Univerza v Ljubljani

Fakulteta
za računalništvo
in informatiko

Omrežna plast

Vsebina

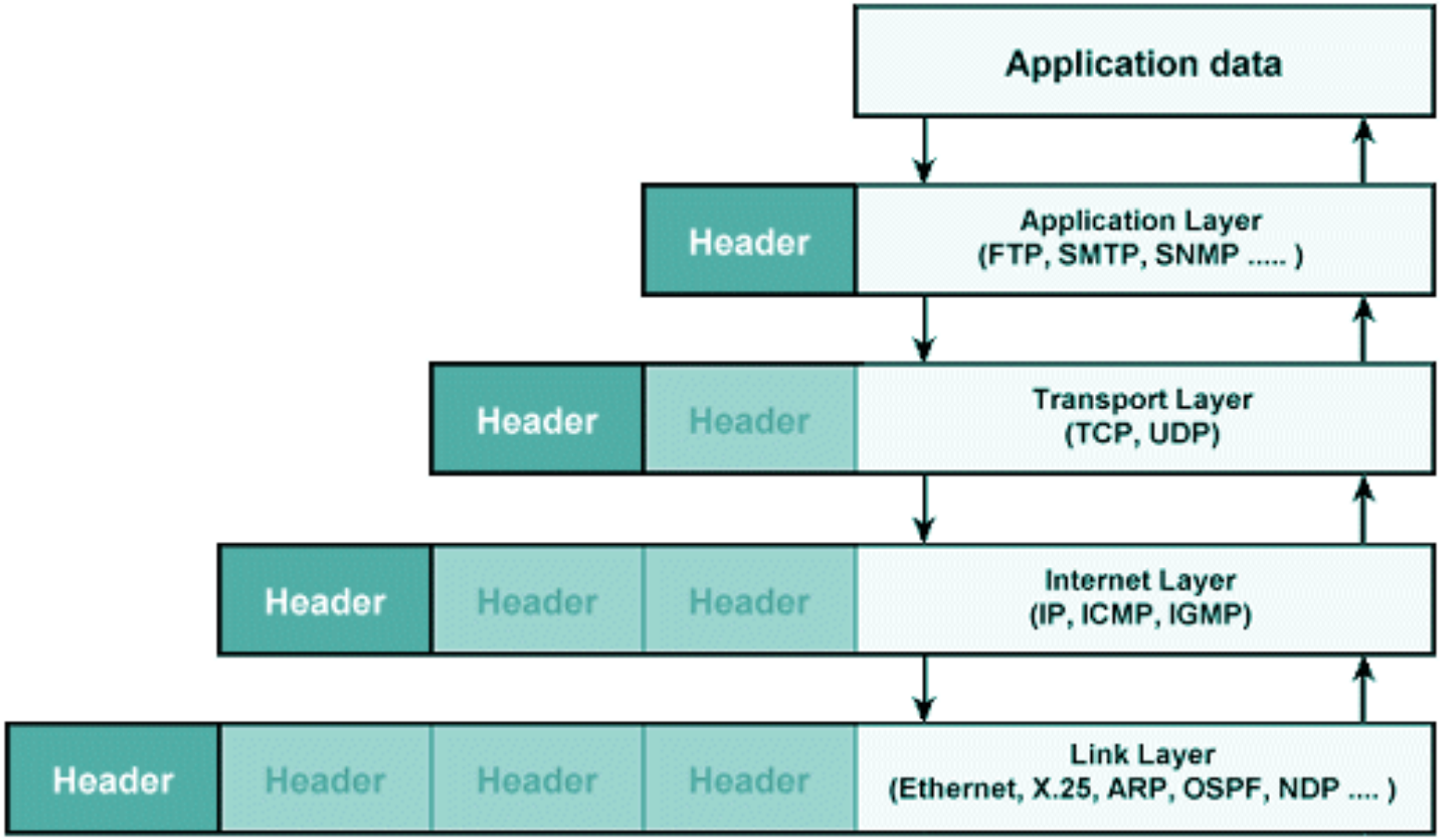
Delovanje storitev omrežne plasti

- Virtualne zveze in datagramske povezave
- Usmerjevalniki
- IP protokol: format, naslavljanje, ICMP, IPv6
- Usmerjevalni algoritmi
- Usmerjanje v Internetu, broadcast in multicast

Omrežna plast

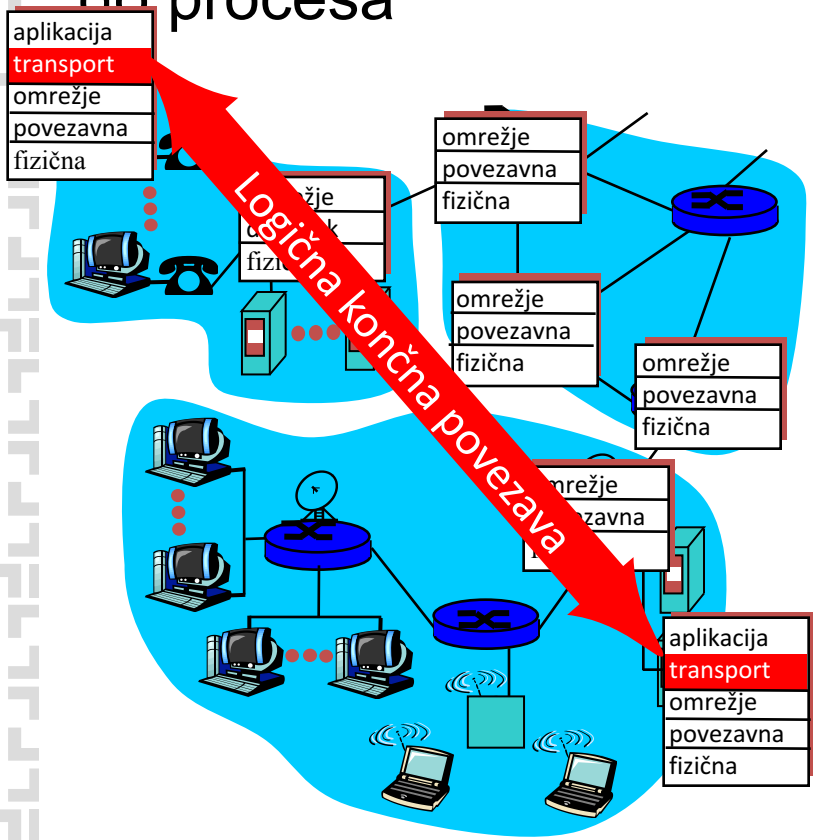
- Omrežni protokoli so v **vsakem** računalniku in usmerjevalniku!
- NALOGE
  - Prenos segmenta transportne plasti od izvornega do ciljnega računalnika.
    - Iskanje poti, naslavljanje, delo z datagrami, obvestila
  - Pošiljatelj: enkapsulacija segmentov **v IP datagrame**
  - Prejemnik: izluščenje in predaja segmentov transportni plasti

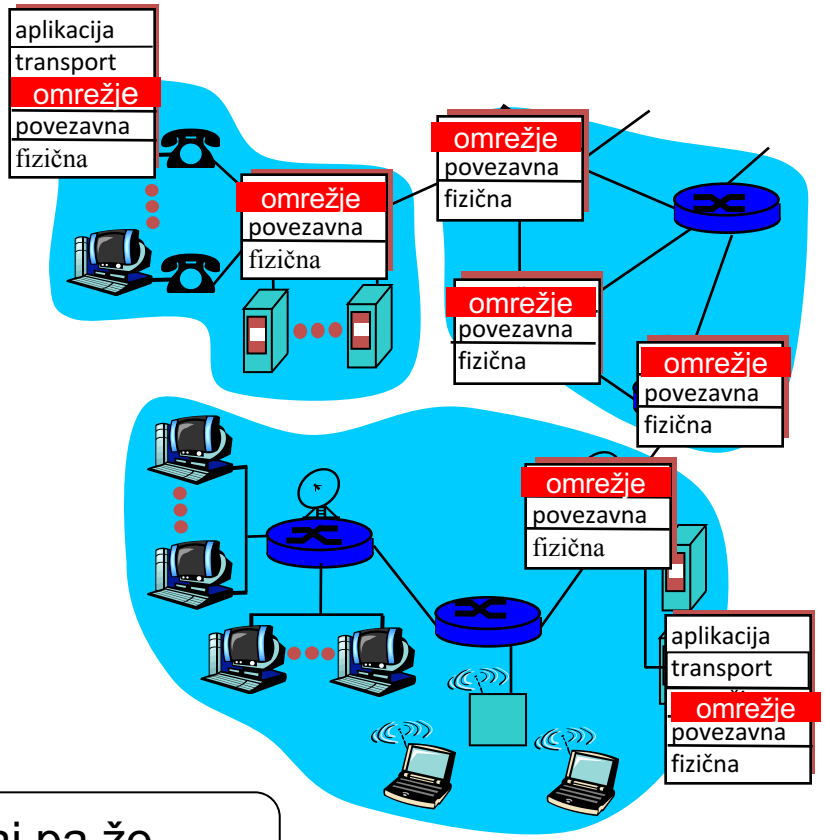
Enkapsulacija



Omrežna plast nudi storitve transportni plasti

- Transportna: od procesa do procesa
- Omrežna: od računalnika do računalnika

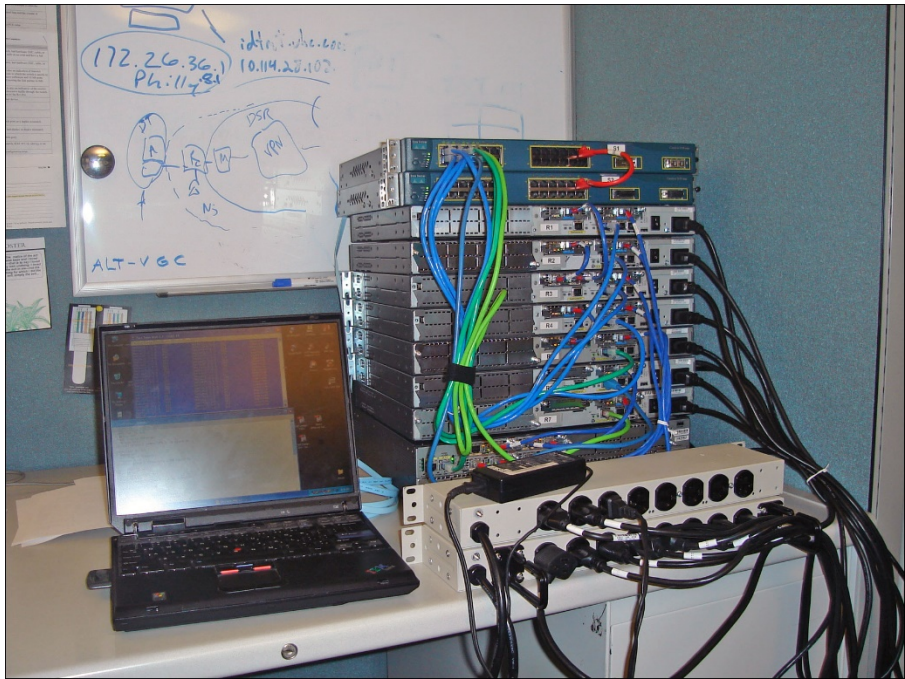




Hm, kaj pa že dela povezavna?

USMERJEVALNIK

- Naprava, ki deluje na omrežni plasti
- Posreduje datagrame iz enega v drugo omrežje
- Prenaša datagrame po hrbtenici omrežja
- Izvaja posredovanje in usmerjanje



Ključni funkciji omrežne plasti

- Posredovanje paketov (forwarding)
  - "Prenos" paketa iz vhoda v usmerjevalnik na ustrezno izhodno povezavo.
    Znotraj enega usmerjevalnika!
- Usmerjanje (routing)
  - Določitev in izvedba poti paketov od izvora do cilja.
    "Kolektivno delo" vseh naprav po pravilih usmerjevalnega protokola.
- Pogosto zamenjavanje teh dveh pojmov
  (npr. usmerjevalna tabela - posredovalna...)
- V nekaterih omrežjih je funkcija omrežne plasti tudi vzpostavljanje povezave (ATM, Frame Relay, X.25)

Model omrežnih storitev

Kaj omrežna plast lahko zagotovi transportni plasti?

- Zagotovljena dostava paketa
- Zgornje, z navzgor omejeno zakasnitvijo

Za posamezen paket

- Dostava paketov v pravem zaporedju
- Zagotovljena spodnja meja pasovne širine
- Čas med prejemom dveh paketov je le malo (navzgor omejeno) različen od časa med njuno oddajo – ***jitter***.

Za zaporedje paketov

Kaj od tega zagotavlja Internet?

best – effort, ni nobenih zagotovil ☹

Primer ATM: več modelov storitev

- Za različne povezave lahko vzamemo različne modele.

Omrežje	Model	Zagotavlja?				Zamašitev-obvestilo
		Pas. širina	Izguba	Vr. red	Čas	
Internet	best effort	ne	ne	ne	ne	Ne (izguba)
ATM	CBR	Konstant.	da	da	da	Ni zamašitev
ATM	ABR	minimalna	ne	da	ne	da (CI)

Available bit rate

Constant bit rate

IPv4 naslavljanje

- Vmesnik: povezuje računalnik ali usmerjevalnik s fizično linijo (network interface, omrežna kartica...).
- IPv4 naslov je 32-bitni **ID vmesnika.**

- Koliko vmesnikov ima navadno računalnik in koliko usmerjevalnik? Koliko pa IP naslovov?
- Koliko je možnih različnih IP naslovov?

