

QoS arhitekture i MPLS

(Multi Protocol Label Switching)

Motiv

■ IP

- Prvi definisani i korišćeni protokol
- *De-facto* protokol za globalno funkcionisanje Interneta

... *međutim ima nedostatke*

■ Mane IP rutiranja:

- *Connectionless*
 - Nema garancije kvaliteta servisa
- Svaki ruter nezavisno donosi odluke o rutiranju na osnovu IP adrese
- Veliko IP zaglavlje
 - Najmanje 20 B
- Rutiranje na mrežnom nivou
 - Sporije od komutacije
- Obično dizajnirano tako da koristi najkraće rute
 - Ne uzima u obzir dodatne metrike

Motiv

- **ATM (Asynchronous Transfer Mode)**
 - Konekcijski orijentisan
 - Podržava QoS
 - Brza komutacija paketa fiksne dužine (ćelija)

... *ali ima nedostatke*

- **Nedostaci ATM-a**
 - Kompleksan
 - Skup
 - Slabo prihvaćen

Motiv

- **Ideja:** Kombinovati najbolje od IP i tehnologija komutacije kola

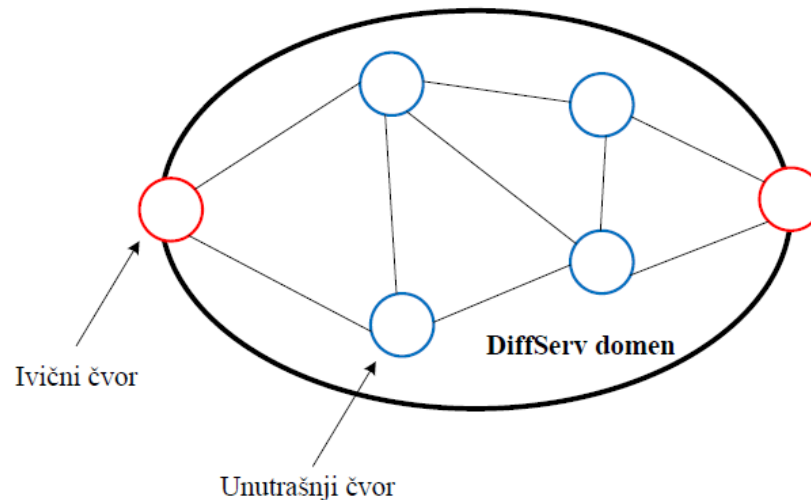


QoS arhitekture

- Dvije osnovne arhitekture
 - **DiffServ** - diferencijalni servisi
 - **IntServ** - integrisani servisi
- DiffServ
 - Korisnički tokovi raspoređuju se u odgovarajuće klase servisa, a ruter potom vrši opsluživanje na nivou agregiranih tokova
 - pri opsluživanju koristi se samo informacija kojoj klasi servisa pripada paket
 - Skalabilnije rešenje
- IntServ
 - Posmatraju se individualni tokovi, za koje se rezervišu resursi na početku veze u svim ruterima na ruti
 - Striktne QoS garancije

DiffServ

- RFC 2474 i RFC 2475
- DiffServ domen:
 - Ruteri sa DiffServ podrškom koji su pod kontrolom istog provajdera.
 - Ruteri na ivici domena se nazivaju **ivičnim** ili graničnim ruterima, a ruteri u unutrašnjosti domena se nazivaju **unutrašnji** ruteri

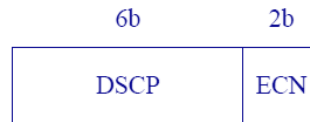


DiffServ

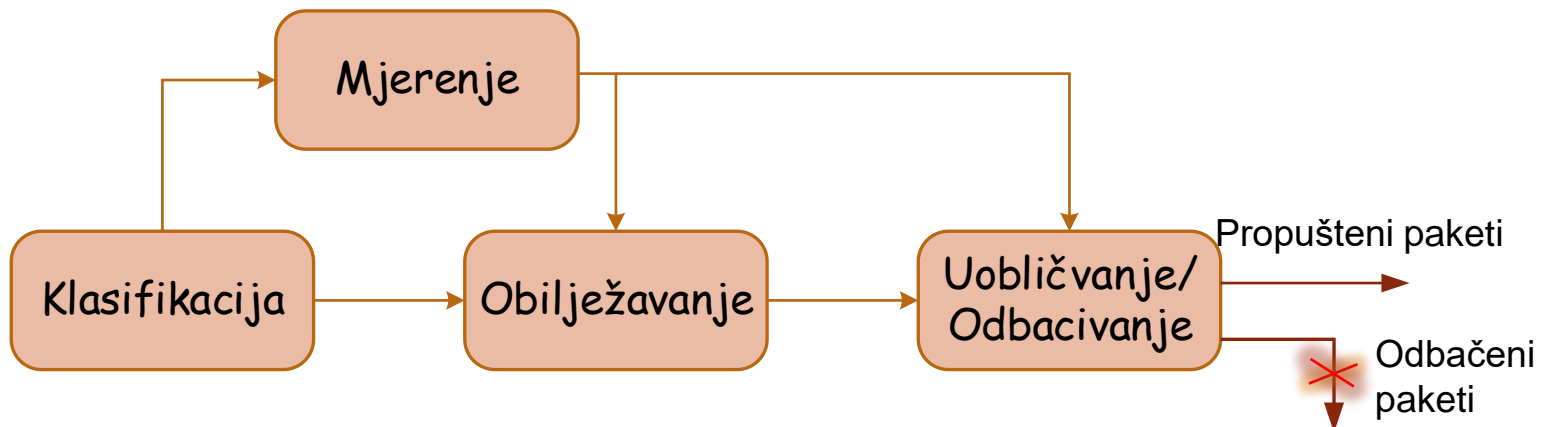
- Korisnici sa provajderom sklapaju dogovor (**SLA** - *Service Level Agreement*)
- SLA definiše parametre saobraćaja korisnika i željeni kvalitet usluge koji mreža treba da pruži korisniku.
 - Saobraćaj se tipično predstavlja u vidu parametara *leaky bucket*-a **b** i **r**, gde **b** predstavlja veličinu bursta, a **r** prosječan protok.
 - Između DiffServ domena se takođe sklapa SLA.
 - SLA može da se podesi statički (administratori domena vrše podešavanje), ali može i dinamički pomoću nekog protokola (na primjer, RSVP).
 - Da li će dogovor biti postignut ili ne zavisi i od trenutnih tokova u mreži, pa ako tok traži resurse koji ne mogu da se zadovolje dogovor neće biti postignut.

