

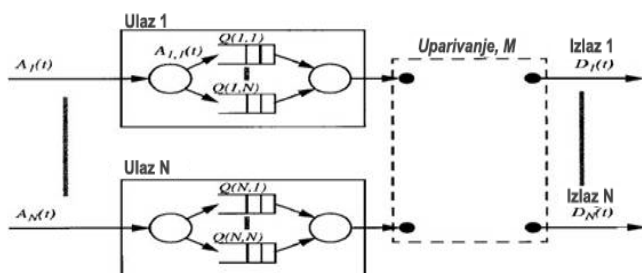
PRILOG ANALIZI PERFORMANSI VOQ KOMUTATORA ČELIJA

Dušan Banović, *Internet Crna Gora*
Igor Radusinović, *Elektrotehnički fakultet u Podgorici*

Sadržaj- U radu je prezentovana analiza performansi VOQ komutatora ćelija u zavisnosti od algoritma za odlučivanje o prosljeđivanju ćelija koji je primjenjen. Pokazano je da iterativni MWM algoritmi daju dobre performanse ali su složeni za implementaciju a da mnogo jednostavniji za implementaciju jednoiterativni MSM algoritmi za realne uslove saobraćaja ne daju idealne performanse.

I. UVOD

Naglo povećanje količine Internet saobraćaja posljednjih godina, uslovlilo je potrebu za komutatorima paketa visokih performansi. To podrazumjeva komutatore paketa koji su dizajnirani tako da za velike brzine prenosa obezbjeđuju veliku propusnost (preko 1Tb/s), prihvatljivo kašnjenje i stabilnost za sve vrste realnog saobraćaja. Naravno, bitno je da taj dizajn bude prihvatljiv sa komercijalnog aspekta. Jedna od najviše korištenih arhitektura kod komutatora paketa visokih performansi je krosbar zbog svoje jednostavnosti i ne-blokirajućih karakteristika. Komutacija paketa fiksne dužine je široko prihvaćena metoda u dizajnu komutatora paketa zbog dobrih performansi koje daje na velikim brzinama. IP paketi promijenjive dužine se segmentiraju na pakete fiksne dužine tzv. ćelije. Komutacija ćelija bazirana na ulaznom baferovanju (input queueing (IQ)) je poželjna za primjenu kod komutatora paketa visokih performansi zbog jednostavnosti implementacije i najmanjih zahtjeva u pogledu memorije i interne brzine komutacionog uređaja. Međutim IQ ima veliko ograničenje u pogledu propusnosti koja je limitirana na 58,6% zbog pojave HOL (Head Of Line) blokiranja [1], [2]. Postoje različite metode za redukciju pojave HOL blokiranja. Virtual Output Queueing (VOQ) solucija omogućava potpunu eliminaciju ove pojave i zato je veoma atraktivna za praktičnu primjenu. VOQ komutator paketa (prikazan na slici 1.) ima na svakom ulazu N redova čekanja, za svaki izlaz poseban red čekanja, zahvaljujući čemu se bez internog ubrzanja komutacionog uređaja eliminiše HOL blokiranje. Prosleđivanje ćelija se obavlja na bazi odgovarajućih algoritama za odlučivanje o uparivanju ulaza i izlaza.



Slika 1. VOQ NxN komutator paketa

Cilj rada je da se uporede performanse (srednje kašnjenje, propusnost i stabilnost) VOQ komutatora ćelija u zavisnosti od algoritma za odlučivanje o prosljeđivanju ćelija koji se primjenjen. Rad je organizovan na sljedeći način. U

dijelu II je data klasifikacija i opisan princip rada najznačajnijih algoritama za odlučivanje o prosljeđivanju ćelija. U dijelu III dato je upoređenje performansi (propusnost, kašnjenje, stabilnost) VOQ komutatora ćelija u zavisnosti od algoritma koji je primjenjen za različite saobraćajne uslove. Konačno, na kraju su dati zaključci i smjernica budućeg rada.

