

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ УКРАИНЫ
ХАРЬКОВСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ
РАДИОЭЛЕКТРОНИКИ

На правах рукописи

СИМОНЕНКО Александр Викторович

УДК 621.391

**ПОТОКОВЫЕ МОДЕЛИ И МЕТОДЫ УПРАВЛЕНИЯ ОЧЕРЕДЯМИ
НА ИНТЕРФЕЙСАХ МАРШРУТИЗАТОРОВ ТРАНСПОРТНЫХ
ТЕЛЕКОММУНИКАЦИОННЫХ СЕТЕЙ**

Специальность: 05.12.02 – Телекоммуникационные системы и сети

ДИССЕРТАЦИЯ НА СОИСКАНИЕ УЧЕНОЙ СТЕПЕНИ
КАНДИДАТА ТЕХНИЧЕСКИХ НАУК

Научный руководитель
ЛЕМЕШКО Александр Витальевич,
доктор технических наук, профессор

Харьков – 2015

СОДЕРЖАНИЕ

ПЕРЕЧЕНЬ УСЛОВНЫХ ОБОЗНАЧЕНИЙ	5
ВВЕДЕНИЕ.....	7
РАЗДЕЛ 1. ОБЗОР ТЕОРЕТИЧЕСКИХ И ТЕХНОЛОГИЧЕСКИХ РЕШЕНИЙ В ОБЛАСТИ УПРАВЛЕНИЯ ОЧЕРЕДЯМИ НА МАРШРУТИЗАТОРАХ ТРАНСПОРТНОЙ ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ	17
1.1. Место задач управления очередями в архитектуре обеспечения качества обслуживания сетей следующего поколения	17
1.1.1. Анализ общей архитектуры сети следующего поколения	17
1.1.2. Классификация основных показателей качества обслуживания и сетевых ресурсов.....	19
1.1.3. Общая характеристика средств обеспечения качества обслуживания на основе управления сетевым ресурсом.....	25
1.2. Анализ известных механизмов управления очередями на маршрутизаторах ТКС.....	30
1.3. Обзор известных математических моделей и методов управления очередями на маршрутизаторах ТКС.....	39
1.4. Формулировка требований к перспективным решениям в области управления очередями на маршрутизаторах ТКС. Постановка научной задачи.....	44
1.5. Выводы по первому разделу.....	47
РАЗДЕЛ 2. РАЗРАБОТКА КОМПЛЕКСНОЙ МАТЕМАТИЧЕСКОЙ МОДЕЛИ УПРАВЛЕНИЯ ОЧЕРЕДЯМИ НА МАРШРУТИЗАТОРАХ ТРАНСПОРТНОЙ ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ.....	50
2.1. Вводимые в рамках предлагаемой модели условные обозначения....	51
2.2. Формализация условий сохранения потока и предотвращения перегрузки очередей на интерфейсе маршрутизатора	56

2.3. Выбор критерия оптимальности управления очередями на интерфейсах маршрутизаторов транспортной телекоммуникационной сети.....	61
2.4. Обзор методов решения сформулированной оптимизационной задачи управления очередями в рамках предложенной модели.....	63
2.5. Методика сведения задачи целочисленной оптимизации к задаче нелинейного программирования	68
2.6. Исследование процессов управления очередями на маршрутизаторах транспортной телекоммуникационной сети с использованием предложенной модели.....	69
2.6.1. Примеры решения задач управления очередями на маршрутизаторах ТКС.....	70
2.6.2. Анализ влияния характеристик потоков и очередей на результат решения задачи по управлению очередями в рамках предложенной модели.....	74
2.7. Выводы по второму разделу.....	77
РАЗДЕЛ 3. РАЗРАБОТКА ПОТОКОВОЙ МОДЕЛИ УПРАВЛЕНИЯ ОЧЕРЕДЯМИ НА ИНТЕРФЕЙСАХ МАРШРУТИЗАТОРОВ АКТИВНОЙ СЕТИ.....	80
3.1. Анализ функциональных особенностей построения активной сети..	80
3.1.1. Обзор целей и задач активных сетей.....	80
3.1.2. Анализ архитектуры активного узла.....	83
3.2. Математическая модель управления очередями на интерфейсах маршрутизаторов активной сети.....	86
3.3. Исследование процессов управления очередями на интерфейсах активных узлов с использованием предложенной модели.....	93
3.3.1. Пример решения задачи управления очередями на интерфейсах активных узлов с обеспечением сбалансированной загруженности очередей по их длине.....	94

3.3.2. Пример решения задачи управления очередями на интерфейсах активных узлов с обеспечением сбалансированной загрузки активных процессоров и очередей.....	96
3.4. Модификация модели управления очередями на интерфейсе активного узла при организации конвейера активных процессоров	99
3.5. Выводы по третьему разделу.....	104
РАЗДЕЛ 4. РАЗРАБОТКА И ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ МЕТОДА УПРАВЛЕНИЯ ОЧЕРЕДЯМИ НА ИНТЕРФЕЙСАХ МАРШРУТИЗАТОРОВ ТРАНСПОРТНОЙ ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ.....	
4.1. Описание метода управления очередями на интерфейсах маршрутизаторов транспортной телекоммуникационной сети.....	108
4.2. Исследование предложенного метода управления очередями на интерфейсах маршрутизаторов транспортной ТКС.....	112
4.2.1. Описание общей схемы проводимого лабораторного эксперимента.....	113
4.2.2. Оценка эффективности процесса управления очередями, организованного в рамках предложенного метода.....	122
4.3. Разработка рекомендаций по практической реализации предложенных моделей и метода управления очередями на маршрутизаторах транспортной телекоммуникационной сети	129
4.4. Выводы по четвертому разделу	134
ВЫВОДЫ ПО РАБОТЕ.....	139
СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ.....	143
АКТЫ ВНЕДРЕНИЯ.....	153
ПРИЛОЖЕНИЕ А. Акт внедрения в учебный процесс ХНУРЭ.....	154
ПРИЛОЖЕНИЕ Б. Акт внедрения в ООО «Сигма СОФТВЕА».....	155

ПЕРЕЧЕНЬ УСЛОВНЫХ ОБОЗНАЧЕНИЙ

ACL	- Access Control List
AN	- Active Network
APCI	- Active Processing Control Information
AQM	- Active Queue Management
ATM	- Asynchronous Transfer Mode
CBQ	- Class-Based Queuing
CBWFQ	- Class-Based Distributed WFQ
CLI	- Command Line Interface
CoS	- Class of Service
CPU	- Central Processing Unit
CQ	- Custom Queuing
DiffServ	- Differentiated Services
D-ITG	- Distributed Internet Traffic Generator
DSCP	- Differentiated Services Code Point
DWFQ	- Flow-Based Distributed WFQ
ECN	- Explicit Congestion Notification
EIGRP	- Extended IGRP
FIFO	- First In, First Out
FQ	- Fair Queuing
IGRP	- Interior Gateway Routing Protocol
IntServ	- Integrated Services Architecture
IP	- Internet Protocol
IPDV	- IP packet delay variation
IPTD	- IP packet transfer delay
IS-IS	- Intermediate System – to – Intermediate System
ITU-T	- International Telecommunications Union-Telecommunications
LLQ	- Low Latency Queuing
MPLS	- MultiProtocol Label Switching

NGN	- Next Generation Network
OSPF	- Open Shortest Path First
PHB	- Per Hop Behavior
PNNI	- Private Network-to-Network Interface
PQ	- Priority Queuing
QoS	- Quality of Service
RED	- Random Early Detection
RIP	- Routing Internet Protocol
RSVP	- Resource Reservation Protocol
SLA	- Service Level Agreement
SPD	- Selective Packet Discard
TCP	- Transmission Control Protocol
TE	- Traffic Engineering
ToS	- Type of Service
WAN	- Wide Area Network
WFQ	- Weighted Fair Queuing
WRED	- Weighted Random Early Detection
АП	- Активный процессор
КС	- Канал связи
ПС	- Пропускная способность
ТКС	- Телекоммуникационная сеть
ТМО	- Теория массового обслуживания
ЭМВОС	- Эталонная модель взаимодействия открытых систем

ВВЕДЕНИЕ

Актуальность темы. Текущий этап развития современного общества характеризуется все более широким внедрением инфокоммуникационных систем и технологий. Эти решения играют ключевую роль при внедрении важных для каждого конкретного человека систем жизнеобеспечения: телемедицины, электронной торговли, городских информационных служб, интеллектуальных систем транспорта, систем локального и глобального позиционирования, электронных библиотек, музеев и т.д.

Системообразующим каркасом инфокоммуникационных решений традиционно являются телекоммуникационные системы и сети, которые в силу отмеченных выше особенностей должны быть мультисервисными, т.е. обеспечивать предоставление одновременно множества сервисов на базе одной транспортной платформы. Именно этим критериям отвечают телекоммуникационные сети (ТКС), построенные в соответствии с рекомендациями ИТУ-Т [1, 2] на принципах сетей следующего поколения (Next Generation Network, NGN).

При выборе транспортной телекоммуникационной технологии (IP, ATM, MPLS/GMPLS) [3-5] в целом и конкретных сетевых протоколов в частности, на которых основывается NGN, важным является учет их возможностей по поддержке и обеспечению качества обслуживания (Quality of Service, QoS). Ввиду особой востребованности сквозного (end-to-end) QoS на первое место в архитектуре управления трафиком выходят решения именно сетевого уровня эталонной модели взаимодействия открытых систем (ЭМВОС). Среди процессов сетевого уровня важную роль играют задачи управления очередями, т.к. именно неэффективное управление очередями приводит к неконтролируемому росту задержек и уровня потерь пакетов.

Как показал проведенный анализ [6, 7], в современном коммутационном оборудовании реализовано множество механизмов

управления очередями как с точки зрения их формирования и обслуживания (FIFO, PQ, CQ, FQ/WFQ, CBQ, LLQ), так и предотвращения перегрузки (RED, WRED, ECN, SPD). Их отличительной особенностью и ключевым недостатком является преобладание ручных настроек в ходе конфигурации оборудования, что не позволяет оперативно реагировать на изменение состояния (загруженности) интерфейсов маршрутизатора и сети в целом, а также на вариацию характеристик трафика. В этой связи множество отечественных и зарубежных ученых активно работают над совершенствованием средств управления трафиком и обеспечения качества обслуживания в ТКС, в частности и в направлении пересмотра моделей, методов и самих механизмов управления очередями на маршрутизаторах сети. К их числу стоит отнести, прежде всего, Беркман Л.Н., Поповского В.В., Лемешко А.В., Романова А.И., Ложковского А.Г., Климаша М.Н., Польщикова К.А., Stein С., Burns А., Ling Wang, Yihan Li, Shivendra Panwar и многих других.

Перспективным направлением повышения эффективности управления очередями является обеспечение согласованности решений отдельных, но взаимосвязанных между собой интерфейсных задач. К их числу, прежде всего, относятся классификация и маркировка пакетов; создание и настройка системы очередей на интерфейсе; агрегирование потоков и распределение пакетов по очередям интерфейса с учетом параметров передаваемых потоков, требований к качеству обслуживания, характеристик создаваемых очередей и интерфейса в целом; определение порядка обслуживания очередей, т.е. установление очередности передачи пакетов из очередей в канал связи; распределение пропускной способности интерфейса между отдельными очередями; превентивное (заблаговременное) ограничение длины очереди. Кроме этого первоочередного требования новые решения в этой области должны удовлетворять следующим требованиям:

- учет потоковой структуры современного сетевого трафика, обусловленного ростом общего числа мультимедиа приложений в перечне инфокоммуникационных услуг;

- оптимизационная постановка и решение задачи управления очередями, связанное с необходимостью сбалансированного использования доступного сетевого (буферного и канального) ресурса;

- поддержка дифференциации обслуживания пакетов на интерфейсах маршрутизаторов ТКС в соответствии с их QoS-требованиями;

- реализация динамических стратегий управления очередями;

- обеспечение справедливости обслуживания пакетов одного и того же потока;

- поддержка распределенных решений по управлению очередями;

- простота алгоритмически-программной и аппаратной реализации.

В этой связи актуальной представляется *научная задача*, связанная с оптимизацией процесса управления очередями на интерфейсе маршрутизаторов на основе обеспечения согласованного решения задач по агрегированию потоков, распределения пакетов между очередями и выделения очередям пропускной способности интерфейса путем разработки соответствующих математических моделей и метода для повышения качества обслуживания в транспортных ТКС в целом.

Связь работы с научными программами, планами и темами.

Диссертационная работа выполнена согласно требованиям положений «Концепции конвергенции телефонных сетей и сетей с пакетной коммутацией в Украине», «Основных принципов развития информационного общества в Украине на 2007 – 2015 года», «Концепции национальной информационной политики» и «Концепции Национальной программы информатизации». Результаты диссертации использованы при модернизации и развитии телекоммуникационной системы ООО «Сигма СОФТВЕА».

Цель и задачи исследования. Цель диссертационной работы состоит в повышении качества обслуживания в транспортной ТКС на основе усовершенствования математических моделей и методов управления очередями на интерфейсах ее маршрутизаторов. Для решения поставленной научной задачи в диссертационной работе решены следующие частные задачи исследования:

- анализ протокольных решений по управлению трафиком и очередями на интерфейсах маршрутизаторов транспортных ТКС;
- обзор известных теоретических решений, математических моделей и методов управления очередями на интерфейсах маршрутизаторов транспортных ТКС;
- формулирование требований к перспективным решениям в области управления очередями на маршрутизаторах ТКС;
- усовершенствование потоковых моделей и метода управления очередями на интерфейсах маршрутизаторов транспортных ТКС;
- анализ адекватности и исследование эффективности предложенных в диссертации решений по управлению очередями;
- разработка рекомендаций по практическому использованию полученных в работе результатов в современных и перспективных транспортных ТКС.

Объект исследования – процессы управления очередями на интерфейсах маршрутизаторов транспортных телекоммуникационных сетей.

Предмет исследования – математические модели и методы управления очередями на интерфейсах маршрутизаторов транспортных телекоммуникационных сетей.

Методы исследования. В диссертации нашла свое применение система аналитических и имитационных методов исследования, а также методы лабораторного эксперимента. При этом в ходе

усовершенствования моделей управления очередями на интерфейсах маршрутизаторов ТКС использованы возможности теории массового обслуживания и аппарата математического программирования. В ходе сравнительного анализа предложенных и ранее известных решений по управлению очередями использованы средства имитационного моделирования. Для подтверждения адекватности предложенных моделей и метода, а также для проверки достоверности получаемых результатов в работе проводился лабораторный эксперимент.

Научная новизна полученных результатов.

1. Получила дальнейшее развитие потоковая модель управления очередями на интерфейсах маршрутизаторов транспортных телекоммуникационных сетей. Новизна модели заключается в том, что в ней в ходе управления очередями осуществляется оптимальное агрегирование и распределение потоков на основании сравнения и минимизации расхождения классов потоков и очередей, в которые они направляются. Использование модели позволило в большинстве случаев сократить число поддерживаемых очередей на интерфейсе без снижения уровня дифференциации качества обслуживания, что способствовало минимизации времени на обработку пакетов на интерфейсе и снижению их межконцевой задержки в сети в целом.

2. Получила дальнейшее развитие модель управления очередями на интерфейсах маршрутизаторов активных сетей. Новизна модели заключается в том, что в ходе обслуживания пакетов учитываются не только количество и параметры потоков и очередей, но и характеристики активных процессоров (тип, загруженность и производительность), входящих в архитектуру маршрутизатора активной сети. Реализация предложенной модели позволила обеспечить более эффективное использование буферного и вычислительного ресурса маршрутизатора активных сетей.

3. Усовершенствован метод управления очередями на интерфейсах

маршрутизаторов транспортных телекоммуникационных сетей. Новизна метода заключается в том, что с его помощью оптимизируется работа интерфейса в ходе согласованного решения задач по агрегированию потоков, распределению пакетов по очередям и выделения очередям необходимой пропускной способности интерфейса, а также задач по задействованию и контролю использования активных процессоров маршрутизатора.

Достоверность и обоснованность полученных новых научных результатов обеспечивалась и подтверждалась корректным применением основных положений известного и хорошо апробированного математического аппарата, тщательным аналитическим и числовым обоснованием принятых приближений и допущений, а также четкой физической интерпретацией полученных результатов расчета. Кроме того, достоверность результатов работы определялась допустимым расхождением результатов аналитического моделирования и проведенного в работе лабораторного эксперимента.

Научное значение результатов работы заключается в обобщении и дальнейшем развитии математических моделей и методов управления очередями на интерфейсах маршрутизаторов и трафиком в ТКС в целом. Использование предложенных моделей позволит придать динамический характер решениям по управлению сетевым (канальным и буферным ресурсом), которые должны прийти на смену существующим квазистатическим схемам обработки пакетов.

Практическая значимость полученных результатов заключается в возможности непосредственного использования предложенных потоковых моделей и метода при внедрении перспективных сетевых технологий и отдельных механизмов управления очередями на маршрутизаторах как обычных (традиционных), так и активных ТКС. Разработанные модели и метод получили свою программную реализацию в среде MatLab и апробированы в ходе

лабораторного эксперимента на реальном сетевом оборудовании транспортных ТКС. Кроме того, результаты диссертации использованы при модернизации и развитии телекоммуникационной системы ООО «Сигма СОФТВЕА» и в учебном процессе кафедры телекоммуникационных систем ХНУРЭ в дисциплине «Системы коммутации и распределения информации, часть 2». По результатам диссертационных исследований также получены два патента на полезную модель (47461, № u200905821 и 70088, № u201113675). По всем реализациям результатов работы составлены необходимые акты.

Личный вклад соискателя. Все основные научные результаты, изложенные в диссертационной работе, соискатель получил самостоятельно. Кроме того, в работе [8] автором предложена потоковая модель управления очередями на интерфейсе маршрутизатора активной сети; в статье [9] соискателем проведено сравнение размерности оптимизационных задач по управлению очередями при статическом и динамическом распределении пропускной способности интерфейса в зависимости от количества очередей и числа приоритетов пакетов; в публикации [10] автором предложена потоковая модель управления очередями с динамическим распределением пропускной способности интерфейса; в работе [11] автором предложена потоковая модель управления очередями с введением условий по недопущению их перегрузки не только по пропускной способности, но и по длине; в статье [12] соискатель сформулировал задачу управления очередями в виде задачи оптимальной балансировки их длины; в публикации [13] автором осуществлен анализ известных механизмов по управлению очередями и сформулированы требования, предъявляемые к перспективным решениям в этой области; в работах [14, 15] соискатель выполнил исследование метода управления очередями, базировавшегося на использовании различных по своей сложности и функциональности математических

моделей; в статье [16] автором проанализированы вопросы взаимодействия процессов маршрутизации и балансировки очередей на интерфейсах маршрутизаторов ТКС; в работе [17] соискатель рассмотрел возможности реализации активного управления очередями с помощью потоковых моделей; в статье [18] автором предложена математическая модель управления очередями на маршрутизаторах телекоммуникационной сети на основе оптимального агрегирования потоков и распределения пакетов по очередям.

Апробация основных положений диссертационной работы проводилась в ходе двадцати двух научных конференций, форумов и семинаров [19-40], среди которых:

– VII и VIII Международные научно-технические конференции "Перспективные технологии в средствах передачи информации", 2007 г., 2009 г. (Владимир, ВГУ);

– второй научно-практический семинар молодых учёных и студентов "Сучасні телекомунікаційні та інформаційні технології", 12-14 декабря 2007 г. (Киев, УНИИС);

– четырнадцатая Международная научно-техническая конференция студентов и аспирантов "Радиоэлектроника, электротехника и энергетика", 28-29 февраля 2008 г. (Москва, МЭИ);

– вторая научно-техническая конференция "Проблеми телекомунікацій", 20-23 мая 2008 г. (Киев, НТУУ "КПИ");

– XII и XIII Международные молодежные форумы "Радиоэлектроника и молодежь в XXI веке", 2008 г., 2009 г. (Харьков, ХНУРЭ);

– третий и четвертый Международные радиоэлектронные форумы "Прикладная радиоэлектроника. Состояние и перспективы развития", 2008 г., 2011 г. (Харьков: АНПРЭ, ХНУРЭ);

- с 5-й по 9-ю научные конференции Харьковского университета Воздушных Сил им. Ивана Кожедуба "Новітні технології – для захисту повітряного простору", 2009-2013 гг. (Харьков, ХУ ВС);
- научно-практическая конференция «Сучасні проблеми телекомунікацій», 28-30 октября 2010 г., (Львов, НУ "Львівська політехніка");
- научно-техническая конференция с международным участием "Компьютерное моделирование в наукоемких технологиях" (КМНТ-2010), Харьковского национального университета имени В.Н. Каразина, 18-21 мая 2010 г. (Харьков, ХНУ);
- V Международная научно-техническая конференция "Сучасні проблеми і досягнення в галузі радіотехніки, телекомунікацій та інформаційних технологій", 22-24 сентября 2010 г. (Запорожье, ЗНТУ);
- научно-практический семинар молодых учёных и студентов "Сучасні телекомунікаційні та інформаційні технології", №1, 2010 г. (Киев, УНИИС);
- восьмая Международная научно-техническая конференция "Сучасні інформаційно-комунікаційні технології" – COMINFO'2012 – Livadia, 1-5 октября 2012 г., (Крым-Ялта-Ливадия: ГУИКТ);
- 23-я Международная Крымская конференция «СВЧ-техника и телекоммуникационные технологии», 8-13 сентября 2013 г. (Севастополь, Севастопольский национальный технический университет);
- 12-я и 15-я Международные научно-технические конференции «The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM'2013)», 2013 г., 2015 г., (Львов, Национальный университет "Львівська політехніка").

Публикации. Основные положения диссертации отражены в 35 научных работах. Среди работ 11 статей, из которых восемь в научных специализированных изданиях, утвержденных МОН Украины [8-15, 18],

три статьи [15-17] в иностранных изданиях телекоммуникационной направленности, входящих в наукометрические базы e-library (РИНЦ) и Google Scholar. Кроме того, материалы диссертации опубликованы в 22 материалах и тезисах докладов на научно-технических конференциях и форумах [19-40], из которых три конференции проходили под эгидой IEEE [37, 39, 40] и выложены в наукометрических базах Scopus и IEEE Xplore Digital Library, а три конференции проходили за рубежом [19, 21, 27]. Соискатель также имеет два патента на полезную модель [41, 42].

РАЗДЕЛ 1

ОБЗОР ТЕОРЕТИЧЕСКИХ И ТЕХНОЛОГИЧЕСКИХ РЕШЕНИЙ В ОБЛАСТИ УПРАВЛЕНИЯ ОЧЕРЕДЯМИ НА МАРШРУТИЗАТОРАХ ТРАНСПОРТНОЙ ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ

1.1. Место задач управления очередями в архитектуре обеспечения качества обслуживания сетей следующего поколения

1.1.1. Анализ общей архитектуры сети следующего поколения

Переход к информационному обществу и связанное с этим развитие и совершенствование соответствующих инфокоммуникационных платформ и технологий привело к созданию мультисервисных телекоммуникационных систем и сетей, способных обеспечить передачу разнородной информации (мультимедиа, данные и др.) с заданными показателями качества обслуживания [2-6]. В этой связи на рынке связи постоянно обновляются образцы коммутационного, серверного и оконечного оборудования, непрерывно усовершенствуется программное обеспечение (ПО) сетевого оборудования и отдельных терминалов.

Расширение номенклатуры предоставляемых инфокоммуникационных услуг обусловило появление и практическую реализацию концепции построения сетей следующего поколения (Next Generation Network, NGN). Отличительной чертой NGN решений является разделение функций предоставления сервисов и транспортировки контента [1, 2]. В рамках рекомендаций ITU-T NGN определена как сеть пакетной коммутации, которая способна предоставлять телекоммуникационные услуги с возможностью использования нескольких высокоскоростных транспортных технологий, обеспечивающих качество обслуживания и независимость функций предоставления услуг и транспортировки пакетов [2]. Одной из

главных задач, которые должна выполнять NGN, является обеспечение качества обслуживания при минимальном использовании (загруженности) сетевых ресурсов, что должно повлечь удешевление стоимости предоставляемых услуг связи в целом, а также способствовать привлечению большего числа пользователей [3].

В общем случае архитектура NGN [2-4] включает в себя следующие иерархические уровни (рис. 1.1):

- уровень управления услугами;
- уровень транспорта;
- уровень доступа.

Уровень управления услугами включает в себя множество серверов услуг и управления сетью для обеспечения эффективного предоставления разнородных сервисов пользователям ТКС. Именно на него возложено решение задач по поддержке существующего перечня услуг, а также по внедрению новых сервисов с минимальными затратами и в минимальные сроки. *Уровень транспорта* функционирует на основе таких технологий территориально-распределенных сетей (Wide Area Network, WAN), как IP (Internet Protocol), MPLS (MultiProtocol Label Switching) и, все же, ATM (Asynchrony Transfer Mode) [4, 5, 43, 44]. Его основной задачей является организация информационного взаимодействия между сетями доступа, к которым подключены непосредственно терминалы пользователей. Технологии этого уровня используют преимущественно режимы пакетной коммутации и оснащаются средствами поддержки сквозного QoS. *Уровень доступа* предназначен для получения конечными пользователями запрашиваемых услуг и их взаимодействия с использованием возможностей и средств транспортного уровня NGN.

Описанная выше градация по уровням (рис.1.1) позволяет реализовать на практике мультисервисные решения с поддержкой различных типов качества обслуживания.

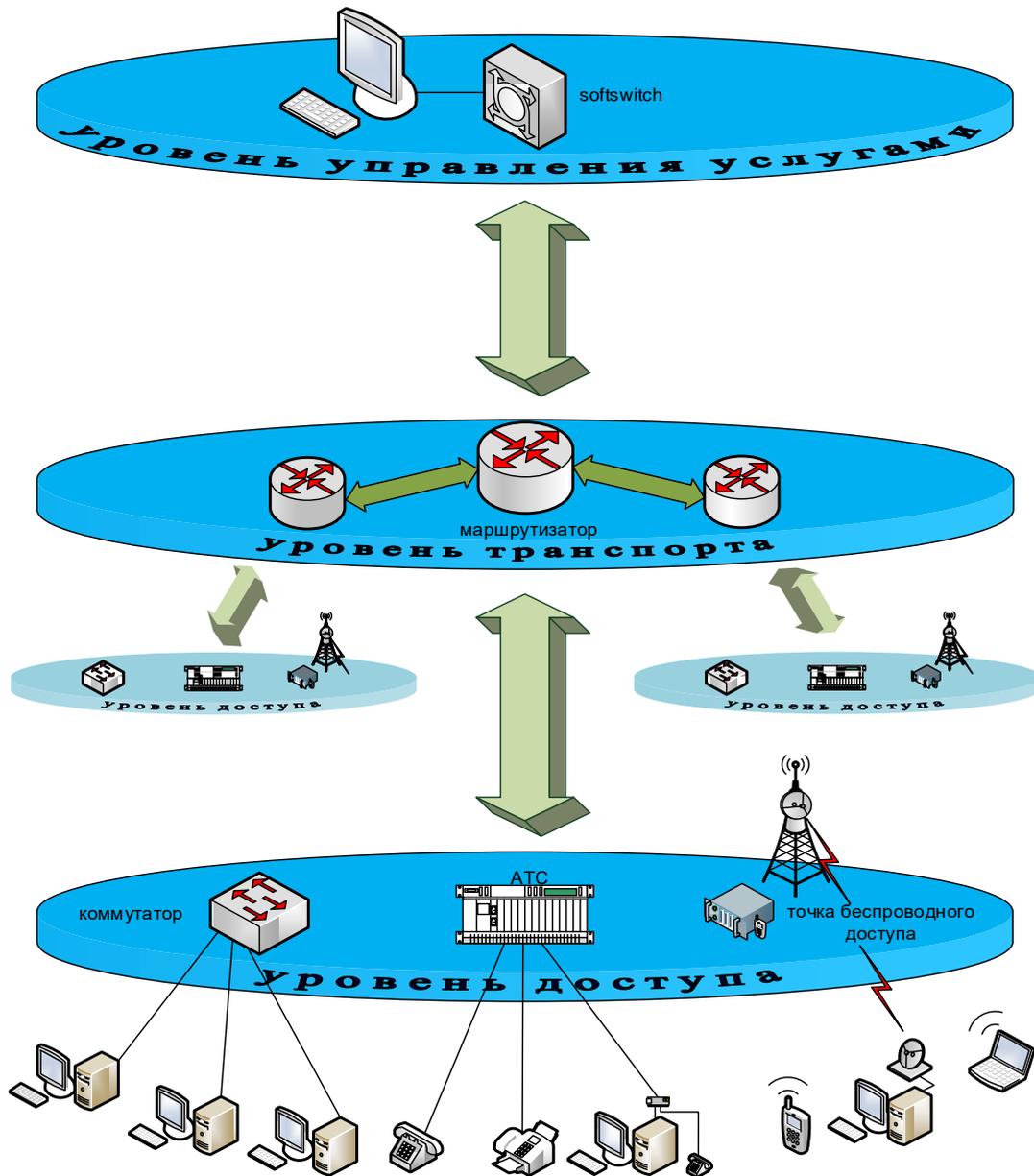


Рис.1.1. Вариант архитектуры сети следующего поколения

1.1.2. Классификация основных показателей качества обслуживания и сетевых ресурсов

В современных IP-сетях, реализующих принципы построения NGN, выделяют следующие основные типы качества обслуживания, используемые для поддержки различных сетевых сервисов (рис. 1.2):

- best effort service, обслуживание по возможности, при котором никаких гарантий предоставления сервиса не предоставляется;
- сервис с предпочтением или дифференцированное обслуживание (Differentiated Service, DiffServ), при котором степень удовлетворения QoS-требований зависит от класса (приоритета) трафика;
- гарантированный сервис (Guaranteed Service) или интегрированное обслуживание (Integrated Service, IntServ), при котором обеспечиваются гарантии в удовлетворении QoS-требований на основе резервирования сетевых ресурсов.

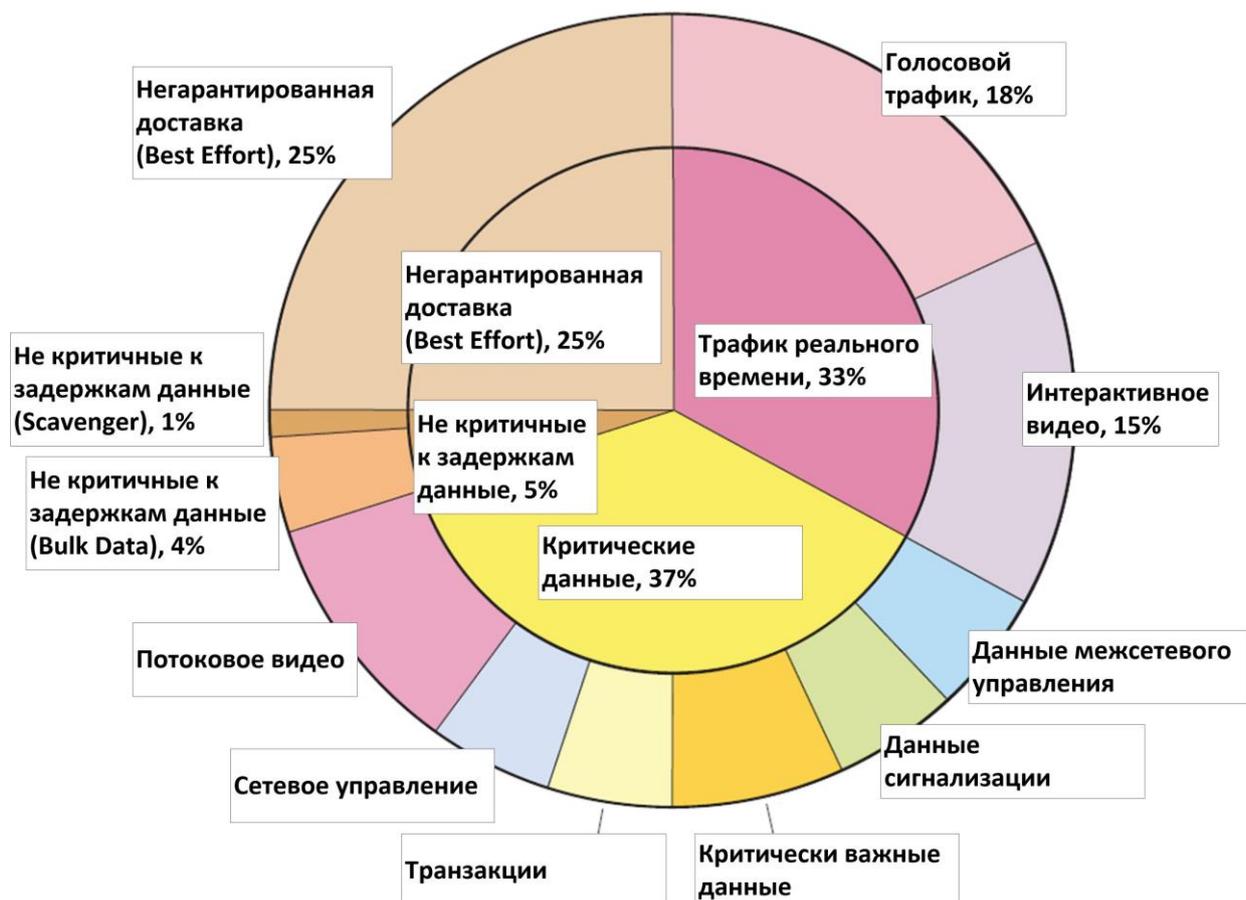


Рис. 1.2. Структура современного сетевого трафика с точки зрения состава сервисных потоков и приложений

Результаты анализа структуры современного сетевого трафика с точки зрения состава сервисных потоков и приложений [4] позволили сделать

вывод о том, что число мультимедиа контента в ТКС постоянно растет, как и увеличивается число приложений, для которых необходимо обеспечить гарантии по качеству обслуживанию.

С точки зрения характеристики основных сетевых ресурсов стоит отметить, что в общем случае имеет место их классификация, представленная на рис. 1.3.



Рис. 1.3. Классификация сетевых ресурсов

К аппаратным ресурсам, прежде всего относят:

- пропускную способность интерфейсов маршрутизаторов, которая может достигать в перспективных решениях десятки терабит за секунду (Тбит/с);

- производительность самого маршрутизатора, которая выросла от десятков пакетов за секунду (70-е годы 20-го века) до миллионов/миллиардов пакетов за секунду (начало 21 века);

- буфер очереди маршрутизатора, в котором ожидают своего обслуживания пакеты при перегрузке канала связи. Его размер варьируется в пределах десятков пакетов, с одной стороны, для минимизации вероятности потерь пакетов, а с другой – для управляемости величиной средней задержки пакетов.

Программные ресурсы – это, прежде всего, программное обеспечение коммутационного оборудования (коммутаторов и маршрутизаторов), серверов услуг и терминалов пользователей. Все управляющие протоколы и механизмы фактически представляют собой часть специального программного обеспечения маршрутизатора или коммутатора. Сетевой трафик, содержание маршрутных таблиц и баз данных о состоянии ТКС – это информационный ресурс, который также перераспределяется с целью обеспечения сбалансированной загруженности ТКС.

Эффективное управление перечисленным сетевым ресурсом, особенно с точки зрения его перераспределения в интересах потоков того или иного приложения или сервиса, направлено, прежде всего, на повышение качества обслуживания основных показателей QoS (рис. 1.4). Основными показателями QoS являются [4-6] показатели производительности (скоростные показатели), показатели временной прозрачности (временные показатели) и показатели семантической прозрачности (показатели надежности и достоверности).

К основным показателям производительности (скоростным показателям) сети обычно относят минимальную, максимальную и среднюю скорость передачи пакетов, измеряемую, как правило, в бит/с (кбит/с. Мбит/с). Эти значения, как правило, указываются в соглашении об уровне сервиса (Service Level Agreement, SLA), заключаемым между пользователем и провайдером. В этом же документе могут прописываться и требуемые

значения временных показателей – средней задержки (IP packet transfer delay, IPTD) и джиттера пакетов (IP packet delay variation, IPDV), но уже при предоставлении мультимедиауслуг.



Рис. 1.4. Классификация показателей QoS

В рекомендациях ITU-T Y.1540 и Y.1541 указаны ориентировочные значения для межконцевых временных показателей, которые могут варьироваться от десятков до сотен миллисекунд в зависимости от класса

потока, подлежащего обслуживанию. К показателям надежности традиционно относят вероятность потерь пакетов или, например, вероятность доставки пакета не по адресу. Как показано в табл. 1.1 современные сетевые приложения по-разному чувствительны к изменению численных значений основных показателей качества обслуживания.

Таблица 1.1

Характер требований сетевых приложений к уровню QoS

Тип приложения	Надежность	Средняя задержка	Джиттер	Скорость передачи
Электронная почта	высокие	низкие	низкие	низкие
Передача файлов	высокие	низкие	низкие	средние
Web доступ	высокие	средние	низкие	средние
Аудио по требованию	низкие	низкие	высокие	средние
Видео по требованию	низкие	низкие	высокие	высокие
Телефония	низкие	высокие	высокие	низкие
Видеоконференция	низкие	высокие	высокие	высокие

В этой связи потокам (сообщениям, пакетам) различных приложений необходимо выделять различные объемы сетевых ресурсов, чтобы обеспечить дифференциацию уровня обслуживания. В современных ТКС реализовано множество технологических средств (протоколов, механизмов) обеспечения качества обслуживания, основанных на управлении доступным сетевым ресурсом и трафиком в частности.

1.1.3. Общая характеристика средств обеспечения качества обслуживания на основе управления сетевым ресурсом

Уровень качества обслуживания, обеспечиваемый той или иной телекоммуникационной сетью, во многом определяется как объемом доступного сетевого ресурса, так и эффективностью его распределения. Поэтому для обеспечения QoS в ТКС задействуются возможности всех уровней эталонной модели взаимодействия открытых систем. При этом физический и канальный уровни непосредственно участвуют в формировании сетевого ресурса, который в дальнейшем необходимо грамотно распределить уже средствами сетевого и транспортного уровней.

На рис. 1.5 перечислены основные средства сетевого уровня ЭМВОС, непосредственно участвующие в распределении доступного канального и буферного ресурса.

