

**МОСКОВСКИЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
СВЯЗИ И ИНФОРМАТИКИ**

На правах рукописи

Углов Иван Валерьевич

**РАЗРАБОТКА ОБОБЩЕННЫХ АНАЛИТИЧЕСКИХ
МОДЕЛЕЙ ПРОЦЕССОВ СИГНАЛЬНОГО ОБМЕНА
В КОНВЕРГЕНТНОЙ СЕТИ**

Специальность 05.12.13 – Системы, сети и устройства телекоммуникаций
(по техническим наукам)

Диссертация
на соискание ученой степени кандидата
технических наук

Научный руководитель
кандидат физико-математических наук,
доцент Ефимушкин Владимир Александрович

Москва – 2015

ОГЛАВЛЕНИЕ

ОГЛАВЛЕНИЕ	2
ВВЕДЕНИЕ	4
ГЛАВА 1. Исследование задач функционирования конвергентных сетей и ее фрагментов	12
1.1. Исследование технологий конвергентных сетей	12
1.2. Анализ функциональной модели конвергентной сети	16
1.3. Анализ актуальности задач аналитического и экспериментального исследования конвергентной сети и ее фрагментов	21
1.4. Выводы	30
ГЛАВА 2. Разработка моделей функционирования фрагментов конвергентной сети при обработке сигнальной информации	32
2.1. Модель функционирования фрагмента конвергентной сети на базе элементов MSC/VLR/HLR	33
2.2. Анализ модели функционирования фрагмента MSC/VLR/HLR	41
2.3. Анализ BBX СМО фрагмента MSC/VLR/HLR	52
2.4. Модель функционирования узла P-/I-CSC.	56
2.5. Анализ модели функционирования узла P-/I- CSC	59
2.6. Анализ BBX СМО P-/I-CSC	62
2.7. Выводы	66
ГЛАВА 3. Исследование и разработка модели сигнального обмена в конвергентной сети	67
3.1. Разработка процедуры построения математической модели предоставления услуг в конвергентной сети	67
3.2. Исследование процессов сигнально обмена в конвергентных сетях на базе СПС 2/3/4 поколений	74
3.3. Численные результаты моделирования и выводы	85
3.4. Решение СеМО для узлов сети типа GI G m	88
3.5. Выводы	95

ГЛАВА 4. Экспериментальное исследование сигнального обмена на магистрали конвергентной сети	96
4.1. Исследования проблем экспериментального изучения магистральных участков конвергентной сети	96
4.2. Исследование трафика и структуры потоков сигнальных SIP сообщений на магистральном канале между сетями агрегации и центром обработки вызов	107
4.3. Выводы	114
ЗАКЛЮЧЕНИЕ	116
СПИСОК ИСТОЧНИКОВ	118
ПРИЛОЖЕНИЕ 1. Список сокращений	133
ПРИЛОЖЕНИЕ 2. Список основных обозначений	139
ПРИЛОЖЕНИЕ 3. Анализ архитектур конвергентных сетей	141
ПРИЛОЖЕНИЕ 4. Соответствие процессов сигнального обмена переходам и классам заявок SeMO конвергентной сети	168
ПРИЛОЖЕНИЕ 5. Соответствие классов и типов сигнальных сообщений для процедуры исходящего вызова в сети LTE с использованием CSFB в сеть 2/3-го поколений без поддержки поддержки пакетного хендовера (PS HO)	173
ПРИЛОЖЕНИЕ 6. Документы, подтверждающие использование результатов диссертации	175

ВВЕДЕНИЕ

При проектировании новой и модернизации существующей конвергентной сети, либо ее фрагмента, включающего несколько логически связанных элементов, возникает задача оценки параметров функционирования конвергентной сети или ее фрагмента, что приводит к необходимости создания аналитических, либо имитационных моделей. Создание аналитических моделей предпочтительнее, т.к. они позволяют получить приемлемые оценки параметров конвергентной сети без задействования больших вычислительных ресурсов. Анализ вероятностно-временных характеристик (ВВХ) модели конвергентной сети дает возможность выявить и устранить ошибки на этапе проектирования, позволяя избежать избыточных затрат при строительстве. Таким образом, исследование ВВХ процессов сигнального обмена в конвергентной сети является **актуальной задачей**.

Исследование вопросов моделирования сетей связи нашло широкое отражение в работах отечественных и зарубежных ученых: Башарина Г.П., Вишневого В.М., Гольдштейна Б.С., Деарта В.Ю., Докучаева В.А., Ефимушкина В.А., Кучерявого А.Е., Назарова А.Н., Нетеса В.А., Печинкина А.В., Пшеничникова А.П., Ромашковой О.Н., Самуйлова К.Е., Семенова Ю.В., Соколова Н.А., Степанова С.Н., Сычева К.И., Шоргина С.Я., Шнепс-Шнеппе М.А., Яновского Г.Г., Baskett F., Chandy K.M., Cox C., Janevski T., Lagenbach-Belz M., Mayer G., Muntz R.R., Palacios F.G., Whitt W., и др.

Наиболее перспективной сетью, в которой возможно оказание широкого спектра услуг связи, в том числе традиционных, является сеть последующего поколения (NGN, Next Generation Network). Переход к NGN инициирует процессы конвергенции фиксированных и мобильных

сетей с различным уровнем взаимопроникновения, приводящие в результате к формированию конвергентной сети.

Исследованию особенностей построения и функционирования подобных сетей посвящены работы Гольдштейна Б.С., Гордиенко В.Н., Кучерявого А.Е., Кучерявого Е.А., Самуйлова К.Е., Соколова Н.А., Степанова С.Н., Шалагинова В.А., Boucherie R.J., Dijk N.M., Kramer W., Poikselka M., и др.

Процессы обмена сигнальной информацией и ее обработки в конвергентной сети являются достаточно сложными, что не позволяет использовать для описания функционирования конвергентной сети и ее фрагментов простые системы массового обслуживания (СМО), для этого требуется рассматривать СМО и сети массового обслуживания (СеМО) сложной структуры.

На этапе проектирования наиболее часто возникают задачи исследования процессов обмена сигнальной информацией и ее обработки во фрагментах КС, состоящих из нескольких элементов, которые следует рассматривать в совокупности, как единый функциональный элемент, что приводит к необходимости построения и анализа моделей фрагментов КС, являющихся, таким образом, обобщенными моделями.

Оценку качества работы конвергентной сети на этапе эксплуатации проводят на основе экспериментального исследования информационных потоков, что одновременно позволяет решать следующие актуальные задачи: проводить оценку качества услуг, использовать экспериментально полученные данные и результаты их обработки в качестве исходных при проектировании новых объектов, создании аналитических и имитационных моделей. Методы экспериментального анализа информационных потоков должны обеспечивать решение следующих актуальных задач: минимизации влияния средств экспериментального изучения конвергентной сети на ее работу, оценки характеристик работы сети и ее отдельных узлов, определения качества предоставляемых услуг, формирования достаточного объема

информации о структуре информационных потоков и режимах работы сетевых узлов.

Полученные в диссертации результаты могут быть использованы для решения перечисленных задач.

Объектом исследования является конвергентная сеть, обеспечивающая аудио- и видео- связь, ее фрагменты, потоки сигнальной информации в ней

К предмету исследования диссертации относятся характеристики функционирования конвергентной сети и ее фрагментов в процессе обмена сигнальной информацией.

Цель исследования заключается в разработке обобщенных аналитических моделей функционирования КС и ее фрагментов в процессе обмена сигнальной информацией при организации аудио и видео вызовов, экспериментальном анализе характеристик процессов сигнального обмена, позволяющих повысить эффективность функционирования КС.

Для достижения цели исследования при решении поставленных **задач** были получены следующие результаты:

- проведен анализ существующих архитектур и основных сценариев предоставления услуг, используемых технологий и протоколов в конвергентной сети;
- разработаны обобщенные аналитические модели фрагментов конвергентной сети, позволяющие определить характеристики работы коммутационного ядра;
- разработан механизм формализации процессов сигнального обмена в конвергентной сети при построении СеМО;
- на основе предложенного механизма формализации построена аналитическая модель процессов сигнального обмена в конвергентной сети при установлении голосовых вызовов с использованием возврата к технологии канальной коммутации (CSFB, Circuit Switched FallBack);

- исследованы экспериментальные способы изучения характеристик конвергентной сети;
- предложена методика и результаты экспериментального исследования характеристик потоков сигнальной информации на транзитных направлениях крупной конвергентной сети.

Научная новизна работы заключается в следующем:

1. Впервые предложены обобщенные аналитические модели фрагментов ядра конвергентной сети и проведено их исследование. Разработаны рекомендации по выбору параметров элементов конвергентной сети в зависимости от интенсивности нагрузки и ее типа, что позволяет при заданной величине потерь на элементах получить соотношения для производительности элементов моделируемых фрагментов и обеспечить их равномерную загрузку.

2. Предложен новый механизм формализации процессов сигнального обмена в конвергентной сети и построения с использованием данного механизма СеМО с входящими потоками общего вида и узлами в виде многопроцессорных систем с произвольными дисциплинами обслуживания. Модель позволяет оценить основные ВВХ конвергентной сети и исследовать процесс организации исходящего вызова с использованием технологии CSFB на сети. Предложены и исследованы типовые конфигурации конвергентной сети, проведен анализ применимости моделей узлов различных типов для описания элементов конвергентной сети. Это позволяет использовать разработанные в диссертации рекомендации в проектных расчетах для повышения эффективности конвергентной сети и ее фрагментов.

3. Экспериментально подтверждено наличие в конвергентной сети потоков сигнальных сообщений высокой интенсивности не пуассоновского типа. Определены характеристики подобного потока между сетью агрегации трафика и контактном центром (КЦ), основные сценарии и особенности установления вызовов, что может быть

использовано при выборе параметров входящего потока сигнальных сообщений в аналитических моделях.

Практическая значимость работы заключается в следующем:

Аналитические модели для оценки характеристик узлов сети и рекомендации по выбору ее элементов, алгоритм анализа потоков сигнальных сообщений и его программная реализация, результаты экспериментального анализа потоков сигнальных сообщений могут использоваться научно-исследовательскими и проектными организациями при проведении исследований и разработок в области конвергентных сетей, операторами связи при оценке характеристик функционирования создаваемых и модернизируемых фиксированных сетей и сетей сотовой подвижной связи.

Методы исследования. В диссертации используются методы теории сетей связи, теории вероятностей, теории сетей массового обслуживания, теории телетрафика, линейной алгебры и модельный эксперимент.

Достоверность и обоснованность полученных результатов обусловлена корректным применением математических и алгоритмических методов исследования, соответствием результатов аналитических моделей и численных экспериментов на реальных сетевых объектах.

Структура диссертационной работы. Диссертация состоит из введения, четырех глав, заключения, списка литературы из 139 наименований. Диссертация содержит 132 страницы, 38 рисунков и 6 таблиц и имеет шесть приложений.

В **первой главе** проведена систематизация и классификация существующих архитектурных и технологических решений по построению сетей связи. Классифицированы технологии передачи информации, протоколы и типовые архитектурные решения, используемые в конвергентных сетях. Определены классы потоков информации в конвергентных сетях и требования к ним. Обозначены

точки агрегации информации различных типов. Проанализированы основные сценарии оказания услуг в конвергентных сетях.

Во **второй главе** приведены обобщенные аналитические модели фрагментов конвергентной сети, для каждой модели приведены результаты численного анализа ВВХ. Сделаны выводы относительно целесообразности использования данных моделей модернизации сети с целью улучшения ее характеристик. Отмечены возможности использования порогового регулирования нагрузки для улучшения работы сети. Предложен алгоритм расчета характеристик моделей с учетом экономии вычислительных ресурсов, что расширяет области применения модели для систем с пространствами состояний, выходящими за пределы возможностей стандартного персонального компьютера.

В **третьей главе** предложена методика формализации задачи моделирования процессов обмена сигнальной информацией в конвергентной сети. На основе методики разработана аналитическая модель конвергентной сети подвижной связи, рассмотрен процесс исходящего вызова с использованием возврата к технологии канальной коммутации. Рассмотрены типовые конфигурации подобных сетей и проведена оценка применимости моделей узлов различных типов для описания элементов конвергентной сети. Проведен анализ основных ВВХ.

В **четвертой главе** проведена классификация, анализ и определение схем записи и последующего анализа сигнальной информации с реальных конвергентных сетей. Приведены результаты анализа характеристик потоков сигнальных сообщений между сетью агрегации трафика и центром обработки вызовов.

В заключение приведена общая характеристика работы, даны основные выводы и рекомендации по применению полученных результатов и предложенных методов.

Апробация работы. Основные положения и результаты диссертационной работы докладывались и обсуждались на следующих конференциях и семинарах:

– 2 – 4-й, 6-й и 7-й Отраслевых научных конференциях-форумах «Технологии информационного общества» в 2008-2010 гг., 2012-2014 гг., г. Москва;

– 4-й Ежегодной конференции «Конвергенция в телекоммуникациях» в 2010 г., г. Москва;

– 17-й Международной конференции DCCN-2013 «Распределенные компьютерные и коммуникационные сети: управление, вычисление, связь», в 2013 г., г. Москва;

– 9-м Всероссийском совещании по проблемам управления, 2014 г., г. Москва.

Публикации. По результатам исследований опубликованы семнадцать работ, в том числе девять работ в журналах, входящих в «Перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные научные результаты диссертации на соискание ученой степени доктора и кандидата наук».

Реализация результатов работы. Полученные в диссертации результаты исследований использованы в спецкурсе «Конструктивно-технологическое обеспечения средств связи», читаемом на кафедре компьютерной и информационной безопасности Московского государственного технического университета радиотехники, электроники и автоматики. Программное обеспечение, реализующее разработанные в диссертации методы анализа сигнальных сообщений, зарегистрировано со свидетельством о государственной регистрации программы для ЭВМ 2013661814 РФ: МПК-2014.01 Н 04 Q 1/10. Данное программное обеспечение использовалось в рамках работ по гранту РФФИ 13-07-00665 А, 2013-2014 гг. Разработанные в диссертации модели и методы были использованы для анализа и улучшения

характеристик сети оператора ОАО «Мобильные ТелеСистемы», что подтверждено соответствующими актами.

На защиту диссертационной работы выносятся:

1. Разработанная в диссертации аналитическая модель P-I-CSC позволяет установить критерий выбора параметра r'_1 , определяющего схему обработки заявок на узле, при котором возможно уменьшение вероятности потерь как приоритетных заявок (ПЗ), так и непериприоритетных заявок (НЗ).

2. Разработанная обобщенная аналитическая модель фрагмента ядра КС позволяет определить критерии выбора параметров его элементов и при заданном уровне потерь обеспечить их равномерную загрузку

3. Разработанный механизм формализации задачи моделирования процессов сигнального обмена в конвергентной сети с пошаговым алгоритмом позволяет осуществить переход от инженерного описания процессов сигнального обмена к СеМО.

4. Разработанная в диссертации модель конвергентной сети позволяет определить системы массового обслуживания (СМО), наиболее близко отражающие изменения временных характеристик обработки сигнальных сообщений. Для моделирования элементов коммутационного ядра предлагается использовать СМО $\bullet/M/m$, либо $GI/G/m$; для базовых станции сетей подвижной связи – СМО $\bullet/G/1$, для множества абонентских устройств и других сетей – СМО $\bullet/G/\infty$, здесь и далее « \bullet » обозначает один или несколько входящих потоков.

5. Полученные характеристики потоков сигнальных сообщений между сетью агрегации вызовов и КЦ подтверждают наличие в конвергентной сети потоков не пуассоновского типа высокой интенсивности и необходимость учета их при моделировании.