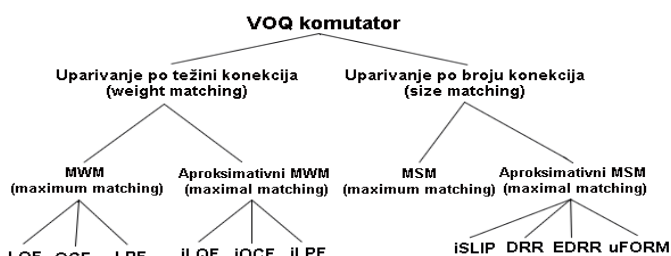
II. KLASIFIKACIJA I PRINCIP RADA ALGORITAMA ZA ODLUČIVANJE

Dosta toga je urađeno u pogledu nalaženja algoritma za odlučivanje o prosljeđivanju ćelija koji će kod VOQ komutatora paketa fiksne dužine dati dobre performanse (propusnost, kašnjenje, vjerovatnoća gubitka) za sve saobraćajne uslove a da pri tom implementacija ne bude previše komplikovana i skupa. Kada govorimo o kompleksnosti pri implementaciji nekog algoritma kod VOQ komutatora ćelija onda moramo uzeti u obzir da sada imamo N^2 redova čekanja na ulazu komutatora, što N puta uvećava kompleksnost u odnosu na IQ komutator. Međutim, danas većina komutatora paketa visokih performansi ima mali broj portova (između 8 i 32) što omogućava da ta kompleksnost bude prihvatljiva pri komercijalnoj implementaciji.

Maximum weight matching (MWM) algoritmi [3], [4] imaju veoma dobre performanse za širok spektar saobraćajnih uslova ali i veliku kompleksnost implementacije $O(N^3 \log N)$ što ih čini najkomplicovanijim za primjenu u praksi. Ako se tome dodaju zahtjevi za multikastom i QoS uslugama onda je ta kompleksnost još veća. MWM metoda se sastoji u uparivanju ulaza i izlaza tako da ukupni težinski faktor svih konekcija bude maksimalan. Taj težinski faktor može biti okupiranost reda ili vrijeme čekanja neke ćelije. Algoritmi koji spadaju u ovu grupu su LQF (Longest Queue First) [4], OCF (Oldest Cell First) [4], i LPF (Longest Port First) [3]. Zbog velike kompleksnosti i spore konvergencije koju imaju MWM algoritmi oni se ne primjenjuju u praktičnim realizacijama. Mnogo jednostavniji za implementaciju su iterativni oblici ovih algoritama iLQF, iOCF i iLPF zbog manje kompleksnosti [3], [4]. Ovi algoritmi koriste metodu uparivanja ulaza i izlaza, koja se sastoji u iterativnom dodavanju konekcija koje nisu postojale u predhodnoj iteraciji, te spadaju u aproksimativne MWM algoritme. Naravno kao posljedica korištenja ove metode i smanjenja kompleksnosti implementacije dobijaju su lošije performanse u odnosu na MWM algoritme. Više iteracija podrazumjeva više konekcija ulaz-izlaz u jednom vremenskom slotu ali i sporiju konvergenciju i težu implementaciju. Ovi algoritmi tek za veći broj iteracija postižu zadovoljavajuće uparivanje te su zbog toga složeni za implementaciju. Postojanje komparatora težinskih faktora uslovljava dodatnu složenost kod svih MWM algoritama.

Idealizovano gledano najefikasniji algoritmi su oni koji nalaze maksimalni broj konekcija ulaz-izlaz tako da

