

Protokoly TCP/IP

Petr Grygárek

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

1

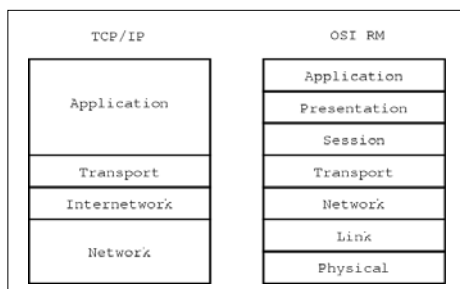
TCP/IP

- standard pro komunikaci v Internetu
 - a stále více i v intranetech
- TCP – protokol 4. vrstvy (spolu s UDP)
- IP – protokol 3. vrstvy

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

2

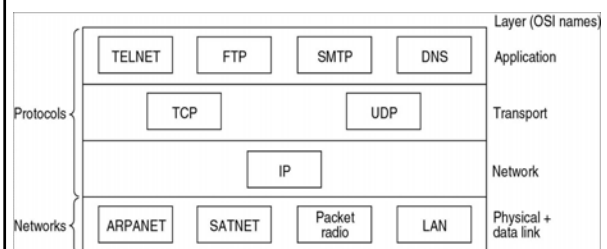
Vrstvený model a srovnání s OSI-RM



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

3

Vrstvený model TCP/IP



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

4

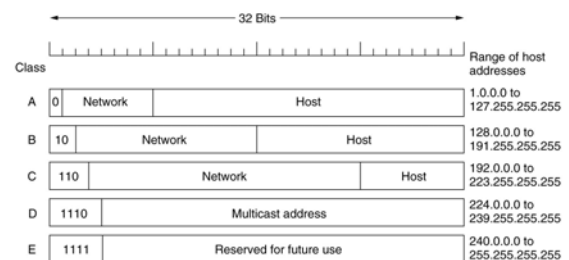
Adresace v protokolu IP

- adresy 32b (X.X.X.X)
 - každé rozhraní prvku rozumějícího 3. vrstvě OSI RM připojené do sítě musí mít jednoznačnou IP adresu
 - (stanice a rozhraní směrovačů)
- dělení na adresu sítě + adresu uzlu v rámci sítě
 - adresy všech stanic na segmentu LAN (broadcast doména 2. vrstvy) mají společnou část IP adresy (adresu sítě, prefix)
 - směrovače nemusí ukládat adresy všech stanic v síti, pouze adresy jednotlivých sítí
 - = > omezení rozsahu směrovacích tabulek

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

5

Třídy IP adres (historie)



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

6

Beztrždní (classless) adresy

- délka prefixu sítě přidělována podle potřeby
- k beztrždní adrese musí být specifikována maska podsítě (subnet mask) určující délku prefixu
- v poslední době se třídy adres prakticky přestaly používat
 - přechod na CIDR-Classless Inter-Domain Routing)
 - možnost agregování záznamů ve směrovací tabulce na základě společného prefixu (bez ohledu na třídy) - supernetting

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

7

Přidělování IP adres

- adresy sítí přiděluje oblastní správce (pro Evropu RIPE)
 - vyřizování elektronickou cestou (zprostředkovává poskytovatel)
- adresy původně přidělovány bez ohledu na topologii a geografickou polohu
- v posledních letech snaha o hierarchickou adresaci (přidělování prefixu sítě s délkou podle potřeby)
 - případné další podsít'ování (subnetting)
- soukromé izolované sítě mají vyhrazené rozsahy adres použitelné opakovaně, nesmí být přímo připojeny k Internetu
 - pokud jsou připojeny, tak přes proxy s překladem adres-NAT
 - 10.0.0.0, 172.16.0.0-172.31.0.0, 192.168.0.0-192.168.255.0

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

8

Speciální IP adresy

0 0	This host
0 0 ... 0 0	Host
1 1	Broadcast on the local network
Network 1 1 1 1 ... 1 1 1 1	Broadcast on a distant network
127 (Anything)	Loopback

- Univerzální broadcast: 255.255.255.255
- Multicast: 224.x.x.x - 239.x.x.x

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

9

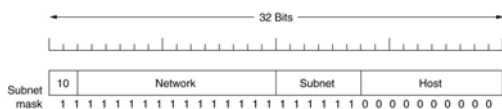
Podsít'ování (subnetting)