ГЛАВА 1

ИССЛЕДОВАНИЕ ЗАДАЧ ФУНКЦИОНИРОВАНИЯ КОНВЕРГЕНТНЫХ СЕТЕЙ И ЕЕ ФРАГМЕНТОВ

1.1. Исследование технологий конвергентных сетей

Конвергенция сетей связи различных типов – это процесс их гармоничной эволюции, направленной на создание единой инфраструктуры для поддержки широкого спектра услуг и приложений.

В соответствии с концепцией построения сети NGN [117, 118] конвергенция может происходить на трех уровнях:

- 1) приложений;
- 2) услуг (включая услуги управления сетью);
- 3) транспорта, включая функцию управления сетевыми ресурсами и уровень доступа.

Развитие технологий передачи информации при помощи сетей коммутации пакетов и их использование при оказании услуг связи привело к появлению конвергентных сетей. Такая сеть состоит из технологически неоднородных сегментов, обеспечивающих единые услуги вне зависимости от конкретной технологии. Одной из главных ее задач является обеспечение качества предоставляемых услуг, которое, в свою очередь, определяется на основе анализа ее параметров при передаче сигнальной информации и пользовательских данных. Передача сигнальной информации с заданным уровнем качества является наиболее приоритетной задачей, т.к. при ее невыполнении функциональные элементы конвергентной сети не могут нормально взаимодействовать. Из-за этого сеансы передачи пользовательских данных организуются с ошибками либо становятся невозможны.

На этапе проектирования новой или модернизации существующей конвергентной сети возникает задача оценки параметров ее функционирования, что приводит к необходимости создания аналитических либо имитационных моделей. Аналитические модели предпочтительнее, т.к. с их помощью удастся провести анализ различных параметров, зачастую без использования больших вычислительных ресурсов. Анализ ВВХ аналитических моделей сети, ее фрагментов и узлов дает возможность выявить ошибки на этапе проектирования и устранить их, тем самым позволяя избежать избыточных затрат при строительстве. Таким образом, исследование функционирования конвергентной сети и анализ процессов информационного обмена, в особенности – сигнального, с целью оценки качества является актуальной задачей.

Далее в разделе будут проанализированы основные способы обеспечения обмена сигнальной информацией в конвергентной сети и определены задачи для исследования. Краткая справка по технологиям и устройству конвергентных сетей дана в Приложении 4 к диссертации.

Анализ технологий и протоколов сигнализации в конвергентных сетях

Основные протоколы сигнализации, используемые в конвергентных сетях, рассмотрены в Приложении 4 к диссертации. В данном разделе остановимся на рассмотрении особенностей протокола установления сеанса (SIP, Session Initiation Protocol) [97] и его подвидов SIP-I/-T [85], поскольку согласно международным стандартам [87] данный протокол является целевым при организации управления соединениями в NGN сетях на базе решения SoftSwitch и мультимедийной платформы (IMS, IP Multimedia Subsystem).

Быстро развивающиеся услуги на базе протокола SIP делают вопрос обеспечения их надлежащего качества актуальным. Большинство услуг, для обеспечения которых используется данный протокол, – услуги, чувствительные к задержкам. Исследование механизмов их формирования при управлении сеансом на основе протокола SIP, а также

механизмов работы SIP-устройств, является на данный момент не завершенным вопросом.

Рассмотрение вопросов управления сеансами в конвергентных сетях с учетом их сложной архитектуры, и обеспечение надлежащего качества услуг представляется нетривиальной задачей. Вопросам исследования работы SIP-устройств и формирования задержки при установлении SIP-сеансов в конвергентных сетях посвящен ряд работ отечественных [45, 62-65] и зарубежных авторов [125, 131, 133].

Состав SIP-услуг постоянно расширяется, изменяются и характеристики потоков SIP-сообщений. Для адекватного моделирования сетевых SIP-устройств необходимо обладать информацией о характеристиках и структуре потоков SIP-сообщений в исследуемых конвергентных сетях, которые могут отличаться от общепринятых норм для ССОП, строившихся на базе технологий КК [2, 65-67].

В процессе исследования SIP-сетей необходимо принимать в расчет следующие моменты [10]:

- эмпирические исследования реальных SIP-сетей и устройств, зачастую сопряжены с большим количеством проблем. К ним можно отнести сложность организации измерений на реальных сетях, обеспечение адекватности данных без их искажения и влияния на качество предоставляемых услуг, организация обработки данных, обеспечение условий конфиденциальности полученной информации;

- исследование современных SIP-сетей и устройств показывает, что аналитические модели, являясь удобным инструментом, не всегда позволяют точно описать исследуемый объект. Объекты реальных сетей зачастую настолько сложны, что рассчитать их вероятностно-временные характеристики (ВВХ) с использованием современных программно-аппаратных комплексов (ПАК) не является возможным. В результате, для инженерных расчетов часто применяют приближенные методы расчетов основных характеристик исследуемого объекта [25].

Согласно архитектурной модели конвергентной сети (подробно рассмотрена в Приложении 3, п. 1.3 [52]), ее можно логически разделить на ядро сети и сети доступа. Отметим, что технологии и протоколы, используемые в телекоммуникационных сетях, и описанные в [13, 38] не являются предметом исследования настоящей диссертации. Кратко остановимся на структуре транспортной сети.

Ядро конвергентной сети обычно представляет собой совокупность магистральных каналов передачи данных и сетевого оборудования, обеспечивающих передачу информации и управляющего работой конвергентной сети в целом. Ядро конвергентной сети использует стек протоколов TCP/IP и обширный диапазон специализированных протоколов уровней 4-7 модели OSI для транспортировки потоков трафика различных типов поверх IP-сети.

Сети доступа организуются на основе сетей КК и КП. Разнообразие технологий, использующихся в сетях доступа, обусловлено процессом естественной эволюции сетей связи. Таким образом, в сетях доступа используется все основные транспортные протоколы и технологии сетей КК и КП.

Протокол SIP, как упоминалось выше, является одним из основных и наиболее перспективных протоколов управления сеансами передачи пользовательской информации. В связи с этим вопросы исследования механизмов обслуживания SIP-сообщений в конвергентных сетях и, на SIP-устройствах, в частности, являются неразрывно связанными с вопросами обеспечения и контроля QoS в конвергентных сетях [14, 18].

Кроме протокола SIP, актуальными являются вопросы, связанные с передачей пользовательских данных и сигнальной информации с использованием других протоколов. Одним из новых актуальных направлений исследований обеспечения QoS для услуг в конвергентных СПС является обеспечение их в сетях LTE [41, 134].

В сети NGN необходимо обеспечивать контроль и соответствие требованиям к качеству передачи информации на всех ее логических

уровнях. Согласно модели NGN [115, 116] различают четыре вида обеспечения QoS [118], одновременный учет которых позволяет судить о качестве предоставления услуг на сети:

- обеспечение качества передачи данных между интерфейсами;
- обеспечение качества контролирующих механизмов;
- обеспечение качества контроля функциональной архитектуры;
- обеспечение качества управления и сигнализации.

Для обеспечения требуемого качества обслуживания на уровнях L2-L4 в сети обычно используются следующие механизмы:

- работа в недогруженном режиме – все коммутаторы и каналы передачи данных работают на 20-30%;
- введение классов обслуживания в нагруженной сети – общая производительность каждого из ресурсов сети делится между разными классами трафика не равномерно.

1.2. Анализ принципов взаимодействия конвергентных сетей и маршрутов прохождения сигнальной и голосовой информации

Взаимодействие сетей КК и сетей КП сводится к организации точек обмена сигнальной и пользовательской информацией с помощью сигнального шлюза (SG, Signaling Gateway) и шлюза данных (MG, Media Gateway), находящегося под управлением контроллера шлюзов (MGC, Media Gateway Controller), рис 1.1.

Взаимодействие сетей 2-го и 3-го поколений в настоящий момент свелось к объединению функционалов узлов радио- и коммутационной подсистемы на единых аппаратно-программных платформах. Так, например, базовые станции (БС) 2-го и 3-го поколений исполняются в виде комбинированных устройств, позволяющих развертывать эти сети одновременно, рис 1.1.

Рассмотрим более подробно актуальный процесс перехода к сетям 4-го поколения (на основе технологий LTE и LTE-Advanced, далее – сеть LTE [8]). Процесс перехода предполагает два последовательных этапа.

Первый этап сводится к процессу построения радиоподсистемы и части элементов коммутационной подсистемы с использованием технологий LTE. На этом этапе возможно предоставление пользователям только услуг пакетной передачи данных.

На этом этапе основные функции по обеспечению мобильности абонента, контроля местоположения абонента организуются оборудованием сети LTE. Для обеспечения возможности совершения голосовых вызовов в сети LTE при этом применяется технология CSFB [84]. Одной из задач первого этапа является приведение сети оператора в состояние готовности к пропуску больших объемов трафика на магистралях, подготовка квалифицированного персонала и приобретение необходимого опыта обслуживания сети LTE в режиме предоставления пакетных услуг, которые менее требовательны к обеспечению качества сигнала, чем услуги голосовых и видео вызовов. Таким образом, предоставление полного спектра услуг при использовании только лишь инфраструктуры сети LTE на первом этапе внедрения невозможно.

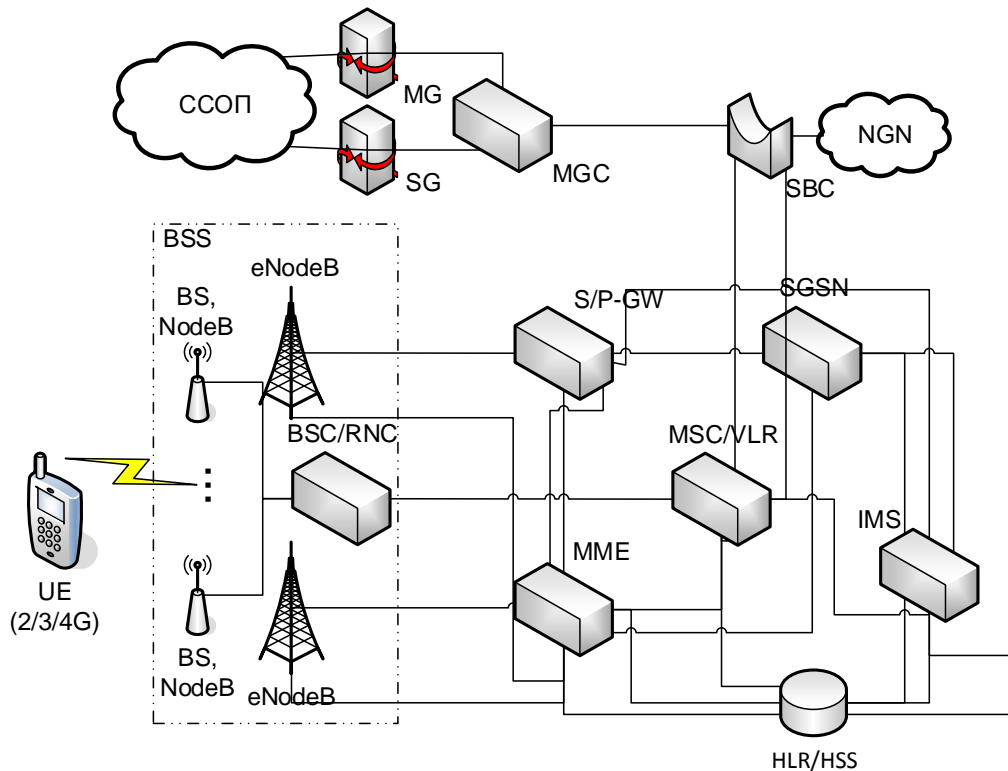


Рис.1.1. Фрагмент конвергентной сети с использованием платформы IMS

Второй этап внедрения технологии LTE проводится, когда радиопокрытие LTE достигает необходимой плотности и охвата, сеть оператора полностью готова к пропуску требуемых объемов трафика с заданным качеством. Необходимым условием перехода ко второму этапу внедрения технологии LTE является также установка центрального коммутационного узла, управляющего вызовами в реальном времени в сети LTE – платформы IMS. После выполнения мероприятий, обозначенных выше, возможна организация услуг передачи голосовых и видео вызовов поверх сети LTE (VoLTE, Voice over LTE) [138].

Основываясь на рекомендациях МСЭ-Т [117, 118, 127] можно определить пути потоков сигнальной и пользовательской информации в конвергентной сети на примере организации узла зоновой связи [37, 42, 51, 56, 98, 103]. В соответствии с принципами построения конвергентной сети в ней можно выделить два класса трафика – сигнальный и пользовательский, который проходит по сетям КК (рис 1.2).

Каналы связи такой сети должны по своим техническим и качественным характеристикам соответствовать минимальным требованиям к пропускной способности, гарантирующим предоставление абонентам услуг с заданным качеством.

Анализируя архитектурные особенности конвергентных сетей, определим узлы агрегации сигнальной и пользовательской информации. Очевидно, что точки агрегации потоков данных, являются одновременно и наиболее вероятными точками отказа сети либо ее сегмента. В зависимости от обширности области, с которой через точку выводятся потоки данных на верхние логические уровни, можно определить влияние корректной работы точки агрегации на работу сети.

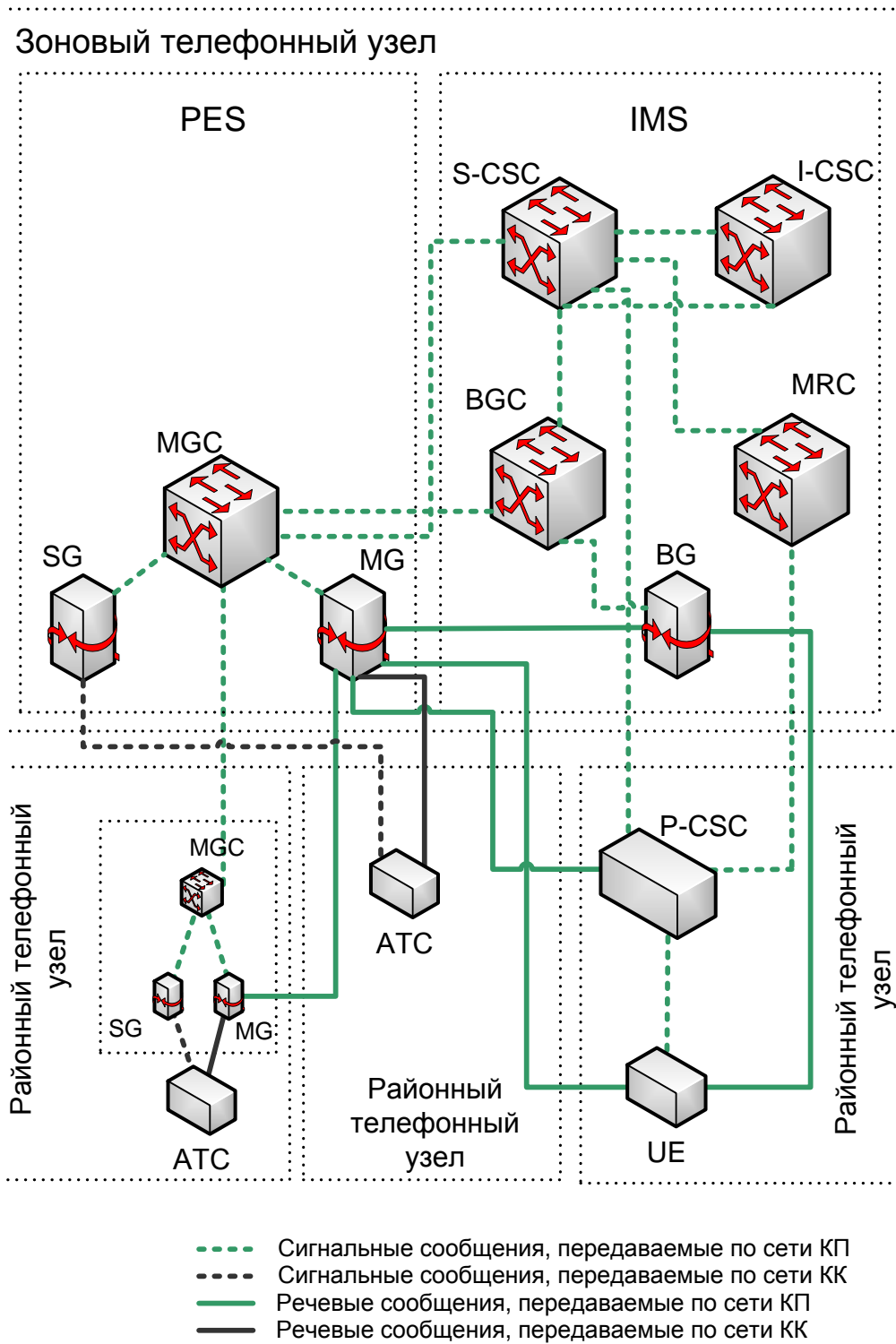


Рис.1.2 Схема путей потоков сигнальной и пользовательской информации в конвергентной сети

В коммутационном ядре сети оператора, основными точками отказа на региональном уровне будут являться:

- контроллеры базовых станций 2/3/4-го поколения СПС;

– магистральные узлы транспортных сетей регионального уровня;

На уровне субъектов федерации точками концентрации трафика являются:

– узлы MSC 2/3/4-го поколения СПС;

– узлы P-I/S-CSC, установленные в рамках одного субъекта федерации;

– магистральные узлы транспортных сетей, агрегирующие потоки на уровне субъекта федерации;

На уровне федеральных округов точками концентрации трафика будут являться:

– транзитные коммутаторы SoftSwitch либо P-I/S-CSC большой мощности;

– магистральные узлы транспортных сетей большой мощности;

На международном уровне - точки концентрации национального трафика:

– транзитные контроллеры MGC либо I-CSC большой мощности;

– магистральные узлы транспортных сетей большой мощности.

