Zálohujeme připojení k IPv6 Internetu

Radek Zajíc

LinuxDays, 7. října 2017

about:me

2008 – 2012: Seznam.cz

IPv6 na webu

2012 – 2017: T-Mobile CZ

IPv6 na DSL a v mobilní síti

2017: Showmax.com

Desítky gigabitů provozu v Africe a Polsku.

Máme rádi open-source. ☺

Od připojení k Internetu očekávám…

spolehlivost
stabilitu
rychlost

Přístup k Internetu se stává samozřejmostí – podobně jako elektřina, teplo, voda. Když vám v Praze na dva týdny vypne teplárna horkou vodu, nejste úplně nadšeni. Podobně když dojde k výpadku elektřiny nebo neteče pitná voda. S Internetem to začíná být podobné. Na připojení se potřebujete spolehnout (protože přes internet vyřizujete pracovní věci, užíváte si zábavu, jste v kontaktu s lidmi), informace a data potřebujete získat v rozumném čase (tedy rychle), připojení by mělo mít určitou kvalitu a být stabilní.

Je rozdíl, když jsem…

velká a střední firma, ISP - $$$

RIPE, AS, PA IP(v6), dvě vlákna, BGP4+, Cisco/Juniper/Brocade/Arista

VS.

malá firma, domácnost - $

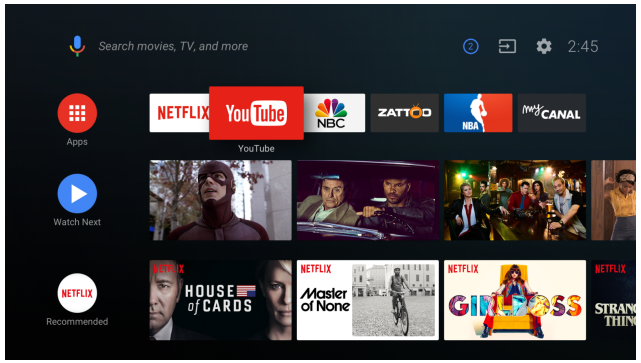
Turris, Linux, kabelovka, DSL, Wi-Fi, IP od providerů, vlastní skripty

Když jste velká firma, za spolehlivost si zaplatíte – ale dostanete ji. Jako koncový uživatel máte situaci o dost složitější. Prostředky velkých firem (finanční ani technické) nemáte k dispozici.

Kdysi byla situace jednodušší

Dokud byl v domácnosti/firmě jeden počítač, byla situace mnohem jednodušší. Jedna telefonní/kabelová linka, jedna Wi-Fi linka, jeden počítač. Jeden uživatel. Neměli jste za zády adolescenty, kteří by křičeli, že nejde internet, protože jste pravděpodobně žádný internet neznali (na rozdíl od adolescentů).

Typická domácí síť dnes

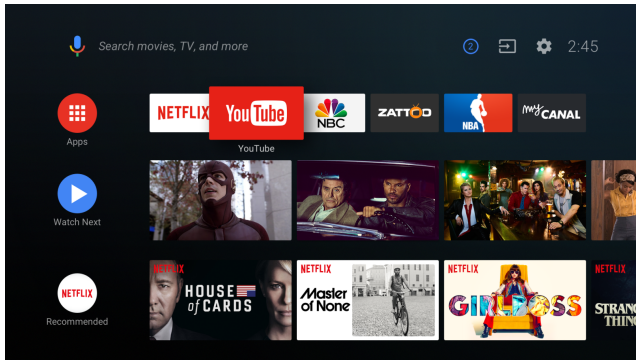


Dnes vypadá domácí síť úplně jinak. Smartphony, tablety, notebooky, připojené televize, ledničky, senzory a další, to vše naškálováno podle počtu členů domácnosti, vyžadují prakticky neustále funkční konektivitu k Internetu. Pokud konektivitu nemáte, dříve nebo později se ozve někdo z rodiny, komu to začne vadit.

Domácí síť, když nefunguje připojení

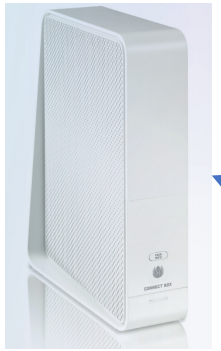
Pokud se připojení rozsype, dřív nebo později se u nás doma ozve „Radku, nejde wi-fi!“ (Co na tom, že Wi-Fi vysílá, to jen spadly uplinky do Internetu. Zbytečné vysvětlovat, uživatel to vidí jinak.) Na obrázcích ukázka, jak by to asi dnes namaloval Edvard Munch a jak ve skutečnosti vypadá ta zákeřná hruška, která vám resetuje spojení…

Typická domácí síť zítra



Pořídíte si druhou přípojku od jiného poskytovatele a na jiné technologii. Budete doufat, že nevypadnou obě naráz. A správným skriptem zařídíte, že uživatelé budou mít vždy funkční konektivitu – budete přepínat mezi linkami.

