



Počítačové a komunikačné siete

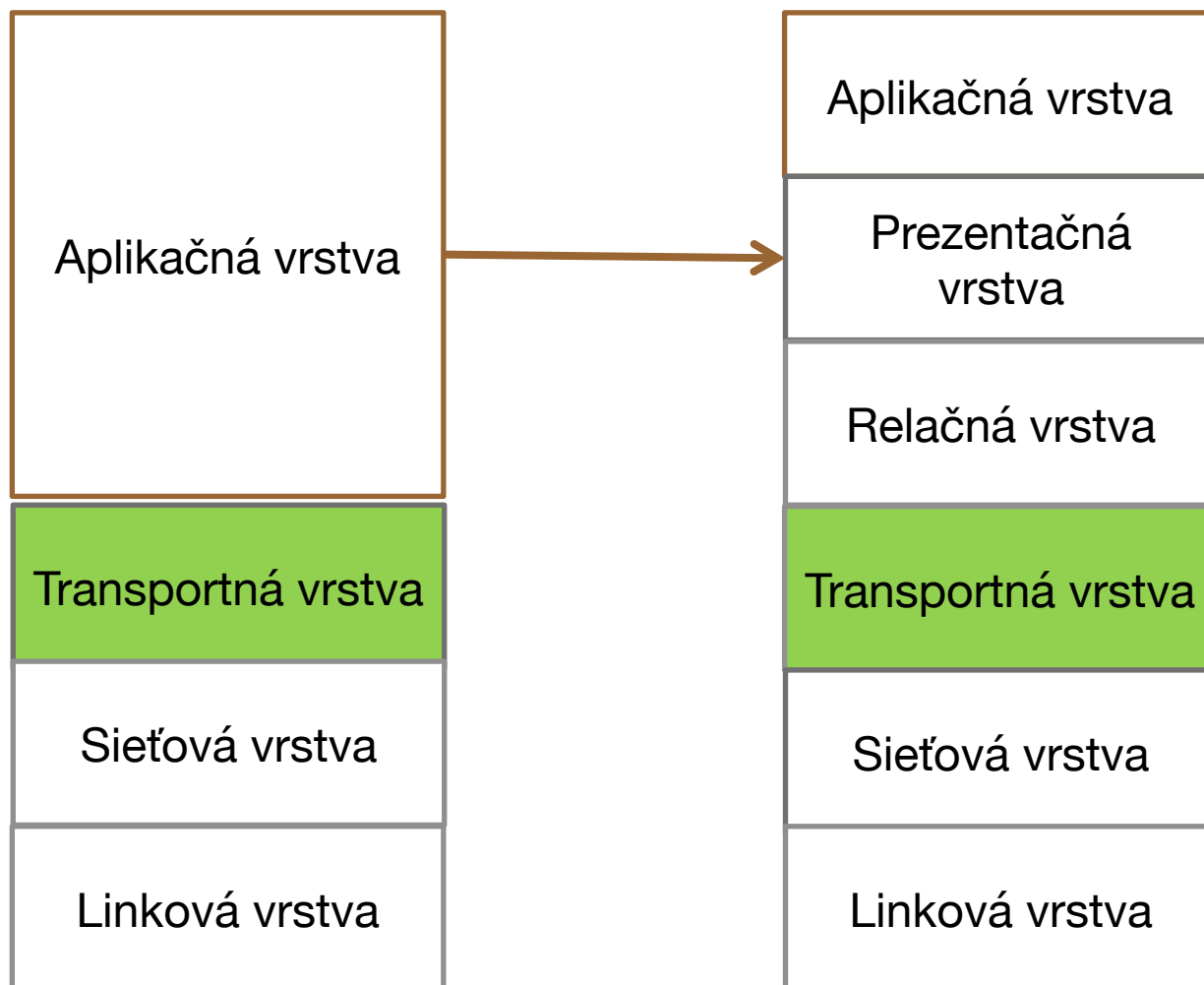
# TCP 2 riadenie toku dát Internet Protocol (IP)

Prednáška 5

# Opakovanie minulej prednášky

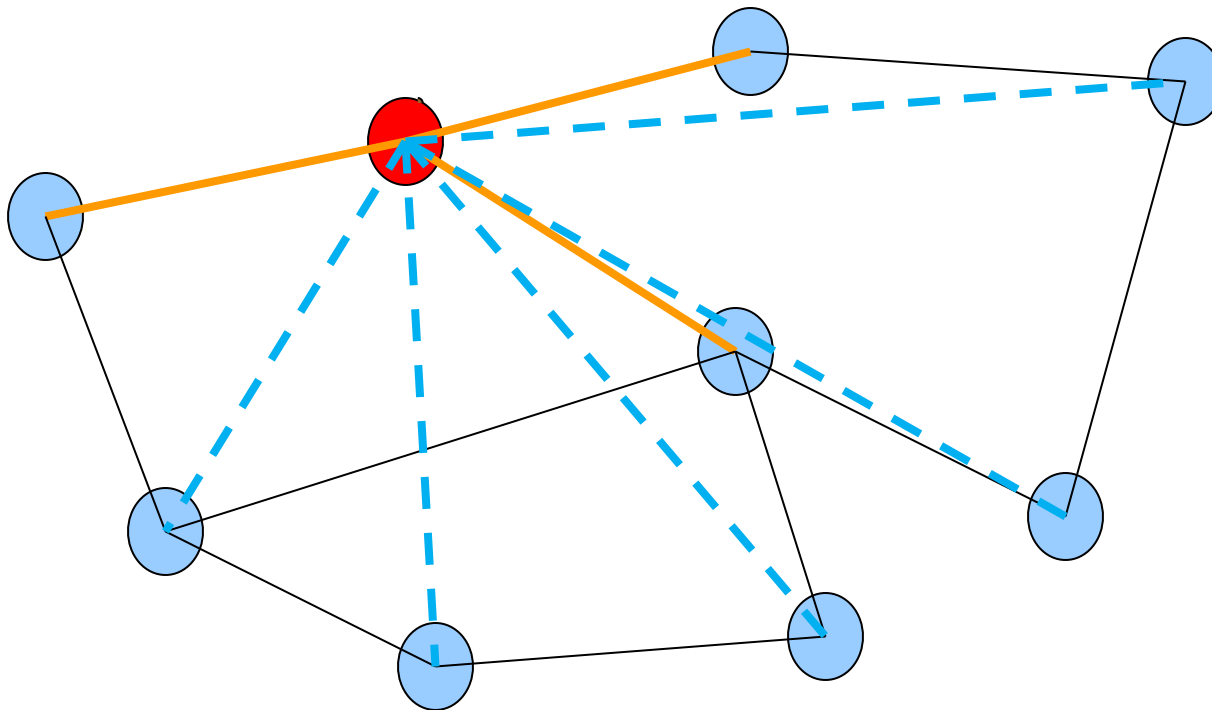
- » TCP (používané HTTP)
  - Riadenie toku (pomalé / rýchle linky)
  - Potvrdzovanie dát (ACK, NACK)
  - Ukončenie spojenia
  - Znovuodoslanie dát

# Transportná vrstva



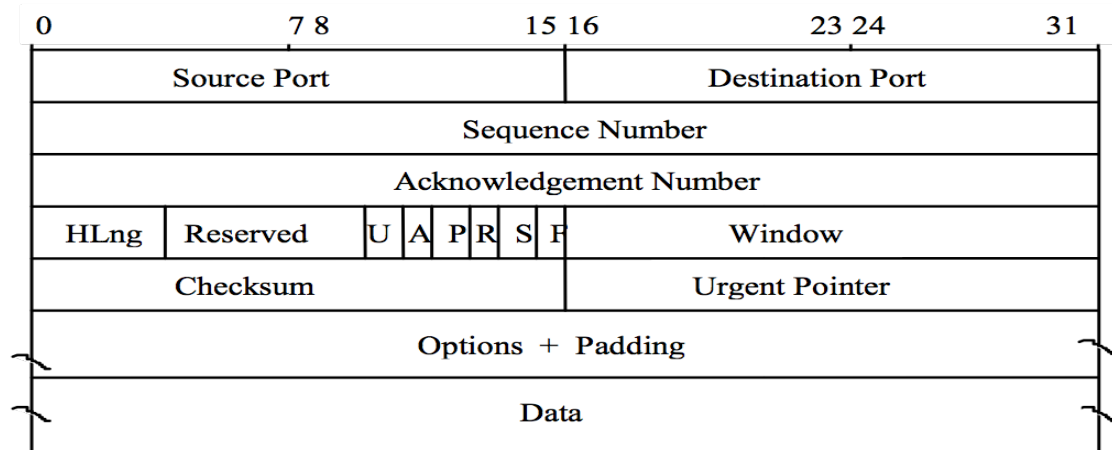
# Pohľad vrstiev na topoloógiu siete

## Transportná vrstva



# Protokol TCP

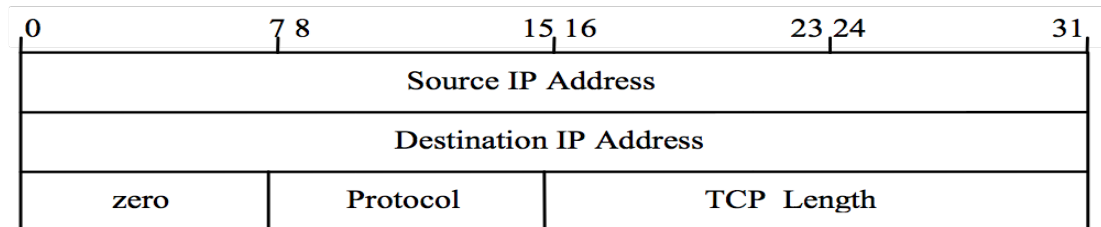
TCP segment



pseudohlavička

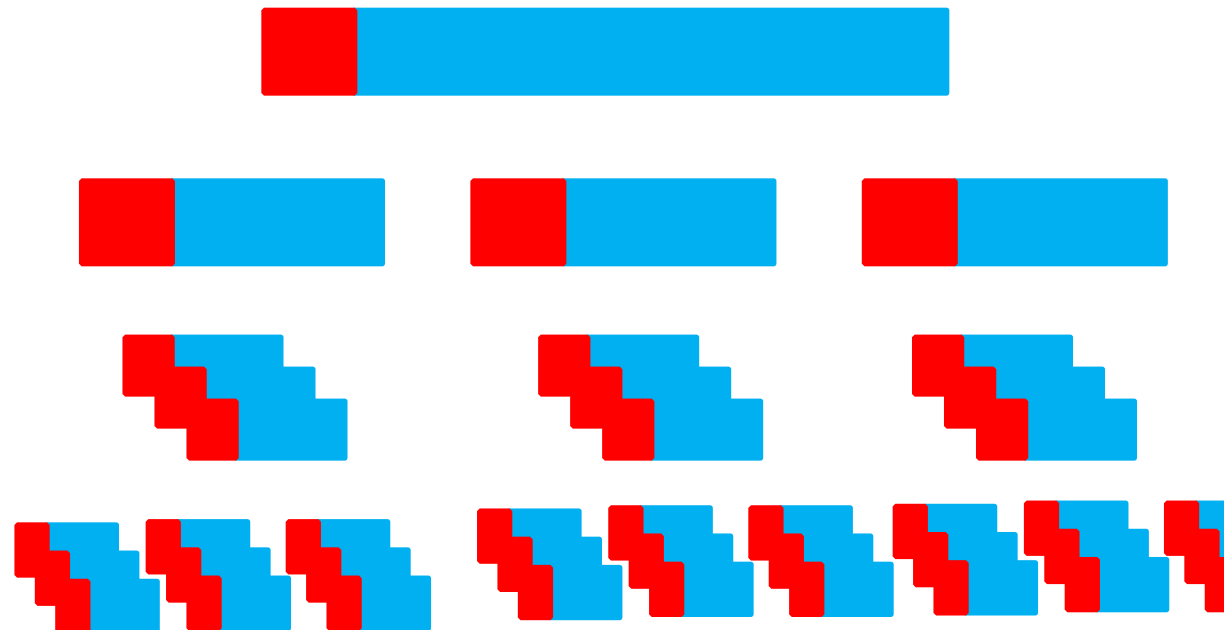
Niektoré voliteľné položky  
(options):

kind	length	význam
2	4	MSS
3	3	zváženie okna
4	2	povolenie SACK
5	prem.	SACK



# Vnáranie stránky do paketov

Aplikácia (Web)
Aplikačná vrstva (HTTP)
Transportná vrstva (TCP)
Sieťová vrstva (IP)
Linková vrstva (Ethernet/WiFi)



