

The Brothers WISP



Route it like it's **HOT**

Greg Sowell Consulting

Using BGP For QoS

MUM 2015

Who Am I

- ▶ Greg Sowell – A+, Network+, CCNA, CCNP, CCIE Written, MTCNA, MTCRE, MTCINE, Mikrotik Certified Trainer
- ▶ Director of Technology FIBERTOWN Datacenters
- ▶ Consultant – GregSowell.com

The Brothers WISP

TheBrothersWISP.com



Greg and Andrew Cox



Justin



JJ



Tomas



Tom

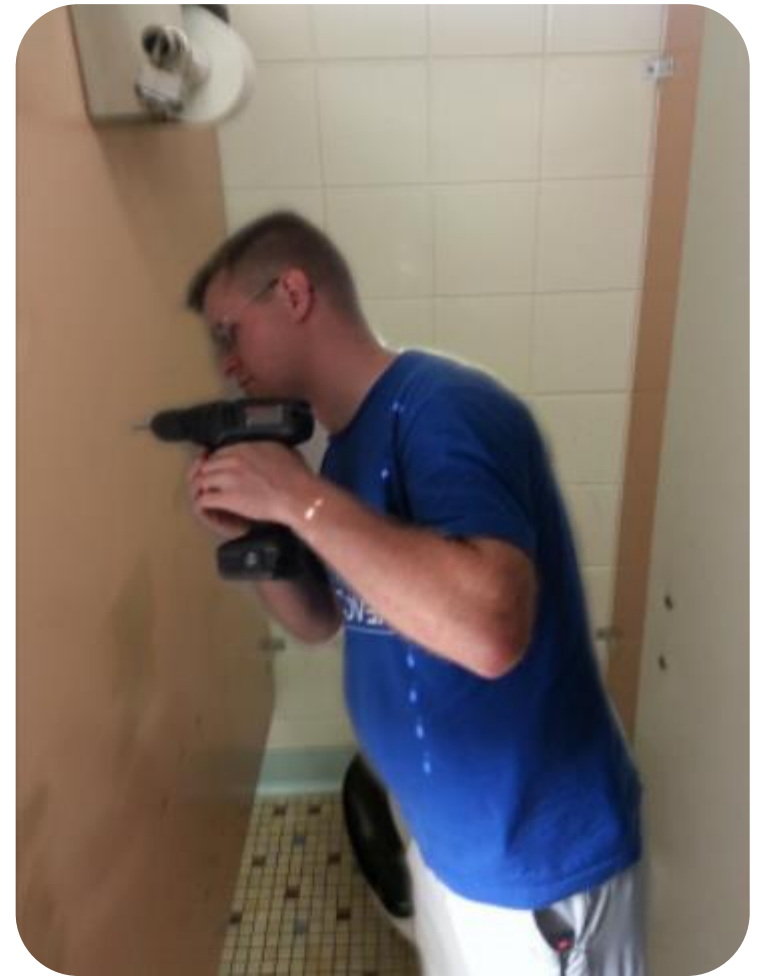
Fabio
Thrift



The Brothers WISP

TheBrothersWISP.com

And then there is Mike



Assumptions

- ▶ You are somewhat familiar with BGP and its configuration.
 - ▶ ASN.
 - ▶ Peers.
 - ▶ Filters.
- ▶ You are somewhat familiar with QoS and its configuration.
 - ▶ Address-lists.
 - ▶ Mangle rules.
 - ▶ Queue trees.
- ▶ But we will review a little anyway as we go along...

What is BGP

- ▶ Border Gateway Protocol is the dynamic routing protocol that carries all of the routing information for the Internet...no big deal.
- ▶ You can choose to accept all routes, partial routes, default route, or some combination of them all.
 - ▶ You can also filter on your side to be even more selective.
 - ▶ We will be accepting all routes /23 or larger.

BGP

Instances VRFs Peers Networks Aggregates VPN4 Routes Advertisements

+ - ✓ ✗ 📄 🔍 Refresh Refresh All Resend Resend All

Name	Instance	Rem...	Re...	M...	R...	TTL	Remot...	Uptime	Prefix Co...	State
peer-ISP	default	216...	19...	no	no	d...	209.1...	05:23:24	242849	established
peer-OIX	default	172...	65...	no	no	d...	74.19...	05:23:23	1	established

