Internet Exchange:

Peering with RouterOS in the IXs

by Lorenzo Busatti

Welcome to the 1st MUM in Denmark!

About me

Lorenzo Busatti

- Founder of Grifonline S.r.l. [ISP] (1997)
- Founder of Linkwave [WISP] (2006)
- MikroTik Trainer [NA,RE,WE,TCE,INE,UME](2010)
- Member of RIPE, AMS-IX, MIX-IT

About me

Lorenzo Busatti

- Worldwide Trainer and Consultant, previous experiences in:
- Italy, USA, UnitedArabEmirates, Brasil, Slovenia, Poland, Croatia, Czech Republic, etc ….

linkwave
2006 2016
10 anni di wireless!

About me

Liechtenstein
Suisse
Svizzera
Switzerland
Graz
Klagenfurt am Wörthersee
Maribor
Slovenija
Slovenia
Zagreb
Pécs
Karlovac
Milano
Verona
Venezia
Padova
Torino
Asti
Parma
Bologna
Genova
Ravenna
Hrvatska
Croatia
Bosna i Hercegovi
Босна и Херцегови
Bosnia an Herzegovi
Zadar
Šibenik
Mostar
San Marino
Monaco
Cannes
Toulon
Livorno
Firenze
Italia
Italy
Grosseto
Viterbo
Pescara
L'Aquila
Roma
Latina
Foggia
Bari
Napoli
Potenza
Salerno
Taranto
Brindi
Lecce
Adriatic Sea
Ajaccio
Olbia
Sassari
Sardegna
Tyrrhenian Sea
Cagliari
Cosenza
Catanzaro
Palermo
Messina
Trapani
Marsala
Catania
Agrigento
Gela
Siracusa
Ragusa
Ionia
عنابه
Annaba
تونس
Tunis
Google
Camaiore
Pistoia
Prato
Viareggio
Lucca
E76
Quarrata
Firenze
San Piero In Bagno
San Marin
Pisa
Empoli
Bibbiena
Cascina
Pontedera
Urbania
Figline Valdarno
Livorno
Sansepolcro
Rosignano Solvay
Poggibonsi
Montevarchi
Arezzo
Colle di Val d'Elsa
Città di Castello
Cecina
Siena
Castiglion Fiorentino
Gubbio
E80
E35
Umbertide
Cortona
Sinalunga
Campiglia Marittima
Massa Marittima
Montepulciano
Perugia
Città della Pieve
Ribolla
Piombino
Follonica
Arcidosso
Marsciano
Portoferraio
A1
Todi
Grosseto
Orvieto
Acquapendente
Castiglione della Pesc
E45
Pitigliano
E80
Montefiascone
Amelia
Orbetello
Viterbo
Narni
Montalto di Castro
Civita Castellana
Vetralla
Tarquinia
Civitavecchia
Bracciano
Cerveteri
Ladispoli
Roma
Fiumicino

I'm a MikroTik *evangelist*

About me

- Founder (2016) of the

ROUTING & WIRELESS® ACADEMY

Non **P**rofit **O**rganization for

High Quality Training Partners

Dedicated to Max

The peering in the Internet Exchange Points

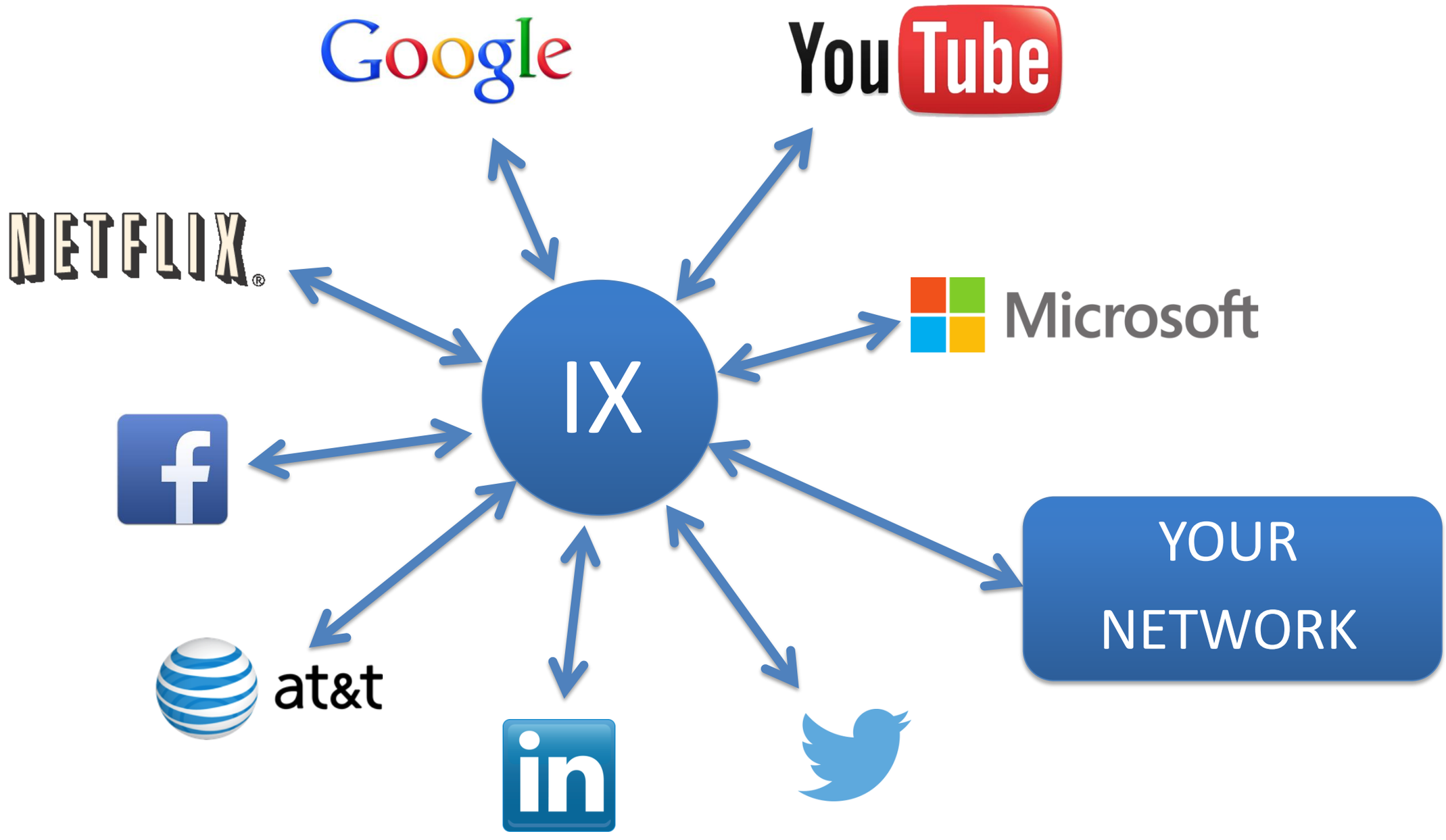
Internet Exchange Points

An Internet exchange point (IX or IXP) is a physical infrastructure where Internet service providers (ISPs) exchange Internet traffic between their networks (autonomous systems).

http://en.wikipedia.org/wiki/Internet_exchange_point

Internet Exchange Points

Carriers, ISPs and Content Provider meet togheter



Internet Exchange (Cont.)

IXPs reduce the portion of an ISP's traffic which must be delivered via their upstream transit providers, thereby reducing the average per-bit delivery cost of their service. Furthermore, the increased number of paths learned through the IXP improves routing efficiency and fault-tolerance.