В реальной сети крупного оператора связи оборудование, выполняющее концентрацию трафика (особенно на верхних уровнях) подлежит обязательному резервированию. Кроме того, обычно, создаются обходные маршруты на случай аварии.

Исходя из сказанного, особый интерес представляет изучение механизмов работы оборудования, выполняющих функции концентрации трафика. Отдельно можно выделить оборудование I-P/S-CSC, MSC, SoftSwitch, поскольку отказ данного оборудования обычно ведет к перегрузкам обходных направлений, потере больших объемов коммерческого трафика, финансовым и имиджевым убыткам оператора.

1.3. Анализ актуальности задач аналитического и экспериментального исследования конвергентной сети и ее фрагментов

В диссертации, целью которой является исследование конвергентной сети и, в особенности, ее коммутационного ядра и потоков сигнальных сообщений при организации сеансов обмена информацией в режиме реального времени, решены задачи аналитического моделирования конвергентной сети, ее фрагментов, предложены результаты исследования потоков сигнальной информации для обоснования актуальности предложенных моделей. Далее в главе ставятся задачи, решению которых посвящена диссертационная работа.

Задача исследования характеристик функционирования фрагмента коммутационного ядра конвергентной сети

Необходимость анализа маршрутов обмена сигнальной информацией в точках ее агрегации является актуальной задачей. Аналитические модели позволяют оценить основные характеристики работы элементов сети в их взаимодействии и могут использоваться при проектировании ее новых участков, либо модернизации существующих. Моделированию отдельных коммутационных элементов сети в настоящее время посвящено значительное число работ (см. [1, 69, 71, 72, 131] и библиографию в них). Публикации, посвященные аналитическому моделированию коммутационного ядра конвергентной сети, учитывающему взаимодействие его узлов, по мнению автора, отсутствуют, судя по анализу открытых научных источников. Рассмотрение коммутационного ядра в качестве комплексного элемента и создание его обобщенной модели может быть целесообразным в следующих случаях:

– коммутационное ядро обслуживает соединения различного типа, например, транзитные вызовы и вызовы от абонентов;

– если при исследовании потоков сигнальных сообщений ядро требуется рассматривать в качестве единого функционального элемента.

Анализ моделей сетевых фрагментов чаще всего осложнен большим количеством состояний, в которых они могут находиться, что приводит к необходимости рассмотрения многомерных Марковских процессов (МП).

Рассмотрим коммутационное ядро конвергентной сети, построенное на основе технологии SoftSwitch, и обеспечивающее конвергенцию фиксированного и мобильного сегментов и организацию сеансов передачи голосового и других видов трафика реального времени. Такое ядро будет состоять из следующих сетевых элементов, рис.1.3:

- MSC на базе MGC, выполняющего функции обслуживания мобильных абонентов и коммутации транзитных вызовов, поступающих от фиксированных абонентов и присоединенных операторов;

- VLR;

- HLR либо HSS.

Рассмотрим механизм функционирования узла MSC сети СПС конвергентного оператора. Отметим, что, несмотря на указания в международных стандартах [82] о функциональном разделении элементов VLR и MSC, на практике производители реализуют эти два функциональных элемента в виде единого ПАК. Данный подход обоснован многими факторами, из которых можно выделить следующие:

- необходимость обмена большими объемами сигнальной информации между VLR и MSC с малым временем задержки;

- невозможность функционирования сети СПС без любого из этих двух элементов.

Отметим, что для целостного описания процессов коммутации на MSC, необходимо также учитывать запросы информации к HLR/HSS [82].

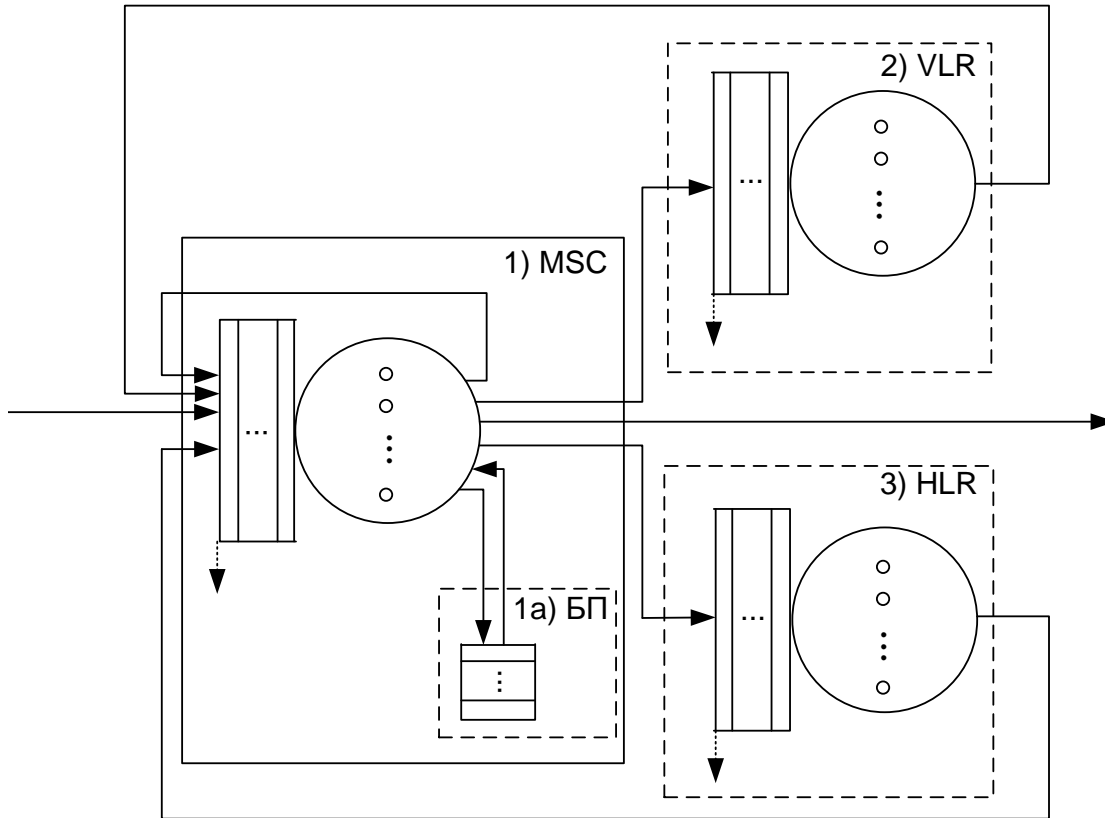


Рис.1.3. Логическая схема работы фрагмента MSC/VLR/HLR

В виду вышеизложенного целесообразно рассматривать MSC, VLR и HLR как фрагмент конвергентной сети. Рассмотрим подробнее логическую схему фрагмента, рис.1.3. Введем нумерацию элементов: MSC – первый элемент; VLR – второй; HLR – третий. Маршруты прохождения пакетов отметим сплошными линиями, возможность потери пакета – пунктирной линией. Каждый из элементов обладает буферным накопителем (БН), элемент 1 обладает также буферной памятью (БП), отмеченной номером 1а.

Рассмотрим два сценария коммутации на MSC фрагмента, сформированных на базе сценариев установления вызовов в СПС сетях [12, 13, 80, 81, 83, 85], а также варианты маршрутов прохождения пакетов и логику их обработки.

1) На вход MSC поступает сигнальное сообщение о начале процедуры установления соединения, например, ISUP IAM или SIP

INVITE. После ожидания в очереди MSC, сообщение поступает на обработку. Информация о поступившем сообщении заносится в БП (рис.1.3), после чего формируется запрос на VLR о местонахождении вызываемого абонента. Поступив в VLR, запрос попадает в очередь и далее на обработку. В результате обработки запроса на VLR возможны следующие (основные) сценарии:

– если в VLR информация о вызываемом абоненте найдена, то к MSC уходит сообщение с указанием идентификатора зоны (LAI, Location Area Identity) [135], в которой находится вызываемый абонент. После получения такого сообщения от VLR MSC инициирует запрос к БП, где хранится информация об исходном сообщении. На основе информации, полученной от VLR и сохраненной в БП, формируется исходящее сообщение об инициации соединения на радиоподсистеме в зоне с идентификатором LAI. Данное сообщение покидает узел MSC/VLR/HLR. Данные, связанные с обработкой сообщения из БП, удаляются, а данные устанавливаемой сессии сохраняются в БП.

– если VLR возвратил информацию о вызываемом абоненте с пометкой об отключении от сети – международный мобильный пользовательский идентификатор отключен (International Mobile Subscriber Identity (IMSI) detach) [135], то MSC немедленно формирует сообщение в сторону вызывающего абонента о завершении соединения. Данное сообщение покидает узел MSC/VLR/HLR. Из БП удаляются данные, связанные с обработкой сообщения. Данные устанавливаемого сеанса сохраняются в БП.

– если в VLR информация о вызываемом абоненте не найдена, то MSC формирует запрос к HLR о местоположении вызываемого абонента. Информация о посланном запросе в сторону HLR сохраняется в кэш-памяти MSC. Получив ответ от HLR с указанием идентификатора MSC, в котором зарегистрирован вызываемый абонент, MSC извлекает из БП информацию о начальном сообщении и сообщении, отосланном в сторону HLR. На основе данной информации формирует сообщение на

указанный идентификатор MSC. Данное сообщение покидает узел MSC/VLR/HLR. Из БП удаляются данные, связанные с обработкой сообщения. Данные устанавливаемого сеанса сохраняются в БП.

2) На вход MSC поступает сигнальное сообщение в рамках инициированного ранее сеанса. В этом случае на основе уникального идентификатора сеанса на MSC делается запрос в БП, из нее извлекается информация о сеансе, в рамках которого пришло очередное сообщение. На основе информации о сессии сообщение второго типа немедленно пересылается через внешние интерфейсы MSC.

Представленная выше задача решается в разделах 2.1 – 2.3 следующей главы.

Задача аналитического моделирования коммутационных элементов, устанавливаемых в точках взаимодействия конвергентных сетей или их сегментов

При организации взаимодействия сетей различных операторов, либо сегментов сети одного оператора часто возникает вопрос о приоритетности и легитимности трафика, поступающего на элемент из различных сетевых сегментов. Трафик, поступающий со стороны коммутационного ядра сети оператора считается более надежным и легитимным (доверенным), чем трафик, генерируемый подключенными абонентами и присоединенными операторами. Со стороны абонентских сетей и сетей присоединенных операторов вероятность организации различного рода попыток нарушения информационной безопасности выше.

Данная проблема особенно актуальна для элементов, осуществляющих взаимодействие двух и более сегментов сети, организованной на основе технологии КП, т.к. атаки на оборудование коммутации пакетов могут привести к потере всего проходящего через него трафика, либо существенному снижению качества обслуживания.

Исходя из проведенного выше анализа маршрутов передачи информации в конвергентной сети и ее архитектуры, можно выделить перечень элементов, соединяющих сегменты сети – это SoftSwitch, в составе MGC, MG, SG, контроллеры P-/I-CSC. В связи с тем, что оборудование сетей КП в большей степени подвержено различного рода атакам отказа в обслуживании, а также в связи с тем, что технологии КП постепенно замещают технологии КК, можно сделать вывод, что исследование механизмов обработки сигнальной информации на оборудовании P-/I-CSC является актуальной задачей.

Рассмотрим механизмы функционирования узлов P-/I-CSC – основных составляющих платформы IMS – нашедших распространение в конвергентных сетях еще до внедрения самой платформы. Рассмотрение модели функционирования узлов S-CSC оставим за рамками исследования, в виду отсутствия узлов данного типа на реальных сетях операторов России. Узлы P-CSC повсеместно используются для обеспечения единой точки регистрации абонентских устройств, узлы I-CSC – для организации точек обмена трафиком между отдельными сегментами сети оператора или сетями различных операторов [87]. В большинстве случаев на практике I-CSC и P-CSC реализуются на базе ПАК. Также, с учетом практических условий применения, в функциональность элементов P-/I-CSC производители зачастую добавляют механизмы защиты от атак уровней L3-L7.

Рассмотрим логику работы P-/I-CSC. На данный момент в сетях конвергентных операторов РФ устройства I-CSC в подавляющем большинстве используются в качестве IBC [87], поэтому далее будем рассматривать аспекты работы I-CSC для организации точек присоединения SIP между различными зонами одной сети одного оператора, либо между сетями различных операторов (рис. 1.4).

При рассмотрении устройств P-CSC (рис.1.5) будем учитывать весь описанный в стандартах функционал [87]. Основной задачей устройств P-/I-CSC является анализ сообщений сигнального трафика, и их

маршрутизация между соединяемыми сетями или их сегментами в соответствии с правилами маршрутизации на основе информации заголовков SIP-сообщений.

Устройства I/P-CSC устанавливаются в пограничных точках соединения различных сетей или их областей (рис.1.4, 1.5), что предполагает наличие двух типов поступающего трафика: «надежного» – от проверенных источников, и «ненадежного» – от неизвестных источников, которые могут являться источниками атак [125].

Проверенными источниками трафика являются внутренние сети операторов, непроверенными – сеть доступа и сеть другого оператора (рис. 1.4, 1.5).

Анализ процесса обработки информации в пакетных сетях на L1-L4 логических уровнях подробно представлен в отечественных и зарубежных работах [32, 75, 126, 130, 132, 136], сосредоточимся на исследовании механизмов работы P-/I-CSC на уровне L7 при анализе SIP-трафика.