DiffServ

- Klasa servisa se stavlja u odgovarajuće polje IP zaglavlja



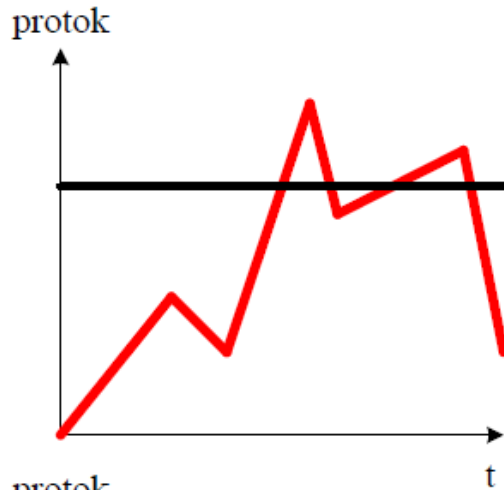
- ToS polje u IPv4 zaglavlju, odnosno Traffic Class polje u IPv6 zaglavlju
- DSCP (*DiffServ CodePoint*) polje označava klasu servisa i dužine je 6 bita, dok preostala dva bita čine ECN (*Explicit Congestion Notification*) polje koje se koristi u mehanizmu eksplicitnog obavještanja o zagušenju u mreži.



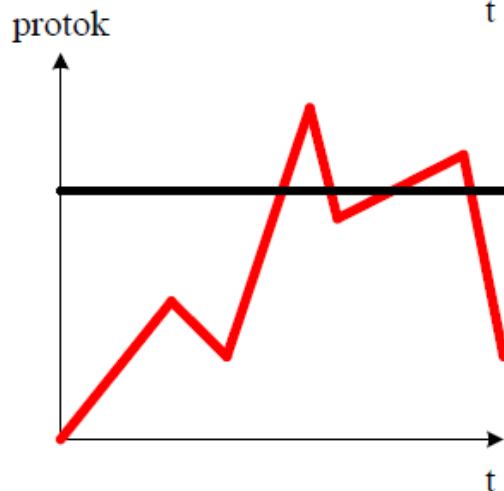
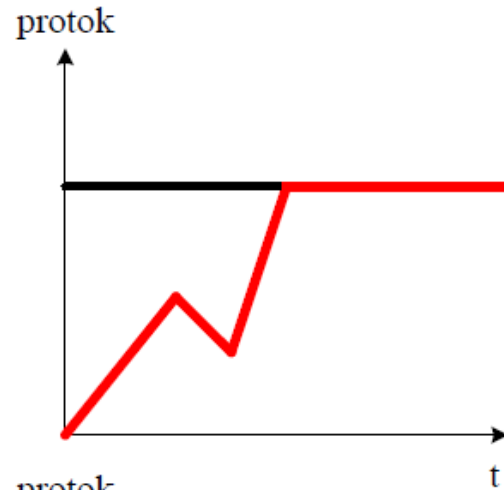
DiffServ

- **Funkcija klasifikacije** vrši ispitivanje zaglavlja paketa da bi odredila kojoj klasi servisa pripada paket
 - prije svega IP zaglavlja, ali može da se ispituje i transportno zaglavlje (TCP, UDP...), pa čak i aplikacioni dio paketa, kao i interfejs na koji je paket pristigao
- **Funkcija obilježavanja** vrši dodjelu odgovarajuće klase paketu u IP zaglavlju.
 - Postavlja odgovarajuću vrijednost DSCP koda u IP zaglavlju paketa
- **Funkcija uobličavanja** vrši uobličavanje toka
 - Implementira bafer kojim se mogu zakasniti paketi koji narušavaju saobraćajne parametre tako da njihovo zakašnjeno slanje bude u skladu sa dogovorenim saobraćajnim parametrima
- **Funkcija odbacivanja** vrši odbacivanje paketa koji su narušili dogovorene saobraćajne parametre
 - Odbacuje svaki paket koji naruši dogovor
- **Funkcija mjerenja** prati saobraćajne parametre korisničkog toka i detektuje pakete koji narušavaju dogovorene parametre

