

**ELEKTROTEHNIČKI FAKULTET UNIVERZITETA U BEOGRADU**



## **BIRKHOFF VON NEUMANN KOMUTATORI**

–Diplomski rad –

Kandidat:

Vladimir Blagojević 2009/331

Mentor:

doc. dr Zoran Čiča

Beograd, Oktobar 2015.

# SADRŽAJ

<b>SADRŽAJ</b> .....	<b>2</b>
<b>1. UVOD</b> .....	<b>3</b>
<b>2. KOMUTATORI, OPŠTE KARAKTERISTIKE I PODELE</b> .....	<b>4</b>
2.1. OPŠTE KARAKTERISTIKE PAKETSKIH KOMUTATORA .....	4
2.2. PODELA PAKETSKIH KOMUTATORA .....	4
2.3. OSOBINE KOMUTATORA BITNE ZA ANALIZU BVN KOMUTATORA .....	5
2.3.1. <i>Baferi</i> .....	5
2.3.2. <i>Krosbar komutatori</i> .....	6
<b>3. BVN KOMUTATOR</b> .....	<b>8</b>
3.1. BVN KOMUTATOR, OSNOVNA VARIJANTA .....	8
3.2. BVN KOMUTATOR SA BALANSIRANJEM SAOBRAĆAJA .....	10
3.3. PROBLEMI KOJI SE JAVLJAJU KOD LB-BVN KOMUTATORA .....	11
<b>4. VARIJANTE LB-BVN KOMUTATORA KOJE REŠAVAJU PROBLEM PAKETA VAN REDOSLEDA</b> .....	<b>14</b>
4.1. FIRST COME FIRST SERVED - FCFS .....	14
4.2. EARLIEST DEADLINE FIRST – EDF .....	16
4.3. EARLIEST DEADLINE FIRST USING THREE-DIMENSIONAL QUEUES, EDF-3DQ .....	16
4.4. FULL FRAMES FIRST – FFF .....	17
4.5. FULL ORDERED FRAMES FIRST – FOFF .....	19
4.6. MAILBOX KOMUTATOR .....	20
4.7. BYTE-FOCAL KOMUTATOR .....	23
<b>5. ZAKLJUČAK</b> .....	<b>25</b>
<b>LITERATURA</b> .....	<b>26</b>

# 1. UVOD

Komutacija korisničkih signala se može vršiti na dva načina, to su komutacija kola i komutacija paketa. Komutacija kola je starija i zasniva se na principu zauzimanja resursa mreže između korisnika koji trebaju da uspostave međusobnu komunikaciju. Korisnici, tj. izvorište i odredište, zauzimaju te resurse dokle god traje veza između njih. Resursi mreže se oslobađaju tek kad izvorište saopšti mreži da je transakcija završena. Prednost ove tehnike je mogućnost garantovanja kvaliteta servisa korisniku jer se parametri veze definišu prilikom uspostave veze i oni važe dokle god ta veza traje. Najveća mana ove tehnike je slaba iskorišćenost kapaciteta mreže. To je u prošlosti bilo prihvatljivo jer su tad dominirale telefonske mreže. Međutim, razvojem računara i nastankom računarskih mreža princip komutacije kola jednostavno postaje neefikasan. To je uočeno još šezdesetih godina prošlog veka i predložena je tehnika komutacije paketa. Kod ove tehnike podaci se na izvorištu dele na pakete različitih ili istih dužina. Ti paketi se zatim šalju ka odredištu pri čemu mogu prolaziti različitim putevima kroz mrežu i samim tim resursi mreže su mnogo bolje iskorišćeni. Paketi se u suštini sastoje iz dva dela, to su korisni deo (podaci koji se šalju) i zaglavlje (u njemu su sadržane informacije o izvorištu, odredištu, dužini paketa...). Paketi na svom putu prolaze kroz mrežne čvorove i prilikom ulaska u mrežni čvor bira se izlaz preko kog treba dalje poslati paket. Očigledan problem koji se ovde javlja je kad više paketa trebaju da napuste čvor preko istog izlaza. Tada može doći do gubljenja paketa ili do zagušenja u mreži jer svaki paket čeka svoj red da bi napustio čvor. Može se desiti da paketi izlaze iz čvora drugačijim redosledom nego što su ušli. Takođe, često se dešava da usled različitih putanja između izvorišta i odredišta, paketi opet stignu na odredište van redosleda u kom su poslani sa izvorišta. Dakle, kod tehnike komutacije paketa postoji velika iskorišćenost mrežnih resursa ali se ne može garantovati kvalitet servisa jer on zavisi od opterećenja mreže, tj. od mrežnih čvorova.

Komutator je mrežni uređaj koji upravlja protokom podataka između pojedinih korisnika u računarskoj mreži. Komutator ima više ulaza i više izlaza. Pakete koji mu stignu na ulaze prosleđuje samo odgovarajućim izlazima na osnovu zaglavlja pristiglih paketa i na osnovu sopstvene arhitekture odnosno načina opsluživanja paketa.

Internet, kao najveća svetska računarska mreža, nastavlja da raste ubrzano. Da bi se održao taj trend rasta nastala su mnoga istraživanja u vezi arhitektura komutatora velikih brzina i velikih kapaciteta koje ostvaruju bolje karakteristike od arhitektura koje se trenutno upotrebljavaju.

Ovaj rad se bavi jednom od tih arhitektura i njenim brojnim varijantama. To su Birkhoff von Neumann (BvN) komutatori. Ostatak rada je organizovan u nekoliko celina, tj. poglavlja. Drugo poglavlje se bavi karakteristikama i podelama samih komutatora. Treće poglavlje se bavi principom rada osnovne verzije BvN komutatora. Pored toga, analizira se i princip rada naprednije varijante ovog komutatora gde se uvodi balansiranje saobraćaja, kao i problemi koji se javljaju u ovim varijantama. Četvrto poglavlje se bavi principima rada naknadnih, složenijih verzija BvN komutatora. Takođe, analiziraju se i problemi koji se kod njih javljaju kao i njihova rešenja. Poslednje, 6. poglavlje, je zaključak u kom je sumiran rad.

## 2. KOMUTATORI, OPŠTE KARAKTERISTIKE I PODELE

Da bi se u potpunosti razumela ideja BvN komutatora, kao i njegovih naknadnih (složenijih) varijanti, u ovom poglavlju su opisane opšte osobine paketskih komutatora koje će pomoći u analizi. Takođe, izvršena je i osnovna podela komutatora da bi se shvatilo gde se, po principu rada i arhitekturi, BvN komutator nalazi.

### 2.1. Opšte karakteristike paketskih komutatora

Neke od najvažnijih karakteristika paketskih komutatora su[1]:

- Podrška multikast saobraćaja. Paket koji se prosleđuje kroz komutator može biti unicast (jedan ulaz na jedan izlaz) i multikast (jedan ulaz na proizvoljan broj izlaza).
- Blokada. Razlikujemo neblokirajuće i blokirajuće komutatore. Neblokirajući je onaj komutator kod kog se može uspostaviti veza između bilo kog ulaza sa bilo kojim izlazom bez obzira na već uspostavljene konekcije, u suprotnom je blokirajući.
- Skalabilnost. Određuje koliko je jednostavno proširivati kapacitet komutatora. Komutator je skalabilniji ako je moguće lakše proširiti njegov kapacitet.
- Ubrzanje komutatora. Predstavlja na koliko puta većoj brzini radi komutator u odnosu na ulazne/izlazne linkove.
- Propusnost komutatora. Određuje efikasnost, tj. kapacitet komutatora. Cilj je da bude 100%.
- Cena. Zavisi od kompleksnosti izrade komutatora. Ako se upotrebljava više jednostavnijih elemenata komutator će biti ekonomičniji.
- Kompleksnost. Zavisi od toga koliko je lako implementirati dati komutator, kao i složenost upravljanja samog komutatora.
- Tip paketa koji se komutira. Paketi mogu biti promenljive ili fiksne dužine. Paketi fiksne dužine se zovu ćelije, i dokazano je da komutatori efikasnije rade sa njima.
- Baferi. Mogu da se postave na ulaznim portovima, izlaznim portovima, ili u paketskom komutatoru. Dozvoljene su sve moguće kombinacije a to gde su pozicionirani zavisi od same arhitekture mrežnog čvora i tipa paketskog komutatora koji se koristi.

### 2.2. Podela paketskih komutatora

Osnovna podela paketskih komutatora je na[1]:

- Vremenske komutatore
- Prostorne komutatore

Vremenski komutatori se dalje dele na:

- Komutatore sa zajedničkim medijumom za prenos
- Komutatore sa zajedničkom memorijom

Prostorni komutatori se dele na:

- Komutatore sa jednostrukim putanjama
- Komutatore sa višestrukim putanjama

Ako postoji samo jedan put od ulaza  $i$  do izlaza  $j$ , u pitanju je komutator sa jednostrukom putanjom. Ako postoji više putanja u pitanju je komutator sa višestrukim putanjama.

Komutatori sa jednostrukim putanjama se dalje dele na:

- Krosbar komutatore
- Potpuno povezane komutatore
- Banyan komutatore

Komutatori sa višestrukim putanjama se dalje dele na:

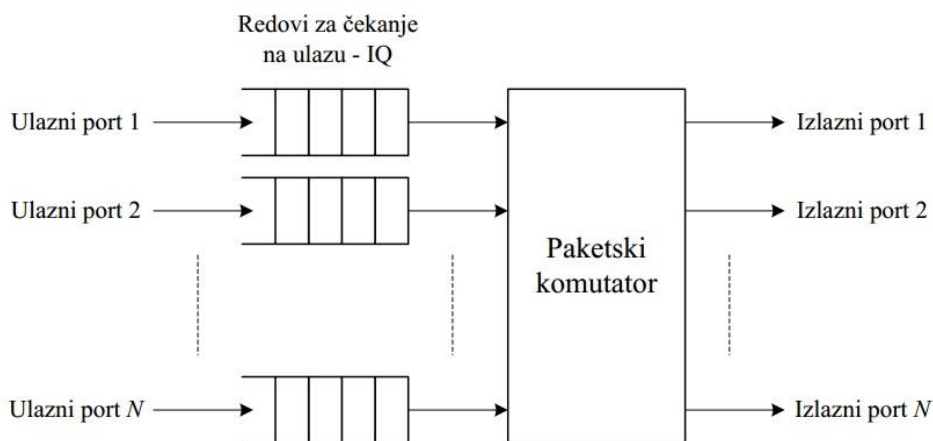
- Umnožene Banyan komutatore
- Klosove komutatore
- Višeravanske komutatore
- Recirkulacione komutatore

## 2.3. Osobine komutatora bitne za analizu BvN komutatora

Osnovna varijanta BvN komutatora se sastoji od jednog krosbar komutatora i bafera na njegovim ulazima. Da bi se bolje razumela osnovna varijanta BvN komutatora, kao i da bi se lakše analizirale njegove naprednije varijante, slede detaljnija objašnjenja nekih opštih osobina komutatora.