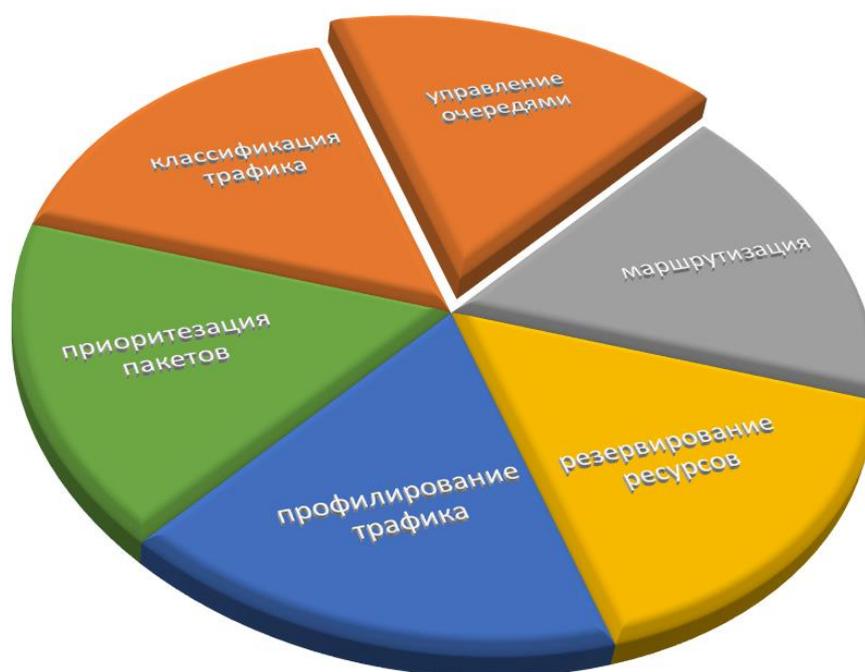


Рис. 1.5 Основные средства сетевого уровня ЭМВОС, участвующие в распределении доступного канального и буферного ресурса

Основой обеспечения качества обслуживания в IP и MPLS-сетях являются процессы классификации и маркировки (приоритезации) пакетов. Лишь эффективное решение данных задач, заключающееся в корректном определении типа (класса) передаваемого контента и присвоении соответствующего приоритета), позволит в дальнейшем остальным средствам управления трафиком обеспечить должное распределение сетевого ресурса в интересах данных потоков.

Задача классификации состоит в определении класса (типа) передаваемого трафика путем анализа содержимого заголовка IP-пакета и/или характеристик потока пакетов в целом. Идентификация класса потока, как правило, основывается на анализе множества параметров, к которым, прежде всего, относят следующие

- IP-адрес источника пакета;
- IP-адрес получателя пакета
- MAC-адреса источника и получателя пакетов;
- стандартный или расширенный список доступа к источникам/адресатам;
- номера портов источника или получателя пакетов;
- содержание полей приоритета IP-пакета: DSCP или IP Precedence;
- код длины пакета;
- 0-99 QoS-group;
- содержание поля CoS.

Приоритезация – основа модели DiffServ, поэтому ошибки при классификации трафика, приведя к неправильному присвоению приоритета, в дальнейшем заметно снизят эффективность распределения сетевого ресурса и обеспечения качества обслуживания. Приоритет пакета записывается в байте типа обслуживания (Type of Service, ToS) в заголовке пакета IPv4 или байте класса трафика (Traffic Class) пакета IPv6. За это отвечают поля DSCP (Differentiated Services Code Point) и IP Precedence (рис. 1.6).



Рис. 1.6. Размещение полей DSCP и IP Precedence в заголовке IP-пакета

Маркировка пакетов может осуществляться на приграничном маршрутизаторе транспортной сети (чаще всего), на коммутаторе с поддержкой функций 3 уровня или на «интеллектуальном» терминальном устройстве, например, на VoIP-телефоне.

Протоколы маршрутизации, основными из которых в IP-сетях являются RIP (Routing Information Protocol), IGRP (Interior Gateway Routing Protocol), OSPF (Open Shortest Path First), IS-IS (Intermediate System to Intermediate System), BGP (Border Gateway Protocol), EGP (Exterior Gateway Protocol) и I-PNNI (Integrated Private Network to Network Interface) [69], отвечают за расчет одного или множества маршрутов между заданной парой (парами) маршрутизаторов ТКС. Подобные протоколы, основываясь на постоянном мониторинге состояния сети (ее топологии и загруженности), формируют маршрутные таблицы, в соответствии с которыми потоки или отдельные пакеты автоматически перенаправляются к получателям.

Положительной чертой маршрутных протоколов является то, что сбор, обработка информации о состоянии сети и формирование маршрутных таблиц (MT) осуществляется при минимальном вмешательстве

администратора сети, чего не скажешь о большинстве других средств управления трафиком. В функции администратора входит корректировка таймеров обновления и констант (variance), отвечающих за балансировку нагрузки. Недостатком большинства протоколов маршрутизации остается использование достаточно примитивных по своим возможностям алгоритмов Дейкстры и Беллмана-Форда расчета кратчайшего пути на графе, которые ориентируют на реализацию однопутевой стратегии маршрутизации и не производят учет характеристик передаваемых потоков пакетов.

Перспективные решения в области маршрутизации, например, протокол CSPF (Constrained Shortest Path First), позволяет учитывать не только топологию ТКС, но и загруженности каналов связи и буферов очередей на ее маршрутизаторах. Это обеспечивается на этапе формирования маршрутных метрик, которые должны учитывать как номинальную, так и доступную пропускную способность каналов связи. Возможность использования композитных метрик открывает потенциал по учету множества показателей качества обслуживания при расчете искомого маршрута.

Важный отпечаток на развитие большинства средств управления трафиком, обеспечения качества обслуживания в целом и протоколов маршрутизации в частности наложила технология Traffic Engineering (TE) [43, 44]. Ее целью является обеспечение сбалансированной загруженности сетевых ресурсов (канальных, буферных, вычислительных) в ходе использования рассчитанного множества путей, т.к. балансировка нагрузки предполагает реализацию многопутевой стратегии маршрутизации. Практическое использование технологии Traffic Engineering несомненно приводит к усложнению маршрутных решений, т.к. необходимо рассчитывать и поддерживать в ходе управления не один, а множество маршрутов. Но это компенсируется улучшением качества обслуживания в ТКС, особенно при высокой нагрузке на сеть.

Стоит отметить, что протоколы маршрутизации осуществляют расчет маршрутных таблиц и их использование, а вот управление трафиком вдоль выбранного маршрута уже возлагается на средства управления очередями и резервирования ресурсов на каждом конкретном маршрутизаторе (рис. 1.7).

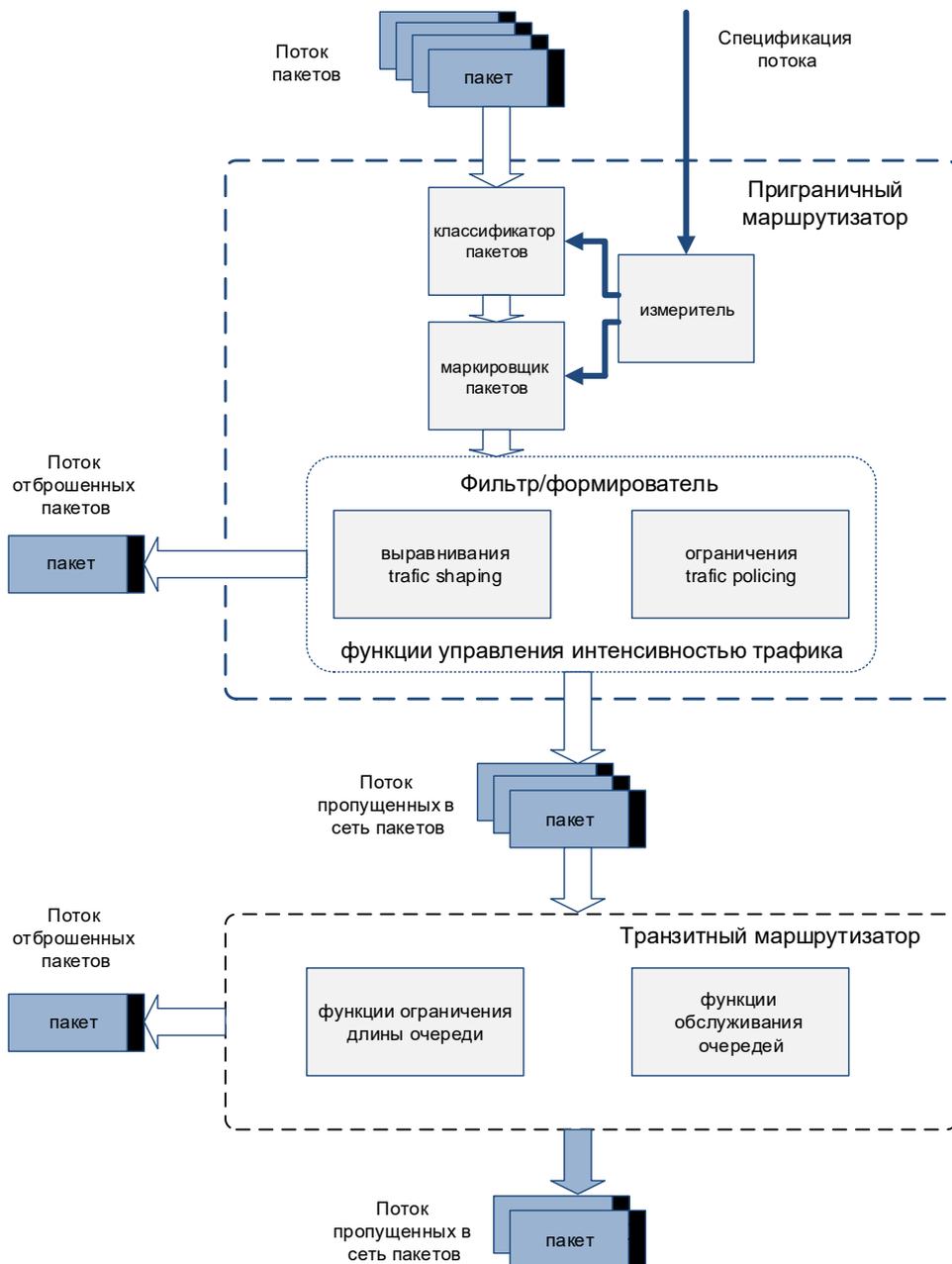


Рис. 1.7. Общая схема обработки пакетов на маршрутизаторах ТКС

Как правило, на приграничных маршрутизаторах выполняется очень важная функция, связанная с профилированием трафика. Она реализуется с

помощью механизмов Traffic Shaping/Policing, основанных на алгоритмах token bucket и leaky bucket, и выполняется с целью ограничения скорости поступления пакетов в сеть ввиду возможной перегрузки или нарушения SLA-договора.

Дополняют функционал средств обеспечения качества обслуживания особенно в условиях возможной перегрузки механизмы управления очередями, которые фактически отвечают и за решение задач по распределению пропускной способности каналов связи между потоками пользователей. Как показал проведенный анализ [6, 7], в современном коммутационном оборудовании реализовано множество механизмов управления очередями как с точки зрения их формирования и обслуживания (FIFO, PQ, CQ, FQ/WFQ, CBQ, LLQ), так и предотвращения перегрузки (RED, WRED, ECN, SPD). Они отличаются своими возможностями по дифференциации обслуживания, возможностями по настройке и эффективностью в целом. Решение по включению того или иного механизма на конкретном интерфейсе принимает администратор сети.

1.2. Анализ известных механизмов управления очередями на маршрутизаторах ТКС

Условно механизмы управления очередями можно разделить на две группы: механизмы обслуживания очередей и механизмы превентивного ограничения длины очереди [6, 7, 45, 46] (рис. 1.8). С их помощью можно определить число организованных на интерфейсе очередей, выставить максимально возможную длину очереди, контролировать загруженность очереди и управлять процессом отбрасывания пакетов в случае перегрузки интерфейса. Ряд механизмов допускают настройку порядка очередности обслуживания пакетов.

Важно отметить, что некоторые механизмы решают возложенные на них задачи в автоматическом режиме, а некоторые из технологических

средств управления очередями требуют постоянного вмешательства администратора сети. Остановимся на достоинствах и недостатках основных механизмов управления очередями, сравнив их и определив область их целесообразного применения.

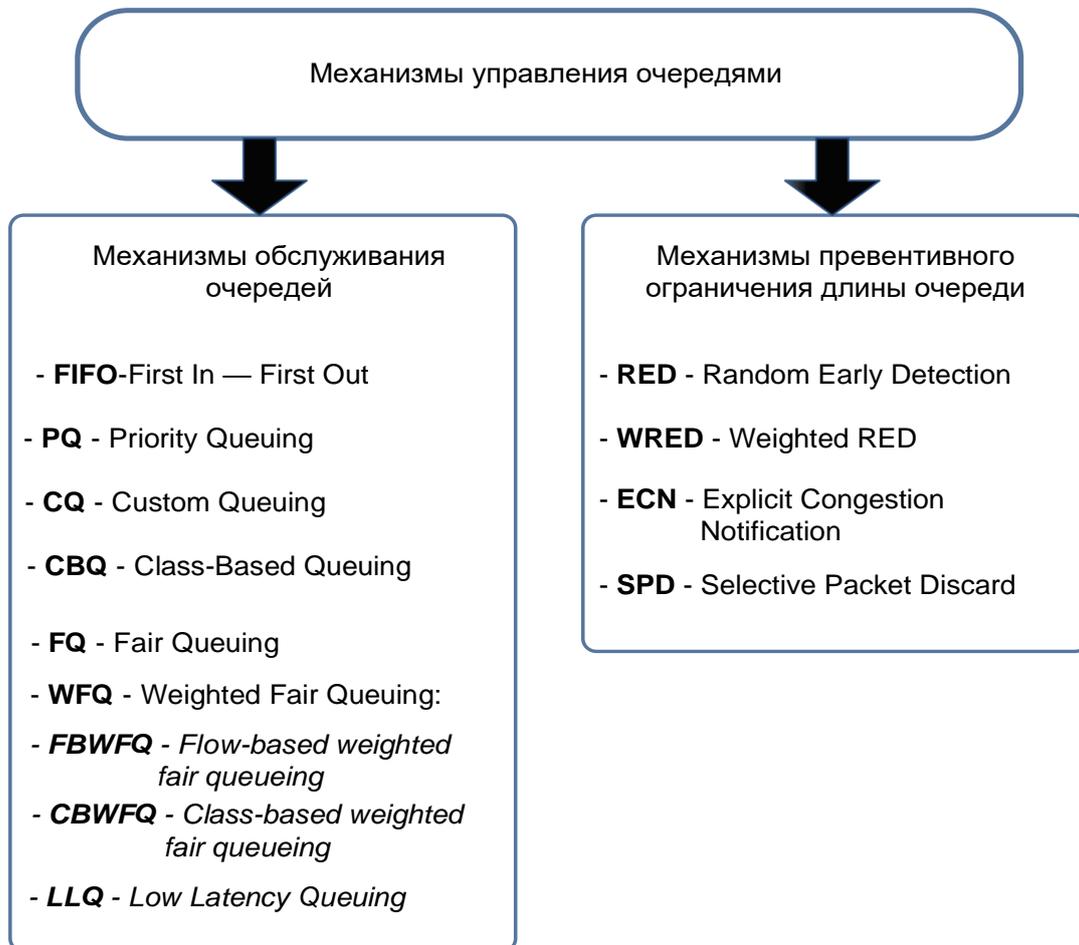


Рис.1.8. Классификация механизмов управления очередями

Механизм FIFO (First In, First Out) основан на одноименном алгоритме. При его использовании на интерфейсе организуется единственная очередь, ее длина по умолчанию составляет 40 пакетов. Поэтому порядок обслуживания пакетов полностью определяется порядком их поступления в очередь, т.е. не зависит от приоритета пакета, его класса или длины. Отсутствие дифференциации в обслуживании является основным недостатком данного решения, однако этот механизм широко используется

на практике, т.к. его реализация вносит минимальные задержки на обработку пакетов и практически не занимает процессорное время маршрутизатора.

Например, в маршрутизаторах, произведенных компанией Cisco, данный механизм включается «по умолчанию» на высокоскоростных интерфейсах, имеющих производительность 2 Мбит/с и выше. Это определяется тем, что вероятность его перегрузки невысока и нет смысла организовывать систему очередей и необоснованно тратить время на обработку пакетов в них. В условиях вероятной перегрузки интерфейса использование механизма FIFO крайне нежелательно. Кроме того, этот механизм используется при обработке пакетов в рамках каждой отдельной очереди, организованных с помощью других механизмов, например, CQ, PQ или WFQ.

Механизм приоритетных очередей (Priority queueing, PQ) организует уже четыре очереди: {high | medium | normal | low}, которые по умолчанию имеют максимальную длину в 20, 40, 60 и 80 пакетов соответственно. Планировщик пакетов в этом механизме работает по следующему принципу: пакеты из данной очереди начнут обслуживаться в том случае, если более приоритетные очереди в этот момент времени пусты. Таким образом, сетевой администратор, как и в случае с FIFO, никак не может повлиять на порядок обслуживания пакетов в очередях – все определяется приоритетом очереди. Однако на человека полностью перекладываются функции по заполнению очередей путем формирования списков доступа (Access Control Lists, ACLs) в режиме командной строки (Command Line Interface, CLI). Именно в этих списках заложена возможность учета параметров передаваемых пакетов (IP-адресов отправителя и получателя, номеров портов, приоритета пакета, типа используемого протокола, кода длины и др.).

Таким образом, возможности по дифференциации обслуживания в механизме PQ выше, чем в механизме FIFO. Если направить чувствительный к задержке поток VoIP в очередь {high}, то этим обеспечится максимально высокий уровень качества обслуживания. Однако неправильно

сконфигурированные списки доступа могут служить причиной полной блокировки обслуживания пакетов в низкоприоритетных очередях, что ограничивает использование PQ на практике.

В оборудовании также реализован механизм обычных очередей (Custom Queuing, CQ), с помощью которого уже организовывается 16 очередей (и 1 системная). Заполнение очередей, как и в случае применения PQ, осуществляется в ходе формирования ACLs. Обслуживание очередей, т.е. определение очередности передачи пакетов из очереди в канал, осуществляется в карусельном режиме (Round Robin), но число отправленных пакетов за один цикл просмотра очереди регулируется с помощью управляющего параметра – счетчика байт (Byte Counter). Этот счетчик «по умолчанию» для каждой очереди равен 1500 байт. Увеличивая или уменьшая значения данного параметра, администратор может регулировать уровень качества обслуживания пакетов настраиваемой очереди. Таким образом, возможности по дифференциации данного механизма выше, чем в FIFO и PQ, т.к. с его помощью можно настроить большее число очередей. Кроме того, с помощью счетчика байт можно достаточно точно распределить пропускную способность интерфейса в заданной пропорции между отдельными очередями. К недостаткам механизма CQ стоит отнести отсутствие учета приоритета пакета и других его параметров при организации процесса обслуживания.

Семейство механизмов «справедливых» или «честных» очередей включает в себя решения FQ (Fair Queuing), WFQ (Weighted FQ) и FB-WFQ (Flow-based WFQ). Неоспоримым достоинством данных механизмов является поддержка «по умолчанию» до 256 очередей (с возможностью административного увеличения до 4096). Это значительно превышает возможности по дифференциации обслуживания выше рассмотренных решений. Формирование и обслуживание очередей пакетов в данном случае осуществляется автоматически, т.е. практически без вмешательства администратора, который дополнительно может еще изменять размер буфера

очереди.

Механизм FQ при определении порядка обслуживания путем расчета и сравнения порядковых номеров пакета отдает предпочтение коротким пакетам, с помощью которых циркулирует мультимедийный трафик, чувствительный к задержкам и джиттеру (табл. 1.1). В механизме WFQ на порядок обслуживания также влияет дополнительно и приоритет пакета, который через вес пакета влияет на расчет порядкового номера пакета.

Ввиду неоспоримых преимуществ механизма WFQ, к числу которых, прежде всего, относят высокую дифференциацию обслуживания (организует 256 очередей); автоматическое функционирование (и при формировании очередей, и при их обработке); учет длины пакета и его приоритета, именно его рекомендуют [6, 7, 45, 46] как решение «по умолчанию» на низкоскоростных интерфейсах, имеющих производительность менее 2 Мбит/с.

В случае обработки мультимедийного трафика, когда важно учитывать не только приоритет пакетов, но и интенсивность самого потока, целесообразно использовать механизм FB-WFQ, т.е. основанные на потоке взвешенные честные очереди. В основу данного решения положено математическое выражение для расчета B_j , т.е. пропускной способности, выделяемой для каждого потока с j -м IP-приоритетом:

$$B_j = B \cdot \frac{IP_{pr}^j + 1}{\sum_{i=0}^7 M_{flow}^i \cdot (IP_{pr}^i + 1)}, \quad (1.1)$$

где B – пропускная способность интерфейса;

IP_{pr}^j – j -й IP-приоритет потока;

M_{flow}^i – общее число потоков с i -м IP-приоритетом.

Все больше внимания в последнее время уделяется технологическим решениям по управлению трафиком, основанных на классах обслуживания. Механизм CBQ (Class-Based Queuing) предполагает организацию классовых очередей и выделение им заданной части пропускной способности интерфейса в явном виде, а не косвенно, как в механизме CQ, через счетчик байт. Классы очередей формируются с помощью все тех же списков доступа. К примеру, в механизме CBWFQ (Class-Based WFQ) организуется 64 классовых очереди, в т.ч. одна очередь «по умолчанию», которая может работать в режиме WFQ, распределяя между неклассифицированными потоками 25% от пропускной способности интерфейса. Оставшаяся часть пропускной способности интерфейса может административно с помощью команды «bandwidth» перераспределяться между классовыми очередями. В механизме LLQ (Low Latency Queuing), также известном как PQ-CBWFQ, в дополнении к очередям CBWFQ вводится дополнительная приоритетная очередь, а выделяемый ей ресурс определяется командой «priority».

Решения, основанные на классах, позволяют организовать достаточно большое число очередей с заранее заданным уровнем дифференциации, т.к. и в классовых очередях можно использовать режим работы WFQ. Но решение задач по формированию (заполнению) очередей путем классификации трафика и определение порядка обслуживания пакетов полностью возлагается на администратора сети. Поэтому эффективность получаемых решений полностью зависит от его опыта и уровня квалификации.

В табл. 1.2 приведены результаты сравнительного анализа и краткая характеристика выше рассмотренных механизмов обслуживания очередей.

На интерфейсах маршрутизаторов транспортных ТКС также настраиваются и функционируют средства предотвращения перегрузки очередей на основе превентивного ограничения длины очереди. Эти средства еще называют механизмами активного управления очередью (Active Queue Management, AQM), в отличие от схемы Tail drop, которая реализует

пассивную стратегию ограничения длины очереди по ее фактическому переполнению.

Таблица 1.2

Результаты сравнительного анализа и краткая характеристика выше рассмотренных механизмов обслуживания очередей

Свойство механизма	Название механизма					
	FIFO	PQ	CQ	FQ/WFQ, FB-WFQ	CBWFQ	LLQ
Число очередей (по умолчанию)	1	4	16	256 (можно увеличить до 4096)	64	64+1
Формирование очереди	Авт.*	Ручн., CLI- ACLs	Ручн., CLI- ACLs	Авт.	Ручн.**, CLI- ACLs	Ручн., CLI- ACLs
Обслуживание очереди (планировщик)	Авт.	Strict Priority	Round Robin, Byte counter	Авт.	Ручн., CLI	Ручн., CLI
Обеспечение QoS-гарантий	Нет	Нет	Нет	Нет	Да	Да
* Авт. – автоматическое решение;						
** Ручн. – ручная конфигурация.						

Наиболее распространенными AQM-механизмами являются [6, 57]:

- механизм случайного раннего обнаружения – RED (Random Early Detection);
- механизм взвешенного случайного раннего обнаружения – WRED (Weighted RED);

- механизм взвешенного случайного раннего обнаружения на основе потока – Flow WRED;
- механизм явного уведомления о перегрузке – ECN (Explicit Congestion Notification);
- механизм избирательного отбрасывания пакетов – SPD (Selective Packet Discard).

К основным целям и задачам AQM-механизмов стоит отнести такие:

1. Обеспечение управляемости процессом борьбы с перегрузкой на интерфейсе маршрутизатора.
2. Минимизация вероятности возникновения эффекта «глобальной синхронизации», приводящего к неэффективному использованию пропускной способности ТСП-соединения.
3. Обеспечение дифференциации в отбрасывании пакетов в соответствии с их приоритетом (WRED) или типом трафика (SPD).
4. Снижение вероятности повторной передачи пакетов/сегментов протоколом ТСП.
5. Улучшение качества обслуживания в сети в целом.

Для достижения поставленных целей механизм RED, основываясь на постоянном мониторинге за загруженностью очереди и последовательной работе алгоритмов определения средней длины очереди и расчета вероятности отбрасывания пакетов, начинает случайным образом отбрасывать пакеты из очереди, не дожидаясь ее полной перегрузки, т.е. заблаговременно (превентивно). Степень превентивности регулируется административно путем задания значения минимального порога для средней длины очереди (Θ_{\min}).

В механизме WRED учет приоритета обрабатываемых пакетов (табл. 1.3) осуществляется путем дифференцированной настройки минимального и максимального порога (Θ_{\max}) для средней длины очереди или знаменателя граничной вероятности (δ). В механизме SPD вместо приоритета пакета критерием отбрасывания пакетов является тип трафика

(управляющий, пользовательский или не прошедший проверку на корректность).

Таблица 1.3

Значения управляющих параметров WRED «по умолчанию»

IP-приоритет	Минимальное значение очереди Θ_{\min}	Максимальное значение очереди Θ_{\max}	знаменатель граничной вероятности δ	Максимальный процент отброшенных пакетов $\frac{1}{\delta}$
0	20	40	10	10%
1	22	40	10	10%
2	24	40	10	10%
3	26	40	10	10%
4	28	40	10	10%
5	31	40	10	10%
6	33	40	10	10%
7	35	40	10	10%
RSVP	37	40	10	10%

В механизме ECN предполагается явное уведомление TCP-источника о перегрузке путем установления специальных ECN-бит (рис. 1.6) в поле ToS тех пакетов, которые механизм WRED решил отбросить из очереди. Это в большинстве случаев несколько улучшает уровень QoS, т.к. уменьшается число отброшенных пакетов и не тратится время на их повторную передачу с помощью транспортного протокола.

На рынке AQM-решений также присутствует множество модификаций механизма RED/WRED [45, 46], отличающиеся, как правило, порядком расчета вероятности отбрасывания пакетов при грядущей перегрузке интерфейса. К подобным решениям, прежде всего, относят RED со стабилизацией (Stabilized RED, SRED), случайное экспоненциальное маркирование (Random Exponential Marking, REM), RED с двоичной линейной функцией отбрасывания пакетов с различными наклонами (Double Slope RED, DSRED), справедливый RED (Fair RED, FRED), RED со справедливой буферизацией (Fair Buffering RED, FBRED), RED с порогом на

основе классов трафика (Class based threshold RED, CBT-RED), RED с балансировкой (Balanced RED, BRED).

Общим недостатком семейства механизмов RED является превалирование ручных настроек, которые заметно снижают оперативность реакции на возможную перегрузку интерфейса и приводят к тому, что результативность управления всецело зависит от квалификации сетевого администратора. Кроме того, к недостаткам известных AQM-механизмов стоит отнести отсутствие согласованности в работе с другими механизмами управления очередями, например, с FIFO, PQ, CQ, FQ/WFQ, LLQ, а также с родственными средствами борьбы с перегрузкой. Причина этого является использование эвристических схем и несовершенных математических моделей обслуживания очередей и борьбы с перегрузкой очереди, что является основной причиной снижения производительности TCP-соединений, роста средней задержки пакетов, джиттера и числа необоснованно отброшенных пакетов.

Данный класс решений постоянно совершенствуется и находит все новые области применения [47-57].

1.3. Обзор известных математических моделей и методов управления очередями на маршрутизаторах ТКС

К настоящему времени известно достаточно много подходов к математическому описанию процессов управления очередями, отличающихся постановкой задачи, типом используемого математического аппарата, ожидаемой точностью расчетов и уровнем сложности вычисления искомых значений. Большую роль в выборе того или иного подхода также играет область применения моделей и методов – для решения задач анализа возможных вариантов и стратегий управления очередями или задач синтеза, связанных с перспективой дальнейшей практической реализации в новых сетевых механизмах.

Наиболее адекватным при описании процессов управления очередями является подход, основанный на использовании аппарата дифференциально-разностных или интегральных уравнений состояния интерфейса. При этом под состоянием интерфейса может пониматься средняя длина очереди, коэффициент загрузки, средняя скорость передачи пакетов и т.д.

Одним из примеров подобных решений является применение нелинейных дифференциальных уравнений для описания динамики изменения загрузки очередей интерфейса, полученных на основе различного рода аппроксимаций (табл. 1.4) [58-60].

Таблица 1.4

Перечень типов аппроксимаций, используемых для описания динамики изменения состояния интерфейса

Тип аппроксимации	Краткая характеристика
The Simple Stationary Approximation (SSA)	Простая стационарная аппроксимация
The Closure Approximation for Nonstationary Queues	Замкнутая аппроксимация для нестационарных очередей
The Pointwise Stationary Approximation (PSA)	Стационарная точечная аппроксимация
The Pointwise Stationary Fluid Flow Approximation (PSFFA)	Стационарная точечная аппроксимация нестационарного потока

Для примера, при использовании аппарата нелинейных дифференциальных уравнений и стационарной точечной аппроксимации нестационарного потока (Pointwise Stationary Fluid Flow Approximation, PSFFA) динамика изменения загрузки очереди может быть представлена в следующем виде:

$$\dot{x}(t) = -\mu \left(\frac{x(t)}{x(t)+1} \right) + \lambda, \quad (1.2)$$

где $x(t)$ – средняя длина очереди на рассматриваемом интерфейсе;

λ – интенсивность потока на входе интерфейса (1/с);

μ – пропускная способность интерфейса (очереди) (1/с).

Использование подобного подхода позволяет отслеживать не только текущее значение средней длины очереди на интерфейсе, но и прогнозировать ее значение на других временных интервалах.

В работах [61, 62] рассмотрен вариант применения аппарата разностных неавтономных уравнений состояния для описания динамики очередей на маршрутизаторах ТКС:

$$x_{i,j}(k+1) = x_{i,j}(k) - \sum_{\substack{l=1, \\ l \neq i}}^M b_{i,l}(k) u_{i,l}^j(k) + \sum_{\substack{m=1, \\ m \neq i,j}}^M b_{m,i}(k) u_{m,i}^j(k) + y_{i,j}(k), \quad (1.3)$$

где $b_{m,i}^{(M)}(k) = c_{m,i}(k) \cdot \Delta t$, $y_{i,j}(k) = \zeta_{i,j}(k) \cdot \Delta t$, ($k=0,1,2,\dots$; $\Delta t = t_{k+1} - t_k$);

$x_{i,j}(k)$ – объем данных, находящихся в очереди на i -м маршрутизаторе и предназначенных для передачи j -му маршрутизатору в момент времени t_k ;

$\varphi_{i,j}(k)$ – пропускная способность канала связи (i, j) в момент времени t_k ;

$u_{i,l}^j(k)$ – доля пропускной способности канала связи (i, j), выделенная данным для j -го маршрутизатора в момент времени t_k ;

$\zeta_{i,j}(k)$ – интенсивность поступления данных от абонентов сети на i -й маршрутизатор для j -го маршрутизатора в момент времени t_k .

В целом данная модель была адаптирована для решения задач маршрутизации на основе балансировки длин очередей на маршрутизаторах ТКС [63].

Примером применения нелинейных интегральных уравнений Вольтерра второго рода для описания задач управления очередями может также служить модель, предложенная в работе [64] и описывающая динамику изменения интенсивности потоков в каналах связи и загрузку очередей на маршрутизаторах ТКС:

$$\lambda_i(t) = \sum_{j \neq i}^z \int_0^t p_{ji}(\lambda_j^0 + \lambda_j(\tau)) \phi_{ji}(t-\tau, \lambda_j^0 + \lambda_j(t-\tau)) \lambda_j(\tau) d\tau + \vartheta_i(t), \quad (1.4)$$

где $\lambda_i(t)$ – интенсивность обслуживания пакетов на выходе очереди без потерь (j,i) в момент времени t ;

z – количество очередей на узле сети;

$\lambda_j(\tau)$ – интенсивность потока в момент времени τ ;

$p_{j,i}(\lambda_j^0)$ – вероятность обслуживания пакетов;

$\vartheta_i(t)$ – детерминированная составляющая потока $\lambda_i(t)$;

λ_j^0 – интенсивность потока пакетов на обслуживание в момент времени t_0 ;

$\phi_{i,j}(\tau, \lambda_j^0) = \frac{d}{dt} \Phi_{j,i}(\tau, \lambda_j^0)$ и $\Phi_{j,i}(\tau, \lambda_j^0)$ – функция распределения времени обслуживания пакетов в очереди (j,i) .

В работе [65] рассмотрены некоторые подходы к математическому описанию процессов контроля перегрузки каналов и очередей сети на основе взаимодействия ТСП/AQM-решений с использованием аппарата дифференциальных уравнений. В качестве переменной состояния могут выступать окно перегрузки, характеризующее скорость передачи данных ТСП-источником. Например, модель, описывающая динамику ТСП-сеанса может быть представлена рекуррентным соотношением:

$$X_{i,n+1} = X_{i,n} + r_i \chi(M_{i,n} = 0) - \eta_i X_{i,n} \chi(M_{i,n} = 0), \quad (1.5)$$

где $X_{i,n} = X(t_{i,n})$ – скорость передачи данных i -м источником;

n – номер момента корректировки процесса;

$t_{i,n}$ – последовательность моментов времени, когда процесс корректируется;

$M_{i,n}$ – число потерянных пакетов за время $t_{i,n+1} - t_{i,n}$,

$\chi(A)$ – индикатор множества.

Учет динамики протекающих на интерфейсе процессов является неоспоримым преимуществом рассмотренных подходов. Однако решение в реальном времени систем дифференциальных (в т.ч. нелинейных) уравнений не всегда представляется возможным. Это существенно повышает требования к вычислительной мощности маршрутизатора, а значит, и его стоимости. Поэтому подобные модели и методы целесообразно использовать при решении задач анализа или для получения своеобразного эталона для сравнения с ним субоптимальных (упрощенных) решений с возможностью количественной оценки расчетной погрешности.

В работах [66, 67] обосновывается возможность применения нейронных сетей и методов нечеткой логики для организации управления очередями, в частности отбрасывания пакетов на транзитных маршрутизаторах телекоммуникационной сети. Применение такого подхода нацелено на повышение уровня обоснованности решений по отбрасыванию пакетов за счет применения постоянно обучаемых нейронных сетей, которые способны адаптироваться к изменению пропускной способности интерфейса и уровня его загруженности. Однако вопросы скорости обучения подобной системы требуют дополнительных исследований, т.к. управление очередью реализуется в миллисекундном масштабе времени ввиду высокой динамики изменения состояния интерфейса. Кроме того, методы, предложенные в [66, 67], ориентируют на получение допустимых, но не оптимальных решений, что может способствовать неэффективному использованию доступного сетевого (канального и буферного) ресурса.

Компромиссным вариантом с точки зрения уровня адекватности математического описания процессов управления очередями и ожидаемой вычислительной сложностью получаемых решений может служить подход,

представленный в работах [9-15]. Он основан на оптимизационной постановке задачи управления очередями с точки зрения сбалансированного использования доступного ресурса и загруженности очередей. Кроме того, эти решения позволяют обеспечить согласованное решение основных интерфейсных задач. Одним из недостатков данных решений является отсутствие учета возможности агрегирования потоков при формировании очередей, что является важным моментом при их практической реализации, т.к. направление в одну очередь потоков различных классов (приоритетов) может нарушить основные принципы дифференцированного обслуживания в ТКС. Кроме того, необходимо адаптировать полученные ранее решения на новый класс сетей – на активные сети, в рамках которых время нахождения пакета в очереди может использоваться в ходе его распределенной обработки при предоставлении новых инфокоммуникационных сервисов.

