

ŽILINSKÁ UNIVERZITA V ŽILINE

**AUTOREFERÁT
DIZERTAČNEJ PRÁCE**

Žilina, Apríl 2013

Ing. Dušan Nemček

Žilinská univerzita v Žiline
Fakulta riadenia a informatiky

Ing. Dušan Nemček

Autoreferát dizertačnej práce

Modelovanie vybraných mechanizmov garancie kvality služieb v komunikačných sieťach

na získanie akademického titulu „**philosophiae doctor**“ (v skratke **PhD.**)
v študijnom programe doktorandského štúdia
aplikovaná informatika

v študijnom odbore:
9.2.9 aplikovaná informatika

Žilina, Apríl 2013

Dizertačná práca bola vypracovaná v dennej forme doktorandského štúdia na katedre Informačných sietí, Fakulte riadenia a informatiky Žilinskej univerzity v Žiline.

Predkladateľ: Ing. Dušan Nemček
Žilinská univerzita v Žiline
Fakulta riadenia a informatiky
Katedra informačných sietí

Školiteľ: doc. RNDr. Katarína Bachratá, PhD.
Žilinská univerzita v Žiline
Fakulta riadenia a informatiky
Katedra informačných sietí

Oponenti: prof. Ing. Stanislav Marchevský, CSc.
FEI Technická univerzita Košice

doc. Ing. František Jakab, PhD.
FEI Technická univerzita Košice

doc. Ing. Ivan Kotuliak, PhD.
prodekan pre domácu a zahraničnú spoluprácu a propagáciu fakulty
Fakulta informatiky a informačných technológií STU v Bratislave

Autoreferát bol rozoslaný dňa:

Obhajoba dizertačnej práce sa koná dňa o h. pred komisiou pre obhajobu dizertačnej práce schválenou odborovou komisiou v študijnom odbore **9.2.9 aplikovaná informatika, v študijnom programe aplikovaná informatika**, vymenovanou dekanom Fakulty riadenia a informatiky Žilinskej univerzity v Žiline dňa

prof. Ing. Martin Klimo, PhD.
predseda odborovej komisie
študijného programu **aplikovaná informatika**
v študijnom odbore **9.2.9 aplikovaná informatika**
Fakulta riadenia a informatiky
Žilinská univerzita
Univerzitná 8215/1
010 26 Žilina

Abstrakt

Na zabezpečenie kvality služieb v súčasných počítačových sieťach sa využíva veľké množstvo mechanizmov garancie kvality služieb (tzv. *QoS mechanizmov*). Základným parametrom sieťovej výkonnosti, ktorý sa podieľa na posudzovaní kvality služieb v komunikačnej sieti je oneskorenie dát pri prenose sieťou. Významný podiel na celkovom oneskorení v sieti má tzv. frontové oneskorenie, ktoré vzniká v dôsledku čakania paketov v softvérových frontoch sieťových zariadení. Do týchto frontov sú pakety radené mechanizmami používanými pre riadenie zahľtenia siete. Mechanizmy riadenia zahľtenia siete predstavujú významnú skupinu QoS mechanizmov používaných na zabezpečenie kvality služieb v konvergovaných počítačových sieťach, ktoré sú založené na architektúre diferencovaných služieb.

Moja práca sa zoberá modelovaním vybraných mechanizmov riadenia zahľtenia. Modelovanie je zamerané na odhad frontového ako aj celkového oneskorenia v IP sieti s Triple-Play tokmi. Konkrétne sa jedná o mechanizmy známe pod označeniami FIFO (ang. *First-In, First-Out*) a WRR (ang. *Weighted Round Robin*). Výstupom práce sú matematické a simulačné modely týchto mechanizmov. Správnosť a presnosť modelov bola overená praktickými experimentmi na reálnych zariadeniach v laboratórnych podmienkach. Dosiahnuté výsledky sú dokladované grafickými závislosťami, pričom vo väčšine prípadov ide o závislosti oneskorenia od rýchlosti najpomalšej linky medzi vysielaciu a prijímacou stanicou. Uvedené grafické závislosti súčasne názorne dokumentujú správanie sa QoS mechanizmov voči Triple-Play tokom, čím umožňujú ľahšie predvídať dôsledky plynúce z ich nasadenia v praxi. Presný odhad celkového ako aj frontového oneskorenia paketov dáva navyše reálny predpoklad správnej konfigurácie zvoleného mechanizmu riadenia zahľtenia, bez nutnosti jej zdĺhavého hľadania a odladovania v reálnych podmienkach.

Kritickou veličinou a to nie len v konvergovaných sieťach je dostupná kapacita prenosových liniek. Jej množstvo býva často obmedzené a bez nasadenia a správnej konfigurácie QoS mechanizmov je často nemožné zabezpečiť požadovanú kvalitu služieb citlivých na oneskorenie a straty ako VoIP a IPTV. Základné parametre zadávané pri nasadzovaní mechanizmov riadenia zahľtenia sú pre každú triedu prevádzky: garantovaná šírka pásma a kapacita paketového frontu. Poznatky nadobudnuté pri tvorbe a overovaní modelov mechanizmov FIFO a WRR sa následne využili na zostavenie a overenie matematického modelu, zameraného na nájdenie hodnôt minimálnej šírky pásma a vhodnej dĺžky frontu, ktoré sú potrebné pre zachovanie kvality príslušnej služby. Model bol laboratórne overený na mechanizme CBWFQ (ang. *Class Based Weighted Fair Queuing*), ktorý patrí v súčasnosti medzi často nasadzované mechanizmy riadenia zahľtenia.

Znalosti z tvorby a overovania zostavených modelov boli v závere zosumarizované do metodiky voľby parametrov mechanizmu CBWFQ v sieťach s Triple-Play prevádzkou.

Obsah

ŽILINSKÁ UNIVERZITA V ŽILINE	1
AUTOREFERÁT	1
DIZERTAČNEJ PRÁCE	1
ABSTRAKT	5
OBSAH	6
ÚVOD	1
1 MECHANIZMY RIADENIA ZAHLTENIA	1
2 CIELE PRÁCE	2
3 MODELOVANÉ MECHANIZMY	3
3.1 FIFO	3
3.2 Mechanizmus WRR.....	4
3.2.1 Dynamické prerozdelenie nevyužitej šírky pásma	5
4 ZOSTAVENÉ MODELY QOS MECHANIZMOV	5
4.1 Modely so plnými frontami.....	6
4.1.1 Mechanizmus FIFO	6
4.2 Mechanizmus WRR.....	8
4.2.1 Výsledky	9
4.3 Modely s premenlivými frontami	10
4.3.1 Mechanizmus FIFO	10
4.3.2 Mechanizmus WRR	12
5 VYUŽITIE FRONTOVÝCH MODELOV NA DIMENZOVANIE PARAMETROV QOS MECHANIZMOV	16
5.1 Vymedzenie modelu.....	17
5.2 Zápis modelu.....	17
5.3 Princíp riešenia modelu.....	18
5.4 Overenie modelu a výsledky	19
5.4.1 Oneskorenie	19
5.4.2 Straty	21
5.4.3 QoS Booster	22

5.4.4	CBWFQ s prioritným frontom	23
6	METODIKA KONFIGURÁCIE MECHANIZMU CBWFQ V SIETI S TRIPLE-PLAY TOKMI	23
6.1.1	VoIP toky	23
6.1.2	IPTV toky	24
6.1.3	Dátové toky	24
7	ZÁVER A ZHODNOTENIE VÝSLEDKOV.....	25
8	PRÍNOSY PRÁCE.....	29
	POUŽITÁ LITERATÚRA	30

Úvod

Súčasný trend vo vývoji komunikačných sietí je konvergencia existujúcich sietí do jednej univerzálnej siete, ktorá zabezpečuje poskytovanie rôznorodých služieb s garanciou požadovanej kvality. Medzi takéto siete môžeme zaradiť aj metropolitné siete pokrývajúce mestské časti alebo celé mestá, v ktorých poskytovatelia služieb poskytujú tzv. *Triple-Play*¹ služby. Tieto zahŕňajú dátové služby, služby prenosu hlasu a služby prenosu obrazu. Základom konvergovaných sietí sa stali počítačové paketové siete. Avšak ich pôvodnou úlohou bolo prenášanie dát, bez nejakého zabezpečenia kvality. Jediná garancia spočívala v prístupe do siete bez odmietnutia a v následnom spracovaní prevádzky v závislosti na aktuálne dostupných prostriedkoch v sieti (tzv. *best-effort* model) [1]. Sieť nerozlišovala medzi tokmi, ktoré prenášala, ku všetkým sa pri prenose správala rovnako bez ohľadu na jej zaťaženie. Šírka pásma, ktorú mal daný tok k dispozícii závisela od požiadaviek ostatných tokov a aktuálneho vytťaženia liniek, čo sa negatívne prejavilo predovšetkým na kvalite služieb prenosu hlasu a obrazu. Intenzita vysielania paketov hlasových a obrazových tokoch je pevne daná a neprispôsobuje sa dostupnej šírke pásma v sieti tak, ako je to v prípade dátovej prevádzky. Ak intenzita hlasovej alebo video prevádzky prevýši aktuálne dostupnú šírku pásma, pakety nad rámec sa nenávratne zahadzujú, čo má za následok degradáciu kvality príslušnej služby. Keďže hlasová a televízna prevádzka nemajú rovnaké požiadavky na sieť ako dátové prevádzka, bolo potrebné toky jednotlivých prevádzok od seba odlíšiť a poskytnúť im adekvátne spracovanie. Z tohto dôvodu bolo v dnešných IP sieťach navrhnutých a aj implementovaných veľké množstvo *mechanizmov garancie kvality služieb* (tzv. *QoS mechanizmov*), ktoré umožňujú rôzne spracovanie paketov rozličných prevádzok na ceste od odosielateľa k príjemcovi. Cieľom nasadzovania QoS mechanizmov v konvergovaných sieťach je poskytnúť pre hlasovú a video prevádzku prednostné spracovanie pred dátovou prevádzkou alebo im aspoň garantovať určitú šírku pásma, a tým im zabezpečiť akceptovateľné hodnoty sieťového oneskorenia, jeho kolísania alebo strát.

1 Mechanizmy riadenia zahltienia

Výsledná kvalita služby závisí predovšetkým na šírke pásma, ktorú má prevádzka pri prechode sieťou k dispozícii. Od dostupnej šírky pásma priamo závisia ďalšie faktory ovplyvňujúce QoS ako stratovosť paketov, oneskorenie a jeho kolísanie. Dôsledky plynúce z nedostatku dostupnej šírky pásma sú z pohľadu kvality služby citelne u všetkých aplikácií, no kritické sú predovšetkým u aplikácií reálneho času (VoIP, IPTV), kde nie je prípustné

¹Triple-Play je marketingové označenie balíka služieb, ktorý ponúkajú sieťoví poskytovatelia svojim zákazníkom. Tento balík zahŕňa dátové služby, telefóniu a televíziu, ktoré sú zákazníkovi poskytované spolu za zvýhodnenú cenu. Nejedná sa o štandardizovaný výraz, no kvôli jednoduchosti bude v dizertačnej práci používaný na súhrnné označenie dátovej, hlasovej a video prevádzky.

a ani možné (protokol UDP), aby sa intenzita vysielania paketov prispôbovala aktuálne dostupnej šírke pásma.

Významnou množinou QoS mechanizmov, ktoré sa podieľajú na zabezpečení kvality služieb v konvergovaných sieťach sú mechanizmy riadenia zahľtenia. Mechanizmy riadenia zahľtenia efektívne rozdeľujú kapacitu linky medzi viaceré triedy služieb. Pomocou mechanizmov riadenia zahľtenia sa dá docieľiť, že službám, ktoré sú na dostupnú šírku pásma obzvlášť citlivé bude vždy garantovaná taká šírka pásma, ktorú požadujú. Súčasťou týchto mechanizmov sú vyrovnávacie pamäte, slúžiace na dočasne uskladnenie paketov, ktoré pred vstupom na kritické linky buď prevyšujú im vyhradenú šírku pásma alebo musia byť nutne pozdržané v dôsledku nižšej priority. Pre pojem „vyrovnávacia pamäť“ v terminológii z oblasti kvality služieb zaužíval výraz *paketový front*, ktorý lepšie a názornejšie vystihuje účel jeho použitia. Aby bolo efektívne rozdelenie kapacity linky medzi jednotlivé triedy možné, každá trieda prevádzky má pridelený vlastný front. Samozrejme ak požiadavky hlasových alebo televíznych aplikácií krátkodobo prevýšia kapacitu linky alebo prednastavenú šírku pásma, budú do paketových frontov radené aj ich pakety. Výhodou takéhoto prístupu je výrazné zníženie potenciálnych strát paketov, ktoré by inak nasledovali, ak by došlý paket nemohol byť odvysielaný po jeho príchode na výstupnú linku, ktorá by v tom čase bola vyhradená pre spracovanie iného paketu. Každé pozdržanie paketu v paketovom fronte na druhej strane predlžuje dobu doručenia paketu cieľovej stanici, čo môže následne zhoršiť kvalitu príslušnej služby.

Oneskorenie, ktoré vzniká v dôsledku uloženia paketov do frontov QoS mechanizmov (skr. *frontové oneskorenie*) tvorí významný podiel na celkovom oneskorení pri prenose prevádzky sieťou.

2 Ciele práce

Pre vhodný návrh a nadimenzovanie siete je dôležité poznať alebo aspoň predvídať hodnotu oneskorenia príslušnej triedy prevádzky, keďže sa jedná o významný faktor ovplyvňujúci výslednú kvalitu služby. Presný odhad celkového ako aj frontového oneskorenia paketov dáva reálny predpoklad správnej konfigurácie zvoleného mechanizmu riadenia zahľtenia, bez nutnosti jej zdĺhavého hľadania a odladovania v reálnych podmienkach. Okrem praktických experimentov v sieti, meraním a vyhodnocovaním je možné jednoduchšie určiť alebo aspoň odhadnúť hodnotu frontového ako aj celkového oneskorenia modelovaním. Z pohľadu modelovacích techník sa dá na tento účel využiť základný spôsob modelovania dynamických systémov - *simulácia*. Medzi sofistikovanejšie a rýchlejšie metódy patrí napr. *matematické modelovanie*. Ich použitie v praxi však nie je moc zaužívané. Keďže v dnešných IP sieťach sa pohybujú oneskorenia v hodnotách blízkych kritickej degradácii kvality služby, je vhodné venovať týmto metódam náležitú pozornosť a výskum.

Prvým cieľom DP bolo vypracovať analytické a simulačné modely mechanizmov riadenia zahľtenia používaných v architektúre diferencovaných služieb. Tvorba modelov bola zameraná na vyjadrenie frontového oneskorenia, ktoré tvorí významnú časť celkového oneskorenia siete.

Druhým cieľom DP bolo zhodnotiť použiteľnosť vytvorených modelov na odhad oneskorenia v reálnej IP sieti s Triple-Play tokmi. Porovnaním výsledkov modelov s hodnotami nameranými v realite sa súčasne overila správnosť zostavených modelov.

Kritickou veličinou a to nie len v konvergovaných sieťach je dostupná kapacita prenosových liniek. Jej množstvo je býva často obmedzené a bez nasadenia a správnej konfigurácie QoS mechanizmov je často nemožné zabezpečiť požadovanú kvalitu služieb citlivých na oneskorenie a straty ako VoIP a IPTV. Základné parametre zadávané pri nasadzovaní mechanizmov riadenia zahltienia sú pre každú triedu prevádzky garantovaná šírka pásma a dĺžka paketového frontu. Tretím cieľom DP bolo využiť poznatky nadobudnuté pri tvorbe modelov na zostavenie a overenie matematického modelu, zameraného na nájdenie hodnôt minimálnej šírky pásma a vhodnej dĺžky frontu, ktoré sú potrebné pre zachovanie kvality príslušnej služby.

