

Y36PSI

Protokolová rodina TCP/IP



Program

- protokol síťové vrstvy
 - IP
- podpůrné protokoly
 - ICMP
 - RARP, BOOTP, DHCP
- protokoly transportní vrstvy
 - UDP
 - TCP

Protokol IP

- InterNet Protokol, rfc791
- podpora fragmentace
- podpora komunikace přes směrovače
- vytváření IP paketů z paketů vyšší vrstvy
- IP-adresa 4B
- koncové uzly, směrovače
- základní protokol rodiny TCP/IP

Fragmentace

- umožňuje vložení IP paketu do kratších rámců nižší vrstvy (MTU)
- fragmentaci provádí libovolný směrovač
- defragmentaci provádí koncový uzel
- možnost zakázání fragmentace
- podpora fragmentace
 - DF (Don't Fragment)
 - MF (More Fragments)
 - identifikace IP datagramu
 - fragment offset

Směrování

- přímé směrování – uzly ve stejné síti
- nepřímé směrování – uzly v různých sítích
- implicitní směrování – defaultní brána
- směrování podle tabulky
 - statické
 - dynamické
- služby směrovače
 - podpora předávání paketů (forwarding)
 - kontrola a snižování TTL
 - přepočítání kontrolního součtu
 - zohlednění ToS
 - prioritita (precedence) 3b
 - nízké zpoždění (delay) 1b
 - vysoká propustnost (throughput) 1b

IP hlavička

verze IP	délka záhlaví	typ služby	celková délka	
identifikace IP datagramu			příznaky	posunutí fragmentu
TTL	protokol vyšší vrstvy		kontrolní součet IP záhlaví	
IP adresa odesílatele				
IP adresa příjemce				
volitelné položky hlavičky				
data				

- verze IP – 4
- délka záhlaví – po 32b
- typ služby (ToS) – opomíjeno
 - priorita 3b
 - nízké zpoždění 1b
 - vysoká propustnost 1b
- celková délka – omezení na 64KB
- identifikace
- příznaky
 - 0
 - DF ... 1 nefragmentovat
 - MF ... 0 poslední fragment
- posunutí fragmentu 0 první
- TTL – 0 likvidace paketu
- protokol vyšší vrstvy
 - 1 ICMP, 2 IGMP, 6 TCP, 17 UDP
 - 4 IP over IP, 97 Eth within IP
- volitelné položky – zahazuje se

Volitelné položky

- max. 40B
- kód
 - kopírovat 1b – kopírování při fragmentaci
 - třída 2b – 0 data, řízení; 2 ladění, měření
 - číslo volby 5b
- zaznamenávej směrovače – max 9 odchozích adres
- zaznamenávej čas – max 4 adresy s časem
- explicitní směrování – přes které směrovače
- striktní explicitní směrování - všechny směrovače
- upozornění pro směrovač – informace pro mezilehlé směrovače

Protokol ICMP

- Internet Control Message Protocol, rfc 777
- přenos uvnitř IP datagramu
- přenos chybových i řídicích informací
 - testování dostupnosti (Echo Req/Rep)
 - řízení zahlcení a toku
 - změna směrovací tabulky
 - informace o masce
 - časová synchronizace
 - ...
- omezená implementace
- zahazování z bezpečnostních důvodů

Typy ICMP paketů

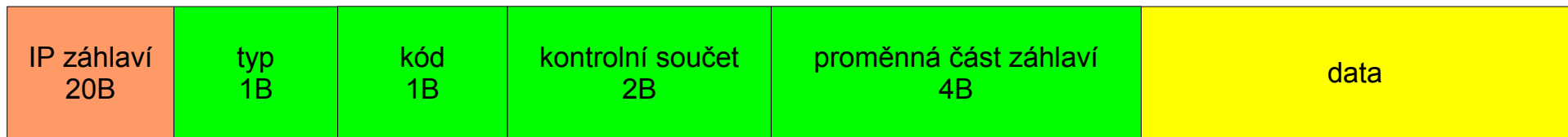
- 0 0 echo reply
- 3 Nedoručitelný IP datagram
 - 0 nedosažitelná síť
 - 1 nedosažitelný uzel
 - 2 nedosažitelný protokol
 - 3 nedosažitelný UDP port
 - 4 fragmentace zakázána
 - 5 chyba explicitního směrování
 - 6 neznámá síť
 - 7 neznámý uzel
 - 9 administrativně uzavřená síť
 - 10 administrativně uzavřený uzel
 - 11 nedosažitelná síť pro službu
 - 12 nedosažitelný uzel pro službu
 - 13 komunikace administrativně uzavřena filtrací
- 4 0 sniž rychlost odesílání

