

UNIVERZITET U BEOGRADU
MATEMATIČKI FAKULTET

Ozren Demonja

**REALIZACIJA P2P PROTOKOLA ZA
DOSTAVU SINHRONIZOVANOG
SADRŽAJA**

master rad

Beograd, 2018.

Mentor:

dr Aleksandar KARTELJ, docent
Univerzitet u Beogradu, Matematički fakultet

Članovi komisije:

dr Saša MALKOV, profesor
Univerzitet u Beogradu, Matematički fakultet

dr Miroslav MARIĆ, profesor
Univerzitet u Beogradu, Matematički fakultet

Datum odbrane: _____

Mami i tati

Naslov master rada: Realizacija P2P protokola za dostavu sinhronizovanog sadržaja

Rezime:

Ključne reči: P2P mreža, deljenje video sadržaja, sinhronizovano deljenje sadržaja, P2P protokol, deljenje u realnom vremenu, decentralizovana razmena podataka

Sadržaj

1	Uvod	1
2	P2P mreža	2
2.1	Osnovni koncepti	3
2.2	Klasifikacija P2P mreža	4
2.3	Topologija	5
2.4	Protok video sadržaja	9
2.5	Modelovanje mreže	11
3	Predloženi Kikkar P2P protokol	13
3.1	Arhitektura mreže	13
3.2	Pridruživanje parnjaka	15
3.3	Distribucija sadržaja	16
3.4	Kontrola zagušenja	18
3.5	Sinhronizacija parnjaka	19
4	Zaključak	21
	Bibliography	22

Glava 1

Uvod

Glava 2

P2P mreža

Osnovna svrha interneta kao globalne mreže u današnje vreme je razmena digitalnog sadržaja. Arhitektura koja se koristi za podršku razmeni često je zasnovana na modelu klijent-server (eng. *Client-Server architecture*). Model klijent-server deli zadatke između servera koji obezbeđuju resurse ili usluge i klijenata koji zahtevaju usluge [7].

Kako svaki pristup sistemu od strane klijenta zahteva veću upotrebu resursa i bolje performanse, jasno se uočava da je server usko grlo sistema. Tokom 1990-ih godina internet se najvećim delom zasnivao na klijent-server modelu. Poslednjih desetak godina dolazi do velikog proboja novih tehnologija usled želje korisnika za interaktivnijim sadržajem, porastom brzine pristupa internetu (eng. *broadband*), rasprostranjenosti pristupa, boljom stabilnosti veze i jačim računarima korisnika.

Navedene činjenice dovode do promene načina korišćenja interneta i pravljenja mreže ravnopravnih članova (eng. *peer-to-peer (P2P) network*) u kojoj korisnici dele svoje resurse [19]. Primena ovakvog pristupa obećavajuće je rešenje za mnoge oblasti.

U poglavlju 2.1 opisani su osnovni koncepti mreže ravnopravnih članova. U poglavlju 2.2 opisane su struktuirane i nestruktuirane mreže ravnopravnih članova. U poglavlju 2.3 opisane su topologija stabla (eng. *tree topology*), mrežna (eng. *mesh topology*) i hibridna (eng. *hybrid topology*) topologija. Način pripreme i slanja video sadržaja opisan je u poglavlju 2.4. U poglavlju 2.5 opisani su osnovni problemi prilikom pravljenja mreže ravnopravnih članova.

2.1 Osnovni koncepti

Korišćenje aplikativnog sloja u formiranju pokrivača (eng. *overlays*) za pristup internet servisima ima dugu istoriju. I pored duge tradicije ovakav pristup se do nedavno koristio samo za dizajn specifičnih protokola i interno povezivanje servisa infrastrukture, bez mogućnosti korišćenja između dva računara koja nisu susedi. Popularizacija ideje za formiranje pokrivača na aplikativnom sloju počinje 1999. godine sa Napster-om [1] i nastavlja se kroz druge sisteme za razmenu digitalnog sadržaja kao što su Gnutella [4], FastTrack [9], KaZaa [12] i BitTorrent [3].

Arhitektura ravnopravnih članova (eng. *peer-to-peer*, *P2P*) definiše se kao distribuirani vid arhitekture u kojoj svaki parnjak (eng. *peer*) može samostalno da obavlja poslove, ali omogućava drugim parnjacima da koriste njegove resurse kao što i sam može po potrebi koristiti resurse drugih parnjaka [16]. Drugim rečima, svaki parnjak komunicira sa ostalim parnjacima u mreži koristeći softver koji se ponaša i kao klijent i kao server. Na osnovu navedenih činjenica zaključuje se da arhitektura ravnopravnih članova omogućava veću autonomiju članova mreže koji sami definišu pravila o deljenju resursa. Glavna primena arhitekture ravnopravnih članova prisutna je u sistemima sa većom tolerancijom greške kod distribuirane obrade. Pored toga, ovaj vid arhitekture koristi se prilikom razmene fajlova, direktne komunikacije, obrade velikih količina podataka i softvera za zabavu.

Prilikom deljenja resursa svaki parnjak doprinosi resursima mreže ravnopravnih članova. Idealno, deljenje resursa je proporcionalno broju parnjaka koji koriste mrežu ravnopravnih članova. Međutim, usled različitih problema u praksi navedeno svojstvo većinom nije ispunjeno.

P2P sistem čini skup međusobno povezanih parnjaka, direktno ili preko drugih parnjaka formirajući povezani graf. Kako u grafu ne postoji centralna tačka kontrole i kolektivno se obavlja posao, P2P sistemi su decentralizovani. U nekim situacijama definiše se centralni server za poslove kao što su identifikacija, autentikacija i različite sigurnosne provere. U ovakvim uslovima, po pravilu server obavlja što je moguće manji deo posla. Pored toga, ponekad se u P2P sistemima iz različitih razloga uvode parnjaci zaduženi za dodatni deo posla. Ovakvi parnjaci se nazivaju super parnjaci (eng. *super peers*).

Učešće parnjaka u obavljanju posla u P2P sistemima određeno je lokalno. Iz ove činjenice proizilazi bolja organizacija P2P sistema vremenom usled korišćenja lokalnih znanja i operacija na svakom parnjaku. Ovim se takođe sprečava neplanirana

dominacija nekog parnjaka sistemom.

Dva važna svojstva P2P sistema su samoskalabilnost i stabilnost. Samoskalabilnost, kao glavna prednost P2P sistema, ispoljava se povećanjem ukupne snage sistema prilikom pristupanja parnjaka. Stabilnost mreže se obično definiše kao njena otpornost na veliku stopu promene strukture mreže usled čestog dolaska i odlaska parnjaka. U literaturi ove promene se nazivaju talasanja (eng. *churn*). P2P sistem bi trebao biti stabilan do određenog nivoa talasanja, u kontekstu održavanja grafa konekcije i upućivanja (eng. *route*) sadržaja deterministički sa prihvatljivim granicama broja skokova (eng. *hop-count bounds*).

P2P sistem, kao tehnologija u zamahu, privlači veliku pažnju običnih i korporativnih korisnika. Takođe, pored pažnje korisnika, značajno interesovanje beleži se i u istraživačkim krugovima [7]. U poslednjih nekoliko godina, veliki deo istraživanja bavi se P2P deljenjem resursa i prenosom podataka.

