Most underused and overused RouterOS features

OR

My “holy war” against masquerade

MUM, Europe 2017

Objectives

- To help you understand and diagnose most common RouterOS configurations issues

- Show the proper application of RouterOS features to avoid configurations issues

- Encourage you to use latest RouterOS versions and newest features

- Reduce the amount of RouterOS configuration issue emails to support@mikrotik.com!

admin@10.5.112.221 (CCR_middle_1st_from_bottom_SERVER) - WinBox v6.39rc38 on CCR1072-1G-8S+ (tile)

Session Settings Dashboard

Safe Mode | Session: 10.5.112.221 | Time: 11:10:54 | Date: Feb/27/2017 | Uptime: 2d 22:00:10 | Memory: 14.6 GiB | CPU: 24%

- Quick Set
- Interfaces
- Bridge
- PPP
- Mesh
- IP
- IPv6
- Routing
- System
- Queues
- Files
- Log
- Radius
- Tools
- New Terminal
- LCD
- Partition
- Make Supout.rif
- Manual
- New WinBox
- Exit

RouterOS WinBox

Profile (Running)

CPU: total

Start | Stop | Close | New Window

Name	CPU	Usage
total		26.8
ethernet		15.6
networking		5.9
firewall		2.3
management		1.4
profiling		0.9
unclassified		0.5
ppp		0.1
queuing		0.1
firewall-mgmt		0.0
routing		0.0
spi		0.0
winbox		0.0

13 items (1 selected)

Torch

Basic
- Interface: sfp-sfpplus2
- Entry Timeout: 00:00:03 s
- Collect: Src. Address, Dst. Address, MAC Protocol, Protocol, DSCP, Src. Address6, Dst. Address6, Port, VLAN Id

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

Start | Stop | Close | New Window

Eth. Protocol	Prot...	Src.	Dst.	VLAN Id	DSCP	Tx Rate	Rx Rate
800 (ip)	255	172.16.3.236	172.16.47.236			0 bps	34.9 kbps
800 (ip)	255	172.16.3.236	172.16.51.236			0 bps	46.5 kbps
800 (ip)	255	172.16.3.218	172.16.43.218			0 bps	0 bps
800 (ip)	6 (tcp)	172.16.3.237:50000	172.16.35.237:50000			0 bps	0 bps
800 (ip)	255	172.16.3.237	172.16.43.237			0 bps	23.2 kbps
800 (ip)	255	172.16.3.191	172.16.43.191			0 bps	0 bps
800 (ip)	255	172.16.3.236	172.16.43.236			0 bps	34.9 kbps
800 (ip)	255	172.16.3.237	172.16.51.237			0 bps	34.9 kbps
800 (ip)	6 (tcp)	172.16.3.238:50000	172.16.35.238:50000			0 bps	0 bps
800 (ip)	255	172.16.3.203	172.16.43.203			0 bps	0 bps
800 (ip)	255	172.16.3.238	172.16.43.238			0 bps	34.9 kbps

9131 items | Total Tx: 0 bps | Total Rx: 0 bps | Total Tx Packet: 0 | Total Rx Packet: 0

CPU

CPU	Load (...	IRQ (%)	Disk (%)
cpu39	63	9	0
cpu63	54	16	0
cpu45	43	24	0
cpu69	39	39	0
cpu33	37	24	0
cpu0	35	34	0
cpu9	33	19	0
cpu4	31	30	0
cpu42	31	31	0
cpu13	30	30	0
cpu47	30	29	0
cpu26	28	26	0
cpu12	27	25	0
cpu35	27	25	0
cpu36	27	27	0
cpu2	26	26	0
cpu3	26	23	0
cpu25	26	26	0
cpu52	26	25	0
cpu58	26	26	0
cpu59	26	23	0
cpu56	24	21	0
cpu15	25	24	0
cpu16	25	20	0
cpu23	25	19	0
cpu37	25	23	0
cpu41	25	25	0
cpu14	24	24	0
cpu22	24	21	0
cpu24	24	20	0
cpu60	24	24	0
cpu1	23	23	0
cpu8	23	23	0
cpu10	23	22	0
cpu20	23	23	0
cpu21	23	17	0
cpu32	23	23	0
cpu40	23	21	0
cpu54	23	23	0
cpu57	23	23	0
cpu68	23	22	0
cpu70	23	23	0
cpu50	17	17	0
cpu6	22	22	0
cpu11	22	22	0
cpu18	22	22	0
cpu48	22	21	0
cpu55	22	19	0
cpu61	22	22	0
cpu64	22	22	0
cpu71	22	22	0
cpu30	21	21	0
cpu34	21	21	0
cpu44	21	18	0
cpu7	20	20	0
cpu27	20	20	0

