

# Pokročilejší síťování v Linuxu

Ondřej Caletka



6. října 2019



Uvedené dílo podléhá licenci Creative Commons Uveďte autora 3.0 Česko.

# O sdružení CESNET



	n × 100 Gb/s		100 Gb/s
	n × 10 Gb/s		10 Gb/s
	uzel (PoP)		1-2,5 Gb/s
	uživatel (user)		<1 Gb/s



MetaCentrum



UltraGrid

- přijďte na stánek CESNETu
  - prohlédněte si dema
  - nastavte správně eduroam
  - zahrajte si hru The Catch Nano
- přihlaste si eNews, ať vám nic neunikne
- hledáme posily
  - vývojář systémů pro sledování provozu a bezpečnosti sítě
  - vývojář cloudové platformy pro rozsáhlé vědecké výpočty
  - pracovník pro správu Linuxových systémů

**12:00** Radko Krkoš – CESNET PassiveDNS: Ako byť archeológom a bezpečnostným analytikom naraz?

# The Catch!

Začínáme zítra 7. října ve 14 hodin!



<https://www.thecatch.cz/>

- 1 Rychlý přehled
- 2 Více sítí zároveň
- 3 Praktický příklad – Wireguard
- 4 Demo

# Rychlý přehled

- zastaralé nástroje: `ifconfig`, `route`, `netstat`, `brctl`
- moderní nástroje: `ip`, `ip route`, `ss`, `ip link`
- všechny změny jsou neperzistentní; zajištění perzistence se liší podle distribuce
- neexistuje žádný *reset* síťového stacku

**ip link** vytváření a editace síťových rozhraní

**ip address** nastavování IP adres

**ip route** práce se směrovacími tabulkami

**ip rule** práce pravidly směrování (policy based routing)

**ss** statistika otevřených soketů

**bridge** nastavení mostů

**tc** nastavení front



## Příklad ruční konfigurace sítě

```
# ip link set dev eth0 up
# ip addr add dev eth0 192.168.1.2/24
# ip rou add default via 192.168.1.1
# echo "nameserver 4.2.2.1" >/etc/resolv.conf
```

- `nameserver <IP>` adresa serveru (max. 3×)
- `domain <d>` místní doména
- `search <d> <d>...` prohledávací seznam
- `options rotate ...` náhodně měnit použitý DNS server
- `options ... edns0` používat EDNS0 (např. DNSSEC)

Soubor může být spravován Network Managerem, utilitou `resolvconf` nebo jinak.

- userspace implementace stub resolveru
- dbus, glibc a DNS API
- resolvování lokálních jmen z /etc/hosts, jména `_gateway`
- jména bez tečky jsou resolvována pomocí LLMNR
- je možné směřovat na různé DNS servery v závislosti na dotazovaném jménu
- podpora validace DNSSEC

# Konfigurace IPv6

- minimální implementace vestavěná v jádře
- dnes často vypnutá ve prospěch komplexnější implementace v userspace
- automatické nastavení směrování podle ohlášení směrovačů
- automatické nastavení IP adres pomocí SLAAC
- vyčištění pomocí `ip -6 addr flush dev eth0 scope global`
- konfigurace pomocí voleb `sysctl`:

`net.ipv6.conf.eth0.accept_ra` povolí zpracování RA

`net.ipv6.conf.eth0.autoconf` povolí automatickou konfiguraci adres  
(=2 i v režimu směrovače)

- `ping` posílá ICMP echo-request
- `traceroute` hledá cestu pomocí UDP, TCP, nebo ICMP
- `mtr` lepší traceroute
- `arping` objevuje stanice na segmentu pomocí ARP
- `host` provádí DNS dotazy
- `iftop` vizualizuje toky na rozhraní
- `tcpdump` zaznamenává a analyzuje obsah přenášených zpráv

# Více sítí zároveň

# Připojení k více sítím zároveň

- pro komunikaci se sousedy není problém
- nelze mít víc výchozích bran
- nelze mít víc DNS serverů

## Jednoduché řešení pomocí `ip route from`

```
# ip link set dev eth1 up
# ip addr add dev eth1 172.17.1.2/24
# ip rout add default from 172.17.1.2 via 172.17.1.1
```

# Která adresa se použije?

## Originující provoz používající anonymní soket

- 1 místní adresu neznáme
- 2 vybereme vzdálenou adresu
- 3 hledáme cestu ke vzdálené adrese, místní stále neznáme
- 4 zvolíme vhodnou místní adresu

## Terminovaný provoz a pojmenovaný soket

- 1 místní adresa je pevně určena a neměnná
- 2 hledáme cestu ke vzdálené adrese



- lze vytvořit víc směrovacích tabulek
- pravidla `ip rule` vybírají, která tabulka bude použita
- tabulky lze pojmenovat čísly nebo v `/etc/iproute2/rt_tables`

## Příklad policy routingu

```
# ip link set dev eth1 up
# ip addr add dev eth1 172.17.1.2/24
# ip rou add default via 172.17.1.1 table 123
# ip rule add from 172.17.1.2 lookup 123
```