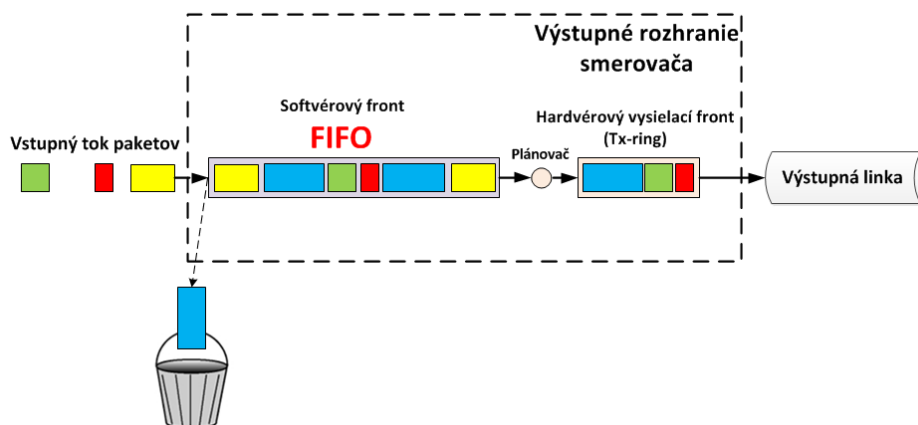
Štvrtým cieľom práce bolo zosumarizovať výsledky a poznatky z tvorby a overovania zostavených modelov do metodiky voľby parametrov mechanizmov predchádzania zahltieniu v IP sieti s Triple-Play službami. Súčasťou uvedenej metodiky bol aj posledne spomínaný model.

3 Modelované mechanizmy

Za modelované QoS mechanizmy v rámci dizertačnej práce boli zvolené dva mechanizmy. Najjednoduchší mechanizmus riadenia zahltienia FIFO a pokročilý mechanizmus WRR.

3.1 FIFO

V mechanizme FIFO zdieľajú pakety všetkých triedy prevádzky jeden front, do ktorého sú došlé pakety radené v poradí príchodu a v rovnakom poradí sú vysielané na výstupnú linku. V dôsledku toho mechanizmus FIFO negarantuje žiadnej z tried dostupnú šírku pásma a jeho nasadenie v konvergovaných sieťach sa z tohto dôvodu nedoporučuje.



Obr. 3.1 Mechanizmus FIFO

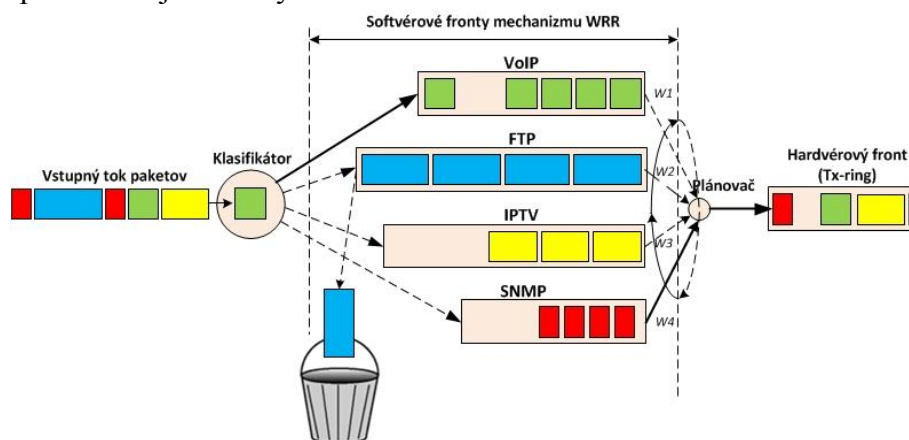
Na mechanizme FIFO bola predovšetkým overená skutočnosť, či sú modely známe z teórie hromadnej obsluhy vôbec použiteľné na odhad oneskorenia pre IPTV a VoIP toky. Ako by modely mechanizmu FIFO nedávali uspokojivé výsledky, ťažko možno očakávať, že dodatočne vytvorené modely podstatne zložitejšieho mechanizmu WRR uspokojivé výsledky dávajú budú.

Známe analytické modely mechanizmu FIFO predpokladajú predovšetkým Poissonove toky paketov, no reálne IPTV a VoIP toky majú iný charakter. Na mechanizme FIFO sa tak súčasne overil vplyv rôznych tokov na vývoj frontového oneskorenia, ktoré tvorí významnú zložku celkového oneskorenia v sieti.

Počas veľkého zaťaženia siete nedochádza k dynamickému prerozdeleniu nevyužitej šírky pásma medzi frontami zložitejších mechanizmov riadenia zahltienia. Dá sa preto očakávať, že na modelovanie jednotlivých frontov bude možné využiť v tomto prípade aj model mechanizmu FIFO.

3.2 Mechanizmus WRR

Mechanizmus váženého výberu frontov (ang. *Weighted Round Robin*, skr. *WRR*) [6][36] zabezpečuje efektívne rozdelenie šírky pásma prevádzky medzi viaceré triedy služieb. Toto dosahuje tak, že rozdeľuje prevádzku podľa prevádzkových tried do viacerých FIFO frontov. Zaradeniu paketu do príslušného frontu nutne predchádza klasifikácia paketu. S činnosťou mechanizmu vážených výberov z frontov je spojený tzv. *obslužný cyklus*, v rámci ktorého striedavo obsluhuje všetky neprázdne fronty. Obslužný cyklus tak predstavuje sériu obslúh neprázdnych frontov, ktorá sa počas činnosti mechanizmu permanentne opakuje. Maximálny počet paketov, ktoré mechanizmus vyberie v rámci obsluhy jedného frontu je dané váhou pridelenou príslušnému frontu. Váhy frontov tak určujú rozdelenie výstupnej šírky pásma linky medzi prevádzku jednotlivých tried.



Obr. 3.2 Činnosť mechanizmu WRR

V reálnych implementáciách často nepredstavujú počty paketov, ale priemerný počty bajtov, ktoré sa z frontu môžu odobrať v rámci jedného obslužného cyklu. Na základe priemerného počtu bajtov, nie je problém odhadnúť presné rozdelenie šírky pásma medzi jednotlivé triedy služieb.

Z mechanizmu WRR vychádzajú ďalšie mechanizmy riadenia zahltenia, ktoré sa podieľajú na efektívnom delení šírky pásma medzi rôzne triedy prevádzky (CBWFQ, DWRR, LLQ...).

3.2.1 Dynamické prerozdelenie nevyužitej šírky pásma

Šírka pásma vyhradená jednotlivým frontom v mechanizmoch riadenia zahltenia predstavuje pre ne garantované minimum, ktoré majú k dispozícii za akýchkoľvek podmienok. Mechanizmus prázdne fronty neobsluhuje. Dôsledkom toho je, že ak sú jeden alebo viacero z frontov mechanizmu WRR prázdne, neprázdne fronty si proporcionálne rozdelia šírku pásma vyhradenú pre tieto front. V dvojfrontovej konfigurácii to znamená asi toľko, že ak je jeden front prázdny a druhý front neprázdny, neprázdny front je počas celej doby trvania tohto stavu oprávnený využívať kompletnú kapacitu výstupnej linky.

4 Zostavené modely QoS mechanizmov

V rámci dizertačnej práce boli zostavené dva typy modelov.

- Modely matematické
- Modely simulačné

Matematické modely boli riešené analytickými alebo numerickými metódami. Simulačné modely boli zostavené buď vo vlastnom simulačnom nástroji *QoS Tester* alebo komerčnom simulačnom nástroji *Opnet Modeler*.

Zostavené modely QoS mechanizmov je možné rozdeliť do dvoch kategórií.

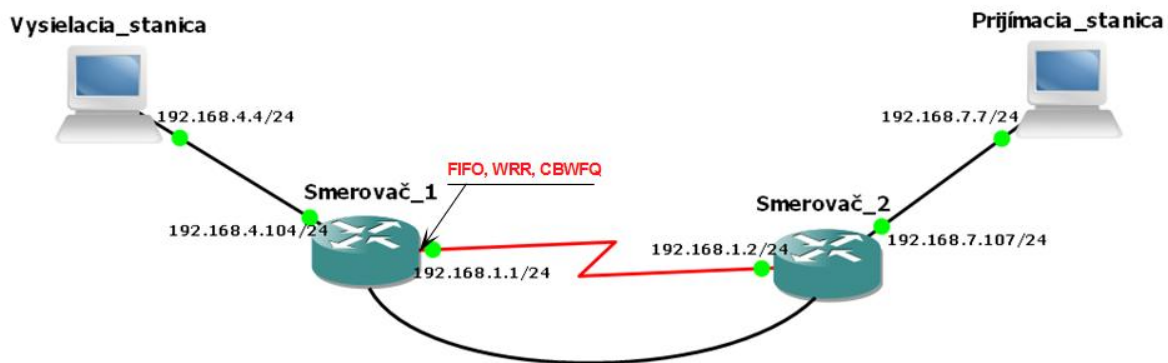
- Modely predpokladajúce plné fronty
- Modely predpokladajúce premenlivé fronty

Modelované mechanizmy	Modely	Fronty	Toky
FIFO	Matematické	Plné	$c < \omega$
		Premenlivé	Poissonove
	Simulačné	Plné	$c < \omega$
		Premenlivé	Totožné s reálnymi
WRR	Matematické	Plné	$c < \omega$
		Premenlivé	Poissonove
	Simulačné	Plné	$c < \omega$
		Premenlivé	Totožné s reálnymi

Obr. 4.1 Zostavené modely v rámci dizertačnej práce

Správnosť a presnosť uvedených modelov bola overená praktickými experimentmi uskutočnenými v laboratórnych podmienkach. Testovacia zostava obsahovala linkou nízkej priepustnosti, na vstupe ktorej vznikalo počas generovania zahltenie. Zahltenie linky bolo obsluhované príslušným mechanizmom riadenia zahltenia. Pri overovaní modelov sa využili VoIP a IPTV toky, ktoré boli emulované prostredníctvom softvérového generátora paketov.

Kritériom hodnotenia modelov bola presnosť odhadu celkového oneskorenia v testovacej topológii, ktoré je kľúčovým parametrom pri posudzovaní kvality služieb. V modeloch ako aj v laboratórnych experimentoch bola kapacita vysielačieho frontu (tx-ring) vždy jeden paket.



Obr. 4.2 Testovacia topológia

4.1 Modely so plnými frontami

Najčastejšou príčinou permanentne plných softvérových frontov mechanizmov riadenia zahŕňajúce je nesprávne nadimenzovaná sieť, ktorá v dôsledku nízkej kapacity v čase maximálnej záťaže nedokáže dlhodobo preniesť všetku prevádzku. Napriek tomu, že k permanentne plným frontom dochádza v správne nadimenzovaných sieťach len výnimočne, znalosť oneskorenia pre tento výnimočný stav dáva užívateľovi možnosť predvídať najväčšie oneskorenie, ktoré môže pre zvolenú konfiguráciu v sieti nastať. Oneskorenie odhadnuté pre zahŕňajúce fronty tvorí tak po prirátaní fixného oneskorenia siete a serializačného oneskorenia kritickej linky tvoria hornú hranicu očakávaného oneskorenia v sieti.1

4.1.1 Mechanizmus FIFO

Z teoretického hľadiska pre vznik a udržanie permanentne plného frontu musí platiť :

$$c \ll \omega, \quad (4.1)$$

kde c je šírka pásma linky a ω je intenzita vstupného toku paketov [*bit/s*].

Na to, aby front ostal stále plný, musí okamžite potom ako prvý paket opustí front, prísť na koniec frontu nový paket. Keďže prakticky sa takýto scenár nedá dosiahnuť, budeme ďalej za permanentne plný front považovať front, ktoré aktuálna dĺžka je rovná alebo o jeden paket menšia než je kapacita tohto frontu.

V laboratórnych podmienkach nie je zložité dosiahnuť permanentne plný FIFO front. Stačí, aby vstupný tok ako aj veľkosti paketov boli deterministické a súčasne medzery medzi paketmi vstupného toku boli väčšie než je doba vysielačieho frontu z tx-ringy na výstupnú linku.

Z laboratórnych experimentov vyplynulo, že pri radení paketov do plného FIFO frontu vzniká v uvažovaných sieťových zariadeniach dodatočné oneskorenie, ktoré v absolútnom vyjadrení

závisí od rýchlosti linky, no nezávisí na dĺžke frontu. Dodatočné oneskorenie v bajtovom vyjadrení *bajtové oneskorenie* (ψ) je rovné pre hodnotu 505 B, čo absolútnom vyjadrení zodpovedajú časovému intervalu

$$\frac{8 \cdot 505}{c} \text{ [s]}. \quad (4.2)$$

Nech L je dĺžka FIFO frontu, λ intenzita paketov toku [pak/s], c je kapacita linky a l je stredná veľkosť paketov [B]. Pre spresnený odhad stredného oneskorenia v zahľtenom FIFO fronte bol odvodený vzťah:

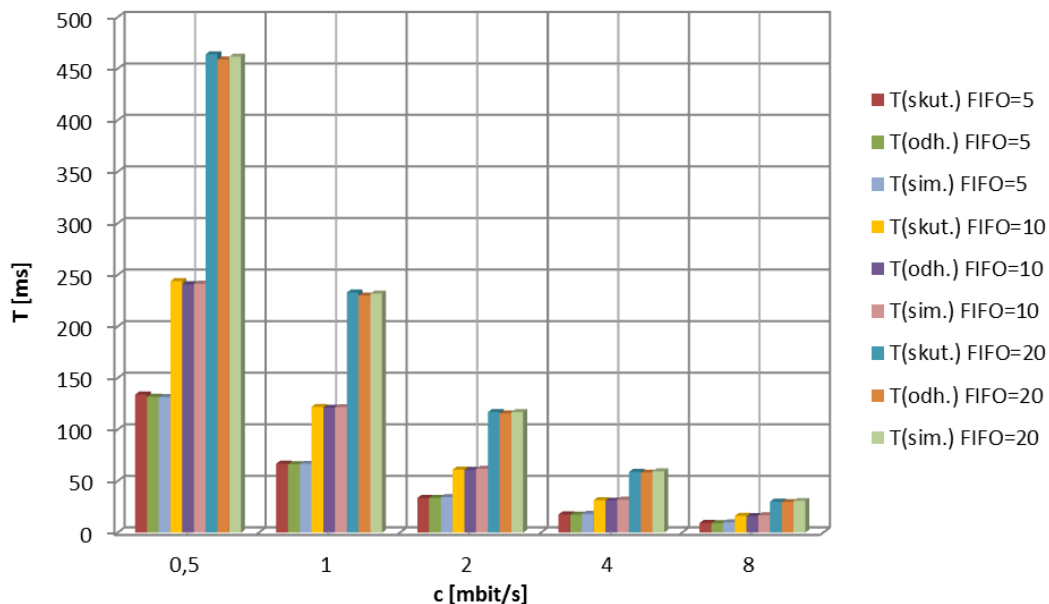
$$T = \frac{8lL}{c} + \frac{8l}{c} + \frac{8\psi}{c} + T_{fix}. \quad (4.3)$$

Prvý člen výrazu (4.3) predstavuje frontové oneskorenie. Druhý člen výrazu (4.3) predstavuje serializačné oneskorenie, ktoré vzniká pri vysielaní paketu z hardvérového frontu na výstupnú linku. Tretí člen výrazu (4.3) predstavuje dodatočné oneskorenie pri spracovaní paketu a štvrtý člen fixné oneskorenie siete.