- možnost rozdělení přiděleného adresního rozsahu mezi více segmentů
 - každý segment musí mít svou vlastní adresu podsítě
- Umožňuje efektivnější rozdělení adres vzhledem k reálným počtům stanic na segmentech
 - nejmarkantnější u třídních adres
- část adresy původně určené pro identifikaci uzlu sítě se rozdělí na adresu „podsítě“ a na adresu uzlu v této podsíti
- dělit možno po bitech s ohledem na skutečné počty uzlů v jednotlivých segmentech a počet segmentů

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

10

Maska podsítě (subnet mask)



- pro každou (podsít'ovanou) adresu nutno udat, kolik bitů zleva představuje síť+subsít' a kolik uzlů.
- jednička na příslušném bitovém místě masky podsítě znamená, že odpovídající bit adresy patří do adresy sítě resp. podsítě, nula zařazuje bit do adresy uzlu

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

11

Podmínky pro podsít'ování

- Minimální počet bitů pro adresu uzlu v podsíti je 2
 - Musíme umět zaadresovat podsít' jako takovou (adresa uzlu nuly) a všechny stanice na podsíti (adresa uzlu jedničky), takže maximální počet uzlů v podsíti je vždy o 2 menší, než odpovídá počtu bitů ponechaných pro adresu uzlu.
- Podsít' určená bitovou kombinací samých nul ("subnet zero") se z historických (formálních) důvodů dříve nepoužívala, dnes se používá běžně.
 - Na některých směrovačích je nutné použít subnet zero explicitně povolit.

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

12

Příklady podsít'ovaných adres

- 151.88.19.103/255.255.255.0:
třída B, podsít' 151.88.19 sítě 151.88.0.0, broadcast pro podsít' 151.88.19.255
- 151.88.19.103/255.255.255.224:
třída B, 8 bitů předposledního a 3 bity posledního oktetu použito pro podsít',
podsít' 151.88.19.96 sítě 151.88.0.0, broadcast pro podsít' 151.88.19.127
- 10.0.0.239/255.255.255.240:
broadcast adresa na síti 10.0.0.224 (!)

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

13

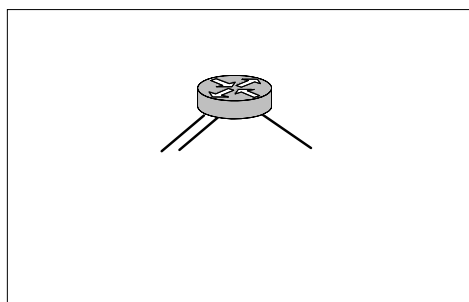
Praktické použití podsít'ování

- Rozdělení prefixu přidělené délky na daný počet podsítí (zadány maximální počty stanic na segmentech)
 - pozor na nepoužitelné adresy a adresu rozhraní routeru
- Stanovení maximální délky pevně přiděleného prefixu (požadovaného od ISP) pro požadovaný počet podsítí a požadované počty stanic na jednotlivých segmentech
- Vytvoření adresního plánu sítě WAN
 - zadaná topologie dvoubodových spojů, u jednotlivých směrovačů připojeny LAN, zadány požadované počty stanic na segmentech LAN

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

14

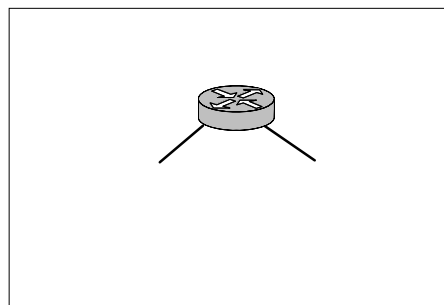
Rozdělení rozsahu (1)



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

15

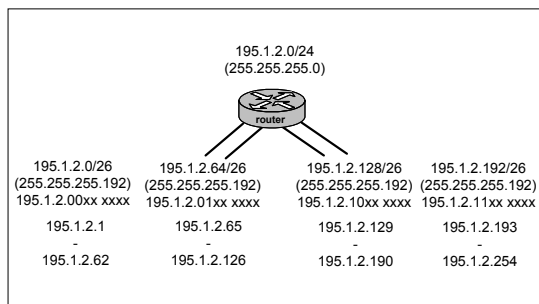
Rozdělení rozsahu (2)



