



UNIVERZITET U BEOGRADU
ELEKTROTEHNIČKI FAKULTET

NEMANJA M. NINKOVIĆ

**NOVA REŠENJA ZA POBOLJŠANJE
KVALITETA SERVISA U INTRADOMENSKOM I
INTERDOMENSKOM MREŽNOM OKRUŽENJU**

DOKTORSKA DISERTACIJA

BEOGRAD 2016.



UNIVERSITY OF BELGRADE
SCHOOL OF ELECTRICAL ENGINEERING

NEMANJA M. NINKOVIĆ

**NOVEL APPROACHES FOR INTRADOMAIN
AND INTERDOMAIN QUALITY OF SERVICE
IMPROVEMENT**

DOCTORAL DISSERTATION

BELGRADE 2016.

Mentor:

doc. dr Slavko Gajin, docent, Elektrotehnički fakultet, Univerzitet u Beogradu

Članovi komisije:

prof. dr Zoran Jovanović, redovni profesor, Elektrotehnički fakultet, Univerzitet u Beogradu

prof. dr Mirjana Stojanović, vanredni profesor, Saobraćajni fakultet, Univerzitet u Beogradu

Datum odbrane:

ZAHVALNICA

Neizmernu zahvalnost dugujem svojim roditeljima koji su me usmerili na pravi put i omogućili da realizujem ovu doktorsku disertaciju.

Veliku zahvalnost upućujem svim svojim saradnicima i prijateljima koji su direktno ili indirektno bili deo mog naučno-istraživačkog rada .

NOVA REŠENJA ZA POBOLJŠANJE KVALITETA SERVISA U INTRADOMENSKOM I INTERDOMENSKOM MREŽNOM OKRUŽENJU

Nemanja M. Ninković

REZIME

U ovoj disertaciji su predloženi i analizirani alternativni pristupi ostvarivanja *Quality of Service* (QoS) na intradomenskom i interdomenskom mrežnom nivou. Kao prateća pojava razvoja mreže i konsolidacije različitih servisa preko jedinstvene mrežne infrastrukture, neizbežna je i pojava specifikacije različitih QoS zahteva. Trenutna QoS rešenja se dominantno zasnivaju na DiffServ arhitekturi koja je usmerena na garantovanje performansi u uslovima zagušenja. Međutim, česta je pojava degradacije performansi kao rezultat različitih grešaka i otkaza pojedinih delova mreže koji ne spadaju u zagušenje. Sa druge strane, primena algoritama QoS rutiranja ima inherentne probleme sa skalabilnošću i kompleksnošću, pa je postizanje determinističkih performansi i dalje otvoren problem. Realizacija interdomenskog QoS predstavlja još kompleksniji problem kada se moraju usaglasiti mere performansi i mapiranje klasa servisa između domena na servisnoj putanji. Aktuelni bilateralni model ugovaranja servisa između domena nije adekvatan za realizaciju interdomenskog QoS s obzirom da proces ugovaranja postaje neefikasan sa porastom broja domena na servisnoj putanji, dok specifikacija interdomenskih relacija predstavlja otvoreno pitanje.

Primena paketske disperzije se u ovom radu detaljno analizira posmatrajući različite strategije disperzije i njihov uticaj u uslovima postojanja rafalnih gubitaka paketa izraženih na više putanja u mreži. Koristeći postojanje više putanja, paketska disperzija vrši slanje uzastopnih paketa preko različitih putanja i posledično utiče na strukturu rafalnih gubitaka paketa. Prethodne analize paketske disperzije su obuhvatale posmatranje rastojanja gubitaka paketa kada se rafalni gubici modeluju Markovljevim modelom sa 2 stanja. Naučni doprinos u ovoj disertaciji sa aspekta paketske disperzije se ogleda u realizaciji komparativne analize uticaja strategija na rafalne gubitke modelovane Markovljevim modelom sa 4 stanja koji predstavlja tačniji model u odnosu na prethodno navedeni. Dodatno, kao verifikaciju superiornih rezultata u odnosu na konvencionalno rutiranje preko jedne putanje, prikazani su rezultati praktičnog testiranja strategija, a takođe su identifikovani i scenariji u okviru kojih primena paketske disperzije ima ograničenu efikasnost na rafalne gubitke. Priložena

analiza potvrđuje ostvarivanje značajnih poboljšanja QoS koja direktno zavise od primenjene strategije disperzije.

U kontekstu realizacije interdomenskog QoS, naučni doprinos ove disertacije se ogleda u predlogu i analizi dva algoritma mapiranja klasa servisa između domena razmatrana u okviru *third party* (3P) modela ugovaranja servisa – *3P Integer Programming* (3PIP) i *Goal Programming-based Mapping Scheme* (GPMS). Proces mapiranja klasa servisa u 3P modelu se ostvaruje posredstvom 3P agenta koji realizuje prethodno navedene algoritme i vrši mapiranje klasa servisa između domena na servisnoj putanji. Predloženi algoritmi unose značajan nivo fleksibilnosti i kontrole procesa mapiranja s obzirom da je moguće specificirati zahtevane QoS parametre i mapiranje bazirati na predodređenim poslovnim i ekonomskim politikama s obzirom da se kao dodatni ulazni parametar posmatra i cena interkonekcije pridružena selektovanoj klasi servisa. Komparativna analiza postojećeg *Policy-Conformance Matching Scheme* (P-CMS) algoritma sa predloženim algoritmima je realizovana kroz posmatranje selekcije klasa servisa sa aspekta QoS parametara i cene interkonekcije. Rezultati ove analize ukazuju da je sa predloženim algoritmima moguće izvršiti tačnije mapiranje, a zahvaljujući *Integer Programming* i *Multi-Goal Programming* modelima je moguće realizovati i poslovne politike sa ciljem maksimizacije profita domena koji učestvuju u realizaciji interdomenskog QoS.

Ključne reči: mapiranje klasa servisa, paketska disperzija, *Voice over IP*, *Quality of service*

Naučna oblast: Elektrotehnika

Uža naučna oblast: Računarska tehnika i informatika

UDK: 621.3

NOVEL APPROACHES FOR INTRADOMAIN AND INTERDOMAIN QUALITY OF SERVICE IMPROVEMENT

Nemanja M. Ninković

ABSTRACT

This dissertation contains proposal and subsequent analysis of alternative approaches to achieving Quality of Service (QoS) on the intradomain and interdomain levels. Specification of various QoS requirements is inevitable as a result of networking growth and service consolidation over a single network infrastructure. Current QoS solutions are generally based on DiffServ architecture, which aims to guarantee performance in the event of bandwidth congestion. However, performance degradation often occurs as a result of various faults and outages of network sections not related to congestion problems. Conversely, application of QoS routing algorithms carries inherent scalability limitations and computational complexity regardless of applied redundancy, making the problem of achieving predictable performance still unresolved. Interdomain QoS delivery presents more complex problem since it is necessary to coordinate multiple domains located on the service path in terms of performance metrics and class mapping. Current bilateral service negotiation model is not suitable for interdomain QoS delivery as it is unable to reconcile efficiency and scalability leaving the interdomain relations problem opened.

The application of packet dispersion is analysed through investigation of dispersion strategies and their immediate effect on burst packet losses expressed on several network paths. Packet dispersion relies on path diversity in order to transport subsequent packets over disjoint paths resulting in pattern modification of burst losses, which have devastating effect on QoS critical applications. Previous packet dispersion analysis encompassed effect on distance between subsequent losses modelled using Markov model with 2 states. The contribution from the aspect of packet dispersion discusses comparative analysis of dispersion strategies and their effect on burst losses modelled using Markov model with 4 states, which is actually more accurate model than its 2-state counterpart. The results of real network testing are provided as a verification of packet dispersion superiority in comparison to conventional single-path routing. Furthermore, scenarios have been identified for which packet

dispersion has limited efficiency on burst packet losses. Conducted analysis and tests indicate significant QoS improvement that is correlated with applied dispersion strategy.

From the aspect of delivering interdomain QoS, mayor contribution refers to proposal and analysis of two mapping algorithms applied in third party (3P) negotiation model - 3P Integer Programming (3PIP) i Goal Programming-based Mapping Scheme (GPMS). Class mapping process in 3P negotiation model is implemented using trusted entity, i.e. 3P agent, which implements proposed mapping algorithms between domains on the service path. Proposed mapping algorithms introduce significant level of flexibility and mapping control as a result of the ability to specify QoS parameters and class mapping based on predetermined business policies. Previous flexibility is achieved since interconnection cost is accounted as an additional input parameter associated with the specific service class. Comparative analysis of the existing Policy-Conformance Matching Scheme (P-CMS) algorithm and proposed mapping algorithms has been carried out in order to investigate service class selection from the perspective of QoS parameters and interconnection cost. The results of this analysis point out that using proposed mapping algorithms it is possible to perform more accurate mappings, while underlying Integer Programming and Multi-Goal Programming models allow incorporation of business policies aiming to maximize profit of domains, which are involved in the interdomain QoS delivery.

Keywords: service class mapping, packet dispersion, quality of service, voice over IP

Scientific field: Electrical Engineering

Specific scientific area: Computer Networks

UDK: 621.3

SADRŽAJ

1	UVOD	1
2	DEFINICIJA KVALITETA SERVISA I MERA PERFORMANSI	5
2.1.	Koncept kvaliteta servisa	6
2.2.	Pregled pokazatelja mrežnih performansi	8
2.2.1.	Kašnjenje.....	9
2.2.2.	Džiter.....	9
2.2.3.	Gubici paketa	10
2.2.4.	Protok.....	11
2.3.	Definisanje klasa servisa.....	11
2.4.	DiffServ arhitektura i implementacija QoS politika	14
2.5.	Klasifikacija paketa i ograničenje protoka.....	15
2.6.	Algoritmi za kontrolu zagušenja	16
2.7.	Algoritmi za kontrolu opsluživanja redova.....	18
2.8.	Algoritmi za izbegavanje zagušenja – <i>Random Early Detection</i>	20
2.9.	Izazovi u implementaciji QoS politike i zaključna razmatranja	22
3	UTICAJ STRATEGIJA PAKETSKE DISPERZIJE NA QOS UNUTAR DOMENA	24
3.1.	Politike rutiranja unutar domena.....	25
3.2.	Značaj mrežne raspoloživosti.....	27
3.3.	Distribucija saobraćaja preko više putanja.....	29
3.3.1.	Mehanizmi deljenja saobraćaja na paketskom nivou	30
3.3.2.	Mehanizmi deljenja saobraćaja na flow nivou	31
3.4.	Modelovanje i relevantni parametri gubitaka paketa	32
3.4.1.	Modelovanje gubitaka paketa posredstvom Bernulijevog modela	36
3.4.2.	Modelovanje gubitaka paketa posredstvom Gilbertovog modela	37
3.4.3.	Modelovanje gubitaka paketa posredstvom Gilbert-Eliotovog modela	39
3.4.4.	Modelovanje gubitaka paketa posredstvom Markovljevog modela sa 4 stanja	40
3.4.5.	Mapiranje Gilbert-Eliotovog modela i Markovljevog modela sa 4 stanja	42
3.5.	Uticaj QoS parametara na kvalitet VoIP servisa.....	43
3.5.1.	Paketizacija govornog signala.....	44
3.5.2.	Detekcija aktivnosti govornog signala (VAD) i algoritmi za sakrivanje gubitaka paketa (PLC)	47
3.6.	Metode ocene kvaliteta VoIP servisa.....	49
3.6.1.	Primena ITU-T E-Modela.....	50

3.6.2.	Uticaj gubitaka paketa u E-Modelu.....	52
3.6.3.	MOS ocena kvaliteta.....	54
3.6.4.	Estimacija rafalnih gubitaka iz dolaznog niza paketa	55
3.7.	Analiza rastojanja gubitaka u VoIP aplikacijama	57
3.7.1.	NLR analiza u prisustvu Gilbertovih i Gilbert-Eliotovih gubitaka	59
3.7.2.	NLR analiza u prisustvu gubitaka modelovanih Markovljevim modelom sa 4 stanja..	62
3.8.	Princip rada paketske disperzije.....	63
3.8.1.	Relevantni radovi iz oblasti paketske disperzije	66
3.9.	Strategije paketske disperzije.....	70
3.9.1.	Round robin strategija.....	71
3.9.2.	Random strategija	72
3.9.3.	Adaptivna strategija	73
3.9.4.	Redundantna strategija.....	74
3.10.	Analiza rastojanja gubitaka kod strategija u prisustvu Gilbert-Eliotovih gubitaka paketa	75
3.10.1.	NLR analiza kod round robin strategije	75
3.10.2.	NLR analiza kod random i adaptivne strategije	78
3.10.3.	NLR analiza kod redundantne strategije	79
3.11.	Analiza rastojanja gubitaka kod strategija u prisustvu gubitaka paketa sa Markovljevim modelom sa 4 stanja.....	80
3.11.1.	NLR analiza kod round robin strategije	80
3.11.2.	NLR analiza kod random i adaptivne strategije	82
3.11.3.	NLR analiza kod redundantne strategije	84
3.12.	Uticaj <i>BurstR</i> parametra na strategije disperzije	85
3.13.	Dinamička aktivacija paketske disperzije	86
3.13.1.	Predložena šema aktivacije paketske disperzije.....	87
3.13.2.	Analiza aktivacije paketske disperzije	90
3.13.3.	Aktivacije paketske disperzije posredstvom zauzetosti bafera	93
4	Realizacija interdomenskog QoS	95
4.1.	Problemi i pristupi garantovanju performansi izvan granica domena.....	96
4.2.	Tehnička realizacija interdomenskog QoS	97
4.3.	Definicija klasa servisa i servisnog zahteva.....	100
4.4.	Modeli ugovaranja servisa	100
4.4.1.	Bilateralni model.....	101
4.4.1.	Kaskadni model.....	102
4.4.2.	Third party (3P) model.....	104
4.5.	Modeli interdomenskog tarifiranja.....	107

4.6. Mapiranje klasa servisa između domena	109
4.7. Algoritmi za mapiranje klasa servisa između domena 3P modela ugovaranja servisa	110
4.8. P-CMS algoritam mapiranja klasa servisa	111
4.9. 3PIP algoritam mapiranja klasa servisa	116
4.10. GPMS algoritam mapiranja klasa servisa	119
4.11. Kvalitativna komparacija algoritama za mapiranje klasa između domena	120
5 Analiza predloženih intradomenskih i interdomenskih QoS mehanizama	122
5.1. Komparativna analiza strategija paketske disperzije	122
5.1.1. Verifikacija rezultata predloženog analitičkog modela.....	124
5.1.2. Analiza rastojanja gubitaka paketa.....	126
5.1.3. Analiza VoIP kvaliteta	129
5.2. Analiza algoritama mapiranja klasa servisa između domena.....	133
6 Zaključak.....	139
Reference	142

1

UVOD

Kao direktna posledica razvoja i zastupljenosti konvergiranih mreža (*converged networks*) javlja se izražena potreba za implementacijom različitih *Quality of Service* (QoS) mehanizama koje suštinski predstavljaju preduslov za korišćenje jedinstvene infrastrukture za realizaciju servisa sa različitim potrebama u pogledu zahtevanih performansi. S obzirom na širok opseg servisnih zahteva i mogućih tehničkih rešenja, QoS arhitektura mora da zadovolji sve postavljene zahteve kako bi se omogućila realizacija različitih servisa preko konvergiranih mreža. U tom smislu, mogu se razlikovati intradomenski QoS mehanizmi, koji se sastoje u preduzimanju tehničkih koraka u cilju uspostavljanja zahtevanog nivoa performansi u granicama posmatranog domena, i interdomenski QoS mehanizmi, koji podrazumevaju implementaciju garancija performansi i izvan granica posmatranog domena kada se pomatra scenario sa više domena.

Intradomska QoS arhitektura obuhvata širok spektar mehanizama za prioritetizaciju i podelu saobraćaja na klase u cilju realizacije determinističkih performansi u okviru jednog domena kada su parametri, kao što su kašnjenje, varijacija kašnjenja i gubici paketa, iznad očekivanog nivoa. Prioritetizacija saobraćaja se po pravilu obavlja u svakom čvoru mreže (ruteru) i podrazumeva definisanje i projektovanje QoS politike koja treba biti dosledno primenjena na čitav domen. Posledično, na nivou domena mora postojati odgovarajuća doslednost prilikom definisanja QoS mehanizama koji će biti realizovani. Upravo je ta potreba za visokim nivoom doslednosti i zahtevom za detaljnim planiranjem QoS politika razlog što administratori i mrežni inženjeri u pojedinim slučajevima preskoče uvođenje prioritetizacije saobraćaja i servisa, pa nekada saobraćaj koji bi trebalo da bude rangiran visokim prioritetom bude tretiran kao *best effort* klasa. Još kada se uzme u obzir da pored QoS mehanizama koji su standardizovani, vendori dodaju i neke dodatne funkcionalnosti, postaje jasno da takva mešavina brzo može da dovede do problema sa

kompleksnošću i da prednosti koje donosi implementacija QoS arhitekture brzo može da bude zamagljena problemima u projektovanju i implementaciji.

Razmatranjem problema koji se mogu pojaviti prilikom realizacije intradomenskog QoS postaje jasno da je garancija performansi u višedomenskim scenarijima još kompleksniji problem. Realizacija interdomenskog QoS mora najpre da bude započeta na administrativnom nivou, sagledavanjem ekonomskih ciljeva i odnosa sa drugim provajderima, odnosno domenima. U tom smislu, potrebno je kombinovati intradomenske QoS mehanizme i interdomenske algoritme za mapiranje klasa servisa koji su posledica QoS politike definisane od strane pojedinačnih domena. Primenom algoritama za mapiranje klasa servisa između pojedinačnih domena moguće je ostvariti zahtevani nivo interdomenskog QoS sa osvrtnom na garanciju performansi od domena inicijatora do odredišnog domena - *end-to-end* (E2E). Dodatni problem interdomenskog QoS javlja se kod saradnje između pojedinačnih domena. Naime, domeni predstavljaju provajdere koji imaju konkurentne ciljeve, pa je i saradnja u ovom kontekstu problematična. Jedan od aktuelnih pristupa u rešavanju ovog izazova jeste uvođenje *third party* (3P) agenta koji bi vršio pomenuto E2E mapiranje klasa servisa. 3P pristup istovremeno rešava probleme interdomenske saradnje, konkurentnosti, a takođe ostavlja prostora za postizanje fer odnosa prema svim domenima koji se nalaze na E2E putanji bez favorizovanja bilo kog na osnovu ponuđenih klasa servisa.

U ovoj disertaciji se posmatra problem QoS na dva nivoa – intradomenski i interdomenski. Najpre, osnovni zahtev za intradomenski QoS jeste rešavanje problema kompleksnosti i realizacije. Imajući u vidu da postojeći mehanizmi zahtevaju doslednu i standardizovanu implementaciju u čitavom domenu, predlaže se modifikovanje politike rutiranja u okviru domena koji će poboljšati performanse servisa koji su osetljivi na gubitke paketa. Naime, u prisustvu gubitaka paketa na jednoj ili više putanja u okviru mreže, dolazi do manje ili više primetnog uticaja na kvalitet servisa kod korisnika. Ovo je naročito izraženo kod *voice* i *video* servisa koje je u određenim implementacijama potrebno realizovati u realnom vremenu. Konkretni pristup koji se koristi prilikom modifikacije politike rutiranja i koji se ispituje u ovom radu je paketska disperzija. Paketska disperzija koristi postojanje više raspoloživih putanja do odredišta da prenese pakete kako bi se u slučaju gubitaka paketa na pojedinačnim putanjama ostvarilo veće rastojanje između gubitaka paketa, i posledično, poboljšao perceptivni kvalitet servisa koji su osetljivi na te gubitke. Tipični predstavnik takvog servisa jeste *Voice over IP* (VoIP) na čijem primeru se pokazuju prednosti koje se ostvaruju primenom različitih strategija paketske disperzije, a pre svega povećanje robusnosti servisa u uslovima rafalnih gubitaka paketa istovremeno izraženih na više putanja.

U cilju razmatranja kompletnog skupa QoS problematike, u ovom radu se analizira i 3P model ugovaranja servisa koji adekvatno može da odgovori na izazov koji multidomenski scenariji postavljaju prilikom potrebe za garancijom E2E performansi. Detaljno su analizirane prednosti i mane 3P modela u odnosu na postojeći bilateralni model ugovaranja saobraćaja, kao i na kooperativni model za koji i dalje nema standardnog pristupa u implementaciji. 3P agent je analiziran sa aspekta funkcionalnih celina koje mora da

podržava kako bi mogao da ostvari kako tehničke, tako i ekonomske ciljeve. U ovom radu je 3P model posmatran sa stanovišta implementacije nekoliko algoritama mapiranja klasa servisa gde je pružen detaljan osvrt na nedostatke i prednosti individualnih pristupa u rešavanju procesa mapiranja klasa servisa između domena. Osim toga, data je komparativna analiza za postojeći algoritam i dva predložena algoritma mapiranja.

Naučni doprinosi ove disertacije se mogu sumariizovati po intradomenskom i interdomenskom QoS. Glavni doprinosi za intradomenski QoS su vezani za primenu paketske disperzije: (1) matematički model za izračunavanje rastojanja gubitaka u slučaju različitih strategija paketske disperzije kada su rafalni gubici modelovani Markovljevim modelom sa 4 stanja; (2) Verifikacija rezultata dobijena praktičnim testiranjem strategija paketske disperzije; (3) Komparativna analiza strategija paketske disperzije sa aspekta VoIP kvaliteta servisa.

Značajni doprinosi za interdomenski QoS se ogledaju u razmatranju procesa mapiranja klasa servisa u 3P modelu ugovaranja servisa: (1) Predlog i analiza 3P *Integer Programming* (3PIP) algoritma mapiranja; (2) Predlog i analiza *Goal Programming Matching Scheme* (GPMS) algoritma mapiranja; (3) Komparativna analiza postojećeg *Policy Conformance Matching Scheme* (P-CMS), 3PIP i GPMS algoritama sa aspekta tačnosti mapiranja klasa servisa posredstvom 3P modela ugovaranja.

Disertacija je organizovana u šest poglavlja. Drugo poglavlje sadrži analizu osnovne QoS definicije kao i pregled postojećih QoS mehanizama koji se primenjuju u današnjim konvergiranim mrežama. Pregledom postojećih QoS mehanizama se daje osvrt na prednosti i nedostatke istih, kao i na eventualna poboljšanja koja se mogu preduzeti u cilju fleksibilnije i efikasnije implementacije QoS na nivou jednog domena.

Treće poglavlje predstavlja jedan od centralnih doprinosa ove disertacije s obzirom da uvodi proces paketske disperzije kao metode izmene politike rutiranja sa jednostavnom implementacijom bez značajnih promena u mreži. Naučni doprinos se ogleda u razmatranju Markovljevog modela sa 4 stanja u odnosu na prethodne analize koje su posmatrale uticaj paketske disperzije na rafalne gubitke modelovane Markovljevim modelom sa 2 stanja, čime se nije mogla u potpunosti sagledati efikasnost paketske disperzije na strukturu rafalnih gubitaka. Posmatra se *Voice over IP* (VoIP) servis se kao reprezentativan pokazatelj uticaja na percepirani kvalitet u slučaju primene paketske disperzije. Dodatni doprinos sadržan u ovom poglavlju se ogleda i predlogu i matematičkoj analizi novih strategija paketske disperzije.

Četvrto poglavlje sadrži detaljan opis problema realizacije interdomenskog QoS kao i modele ugovaranja servisa između domena. Apostrofiraju se prednosti i mane postojećih modela ugovaranja servisa kao što su bilateralni i kooperativni model, dok se 3P model analizira sa aspekta rešavanja nedostataka njegovih prethodnika. Naučni doprinos sa aspekta realizacije interdomenskog QoS se ogleda u analizi zahteva algoritama mapiranja klasa servisa i predlogu dva algoritma mapiranja klasa servisa između domena

namenjena za primenu u 3P modelu ugovaranja. Predloženi algoritmi koriste matematičke modele kojima se može postići granularnija kontrola, veća tačnost i fleksibilnost procesa mapiranja klasa servisa u odnosu na postojeći 3P algoritam mapiranja.

Peto poglavlje sadrži rezultate komparativnih analiza strategije paketske disperzije sa aspekta rastojanja gubitaka paketa kao i uticaj pojedinačnih strategija disperzije na objektivni kvalitet VoIP servisa. Sa druge strane, data je i komparativna analiza algoritma mapiranja klasa servisa posmatrajući pojedinačne QoS parametre i cenu interkonekcije. Ovim komparativnim analizama se lako može uočiti superiornost paketske disperzije u odnosu na konvencionalno rutiranje u prisustvu rafalnih gubitaka paketa, kada se razmatra intradomenski QoS, odnosno superiornost predloženih algoritma mapiranja klasa servisa u odnosu na postojeći 3P algoritam kada se razmatra interdomenski QoS.

Šesto poglavlje sadrži zaključna razmatranja kao i eventualne pravce budućeg istraživanja koji se otvaraju na osnovu dobijenih rezultata analiza koje su bile predmet ovde disertacije.

2

DEFINICIJA KVALITETA SERVISA I MERA PERFORMANSI

Koncept kvaliteta servisa (QoS) u modernim telekomunikacionim mrežama se vezuje za ostvarivanje određenog nivoa performansi kroz posmatranu mrežu ili mreže. Potreba da se korisnicima osigura odgovarajući nivo performansi predstavlja značajan aspekt upravljanja mrežama, pa je i samim tim uloga QoS mehanizama postala presudna u mrežama naredne generacije (*Next Generation Networks* - NGN). Analiza ostvarenih performansi u mreži će imati presudnu ulogu za razvijanje novih servisa kao i definisanje poslovnih i ekonomskih procesa u okviru servis provajderskih mreža. U cilju realizacije neophodnog QoS, administratori moraju da se oslanjaju na skalabilnost postojeće infrastrukture. Ukoliko se pokaže da posmatrana infrastruktura nije u stanju da podrži dodatni saobraćaj, servisne zahteve i povećanje broja korisnika, potrebno je dodatno investirati u implementaciju proširenja infrastrukture (mrežni uređaji, zakup dodatnih veza i kapaciteta), novih tehnologija (u cilju boljeg iskorišćenja postojeće infrastrukture) kao i dodatno ulaganje u procese koji se odnose na upravljanje mrežom. Posledično, svaka nova implementacija će zahtevati značajna ulaganja kako bi se uspešno odgovorilo na nove servisne i korisničke zahteve.

U cilju zaštite investicija, servis provajderi i administratori velikih korporativnih mreža (*enterprise networks*) se okreću efikasnijim upravljanjem kako bi se ostvarilo veće iskorišćenje postojećih resursa. To upravljanje se ogleda kroz postavljanje ograničenja količine saobraćaja koji se može primiti u mrežu kao i

kroz prioritetizaciju pojedinih saobraćajnih klasa odnosno servisa. Na taj način se postiže da vrednosti pojedinih parametara kojima se opisuju performanse u paketskim mrežama drže iznad određenog praga u cilju ostvarivanja adekvatnog QoS za pojedine korisnike i raspoloživosti posmatranog servisa. Prioritetizacije paketa i ostvarivanja planiranog nivoa performansi je kompleksan proces koji zavisi od primene velikog broja mehanizama sa komplementarnim funkcijama, a sve u cilju postizanja determinističkih performansi paketa.

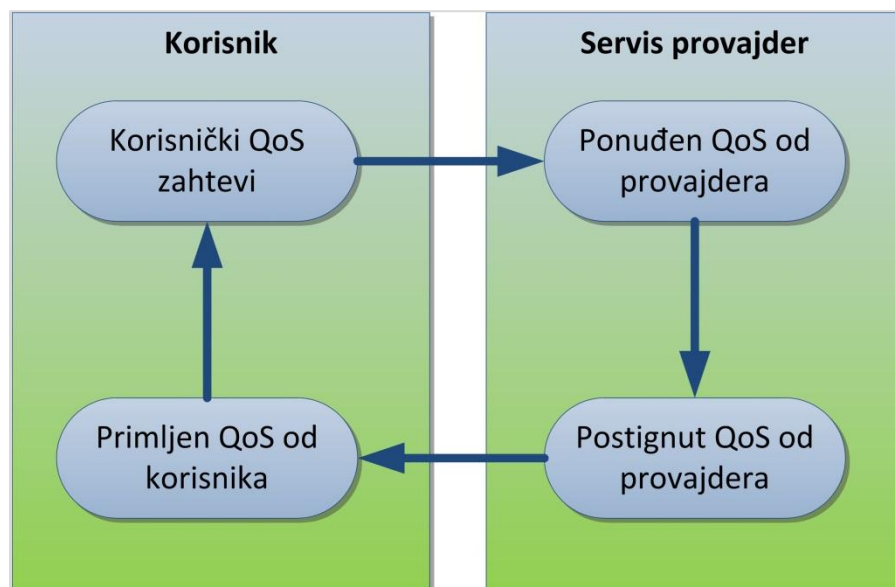
Iako se iz prethodnog moglo pretpostaviti da je implementacija QoS dovoljan preduslov za postizanje potrebnih *end-to-end* (E2E) performansi, postoje značajni izazovi koji sprečavaju efikasnu implementaciju QoS u mrežama. Sa E2E aspekta i različitih mrežnih uređaja u okviru kojih se ostvaruje implementacija QoS mehanizama, javlja se značajan problem kontrole čitavog QoS procesa. Naime, na mrežnim uređajima (ruteri, svičevi, itd.) se mogu implementirati neadekvatni QoS mehanizmi kao posledica nepoznavanja karakteristike saobraćaja ili pojedinih servisa. Dodatno, usled odsustva doslednosti u implementaciji QoS politike može doći do neočekivanih rezultata QoS parametara kao što su kašnjenje, varijacija kašnjenja i gubici paketa, što može imati negativan uticaj na deo ili čitavu mrežu. Takođe, primena ovih QoS mehanizama omogućava bolje iskorišćenje postojećih resursa, ali se ne sprečavaju greške u prenosu paketa koji se mogu javiti kao posledica velikog broja uzroka, što se konačno manifestuje degradacijom performansi.

Kako bi se bolje sagledao QoS koncept potrebno je najpre definisati parametre koji se koriste kako bi se opisale performanse servisa, a na osnovu čijih vrednosti se može analizirati da li je kroz mrežu realizovan potreban nivo performansi. Takođe, u nastavku će biti opisani aktuelni QoS mehanizmi u okviru mrežnih uređaja koji se koriste za postizanje diferencijacije servisa i ostvarivanje ciljnih performansi u uslovima zagušenja.

2.1. Koncept kvaliteta servisa

Posmatrajući QoS može se reći da postoje dve ciljne grupe koje poseduju različitu perspektivu na nivo mrežnih performansi kao što je definisano u [1]. Relacija tih ciljnih grupa i njihova perspektiva na QoS je prikazana na Slika 2.1. Naime, prvi je korisnik koji postavlja servisne zahteve i koje generalno zavise od aplikacije i samog servisa koji korisnik želi da koristi. Osim toga, korisnik će na osnovu QoS koji je dobio od provajdera da utvrdi da li je nivo performansi dovoljan imajući u vidu inicijalne servisne zahteve. Drugoj ciljnoj grupi pripada servis provajder koji na osnovu svojih resursa, tehnologije i raspoloživih kapaciteta nudi QoS u zavisnosti od servisnih zahteva. Provajder će na osnovu ove ponude i formirati cenu koja će biti ispostavljena korisniku. Na osnovu operativnog stanja i raspoloživosti mreže, servisa provajder će ponuditi QoS koji se generalno može razlikovati od očekivanog QoS nivoa od strane korisnika, a i od planiranog nivoa od strane provajdera. Ono što je kritično i za korisnika i provajdera jeste da QoS nivo koji je postignut od strane provajdera i QoS nivo koji je korisnik očekivao budu bliski.

Ako se dalje razmatra korisnička perspektiva, činioci u QoS kontekstu koji će značajno uticati da korisnik plati provajderu za odgovarajući servis mogu se posmatrati sa tri strane. Prvi činilac je unutrašnji QoS koji se postiže odgovarajućim planiranjem i projektovanjem same mrežne infrastrukture, tehnologije i kapaciteta. Adekvano planirana skalabilnost i izbor transportne tehnologije u okviru mreže provajdera određuje da li postoji dovoljno resursa da se odgovori na servisne zahteve korisnika.



Slika 2.1. Relacija korisnika i servis provajdera iz perspektive QoS [1].

Drugi činilac predstavlja tzv. perceptivni QoS koji u zavisnosti od konteksta servisa za koji se posmatra se može opisivati različitim parametrima i njihovim graničnim vrednostima. Ovaj QoS suštinski zavisi od korisničke percepcije unutrašnjeg QoS koji je realizovan od strane provajdera. Dodatno, u oceni perceptivnog QoS može doći do izražavanja subjektivnog očekivanja korisnika po pitanju performansi, a koja su po pravilu u korelaciji sa performansama sličnih servisa. Ova subjektivnost može biti i određena od strane samog proizvođača servisa i komparacije istog sa sličnim servisima sa kojima je korisnik prethodno imao iskustvo.

Treći činilac se naziva ostvareni QoS koji diktira da li će korisnik koji koristi servis smatrati da je nivo performansi dovoljno dobar za nastavak realizacije posmatranog servisa. Ostvareni QoS pre svega zavisi od perceptivnog QoS, ali i takođe od eventualnog odnosa servis provajdera i dizajnera servisa prema eventualnim problemima sa samim servisom, ukoliko postoje.

Ono što se mora naglasiti jeste da je unutrašnji QoS odgovoran da li će posmatrani servis biti atraktivan korisniku, i na kraju, da li će za isti servis korisnik platiti. Sa druge strane, percepirani QoS će diktirati da li je nivo performansi prihvatljiv za korisnika. Ostvareni QoS će odrediti da li će korisnik zadržati ponuđeni servis. Istovremeno, dobar pokazatelj postignutog QoS jeste upravo dobijeni nivo performansi

koji ocenjuje korisnik i koji generalno treba da bude prihvatljiv svim korisnicima koji koriste posmatrani servis.

U kontekstu prethodne priče o QoS, najvažnije je naći adekvatne mere kojima se može opisati kvalitet servisa, tačnije nivo performansi na koji se oslanja raspoloživost jednog servisa. U tom smislu, potrebno je da se kvantitativno opišu ove pomenute vrste QoS kako bi se mogli uporediti unutrašnji, postignuti i percipirani QoS. Naravno, kvantitativni parametri koji opisuju nivo performansi za pojedinačne servise moraju da adekvatno reflektuju kvalitativne osobine percipiranog QoS kako bi se isti mogao dovesti u vezu sa unutrašnjim QoS, a na osnovu koga će se i odrediti ostvareni QoS. Kada se definišu relevantni kvantitativni pokazatelji treba se analizirati skup servisnih atributa koji direktno utiču na formiranje percepcije o ostvarenom, i na kraju, percipiranom kvalitetu jednog servisa.

Imajući u vidu veliki broj različitih servisa i tehnologija na koji se oslanjaju, definicija mera performansi često nije trivijalna stvar. Određivanje mera performansi, a samim tim i njihovih kvantitativnih vrednosti u velikoj meri zavisi od tipa servisa za koje se razmatraju potrebne performanse (na primer, *voice* i video servisi zahtevaju različite performanse u poređenju sa *data* servisima), kao i od međusobnog odnosa vrednosti kvalitativnih pokazatelja, tj. relevantnih mera performansi. U tom smislu, najpre se mora pristupiti definiciji mera performansi, a zatim identifikaciji njihovih graničnih vrednosti koje zavise od konkretnog servisa.

2.2. Pregled pokazatelja mrežnih performansi

Kada se razmatraju mrežne performanse parametri koji na adekvatan način opisuju očekivani nivo performansi jesu kašnjenje, varijacija kašnjenja (džiter), gubitak paketa i protok (*bandwidth*). Upravo ovi parametri predstavljaju glavne pokazatelje mogućnosti jedne mreže da odgovori na servisne zahteve. Dodatno, cilj QoS mehanizama jeste da drži vrednosti ovih parametara kontrolisanim s obzirom da su presudni u određivanju ukupnih mrežnih performansi i raspoloživosti servisa. U zavisnosti od specifikacije servisa, svaki od ovih mera performansi je manje ili više značajan, pa se u tu svrhu često koristi termin *Key Performance Indicator* (KPI) kako bi se označio važnost pojedinog parametra za posmatrani servis.

Svaki od ovih parametara mrežnih performansi se treba posmatrati zasebno i izuzetno je bitno da se sagleda uticaj pojedinačnog parametra na QoS. U tom smislu, u nastavku je dat pregled pojedinačnih parametara kao i njihov direktan i indirektan uticaj na mrežne performanse kao i servisne zahteve. Glavni parametri od kojih performanse servisa zavise jesu upravo kašnjenje, džiter, verovatnoća gubitaka paketa i protok. Ove mere performansi su formalno definisane u [2].

2.2.1. Kašnjenje

Kašnjenje paketa u mrežama se po definiciji određuje kao vreme kada se paket pošalje sa izvorišta do trenutka kada stigne na odredište. Većina QoS mehanizama koja se primenjuje u današnjim mrežama se u većoj ili manjom meri odnose na kontrolisanje kašnjenja paketa, a to je posledica važnosti ovog parametara u opisivanju postignutog i percepiranog QoS.

Kašnjenje paketa je uzrok mnogobrojnih aktivnosti u mreži i uglavnom se vezuje za mrežne čvorove, tačnije uređaje u okviru kojih se paketi obrađuju. Kašnjenje je inherentno vezano za obradu paketa, pa je za minimizaciju kašnjenja paketa prilikom njegove obrade potrebno da procesi koji se realizuju unutar rutera i svičeva budu svesni prioriteta paketa i u slučaju zagušenja da se primeni prioritetizacija paketa koja je prethodno usvojena u okviru QoS politike jedne mreže.

Ukupno kašnjenje se dalje može razložiti na serijsko, propagaciono, *queueing*, procesorsko, *shaping*, mrežno, kodno i kompresiono kašnjenje. Svaki od ovih pojedinačnih aspekata kašnjenja zavisi od određenog dela mreže i može se smatrati za jednu putanju od izvorišta do odredišta fiksnim ili promenljivim. QoS mehanizmi koji se primenjuju u okviru mrežnih uređaja (mehanizmi upravljanja zagušenjem, prioritetizacija paketa, ograničenje protoka paketa, itd.) imaju za cilj da smanje *queueing* i *shaping* kašnjenje. Naravno, na ostale vrste kašnjenja administratori često ne mogu da utiču bez unošenja većih promena kao što je na primer promena politike rutiranja u cilju postizanja manjeg kašnjenja preko druge putanje. Formalna definicija kašnjenja data kao *IP packet transfer delay* (IPTD) je data u [2].

Smanjenje kašnjenja se može postići povećanjem kapaciteta mreže, odnosno povećanjem protoka. Mada, takav pristup obično nije isplativ i često je tehnički neopravdan. Međutim, povećanje kapaciteta mreže ne može uvek da reši problem pojedinih servisa sa kašnjenjem jer se povećanjem kapaciteta mogu maskirati neki drugi problemi koji dovode do problema sa servisima koji su osetljivi na QoS (*voice* i *video*). U tom smislu, daleko je efikasnije izvršiti implementaciju QoS mehanizama u cilju dobijanja determinističkih performansi pojedinih klasa saobraćaja.

2.2.2. Džiter

Varijacija kašnjenja ili džiter predstavlja značajan pokazatelj stabilnosti mreže s obzirom da se posmatra razlika u kašnjenju susednih paketa. U okviru paketskih mreža ako pomenute komponente kašnjenja imaju previse promenljiv, ta varijabilnost će dovesti i do izraženog džitera koji u određenim slučajevima može imati uticaja na rad aplikacije. Džiter ima izražen uticaj na aplikacije u realnom vremenu kao što je *Voice over IP* (VoIP) s obzirom da je za nesmetanu reprodukciju zvuka potrebno da *voice* paketi dolaze do odredišta sa jednakim vremenskim rastojanjem. Upravo varijacija u kašnjenju tih paketa može da promeni VoIP i performanse drugih multimedijalnih servisa koji su generalno osetljivi na varijabilnost kašnjenja paketa. Formalna definicija džitera je data kao *IP delay variation* (IPDV) u [2].

Iako je analizom kašnjenja utvrđeno da je postojanje promenljivih komponenti kašnjenja neizbežno, razvoj multimedijalnih aplikacija je doveo do pojave mehanizama koje eliminišu negativan uticaj džitera na ove aplikacije. To pre svega podrazumeva implementaciju tzv. džiter bafera, a sam process se naziva dedžiterizacija. Baferovanjem susednih paketa se postiže eliminacija nejednakog kašnjenja između paketa, ali se istovremeno i povećava ukupno kašnjenje. Dok god su i kašnjenje i džiter kontrolisani, neznatna degradacija ovih parametara neće dovesti do neraspoloživosti ovih servisa.

Osim dedžiterizacije, implementacija QoS mehanizama u ruterima može znatno da smanji *queueing* kašnjenje, dok se prioritizacijom postiže da paketi aplikacija koji su osetljivi na kašnjenje budu obrađeni značajno brže u ruterima, pa samim tim i promenljiva komponenta kašnjenja kao posledica čekanja u odlaznim redovima čekanja u ruterima nije značajna. Međutim, kao i kod kašnjenja povećanje kapaciteta u mrežama može dovesti do smanjenja džitera, ali ga ne može u potpunosti eliminisati. Iz tog razloga, krajnje tačke u kojima se terminira saobraćaj namenjen za aplikacije osetljive na kašnjenje i džiter takođe implementiraju džiter bafere.

2.2.3. Gubici paketa

Gubici paketa generalno mogu da dovedu do osetne promene u performansama i raspoloživosti servisa. Mnogo je razloga koji mogu dovesti do gubitaka paketa. Najčešći razlozi jesu odbacivanje pojedinačnih paketa u okviru rutera kao posledica zagušenja ili greške u prenosu. Osim toga, mogu se desiti različite konfiguracione greške kao i greške u rutiranju koje rezultiraju odbacivanjem ili gubljenjem paketa.

Kod odbacivanja paketa u ruterima usled zagušenja mogu se koristiti QoS mehanizmi u cilju smanjenja uticaja ovog odbacivanja paketa i vršenje selektivnog odbacivanja kako bi se ispoštovala primenjena prioritizacija. Sa odbacivanjem paketa u okviru rutera kao posledica zagušenja često se vezuje termin *tail drop* koji podrazumeva istovremeno odbacivanje više susednih paketa koji pripadaju različitim sesijama i saobraćajnih tokova usled nedostatka memorijskog prostora u baferu – odlaznom redu čekanja. Ovakva vrsta odbacivanja paketa je izuzetno pogubna kako za TCP tako i za UDP aplikacije, a posebno negativan uticaj ima na multimedijalne aplikacije koje se odvijaju u realnom vremenu (VoIP, video konferencije, *streaming*).

Gubici paketa se izražavaju verovatnoćom pri čemu od same zavisi aplikacije i transportnog protokola (TCP ili UDP) koliko će moći da se tolerišu gubici paketa. Dodatni problem sa gubicima paketa jeste kako ih izmeriti s obzirom da se pojedini gubici dešavaju sporadično i nekada ih je teško detektovati i konačno izmeriti kako bi se analitički utvrdio efekat ovih gubitaka. Ova činjenica značajno komplikuje realizaciju nekih servisa, i implementaciju mrežnog nagledanja i upravljanja. Danas se u većini mreža, kao alat za merenje performansi koristi ICMP protokol koji daje povratnu informaciju o E2E kašnjenju i eventualnim gubicima paketa. Međutim, ovakav pristup nije adekvatan, pre svega jer nije u stanju da tačno i na pravi način reflektuje gubitke koji se dešavaju u sesijama koje su uspostavljene kroz mrežu.

Osim verovatnoće gubitaka paketa, koja je pre svega pokazatelj srednje vrednosti gubitaka, u cilju bolje karakterizacije šablona po kom se javljaju gubici što će detaljno biti analizirano u ovom radu. ITU-T je osim verovatnoće gubitaka date kao *IP packet loss ratio* (IPLR) u [2] dao i definicije *IP packet error ratio* (IPER), *IP packet reordered ratio* (IPRR), *IP packet duplicate ratio* (IPDR), *replicated IP packet ratio* (RIPR) koje takođe učestvuju u realizaciji performansi, ali u nešto manjem obimu.

2.2.4. Protok

Protok (*bandwidth*) se definiše kao broj bita u sekundi koji se može postići preko odgovarajućeg mrežnog medijuma ili generalno gledano, mrežnog uređaja. Određivanje tačnog protoka je u određenim slučajevima težak zadatak s obzirom da zavisi od činjenice da li na tom medijumu već postoji saobraćaj, pa samim tim izmereni protok je manji od realnog protoka koji može da se u datom slučaju postigne.

Poznavanje protoka na mrežnom segmentu je od izuzetnog značaja kako bi se utvrdilo da li postojeća infrastruktura može da prenese pomenutu količinu saobraćaja ili je potrebno da se izvrši odgovarajuća modifikacija politike rutiranja, tačnije prebacivanje dela saobraćaja na druge putanje sa adekvatnim kapacitetom. U cilju izbegavanja problema sa performansama i postizanja adekvatnih vrednosti prethodno navedenih parametara kojima se opisuje QoS, karakteristike saobraćaja se moraju uzeti u obzir. Jaki rafali i izražavanje multifraktalnih osobina saobraćaja mogu da dovedu da se u pojedinim intervalima, kratkotrajnim ili dugotrajnim, mogu desiti prekoračenja kapaciteta koja će rezultirati kašnjenjem ili gubicima paketa. Dakle, preduslov za realizaciju bilo kog servisa preko posmatranog mrežnog segmenta predstavlja dovoljan protok. Ukoliko dođe do kratkotrajnog zagušenja usled nedostatka kapaciteta, primenom QoS mehanizama i klasifikacijom klasa saobraćaja se donekle može postići predvidiv nivo performansi za klase koje imaju viši prioritet u odnosu na *best effort* saobraćaj.

Ono što se treba naglasiti jeste upravo međusobna zavisnost QoS parametara. Na primer, usled pojave zagušenja, odnosno smanjenja protoka, povećava se kašnjenje paketa, a u slučaju dugotrajnog zagušenja moguća je pojava i gubitaka paketa (naročito rafalni gubici čiji će negativni uticaj biti analiziran u ovom radu). Tokom ovog procesa, kašnjenje paketa će očekivano biti promenljivo, pa će na odredištu biti izražena i pojava džitera. Posledično, kada se posmatra jedna aplikacija, specificira se nekoliko QoS parametara koji su kritični za konkretnu aplikaciju pri čemu se mora imati u vidu i njihov međusoban uticaj.

2.3. Definisanje klasa servisa

Planiranje QoS politike i definisanje klasa servisa upravo polazi od identifikacije zahteva za performansama servisa i aplikacija u jednoj mreži. Naime, nakon što se utvrde različite potrebe za QoS, pristupa se definisanju klasa servisa kojima će se pridruživati saobraćaj pojedinačnih aplikacijama sa

specificiranim performansama. Definisanje klasa servisa je dato ITU-T preporukom [3]. U skladu sa tom preporukom, kao posledica potrebe zahteva za ispunjenjem različitih nivoa performansi, servis provajderi treba da definišu odgovarajući broj klasa servisa koje će garantovati QoS kako bi saobraćaj koji generišu aplikacije mogao da budu tretiran u skladu sa zahtevanim performansama.

Analiziranjem različitih servisa jasno je da se mora definisati nekoliko različitih nivoa performansi kako bi se na adekvatan način obezbedile QoS garancije za servise sa različitim servisnim zahtevima. Servisni zahtevi se generalno mogu posmatrati u kontekstu performansi od najstrožijih koji su svojstveni servisima koje je potrebno realizovati u realnom vremenu do servisa koje nemaju zahteve po pitanju kašnjenja i strogih vrednosti gubitaka paketa. Ono što se naročito mora naglasiti jeste da ne postoji minimalni broj klasa servisa koji se mora definisati, već se najpre identifikuje broj nivoa performansi koji se mora obezbediti u mreži za potrebe servisa sa različitim zahtevima. U kontekstu ITU-T preporuke [3] početni korak uvek predstavlja prepoznavanje servisnih potreba i mapiranje tih potreba u tehničke parametre čijom realizacijom se mogu postići nivoi performansi.

Tabela 2.1. Tehnička specifikacija klasa servisa [3]

Mera performanse	Određivanje praga mere performansi	QoS klase					
		Klasa 0	Klasa 1	Klasa 2	Klasa 3	Klasa 4	Klasa 5
IPTD	Gornja granica srednje vrednosti IPTD	100 ms	400 ms	100 ms	400 ms	1 s	U
IPDV	Gornja granica razlike $1-10^{-3}$ kvantila IPTD i $\text{IPTD}_{\min}$	50 ms	50 ms	U	U	U	U
IPLR	Gornja granica verovatnoće gubitaka paketa	1×10^{-3}	1×10^{-3}	1×10^{-3}	1×10^{-3}	1×10^{-3}	U
IPER	Gornja granica	1×10^{-4}					U

Formiranje adekvatnih klasa servisa igra značajnu ulogu prilikom uspostavljanja E2E performansi. U slučaju interprovajderskih scenarija koji obuhvataju garancije QoS i izvan granica jednog domena potrebno je i izvršiti redefinisane klasa servisa ukoliko se za to pokaže potreba. Često se umesto redefinisane klasa servisa vrši jednostavno dodavanje novih klasa koji odgovaraju novim servisnim zahtevima koji prethodno nisu bili na adekvatan način rešavani.

U okviru Tabele 2.2 su prikazane pet definisanih klasa servisa sa detaljno specificiranim gornjim vrednostima kašnjenja, džitera, gubitaka paketa i grešaka u prenosu. Ovih pet klasa servisa predstavljaju

samo preporuku datu u [3] i suštinski odgovaraju po performansama servisima datim u Tabeli 2.2. Primetno je da najstrožiju garanciju performansi zahtevaju prve tri klase servisa koje obuhvataju aplikacije čija se realizacija zahteva u realnom vremenu. Generalno gledano, za implementaciju ovih servisa je potrebno prethodno planiranje rutiranja kao i sprečavanje pojave zagušenja na putanjama gde bi se rutirao ovaj saobraćaj kojem se po pravilu dodaje visok prioritet. Bitno je napomenuti da se u skladu sa QoS politikom o garanciji nivoa performanse mora voditi računa i o politikama rutiranja, pa se kao deo realizacije ovih klasa servisa specificira i eventualno QoS rutiranje. QoS rutiranje podrazumeva da se kao ulazni parametri algoritma rutiranja koriste i QoS parametri na individualnim mrežnim segmentima. Dodatno, alternativa je i upotreba *traffic engineering* (TE) mehanizama što je čest pristup u servis provajderskim mrežama sa MPLS tehnologijom. Ostale klase servisa podrazumevaju uglavnom *data* servise za koje kašnjenje nije presudan parametar u određivanju E2E performansi. Takođe, za ove servise koji se u mrežama sa slabo regulisanom QoS politikom klasifikuju kao *best effort* saobraćaj ne koristi specijalna implementacija rutiranja niti TE. U Tabeli 2.2 je sa „U“ označen parametar koji nije kritičan, odnosno koji se ne mora garantovati u okviru date klase.

Tabela 2.2. Definisane klase servisa prema servisnim zahtevima [3]

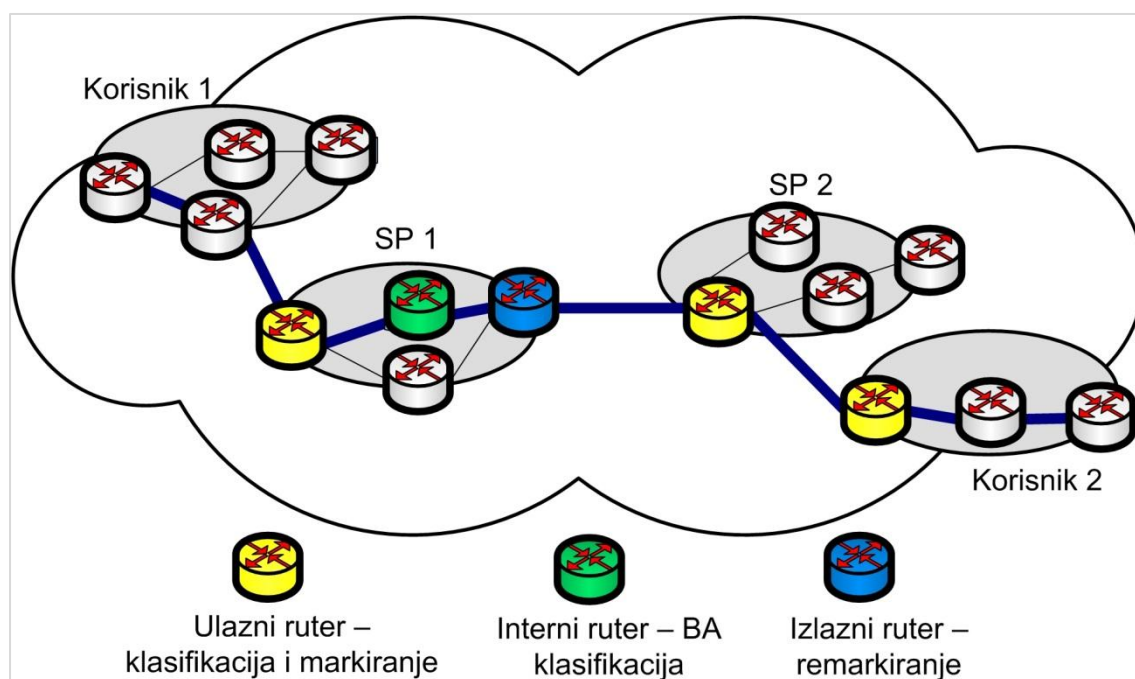
QoS klasa	Aplikacije (primeri)	Mrežne tehnike
0	Real-time, osetljiv džiter, visok nivo interakcije (VoIP, VTC)	Strogo QoS rutiranje i metrika rastojanja
1	Real-time, osetljiv džiter, Interaktivni servisi (VoIP, VTC)	QoS rutiranje i metrika rastojanja
2	Transakcioni servisi, visoka transakcija (signalizacija)	QoS rutiranje i metrika rastojanja
3	Transakcioni i interaktivni servisi,	QoS rutiranje i metrika rastojanja
4	Servisi sa malim gubicima paketa (kratke transakcije, data, video streaming)	Bilo koja ruta/putanja
5	Konvencionalne aplikacije i servisi u klasičnim IP mrežama	Bilo koja ruta/putanja

Klasifikacija saobraćaja u skladu sa klasama saobraćaja na paketskom nivou zavisi od primenjenih mehanizama klasifikacije paketa s obzirom da se u skladu sa DiffServ [4] arhitekturom, za svaki paket koristi odgovarajući metod klasifikacije pojedinačnog paketa. Pomenuta paketska klasifikacija je prisutna na drugom sloju u okviru *Class of Service* (CoS) polja i na trećem IP sloju u skladu sa specifikacijama definisanim u kao i za *Type of Service* (ToS) polja u IP zaglavlju [5], [6], [7].

2.4. DiffServ arhitektura i implementacija QoS politika

DiffServ model formalno obezbeđuje mehanizme za implementaciju QoS politike u skladu sa trenutnim tendencijama o garancijama performansi u posmatranoj mreži. DiffServ podrazumeva definiciju mnogih mehanizama i termina koji su vezani za QoS i s obzirom da se radi o standardu, vendori se trude da ispoštuju odgovarajuće termine iako se neretko dešava da se naziv nekog DiffServ termina promeni prema vendorskim marketinškim i tehničkim težnjama. DiffServ je vezan za polje u IP zaglavlju – *DiffServ Code Point* (DSCP) [5] i podrazumeva klasifikaciju paketa na osnovu ovog polja koje će odrediti kakav tretman će paket dobiti u ruteru, tačnije u mreži.

DiffServ pristup se koristi za implementaciju QoS politike koja može skalabilno da garantuje performanse u mrežama različite veličine. Na slici 2.2 je prikazan scenario kada je više različitih mreža (koja su pod različitom administrativnom upravom) međusobno povezano. Cilj je da se između korisnika preko provajderskih mreža obezbedi odgovarajući QoS tako što će se koristiti DiffServ model. Korišćenje DiffServ modela u ovom slučaju podrazumeva da će se kao deo definisanja QoS politike identifikovati nekoliko klasa servisa što se naziva diferencijacija klasa servisa. Ovom diferencijacijom se obezbeđuje u okviru jednog regiona (unutar mreže ili van mreže) nivo performansi za servise sa različitim zahtevima. U slučaju interdomenskih scenarija, svaki autonomni sistem predstavlja individualni DiffServ domen. Takođe, u okviru samih domena, u zavisnosti od pozicije rutera u mreži (ulazni, interni i izlazni ruter) implementiraju se odgovarajuće metode klasifikacije i prioritetizacije što je prikazano na slici 2.2.



Slika 2.2. DiffServ model u interdomenskom okruženju

Osnovni koncept podrazumeva da se u ulaznim tačkama, tačnije ulaznim ruterima, vrši klasifikacija paketa prema servisu kojem taj paket pripada. Klasifikacija paketa se realizuje markiranjem u skladu sa [5] tako što se vrednost DSCP polja u IP zaglavlju postavlja na predefinisanu vrednost koja odgovara klasi servisa. Dodatno, DiffServ razlikuje dve vrste tretiranja paketa u zavisnosti od vrednosti DSCP polja. U DiffServ terminologiji različiti tretman nad paketima se naziva *Per-Hop Behavior* (PHB), pa su prema DiffServ modelu definisana dva PHB – *assured forwarding* (AF) i *expedited forwarding* (EF). Oba PHB su namenjena za različite servise kod kojih su presudni različiti mrežni parametri.

EF PHB [8] podrazumeva pakete kod kojih je potrebno da se prenesu od izvora do odredišta sa minimalnim kašnjenjem, pri čemu saobraćaj koji ima ovakve zahteve mora da ima ograničen protok kako se ne bi ostale klase servisa dodatno degradirale. Specifikacijom EF DSCP vrednosti u IP zaglavlju paketa u mreži sa implementiranom QoS politikom se minimizuje kašnjenje, džiter i gubici paketa do odgovarajućeg nivoa protoka.

Sa druge strane, AF PHB [9] specificira postojanje čak četiri različite klase servisa u okviru kojih se može definisati 4 prioriteta čekanja u redovima i za svaki od tih redova se mogu specificirati verovatnoće odbacivanja paketa u slučaju pojave zagušenja usled nedostatka kapaciteta. PHB funkcija je vezana za mrežni čvor, koji generalno predstavlja ruter, pri čemu su PHB procesu pridružene dve funkcije i to postavljanje paketa u redove čekanja u skladu sa klasifikacijom (*queuing*) i pokretanje mehanizama za izbegavanje zagušenja (*congestion avoidance*). Za obe pomenute funkcije se koriste različiti algoritmi kako bi se postigle odgovarajuće performanse i izdiferencirale klase saobraćaja koje imaju različite zahteve u pogledu performansi.

2.5. Klasifikacija paketa i ograničenje protoka

Kao prvi korak implementacije QoS, realizuje se klasifikacija paketa na osnovu predefinisanih kriterijuma. Kao što je prethodno navedeno, a u skladu sa DiffServ modelom, klasifikacija se obavlja podešavanjem ili pregledom vrednosti DSCP polja. Klasifikacija se generalno vrši u ulaznim ruterima, i deli se na *Behavior Aggregate* (BA) i *MultiField* (MF) klasifikaciju. Osim klasifikacije, u ulaznim ruterima radi se i *policing* saobraćaja, tačnije ograničenje protoka u skladu sa predefinisanom vrednošću koja je dogovorena za dati saobraćaj odnosno korisnika koji generiše taj saobraćaj.

BA klasifikacija obuhvata isključivo pregledanje vrednosti DSCP polja i podrazumeva da je process klasifikacije i *policing*-a prethodno obavljen. Samim tim, BA klasifikacija se obavlja u okviru rutera unutar jedne mreže jer praktično verodostojno uzima DSCP vrednosti dolaznih paketa s obzirom da su prethodno podešene u skladu sa QoS politikom i klasama servisa. Nasuprot BA klasifikaciji, MF klasifikacija se uglavnom realizuje na ulaznim ruterima u mreži i vrši dodelu vrednosti DSCP polja paketa na osnovu drugih parametara kao što su izvorišna ili odredišna IP adresa, broj porta, ulazni port rutera, itd. Jasno je da

MF klasifikacija polazi od toga da se paketima koji su eventualno markirani u skladu sa korisničkom QoS politikom ne može verovati, pa se vrši remarkiranje na ulaznom ruteru. Na osnovu MF klasifikacije i posledičnog markiranja, BA klasifikacija u ostalim ruterima u mreži dodeljuje paket odgovarajućoj klasi i prosleđuje posmatrani paket u skladu sa prioritetom koji ta klasa poseduje. MF klasifikacija takođe služi da se izvrši, ukoliko je potrebno, remarkiranje kako bi unutar mreže BA klasifikacija mogla efikasnije da se obavlja.

Ograničenje saobraćaja predstavlja drugu funkciju ulaznog rutera koja se realizuje nakon klasifikacije saobraćaja. Ograničenje protoka definiše situaciju u kojoj saobraćaj posmatranog korisnika prekoračuje dogovorenu vrednost. Dodatno, u okviru mehanizama ograničenja se definiše šta se dešava sa saobraćajem ako se to prekoračenje dogodi.

Pod ograničenjem protoka (*policing-om*) se definišu dodatne radnje koje imaju cilj da spreče pojavu zagušenja usled prekomerno generisanog saobraćaja koji prethodno nije definisan ugovorom između provajdera i korisnika. Da bi se primenilo ograničenje, potrebno je da se stalno prati nivo ulaznog saobraćaja i da se u svakom trenutku utvrdi da li je generisani saobraćaj u skladu sa ugovorom. Ukoliko jeste, normalno se prosleđuje u ostatak servis provajderske mreže. U slučaju prekoračenja, u zavisnosti od QoS politike, može se desiti da paketi koji prekoračuju protok bivaju odbačeni. Takođe, može se primeniti i oblikovanje saobraćaja (*shaping*) čime se prekomerni paketi baferuju privremeno dok saobraćaj ne postane ponovo konforman kako bi se baferovani paketi poslali. Posledično, kašnjenje baferovanih paketa je veće i za njih postoji veća verovatnoća da će biti odbačeni ukoliko saobraćaj ne postane konforman. U skladu sa tim, baferovanje je proces koji nije primenljiv na *real time* servise gde je kašnjenje kritičan parametar. U praksi se *shaping* koristi kod saobraćaja koji u određenim intervalima može imati *peak-ove* u smislu velikog broja paketa u kratkom intervalu. Osim odbacivanja i *shaping-a* saobraćaja, moguće je izvršiti i remarkiranje čime se *de-facto* menja verovatnoća odbacivanja paketa u slučaju zagušenja. Dakle, ovim se postiže da ukoliko paketi koji nisu konformni mogu da se obrade jer ima raspoloživog kapaciteta, moguće ih je markirati tako da ukoliko eventualno dođe do zagušenja u nekom drugom delu mreže, ti paketi na osnovu tog markiranja postaju prvi kandidati za odbacivanje.

2.6. Algoritmi za kontrolu zagušenja

Osim što ruteri vrše rutiranje i prosleđivanje paketa, oni su kritična tačka za pojavu zagušenja, pa samim tim se u okviru arhitekture rutera implementiraju i algoritmi kojima se ublažava ili, ukoliko je moguće, eliminiše zagušenje. S obzirom da svaki paket koji se prosleđuje na odlazni port se smešta u odgovarajuće redove, od izuzetnog značaja je da algoritmi koji upravljaju opsluživanjem redova budi adekvatno prilagođeni servisnim potrebama.

Algoritmi za kontrolu zagušenja upravljaju dužinom paketskih redova i ako je potrebno odbacuju pakete iz redova, dok opsluživanje redova podrazumeva određivanje koji će se naredni paket poslati kroz odlazni port. Oba tipa algoritama se pre svega koriste kako bi se upravljalo slobodnim kapacitetom između različitih klasa saobraćaja, odnosno paketima koji pripadaju različitim klasama.

Najjednostavniji algoritam za kontrolu zagušenja koji upravlja procesom smeštanja paketa u redove čekanja jeste *first-in first-out* (FIFO) mehanizam. Ovaj pristup podrazumeva scenario u kome ne postoji klasifikacija saobraćaja niti prioritizacija. U tom slučaju, svaki paket se posmatra na isti način i opslužuju se po redosledu dolaska u red. Ovakvom redosledu opsluživanja je i pridružena primitivna politika odbacivanja paketa u slučaju zagušenja koja podrazumeva da se novopristigli paketi neselektivno odbacuju čim se red čekanja popuni. Ovakvo odbacivanje se naziva *tail drop* i često u slučaju zagušenja dovodi do pojave rafalnih gubitaka koji su pogubni po kvalitet servisa koji su osetljivi na gubitke.

