

Analiza CQ komutatora paketa u uslovima saobraćajnog preopterećenja

Milutin Radonjić, *Member, IEEE*, Igor Radusinović, *Member, IEEE*, Ivo Maljević, *Member, IEEE*, Dušan Banović

Sadržaj — U ovom radu je prikazana analiza 2x2 krosbar komutatora paketa sa baferima u ukrsnim tačkama komutacione matrice, kod koga se raspoređivanje paketa vrši na bazi *round robin* algoritma. Analiza je izvršena za slučaj tzv. *non-admissible* saobraćaja, kod kojeg su izlazni portovi preopterećeni. Posmatran je slučaj punog opterećenja dolaznog saobraćaja, pri čemu su izlazni portovi opterećeni sa paketima koji imaju različite vjerovatnoće dolaska. Pored parametara koji se uobičajeno posmatraju u ovakvoj analizi (propusnost i srednje kašnjenje paketa), analizirani su i memorijski zahtjevi za realizaciju bafera, kao i pravična zastupljenost prilikom opsluživanja bafera – tzv. *fernost*. Rezultati su pokazali da čak i kod komutatora sa malim brojem portova treba implementirati jako velike bafere, ako se žele postići zadovoljavajuće performanse u uslovima saobraćajnog preopterećenja.

Ključne reči — Krosbar komutator paketa, srednje kašnjenje, propusnost, *fernost*, raspoređivač.

I. UVOD

SA današnjim stepenom razvoja računarskih komunikacija, javlja se stalna potreba za unapređivanjem performansi ključnih komponenti za prenos podataka, prvenstveno svičeva i ruteru. Već duže vrijeme se u konstrukciji ovih uređaja koristi krosbar arhitektura, zbog svoje jednostavnosti i osobine da nema unutrašnjeg blokiranja saobraćaja [1]. U literaturi su poznata mnoga rješenja na bazi krosbar komutacione matrice. Najčešće je zastupljeno rješenje sa baferima na ulaznim portovima, organizovanim u tzv. virtualne izlazne redove (Virtual Output Queueing – VOQ) [2].

VOQ komutatori predstavljaju veoma dobro rješenje sve dok se linijske kartice (u kojima se nalaze baferi sa paketima koji čekaju na prosleđivanje) nalaze blizu komutacione matrice. Naime, raspoređivač koji odlučuje o paketima koji će biti prosleđeni u sljedećem vremenskom slotu mora imati ažurnu informaciju o zauzetosti bafera. To uzrokuje intenzivnu kontrolnu komunikaciju između bafera i raspoređivača paketa koji je sastavni dio krosbar komutatora. Ukoliko se baferi nalaze blizu komutatora može se smatrati da se ova kontrolna komunikacija obavlja

trenutno, sa stanovišta brzine prenosa paketa. Međutim, u modernim saobraćajnim čvorištima postaje uobičajeno da se baferi nalaze prilično udaljeni od komutacionog polja (distribuirani sistemi) što dovodi do toga da se kontrolna komunikacije više ne može smatrati trenutnom, već utiče na performanse čitavog uređaja.

Ovom se problemu može pristupiti na više načina. Najjednostavnije je povećati trajanje vremenskog slotu predviđenog za prenos jednog paketa kroz komutacionu matricu. Time se rješava problem vremena potrebnog za kontrolnu komunikaciju, ali se direktno degradira brzina prenosa komutacionog uređaja. Drugi pristup, poznat iz dostupne literature, bazira se na implementaciji malih bafera u ukrsnim tačkama komutacione matrice, pored postojećih bafera u linijskim karticama na ulazu [3]. Na ovaj način se ublažava problem sa kontrolnom komunikacijom, ali se ne eliminiše u potpunosti, jer je i dalje potrebno da raspoređivač ima informaciju o zauzetosti ulaznih bafera.