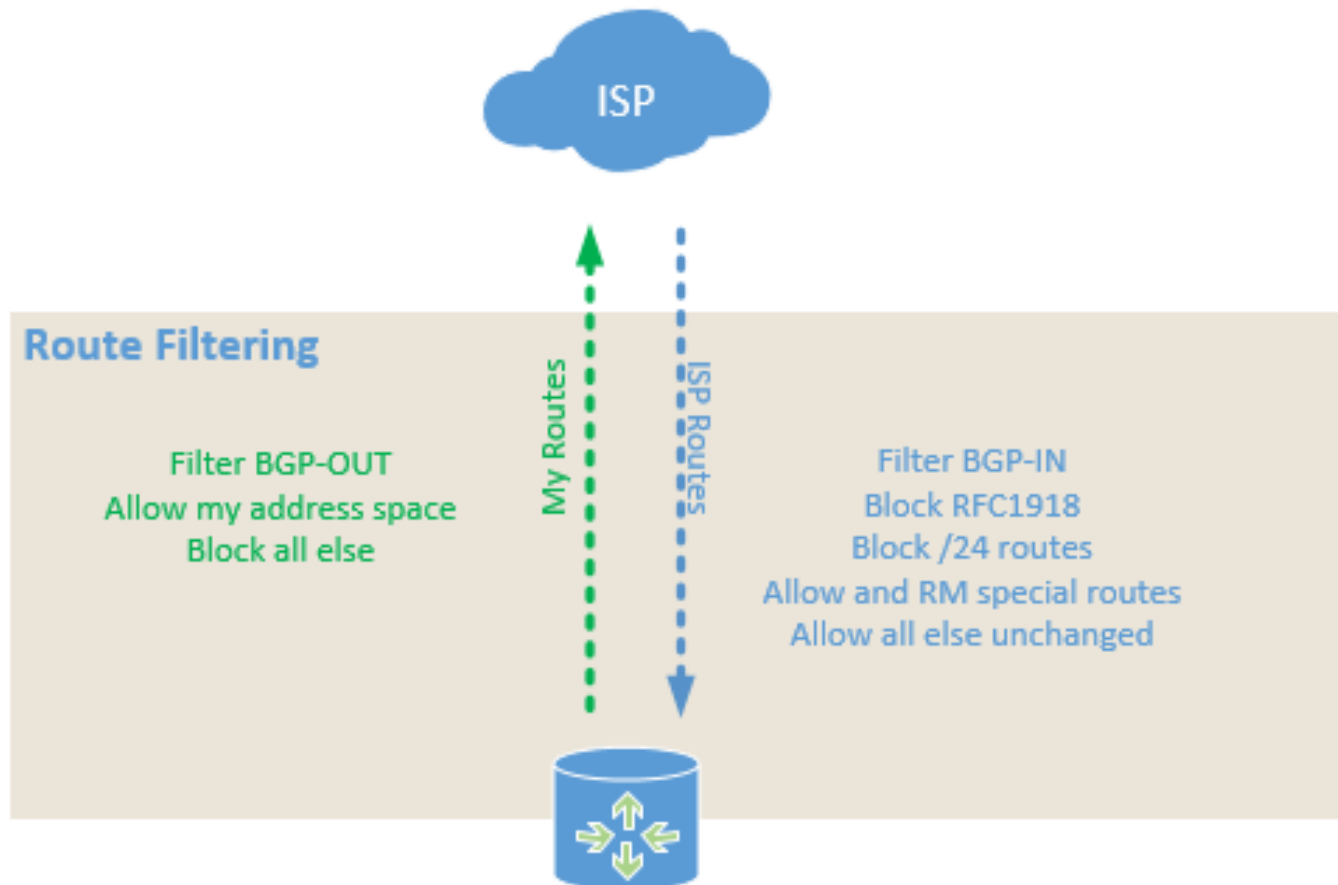
Open Internet Exchange

- ▶ An OIX is a network connection where you peer with multiple other organizations.
- ▶ There is, generally, no charge from the other organizations for this peering.
- ▶ Local transit from this area.
- ▶ Content providers (Netflix), CDNs(CloudFlare)

Steps

- ▶ Route Filters
 - ▶ Create
 - ▶ Apply
 - ▶ Verify
- ▶ BGP QoS Script
- ▶ Address-lists
- ▶ Mangle Rules
- ▶ Queue Trees
- ▶ Verification of Mangle/Queues

Route Filters

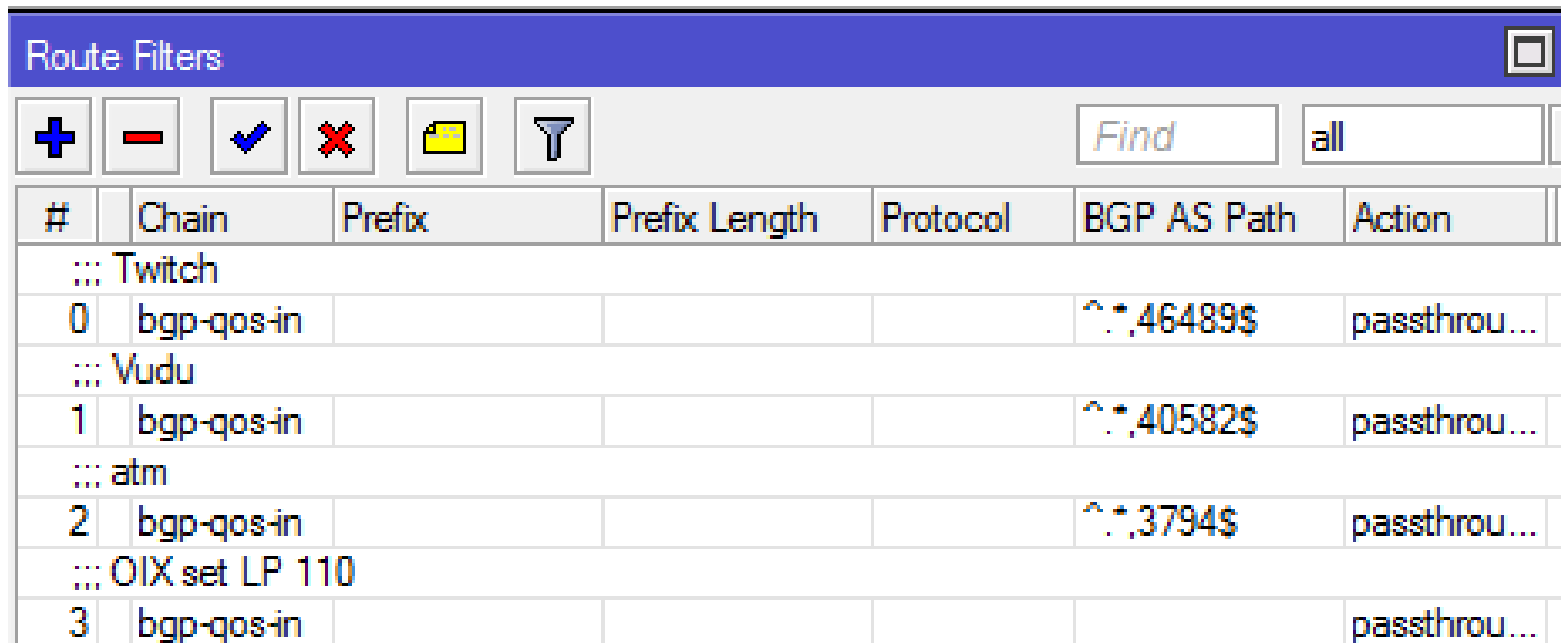


Route Filters

- ▶ Route filters (/routing filter) are used to identify and act upon routes received or sent via various dynamic routing protocols.
- ▶ Match on various BGP attributes, prefix, prefix lengths, OSPF type, etc.
- ▶ Simple actions can be taken like accept the route or reject the route.
- ▶ Complex actions can be taken like set next hop, set distance, set BGP prepending, set BGP community.