4.1.1.1 Výsledky

Na obrázku Obr. 4.3 sú porovnané hodnoty oneskorenia nameraného v laboratórnych podmienkach, odhadnutého prostredníctvom vzťahu (4.3) a odhadnutého prostredníctvom simulácie v nástroji Opnet Modeler pre rôzne kapacity linky a dĺžky frontov

Priemerná relatívna chyba analytického odhadu všetkých experimentoch predstavovala len 3,5 %.



Obr. 4.3 Porovnanie nameraných a odhadovaných oneskorení IPTV paketov po zohľadnení priemerného bajtového oneskorenia

4.2 Mechanizmus WRR

Podmienka permanentne plných frontov mechanizmu WRR je podobná ak v prípade mechanizmu FIFO. Pre každý z frontov mechanizmu WRR musí platiť, že intenzita vstupného dátového toku je oveľa väčšia než garantovaná šírka pásma prevádzky daného frontu.

$$\omega_i \gg c_i. \quad (4.4)$$

Ak pripúšťame, že aktuálna dĺžka v permanentne plnom fronte môže byť aj o jeden paket menšia ako je jeho kapacita (stav medzi odchodom paketu z frontu a príchodom ďalšieho paketu do frontu), potom stačí, aby príchody boli deterministické a platila podmienka:

$$\omega_i > c_i. \quad (4.5)$$

Pre oneskorenie paketov j -tej triedy v n -frontovej konfigurácii mechanizmu WRR, ktorý je súčasťou zariadení CISCO bol odvodený vzťah:

$$E[T_j] = \frac{\delta l_j L_j \sum_{i=1}^n w_i}{c w_j} + \frac{\delta \psi}{c} + \frac{\delta l_j}{c} + T_{fix}^j. \quad (4.6)$$

w_j je priemerný počet bajtov vybraných z frontu j -tej triedy v rámci jedného obslužného cyklu mechanizmu WRR, L_j je dĺžka j -tého frontu, l_j je veľkosť paketu j -tého frontu a T_{fix}^j je fixné oneskorenie paketov j -tej triedy v danej sieti.

Pre dvojfrontovú konfiguráciu mechanizmu WRR v zariadeniach CISCO s IPTV a VoIP prevádzkou zo vzťahu (4.6) platí:

$$E[T_{iptv}] = \frac{\delta l_{iptv} L_{iptv} (w_{iptv} + w_{voip})}{c w_{iptv}} + \frac{\delta l_{iptv}}{c} + \frac{\delta \psi}{c} + T_{fix}^{iptv} \quad (4.7)$$

$$E[T_{voip}] = \frac{\delta l_{voip} L_{voip} (w_{iptv} + w_{voip})}{c w_{voip}} + \frac{\delta l_{voip}}{c} + \frac{\delta \psi}{c} + T_{fix}^{voip} \quad (4.8)$$

V kapitole 4.2.1 sa ukázalo, že vážené výbery paketov z frontov, ktoré prevádza mechanizmus WRR je možné úspešne modelovať pomocou pravdepodobností výberov z frontov. Zavedenie pravdepodobností výberov paketov z frontov sa využilo predovšetkým v analytických modeloch mechanizmu WRR s premenlivými dĺžkami frontov.

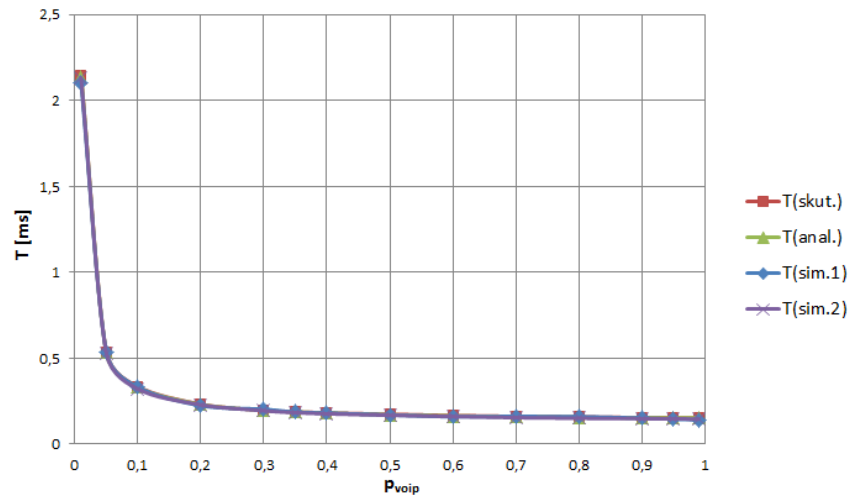
Pre dvojfrontovú konfiguráciu s pravdepodobnostnými výbermi boli odvodené vzťahy:

$$E[T_{iptv}] = \left(\frac{\delta l_{voip}}{c} \frac{P_{voip}}{P_{iptv}} + \frac{\delta l_{iptv}}{c} - \frac{\delta l_{voip}}{c} \right) L_{iptv} + \frac{\delta l_{iptv}}{c} + \frac{\delta \psi}{c} + T_{fix}^{iptv}, \quad (4.9)$$

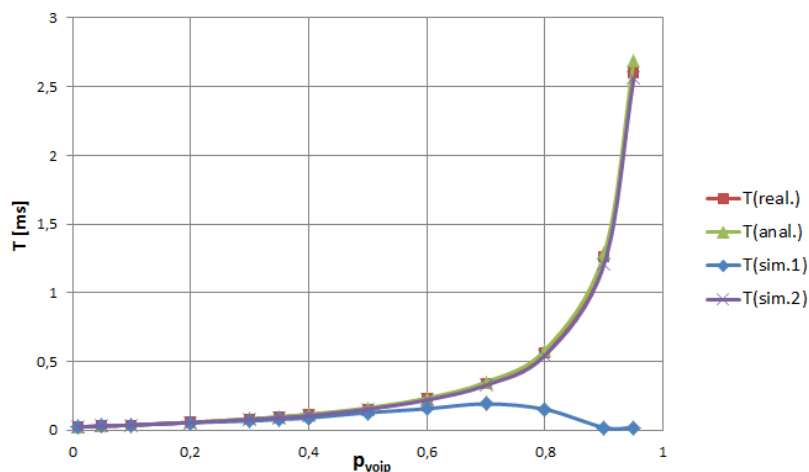
$$E[T_{voip}] = \left(\frac{\delta l_{iptv}}{c} \frac{P_{iptv}}{P_{voip}} + \frac{\delta l_{voip}}{c} - \frac{\delta l_{iptv}}{c} \right) L_{voip} + \frac{\delta l_{voip}}{c} + \frac{\delta \psi}{c} + T_{fix}^{voip}. \quad (4.10)$$

4.2.1 Výsledky

Na obrázkoch Obr. 4.4, Obr. 4.5 sú uvedené závislosti oneskorenia IPTV a VoIP paketov od zmeny pomeru váh pridelených frontom, ktorý je v grafe vyjadrený pomocou pravdepodobnosti výberu paketu z VoIP frontu. Uvedené závislosti boli dosiahnuté reálnym meraním ($T(skut.)$), analytickým výpočtom pomocou vzťahov (4.7) a (4.8) ($T(anal.)$), simuláciou v nástroji *Opnet[®] Modeler* ($T(sim.)$) a simuláciou v nástroji *QoS Tester* pri použití pravdepodobnostných výberov z frontov ($T(sim.2)$).



Obr. 4.4 Závislosť oneskorenia VoIP paketov od zmeny pravdepodobnosti výberu paketov z frontov v dvojfrontovom mechanizme WRR



Obr. 4.5 Závislosť oneskorenia IPTV paketov od zmeny pravdepodobnosti výberu paketov z frontov v dvojfrontovom mechanizme WRR

Podľa predpokladov analytický odhad oneskorenia verne kopíruje výsledky meraní pre všetky hodnoty pravdepodobnosti výberu tak pre IPTV ako aj VoIP prevádzku. Priemerná relatívna chyba analytického odhadu sa pohybuje pod hranicou 0,1 %, a tak ho možno pokladať za veľmi presný.

4.3 Modely s premenlivými frontami

V prípade zahŕňaných softvérových frontov je možné oneskorenie odhadnúť relatívne presne. V praxi sa však vyskytujú predovšetkým prípady premenlivých paketových frontov, kedy je presný odhad oneskorenia podstatne komplikovanejší. Problémom pri vyjadrovaní alebo odhade frontového oneskorenia je v prípade premenlivých frontov skutočnosť, že dĺžka paketového frontu, s ktorou veľkosť frontového oneskorenia úzko súvisí, nie je konštantná veličina, ale sa s časom náhodne a rýchlo mení. Hlavným faktorom, ktorý výrazne ovplyvňuje stav frontu v čase a komplikuje tak odhad oneskorenia, je charakter vstupného toku resp. variabilita medzier vstupného toku. Vyjadrením charakteru toku resp. jeho variability je napríklad variačný koeficient alebo Hurstov parameter. Zohľadnenie uvedených veličín je v analytických modeloch za účelom odhadu oneskorenia veľmi problematické. Vo väčšine prípadov sa medzery vo vstupnom toku paketov modelujú bezpamäťovým exponenciálnym rozdelením. Tok s exponenciálnymi príchodmi paketov je označovaný ako tok Poissonov. Všetky zostavené analytické modely budú preto predpokladať Poissonove vstupné toky paketov.

4.3.1 Mechanizmus FIFO

Pri odhade oneskorenia v mechanizme FIFO bol využitý známy frontový model v THO označovaný aj ako $M/D/1/\infty$.

Pre odhad frontové oneskorenia bol odvodený vzťah:

$$E[W] = \frac{32\lambda^2}{c(1-\varepsilon)(c-8\lambda)}, \quad (4.11)$$

kde L je veľkosť paketu a λ je intenzita vstupného toku paketov.

Laboratórnymi experimentmi sa potvrdilo, že pri spracovaní paketov mechanizmom FIFO vzniká vo fronte dodatočné oneskorenie, ktoré bolo vyjadrené prostredníctvom relatívnej chyby odhadu frontového oneskorenia ε . Hodnota relatívnej chyby nebola závislá na hodnote kapacity linky. Pre VoIP pakety bola táto hodnota $\varepsilon_{voip} = 0,07$ a pre IPTV pakety $\varepsilon_{iptv} = 0,16$.

Pre frontové oneskorenie v mechanizme FIFO pre prostredie s viacerými typmi tokov bol odvodený vzťah:

$$E(W) = \frac{\sum_{k=1}^m \frac{\lambda_k \tau_k^2}{1-\varepsilon_k}}{2(1-\rho)} = \frac{32 \sum_{k=1}^m \frac{l_k^2 \lambda_k}{1-\varepsilon_k}}{c^2(1-\rho)}, \quad (4.12)$$

kde ε_k je relatívna chyba odhadu frontového oneskorenia vykonaného prostredníctvom štandardného modelu $M/D/1/\infty$ s jedným tokom.

Keďže v tomto prípade zdieľajú všetky toky paketov spoločný front, frontové oneskorenie všetkých tokov je rovnaké.

Po zohľadnení serializačného a fixného oneskorenia paketov j -tého toku paketu dostávame spresnený odhad celkového oneskorenia v sieti pre j -ty tok:

$$E(T_j) = \frac{32 \sum_{k=1}^m \frac{l_k^2 \lambda_k}{1 - \varepsilon_k}}{c^2 (1 - \rho)} + \frac{8l_j}{c} + T_{fix}^j. \quad (4.13)$$

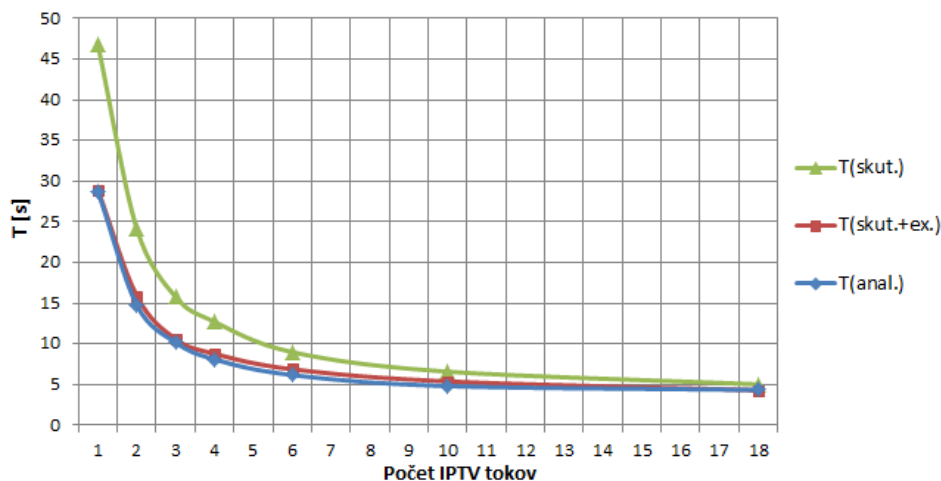
Podľa teoretických predpokladov platí, že po zlúčení n tokov s ľubovoľným rozdelením pravdepodobnosti medzipaketových medzier so strednou hodnotou τ , konverguje výsledný tok s rastúcim n k toku s exponenciálnymi medziami so strednou hodnotou τ/n . Keďže analytické modely predpokladajú práve toky s exponenciálnymi príchodmi, zaujímavé bude zistiť, pre aké veľké n možno pokladať odhad oneskorenia prostredníctvom týchto modelov v prípade IPTV a VoIP tokov za presný (chyba menšia ako 5 %).

4.3.1.1 Výsledky

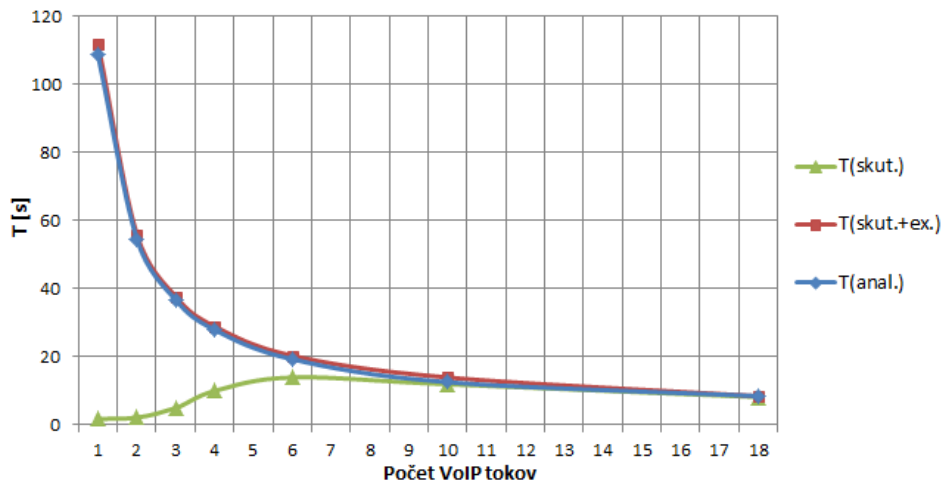
Na obrázkoch Obr. 4.6 a Obr. 4.7 sú zachytené závislosti oneskorenia od počtu zlúčených IPTV a VoIP tokov pri zaťažení linky $\rho=0,9$. Jednotlivé závislosti boli dosiahnuté odhadom prostredníctvom vzťahu (4.13) ($T(anal.)$) a meraním pri použití tak autentických ($T(skut.)$) ako aj exponenciálnych medzier toku ($T(skut.+ex.)$).