IXPs in the world

Many Internet Exchange in the continents. The largest ones in the world are DE-CIX in Frankfurt, AMS-IX in Amsterdam, LINX in London, the Moscow Internet Exchange, Equinix Ashburn in Washington D.C., and JPNAP in Tokyo, and so on.

http://en.wikipedia.org/wiki/List_of_Internet_exchange_points

http://en.wikipedia.org/wiki/List_of_Internet_exchange_points_by_size

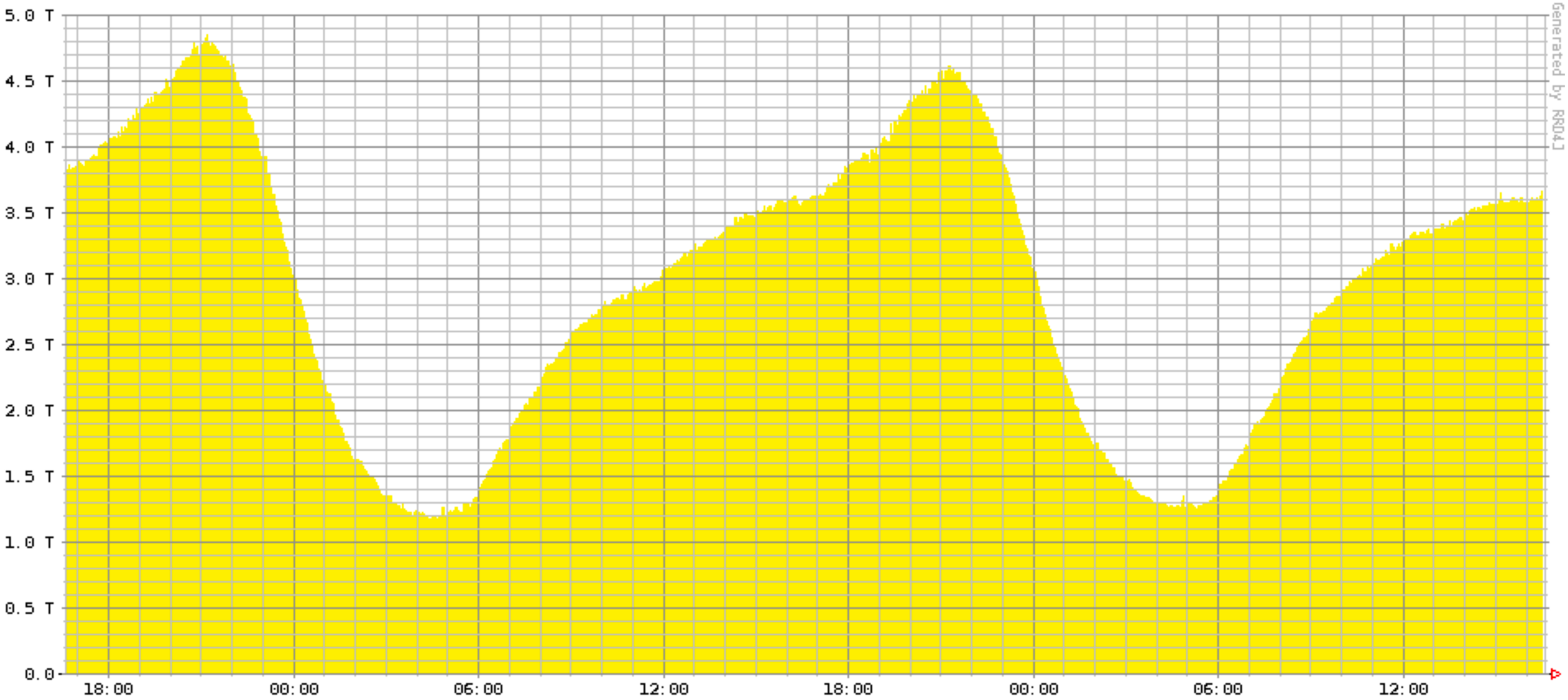
IXPs in the world

DE-CIX	Frankfurt	**5.2 Tbps**	660 ASNs
AMS-IX	Amsterdam	4.7 Tbps	**792 ASNs**
LINX	London	3.3 Tbps	690 ASNs

DE-CIX, Frankfurt

1st in the world, for traffic exchange.

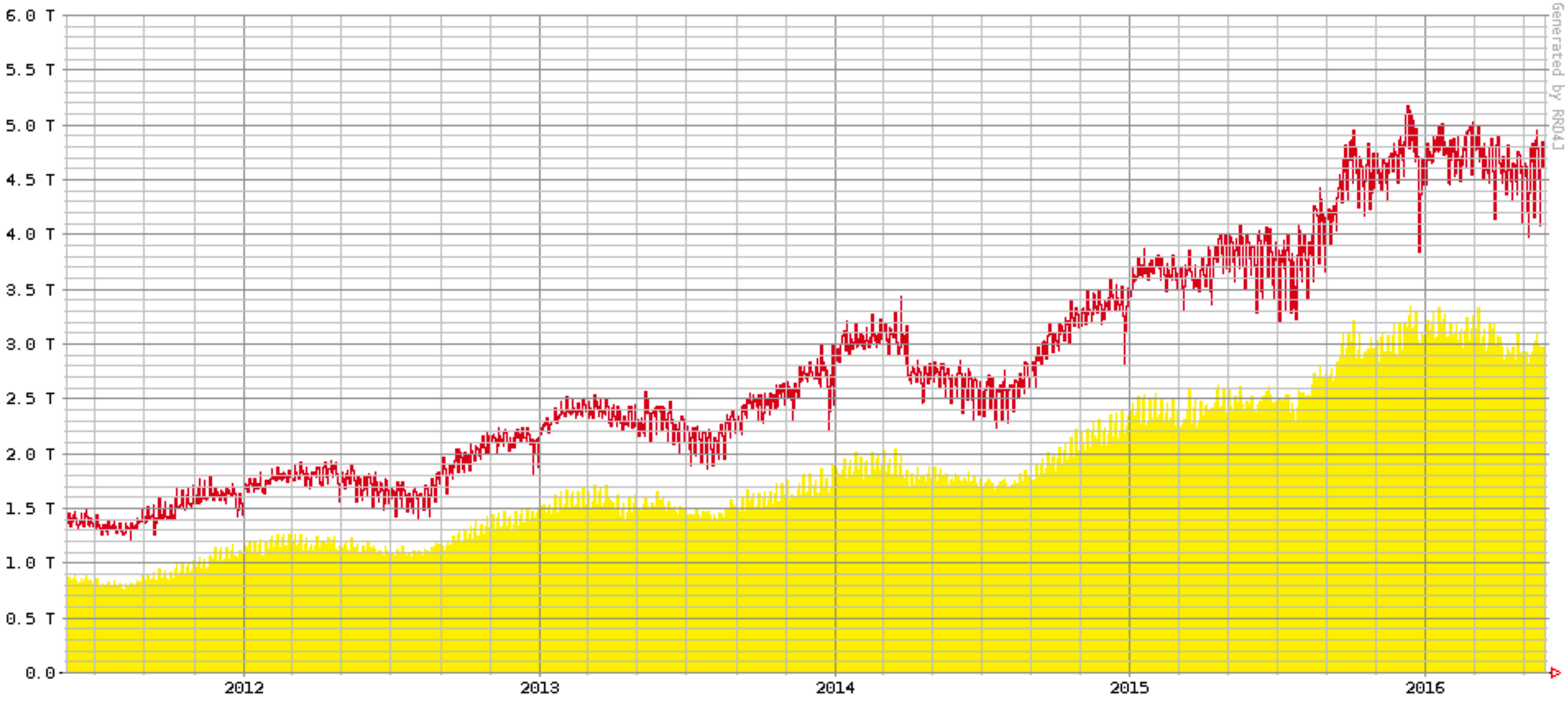
The last 2 days traffic average:



DE-CIX, Frankfurt

1st in the world, for traffic exchange.