Primer IPv4 naslova:

11011111 00000001 00000001 00000001

Desetiški zapis: 223.1.1.1

Podomrežje

- IP naslov vsebinsko pomeni dvoje: naslov omrežja (predpona) | naslov naprave znotraj tega omrežja
  - (analogija: hišne številke na ulici)
- (Pod)omrežje je množica vmesnikov,
  - ki imajo enak naslov (pod)omrežja,
  - med seboj so dosegljivi brez posredovanja usmerjevalnika.
- Maska podomrežja določa dolžino naslova (pod)omrežja.
  - Maska je 32-bitni niz, ki ima enice na mestih, ki označujejo naslov omrežja, na ostalih so ničle.
  - Npr. maska /25 pomeni, da je prvih 25 bitov naslov omrežja, zadnjih (desnih) 7 pa naslov naprave.
  - Primer: 11111111 11111111 11111111 10000000 ali 255.255.255.128

Primer - naslavljanje

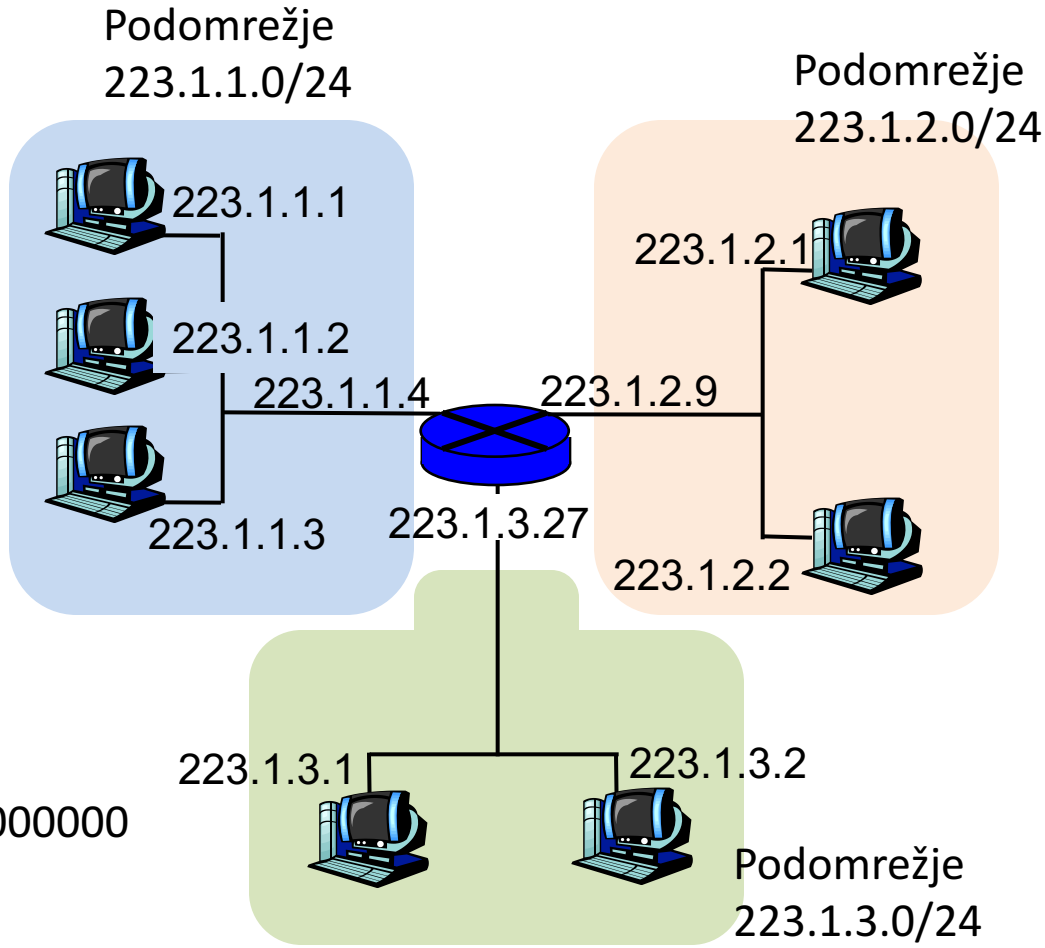
Usmerjevalnik ima na vsakem vmesniku drugo (pod)omrežje.

223.1.1.0/24 : levih 24 bitov (MSB) označuje naslov (prefix) omrežja, ostali naslov naprave.

Maska podomrežja /24:

111111111 11111111 11111111 00000000
oziroma
255.255.255.0

Maska je lahko poljubno dolga, npr. /17 ali /25...

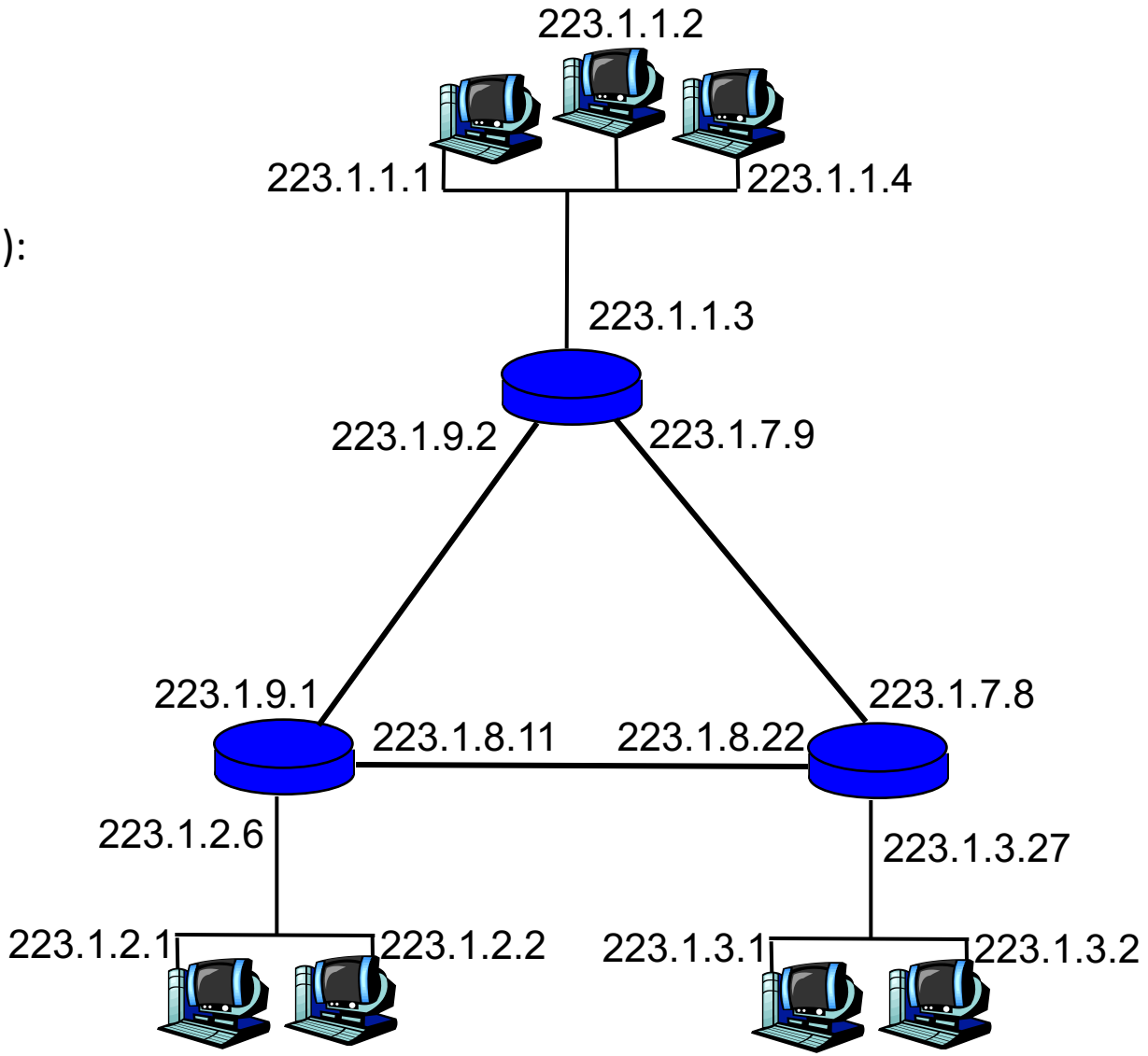


Znotraj (pod)omrežja ni usmerjevalnikov, so pa lahko stikala (switch) in razdelilniki (hub).

Koliko je podomrežij?

Prefiksna ali CIDR notacija (classless inter-domain routing): 223.1.1.0/24

Broadcast naslov: same enice. Velja za omrežje in napravo. Pošilja se vsem v omrežju, usmerjevalnik ga ne posreduje naprej. 233.1.1.255 255.255.255.255



Kako določati podomrežja?

- Nekoč so bili definirani razredi omrežij z masko 8, 16 ali 24 bitov.
  - Težava: prevelika ali premajhna podomrežja, neizkoriščenost naslovnega prostora.
  - Classful usmerjanje: maske ne potrebujemo.
    - razred A – prvi bit =0
    - Razred B – prva dva bita 10
    - Razred C – prvi 3 biti 110
- Kasneje: brezrazredno usmerjanje (classless) – CIDR ali prefiksna notacija, potrebujemo masko
- Broadcast naslov: razpošiljanje vsem adapterjem znotraj podomrežja (same enice v naslovu naprave).

Kako poteka dodeljevanje IP naslovov

- Naprava:
  - Administrator vpiše naslov (fiksen) ali
  - DHCP strežnik dodeli naslov (dinamičen) – admin prej strežniku dodeli ustrezen rang naslovov
- Omrežje podjetja:
  - Ponudnik dostopa do interneta (ISP) dodeli del svojega naslovnega prostora.

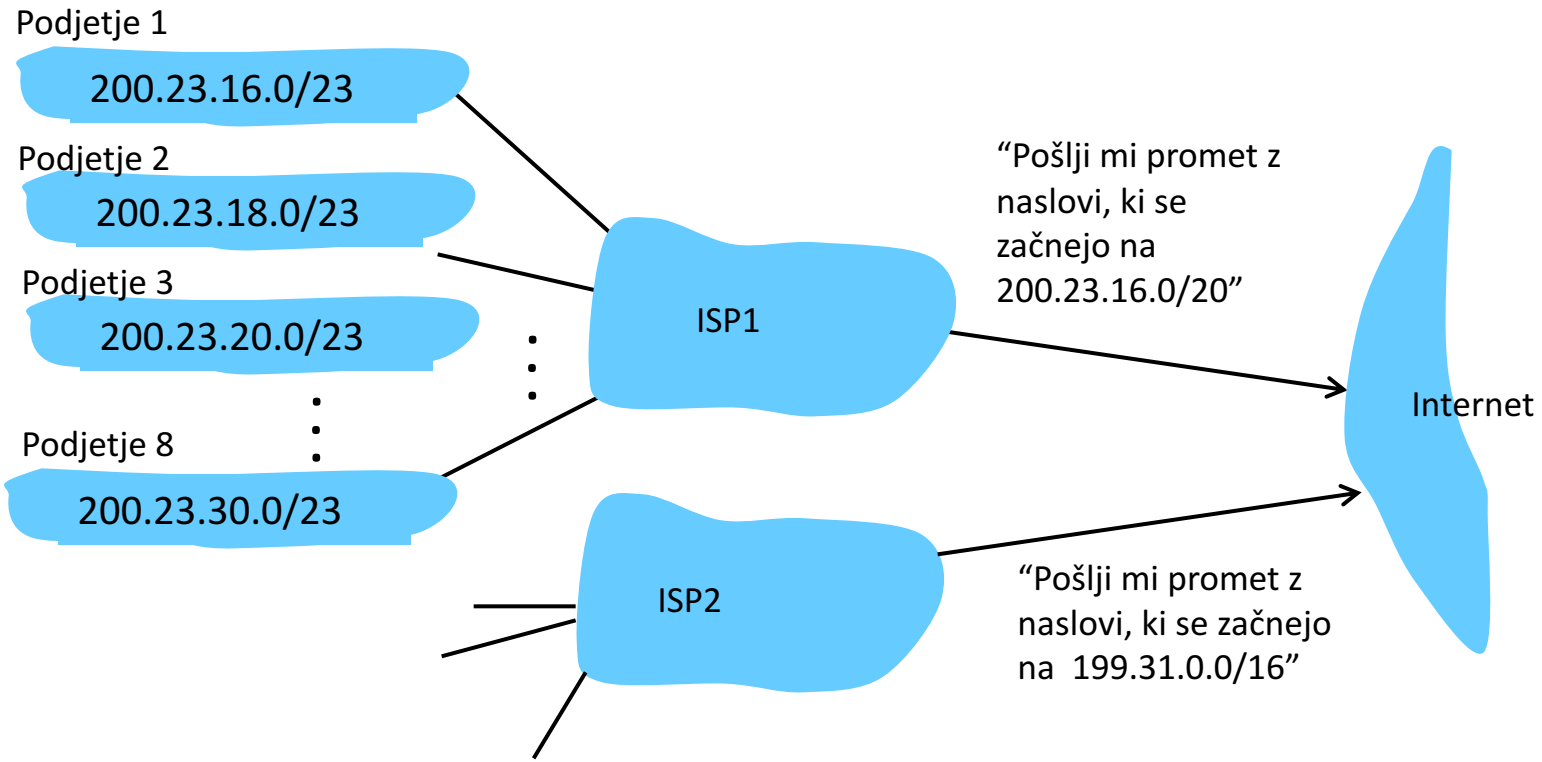
ISP-jev blok:	11001000 00010111 00010000 00000000	200.23.16.0/20
Podjetje1:	11001000 00010111 00010000 00000000	200.23.16.0/23
Podjetje2:	11001000 00010111 00010100 00000000	200.23.18.0/23
...	...	...