# Číslo segmentu, číslo potvrdenia

Aplikácia (Web)
Aplikačná vrstva (HTTP)
Transportná vrstva (TCP)



0

4999

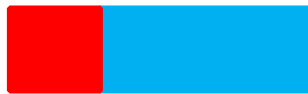


0 1499

1500 2999

3000 4499

SN: 111



SN: 111

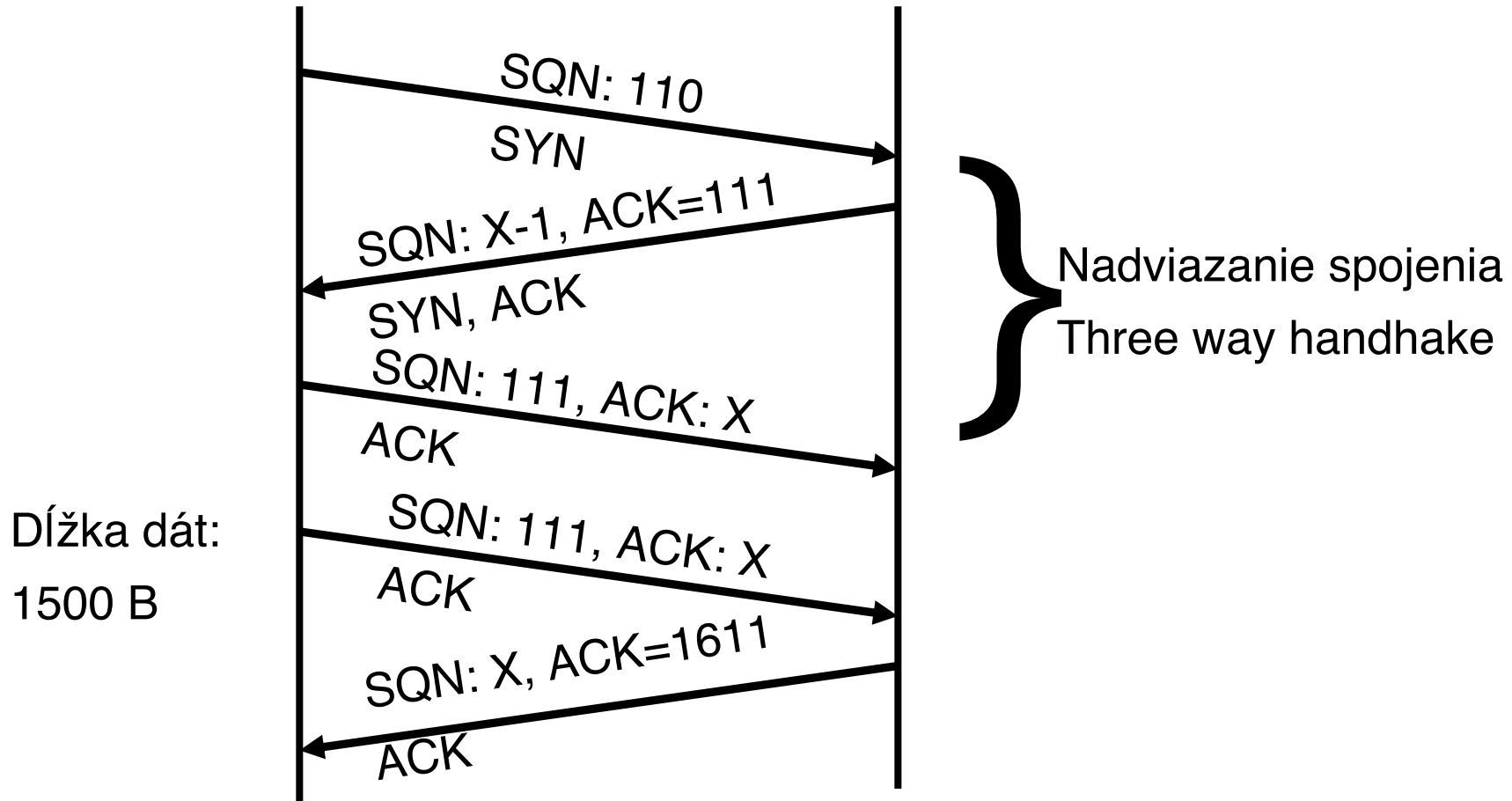
SN:1611

SN: 3111

ACK: 1611

ACK:3111. ACK:4611

# Segment number, ACK number



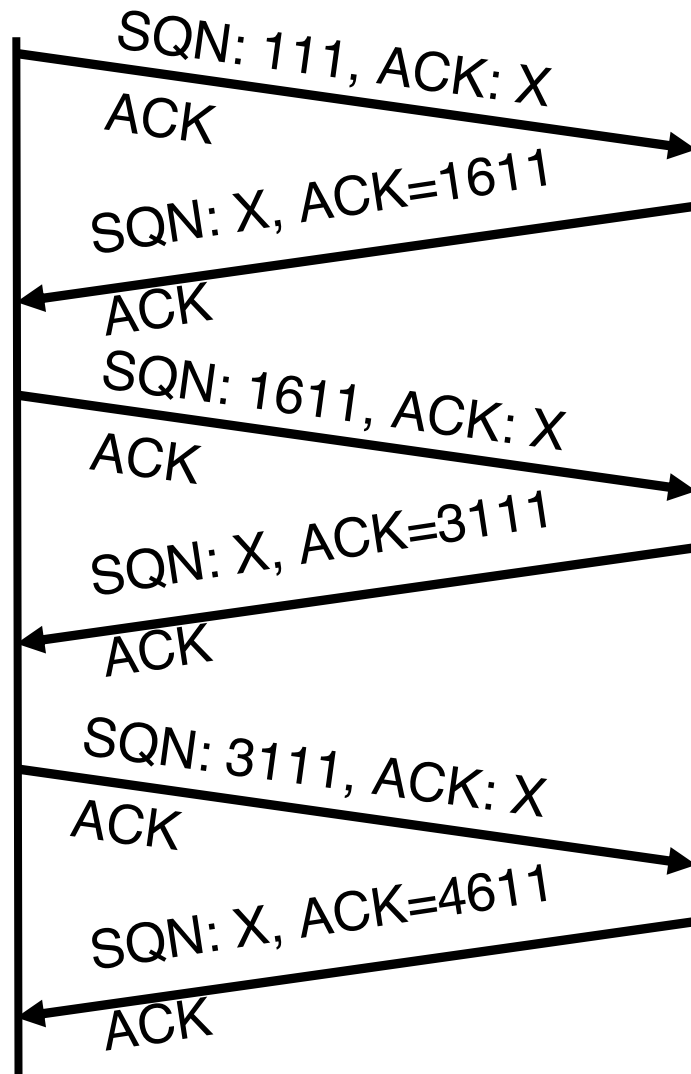


# Segment number, ACK number

Dĺžka dát:  
1500 B

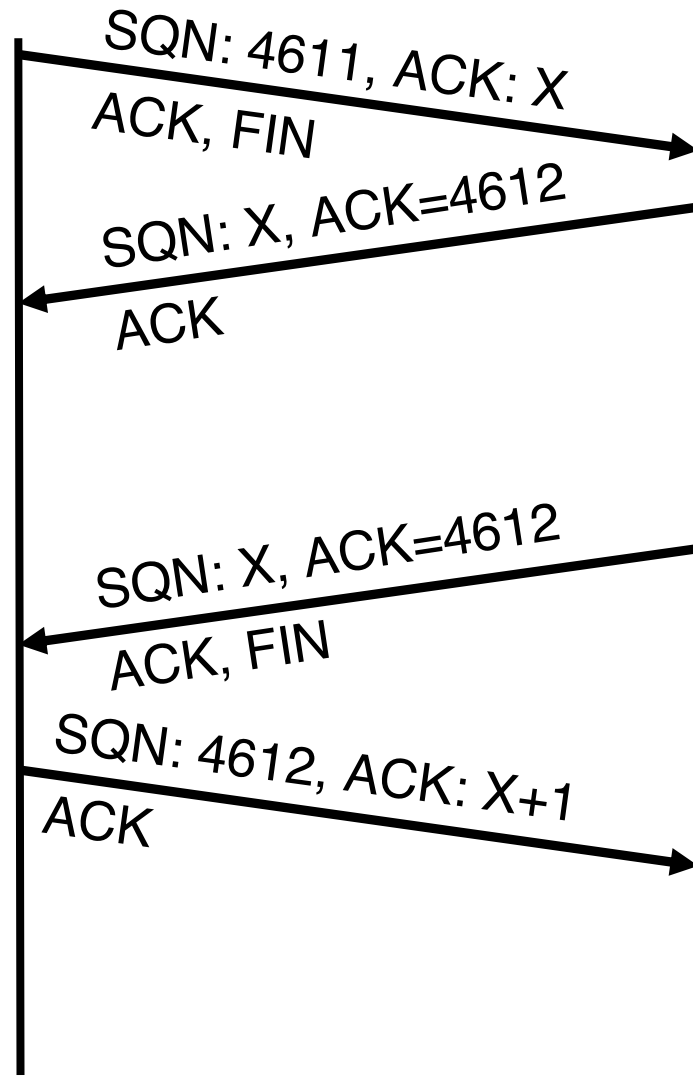
Dĺžka dát:  
1500 B

Dĺžka dát:  
1500 B



Opakovaná správa  
z predchádzajúceho slajdu

# Segment number, ACK number



# Čo nás čaká na prednáške

- » Kontrola toku dát a zahltenia
- » IP protokol

# Kontrola toku dát

- » Cieľ: čo najefektívnejšie preniesť dáta cez sieť bez straty
- » Čo k tomu potrebujem?

# Kontrola toku dát

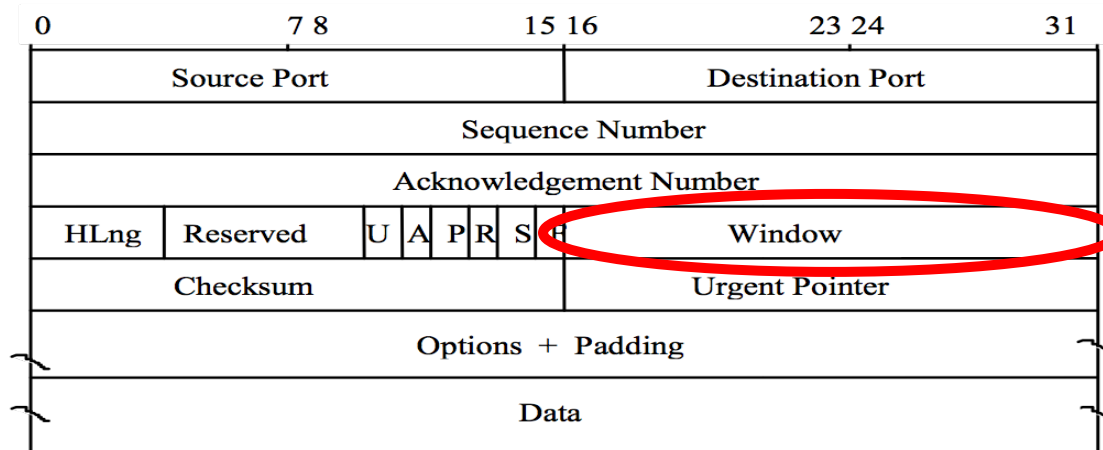
- » Cieľ: čo najefektívnejšie preniesť dáta cez sieť bez straty
- » Čo k tomu potrebujem?
  - Informáciu o stave:
    - Vysielača (Okno vysielača)
    - Prijímača (Okno prijímača)
    - Informáciu o stave siete

# Protokol TCP – kontrola zahltenia

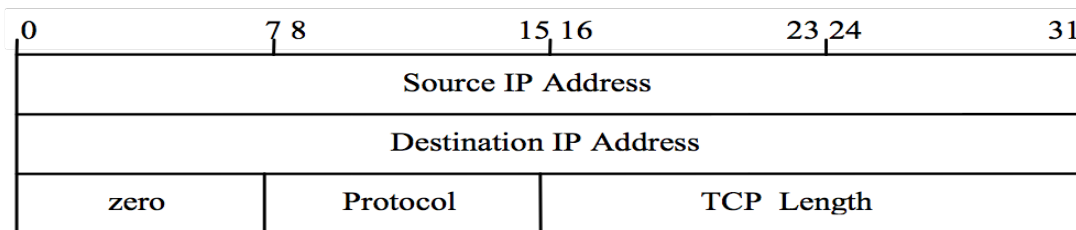
## Okno prijímača

- » Určuje, koľko dát môže prijímač prijať
- » vysielateľ nemusí poslať celé okno
- » prijímač môže meniť veľkosť okna
- » prijímač nemusí čakať na naplnenie okna pred vyslaním ACK
- » kumulované potvrdzovanie
- » test nulového okna
- » potvrdenie – ARQ s návratom, selektívne ARQ