2.2 Klasifikacija P2P mreža

Usled velikog broja kriterijuma dizajniranja P2P mreža, ne postoji jedinstven način klasifikacije [16]. Na primer, sistemi za deljenje fajlova često se klasifikuju po generacijama. Prvu generaciju karakteriše hibridni dizajn. Glavnu karakteristiku ove generacije predstavlja kombinacija servera sa P2P upućivanjem sadržaja. Druga generacija je zasnovana na decentralizovanoj arhitekturi. Trećom generacijom smatraju se anonimni P2P sistemi kao što su Freenet [5] i I2P [11]. Ovakav način kategorizacije prema generacijama ima značajnih nedostataka. Glavni nedostatak ogleda se u nedefinisanim važnim dimenzijama problema i nejasnoći dodatnih mogućnosti sledeće generacije. Takođe, sistemi sve tri generacije su korišćeni u isto vreme što se ne može zaključiti na osnovu njihovih imena. Zbog navedenih nedostataka, u literaturi se češće sreće podela prema P2P pokrivaču. Prema ovoj podeli, razlikujemo struktuiranu i nestruktuiranu P2P mrežu [10].

U struktuiranim P2P mrežama topologija je precizno definisana. Takođe, neophodna je striktna kontrola protoka sadržaja. Sadržaj se ne sme nalaziti kod proizvoljnih parnjaka već mora biti na strogo definisanim lokacijama koje će omogućiti efikasniju pretragu. Većina struktuiranih P2P mreža je bazirana na distribuiranim heš tabelama (eng. *Distributed Hash Table (DHT)*). Kao primere struktuiranih P2P mreža navodimo Tapestry [21], Chord [17], Pastry [14] i Kademlia [13].

Ne struktuirani P2P sistem je sačinjen od parnjaka koji se pridružuju mreži sa

nekim labavim skupom pravila, bez ikakvog predznanja o njenoj topologiji. Freenet, Gnutella, FastTrack, KaZaA i BitTorrent su primeri ovakvih sistema. Ovakve mreže se zbog svojih svojstava tipično koriste za udruživanje snage parnjaka nebi li se neki posao brže obavio ili razmenu fajlova kod kojih nije od presudnog značaja brzina pronalaska fajla.

U nestruktuiranim P2P mrežama, pretraga je bazirana na prosleđivanju upita. Svaki parnjak upit prosleđuje komšijama, osim u slučaju kada parnjak može odgovoriti na traženi upit ili se brojač skokova upita smanji na 0 [16]. Sa ovakvim načinom prosleđivanja postoje različita pravila kontrole prosleđivanja upita. Način prosleđivanja može se kontrolisati tako što se upit neće proslediti svim komšijama nego samo određenim. Izbor kome će upit biti prosleđen se vrši prema informacijama koje parnjak čuva o komšijama. Te informacije predstavljene su istorijom ili indeksima sadržaja koji komšija može obezbediti. Takođe, često se prosleđivanje vrši slučajno bez ikakvog znanja. Naravno, ovo svojstvo smanjuje šansu da se pronađe sadržaj. Opravdanje nestruktuiranog pristupa zasniva se na promenljivosti broja članova mreže i ispravnosti prvog izbora komšije.

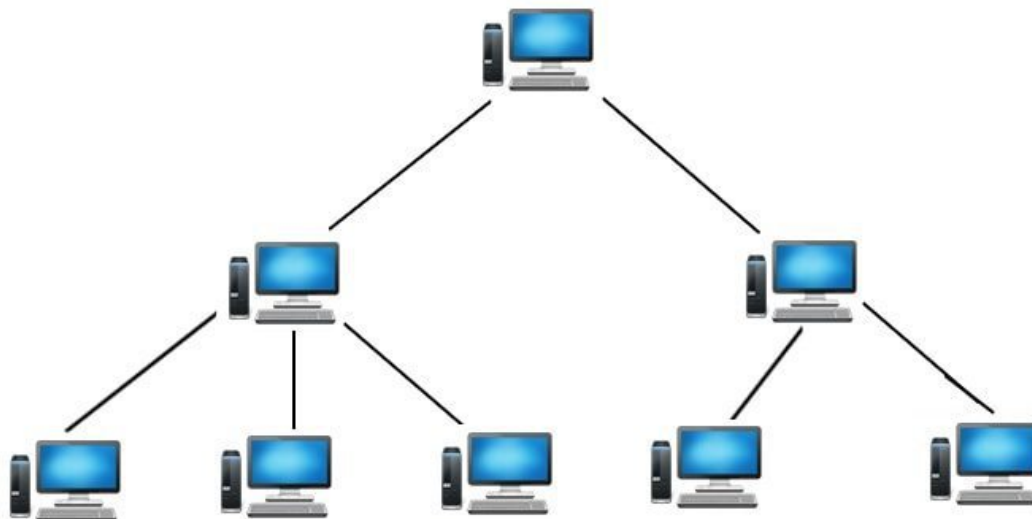
Nestruktuirane mreže ravnopravnih računara uglavnom su sastavljene od parnjaka koji su kućni PC računari. Oni su najčešće povezani sa mrežom slabim propusnim opsegom (eng. *bandwidth*). Takođe, većinom je propusni opseg nesrazmeran u odnosu na dolazni i odlazni saobraćaj u korist dolaznog saobraćaja. Navedena činjenica blago narušava samoskalabilnost mreže. Postoje različiti načini da se spreči narušavanje samoskalabilnosti. Neki od njih su zasnovani na kažnjavanju parnjaka koji ne doprinose ili jako malo doprinose resursima sistema.

2.3 Topologija

P2P sistem je sačinjen od dva podsistema koji se izvršavaju paralelno. To su konstrukcija i održavanje izabrane topologije kao i prosleđivanje bloka (eng. *chunk*) podataka. Algoritmi za konstrukciju i prosleđivanje zajedno formiraju P2P protokol. U zavisnosti od odluka donešenih pri dizajnu sistema tj. od izabrane topologije i algoritma prosleđivanja može se napraviti P2P sistem sa različitim karakteristikama.

Topologija drveta

Među prvim predloženim topologijama P2P sistema za razmenu sadržaja predstavljena je topologija drveta. U okviru ove topologije, svi parnjaci su organizovani tako da čine strukturu drveta. Primer topologije drveta prikazan je na slici 2.1.



Slika 2.1: Primer topologije drveta

U korenu drveta nalazi se parnjak koji deli sadržaj-izvor. Svaki parnjak u sistemu može imati tačno određen broj dece koji je ili unapred definisan ili određen kapacitetom parnjaka. Deca su komšije koje preuzimaju sadržaj od roditelja. Konstrukcija drveta i relacija roditelj-dete može biti određena različitim faktorima kao što su ukupno vreme potrebno da podatak stigne od izvora do svih parnjaka u mreži, dostupan propusni opseg i fizička topologija same mreže. Podaci koji se šalju na ovakav način se najpre podele na manje delove koji se nazivaju paketi. Paket se zatim šalje od izvora do parnjaka koji se nalaze na prvom nivou drveta. Nakon toga parnjaci sa prvog nivoa drveta šalju podatke parnjacima na drugom nivou drveta i tako redom sve dok paket ne dođe do svih parnjaka u mreži [7].

Iako jednostavna za konstrukciju, topologija drveta ima dosta nedostataka. Najveći nedostatak je neravnomerna raspodela tereta prosleđivanja koja se javlja kao posledica činjenice da veliki deo parnjaka čine listovi koji ne daju nikakav doprinos sistemu. Dodatno, parnjak koji nema odlazni (eng. *upload*) kapacitet proporcionalan proizvodu broja dece i veličine bloka koji prenosi, konstantno će nagomilavati

blokove podataka i usporavati ceo sistem. Ovakav parnjak treba da bude lociran jedino u listovima stabla ili veoma blizu listova. Pored navedenih osobina, važno je napomenuti da je ovakva struktura ranjiva na napuštanje parnjaka. Prilikom napuštanja parnjaka na višem nivou, njegova podstabla privremeno ostaju odsečena od ostatka mreže.

