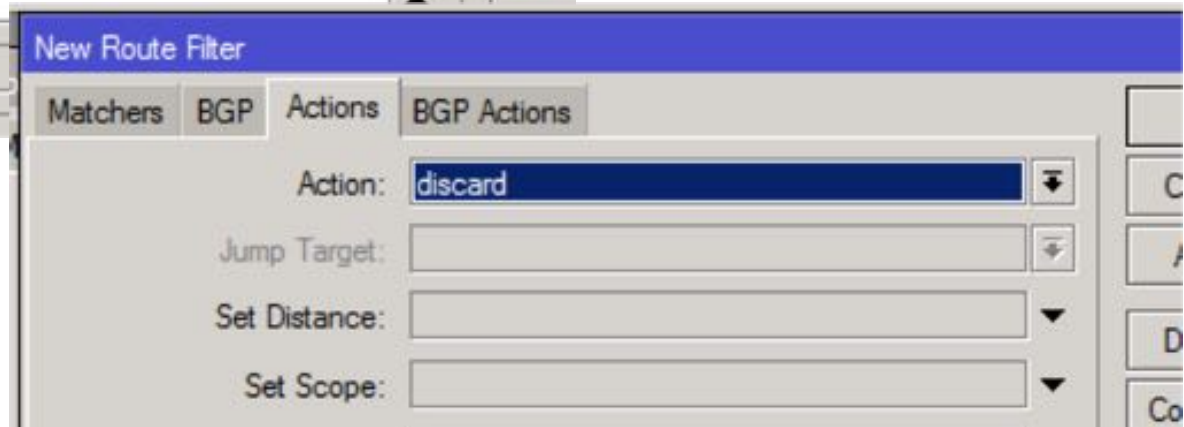
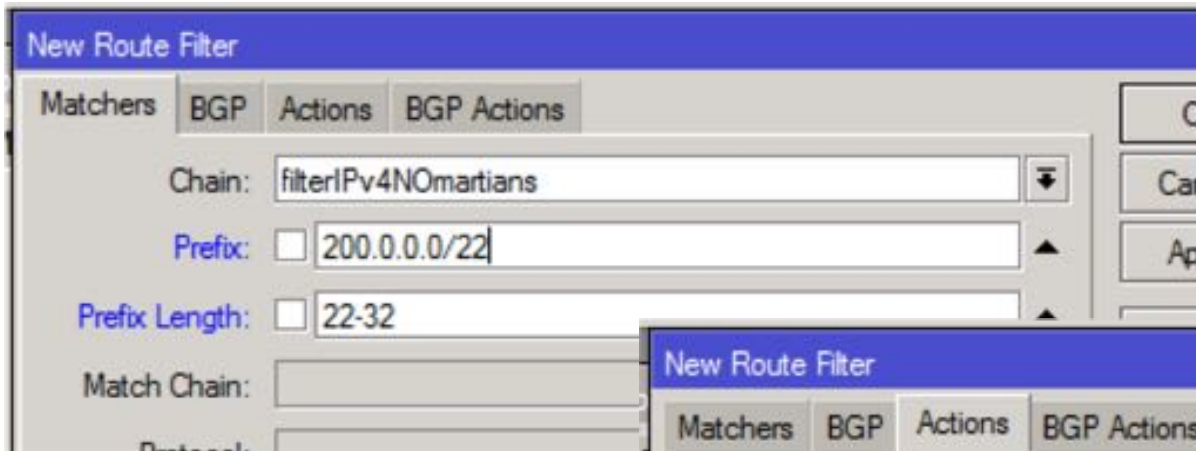


The screenshot shows the 'Route Filters' configuration window in Mikrotik WinBox. It contains a table with 6 columns: #, Chain, Prefix, Prefix Length, Action, and an empty column. The table lists 10 filters (rows 23-32) for the 'filterIPv4NOmartians' chain. Each filter from row 23 to 31 has an action of 'discard', while row 32 has an action of 'accept'. The filters target various IP ranges associated with Martian routes.

#	Chain	Prefix	Prefix Length	Action	
23	filterIPv4NOmartians	192.0.2.0/24	24-32	discard	
24	filterIPv4NOmartians	10.0.0.0/8	8-32	discard	
25	filterIPv4NOmartians	172.16.0.0/12	12-32	discard	
26	filterIPv4NOmartians	192.168.0.0/16	16-32	discard	
27	filterIPv4NOmartians	192.168.0.0/15	15-32	discard	
28	filterIPv4NOmartians	168.254.0.0/16	16-32	discard	
29	filterIPv4NOmartians	240.0.0.0/4	4-32	discard	
30	filterIPv4NOmartians	0.0.0.0/8	8-32	discard	
31	filterIPv4NOmartians	127.0.0.0/8	8-32	discard	
32	filterIPv4NOmartians			accept	

# Filtros BGP: IN & OUT

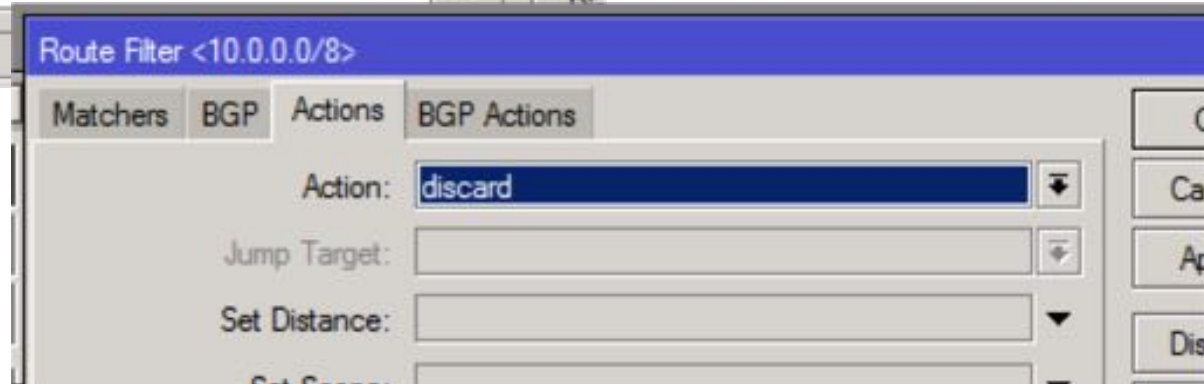
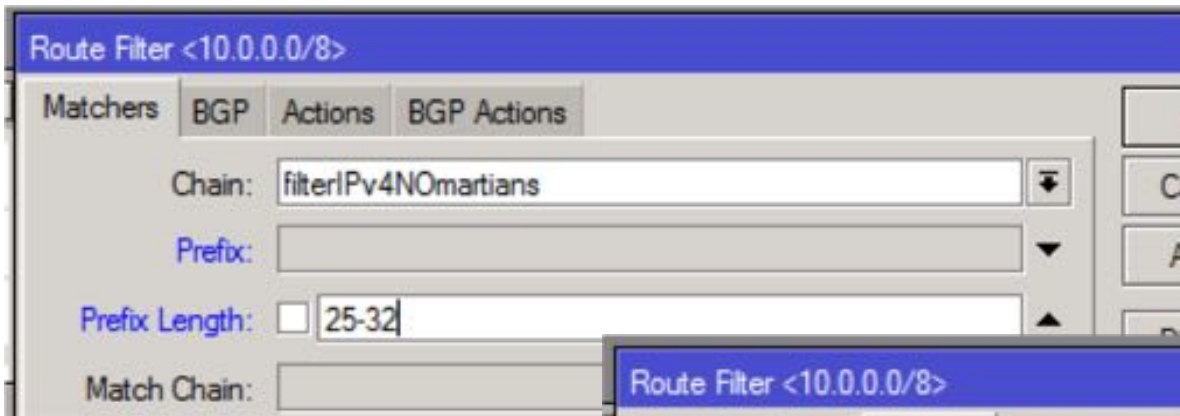
- Filtros IN: (filtrar la red local)
- Routing->Filters -> +