Sa druge strane, *priority queing* (PQ) je mehanizam po kome se paketi klasifikuju u nekoliko redova kojima su dodeljeni prioriteti, na primer *high*, *medium*, *low* prioriteti. U slučaju PQ kontrole zagušenja, svakoj klasi paketa se dodeljuje po jedan red koji kada se napuni podrazumeva se da se vrši odbacivanje. Karakteristično za PQ jeste redosled opsluživanja redova koji se opslužuju u skladu sa prioritetom pri čemu se šalju paketi iz reda višeg prioriteta i tek nakon što se opsluže svi paketi u tom redu, prelazi se na naredni red nižeg prioriteta. Ovaj pristup se primenjuje sve do reda najnižeg prioriteta koji dobija pravo da pošalje pakete tek kada svi ostali redovi višeg prioriteta nemaju više paketa za slanje. PQ se uglavnom upotrebljuje kada postoji klasa servisa koja mora da ima malo kašnjenje i gubitke paketa. Toj klasi se dodeljuje najviši prioritet. Međutim, mora se voditi računa kod primene ovog mehanizma kontrole zagušenja jer ukoliko postoji veliki saobraćaj najvišeg prioriteta, može se desiti da se zbog redosleda opsluživanja ostali redovi neće opslužiti adekvatno što će rezultirati odbacivanjem paketa iz više klasa nižeg prioriteta. Ova pojava se još naziva „izgladnjivanje redova“ (*queue starvation*).

Kao korak prema fer pristupu u odnosu na PQ algoritam opsluživanja, definiše se i *fair queing* (FQ) [10]. Sličan mehanizam FQ je *Custom Queing* (CQ) koji deli klase servisa po redovima i opslužuje ih u *round robin* režimu prema kome se iz svakog reda šalju paketi u specificiranom intervalu. FQ je polazna tačka za naprednije algoritme upravljanja zagušenjem koji se baziraju na činjenici da svaka klasa treba da dobije fer pristup mrežnom resursu. Raspoloživ kapacitet se deli na klase saobraćaja, tako da svaka klasa dobije određeni deo kapaciteta u intervalu vremena. Kod FQ ne postoji klasa koja će monopolizovati kapacitet, već se se paketi opsluživati u skladu sa delom kapaciteta koji je posmatrani red dobio. Ovakva vrsta kontrole zagušenja poseduje i naprednije varijante opsluživanja u okviru kojih se dalje pojedini redovi mogu prioritizovati. Takođe, FQ je pogodan kada postoji nekoliko klasa servisa sa sličnim servisnim zahtevima gde nijednoj klasi nije potrebno dodeliti prioritet kako bi se obezbedio E2E QoS.

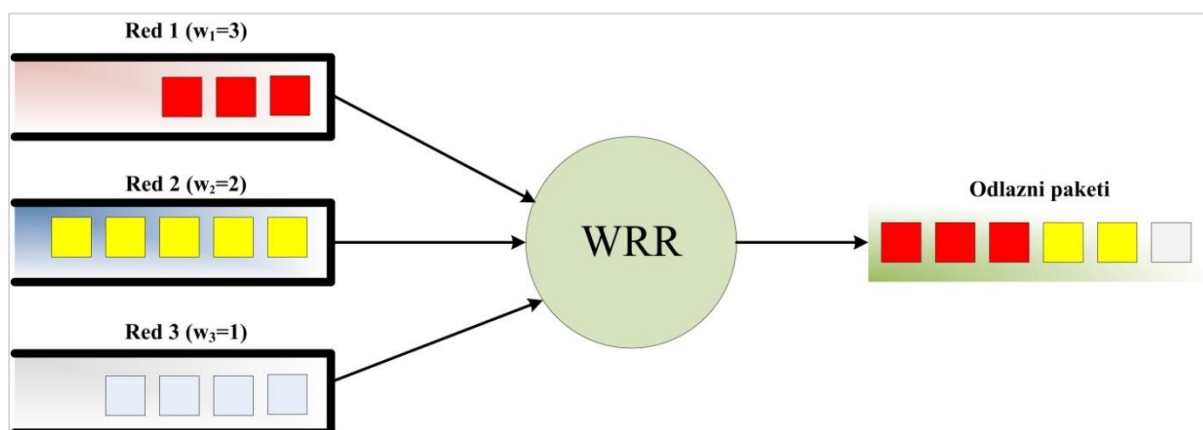
Weighted Fair Queuing (WFQ) predstavlja mehanizam u okviru koga se redovima čekanja dodeljuje kapacitet u skladu sa pridruženim težinskim faktorom. WFQ podrazumeva da se paketi tretiraju podjednako imajući u vidu dodeljenu težinu individualnom redu. Kao specijalni slučajevi WFQ definišu se *Class-Based Weighted Fair Queuing (CBWFQ)* i *Low Latency Queuing (LLQ)*.

CBWFQ se u današnjim mrežama najčešće primenjuje i po principu rada najsličniji je CQ algoritmu. U poređenju sa CQ je dosta efikasniji s obzirom da se kod CBWFQ konfigurišu klase servisa i saobraćaj se dodeljuje tim klasama. Nakon toga, klasama se dodeljuje deo kapaciteta kojim se svaki red opslužuje. Na ovaj način, svaki red (klasa koja se nalazi u tom redu) dobija minimalni zagarantovani kapacitet koji će dobiti u slučaju zagušenja.

LLQ predstavlja mehanizam koji se dobija kada se CBWFQ klasama dodeli još jedna klasa sa najvišim prioritetom koja se opslužuje po principu PQ. LLQ je uveden zbog potrebe VoIP saobraćaja za minimumom kašnjenja i džitera, dok ostale klase servisa imaju minimalni zagarantovani kapacitet. Na ovaj način se dobija na fleksibilnosti definisanja dela raspoloživog kapaciteta za ostale klase saobraćaja, dok prioritet ima PQ red koji će dobiti minimalno kašnjenje i džiter u uslovima zagušenja.

2.7. Algoritmi za kontrolu opsluživanja redova

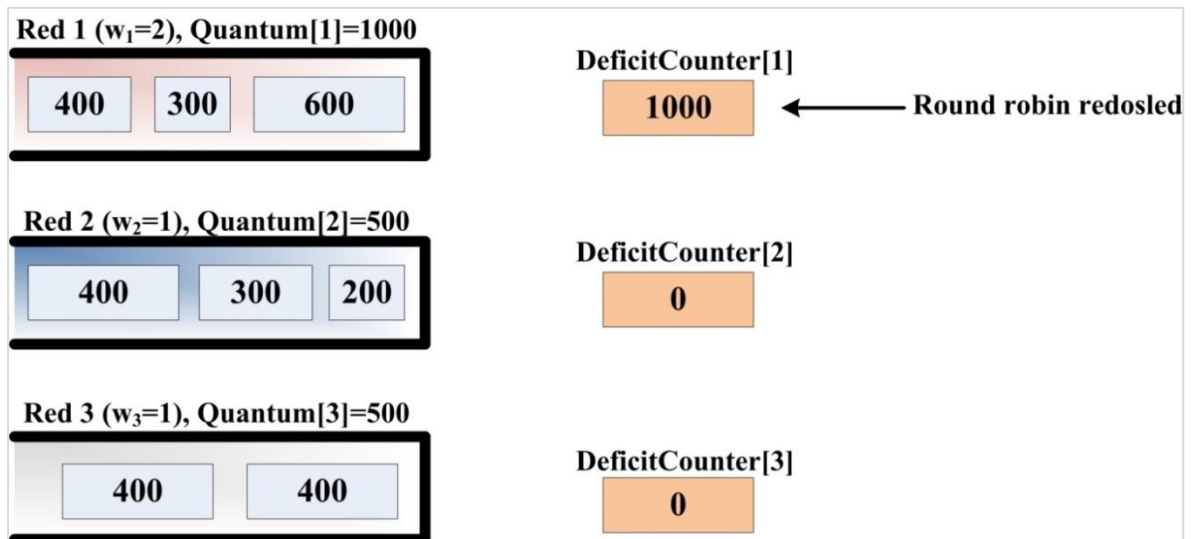
Algoritmi za kontrolu opsluživanja redova (*scheduling*) podrazumevaju određivanje redosleda opsluživanja paketa u okviru redova. Kao i kod algoritama za kontrolu zagušenja, postoji nekoliko algoritama koji određuju redosled opsluživanja paketa. U nastavku će biti pomenuto nekoliko bitnijih *scheduling* algoritama.



Slika 2.3. Princip rada WRR scheduling algoritma

Weighted Round Robin (WRR) predstavlja jednostavan *scheduling* algoritam u okviru koga se dodeljuju statički težinski faktori redovima. Algoritam postupa ciklično od reda do reda i opslužuje onoliko paketa po svakom redu u zavisnosti od dodeljenog dela kapaciteta odlazne veze. Na taj način se postiže da redovi koji imaju niži prioritet nikada ne mogu ostati bez svog kapaciteta jer se ovakvim opsluživanjem redova

garantuje minimalni deo kapaciteta koji poseduje svaki red čekanja. Ozbiljno ograničenje WRR algoritma se ogleda u tome što deo kapaciteta odlazne veze koji se dodeljuje pojedinačnim redovima zavisi od veličine paketa koji se nalaze u tim redovima, pa je za efikasnu upotrebu WRR potrebno unapred poznavati statističke osobine paketa koji će se nalaziti u pojedinačnim klasama saobraćaja, odnosno redovima čekanja. Na slici 2.3 je prikazan primer opsluživanja paketa kada postoji tri reda sa različitim težinskim faktorima, w_1 , w_2 i w_3 .

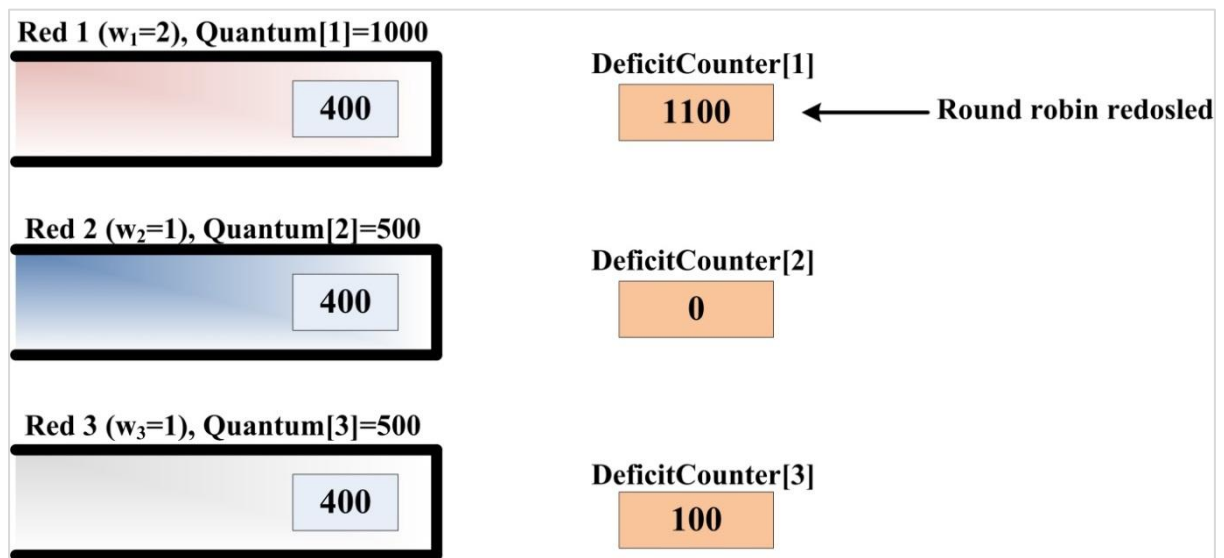


Slika 2.4. Princip rada DWRR scheduling algoritma (opsluživanje prvog reda u prvom ciklusu)

Deficit Weighted Round Robin (DWRR) algoritam rešava prethodno navedeno ograničenje WRR tako što podržava fer distribuciju kapaciteta kod redova sa različitim težinskim faktorima kada su paketi promenljive dužine. U odnosu na WRR, sa stanovišta kompleksnosti, lakše se realizuje hardverski nego njegov prethodnik. DWRR opslužuje svaki red koji sadrži pakete koji čekaju da budu opsluženi i utvrđuje broj bajtova paketa koji je prvi u redu. Ukoliko je veličina paketa veća u odnosu na preostali kapacitet koji posmatrani red može da iskoristi, DWRR prelazi na sledeći red pri čemu se ostatak kapaciteta tog reda koristi za opsluživanje u narednom ciklusu. Na taj način se eliminiše problem sa pojavom paketa različite dužine, pa se individualnim redovima se dodeljuje tačan procenat ukupnog kapaciteta koji mogu da koriste.

Primer sa tri reda čekanja koja se opslužuju u skladu sa WRR je prikazan na Slika 2.4 i Slika 2.5. Naime, na slici 2.4 prikazana su tri reda pri čemu prvi red ima dodeljen 50% kapaciteta odlaznog linka, a druga dva po 25%. Takođe, *quantum* parametar određuje koliko bajtova u jednom ciklusu DWRR može da opsluži iz jednog reda i taj parametar je srazmeran težinskim faktorima. Za potrebe ovog primera, *quantum* parametar je inicijalizovan na 1000B za red 1, odnosno po 500B za redove 2 i 3. Parametar *deficit counter* preostali broj bajtova koji se može opslužiti u okviru posmatranog reda. Početna vrednost je jednaka *quantum* parametru plus preostala vrednost ovog parametra iz prethodnog ciklusa (s obzirom da se radi o

incijalizaciji, prethodna vrednost je 0). Prvi paket od 600B se opslužuje i *deficit counter* je nakon toga 400B. Drugi paket od 300B se takođe opslužuje i sada *deficit counter* iznosi 100B. S obzirom da je ostalo 100B kapaciteta da se pridruži redu 1, a naredni paket od 400B se ne može smestiti u preostali kapacitet, DWRR prelazi na opsluživanje reda 2, ali *deficit counter* iznosi 100B za red 1 kome se u narednom ciklusu (nakon primene istog principa na red 2 i 3) dodaje *quantum* vrednost, pa sada u tom ciklusu ukupan kapacitet iznosi 1100B kao što je prikazano na slici 2.5. Treba se eventualno naglasiti i varijacija rada DWRR koja podrazumeva da se naredni paket za koji nema dovoljnog kapaciteta opsluži, ali da se prilikom sledećeg ciklusa ta vrednost oduzme od raspoloživog kapaciteta reda. Analitički se to može sagledati ako se smatra da *deficit counter* parametar može biti negativan.



Slika 2.5. Princip rada DWRR scheduling algoritma (opsluživanje prvog reda u drugom ciklusu)

2.8. Algoritmi za izbegavanje zagušenja – *Random Early Detection*

Sastavni deo QoS mehanizama predstavljaju i algoritmi za izbegavanje zagušenja čija je funkcija da u slučaju zagušenja preventivno vrše odbacivanja paketa kako bi pomenuto zagušenje imalo manji uticaj na pojedinačne servise. Naime, u slučaju da se desi zagušenje i da su redovi čekanja puni, svaki novi pristigli paketi će biti odbačeni usled nedostatka mesta u okviru redova. Ovakva odbacivanja se dešavaju ukoliko prethodno nisu korišćeni algoritmi za izbegavanje zagušenja. Neselektivno *tail drop* odbacivanje u okviru rutera dovodi potencijalno do mnogo problema koji se dalje manifestuju u dodatnoj degradaciji performansi, pre svega aplikacija koje se oslanjaju na TCP protokol.

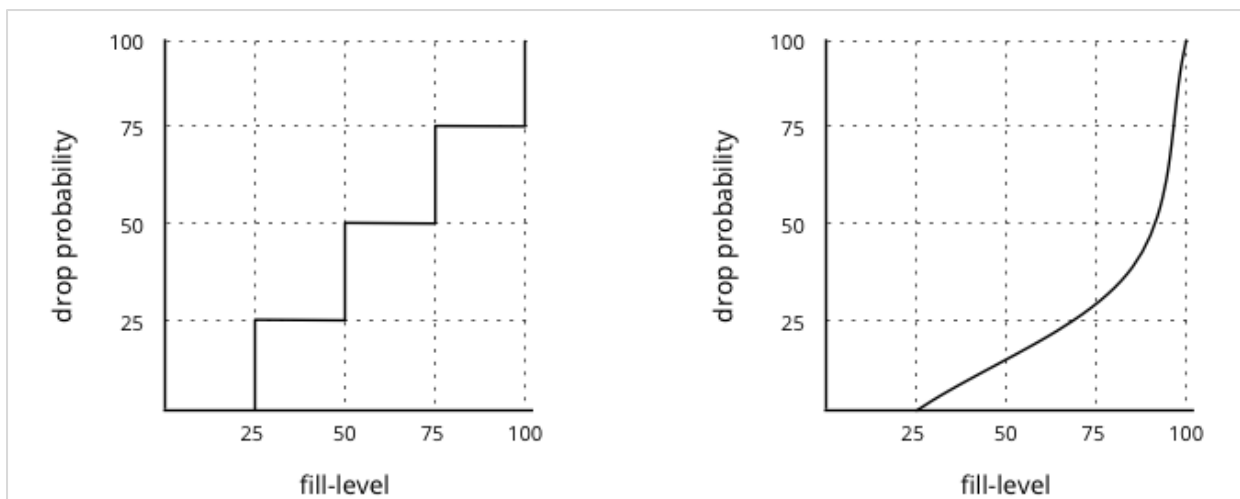
Kada se usled *tail drop* odbacivanja paketa odbace paketi koji pripadaju većem broju različitih TCP sesija, TCP će kao odgovor na ovaj gubitak smanjiti svoj *congestion window* što će u slučaju svih TCP sesija dovesti do smanjenja performansi. Nakon toga, doći će to pojave koja se naziva globalna sinhronizacija koja u suštini dovodi ponovo do naglog porasta saobraćaja i broja paketa u redovima čekanja s obzirom da će se sve te TCP sesije sinhronizovati i istovremeno povećavati širinu *congestion window* parametra što će

ponovo dovesti do novog zagušenja i odbacivanja. Treba se naglasiti da UDP aplikacije usled nedostatka retransmisionog mehanizma i kontrole zagušenja nemaju problema sa ovom pojavom, pa gubitak UDP paketa ne nosi opasnost po mrežu u širem obimu.

Kako bi se sprečio problem sa dodatnim zagušenjima usled TCP saobraćaja, koristi se sada već standardan mehanizam za izbegavanje zagušenja – *Random Early Detection* (RED), odnosno njegova naprednija i fleksibilnija varijanta *Weighted Random Early Detection* (WRED).

RED mehanizam rešava prethodno navedeni problem zagušenja u redovima tako što se vrši preventivna selekcija odbacivanja paketa određenih TCP sesija tako da te sesije usled inherentne kontrole zagušenja sada šalju manje paketa u mrežu. Dodatno, kao posledica ovog selektivnog odbacivanja, sada ne postoji opasnost od TCP globalne sinhronizacije s obzirom da RED vrši ranu detekciju zagušenja pojedinačnih redova čekanja. Primetno je da se paket selektivno odbacuju i više nema pojavljivanja *tail drop* odbacivanja paketa, a samim tim ni globalne TCP sinhronizacije..

RED mehanizam sadrži dve funkcionalne celine i to: funkciju da detektuje kada se zagušenje dešava i funkciju odluke po kojoj se određuje koliko paketa treba odbaciti kako nadolazeće zagušenje ne bi napravilo potpunu degradaciju TCP aplikacija. Detekcijom zagušenja se određuje pod kojim uslovima je optimalno da se vrši odbacivanje paketa.



Slika 2.6. RED profili odbacivanja paketa: (a) stepenasta funkcija; (b) interpolirana funkcija.

Detekcija zagušenja se vrši na osnovu posmatranja popunjenosti jednog reda. Naime, definiše se nekoliko pragova popunjenosti reda na osnovu koga se statistički određuje da li je paket kandidat za odbacivanje ili nije. Takođe, kao deo QoS politike se svaki paket je markiran tako da se na osnovu toga može izvesti verovatnoća kojom će biti odbačen u zavisnosti od popunjenosti reda – *Packet Loss Probability* (PLP). Vendori često implementiraju u svojim uređajima u cilju postizanja fleksibilnosti mogućnost zadavanja različitih zavisnosti verovatnoće odbacivanja od popunjenosti reda. Na slici 2.6 su prikazana dva profila

koja omogućavaju da se formiraju dva RED profila odbacivanja u zavisnosti od klase saobraćaja i eventualne opasnosti od zagušenja. Na slici 2.6(a) je prikazana stepenasta funkcija koja definiše u kojim slučajevima će paket biti odbačen. U nekim vendorskim implementacijama paket sa podešenim PLP se izračunava slučajni broj koji ukoliko se nalazi iznad funkcije odbacivanja podrazumeva da se paket preventivno odbacuje, dok ukoliko je ispod funkcije paket ostaje u redu čekanja. Na slici 2.6(b) prikazana je interpolirana funkcija koja u poređenju sa stepenastom funkcijom na slici 2.6(a) predstavlja agresivniji profil odbacivanja s obzirom da je površina ispod krive mnogo manja.

WRED predstavlja specijalni slučaj RED mehanizma sa tom razlikom da odluku o odbacivanju paketa donosi na osnovu vrednosti DSCP polja u IP zaglavlju paketa. WRED u današnjim mrežama predstavlja standardni mehanizam s obzirom da se prethodno pomenuti profile odbacivanja upravo odnose na markiranje paketa i dodeljivanje PLP vrednosti svakom paketu u skladu sa QoS politikom. Na ovaj način omogućena je potpuna kontrola procesa izbegavanja zagušenja pri čemu se u okviru QoS politike jedne mreže može detaljno isplanirati selektivno odbacivanje paketa kao i mapiranje između primene određenih profila odbacivanja i klasa saobraćaja.

2.9. Izazovi u implementaciji QoS politike i zaključna razmatranja

U današnjim servis provajderskim, pa čak i *enterprise* mrežama implementacija QoS arhitekture predstavlja jedan od preduslova za diferencijaciju servisa i ostvarivanje poslovnih ciljeva kroz realizaciju „*state of the art*“ servisa. Međutim, prethodno predstavljeni QoS mehanizmi su generalno kompleksni za implementaciju i zahtevaju detaljno poznavanje statistike saobraćaja u mreži. Iako postoje predlozi za skalabilnu implementaciju QoS arhitekture, problem kod skalabilne QoS implementacije podrazumeva da se nedosledno primenjuje u pojedinim delovima iste mreže. Dodatni problem jesu interprovajderski scenariji i različite QoS politike u njima. Naime, svaki provajder kreira u skladu sa svojim ekonomskim i poslovnim ciljevima QoS politika koja se pre svega zasniva na ponuđenim servisima. Kada se posmatra interprovajderski scenario, problem uglavnom nastaje usled različite implementacije QoS politike i nejednakih kriterijuma na osnovu kojih se ta politika primenjuje.

Primarni cilj QoS politike jeste upravo diferencijacija servisa kako bi se garantovale performanse za pojedine klase servisa. Tehnički gledano, svaka klasa servisa treba da ima definisane granične vrednosti QoS parametara kao što su kašnjenje, džiter i gubici paketa. U slučaju zagušenja na pojedinim linkovima ili mrežnim uređajima, primena QoS mehanizama realizuje da se ciljne vrednosti za specificirane klase servisa održe. Međutim, operativno stanje mreže može da dovede do pojave degradacije QoS parametara, a koja nije uzrokovana zagušenjem. U tom slučaju, postojeći QoS mehanizmi nisu sposobni da adekvatno reaguju na datu degradaciju, pa samim tim čak i sa dosledno implementiranom QoS arhitekturom nije moguće popraviti povećano kašnjenje ili gubitke paketa koji se javljaju, na primer, usled nestabilnosti fizičke veze ili procesorske opterećenosti rutera.

Kao jedna od trivijalnih metoda koja ne bi zahtevala upotrebu QoS jeste da se kapacitet mreže višestruko uveća iznad realnih potreba zajedno sa kapacitetom interkonekcija između provajdera [11]. Ovakav pristup ne bi doveo do zagušenja, pa samim tim ne bi bilo potrebe za prioritetizacijom servisa jer postoji dovoljno resursa da se paketi obrade bez degradacije performansi. Međutim, ovakav pristup nije ekonomski isplativ, pa samim je opravdano da neka forma QoS implementacije treba da postoji.

Definisanje QoS politike često treba da bude koordinisano sa poslovnim ciljevima. U implementacionom kontekstu, QoS često biva implementiran nedosledno što naknadno može da stvori operative probleme u mreži koji dodatno mogu da uzrokuju probleme sa performansama i otežan rad pojedinih servisa. Ova pojava je uglavnom prisutna u mrežama sa opremom od više vendora koje nude mogućnost za različite QoS mehanizme koji nemaju iste kriterijume u pogledu performansi.

Kao još jedan aktuelni problem kod implementacije QoS mehanizama jeste problem Internet neutralnosti koji predstavlja potencijalnu opasnost od korišćenja prioritetizacije na diskriminatoran način kako bi se favorizovao skup provajdera sadržaja u pogledu performansi [12], [13], [14]. Današnji Internet se uveliko zasniva na slobodnom pristupanju sadržajima gde je osnovna pretpostavka da će korisnik dobiti podjednake performanse bez ikakve diskriminacije provajdera sadržaja (*content provider*). Iako servisna diferencijacija može korisnicima ponuditi prednosti u pogledu ostvarivanja i donekle garantovanja performansi, ona mora biti definisana na način kojim se poštuje konkurencija na Internetu i čime se izbegava štetno diskriminisanje pojedinih kontenta provajdera od strane većih ili manjih servis provajdera. Regulative koje se propisuju i bave ovim problemom sigurno će postaviti QoS kao predmet detaljne analize u cilju utvrđivanja potencijalnih mogućnosti kako bi se isključila mogućnost upotreba ovih mehanizama protiv Internet neutralnosti.

Ono što bi mogla biti alternativa postojećim QoS mehanizmima opisanim u ovom poglavlju jeste upravo QoS rutiranje koje će rešiti neke od važnih problema sa kojima se sreću aktuelno korišćeni mehanizmi za garantovanje performansi. Modifikacije trenutnih politika rutiranja predstavljaju perspektivan pristup u rešavanju problema sa degradacijom performansi u okviru mreže (zagušenje ili greške u funkcionisanju linkova i/ili mrežnih čvorova), dok se implementacijom naprednih mehanizama (kombinovanje servisnih ugovaranja sa ispunjavanjem poslovnih i ekonomskih ciljeva) koje često prevazilaze tehnički nivo može implemenirati garantovanje performansi i izvan granica jedne mreže.

3

UTICAJ STRATEGIJA PAKETSKE DISPERZIJE NA QOS UNUTAR DOMENA

Potreba za diferencijacijom klasa servisa u unutar domena uzrokovala je nastanak čitavog spektra QoS mehanizama koji omogućavaju postizanje specificiranog nivoa performansi za pojedine klase servisa u uslovima zagušenja. Međutim, definisanje QoS politike ne sprečava da pojedine degradacije performansi neminovno utiču na sve klase saobraćaja, pa i na prioritetizovane klase koje po pravilu dobijaju preferencijalni tretman u mrežnim uređajima. Dodatni problem nastaje kada primenom QoS politike nastaje heterogenost u performansama koja se ogleda u nepoznatim graničnim vrednostima mera performansi relevantnih za servise koji se pružaju. U tom slučaju se ne može na nivou celog domena utvrditi nivo performansi koje pojedine klase servisa mogu da ostvare usled nedosledno primenjene QoS arhitekture.

Uzimajući u obzir navedenu kompleksnost u implementaciji i planiranju, diferencijacija klasa servisa je tesno povezana sa QoS politikom. Osim postojećih QoS mehanizama, kao deo planiranja QoS politike treba se utvrditi mogućnost za upotrebom QoS protokola rutiranja koji mogu da određuju putanje u zavisnosti od specificiranih performansi. U servis provajderskim mrežama, upotreba ovakvih protokola se

primenjuje kao deo TE, dok upotreba QoS *Interior Gateway Protocol* (IGP) protokola se često izbegava zbog loše skalabilnosti i dužeg vremena konvergencije.

Imajući u vidu različite nivoe performansi koje treba obezbediti servisima potrebno je da se sagledaju mogući uzroci pojave degradacije performansi koji nisu samo izazvani pojavom zagušenja. Te degradacije se manifestuju kao nekontrolisano menjanje vrednosti mera performansi, pre svega kašnjenja, džitera i verovatnoće gubitaka paketa. U zavisnosti od posmatrane aplikacije, svaka od ovih mera performansi ima drugačiji efekat. Kako bi se detaljno ispitao uticaj ovih parametara potrebno je poznavati njihova statistička svojstva i načine modelovanja kako bi se mogla u potpunosti ispitati dubina uticaja individualnih QoS parametara na pojedine servise.

Kako bi se vrednosti QoS parametara držale pod kontrolom i performanse bile na očekivanom nivou, potrebno je upotrebiti osim standarnih mrežnih mehanizama i alternativne pristupe u rešavanju problema sa pojavom degradacije performansi. Jedna od tih metoda predstavlja promenu politike rutiranja čijom se implementacijom postiže da u slučaju pojave problema sa performansama, može se izvršiti rutiranje QoS kritičnog saobraćaja i servisa na druge raspoložive putanje. Osim toga, u današnjim mrežama je prisutan visok stepen redundanse, pa se odabirom nekoliko različitih putanja može ostvariti ciljani nivo performansi.

U okviru ovog rada će se razmatrati mogući pristupi promena politike rutiranja koje za primarni cilj imaju ostvarivanje zahtevanog nivoa performansi, dok će se kao relevantni pokazatelji delotvornosti posmatrati QoS parametri. Centralni deo analize politika rutiranja jeste paketska disperzija u okviru koje se predlaže analitički model čijom primenom je moguće odrediti način na koji različite strategije paketske disperzije mogu da utiču na poboljšanje kvaliteta *voice* aplikacija, kao aplikacija koje su osetljive na degradaciju performansi, naročito na gubitke paketa.

3.1. Politike rutiranja unutar domena

Današnje politike rutiranja, bilo na intradomenskom ili interdomenskom nivou, se dominantno zasnivaju na *single path* rutiranju, što znači da se do specificiranog odredišta koristi jedna putanja odabrana od strane protokola rutiranja po odgovarajućem kriterijumu. Rutiranje po jednoj putanji, iako je prihvaćeno kao konvencionalno rutiranje u današnjem Internetu, ipak nije dovoljno fleksibilno da obezbedi sve veće zahteve po pitanju raspoloživosti i QoS, pri čemu se i pitanje iskorišćenosti mrežne infrastructure u ovom slučaju može dovesti u pitanje [15]. Visok nivo redundanse u današnjim mrežama je posledica upravo većeg zahteva za raspoloživošću servisa koja se oslanja na fleksibilnu arhitekturu rutiranja koja mora biti sposobna da uzme u obzir sve raspoložive putanje i na odgovarajući način usmeri pakete preko istih. Paradigma *multipath* rutiranja rešava prethodno pomenute zahteve tako što koristi protokole rutiranja koji

su sposobni da odrede raspoložive putanje koje ispunjavaju određeni kriterijum i da se različite putanje istovremeno koriste za rutiranje saobraćaja prema jednom odredištu.

Iako je doprinos *multipath* rutiranja višestruk, razlog što ovakav pristup rutiranju u većoj meri zastupljen je tesno vezan sa kompleksnošću, vremenom konvergencije i skalabilnošću. Naime, protokoli rutiranja koriste odgovarajući algoritam na osnovu koga se proračunava jedna ili više raspoloživih putanja do odredišta. U slučaju *multipath* algoritama rutiranja, algoritam je dosta kompleksniji jer mora da sagleda ne jednu „najbolju“ putanju, već više kandidata putanja do jednog odredišta koji su prihvatljivi sa aspekta implementiranog protokola rutiranja. Sa tim u vezi, procesorski resursi koji se koriste za ova izračunavanja su sada veći u odnosu na *single path* algoritme, pa i ruteri u okviru kojih se implementiraju protokoli rutiranja moraju da poseduju značajnije resurse kako bi mogli da odgovore na ovaj zahtev. Dodatno, s obzirom da se između rutera mora realizovati signalizaciona razmena koju diktira protokol rutiranja, mora se imati u vidu da će se generisati i veći kontrolni saobraćaj, koji u određenom slučaju može da smanji kapacitet koji je namenjen za produkcionni saobraćaj. Sve ove prethodno navedene činjenice počinju itekako da utiču sa povećanjem same mreže koje se ogleda u povećanju broja rutera, saobraćaja, linkova i broja raspoloživih putanja između dve tačke.

Ako se posmatra intradomensko rutiranje, administrator domena ima potpunu kontrolu svoje mreže, pa samim tim je u mogućnosti da kroji politiku rutiranja u skladu sa servisnim zahtevima i poslovnim ciljevima entiteta koji je vlasnik posmatrane infrastrukture, odnosno domena. U današnjim mrežama kao IGP protokoli uglavnom koriste *Open Shortest Path First (OSPF)* [16] i *Intermediate System to Intermediate System (IS-IS)* [17], [18] koji su *link state* tipa. *Link state* tipovi protokola rutiranja podrazumevaju da je svaki ruter u okviru jednog domena svestan kompletne topologije, pri čemu se putanje do odredišta određuju na osnovu baze podataka koja je identična kod svih rutera domena i koja je kreirana razmenom kontrolnih informacija između rutera. Bez obzira na topološko poznavanje, algoritmi rutiranja i dalje ne uzimaju u obzir postojanje više putanji prema odredištima, pa je i nivo implementirane redundanse posledično neiskorišćen što je i pokazano na primeru *tier 1* servis provajderske mreže u [19].

Implementacijom *multipath* protokola rutiranja se javlja dodatni aspekt koji se mora sagledati sa više stanovišta. Naime, činjenica da postoji više od jedne putanje do odredišta implicira da se saobraćaj do tog odredišta mora na odgovarajući način podeliti po pomenutim putanjama. Ostim toga, postojanjem više putanja je proces konvergencije nakon otkaza rutera ili linka do odredišta brži s obzirom da de facto nema prekida servisa čiji saobraćaj nastavlja da se prenosi nekom od preostalih putanja. *Multipath* pristupom je sada moguće da se dalje kreira QoS politika u zavisnosti od putanje, a da se balansiranje saobraćaja upravo vrši na osnovu ove aplikacionih zahteva u pogledu performansi. Upravo ovom premisom se vodi paketska disperzija kada se definiše nivo QoS koju jedna aplikacija treba da ima kada postoji više putanja u mreži.

Direktne prednosti koje se ostvaruju primenom *multipath* rutiranja se ogledaju u robusnosti, mogućnosti balansiranja saobraćajnog opterećenja, smanjenom kašnjenju i izbegavanju zagušenja.

Robusnost je direktna posledica raspoloživosti više putanja do jednog odredišta. U slučaju otkaza, *multipath* protokol rutiranja će usmeriti saobraćaj na jedan ili više preostalih putanja ka originalnom odredištu. U zavisnosti od kompleksnosti samog protokola, moguće je da se smanji opterećenje u vidu signalizacionog saobraćaja koji je posledica raspoloživosti redundantnih putanja do istog odredišta pri čemu se istovremeno zadržava isti nivo pouzdanosti.

Balansiranje saobraćaja podrazumeva da će saobraćaj prema istom odredištu biti podeljen u skladu sa pravilom o raspodeli koje će odrediti kojim putanjama će se rutirati pojedinačni delovi saobraćajnog toka. Implementacija strategija distribucije zavisi od većeg broja faktora, kao što je na primer kapacitet putanja, QoS parametri na tim putanjama, kao i servisni zahtevi kojima se mora prilagoditi saobraćaj nad kojim se primenjuje *multipath* rutiranje. Dodatno, balansiranje saobraćaja može da spreči i pojavu zagušenja tako što protokol rutiranja aktivira prerutiranje saobraćaja sa zagušenih putanja na manje zagušene, pa na taj način povećava raspoloživost servisa i bolja iskorišćenost postojećih kapaciteta mreže.

Smanjenje kašnjenja se ogleda pre svega u vremenu koje je potrebno protokolu rutiranja da pronade alternativnu putanju nakon otkaza trenutne putanje preko koje se rutira saobraćaj. Vreme konvergencije je parametar za koga se generalno vezuje sposobnost protokola rutiranja da odredi putanju i da je ubaci u tabelu rutiranja. Naravno, vreme konvergencije zavisi od samog algoritma izračunavanja putanja. Iako *multipath* protokoli rutiranja imaju veće inicijalno vreme konvergencije koje je potrebno da se odrede sve putanje do jednog odredišta, u slučaju otkaza jedne ili više, nije potrebno da se ponovo izračunavaju postojeće jer će saobraćaj sada biti rutiran preko tih putanja. Ovakav način rutiranja koji je posledica osobine raspoloživosti *multipath* rutiranja čini vreme konvergencije neznatnim u odnosu na *single path* rutiranje. Kratko vreme konvergencije je osobina koja je od velike važnosti za aplikacije koje su osetljive na kašnjenje i gubitke paketa – *real time* servisi.

3.2. Značaj mrežne raspoloživosti

Kao jedna od najvažnijih osobina današnjih paketskih mreža jeste raspoloživost koja predstavlja relevantno merilo pouzdanosti i dobrog dizajna mreže. U zavisnosti od toga kog tipa je posmatrana mreža, servis provajderska ili *enterprise* tipa, na raspoloživost pre svega najviše utiče fizička topologija jedne mreže. Prethodni radovi iz oblasti određivanja uticaja i zaštite od višestrukih otkaza [20], [21] uzimaju u obzir samo dva istovremena otkaza pri čemu se otkazi biraju po slučajnom zakonu. Dodatno, istovremeni otkazi linkova ne moraju obavezno da znače da su se desili u isto vreme, već se u okviru takvog razmatranja može desiti da tokom vremenskog prozora određenog za popravku linka koji je otkazao, desi još jedan otkaz. Za razliku od sličnih radova, u [22] se analizira istovremeni otkaz tri linka. U tom

slučaju se definitivno zahteva specijalno planiranje, identifikacija i sagledavanje potencijalnih posledica koje takvi otkazi mogu imati na kritičan QoS saobraćaj. Dodatno, predlaže se i uvođenje odgovarajućih signalizacionih protokola koji će vršiti proračun putanja pre otkaza kako bi date putanje bile spremne u slučaju katastrofalnog otkaza.

Tabela 3.1 Prosečno godišnje vreme otkaza u zavisnosti od broja otkazanih linkova [22]

		Mreža 1	Mreža 2	Mreža 3
Otkaz jednog linka u mreži	Bez katastrofe	404h	361h	263h
	Sa katastrofom	425h	378h	272h
Otkaz dva linka u mreži	Bez katastrofe	10h	8h	4h
	Sa katastrofom	19h	15h	7h
Otkaz tri linka u mreži	Bez katastrofe	0.1h	0.1h	0.04h
	Sa katastrofom	3h	2h	0.6h

U Tabeli 3.1 su prikazane prosečne vrednosti trajanja istovremenog prekida više linkova u okviru tri različite mreže. *Mreža 1* [25] predstavlja prikaz *carrier* mreže koja je dizajnirana sa posebnim osvrtom na robusnost i veći broj različitih putanja u odnosu na ostale. Ova mreža se još i razlikuje po tome što poseduje čak četiri različite kontinentalne putanje što prevazilazi trenutne mogućnosti *carrier* mreže. *Mreža 2* [26] podrazumeva klasičnu *carrier* mrežu sa nešto manjim brojem čvorova i linkova u odnosu na *mrežu 1*, dok *mreža 3* [26] karakteriše mrežu manjeg servis provajdera sa dve kontinentalne putanje što je manje u odnosu na *mrežu 2* koja ima tri putanje ovog tipa. Dodatna karakteristika koja može da pruži indicaciju fizičke topologije i implementirane redundanse jeste broj različitih putanja između para izvor-odredište i koja je prikazana u Tabeli 3.2. Upoređivanjem Tabela 3.1 i 3.2 se primećuje da i pored visokog nivoa redunanse (određene po broju različitih putanja između para izvor-odredište) može se zaključiti da vreme otkaza na godišnjem nivou nije zanemarljivo.

U skladu sa simulacijama koje su rađene u [22] autori su došli do zaključka da čak i postojanje 2 različite putanje između para izvor-odredište u slučaju istovremenog prekida ili katastrofe dovodi do grubog kršenja *Service Level Agreement* (SLA) specifikacija o raspoloživosti čime se moraju razmatrati dodatni mehanizmi zaštite. Kada se uzmu u obzir i različiti nivoi servisa u mrežama kao i eventualne pojave katastrofe kod servis provajdera višeg nivoa (poplave, zemljotresi, požari, itd.), postojanje redundantne fizičke topologije nije dovoljan uslov za postizanje visokog nivoa raspoloživosti, već se komplementarno

moraju koristiti i druge metode zaštite kao i napredni protokoli rutiranja sa malim vremenom konvergencije.

Prilikom analize raspoloživosti jedne topologije, upotrebljava se i *Shared-Risk Link Groups* (SRLG). SRLG pristup tretira grupu linkova koji su tipično smešteni u okviru istog optičkog kabla (cevi) kao grupu koja ima istu verovatnoću usled neke havarije kao što je presecanje kabla. Problem mreža sa SRLG je analiziran i rešavan u mnogim radovima pri čemu većina koristi neku varijantu heurističkog algoritma za pronalaženje zaštite od eventualnih višestrukih otkaza [23], [24].

Tabela 3.2 Broj istovremenih putanja odvojenih linkovima [22]

	Mreža 1	Mreža 2	Mreža 3
2 različite putanje	2145	1493	427
3 različite putanje	575	271	8
4 različite putanje	55	5	0
5 različite putanje	0	1	0

Iako fizička topologija predstavlja preduslov na osnovu kog će se odrediti nivo raspoloživosti, mora se dodatno implementirati i rutiranje, tačnije protokoli rutiranja koji će odrediti kako se vrši rutiranje preko raspoloživih putanja. U zavisnosti od balansa između servisnih zahteva (QoS) i nivoa raspoloživosti u kontekstu *multipath* rutiranja, važnu ulogu imaju algoritmi distribucije saobraćaja i QoS protokoli rutiranja.

3.3. Distribucija saobraćaja preko više putanja

U cilju efikasnog iskorišćenja raspoloživih kapaciteta mreže u prisustvu *multipath* rutiranja trebaju se odrediti pravila distribucije saobraćaja na osnovu kojih će pojedinačni paketi ili *flow* saobraćaja raspodeliti po raspoloživim putanjama. Distribucija saobraćaja se može obaviti na mnogo načina, pa u skladu sa tim relevantno je pomenuti generalnu klasifikaciju algoritama za distribuciju saobraćaja preko više putanja [27].

Multipath prosleđivanje se zasniva na dve funkcije: deljenje saobraćaja i odabir putanje. Deljenje saobraćaja se realizuje na nivou paketa ili na nivou *flow*-a. Posmatrano iz perspektive *multipath* rutiranja, jedinica saobraćaja biće konzistentno i nezavisno transportovana preko jedne putanje koja je prethodno određena u okviru procesa odabira putanje.

Deljenje saobraćaja podrazumeva deobu agregiranog saobraćaja na jedinice koje mogu biti paketi ili *flow*-ovi. U slučaju deljenja saobraćaja na pakete, pretpostavlja se da se svaki paket pojedinačno i nezavisno

rutira preko raspoloživih putanja pri čemu se sama odluka u vezi putanje preko koje će se prenos obaviti donosi nezavisno od ostalih paketa. Paket predstavlja najmanju jedinicu saobraćaja u paketskim mrežama, pa samim tim se u slučaju balansiranja saobraćaja (*load balancing*) može preciznije realizovati raspodela preko više putanja.

Za razliku od distribucije saobraćaja na paketskom nivou, *flow* distribucija podrazumeva da se uzimaju u obzir odgovarajući identifikatori (na primer, izvorišna i/ili odredišna IP adresa) koji se nalaze u zaglavlju paketa i da se na odnosu njihovih vrednosti realizuje odgovarajući model deljenja saobraćaja. Konkretni primeri podrazumevaju da jednom *flow* pripada saobraćaj koji ide ka jednom odredištu ili pripada jednom tipu servisa. Važno je da se primeti, da u slučaju *flow* distribucije saobraćaja ne postoji mogućnost da paketi koji pripadaju individualnom *flow*-u stignu pogrešnim redosledom. Za svaki *flow* odluka o putanju preko koje će biti preneti paketi se donosi za svaki paket zasebno. Specijalni slučajevi *flow* distribucije saobraćaja pretpostavljaju agregaciju više *flow*-ova u jedan *superflow*, dok nasuprot ovome, postoji i *flowlet* koji predstavlja *flow* podskup pri čemu se u okviru jednog *flowlet*-a nalaze paketi sa istim odredištem [27].

Selekcija putanja predstavlja integralni deo *multipath* rutiranja koji će konkretno odrediti način distribucije saobraćaja preko raspoloživih putanja u zavisnosti od jedinice saobraćaja. U ovom radu se kroz analizu paketske disperzije upravo posmatra paketska selekcija putanja što podrazumeva da se prema modelu selekcije odabira putanja koju će naredni paket koristiti. U nastavku će biti pomenuti paketski i *flow* modeli selekcije putanja.

3.3.1. Mehanizmi deljenja saobraćaja na paketskom nivou

Round robin model predstavlja najjednostavniji model paketskog odabira putanja tako što se *round robin* režim primenjuje prilikom odabira putanje za naredni paket u redu. Najvažnije osobine ovog modela selekcije jesu jednostavnost u implementaciji kao i odustvo eventualnog zagušenja usled čekanja u redu s obzirom da ne postoji putanja odnosno red koji se preferira u odnosu na druge putanje. Međutim, jedna od glavnih mana jeste nemogućnost da se kontroliše koliko paketa u *round robin* ciklusu se preraspodeli po putanjama, pa samim tim putanje sa manjim kapacitetom primenom ovog načina selekcije bivaju brže opterećene saobraćajem. Drugi problem sa *round robin* selekcijom putanja jeste pristizanje paketa na odredište drugačijim rasporedom u odnosu na raspored kojim su poslani – *out of order*. Ovo je naravno posledica različitog kašnjenja koje se iskazuje na različitim putanjama. Pojedine aplikacije i mehanizmi kao što je *Equal-Cost Multi-Path* (ECMP) rutiranje može da koristi ovu vrstu selekcije putanja radi balansiranja saobraćaja [28].

Weighted Round Robin (WRR) selekcija putanja se pojavljuje kao rezultat problema preraspodele saobraćaja po putanjama sa različitim kapacitetom. WRR koncept selekcije putanja podrazumeva da se pojedinim putanjama dodele odgovarajuće težinske vrednosti koje će u određenoj srazmeri odgovarati

kapacitetu posmatrane putanje. Na taj način će se odrediti broj paketa koji će biti poslat u okviru jednog *round robin* ciklusa preko određene putanje. Veći težinski faktor će srazmerno odrediti koliko će paketa više biti poslato preko posmatrane putanje u odnosu na ostale putanje sa manjim težinskim faktorom. Problemi kod WRR primena se mogu dogoditi ukoliko se nepravilno dodele prioriteta koji ne odražavaju tačno kapacitete linkova odnosno putanja. Nasuprot ovome, mogu se desiti i problemi sa promenljivom veličinom paketa koja će uzrokovati da kratkovremeno pojedine putanje premaše veličinu saobraćaja koji je diktiran od strane dodeljenih težinskih faktora. O sličnom problemu u ovom radu je već diskutovano u okviru algoritama za opsluživanje paketa.

Osim pomenute dva mehanizma selekcije putanja koji se baziraju na *round robin* režimu, postoje i nekoliko specijalnih slučajeva kao što su *Weighted Interleaved Round Robin* (WIRR) [29] i *Surplus Round Robin* (SRR) [30].

U ovom radu termini selekcija putanja na paketskom nivou i paketska disperzija označavaju proces određivanja algoritma odnosno strategije u skladu sa kojom će se paketi slati po raspoloživim putanjama do odredišta. Pod paketskom disperzijom se podrazumevaju samo algoritmi koji kao jedinicu saobraćaja podrazumevaju paket, a ne *flow*.

3.3.2. Mehanizmi deljenja saobraćaja na *flow* nivou

Što se tiče selekcije putanja na *flow* nivou, postoji takođe nekoliko značajnih algoritama po kojima se vrši raspored pojedinačnih *flow*-ova po putanjama. Neki od značajnih mehanizama selekcije putanja jesu *Fast Switching* (FS), *Direct Hashing* (DH), *Table-Based Hashing* (TH), *Hash Threshold* (HT), *Highest Random Weight* (HRW), *Primary Number Modulo-N Load Balance* (PMN-LB).

FS predstavlja mehanizam prosleđivanja saobraćaja koji je implementiran u hardveru. Funkcioniše tako što se svi paketi koji pripadaju istom *flow* šalju na istu putanju, dok pojava novog *flow*-a rezultuje slanjem paketa na novu putanju u skladu sa *round robin* ciklusom. FS mehanizam teži da postigne da podjednak broj *flow*-ova bude zastupljen po jednoj putanji, ali ne uzima u obzir veličinu paketa u okviru posmatranih *flow*-ova što može da dovede do eventualnog disbalansa između opterećenosti pojedinih putanja. Značajni parametar kod FS *flow* selekcije putanja jeste veličina memorije u kojoj je sadržano mapiranje *flow*-putanja, pa samim tim su izraženi problemi sa skalabilnošću u slučaju velikog broja *flow*-ova i neadekvatnog memorijskog prostora za ove potrebe. Dodatno, u slučaju nedostatka memorijskog prostora, najstariji *flow* se briše iz memorije, dok se pristizanjem novih paketa tog *flow*-a on ubacuje ponovo u memoriju što može rezultirati odabirom nove putanje i pojavom *out of order* paketa.

DH predstavlja mehanizam koji se uveliko koristi za proces selekcije putanja na nivou *flow*-a. Koriste ga mnogi protokoli rutiranja koji su sposobni za *multipath* rutiranje. Funkcioniše tako što se primeni *hash*

funkcija nad identifikatorom paketa (najčešće određenoj IP adresom, ali se mogu koristiti i više parametara iz zaglavlja paketa). Rezultat te *hash* funkcije se dalje koristi u moduo funkciji gde se traži moduo ove vrednosti od broja raspoloživih putanja. DH suštinski predstavlja algoritam sa malom računarskom kompleksnošću i procesorskim zahtevima. Međutim, kao i njegov prethodnik FS, nije u stanju da uzme u obzir promenljivu veličinu pojedinačnih *flow*-ova, pa samim tim je moguće da primenom ovog algoritma pojedini odlazni interfejsi budu zagušeni usled neadekvatnog balansiranja opterećenja. Naravno, inherentna mana je i ta da ukoliko različiti *flow*-ovi kao rezultat *hash* i moduo funkcije daju istu vrednost, koristiće se samo jedna putanja za sve posmatrane pakete koji pripadaju različitim *flow*-ovima.

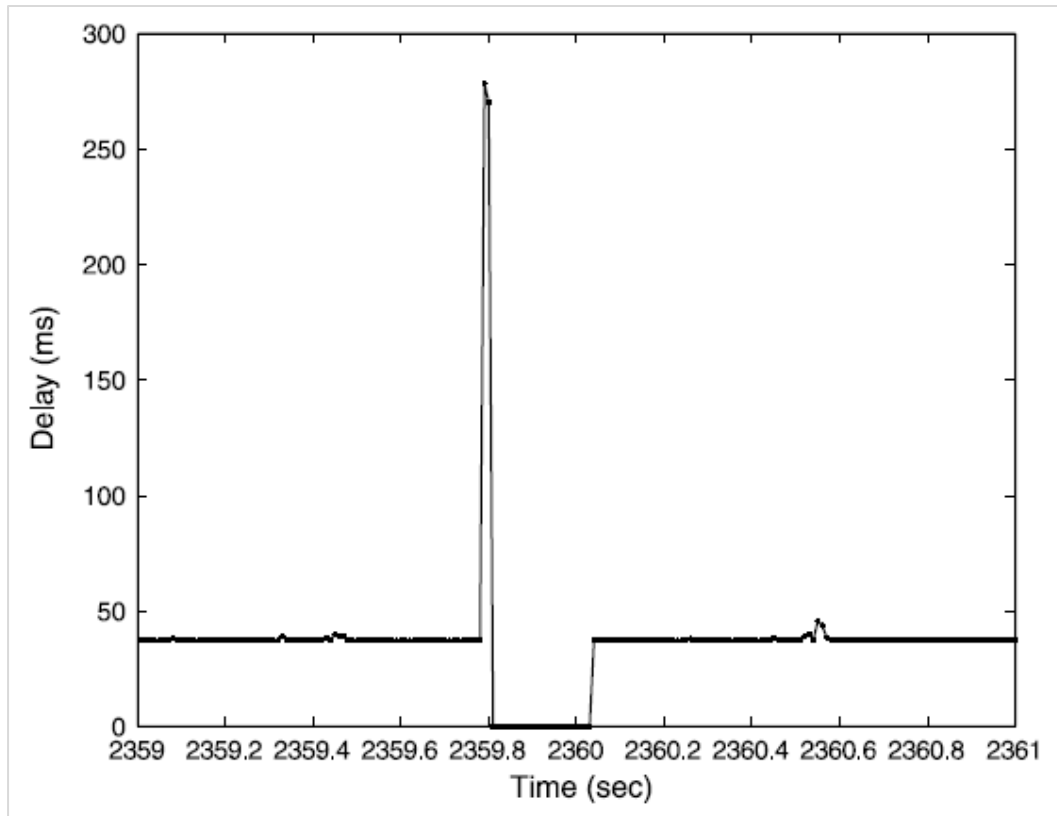
TH [31] i HT [32] mehanizmi se često koriste u okviru ECMP rutiranja pri čemu je primenom *hash* funkcija omogućeno donekle fleksibilno mapiranje rezultata ove funkcije na raspoložive putanje. TH mehanizam se često poistovećuje sa DH jer je implementirano 1:1 mapiranje rezultata *hash* funkcije i putanje. Međutim, glavna razlika između TH i DH mehanizma jeste upravo mogućnost da se funkcija mapiranja prilagodi osobinama raspoloživih putanja u odgovarajućem odnosu. Dodatno, varijabilnost u *flow* veličini može da dovede do neefikasne raspodele saobraćaja između putanja. Sa druge strane, HT mehanizam podrazumeva uvođenje veće granularnosti u raspodelu saobraćaja tako što se uvodi pojam regiona koji označava broj *flow*-ova koji se mogu pridružiti jednoj putanji. Putanja se bira tako što se za svaki paket određuje *hash* vrednost koja se smešta u odgovarajući region koji je pridružen određenoj putanji. Bolja granularnost omogućava tačnije i bolje balansiranje opterećenja u odnosu na TH mehanizam.

3.4. Modelovanje i relevantni parametri gubitaka paketa

Pojava gubitaka paketa predstavlja jedan od najznačajnijih faktora koji utiču na kvalitet servisa koji su osetljivi na QoS u okviru servis provajderskih mreža, a naročito u slučaju VoIP servisa. Merenja obavljena u [33] pokazuju da u servis provajderskim mrežama za VoIP servis u pogledu kvaliteta najveći uticaj ima gubitak paketa, dok ostali QoS parametri kao što su kašnjenje i džiter skoro i da nemaju na uticaj subjektivni kvalitet koji je percipiran od strane korisnika. Dodatno, VoIP kao servis koji je sada već postao deo standardne ponude servis provajdera treba da bude dovoljno robusan kako bi se uspeo da se orealizuje kvalitet koji korisnici očekuju.

Iako je VoIP servis reprezentativan servis u pogledu QoS osetljivosti, ostali servisi koji nisu *real time* se oslanjaju na TCP da reguliše eventualne gubitke paketa posredstvom retransmisija. Naravno, TCP nije adekvatan transportni protokol za *real time* servise kakav je VoIP s obzirom da retransmisija paketa usled strogih zahteva po pitanju kašnjenja nije opcija. Što se tiče servisa koji se oslanjaju na TCP, potrebno je naglasiti da iako poseduju mehanizme retransmisije koji u slučaju gubitaka paketa sprečavaju da isti budu nepovratno izgubljeni, uticaj izgubljenih paketa deluje na ukupne TCP performanse. TCP performanse direktno zavise od mehanizma kontrole zagušenja – *congestion window*, koji u slučaju gubitaka paketa se

smanjuje i direktno utiče na propusnost u okviru TCP sesije. U zavisnosti od primenjenog TCP algoritma [34], [35], [36], [37] mehanizam zagušenja će imati različitu dinamiku menjanja veličine TCP segmenta što znači da će odabir algoritma imati uticaja na brzinu oporavka TCP sesije od gubitka paketa.



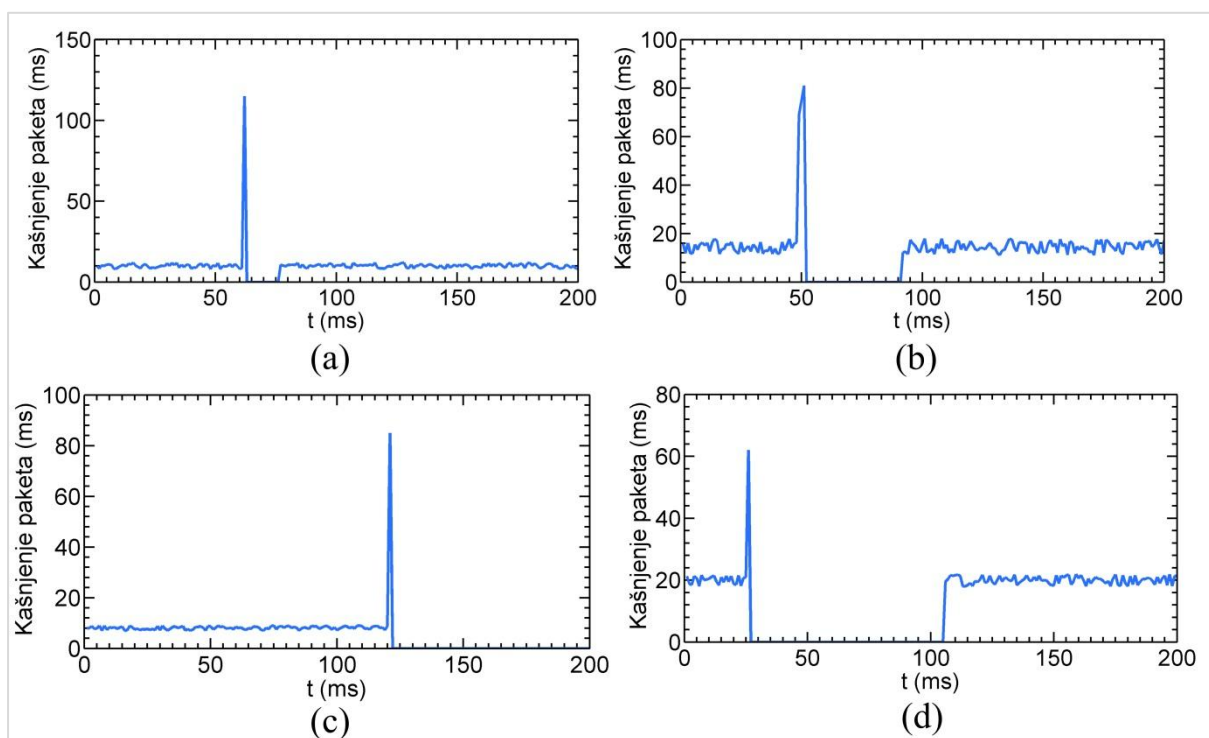
Slika 3.1 Primer elementarnog događaja od 23 uzastopna gubitka [40].

Ranija istraživanja obavljena na temu gubitaka paketa na Internetu [38], [39], [40], [41] ukazuju da se gubici uglavnom pojavljuju u rafalnom obliku. Pod rafalnim gubicima se podrazumeva da postoje periodi u okviru kojih se gubici paketa češće ispoljavaju nakon čega ponovo sledi stanje bez ili sa malim gubicima paketa. Takođe, podrazumeva se da je pojedinačni gubitak paketa statistički nezavisan od svih prethodnih i budućih gubitaka. Međutim, posmatrajući uzorke gubitaka saobraćaja koji su analizirani u [41], jasan je zaključak da su rafalni gubici paketa mnogo češća pojava u realnim mrežama.

Rafalni gubici paketa ispoljavaju *Long Range Dependence* (LRD) i *Short Range Dependence* (SRD), što podrazumeva da događaj gubitka ima veći ili manji efekat na pojavu narednih gubitaka paketa. Sa stanovišta QoS, rafalni gubici imaju izuzetno poguban efekat imajući u vidu da struktura gubitaka paketa može da utiče na performanse servisa i posledično na kvalitet percepiran od strane krajnjeg korisnika, što je od izuzetnog značaja za *real time* servise.

Treba se i naglasiti korelacija između gubitaka paketa i kašnjenja. Naime, analizi gubitaka paketa realizovanoj u [40] pristupilo se tako što su se identifikovala dva tipa događaja gubitaka: (1) elementarni i

(2) kompleksni. Elementarni događaji podrazumevaju postojanje uzastopnih gubitaka paketa. Ti gubici se mogu sastojati od jednog, pa do nekoliko uzastopnih paketa koji mogu trajati i nekoliko sekundi. Primer od 23 uzastopna gubitka u korelaciji sa kašnjenjem prikazan je na slici 3.1 [40]. Prikazano je kašnjenje mernih paketa pri čemu paketi koji imaju nulto kašnjenje odgovaraju izgubljenim paketima. Ono što je naročito interesantno jeste pojava velikog kašnjenja paketa koji prethode gubicima paketa i koji se javljaju u nekoliko provajderskih mreža za šta autori ove analize nisu imali objašnjenje.



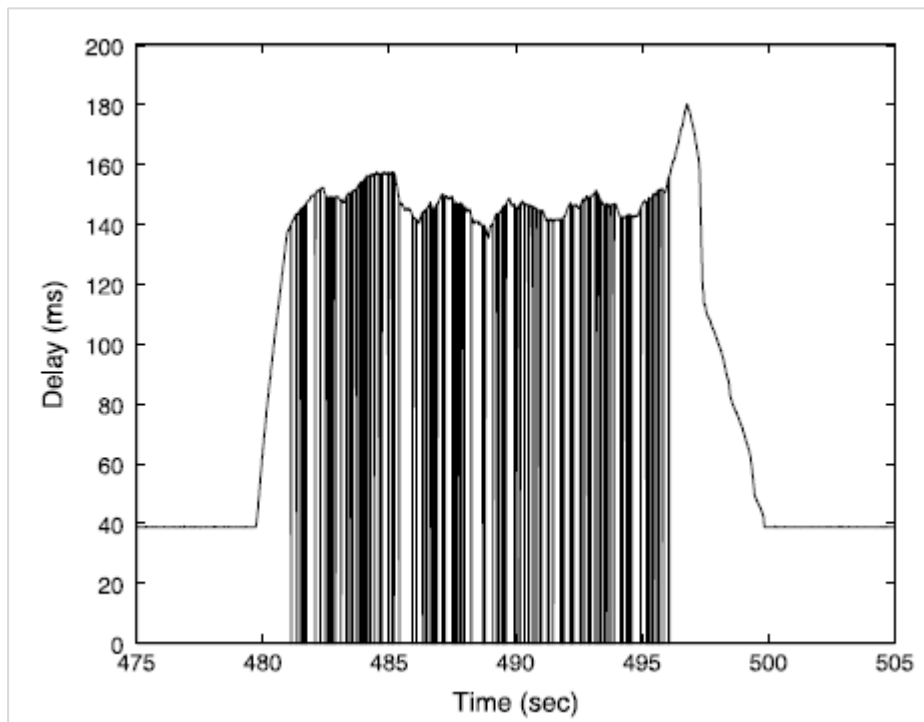
Slika 3.2 Kašnjenje paketa u AMRES2 mreži na degradiranim linkovima: (a)-(b) neispravan transiver; (c)-(d) degradirano vlakno.

Za potrebe objašnjenja prethodnog fenomena koji uključuje povećano kašnjenje koje prethodi uzastopnim gubicima, slični testovi su rađeni u AMRES2¹ mreži, koji su prikazani na slici 3.2. Slika 3.2, pri čemu se ispostavilo da se ovakvi gubici ispoljavaju kada postoje fizički nedostaci na vezi između rutera. Konkretno, na optičkom vlaknu na kome su vršena merenja kašnjenja su uočena značajna slabljenja koja su prevazilazila optički domet transivera (optičkog primopredajnika). Rezultati merenja na nekoliko sličnih segmenata ukazuju da sa velikom verovatnoćom možemo zaključiti da je uzrok izraženog kašnjenja koji prethode gubicima problem na fizičkom nivou.

Kompleksni događaji gubitaka paketa pretpostavljaju pojavu više elementarnih događaja gubitaka kada se posmatra kratak period (u [40] je naveden period od 50s) tokom koga je verovatnoća gubitaka paketa do 80%. Primer kompleksnog događaja koji se sastoji od elementarnih događaja pri čemu se svaki od tih

¹ Akademska mreža Republike Srbije

događaja sastoji od gubitka jednog paketa. Na slici 3.3 je prikazan kompleksan događaj koji traje oko 15 s i u okviru koga se paketi gube sa verovatnoćom od 9.4%.



Slika 3.3 Primer kompleksnog događaja koji se sastoji od pojedinačnih gubitaka paketa [40].

