



IPv6 v Sokolovské uhelné

Michal Žejdl, zejdl@suas.cz



rozsah sítě

- tisíce portů a na nich živých prvků
- stovky *přepínačů*, **stanic** a dalších specifických zařízení
- desítky **směrovačů**, **serverů**, *zákazníků*
- jeden důl, tři elektrárny

historie

- 2009 - 2011 - SixXS /48
 - AAAA z DNS zóny google.com (do 2012)
 - Maribor (Czech users can only use the Slovenian PoP)
- 2011 - 2012 - TO2, GTS /48 (negarantovaná služba, rizika)
 - první přečíslování /64 prefixů ze SixXS na TO2 přes RA
- 2012 - RIPE, LIR /32
 - nezávislost na ISP (failover, adresy)



adresace sítí

- prefixy /64 z jednoho rozsahu /48
 - stejná čísla jako v IPv4 (1:1)
- 2a00:e140:oblast:síť::/64
 - PD /56 v rámci každé oblasti

adresace zařízení

- zpočátku autokonfigurace (SLAAC)
 - MS Win privacy extensions 1/týden, 1/den
 - vypnutí někdy nevydrželo (reboot, update)
- potřeby
 - omezit adresy, registrovat jména (DDNS)
 - udržovat dlouhodobá spojení (PuTTY)
- řešení na většině sítí
 - RA managed flag, no autoconf prefix (pouze DHCP)

DHCP

- klienti s rychlým renew (až 200 pps)
 - tiskárny, UPSky, kamery, převodníky
 - stejný stack? firmware UPS to řeší
- DDNS
 - dvě zóny, DHCP server si nespojí IPv4 a IPv6 klienta
 - přejmenování MS Windows nepošle FQDN
 - nutný ipconfig /renew6 (reboot nestačí)
- pozor na jednoduché klonování, které nezmění DUID



open-source software

- Linux kernel
- quagga (BGP, OSPF, RA), bird (testujeme)
- ISC DHCP - NA i PD (s výhradami)
- nana - evidence adres, aktuální stav sítě
- Argus - toky (flow)
- squid (caching proxy)

Linux kernel

- offloading Intel ET (igb) + pcap > 1500 B
 - ethtool gro off
- net.ipv6.route.max_size < BGP (50.000)
- první scan - Neighbour table overflow
 - net.ipv6.neigh.default.gc_thresh[123]
- netlink-cmd error: Cannot allocate memory
 - přidávání IPv6 cest brzy po startu stroje, viděno na VM

směrování

- quagga
 - slouží dlouho - BGP, OSPF, RA
 - jednou za více měsíců pád ospf6d na směrovači vzdáleném změně (LS flood?)
 - neřeší netlink error (neúplná směrovací tabulka v jádře)
- bird
 - dokáže se vypořádat s netlink error - doplní cesty později
 - nelze nastavovat pomocí příkazů, ale pouze souborem



ISC DHCP

- problémy se souborem pronájmů (leases)
 - ve starších verzích pouze rostl (rychlý renew dělal GB)
 - po změně nastavení je ustřižen po prvním neznámém prefixu (řešíme skriptem)
- prefixy deleguje, ale adresu pro next-hop neodhalí
 - možno použít NA se stejným DUID jako PD



nana

- vlastní udělátko pro automatickou evidenci adres a SNMP zařízení
- usnadňuje každodenní práci nejen správcům sítě
- vstupem je pouze sonda (sw pcap), ta chytá ARP a ND
- výstupy přes web frontend - vyhledávání, grafy, L2 topologie, užitečné údaje z SNMP
- web: nanyinka.org



Argus

- záznamy toků v síti
- velké možnosti sbírání i zpracování
- web: qosient.com/argus

squid

- 70% provozu z internetu je po IPv6
- zatím nemá Happy Eyeballs
 - pokud má někdo AAAA a nejde mu web, naši uživatelé si toho všimnou
 - děláme výhybky, píšeme správcům
- 2014 - stackoverflow.com - Cloudflare
 - úplné ztráty nejen našich src `:[ef]xxx:xxxx` (EUI-64 + 1/8 náhodných) na většinu testovaných Cloudflare serverů

NTP (pool.ntp.org)

- server ntp.suas.cz - 400 pps, 1% IPv6
- monitor v LA hlásil nedostupnost po IPv6
 - už kdysi přes TO2; v roce 2017 přes TPS trvale
 - RIPE Atlas sondy nepomohly
 - hrátky s BGP ukázaly na přechod HE/DT
 - komunikace s DT&HE - jen HE/LINX/DT, IPv6 UDP dst 123
 - DT našel viníka (agregační L2 switch)

další problémy

- za 9 let desítky různých, většina vyřešena
- přepínače Westermo zahazovaly link-local adresy se zapnutým IPv4 IGMP snooping
- NIC Intel i217-LM v MS Windows - IPv6 multicast flood při uspání dvou NIC na jedné broadcast doméně
- telefony Grandstream zjišťují IPv4 adresu NTP serveru: Standard query A 2a00:e140::123
- Supermicro IPMI má IPv6, ale bez omezení



otázky

- děkuji za pozornost
- otázky?