Route Filters

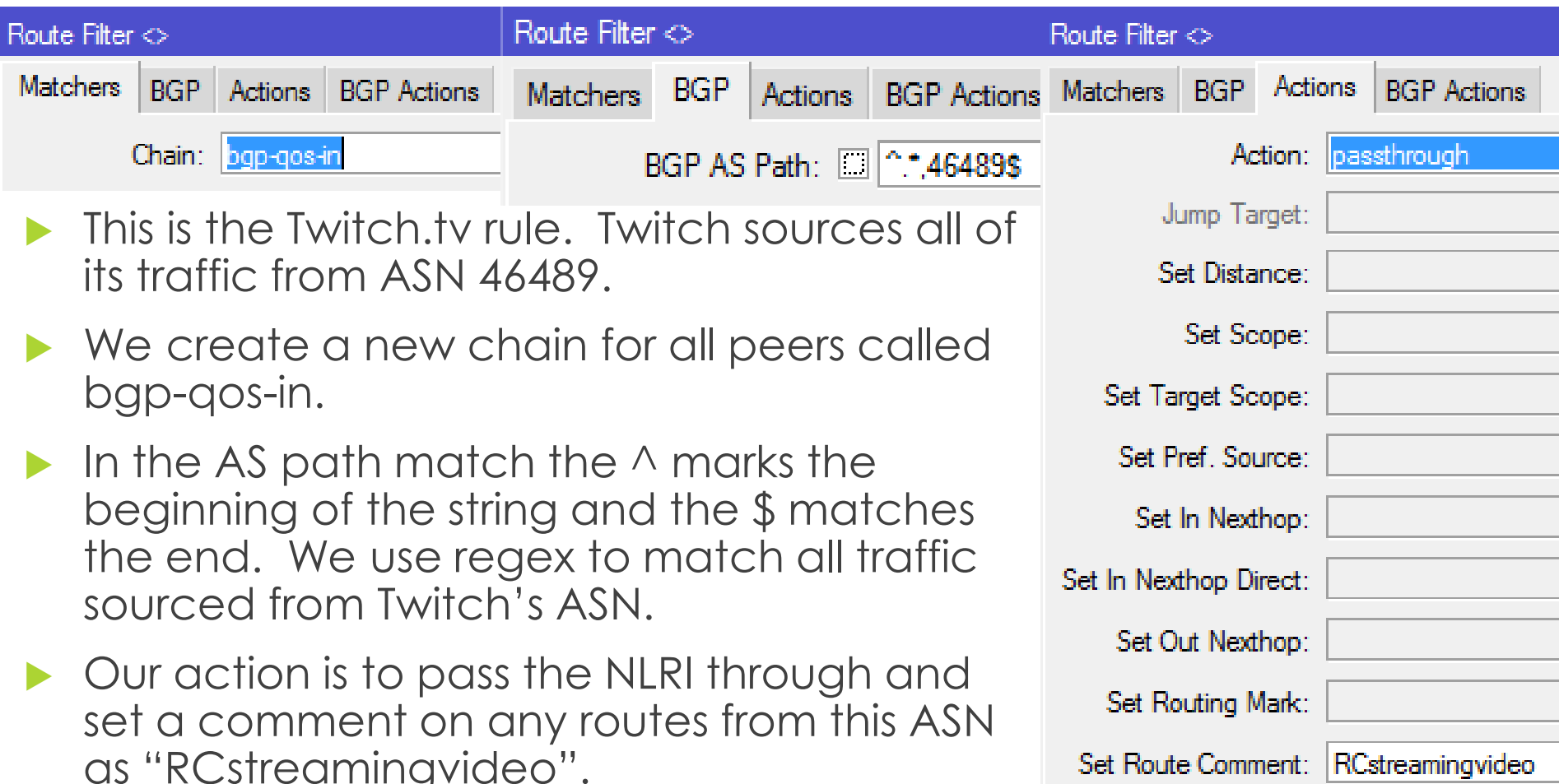
- ▶ Our example includes 4 filter statements.
 - ▶ 3 match based on BGP AS Path
 - ▶ 1 matches on BGP community



The screenshot shows a window titled "Route Filters" with a toolbar containing icons for adding (+), removing (-), enabling (checkmark), disabling (X), saving (floppy disk), and filtering (funnel). A search bar contains the text "Find" and "all". Below the toolbar is a table with the following columns: #, Chain, Prefix, Prefix Length, Protocol, BGP AS Path, and Action.

#	Chain	Prefix	Prefix Length	Protocol	BGP AS Path	Action
::: Twitch						
0	bgp-qos-in				^*,46489\$	passthrou...
::: Vudu						
1	bgp-qos-in				^*,40582\$	passthrou...
::: atm						
2	bgp-qos-in				^*,3794\$	passthrou...
::: OIX set LP 110						
3	bgp-qos-in					passthrou...

Twitch.tv Route Filter



The image shows three screenshots of Mikrotik WinBox configuration for route filters. The first screenshot shows a chain named 'bgp-qos-in'. The second screenshot shows a BGP AS Path match configuration with the regex '^*.46489\$'. The third screenshot shows the action configuration set to 'passthrough' and 'Set Route Comment' set to 'RCstreamingvideo'.

- ▶ This is the Twitch.tv rule. Twitch sources all of its traffic from ASN 46489.
- ▶ We create a new chain for all peers called bgp-qos-in.
- ▶ In the AS path match the ^ marks the beginning of the string and the \$ matches the end. We use regex to match all traffic sourced from Twitch's ASN.
- ▶ Our action is to pass the NLRI through and set a comment on any routes from this ASN as "RCstreamingvideo".

Apply Filter

- ▶ *Remember that when you apply a filter to a peer, it resets the peer completely.
- ▶ **Remember that when you adjust these lists, all of your routes from this peer become momentarily disabled while they run through the adjusted filter.