U pogledu raspoloživosti *core* dela mreže, može se smatrati da se ispoljavaju dosta dobre performanse u vidu kašnjenja, džitera i gubitaka paketa. Međutim, merenjima se mogu identifikovati putanje koje su konzistentno loših performansi i raspoloživosti usled otkaza, kompleksnih događaja gubitaka i visokih kašnjenja. Na osnovu prethodnog se može zaključiti da se u slučaju QoS aplikacija moraju koristiti mehanizmi kako na nivou mreže tako i na nivou krajnjih uređaja koji će pomoći da se ostvari zahtevani QoS nivo kako bi aplikacije funkcionisale na očekivani način.

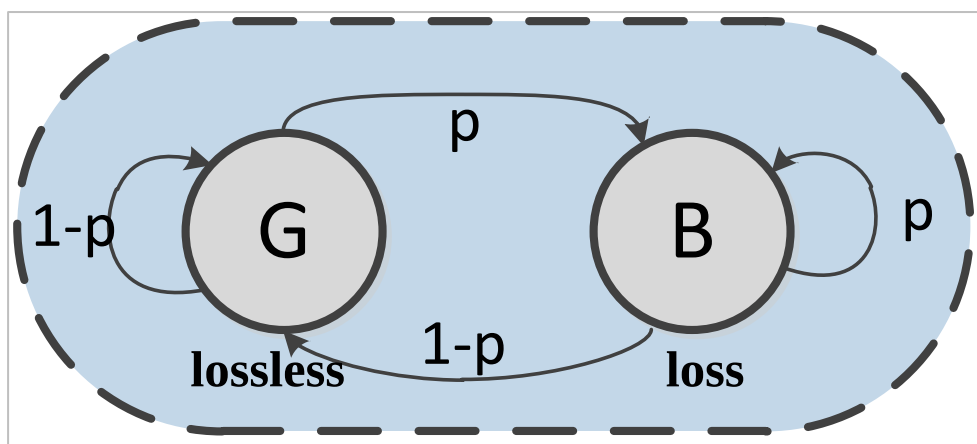
Razlozi pojave gubitaka, kako pojedinačnih, tako i elementarnih i kompleksnih mogu biti višestruki pri čemu identifikacija uzroka gubitaka podrazumeva detaljno poznavanje detaljne strukture jedne mreže, od fizičkog, pa do aplikacionog nivoa. Pojedinačni gubici paketa, tačnije elementarni događaji gubitaka su najčešće posledica odbacivanja rutera usled *tail drop* discipline, odnosno neadekvatnog kapaciteta bafera. Dodatno, primenom RED mehanizma, može sporadično doći do preventivnog odbacivanja paketa koji uzorkuje elementarne događaje. Dodatno, duži prekidi mogu se smatrati kao posledica promena u politici rutiranja, pri čemu samo vreme prekida zavisi od sposobnosti vremena konvergencije korišćenog protokola rutiranja. Što se tiče ostalih razloga prekida na duže vreme, oni se sa velikom verovatnoćom mogu pripisati prekidima veza ili redovnim radovima u cilju održavanja mreže. Bitno je naglasiti da je moguća pojava elementarnih i kompleksnih događaja istovremeno na više putanja što može dati kao indicaciju isti

problem na deljenom linku odnosno resursu. Upravo je istovremena pojava odgovarajuće strukture gubitaka paketa na različitim putanjama razlog implementacije paketske disperzije.

Za razliku od rafalnih gubitaka, biće analizirana i pojava Bernulijevih gubitaka paketa, tačnije paketa čija je pojava statistički nezavisna od prethodnih i budućih gubitaka paketa. Po pitanju analize, kompleksnija je naravno analiza rafalnih gubitaka, dok se za opisivanje rafalnih gubitaka moraju uvesti dodatni parametri kako bi se sama struktura gubitaka paketa mogla analitički opisati i staviti u kontekst uticaja na kvalitet servisa. Analiza će obuhvatiti matematičke modele koji se koriste kako bi se okarakterisali Bernulijevi i rafalni gubici, a potom će za te modele biti definisane odgovarajući parametri za opisivanje strukture gubitaka paketa. U ovom radu, posebno se apostrofira analiza Markovljevog modela sa 4 stanja kao modela sposobnog da modeluje rafalne gubitke istovremeno poštujući balans između tačnosti modelovanja i računske kompleksnosti.

3.4.1. Modelovanje gubitaka paketa posredstvom Bernulijevog modela

Primenom Bernulijevog modela gubitaka paketa moguće je modelovati gubitke čija je pojava nezavisna. Bernulijev model je prikazan na slici 3.4. Generalno gledano, njega je moguće predstaviti kao Markovljev model sa dva stanja, tj. dobro stanje (*good state*) u kome nema gubitaka i loše stanje (*bad state*) u kome se nalaze gubici paketa. Tranziciona verovatnoća p označava verovatnoću prelaska iz jednog stanja u drugo i kao što se sa slike 3.4. može primetiti, upravo je ova tranziciona verovatnoća jedini relevantan parametar za prelazak iz jednog u drugo stanje što implicira da je verovatnoća gubitaka $P_{loss}=p$, dok je verovatnoća prelaska iz dobrog u loše stanje $1-P_{loss}=1-p$.

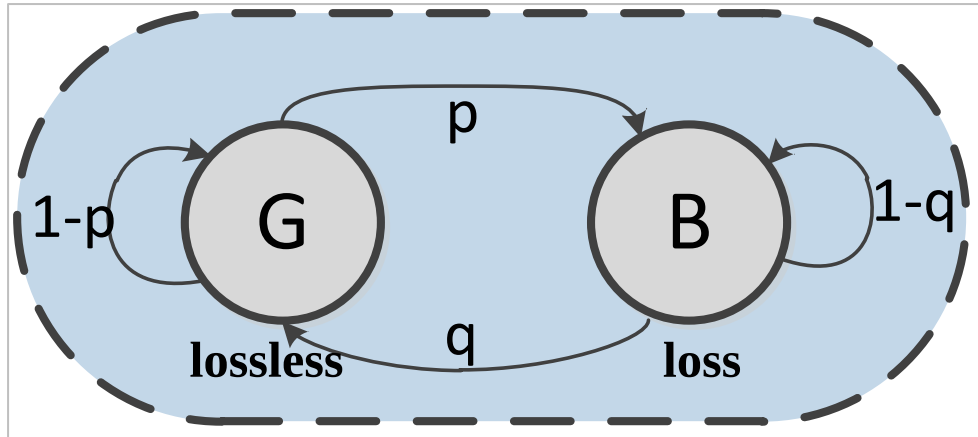


Slika 3.4 Bernulijev model gubitaka paketa.

Bernulijev model generiše u potpunosti nekorelisane gubitke što može da odgovara realnim gubicima u pojedinim slučajevima. Ovaj model predstavlja najjednostavniji model koji se koristi za generisanje gubitaka paketa s obzirom da nema zahteva za kreiranjem odgovarajuće strukture pojave gubitaka paketa, kao što je na primer dužina rafala gubitaka, određena verovatnoća gubitaka rafala i sl.

3.4.2. Modelovanje gubitaka paketa posredstvom Gilbertovog modela

Gilbertov model [42] se koristi za modelovanje rafalnih gubitaka paketa i poseduje dva stanja i dve tranzicione verovatnoće koje označavaju verovatnoću prelaska iz dobrog u loše stanje p i verovatnoću prelaska iz lošeg u dobro stanje q . Ovaj Gilbertov model je adekvatan za modelovanje uzastopnih gubitaka paketa pri čemu se odabirom vrednosti tranzicionih verovatnoća može granularno odrediti koliki niz uzastopnih gubitaka je potreban. Gilbertov model prikazan je na slici 3.5.



Slika 3.5 Modelovanje uzastopnih gubitaka posredstvom Gilbertovog modela.

Pretpostavljajući da verovatnoća da se model nađe u dobrom stanju i lošem stanju iznosi S_G i S_B , respektivno, može se postaviti sledeća jednakost:

$$S_G \cdot p = S_B \cdot q \quad (3.1)$$

po kojoj postoji verovatnoća da sistem iz dobrog stanja može da pređe u loše stanje bude jednaka verovatnoći da sistem pređe iz lošeg u dobro stanje. Definisanjem naredne trivijalne jednakosti, može se pojedinačno odrediti verovatnoća koja određuje da je sistem u dobrom ili lošem stanju:

$$S_G + S_B = 1 \quad (3.2)$$

Rešavanjem prethodne dve jednakosti dobija se verovatnoća da je sistem u dobrom stanju S_G odnosno lošem stanju S_B :

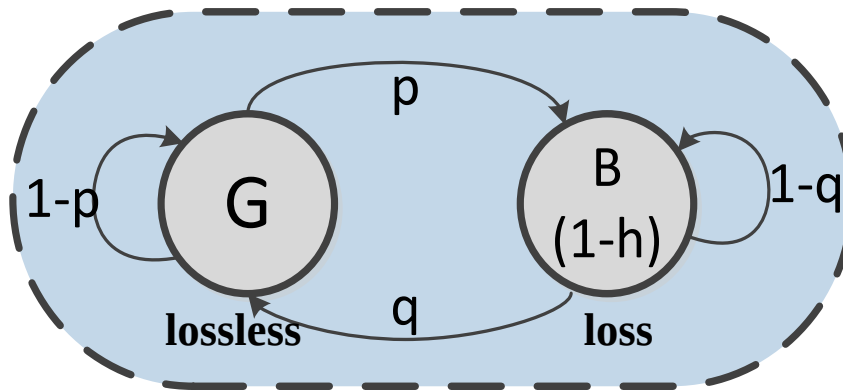
$$S_G = \frac{q}{p+q} \quad (3.3)$$

$$S_B = \frac{p}{p+q} \quad (3.4)$$

Imajući u vidu ovako generisanu strukturu rafalnih gubitaka, u nastavku su definisani parametri kojima se može detaljnije opisati struktura pojave gubitaka. Posmatrajući verovatnoću gubitaka paketa P_{loss} može se zaključiti da se gubici prema Gilbertovom modelu događaju samo kada je sistem u lošem stanju, pa samim tim važi jednakost $P_{loss} = S_B$. Rafalni period B predstavlja vreme koje sistem provede u lošem stanju, i prema Gilbertovom modelu on je jednak:

$$B = \frac{1}{q} \quad (3.5)$$

Rafalni period je relevantan parametar za opisivanje nivoa rafalnosti koja je izražena u okviru niza gubitaka paketa. Prilikom ocene kvaliteta servisa i komparacije različitih struktura gubitaka paketa, ovaj parametar doprinosi adekvatnoj oceni uticaja strukture gubitaka na performanse i kvalitet aplikacije.



Slika 3.6 Modelovanje korelisanih gubitaka posredstvom Gilbertovog modela.

Osim prethodno opisanog Gilbertovog modela, postoji i Gilbertov model koji je sposoban da generiše korelisane gubitke paketa koji su u mogućnosti da tačnije reprodukuju događanja u kontekstu elementarnih i kompleksnih događaja kakva su definisana u [40]. Ovaj model prikazan je na slici 3.6.

Ovaj oblik Gilbertovog modela ima takođe dva stanja, ali sada osim tranzicionih verovatnoća p i q uvodi i novu verovatnoću koja određuje gustinu gubitaka $1-h$. Gustina predstavlja verovatnoću da se paket izgubi kada se sistem nađe u lošem stanju. Dakle, u odnosu na „prostiju“ verziju Gilbertovog modela, kada je $h=0$ i kada se uvek paket gubi kada se sistem nađe u lošem stanju, sada je moguće specificirati verovatnoću gubitaka u lošem stanju što predstavlja gustinu gubitaka paketa. Dakle, moguće je da se detaljnije specificira struktura gubitaka u okviru lošeg stanja tako što će između uzastopno izgubljenih paketa neki od paketa biti uspešno primljeni. Ovakav pristup modelovanju tačnije reflektuje strukturu gubitaka u realnim mrežama u odnosu na Bernulijev i prethodni Gilbertov model koji je podrazumevao da su svi paketi u lošem stanju izgubljeni.

Što se tiče prethodno definisanih parametara za opisivanje strukture gubitaka paketa, potrebno je uvesti odgovarajuće korekcije u odnosu na relacije izvedene za prethodni Gilbertov model. Naime, verovatnoća gubitaka paketa P_{loss} mora da uzme u obzir gustinu gubitaka u lošem stanju, pa je jednaka:

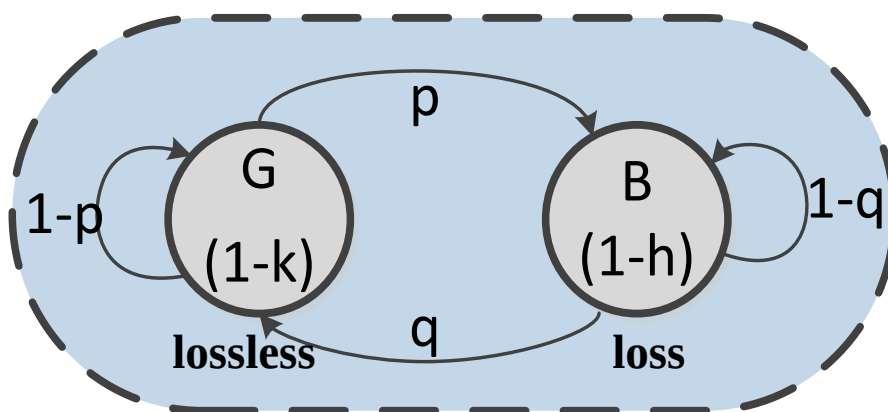
$$P_{loss} = S_B \cdot (1 - h) = \frac{p}{p+q} \cdot (1 - h) \quad (3.6)$$

gde je sada gustina gubitaka $1-h$ pomnožena sa verovatnoćom sistema da bude u lošem stanju. Prema definiciji rafala gubitaka u [44], rafalni period B neće zavisiti od gustine gubitaka u lošem stanju, pa ostaje da važi izraz (3.5).

3.4.3. Modelovanje gubitaka paketa posredstvom Gilbert-Eliotovog modela

Gilbert-Eliotov [43] modela predstavlja Markovljev model sa dva stanja i u odnosu na prethodne Gilbertove modele pretpostavlja da se gubici paketa mogu desiti i kada je sistem u dobrom stanju (slika 3.7 **Slika 3.7**). Osim tranzicionih verovatnoća p i q kao i gustine gubitaka $1-h$ u lošem stanju, uvodi se sada i termin izolovanih gubitaka u dobrom stanju $1-k$. Dakle, u oba stanja Gilbert-Eliotovog modela su mogući gubici paketa pri čemu se ovaj model svodi na prethodni Gilbertov model ako se pretpostavi da je $k=0$.

Treba naglasiti da su gubici paketa koji se dešavaju u dobrom i u lošem stanju potpuno nezavisni događaji. Odabir ovih vrednosti treba da bude takav da je verovatnoća gubitaka u lošem stanju nekoliko puta veća od verovatnoće u dobrom stanju kako bi bilo moguće razlikovati pojavu rafala gubitaka u odnosu na izolovane gubitke. Izolovani gubici se ne bi trebali zameniti sa rafalima iz lošeg stanja s obzirom da je njihova struktura takva da se po definiciji ne bi trebalo podvesti pod rafalne gubitke.



Slika 3.7 Modelovanje gubitaka posredstvom Gilbert-Eliotovog modela.

Verovatnoća gubitaka paketa P_{loss} je slična izrazu (3.6) kada se uzme u obzir i postojanje izolovanih gubitaka $1-k$:

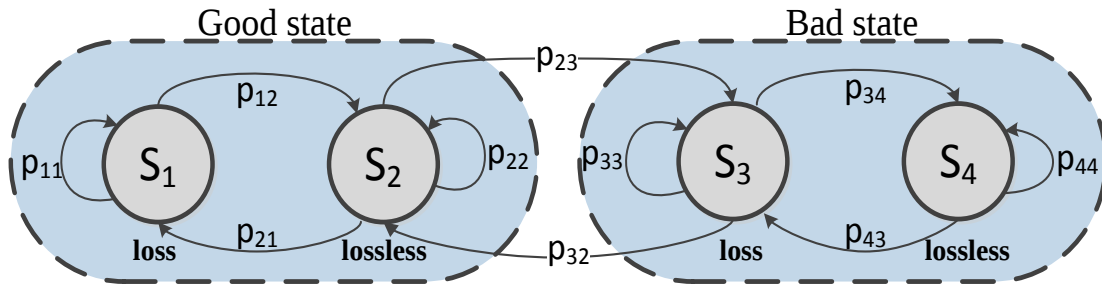
$$P_{loss} = S_B \cdot (1 - h) + S_G \cdot (1 - k) = \frac{p}{p+q} \cdot (1 - h) + \frac{q}{p+q} \cdot (1 - k) \quad (3.7)$$

Što se tiče ostalih parametara, rafalni period B ostaje kao u izrazu (3.5), dok je gustina gubitaka jednaka $1-h$ i verovatnoća izolovanih gubitaka paketa je jednaka $1-k$.

3.4.4. Modelovanje gubitaka paketa posredstvom Markovljevog modela sa 4 stanja

Uvođenje Markovljevog modela sa četiri stanja je rezultat potrebe da se struktura gubitaka paketa što tačnije modeluje u poređenju sa realnim rafalnim gubicima. Analiza VoIP gubitaka koja je urađena u [45] i [46] dovela je do zaključka da se realni gubici paketa koji su izraženi kod VoIP servisa tačnije modeliraju primenom Markovljevog modela sa četiri stanja u odnosu na modele sa dva stanja, dok je ovaj model prvobitno predložen za primenu modelovanja VoIP gubitaka u [47].

Markovljev model sa četiri stanja je prikazan na slici 3.8. Dobro stanje se sastoji od S_1 i S_2 stanja, dok se loše stanje sastoji od S_3 i S_4 stanja. Svaka tranzicija u S_1 i S_3 stanje rezultuje izgubljenim paketom, dok tranzicije ka S_2 i S_4 stanju podrazumevaju uspešno primljen paket.



Slika 3.8 Modelovanje gubitaka posredstvom Markovljevog modela sa 4 stanja.

Imajući u vidu postojanje četiri stanja, kod ovog Markovljevog modela sada postoji veći broj tranzicionih verovatnoća koje je najpreglednije prikazati u matričnom obliku. U tom smislu, definiše se tranziciona matrica koja je jednaka:

$$T = \begin{bmatrix} p_{11} & p_{12} & 0 & 0 \\ p_{21} & p_{22} & p_{23} & 0 \\ 0 & p_{32} & p_{33} & p_{34} \\ 0 & 0 & p_{43} & p_{44} \end{bmatrix} \quad (3.8)$$

u okviru koje su definisane odgovarajuće tranzicione verovatnoće u skladu sa slikom 3.8. Polazeći od sledećih jednakosti verovatnoća stanja sistema:

$$S_1 \cdot p_{12} = S_2 \cdot p_{21} \quad (3.9)$$

$$S_4 \cdot p_{43} = S_3 \cdot p_{34} \quad (3.10)$$

$$S_4 \cdot p_{43} = S_3 \cdot p_{34} \quad (3.11)$$

$$S_2 \cdot (p_{21} + p_{23}) = S_3 \cdot p_{32} + S_1 \cdot p_{12} \quad (3.12)$$

$$S_1 + S_2 + S_3 + S_4 = 1 \quad (3.13)$$

kao i relacije između tranzicionih verovatnoća koje direktno proizilaze iz modela:

$$p_{11} = 1 - p_{12} \quad (3.14)$$

$$p_{22} = 1 - p_{23} - p_{21} \quad (3.15)$$

$$p_{33} = 1 - p_{32} - p_{34} \quad (3.16)$$

$$p_{44} = 1 - p_{43} \quad (3.17)$$

dobija se verovatnoća sistema da bude u pojedinačnim stanjima:

$$S_1 = \frac{p_{21} \cdot p_{32} \cdot p_{43}}{p_{21} \cdot p_{32} \cdot p_{43} + p_{12} \cdot p_{32} \cdot p_{43} + p_{12} \cdot p_{23} \cdot p_{43} + p_{12} \cdot p_{23} \cdot p_{34}} \quad (3.18)$$

$$S_2 = \frac{p_{12} \cdot p_{32} \cdot p_{43}}{p_{21} \cdot p_{32} \cdot p_{43} + p_{12} \cdot p_{32} \cdot p_{43} + p_{12} \cdot p_{23} \cdot p_{43} + p_{12} \cdot p_{23} \cdot p_{34}} \quad (3.19)$$

$$S_3 = \frac{p_{12} \cdot p_{23} \cdot p_{43}}{p_{21} \cdot p_{32} \cdot p_{43} + p_{12} \cdot p_{32} \cdot p_{43} + p_{12} \cdot p_{23} \cdot p_{43} + p_{12} \cdot p_{23} \cdot p_{34}} \quad (3.20)$$

$$S_4 = \frac{p_{12} \cdot p_{23} \cdot p_{34}}{p_{21} \cdot p_{32} \cdot p_{43} + p_{12} \cdot p_{32} \cdot p_{43} + p_{12} \cdot p_{23} \cdot p_{43} + p_{12} \cdot p_{23} \cdot p_{34}} \quad (3.21)$$

Na osnovu izvedenih verovatnoća pojedinačnih stanja, dalje se mogu odrediti i ostali relevantni parametri neophodni za tačan opis strukture gubitaka paketa. Verovatnoća gubitaka paketa P_{loss} je jednaka zbiru verovatnoća stanja u kojima se dešavaju gubici, dakle S_1 i S_3 :

$$P_{loss} = S_1 + S_3 \quad (3.22)$$

dok je gustina gubitaka u lošem stanju P_{burst} sada jednaka:

$$P_{burst} = \frac{S_3}{S_3 + S_4} \quad (3.23)$$

a izolovani gubici P_{isol} iznosi:

$$P_{isol} = \frac{S_1}{S_1 + S_2} \quad (3.24)$$

Određivanje dužine rafalnog perioda B je nešto komplikovaniji proces koji zahteva razumevanje definicije rafala gubitaka. Naime, rafalni period predstavlja dužinu vremena koje sistem provede u okviru lošeg stanja koje se sastoji od pojedinačnih stanja sistema S_3 i S_4 . U skladu sa tim, rafalni period se određuje kao srednji broj paketa (vreme se određuje tako što se trajanje paketa množi sa brojem paketa, pri čemu jedan paket odgovara jednoj tranziciji sistema) prema sledećem izrazu:

$$B = \sum_{k=0}^{\infty} (E(S_3) + k \cdot (E(S_3) + E(S_4))) \cdot \frac{p_{32}}{p_{32} + p_{34}} \cdot \left(\frac{p_{34}}{p_{32} + p_{34}} \right)^k \quad (3.25)$$

pri čemu k određuje trenutni korak tranzicije u okviru lošeg stanja, dok $E(S_3)$ i $E(S_4)$ predstavljaju srednje vreme provedeno u stanjima S_3 i S_4 , respektivno, i oni su jednaki:

$$E(S_3) = \frac{1}{p_{32} + p_{34}} \quad (3.26)$$

$$E(S_4) = \frac{1}{p_{43}} \quad (3.27)$$

Oblik izraza (3.25) se može dodatno uprostiti tako što se iskoriste prethodni izrazi (3.26) i (3.27) pri čemu će se iskoristiti poznata jednakost:

$$\frac{d}{dP} \left(\sum_{k=0}^{\infty} P^k \right) = \frac{1}{(1-P)^2} \quad (3.28)$$

pa će se nakon sređivanja dobiti kompaktan izraz za rafalni period B kod Markovljevog modela sa 4 stanja:

$$B = \frac{p_{43} + p_{34}}{p_{32} \cdot p_{43}} \quad (3.29)$$

3.4.5. Mapiranje Gilbert-Eliotovog modela i Markovljevog modela sa 4 stanja

Prethodno dati načini modelovanja gubitaka paketa se razlikuju po svojoj tačnosti gde je uvek prisutan balans između tačnosti modelovanja i kompleksnosti računanja modela. Naime, Markovljevi modeli višeg reda tačnije modeluju gubitke paketa koji su uočeni u realnim uzorcima IP saobraćaja.

Modeli sa 2 stanja mogu biti opisani posredstvom Markovljevog modela sa 4 stanja, pri čemu se kod Gilbert-Eliotovog modela dobro stanje (sa verovatnoćom gubitaka $1-k$) mapira u dobro stanje Markovljevog modela koji se sastoji iz stanja S_1 i S_2 . Za mapiranje lošeg stanja važi obrnuto kao što je i naznačeno na slici 3.9. Na osnovu ovih pretpostavki, mogu se postaviti sledeće jednakosti za verovatnoću dobrog stanja Gilbert-Eliotovog modela S_B izražene posredstvom tranzicionih verovatnoća Markovljevog modela sa 4 stanja:

$$S_B = \frac{p_{23}}{p_{23} + p_{32}} \quad (3.30)$$

pri čemu slično važi i za dobro stanje S_G :

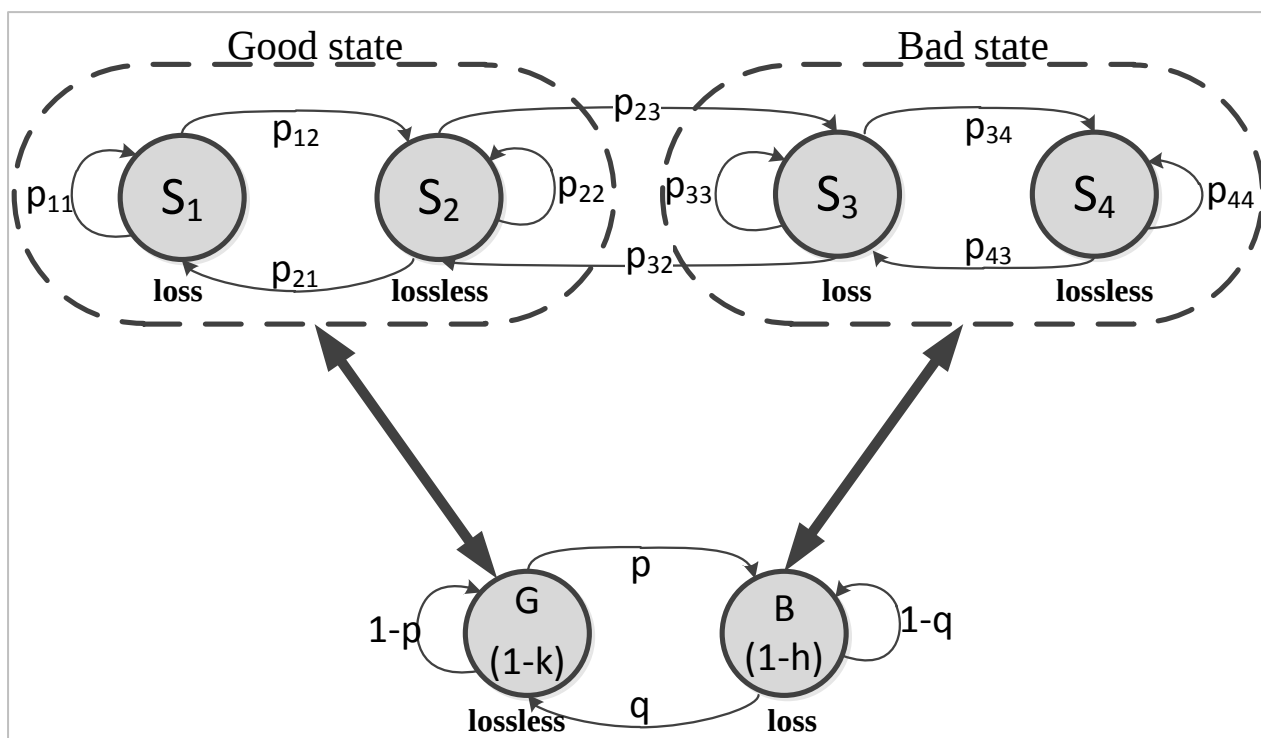
$$S_G = \frac{p_{32}}{p_{23} + p_{32}} \quad (3.31)$$

Verovatnoća gubitaka paketa u lošem stanju Gilbert-Eliotovog modela P_{burst} izražena preko tranzicionih verovatnoća modela sa 4 stanja se izračunava prema sledećem izrazu:

$$P_{burst} = 1 - h = \frac{p_{43}}{p_{43} + p_{34}} \quad (3.32)$$

dok se izolovani gubici određuju prema:

$$P_{isol} = 1 - k = \frac{p_{21}}{p_{21} + p_{12}} \quad (3.33)$$



Slika 3.9 Modelovanje gubitaka posredstvom Markovljevog modela sa 4 stanja.

3.5. Uticaj QoS parametara na kvalitet VoIP servisa

Popularnost korišćenja aplikacija koje zahtevaju prenos glasa i videa preko Interneta je u poslednjoj deceniji dovela je do redefinisivanja servisnih zahteva korisnika kao i do redizajniranja mreža preko kojih se isti prenose. Posledično, IP mreže koje su do tada imale malu količinu saobraćaja sa QoS zahtevima sada moraju da prilagode svoje mogućnosti i da ponude servise kod kojih treba da budu garantovane

performanse (kašnjenje, gubitak paketa, *bandwidth*, džiter) jer je obim takvog saobraćaja sada višestruko veći. U Tabeli 3.3 **Tabela 3.3** su sadržani tipični zahtevi po mrežnim resursima za različite tipove servisa.

Tabela 3.3 Tipične vrednosti *QoS* parametara za različite servise

	<i>Bandwidth</i>	Kašnjenje	Džiter	Gubici paketa
<i>Voice</i>	Niske	Niske	Niske	Niske
Video – interaktivan (two-way)	Visoke	Niske	Niske	Niske
Video streaming (one way)	Visoke	Visoke	Visoke	Niske
Video signalizacija	Niske	Niske	Srednje	Srednje
Voice signalizacija	Niske	Niske	Srednje	Srednje
Interaktivni data servisi - kritični	Promenljive, obično srednje	Srednje	Srednje	Srednje
Data servisi - kritični	Promenljive, obično visoke	Visoke	Visoke	Srednje
Interaktivni data servisi	Promenljive, obično srednje	Visoke	Visoke	Srednje
Data servisi	Promenljive, obično visoke	Visoke	Visoke	Visoke

Sa stanovišta vremenski kritičnih servisa, najrasprostranjeniji servis predstavlja VoIP koji podrazumeva prenos paketizovanog govora preko IP mreže [48]. Kao servis kod koga je od izuzetnog značaja ograničenje kašnjenja, džitera i gubitaka paketa, potrebno je realizovati prenos VoIP paketa preko mreže putanjama koje neće pokvariti performanse tako da se E2E servis može uspešno uspostaviti i realizovati.

Osim VoIP servisa, video servis je takođe dobio na popularnosti sa razvojem procesorskih mogućnosti mobilnih uređaja. Video prenos zahteva višestruko bolje performanse u poređenju sa VoIP servisom i to u zavisnosti od kvaliteta video informacije koji se šalje i primenjenog kodeka koji vrši kodovanje video informacije. Termin „bolje performanse“ u ovom smislu podrazumeva manje kašnjenje, džiter i gubitke paketa kao i da se moraju pružiti jače garancije jer je video po svojoj prirodi mnogo osetljiviji na pad performansi zbog svoje multidimenzionalnosti i potrebe za sinhronizacijom zvuka i slike.

3.5.1. Paketizacija govornog signala

Da bi se govorni signal generisan od strane govornika preneo od izvorišta do odredišta, potrebno je obaviti čitav niz operacija kako bi se došlo do njegove paketizacije, prenosa i konačno, prijema. Konverzija promene zvučnog pritiska govornika obavlja se u mikrofону na čijem izlazu nastaje analogni signal, kontinualan u vremenskom domenu koji dalje mora da se digitalizuje. Digitalizacija se dalje obavlja odabiranjem sa frekvencijom od 8kHz koja predstavlja dva puta veću frekvenciju od maksimalne frekvencije u spektru govornog signala u skladu sa teoremom o odabiranju [49]. Nakon odabiranja, amplitude odbiraka se kvantizuju prema nekom od predefinisanih pravila, μ -law ili A-law, nakon čega se svaka od tako dobijenih kvantizovanih vrednosti koduje sa 8 bita. Ovako kodovani signal se još naziva i

PCM (*Pulse Code Modulation*) signal. Kodek G.711 [50] koristi PCM kodnu modulaciju i to u dve verzije u zavisnosti od primenjenog pravila o kvantizaciji, μ -law ili A-law. Uredni pregled aktuelnih VoIP kodeka je dat u Tabeli 3.4.

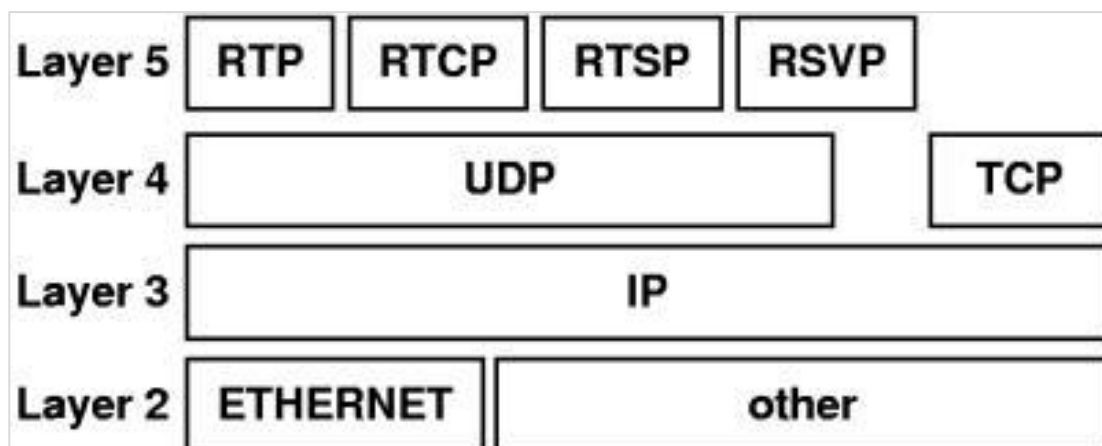
Tabela 3.4 Komparacija različitih kodeka sa pratećim parametrima

Tip kodeka i protok (Kbps)	Veličina odbiraka (Bajtovi)	Interval odabiranja (ms)	Mean Opinion Score (MOS)	Veličina voice paketa (Bajtovi)	Trajanje voice paketa (ms)
G.711 (64 Kbps)	80	10	4.1	160	20
G.729 (8 Kbps)	10	10	3.92	20	20
G.723.1 (6.3 Kbps)	24	30	3.9	24	30
G.723.1 (5.3 Kbps)	20	30	3.8	20	30
G.726 (32 Kbps)	20	5	3.85	80	20
G.726 (24 Kbps)	15	5		60	20
G.728 (16 Kbps)	10	5	3.61	60	30

Svaki od navedenih kodeka u tabeli se razlikuju po veličini *payload*-a kao i protoku. Najkvalitetniji, ali i nazahtevniji jeste upravo G.711 kodek sa veličinom *payload* dela od 160B (kada je trajanje *voice* dela paketa 20ms) i protokom od 64kb/s, dok kvalitet, opisan MOS (*Mean Opinion Score*) parametrom koji standardno služi za ocenu kvaliteta telefonskog signala i iznosi 4.1.

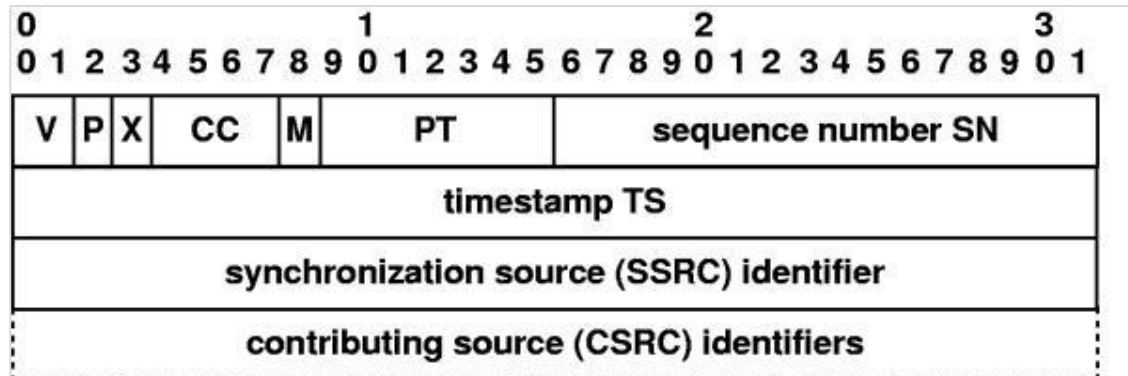
Standardno se paketizovani govor prenosi preko UDP protokola kao transportnog protokola (kao i ostali vremenski kritični servisi). Razlog tome je što UDP ne zahteva prethodnu uspostavu sesije kao što je to slučaj kod TCP protokola. Samim tim, kao posledica upotrebe UDP protokola, ne postoji retransmisija paketa koji nisu stigli do odredišta kao posledica tajm-auta ili gubitka u prenosu tako da se na višim slojevima ovaj gubitak mora rešiti. Upotreba TCP za rešavanje gubitaka paketa vremenski kritičnih servisa ne bi bila rešenje jer, kao što je već naglašeno, kašnjenje koje bi se izazvalo retransmisijom bi onemogućilo funkcionisanje vremenski kritičnih servisa.

Ključni element za prenos paketizovanog govora i zadovoljenje njegovih zahteva po pitanju kašnjenja i varijacije kašnjenja jeste upravo RTP (*Real-time Transport Protocol*) protokol [51]. Pozicija RTP protokola u protokolskom steku je prikazana na slici 3.10. RTP je protokol koji specificira na koji način aplikacija upravlja slanjem multimedijalnih podataka preko *unicast* i *multicast* mrežnih servisa. RTP sam po sebi ne garantuje dostavljanje paketa u realnom vremenu jer je ono zavisno od mrežnih karakteristika i stanja u mreži preko koje se prenose multimedijalni podaci.



Slika 3.10 Pozicija RTP protokola u protokolskom steku.

RTP kombinuje prenos podataka sa kontrolnim protokolom RTCP (*RTP Control Protocol*) koji omogućava upravljanje prijemom podataka. Ovo upravljanje omogućava prijemniku da detektuje potencijalni gubitak paketa i da izvrši kompenzaciju kašnjenja ili džitera. Oba protokola, RTP i RTCP, rade nezavisno od protokola na transportnom i mrežnom sloju. Informacije koje se nalaze u RTP zaglavlju govore prijemniku kako da izvrši rekonstrukciju podataka i da opiše kako su paketizovani strimovi bita. Iako se koristi samo UDP kao transportni protokol, što se tiče RTP, ne postoji neko ograničenje po pitanju drugog transportnog protokola. Izgled RTP zaglavlja je prikazan na slici 3.11.

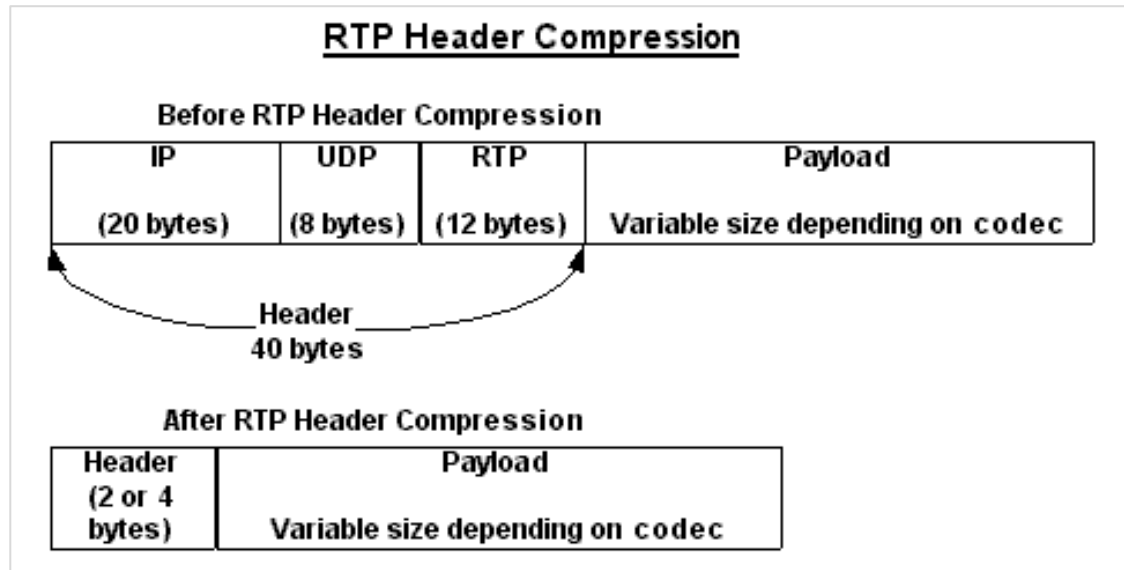


Slika 3.11 Izgled RTP zaglavlja.

RTP komponente podrazumevaju *sequence number* (SN) koji služi za detekciju izgubljenih paketa, *payload type* (PT) koji identifikuje konkretno kodovanje koje je primenjeno na multimedijalne informacije koje se primaju, *timestamp* (TS) koji detektuje različit džiter u okviru pojedinačnog strima i kompenzuje ga.

Dužina pojedinačnih zaglavlja i polja u govornom paketu zavisi od primenjenog kodeka i u opštem slučaju može biti promenljiva (Tabela 3.4). Fiksni deo veličine paketa bez obzira od primenjenog kodeka čine RTP, UDP i IP zaglavlje, i to, 12B, 8B i 20B, respektivno. S obzirom da je od paketa govora potrebno da budu što kraći i da zaglavlje bude što manje kako bi se postiglo manje kašnjenje paketa u mreži, pa samim

tim i u ruterima, realizuje se kompresija zaglavlja posredstvom CRTP (*Compressed RTP*) protokola [52]. CRTP je razvijen kako bi se smanjila veličina IP, UDP i RTP zaglavlja u govornim paketima. Međutim, on je razvijen da radi sa pouzdanim i brzim *point-to-point* linkovima. U nešto manje nego optimalnim uslovima, gde se mogu ostvariti nešto veća kašnjenja, CRTP nema tako dobre karakteristike. Kompresija zaglavlja posredstvom CRTP je prikazana na slici 3.12.



Slika 3.12 Kompresija zaglavlja paketa posredstvom CRTP.

RTCP komponente uključuju takođe i QoS *feedback* koji sadrži broj izgubljenih paketa, RTT (*round-trip time*) i džiter tako da izvorište može da adaptivno prilagodi protok paketa. Dodatno je realizovan i kontrolni mehanizam koji se koristi da se obezbedi indikacija izlaska iz sesije.

3.5.2. Detekcija aktivnosti govornog signala (VAD) i algoritmi za sakrivanje gubitaka paketa (PLC)

VAD (*Voice Activity Detector*) predstavlja mehanizam detekcije odsustva govora kojim se postiže razlikovanje govora od pozadinskog šuma koji se smatra tišinom i kojeg nije potrebno preneti do odredišta. Na taj način se štedi na mrežnom *bandwidth*-u jer se u trenucima tišine ne generišu paketi. Mnoge analize su pokazale da više od 50% govora predstavlja tišinu (pauzu između izgovorenih reči, slova, i sl.) tako da se može postići značajna ušteda ako se primenjuje VAD mehanizam.

Princip rada VAD algoritama se razlikuje u odnosu na elemente i parametre koje koriste da utvrde prisustvo odnosno odsustvo govora. U tom kontekstu, postoje algoritmi koji rade na principu detekcije određenog energetskog praga *voice* signala [53] i [54], [55]. Algoritmi na principu detekcije energetskog praga određuju prisustvo govora na osnovu energije signala u posmatranom vremenskom trenutku, dok spektralno bazirani algoritmi određuju prisustvo govora na osnovu spektralnih karakteristika govornog signala i vrše razlikovanje tog spektra od pozadinskog šuma. Od pozadinskog šuma zavisi koliko će

pojedini VAD tipovi algoritma da budu efikasni. U zavisnosti od prisustva dominantnog pozadinskog šuma, teže je identifikovati periode odsustva govora. Komparacija VAD algoritama koji se baziraju na ova dva pristupa je obavljena u [56].

Robusnost VoIP servisa u velikoj meri zavisi od implementiranih *loss recovery* mehanizama koji se generalno mogu podeliti na *sender-based* i *receiver-based* [57]. Korišćenjem ovih mehanizama se omogućava lakša implementacija VoIP servisa i do određenog nivoa se smanjuju strogi zahtevi po pitanju raspoloživih performansi u mreži. *Sender-based loss recovery* mehanizmi se zasnivaju na pretpostavci da će se viši nivo robusnosti *voice* paketa poslati ako se dodatno šalju redundantni podaci koji će povećati nivo robusnosti. Sa druge strane, *receiver-based* mehanizmi implementiraju algoritme koji su zastupljeni na prijemnoj strani i koji se zasnivaju na zahtevu da se izgubljeni paketi na neki način moraju maskirati. U ovom radu, od interesa za posmatranje ponašanja u prisustvu gubitaka paketa jesu *receiver-based* mehanizmi koji predstavljaju nezamenljiv mehanizam povećanja robusnosti u prisustvu rafalnih gubitaka paketa.

PLC (*Packet Loss Concealment*) [58] predstavlja tehniku koja ima ulogu da minimizira praktične efekte izgubljenih paketa. PLC se pre svega koristi u VoIP komunikaciji. PLC smanjuje primetnost izgubljenog ili izgubljenih paketa. Paketi koji stižu do odredišta nešto kasnije u odnosu na druge pakete mogu da budu postavljeni u bafer gde će ostati određen interval nakon čega se vrši ponovno asembliranje paketa i prosleđivanje na više slojeve. Baferovanjem paketa se povećava kašnjenje, koje je dosta bolja alternativa u odnosu na odbacivanje zakasnelog paketa. Baferovanjem se rešava varijacija kašnjenja na račun povećavanja E2E kašnjenja paketa od izvorišta do odredišta tako što se postiže da se slanje paketa prema višim slojevima obavlja konstantnim protokom bez promenljivog kašnjenja. Baferovanje je moguće primeniti ako se njime ne prouzrokuje povećano kašnjenje koje bi dovelo do uvećane degradacije VoIP servisa. Inače, u sadašnjim VoIP implementacijama upotreba bafera za potrebe ispravljanja varijacije kašnjenja je redovna praksa koja u normalnim uslovima neprimetno utiče na povećanje kašnjenja.

Ozbiljniji problem predstavljaju paketi čije je kašnjenje iznad dozvoljenog ili paketi koji su izgubljeni zbog odbacivanja kao posledice lošeg kvaliteta putanje ili zagušenja u redovima opsluživanja. Jedan metod koji koristi PLC u ovakvim slučajevima jeste veštačka regeneracija paketa koji su prethodno primljeni koji se koriste za popunjavanje izostalih paketa. Drugi metod predstavlja relizaciju interpolacije i veštačke generacije paketa na mestu izgubljenog paketa.

Sakrivanje gubitaka paketa može da bude teško za implementaciju ukoliko je PLR veliki ili ukoliko se radi o rafalnim gubicima manifestovanim kroz uzastopne ili bliske gubitke paketa. U slučajevima rafalnih gubitaka, PLC nije dovoljno efikasan da prikrije te gubitke i upravo je ova problematika jedan od predmeta analize u ovom radu. Jedan od doprinosa ovog rada jeste da se primenom paketske disperzije VoIP paketa

utiče na povećanje efikasnosti PLC algoritama u sakrivanju izgubljenih paketa tako što će se efektivno promeniti struktura rafalnih gubitaka koji se javljaju.

3.6. Metode ocene kvaliteta VoIP servisa

U cilju utvrđivanja kako VoIP aplikacije reaguju na promenu operativnog stanja mreže, tačnije nivoa QoS opisanog merama performansi potrebno je sagledati odgovarajuće modele koji pružaju uvid kako i u kojoj meri pojedini QoS parametri utiču na detektovan kvalitet od strane korisnika. Razmatrajući modele ocene kvaliteta VoIP servisa, razlikuju se dva pristupa: subjektivni i objektivni testovi kvaliteta.

Subjektivne metode određivanja kvaliteta *voice* aplikacija određuju nivo kvaliteta iz ljudske perspektive i kao takvi daju odgovarajuću količinu subjektivnosti u smislu ljudske percepcije. Ove tehnike određivanja kvaliteta podrazumevaju da testovi rezultiraju ocenom koja se naziva *Mean Opinion Score* (MOS) koja rangira kvalitet u opsegu od 1 do 5. ITU-T predviđa da ocena 1 podrazumeva veoma loš, dok ocena 5 podrazumeva odličan kvalitet. ITU-T preporuke P.920 [59] i P.800 [60] definišu način i daju smernice oko samog procesa obavljanja procesa subjektivnog određivanja kvaliteta. Ove smernice obuhvataju specifikaciju prostornog i fizičkog okruženja u okviru koga se merenja realizuju, procedure ocenjivanja, detalji oko starosnih ograničenja subjekata koji su podvrgnuti testovima, i sl. Subjektivni pristupi u merenjima kvaliteta *voice* i ostalih multimedijalnih servisa podrazumevaju uključivanje većeg broja subjekata u cilju postizanja veće tačnosti merenja kvaliteta.

Iako subjektivni pristup merenju kvaliteta predstavlja pouzdan način određivanja uticaja QoS parametara, ovi testovi su prilično teški za realizaciju s obzirom da zahtevaju značajne resurse za izvođenje. Naime, da bi merenja bila tačna, veći broj subjekata mora biti uključen, dok procedure samog testiranja moraju biti prema važećim standardima iz relevantne oblasti. Posledično, ovi testovi nisu adekvatni sa aspekta skalabilnosti i za primenu kod *real time* servisa, pa se zbog toga značajni trud ulaže u razvijanje objektivnih pristupa određivanju kvaliteta koji su skalabilniji i mogu se lako prilagoditi *real time* servisnim zahtevima.

Imajući u vidu nedostatke subjektivnih pristupa određivanju kvaliteta, objektivne metode podrazumevaju komparaciju primljenog signala u odnosu na originalni. Dodatno, ove metode uglavnom određuju kvalitet od strane slušaoca, dok se ne konverzacijom kvalitetom. Kvalitet na strani slušaoca uzima u obzir samo kvalitet primljenog signala, dok konverzacijom obuhvata ukupan kvalitet servisa i posmatra interaktivni aspekt koji je od izuzetnog značaja u *real time* dvosmernim servisima.

Većina modela za objektivno određivanje kvaliteta posmatra kvalitet na strani slušaoca pri čemu su neki od dominantnih modela *Perceptual Speech Quality Measure* (PSQM) [60], E-Model [62] i PESQ [63]. U

ovom radu se kao metoda određivanja kvaliteta VoIP aplikacija koristi ITU-T E-Model kako bi se kvantitativno iskazao dobitak u kvalitetu dobijen primenom paketske disperzije.

3.6.1. Primena ITU-T E-Modela

ITU-T E-Model predstavlja jedan od najrasprostranjenijih metoda za određivanje objektivnog kvaliteta voice aplikacija. Kao rezultat se određuje *transmission rating factor* R koji obuhvata kombinovane efekte parametara koji su bitni u procesu prenošenja voice signala. Određuje se prema sledećem izrazu:

$$R = R_0 - I_s - I_d - I_{e-eff} + A \quad (3.34)$$

gde R_0 označava osnovni odnos signal-šum, factor I_s predstavlja kombinaciju svih parametara koji manje ili više utiču istovremeno na voice signal, I_d određuje meru pogoršanja voice kvaliteta koja potiče od kašnjenja, dok I_{e-eff} se vezuje za pogoršanja kvaliteta uneta od strane kodeka sa malim protokom i A koji se naziva *advantage factor* i koji uzima u obzir tip pristupa koji se realizuje kod voice aplikacije.

Tabela 3.5 Trenutno važeće vrednosti *advantage factor* A parametra

Način pristupa servisu	Maksimalna vrednost faktora A
Konvencionalni pristup (žični)	0
Mobilni servis posredstvom ćelijske mreže u okviru zgrade	5
Mobilnost u okviru određene geografske oblasti ili kretanje vozilom	10
Pristup teško pristupačnim lokacijama (npr. posredstvom satelitskih veza)	20

Iako u skladu sa ITU-T specifikacijom E-Modela za svaki od prethodno navedenih parametara postoji analitički način da se odredi koji zavisi od različitih faktora, postoji i aproksimacija modela koja se često koristi kako bi se ostvario uvid u VoIP kvalitet u odnosu na raspoložive mrežne performanse opisane kašnjenjem i gubicima paketa. Jedna od čestih aproksimacija u ovom kontekstu je analizirana u [64] prema kojoj se određuju granične vrednosti mrežnih parametara od kojih dati parametri kreću u velikoj meri da utiču na voice kvalitet. U skladu sa tipičnim vrednostima određenim u skladu sa [62] može se posmatrati naredni izraz za određivanje transmisionog faktora R :

$$R = 93.2 - I_d - I_{e-eff} \quad (3.35)$$

iz koga se izuzima *advantage factor* s obzirom da razmatramo konvencionalni pristup VoIP servisu. U okviru Tabele 3.5 su prikazane važeće vrednosti u skladu sa ITU-T specifikacijom, dok su u ITU-T preporuci G.113 [65] predstavljeni relevantni aspekti prilikom određivanja vrednosti ovog parametra.

Kao nastavak analitički prikazane aproksimacije u izrazu (3.35) dalje se može aproksimirati uticaj kašnjenja na kvalitet. Naime, posmatra se kašnjenje paketa u jednom smeru d i sa slike 3.13 se može

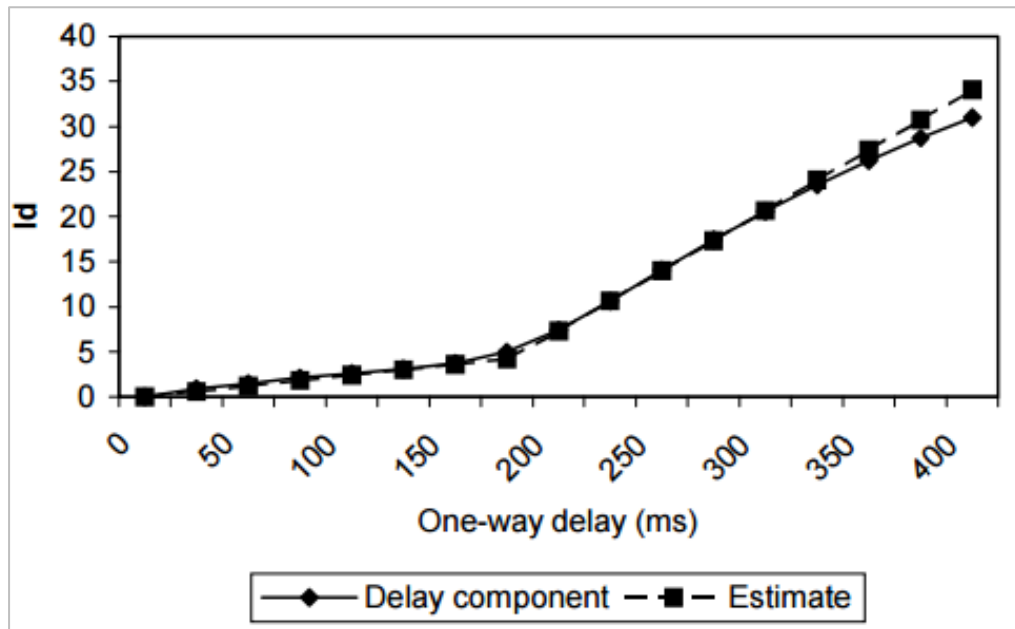
primetiti ograničen uticaj parametra d na meru pogoršanja kvaliteta posredstvom uticaja kašnjenja opisanom parametrom I_d [64].

Sa slike 3.13 se uočava promena parametra I_d za vrednost $d=177\text{ ms}$ na osnovu čega se može zaključiti da vrednosti jednosmernog kašnjenja manjih od ove vrednosti nemaju značajnog uticaja na voice kvalitet. Prethodno predstavljena zavisnost se može i analitički opisati:

$$I_d = 0.024 \cdot d + 0.11(d - 177.3) \cdot H(d - 177.3) \quad (3.36)$$

pri čemu je H predstavlja *Heavyside*-ovu funkciju za koju važi sledeće:

$$H(x) = \begin{cases} 0, & \text{za } x < 0 \\ 1, & \text{za } x \geq 0 \end{cases} \quad (3.37)$$



Slika 3.13 Zavisnost vrednost parametra I_d od jednosmernog kašnjenja [64].

Bitno je naglasiti da E-Model posmatra uticaj parametra na kumulativan način tako što su pojedinačni uticaji mrežnih i parametara okruženja opisani uvođenjem odgovarajućih parametara u neaproksimiranom izrazu (3.34). Naravno, iako se uticaji razmatraju kao odvojeni, mrežni parametri kao što su kašnjenje i gubici paketa su u korelaciji i generalno zavise od mrežnih uslova i transmisijnog medijuma. Upravo je ovaj način analize uticaja pojedinačnih parametara najveća mana E-Modela [66]. Međutim, za potrebe ovog rada, tačnost određivanja kvaliteta prevazilazi okvire istog i smatra se da je E-Model dovoljno tačan da pruži uvid u VoIP kvalitet kada se razmatra uticaj mrežnih parametara, pre svega gubitaka paketa.

3.6.2. Uticaj gubitaka paketa u E-Modelu

Uticaj gubitaka paketa prema E-Modelu je obuhvaćen parametrom I_{e-eff} u skladu sa kojim se razmatra pogoršanja koje unosi sama oprema, od gubitaka paketa i njihove strukture pojave, pa do tipa korišćenog VoIP kodeka. Parametar I_{e-eff} se može izraziti preko sledećeg izraza:

$$I_{e-eff} = I_e + (95 - I_e) \cdot \frac{P_{pl}}{\frac{P_{pl}}{BurstR} + B_{pl}} \quad (3.38)$$

gde je P_{pl} verovatnoća gubitaka paketa, B_{pl} faktor robusnosti, $BurstR$ je parametar koji opisuje rafalnost gubitaka paketa (*burst ratio*), dok je I_e faktor pogoršanja koji zavisi od primenjenog tipa kodeka.

Vrednosti faktora pogoršanja I_e su definisane u [65] i prikazane su u Tabeli 3.6. Navedene vrednosti su određene bez prisustva gubitaka paketa ili grešaka u prenosu.

Prethodne vrednosti su pridružene odgovarajućem tipu kodeka koji se koristi i važe u slučajevima kada nema grešaka niti gubitaka paketa u prenosu. U slučaju pojave gubitaka paketa, vrednosti parametara I_e i B_{pl} su prikazane u Tabeli 3.7 i važe u zavisnosti od veličine paketa, postojanja PLC šema i samog tipa kodeka.

Tabela 3.6 Trenutno važeće vrednosti parametra I_e bez grešaka i gubitaka paketa [65]

Tip kodeka	Specifikacija kodeka	Protok	I_e
PCM	G.711	64	0
ADPCM	G.726, G.727	40	2
	G.721, G.726, G.727	32	7
	G.726, G.727	24	25
	G.726, G.727	16	50
LD-CELP	G.728	16	7
		12.8	20
CS-ACELP	G.729	8	10
	G.729-A+VAD	8	11
VSELP	IS-54	8	20
ACELP	IS-641	7.4	10
QCELP	IS-96a	8	21
RCELP	IS-127	8	6
VSELP	Japanski PDC	6.7	24
RPE-LTP	GSM 06.10, <i>full-rate</i>	13	20
VSELP	GSM 06.20, <i>half-rate</i>	5.6	23
ACELP	GSM 06.60, <i>enhanced full rate</i>	12.2	5
ACELP	G.723.1	5.3	19
MP-MLQ	G.723.1	6.3	15

U okviru izraza (3.38) obuhvaćeni su i slučajne pojave gubitaka paketa, koje generalno ne zavise od prethodnih događaja gubitaka, kao i rafalni gubici koji imaju memorijski efekat kada verovatnoća narednih gubitaka paketa zavisi od prethodnih događaja gubitaka. Kao što je i ranije analizirano, nezavisna pojava gubitaka paketa u realnim mrežama je izuzetno retka, naročito kod VoIP servisa. Dodatno, svaki tip kodeka poseduje određeni nivo robusnosti, pa je neretka pojava da pojedine specifične implementacije kodeka budu robusnije nego što je to navedeno za kodeke u Tabeli 3.7.

Pojava rafalnih gubitaka se uzima u obzir analizom *BurstR* parametra koji dovodi u vezu nivo rafalnosti gubitaka paketa sa ukupnim VoIP kvalitetom. *BurstR* parametar je jednak količniku verovatnoće pojave gubitka paketa i tranzicione verovatnoće prelaska Markovljevog modela iz dobrog u loše stanje. Kod Markovljevog modela sa 2 stanja, *BurstR* parametar je jednak:

$$BurstR = \frac{P_{loss}}{p} = \frac{1}{p+q} \quad (3.39)$$

u skladu sa modelom prikazanim na slici 3.2. *BurstR* parametar kod Gilbert-Eliotovog modela je jednak:

$$BurstR = \frac{P_{loss}}{p} = \frac{1}{p+q} \cdot (1-h) + \frac{q}{p \cdot (p+q)} \cdot (1-k) \quad (3.40)$$

Određivanje *BurstR* parametara kod Markovljevog modela sa 4 stanja je nešto komplikovanije u odnosu na prethodna dva modela. Naime, težina u određivanju ovog parametra se ogleda u problemu određivanja tranzicione verovatnoće p_{g-b} koja se izračunava na osnovu sledećeg izraza:

$$p_{g-b} = \frac{P_{g-b}}{1-P_{loss}} \quad (3.41)$$

gde je P_{g-b} verovatnoća gubitaka koji je nastao kao posledica usled tranzicije iz stanja u kojima nema gubitaka, S_2 i S_4 u stanja u kojima se dešavaju gubici paketa, S_1 i S_3 . Verovatnoća P_{g-b} se određuje tako što se najpre definišu matrica stanja S_{good} koja sadrži samo dobra stanja i tranziciona matrica koja sadrži isključivo tranzicione verovatnoće koje podrazumevaju prelaze između prethodno navedenih stanja T_{g-b} u skladu sa sledećim izrazima:

$$S_{good} = [0 \quad S_2 \quad 0 \quad S_4] \quad (3.42)$$

$$T_{g-b} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ p_{21} & 0 & p_{23} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & p_{43} & 0 \end{bmatrix} \quad (3.43)$$

Tabela 3.7 Trenutno važeće vrednosti parametara I_e i B_{pl} u slučaju grešaka i gubitaka paketa [65]

Tip kodeka	Veličina paketa	PLC	I_e	B_{pl}
G.723.1+VAD	30 ms	Postojeći	15	16.1
G.729A+VAD	20 ms (u 2 frejma)	Postojeći	11	19.0
GSM-EFR	20 ms	Postojeći	5	10.0
G.711	10 ms	-	0	4.3
G.711	10 ms	ITU-T G.711 specifikacija	0	25.1

Nakon određivanja matrice stanja S_{good} i tranzicione matrice T_{g-b} verovatnoća gubitaka paketa jednaka je matričnom proizvodu prethodno pomenutih matrica prema sledećem izrazu:

$$P_{g-b} = S_{good} \times T_{g-b} \times I \quad (3.44)$$

gde je I matrica dimenzija 4x1 koja se sastoji od jediničnih elemenata kako bi se dobila odgovarajuća vrednost tražene verovatnoće. Uključivanjem ove vrednosti u izraz (3.41) dobija se tranziciona verovatnoća, a $BurstR$ se zatim određuje kao količnik ukupne verovatnoće gubitaka paketa P_{loss} i tranzicione verovatnoće p_{b-g} .

3.6.3. MOS ocena kvaliteta

Izlaz E-Modela predstavlja transmisioni faktor R čije su vrednosti u opsegu 1-100 i koji se često izražava MOS ocenu čije vrednosti rangiraju kvalitet voice signala na prijemnoj strani u opsegu 1-5. U okviru samog E-Modela nije navedena specifikacija kako se ove dve ocene mapiraju jedna u drugu. Relacija između ovih ocena se obavlja posredstvom jednakosti [67]:

$$MOS = \begin{cases} 1, & R < 0 \\ 1 + 0.35 \cdot R + R \cdot (R - 60) \cdot (100 - R) \cdot 7 \cdot 10^{-6}, & 0 \leq R < 100 \\ 4.5, & R = 100 \end{cases} \quad (3.45)$$

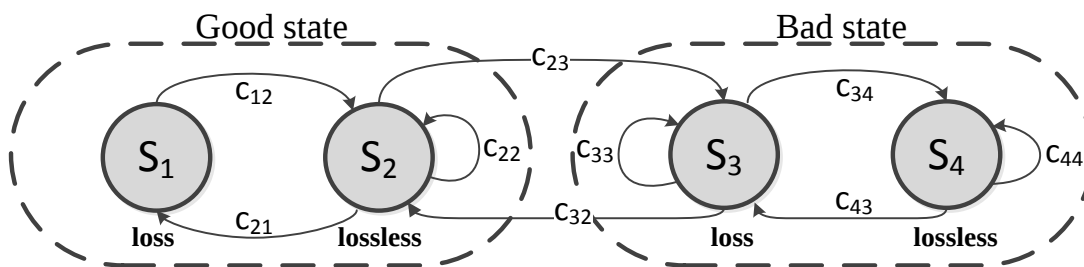
Tabela 3.8 sadrži mapiranje MOS ocene i faktora R kao i nivo percepiranog kvaliteta od strane korisnika.