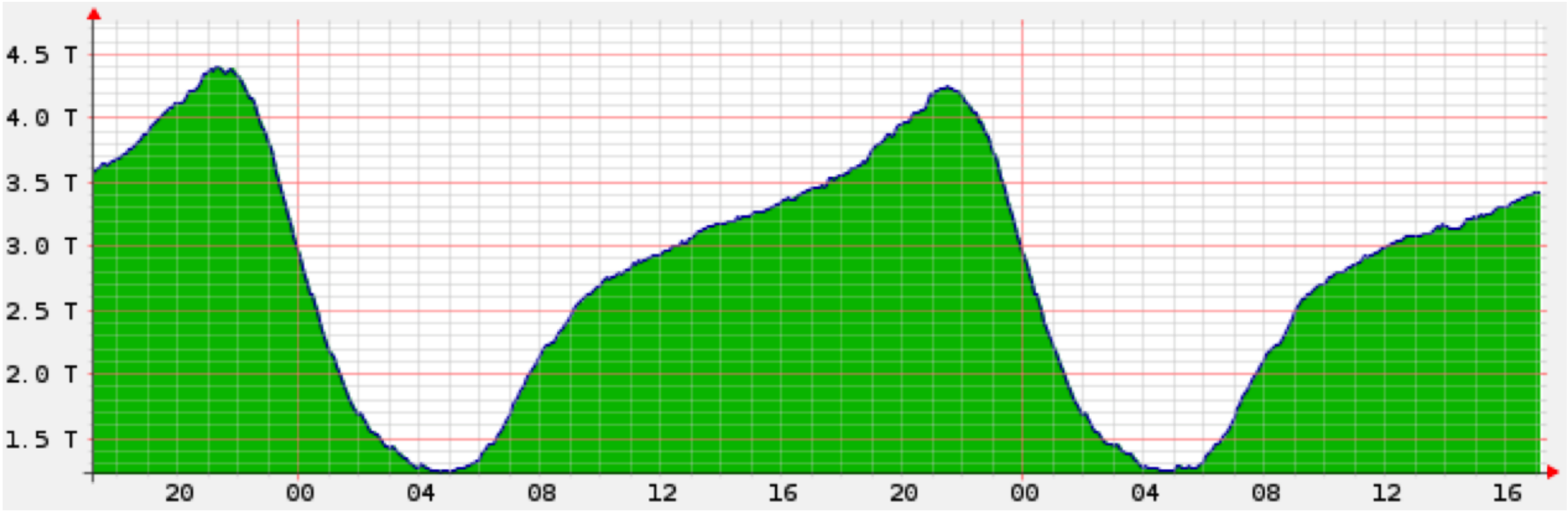
The last 5 years traffic average:



AMS-IX, Amsterdam

2nd in the world, for traffic exchange. The 1st for ASNs

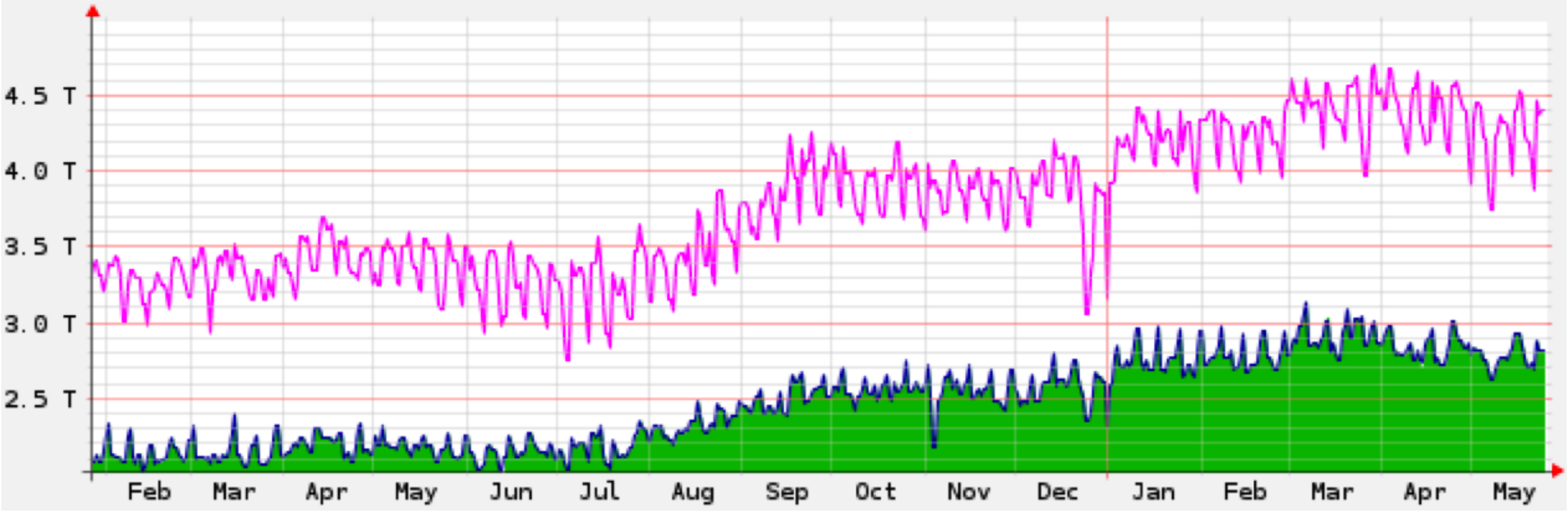
The last 2 days traffic average:



AMS-IX, Amsterdam

2nd in the world, for traffic exchange. The 1st for ASNs

The last 1 year traffic average:



AMS-IX, Amsterdam

Few numbers:

ASNs	Ports	Peak (Tb/s)	Cur (Tb/s)	Capacity (Tb/s)
792	1481	4.711	3.386	19.9729

Peering around the globe

AMS-IX, Amsterdam

2014, November:

ASNs	Ports	Peak (Tb/s)	Cur (Tb/s)	Capacity (Tb/s)
685	1324	3.339	3.142	12.907

2015, March:

ASNs	Ports	Peak (Tb/s)	Cur (Tb/s)	Capacity (Tb/s)
711	1404	3.52	1.724	14.4314

2016, May:

ASNs	Ports	Peak (Tb/s)	Cur (Tb/s)	Capacity (Tb/s)
792	1481	4.711	3.386	19.9729

IXPs in the world

- The grow is impressive
- The more they grow the less I have to use an upstream ISP

Public & Private Peering

Each entity that participates in a IXP define if:

- Participate in the "public peering" (Route Server) [many-to-many]
- Only do private “peering” [one-to-one]
- Both of them

Direct Peer

To the ASNs with I have a direct peer the AS PATH will be the shorter one. ☺

Direct Peer

The consequence will be:

You'll start to peer like a hell ☺

About the router

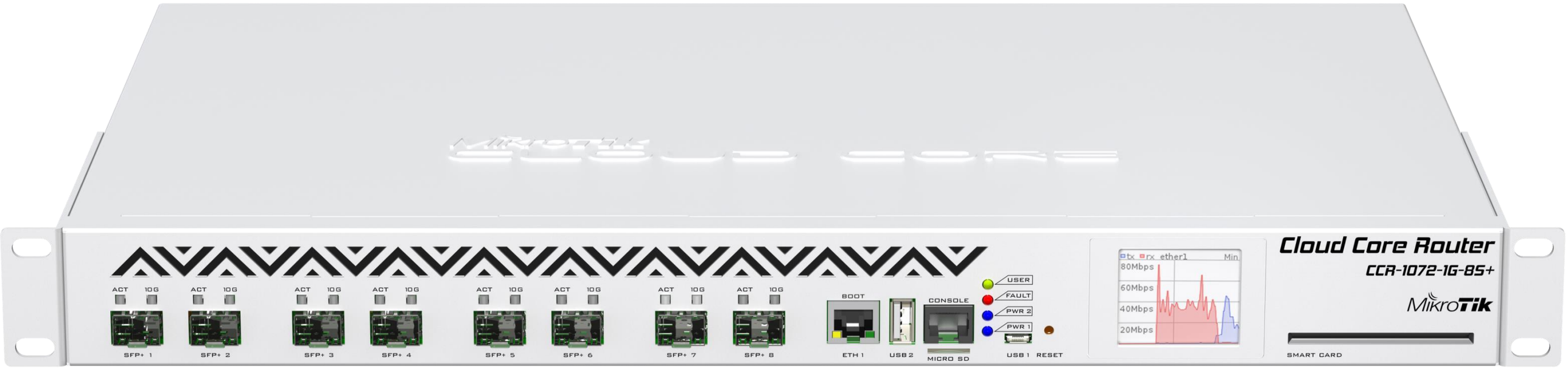
MIkroTik for BGP

And now **the** question:

How many are using RouterOS for their BGP routers?

CCR1072-1G-8S+

1U rackmount, 8xSFP+ cages, 72 cores x 1GHz CPU, 16GB RAM, two removable (hotplug) power supplies



up to **120 million packets per second**

up to **80Gbps throughput**

Router prices

CCR1072-1G-8S+ List price: $ 3,050.00

(**NO** options or **additional licenses** required) [SFP+]

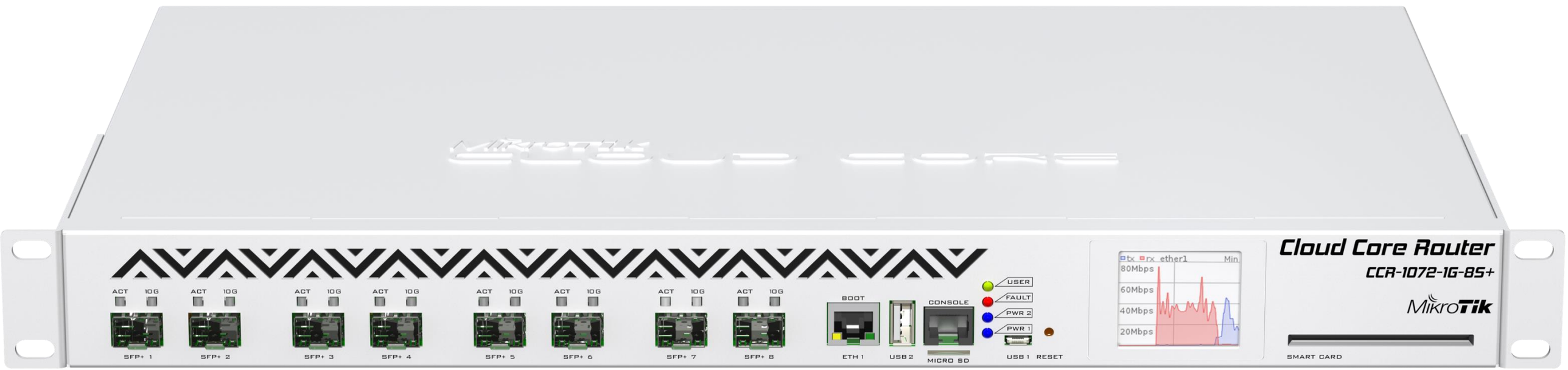
A Cisco equivalent **XXX times** the CCR1072

A Juniper equivalent **XXX times** the CCR1072

MIkroTik for BGP

And now another question:

There are reasons for not to use RouterOS as BGP router?



My experience with RouterOS and BGP

RouterOS in real life BGP

For my BGP routers I'm using RouterOS since the **version 3.XX, in 2007**.

Then upgraded to version 4.XX.

Then upgraded to version 5.XX.

Then upgraded to version 6.XX.

…….

Then I will upgrade to version 7.XX (soon??)

RouterOS in real life BGP

- Never crashed
- Never rebooted (except few secs for upgrades)

I can state: yes, it’s working fine!

My BGP router

BGP

Instances | VRFs | Peers | Networks | Aggregates | VPN4 Routes | Advertisements

Refresh | Refresh All | Resend | Resend All

Name	Remote AS	TTL	Prefix Co...	State
AMS-IX: Akamai 1 IPv4	20940	1	6	established
AMS-IX: Akamai 2 IPv4	20940	1	26	established
AMS-IX: Akamai 2 IPv6	20940	1	6	established
AMS-IX: Apple 1 IPv4	714	1	21	established
AMS-IX: Apple 1 IPv6	714	1	6	established
AMS-IX: Apple 2 IPv4	714	1	21	established
AMS-IX: Apple 2 IPv6	714	1	6	established
AMS-IX: BIT IPv4	12859	1	81	established
AMS-IX: BIT IPv6	12859	1	36	established
AMS-IX: CloudFlare IPv4	13335	1	249	established
AMS-IX: CloudFlare IPv6	13335	1	3	established
AMS-IX: Computerline IPv4	41913	1	11	established
AMS-IX: Computerline IPv6	41913	1	5	established
AMS-IX: Dropbox IPv4	19679	1	39	established
AMS-IX: Dropbox IPv6	19679	1		established
AMS-IX: Facebook 1 IPv4	32934	1	29	established
AMS-IX: Facebook 1 IPv6	32934	1	15	established
AMS-IX: Facebook 2 IPv4	32934	1	30	established
AMS-IX: Facebook 2 IPv6	32934	1	15	established
AMS-IX: GoDaddy IPv4	26496	1	36	established
AMS-IX: GoDaddy IPv6	26496	1	8	established
AMS-IX: HE IPv4	6939	1	88731	established
AMS-IX: HE IPv6	6939	1	14449	established
AMS-IX: Linkedin IPv4	14413	1	5	established
AMS-IX: Linkedin IPv6	14413	1	4	established
AMS-IX: MANDA IPv4	8365	1	45	established
AMS-IX: MANDA IPv6	8365	1	11	established
AMS-IX: Microsoft 1 IPv4	8075	1	215	established
AMS-IX: Microsoft 1 IPv6	8075	1	35	established
AMS-IX: Microsoft 2 IPv4	8075	1	215	established
AMS-IX: Microsoft 2 IPv6	8075	1	35	established
AMS-IX: Netflix IPv4	2906	1	2	established
AMS-IX: Netflix IPv4	2906	1	2	established
AMS-IX: Netflix IPv6	2906	1	2	established
AMS-IX: Netflix IPv6	2906	1	2	established
AMS-IX: OVH 1 IPv4	16276	1	66	established
AMS-IX: OVH 1 IPv6	16276	1	3	established
AMS-IX: OVH 2 IPv4	16276	1	66	established
AMS-IX: OVH 2 IPv6	16276	1	3	established
AMS-IX: Open Carrier IPv4	41692	1	62	established
AMS-IX: Open Carrier IPv6	41692	1	10	established
AMS-IX: Packet Clearing House 1 IP...	42	1	48	established
AMS-IX: Packet Clearing House 1 IP...	42	1	45	established