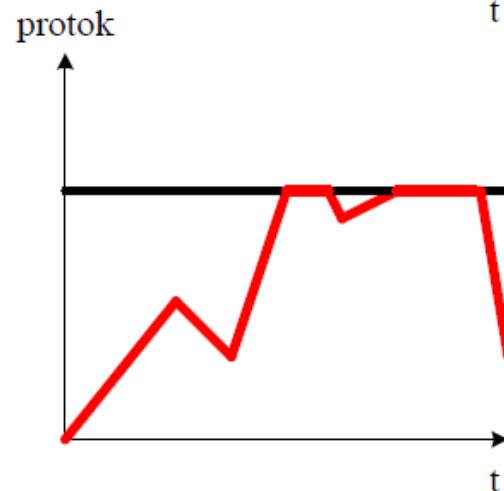
DiffServ



Uobličavanje
→



Odbacivanje
→



DiffServ klase servisa

Definisane DSCP vrijednosti

| DSCP | PHB (<i>Per Hop Behavior</i>) klasa servisa |
|--|---|
| 000000 | Difolt klasa (<i>best effort</i> saobraćaj) |
| 101110 | EF (<i>Expedited Forwarding</i>) klasa |
| 001010, 001100, 001110, 010010, 010100, 010110, 011010, 011100, 011110, 100010, 100100, 100110 | AF (<i>Assured Forwarding</i>) klasa |
| xxx000 | CS (<i>Class Selector</i>) klasa |

AF klasa

| | Potklasa 1 | Potklasa 2 | Potklasa 3 | Potklasa 4 |
|-------------------|---------------|---------------|---------------|---------------|
| Nizak prioritet | 001010 (AF11) | 010010 (AF21) | 011010 (AF31) | 100010 (AF41) |
| Srednji prioritet | 001100 (AF12) | 010100 (AF22) | 011100 (AF32) | 100100 (AF42) |
| Najviši prioritet | 001110 (AF13) | 010110 (AF23) | 011110 (AF33) | 100110 (AF43) |

DiffServ klase servisa

- **Difolt klasa** predstavlja *best effort* saobraćaj.
- **EF klasa** se koristi za saobraćaj koji je veoma osjetljiv na kašnjenja i varijacije kašnjenja (džiter)
 - Npr. telefonski saobraćaj
 - Ima najviši prioritet pri opsluživanju
 - Ako agregirani EF saobraćaj postane prevelik može doći do problema u kvalitetu opsluživanja
 - Za ovaj saobraćaj se obavezno vrši procjena da li može da se prihvati ili ne na osnovu trenutnih EF tokova u mreži
- **AF klasa** pokušava da ostvari željene protoke korisnika, ali uz varijacije u opsluživanju agregiranih tokova u zavisnosti od trenutnog stanja u mreži
 - Ideja je da ovi tokovi dobiju bolje opsluživanje od *best effort* saobraćaja.
 - Definišu se 4 potklase (što je veći redni broj potklase to je ona prioritetnija), pri čemu se za svaku potklasu dodatno definišu 3 nivoa prioriteta
- **CS klasa** je uvedena radi kompatibilnosti sa starijim sistemima

| Tip saobraćaja | Klasa servisa | DSCP vrijednost | Primjer aplikacije |
|----------------------------------|----------------------|------------------------|--|
| Mrežna kontrola | CS6 | 110000 | Protokoli rutiranja |
| Telefonija | EF | 101110 | VOIP |
| Signalizacija | CS5 | 101000 | VOIP signalizacija |
| Multimedijalna konferencija | AF41, AF42, AF43 | 100010, 100100, 100110 | H.323 video konferencija |
| Interaktivnost u realnom vremenu | CS4 | 100000 | Interaktivne igrice |
| Multimedijalni striming | AF31, AF32, AF33 | 011010, 011100, 011110 | Audio/video striming na zahtjev |
| Brodcast video | CS3 | 011000 | Brodcast TV |
| Podaci osjetljivi na kašnjenje | AF21, AF22, AF23 | 010010, 010100, 010110 | Klijent/server transakcije |
| OAM | CS2 | 010000 | OAM |
| Podaci visokog protoka | AF11, AF12, AF13 | 001010, 001100, 001110 | Uskladišti i prosljedi aplikacije |
| Podaci niskog prioriteta | CS1 | 0010000 | Aplikacije za prenos podatka koje nemaju garancije protoka |

DiffServ

- Unutrašnji čvorovi vrše opsluživanje paketa na bazi klasa servisa
 - Tipično paketi klase servisa višeg prioriteta budu prvi opsluženi, a poslednji paketi *best effort* klase
- PHB - *per hop behavior*
 - Opsluživanje sa stanovišta paketa je po hopu
 - U svakom ruteru se vrši identičan princip opsluživanja
 - Princip opsluživanja zavisi od konkretne implementacije (npr. WFQ)
 - Ovo omogućava jednostavniju realizaciju unutrašnjih rutera jer je najkomplicovaniji proces opsluživanja smešten u ivične rutere.
 - I unutrašnji ruteri mogu da vrše slične funkcije kao ivični ruteri, **na nivou agregiranih tokova**, i da mijenjaju vrednosti DSCP koda u slučaju paketa koji narušava saobraćajne parametre (ali sada sa stanovišta agregiranih tokova).
 - Mijenjanje DSCP koda uvijek je na lošiju klasu servisa, tipično *best effort* klasu saobraćaja.