72 items (1 selected)

Interface List

Interface | Interface List | Ethernet | PWR | EoIP Tunnel | IP Tunnel | GRE Tunnel | VLAN | VRRP | Bonding | LTE

	Name	Type	Actual MTU	L2 MTU	Tx	Rx	Tx Packet (p/s)	Rx Packet (p/s)	FP Tx
R	ether1	Ethernet	1500	1600	2.3 Mbps	27.2 kbps	203	40	0
R	sfp-sfpplus1	Ethernet	1500	1580	1166.2 Mbps	1208.1 Mbps	99 071	102 597	1166.2 M
R	vlan1	VLAN	1500	1576	1167.1 Mbps	1208.5 Mbps	99 454	102 907	0
DR	<pppoe-a1>	PPPoE Server Binding	1480		238.1 kbps	237.5 kbps	21	21	0
DR	<pppoe-a2>	PPPoE Server Binding	1480		229.0 kbps	239.8 kbps	20	21	0
DR	<pppoe-a3>	PPPoE Server Binding	1480		240.5 kbps	239.8 kbps	21	21	0
DR	<pppoe-a4>	PPPoE Server Binding	1480		233.6 kbps	244.7 kbps	20	21	0
DR	<pppoe-a5>	PPPoE Server Binding	1480		242.9 kbps	230.7 kbps	21	20	0
DR	<pppoe-a6>	PPPoE Server Binding	1480		268.0 kbps	239.8 kbps	27	21	0
DR	<pppoe-a7>	PPPoE Server Binding	1480		242.9 kbps	230.7 kbps	21	20	0
DR	<pppoe-a8>	PPPoE Server Binding	1480		240.5 kbps	239.8 kbps	21	21	0
DR	<pppoe-a9>	PPPoE Server Binding	1480		229.0 kbps	239.8 kbps	20	21	0
DR	<pppoe-a10>	PPPoE Server Binding	1480		240.5 kbps	228.4 kbps	21	20	0
DR	<pppoe-a11>	PPPoE Server Binding	1480		236.0 kbps	235.4 kbps	20	20	0
DR	<pppoe-a12>	PPPoE Server Binding	1480		250.4 kbps	237.8 kbps	22	21	0
DR	<pppoe-a13>	PPPoE Server Binding	1480		233.6 kbps	244.7 kbps	20	21	0
DR	<pppoe-a14>	PPPoE Server Binding	1480		231.3 kbps	230.7 kbps	20	20	0
DR	<pppoe-a15>	PPPoE Server Binding	1480		229.0 kbps	239.8 kbps	20	21	0
DR	<pppoe-a16>	PPPoE Server Binding	1480		240.5 kbps	239.8 kbps	21	21	0

5009 items (1 selected)

Presentation plan

- This presentation will consist of most popular configuration issues sent to support@mikrotik.com
- Examples are compressed/combined/simplified for the purpose of presentation
- Presentation will show problematic configuration and corrected configuration (PLEASE!!! DON'T CONFUSE THEM)

“High Layer7 load”

“High Layer7 load”

- ```
  /ip firewall layer7-protocol
     add name=youtube regexp="^.+(youtube).*\$"
     add name=facebook regexp="^.+(facebook).*\$"
  ```
- ```
  /ip firewall filter
     add action=drop chain=forward layer7-
  protocol=facebook
     add action=drop chain=forward layer7-
  protocol=youtube
  ```

WRONG!!!

Analysis of the problem

- Problem:
  - High CPU load, increased latency, packet loss, jitter, youtube and facebook is not blocked
- Diagnosis:
  - “/tool profile” high layer7 load
- Reason:
  - Each connection is rechecked over and over again
  - Layer7 is checked in the wrong place and against all traffic

Layer7

- Layer7-protocol is a method of searching for patterns in **ICMP/TCP/UDP** streams

- On trigger Layer7 collects next 10 packets or 2KB of a connection and searches for the pattern in the collected data

- All Layer7 patterns available on the Internet are designed to work only for the first 10 packets or 2KB of a connection.

