Jak funguje bezstavový NAT64 překladač

Vít Labuda • *vit (at) vitlabuda (dot) cz* • https://vitlabuda.cz/

Seminář IPv6: deset let poté • 6. června 2022

Tundra

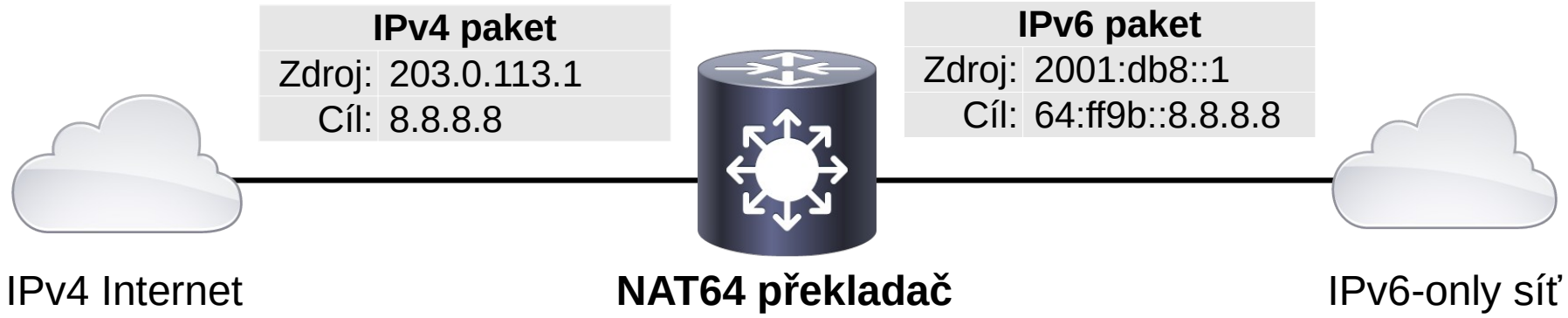
- **bezstavový NAT64 a CLAT překladač**
- **pro Linux**, běží zcela v uživatelském prostoru
- vícevláknový program
  - schopný využít moderní vícejádrové procesory
- naprogramovaný v jazyce C
- **open-source** (BSD licence)
- https://github.com/vitlabuda/tundra-nat64

Obsah přednášky

- NAT64 – úvod
- SIIT – *Stateless IP/ICMP Translation Algorithm*:
  - překlad IP hlaviček
  - překlad ICMP zpráv
- NAT64 – překlad adres
- Fragmentace a *Path MTU Discovery* (PMTUD)

NAT64 – úvod

- přechodový mechanismus umožňující zařízením v *IPv6-only* síti komunikovat s IPv4 Internetem
- dle pravidel algoritmu SIIT překládá pakety mezi IPv4 a IPv6 a mapuje IP adresy v nich obsažené



SIIT

- = *Stateless IP/ICMP Translation Algorithm*
- RFC 7915 (dříve RFC 6145 a 2765)
- definuje přesná pravidla pro:
  - překlad IP hlaviček
  - překlad ICMP zpráv
  - aktualizaci kontrolních součtů transportních protokolů
    - povinně u TCP a UDP, u jiných protokolů podpora volitelná
- nezabývá se překladem adres
  - ponechán na konkrétním přechodovém mechanismu
- každý paket je překládán individuálně

IPv6 hlavička (nejméně 40 bajtů)

Verze = 6	Třída provozu (TC)	Značka toku (20 bitů)	
Délka dat (zahrnuje rozšiřující hlavičky)		Další hlavička (NH)	Max. skoků (HL)
Zdrojová IPv6 adresa (128 bitů)			
Cílová IPv6 adresa (128 bitů)			
Rozšiřující hlavičky (libovolný počet, nemusí být žádná)			

IPv4 hlavička (20–60 bajtů)

Verze = 4	IHL	Typ služby (ToS)	Celková délka			
Identifikace (2 bajty)			R=0	DF	MF	Posun fragmentu (offset; 13 bitů)
Životnost (TTL)		Protokol	Kontrolní součet hlavičky			
Zdrojová IPv4 adresa (32 bitů)						
Cílová IPv4 adresa (32 bitů)						
Volby (pokud IHL > 5)						

Příznaky: R = rezervováno; DF = nefragmentovat (*Don’t Fragment*); MF = více fragmentů (*More Fragments*)

Rozšiřující IPv6 hlavičky procházené SIIT překladači

Volby pro všechny uzly (*Hop-by-hop options*; NH v předchozí hlavičce = **0**),
Volby pro cílový uzel (*Destination options*; NH v předchozí hlavičce = **60**)

Další hlavička (NH)	Délka dalších voleb	Volby
Volby		
Další volby (volitelné)		