Jedna přípojka – IPv4



PRIMARY

192.0.1.2

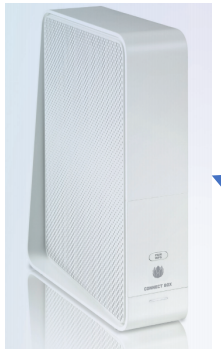
GATEWAY
S NATEM

192.168.1.2

Ve světě IPv4 jsme si zvykli na určitou pohodu a pohodlí. Když běží primární přípojka, máme nastavenou default gateway na ni. Při přepnutí na záložní linku prostě jen změníme default gateway – a provoz se díky NATům přelije na nového poskytovatele. Na zařízeních v lokální síti se nic nemění (DNS určitě nepoužíváme operátorské.)

Ve světě IPv4 jsme si zvykli na určitou pohodu a pohodlí. Když běží primární přípojka, máme nastavenou default gateway na ni. Při přepnutí na záložní linku prostě jen změníme default gateway – a provoz se díky NATům přelije na nového poskytovatele. Na zařízeních v lokální síti se nic nemění (DNS určitě nepoužíváme operátorské.)

Dvě přípojky – IPv4



PRIMARY

192.0.1.2

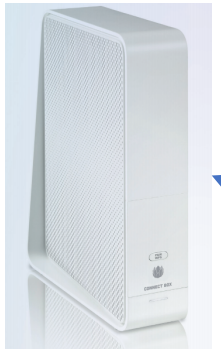
192.168.1.2

GATEWAY
S NATEM

192.2.8.5

SECONDARY

Dvě přípojky – IPv6



WAN: 2001:db8:dead:beef::/64
Delegovaná: 2001:db8:face:b000::/56

PRIMARY

GATEWAY
s DHCPv6 PD

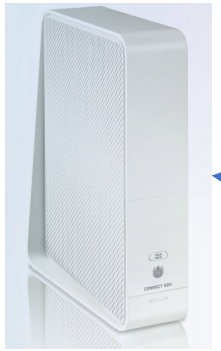
?

WAN: 2001:db8:bad:babe::/64
Delegovaná: 2001:db8:babe:1000::/56

SECONDARY

S IPv6 se situace stává složitější. Od každého poskytovatele získáte vlastní IPv6 prefix pro WAN (rozhraní k operátorovi) a zároveň si můžete požádat o prefix adres pro LAN (vaši síť). Pokud máte operátorů víc, musíte si pro lokální síť žádat o adresy u každého z nich. To je naštěstí řeší jedna z funkcí DHCPv6 – delegace prefixů (Prefix Delegation). Jenže – jaké adresy mají používat zařízení v lokální síti?

Dvě přípojky – každý chvilku tahá pilku



WAN: 2001:db8:dead:beef::/64
Delegovaná: 2001:db8:face:b000::/64

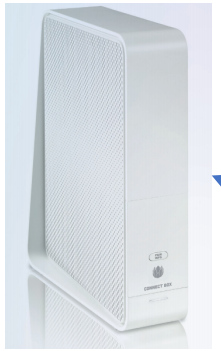
PRIMARY

WAN: 2001:db8:bad:babe::/64
Delegovaná: 2001:db8:babe:1000::/56

SECONDARY

Můžete vypustit router z cesty a nechat modemy/routery operátorů, aby vám samy oznamovaly do sítě prefixy. Pokud ztratí konektivitu, měly by podle RFC 7084 (bod 3.2.1) přestat oznamovat samy sebe jako gateway a zařízení v LAN by skrz ně tudíž neměla posílat data. (Nevýhoda: routery jsou obvykle tupé a nejde jim nastavit priorita. To vadí v případě, kdy je jedna linka o dost slabší. Nevýhoda 2: tohle řešení nejde použít pro LAN, kde chcete mít i IPv4.)