1.4. Формулировка требований к перспективным решениям в области управления очередями на маршрутизаторах ТКС. Постановка научной задачи

Как показал проведенный анализ, для повышения уровня качества обслуживания и эффективного распределения доступного сетевого ресурса к перспективным моделям управления очередями выдвигается ряд важных требований, к основным из которых стоит отнести следующие (рис. 1.9):

- учет потоковой структуры современного сетевого трафика, обусловленного ростом общего числа мультимедиа приложений в перечне инфокоммуникационных услуг;
- оптимизационная постановка и решение задачи управления очередями, связанная с необходимостью сбалансированного использования доступного сетевого (буферного и канального) ресурса;
- поддержка дифференциации обслуживания пакетов на интерфейсах маршрутизаторов ТКС в соответствии с их QoS-требованиями;

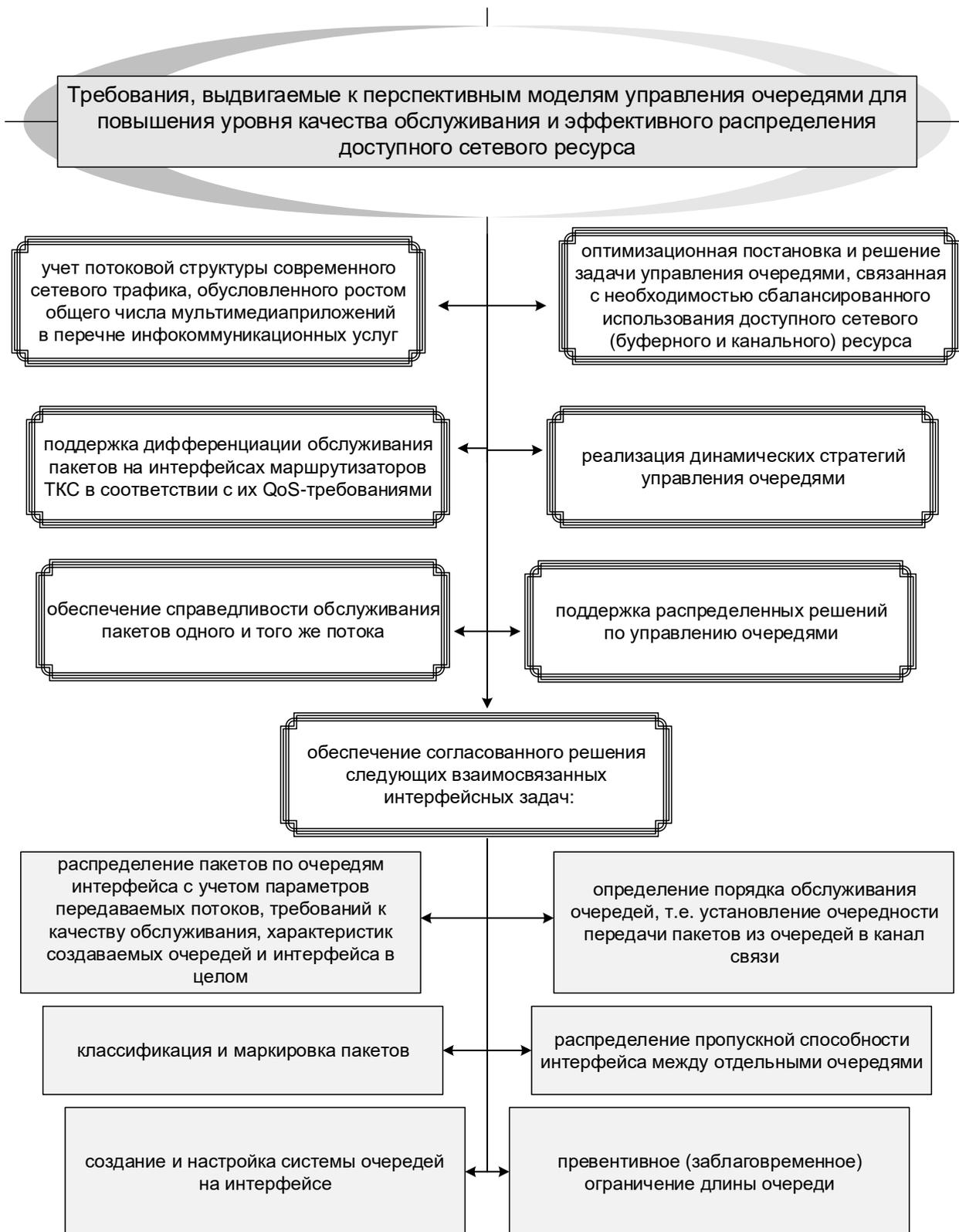


Рис. 1.9. Требования, выдвигаемые к перспективным моделям управления очередями для повышения уровня качества обслуживания и эффективного распределения доступного сетевого ресурса

- реализация динамических стратегий управления очередями;
- обеспечение справедливости обслуживания пакетов одного и того же потока;
- поддержка распределенных решений по управлению очередями;
- простота алгоритмически-программной и аппаратной реализации;
- обеспечение согласованного решения следующих взаимосвязанных интерфейсных задач:

- классификация и маркировка пакетов;
- создание и настройка системы очередей на интерфейсе;
- агрегирование потоков и распределение пакетов по очередям интерфейса с учетом параметров передаваемых потоков, требований к качеству обслуживания, характеристик создаваемых очередей и интерфейса в целом;
- определение порядка обслуживания очередей, т.е. установление очередности передачи пакетов из очередей в канал связи;
- распределение пропускной способности интерфейса между отдельными очередями;
- превентивное (заблаговременное) ограничение длины очереди.

В этой связи актуальной представляется *научная задача*, связанная с оптимизацией процесса управления очередями на интерфейсе маршрутизаторов на основе обеспечения согласованного решения задач по агрегированию потоков, распределения пакетов между очередями и выделения очередям пропускной способности интерфейса путем разработки соответствующих математических моделей и метода для повышения качества обслуживания в транспортных ТКС в целом.

Для решения поставленной научной задачи в диссертационной работе решены следующие частные задачи исследования:

- анализ протокольных решений по управлению трафиком и очередями на интерфейсах маршрутизаторов транспортных ТКС;

- обзор известных теоретических решений, математических моделей и методов управления очередями на интерфейсах маршрутизаторов транспортных ТКС;
- формулирование требований к перспективным решениям в области управления очередями на маршрутизаторах ТКС;
- усовершенствование потоковых моделей и метода управления очередями на интерфейсах маршрутизаторов транспортных ТКС;
- анализ адекватности и исследование эффективности предложенных в диссертации решений по управлению очередями;
- разработка рекомендаций по практическому использованию полученных в работе результатов в современных и перспективных транспортных ТКС.

1.5. Выводы по первому разделу

1. В ходе анализа принципов построения и функционирования современных телекоммуникационных систем и сетей установлено, что они развиваются по пути совершенствования NGN-решений, основанных на поддержке качества обслуживания. Показано, что сетевые и в особенности мультимедийные приложения требуют дифференциации качества обслуживания одновременно по множеству QoS-показателей: скорости, средней задержке, вероятности потерь пакетов.

2. Обзор сетевых протоколов и механизмов обеспечения качества обслуживания показал, что наряду с маршрутизацией, резервированием ресурсов, маркировкой и классификацией пакетов, профилированием трафика, важное место в архитектуре QoS отводится технологическим средствам управления очередями, которые отвечают за их формирование и контроль перегрузки. Именно неэффективное управление очередями приводит к неконтролируемому росту задержек, джиттера и уровня потерь пакетов. Установлено, что ключевым недостатком известных механизмов

управления очередями является преобладание ручных настроек в ходе конфигурации оборудования, что не позволяет оперативно реагировать на изменение состояния (загруженности) интерфейсов маршрутизатора и сети в целом, а также на вариацию характеристик трафика.

3. Важность и актуальность проблематики обеспечения эффективного управления очередями (буферным ресурсом) определила тот факт, что многие отечественные и зарубежные ученые активно работают над совершенствованием средств управления трафиком и обеспечения качества обслуживания в ТКС, в частности и в направлении пересмотра моделей, методов и самих механизмов управления очередями на маршрутизаторах сети. В результате анализа перспективных математических моделей и методов управления очередями установлено, что будущее за потоковыми и оптимизационными решениями, которые бы являлись компромиссом между высоким уровнем адекватности описания и получаемой вычислительной сложностью.

4. На основе проведенного анализа сформулирована система требований, которые выдвигаются к перспективным моделям и методам управления очередями, к числу которых, прежде всего, относятся учет потоковой структуры современного сетевого трафика; оптимизационная постановка и решение задачи управления очередями; поддержка дифференциации обслуживания пакетов на интерфейсах маршрутизаторов ТКС в соответствии с их QoS-требованиями; реализация динамических стратегий управления очередями; простота алгоритмически-программной и аппаратной реализации; а также обеспечение согласованного решения взаимосвязанных интерфейсных задач (классификация и маркировка пакетов; создание и настройка системы очередей на интерфейсе; агрегирование потоков и распределение пакетов по очередям интерфейса с учетом параметров передаваемых потоков, требований к качеству обслуживания, характеристик создаваемых очередей и интерфейса в целом; определение порядка обслуживания очередей; распределение пропускной

способности интерфейса между отдельными очередями; превентивное (заблаговременное) ограничение длины очереди).

5. В этой связи обоснована к решению в настоящей работе научная задача, которая связана с оптимизацией процесса управления очередями на интерфейсе маршрутизаторов на основе обеспечения согласованного решения задач по агрегированию потоков, распределения пакетов между очередями и выделения очередям пропускной способности интерфейса путем разработки соответствующих математических моделей и метода для повышения качества обслуживания в транспортных ТКС в целом. Произведена декомпозиция сформулированной научной задачи на частные задачи исследования.

РАЗДЕЛ 2

РАЗРАБОТКА КОМПЛЕКСНОЙ МАТЕМАТИЧЕСКОЙ МОДЕЛИ УПРАВЛЕНИЯ ОЧЕРЕДЯМИ НА МАРШРУТИЗАТОРАХ ТРАНСПОРТНОЙ ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ

Как показал проведенный в первом разделе анализ, для повышения уровня качества обслуживания и эффективного распределения доступного сетевого ресурса к перспективным механизмам и моделям управления очередями выдвигается ряд важных требований, к основным из которых стоит отнести учет потоковой структуры современного сетевого трафика; оптимизационную постановку и решение задачи управления очередями; поддержку дифференциации обслуживания пакетов на интерфейсах маршрутизаторов ТКС в соответствии с их QoS-требованиями; реализацию динамических стратегий управления очередями; простоту алгоритмически-программной и аппаратной реализации; обеспечение согласованного решения взаимосвязанных интерфейсных задач.

В этой связи в данном разделе предложена комплексная математическая модель управления очередями на маршрутизаторах транспортной телекоммуникационной сети, в рамках которой максимально удовлетворяются вышеперечисленные требования. Продемонстрированы ее преимущества относительно обеспечения оптимального агрегирования потоков и сбалансированной загруженности интерфейса в целом и каждой очереди в частности в соответствии с ее классом. Работа модели продемонстрирована на ряде численных примеров.

Материалы второго раздела опубликованы в работах [9-18].

2.1. Вводимые в рамках предлагаемой модели условные обозначения

В основу предлагаемой комплексной модели будут положены результаты, полученные в работах [13, 14, 70]. Тогда пусть на интерфейс маршрутизатора поступает M потоков со следующими известными характеристиками:

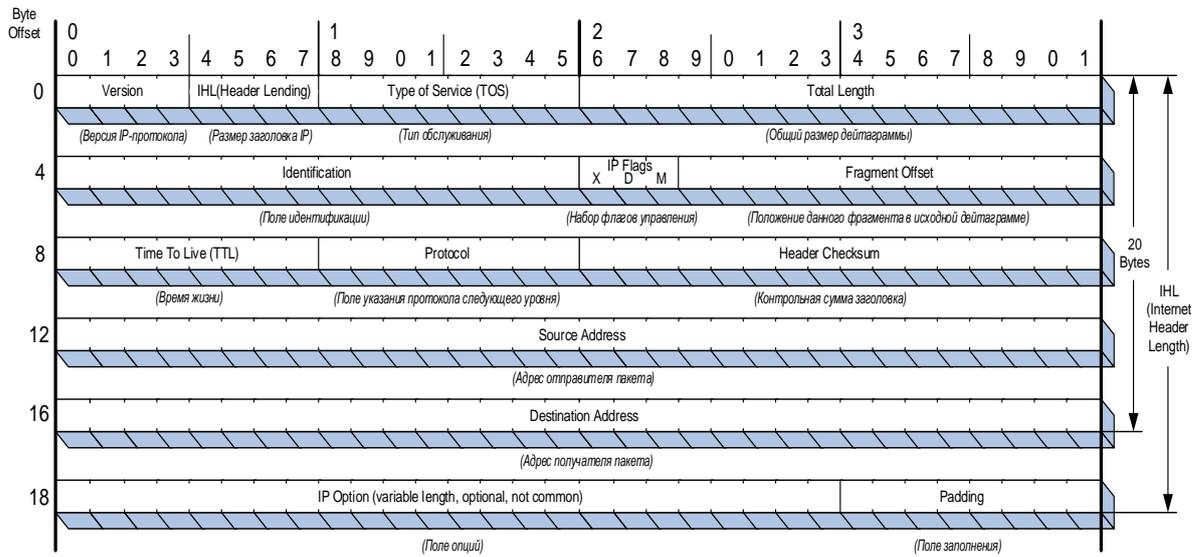
a_i – величина средней интенсивности i -го потока, измеряемая в пакетах в секунду (1/с);

$K_i = \{k_i^l, l = \overline{1, L}\}$ – множество параметров пакетов i -го потока, которые используются для классификации сетевого трафика в рамках реализованной в ТКС транспортной технологии;

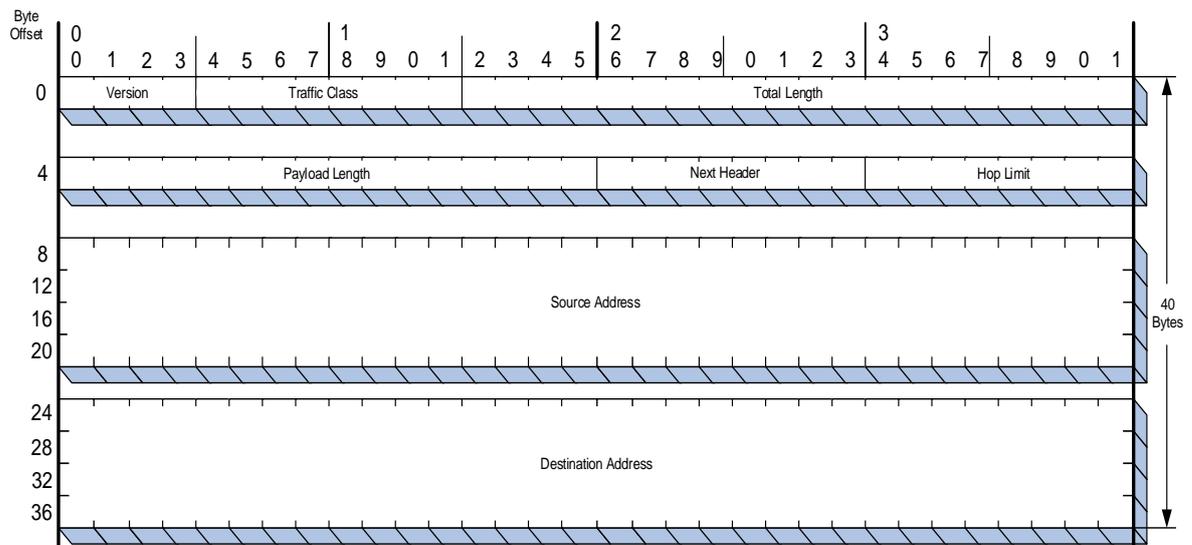
L – общее число параметров для классификации трафика.

Примером подобных параметров в IP-технологии являются приоритет пакета (содержимое полей IP-пакета: IP-precedence или DSCP) или значения QoS-group; IP-адреса источника и (или) получателя пакетов; MAC-адреса источника и (или) получателя пакетов; стандартный или расширенный список доступа к источникам/получателям; TCP/UDP-порты источника и (или) получателя пакетов; код длины пакета (рис. 2.1). В других технологиях пакетной коммутации к подобным параметрам также могут быть отнесены экспериментальные биты MPLS, бит Frame Relay DE.

Пусть для каждого i -го потока на основе анализа множества параметров $\{k_i^l, l = \overline{1, L}\}$ определяется его класс k_i^{Π} , который является некоторой функцией от элементов множества K_i .



а) для IPv4



б) для IPv6

Рис. 2.1. Указание основных полей заголовка IP-пакета

В общем случае данная функция может носить нелинейный характер, так, например, в работе [13] аналог данной зависимости имел форму

$$k_i^{\text{II}}(p_i, d_i) = \frac{p_i}{v \cdot d_i}, \quad (i = \overline{1, M}), \quad (2.1)$$

где p_i – приоритет пакетов i -го потока;

d_i – средняя длина пакетов i -го потока;

v – некоторый нормировочный коэффициент, который должен сглаживать различие в порядке значений приоритета ($0 \div 7$ или $0 \div 63$) и длины пакета в байтах.

В общем случае величина k_i^{II} является безразмерной, и для удобства последующего изложения будем считать, что k_i^{II} нормируется в пределах от единицы до десяти. Самый важный поток будет иметь максимальное значение класса, т.е. $k_i^{\text{II}} = 10$. Чем меньше класс, тем значение k_i^{II} будет ближе к единице.

Кроме того, для определения класса потока можно использовать аналог маршрутной метрики протокола IGRP (Interior Gateway Routing Protocol) [6, 68, 69]:

$$k^{\text{II}} = v_r [K_1 M_b + K_2 M_d] \frac{K_3}{M_r + K_4}, \quad (2.2)$$

где K_1, K_2, K_3, K_4 – коэффициенты, задаваемые административно и определяющие влияние того или иного QoS-показателя на класс потока;

M_b, M_d, M_r – величины, характеризующие отдельные показатели качества обслуживания, причем:

M_b – величина, обратная требованиям по пропускной способности;
 M_d – величина, характеризующая требования к средней задержке пакетов;
 M_r – величина, определяющая требуемый уровень надежности доставки пакетов;
 v_r – нормировочный коэффициент, вводимый для обеспечения выполнения условия $1 \leq k^{\text{II}} \leq 10$.

При использовании выражения (2.2) удастся обеспечить учет в явном виде требований к качеству обслуживания при управлении очередями на интерфейсе маршрутизатора ТКС.

Обозначим через N общее число организуемых на интерфейсе очередей (табл. 2.1). В рамках последующего изложения условимся, что число потоков превышает количество поддерживаемых на интерфейсе очередей, т.е. имеет место неравенство

$$M > N. \quad (2.3)$$

По аналогии с классификацией потоков (пакетов) будем считать, что в системе управления очередями установлены также классы и самих очередей, как это сделано, например, в механизмах CBQ, CBWFQ и LLQ. Тогда каждой j -й очереди сопоставим класс k_j^{O} ($j = \overline{1, N}$), который по аналогии с классом потоков k_i^{II} также является безразмерной величиной, варьируемой в пределах от 1 до 10. Классы k_j^{O} ($j = \overline{1, N}$) распределяются равномерно между очередями в зависимости от общего числа самих очередей.

**Число очередей, поддерживаемых по умолчанию в основных
механизмах обслуживания очередей пакетов**

Название планировщика	Число поддерживаемых очередей
FIFO	1
PQ	4
CQ	16 + 1 (системная)
FQ/WFQ	256 (с возможностью административного увеличения до 4096)
CBWFQ	64 класса очередей (каждая классовая очередь и очередь «по умолчанию» может настраиваться WFQ, т.е. разбиваться на 256 (до 4096) подочерей)
LLQ	К очередям CBWFQ добавляется одна приоритетная очередь

Таким образом, пакеты M потоков, поступивших на интерфейс в соответствии с содержанием актуальной таблицы маршрутизации, должны распределяться между N очередями в ходе решения задач Congestion Management. Для этого введем в рассмотрение множество переменных $x_{i,j}$ ($i = \overline{1, M}$, $j = \overline{1, N}$), каждая из которых характеризует долю i -го потока, направленного на обслуживание в j -ю очередь.

Для решения задач распределения пропускной способности интерфейса между очередями (Resource Allocation) в рамках предлагаемой модели необходимо рассчитать множество переменных b_j , каждая из которых характеризует величину пропускной способности интерфейса, выделенную для обслуживания пакетов из j -й очереди. Пусть b – общая

пропускная способность рассматриваемого интерфейса, на котором организуются данные очереди.

Кроме того, по аналогии с работой [70] в структуру предлагаемой модели введем множество (по числу потоков) управляющих переменных α_i ($i = \overline{1, M}$). Каждая из данных переменных характеризует долю i -го потока пакетов, получивших отказ в обслуживании (т.е. отброшенных из очереди) на интерфейсе маршрутизатора при реализации функций Congestion Avoidance. Фактически переменная α_i численно определяет вероятность отбрасывания пакетов i -го потока на рассматриваемом интерфейсе маршрутизатора.

2.2. Формализация условий сохранения потока и предотвращения перегрузки очередей на интерфейсе маршрутизатора

Известно, что организация нескольких очередей на одном интерфейсе производится для обеспечения дифференциации обслуживания потоков пакетов, имеющих различные требования к уровню качества обслуживания. Поэтому пакеты одного потока, как правило, направляются в одну очередь. Также в одну и ту же очередь могут направляться пакеты с близкими QoS-требованиями (классами). Поэтому в соответствии с физическим смыслом решаемой задачи переменная $x_{i,j}$ должна быть булевой:

$$x_{i,j} \in \{0,1\}, \quad (i = \overline{1, M}, j = \overline{1, N}). \quad (2.4)$$

Учитывая особенности постановки задачи управления очередями на управляющие переменные (2.4) должны быть наложены ограничения, представленные в виде условий сохранения потока на интерфейсе маршрутизатора: [13, 18, 70]:

$$\sum_{j=1}^N x_{i,j} = 1, \quad (i = \overline{1, M}). \quad (2.5)$$

Выполнение условия (2.5) совместно с (2.4) гарантирует, что пакеты i -го потока гарантированно будут направлены лишь в одну из очередей, организованных на рассматриваемом интерфейсе.

Переменные b_j , определяющие размер пропускной способности интерфейса, выделенной для обслуживания пакетов j -й очереди, являются положительными вещественными величинами, на которые также накладывается система ограничений вида

$$0 \leq b_j, \quad \sum_{j=1}^N b_j \leq b. \quad (2.6)$$

В ходе выполнения условия (2.6) соблюдается корректность распределения пропускной способности интерфейса между отдельными очередями.

Для обеспечения управляемости процессом предотвращения перегрузки (Congestion Avoidance) интерфейса необходимо выполнить условие:

$$\sum_{i=1}^M a_i (1 - \alpha_i) x_{i,j} < b_j, \quad (j = \overline{1, N}), \quad (2.7)$$

которое несколько отличается по своей форме и смыслу от подобных условий, предложенных в работе [70]. Смысл выражений (2.7) состоит в том, чтобы суммарная интенсивность потоков, направленных на обслуживание в j -ю очередь, не превышала пропускную способность интерфейса, которая выделена данной очереди.

Согласно физическому смыслу переменных α_i , на них также накладываются ограничения следующего вида:

$$0 \leq \alpha_i \leq 1. \quad (2.8)$$

В условиях стохастического и нестационарного характера процессов поступления и обслуживания пакетов на интерфейсе маршрутизатора выполнение условия (2.7) не является достаточным для предотвращения перегрузки буфера очереди по ее длине. Поэтому для каждой j -й очереди обозначим через \bar{n}_j и n_j^{\max} ($j = \overline{1, N}$) ее текущую длину (в пакетах) и максимальную емкость. Тогда условия управляемости процессом предотвращения перегрузкой (2.7) дополняются условиями предотвращения перегрузки очередей по их длине:

$$\bar{n}_j \leq n_j^{\max} \quad (j = \overline{1, N}), \quad (2.9)$$

где значения \bar{n}_j зависят от статистических характеристик потока, выбранной дисциплины обслуживания пакетов и выделенной данной очереди ПС интерфейса. Варианты расчетных выражений для оценки \bar{n}_j в зависимости от используемого типа системы массового обслуживания при моделировании работы интерфейса представлены в работах [71-73] и имеют вид:

M/M/1:

$$\bar{n}_j = \frac{\rho_j^2}{1 - \rho_j}, \quad (2.10)$$

M/D/1:

$$\bar{n}_j = \frac{\rho_j^2}{2(1 - \rho_j)}, \quad (2.11)$$

M/M/1/ N_{\max} :

$$\bar{n}_j = \frac{\rho_j}{1 - \rho_j^{(N_{\max} + 1)}} \cdot \frac{1 - (N_{\max} + 1) \cdot \rho_j^{N_{\max}} + N_{\max} \cdot \rho_j^{N_{\max} + 1}}{1 - \rho_j} - \rho_j, \quad (2.12)$$

fBM/M/1/ N_{\max} :

$$\bar{n}_j = \frac{\frac{\rho_j}{p_n} f(H)}{1 - \left[\frac{\rho_j}{p_n} f(H) \right]^{N+2}} \cdot \frac{\left\{ 1 - (N_{\max} + 1) \left[\frac{\rho_j}{p_n} f(H) \right]^{N_{\max}} \right\} + N_{\max} \left[\frac{\rho_j}{p_n} f(H) \right]^{N_{\max} + 1}}{1 - \frac{\rho_j}{p_n} f(H)} - \frac{\rho_j}{p_n} f(H), \quad (2.13)$$

где $\rho_j = \frac{\sum_{i=1}^M a_i (1 - \alpha_i) x_{i,j}}{b_j}$ – коэффициент использования пропускной

способности, выделенной для j -й очереди;

p_n – вероятность отсутствия повторной передачи пакетов;

H – коэффициент Херста;

$f(H)$ – функция, которая учитывает влияние самоподобия поступающего в узел потока [73].

Условия (2.7), (2.9)-(2.13) связывают между собой все три типа управляющих переменных ($x_{i,j}$, b_j и α_i), придавая согласованный характер решениям основных интерфейсных задач, связанных с организацией очередей, их обслуживанием, определением порядка обслуживания пакетов в разных очередях, распределения пропускной способности интерфейса между очередями и превентивного ограничения длины очереди.

Развивая и дополняя идеи концепции Traffic Engineering Queues [12, 13, 74], регламентирующей вопросы обеспечения сбалансированной загрузки очередей, в структуру модели введем ряд дополнительных условий-ограничений

$$k_j^0 \cdot \bar{n}_j \leq \beta \cdot n_j^{\max}, \quad (j = \overline{1, N}), \quad (2.14)$$

где β – верхний динамически управляемый порог загруженности очередей по их длине на интерфейсе маршрутизатора. Физический смысл условий (2.14) состоит в том, чтобы очереди, создаваемые на интерфейсе, загружались сбалансированно. При этом, чем выше класс очереди (k_j^0), тем меньшую длину она должна иметь.

По сути β – дополнительная управляющая переменная, на которую накладываются ограничения следующего вида:

$$0 \leq \beta \leq 1. \quad (2.15)$$

2.3. Выбор критерия оптимальности управления очередями на интерфейсах маршрутизаторов транспортной телекоммуникационной сети

Важным требованием к перспективным решениям в области управления очередями является то, чтобы они обеспечивали нахождение решения, способствующего повышению уровня качества обслуживания при минимизации используемого канального и буферного ресурса. Это может быть достигнуто лишь в рамках оптимизационной постановки задачи управления очередями, важным этапом которой является выбор критерия оптимальности получаемых решений. Традиционно выбор критерия осуществляется в ходе нахождения компромисса между максимальным учетом физики моделируемых процессов и возможностью получения искомых решений с приемлемой точностью и, что немаловажно, в реальном времени.

В работах [9-11] в качестве критерия оптимальности предлагается использовать минимум стоимостной функции, характеризующей условную стоимость решения отдельных интерфейсных задач, связанных, например, с распределением потоков по очередям или выделениям им пропускной способности интерфейса. Наряду с высокой наглядностью данный подход во многом зависит от порядка назначения метрики использования очередей или выделения им пропускной способности интерфейса.

В работах [12, 14, 74] предложены критерии, основанные на обеспечении сбалансированного использования очередей, в т.ч. с учетом их приоритетов [13]. Однако использование подобного подхода в ряде случаев сопровождалось тем, что сбалансированность загруженности отдельных очередей могла обеспечиваться за счет перенаправления потоков пакетов с загруженных очередей в менее загруженные без учета

их различных требований к дифференцированному уровню качества обслуживания.

В этой связи в данной работе согласованный расчет управляющих переменных $x_{i,j}$, b_j , α_i и β будет обеспечен в ходе решения оптимизационной задачи, связанной с минимизацией целевой функции вида:

$$F = \sum_{i=1}^M \sum_{j=1}^N h_{i,j}^x \cdot x_{i,j} + \sum_{j=1}^N h_j^b \cdot b_j + \sum_{i=1}^M h_i^\alpha \cdot \alpha_i + h^\beta \cdot \beta, \quad (2.16)$$

где $h_{i,j}^x$ – условная стоимость (метрика) обслуживания пакетов i -го потока с помощью j -й очереди;

h_j^b – условная стоимость (метрика) выделения единицы пропускной способности интерфейса пакетам из j -й очереди;

h_i^α – условная стоимость (метрика) отказов в обслуживании пакетам i -го потока;

h^β – условная стоимость реализации процесса балансировки длин очередей на интерфейсе маршрутизатора.

Таким образом, общее число управляющих переменных $x_{i,j}$, b_j , α_i и β составит $M \cdot N + N + M + 1$.

Физический смысл сформулированной задачи в целом состоит в том, что расчет управляющих переменных должен приводить к минимизации суммарной стоимости использования сетевых ресурсов: первое слагаемое отвечает за порядок использования буфера очереди (задача Congestion Management), второе – пропускной способности интерфейса (задача Resource Allocation), а третье – за решение задачи Congestion Avoidance. Четвертое слагаемое отвечает за качество решения всех задач в целом.

Важно отметить, что критерием направления того или иного потока в определенную очередь является максимальное совпадение (соизмеримость) их классов $k_i^{\mathbf{n}}$ и $k_j^{\mathbf{o}}$. Тогда в рамках усовершенствования модели, описанной в работах [12-14, 18] предлагается, чтобы метрика, отвечающая за распределение потоков пакетов по очередям, а при выполнении условия (2.3) и за агрегирование потоков, рассчитывалась согласно формуле

$$h_{i,j}^x = w_x^b (k_i^{\mathbf{n}} - k_j^{\mathbf{o}})^2 + 1, \quad (i = \overline{1, M}, j = \overline{1, N}). \quad (2.17)$$

Метрика $h_{i,j}^x$ является неотрицательной величиной и напрямую зависит от квадрата расстояния между классами отдельных потоков и очередей. С помощью положительного параметра w_x^b можно регулировать влияние на конечное численное значение целевой функции (2.16) ее первого слагаемого.

Предложенная модель с вновь вводимыми формализмами (2.2), (2.7), (2.14), (2.16) и (2.17) позволяет обеспечить согласованное решение задач по агрегированию и распределению потоков по очередям, по распределению пропускной способности интерфейса между системой поддерживаемых на нем очередей, а также организовать ограничение интенсивности потоков, поступающих на интерфейс, в случае возможного переполнения очереди.

2.4. Обзор методов решения сформулированной оптимизационной задачи управления очередями в рамках предложенной модели

В подразделе 2.3 сформулирована оптимизационная задача по управлению очередями на интерфейсе маршрутизатора, которая относится

к классу задач нелинейного программирования, т.к. система условий (2.5), (2.6) и (2.15), накладываемых на управляющие переменные, относится к *линейным* ограничениям, а условия (2.7), (2.9)-(2.14), связывающие между собой величины введенных управляющих переменных, относятся к *нелинейным* ограничениям. Выбранная целевая функция (2.16), подлежащая минимизации, является линейной формой.

Стоит также учесть, что переменные (2.4), отвечающие за решение задачи по распределению пакетов потоков между сформированными очередями (задачи Congestion Management), носят булевый характер; переменные (2.6), отвечающие за решение задачи по распределению пропускной способности интерфейса между очередями (задачи Resource Allocation), вещественные; переменные (2.8), отвечающие за решение задачи по ограничению интенсивности потоков, поступающих на интерфейс, (задачи Congestion Avoidance), вещественные; переменная (2.15), определяющая верхний динамически управляемый порог загрузки очередей на интерфейсе маршрутизатора, также вещественна. Поэтому оптимизационная задача с критерием (2.16) и ограничениями (2.5)-(2.7), (2.9)-(2.15) в силу перечисленных особенностей относится к подклассу задач смешанного целочисленного нелинейного программирования (Mixed Integer Nonlinear Programming, MINLP).

Несмотря на достаточную сложность подобного класса задач, определяемую, прежде всего, нелинейностью ограничений и булевым характером части искомым переменных, для их решения учеными разработано ряд эффективных методов и алгоритмов, среди которых методы округления, «ветвей и границ», имитации отжига, последовательной линеаризации, множителей Лагранжа, штрафных функций (штрафа), а также их многочисленные комбинации [75-81].

Особенностью метода округления является то, что при округлении получаемого решения, как правило, к ближайшему целому, необходимо

контролировать суммарную погрешность решения, что особенно важно при работе с большим числом переменных, что имеет место в рассматриваемой оптимизационной задаче. Кроме того, зачастую округление к целому значению может привести к нарушению условий-ограничений, накладываемых на искомые переменные, что приведет к утрате их физического смысла, снижению достоверности получаемых решений и потере адекватности используемой модели в целом.

При использовании комбинаторного метода «ветвей и границ», основанного на последовательном исключении заведомо неоптимальных вариантов решений, важно обладать эффективной процедурой выбора и постоянного перерасчета верхних и нижних границ значений целевой функции. От этого зависит суммарный объем вычислений и время решения задачи в целом, определяя тем самым требования к производительности спецвычислителя маршрутизатора ТКС для обеспечения получения искомого результата по управлению очередями в реальном масштабе времени даже при настройке множества очередей и наличии большого числа разнотипных потоков, т.е. при высокой размерности оптимизационной задачи.

Для поиска оптимальных значений рассчитываемых переменных в условиях наличия нелинейных ограничений применяются метод имитации отжига (Simulated Annealing) и метод последовательной линеаризации. Первый метод основан на случайном выборе направления поиска, что должно привести к более оперативному пошаговому поиску глобального оптимума на фоне множества локальных минимумов, чем при использовании обычного градиентного поиска. Однако случайный характер выбора направления поиска может и привести к обратным результатам: процесс расчета может затянуться во времени, особенно для задач высокой размерности.

При использовании метода последовательной линеаризации (особенно в сочетании с другими численными методами), который во многом является аналогом метода наискорейшего спуска, в ряде случаев удается получить искомое оптимальное решение с более высокой точностью или сократив общее время расчетов.

Традиционно для решения задач на условный оптимум (в данном случае минимум) используются методы множителей Лагранжа и штрафных функций (метод штрафа). Метод множителей Лагранжа основан на введении дополнительных переменных, количество которых зависит от числа ограничений в задаче, что способствует росту результирующей размерности оптимизационной задачи и времени ее решения. При использовании метода штрафных функций размерность исходной оптимизационной задачи остается прежней, но процесс поиска оптимума может быть существенно замедлен, т.к. в отличие от постоянно пересчитываемых множителей Лагранжа штрафы за нарушение ограничений остаются прежними. Для обоих методов важным является выбор и задание начальной точки поиска, которая бы удовлетворяла всему множеству введенных в задаче ограничений, что является для размерных задач достаточно сложной проблемой.

В данной работе для решения поставленной оптимизационной задачи из класса MINLP использовался решатель программы «fminconset», функционирующий в среде MatLab. Он хоть и не входит в состав Optimization Toolbox официального релиза MatLab, но активно используется исследователями для решения оптимизационных задач подобного класса. В основу программы «fminconset» положена комбинация различных численных методов, например, метода «ветвей и границ», а также метода последовательного квадратичного программирования. Пример использования программы «fminconset» для

решения сформулированной в подразделе 2.2. задачи управления очередями на интерфейсе маршрутизатора показан на рис. 2.2.

```

1 - clear all;
2 - global Ko Kp N M alfa
3 - clc;
4 - bb=100; % PC интерфейса
5 - M=3; % ocheredi
6 - N=5; % potoki
7 - Aeq(1:N,1:M*N)=0;
8 - for i=1:N
9 -     Aeq(i,M*(i-1)+1:M*(i-1)+M)=1;
10 - end
11 - Aeq=[Aeq zeros(N,M) zeros(N,1)]; % условия сохранения потока
12 - Aeq(N+1,:)=[zeros(M*N,1); ones(M,1); 0]'; % условия отсутствия перегрузки КС по PC (сумма равна b)
13 - beq=[ones(N,1);bb];
14 - disp('типы очередей'); Ko=rand(M,1) % тип очереди
15 - disp('типы потоков'); Kp=rand(N,1) % тип потока
16 - ll=0;
17 - for i=1:N
18 -     for j=1:M
19 -         ll=ll+1;
20 -         f(ll)=(Kp(i)-Ko(j))^2;
21 -         f2(j)=1/Ko(j);
22 -     end
23 - end
24 - f=[100*f; ones(M,1); 1]; % если минимизируем используемую PC интерфейса
25 - bN=bb/N; % средняя интенсивность потоков
26 - ap=1.5*bN*rand(N,1); % случайная интенсивность потоков
27 - A=[];
28 - for i=1:N
29 -     A=[A ap(i)*diag(ones(M,1))]; % условия отсутствия перегрузки очереди по PC
30 - end
31 - A=[A diag(-1*ones(M,1)) zeros(M,1)]; % условия отсутствия перегрузки очереди по PC (с упр переменными bj)
32 - A(M+1-M+1+M-1,:)=[zeros(M,M*N) -1*diag(ones(M,1)) bb*ones(M,1)]; % условия балансировки выделения PC
33 - b=[zeros(M,1); zeros(M,1)];
34 - d=ones(M*N,1);
35 - for i=1:(M*N-1)
36 -     d(i+1)=i+1;
37 - end
38 - intcon=d';
39 - lb=zeros(M*N+M+1,1);
40 - ub=[ones(M*N,1); bb*ones(M,1); 1];
41 - [x3,fval3] = intlinprog(f,intcon,A,b,Aeq,beq,lb,ub);
42 - x6 = fminconset(@myfun,x3,A,b,Aeq,beq,lb,ub,@mycon)

```

Рис. 2.2. Пример использования программы «fminconset» для решения сформулированной оптимизационной задачи управления очередями на интерфейсе маршрутизатора класса MINLP

2.5. Методика сведения задачи целочисленной оптимизации к задаче нелинейного программирования

С целью снижения вычислительной сложности получаемых решений по управлению очередями с использованием предложенной модели (2.1)-(2.17) при решении задачи MINLP была предложена методика сведения задачи целочисленной оптимизации к задаче нелинейного программирования. Суть данной методики заключается в том, чтобы отказаться в явном виде от целочисленности переменных, заменив (2.4) на условия вида:

$$0 \leq x_{i,j} \leq 1, \quad (i = \overline{1, M}, j = \overline{1, N}), \quad (2.18)$$

а их булевость удалось обеспечить в ходе совместного выполнения условий (2.15) и дополнительно вводимой системы нелинейных условий (по числу потоков):

$$\sum_{j=1}^N \sum_{\substack{m=1, \\ m \neq j}}^N x_{i,j} x_{i,m} = 0, \quad (i = \overline{1, M}). \quad (2.19)$$

Выполнение условий (2.19) гарантирует, что одновременно пакеты i -го потока не будут направляться в две и более очереди, а условие (2.5) определяет, что этот же поток будет направлен в обязательном порядке лишь в одну из очередей. Таким образом, условия (2.18), (2.5) и (2.19) позволяют формально отказаться от выполнения условий булевости управляющих переменных (2.4), но физический смысл, по-прежнему, остается неизменным.