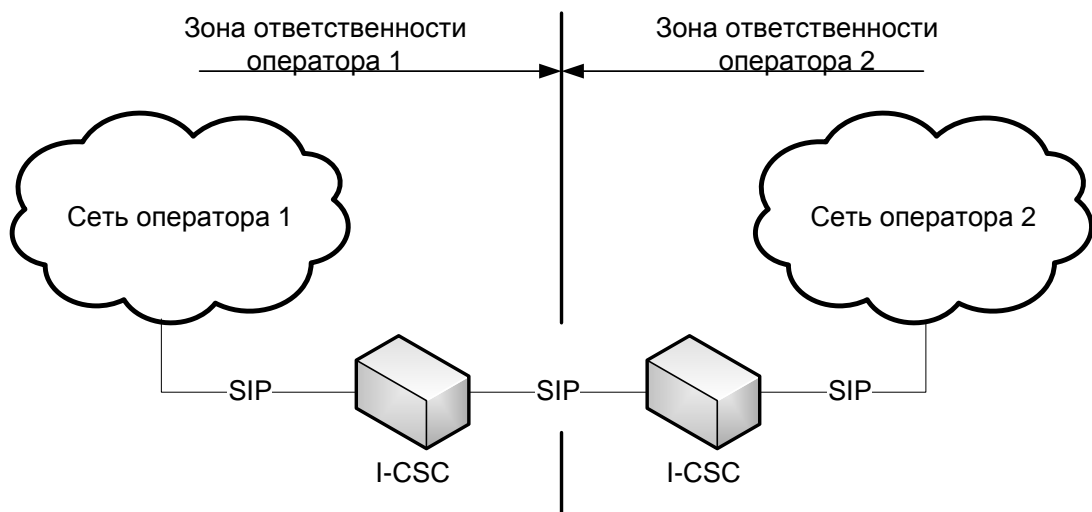


Рис.1.4. Схема организации обмена трафиком между сетями операторов с использованием I-CSC

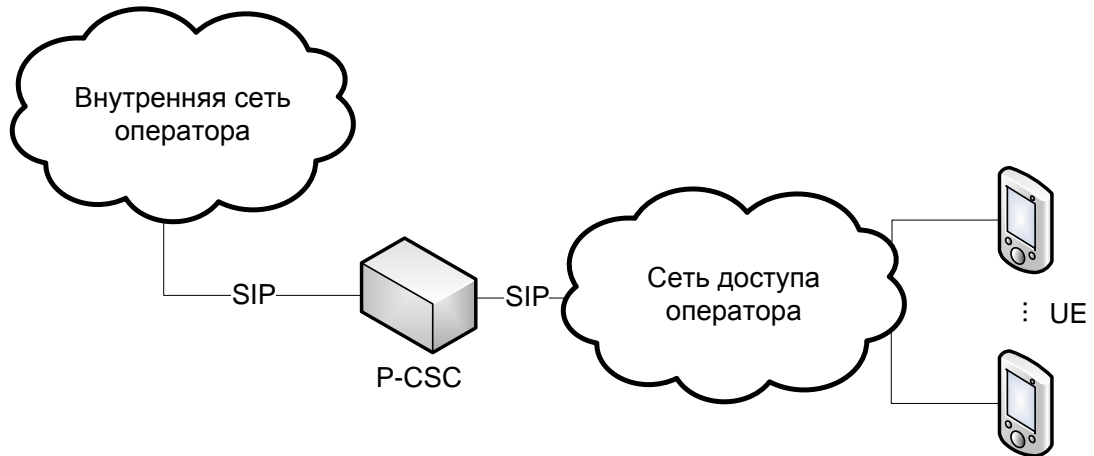


Рис.1.5. Схема организации сети с использованием P-CSC

Механизмы обработки SIP-трафика на узлах P-/I-CSC рассматриваются в работах [1, 26, 40, 63, 123, 131], где также обсуждаются механизмы регулирования нагрузки. Известные по литературе модели, однако, не учитывают наличие порогового механизма управления режимами работы узла, присутствующего в реальных устройствах P-/I-CSC. Модель, предложенная далее в диссертации, отличается от рассматриваемых ранее и учитывает данный механизм. Также, предложенная модель учитывает функциональность двух функциональных элементов конвергентной сети P-CSC и I-CSC одновременно и в этом смысле является обобщенной.

Данная задача дополняет первую задачу в части анализа характеристик пограничных элементов, а в совокупности решения первых двух задач предоставляют возможности комплексного исследования характеристик функционирования основного функционального фрагмента ядра. Решению второй задачи посвящены разделы 2.4 – 2.6 диссертации.

Задача моделирования конвергентной сети и процессов установления соединений в ней

Аналитические модели фрагментов конвергентной сети, предложенные в главе 2 диссертационной работы, как и модели узлов,

предложенные другими исследователями, не дают понимания характеристик работы конвергентной сети как единого объекта. Для проектировщиков конвергентной сети, и, в итоге, для потребителей услуг, являются важными такие характеристики как, например, время установления соединения, вероятность потерь вызовов в сети, что делает актуальным создание средств оценки работы сети в виде аналитических моделей. Для целей моделирования процессов, возникающих в конвергентной сети, необходимо определить:

- алгоритм перехода от инженерных диаграмм сигнального обмена в сети к описанию ее и рассматриваемых процессов передачи информации в ней в терминах теории массового обслуживания (ТМО);
- типовые конфигурации конвергентной сети;
- рекомендуемые СМО для описания работы типовых узлов сети;
- расчетные формулы для оценки характеристик установления соединения.

Поскольку в настоящее время существует выраженная тенденция к переходу абонентов из сегмента фиксированной телефонной сети в сети подвижной связи, то, с учетом целевой технологии для построения сетей 4-го поколения СПС [41], представляется целесообразным изучение конвергентной сети на основе сегментов СПС 2, 3 и 4-го поколений и фиксированного сегмента. Учитывая, что в настоящее время развитие сетей LTE на территории Российской Федерации находится на первом этапе, целевой технологией организации вызовов для абонентов сетей LTE является CSFB. Определение средних временных характеристик проектируемых конвергентных сетей необходимо, т.к. при сложных сценариях установления вызова с использованием CSFB время установления соединения может достигать до 15-20 с, что является серьезным фактором, влияющим на качество предоставляемых услуг.

Для расчета средних временных характеристик сети можно использовать СеМО типа ВСМР [70], а также СеМО с узлами общего

вида $GI/G/m$. Необходимость использования СМО общего вида для моделирования узлов оправдана при наличии непуассоновских потоков.

Решению третьей задачи посвящена глава 3 диссертации.

Задача экспериментального исследования конвергентных сетей на примере вызовов в контактный центр

В общем случае при большом количестве абонентов и стандартном типе вызова можно предполагать пуассоновский характер потоков, однако в некоторых случаях данное предположение не будет корректным [2, 65, 66, 126, 133], что приводит к необходимости обобщения модели конвергентной сети в виде СеМО с узлами $GI/G/m$, исследуемой в главе 3. В главе 4 экспериментально показано отсутствие свойства пуассоновости потоков сигнальных сообщений в конвергентной сети для некоторых услуг. К ним можно отнести динамично развивающиеся услуги контактных центров и виртуальных УПАТС, организуемых на сети оператора для корпоративных клиентов. Общий объем трафика от таких клиентов составляет на некоторых транзитных направлениях почти 100%, при этом потоки сообщений на таких направлениях не является пуассоновским. Для демонстрации этого была поставлена задача по анализу характера потока сообщений между сетью агрегации вызовов и оборудованием контактного центра на сети конвергентного оператора при оказании услуг.

1.4. Выводы

1. Проанализированы технологии передачи информации в конвергентных сетях, рассмотрены механизмы обеспечения и контроля качества передачи информации и маршруты ее передачи, включая точки концентрации, что позволило определить фрагменты конвергентной сети для исследования.

2. Предложено описание логики работы ядра конвергентной сети на базе элементов MSC/VLR/HLR и показана необходимость их совместного рассмотрения в качестве отдельного фрагмента и

разработки и исследования обобщенной аналитической модели данного фрагмента.

3. Предложено описание логики работы P-I-CSC, обоснована необходимость учета двух типов трафика – от доверенных источников и от потенциально опасных. Определена необходимость и поставлена задача разработки и исследования модели P-I-CSC, как дополняющей модель фрагмента MSC/VLR/HLR конвергентной сети.

4. Показана актуальность разработки и поставлена задача исследования модели конвергентной сети и определения ее средних временных характеристик при обработке вызовов от абонентов сети LTE с использованием технологии CSFB. Обоснована необходимость рассмотрения модели конвергентной сети, как для пуассоновского вида входящих потоков, так и для потоков общего вида.

5. Поставлена задача экспериментального исследования конвергентных сетей, обосновывающего необходимость использования сетей массового обслуживания с узлами общего вида для моделирования.

ГЛАВА 2

РАЗРАБОТКА МОДЕЛЕЙ ФУНКЦИОНИРОВАНИЯ ФРАГМЕНТОВ КОНВЕРГЕНТНОЙ СЕТИ ПРИ ОБРАБОТКЕ СИГНАЛЬНОЙ ИНФОРМАЦИИ

В главе рассматриваются вопросы моделирования узлов конвергентной сети. Вопросы моделирования элементов телекоммуникационных сетей исследовались в работах отечественных и зарубежных авторов на всем пути развития теории массового обслуживания (ТМО) [5-9, 25, 27, 29, 30, 32, 33, 35, 59, 70].

Согласно главе 1 диссертации, на основе анализа точек концентрации трафика и уровней агрегации можно сделать вывод о том, что для коммутации и управления соединениями в конвергентных сетях наиболее часто используются программные коммутаторы MGC, которые могут работать совместно со шлюзовым оборудованием MG и SG. На основе архитектуры программного коммутатора строится большинство узлов связи сети конвергентного оператора и более сложные функциональные элементы сети конвергентного оператора СПС как MSC/VLR [82].

Широкое распространение получили технологии организации связи с использованием протокола SIP. Точки концентрации сигнальной информации на различных уровнях агрегации могут создаваться на базе прокси серверов SIP с функционалом SBC. На базе прокси-серверов SIP различной мощности могут быть организованы элементы конвергентной сети оператора, такие как P-/S-/I-CSC, MGC. От режимов и характеристик работы данных устройств напрямую зависит качество предоставления услуг в конвергентной сети. Таким образом, изучение характеристик функционирования прокси-серверов SIP, контроллеров

SBC, программных коммутаторов, а также MSC/VLR представляется актуальной задачей. В виду широкого распространения и активного развития услуг с использованием протокола SIP, остановимся на изучении моделей функционирования узлов конвергентной сети, построенных на базе прокси-серверов SIP с функционалом SBC, и узлов MSC, построенных на базе программных коммутаторов. Рассмотрены случаи для фрагментов конвергентной сети, состоящих из узлов MSC, VLR, SBC, P-/S-/I-CSC [55].

2.1. Модель функционирования фрагмента конвергентной сети на базе элементов MSC/VLR/HLR

Определив логику работы фрагмента MSC/VLR/HLR в первой главе, перейдем к постановке задачи моделирования в терминах ТМО [30].

На практике для увеличения пропускной способности MSC, VLR, HLR используют масштабируемые решения с различным количеством идентичных обслуживающих блоков (ОБ), рис. 2.1, а, б. Здесь и далее количество ОБ на элементе (далее для краткости – элемент) будем обозначать m_i , где i – номер элемента рис. 1.5. Поступающие пакеты будем рассматривать в качестве заявок [7], заявки могут относиться к тому или иному классу.

Поступающие на элемент заявки равномерно распределяются по имеющимся на элементе ОБ, интенсивность поступающего на элемент потока обозначим λ_i , где i – номер элемента. Исходя из того, что на практике количество источников нагрузки велико (много больше m_i), предположим характер входящего потока заявок пуассоновским [6]. Интенсивность потока заявок на ОБ i -го элемента обозначим λ'_i , которая, в силу равномерного распределения входящего потока заявок по все ОБ элемента есть $\lambda'_i = \lambda_i / m_i$.

Каждый ОБ представляет собой однолинейную СМО с БН конечной емкости и, в случае первого элемента – БП для хранения обработанных

данных. Обозначим интенсивность обслуживания одного прибора μ'_i , емкость БН – r'_i .

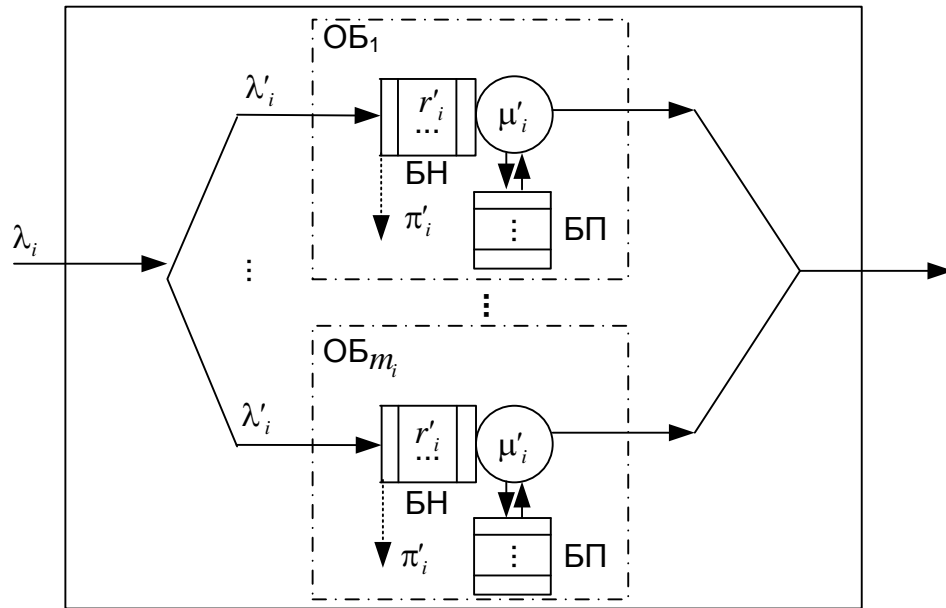


Рис.2.1, а. Модель 1-го элемента исследуемого фрагмента

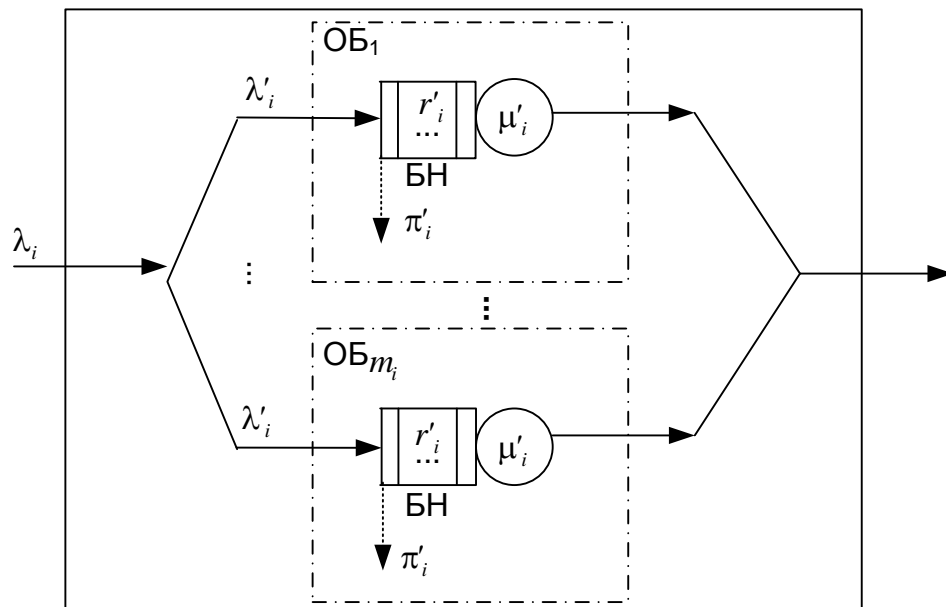


Рис.2.1, б. Модель 2-го и 3-го элемента исследуемого фрагмента

Вероятность потерь на БН i -го ОБ обозначим π'_i . Предположим, что длительность обслуживания на приборе распределена по экспоненциальному закону. Будем рассматривать дисциплину «пришедший первым обслуживается первым» (FCFS, First Come – First Served) [7], что соответствует работе MSC, HLR, VLR, рис. 2.1, а, б.