### 2.3.1. Baferi

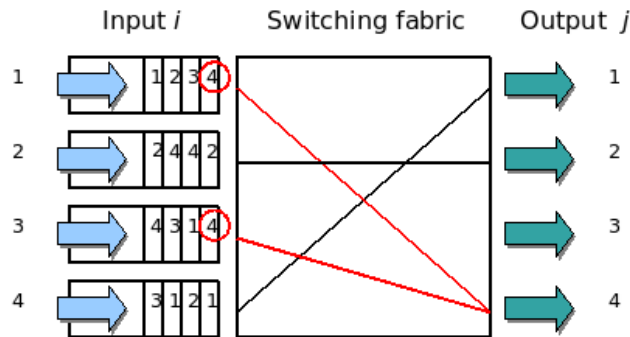
Pošto je paketski saobraćaj po prirodi sporadičan (*bursty*) povremeno dolazi do preopterećenja izlaznih portova. To se dešava jer dva paketa mogu biti namenjena istom izlaznom portu, ali ne mogu fizički da ga napuste istovremeno. Da ne bi došlo do nepotrebnih gubitaka paketa potrebna je upotreba bafera. Gde će baferi biti pozicionirani zavisi od same arhitekture paketskog komutatora koji se koristi. BvN komutator koristi baferne na ulaznim portovima (IQ – *Input Queuing*).



Slika 2.3.1.1. Baferi na ulaznim portovima[1]

U ovom slučaju paketi se čuvaju u baferima koji su na ulaznim portovima. Koriste se raspoređivači (*scheduler*) koji za svaki slot vrše uparivanje ulaznih i izlaznih portova. Par  $(i, j)$  označava da će sa ulaznog porta  $i$  paket biti poslat kroz komutator na izlazni port  $j$ . Za svaki ulazni port postoji po jedan bafer. Za slučaj da ne postoji ubrzanje, vrši se jedan upis i jedno čitanje paketa. U suprotnom, broj upisa/čitanja zavisi od ubrzanja  $s$  samog komutatora. Najjednostavnija realizacija bafera na ulaznom portu se ostvaruje upotrebom FIFO (*First In First Out*) bafera, ali njegova upotreba

dovodi do tzv. HOL (*Head Of Line*) blokade. Naime, u najgoroj situaciji, na vrhu FIFO bafera se na svim ulaznim portovima može naći paket koji je namenjen istom izlaznom portu. To znači da će za neki ulazni port trebati da se proslede paketi svih ostalih ulaznih portova, pre nego što se prosledi i njegov paket.

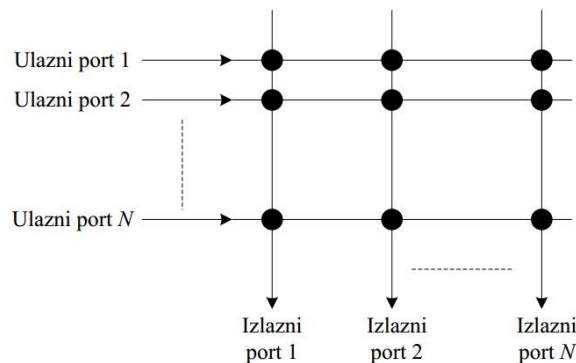


Slika 2.3.1.2. Ilustracija HOL blokade[6]

Usled HOL blokade može doći do značajnog obaranja propusnosti paketskog komutatora, tako da se FIFO baferi više ne koriste. Bafer sa VOQ (*Virtual Output Queuing*) redovima za čekanje je nešto komplikovaniji ali rešava problem HOL blokade, jer se paketi nalaze u listama (virtuelni redovi za čekanje) pa se paket iz bilo koje liste može proslediti. Mana koja se javlja kod bafera sa VOQ redovima je da se komplikuju algoritmi za raspoređivanje. Kod FIFO bafera algoritmi za raspoređivanje su radili sa  $N$  podataka, dok kod VOQ bafera rade sa  $N^2$  podataka.

### 2.3.2. Krosbar komutatori

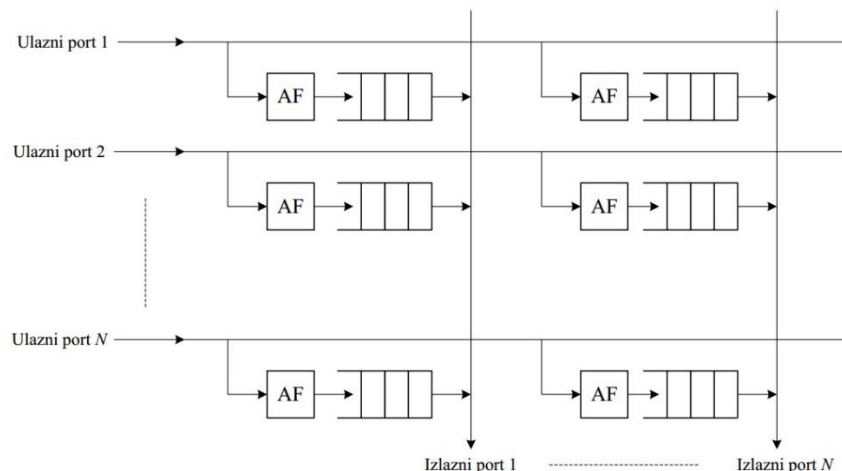
Ovaj komutator predstavlja rešetkastu strukturu gde svaki red (horizontalna linija) odgovara jednom ulaznom portu, a svaka kolona (vertikalna linija) odgovara jednom izlaznom portu. U preseku linija se nalaze prekidači koji kontrolišu rad krosbar komutatora.



Slika 2.3.2.1. Krosbar komutator[1]

Prekidači su u startu otvoreni (*cross* stanje). Ako se želi spojiti ulaz  $i$  sa izlazom  $j$ , tada će se zatvoriti prekidač u preseku horizontalne linije  $i$  i vertikalne linije  $j$ . Stanje zatvorenog prekidača se naziva terminom *bar* stanje. Da ne bi došlo do grešaka u prosleđivanju paketa u jednoj koloni (ili vrsti) ne sme biti više zatvorenih prekidača. Krosbar komutator ima osobinu samorutiranja, odnosno pored mogućnosti da se konfigurira spoljašnjom kontrolom ima sposobnost da se samostalno konfigurira. To se vrši tako što se paketu koji se prosleđuje doda interno zaglavlje koje sadrži identifikaciju izlaznog porta na koji paket treba da se prosledi. Svaka tačka ukrštanja ima jednostavnu

logiku kojom vrši otvaranje ili zatvaranje prekidača na osnovu vrednosti zaglavlja paketa, čime se vrši samorutiranje paketa kroz krosbar komutator.



**Slika 2.3.2.2. Krosbar komutator sa baferima u ukrsnim tačkama[1]**

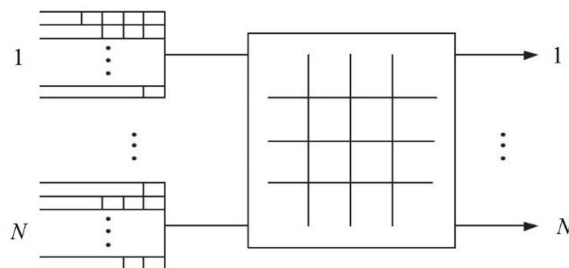
U varijanti kada se u krosbar komutatoru koriste baferi, oni se stavljaju u unakrsne tačke kao što je prikazano na slici 2.3.2.2. Ispred svakog bafera se postavlja adresni filter koji određuje koji će se paket upisati u bafer a koji ne. Zahvaljujući tome u ovom slučaju je multikast saobraćaj podržan jer je moguć upis u više bafera odjednom (na istoj horizontalnoj liniji) za paket sa jednog ulaznog porta.

Krosbar komutator je neblokirajući pa mu nije potrebno unutrašnje ubrzanje za normalno funkcionisanje. Ovaj komutator nije skalabilan jer mu broj prekidača raste sa  $N^2$ , pa se najčešće koristi za manje i srednje veličine komutatora. Takođe, krosbar komutator se može koristiti kao gradivna jedinica u višestepenim komutatorima. Ova njegova osobina je primenjena kod višestepenih BvN komutatora, što će se opisati u narednim poglavljima.

# 3. BvN KOMUTATOR

## 3.1. BvN komutator, osnovna varijanta

BvN komutator je komutator sa baferima na ulazu. Ovaj komutator koristi VOQ (*Virtual Output Queuing*) tehniku da reši HOL (*Head Of Line*) problem blokiranja. Princip koji stoji iza rešavanja problema nadmetanja na izlaznim portovima u BvN komutatoru je da se koristi kapacitetska dekompozicija opisana od strane Birkhoff-a i von Neumann-a, za uređivanje konekcionih paterna (načina obrazovanja veza).



Slika 3.1.1. BvN komutator[3]

Neka je  $\mathbf{r} = [r_{i,j}]$  matrica toka saobraćaja a njen koeficijent  $r_{i,j}$  protok izdvojen za saobraćaj od ulaza  $i$  do izlaza  $j$ , za komutator dimenzija  $N \times N$  sa ulaznim baferom. Neka su zadovoljeni sledeći uslovi[2]:

$$\sum_{i=1}^N r_{i,j} \leq 1, \quad j = 1, 2, \dots, N, \quad (3.1.1)$$

i

$$\sum_{j=1}^N r_{i,j} \leq 1, \quad i = 1, 2, \dots, N. \quad (3.1.2)$$

Onda, postoji skup pozitivnih brojeva  $\Phi_k$  i permutacionih matrica  $\mathbf{P}_k$ ,  $k = 1, 2, \dots, K$ , za neko  $K \leq N^2 - 2N + 2$  koje zadovoljava uslove:

$$\mathbf{r} \leq \sum_{k=1}^K \Phi_k \mathbf{P}_k, \quad (3.1.3)$$

i

$$\sum_{k=1}^K \Phi_k = 1. \quad (3.1.4)$$

Kada je takva dekompozicija ostvarena, onda je samo potrebno obrazovati konekcionu matricu  $\mathbf{P}_k$  proporcionalnu njegovoj težini  $\Phi_k$ . Na primer[3], 4x4 komutator sa ulaznim baferima čija je matrica saobraćaja:

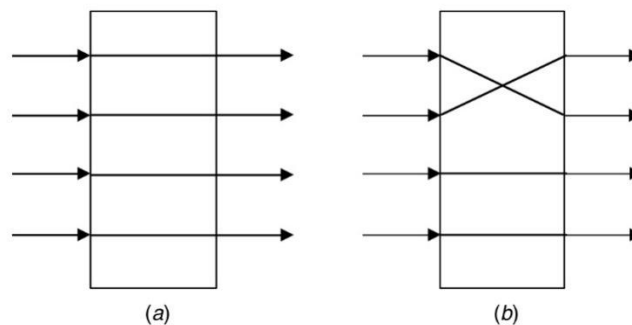


$$\begin{pmatrix} 1 & 2 & 0 & 0 \\ \frac{1}{3} & \frac{2}{3} & 0 & 0 \\ \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Upotrebom BvN dekompozicije, dolazimo do dve permutacione matrice  $\mathbf{P}_1$  i  $\mathbf{P}_2$ , čije su težine  $\Phi_1$  i  $\Phi_2$ :

$$\frac{1}{3} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \frac{2}{3} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Sa ove dve permutacione matrice lako je organizovati komunikacione paterne na osnovu njihovih težina. Konekcionni patern za dati primer je prikazan na slici 3.1.2. Patern (a) se koristi u 1/3 konfiguraciji komutatora, dok se patern (b) koristi u 2/3 konfiguraciji komutatora.