propusnost bude maksimalna. To bi značilo da se dobija najveća propusnost nezavisno od opterećenja na ulazu. Ovi algoritmi se nazivaju jednim imenom maximum size matching (MSM) algoritmi [4] i predstavljaju na neki način specijalni slučaj MWM algoritma za slučaj kada svaka konekcija ulaz-izlaz ima jedinični težinski faktor. Daju dobre performanse kada je u pitanju propusnost ali vrlo sporo konvergiraju i nisu imuni na pojavu zapostavljenja ulaz-izlaz konekcija čak i za proste saobraćajne uslove [4]. Kada je u pitanju kompleksnost implementacije ovih algoritama ona je manja u odnosu na MWM algoritme ali je i dalje velika. Najmanja postignuta kompleksnost MSM algoritma iznosi $O(N^{2.5})$ [5]. Upravo zbog toga u praksi primjenu nalaze aproksimativni MSM algoritmi. Naravno smanjenje kompleksnosti ima za posljedicu degradaciju performansi VOQ komutatora. U ovu grupu algoritama koji nalaze praktičnu primjenu spadaju: iSLIP[3], [4], DRR[6], EDRR[7] i u-FORM[8]. Klasifikacija opisanih algoritama prema metodi uparivanja ulaza i izlaza prikazana je na slici 2.[9]



Slika 2. Klasifikacija algoritama za odlučivanje o proslijedivanju ćelija kod VOQ komutatora

iSLIP je iterativna forma SLIP algoritam [4]. Svaka iteracija se sastoji od tri koraka:

- svaki neupareni ulaz šalje zahtjev svakom izlazu za koji ima ćeliju

- kada neupareni izlaz primi zahtjeve on bira prvi u fiksnom round-robin (RR) rasporedu polazeći od elementa sa najvećim prioritetom. Zatim, izlaz obavještava sve ulaze da li je njihov zahtjev prihvaćen ili ne. Pointer g_{se} pomjera sa elementa najvišeg prioriteta ako je poslata potvrda prihvatanja prihvaćena od strane tog ulaza u trećem koraku iteracije, a u protivnom pointer ostaje na istoj poziciji.

- kada uzlaz primi potvrde prihvatanja (grantove) od izlaza on bira prvi u fiksnom RR rasporedu polazeći od elementa sa najvećim prioritetom. Pointer a_i se pomjera na sljedeći element u fiksnom RR redosledu.

Princip funkcionisanja **iLQF** i **iOCF** algoritma [4] je sličan kao kod opisanog iSLIP algoritama:

- svaki neupareni ulaz šalje zahtjev ka svim izlazima za koje ima ćeliju i obavještava ih koliko ćelija ima za taj izlaz (iLQF) odnosno koliko vremena te ćelije čekaju u redu (iOCF)

- kada neupareni izlaz primi zahtjeve od svih ulaza on bira jedan koji ima najveću težinu i šalje potvrdu prihvatanja ka tom ulazu, a ukoliko imamo više zahtjeva sa istom težinom onda bira jedan slučajnim procesom

- kada ulaz primi sve potvrde prihvatanja on bira jednu koji ima najveću težinu i šalje pozitivan odgovor, a ukoliko imamo više potvrda prihvatanja sa istom težinom onda bira jednu slučajnim procesom i to se naziva proces pozitivnog odgovora.

iLPF algoritam teorijski spada u MWM algoritme ali praktično je MSM algoritam zbog MSM metode uparivanju ulaza i izlaza [3]. Svaka iteracija se sastoji od dva koraka:

- vrši se raspoređivanje svih ulaza i izlaza prema njihovoj opterećenosti

- vrši se uparivanje ulaza i izlaza MSM metodom, pa tako za svaki izlaz krećući se od najopterećenijeg ka najmanje opterećenom i svaki ulaz krećući se od najopterećenijeg ka najmanje opterećenom, ako postoji zahtjev i ako ulaz i izlaz nisu već upareni ostvaruje se njihova konekcija

Zahvaljujući MSM metodi iLPF algoritam se jednostavnije implementira od ostalih MWM algoritama [3].

Kod Dual Round Robin (DRR) algoritma izbor na izlazu potpuno nezavisan od izbora na ulazu [6]. Proces odlučivanja DRR algoritma se odvija u dvije faze:

- svaki ulaz vrši selekciju svih svojih VOQ redova koji nisu prazni i bira jedan na osnovu RR rasporeda, pointer ostaje na istoj poziciji ukoliko na zahtjev ulaza nije poslata potvrda prihvatanja u koraku dva a pomjera se na sljedeću poziciju u RR rasporedu ako i samo ako je zahtjev odobren od strane izlaza (poslata potvrda prihvatanja)

- svaki izlaz vrši odabir svih dospjelih zahtjeva i bira jedan na osnovu RR rasporeda i obavještava sve ulaze da li je njihov zahtjev odobren ili ne, pointer se pošto pošalje potvrdu prihvatanja pomjera na sljedeću poziciju u RR rasporedu a ukoliko nema zahtjeva pointer ostaje na istoj poziciji.