Учитывая равномерное распределение входящего пуассоновского потока заявок по обслуживающим блокам, идентичность ОБ, а также пренебрежимо малое время доступа к БП, отметим следующее:

– временной задержкой при доступе к БП при анализе можно пренебречь. Можно пренебречь также конечной емкостью БП и количеством заявок в ней, поскольку в момент поступления заявки в БП прибор ОБ освобождается для обработки следующей заявки. Размер буферной памяти на реальных обслуживающих элементах достигает 512 кбайт и более, что при учете среднего размера сигнального сообщения протокола Session Initiation Protocol Internetworking (SIP-I) в 700-800 байт и средних нагрузочных условиях [19] позволяет пренебречь потерями на БП и предполагать при расчетах бесконечную емкость БП. Для сообщений подсистем ОКС 7 [89], где размер сигнального сообщения значительно меньше ввиду бинарного формата, данное предположение является еще более приемлемым. Ввиду пренебрежимо малого времени, требующегося для обращения к БП, а также отсутствия потерь на ней вследствие принятой неограниченной емкости, можно исключить из модели рассмотрение БП и предположить, что все процессы, связанные с обращением к БП, учитываются при описании функционирования первого прибора. Это предположение не будет вносить существенных искажений при оценке работы устройства в отсутствие перегрузок;

– учитывая идентичность ОБ в рамках одного элемента исследуемого фрагмента, и равномерное распределение заявок между ОБ, можно сделать следующее приближение: объединим БН всех ОБ элемента в единый БН, емкостью $r_i = r'_i m_i$, при сохранении общего числа приборов. Таким образом, мы переходим от схемы полного разделения

БП (CP, Complete Partitioning) к полнодоступной схеме (CS, Complete Sharing) [8]. Отметим, что данный переход приводит к получению оптимистических оценок для вероятностно-временных характеристик (BBX) функционирования всей системы. Принимая сделанные предположения, от схемы элемента, представленной на рис. 2.1, а, б, переходим к более простой схеме элемента (рис. 2.2).

Таким образом, используя упрощающие предположения относительно режимов работы частей фрагмента, а также идею об идентичности обслуживания сообщений на ОБ, можно привести схему исследуемого фрагмента (рис. 1.5) к виду, представленному на рис. 2.3. Отдельно остановимся на режиме обработки заявок на первом элементе исследуемого фрагмента. Строгое описание процесса обработки заявок на первом элементе требует ввода трех классов заявок [7] – обозначим их как «новые» или приходящие во внешнем входящем потоке, а также «ответы от VLR» – пришедшие от второго элемента исследуемого фрагмента, и «ответы от HLR» – пришедшие от третьего элемента.

Предполагается, что «новые» заявки, равнозначны по своим свойствам, поэтому можно считать, что любая обслуженная на первом элементе «новая» заявка с вероятностями c_1 , c_2 , $c_1 + c_2 = 1$ может либо выйти из системы (если сообщение, поступившее на MSC не является начальным в процессе установления нового сеанса – оно будет обработано без участия VLR), либо попасть на второй элемент исследуемого фрагмента, соответственно (если поступающее на MSC сообщение является первым в процессе установления нового сеанса, оно вызовет обращение к VLR), рис. 2.3.

Заявки – ответы от второго элемента после повторной обработки на первом могут с вероятностью \tilde{c}_1 выйти из системы (т.е. информации, полученной от VLR достаточно и обращение к HLR не требуется), либо попасть на третий элемент с вероятностью \tilde{c}_3 (в VLR информация о вызываемом абоненте не найдена и требуется запрос к HLR), $\tilde{c}_1 + \tilde{c}_3 = 1$.

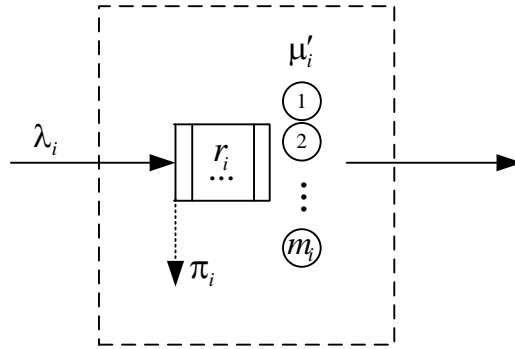


Рис.2.2. Модель работы i -й компоненты, используемой в исследовании

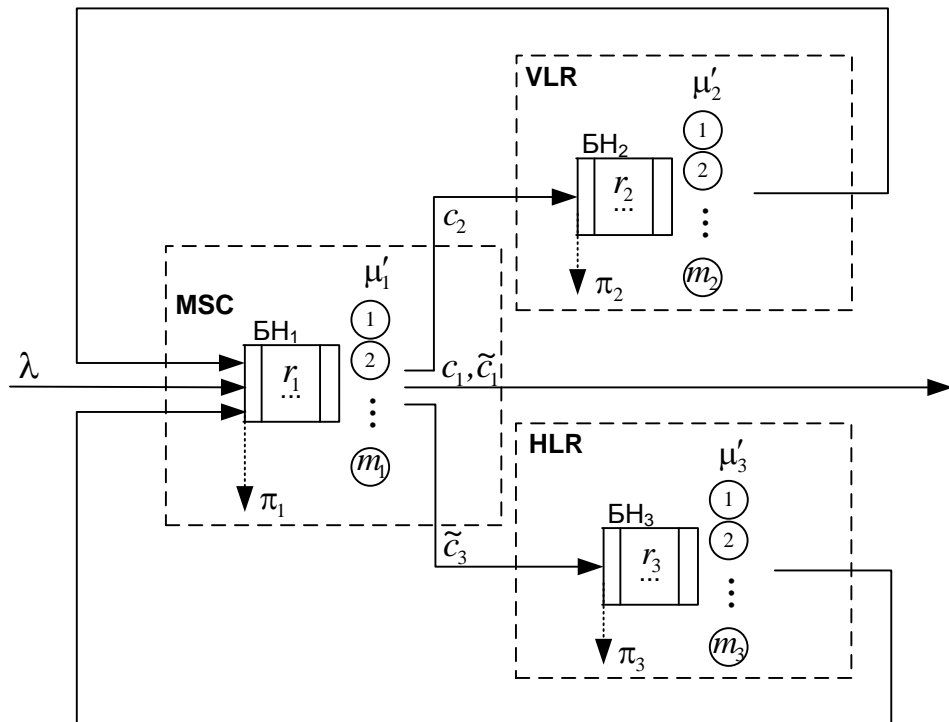


Рис.2.3. Схема работы фрагмента MSC/VLR/HLR

Все заявки – ответы от третьего элемента после обработки на первом покидают исследуемый фрагмент.

Значения c_1 , c_2 , \tilde{c}_1 , \tilde{c}_3 выбираются исходя из практики и экспериментальных характеристик изучаемой сети. Принимая во

внимание, что стандартное количество сообщений в процессе установления соединения по протоколу ISUP [11] (также как и по протоколам SIPv2.0/-I/-T [12]) равно семи, вероятность появления сообщения об установлении нового соединения будет равно $1/7 \approx 0.1429$. Оценочное значение c_1 можно положить равным $6/7 \approx 0.8571$ для стандартных случаев. Для выбора среднего значения запросов от MSC на HLR (в рамках процесса локализации местоположения вызываемого абонента [135]) можно руководствоваться количеством вызовов на абонентов, зарегистрированных в других MSC, оно колеблется в реальных сетях в зависимости от назначения коммутатора от 15 до 50 %. Для стандартных сетей связи данный параметр положим $\tilde{c}_3 = 0.2$. Безусловно, в особых условиях (например, MSC, обслуживающие линии метро, аэропорты) это значение может быть иное.

Вероятности потери заявок на БН элементов обозначим – π_1, π_2, π_3 , соответственно. Здесь и далее точка вместо индекса означает полную сумму по этому индексу.

Отметим, что получение распределения вероятностей для такой системы потребует хранения полного снимка очереди БН первого элемента для описания состояния исследуемого фрагмента.

Для описания состояний потребуется рассмотрение многомерного МП [31], размерности $3^{(r+m_1)} + 2$, что существенно осложняет проведение анализа. Для больших значений емкости БН первого элемента нахождение распределения вероятностей сильно затруднено в связи с ограничениями оперативной памяти компьютеров. В связи с этим, применение вышеописанного подхода к расчетам с практически значимыми значениями структурных параметров в большинстве случаев не применимо.

В связи с этим проведем дальнейшее упрощение модели. Опишем состояние первого элемента трехмерным вектором $\vec{q}_1^T = (q_{10}, q_{11}, q_{12})$.
Здесь:

– q_{10} – количество «новых» заявок на первом элементе (пришедших во внешнем потоке);

– q_{11} – количество заявок-ответов от второго элемента;

– q_{12} – количество заявок-ответов от третьего элемента.

Общее количество заявок всех типов обозначим

$$q_1 = q_{1\bullet}, \quad 0 \leq q_1 \leq R_1, \quad R_1 = r_1 + m_1. \quad (2.1)$$

В момент окончания обслуживания заявки на первом элементе будем разыгрывать тип заявки следующим образом:

– с вероятностью q_{10}/q_1 считаем, что была обслужена «новая» заявка, для которой возможны следующие события:

– с вероятностью c_1 заявка покидает исследуемый фрагмент;

– с вероятностью c_2 заявка уходит на второй элемент;

– с вероятностью q_{11}/q_1 считаем, что была обслужена заявка-ответ от второго элемента, для которой возможны следующие события:

– с вероятностью \tilde{c}_1 заявка покидает исследуемый фрагмент;

– с вероятностью \tilde{c}_3 заявка уходит на третий элемент;

– с вероятностью q_{12}/q_1 считаем, что была обслужена заявка-ответ от третьего элемента, которая с вероятностью 1 покидает СМО.

Для упрощения записи введем дополнительные обозначения:

– вероятность ухода «новой» заявки из исследуемый фрагмент $f_{10} = q_{10}c_1/q_1$;

– вероятность ухода из исследуемого фрагмента заявки-ответа от второго элемента $f_{11} = q_{11}\tilde{c}_1/q_1$;

– вероятность ухода из исследуемого фрагмента заявки-ответа от третьего элемента $f_{12} = q_{12}/q_1$;

– вероятность ухода заявки из исследуемого фрагмента $f_1 = f_{1\bullet}$;

– вероятность перехода заявки, обслуженной на первом элементе, на второй элемент $f_2 = q_{10}c_2/q_1$;

– вероятность перехода заявки, обслуженной на первом элементе, на третий элемент $f_3 = q_{11}\tilde{c}_3/q_1$.

Отдельно определим $f_i = 0, i = \{1,2,3,10,11\}$ при $q_1 = 0$, отметим, что $f_{\bullet} = 1$ [31].

На рис. 2.4 приведена обобщенная модель фрагмента MSC/VLR/HLR с учетом введенных обозначений и упрощений.

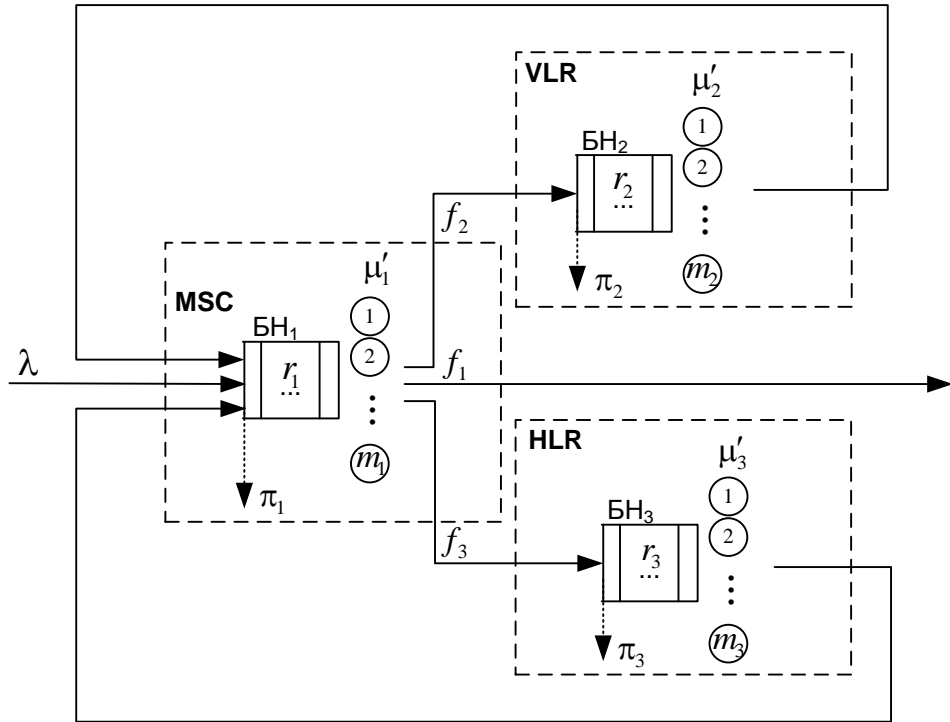


Рис.2.4. Обобщенная модель работы фрагмента MSC/VLR/HLR

Несмотря на введенные упрощения, изучение модели исследуемого фрагмента потребует рассмотрение пятимерного МП, однако вычисление вектора стационарных вероятностей \vec{p}^T при решении уравнения:

$$\vec{p}^T \mathbf{Q} = \mathbf{0}^T, \tag{2.2}$$

где \mathbf{Q} – инфинитезимальный оператор, $\mathbf{0} = \text{row}(0)$, будет сводиться к задаче, поддающейся решению с помощью стандартных вычислительных средств.

2.2. Анализ модели функционирования фрагмента MSC/VLR/HLR

Основываясь на предложенной в п. 2.1 модели фрагмента MSC/VLR/HLR перейдем к аналитическому решению. Рассмотрим МП:

$$\xi(t) = (\vec{\xi}(t))^T = (\xi_{10}(t), \xi_{11}(t), \xi_{12}(t), \xi_2(t), \xi_3(t)), \quad t > 0, \quad (2.3)$$

где $\xi_{10}(t)$, $\xi_{11}(t)$, $\xi_{12}(t)$ – количество заявок в первом элементе исследуемого фрагмента (реализациями процессов $\xi_{10}(t)$, $\xi_{11}(t)$, $\xi_{12}(t)$ в некий фиксированный момент времени t будем считать случайные переменные q_{10} , q_{11} , q_{12} , соответственно), $\xi_i(t)$ – количество заявок в i -м, $i = 2, 3$ элементе (реализации процессов – случайные переменные q_i); МП $\xi(t)$ определен над пространством состояний:

$$X = \left\{ \vec{x}^T = (x_1, x_2, \dots, x_5) : \dots \dots x_i = \overline{0, R_1}, i = \overline{1, 3}; x_i = \overline{0, R_i}, i = \{4, 5\} / \sum_{i=1,3} x_i \leq R_1 \right\}, \quad (2.4)$$

где $R_i = r_i + m_i$, $i = \overline{1, 3}$. Определим множество переходов над пространством состояний X :

$$H = \left\{ \vec{h} = (\vec{x}; \vec{y}) : \vec{x}; \vec{y} \in X \right\}, \quad (2.5)$$

Определим реализацию процесса $\xi(t)$, выражение (2.3), в некий фиксированный момент времени t , как $\vec{q}^T = (q_{10}, q_{11}, q_{12}, q_2, q_3)$, $\vec{q} \in X$, а стационарную вероятность нахождения СМО в состоянии \vec{q} : $\forall \vec{q} \in X : p(\vec{q}) = P\{\vec{q}\}$.