Tabela 3.8 Mapiranja R faktora i MOS ocena

R	Nivo percepcije	MOS
90-100	Veoma zadovoljni	4.3+
80-90	Zadovoljni	4.0 – 4.3
70-80	Neki korisnici nisu zadovoljni	3.6 – 4.0
60-70	Mnogi korisnici nisu zadovoljni	3.1 – 3.6
50-60	Skoro svi korisnici nisu zadovoljni	2.6 – 3.1
0-50	Nije preporučljivo	1.0 – 2.6

3.6.4. Estimacija rafalnih gubitaka iz dolaznog niza paketa

Za niz voice paketa koji pristiže na određite mora da postoji metodologija određivanja karakteristike izraženih gubitaka paketa. Ovo je naročito bitno u slučaju pojave rafalnih gubitaka kada je potrebno detaljnije odrediti nivo rafalnosti opisan prethodno definisanim parametrima kako bi se stekao tačan uvid u uticaj koji rafalni gubici mogu da imaju na VoIP kvalitet. Relevantni parametri u slučaju pojave rafalnih gubitaka jesu verovatnoće gubitaka u okviru rafala (gustina gubitaka - P_{burst}), rafalni period (dužina trajanje rafala - B), dužina trajanja dobrog stanja i verovatnoća izolovanih gubitaka P_{isol} .



Slika 3.14 Model gubitaka sa 4 stanja.

Metodologija estimacija karakteristike rafalnih gubitaka iz niza paketa predložena je u [44] i podrazumeva da se posmatra Markovljev model sa 4 stanja prethodno analiziran sa malim promenama kakve su prikazane na slici 3.14. Metodologija podrazumeva umesto tranzicionih verovatnoća uvođenje brojača prelazaka iz jednog stanja modela u drugo. U skladu sa tim, stanja su definisana prema Tabeli 3.9.

Prema metodologiji estimacije strukture gubitaka svaki paket podrazumeva inkremenaciju odgovarajućeg brojača u zavisnosti od trenutnog stanja modela. To podrazumeva da ako je sistem u stanju S_2 da prelazak sistema u stanje S_3 rezultuje inkrementacijom brojača c_{23} . Isto važi i za ostale tranzicije iz drugih stanja.

Celokupna metodologija se može opisati prema sledećem kodu:

```
Za svaki paket:
if (packet_lost) {
    loss_count++;
}
if (packet_discarded) {
    discard_count++;
}
if (!packet_lost && !packet_discarded) {pkt++;}
else {
    if (pkt >= gmin) {
        if (lost == 1) c21++;
        else c23++;
        lost=1;
        c22+=pkt;
    }
}
else {
    lost++;
}
```

```

    if (pkt == 0) c33++;
    else {
        c43++;
        c44+=(pkt-1);
    }
}
pkt=0;
}

```

Tabela 3.9 Stanja modela gubitaka paketa

S_1	Gubitak u dobrom stanju
S_2	Bez gubitka u dobrom stanju
S_3	Gubitak u lošem stanju
S_4	Bez gubitka u lošem stanju

Promenljiva pkt predstavlja broj paketa koji su primljeni od kada je poslednji paket izgubljen ili odbačen, dok je $lost$ broj izgubljenih paketa unutar trenutnog rafala. Bulove promenljive $packet_discarded$ i $packet_lost$ označavaju promenljive koje se vezuju za pakete koji su odbačeni (kao posledica kasnijeg dolaska ili zbog postojanja greške u prenosu paketa) ili izgubljeni (uopšte se nisu pojavili na prijemu), respektivno. Na osnovu ovih promenljivih i brojača koji su nakon dovoljnog broja paketa inkrementirani se vrši određivanje relevantnih parametara koji jasno opisuju strukturu rafalnih gubitaka.

Za brojače važe sledeće pretpostavke:

$$c_{32} = c_{23} \quad (3.46)$$

$$c_{43} = c_{34} \quad (3.47)$$

Ukupni broj paketa tokom jednog ciklusa posmatranja rafalnih gubitaka c_{total} :

$$c_{total} = c_{22} + c_{21} + c_{44} + c_{43} + c_{32} + c_{34} + c_{33}. \quad (3.48)$$

Na osnovu ovih brojača se dalje određuju verovatnoće gubitaka u dobrom i lošem stanju kao i rafalni period. Gustina gubitaka $P_{density}$ je jednaka:

$$P_{density} = \frac{c_{23} + c_{33} + c_{43}}{c_{32} + c_{33} + c_{34} + c_{44}} \quad (3.49)$$

Verovatnoća izolovanih gubitaka paketa P_{gap} se može odrediti u skladu sa sledećim izrazom:

$$P_{gap} = \frac{c_{21}}{c_{22} + c_{21}} \quad (3.50)$$

dok su periodi sistema provedenih u dobrom i lošem stanju određenih posredstvom respektivnih relacija:

$$B_{gap} = \frac{c_{22}+c_{21}}{c_{23}+1} \quad (3.51)$$

$$B_{burst} = \frac{c_{21}+c_{33}+c_{34}+c_{44}+c_{43}}{c_{23}} \quad (3.52)$$

Potrebno je naglasiti i činjenicu da se model gubitaka paketa koji koristi ova metodologija i koji je prikazan na slici 3.14 razlikuje od Markovljevog modela sa 4 stanja s obzirom da predstavljena metodologija ne prepoznaje uzastopne gubitke u dobrom stanju modela. Dodatno, početne pretpostavke uvedene i opisane izrazima (3.46) i (3.47) doprinose da će se gustina gubitaka paketa i verovatnoća izolovanih paketa dobijenih ovom metodologijom razlikovati u odnosu na analizu obavljenju u okviru Markovljevog modela sa 4 stanja. Osim toga, periodi provedeni u dobrom i lošem stanju će se razlikovati od vrednosti za rafalni period i period trajanja sistema u dobrom stanju dobijen analitičkim pristupom Markovljevog modela sa 4 stanja sa slike 3.8.

Međutim, i pored predočene razlike u pristupu analizi strukture gubitaka paketa i modelovanju gubitaka paketa, opisana metodologija predstavlja nezamenljivi deo u analizi strukture gubitaka VoIP paketa. Na osnovu primene pomenute metodologije se mogu izvesti značajni zaključci o uticajima različite strukture gubitaka. Takođe, primena ove metodologije može da pruži komparativni značaj za predmet ovog rada kada se analiziraju uticaji različitih strategija paketskih disperzija.

3.7. Analiza rastojanja gubitaka u VoIP aplikacijama

Prethodno definisani parametri rafalnih gubitaka paketa se mogu uspešno primeniti na analizu uticaja gubitaka na multimedijalne aplikacije. Zbog postojećih PLC algoritama za sakrivanje gubitaka, neophodno je uvesti parametar koji će kvantifikovati rastojanje uzastopnih gubitaka paketa. Imajući u vidu da PLC algoritmi i njihova efikasnost u skrivanju gubitaka paketa zavise od rasporeda uzastopnih gubitaka, jasno je da je postojanje odgovarajućeg parametra koji uzima u obzir rastojanje gubitaka paketa primenljivo na multimedijalne aplikacije sa implementiranim PLC algoritmima.

Sastavni deo analize strukture gubitaka paketa podrazumeva razmatranje rastojanja međusobnih uzastopnih gubitaka paketa kako bi se dobio precizniji uvid kako struktura gubitaka paketa može da utiče na multimedijalne aplikacije. U [68] se predlaže razmatranje strukture gubitaka paketa kroz posmatranje rastojanja između gubitaka tako što se definišu dodatni parametri kojima se prethodno pomenuta karakteristika može opisati. Uvođenje ovih dodatnih parametara je posledica rafalnih gubitaka paketa koji su dominantni oblik strukture gubitaka na Internetu i kod koga je kritično poznavati rastojanje paketa kako bi se opisao nivo rafalnosti. Pojava bliskih gubitaka paketa može da ima izrazito negativan uticaj na

performanse pojedinih aplikacija, a naročito na *real-time* multimedijalne aplikacije čiji kvalitet biva degradiran iz perspektive krajnjih korisnika.

U cilju uvođenja mere rastojanje gubitaka paketa, potrebno je prethodno uvesti odgovarajuće definicije termina koji su kritični za posmatranje niza paketa u okviru kojih su izraženi rafalni gubici.

- **Sekvenciranje** - kako bi se utvrdila tačna pozicija primljenih i izgubljenih paketa u nizu, svakom paketu se dodeljuje odgovarajući broj. Postojanje ovih brojeva može da bude deo protokolske specifikacije (na primer, u okviru RTP paketa u odgovarajućem polju) ili ukoliko se posmatraju test paketi kojima se mere performanse način sekvenciranja treba da bude unapred određen.
- **Rafalni gubici** - po definiciji iz [68] pod rafalnim gubicima se podrazumeva niz uzastopnih gubitaka paketa. Ova definicija je u koliziji sa definicijom rafalnih gubitaka iz [44] u kojoj se specificira minimalno rastojanje između paketa na osnovu kojih se utvrđuje postojanje rafalnih gubitaka. Za potrebe analize strukture gubitaka paketa u kontekstu paketske disperzije biće primenjena definicija rafalnih gubitaka iz [44].
- **Rastojanje između gubitaka** – formalna definicija ovog rastojanja podrazumeva razliku sekvencionih brojeva dva uzastopna gubitka koja mogu biti razdvojeni uspešno primljenim paketima.
- **Period gubitaka** – podrazumeva niz izgubljenih paketa nakon uspešno primljenog paketa.

Uzimajući u obzir prethodne definicije, primenom sličnog pristupa na dolazni niz paketa je moguće korišćenjem jednog parametra analizirati rastojanje između uzastopnih paketa. S obzirom da rafalni gubici sa manjim rastojanjem između gubitaka imaju veći uticaj na performanse pojedinih aplikacija, potrebno je identifikovati kritično rastojanje koje PLC algoritmi implementirani u okviru aplikacija nisu sposobni da eliminišu. Upravo ovi gubici rezultuju opadanjem nivoa kvaliteta koji je percipiran od strane korisnika i prigodno se mogu nazvati „primetnim“ gubicima. U skladu sa ovim pristupom, sada se može definisati parametar koji će kvantitativno opisati pojavu i intenzitet primetnih gubitaka paketa.

Svi gubici paketa čije je rastojanje manje od nekog predefinisano rastojanja se mogu smatrati primetnim gubicima, pa i ime parametra koji opisuje iste je deskriptivno dato kao *Noticeable Loss Rate* (NLR) [44]. Bliskost gubitaka paketa je određena uvođenjem parametra koji definiše granicu bliskosti gubitaka ispod koje su gubici primetni. Ovaj parametar se naziva granica primetnih gubitaka δ . Parametar δ se suštinski određuje iz efikasnosti implementiranih PLC algoritama. Ukoliko su PLC šeme sposobne da sakriju bliske gubitke, parametar δ će biti manji. Na primer, ako implementiran PLC algoritam može najviše da sakrije dva gubitka između kojih se prime dva paketa, onda je $\delta=2$. U tom smislu, određivanje adekvatnog parametra δ zavisi od efikasnosti PLC šeme, ali posmatranjem šireg opsega vrednosti ovog parametra se

može obezbediti uvid u bliskost gubitaka paketa čak i ako se ne poznaje efikasnost trenutnih PLC algoritama.

U kontekstu bliskih gubitaka paketa NLR parametar uzima u obzir sve gubitke paketa koji se pojavljuju na rastojanju manjem od δ . Striktna definicija NLR parametra podrazumeva količnik gubitaka čije je rastojanje manje od δ i ukupnog broja izgubljenih paketa. U skladu sa [68] moguće je i alternativno posmatrati količnik primetnih gubitaka paketa i uspešno primljenih paketa. U svakom slučaju, NLR predstavlja verovatnoću pojave primetnih gubitaka paketa koji zavise od osetljivosti primenjene PLC šeme koja kvantitativno uzima u obzir rastojanje gubitaka posredstvom parametra δ . U zavisnosti od konkretne analitičke definicije NLR parametra, on može da posmatra koji deo ukupnih gubitaka paketa je primetan ili alternativno može da označava odnos primetnih gubitaka i uspešno primljenih paketa čime se može ostvariti generalno saznanje o potencijalnom kvalitetu multimedijalnih sesija i servisa.

Kako je moguće posmatrati NLR parametar kroz nekoliko definicija, za potrebe analize rafalnih gubitaka opisanim Markovljevim modelima sa dva i četiri stanja, ovde će se koristiti NLR definicija koja posmatra ovaj parametar kao odnos primetnih gubitaka i ukupnog broja poslatih paketa. Korišćenom definicijom se neznatno menja vrednost NLR parametra pri čemu se zadržava osnovno svojstvo deskripcije primetnih gubitaka paketa zbog čega je NLR parametar inicijalno i uveden u analizu bliskosti i strukture gubitaka paketa.

Narednom analizom rastojanja gubitaka paketa kroz posmatranje NLR parametara se ostvaruje uvid u srednju vrednost rastojanja gubitaka u nizu primljenih paketa. Iz definicije NLR mere rastojanja gubitaka je takođe potrebno odrediti i rastojanje δ . Ova analiza podrazumeva da se realizuje konvencionalno rutiranje, tačnije da se niz paketa prenosi preko jedne putanje na kojoj su izraženi rafalni gubici koji se mogu modelovati Gilbertovim, Gilbert-Eliotovim i Markovljevim modelom sa četiri stanja. Analiza ovih modela je već realizovana u ovom radu, a takođe su i predložene prednosti i mane svakog pojedinačnog modela gubitaka paketa.

3.7.1. NLR analiza u prisustvu Gilbertovih i Gilbert-Eliotovih gubitaka

Kod rafalnih gubitaka paketa koji su modelovani Gilbertovim modelom na jednoj putanji između izvorišta i odredišta NLR parametar se može odrediti tako što se niz paketa koji se prima na odredišnoj strani posmatra kao dva različita događaja – primljeni i izgubljeni paket. Niz događaja se opisuje funkcijom događaja $L(n)$ za paket n . Za funkciju $L(n)$ važi da je $L(n) \in \{1, 0\}$ što podrazumeva ako je $L(n)=0$ onda je paket n uspešno primljen, dok ako je $L(n)=1$ onda je paket n izgubljen.

Kao što je već pomenuto, za potrebe analize NLR parametara u kontekstu rafalnih gubitaka biće usvojena definicija po kojoj NLR predstavlja odnos primetnih gubitaka i ukupnog broja paketa u nizu, tačnije

izgubljenih i primljenih. Koristeći ovu definiciju u okviru naredne analize biće potrebno da se odredi analitički izraz čijom primenom će moći da se odrede приметni gubici koji zavise od predefinisanoг ograničenja δ .

Niz od k paketa se analitički može predstaviti kao $(L(n), L(n+1), L(n+2), L(n+3), \dots, L(n+k-1))$. Verovatnoća pojavljivanja prethodno pomenutog niza kao i ishod događaja za pojedinačne pakete u okviru niza će zavistiti od strukture gubitaka koja je izražena na posmatranoj putanji. U ovom slučaju, razmatraju se rafalni gubici koji se modeluju koristeći Gilbertov model koji je analiziran u sekciji 3.4.2. Verovatnoća pojavljivanja prethodnog niza prilikom uticaja Gilbertovih gubitaka se određuje u skladu sa sledećim izrazom:

$$P[L(n), L(n+1), L(n+2), L(n+3), \dots, L(n+k-1)] = S \times \prod_{i=0}^{k-1} T_{trans} \times I \quad (3.53)$$

pri čemu je S matrica stanja koja sadrži verovatnoću dobrog i lošeg stanja i dimenzija je 1×2 , T_{trans} je tranziciona matrica koja identifikuje pakete koji su primljeni ili izgubljeni u skladu sa Gilbertovim modelom, dok je I matrica dimenzije 2×1 i oba elementa su jednaka 1. Tranziciona matrica T_{trans} je jednaka

$$T_{trans} = \begin{cases} T \times \pi_{loss}, & \text{ako } L(n+i) = 1 \\ T \times \pi_{good}, & \text{ako } L(n+i) = 0 \end{cases} \quad (3.54)$$

gde je T tranziciona matrica Gilbertovog modela, dok je π_{loss} i π_{good} predstavljaju verovatnoće da paket bude izgubljen odnosno primljen u trenutnom stanju Gilbertovog modela. Tranziciona matrica T kao i verovatnoće π_{loss} i π_{good} su date u narednim izrazima:

$$T = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix} \quad (3.55)$$

$$\pi_{loss} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad (3.56)$$

$$\pi_{good} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad (3.57)$$

Iz prethodno navedenih relacija, jasno je da važi sledeće:

$$\pi_{loss} = I - \pi_{good} \quad (3.58)$$

gde je I jedinična matrica dimenzija 2×2 .

Kod Gilbert-Eliotovog modela, gubici paketa su mogući i u dobrom i u lošem stanju za razliku od prethodno analiziranog Gilbertovog modela kod koga se paketi gube samo u lošem stanju. Prethodna

pristup NLR analizi se primenjuje i za Gilbert-Eliotov model gubitaka pri čemu je sada samo potrebno uzeti prethodno navedenu činjenicu u obzir. To će se realizovati promenom π_{loss} i π_{good} parametara kod kojih će se uključiti verovatnoće gubitaka u dobrom i lošem stanju, $1-k$ i $1-h$, respektivno. U skladu sa tim, za π_{loss} i π_{good} će sada važiti izrazi:

$$\pi_{loss} = \begin{bmatrix} 1-k & 0 \\ 0 & 1-h \end{bmatrix} \quad (3.59)$$

$$\pi_{good} = \begin{bmatrix} k & 0 \\ 0 & h \end{bmatrix} \quad (3.60)$$

Određivanje verovatnoće pojavljivanja odgovarajućeg niza paketa predstavlja preduslov za uspostavljanje analitičkog izraza za NLR parametar. Polazeći od NLR definicije, verovatnoća primetnih gubitaka se najlakše može odrediti tako što se izračuna razlika verovatnoće gubitka paketa i verovatnoće da izgubljeni paket nije primetan. Ova razlika predstavlja verovatnoću pojavljivanja paketa u skladu sa narednom relacijom stavljenom u kontekst funkcije događaja $L(n)$:

$$NLR^{(\delta)} = P[L(n) = 1] - P[L(n) = 1, L(n+1) = 0, \dots, L(n+\delta) = 0] \quad (3.61)$$

Iz prethodnog izraza se može zaključiti da se NLR parametar kome je pridruženo rastojanje δ može odrediti kao verovatnoća da je paket n izgubljen i da narednih δ paketa ne bude izgubljeno. Razlika verovatnoća ova dva događaja rezultuje verovatnoćom pojave primetnih gubitaka paketa što je i bio cilj ove analize. Dalje je potrebno ove verovatnoće izraziti u matricnoj formi u skladu sa modelom rafalnih gubitaka koji su izraženi na posmatranoj putanji.

$$NLR^{(\delta)} = P_{loss} - S \times \prod_{i=0}^{\delta-1} T_{trans} \times I \quad (3.62)$$

pri čemu je P_{loss} verovatnoća gubitaka paketa, dok ostali parametri u gore navedenom izrazu zavise od toga da li se radi o Gilbertovom ili Gilbert-Eliotovom modelu gubitaka paketa. Verovatnoća gubitaka P_{loss} se određuje posredstvom izraza (3.4) i (3.6) za Gilbertov i Gilbert-Eliotov model, respektivno.

Treba još naglasiti da se ovde posmatraju Gilbert i Gilbert-Eliotovi modeli kao stabilni Markovljevi modeli, pa se samim tim može smatrati da matrica stanja S u trenutku pre pristizanja paketa n je ista kao i nakon stizanja tog paketa, tačnije da važi jednakost $S(n-1)=S(n)$. Ova pretpostavka će važiti i kod naredne analize kada se rafalni gubici modeluju Markovljevim modelom sa 4 stanja.

3.7.2. NLR analiza u prisustvu gubitaka modelovanih Markovljevim modelom sa 4 stanja

Sličan pristup određivanju NLR parametra se može upotrebiti kada se razmatraju rafalni gubici koji se modeluju Markovljevim modelom sa četiri stanja. U skladu sa analizom obavljenom u delu 3.4.4. potrebno je najpre definisati odgovarajuće matrice čijom upotrebom se određuje razlika verovatnoće izgubljenog paketa i verovatnoće da narednih δ paketa ne bude izgubljeno. S obzirom da se Markovljev model razlikuje u odnosu na Gilbertov i Gilbert-Eliotov model koji su modeli sa 2 stanja, ovaj model sa 4 stanja se sada mora prilagoditi prethodno upotrebljenom pristupu.

Za potrebe određivanja NLR parametra kada se razmatra rastojanje δ najpre se definiše tranziciona matrica koja sadrži tranzicione verovatnoće prelaska iz bilo kog stanja u stanje u kome postoje gubici. U skladu sa definisanim Markovljevim modelom sa 4 stanja, tranziciona matrica T_{loss} koja će uračunati tranzicije u stanje sistema u kome će doći do gubitaka paketa je jednaka

$$T_{loss} = \begin{bmatrix} p_{11} & 0 & 0 & 0 \\ p_{21} & 0 & p_{23} & 0 \\ 0 & 0 & p_{33} & 0 \\ 0 & 0 & p_{43} & 0 \end{bmatrix} \quad (3.63)$$

Na sličan način se može definisati tranziciona matrica T_{good} koja sadrži isključivo tranzicione verovatnoće prelaska iz bilo kog stanja u stanje u kojima neće doći do gubitka paketa. Tranziciona matrica T_{good} je jednaka

$$T_{good} = I - T_{loss} \quad (3.64)$$

pri čemu se u matričnom obliku ova tranziciona matrica se može napisati kao

$$T_{good} = \begin{bmatrix} 0 & p_{12} & 0 & 0 \\ 0 & p_{22} & 0 & 0 \\ 0 & p_{32} & 0 & p_{34} \\ 0 & 0 & 0 & p_{44} \end{bmatrix} \quad (3.65)$$

S obzirom da se analizira Markovljev model sa četiri stanja, matrica stanja S je sada jednaka:

$$S = [S_1 \ S_2 \ S_3 \ S_4] \quad (3.66)$$

gde elementi matrice S označavaju verovatnoću pojedinačnog stanja u kome se sistem može naći.

Polazeći od razlike verovatnoće izgubljenih paketa i verovatnoće da narednih δ paketa ne bude izgubljeno, sada je moguće odrediti NLR parametar koristeći prethodno određene matrice stanja i odgovarajuće tranzicione matrice. NLR parametar je sada jednak:

$$NLR^{(\delta)} = P_{loss} - (S \times T_{loss} \times (T_{good})^{\delta} \times I) \quad (3.67)$$

pri čemu je P_{loss} verovatnoća gubitaka paketa koja je prethodno određena izrazom (3.22), I je matrica dimenzija 4×1 čiji su svi elementi jednaki 1 , dok se izrazom $S \times T_{loss} \times (T_{good})^{\delta} \times I$ određuje verovatnoća da nakon jednog izgubljenog paketa bude δ uspešno primljenih paketa. Ishod događaja će zavisiti od tranzicione matrice kao i od stanja u kome se sistem trenutno nalazi. U gore navedenom izrazu, kao i kod prethodnih modela gubitaka paketa, i kod Markovljev model sa četiri stanja se podrazumeva stabilnost modela čime se postiže pojednostavljenje prethodno predstavljene analize.

3.8. Princip rada paketske disperzije

Paketska disperzija predstavlja „raspršivanje“ paketa po različitim putanjama do odredišta koje su po pravilu razdvojene i koje imaju različitu strukturu gubitaka paketa (rafalnu i slučajnu). Cilj paketske disperzije jeste da se upotrebi postojeći nivo redundantnih putanja do odredišta kako bi se postigla promena strukture pojavljivanja paketa u pojedinačnim tokovima saobraćaja (servisima). Promenom strukture gubitaka paketa koja se analitički može opisati na različite načine moguće je direktno uticati na QoS koji se detektuje na prijemnoj stani u zavisnosti od osetljivosti posmatranog servisa na gubitke paketa.

Generalno gledano, disperzija saobraćaja je itekako zastupljena u telekomunikacionim mrežama. Neki od primera podrazumevaju balansiranje saobraćajnog opterećenja, povećanje nivoa sigurnosti mreže i saobraćaja koji se prenosi, zatim primena u pojedinim tehnologijama (*Code Division Multiple Access* - CDMA) kao i zbog poboljšanja kvaliteta prenosa u radio sistemima kada se koristi više kanala preko bežičnog medijuma u okviru koga su izražene smetnje. Za razliku od saobraćajne disperzije, paketska disperzija vrši slanje pojedinačnih paketa preko raspoloživih disperzionih putanja do specificiranog odredišta. Ova vrsta disperzije se može koristiti za poboljšanje iskorišćenja raspoloživih mrežnih resursa kao i za promenu strukture rafalnih gubitaka, što je i predmet u ovom radu. Međutim, u današnjim mrežama, paketska disperzija se često koristi za potrebe balansiranja saobraćaja tako što se na nivou paketa određuju putanje preko kojih će paketi biti prenošeni. Na taj način se poboljšava iskorišćenje, a ujedno i ostvaruje odgovarajući nivo redundanse. Primena paketske disperzije u cilju poboljšavanja performansi pojedinih aplikacija nije česta, dok se u [69] sugerše da se upotrebom paketske disperzije mogu u pojedinim slučajevima za VoIP aplikacije postići bolje performanse u odnosu na *Forwarding Error Correction* (FEC) tehnike.

Predulsov za realizaciju koncepta paketske disperzije predstavlja raspoloživost više putanja do jednog odredišta koje bi trebalo da budu fizički razdvojene i da nemaju zajedničkih tačaka. Na ovaj način se postiže statistička nezavisnost izraženih rafalnih gubitaka koji se mogu ispoljiti na ovim putanjama kada je cilj primeniti disperziju u cilju promene strukture gubitaka paketa. Primenom paketske disperzije je potencijalno moguće ostvariti povećanje rastojanja između gubitaka i promena strukture rafalnih gubitaka

kako bi se omogućio efikasniji rad PLC algoritama na odredištu. Ukoliko postoji korelacija rafalnih gubitaka između putanja, efikasnost paketske disperzije neće biti potpuna s obzirom da se povećanjem broja putanja ne postiže statistička nezavisnost izražene rafalnosti i disperzija nije u mogućnosti da u velikoj meri promeni strukturu rafalnih gubitaka. Statistička nezavisnost izraženih rafalnih gubitaka se postiže tako što se kao putanje za paketsku disperziju koriste putanje koje su međusobno fizički razdvojene, pri čemu je poželjno da nemaju ni jedan zajednički čvor (mrežni uređaj) koji može dovesti do korelacije strukture gubitaka paketa. Alternativno, mogu se posmatrati putanje koje imaju zajedničke segmente (čvorove i linkove), ali je potrebno da se obezbedi, ili sa visokom sigurnošću tvrdi, da na tim zajedničkim elementima neće biti izraženih rafalnih gubitaka. Ista pretpostavka važi i za Bernulijeve gubitke paketa za koje važi da međusobno nemaju statističku zavisnost.

Za pronalaženje fizički razdvojenih trebalo bi koristiti *multipath* protokole rutiranja pri čemu algoritam rutiranja kao rezultat pronalazi fizički razdvojene putanje između para izvorište-odredište. Po prirodi, ovi *disjoint* protokoli rutiranja su matematički kompleksniji s obzirom da su zahtevani uslovi za pronalaženje više putanja složeniji od konvencionalnih protokola rutiranja preko jedne putanje. Dodatno, vreme konvergencije ovih protokola je znatno duže za mreže sa velikim brojem linkova i čvorova u odnosu na postojeće IGP protokole.

Upotreba paketske disperzije odgovara povećanom nivou *multihoming*-a u modernim mrežama u skladu sa kojim su periferne tačke povezane sa dva i više linka kako bi se povećala raspoloživost krajnjih korisnika i obezbedio adekvatan nivo redundanse. Takođe, opravdanje realizacije većeg broja putanja između svake dve tačke u mreži svoje opravdanje nalazi ne samo u nivou redundanse, već i u balansiranju saobraćaja u skladu sa planiranom politikom rutiranja i/ili raspoloživim kapacitetima u pojedinim delovima mreže.

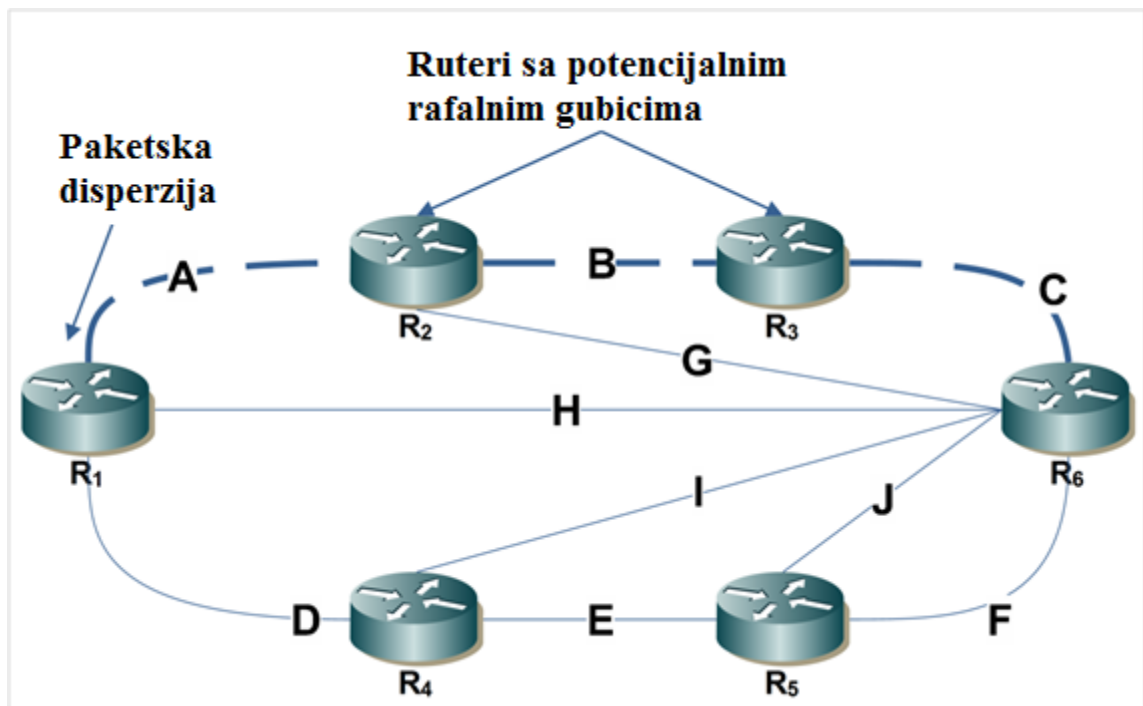
Integralni deo paketske disperzije jeste odabir strategije koja će odrediti kojim redosledom će susedni paketi biti prosleđivani na putanje sa različitim QoS. U kontekstu prethodne priče oko podele saobraćaja i selekcije putanja, paketska disperzija potpada pod mehanizme deljenja saobraćaja na paketskom nivou, dok fundamentalna razlika leži u cilju ovih mehanizama. Naime, paketska disperzija teži da promeni strukturu (šablon) pojavljivanja gubitaka paketa kako bi sa stanovišta QoS osetljivih servisa (*voice* i *video*) mogao da se smanji uticaj ove štetne pojave. Sa druge strane, cilj prethodno spomenutih mehanizama distribucije saobraćaja podrazumeva da se ostvari adekvatno balansiranje saobraćaja bez ikakvog osvrta na QoS zahteve servisa. U slučaju postojanja QoS zahteva, ta funkcionalnost se delegira QoS protokolima rutiranja koje vrše kvalitativno pronalaženje putanja koje su u mogućnosti da odgovore na postavljene zahteve.

Imajući u vidu da percepirani kvalitet *voice* aplikacija značajno opada u prisustvu rafalnih gubitaka paketa i da je efikasnost implementiranog PLC algoritma ograničena i zavisna od rafalnosti gubitaka paketa, paketska disperzija predstavlja adekvatan pristup koji treba preduzeti u mrežama sa visokim stepenom

redundanse. Posledično je moguće smanjiti nivo rafalnosti gubitaka paketa i posledično povećati efikasnost PLC algoritma. Sličan pristup se može koristiti i za ostale multimedijalne aplikacije kao što su različiti video servisi pri čemu su oni znatno osetljiviji na rafalne gubitke u odnosu na *voice* aplikacije.

Paketska disperzija podrazumeva uključivanje same mreže u poboljšanje efikasnosti pojedinih aplikacija pri čemu zajedno sa *loss recovery* tehnikama za poboljšanje robusnosti na izvorištu (FEC) i odredištu (PLC) može da dovede to očekivanog nivoa kvaliteta i zahtevanog nivoa performansi. Prethodno pomenute tehnike poboljšanja robusnosti zavise od krajnjih tačaka, odnosno terminala koji implementiraju ove mehanizme. Paketska disperzija se obavlja nezavisno od ovih tehnika u slučaju pojave rafalnih gubitaka na jednoj ili istovremeno na više putanja do odredišta.

Princip funkcionisanja paketske disperzije je prikazan na slici 3.15. Ruter R_1 do rutera R_6 šalje pakete putanjom A-B-C određenom od strane IGP protokola ili statičkim rutiranjem koje je u ovom trenutku sporedno sa stanovišta analize paketske disperzije i pridruženih strategija disperzije. Postojanje alternativnih i fizički odvojenih putanja do rutera R_6 daje mogućnost da se primeni paketska disperzija u ruteru R_1 tako da se svaki paket ka specifikiranom odredištu šalje preko različitih putanja u skladu sa strategijom disperzije.



Slika 3.15. Osnovni koncept paketske disperzije u redundantnoj topologiji.

Na ovaj način se koristi nivo redundanse primenjene u mreži, pri čemu se posmatraju putanje sa različitim QoS koji se ogleda u pojavi gubitaka paketa različite strukture. Paketskom disperzijom se postiže da se paketi iz pojedinačnih sesija koje se, na primer, prenose putanjom A-B-C na kojoj postoje rafalni gubici,

moгу poslati i na ostale putanje na kojima su gubici paketa drugačije strukture. U zavisnosti od strategije raspoređivanja paketa po putanjama zavisiće i koliko će se ostvariti dobitak na kvalitetu servisa jer će rezultujuća struktura gubitaka po jednoj sesiji biti sada drugačiji u odnosu na slučaj kada se šalje samo preko jedne putanje.

U zavisnosti od primenjene strategije paketske disperzije može se postići da se smanji srednja dužina gubitaka u okviru rafala, a samim tim i rastojanje uzastopnih gubitaka u pojedinačnim sesijama čime se postiže da mehanizmi koji vrše “sakrivanje” gubitaka paketa mogu efikasnije da funkcionišu s obzirom da se ne mogu uspešno sakriti gubici paketa koji su bliski ili se pojavljuju u rafalu.

Na slici 3.15 je na primarnoj putanji A-B-C prikazano da se gubici paketa dešavaju kao posledica *tail drop* discipline odbacivanja. Međutim, pojava rafalnih gubitaka je moguća i u prisustvu implementirane QoS politike s obzirom da se selektivno odbacivanje primenjuje na TCP saobraćaj, dok UDP paketi mogu biti u pojedinim slučajevima biti neselektivno odbačeni. Ovakvo odbacivanje neizbežno dovodi do pojave rafalnih gubitaka s obzirom da se *real time* servisi oslanjaju na UDP kao transportni protokol. Sa visokim porastom *multihoming*-a u današnjem Internetu [70], [71], primena paketske disperzije može u velikoj meri da nađe svoju ulogu u slučaju pojave rafalnih gubitaka u uslovima visoke raspoloživosti koja se ogleda u postojanju redundantnih putanja.

3.8.1. Relevantni radovi iz oblasti paketske disperzije

Postizanje očekivanog nivoa performansi se postiže posredstvom QoS tehnika i mehanizama koje implementiraju u mrežama i u pojedinačnim mrežnim uređajima. Ovi mehanizmi su u stanju da u većini slučajeva postignu odgovarajuću garanciju kvaliteta i nivoa performansi kroz prioritizaciju saobraćaja. U kontekstu *voice* servisa, prioritizacijom se postižu determinističke performanse ove klase saobraćaja koja je generalno osetljiva na kašnjenje i gubitke paketa. Međutim, gubici paketa mogu da budu posledica brojnih događaja u mreži koji ne moraju biti povezani sa zagušenjem što QoS mehanizmi nisu u stanju da reše. Adekvatan pristup rešavanju ovih gubitaka paketa bi bila modifikacija politika rutiranja kako bi se izmenom rutiranja problematični linkovi ili delovi mreže zaobišli.

Trenutna tendencija planiranja i projektovanja mreže jeste da se poveća nivo redundanse i raspoloživosti mrežnih resursa [72]. Projektovanjem mreže tako da postoji više od jedne putanje do većine odredišta u mreži ostavlja prostor za implementaciju *multipath* rutiranja nasuprot konvencionalnom rutiranju preko jedne putanje. Primena *multipath* rutiranja ima smisla imajući u vidu visoke zahteve po pitanju *multihominga* i mobilnosti u modernim mrežama [70], [71].

Nasuprot svojoj jednostavnosti primene i konceptu funkcionisanja, paketska disperzija se u praksi često izbegava zbog različitog kašnjenja paketa u okviru jedne sesije preko različitih putanja između jednog izvorišta i odredišta. Ova razlika u kašnjenju može da negativno utiče na TCP performanse jer kao

posledica različitog kašnjenja i varijacije kašnjenja na više putanja paketi mogu da stignu neispravnim redosledom što uzrokuje pogoršanje TCP performanse [73], [74].

U odnosu na TCP performanse, *real-time* aplikacije sa UDP protokolom se oslanjaju na RTP protokol koji će realizovati funkciju sekvenciranja paketa i identifikacije ispravnog redosleda paketa. U slučaju *real-time* servisa, konkretno VoIP servisa koji se koristi u ovom radu da pokaže efikasnost paketske disperzije, bitno je da paketi stignu na odredište konstantnim tokom i da se proslede gornjim slojevima. Kao razlog prethodne činjenice, različita kašnjenja na putanjama i visoka varijacija kašnjenja može da prouzrokuje dodatne gubitke paketa.

Iz perspektive balansiranja saobraćaja, paketska disperzija je sposobna da realizuje tačnije balansiranje u odnosu na balansiranje saobraćaja na nivou *flow*-a. U skladu sa rezultatima u [75] primena paketske disperzije može značajno da popravi paketsko kašnjenje usled zadržavanja u redovima čekanja. Međutim, bez obzira na prethodnu analizu, u slučaju značajne razlike kašnjenja između putanja može doći do primetne degradacije VoIP kvaliteta.

Kada se razmatra balansiranje opterećenja, ECMP rutiranje se najčešće koristi u mrežama kada protokoli rutiranja određuju putanje sa istom cenom [76], [77], [78]. Većina protokola rutiranja ne uzima u obzir kašnjenja ili gubitke paketa koji se ispoljavaju na pojedinim putanjama koje imaju istu cenu. Ovo može da dovede do različitih ostvarenih performansi u zavisnosti od putanje preko koje se servis realizuje. Standardno se u slučaju putanja sa istom cenom, nakon identifikacije putanja od strane algoritma rutiranja, primenjuje algoritam raspodele saobraćaja: *round robin*, *hashing*, *adaptivan*, itd. U zavisnosti od izraženih QoS parametara na ovim putanjama, primena ovih mehanizama distribucije saobraćaja će rezultovati različitim nivoom performansi. ECMP modeli raspodele saobraćaja nisu detaljno analizirani u literaturi sa aspekta eventualnog poboljšanja kvaliteta multimedijalnih aplikacija kao što je na primer VoIP, a posebno sa aspekta QoS parametara kao što je uticaj rafalnih gubitaka.

Koncept prebacivanja saobraćaja sa jedne putanje na drugu (*path switching*) je analiziran u [79], [80] i [81]. Ovaj koncept podrazumeva mehanizam rutiranja u skladu sa kojim se ostvaruju bolje performanse kroz korišćenje postojećih putanja do specificiranog odredišta tako što se saobraćaj prebacuje na jednu od postojećih putanja u zavisnosti od trenutnih karakteristika putanja. Termin „performanse“ pre svega podrazumeva perцепirani kvalitet koji se sa tehničke strane opisuje QoS parametrima. U kontekstu VoIP servisa, QoS parametri koji utiču na perцепirani kvalitet VoIP-a jesu kašnjenje, džiter i gubici paketa. Osnovni preduslov za realizaciju prebacivanja saobraćaja sa postojeće na drugu putanju do jednog odredišta jeste postojanje različitih putanja do odredišta – *path diversity* [83], [84], [85]. Diverzifikacija putanja zahteva implementaciju odgovarajućeg mehanizma koji će realizovati selekciju putanja u

zavisnosti od ispoljenih performansi kako bi se ostvario ciljani QoS nivo. Selekcija putanje je zasnovana na ideji da će se ostvariti željeni QoS u zavisnosti od zahtevanih performansi i aplikacija koje se koriste.

Za razliku od *path switching*-a, paketska disperzija koristi diverzifikaciju putanja kako bi se slali paketi preko raspoloživih putanja pri čemu se pretpostavlja nezavisnost karakteristika na pojedinačnim putanjama. Bitna razlika je što paketska disperzija koristi sve putanje za disperziju u skladu sa određenom disperzionom strategijom, dok *path switching* koristi samo jednu putanju dok god su karakteristike te putanje zadovoljavajuće za trenutne potrebe. Kao što je već rečeno, statistička nezavisnost performansi putanja podrazumeva da su te putanje fizički razdvojene (*link* i *node disjoint*). Kao sredstvo ostvarivanja ovog potrebnog uslova implementacije paketske disperzije koriste se *disjoint* protokoli rutiranja. Algoritmi za detekciju i određivanje *disjoint* putanja su prilično detaljno analizirani u literaturi pri čemu se predlaže nekoliko pristupa rešavanju problema [86].

Paketska disperzija i diverzifikacija putanja su mehanizmi koji se često koriste sa aspekta postizanja sigurnosti mreža, tačnije servisa i aplikacija u okviru mreža. Ovi mehanizmi koriste redudansu mreže kako bi sprečili presretanje određenih paketa koji pripadaju aplikacijama visokog prioriteta čiji paketi i njihov sadržaj ne sme biti otkriven malicioznim entitetima. Praktično, povećanjem broja putanja koji se koristi za prenos ovog tipa saobraćaja se smanjuje verovatnoća presretanja. Konkretno, paketska disperzija se može koristiti za potrebe obezbeđivanja VoIP servisa bez potrebe za enkripcijom tako što će se po određenoj strategiji paketi slati preko više putanja. Ovakvom disperzijom je cilj da presretač ne “uhvati” bliske pakete čijim će dekodovanjem moći da prisluškuje celokupan ili deo razgovora. Primenom paketske disperzije, ukoliko presretač uhvati sve pakete sa jedne putanje, ti paketi su u *voice* strimu previše udaljeni kako bi se adekvatno dekodovao i prisluškivao jedan VoIP razgovor. Naravno, sigurnost se postiže tako što se odredi adekvatna strategija disperzije u zavisnosti od pridružene kvantitativne mere kojom se opisuju verovatnoća da presretač ugrozi jednu ili više putanja. Detaljna analiza ovih strategija je obavljena u [87] i relevantna je u pogledu razmatranja širokog upotrebnog spectra paketske disperzije.

Aspekt sigurnosti mreža ostvaren paketskom disperzijom i diverzifikacijom putanja se korisnicima može nuditi kao dodatni servis koji ne zahteva implementaciju enkripcije, provere integriteta ili autentifikacije. Ove funkcije generalno podrazumevaju implementaciju na perifernim uređajima što dodatno komplikuje proces implementacije. Primena politike rutiranja u cilju povećanja sigurnosti predstavlja primitivan i jednostavan način implementacije sigurnosti kojom se može postići nizak ili umeren nivo zaštite. Takođe, *multipath* rutiranje i balansiranje saobraćaja se može upotrebiti u kontekstu sigurnosti servisa i saobraćaja. Mogućnosti ovih mehanizama po pitanju bezbednosti mreža su prikazani u [88], [89], [90]. Pristup koji se ovde koristi podrazumeva da se koristi visok nivo raspoloživosti putanja kako bi se smanjila verovatnoća presretanja pojedinačnih paketa koji pripadaju individualnim sesijama što ne podrazumeva konvencionalni pristup kada se koristi procesorski zahtevna enkripcija.

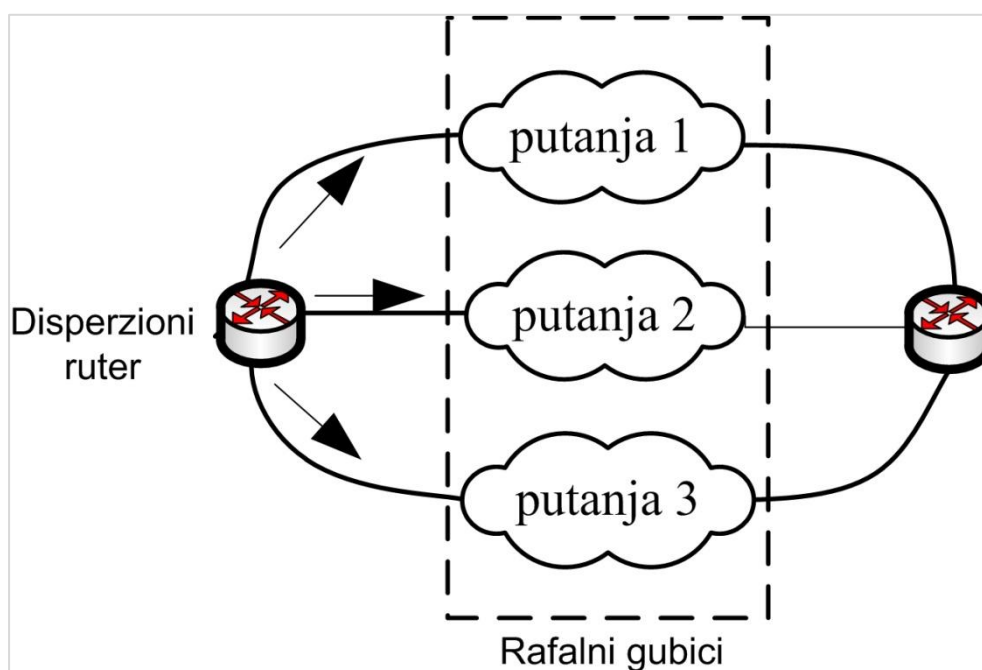
Korišćenje redundantnih putanja i slanje redundantnih paketa preko tih putanja za poboljšanje VoIP QoS je analizirano u [91] i [92]. Naime, u ovim radovima je predstavljen uticaj slanja redundantnih paketa preko putanja sa izraženim rafalnim gubicima pri čemu se posmatra dobijeni kvalitet ocenjen objektivnim E-Modelom. Zaključak je da se primenom redundantnih putanja i duplikacije paketa dobija značajno na poboljšanju kvaliteta čak i prilikom visokih rafalnih gubitaka na pojedinačnim putanjama. Iako se ostvaruju dobri rezultati u pogledu QoS, prilikom implementacije ovog pristupa redundantne disperzije paketa, potrebno je imati u vidu raspoloživost kapaciteta disperzionih putanja s obzirom da se radi o duplikaciji paketa, pa i sam VoIP saobraćaj se uvećava preko ovih putanja. Sa druge strane, moraju se posmatrati i mogućnosti krajnjih tačaka, odnosno da li su krajnji VoIP terminali sposobni da izađu na kraj sa velikim brojem dupliciranih paketa što podrazumeva implementaciju specijalnog *de-dispersion* bafera koji odbacuje sve pristigle duplicirane pakete osim prvog. Međutim, imajući u vidu veoma dobre rezultate, u ovom radu je uključena detaljna analiza redundantne strategije paketske disperzije.

Najznačajniji doprinos vezano za primenu paketske disperzije u cilju poboljšanja QoS je predložen i analiziran u [93]. Pokazano je da koristeći *round robin* i *random* strategije paketske disperzije moguće značajno uticati na rastojanje gubitaka paketa kada su na jednoj ili više putanja izraženi Bernulijevi ili rafalni gubici paketa. Analiza rastojanja gubitaka paketa u [93] predstavlja osnovni polazak za doprinos u ovom radu s obzirom da ova analiza podrazumeva samo ispitivanje prethodno pomenutih strategija, dok se generalno ostavlja prostora za predlaganje i analiziranje dodatnih strategija. Takođe, evaluacioni deo je uzeo u obzir samo postojanje dve putanje, dok se ponašanje predloženog modela samo pretpostavlja, ali ne i dokazuje. Suštinski, uticaj postojanja više od dve disperzione putanje u zavisnosti od primenjene strategije nije analiziran. Osim toga, analitički model je samo analizirao NLR parametar koji ocenjuje rastojanje, a sam model se nije stavio u kontekst ocene dobijenog kvaliteta VoIP aplikacija pri čemu je potrebno napomenuti da se za modelovanje rafalnih gubitaka koristio Gilber-Eliotov model, dok su rezultati posmatrani za rafalne gubitke modelovane Markovljevim modelom sa dva stanja koji je kada je u pitanju VoIP saobraćaj manje tačan u odnosu na Markovljev model sa 4 stanja.

Imajući u vidu prethodne nedostatke analize u [93], u ovom radu je predstavljen kompletniji i matematički tačniji model kojim se može opisati rastojanje gubitaka posredstvom NLR parametara. Dodatno, ostvareni QoS posredstvom odgovarajuće strategije paketske disperzije se stavlja u kontekst objektivne ocene VoIP kvaliteta koji se realizuje E-Modelom. Predloženi model se testira na realnoj opremi i VoIP servisu kako bi se dobio uvid u ostvareni QoS primenom individualnih strategija paketske disperzije. Testiranjem se vrši verifikacija predloženog matematičkog modela i istovremeno proverava njegova tačnost. Osim postojećih *round robin* i *random* paketskih disperzija biće predložene i adaptivne i redundantne paketske disperzije.

3.9. Strategije paketske disperzije

Kao što je već navedeno, cilj paketske disperzije može da bude višestruk u zavisnosti od scenarija upotrebe. Dakle, ukoliko se radi o balansiranju saobraćaja, cilj je da se uvede mehanizam razvrstavanja paketa koji će diktirati redosled slanja paketa po putanjama kako bi se obezbedio adekvatan nivo saobraćaja na ovim putanjama. Sa druge strane, ako se radi o primeni paketske disperzije kao sigurnosnog mehanizma, potrebno je pristupiti odgovarajućem algoritmu koji će upravljati redosledom slanja paketa po putanjama u zavisnosti od servisa kao i metrike koja je dodeljena pojedinačnim putanjama i koja opisuje verovatnoću presretanja paketa.



Slika 3.16 Generalizacija strategija paketske disperzije.

Za razliku od prethodnih ciljeva, kada se paketska disperzija upotrebljava za poboljšanje QoS performansi, tada je potrebno realizovati strategije koje će odrediti redosled slanja pojedinačnih paketa na putanje u zavisnosti od QoS karakteristika koje su izražene na svakoj od ovih disperzionih putanja. Naravno, raspoloživost putanja zavisi od samog dizajna mrežne topologije pri čemu je za veću efikasnosti paketske disperzije potrebno da bude više putanja i da one budu fizički razdvojene od izvora do odredišta – *disjoint* putanje. Primer upotrebe paketske disperzije iz perspektive primenjenih strategija prikazan je na slici 3.16.

Funkcija disperzionog rutera podrazumeva da on u svojoj tabeli rutiranja ima informaciju o razdvojenim putanjama koje su unete statički ili posredstvom odgovarajućeg dinamičkog protokola rutiranja. Ovaj ruter predstavlja najznačajniji deo procesa paketske disperzije s obzirom da se u okviru njega implementiraju strategije paketske disperzije koje će diktirati redosled slanja paketa preko raspoloživih disperzionih putanja. Naime, svaki paket koji pripada određenom servisu ili sesiji (na primer, VoIP) će biti procesiran u

skladu sa strategijom paketske disperzije koja određuje putanju do odredišta kojom će posmatrani paket biti poslat ka odredištu.

Ako se posmatraju VoIP paketi, to znači da ukoliko se koristi *round robin* strategija, svaki naredni paket će biti poslat preko druge putanje u odnosu na prethodni paket. Imajući u vidu da efikasnost strategija paketske disperzije zavisi od statističke nezavisnosti rafalnih gubitaka izraženih na putanjama, gubljenje uzastopnih paketa na jednoj putanji podrazumeva da ti paketi ne pripadaju istoj sesiji i da se na nivou sesije ne nalaze jedan do drugog. Ovakvo delovanje strategija paketske disperzije proređuje gubitke paketa i čini ih da budu na većem rastojanju. Posledično, povećanje rastojanja između gubitaka omogućava da PLC algoritmi efikasnije sakriju izgubljene pakete u okviru sesije pri čemu se na taj način povećava percepirani kvalitet na prijemnoj strani.

Evidentno je da neće sve strategije imati isti uticaj na menjanje strukture gubitaka paketa u okviru jedne VoIP sesije, a takođe će rezultatna struktura gubitaka zavisiti i od strukture gubitaka izražne na disperzionim putanjama. Takođe, postoji dosta varijanti prilagođavanja primenjene strategije paketske disperzije trenutnom QoS stanju na putanjama. Imajući ovo u vidu, potreban je model koji će na osnovu parametara kojima se opisuju rafalni gubici na putanjama deterministički da odredi kolika je očekivana vrednost NLR parametra. Rezultat ovog modela treba da pokaže koja će strategija za određeni scenario sa više raspoloživih putanja i struktura gubitaka paketa dati najbolji rezultat, odnosno ostvariti najbolji E2E QoS.

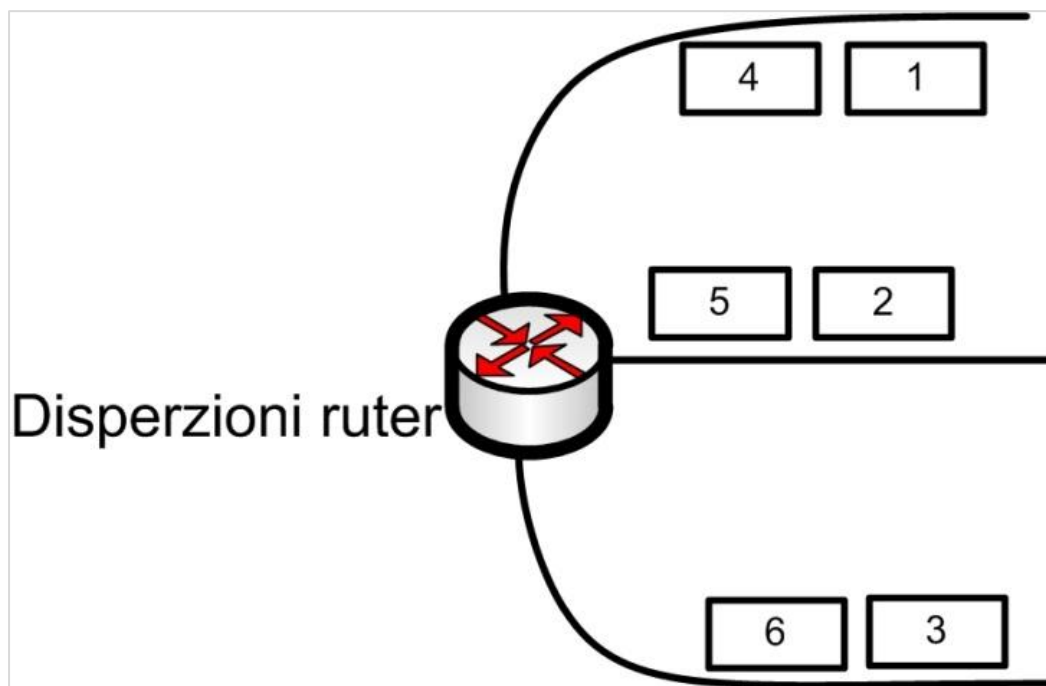
Najpre će se analizirati *round robin*, *random*, adaptivna i redundantna strategija za Gilbert-Eliotov model, a zatim će se ovaj model proširiti za svaku od ovih strategija na Markovljev model sa 4 stanja.

3.9.1.Round robin strategija

Round robin strategija predstavlja mehanizam paketske disperzije prema kojoj se paketi raspoređuju po putanjama u *round robin* režimu. Na slici 3.17 je prikazan princip funkcionisanja *round robin* strategije na primeru šest uzastopnih paketa koji su numerisani kada se vrši disperzija preko tri putanje. Paketi se sada šalju u skladu sa *round robin* redosledom čime se omogućava da uzastopni paketi budu poslani različitim putanjama na kojima je izražena statistički različita struktura gubitaka paketa.

Round robin strategija suštinski predstavlja specijalni slučaj periodične strategije disperzije koja podrazumeva periodični raspored paketa preko disperzionih putanja. U skladu sa ovom strategijom je moguće proizvoljno odrediti raspored paketa po putanjama ako se uzme u obzir periodičan redosled rasporeda paketa. Takođe, periodičnom strategijom je moguće da se isplanira raspoređivanje paketa u skladu sa očekivanom strukturom gubitaka paketa na posmatranoj putanji.

U ovom radu će se posmatrati samo *round robin* strategija kao predstavnik periodične strategije disperzije. Ova strategija podrazumeva da se svaka putanja posmatra nezavisno u smislu izraženih gubitaka paketa, pa se svaki uzastopni paket šalje po narednoj raspoloživoj putanji. U tom smislu, sa *round robin* strategijom nije moguće prioritetizovati putanje u skladu sa rafalnim gubicima paketa, pa samim tim i dobijeni nivo performansi zavisi od konkretne strukture rafalnih gubitaka na svakoj od putanja.

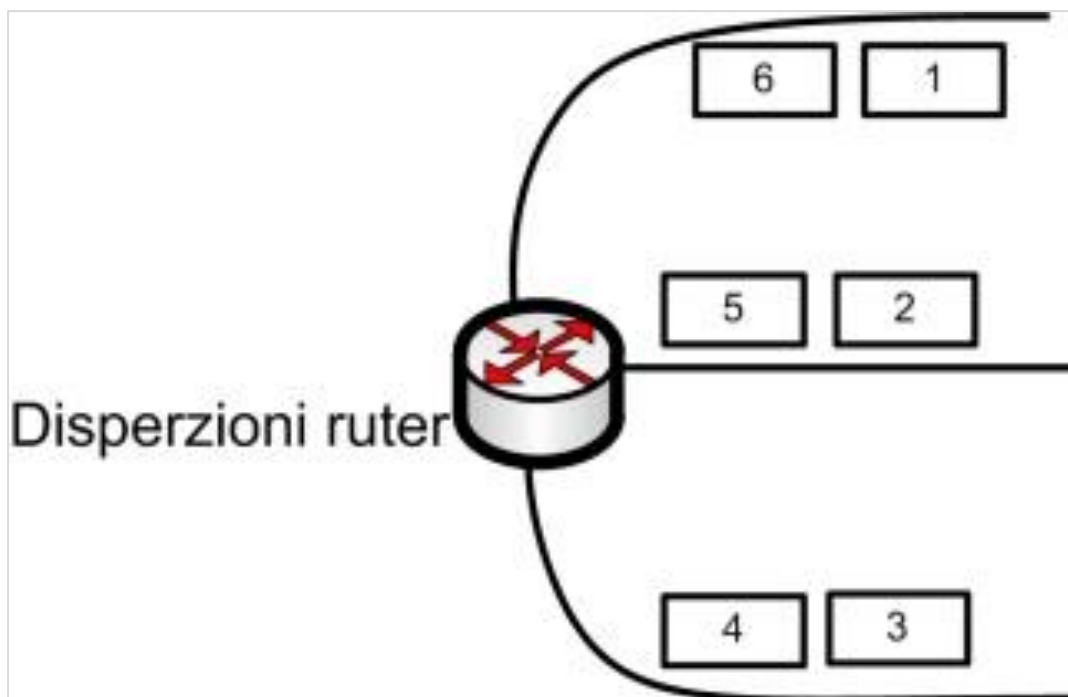


Slika 3.17 Round robin strategija paketske disperzije.

3.9.2. Random strategija

Primenom *random* strategije se omogućava određeni nivo prioritetizacije pojedinih putanja u zavisnosti od rafalnih gubitaka na putanjama tako što se putanjama sa većim rafalnim gubicima paketa dodele manje verovatnoće, dok putanjama sa boljim performansama u vidu rafalnih gubitaka paketa bi se dodeljivale veće verovatnoće. Na ovaj način je moguće da se na putanje sa boljim performansama pošalje veći broj paketa. Princip rada *random* strategije je prikazan na slici 3.18.

Pridruživanjem različitih verovatnoća putanjama je sada moguće da pojedini uzastopni paketi budu poslani preko iste putanje, čak i kada je podjednaka verovatnoća pridružena različitim putanjama. Za potrebe analize *random* strategije, u ovom radu će biti analizirana *random* strategija kada su jednake verovatnoće pridružene disperzionim putanjama.



Slika 3.18 Random strategija paketske disperzije.

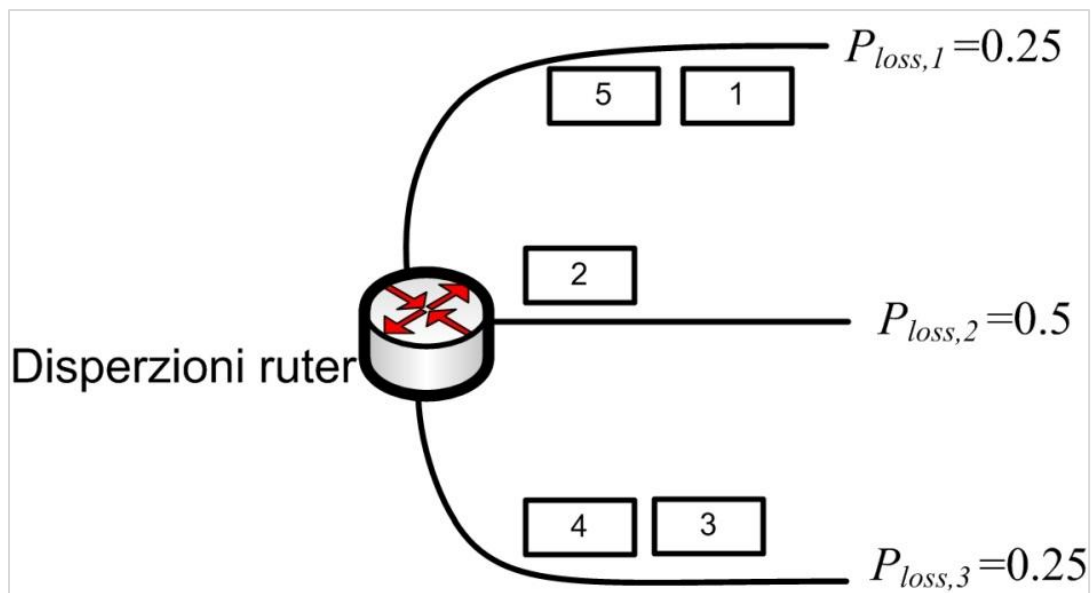
3.9.3. Adaptivna strategija

Adaptivna strategija podrazumeva specijalni slučaj *random* strategije koja uzima u obzir verovatnoću gubitaka paketa koja je izražena na disperzionim putanjama. Naime, verovatnoća slanja paketa po putanji k je veća ukoliko je verovatnoća gubitaka manja u odnosu na ostale disperzione putanje. Verovatnoća φ_k označava verovatnoću slanja paketa preko putanje k i određuje se u skladu sa sledećom relacijom:

$$\varphi_k = \frac{1}{N-1} \cdot \frac{\sum_{i=1, i \neq k}^N P_{loss,i}}{\sum_{i=1}^N P_{loss,i}} \quad (3.68)$$

gde je N ukupan broj putanja raspoloživih za proces disperzije paketa, dok je $P_{loss,i}$ verovatnoća gubitaka paketa na putanji i .