75 items

My BGP router

```
routing bgp peer print count-only
78
```

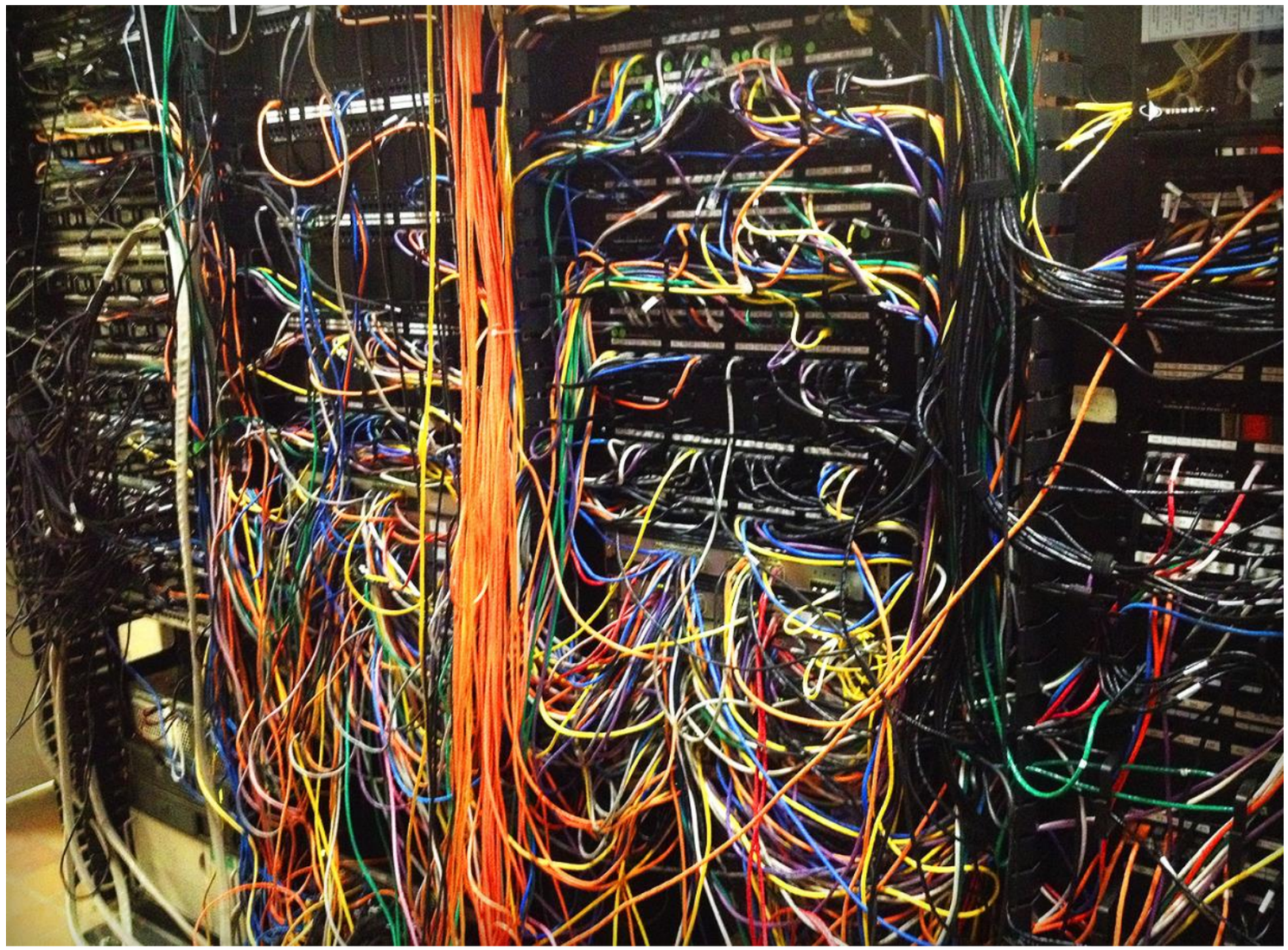
```
ip route print count-only
1587293
```

```
ip route print count-only where active
591582
```

BGPocalypse

"Other" brands crashed reaching 512K routes

(and not a long time ago ….)



Other key points

Other key point

Booting time:

Less than 30 **seconds**

Upgrade and rebooting time:

Less than 60 **seconds**

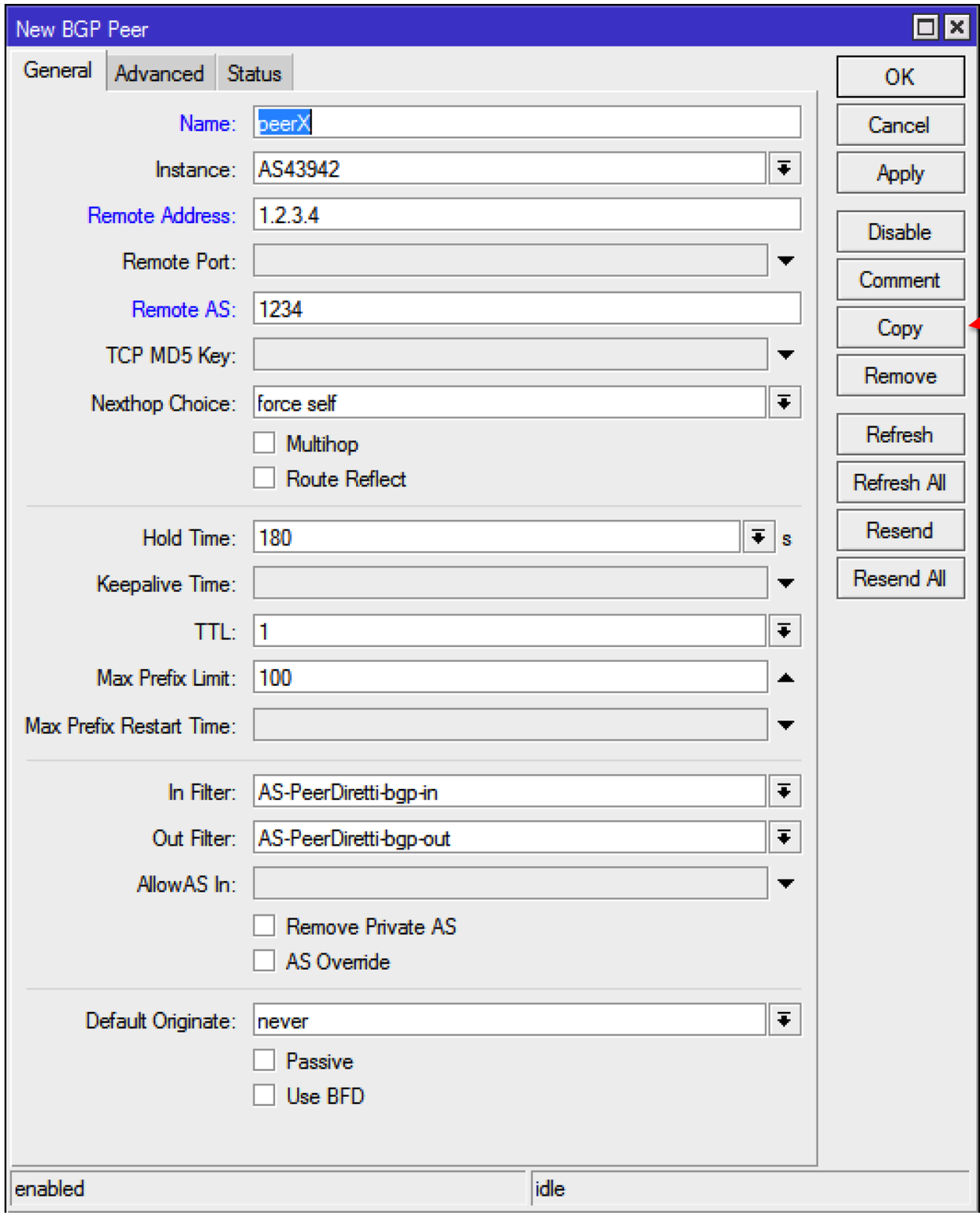
Other key point

Easy **Edit**

```
routing bgp peer set ”peer name"
remote-address=X.X.X.X
```

Other key point

The **Copy**



Other key point

The **Undo**

Need more performances?