Najbolje rješenje bi bilo ono koje u potpunosti eliminiše kontrolnu komunikaciju. To bi značilo da na ulazu nema bafera, već se oni nalaze samo u ukrsnim tačkama komutacione matrice. Ovo rješenje se ranije nije moglo implementirati uslijed tehnoloških ograničenja. Naime, bilo je teško na istom čipu implementirati i komutacionu matricu sa raspoređivačem i bafere velikih kapaciteta. Međutim, nedavno je pokazano da je, korišćenjem moderne tehnologije, danas moguće napraviti velike bafere u ukrsnim tačkama i da je veći problem obezbijediti dovoljan broj pinova za dovođenje/odvođenje podataka na čipu [4].

Komutator koji ima bafere samo u ukrsnim tačkama komutacione matrice zove se CQ (Crosspoint Queued) komutator (Sl. 1). Paketi koji stižu preko ulaza i i namijenjeni su izlazu j smještaju se u bafer B_{ij} . U svakom vremenskom slotu raspoređivač bira jedan od zauzetih bafera koji pripadaju istom izlazu i prosleđuje njegov vodeći paket na taj izlaz. Svaki izlaz se posmatra nezavisno od ostalih. Izbor bafera koji će biti poslužen vrši se po nekom od unaprijed utvrđenih algoritama [5]. Ukoliko je bafer prilikom dolaska novog paketa popunjen, taj se paket odbacuje.

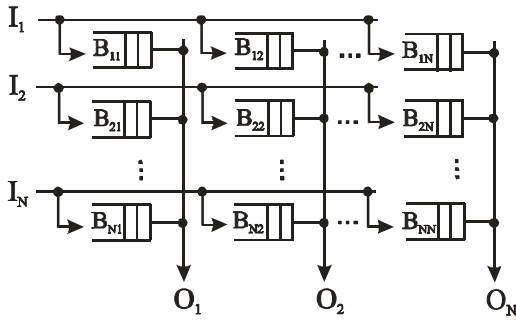
Analiza performansi CQ komutatora paketa je, u ranijim istraživanjima, vršena pod različitim uslovima saobraćaja i različitim konfiguracijama komutatora [6]-[8]. Vršene su simulacije za komutatore sa različitim brojem portova i sa različitim veličinama bafera. Performanse su posmatrane u slučaju uniformnog, nebalansiranog, IBP (Interrupted

Milutin Radonjić, Elektrotehnički fakultet u Podgorici, Crna Gora; (telefon/faks: +382-20-245873; e-mail: m.radonjic@ieee.org).

Igor Radusinović, Elektrotehnički fakultet u Podgorici, Crna Gora; (e-mail: igor@ieee.org).

Ivo Maljević, TELUS Mobility, 200 Consilium Place, Suite 1300 Scarborough, Ontario, Canada, M1H 3J3; (e-mail: ivom@ieee.org).

Dušan Banović, T-com, Moskovska bb, Podgorica, Crna Gora; (e-mail: ban@t-com.me)



Sl. 1. Arhitektura CQ komutatora

Bernoulli Process) i još nekoliko varijanti neuniformnog saobraćaja. Međutim, u svim tim slučajevima radilo se o tzv. *admissible* saobraćaju, što znači da nijedan ulaz odnosno izlaz nije preopterećen saobraćajem. Drugim riječima, statistički gledano, nijedan port nema opterećenje veće od jedinice. U ovom radu posmatrano je ponašanje CQ komutatora paketa u uslovima *non-admissible* saobraćaja, gdje je moguće preopterećenje nekog od portova. Rezultati pokazuju da je potrebno implementirati bafere jako velikog kapaciteta da bi se postigle zadovoljavajuće performanse. Ovo je naročito izraženo u uslovima velikog saobraćajnog preopterećenja nekog od izlaznih portova.

U poglavlju 2, nakon uvodnih napomena, objašnjen je simulacioni model komutatora paketa. Rezultati simulacija i njihova analiza prikazani su u poglavlju 3. Na kraju rada data su zaključna razmatranja.