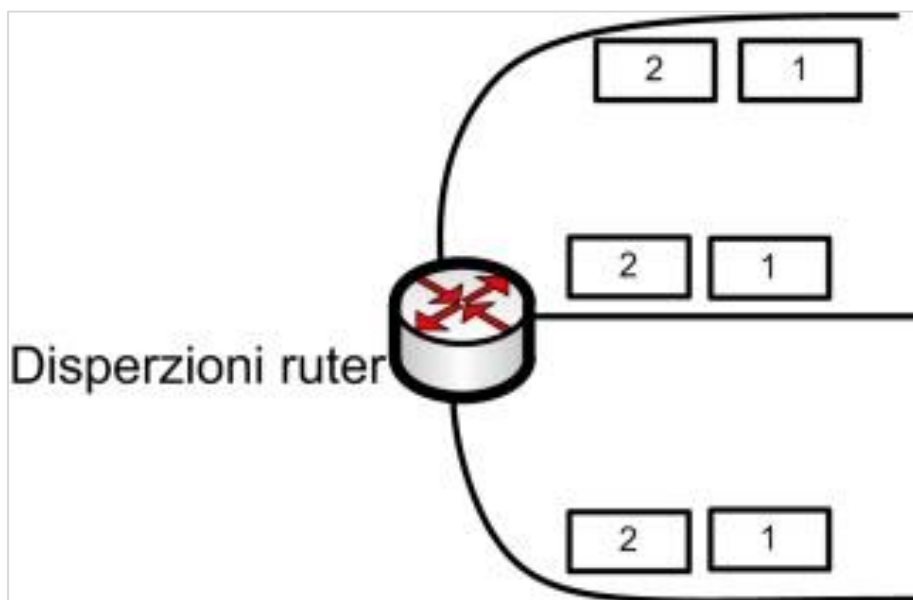
Na slici 3.19 je prikazan primer funkcionisanja adaptivne strategije kada se posmatraju tri putanje pri čemu imaju verovatnoću gubitaka paketa 0.25, 0.5 i 0.25, respektivno. Kao što se vidi, na putanjama 1 i 3, koje imaju dva puta manju verovatnoću gubitaka paketa u odnosu na putanju 2, šalju se dva puta više paketa. Na ovaj način se postižu bolje performanse u pogledu blizine gubitaka paketa jer se manje paketa šalje preko problematičnih putanja, pa samim tim se i očekuje veće rastojanje između susednih gubitaka paketa, odnosno bolja vrednost NLR parametra.



Slika 3.19 Adaptivna strategija paketske disperzije.

3.9.4.Redundantna strategija

Redundantna strategija se razlikuje od prethodno pomenutih strategija po tome što se ne zasniva na konceptu odabira slanja pojedinačnog paketa po disperzionim putanjama, već se u skladu sa ovom strategijom svaki pojedinačan paket šalje po svim raspoloživim putanjama. Na ovaj način se praktično pravi duplikacija paketa, pa se kopije paketa šalju po svakoj putanji.



Slika 3.20 Redundantna strategija paketske disperzije.

Redundantna strategija se zasniva na principu da postoji dovoljno kapaciteta na putanjama koje sada moraju da prenesu veći broj paketa jer u zavisnosti od broja putanja koje učestvuju u redundantnoj paketskoj disperziji, potreban kapacitet se višestruko povećava. Takođe, redundantna strategija je sposobna

da se nosi sa visokim gubicima paketa koji su izraženi istovremeno na više putanja kao rezultat povećane verovatnoće da jedan paket stigne do odredišta kada se istovremeno pošalje preko nekoliko putanja sa različitim QoS.

Koncept redundante strategije na primeru slanja dva paketa je prikazan na slici 3.20. Kao što se vidi, svaki paket se istovremeno šalje preko tri putanje čime se u ovom slučaju tri puta povećava zahtev za kapacitetom. Međutim, redundantna strategija predstavlja odnos kvaliteta i kapaciteta i predstavlja adekvatnu strategiju kada postoje istovremeno izraženi rafalni gubici preko više putanja.

3.10. Analiza rastojanja gubitaka kod strategija u prisustvu Gilbert-Eliotovih gubitaka paketa

Analiza rastojanja gubitaka paketa za pojedinačne strategije paketske disperzije je urađena u [93] i zasniva se na posmatranju NLR parametra kada se koriste *round robin* i *random* strategije. Konkretno se u okviru ove analize razmatra postojanje Gilbert-Eliotovih gubitaka paketa na putanjama. Takođe, analiza u [93] će biti proširena dodatnim strategijama – adaptivnom i redundantnom strategijom.

3.10.1. NLR analiza kod round robin strategije

Analiza se započinje definisanjem relevantnih parametara kojima će se adekvatno opisati *round robin* strategija. Kao što je već rečeno, *round robin* strategija predstavlja specijalni slučaj periodične strategije, pa je samim tim potrebno definisati periodičnu funkciju $Q(i)$ sa periodom K , pri čemu za paket i , ova funkcija ima vrednost koja pripada skupu $Q(i) \in \{1, 2, \dots, k, \dots, N\}$, gde je N broj putanja. Upravo rezultat ove periodične funkcije $Q(i)$ će odlučiti kojom putanjom će posmatrani paket i biti poslat do odredišta.

Imajući u vidu NLR definiciju koja je data u relaciji (3.61), primetni gubici se neće desiti ako se paket i izgubi na putanji P_k , dok narednih δ uzastopnih paketa ne bude izgubljeno na putanjama koje su određene u skladu sa periodičnom funkcijom $Q(i+l)$ tako da je $l \in \{1, 2, \dots, \delta\}$. Za pakete za koje važi nejednakost $Q(i+l) \neq k$, tačnije za pakete koji se prenose na ostalim putanjama, nije od važnosti da li su primljeni ili izgubljeni. Samim tim, funkcija događaja $L(i)$ se mora proširiti kako bi se uračunao i događaj u okviru koga nije bitno da li se radi o gubitku ili prijemu paketa. Pomenuti događaj će biti nazvan „nebitnim“ događajem, i za paket i za koga je događaj nebitan važi jednakost $L(i)=2$. Istovremeno, za narednih δ uzastopnih paketa za koje važi jednakost $Q(i+l)=r$ posmatraju se samo događaji u kojima nema gubitaka paketa na putanji $P_r \neq P_k (r \in \{1, 2, \dots, N\}, r \neq k)$.

U cilju analitičkog opisivanja prethodne pretpostavke potrebno je definisati novu tranzicionu matricu $A_i^{L(i)}$ za putanju $P_{Q(i)}$ kojom se uračunavaju događaji opisani funkcijom događaja $L(i)$:

$$A_{Q(i)}^{L(i)} = \begin{cases} A_{Q(i)} \cdot (1 - B_{loss, Q(i)}), & \text{za } L(i) = 0 \\ A_{Q(i)} \cdot B_{loss, Q(i)}, & \text{za } L(i) = 1 \\ A_{Q(i)}, & \text{za } L(i) = 2 \end{cases} \quad (3.69)$$

gde je $A_{Q(i)}$ tranziciona matrica Gilbert-Eliotovog modela za putanju $P_{Q(i)}$, dok je $B_{loss, Q(i)}$ matrica verovatnoće gubitaka paketa koja je jednaka:

$$B_{loss, i} = \begin{bmatrix} P_G & 0 \\ 0 & P_B \end{bmatrix} \quad (3.70)$$

gde je P_G i P_B verovatnoća gubitka paketa u dobrom i lošem stanju, respektivno.

Ako se uzmu u obzir NLR definicija po kojoj je NLR razlika verovatnoće izgubljenog paketa i verovatnoće da narednih δ uzastopnih paketa ne bude izgubljen, kao i parametri periodična funkcije $Q(i)$ i funkcija događaja $L(i)$ na N raspoloživih putanja, može se posmatrati sumariзована kombinacija događaja na putanjama [93]:

1. Na putanji P_k :
 - a. Dešava se gubitak paketa i , funkcija događaja je $L(i)=1$;
 - b. Nema gubitaka paketa na narednih δ paketa ($i+1, i+2, \dots, i+\delta$) koji se prenose preko putanje P_k u skladu sa periodičnom funkcijom $Q(i)=k$ pri čemu $i \in \{1, 2, \dots, \delta\}$;
 - c. Za sve ostale pakete posmatra se događaj $L(i)=2$.
2. Na putanji P_r :
 - a. Paket i predstavlja nevažan događaj za putanju P_r , tj. važi $L(i)=2$;
 - b. Nema gubitaka paketa na narednih δ paketa ($i+1, i+2, \dots, i+\delta$) koji se prenose preko putanje P_r u skladu sa periodičnom funkcijom $Q(i)=r$ pri čemu $i \in \{1, 2, \dots, \delta\}$;
 - c. Za sve ostale pakete posmatra se događaj $L(i)=2$.

Kako bi se uračunali funkcija događaja $L(i)$ i periodična funkcija $Q(i)$ koja određuje putanju kojom će paket i biti prosleđen prema odredištu, potrebno je definisati još jednu matricu $T_{Q(i)}$ koja će zavisiti od periodične funkcije $Q(i)$:

$$T_{Q(i)} = \begin{cases} A_k^0, & Q(i) = k \\ A_k^2, & Q(i) \neq k \end{cases} \quad (3.71)$$

gde su matrice A_i^0 i A_i^2 definisane relacijom (3.69) za $L(i)=0$ i $L(i)=2$, respektivno. Periodičnom matricom $T_{Q(i)}$ se suštinski stavlja u kontekst funkcija događaja sa periodičnom funkcijom koja predstavlja integralni deo periodične strategije disperzije, odnosno *round robin* strategije.

Ako se prvi izgubljeni paket desio na putanji P_i , onda je verovatnoća pojave primetnog gubitka jednaka razlici verovatnoće izgubljenog paketa na putanji P_i i razlici verovatnoće da narednih δ uzastopnih paketa na bude izgubljeno u skladu sa gore navedenim uslovima za periodičnu strategiju disperzije. Iz analitičke perspektive, verovatnoća pojave primetnih gubitaka kada se razmatra inicijalni gubitak paketa na putanji P_k se određuje prema relaciji:

$$NLR_k^\delta = S_{Q(i)} \cdot B_{loss, Q(i)} - \left((S_{Q(i)} \cdot B_{loss, Q(i)} \cdot (\prod_{i=1}^\delta T_{Q(i)}) \cdot I) \cdot \prod_{r=1, r \neq Q(i)}^N (S_r \cdot \prod_{i=1}^\delta T_r) \cdot I \right) \quad (3.72)$$

gde je $S_{Q(i)}$ matrica stanja dimenzija 1×2 koja sadrži verovatnoću da putanja $P_{Q(i)}$ bude u dobrom stanju P_G i u lošem stanju P_B , dok je I matrica dimenzije 4×1 čiji su svi elementi jednaki 1.

Ukupna verovatnoća primetnih gubitaka NLR^δ se određuje kada se uzmu u obzir da se inicijalni gubici paketa dešavaju na svim raspoloživim putanjama. U tom slučaju, prethodna relacija se modifikuje kako bi se uračunao prethodni iskaz:

$$NLR^\delta = \frac{1}{K} \sum_{i=1}^K \left(S_{Q(i)} \cdot B_{loss, Q(i)} - \left((S_{Q(i)} \cdot B_{loss, Q(i)} \cdot (\prod_{i=1}^\delta T_{Q(i)}) \cdot I) \cdot \prod_{r=1, r \neq Q(i)}^N (S_r \cdot \prod_{i=1}^\delta T_r) \cdot I \right) \right) \quad (3.73)$$

gde je K period funkcije $Q(i)$.

Kompleksnost izračunavanja primetnih gubitaka za periodičnu strategiju (*round robin*) se može izraziti kao $O(N \cdot K \cdot \delta)$ za razliku od direktnog izračunavanja NLR parametra u okviru koga bi se proračun obavio u N -dimenzionom prostoru [93]. Korišćenjem prethodnog pristupa problemu izračunavanja NLR parametra je moguće u velikoj meri uprostiti računsku kompleksnost, pa se samim tim ovaj pristup koristi i kod složenijih Markovljevih modela koji su sposobni da tačnije modeluju rafalne gubitke, na primer kod Markovljevog modela sa 4 stanja.

Treba se naglasiti da se prethodna relacija odnosi na NLR parametar kada se posmatra jedna sesija. U slučaju postojanja F sesija, određuje se očekivana NLR vrednost po sesiji koja se određuje kao srednja vrednost NLR parametra po pojedinačnoj sesiji u skladu sa sledećom relacijom:

$$\overline{NLR}^\delta = \frac{1}{N} \sum_{f=1}^F NLR_f^\delta \quad (3.74)$$

gde je sada NLR_f^δ verovatnoća primetnih gubitaka koja se određuje za sesiju f . U daljoj analizi će biti izostavljeno posmatranje NLR parametra po sesiji s obzirom da ne nosi preveliki značaj u određivanju resultantne strukture gubitaka u okviru jedne sesije kada se razmatraju različite strategije paketske disperzije.

3.10.2. NLR analiza kod random i adaptivne strategije

Određivanje NLR parametra kada se primenjuje *random* strategije se realizuje primenom koncepta ekvivalentne putanje. Kao što je već rečeno, za ostvarivanje dobrih performansi paketske disperzije potrebno je da disperzione putanje izražavaju statistički nezavisne osobine rafalnih gubitaka. Samim tim je moguće realizovati putanju koja će imati strukturu rafalnih gubitaka koja je ekvivalentna primeni *random* disperzije nad N putanja sa odgovarajućom strukturom rafalnih gubitaka paketa.

Postupak NLR analize kod *random* strategije podrazumeva određivanje strukture rafalnih gubitaka na ekvivalentnoj putanji nakon čega se može direktno primeniti definicija po kojoj se određuje NLR preko jedne putanje. Kada se razmatra Gilbert-Eliotov model na svakoj od putanja, ekvivalentna putanja će tada imati strukturu gubitaka paketa koja se može opisati Markovljevom modelom sa 2^N stanja i tranzicionom matricom dimenzija $2^N \times 2^N$ s obzirom da se u slučaju postojanja N disperzionih putanja mora razmatrati kombinacija svih mogućih stanja.

Matrica stanja ekvivalentne putanje π je dimenzija 1×2^N pri čemu je svaki element jednak proizvodu verovatnoće stanja na pojedinačnim putanjama. Tranziciona matrica Γ je dimenzija $2^N \times 2^N$ i predstavlja tranziciju verovatnoću iz jednog od 2^N mogućih stanja u kojima se ekvivalentna putanja može naći. Primer matrice stanja i tranzicione matrice ekvivalentne putanje za $N=2$ je dat sledećim relacijama, respektivno:

$$\pi = [S_{G1} \cdot S_{G2} \quad S_{G1} \cdot S_{B2} \quad S_{B1} \cdot S_{G2} \quad S_{B1} \cdot S_{B2}] \quad (3.75)$$

$$\Gamma = \begin{bmatrix} (1-p_1)(1-p_2) & (1-p_1)p_2 & p_1(1-p_2) & p_1p_2 \\ (1-p_1)q_2 & (1-p_1)(1-q_2) & p_1q_2 & p_1(1-q_2) \\ q_1(1-p_2) & q_1p_2 & (1-q_1)(1-p_2) & (1-q_1)p_2 \\ q_1q_2 & q_1(1-q_2) & (1-q_1)q_2 & (1-q_1)(1-q_2) \end{bmatrix} \quad (3.76)$$

gde su S_{Gk} i S_{Bk} verovatnoća dobrog i lošeg stanja na putanji P_k , respektivno, dok su elementi matrice Γ direktno izvedeni iz mogućih kombinacija stanja Gilbert-Eliotovih tranzicionih verovatnoća p i q na primeru dve putanje.

Dalje je potrebno da se u ekvivalentnu putanju implementiraju verovatnoće φ_i koje u suštini određuju sa kojom verovatnoćom će se paket proslediti po određenoj putanji. Ekvivalentna putanja takođe mora da uzme u obzir ove verovatnoće, a to će biti urađeno tako što će iste biti pridružene verovatnoći gubitaka

paketa u okviru jedne putanje jer će φ_i parametar direktno uticati na verovatnoću gubitaka paketa kada je ekvivalentna putanja u nekom od stanja, bilo dobrom ili lošem.

Za verovatnoće pridružene putanjama važi sledeća jednakost:

$$\sum_{i=1}^N \varphi_i = 1 \quad (3.77)$$

Neposredno zatim se definiše matrica koja uračunava verovatnoću gubitaka paketa kao i verovatnoće φ_i prema sledećoj relaciji na primeru dve putanje:

$$B_{loss,\varphi} = \begin{bmatrix} \varphi_1 \cdot P_{G1} + \varphi_2 \cdot P_{G2} & 0 & 0 & 0 \\ 0 & \varphi_1 \cdot P_{G1} + \varphi_2 \cdot P_{B2} & 0 & 0 \\ 0 & 0 & \varphi_1 \cdot P_{B1} + \varphi_2 \cdot P_{G2} & 0 \\ 0 & 0 & 0 & \varphi_1 \cdot P_{B1} + \varphi_2 \cdot P_{B2} \end{bmatrix} \quad (3.78)$$

gde su P_{Gk} i P_{Bk} verovatnoće gubitaka u dobrom i lošem stanju na putanji P_k , respektivno.

Nakon određivanja potrebnih matrica Γ , π , $B_{loss,\varphi}$ sada je moguće odrediti NLR parametar u skladu sa definicijom kada postoji jedna putanja sa izraženim rafalnim gubicima paketa prema sledećoj formuli

$$NLR^{(\delta)} = \pi \cdot B_{loss,\varphi} \cdot I - \pi \cdot (\Gamma \cdot B_{loss,\varphi} (\Gamma \cdot (I_d - B_{loss,\varphi})^\delta) \cdot I) \quad (3.79)$$

u kojoj je identitet matrica I_d dimenzija 4×4 , a I matrica dimenzija 1×4 gde su svi elementi jednaki 1.

Što se tiče adaptivne strategije, NLR analiza se poklapa sa prethodno izvršenom *random* strategijom pri čemu je verovatnoća φ_i promenljiv parametar i generalno zavisi od verovatnoće gubitaka paketa na pojedinačnim putanjama. NLR vrednost za adaptivnu strategiju se izračunava po relaciji (3.79) pri čemu se parametar φ_i određuje prema izrazu (3.68).

3.10.3. NLR analiza kod redundantne strategije

Redundantna strategija podrazumeva replikaciju pojedinačnih paketa preko raspoloživih putanja disperzije. Pristup koji se kod ove strategije može iskoristiti je kombinacija određivanja NLR parametra kada se koristi jedna putanja i „ekvivalentne“ putanje koja se koristi u slučaju *random* strategije disperzije. Naime, određuje se matrica stanja „ekvivalentne“ putanje koja predstavlja proizvod verovatnoće stanja pojedinačnih putanja s obzirom da se radi redundantna disperzija. U ovom slučaju, matrica stanja Γ „ekvivalentne“ putanje kada se Gilbert-Eliotov model primenjuje na pojedinačnim putanjama je ista kao i u (3.76).

Verovatnoća gubitaka u ovim novonastalim stanjima „ekvivalentne“ putanje je sada proizvod verovatnoće gubitaka paketa u pomenutim pojedinačnim stanjima. Sada je matrica kojom se određuju gubici paketa na „ekvivalentnoj“ putanji $B_{loss,equi}$ jednaka:

$$B_{loss,equi} = \begin{bmatrix} P_{G1} \cdot P_{G2} & 0 & 0 & 0 \\ 0 & P_{G1} \cdot P_{B2} & 0 & 0 \\ 0 & 0 & P_{B1} \cdot P_{G2} & 0 \\ 0 & 0 & 0 & P_{B1} \cdot P_{B2} \end{bmatrix} \quad (3.80)$$

Ova matrica za razliku od matrice kojom se određuje verovatnoća gubitaka paketa $B_{loss,q}$ kod *random* strategije se razlikuje po tome što su pojedinačni elementi matrice jednaki proizvodu kombinacija stanja kao posledica toga što se paket replicira na sve putanje, pa je i *de facto* verovatnoća gubitaka paketa kada se redundantna strategija primenjuje značajno manja. Naravno, efikasnost i detaljna analiza rastojanja gubitaka je pokazana u okviru dela sa rezultatima čime se može videti prednosti, a i mane ove strategije. Konačno, NLR parametar za redundantnu strategiju se određuje prema sledećem izrazu:

$$NLR^{(\delta)} = \pi \cdot B_{loss,equi} \cdot I - \pi \cdot (\Gamma \cdot B_{loss,equi} (\Gamma \cdot (I_d - B_{loss,equi})^\delta) \cdot I) \quad (3.81)$$

3.11. Analiza rastojanja gubitaka kod strategija u prisustvu gubitaka paketa sa Markovljevim modelom sa 4 stanja

Detaljne analize rafalnih gubitaka paketa koje se javljaju kod VoIP servisa su pokazale visok nivo multifraktalnosti VoIP saobraćaja kao i *long range dependency* (LRD) kod rafalnih gubitaka koji se javljaju kod VoIP servisa [46]. Dodatno, analiza [45] je pokazala da se optimalan odnos računске kompleksnosti i tačnosti modelovanja rafalnih gubitaka paketa dobija primenom Markovljevog modela sa 4 stanja. Trenutno u literaturi u kontekstu analize rastojanja gubitaka paketa nije urađena analiza u slučaju primene Markovljevog modela sa 4 stanja, čime se otvaraju dodatne mogućnosti verifikacije ovih rezultata kada se koristi tačnije modelovanje rafalnih gubitaka paketa. Doprinos ovog rada jeste predlog i analiza analitičkog modela određivanja NLR parametra kada se koristi tačniji Markovljev model sa 4 stanja.

Za pojedine strategije će se primeniti sličan pristup kao i kod Gilbert-Eliotovog modela imajući u vidu pravila koja važe kod Markovljevog modela sa 4 stanja. Takođe, za pojedine strategije potrebno je koristiti drugi način određivanja pojedinih matrica kako bi se konačno došlo do izraza za određivanje NLR kao posledica drugačije strukture Markovljevog modela.

3.11.1. NLR analiza kod round robin strategije

Round robin strategija predstavlja samo podskup periodične strategije koja se određuje periodičnom funkcijom $Q(i)$. Kako bi se koristio pristup koji je računski manje kompleksan $O(N \cdot K \cdot \delta)$ primenjuje se analiza kao i kod Gilbert-Eliotovog modela. U cilju preglednosti NLR analize kod Markovljevog modela

sa 4 stanja, pravila za kombinaciju događaja na putanjama kada se razmatra periodična strategija će biti ponovljena:

1. Na putanji P_k :
 - a. Dešava se gubitak paketa i , funkcija događaja je $L(i)=1$;
 - b. Nema gubitaka paketa na narednih δ paketa koji se prenose preko putanje P_k u skladu sa periodičnom funkcijom $Q(i)=k$ pri čemu $i \in \{1,2,\dots, \delta\}$;
 - c. Za sve ostale pakete posmatra se događaj $L(i)=2$.
2. Na putanji P_r :
 - a. Paket i predstavlja nevažan događaj za putanju P_r , tj. važi $L(i)=2$;
 - b. Nema gubitaka paketa na narednih δ paketa koji se prenose preko putanje P_r u skladu sa periodičnom funkcijom $Q(i)=r$ pri čemu $i \in \{1,2,\dots, \delta\}$;
 - c. Za sve ostale pakete posmatra se događaj $L(i)=2$.

U cilju računanja prethodno nabrojanih događaja i imajući u vidu koncept Markovljevog modela sa 4 stanja, potrebno je definisati novu tranzicionu matricu $T_{good,k}^{(k)}$ za putanju k :

$$T_{good,k}^{(k)}(i) = \begin{cases} T_{good,k}, & \text{za putanju } P_k \text{ u skladu sa } Q(i) = k \\ T_k, & \text{za putanju } P_k \text{ u skladu sa } Q(i) \neq k \end{cases} \quad (3.82)$$

gde je $T_{good,k}$ tranziciona matrica T_{good} za putanju P_k koja treba da uračuna pakete u skladu sa $Q(i)=k$, dok je matrica T_k matrica stanja za putanje u skladu sa $Q(i) \neq k$.

U tom slučaju, verovatnoća da se pojavi приметni gubitak $P_k[NL^\delta]$ kada je inicijalni paket izgubljen na putanji P_k se određuje kao razlika verovatnoće da je paket i izgubljen na putanji P_k i verovatnoće da izgubljeni paket nije приметан gubitak. Posledično, ova razlika upravo predstavlja verovatnoću pojavljivanja приметnih gubitaka paketa.

U skladu sa analizom u [94] dalje se razmatra *round robin* strategija disperzije kao validan predstavnik periodične strategije. Smatra se da funkcija $Q(i)$ odgovara *round robin* strategiji za $k=1$, $N=2$ i $K=2$. Ako se inicijalni gubitak paketa i desi na putanji P_1 verovatnoća da narednih δ paketa na putanji P_k ne budu приметni gubici je jednaka združenoj verovatnoći $P_1[L_{Q(i)=1}(i) = 1, L_{Q(i+1)=2}(i+1) = 2, \dots, L_{Q(i+\delta)=1}(i+\delta) = 0]$. Ova združena verovatnoća podrazumeva da su paketi $i+1, i+2, i+3, \dots, i+\delta$ na putanji P_1 ili uspešno primljeni ili su „nebitni“ (u skladu sa funkcijom događaja $L(i)=2$) kada je $Q(i) \neq 1$. Tako da je verovatnoća da se приметni gubici paketa ne javljaju na putanjama različitim od P_1 iznosi

$\prod_{r=1, r \neq k}^{N=2} P [L_{Q(i)=1}(i) = 1, L_{Q(i+1)=2}(i+1) = 2, \dots, L_{Q(i+\delta)=1}(i+\delta) = 0]$. Konačno, verovatnoća primetnog gubitka na putanji P_k za $k=1$ je jednaka:

$$P[NL^{(\delta)}] = P_1 [L_{Q(i)=1}(i) = 1] - (P_1 [L_{Q(i)=1}(i) = 1, L_{Q(i+1)=2}(i+1) = 2, \dots, L_{Q(i+\delta)=1}(i+\delta) = 0]) \cdot \prod_{r=1, r \neq k}^{N=2} P [L_{Q(i)=1}(i) = 1, L_{Q(i+1)=2}(i+1) = 2, \dots, L_{Q(i+\delta)=1}(i+\delta) = 0]) \quad (3.83)$$

Generalizacijom gornjeg izraza i stavljanjem istog u matričnu formu, verovatnoća da je paket i primetni gubitak paketa:

$$P_k[NL^{(\delta)}] = P_{loss,k} - \left(S_k \cdot T_{loss,k} \cdot \prod_{l=1}^{\delta} T_{good,k}^{(k)}(i+l) \cdot I \right) \cdot \prod_{r=1, r \neq k}^N (S_r \cdot \left(\prod_{l=1}^{\delta} T_{good,r}^{(r)}(i+l) \right) \cdot I) \quad (3.84)$$

gde je $P_{loss,k}$, $T_{loss,k}$, S_k PLR, tranziciona matrica koja sadrži verovatnoće prelaska u loša stanja i matrica stanja za putanju P_k , respektivno. Prethodni izraz podrazumeva samo kada je inicijalni gubitak na jednoj putanji pri čemu je potrebno da se uzmu u obzir sve mogućnosti, tačnije putanje na kojima se može javiti pomenuti inicijalni gubitak paketa. Svaka raspoloživa putanja ima podjednaku verovatnoću za pojavljivanje inicijalnog gubitka paketa, pa samim tim se izraz za određivanje ukupnog NLR parametra za parametar δ određuje kao srednja vrednost u skladu sa sledećim izrazom:

$$NLR^{(\delta)} = \frac{1}{K} \sum_{i=1}^K \left\{ \left(P_{loss,Q(i)} - \left(S_{Q(i)} \cdot T_{loss,Q(i)} \cdot \prod_{l=1}^{\delta} T_{good,Q(i)}^{(Q(i))}(i+l) \right) \cdot I \right) \cdot \prod_{r=1, r \neq Q(i)}^N (S_r \cdot \left(\prod_{l=1}^{\delta} T_{good,r}^{(r)}(i+l) \right) \cdot I) \right\} \quad (3.85)$$

3.11.2. NLR analiza kod random i adaptivne strategije

Random disperzija podrazumeva da svaka putanja koja je namenjena za proces disperzije paketa ima dodeljenu verovatnoću φ kojom se određuje verovatnoća da naredni paket bude poslat preko posmatrane putanje. Za prethodno navedeni parametar, s obzirom da se radi o verovatnoći, važi jednakost $\sum_{i=1}^N \varphi_i$. Kao i u slučaju *random* strategije kada je analiziran Gilbert-Eliotov model gubitaka paketa, tako i za Markovljev model sa 4 stanja se koristi pristup sa „ekvivalentnom“ putanjom.

U slučaju *random* strategije, koncept „ekvivalentne“ putanje se ne razlikuje od Gilbert-Eliotovog modela i podrazumeva jednu putanju na kojoj su izražene karakteristike rafalnih gubitaka paketa koje predstavljaju kombinacije rafalnih gubitaka izraženih na pojedinačnim putanjama sa odgovarajućim verovatnoćama. U tom smislu, matrica stanja „ekvivalentne“ putanje S_{equi} se sastoji od kombinacija stanja na pojedinačnim

putanjama, dok tranziciona matrica T_{equi} predstavlja kombinaciju tranzicionih verovatnoća koje potiču od pojedinačnih putanja. Matrice S_{equi} i T_{equi} su dimenzija $1 \times N$ i $N \times N$, respektivno.

Imajući u vidu prethodno analizirani pristup „ekvivalentne“ putanje, Kronekerov proizvod je primenljiv u ovom slučaju čime se iz matrica stanja i tranzicionih matrica dobija specifikacija rafalnih gubitaka koji su izraženi na „ekvivalentnoj“ putanji kada se primenjuje *random* disperzija. Posledično, za matricu stanja važi da je $S_{equi} = S_1 \otimes S_2 \otimes S_3 \dots \otimes S_N$ i za tranzicionu matricu važi $T_{equi} = T_1 \otimes T_2 \otimes T_3 \dots \otimes T_N$.

Određivanjem prethodnih matrica stanja i tranzicije za „ekvivalentnu“ putanju nije uračunata verovatnoća gubitaka paketa kao ni parametar pridružene verovatnoće putanjama disperzije, pa je samim tim potrebno uvesti dodatni parametar koji će regulisati prethodno navedeno. Naime, potrebno je uvesti matricu indikacije gubitaka A_{loss} kako bi se uračunali gubici paketa kao i parametar φ . Ova matrica je dimenzija $4^N \times 4^N$ i praktično na glavnoj dijagonali sadrži verovatnoću gubitaka paketa za sve kombinacije stanja u kojoj se „ekvivalentna“ putanja može naći, što naravno, zavisi od verovatnoće stanja i u okviru pojedinačnih putanja disperzije. Primer matrice A_{loss} za 2 putanje i pridružene verovatnoće φ_1 i φ_2 je dato u sledećoj jednakosti:

$$A_{loss} = \begin{bmatrix} \varphi_1 + \varphi_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \varphi_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \varphi_1 + \varphi_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \varphi_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \varphi_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \varphi_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \varphi_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \varphi_1 + \varphi_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \varphi_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \varphi_1 + \varphi_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \varphi_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \varphi_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \varphi_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \varphi_2 & 0 & 0 \end{bmatrix} \quad (3.86)$$

Elementi na glavnoj dijagonali matrice A_{loss} se određuju prema sledećem pravilu:

- Za kombinaciju stanja kada se obe putanje nalaze u „lošem“ stanju, verovatnoća gubitaka je $\varphi_1 + \varphi_2$;
- Za kombinaciju stanja kada se prva putanja nalazi u „lošem“, a druga u „dobrom“ stanju, verovatnoća gubitaka je φ_1 . Takođe, važi suprotno za drugu putanju;
- Za kombinaciju stanja kada se obe putanje nalaze u „dobrom“ stanju, verovatnoća gubitaka je 0.

U skladu sa prethodnim, matrica koja određuje verovatnoću uspešno primljenih paketa je $A_{no-loss} = I_d - A_{loss}$, gde je I_d identitet matrica, tačnije matrica koja na glavnoj dijagonali ima sve elemente koji su jednaki 1. Kada su određene sve potrebne matrice, može se pristupiti određivanju NLR parametra za „ekvivalentnu“ putanju, odnosno za *random* disperziju u skladu sa sledećim izrazom:

$$NLR^{(\delta)} = P_{loss,equi} - (S_{equi} \cdot (T_{equi} \cdot A_{loss}) \cdot (T_{equi} \cdot A_{no-loss})^\delta \times I) \quad (3.87)$$

gde je I matrica dimenzija $4^N \times 1$ čiji su svi elementi jednaki 1, dok je $P_{loss,equi} = \sum_{i=1}^N \varphi_i \cdot P_{loss,i}$ PLR parametar izražen na “ekvivalentnoj” putanji.

Adaptivna strategija predstavlja specijalni slučaj *random* strategija sa tom razlikom što se kod nje verovatnoće koje su pridružene putanjama raspoloživim za disperziju određuju na osnovu *a priori* saznanja o karakteristikama rafalnih gubitaka na istima. U tom smislu, potrebno je imati predznanje o rafalnim gubicima koji se javljaju ili je moguće implementirati na dinamički način mehanizam koji bi mogao da određuje trenutni nivo rafalnosti gubitaka na više putanja i u skladu sa vrednostima određuje i pomenute verovatnoće. Kao i kod Gilbert-Eliotovog modela, verovatnoće φ_i se određuju prema relaciji (3.68).

3.11.3. NLR analiza kod redundantne strategije

Repliciranje paketa preko raspoloživih putanja podrazumeva da je sada za prenos posmatranog saobraćaja potrebno više kapaciteta, tačnije onoliko puta više koliko ima putanja preko kojih se realizuje pomenuta replikacija paketa. Međutim, implementacijom redundantne strategije omogućava se da prenos paketa preko putanja na kojima su izraženi izuzetno visoki gubici rafalnog oblika. Analiza rastojanja gubitaka paketa prilikom primene redundantne disperzije je razmatrana u [94].

Gubitak paketa u slučaju redundantne strategije preko N putanja se događa kada svaka raspoloživa putanja u tom trenutku prelazi u „loše“ stanje. Kao i kod Gilbert-Eliotovog modela i ovde se koristi primena „ekvivalentne“ putanje. U ovom slučaju koristi se matrica stanja S_{equi} i tranziciona matrica T_{equi} . Ovde se takođe koristi Kronekerov proizvod kako bi se dobile matrice stanja i tranziciona matrica „ekvivalentne“ putanje, pa NLR određivanje se realizuje u skladu sa izrazom:

$$NLR^{(\delta)} = S_{equi} \cdot (T_{loss,1} \otimes T_{loss,2} \dots \otimes T_{loss,N}) \cdot I - \left[S_{equi} \cdot (T_{loss,1} \otimes T_{loss,2} \dots \otimes T_{loss,N}) \cdot (T_{good,1} \otimes T_{good,2} \dots \otimes T_{good,N})^\delta \right] \cdot I \quad (3.88)$$

gde su $T_{loss,i}$ i $T_{good,i}$ tranzicione matrice putanje i koje sadrže tranzicije u “loše” i “dobro” stanje, respektivno.

3.12. Uticaj *BurstR* parametra na strategije disperzije

BurstR predstavlja jedan od mnogobrojnih ulaznih parametara ITU-T E-Modela koji je najrasprostranjeniji u smislu objektivne ocene kvaliteta VoIP servisa. U kontekstu promene strukture rafalnih gubitaka primenom strategija paketske disperzije, osim analize rastojanja gubitaka kada se posmatra NLR parametar, potrebno je i razmatrati MOS ocenu koju VoIP dobija za ostvarene QoS parametre kao što su kašnjenje, džiter i gubici paketa. S obzirom da gubici paketa dominantno utiču na kvalitet VoIP servisa, bitno je utvrditi kako struktura rafalnih gubitaka utiče na objektivnu ocenu kvaliteta. *BurstR* parametar uzima u obzir strukturu rafalnih gubitaka, pa je samim tim bitno utvrditi kako se za pojedinačne strategije disperzije utvrđuje pomenuti parametar. Analiza *BurstR* parametra za Markovljev model sa 4 stanja je realizovana u [94].

U skladu sa definicijom iz [62] *BurstR* parametar se određuje prema sledećoj relaciji:

$$BurstR = \frac{P_{loss}}{p_{g-b}} \quad (3.89)$$

gde je p_{g-b} verovatnoća tranzicije iz „dobrih“ stanja (S_2 i S_4) u „loša“ stanja (S_1 i S_3). Jasno je da bi se odredio *BurstR*, potrebno je najpre odrediti p_{g-b} tranzicionu verovatnoću. Prethodno se određuje verovatnoća P_{g-b} koja proističe da se sistem nalazi u lošem stanju kao posledica tranzicije iz „dobrog“ u „loše“ stanje i ova verovatnoća kod periodične strategije je jednaka:

$$P_{g-b} = \frac{1}{K} \sum_{i=1}^K S_{good,Q(i)} \cdot T_{g-b,Q(i)} \cdot I \quad (3.90)$$

gde je $S_{good,Q(i)}$ matrica stanja koja sadrži samo „dobra“ stanja, $T_{g-b,Q(i)}$ tranziciona matrica koja sadrži samo tranzicione verovatnoće koje podrazumevaju tranziciju iz S_2 i S_4 u S_1 i S_3 . Nakon toga, p_{g-b} se određuje prema sledećem izrazu:

$$p_{g-b} = \frac{P_{g-b}}{1 - P_{loss}} \quad (3.91)$$

Određivanje verovatnoće P_{g-b} kod *random* strategije se realizuje kao:

$$P_{g-b} = S_{good,equi} \cdot T_{g-b,equi} \cdot A_{g-b,equi} \cdot I \quad (3.92)$$

gde je $S_{good,equi} = S_{good,1} \otimes S_{good,2} \otimes \dots \otimes S_{good,N}$, dok $T_{g-b,equi} = T_{g-b,1} \otimes T_{g-b,2} \otimes \dots \otimes T_{g-b,N}$. Rafalna matrica $A_{g-b,equi}$ sadrži elemente koji predstavljaju tranzicije iz „dobrog“ u „loše“ stanje na „ekvivalentnoj“ putanji, dok su ostali elementi matrice jednaki 0. Kako bi se „ekvivalentna“ matrica našla

u „lošem“ stanju, makar jedna putanja mora biti u stanju S_1 ili S_3 . Slično važi i za „dobro“ stanje „ekvivalentne“ putanje kada se makar jedna putanja mora nalaziti u „dobrom“ stanju S_2 ili S_4 . Takođe, nenulti elementi se nalaze na odgovarajućim pozicijama u matrici koje odgovaraju gore navedenim pretpostavkama. Ti nenulti elementi se određuju kao zbir φ_k verovatnoća, ako i samo ako putanja P_k prelazi iz „dobrog“ u „loše“ stanje. Nakon određene vrednosti P_{g-b} , p_{g-b} se određuje prema relaciji (3.91). Na primer, za putanju 1 i 2, i za kombinaciju istovremene tranzicije u loše stanje, element u rafalnoj matrici bi bio $\varphi_1 + \varphi_2$, dok ukoliko putanja 1 prelazi iz dobrog u loše, a putanja 2 ostaje u lošem stanju, odgovarajući element bi bio jednak samo φ_1 .

Određivanje verovatnoće P_{g-b} kod redundantne strategije se realizuje kao:

$$P_{g-b} = S_{good, equi} \cdot T_{g-b, equi} \cdot I \quad (3.93)$$

gde je I matrica sa elementima jednakim 1 dimenzija $4^N \times 1$.

3.13. Dinamička aktivacija paketske disperzije

Paketska disperzija podrazumeva modifikaciju politike rutiranja kojom se može postići promena strukture rafalnosti gubitaka paketa koji se javljaju između izvorišta i odredišta. Međutim, postoji nekoliko nedostataka koji se javljaju kada se primenjuje paketska disperzija. Naime, odvojene putanje koje se koriste za disperziju paketa često mogu da imaju različito kašnjenje i džiter što potencijalno može da dovede do ostvarivanja i lošijih performansi u odnosu na konvencionalno rutiranje preko jedne putanje. Dodatno, čest je zahtev da bude tačan redosled paketa na odredištu što razlika u kašnjenju i razlika u performansama pojedinačnih putanja može da poremeti. Na primer, pojedini paketi koji su se prenosili putanjama sa visokim kašnjenjem mogu kasno da stignu na odredište i da budu posledično odbačeni.

Problemi sa primenom paketske disperzije mogu najviše uticati na saobraćaj QoS kritičnih aplikacija, konkretno na multimedijalne aplikacije. Promena performansi na putanjama može biti u određenoj meri rešena implementacijom *play-out* bafera kojim se unosi dodatno kašnjenje, ali se omogućava konsantan tok paketa u okviru odredišta. Načelno, pojedine strategije paketske disperzije mogu dovesti do manjih ili većih problema sa performansama u smislu odbačenih paketa kao posledica uticaja različitog kašnjenja i džitera na putanjama.

Osim problema sa QoS, paketska disperzija potencijalno može da dovede do povećanog opterećenja kapaciteta pojedinih putanji što dalje može uticati na ostale klase saobraćaja i dovesti do novih problema koje paketska disperzija namenjena inicijalno da reši. Međutim, ovaj problem nije kritičan s obzirom da u današnjim mrežama kapacitet se prilično dobro planira gde u određenim slučajevima izuzetno jaki rafali saobraćaja mogu da dovedu do kratkotrajnih povećanja opterećenja i u ekstremnim slučajevima gubitaka paketa kao posledica ispunjenja kapaciteta ruterske bafer memorije.

U cilju poboljšanja efikasnosti strategija paketske disperzije i njenog neposrednog uticaja na performanse, konkretno na izražene rafalne gubitke paketa, u ovom radu se razmatra primena dinamičke aktivacije paketske disperzije. Koncept dinamičke aktivacije paketske disperzije podrazumeva da bi se ista koristila u slučajevima pojave rafalnih gubitaka paketa kojima bi se bavile specificirane strategije disperzije imajući u vidu da u prisustvu gubitaka paketa za QoS kritične servise (VoIP) je osetan pad performansi u odnosu na efekat ostalih QoS parametara, na primer kašnjenje i džiter [95]. Dalje se predlaže implementacija dinamičke aktivacije paketske disperzije što potencijalno predstavlja način na koji se mogu prevazići inherentni nedostaci procesa disperzije paketa. Aktivacija paketske disperzije predložena u [96] koristi sintetičke pakete koji vrše merenja na osnovu kojih se određuje kada je disperziju potrebno aktivirati. Ovi sintetički paketi mere gubitke paketa i na osnovu izraženih gubitaka paketa aktiviraju disperziju preko određenog broja putanja. Međutim, ovakva implementacija aktivacije disperzije posredstvom sintetičkih paketa predstavlja nepouzdanu rešenje jer je prethodno potrebno utvrditi adekvatan interval merenja i upotrebnu vrednost mernih rezultata u službi aktivacije disperzije. U ovom radu se predlaže mehanizam dinamičke aktivacije u okviru koga se potreba za paketskom disperzijom utvrđuje na osnovu zauzetosti reda čekanja visokog prioriteta. Istovremeno, u skladu sa predloženim mehanizmom se utvrđuje i broj raspoloživih putanja posredstvom kojih se vrši disperzija paketa na osnovu čega se može predvideti zagušenost i performanse na posmatranim putanjama.

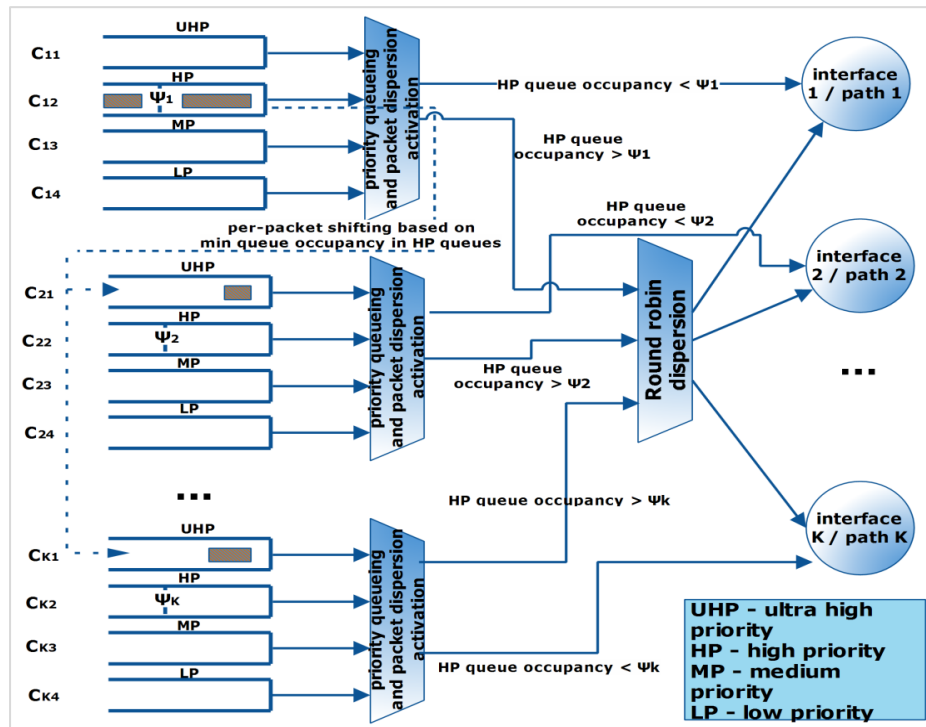
3.13.1. Predložena šema aktivacije paketske disperzije

Predloženi model aktivacije paketske disperzije je zasnovan na prebacivanju paketa iz jednog odlaznog reda interfejsa u jedan ili više redova najviših prioriteta koji su pridruženi interfejsima određenih da budu korišćeni kao disperzione putanje [97]. Pomenute disperzione putanje su određene kao validne do specificiranog odredišta. Kao integralni element predložene aktivacije disperzije uvodi se red sa najvišim prioritetom koji se koristi tokom procesa paketske disperzije.

Pretpostavlja se model koji se sastoji od K putanja do specificiranog odredišta. Putanje su sa stanovišta rutera označene kao K odlaznih interfejsa imajući u vidu da je preduslov efikasnog procesa paketske disperzije fizička separacija putanja od mesta aktivacije paketske disperzije, pa do posmatranog odredišta. U skladu sa predloženim modelom, svaki od ovih interfejsa poseduje odlazne redove čekanja paketa i ti redovi su klasifikovani u skladu sa prioritetima (ultra-visoki, visoki, srednji i niski prioriteti) pri čemu za svaki od njih primenjuje *priority queueing* kao mehanizam upravljanja redovima. Model je prikazan na slici 3.21.

U cilju analize pretpostavlja se da se zagušenje dešava samo u redu visokog prioriteta na putanji 1. Paketi koji su klasifikovani u skladu sa QoS politikom kao saobraćaj visokog prioriteta se smeštaju u red visokog prioriteta. Paketi u tom redu su normalno opsluživani od strane odgovarajućeg odlaznog interfejsa 1 dok god zauzetost reda visokog prioriteta ne dostigne prag Ψ_1 . Prethodno navedeni proces predstavlja normalno

ponašanje u slučaju da ne postoji zagušenje. Putanja 1 za odgovarajuće pakete se smatra najboljom putanjom do posmatranog odredišta iako postoji još $K-1$ dodatnih putanja koje su određene da su sposobne da ispune QoS zahteve visoko prioritetnog saobraćaja. Suštinski, u odsustvu zagušenja realizuje se konvencionalno rutiranje preko jedne putanje prethodno određene kao najbolje u skladu sa politikom rutiranja.



Slika 3.21 Predloženi model aktivacije paketske disperzije.

Kada zauzetost reda visokog prioriteta pređe prag ψ_i paket koji se nalaze na čelu reda visokog prioriteta se prosleđuju na kraj redova ultra visokih prioriteta koji su pridruženi odlaznim interfejsima i koji označavaju QoS putanje do specificiranog odredišta u paketima. Predloženi model po svakom paketu donosi odluku o prosleđivanju paketa u druge redove ultra visokog prioriteta. Posledično, predloženi model radi na *per-packet* principu za razliku od *per-flow* mehanizama koji svaki *flow* saobraćaja tretiraju različito što može da dovede do poremećaja performansi pojedinačnih *flow*-ova [98]. Prosleđeni paketi su smešteni u neki od $K-1$ dodatnih redova ultra visokih prioriteta pridruženih odlaznim interfejsima čiji redovi visokog prioriteta treba da imaju minimalnu zauzetost. Pomenuto prosleđivanje paketa se realizuje implementacijom *switching* mehanizma koji izvršiti potrebno prosleđivanje iz redova visokih prioriteta u redove ultra visokih prioriteta.

Zauzetost redova visokih prioriteta je podložna dinamičkoj promeni u zavisnosti od saobraćaja visokog prioriteta, pa samim tim se može desiti da više pojedinačnih redova ultra visokog prioriteta mogu da prihvate prosleđene pakete iz jednog posmatranog reda visokog prioriteta. To podrazumeva da će određena putanja učestvovati u procesu disperzije dok god postoje paketi u okviru reda ultra visokog prioriteta za tu

putanju. Na taj način se postiže da se više putanja koristi za prosleđivanje paketa do odredišta. Adekvatan aspekt praćenja i upravljanja resursima rutera je opisan u [99] koji bi se efikasno mogao primeniti za praćenje zauzetosti različitih redova u okviru rutera koji implementira dinamičku aktivaciju paketske disperzije.

Zagušenje i eventualno odbacivanje paketa je smanjeno kao posledica implementacije prosleđivanja paketa između redova. Kada se realizuje paketsko prosleđivanje, aktivira se paketska disperzija i paketi prestaju da se šalju direktno na odlazni interfejs, već preko više putanja čiji broj zavisi od broja redova ultra visokog prioriteta koji su zauzeti paketima koji su premešteni iz redova visokih prioriteta. *Priority queueing* se koristi za upravljanje opsluživanja predovima i na osnovu ovog mehanizma se najpre opslužuju redovi ultra visokog prioriteta pre ostalih redova tog odlaznog interfejsa. Kada ne postoji više paketa u redovima ultra visokog prioriteta, ta putanja se uklanja iz procesa paketske disperzije i ne koristi se dok god je pomenuti red prazan.

Kada se desi promena u zauzetosti redova visokog prioriteta, u red ultra visokog prioriteta se sada prosleđuju paketi čime se dodatno aktivira ta putanja koja *de facto* sada učestvuje u procesu paketske disperzije. U predloženom mehanizmu aktivacije se primenjuje *round robin* šema kao politika prosleđivanja paketa. Paketi koji se prosleđuju iz odgovarajućeg reda visokog prioriteta kao i paketi koji se nalaze u redovima ultra visokog prioriteta se opslužuju u *round robin* režimu.

S obzirom da su redovi ultra visokog prioriteta ograničeni po kapacitetu (kako bi se sprečilo dodatno zagušenje redova manjeg prioriteta usled *priority queueing* mehanizma – *queue starvation*) moguće je da nestane kapaciteta koji je potreban za prosleđivanje paketa. U tom slučaju, paketi će nastaviti da pune red visokog prioriteta i iznad praga Ψ_1 do trenutka kada će kapacitet ovog reda biti dostignut. Nakon toga, ako se protok paketa ne smanji ili ukoliko se neki red ultra visokog prioriteta ne oslobodi krenuće se sa odbacivanjem paketa. Verovatnoća da će se ovo odbacivanje desiti će biti manje ukoliko postoji više redova ultra visokog prioriteta, odnosno putanja koji će biti raspoloživi za proces paketske disperzije. Adaptivno i dinamičko aktiviranje procesa paketske disperzije pomaže u problemu pojave zagušenja saobraćaja visokog prioriteta pri čemu se primena ovog mehanizma odnosi i na rešavanje problema pojave rafalnih gubitaka paketa na putanjama do odredišta. Dodatno, prethodno pomenuti inherentni problemi sa primenom paketske disperzije su svedeni na minimum s obzirom da se koristi dinamički broj putanja u zavisnosti od nivoa izraženog zagušenja, a samim tim i pojavom rafalnih gubitaka paketa.

Potrebno je naglasiti da je za efikasnost mehanizma paketske disperzije potrebno odrediti adekvatan kapacitet pojedinačnih redova koji su ključni sastojak u funkcionisanju dinamičke aktivacije disperzije [100], [101], [102], [103]. Dodatno je potrebno odrediti adekvatnu vrednost praga kako bi se na vreme realizovala aktivacija paketske disperzije. Određivanje kapaciteta pojedinačnih redova je potrebno kako bi

se izbeglo *tail drop* odbacivanje paketa. Suštinski, na osnovu rada predloženog mehanizma potrebno je da se obezbedi veći kapacitet redova sa visokim prioritetom u odnosu na redove ultra visokog prioriteta.

3.13.2. Analiza aktivacije paketske disperzije

Imajući u vidu da se aktivacija realizuje kroz zauzetost redova čekanja paketa u ruterima, analiza će biti sprovedena upotrebom teorije redova. Svaki red u predloženom modelu pre aktivacije paketske disperzije se može posmatrati kao konačan M/M/1/C red sa kapacitetom C , Poasonovom raspodelom dolaska paketa, eksponencijalnim vremenom opsluživanja i *First Come First Served* (FCFS) disciplinom opsluživanja. Svaki red u modelu ima kapacitet C_{ij} gde je i indeks koji označava pridruženi odlazni interfejs, dok j indeks označava konkretan prioritet posmatranog reda ($j=1$ je ultra visoki, $j=2$ is visoki, itd.). Za potrebe analize će biti posmatran red visokog prioriteta za interfejs 1 odnosno putanju 1 . Dalje se pretpostavlja da samo paket visokog prioriteta dostiže prag aktivacije paketske disperzije, dok drugi redovi ovog prioriteta ne dostižu ovaj prag zauzetosti redova.

Pretpostavljajući da je λ brzina dolaska paketa u red visokog prioriteta, μ brzina opsluživanja, dok je $\rho=\lambda/\mu$ intenzitet saobraćaja. Pre nego što zauzetost reda visokog prioriteta koji pripada interfejsu 1 dostigne prag Ψ analiza tog reda se svodi na M/M/1/ Ψ sa kapacitetom Ψ . U skladu sa teorijom redova za konačni red M/M/1/ Ψ verovatnoća da će zauzetost reda biti na pragu se određuje prema izrazu:

$$P_{\psi} = \rho^{\psi} \cdot \frac{1-\rho}{1-\rho^{\psi+1}} \quad (3.94)$$

Dok je verovatnoća da je red visokog prioriteta prazan:

$$P_0 = \rho^0 \cdot \frac{1-\rho}{1-\rho^{\psi+1}} \quad (3.95)$$

Srednji broj paketa u redu, uključujući i opslužene paketet kao i pakete koji se nalaze u redu se određuju prema sledećim izrazima, respektivno:

$$L_{s,\psi} = \sum_{n=1}^{\psi} n \cdot P_n = \frac{\rho \cdot (1-\rho^{\psi(\psi+1)+\psi\rho^{\psi+1}})}{(1-\rho) \cdot (1-\rho^{\psi+1})} \quad (3.96)$$

$$L_{q,\psi} = \sum_{n=1}^{\psi} (n-1) \cdot P_n = L_{s,\psi} - (1-P_0) \quad (3.97)$$

Srednje vreme čekanja paketa u redu visokog prioriteta pre aktivacije paketske disperzije je jednako:

$$W_{q,\psi} = \frac{L_{q,\psi}}{\lambda_{\psi}} = \frac{L_{q,\psi}}{(1-P_{\psi})\lambda} \quad (3.98)$$

gde je λ_ψ efektivna brzina dolaska paketa pre aktivacije paketske disperzije.

Kada zauzetost reda visokog prioriteta dostigne prag ψ brzina dolaska paketa će biti promenjena kao posledica paketa koji se prosleđuju sa početka reda prema redovima ultra visokog prioriteta koji pripadaju odlaznom interfejsu čiji redovi visokog prioriteta imaju minimalnu zauzetost. Kada se aktivira process paketske disperzije *round robin* opsluživanje redova se primenjuje na svakom redu, tj. na redu visokog prioriteta zbog koga je aktivirana disperzija i na ultra visokim redovima koji sadrže pakete koji su prethodno bili u pomenutom redu visokog prioriteta. *Round robin* opsluživanje redova se modeluje upotrebom M/G/1/C reda koji se opslužuje *processor sharing* disciplinom [104]. Brzina opsluživanja $\mu_{dispersion}$ u ovom slučaju zavisi od distribucije veličine paketa i kapaciteta interfejsa koji opslužuje ovaj red. S obzirom da interfejsi mogu imati različite kapacitete uzastopni paketi iz jednog reda mogu biti opsluženi od strane interfejsa sa različitim kapacitetom, pa se posledično uzima u obzir srednji kapacitet na koji redovi mogu da računaju. U ovom slučaju je brzina opsluživanja kada je paketska disperzija aktivirana

$$\mu_{dispersion} = \frac{1}{\frac{1}{K} \sum_{k=1}^K \frac{L_k}{B}} \quad (3.99)$$

gde je L_k veličina paketa sa opštom distribucijom i $B = \frac{1}{K} \sum_{k=1}^K B_i$ je srednji kapacitet interfejsa koji se koriste za process paketske disperzije. Ova aproksimacija je dovoljno tačna da se modeluje *round robin* opsluživanja redova kada se red opslužuje interfejsima različitih kapaciteta.

Brzina dolaska paketa pre aktivacije paketske disperzije je prikazana na slici 3.22(a). Ukoliko je zauzetost reda visokog prioriteta visa od praga ψ , paketi će biti prosleđeni prema redovima ultra visokih prioriteta sa brzinom λP_ψ dok će dolazak paketa u pojedinačne redove zavisiti od zauzetosti respektivnih redova visokog prioriteta. Posledično, dolazak prosleđenih paketa u red ultra visokog prioriteta se određuje prema izrazu:

$$\lambda_{i1} = \lambda \cdot P_\psi \cdot P_{i1} \quad (3.100)$$

gde je λ_{i1} brzina dolaska u red ultra visokog prioriteta, dok je P_{i1} verovatnoća da odgovarajući red visokog prioriteta ima minimalnu zauzetost u odnosu na ostale redove istog prioriteta u ruteru. Ovaj scenario je prikazan na slici 3.22(b).

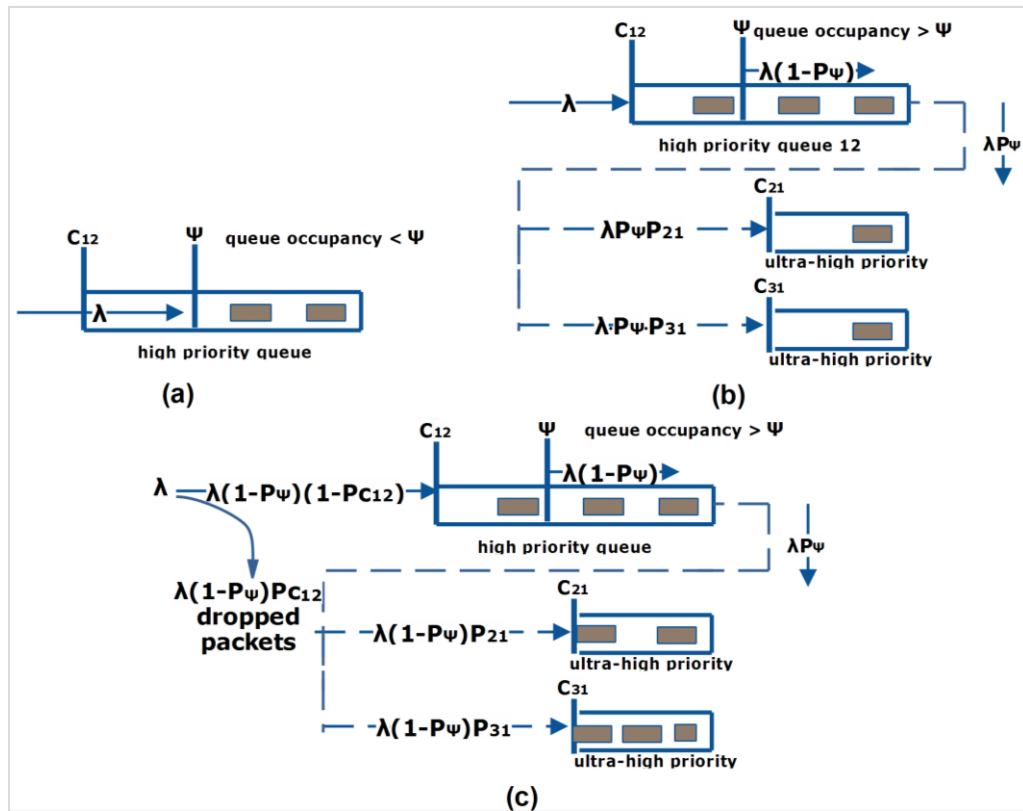
Srednje vreme čekanja u redu paketa koji su prosleđeni se sastoji od dve komponente – srednjeg vremena utrođenog u redu visokog prioriteta pre prosleđivanja i srednjeg vremena utrošenog u čekanju u okviru ultra visokog prioriteta pre opsluživanja. Intenzitet saobraćaja u redu visokog prioriteta koji aktivira paketsku disperziju je $\rho_{i2} = \lambda_\psi / \mu_{dispersion}$ i intenzitet saobraćaja u okviru reda ultra visokog prioriteta $\rho_{i1} =$

$\lambda_{i1}/\mu_{dispersion}$. Samim tim, ukupno vreme čekanja paketa koji su servisirani u okviru reda visokog prioriteta $W_{q,Ci2,dispersion}$ (paketa koji nisu prosleđeni i koji su ostali u tom redu) i paketa koji su prosleđeni prema redu ultra visokog prioriteta $W_{q,Ci1,shifted}$ je respektivno dato narednim izrazima:

$$W_{q,Ci2,dispersion} = \frac{L_{q,Ci2,dispersion}}{\lambda_{\psi}(1-P_{Ci2})} \quad (3.101)$$

$$W_{q,Ci1,shifted} = \frac{L_{q,Ci2,dispersion}}{\lambda_{\psi}(1-P_{Ci2})} + \frac{L_{q,Ci1,shifted}}{\lambda_{i1}(1-P_{Ci1})} \quad (3.102)$$

gde su $L_{q,Ci2,dispersion}$ i $L_{q,Ci1,shifted}$ dužine redova čekanja za relaciju datu u (3.96) kada je intenzitet saobraćaja ρ_{i2} i ρ_{i1} , respektivno. Verovatnoće P_{Ci1} i P_{Ci2} označavaju da su redovi $i1$ i $i2$ dostigli kapacitet. Predloženi mehanizam aktivacije paketske disperzije podrazumeva da se prosleđivanje paketa ne realizuje prema ultra visokim redovima koji su dostigli kapacitet, već se interfejs kome je pridružen taj red izbacuje iz procesa paketske disperzije.



Slika 3.22. Brzina dolaska paketa u respektivnim redovima: (a) zauzetost reda visokog prioriteta < Ψ ; (b) zauzetost reda visokog prioriteta > Ψ ; (c) zauzetost reda visokog prioriteta > Ψ i odbacivanje paketa.

Ostaje da se analizira slučaj kada se dešava odbacivanje paketa usled dostizanja kapaciteta odgovarajućeg reda. Odbacivanje paketa u okviru reda visokog prioriteta se događa kao posledica dostizanja kapaciteta svih raspoloživih redova ultra visokog kapaciteta. Posledično, red visokog prioriteta se puni do svog

kapaciteta kada se vrši odbacivanje paketa što je prikazano na slici 3.22(c). Verovatnoća da će doći do odbacivanja u redu visokog prioriteta se određuje kao

$$P_{C_{i2}} = \rho_{i2}^{C_{i2}} \cdot \frac{1 - \rho_{i2}}{1 - \rho_{i2}^{C_{i2}+1}} \quad (3.103)$$

gde je C_{i2} kapacitet reda visokog prioriteta koji je inicirao paketsku disperziju. Dodatno, brzina odbacivanja paketa λ_d se definiše prema sledećem izrazu:

$$\lambda_d = \lambda_\psi \cdot P_{C_{i2}} = \lambda \cdot (1 - P_\psi) \cdot P_{C_{i2}} \quad (3.104)$$

Prethodna analiza je podrazumevala da je samo jedan red visokog prioriteta inicirao paketsku disperziju. Takođe, predloženi model implementira *priority queueing* mehanizam opsluživanja redova prema kome se određuje brzina opsluživanja redova:

$$\mu_{eff} = \begin{cases} \mu \cdot P_{UHP0}, & \text{zauzetost reda} < \psi \\ \mu_{dispersion} \cdot P_{UHP0}, & \text{zauzetost reda} < \psi \end{cases} \quad (3.105)$$

gde je P_{UHP0} verovatnoća da respektivni red ultra visokog prioriteta bude prazan.

Dok god postoje paketi u nekoliko redova ultra visokog prioriteta više putanja će biti uključeno u okviru paketske disperzije i sam proces disperzije će biti efikasniji u borbi protiv uticaja rafalnih gubitaka paketa. Međutim, postoji i inherentni problem disperzije koji se tiče kašnjenja i džitera i koji se mora minimizovati. Kako bi se sprečila pojava preteranog opsluživanja redova ultra visokog prioriteta, kapacitet redova ultra visokog je manji u odnosu na redove visokog prioriteta, pa je kašnjenje i džiter koji proističu iz prosleđivanja paketa minimalan. Uticaj na rafalane gubitke kao i pregled dinamičke aktivacije paketske disperzije po pitanju broja uključenih putanja će biti pokazani u okviru dela sa rezultatima.

3.13.3. Aktivacije paketske disperzije posredstvom zauzetosti bafera

Alternativa prethodnom mehanizmu aktivacije paketske disperzije podrazumeva da se posmatra zauzetost kapaciteta bafera s obzirom da će u okviru ruterskog bafera biti više paketa kada je odgovarajuća putanja zagušena. Posledično, prag zauzetosti baferskog kapaciteta [98] predstavlja adekvatan način da se aktivira paketska disperzija, tačnije da se aktiviraju disperzione putanje. Takođe, u zavisnosti od nivoa zauzetosti bafera može se odrediti koliko putanja treba da se aktivira. Ovakav pristup dinamičkog određivanja broja putanji koje će biti deo procesa paketske disperzije se direktno suprostavlja inherentnim problemima paketske disperzije kao što je promenljivo kašnjenje i varijacija kašnjenja na različitim putanjama.