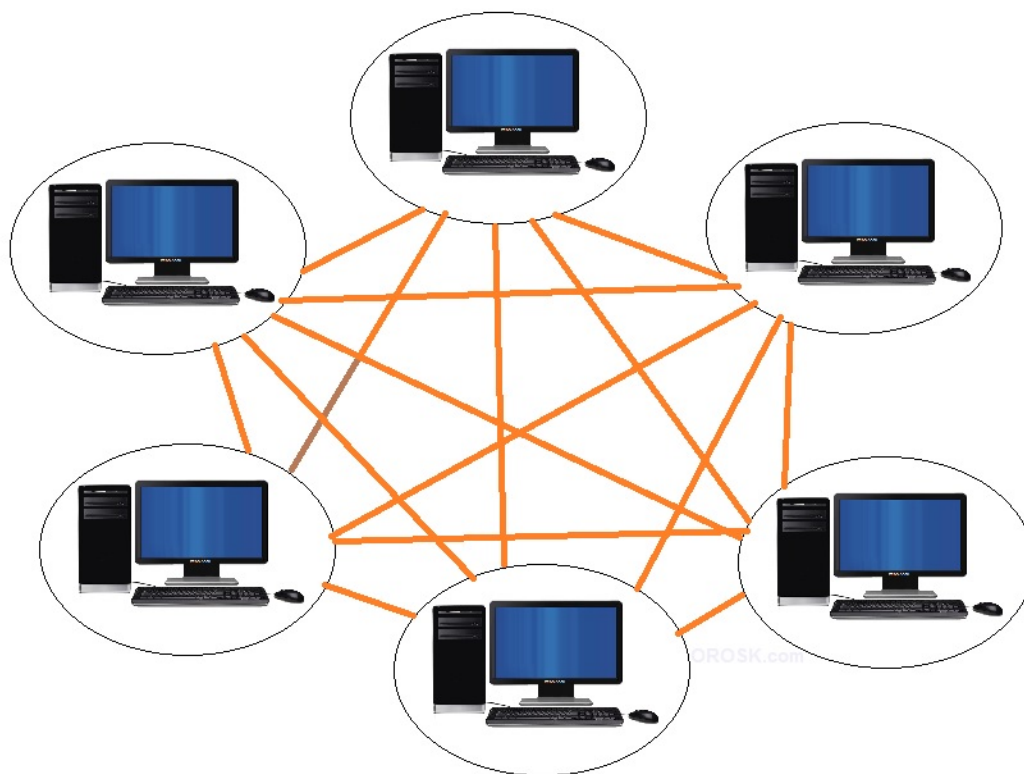
Da bi se povećala otpornost na napuštanja parnjaka i povećao kapacitet mreže koristi se rešenje bazirano na višestrukim stablima. U ovakvim sistemima parnjaci se nalaze u jednom ili više stabala. Svaki parnjak prima sadržaj od svih stabala čiji je član. Izvorni čvor je izvor za sva stabla. Dobijen sadržaj parnjak prosleđuje unapred određenim stablima. Sistem se na ovaj način deli u podsisteme u kojima svaki od podsistema predstavlja jedno stablo. Radi bržeg distribuiranja sadržaja svakom podsistemu se dodeljuje jedinstven paket.

Topologija zasnovana na višestrukim stablima rešava neke od problema upotrebe samo jednog stabla, ali uvodi dodatne probleme [7]. Prvo, na ovakav način postoji više putanja od izvora do lista pa je potreban značajno složeniji algoritam prosleđivanja. Drugo, kako stabla imaju različite podatke potrebno je obezbediti da svi podaci dođu do svakog parnjaka. Treće, prilikom konstrukcije mreže treba kreirati stabla koja su što je moguće manje dubine čime bi se obezbedio brži protok podataka. Dodatno, javlja se povremena potreba prebacivanja parnjaka iz jednog stabla u drugo uz dodatni zadatak balansiranja dubina svih stabala. Četvrti problem leži u balansiranju kapaciteta stabala.

Topologija mreže

Sistem zasnovan na mreži se fokusira na kreiranje topologije koja se brzo adaptira na promene. Glavna ideja pri pravljenju ove topologije bila je da parnjaci provode što je moguće manje vremena za održavanje pokrivača i brzi oporavak mreže pri njihovom napuštanju. Prednost ove topologije je brža i jednostavnija konstrukcija pokrivača. Sa druge strane, u ovakvoj konstrukciji potrebno je uložiti značajno vreme za implementaciju algoritma za prosleđivanje sadržaja između parnjaka. Primer topologije mreže nalazi se na slici 2.2.

Značajan kontrast se može primetiti ukoliko se posmatraju razlike između topologije mreže i topologije drveta 2.3. Kod topologije drveta kreiranje same topologije je složenije i zahteva više vremena, ali jednom kad se postigne željeni dizajn mreže prosleđivanje podataka je jednostavno. Kod topologije mreže situacija je potpuno



Slika 2.2: Primer topologije mreže

drugačija. Sama izrada pokrivača je krajnje jednostavna, jeftina i brza, ali je prosleđivanje podataka značajno kompleksnije.

U mrežnom pokrivaču glavni cilj parnjaka je da održavaju veliki broj dolaznih konekcija [16]. U slučaju izlaska komšije iz mreže, njegov uticaj na parnjake je minimalan. Pokrivač se obično gradi na distribuirani način i svaki parnjak je svestan samo malog dela učesnika, njegovih komšija. Način biranja komšija se razlikuje od politike (eng. *policy*) do politike, pri čemu je u svim slučajevima ovaj proces brz i jednostavan. Primer brze i jednostavne politike je slučajno biranje komšija.

U odnosu na topologiju drвета, prosleđivanje podataka u mreži ne odvija se pravolinijski. Nedostatak konkretne strukture koja bi definisala put kojim paket treba ići čini nemogućim određivanje unapred putanje kojom paket treba proći. Zbog navedenog svojstva prosleđivanje paketa je bazirano na lokalnim odlukama koje svaki parnjak donese. Parnjaci donose odluke prema informacijama koje imaju o komšijama. Paket se može razmeniti primenom guranja (eng. *push*) ili povlačenja (eng. *pull*). Kod guranja parnjak određuje koji paket šalje svakom komšiji, a kod povlačenja određuje koji paket traži od svakog komšije. Oba pristupa imaju dobre

i loše strane.

Pristup sa guranjem je efikasniji u mreži sa ograničenim odlaznim kapacitetom jer izbegava višestruke zahteve za pakete. Pristup sa povlačenjem je dobar izbor za mreže sa ograničenim dolaznim kapacitetom jer parnjak može da kontroliše koliko brzo će primati pakete od komšija. Kod obe šeme zbog lokalne koordinacije propagacije paketa dolazi do blage neefikasnosti u prosleđivanju. Kod sistema guranja ova činjenica se ogleda u višestrukim kopijama koje parnjak može da dobije ako mu više komšija šalje isti paket. Kod sistema povlačenja može se desiti da komšije zaguše parnjaka sa previše zahteva za paketima. Takođe sistemi povlačenja unose dodatan sadržaj u mrežu (eng. *overhead*) s obzirom da se paket mora tražiti.