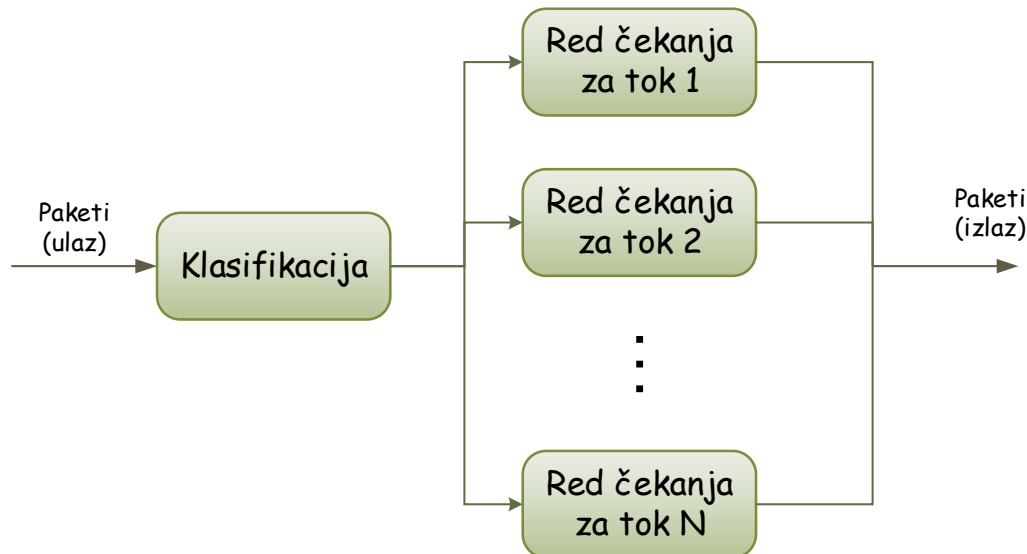
DiffServ

Zaključak

- DiffServ je skalabilan jer njegova kompleksnost zavisi samo od broja klasa servisa.
- Najveće procesiranje se odvija u ulaznim ivičnim čvorovima
- U unutrašnjim čvorovima se vrši opsluživanje agregiranih tokova sa stanovišta klasa servisa čime je postignuta ekonomična implementacija unutrašnjih rutera.
- Ne može da pruži striktne garancije jer ne razlikuje individualne tokove, a takođe, granularnost QoS nivoa nije velika pošto se opslužuju agregirani tokovi podijeljeni u nekoliko klasa servisa.
- Problematična je podrška multikast tokovima jer se takvi tokovi računaju u mreži, a i broj korisnika se tipično mijenja tokom vremena.

IntServ

- RFC 1633
- Individualno opsluživanje tokova
- Vršiti se uspostavljanje konekcije za tok
 - Rezervisanje resursa u ruterima duž kojih će prolaziti paketi dotičnog toka
 - RSVP protokol se koristi u procesu rezervacije resursa
 - Nakon rezervisanja resursa, paketi toka mogu početi da se šalju, a ruteri na putu će obezbediti da tok primi željeni kvalitet servisa



IntServ

- **Klasifikacija** ispituje pristigli paket da bi odredila kom toku pripada dotični paket i potom ga prosleđuje u odgovarajući red za čekanje.
- **Raspoređivanje** paketa iz svih tokova (redova za čekanje) za slanje na izlazni link se vrši po odgovarajućem algoritmu raspoređivanja.
 - Koji algoritam raspoređivanja će se koristiti zavisi od same implementacije.
- Željeni kvalitet servisa se opisuje sa stanovišta prijemne strane koja definiše tzv. *RSPEC (Request Specifications)* specifikacije željenog kvaliteta servisa
 - Na primjer, maksimalno kašnjenje s kraja na kraj.
 - Karakteristike saobraćaja opisuje predajna strana, koja definiše *TSPEC (Traffic Specifications)* specifikacije saobraćaja, tipično u vidu karakteristika *leaky bucket*-a (veličina bursta **b** i protok **r**).
 - Na osnovu ovih parametara se procjenjuje da li mreža može da prihvati ovakav tok ili ne.

IntServ klase saobraćaja

- **Garantovani servis** (*guaranteed service*)
 - Namijenjen *real-time* aplikacijama
 - Korisnik specificira karakteristike saobraćaja i servisne zahtjeve
 - Teoretski garantuje propusnost, kašnjenje i džiter
 - Definišu se RSPEC i TSPEC zahtjevi
 - Zahtjeva kontrolu pristupa na nivou svakog rutera
- **Kontrolisano opterećenje** (*controlled load*)
 - Tokovi će primati veoma sličan kvalitet servisa i u slučaju kada je mreža slabo opterećena i kada je mreža preopterećena (zagušena)
 - Korisnik definiše karakteristike saobraćaja i potrebni kapacitet
 - Kontrola pristupa bazirana na mjerenjima
 - QoS garancije nisu jake
 - Ne definiše se RSPEC
 - Paketi koji naruše TSPEC specifikaciju se ili odbacuju ili se prosleđuju kao best-effort paketi
- **Najbolji pokušaj** (*best-effort*)

IntServ

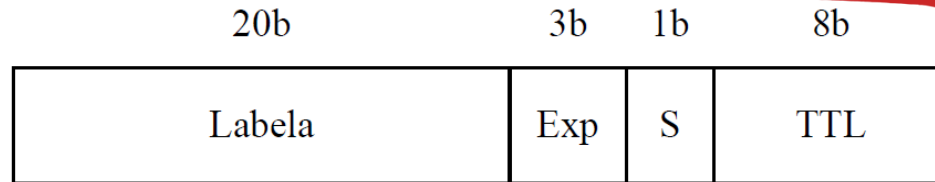
Zaključak:

- IntServ arhitektura omogućava specificiranje kvaliteta servisa na nivou samih tokova
 - TSPEC i RSPEC specifikacijama mogu se fino podesiti parametri kvaliteta servisa.
- QoS garancije od kraja do kraja
 - Kod DiffServ-a individualni QoS zavisi od trenutne količine agregiranih tokova i njihovog trenutnog saobraćaja
- Nije skalabilno rešenje
 - Svi ruteri na ruti moraju da rade sa individualnim tokovima što predstavlja značajno opterećenje
- Problem i DiffServ i IntServ arhitekture je zahtjev da svi ruteri na ruti toka moraju da podržavaju DiffServ ili IntServ arhitekturu
 - Dovoljno je da samo jedan ruter na ruti ne podržava QoS (DiffServ ili IntServ) pa da se izgube garancije kvaliteta servisa toku

MPLS

- MPLS domen:
 - Čine ga ruteri koji imaju sposobnost prosleđivanja na bazi MPLS labela
- Ruteri koji prosleđuju pakete na bazi MPLS labela se nazivaju LSR (*Label Switching Routers*) ruteri
- Ivični LSR ruteri, koji se nalaze na granici MPLS domena, nazivaju i LER (*Label Edge Routers*) ruteri
- LSR ruteri vrše prosleđivanje na bazi MPLS labela
- Gdje se smješta MPLS labela?
 - U zaglavlje frejma ukoliko ima mjesta za nju
 - Ukoliko nema prostora za smještanje MPLS labele, onda se koristi tzv. *shim* zaglavlje koje u suštini predstavlja polje MPLS labele koje se smiješta između zaglavlja drugog i trećeg sloja
- Poznat kao protokol 2.5 sloja

MPLS zaglavlje



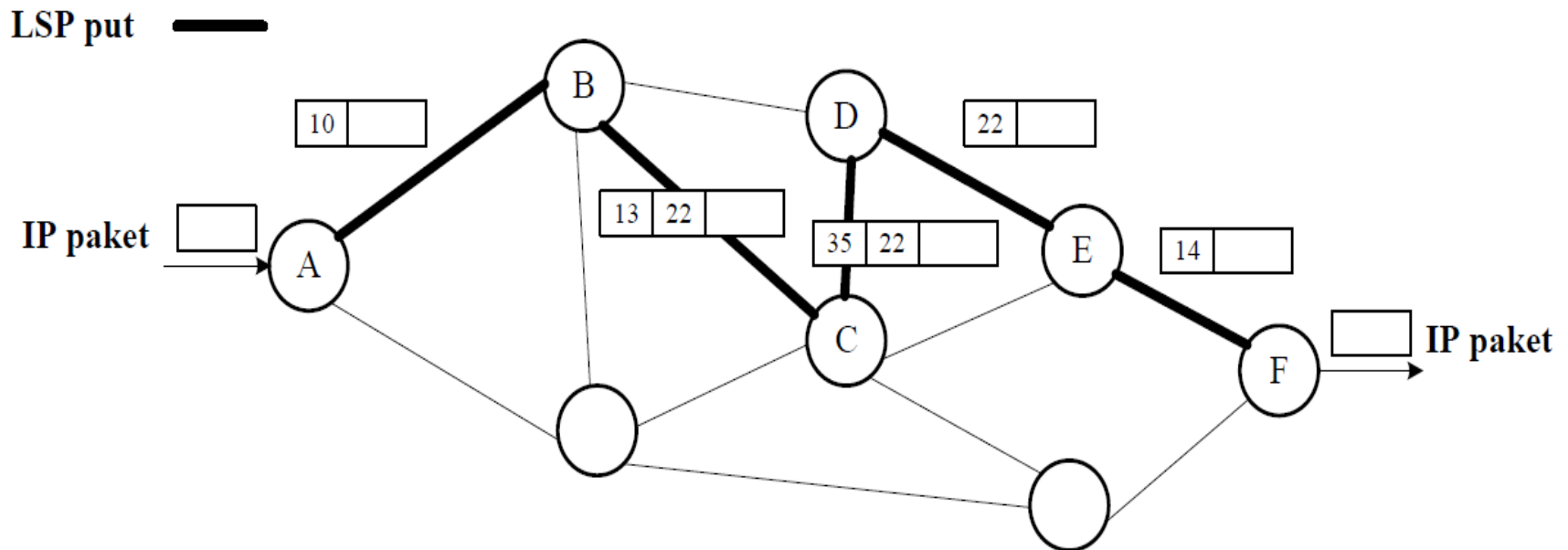
- Eksperimentalno polje je rezervisano u eksperimentalne svrhe i može se iskoristiti za definisanje QoS prioriteta i ECN obavještavanje o zagušenju.
- **S** bit označava da li je u pitanju hijerarhijski najniža labela (vrijednost 1) ili ne (vrijednost 0).
 - Dozvoljeno je kreiranje hijerarhije MPLS domena, gdje svaki hijerarhijski nivo dodaje svoju labelu (pod labelom se podrazumijeva kompletno *shim* zaglavlje) pa se stoga kreira tzv. stek labela
 - Stek labela može imati proizvoljnu dubinu
 - **S** bit označava da li je u pitanju hijerarhijski najniža labela u steku (ona koja je prva dodata na paket)
- TTL (*Time To Live*) polje se koristi kao zamjena za TTL polje iz IP zaglavlja

MPLS prosleđivanje

- Usmjeravanje paketa kroz MPLS domen se vrši na bazi labela
 - Labele imaju lokalni značaj
- LER ruteri na ulazu u MPLS domen dodaju labelu paketu
- LSR ruteri prosleđuju pakete na bazi labela tako što za primljenu labelu traže odgovarajući zapis u tabeli usmjeravanja na bazi labela
 - U okviru tog zapisa se nalazi i podatak o novoj vrijednosti labela kojom treba zamijeniti labelu u primljenom paketu pre nego što se paket prosljedi na odgovarajući izlaz LSR rutera.
- LER ruter na izlazu iz MPLS domena skida labelu sa paketa
- U suštini MPLS ruteri (LSR i LER) mogu vršiti tri osnovne operacije nad labelama:
 - dodavanje (**push**)
 - skidanje (**pop**)
 - zamjena (**swap**)
- Put paketa kroz MPLS domen se naziva LSP (*Label Switched Path*) put.