II. SIMULACIONI MODEL

U ovom radu izvršena je simulacija funkcionisanja 2x2 krosbar komutatora paketa sa baferima u ukrsnim tačkama. Uzeto je da u svakom vremenskom slotu na oba ulaza stiže po jedan paket (ulazno opterećenje $\rho=1$). Vjerovatnoća da je paket sa prvog ulaza namijenjen prvom izlazu označena je sa p_1 , što znači da je vjerovatnoća da je paket sa prvog ulaza namijenjen drugom izlazu jednaka $1-p_1$. Vjerovatnoća da je paket sa drugog ulaza namijenjen prvom izlazu označena je sa p_2 , a da je namijenjen drugom izlazu iznosi $1-p_2$.

Simulacije su vršene na 10 miliona vremenskih slotova za različite vrijednosti parametara p_1 i p_2 u opsegu (0, 1], sa korakom 0,1. Posmatrani su komutatori sa različitim veličinama bafera u ukrsnim tačkama (L): 1, 2, 3, 4, 8, 16, 32, 64, 128, 256 i 512. Pod veličinom bafera se podrazumijeva broj paketa koje je moguće u njemu smjestiti. U svrhu teorijskog razmatranja posmatran je i slučaj kada dužina bafera nije ograničena.

Korišćeno je nekoliko algoritama raspoređivanja, a ovdje će biti predstavljeni rezultati za *round robin* (RR) algoritam, kao izbor koji je najjednostavniji za implementaciju [5], [8].

Uobičajeni parametri za ocjenu performansi komutatora su propusnost i srednje kašnjenje paketa. Propusnost komutatora je definisana kao odnos ukupnog broja prihvaćenih paketa tokom simulacije i maksimalnog mogućeg broja pristiglih paketa (broj portova pomnožen sa brojem vremenskih slotova u simulaciji). Srednje kašnjenje paketa je definisano kao odnos ukupnog

vremena koje su svi prihvaćeni paketi proveli u baferima (izraženo u broju vremenskih slotova) i broja prihvaćenih paketa tokom simulacije.

Pored ovih uobičajenih parametara za evaluaciju performansi komutatora paketa, interesantno je utvrditi kako algoritam raspoređivanja utiče na pravičnu zastupljenost pojedinih portova (bafera) prilikom prenosa paketa. U literaturi se ova karakteristika komutatora naziva fernost (*fairness*). Jedna od najpopularnijih mjera za procjenu fernosti tokova mrežnog saobraćaja je tzv. *Jain's Fairness Index* (*JFI*) [9]. *JFI* se, za slučaj 2x2 komutatora, može definisati na sledeći način:

$$JFI = \frac{\left(\frac{t_1}{p_1} + \frac{t_2}{p_2}\right)^2}{2\left(\left(\frac{t_1}{p_1}\right)^2 + \left(\frac{t_2}{p_2}\right)^2\right)}, \quad (1)$$

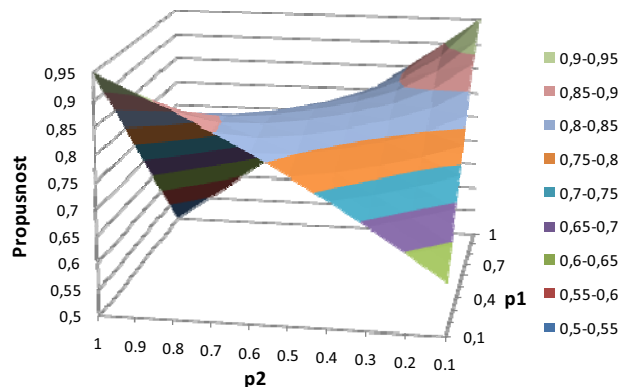
gdje je t_i ($i=1,2$) odnos broja prihvaćenih paketa pristiglih preko ulaza i a namijenjenih prvom izlazu i i broja posmatranih vremenskih slotova (propusnost). Dakle, parametri kod *JFI* se posmatraju gledano u odnosu na jedan od izlaza (u ovom slučaju prvi izlaz).

Parametar *JFI* se može kretati u opsegu [0,1], pri čemu *JFI*=1 znači potpunu fernost između mrežnih tokova, dok *JFI*=0 označava da nema fernosti.