# Filtros BGP: IN & OUT

- Filtros IN: (filtrar prefijos > /24)
- Routing->Filters -> +



# Filtros BGP: IN & OUT

- Filtros IN: Aplicado al peer BGP.
- Routing->Filters -> +

General | Advanced | Status

Name: peer\_ISP-A

Instance: bgpMiRed

Remote Address: 10.1.1.1

Remote Port:

Remote AS: 7094

TCP MD5 Key:

Nexthop Choice: default

Multihop

Route Reflect

Hold Time: 180

Keepalive Time:

TTL: default

Max Prefix Limit:

Max Prefix Restart Time:

In Filter: filterIPv4NoMartians

Out Filter:

AllowAS In:

- Introducción
- creando peers BGP4
- Filtros BGP (in & out)
- **Usando Comunidades BGP**
- Filtros de seguridad mínimos
- Tips para Layer2
- Q & A

# Comunidades BGP

- Que son las comunidades BGP
  - Son un atributo BGP y es un atributo transitivo opcional de longitud variable.
  - Las Comunidades se traspasan de Router a Router, aunque el Router no las “entienda”.
  - A cada prefijo anunciado se le pueden incorporar comunidades.

# Comunidades BGP

- Para que sirven las comunidades BGP?
  - Para realizar Ingeniería de tráfico (o TE)
  - Para identificar un grupo de prefijos con una o mas propiedades en común
  - El o los router upstream pueden ser influidos en su proceso de decisión de rutas.

# Comunidades BGP

- Como ver las comunidades BGP en Mikrotik
  - IP->Routes->Detail      Ejemplo: 61522:65023

The screenshot shows the Mikrotik WinBox interface. On the left, a list of routes is displayed with columns for 'Db' and 'DAb'. The route '200.29.16.0/24' is selected. The right pane shows the 'Route <200.29.16.0/24>' configuration. The 'General' tab is active, showing the following BGP attributes:

Attribute	Value
BGP AS Path:	61522,10778
BGP Weight:	
BGP Local Pref.:	
BGP Prepend:	
BGP MED:	
BGP Atomic Aggregate:	
BGP Origin:	igp
BGP Communities:	61522:65023

# Comunidades BGP

- Como se aplican las Comunidades en Mikrotik

The screenshot displays the Mikrotik WinBox interface for configuring Route Filters. The main window shows a table of filters:

#	Chain	Prefix	Prefix Length	Action
56	ipV4out	200.0.0.0/22		accept
58	ipV4out			discard

A secondary window titled "Route Filter <200.0.0.0/22>" is open, showing the configuration for the filter. The "BGP" tab is selected, and the "Set BGP Communities" field is highlighted with a red line, showing the value "15169:13000".

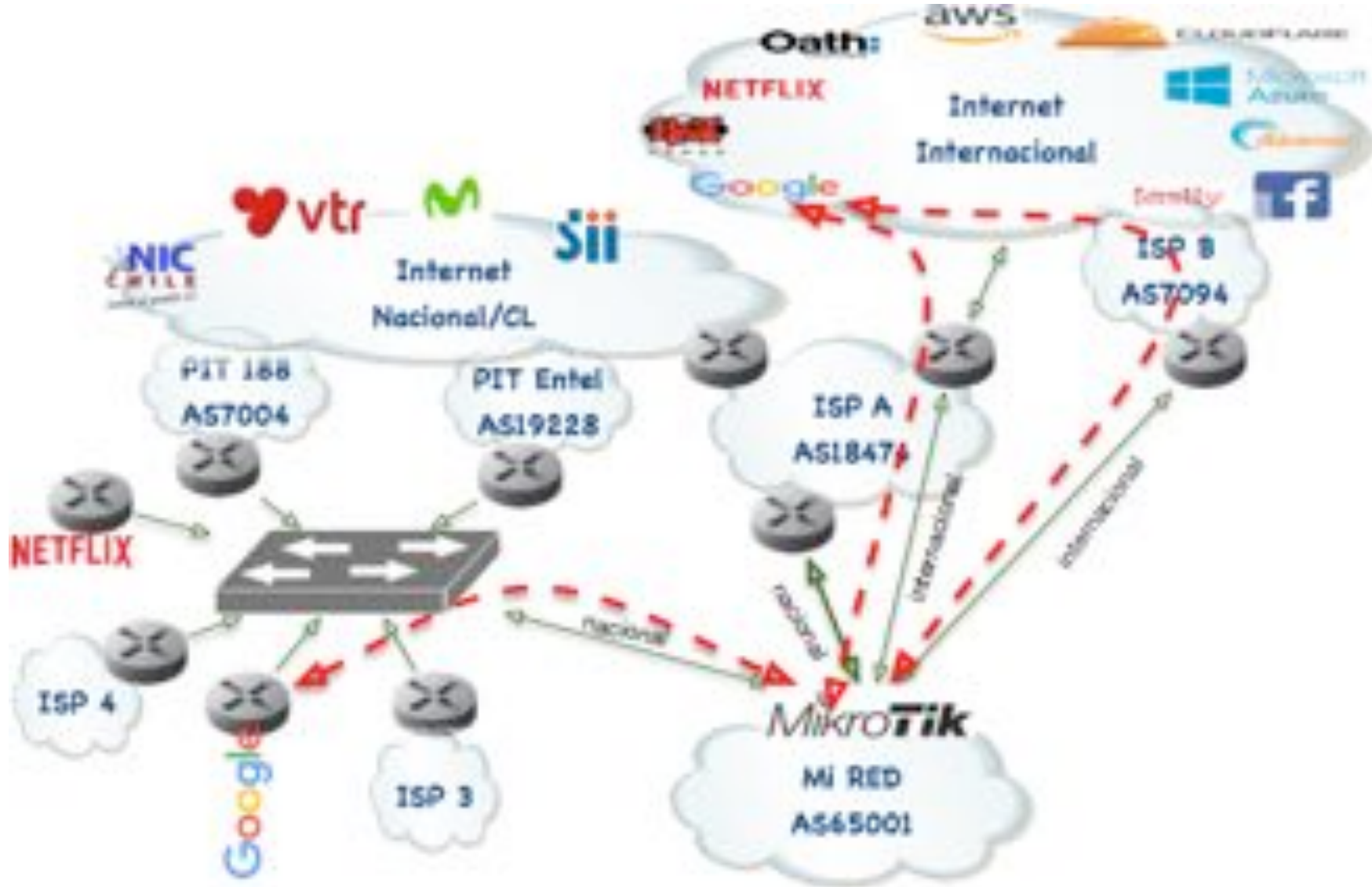
# Comunidades BGP

- Ejemplo Peering con Google (AS15169):