Cloud Hosted Router

Cloud Hosted Router (CHR) is a special distribution, and license, for using RouterOS in a virtual environment.

It supports the x86 64-bit and can be used on most of the popular hypervisors such as VMWare, Hyper-V, VirtualBox, KVM and others (Proxmox, etc.).

Cloud Hosted Router

The licensing model is based on maximum interface speed only, all the **others** features are **full enabled** and **unlimited**:

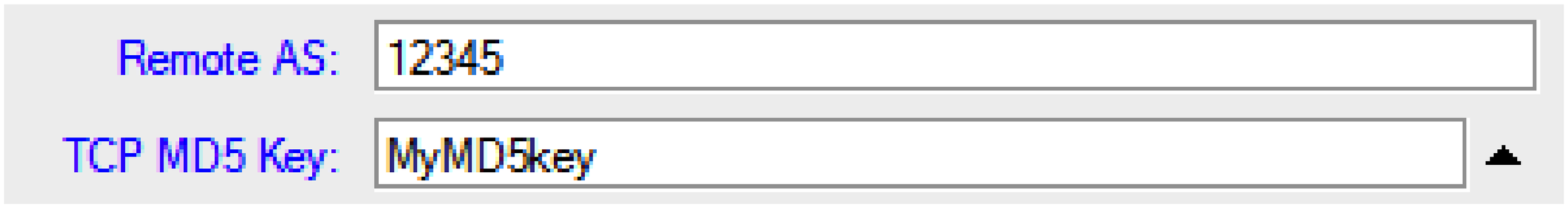
Free	1Mbps	FREE
P1	1Gbps	$ 45
P2	10Gbps	$ 95
P-unlimited	**Unlimited**	**$ 250**

Tips & Tricks

(for peering “safely”)

MD5

When planning a peer you should consider the use of the MD5 password.



It’s free ☺

MAC Address

The MAC Address on the interfaces you're using tothe IXP should not change: their router doesn't tolerate this action and will close all the connections.

IPv4 and IPv6

You should use **two separate** sessions:

- One for **IPv4**
- One for **IPv6**

AMS-IX: Akamai 1 IPv4	AS43942	:	20940
AMS-IX: Akamai 1 IPv6	AS43942	:	20940
AMS-IX: Akamai 2 IPv4	AS43942	:	20940
AMS-IX: Akamai 2 IPv6	AS43942	:	20940
AMS-IX: Apple 1 IPv4	AS43942	:	714
AMS-IX: Apple 1 IPv6	AS43942	:	714
AMS-IX: Apple 2 IPv4	AS43942	:	714
AMS-IX: Apple 2 IPv6	AS43942	:	714

Filtering

Use the filters for:

- Avoid to announce private addresses;
- Avoid to receive private addresses;
- Avoid to receive your own addresses;
- Avoid to be used as a “transit”
- Manipulate upload and download preferences

“Shit Happen” (1)

Subject [redacted] Networks AS[redacted] default route

To tech-l@ams-ix.net <tech-l@ams-ix.net> 21/09/15 13:44

Dear AMS-IX peers,

Currently I notice in my monitoring that AS[redacted] is advertising a default route originated by [redacted] to its peers.

We have this filters so we don’t have an issue but just to inform the others that they might have an issue it they accept the route.

*> 0.0.0.0/0 80.249.[redacted] 90 0 [redacted] i

I also notified them directly.

Best regards,

Filtering

Use the filters for:

- Avoid to receive the default route;

Neighbor Discovery Protocol

Disable the MikroTik Neighbor Discovery protocol (MNDP) on the interfaces you're using verso the IXP: their router doesn't tolerate such broadcast traffic, they will treat you like a threat.

BGP Community

The BGP Community will let you manage “things” of your peering partner without calling him. You can manage these “things” yourself.

And they’re very useful in the real life.

BGP Community

You can know the available BGP Community settings from your peer, or querying public DB like RIPE or anothers.

For example let’s have a look at Cogent, AS174 . Just for illustration purpose!

AS174 Community

BGP Community String	Local Pref	Effect
174:10	10	Set customer route local preference to 10 (below everything-least preferred)
174:70	70	Set customer route local preference to 70 (below peers)
174:120	120	Set customer route local preference to 120 (below customer default)
174:125	125	Set customer route local preference to 125 (below customer default)
174:135	135	Set customer route local preference to 135 (above customer default)
174:140	140	Set customer route local preference to 140 (above customer default)

AS174 Community

BGP Community String	Effect
174:970	Do not send route to NA (North America) - not accepted in NA
174:971	Do not send route to NA peers.
174:975	Set local preference to 10 in NA.
174:980	Do not send route to EU (Europe) - not accepted in EU
174:981	Do not send route to EU peers.
174:985	Set local preference to 10 in EU.
174:990	Do not send route to BGP customers, or peers.
174:991	Do not send route to peers.

AS174 Community

BGP Community String	Effect
174:3000	Do not announce.
174:3001	Prepend 174 1 time.
174:3002	Prepend 174 2 times.
174:3003	Prepend 174 3 times