BGP Peer <peer-ISP>

General | Advanced | Status

Name:

In Filter:

Out Filter:

Route Filter Verification - Twitch

- ▶ You can see the route comment of RCstreamingvideo
- ▶ You can see the full AS path.
 - ▶ Note source is on the right and each new AS is added to the left as it traverses ASNs.

```
[admin@BGP-QOS] /ip route> print detail where dst-address=199.9.248.0/21
Flags: X - disabled, A - active, D - dynamic,
C - connect, S - static, r - rip, b - bgp, o - ospf, m - mme,
B - blackhole, U - unreachable, P - prohibit
0 ADb   ;;; RCstreamingvideo
        dst-address=199.9.248.0/21 gateway=216.81.32.81
        gateway-status=216.81.32.81 reachable via ether3 distance=20 scope=40
        target-scope=10 bgp-as-path="19366,1299,46489" bgp-origin=igp
        received-from=peer-ISP
```

Open Internet Exchange Route Filter

Route Filter <>

Chain:

Route Filter <>

Set BGP Weight:

Set BGP Local Pref.:

Route Filter <>

Matchers BGP Actions BGP Act

BGP AS Path:

BGP AS Path Length:

BGP Weight:

BGP Local Pref.:

BGP MED:

BGP Atomic Aggregate:

BGP Origin:

Locally Originated BGP:

BGP Communities:

Route Filter <>

Matchers BGP Actions BGP Act

Action:

Jump Target:

Set Distance:

Set Scope:

Set Target Scope:

Set Pref. Source:

Set In Nexthop:

Set In Nexthop Direct:

Set Out Nexthop:

Set Routing Mark:

Set Route Comment:

- ▶ The peering OIX router is tagging everything sent to us with community 65101:10.
- ▶ In the BGP tab we simply put 65101:10 in the communities column.
- ▶ Our action is to pass the NLRI through and set a comment on any routes from this ASN as "RCoix".
- ▶ Optionally we are setting the local preference to 110 in the BGP actions tab. This will prefer routes learned here above all others.

Route Filter Verification - OIX

- ▶ You can see the route comment of RCoix
- ▶ You can see the BGP community attached 65101:10.
- ▶ You will also notice our adjusted local preference of 110.

```
[admin@BGP-QOS] /ip route> print detail where dst-address=4.4.4.0/24
Flags: X - disabled, A - active, D - dynamic,
C - connect, S - static, r - rip, b - bgp, o - ospf, m - mme,
B - blackhole, U - unreachable, P - prohibit
0 ADb   ;;; RCoix
      dst-address=4.4.4.0/24 gateway=172.17.1.2
      gateway-status=172.17.1.2 reachable via gre-oix distance=20 scope=40
      target-scope=10 bgp-as-path="65101" bgp-local-pref=110 bgp-origin=igp
      bgp-communities=65101:10 received-from=peer-OIX
```


Steps

- ▶ Route Filters
- ▶ BGP QoS Script
 - ▶ Create
 - ▶ Explain
 - ▶ Verify
 - ▶ Schedule
- ▶ Address-lists
- ▶ Mangle Rules
- ▶ Queue Trees
- ▶ Verification of Mangle/Queues

BGP Script

```
:log info "BGP QoS script start";
#Define Local Var and load data
#loop variables
:local i 0;
#route ip address
:local ipAddress;
#is it marked for us
:local routeMark "null";
#route comment
:local routeComment "null";
#check the beginning of our routeComment
:local listName "null";
#loop to check the entire routing table
:foreach i in=[/ip rou find] do={
  #grab the route's comment
  :set routeComment [/ip route get $i comment]
  #check if to make sure the route comment isn't null
  :if ($routeComment!="") do={
    #grab the first two letters off of the route comment
    set listName [:pick $routeComment 0 2]
    #make sure the first two letters are RC
    :if ($listName="RC") do={
      #get the IP address of the route
      :set ipAddress [/ip route get $i dst-address]
      #log debug info to the log
      #if it is the default gateway don't add it, otherwise add it to the addresslist for 24 hours and 30 seconds
      :if ($ipAddress!=0.0.0.0/0) do={
        /ip firewall address-list rem [find where list=$routeComment address=$ipAddress];
        /ip firewall address-list add list=$routeComment address=$ipAddress timeout=88200;}
    }
  }
}
:log info "BGP QoS script complete";
```

What it does

- ▶ Loop through routing table.
- ▶ Identify routes with a route comment starting with RC.
- ▶ Delete any old address-list entries that are the same.
- ▶ Create a new address-list entry that has the route's address with a 24.5 hour timeout.
- ▶ Name the address-list entry that of the route comment "Rcstreamingvideo".

Running Our Script - CPU

- ▶ Dual 3.5Ghz Xeon
 - ▶ 75 seconds to run
 - ▶ CPU ~80%
- ▶ Quad 3Ghz Xeon
 - ▶ 45 seconds to run
 - ▶ CPU ~40%

Firewall Filter Rule configuration window. The window title is "Firewall Filter Rule". The tabs include "Filter Rule", "NAT", "Mangle", "Service Ports", "Connections", "Address Lists", and "Layer7 Protocols". The "Filter Rule" tab is active. The window contains a table with columns "Name", "Address", and "Timeout". Below the table, it says "0 items".

System status window showing various metrics. It includes a "1 item" header, "Uptime: 00:36:38", "Free Memory: 1761.0 MiB", "Total Memory: 1894.1 MiB", "CPU: Intel(R)", "CPU Count: 4", and "CPU Frequency: 2992 MHz". A mouse cursor is visible over the window.

System status window showing hardware details. It includes "USB", "CPU", "IRQ", and "RPS" sections.

CPU monitoring window showing a table of CPU statistics.

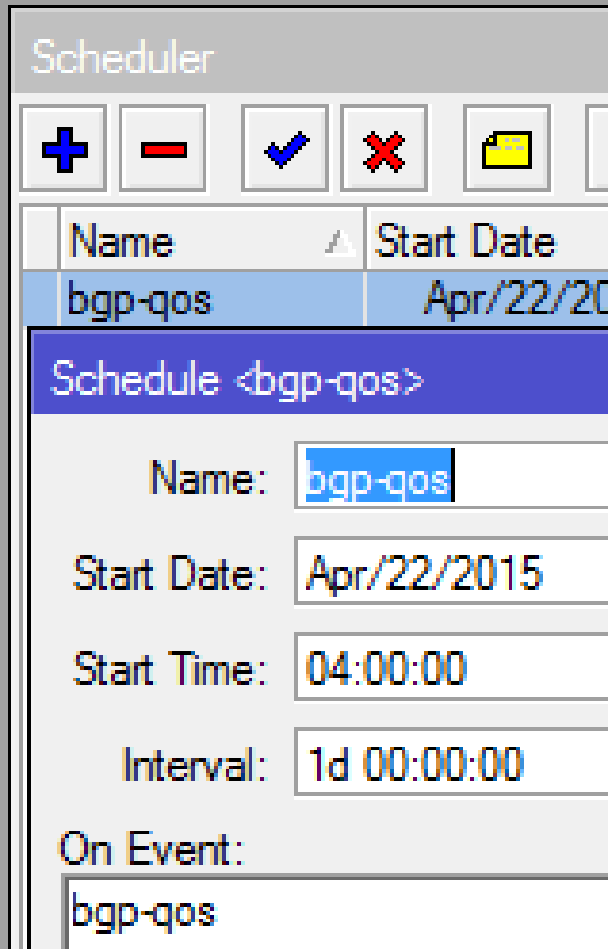
CPU	Load (%)	IRQ (%)	Disk (%)
cpu0	0	0	0
cpu1	0	0	0
cpu2	0	0	0
cpu3	0	0	0

Script List window showing a table of scripts.