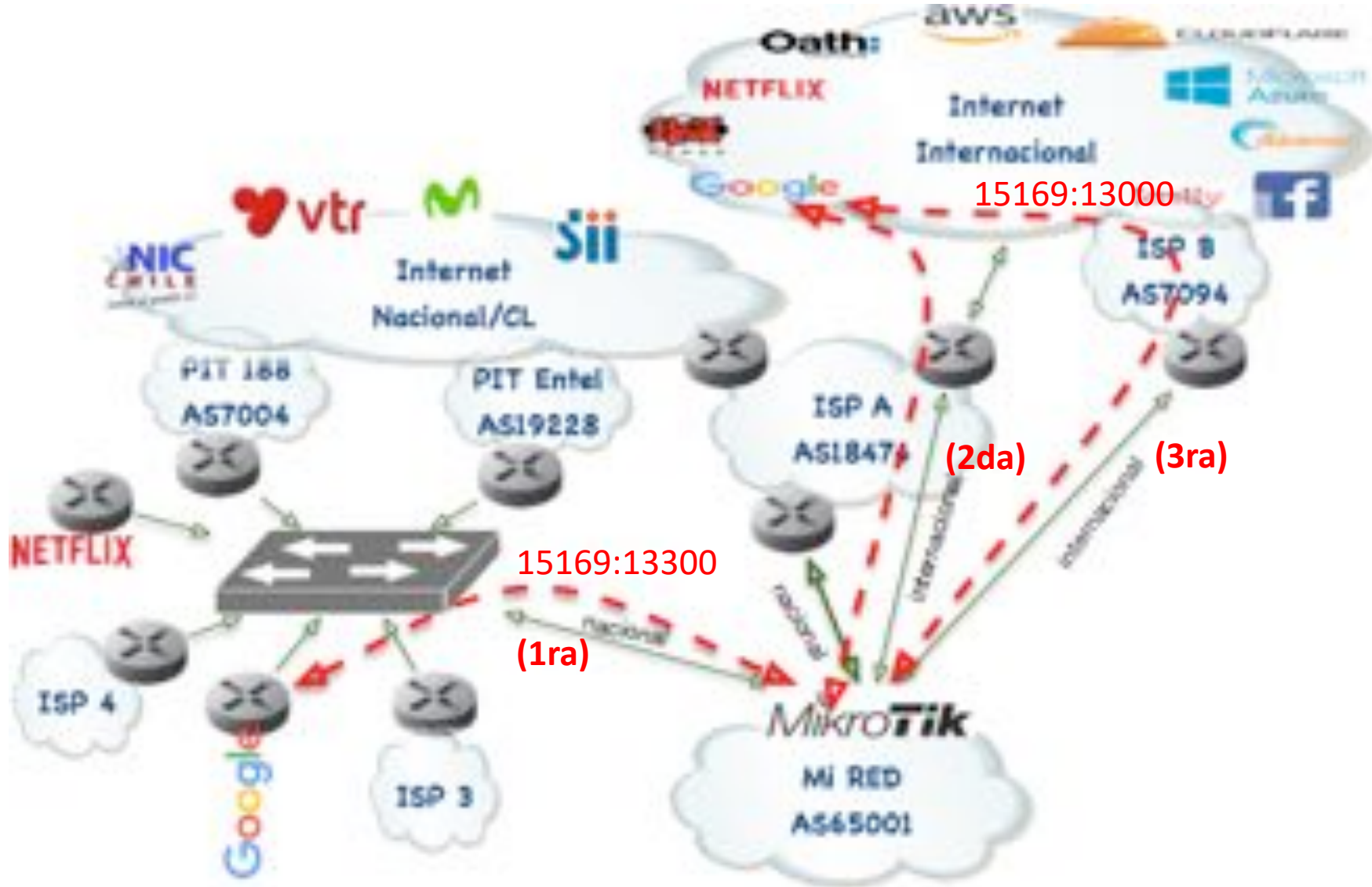
Community	Preferred Ingress Signalling Range
15169:13000	Lowest preference to receive traffic for this prefix at this interconnection point (try to not serve traffic here). Attempt to serve traffic on an indirect path (through other upstreams or peers) before using this prefix.
15169:13100	Default priority of traffic on an indirect path. Tagging with this community indicates that the preference is equal to receiving traffic over an indirect path.
15169:13200	Default priority to receive traffic for this prefix at this interconnection point (the same as if the prefix is untagged).
15169:13300	Highest preference to receive traffic for this prefix at this interconnection point (try to serve traffic here).



# Comunidades BGP



# Comunidades BGP



# Comunidades BGP

- Ejemplo Peering con Netflix (AS2906):

<https://www.netflix.com>

– AWS

- Streaming
  - OCA



# Comunidades BGP

- Algoritmo de selección de OCA mas cercano
  - (1) Prefix (preferencia prefijo mas especifico)  
Un prefijo /24 será preferido sobre un /22
  - (2) AS Path (contabiliza AS Path diferentes)  
A prefijos iguales prefiere menor AS Path
  - (3) MED (prefiere el menor MED)  
Multi-exit Discriminator

# Comunidades BGP

- Como verificar desde cual OCA recibo una pelicula?

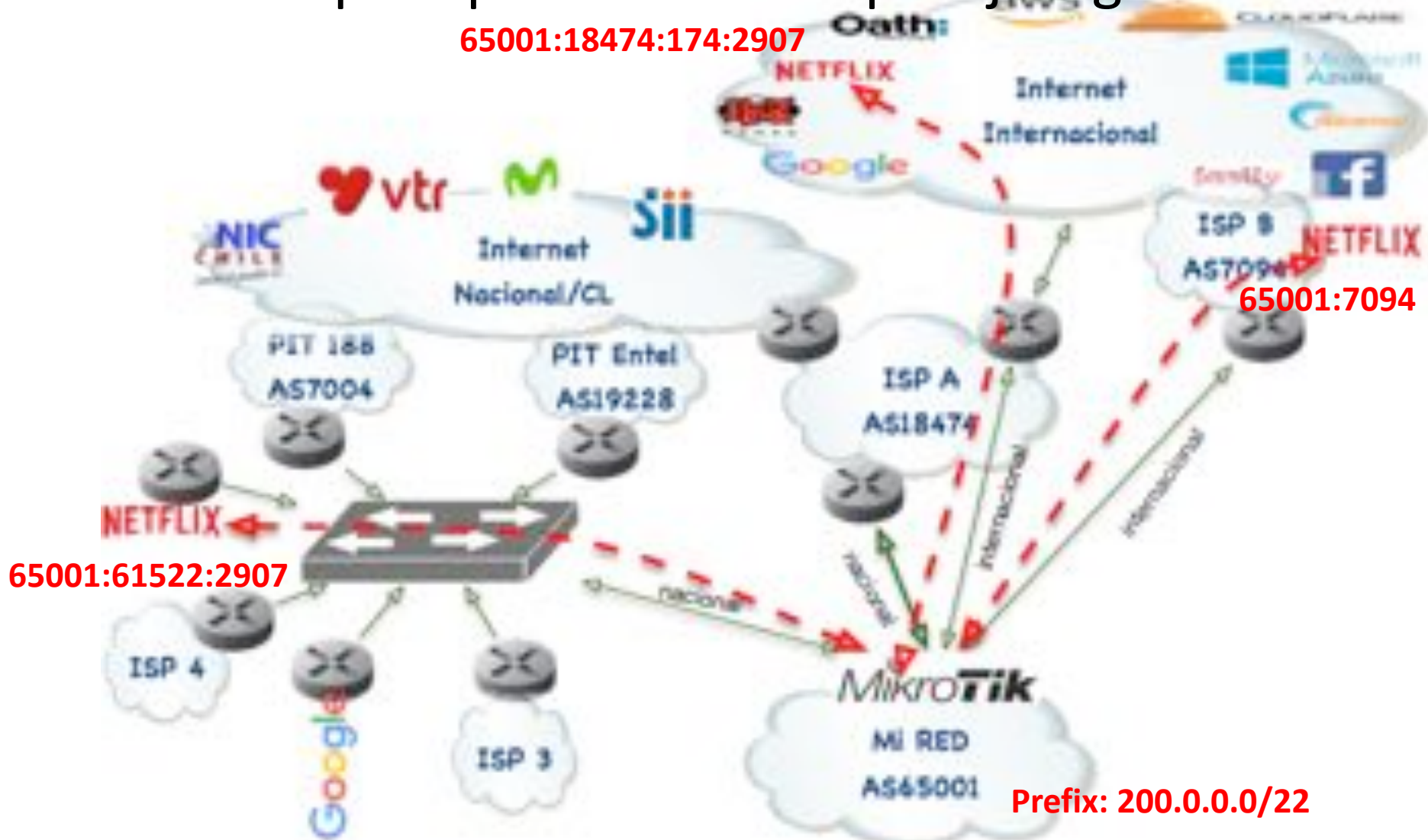
Shift + Ctrl + Alt + Cmd + D



Buffering bitrate (a/v): 96 / 846  
Buffer size in Bytes (a/v): 3167429 / 35612411  
Buffer size in Bytes: 38779840  
Buffer size in Seconds (a/v): 249.604 / 244.672  
**Current CDN (a/v): c001.sdc002.telefonica-cl.isp.nflxvideo.net, I I: 50151 / c001.sdc0**  
Audio Track: en, Id: A:1:1;2;en;1;, Channels: 2.0, Codec: audio/mp4;codecs=mp4a.4  
Video Track: Codec: video/mp4;codecs=vp09.00.11.08.02 (vp9)  
Timed Text Track: es, Profile: dfxp-ls-sdh, Id: T:1:0;1;es;0;0;  
Framerate: 23.977  
Current Dropped Frames: 0