Primjer MPLS usmjeravanja

- **Push** operaciju tipično rade LER ruteri na ulazu u MPLS domen
- **Pop** operaciju tipično rade LER ruteri na izlazu iz MPLS domena
- **Swap** labela tipično rade LSR ruteri, ali i ulazni LER ruteri na granicama između dvije hijerarhije MPLS domen
- Poslednji ruter na LSP putu vrši dvije pretrage tabela usmjeravanja - na bazi labela i na bazi IP adresa
 - Alternativa je *penultimate hop popping*



FEC (*Forwarding Equivalence Class*)

- FEC dešiniše logiku za dodjeljivanje labele.
- U preporuci RFC 3031 se FEC klasa definiše kao mrežni prefiks, ali dozvoljeno je korišćenje i drugih definicija koje bi uključivale polja IP ili transportnih zaglavlja, kao i interfejse po kojima pristižu paketi.
 - Na primer, FEC klase mogu da budu formirane imajući u vidu i QoS zahtjeve
- Paketi iste FEC klase prolaze kroz identičan LSP put i imaju identično opsluživanje ako ulaze u MPLS domen na istom mjestu.
- Određivanje pripadnosti paketa FEC klasi se vrši samo na ulazu u MPLS domen.
- Vezivanje labele za FEC klasu može biti na nivou platforme (uređaja) ili na nivou interfejsa.
- LSR/LER ruter kada izvrši vezivanje labele za FEC klasu formira automatski i odgovarajući zapis u tabeli usmjeravanja na bazi labele.

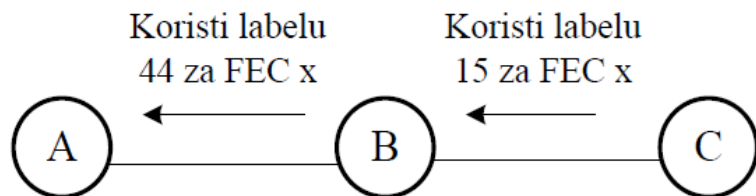
FEC

- Veživanje labele može biti **nezavisno** ili **uređeno**
 - **Nezavisno veživanje:** ruter čim prepozna novu FEC klasu vrši veživanje slobodne labele iz skupa labele.
 - **Uređeno veživanje:** ruter vrši veživanje labele za FEC klasu samo ako je on izlazni LER ruter za dotičnu FEC klasu ili ako je primio oglašenu labelu za tu FEC klasu od rutera koji predstavlja sledeći hop za dotični ruter
- **Bezuslovno oglašavanje labele** (*Unsolicited Downstream*) :
 - Ruter oglašava labelu koju je vezao za FEC klasu svojim susjedima.
 - Uzvodni (*upstream*) susjed će da prihvati tu labelu i upiše je u odgovarajuće mjesto u svojoj tabeli usmjeravanja na bazi labele (ili IP adresa ako je u pitanju LER ruter na ulazu u MPLS domen).
 - Uzvodni susjed je onaj čiji je sledeći hop za dotičnu FEC klasu upravo dotični ruter koji je oglasio labelu.
 - U uzvodnom ruteru u zapisu koji odgovara dotičnoj FEC klasi će biti unijeta vrijednost labele koja se mora staviti na vrh steka labele (ulaz u MPLS domen - LER) ili koja mora zamijeniti labelu sa vrha steka (LSR) paketa.
 - Susjedni ruteri koji nisu uzvodni ruteri u zavisnosti od konfigurisanog moda rada ili ignorišu oglašavanje ili ga prihvataju.

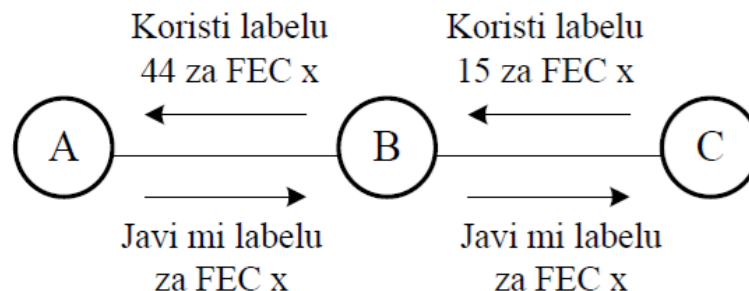
FEC

- **Oglašavanje na zahtjev** (*Downstream on demand*)
 - Oglašavanje se vrši samo ka (uzvodnim) susjedima koji pošalju eksplicitan zahtjev za labelom