Correct implementation

- ```
  /ip firewall mangle
  add action=mark-connection chain=prerouting protocol=udp
   dst-port=53 connection-mark=no-mark layer7-
  protocol=youtube new-connection-mark=youtube_conn
  passthrough=yes

  add action=mark-packet chain=prerouting connection-
  mark=youtube_conn new-packet-mark=youtube_packet
  ```
- ```
  /ip firewall filter
  add action=drop chain=forward packet-mark=youtube_packet
  add action=drop chain=input packet-mark=youtube_packet
  ```

  (and same set for facebook)

“Queues don’t work properly”

“Queues don’t work properly”

- ```
  /ip address
  add address=10.0.0.1/24 interface=local-one
  add address=10.0.1.1/24 interface=local-two
  ```
- ```
  /ip firewall filter
  add chain=forward action=fasttrack-connection
  connection-state=established,related
  add chain=forward action=accept connection-
  state=established,related
  ```
- ```
  /queue simple
  add max-limit=10M/10M dst=10.0.0.2/32
  add max-limit=10M/10M dst=10.0.0.3/32
  add max-limit=10M/10M dst=10.0.0.4/32
  ```

WRONG!!!
  ```

Analysis of the problem

- Problem:
  - Queues works only when “/tool torch” is running, or when fasttrack is disabled, but then captures only download traffic, traffic between local networks are also limited
- Diagnosis:
  - Counters on queues, and fasttrack-connection rule
- Reason:
  - Fasttrack rule is specified for all traffic
  - Simple queue target must be specified

FastTracked

- Conntrack entries now have “Fasttracked” flag
- Implemented as “fasttrack-connection” action for firewall filter/mangle
- Packets from “Fasttracked” connections are allowed to travel in FastPath
- Works only with IPv4/TCP and IPv4/UDP
- Traffic traveling in FastPath will be invisible to other router facilities (firewall, queues, etc)
- Some packets will still follow the regular path to maintain conntrack entries

Simple queue “target”

- “target” option is the only option that determines direction of a simple queue
- If target is not specified (is 0.0.0.0/0) all traffic will be captured in download part of the queue, as everything is download for 0.0.0.0/0
- “dst” option is only an additional filter, it doesn’t determine the direction

Correct implementation

- /ip firewall filter
  add chain=forward action=fasttrack-connection connection-state=established,related **in-interface=local-one out-interface=local-two**
  add chain=forward action=fasttrack-connection connection-state=established,related **in-interface=local-two out-interface=local-one**
  add chain=forward action=accept connection-state=established,related
- /queue simple
  add max-limit=10M/10M **target=10.0.0.2/32**
  add max-limit=10M/10M **target=10.0.0.3/32**
  add max-limit=10M/10M **target=10.0.0.4/32**

“High CPU load on PPPoE server”

“High CPU load on PPPoE server”

- 3000 pppoe-clients in 10.0.0.0/20 network
- Connected via 172.16.x.0/24 networks to other PPPoE servers with 10.x.0.0/20 PPPoE client network.
- All PPPoE servers and gateway in the same backbone area with redistribute connected routes