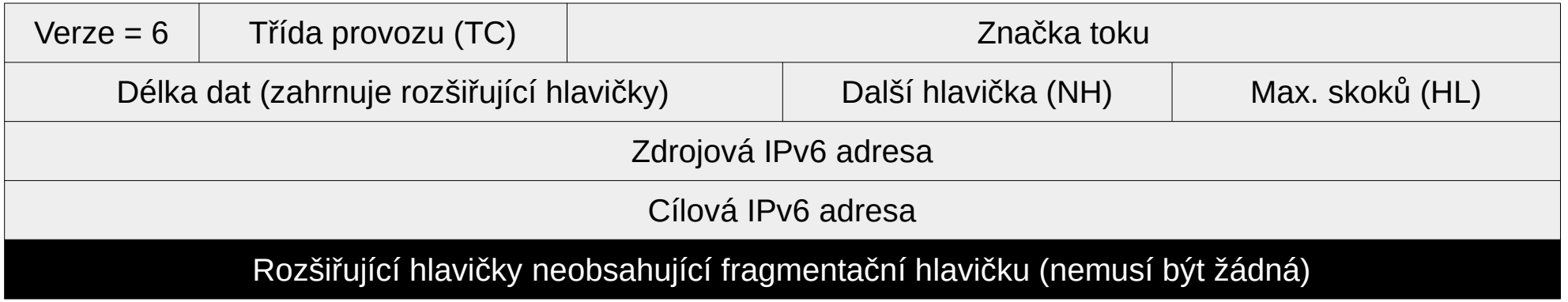
Směrování (*Routing*; NH v předchozí hlavičce = **43**)

Další hlavička (NH)	Délka dalších dat	Typ směrování	Zbývající segmenty
Data			
Další data (volitelná)			

Fragmentace (*Fragment*; NH v předchozí hlavičce = **44**)

Další hlavička (NH)	Rezervováno = 0	Posun fragmentu (offset; 13 bitů)	R=0	MF
Identifikace (4 bajty)				

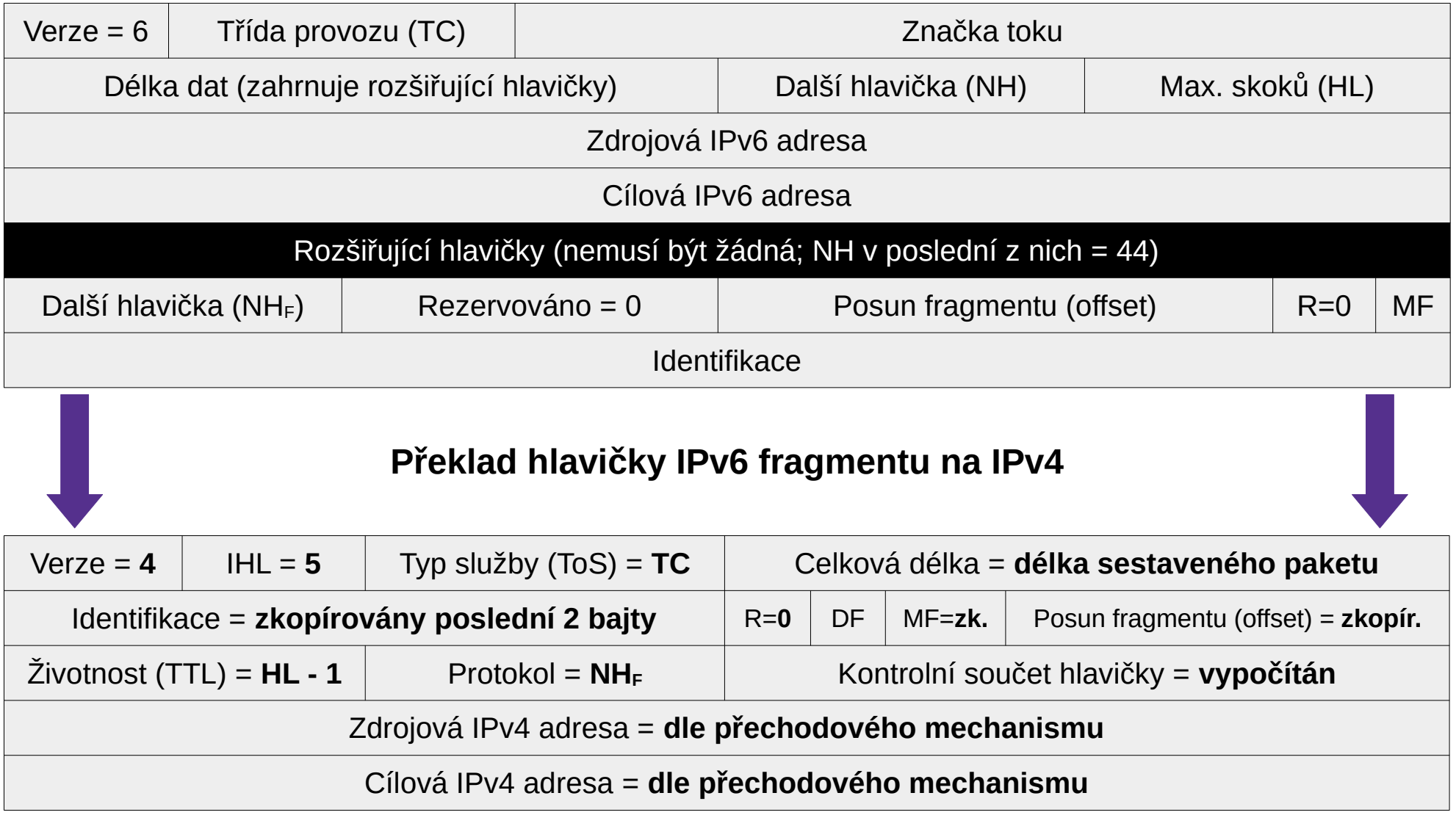
Příznaky: R = rezervováno (2 bity); MF = více fragmentů (*More Fragments*)