Name	Owner	Last Time Started	Run Count
bgo-qos	admin	Apr/14/2015 10:17:12	4
bgo-qos2	admin	Apr/14/2015 10:30:12	5

Schedule Script

- ▶ Create the script to run every 24 hours.
- ▶ Schedule it to run at a time of low network utilization so it won't impact services.



The screenshot shows a 'Scheduler' window with a toolbar containing icons for adding (+), deleting (-), enabling (checkmark), disabling (X), and a folder icon. Below the toolbar is a table with two columns: 'Name' and 'Start Date'. The first row is highlighted in blue and contains the text 'bgp-qos' and 'Apr/22/2015'. Below the table is a section titled 'Schedule <bgp-qos>' with several input fields: 'Name' (bgp-qos), 'Start Date' (Apr/22/2015), 'Start Time' (04:00:00), and 'Interval' (1d 00:00:00). At the bottom, there is a label 'On Event:' followed by a text box containing 'bgp-qos'.

Name	Start Date
bgp-qos	Apr/22/2015

Schedule <bgp-qos>

Name:

Start Date:

Start Time:

Interval:

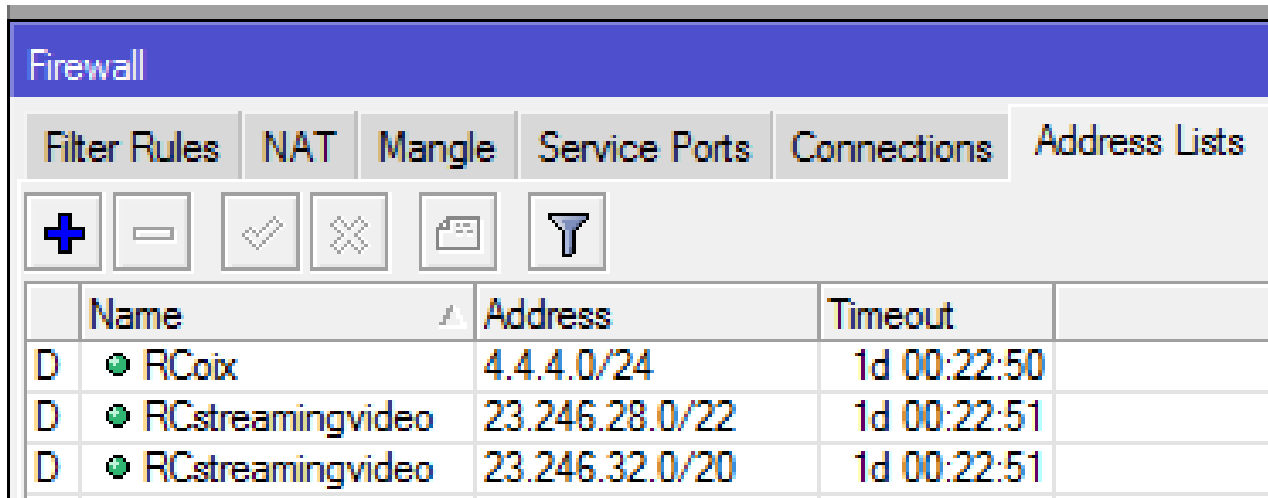
On Event:

Steps

- ▶ Route Filters
- ▶ BGP QoS Script
- ▶ Address-lists
 - ▶ Review a few created
- ▶ Mangle Rules
- ▶ Queue Trees
- ▶ Verification of Mangle/Queues

Address-lists

- ▶ Our script builds address lists.
- ▶ If our route comment starts with “RC”, then it takes the destination address and creates an address list entry with that comment name.
- ▶ All entries are set for 24.5 hours.



The screenshot shows the Mikrotik WinBox Firewall configuration window. The 'Address Lists' tab is selected. The table below displays the configured address lists.

	Name	Address	Timeout
D	RCcoix	4.4.4.0/24	1d 00:22:50
D	RCstreamingvideo	23.246.28.0/22	1d 00:22:51
D	RCstreamingvideo	23.246.32.0/20	1d 00:22:51