```
/routing ospf network
   add network=172.16.1.0/24 area=backbone
   add network=10.0.0.0/20 area=backbone
```

WRONG!!!

Analysis of the problem

- Problem:
  - CPU overloaded, PPPoE clients disconnect, clients can’t reach target speeds, sometimes can’t connect to the device
- Diagnosis:
  - /tool profile shows “routing” process holding one CPU core 100% all the time, all other cores sometimes can also reach 100% with “ppp” and “networking” processes
- Reason:
  - OSPF is spammed with PPPoE client /32 route updates

OSPF and PPPoE

- All dynamic routing protocols (more precisely - routing table updates and protocol calculations) are limited to a single core
- Every time a pppoe-client connects or disconnects it creates or deletes a /32 route. If that route is a part of an OSPF network, OSPF update is initiated
- Every time a pppoe-client connects or disconnects pppoe-interface is added to or removed from OSPF interfaces, that also initiates OSPF update

Passive OSPF interfaces and stub areas

- Stub areas allow to reduce the amount of routing information flooded into areas - external routes are not flooded into and throughout a stub area, default route is used
- Area ranges are used to aggregate routing information on area boundaries, allows to create only one summary LSA for multiple routes and send only single advertisement into adjacent areas
- Passive interface flag if enabled, excludes interface from OSPF protocol communication

Correct implementation

- /routing ospf area
  add area-id=0.0.0.1 authentication=none name=pppoe1 type=stub
- /routing ospf network
  add area=pppoe1 network=10.0.0.0/20
- /routing ospf area range
  add advertise=yes area=pppoe1 range=10.0.0.0/20
- /routing ospf interface
  add interface=all passive=yes

“High CPU load on PPPoE server”

“High CPU load on PPPoE server”

- 3000 pppoe-clients in 10.0.0.0/20 network
- Static public IP address on public interface
- Masquerade rule
- No other firewall

WRONG!!!

Analysis of the Problem

- Problem:
  - CPU overloaded, PPPoE clients disconnect, clients can’t reach target speeds, sometimes can’t connect to boards.
- Diagnosis:
  - /tool profile shows “firewall” process dominating CPU load
- Reason:
  - Improper use of masquerade

Masquerade

- Firewall NAT action=masquerade is unique subversion of action=srcnat, it was designed for specific use in situations when public IP can randomly change - when public IP is dynamic.
- Every time an interface disconnects and/or its IP address changes, router will search and purges connection tracking from connections related to that interface, to improve recovery time

Correct implementation

- ```
  /ip firewall nat
  add action=src-nat chain=srcnat out-
  interface=<Public> to-addresses=<Public_IP>
  ```
  ```

“Local IP leaking to public network”

“Local IP leaking to public network”

- Multi gateway device with policy routing and failover
- Static public IP addresses on public interfaces
- Masquerade rules on every public interface

WRONG!!!

Analysis of the problem

- Problem:
  - After failover happens packets with private IP as source address leak out to public network.
- Diagnosis:
  - /tool sniffer
- Reason:
  - Improper use of masquerade or insufficient amount of safeguards

Masquerade

- On disconnect, all related connection tracking entries are purged
- Next packet from every purged connection will come into firewall as connection-state=new, and, packet will be routed out via alternative route thus creating new connection entry
- When primary link comes back, routing is restored over primary link, so packets that belong to existing connections are sent over primary interface without being masqueraded

Correct implementation

- Use action=src-nat instead of action=masquerade where it is possible
- Drop connection-state=invalid packets
- Drop connection-state=new connection-nat-state=!dstnat packets from public interface
- Creating backup “blackhole” route for each routing-mark

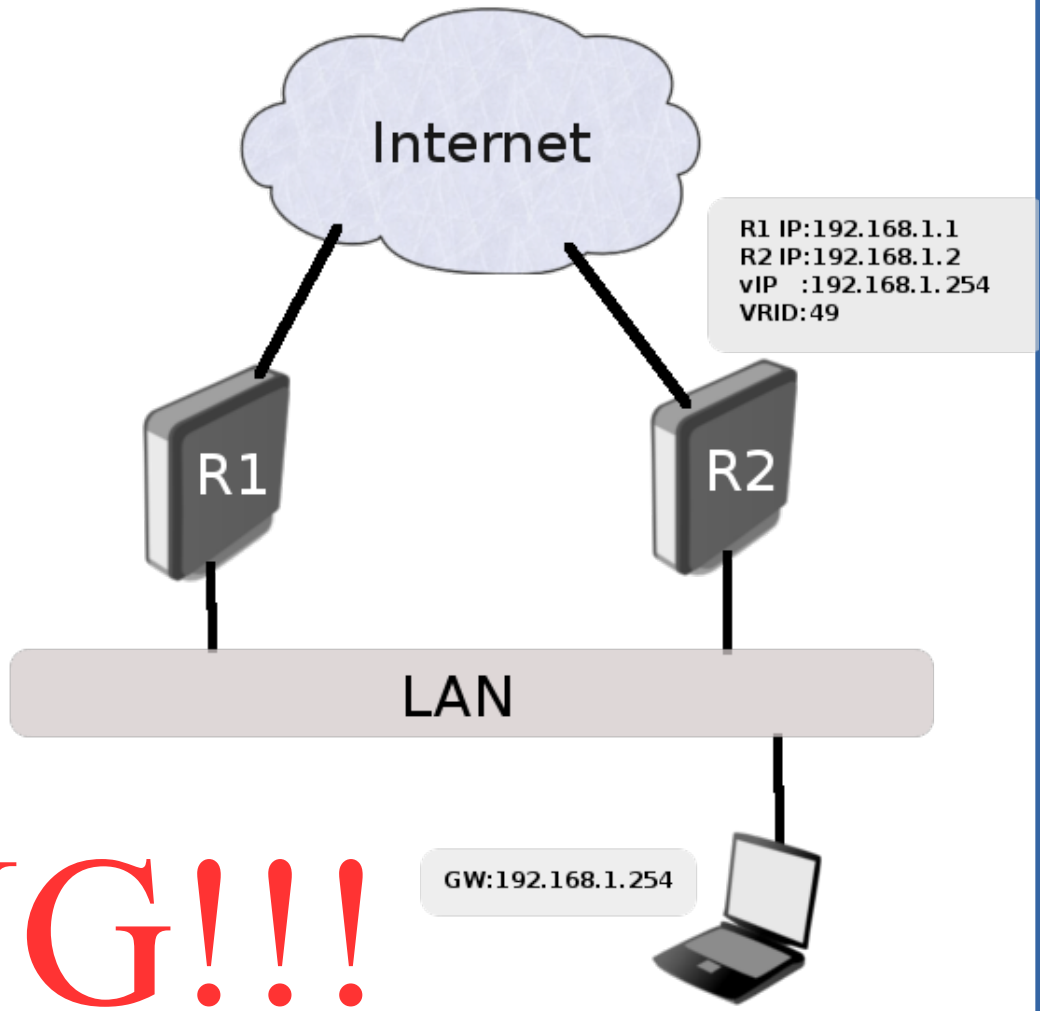
“VRRP and routing problems”