## Comunidades BGP

- Calculo para publicación de prefijos iguales



# Comunidades BGP

- Como preferir el OCA en PIT Chile?
  - Utilizaremos las Comunidades disponibles en PIT Chile:
    - 0:61522 -> No anunciar este prefijo a ningún miembro del IXP.
    - 61522:{ASN} -> Anunciar este prefijo solo al peer ASN: {ASN}

# Comunidades BGP

- Como preferir el OCA en PIT Chile?
  - Publicar 200.0.0.0/22 a todos los Peer.
  - Publicar 200.0.0.0/22 a PitChile con la comunidad:
    - 0:2906 (No anunciar al peer AS2906 (Netflix))
  - Publicar solo a PitChile 200.0.0.0/23 y 200.0.2.0/23 con las comunidades:
    - 0:61522 (No anunciar a ningún miembro de PitChile)
    - 61522:2906 (SI anunciar al peer 2906 (Netflix))



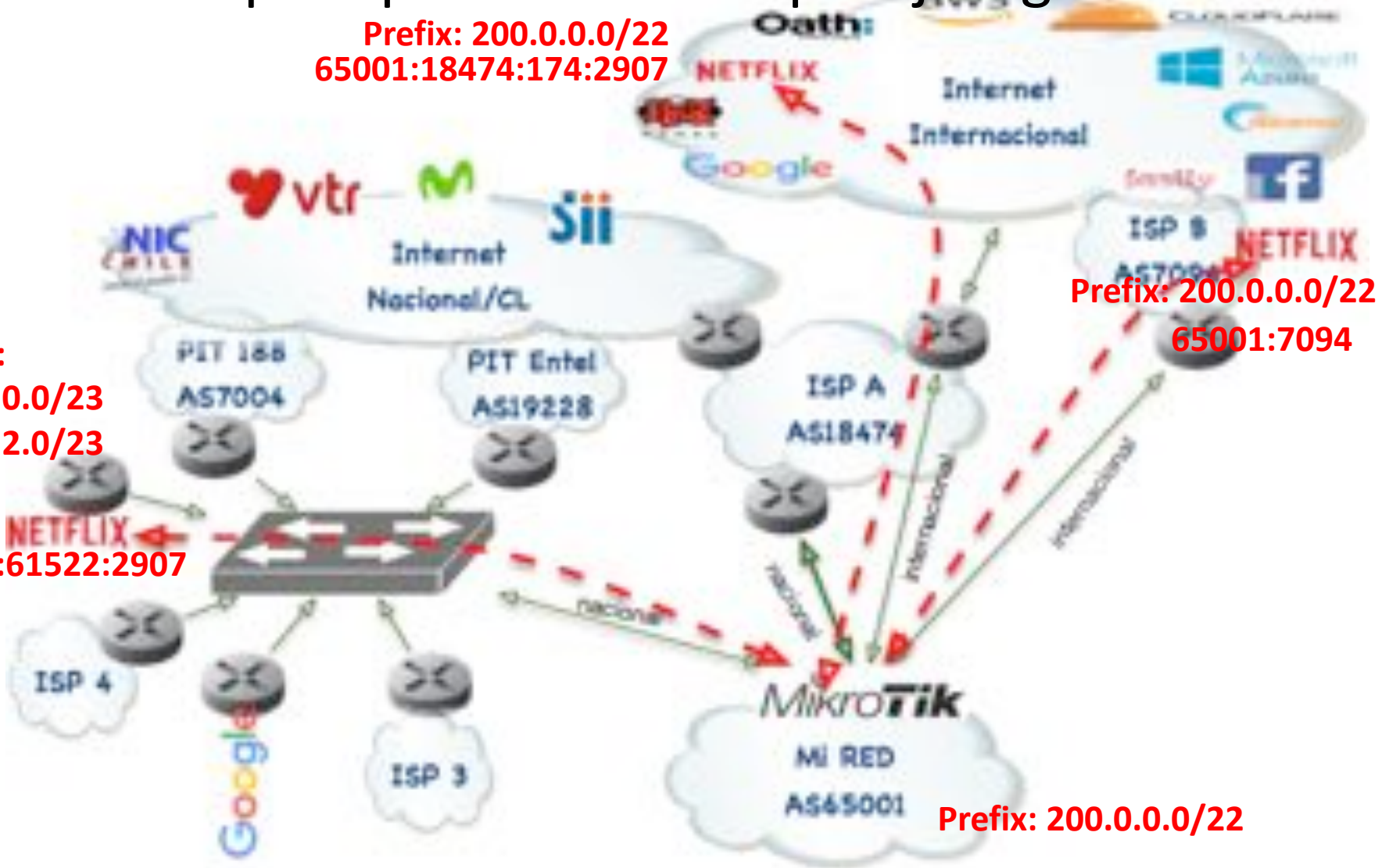
## Comunidades BGP

- Calculo para publicación de prefijos iguales

Prefix: 200.0.0.0/22  
65001:18474:174:2907

Prefix:  
200.0.0.0/23  
200.0.2.0/23

Prefix:  
65001:61522:2907

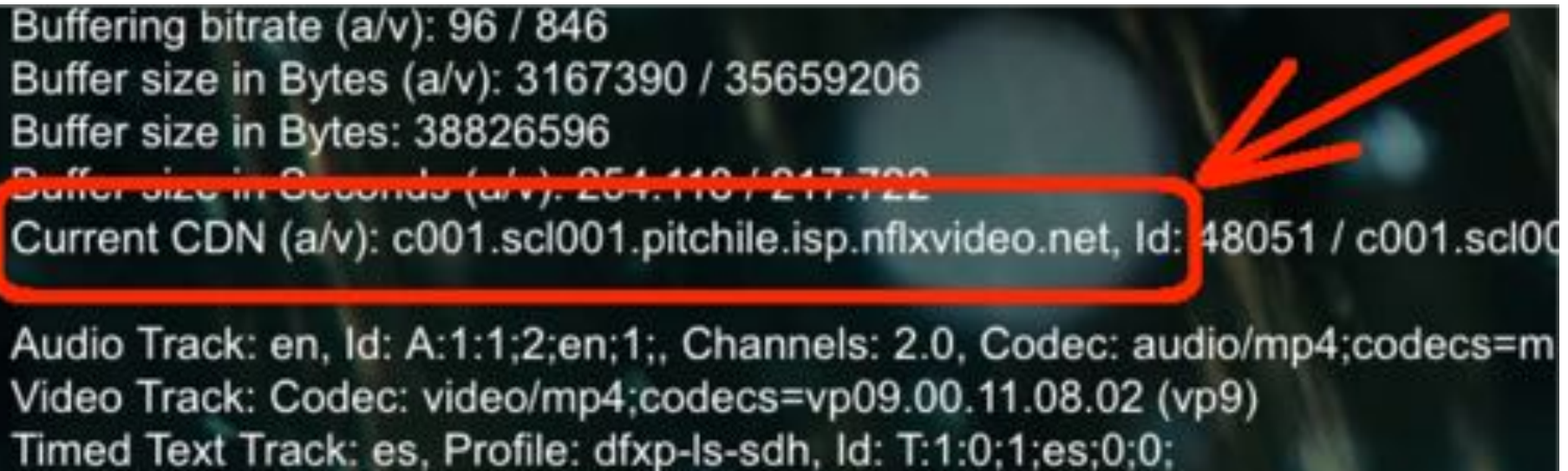


Prefix: 200.0.0.0/22

# Comunidades BGP

- Como verificar desde cual OCA recibo una pelicula?