Ввиду отказа от целочисленных условий (2.4) путем их замены на (2.18) и введения новых условий (2.19) оптимизационная задача управления очередями теперь будет относиться к классу задач

нелинейного программирования, т.к. наряду с линейными ограничениями (2.5), (2.6), (2.15) и (2.18) будут учтены и нелинейные условия (2.7), (2.9)-(2.14) и (2.19). Для решения оптимизационной задачи нелинейного программирования использован пакет Optimization Toolbox среды MatLab [82-84], представленный программой «fmincon». Этот решатель базируется на применении метода последовательного квадратичного программирования (Sequential quadratic programming, SQP) и итерационного метода численной оптимизации BFGS (Бройдена-Флетчера-Гольдфарба-Шанно).

2.6. Исследование процессов управления очередями на маршрутизаторах транспортной телекоммуникационной сети с использованием предложенной модели

В работе проводился анализ процесса управления очередями на интерфейсе маршрутизатора с использованием предложенной модели. В ходе исследования установлено, что на число очередей, которые в ходе оптимальной агрегации потоков не использовались, влиял ряд факторов, среди которых общее число потоков пакетов, количество очередей, организованных на интерфейсе, и закон распределения между ними (потоками и очередями) классов обслуживания.

Общее число потоков и очередей считалось известным. В ходе моделирования число потоков изменялось от 8 до 500, а число очередей – от 4 до 250. Классы очередей детерминированно и равномерно устанавливались в пределах от 1 до 10. Классы потоков определялись по случайному закону также в пределах от единицы до десяти. В работе использовались два закона распределения классов обслуживания между потоками: равномерный и усеченный нормальный. Первый – характерный для случая, когда вероятности появления на интерфейсах маршрутизаторов

мультисервисной сети пакетов того или иного класса приблизительно одинаковы. Усеченный нормальный закон используется по причине того, что классы потоков принимают лишь ограниченный набор значений, и справедлив для случая, когда в сети превалирует определенное множество классов, при этом мощность этого множества задавалась косвенно через дисперсию распределения σ^2 .

2.6.1. Примеры решения задач управления очередями на маршрутизаторах ТКС

Особенности решения поставленной задачи для повышения наглядности излагаемого материала продемонстрируем на двух конкретных примерах. Пусть в табл. 2.2 представлены исходные данные и результаты расчета для первого примера – для двадцати пяти потоков ($M = 25$) и десяти очередей ($N = 10$), пропускной способности интерфейса 100 1/c ($b = 100$) и $w_x^b = 100$. Принималась гипотеза об усеченном нормальном законе распределения классов обслуживания между потоками с $\sigma^2 = 0,3$. Максимальные длины очередей выбирались от 10 до 100 пакетов равномерно по возрастанию: от очереди с минимальным классом до очереди с максимальным классом.

В ходе расчетов средняя длина очереди определялась при условии, что работа интерфейса моделировалась, для примера, системой массового обслуживания М/М/1 (2.10). Минимизируемый верхний динамический порог загрузки очередей составлял $\beta = 0,8172$. Как показали результаты моделирования, агрегация потоков осуществлялась в соответствии с близостью их классов и класса очереди, в которую они направлялись (2.17).

Таблица 2.2

Исходные данные и результаты расчета для первого примера

Характеристика потоков			Параметры очередей				
номер потока	класс потока	средняя интенсивность потока (1/с)	номер используемой очереди	класс очереди	средняя интенсивность потока в очереди (1/с)	выделенная ПС (1/с)	средняя/максимальная длина очереди
10	0,6374	2,7336	1	1	11,7254	11,8671	81,7201/100
14	1,1192	4,1350					
23	0,2286	4,8567					
1	1,5187	5,5946	2	2	24,3278	24,9723	36,7748/90
4	1,7136	6,2721					
8	2,2219	3,8225					
18	2,1990	0,5468					
22	1,8839	4,5273					
25	2,4364	3,5644					
5	3,2998	3,7701	3	3	6,8321	7,1324	21,7923/80
20	2,6624	1,6023					
21	2,9150	1,4596					
2	3,8066	3,4043	4	4	19,3050	20,57717	21,7923/70
11	4,2548	3,9107					
12	4,0434	0,4557					
13	4,0029	5,9555					
15	4,2431	0,6730					
19	3,5408	1,3812					
24	4,4940	3,5245					
-	-	-	5	5	-	-	-
16	6,1355	4,5094	6	6	4,5094	5,0955	6,8101/50
-	-	-	7	7	-	-	-
3	8,2102	5,8870	8	8	11,8539	14,9258	3,0646/30
7	8,0629	5,9669					
-	-	-	9	9	-	-	-
6	9,6647	0,7662	10	10	9,0057	15,4352	0,8172/10
9	9,9977	5,6905					
17	9,8806	2,5490					

Достоинством предложенной модели также является то, что использовалось лишь минимально необходимое число очередей, достаточное для обеспечения дифференцированного обслуживания на интерфейсе.

В рамках первого рассмотренного примера, как показано в табл. 2.2, не задействовались пятая, седьмая и девятая очереди, т.к. не было потоков с соизмеримыми для них классами. Минимизация количества поддерживаемых очередей без снижения уровня QoS приводит к пропорциональному снижению времени, затрачиваемого на обработку пакетов на интерфейсе в конкретной очереди. В рамках рассмотренного примера удалось сократить число используемых очередей на 30%.

В рамках второго примера исходные данные остались прежними, но изменился закон распределения классов обслуживания между потоками на равномерный. Согласно данным, представленным в табл. 2.3, минимизируемый верхний динамический порог загрузки очередей составлял $\beta=0,9963$. Кроме того, из десяти возможных не задействовались седьмая и десятая очереди, т.е. в рамках рассмотренного примера удалось сократить число используемых очередей и время на их обработку на 20%.

Как следует из решений в рамках рассмотренных примеров, предложенная модель обеспечивает комплексное решение основных интерфейсных задач, связанных с агрегированием потоков согласно их классификационным признакам и их распределению по очередям, выделение каждой из очередей пропускной способности с обеспечением сбалансированной их загрузки с учетом класса потоков и очереди.

На каждом конкретном интерфейсе выделяемая пропускная способность очереди фактически определяет время, которое уделяется данной очереди планировщиком пакетов. Чем больше значение выделяемой ПС, тем больше времени из данной очереди планировщик передает пакеты в канал связи на соседний маршрутизатор.

Таблица 2.3

Исходные данные и результаты расчета для второго примера

Характеристика потоков			Параметры очередей				
номер потока	класс потока	средняя интенсивность потока (1/с)	номер используемой очереди	класс очереди	средняя интенсивность потока в очереди (1/с)	выделенная ПС (1/с)	средняя/максимальная длина очереди
10	1,1709	0,8042	1	1	2,7451	2,7724	99,6200/100
15	1,2631	1,9410					
5	1,5041	1,4065	2	2	10,8439	11,0806	44,8328/90
9	1,6042	5,5167					
13	1,8500	3,9207					
3	3,3355	4,2329	3	3	8,0895	8,3833	26,5676/80
6	3,3370	3,8566					
4	3,8974	3,0058	4	4	11,3701	11,9886	17,4351/70
11	3,9855	4,1317					
19	3,8957	4,2326					
14	5,0079	4,7179	5	5	20,8570	22,4758	11,9555/60
20	4,9014	5,7326					
23	4,9839	5,3399					
24	5,2923	5,0646					
7	5,5368	1,1793	6	6	4,8069	5,3291	8,3024/50
8	5,5029	1,2641					
18	5,6435	2,3636					
-	-	-	7	7	-	-	-
1	7,6393	3,7030	8	8	11,8491	14,4498	3,7361/30
12	8,3138	2,7988					
17	7,6665	5,3472					
2	9,4764	4,9650	9	9	17,5831	23,5202	2,2140/20
16	8,6463	0,2750					
21	8,8690	1,7542					
22	9,0505	6,3866					
25	9,0970	4,2023					
-	-	-	10	10	-	-	-

2.6.2. Анализ влияния характеристик потоков и очередей на результат решения задачи по управлению очередями в рамках предложенной модели

Как было отмечено выше, на число незадействованных в ходе агрегации потоков очередей влияло общее число организованных на интерфейсе очередей (N), суммарное число потоков (M), а также закон распределения классов обслуживания между этими потоками. К рассмотрению принималось два закона распределения классов между потоками: равномерный и усеченный нормальный. В ходе моделирования для усеченного нормального закона распределения классов изменялась дисперсия распределения σ^2 , которая определяла фактически преобладание в сетевом трафике потоков определенных классов.

При распределении классов обслуживания между потоками оценивалось влияние соотношения числа потоков и очередей (M/N) на процент незадействованных очередей в ходе оптимальной агрегации потоков в рамках предложенной модели:

$$P_{\%} = \frac{N_{нз}}{N} \cdot 100\%, \quad (2.20)$$

где $N_{нз}$ – число незадействованных очередей, получаемое в результате решения задачи управления очередями с оптимальной агрегацией потоков.

Судя из рис. 2.3, в условиях принятия гипотезы о равномерном распределении классов обслуживания потоков применение модели (2.1)-(2.19) позволило таким образом агрегировать потоки по очередям, чтобы уменьшить число задействованных очередей, а значит и время на их обработку, в среднем от 5-7% до 20-25%.

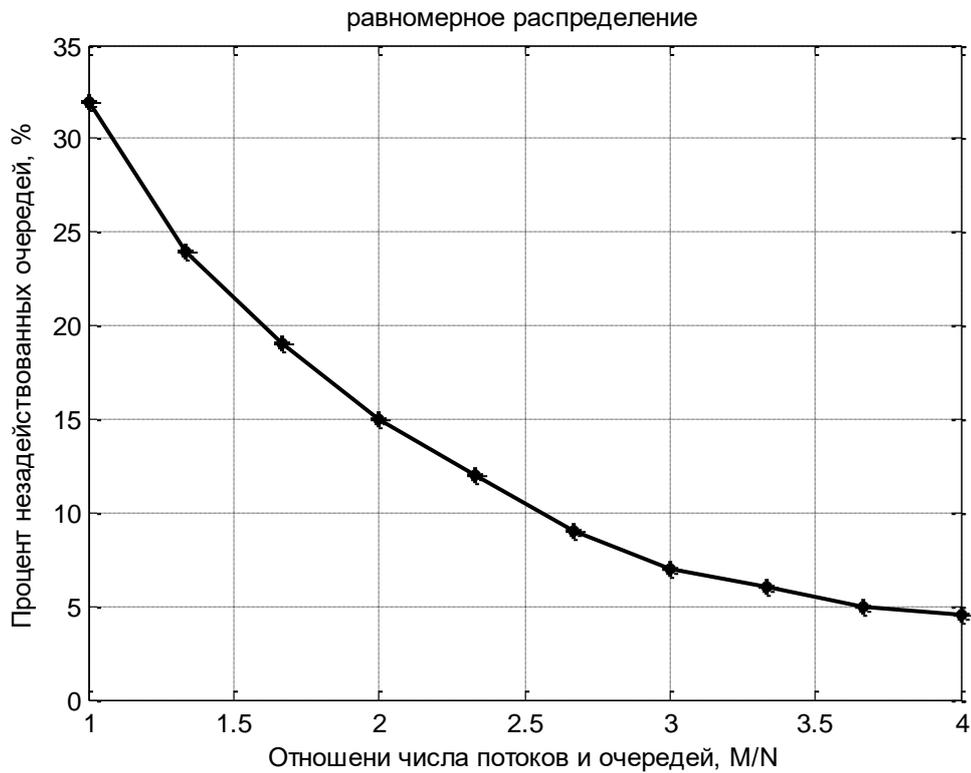


Рис. 2.3. Результаты расчета процента недействующих очередей в ходе оптимальной агрегации потоков в рамках предложенной модели (при равномерном законе распределения классов обслуживания между потоками)

В случае, если принять гипотезу об усеченном нормальном распределении классов обслуживания между очередями, то результаты решения задачи управления очередями с точки зрения расчета процента недействующих очередей представлен на рис. 2.4. На этом рисунке показаны результаты расчета для различных вариантов дисперсии распределения $\sigma^2 = 0,3$, $\sigma^2 = 1$ и $\sigma^2 = 5$.

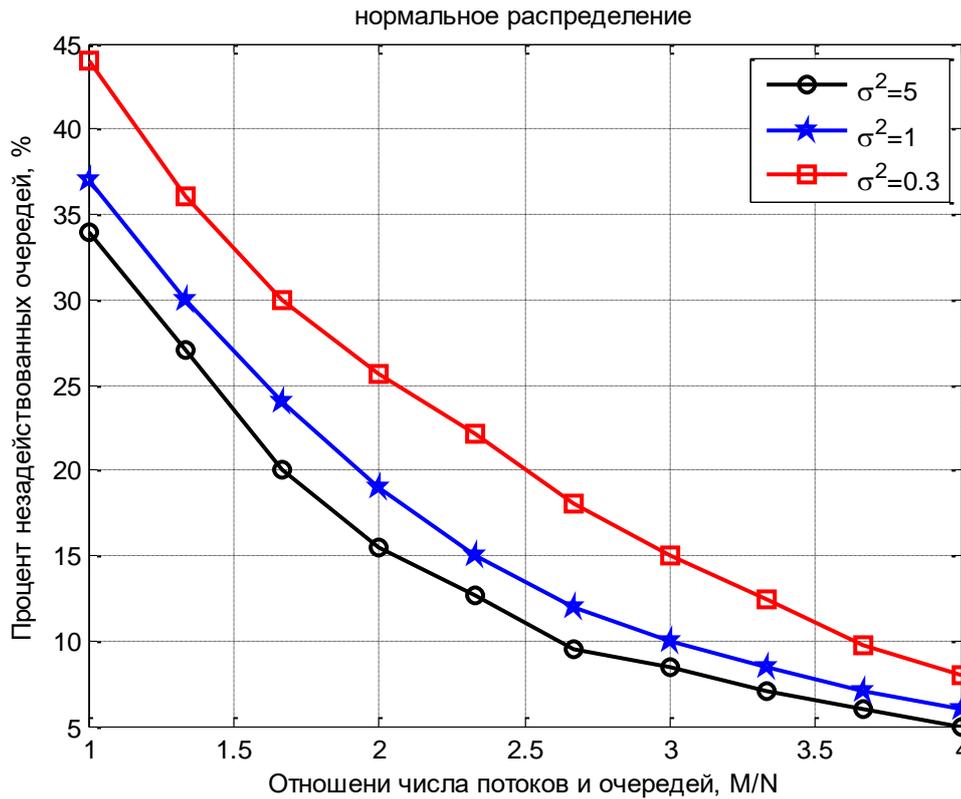


Рис. 2.4. Результаты расчета процента неактивных очередей в ходе оптимальной агрегации потоков в рамках предложенной модели (при нормальном законе распределения классов обслуживания между потоками)

При условии принятия гипотезы об усеченном нормальном распределении классов обслуживания потоков, т.е. когда на интерфейсе наблюдается преобладание пакетов одних классов над другими, применение модели (2.1)-(2.19) позволило таким образом агрегировать потоки по очередям, что уменьшилось число задействованных очередей и время на их обработку в среднем

от 12-15% до 30-35% при $\sigma^2 = 0,3$;

от 8-12% до 28-32% при $\sigma^2 = 1$;

от 7-10% до 25-30% при $\sigma^2 = 5$.

При дальнейшем росте дисперсии распределения полученные результаты все ближе соответствовали решениям, полученным при использовании равномерного закона распределения классов обслуживания между потоками.

2.7. Выводы по второму разделу

1. Как показывает практика, в ходе управления очередями при настройке интерфейса маршрутизатора исходят из двух противоречивых требований. С одной стороны, число используемых очередей на интерфейсе должно быть максимальным и в пределах стремиться к количеству передаваемых через него потоков для обеспечения максимально возможной дифференциации в обслуживании. Но, с другой стороны, с ростом числа поддерживаемых очередей заметно растут и задержки, вносимые планировщиком пакетов, которые связаны с формированием очередей, анализом их состояния и т.д. Рациональный выход из создавшейся ситуации видится в том, чтобы потоки агрегировались по ряду ключевых параметров, связанных с их классификационными признаками и требованиями к уровню качества обслуживания. Это нацелено на обоснованное сокращение числа используемых очередей с обеспечением приблизительно равного уровня качества обслуживания для агрегированных потоков в одной очереди.

2. В этой связи получила дальнейшее развитие потоковая модель управления очередями на маршрутизаторах ТКС на основе оптимального агрегирования потоков и распределения пакетов по очередям. Новизна модели состоит в том, что при распределении потоков по очередям осуществляется их агрегирование на основе сравнения классов потоков и очередей в ходе анализа множества классификационных признаков. Кроме того, в модели сохранено важное свойство относительно реализации

требований концепции *Traffic Engineering Queues* и связанное с обеспечением сбалансированной загруженности очередей по их длине.

3. Использование предложенной модели позволяет оптимизировать процесс агрегирования потоков, распределения пакетов по отдельным очередям и выделения очередям пропускной способности интерфейса, придав решению этих задач высокий уровень согласованности. В основу полученных решений по управлению очередями положена оптимизационная задача класса смешанного целочисленного нелинейного программирования – MINLP, т.к. управляющие переменные носят как целочисленный (2.4), так и вещественный характер (2.6) и (2.8), а среди накладываемых на них ограничений есть как линейные (2.5), (2.6) и (2.15), так и нелинейные условия (2.7), (2.9)-(2.14).

4. Выбранная целевая функция (2.16), подлежащая минимизации, является линейной формой и определяет условную стоимость согласованного решения задач по распределению пакетов потоков между сформированными очередями (Congestion Management), по распределению пропускной способности интерфейса между очередями (Resource Allocation), по ограничению интенсивности потоков, поступающих на интерфейс (Congestion Avoidance).

5. Проведен анализ возможных методов решения сформулированной оптимизационной задачи по управлению очередями в рамках предложенной модели (2.1)-(2.17). Предложена методика сведения задачи целочисленной оптимизации к задаче нелинейного программирования с целью снижения вычислительной сложности получаемых решений по управлению очередями без потери их точности и достоверности. Показано, что для решения оптимизационных задач данного класса целесообразно использовать возможности пакета Optimization Toolbox среды MatLab, а именно программ «fminconset» и «fmincon».

6. Результаты исследований подтвердили работоспособность предложенной модели, а также ее эффективность с точки зрения корректного решения задач по управлению очередями. Оптимизация процесса агрегирования потоков позволяет в зависимости от соотношения числа потоков и очередей, а также их классов сократить количество поддерживаемых очередей от 15-18% до 25-33% без снижения уровня дифференциации качества обслуживания потоков пользователей, что способствует сокращению времени на обработку пакетов на интерфейсе и минимизации межконцевой задержки пакетов в сети в целом.

РАЗДЕЛ 3

РАЗРАБОТКА ПОТОКОВОЙ МОДЕЛИ УПРАВЛЕНИЯ ОЧЕРЕДЯМИ НА ИНТЕРФЕЙСАХ МАРШРУТИЗАТОРОВ АКТИВНОЙ СЕТИ

3.1. Анализ функциональных особенностей построения активной сети

3.1.1. Обзор целей и задач активных сетей

Обзор известных алгоритмов управления очередями [6, 7], к числу которых относятся, прежде всего, алгоритмы обслуживания очередей CQ, CBQ, PQ, FQ и различные их модификации (WFQ, CBWFQ, HWFQ, LLQ и др.), алгоритмы ограничения очередей RED и WRED, позволил констатировать ряд присущих им недостатков, отмеченных в первом разделе. К их числу стоит отнести также реализацию пассивной стратегии обработки пакетов, сводящейся к их элементарной пересылке с входного порта сетевого узла на выходной. Таким образом, время, в течение которого пакет «простаивает» в очереди, не используется в целях его обработки.

Основные недостатки существующих средств управления очередями обусловлены тем, что в основу большинства существующих и вновь проектируемых ТКС положена парадигма «Stupid network», предполагающая «вынос» большинства «интеллектуальных» функций сети на ее периметр, т.е. на приграничные сетевые узлы, а то и на серверы сетевой операционной системы SDN. Подобную ситуацию также наглядно демонстрирует технология многопротокольной коммутации меток (MPLS) [4], в рамках которой к функциям приграничных узлов LER (Label Edge Router), прежде всего, относятся классификация, маркировка и

назначение приоритетов пакетам того или иного потока; маршрутизация от источника с расчетом таблиц и путей коммутации меток (Label Switch Path, LSP) и др. В функции же внутрисетевых (транзитных) узлов (Label Switch Router, LSR) входит лишь решение задач по продвижению пакетов вдоль предварительно рассчитанных на приграничных узлах путей (LSP) на основе анализа таблиц коммутации меток, при этом содержание передаваемых данных обычно игнорируется.

Придать большую гибкость и адаптивность процессам управления сетевым ресурсом в ТКС призвана технология активных сетей (Active Network, AN) [85-89], в рамках которой «интеллект» сети равномерно распределен по всем сетевым узлам – активным узлам. В свою очередь, функции транзитных узлов значительно расширяются за счет возможности решения ими таких важных сетевых задач, как адаптивное изменение приоритетов, фрагментации (дефрагментации), сжатия и маршрутизации пакетов, шифрования трафика и пр. Одним из примеров реализации на практике базовых идей построения AN является построение беспроводных самоорганизующихся Ad-Hoc сетей, в которой узлы сети могут выполнять одновременно функции терминала, коммутационного оборудования, устройства управления сетью и сервера предоставляемых услуг, осуществляющего обработку пакетов.

Основными целями и потенциальными преимуществами реализации концепции построения активных сетей являются следующие [85-89]:

- возможность более оперативного создания и внедрения новых сервисов, минуя или минимизируя достаточно продолжительный этап стандартизации, что особенно важно с точки зрения сетевого сервис-провайдера и достигается за счет динамического изменения состояния сетевых узлов и функций обработки пакетов на них;

- предоставление возможности пользователям или вспомогательным системам на более низком уровне детализации создавать и внедрять сервисы под их специфичные приложения, под текущие сетевые условия функционирования и состояния сети;

- повышение масштабируемости и гибкости процессов обеспечения качества обслуживания при предоставлении распределенных сервисов;

- предоставление эффективной платформы для исследования и тестирования новых сетевых сервисов, в т.ч. в реальном масштабе времени, минимизируя негативное влияние на рабочий сетевой сервис.

Заявленные цели и задачи в AN реализуются на активных узлах (АУ), как правило, с помощью протокола инкапсуляции (Active Network Encapsulation Protocol, ANEP) в рамках среды выполнения (Execution Environments, (EEs) и операционной системы узла (Node Operating System, NodeOS) (рис. 3.1) [85-88].

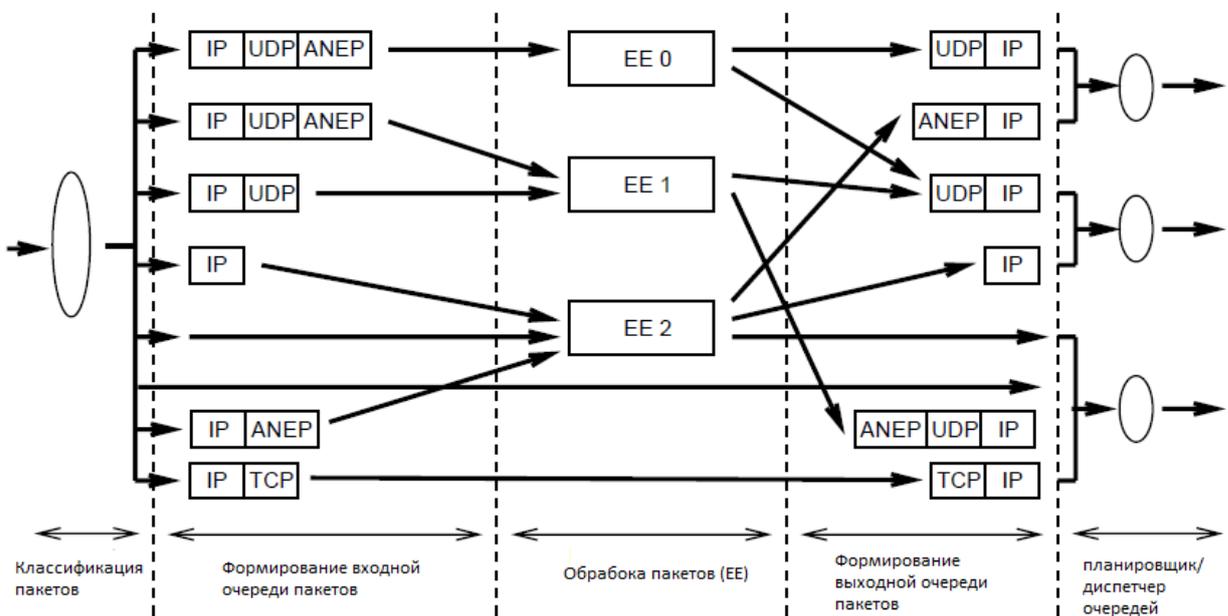


Рис. 3.1. Порядок обработки пакетов на активном узле

3.1.2. Анализ архитектуры активного узла

Основным элементом сетевого узла (рис.3.2), поддерживающим технологию AN и выполняющим задачи ЕЕs, является активный процессор (АП), в функции которого дополнительно входит анализ управляющей информации APCI (Active Processing Control Information) в заголовке каждого пакета, содержащей идентификатор функции и дескриптор (рис. 3.3) [89]. Идентификатор определяет вызываемую функцию обработки, а дескриптор указывает на параметры состояния активного узла, которые должны использоваться в процессе обработки.

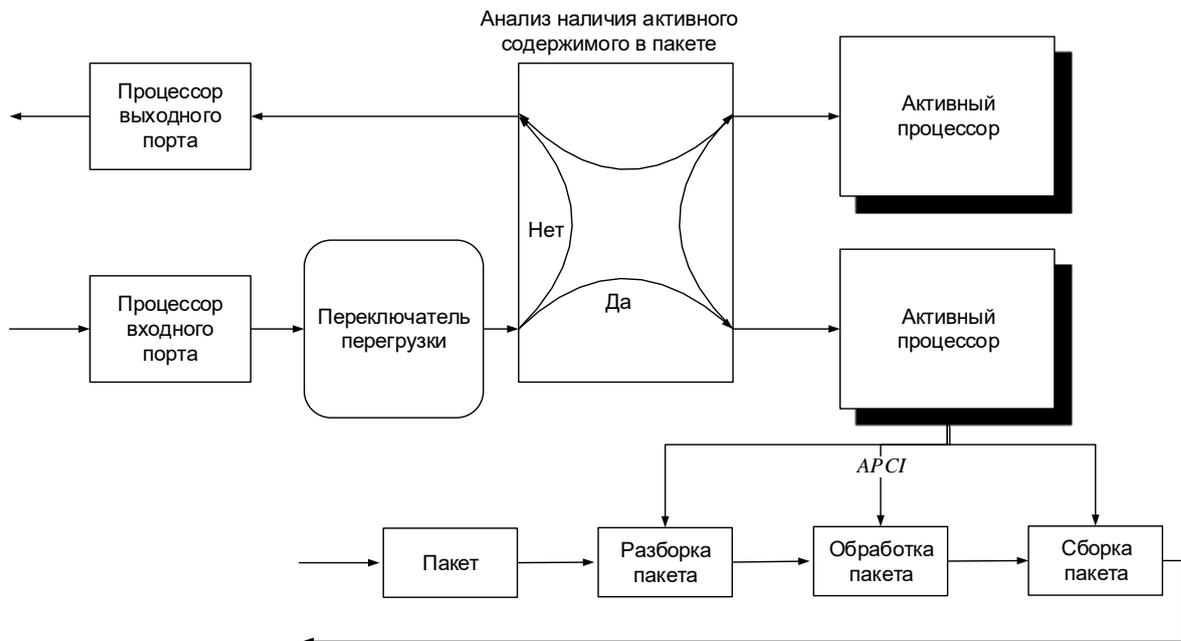


Рис. 3.2. Варианты обработки пакета на активном узле AN

Таким образом, в функции активного узла входит (рис. 3.1) [85-89]:

- во-первых, определение выходного порта,
- во-вторых, в случае наличия управляющей информации пакет переадресуется активному процессору,

- в-третьих, в соответствии с содержимым APCI вызывается функция обработки с параметрами, соответствующими текущему состоянию сети,
- в-четвертых, с обновленным заголовком пакет передается на следующий узел.

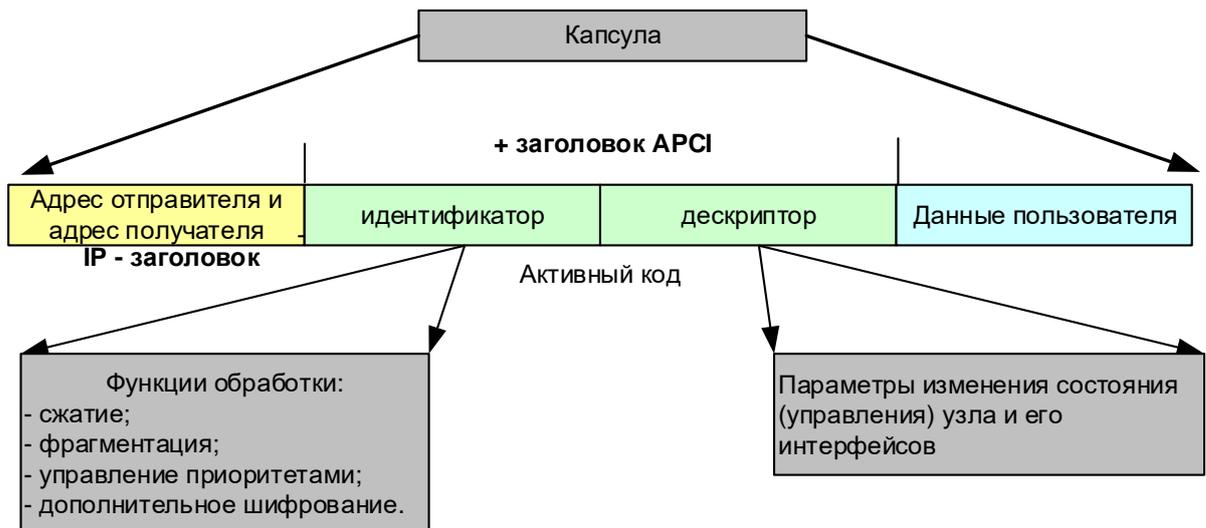


Рис. 3.3. Формат капсулы с активным кодом

Примерами функций, выполняемых на активных узлах сети, могут служить буферизация пакетов, управление пропускной способностью выходных интерфейсов в соответствии с их загрузкой, дополнительное сжатие трафика до передачи его по перегруженному участку сети, шифрование с целью повышения уровня безопасности передаваемых данных и т.д.

Характерной особенностью активного узла (рис.3.2) также является использование многопроцессорной архитектуры ввиду высоких требований к скорости обработки пакетов, которая может достигать до десятков терабит в секунду [90, 91]. Современные высокопроизводительные маршрутизаторы обычно являются гигабитными с возможностью масштабирования до терабитного диапазона. На

сегодняшний день терабитные скорости обеспечивает, например, маршрутизатор TSR-40 от компании Avici – 5,6 Тбит/с. Одно шасси маршрутизатора TeraPlex 20 производства компании Pluris обеспечивает пропускную способность в 150 Гбит/с с масштабированием до суммарной величины 1,44 Тбит/с, а Pluris 2000 масштабируется до 149 Гбит/с в случае одного устройства и до 19.2 Тбит/с при объединении нескольких устройств. Маршрутизирующая платформа Routing Core Platform 7770 компании Alcatel поддерживает скорость до 640 Гбит/с на шасси, причем восемь шасси можно объединить в единое логическое устройство, в результате чего общая производительность составит 20 Тбит/с. Модель 8812 компании Procket достигает, как утверждают разработчики, производительности 960 Гбит/с и, по предварительной оценке, 1,2 млрд пакетов в секунду [90, 91]. Это достигается за счет применения специализированных интегральных схем ASIC (Application-Specific Integrated Circuits), а также микросхем сверхбольшой степени интеграции VLSI (Very Large-Scale Integration), которые являются полностью программируемыми сетевыми процессорами. За счет применения специальных микросхем ASIC маршрутизатор в состоянии обрабатывать на каждой линейной карте миллионы новых потоков в секунду.

Таким образом, механизмы управления очередями на интерфейсах активных узлов должны учитывать перечисленные технологические особенности, свойственные структуре и функциям узлов активной сети, что подразумевает пересмотр самих математических моделей и методов управления в сторону повышения степени их адаптируемости к возможным изменениям характеристик трафика, требований к качеству обслуживания, составу и содержанию реализуемых сетевых сервисов.

3.2. Математическая модель управления очередями на интерфейсах маршрутизаторов активной сети

Условимся, что основу архитектуры сетевого узла образуют K параллельно функционирующих активных процессора, тогда под c_k ($k = \overline{1, K}$) обозначим номинальную производительность k -го АП ($1/c$). Активные процессоры могут быть как однородными, реализуя одну и ту же функцию обработки пакетов на активном узле, так и разнородными, когда за обеспечение выполнения той или иной обрабатываемой функции отвечает некоторое множество активных процессоров.

Тогда с целью математического описания постановки и решения задачи управления очередями на узлах активной сети, расширяя смысл выражений (2.4), введем множество управляющих переменных $x_{i,j}^k$, каждая из которых характеризует долю i -го потока пакетов, который будет направлен для обслуживания в j -ю очередь через k -й активный процессор. Согласно физическому смыслу переменной $x_{i,j}^k$, имеют место следующие условия:

$$x_{i,j}^k \in \{0;1\} \quad (i = \overline{1, M}, j = \overline{1, N}, k = \overline{1, K}). \quad (3.1)$$

Ввиду наличия в структуре активного узла множества активных процессоров условия сохранения потока на интерфейсе маршрутизатора (2.5) несколько модифицируются:

$$\sum_{k=1}^K \sum_{j=1}^N x_{i,j}^k = 1, \quad (i = \overline{1, M}), \quad (3.2)$$

гарантируя, что пакеты каждого i -го потока будут направлены лишь в одну очередь и только через один активный процессор.

Для обеспечения управляемости процессом предотвращения перегрузки (Congestion Avoidance) интерфейса условия (2.7) модифицируются к виду:

$$\sum_{i=1}^M \sum_{k=1}^K a_i (1 - \alpha_i) x_{i,j}^k < b_j, (j = \overline{1, N}). \quad (3.3)$$

Смысл условий (3.3) состоит в том, чтобы суммарная интенсивность потоков, направленных через любой из активных процессоров на обслуживание в j -ю очередь, не превышала пропускную способность интерфейса, которая выделена данной очереди.

На переменные b_j ($j = \overline{1, N}$), определяющие размер пропускной способности интерфейса, выделенной для обслуживания пакетов j -й очереди, и переменные α_i ($i = \overline{1, M}$), характеризующие долю i -го потока, получившего отказ в обслуживании, также накладывается система ограничений (2.6) и (2.8) соответственно их физическому смыслу.

С целью предотвращения перегрузки интерфейса условия обеспечения управляемости этим процессом (3.3) также дополняются условиями предотвращения перегрузки очередей по их длине (2.9) совместно с (2.10)-(2.13) с той лишь оговоркой, что коэффициент использования каждой j -й очереди рассчитывается согласно выражению:

$$\rho_j = \frac{\sum_{i=1}^M \sum_{k=1}^K a_i (1 - \alpha_i) x_{i,j}^k}{b_j}, (j = \overline{1, N}). \quad (3.4)$$

Условия (3.3), (2.9)-(2.13), (3.4) отражают взаимосвязь всех трех типов управляющих переменных ($x_{i,j}^k$, b_j и α_i), тем самым обеспечивая высокий уровень согласованности при решении основных интерфейсных

задач, связанных с организацией очередей, их обслуживанием, определением порядка обработки пакетов активными процессорами и обслуживания в разных очередях, распределения пропускной способности интерфейса между отдельными очередями и превентивного ограничения длины очереди.

Важным дополнением модели управления очередями на интерфейсе активного узла является условие предотвращения перегрузки активных процессоров по производительности

$$\sum_{i=1}^M \sum_{j=1}^N a_i (1 - \alpha_i) x_{i,j}^k \leq c_k, \quad (k = \overline{1, K}). \quad (3.5)$$

Смысл данных условий заключается в том, чтобы результирующая пакетная нагрузка на каждый из активных процессоров не превышала его производительность. Для обеспечения сбалансированной загруженности очередей, как и в случае модели, рассмотренной во втором разделе, вводятся условия (2.14).

На рис. 3.4 представлена функциональная архитектура интерфейса активного узла с обозначениями характеристик потоков, параметров активных процессоров и очередей, а также управляющих переменных, введенных в рамках модели (3.1)-(3.5), (2.6), (2.8)-(2.14).

В связи с тем, что в общем случае вариантов выбора $x_{i,j}^k$, b_j , α_i и β в рамках ограничений (3.1)-(3.5), (2.6), (2.8)-(2.14) может оказаться достаточно много, то целесообразно сформулировать задачу расчета этих управляющих переменных в оптимизационной постановке.