## Základní směrovací tabulky

**local** cache cest, např. pro objevování MTU trasy

**main** výchozí směrovací tabulka

**default** prázdná tabulka s nejnižší prioritou

## Možná pravidla policy routingu

```
[ not ] [ from PREFIX ] [ to PREFIX ] [ tos TOS ] [ fwmark  
FWMARK[/MASK] ] [ iif STRING ] [ oif STRING ] [ pref NUMBER  
] [ uidrange NUMBER-NUMBER ] [ ipproto PROTOCOL ] [ sport [  
NUMBER | NUMBER-NUMBER ] ] [ dport [ NUMBER | NUMBER-NUMBER  
] ]
```

# Virtual Routing and Forwarding (-lite)

- možnost rozdělit jednotlivá rozhraní do virtuálních směrovačů
- provádí se pomocí speciálního *master* rozhraní typu *vrf*
- příslušná *slave* rozhraní používají společnou samostatnou směrovací tabulku
- IP adresy se mohou opakovat v různých doménách VRF
- běžné síťové služby provoz VRF nevidí
- týká se pouze IP; linkové protokoly nadále používají všechna rozhraní

---

<https://www.kernel.org/doc/Documentation/networking/vrf.txt>

# Network namespaces

- stavební prvek linuxových kontejnerů
- lze použít i samostatně
- každá síťová karta je právě v jednom NS

## Přestěhování eth0 do vlastního NS

```
# ip netns add testovani
# ip link set eth0 netns testovani
# ip netns exec testovani bash
...
# ip netns delete testovani
```

# Na co se často zapomíná

- směrování je ve výchozím stavu vypnuto
- privátní adresy na internet nepatří
- nasměrované, ale nepoužívané adresy zahazujeme

```
# sysctl net.ipv4.ip_forward=1
# ip route add unreachable 192.168.0.0/16
# ip -6 route add unreachable 2001:db8::/32
```

# Praktický příklad – WireGuard

- nový VPN protokol přenášející IP datagramy v UDP zprávách
- mimo jiné velmi efektivní implementace v jádře Linuxu
- autentizace veřejným klíčem (jako SSH), vazba klíče a IP adres
- spolupráce s *namespaces*
- elegantní řešení problému „cyklu VPN“
- nově veřejná bezplatná WireGuard VPN – Cloudflare Warp

## Problém „cyklu VPN“

- většina uživatelů chce do VPN směřovat veškerý provoz
- data samotné VPN však musí procházet mimo VPN

## Obvyklé řešení

- výchozí brána zůstane původní
- přidají se dva záznamy  $0/1$  a  $128/1$  směrem do VPN
- přidá se záznam s cílovou adresou VPN koncentrátoru a původní výchozí branou



## Problém obvyklého řešení

- při výpadku fyzického rozhraní se VPN zacyklí
- obnovení neproběhne automaticky
- provoz k IP adrese VPN koncentrátoru neprochází VPN

## Řešení použité ve WireGuard

```
# ip -6 route add ::/0 dev wgcf table 1234
# wg set wgcf fwmark 51820
# ip -6 rule add not fwmark 51820 table 1234
# ip -6 rule add table main suppress_prefixlength 0
```

- UDP soket pro předávání šifrovaného provozu je vytvořen při založení tunelu
- rozhraní tunelu je možné přesunout do jiného *namespace*
- dva scénáře použití:
  - 1 chráněný *namespace* s přístupem pouze k tunelu
  - 2 veřejný *namespace* s přístupem k fyzickým rozhraním

- veřejná VPN postavená na (drobně upraveném) WireGuardu
- dostupná prostřednictvím aplikace 1.1.1.1 pro Android a iOS
- plně podporuje IPv6 vně i uvnitř tunelu
- používá NAT pro IPv4 i IPv6 😬
- kvůli unikátním lokálním IPv6 adresám **není IPv6 nikdy preferováno**

---

Root.cz: Cloudflare spustil bezplatnou VPN pro všechny. IPv6 používá NAT

# Cloudflare Warp na počítači



**Jay Freeman (saurik)**

@saurik



Today, Cloudflare made WARP, their VPN service with an unlimited free tier, available to everyone; it only has clients for iOS and Android, but the protocol they are using seems to be off-the-shelf Wireguard, so you can connect from macOS! Run this script: [cache.saurik.com/twitter/wgcf.sh](https://cache.saurik.com/twitter/wgcf.sh)

[Přeložit Tweet](#)

6:17 odp. · 25. 9. 2019 · [Twitter for iPhone](#)



link

cesnet  
.....



**Ondřej Caletka**

@Oskar456



Odpověď uživateli [@saurik](#)

Thank you for your great work! I have forked your script to Linux, where it only generates config file for the wg-quick(8) command. I also added [#IPv6](#) support. Works pretty well!  
[gist.github.com/oskar456/594f1...](https://gist.github.com/oskar456/594f1...)

link

# Demo

Děkuji za pozornost

**Ondřej Caletka**  
**Ondrej.Caletka@cesnet.cz**  
**[https://Ondřej.Caletka.cz](https://Ondrej.Caletka.cz)**