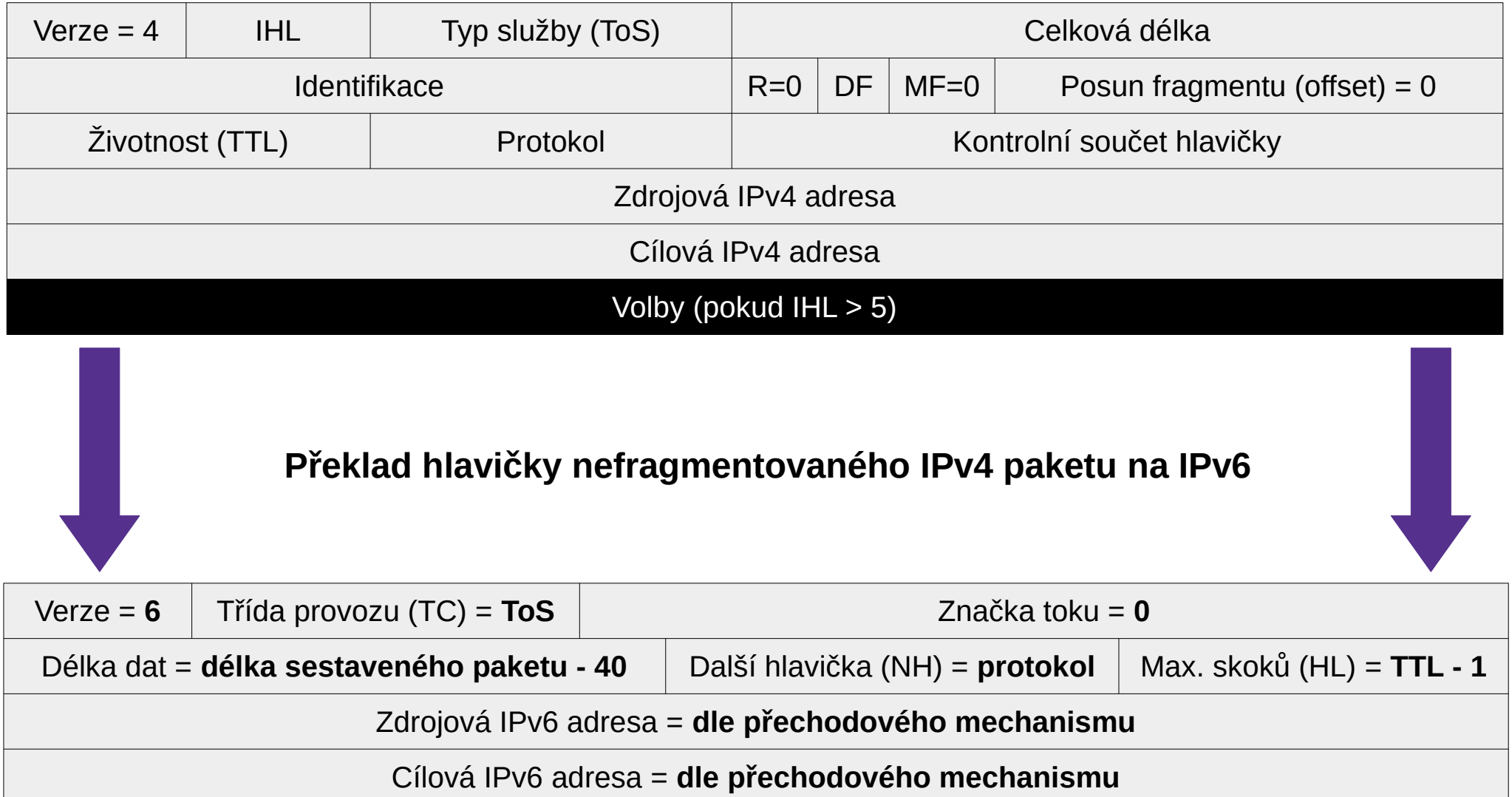


Překlad hlavičky nefragmentovaného IPv6 paketu na IPv4

Verze = **4**	IHL = **5**	Typ služby (ToS) = **TC**	Celková délka = **délka sestaveného paketu**			
Identifikace = **vygenerována**			R=**0**	DF	MF=**0**	Posun fragmentu (offset) = **0**
Životnost (TTL) = **HL - 1**		Protokol = **NH***	Kontrolní součet hlavičky = **vypočítán**			
Zdrojová IPv4 adresa = **dle přechodového mechanismu**						
Cílová IPv4 adresa = **dle přechodového mechanismu**						