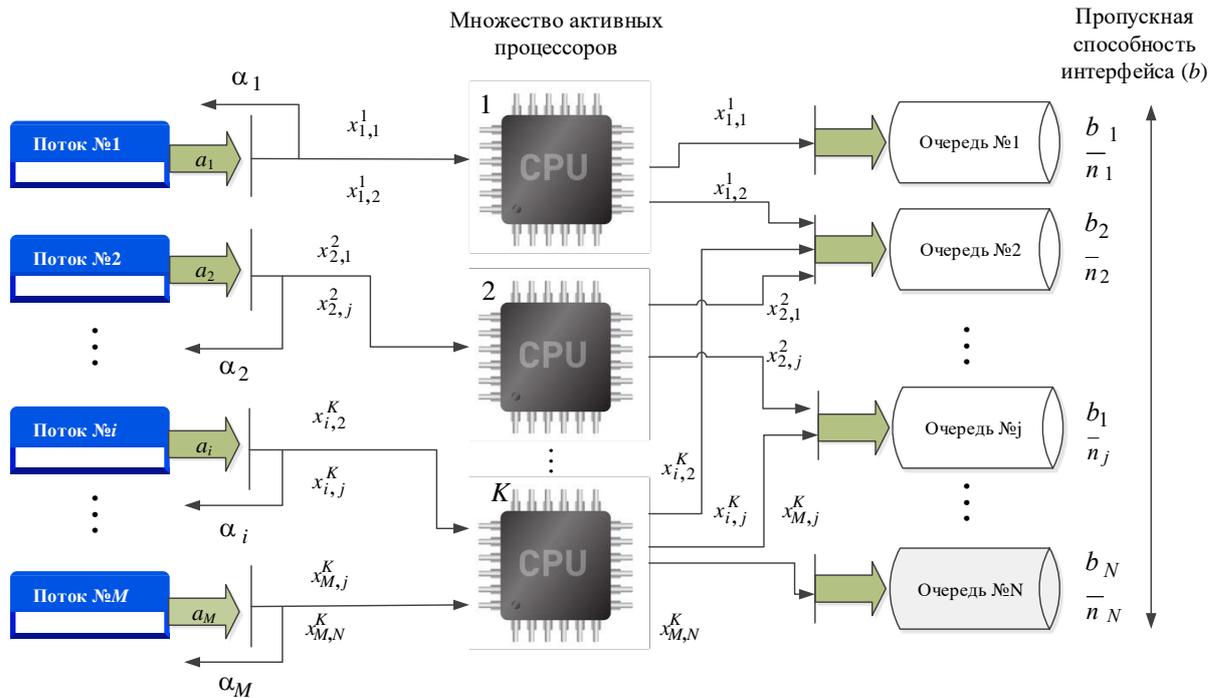


Рис. 3.4. Функциональная архитектура интерфейса активного узла с однородной архитектурой

В этой связи, в данной работе согласованный расчет управляющих переменных будет обеспечен в ходе решения оптимизационной задачи, связанной с минимизацией целевой функции вида:

$$F = \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^K h_{i,j}^k \cdot x_{i,j}^k + \sum_{j=1}^N h_j^b \cdot b_j + \sum_{i=1}^M h_i^\alpha \cdot \alpha_i + h^\beta \cdot \beta, \quad (3.6)$$

в которой $h_{i,j}^k$ – условная стоимость (метрика) обслуживания пакетов i -го потока с помощью k -го активного процессора и j -й очереди интерфейса активного узла. Метрики $h_{i,j}^k$ рассчитываются с учетом выражения (2.17) и могут быть представлены в виде:

$$h_{i,j}^k = h_k w_x^b (k_i^\Pi - k_j^0)^2 + 1, \quad (i = \overline{1, M}, j = \overline{1, N}, k = \overline{1, K}), \quad (3.7)$$

где h_k – условная стоимость (метрика) использования k -го активного процессора на рассматриваемом активном узле.

В случае, если активный узел имеет однородную архитектуру (рис. 3.5), включая в себя активные процессоры лишь одного типа, т.е. реализующих одну и ту же функцию обработки пакетов, то общее число управляющих переменных $x_{i,j}^k$, b_j , α_i и β составит $M \cdot N \cdot K + N + M + 1$.

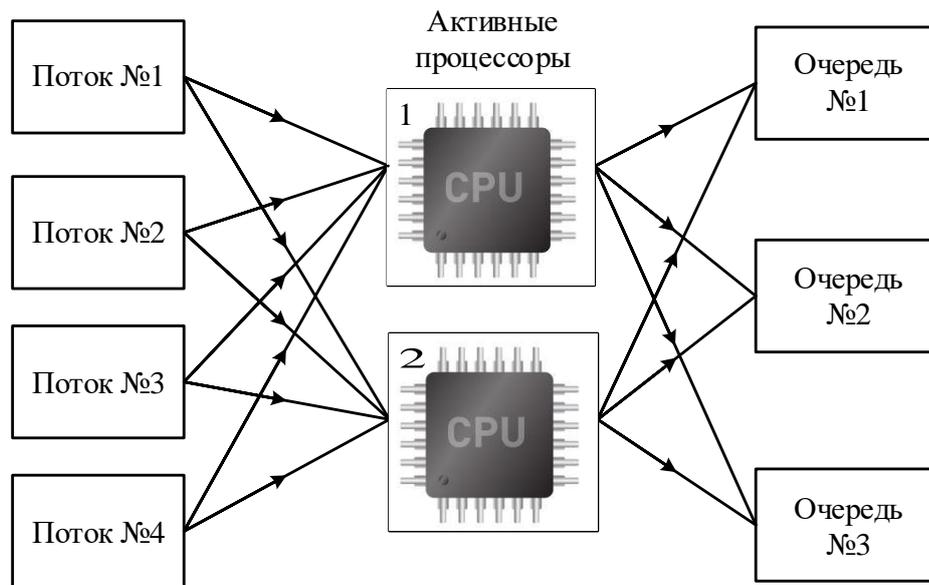


Рис. 3.5. Упрощенная схема активного узла с однородной архитектурой

Если поддерживаемых активным узлом функций обработки пакетов больше одной, например, шифрование и сжатие данных, то в его гибридной архитектуре будут присутствовать активные процессоры различных типов (рис. 3.6).

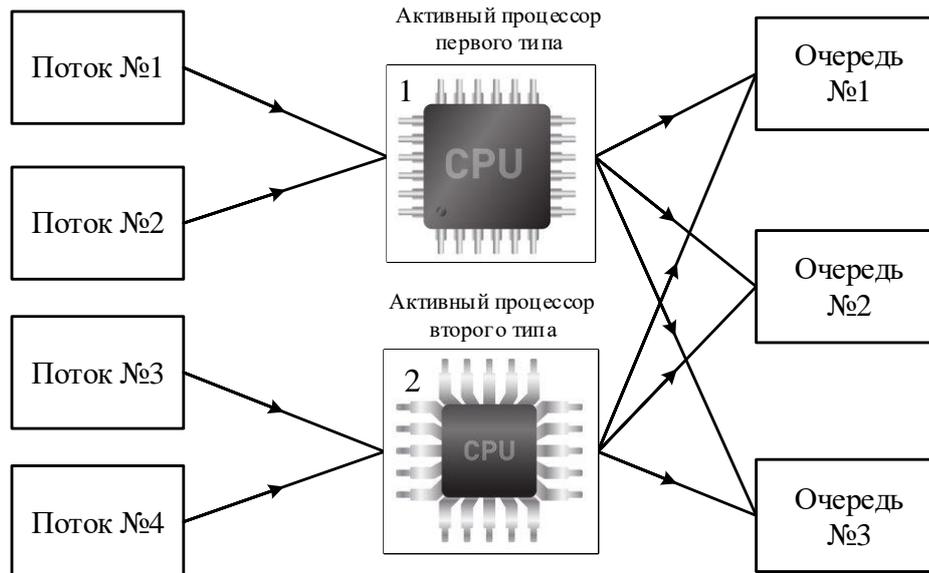


Рис. 3.6. Упрощенная схема активного узла с гибридной архитектурой

В зависимости от типа процессора реализуемая им функция обработки может быть применена, как правило, к некоторому множеству потоков пакетов. Тогда размерность оптимизационной задачи по переменной $x_{i,j}^k$ будет меньше $M \cdot N \cdot K$, определяясь полностью множеством типов активных процессоров и их возможностями по обработке пакетов конкретных потоков. Например, на рис. 3.6 показан случай, когда активный процессор №1 должен реализовать функции по обработке пакетов лишь первого и второго потоков, а активный процессор №2 – пакетов третьего и четвертого потоков.

В качестве эффективного варианта решения задачи, связанной с выбором критерия оптимальности управления очередями на интерфейсе активного узла, может выступать минимум целевой функции

$$F = \sum_{i=1}^M \sum_{j=1}^N \sum_{k=1}^K h_{i,j}^k \cdot x_{i,j}^k + \sum_{j=1}^N h_j^b \cdot b_j + \sum_{i=1}^M h_i^\alpha \cdot \alpha_i + h^\gamma \cdot \gamma + h^\beta \cdot \beta, \quad (3.8)$$

где h^γ – условная стоимость (метрика) балансировки загруженности активных процессоров;

γ – верхний динамически управляемый порог загруженности активных процессоров на интерфейсе активного узла, который входит в ограничение – условие балансировки, модифицируя выражение (3.5)

$$\sum_{i=1}^M \sum_{j=1}^N a_i (1 - \alpha_i) x_{i,j}^k \leq \gamma \cdot c_k, \quad (k = \overline{1, K}). \quad (3.9)$$

Предложенная математическая модель управления очередями на узлах активной сети (3.1)-(3.9), (2.6), (2.8)-(2.14) представлена в виде многоиндексной задачи математического программирования. Наличие ограничений вида (3.2), (3.3) и (3.5), (3.9) позволяет отнести ее к подклассу целочисленных трипланарных задач [92].

Таким образом, в работе предложена математическая модель управления очередями, на основании которой сама задача распределения пакетов по очередям представлена в виде оптимизационной задачи, относящейся к классу трипланарных транспортных задач линейного программирования. В рамках разработанной модели в отличие от ранее известных решений обеспечивается учет характеристик трафика, производительности активных процессоров сетевых узлов и пропускной способности очередей, что позволило придать адаптивный характер процессу распределения поступающих пакетов по отдельным очередям сетевого узла. На рис. 3.7 представлена функциональная архитектура интерфейса активного узла с гибридной архитектурой, определяемая содержанием предложенной в разделе математической модели.

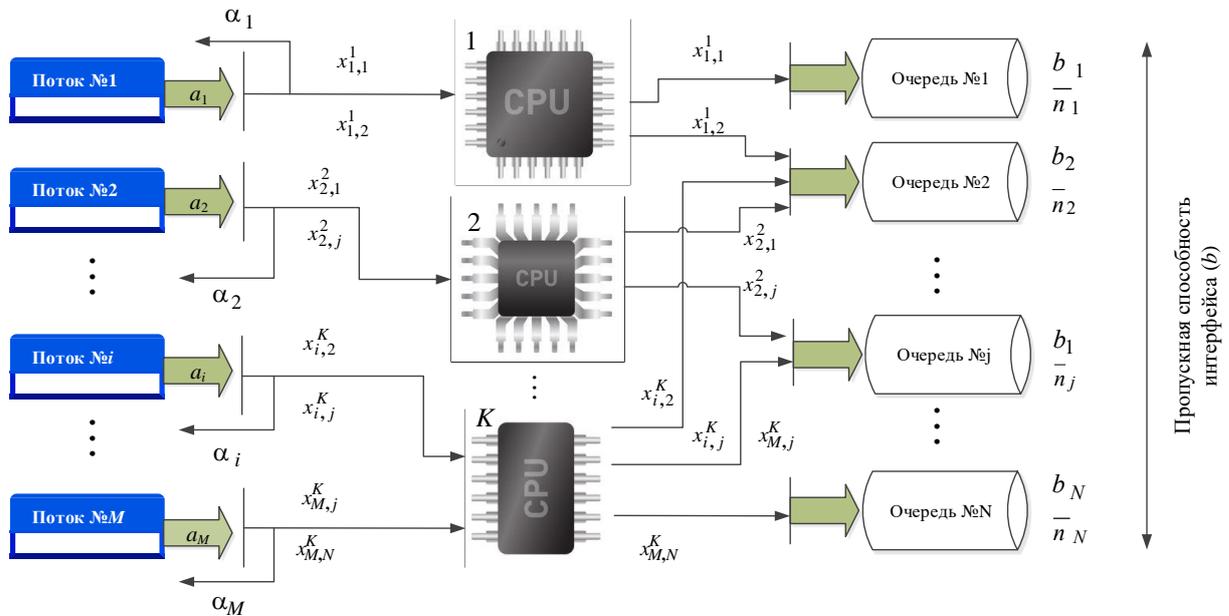


Рис. 3.7. Функциональная архитектура интерфейса активного узла с гибридной архитектурой

Область преимущественного использования предложенной модели управления очередями – сетевые узлы мультисервисных сетей следующего поколения, функционирующие на основе активных технологий.

3.3. Исследование процессов управления очередями на интерфейсах активных узлов с использованием предложенной модели

В работе проведен анализ процессов управления очередями на интерфейсе активного узла с использованием предложенной в подразделе 3.2 модели. В ходе исследования подлежал анализу результат решения задачи агрегирования потоков, выбора активного процессора для реализации определенной функции обработки пакетов, определения номера выходной очереди, контроль за ее перегрузкой на основе балансировки длин очередей с учетом ее типа (приоритета). Рассматривались варианты активного узла как с однородной, так и с гибридной архитектурой. В ходе расчетов принималась гипотеза о

равномерном распределении классов обслуживания между потоками, а средние длины очередей рассчитывались, для примера, в соответствии с системой $M/M/1$ (2.10), с помощью которой моделировалась работа интерфейса. Продемонстрируем работоспособность модели и адекватность получаемых на ее основе результатов решения основных интерфейсных задач по управлению очередями активного узла на ряде примеров.

3.3.1. Пример решения задачи управления очередями на интерфейсах активных узлов с обеспечением сбалансированной загруженности очередей по их длине

В первом случае рассматривался вариант с однородной архитектурой активного узла, в состав которого входило три однотипных активных процессора. Число потоков и очередей принималось равным десяти ($M = 10$) и пяти ($N = 5$) соответственно. Пропускная способность рассматриваемого интерфейса (b) активного узла составляла, например, 100 пакетов в секунду ($1/c$). Более подробно исходные данные по типам потоков и очередей, а также результаты решения задачи по управлению очередями на активном узле с однородной архитектурой представлены в табл. 3.1 и на рис. 3.8.

В ходе решения (рис. 3.8), полученного на основе использования критерия оптимальности (3.6), на первом активном процессоре обрабатывались пакеты первого, третьего и четвертого потоков, на втором АП – пакеты второго, пятого, шестого и десятого потоков, на третьем – пакеты оставшихся седьмого, восьмого и девятого потоков. В соответствии с классами потоков и очередей не задействовались вторая и пятая очереди.

В результате распределения потоков активные процессоры не были перегружены по своей производительности, а очереди не были перегружены по своей длине агрегированными потоками (табл. 3.1).

Подлежащий минимизации верхний порог загруженности очередей, взвешенный относительно классов потоков и очередей, составил $\beta = 0,6427$.

Таблица 3.1

**Исходные данные и результаты расчета для активного узла с
однородной архитектурой ($N=5, M=10, K=3$ и $b=100$)**

Характеристика потоков			Параметры очередей				
номер потока	класс потока	средняя интенсивность потока (1/с)	номер используемой очереди	класс очереди	средняя интенсивность потока в очереди (1/с)	выделенная ПС	средняя/максимальная длина очереди
1	0,596	10,6196	1	1	48,1835	46,8912	64,2722/100
3	0,424	15,6649					
4	0,714	10,4488					
6	0,967	7,3061					
10	1,498	2,1441					
-	-	-	2	3,25	-	-	-/78
5	5,216	12,8853	3	5,50	12,8853	14,6488	6,4272/55
2	6,819	8,3494	4	7,75	29,9423	38,4600	2,7368/33
7	8,181	6,9615					
8	8,175	13,2876					
9	7,224	1,3439					
-	-	-	5	10	-	-	-/10

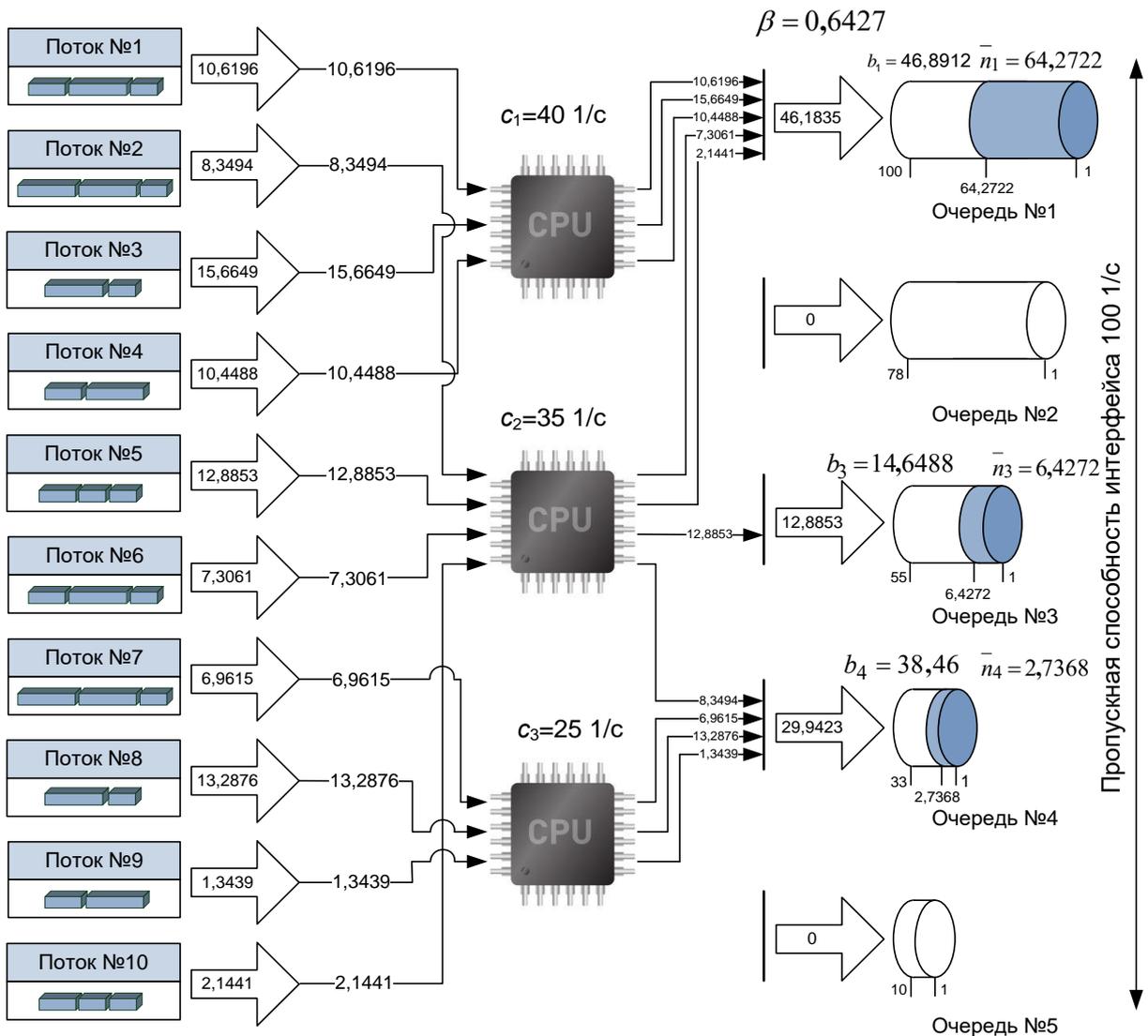


Рис. 3.8. Результаты решения задачи по управлению очередями на активном узле с однородной архитектурой

3.3.2. Пример решения задачи управления очередями на интерфейсах активных узлов с обеспечением сбалансированной загрузки активных процессоров и очередей

Во втором случае рассматривалась гибридная архитектура активного узла, в состав которого входило четыре активных процессора ($K = 4$) – по два каждого типа (табл. 3.2). На рассматриваемый интерфейс с пропускной способностью (b) 100 1/с поступало четырнадцать потоков

($M=14$) с заданными классами, было организовано семь очередей ($N=7$). Результаты расчетов представлены в табл. 3.2 и на рис. 3.9.

Таблица 3.2

Исходные данные и результаты расчета для активного узла с гибридной архитектурой ($N=7$, $M=14$, $K=4$ и $b=100$)

Характеристика потоков			Параметры очередей				
номер потока	класс потока	средняя интенсивность потока (1/с)	номер используемой очереди	класс очереди	средняя интенсивность потока в очереди (1/с)	выделенная ПС	средняя/максимальная длина очереди
4	1,5272	6,2858	1	1	6,2858	6,4346	41,2618/100
1	1,7587	1,0478	2	2,5	22,2994	23,7893	14,0300/85
7	1,9175	7,4476					
9	2,4285	7,3115					
11	2,6906	2,4028					
13	1,8866	2,7167					
14	2,8750	1,3731	3	4	12,7546	14,3270	7,2213/70
3	4,7349	7,8587					
5	3,4112	4,8959	4	5,5	7,4111	8,9057	4,1265/55
6	6,0739	7,4111					
2	7,2176	6,6264					
8	7,3843	7,8087	5	7	22,5918	29,8445	2,3580/40
12	7,6550	8,1567					
10	9,1742	10,8695	6	8,5	10,8695	16,6990	1,2137/25
-	-	-	7	10	-	-	-/10

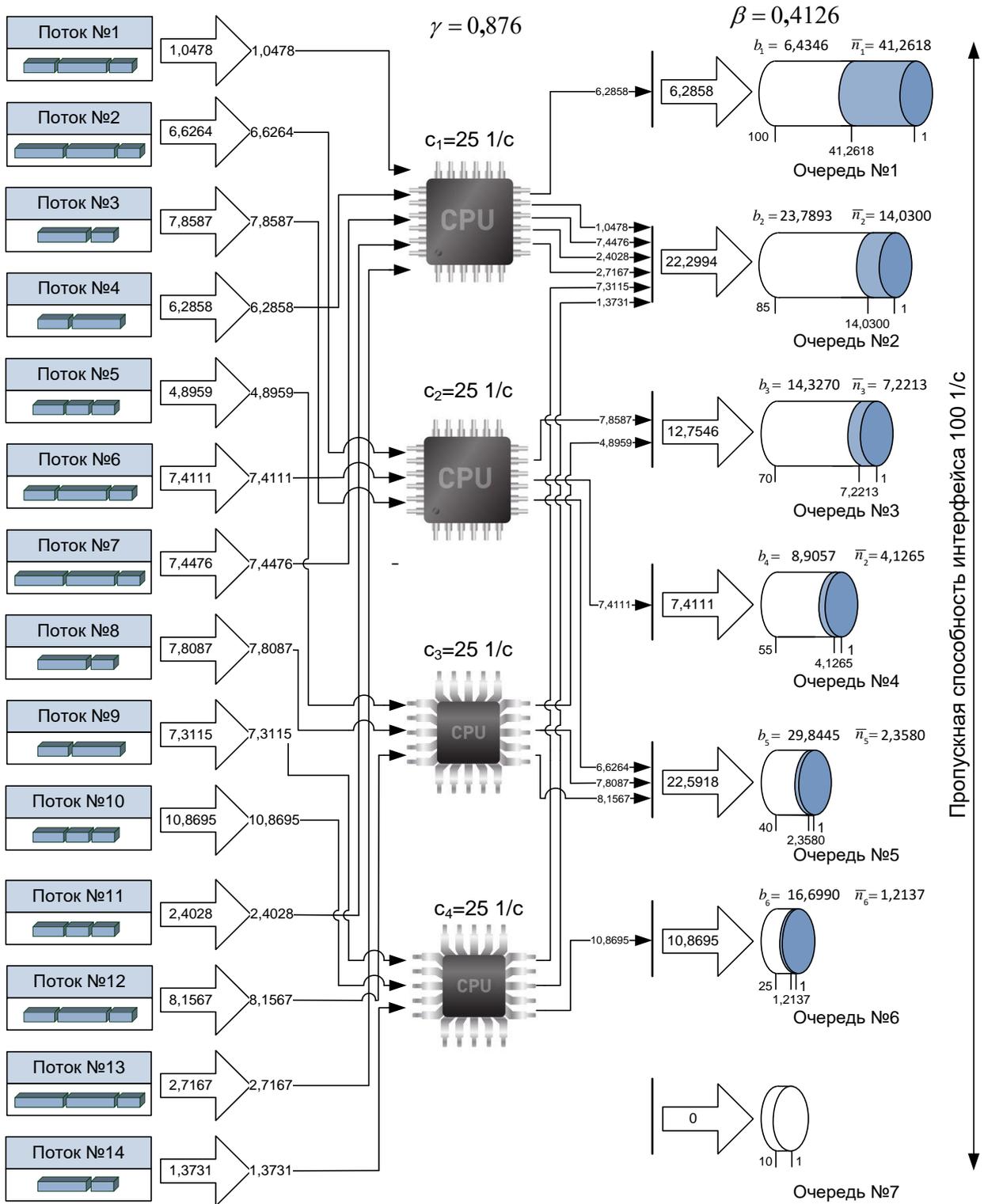


Рис. 3.9. Результаты решения задачи по управлению очередями на активном узле с гибридной архитектурой

В ходе решения оптимизационной задачи (3.1)-(3.5), (3.7), (3.9), (2.6), (2.8)-(2.14), но уже с критерием (3.8), был определен сбалансированный характер решения по распределению потоков не только по очередям, но и по активным процессорам разных типов.

При этом верхний порог загрузки активных процессоров составил $\gamma = 0,876$, а верхний взвешенный по классам обслуживания порог загрузки очередей был равен $\beta = 0,4126$. В ходе агрегирования потоков по очередям десятая очередь остается свободной, т.к. на анализируемом интерфейсе отсутствовали потоки с близким ей классом обслуживания. Напомним, что различие в классах потока и очереди оценивалось с помощью выражения (3.7).

3.4. Модификация модели управления очередями на интерфейсе активного узла при организации конвейера активных процессоров

С целью реализации на активном узле АН сложных (составных) функций по обработке поступающих на него пакетов может потребоваться организация конвейерной архитектуры активных процессоров. Тогда порядок выполнения сложной функции обработки пакетов, например, сжатие и шифрование, полностью определяется топологией соединения подобных процессоров, каждый из которых реализует простую функцию обработки пакетов. Каждой стадии обработки, т.е. реализации одной простой функции, соответствует одна ступень конвейера, образованная параллельно соединенными активными процессорами (рис. 3.10).

На рис. 3.10 показан пример схемы активного узла с конвейерной архитектурой, в рамках которого организовано две ступени обработки пакетов. Первую ступень образуют два разнотипных активных процессора, а вторую – два однотипных АП.

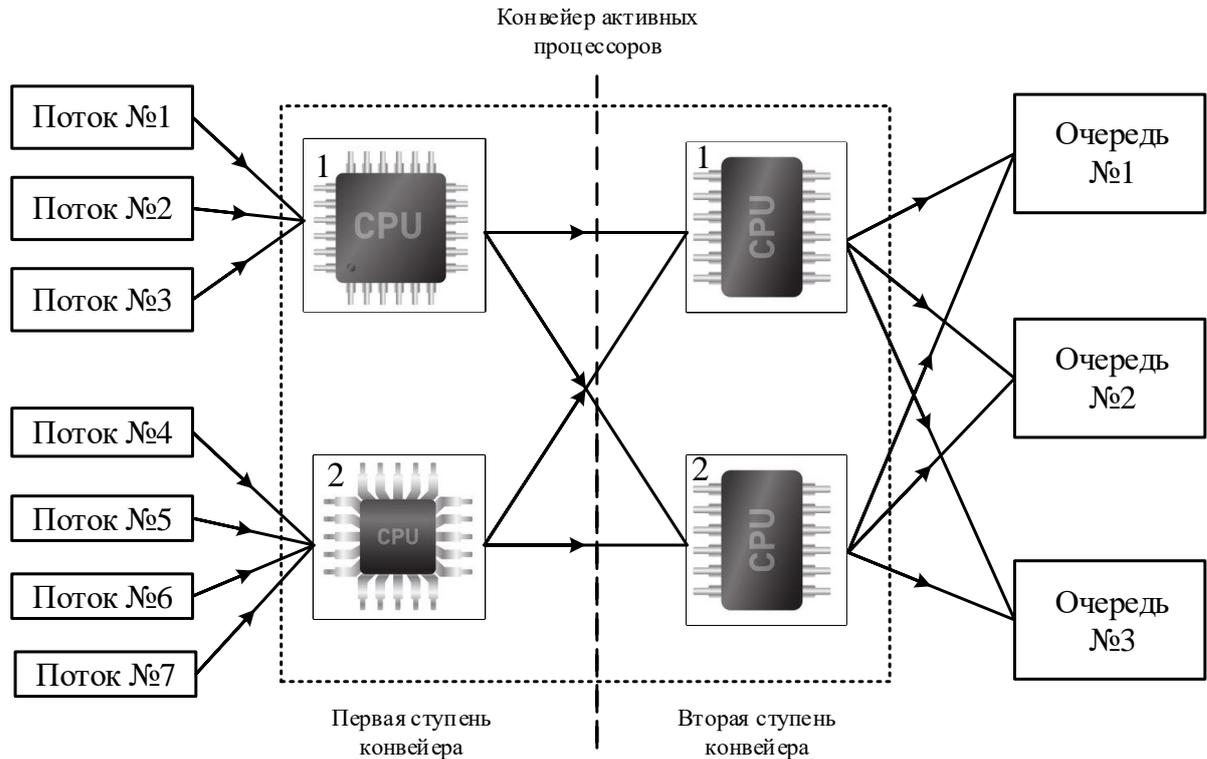


Рис. 3.10. Упрощенная схема активного узла с конвейерной архитектурой

С целью математического описания постановки и решения задачи управления очередями на активном узле с конвейерной архитектурой, расширяя смысл выражений (2.4) и (3.1), вводится множество управляющих переменных $x_{i,j}^{k_1, \dots, k_s, \dots, k_S}$, каждая из которых характеризует долю i -го потока пакетов, который будет направлен для обслуживания в j -ю очередь через цепочку активных процессоров $k_1, \dots, k_s, \dots, k_S$, в которой k_s – номер k -го активного процессора на s -й ступени конвейера ($s = \overline{1, S}$), а S – общее число ступеней в конвейере АП. Тогда согласно физическому смыслу переменных $x_{i,j}^{k_1, \dots, k_s, \dots, k_S}$ справедливы такие условия:

$$x_{i,j}^{k_1, \dots, k_s, \dots, k_S} \in \{0; 1\} \quad (i = \overline{1, M}, j = \overline{1, N}, k_s = \overline{1, K^s}), \quad (3.10)$$

где K^s – общее число активных процессоров на s -й ступени конвейера.

По причине организации конвейера активных процессоров и наличия нескольких ступеней обработки в архитектуре активного узла условия сохранения каждого i -го потока на интерфейсе маршрутизатора (3.2) изменятся в виду:

$$\sum_{k_1=1}^{K^1} \dots \sum_{k_s=1}^{K^s} \sum_{j=1}^N x_{i,j}^{k_1, \dots, k_s, \dots, k_s} = 1, \quad (i = \overline{1, M}). \quad (3.11)$$

Выполнение условий (3.11) гарантирует, что пакеты каждого i -го потока будут направлены лишь в одну очередь через цепочку активных процессоров различных ступеней конвейера. При этом АП каждой отдельной ступени конвейера будет входить в эту цепочку обработки пакета лишь один раз.

Для управляемости процессом предотвращения перегрузки очередей и интерфейса в целом условия (3.3) модифицируются к виду:

$$\sum_{i=1}^M \sum_{k_1=1}^{K^1} \dots \sum_{k_s=1}^{K^s} a_i (1 - \alpha_i) x_{i,j}^{k_1, \dots, k_s, \dots, k_s} < b_j, \quad (j = \overline{1, N}). \quad (3.12)$$

Смысл условий (3.12), как и в случае с (3.3), состоит в том, чтобы суммарная интенсивность потоков, направленных через любую цепочку активных процессоров на обслуживание в j -ю очередь, не превышала пропускную способность интерфейса, которая выделена данной очереди.

На переменные b_j ($j = \overline{1, N}$), определяющие размер пропускной способности интерфейса, выделенной для обслуживания пакетов j -й очереди, и переменные α_i ($i = \overline{1, M}$), характеризующие долю i -го потока, получившего отказ в обслуживании, все также накладывается система ограничений (2.6) и (2.8) соответственно их физическому смыслу.

С целью предотвращения возможной перегрузки интерфейса и очередей, организованных на нем, условия (3.12) также дополняются условиями предотвращения перегрузки очередей по их длине (2.9) совместно с (2.10)-(2.13), но коэффициент использования каждой j -й очереди рассчитывается согласно выражению:

$$\rho_j = \frac{\sum_{i=1}^M \sum_{k_1=1}^{K^1} \dots \sum_{k_s=1}^{K^s} a_i (1 - \alpha_i) x_{i,j}^{k_1, \dots, k_s, \dots, k_s}}{b_j}, \quad (j = \overline{1, N}). \quad (3.13)$$

Условия (3.12), (2.9)-(2.13), (3.13) связывают все три типа управляющих переменных ($x_{i,j}^{k_1, \dots, k_s, \dots, k_s}$, b_j и α_i). Этим обеспечивается столь необходимая согласованность при решении основных интерфейсных задач, связанных с организацией очередей, их обслуживанием, определением порядка обработки пакетов активными процессорами различных ступеней конвейера и обслуживания в разных очередях, распределения пропускной способности интерфейса между отдельными очередями и превентивного ограничения длины очереди.

Для предотвращения перегрузки по производительности каждого из активных процессоров той или иной ступени конвейера условия (3.5) заменяются на выражения вида:

$$\sum_{i=1}^M \sum_{j=1}^N a_i (1 - \alpha_i) x_{i,j}^{k_1, \dots, k_s, \dots, k_s} \leq c_{k_s}, \quad (k_s = \overline{1, K^s}), \quad (3.14)$$

где c_{k_s} – производительность k -го активного процессора на s -й ступени конвейера ($s = \overline{1, S}$).

Условия (3.14) вводятся для того, чтобы результирующая пакетная нагрузка, поступающая на каждый активный процессор любой ступени

конвейера, не превышала его производительности. Для обеспечения сбалансированной загруженности очередей, как и в случае модели, рассмотренной во втором разделе, вводятся условия (2.14).

Предлагается осуществлять расчет переменных $x_{i,j}^{k_1, \dots, k_s, \dots, k_S}$, b_j , α_i и β путем решения оптимизационной задачи в рамках ограничений (3.10)-(3.14), (2.6), (2.8)-(2.14) в ходе минимизации целевой функции вида:

$$F = \sum_{i=1}^M \sum_{j=1}^N \sum_{k_1=1}^{K^1} \dots \sum_{k_S=1}^{K^S} \sum_{j=1}^N h_{i,j}^{k_1, \dots, k_s, \dots, k_S} \cdot x_{i,j}^{k_1, \dots, k_s, \dots, k_S} + \sum_{j=1}^N h_j^b \cdot b_j + \sum_{i=1}^M h_i^\alpha \cdot \alpha_i + h^\beta \cdot \beta, \quad (3.15)$$

в которой $h_{i,j}^{k_1, \dots, k_s, \dots, k_S}$ – условная стоимость (метрика) обслуживания пакетов i -го потока с помощью цепочки активных процессоров $k_1, \dots, k_s, \dots, k_S$ и j -й очереди интерфейса активного узла. Метрики $h_{i,j}^k$ рассчитываются с учетом выражения (2.17) и (3.7):

$$h_{i,j}^{k_1, \dots, k_s, \dots, k_S} = h^{k_1, \dots, k_s, \dots, k_S} w_x^b (k_i^{\mathbf{II}} - k_j^{\mathbf{O}})^2 + 1, \quad (i = \overline{1, M}, j = \overline{1, N}), \quad (3.16)$$

где $h^{k_1, \dots, k_s, \dots, k_S}$ – условная стоимость (метрика) использования цепочки активных процессоров $k_1, \dots, k_s, \dots, k_S$ на рассматриваемом активном узле.

При необходимости обеспечения сбалансированной загруженности не только очередей, но и активных процессоров разных ступеней конвейера, в качестве критерия оптимальности может выступать минимум выражения (3.8), адаптированного по примеру (3.16) к конвейерной обработке пакетов.

3.5. Выводы по третьему разделу

1. На основе проведенного анализа в разделе установлено, что перспективным направлением усовершенствования современных телекоммуникационных технологий является внедрение концепции построения активных сетей. К основным преимуществам данной концепции относят появление дополнительных возможностей относительно оперативного создания и внедрения новых сервисов; предоставление пользователям функций по созданию и внедрению сервисов под их специфичные приложения, текущие сетевые условия функционирования и состояния сети; повышение масштабируемости и гибкости процессов обеспечения качества обслуживания при предоставлении распределенных сервисов; предоставление эффективной платформы для исследования и тестирования новых сетевых сервисов.

2. На основе анализа целей и задач активных сетей и их основных компонентов установлено, что в архитектуре активных узлов выделяется множество активных процессоров, в функции которых входит решение дополнительных задач по обработке пакетов, связанных с их фрагментацией, сжатием данных, управлением приоритетами, шифрованием для обеспечения более высоких показателей качества обслуживания и безопасности сети. Таким образом, задержку пакетов, которую ранее испытывали пакеты, находясь в очереди по причине перегрузки каналов связи, теперь эффективно используют для предоставления распределенных сервисов и сетевого управления.

3. Реализация парадигмы построения активных сетей предопределила необходимость пересмотра математических моделей, методов и самих сетевых протоколов управления трафиком, в частности и механизмов управления очередями на интерфейсах активных узлов. В этой связи в разделе получила дальнейшее развитие модель управления

очередями на интерфейсах маршрутизаторов активных сетей (3.1)-(3.9). Новизна модели заключается в том, что в ходе обслуживания пакетов учитываются не только количество и параметры потоков и очередей, но и характеристики активных процессоров (тип, загруженность и производительность), входящих в архитектуру маршрутизатора активной сети. Реализация предложенной модели позволила обеспечить более эффективное использование буферного и вычислительного ресурса маршрутизатора активных сетей.

4. В рамках предложенного усовершенствования математической модели задача управления очередями на интерфейсах активных сетей была сформулирована как оптимизационная задача математического программирования, относящаяся к подклассу целочисленных трипланарных задач. С помощью предложенной модели на основе введения соответствующих условий-ограничений и вида целевой функции удалось обеспечить согласованное решение как основных интерфейсных задач, связанных с агрегированием потоков, их распределением по очередям, выделением очередям необходимой пропускной способности интерфейса, балансировки длин очередей в соответствии с их классами обслуживания, так и задач по выбору активных процессоров при обработке пакетов и оптимизации уровня их загруженности по производительности.

5. Предложенная модель охватывает случай управления очередями на интерфейсах активного узла как с однородной архитектурой, основанной на множестве однотипных активных процессоров, так и гибридной архитектурой, когда множество АП образовывалось процессорами разных типов, выполняющих различные задачи по обработке пакетов, находящихся в очереди.