# Protokol TCP



Okno prijímača v B



# Protokol TCP – kontrola zahltenia

## Okno vysielaca

- » Kde ho nájdem?
  - Je iba virtuálne
- » Určuje, koľko dát očakávam, že prejde sieťou



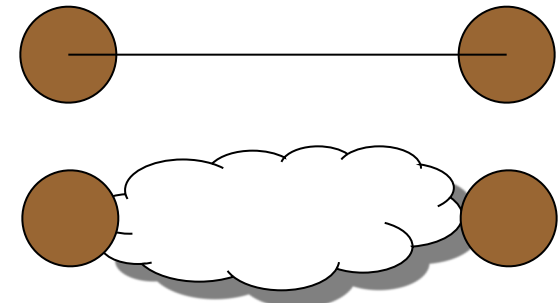
# Protokol TCP – kontrola zahltenia

## Stav siete

### Riadenie toku dát (flow control)

- zabezpečiť, aby rýchlosť prenosu dát medzi dvoma prepojenými uzlami neviedla k zahlteniu prijímajúceho uzla

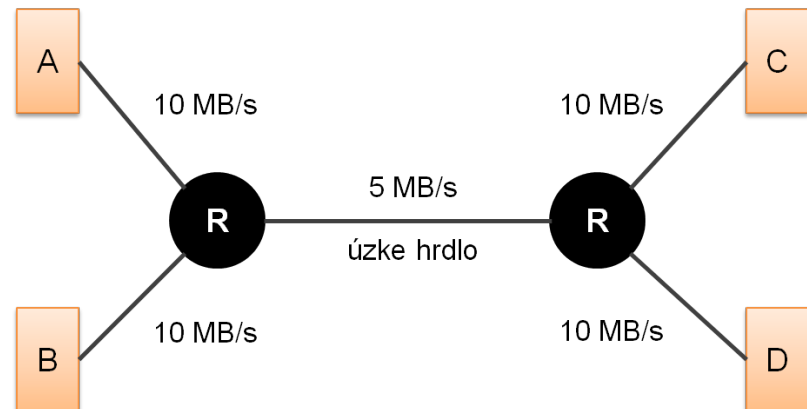
Riešenie: väčšinou spätná väzba



### Riadenie zahltenia (congestion control)

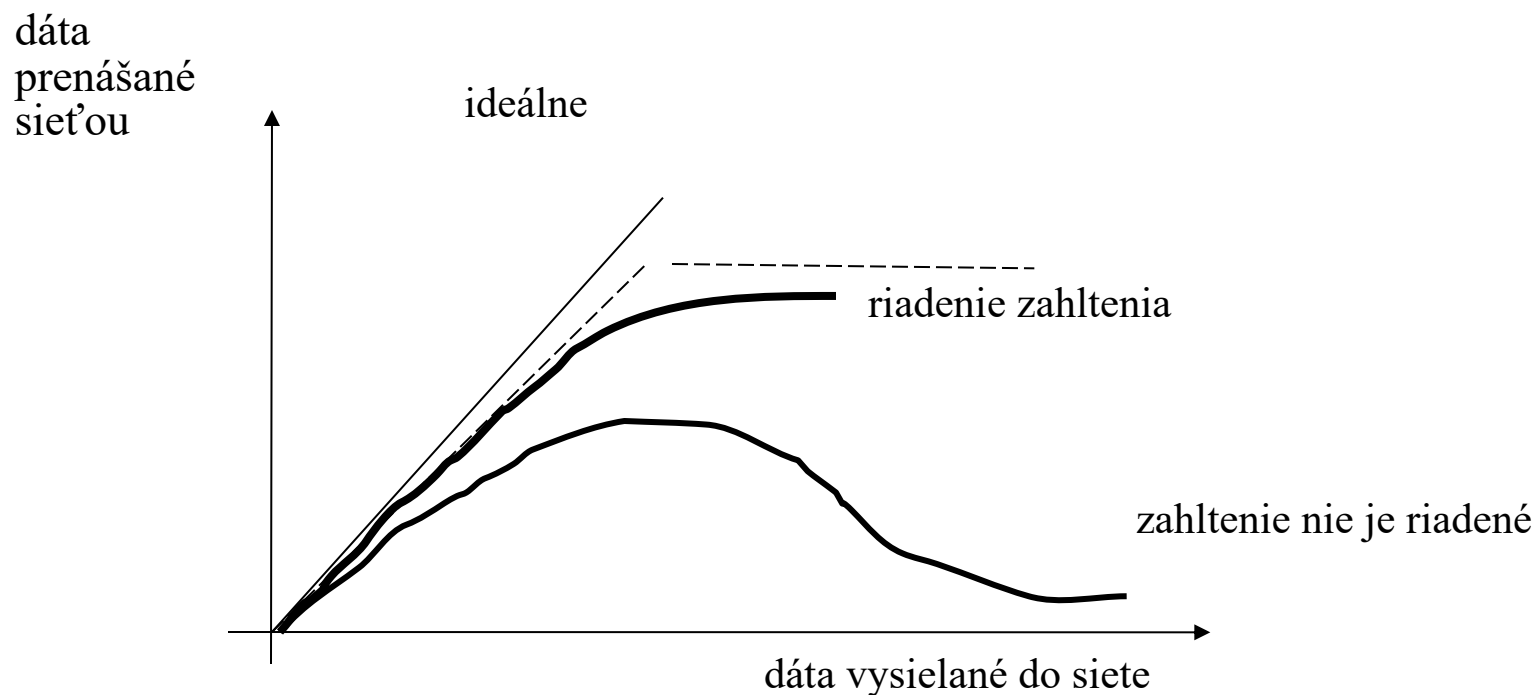
- zabezpečiť, aby v komunikačnej podsieti (sieti) nevzniklo blokovanie prenosu dát (zníženie celkovej prenosovej kapacity)

Riešenie: prevencia a spätná väzba



# Riadenie toku dát a zahltenia /1

zahltenie prijímajúceho uzla =>  
lokálne zahltenie siete =>  
globálne zahltenie siete

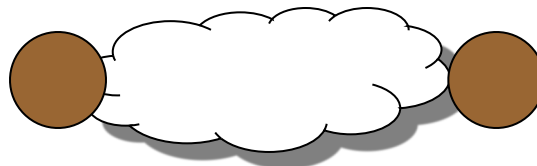
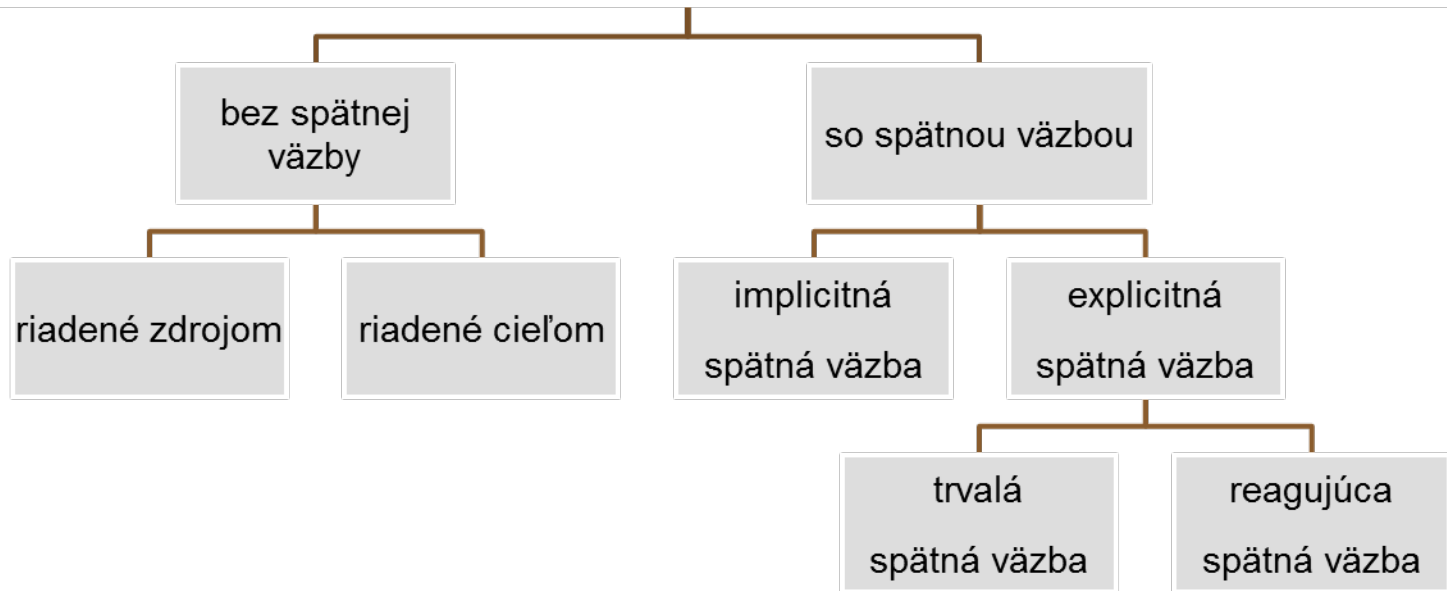


# Protokol TCP – riadenie toku dát a zahltenia

- problémy
  - kapacita siete
  - kapacita prijímača
- množstvo vysielaných dát určované oknami
  - rwnd - okno (dynamické) na strane prijímača (receiver window), veľkosť určuje prijímač v priebehu spojenia
  - cwnd – okno na strane vysielacza (congestion window), predchádzanie zahlteniu
  - aktuálne okno =  $\min(\text{rwnd}, \text{cwnd})$  - max. počet vyslaných nepotvrdených dát