* pokud IPv6 paket obsahuje rozšiřující hlavičky, číslo neseného protokolu je v té poslední


Verze = 6
Třída provozu (TC)
Značka toku
Délka dat (zahrnuje rozšiřující hlavičky)
Další hlavička (NH)
Max. skoků (HL)
Zdrojová IPv6 adresa
Cílová IPv6 adresa
Rozšiřující hlavičky (nemusí být žádná; NH v poslední z nich = 44)
Další hlavička (NHF)
Rezervováno = 0
Posun fragmentu (offset)
R=0
MF
Identifikace
Překlad hlavičky IPv6 fragmentu na IPv4
Verze = 4
IHL = 5
Typ služby (ToS) = TC
Celková délka = délka sestaveného paketu
Identifikace = zkopírovány poslední 2 bajty
R=0
DF
MF=zk.
Posun fragmentu (offset) = zkopír.
Životnost (TTL) = HL - 1
Protokol = NHF
Kontrolní součet hlavičky = vypočítán
Zdrojová IPv4 adresa = dle přechodového mechanismu
Cílová IPv4 adresa = dle přechodového mechanismu


Verze = 4
IHL
Typ služby (ToS)
Celková délka
Identifikace
R=0
DF
MF=0
Posun fragmentu (offset) = 0
Životnost (TTL)
Protokol
Kontrolní součet hlavičky
Zdrojová IPv4 adresa
Cílová IPv4 adresa
Volby (pokud IHL > 5)
Překlad hlavičky nefragmentovaného IPv4 paketu na IPv6
Verze = 6
Třída provozu (TC) = ToS
Značka toku = 0
Délka dat = délka sestaveného paketu - 40
Další hlavička (NH) = protokol
Max. skoků (HL) = TTL - 1
Zdrojová IPv6 adresa = dle přechodového mechanismu
Cílová IPv6 adresa = dle přechodového mechanismu

<table>
<tr><td>Verze = 4</td><td>IHL</td><td>Typ služby (ToS)</td><td colspan="4">Celková délka</td></tr>
<tr><td colspan="3">Identifikace</td><td>R=0</td><td>DF</td><td>MF≠0*</td><td>Posun fragmentu (offset) ≠ 0*</td></tr>
<tr><td colspan="2">Životnost (TTL)</td><td>Protokol</td><td colspan="4">Kontrolní součet hlavičky</td></tr>
<tr><td colspan="7">Zdrojová IPv4 adresa</td></tr>
<tr><td colspan="7">Cílová IPv4 adresa</td></tr>
<tr><td colspan="7">Volby (pokud IHL > 5)</td></tr>
</table>

* IPv4 paket je za fragment považován tehdy, je-li příznak MF **nebo** posun fragmentu nenulový

Překlad hlavičky IPv4 fragmentu na IPv6

<table>
<tr><td>Verze = 6</td><td>Třída provozu (TC) = ToS</td><td colspan="5">Značka toku = 0</td></tr>
<tr><td colspan="2">Délka dat = délka sestaveného paketu - 40</td><td colspan="2">Další hlavička (NH) = 44</td><td colspan="3">Max. skoků (HL) = TTL - 1</td></tr>
<tr><td colspan="7">Zdrojová IPv6 adresa = dle přechodového mechanismu</td></tr>
<tr><td colspan="7">Cílová IPv6 adresa = dle přechodového mechanismu</td></tr>
<tr><td colspan="2">Další hlavička (NH) = proto.</td><td>Rezervováno = 0</td><td colspan="2">Posun fragmentu (offset) = zkopírován</td><td>R=0</td><td>MF=zk.</td></tr>
<tr><td colspan="7">Identifikace = první 2 bajty nulové, poslední 2 bajty zkopírované</td></tr>
</table>

Překlad ICMP zpráv

- struktura ICMPv4 a ICMPv6 hlavičky je stejný, liší se obsah
- informační zprávy:
  - překládají se pouze zprávy *Echo Request* a *Echo Reply* (ping)
- chybové zprávy:
  - obsahují část IPv4/v6 paketu, který chybu způsobil – nutno přeložit
- v IP hlavičce se mění číslo protokolu
  - 1 (ICMPv4) <—> 58 (ICMPv6)
- fragmentované ICMP zprávy se nepřekládají