DRR with Exhaustive Service ili skraćeno **EDRR** kreiran je kao modifikacija DRR algoritma sa ciljem da se poboljšaju karakteristike VOQ komutatora za ne-uniformne uslove saobraćaja [7]. To se postiglo prelaskom sa metode limitiranog servisa (kod DRR algoritma) na metodu iscrpljujućeg servisa (kod EDRR algoritma). Detaljni opis oba koraka EDRR algoritma ima sljedeći izgled:

1. Zahtjev: Svaki ulaz pomjera pointer na prvi VOQ red koji nije prazan u fiksnom RR rasporedu i šalje zahtjev izlazu koji odgovara izabranom VOQ redu. Pointer se pomjera na sljedeću poziciju u fiksnom RR rasporedu ako u koraku 2. zahtjev nije odobren od traženog izlaza ili ukoliko je odobrena posljednja ćelija iz tog VOQ reda te on postaje prazan. U slučaju kada je potvrda prihvatanja poslata i postoji još ćelija u izabranom VOQ redu pointer se ne pomjera tj. zadržava svoju poziciju.

2. Potvrda prihvatanja: Ako izlaz primi jedan ili više zahtjeva bira prvi koji se pojavi u fiksnom RR rasporedu. Izlaz obavještava svaki ulaz koji je poslao zahtjev da li je ili nije dobio potvrdu prihvatanja na svoj zahtjev. Pointer ostaje na ulazu sa potvrdom prihvatanja. Takođe ukoliko nema zahtjeva pointer ne mijenja mjesto.

uFORM je noviji algoritam odlučivanja kod VOQ komutatora koji je takođe baziran na RR metodi [8]. uFORM algoritamska šema ima tri faze kao i iSLIP algoritam:

1. Zahtjev- VOQ redovi koji nisu prazni i sa statusom ON šalju zahtjeve ka traženim izlazima, dok VOQ redovi koji nisu prazni i sa statusom OFF šalju zahtjeve jedino ako je ulaz kome oni pripadaju sa statusom OFF.

2. Potvrda prihvatanja- Kada izlaz primi zahtjeve on bira jedan zahtjev poslat od strane VOQ reda koji ima status ON na osnovu RR rasporeda. Ukoliko ne postoji ON status zahtjeva onda izlaz bira OFF status zahtjev na osnovu RR rasporeda.

3. Pozitivan odgovor- Kada ulaz primi sve potvrde prihvatanja on bira jednu potvrdu prihvatanja poslata od ON izlaza na osnovu RR rasporeda i šalje pozitivan odgovor. Ako ne postoji ON potvrda prihvatanja onda ulaz bira jednu OFF potvrdu prihvatanja na osnovu RR rasporeda i šalje pozitivan odgovor. Pointeri na ulazu i izlazu se poslije uspješnog uparivanja ulaza i izlaza pomjeraju na sljedeću poziciju u RR rasporedu.