Bezuslovno oglašavanje



Oglašavanje na zahtjev



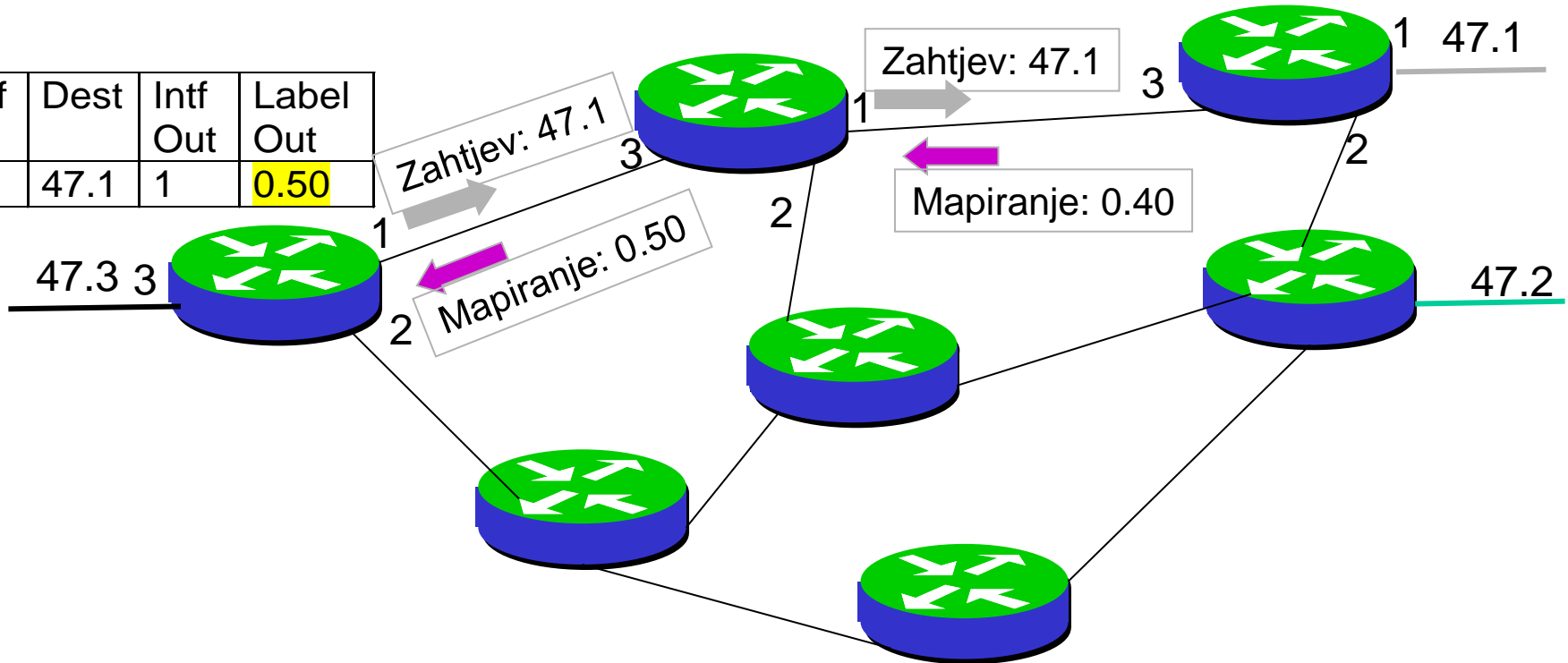
- Za oglašavanje labela neophodan je odgovarajući signalizacioni protokol
 - MPLS ne specificira određeni protokol
 - Dva načina realizacije:
 - signalizacioni protokol razvijen specijalno za ovu namjenu (npr. LDP)
 - postojeći protokoli rutiranja i signalizacije koji se proširuju tako da podrže i oglašavanje labela (npr. RSVP i BGP)

Distribucija labela

| Intf In | Label In | Dest | Intf Out | Label Out |
|---------|----------|------|----------|-----------|
| 3 | 0.50 | 47.1 | 1 | 0.40 |

| Intf In | Label In | Dest | Intf Out |
|---------|----------|------|----------|
| 3 | 0.40 | 47.1 | 1 |

| Intf In | Dest | Intf Out | Label Out |
|---------|------|----------|-----------|
| 3 | 47.1 | 1 | 0.50 |