**Slika 3.1.2. Konekcionni patern komutatora iz primera. (a) Konekcionni patern sa težinom 1/3; (b) konekcionni patern sa težinom 2/3. [2]**

Ako je izdvojeni propusni opseg za saobraćaj veći od pristižućeg saobraćaja, pokazano je da se za svaki ulazno – izlazni par može postići propusnost od 100%, u slučaju BvN komutatora sa ulaznim baferima i to bez potrebe za unutrašnjim ubrzanjem ili frejmingom. Ali, složenost BvN algoritma dekompozicije može da dostigne i  $O(N^{4.5})$ , a broj permutacionih matrica koje mogu nastati iz BvN algoritma dekompozicije je  $O(N^2)$ , što nije skalabilno za komutatore sa velikim brojem ulaznih i izlaznih portova. Prema tome, BvN komutator sa ulaznim baferima ima poželjne sistemske performanse, ali mu je mana velika složenost i prema tome loša skalabilnost. Ipak, ako bi razmotrili specijalan slučaj kada bi na sve ulaze komutatora doveli saobraćaj sa uniformnom raspodelom dobilo bi se sledeće:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 1 & 1 & 1 & 1 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 1 & 1 & 1 & 1 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 1 & 1 & 1 & 1 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix},$$

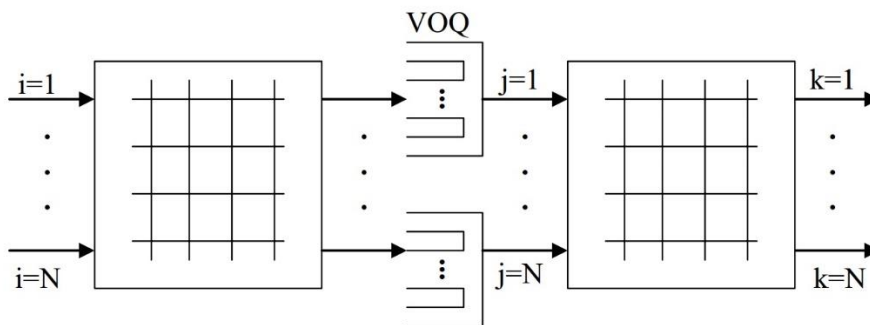
a permutacione matrice i njihove odgovarajuće težine za ovaj komutator sa balansom saobraćaja su:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Primećuje se da će za komutator sa ulaznim baferima i sa uniformnom distribucijom saobraćaja, BvN dekompozicija dati jednostavan TDM (*Time Division Multiplex*) konekcioni patern. Odnosno, ova osobina može omogućiti saobraćajno balansirani komutator koji postiže propusnost od 100% sa složenošću koja iznosi  $O(1)$ . Zbog ove osobine došlo se do ideje za dizajniranje dvostepenog BvN komutatora sa balansiranjem saobraćaja.

### 3.2. BvN komutator sa balansiranjem saobraćaja

BvN komutator sa balansiranjem saobraćaja (LB - *Load Balance*) se sastoji od dva stepena koje čine krosbar komutatori i jednog skupa VOQ redova između njih. Predložen je od strane Chang-a i ostalih [9]. Prvi stepen obavlja balansiranje saobraćaja a drugi stepen obavlja samu komutaciju saobraćaja.



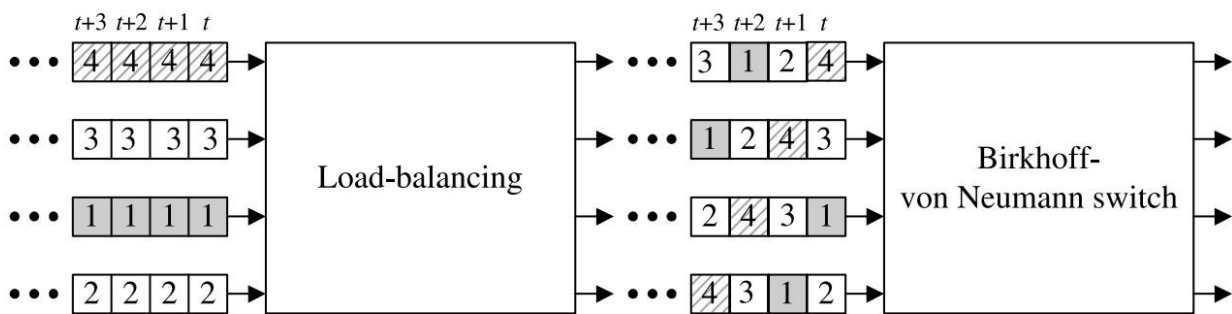
Slika 3.2.1. LB-BvN komutator[4]

Konekcioni paterni se biraju tako da u svakih  $N$  uzastopnih vremenskih slotova, svaki ulaz treba da se poveže sa svakim izlazom tačno jednom tokom trajanja jednog vremenskog slot. U prednosti LB-BvN komutatora spadaju:

- Skalabilnost. Složenost algoritma za raspoređivanje saobraćaja u komutatoru je  $O(1)$ .
- Mala hardverska složenost. Samo dva krosbar komutatora sa baferima su potrebna. Oni mogu biti realizovani preko banyan mreža, zahvaljujući determinističkoj i periodičnoj prirodi konekcionih paterni. Nije potrebno unutrašnje ubrzanje komutatora.
- Propusnost. LB-BvN komutator postiže propusnost od 100%.

- Malo prosečno kašnjenje pri velikom opterećenju i za *bursty* saobraćaj. U slučaju *bursty* saobraćaja, balansiranje saobraćaja je veoma efektivno u smanjenju kašnjenja.
- Efikasna upotreba bafera.

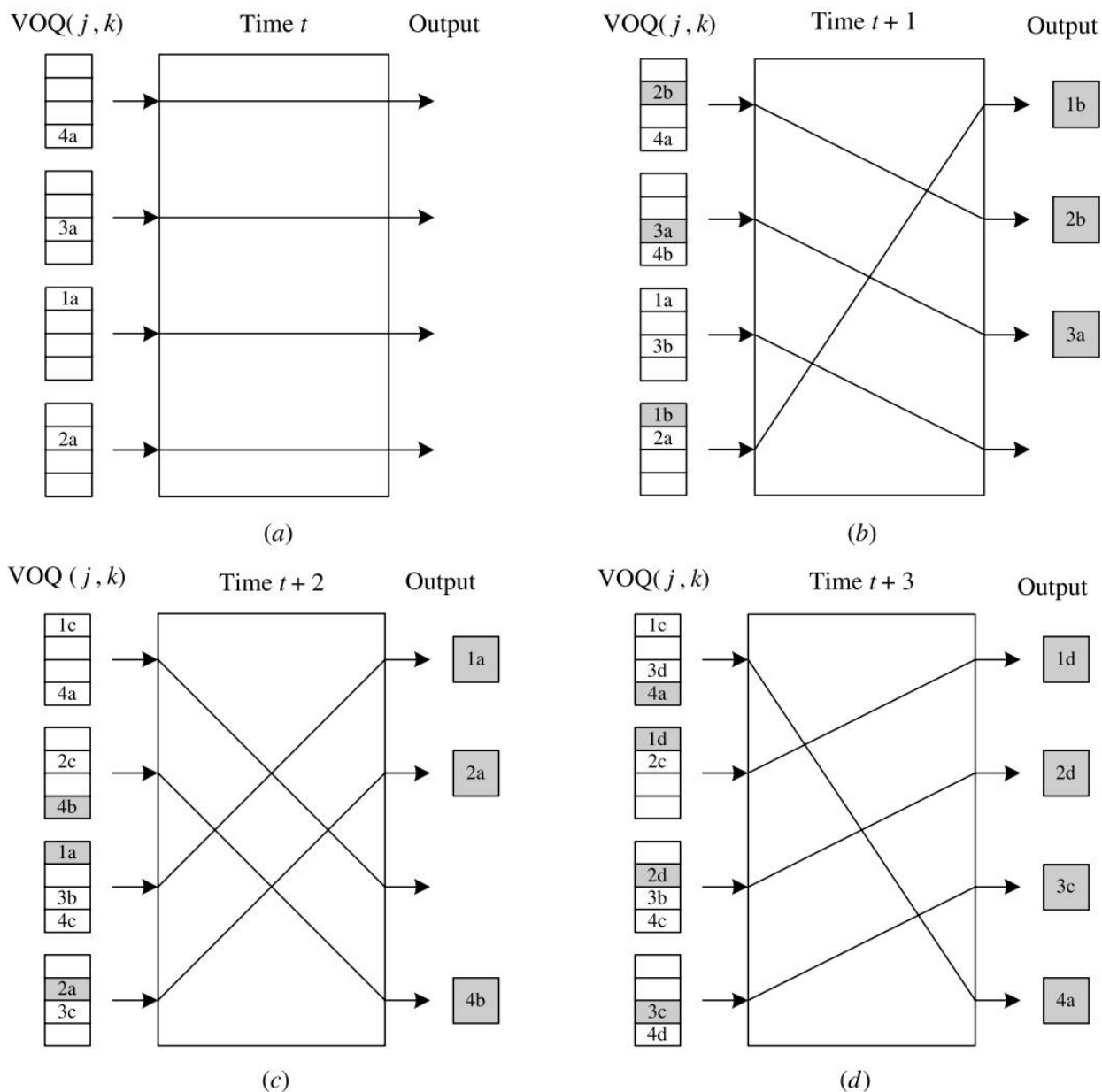
Kao što je pomenuto u potpoglavlju 3.1, jednostepeni komutator sa ulaznim baferima koji koristi determinističke sekvence postiže propusnost od 100% ako mu se na ulaz dovede uniformni Bernulijev i.i.d. (*independent and identically distributed*) saobraćaj. U LB-BvN komutatoru, prvi stepen obezbeđuje drugom stepenu baš takav saobraćaj tako što izvrši balansiranje opterećenja koristeći determinističke konekcione paterne. S obzirom na to da drugi stepen vrši komutaciju Bernulijevog i.i.d. saobraćaja, celokupan sistem može postići propusnost od 100% za skoro sve ulazne paterne saobraćaja. Prosečno kašnjenje paketa kod LB-BvN komutatora je primetno manje nego kod običnog jednostepenog BvN komutatora. Odlične performanse po pitanju kašnjenja kod LB-BvN komutatora se postižu zahvaljujući prvom stepenu, odnosno ulaznom komutatoru koji vrši balansiranje saobraćaja. On efikasno smanjuje sporadičnost (*burst*) ulaznog saobraćaja do VOQ redova koji se nalaze između dva stepena komutatora. Kao rezultat toga, drugi stepen komutatora uvek dobija na svojim ulazima uniformni Bernulijev i.i.d. saobraćaj. Sledeća slika ilustruje efekat koji izaziva prvi stepen LB-BvN komutatora.