III. REZULTATI SIMULACIJE

A. Propusnost

Na sl. 2 prikazana je propusnost u funkciji vjerovatnoća p_1 i p_2 , za slučaj bafera u koje se može smjestiti samo po jedan paket ($L=1$). Propusnost se kreće od 0,5 (kada je $p_1=p_2=0$, odnosno $p_1=p_2=1$) do 0,95 (kada je $p_1=1$ i $p_2=0,1$, odnosno $p_1=0,1$ i $p_2=1$). Generalno, dijagram je simetričan u odnosu na obje dijagonale kvadranta određenog sa p_1 i p_2 . Najinteresantniji dio dijagrama je onaj koji definišu vrijednosti vjerovatnoća za koje važi da je $p_1+p_2=1$. Taj dio predstavlja granicu tzv. *admissible* saobraćaja i na njemu se najbolje vidi pad propusnosti kao posljedica malih bafera. Najmanja propusnost, u tom slučaju, je kod vrlo bliskih vrijednosti vjerovatnoća. Što se vjerovatnoće više među sobom razlikuju, to je veća i propusnost. Obzirom na način definisanja posmatranih vjerovatnoća, ovakvi rezultati su se mogli i očekivati.



Sl. 2. Propusnost za bafer dužine $L=1$

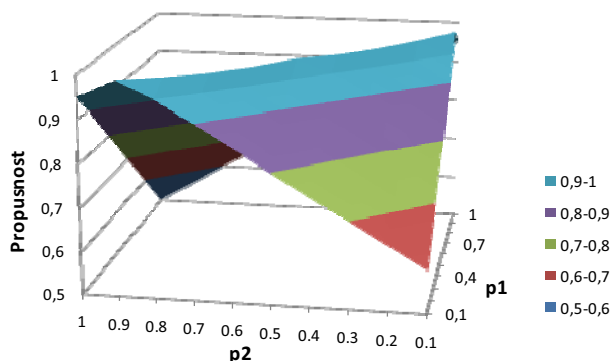
Kada se vjerovatnoće p_1 i p_2 udaljavaju od linije $p_1+p_2=1$, propusnost drastično opada. Ovo je posljedica prirode saobraćaja, jer je tada jedan od izlaznih portova preopterećen.

Sa povećanjem dužine bafera, pokazuje se da propusnost raste u uslovima bliskim *admissible* saobraćaju. Međutim, za saobraćaj koji preopterećuje neki od izlaznih portova, povećanje dužine bafera ne donosi značajno povećanje propusnosti. Kao ilustracija rečenog, na Sl. 3 je prikazana propusnost za slučaj bafera dužine $L=8$. U zoni granice *admissible* saobraćaja ovaj komutator ima gotovo istu propusnost za sve vrijednosti vjerovatnoća p_1 i p_2 , koja iznosi približno 0,98. Sa daljim povećanjem dužine bafera, oblik dijagrama ostaje isti, a mijenja se samo maksimalna propusnost koja postaje bliža jedinici.

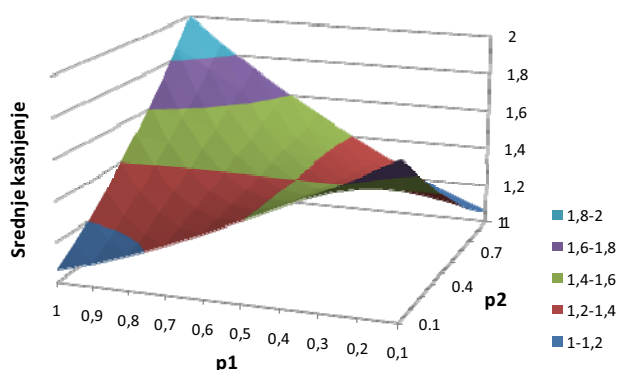
B. Srednje kašnjenje

Dijagram srednjeg kašnjenja paketa u funkciji vjerovatnoća p_1 i p_2 , za slučaj bafera u koje se može smjestiti samo po jedan paket ($L=1$), prikazan je na Sl. 4. Zbog jako male dužine bafera i srednje kašnjenje ima jako male vrijednosti. Međutim, može se uočiti trend koji važi i za bafere većih dužina. U slučaju kad je neki izlazni port preopterećen, tada i srednje kašnjenje ima veću vrijednost. Veliko srednje kašnjenje se poklapa sa niskom propusnošću komutatora. U opsegu gdje je $p_1+p_2=1$ paketi se najmanje zadržavaju u baferima, a to se poklapa sa velikom propusnošću.