Dvě přípojky – IPv6 a NPT



WAN: 2001:db8:dead:beef::/64
Delegovaná: 2001:db8:face:b000::/56

PRIMARY

fde4:8dba:82e1::/64

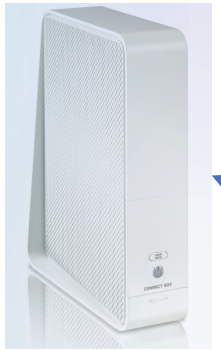
GATEWAY
s DHCPv6 PD + NPT

WAN: 2001:db8:bad:babe::/64
Delegovaná: 2001:db8:babe:1000::/56

SECONDARY

NPT: RFC 6296
ULA: RFC 4193

Dvě přípojky – oznamování obou prefixů



WAN: 2001:db8:dead:beef::/64
Delegovaná: 2001:db8:face:b000::/56

PRIMARY

GATEWAY
s DHCPv6 PD

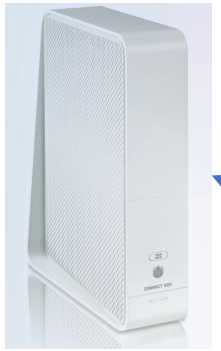
2001:db8:face:b000::/64
2001:db8:babe:1000::/64

Můžete z routeru oznamovat obě delegované sítě - počítače v LAN si pak budou vybírat samy zdrojovou adresu a data vám nekontrolovatelně potečou jednou nebo druhou stranou.

WAN: 2001:db8:bad:babe::/64
Delegovaná: 2001:db8:babe:1000::/56

SECONDARY

Dvě přípojky – vlastní řízení provozu



WAN: 2001:db8:dead:beef::/64

Delegovaná: 2001:db8:face:b000::/56

PRIMARY

2001:db8:face:b000::/64

GATEWAY
s DHCPv6 PD

Můžete z routeru oznamovat ale i jen jednu síť - tu, kterou podle svých kritérií označíte jako vhodnější. Pokud je funkční primární linka, budete oznamovat její prefix!

WAN: 2001:db8:bad:babe::/64

Delegovaná: 2001:db8:babe:1000::/56

SECONDARY

Dvě přípojky – vlastní řízení provozu

PRIMARY

WAN: 2001:db8:dead:beef::/64

Delegovaná: 2001:db8:face:b000::/56

GATEWAY
s DHCPv6 PD

2001:db8:babe:1000::/64

WAN: 2001:db8:bad:babe::/64

Delegovaná: 2001:db8:babe:1000::/56

SECONDARY

Pokud hlavní linka selže, na routeru zjistíte nefunkčnost pomocí skriptů. Do lokální sítě oznámíte, že původní IPv6 prefix už není dostupný a naopak začnete oznamovat nový prefix. Počítače **časem** původní adresy zapomenou a **jiným časem** začnou používat adresy z rozsahu záložní linky.

Dvě přípojky – source routing

```
# ip -6 rule
0:    from all lookup local
32352:      from 2001:db8:bad:babe::/64 lookup 4
32353:      from 2001:db8:babe:1000::/56 lookup 4
32358:      from 2001:db8:face:b000::/60 lookup 5
32590:      from 2001:db8:dead:beef::/64 lookup 5
32766:      from all lookup main
# ip -6 route show table 4
2001:db8:babe:1000::/64 dev LAN
2001:db8:bad:babe::/64 dev WAN1
# ip -6 route show table 5
2001:db8:face:b000::/64 dev LAN
2001:db8:dead:beef::/64 dev WAN2
# ip -6 route
default via fe80::a25c:ccff:ff08:ff24 dev WAN1
# grep script /etc/dibbler/dibbler.conf
script "/etc/dibbler/client-notify.sh"
```

Aby vám routování správně fungovalo, je nejprve potřeba nakonfigurovat source routing. Protože se adresy přidělené pomocí prefix delegation mohou měnit, doporučuji přidat si DHCPv6 PD hook a při delegaci provést rekonfiguraci ip rule a ip route pro konkrétní routovací tabulku. Pokud pro prefix delegation používáte dibbler-client, použijte pro konfiguraci hookovacího skriptu direktivu script v dibbler.conf. Pokud chcete routeru explicitně říct, které rozhraní si pro svá připojení má vybírat, nastavte i default gateway v hlavní routovací tabulce (na slajdu na konci).

Dvě přípojky – router advertisement

```
interface LAN
{
        AdvSendAdvert on;

        prefix %PREFIX%/64
        {
                # Při shutdownu informuj klienty, že prefix
                # nemají dále používat (AdvPreferredLifetime 0)
                DeprecatePrefix on;
                # Doba, po kterou se prefix používá (sekundy)
                AdvValidLifetime 120;
                # Doba, po jakou prefix zůstává preferovaný (sekundy)
                AdvPreferredLifetime 60;
        };
};
```

Oznamovaný prefix bude na klientech preferován po dobu 60 sekund a platný celkem 120 sekund (obojí od okamžiku přijetí informace od routeru). Při restartu radvd se pošle klientům oznámení, že prefix mají přestat používat (prefix už nebude preferován, ale bude stále platný po dobu dalších 120 sekund od okamžiku přijetí informace o ukončení používání).