Za potrebe aktivacije paketske disperzije pretpostaviće se da je C kapacitet bafera u okviru rutera koji inicira paketsku disperziju. Predloženi mehanizam aktivacije disperzije podrazumeva da se definišu pragovi

bafera tako što će se baferska memorija podeliti na N delova koliko i ima raspoloživih putanja. Svaki put kada kapacitet bafera dostigne odgovarajući deo, određeni broj putanja se aktivira i postaje deo procesa paketske disperzije. Baferski prag kojim se aktivira k putanja se određuje prema jednačini:

$$W_k = (k - 1) \cdot \frac{C}{N} \quad (3.106)$$

gde je W_k prag kapaciteta bafera nakon čega će se aktivirati k putanja, dok se smatra da je putanja označena kao $k=1$ najbolja putanja koja se koristi za konvencionalno rutiranje pre aktivacije paketske disperzije. Prethodna formulacija podrazumeva da deljenje bafera u jednake delove. Na primer, za bafer u koji mogu da stanu 300 paketa za $N=3$ putanja, konvencionalno rutiranje preko jedne putanje će se koristiti sve dok trenutna popunjenost bafera ne dostigne 100 paketa, nakon čega se aktivira i dodatna disperziona putanja. Kada $W_2=200$ paketa biće aktivirana još jedna putanja što predstavlja ukupno tri aktivirane putanje koje učestvuju u paketskoj disperziji.

Predložena šema aktivacije paketske disperzije može generalno biti predstavljena sledećom relacijom:

$$W_k = \chi(k) \cdot C \quad (3.107)$$

gde je $\chi(k)$ ograničena funkcija koja određuje veličinu bafera u okviru koje je konstantan broj disperzionih putanja aktiviran. Formalna definicija prethodne funkcije se može izraziti kao

$$\chi(k) = f(k) \leq 1, \quad \forall k \geq 1 \quad (3.108)$$

gde je $f(k)$ opšta funkcija koja potencijalno može da predstavlja eksponencijalnu ili neku drugu proizvoljnu funkciju u skladu sa ograničenjima navedenim u (3.108). Izbor $f(k)$ funkcije će suštinski odrediti veličinu baferskih delova koji generalno ne moraju da budu jednaki i koji trebaju biti adekvatno planirani u skladu sa statističkim osobinama saobraćaja koji će biti podvrgnut paketskoj disperziji.

Planiranje kapaciteta bafera C je od izuzetne važnosti za efikasnost mehanizma aktivacije paketske disperzije i generalno zavisi od statističke distribucije koju saobraćaj izražava [105], [106], [101]. Zaključak prethodnih istraživanja pokazuje da saobraćaj sa *Long Range Dependence* (LRD) osobinama može da bude poguban za kapacitet bafera s obzirom da rafalne karakteristike pomenute vrste saobraćaja mogu da dovedu do velikih gubitaka paketa ako se prilikom planiranja bafeskog kapaciteta ne uzmu realne statističke osobine. U literaturi su predloženi načini kojima se delimično može rešiti pomenuti problem [102], [103].

4

REALIZACIJA INTERDOMENSKOG QOS

Proliferacija multimedijalnih servisa i rapidno povećanje zahteva za pristupom digitalnim sadržajima predstavlja sve veće opterećenje današnjoj Internet strukturi. Aktuelne Internet tehnologije i poslovni modeli koji se koriste neće biti u mogućnosti da odgovore na postavljene zahteve po pitanju kvaliteta servisa koji je povezan kako sa *content service* provajderima tako i mrežnim servis provajderima. Trenutno, Internet obezbeđuje *best-effort* servis koji zajedno sa konvencionalnim *peer-ingom* između provajdera, ugovaranjem kapaciteta tranzitnog saobraćaja (*wholesale agreement*) značajno ograničava ponudu pojedinačnih servis provajdera, kako mrežnih tako i *content* provajdera. Posledično, korisnicima se ne mogu garantovati performanse ukoliko se servis ne realizuje u okviru jednog domena koji je pod kontrolom posmatranog provajdera. Samim tim, *best-effort* servis se i dalje nudi korisnicima, dok koncept servisa sa garantovanim performansama i dalje nije realizovan uprkos objektivnim potrebama i jasnim tendencijama razvoja današnjeg Interneta.

Problem realizacije interdomenskog² QoS upravo ima za cilj da reši problem zagarantovanog servisa kako sa tehničke strane tako i sa administrativne. Naravno, primenom različitih QoS tehnika i politika o kojima je bilo reči u prethodnim poglavljima moguće je ostvariti nivo performansi u okviru jednog domena. Međutim, u interdomenskim uslovima scenario realizacije QoS se značajno razlikuje imajući u vidu relacije između domena i nivo saradnje (kooperativnosti) između istih. Dodatno, interdomenski QoS

² Termini “domen” i “provajder” u ovom radu su ekvivalentni i u kontekstu realizacije QoS podrazumevaju infrastrukturu jednog entiteta nad kojom isti ima apsolutnu kontrolu i slobodu da kreira QoS politiku u skladu sa svojim potrebama i mogućnostima.

značajno otežava i heterogenost QoS tehnologija koje pojedinačni domeni implementiraju, pa samim tim se i ostavlja malo mesta za uspostavljanje jasnih poslovnih modela i relacija između provajdera. Uspostavljanje interdomenskog QoS bi trebalo realizovati tako da bude komplementarno trenutnom *best effort* Internetu i odgovarajućim *peer-ing* konekcijama između domena. Predlaganje poslovnih modela kao i tehničkih relacija u cilju ostvarivanja QoS i izvan granica jednog domena predstavlja jedan od ključnih izazova u daljem razvoju Interneta.

4.1. Problemi i pristupi garantovanju performansi izvan granica domena

Jasno je da realizacija interdomenskog QoS ne može biti obavljena *clean slate* principom, tačnije redefinisanjem Internet arhitekture kako bi se obezbedio način garantovanja E2E performansi. Takođe, predlog takve arhitekture je u direktnoj koliziji sa zaštitom investicije u vidu tehnologija i primenjenih poslovnih modela koji trenutno važe što predstavlja ogromnu prepreku. Imajući prethodno u vidu dolazi do kristalizacije novih problema koji dodatno mogu usporiti realizaciju rešenja interdomenskog QoS, a o kojima se mora voditi računa prilikom predlaganja adekvatnog rešenja.

Nepostojanje adekvatne informacije o nivou kvaliteta servisa na tržištu onemogućava korisniku da proceni koji je to kvalitet koji se konkretno nudi i kako da isti razlikuje između servisa koje nude različiti provajderi. U odsustvu adekvatne mere kvaliteta, korisnik će se odlučiti da bez obzira na nivo kvaliteta plati tržišnu cenu što praktično ne pruža servis provajderu nikakav podsticaj da podigne nivo kvaliteta servisa ukoliko korisnik nema mehanizam da oceni isti. Dalekosežna posledica je što se ovakvim pristupom postiže da se kvalitet kao parametar istiskuje iz ponude, pa se ponovo dolazi do *best effort* koncepta kao dominantnog servisa koji se nudi korisnicima.

Nedostatak koordinacije između domena je značajna prepreka kod uspostave interdomenskog QoS s obzirom da je određen nivo saradnje između istih potreban kako bi se uspostavili dogovoreni poslovni modeli i tehnički parametri posmatrane veze. Kao posledica povećanja *downstream* saobraćaja i rasprostranjenost *video streaming*-a može se očekivati da koordinacija između domena bude još manja.

Neefikasnost rutiranja prilikom pojave zagušenja kao posledica lošeg planiranja infrastrukture kao i želja za bržom pozicioniranošću i ostvarivanjem većeg udela na tržištu, potencijalno dovodi do nepredvidivog nivoa kvaliteta servisa koji se može ostvariti. Posledično, ne postoji saobraćajna diferencijacija, pa saobraćaj visokog prioriteta često može biti poistovećen sa saobraćajem nižeg prioriteta. Ovaj problem se uglavnom rešavao značajnim povećanjem kapaciteta tako da se pojava zagušenja dešava retko ili u ekstremnim slučajevima male verovatnoće. Međutim, povećanje kapaciteta na intekonekcionim linkovima često može da dovede do problema oko ekonomskog regulisanja pitanja pokrivanja troškova istog.

Različiti učesnici u realizaciji E2E QoS mogu imati različite prioritete u rešavanju prethodno navedenih problema i nedostataka. U tom kontekstu, potencijalno rešenje bi trebalo da obuhvati što širi skup

problema koji se mogu eliminisati kako bi sveobuhvatno rešenje bilo široko prihvaćeno. Koordinacija kao jedan od preduslova za ostvarivanje E2E QoS bi trebalo da obuhvati sledeće principe i preporuke u cilju realizacije garantovanih performansi između domena:

- Zajedničku definiciju za nivo kvaliteta kako bi se povećala transparentnost i postigli skalabilni poslovni modeli;
- Zajednički koncept tarifiranja saobraćaja;
- Propagacija kontrolnih informacija koje se tiču rutiranja i deljenja prihoda;
- Koordinacija implementacije garantovanja nivoa performansi E2E servisa između mreža koje implementiraju heterogene tehnologije.

Podrazumeva se da do ovih promena u cilju postizanja E2E QoS ne može doći za kratko vreme, pa je potrebno definisati i prelazna rešenja u okviru kojih je formiranje federacija i skupa kooperativnih domena jedan od načina lakog postizanja cilja koji podrazumeva omogućavanje garancije performansi E2E servisa i izvan granica posmatranog domena. Takođe, trenutne *best-effort* interdomenske konekcije bi i dalje nastavile da postoje samo što bi bile primenljive za klase servisa koje ne zahtevaju garanciju performansi i čija raspoloživost nije osetljiva na QoS. Kod formiranja federacija je potrebno uspostaviti pravila koja se tiču tarifiranja i tehnoloških realizacija alokacije resursa kako bi se obezbedio skalabilan koncept u cilju garancije performansi.

4.2. Tehnička realizacija interdomenskog QoS

Glavni nedostatak u realizaciji interdomenskog QoS predstavlja odsustvo primene adekvatnih signalizacionih mehanizama (protokola) kako bi se obezbedile putanje koje se sastoje od lanca domena preko kojih se uspostavlja E2E QoS. Identifikacija raspoloživih putanja u multidomenskim scenarijima je ključan sastojak kako bi se obezbedila E2E veza preko koje se može garantovati zahtevni nivo performansi bilo da je taj zahtev formulisan od strane krajnjeg korisnika ili predstavlja zahtev za agregirani (*wholesale*) saobraćaj.

Interdomenski QoS se suštinski oslanja na model koji će obezbediti adekvatnu razmenu informacija između domena, a koje se tiču putanja, specifikacija klasa servisa, tarifiranja, itd. Takođe, kroz odgovarajući signalizacioni mehanizam se moraju preneti i QoS zahtevi na osnovu kojih se mora formulirati adekvatna ponuda i konačno formirati odgovarajući SLA. Interdomenski modeli moraju da specificiraju i da se bave širokom problematikom koja prevazilazi granice pojedinačnog domena i koji se bave kako signalizacijom tako i nivoom kooperativnosti između pojedinih domena. U zavisnosti od

primenjenog modela, potrebno je ostvariti različit nivo saradnje između domena, što nekada može da bude problematično ako se razmatra konkurentnost na servis provajderskom tržištu.

Poznavanje i otkrivanje informacija koje se tiču mogućnosti pojedinačnih domena predstavlja važan elemenat u kreiranju interdomenskih relacija na osnovu kojih se vrši ugovaranje potrebnih tehničkih parametara u cilju uspostavljanja E2E QoS. Adekvatni signalizacioni protokoli su *Resource Reservation Protocol* (RSVP) [107] za otkrivanje raspoloživih resursa i *Border Gateway Protocol* (BGP) [108] za otkrivanje informacija o rutiranju između domena. Međutim, trenutna primena BGP protokola podrazumeva samo razmenu informacija između domena bez ikakve dodatne razumene QoS mogućnosti i parametara. Samim tim, potrebno je realizovati odgovarajuću QoS ekstenziju BGP protokola koji bi i pored informacija o rutiranju prenosila i dodatne QoS parametre pridružene pojedinačnim putanjama. Problem sa QoS proširenjem BGP protokola se odražava kroz smanjenje skalabilnosti imajući u vidu da će implementacija QoS interdomenskog rutiranja povećati vreme konvergencije protokola rutiranja kao i da će primena proširenog BGP protokola uticati na performanse ulaznih rutera preko kojih se vrši *peering*. Analiza problema algoritama koji se koriste za QoS rutiranje je predstavljena u [109], dok je analiza sa stanovišta otvorenih problema u kontekstu QoS rutiranja sadržana u [110]. Postoji dosta predloga po pitanju QoS ekstenzije BGP protokola pri čemu i dalje ne postoji konkretan plan implementacije ovakve BGP varijante [111], [112], [113].

Najveći izazov u realizaciji interdomenskog QoS predstavlja odsustvo zajedničke definicije klase servisa koja će obuhvatati različite domene. Ova raznolikost u servisnoj diferencijaciji se ogleda u potrebi za mapiranjem klasa servisa između domena u zavisnosti od primenjenog modela ugovaranja između domena kao i E2E servisnih specifikacija. Naravno, nije samo broj definisanih klasa servisa problematičan, već sama specifikacija. Naime, potrebno je da u okviru svake klase servisa postoje QoS parametri koji su pokazatelji performansi koje saobraćaj pridružen posmatranoj klasi servisa može da očekuje. QoS parametri su različiti od klase do klase i od domena do domena pri čemu se razlikuju u zavisnosti od primenjene mrežne tehnologije u domenu, domenske hijerarhije i nivoa servisne diferencijacije (QoS granularnosti). Međutim, i pored usvajanje QoS parametara može se očekivati da nemaju sve klase servisa iste definisane parametre, pa posledično, u tim slučajevima i nije moguće realizovati mapiranje klasa servisa jer nije jasno koje E2E performanse se dobijaju mapiranjem klasa servisa sa različitom specifikacijom QoS parametara.

Problem definicije odnosno specifikacije klase servisa postaje složeniji sa uvođenjem većeg broja domena prilikom uspostavljanja E2E servisa. Posledično, kako bi se povećala verovatnoća uspešnog mapiranja klasa servisa, potrebno je koncizno i dosledno definisati po domenu klase servisa sa preciznim QoS parametrima i njihovim graničnim vrednostima. Ugovaranje tehničkih parametara, u ovom slučaju QoS parametara od kojih *de facto* i zavise E2E performanse mora biti precizno što je formalizavano u okviru SLA dokumenta između domena.

Naredni problem sa tehničkog aspekta podrazumeva verifikaciju performansi E2E servisa koja se mora realizovati posredstvom procesa monitoringa performansi. Kompleksnost u monitoringu se ogleda u interdomenskom scenariju s obzirom da se moraju nadgledati performanse koje pojedinačni domen realizuje kako bi se uporedile sa dogovorenim performansama koje diktira prethodno ugovoreni SLA. Kao i kod servisne definicije, problem monitoringa je teži za realizaciju i kontrolu sa učestvovanjem većeg broja domena u uspostavi E2E servisa, pa su alternative da se domen koji kupuje servis od drugog ili drugih domena osloni na njihov monitoring ili da monitoring realizuje nezavisno treće lice. Realizacija praćenja E2E performansi mora da bude takva kako bi se u slučaju kršenja QoS parametara definisanih u SLA mogao identifikovati domen koji je „podbacio“ sa performansama pri čemu identifikacija tog domena mora da bude transparentna kako se ne bi došlo do scenarija problema sa performansama, dok domen koji krši SLA neće da prizna problem.

Značajnu ulogu u definiciji klase servisa, pa i samog SLA imaju mere performansi ili QoS parametri. Aktuelne mere performansi se uobičajno realizuju posredstvom *Internet Control Message Procol* (ICMP) protokola i to kao srednje vrednosti kašnjenja u oba smera – *Round Trip Time* (RTT). Dodatno, posredstvom ICMP je moguće odrediti i procenat izgubljenih paketa. Važna činjenica kod korišćenja ICMP kao dijagnostičkog alata jeste period u kome se vrši merenje i broj paketa koji se koriste za merenje. Dodatno, džiter je takođe sastavni deo SLA dokumenta iako se rede pojavljuje u odnosu na prethodna dva QoS parametra, kašnjenje i procenat izgubljenih paketa.

U kontekstu interdomenskih merenja kako bi merni rezultati bili svrsishodni, mora se za njih primeniti odgovarajuća kompozicija mera performansi. Da bi ta kompozicija bila moguća, definicija QoS parametara, period merenja, agregacija rezultata merenja i referentna vremena u kojima se merenje realizuje moraju da budu usklađena između domena. Kao što se primećuje, nije samo dovoljna dosledna definicija QoS parametara, već i dodatni parametri kako bi merni rezultati bili uporedivi i kako bi kompozicija metrika bila moguća [114]. Na ovaj način je moguće identifikovati problematičan domen u lancu domena preko kojih se uspostavlja E2E servis.

Prilikom odabira perioda merenja mora se voditi računa da merenja ne budu česta kako ne bi unela preveliku količinu mernog saobraćaja koji bi potencijalno mogao da poremeti performanse korisničkog saobraćaja. U cilju vršenja agregacije merenja potrebno je da se obuhvati dovoljan broj pojedinačnih merenja kako bi se postigao visok nivo statističke tačnosti. Period agregacije mora da bude dovoljno veliki kako bi se postigla sinhronizacija između domena, a i obuhvatile početak i kraj velikog broja saobraćajnih tokova. Kod odabira perioda agregacije, dodatno je potrebno obratiti pažnju i na eventualne promene u mreži, kao što su rafali saobraćaja ili otkazi i procesi promene rutiranja koji mogu da utiču na E2E performanse kako se ovi događaji ne bi usrednjili i „zamaskirali“.

4.3. Definicija klasa servisa i servisnog zahteva

U kontekstu intedomenskog QoS definisanje nekoliko zajedničkih klasa servisa od strane više domena bi značajno umanjilo kompleksnost i izbegla bi se potreba za mapiranjem klasa servisa između domena. Mapiranje je potrebno u slučaju QoS heterogenosti i različite QoS diferencijacije između domena što je čest slučaj. Osim toga, nepravilno mapiranje potencijalno može da dovede da se E2E saobraćaj realizuje preko putanja koje ili nisu u stanju da ispune E2E servisne zahteve (što dovodi do nemogućnosti garancija performansi) ili preko putanja koje su predimenzionisane u vidu performansi (što dovodi do neefikasnog iskorišćenja resursa).

Kategorizacija aplikacija i pojedinačnih servisa u okviru klasa servisa se realizuje na osnovu servisnih zahteva koju formuliše ta aplikacija ili servis. Servisni zahtev se definiše zahtevanjem maksimalne vrednosti jednog ili više QoS parametara. Saobraćaj koji je pridružen pojedinačnoj klasi servisa se tretira tako da se postignu ciljane vrednosti QoS parametara definisanih servisnim zahtevom. Adekvatan nivo servisne diferencijacije bi omogućio da se pokrije širok spektar servisnih zahteva.

Broj definisanih klasa servisa bi trebalo da bude nizak, u skladu sa potrebama, pri čemu je potrebno definisati klasu servisa za servise koji su osetljivi na promenu QoS (VoIP, video), dok specifikacija ostalih klasa servisa može da se definiše u skladu sa potrebama pojedinačnog domena kao i potrebama drugih domena za koji posmatrani domen realizuje tranzitni servis.

Servisna specifikacija suštinski formuliše nivo E2E performansi koji se očekuje na putanji od izvora do odredišta. Ove performanse se generalno definišu posredstvom QoS parametara pri čemu je ponovo akcenat na jedinstvenoj definiciji QoS parametara kao i periodima u okviru kojih se isti mere. Takođe, u cilju uspostavljanja E2E QoS u interdomenskom scenariju, potrebno je da klase servisa specificiraju zahtevane parametre kako bi se mogle garantovati performanse preko lanca domena. Scenario u kome postoji specifikacija servisnog zahteva, kao i specifikacija klasa servisa u okviru svakog pojedinačnog domena, dovodi neminovno do potrebe za mapiranjem klasa servisa kako bi se kao rezultat tog mapiranja realizovao E2E QoS koji je potrebno da ispunjava servisne zahteve. Naravno, mapiranje klasa servisa između domena ne bi bilo potrebno da su specifikacije klasa servisa u okviru domena iste. Međutim, to nije realan scenario imajući u vidu da postoje odgovarajuća domenska hijerarhija, tehnološka heterogenost i težnja pojedinačnih domena da se probiju na tržištu sa boljim kvalitetom servisa. Kao posledica ovakve definicije servisnog zahteva i specifikacije klasa servisa, mapiranje klasa servisa predstavlja sastavni deo uspostave intedomenskog QoS i kao takav se mora dodatno analizirati.

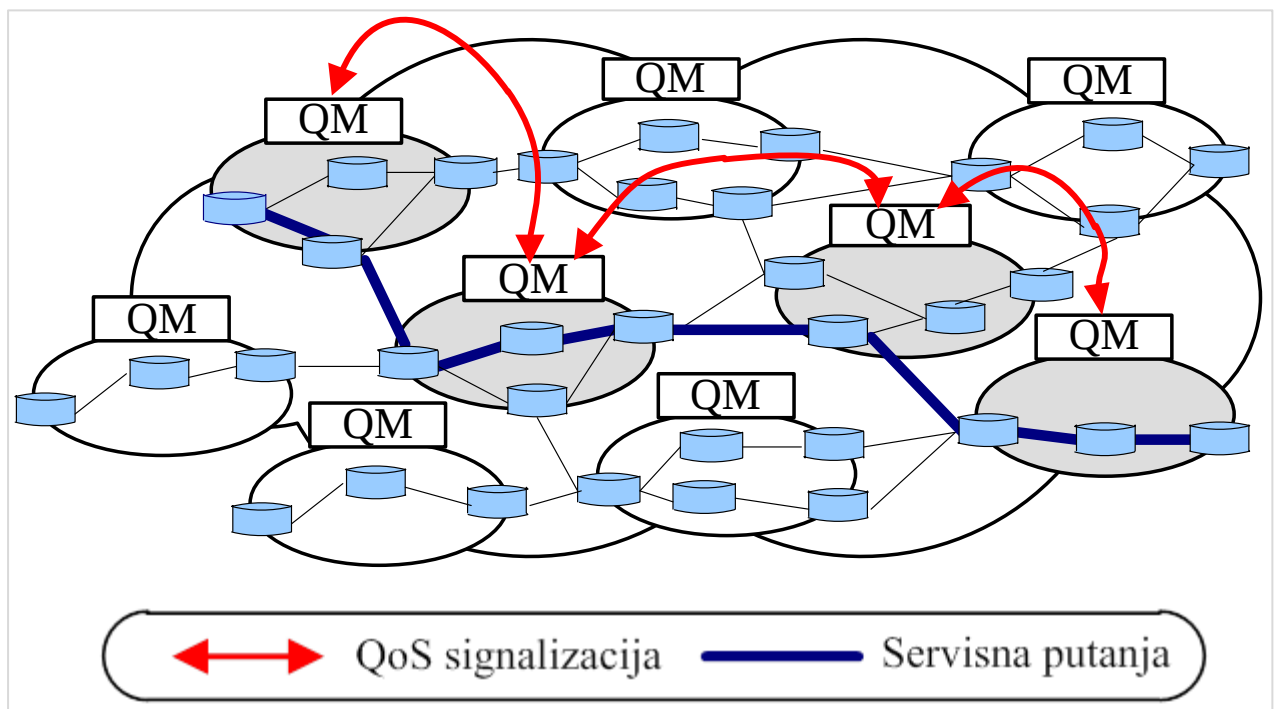
4.4. Modeli ugovaranja servisa

Ugovaranje E2E servisa preko više domena se realizuje posredstvom modela koji specificiraju način ugovaranja servisa kao i smer signalizacije kako bi došlo do uspostavljanja E2E QoS. Ovi modeli

ugovaranja servisa između domena diktiraju način i nivo koordinacije između domena kako bi došlo do uspostave zahtevanog servisa čije su performanse specificirane E2E servisnim zahtevima.

4.4.1. Bilateralni model

U današnjem Internetu, ugovaranje servisa između domena se uglavnom realizuje posredstvom osnovnog bilateralnog modela, koji podrazumeva da svaki domen ugovara servis samo sa susednim domenima [115]. U skladu sa bilateralnim modelom dva domena se povezuju posredstvom jedne ili više veza nakon čega se implementira odgovarajuće mapiranje klasa servisa u skladu sa prethodno određenom specifikacijom. Poslovni procesi koji prate ugovaranje E2E servisa podrazumevaju definisanje SLA čime se definišu penali za eventualno kršenje garantovanih performansi. Dodatno, svaki domen specificira odgovarajuće klase servisa, pa se u okviru bilateralnog modela ugovaranja servisa uglavnom realizuje statičko mapiranje kome prethodi dogovor između domena. Prilikom specifikacije klasa servisa, dogovaraju se mere performansi, kao i ostali tehnički detalji koji su potrebni za uspostavljanje E2E QoS. Bilateralni model ugovaranja u interdomenskom scenariju je prikazan na slici 4.1.



Slika 4.1 Bilateralni model ugovaranja servisa.

Bilateralni model kao najprostiji model ugovaranja servisa između domena zahteva odgovarajući nivo saradnje između domena koji u velikom broju slučajeva može biti prepreka uspostavi interdomenskog QoS. Statičko mapiranje klasa servisa kao i ugovaranje pratećih tehničkih i ekonomskih parametara kod ovog modela je statički, što može da dovede do dužeg vremena uspostave E2E servisa. Statička ugovaranja između susednih domena se realizuju posredstvom QoS Manager (QM) kao što je prikazano na slici 4.1.

Takođe, na slici je označena adekvatna signalizacija koja podrazumeva samo razmenu potrebnih informacija između susednih domena. Realizacija E2E servisa korišćenjem bilateralnog modela preko više domena može da dovede do neefikasne alokacije resursa u pojedinačnim domenima s obzirom da za uspešnu uspostavu E2E servisa mora da svako bilateralno ugovaranje servisa između susednog domena bude uspešno. Dovoljno je da se ne realizuje samo jedno ugovaranje između para domena što uzrokuje neuspelo E2E ugovaranje servisa na čitavoj putanji.

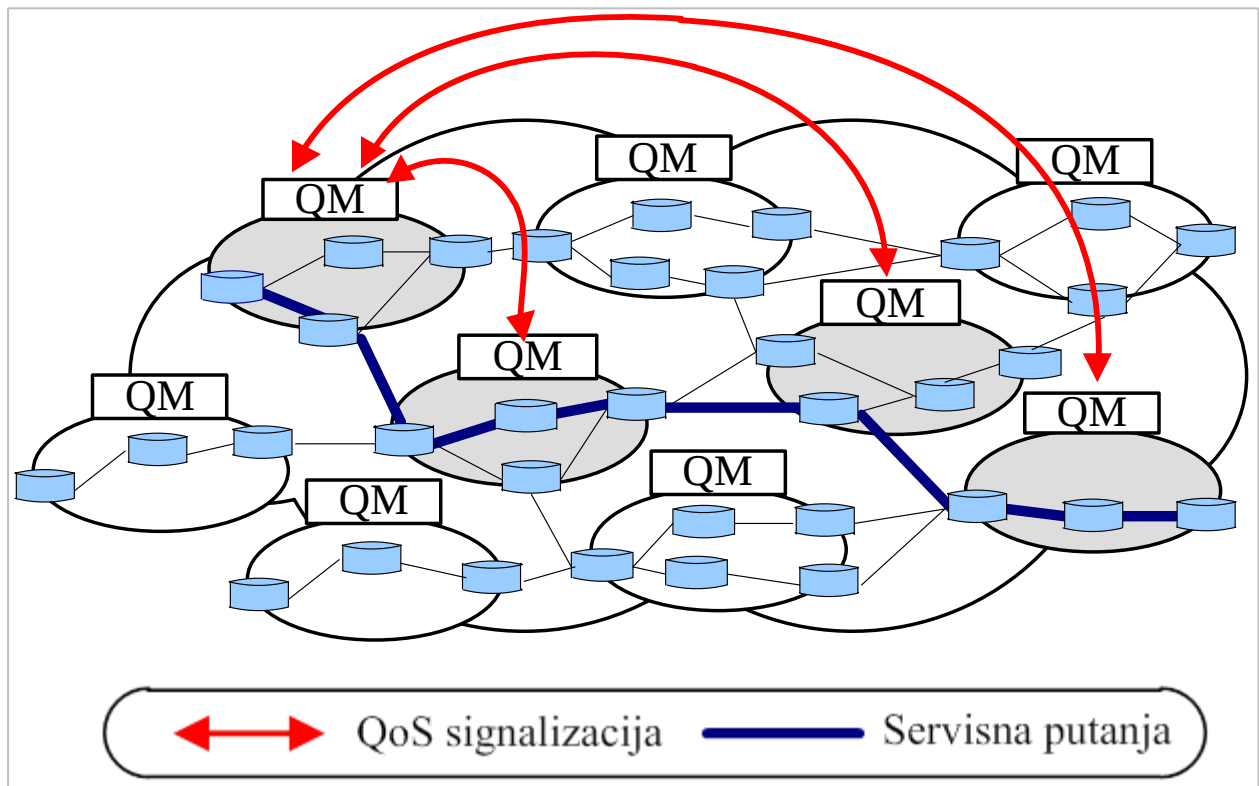
Nivo saradnje između domena kod bilateralnog modela predstavlja još jedan problematičan aspekt ovog načina ugovaranja. Naime, usled statičkog ugovaranja servisa između pojedinačnih parova susednih domena potreban je visok nivo saradnje u razmeni tehničkih parametara i tehničkih mogućnosti domena, koje podrazumeva potencijalnu detekciju topologije i primenjenih tehnologija, što dalje može dugoročno uticati na konkurentnost domena. Veliko ograničenje bilateralnog ugovaranja servisa između susednih domena može biti nemogućnost uspešne realizacije E2E QoS ako u realizaciji istog učestvuje veći broj domena. Kako bi se rešila pomenuta ograničenja razmatraju se alternativni modeli ugovaranja servisa između domena.

4.4.1. Kaskadni model

Kaskadni model predstavlja ekstenziju bilateralnog modela u okviru koga domen koji vrši iniciranje uspostave E2E servisa i koji zahteva realizaciju E2E QoS direktno ugovara servis sa domenima koji se nalaze na putanji preko koje se uspostavlja E2E servis. Ovaj model za ugovaranje E2E servisa koristi drugačiji pristup od bilateralnog s obzirom da domen inicijator stupa u ugovaranje servisa sa više učesnika (domena) kako bi se ustanovilo da li je kompozicija performansi u stanju da odgovori na E2E servisne zahteve. Ovakav pristup unosi odgovarajući nivo implementacione kompleksnosti s obzirom da je inicirajući domen odgovoran za upravljanje nivoom performansi i dinamičkim ugovaranjem sa ostalim domenima na putanji. Na slici 4.2 je prikazan koncept kaskadnog modela, kao i signalizacioni saobraćaj koji se razmenjuje između domena koji realizuju E2E servis.

Kaskadni model ugovaranja servisa između domena omogućava implementaciju dinamičkog ugovaranja koje je dosta fleksibilnije i skalabilnije u odnosu na pristup koji se koristi kod bilateralnog modela. Posredstvom signalizacije dinamički se razmenjuju tehnički parametri koji su potrebni za uspostavljanje interdomenskog QoS pri čemu vreme ugovaranja servisa traje kraće u odnosu na bilateralni model imajući u vidu dinamički pristup ugovaranju servisa. Signalizacioni protokoli koji su adekvatni za razmenu QoS parametara između domena su RSVP i *Next Steps in Signaling* (NSIS) [116]. Osim razmene signalizacije za potrebe realizacije E2E QoS kojom će se razmeniti tehnički parametri, potrebno je i da se realizuje odgovarajući *discovery* mehanizam koji će biti u stanju da utvrdi sa stanovišta inicirajućeg domena lanac domena kroz koje će se E2E servis uspostaviti. Ovaj *discovery* mehanizam je potrebno da dinamički utvrdi

informacije o rutiranju kako bi se nakon toga razmenile informacije o QoS mogućnostima koje su pojedinačni domeni na identifikovanoj putanji u stanju da ponude.



Slika 4.2 Kaskadni model ugovaranja servisa.

Nedostatak procesa signalizacije kod kaskadnog modela se ogleda u tome što se signalizacioni saobraćaj mora razmeniti sa udaljenim domenima, što ponovo zahteva poznavanje informacija o rutiranju između domena. Takođe, mora se dalje izvršiti i analiza kojom bi se utvrdilo na koji način bi se mogla obaviti bezbedna razmena signalizacije imajući u vidu da se signalizacione poruke razmenjuju između inicirajućeg domena i svakog pojedinačnog domena na putanji posredstvom tranzitnih domena. Praktično se signalizacija realizuje *in-band*, odnosno prenosi se putanjama kao i ostale klase saobraćaja što opet ostavlja mesta za razmatranje kako će tranzitni domeni tretirati ovu klasu saobraćaja. Posledično, iako kaskadni model rešava nedostatak skalabilnosti koji je postojao kod bilateralnog modela, dinamički pristup ugovaranju E2E servisa posredstvom signalizacionih protokola stvara nove probleme u pogledu signalizacije i implementacije.

Nivo saradnje između domena kod kaskadnog modela se ogleda u definiciji tehničkih parametara koje je potrebno dinamički razmeniti. Na ovaj način je problem saradnje između domena sveden na potreban nivo kako bi se razmenili potrebni parametri na način koji diktira proces signalizacije. Imajući u vidu da se ugovaranje realizuje na dinamički način, česta praksa između domena jeste i formiranje odgovarajućih dogovora koji se ogledaju u realizaciji alijansi i federacija kako bi više domena mogli da deluju jedinstveno

i da zajednički ponude odgovarajući servis drugim domenima i korisnicima. Ovakav pristup je analiziran u [117] i kvalifikovan je kao kooperativni model. Kooperativni model suštinski se može razmatrati i kao generalizacija osnovnog bilateralnog modela u okviru koga domeni vrše razmenu potrebnih tehničkih i ekonomskih informacija u cilju uspostave određenog nivoa QoS.

Federacija podrazumeva skup domena između kojih se razmenjuju dodatne informacije i definiše skup ciljeva i tehnologija koje će se koristiti za ostvarivanje tih ciljeva. Sa druge strane, alijanse podrazumevaju skup domena koji osim prethodnog specificiraju jedinstvene poslovne politike i načine deljenje prihoda od tarifiranja. Naime, federacija i alijansa predstavljaju skup domena koji su odlučili da se na tržište pojave sa određenim servisom jedinstveno pri čemu se njihova interna organizacija u okviru ove dve strukture razlikuje. Kod federacija, akcenat je na definiciji jedinstvenih tehničkih pristupa, dok kod alijansi se podrazumeva i usklađivanje poslovnih politika i način realizacija profita od posmatranog servisa.

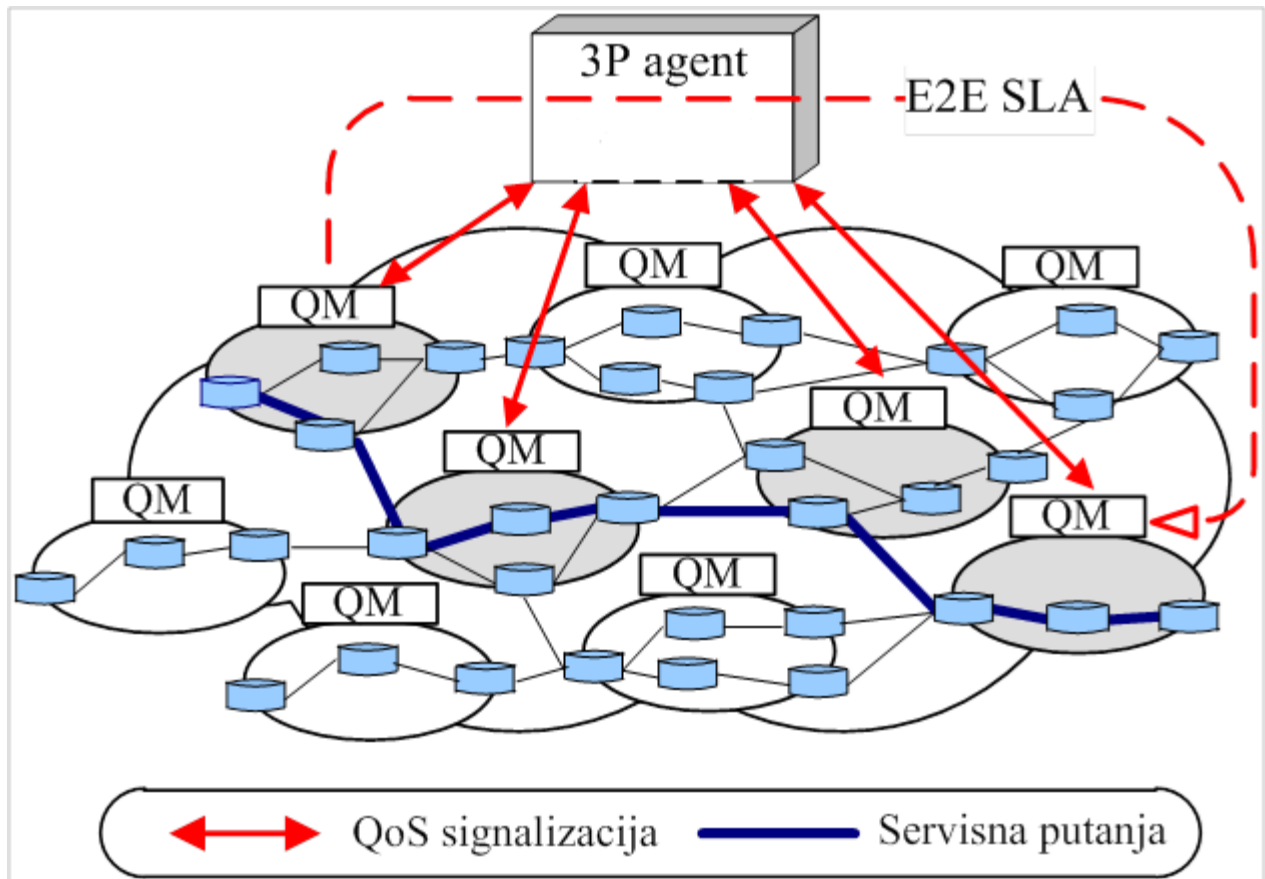
Formiranje federacija i alijansi u okviru kaskadnog modela smanjuje kompleksnost realizacije E2E QoS pri čemu je lakše implementirati *discovery* mehanizme i sprovesti odgovarajuće poslovne politike. Važno je napomenuti da se kroz signalizaciju može realizovati i odgovarajuće mapiranje servisa, pa je stoga potrebno i uposliti adekvatan algoritam koji će moći da obavi ispravno mapiranje klasa servisa između skupa domena sa odgovarajućim ciljevima.

Praktično se povećavanjem implementacione kompleksnosti kod kaskadnog modela ugovaranja servisa rešava problem skalabilnosti statičkog ugovaranja čime se značajno smanjuje vreme ugovaranja i povećava verovatnoća uspešnog ugovaranja E2E servisa kada učestvuje veći broj domena. Implementacijom signalizacionog protokola se mora rešiti način rada *discovery* mehanizma, kao i problem *in-band* signalizacije. Imajući u vidu prethodne nedostatke, moguće je definisati novi model ugovaranja servisa između domena koji će uvođenjem nezavisnog agenta posredovati u ugovaranju tehničkih detalja između domena. Aspekti koji utiču na balans između fleksibilnosti i kompleksnosti implementacije se mogu sagledati u okviru [118].

4.4.2. *Third party (3P) model*

Nedostaci ispoljeni kod prethodna dva modela zahtevaju uvođenje novog modela ugovaranja koji bi potencijalno mogao da eliminiše pomenute nedostatke i koji bi bio adekvatan za upotrebu u cilju uspostavljanja interdomenskog QoS. Uvođenje *third party* (3P) agenta kao nezavisnog učesnika ugovaranja E2E servisa podrazumeva da će svaki pojedinačni domen na E2E servisnoj putanji direktno razmenjivati signalizacione poruke sa 3P agentom pri čemu nema direktne interakcije niti saradnje sa susednim (bilateralni) niti sa ostalim domenima (kaskadni). Ugovaranje između domena i 3P agenta se realizuje dinamički, koristeći adekvatan signalizacioni protokol. Ugovaranje od strane svakog domena se obavlja kroz QM koji razmenjuje neophodne parametre sa 3P agentom (Slika 4.3).

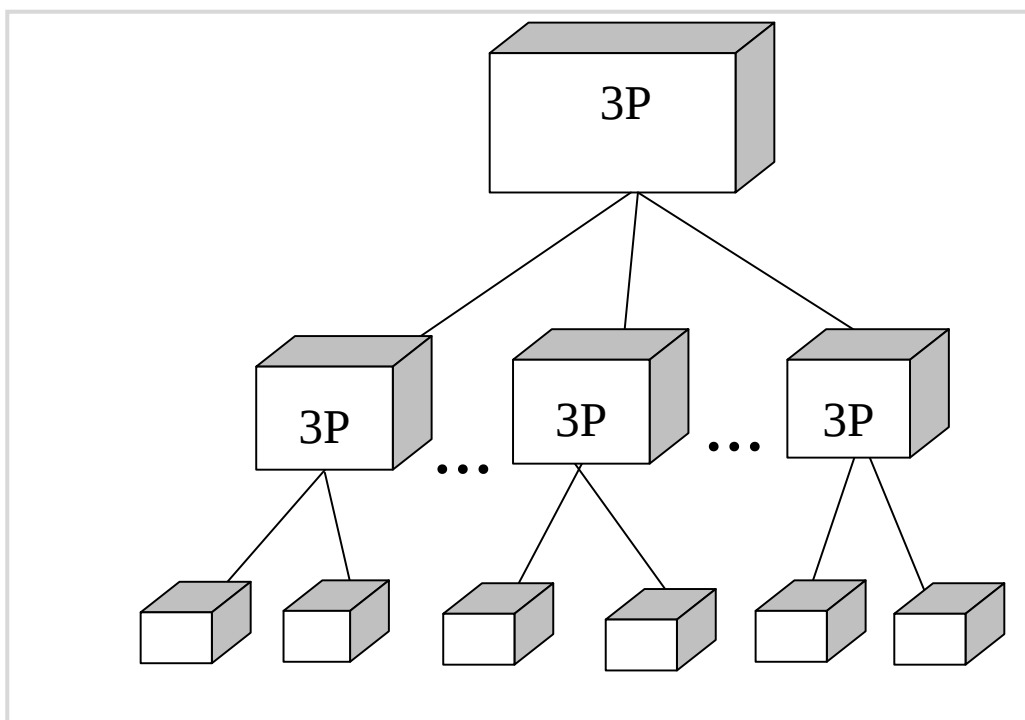
Nivo saradanje između domena je posredstvom 3P modela ugovaranja u potpunosti ukinut s obzirom da postoji posrednik u ugovaranju i sva signalizacija koja prethodi uspostavi E2E servisa se prenosi do 3P agenta. U okviru 3P agenta se nalaze sve informacije koje su definisane od strane pojedinačnih domena, a tiču se specifikacije klasa servisa, QoS parametara, QoS granularnosti, tarifnih informacija i informacija o rutiranju. Centralizovano uređenje 3P modela obezbeđuje da se na jednom mestu nalaze potrebne informacije o uspostavi E2E servisa, pa je imajući ovo u vidu, potrebno da se realizuju adekvatni algoritmi i protokoli za proces razmene signalizacije, tarifiranja i mapiranja klasa servisa.



Slika 4.3 3P model ugovaranja servisa.

Adekvatan signalizacioni protokol za primenu u okviru 3P modela je *Common Open Policy Service* (COPS) [119]. Ovaj protokol je zasnovan na klijent-server modelu i podrazumeva upravljanje politikama kontrole QoS. Ove politike su smeštene na serveru, dok se iste sprovode na klijentima. Analizirajući centralizovanu strukturu 3P modela kao i osnovne elemente koji ga čine, jasno je da je 3P agent u COPS terminologiji server i kao takav predstavlja *Policy Decision Point* (PDP), dok su QM u okviru pojedinačnih domena klijenti koji suštinski predstavljaju *Policy Enforcement Point* (PEP) i mesto gde se definisane politike sprovode. Praktično se PEP i PDP uređenjem postiže realizacija algoritama tarifiranja, rutiranja, mapiranja pri čemu se rezultati tih algoritama šalju na PEP pozicije koji te politike sprovode.

3P predstavlja posrednika u uspostavi E2E QoS pri čemu je potrebno da postoji generalna saglasnost domena koji se nalaze pod ingerencijom 3P agenta o potrebnim informacijama koje pojedinačni domen treba da dostavi 3P agentu na osnovu kojih bi bilo moguće da se uspostavi interdomenski QoS. U kontekstu performansi E2E servisa svaki domen definiše svoju diferencijaciju klasa servisa kao i cenu za svaku pojedinačnu klasu servisa. Na osnovu podataka o specifikaciji klasa servisa (posredstvom definisanja vrednosti QoS parametara), 3P agent treba da izvrši selekciju klase servisa u okviru svakog domena kako bi E2E servisni zahtev bio ispunjen i kako bi E2E QoS bio uspostavljen. Fleksibilnost u upotrebi 3P modela se upravo ogleda u centralizovanoj strukturi i kroz proces mapiranja klasa servisa koji se može prilagoditi različitim poslovnim politikama. Primena algoritama za mapiranje klasa servisa predstavlja centralno težište ovog rada u kontekstu uspostave interdomenskog QoS, s obzirom da u literaturi ne postoje adekvatna rešenja za problem mapiranja klasa servisa između domena u 3P modelu ugovaranja.



Slika 4.4 Hijerarhijska organizacija 3P agenata.

Vreme ugovaranja E2E servisa kod 3P modela zavisi od vremena koje je potrebno 3P modelu da izvrši potrebne procese koji se tiču razmene signalizacije sa pojedinačnim domenima, mapiranje klasa servisa i eventualno primene dogovorene poslovne politike. Bez obzira na broj procesa, vreme ugovaranja servisa preko 3P agenta je višestruko kraće u odnosu na bilateralni i kaskadni model.

Skalabilnost 3P modela je ograničena kapacitetom jednog 3P agenta da posreduje u ugovaranju servisa za određeni broj domena koji se nalazi pod autoritetom posmatranog 3P agenta. Uvođenje hijerarhijske organizacije 3P agenata bi rešio problem skalabilnosti tako što bi svaki 3P agent imao svoje mesto u hijerarhijskoj strukturi pod čijom kontrolom bi bio odgovarajući profil domena (na primer, *access domeni*), dok bi na sledećem nivou bili 3P agenti koji bi ugovarali servis za domene koji se nalaze na višem hijerarhijskom nivou [120]. Relacija 3P agenata u hijerarhijskoj organizaciji je prikazana na slici 4.4. Signalizacija bi se razmenjivala, ako je potrebno uključivanje domena koji se nalaze pod autoritetom drugog 3P agenta, direktno između 3P agenata koji se nalaze na različitom ili istom nivou. Takođe, zbog skalabilnosti, preporuka je da komunikacija 3P agenata koji pripadaju istom nivou u hijerarhiji ide preko 3P agenta višeg nivoa.

Proučavanje i analiza na 3P modelu ugovaranja servisa nije zastupljena u meri kao što je to slučaj sa bilateralnim i kaskadnim modelima. Ovom činjenicom je ostavljen prostor za razvijanje novih i fleksibilnih poslovnih politika koje se otvaraju u prisustvu 3P agenata kao nepristrasnih entiteta u procesu ugovaranja E2E servisa i posledično interdomenskog QoS. Vodeći računa da ovaj rad podrazumeva analizu tehničkih aspekata realizacije interdomenskog QoS, proces alokacije odgovarajućih performansi odnosno mapiranje klasa servisa u 3P modelu ugovaranja predstavlja cilj ovog rada u kontekstu interdomenskog QoS. Predlaganje algoritama za mapiranje klasa servisa može doprineti razvoju fleksibilnih poslovnih politika koji bi doprineli olakšavanju uspostave E2E QoS preko 3P agenta.

4.5. Modeli interdomenskog tarifiranja

Do sada je 3P model razmatran u literaturi samo u kontekstu tehničkih problema u realizaciji interdomenskog QoS [117]. Signalizacioni zahtevi kod ugovaranja servisa su analizirani u [121]. Doprinosi predstavljeni u tom radu se odnose na efikasnost i fer raspodelu sredstava kada je u pitanju 3P model u poređenju sa kaskadnim modelom.

U *New Generation Networks* (NGN) okruženju zahtev za kontrolu E2E QoS putanje neminovno utiče na primenu modela ugovaranja servisa. Umesto za slobodnim ugovaranjem tranzitnog ili *peering* servisa između susednihdomena, izražena je jaka potreba za postojanjem jednog učesnika (provajdera ili nezavisnog entiteta) koji kontroliše plaćanje servisa (na primer, transportnih kapaciteta) od strane svih domena koji učestvuju u realizaciji E2E servisa. Tarifa interkonekcije treba da bude određena na način koji promoviše ekonomsku efikasnost i podstiče investicije provajdera u implementaciju novih tehnologija i to sa ciljem smanjenja troškova ili proširenja skupa servisa koji se nude korisnicima.

Sveobuhvatna ekonomska analiza različitih modela povezivanja su dostupni u literaturi [122], [123], [124]. *Bill-and-keep* model podrazumeva uspostavljanje bidirekcionog ugovora bez eksplicitnog naplaćivanja interkonekcionog servisa između domena i predstavlja ekonomski razumnu politiku ukoliko je saobraćaj

simetričan (*upstream* i *downstream* su približno jednaki). Ekonomski modeli zasnovani na ceni podrazumevaju maloprodajnu cenu servisa koji se nudi krajnjim korisnicima i ta cena se određuje kao suma svih troškova (uključujući i inkrementalne troškove) kao i očekivani profit. Dva principa su dominantna kada se radi o ovom tarifnom pristupu: (1) *initiating party's network pays* (IPNP) koncept u kome strana koja je inicirala saobraćaj plaća sve troškove terminacije; (2) *receiving party's network pays* (RPNP) koncept u kome strana koja terminira servis plaća sve troškove. Model deljenja prihoda (*revenue sharing*) pretostavlja da se profit i gubici dele između grupe domena u skladu sa predefinisanim dogovorom.

Komparativna analiza kooperativnih i nekooperativnih strategije za deljenje prihoda su analizirane u [125]. Sa aspekta teorije igara dokazano je da nekooperativna strategija može dovesti do nepravedne raspodele prihoda između domena. Međutim, deljenje prihoda na osnovu fer principa koji je težinski proporcionalan podstiče saradnju između provajdera i može da dovede do viših profita za sve provajdere koji učestvuju u saradnji. U [126] analizirani su četiri scenarija ugovaranja: (1) *best effort*; (2) bilateralno pregovaranje o ceni između susednih domena; (3) kaskadna ugovaranja koja podrazumevaju konkatenciju bilateralnih ugovaranja i (4) formiranje velike koalicije između svih domena u lancu. Finansijski dogovori u ovakvim dogovorima su analizirani posmatrajuću različite QoS nivoe koji su definisani u skladu sa očekivanim kvalitetom servisa od strane korisnika. Formiranjem domenskih federacija bi se obezbedila tehnološka konvergencija što je razmatrano u [127]. Naročito se naglašava scenario ugovaranja servisa između domena preko 3P agenta i posledično se analizira formulacija teorije igara pridružena ovom modelu.

Tarifiranje interkonekcije koje se kontroliše posredstvom E2E SLA i koji se ugovara posredstvom kaskadnog modela je predmet analize predstavljene u [128]. Kompleksne relacije između domena su analizirane uključujući i koegzistenciju domena koji ne isporučuju dogovorene performanse kao i domene koji kompenzuju lošije performanse razmatrajući kooperativni scenario. Prva faza podrazumeva kvantifikovanje odnosa između potrebnog QoS, ponuđenog QoS i realno postignutog QoS na oba nivoa – E2E i domenskom nivou. Nominalna cena koja se nudi korisnicima, kao i tarifa interkonekcije za posmatrane klase servisa se koriguju ukoliko je to potrebno. Korekcija se realizuje na osnovu prethodno pomenute kvantifikacije. Predloženo je i analizirano pet politika pod različitim QoS scenarijima. Istraživanje je pokazalo da ispravno odabrana SLA politika treba da podstakne saradnju između domena i da stimuliše svaki domen da ponudi servis koji približno ili tačno odgovara servisnim zahtevima kako bi se ispunile ugovorne obaveze diktirane posredstvom SLA.

Primena adekvatnih politika tarifiranja je ključna u oceni prihvatanja pojedinih modela ugovaranja servisa. Analize su pokazale da 3P model pokazuje visok potencijal u primeni fleksibilnih politika koje mogu, a i ne moraju da se zasnivaju na saradnji domena u uspostavi E2E QoS. Bitna činjenica je da nisu identifikovane sve prednosti i nedostaci 3P modela, a samim tim i nije otkriven pun potencijal 3P agenta u fleksibilnoj realizaciji interdomenskog QoS. Samim tim, budući rad na istraživanju fleksibilnih i

adekvatnih politika tarifiranja u okviru 3P modela bi trebalo da dovede do predloga koje će podstaći domene da ponude bolje servise sa garantovanim performansama.

4.6. Mapiranje klasa servisa između domena

Proces mapiranja klasa servisa između domena predstavlja deo ugovaranja QoS od čijeg ishoda zavise E2E performanse. Za realizaciju mapiranja klasa je potrebno balansirati između kompleksnosti i odgovarajuće selekcije klasa servisa između domena kako bi se uspostavio E2E QoS u skladu sa dogovorenom poslovnom politikom. Naravno, statičko mapiranje je najjednostavnije sa aspekta implementacije, ali ovakav pristup ima ozbiljna ograničenja u pogledu fleksibilnosti i skalabilnosti. Takođe, statičko mapiranje zahteva i definiciju zajedničke klase servisa između domena kako bi se omogućilo predodređeno mapiranje koje je uglavnom primenljivo na individualnom paru domena. Posledično, problem mapiranja klasa servisa ostaje otvoren problem jer trenutno ne postoji dovoljno skalabilno i jednostavno rešenje koje bi bilo rasprostranjeno.

Jedan pristup rešavanju problema mapiranja klasa servisa podrazumeva definisanje poznatog skupa klasa servisa koje su specificirane odgovarajućim QoS parametrima. Ove klase servisa, njihov broj i specifikacija su poznati svim domenima kako bi se postigla adekvatna konvergencija poslovne politike i implementiralo mapiranje niske kompleksnosti. Bez obzira na predloženo rešenje mapiranja i njegovu implementaciju, svaki domen treba da bude nezavisan u planiranju svoje poslovne politike i specifikacije klasa servisa koje nudi.

Alternativno se može okrenuti definiciji generičkog šablona specifikacije klase servisa i definiciji odgovarajućeg algoritma za mapiranje ovih generičkih šablona između domena. Generički šabloni mogu da budu modifikovani od strane administratora domena sa vrednostima u skladu sa tehnološkim mogućnostima i definisanom poslovnom politikom. Međutim, ovakav pristup je uglavnom primenljiv kod bilateralnog modela. Mapiranje klasa servisa kod bilateralnog modela može biti realizovano na način koji nije fer prema svim domenima koji učestvuju, a kao razlog toga je nejednaka distribucija resursa u lancu domena koju nije lako kontrolisati. Ova distribucija resursa se ogleda u tome što neki domeni u lancu mogu da alociraju manje resursa za razliku od drugih koji moraju da ponude više resursa kako bi E2E SLA bio ispunjen. Naravno, ovakav pristup nije adekvatan, pa je fer selekcija klasa servisa jedan od bitnih parametara koji se mora sagledati prilikom ocene algoritma za mapiranje klasa servisa.

QoS specifikacije obuhvataju različite mere performansi pri čemu je prema [129] za IP mreže dovoljna specifikacija četiri parametra koji kvantifikuju performanse i to: kašnjenje, *throughput*, džiter i procenat izgubljenih paketa. Svaka klasa servisa je specificirana definisanjem graničnih vrednosti ovih parametara koje je moguće ostvariti u okviru te klase. U zavisnosti od specifikacije nije obavezno definisati sve QoS parametre u okviru jedne klase. Najbolje klase servisa koje domeni nude (*premium* klase servisa)

podrazumevaju specifikaciju u okviru kojih se specificiraju kašnjenje, džiter i procenat izgubljenih paketa. Ove klase su uglavnom namenjene za QoS kritičan saobraćaj kao i za multimedijalne servise koji su po pravilu osetljivi na pogoršanje performansi. Ostale klase domeni mogu da specificiraju u skladu sa svojim potrebama i mogućnostima.

4.7. Algoritmi za mapiranje klasa servisa između domena 3P modela ugovaranja servisa

Do sada 3P pristup nije detaljno analiziran sa aspekta interdomenskog QoS. Predlog realizacije interdomenskog SLA je dat u okviru [130] pri čemu se pretpostavlja delegiranje parcijalne odgovornosti 3P agentu u vezi sa nadgledanjem i merenjem performansi. Predlog multidimenzione i autonomne QoS evaluacije koja se zasniva na 3P serverima dat je u [131].

Sa druge stranem algoritmi za mapiranje klasa servisa u 3P modelu nisu razmatrani u literaturi. Proces mapiranja klasa servisa može da uzrokuje da se E2E servis realizuje sa neadekvatno dimenzionisanim performansama. Međutim, odsustvo uniformnog rešenja za problem mapiranja klasa servisa kao i nepostojanje opšte prihvaćenosti primene mapiranja klasa servisa je realnost. Imajući u vidu da je mapiranje neizbežno u procesu uspostave E2E QoS kao i da primena 3P modela dovodi do visokog nivoa fleksibilnosti, važno je implementirati fer i fleksibilnos mapiranja klasa servisa.

Sa aspekta algoritama za mapiranje klasa servisa u 3P modelu ugovaranja predloženi algoritmi u [132], [133], [134] i [135] otkrivaju prednosti mapiranja baziranog na 3P modelu posmatrajući fleksibilnost u selekciji klasa, kao i saglasnost sa različitim QoS modelima. U ovom radu svaki od predloženih algoritama u ovim radovima će biti detaljno analitički analiziran i biće data komparativna analiza E2E performanse u zavisnosti od selekcije klasa servisa.

3P agent sakuplja i čuva informacije vezane za ponude pojedinačnih domena koje su formulisane kroz diferencijaciju klasa servisa. Svaka klasa servisa je definisana u skladu sa sledećim parametrima: (1) mere performansi kao što su kašnjenje, džiter i PLR; (2) cena interkonekcije koja podrazumeva operativne troškove generisanja, transporta i terminiranja saobraćaja (može se iskazati kao cena po jedinici kapaciteta).

Pretpostavlja se da su $\mu_{j.req}$ i $\mu_{j.off}$ zahtevane i ponuđene E2E vrednosti mere performansi j , respektivno. U ovom radu se usvaja konvencija da indeks $j=1,2,3$ označava kašnjenje, džiter i PLR, respektivno. Osim konvencije za QoS parametre, koji su osnovni ulazni parametri u procesu selekcije klasa servisa i njihovog mapiranja, dalje se definiše parametar kojim se ocenjuju algoritmi klasa servisa - *Degree of Correspondence (DC)*. DC parametar se može definisati na E2E i na domenskom nivou u zavisnosti od prirode QoS parametra. Konkretno, kašnjenje predstavlja aditivnu meru, džiter je u najgorem slučaju

aditivan, dok je PLR indirektno multiplikativan. DC za QoS parametar j na E2E nivou se računa u skladu sa sledećim jednakostima:

$$DC_j^{E2E} = \frac{\mu_{j,req}}{\mu_{j,off}}, j = 1,2 \quad (4.1)$$

$$DC_j^{E2E} = \left| \frac{\log(\mu_{j,off})}{\log(\mu_{j,req})} \right|, j = 3 \quad (4.2)$$

Vrednost $DC_j^{E2E} = 1$ predstavlja potpuno poklapanje zahtevane i ponuđene E2E vrednosti za meru performanse j . Sa druge strane, $DC_j^{E2E} < 1$ predstavlja scenario kada su ponuđene performanse lošije u odnosu na zahtevane i slično važi za $DC_j^{E2E} > 1$ gde se podrazumeva da su ponuđene bolje u odnosu na zahtevane performanse.

Razmatra se generalni slučaj kada 3P agent upravlja ugovaranjem grupe od N domena. Domen d ($d=1,2,...N$) specificira K^d klasa servisa. Ponuda domena d u kontekstu klase k ($k=1,2,...,K^d$) je izražena skupom parametara $\mu_j^d \mid j = 1,2,3$, dok je cena interkonekcije za klasu k data parametrom c_k^d . U skladu sa prethodnim definicijama $\mu_{j,k}^d$ označava meru performansi j , klasu servisa k u okviru domena d .

Slično E2E definiciji DC parametara, definiše se DC parametar na nivou pojedinačnog domena koji je jednak:

$$DC_{j,k}^d = \frac{\mu_{j,req}^d}{\mu_{j,k}^d}, j = 1,2 \quad (4.3)$$

$$DC_{j,k}^d = \left| \frac{\log(\mu_{j,k}^d)}{\log(\mu_{j,req}^d)} \right|, j = 3 \quad (4.4)$$

gde je $\mu_{j,req}^d$ zahtevani nivo performansi j od domena d .

U skladu sa uvedenim definicijama QoS parametara, cene interkonekcije kao i odgovarajućih indeksa koji označavaju domene u lancu, klase servisa u okviru domena i odgovarajući parametar dalje se primenjuju algoritmi mapiranja klasa servisa.

4.8. P-CMS algoritam mapiranja klasa servisa

Policy-based Conformance Matching Scheme (P-CMS) je predložen kao algoritam mapiranja klasa servisa za upotrebu u 3P modelu [132], [133] i predstavlja trenutno jedini predloženi algoritam u literaturi namenjen za upotrebu u 3P modelu ugovaranja. Ovaj algoritam omogućava primenu nekoliko politika mapiranja koji uzimaju u obzir zahtevane QoS parametre.

Bitan aspekt P-CMS algoritma mapiranja predstavlja marigna pogoršanja (*impairment budget*) koja se definiše kao ukupna vrednost dozvoljenog pogoršanja E2E performansi kada se razmatraju kašnjenje, džiter i PLR. Ova margina se može na odgovarajući način podeliti između mrežnih segmenata, odnosno u slučaju interdomenskog QoS, između domena. ITU-T preporuka Y.1542 [136] pretpostavlja postojanje nekoliko pristupa za alokaciju margine pogoršanja performansi i to statički, pseudo-statički, signalizacioni i akumulacioni pristup.

Statički pristup podrazumeva alokaciju fiksne vrednosti margine pogoršanja performansi između domena. Razlika između hijerarhijskog nivoa domena se obično uzima u obzir prilikom alokacije pogoršanja performansi. Bez obzira na jednostavnost statičkog pristupa alokacije, statička alokacija margine pogoršanja može rezultirati lošim dimenzionisanjem resursa na E2E putanji s obzirom na različiti broj domena koji učestvuju u E2E servisu.

Kod pseudo-statičkog pristupa alokaciji, informacije o tačnom broju domena na putanji se moraju signalizirati između domena. Pseudo-statički pristup podrazumeva određeni stepen fleksibilnosti jer se u zavisnosti od servisnih zahteva može izvršiti realokacija vrednosti pogoršanja performansi.

Signalizacioni pristup podrazumeva dinamičku alokaciju margine pogoršanja između domena, pa je stoga potrebna implementacija odgovarajućeg signalizacionog protokola za te potrebe koji bi razmenio neophodne informacije o alokaciji margine između domena koji učestvuju u uspostavi E2E servisa.

Akumulacija margine pogoršanja podrazumeva da se alokacija oslanja na domenske ponude u kontekstu performansi, a ne na eksplicitnu alokaciju pojedinačnih vrednosti margina pogoršanja performansi svakom domenu. Uspostava poznatog skupa klasa servisa između domena doprinela bi realizaciji ovog pristupa alokaciji margine pogoršanja i načinila bi da proces mapiranja klasa servisa ne bi bio neophodan. Međutim, iako postoji otvorena potreba za definisanjem poznatog skupa klasa servisa u okviru domena, ovaj pristup nije izvodljiv imajući u vidu heterogenu strukturu i uređenje domena kao i dinamičko tržište telekomunikacionih servisa.