Shift + Ctrl + Alt + Cmd + D



Buffering bitrate (a/v): 96 / 846  
Buffer size in Bytes (a/v): 3167390 / 35659206  
Buffer size in Bytes: 38826596  
Buffer size in Seconds (a/v): 254.110 / 217.722  
**Current CDN (a/v): c001.scl001.pitchile.isp.nflxvideo.net, Id: 48051 / c001.scl00**  
Audio Track: en, Id: A:1:1;2;en;1;, Channels: 2.0, Codec: audio/mp4;codecs=m  
Video Track: Codec: video/mp4;codecs=vp09.00.11.08.02 (vp9)  
Timed Text Track: es, Profile: dfxp-ls-sdh, Id: T:1:0;1;es;0;0;

- Introducción
- creando peers BGP4
- Filtros BGP (in & out)
- Usando Comunidades BGP
- **Filtros de seguridad mínimos**
- Tips para Layer2
- Q & A

# Minimo Security Filters

- Filtros de paquetes mínimos de seguridad:
  - Martian Routes filter
    - Using Address List
  - Spoofing IP Filter (IN & OUT)
    - (<https://www.caida.org/projects/spoofers> )
  - Transit IP Filters
  - Permitir puerto BGP tcp/179 Solo desde Peers

# Minimo Security Filters

- Martian Routes Filter:

The screenshot shows the Mikrotik WinBox Firewall Filter Rules configuration window. The 'Filter Rules' tab is active, and a search filter 'Martian-routes' is applied. The table below lists the configured rules:

Name	Address	Time...	Creation Time
Martian-routes	10.0.0.0/8		Nov/13/2018 22:11:16
Martian-routes	172.16.0.0/12		Nov/13/2018 22:11:16
Martian-routes	192.168.0.0/16		Nov/13/2018 22:11:16
Martian-routes	127.0.0.0/8		Nov/13/2018 22:11:16
Martian-routes	0.0.0.0/8		Nov/13/2018 22:11:16

The screenshot shows the Mikrotik WinBox Firewall Rule configuration window, General tab. The 'Dst. Address List' is set to 'Martian-routes'.

Src. Address List:

Dst. Address List:  Martian-routes

Layer7 Protocol:

Content:

Connection Rule:

The screenshot shows the Mikrotik WinBox Firewall Rule configuration window, Action tab. The 'Action' is set to 'reject'.

Action: reject

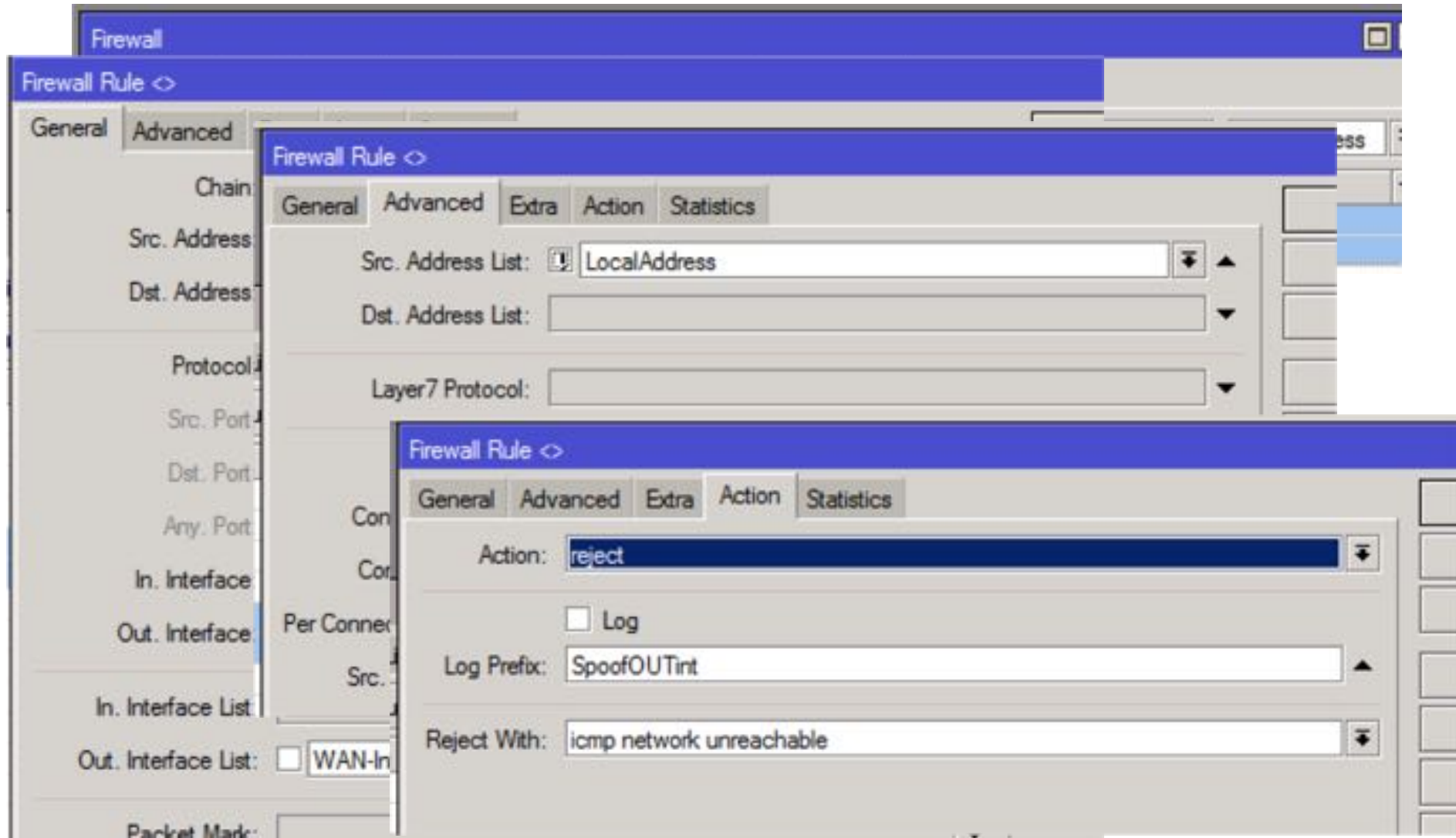
Log

Log Prefix:

Reject With: icmp network unreachable

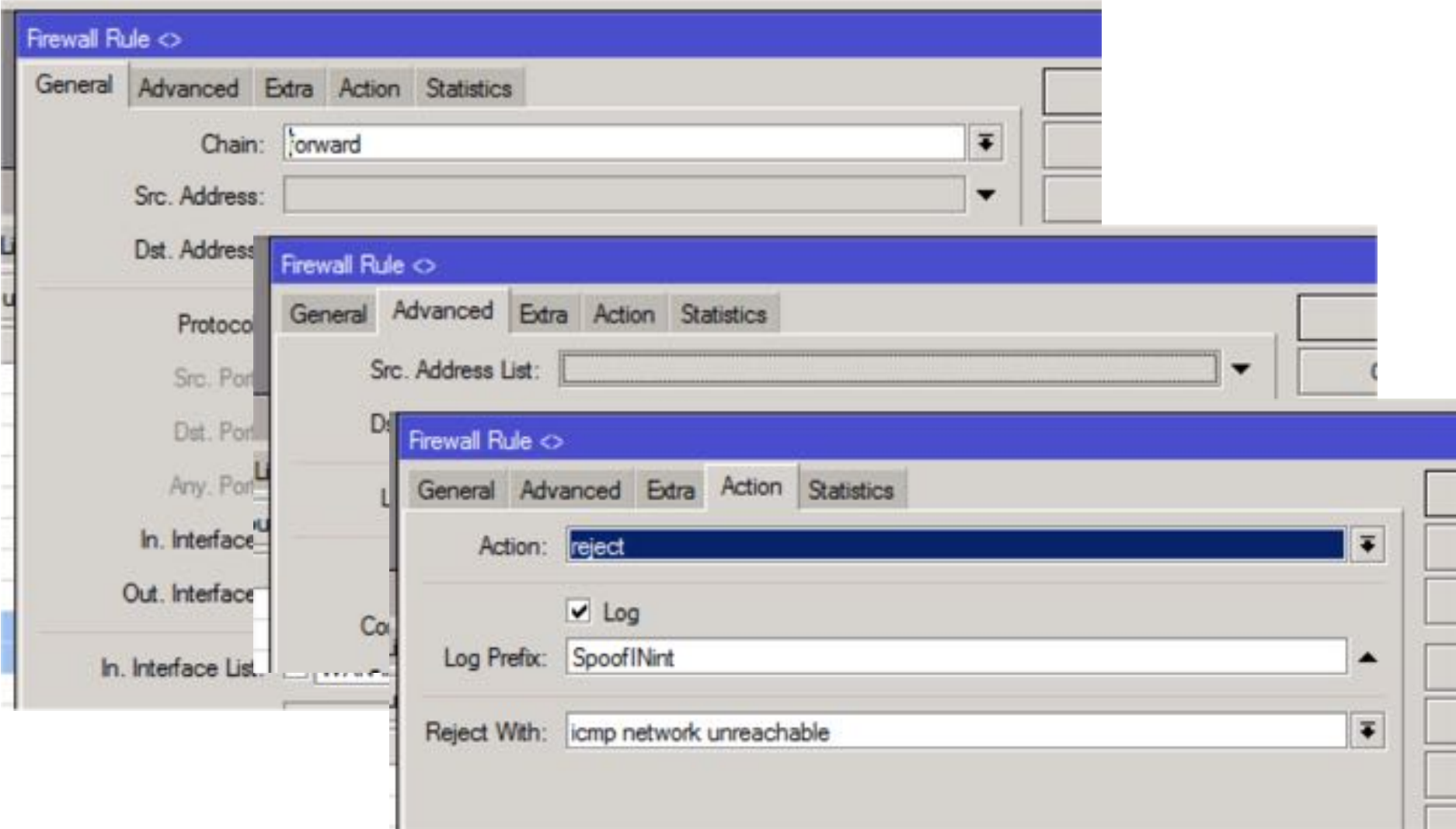
# Minimo Security Filters

- Spoofing IP Filter (IN & OUT)



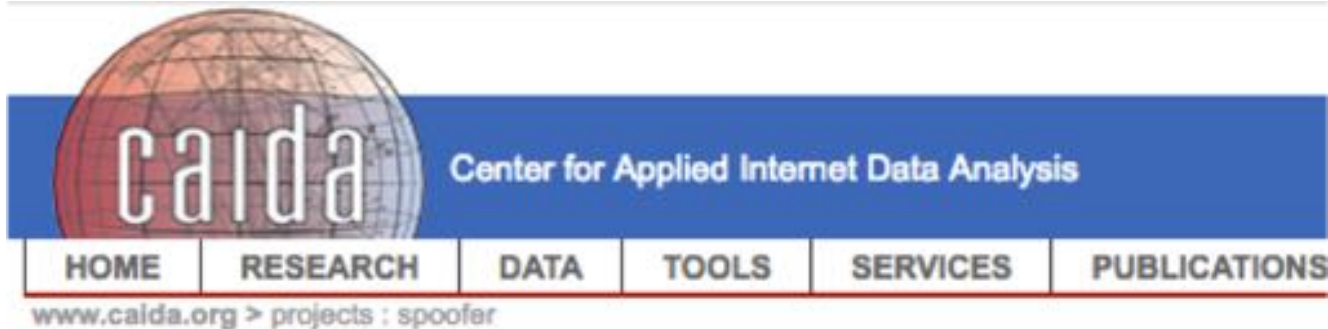
# Minimo Security Filters

- Spoofing IP Filter (IN & OUT)



# Minimo Security Filters

- Spoofing IP Filter (IN & OUT)
  - Como verificar mi red?



## Spoofing

Seeking to minimize Internet's susceptibility to spoofed DDoS attacks, we are developing validation (SAV) best anti-spoofing practices. This project includes applied research, interactive analysis and reporting service.

We have developed and support a new client-server system for Windows, MacOS, and Linux. We are (in the process of) producing reports on source IP addresses (spoofed packets). We are (in the process of) producing reports on source IP addresses (spoofed packets).

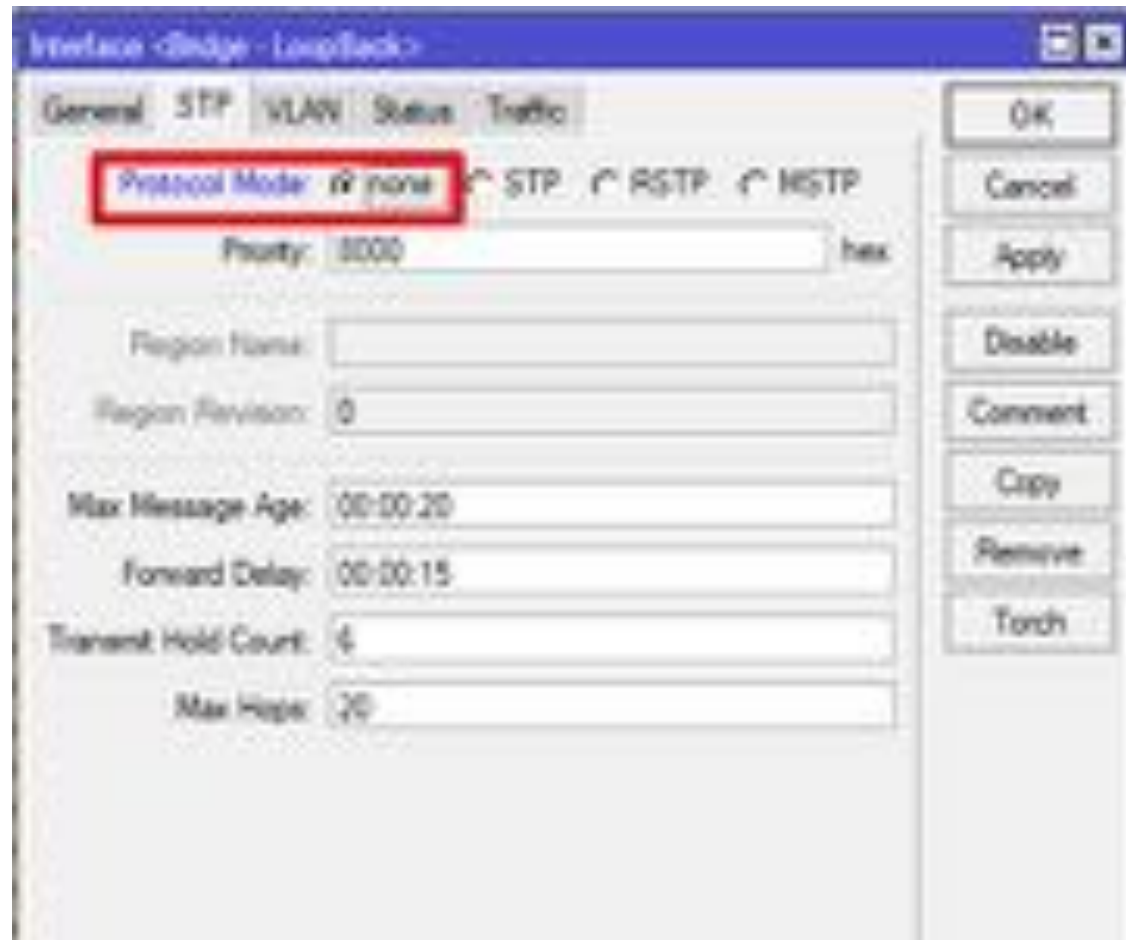
- <https://www.caida.org/projects/spoofing>