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

16

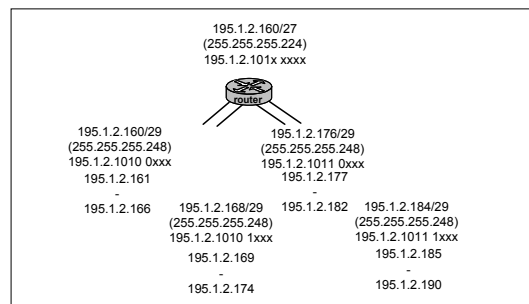
Rozdělení rozsahu (3)



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

17

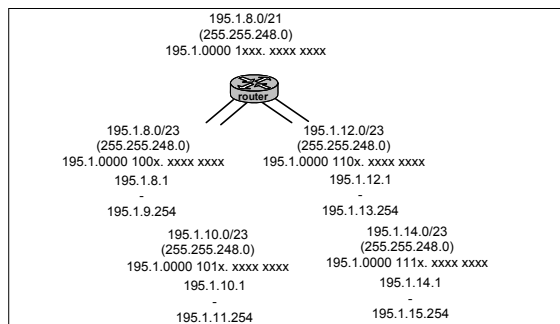
Rozdělení rozsahu (4)



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

18

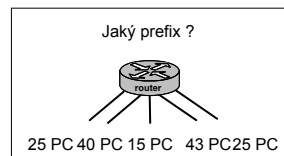
Rozdělení rozsahu (5)



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

19

Jak dlouhý prefix vyžádat od ISP ?

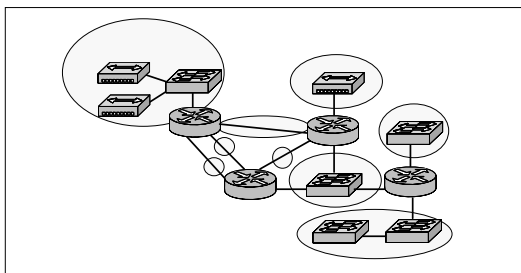


- Maximální počet stanic 43
 - +1 adresa na rozhraní směrovače = 44
- K adresování 44 kombinací nutných 6b (64)
- 5 podsítí – k jejich adresování nutné 3b (8)
- Potřebujeme 6+3=9b, vyžádáme prefix 32-9-23b (/23)
 - Použijeme masku podsítě /26

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

20

Adresní plán WAN



- Podsítě omezeny zařízeními pracujícími na 3. vrstvě OSI RM
 - směrovače, stanice – ne přepínače a rozbočovače
- Adresní prostor pro jednotlivé podsítě rozdělíme stejně jako v předchozím případě

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

21

Překlad adres (NAT)

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

22

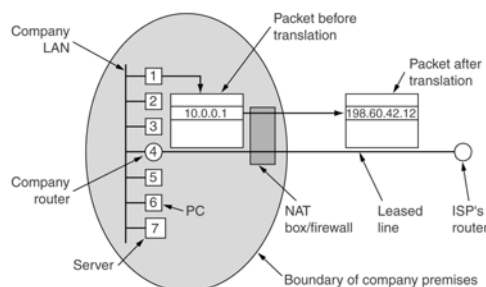
Network Address Translation (NAT)

- překlad zdrojové nebo cílové IP adresy
 - probíhá obvykle na směrovačích (L3 prvcích)
- použití překladové tabulky
 - záznamy buďto konfigurovány staticky nebo se vytvářejí dynamicky automaticky
- typicky se překládá mezi "vnitřní" sítí s privátními adresami a "vnější" sítí s veřejnými (globálně jednoznačnými) adresami

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

23

Scénář použití NAT



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

24

Statický a dynamický NAT

- Statický NAT
 - překladová tabulka konfigurována staticky
- Dynamický NAT
 - překladová tabulka vzniká za provozu dynamicky
 - adresy se propůjčují z rezervuáru adres (pool)