Slika 3.2.2. Ilustrovanje smanjenja bursta od strane prvog komutatora[2]

### 3.3. Problemi koji se javljaju kod LB-BvN komutatora

Tok paketa kroz komutator je definisan kao skup svih paketa koji idu od jednog ulaza  $i$  do jednog izlaza  $k$ , i označavaćemo ga sa  $S(i, k)$ . S obzirom na to da na ulaz komutatora ne dolazi samo uniformni saobraćaj, broj paketa u svakom toku može varirati. Ovo se direktno odražava na razlike u dužinama VOQ redova između dva stepena komutatora. S obzirom na to da se VOQ redovi opslužuju uniformno, tj. nezavisno od njihovih dužina, kašnjenje i problem paketa koji izlaze van redosleda su neizbežni u osnovnoj varijanti LB-BvN komutatora. Paketi koji su van redosleda uslovljavaju TCP (*Transmission Control Protocol*) da aktivira *fast recovery*, i klizni prozor TCP protokola se smanjuje za pola. Ovo uslovljava smanjenje propusnosti od jednog kraja do drugog takođe za pola. Da bi smo detaljnije analizirali ovaj problem paketa van redosleda koristićemo sledeći primer[2]:



**Slika 3.3.1. Koraci komutiranja kod LB-BvN komutatora. (a) u trenutku  $t$ ; (b) u trenutku  $t+1$ ; (c) u trenutku  $t+2$ ; (d) u trenutku  $t+3$ . [2]**

Imamo četiri toka  $S(1, 4)$ ,  $S(2, 3)$ ,  $S(3, 1)$ ,  $S(4, 2)$ , i svaki od njih sadrži po četiri paketa (a, b, c, d), a vreme njihovog pristizanja na ulaz komutatora je prema redosledu  $a < b < c < d$ . Na slici su prikazani koraci komutiranja u različitim trenucima vremena, tokom jednog ciklusa konekcije.

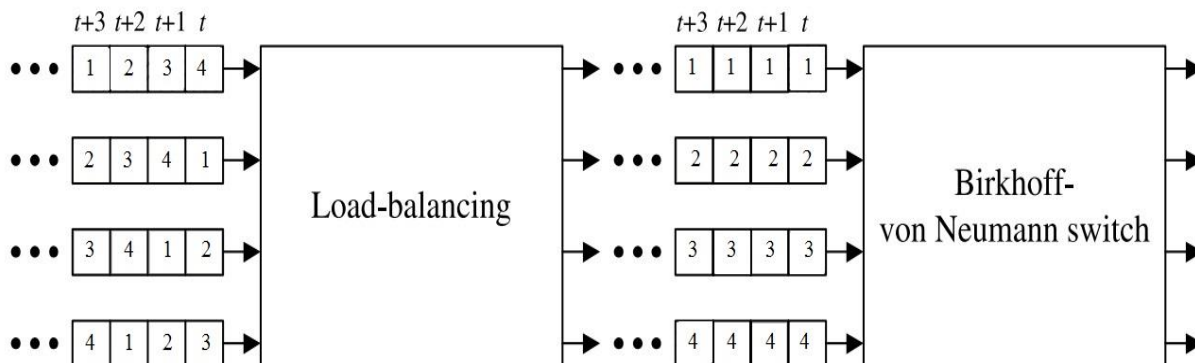
U trenutku  $t$ , četiri paketa 4a, 3a, 1a i 2a, stižu na svaki od četiri VOQ reda. S obzirom na to da nema paketa u redovima za izlaze određenih ovim konekcionim paternom, neće biti slanja paketa.

U trenutku  $t+1$ , još četiri paketa (2b, 4b, 3b, i 1b) stižu na različite VOQ. Novi konekcioni patern dozvoljava da tri paketa sa ulaza budu prosleđena. Već počinje da se primećuje problem paketa van redosleda, jer su paketi 1b i 2b prosleđeni dok paketi 1a i 2a, koji su stigli jedan vremenski slot pre, još uvek čekaju u baferima.

U trenutku  $t + 2$ , četiri nova paketa stižu i trenutni konekcioni patern prosleđuje još tri paketa. Primećuje se da paket 1a stiže na izlaz 1 dok je u prethodnom vremenskom slotu paket 1b već bio prosleđen.

U trenutku  $t + 3$ , poslednji paketi sva četiri toka stižu na ulaze drugog stepena komutatora. Konekcioni patern zatim počinje da se ponavlja periodično, dozvoljavajući da se paketi na baferima prosleđuju kada su povezani sa svojim željenim izlaznim portovima.

Drugi primer, u kome se obara propusnost komutatora dat je na slici 3.3.2.



Slika 3.3.2. Primer ilustrovanja obaranja protoka komutatora na  $\frac{1}{4}$

U datom primeru, za komutator dimenzija  $4 \times 4$ , vidi se da će svi paketi koji su namenjeni izlazu 1 svi završiti na izlazu 1 prve kaskade, oni namenjeni izlazu 2 će završiti na izlazu 2 prve kaskade itd., čime se zaključuje da dolazi do obaranja protoka na  $\frac{1}{4}$ . U opštem slučaju, za komutator dimenzija  $N \times N$ , dolazi do obaranja protoka na  $\frac{1}{N}$ .

U sledećem poglavlju će biti opisane varijante, tj. šeme koje su predložene za rešavanje problema paketa van redosleda kod LB-BvN komutatora.

## 4. VARIJANTE LB-BvN KOMUTATORA KOJE REŠAVAJU PROBLEM PAKETA VAN REDOSLEDA

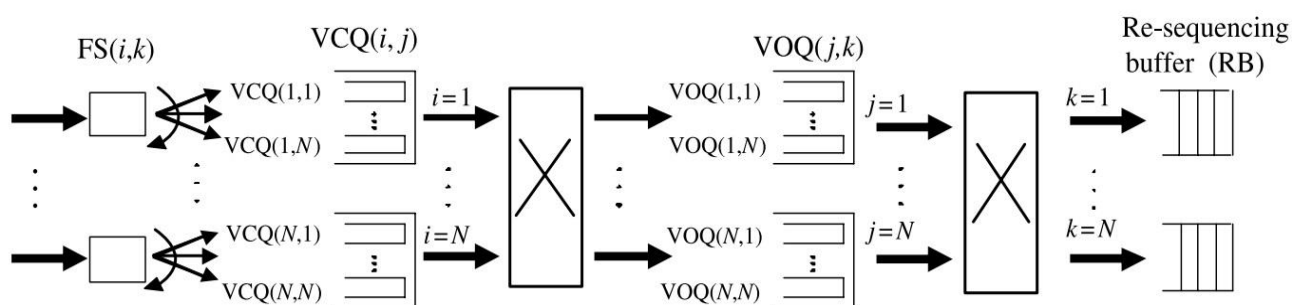
Sedam šema je predloženo za rešavanje problema van redosleda kod LB-BvN komutatora, i one se opisane u ovom poglavlju[2]:

- 1) FCFS (*First Come First Served*)
- 2) EDF (*Earliest Deadline First*)
- 3) EDF-3DQ (*Earliest Deadline First using Three-Dimensional Queue*)
- 4) FFF (*Full Frames First*)
- 5) FOFF (*Full Ordered Frames First*)
- 6) Mailbox
- 7) Byte-Focal

Ove šeme se mogu razvrstati na dve kategorije na osnovu pristupa rešavanju problema. Prvi način je da se spreče paketi da dolaze van redosleda na izlaze, npr. FFF i Mailbox komutator. Drugi pristup je da se limitira broj paketa koji su van redosleda na neku gornju granicu, i onda se na izlaz dodaje RB (*Resequencing Buffer*) koji reorganizuje pakete u pravilan redosled. Šeme koje to rade su: FCFS, EDF, EDF-3DQ, FOFF i Byte-Focal. RB bafer se konceptualno sastoji od dva virtuelna bafera: bafera za preraspodelu i izlaznog bafera. Svrha prvog je da reorganizuje pakete tako da paketi jednog toka izlaze u istom redosledu kako su i došli. Posle ponovnog sastavljanja paketi se čuvaju u izlaznom baferu gde čekaju transmisiju ka izlaznom linku.

### 4.1. First Come First Served - FCFS

Da bi rešio problem paketa van redosleda koji se javlja kod LB-BvN komutatora, FCFS šema koristi dva dodatna bafera. Ova šema je predložena od strane Chang-a i ostalih[7]. Prvi bafer služi za balansiranje saobraćaja i on se stavlja na svaki ulaz komutatora. Drugi bafer je RB i stavlja se na izlaz komutatora.



Slika 4.1.1. FCFS varijanta LB-BvN komutatora[2]

Objašnjenja oznaka sa slike su:

- $FS(i, k)$ . *Flow splitter* – deli tok  $S(i, k)$ .
- $VCQ(i, j)$ . *Virtual Central Queue*, red na ulazu  $i$  za odgovarajući izlaz  $j$  prvog stepena.
- $VOQ(j, k)$ . *Virtual Output Queue*, red na ulazu  $j$  za odgovarajući izlaz  $k$  drugog stepena.

Rad komutatora je opisan na sledeći način. Paketi koji stižu na komutator se šire na više VCQ redova da bi se izvršilo balansiranje saobraćaja. Ovo širenje toka se vrši na sledeći način. FCFS šema koristi *flow splitter* za svaki tok na svakom ulaznom portu koji označava pakete na svakom ulazu kao pripadajuće određenom toku  $S(i, k)$ . S obzirom na to da postoji  $N$  mogućih izlaza postoji i  $N$  mogućih tokova, po jedan za svaki od  $N$  izlaznih portova. Zatim se vrši raspodela saobraćaja svih paketa pristiglog toka ka  $N$  VCQ redovima, prema round-robin principu. Postoji različito balansiranje saobraćaja za svaki tok. Za svaki *flow splitter* postoji pokazivač koji prati putanju  $VCQ(i, j)$  reda na koju je poslat prethodni paket trenutnog toka. Prvi stepen komutatora prema periodičnom i determinističkom paternu povezuje ulaz  $i$  sa izlazom  $j$  za jedan vremenski slot u svakom frejm slotu. Frejm slot je definisan kao  $N$  vremenskih slotova. U vremenskom slotu  $t$ , konekcioni patern  $(i, j)$  zadovoljava:

$$i - 1 = (j - 1 + t) \bmod N \quad (4.1.1)$$

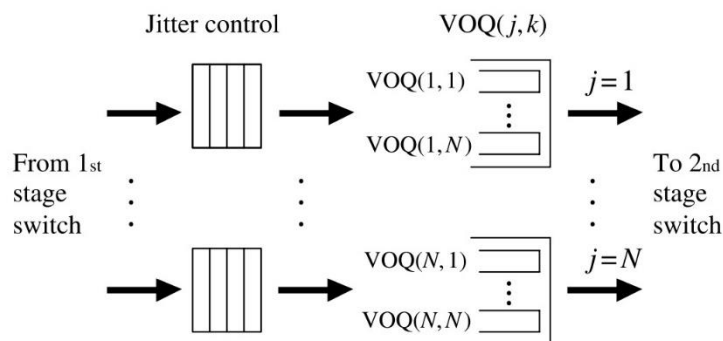
gde  $t = 1, 2, \dots, N$  i  $j = 1, 2, \dots, N$ .

Drugi stepen komutatora radi sličnu stvar kao i prvi stepen, u vremenskom slotu  $t$  konekcioni patern  $(j, k)$  zadovoljava:

$$k - 1 = (j - 1 + t) \bmod N \quad (4.1.2)$$

gde  $t = 1, 2, \dots, N$ ,  $j = 1, 2, \dots, N$ , i  $k = 1, 2, \dots, N$ .

FCFS ima mehanizam za kontrolu džitera ispred bafera drugog stepena komutatora. Džiter je varijacija kašnjenja paketa, tj. vreme kašnjenja može varirati od paketa do paketa.