Hibridna topologija

Hibridna topologija kombinuje prednosti topologija drveta i mreže. Preciznije, predstavlja kombinaciju robusnosti mrežne topologije sa jednostavnošću i efikasnošću prosleđivanja paketa koju nudi topologija drveta.

U hibridnoj topologiji, mreža je podeljena na više podmreža slično kao kod topologije drveta [16]. Da bi parnjak mogao da prima pakete, mora pronaći roditelje koji će mu prosleđivati pakete iz svih podmreža. Kada se novi parnjak pridružuje mreži, on dobija listu svih parnjaka na koje može da se poveže. Zatim parnjak među njima pokušava da odabere komšije tako da pokrije sve podmreže. Nakon što je pronašao takve parnjake povezuje se na njih. Glavna mana ovakvog sistema je ta što je dobijanje pravilne topologije težak zadatak. Često se može desiti da mala greška dovodi do situacije da parnjak ili grupa parnjaka bude izolovana od ostatka mreže.

2.4 Protok video sadržaja

Od samog nastanka, video je važan mediji za komunikaciju i zabavu. Sačinjen je od serije slika koje se smenjuju odgovarajućom brzinom dajući privid neprekidnog kretanja. Ovaj trik bio je poznat još u drugom veku u Kini, ali je ostao nepoznat ostatku sveta sve do 19. veka. Otkrićem kamere 1888. godine omogućeno je „automatsko” hvatanje i čuvanje pojedinačnih komponenti slike na filmskoj traci. Emitovanje televizijskog signala, nakon izuma 1928. godine, omogućilo je bilionima ljudi širom sveta uživo praćenje snimljenog sadržaja. Zahvaljujući širokoj dostup-

nosti signala, primarni vid zabave i dobijanja informacija umesto novina i radija postala je televizija.

Tokom većeg dela 20. veka, jedini način dobijanja televizijskog signala bio je preko vazduha ili kabla. Ranih 2000-tih internet doživljava veliki rast propusnog opsega. Pored rasta propusnog opsega svakodnevno se razvijaju bolji algoritmi kompresije video snimaka čime se omogućava da se sa značajno manje podataka predstavi isti snimak bez velikih gubitaka na kvalitetu. Takođe, značajnu ulogu u svemu igra i Murov zakon [15] činjenicom da eksponencijalni porast snage računara omogućava rešavanje raznovrsnijih problema. Navedena svojstva omogućavaju da sistem za dostavu video sadržaja u realnom vremenu (eng. *streaming*) putem interneta postane moguć. Dostava video sadržaja u realnom vremenu omogućava emitovanje pre preuzimanja celokupnog sadržaja. Video se neprekidno šalje i emitovanje se omogućava odmah po pristizanju ili sa određenom zadržkom [18].

Internet je dizajniran za slanje paketa bez mogućnosti kontrole toka. Ovakav način transfera ne pogoduje slanju vremenski neprekidno baziranog saobraćaja kao što su video i audio sadržaji [19]. Glavni razlog je što dostava u realnom vremenu ima određeni poredak kojim se mora emitovati kao i vremenska ograničenja. Na primer, video sadržaj se mora emitovati neprekidno slika za slikom. U slučaju da podatak ne stigne na vreme, proces emitovanja će biti prekinut.

Danas se internet sve više koristi za razmenu multimedijalnog sadržaja umesto statičnog teksta i grafika [2]. Aplikacije koje zahtevaju dostavu u realnom vremenu velikom broju korisnika su, između ostalog, Internet TV, prenos sportskih događaja, online igrice i edukacija preko daljine. Usled ovako širokog spektra primena istraživači već skoro trideset godina istražuju odgovarajuću podršku aplikacijama u obliku IP višesmernog emitovanja (eng. *multicast*). IP višesmerno emitovanje je vrsta slanja paketa u kome se dodaje posebna zastavica (eng. *flag*) u IP pakete kojom se signalizira ruterima da je paket potrebno poslati grupi računara paralelno [8]. Međutim, usled ozbiljnih problema sa skaliranjem i podrškom na višim nivoima funkcionalnosti, višesmerno emitovanje nije zaživelo šire. Visoki troškovi mreža za isporuku sadržaja (eng. *Content Delivery Networks (CDN)*) i obezbeđivanja odgovarajućeg propusnog opsega su dva glavna faktora koji ograničavaju ovakav vid slanja samo na mali deo izdavača Internet usluga, koji višesmerno emitovanje prevashodno koriste za dostavu kvalitetnog video sadržaja IPTV. Za dostavu sadržaja pretplatnicima izdavači koriste usmerivače paketa (eng. *packet switching*).

U nadolazećim godinama, postoji značajni interes za korišćenje mreže ravno-

pravnih računara za dostavu sadržaja u realnom vremenu. Dve ključne stavke čine pomenuti pristup primamljivim. Prvo, ovakva tehnologija ne zahteva podršku mrežne infrastrukture što je čini jeftinom i jednostavnom za izradu. Drugo, u takvoj tehnologiji, svi koji primaju sadržaj isti taj sadržaj dele dalje, što obezbeđuje visok nivo skalabilnosti. Pored velikog interesa i uloženog truda u pronalaženje odgovarajućeg P2P sistema, ova tema je još uvek otvorena. Glavna tačka spoticanja P2P mreže za dostavu sadržaja u realnom vremenu je susret sa drugačijim poteškoćama u odnosu na uobičajne probleme mreže. Kod dostave sadržaja u realnom vremenu posebno se mora voditi računa o kašnjenju signala, dok u slučaju preuzimanja sadržaja kašnjenje nije kritično. Zapravo, prihvatljivo je preuzimanje sadržaja u trajanju od nekoliko sati pa čak i dana. Različitost i striktniji zahtevi prenosa video sadržaja u realnom vremenu zahteva fundamentalno drugačiji dizajn i pristup.

2.5 Modelovanje mreže

U idealnom slučaju, P2P sistem za dostavu sinhronizovanog sadržaja bi trebao postići jednako dobre performanse uzimajući u obzir sve bitne metrike. Pri postizanju ovog cilja neophodno je prevazići različite izazove povezane sa okruženjem u kojem P2P sistem radi i korisničkim ponašanjem. Izazovi su mnogobrojni i različiti pri čemu u nastavku navodimo četiri glavna.

Prvo, P2P sistem treba biti što je moguće otporniji na talasanje. Talasanje može da utiče na kontinualnost signala koji se prikazuje korisniku uzimajući u obzir činjenicu da napuštanje parnjaka može usporiti ili potpuno onemogućiti distribuciju podataka u mreži. Kada parnjak napusti mrežu, svi parnjaci kojima je on bio dobavljač sadržaja moraju pronaći novog dobavljača u što kraćem roku.

Drugo, ovakav sistem treba imati dobro skaliranje. Kako mreža raste, povećava se broj parnjaka kao i ukupno vreme potrebno da se sadržaj pošalje svim parnjacima. Za nesmetanu reprodukciju ovakvog sadržaja, kašnjenje reprodukcije takođe treba da raste. Kada se u obzir uzme i skalabilnost, izazov je dvostruki. Dodatno, porast prosečnog kašnjenja reprodukcije sadržaja treba biti takav da nema veliki uticaj na ukupno iskustvo korisnika (eng. *user experience*). Drugo, sistem treba održavati što manje kašnjenje čime se obezbeđuje bolji privid prenosa uživo.