RA pod pokličkou

Na LinuxDays jsem se spletl a uváděl, že RFC stanovuje klientům povinnost nastavit si při refreshi router advertisementu AdvPreferredLifetime (preference) na NOW()+2 hod. Ve skutečnosti se to vztahuje na AdvValidLifetime (platnost), nicméně klienti to stejně ignorují a poslušně nastavují i nižší hodnoty, např. NOW()+120 sekund.

1. **Start radvd**
2. Oznámení prefixu PRIMARY/64 do LAN
   - Klienti začnou okamžitě používat prefix PRIMARY/64
3. Radvd periodicky (v čase ~3/4 AdvPreferredLifetime) oznamuje znovu PRIMARY/64 do LAN
   - Klienti si po každém přijatém oznámení obnoví čas platnosti prefixu na NOW()+120s a preference na NOW()+60s
   - Obnova **platnosti** na NOW()+120s je v **rozporu** s RFC 4862, 5.5.3, e), ale nikomu to asi nevadí ☺
   - RFC vyžaduje, aby minimální platnost u neautentizovaných obnov byla 2 hodiny
4. **Vyměníme prefix (přepínáme poskytovatele) - restart radvd**
5. Oznámení prefixu PRIMARY/64 s AdvPreferredLifetime 0 v čase T
   1. Klienti označí starý prefix jako Deprecated (resetují countery na Preferred=0, Valid=120s) a neiniciují z něj připojení. Existující spojení pokračují.
6. Oznámení prefixu SECONDARY/64 do LAN
   1. Klienti začnou okamžitě používat prefix SECONDARY/64
7. Radvd periodicky (v čase ~3/4 AdvPreferredLifetime) oznamuje znovu SECONDARY/64 do LAN
8. Klienti po uplynutí T+120 sekund (AdvValidLifetime) smažou PRIMARY/64 adresy z rozhraní síťovky i routovací tabulky
9. **Vyměníme prefix (přepínáme poskytovatele) - restart radvd, a tak dále**

Kde všude jsi to testoval? Fakt to funguje?

Android 5.1

Linux 4.8, 4.10

Windows 7

Windows 10

iOS 10

iOS 11

Mac OS 10.12

Uvedený způsob přepínání providerů jsem úspěšně otestoval na uvedených operačních systémech. Všechny se chovají naprosto stejně – ihned po obdržení AdvPreferredLifetime „0“ přestanou z adres smazaného prefixu iniciovat odchozí spojení a po vypršení AdvValidLifetime (120 sekund) si ze svého rozhraní adresy patřící do starého prefixu smažou.

Linux

Běžný provoz

```
2: enp0s25: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 state UP qlen 1000
    inet6 2001:db8:babe:1000:7a2b:cbff:fe8d:751a/64 scope global dynamic
       valid_lft 114sec preferred_lft 54sec
```

Výměna providera

```
2: enp0s25: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 state UP qlen 1000
    inet6 2001:db8:face:b000:7a2b:cbff:fe8d:751a/64 scope global tentative dynamic
       valid_lft 120sec preferred_lft 60sec
    inet6 2001:db8:babe:1000:7a2b:cbff:fe8d:751a/64 scope global deprecated dynamic
       valid_lft 119sec preferred_lft 0sec
```

Běžný provoz po exspiraci staré adresy

```
2: enp0s25: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 state UP qlen 1000
    inet6 2001:db8:face:b000:7a2b:cbff:fe8d:751a/64 scope global dynamic
       valid_lft 107sec preferred_lft 47sec
```

Ukázka chování v Linuxu. Při běžném provozu se prodlužuje valid_lft a preferred_lft o 60/120 sekund. Po rekonfiguraci a restartu radvd se stará adresa stane depreferovanou, ale po dvě minuty je ještě validní (pro nová příchozí spojení a existující stará odchozí spojení). Po dvou minutách se ztratí.

Záludnosti DHCPv6 delegace

Restart nadřazeného routeru způsobí ztrátu DHCPv6 delegace (nejsou k vám směrované pakety). Musíte znovu provést DHCPv6 request. Proto je vhodné testovat konektivitu s využitím **delegovaných** adres, např.:

```
ping -I 2001:db8:face:b000::1 www.seznam.cz
```

Dotazy?

Děkuji za pozornost

@zajDee