Slika 4.1.2. Mehanizam za kontrolu džitera i baferi drugog stepena komutatora[2]

Paketi istog toka se jednako raspoređuju u prvom stepenu i stižu na mehanizam za kontrolu džitera. S obzirom na to da paketi istog toka mogu imati različita kašnjenja u baferima prvog stepena usled različite dužine VCQ redova, mehanizam za kontrolu džitera se koristi da organizuje pristizanje paketa na bafere drugog stepena tako što će izazivati kašnjenja svakog paketa pre nego što se pridruže VOQ redovima. Ipak, ovo ne mora da znači da će odlazak paketa sa bafera drugog stepena biti u istom redosledu kao pri prispeću. Ovo je najčešće posledica različitih dužina paketa koji se pridružuju VOQ redovima. Zbog toga je potrebna upotreba RB na izlazima.

Naredni primer[2] prikazuje kako šema FCFS rešava problem paketa van redosleda. Posmatramo tok  $S(1, 1)$ , koji se odnosi na tok paketa od ulaza 1 do izlaza 1. Paketi  $P_a$  i  $P_b$  su dva uzastopna paketa iz ovog toka i raspoređeni su tako da je  $P_a < P_b$ .  $P_a$  i  $P_b$  se dele na dva uzastopna toka  $VCQ(1, j)$  i  $VCQ(1, j + 1)$  prema round-robin principu po njihovom prispeću. S obzirom da dužine svakog od  $VCQ(i, j)$  nisu iste, paketi koji pripadaju istom toku mogu napustiti  $VCQ(i, j)$  prema neuređenom redosledu. Ovo se rešava korišćenjem mehanizma za kontrolu džitera koji se dodaje ispred  $VOQ(j, k)$ . Ipak, nejednake dužine redova  $VOQ(j, k)$  nisu ravnomerno raspoređene od vremenskog slota do vremenskog slota. Iako mehanizam za kontrolu džitera rešava problem paketa van redosleda izazvanog od strane ulaznog bafera  $VCQ(i, j)$ , ovaj problem se može ponovo javiti zbog  $VOQ(j, k)$ . Zbog toga, RB se dodaju na svaki izlaz. Oni vrše reorganizaciju paketa tako da paketi toka napuštaju komutator u istom redosledu kao što su došli. Oni se zatim čuvaju u izlaznom baferu gde čekaju svoj red za slanje.

FCFS šema LB-BvN komutatora rešava problem paketa van redosleda po cenu:

- 1) Dve dodatne grupe bafera, na ulazima i izlazima;
- 2) Mehanizma za kontrolu džitera, koji izaziva kašnjenje paketa;
- 3) Zbog 1) i 2) se povećava složenost samog komutatora, a samim tim i njegova cena što uslovljava mogućnost implementacije.

Još jedna od mana osnovne varijante LB-BvN komutatora je da obezbeđuje samo *best effort service*[3]. To znači da je kompletan saobraćaj u mreži ravnopravan i da nema mogućnosti kontrole parametara kašnjenja, protoka, pouzdanosti pa sistem ne može garantovati da će servis imati zahtevani kvalitet (*guaranteed rate service*). Da bi se obezbedio i *guaranteed rate service* uvedene su šeme FCFS i EDF.

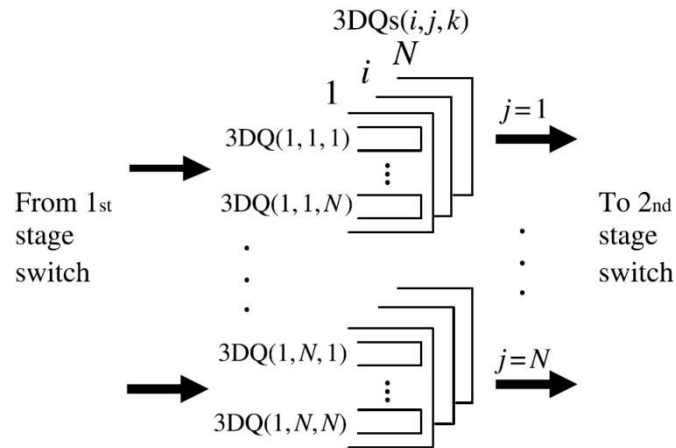
## 4.2. Earliest Deadline First – EDF

EDF šema je predložena od strane Chang-a i ostalih [7]. EDF koristi istu arhitekturu komutatora kao FCFS. Međutim, EDF šema izbacuje mehanizam za kontrolu džitera. Umesto njega, svakom paketu se dodeljuje krajnji rok (*deadline*) da bi se utvrdilo njegovo vreme odlaska iz bafera drugog stepena,  $VOQ(j, k)$ . Paketi na izlaznom baferu drugog stepena se opslužuju prema vrednostima njihovih krajnjih rokova. Što je krajnji rok bliži to će se paket pre opslužiti na komutatoru drugog stepena. S obzirom na to da paketi stižu na  $VOQ(j, k)$  bez nekog određenog redosleda, ne mora da znači da će HOL paket sadržati ovaj krajnji rok. Traženje najmanje vremenske oznake (ovih krajnjih rokova) u svakom  $VOQ$  redu je veoma složeno i skupo. Takođe, EDF šema zahteva RB veličine  $2N^2 - 2N$  za reorganizaciju paketa van redosleda.

## 4.3. Earliest Deadline First using Three-Dimensional Queues, EDF-3DQ

I. Keslassy i N. McKeown su u [10] predložili EDF-3DQ šemu koja unapređuje EDF šemu tako što se  $VOQ$  redovi u baferima drugog stepena menjaju sa trodimenzionalnim redovima (3DQ - Three Dimensional Queue).



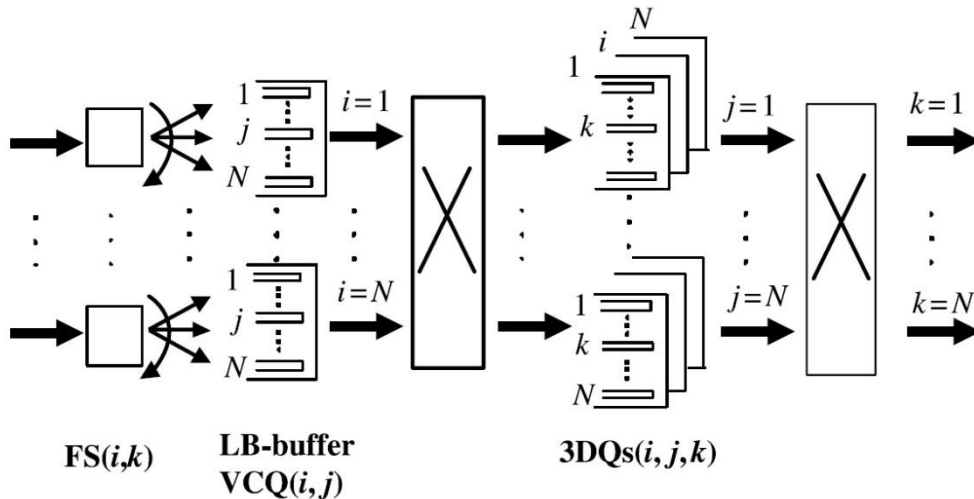


Slika 4.3.1. Trodimenzionalni redovi kod EDF[2]

U 3DQ strukturi postoje po tri različita reda za  $(i, j, k)$ , zbog toga postoji  $N^3$  logičkih redova između prvog i drugog stepena komutatora. Kod originalne VOQ strukture, svaki  $VOQ(j, k)$  red sadrži pakete od više tokova koji su određeni za izlazni port  $k$ . U 3DQ strukturi, svaki  $VOQ(j, k)$  red ima u sebi celokupan broj od  $N$  redova obeleženih sa  $3DQ(i, j, k)$ . Svaki od ovih  $N$  redova sadrži pakete istog toka. Postoji ukupno  $N$  mogućih tokova, što znači da postoji  $N$  redova za svaki  $VOQ(j, k)$ . Na primer,  $3DQ(1,1,1)$  sadrži samo pakete toka  $S(1,1)$ ;  $3DQ(2,1,1)$  sadrži samo pakete toka  $S(2,1)$ . U originalnoj VOQ strukturi, oba ova toka mogu biti smeštena u jedan red. Pa je tako cilj 3DQ strukture da odvoji pakete na svakom VOQ redu u njihove individualne tokove. Kod 3DQ, najraniji (prvi) paket za  $(j, k)$  je uvek HOL paket za svoj red. Kao rezultat, moramo da uporedimo samo  $N$  HOL vremenskih oznaka paketa da bi smo pronašli najraniji krajnji rok paketa, umesto upoređivanja  $Q_{max}$  vremenskih oznaka u originalnoj EDF šemi, gde je  $Q_{max}$  maksimalna dužina reda VOQ. Iako EDF-3DQ smanjuje opterećenje koje iziskuje potraga za najmanjom vremenskom oznakom u celokupnom redu, i dalje je potrebno izvršiti poređenje  $N$  vremenskih oznaka  $N$  HOL paketa svakog 3DQ u svakom vremenskom slotu, što utiče na veličinu komutatora kao i na njegovu brzinu. Kao kod originalne EDF šeme, i kod EDF-3DQ šeme je potrebno da veličina RB bude  $2N^2 - 2N$ .

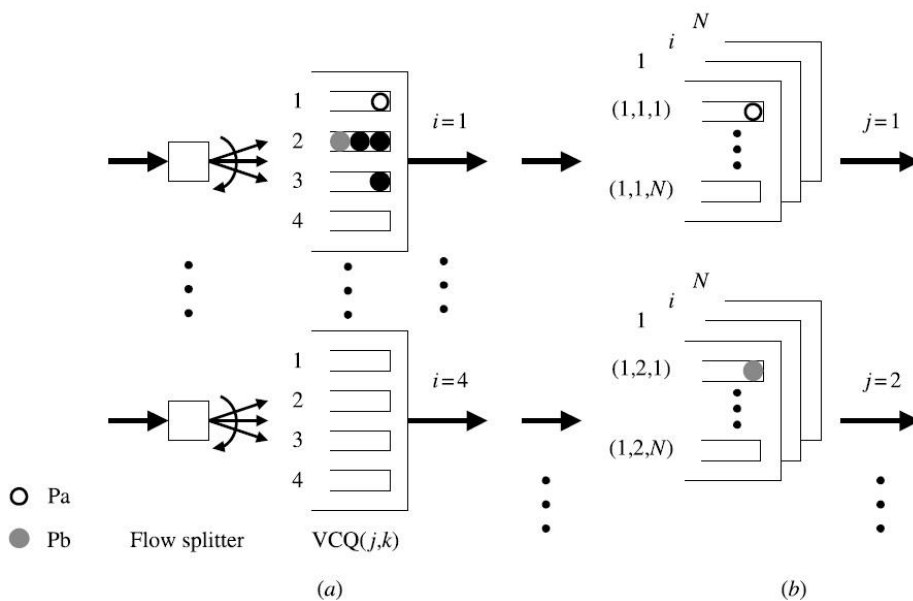
#### 4.4. Full Frames First – FFF

Ova šema koristi pristup na bazi frejmova da bi rešila problem paketa koji izlaze van rasporeda kod LB-BvN komutatora. Predložena je u [10]. Ova šema se razlikuje od drugih zato što kompletno reguliše problem ispadanja paketa iz redosleda, a da pritom ne koristi RB bafere na izlazima. Arhitektura komutatora je slična kao kod FCFS i EDF šema, samo što FFF varijanta ima 3DQ između dva komutatora umesto VOQ, koji je prisutan kod FCFS i EDF šema. Najveća prednost je što ne zahteva dodatne RB.