Podľa očakávaní odhady prostredníctvom vzťahu (4.13), ktorý predpokladal exponenciálne medzery, korešpondujú s výsledkami meraní, ktoré boli prevedené s IPTV a VoIP tokmi s exponenciálnymi medziami.

Takisto sa potvrdil klesajúci trend rozdielu oneskorenia medzi emulovanými tokmi s exponenciálnym a autentickými medziami paketov. Rozdiel je hlavne v spôsobe konvergenzie tokov jednotlivých tried. IPTV tok je zhlukovejší ako tok Poissonov, a tak sa oneskorenie kumulovaných IPTV tokov blíži k oneskoreniu Poissonových tokov zdola. Samotný deterministický tok nemá tendenciu generovať žiadne frontové oneskorenie ($\rho < 1$), a tak sa oneskorenie kumulovaných VoIP tokov blíži k oneskoreniu Poissonových tokov zdola.



Obr. 4.6 Porovnanie nameraných a odhadnutých oneskorení paketov v závislosti od počtu IPTV tokov pri zachovanom zaťažení linky $\rho=0,9$



Obr. 4.7 Porovnanie nameraných a odhadnutých oneskorení paketov v závislosti od počtu VoIP tokov pri zachovanom zaťažení linky $\rho=0,9$

Kým pre jeden IPTV tok relatívny rozdiel v oneskoreniach predstavuje 34 %, pre 10 tokov je to už len 5,4 % a pre 18 tokov 3,2 %. Z určitosťou tak možno tvrdiť, že na dosiahnutie menšej chyby ako 5 % je nutné pri zaťažení $\rho_{iptv}=0,9$ zlúčiť viac ako 10 tokov.

Kým pre šesť VoIP tokov relatívny rozdiel v oneskoreniach predstavuje 38 %, pre 10 tokov je to už len 5,1 % a pre 18 tokov 3,4 %. Z určitosťou tak možno tvrdiť, že na dosiahnutie menšej chyby ako 5 % je nutné pri zaťažení $\rho_{voip}=0,9$ zlúčiť viac ako 10 tokov.

4.3.2 Mechanizmus WRR

Najväčšie úsilie v rámci dizertačnej práce bolo venované modelovaniu mechanizmu WRR, ktorý efektívne rozdeľuje šírku pásma linky medzi jednotlivé triedy služieb. Okrem problému s premenlivým frontom je ďalším významným problémom pri matematickom modelovaní mechanizmu WRR skutočnosť, že šírka pásma vyhradená prevádzkam jednotlivých tried nie je konštantná veličina, ale sa časom náhodne mení (kap. 3.2.1).

V dizertačnej práci bolo zostavených niekoľko modelov, ktoré dynamické prerozdelenie nevyužitej šírky pásma snažili rôznymi spôsobmi zohľadniť a tým prispieť k vylepšeniu odhadu frontového oneskorenia paketov v mechanizme WRR.

Medzi simulačné modely a matematické modely s numerickým riešením patrili:

- simulačný model predpokladajúci pravdepodobnostné výbery z frontov, exponenciálne príchody paketov a exponenciálne veľkosti paketov,
- simulačný model predpokladajúci pravdepodobnostné výbery z frontov a rovnaké toky ako toky v realite,
- simulačný model predpokladajúci vážené výbery a rovnaké toky ako toky v realite,
- Markovov model vychádzajúci z numerického riešenia Kolmogorovových diferenciálnych rovníc,

- zjednodušený Markovov model vychádzajúci z numerického riešenia Kolmogorovových, diferenciálnych rovníc.

Medzi matematické modely s analytickým riešením patrili:

- Analytický model vychádzajúci z Markovových reťazcov s významným zjednodušením
- Analytický model vychádzajúci z Bernoulliho systému
- Alternatívny analytický model mechanizmu WRR
- Model mechanizmu WRR zanedbávajúci možnosť prázdnych frontov

4.3.2.1 Pravdepodobnostné výbery paketov z frontov

Pre účely matematického modelovania bol obslužný cyklus mechanizmu WRR spolu s váženými výbermi paketov nahradený tzv. pravdepodobnostnými výbermi z frontov. Jednotlivým frontom tak boli pridelené pravdepodobnosti, ktoré vyjadrujú šancu výberu paketu z daného frontu. Ich hodnoty boli úmerné počtu paketov, ktoré vyberie mechanizmus WRR v rámci jedného obslužného cyklu. Z pravdepodobností výberov paketov z frontov však nie je problém pri zohľadnení veľkostí paketov jednotlivých frontov dopočítať odpovedajúce váhy vyjadrujúce priemerný počet dát vybraných z frontu pri jeho obsluhu.

Implementácia mechanizmu WRR spoločnosti CISCO tzv. *Custom Queuing* (CQ) zahŕňa cyklickú obsluhu neprázdnych frontov, pričom priemerné množstvo dát vybraných v rámci jednej obsluhy frontu je pri konfigurácii zadávané v bajtoch. Analytické modely mechanizmu WRR predpokladajú pravdepodobnostný výber paketov z frontov. Pri návrhu adekvátnej konfigurácie mechanizmu CQ v testovacej zostave k zvoleným pravdepodobnostiam matematických modelov, ako aj pri porovnávaní výsledkov zostavených modelov s výsledkami laboratórnych experimentov je nutné urobiť prepočet pravdepodobností výberov na zodpovedajúce váhy frontov, ktoré sú uvádzané bajtoch.

Ďalej budeme predpokladať model mechanizmu CQ s dvoma frontami. Nech w_i bude množstvo dát vybraných pri jednej obsluhu i -tého frontu udávané v bajtoch. Nech l_i bude veľkosť paketu i -tého frontu (nie rámca). Nech N_i bude priemerné množstvo paketov vybraných pri obsluhu i -tého frontu.

Pre pravdepodobnosť výberu paketu z prvého resp. druhého frontu bude platiť:

$$p_1 = \frac{E(N_1)}{E(N_1) + E(N_2)} = \frac{w_1/l_1}{w_1/l_1 + w_2/l_2} = \frac{w_1 l_2}{w_1 l_2 + w_2 l_1} \quad (4.14)$$

$$p_2 = \frac{E(N_2)}{E(N_1) + E(N_2)} = \frac{w_2/l_2}{w_1/l_1 + w_2/l_2} = \frac{w_2 l_1}{w_1 l_2 + w_2 l_1} \quad (4.15)$$

Pre váhy w_1 resp. w_2 zo vzťahov (4.14), (4.15) dostávame:

$$w_1 = \frac{p_1 w_2 l_1}{(1 - p_1) l_2} \quad (4.16)$$

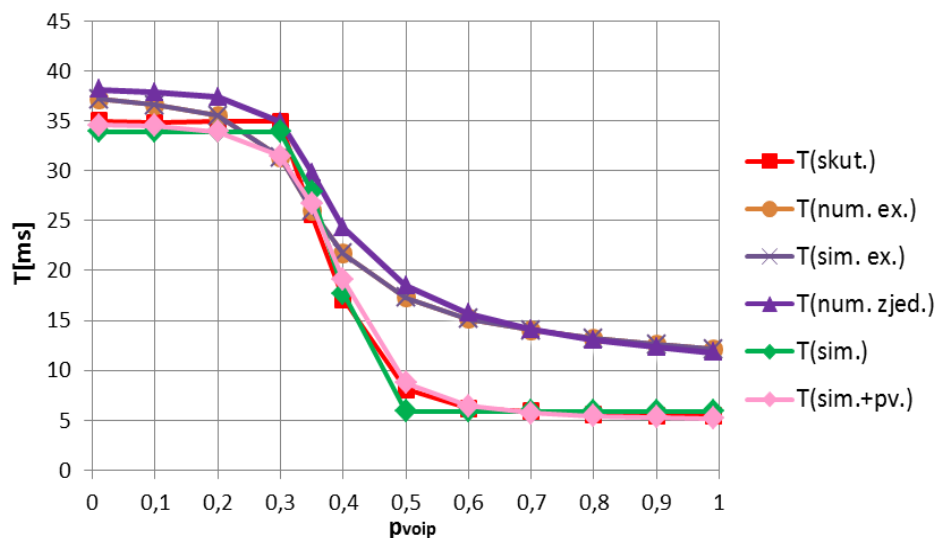
$$w_2 = \frac{p_2 w_1 l_2}{(1 - p_2) l_1} \quad (4.17)$$

4.3.2.2 Výsledky

Výsledky uvažovaných modelov skorigovaných rovnakým spôsobom ako v prípade mechanizmu FIFO sú zobrazené na obrázkoch Obr. 4.8 a Obr. 4.9. Zaťaženie linky bolo vo všetkých scenároch $\rho=0,9$.

Výsledky získané simulačným modelom predpokladajúcim exponenciálne príchody a veľkosti paketov presne kopírujú priebeh pôvodného numerického modelu, čím modely vzájomne potvrdzujú správnosť ich implementácie. Zjednodušený numerický model je pre IPTV pakety len mierne nadhodnotený, zatiaľ čo pre VoIP pakety mierne podhodnotený. Relatívna chyba však s rastom pravdepodobnosti výberu paketu z frontu klesá. V celku tak zjednodušený model možno pokladať za relatívne presný.

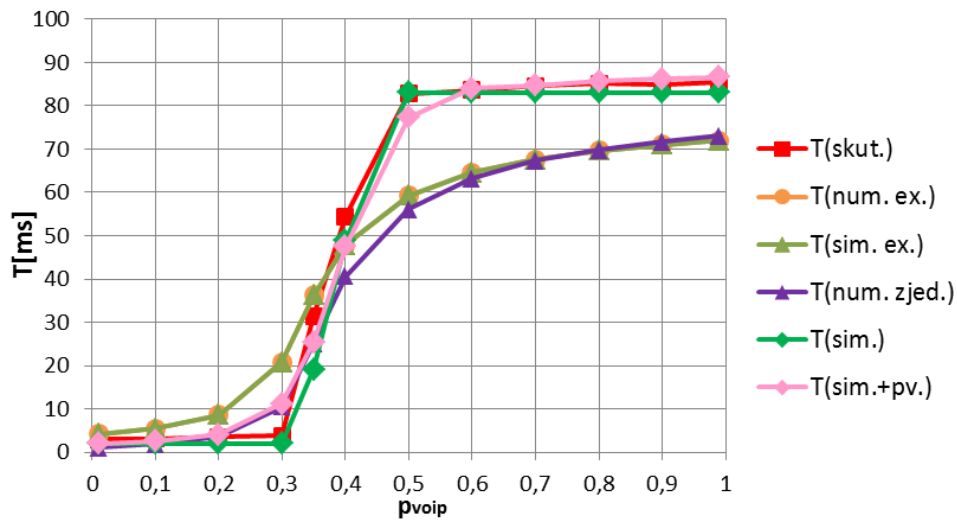
Výsledky modelov s exponenciálnymi príchodmi paketov, exponenciálnymi veľkosťami paketov a pravdepodobnostnými výbermi majú podľa očakávania plynulý priebeh pre všetky hodnoty pravdepodobnosti výberu paketov. V porovnaní s výsledkami meraní a modelmi so zachovaným charakterom tokov sa ich výsledky líšia a to predovšetkým pre $p_I > 0,4$. Kým pre IPTV pakety dávajú modely s Poissonovými tokmi na intervale $p_I \in (0,4; 1)$ podhodnotené výsledky, pre VoIP pakety sú ich výsledky výrazne nadhodnotené. Takýto vývoj je spôsobený predovšetkým rôznym charakterom reálnych a modelovaných tokov.



Obr. 4.8 Závislosti oneskorenia od zmeny pravdepodobnosti výberov pre IPTV pakety

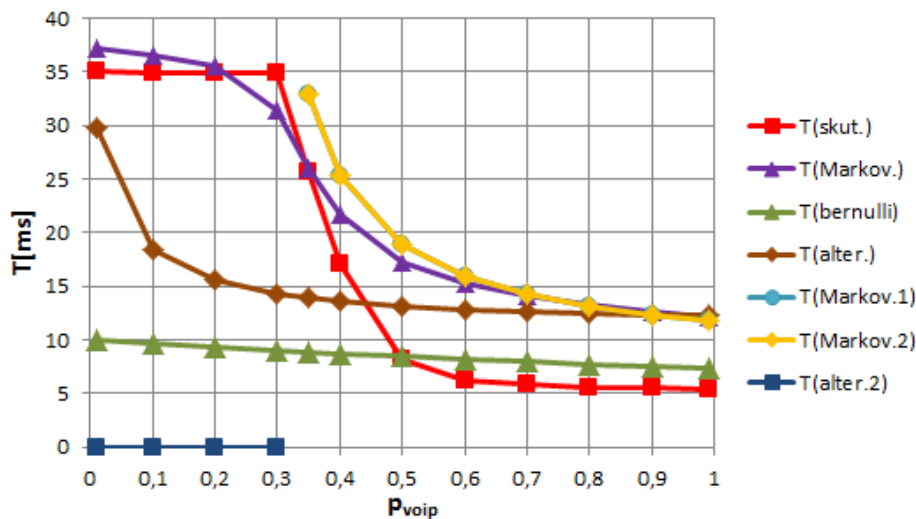
Po aplikácii korekčnej hodnoty na výsledky simulácie dostávame aj v prípade mechanizmu WRR relatívne presný odhad reálneho oneskorenia tak pre VoIP ako aj IPTV pakety. Uvedený fakt dokazuje univerzálnosť korekčnej hodnoty dodatočného oneskorenia. Obdobne sa správa aj simulačný model s pravdepodobnostnými výbermi a tokmi rovnakými ako toky v realite. Rozdiel medzi výsledkami spomínaných modelov však spočíva v plynulosti priebehu oneskorenia. K zmenám oneskorenia pri pravdepodobnostných výberoch dochádza podľa predpokladov na celom intervale $p_I \in (0; 1)$ plynule. Priebeh oneskorenia získaný prostredníctvom simulácie, ktorá predpokladá vážené výbery, sa podobne ako priebeh nameraného oneskorenia láme pre váhy zodpovedajúce pravdepodobnostiam 0,3 a 0,5.

Súčasne sú hodnoty oneskorenia na intervaloch $p_1 \in (0;0,3)$ a $p_1 \in (0,5;1)$ tak v simulačnom modeli ako aj v realite takmer konštantné tak pre IPTV ako aj VoIP pakety.

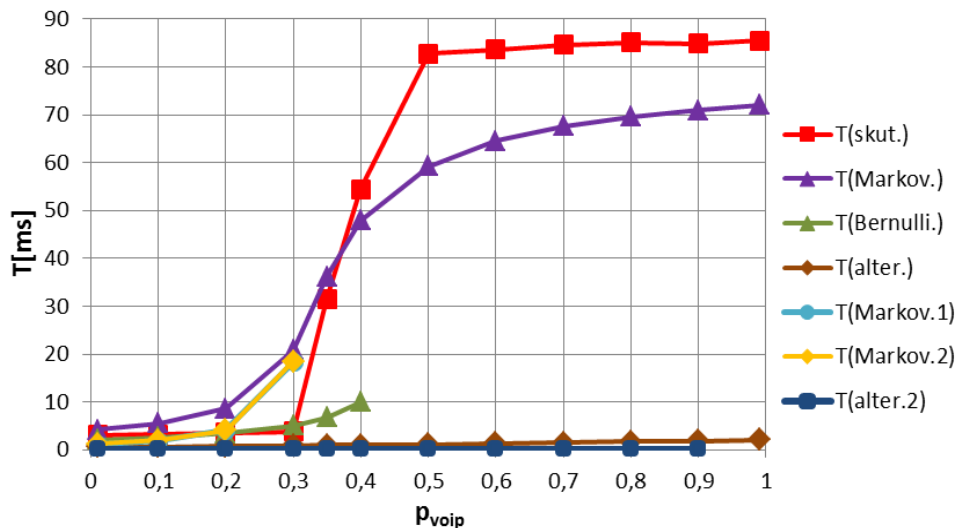


Obr. 4.9 Závislosti oneskorenia od zmeny pravdepodobnosti výberov pre VoIP pakety

Na obrázkoch Obr. 4.10, Obr. 4.11 sú porovnané oneskorenia, ktoré boli namerané v laboratórnych podmienkach s oneskoreniami odhadnutými prostredníctvom hore uvedených analytických modelov pre VoIP a IPTV pakety pri celkovom zaťažení výstupnej linky $\rho = 0,9$. Výsledky sú uvádzané len pre tie hodnoty p_1 , pre ktoré je príslušný model definovaný. Aby bolo možné určiť ako zjednodušovanie modelov vychádzajúcich z Markovových reťazcov v konečnom dôsledku ovplyvnilo výsledky modelov, boli do grafov zakomponované aj výsledky pôvodného Markovovho model s pravdepodobnostnými výbermi, ktorý bol riešený numerickými metódami.