Для некоторых фиксированных \vec{q} и $\hat{\vec{q}}^T = (\hat{q}_{10}, \hat{q}_{11}, \hat{q}_{12}, \hat{q}_2, \hat{q}_3)$, $\hat{\vec{q}} \in X$ обозначим переход $\vec{s} = (\vec{q}; \hat{\vec{q}})$, $\vec{s} \in H$ тогда матрица интенсивностей переходов над множеством переходов H будет иметь вид:

$$\mathbf{Q} = \left\{ \gamma(\vec{s}) : \vec{s} \in H \right\}, \quad (2.6)$$

где $\gamma(\vec{s})$ – интенсивность перехода \vec{s} . Согласно [30], метод нахождения стационарного распределения вероятностей состоит в решении системы

линейных уравнений (СЛУ), представленной выражением (2.2) при учете нормировочного условия:

$$\bar{\mathbf{p}}^T \bar{\mathbf{1}} = 1, \quad (2.7)$$

где $\bar{\mathbf{p}} = \{p(\bar{\mathbf{q}}) : \bar{\mathbf{q}} \in X\}$ и $\bar{\mathbf{1}} = \text{row}(1, \dots, 1)$. Положение элемента $\gamma(\bar{\mathbf{s}})$ в матрице \mathbf{Q} определим согласно лексикографическому порядку, представленному в табл. 2.1. Элементы столбцов вычисляются в соответствии алгоритмом 2.1, представленным на рис. 2.5.

Для оптимизации изложения определим сокращенную запись для перехода $\bar{\mathbf{s}}$ между состояниями $\bar{\mathbf{q}}$ и $\hat{\bar{\mathbf{q}}}$, при котором часть координат $\bar{\mathbf{s}}$ остается неизменной (например, $q_{10} = \hat{q}_{10}$), следующим образом:

$$\bar{\mathbf{s}} = (q_{11}, q_{12}, q_2, q_3; \hat{q}_{11}, \hat{q}_{12}, \hat{q}_2, \hat{q}_3 / q_{10}).$$

Введем следующие обозначения [31]:

– функция Хэвисайда: $u(x) = \begin{cases} 1, x > 0, \\ 0, x \leq 0; \end{cases}$

– символ Кронекера: $\delta(i, j) = \begin{cases} 1, i = j, \\ 0, i \neq j; \end{cases}$

– единичная диагональная матрица $\mathbf{E} = \text{diag}(1, 1, \dots, 1)$;

– диагональная матрица со смещенной единичной диагональю

$$\mathbf{E}^{(n)} = (\delta_{i, j-n})_{i=1, j=1}^{M, N}, |n| \leq \max\{M, N\} \text{ (отметим, что } \mathbf{E} = \mathbf{E}^{(0)});$$

– $M_i = \mu'_i \min\{m_i, q_i\}, i = \overline{1, 3}$.

Табл. 2.1. Лексикографический порядок состояний $\bar{\mathbf{q}}_i, \bar{\mathbf{q}}_i \in X$

Порядковый номер i	q_{10}	q_{11}	q_{12}	q_2	q_3
1	0	0	0	0	0
2	0	0	0	0	1
...
$R_3 + 1$	0	0	0	0	R_3
$R_3 + 2$	0	0	0	1	0

$R_3 + 3$	0	0	0	1	1
...
$2(R_3 + 1)$	0	0	0	1	R_3
...
$I = \left(\sum_{i=0}^{R_1} \sum_{j=0}^{R_1-i} R_1 + 1 - j - i \right) \prod_{i=2,3} (R_i + 1)$	R_1	0	0	R_2	R_3

С учетом введенных обозначений и фрагмента диаграммы интенсивности переходов для $0 < q_i < R_1$, $i = \overline{10,12}$; $0 < q_i < R_i$, $i = 2,3$, представленного на рис. 2.6, интенсивности исходящих потоков для всех ненулевых переходов будут иметь вид:

$$\begin{aligned}
 \gamma(q_{10}; q_{10} - 1 / q_{11}, q_{12}, q_2, q_3) &= u(q_{10}) M_1 (f_{10} + \delta(R_2, q_2) f_2), \\
 \gamma(q_{11}; q_{11} - 1 / q_{10}, q_{12}, q_2, q_3) &= u(q_{11}) M_1 (f_{11} + \delta(R_3, q_3) f_3), \\
 \gamma(q_{12}; q_{12} - 1 / q_{10}, q_{11}, q_2, q_3) &= u(q_{12}) M_1 f_{12}, \\
 \gamma(q_2; q_2 - 1 / q_{10}, q_{11}, q_{12}, q_3) &= \delta(R_1, q_1) M_2, \\
 \gamma(q_3; q_3 - 1 / q_{10}, q_{11}, q_{12}, q_2) &= \delta(R_1, q_1) M_3, \\
 \gamma(q_{11}, q_2; q_{11} + 1, q_2 - 1 / q_{10}, q_{12}, q_3) &= u(R_1 - q_1) u(q_2) M_2, \\
 \gamma(q_{12}, q_3; q_{12} + 1, q_3 - 1 / q_{10}, q_{11}, q_2) &= u(R_1 - q_1) u(q_3) M_3, \\
 \gamma(q_{10}, q_2; q_{10} - 1, q_2 + 1 / q_{11}, q_{12}, q_3) &= u(q_{10}) u(R_2 - q_2) M_1 f_2, \\
 \gamma(q_{11}, q_3; q_{11} - 1, q_3 + 1 / q_{10}, q_{12}, q_2) &= u(q_{11}) u(R_3 - q_3) M_1 f_3, \\
 \gamma(q_{10}; q_{10} + 1 / q_{11}, q_{12}, q_2, q_3) &= u(R_1 - q_1) \lambda.
 \end{aligned} \tag{2.8}$$

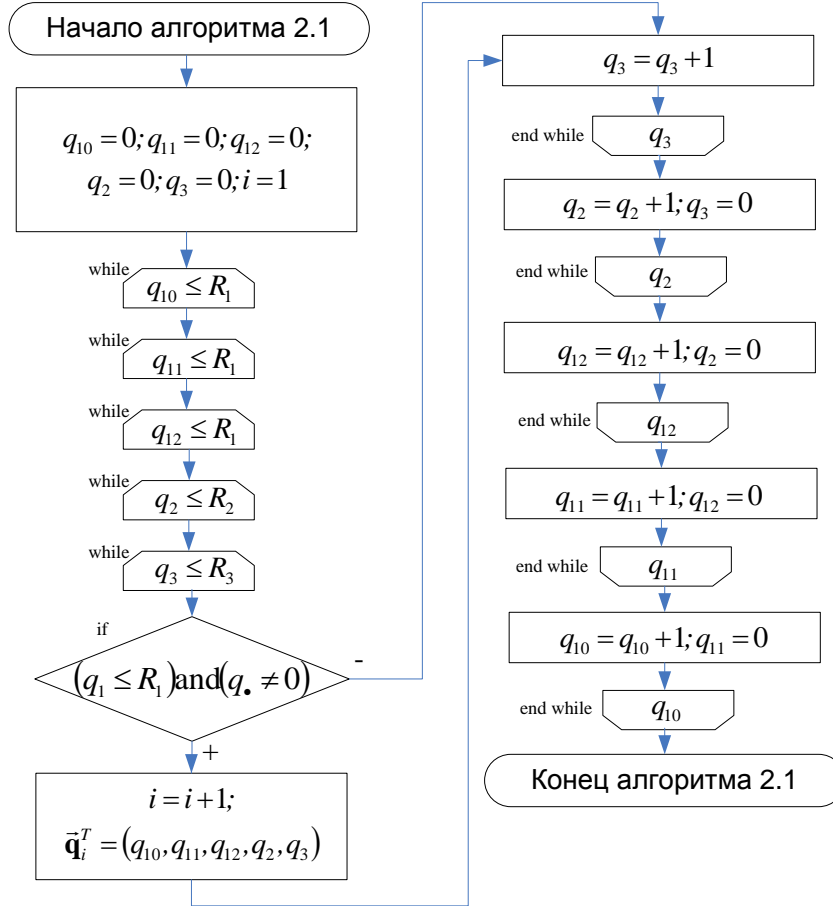


Рис.2.5. Алгоритм 2.1 порядковой нумерации состояния системы
(заполнение табл. 2.1)

Интенсивности входящих потоков для всех ненулевых переходов будут иметь вид:

$$\begin{aligned}
 \gamma(q_{10} - 1; q_{10} / q_{11}, q_{12}, q_2, q_3) &= u(q_{10})\lambda, \\
 \gamma(q_{10} + 1, q_2 - 1; q_{10}; q_2 / q_{11}, q_{12}, q_3) &= u(R_1 - q_1)u(q_2)M_1 f_2, \\
 \gamma(q_{11} + 1, q_3 - 1; q_{11}; q_3 / q_{10}, q_{12}, q_2) &= u(R_1 - q_1)u(q_3)M_1 f_3, \\
 \gamma(q_{11} - 1, q_2 + 1; q_{11}; q_2 / q_{10}, q_{12}, q_3) &= u(q_{11})u(R_2 - q_2)M_2, \\
 \gamma(q_{12} - 1, q_3 + 1; q_{12}; q_3 / q_{10}, q_{11}, q_2) &= u(q_{12})u(R_3 - q_3)M_3, \\
 \gamma(q_{10} + 1; q_{10} / q_{11}, q_{12}, q_2, q_3) &= u(R_1 - q_1)M_1(f_{10} + \delta(R_2, q_2)f_2), \\
 \gamma(q_{11} + 1; q_{11} / q_{10}, q_{12}, q_2, q_3) &= u(R_1 - q_1)M_1(f_{11} + \delta(R_3, q_3)f_3), \\
 \gamma(q_{12} + 1; q_{12} / q_{10}, q_{11}, q_2, q_3) &= u(R_1 - q_1)M_1 f_{12}, \\
 \gamma(q_2 + 1; q_2 / q_{10}, q_{11}, q_{12}, q_3) &= u(R_2 - q_2)\delta(R_1, q_1)M_2, \\
 \gamma(q_3 + 1; q_3 / q_{10}, q_{11}, q_{12}, q_2) &= u(R_3 - q_3)\delta(R_1, q_1)M_3.
 \end{aligned} \tag{2.9}$$

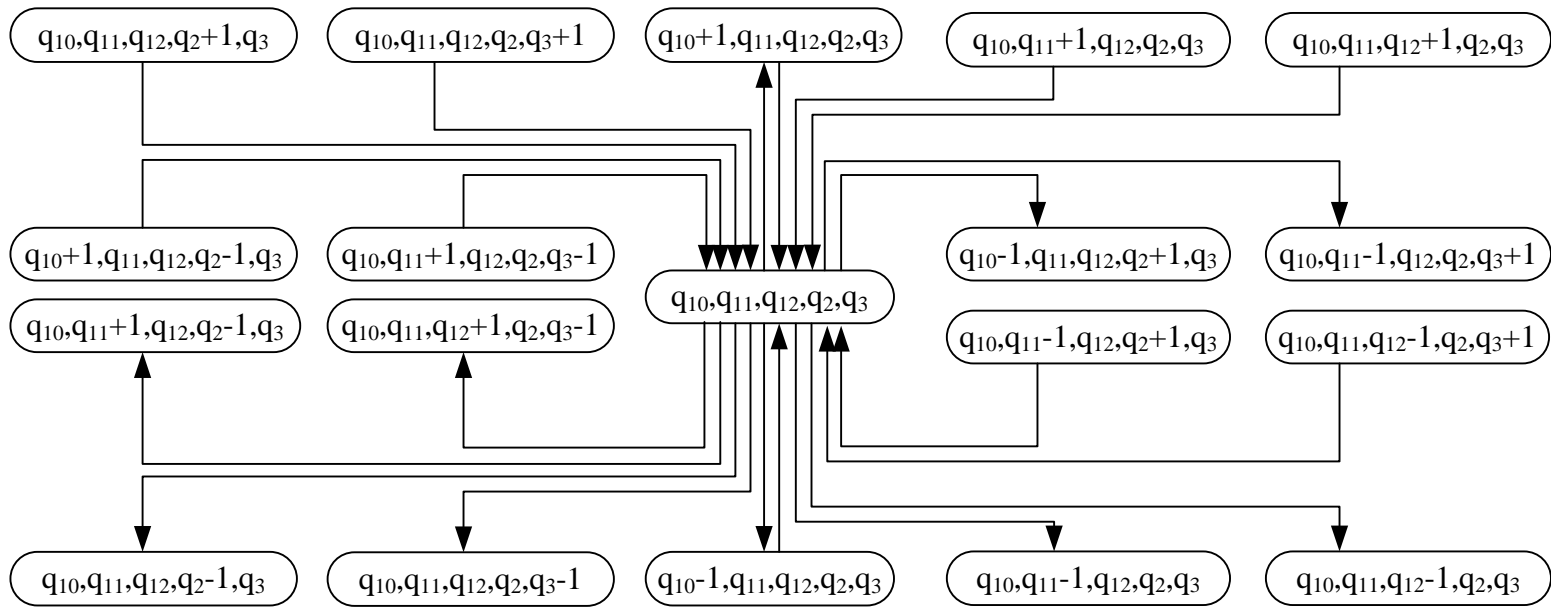


Рис.2.6. Фрагмент диаграммы интенсивностей переходов фрагмента MSC/VLR/HLR

В соответствии с порядком табл. 2.1 в дальнейшем стационарную вероятность состояния системы будем обозначать:

$$p_i = p(\vec{q}_i^T), \vec{q}_i^T \in X. \quad (2.10)$$

В соответствии с табл. 2.1, формат разбиения матрицы \mathbf{Q} имеет блочную пятиуровневую структуру: элементарный блок или блок первого уровня матрицы получается при фиксации значений параметров $q_i, i = \{10,11,12,2\}$. При фиксации первых трех параметров, $i = \{10,11,12\}$, будет образован блок второго уровня, который, в свою очередь, возможно разбить на элементарные блоки. Аналогичным образом определим блоки третьего и четвертого уровней на основе табл. 2.1. Блоком пятого уровня, очевидно, является сама матрица \mathbf{Q} .

Размеры блоков третьего и четвертого уровней не будут постоянны. Из табл. 2.1 и алгоритма 2.1 заключаем, что все блоки четвертого уровня будут иметь различный размер, что не позволяет выдвинуть предположение о квазиплоскости блочной матрицы \mathbf{Q} [28].

Используя нормировочное условие (2.7) запишем уравнение (2.2) следующим образом:

$$\vec{p}^T \hat{\mathbf{Q}} = (\vec{1}, \vec{0}^T), \quad (2.11)$$

где $\hat{\mathbf{Q}}$ – матрица \mathbf{Q} , первый столбец которой заменен на единичный, в силу условия (2.7). Решение данного уравнения при учете формата заполнения матрицы \mathbf{Q} является тривиальной задачей при малых значениях $r_i, i = \overline{1,3}$ и сопряжено с проблемами при $r_i \geq 10$. Остановимся на вычислительных аспектах решения.

Оптимизация решения задачи с использованием компактных схем хранения данных

Поскольку $|X|$ уже при $r_i + m_i = 10, i = \overline{1,3}$ достигает значения 10^3 , прямое решение уравнения путем обращения матрицы $\hat{\mathbf{Q}}$ является трудоемкой с вычислительной точки зрения задачей [39]. При дальнейшем увеличении максимальной емкости БН возникают

сложности с хранением \hat{Q} , а процесс ее обращения уже выходит за рамки возможностей стандартных персональных компьютеров. Учитывая эти факты, рассмотрим метод решения СЛУ (2.11) с использованием компактных методов хранения данных, и опишем более экономичный метод нахождения \bar{p} .