Treći izazov sa kojim se susreću P2P sistemi su različiti propusni opsezi parnjaka. Uspešnost samog sistema zavisi od prosečnog odlaznog kapaciteta parnjaka i jačine enkodovanja video snimka. Danas, propusni opsezi parnjaka mogu biti ve-

oma različiti. Dodatnu poteškoću predstavlja asimetrična priroda pristupa internetu koja dovodi do toga da parnjaci značajno više sadržaja mogu preuzeti nego poslati. Zbog ovih razloga, P2P sistem treba biti izgrađen tako da se heterogenost propusnog opsega koristi efektivno.

Četvrti izazov je da prosečna brzina slanja parnjaka može biti smanjena usled nevoljnosti parnjaka za doprinos mreži. Parnjaci koji žele da prime podatke ali ne i da ih dalje dele sa ostatkom mreže se nazivaju grebatori (eng. *free-riders*). Grebatori su opasnost za sistem jer koriste resurse bez daljeg deljenja pa samim tim smanjuju prosečnu brzinu deljenja u mreži. Stoga, važno je da se P2P sistem dizajnira tako da podstiče doprinos mreži.

Glava 3

Predloženi Kikkar P2P protokol

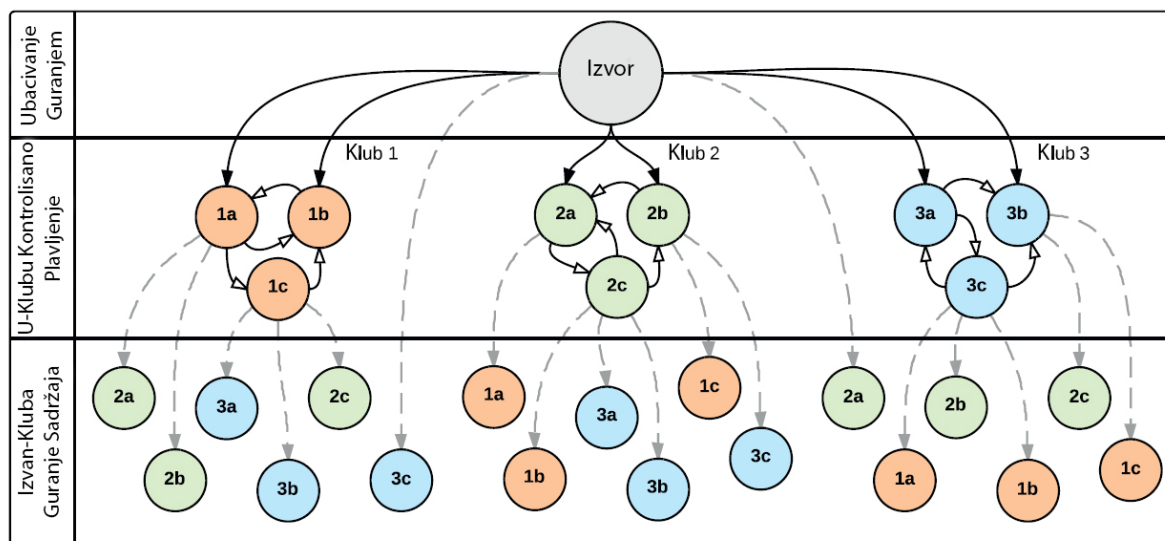
U poglavlju 2 navedeno je da distribucija medijskog sadržaja preko interneta privlači veliku pažnju u poslednjoj deceniji. Navedena činjenica dovela je do detaljnijeg istraživanja teme od strane naučnika i industrije. Veliki broj različitih P2P protokola za dostavu video sadržaja u realnom vremenu predstavljen je u prethodnim godinama [20]. Pored velikog broja predloženih protokola, samo pojedini imaju neki vid sinhronizacije. Ostali protokoli vode se logikom da će sadržaj biti relativno sinhronizovan. Relativna sinhronizacija uzrokuje da se sadržaj koji se prikazuje na parnjacima može značajno razlikovati.

Protokol Kikkar ¹ je baziran na Screamer protokolu [6]. Screamer koristi hibridnu topologiju koja omogućava dostavu video sadržaja u realnom vremenu sa malim zastojem (eng. *delays*) i malim dodatnim saobraćajem u mreži (eng. *overhead*). Zbog ovakvih karakteristika posebno je pogodan za dostavu video sadržaja u realnom vremenu. Važno je napomenuti da sam Screamer protokol nema nikakav mehanizam kojim bi se kontrolisala sinhronizacija sadržaja među parnjacima, te Kikkar menja neke delove i donosi nove kako bi omogućio sinhronizaciju.

3.1 Arhitektura mreže

Protokol Kikkar je baziran na protokolu Screamer. Odavde zaključujemo da se Kikkar klasifikuje kao hibridni pokrivač koji koristi brojne mehanizme kako bi postigao što manji zastoj i dodatni saobraćaj u mreži. Na slici 3.1 prikazana je struktura mreže za dostavu video sadržaja u realnom vremenu za jednu konfiguraciju protokola. Uočavamo da su parnjaci podeljeni u klubove. Jedan parnjak može da

¹Kikkar (hebrejski kružni okrug, kompletna težina)



Slika 3.1: Screamer topologija

pripada najviše jednom klubu. Ukupan broj klubova je 6. Proces dostave video sadržaja se može apstrahovati na tri faze: (1) ubacivanje guranjem video blokova od izvora (eng. *source*) u ostale klubove, (2) kontrolisano plavljenje (eng. *flooding*) u klubu, i (3) guranje sadržaja u listove mreže, tj. guranje sadržaja izvan kluba.

Kikkar protokol razlikuje tri glavne uloge u sistemu: tragač (eng. *tracker*), parnjak koji proizvodi i distribuira video snimak (eng. *source*) i parnjaci koji traže video snimak. Tragač služi da pomogne parnjacima u bržem i lakšem pronalaženju ostalih parnjaka. Bitno je napomenuti da se tragač koristi samo za otkrivanje novih parnjaka i sinhronizaciju časovnika u mreži.

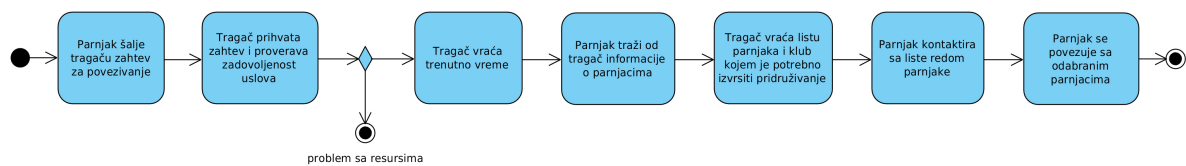
Kod Kikkar protokola, svi parnjaci su podeljeni u klubove, tj. svaki parnjak pripada jednom klubu. Dodeljivanje veze parnjak-klub vrši tragač prilikom pristupanja parnjaka sistemu. Parnjak doprinosi klubu kome pripada svojim resursima. Ostalim klubovima parnjak je list tj. ništa ne doprinosi već samo dobija sadržaj. Parnjak koji šalje video igra posebnu ulogu u sistemu, tj. pripada svim klubovima. Takođe, on je zadužen za podelu video snimaka na manje delove i dodelu delova određenim klubovima. Ovakav način slanja je čest kod topologije višestrukih stabala.

Parnjak komunicira sa tragačem upotrebom HTTP protokola. Parnjaci komuniciraju međusobom koristeći protokol korisničkih datagrama (eng. *User Datagram Protocol, UDP*). Veze između parnjaka se formiraju tako što se povezuju parnjak

koji šalje i parnjak koji prima. Važi pravilo da parnjak koji je na početku veze odabran da šalje video ne prima sadržaj od parnjaka primaoca. Drugim rečima veze su jednosmerne.