Obr. 4.10 Porovnanie skutočného a s využitím rôznych analytických modelov odhadovaného oneskorenia VoIP paketov v závislosti od pravdepodobnosti výberu paketov z VoIP frontu pri zaťažení linky $\rho = 0,9$



Obr. 4.11 Porovnanie skutočného a s využitím rôznych analytických modelov odhadovaného oneskorenia IPTV paketov v závislosti od pravdepodobnosti výberu paketu z VoIP frontu pri zaťažení linky $\rho=0,9$

Na základe dosiahnutých výsledkov možno skonštatovať, že z uvažovaných modelov sa najviac približujú k oneskoreniu nameranému v laboratórnych podmienkach model vychádzajúci z Markovových reťazcov. Toto platí pre VoIP ako aj IPTV. Výsledky ostatných modelov sú tak voči realite ako aj voči numerickému modelu takmer na celom uvažovanom intervale $p_i \in (0; 1)$ výrazne nepresné, pričom ťažko možno očakávať, že prípadná korekcia by viedla k uspokojivým výsledkom.

5 Využitie frontových modelov na dimenzovanie parametrov QoS mechanizmov

Z pohľadu zabezpečenia kvality služieb je rovnako dôležitou vecou ako predvídanie oneskorenia v sieti, vedieť vhodne voliť kľúčové parametre QoS mechanizmov, ktoré na oneskorenie a straty vplyvajú. Presný odhad celkového ako aj frontového oneskorenia paketov dáva reálny predpoklad správnej konfigurácie zvoleného mechanizmu riadenia zahŕňajúceho, bez nutnosti jej zdĺhavého hľadania a odlaďovania v reálnych podmienkach. Optimálna voľba parametrov by mala na jednej zabezpečiť, že ani v čase najväčšej záťaže siete nedôjde k prekročeniu maximálnej prípustnej hranice koncového oneskorenia a strát pre danú prevádzku, no na druhej strane by nemala k tomuto účelu zbytočne plyvať kritickou sieťovou veličinou, ktorou je šírka pásma výstupnej linky.

Znalosti nadobudnuté z modelovania frontových QoS mechanizmov využité pri zostavení a overení matematického modelu, zameraného na odhad základných parametrov frontových QoS mechanizmov, ktorými sú garantovaná šírka pásma a kapacita paketového frontu.

5.1 Vymedzenie modelu

Model sa snaží minimalizovať potrebnú šírku pásma i -tej triedy prevádzky v mechanizmoch riadenia zahľtenia pri súčasnom dimenzovaní dĺžky i -tého frontu tak, aby sa neprekročili maximálne hodnoty strát a oneskorení, ktoré garantujú akceptovateľnú kvalitu príslušnej služby. Model vychádza so situácie veľkého zaťaženia siete, kedy sú všetky fronty okrem i -tého frontu dlhodobo neprázdne, a tak nedochádza k prerozdeleniu garantovanej šírky pásma tried s prázdnyimi frontami.

V laboratórnych experimentoch ukázalo, že ak paket prichádza do prázdneho frontu (paket i -tého frontu) musí čakať vo fronte na obsluhu susedných frontov dobu, ktorá je približne rovná dobe jeho vysielania na linku s kapacitou garantovanej šírky pásma i -tej triedy (c_i).

Uvedené predpoklady umožňuje modelovať i -ty front uvažovaného mechanizmu riadenia zahľtenia prostredníctvom ohraničeného systému $M/D/1/n$ s tým, že prvý paket v i -tom fronte je paket, ktorý sa v systéme $M/D/1/n$ práve vysielajú.

Model systému $M/D/1/\infty$ s v kapitole ukázal byť pre väčší počet tokov použiteľný na odhad frontového oneskorenia v mechanizme *FIFO*, a tak je tu predpoklad aj úspešného využitia modelu systému $M/D/1/n$ na modelovanie frontu i -tej triedy prevádzky vo viacfrontovom mechanizme riadenia zahľtenia.

Zostavený model tak predpokladá Poissonove toky paketov a deterministické doby vysielania paketov. Za dobu vysielania paketu i -tej triedy je pokladaná doba

$$\tau_i = \frac{8l_i}{c_i}, \quad (5.1)$$

kde l_i je veľkosť paketu a c_i je šírka pásma nastavovaná v mechanizme CBWFQ pre i -ty front alebo l_i je veľkosť rámca a c_i je skutočný podiel danej i -tej triedy na šírke pásma výstupnej linky.

Model predpokladá znalosť maximálnych hodnôt strát (T_i^{max}) a oneskorení (Ps_i^{max}), ktoré by sa v záujme zachovania kvality príslušnej služby nemali presiahnuť.

Model predpokladá znalosť maximálnej očakávanej intenzity vstupných tokov paketov do príslušného frontu.

Nech N je maximálny očakávaný počet tokov i -tej triedy prevádzky. Nech λ_i je stredná intenzita paketov jedného toku tejto triedy prevádzky [*pak/s*]. Pre maximálnu očakávanú intenzitu paketov λ_i^{max} platí:

$$\lambda_i^{max} = N\lambda_i. \quad (5.2)$$

5.2 Zápis modelu

Zápis uvedeného modelu pre i -tu triedu prevádzky predstavoval nelineárnu optimalizačnú úlohu matematického programovania.

$$\max(\rho_i) \quad (5.3)$$

$$E[T_i^{max}] \geq \left[\frac{\zeta_i}{P_{S_i}^{max} \lambda_i^{max}} \sum_{i=1}^L i p(i) + \frac{L_i + 1}{P_{S_i}^{max} \lambda_i^{max}} \left(1 - \frac{1 - \zeta_i (1 - \rho_i)}{\rho_i} \right) \right] / (1 - \varepsilon_i) + T_i^{fix}, \quad (5.4)$$

$$P_{S_i}^{max} \geq \frac{(1 - \rho_i) \rho_i}{e^{-2L_i (\rho_i - 1) (\rho_i + 1)^{-1}} - \rho_i^2} / (1 - \varepsilon_i), \quad (5.5)$$

$$\rho_i \in (0; 1), L_i \in \mathbb{Z}_0^+. \quad (5.6)$$

Šírku pásma bola z praktických dôvodov nahradená zaťažením i -tého frontu (ρ_i). Zaťaženie poklesom šírky pásma rastie, preto je snaha o jeho maximalizáciu. Pre zaťaženie linky i -tej triedy platí:

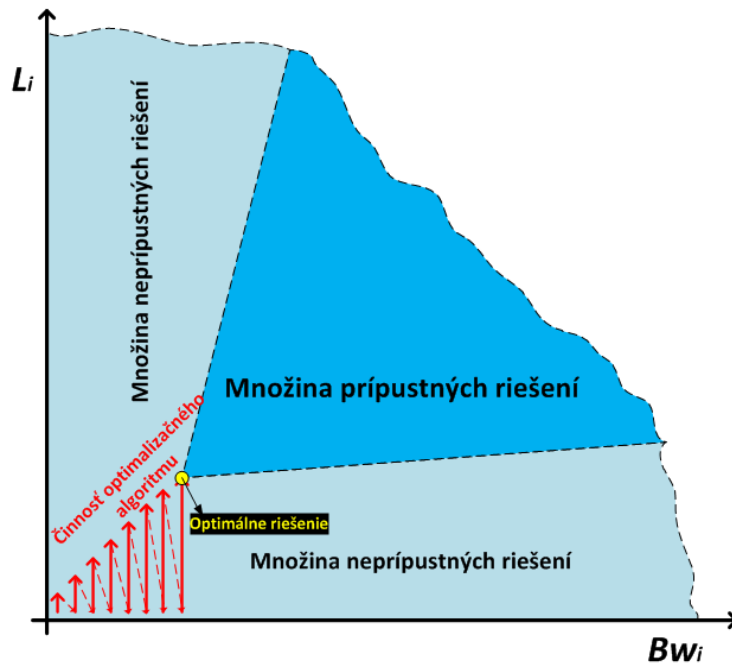
$$\rho_i = \frac{8l_i \lambda_i^{max}}{c_i}. \quad (5.7)$$

Po získaní zaťaženia tak nie je problém zo vzťahu (5.7) nájsť hľadanú šírku pásma. Zostaveniu uvedeného modelu (podmienky (5.4), (5.5)) predchádzalo odvodenie strát a oneskorení systému $M/D/1/n$ prostredníctvom difúznej aproximácie [46][47], ktorej súčasťou je parameter ζ_i . Vzťah (5.7) zohľadňuje aj existenciu fixného oneskorenia siete (T_i^{fix}) ako ja korekčnú hodnotu ε_i pri odhade oneskorenia podľa modelu $M/D/1/\infty$.

5.3 Princíp riešenia modelu

Vzhľadom na enormnú zložitosť optimalizačného modelu zostaveného v kapitole 5.2 bol model riešený numerickou metódou.

Algoritmus numerického riešenia vychádza z neprípustného riešenia daného minimálnou šírkou pásma a minimálnou dĺžkou frontu. Algoritmus postupne zväčšuje šírku pásma o vopred stanovený krok. Pre testovanú šírku pásma zväčšuje dĺžku front až kým sa nedosiahne hranica maximálneho prípustného oneskorenia. Ak vypočítané straty pre túto kombináciu šírky pásma a dĺžky frontu nepresiahli maximálnu prípustnú hranicu strát, algoritmus končí s tým, že nájdené riešenie je riešenie optimálnym. Ak straty pre testovanú kombináciu šírky pásma a dĺžky frontu presiahli maximálnu hranicu strát, šírka pásma sa zväčší o jeden vopred stanovený krok, kapacita frontu sa nastaví na jeden paket a celý proces s postupným zväčšovaním dĺžky frontu sa opakuje.



Obr. 5.1 Zobrazenie množiny prípustných a neprípustných riešení a spôsobu hľadania optimálneho riešenia

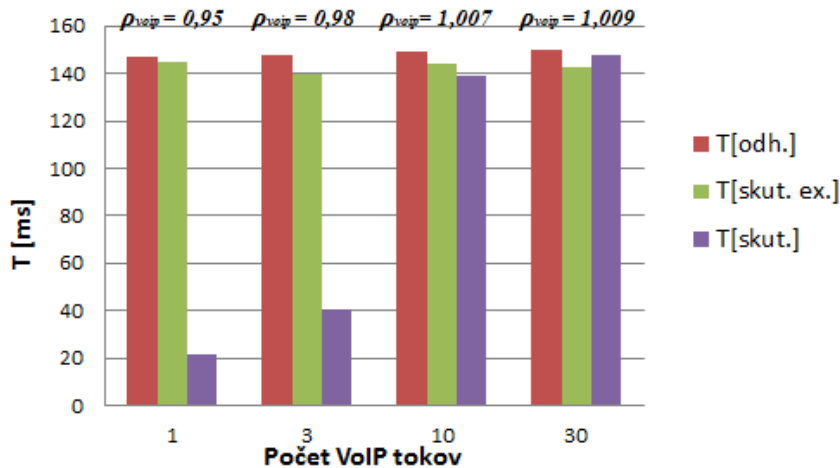
5.4 Overenie modelu a výsledky

Model bol overený na mechanizme CBWFQ (ang. *Class Based Weighted Fair Queuing*), ktorý patrí v súčasnosti medzi často nasadzované mechanizmy riadenia zahŕňajúce. CBWFQ bol nasadený v testovacej topológii s VoIP, IPTV a FTP tokmi. Konfigurácia šírky pásma a dĺžky frontu pre VoIP a IPTV toky prebiehala podľa zostaveného modelu. Z dôvodu rýchleho prevedenia numerického výpočtu bola pre tento účel zostavená aplikácia *QoS Booster*. Pre VoIP prevádzku bola maximálna hranica oneskorenia stanovená na 150 ms a maximálna hranica strát na 1% . Pre IPTV prevádzku to bolo 250 ms a $3,5\%$. Experimenty boli vykonávané s rôznym počtom tokov.

Keďže model predpokladal exponenciálne príchody paketov, bol porovnávaný s meraniami vykonanými tak s Poissonovými IPTV a VoIP tokmi, ako aj s reálnymi VoIP a IPTV tokmi. Porovnávané boli odhadnuté hodnoty strát a oneskorenia získané na základe vypočítaných parametrov z modelu a hodnoty namerané pre Poissonove a reálne toky v laboratórnych podmienkach.

5.4.1 Oneskorenie

Na obrázku Obr. 5.2 sú uvedené výsledky vzťahujúce sa na oneskorenie jedného, troch, desiatich a tridsiatich VoIP tokov. Hodnoty odhadovaného oneskorenia $T[\text{odh.}]$ a skutočného oneskorenia s exponenciálnymi medzermi $T[\text{skut. ex}]$ sú podobné, pričom podľa očakávania dosahujú hodnotu blízku 150 ms .

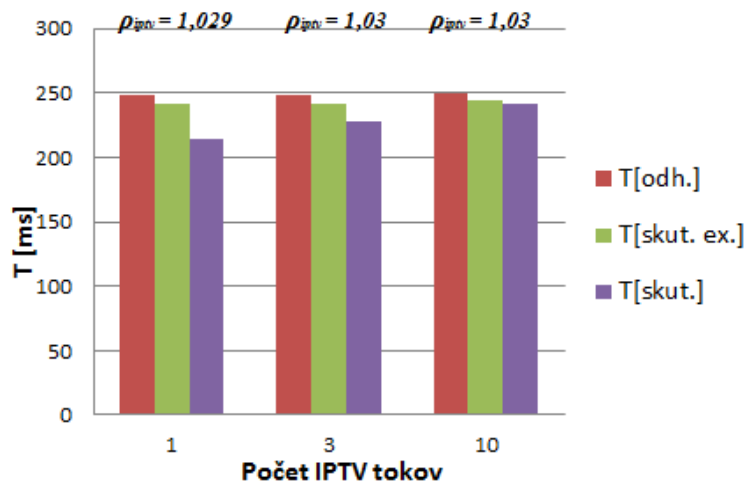


Obr. 5.2 Oneskorenia VoIP tokov v mechanizme CBWFQ

Skutočné oneskorenie scenáru s jedným VoIP tokom s deterministickými medzermi je v porovnaní s maximálnou hodnotou oneskorenia 150 ms výrazne nižšie, pričom dosahuje hodnotu približne 20 ms . Toto nadhodnotenie oneskorenia nie je nijako prekvapujúce, nakoľko pakety prichádzajú na rozdiel od Poissonovho toku pravidelne.