Одним из основных препятствий для расчета задачи для $r_i > 10$ является отсутствие необходимого объема оперативной памяти для хранения матрицы \hat{Q} . Отметим, что матрица Q является разреженной ленточной несимметричной матрицей с шириной ленты [39]:

$$\beta = 2 \left(1 + R_2 R_3 \sum_{i=1}^{R_1} (R_1 - i) i \right). \quad (2.12)$$

Поскольку ширина ленты достаточно велика, по сравнению с количеством ненулевых элементов, целесообразно для хранения подобной структуры использовать разреженный строчный формат, предложенный Ченгом и Густавсоном [39, п.1.8]. Использование данной схемы хранения позволит сократить объем занимаемой данными памяти. Обозначим максимальное количество ненулевых элементов в строке матрицы \hat{Q} как N_{nz} , тогда верхняя оценка необходимого для хранения матрицы \hat{Q} объема памяти складывается из следующих составляющих:

– $8 N_{nz} I$ байт – для хранения ненулевых значений $q(\bar{s})$ в 8-байтных ячейках памяти с двойной точностью;

– $4 N_{nz} I$ байт – для хранения массивов указателей и индексов столбцов.

Итого, верхняя оценка необходимого объема памяти для хранения матрицы \hat{Q} будет равна $12 N_{nz} I$ байт. Из выражений (2.8) и при учете (2.7) следует, что максимальное значение N_{nz} не превышает 12. Следовательно, верхняя оценка необходимого объема оперативной памяти для хранения \hat{Q} при использовании разреженного строчного формата равна $1584 I$ байт. Соотношение между объемом памяти,

требуемым для хранения матрицы $\hat{\mathbf{Q}}$ с использованием стандартной схемы и объемом при использовании разреженного строчного формата выражается соотношением $2I/396$. По сравнению со стандартной схемой хранения матриц (объем $8I^2$ байт), предложенный формат дает преимущества уже при $r_i > 3, i = \overline{1,3}$.

Используя разреженный строчный формат для хранения $\hat{\mathbf{Q}}$, предлагается применить к матрице Гауссова исключение по строкам [39] с пороговым выбором главного элемента для приведения ее к нижнетреугольному виду и последующего вычисления вектора (алгоритм 2.2). В качестве критерия разреженности при выборе главного элемента можно использовать критерий Марковица [39, гл. 5].

Приведем оценку ошибки в результате работы алгоритма. Условимся, что все операции над числами матрицы выполняются в формате чисел с двойной точностью [92]. Вследствие этого, машинная ошибка будет равна $\varepsilon_M = 2^{-t}$, где t – мантисса числа. Для чисел с двойной точностью $\varepsilon_M < 10^{-15}$ [39]. Один шаг Гауссова исключения будет давать следующую ошибку для приведенного элемента матрицы с позицией (i, j) [39]:

$$|e_{ij}^{(k)}| \leq (3 + 10^{-5}) \alpha_M(i, j) \varepsilon_M, \quad (2.13)$$

где $\alpha_M(i, j) > \max_k |q_{ij}^{(k)}|$.

Для верхней оценки суммарной ошибки алгоритма, введем обозначения для произвольной вещественной $n \times n$ матрицы \mathbf{X} :

– 1-норма $\|\mathbf{X}\|_1 = \max_j \sum_{i=1}^n |x_{ij}|$;

– ∞ -норма $\|\mathbf{X}\|_\infty = \max_i \sum_{j=1}^n |x_{ij}|$;

– \mathbf{E}_r – матрица ошибок факторизации;

– α_M – верхняя граница всех элементов матриц, редуцированных в процессе Гауссова алгоритма;

$$- n_{ij} = \sum_{k=1}^m n_{ij}^{(k)}, \quad (2.14)$$

где $n_{ij}^{(k)} = (l_{ik} \neq 0) \cap (u_{kj} \neq 0)$ – булева переменная.

Тогда, согласно [39], для сильноразреженной ленточной матрицы целесообразно применить критерий оценки ошибки без учета свойства ленточности, т.к. в общем случае ширина ленты много больше, чем количество ненулевых элементов в ней. Оба критерия (для ленточной и для сильно разреженной матрицы) применимы, однако второй дает для рассматриваемой задачи более точную оценку для общего случая:

$$\|\mathbf{E}_r\|_1 \leq (3 + 10^{-15}) \varepsilon_M \alpha_M \max_j \sum_{i=1}^I n_{ij}; \quad (2.15)$$

$$\|\mathbf{E}_r\|_\infty \leq (3 + 10^{-15}) \varepsilon_M \alpha_M \max_i \sum_{j=1}^I n_{ij}. \quad (2.16)$$

Алгоритм 2.2. Итерационный алгоритм вычисления значений $\bar{\mathbf{r}}$ с использованием компактных схем хранения данных.

Шаг 1. Определить переменные и константы:

- константу I согласно табл. 2.1;
- $k = I + 1$;
- матрицу $\hat{\mathbf{Q}}$ и вектор $\bar{\mathbf{b}}^T = (\mathbf{1}, \bar{\mathbf{0}}^T)$ согласно выражениям (2.8), (2.11);
- блочную матрицу $\mathbf{Z} = (\hat{\mathbf{Q}} | \bar{\mathbf{b}})$;
- ширину ленты β согласно (2.12);
- значение параметра допустимости u , $0 < u \leq 1$ (для введения критерия численной устойчивости).

Шаг 2. Определить:

- $K = \min(I, k + \beta / 2)$ - размер активной подматрицы в $\hat{\mathbf{Q}}$, в которой будет осуществляться поиск главного элемента;
- $k = k - 1$;
- $S_{pc} = \{q_{ij} = q(\bar{\mathbf{q}}_i^T; \bar{\mathbf{q}}_j^T) : q \neq 0; i, j = \overline{k, K}\}$ – подмножество первоначальных кандидатов – ненулевых элементов, среди которых будет осуществляться поиск главного элемента.

Шаг 3. По критерию Марковица [39] определить наилучший с точки зрения сохранения разреженности матрицы элемент q_m .

Шаг 4. Проверить выбранный элемент q_m удовлетворению по пороговому критерию численной устойчивости:

$$|q_m| \geq u \max_{k \leq p \leq K} |q_{pj}|;$$

$$|q_m| \geq u \max_{k \leq p \leq K} |q_{ip}|.$$

Шаг 5. Если ни одно из условий не удовлетворяется, то необходимо вернуться на шаг 3 и произвести повторный выбор наилучшего с точки зрения сохранения разреженности элемента, кандидатами в которые являются все элементы множества $S_{pc} = S_{pc} \setminus \{q_m\}$ (S_{pc} переопределяется). Если S_{pc} – пустое, то происходит выход из алгоритма и необходимо задать более мягкое значение параметра u .

Шаг 6. Если одно из условий, проверяемых на шаге 4, выполняется, элемент q_m выбирается главным элементом, и осуществляются необходимые перестановки строк и столбцов в матрице $\mathbf{Z}^{(k+1)}$ для перевода его на главную диагональ в позицию (k, k) . Здесь $\mathbf{Z}^{(k+1)}$ – матрица, полученная на предыдущем этапе (этап $k+1$) Гауссова приведения матрицы.

Шаг 7. Выполняется процедура Гауссова исключения по строкам для строки k в обратном порядке (снизу вверх) над элементами блочной матрицы \mathbf{Z} , где элементы матрицы на $(k-1)$ этапе вычисляются согласно выражению [117]:

$$z_{ij}^{(k-1)} = z_{ij}^{(k)} - u_{kj} l_{ik}, \quad (2.17)$$

$$u_{kj} = \frac{z_{kj}^{(k)}}{z_{ij}^{(k)}}, \quad l_{ik} = z_{ik}^{(k)}.$$

Отметим, что на шаге 7 последний столбец матрицы \mathbf{Z} (т.е. вектор $\bar{\mathbf{b}}$) остается неизменным из-за условия $b_i = 0$ при $i > 1$.

Шаг 8. Выполняется вычисление соответствующих членов матрицы ошибок \mathbf{E}_r согласно выражению (2.13). Далее:

- если $k > 1$: переход к **шагу 2**;
- иначе: переход к **шагу 9**.

Шаг 9. Вычислить булеву матрицу $\mathbf{N} = \{n_{ij}\}$, и нормы $\|\mathbf{E}_r\|_1$, $\|\mathbf{E}_r\|_\infty$ согласно выражениям (2.15) – (2.16).

Шаг 10. Определить $\mathbf{Q}_L = \{q_{ij}^{(L)}\} = \{z_{ij}^{(t)} : 1 \leq i, j \leq I\}$. Отметим, что вектор $\vec{\mathbf{b}}$ в результате процедуры приведения остался неизменным. Вычислить значения вектора $\vec{\mathbf{p}}^T$ из $\vec{\mathbf{p}}^T \mathbf{Q}_L = \vec{\mathbf{b}}$ используя прямую подстановку: $p_1 = b_1 / q_{1,1}^{(L)}$; $p_i = b_i - \sum_{j=1}^{i-1} b_j p_j$. Процедура сводится к последовательному вычислению векторов $\vec{\mathbf{b}}^{(1)} \equiv \vec{\mathbf{b}}$, $\vec{\mathbf{b}}^{(2)}$, ..., $\vec{\mathbf{b}}^{(t)}$. На k -м ($k = \overline{1, I}$) шаге выполняются следующие вычисления:

$$p_k = \left(\begin{array}{c} b_k^{(k)} \\ q_{kk}^{(L)} \end{array} \right); \quad (2.18)$$

$$b_i^{(k+1)} = b_i^{(k)} - q_{ik}^{(L)} p_k. \quad (2.19)$$

Шаг 11. Для оценки ошибки вычислений шага 10 возможно использовать следующее выражение [39]:

$$|\delta b_i| \leq (3 + 10^{-15}) \epsilon_M b_{Mi} (r_i^L + 1),$$

где r_i^L – количество ненулевых элементов строки i в \mathbf{Q}_L , $b_{Mi} = \max_k |b_i^{(k)}|$.

Шаг 12. Вычислить невязку численного и точного результатов:

$$\vec{\mathbf{r}}^T = \vec{\mathbf{p}}^T \mathbf{Q}_L - \vec{\mathbf{b}};$$

$$\vec{\mathbf{r}}^T = \vec{\mathbf{p}}^T (\mathbf{Q}_L - \mathbf{E}_r) - \vec{\mathbf{b}} = \vec{\mathbf{b}} + \delta \vec{\mathbf{b}} + \vec{\mathbf{p}}^T \mathbf{E}_r - \vec{\mathbf{b}} = \delta \vec{\mathbf{b}} + \vec{\mathbf{p}}^T \mathbf{E}_r;$$

$$\|\vec{\mathbf{r}}^T\| \leq \|\delta \vec{\mathbf{b}}\| + \|\vec{\mathbf{p}}^T\| \|\mathbf{E}_r\|. \quad (2.20)$$

Шаг 13. Окончание алгоритма. Сохранение результатов:

- вектор $\vec{\mathbf{p}}^T$;
- невязка $\|\vec{\mathbf{r}}^T\|$.

Алгоритм 2.2 предполагает использование разреженного строчного формата для хранения разреженной матрицы $\hat{\mathbf{Q}}$, что позволяет получить значительную экономию объемов памяти, а также сократить количество арифметических операций при решении уравнения (2.11). Экономия необходимой памяти и сокращения количества арифметических операций позволяет применять алгоритм для решения систем с широким диапазоном емкостей БН. При необходимости дальнейшей оптимизации вычислительного процесса возможен расчет не всего вектора $\bar{\mathbf{p}}^T = (p_0, p_1, \dots, p_l)$, а лишь его части, определенной над модифицированным пространством состояний \check{X} , полученным после удаления из X маловероятных состояний [49, 50], $\check{X} \subset X$.

Параметр порогового критерия Марковица позволяет влиять на численную устойчивость процесса решения (2.11). Наряду с вектором $\bar{\mathbf{p}}^T$ вычисляется также вектор невязки $\|\bar{\mathbf{r}}^T\|$, что позволяет оценить диапазоны нахождения точных значений $\bar{\mathbf{p}}^T$.

2.3. Анализ ВВХ СМО фрагмента MSC/VLR/HLR

Проведем анализ характеристик работы узла MSC/VLR/HLR. Для телекоммуникационных сетей критичным является процент потерь на сетевых элементах. В результате потери хотя бы одного сигнального сообщения сеанс может быть завершён некорректно, либо возникнет существенная задержка. Назначение данной модели – оценить загрузку узлов и вероятность потерь БН. На данных показателях остановимся при численном анализе.

Обозначим как $q_k^{(i)}$ количество заявок типа k в i -м состоянии, описываемом вектором $\bar{\mathbf{q}}_i^T$. Вероятность пребывания СМО в состоянии блокировки по одному из БН равна [30]:

$$\pi_1 = p \cdot \left(\bar{\mathbf{q}}_i : \sum_{i=10,12} q_i = r_1 + m_1 \right); \quad (2.21)$$

$$\pi_2 = p \cdot \left(\bar{\mathbf{q}}_i : q_2 = r_2 + m_2 \right); \quad (2.22)$$

$$\pi_3 = p_0(\bar{\mathbf{q}}_i : q_3 = r_3 + m_3) \quad (2.23)$$

Среднее количество заявок в СМО вычисляется по формулам:

$$N_1 = \sum_{j=10,12} \sum_{i=1}^I p(\bar{\mathbf{q}}_i) q_j \quad (2.24)$$

$$N_j = \sum_{i=1}^I p(\bar{\mathbf{q}}_i) q_j, \quad j = 2,3. \quad (2.25)$$

Среднее количество заявок на приборах вычисляется по формулам:

$$M_1 = \sum_{i=1}^I p(\bar{\mathbf{q}}_i) \min \left\{ m_1, \sum_{j=10,12} q_j \right\}; \quad (2.26)$$

$$M_j = \sum_{i=1}^I p(\bar{\mathbf{q}}_i) \min \{ m_j, q_j \}, \quad j = 2,3. \quad (2.27)$$

Результаты расчетов ВВХ СМО согласно выражениям (2.21) – (2.27) и для значений $m_i = 2$, $r_i = 2,4$, $i = \overline{1,3}$, показаны при изменении интенсивности входящего потока.

График $p_0(\lambda)$ отражает вероятность простоя системы: $p_0 = p(\bar{\mathbf{q}}_i : \mathbf{q} = 0)$. На графиках показаны кривые $\pi(\lambda)$, $N(\lambda)$, $M(\lambda)$ для каждого элемента системы: синим цветом – для первого элемента, зеленым – для второго, красным – для третьего. Зависимости, обозначенные пунктирными линиями на рис. 2.7, соответствуют случаю $r_i = 4$, сплошными – случаю $r_i = 2$, $i = \overline{1,3}$, $m_i = 2$. На графике для вероятности потерь на элементах исследуемого фрагмента хорошо видно влияние увеличения емкостей БН₁, БН₂, БН₃, на суммарное количество потерь при обработке сообщений на фрагменте и на каждом из его элементов. Вероятности потерь существенно снижаются при увеличении емкости БН_{*i*}, $i = \overline{1,3}$.

Исходные параметры взяты из практических соображений: для современных магистральных узлов MSC, VLR, HLR. Интенсивность обслуживания для одного прибора взята равной 500 сообщений в секунду, средняя продолжительность соединения – 120 с, среднее количество сообщений в рамках одного соединения – 10. При $m_i = 2$

верхняя оценка производительности такого сетевого элемента составит

около $\frac{120}{10} * 2 * 500 = 12000$ одновременных соединений.