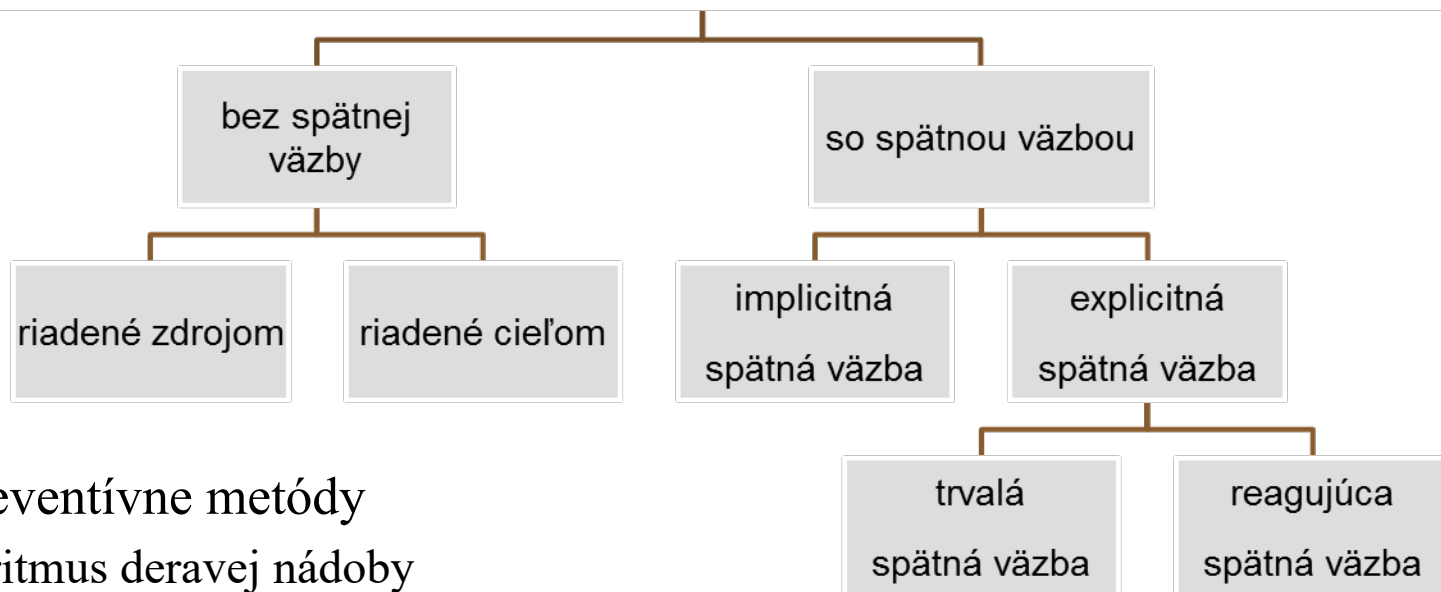
# Riadenie toku dát a zhltenia /4

## Klasifikácia prístupov riadenia toku dát a zhltenia



# Riadenie toku dát a zahľtenia /4

## Klasifikácia prístupov riadenia toku dát a zahľtenia



napr. preventívne metódy

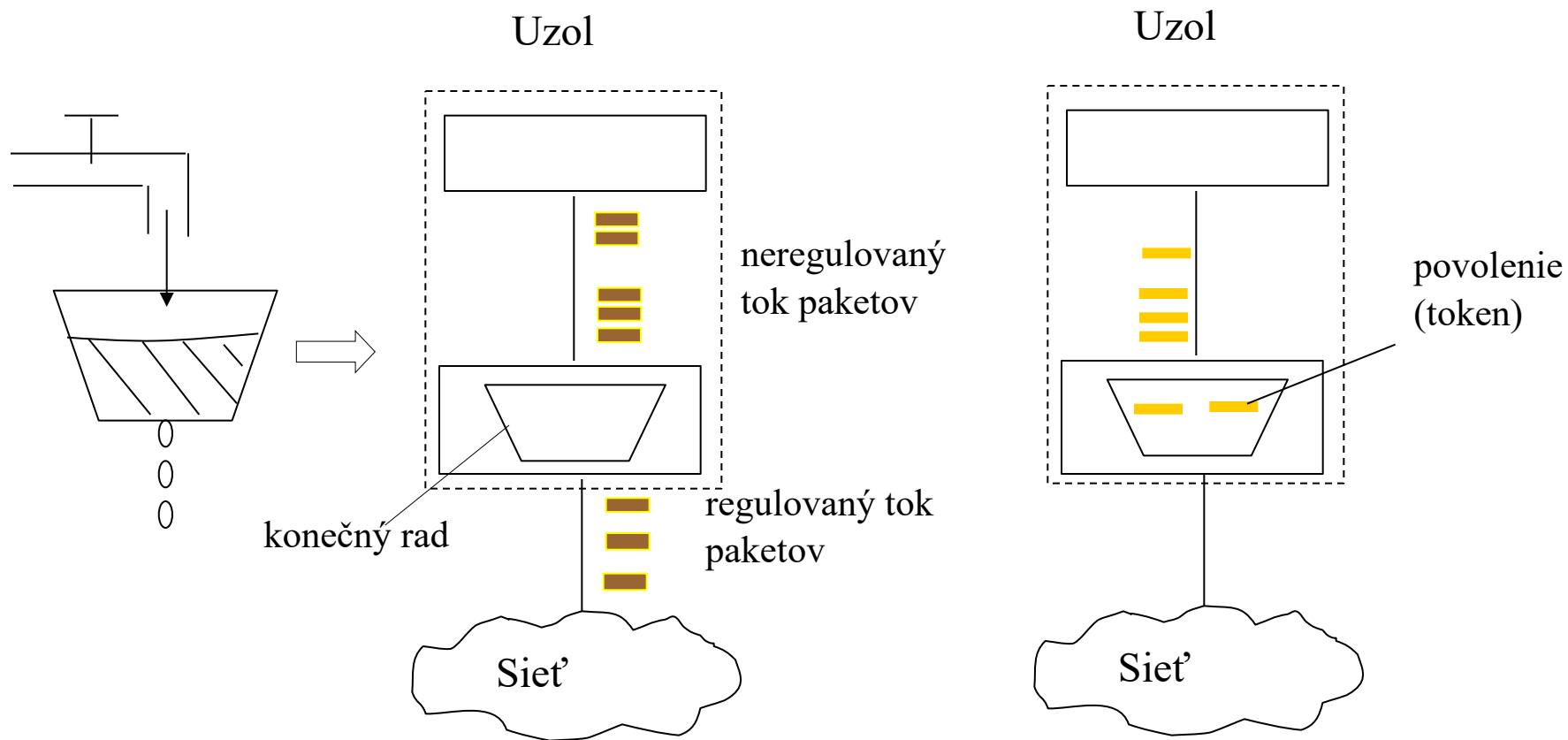
- algoritmus deravej nádoby  
( the leaky bucket alg. )
- algoritmus nádoby povolení  
( the token bucket alg. )

# Preventívne metódy

algoritmus deravej nádoby ( the leaky bucket alg. )

algoritmus nádoby povolení ( the token bucket alg. )

princíp:

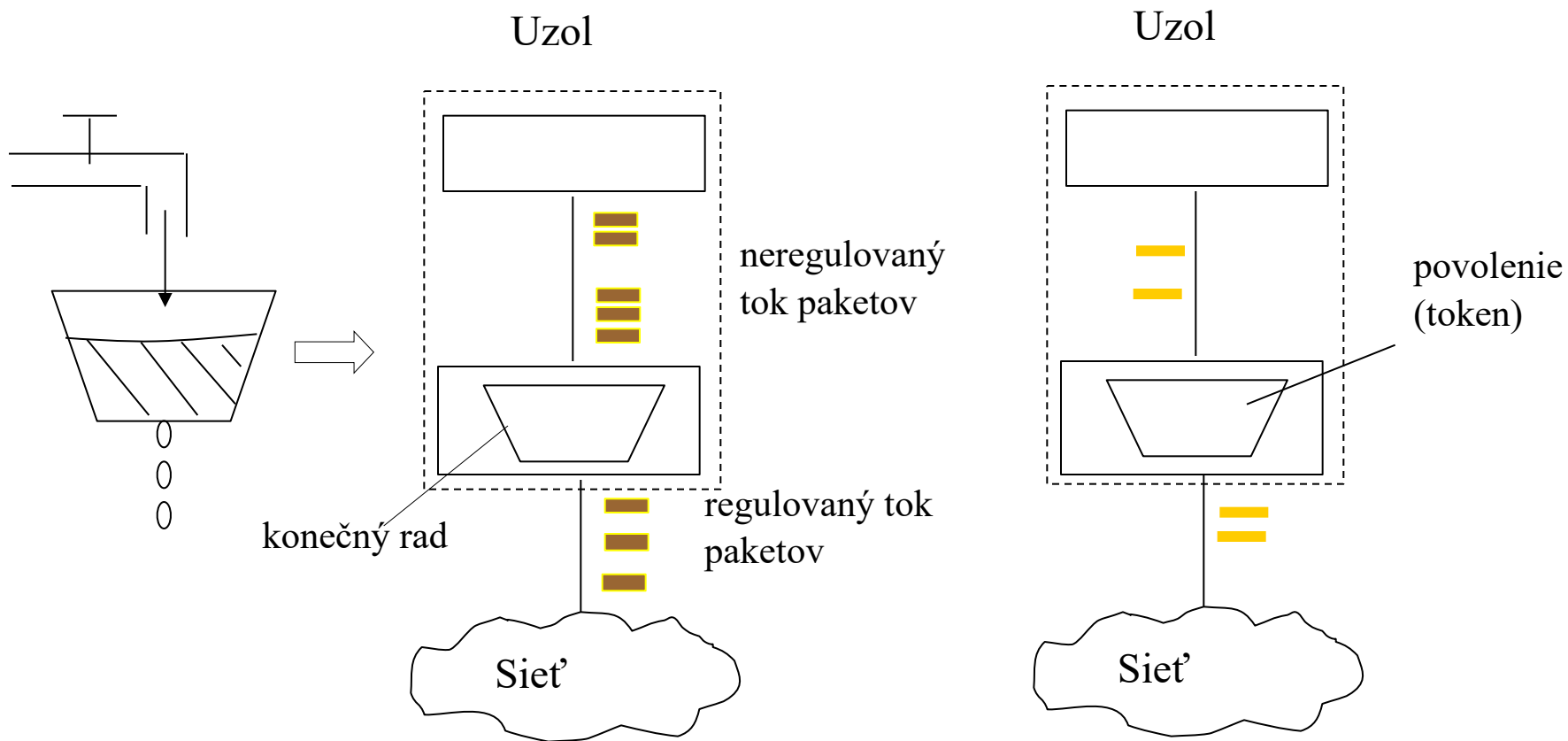


# Preventívne metódy

algoritmus deravej nádoby ( the leaky bucket alg. )

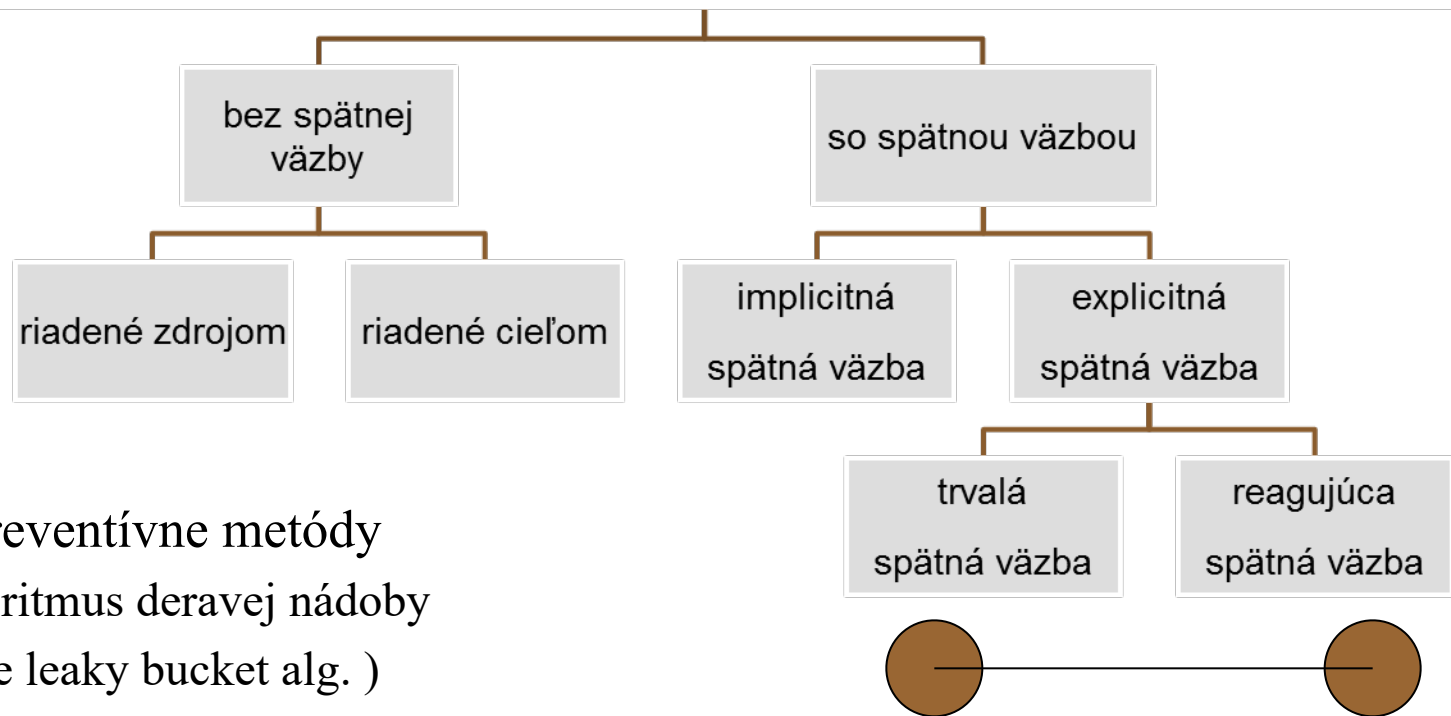
algoritmus nádoby povolení ( the token bucket alg. )

princíp:



# Riadenie toku dát a zahltenia /4

## Klasifikácia prístupov riadenia toku dát a zahltenia



napr. preventívne metódy

- algoritmus deravej nádoby  
( the leaky bucket alg. )
- algoritmus nádoby povolení  
( the token bucket alg. )

napr.

stop-and-wait, static window ..... 1. generácia  
dynamic window, dynamic rate... 2. generácia  
(end-to-end, hop-by-hop)

napr. choke packets (ICMP – source quench)



# Protokol TCP – prenos dát

## Interaktívne vysielanie

- problém malých IP datagramov (tiny datagrams) - Nagle\_ov algoritmus
  - TCP spojenie môže mať len jeden vyslaný nepotvrdený segment
  - použitie algoritmu