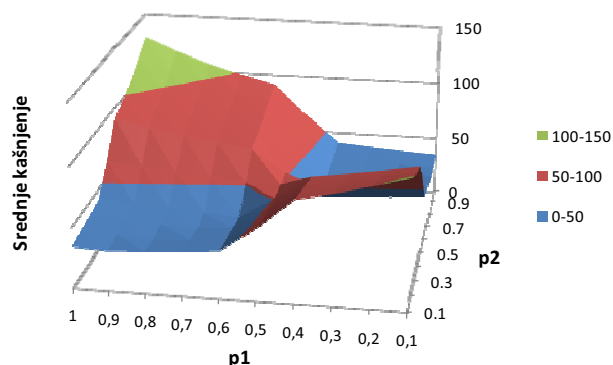
Povećanjem veličine bafera povećava se i srednje kašnjenje paketa. Oblik dijagrama se značajnije ne mijenja, što znači da i dalje važe raniji zaključci. Kao ilustracija dijagrama srednjeg kašnjenja za bafere većeg kapaciteta, na Sl. 5 je prikazan dijagram za komutator sa baferima dužine $L=64$.



Sl. 3. Propusnost za bafer dužine $L=8$



Sl. 4. Srednje kašnjenje za bafer dužine $L=1$



Sl. 5. Srednje kašnjenje za bafer dužine $L=64$

C. Maksimalne potrebne dužine bafera

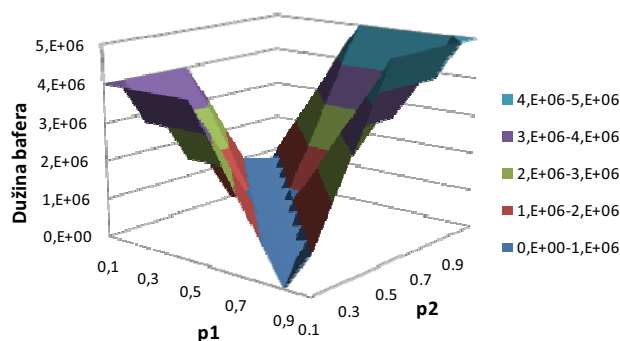
Tokom simulacija je posmatran i slučaj kada baferi u ukrasnim tačkama nisu imali ograničenje u dužini. U ovoj situaciji je interesantno vidjeti koliko će biti maksimalno zauzeće bafera tokom simulacije. Drugim riječima, posmatra se kolika je potrebna dužina bafera pri kojoj neće biti odbačenih paketa. Rezultati su prikazani na sl. 6.

Kao što se iz rezultata za propusnost moglo očekivati, najmanji memorijski zahtjevi su u zoni bliskoj *admissible* saobraćaju. Za slučajeve kad je $p_1+p_2=1$ potrebni su baferi čija se dužina kreće do 3000 paketa. Sa udaljavanjem od te zone dolazi do drastičnog povećanja zahtjeva za memorijskim kapacitetom. Već za slučaj kada je $p_1+p_2=1,1$ potrebni su baferi čiji je kapacitet reda veličine milion paketa. U slučaju jako velikog opterećenja nekog izlaznog porta (p_1+p_2 jako udaljeno od jedinice), značajno rastu i memorijski zahtjevi za baferevanjem paketa. U prikazanim simulacijama dolazilo se i do zauzeća bafera sa 5 miliona paketa.