6. Расчетные примеры и представленные результаты исследований подтвердили адекватность и эффективность предложенной модели по управлению очередями на интерфейсах активных узлов. В рамках

рассмотренных примеров продемонстрированы возможности предложенной модели по обеспечению сбалансированной загруженности как непосредственно организованных на интерфейсе очередей пакетов, так и активных процессоров, составляющих однородную или гибридную архитектуру активного узла AN.

7. Предложенная в разделе модель (3.1)-(3.9) получила свое обобщение (3.10)-(3.16) для случая оптимизации процессов управления очередями на интерфейсе активного узла, на котором организован конвейер АП, характеризующийся наличием нескольких последовательных ступеней для реализации сложной (составной) функции обработки пакетов.

РАЗДЕЛ 4

РАЗРАБОТКА И ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ МЕТОДА УПРАВЛЕНИЯ ОЧЕРЕДЯМИ НА ИНТЕРФЕЙСАХ МАРШРУТИЗАТОРОВ ТРАНСПОРТНОЙ ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ

Потоковые математические модели, представленные во втором и третьем разделах данной диссертационной работы, описывают общие закономерности организации процесса управления очередями на маршрутизаторах транспортной телекоммуникационной сети. В рамках данных моделей определен характер взаимосвязи между параметрами интерфейса, активных процессоров и характеристиками потока, с одной стороны, и управляющими переменными, отвечающими за согласованное решение основных интерфейсных задач – с другой.

Однако порядок использования математических выражений, составляющих основу рассмотренных моделей, может определить лишь соответствующий метод управления очередями на интерфейсах маршрутизаторов транспортной сети. В методе должны быть учтены особенности использования тех или иных математических зависимостей в конкретной сложившейся ситуации по состоянию загруженности интерфейса или изменению характеристик потока.

В этой связи в данном разделе будет описан предлагаемый метод управления очередями на интерфейсах маршрутизаторов транспортной телекоммуникационной сети, проведено его исследование, в том числе с помощью лабораторного эксперимента на реальном сетевом оборудовании, и оценена эффективность его применения. Результаты исследования определили рекомендации по его практической реализации в современных и перспективных технологиях управления трафиком в целом и очередями в частности.

4.1. Описание метода управления очередями на интерфейсах маршрутизаторов транспортной телекоммуникационной сети

В соответствии с усовершенствованными в предыдущих разделах моделями предлагаемый в работе метод управления очередями на интерфейсах маршрутизаторов транспортной телекоммуникационной сети будет включать в себя следующие основные этапы и процедуры:

1. Мониторинг и анализ следующих параметров и характеристик, которые выступают в качестве исходных данных для последующего расчета:

- параметров интерфейса, к которым, прежде всего, относится его тип и номинальная пропускная способность (b), а также число настроенных «по умолчанию» очередей (N);

- характеристик потоков пакетов, которые в соответствии с содержанием маршрутной таблицы маршрутизатора поступают на выбранный интерфейс: общее число потоков (M), характеристики каждого из потоков – средняя интенсивность (скорость поступления пакетов) a_i , длина пакета, значения IP-приоритета, поле DSCP, IP-адрес, MAC-адрес и номера TCP/UDP-портов отправителя/получателя и др.;

- параметров активных процессоров, если маршрутизатор является элементом AN, т.е. активным узлом: общего числа (K) и типов активных процессоров, их производительности (c_k , $k = \overline{1, K}$). Если организован конвейер активных процессоров при реализации сложных функций обработки пакетов, то уточняется и число ступеней конвейера, а также количество активных процессоров, параллельно функционирующих в рамках каждой из ступеней.

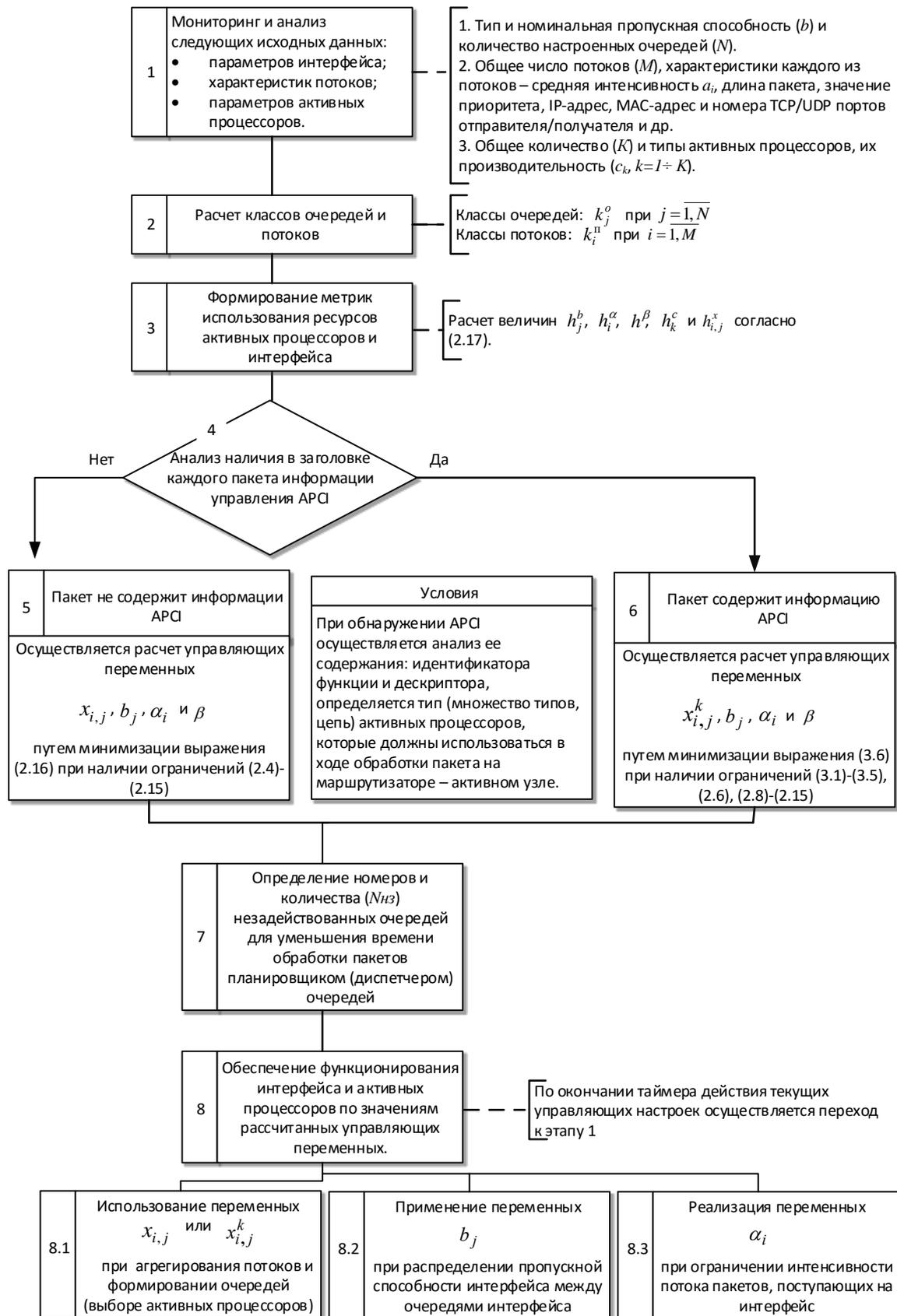


Рис. 4.1. Структура предлагаемого метода управления очередями на интерфейсах маршрутизаторов транспортной телекоммуникационной сети

2. Исходя из анализа параметров активных процессоров маршрутизатора, состояния интерфейса и характеристик потоков пользователей, которые на него поступают, осуществляется расчет классов очередей k_j^y ($j = \overline{1, N}$) и потоков k_i^p ($i = \overline{1, M}$) соответственно. При этом для расчета классов потоков могут использоваться выражения (2.1) или (2.2). Использование формулы (2.1) целесообразно при дифференциации качества обслуживания, а выражение (2.2) по сравнению с (2.1) позволяет обеспечить более полный учет в явном виде показателей качества обслуживания, но предполагает, что эта информация доступна по каждому потоку на данном маршрутизаторе.

3. Формирование метрик использования ресурсов активных процессоров и интерфейса путем определения величин h_j^b , h_i^α , h^β , h_k^c и $h_{i,j}^x$ в соответствии с выражением (2.17). Путем определения соотношения между метриками h^β и h_i^α можно регулировать уровень превентивности (заблаговременности) реакции управления на возможную перегрузку очереди путем ограничения интенсивности потока, поступающего в очередь.

4. Анализ наличия в заголовке каждого пакета информации управления APСI. При идентификации поля APСI осуществляется анализ его содержания: идентификатора функции и дескриптора, определяется тип (множество типов, цепочка) активных процессоров, которые могут использоваться в ходе обработки пакета на маршрутизаторе – активном узле.

5. Если пакет не содержит информацию APСI, то осуществляется расчет управляющих переменных $x_{i,j}$, b_j , α_i и β путем минимизации выражения (2.16) при наличии ограничений (2.4)-(2.15).

6. В случае, если в пакете содержится информация APCT, то производится согласованный расчет управляющих переменных $x_{i,j}^k$, b_j , α_i и β в ходе минимизации целевой функции (3.6) при выполнении условий-ограничений (2.6), (2.8)-(2.15), (3.1)-(3.5). При обеспечении дополнительной балансировки загруженности активных процессоров минимизации подлежит функция (3.8) с заменой выражения (3.5) на (3.9). Если организован конвейер обработки пакетов, то расчету подлежат также переменные $x_{i,j}^{k_1, \dots, k_s, \dots, k_s}$ (3.10), но уже за счет минимизации выражения (3.15) при наличии модифицированных условий (3.11)-(3.14).

7. По результатам расчета, анализируя значения $x_{i,j}$, $x_{i,j}^k$ и b_j , определяются номера и количество ($N_{нз}$) незадействованных в процессе обслуживания пакетов очередей для снижения времени обработки пакетов планировщиком (диспетчером) очередей.

8. Обеспечение функционирования интерфейса и активных процессоров по значениям рассчитанных управляющих переменных:

- переменные $x_{i,j}$ используются для агрегирования потоков и формирования очередей, а значения $x_{i,j}^k$ применяются еще и для выбора активных процессоров, участвующих в обработке пакетов. Фактически данные переменные служат заменой списков доступа (ACLs), формируемых администратором вручную, что также способствует повышению оперативности управления очередями;

- переменные b_j применяются в ходе распределения пропускной способности интерфейса между настроенными очередями. На практике значения данных управляющих переменных могут определять параметры команд «bandwidth» или «priority» для механизмов CBWFQ и LLQ соответственно;

- значения управляющих переменных α_i по своему физическому смыслу совпадают с вероятностью отбрасывания пакетов каждого i -го потока. Поэтому данный тип управляющих переменных отображает результат работы AQM-механизмов, определяя долю интенсивности потока пакетов, отброшенных из очереди ввиду вероятной перегрузки очереди на интерфейсе.

После окончания таймера действия текущих значений управляющих настроек осуществляется переход к этапу 1 (рис. 4.1).

4.2. Исследование предложенного метода управления очередями на интерфейсах маршрутизаторов транспортной ТКС

Ключевым моментом при разработке математических моделей и методов управления очередями на интерфейсах маршрутизаторов ТКС является оценка эффективности получаемых на их основе решений. Именно по результатам оценки эффективности предложенных моделей и метода можно уточнить границы области применения принятых в ходе исследований допущений и предположений, обосновать наиболее рациональную их область практического использования с точки зрения состояния интерфейса, характеристик потоков и требований к уровню качества обслуживания. Кроме того, в ходе проводимого эксперимента становится доступной проверка адекватности расчетных решений, ранее полученных с помощью методов аналитического моделирования.

Наиболее точные результаты исследования возможно получить, лишь используя возможности натурального эксперимента, который позволяет максимально учесть реальные условия функционирования и эксплуатации современных телекоммуникационных систем и сетей, коммутационного оборудования и сетевых протоколов информационного обмена. Ввиду того, что тестировать новые технологические решения на базе реально

функционирующих сетей не всегда представляется возможным, то некоторым выходом из сложившейся ситуации является проведение лабораторного эксперимента, т.е. эксперимента на лабораторном оборудовании, состоящем из серийных коммутационных устройств маршрутизаторов и коммутаторов, нагружаемых с помощью генераторов разнотипного трафика.

По причине достаточно большого числа возможных вариантов исходных данных, для которых необходимо получить результаты исследования, натурный и лабораторный эксперимент может сопровождаться достаточно большими временными, иногда и материальными затратами. Поэтому лабораторный эксперимент целесообразно проводить на финальной стадии исследований в ходе проверки ограниченного числа предварительно отобранных решений, полученных с помощью средств математического моделирования – аналитического и/или имитационного.

4.2.1. Описание общей схемы проводимого лабораторного эксперимента

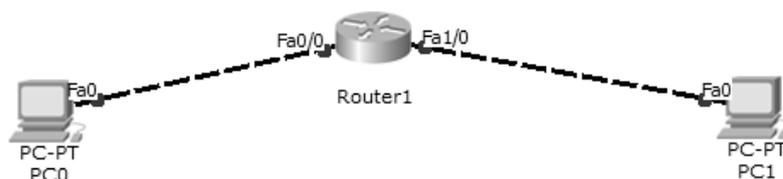
При организации лабораторного эксперимента использовалась методика исследований, которая была предложена и успешно апробирована в лаборатории систем распределения информации кафедры телекоммуникационных систем ХНУРЭ, как в учебном процессе при выполнении практических и лабораторных работ, так и в ходе диссертационных исследований аспирантов и докторантов кафедры [93-95]. В ходе исследования использовалось коммутационное оборудование, представленное маршрутизаторами производства компании Cisco серии 2800 (рис. 4.2), в рамках которых было предусмотрено наличие высокоскоростных интерфейсов Fast Ethernet с пропускной способностью

10/100 Мбит/с, а также настраиваемых синхронных последовательных интерфейсов Serial, на которых устанавливалась скорость до 2 Мбит/с.

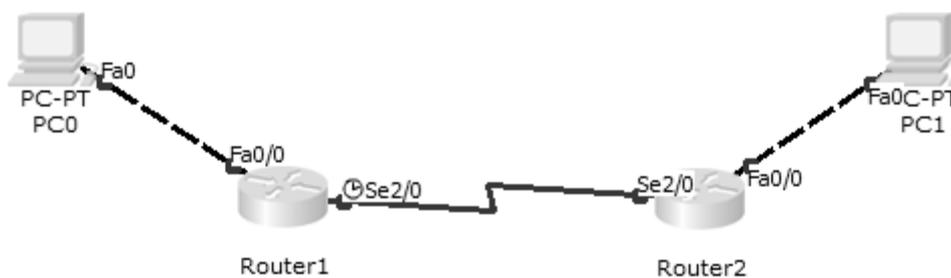


Рис. 4.2. Внешний вид маршрутизатора Cisco серии 2800 с Fast Ethernet и Serial интерфейсами

Ввиду того, что в ходе исследования интересовала работа лишь отдельного интерфейса, то для исследования работы высокоскоростного интерфейса Fast Ethernet использовалась схема лабораторной установки, содержащая два терминала и один маршрутизатор (рис. 4.3 а).



а)

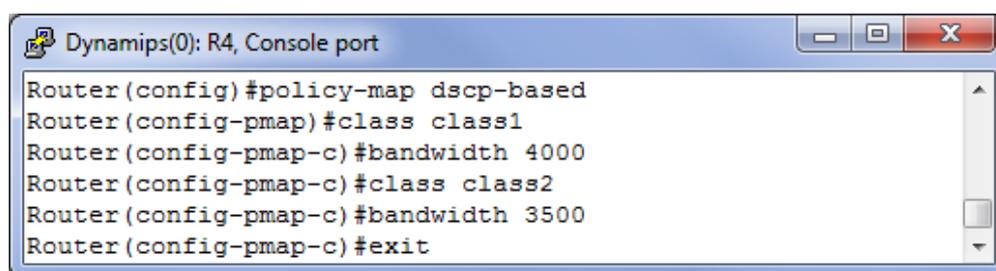


б)

Рис. 4.3. Варианты схем организации лабораторного эксперимента

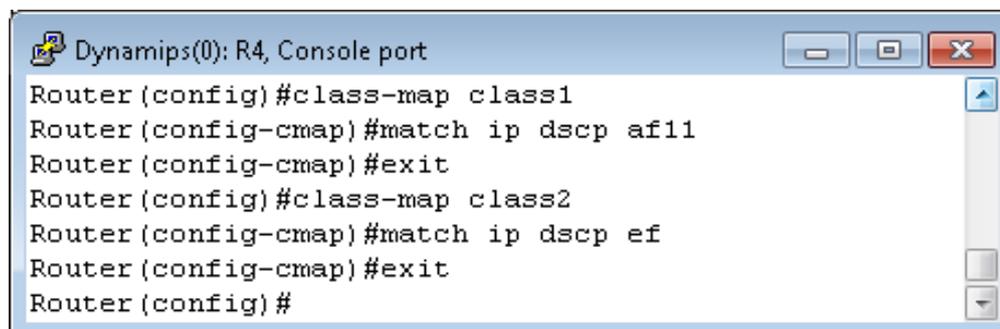
При анализе работы низкоскоростного интерфейса схема (рис. 4.3 б), что особенно было востребовано при исследовании режимов высокой и критической загруженности очередей (т.к. перегрузить высокоскоростной интерфейс было достаточно трудно), содержала два терминала и два маршрутизатора – «Router 1» и «Router 2», соединенные посредством Serial интерфейса.

При настройке очередей под результаты расчета, полученные в рамках предложенного метода, использовался механизм обслуживания очередей на основании классов CBWFQ [96-98], с помощью которой командой «bandwidth» в явном виде задавалась пропускная способность интерфейса (b_j), выделенная для каждой из очередей (рис. 4.4 а).



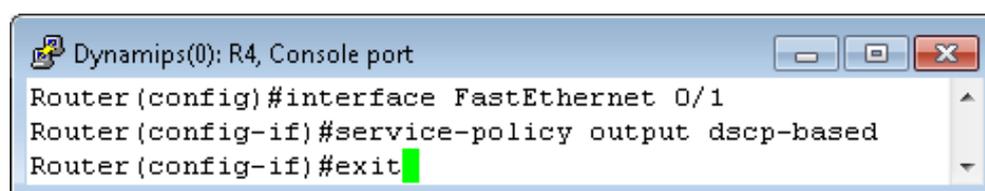
```
Dynamips(0): R4, Console port
Router(config)#policy-map dscp-based
Router(config-pmap)#class class1
Router(config-pmap-c)#bandwidth 4000
Router(config-pmap-c)#class class2
Router(config-pmap-c)#bandwidth 3500
Router(config-pmap-c)#exit
```

а)



```
Dynamips(0): R4, Console port
Router(config)#class-map class1
Router(config-cmap)#match ip dscp af11
Router(config-cmap)#exit
Router(config)#class-map class2
Router(config-cmap)#match ip dscp ef
Router(config-cmap)#exit
Router(config)#
```

б)



```
Dynamips(0): R4, Console port
Router(config)#interface FastEthernet 0/1
Router(config-if)#service-policy output dscp-based
Router(config-if)#exit
```

в)

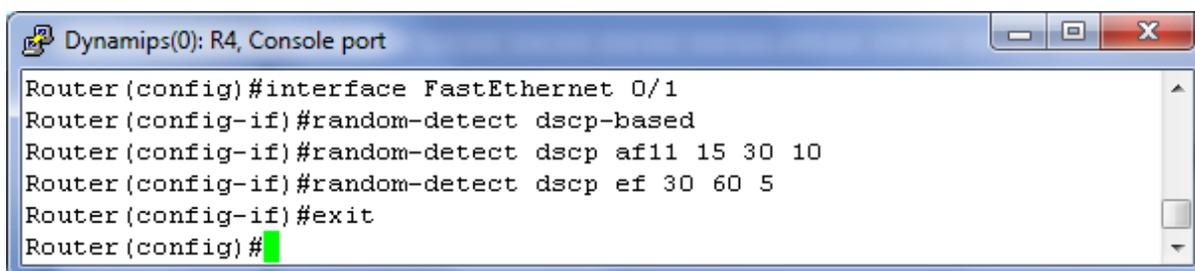
Рис. 4.4. Пример настройки механизма CBWFQ под результаты расчета, полученные с помощью предлагаемого метода управления очередями

На рис. 4.4 б) показан пример классификации и агрегирования потоков, например, по полю DSCP (AF11 и EF) в ходе создания отмеченных на рис. 4.4 а) двух классов обслуживания: class1 и class2. Пример применения созданной политики на конкретном сетевом интерфейсе (Fast Ethernet) показан на рис. 4.4 в).

В ходе проводимого эксперимента также возникла необходимость в реализации на каждом конкретном интерфейсе такой вероятности отбрасывания пакетов для каждого i -го потока, которая была определена в соответствии с рассчитанными в рамках предложенного метода величинами α_i , $i = \overline{1, M}$. Это достигалось соответствующим заданием верхнего и нижнего порога для средней длины очереди, а также знаменателя граничной вероятности для обеспечения наперед заданного значения вероятности отбрасывания пакетов [6]

$$P_{drop} = \frac{q_{av} - \Theta_{\min}}{\Theta_{\max} - \Theta_{\min}} \cdot \frac{1}{\delta}.$$

Пример задания верхнего и нижнего порога для средней длины очереди, а также знаменателя граничной вероятности в рамках механизма WRED на интерфейсе Fast Ethernet представлен на рис.4.5.

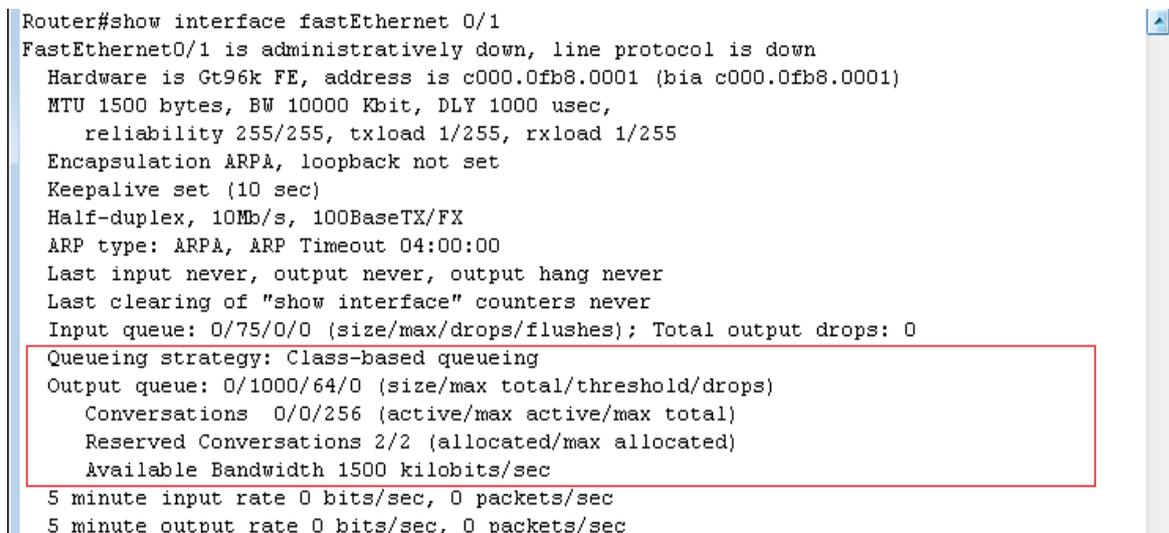


```

Dynamips(0): R4, Console port
Router(config)#interface FastEthernet 0/1
Router(config-if)#random-detect dscp-based
Router(config-if)#random-detect dscp af11 15 30 10
Router(config-if)#random-detect dscp ef 30 60 5
Router(config-if)#exit
Router(config)#
  
```

Рис. 4.5. Пример настройки механизма WRED на интерфейсе Fast Ethernet для обеспечения рассчитанной методом вероятности отбрасывания пакетов

С помощью команды `show interface` (рис. 4.6) анализировалось состояние интерфейса с точки зрения текущей и максимальной длины очереди, а также статистика о количестве отброшенных пакетов за время выполнения команды.



```

Router#show interface fastEthernet 0/1
FastEthernet0/1 is administratively down, line protocol is down
Hardware is Gt96k FE, address is c000.0fb8.0001 (bia c000.0fb8.0001)
MTU 1500 bytes, BW 10000 Kbit, DLY 1000 usec,
    reliability 255/255, txload 1/255, rxload 1/255
Encapsulation ARPA, loopback not set
Keepalive set (10 sec)
Half-duplex, 10Mb/s, 100BaseTX/FX
ARP type: ARPA, ARP Timeout 04:00:00
Last input never, output never, output hang never
Last clearing of "show interface" counters never
Input queue: 0/75/0/0 (size/max/drops/flushes); Total output drops: 0
Queueing strategy: Class-based queueing
Output queue: 0/1000/64/0 (size/max total/threshold/drops)
    Conversations 0/0/256 (active/max active/max total)
    Reserved Conversations 2/2 (allocated/max allocated)
    Available Bandwidth 1500 kilobits/sec
5 minute input rate 0 bits/sec, 0 packets/sec
5 minute output rate 0 bits/sec, 0 packets/sec

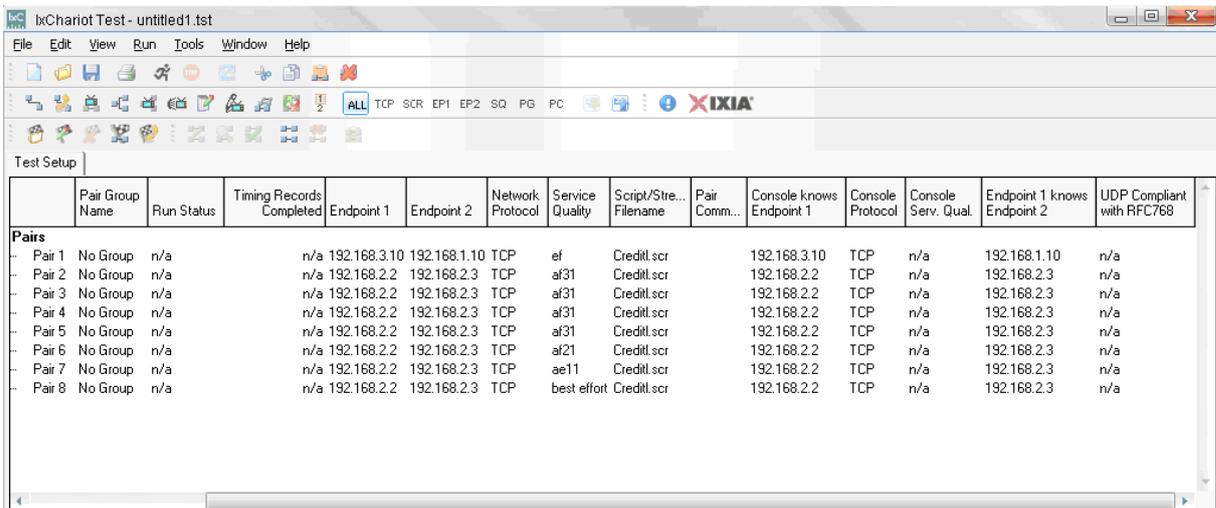
```

Рис. 4.6. Результаты анализа состояния интерфейса

Для генерации трафика на стороне отправителя пакетов и контроля за численными значениями основных показателей качества обслуживания, в общем случае, могут использоваться различные пакеты нагрузочного тестирования, относящиеся к средствам имитационного моделирования. Среди них основными являются следующие: D-ITG (Distributed Internet Traffic Generator), Mtools, Rude/Crude, Netperf, Mgen, Iperf, UDPgenerator, IxChariot, IP Traffic [99-101], отличающиеся числом поддерживаемых функций, возможностями по настройке, условиями распространения (лицензирования) и др.

Пакеты Iperf и Netperf, как правило, конфигурируются в режиме командной строки (CLI), что предполагает соответствующий уровень квалификации и знание команд настройки исследователем. В отличие от этих пакетов генераторы IP Traffic и IxChariot поддерживают графический

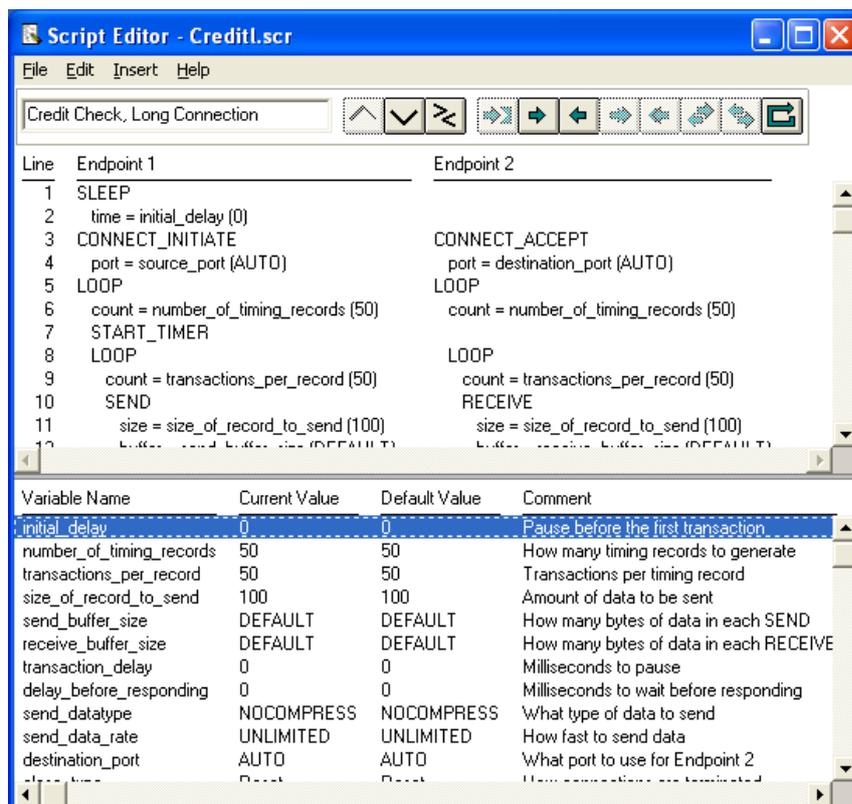
интерфейс (рис. 4.7), что способствует повышению наглядности результатов настроек.



The screenshot shows the IxChariot Test interface with a table titled "Test Setup". The table lists various test pairs with their configurations.

	Pair Group Name	Run Status	Timing Records Completed	Endpoint 1	Endpoint 2	Network Protocol	Service Quality	Script/Stream Filename	Pair Comm...	Console knows Endpoint 1	Console Protocol	Console Serv. Qual.	Endpoint 1 knows Endpoint 2	UDP Compliant with RFC768
Pairs														
-	Pair 1	No Group	n/a	n/a	192.168.3.10	192.168.1.10	TCP	ef	Credit.scr	192.168.3.10	TCP	n/a	192.168.1.10	n/a
-	Pair 2	No Group	n/a	n/a	192.168.2.2	192.168.2.3	TCP	af31	Credit.scr	192.168.2.2	TCP	n/a	192.168.2.3	n/a
-	Pair 3	No Group	n/a	n/a	192.168.2.2	192.168.2.3	TCP	af31	Credit.scr	192.168.2.2	TCP	n/a	192.168.2.3	n/a
-	Pair 4	No Group	n/a	n/a	192.168.2.2	192.168.2.3	TCP	af31	Credit.scr	192.168.2.2	TCP	n/a	192.168.2.3	n/a
-	Pair 5	No Group	n/a	n/a	192.168.2.2	192.168.2.3	TCP	af31	Credit.scr	192.168.2.2	TCP	n/a	192.168.2.3	n/a
-	Pair 6	No Group	n/a	n/a	192.168.2.2	192.168.2.3	TCP	af21	Credit.scr	192.168.2.2	TCP	n/a	192.168.2.3	n/a
-	Pair 7	No Group	n/a	n/a	192.168.2.2	192.168.2.3	TCP	ae11	Credit.scr	192.168.2.2	TCP	n/a	192.168.2.3	n/a
-	Pair 8	No Group	n/a	n/a	192.168.2.2	192.168.2.3	TCP	best effort	Credit.scr	192.168.2.2	TCP	n/a	192.168.2.3	n/a

а) пример установки числа потоков, их приоритетов, IP-адресов отправителя и получателя, типа транспортного протокола



The screenshot shows the Script Editor for "CreditL.scr" with a script for a credit check and a table of variables.

Line	Endpoint 1	Endpoint 2
1	SLEEP	
2	time = initial_delay (0)	
3	CONNECT_INITIATE	CONNECT_ACCEPT
4	port = source_port (AUTO)	port = destination_port (AUTO)
5	LOOP	LOOP
6	count = number_of_timing_records (50)	count = number_of_timing_records (50)
7	START_TIMER	
8	LOOP	LOOP
9	count = transactions_per_record (50)	count = transactions_per_record (50)
10	SEND	RECEIVE
11	size = size_of_record_to_send (100)	size = size_of_record_to_send (100)
12	buffer_size = buffer_size (DEFAULT)	buffer_size = buffer_size (DEFAULT)

Variable Name	Current Value	Default Value	Comment
initial_delay	0	0	Pause before the first transaction
number_of_timing_records	50	50	How many timing records to generate
transactions_per_record	50	50	Transactions per timing record
size_of_record_to_send	100	100	Amount of data to be sent
send_buffer_size	DEFAULT	DEFAULT	How many bytes of data in each SEND
receive_buffer_size	DEFAULT	DEFAULT	How many bytes of data in each RECEIVE
transaction_delay	0	0	Milliseconds to pause
delay_before_responding	0	0	Milliseconds to wait before responding
send_datatype	NOCOMPRESS	NOCOMPRESS	What type of data to send
send_data_rate	UNLIMITED	UNLIMITED	How fast to send data
destination_port	AUTO	AUTO	What port to use for Endpoint 2

б) пример задания длин пакетов и интенсивности потока

Рис. 4.7. Примеры задания параметров сетевого трафика в пакете IxChariot

Пакет D-ITG может конфигурироваться как в режиме командной строки, так и с помощью графического интерфейса, что является его достоинством. Все перечисленные пакеты нагрузочного тестирования, как правило, поддерживают генерирование потоков пакетов, передаваемых в рамках транспортных протоколов TCP и UDP. Кроме того, генераторы D-ITG и IxChariot позволяют задавать характеристики ICMP и VoIP потоков.

Пакеты D-ITG, IP Traffic и IxChariot обеспечивают достаточно высокую гибкость при задании параметров потоков, реализуя тот или иной закон распределения. Например, генератор D-ITG позволяет управлять выбором закона распределения длины пакета (Packet Size) и времени между отправкой пакетов (Inter Departure Time), среди которых экспоненциальное, равномерное, распределение Коши, нормальное, распределение Парето, Пуассоновское распределение, гамма-распределение.

Большинство пакетов позволяют оценивать в реальном времени основные показатели качества обслуживания: скорость, среднюю задержку, джиттер и уровень потерь пакетов. IxChariot, кроме того, способен обеспечить расчет MOS-показателей воспринимаемого качества обслуживания. Но в отличие от генератора IxChariot пакет D-ITG является свободно распространяемым программным обеспечением [99-101], что и определило его выбор для использования в ходе проводимого лабораторного эксперимента.

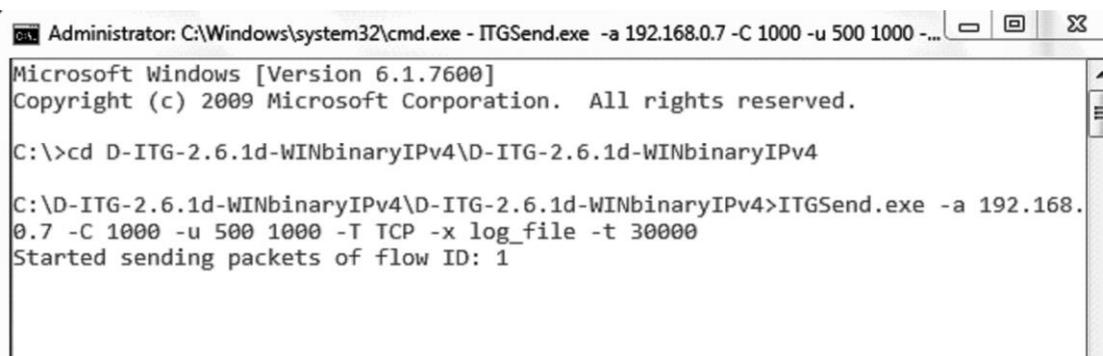
С помощью пакета D-ITG задавалось число и параметры потоков пакетов, а также непосредственно производились измерения средней межконцевой задержки пакетов (one-way-delay, OWD), определялись количество потерянных пакетов (packet loss), джиттер (jitter) и пропускная способность (throughput) соединения (пути). На стороне отправителя запускалась подпрограмма D-ITGSender, а на стороне получателя пакетов – подпрограмма D-ITGReceiver, а также подпрограмма обработки

результатов D-ITGDecoder. Все подпрограммы запускались в режиме CLI и настраивались с помощью изменения «ключей».