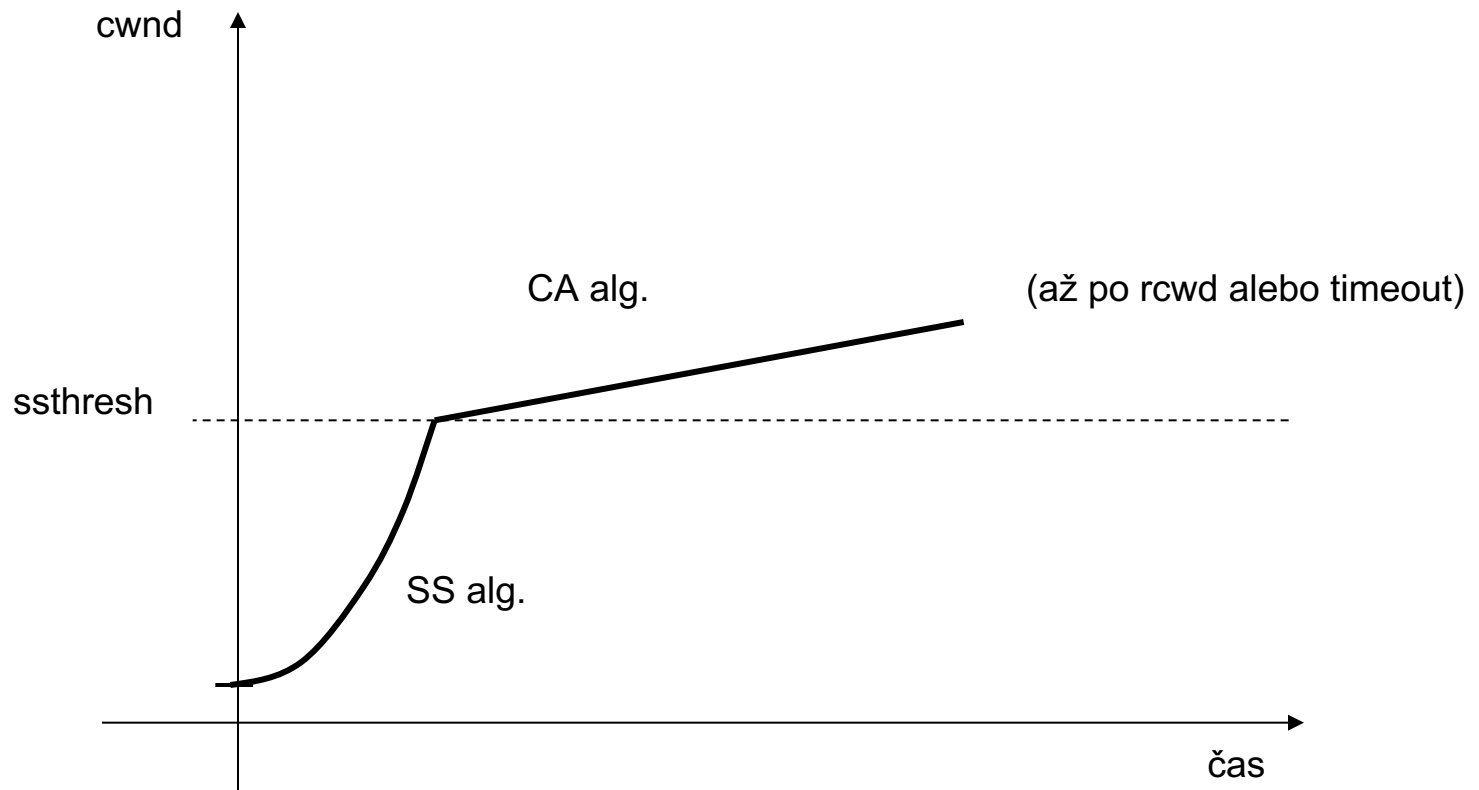
## Neinteraktívne vysielanie

- rýchly vysielateľ, pomalý prijímač
  - syndróm SWS (Silly Window Syndrom)
  - Clarke\_ov algoritmus – zabrániť vysielateľu posielat' „malé“ okno

# Protokol TCP – riadenie toku dát a zahltenia

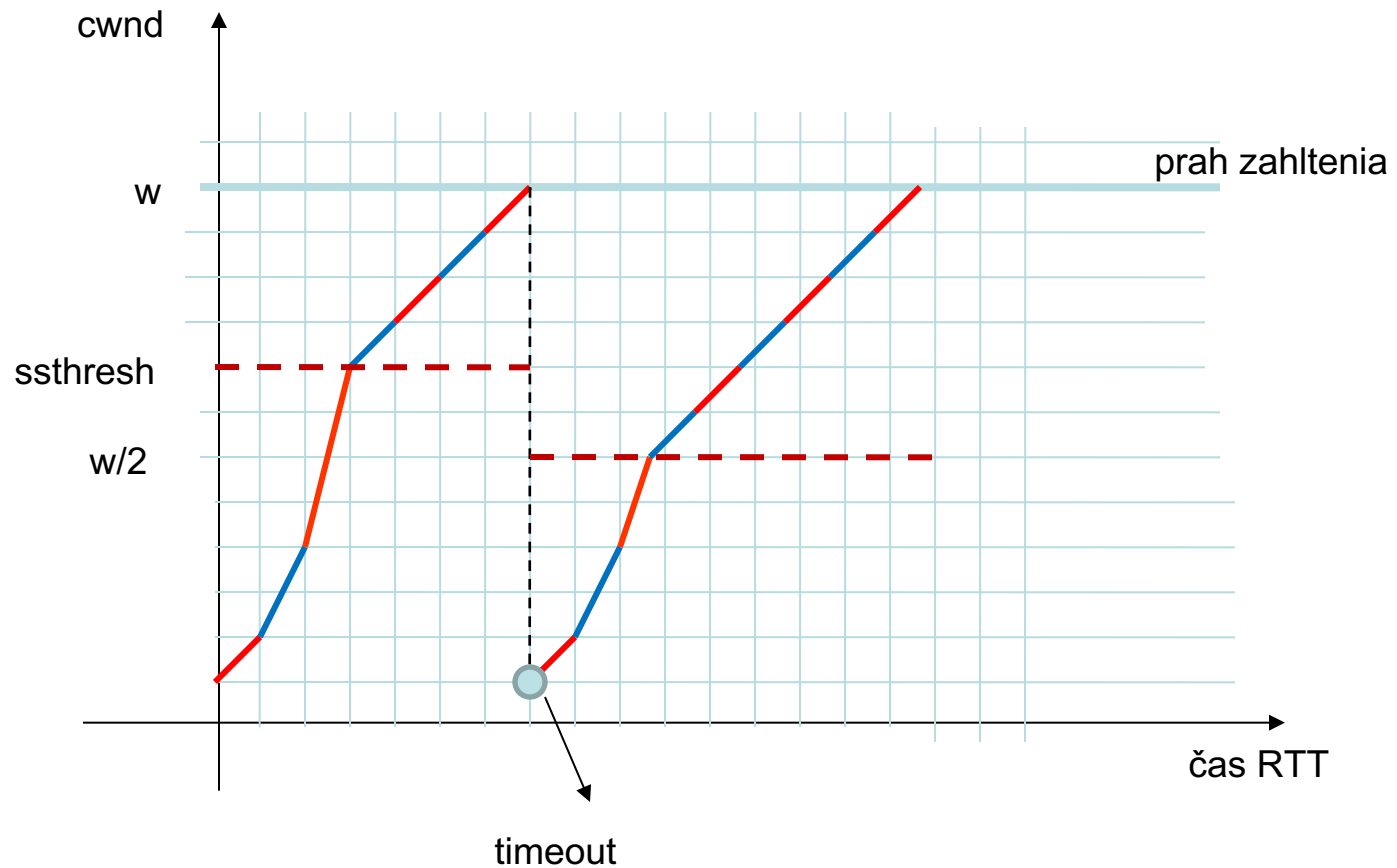
- cwnd – určuje dynamicky vysielateľ
  - cwnd = 1, vyšle 1 segment a čaká na potvrdenie
  - cwnd = 2, vyšle 2 segmenty a čaká na potvrdenie
  - cwnd = 4, .....
  - postupne vysielateľ  $2^n$  segmentov (pomalý štart)
- prah zvyšovania okna – ssthresh (slow start threshold), po prekročení prahu vysoká pravdepodobnosť vzniku zahltenia
- algoritmus pomalého štartu (Slow Start - SS) a algoritmus vyhýbania sa zahlteniu (Congestion Avoidance - CA)
  - mss, cwnd, ssthresh, rwnd
  - pomalý štart:  $\text{cwnd} \leq \text{ssthresh}$
  - vyhýbanie sa zahlteniu:  $\text{cwnd} > \text{ssthresh}$

# Protokol TCP – SS a CA



# Protokol TCP – algoritmus SS a CA

## TCP Tahoe



# TCP - riadenie toku dát a zahltenia

Ako je limitovaná rýchlosť vysielania ?

$$\text{rýchlosť (rate)} = \text{cwnd} / \text{RTT}$$

Ako pozná, že je zahltenie ?

implicitná- , explicitná detekcia

detekcia na základe straty paketu (loss-based detection), napr. TCP Reno, HS-TCP, S-TCP

detekcia na základe oneskorenia paketu (delay-based detection), napr. TCP Vegas, Fast TCP

Aký algoritmus sa použije na riadenie zahltenia ?

aditívny-, multiplikatívny-, iný princíp

napr. AIMD (TCP Reno, HS-TCP), MIMD (Fast TCP),

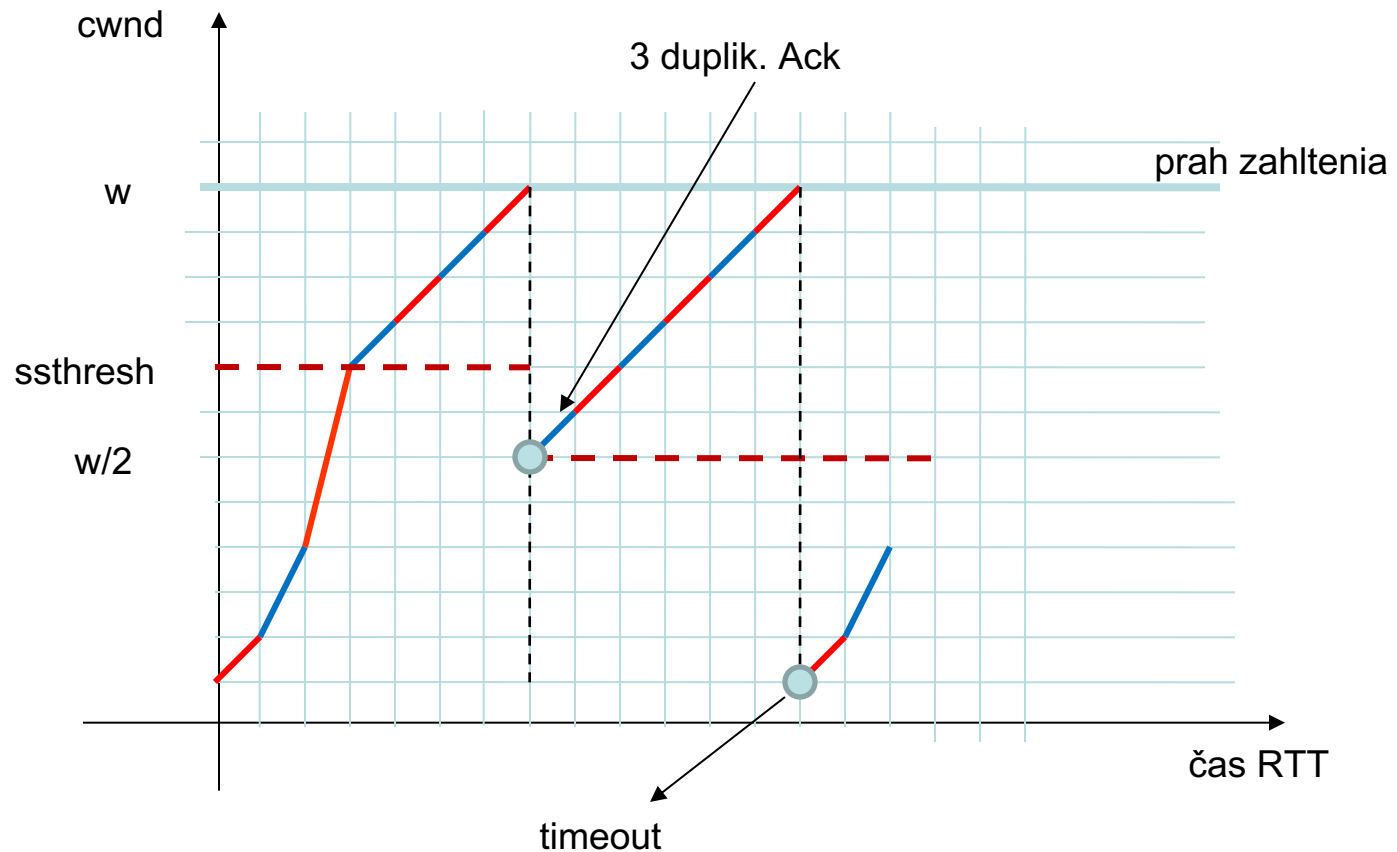
iné (BIC, CUBIC)