S obzirom da se kao jedan od većih problema u procesu mapiranja klasa servisa kao i alokacije margine pogoršanja performansi pominje neadekvatno dimenzionisanje resursa u okviru pojedinačnih domena, P-CMS algoritam je predložen sa ciljem da reši pomenuti problem. P-CMS ispunjava sledeće karakteristike bitne za primenu u okviru 3P modela i za proces mapiranja klasa servisa između domena:

- Nezavisnost od QoS arhitekture predstavlja potencijal za upotrebu algoritma mapiranja u velikom broju scenarija;
- Nizak nivo implementacione kompleksnosti podrazumeva da algoritam mapiranja ne troši previše resursa na rad, a i za njegovu primenu nije potrebno izvršiti redefiniciju arhitekture ugovaranja;

- Algoritam mapiranja treba da sprovede fer uslove prema svim domenima koji učestvuju u realizaciji E2E servisa;
- Dinamička adaptacija na mehanizme selekcije putanja podrazumeva da se poslovne politike koje proističu iz primene ovog algoritma mapiranja mogu i odnositi na upotrebu mehanizama posredstvom kojih se putanja E2E servisa menja;
- Skalabilnost kao jedna od najbitnijih karakteristika podrazumeva da algoritam nije ograničen brojem domena koji učestvuje u E2E realizaciji iako ovo ograničenje nije rešeno primenom ovog algoritma mapiranja već centralizovanjem modela ugovaranja.

Analitički posmatrana margina pogoršanja za meru performanse j i zahtevanu vrednost iste $\mu_{j,req}$ na putanji od N domena se daje sledećim izrazom za aditivnu i indirektno multiplikativnu meru performanse, respektivno:

$$\mu_{j,req} = \sum_{d=1}^N \mu_{j,req}^d \quad (4.5)$$

$$\mu_{j,req} = 1 - \prod_{d=1}^N (1 - \mu_{j,req}^d) \quad (4.6)$$

Kako bi se ispunili gore navedeni ciljevi, a naročito ostvarila alokacija margine pogoršanja performansi po domenu, pretpostavlja se uvođenje generičke politike za alokacije margine po domenu $\mu_{j,req}^d$. Ove margine se određuju u skladu sa sledećim izrazima u zavisnosti od aditivnosti ili indirektno multiplikativnosti mere performansi za koju se određuje, respektivno:

$$\mu_{j,req}^d = \mu_{j,req} \cdot \phi_j^d \quad (4.7)$$

$$\mu_{j,req}^d = 1 - (1 - \mu_{j,req})^{\phi_j^d} \quad (4.8)$$

U prethodnim jednačinama ϕ_j^d predstavlja faktor pogoršanja za domen d i meru j koja mora da zadovolji trivijalnu nejednakost $\sum_{d=1}^N \phi_j^d \leq 1$. Generalizovani format jednačine za faktor ϕ_j^d glasi

$$\phi_j^d = \frac{1}{N} \rho_j^d \sum_{k=1}^N w_k^d \varphi_{j,k}^d, \quad K \geq 1 \quad (4.9)$$

gde su $k=1$ i $k=2$ indeksi koji određuju tip domena i QoS granularnost, respektivno. Koeficijenti $\varphi_{j,k}^d$ za domen d koji mogu da određuju tip domena (pristupni ili tranzitni domen), QoS granularnost ili neku administrativnu karakteristiku odnosno ograničenje pri čemu se mora zadovoljiti uslov $\sum_{d=1}^N \varphi_{j,k}^d = N$.

Dalje je w_k^d težinski faktor $\sum_{k=1}^{K^d} w_k^d = 1$ koji određuje doprinos pojedinačnog koeficijenta $\varphi_{j,k}^d$ ukupnoj vrednosti faktora pogoršanja ϕ_j^d . I konačno, ρ_j^d predstavlja regulacioni faktor za koji važi $\sum_{d=1}^N \rho_j^d \leq 1$ i koji odražava kratkoročne karaktersitike domena d na primer nivo iskorišćenja resursa i sl..

Na osnovu prethodno definisanog faktora pogoršanja ϕ_j^d koji suštinski uzima u obzir alokaciju margine pogoršanja performansi u zavisnosti od zahtevanih QoS parametara, dalje se u okviru P-CMS algoritma mogu specificirati različite politike alokacije i to *Equal Request* (ER), *Domain Hierarchy* (DH) i *Domain Hierarchy and QoS Granularity* (DH+QG).

ER politika predstavlja najjednostavniju P-CMS politiku alokacije za koju važi sledeće: $\rho_j^d = 1$, $K = 1$, $\varphi_{j,1}^d = 1$ i $w_1^d = 1$ za svaki domen d . Posledično, sledi da je faktor pogoršanja $\phi_j^d = \frac{1}{N}$.

DH politika uzima u obzir hijerarhijsku organizaciju domena, pa se posledično vodi računa o alokaciji vrednosti margine pogoršanja u skladu sa nivoom koji domen zauzima u hijerarhiji. U okviru ove politike, $\rho_j^d = 1$, $K = 1$, $\varphi_{j,1}^d = 1$ i $w_1^d = 1$ za svaki domen d dobija se da je $\phi_j^d = \frac{1}{N} \varphi_{j,1}^d$. Jedan od mogućih načina određivanja $\varphi_{j,1}^d$ može da se izrazi preko sledećeg generalizovanog izraza:

$$\varphi_{j,1}^d = \frac{N \cdot t_j^d}{\sum_{d=1}^N N_t \cdot T_{t,j}} \quad (4.10)$$

gde je T ukupan broj različitih tipova domena, N_t broj domena istog tipa t i koji istovremeno zadovoljava uslov $\sum_{t=1}^T N_t = N$, dok je $T_{t,j}$ faktor koji je zavisan od tipa domena t i mere performansi j . Promenljiva t_j^d uzima vrednosti iz skupa $\{T_{t,j}\} | 1 \leq t \leq T$.

DH+QG politika istovremeno uzima u obzir domensku hijerarhiju kao i QoS granularnost koja se izražava posredstvom $\varphi_{j,2}^d$ parametra. Pretpostavljajući da je $\rho_j^d = 1$, $K = 2$, dobija se da je faktor jednak

$$\phi_j^d = \frac{1}{N} (w_1 \cdot \varphi_{j,1}^d + w_2 \cdot \varphi_{j,2}^d) \quad (4.11)$$

P-CMS pretpostavlja da se treba alocirati veća margina pogoršanja domenima koji imaju bolju QoS granularnost. Drugim rečima, relaksiranjem servisnog zahteva doemnima sa boljom servisnom granulacijom se omogućava bolja efikasnost u korišćenju raspoloživih resursa tih domena. Jednostavnim analitičkim pristupom se može odrediti korelacija između dodeljene margine pogoršanja i broja klasa servisa koji su definisani. Neka je C_{tot} ukupan broj klasa servisa u svim domenima, dok je C_{avg} prosečan broj klasa po domenu koji je jednak $C_{avg} = C_{tot}/N$. Ako se pretpostavi da je K^d tačan broj klasa servisa u domenu d , tada se koeficijent $\varphi_{j,2}^d$ određuje u skladu sa sledećim izrazom:

$$\varphi_{j,2}^d = \frac{C_d}{C_{avg}} = \frac{N \cdot C_d}{C_{tot}}, \quad \forall j \quad (4.12)$$

Regulacioni faktor ρ_j^d bi trebalo da se koristi u cilju korigovanja margine pogoršanja koja se pridružuje dugoročnoj politici imajući u vidu dinamičke karakteristike domena. Ovo se odnosi na nivo iskorišćenja resursa u okviru domena i uglavnom se vezuje za kapacitet. Primena ovog parametara nikako ne može da reši problem nedostatka kapaciteta u okviru jednog ili više domena, pa se upravo ova pojava rešava promenom politike rutiranja, odnosno promenom lanca domena što rezultuje ponovnim ugovaranjem E2E servisa.

Nakon definicije načina upotrebe margine pogoršanja, P-CMS dalje vrši proces selekcije klasa servisa u okviru domena imajući u vidu odgovarajuću primenu politike alociranja margine pogoršanja. Ulazni parametri P-CMS su QoS parametri koji su specificirani u okviru servisnog zahteva, informacija o lancu domena preko kojih će se E2E servis uspostaviti i na kraju, odabrana politika alokacije (ER, DH, DH+QG). P-CMS radi tako što najpre odredi, u skladu sa predodređenom politikom, alokaciju margine pogoršanja po domenima, a na osnovu toga identifikuje klase koje mogu da budu kandidati, odnosno klase koje specificiraju parametre koji su specificirani servisnim zahtevom.

Koncept rada P-CMS se zasniva da se odabere u svakom domenu klasa servisa koja će uspostaviti odgovarajuću vrednost DC parametra između QoS zahteva i ponuđenih performansi koje mogu da se realizuju posredstvom klasa kandidata. DC parametrar za klasu k klasu i meru performansi j u domenu d se određuje na osnovu izraza za aditivne i indirektno multiplikativne QoS parametre, respektivno:

$$DC_{j,k}^d = \frac{\mu_{j,req}^d}{\mu_{j,off}^{d,i}} = \frac{\mu_{j,req} \cdot \phi_j^d}{\mu_{j,off}^{d,i}} \quad (4.13)$$

$$DC_{j,k}^d = \left| \frac{\log(\mu_{j,off}^{d,i})}{\log(\mu_{j,req}^d)} \right| = \left| \frac{\log(\mu_{j,off}^{d,i})}{\log(\mu_{j,req} \cdot \phi_j^d)} \right| \quad (4.14)$$

Ukoliko mera performanse j nije specificirana u okviru klase, P-CMS elimiše tu klasu iz daljeg razmatranja i kreće da razmatra sledeću klasu. Nakon što su analizirane sve klase servisa, ukupan DC parametar se određuje za svaku klasu tako što se uzima DC parametar određen za svaku meru j :

$$DC_k^d = \frac{1}{M} \sum_{j=1}^M DC_{j,k}^d \quad (4.15)$$

U okviru koga je M ukupan broj specificiranih QoS parametara u okviru servisnog zahteva.

Konačno, selekcija klasa servisa se obavlja tako što se u okviru domena d koji ima C_d klasa odabira klasa k za koju važi sledeća jednakost:

$$|1 - DC_k^d| = \min\{|1 - DC_1^d|, |1 - DC_2^d|, \dots, |1 - DC_{C_d}^d|\} \quad (4.16)$$

P-CMS praktično vrši selekciju klasa servisa tako što u okviru domena traži klasu servisa koja je po apsolutnim vrednostima što bliža zahtevanim. Iako ovakav koncept dovodi do optimalnog mapiranja u kome se minimizuje da se na putanji dešava predimenzionisanje resursa, postoji velika verovatnoća da se ovako realizovana selekcija mapiranja klasa servisa dovesti do E2E performansi koje ne ispunjavaju servisne zahteve, što će dovesti do ponovnog ugovaranja servisa. Imajući u vidu ove nedostatke P-CMS algoritma mapiranja, u ovom radu su predloženi novi algoritmi koji rešavaju nedostatke prethodno predstavljenog algoritma mapiranja i uvode viši nivo fleksibilnosti u proces mapiranja klasa servisa između domena.

4.9. 3PIP algoritam mapiranja klasa servisa

3P-based Integer Programming (3PIP) algoritam mapiranja klasa servisa je predložen u [134]. Njegova specifikacija je zasnovana na osvrtu na nedostatke identifikovane kod P-CMS algoritma kao i na daljem unapređenju procesa mapiranja klasa servisa u 3P modelu ugovaranja servisa. Naime, preloženi metod uvodi u proces mapiranja klasa servisa i ocenu cene interkonekcije koja se razmatra kao dodatni ulazni parametar pridružen kandidat klasi servisa po domenu. Na ovaj način je moguće realizovati značajno viši nivo fleksibilnosti u implementaciji poslovnih politika.

Pre formulacije 3PIP algoritma, potrebno je da se identifikuju zahtevi koji treba da se ispune, pre svega mere performansi i cena interkonekcije. 3PIP se zasniva na pretpostavci da postoji jedan cilj koji treba da se ispuni, dok se istovremeno moraju poštovati višestruka ograničenja. Ova ograničenja se odnose na ispunjenje QoS zahteva, dok je ciljna funkcija specificira da E2E cena interkonekcije za svako mapiranje treba da bude što manja. U skladu sa ovim 3PIP predstavlja adekvatan algoritam za primenu kod selekcije klasa servisa prilikom uspostave E2E servisa koji su kritični u kontekstu performansi.

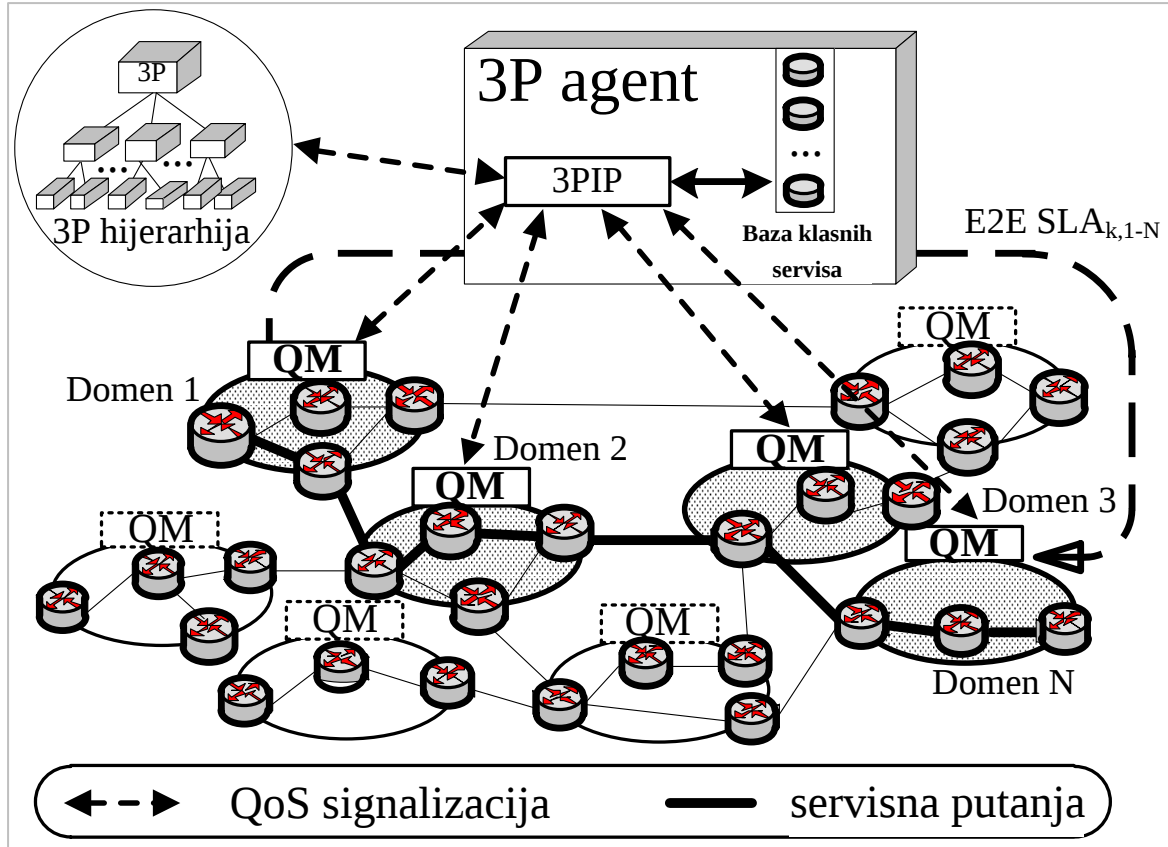
Nivo diferencijacije servisa kao i postavljeni zahtevi za performansama mogu da uzrokuju scenarije u kojima 3PIP neće moći da realizuje mapiranje kao posledica definisanih ograničenja i niskog nivoa QS granularnosti. Međutim, ova pojava se ne može smatrati nedostatkom algoritma mapiranja s obzirom da direktno proističe iz matematičkog modela na koji se algoritam naslanja i koji daje tačno rešenje, tačnije selekciju klasa servisa koja proističe iz E2E performansi.

3PIP je baziran na *Integer programming* matematičkom modelu [137] koji predstavlja optimizacioni metod koji je rešiv u polinomijalnom vremenu. Za potrebe mapiranja klasa se primenjuje ovaj model imajući u vidu ograničenja koja su data u [138]. Imajući u vidu smernice date u [139] kao i realne multidomenske scenarije [136], 3PIP algoritam predstavlja skalabilnu šemu za mapiranje klasa servisa.

U analizu se uvode promenljive y_d i x_k^d pri čemu je $d \in \{1, 2, \dots, N\}$, dok je $k \in \{1, 2, \dots, K^d\}$:

$$y_d = \begin{cases} 1, & \text{ako se domen } d \text{ nalazi na putanji} \\ 0, & \text{ako se domen } d \text{ ne nalazi na putanji} \end{cases} \quad (4.17)$$

$$x_k^d = \begin{cases} 1, & \text{ako se bira klasa } k \text{ u domenu } d \\ 0, & \text{ako se ne bira klasa } k \text{ u domenu } d \end{cases} \quad (4.18)$$



Slika 4.5 Primena 3PIP algoritma u 3P modelu ugovaranja servisa.

3P agent bira jednu klasu k za koju važi $x_k^d = 1$. Kao što je prethodno rečeno, pojedinačni domen definišu cenu interkonekcije po jedinici kapaciteta c_k^d za svaku klasu servisa. Specifikacija E2E servisnih zahteva definiše vrednosti ograničenja za pojedine QoS parametre na osnovu čega je i potrebno realizovati mapiranje klasa. Ove vrednosti su sadržane u okviru SLA za određeni E2E servis uključujući i eventualne penale koji će biti aktivirani u slučaju da dođe do potencijalne degradacije performansi. Realizacija 3PIP algoritma u 3P modelu je prikazana na slici 4.5.

Kompozicija mera performansi se koristi kako bi se utvrdio E2E kašnjenje, džiter i PLR iz ponuđenih vrednosti koje su specificirane u okviru selektovane klase u pojedinačnim domenima. Kašnjenje i džiter predstavljaju aditivne mere, dok je PLR indirektno multiplikativna mera performansi. Međutim, PLR se može takođe transformisati u aditivnu meru što će biti urađeno za potrebe ove analize.

Ograničenja se definišu u kontekstu kašnjenja, džitera i PLR parametara u zavisnosti od E2E servisnih performansi u skladu sa sledećim izrazima:

$$\sum_{d=1}^N \sum_{k=1}^{K^d} \mu_{1,k}^d x_k^d y_d \leq \mu_{1,req} \quad (4.19)$$

$$\sum_{d=1}^N \sum_{k=1}^{K^d} \mu_{2,k}^d x_k^d y_d \leq \mu_{2,req} \quad (4.20)$$

$$\sum_{d=1}^N \sum_{k=1}^{K^d} (-\log(1 - \mu_{3,k}^d)) x_k^d y_d \leq -\log(1 - \mu_{3,req}) \quad (4.21)$$

Nakon definisanja ograničenja, 3PIP vrši selekciju klasa koje odgovaraju definisanim ograničenjima pri čemu se te klase servisa biraju u skladu sa funkcijom cilja koja teži da pronade klase servisa koje će ispuniti ograničenja, a pri tome istovremeno i minimizirati cenu interkonekcije. Funkcija cilja se definiše u skladu sa sledećom jednakosti:

$$C_k(SLA_{k,1-N}) = \min \left\{ \sum_{d=1}^N \sum_{k=1}^{K^d} c_k^d \cdot x_k^d \cdot y_d \right\} \quad (4.22)$$

Minimizacija cene interkonekcije $C_k(SLA_{k,1-N})$ dovodi do toga da se projektovani profit provajdera povećava, pa stoga se fleksibilno mogu prilagoditi ekonomske politike u skladu sa strateškim ciljevima domena koji učestvuju u realizaciji E2E servisa.

Princip rada 3PIP algoritma se može formulisati u naredna tri koraka:

1. E2E servisni zahtevi i specifikacije klasa servisa u okviru pojedinačnih domena se inicijalizuju;
2. Selekcija klasa servisa se vrši posredstvom *branch and bound* matematičke relaksacije *integer programming* modela čijom se primenom povećava skalabilnost 3PIP algoritma. Rešava se funkcija cilja definisana u (4.22) imajući u vidu ograničenja formulisana u (4.19)-(4.21).
3. Ukoliko postoje mapiranja koje kao rezultat daju istu E2E cenu interkonekcije, mapiranje koje daje najbolje performanse se bira kao rezultat tako što se odabira mapiranje koje ima najveću vrednost izraza $\sum_{j=1}^3 (\mu_{j,req} / \sum_{d=1}^N \mu_{j,k}^d)$ što rezultuje mapiranjem koje daje najbolje performanse.

Potencijalna prioritizacija pojedinačnih mera performansi se može primeniti kao sredstvo razrešenja više mapiranja koje daju iste performanse. Međutim, prethodna tri navedena koraka su dovoljna za praktične scenarije.

4.10. GPMS algoritam mapiranja klasa servisa

Goal Programming-based Mapping Scheme (GPMS) algoritam je predložen u [135]. GPMS se oslanja na *multi-goal programming* model kojim se postiže da se može definisati više ciljeva uz višestruka ograničenja kao kod 3PIP. Međutim, za razliku od 3PIP, ciljna funkcija se sastoji od više ciljeva čime je povećan nivo fleksibilnosti algoritma mapiranja i samog procesa mapiranja klasa servisa. *Goal programming* (GP) metod [140] multikriterijumskog odlučivanja koji može da obezbedi rešenje u slučajevima suprostavljenih ciljeva i višestrukih ograničenja. Specifikacija ovih ograničenja i više ciljeva realizuje deterministički proces mapiranja pri čemu se postiže veća tačnost samog mapiranja.

Višestruki GP pristup podrazumeva specifikaciju više ciljeva pri čemu se moraju u svakom momentu ispuniti sva postavljena ograničenja. Predloženi metod na koji se oslanja GPMS podrazumeva pretraživanje mogućih mapiranja na posmatranoj E2E putanji. Ciljna funkcija Z sadrži zahtev za ispunjenjem svih postavljenih QoS parametara kao i minimizaciju E2E cene interkonekcije. U pogledu QoS zahteva idealno rešenje kako bi se postiglo bolje mapiranje, a samim tim i bolje poklapanje između zahtevanih i ponuđenih performansi jeste da vrednost DC parametra bude jednaka 1.

GPMS vrši pretraživanje kandidata klasa servisa na osnovu sledećih pravila:

$$\sum_{d=1}^N \sum_{k=1}^{K^d} \mu_{j,k}^d x_k^d y_d + \delta_j^- + \delta_j^+ = \mu_{1,req} \quad , j = 1, 2 \quad (4.23)$$

$$\sum_{d=1}^N \sum_{k=1}^{K^d} -\log(1 - \mu_{j,k}^d) x_k^d y_d + \delta_j^- + \delta_j^+ = -\log(1 - \mu_{1,req}) \quad , j = 3 \quad (4.24)$$

Jednakosti (4.23) i (4.24) važe za aditivne i multiplikativne mere performansi, respektivno. U ovim jednačinama δ_j^- i δ_j^+ ($\delta_j^-, \delta_j^+ \geq 0$) su pomoćne promenljive koje se nazivaju pozitivna i negativna devijacija, respektivno. Ove devijacije kvantifikuju rastojanje od zahtevane vrednosti, tj. δ_j^- označava bolju ponudu, dok δ_j^+ označava ponudu koja je lošija u odnosu na zahtevanu vrednost mere performansi j .

GPMS algoritam obuhvata ograničenja koja su definisana u zavisnosti od QoS zahteva koji su specificirani u jednačinama (4.23) i (4.24) tako što ih objedinjuje u ciljnoj funkciji i ovim ograničenjima se pridružuju težinski faktori koji su dodeljeni u skladu sa prioritetom individualnih ciljeva. Selekcija klasa servisa se realizuje tako što se teži da se minimizuju vrednosti pomoćnih promenljivih δ_j^- i δ_j^+ za sve mere performansi j . Istovremeno, selektiraju se klase koje minimizuju cenu interkonekcije preko N domena. Ciljna funkcija Z se definiše koristeći narednu jednakost:

$$Z = \min \left\{ \sum_{j=1}^3 \alpha_j \cdot (\delta_j^- + \delta_j^+) + \alpha_c \cdot \sum_{d=1}^N \sum_{k=1}^{K^d} c_k^d \cdot y_d \right\} \quad (4.25)$$

gde su α_j i α_c težinski faktori koji su dodeljeni merama performansi i ceni interkonekcije, respektivno. GPMS realizuje dodatni nivo fleksibilnosti procesu mapiranja klasa servisa posredstvom prioritizacije. Na primer, svaki domen koji inicijalizuje E2E servisne zahteve može da specificira važnost aspekta pojedinačnih mera performansi i/ili cene interkonekcije.

Osim specificikacije težinskih faktora α_j i α_c moguće je i da se definišu ograničenja vrednosti parametara devijacije δ_j^- i δ_j^+ čime se realizuje mapiranje kojim se dobija tačno određeni bolji ili lošiji opseg u odnosu na zahtevani. Uspostavljanjem fleksibilnog okruženja u kontekstu mapiranja klasa servisa 3P agent može da vrši mapiranje klasa servisa koje bolje odgovara domenskim poslovnim ciljevima držeći se administrativno definisanih ograničenja.

4.11. Kvalitativna komparacija algoritama za mapiranje klasa između domena

Algoritmi za mapiranje klasa servisa kod 3P modela ugovaranja servisa se mogu uporediti sa različitim aspektima imajući u vidu pristupe koje zauzimaju prilikom procesa mapiranja. U Tabeli 4.1 su prikazani relevantni parametri za ocenu algoritama mapiranja klasa servisa.

Razmatranje cene interkonekcije prilikom mapiranja predstavlja važan aspekt evaluacije pojedinačnih algoritama mapiranja. Naime, uključivanjem cene interkonekcije kao dodatnog ulaznog parametra se omogućava planiranje poslovnih politika s obzirom da postoji korelacija između cene interkonekcije i E2E performansi. Predloženi 3PIP i GPMS algoritmi razmatraju cenu interkonekcije i kao cilj mapiranja vrše selekciju klasa kako bi se postigla minimalna cena interkonekcije. Za razliku od ova dva algoritma, P-CMS algoritam ne podržava uključivanje cene interkonekcije u proces selekcije klasa servisa što predstavlja značajan nedostatak ovog algoritma.

Tabela 4.1 Relevantni parametri ocene algoritama mapiranja klasa servisa

Algoritam mapiranja	Cena interkonekcije	Alokacija margine pogoršanja performansi	Formulacija zahteva na nivou domena	Nivo fleksibilnosti
P-CMS	✗	Signalizacioni	✓	Srednji
3PIP	✓	Akumulacija	✗	Nizak
GPMS	✓	Akumulacija	✗	Visok

Planiranje alokacije margine pogoršanja predstavlja unošenje dodatnog nivoa robusnosti u proces mapiranja s obzirom da se u slučaju eventualnog podbacivanja u garanciji performansi od strane jednog domena u lancu moguće da zbog uračunavanja viška performansi E2E servis ne bude ugrožen. P-CMS

algoritam realizuje signalizacioni pristup alokaciji margine pogoršanja. Signalizacioni pristup je potreban kako bi 3P agent obavestio svaki domen u lancu koji je deo margine alociran posmatranom domenu. Za razliku od P-CMS, 3PIP i GPMS algoritmi podrazumevaju akumulacioni pristup gde se alokacija obavlja na E2E nivou i kao takva se može specificiranom politikom dodeliti pojedinačnim domenima u skladu sa administrativnom politikom. Posmatrajući alokaciju margine pogoršanja, P-CMS ostvaruje prednost po pitanju alokacije margine pogoršanja jer se alokacija određuje u skladu sa domenskom hijerarhijom i nivoom QoS granularnosti.

Formulacija zahteva na nivou domena predstavlja način rada algoritma mapiranja koji može da definiše servisne zahteve na nivou domena kako bi se tačno utvrdio koje performanse se zahtevaju od posmatranih domena. Na ovaj način je moguće implementirati odgovarajuće ekonomske politike i uvesti mehanizam kazni i nagrada za kršenje i prekoračenje performansi E2E servisa.

Aspektom fleksibilnosti se opisuje koliko je lako modifikovati algoritam mapiranja klasa servisa u cilju prilagođenja za poslovne politike. 3PIP i GPMS kao posledica oslanjanja na ciljne funkcije moguće je da definišu jedan i više ciljeva, respektivno, koji se može prilagoditi ekonomskim ciljevima i različitim multidomenskim scenarijima. Kod GPMS se prioritizacijom parametara može uračunati važnost pojedinih aspekata i njihova korelacija sa planiranim ciljevima mapiranja. P-CMS sadrži brojne parametre čijim se modifikacijama može kontrolisati proces mapiranja i alokacije margine pogoršanja. Takođe, identifikacijom nedostataka P-CMS algoritam, predloženi su algoritmi mapiranja koji su rešili postojeće nedostatke i uveli dodatne funkcionalnosti i prioritizaciju koja dozvoljava planiranje poslovnih ciljeva u uspostavi E2E servisa i kontrolisani proces mapiranja klasa servisa.

5

ANALIZA PREDLOŽENIH INTRADOMENSKIH I INTERDOMENSKIH QoS MEHANIZAMA

U ovom poglavlju se predstavljaju rezultati dobijeni praktičnim i analičkim ispitivanjima predloženih mehanizama za popravljjanje QoS nivoa u intradomenskim i interdomenskim orkuženjima. Verifikacija doprinosa QoS će biti pokazana tako što će biti izvršena komparacija postojećih i predloženih rešenja kako bi se pokazala dominantnost mehanizama koji su analizirani i predloženi u ovom radu.

U kontekstu intedomenskih mehanizama, komparativna analiza predloženih i postojećeg algoritma mapiranja klasa servisa će biti pokazana sa aspekta posmatranja DC parametra u multidomenskim scenarijima.

5.1. Komparativna analiza strategija paketske disperzije

Strategije paketske disperzije koje su analizirane u trećem poglavlju se evaluiraju iz perspektive strukture gubitaka paketa posredstvom realizacije praktičnih testova. Testiranja se obavljaju na realnoj mrežnoj opremi analizirajući *voice* pakete iz 250 VoIP poziva različite dužine trajanja koji su podvrgnuti različitoj strukturi rafalnih gubitaka paketa. Ispituje se uticaj pojedinačnih strategija paketske disperzije kada se

koristi promenljivi broj disperzionih putanja na strukturu gubitaka paketa i na VoIP kvalitet. Takođe, kako bi se verifikovala tačnost prethodno predloženog analitičkog modela za utvrđivanje rastojanja gubitaka paketa upoređuju se rezultati dobijeni praktičnim testiranjem i korišćenjem predloženog modela.

Paketska disperzija se obavlja na ruteru koji je realizovan kao *Linux* server na kome se posredstvom izmene politike rutiranja primenjuju različite strategije rutiranja – *round robin*, *random*, adaptivna i redundantna. VoIP pozivi su realizovani korišćenjem različitih VoIP terminala, kako softverskih tako i hardverskih rešenja. Naime, za potrebe testova korišćeni su VoIP terminali tipa *Cisco IP phone SPA450G*, dok se kao softversko rešenje koristio *Cisco IP Communicator*. Oba modela implementiraju PLC algoritme sa adaptivnim dedžiterizacionim baferom koji smanjuje uticaj gubitaka paketa i džitera na kvalitet, respektivno. U cilju postizanja realnijih uslova testiranja, testirani VoIP terminali su registrovani na *Cisco 3725 Multiservice Access Router* sa VoIP servisom. Ovaj ruter je sposoban da se ponaša kao VoIP centrala, koji je sposoban da terminira i rutira VoIP signalizaciju.

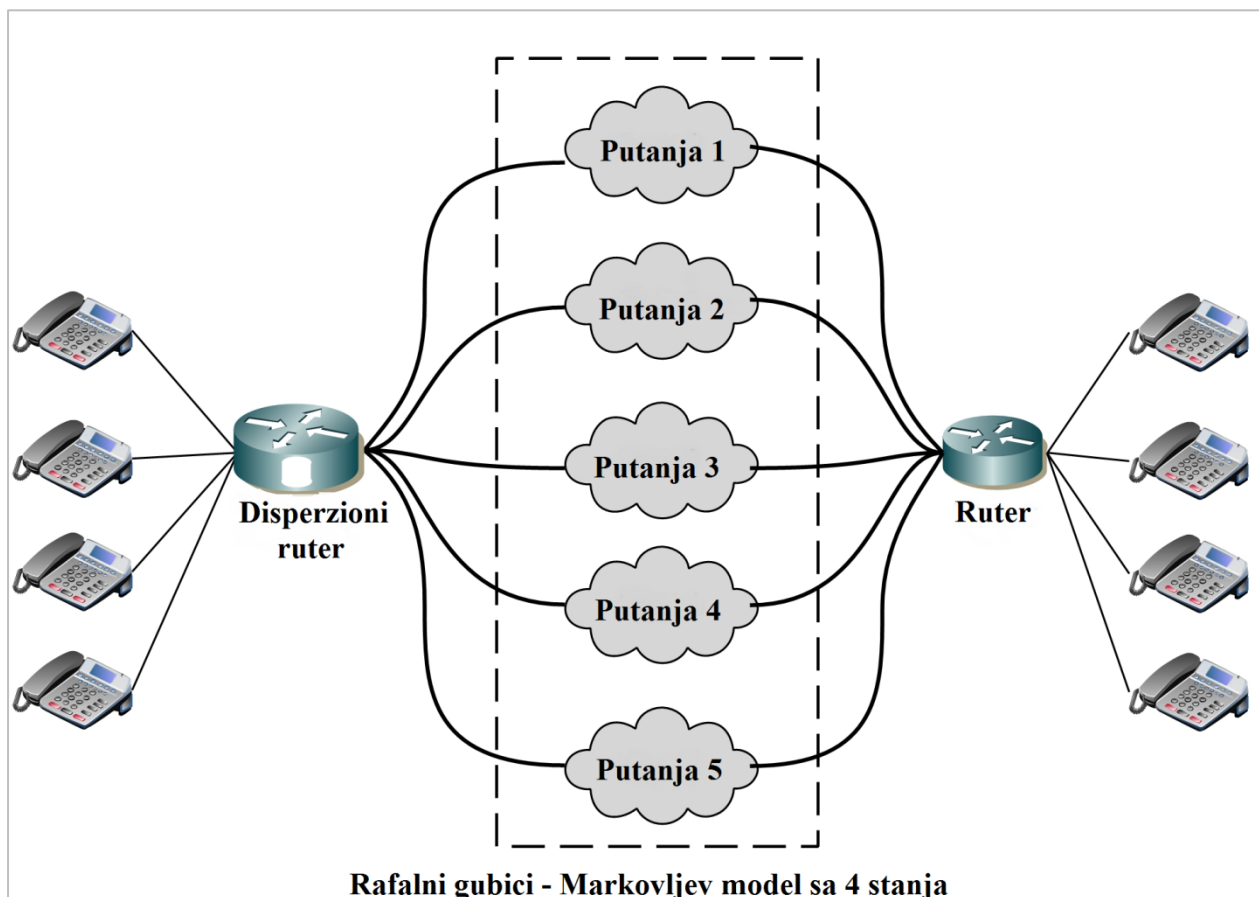
Uspostavljeni VoIP pozivi koriste G.711 kodek koji važi za visoko kvalitetni kodek u VoIP terminologiji i koji se često koristi za realizaciju VoIP servisa. Analiza je obuhvatila primenu samo ovog tipa kodeka pri čemu se ne smanjuje tačnost ove studije s obzirom da se posmatra struktura gubitaka paketa i određuje VoIP kvalitet. Slična analiza bi mogla da se realizuje i sa ostalim tipovima kodeka i zaključci studije bi bili slični, ako ne i isti.

Analiza VoIP kvaliteta je bazirana na razmatranju ITU-T E-modela gde se ulazni parametri određuju na osnovu rezultujuće strukture gubitaka paketa i predodređenih vrednosti pridruženih specifičnom tipu kodeka, konkretno u slučaju ovog ispitivanja, kodeka G.711 sa implementiranim VAD i PLC algoritmom.

Skorašnja modifikacija NETEM modula [141] je integrisana u okviru disperzionog rutera i koristi se za realizaciju Markovljevog modela sa 4 stanja kao i za unošenje kašnjenja na disperzionim putanjama do odredišta. Na ovaj način je moguće menjati QoS mrežnog okruženja i na fleksibilan način testirati različite scenarije do kojih može doći na putanjama. Predloženi analitički model određivanja rastojanja gubitaka paketa se testira posredstvom MATLAB paketa pri čemu se odgovarajućim skriptama koje se zasnivaju na predloženom modelu, određuje i struktura gubitaka paketa iz snimljenih VoIP poziva. Testno okruženje je prikazano na slici 5.1 u okviru koga se može videti način realizacije VoIP poziva i implementacija paketske disperzije.

Kao što je već navedeno, razmatranje samo jednog parametra kojim bi se opisali gubici paketa nije dovoljno za tačnu analizu i uvid u uticaj pojave gubitaka paketa na VoIP servis. Posledično, sprovedena analiza se sastoji od nekoliko delova: (1) verifikacija predloženog modela određivanja rastojanja gubitaka za Markovljev model sa 4 stanja; (2) komparacija rezultata dobijenih praktičnim testiranjima u kontekstu

uticaja strategija paketske disperzije na povećanje rastojanja gubitaka paketa; (3) komparacija VoIP kvaliteta za različite strategije paketske disperzije.



Slika 5.1 Testno okruženje za testiranje uticaja strategija paketske disperzije na rafalne gubitke.

5.1.1. Verifikacija rezultata predloženog analitičkog modela

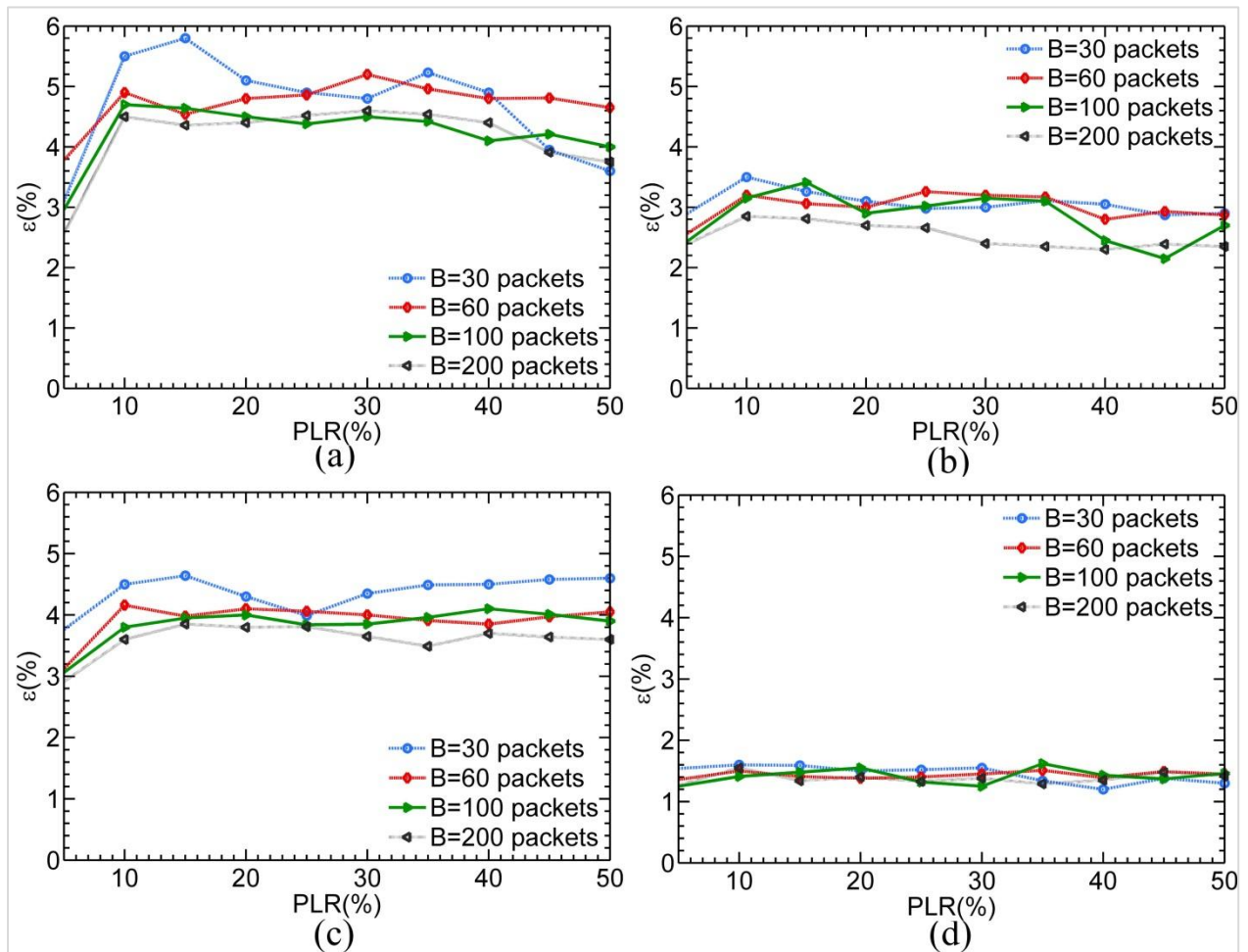
U cilju sagledavanja tačnosti i podobnosti analitičkog pristupa za različite strategije, realizuje se komparacija NLR parametra čije vrednosti se dobijaju iz predloženog analitičkog modela i praktičnih testova.

Relativna greška ε se definiše kao mera tačnosti i to kao

$$\varepsilon = \left| \frac{\frac{1}{F} \sum_{i=1}^F NLR_{msrm,i}}{NLR_{analyt}} \right| \quad (5.1)$$

gde $NLR_{msrm,i}$ i NLR_{analyt} označavaju izmerene i izračunate NLR vrednosti, respektivno, dok F označava broj realizovanih merenja kako bi se dobila srednja vrednost NLR parametra. Za potrebe verifikacije rezultata i određivanje nivoa tačnosti predloženog modela, uzima se pet merenja, tačnije $F=5$.

Pretpostavlja se da postoje dve disperzione putanje. Prva putanja ima karakteristike $PLR=5\%$ i $B=30$ paketa, dok struktura gubitaka paketa na drugoj putanji se menja u širokom opsegu kako bi se ispitao uticaj različite strukture gubitaka paketa na tačnost modela određivanja rastojanja gubitaka paketa. Komparacija analitičke tačnosti i samim tim verifikacija tačnosti modela je prikazana na slici 5.2 za analizirane strategije. Međutim, kao što se vidi sa slike, postoje značajne razlike u tačnosti modela između posmatranih strategije disperzije. Relativna greška za sve strategije je manja od 6% što je dovoljna tačnost za dokazivanje validnosti analitičkog modela.



Slika 5.2 Nivo tačnosti opisan posredstvom relativne greške ε : (a) Round robin; (b) Random; (c) Adaptivna; (e) Redundantna strategija.

Round robin strategija ima najvišu vrednost relativne greške što je posledica korišćenog pristupa u analitičkom modelu za koji se pokazalo da je problem rešiv u N -dimenzionom prostoru što je iz prikazanog imalo uticaj na tačnost modela. U kontekstu random i adaptivne strategije pristup koji koristi „ekvivalentnu putanju“ zahteva razmatranje 4^N -dimenzionog prostora što posledično dovodi do eksponencijalnog povećanja računске kompleksnosti sa povećanjem broja putanja. Na osnovu rezultata prikazanih na slici 5.2 primetno je da viša računska kompleksnosti dovodi do veće tačnosti. Sa druge

strane, najviši nivo tačnosti ima redundantna strategija što je rezultat jednostavnog koncepta funkcionisanja ove strategije. Ovakva implementacija se svodi na množenje tranzicionih verovatnoća koji su pridruženi pojedinačnim putanjama koristeći Kronekerov proizvod. Ovakav pristup dovodi do relativne greške koja je manja od 2%. Takođe, redundantna strategija ne zahteva uvođenje analitički kompleksnih parametara u model kao što je to slučaj sa prethodnim strategijama disperzije.

Bitno je naglasiti da rezultati sa slike 5.2 pokazuju da se predloženi analitički model može koristiti sa visokom pouzdanošću u određivanju rezultatne strukture gubitaka paketa u zavisnosti od primenjene strategije disperzije. Takođe, kao što se vidi, ne postoji jak uticaj rafalnih gubitaka paketa na tačnost modela.

5.1.2. Analiza rastojanja gubitaka paketa

NLR se određuje iz snimljenog VoIP saobraćaja u zavisnosti od broja putanja i primenjene strategije disperzije. Za potrebe analize rastojanja gubitaka paketa se usvaja vrednost granice primetnih gubitaka $\delta=3$ paketa jer se ovim usvajanjem razmaraju PLC algoritmi koji ne poseduju visok nivo efikasnosti skrivanja gubitaka paketa. Pretpostavljajući postojanje dve putanje sa različitim PLR parametrima, na slici 5.3 je prikazana komparacija NLR parametara kada se primenjuje *round robin*, *random*, adaptivna i redundantna strategija. Svaka pojedinačna strategija koja je ispitivana je poređena sa konvencionalnim rutiranjem preko jedne putanje.

Jasno se vidi iz rezultata analize da se primenom paketske disperzije postižu primetno bolji rezultati u povećanju rastojanja između gubitaka paketa u odnosu na rutiranje preko jedne putanje (slika 5.3 (a) i slika 5.3(b)). Komparacija rezultata relevantnih za posmatranje rastojanja gubitaka paketa kada se primenjuje *round robin*, *random* i adaptivna strategija je prikazana na slici 5.3(c)-(e). *Random* disperzija realizuje veća rastojanja između gubitaka paketa u odnosu na *round robin* strategiju što je posledica činjenice da *random* disperzija u skladu sa pridruženom verovatnoćom po jednoj putanji dozvoljava prenošenje uzastopnih paketa, pa se samim tim izbegava pojava dužih rafala gubitaka paketa kada se primenjuje ova strategija. Očekivano je da adaptivna strategija izražava bolje performanse u pogledu rastojanja gubitaka u odnosu na *round robin* i *random* strategije imajući u vidu dinamičku promenu verovatnoće pridružene putanjama koja zavisi od PLR karakteristika izraženih na istim. Po rezultatima sa slike 5.3 se vidi da primenom adaptivne strategije je moguće ostvariti bolje performanse i do 85% u odnosu na *random* strategiju.

S obzirom da redundantna strategija ima mogućnost da se izbori sa najgorijim scenarijima gubitaka paketa na putanjama, sprovedena analiza podrazumeva posmatranje zavisnosti NLR parametra kada se koriste dve putanje sa jednakim PLR karakteristikama od promene vrednosti granice primetnih gubitaka paketa δ . Razmatranje šireg opsega vrednosti granice primetnih gubitaka praktično omogućava ispitivanje ponašanja PLC algoritama različitih efikasnosti skrivanja bliskih gubitaka paketa. Značajno poboljšanje u vidu

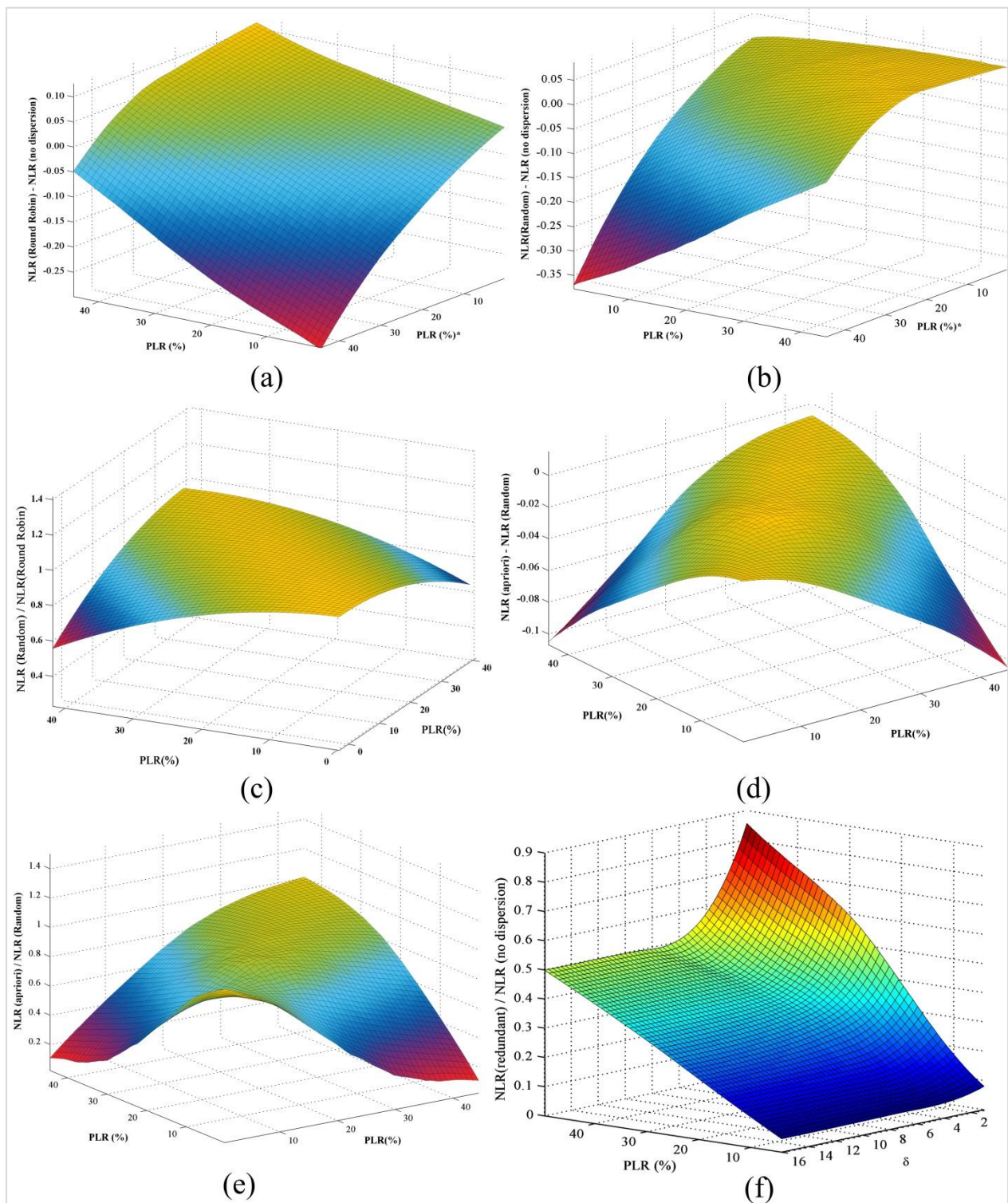
strukture gubitaka paketa se realizuje primenom ove redundantne strategije, što je primetno na slici 5.3(f), postižući 10-90% niži NLR u odnosu na konvencionalno rutiranje.

Detaljnijim uvidom rezultata dobijenih praktičnim testiranjem VoIP poziva sa slike 5.3(a)-(e) vidi se da se ostvaruje mali ili zanemarljiv NLR dobitak kada je struktura gubitaka paketa na putanjama slična ili ista. Posledično se identifikuje problematičan scenario u okviru koga se ispoljava ozbiljan nedostatak koji predstavlja ograničenje efikasnosti ispitivanih strategija paketske disperzije. Samim tim je potrebno dodatno analizirati domen delovanja ovog ograničenja i identifikovati u okviru kojih strategija dati problem može da utiče na eventualnu praktičnu primenu paketske disperzije. Pitanje upotrebe paketske disperzije u slučaju pojave ovog scenarija je naročito ozbiljan kada se uzme u obzir da u uslovima sličnih struktura gubitaka paketa na putanjama kao rezultat primene paketske disperzije mogu dobiti i nepovoljnija struktura u odnosu na konvencionalno rutiranje. U skladu sa tim, u nastavku se realizuje analiza ovog ograničenja kada se primenjuju prethodno pomenute strategije u prisustvu sličnih karakteristika gubitaka na putanjama. Kao posledica defincije adaptivne strategije, u uslovima sličnih karakteristika gubitaka paketa, ova strategija se svodi na *random* strategiju. Na slici 5.4 su prikazani NLR rezultati kada se koristi i do 5 putanja sa istom rafalnom strukturom gubitaka paketa za *round robin*, *random* i redundantnu strategiju.

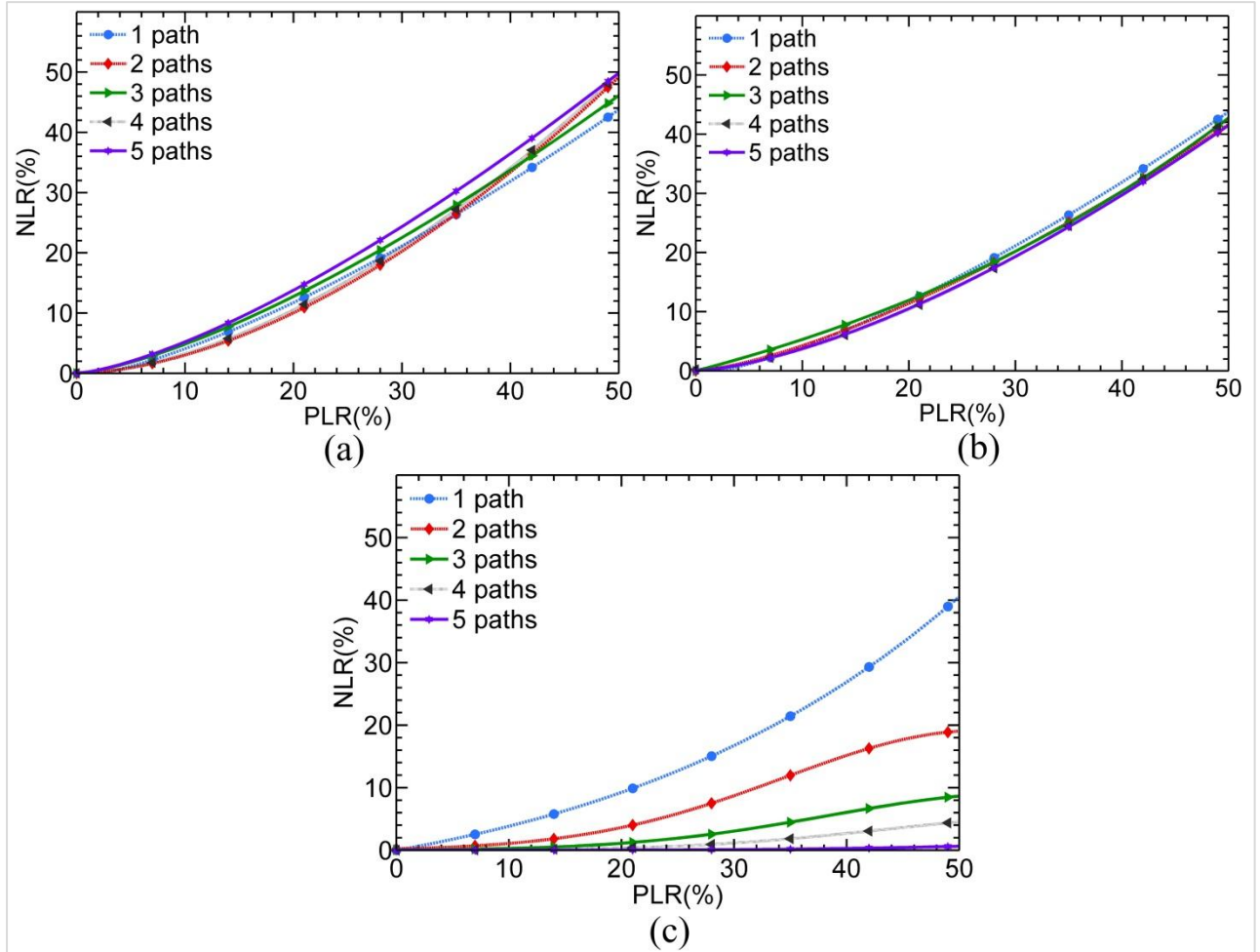
Na osnovu prikazanih rezultata može se primetiti da *round robin* i *random* strategije ne ostvaruju značajno poboljšanje u vidu rastojanja gubitaka paketa. Takođe, *round robin* strategija može da dovede i do lošijih rezultata u odnosu na konvencionalno rutiranje u slučaju ovog scenarija. Imajući ovo u vidu, predlaže se da se u slučajevima sličnih karakteristika gubitaka paketa na putanjama, izbegne primena *round robin* strategije. Sa druge strane, primenom *random* strategije se ostvaruje minimalno poboljšanje NLR vrednosti u odnosu na *round robin* i to za visoke vrednosti PLR, dok za niske vrednosti se ne izražava poboljšanje u cilju razdvajanja i povećavanja rastojanja između gubitaka.

Sprovedena analiza u cilju identifikovanja nedostataka paketske disperzije treba da pokaže da li se povećanjem broja disperzionih putanja eliminiše pomenuto ograničenje i da se vidi kako promena broja putanja utiče na NLR vrednost. Iz rezultata sa slike 5.4 se vidi da za *round robin* i *random* strategije ne postoji nikakav uticaj povećanja broja disperzionih putanja na NLR vrednost.

Za razliku od *round robin* i *random* strategija, redundantna strategija podrazumeva replikaciju paketa preko disperzionih putanja bez obzira na sličnost karakteristika gubitaka na tim putanjama. Kod redundantne strategije posledično dolazi do smanjivanja PLR vrednosti kada se poveća broj putanja, pa samim tim i rastojanje gubitaka se povećava jer ih posledično ima manje. Nivo kvaliteta koji se planira za VoIP servis će se dobiti primenom redundantne strategije i upotrebom odgovarajućeg broja disperzionih putanja.



Slika 5.3 NLR analiza rastojanja gubitaka: (a) Razlika round robin i konvencionalnog rutiranja; (b) Razlika random i konvencionalnog rutiranja; (c) Odnos random i konvencionalnog rutiranja; (d) Razlika adaptivne and random strategije; (e) Odnos adaptivne i random strategije; (f) Redundantna strategija za širok opseg δ .



Slika 5.4 Komparacija strategija disperzije u pogledu rastojanja gubitaka paketa u uslovima postojanja dve putanje sa istim karakteristikama gubitaka ($B=100$ paketa): (a) Round Robin; (b) Random; (c) Redundantna.

5.1.3. Analiza VoIP kvaliteta

Kvalitativna analiza obuhvata posmatranje individualnih strategija i njihov efekat na VoIP kvalitet u skladu sa E-modelom. Prethodno realizovane analize su posmatrale isključivo tehničke parametre i to parametre koji opisuju gubitke paketa. Posledično, razumno je da se kao metod ocene VoIP kvaliteta odabere objektivna metoda određivanja kvaliteta u okviru koje se koriste prethodno analizirani parametri. Rezultat E-modela je izražen kao R faktor čija je vrednost u opsegu 0-100. Ovaj faktor se konvertuje u MOS ocenu VoIP kvaliteta koja je izražena u opsegu od 1 do 5 koji određuje ocenu lošeg odnosno visokog kvaliteta. Relacija između R faktora i MOS parametra je data u nastavku [67]:

$$MOS = \begin{cases} 1, & R < 0 \\ 1 + 0.035 \cdot R + R \cdot (R - 60) \cdot (100 - R) \cdot 7 \cdot 10^{-6}, & 0 \leq R \leq 100 \\ 4.5, & R > 100 \end{cases} \quad (5.2)$$

VoIP kvalitet se usko povezuje sa QoS parametrima, konkretno, kašnjenjem, džiterom i procentom izgubljenih paketa. Nivo uticaja pojedinačnih QoS parametara na kvalitet VoIP servisa je različit, pa samim tim pojedinačni QoS parametri na različit način utiču na percepivani kvalitet. Realizovana studija o zajedničkom i pojedinačnom uticaju na kvalitet je prikazana u [95]. Na osnovu te studije se zaključuje da u prisustvu gubitaka paketa degradacija VoIP kvaliteta kao posledica pojave kašnjenja i džitera je skoro neprimetna u odnosu na degradaciju kvaliteta kao posledica pojave izgubljenih paketa. Posledično se može zaključiti da pojava gubitaka paketa ima dominantan uticaj na promenu nivoa percepiranog kvaliteta što ide u prilog primeni paketske disperzije koja ima tendenciju da smanji uticaj gubitaka paketa na kvalitet na osnovu pogoršanja drugih QoS parametara čiji je uticaj u poređenju sa izgubljenim paketima zanemarljiv. U okviru rada u [64] je pokazano da kašnjenje kao QoS parametar ima ograničenu vrednost ispod koje je uticaj kašnjenja na VoIP kvalitet niskog nivoa. Naime, jednosmerno kašnjenje manje od 177 ms ne degradira kvalitet razgovora dozvoljavajući realizaciju VoIP servisa bez značajnog uticaja na kvalitet. Iznad navedene vrednosti kvalitet postaje degradiran što je primetno sa slike 3.13. Što se tiče uticaja džitera na VoIP kvalitet, on se eliminiše implementacijom dedžiterizacionog bafera koga sadrži svaki moderni VoIP terminal. Primena dedžiterizacionog bafera normalno vrši „peglanje“ promenljive komponente kašnjenja tako da se paketi dalje prosleđuju VoIP aplikaciji u konstantnom protoku čime se eliminiše uticaj promenljivog kašnjenja na kvalitet. Međutim, baferovanjem paketa se unosi dodatno kašnjenje čime se povećava degradacija kvaliteta kao rezultat kašnjenja. Međutim, ukoliko je jednosmerno kašnjenje ispod kritične vrednosti, primena baferovanja paketa neće značajno uticati na kašnjenje. Kao što je rečeno, današnji VoIP terminali obavezno implementiraju ove bafere čime se eliminiše degradacija kvaliteta koja potiče od pojave promenljivog kašnjenja na putanjama dodajući odgovarajuću vrednost kašnjenja na već postojeće jednosmerno kašnjenje VoIP paketa.

Za potrebe analize VoIP kvaliteta u ovom radu se primenjuje ITU-T E-model kao objektivni metod ocene kvaliteta VoIP servisa. Takođe, u ovom radu se uzimaju i preporučene vrednosti mnogobrojnih ulaznih parametara koji ne uključuje QoS parametra i koje su specificirane u [62]. S obzirom da je cilj ove analize i generalno, ovog rada, da se pokaže uticaj paketske disperzije na strukturu gubitaka paketa, razmatranje kašnjenje i uticaja džitera u kontekstu strategija paketske disperzije je u drugom planu. U okviru analize kvaliteta, uzima se da je jednosmerno kašnjenje na disperzionim putanjama između 90 i 100 ms, sa džiter vrednostima koje ne prelaze 5ms. U skladu sa prethodnim iskazima o uticaju QoS parametara na VoIP kvalitet, ove vrednosti ne utiču na značajnu degradaciju percepiranog kvaliteta.

Scenario u okviru koga se vrši analiza VoIP kvaliteta i posledični uticaj strategija paketske disperzije podrazumeva postojanje dve disperzione putanje sa strukturom gubitaka paketa koje se međusobno značajno razlikuju. Na osnovu serije snimljenih VoIP paketa unutar testnog okruženja moguće je odrediti strukturu gubitka paketa koji se dobija direktnom primenom strategija paketske disperzije. Estimacija parametara Markovljevog modela sa 4 stanja iz snimljene serije VoIP paketa se dobija primenom

mehanizma opisanog u [44]. U skladu sa ovim pristupom, izgubljeni paket se smatra da je deo rafala ukoliko je rastojanje od prethodno izgubljenog paketa g_{min} pri čemu je rastojanje izraženo u paketima. Ovaj parametar se koristi za definiciju rafala na osnovu koje se izvode ostale relevantne tranzicione verovatnoće u skladu sa Markovljevim modelom sa 4 stanja. U cilju jasne raspodele između dobrog kvaliteta, kada nema rafalnih gubitaka, odnosno lošeg kvaliteta, kada je primetna pojava rafalnih gubitaka paketa, potrebno je usvojiti vrednost g_{min} koji će specificirati da li posmatrani gubitak paketa pripada rafalnim ili izolovanim gubicima. Za VoIP servis se usvaja vrednost $g_{min}=16$ paketa, čime se obezbeđuje jasna pojava rafala. Usvajanje visokih vrednosti ovog parametara odgovara servisima koji su osetljivi na gubitke paketa kao što je to slučaj sa video servisima kod kojih se usvaja vrednost g_{min} parametara koji je jednak 128 paketa. U slučaju ove analize, imajući u vidu rad sa VoIP servisom, usvaja se $g_{min}=16$.

Za komparaciju kvaliteta VoIP servisa se posmatra uticaj strategija paketske disperzije kada su raspoložive dve putanje za proces disperzije. Struktura gubitaka paketa na ove dve putanje (*Putanja 1* i *Putanja 2*) je izražena preko verovatnoće gubitaka paketa P_{loss} , broja rafalnih pojava (koji direktno zavisi od usvojene vrednosti g_{min} parametra), srednja dužina rafala B i verovatnoća gubitka tokom lošeg stanja Markovljevog modela P_{burst} . Nakon što se odrede ovi parametri, dalje se određuje $BurstR$ parametar koristeći matematičke relacije razvijene i predložene u ovom radu. Posledično, moguće je posredstvom E-modela realizovati kvalitativnu komparaciju VoIP servisa posmatrajući respektivno strategije paketske disperzije. Ova komparacija će biti kvantifikovana u MOS oceni kvaliteta pri čemu se konverzija iz R vrednosti parametra E-modela vrši u skladu sa izrazom (5.2). Ista metodologija se koristi za snimanje VoIP paketa kada se primenjuju pojedinačne strategije paketske disperzije preko navedene dve putanje.