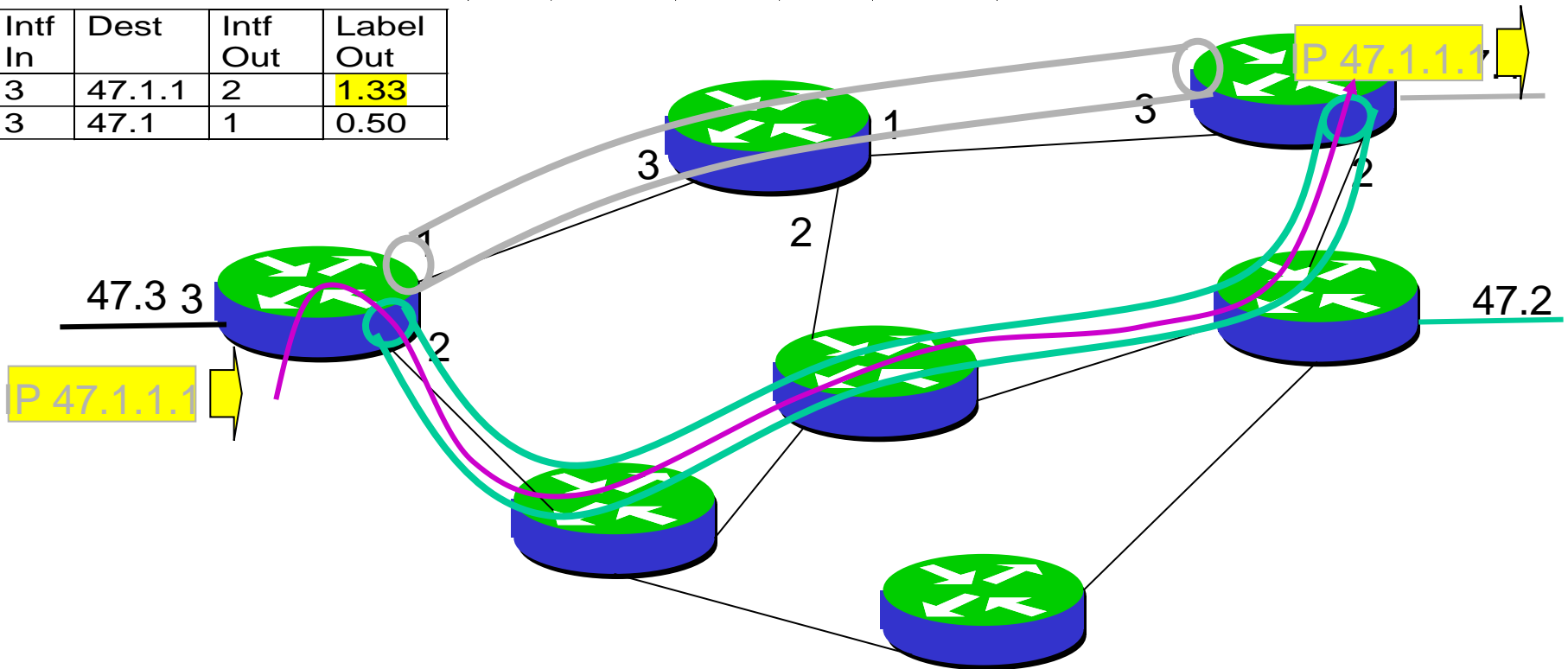


Distribucija labela

| Intf In | Label In | Dest | Intf Out | Label Out |
|---------|----------|------|----------|-----------|
| 3 | 0.50 | 47.1 | 1 | 0.40 |

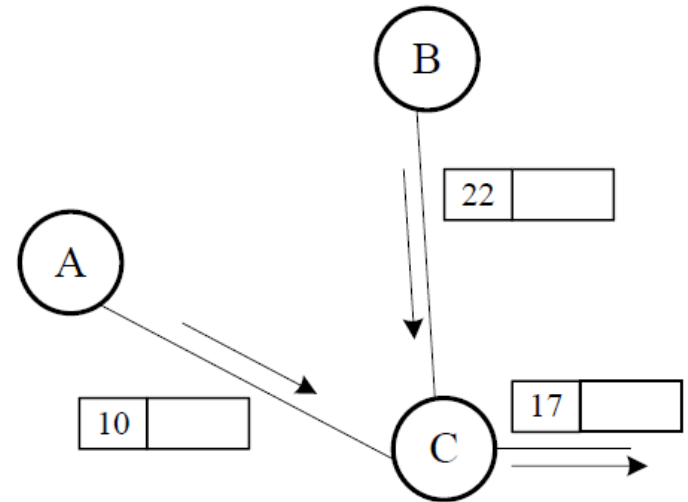
| Intf In | Label In | Dest | Intf Out |
|---------|----------|------|----------|
| 3 | 0.40 | 47.1 | 1 |

| Intf In | Dest | Intf Out | Label Out |
|---------|--------|----------|-----------|
| 3 | 47.1.1 | 2 | 1.33 |
| 3 | 47.1 | 1 | 0.50 |



Spajanje MPLS labela

- Paketi koji pripadaju istoj FEC klasi mogu ući na različitim mjestima u MPLS domen, a izaći na istom mjestu.
- LSP putevi ovih paketa će se spajati u MPLS domenu.
- Otuda je dozvoljeno tzv. spajanje labela (*label merging*).
 - Paketi iste FEC klase koji dolaze sa različitih ulaznih interfejsa i/ili sa različitim labelama u LSR ruter će se u slučaju spajanja labela prosleđivati na isti izlaz LSR rutera, pri čemu će nova vrijednost labele (na izlazu) biti identična za sve te pakete.



MPLS odabir ruta

- **Hop-by-hop** rutiranje: koristi se ruta koju određuju dinamički protokoli rutiranja
- **Explicit routing** (ER): LSR na ulazu MPLS domena specificira eksplicitnu rutu za LSP
 - Eksplicitna ruta može biti izračunata unaprijed ili dinamički
 - Pošto LSP put formiran eksplicitnim rutiranjem ne mora da bude najkraći put onda se ne mogu koristiti samo klasični protokoli rutiranja
 - Definisana su proširenja protokola rutiranja (na primer, OSPF-TE) koja omogućavaju razmjenu detaljnijih informacija o mreži čime se omogućava ulaznom LER ruteru da lakše odabere eksplicitnu putanju.
 - Eksplicitno rutiranje je veoma važno u slučaju kada se želi formirati VPN tunel, kao i u slučaju implementacije QoS podrške.
- Uvijek se uspostavlja jednosmjernan LSP put.

DiffServ i MPLS

- MPLS se može koristiti u kombinaciji sa DiffServ-om u cilju pružanja QoS-a.
- LSP putevi se konfiguriraju između svakog *ingress-egress* (IE) para.
 - Za svaki IE par, različit LSP može se koristiti za svaku od saobraćajnih klasa, ili
 - Isti LSP mogu dijeliti sve klase IE para, pri čemu se Exp biti koriste za diferenciranje klase paketa.
- **Skalabilno**: sa povećanjem broja tokova ne povećava se broj LSP puteva.