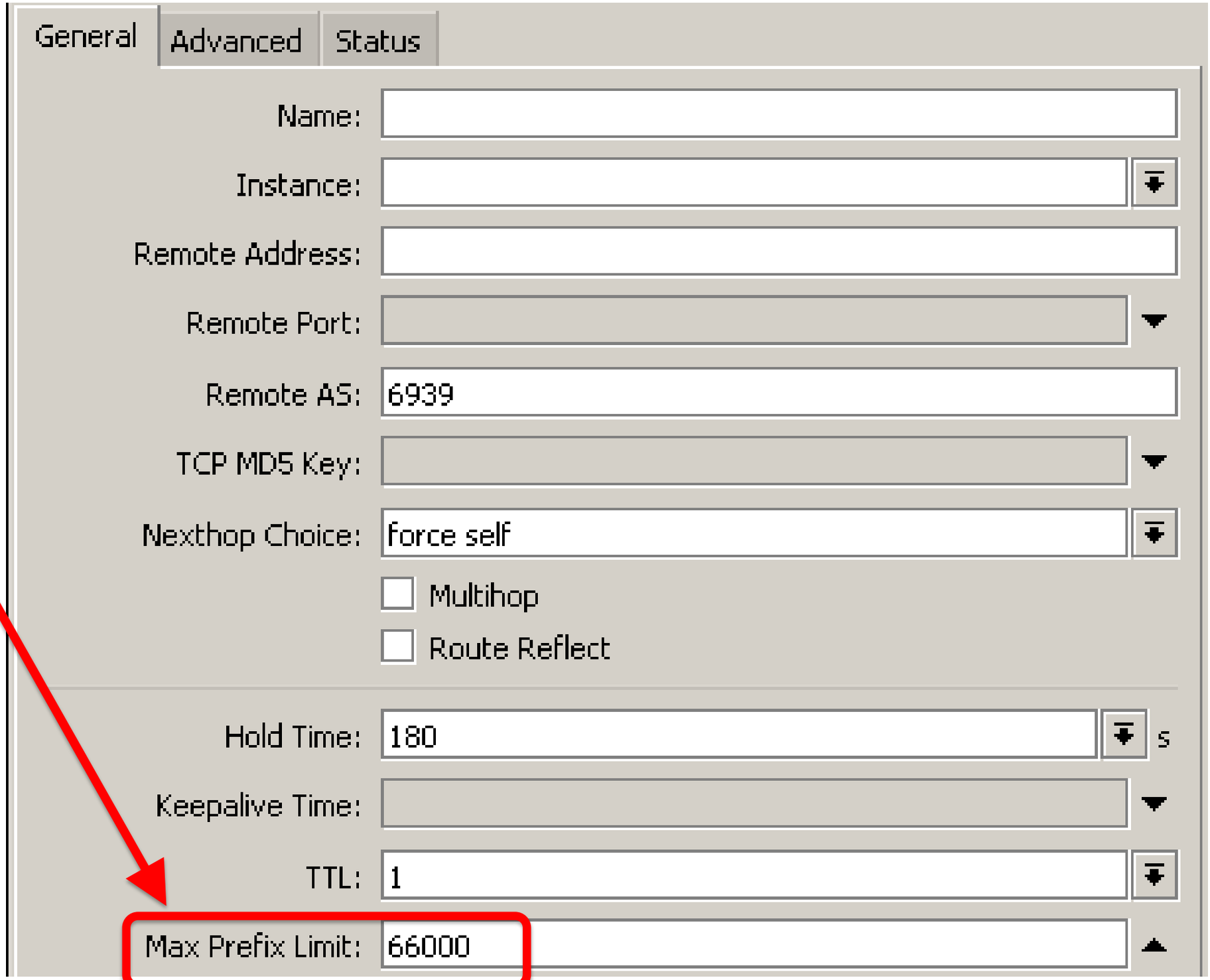
BGP Community String	Description
174:21000	Route is learned from NA (North America) non-customer.
174:21001	Route is NA internal or customer route.
174:21100	Route is learned from EU (Europe) non-customer.
174:21101	Route is a EU internal or customer route.

Prefix limit

For every peer should be a very good idea to set up a prefix limit. It should be limit the damage from receiving too many routes, i.e. in case of some “leaks”.

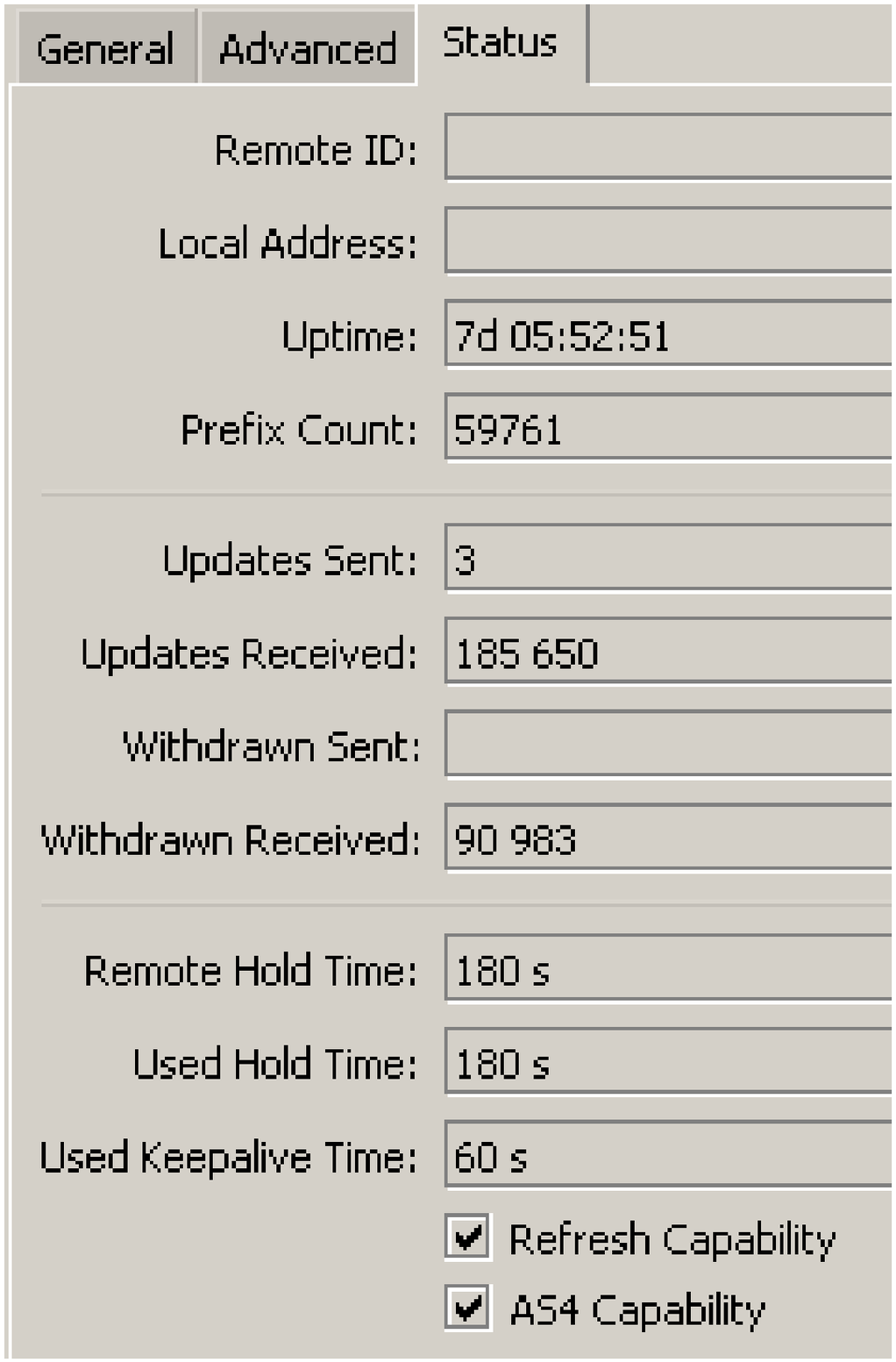
When according for a peer this information should be exchanged!

Prefix limit



Peer Status

Check the uptime, the prefix count (maybe you'll update the "prefix limit"), and the updates that happened.



“Shit Happen” (2)

Subject: ASXXXXX - IPv4 bogon prefixes leak

Hi AMS-IX peers,

I would like to apologise because between 12:11PM and 12:13PM (London time) we accidentally leaked about 3k IPv4 bogon prefixes.

That was caused by a human mistake and it was immediately corrected.

We took necessary actions to make sure it will not happen again in the future.

If you peer with us at AMS-IX could you please reset the peering session(s), if it is still down:

ASN: XXXXX

IP: XX.XX.XX.XX

Please accept our apology for problems the issue may caused.

Team Cymru

If you don’t know them:

A no-profit organization that provide you a free BGP peer, feeding you the daily updated bogons prefixes.