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

25

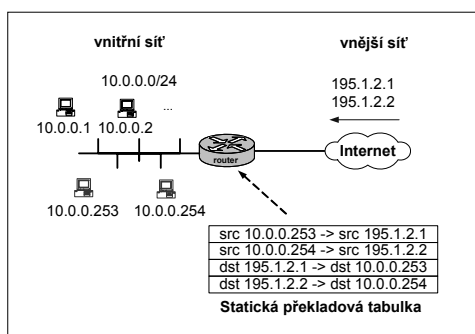
Způsob použití statického NAT

- statický překlad konkrétní zdrojové adresy vnitřní sítě na konkrétní adresu směrovatelnou ve vnější síti
- statický překlad konkrétní cílové adresy (směrovatelné ve vnější síti) na konkrétní adresu vnitřní sítě

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

26

Statický NAT



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

27

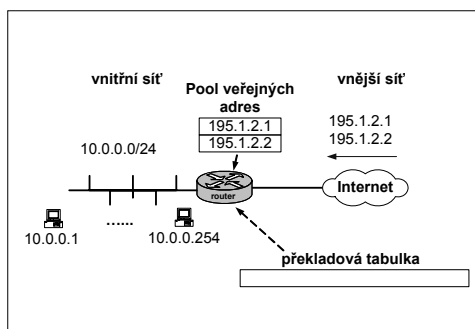
Způsob použití dynamického NAT

- uživatel je přiděleno M veřejných adres
- uživatel chce ve vnitřní síti provozovat $N > M$ strojů a umožnit jim přístup do vnější sítě (vždy nejvýše M strojům současně)
- dosud nevyužité veřejné adresy směrovač udržuje v poolu
- jestliže stanice S z vnitřní sítě pošle paket do vnější sítě, je jí dočasně přidělena některá adresa V z poolu veřejných adres (pokud v něm nějaká zbývá)
 - v překladové tabulce se vytvoří záznam mapující IP adresu stanice S na adresu V
 - v odchozím paketu se přepíše (zdrojová) adresa stanice S na adresu V (ta je ve vnější síti jednoznačná a směrovatelná)
 - při příchodu odpovědi na adresu V se v překladové tabulce najde, že se cílová adresa V má přeložit na adresu S, což se provede a paket se odešle do vnitřní sítě

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

28

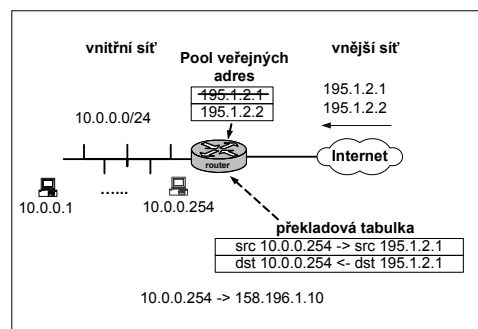
Dynamický NAT (1)



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

29

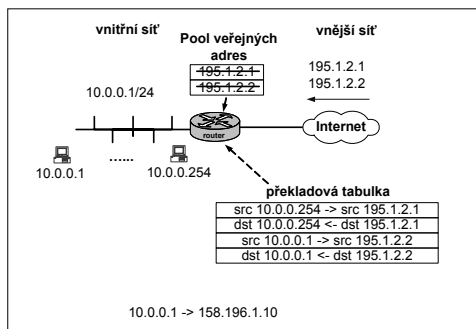
Dynamický NAT (2)



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

30

Dynamický NAT (3)



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

31

Časové omezení dynamického NAT

- aby mohlo N strojů sdílet M adres, mají dynamicky vytvořené záznamy překladačové tabulky časově omezenou platnost (timeout od posledního použití)
- při odstranění expirované položky se veřejná adresa vrátí zpět do poolu

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

32

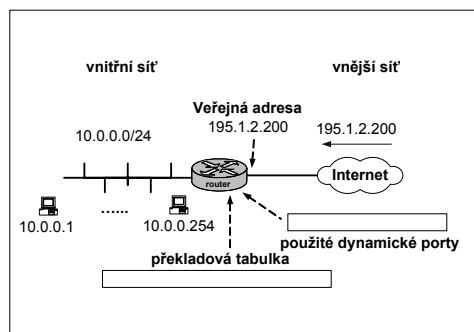
Port Address Translation

- V terminologii Linuxu „Masquarading“
- Ukrytí více stanic za jedinou IP adresu, rozlišení pomocí různých zdrojových portů
 - Zdrojové porty přidělovány dynamicky, přičemž vzniká tabulka mapující jednotlivé porty na vnitřní IP adresy