Slika 4.4.1. Full Frame First varijanta LB-BvN komutatora[2]

Slično kao i kod FCFS i EDF varijanti, svaki tok saobraćaja  $S(i, k)$  je podeljen na  $N$   $VOQ(i, j)$  redova i to na osnovu round-robin principa prema njihovim dolascima na ulaz. Prvi stepen komutatora deterministički isporučuje pakete od strane reda  $VCQ(i, j)$  do  $3DQ(i, j, k)$ , periodično. Slično tome, u drugom stepenu se vrši deterministička konekcija paterna od paketa sa  $3DQ(i, j, k)$  na njihove odgovarajuće izlaze. Da bi se sprečilo ispadanje paketa iz redosleda usled različitih dužina tokova 3DQ, algoritam mora da prosledi pakete od 3DQ do izlaza. U FFF, skup kandidata 3DQ za  $(i, k)$  se sastoji od paketa od tokova  $(i, 1, k)$ ,  $(i, 2, k)$  i  $(i, N, k)$ . Važno je zapamtiti da se svaki  $3DQ(i, j, k)$  sastoji od paketa koji dolaze od jedinstvenih tokova (paketi sa različitih tokova neće biti u istom redu). Zbog balansiranja saobraćaja, tok je uniformno raspoređen između svih  $N$  3DQ. Pretpostavimo da je poslednji uslužen paket u skupu kandidata došao od  $3DQ(i, j_{last}, k)$ . Zbog osobina balansiranja opterećenja i 3DQ, znamo da će sledeći paket koji dolazi na red za tok  $S(i, k)$  doći od  $3DQ(i, j_{(last+1) \bmod N}, k)$ . Neka je  $p_{ik}$  pokazivač na 3DQ  $(i, j, k)$  i to na paket koji je sledeći na redu:  $p_{ik} = j_{(last+1) \bmod N}$ . Na primer, neka su  $P_a$  i  $P_b$  dva paketa koja pripadaju istom toku  $S(1, 1)$  i to po redosledu  $P_a < P_b$ . Oni će biti u redu kod  $VCQ(1, 1)$  i  $VCQ(1, 2)$ , respektivno, prema njihovim dolascima na ulaze.



Slika 4.4.2. Osobine flow splittera i 3DQ-ova. (a)  $P_a$  i  $P_b$  na ulazu  $VCQ(j, k)$ ; (b)  $P_a$  i  $P_b$  na  $3DQ(i, j, k)$ [2]

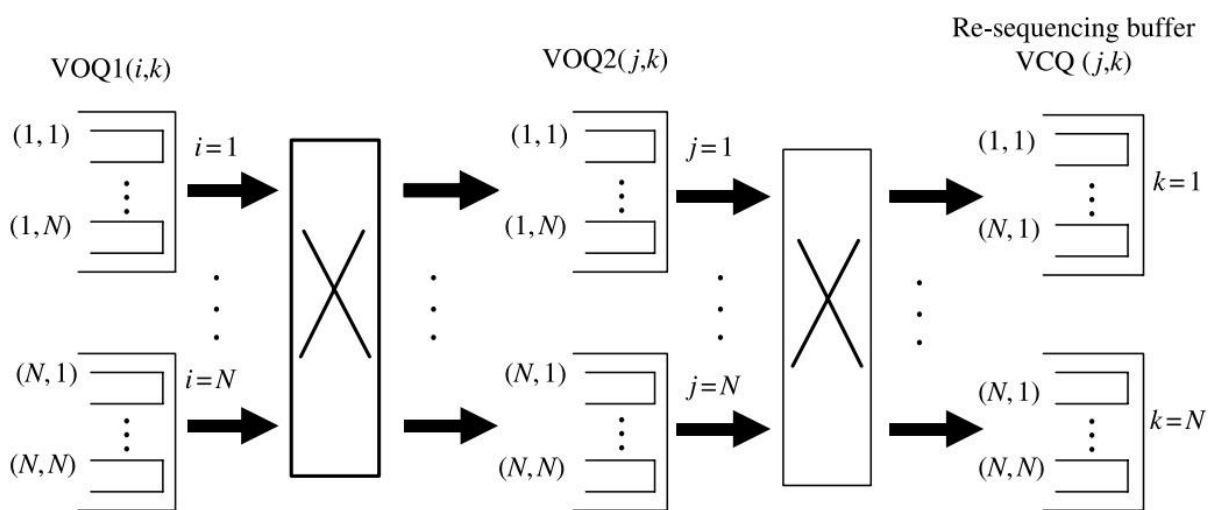
Iako će Pa i Pb biti prosleđeni 3DQ u različitim vremenskim slotovima, može se očekivati da će biti u redovima 3DQ(1, 1, 1) i 3DQ(1, 2, 1), respektivno. Drugim rečima, ako je Pa iz 3DQ(1, 1, 1), onda će paket koji je sledeći na redu obavezno biti učitav iz 3DQ(1, 2, 1) sa  $p_{ik} = 1 + 1 = 2$ , koje je Pb u ovom primeru. Kod FFF, ciklus je definisan kao  $N$  uzastopnih vremenskih slotova. Frejm toka  $(i, k)$  definišemo kao  $f(i, k) = (i, p_{ik}, k), (i, p_{ik} + 1, k), \dots, (i, N, k)$ . Frejm je pun ako je svaki 3DQ( $i, j, k$ ) red za  $j = p_{ik}, \dots, N$  neprazan. Može se primetiti da ako je frejm pun, drugi stepen komutatora može neprekidno da premešta pakete koji su u tačnom redosledu od 3DQ( $i, p_{ik}, k$ ) do 3DQ( $i, N, k$ ). Ovo je ključno da bi se sprečio problem paketa van redosleda kod FFF. Potraga za punim frejmovima se vrši po jednom u svakom ciklusu. Izlaz učitava sve pakete iz jednog ulaza u pun frejm, pre nego što nastavi da učitava pun frejm sledećeg ulaza. Svaki izlaz  $k$  koristi round-robin pokazivač  $p_{ff}(k)$  da bi zapamtio sa kojeg ulaza je došao poslednji pun frejm. Ako nema punih frejmova, izlaz  $k$  opslužuje nepune frejmove prema round-robin principu koristeći pokazivač  $p_{nff}(k)$ . Preciznije rečeno, FFF šema koristi tri koraka za svaki izlazni port  $k$  na početku svakog ciklusa. To su koraci:

- 1) Određuje se koji od frejmova  $f(i, k)$  je pun;
- 2) Počevši od  $p_{ff}(k)$ , pronalazi se prvi pun frejm i zatim se ažurira pokazivač na pune frejmove  $p_{ff}(k)$  na taj ulaz kom je pridružen taj pun frejm;
- 3) Ako nema punih frejmova, polazeći od  $p_{nff}(k)$ , pronalazi se prvi nepun frejm, a zatim se ažurira pokazivač prema njemu.

FFF najbolje koristi prednosti LB-BvN komutatora. Ipak, potrebna mu je upotreba složenih 3DQ redova između dva komutatora, kao i velika količina komuniciranja na tokovima da bi se tražili puni frejmovi.

#### 4.5. Full Ordered Frames First – FOFF

FOFF je još jedna šema bazirana na frejmovima koja rešava problem paketa van redosleda kod LB-BvN komutatora. I. Keslasy i ostali su je predložili u [11]. Za razliku od FFF šeme, FOFF dozvoljava paketima da budu van redosleda kroz oba stepena komutatora.



Slika 4.5.1. Arhitektura FOFF komutatora[2]

Ovaj komutator se sastoji od tri grupe bafera i dva deterministička TDM komutatora. Tri grupe bafera su:

- 1) VOQ1( $i, k$ ) na svakom ulazu, svaki je povezan sa svojim izlaznim portom  $k$ ;
- 2) VOQ2( $j, k$ ) između dva komutatora, svaki od njih je povezan sa svojim izlaznim portom  $k$ ;
- 3) RB VCQ( $j, k$ ) na izlazima, svaki je dodeljen ulazu komutatora drugog stepena.

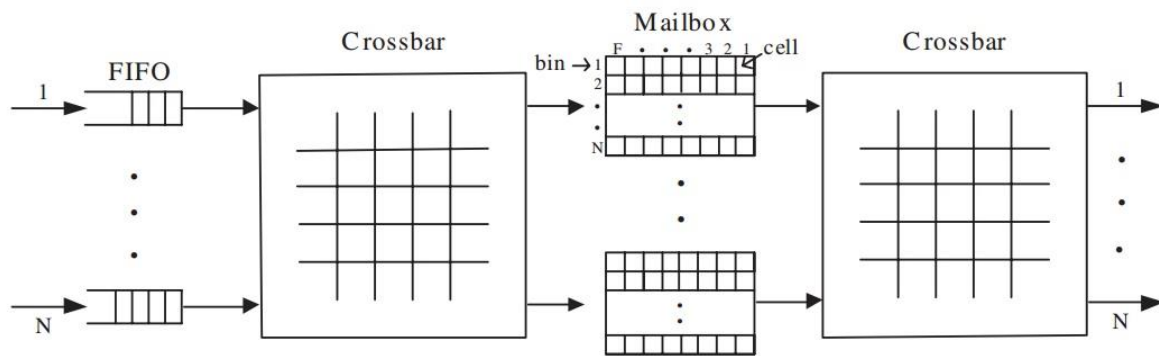
Paketi se šalju u red na VOQ1( $i, k$ ) nakon njihovog pristizanja, kao i kod tradicionalnih komutatora sa ulaznim baferima. Na početku svakog frejma (1 frejm =  $N$  vremenskih slotova), svaki ulaz bira VOQ1 koji će slati pakete u sledećem frejm slotu. Puni frejmovi u različitim VOQ1 se prvo biraju i opslužuju prema round-robin principu. Ako nema punih frejmova, parcijalni frejmovi se biraju i opslužuju prema round-robin principu. Kada se parcijalni frejmovi biraju za slanje u sledeći frejm slot, biće određenog gubitka protoka na komutatoru prvog stepena. Ako se bar jedan puni frejm može naći na svakom ulazu u svakom frejm slotu, neće biti problema sa gubitkom redosleda paketa. Ali, ako ulaz može da pošalje samo parcijalni frejm, onda će postojati razlika u zauzetosti na VOQ2 i paketi će biti van redosleda na izlazu. Ali ovaj problem je ograničen. Dokazano je da je RB dimenzije  $N^2$  na svakom izlazu komutatora dovoljan da izvrši ponovnovraćanje paketa u redosled. Drugim rečima, ako postoji  $N^2+1$  paketa u RB, bar jedan od HOL paketa sa VCQ je HOF (*Head Of Flow*) paket i može se izabrati za prenos. Parcijalni frejmovi mogu izazvati drugi problem kod FFFF šeme, pored problema paketa van redosleda na izlazu. Oni mogu izazvati gubitak protoka na komutatoru prvog stepena i samim tim izazvati povećanje prosečnog kašnjenja paketa. Ponovno vraćanje paketa u redosled na izlazu može biti veoma zahtevno. Dokazano je da je bar jedan paket van redosleda ako se šalje  $N^2 + 1$  paketa. U svakom vremenskom slotu moramo tražiti i do  $N$  HOL paketa od VCQ redova pre nego što nađemo HOF paket. HOF paket toka je određen tako što se uporedi njegov redni broj paketa sa očekivanim rednim brojem paketa (istog toka). U svakom vremenskom slotu, svaki izlaz vrši sledeće operacije da bi pronašao HOL paket:

- 1) Određuje kom toku pripada HOL paket
- 2) Određuje broj sekvenci HOF paketa toka
- 3) Upoređuje se 2) sa brojem sekvenci HOL paketa
- 4) Ako se podudara, šalje se HOL paket i ažurira se sledeći broj sekvenci toka
- 5) U suprotnom, ponavljaju se koraci 1) -4) dok se ne nađe sledeći VCQ HOL paket.