- ISP: ICANN dodeli naslovni prostor
  - Internet Corporation for Assigned Names and Numbers, www.icann.org
  - Oddelek IANA, Internet Assigned Numbers Authority, www.iana.org

Hierarhično naslavljanje

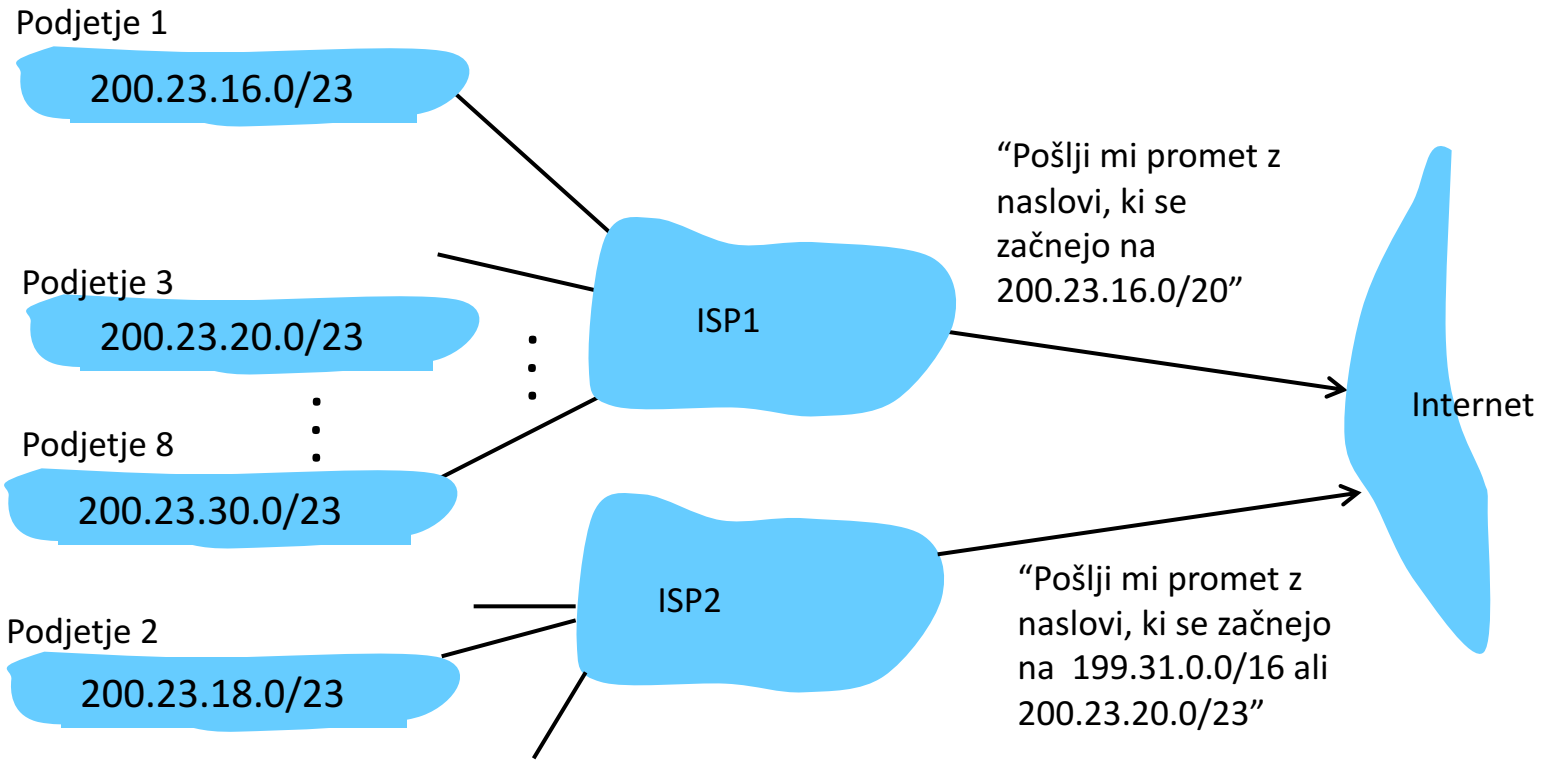
Pravilno dodeljevanje CIDR naslovov olajša usmerjanje!

Agregiranje ali sumarizacija naslovov – en prefiks za usmerjanje v več omrežij.

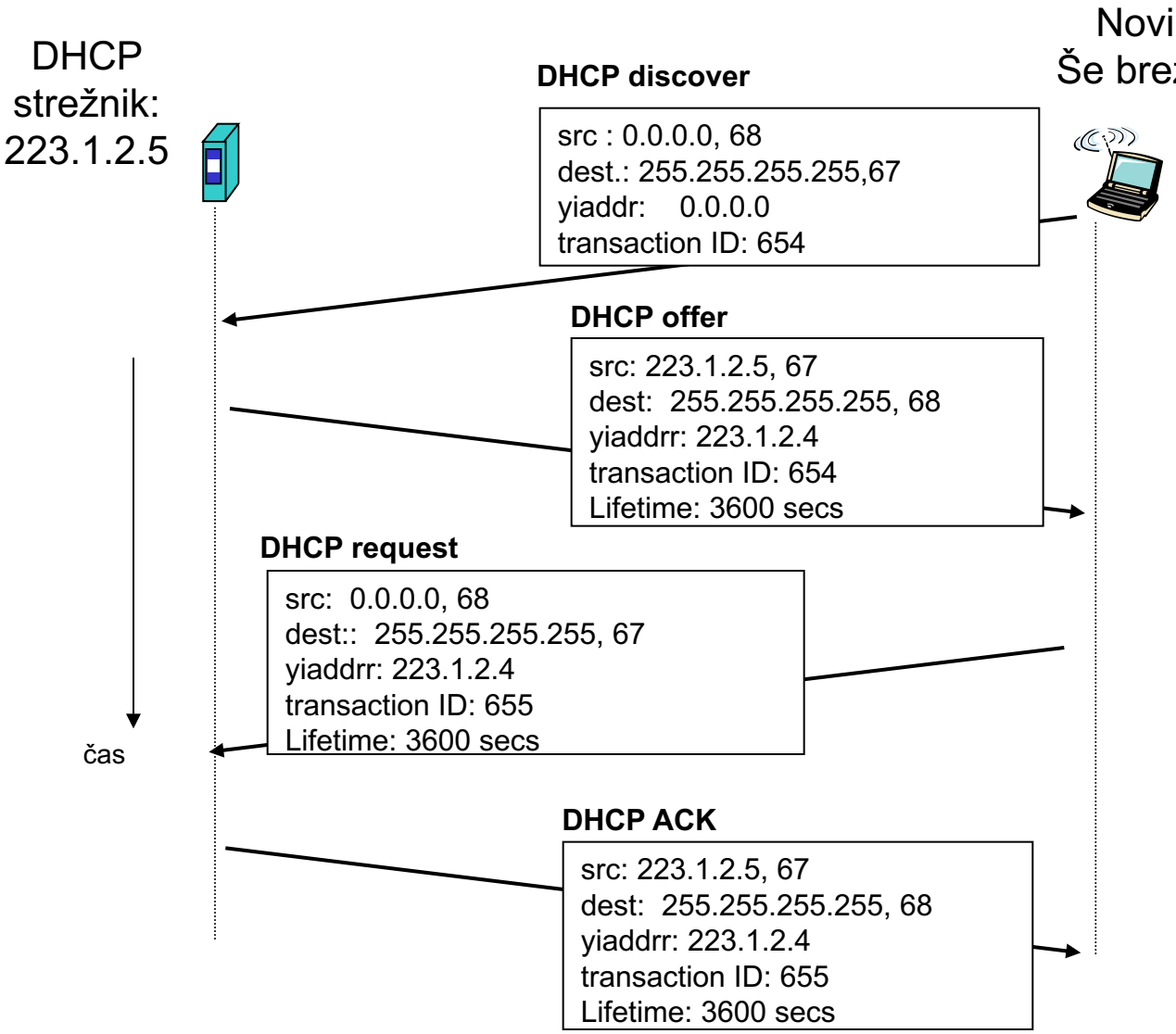


Manj učinkovito naslavljanje

ISP2 ima bolj specifičen naslov (daljši prefiks se ujema) za usmerjanje v Podjetje2. Usmerjevalne tabele so daljše.



Dodelitev naslova s pomočjo DHCP



Je tu kak DHCP strežnik?

Da, lahko bi ti dodelil tak naslov.

Lahko dobim ta naslov?

Potrditev.

Yiaddr = Your IP Address

DHCP strežnik pošlje tudi ostale omrežne nastavitve (privzeti prehod, DNS strežnik)

NAT – Network Address Translation (RFC 2663,3022)

- Motivacija: pomanjkanje IPv4 naslovnega prostora
- Zasebni naslovni prostor, RFC 1918

Naslovi	Omrežje/maska	Št. naslovov
10.0.0.0 - 10.255.255.255	10.0.0.0/8	2^{24}
172.16.0.0 - 172.31.255.255	172.16.0.0/12	2^{20}
192.168.0.0 - 192.168.255.255	192.168.0.0/16	2^{16}

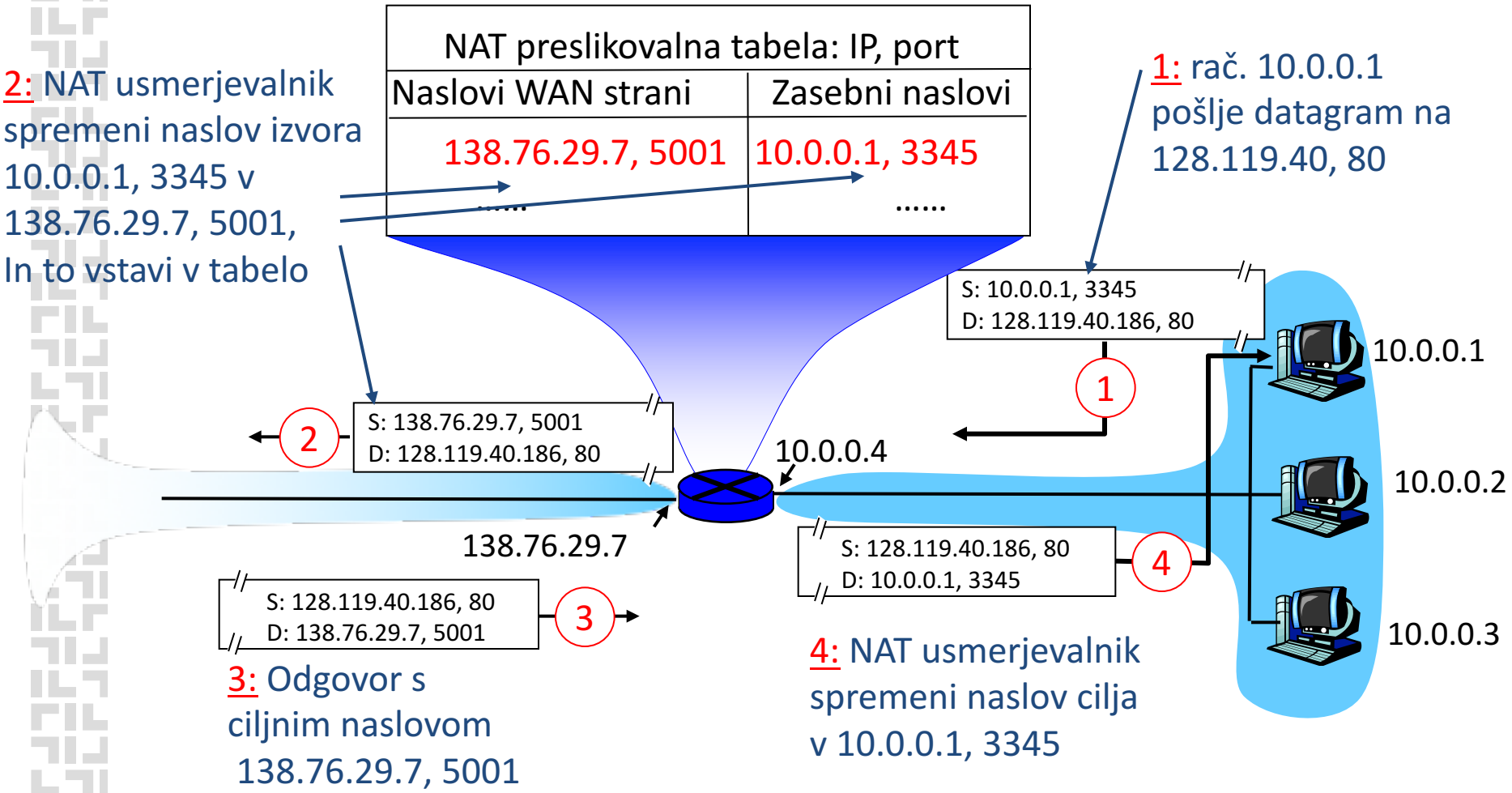
- Zasebni (notranji, interni) naslovi se uporabljajo le znotraj omrežja.
- Na NAT usmerjevalniku se naslov preslika v zunanji naslov.

NAT – Network Address Translation

- NAT usmerjevalnik in celo omrežje za njim navzven izgleda kot ena naprava.

- NAT usmerjevalnik:
  - Zamenja naslov izvora izhodnim datagramom
  - Zapomni si preslikavo (par notranji + zunanji naslov)
  - Zamenja naslov cilja vhodnim datagramom

NAT/PAT


NAT preslikovalna tabela: IP, port
Naslovi WAN strani
Zasebni naslovi
138.76.29.7, 5001
10.0.0.1, 3345
.....
.....
2: NAT usmerjevalnik spremeni naslov izvora 10.0.0.1, 3345 v 138.76.29.7, 5001, In to vstavi v tabelo
1: rač. 10.0.0.1 pošlje datagram na 128.119.40, 80
S: 10.0.0.1, 3345
D: 128.119.40.186, 80
1
2
S: 138.76.29.7, 5001
D: 128.119.40.186, 80
10.0.0.4
10.0.0.1
10.0.0.2
10.0.0.3
138.76.29.7
S: 128.119.40.186, 80
D: 10.0.0.1, 3345
4
S: 128.119.40.186, 80
D: 138.76.29.7, 5001
3
3: Odgovor s ciljnim naslovom 138.76.29.7, 5001
4: NAT usmerjevalnik spremeni naslov cilja v 10.0.0.1, 3345

Prednosti uporabe NAT

- Za celotno omrežje zadošča le en javni IP naslov
- V omrežju je laho preko 65000 naprav (port – št. vrat je 16 bitna številka)
- Notranje naprave niso neposredno dostopne od zunaj, zato so manj varnostno izpostavljene
- Naslove notranjih naprav lahko spreminjamo neodvisno od zunanjega naslova
- Laho zamenjamo ponudnika dostopa do interneta brez spreminjanja notranjih naslovov

Kritika NAT-a

- Usmerjevalniki – 3.plast: naj ne bi imeli opravka s 4. plastjo (vrata - porti)!!! Št.vrat je namenjena za naslavljanje procesov, ne računalnikov.
- Težava s strežniki na notranji strani (poslušajo na dogovorjenih vratih – well known port, NAT to številko zamenja).
- Pomanjkanje naslovov: raje uporabimo IPv6!
- Krši „end-to-end argument“ (za aplikacije naj bi bilo omrežje transparentno): npr. P2P načrtovalci morajo programirati tudi za primer NATa.
- Računalnik za NAT-om ne more sprejemati povezav, ker nima fiksnega naslova in ga ne more objaviti. Lahko le sam zahteva povezave (NAT traversal).