- Introducción
- creando peers BGP4
- Filtros BGP (in & out)
- Usando Comunidades BGP
- Filtros de seguridad mínimos
- **Tips para Layer2**
- Q & A

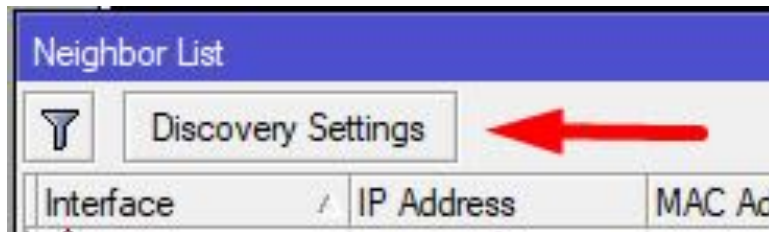
# Tips Layer2 IXP

- Mejorando la seguridad de MKT en el borde:
  - Desactivando STP (Spanning tree protocol) si usas bridge



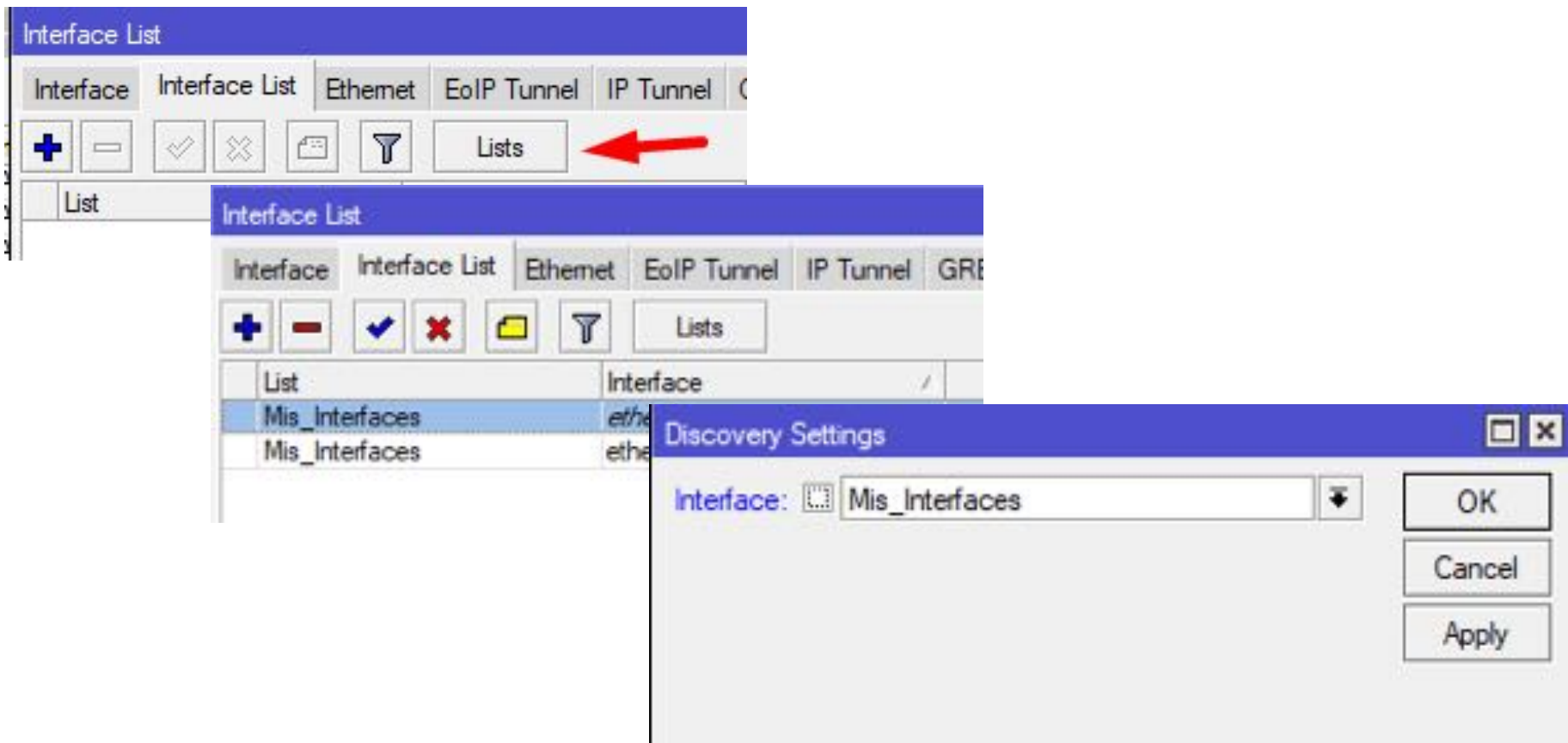
# Tips Layer2 IXP

- MNDP (Mikrotik Neighbor Discovery Protocol)
  - IP -> Neighbors



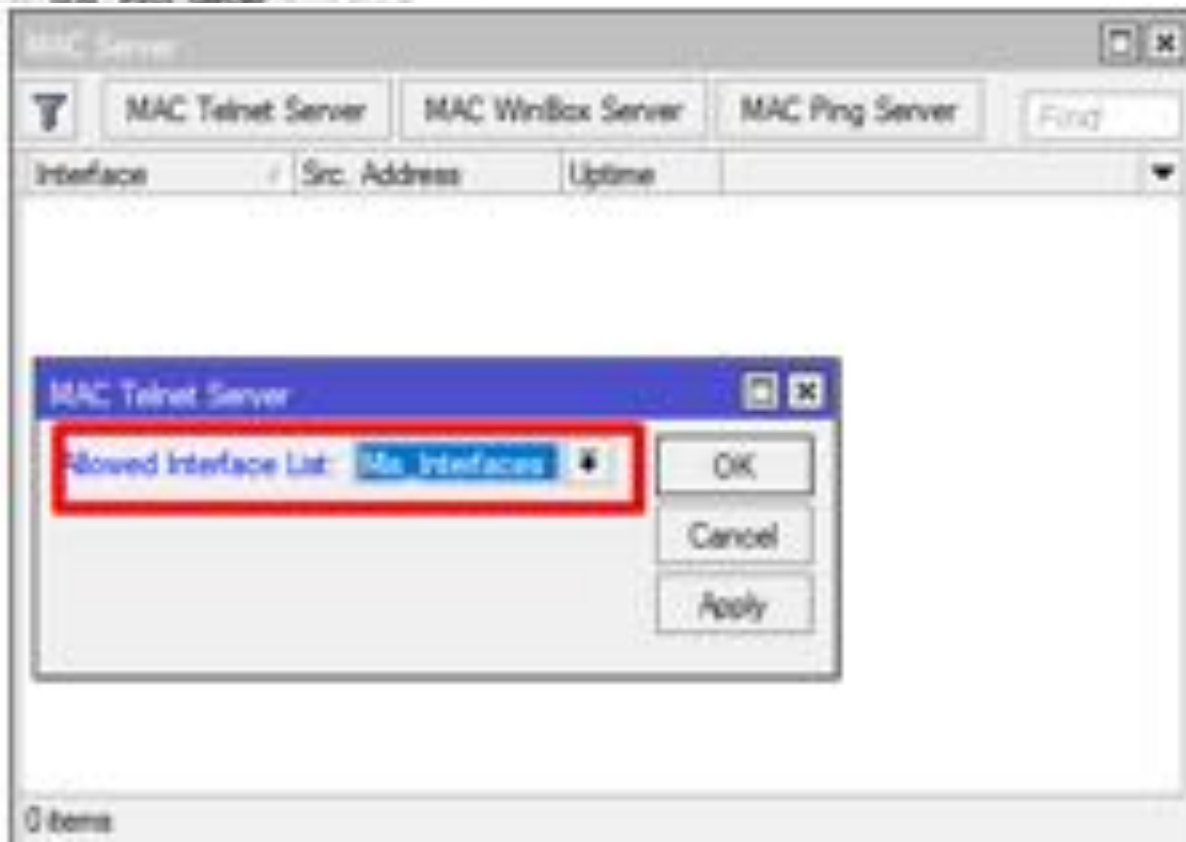
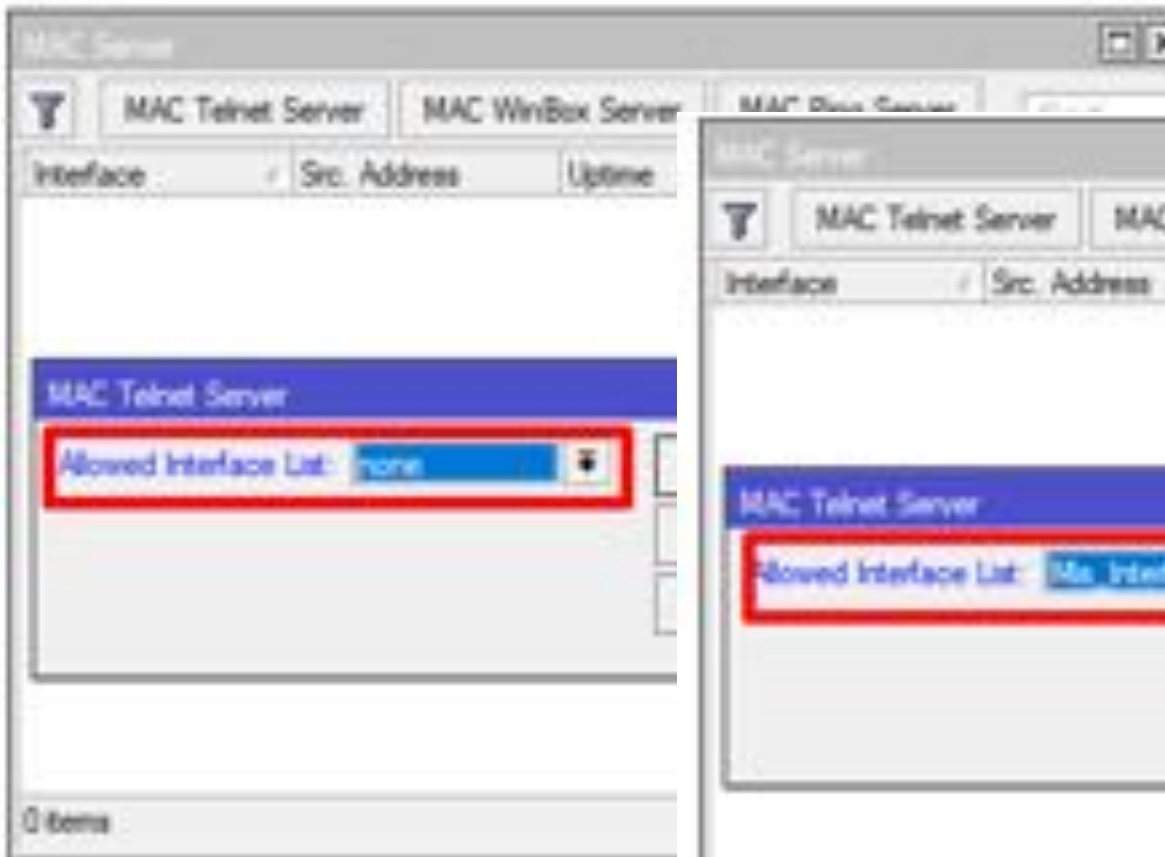
# Tips Layer2 IXP

- MNDP (Mikrotik Neighbor Discovery Protocol)
  - Solo dejarlo para una interfaz particular?