Struktura ICMPv4 / ICMPv6 zprávy ***Echo Request*** / ***Echo Reply***

Typ*	Kód = 0	Kontrolní součet
Identifikátor		Pořadí
Data		

* *Echo Request* – **8** (ICMPv4) / **128** (ICMPv6)
Echo Reply – **0** (ICMPv4) / **129** (ICMPv6)

Obecná struktura **chybových** ICMPv4 / ICMPv6 zpráv

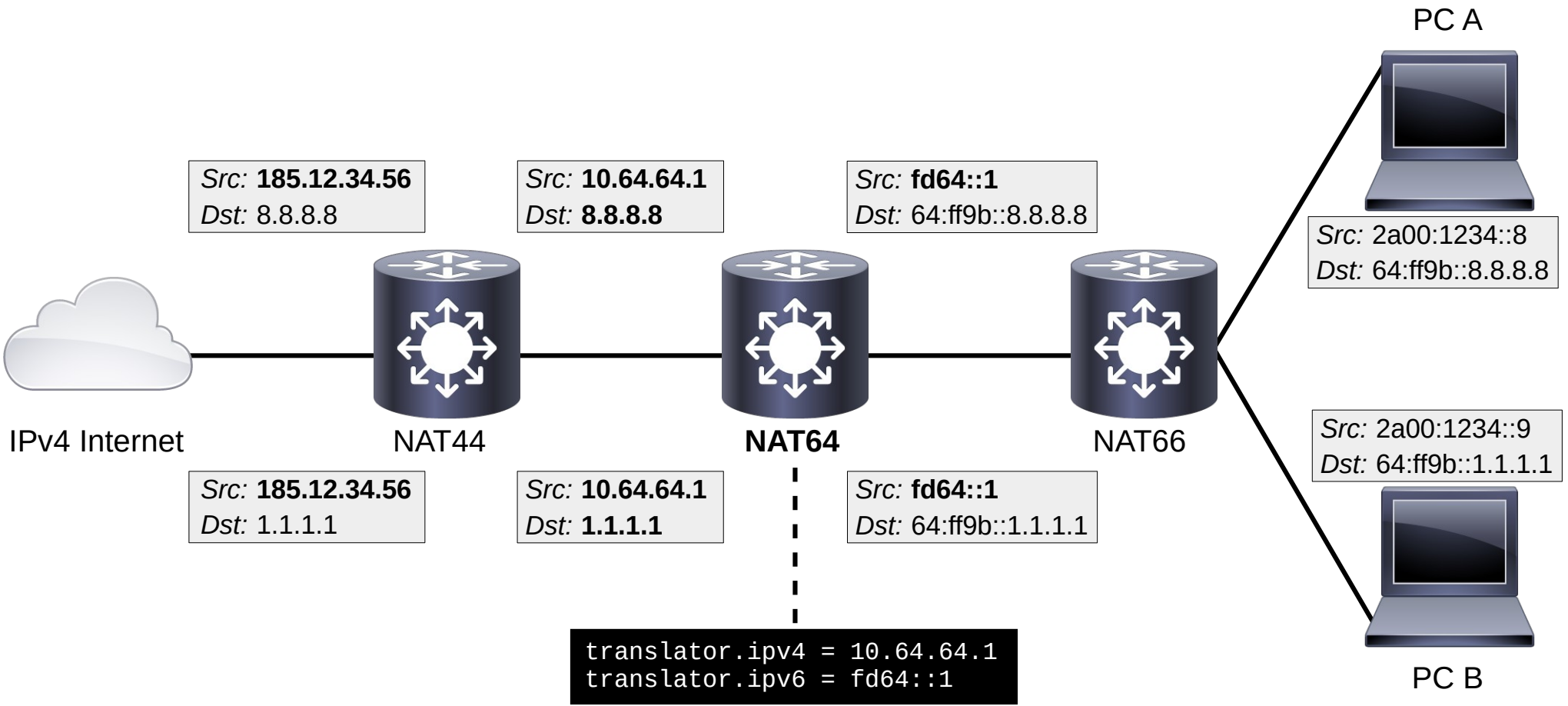
Typ	Kód	Kontrolní součet
Doplňující informace (dle typu a kódu zprávy)*		
Část IPv4 / IPv6 paketu, který způsobil chybu		

* např. v případě ICMPv4 „*Fragmentation needed and DF set*“ / ICMPv6 „*Packet too big*“ je zde **MTU**