Na obrázku Obr. 5.3 sú uvedené výsledky vzťahujúce sa na oneskorenie jedného, troch, desiatich IPTV tokov. V prípade jedného IPTV toku sú hodnoty odhadovaného oneskorenia $T[\text{odh.}]$ a skutočného oneskorenia s exponenciálnymi medzermi $T[\text{skut. ex.}]$ veľmi podobné, pričom sa podľa očakávania blížia hodnote maximálneho prípustného oneskorenia (250 ms).

Z grafu na Obr. 5.3 si taktiež možno všimnúť, že skutočné oneskorenie pre jeden tok je voči odhadovanému oneskoreniu mierne podhodnotené. Uvedený fakt je v kontraste s tvrdením, že zhlukový charakter IPTV toku generuje väčšie oneskorenie než Poissonov tok, ktorý predpokladal zostavený model. Toto tvrdenie je však platné pre nekonečné dĺžky frontov. Pri pohľade na odpovedajúce straty (Obr. 5.4) je ale zrejme, že zníženie oneskorenia je podmienené vysokými stratami. Veľká časť paketov, ktoré by sa v prípade nekonečného frontu podieľali na navýšení priemerného oneskorenia sú v dôsledku ohraničeného frontu predčasne zahadzované, čo v konečnom dôsledku vedie k nečakanému poklesu oneskorenia.



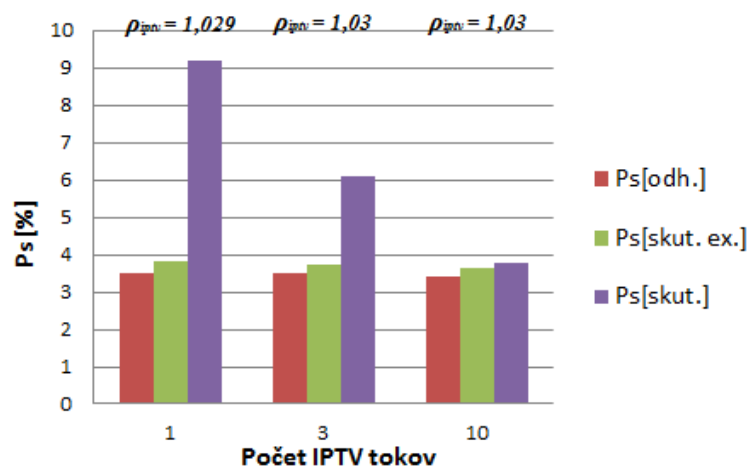
Obr. 5.3 Oneskorenia VoIP tokov v mechanizme CBWFQ

Minimálne šírky pásma získané z modelu odpovedali pre všetky počty tokov zaťaženiám blízky jednej. Relatívne vysoké horné hranice prípustných oneskorení (150 ms, resp. 250 ms) sú schopná tolerovať aj veľké priemerné dĺžky frontov a dovoľujú tak minimalizovať potrebnú šírku pásma pri spracovaní paketov k hodnotám blízky intenzite vstupného toku [bit/s]. Keďže sú prípustné aj straty paketov, šírka pásma sa vo väčšine scenárov dostala dokonca pod hranicu maximálnej vstupnej intenzity. Z toho následne rezultovali zaťaženia väčšie ako jedna. Pre navýšenie minimálnej šírky pásma by bolo potrebné znížiť maximálnu hranicu oneskorenia alebo výrazne zvýšiť fixné oneskorenie siete.

5.4.2 Straty

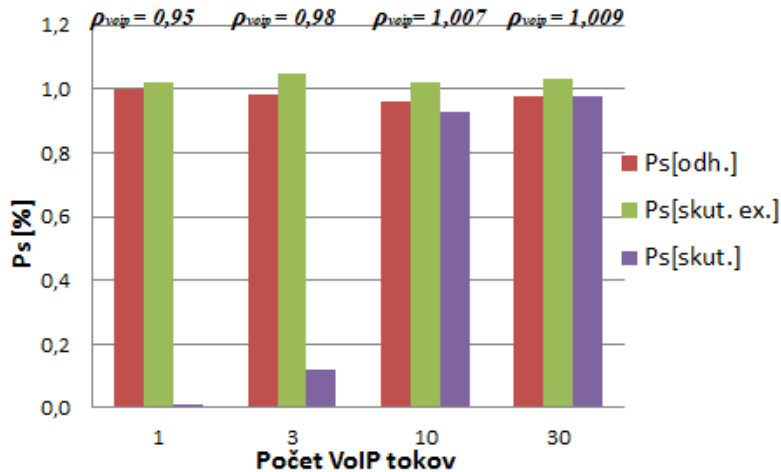
V prípade IPTV toku, ktorý je zhukovejší sú reálne straty výrazne vyššie než straty odhadované (Obr. 5.4). Podobne ako pri vývoji oneskorenia aj v prípade strát však platí, že s rastúcim počtom tokov sa nadhodnotenie nameraných strát voči stratám odhadovaným znižuje. Kým podhodnotenie odhadu pre jeden tok predstavuje 166 %, pre desať IPTV tokov je to len 7,8 %. Opäť sa potvrdilo, že kumuláciou viacerých tokov pochádzajúcich z akéhokoľvek rozdelenia konverguje výsledný tok k toku Poissonovmu. Kumuláciou desiatich IPTV tokov vznikol tok, ktorý má výrazné črty toku Poissonovho, čo vysvetľuje výrazné priblíženie nameraných strát k stratám odhadnutým prostredníctvom modelu.

Pri porovnaní výsledkov strát s výsledkami dosiahnutými pre oneskorenia (Obr. 5.3) sa dá skonštatovať, že pre uvažovanú kombináciu maximálnych prípustných strát a oneskorení sú straty na zmenu charakteru toku viac citlivé než oneskorenie.



Obr. 5.4 Porovnanie strát paketov pre IPTV toky v mechanizme CBWFQ

V prípade deterministických VoIP tokov (Obr. 5.5) podľa očakávania reálne straty nikdy nepresiahli maximálne prípustné, resp. odhadované straty. Keďže aj v prípade deterministických tokov platí, že ich kumuláciou vzniká tok Poissonov, s rastúcim počtom tokov sa skutočné straty blížili k stratám odhadovaným resp. maximálnym prípustným (1 %).



Obr. 5.5 Porovnanie strát paketov pre IPTV toky v mechanizme CBWFQ

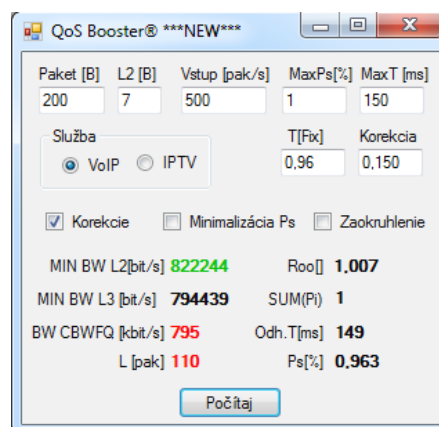
Optimalizačný model predpokladal Poissonove vstupné toky paketov. V experimentoch, v ktorých boli tak VoIP toky ako aj IPTV toky generované ako toky Poissonove, boli hodnoty odhadovaného oneskorenia a skutočného oneskorenia veľmi podobné. Uvedené tvrdenie platilo aj pre straty, čím sa potvrdila správnosť zostaveného modelu.

Experimentmi sa taktiež potvrdil fakt, že kumuláciou viacerých tokov s rôznym charakterom konverguje výsledný tok, k toku Poissonovmu. V experimentoch s desiatimi reálnymi VoIP tokmi ako aj IPTV tokmi dosahovali výsledné oneskorenia a straty hodnoty blízke hodnotám s Poissonovými tokmi. Priemerná chyba odhadu sa pohybovala na úrovni 5 %. Zostavený model je tak použiteľný predovšetkým pre počty tokov väčšie ako desať.

5.4.3 QoS Booster

Aplikácia *QoS Booster* implementuje algoritmus pre nájdenie minimálnej šírky pásma a odpovedajúcej dĺžky frontu v mechanizme CBWFQ podľa užívateľom zadaných parametrov. Parametre, ktoré zadáva užívateľ pri spustení výpočtu sú:

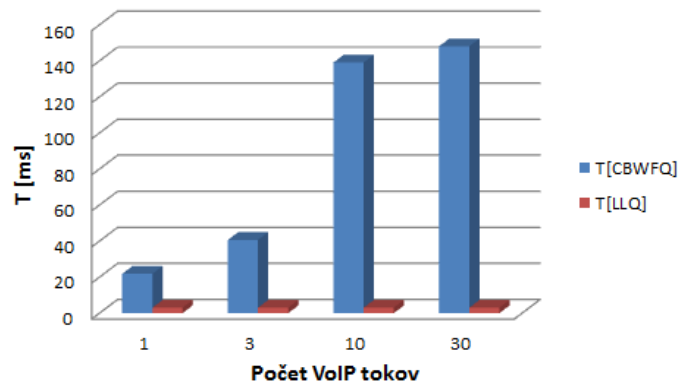
maximálna očakávaná vstupná intenzita paketov odvodená od počtu tokov a ich intenzít [pak/s], veľkosť paketu [B], veľkosť hlavičky rámca [B], maximálne prípustné oneskorenie [s], fixné oneskorenie [ms], korekčná hodnota (ϵ_{voip} , ϵ_{iptv}).



Obr. 5.6 Užívateľské rozhranie aplikácie QoS Booster

5.4.4 CBWFQ s prioritným frontom

Mechanizmus CBWFQ s prioritným frontom sa označuje aj ako LLQ. Prioritný front tohto mechanizmu je obvykle vyhradený pre prevádzku do ktorej sa očakáva nízke oneskorenie. Z Triple-Play služieb patrí medzi túto prevádzku VoIP prevádzka.



Obr. 5.7 Porovnanie oneskorenie v mechanizmoch LLQ a CBWFQ pri šírke pásma rovnaj vstupnému toku paketov

V rámci dizertačnej práce bola vykoná séria laboratórnych meraní s mechanizmom CBWFQ s prioritným frontom, do ktorého boli smerované VoIP toky. Z dosiahnutých výsledkov vyplynulo, že prioritnému frontu stačí vyhradiť šírku pásma rovnakú šírku pásma než je sumárna intenzita vstupných tokov na to, aby sa oneskorenie nepresiahlo 3 ms a straty nepresiahli 1%.

6 Metodika konfigurácie mechanizmu CBWFQ v sieti s Triple-Play tokmi

Na základe výsledkov dosiahnutých v kapitole 5 možno sformulovať odporúčania, týkajúce sa konfigurácie mechanizmu CBWFQ v prostredí Triple-Play služieb. Odporúčania je možné aplikovať aj na iné mechanizmy riadenia zahŕňajúce, ktoré umožňujú garantovať vybraným prevádzkam istú šírku pásma. Metodika vychádza z predpokladu minimalizácie potrebnej kapacity linky pri súčasnom zachovaní prípustných strát a oneskorení služieb VoIP a IPTV. Metodika využíva aplikáciu QoS Booster (kap. 5.4.3), ktorá implementovala numerické riešenie optimalizačného modelu z kapitoly 5.

6.1.1 VoIP toky

Na zaistenie prípustného oneskorenia a strát stačí VoIP tokom garantovať šírku pásma, ktorá je rovná súčtu intenzít maximálne možného počtu VoIP tokov v sieti. Pre viac ako 10 tokov je možné túto hodnotu ešte znížiť. Hodnoty minimálnej šírky pásma a vhodnej kapacity frontu sa dajú zistiť prostredníctvom aplikácie *QoS Booster*.

Ak je k dispozícii mechanizmus riadenia zahľtenia s možnosťou prioritizácie (napr. LLQ), je vhodné túto možnosť aplikovať práve na VoIP prevádzku. Minimalizuje sa tak jej oneskorenie.

Na zaistenie minimálneho oneskorenia a prípustných strát stačí VoIP tokom vyhradiť v mechanizme LLQ prioritný front so šírkou pásma, ktorá je rovná súčtu intenzít maximálneho počtu VoIP tokov v sieti. Kapacita frontu by v tomto prípade nemala presiahnuť hodnotu, ktorá by znamenala, že v prípade nečakaného zahľtenia frontu by došlo k prekročeniu maximálnej prípustnej hranice oneskorenia. Pre určenie maximálnej dĺžky prioritného frontu L_{voip} je možné využiť vzťah pre permanentne plný front.

$$T_{voip}^{max} = \frac{8L_{voip}l_{voip}}{bw_{voip}} + \frac{8(l_{voip} + h)}{c} + T_{voip}^{fix}. \quad (6.1)$$

Odkiaľ pre dĺžku frontu dostávame:

$$L_{voip} = \frac{bw_{voip}(cT_{voip}^{max} - cT_{voip}^{fix} - 8h - 8l_{voip})}{8cl_{voip}}, \quad (6.2)$$

kde c je šírka pásma linky, T_{voip}^{max} je maximálne prípustné oneskorenie, h je veľkosť hlavičky rámca, bw_{voip} je šírka pásma vyhradená pre prioritný front.

Okrem VoIP prevádzky by mal byť ďalší prioritný front s adekvátnou šírkou pásma vyhradený pre prevádzku správy siete a prevádzku smerovacích protokolov siete.

6.1.2 IPTV toky

Pre IPTV toky je vhodné vyhradiť v mechanizmoch riadenia zahľtenia samostatný front bez možnosti prioritizácie. Voľbu garantovanej šírky pásma a kapacity frontu pre viac ako desať tokov je vhodné realizovať prostredníctvom aplikácie *QoS Booster*.

6.1.3 Dátové toky

Po konfigurácii parametrov pre VoIP a IPTV prevádzku by sa mala ďalej realizovať voľba parametrov pre dátovú prevádzku. Pre dátovú prevádzku sa vyhradí zostávajúca časť kapacity linky. Šírku pásma vyhradenú dátovej prevádzke je vhodné rozdeliť podľa dôležitosti medzi rôzne triedy dátovej prevádzky. V tom prípade je ale nutné vyhradiť v rámci mechanizmu riadenia zahľtenia každej takejto triede vlastný front. Za jednoduché, no pomerne efektívne riešenie sa dá považovať smerovanie dátovej prevádzky do implicitného frontu mechanizmu CBWFQ (tzv. *class default*). Pri obsluhu prevádzky tejto triedy sa dá uplatniť mechanizmus WFQ (férovo-vážené výbery z frontov, ktorý nie je nutné dodatočne konfigurovať). Mechanizmus WFQ rozdelí garantovanú šírku pásma prevádzky implicitného frontu mechanizmu CBWFQ spravodlivo medzi jednotlivé dátové toky. Prípadné uprednostnenie tokov určitých tried je možné dodatočne realizovať zmenou hodnoty DSCP resp. ToS alebo vyhradením vlastného frontu tejto prevádzke.

Dátová prevádzka využíva na svoj prenos protokol TCP, dôsledkom čoho zdrojové stanice prispôsobujú intenzitu vysielania paketov v závislosti na aktuálnych podmienkach v sieti. Experimentmi s FTP tokom a vlastnou aplikáciou SNMP Tester sa potvrdilo, že intenzita vysielania prevádzky protokolu TCP sa riadi podľa aktuálnej hodnoty celkového oneskorenia. Z výsledkov dizertačnej práce vyplynulo, že celkové oneskorenia závisí predovšetkým od aktuálneho stavu paketových frontov medzi vysielacou a prijímacou stanicou. Intenzita vysielania sa tak prispôsobuje aktuálnym dĺžkam paketových frontov. Ak je teda dátovej prevádzke vyhradený vlastný front, nemá zmysel zahadzovať pakety predčasne skracovaním jeho kapacity.