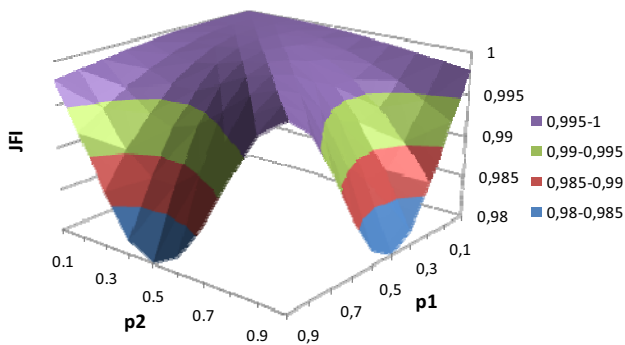
Jasno je da baferi takvih dimenzija nisu pogodni za implementaciju. Međutim, informacija o maksimalnom zauzeću bafera može da posluži prilikom procjene koliku dužinu bafera odabrati prilikom implementacije, za pojedine uslove saobraćaja.

D. Fernost

Na sl. 7 su prikazane vrijednosti parametra *JFI* u funkciji vjerovatnoća p_1 i p_2 , za slučaj komutatora sa baferima dužine $L=1$. Round robin algoritam pokazuje potpuno ravnopravno zastupanje portova (potpuna fernost) u oblasti gdje je $p_1=p_2$. Ovakav rezultat se i mogao očekivati obzirom na činjenicu da su kod takvog saobraćaja izlazni portovi ravnomjerno opterećeni. Pošto RR algoritam naizmenično opslužuje zauzete bafere, jasno je da su oni ravnopravno zastupljeni i što se tiče



Sl. 6. Maksimalna zauzeća bafera tokom simulacije



Sl. 7. Dijagram fernosti za slučaj bafera dužine $L=1$

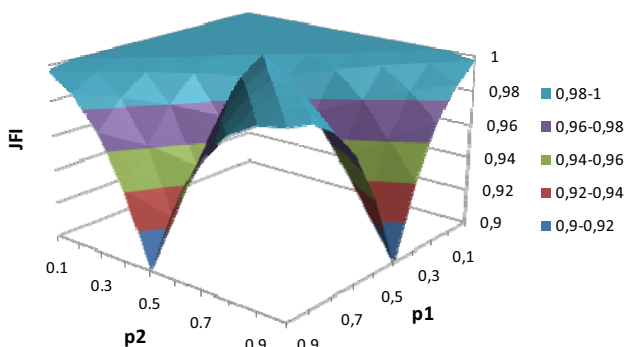
opterećenja i što se tiče opsluživanja. Time je i fernost, koja se računa na bazi propusnosti, potpuna.

Zato je interesantnije posmatrati fernost u opsegu u kojem je posmatrani izlazni port neravnomjerno zastupljen od strane ulaznih portova. Najmanja fernost je u slučaju kada je jedna od vjerovatnoća jednaka jedinici a druga jednaka 0,5. Tada će se jedan bafer puniti u svakom vremenskom slotu a drugi, statistički gledano, u svakom drugom. Pošto se kod round robin algoritma baferi naizmjenično opslužuju, to je jasno da će (sa stanovišta propusnosti) bafer koji se stalno puni biti neravnomjerno zastupljen. Smanjivanjem vjerovatnoća p_1 i p_2 , kao i prilikom približavanja njihovih vrijednosti, dolazi se do povećanja fernosti.

Dalja analiza vršena je za slučaj bafera većih dužina. Kod svih posmatranih bafera dijagrami imaju veoma sličan oblik i razlikuju se samo u konkretnim vrijednostima parametra JFI , za pojedine vrijednosti vjerovatnoća p_1 i p_2 . Kod svih posmatranih komutatora najmanji JFI je za slučaj kad je jedna vjerovatnoća jednaka jedinici a druga jednaka 0,5. Sa povećanjem dužine bafera smanjuje se vrijednost minimalnog JFI , da bi kod bafera dužine $L=512$ imala vrijednost od približno 0,9 (Sl. 8). Takođe, sa povećanjem dužine bafera povećava se opseg vjerovatnoća kod kojih se postiže potpuna fernost ($JFI=1$). Ovo je posljedica činjenice da se kod dužih bafera povećava propusnost, na osnovu koje je JFI definisan. Kao što se sa Sl. 8 može vidjeti, u slučaju bafera dužine $L=512$ opseg vjerovatnoća za koje komutator pokazuje maksimalnu fernost obuhvata vrijednosti za koje važi da je $p_1+p_2 \leq 1$.