Rešitve za prehod čez NAT (NAT traversal)

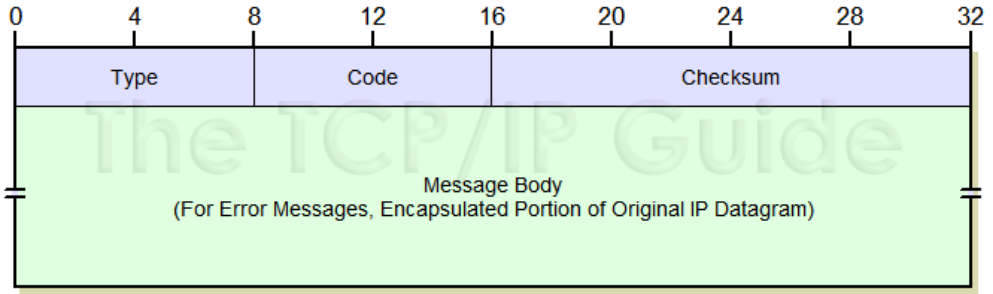
1. Statično konfiguriramo NAT – dodamo zapis v NAT tabelo (npr. 123.76.29.7, 2500 gre vedno v 10.0.0.1, 25000).
2. Universal Plug and Play (UPnP) Internet Gateway Device (IGD) protokol omogoča napravi za NATom ugotoviti zunanji IP naslov, in statične vnose v tabelo.
3. Prek posrednika: rač. za NAT-om ima stalno povezavo do posrednika, ki niza NAT-om. Sogovornik vzpostavi povezavo s posrednikom. Posrednik posreduje promet med njima, ali pa signalizira prvemu, da vzpostavi povezavo do drugega.
   - *Connection reversal*: Peer A se poveže z B prek C-ja, s katerim ima B trenutno aktivno povezavo, in ga prosi, naj B vzpostavi povezavo z A.

ICMP (RFC 792)

- Internet Control Message Protocol
- Sporočila v zvezi z omrežjem – napake, ...
- Pod-plast v omrežni plasti, leži rahlo nad IP (uporablja IP datagram za prenos ICMP sporočila, kot protokol višje plasti v glavi je naveden ICMP)
- Polja ICMP sporočila: tip, koda, glava in del IP datagrama, ki je povzročil napako (če je bila...)

ICMP sporočila

Tip	Koda	Pomen
0	0	echo reply (ping)
3	0	dest network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control – ni v uporabi)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header



Traceroute

- Po kateri pot igre promet do določenega IP-ja?
- Izvor pošilja serijo UDP paketov na (redek) port
  - Prvi: TTL=1, drugi: TTL=2, itd.
- Usmerjevalnik prejme datagram s TTL=0
  - Ga zavrže
  - Izvoru pošlje obvestilo – ICMP tip 11, koda 0
  - Obvestilo vključuje ime in IP usmerjevalnika
- Izvor izračuna čas vrnitve
- STOP: ko naslednji UDP paket doseže cilj, ali pa izvor dobi sporočilo “host unreachable” – tip3, koda 3.

Napadi na ICMP

- Ponarejen ICMP "Time exceeded" ali "Destination unreachable" povzroči, da takoj pade TCP povezava.
- Ping of Death – napad s fragmentacijo – pošljemo fragmentiran ping paket, daljši kot 65635 bytov (razlaga pri fragmentaciji!)
  - Obramba: kontrola odmikov in dolžin fragmentov (polje odmik: 13 bitov -> zadnji fragment z max. odmikom je lahko dolg max 7 bytov, sicer je datagram predolg).
- Smurf – napadalec pošlje ping s ponarejenim naslovom izvora na broadcast naslov v omrežju. Vsi odgovorijo napadenemu – DoS. Tako omrežje je smurf amplifier. Obramba:
  - Blokirati ping promet / broadcast promet
  - Usmerjevalniki (prehodi) ne spustijo v omrežje paketov na broadcast naslov.

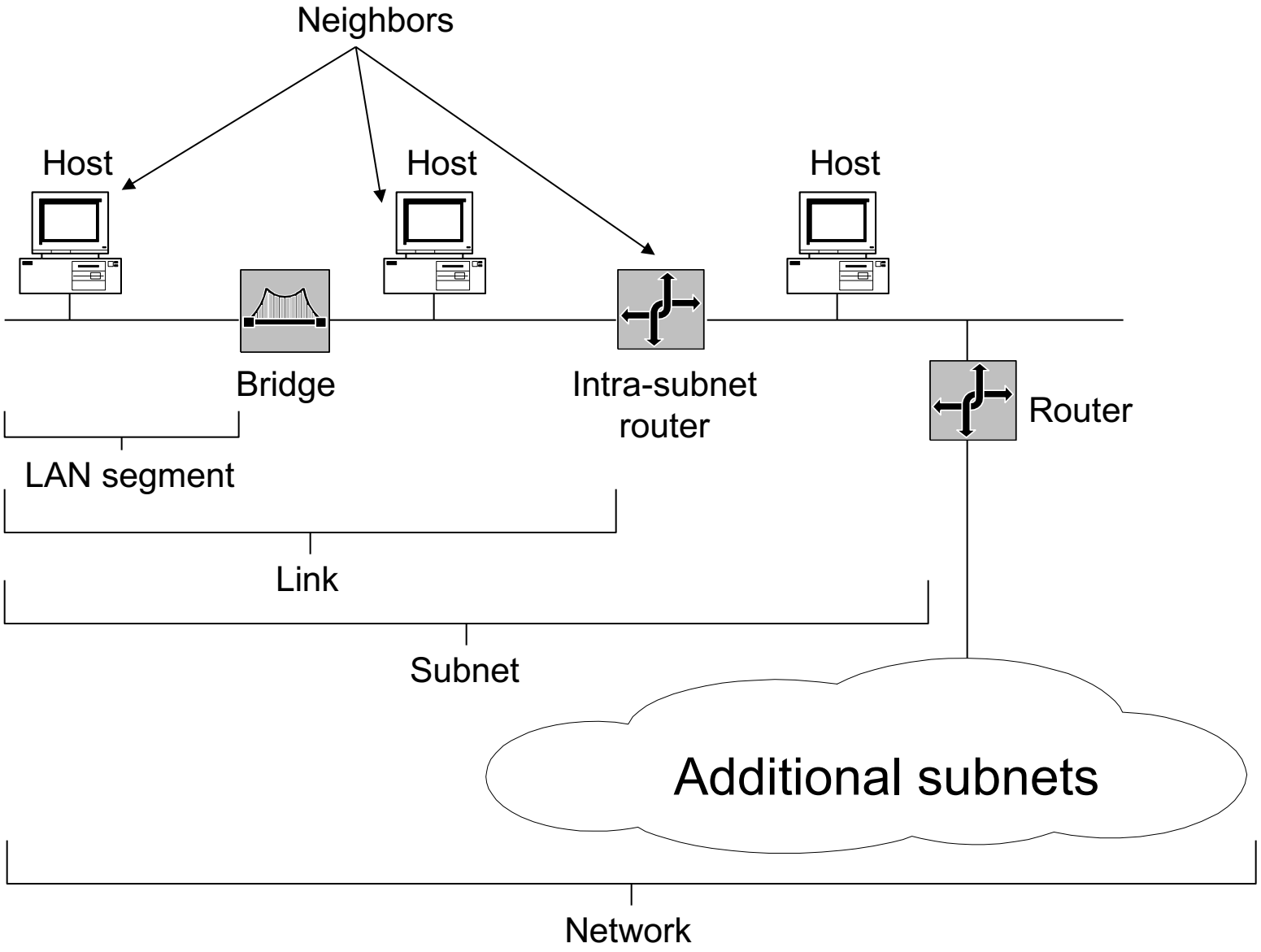
IPv6

- Motivacija:
  - večji naslovni prostor je potreben – 128 bitov
  - Format glave – hitrejše usmerjanje
  - Glava – omogoča QoS
- Ipv6 datagram:
  - Fiksna glava 40 bytov
  - Fragmentacija ni dovoljena

Prednosti IPv6

- Dovolj velik naslovni prostor
- Mednarodno uravnoteženje
- End-to-end komunikacija (P2P)
- Strukturirano izbiranje naslovov
- Razširljivost
- Hitro usmerjanje in posredovanje
- Vgrajeno: varnost in mobilnost, QoS

IP v6 terminologija



Naslovni prostor IPv6

- 128-bitni naslovni prostor
  - 340,282,366,920,938,463,463,374,607,431,768,211,456 naslovov (3.4×10^{38})
  - 6.65×10^{23} naslovov na m^2 zemljine površine !!!
- Zato imamo lahko fleksibilno večnivojsko hierarhijo (naslavljanje, usmerjanje)
- Tipičen unicast naslov:
  - 64 bitov: ID podomrežja
  - 64 bitov: ID vmesnika

Sintaksa IPv6 naslova

- IPv6 naslov v binarni obliki :

```
0010000111011010000000001101001100000000000000000010111100111011
0000001010101010000000001111111111111110001010001001110001011010
```

- Razdeljen na osem 16-bitnih skupin:

```
0010000111011010   0000000011010011   0000000000000000   0010111100111011
0000001010101010   0000000011111111   1111111000101000   1001110001011010
```

- Zapisan šestnajstiško, ločeno z dvopičji

  21DA:00D3:0000:2F3B:02AA:00FF:FE28:9C5A

- Vodilne ničle v vsaki skupini lahko izpustimo:

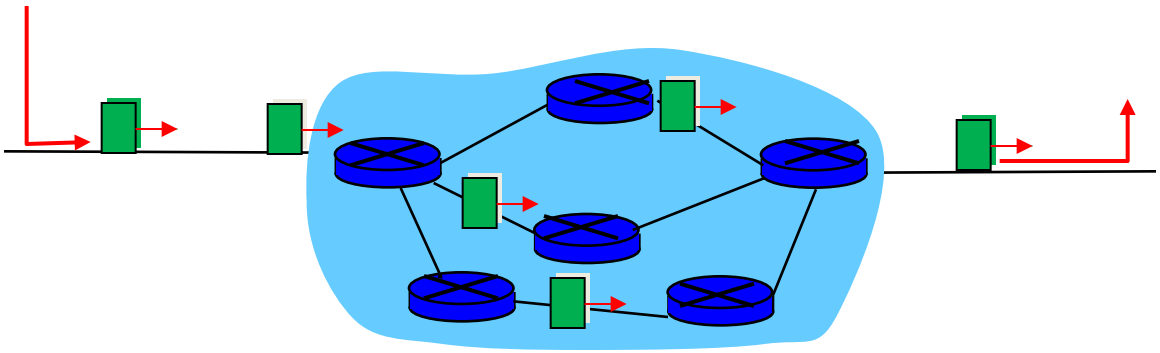
  21DA:D3:0:2F3B:2AA:FF:FE28:9C5A

Kompresija ničel v zapisu naslova

- Dolga zaporedja samih ničel
- Zaporedje 16-bitnih blokov iz samih ničel lahko zapišemo kot dve dvopičji ::
- Primer
  - FE80:0:0:0:2AA:FF:FE9A:4CA2 ali krajše FE80::2AA:FF:FE9A:4CA2
  - FF02:0:0:0:0:0:0:2 ali krajše FF02::2
- To ne velja za dele blokov – cel blok mora biti 0
  - FF02:30:0:0:0:0:0:5 ni isto kot FF02:3::5,
  - lahko pa zapišemo FF02:30::5.
- Kompatibilnost z v4 naslovi: spredaj dodamo ničle
  - 193.2.72.1 → ::193.2.72.1
  - Lahko pustimo tudi pike iz v4 naslova!

Datagramsko omrežje

- Na omrežni plasti ni vzpostavljanja klica.
- Usmerjevalniki ne vedo nič o končnih povezavah.
- Paket se posreduje glede na naslov cilja.
- Med istim izvorom in ciljem lahko po več poteh.



Posredovalna tabela v datagramskem omrežju

- Če je 32-bitni naslov: 4 mrd. različnih naslovov!
- V tabeli uporabimo rang naslovov, npr:

Ciljni naslov	Vmesnik povezave
Od 11001000 00010111 00010000 00000000 Do 11001000 00010111 00010111 11111111	0
Od 11001000 00010111 00011000 00000000 Do 11001000 00010111 00011000 11111111	1
Od 11001000 00010111 00011001 00000000 Do 11001000 00010111 00011111 11111111	2
sicer	3

Posredovalna tabela v datagramskem omrežju

- Če je 32-bitni naslov: 4 mrd. različnih naslovov!
- V tabeli uporabimo rang naslovov, npr:

Ciljni naslov	Vmesnik povezave
Od 11001000 00010111 00010000 00000000 Do 11001000 00010111 00010111 11111111	0
Od 11001000 00010111 00011000 00000000 Do 11001000 00010111 00011000 11111111	1
Od 11001000 00010111 00011001 00000000 Do 11001000 00010111 00011111 11111111	2
sicer	3

- Ujemanje najdaljše predpone (longest prefix match)

Primer: ujemanje najdaljše predpone

- Na kateri vmesnik posredovati
  - 11001000 00010111 00010110 10100001 ?
  - 11001000 00010111 00011000 10101010 ?

Ciljni naslov	Vmesnik povezave
11001000 00010111 00010	0
11001000 00010111 00011000	1
11001000 00010111 00011	2
sicer	3

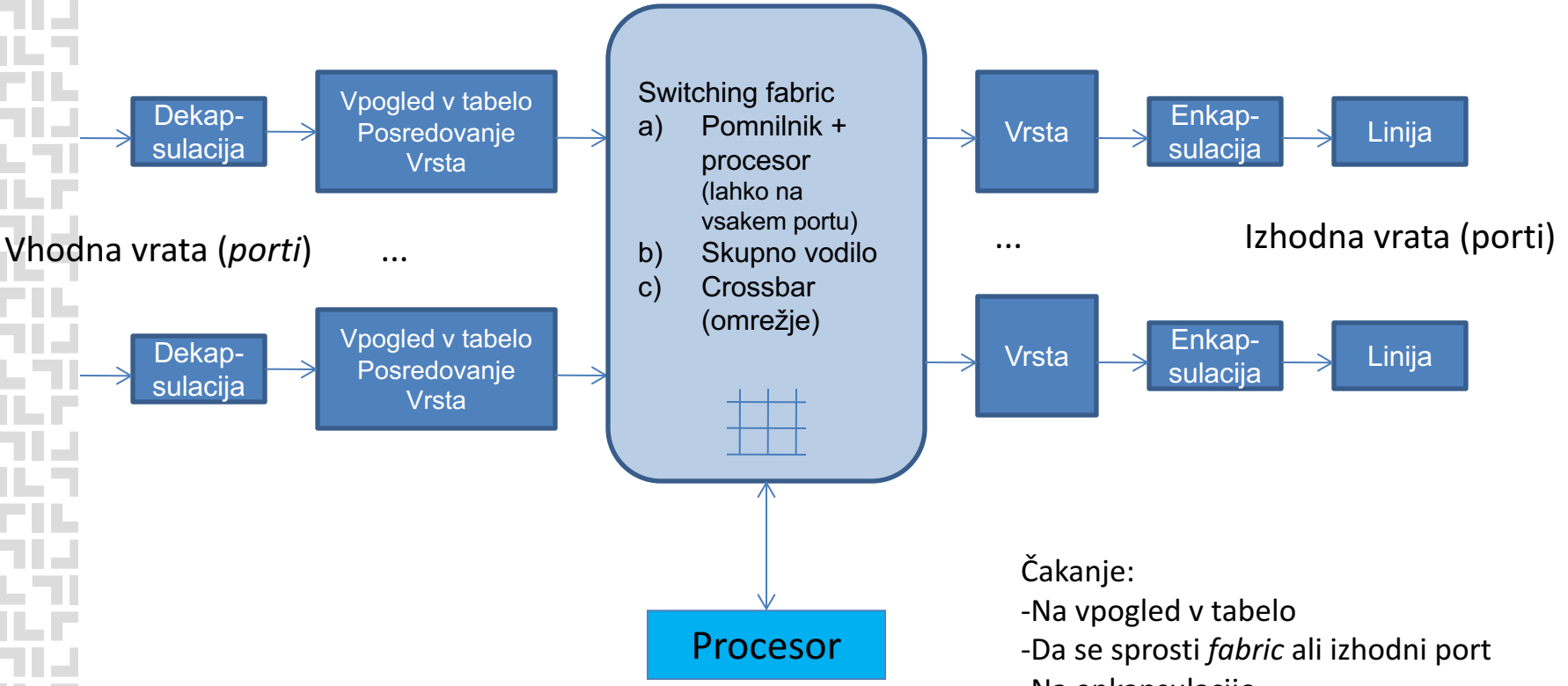
Učinkovitost: če so veliki bloki zaporednih naslovov za posamezno izhodno povezavo (krajši prefiksi, krajše tabele)

Kaj dela usmerjevalnik?