# Protokol TCP – prenos a zotavenie

- **zrýchlený prenos** (Fast Retransmission, FRet):  
ak počet duplikátov ACK je  $\geq 3$  – veľká pravdepodobnosť, že segment je stratený a vysielateľ opakuje prenos segmentu bez čakania na vypršanie časovača
- **rýchle zotavenie** (Fast Recovery, FRec):  
nezačína sa pri prenose s  $cwnd=1$

# Protokol TCP – algoritmus SS a CA

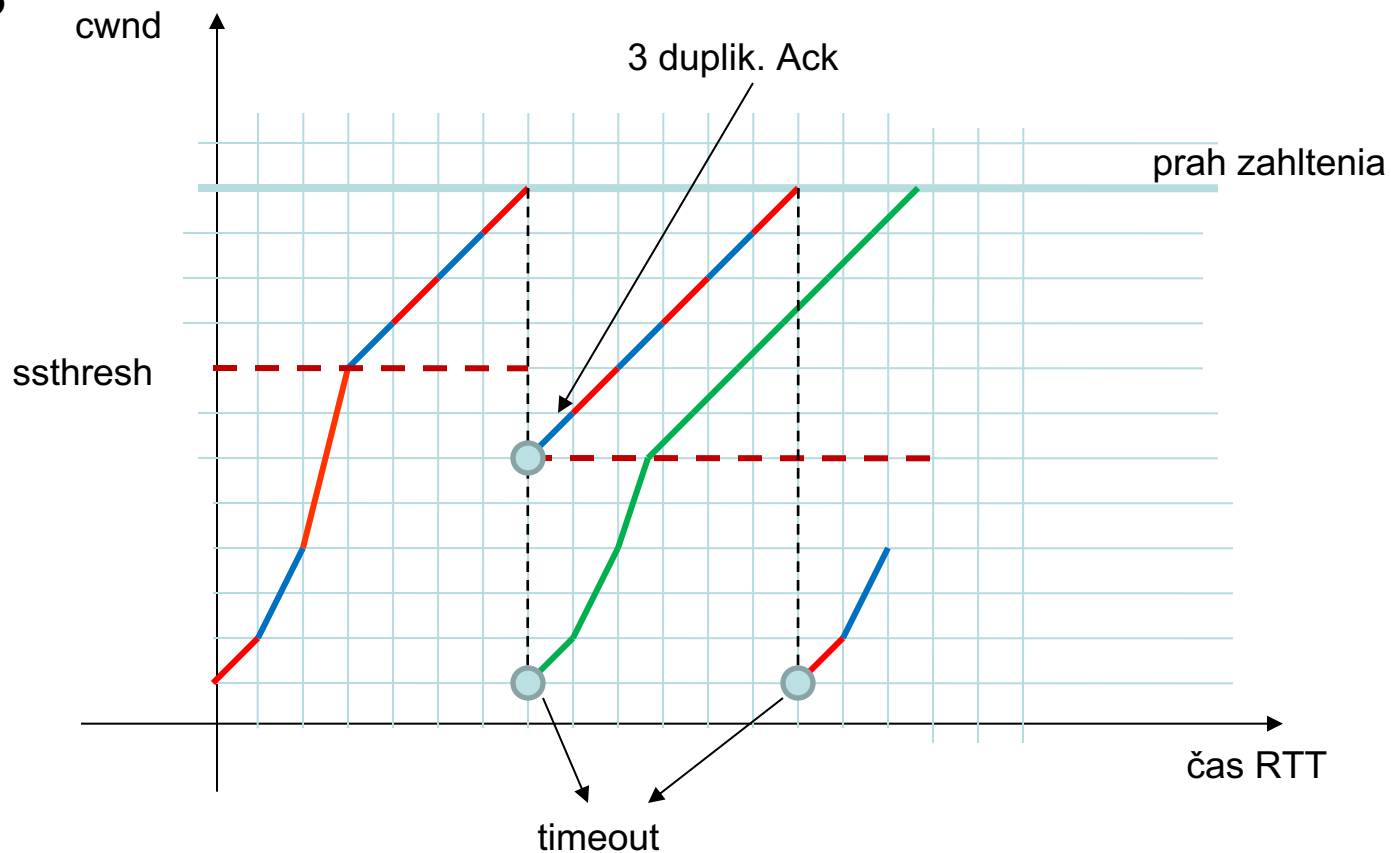
## TCP Reno



# Protokol TCP – algoritmus SS a CA

TCP Tahoe

TCP Reno





## Modifikácia veľkosti okna

- Základné prístupy
  - znižovanie, zvyšovanie veľkosti okna
- Aditívny, multiplikatívny prístup

AI- , MI- , AD- , MD-

$x(t+1) = a_i + b_i * x(t)$ , ak  $y(t) = 0$  (nie je zahltenie)

$x(t+1) = a_d + b_d * x(t)$ , ak  $y(t) = 1$  (je zahltenie)

# Modifikácia veľkosti okna

$x(t+1) = a_i + b_i * x(t)$ , ak  $y(t) = 0$  (nie je zahltenie)

$x(t+1) = a_d + b_d * x(t)$ , ak  $y(t) = 1$  (je zahltenie)

- AIAD (Additive Increase, Additive Decrease )

$$a_i > 0, \quad a_d < 0, \quad b_i = 1, \quad b_d = 1$$

- MIMD (Multiplicative Increase, Multiplicative Decrease)

$$a_i = 0, \quad a_d = 0, \quad b_i > 1, \quad 0 < b_d < 1$$

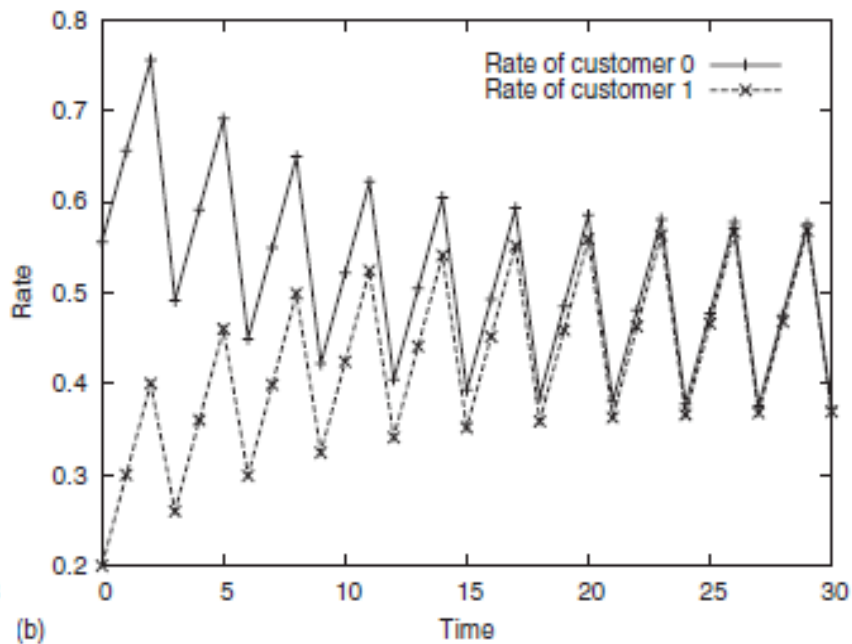
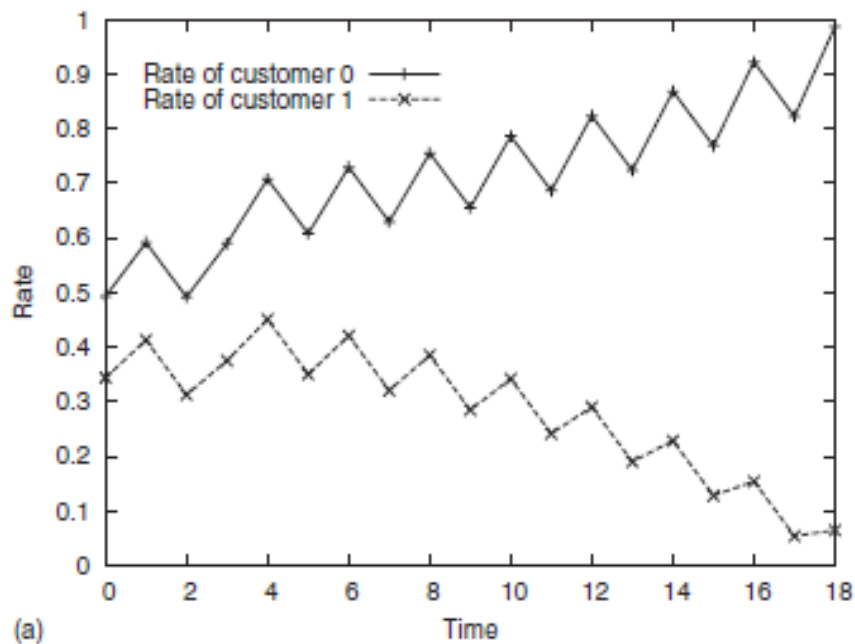
- MIAD (Multiplicative Increase, Additive Decrease)

$$a_i = 0, \quad a_d < 0, \quad b_i > 1, \quad b_d = 1$$

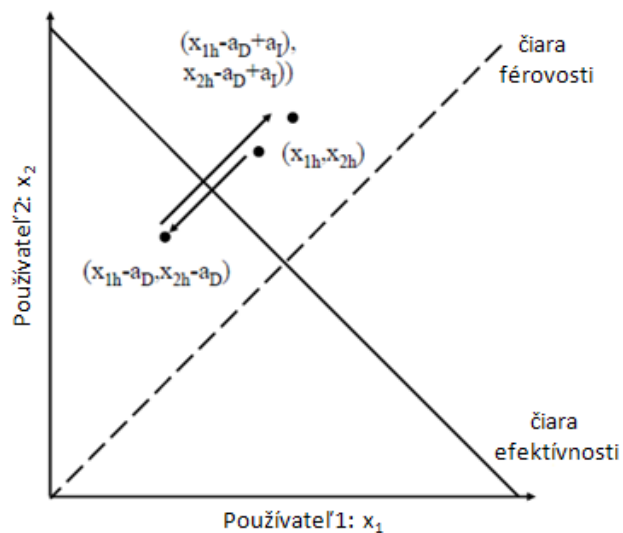
- AIMD (Additive Increase, Multiplicative Decrease)

$$a_i > 0, \quad a_d = 0, \quad b_i = 1, \quad 0 < b_d < 1$$

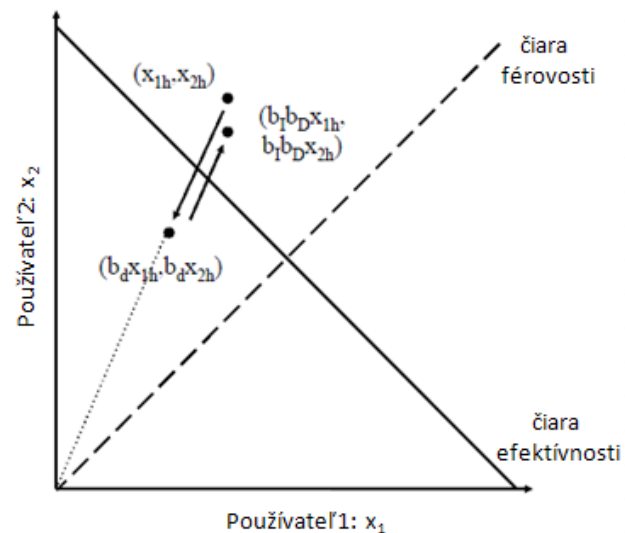
# Porovnanie prístupov MIAD a AIMD



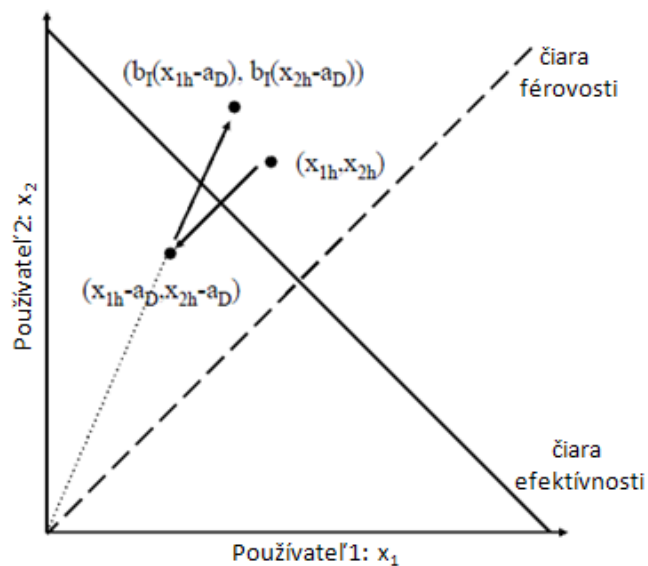
## AIAD



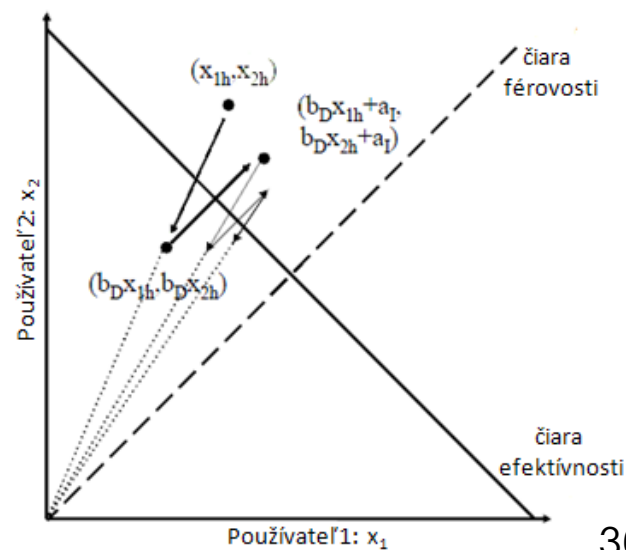
## MIMD



## MIAD



## AIMD



# Niektoré varianty TCP

## TCP Tahoe

- detekcia straty (len na základe časovača RTO)
- pomalý štart SS, - vyhýbanie sa zahlteniu CA, - zrýchlený prenos FRet