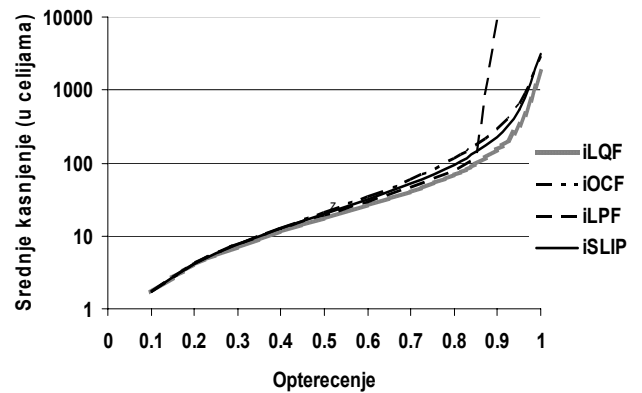
III. UPOREĐENJE PERFORMANSI VOQ KOMUTATORA ČELIJA

Uzimajući u obzir dosad objavljene radove i analize pokazano je da MWM algoritmi [3], [4] a to su LQF, OCF i LPF imaju dobre performanse kada su u pitanju kašnjenje i stabilnost, kao i da obezbjeđuju 100% propusnost i za uniformni i za ne-uniformni saobraćaj. Sa praktičnog aspekta mnogo su interesantniji iterativni oblici ovih algoritama zbog manje kompleksnosti. iLQF, iOCF i iLPF imaju malo lošije performanse za ne-uniformne uslove u pogledu propusnosti i srednjeg kašnjenja u odnosu na MWM algoritme ali su njihove karakteristike bolje u odnosu na karakteristike MSM algoritama (iSLIP, DRR, EDRR, uFORM) ne samo u pogledu propusnosti nego i u pogledu stabilnosti. Iako aproksimativni MWM algoritmi predstavljaju alternativu MWM algoritmima i imaju dobre performanse njihova velika mana je to što zahtijevaju veliki broj iteracija da bi se dobilo zadovoljavajuće uparivanje ulaza i izlaza kao i postojanje komparatora što uzrokuje veliku kompleksnost pri implementaciji kod komutatora visokih performansi. Na slikama 3., 4. i 5. slikovito je prikazano kako se ponašaju VOQ komutatori ćelija 32x32 u zavisnosti od primjenjenog algoritma za odlučivanje za bursty i neuniformne uslove saobraćaja (uzeti su u obzir uslovi kada je opterećenje od ulaza i ka izlazu i 2p, ka izlazu i+1 p dok je ka ostalim izlazima 0, gdje p uzima vrijednosti između 0 i 1 što je prikazano matricom saobraćaja (1)).

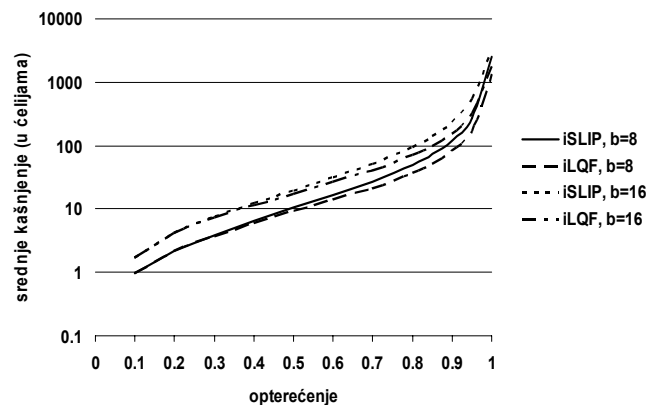
$$\begin{bmatrix} 2p & p & 0 & \dots & 0 \\ 0 & 2p & p & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 2p & p \\ p & 0 & 0 & \dots & 2p \end{bmatrix} \quad (1)$$

Poređenje je vršeno između iterativnih MWM algoritama i iSLIP algoritma koji zahtjevaju određeni broj iteracija da bi se dobile zadovoljavajuće performanse (uzet je slučaj logN iteracija, gdje je N veličina komutatora zato što je pokazano da je to optimalan broj iteracija za iSLIP [3] i da dalje povećanje broja iteracija neznatno utiče na performanse). Za simulaciju su kreirani konfiguracioni fajlovi sa zadatim ulaznim parametrima i korišten je SIM simulator [10] za dobijanje konkretnih vrijednosti izlaznih parametara na bazi kojih su isctavani odgovarajući grafici performansi. Kada je u pitanju bursty saobraćaj prikazano je na slici 3. srednje kašnjenje za iterativne MWM algoritme i iSLIP algoritam za dužinu burst-a 16 i log32 iteracija. Potom, na slici 4. je pokazano kako se srednje kašnjenje ponaša sa promjenom dužine bursta kod iLQF algoritma kao predstavnika MWM algoritma i kod iSLIP kao predstavnika MSM algoritama. Sa slika se vidi da za slučaj log32 iteracija koje predstavljaju optimalan broj iteracija za iSLIP algoritam kada su u pitanju