- Izvaja usmerjevalni protokol (npr. OSPF, BGP, RIP, EIGRP...)
- Posreduje datagrame iz vhodnih na izhodne povezave.
- Vhod:
  - Sprejem na nivoju bitov
  - Povezavna plast – dekapsulacija
  - Omrežna plast, ciljni IP naslov, vpogled v posredovalno tabelo, po potrebi čakanje v vrsti za prenos na izhod
  - Težave, če je jedro počasnejše kot kombinacija vseh vhodov! Vrste, zakasnitve, izgube paketov.
  - HOL (Head of the Line) blokiranje – prvi datagram v vrsti blokira tiste za njim, ki bi sicer lahko napredovali prek jedra.

Usmerjevalnik

- Težava: posredovanje je potrebno izvajati s hitrostjo vhodne povezave.



Čakanje:
-Na vpogled v tabelo
-Da se sprosti *fabric* ali izhodni port
-Na enkapsulacijo
-Kaj če vsi vh. porti pošiljajo na istega izh.?

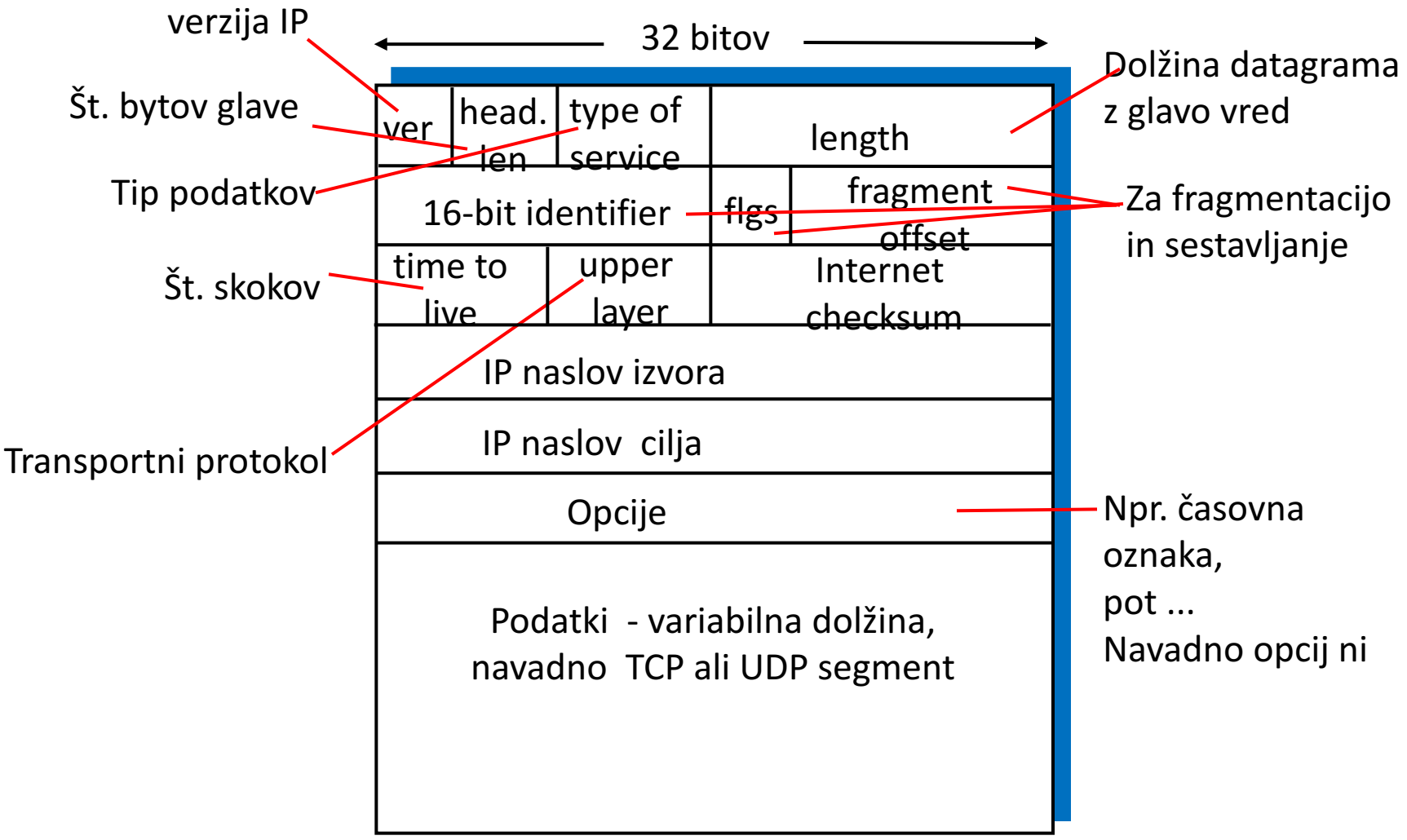
Preklapljanje med vhodnimi in izhodnimi vmesniki (porti)

- Preko pomnilnika:
  - Preklapljanje je pod nadzorom CPU
  - Datagram se prebere v sistemski pomnilnik, nato CPU naredi vpogled v posredovalno tabelo
  - Datagram se skopira na ustrezni izhod
  - Za vsak datagram sta potrebna 2 prehoda po vodilu
  - Največja prepustnost?
- Preko vodila:
  - Datagram se prebere iz pomnilnika na vhodu direktno v pomnilnik na izhodu
  - Za vsak datagram le en prehod po vodilu
  - Največja prepustnost?

Kaj se dogaja na izhodu?

- Jedro je hitrejše kot izhod: potrebna je vrsta.
- Čakanje – zakasnitve
- Datagrami se lahko izgubijo (zamašitev)
- Kaj če je intenzivnost prihajanja višja kot intenzivnost odhajanja?
- Razvrščanje datagramov v vrsti za izhod (prioritete – kdo dobi najboljše performanse)
- „Output Port Buffer Overflow“ – zakasnitve in izgube
- Koliko prostora v izhodnem bufferju?
  - (RFC 3439) : RTT (npr. 150 ms) * hitrost povezave
  - Če je N tokov: RTT*C / sqrt(N)

IP – internet protokol: format IPv4 datagrama



Režija

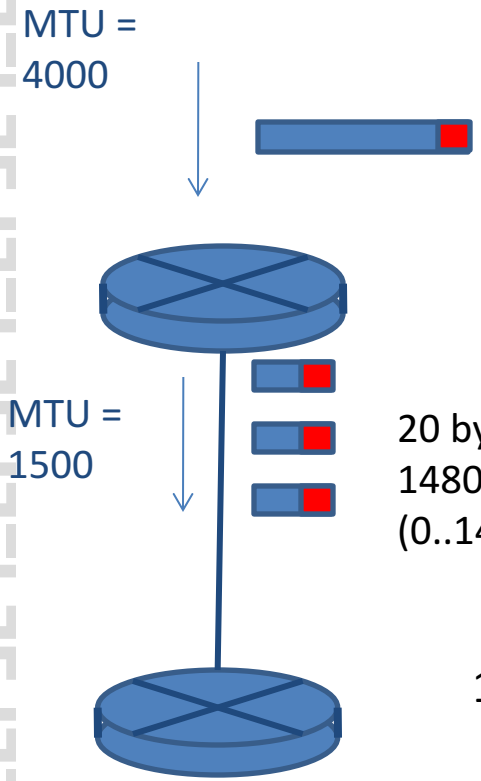
- 20 bytov: TCP glava
- Plus 20 bytov IP glava
- Plus režija aplikacijske plasti
- Delež režije v paketu je odvisen od dolžine podatkovnega dela!

Fragmentacija

- Povezavna plast: omejena dolžina okvirja (MTU), odvisna od tehnologije.
- V omrežju je lahko več tehnologij, zato se “med potjo” spreminja MTU!
- Fragmentacija: velik IP datagram z vhoda se razbije na več manjših IP datagramov-fragmentov.
- Fragmentira lahko usmerjevalnik sredi poti.
- Nazaj sestavlja vedno šele omrežna plast na cilju, pred predajo transportni plasti.

Fragmentacija: primer

- Datagram 4000 bytov, MTU 1500 bytov



	length =4000	ID =x	fragflag =0	offset =0	

Sledijo še fragmenti

Enota za odmik je 8 **bytov**!

20 bytov glave in
1480 bytov podatkov
(0..1479)

	length =1500	ID =x	fragflag =1	offset =0	

185 =1480/8

	length =1500	ID =x	fragflag =1	offset =185	

	length =1040	ID =x	fragflag =0	offset =370	

zadnji

Dolžina telesa fragmentov (razen zadnjega) je večkratnik od 8 v bytih!

Napadi na fragmentacijo

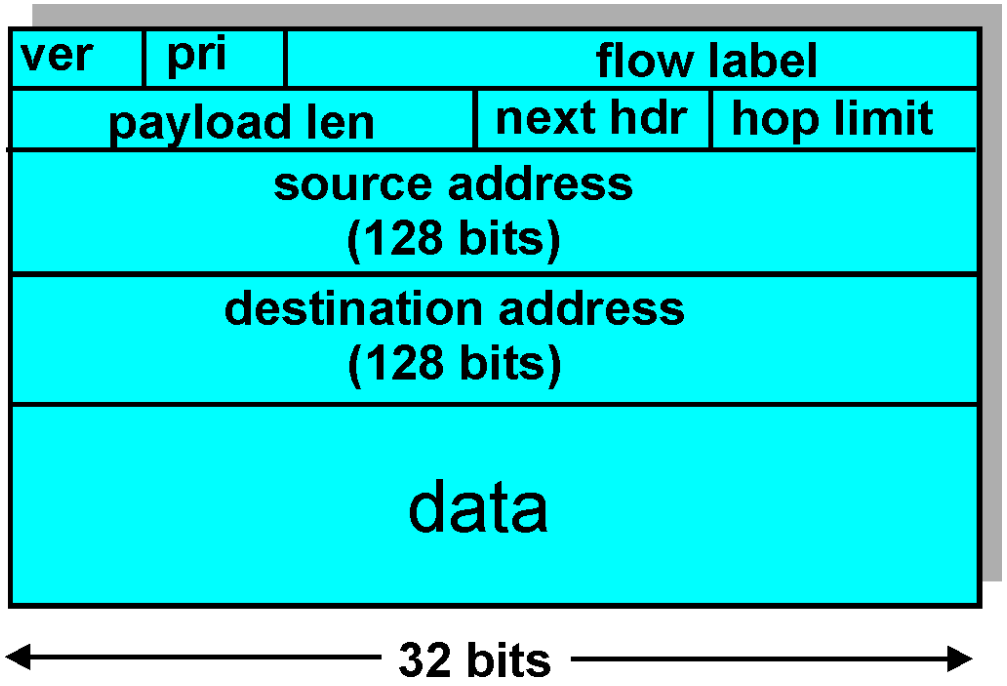
- Klasika: TEARDROP napad - spada med DoS napade
  - Napadalec: fragmentirani paketi z namerno napačnimi odmiki/dolžinami (prekrivanje) – *„fragment overlapped“*
  - Pri sestavljanju se ciljni sistem zmede in (lahko) sesuje – napaka v kodi TCP/IP sklada!
  - Občutljivi: Win 3.1, Win95, WinNT, Linux do 2.1.63
- Ping of Death – pošljemo fragmentiran ping paket, daljši kot 65635 bytov (že omenjeno pri ICMP napadih)
  - Obramba: kontrola odmikov in dolžin fragmentov (polje odmik: 13 bitov -> zadnji fragment z max. odmikom je lahko dolg max 7 bytov, sicer je datagram predolg).