“VRRP and routing problems”

- /ip address add
  address=192.168.1.1/24
  interface=ether1

  /interface vrrp add
  interface=ether1 vrid=49
  priority=254

  /ip address add
  address=192.168.1.254/24
  interface=vrrp1



WRONG!!!

Analysis of the problem

- Problem:
  - Routing doesn’t work properly, Fastpath/fasttrack doesn’t work, networking process have high load
- Diagnosis:
  - Routing table, interface statistics counters
- Reason:
  - VRRP interface creates routing conflict, by having 2 interfaces with 2 identical subnets on them

Correct implementation

- `/ip address add address=192.168.1.1/24 interface=ether1`

  `/interface vrrp add interface=ether1 vrid=49 priority=254`

  `/ip address add address=192.168.1.254/32 interface=vrrp1`

“DNS cache”

“DNS cache”

- `/ip dns`
  `set allow-remote-requests=yes servers=8.8.8.8`
- `/ip firewall nat`
  `add action=masquerade chain=srcnat out-interface=Internet`
- `/ip firewall filters`
  `add action=fasttrack-connection chain=forward connection-state=established,related`
  <nothing more>
- Public IP on the Internet interface

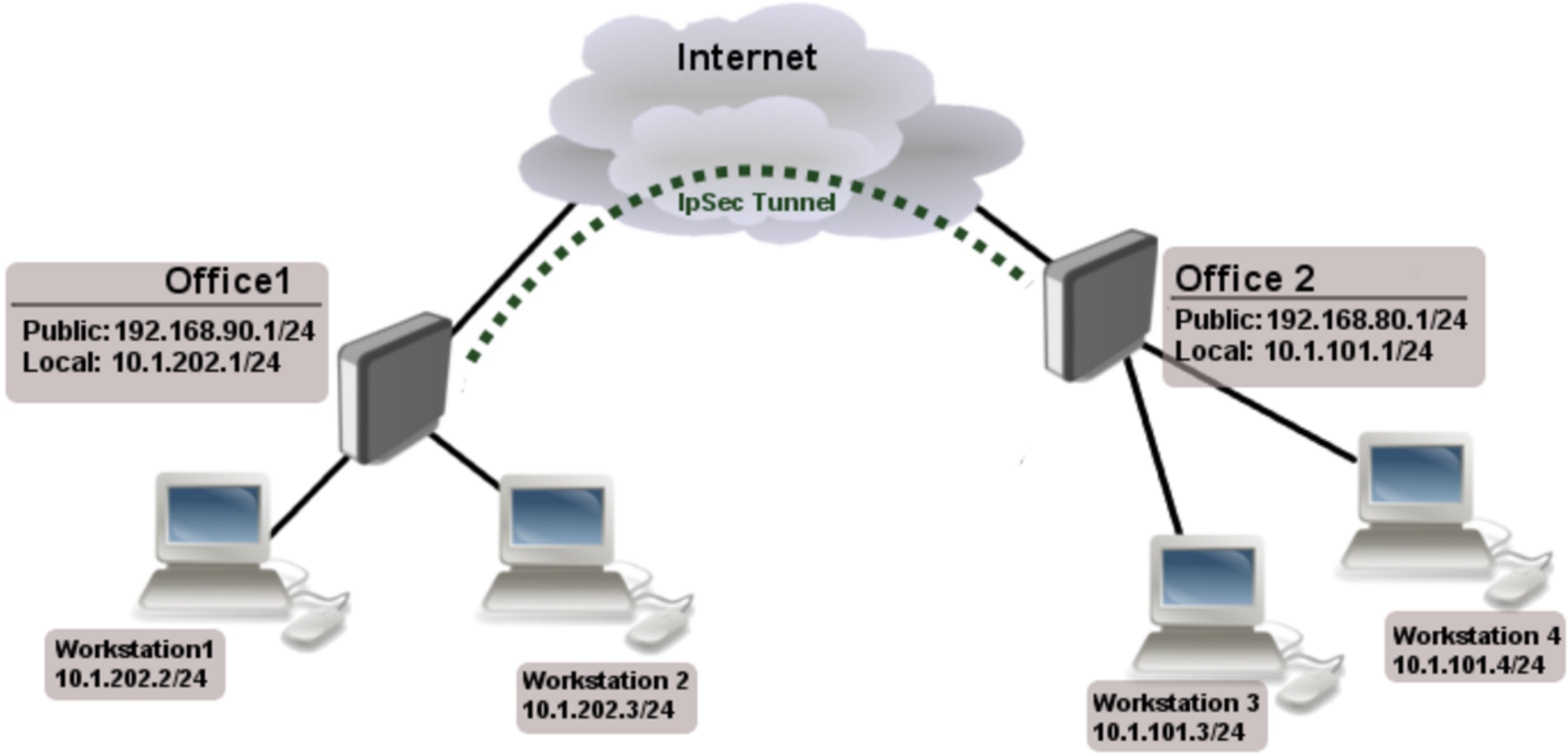
WRONG!!!

Analysis of the problem

- Problem:
  - High CPU load, high amount of unknown traffic on public interface
- Diagnosis:
  - /tool torch, /tool profile “dns” load
- Reason:
  - Your router is used as Open DNS resolver. It answers recursive queries for hosts outside of its domain and is utilized in DNS Amplification attacks

