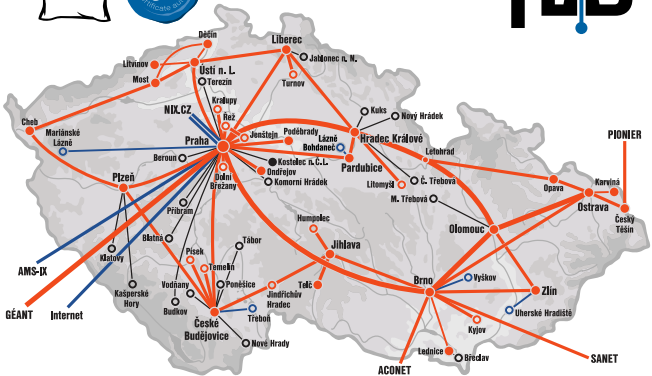
Digitální stopy v internetové komunikaci

Ondřej Caletka

15. ledna 2018

O sdružení CESNET



Jak vlastně funguje internet

Naivní představa

- noviny
- televize
- pošta

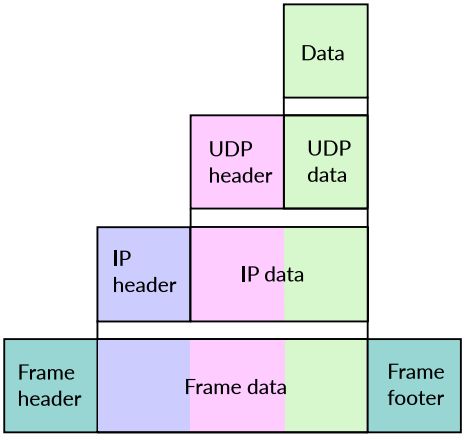
Realita

- telefon
- nelze *pasivně* konzumovat

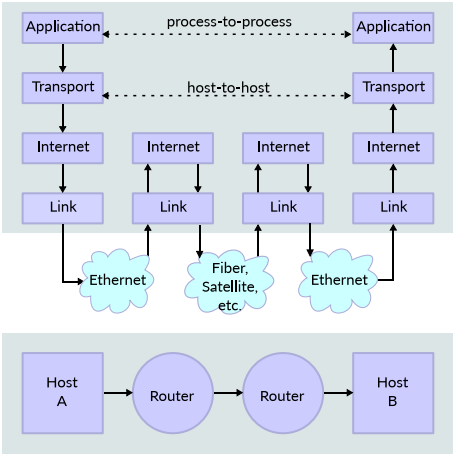
- datová komunikace jeden-na-jednoho
- každé použití zanechává stopy

cesnet

Vrstvový model



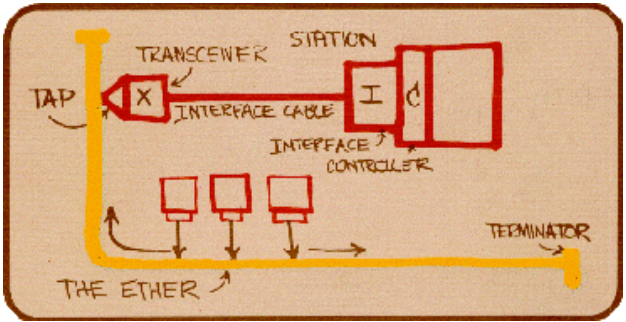
Autor: Cburnet, Kbrose, Wikimedia commons, GFDL



cesnet

Linková vrstva

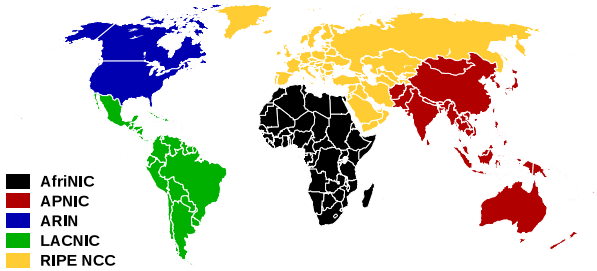
- pouze lokální komunikace
- nejčastěji Ethernet
- adresování MAC adresami
- 48bitové identifikátory
- **stabilní**



cesnet

Internetová vrstva

- s globálním dosahem
- hierarchické adresování
- identifikátor i lokátor
- 32bitové IPv4
- 128bitové IPv6
- bez pevného vztahu k hardware



cesnet

Transportní vrstva

- spojení z konce na konec
- přizpůsobuje přenosový kanál
- nejčastěji TCP a UDP
- adresace pomocí portů