Пример задания параметров потока с помощью подпрограммы D-ITG Sender представлен на рис. 4.8. С помощью «ключей» указывались следующие основные характеристики потока:

- IP-адрес получателя пакетов (-a 192.168.0.7);
- пакетная скорость передачи (1/с) (-C 1000);
- размер пакета изменялся по равномерному закону от 500 до 1000 байт (-u 500 1000);
- использовался протокол транспортного уровня TCP (-T TCP);
- время генерирования пакетов составляло 30 секунд (-t 30000)

Кроме того, с помощью «ключа» -b (ToS byte) задавалось значение поля DSCP, которое принимало значения от 0 до 255 (табл. 4.1).



```

Administrator: C:\Windows\system32\cmd.exe - ITGSend.exe -a 192.168.0.7 -C 1000 -u 500 1000 -...
Microsoft Windows [Version 6.1.7600]
Copyright (c) 2009 Microsoft Corporation. All rights reserved.

C:\>\cd D-ITG-2.6.1d-WINbinaryIPv4\D-ITG-2.6.1d-WINbinaryIPv4

C:\D-ITG-2.6.1d-WINbinaryIPv4\D-ITG-2.6.1d-WINbinaryIPv4>ITGSend.exe -a 192.168.
0.7 -C 1000 -u 500 1000 -T TCP -x log_file -t 30000
Started sending packets of flow ID: 1
  
```

Рис. 4.8. Пример задания параметров потока с помощью подпрограммы D-ITG Sender

Таблица 4.1

Соответствие кода поля DSCP его десятичному значению

Код поля DSCP	AF11	AF12	AF13	AF21	AF22	AF23	AF31
Десятичное значение	40	48	56	72	80	88	104
Код поля DSCP	AF32	AF33	AF41	AF42	AF43	EF	
Десятичное значение	112	120	136	144	152	184	

На приемной стороне с помощью подпрограмм D-ITGReceiver и D-ITGDecoder формировались, обрабатывались и отображались результаты эксперимента (рис. 4.9) с указанием численных значений общего числа переданных пакетов; значений минимальной, максимальной и средней задержки пакетов; джиттера (вариации задержки); средней битовой и пакетной скорости передачи потока, а также процента потерянных пакетов.

```

C:\D-ITG-2.6.1d-WINbinaryIPv4\D-ITG-2.6.1d-WINbinaryIPv4>ITGDec.exe log_file -d
1000 -j 1000 -p 1000
/
C:\D-ITG-2.6.1d-WINbinaryIPv4\D-ITG-2.6.1d-WINbinaryIPv4>ITGDec.exe log_file -d
1000 -j 1000 -p 1000 -v
/-----
Flow number: 1
From 192.168.0.7:51330
To 192.168.0.7:8999
-----
Total time           = 29.983000 s
Total packets       = 30000
Minimum delay       = 0.000000 s
Maximum delay       = 0.300000 s
Average delay       = 0.000927 s
Average jitter      = 0.000158 s
Delay standard deviation = 0.010089 s
Bytes received      = 22462714
Average bitrate     = 5993.453357 Kbit/s
Average packet rate = 1000.566988 pkt/s
Packets dropped     = 0 (0.00 %)
-----

***** TOTAL RESULTS *****
-----
Number of flows     = 1
Total time         = 29.983000 s
Total packets      = 30000
Minimum delay      = 0.000000 s
Maximum delay      = 0.300000 s
Average delay      = 0.000927 s
Average jitter     = 0.000158 s
Delay standard deviation = 0.010089 s
Bytes received     = 22462714
Average bitrate    = 5993.453357 Kbit/s
Average packet rate = 1000.566988 pkt/s
Packets dropped    = 0 (0.00 %)
Error lines        = 0
-----
C:\D-ITG-2.6.1d-WINbinaryIPv4\D-ITG-2.6.1d-WINbinaryIPv4>

```

Рис. 4.9. Пример отображения значений основных QoS-показателей с использованием пакета D-ITG

4.2.2. Оценка эффективности процесса управления очередями, организованного в рамках предложенного метода

Для оценки эффективности решений по управлению очередями, полученных с помощью предложенного в подразделе 4.1 метода (**Метод 1**), проведено их сравнение с механизмами, которые автоматически настраиваются на интерфейсах маршрутизаторов (FIFO/WRED, WFQ/WRED), а также методом оптимальной балансировки длин очередей (**Метод 2**), предложенным в работе [13], и методом балансировки загруженности интерфейса (**Метод 3**) [12], в рамках которых агрегирование потоков по очередям не предусматривалось. Механизмы CQ, CBQ, CBWFQ и LLQ не входили в число сравниваемых решений, т.к. эффективность их функционирования целиком и полностью определялась содержанием команд административной настройки.

Эффективность предложенного метода и ранее известных решений оценивалась по тому, насколько был улучшен тот или иной QoS-показатель по сравнению с применением решения FIFO/WRED. Например, выигрыш по средней задержке пакетов рассчитывался для каждого m -го сравниваемого решения (WFQ/WRED, Метод 1, Метод 2, Метод 3) с помощью выражения

$$V_{\tau}^m = \left(1 - \frac{\tau_m}{\tau_{fifo}} \right) \cdot 100\%, \quad (4.1)$$

где τ_{fifo} – средняя задержка пакетов на интерфейсе при использовании решения FIFO/WRED;

τ_m – средняя задержка пакетов на интерфейсе при использовании m -го сравниваемого решения.

В ходе исследования оценивалось влияние на среднюю задержку пакетов, имеющих разный приоритет, загруженности интерфейса ($0 < \rho < 1$), который традиционно определялся из отношения суммарной пакетной нагрузки на интерфейс к его пропускной способности. На вход интерфейса подавался агрегированный трафик, число потоков в котором изменялось от десяти до двадцати пяти. Половина потоков имела нулевой приоритет (потоки данных), приоритеты остальных потоков выбирались равномерно с соответствием РНВ-политиками – AF и EF (табл. 4.1).

Результаты эксперимента показали, что предлагаемый метод позволил в целом улучшить значения средней задержки пакетов для приоритетных потоков. При этом в области малой загруженности интерфейса ($0 \leq \rho \leq 0,55$) вероятность возникновения очереди была крайне мала, поэтому все сравниваемые средства фактически реализовывали принцип работы алгоритма FIFO.

В области средних и высоких нагрузок ($0,6 \leq \rho \leq 0,95$) использование предлагаемого в работе метода (Метод 1) позволило дифференцированно улучшить среднюю задержку пакетов:

- для потока пакетов, передаваемых с приоритетом AF11, от 5-15% до 30-45% по сравнению с FIFO/WRED; от 3-8% до 12-20% по сравнению с WFQ/WRED; от 3-5% до 11-17% по сравнению с методом 3; от 2-4% до 6-12% по сравнению с методом 2 (рис. 4.10);

- для потока пакетов, передаваемых с приоритетом AF32, от 4-9% до 32-47% по сравнению с FIFO/WRED; от 2-6% до 15-22% по сравнению с WFQ/WRED; от 2-4% до 12-16% по сравнению с методом 3; от 1-3% до 6-7% по сравнению с методом 2 (рис. 4.11);

- для потока пакетов, передаваемых с приоритетом EF, от 5-10% до 35-50% по сравнению с FIFO/WRED; от 2-5% до 17-23% по сравнению с WFQ/WRED; от 2-4% до 13-17% по сравнению с методом 3; от 1-3% до 7-8% по сравнению с методом 2 (рис. 4.12).

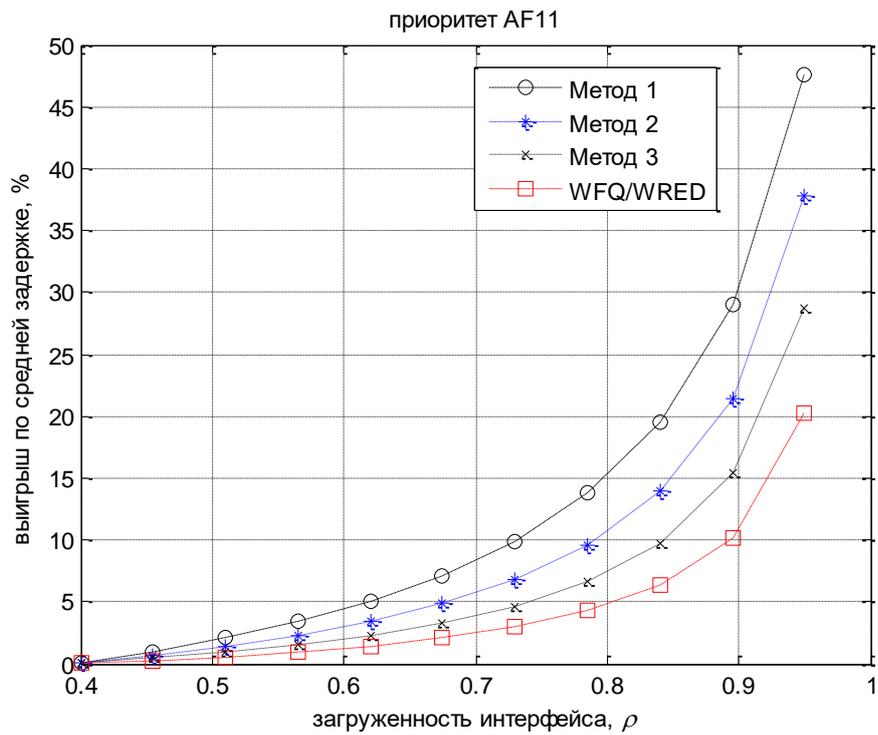


Рис. 4.10. Результаты сравнения решений по средней задержке пакетов с приоритетом AF11

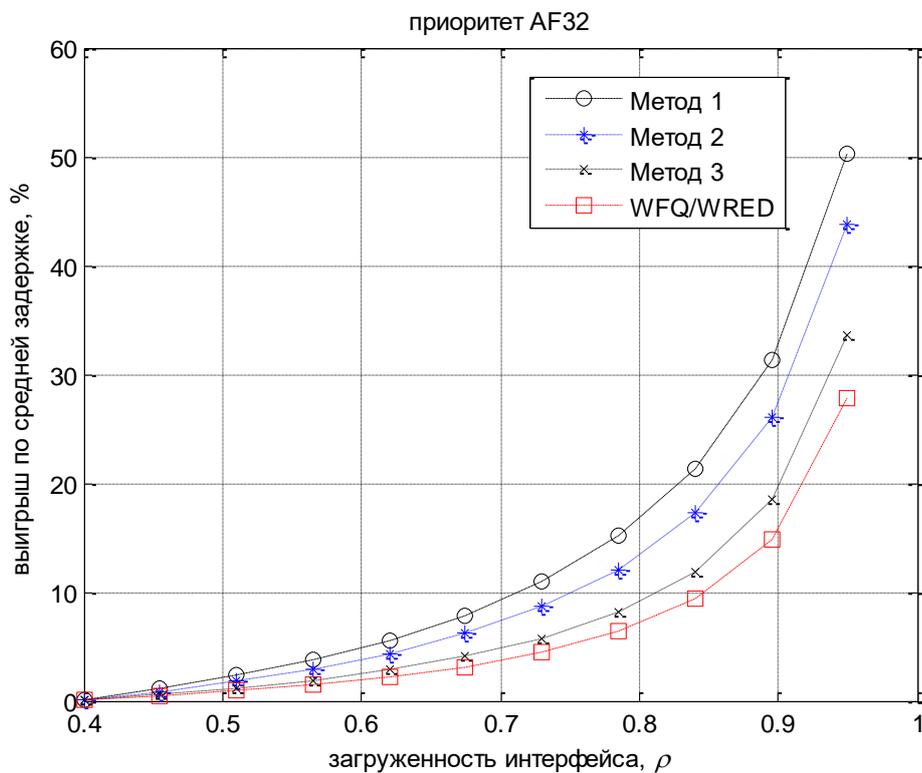


Рис. 4.11. Результаты сравнения решений по средней задержке пакетов с приоритетом AF32

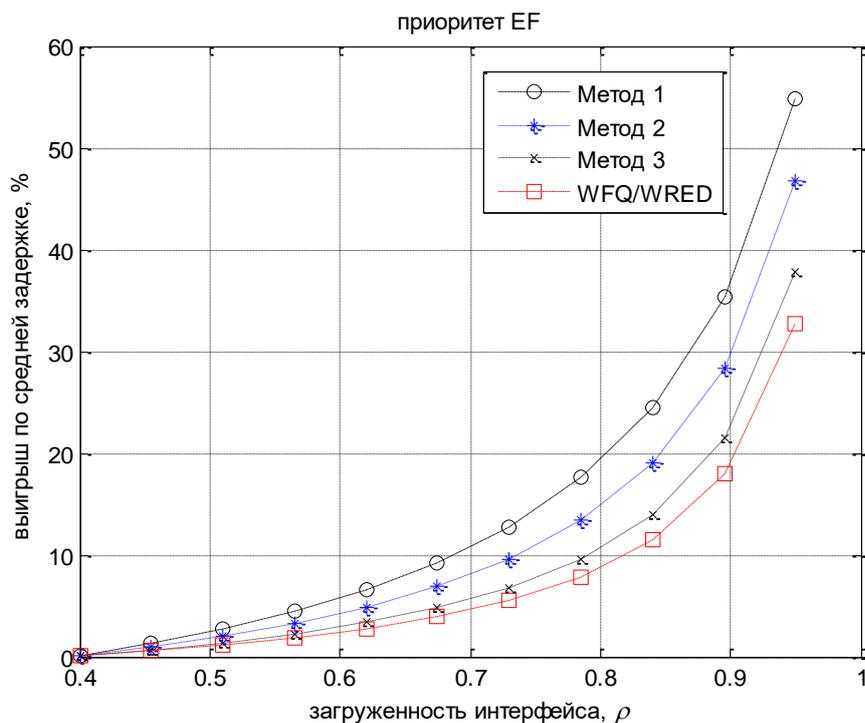


Рис. 4.12. Результаты сравнения решений по средней задержке пакетов с приоритетом EF

С другой стороны, использование предложенного метода позволяет повысить скорость передачи пакетов (при обеспечении одинаковых значений средней задержки):

- для потока AF11 в среднем на 15% по сравнению с WFQ/WRED; на 12-14% по сравнению с методом 3; на 5-6% по сравнению с методом 2 (рис. 4.10);

- для потока AF32 в среднем на 12-13% по сравнению с WFQ/WRED; на 10-11% по сравнению с методом 3; на 4-5% по сравнению с методом 2 (рис. 4.11);

- для потока EF в среднем на 11-12% по сравнению с WFQ/WRED; на 9-10% по сравнению с методом 3; на 4-5% по сравнению с методом 2 (рис. 4.12).

В области критических нагрузок ($\rho > 0,95$) при анализе UDP-потоков качество обслуживания для всех сравниваемых решений резко

ухудшалось ввиду перегрузки интерфейса. Для TCP-потоков за счет адаптивного изменения окна передачи (tcp window size) нагрузка на интерфейс регулировалась автоматически, стабилизируя значения средней задержки для всех рассматриваемых решений.

В ходе проведенного исследования выигрыш по вероятности потерь пакетов также для каждого m -го сравниваемого решения (WFQ/WRED, Метод 1, Метод 2. Метод 3) оценивался (по аналогии с выражением (4.1)) с помощью выражения

$$V_p^m = \left(1 - \frac{P_m}{P_{fifo}} \right) \cdot 100\%, \quad (4.2)$$

где P_{fifo} – вероятность потерь пакетов на интерфейсе при использовании решения FIFO/WRED;

P_m – вероятность потерь пакетов на интерфейсе при использовании m -го сравниваемого решения.

В ходе данного исследования также оценивалось влияние на вероятность потерь пакетов, имеющих разный приоритет, загруженности интерфейса. Результаты эксперимента показали, что предлагаемый метод позволил несколько улучшить значения вероятности потерь пакетов для приоритетных потоков. При этом в области малой и средней загруженности интерфейса ($0 \leq \rho \leq 0,7$) вне зависимости от используемого решения по управлению очередью она практически не переполнялась, а потери пакетов не возникали.

В области высоких нагрузок ($0,7 \leq \rho \leq 0,95$) использование предлагаемого в работе метода (Метод 1) позволило дифференцированно уменьшить вероятность потерь:

- для потока пакетов, передаваемых с приоритетом AF11, от 7,4 до 12% по сравнению с FIFO/WRED; от 4 до 6,4 % по сравнению с

WFQ/WRED; от 3,5 до 5% по сравнению с методом 3; от 2,4 до 2,7% по сравнению с методом 2 (рис. 4.13);

- для потока пакетов, передаваемых с приоритетом AF32, от 8,5 до 13,5% по сравнению с FIFO/WRED; от 4 до 6% по сравнению с WFQ/WRED; от 3 до 4,5% по сравнению с методом 3; от 1,5 до 2,5% по сравнению с методом 2 (рис. 4.14);

- для потока пакетов, передаваемых с приоритетом EF, от 8 до 14% по сравнению с FIFO/WRED; от 2,7 до 4,5% по сравнению с WFQ/WRED; от 2 до 3,5% по сравнению с методом 3; от 1,8 до 2,5% по сравнению с методом 2 (рис. 4.15).

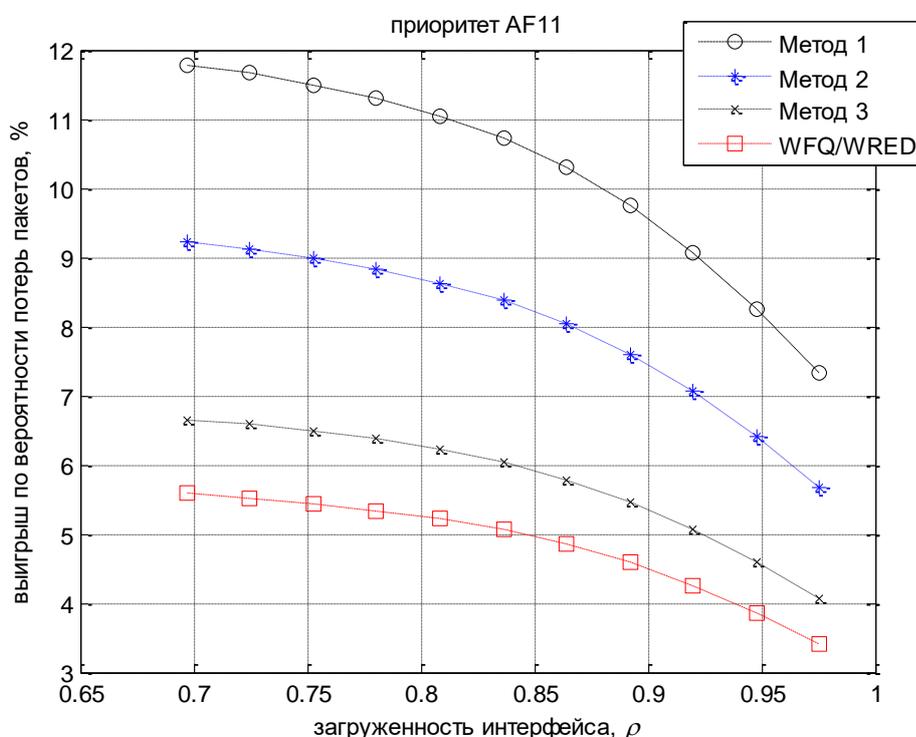


Рис. 4.13. Результаты сравнения решений по вероятности потерь пакетов с приоритетом AF11

Как и в предыдущем исследовании применение предложенного в подразделе 4.1 метода позволяет при обеспечении одной и той же вероятности потерь пакетов повысить скорость передачи пакетов:

- для потока AF11 в среднем от 8 до 15% по сравнению с методом 2 (рис. 4.13);

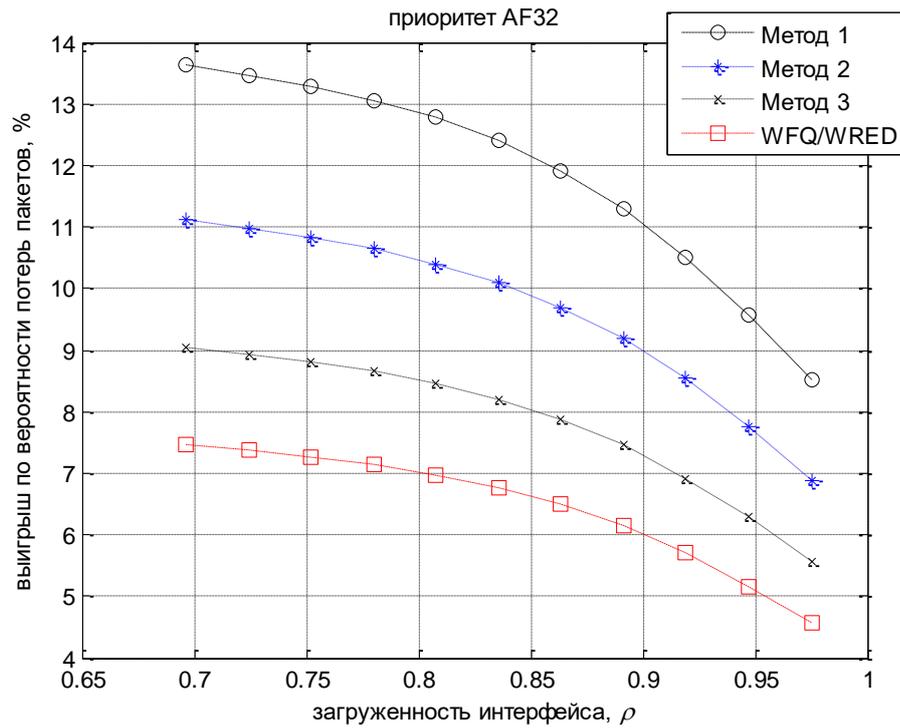


Рис. 4.14. Результаты сравнения решений по вероятности потерь пакетов с приоритетом AF32

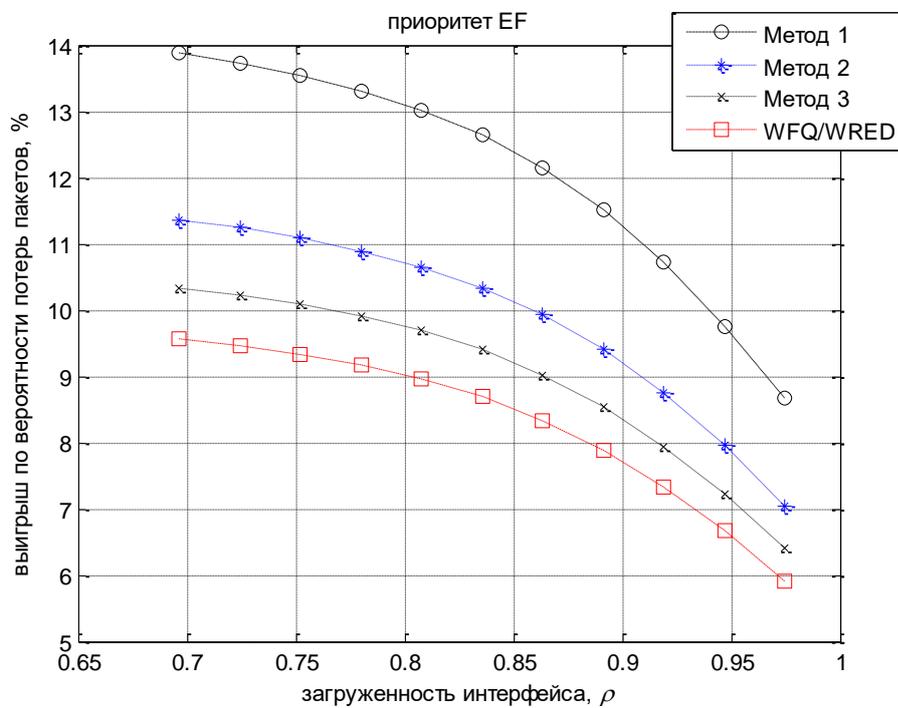


Рис. 4.15. Результаты сравнения решений по вероятности потерь пакетов с приоритетом EF

- для потока AF32 в среднем на 10-17% по сравнению с методом 2 (рис. 4.14);

- для потока EF от 8 до 18% по сравнению с методом 2 (рис. 4.15).

При разработке математических моделей ключевых сетевых процессов, к которым относятся и процессы управления очередями на интерфейсах маршрутизаторов, обязательным моментом является проверка их адекватности, т.е. степени соответствия аналитических результатов расчета тем показателям, которые получены (измерены) в рамках проводимого лабораторного эксперимента на реальном сетевом оборудовании.

В ходе экспериментальных исследований подтверждена адекватность предлагаемых моделей по показателям сходимости результатов аналитического моделирования и лабораторного эксперимента. Минимальное расхождение (в среднем до 6-13%) наблюдалось в области средних и высоких нагрузок (при $0,6 \leq \rho \leq 0,95$), где как раз и целесообразно повышение эффективности управления очередями для повышения качества обслуживания. В области низких нагрузок (при $0 \leq \rho \leq 0,55$) очереди и необходимости у управления ими отсутствовала. При критической загрузке интерфейса (при $\rho > 0,95$) происходила перегрузка интерфейса, которая должна устраняться такими средствами, как маршрутизация и профилирование трафика.

4.3. Разработка рекомендаций по практической реализации предложенных моделей и метода управления очередями на маршрутизаторах транспортной телекоммуникационной сети

В основу выработки рекомендаций по практическому использованию предлагаемых в работе моделей и метода управления очередями всецело положены результаты аналитических расчетов и

лабораторных исследований. Это продиктовано необходимостью обоснования области наиболее эффективного применения результатов диссертационной работы с точки зрения обеспечения наиболее высоких показателей качества обслуживания при усовершенствовании существующих и (или) проектировании новых механизмов управления очередями на маршрутизаторах ТКС.

В общем случае научно-практические рекомендации непосредственно связаны с необходимостью проведения комплекса организационных и технических мероприятий, затрагивающих изменение как самой архитектуры маршрутизатора, так и некоторых базовых принципов их функционирования на уровне алгоритмического и программного обеспечения.

К перечню **общих рекомендаций**, прежде всего, стоит отнести:

- реализацию динамических стратегий управления очередями с адаптацией к возможному изменению состояния интерфейса и характеристик потока;
- использование метрик, максимально учитывающих основные показатели качества обслуживания;
- поддержка политик дифференцированной обработки пакетов с учетом их классов и, прежде всего, приоритета;
- реализация стратегии маршрутизации «от источника»;
- поддержка иерархических решений (иерархических очередей) для повышения масштабируемости процесса управления очередями в целом.

Рекомендации по усовершенствованию программно-аппаратной архитектуры маршрутизатора.

Как показано в работах [102], архитектура современного маршрутизатора транспортной сети может быть условно представлена в виде трех основных функциональных уровней (рис. 4.16):

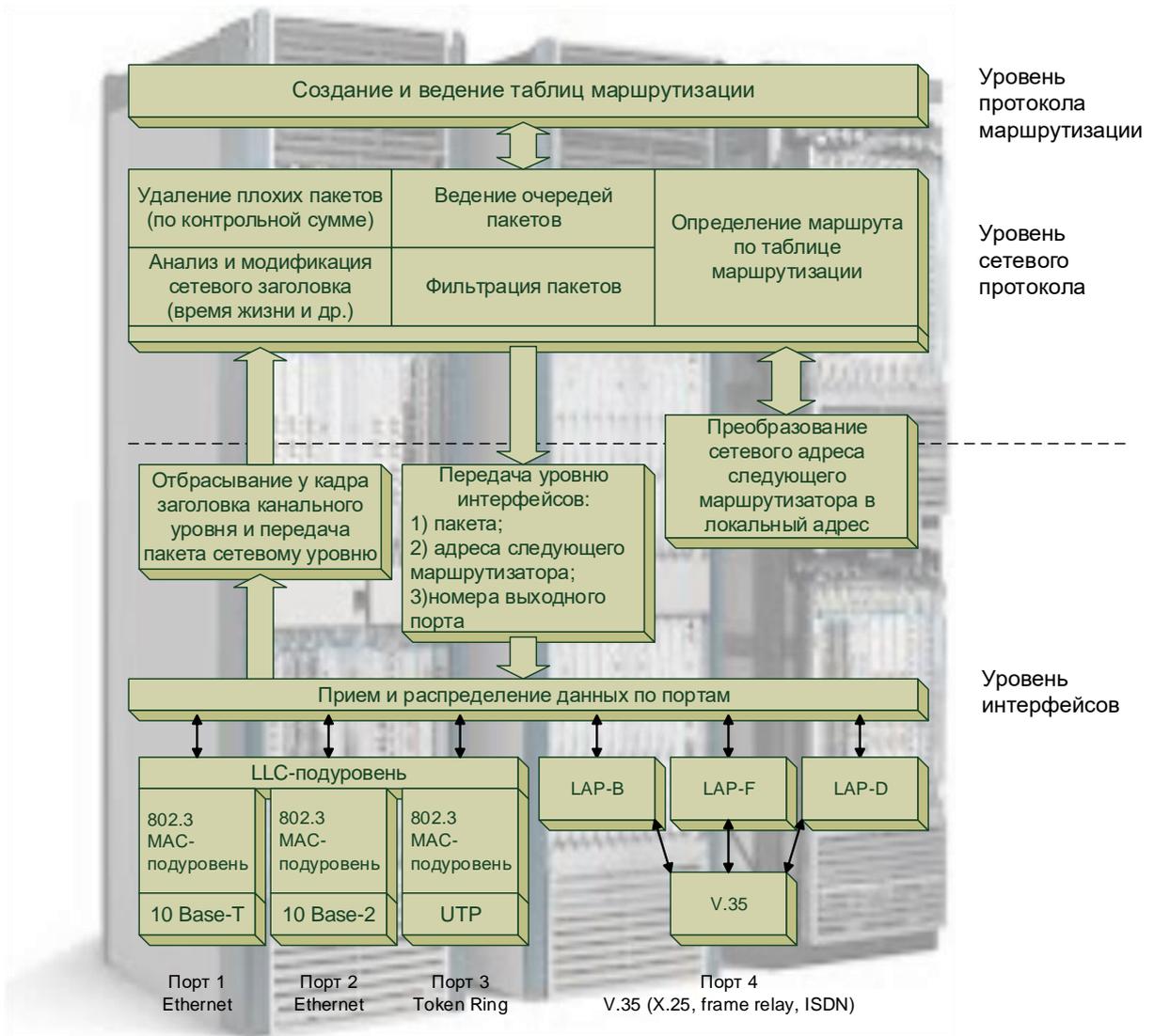


Рис. 4.16. Функциональная модель маршрутизатора

1. Уровень интерфейсов.
2. Уровень сетевого протокола.
3. Уровень протокола маршрутизации.

На *уровне интерфейсов* маршрутизатор обеспечивает физическое взаимодействие со средой передачи, организуя согласование уровней электрических сигналов, кодирование (линейное, логическое) и пр. В разных моделях маршрутизаторов могут предусматриваться различные наборы физических интерфейсов, представляющих собой комбинацию портов для соединения локальных и территориально-распределенных сетей. В свою очередь с каждым интерфейсом для подключения локальной сети ассоциирован определенный протокол канального уровня (Ethernet, Token Ring, FDDI), использование которого и определяет пропускную способность интерфейса – объем канального ресурса, подлежащего последующему распределению в ходе решения задач управления очередями. Перечень, тип и производительность физических интерфейсов, поддерживаемых на маршрутизаторе, определяет его область применения, возможности по подключению и коммутации других маршрутизаторов и сетей доступа.

На *уровне сетевого протокола* обеспечивается решение множества задач, отмеченных на рис. 4.16, ключевыми среди которых является автоматическая (а не ручная) классификация потоков, их агрегация на основе близости классов обслуживания, а также поддержка и обслуживание (ведение) очередей. При решении перечисленных задач с помощью предложенных моделей и метода удастся придать высокую автоматичность процессу управления очередями и обеспечить согласованность работы средств классификации, агрегирования потоков, организации и обслуживания очередей, т.к. управляющие переменные, отвечающие за данные задачи, будут определяться совместно.

За подобные решения должно отвечать не множество отдельных механизмов, а единый программный модуль, который бы, основываясь на результатах постоянного мониторинга состояния интерфейса и характеристик потока, в области малой загрузки ($0 \leq \rho \leq 0,55$) интерфейса включал алгоритм FIFO, т.к. вероятность возникновения очереди пакетов достаточно мала, а раз отсутствует объект управления, то и незачем вносить дополнительную задержку, связанную с организацией очередей и их поддержкой. В этой области загрузки интерфейса, как показали результаты исследования, дисциплина обслуживания пакетов на QoS-показатели практически не влияет.

В области средних и высоких нагрузок ($0,6 \leq \rho \leq 0,95$) целесообразно автоматически включать механизм управления очередями, работа которого должна быть основана на предлагаемом методе. Важно отметить, что численное значение «по умолчанию» порога включения такого механизма рекомендуется включить в перечень параметров, которые можно изменять административно. Кроме порога к числу таких параметров, при необходимости модифицируемых администратором, рекомендуется включить и метрики, входящие в выражение (2.17).

В области критичных нагрузок ($\rho > 0,95$) дисциплина обслуживания пакетов снова не играет определяющей роли, а борьба с перегрузкой целиком и полностью должна быть переложена на средства активного управления очередью, реализуемых в рамках предложенного метода с помощью предварительно рассчитанных переменных α_i . При этом степень превентивности отбрасывания пакетов рекомендуется регулировать через соотношение метрик в выражении (2.17), а также с учетом классов пакетов и очередей.

Для повышения эффективности решения задач управления очередями, организованного с помощью предложенного метода, целесообразно на уровне сетевого протокола использовать программно

реализованные анализаторы трафика с поддержкой функции прогнозирования числа и характеристик потоков, поступающих на тот или иной интерфейс. Для этого необходимо собирать и анализировать статистику о параметрах обслуживаемых потоков на определенном временном интервале, а также учитывать особенности реализуемой стратегии маршрутизации с помощью настроенного на данном маршрутизаторе соответствующего протокола.

На верхнем уровне маршрутизатора, т.е. на *уровне создания и ведения таблиц маршрутизации* рекомендовано включить в перечень маршрутных метрик параметры не только номинальной пропускной способности интерфейса (как это сделано в большинстве существующих протоколов), но и данные о загруженности интерфейса, обеспечив, таким образом, реализацию многопутевой стратегии маршрутизации на основе балансировки очередей [16]. Это позволит согласовать решения по маршрутизации и управлению очередями, создав условия для улучшения межконцевого качества обслуживания в целом.

К числу рекомендаций, касающихся к усовершенствованию аппаратной части маршрутизатора, стоит отнести и поддержку многопроцессорной (многоядерной) архитектуры, которая уже интенсивно используется в новейших образцах маршрутизирующих устройств [36, 39].

4.4. Выводы по четвертому разделу

1. В разделе предложено усовершенствование метода управления очередями на интерфейсах маршрутизаторов транспортных телекоммуникационных сетей, основанного на предложенных во втором и третьем разделах математических моделях. Новизна метода заключается в том, что с его помощью оптимизируется работа интерфейса в ходе согласованного решения задач по агрегированию потоков, распределению

пакетов по очередям и выделению очередям необходимой пропускной способности интерфейса, а также задач по задействованию и контролю использования активных процессоров маршрутизатора. Предложенный метод управления очередями на маршрутизаторах транспортных ТКС в целом отвечает требованиям концепции Traffic Engineering Queues, нацеленных на обеспечение сбалансированного и дифференцированного использования канального и буферного ресурса.

2. С целью проверки адекватности предложенных потоковых моделей и эффективности решений, получаемых с помощью усовершенствованного метода управления очередями на маршрутизаторах ТКС, проведено экспериментальное исследование на лабораторном сетевом оборудовании – маршрутизаторах производства компании Cisco серии 2800. В ходе эксперимента задействованы широкие возможности пакета D-ITG, который использовался как генератор и анализатор трафика. Реализация на оборудовании управляющих параметров, рассчитанных с помощью предложенного метода, осуществлялась с помощью механизма CBWFQ (для организации и обслуживания очередей) с применением списков доступа (ACL) – для агрегирования потоков и распределения их по очередям интерфейса, а также с помощью механизма WRED – в ходе задания рассчитанных значений вероятности отбрасывания пакетов.

3. В ходе экспериментальных исследований подтверждена адекватность предлагаемых моделей по показателям сходимости результатов аналитического моделирования и лабораторного эксперимента. Минимальное расхождение (в среднем до 6-13%) наблюдалось в области средних и высоких нагрузок (при $0,6 \leq \rho \leq 0,95$), где как раз и целесообразно повышение эффективности управления очередями для повышения качества обслуживания. В области низких нагрузок (при $0 \leq \rho \leq 0,55$) очереди и необходимости в управлении ими отсутствовала. При критической загруженности интерфейса (при $\rho > 0,95$)

происходила перегрузка интерфейса, которая должна устраняться такими средствами, как маршрутизация и профилирование трафика.

4. В области средних и высоких нагрузок ($0,6 \leq \rho \leq 0,95$) использование предлагаемого в работе метода (Метод 1) позволило дифференцированно улучшить среднюю задержку пакетов:

- для потока пакетов, передаваемых с приоритетом AF11, от 5-15% до 30-45% по сравнению с FIFO/WRED; от 3-8% до 12-20% по сравнению с WFQ/WRED; от 3-5% до 11-17% по сравнению с методом 3; от 2-4% до 6-12% по сравнению с методом 2;

- для потока пакетов, передаваемых с приоритетом AF32, от 4-9% до 32-47% по сравнению с FIFO/WRED; от 2-6% до 15-22% по сравнению с WFQ/WRED; от 2-4% до 12-16% по сравнению с методом 3; от 1-3% до 6-7% по сравнению с методом 2;

- для потока пакетов, передаваемых с приоритетом EF, от 5-10% до 35-50% по сравнению с FIFO/WRED; от 2-5% до 17-23% по сравнению с WFQ/WRED; от 2-4% до 13-17% по сравнению с методом 3; от 1-3% до 7-8% по сравнению с методом 2.

С другой стороны, использование предложенного метода позволяет повысить скорость передачи пакетов (при обеспечении одинаковых значений средней задержки):

- для потока AF11 в среднем на 15% по сравнению с WFQ/WRED; на 12-14% по сравнению с методом 3; на 5-6% по сравнению с методом 2;

- для потока AF32 в среднем на 12-13% по сравнению с WFQ/WRED; на 10-11% по сравнению с методом 3; на 4-5% по сравнению с методом 2;

- для потока EF в среднем на 11-12% по сравнению с WFQ/WRED; на 9-10% по сравнению с методом 3; на 4-5% по сравнению с методом 2.

5. В области высоких нагрузок ($0,7 \leq \rho \leq 0,95$) использование предлагаемого в работе метода (Метод 1) позволило дифференцированно уменьшить вероятность потерь:

- для потока пакетов, передаваемых с приоритетом AF11, от 7,4 до 12% по сравнению с FIFO/WRED; от 4 до 6,4 % по сравнению с WFQ/WRED; от 3,5 до 5% по сравнению с методом 3; от 2,4 до 2,7% по сравнению с методом 2;

- для потока пакетов, передаваемых с приоритетом AF32, от 8,5 до 13,5% по сравнению с FIFO/WRED; от 4 до 6% по сравнению с WFQ/WRED; от 3 до 4,5% по сравнению с методом 3; от 1,5 до 2,5% по сравнению с методом 2;

- для потока пакетов, передаваемых с приоритетом EF, от 8 до 14% по сравнению с FIFO/WRED; от 2,7 до 4,5% по сравнению с WFQ/WRED; от 2 до 3,5% по сравнению с методом 3; от 1,8 до 2,5% по сравнению с методом 2.