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

33

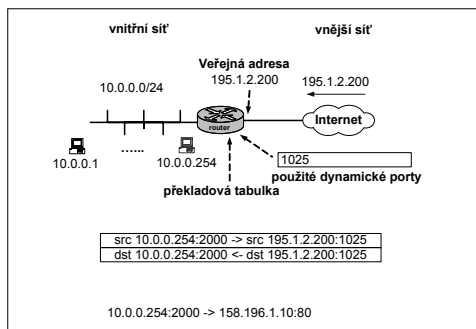
PAT (1)



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

34

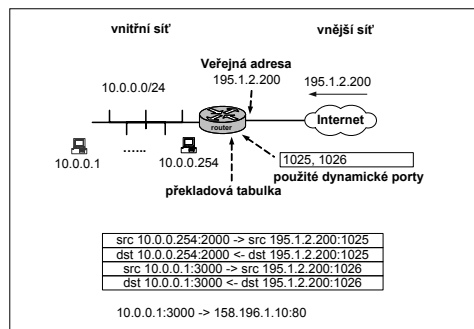
PAT (2)



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

35

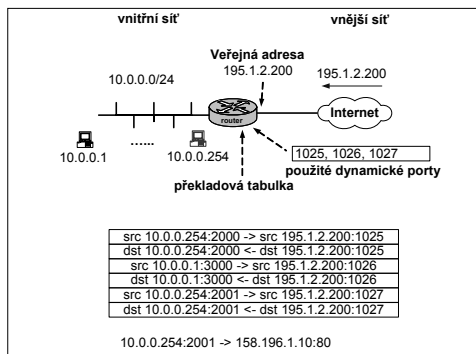
PAT (3)



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

36

PAT (4)



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

37

Protokol IP

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

38

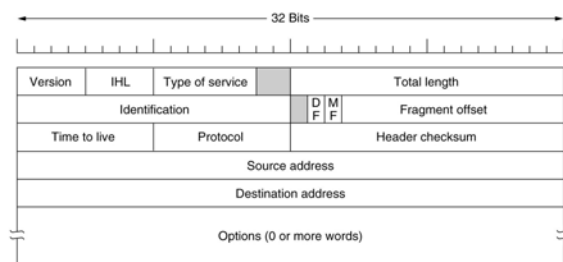
IP - Internet Protocol

- 3. vrstva, síťová služba posílání nezávisle směrovaných paketů bez spojení
- RFC 791, 1042, 894, v současné době verze 4, chystá se verze 6

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

39

Hlavička IP



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

40

Fragmentace paketů

- Rozdělení paketu při průchodu linkami s nedostatečným MTU (Maximum Transfer Unit = max. délka datové části rámce)
 - fragmentace ve směrovačích nebo na zdroji
 - skládání až v cílovém uzlu
 - fragmenty mohou jít různými cestami
 - Skládání podle Identification, pořadí dle Fragment Offset, poslední fragment nemá nastaven More Fragments flag
- Podle konvence musí každý segment Internetu být schopen přenést paket o délce 576 B

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

41

Podpůrné protokoly IP

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

42

ARP - Address Resolution Protocol

- RFC 826, 1027
- mapování IP adres na MAC adresy
- Při potřebě zjistit MAC adresu k IP adrese se generuje ARP request (broadcast), ten obsahuje požadovanou IP adresu. Stanice s touto adresou odpoví svou MAC adresou (ARP reply).
- Zdroj ARP dotazu si výsledek uloží do ARP cache
 - (lokální cache jednotlivých stanic udržující známá mapování IP-MAC adres)
- Navíc se do requestu vkládá dvojice < zdrojová IP, zdrojová MAC >, každý počítač sleduje všechny ARP broadcasty a doplňuje informace ve své ARP cache..