Typy ICMP paketů II

- 5 změň směrování
 - 0 pro síť
 - 1 pro uzel
 - 2 pro síť pro daný typ služby
 - 3 pro uzel pro daný typ služby
- 8 0 echo request
- 9 0 odpověď na žádost o směrování
- 10 0 žádost o směrování
- 11 čas vypršel
 - 0 TTL
 - 1 defragmentace
- 12 chybný parametr
 - 0 chybné IP záhlaví
 - 1 schází volitelný parametr
- 13 0 časová synchronizace – požadavek
- 14 0 časová synchronizace – odpověď

Typy ICMP paketů III

- 17 0 žádost o masku
- 18 0 odpověď na žádost o masku



Přidělení IP adresy

- ruční x automatické
- statické x dynamické
- RARP – Reverse Address Resolution Protocol, rfc 903
 - přidělení adresy bezdiskové stanici
 - nepoužívá se
- BOOTP – Bootstrap Protocol, rfc 951, rfc 2132
 - využívá protokol UDP (67 server, 68 klient)
 - statické přidělení parametrů
 - adresa, jméno, maska, směrovače, time server, DNS, boot, log, lpr, forwarding, MTU, ...
- DHCP – Dynamic Host Configuration Protocol, rfc 2131
 - podobné BOOTP
 - více DHCP serverů
 - dynamické přidělení parametrů

Transportní protokoly

- slouží k přenosu dat mezi aplikacemi
- podpora aplikačního multiplexu - porty
- podpora řízení toku
- nepotvrzovaný UDP
- potvrzovaný TCP

Protokol UDP

- User Datagram Protocol, rfc 768
- nepotvrzovaná služba bez spojení
- umožňuje broadcast
- DNS, BootP, TFTP, SNMP ...

UDP hlavička

verze IP	délka záhlaví	typ služby	celková délka	
identifikace IP datagramu			příznaky	posunutí fragmentu
TTL	protokol vyšší vrstvy		kontrolní součet IP záhlaví	
IP adresa odesílatele				
IP adresa příjemce				
volitelné položky IP hlavičky				
zdrojový port UDP			cílový port UDP	
délka dat			kontrolní součet UDP záhlaví	
data				

- zdrojový, cílový port – 64K
- délka UDP dat – min 8
- kontrolní součet
 - nepovinný
 - aplikace mohou vyžadovat
 - výpočet z pseudohlavičky

UDP pseudohlavička

IP adresa odesílatele		
IP adresa příjemce		
binární nuly	protokol vyšší vrstvy	celková délka IP datagramu
zdrojový port UDP		cílový port UDP
délka dat		kontrolní součet UDP záhlaví
data		

- rfc 1071
- data zarovnána na sudý počet bajtů

Protokol TCP

- Transmission Control Protocol, rfc 793
- spolehlivá spojovaná služba
- multiplex
- duplexní přenos (piggybacking)
- proudový přenos (stream)
- potvrzování a opakování segmentů
- adaptivní přizpůsobení parametrů
- klouzavé okénko
- vytváření a rušení transportních spojení
- příjem dat z vyšší vrstvy a vytváření TCP paketů
- koncové řízení toku
- http, ssh, telnet, ...

TCP hlavička

verze IP	délka záhlaví	typ služby					celková délka				
identifikace IP datagramu						příznaky	posunutí fragmentu				
TTL		protokol vyšší vrstvy				kontrolní součet IP záhlaví					
IP adresa odesílatele											
IP adresa příjemce											
volitelné položky IP hlavičky											
zdrojový port TCP						cílový port TCP					
pořadové číslo odesílaného bajtu											
pořadové číslo očekávaného bajtu											
délka záhlaví	rezerva	U	A	P	R	S	F	délka okna			
kontrolní součet TCP						ukazatel naléhavých dat					
volitelné položky TCP hlavičky											
data											