Svaki parnjak u klubu teži da ima bar dve i ne više od tri odlazne i dolazne konekcije sa parnjacima koji pripadaju njegovom klubu. Ovim svojstvom se postiže eksponencijalna brzina slanja video snimka. Takođe, pored parnjaka iz njegovog kluba teži i da postigne po jednu dolaznu konekciju sa parnjacima koji ne pripadaju njegovom klubu. Ova konekcija je istog značaja kao i konekcija u klubu. Bez njenog postojanja, ne mogu se obezbediti svi potrebni podaci. Broj odlaznih konekcija sa parnjacima koji ne pripadaju njegovom klubu nije bitan. Konekcije će detaljnije biti opisane u pogavlju 3.4.

3.2 Pridruživanje parnjaka



Slika 3.2: Pridruživanje novog parnjaka

Prilikom pridruživanja parnjaka jatu (eng. *swarm*) kako bi preuzeo sadržaj u mreži, neophodno je slanje zahteva za povezivanje tragaču. Navedeno svojstvo ilustrovano je dijagramom aktivnosti 3.2. Kako bi se jato zaštitilo od pridruživanja parnjaka koji ne može obezbediti odgovarajuće resurse, tragač proverava brzinu dolaznog i odlaznog saobraćaja novog parnjaka. Ukoliko je brzina dolaznog i odlaznog saobraćaja ispod granice, tragač prekida kontakt sa parnjakom. Ukoliko je brzina dolaznog i odlaznog saobraćaja zadovoljavajuća tragač vraća trenutno vreme parnjaku korišćenjem protokola mrežnog vremena (eng. *Network Time Protocol, NTP*). Nakon uspešne sinhronizacije vremena, parnjak od tragača zahteva informaciju o parnjacima koji pripadaju jatu. Tragač izdvaja iz svakog kluba određeni broj parnjaka i prosleđuje njihove informacije. Informacija sadrži IP adresu parnjaka, port na kome parnjak osluškuje i pripadnost parnjaka klubu. Dodatno, pored informacije o parnjacima tragač prosleđuje parnjaku klub u koji je obavezan da uđe. Kako bi se obezbedio jednak broj parnjaka po klubovima, novi parnjak se uvek upućuje na prvi sledeći klub na koji je poslednji parnjak prosleđen.

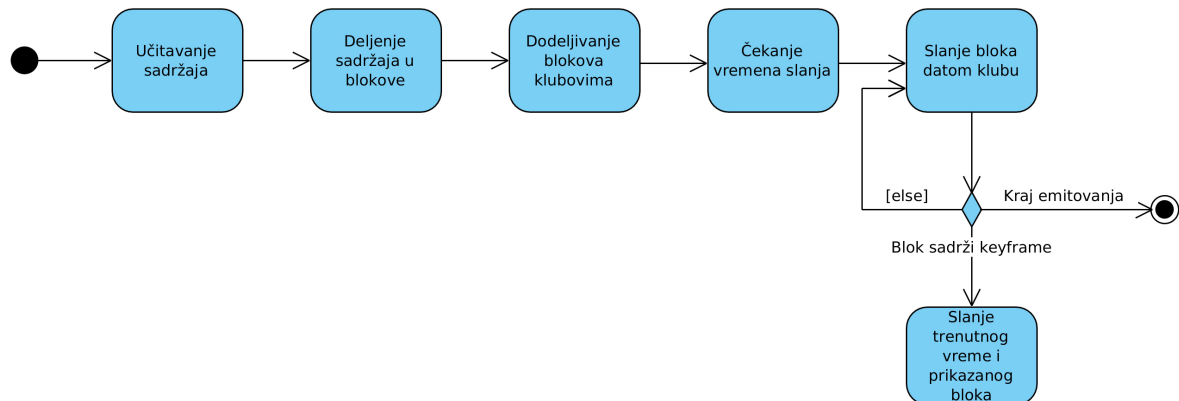
Parnjak zatim koristi dobijenu listu parnjaka kako bi odlučio sa kojima od njih će se povezati. Najpre se šalje **ping** poruka za povezivanje svim parnjacima sa liste. Ping poruka sadrži broj porta na kojem parnjak osluškuje, broj kluba u kome se parnjak nalazi i tip konekcije koju parnjak želi da ostvari tj. odlazna ili dolazna konekcija. Zatim se čekaju odgovori određeno vreme. Nakon pristizanja ping poruke, parnjak odgovara **pong** porukom koja obavezno sadrži informacije o broju veza koje ima ping-ovani parnjak i količini uskladištenih podataka. Nakon određenog vremena, parnjak pregleda primljene pong poruk i sortira parnjake prema poželjnosti. Uslov koji se koristi da bi se poredili parnjaci prema poželjnosti je količina uskladištenih podataka. U slučaju postojanja više parnjaka sa istom količinom uskladištenih podataka bira se parnjak sa što manjim brojem dolaznih ili odlaznih veza kako bi se što manje remetilo jato.

Zatim se odgovarajućim parnjacima šalje **request** poruka. Ova poruka sadrži informaciju o klubu u kome se parnjak trenutno nalazi kao i tip konekcije. Trenutni klub i tip konekcije su prosleđeni ranije u okviru ping poruke ali kako je moguće da parnjak u međuvremenu dobije neki drugi klub ili port ponovo se šalju. Na request poruku odgovara se **response** porukom. Response poruka predstavlja potvrdu o povezivanju. Ukoliko izostane response poruka povezivanje se i dalje smatram uspešnim pri čemu sistem za kontrolu zagušenja ima zadatak da neaktivne parnjake izbaci iz mreže.

Svi parnjaci koji su se javili pong porukom i nisu izabrani se čuvaju u specijalnoj strukturi podataka. Glavni razlog čuvanja je brže kasnije povezivanje i manje opterećivanje servera ukoliko se kasnije ukaže potreba za povezivanjem. Parnjaci koji nisu poslali pong poruku se brišu jer se smatraju neaktivnim ili zagušenim.

3.3 Distribucija sadržaja

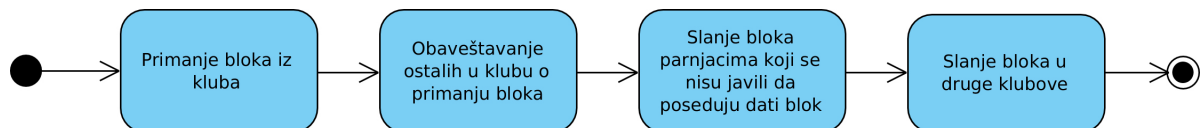
U delu 3.1 izvor se definiše kao parnjak koji generiše i šalje video sadržaj. Video sadržaj može biti prethodno snimljen pa samo emitovan ili uživo snimljen i emitovan. Na slici 3.4 predstavljen je dijagram aktivnosti izvora. Najpre je neophodno da izvor učita sadržaj. Zatim se učitani sadržaj deli na manje blokove koji su pogodniji za prenošenje putem interneta. Svaki od blokova sadrži svoj redni broj i prema tom broju se dodeljuje jednom od klubova. Dodeljivanje se vrši tako što se redni broj podeli po modulu sa brojem klubova. Dobijeni broj odgovara broju kluba kome dati blok pripada. Dodatno, blok sadrži redni broj dela video snimka koji se šalje,



Slika 3.3: Distribucija sadržaja od izvora ka korisnicima

informaciju da li je blok keyframe kao i sam video snimak.