7 Záver a zhodnotenie výsledkov

Z výsledkov vyplynulo, že odhad oneskorenia v prípade nezahltenej linky, kedy je frontové oneskorenie nulové je možné uskutočniť relatívne presne aj bez ďalších korekcií.

Ak frontové oneskorenie nie je nulové, tvorí významnú časť koncového oneskorenia, pričom podiel frontového oneskorenia na celkovom oneskorení závisí predovšetkým od zaťaženia linky na ktorej sa v dôsledku jej zahltienia front paketov vytvoril.

V prípade plných frontov mechanizmov FIFO (kap. 4.1.1) alebo WRR (kap. 4.3.2.1) je matematický odhad celkového oneskorenia v sieti taktiež možné realizovať prostredníctvom zostavených analytických modelov veľmi presne. Relatívna chyba odhadu sa pohybuje do 5 %. Napriek tomu, že k permanentne plným frontom dochádza v správne nadimenzovaným sieťam len výnimočne, znalosť oneskorenia pre tento výnimočný stav dáva možnosť predvídať najväčšie oneskorenie, ktoré môže pre zvolenú konfiguráciu v sieti nastať. Oneskorenie odhadnuté pre zahltené fronty tvorí tak po prirátaní fixného oneskorenia siete a serializačného oneskorenia kritickej linky hornú hranicu očakávaného oneskorenia v sieti.

Najobsiahlejšiu časť dizertačnej práce bola vená problematike premenlivých frontov, ktoré sa v praxi vyskytujú najčastejšie (kap. 4.3). Problémom pri analytickom vyjadrovaní alebo odhade frontového oneskorenia je v prípade premenlivých frontov skutočnosť, že dĺžka paketového frontu, s ktorou veľkosť frontového oneskorenia úzko súvisí, nie je konštantná veličina, ale sa s časom náhodne a rýchlo mení. Hlavným faktorom, ktorý výrazne ovplyvňuje stav frontu v čase a komplikuje tak odhad stredného oneskorenia, je charakter vstupného toku. Ak je známa stredná intenzita vstupného toku paketov, veľkosti paketov a šírka pásma kritickej linky, je možné jednoznačne určiť len skutočnosť, či ide o premenlivý front alebo o permanentne plný front. Aj napriek znalosti konkrétnych hodnôt strednej intenzity vstupného toku paketov, veľkosti paketov a šírky pásma kritickej linky nie je možné presne odhadnúť stredné oneskorenie, ktoré môže závislosti od charakteru toku nadobúdať širokú škálu hodnôt. Čiastočným vyjadrením charakteru toku resp. jeho variability je napríklad variačný koeficient alebo Hurstov parameter. Zohľadnenie uvedených veličín je však v matematických modeloch za účelom odhadu oneskorenia veľmi problematické. Vo väčšine prípadov sa variabilné medzery paketov vo vstupnom toku modelujú bezpomäťovým exponenciálnym rozdelením, ktoré tvorí akýsi kompromis medzi deterministickými

a zhlukovými tokmi. Tok s exponenciálnymi medzerami medzi príchodmi paketov je označovaný ako Poissonov.

Reálne VoIP toky majú deterministické medzery medzi paketmi ($\nu = 0$), a tak v porovnaní s exponenciálnymi tokmi ($\nu = 1$) majú tendenciu generovať menšie dĺžky frontov a tým aj frontové oneskorenie. IPTV toky naopak pochádzajú z rozdelení s ťažkými chvostmi ($\nu > 1$) a v porovnaní s exponenciálnymi tokmi ($\nu = 1$) majú tendenciu generovať väčšie dĺžky frontov a tým aj frontové oneskorenie. Z uvedených dôvodov sú analytické odhady oneskorenia pre VoIP toky nadhodnotené, zatiaľ čo pre IPTV toky podhodnotené. Poľahčujúcou okolnosťou pri modelovaní tokov a odhade oneskorenia je fakt, že zlúčením viacerých tokov z ľubovoľným charakterom konverguje výsledný tok, k toku Poissonovmu so strednou intenzitou rovnou súčtu stredných intenzít jednotlivých tokov. Keďže drvivá väčšina analytických modelov predpokladá kvôli jednoduchosti práve Poissonove toky paketov, s rastúcim počtom tokov sa ich presnosť pri odhade frontového oneskorenia zvyšuje. Analytický odhad oneskorenia v prípade mechanizmu FIFO bol pri kumulácii desiatich deterministických VoIP podhodnotený (kap. 4.3.1.1), pričom predstavoval chybu odhadu 5,1 %. V scenári s desiatimi IPTV tokmi (kap. 4.3.1.1), ktoré sú zhlukovejšie ako tok exponenciálny, bolo analytický odhad podľa očakávania nadhodnotený, pričom predstavoval chybu 5,4 %. Simulačný odhad vykonaný prostredníctvom nástroja Opnet Modeler so zachovaným charakterom VoIP a IPTV tokov bol vo všetkých vykonaných scenároch s premenlivým FIFO frontom presný, relatívna chyba odhadu sa pohybovala pod hranicou 2,5 %. Keďže zostavené matematické modely predpokladali Poissonove vstupné toky paketov, za účelom overenia správnosti zostavených modelov boli v laboratórnych experimentoch generované aj toky s exponenciálnymi medzerami. Po menších korekciách, ktoré súviseli s existenciou dodatočného oneskorenia pri zaraďovaní paketov do frontov v zariadeniach CISCO, výsledky modelov presne korešpondovali s výsledkami laboratórnych meraní. Tým sa potvrdila ich správnosť a prípustnosť pri odhade frontového oneskorenia. Uvedené konštatovanie platilo aj pri kumulácii IPTV toku a niekoľkých VoIP tokov. Korekčné hodnoty (ϵ_{voip} , ϵ_{iptv}) pre spresnenie odhadu frontového oneskorenia prostredníctvom zostavených modelov boli pre VoIP a IPTV toky stanovené v kapitole 4.3.1.

Najväčšie úsilie v rámci dizertačnej práce bolo venované modelovaniu mechanizmu WRR 3.2, ktorý efektívne rozdeľuje šírku pásma linky medzi jednotlivé triedy služieb. Okrem problému s charakterom vstupného toku paketov je ďalším významným problémom pri matematickom modelovaní mechanizmu WRR skutočnosť, že šírka pásma vyhradená prevádzkam jednotlivých tried nie je konštantná veličina, ale sa časom náhodne mení. Pokiaľ je totiž jeden alebo viaceré fronty mechanizmu prázdne, neprázdne fronty si úmerne podľa pridelených váh (garantovaných širok pásma) rozdelia šírku pásma vyhradenú pre tieto fronty. V dvojfrontovej konfigurácii to znamená asi toľko, že ak je jeden front prázdny a druhý front neprázdny, neprázdny front je počas celej doby trvania tohto stavu oprávnený využívať okrem svojej vyhradenej šírky pásma aj šírku pásma frontu susedného. Na aktuálny stav paketov vo frontoch a tým pádom aj na oneskorenie majú tak zásadný vplyv aj veličiny, ktoré charakterizujú prevádzku a konfiguráciu frontov susedných (garantovaná šírka, veľkosť paketov, charakter vstupného toku, intenzita vstupného toku). Keďže početnosť týchto veličín sa s pribúdajúcim počtom frontov násobí, stáva sa odhad oneskorenia o to ťažším. Táto

skutočnosť výrazne komplikuje zostavenie presného analytického modelu aj za predpokladu Poissonových vstupných tokov. V rámci dizertačnej práce bolo zostavených niekoľko modelov, ktoré sa snažili dynamické prerozdelenie nevyužitej šírky pásma rôznymi spôsobmi zohľadniť a prispieť tým k vylepšeniu odhadu frontového oneskorenia paketov v mechanizme WRR. Podľa očakávania bol najpresnejší vlastný simulačný model (QoS Tester), ktorý okrem autentických vstupných tokov presne simuloval aj činnosť mechanizmu WRR tak, ako je implementovaný v zariadeniach firmy CISCO. Priemerná relatívna chyba odhadu koncového oneskorenia po vykonaní dodatočnej korekcie predstavovala len 3 %.

Pre účely matematického modelovania bol obslužný cyklus mechanizmu WRR spolu s váženými výbermi paketov nahradený tzv. pravdepodobnostnými výbermi z frontov (kap. 4.3.2.1). Jednotlivým frontom tak boli pridelené pravdepodobnosti, ktoré vyjadrujú šancu výberu paketu z daného frontu. Hodnoty boli úmerné počtu paketov, ktoré vyberie mechanizmus WRR v rámci jedného obslužného cyklu. Konverzii váh jednotlivých frontov udávaných v bajtoch na pravdepodobnosti výberov paketov z týchto frontov a späť bola venovaná kapitola kap. 4.3.2.1. Simulačný model s pravdepodobnostnými výbermi a autentickým tokmi dosahoval podobné výsledky ako pôvodný simulačný model s váženými výbermi paketov z frontov. Priebehy oneskorenia v závislosti od zmeny váh mali však plynulejší priebeh (kap. 4.3.2.2).

Z matematických modelov sa ako najpresnejší ukázal byť model vychádzajúcich z Markovových reťazcov, ktorý predpokladal pravdepodobnostné výbery z frontov. Napriek tomu, že sa jednalo len od dvojfrontový model, analytické riešenie vychádzajúce z rezov prechodového grafu a následnej aplikácie vytvárajúcich funkcií nebolo z dôvodu veľkého množstva neznámych úspešne dokončiteľné. Následne bol model popísaný systémom diferenciálnych rovníc, ktorý bol kvôli enormnej zložitosti riešený numericky. Pri riešení bola využitá Eulerova aproximácia metóda, pričom samotný numerický výpočet vykonávala vlastná aplikácia *QoS Tester*. Znížená presnosť voči už spomínaným simulačným modelom vyplývala predovšetkým z toho, že model nutne uvažoval tak exponenciálne medzery medzi príchodmi paketov, ako aj exponenciálne veľkosti paketov. Priemerná chyba modelu voči realite predstavovala 35 %

Aby mohol byť pôvodný Markovov model s pravdepodobnostnými výbermi paketov riešený aj analytickými metódami, bolo nutné vykonať dvojnásobné zjednodušenie. Prvé zjednodušenie spočívalo v zavedení strednej intenzity vysielenia všetkých paketov v modeli bez ohľadu na ich typ. To opatrenie viedlo k výraznému zjednodušeniu prechodového grafu popisujúceho modelovaný systém. No aj napriek tomu analytické riešenie vychádzajúce z rezov prechodového grafu a následnej aplikácie vytvárajúcich funkcií nebolo z dôvodu vzniku dodatočných neznámych úspešne dokončiteľné. Numerického riešenie zjednodušeného modelu nevykazovalo výrazný odklon od numerického riešenia pôvodného modelu.

Druhé vynútené zjednodušenie spočívalo v odstránení predpokladu dynamického prerozdelenia nevyužitej šírky pásma prázdnych frontov. To sa v prechodovom grafe navonok prejavilo tak, že intenzity vysielenia paketov boli vo všetkých stĺpcoch resp. riadkoch grafu rovnaké. Toto zjednodušenie v kombinácii so zjednodušením predchádzajúcim už viedlo k nájdeniu hľadaných vytvárajúcich funkcií pre pravdepodobnosti stavov modelu v uzavretom tvare, z ktorých nebolo zložitú získať vhodnými úpravami výsledné vzťahy pre

hľadané oneskorenie paketov jednotlivých typov. Vzhľadom na koncepciu modelu výsledné vzťahy zahŕňajú už aj serializačné oneskorenie pri vysielaní paketu z tx-ringu na výstupnú linku. Pre získanie celkového oneskorenia k nim tak stačilo pripočítať fixné oneskorenie v sieti. Výsledky dosiahnuté týmto modelom sú voči pôvodnému nezjednodušenému modelu pochopiteľne mierne nadhodnotené, keďže model neuvažoval prerozdelenie šírky pásma neprázdného frontu. S rastom zaťaženia tak konvergujú výsledky zjednodušeného modelu k výsledkom originálneho modelu. Nevýhodnou modelom je aj značne obmedzený definičný obor, ktorý znemožňuje model využiť pre všetky pravdepodobnosti výberu paketu resp. váhy.

Výsledky ostatných analytických modelov predstavených v kapitole boli skôr sklamaním ako perspektívou pre odhad oneskorenia v IP sieťach.

Na základe nadobudnutých výsledkov možno skonštatovať, že na odhad oneskorenia v sieti s mechanizmom WRR alebo iným mechanizmom jemu podobným (napr. CBWFQ) je najlepšie využiť simuláciu, ktorá je schopná zohľadniť tak charakter tokov ako aj dynamické prerozdelenie šíry pásma medzi frontami. Horný odhad oneskorenia je však možné uskutočniť prostredníctvom modelu, ktorý predpokladal permanentne zahľtené fronty paketov (kap. 4.2). Do úvahy taktiež prichádza aj možnosť použitia modelu $M/D/1/n$. Overenie tejto alternatívy sa však z časových dôvodov nepodarilo zrealizovať, a tak sa stáva predmetom ďalšieho výskumu v budúcnosti.

Z pohľadu zabezpečenia kvality služieb je rovnako dôležitou vecou ako predvídanie oneskorenia v sieti, vedieť vhodne voliť kľúčové parametre QoS mechanizmov, ktoré na oneskorenie a straty vplyvajú. Optimálna voľba parametrov by mala zabezpečiť, že aj v čase najväčšej záťaže siete nedôjde k prekročeniu maximálnej prípustnej hranice koncového oneskorenia a strát pre danú prevádzku, no na druhej strane by nemala k tomuto účelu zbytočne plytvať kritickou sieťovou veličinou, ktorou je šírka pásma výstupnej linky.

Znalosti a skúsenosti nadobudnuté z tvorby a overovania modelov QoS mechanizmov boli v kapitole 6 využité pri zostavení a overení matematického modelu, ktorý bol zameraný na odhad základných parametrov frontových mechanizmov pre IPTV a VoIP toky. Model sa snaží minimalizovať potrebnú šírku pásma a vhodne dimenzovať dĺžku frontu tak, aby bola zaistené prípustné hodnoty strát a oneskorenia pre garanciu kvality príslušnej služby. Vyjadrením kvality služby v modeli sú hodnoty maximálneho prípustného oneskorenia a strát, ktoré by sa nemali v záujme udržania akceptovateľnej kvality služby presiahnuť. V praxi by sa dal zostavený model využiť aj pri zavádzaní služieb VoIP a IPTV v sieťach pôvodne dimenzovaných na poskytovanie dátových služieb. Prostredníctvom zostaveného modelu by sa odhadla minimálna šírka pásma, o ktorú je nutné rozšíriť existujúcu kapacitu kritických liniek bez toho, aby bola ovplyvnená kvalita dátových služieb. Model bol overený na mechanizme CBWFQ, ktorý patrí v súčasnosti medzi často nasadzované mechanizmy riadenia zahľtenia.

Keďže sa ukázalo, že kumuláciou tokov pochádzajúcich z ľubovoľného rozdelenia pravdepodobnosti konverguje výsledný tok k toku Poissonovmu (kap. 4.3.1.1), presnosť výsledkov modelu sa pre akékoľvek toky zvyšuje s rastúcim počtom tokov. V prípade deterministických VoIP tokov, ktoré majú tendenciu generovať kratšie paketové fronty než toky exponenciálne, je model použiteľný pre ľubovoľný počet tokov tak, aby boli splnené

kvalitatívne predpoklady dané hodnotami maximálny strát a oneskorení. Pre IPTV toky, ktoré majú tendenciu generovať väčšie oneskorenie a straty ako toky Poissonove, je model použiteľný pre viac ako 10 tokov. Pri desiatich IPTV tokoch dosahovala relatívna chyba odhadu strát hodnotu približne 7,8 % (kap. 5.4.2). V prípade oneskorenia bola relatívna chyba odhadu oneskorenia len 5,1 % (kap. 5.4.1). Možnosť zlepšenia odhadu hľadaných veličín pre nižší počet IPTV tokov bola naznačená v dizertačnej práci. Jej realizácia sa stáva predmetom ďalšieho výskumu v budúcnosti.