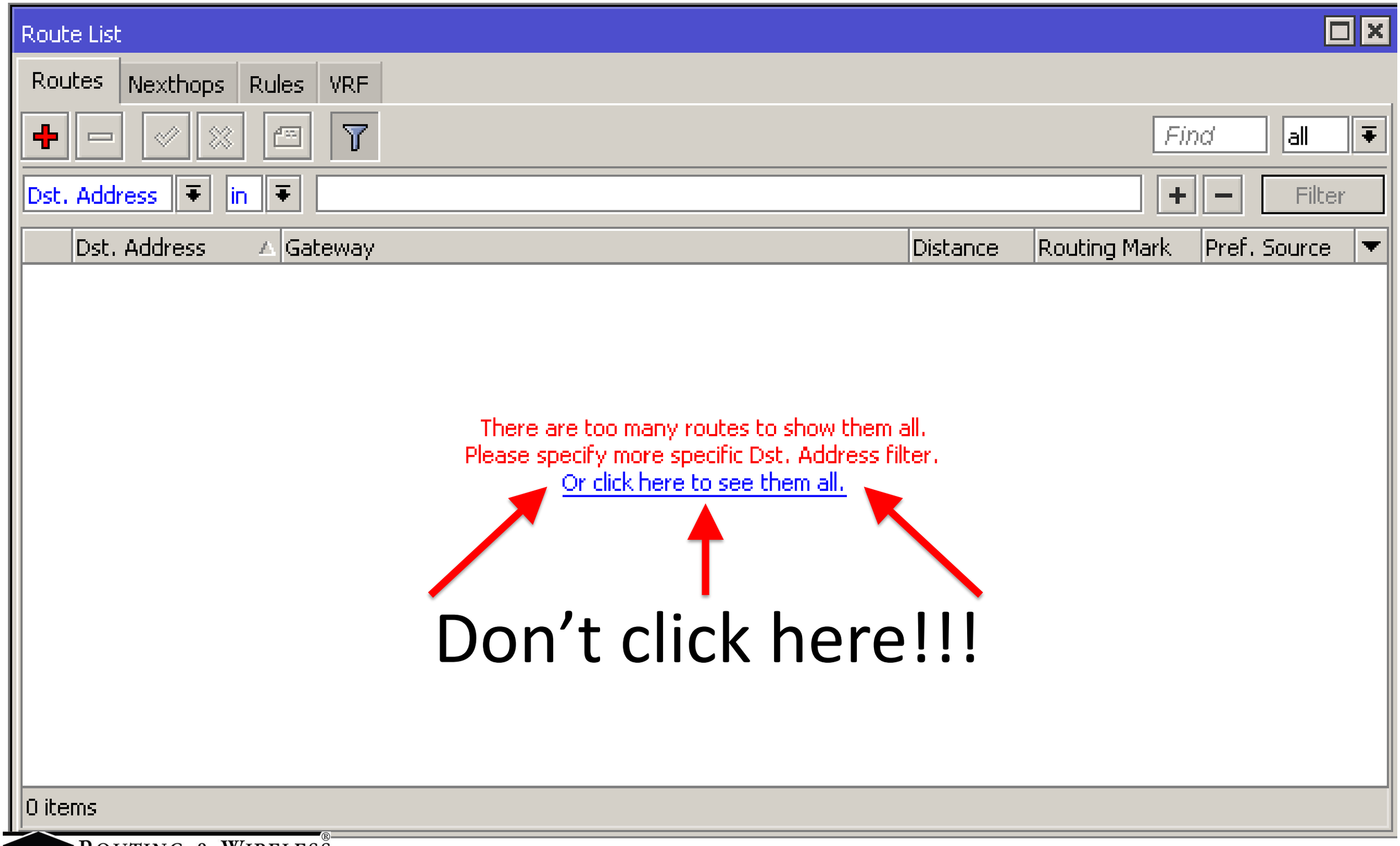
So you can filter them!

http://www.team-cymru.org

Winbox

Do you want to check your routing table using Winbox on your BGP router?

Winbox


Route List
Routes
Nexthops
Rules
VRF
Find
all
Dst. Address
in
Filter
Dst. Address
Gateway
Distance
Routing Mark
Pref. Source
There are too many routes to show them all.
Please specify more specific Dst. Address filter.
Or click here to see them all.
Don’t click here!!!
0 items

Winbox

You’re surely don’t want your router will upload to your winbox client all the “thousand of thousands” routes.

The CLI (Command Line Interface) will do the task better!

Winbox

```
ip route print count-only

routing bgp advertisements print <peer_name>

ip route print detail where <ip_addresss> in
dst-address and bgp

ip route print where received-
from=<peer_name>
```

[admin@BGP] > ip route print detail where 90.189.192.1 in dst-address and bgp

```
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  dst-address=90.189.128.0/17 gateway=xxx.yyy.zzz.kkk
        gateway-status=xxx.yyy.zzz.kkk reachable via  ether5 distance=20
        scope=40 target-scope=10 bgp-as-path="8928,12389,41440" bgp-med=10
        bgp-communities=8928:10403,8928:10901,8928:10902,
                8928:11003,8928:65103,8928:65152,8928:65193,8928:65194,
                41440:1501
        received-from=Peer1

 1  Db  dst-address=90.189.128.0/17 gateway=xxx.yyy.zzz.kkk
        gateway-status=xxx.yyy.zzz.kkk reachable via  ether1 distance=20
        scope=40 target-scope=10 bgp-as-path="28716,1299,12389,41440"
        received-from=Peer2

 2 ADb  dst-address=90.189.192.0/23 gateway=xxx.yyy.zzz.kkk
        gateway-status=xxx.yyy.zzz.kkk reachable via  ether5 distance=20
        scope=40 target-scope=10 bgp-as-path="8928,12389" bgp-med=10
        bgp-communities=8928:10403,8928:10901,8928:10902,
                8928:11003,8928:65103,8928:65152,8928:65193,8928:65194
        received-from=Peer1

 3  Db  dst-address=90.189.192.0/23 gateway=xxx.yyy.zzz.kkk
        gateway-status=xxx.yyy.zzz.kkk reachable via  ether3 distance=20
        scope=40 target-scope=10 bgp-as-path="6939,12389" bgp-med=1
        received-from=Peer3

 4  Db  dst-address=90.189.192.0/23 gateway=xxx.yyy.zzz.kkk
        gateway-status=xxx.yyy.zzz.kkk reachable via  ether1 distance=20
        scope=40 target-scope=10 bgp-as-path="28716,6939,12389"
        received-from=Peer2
```

Additional BGP lectures

My NetFlow presentation:

MUM Ljubljana 2016

NetFlow: what happens in your network?

https://www.youtube.com/watch?v=KEHNxH5Xk3c

http://mum.mikrotik.com/presentations/EU16/presentation_3049_1456752471.pdf

Additional BGP lectures

DDoS attacks detection and mitigation

MUM Ljubljana 2016

by my friend and colleague Wardner Maia

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

Additional BGP lectures

DDoS Mitigation Playbook

MUM Dallas 2016

by my friend and colleague Tom Smyth

Additional BGP lectures

Advanced BGP Filtering with RouterOS

MUM Zagreb 2013

by my friend and colleague Wardner Maia

http://mum.mikrotik.com/presentations/HR13/maia.pdf

What we need more?

Hoping to find some news in ROS 7.XX:

- RPKI implementation;
- A “real” DB for the query into the routes
- 32bit community support

Wrap up

- ✓ **Use RouterOS for BGP**, it's working very well.
- ✓ Improve your **BGP knowledge**.
- ✓ And quoting my friend Tom Smith:

“Buy bigger routers, bitches!”

ROUTING & PEERING THE WORLD
MikroTik™
www.mikrotik.com

Thank you!

Q & A

http://training.grifonline.it
training@grifonline.it