## TCP Reno

- detekcia straty (na základe časovača RTO a duplicitných potvrdení)
- pomalý štart SS, - vyhýbanie sa zahlteniu CA, - zrýchlený prenos FRet, - rýchle zotavenie FRec
- AIMD    AI:  $w \leftarrow w + 1/w$     MD:  $w \leftarrow w - \frac{1}{2} * w$

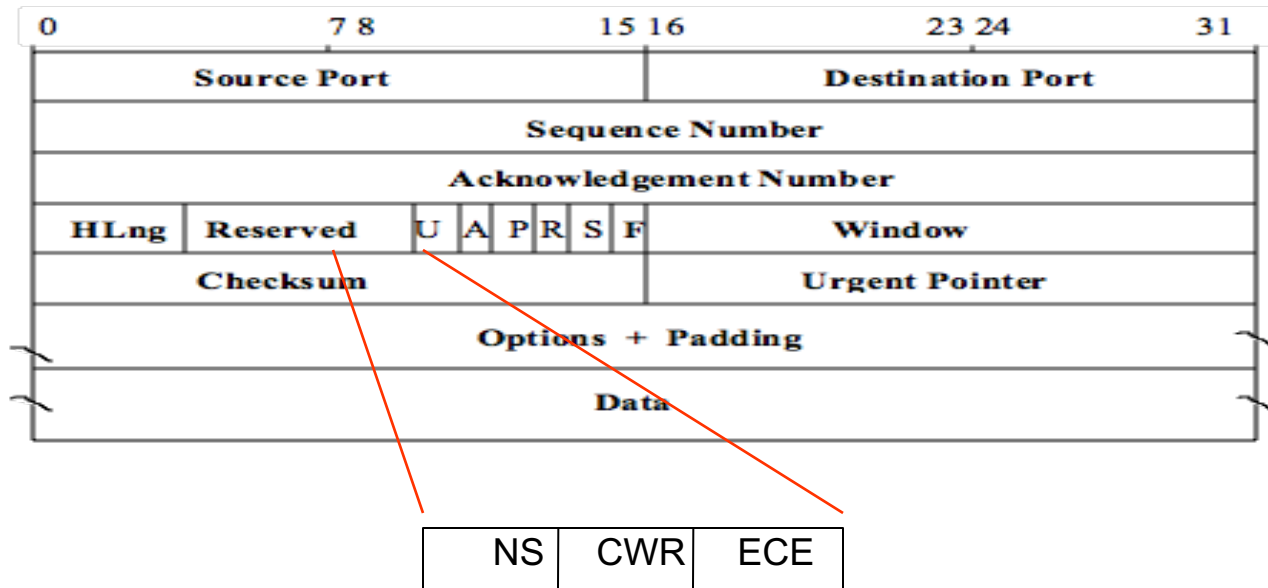
## TCP Vegas

- detekcia oneskorenia
- modifikácia pomalého štartu (expon. zvýšenie cwnd len počas každého 2. RTT)
- AIAD

očakávana a skutočná rýchlosť vysielania

# Protokol TCP - ECN

explicitné riadenie zahltenia



ECE - ECN-Echo

CWR - Congestion Window Reduced

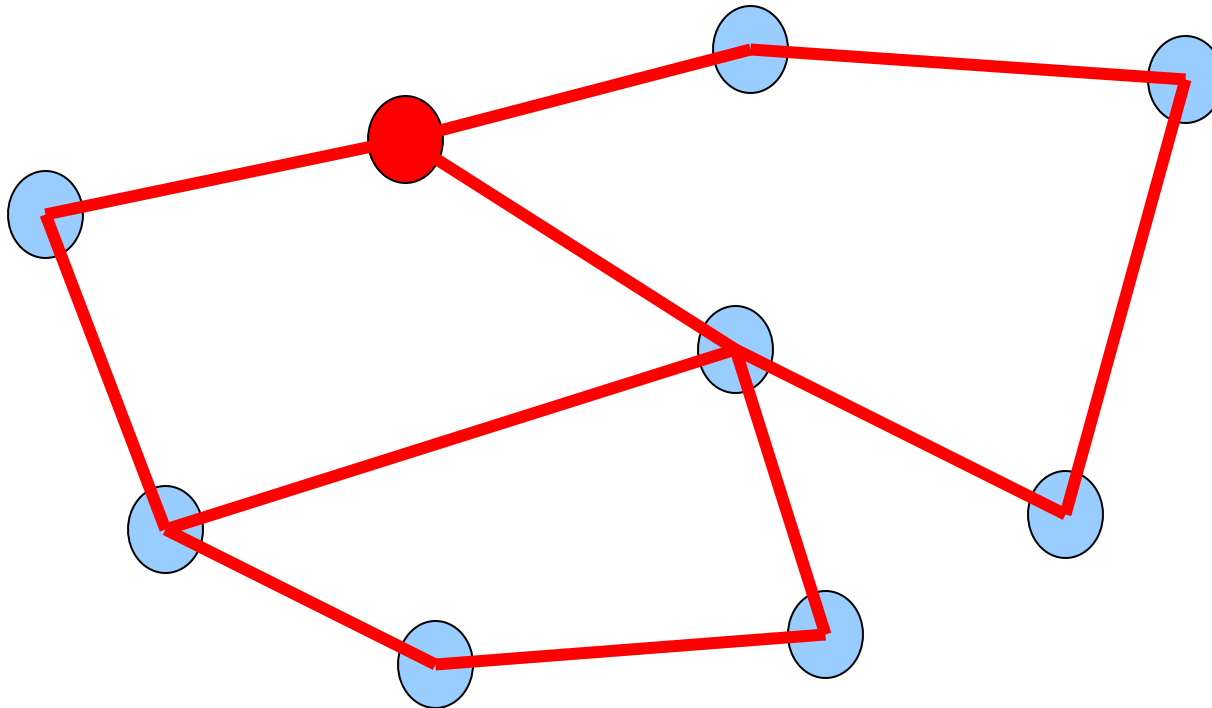
NS - Nonce Sum

# TCP rozšírenia

- zväčšenie okna (window)
- spresnenie merania RTT
- selektívne potvrdzovanie
- reakcia na zahltenie, okno cwnd

# “Pohl’ad vrstiev” na topológiu siete

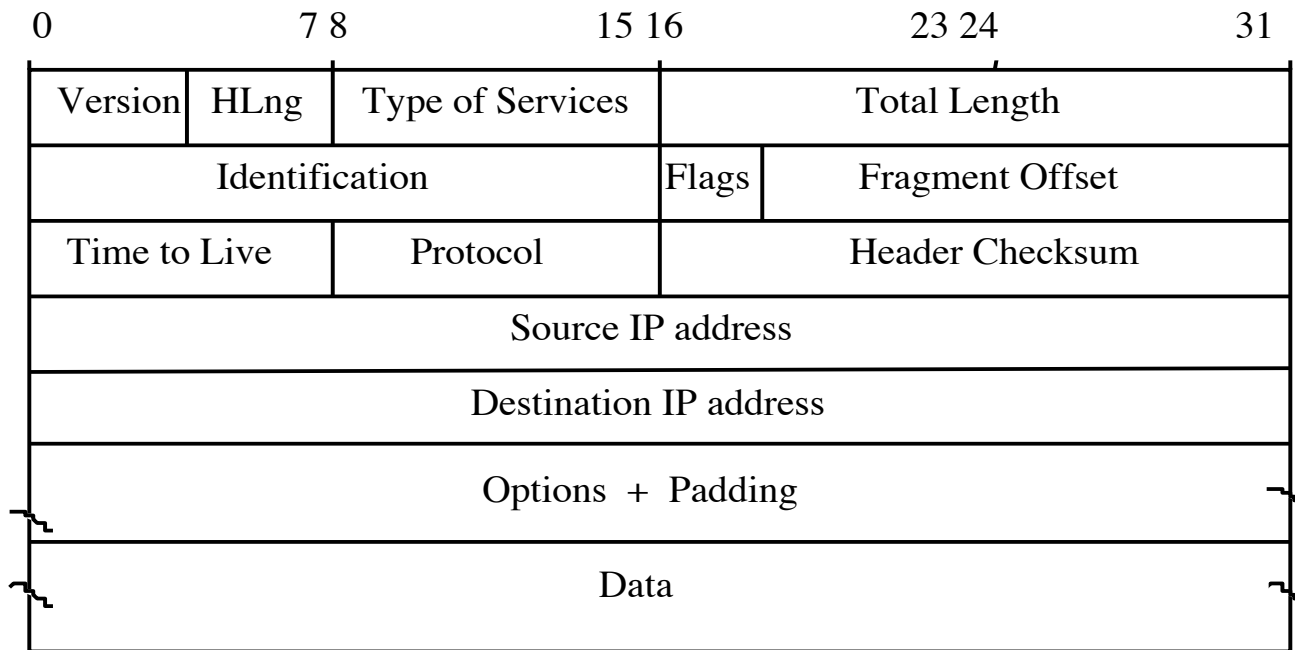
## sieťová vrstva





# Formát IPv4 paketu

## IP protokol, PDU - datagram



pole

Type of Services :

8

13 14 15

D	S	C	P	ECN
---	---	---	---	-----

pole


Flags :

16 17 18

0	DF	MF
---	----	----

# RFC 5735: Special Use IPv4 Addresses

Address Block	Present Use	Reference
0.0.0.0/8	"This" Network	RFC 1122
10.0.0.0/8	Private-Use Networks	RFC 1918
127.0.0.0/8	Loopback	RFC 1122
169.254.0.0/16	Link Local	RFC 3927
172.16.0.0/12	Private-Use Networks	RFC 1918
192.0.0.0/24	IETF Protocol Assignments	RFC 5736
192.0.2.0/24	TEST-NET-1	RFC 5737
192.88.99.0/24	6to4 Relay Anycast	RFC 3068
192.168.0.0/16	Private-Use Networks	RFC 1918
198.18.0.0/15	Network Interconnect Device Benchmark Testing	RFC 2544
198.51.100.0/24	TEST-NET-2	RFC 5737
203.0.113.0/24	TEST-NET-3	RFC 5737
224.0.0.0/4	Multicast	RFC 3171
240.0.0.0/4	Reserved for Future Use	RFC 1112
255.255.255.255/32	Limited Broadcast	RFC 919, RFC 922



192.168.1.100/16

1100 0000.1010 1000.0000 0001.0110 0100

1111 1111.1111 1111.0000 0000.0000 0000

Adresa siete:

1100 0000.1010 1000.0000 0000.0000.0000

192.168.0.0/16

0.0.1.100



Adresa počítača: 192.168.1.100/16

Adresa siete: 192.168.0.0

Broadcast: 192.168.255.255

Router: zvyčajne: 192.168.0.1



192.168.0.0/20

Druhá sieť

Adresa siete: 192.168.16.0

Broadcast: 192.168.31.255

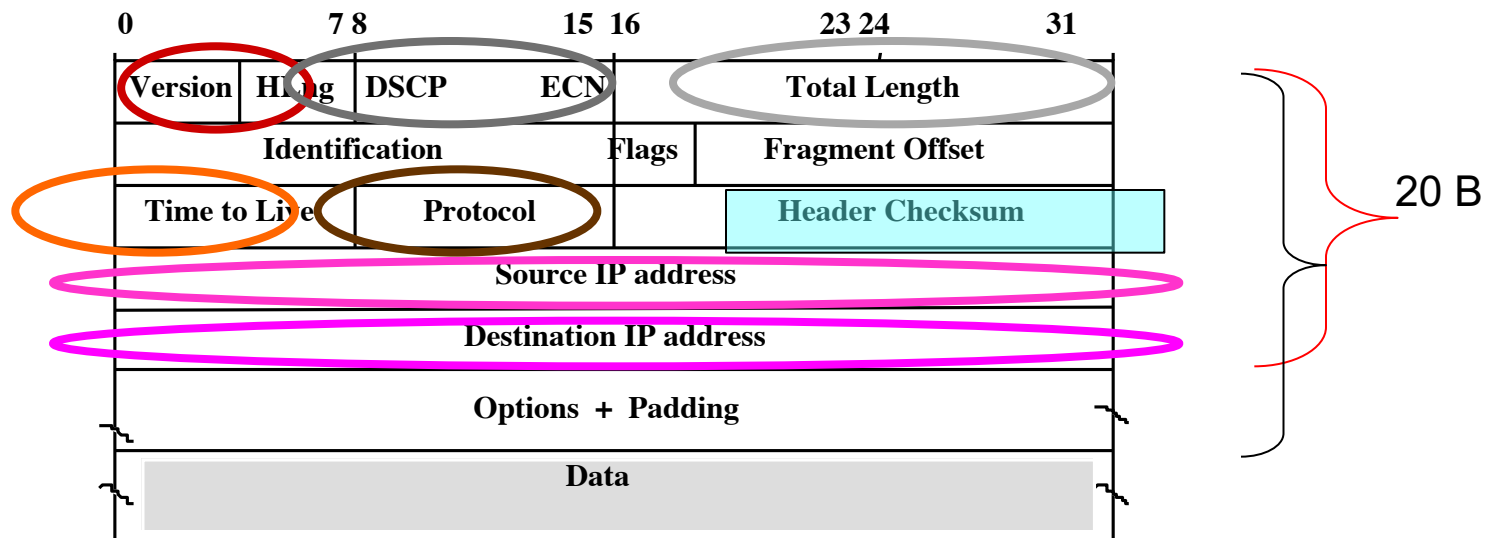
Príklad adresy počítača: 192.168.28.28

1100 0000.1010 1000.0001 0000.0000 0000

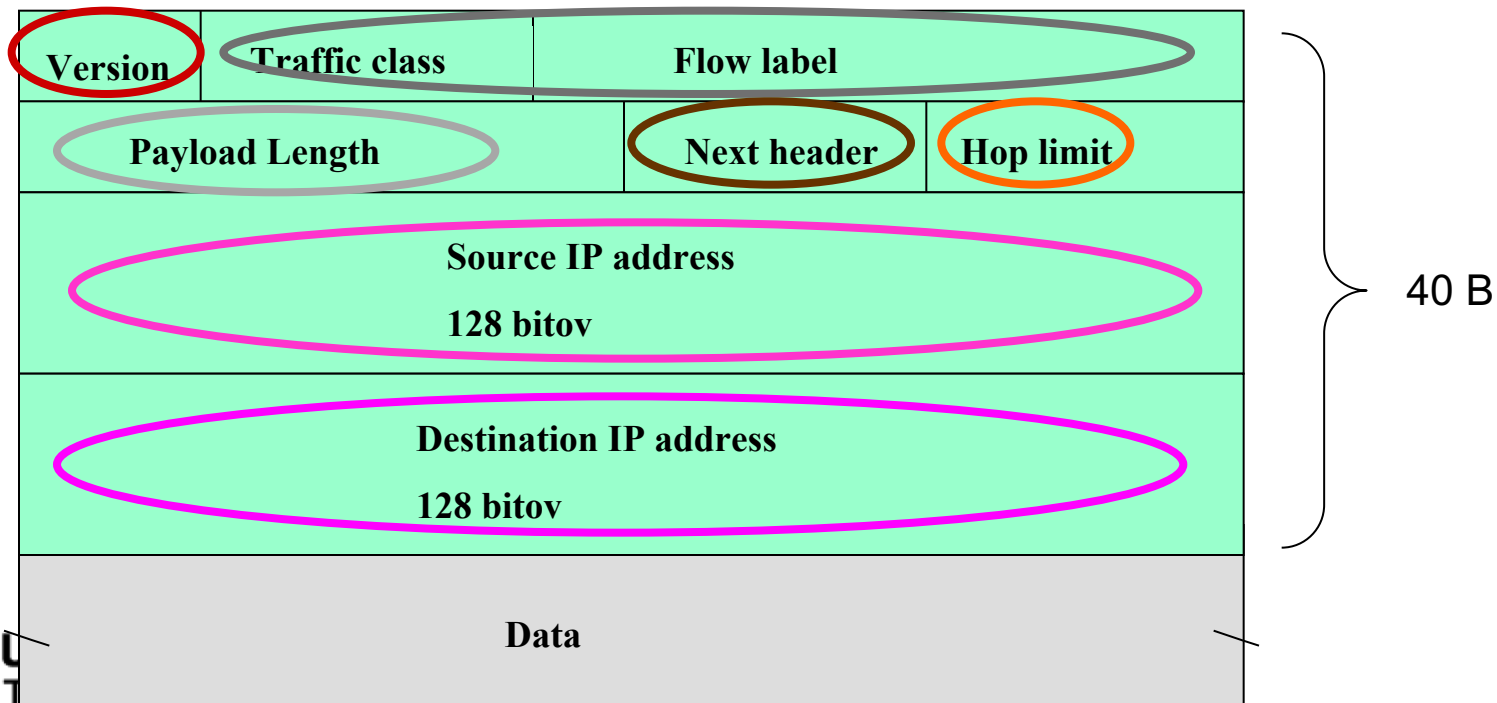
1100 0000.1010 1000.0001 1111.1111 1111

1111 1111.1111 1111.1111 0000.0000 0000

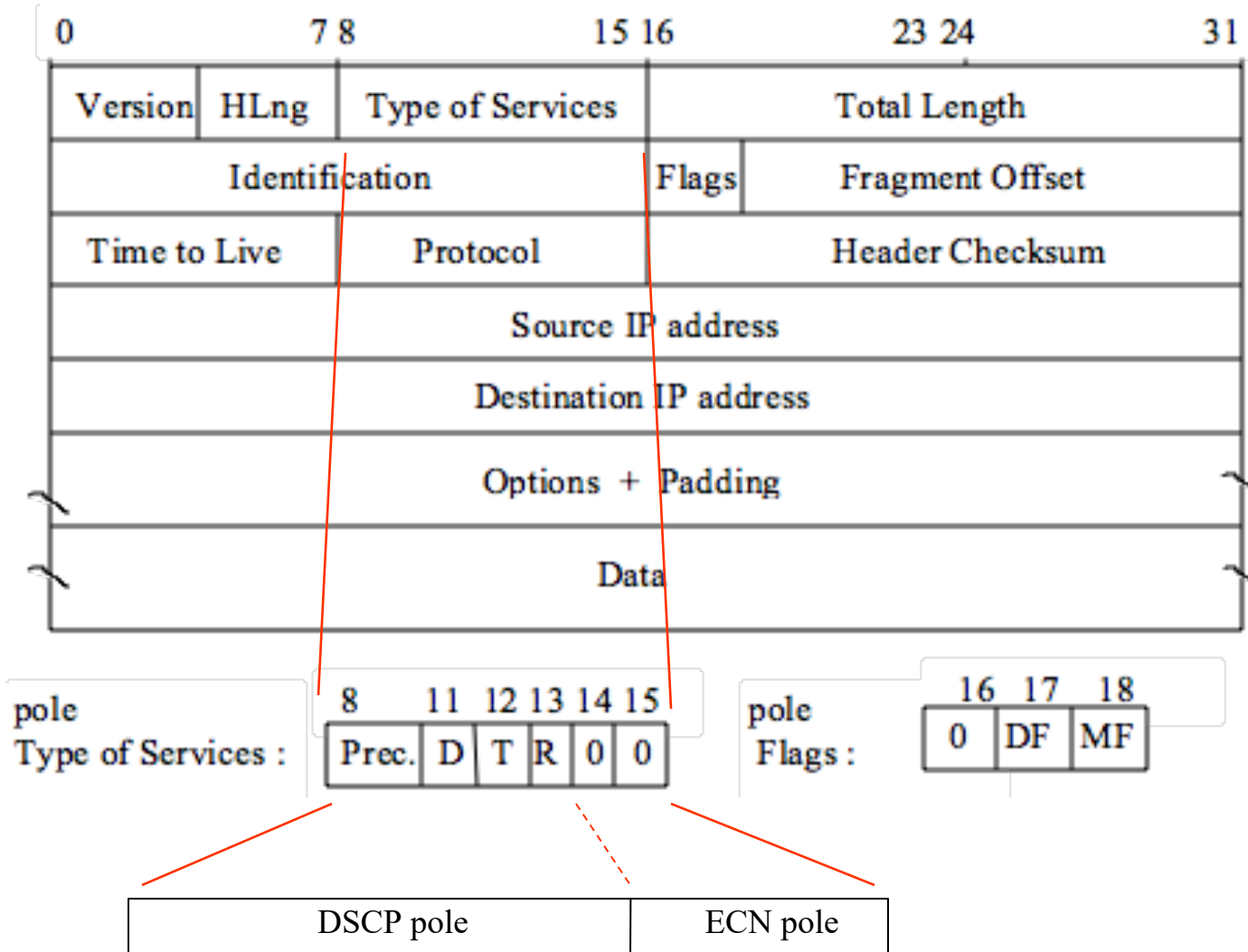
# IPv4



# IPv6



# IP protokol - ECN



DSCP - Differentiated Services CodePoint  
ECN - Explicit Congestion Notificaton

ECN pole:  
ECT - ECN Capable Transpor  
CE – Congestion Experienced

# Zhrnutie prednášky

- » TCP - dokončenie
- » IP protokol – formát paketu



# Čo nás čaká na budúcej prednáške

- » Subnetting (IP adresovanie)
- » Skutočné funkcie sieťovej vrstvy (smerovanie) budú neskôr

## DCCP (Datagram Congestion Control Protocol)

- nespoľahlivý tok datagramov
- spoľahlivé nadväzovanie a ukončenie spojenia
- spoľahlivé dohadovanie volieb, zahŕňa aj voľbu riadiaceho mechanizmu zahltenia
- spôsob prenosu potvrdení ACK závisí od mechanizmu riadenia zahltenia
- výber modulárneho mechanizmu riadenia zahltenia
  - CCID2, TCP-like congestion control (RFC 4341)
  - CCID3, TCP-Friendly Rate Control, TFRC (RFC 4342)
  - CCID4, TCP-Friendly Rate Control for Small Packets, TFRC-SP (RFC 4828)
- explicitné riadenia zahltenia (ECN)
- vhodné napr. pre streaming video

## Protokol TCP

- multiplexovanie a demultiplexovanie segmentov
- MSS (Maximum Segment Size)
- číslovanie bajtov
- potvrdzovanie príjmu, kontinuálna ARQ metóda s návratom, selektívna
- riadenie toku a zahltenia
- pohyblivé okno (def. 4096, rozšírenie okna na 32 b)
- časovače

# Stručný úvod do TCP

- » protokol so spojením, s potvrdením, spoľahlivý prenos
- » prenos dát – prúd bajtov, počet vyslaných bajtov aplikáciou a TCP entitou môže byť rôzny
- » vyrovnávacie pamäte – segmentácia prúdu bajtov
- » interaktívny a neinteraktívny prenos dát (typ aplikácie)
- » TCP spojenie – plný duplex, dvojbodové
- » urgentné dáta
- » príjem dát aplikáciou – príznak PUSH

# Riadenie toku dát a zahltenia /3

Príčiny zahltenia na rôznych vrstvách

## **spojová (dátová) vrstva**

- potvrdzovanie
- riadenie toku
- správa vyrovnávacích pamätí prijímača

.....

## **sieťová vrstva**

- správa vyrovnávacích pamätí
- zničenie paketov
- životnosť paketov
- smerovacie algoritmy

transportná
sieťová
spojová
fyzická

## **transportná vrstva**

potvrdzovanie

správa vyrovnávacích pamätí

určovanie host-host oneskorenia

(„timeout\_u“ cez sieť )