Nakon dodeljivanja bloka klubu čeka se vreme slanja. Vreme čekanja odgovara ukupnom broju blokova koje je potrebno poslati podeljeno sa vremenom potrebnim za slanje blokova. Po isticanju vremena, blok se šalje u klub. Ukoliko poslati blok sadrži keyframe, pre njega se šalje **kontrolna** poruka. Kontrolna poruka sadrži broj video snimka koji je prikazan, dužinu do sada prikazanog video snimka izvora kao i vreme kada je kontrolna poruka napravljena. Osnovna svrha kontrolne poruke je da obavesti ostale parnjake u jatu o trenutno prikazanom sadržaju, tj. da omogućiti globalnu sinhronizaciju jata.



Slika 3.4: Distribucija sadržaja od izvora ka korisnicima

Kod distribucije sadržaja među korisnicima razlikujemo dva načina distribucije. Podela je izvršena u zavisnosti da li dobijeni sadržaj pripada klubu ili ne. Kada parnjak primi blok podataka koji je dodeljen njegovom klubu, on obaveštava o isporuci sve parnjake u klubu sa kojima ima odlaznu vezu a zatim snima dati blok u skladište. Obaveštenje se vrši preko **have** poruke koja sadrži broj primljenog bloka. Obaveštavanje ostalih parnjaka ima za cilj da smanji mogućnost višestruke isporuke istog sadržaja pojedinim parnjacima. Posle određenog vremena parnjak šalje sadržaj svim parnjacima od kojih nije dobio nikakvu poruku koja bi ga obavestila da već ima taj sadržaj. U poruci koja služi za obaveštenje o postojanju sadržaja nalazi se

broj prisutnog bloka. Nakon što je sadržaj poslat u klub, sadržaj se šalje ostalim klubovima. Kako parnjak ima po jednu vezu sa parnjacima iz drugih klubova, nema potrebe za razmenom kontrolne poruke. Kada parnjak dobije blok koji ne pripada njegovom klubu, snima dati blok u skladište i ne distribuira blok dalje.

Prethodno opisan metod distribucije blokova u jatu funkcioniše bez greške u idealnim uslovima. Kako se u okviru mreže javljaju zagušenja i gube paketi, postoji mogućnost da deo snimka ne stigne do parnjaka. Da bismo obezbedili što bolji kvalitet snimka koristi se mehanizam traženja izgubljenih blokova. Pri primanju kontrolne poruke, parnjak proverava stanje prethodne grupe blokova. Za svaki nedostajući blok parnjak šalje zahtev za slanje svim komšijama u klubu. Zahtev sadrži listu brojeva nedostajućih blokova. Kako bi se izbeglo zatrpavanje komšija parnjak može tražiti najviše 20 nedostajućih blokova.

3.4 Kontrola zagušenja

U Kikkar protokolu, radi kontrolisanja zagušenja može doći do raskidanja konekcije pri određenim uslovima. Za raskidanje veze se koristi **terminirajuća** poruka. Ova poruka sadrži razlog terminacije (na primer dolazak novog parnjaka, prespora veza, izlazak iz jata). Jedna od okolnosti pri kojoj može doći do raskidanja konekcije je preveliko kašnjenje paketa. Ukoliko se preko dolazne konekcije ne dobije ni jedna poruka za unapred definisano vreme veza se smatra mrtvom i prekida se. Tačno vreme čekanja zavisi od same mreže i postavlja ga administrator mreže. Merenje vremena počinje dolaskom poslednje poruke. Pri svakoj dobijenoj poruci proveravaju se svi parnjaci. U slučaju da neki od parnjaka nije poslao poruku određeno vreme, veza se raskida. Ukoliko parnjak nema sadržaj za slanje a želi da spreči raskidanje veze, on šalje **keep-alive** poruku kojom obaveštava dobavljača sadržaja da je sa vezom sve u redu. Sa ciljem da bude što manja, keep-alive poruka sadrži samo redni broj koji je jedinstveno identifikuje.

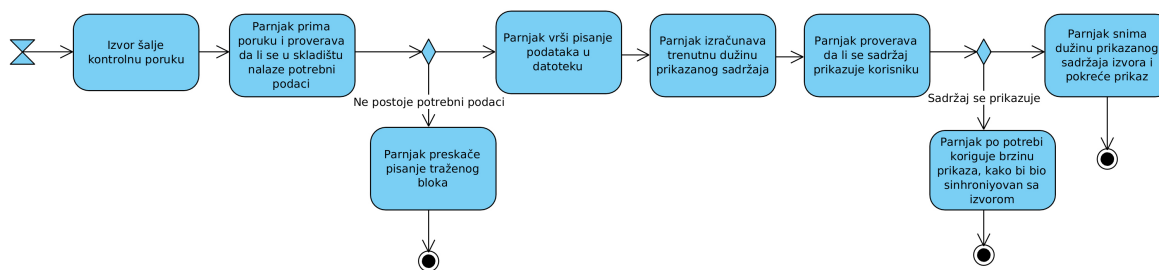
Druga okolnost pod kojom je dozvoljeno raskidanje veze je dolazak paketa u pogrešnom redosledu. Svaki paket ima svoj broj. Ukoliko paketi pristižu u pogrešnom redosledu to može ukazivati na postojanje zagušenja pa će paketi ići alternativnim putevima. U tom slučaju, radi rasterećenja mreže veza se raskida. Dodatno, mali broj paketa koji pristižu u pogrešnom redosledu za veći vremenski period ne mora ukazivati na kritično zagušenje. Treći način rešavanja problema sa zagušenjem je odbacivanje svih veza u kojima se blok podataka predugo čeka. Predugo čekanje

obično ukazuje na usko grlo u sistemu. Radi određivanja trajanja pauze u pristizanju paketa, čuvamo vreme kada je poslednji paket primljen.

Očigledno je da odbacivanjem neke veze možemo pokvariti topologiju i da je potrebno vršiti nova povezivanja. U zavisnosti od veze koja je odbačena, zavisi da li će doći do ponovnog povezivanja. Ukoliko je parnjak izgubio dolaznu vezu van kluba, onda od tragača traži spisak parnjaka i povezuje se sa parnjakom iz istog kluba kome je pripadao parnjak čija je veza izgubljena. Povezivanje se vrši postupkom koji je opisan u delu 3.2. Ukoliko je parnjak izgubio odlaznu vezu van kluba, ne vrši se nikakvo povezivanje. U slučaju da su parnjaci u istom klubu, stvari su komplikovanije. U slučaju da parnjak izgubi dolaznu vezu u klubu, proverava da li su mu ostale najmanje još dve dolazne veze. Ukoliko nisu pronalazi se novi parnjak i vrši se povezivanje koje je opisano u delu 3.2. Isti princip važi i u slučaju gubljenja dolazne veze.

Pored problema sa manjkom veza, može se javiti i višak veza parnjaka. Kada parnjak dobije zahtev za povezivanje, povezivanje se ostvaruje ne uzimajući u obzir prethodne veze. Ukoliko parnjak ima više od tri dolazne ili tri odlazne veze u klubu potrebno je neke veze odbaciti. Odbacuju se one veze koje najduže nisu prenele paket.