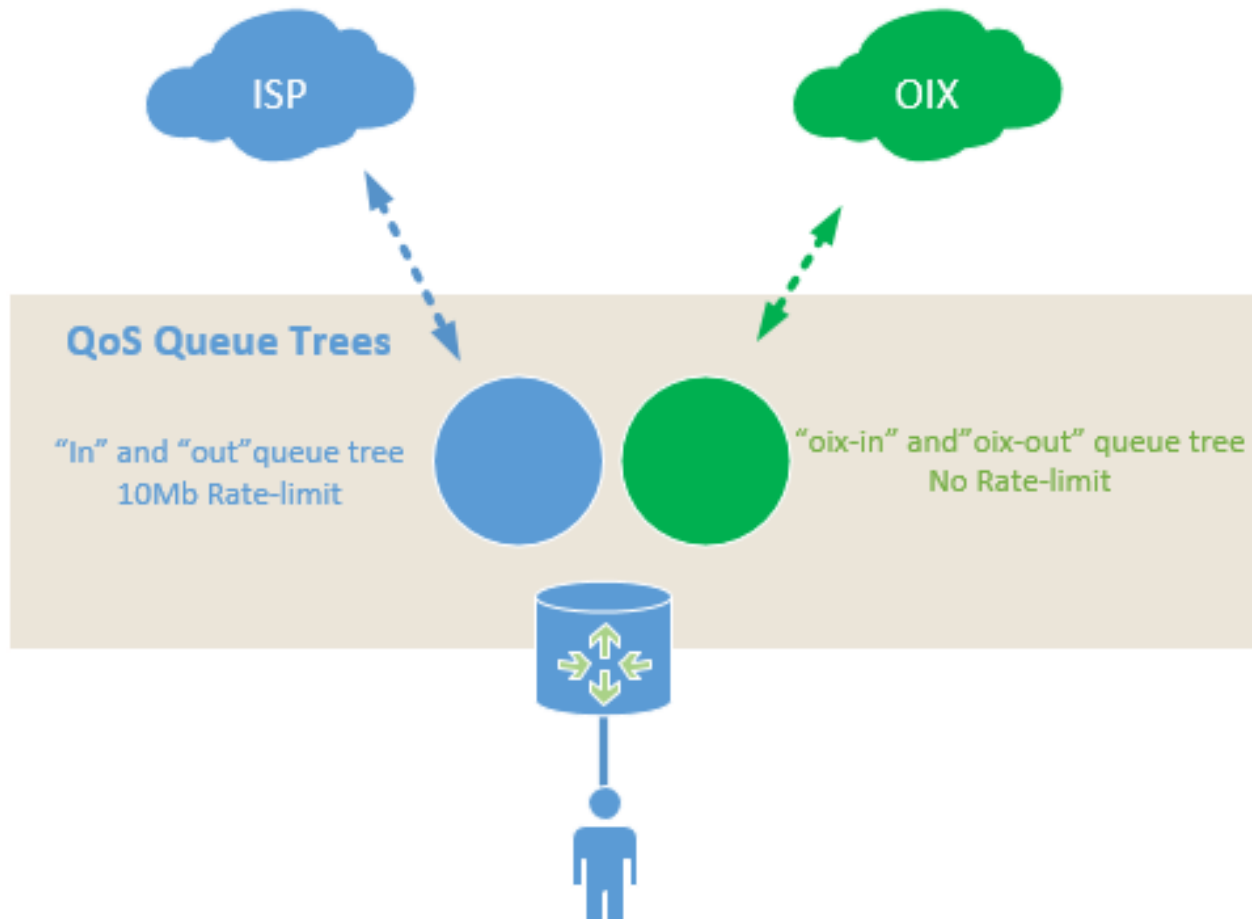
Now We Can Take Over The World

- ▶ We can use our Address-list entries in:
 - ▶ Filter rules
 - ▶ NAT rules
 - ▶ Mangle rules
- ▶ We will be using them in mangle to classify traffic for use in QoS

Steps

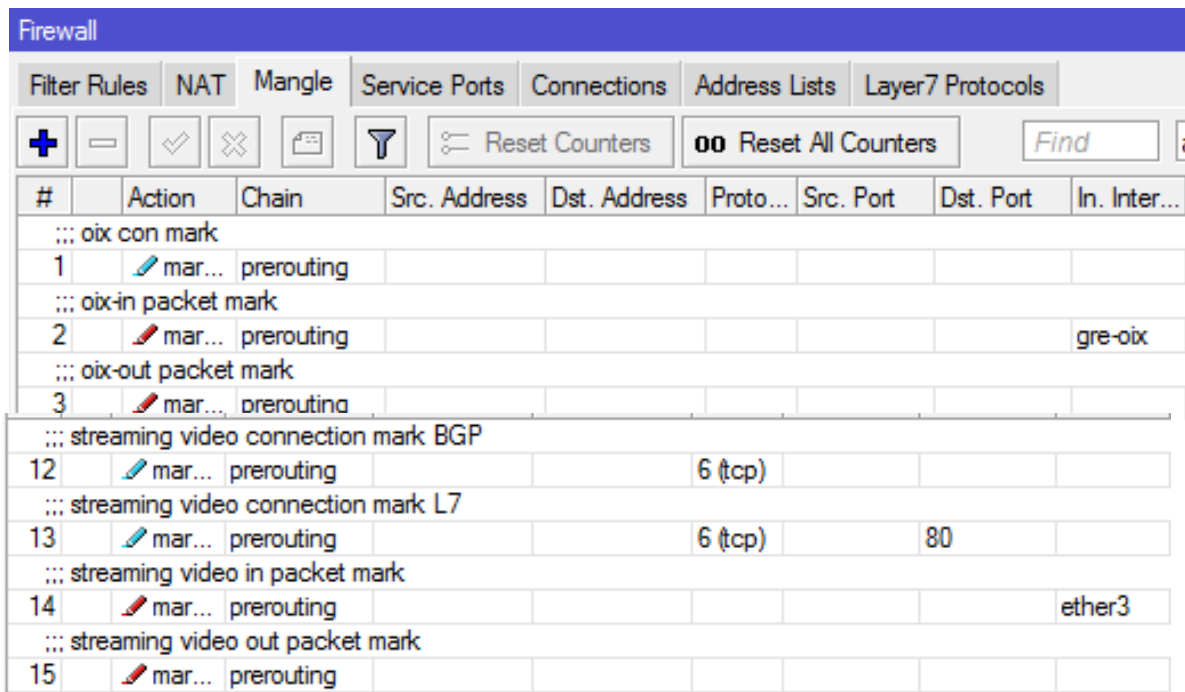
- ▶ Route Filters
- ▶ BGP QoS Script
- ▶ Address-lists
- ▶ Mangle Rules
 - ▶ Creation OIX/Twitch
- ▶ Queue Trees
- ▶ Verification of Mangle/Queues

No Rate Limit On OIX Traffic



Mangle Rules

- ▶ We connection mark the traffic based on address-lists built from the BGP-QoS script.
- ▶ Using the connection mark we packet mark traffic inbound and outbound.



The screenshot shows the Mikrotik WinBox Firewall configuration window, specifically the Mangle Rules tab. The window title is "Firewall". The tabs at the top are "Filter Rules", "NAT", "Mangle", "Service Ports", "Connections", "Address Lists", and "Layer7 Protocols". The "Mangle" tab is selected. Below the tabs are several icons for adding, deleting, and editing rules, along with buttons for "Reset Counters" and "Reset All Counters". A search box labeled "Find" is also present. The main area displays a table of mangle rules.