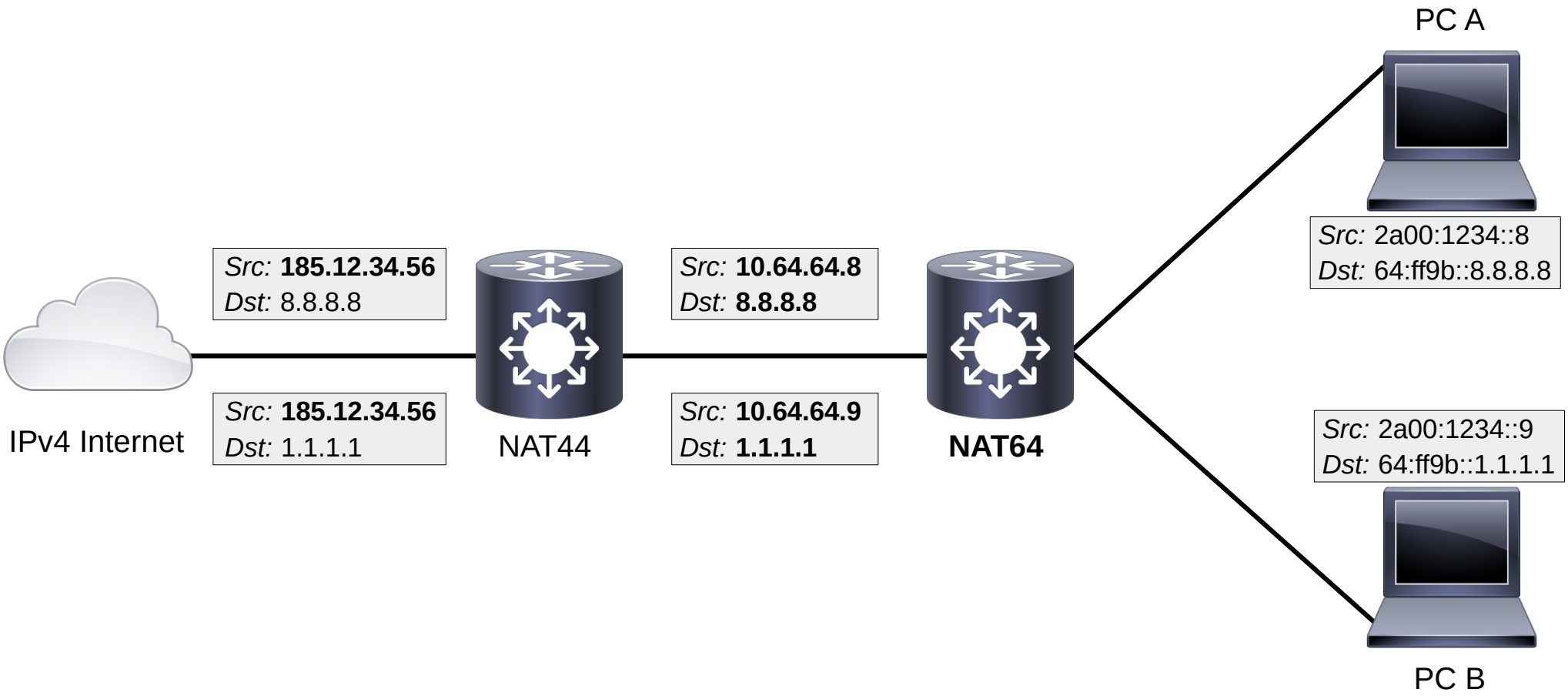
NAT64 – překlad adres

- **„Překlad pouze jedné adresy“:**
  - v konfiguraci překladače je specifikována jedna IPv4 a jedna IPv6 adresa
  - provoz z více IPv6 adres je možné překládat ve spolupráci s NAT66 překladačem
  - open-source implementace **Tundra**
- **„Překlad 1:1 s mapovací tabulkou“:**
  - každá jedna IPv6 adresa je mapována na jednu IPv4 adresu
  - překladač sleduje procházející IP adresy, ale nesleduje stav probíhajících spojení
  - open-source implementace **Tayga**
- **Stavový NAT64:**
  - standardizován v RFC 6146 (*Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers*)
  - open-source implemenace **Jool**