3.5 Sinhronizacija parnjaka



Slika 3.5: Sinhronizacija sadržaja

U Kikkar protokolu izvor je pored slanja video sadržaja zadužen i za slanje kontrolnih poruka. Na slici 3.5 ilustrovan je način obrade kontrolne poruke. Kontrolna poruka se šalje u pravilnim razmacima pre svakog keyframe-a. Za razliku od ostalih poruka, kontrolna poruka se zbog svoje važnosti šalje u sve klubove. Nakon primanja kontrolne poruke, parnjak je prosleđuje svim komšijama sa kojima ima odlaznu

vezu, bez obzira da li su je oni već primili i da li su u njegovom klubu. Na ovaj način se osigurava da poruka stigne do svakog parnjaka. Nakon primanja kontrolne poruke koja ima isti sadržaj kao i neka od prethodnih poruka parnjak prekida njeno dalje prosleđivanje. Na ovaj način se sprečava beskonačno slanje poruke kroz jato. Nakon primanja kontrolne poruke, parnjak proverava da li skladište sadrži odgovarajuće video blokove koji su porukom obuhvaćeni. Ukoliko skladištu nedostaje prvi obuhvaćeni blok, preskače se ceo opseg. Kako se sve značajne informacije za prikazivanje video snimka nalaze u prvom bloku, prethodna činjenica je potpuno opravdana. U slučaju postojanja prvog bloka, obuhvaćeni blokovi se šalju video plejeru. Kako bi se sprečilo prekidanje video snimka, u slučaju preskakanja nekog video bloka prikazuje se prethodni video blok.

Ukoliko je uspešno izvršeno slanje sadržaja video plejeru, prelazi se na korak sinhronizacije. U koraku sinhronizacije se proverava da li video plejer prikazuje sadržaj. U slučaju neaktivnosti multimedijalnog plejera, snima se vrednost prikazanog sadržaja izvora sabrana sa kašnjenjem paketa. Ova vrednost će se koristiti kasnije kao reper za eventualne korekcije. Nakon što je vrednost snimljena pokreće se multimedijalni plejer. Ukoliko je plejer već aktivan, izračunava se dužina do sada prikazanog sadržaja. Ukoliko je ta dužina manja od zbira dužine emitovanja sadržaja izvora i kašnjenja paketa, potrebno je povećati brzinu snimka određeni vremenski interval dok se ne sinhronizuje sa izvorom. Slično, u slučaju da je pomenuta dužina veća od zbira dužine emitovanja sadržaja izvora i kašnjenja paketa, brzina se smanjuje određeni vremenski interval dok se ne postigne sinhronizacija sa izvorom.

Glava 4

Zaključak

Bibliography

- [1] K. Aberer, A. Datta, and M. Hauswirth. „Efficient, Self-Contained Handling of Identity in Peer-to-Peer Systems”. In: *IEEE Transactions on Knowledge and Data Engineering* 16.07 (July 2004), pp. 858–869. DOI: 10.1109/tkde.2004.1318567. URL: <https://doi.org/10.1109%2Ftkde.2004.1318567>.
- [2] Josh Beggs and Dylan Thede. *Designing Web Audio*. Ed. by Richard Koman. 1st. Sebastopol, CA, USA: O'Reilly & Associates, Inc., 1999. ISBN: 1565923537.
- [3] *Bittorrent Opis*. <http://www.bittorrent.com/>.
- [4] Fernando Bordignon and Gabriel Tolosa. „Gnutella: Distributed System for Information Storage and Searching Model Description”. In: (Aug. 2001).
- [5] Ian Clarke et al. „Freenet: A Distributed Anonymous Information Storage and Retrieval System”. In: *Designing Privacy Enhancing Technologies*. Springer Berlin Heidelberg, 2001, pp. 46–66. DOI: 10.1007/3-540-44702-4_4. URL: https://doi.org/10.1007%2F3-540-44702-4_4.
- [6] Bram Cohen. *Peer-to-Peer live streaming*. US Patent 4/801,778. May 2011. URL: <https://patents.google.com/patent/US20150326657>.
- [7] Jorn De Boever. „Peer-to-Peer Networks as a Distribution and Publishing Model”. In: (Jan. 2007).
- [8] Stephen E. Deering and David R. Cheriton. „Multicast Routing in Datagram Internetworks and Extended LANs”. In: *ACM Trans. Comput. Syst.* 8.2 (May 1990), pp. 85–110. ISSN: 0734-2071. DOI: 10.1145/78952.78953. URL: <http://doi.acm.org/10.1145/78952.78953>.
- [9] *FastTrack Opis*. <http://ntrg.cs.tcd.ie/undergrad/4ba2.02/p2p/index.html>.

- [10] Prasanna Ganesan, Krishna Gummadi, and H. Garcia-Molina. „Canon in G major: designing DHTs with hierarchical structure”. In: *24th International Conference on Distributed Computing Systems, 2004. Proceedings.* IEEE, 2004. DOI: 10.1109/icdcs.2004.1281591. URL: <https://doi.org/10.1109%2Ficdcs.2004.1281591>.
- [11] *I2P Dokumentacija*. <https://geti2p.net/en/docs>.
- [12] Jian Liang, Rakesh Kumar, and Keith W Ross. „The kazaa overlay: A measurement study”. In: (Oct. 2004).
- [13] Petar Maymounkov and David Mazières. „Kademlia: A Peer-to-Peer Information System Based on the XOR Metric”. In: *Peer-to-Peer Systems*. Springer Berlin Heidelberg, 2002, pp. 53–65. DOI: 10.1007/3-540-45748-8_5. URL: https://doi.org/10.1007%2F3-540-45748-8_5.
- [14] Antony Rowstron and Peter Druschel. „Pastry: Scalable, Decentralized Object Location, and Routing for Large-Scale Peer-to-Peer Systems”. In: *Middleware 2001*. Springer Berlin Heidelberg, 2001, pp. 329–350. DOI: 10.1007/3-540-45518-3_18. URL: https://doi.org/10.1007%2F3-540-45518-3_18.
- [15] Robert R. Schaller. *Moore’s law: Past, present, and future*. IEEE Spectr.,34(6):52–59, 1997.
- [16] Xuemin Shen et al. *Handbook of Peer-to-Peer Networking*. 1st. Springer Publishing Company, Incorporated, 2009. ISBN: 0387097503, 9780387097503.
- [17] I. Stoica et al. „Chord: a scalable peer-to-peer lookup protocol for internet applications”. In: *IEEE/ACM Transactions on Networking* 11.1 (Feb. 2003), pp. 17–32. DOI: 10.1109/tnet.2002.808407. URL: <https://doi.org/10.1109%2Ftnet.2002.808407>.
- [18] „Telecom, ATIS Telecom Glossary 2007”. In: *American National Standard for Telecommunications* (2010). URL: <http://www.atis.org>.
- [19] David J. Wetherall and Andrew S. Tanenbaum. *Computer Networks*. Pearson Education, 2013. ISBN: 1292024224, 9781292024226.
- [20] Xiangyang Zhang and Hossam Hassanein. „A Survey of Peer-to-peer Live Video Streaming Schemes - An Algorithmic Perspective”. In: *Comput. Netw.* 56.15 (Oct. 2012), pp. 3548–3579. ISSN: 1389-1286. DOI: 10.1016/j.comnet.2012.06.013. URL: <http://dx.doi.org/10.1016/j.comnet.2012.06.013>.

BIBLIOGRAPHY

- [21] B.Y. Zhao et al. „Tapestry: A Resilient Global-Scale Overlay for Service Deployment”. In: *IEEE Journal on Selected Areas in Communications* 22.1 (Jan. 2004), pp. 41–53. DOI: 10.1109/jsac.2003.818784. URL: <https://doi.org/10.1109%2Fjsac.2003.818784>.