Transport Layer Security

- mezivrtsva nad transportní
- šifruje data vyšších vrstev
- brání nižším vrstvám v sledování a **pozměňování** obsahu

cesnet

Aplikační vrstva

- mnoho protokolů, nejčastěji HTTP/HTTPS
- vlastní adresování
- často utajené nižším vrstvám pomocí šifrování TLS
- spousta **metadat**

cesnet

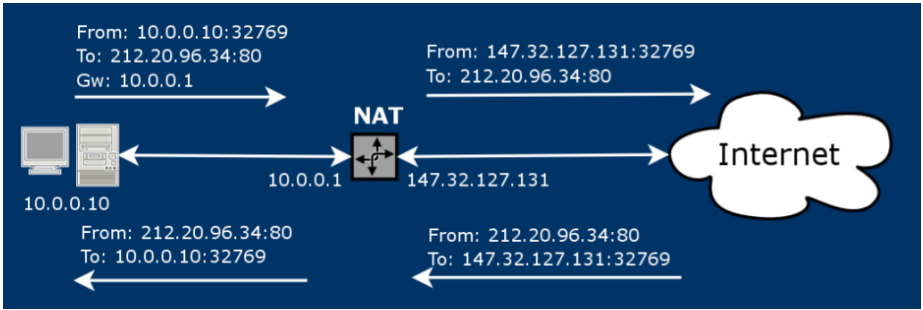
Vazba MAC a IP adres

- v reálném čase protokolem ARP/NDP
- u IPv4 často *semi*-permanentně pomocí DHCP serveru
- u IPv6 nejčastěji náhodná IP adresa proměnná v čase
- v malých sítích bez trvalých stop
- u větších sítí záznamenáváno a uchováváno

cesnet

Sdílení a překlad adres

- běžná praxe v IPv4
- zásah do komunikace porušující vrstvový model
- privátní a veřejné adresy, mapování čísel portů
- proměnlivé v čase
- zaznamenávání historie mapování jen u velkých sítí (Carrier Grade NAT)



cesnet

Data retention

- povinnost zaznamenávat provozní a lokalizační údaje
- záznam informací o datových tocích
- data internetové a transportní vrstvy
- *nevidí* za NAT koncového zákazníka
- možnost korelace záznamů z více zdrojů

cesnet

OS fingerprinting

- drobné odlišnosti v implementacích TCP/IP
- aktivní i pasivní

Příklad

operační systém	hodnota TTL	velikost TCP okna
Linux	64	5840
FreeBSD	64	65535
Windows XP	128	65535
Windows 7	128	8192

cesnet

Záznamy aplikační vrstvy

- jsou k dispozici pouze na koncových bodech
- obsáhlé HTTP hlavičky
- zejména User-Agent, Referrer, Cookies

▼ Request Headers

```
:authority: en.wikipedia.org
:method: GET
:path: /wiki/P3P
:scheme: https
accept: text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,*/*;q=0.8
accept-encoding: gzip, deflate, br
accept-language: cs,en-US;q=0.8,en;q=0.6,sk;q=0.4
referer: https://www.google.cz/
upgrade-insecure-requests: 1
user-agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/61.0.3163.100 Safari/537.36
```

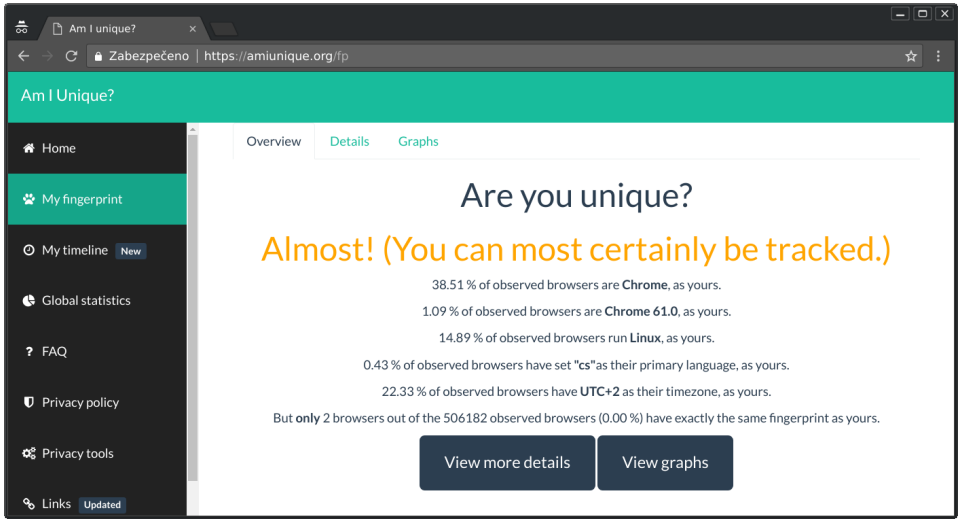
cesnet

Cookies

- informace, které může webová služba uložit a později přečíst
- nezbytná např. pro funkci přihlášení, stav nákupního košíku, apod.
- zneužívaná pro sledování chování uživatelů
- zajímavé zejména cookies významných třetích stran
  - Google Analytics
  - Facebook Like tlačítko
  - reklamní sítě

cesnet

Je prohlížeč unikátní?



cesnet

Shrnutí

- MAC jsou stabilní, ale nepřenáší se
- Vazbu IP adresy na MAC adresu dlouhodobě evidují jen velké sítě
- Překlady adres a náhodné IPv6 adresy dohledání majitelů komplikují
- Webový prohlížeč odesílá spoustu metadat, často unikátních
- Zajímavé informace mají mimo jiné provozovatelé reklamních systémů

cesnet

Děkuji za pozornost

Ondřej Caletka
Ondrej.Caletka@cesnet.cz
https://**Ondřej.Caletka.cz**

Prezentace je již nyní k dispozici ke stažení.