#	Action	Chain	Src. Address	Dst. Address	Proto...	Src. Port	Dst. Port	In. Inter...
::: oix con mark								
1	mar...	prerouting						
::: oix-in packet mark								
2	mar...	prerouting						gre-oix
::: oix-out packet mark								
3	mar...	prerouting						
::: streaming video connection mark BGP								
12	mar...	prerouting			6 (tcp)			
::: streaming video connection mark L7								
13	mar...	prerouting			6 (tcp)		80	
::: streaming video in packet mark								
14	mar...	prerouting						ether3
::: streaming video out packet mark								
15	mar...	prerouting						

Marking OIX

Mangle Rule <>

General Advanced Extra Action

Chain: prerouting

Mangle Rule <>

General Advanced Extra Action

Src. Address List:

Dst. Address List: RCoix

Mangle Rule <>

General Advanced Extra Action State

Action: mark connection

Log

Log Prefix:

New Connection Mark: oix

Passthrough

Mangle Rule <>

General Advanced Extra Action State

Chain: prerouting

In. Interface: gre-oix

Out. Interface:

Packet Mark:

Connection Mark: oix

Mangle Rule <>

General Advanced Extra Action

Action: mark packet

Log

Log Prefix:

New Packet Mark: oix-in

Passthrough

Mangle Rule <>

General Advanced Extra Action

Chain: prerouting

Connection Mark: oix

Mangle Rule <>

General Advanced Extra Action State

Action: mark packet

Log

Log Prefix:

New Packet Mark: oix-out

Passthrough

Marking Twitch.TV Traffic

Mangle Rule <>

General Advanced Extra Action Statistics

Src. Address List: internal-nets

Dst. Address List: RCstreamingvideo

Mangle Rule <>

General Advanced Extra Action Statistics

Action: mark connection

Log

Log Prefix:

New Connection Mark: streaming-video

Passthrough

Mangle Rule <>

General Advanced Extra Action

Chain: prerouting

In. Interface: ether3

Connection Mark: streaming-video

Mangle Rule <>

General Advanced Extra Action Statistics

Action: mark packet

Log

Log Prefix:

New Packet Mark: streaming-video-in

Passthrough

Mangle Rule <>

General Advanced Extra Action

Chain: prerouting

Connection Mark: streaming-video

Mangle Rule <>

General Advanced Extra Action Statistics

Action: mark packet

Log

Log Prefix:

New Packet Mark: streaming-video-out

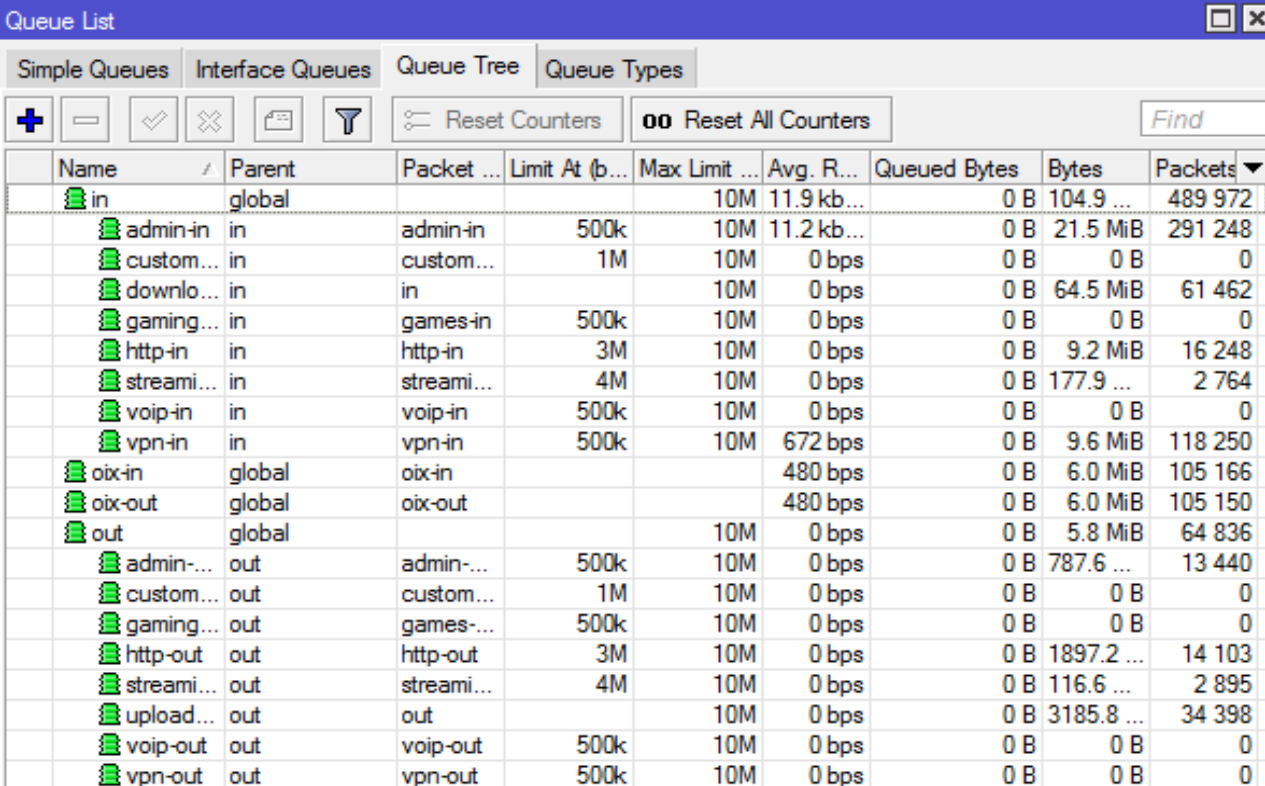
Passthrough

Steps

- ▶ Route Filters
- ▶ BGP QoS Script
- ▶ Address-lists
- ▶ Mangle Rules
- ▶ Queue Trees
 - ▶ Creation
- ▶ Verification of Mangle/Queues

Queue Trees

- ▶ Single direction HTBs.
- ▶ Allows for ingoing and outgoing queues.
- ▶ Break down each via packet marks.
- ▶ We have different services separated and prioritized.



The screenshot shows the Mikrotik WinBox 'Queue List' window. The 'Queue Tree' tab is selected, displaying a hierarchical list of queues. The table columns include Name, Parent, Packet Limit, Limit At (bps), Max Limit, Avg. Rate, Queued Bytes, Bytes, and Packets. The tree structure shows a root 'in' queue under 'global', which branches into various service-specific queues like 'admin-in', 'custom...', 'downlo...', 'gaming...', 'http-in', 'streami...', 'voip-in', and 'vpn-in'. There are also 'oix-in' and 'oix-out' queues under 'global', and an 'out' queue under 'global' which branches into 'admin-...', 'custom...', 'gaming...', 'http-out', 'streami...', 'upload...', 'voip-out', and 'vpn-out'.