Napadi na fragmentacijo

- Fragment overlapped – prekrivanje
  - Sistem se zmede, lahko crash (DoS)
  - Fragment se (delno) prepiše – želimo zaobiti IDS, da ne prepozna napada (če IDS ne defragmentira) (Fragment overwrite)
- Fragmentation Buffer Full - posledica (želimo zaobiti IDS)
  - Veliko datagramov z manjkajočimi fragmenti (Incomplete Datagram)
  - Posamezni datagrami z veliko/velikimi fragmenti
- Fragment Overrun – sestavljen datagram je večji kot dovoljuje dolžina polja length v glavi. (crash – DoS napad)
- Fragment Too Small – ne-zadnji fragment krajši od 400 bytov (zaobiti IDS ali druge filtre)

IPv6 format datagrama

Pri – prioriteta med datagrami (razred prometa)
Flow label – omogoča identificirati datagrame, ki pripadajo istemu toku (npr. video)
Next header – protokol višje plasti ali lokacija razširitve glave
Hop limit = TTL



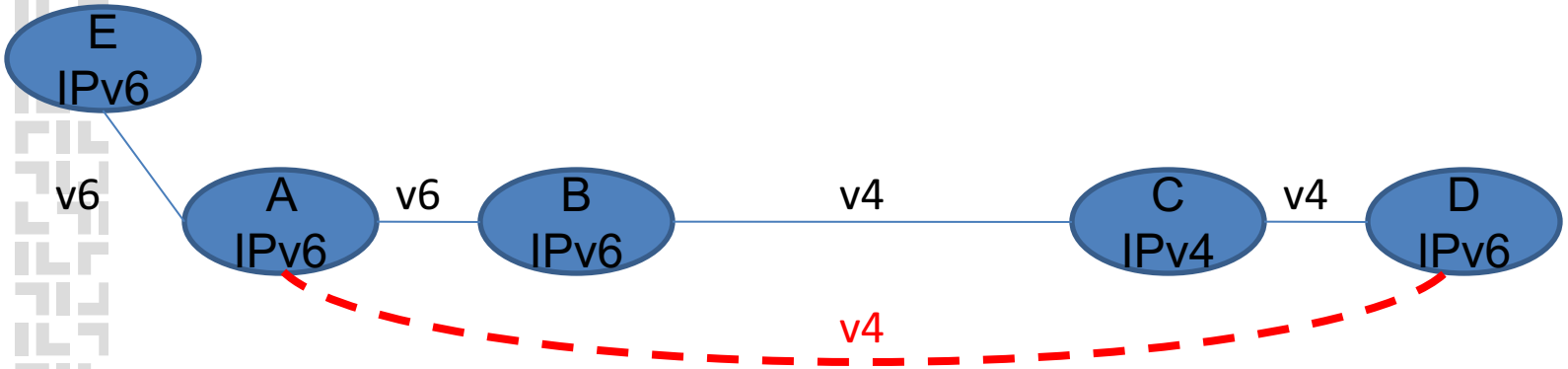
Ni polj fragmentacije, kontrolne vsote, opcij (le kot razširitev glave).

ICMP v6 – dodatne funkcije, npr. sporočilo Packet Too Big.

Prehod IPv4 – IPv6

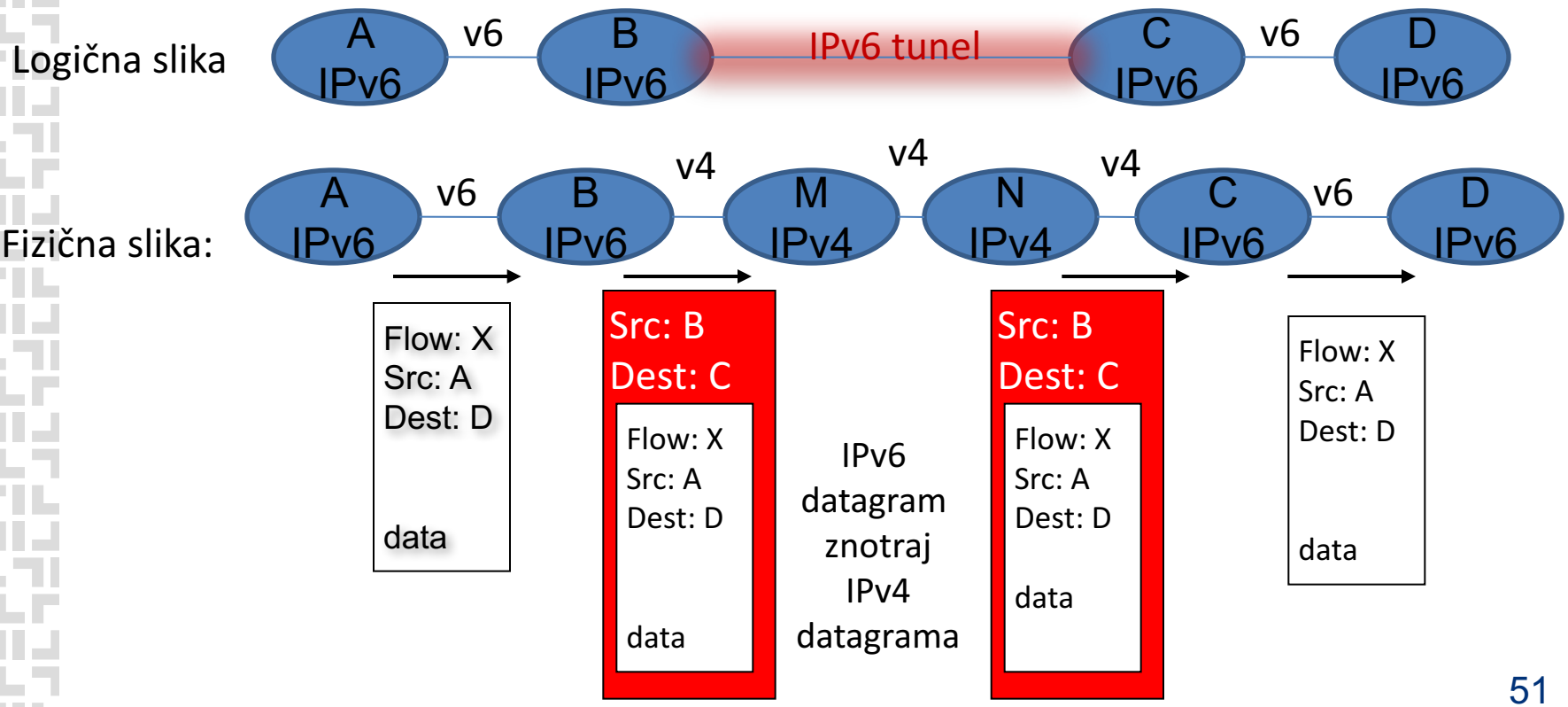
Ne bo se zgodil čez noč!

- Dual-stack: usmerjevalnik pozna v4 in v6. Z v6-enabled “govori” v6, z ostalimi pa v4.
  - Kako to ugotovi? DNS vrne v6 ali v4 naslov. DNS mora biti že na v6!
  - Če je na poti med dvema v6 vozliščema kakšno v4, se bo promet vmes pretvarjal v v4; v6-specifična polja se bodo izgubila!



Prehod IPv4-IPv6

- Tuneliranje: IPv6 datagram zapakiramo v enega ali več IPv4 datagramov kot podatke.
- Translacija naslovov (prevajanje).



Usmerjanje

- Abstraktni model:
  - teorija grafov, vozlišča, povezave.
- Algoritmi za iskanje najkrajše (najcenejše) poti
  - to je naloga usmerjevalnih algoritmov – prilagoditi posredovalne tabele tako, da bodo šli paketi po najkrajši poti.

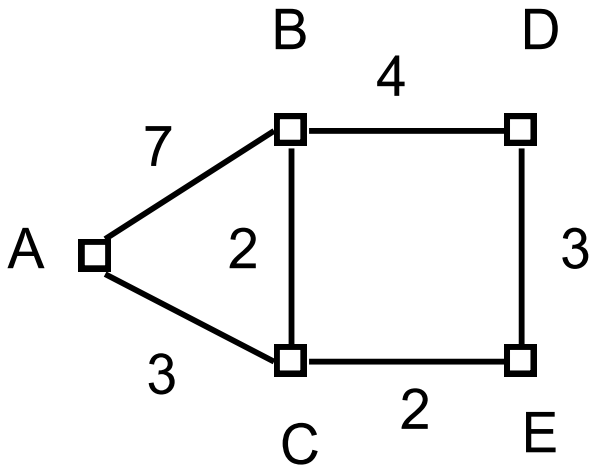
Principi

- Statično (neadaptivno) ali dinamično (adaptivno)
  - ali upošteva trenutne razmere v omrežju in jim prilagaja usmerjanje prometa?
- Po eni poti ali po več poteh
  - ali gredo v nekem trenutku vsi paketi z istim ciljem po isti poti?
- Globalno (centralizirano) ali porazdeljeno
  - Ali so pri izračunu poti znani podatki za celo omrežje?
- Prilagodljivi in neprilagodljivi na obremenitev povezav
  - prilagodljivi avtomatsko prilagajajo cene povezav glede na zasičenost povezave, s čimer dobijo manjšo ceno bolj proste poti
- Možne so vse kombinacije.
- OPTIMALNO usmerjanje:
  - Vsebovanost krajših optimalnih poti v daljši
  - Drevo ponora (sink tree)

Usmerjanje po najkrajši poti

Glede na

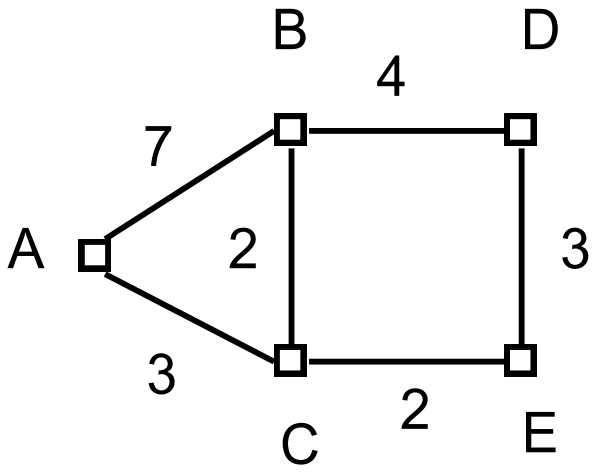
- čas,
- ceno,
- število skokov...



Usmerjanje po najkrajši poti

Glede na

- čas,
- ceno,
- število skokov...



Usmerjevalna tabela za vozlišče A
-zakasnitev
-št. skokov

- statično
- po eni poti

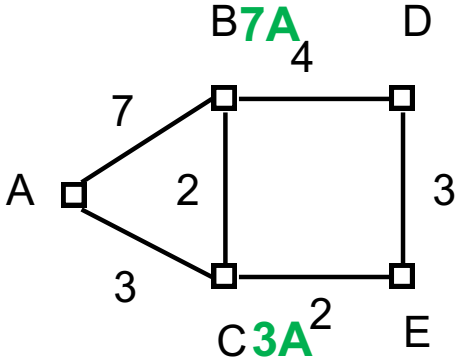
AB	AB (1)
AC	AC (1)
AD	ABD (2)
AE	ACE (2)

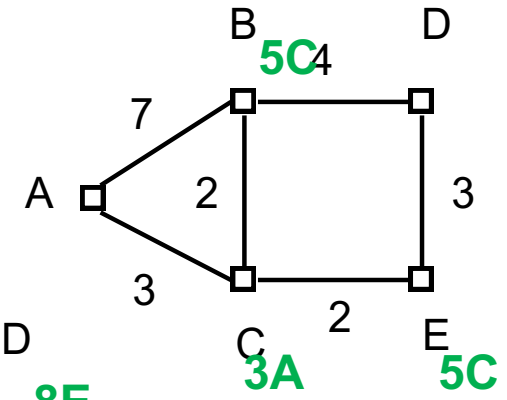
AB	ACB (5)
AC	AC (3)
AD	ACED (8)
AE	ACE (5)

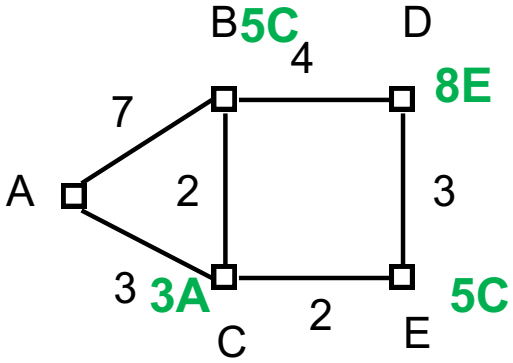
Usmerjanje po najkrajši poti: Dijkstrin algoritem za iskanje najkr. poti

Iščemo pot A-E.

- Začnemo v A
- Vsako vozlišče dobi oznako: ceno +zadnjo postajo do sedaj najboljše poti.
- V vsaki iteraciji smo korak bližje cilju.
- Nehamo, ko so vsa vozlišča označena.







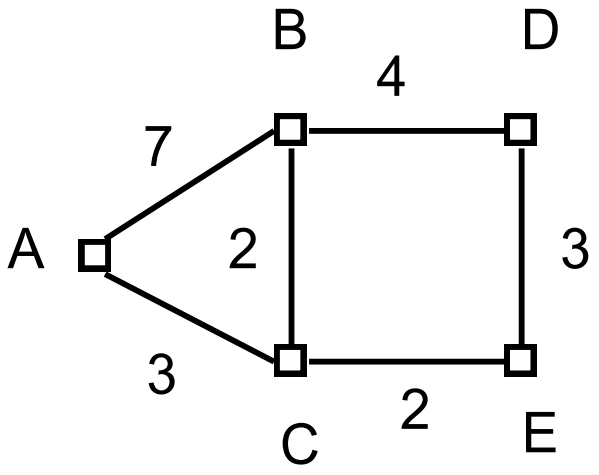
Usmerjanje po več poteh

- Določen je delež paketov za vsako izmed možnih poti.
- Uporaba npr. za uravnoteženo obremenitev (load balancing).
- Ponekod je lahko možna le ena pot.
- Paketi lahko blodijo – preprečiti!

Usmerjevalna tabela za vozlišče A:

A➔B: B 33%, C 67%
A➔D: B 50%, C 50%
A➔C: B 12%, C 88%
A➔E: C 100%

- statično
- po več poteh



Centralizirano usmerjanje

- Glavno vozlišče (*master, koordinator*)
- Zbira podatke o razmerah v omrežju
- Izračuna tabele in jih razpošlje
- Alternativa: vsi razpošiljajo podatke, vsak zase izračunava globalno usmerjanje (link state routing)
- TEŽAVA: velika omrežja s hitrimi spremembami

• dinamično
• lahko po eni ali po več poteh

Izolirano usmerjanje

- NE UPOŠTEVA razmer v omrežju
- *Hot potato*: vozlišče (usmerjevalnik) se hoče čimprej znebiti paketa, zato ga vrže
  - V najkrajšo izhodno vrsto
  - Dolžina vrste × utež
- Poplavljanje – v vse izhodne vrste
- Selektivno poplavljanje – tiste, ki so približno v pravi smeri

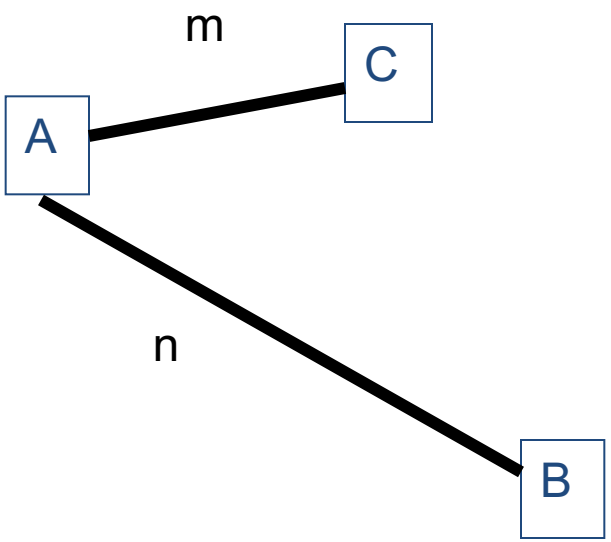
Porazdeljeno usmerjanje

- Vsako vozlišče pozna razdaljo do svojih sosedov.
- Med seboj si izmenjujejo usmerjevalne tabele (asinhrono, ob spremembah lokalnih povezav ali ob prejemu drugih sprememb)
- Potem pregledajo in prilagodijo svoje tabele.