U najgorem slučaju može biti i do  $N$  ponavljanja gore navedenih operacija pre nego što se pronađe HOF paket za slanje.

#### 4.6. Mailbox komutator

Predložena je od strane Chang-a i ostalih u [8]. Pretpostavka je da su svi paketi jednake dužine. Takođe, pretpostavićemo da je vreme podeljeno u slotove i sinhronizovano tako da se jedan paket može poslati u jednom vremenskom slotu. Komutator se sastoji od dva krosbar komutatora dimenzija  $N \times N$ , i bafera između njih. Ti baferi se nazivaju poštanski sandučići, tj. *mailboxes*. Postoji  $N$  *mailbox*-ova i oni su numerisani od 1 do  $N$ . Svaki *mailbox* se sastoji od  $N$  podeoka i oni su takođe numerisani od 1 do  $N$ . Svaki podeok se sastoji od  $F$  ćelija, koje su numerisane od 1 do  $F$ . U svaku ćeliju može stati tačno jedan paket. Ćelije koje se nalaze u  $i$ -tom podeoku *mailbox*-a služe čuvanju paketa koji su namenjeni  $i$ -tom izlaznom portu drugog komutatora. Kao dodatak, postoje i FIFO redovi ispred svakog ulaznog porta prvog stepena komutatora.



Slika 4.6.1. Arhitektura Mailbox komutatora[7]

Sledi objašnjenje za formiranje konekcionih parametara između dva komutatora. U svakom vremenskom slotu, oba komutatora sa slike 4.6.1. imaju isti konekcioni patern. Tokom  $t$ -tog vremenskog slotu, ulazni port  $i$  je povezan sa izlaznim portom  $j$  ako je:

$$(i + j) \bmod N = (t + 1) \bmod N \quad (4.6.1)$$

Konkretno, u trenutku  $t = 1$ , ulazni port 1 je povezan sa izlaznim portom 1, ulazni port 2 je povezan sa izlaznim portom  $N - 1, \dots$ , i ulazni port  $N$  je povezan sa izlaznim portom 2. Očigledno, ovi konekcioni paterni su periodični i njihova perioda je  $N$ . Takođe, svaki ulazni port je povezan sa svakim od  $N$  izlaznih portova tačno jednom u svakom od  $N$  vremenskih slotova. Konkretno, ulazni port  $i$  je povezan sa izlaznim portom 1 u trenutku  $i$ , sa izlaznim portom 2 u trenutku  $i + 1, \dots$ , sa izlaznim portom  $N$  u trenutku  $i + N - 1$ . Takođe, na osnovu formule 4.6.1 se vidi da su ovi konekcioni paterni simetrični, tj. ulazni port  $i$  i izlazni port  $j$  su povezani samo ako su povezani ulazni port  $j$  i izlazni port  $i$ . Iz formule 4.6.1 možemo izračunati  $j$  kao:

$$j = h(i, t) = ((t - i) \bmod N) + 1 \quad (4.6.2)$$

gde je sa  $h(i, t)$  obeležen mailbox ulaznog porta i prvog komutatora u trenutku  $t$ .

Kako su ulazni port  $i$  prvog komutatora i izlazni port  $i$  drugog komutatora na istom linijskom modulu, ova osobina simetričnosti omogućava uspostavljanje bidirekcionog komunikacionog linka između linijskog modula i mailbox-a. Kako su konekcioni paterni u mailbox komutatoru specijalan slučaj jednostepenog LB-BvN [7], moglo bi se pretpostaviti da će mu i propusnost težiti 100% ako se koristi FIFO princip za svaki podeok i ako se povećava veličina podeoka  $F$  ka beskonačnosti. Ipak, doći će do problema paketa van redosleda. Može se desiti da paketi koji imaju isti ulazni port na prvom komutatoru i isti izlazni port na drugom komutatoru budu rutirani na različite mailbox-ove i da napuste drugi stepen komutatora u drugačijem redosledu nego kad su stigli na ulazni port prvog komutatora.

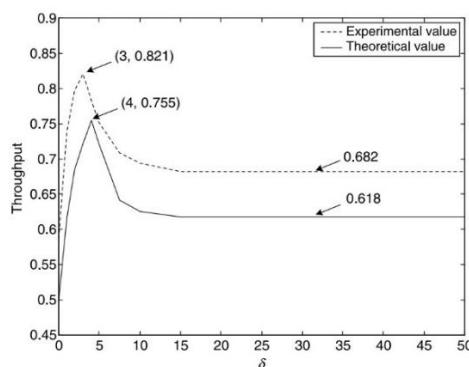
Za rešavanje problema paketa van redosleda koristi se sledeći princip. Kada se paket smesti u mailbox potrebno je da znamo njegovo vreme odlaska, s obzirom da su konekcioni paterni deterministički i periodični. S obzirom na to da se ulazni port prvog komutatora i njemu odgovarajući izlazni port drugog komutatora najčešće nalaze na istom linijskom modulu, informacija o odlasku paketa se može poslati ulazima, pa se paketi mogu poređati tako da budu u istom redosledu kao kad su stigli na ulaz prvog komutatora.

Konkretno, neka je tok  $(i, j)$  definisan kao sekvenca paketa koji stižu na ulazni port  $i$  prvog komutatora i koji su namenjeni izlaznom portu  $j$  drugog komutatora. Kada paket toka  $(i, j)$  postane HOL (Head Of Line) paket FIFO reda  $i$ -tog ulaznog porta prvog stepena komutatora u trenutku  $t$ , on

mora da sačeka određen broj vremenskih slotova, koji definišemo sa  $V_{i,j}(t)$ .  $V_{i,j}(t)$  nazivamo virtuelno vreme čekanja toka  $(i, j)$ . Sada će se opisati kako *mailbox* komutator vraća pakete u redosled u kom su stigli na ulaze prvog stepena komutatora.

- **Primanje pošte:** u trenutku  $t$ ,  $i$ -ti izlazni port drugog komutatora je povezan sa  $h(i, t)$ -im *mailbox*-om. Paket u prvoj ćeliji  $i$ -tog podeoka se šalje  $i$ -tom izlaznom portu. Paketi u ćelijama 2, 3,...,  $F$   $i$ -tog podeoka se pomeraju za jednu ćeliju unapred, tj. u ćelije 1, 2,...,  $F - 1$ .
- **Slanje pošte:** pretpostavi se da HOL paket  $i$ -tog ulaznog porta prvog komutatora potiče od toka  $(i, j)$ . Primećuje se da je  $i$ -ti izlazni port prvog komutatora takođe povezan sa  $h(i, t)$ -im *mailbox*-om. Da bi se održao pravilan redosled paketa, ovaj HOL paket se smešta u prvu praznu ćeliju  $j$ -tog podeoka  $h(i, t)$ -og *mailbox*-a tako da neće otići pre trenutka  $t + V_{i,j}(t)$ . Ako ne postoji takva prazna ćelija, HOL paket se blokira i zadržava status HOL paketa tog FIFO reda.
- **Ažuriranje virtuelnih vremena čekanja:** svi tokovi koji ne šalju pakete u trenutku  $t$  smanjuju svoje virtuelno vreme čekanja za 1. Ovde se svrstavaju i tokovi koji imaju blokirane transmisije. Da bi se ažuriralo virtuelno vreme čekanja toka  $(i, j)$ , pretpostavlja se da je HOL paket smešten u  $f$ -tu ćeliju  $j$ -tog podeoka  $h(i, t)$ -og *mailbox*-a. S obzirom da su konekcioni paterni deterministički i periodični, jednostavno se zaključuje da će  $h(i, t)$ -i *mailbox* biti povezan sa  $j$ -tim izlaznim portom drugog komutatora u trenutku  $t + ((j - i - 1) \bmod N) + 1$ . Stoga, vreme odlaska ovog paketa će biti  $t + (f - 1)N + ((j - i - 1) \bmod N) + 1$ . Prema tome, broj vremenskih slotova koji moraju da se sačekaju u trenutku  $t + 1$  za tok  $(i, j)$  iznosi  $(f - 1)N + ((j - i - 1) \bmod N)$ . Prema tome, dobija se da virtuelno vreme čekanja u trenutku  $t + 1$  iznosi  $V_{i,j}(t + 1) = (f - 1)N + ((j - i - 1) \bmod N)$ .

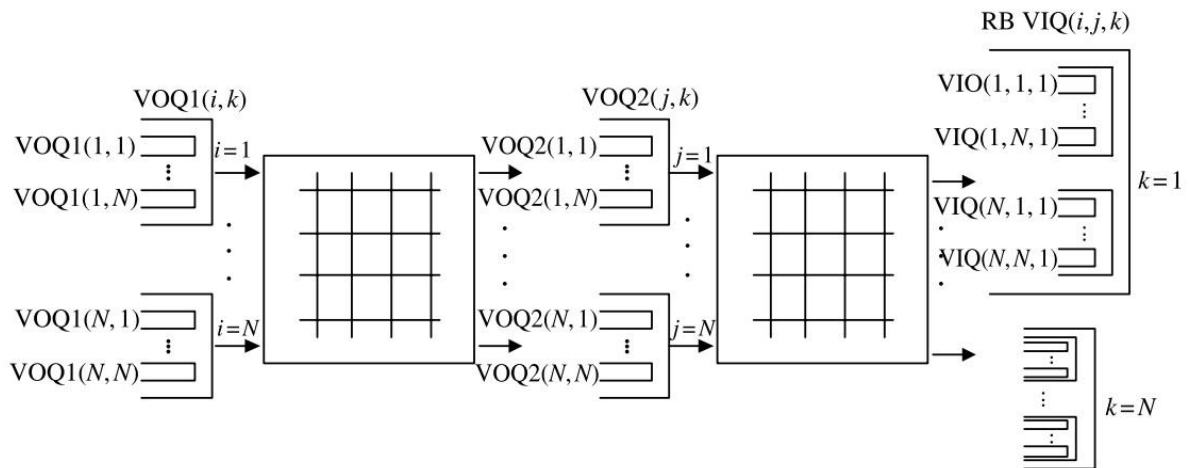
S obzirom na to da je dužina svakog *mailbox*-a ograničena, i da potraga za odgovarajućom ćelijom za smeštanje paketa zahteva veći broj pokušaja, svaki bezuspešni pokušaj može rezultirati odbijanjem  $N$  vremenskih paketa za paketsko vreme odlaska. Takav gubitak ne utiče samo na pakete koji se smeštaju, već i na sve sledeće pakete koji pripadaju istom toku zbog virtuelnog vremena čekanja tog toka koji je takođe povećan za  $N$  vremenskih slotova. Da bi se izbegli ti problemi, bolje je da se ograniči broj pokušaja slanja ( $\delta$ ) koje ćelija može pokušati, tako da posle potrage  $\delta$  ćelija izvan virtuelnog vremena čekanja, paket jednostavno odustane od tog vremenskog slota. Ovaj parametar  $\delta$  je veoma bitan kod *Mailbox* komutatora. Kao što je prikazano na slici 4.6.2. kad je  $\delta$  malo, propusnost sistema je ograničena HOL blokiranjem na FIFO redovima prvog komutatora. S druge strane, kada je  $\delta$  veliko, propusnost je ograničena stabilnošću virtuelnih vremena čekanja. Najveća propusnost koju *Mailbox* komutator može da postigne je 95%.