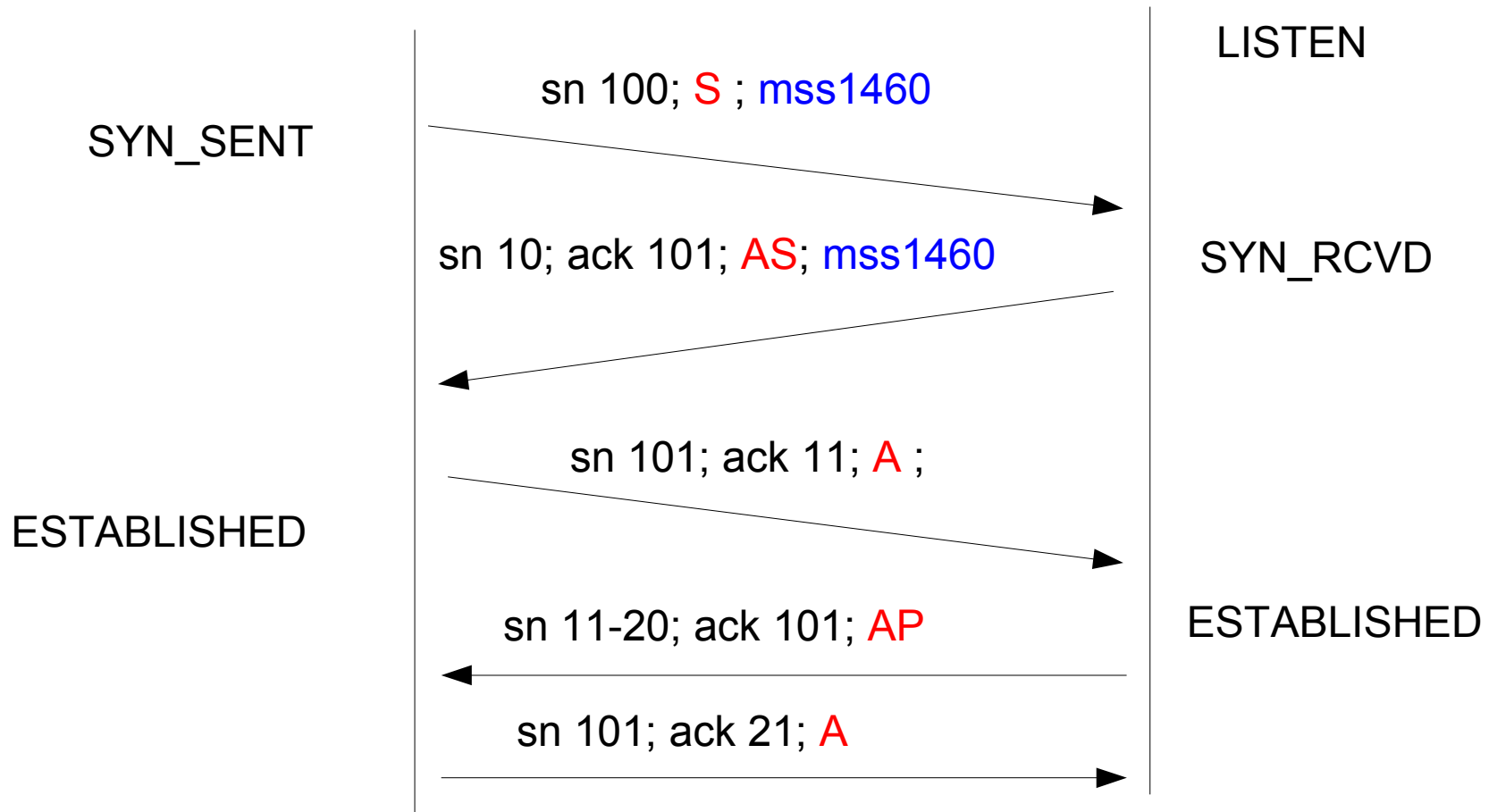
Položky TCP hlavičky

- pořadové číslo odesílaného bajtu – číslo prvního bajtu
- pořadové číslo očekávaného bajtu – poslední dobře přijatý bajt + 1
- číslování je v obou směrech nezávislé
- délka záhlaví v násobcích 32b
- délka okna – maximální akceptovaný přírůstek čísla očekávaného bajtu
- ukazatel urgentních dat – spolu s příznakem U ukazuje na konec naléhavých dat
- příznaky
 - U urgentní data
 - A potvrzení
 - P aplikační data
 - R odmítnutí spojení
 - S nová sekvence číslování
 - F ukončení spojení
- kontrolní součet – pseudohlavička a data, vyplnění dat na sudý počet