Znalosti a skúsenosti z tvorby a overovania zostavených modelov boli v kapitole 6 zosumarizované do metodiky voľby parametrov mechanizmu CBWFQ v sieťach s Triple-Play tokmi. Napriek tomu, že metodika je koncipovaná na mechanizmus CBWFQ, nevyklučuje to jej použitie aj v prípade iných mechanizmov riadenia zahltenia, ktoré efektívne delia šírku pásma linky medzi jednotlivé triedy prevádzky.

8 Prínosy práce

- Analýza a zhodnotenie súčasného stavu v oblasti mechanizmov garancie kvality služieb v architektúre diferencovaných služieb.
- Návrh a implementácia simulačných a analytických modelov mechanizmov riadenia zahltenia FIFO a WRR so zameraním na odhad frontového a celkového oneskorenia v IP sieťach s Triple-Play tokmi.
- Návrh metodiky testovania a korekcie modelov mechanizmov riadenia zahltenia v IP sieťach.
- Rozšírenie funkcionality aplikácie *D-ITG* o možnosť generovania paketov podľa vopred vytvoreného scenára.
- Implementácia aplikácie *SNMP Tester*, ktorá umožňuje vyhodnocovanie stavu frontov mechanizmov riadenia zahltenia s extrémne vysokou frekvenciou.
- Návrh modelu pre voľbu parametrov mechanizmov riadenia zahltenia so zameraním na minimalizáciu potrebnej šírky pásma pre Triple-Play toky.
- Zostavenie aplikácie *QoS Booster* implementujúcej algoritmus modelu pre voľbu parametrov mechanizmov riadenia zahltenia.
- Návrh metodiky konfigurácie mechanizmu CBWFQ s využitím zostavenej aplikácie v prostredí s Triple-Play tokmi.

Použitá literatura

- [1] Pužmanová R.: *Moderní komunikační sítě*, Brno, Computer Press, a.s., 2006, ISBN: 80-251-1278-0
- [2] Sheldon, T.: *Encyclopedia of Networking and Telecommunication*, McGraw-Hill Media, Osborne, 2001, ISBN: 0-07-212005-3
- [3] Odom, W., Healy, R., Donohue, D.: *CCIE Routing and Switching Certification Guide*, Cisco Press, Indianapolis, 2009, Four Edition.
- [4] Sziget, T., Hattingh, Ch.: *End-To-End QoS Network Design*, Cisco Press, 2004, ISBN: 1-58705-176-1
- [5] Cisco Systems: *Implementing CISCO Quality of Service – Student Guide, Volume 1*, Cisco Press, 2006.
- [6] Cisco Systems: *Implementing CISCO Quality of Service– Student Guide, Volume 2*, Cisco Press, 2006.
- [7] Donahue, G.A.: *Kompletní průvodce síťového experta, Brno, Computer Press, a.s., 2009, ISBN: 978-80-251-2247-1*
- [8] Cisco Systems: *Cisco IOS Quality of Service Solutions Configuration Guide*, Cisco Systems, 2008-2009, dostupné na:
http://www.cisco.com/en/US/docs/ios/12_2/qos/configuration/guide/fqos_c.html
- [9] Cisco Systems: *Quality of Service Solutions Configuration Guide*, Cisco IOS Release 15.1M&T, 2010, dostupné na:
http://www.cisco.com/en/US/docs/ios/qos/configuration/guide/15_1/qos_15_1_book.pdf
- [10] Flannagan, B., Froom, R., Turek, K.: *Quality of Service in Campus Network*, Cisco Press, 2003, ISBN: 1-58705-120-6
- [11] Komosný, D.: *Kvalita Služeb IP sítí*, v *Electroreue*, 6/2004
- [12] Lapukhov, P.: *Understanding Custom Queuing*, dostupné na:
<http://blog.ine.com/2008/08/12/understanding-custom-queuing/>
- [13] Klimo, M. a kol.: *Teória IP telefónie*, EDIS, Žilina, 2009, 978-80-8070-915-0
- [14] FITCE Study Commission: *The Study of the Network Performance Considering Customer Requirements*, Final Report 1991-1993
- [15] Song, J., Chang, Y.M., Lee, S.S.: *Overview of ITU-T NGN QoS Control*, v: *IEEE Communications Magazine*, 2007
- [16] Sulman, Ch.: *Final Text of X.641 for Acceptance at the SG7 December 1997 Plenary*, Geneva, 1 – 12 December 1997
- [17] Halteren, A., Franken, L., Vries, D.: *QoS architectures and mechanisms*, State of the art, 1999, s. 1
- [18] IETF: *Multiprotocol Label Switching Architecture (RFC 3031)*, dostupné na:
<http://tools.ietf.org/html/rfc3031>
- [19] IETF: *A Protocol for RSVP-based Admission Control over IEEE 802-style networks (RFC 2814)*, 2000, dostupné na: <http://www.ietf.org/rfc/rfc2814.txt>
- [20] S. Oueslati and J. Roberts, “A New Direction for Quality of Service: Flow-Aware Networking,” *Proc. Conf. Next Generation Internet Network (NGI)*, Apr. 2005.

- [21] ITU-T Draft Rec. Y.flowreq, “Requirements for the Support of Stateful Flow-Aware Transport Technology in an NGN.”
- [22] Hernandez, A., Magana, E.: *One-way delay measurement and characterization*, v: IEEE ICNS '07, Washington, DC, USA, s. 114.
- [23] ITU-T Recommendation Y.1540: *Internet Protocol Data Communication Service - IP Packet Transfer and Availability Performance Parameters*, 02/2002.
- [24] Koodli.R, Ravikanth.R: *One-way Loss Pattern Sample Metrics* (RFC 3357), 2002.
- [25] ITU-T Recommendation G.114: *One-way transmission time*, 05/2003.
- [26] *Fiber to the X*, dostupné na: http://en.wikipedia.org/wiki/Fiber_to_the_x
- [27] JAKAB, F. a kol.: *Možnosti zabezpečenia prenosu hlasu s požadovanou kvalitou – príspevok na konferenciu ECI*, dostupné na: <http://voip.cnl.tuke.sk/node/4>
- [28] FLOYD, S., JACOBSON, V.: *Random Early Detection Gateways for Congestion Avoidance*, Lawrence Berkeley Laboratory: University of California, August 1993, dostupné na: <http://www.icir.org/floyd/papers/red/red.html>
- [29] HADEN, R.: *VoIP – Data network resource*, dostupné na: <http://www.rhyshaden.com/voice.htm>
- [30] ZEMAN, Otto: *Modelování chování páteřních směrovačů DiffServ domény*, Bakalárska práca, Vysoké učení technické v Brne, máj 2006, Brno
- [31] Janoňák, M.: *Simulačný model prístupovej siete v programe OPNET Modeler*, 2010, Žilina, diplomová práca (r.č.71/2010)
- [32] International Cablemakers Federation, *How Much Bandwidth*, dostupné na: http://www.icf.at/en/6000/how_much_bandwidth.html
- [33] PUŽMANOVÁ, R.: *IPTV vyžaduje nejen šířku pásma, ale vysokou kvalitu*, dostupné na: <http://www.dsl.cz/clanek/891-iptv-vyzaduje-nejen-sirku-pasma-ale-vysokou-kvalitu-i>
- [34] Wikipedia: *IPTV*, verzia článku 12.11.2010, dostupné na: <http://en.wikipedia.org/wiki/IPTV>
- [35] Internetworking Technology Overview, June 1999.
- [36] Semeria,Ch.: *Supporting differentiated service classes: Queue scheduling principles*, dostupné na: http://cn.juniper.net/solutions/literature/white_papers/200020.pdf
- [37] Semeria,Ch.: *Supporting differentiated service classes: Active queue memory management*, dostupné na: http://cn.juniper.net/solutions/literature/white_papers/200021.pdf
- [38] Semeria,Ch.: *Supporting differentiated service classes: TCP Congestion control mechanisms*, dostupné na: http://cn.juniper.net/solutions/literature/white_papers/200022.pdf
- [39] Herbert, H.: *Queuing Methods from FIFO to CB-WFQ/PQ*, 2009, dostupné na : <http://www.perihel.at/2/rno/03-QoS-Queuing-Methods.pdf>
- [40] Congestion Management: *Configuring FIFO and WFQ*, dostupné na: http://www.cisco.com/en/US/docs/ios-xml/ios/qos_conmgt/configuration/15-mt/qos-conmgt-15-mt-book.pdf
- [41] Ethernet, dostupné na: <http://cs.wikipedia.org/wiki/Ethernet>
- [42] Ethernet Frame , dostupné na: http://en.wikipedia.org/wiki/Ethernet_frame

- [43] Management information base, dostupné na : Janáček, J.: Matematické programování, Edis, 2003, Žilina, ISBN: 80-8070-054-0
- [44] Takagi, H.: *Queueing Analysis, Volume 1: Vacation and Priority Systems*, Amsterdam, North-Holland, 1991, ISBN: 0-444-88910-8
- [45] Hassan, M., Sarker, R., Atiquzzaman, M.: *Modeling IP-ATM gateway using M/G/1/N queue*, dostupné na: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.53.4330>
- [46] Medhi, J.: *Stochastic Models in Queueing Theory*, Academic Press, 1991
- [47] Gelenbe, E., Pujolle, G.: *Introduction to Queueing Networks*, John Wiley and Sons, 1987
- [48] Gebali, F.: *Analysis of Computer and Communication Networks*, Springer, New York, 2008, ISBN: 978-0-387-74436-0
- [49] Smieško, J., Peško, Š.: *Stochastické modeli operačnej analýzy*, Žilina, 2006
- [50] Peško, Š.: *Operačná analýza 2*, Žilina, 2000- 2001, skriptá
- [51] Hottmar, V., Adamec B.: *Analytical Model of Weighted Round Robin Service System*; In: Journal of Electrical and Computer Engineering: Volume 2012, Article ID 374961, dostupné na: <http://www.hindawi.com/journals/jece/2012/374961/>
- [52] Jánošíková, L.: *Modelovanie a optimalizácia*, dostupné na: http://frdsa.fri.uniza.sk/~janosik/Vyuka/MO_pred1.pdf
- [53] Markechová, D., Stehlíková, B., Tirpáková, A.: *Štatistické metódy a ich aplikácie*, Nitra, 2011, ISBN 978-80-8094-807-8
- [54] Michalíková, J.: *Rozdelenia s ťažkými chvostami v neživotnom poistení*, Bratislava, 2009, diplomová práca
- [55] Valasek, J.: *Návrh a realizácia sieťového segmentu na báze protokolu IPv6*, Košice, 2004, dostupné na: http://www.cnl.tuke.sk/files/DP2004_Valasek_Narh_a_realizacia_experimentalneho_sietoveho_segmentu_na_baze_protokolu_IPv6.pdf
- [56] Walter, J: *Stochastické modely v ekonómii*, SNTL – Nakladatelství technické literatury, Praha, 1970
- [57] Klimo, M.: *Modelovanie IP prevádzky*, dostupné na ŽU FRI KIS
- [58] T-COM: Project: *Modeling of IP traffic*, Research report, 2007-2008, dostupné na KIS FRI ŽU
- [59] Giambene, G.: *Queueing theory and telecommunications networks and application*, vyd. Springer, USA, 2005, ISBN 978-0378-24065-7
- [60] D-ITG, Distributed Internet Traffic Generator, dostupné na: <http://www.grid.unina.it/software/ITG>

Vlastné publikácie

- [61] *Analytical model of QoS mechanism - Priority Queueing using Markov's chains* / Dušan Nemček. In: AD ALTA : journal of interdisciplinary research. - ISSN 1804-7890. - Vol. 1, Iss. 1 (2011), s. 113-114.
- [62] *Modelling of QoS mechanisms* / Dušan Nemček. In: Zimná škola MICT [elektronický zdroj] : mathematics for information and communication technologies : 6th winter school of mathematics for ICT : Šachtičky

- 3.-8.1.2011. - Banská Bystrica: Science and Research Institute, MBU, 2011. - ISBN 978-80-557-0252-0. - CD-ROM, s. 78-80.
- [63] *Inginiering approach to modelling of QoS mechanism - weighted round robin Markov's chain* / Dušan Nemček. In: TRANSCOM 2011 : 9-th European conference of young research and scientific workers : Žilina, June 27-29, 2011, Slovak Republic. - Žilina: University of Žilina, 2011. - ISBN 978-80-554-0372-4. - S. 181-184.
- [64] *Models of QoS mechanism - Priority Queuing* / Dušan Nemček. In: Journal of information, control and management systems. - ISSN 1336-1716. - Vol. 8, no. 4 spec. iss. (2010), s. 383-392.
- [65] *Models of QoS mechanism - Weighted Round Robin* / Dušan Nemček. In: QUAERE 2011 [elektronický zdroj] : recenzovaný sborník příspěvků interdisciplinární mezinárodní vědecké konference doktorandů : 25.-29. dubna 2011, Hradec králové, Česká republika. - Hradec Králové: Magnanimitas, 2011. - ISBN 978-80-904877-3-4. - S. 912-921. - Požadavky na systém: CD-ROM mechanika.
- [66] *Analytical model of QoS mechanism - weighted round Robin using Markov's chains* / Dušan Nemček. In: QUAERE 2011 [elektronický zdroj] : recenzovaný sborník příspěvků interdisciplinární mezinárodní vědecké konference doktorandů : 25.-29. dubna 2011, Hradec králové, Česká republika. - Hradec Králové: Magnanimitas, 2011. - ISBN 978-80-904877-3-4. - S. 922-927. - Požadavky na systém: CD-ROM mechanika.
- [67] *Analytický model QoS mechanizmu - Priority Queuing s využitím Markovových řetězců = Analytical model of QoS mechanism - Priority Queuing using Markov's chains* / Dušan Nemček. In: MMK 2010 [elektronický zdroj] : mezinárodní Masarykova konference pro doktorandy a mladé vědecké pracovníky : 6.-10. prosince 2010, Hradec Králové, Česká republika : sborník příspěvků z mezinárodní vědecké konference. - Hradec Králové: Olga Čermáková, grafické a reklamní studio, 2010. - ISBN 978-80-86703-41-1. - S. 920-925. - Požadavky na systém: CD-ROM mechanika.
- [68] *QoS mechanizmus Priority Queuing aj jeho modelovanie* / Dušan Nemček. In: ITAT 2010 = Informačné technológie - aplikácie a teória : zborník príspevkov prezentovaných na konferencii ITAT Smrekovica, Slovensko, September 2010. - Seňa: PONT s. r. o., [2010]. - ISBN 978-80-970179-3-4. - S. 123-124.
- [69] *Modelovanie QoS mechanizmu Weighted Round Robin s využitím teórie hromadnej obsluhy* / Dušan Nemček. In: ITAT 2009 = Informačné technológie - aplikácie a teória : zborník príspevkov prezentovaných na pracovnom seminári ITAT : 25. -29. september 2009, Horský hotel Kráľová studňa, Slovensko. - Seňa: PONT s.r.o., 2009. - ISBN 978-80-970179-1-0. - [2] s.