Кроме того, применение предложенного в работе метода позволяет при обеспечении одной и той же вероятности потерь пакетов повысить скорость передачи пакетов:

- для потока AF11 в среднем от 8 до 15% по сравнению с методом 2;
- для потока AF32 в среднем на 10-17% по сравнению с методом 2;
- для потока EF от 8 до 18% по сравнению с методом 2.

6. На основе полученных в разделе результатов экспериментального исследования, предложенных в диссертационной работе решений, представленных потоковыми моделями и методом управления очередями, сформулированы рекомендации по их практическому использованию в существующих и перспективных телекоммуникационных технологиях и сетях. Применение на практике разработанных рекомендаций связано с проведением ряда технологических мероприятий, которые несколько затрагивают саму архитектуру, а также базовый функционал

алгоритмического и программного обеспечения современных и вновь проектируемых маршрутизаторов ТКС. Выполнение представленных рекомендаций должно способствовать улучшению уровня качества обслуживания в ТКС в целом.

7. Полученные в данном разделе результаты опубликованы в работах [9, 14, 15].

ВЫВОДЫ ПО РАБОТЕ

В диссертационной работе решена актуальная научно-прикладная задача, которая заключалась в оптимизации процесса управления очередями на интерфейсах маршрутизаторов на основе обеспечения согласованного решения задач по агрегированию потоков, распределения пакетов между очередями и выделения очередям пропускной способности интерфейса путем разработки соответствующих математических моделей и метода для повышения качества обслуживания в транспортных ТКС в целом. По результатам решения поставленной научной задачи можно сделать ряд важных выводов.

1. В работе показано, что важнейшая роль в архитектуре обеспечения качества обслуживания в современных телекоммуникационных сетях отводится средствам управления трафиком: протоколам маршрутизации и резервирования ресурсов, алгоритмам классификации и маркировки пакетов, механизмам управления очередями и профилирования трафика. Несмотря на важность всех слагаемых технологий управления трафиком, ключевое место среди них занимают механизмы управления очередями, т.к. именно неэффективное управление очередями приводит к снижению качества обслуживания, проявляющегося в неконтролируемом росте задержек и уровня потерь пакетов на интерфейсах маршрутизаторов ТКС.

2. Проведен анализ основных известных математических методов и технологических механизмов управления очередями. На основе выявленных недостатков существующих решений сформулирован перечень требований, которые предъявляются к перспективным моделям, методам и механизмам по управлению. К основным из них относятся: учет потоковой структуры современного сетевого трафика; оптимизационная постановка и решение задачи управления очередями; поддержка

дифференциации обслуживания пакетов на интерфейсах маршрутизаторов ТКС в соответствии с их QoS-требованиями; реализация динамических стратегий управления очередями; простота алгоритмически-программной и аппаратной реализации; а также обеспечение согласованного решения взаимосвязанных интерфейсных задач (классификация и маркировка пакетов; создание и настройка системы очередей на интерфейсе; агрегирование потоков и распределение пакетов по очередям интерфейса с учетом параметров передаваемых потоков, требований к качеству обслуживания, характеристик создаваемых очередей и интерфейса в целом; определение порядка обслуживания очередей; распределение пропускной способности интерфейса между отдельными очередями; превентивное (заблаговременное) ограничение длины очереди).

3. Получила дальнейшее развитие потоковая модель управления очередями на интерфейсах маршрутизаторов транспортных телекоммуникационных сетей. Новизна модели заключается в том, что в ней в ходе управления очередями осуществляется оптимальное агрегирование и распределение потоков на основании сравнения и минимизации расхождения классов потоков и очередей, в которые они направляются. Использование модели позволило в большинстве случаев сократить число поддерживаемых очередей на интерфейсе без снижения уровня дифференциации качества обслуживания, что способствовало минимизации времени на обработку пакетов на интерфейсе и снижению их межконцевой задержки в сети в целом.

4. Получила дальнейшее развитие модель управления очередями на интерфейсах маршрутизаторов активных сетей. Новизна модели заключается в том, что в ходе обслуживания пакетов учитываются не только количество и параметры потоков и очередей, но и характеристики активных процессоров (тип, загруженность и производительность), входящих в архитектуру маршрутизатора активной сети. Реализация предложенной модели позволила обеспечить более эффективное

использование буферного и вычислительного ресурса маршрутизатора активных сетей.

5. Усовершенствован метод управления очередями на интерфейсах маршрутизаторов транспортных телекоммуникационных сетей. Новизна метода заключается в том, что с его помощью оптимизируется работа интерфейса в ходе согласованного решения задач по агрегированию потоков, распределению пакетов по очередям и выделению очередям необходимой пропускной способности интерфейса, а также задач по задействованию и контролю использования активных процессоров маршрутизатора.

6. В ходе проведенного исследования предложенных и усовершенствованных в работе моделей и метода управления очередями установлено, что их применение позволяет, во-первых, в зависимости от соотношения числа потоков и очередей, а также их классов сократить количество поддерживаемых очередей от 15-18% до 25-33% без снижения уровня дифференциации качества обслуживания потоков пользователей, что способствует сокращению времени на обработку пакетов на интерфейсе и минимизации межконцевой задержки пакетов в сети в целом; во-вторых, улучшить качество обслуживания по показателям средней задержки пакетов от 7-10% до 14-18% или вероятности потерь пакетов от 11 до 15% в зависимости от характеристик потоков пакетов, числа организованных очередей и параметров интерфейса.

7. В ходе исследований был организован лабораторный эксперимент на реальном сетевом оборудовании компании Cisco Systems лаборатории распределения информации кафедры телекоммуникационных систем ХНУРЭ. Результаты экспериментального исследования предложенных в диссертационной работе научных результатов в целом подтвердили их адекватность и эффективность с точки зрения практической реализации. Научные и практические результаты диссертационных исследований также использованы при модернизации и

развитии телекоммуникационной системы ООО «Сигма СОФТВЕА» и в учебном процессе кафедры телекоммуникационных систем ХНУРЭ в дисциплине «Системы коммутации и распределения информации, часть 2». По всем реализациям результатов работы составлены необходимые акты.

СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. Рекомендация МСЭ-Т Y.110 «Принципы и архитектура глобальной информационной инфраструктуры» (ITU-T Rec. Y.110 «Global Information Infrastructure principles and framework architecture»). – 1998. – 47 с.
2. Сети следующего поколения NGN / Под ред. А.В. Рослякова. // М: Эко-Трендз, 2008. – 424 с.
3. Багатоканальний електрозв'язок та телекомунікаційні технології: Підручник / О.В. Лемешко, В.В. Поповський, В.А. Лошаков, за ред. В.В. Поповського // Харків: ТОВ "Компанія СМІТ", 2010. – 470 с.
4. Álvarez S. QoS for IP/MPLS networks / S. Álvarez. // Cisco Press, 2006. – 299 p.
5. Мультисервисные АТМ сети / Денисова Т.Б., Лихтциндер Б.Я., Назаров А.Н., Симонов М.В., Фомичев С.М. – М.: Эко Трендз, 2005. – 320 с.
6. Вегешна Ш. Качество обслуживания в сетях IP: пер. с англ. / Ш. Вегешна – М.: Изд. дом «Вильямс», 2003. – 368 с.
7. Tsai T.-Y., Chung Y.-L., Tsai Z. Introduction to Packet Scheduling Algorithms for Communication Networks // Communications and Networking, Jun Peng. – 2010. – 434 p.
8. Лемешко А.В. Управление очередями на узлах активной сети / А.В. Лемешко, М. Ватти, А.В. Симоненко // Радиотехника: Всеукр. межвед. науч.- техн. сб. 2007. – Вып. 151. – С. 92-97.
9. Симоненко А.В. Модель динамического управления очередями и пропускной способностью канала связи на маршрутизаторах мультисервисной сети / А.В. Симоненко, Ахмад Хайлан, Али Али // Радиотехника: Всеукр. межвед. науч.- техн. сб. 2008. – Вып. 155. – С. 164-168.
10. Лемешко А.В. Поточковая модель управления очередями с динамическим распределением пропускной способности исходящего канала связи / А.В. Лемешко, А.В. Симоненко, Махмуд Ватти // Наукові записки УНДІЗ. – 2008. – №3 (5). – С.34-39.
11. Лемешко А.В. Математическая модель динамического управления каналным и буферным ресурсом на узлах телекоммуникационной сети / А.В. Лемешко, А.В. Симоненко // Радиотехника: Всеукр. межвед. науч.- техн. сб., 2009. – Вып. 156. – С. 36-41.
12. Лемешко А.В. Поточковая модель динамической балансировки очередей на узлах телекоммуникационной сети / А.В. Лемешко, А.В. Симоненко, С.И. Сивашенко // Радиотехника: Всеукр. межвед. науч.- техн. сб., 2009. – Вып. 159. – С. 46-49.

13. Али С. Али. Поточковая модель динамической балансировки очередей в MPLS-сети с поддержкой Traffic Engineering Queues: [Электронный ресурс] / Али С. Али, А.В. Симоненко // Проблемы телекоммуникаций. – 2010. – № 1 (1). – С. 59-67. – Режим доступа до журн.: http://pt.journal.kh.ua/2010/1/1/101_ali_balancing.pdf.

14. Семеняка М.В. Разработка и исследование метода динамической балансировки очередей на маршрутизаторах мультисервисной телекоммуникационной сети: [Электронный ресурс] / М.В. Семеняка, А.В. Симоненко, Али С. Али // Проблемы телекоммуникаций. – 2012. – № 3(8). – С. 66-78. – Режим доступа до журн.: http://pt.journal.kh.ua/2012/3/1/123_semenyaka_balancing.pdf.

15. Семеняка М.В. Экспериментальное исследование метода динамической балансировки очередей на маршрутизаторах мультисервисной телекоммуникационной сети / М.В. Семеняка, А.В. Симоненко // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. – Ростов-на-Дону: ПЦ "Университет" СКФ МТУСИ, 2013. – С.220-223.

16. Вавенко Т.В. Поточковая модель маршрутизации с балансировкой нагрузки по длине очереди в программно-конфигурируемых сетях / Т.В. Вавенко, В.Л. Стерин, А.В. Симоненко // Научно-технический вестник информационных технологий, механики и оптики. – 2013. – № 4 (86). – С. 38-45.

17. Лемешко А.В. Модель активного управления очередями на маршрутизаторах телекоммуникационной сети / А.В. Лемешко, М.В. Семеняка, А.В. Симоненко // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. – Ростов-на-Дону: ПЦ "Университет" СКФ МТУСИ'2015. – Ч.1 – С. 244-248.

18. Симоненко А.В. Математическая модель управления очередями на маршрутизаторах телекоммуникационной сети на основе оптимального агрегирования потоков и распределения пакетов по очередям [Электронный ресурс] / А.В. Симоненко, Д.В. Андрушко // Проблемы телекоммуникаций. – 2015. – № 1 (16). – С. 94 - 102. – Режим доступа до журн.: http://pt.journal.kh.ua/2015/1/1/151_simonenko_queue.pdf.

19. Лемешко А.В. Математическая модель управления сетевыми ресурсами с адаптивным ограничением абонентского трафика / А.В. Лемешко, Ю.Н. Добрышкин, А.В. Симоненко // Перспективные технологии в средствах передачи информации: VII междунар. науч.-техн. конф., 10-12 октября 2007 г. – Владимир: РОСТ, 2007. – С. 61-64.

20. Симоненко О.В. Модель управління чергами на вузлах активної мережі / О.В. Симоненко, М. Ватті // Сучасні телекомунікаційні та інформаційні технології: 2-й наук.-

практ. семінар молодих науковців та студентства., 12-14 грудня 2007 р. – Київ.: Український науково-дослідний інститут зв'язку, 2007. – С.73-76.

21. Добрышкин Ю.Н. Модель адаптивного ограничения абонентского трафика в мультисервисных сетях / Ю.Н. Добрышкин, А.В. Симоненко // Радиоэлектроника, электротехника и энергетика: четырнадцатая Междунар. науч.-техн. конф. студентов и аспирантов, 28-29 февраля 2008 г.: тезисы докл. – М.: Издательский дом Московского Энергетического института, 2008. – Т.1 – С. 129-130.

22. Добрышкин Ю.Н., Модель согласованного управления трафиком и буферным ресурсом в телекоммуникационной сети / Ю.Н. Добрышкин, А.В. Симоненко // Проблемы телекоммуникаций: друга Міжнар. наук.-техн. конф., 20-23 травня 2008 р. : збірник тез. – К.: НТУУ "КПІ", 2008. – С.207-208.

23. Симоненко А.В. Математическая модель управления очередями на транзитных узлах активной сети / А.В. Симоненко, Махмуд Ватти // Радиоэлектроника и молодежь в XII веке: 12-й Междунар. молодежный форум, 1-3 апреля 2008 г.: сб. материалов форума – Х.: ХНУРЭ, 2008. – Ч.1 – С.161.

24. Лемешко А.В. Модели управления очередями с динамическим распределением пропускной способности исходящего канала связи / А.В. Лемешко, А.В. Симоненко, М. Ватти // Прикладная радиоэлектроника. Состояние и перспективы развития: 3-й Междунар. радиоэлектронный форум МРФ-2008, 22-24 октября 2008 г.: сб. науч. трудов – Х.: АНПРЭ, ХНУРЭ, 2008. – Т. II – С. II-69–II-72.

25. Симоненко А.В. Модель управления буферным и канальным ресурсом мультисервисной сети / А.В. Симоненко, Ахмад Хайлан, Али Али // Радиоэлектроника и молодежь в XXI веке: 13-й Междунар. молодежный форум, 30 марта-1 апреля 2009 г.: сб. материалов форума – Х.: ХНУРЭ, 2009. – Ч.1 – С.161.

26. Лемешко А.В. Математическая модель динамического управления канальным и буферным ресурсом на узлах телекоммуникационной сети / А.В. Лемешко, А.В. Симоненко // Новітні технології – для захисту повітряного простору: п'ята наук. конф. Харківського університету Повітряних Сил імені Івана Кожедуба, 15-16 квітня 2009 р.: тези доповідей – Х.: ХУ ПС, 2009. – С.98-99.

27. Симоненко А.В. Поточковая модель внутриузлового управления сетевыми ресурсами в мультисервисной телекоммуникационной системе / А.В. Симоненко, Ахмад М. Хайлан, Али С. Али // Перспективные технологии в средствах передачи информации: VIII Междунар. науч.-техн. конф., 21-22 мая 2009 г.: материалы конф. – Владимир: Владим. гос. ун-т., 2009. – Т. 2. – С. 94-97.

28. Алі С. Алі. Потокова модель динамічного балансування завантаженості черг на вузлах MPLS-мережі / Алі С. Алі, О.В. Симоненко // Сучасні проблеми телекомунікацій – 2010: наук.-практ. конф., 28-30 жовтня 2010 р.: матеріали конф. - Львів: НУ "Львівська політехніка", 2010. - С. 40-43.

29. Женжера С.В. Математическая модель с динамическим управлением сетевыми ресурсами на узлах телекоммуникационной сети / С.В. Женжера, А.В. Симоненко // Новітні технології – для захисту повітряного простору: шоста наук. конф. Харківського університету Повітряних Сил імені Івана Кожедуба, 14-15 квітня 2010 р.: тези доповідей – Х.: ХУ ПС, 2010. – С.116-117.

30. Симоненко А.В. Потоковая модель динамической балансировки загруженности очередей на узлах IP-сети / А.В. Симоненко, Али С. Али, А.С. Бильк, // Компьютерное моделирование в наукоемких технологиях - КМНТ-2010: науч.-техн. конф. с междунар. участием, 18-21 мая 2010 г.: сб. науч. трудов – Х.: ХНУ имени В.Н. Каразина, 2010. – Ч. 2. – С.231-234.

31. Лемешко А.В. Математическая модель динамической балансировки загруженности очередей на узлах телекоммуникационной сети / А.В. Лемешко, А.В. Симоненко // Сучасні проблеми і досягнення в галузі радіотехніки, телекомунікацій та інформаційних технологій: V міжнар. наук.-техн. конф., 22-24 вересня 2010 р. – Запоріжжя: ЗНТУ, 2010. – С.111-112.

32. Симоненко А.В. Динамическая балансировка загруженности очередей на узлах телекоммуникационной сети / А.В. Симоненко, А.Г. Беленков // Сучасні телекомунікаційні та інформаційні технології: наук.-практ. семінар молодих науковців, травень-червень 2010 р.: наук.-виробн. збірник – Київ.: Український науково-дослідний інститут зв'язку, 2010. – №1 – С.16-18.

33. Лемешко А.В. Потоковая модель динамической балансировки очередей в MPLS-сети с поддержкой Traffic Engineering Queues / А.В. Лемешко, А.В. Симоненко // Новітні технології – для захисту повітряного простору: 7 наук. конф. Харківського університету Повітряних Сил імені Івана Кожедуба, 13-14 квітня 2011 р.: тези доповідей – Х.: ХУ ПС, 2011. – С.158-159.

34. Симоненко А.В. Потоковая модель динамической балансировки загруженности очередей в MPLS-сети с поддержкой Traffic Engineering Queues // Прикладная радиоэлектроника. Состояние и перспективы развития: 4-й Междунар. радиоэлектронный форум МРФ-2011, 18-21 октября 2011 г.: сб. науч. трудов – Х.: АНПРЭ, ХНУРЭ. 2011. – Т. II – С. II-76–II-80.

35. Женжера С.В. Управление очередями на узлах телекоммуникационной сети / С.В. Женжера, А.В. Симоненко // Новітні технології – для захисту повітряного простору: восьма наук. конф. Харківського університету Повітряних Сил імені Івана Кожедуба, 18-19 квітня 2012 р.: тези доповідей – Х.: ХУ ПС, 2012. – С.195-196.

36. Семеняка М.В. Исследование концепции иерархических очередей для решения задач управления перегрузками в телекоммуникационной сети / М.В. Семеняка, А.В. Симоненко // Сучасні інформаційно-комунікаційні технології (COMINFO'2012 – Livadia): VIII міжнар. наук.-техн. конф., 1-5 жовтня 2012 р.: зб. тез – К: ДУІКТ, 2012. – С. 68-70.

37. Lemeshko O. Researching and designing of the dynamic adaptive queue balancing method on telecommunication network routers / O. Lemeshko, M. Semenyaka, O. Simonenko // The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM'2013): XII international conference, 19-23 Febr. 2013. – Lviv: Lviv Polytechnic Publishing House, 2013. – P. 204-207.

38. Женжера С.В. Динамическая балансировка очередей в MPLS-сети с поддержкой Traffic Engineering Queues / С.В. Женжера, А.В. Симоненко // Новітні технології – для захисту повітряного простору: дев'ята наук. конф. Харківського університету Повітряних Сил імені Івана Кожедуба, 17-18 квітня 2013 р.: тези доповідей – Х.: ХУ ПС, 2013. – С.199.

39. Семеняка М.В. Двухуровневая иерархически-координационная модель распределения трафика пользователей на узлах телекоммуникационной сети / М.В. Семеняка, А.В. Симоненко // СВЧ-техника и телекоммуникационные технологии: 23-я междунар. Крымская конф., 8-13 сентября 2013 г.: материалы конф. – Севастополь: СевНТУ, 2013. – С. 527-528.

40. Lemeshko O. A queue management model on router of active network / O. Lemeshko, Ali Salem. Ali, O. Simonenko // The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM'2015): XIII international conference, 24-27 Febr. 2015: proceedings – Lviv: Lviv Polytechnic Publishing House, 2015. – P. 419-421.

41. Патент на корисну модель 47461 Україна, МПК H04L12/56. Спосіб управління трафіком в мультисервісній телекомунікаційній мережі / О.В. Лемешко, Ю.М. Добришкін, О.В. Симоненко, К.С. Васюта; власник Харківський університет Повітряних Сил ім. Івана Кожедуба. – № u 2009 05821; Заявл. 09.06.2009, Опубл. 10.02.2010. Бюл. №3.

42. Патент на корисну модель 70088 Україна, МПК G06G 3/00 (2012.01). Спосіб управління чергами на маршрутизаторах мультисервісної телекомунікаційної мережі / О.В. Лемешко, Алі Салем Алі, М.В. Семеняка, О.В. Симоненко; власник Харківський

національний університет радіоелектроніки. – № и 2011 13675 Заявл. 21.11.2011, Опубл. 25.05.2012. Бюл. №10.

43. Олвейн В. Структура и реализация современной технологии MPLS: Пер. с англ. М.: Издательский дом «Вильямс», 2004. 480 с.

44. Osborn E. Traffic Engineering with MPLS / E. Osborn, A. Simba. – Cisco Press, 2002. – 608 p.

45. Thiruchelvi G. A Survey On Active Queue Management Mechanisms / G. Thiruchelvi, J. Raja // IJCSNS International Journal of Computer Science and Network Security. – 2008. – Vol. 8, №. 12. – Pp. 130 – 145.

46. Gwyn Chatranona A survey of TCP-friendly router-based AQM schemes / Gwyn Chatranona, Miguel A. Labradorb, Sujata Banerjee. – Computer Communications. – 2004. – PP. 1424 – 1440.

47. Socrates C. Congestion Control for Packet Switched Networks: A Survey / C. Socrates, P.M. Beulah Devamalar, R. Kannamma Sridharan // International Journal of Scientific and Research Publications. – 2014. – Vol.4. – P. 1-6;

48. Jianyong Chen. Self-Tuning Random Early Detection Algorithm to Improve Performance of Network Transmission / Jianyong Chen, Cunying Hu, and Zhen Ji // Mathematical Problems in Engineering. – 2012. – P. 1-17.

49. Jui-Pin Yang. Pushout with Differentiated Dropping Queue Management for High-Speed Networks / Jui-Pin Yang // Applied Mathematics & Information Sciences. – 2015. - №4. – P. 1961-1969.

50. Soumen Sarkar. Investigation on the Influence of Packet Sizes on WRR Queue Management in Ethernet switch / Soumen Sarkar, Nikolay Kakanakov, Shyam Sundar Prasad // Mobile Communication and Power Engineering. – 2013. – P. 50-55.

51. Faritha Banu J. Multipath virtual queue management system for effective packet scheduling in MPLS networks / J. Faritha Banu, V. Ramachandran // Indian Journal of Computer Science and Engineering. – 2012. - №4. – P. 588-594.

52. Fang Y. Buffer Management Algorithm Design and Implementation Based on Network Processors / Yechang Fang, Kang Yen, Deng Pan, Zhuo Sun // International Journal of Computer Science and Information Security.- 2010.– Vol.8, №1. – P. 1-8;

53. Sumalatha S. Multicast Node Communication Using Virtual Queue with Process Scheduling / S. Sumalatha, G. Sreenivasulu, G. Uday Kiran Bhargava // Journal of Computer Science. – 2015. - №11 (2) – P.337-343.

54. Jiao Zhenzhen Jiao. A virtual queue-based back-pressure scheduling algorithm for wireless sensor networks / Jiao Zhenzhen Jiao, Baoxian Zhang¹, Wei Gong, Hussein Mouftah // *EURASIP Journal on Wireless Communications and Networking*. – 2015. - P. 2-9;
55. Ramtin Pedarsani. Stability of multiclass queueing networks under longest-queue and longest-dominating-queue scheduling / Ramtin Pedarsani, Jean Walrand // *Applied Probability Trust*. – 2015. – P. 1-18;
56. Sonagara Payal T. Round Robin CPU Scheduling Using Dynamic Time Quantum with Multiple Queue / Sonagara Payal T. // *International Journal of Science and Research*. – 2015. – Vol.4 – P.997-999.
57. Annadurai C. Review of Packet Scheduling Algorithms in Mobile Ad Hoc Networks / C. Annadurai // *International Journal of Computer Applications*. – 2011. – Vol. 15, No. 1. – P. 7-10.
58. Alnowibet K.A. Nonstationary Erlang Loss Queues and Networks // PhD Thesis, North Carolina State University, 2004. – 140 p.
59. Wei-Ping Wang, David Tipper, Sujata Banerjee. A Simple Approximation for modeling Nonstationary Queues // *Proceedings of the Fifteenth Annual Joint Conference of the IEEE Computer Societies (INFOCOM '96)*. – 1996. – V.1. – P. 255–262.
60. Еременко А.С., Лебедеко Т.Н., Старкова Е.В. Анализ влияния состояния интерфейса на динамику загрузки очереди на маршрутизаторе телекоммуникационной сети // *Телекомунікаційні та інформаційні технології*. – 2015. – №3. – С. 45-51.
61. Segall A. The modeling of adaptive routing in data communications networks / A. Segall // *IEEE Trans. on communications*. – 1975. –Vol. 25, № 1. – P.85 – 95.
62. Лемешко А.В. Динамическая модель комплексного решения задач маршрутизации и абонентского доступа в территориально-распределенных телекоммуникационных сетях / А.В. Лемешко, А.Г. Беленков // *Открытые информационные и компьютерные интегрированные технологии*. – 2003. – Вып. 18. – С. 134 – 139.
63. Лемешко О.В., Євсєєва О.Ю., Симоненко Д.В. Модель динамічного балансування мережних ресурсів у телекомунікаційній мережі // *Системи обробки інформації*. – 2008. – Вип. 5(72). – С.71-74.
64. Гуревич И.М. Динамическая модель сети связи / И.М. Гуревич // *Теория телетрафика в системах информатики*. – М.: Наука, 1989. – С.54-64.

65. Васенин В.А. Математические модели управления трафиком в интернет: новые подходы на основе схем TCP/AQM / В.А. Васенин, Г.И. Симонова // Автоматика и телемеханика. – 2005. – № 8. – С. 94 – 107.
66. Польщиков К.О. Метод нейро-нечёткого активного управления пакетными чергами в мобильной радиомережи специального назначения // Збірник наукових праць ВІПІ НТУУ «КПІ». – 2010. – Вип. 1. – С. 52 – 59.
67. Польщиков К.А. Усовершенствованный метод нейро-нечеткого управления отбрасыванием пакетов в транзитных маршрутизаторах телекоммуникационной сети [Электронный ресурс] / К.А. Польщиков, Ю.Н. Здоренко // Проблемы телекоммуникаций. – 2014. – № 2 (14). – С. 76 - 90. – Режим доступа: http://pt.journal.kh.ua/2014/2/1/142_polshikov_fuzzy.pdf.
68. Marchese M. QoS Over Heterogeneous Networks // John Wiley & Sons, Ltd, 2007. – 328 p.
69. Остерлох Х. Маршрутизация в IP-сетях. Принципы, протоколы, настройка / Х. Остерлох– Киев: «ДиаСофт», 2002. – 512 с.
70. Лемешко А.В. Модель и метод предотвращения перегрузок с активным управлением длинной очереди на узлах телекоммуникационной сети / А.В. Лемешко, М.В. Семеняка // Проблемы телекоммуникаций. 2014. – Вип 9(14). – С. 15-25. [Электронный ресурс] – Режим доступа до журн.: http://pt.journal.kh.ua/2014/3/1/123_model_peregruzka.pdf.
71. Клейнрок Л. Теория массового обслуживания. – М.: Машин-е, 1979. – 432 с.
72. Петров М.Н. Вероятностно-временные характеристики в сетях и системах передачи интегральной информации: КГТУ, Красноярск, 1997. – 220 с.
73. Петров М.Н. Самоподобие в системах массового обслуживания с ограниченным буфером/ М.Н.Петров, Д.Ю.Пономарев // Электросвязь. - 2002. - № 2. – С. 35-39.
74. Li Y. Panwar S. Liu C.J. On the Performance of MPLS TE Queues for QoS Routing // Simulation series. – 2004. – Vol. 36; part 3. – P. 170–174
75. Костевич Л.С. Математическое программирование: Информ. технологии оптимальных решений / Л.С. Костевич. – Мн.: Новое знание, 2003. – 150 с.
76. Tawarmalani M. Global optimization of mixed-integer nonlinear programs: A theoretical and computational study / М. Tawarmalani, N.V. Sahinidis // Mathematical Programming. – 2004. – Vol. 99(3). – P. 563-591.
77. Tawarmalani M. Convexification and global optimization in continuous and mixed-integer nonlinear programming: theory, algorithms, software, and applications, nonconvex

optimization and its applications / M. Tawarmalani, N.V. Sahinidis. Series, 65. – Boston MA: Kluwer Academic Publishers, 2002. – 499 p.

78. Kocis G.R. Computation Experience with DICOPT: solving MILNP problems in process systems engineering / G.R. Kocis, I.E. Grossmann // Computers Chem. Eng. – 1989. – Vol. 13. – P. 307-315.

79. Бертсекас Д. Условная оптимизация и методы множителей Лагранжа. – М.: Радио и связь, 1987. – 400 с.

80. Сейдж Э.П. Оптимальное управление системами. Пер. с англ. под ред. Б.Р.Левина / Э.П. Сейдж, Ч.С. Уайт. – М.: Радио и связь, 1982. – 392 с.

81. Еремин И.И. О методе "штрафов" в выпуклом программировании // Кибернетика. – 1967. – № 4. – С.63-67.

82. Чен К. МАТЛАВ в математических исследованиях: Пер. с англ. / К. Чен, П. Джиблин, А. Ирвинг. – М.: Мир, 2001. – 346 с.

83. Ануфриев И.Е. МАТЛАВ 7 / И.Е. Ануфриев, А.Б. Смирнов, Е.Н. Смирнова – СПб.: БХВ-Петербург, 2005. – 1104 с.

84. Дьяконов В. П. МАТЛАВ 7.*/R2006/R2007: Самоучитель / В. П. Дьяконов – М.: ДМК Пресс, 2008. – 768 с.

85. Wetherall D. Active network vision and reality: lessons from a capsule-based system // 17th ACM Symposium on Operating Systems Principles (SOSP '99). Published as Operating Systems Review. – 1999. – 34(5). – PP. 64–79.

86. Tennenhouse D. L., Smith J. M., Sincoskie W. D., Wetherall D. J., Minden G. J. A Survey of Active Network Research // IEEE Communications Magazine. – 1997. – Vol. 35, №1. – P. 80-86.

87. Psounis K. Active networks: applications, security, safety, and architectures // Proc. IEEE Communications Surveys. – 1999. – Vol.1. – P. 2-16.

88. Raz D., Shavitt Y. An Active Network Approach to Efficient Network Management // Technical Report. DIMACS, 1999. – P. 99-25.

89. Иванов П. АКТИВНЫЕ СЕТИ // Сети. 1999. №10. С. 14-24. 8. Amit S., Raj J. Terabit switching: a survey of techniques and current products // Comput. Commun. – 2002. – Vol.25, № 6. – P. 547-556.

90. Ti-Shiang W. Architectural evolution and principles of optical terabit packet switches (OTPS) // Comput. Commun. 2002. Vol.25, № 6. P. 557-576.

91. Гринфилд Д. Терабитные маршрутизаторы: поучительная путаница // Журнал сетевых решений – LAN. 2000. №4. С.21-23. 11. Раскин Л.Г., Кириченко И.О. Многоиндексные задачи линейного программирования. М.: Радио и связь, 1982. 240 с.
92. Раскин Л.Г., Кириченко И.О. Многоиндексные задачи линейного программирования. М.: Радио и связь, 1982. 240 с.
93. Лемешко А.В. Особенности проведения лабораторного эксперимента по исследованию процессов управления очередями в телекоммуникационных сетях / А.В. Лемешко, Али С. Али, М.В. Семеняка // Цифрові технології. – 2011. – Вип. 10. – С.5-17.
94. Методические указания к лабораторным работам по дисциплине "Системы коммутации и распределения информации. Часть 2" для студентов специальности 7.092401 – Телекоммуникационные системы и сети / Сост.: М.В. Семеняка, Е.В. Старкова, А.В. Лемешко. – Харьков: ХНУРЭ, 2014. 92 с.
95. Евсеева О.Ю. Методика экспериментального исследования методов управления трафиком в сети [Электронный ресурс] // Проблемы телекомунікацій. – 2011. – № 3 (5). – С. 89 – 104. – Режим доступа до журн.: http://pt.journal.kh.ua/2011/3/1/113_yevseyeva_experiment.pdf.
96. Class-Based Weighted Fair Queueing [Электронный ресурс] / Cisco IOS Software Releases 12.0.
97. Cisco IOS Configuration Fundamentals and Network Management Configuration Guide, Release 12.3..
98. Руденко И. Tsunami Computing, Маршрутизаторы CISCO для IP-сетей / И. Руденко. – М.: Кудиц- Образ, 2003. – 656 с.
99. Avallone S. D-ITG Distributed Internet Traffic Generator / S. Avallone, S. Guadagno, D. Emma, A. Pescape, G. Ventre // Proceedings of the First International Conference on the Quantitative Evaluation of Systems (QEST'04) – 2004. – Vol. 4. – P 58-59.
100. Avallone S. Performance evaluation of an open distributed platform for realistic traffic generation / S. Avallone, S. Guadagno, D. Emma, A. Pescape, G. Ventre // Performance evaluation. International journal – 2005. – P 359-392.
101. Kolahi S. S. Performance Monitoring of Various Network Traffic Generators / S. S.Kolahi, S.Narayan, Du. D. T. Nguyen, Y. Sunarto // UKSIM '11 Proceedings of the 2011 UKSim 13th International Conference on Computer Modelling and Simulation – 2011. – P 501-506.
102. Олифер В.Г., Компьютерные сети. Принципы, технологии, протоколы / В.Г. Олифер, Н.А. Олифер: Учебник для вузов. 3-е изд. — СПб.: Питер, 2006. — 958 с.

АКТЫ ВНЕДРЕНИЯ

ПРИЛОЖЕНИЕ А

АКТ ВНЕДРЕНИЯ В УЧЕБНЫЙ ПРОЦЕСС ХНУРЭ

ЗАТВЕРДЖУЮ

Проректор з науково-педагогічної роботи
Харківського національного університету
радіоелектроніки



проф. Лесна Н.О.



« 21 » 8 2015 р.

АКТ

про впровадження результатів дисертаційної роботи Симоненка О.В. за темою «Потокові моделі та методи управління чергами на інтерфейсах маршрутизаторів транспортних телекомунікаційних мереж», представлену на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.12.02 – телекомунікаційні системи та мережі.

Комісія у складі:

голови – д.т.н., проф., зав. каф. ТКС Поповського В.В.;

членів – к.т.н., проф. каф. ТКС Пастушенко М.С.;

– к.т.н., доц. каф. Єременко О.С.

склала даний акт у тому, що результати дисертаційної роботи Симоненка О.В., а саме:

1. Потокова модель управління чергами на інтерфейсах маршрутизаторів транспортних телекомунікаційних мереж, що заснована на оптимальному агрегуванні та розподілі потоків на підставі порівняння та мінімізації розбіжності класів потоків і черг, в які вони направляються.
2. Потокова модель управління чергами на інтерфейсах маршрутизаторів активних мереж, що враховує не тільки кількість та параметри потоків і черг, але й характеристики (тип, завантаженість та продуктивність) активних процесорів, що входять до архітектури маршрутизатора активної мережі.

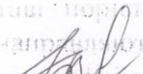
впроваджені у навчальному процесі кафедри телекомунікаційних систем ХНУРЕ в курсі лекцій та на практичних заняттях дисципліни «Системи розподілу інформації, частина 2».

Голова комісії



В.В. Поповський

Члени комісії



М.С. Пастушенко



О.С. Єременко

ПРИЛОЖЕНИЕ Б

АКТ ВНЕДРЕНИЯ В ООО «СИГМА СОФТВЕА»



№ 445
14 / 04 / 2015

«УТВЕРЖДАЮ»
Генеральный директор ООО «Сигма Софтвеа»
Вартанян Д.В.
« _____ » _____ 2015 г.



АКТ

об использовании ООО «Сигма Софтвеа» результатов диссертационной работы Симоненко Александра Викторовича на тему «Потоковые модели и методы управления очередями на интерфейсах маршрутизаторов транспортных телекоммуникационных сетей», представленную на соискание научной степени кандидата технических наук по специальности 05.12.02 – телекоммуникационные системы и сети

Комиссия в составе:

Председатель - начальник отдела	Туревский В.Л.
Члены комиссии - Инженер по тестированию ПО, к.ф.-м.н.	Уваров Р.А.
- Инженер по разработке ПО, к.т.н.	Головань К.В.

рассмотрела результаты диссертационной работы Симоненко А.В. и пришла к заключению, что материалы диссертации использованы при модернизации и развитии телекоммуникационной системы ООО «Сигма Софтвеа», а именно:

1. Математическая модель управления очередями на маршрутизаторах телекоммуникационной сети на основе оптимального агрегирования потоков и распределения пакетов по очередям с учетом их классов.
2. Инженерная методика оценки основных показателей качества обслуживания (средней задержки и вероятности потерь пакетов) при внедрении предложенной модели управления очередями на маршрутизаторах телекоммуникационной сети.

Председатель комиссии _____ Туревский В.Л.

Члены комиссии _____ Уваров Р.А.

_____ Головань К.В.

SIGMA SOFTWARE LLC
Postal address: 1, Ac. Proskura str., 61070, Kharkiv, Ukraine
Legal address: 47, Sobornaya str., k.12, 64300, Izium, Ukraine
ID 31935930
С/А 2600410462 in Public Joint Stock Company "First Ukrainian International Bank",
4 Andriivska St., 04070, Kyiv, Ukraine, bank code 334851
Phone/fax: +380 (57) 716 45 85, 766 01 59
E-mail info@sigma.software, Site: <http://ua.sigma.software/>

ТОВ «СИГМА СОФТВЕА»
Поштова адреса: вул. Ак. Проскури 1, 61070, м. Харків, Україна
Юридична адреса: вул. Соборна 47, к. 12, 64300, м. Ізюм, Україна
Код ЄДРПОУ 31935930
П/р 2600410462 в ПАТ «Перший Український Міжнародний Банк»
Вул. Андріївська, 4, 04070, м. Київ, Україна, МФО 334851
Тел/факс +380 (57) 716 45 85, 766 01 59
E-пошта: info@sigma.software, Сайт: <http://ua.sigma.software/>