Correct implementation

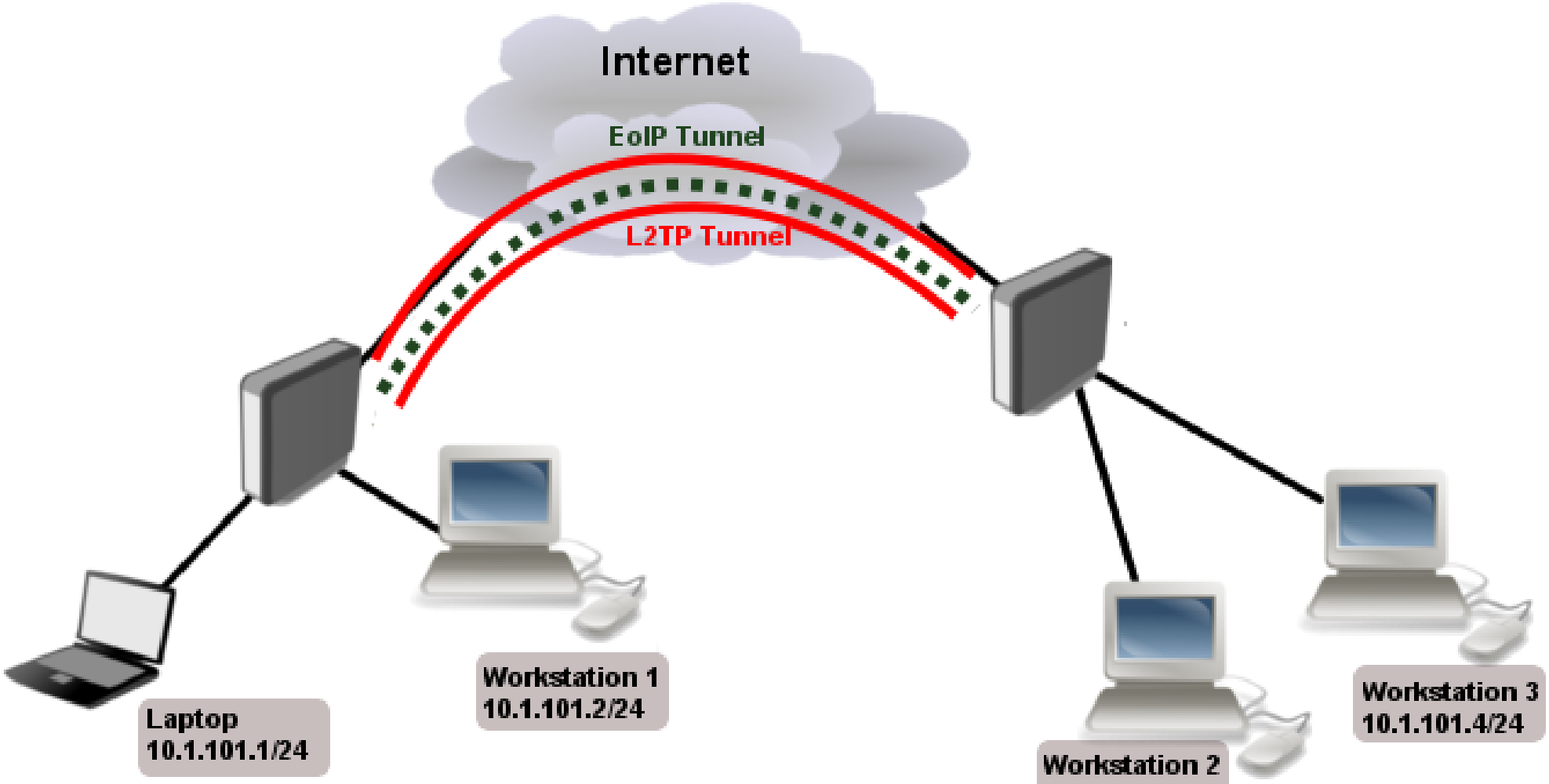
- ```
  /ip firewall filter
  add action=reject chain=input dst-port=53
  protocol=udp reject-with=icmp-port-unreachable
  add action=reject chain=input dst-port=53
  protocol=tcp reject-with=icmp-port-unreachable
  ```

  (and rest of the firewall filter)
  ```

“IPSec tunnel doesn’t work”

“IPSec tunnel doesn’t work”



- Simple masquerade on both routers

WRONG!!!

Analysis of the problem

- Problem:
  - IPsec packets are rejected, tunnel cannot be established
- Diagnosis:
  - /tool sniffer
- Reason:
  - NAT rules are changing src-address of encrypted packets, scr-address doesn’t correspond to IPsec policy on opposite end

Raw table

- Firewall RAW table allows to selectively bypass or drop packets before connection tracking that way significantly reducing load on CPU
- If packet is marked to bypass connection tracking
  - packet de-fragmentation will not occur
  - NAT will be skipped
  - matchers that depend on connection tracking will not trigger (fasttrack-connection, mark-connection, layer7 etc.)
  - will have connection-state=untracked

Correct implementation

- ```
  /ip firewall raw
  add action=notrack chain=prerouting src-
  address=10.1.101.0/24 dst-address=10.1.202.0/24

  add action=notrack chain=prerouting src-
  address=10.1.202.0/24 dst-address=10.1.101.0/24
  ```
  ```

“Securely bridge two local networks”

“Securely bridge two local networks”