- Lastnosti:
  - Dobre novice se širijo hitro, slabe počasi (počasi konvergira).
  - Problem štetja do neskončnosti (pomagamo si s poisoned reverse)

Porazdeljeno usmerjanje

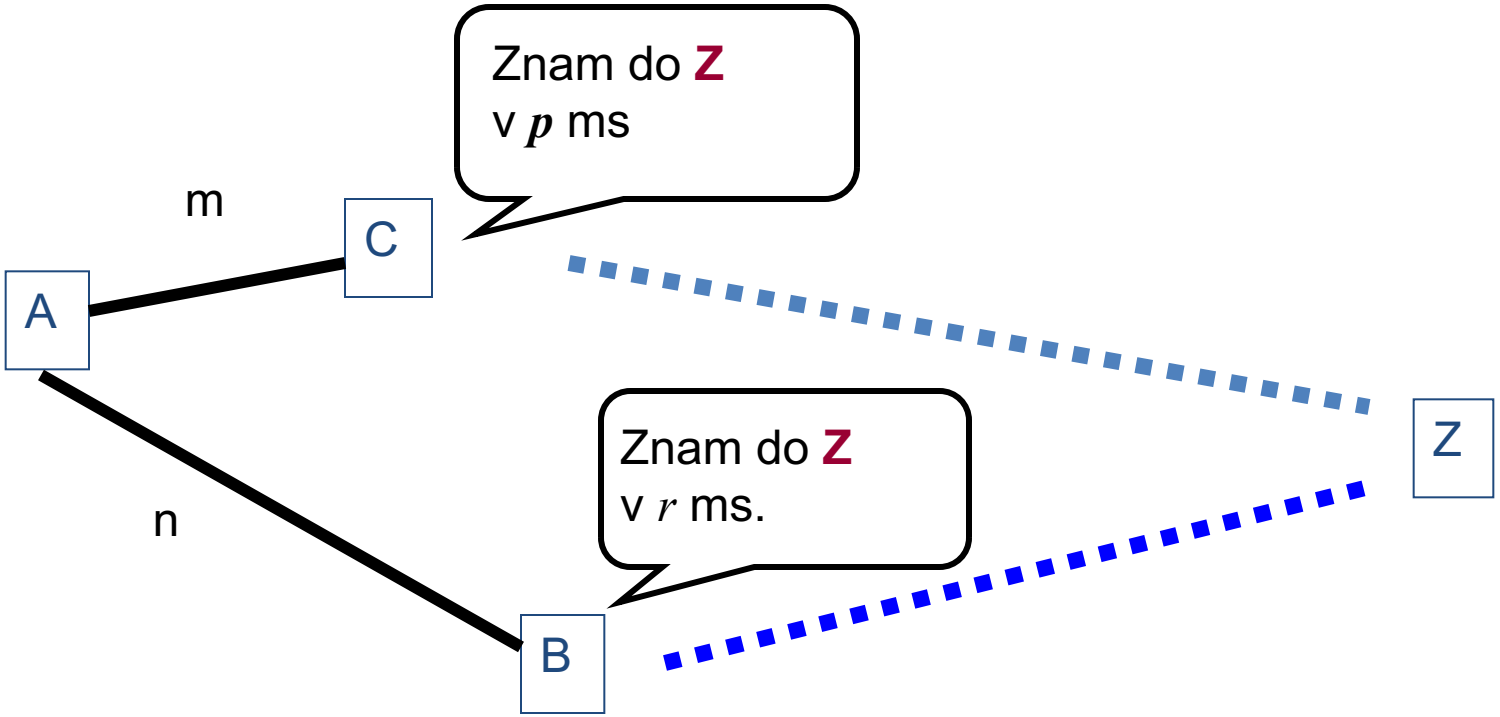
Vozlišče A pozna razdaljo (čas) do C in B.



Z

Porazdeljeno usmerjanje

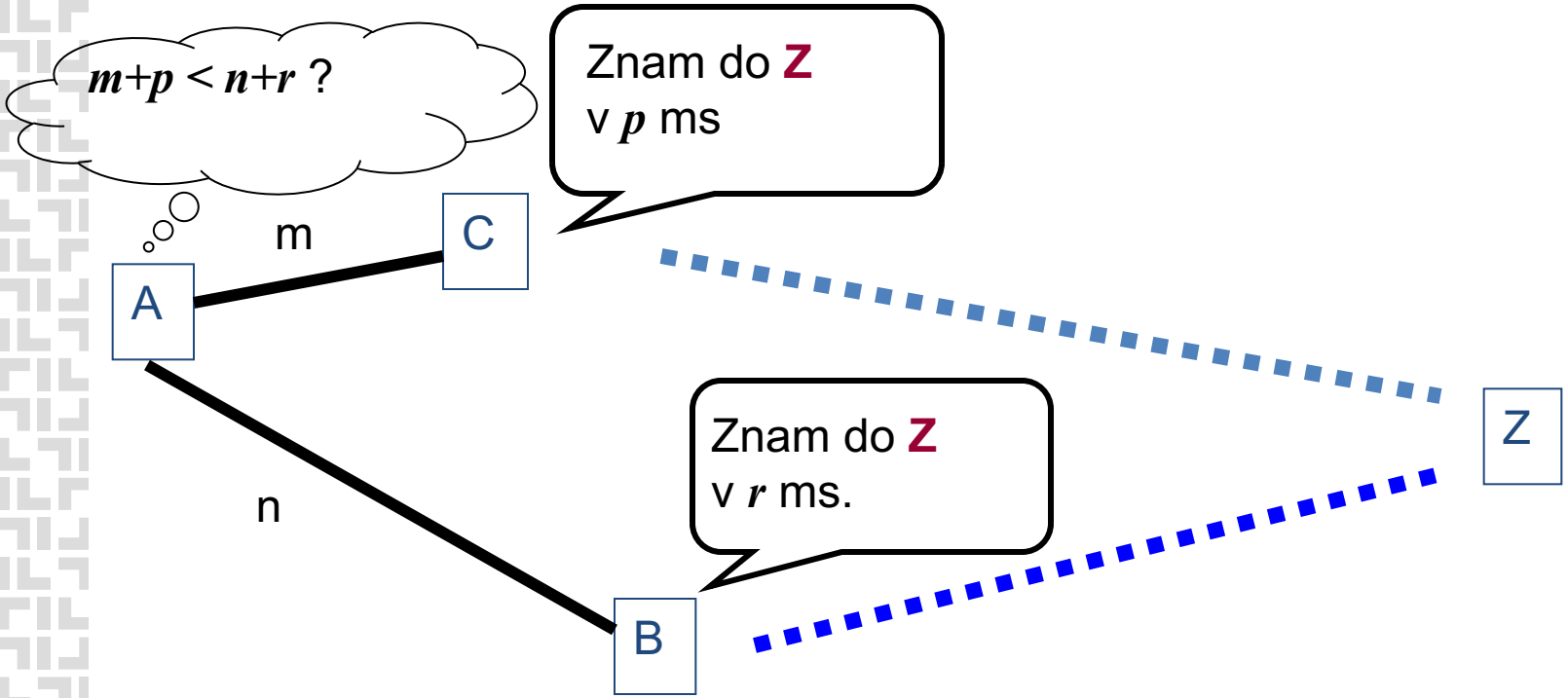
Vozlišče A pozna razdaljo (čas) do C in B.



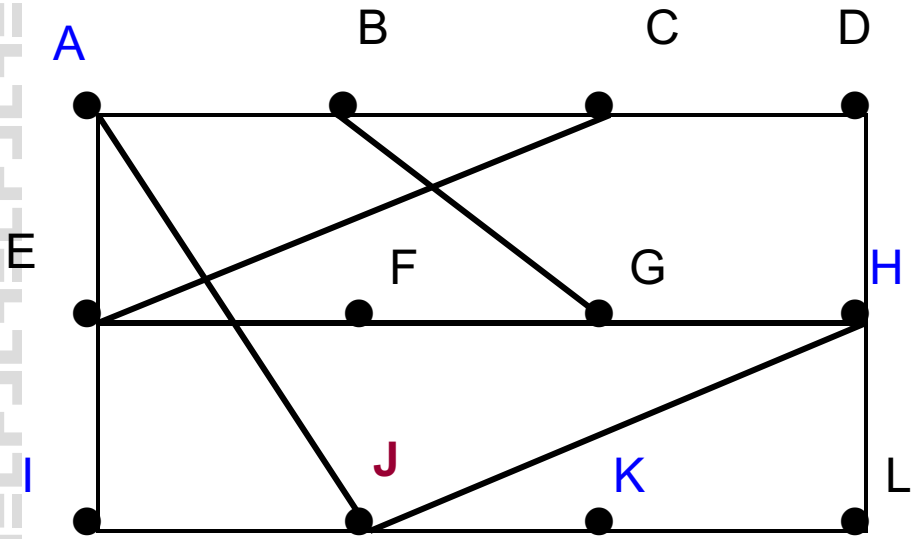
Porazdeljeno usmerjanje

Vozlišče A pozna razdaljo (čas) do C in B.

Bellman-Ford enačba za razdaljo (čas) od A do Z:



Porazdeljeno usmerjanje

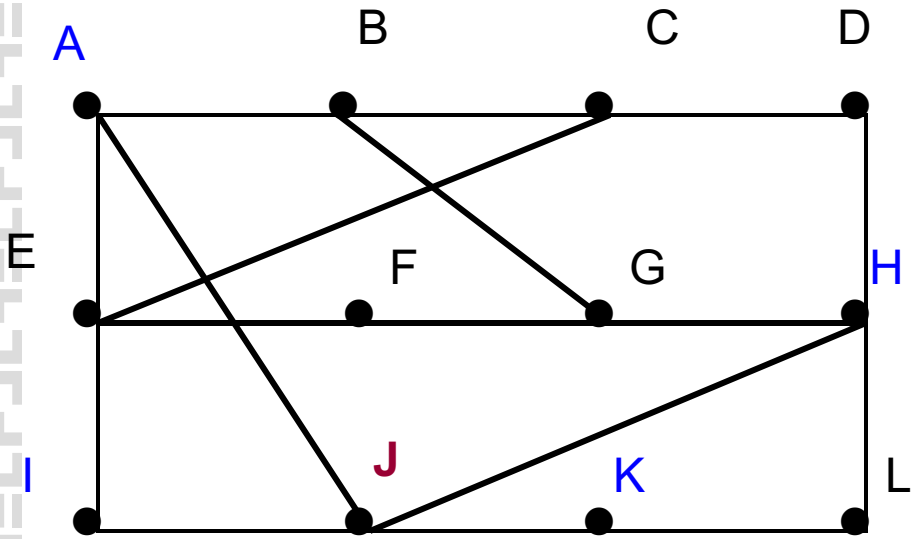


JA: 8 JH: 12

JI: 10 JK: 6

Porazdeljeno usmerjanje

J dobi tabele od sosedov. Poišči njegovo novo tabelo!

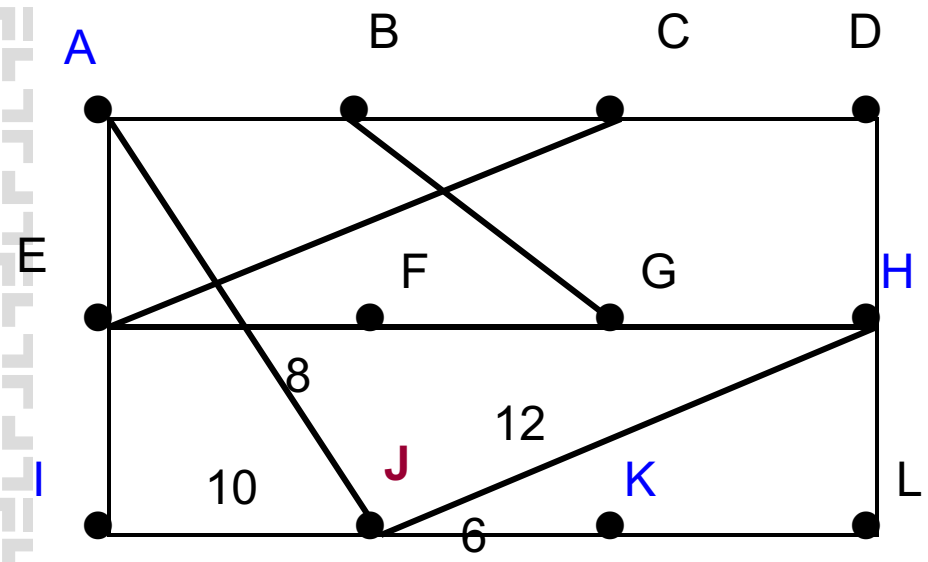


JA: 8 JH: 12

JI: 10 JK: 6

Dobi od	A	I	H	K
Smer A	-	24	20	21
B	12	36	31	28
C	25	18	19	26
D	40	27	8	24
E	14	7	30	22
F	23	20	19	40
G	18	31	6	31
H	17	20	-	19
I	21	-	14	22
J	9	11	7	10
K	24	22	22	-
L	29	33	9	9

Porazdeljeno usmerjanje -rešitev



Smer	Ocena	Sosed
A	8	A
B	20	A
C	28	I
D	20	H
E	17	I
F	30	I
G	18	H
H	12	H
I	10	I
J	-	-
K	6	K
L	15	K

Kaj če se ob naslednji izmenjavi podatki od sosedov spremenijo?

- Daljši čas (ali pretrgana pov.)
- Krajši čas

Vnos zamenjamo, če najdemo boljšo pot.