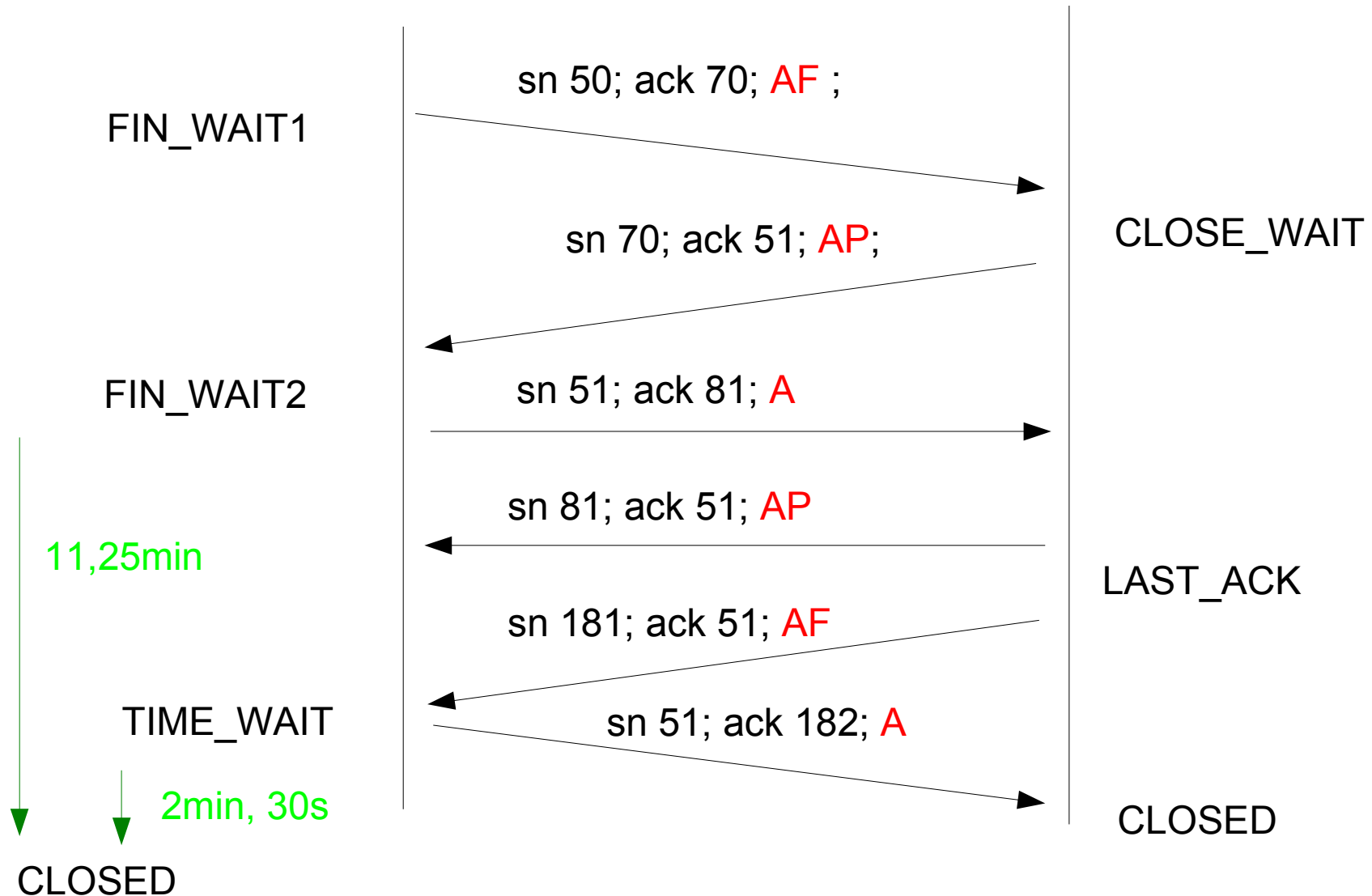
Volitelné položky TCP

- oproti volitelným položkám UDP se používají a přenáší
- maximální délka segmentu – přizpůsobení MTU (SYN)
- zvětšení okna – více než 64KB
- časové razítko
- čítač spojení
- ...

Navázání spojení



Ukončení spojení



Vylepšení provozu

- zpoždění odpovědi - sloučení dat a potvrzení
- technika okna
 - délka okna WIN určuje velikost povolených dat
- zahlcení sítě
 - přidání parametru CWND na straně odesílatele
 - $\min(\text{WIN}, \text{CWND})$
 - pomalý start – 1,2,4... segmenty; po výpadku potvrzení zpět na polovinu = Ssthresh; po úplném výpadku nový start
 - vyhýbání zahlcení – pomalejší start nad Ssthresh
- výpadek segmentu
 - segment mimo pořadí – opakování potvrzení
 - třetí duplikát – $\text{Ssthresh} = \text{CWND} / 2$, zopakování segmentu, $\text{CWND} = \text{Ssthresh} + 3 * \text{MSS}$
 - další duplikáty – zvětšení CWND o MSS
 - korektní potvrzení $\text{CWND} = \text{Ssthresh}$

Vylepšení provozu II

- nastavení timeoutu

- $SRTT' = (1 - \alpha) * SRTT + \alpha * RTT$
- $DEV = |RTT - SRTT|$
- $SDEV' = (3/4) * SDEV + (1/4) * DEV$
- $T = SRTT' + 2 * SDEV'$
- α (0-1), 0.5
- pokud RTT skokově překročí T , $T' = 2 * T$

existují i jiné implementace

...