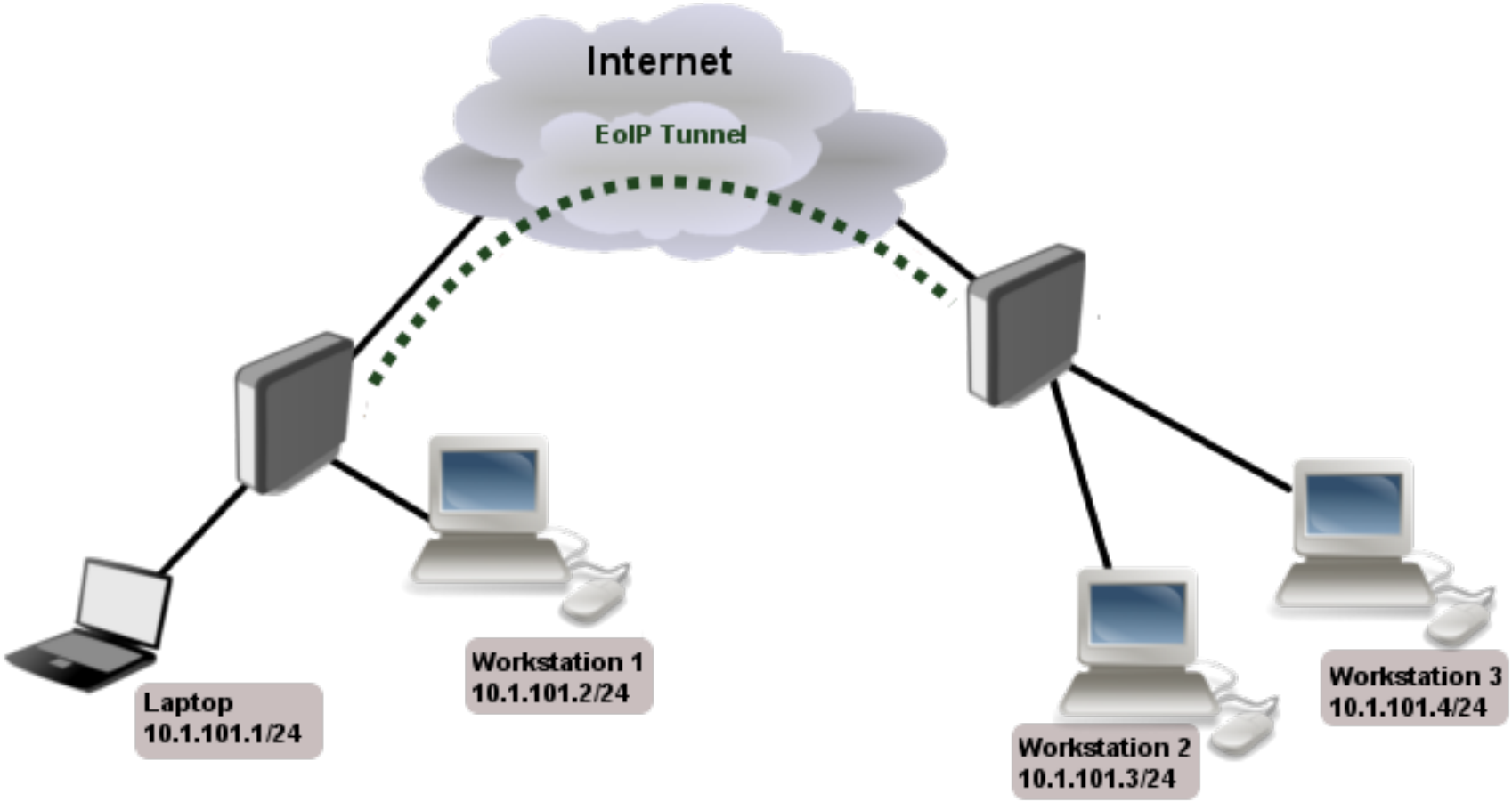


WRONG!!!

Analysis of the problem

- Problem:
  - Web pages very slow to open, slow download speeds, strange suspicion that competition knows your secret information :)
- Diagnosis:
  - /tool bandwidth-test, /tool ping with different packet sizes
- Reason:
  - PPTP/L2TP is not secure anymore, severe packet overhead from two tunnel overheads, fragmentation, because of reduced MTU

Correct implementation



- `/interface eoip set ipsec-secret=`

CCR HW encryption acceleration

- Completely new driver for hardware encryption accelerator in RouterOS v6.39 for CCR devices
- Solves out-of-order issue for encrypted traffic and improves performance (1400 byte UDP packets):
  - CCR1072 from up to 9,2Gbps to up to 13,8Gbps
  - CCR1036 from up to 3,4Gbps to up to 7Gbps
  - CCR1009 from up to 1,5Gbps to up to 2,2Gbps

Questions!!!