Porazdeljeno usmerjanje ali usmerjanje z vektorjem razdalj

- Ang. Distance vector routing
- Internet:
  - RIP: se opušča
  - Cisco: IGRP (ne podpira VLSM - variabilne maske) - se opušča (Interior Gateway Routing Protocol)
  - Cisco: EIGRP – Enhanced IGRP (podpira VLSM)
  - BGP – Border Gateway Protocol
- DSDV - Destination-Sequenced Distance-Vector Routing (za ad hoc mobilna omrežja)

Usmerjanje glede na stanje povezav

- Usmerjevalnik: odkrivanje sosedov, naslovi
  - HELLO paket
- Meritev zakasnitve (cene) do sosedov
  - ECHO paket: upoštevati tudi čas v vrsti ali ne?
- Izdelava paketa z vsemi temi podatki (+zap.št. in TTL)
  - Pošiljati periodično ali ob večjih spremembah?
- Pošiljanje tega paketa ostalim + sprejem ostalih (broadcast)
  - Poplavljanje, detekcija duplikatov
- Vsak: izračun najkrajših poti (npr. Dijkstra, Prim) – celotnih.

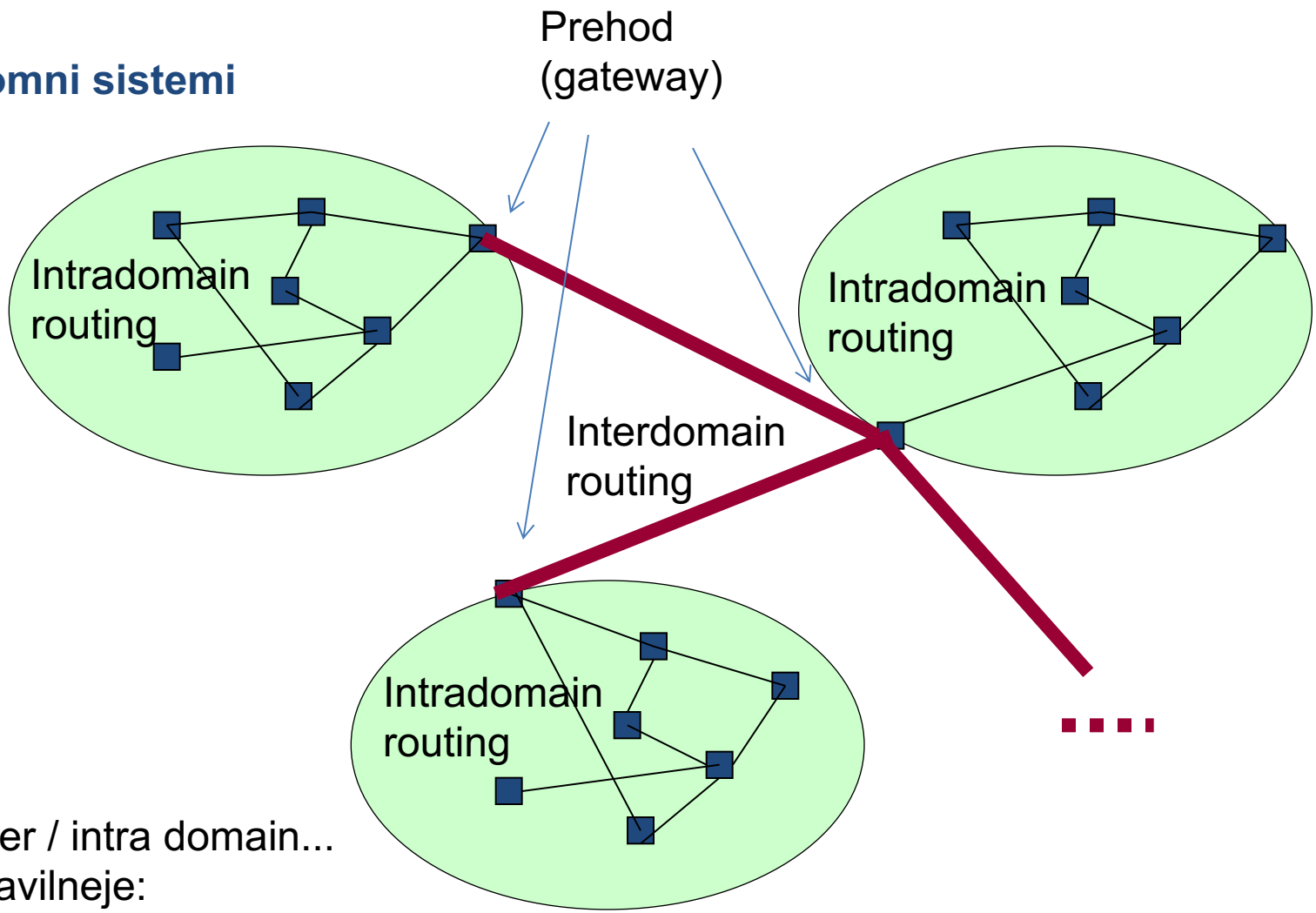
- **V praksi**: OSPF; IS-IS (interdomain).

Usmerjanje v internetu

- Usmerjanje z vektorjem razdalj (distance vector routing): porazdeljeno. RIP (se opušča).
  - Algoritem: Bellman-Ford oz. Ford-Fulkerson.
- Usmerjanje glede na stanje povezav (link state routing) – temelji na najkrajših poteh (alg.: Dijkstra).
- Usmerjanje broadcastov in multicastov
  - Vpeto drevo ali “sink tree”; Reverse path forwarding
  - Multicast: usm. mora vedeti, kateri naslovi so v grupi.
- Hierarhično usmerjanje (podomrežja, agregacija)
- Znotraj domene (AS) / med domenami (AS) - (intradomain, interdomain routing)

Internet

Avtonomni sistemi



Inter / intra domain...
Pravilneje:
Inter/ intra AS ...

Usmerjanje med avtonomnimi sistemi

- Interdomain routing, v internetu BGP4 (RFC 1771).
- ZAKAJ dve vrsti usmerjanja?
  - politika, velikost interneta, zmogljivost znotraj AS
- Medsebojno informiranje
  - AS oglašuje naslove, ki jih premore.
  - AS oglašuje (nekatere) naslove, do katerih zna usmerjati (politika).

Usmerjanje iz AS

- Če je v AS le en prehod:
  - promet, namenjen iz AS, se usmerja na ta prehod
- Če je več kot en prehod:
  - Na katerega naj se usmerja promet, namenjen iz AS?
    1. Usmerjevalnik ugotovi, da je več prehodov do X.
    2. Iz intra-AS ugotovi, do katerega prehoda pride najceneje
    3. Hot potato: promet usmeri na najcenejšega
    4. Doda ta podatek v svojo posredovalno tabelo

BGP

- BGP seje: med usmerjevalniki znotraj AS in med AS-ji
- Različna omrežja (stub – ne posreduje prometa v druge AS, multihome – več kot 1 prehod, ...)
- Ogromne tabele (več 10.000 zapisov). Naslovi v tabelah predstavljajo omrežja – CIDR prefiksi.
- BGP: prehodi usmerjajo po ideji vektorja razdalj, merilo je št. skokov.
- Oglaševanje poti gre tudi v ne-sosedne AS.
- BGP mora upoštevati še politiko, ki ni odvisna od tehnologije. Npr:
  - Promet z izvorom ali ponorom v IBM ne sme prek Microsofta.
  - Čez Sirijo pošiljamo le, če ne gre nikjer drugod.

Povezavne in nepovezavne storitve

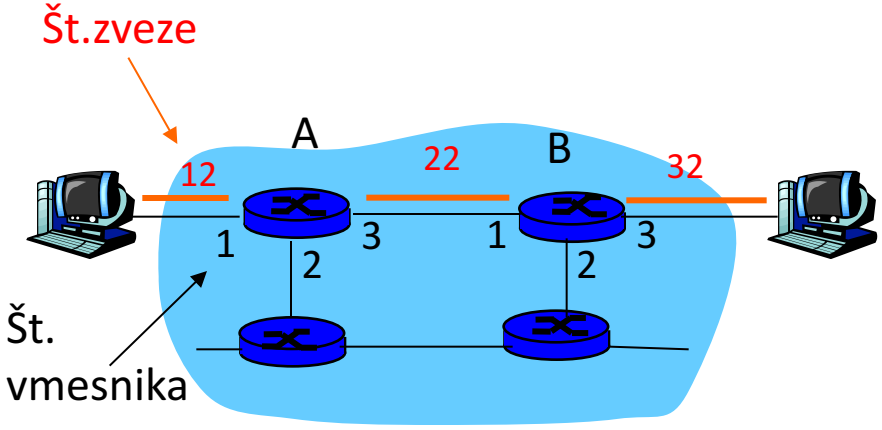
- Podobno kot pri transportnih storitvah (vendar) :
  - Storitev od izvornega do ciljnega **računalnika**
  - **Ni izbire** (le eno ali drugo - kar ponuja omrežje)
  - Izvedba je v jedru omrežne hrbtenice - usmerjevalnikih
- Datagramsko omrežje: nepovezavna storitev
- Virtualne zveze (virtual circuit): povezavna storitev, npr. ATM, Frame Relay, X.25

Virtualne zveze

- Podobno kot telefonske zveze
  - Vzpostavljanje in rušenje povezave: sodelujejo vsi usmerjevalniki na poti.
    - Signalizacija: protokoli vzp. in rušenja (klic prihaja, sprejem klica)...
  - Vsak paket ima identifikator zveze (ne naslov cilja)
  - Ob vsakem hopu se št. zveze zamenja
  - Vsak usmerjevalnik na poti: vodi stanje vsake aktivne zveze
  - Za zvezo se lahko se rezervirajo viri (vmesniki, pasovna širina)
- Elementi virtualne zveze: celotna pot, številke zveze (po ena za vsak hop), zapisi v tabelah na poti.

Virtualna zveza: primer

Povezovalna tabela: usmerjevalnik ima podatke o stanju zvez – podatke, potrebne za posredovanje.



A

Vhodni vmesnik	Vhodna št.zveze	Izhodni vmesnik	Izhodna št. zveze
1	12	3	22
2	63	1	18
3	7	2	17
...	...	...	...

B

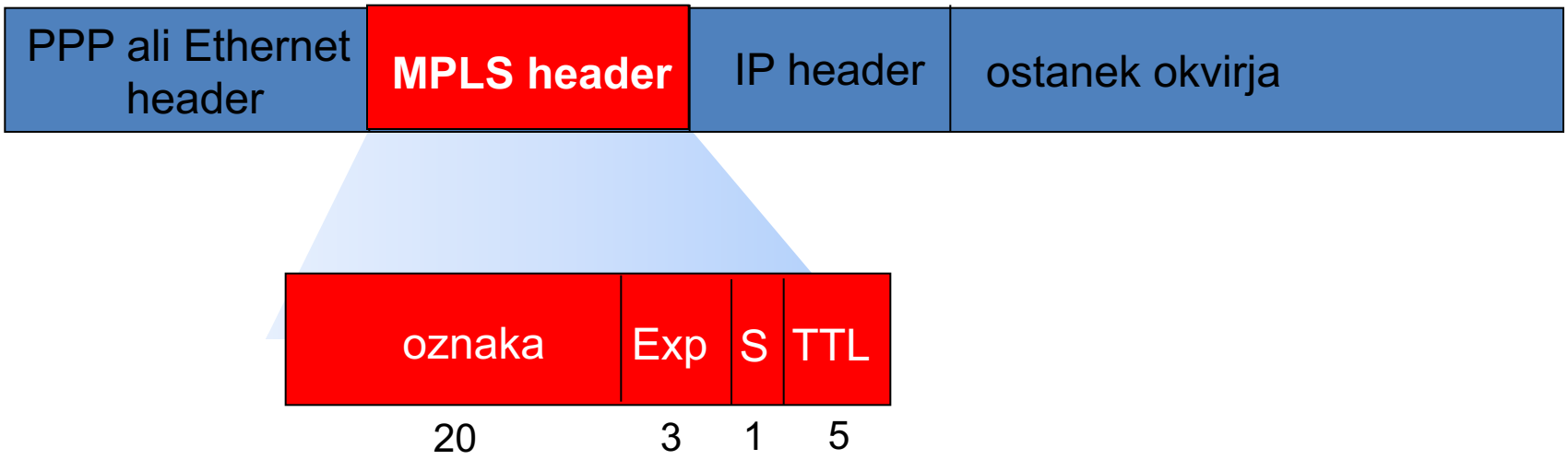
Vhodni vmesnik	Vhodna št.zveze	Izhodni vmesnik	Izhodna št. zveze
1	22	3	32
1	34	2	23
2	4	1	55
...	...	...	...

Primerjava datagramskega in omrežja z virtualnimi zvezami

Internet	ATM
Komunikacija med računalniki. Elastične storitve, čas ni tako pomemben	Izvira iz telefonije. Zakasnitev in zanesljivost sta pomembna
“Pametni” končni sistemi (računalnik)	“Neumni” končni sistemi (telefon)
Preprostejše omrežje (usmerjevalnik)	Kompleksnejše omrežje (usmerjevalnik)
Lažje dodajati nove storitve (aplikacija). Lažje povezovati heterogena omrežja.	Težje dodajati nove storitve (infrastruktura)

MPLS

- Multiprotocol label switching
- Namen: pospešiti IP usmerjanje (posredovanje):
  - Na podlagi oznake (fiksne dolžine) namesto IP naslova
  - Ideja je sposojena iz virtualnih zvez, vendar datagram obdrži IP naslov

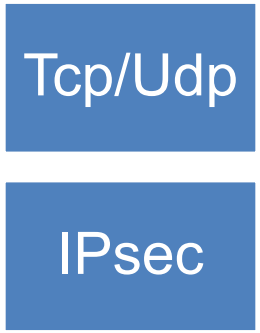


MPLS usmerjanje

- “Label-switched” usmerjevalnik
- Posredovanje paketov glede na oznako
- MPLS tabela je drugačna od usmerjevalne

  (v IP usmerjanju npr. določanje poti glede na izvor prometa ne bi bilo možno)
- Potreben je signalizacijski protokol za vzpostavljanje poti (RSVP – Resource ReSerVation Protocol, RFC 2205)
- Dobra združljivost z IP-usmerjevalniki

Varnost in IP: IPsec



- Uporaba za VPN:
  - kriptiranje prometa od izvora do ponora
  - IPsec funkcionalnost potrebna le na izvoru in ponoru
  - IPsec plast vzame transportni segment, ga kriptira, doda svojo glavo in vse to zapakira kot telo v navaden IP datagram.
- Storitve:
  - Dogovor o kriptografiji in ključih
  - Enkripcija in dekripcija
  - Integriteta podatkov
  - Avtentikacija izvora