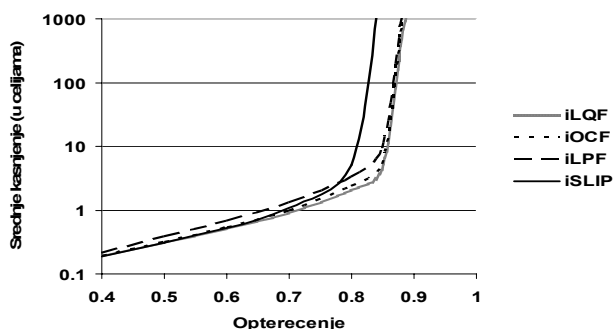
performanse sistema ovaj algoritam ne ponaša mnogo lošije od MWM algoritama. Međutim, dalje povećanje broja iteracija kod MWM dovodi do poboljšanja performansi u odnosu na iSLIP algoritma čije se performanse neznatno mijenjaju. Naravno povećanje broja iteracija na već dovoljno veliku kompleksnost MWM algoritama uslovljava njihovu praktičnu neprimjenjivost. Kada su u pitanju opisani neuniformni uslovi saobraćaja već za logN iteracija se pokazuje superiornost MWM algoritama u odnosu na MSM algoritme što se vidi na slici 5. Upravo zbog toga što povećanje broja iteracija povećava vrijeme potrebno za odlučivanje koje treba da bude reda 10ns mnogo interesantniji za praksu su jedno-iterativni algoritmi kao što su DRR i EDRR. Na slikama 6. i 7. vršeno je upoređenje performansi DRR i EDRR algoritma sa 1-SLIP algoritmom za slučaj bursty i neuniformnog saobraćaja (uzeti su isti uslovi kao i u predhodnom razmatranju) na bazi rezultata dobijenih pomoću SIM simulatora i rezultata iz literature [6], [7]. Sa slika se vidi da kada su u pitanju bursty i neuniformni uslovi da EDRR ima mnogo bolje performanse upravo zbog načina inkrementiranja pointera i politike iscrpljujućeg servisa. Nedostatak EDRR algoritma je nešto lošije ponašanje za slučaj uniformnog saobraćaja. U ovoj analizi akcenat je stavljen na anlizu performansi sistem za bursty i neuniformne uslove saobraćaja pošto oni odgovaraju uslovima realnog saobraćaja. Kada su u pitanju uniformni uslovi svi opisani algoritmi pokazuju dobre performanse s tim da su performanse MWM algoritama bolje u odnosu na MSM algoritama sa aspekta srednjeg kašnjenja i stabilnosti ali su i neuporedivo kompleksniji za implementaciju tako da imaju više teorijski značaj.



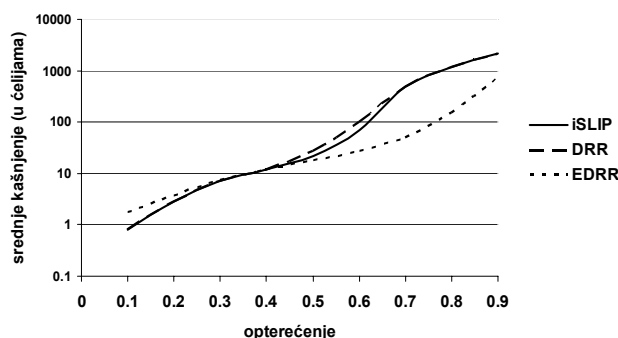
Slika 3. Srednje kašnjenje za slučaj dužine burst-a 16 i log32 iteracija pri primjeni iLQF, iOCF, iLPF i iSLIP algoritma