IV. ZAKLJUČAK

Rezultati prikazani u ovom radu pokazuju da se u uslovima visokog intenziteta dolaznog saobraćaja javlja



Sl. 8. Dijagram fernosti za slučaj bafera dužine $L=512$

potreba za jako velikim baferima čak i kod komutatora sa malim brojem portova. Najveća propusnost se dobija u zoni bliskoj tzv. *admissible* saobraćaju, a udaljavanjem od te zone se propusnost jako brzo smanjuje. Pored toga, prilikom preopterećenja nekog od izlaznih portova, javlja se i veliko povećavanje srednjeg kašnjenja paketa.

Round robin algoritam pokazuje veliku fernost kod opsluživanja bafera u slučaju ravnomjernog opterećenja izlaznih portova. Sa povećanjem veličine bafera povećava se i oblast u kojoj je fernost potpuna, ali se zadržava u granicama vjerovatnoća koje zadovoljavaju uslov da je $p_1+p_2 \leq 1$.

Dalja istraživanja se planiraju u domenu analize performansi za druge algoritme raspoređivanja, kao i za komutatore sa većim brojem portova.

LITERATURA

- [1] A. Mekikittikul, N. McKeown, "A Practical Scheduling Algorithm to Achieve 100% Throughput in Input-queued Switches", in *Proc. of INFOCOM '98*, San Francisco, USA, 1998, pp. 792-799.
- [2] N. H. Liu, K. L. Yeung, D. C. W. Pao, "Scheduling Algorithms for Input-queued Switches with Virtual Output Queueing", in *Proc. of ICC'2001*, Helsinki, Finland, June 2001, pp. 2038 - 2042.
- [3] R. Rojas-Cessa, E. Oki, Z. Jing, H. J. Chao, "CIXB-1: Combined Input-One-cell-crosspoint Buffered Switch", in *Proc. of IEEE HPSR '01*, Dallas, Texas, USA, 2001, pp. 324-329.
- [4] Y. Kanizo, D. Hay, I. Keslassy, "The Crosspoint-queued Switch", in *Proc. of INFOCOM '09*, Rio de Janeiro, Brazil, 2009, pp. 729-737.
- [5] I. Radusinović, M. Pejanović, Z. Petrović, "Impact of Scheduling Algorithms of Performances of Buffered Crossbar Switch Fabric", *IEEE ICC 2002*, New York, USA, 2002, pp. 2416-2420.
- [6] M. Radonjic, I. Radusinovic, "Buffer Length Impact to Crosspoint Queued Crossbar Switch Performance", in *Proc. of the 15th IEEE MELECON*, Valleta, Malta, 2010, pp. 119-124.
- [7] J. Cvorovic, I. Radusinovic, M. Radonjic, "Buffering in Crosspoint-queued Switch", in *Proc. of the 17th TELFOR 2009*, Belgrade, Serbia, 2009, pp.198-201.
- [8] M. Radonjic, I. Radusinovic, "Impact of scheduling algorithms on performance of crosspoint queued switch", *Annals of Telecommunications*, to be published.
- [9] S. Bhatti, M. Bateman, D. Rehumanathan, T. Henderson, G. Bigwood, D. Miras, "Revisiting inter-flow fairness", in *Proc. of BROADNETS 2008*, London, 2008, pp. 585-592.

ABSTRACT

In this paper, performance analysis of 2×2 crosspoint queued crossbar switch, based on the round robin scheduling algorithm, is presented. The analysis is performed for non-admissible traffic, where the output ports are overloaded. A case of fully loaded incoming traffic, where output ports receive packets with a range of packet arrival probabilities, is observed. In addition to the parameters normally observed in such an analysis (throughput and average cell latency), memory requirements for buffer implementation are also analyzed, as well as inter-flow fairness. The results show that, even for switches with small number of ports, very long buffers are required in order to achieve good performance under the traffic overload.

CQ SWITCH ANALYSIS UNDER THE TRAFFIC OVERLOAD

Milutin Radonjić, Igor Radusinović, Ivo Maljević,
Dušan Banović