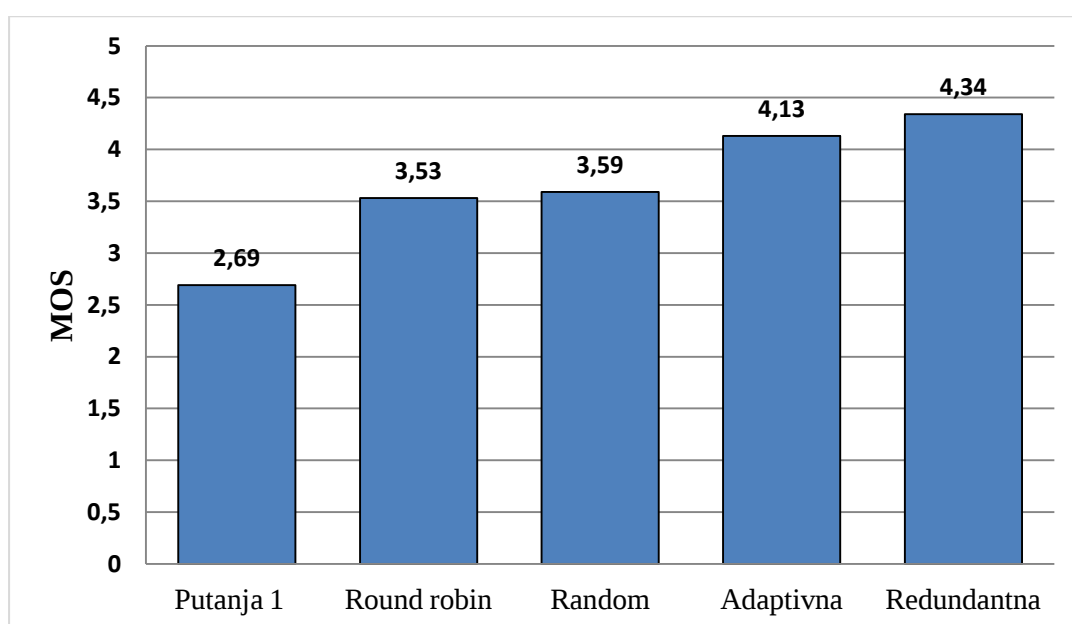
Tabela 5.1 Specifikacija rafalnih gubitaka dobijenih primenom paketske disperzije

Disperziona strategija	P_{loss} (%)	Broj rafalnih događaja	B (packets)	P_{burst} (%)
Putanja 1 (Konvencionalno rutiranje)	13.60	162	41.85	10.25
Putanja 2	1.38	13	19.03	2.56
Round robin	7.43	154	27.68	7.02
Random	6.71	140	26.82	6.52
Adaptivna	2.61	46	22.89	1.85
Redundantna	0.15	11	10.95	0.08

Kvalitativna komparacija VoIP servisa izražena u MOS vrednostima je prikazana na slici 5.5 gde se posmatraju poboljšanja koja se ostvaruju primenom pojedinačnih strategija paketske disperzije u odnosu na konvencionalno rutiranje kada se koristi jedna putanja, u ovom slučaju *Putanja 1*. Na osnovu specificiranih rafalnih gubitaka koji su izraženi na disperzionim putanjama (*Putanja 1* i *Putanja 2*) primećuje se da se ostvaruje sličan rezultat u pogledu MOS ocene kada se primenjuje *round robin*,

odnosno *random* strategija. Rezultati pokazuju da je u blagoj kvalitativnoj prednosti *random* strategija, što je takođe važno i za analizu rastojanja gubitaka paketa. Sa druge strane, adaptivna i redundantna strategija se u pogledu kvaliteta značajno razlikuju u odnosu na prethodne dve strategije, a takođe je i izražena međusobna razlika.

Dobre kvalitativne osobine adaptivne strategije su ovde posledica veće razlike u strukturi gubitaka paketa između disperzionih putanja. Naime, u specifikaciji strukture gubitaka paketa, *Putanja 2* kao dodatna putanja ima značajno bolje karakteristike gubitaka paketa u odnosu na *Putanju 1*, što posledično dovodi do činjenice da se primenom adaptivne strategije mogu ostvariti mnogo bolji kvalitativni rezultati jer adaptivna strategija izražava dobre osobine kod eliminacije bliskih i rafalnih gubitaka paketa kada je razlika strukture gubitaka paketa visoka.



Slika 5.5 Kvalitativna komparacija strategija paketske disperzije.

Redundantna strategija je superiornija u kvalitativnom smislu u odnosu na adaptivnu pri čemu se njenom primenom eliminišu skoro svi rafalni gubici paketa i *de facto* se povećava rastojanje između gubitaka paketa, što je i pokazano prethodnom analizom rastojanja gubitaka. Eliminacijom gubitaka paketa, a posledično i rafalnih gubitaka, omogućava se efikasniji rad PLC algoritama, što dovodi do povećanja kvaliteta VoIP servisa u slučaju primene redundantne strategije. Komparacija između adaptivne i redundantne strategije dovodi do MOS razlike od 0.21 što je posledica zavisnosti procenta izgubljenih paketa i MOS ocene jer u opsegu nižih vrednosti gubitaka paketa (manje od 5%) MOS ima blagu tendenciju rasta prema oceni kojom se manifestuje odličan kvalitet [142]. Veoma bitan zaključak ove analize je da sve strategije paketske disperzije ostvaruju značajno bolji VoIP kvalitet u odnosu na konvencionalno rutiranje u prisustvu rafalnih gubitaka, što proističe direktno iz velike razlike u strukturi

gubitaka paketa između disperzionih putanja, a naročito zbog rafalnih gubitaka paketa izraženih na *Putanji 1*.

Za scenarije koje uključuju slabije rafalne gubitke, optimalan izbor bi bio da se koristi adaptivna strategija ili eventualno *round robin* i *random* strategije kojim se ostvaruje značajno bolji kvalitet VoIP servisa u odnosu na odsustvo paketske disperzije. U slučajevima visokih rafalnih gubitaka na nekoliko putanja istovremeno se zahteva primena redundantne strategije koja zahteva više kapaciteta za prenos repliciranih VoIP paketa. Scenario koji uključuje razliku u strukturi gubitaka paketa na nekoliko putanja je adekvatan za primenu adaptivne strategije, pri čemu se moraju razmatrati i problemi koji proističu iz implementacione kompleksnosti. Na osnovu ciljanog nivoa kvaliteta i broja raspoloživih putanja za disperziju, selekcija strategije paketske disperzije bi trebalo da se bazira na balansu između kompleksnosti, zahtevanog kapaciteta i robusnosti individualne strategije na odgovarajuće scenarije rafalnih gubitaka (na primer, slična struktura gubitaka na disperzionim putanjama).

5.2. Analiza algoritama mapiranja klasa servisa između domena

Algoritmi mapiranja klasa servisa se posmatraju sa aspekta DC parametra koji je u prethodnom poglavlju definisan i kao takav predstavlja adekvatan parametar posredstvom koga se ocenjuje nivo tačnosti mapiranja kada se posmatraju ponuđene i zahtevane performanse. Analiziraju se postojeći algoritam P-CMS kao i predloženi algoritmi – 3PIP i GPMS.

Tabela 5.2 Specifikacija klasa servisa u okviru pojedinačnih domena

Domen	Klasa	Kašnjenje (ms)	Džiter (ms)	PLR	Cena (MU)
A1 (pristupni)	1	40	10	10^{-5}	40
	2	80	30	10^{-4}	20
	3	120	–	10^{-4}	15
A2 (pristupni)	1	20	15	10^{-6}	55
	2	50	20	10^{-5}	30
	3	70	30	5×10^{-5}	20
	4	120	–	10^{-4}	17
R1 (regionalni)	1	15	10	10^{-6}	75
	2	50	30	10^{-5}	60
R2 (regionalni)	1	12	6	10^{-5}	80
	2	–	–	10^{-4}	70
C (kontinentalni)	1	45	5	10^{-5}	90
	2	100	15	5×10^{-4}	65
	3	120	40	10^{-4}	55

Posmatra se E2E putanja koja je tipična za međunarodni saobraćaj i koja se sastoji od 5 hijerarhijski organizovanih domena. Dva od ta domena su pristupna, dok su ostali domeni označeni kao tranzitni i to su dva regionalna i jedan kontinentalni domen. Kako bi se napravilo adekvatno poređenje sa prethodnim 3P

baziranim algoritmima mapiranja klasa servisa, posmatraju se specifikacije klasa servisa po domenu u skladu sa Tabelom 5.2 koja je specificirana u [134]. Posledično, E2E putanja je ustvari A1-R1-C-R2-A2. Dodatno, cena interkonekcije se određuje za svaku klasu servisa od strane pojedinačnog domena i ova cena je izražena u monetarnim jedinicama (MU). Cena interkonekcije predstavlja cenu koja se određuje po jedinici kapaciteta koji je alocirani u okviru posmatranog domena.

Za realizaciju komparacije algoritama mapiranja klasa servisa između domena potrebno je dodatno specificirati E2E zahteve imajući u vidu tipične QoS zahteve i preporučene ciljne vrednosti specificirane u [3]. E2E zahtevi su specificirani u Tabeli 5.3.

Tabela 5.3 Specifikacija E2E zahteva

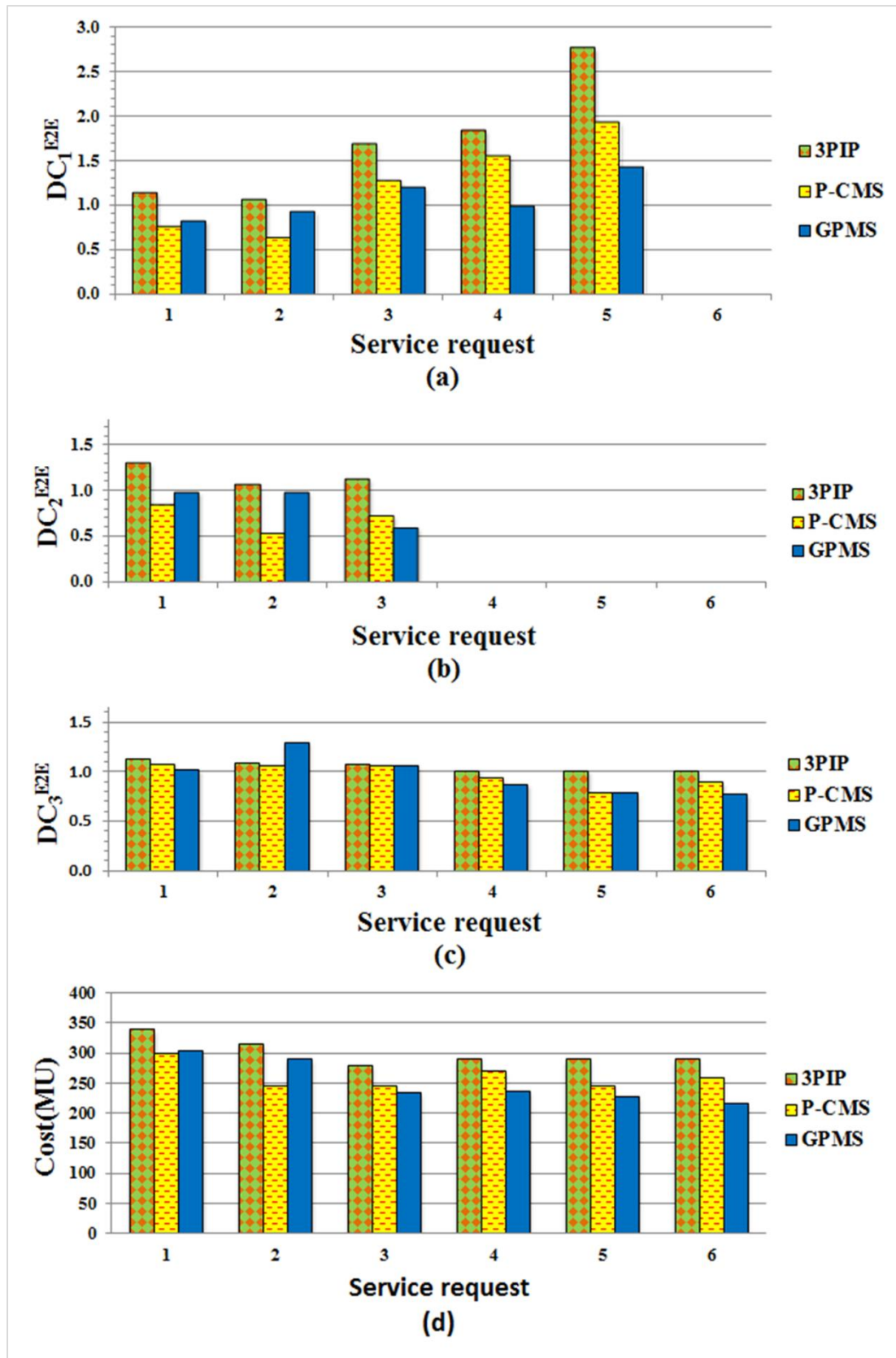
QoS parametar	Servisni zahtev					
	1	2	3	4	5	6
Kašnjenje (ms)	150	200	400	400	600	–
Džiter (ms)	60	60	80	–	–	–
PLR	10^{-4}	10^{-3}	10^{-3}	10^{-4}	10^{-4}	10^{-4}

Tabela 5.4 sadrži kvalitativnu komparaciju pomenuta tri algoritma mapiranja klasa servisa. P-CMS ne uzima u obzir cenu interkonekcije, dok 3PIP i GPMS podrazumevaju cenu interkonekcije kao ulazni parametar procesa mapiranja klasa. Margina pogoršanja performansi se odnosi na pogoršanje E2E performansi kome doprinosi jedan ili više domena koji se nalaze na E2E putanji. U skladu sa [136] razlikuju se statički, pseudo-statički, signalizacioni i akumulacioni pristupi alokaciji margine pogoršanja.

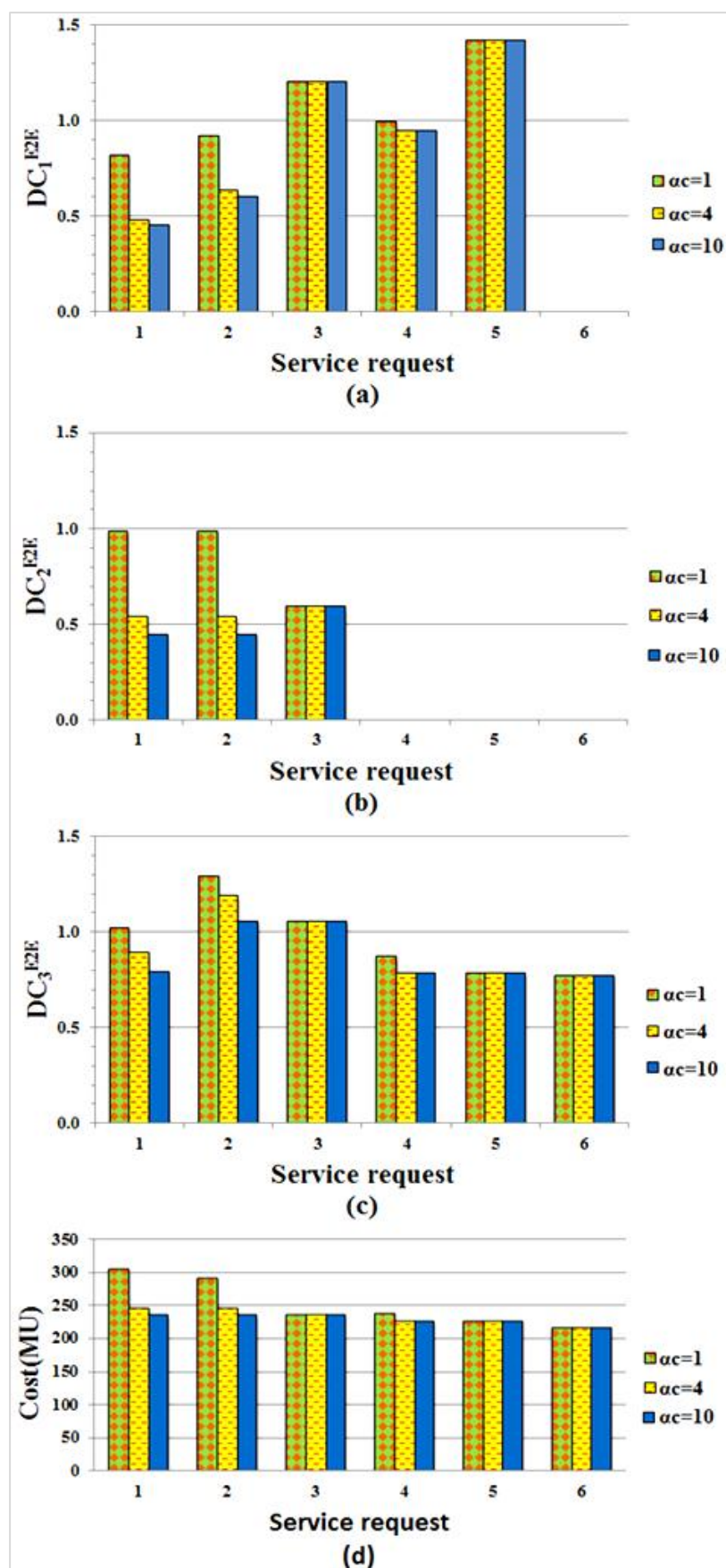
Tabela 5.4 Kvalitativna komparacija analiziranih algoritama mapiranja klasa servisa

Algoritam mapiranja	Cena interkonekcije	Alokacija margine pogoršanja	Servisni zahtevi na nivou domena	Fleksibilnost
P-CMS	Ne	Signalizacioni	Da	Srednja
3PIP	Da	Akumulacioni	Ne	Niska
GPMS	Da	Akumulacioni	Ne	Visoka

P-CMS koristi signalizacioni pristup alokacije margine pogoršanja kako bi se odredila margina koja se pridružuje pojedinačnom domenu, dok 3PIP i GPMS algoritmi se oslanjaju na akumulacioni pristup alokacije i podrazumevaju da se margina pogoršanja određuje na E2E nivou nezavisno od broja domena na servisnoj putanji. U skladu sa prethodnom činjenicom, P-CMS računa servisne zahteve na nivou pojedinačnog domena u skladu sa politikom kojom se specificira način alokacije margine pogoršanja. Na taj način se postiže odgovarajući fer odnos prema domenima koji se nalaze na E2E putanji na račun povećanja implementacione i računске kompleksnosti.



Slika 5.6 E2E QoS performanse i cena interkonekcije za različite servisne zahteve i algoritme mapiranja: (a) Kašnjenje; (b) Džiter; (c) PLR; (d) E2E cena interkonekcije.



Slika 5.7 Efekat modifikacije GPMS α_c parametra na: (a) kašnjenje; (b) džiter; (c) PLR; (d) E2E cenu interkonekcije.

Sa druge strane, algoritmi 3PIP i GPMS uzimaju u obzir cenu interkonekcije i oslanjaju se isključivo na ponude pojedinačnih domena u okviru E2E putanje. GPMS nudi visok nivo fleksibilnosti kroz mogućnost prioritizacije ciljeva mapiranja kao i modifikacije deviacionih promenljivih. Za razliku od GPMS, 3PIP algoritam podrazumeva samo specifikaciju E2E QoS nivoa, dok P-CMS može opcionalno da dodeli težinske faktore odgovarajućim karakteristikama pridruženih domenima.

Komparativna analiza se dalje sprovodi tako što se određuju E2E DC parametri za pojedinačne mere performansi (DC_j^{E2E} , $j=1, 2, 3$) i E2E cenu interkonekcije. Kao deo ove komparativne analize, jednaki težinski faktori kod GPMS algoritma su dodeljeni QoS parametrima i ceni interkonekcije, odnosno važi jednakost $\alpha_c = \alpha_j = 1$. Dobijeni rezultati su prikazani na slici 5.6 za 6 servisnih zahteva koji su specificirani u Tabeli 5.3.

Primetno je da GPMS algoritam realizuje najtačnije mapiranje u smislu poklapanja sa E2E servisnim zahtevima s obzirom da su vrednosti DC_j^{E2E} bliske vrednosti 1 u većini slučajeva. Mapiranje realizovano na ovaj način minimizuje verovatnoću pojavljivanja predimenzionisanja ili nedovoljnog dimenzionisanja performansi u zavisnosti od E2E servisnih zahteva. U slučajevima kada P-CMS i 3PIP realizuju bolje poklapanje performansi sa zahtevanim, GPMS dodatno unosi balans između cene i performansi. Za takve zahteve, GPMS u skladu sa modelom na koji se oslanja, može da utvrdi da je za ponuđenu cenu interkonekcije bolje uzeti nešto slabije performanse, a ostvariti bolju cenu. Ovaj balans između cene interkonekcije i performansi se kod GPMS algoritma može dodatno kontrolisati modifikacijom parametara α_c i α_j parametara.

U cilju analize nivoa fleksibilnosti GPMS algoritma dalje se utvrđuje kako na mapiranje utiče povećanje težinskog faktora α_c kojim se vrši minimizacija cene interkonekcije kada se vrednost težinskog parametra $\alpha_j=1$ održava konstantnim. Slika 5.7 prikazuje dobijene rezultate kada se posmatra DC_j^{E2E} parametar i E2E cena interkonekcije za 6 servisnih zahteva.

Činjenica da minimizacija cene interkonekcije ima veći prioritet utiče na GPMS mapiranje klasa servisa i u tom slučaju se dobijaju lošije E2E performanse. Primetno je da za servisni zahtev 4, 5 i 6, mapiranje klasa servisa ostaje isto kada je $\alpha_c=4$ i $\alpha_c=10$. Ovo je direktna posledica odluke modela na koji se oslanja GPMS algoritam u cilju očuvanja nivoa ponuđenih E2E performansi s obzirom da ne postoji mapiranje koje bi dovelo do boljeg balansa između cene i performansi. Slično važi i za α_j parametar koji može biti modifikovan kako bi se sprovelo mapiranje klasa servisa koje što bliže odgovara E2E servisnim zahtevima. Modifikacijom ovih parametara se može dosta granularno upravljati procesom mapiranja klasa servisa, a samim tim primena GPMS algoritma mapiranja obezbeđuje superiornije rezultate u poređenju sa P-CMS i 3PIP algoritmima mapiranja.

Komparacija E2E performansi koji se dobijaju 3PIP mapiranjem klasa servisa dovodi do bitnog zaključka, a to je da 3PIP mapiranja uvek ispunjavaju postavljene E2E servisne zahteve. Primetno je sa slike 5.6 da je DC parametar za svaki servisni zahtev veći od 1 što podrazumeva da je 3PIP mapiranje dovelo do boljih ponuđenih E2E performansi u odnosu na zahtevane. Sa druge strane, P-CMS vrši mapiranje klasa tako što nalazi najbliže poklapanje sa DC parametrom, odnosno da P-CMS mapiranje teži da DC parametar bude jednak 1. Posledično, P-CMS mapiranja zbog svog metoda rada dovode do toga da u slučaju striktnih zahtevanih QoS parametara, ponuđene performanse mogu da ne ispunjavaju postavljene zahteve što dovodi do neuspostavljanja servisa. 3PIP mapiranje uzima u obzir prethodno postavljene E2E zahteve i ukoliko postoji mapiranje koje može da ispuni E2E zahteve, 3PIP vrši selekciju tih klasa servisa u okviru domena na putanji. Kao posledica principa koji koristi 3PIP algoritam, selekcija klasa često dovodi do predimenzionisanja performansi kao što je to slučaj kod kašnjenja za servisni zahtev 5 u okviru koga je ponuđenog kašnjenje čak tri puta manje od zahtevanog, dok je za PLR parametar postignuto skoro idealno mapiranje.

Cene interkonekcije pridružene po klasi servisa specificirane u Tabeli 5.2 i definisane su u skladu sa hijerarhijskom organizacijom domena imajući u vidu smernice predstavljene u [128]. Ispunjenje svih zahtevanih QoS parametara direktno utiče na cenu interkonekcije koja je sada viša kao što je to prikazano na slici 5.6(d) za 3PIP mapiranje u poređenju sa P-CMS i GPMS algoritmima. Iako je u osnovi 3PIP algoritma ciljna funkcija koja teži da cena interkonekcije bude niža, niža cena P-CMS selekcije klasa je direktno uzrokovana selekcijom klasa slabijih performansi. Međutim, bez obzira na višu cenu, 3PIP je sposoban da izvrši selekciju klasa za servise koje ispoljavaju stroge QoS zahteve pri čemu se kod 3PIP stalno traži balans, odnosno traži se „najgore“ mapiranje klasa koje će ispuniti E2E zahtev za minimalnu cenu interkonekcije.

Slaba korelacija cene interkonekcije i ponuđenih performansi direktno proističe iz činjenice da P-CMS ne uzima cenu klase servisa kao ulazni parametar. Dodatno, veliki nedostatak P-CMS mapiranja predstavlja i težnja selekcije klasa da E2E performanse budu što bliže vrednosti 1 DC parametra pri čemu se posmatra srednja vrednost DC parametra, a ne pojedinačni QoS parametri koji su specificirani u okviru E2E servisnog zahteva. Međutim, P-CMS bi se mogao uz dodatne modifikacije uskladiti sa potrebama šireg dijapazona servisa različitih QoS zahteva, pa samim tim bi se mogla uspostaviti i bolja kontrola procesa mapiranja klasa servisa kada se primenjuje P-CMS.

6

ZAKLJUČAK

Rapidna proliferacija QoS zahtevnih aplikacija je uslovila neophodan razvoj i širenje i mreža. Ova činjenica neminovno dovodi do povećanja kompleksnosti usled implementiranih mrežnih mehanizama u cilju postizanja determinističkih performansi servisa. Degradacija performansi se manifestuje kroz pogoršanje vrednosti QoS parametara pri čemu uzroci degradacije mogu biti višestruki. Čest slučaj je pojava zagušenja gde se primenom postojećih QoS mehanizama i DiffServ arhitekture mogu postići determinističke performanse QoS kritičnih servisa. Međutim, ovi QoS mehanizmi podrazumevaju temeljno planiranje QoS politike koja usled kompleksnosti i potencijalnih grešaka se ne primenjuje često na dosledan način, kako unutar jedne mreže (domena), tako i izvan domena, gde se pojavljuje dodatni problem neusklađenosti QoS arhitekture između domena.

Imajući u vidu konkretne nedostatke u kontekstu QoS mehanizama u intradomenskim i interdomenskim scenarijima, u ovom radu je detaljno predstavljen uticaj paketske disperzije na poboljšanje QoS parametra, tačnije na uticaj strukture rafalnih gubitaka koji su označeni kao izuzetno štetni na percipirani kvalitet servisa. Detaljnom analizom nekoliko parametara gubitaka paketa kojima se može detaljno opisati struktura rafalnih gubitaka paketa, ustanovljen je nivo uticaja strategija paketske disperzije na promenu rafalnih karakteristika što rezultuje boljim percipiranim kvalitetom. Analizirane su postojeće strategije (*random* i *round robin*), a takođe su u ovom radu i predložene nove strategije (redundantna i adaptivna).

Razvijanje analitičkog modela koji uzima u obzir rafalne gubitke koji su modelovani Markovljevom modelom sa 4 stanja predstavlja značajan doprinos imajući u vidu da ovaj tip modelovanja predstavlja najtačniju varijantu modelovanja rafalnih gubitaka paketa kojim se dodatno postiže dobar odnos računске kompleksnosti i tačnosti. Analitički model je razvijen za sve četiri analizirane strategije, dok su rezultati verifikovani praktičnim testiranjem u realnom okruženju posmatrajući VoIP servis kao reprezentativni

servis koji ima stroge QoS zahteve. Rezultati su pokazali da primena strategija paketske disperzije značajno doprinosi poboljšanju QoS naročito u prisustvu rafalnih gubitaka. Potrebno je dodatno istaći i činjenicu da je praktično testiranje pokazalo i slučajeve u kojima je ograničena efikasnost delovanja paketske disperzije na rafalne gubitke paketa, što generalno i zavisi od samih karakteristika rafalnih gubitaka koji se javljaju na disperzionim putanjama. Na osnovu rezultata je zaključeno da svaka strategija ostvaruje bolje performanse u odnosu na konvencionalno rutiranje. Sa druge strane, primena *random* ostvaruje nešto bolje karakteristike u odnosu na *round robin*, dok efikasnost adaptivne strategije zavisi od toga koliko se razlikuje struktura izraženih gubitaka na putanjama. Primena redundantne strategije se predlaže u ekstremnim situacijama kada su visoki rafalni gubici istovremeno izraženi na nekoliko putanja, pa primena prethodno navedenih strategija ne bi imala osetnog uticaja na poboljšanje QoS.

Paketska disperzija predstavlja perspektivnu tehniku kojom je moguće sa minimalnom izmenom politike rutiranja ostvariti bolji QoS u uslovima rafalnih gubitaka bez upotrebe dodatne QoS arhitekture. Predloženi analitički model i testiranja paketske disperzije ukazuju na važnost daljeg razvoja QoS adaptivnih *multipath* protokola rutiranja koji bi obezbedili sveobuhvatno intradomensko QoS rešenje. Primena predstavljenog analitičkog modela se može ostvariti prilikom planiranja ciljnih vrednosti QoS parametara i samim tim odabirom adekvatne strategije u skladu sa očekivanim nivoom QoS i ispoljenim rafalnim gubicima paketa na putanjama, a sve u cilju maksimizacije nivoa percepiranog kvaliteta i postizanja determinističkih performansi.

Primena paketske disperzije se predlaže na nivou jednog domena s obzirom da algoritmi rutiranja koji daju fizički razdvojene putanje imaju visoku računsku kompleksnost, pa samim tim proširenje dinamičkog protokola rutiranja koji bi nalazio razdvojene putanje u interdomenskom scenariju bi bio previše složen i posledično bi imao dugačko trajanje procesa konvergencije. U ovom radu se pitanje realizacije interdomenskog QoS razmatra posmatranjem 3P modela ugovaranja pri čemu se kao centralna figura procesa ugovaranja razmatra mapiranje klasa servisa između domena.

S obzirom da 3P model ugovaranja servisa donosi dosta prednosti u odnosu na postojeći bilateralni model ugovaranja servisa između domena, implementacija algoritama mapiranja klasa servisa je sada moguća preko 3P domena koji direktno razmenjuje signalizaciju sa domenima koji učestvuju u realizaciji E2E QoS. Prateći centralizovano ugovaranje posredstvom 3P agenta, u ovom radu se, pored postojećeg P-CMS algoritma, predlažu 3PIP i GPMS algoritmi mapiranja koji su usmereni u rešavanje nedostataka P-CMS selekcije klasa servisa.

Uvođenjem predloženih algoritama se postiže ostvarivanje veće kontrole procesa mapiranja, odnosno definisanjem ulaznih parametara posredstvom 3P agenta je moguće specificirati što tačnije mapiranje klasa servisa između domena kako bi se postigao interdomenski QoS. Komparativna analiza algoritama mapiranja klasa servisa između domena pokazala je da je fleksibilnost u cilju povećanja kontrole procesa

mapiranja jedna od najvažnijih karakteristika kako bi se mogla izvršiti selekcija klasa koja ispunjava E2E servisne zahteve. Takođe, predložene metode mapiranja, 3PIP i GPMS, se oslanjaju na ciljne funkcije u okviru kojih se mogu prioritetizovati ekonomski ili tehnički parametri koji su relevantni za proces mapiranja. Ovakav pristup omogućava selekciju klasa servisa koja može da realizuje E2E QoS i za najstrožije zahteve, a pri tome se dodatno mogu implementirati poslovne politike kojima je cilj povećanje profita pojedinačnih domena.

Razmatranjem predmeta proučavanja i ostvarenih rezultata u ovoj disertaciji je moguće sagledati potencijalne pravce budućeg istraživanja. Najpre matematički model pridružen strategijama paketske disperzije se može integrisati u okviru budućeg algoritma rutiranja koji nalazi fizički razdvojene putanje adekvatne za implementaciju procesa paketske disperzije. Algoritam rutiranja može da primeni matematički model izračunavanja rastojanja i VoIP kvaliteta u zavisnosti od predloženih putanja i na taj način da dinamički odabere strategiju koja daje optimalne performanse.

Sa druge strane, ugovaranja servisa posredstvom 3P modela zahtevaju dodatnu specifikaciju i standardizaciju oko toga koje parametre treba razmenjivati između individualnih domena i 3P agenta. Buduće ostvarenje doprinosa na ovom polju podrazumeva kompletnu automatizaciju procesa ugovaranja servisa u okviru koga je takođe i proces mapiranja klasa servisa koji se realizuje posredstvom algoritma mapiranja. Kompletna automatizacija ugovaranja E2E QoS servisa teži ka integraciji ekonomskih ciljeva što se potencijalno može postići primenom mapiranja koji se oslanja na matematičke modele koji su sagledani u ovoj disertaciji. Implementacijom ovakvih algoritama je moguće u potpunosti iskontrolisati proces E2E mapiranja klasa servisa između domena, a odabirom ulaznih parametara se mogu postići specificirani ciljevi kroz automatizovani proces ugovaranja servisa.

REFERENCE

- [1] ITU-T Rec. G.1000, “Communications quality of service: A framework and definitions”, November 2001.
- [2] ITU-T Rec. Y.1540, “Internet protocol data communication service – IP packet transfer and availability performance parameters”, March 2011.
- [3] ITU-T Rec. Y.1541, “Network performance objectives for IP-based services”, December 2011.
- [4] F. Baker, “A Differentiated Services Code Point (DSCP) for Capacity-Admitted Traffic”, IETF RFC5868, May 2010.
- [5] K. Nichols, S. Blake, F. Baker, D. Black, “Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers”, IETF RFC 2474, December 1998.
- [6] K. Ramakrishnan, S. Floyd, D. Black, “The Addition of Explicit Congestion Notification (ECN) to IP”, IETF RFC 3168, September 2001.
- [7] D. Grossman, “New Terminology and Clarifications for Diffserv”, IETF RFC3260, April 2002.
- [8] V. Jacobson, K. Nichols, K. Poduri, “An Expedited Forwarding PHB”, IETF RFC 2598, June 1999.
- [9] J. Heinanen, F. Baker, W. Weiss, J. Wroclawski, “Assured Forwarding PHB Group”, IETF RFC 2597, June 1999.
- [10] A. Demers, S. Keshav, S. Shenker, “Analysis and simulation of a fair queueing algorithm”, In Symposium proceedings on Communications architectures & protocols (SIGCOMM '89) ACM, New York, NY, USA, pp.1-12.
- [11] A. C. Odinma, L. Oborkhale, “Quality of Service Mechanisms and Challenges for IP Networks”, The Pacific Journal of Science and Technology, Vol. 7(1), 2006.

-
- [12] J. Krämer, L. Wiewiorra, C. Weinhardt, “Net neutrality”, *Telecommunication Policy*, Vol. 37, No. 9, October 2013.
 - [13] T. Wu, C.Yoo, “Keeping the Internet Neutral?”, *Federal Communications Law Journal*, Vol. 59 , pp. 575–592, 2013.
 - [14] J. Scott Marcus, “Net Neutrality Revisited: Challenges and Responses in the EU and in the US”, Study for the IMCO Committee, December 2014.
 - [15] J. He, J. Rexford, "Toward internet-wide multipath routing", *IEEE Network*, Vol.22, No.2, pp.16-21, March-April 2008.
 - [16] J. Moy, “OSPF Version 2”, IETF RFC 2328, April 1998.
 - [17] M. Shand, L. Ginsberg, “Reclassification of RFC 1142 to Historic”, IETF RFC 7142, February 2014.
 - [18] T. Li, H. Smit, T. Przygienda , “Domain-Wide Prefix Distribution with Two-Level IS-IS”, IETF RFC 5302, October 2008.
 - [19] R. Teixeira, K. Marzullo, S. Savage, G. M. Voelker "Characterizing and Measuring Path Diversity of Internet Topologies," In *Proceedings of ACM SIGMETRICS*, June 2003.
 - [20] D. A. Schupke, A. Autenrieth, T. Fischer, ”Survivability of multiple fiber duct failures”, *Third International Workshop on the Design of Reliable Communication Networks (DRCN)*, Oct. 2001, Budapest.
 - [21] S. Kim and S. S. Lumetta, ”Evaluation of protection reconfiguration for multiple failures in WDM mesh networks,” In *Proceedings of OFC 2003*, Vol. 1, pp. 210–211, Mar. 23-28, 2003, Atlanta, GA.
 - [22] J. M. Simmons, “Catastrophic failures in a backbone network”, *IEEE Communications Letters*, Vol. 16(8), pp. 1328 – 1331, August 2012.
 - [23] D. Xu, Y. Xiong, C. Qiao, G. Li, ”Trap avoidance and protection schemes in networks with shared risk link groups,” *IEEE/OSA J. Lightw. Technol.*, Vol. 21, No. 11, pp. 2683–2693, Nov. 2003.

- [24] D. Xu, Y. Xiong, "A Novel Segment Protection Approach for SRLG Networks", IEEE Journal on Selected Areas in Communications: pp. 1320-1331, 2003.
- [25] A. A. M. Saleh, "Dynamic multi-terabit core optical networks: architecture, protocols, control and management (CORONET)," DARPA BAA 06-29, Proposer Information Pamphlet.
- [26] J. M. Simmons, Optical Network Design and Planning, New York, NY: Springer, 2008.
- [27] S. Prabhavat, H. Nishiyama, N. Ansari, N. Kato, "On Load Distribution over Multipath Networks," IEEE Communications Surveys & Tutorials, Vol.14, No.3, pp.662-680, 2012.
- [28] J. Cao, R. Xia, P. Yang, C. Guo, G. Lu, L. Yuan, Y. Zheng, H. Wu, Y. Xiong, D. Maltz. "Per-packet load-balanced, low-latency routing for clos-based data center networks", In Proceedings of the ninth ACM conference on Emerging networking experiments and technologies (CoNEXT '13). pp. 49-60, ACM, New York, USA.
- [29] M. Lengyel, J. Sztrik, "Performance comparison of traditional schedulers in DiffServ architectures Using NS," In Proceedings of the 16th European Simulation Symposium (ESS), Budapest, Hungary, October 2004.
- [30] H. Adishesu, G. Parulkar, and G. Varghese, "A reliable and scalable striping protocol," ACM SIGCOMM Computer Communication Review, Vol. 26, No. 4, pp. 131–141, Oct. 1996.
- [31] Z. Cao, Z. Wang, and E. Zegura, "Performance of hashing based schemes for Internet load balancing," In Proceedings IEEE INFOCOM, Tel Aviv, Israel, March 2000, pp. 332–341.
- [32] D. Thaler, C. Hopps, "Multipath issues in unicast and multicast nexthop selection," IETF RFC 2991, Nov. 2000.
- [33] R. Birke, M. Mellia, M. Petracca, D. Ross, "Experiences of VoIP traffic monitoring in a commercial ISP", International Journal of Network Management, Vol. 20, No.5, pp. 339-359, 2010.
- [34] J. F. Kurose, K. W. Ross, Computer Networking: A Top-Down Approach, Pearson, 6th Edition, 2012.

-
- [35] S. Low, L. Peterson, L. Wang, "Understanding TCP Vegas: Theory and Practice", Technical report 616-00, Princeton University, January 2000.
 - [36] T. Henderson, S. Floyd, A. Gurtov, Y. Nishida, "The New Reno Modification to TCP's Fast Recovery Algorithm", IETF RFC 6582, April 2012.
 - [37] V. Jacobson, R. Braden, "TCP Extensions for Long-Delay Paths", IETF RFC 1072, October 1998.
 - [38] M.S. Borella, D. Swider, S. Uludag, G.B. Brewster, "Internet packet loss: Measurement and implications for end-to-end qos", In Proceedings of ICPP'98, 1998.
 - [39] W. Jiang, H. Schulzrinne, "Modeling of packet loss and delay and their effect on real-time multimedia service quality", In Proceedings of NOSSDAV'00, 2000.
 - [40] A. Markopoulou, F. Tobagi, M. Karam, "Loss and Delay Measurements of Internet Backbones", Computer Communications, Vol. 29, pp. 1590-1604, 2006.
 - [41] L. Tang, J. Li, Y. Li, S. Shenker, "An investigation of the Internet's IP-layer connectivity", Computer Communications, Vol. 32, No. 5, pp. 913-926, 2009.
 - [42] E. N. Gilbert, "Capacity of a Burst-Noise Channel", Bell System Technical Journal, Vol. 39, No. 5, pp. 1253-1265, 1960.
 - [43] E. O. Elliott, "Estimates of Error Rates for Codes on Burst-Noise Channels", Bell System Technical Journal, Vol. 42, No. 5, pp. 1977-1997, 1963.
 - [44] T. Friedman, R. Caceres, A. Clark, "RTP Control Protocol Extended Reports (RTCP XR)", IETF RFC 3611, November 2003.
 - [45] L. Estrada-Vargas, D. Torres-Roman, H. Toral-Cruz, "Characterization and Modeling of Packet Loss of a VoIP Communication", Proceedings of the International Conference on Computer, Electrical, and Systems Science, and Engineering (ICCESSE), pp. 926-930, 2010.
 - [46] H. Toral-Cruz, A-S Khan Pathan, J. Ramírez-Pacheco, "Accurate Modeling of VoIP Traffic QoS Parameters in Current and Future Networks with Multifractal and Markov Models", Mathematical and Computer Modelling Journal, Vol. 57, No. 11-12, pp. 2832-2845, June 2013.

- [47] A. Clark, "Modeling the effects of burst packet loss and recency on subjective voice quality", Internet Telephony Workshop (IPtel 2001), Columbia University, 2001.
- [48] H. P. Singh, S. Singh, J. Singh, S. A. Khan, "Review: VoIP: State of art for global connectivity-A critical review", Journal of Network and Computer Applications, Vol. 37, pp. 365-379, 2014.
- [49] A. J. Jerri, "The Shannon sampling theorem—Its various extensions and applications: A tutorial review", Proceedings of the IEEE, Vol. 65, No. 11, pp. 1565 – 1596, November 1977.
- [50] ITU-T Rec. G.711, "Pulse code modulation (PCM) of voice frequencies", November 1988.
- [51] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", IETF RFC 3550, July 2003.
- [52] T. Koren, S. Casner, C. Bormann, "IP Header Compression over PPP", IETF RFC 3544, July 2003.
- [53] A. Sangwan, H. S. Jamadagni, M. C. Chiranth, R. Sah, V. Gaurav, "Second and Third Order Adaptable Threshold for VAD in VoIP", In Proceedings of the 6th International Conference on Signal Processing, Vol. 2, pp. 1693-1696, August 2002.
- [54] C. Kasap, M. Arslan, "A unified approach to speech enhancement and voice activity detection", Turkish Journal of Electrical Engineering and Computer Sciences, Vol. 21, pp. 527-547, February 2012.
- [55] B. Wu, K. Wang, "Noise Spectrum Estimation with Entropy-Based VAD in Non-stationary Environments", IEICE TRANS. FUNDAMENTALS, Vol. E89–A, No.2, February 2006.
- [56] R. Prasad, A. Sangwan, H. S. Jamadagni, M. C. Chiranth, R. Sah, V. Gaurav, "Comparison of Voice Activity Detection Algorithms for VoIP", In Proceedings of the Seventh International Symposium on Computers and Communications (ISCC'02), pp. 530-535, 2002.
- [57] T. Chua, D. C. Pheanis, "QoS evaluation of sender-based loss-recovery techniques for VoIP", IEEE Network, Vol.20, No.6, pp. 14-22, 2006.

- [58] A. Janicki, B. Ksiezak, “Packet Loss Concealment Algorithm for VoIP Transmission in Unreliable Networks”, In Proceedings of the 2008 conference on New Trends in Multimedia and Network Information Systems, pp. 23-33, 2008.
- [59] ITU-T Rec. P.920, “Interactive test methods for audiovisual communications”, May 2000.
- [60] ITU-T Rec. P.800, “Methods for subjective determination of transmission quality”, August 1998.
- [61] J. Beerends, Improvement of the P.861 Perceptual Speech Quality Measure, ITU-T SG12 COM-34E, December 1997.
- [62] ITU-T Rec. G.107, “The E-model: a computational model for use in transmission planning”, June 2015.
- [63] ITU-T Rec. P.862, “Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs”, February 2002.
- [64] R. G. Cole, J. H. Rosenbluth, “Voice over IP performance monitoring”, ACM SIGCOMM Computer Communication Review, Vol. 31, No. 2, 2001.
- [65] ITU-T Rec. G.113, “Transmission impairments due to speech processing”, November 2011.
- [66] T.A. Hall, “Objective speech quality measures for Internet telephony”, in Proceedings of SPIE Voice over IP VoIP Technology, Vol. 4522, pp. 128–136, Denver, CO, USA, 2001.
- [67] H. Assem, D. Malone, J. Dunne, P. O'Sullivan, “Monitoring VoIP call quality using improved simplified E-model”, In the Proceedings of International Conference on Computing, Networking and Communications (ICNC), pp. 927 – 931, January 2013.
- [68] R. Koodli, R. Ravikanth, “One-way loss pattern sample metrics”, IETF RFC 3357, August 2002.
- [69] W. Jiang , H. Schulzrinne, “Comparison and optimization of packet loss repair methods on VoIP perceived quality under bursty loss”, In Proc. International Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV), 2002.

- [70] A. Gladisch, R. Daher, D. Tavangarian, "Survey on Mobility and Multihoming in Future Internet", *Wireless Personal Communications*, Vol. 74, No. 1, pp. 45-81, 2014.
- [71] A. Habib, J. Chuang, "Improving application QoS with residential multihoming", *Computer Networks*, Vol. 51, No. 12, pp. 3323-3337, 2007.
- [72] D. Tipper, "Resilient network design: challenges and future directions", *Telecommunication Systems*, Vol. 56, No. 1, pp. 5-16, 2014.
- [73] D. Wischik, C. Raiciu, A. Greenhalgh, M. Handley, "Design, implementation and evaluation of congestion control for multipath TCP", In *Proceedings of the 8th USENIX conference on Networked systems design and implementation (NSDI'11)*, pp. 99-112, 2011.
- [74] B. Augustin, T. Friedman, R. Teixeira, "Measuring load-balanced paths in the internet", In *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement (IMC '07)*, pp. 149-160, 2007.
- [75] F. Ishizaki, "Analysis of Performance Improvement with Packet Dispersion," *TENCON 2005*, pp.1-6, 2005.
- [76] I. Gojmerac, P. Reichl, L. Jansen, "Towards low-complexity Internet traffic engineering: The Adaptive Multi-Path algorithm", *Computer Networks*, Vol. 52, No. 15, pp. 2894-2907, 2008.
- [77] M. Chiesa, G. Kindler, M. Schapira, "Traffic engineering with Equal-Cost-Multipath: An algorithmic perspective", *IEEE Conference on Computer Communications (INFOCOM 2014)*, pp. 1590-1598, 2014.
- [78] B. Choi, S. Moon, Z. Zhang, K. Papagiannaki, C. Diot, "Analysis of point-to-point packet delay in an operational network", *Computer Networks*, Vol. 51, No.13, pp. 3812-3827, 2007.
- [79] S. Tao, K. Xu, Y. Xu, T. Fei, L. Gao, R. Guerin, J. Kurose, D. Towsley, Z. Zhang, "Exploring the performance benefits of end-to-end path switching", *Proceedings of the 12th IEEE International Conference on Network protocols (ICNP 2004)*, pp.304-315, 2004.

-
- [80] C. Lin-huang, L. Tsung-Han, C. Hung-Chi, L. Yu-Lung, C. Yu-Jen, "QoS-aware path switching for VoIP traffic using SCTP", *Computer Standards & Interfaces*, Vol. 35, No. 1, pp. 158-169, 2013.
- [81] S. Tao, K. Xu, A. Estepa, T. F. L. Gao, R. Guerin, J. Kurose, D. Towsley, Z. Zhang, "Improving VoIP quality through path switching", *Proceedings of the IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2005)*, Vol.4, pp. 2268-2278, 2005.
- [82] Y. J. Liang, E. G. Steinbach, B. Girod, "Multi-stream voice over IP using packet path diversity", *IEEE Fourth Workshop on Multimedia Signal Processing*, pp.555-560, 2001.
- [83] J. Apostolopoulos, M. Trott, W. Tan, *Multimedia over IP and Wireless Networks*, Press, Burlington, Chapter 17, pp. 559-590, 2007, ISBN 9780120884803.
- [84] S. Qazi, T. Moors, "Finding Alternate Paths in the Internet: A Survey of Techniques for End to End Path Discovery", *International Journal of Current Engineering and Technology*, Vol. 2 No.4, pp. 328-339, 2012.
- [85] X. Yang, D. Wetherall, "Source selectable path diversity via routing deflections", *SIGCOMM Computer Communications Review*, Vol.36, No. 4, pp. 159-170, 2006.
- [86] Y. O. Lee, A. L. Narasimha Reddy, "Constructing disjoint paths for failure recovery and multipath routing", *Computer Networks*, Vol. 56, No. 2, pp. 719-730, 2012.
- [87] H. Zlatokrilov, H. Levy, "Session Privacy Enhancement by Traffic Dispersion", In *Proceedings 25th IEEE International Conference on Computer Communications INFOCOM 2006* , pp.1-12, 2006.
- [88] S. Avallone, V. Manetti, M. Mariano, S. P. Romano, "A splitting infrastructure for load balancing and security in an MPLS network", *3rd International Conference on Testbeds and Research Infrastructure for the Development of Networks and Communities, TridentCom 2007*, pp.1-6, 2007.

- [89] A. Bagchi , A. Chaudhary , M. T. Goodrich , S. Xu, “Constructing Disjoint Paths for Secure Communication”, In Proceedings of 17th Intl. Conference on Distributed Computing (DISC '03), Vol. 2848 of Lecture Notes in Computer Science, 2003.
- [90] H. Alwan, A. Agarwal, “A Multipath Routing Approach for Secure and Reliable Data Delivery in Wireless Sensor Networks,” International Journal of Distributed Sensor Networks, Vol. 2013, Article ID 232798, 10 pages, 2013.
- [91] S. Bettermann, Y. Rong, "Estimating the deliverable quality of a fully redundant dispersity routing system", In Proc. 17th Asia-Pacific Conference on Communications (APCC) 2011, pp.512-517, 2-5 Oct. 2011.
- [92] S. Bettermann, Y. Rong, "Effects of fully redundant dispersity routing on VoIP quality", In Proc. IEEE International Workshop Technical Committee on Communications Quality and Reliability (CQR) 2011, pp.1-6, 10-12 May 2011.
- [93] H. Levy, H. Zlatokrilov, “The effect of packet dispersion on Voice applications in IP Networks”, IEEE/ACM Transactions on Networking, Vol. 14, No. 2, pp. 277-288, 2006.
- [94] N. Ninkovic, S. Gajin, I. Reljin, “Packet dispersion strategy evaluation from the perspective of packet loss pattern and VoIP quality”, Computer Science and Information Systems, to be published.
- [95] P. Couto da Silva, M. Varela, E. de Souza e Silva, R. Leão, G. Rubino, “Quality assessment of interactive voice applications”, Computer Networks, Vol. 52, No. 6, pp. 1179-1192, 2008.
- [96] J. F. Banu, V. Ramachandran, “Multipath Adaptive Packet Dispersion for Voice Applications”, Journal of Computer Science, Vol. 8(4), pp. 454-459, 2012.
- [97] N. Ninkovic, Z. Bojovic, S. Gajin, “A Novel Scheme for Dynamic Triggering Of Packet Dispersion”, Elektronika ir Elektrotechnika, Vol. 20, No. 5, May 2014.
- [98] W. Shi, M. H. MacGrefor, P. Gburzynski, “Load Balancing for Parallel Forwarding”, IEEE/ACM Trans. on Networking, Vol. 13(4), pp. 790-801, August 2005.

-
- [99] G. Cincikas, S. Japertas, "Router Resource Management Research", *Elektronika ir Elektrotechnika*, Vol. 122, No. 6, 2012.
- [100] A. Vishwanath, V. Sivaraman, M. Thottan, "Perspectives on router buffer sizing", *SIGCOMM Comput. Commun. Rev.* Vol 39(2), pp. 34-39, April 2009.,
- [101] L. Sequeira, J. Fernandez-Navajas, J. Saldana, L. Casadesus, "Empirically Characterizing the Buffer Behaviour of Real Devices", *Performance Evaluation of Computer and Telecommunication Systems (SPECTS)*, pp. 1-6, July 2012.
- [102] Y. Zang, J. YanTuning, "Self-Similar Traffic to Improve Loss Performance in Small Buffer Routers", *The Tenth International Conference on Networks*, January 2011.
- [103] Y. S. Hanay, T. Wolf, "Practical Packet Pacing in Small-Buffer Networks", *IEEE International Conference on Communications ICC '09*, pp. 1-6, June 2009.
- [104] A. R. Ward , W. Whitt, "Predicting response times in processor-sharing queues", In *Proc. of the Fields Institute Conf. on Comm. Networks*, pp. 1-29, 2000.
- [105] J. Saldana, J. Fernandez-Navajas, J. Ruiz-Mas, E. Viruete Navarro, "The Utility of Characterizing Packet Loss as a Function of Packet Size in Commercial Routers", *Consumer Communications and Networking Conference (CCNC)*, pp. 346-347, January 2012.
- [106] G. Appenzeller , I. Keslassy , N. McKeown, "Sizing Router Buffers", In *Proc. of ACM SIGCOMM*, pp. 281—292, 2004.
- [107] K. Kompella, J. Lang, "Procedures for Modifying the Resource reSerVation Protocol (RSVP)", *IETF RFC 3936*, October 2004.
- [108] Y. Rekhter, T. Li, S. Hares, "A Border Gateway Protocol 4 (BGP-4)", *IETF RFC 4271*, January 2006.
- [109] S. Norden, "Inter-domain routing: Algorithms for QoS guarantees", *Computer Networks*, Vol. 49, No. 4, pp. 593-619 November 2005.

- [110] X. Masip-Bruin, M. Yannuzzi, J. Domingo-Pascual, A. Fonte, M. Curado, E. Monteiro, F. Kuipers, P. Van Mieghem, S. Avallone, G. Ventre, P. Aranda-Gutiérrez, M. Hollick, R. Steinmetz, L. Iannone, K. Salamatian, "Research challenges in QoS routing", *Computer Communications*, Vol. 29, No. 5, pp. 563-581, March 2006.
- [111] Z. Tielei, Y. Cui, Y. Youjian, F. Lizheng, T. Korkmaz, "Scalable BGP QoS Extension with Multiple Metrics," In *Proc. International conference on Networking and Services ICNS '06*, pp.80-80, July 2006.
- [112] Y. Cui , Z. YouJian, T. Korkmaz, Z. Tielei, "Supporting multiple metrics in QoS-aware BGP", *Science China Information Sciences*, Vol. 53, No. 10, pp 1947-1962, October 2010.
- [113] A. Fonte, M. Curado, E. Monteiro, "Interdomain quality of service routing: setting the grounds for the way ahead", *Annals of telecommunications*, Vol. 63, No. 11, pp. 683-695, December 2008.
- [114] A. Morton, S. Van den Berghe, "Framework for Metric Composition", *IETF RFC 5835*, April 2010.
- [115] P. Levis, M. Boucadair, "Considerations of provider-to-provider agreements for Internet-scale quality of service (QoS)", *IETF RFC 5160*, 2008.
- [116] R. Hancock, G. Karagiannis, J. Loughney, S. Van den Bosch, "Next Steps in Signaling (NSIS): Framework", *IETF RFC 4080*, June 2005.
- [117] P. Jacobs, B. Davie, "Technical challenges in the delivery of interprovider QoS". *IEEE Communications Magazine*, Vol. 43, No. 6, pp. 112–118, 2005.
- [118] D. Clark, J. Wroclawski, K. R. Sollins, R. Braden, "Tussle in Cyberspace: Defining Tomorrow's Internet", *IEEE/ ACM Trans. Networking*, Vol. 13, No. 3, pp. 462-475, June 2005.
- [119] J. Walker, A. Kulkarni, "Common Open Policy Service (COPS) Over Transport Layer Security (TLS)", *IETF RFC 4261*, December 2005.
- [120] G. Bruns, M. Cortes, "A hierarchical approach to service negotiation", In *Proceedings of the 2011 IEEE International Conference on Web Services*: pp. 460–4677, 2011.

-
- [121] C. A. Kamienski, D. Sadok, “The case of interdomain dynamic QoS-based service negotiation in the internet”, *Computer Communications* Vol. 27, pp. 622–37, 2004.
- [122] M. Dodd, A. Jung, B. Mitchell, P. Paterson, P. Reynolds, “Bill-and-keep and the economics of interconnection in next-generation networks”, *Telecommunications Policy*; Vol. 33, pp. 241–252, 2009
- [123] V. Siris, I. Ganchev, M. O’Droma, B. Stiller, Services, optimization, and economic aspects. In Koucheryavy Y et al. (Eds.), *Traffic and QoS Management in Wireless Multimedia Networks: COST 290 Final Report. Lecture Notes in Electrical Engineering* Vol. 31, pp. 267–304. Springer Science+Business Media, LLC; 2009.
- [124] J. L. Hougaard, *An Introduction to Allocation Rules*, Springer-Verlag, Berlin Heidelberg; 2009.
- [125] L. He, J. Walrand, “Pricing and revenue sharing strategies for internet service providers”, *IEEE Journal on Selected Areas in Communications*, Vol. 24, pp. 942–51, 2006.
- [126] H. Le Cadre, D. Barth, H. Pouyllau, “QoS commitment between vertically integrated autonomous systems”, *European Journal of Operational Research*, pp.214: 627–643, 2011.
- [127] H. Pouyllau, R. Douville, “End-to-end QoS negotiation in network federations”, In *Proceedings of the 2010 IEEE/IFIP Network Operations and Management Symposium Workshops*, pp.173–176, 2010.
- [128] M. Stojanovic, A. Kostic-Ljubisavljevic, V. Radonjic Djogatovic, “SLA-controlled inter-connection charging in next generation networks”, *Computer Networks*, Vol. 57, No. 11, pp. 2374–2394, 2013.
- [129] J. Gozdecki, A. Jajszczyk, R. Stankiewicz, “Quality of service terminology in IP networks”, *IEEE Communications Magazine*, Vol. 41, No. 3, pp. 153–159, 2003.
- [130] E. Boschi, M. Bossardt, and T. Dübendorfer, “Validating inter-domain SLAs with a programmable traffic control system”, in *Active and Programmable Networks, Lecture Notes in Computer Science*, vol. 4388, pp. 1–12, Springer, 2009.

- [131] A. Kajackas, V. Batkauskas, A. Saltis, and D. Gursnys, “Autonomous system for observation of QoS in telecommunications networks”, *Elektronika Ir Elektrotechnika*, vol. 111, no. 5, pp. 15–18, 2011.
- [132] M. Stojanovic, S. Bostjancic Rakas, and V. Acimovic-Raspopovic, “End-to-end quality of service specification and mapping: the third party approach”, *Computer Communications*, vol. 33, pp.1354–1368, 2010.
- [133] M. Stojanovic and S. Bostjancic Rakas, “Policies for allocating performance impairment budgets among multiple IP providers”, *AEÜ-International Journal of Electronics and Communications*, vol. 67, no. 3, pp. 206–216, 2013.
- [134] B. Mali, N. Ninkovic, M. Stojanovic, and G. Savic, “Service class mapping based on integer programming algorithm in the third party agent”, in *Proceedings of the 22nd Telecommunications Forum TELFOR 2014*, pp. 170–173, Belgrade, Serbia, November 25-27, 2014.
- [135] Nemanja Ninkovic, Bozidar J. Mali, Mirjana D. Stojanovic, Gordana I. Savic, “Multi-Objective Third-party Approach for Service Class Mapping among Multiple Providers in the Internet”, *Elektronika ir Elektrotechnika*, Vol. 21, No. 2, May 2015.
- [136] ITU-T Rec. Y.1542, “Framework for achieving end-to-end IP performance objectives”, July 2006.
- [137] B. Korte, J. Vygen, *Combinatorial Optimization: Theory and Algorithms*, 5th edition. Springer Berlin Heidelberg, 2012, ch. 5.
- [138] P. H. Williams, *Model Building in Mathematical Programming*, 5th edition. Chichester, UK: John Wiley & Sons, 2013, ch. 3.
- [139] E. Mingozzi, G. Stea, M. A. Callejo-Rodríguez., “EuQoS: Endto-end quality of service over heterogeneous networks”, *Computer Communications*, vol. 32, no. 12, July 2009, pp. 1355–1370.
- [140] W. L. Winston, *Operations Research: Applications and Algorithms*, 4th edition. Belmont, CA: Duxbury Press, 2003, ch. 4.

- [141] S. Hemminger, “Network Emulation with NetEm”, In Proceedings of the 6th Australia’s National Linux Conference (LCA2005). 2005.
- [142] A. Raake, Speech Quality of VoIP: Assessment and Prediction, John Wiley & Sons, 2006.

Biografija

Nemanja Ninković je rođen 11. juna 1987. godine u Beogradu. Osnovnu školu i srednju školu je završio u Beogradu, 2002. i 2006, respektivno. Nosilac je Vukove diplome i bronzane medalje sa republičkog takmičenja iz elektronike 2005. godine, a bio je i đak generacije Srednje tehničke PTT škole u Beogradu.

Osnovne akademske studije je završio na Elektrotehničkom fakultetu, Univerziteta u Beogradu 2010. godine, na odseku Telekomunikacije, smer Sistemsko inženjerstvo sa prosekom 8,87. Diplomirao je kod prof. dr Mirjane Stojanović sa ocenom 10, sa temom „Implementacija QoS u višedomenskim IP mrežama“. Master studije je takođe završio na Elektrotehničkom fakultetu, modul Sistemsko inženjerstvo i radio komunikacije sa prosekom 10 tokom kojih je diplomirao kod prof. dr Mirjane Stojanović sa ocenom 10, sa temom „Analiza tehničkih aspekata ugovaranja servisa između IP provajdera“.

U oktobru 2012. godine upisao je doktorske studije na Elektrotehničkom fakultetu, Univerziteta u Beogradu, studijski modul Računarska tehnika i informatika. Položio je sve ispite na doktorskim studijama sa prosečnom ocenom 10.

Od juna 2012. godine se nalazi u stalnom radnom odnosu u okviru Informacione-komunikacione ustanove „Akademska mreža Srbije – AMRES“ gde trenutno obavlja poslove na poziciji višeg stručnog saradnika u okviru sektora za infrastrukturu i resurse. Poslovi podrazumevaju planiranje, implementaciju i održavanje mrežne infrastrukture koja je osnova Internet provajdinga akademskim i naučno-istraživačkim institucijama u Srbiji, kao i konfiguraciju mrežne opreme vodećih proizvođača i implementaciju nadgledanja mreže.

Tokom doktorskih studija je učestvovao u GEANT3+ i trenutno radi na GEANT4 projektu na razvoju i implementaciji rešenja za nadgledanje i verifikaciju performansi za potrebe multidomenskih servisa. Svoje stručno usavršavanje je materijalizovao kroz industrijske sertifikate mrežne opreme, a trenutno je angažovan u RCUB Cisco akademiji kao instruktor.

Прилог 1.

Изјава о ауторству

Потписани-а Немања Нинковић

Број индекса или пријаве докторске дисертације 5051/12

Изјављујем

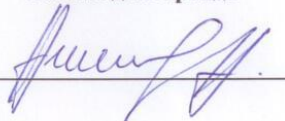
да је докторска дисертација под насловом:

Нова решења за побољшање квалитета сервиса у интрадоменском и
интердоменском мрежном окружењу

- резултат сопственог истраживачког рада,
- да предложена докторска дисертација у целини ни у деловима није била предложена за добијање било које дипломе према студијским програмима других високошколских установа,
- да су резултати коректно наведени и
- да нисам кришио/ла ауторска права и користио/ла интелектуалну својину других лица.

У Београду, 5.04.2016.

Потпис докторанда



Прилог 2.

**Изјава О истоветности штампане и електронске
верзије докторске дисертације**

Име и презиме аутора Немања Нинковић

Број индекса или пријаве докторске дисертације 5051/12

Студијски програм Рачунарска техника и информатика

Наслов докторске дисертације Нова решења за побољшање квалитета
сервиса у интрадоменском и интердоменском мрежном окружењу

Ментор др Славко Гајин, доцент

Потписани/а Немања Нинковић

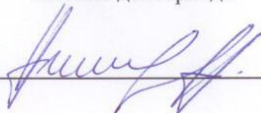
Изјављујем да је штампана верзија моје докторске дисертације истоветна електронској верзији коју сам предао/ла за објављивање на порталу **Дигиталног репозиторијума Универзитета у Београду**.

Дозвољавам да се објаве моји лични подаци везани за добијање академског звања доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

У Београду, 5.04.2016.

Потпис докторанда



Прилог 3.

Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

Нова решења за побољшање квалитета сервиса у интрадоменском

и интердоменском мрежном окружењу

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигитални репозиторијум Универзитета у Београду могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство
2. Ауторство - некомерцијално
- ☒ 3. Ауторство – некомерцијално – без прераде
4. Ауторство – некомерцијално – делити под истим условима
5. Ауторство – без прераде
6. Ауторство – делити под истим условима

(Молимо да заокружите само једну од шест понуђених лиценци, кратак опис лиценци дат је на крају).

У Београду, 5.04.2016.

Потпис докторанда