Slika 4.6.2. Propusnost Mailbox komutatora kao funkcija od  $\delta$ [2].

## 4.7. Byte-Focal komutator

U ovom poglavlju je opisana praktična šema komutatora sa balansiranjem saobraćaja, predložena od strane Y.Shen-a i ostalih u [12]. Naziva se *Byte-Focal* komutator i koristi raspoređivanje paket po paket da bi se značajno poboljšale performanse što se tiče kašnjenja u odnosu na komutatore slične složenosti. Nazvan je tako jer mu se paketi toka (saobraćaj od ulaza do izlaza) šire preko svih linijskih modula i zatim se dovode do fokalne tačke(određenog izlaza). *Byte-Focal*-ni komutator je jednostavan za implementaciju i veoma je skalabilan. Ne koristi kompleksni algoritam organizacije, ili bilo kakvu komunikaciju između linijskih modula. Ovaj komutator postiže propusnost od 100%.



Slika 4.7.1 Arhitektura *Byte-Focal*-nog komutatora[2]

Sastoji se od dva deterministička komutatora i tri stepena redova, koje nazivamo ulazni red  $i$ , centralni red  $j$ , i izlazni red  $k$  sa RB, gde  $i, j, k=1, 2, \dots, N$ . Komutatori rade na isti način kao i kod LB-BvN komutatora, gde oba stepena koriste determinističke i periodične konekcione paterne. Postoje dva stepena VOQ, VOQ1 i VOQ2 za prvi i drugi stepen, respektivno. Tok  $f_{ik}$  je definisan paketima koji dolaze sa ulaznog porta  $i$  i određeni su za izlazni port  $k$ . Kao što je prikazano na slici paketi od  $f_{ik}$  su smešteni u VOQ1( $i, k$ ). Zbog toga što je u svakom vremenskom slotu ulazni port na prvom stepenu povezan sa drugim stepenom ciklično, paketi u VOQ1( $i, k$ ) su poslani na  $N$  ulaznih portova drugog komutatora i po round-robin principu se raspoređuju po VOQ2(1, $k$ ), VOQ2(2, $k$ ), ..., VOQ2( $N, k$ ) prema njihovim konačnim odredištima. *Byte-Focal*-ni komutator garantuje da će se (kumulativno) broj paketa koji su poslani na svaki od izlaznih portova drugog komutatora za svaki tok razlikovati za najviše jedan. VOQ2 se onda opslužuje od strane drugog fiksiranog, ravnomerno opterećenog komutatora. S obzirom na to da paketi, generalno, pate od različitih kašnjenja u drugom stepenu, oni stižu na izlaz van redosleda. *Byte-Focal*-ni komutator koristi VIQ (*Virtual Input Queue*) strukturu za RB. Na svakom izlazu, postoji  $N$  skupova VIQ redova gde svaki skup odgovara određenom ulaznom portu  $i$ . Unutar svakog skupa VIQ, postoji  $N$  logičkih redova i svaki od tih redova odgovara određenom ulazu  $j$  na drugom stepenu. VIQ( $i, j, k$ ) razdvaja svaki tok ne samo na osnovu ulaznih portova  $i$ , već i po redovima  $j$  na drugom stepenu. Paketi sa ulaza  $i$  koji trebaju da stignu do izlaza  $k$  preko ulaza  $j$  drugog stepena su sačuvani u VIQ( $i, j, k$ ). Očigledno je da su paketi u istom VOQ( $i, j, k$ ) u dobrom redosledu. HOF paket se definiše kao prvi paket u datom toku koji još nije napustio komutator, i HOL paket kao prvi paket datog VIQ( $i, j, k$ ) reda. U svakom VIQ skupu, pokazivač pokazuje na VIQ( $i, j, k$ ) na koji će sledeći očekivani HOF paket stići. Zbog ove osobine usluživanja komutatora prvog stepena, svaki ulazni port jednako raspoređuje pakete prema round-

robin principu na red  $j$  drugog stepena. Ovo garantuje da će HOF paket da se prikaže kao HOL paket skupa VIQ prema round-robin principu. Prema tome, u svakom vremenskom slotu, ako je HOF paket na izlazu, on se opslužuje i pokazivač se pomera na sledeću lokaciju  $VIQ(i, (j+1) \bmod N, k)$  HOF paketa. S obzirom na to da ima  $N$  tokova po izlazu, više od jednog HOF paketa može biti pogodno za opsluživanje u datom vremenskom slotu. Prema tome, kao dodatak VIQ strukturi, postoji DP (*Departure Queue*) čija je dužina najviše  $N$ . DQ je najjednostavnije rečeno FIFO logički red. On čuva indekse VIQ skupova, ali samo jedan od svakog VIQ skupa. Kada HOF paket VIQ skupa  $i$  stigne, index  $i$  se pridružuje kraju DQ. Kada paket ode sa DQ, njegov indeks se pomera sa čela DQ i pridružuje se kraju DQ ako je stigao sledeći HOF paket. Prednost korišćenja VIQ i DQ strukture je da vremenska složenost pronalaženja i opsluživanja paketa koji su sinhronizovani iznosi  $O(1)$ . Na svakom vremenskom slotu, svaki VIQ skup koristi pokazivač da proveriti da je li HOF paket stigao, dok izlazni port opslužuje jedan paket sa čela DQ. Kao što je objašnjenje, VIQ struktura omogućava da *Byte-Focal*-ni komutator emituje pakete u pravilnom redosledu.



## 5. ZAKLJUČAK

Internet, kao najveća računarska mreža, postaje integralan deo ljudskih života i koristi se za veoma širok spektar aktivnosti. Sve je veći broj korisnika ove mreže, samim tim i uređaja koji se umrežavaju. To dovodi do razvijanja sve većeg broja servisa koji zahtevaju sve veće protoke i samim tim je potreba za efikasnim organizovanjem internet saobraćaja sve izraženija. Zbog toga se u poslednje vreme vrše istraživanja u velikom obimu na ovom polju. Potrebno je sve više povećavati protok podataka korisnika, i pritom unapređivati kvalitet servisa. Uz sve to potrebno je obratiti pažnju i na ekonomske aspekte, tj. efikasnost ovih rešenja.

U ovom radu je analizirana arhitektura komutatora predložena od strane Birkhoff-a i von Neumann-a. Prvo su opisane tehnike komutacije podataka i njihove najveće prednosti i mane. Zatim, da bi se što bolje razumela ideja ovog komutatora, prvo je analizirana osnovna podela kao i opšte osobine komutatora. Nakon toga je, ukratko, opisana matematika koja stoji iza ove ideje Birkhoff-a i von Neumann-a. Analizirana je originalna varijanta ovog komutatora, kao i problemi koji se javljaju u njoj. Primetilo se da kada se na ulaz BvN komutatora dovede uniformni saobraćaj, dolazi do smanjenja složenosti algoritma dekompozicije. Predložena je upotreba balansa saobraćaja tj. dodavanje još jednog stepena osnovnoj verziji BvN komutatora. Sada se u ovom, dvostepenom, komutatoru u prvom stepenu vršilo balansiranje saobraćaja dok se u drugom stepenu vršila komutacija paketa. Primetilo se da postoji problem ispadanja paketa iz redosleda pa se pristupilo njegovom rešavanju. Zatim, u ovom radu je opisano sedam različitih šema, odnosno varijacija BvN komutatora koje rešavaju ovaj problem. Opisan je njihov princip rada, kao i njihove prednosti i mane.

U poslednje vreme se sve više upotrebljavaju optički kablovi za prenos podataka, zbog svojih izuzetnih osobina, najviše na polju protoka. Pošto „klasični“ komutari (među kojima je i BvN) koji su trenutno dominantni ne mogu da obavljaju komutaciju podataka u svetlosnom domenu, trenutno je celokupna arhitektura računarskih mreža (samim tim i interneta) dodatno zakomplikovana upotrebom raznih opto-električnih konvertora i sl. S obzirom na to da se sve više teži razvijanju optičkih mreža, da bi njihova primena bila najefektivnija potrebna je upotreba optičkih komutatora i smatram da je u njima budućnost. Istraživanja na temu optičkih komutatorasu u zamahu i nije isključeno da će jednog dana celokupan prenos podataka preći u optički domen.

## LITERATURA

- [1] Zoran Čiča, „*Komutacioni sistemi*“, 2013.
- [2] H. Jonathan Chao And Bin Liu, *High Performance Switches And Routers*, 2007 by John Wiley & Sons, Inc.
- [3] Cheng-Shang Chang, Duan-Shin Lee, and Chi-Yao Yue, „Providing Guaranteed Rate Services in the Load Balanced Birkhoff-von Neumann Switches“ IEEE INFOCOM 2003.
- [4] Yanming Shen, Shivendra S. Panwar, and H. Jonathan Chao, „Design and Performance Analysis of a Practical Load-Balanced Switch“ , *IEEE TRANSACTIONS ON COMMUNICATIONS*, PAGE 2420, VOL. 57, NO. 8, AUGUST 2009.
- [5] Cheng-Shang Chang, Duan-Shin Lee, and Ying-Ju Shih „Mailbox Switch: A Scalable Two-stage Switch Architecture for Conflict Resolution of Ordered Packets“ IEEE INFOCOM 2004.
- [6] [https://en.wikipedia.org/wiki/Head-of-line\\_blocking](https://en.wikipedia.org/wiki/Head-of-line_blocking)
- [7] C.-S. Chang, D.-S. Lee and C.-M. Lien, “Load balanced Birkhoff-von Neumann switch, part II: Multi-stage buffering”, *Computer Communications*, Vol. 25, pp. 623-634, 2002.
- [8] Cheng-Shang Chang, Duan-Shin Lee, and Ying-Ju Shih, „Mailbox Switch: A Scalable Two-stage Switch Architecture for Conflict Resolution of Ordered Packets“, IEEE INFOCOM 2004.
- [9] C. S. Chang, D. Lee, and Y. Jou, “Load balanced Birkhoff-von Neumann switches. Part I: One-stage buffering,” *Computer Communications*, vol. 25, no. 6, pp. 611–622 (2002).
- [10] I. Keslassy and N. McKeown, “Maintaining packet order in two-stage switches,” in *Proc. IEEE INFOCOM’02*, New York, vol. 2, pp. 1032–1041 (June 2002).
- [11] I. Keslassy, S. Chuang, K. Yu, D. Miller, M. Horowitz, O. Solgaard, and N. McKeown, “Scaling Internet routers using optics,” in *Proc. ACM SIGCOMM*, Karlsruhe, Germany, pp. 189–200 (Aug. 2003).
- [12] Y. Shen, S. Jiang, S. Panwar, and H. J. Chao, “Byte-Focal: a practical load balanced switch,” in *Proc. IEEE Workshop on High Performance Switching and Routing (HPSR’05)*, Hong Kong, pp. 6–12 (May 2005).