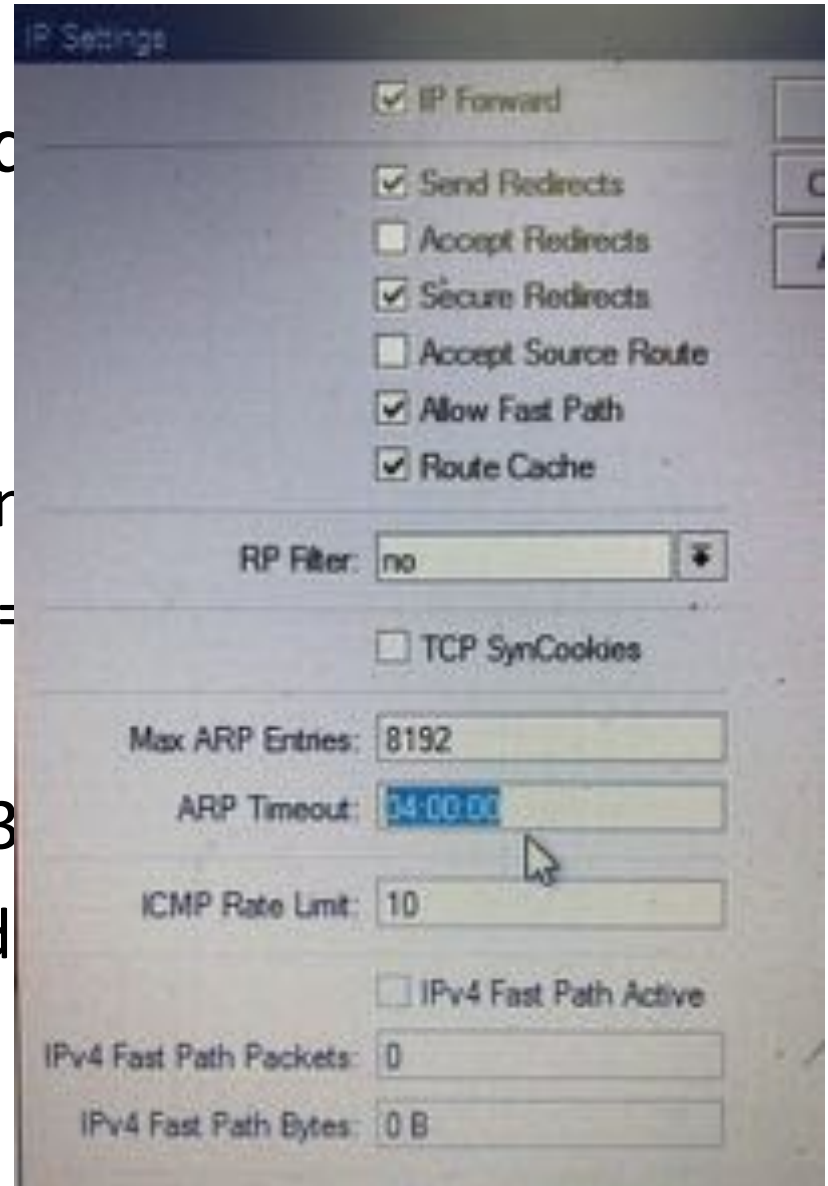
## Tips Layer2 IXP

- Mejorando la seguridad de MKT en el borde:
- MAC telnet server // MAC winbox server



# Tips Layer2 IXP

- Mejorando la seguridad de
- RTP Timeout
  - Mikrotik Default: 30 segundos
  - `/ip settings set arp-timeout=4h`
  - Por interfaz (versión  $\geq 6.3$ )
  - `/interface ethernet set [find timeout=4h`



# Agenda

- Introducción
- creando peers BGP4
- Filtros BGP (in & out)
- Usando Comunidades BGP
- Filtros de seguridad mínimos
- Tips para Layer2
- **Q & A**

Q & A



# Tips & Tricks to connect to a Layer2 IXP

MUM Chile 2019