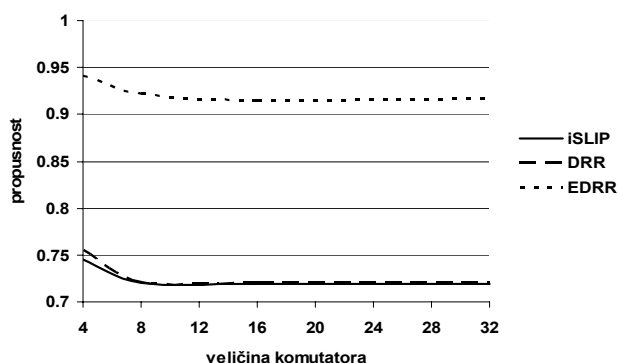
Slika 4. Srednje kašnjenje za slučaj različite dužine burst-a i log32 iteracija pri primjeni iLQF i iSLIP algoritma



Slika 5. Propusnost za neuniformne uslove saobraćaja opisane matricom [1] i log32 iteracije pri primjeni iLQF, iOCF, iLPF i iSLIP algoritma



Slika 6. Srednje kašnjenje za slučaj dužine burst-a 16 kod 1-SLIP, DRR i EDRR algoritma



Slika 7. Propusnost za neuniformne uslove saobraćaja opisane matricom [1] i različite veličine komutatora pri primjeni 1-SLIP, DRR i EDRR algoritma

IV. ZAKLJUČAK I BUDUĆI RAD

Korištenje VOQ komutatora omogućava potpunu eliminaciju pojave HOL blokiranja što omogućava pravazilaženje problema koji izaziva ova pojava u pogledu ograničenja propusnosti na 58,6% kod IQ komutatora [1], [2].

U ovom radu izvršena je analiza performansi VOQ komutatora ćelija na bazi algoritama za odlučivanje o prosljeđivanju ćelija. Izvršena je i klasifikacija i pregled najznačajnijih algoritama koji se primjenjuju kod VOQ komutatora ćelija.

Pravac budućeg rada je simulacija performansi VOQ komutatora ćelija pri primjeni opisanih i drugih algoritama za različite saobraćajne uslove i sveobuhvatna analiza performansi (propusnost, srednje kašnjenje, vjerovatnoća gubitka).

LITERATURA

- [1] M. J. Karol, M. Hluchyj, and S. Morgan, "Input vs. output queuing on a space-division packet switch", 1986
- [2] M. J. Karol, M. Hluchyj, and S. Morgan, "Input versus output queuing on a space-division packet switch," IEEE Trans. on Communications, vol.35, pp. 1347-1356, 1987.
- [3] Adisak Mekittikul, "Scheduling Non-uniform Traffic in High Speed Packet Switches and Routers," PhD Thesis, Stanford University, November 1998..
- [4] N. McKeown, "Scheduling algorithms for input-queued cell switches", Ph.D. Thesis, UC Berkeley, 1995
- [5] J. E. Hopcroft and R. M. Karp, "An algorithm for maximum matching in bipartite graphs," 1973.
- [6] Yihan Li, Shivendra Panwar and H. Jonathan Chao, "On the Performance of a Dual Round-Robin Switch," IEEE INFOCOM 2001
- [7] Yihan Li, Shivendra Panwar and H. Jonathan Chao, "The Dual Round-Robin Matching Switch with Exhaustive Service,"
- [8] Roberto Rojas-Cessa and Chuan-bi Lin, "Captured-Frame Eligibility and Round-robin Matching for Input-queue Packet Switches," IEEE Communication Letters, Sep.2004
- [9] Dušan Banović i Igor Radusinović "Sistematizacija VOQ komutatora ćelija" IT Žabljak, Mart 2005
- [10] <http://klamath.stanford.edu/tools/SIM/>

Abstract – This paper presents performance analysis of VOQ switches depending on scheduling algorithm. It is showed that iterative MWM algorithms provide good performances, yet are too complex from the aspect of implementation, while single-iterative MSM algorithms are easier to be implemented but do not provide ideal performances for real traffic requirements.

APPENDIX TO THE PERFORMANCE ANALYSIS OF VOQ SWITCHES

Dušan Banović, Igor Radusinović