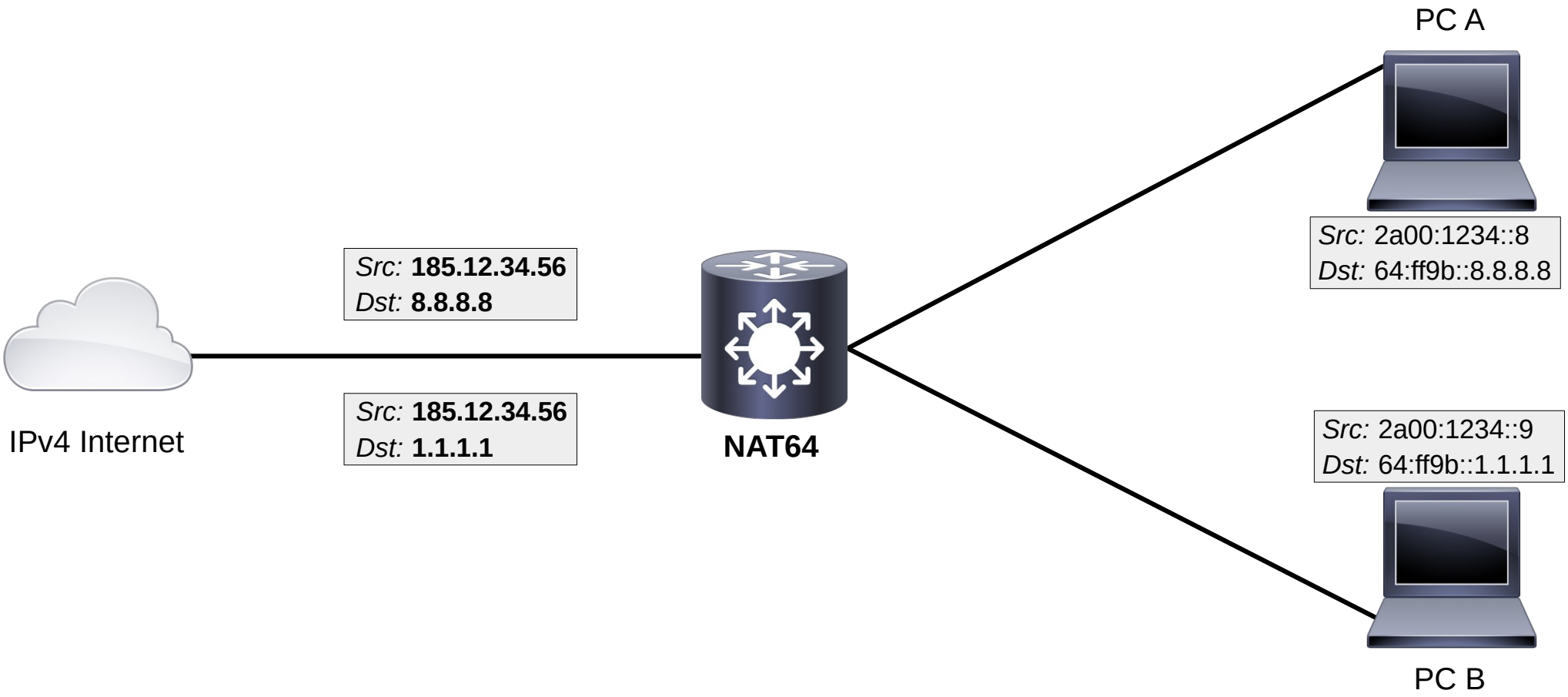
„Překlad pouze jedné adresy“



„Překlad 1:1 s mapovací tabulkou“



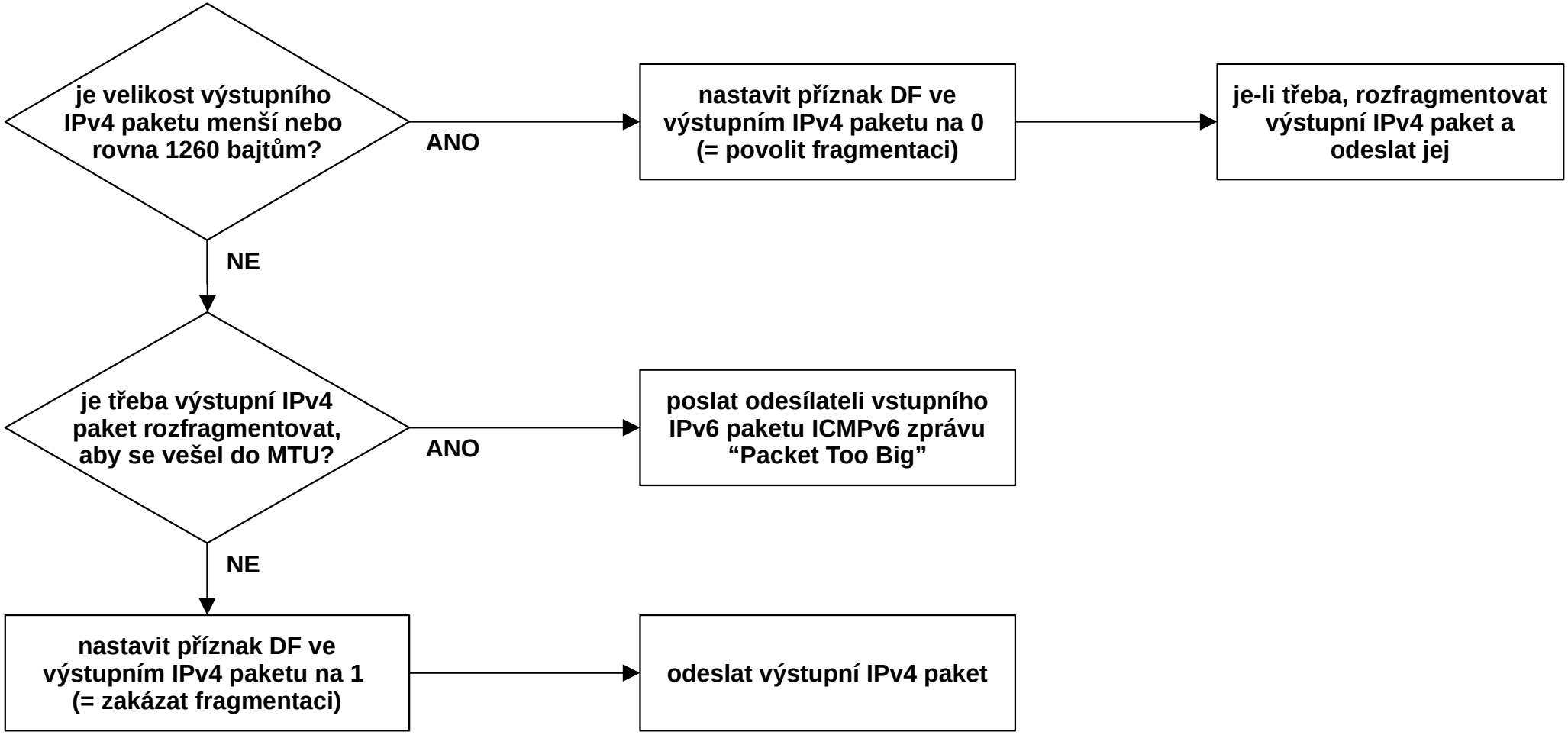
Stavový NAT64



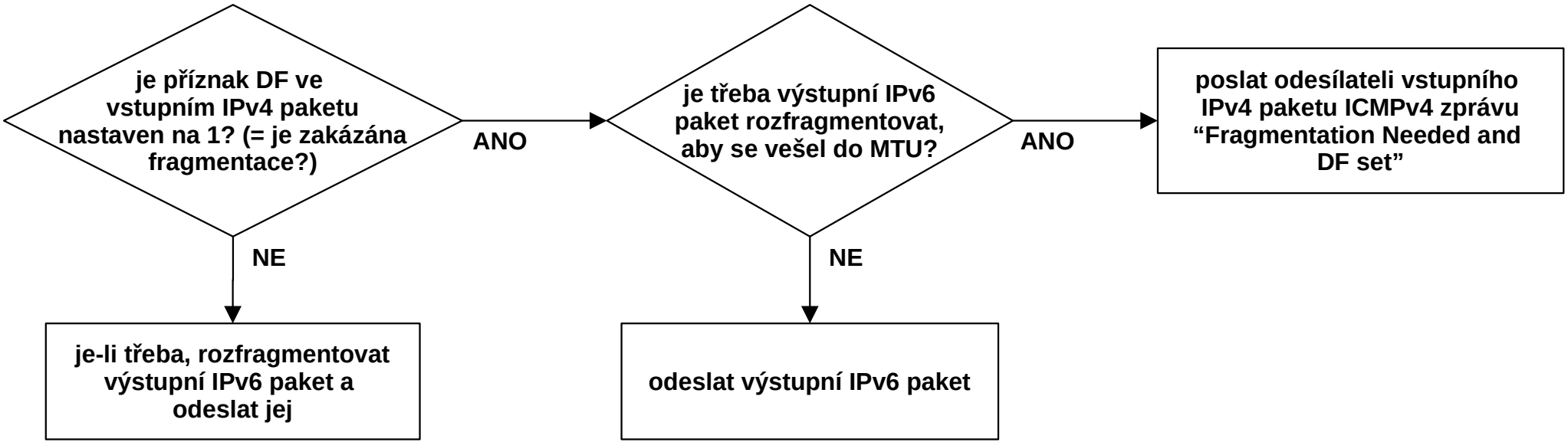
Fragmentace a *Path MTU Discovery*

- *nejsložitější* část SIIT překladu
- IPv4 požaduje MTU nejméně **68 bajtů**, zatímco IPv6 **1280 bajtů**
- v IPv4 je PMTUD **volitelné**, zatímco v IPv6 je **povinné**
  - v IPv4 ovládané příznakem DF (= nefragmentovat) v hlavičce
  - IPv6 pakety smí fragmentovat pouze odesílatel (routery po cestě nikoliv)
- chybové zprávy ICMPv4 „*Fragmentation needed and DF set*“ / ICMPv6 „*Packet too big*“ obsahují **MTU**, které bylo překročeno
  - při překladu je přizpůsobeno rozdílným velikostem IP hlaviček a vlastnostem obou IP protokolů

Fragmentace – IPv6 –> IPv4



Fragmentace – IPv4 –> IPv6



Děkuji za pozornost

Vít Labuda • *vit (at) vitlabuda (dot) cz* • https://vitlabuda.cz/