Name	Parent	Packet ...	Limit At (b...	Max Limit ...	Avg. R...	Queued Bytes	Bytes	Packets
in	global			10M	11.9 kb...	0 B	104.9 ...	489 972
admin-in	in	admin-in	500k	10M	11.2 kb...	0 B	21.5 MiB	291 248
custom...	in	custom...	1M	10M	0 bps	0 B	0 B	0
downlo...	in	in		10M	0 bps	0 B	64.5 MiB	61 462
gaming...	in	games-in	500k	10M	0 bps	0 B	0 B	0
http-in	in	http-in	3M	10M	0 bps	0 B	9.2 MiB	16 248
streami...	in	streami...	4M	10M	0 bps	0 B	177.9 ...	2 764
voip-in	in	voip-in	500k	10M	0 bps	0 B	0 B	0
vpn-in	in	vpn-in	500k	10M	672 bps	0 B	9.6 MiB	118 250
oix-in	global	oix-in			480 bps	0 B	6.0 MiB	105 166
oix-out	global	oix-out			480 bps	0 B	6.0 MiB	105 150
out	global			10M	0 bps	0 B	5.8 MiB	64 836
admin-...	out	admin-...	500k	10M	0 bps	0 B	787.6 ...	13 440
custom...	out	custom...	1M	10M	0 bps	0 B	0 B	0
gaming...	out	games-...	500k	10M	0 bps	0 B	0 B	0
http-out	out	http-out	3M	10M	0 bps	0 B	1897.2 ...	14 103
streami...	out	streami...	4M	10M	0 bps	0 B	116.6 ...	2 895
upload...	out	out		10M	0 bps	0 B	3185.8 ...	34 398
voip-out	out	voip-out	500k	10M	0 bps	0 B	0 B	0
vpn-out	out	vpn-out	500k	10M	0 bps	0 B	0 B	0

Queue Trees

- ▶ Incoming and outgoing based on packet marks from our mangle rules.

Queue List

Simple Queues | Interface Queues | Queue Tree

+ - ✓ ✕ [icon] [icon] Reset Co...

Name [v] contains [v] oix

Name	Parent	Packet ...	Lim
oix-in	global	oix-in	
oix-out	global	oix-out	

Queue <oix-in>

General | Statistics

Name: oix-in

Parent: global [v]

Packet Marks: oix-in [v] [d]

Queue Type: default [v]

Priority: 8

Limit At: [] bits/s

Max Limit: [] bits/s

Burst Limit: [] bits/s

Burst Threshold: [] bits/s

Burst Time: [] s

Queue <oix-out>

General | Statistics

Name: oix-out

Parent: global [v]

Packet Marks: oix-out [v] [d]

Queue Type: default [v]

Priority: 8

Limit At: [] bits/s

Max Limit: [] bits/s

Burst Limit: [] bits/s

Burst Threshold: [] bits/s

Burst Time: [] s

Steps

- ▶ Route Filters
- ▶ BGP QoS Script
- ▶ Address-lists
- ▶ Mangle Rules
- ▶ Queue Trees
- ▶ Verification of Mangle/Queues
 - ▶ OIX
 - ▶ Twitch

OIX Testing

Queue List

Simple Queues Interface Queues Queue Tree Queue Types

Torch (Running)

- Basic -
 Interface: ether4
 Entry Timeout: 00:00:03 s

- Collect -
 Src. Address Src. Address6
 Dst. Address Dst. Address6
 MAC Protocol Port
 Protocol VLAN Id
 DSCP

- Filters -
 Src. Address: 0.0.0.0/0
 Dst. Address: 0.0.0.0/0
 Src. Address6: ::/0
 Dst. Address6: ::/0
 MAC Protocol: all
 Protocol: any
 Port: any
 VLAN Id: any
 DSCP: any

Et...	Prot...	Src.	Dst.	VLAN Id	DSCP	Tx Rate	Rx Rate	Tx Pac
800 (ip)	1 (ic...	172.22.0.254	4.4.4.4			394 bps	592 bps	

1 item Total Tx: 394 bps Total Rx: 592 bps Total Tx Packet: 0 Total Rx Packet: 1

bgp-qos2	admin	Apr/18/2015 21:09:47	4	cpu3	0	0
----------	-------	----------------------	---	------	---	---

Twitch Testing

Queue List

Simple Queues | Interface Queues | Queue Tree | Queue Types

Buttons: +, -, ✓, ✗, 📁, 🗑️

00 Reset Counters 00 Reset All Counters Find

Name	Parent	Packet ...	Limit At (b...	Max Limit ...	Avg. R...	Queued Bytes	Bytes	Packets
in	global			10M	995.9 k...	0 B	1264.3 ...	1 15
admin-in	in	admin-in	500k	10M	9.4 kbps	0 B	11.3 KiB	14
custom...	in	custom...	1M	10M	0 bps	0 B	0 B	
downlo...	in	in		10M	1632 bps	0 B	1036 B	
gaming...	in	games-in	500k	10M	0 bps	0 B	0 B	
http-in	in	http-in	3M	10M	983.9 k...	0 B	1250.5 ...	98

Firewall

Filter Rules | NAT | Mangle | Service Ports | Connections | Address Lists | Layer7 Protocols

Buttons: +, -, ✓, ✗, 📄, 🗑️


Find all

Name	Address	Timeout
external-n...	216.81.32.80/29	
internal-nets	172.22.0.0/16	

2 items

	Load (%)	MEM (%)	DISK (%)
	0	0	0
	0	0	0
	0	0	0
	0	0	0

2 items (1 selected)



One last thing, shake my hand
and buy the brothers a beer!

Thanks and happy routing!

Resources

- ▶ Greg's Blog
 - ▶ <http://GregSowell.com>
- ▶ TheBrothersWISP
 - ▶ <http://thebrotherswisp.com/>
- ▶ Greg Sowell Routing
 - ▶ <http://gregsowell.com/?p=1611>
- ▶ Greg Sowell QoS
 - ▶ <http://gregsowell.com/?p=4665>