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

43

ICMP - Internet Control Message Protocol

- RFC 792
- protokol služebních řídicích zpráv
- ohlašování chyb a zvláštních stavů při přenosu paketů
- šíří se v datové části IP paketů

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

44

Zprávy ICMP

- Echo request, echo reply
- Destination unreachable (network, host, port, protocol unreachable, zakázána, ale nutná fragmentace)
 - + administratively prohibited
- Time exceeded (TTL=0 nebo vypršel čas pro refragmentaci)
- Redirect
- Parameter problem

Novější (a ne vždy podporované) zprávy

- Source quench - žádost cílové stanice o snížení rychlosti generování zpráv zdrojem (přepřlují se buffery)
- Address mask request, Address mask reply - zjištění síťové masky rozhraní
- Router solicitation, Router advertisement

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

45

Zjišťování cesty sítí - traceroute

- většina OS
- zjištění všech směrovačů na cestě k cílové stanici
- využívá pole TTL, začíná se od 1, stále se zvyšuje, sledují se IP adresy, ze kterých přijde ICMP Time Exceeded
- testovací paket buďto ICMP (Microsoft) nebo UDP na neexistující port (Unix)

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

46

Transportní vrstva TCP/IP: UDP a TCP

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

47

Porty

- Spolu s IP adresou identifikují konkrétní proces (službu) na konkrétním zařízení v Internetu
 - (transportní entitu)
- 16bit (0-65535), zvlášť pro TCP a UDP
 - 0-1023: Veřejně definované služby (well-known)
 - >1024 (4096) – klientské porty, obvykle přidělování volných portů operačním systémem
- Vždy cílový i zdrojový port

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

48

UDP - User Datagram Protocol

- nepotvrzovaná datagramová služba
- podpora všesměrového a skupinového vysílání (na daném portu)
- porty identifikují proces odesílatele, resp. příjemce na vysílající, resp. přijímající stanici
- kontrolní součet zahrnuje datovou část (na rozdíl od IP, tam jen hlavičku)
 - není však povinný

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

49

(Pseudo)hlavička UDP



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

50

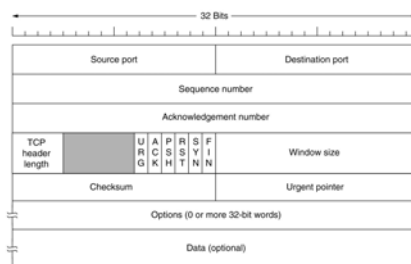
TCP: Transmission Control Protocol

- duplexní spolehlivý logický kanál
 - v prostředí se ztrácením, duplikací, a přehazováním pořadí
- segmentování dat (rozdělení proudu dat do částí vhodných pro přenos v paketech), číslování oktetů proudu dat
- algoritmus sliding window (go-back-N), pozitivní (inkluzivní) potvrzování, piggybacking, adaptivní změna časového limitu pro retransmisi
- řízení toku dat inzerováním aktuální kapacity přijímacích bufferů, vysílací okno se dynamicky přizpůsobuje přijímacímu
- robustní protokol navazování spojení a ukončování spojení

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

51

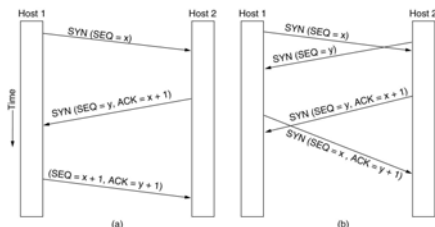
(Pseudo)hlavička TCP



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

52

Navazování TCP spojení



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

53

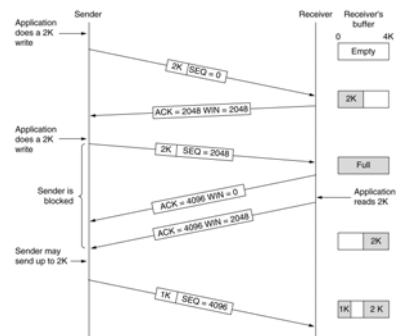
Navazování TCP spojení

- three way handshake: SYN, SYN+ACK, ACK
 - dohoda o startovacím sekvenčním čísle (zvláště pro oba směry)
 - počáteční sekvenční čísla náhodná, aby se zabránilo případnému ovlivnění zbloudilými pakety ze zavřeného a poté brzy opět znovu otevřeného spojení mezi týmiž entitami
- řeší i problémy pokusu o aktivní navázání spojení oběma stranami současně

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

54

Průběh TCP spojení - řízení toku dat



© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

55

Uzavření spojení

- Uzavírá se zvlášť z obou stran
 - Možnost "polovičního" uzavření spojení (half-close)
 - FIN+ACK z obou stran
- První může uzavřít kterákoli strana

© 2005 Petr Grygárek, FEI VŠB-TU Ostrava, Počítačové sítě (Bc.)

56