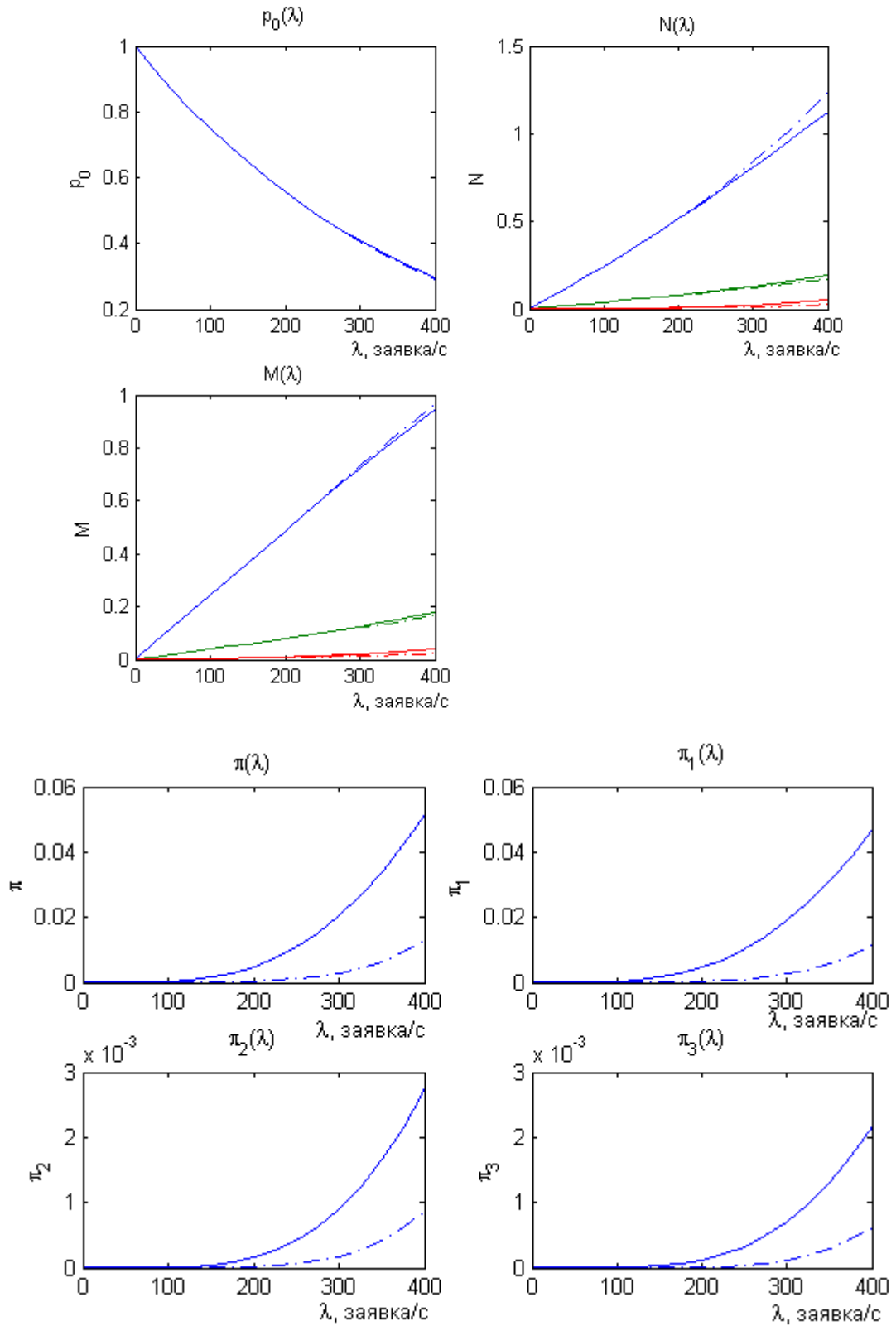


Рис.2.7. Графики для BBX системы MSC/VLR/HLR

Рост потерь на элементах исследуемого фрагмента при достижении $\lambda = 200, 300$ заявок/с, $r_i = 2, 4$, $m_i = 2$, $i = \overline{1, 3}$, соответственно, обусловлен описанными в главе 1 сценариями обработки вызовов: помимо входящего потока сообщений на элемент MSC приходят потоки обработанных заявок от VLR и HLR. Таким образом, для оптимального использования ресурсов сети и обеспечения равномерной загрузки элементов исследуемого фрагмента, целесообразно производительность $m_1\mu_1, m_2\mu_2, m_3\mu_3$ элементов находить из соотношения вероятностей, позволяющего обеспечить вероятность потерь на элементах исследуемого фрагмента, отличающуюся не более чем на порядок, а именно:

$$(1 + c_1 + c_1\tilde{c}_3) / c_2 / \tilde{c}_3 = m_1\mu_1 / m_2\mu_2 / m_3\mu_3. \quad (2.28)$$

Таким образом, использование данного соотношения при расчете мощности элементов фрагмента позволит обеспечить одинаковый порядок величины вероятности потерь на всех элементах. Исходя допустимого для данной сети уровня потерь на каждом из элементов с учетом (2.28) и результатов численного анализа модели, можно определить количественную величину значения $m_i\mu_i$ для элементов фрагмента [44, 61].

Для случая, когда при расчете необходимой производительности элементов исследуемого фрагмента использовалось соотношение (2.28), результаты численного моделирования представлены на рис. 2.8. В качестве максимально допустимой величины вероятности потерь сообщений на элементе бралось значение, равное 10^{-2} . Производительность элементов принималась следующей: $m_1\mu_1 = 1040$; $m_2\mu_2 = 160$; $m_3\mu_3 = 80$, $r_i = 2$, $i = \overline{1, 3}$. На рис. 2.8 видно, что выбранной производительности всех элементов фрагмента достаточно для обработки нагрузки с заданными значениями вероятности потерь на элементах.

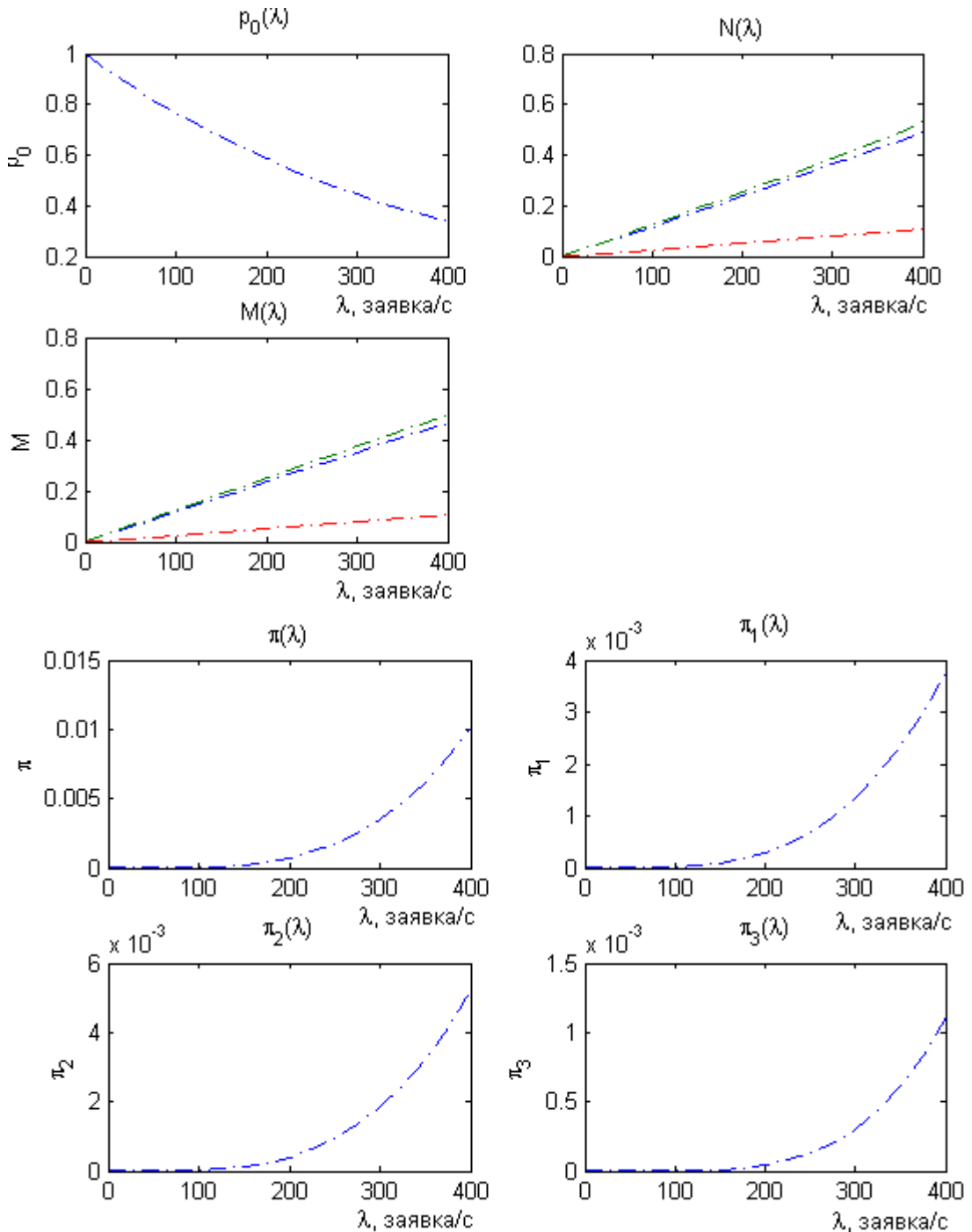


Рис.2.8. Графики для BBX системы MSC/VLR/HLR с предварительным выбором мощности элементов фрагмента

2.4. Обобщенная модель функционирования узла P-I-CSC

Устройства P-I-CSC имеют, обычно, несколько сетевых интерфейсов, центральный процессорный модуль с несколькими

физически или логически независимыми процессорами и две очереди для «надежного» и «ненадежного» трафика (рис. 2.9).

На основе анализа логики работы P-I-CSC была предложена обобщенная аналитическая модель P-I-CSC (рис. 2.10). В соответствии с наличием двух аппаратно-независимых очередей в предлагаемой модели предусмотрены два БН для поступающих ПЗ и НЗ.

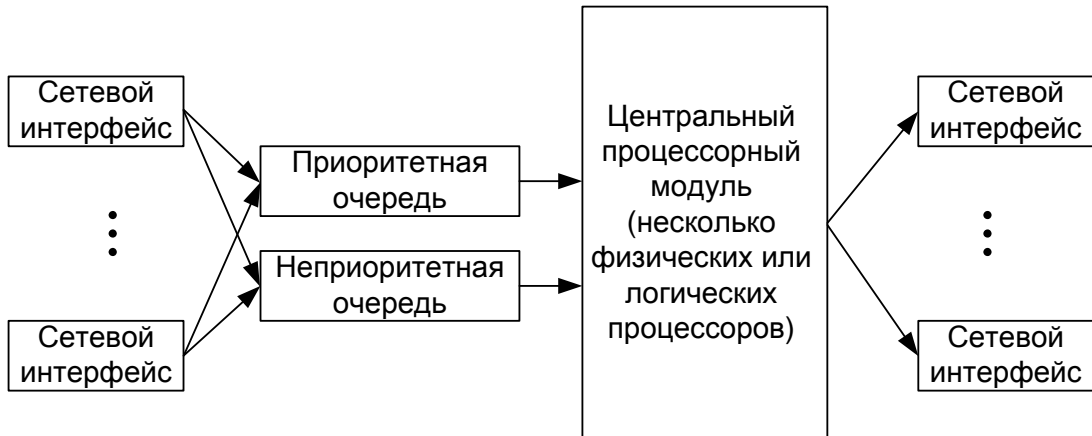


Рис.2.9. Логическая схема обработки сообщений в P-I-CSC

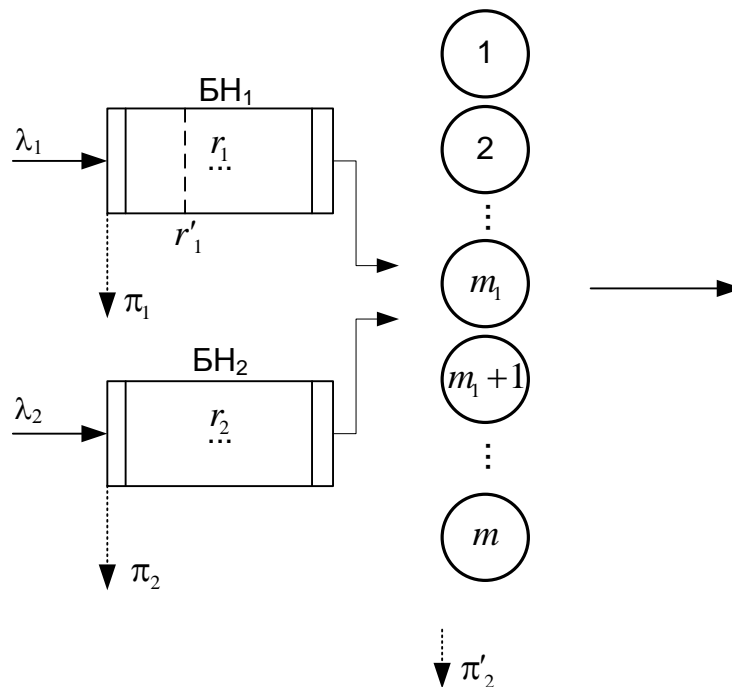


Рис.2.10. Обобщенная модель работы узла P-I-CSC

Будем считать, что на вход фрагмента Р-И-CSC поступает два пуассоновских потока заявок:

– поток приоритетных заявок (ПЗ) от «надежных» источников с интенсивностью λ_1 попадает в приоритетную очередь (БН₁);

– поток неприоритетных заявок (НЗ) от «ненадежных» источников с интенсивностью λ_2 – в не приоритетную очередь (БН₂).

Емкости БН₁ и БН₂ для ПЗ и НЗ обозначим r_1 и r_2 , $0 < r_i < \infty$, $i = 1, 2$, потери – π_1 , π_2 , соответственно. Будем рассматривать дисциплину абсолютного приоритета при обслуживании. Вероятность потери НЗ при обслуживании обозначим π'_2 (при вытеснении НЗ приоритетной). Заявки обоих приоритетов обслуживаются m приборами с экспоненциальным временем обслуживания (интенсивности μ). Выбор заявок на обслуживание будет осуществляться по следующему алгоритму:

– если количество заявок в приоритетной очереди меньше r'_1 , ПЗ обслуживаются m_1 приборами, $m_1 < m$, $0 < r'_1 \leq r_1$;

– если количество заявок в приоритетной очереди больше либо равно r'_1 , то ПЗ обслуживаются m приборами. При этом сперва очередь занимают свободные приборы, и если свободных нет, ПЗ занимает занятый НЗ прибор, при этом НЗ теряется. Процесс происходит указанным образом до тех пор, пока длина приоритетной очереди не станет меньше r'_1 ;

– распределение НЗ происходит всегда только на свободные приборы с номерами от $m_1 + 1$ до m , НЗ могут занимать максимум m_2 приборов, где $m_2 = m - m_1$.

Отметим, что при уменьшении параметра r'_1 , СМО входит в состояние, в котором возможно вытеснение НЗ тем раньше, чем меньше r'_1 . Данное свойство СМО будем называть агрессивностью по отношению к НЗ.

2.5. Анализ модели функционирования узла P-I-CSC

Перейдем к аналитическому решению модели P-I-CSC, предложенной в п. 2.4. Рассмотрим МП:

$$\xi(t) = \bar{\xi}(t)^T = (\xi_1(t), \xi_2(t), \xi_3(t), \xi_4(t)), \quad t > 0, \quad (2.29)$$

где $\xi_1(t)$ – число ПЗ в БН₁, $\xi_2(t)$ – число ПЗ на первых m_1 приборах (рис. 2.10), $\xi_3(t)$ – число ПЗ на приборах с $m_1 + 1$ по m , $\xi_4(t)$ – число НЗ в СМО (реализации процессов – случайные переменные q_1, q_2, q_3, q_4).

Процесс $\xi(t)$ определен над пространством состояний:

$$X = \{\bar{\mathbf{x}}^T = (x_1, x_2, x_3, x_4)\}, \quad (2.30)$$

где $0 \leq x_1 \leq r_1, 0 \leq x_4 \leq r_4 + m_2; 0 \leq x_i \leq m_i, i = 2, 3$.

Множество переходов над X есть:

$$H = \{(\bar{\mathbf{x}}; \bar{\mathbf{y}}) : \bar{\mathbf{x}}, \bar{\mathbf{y}} \in X\}. \quad (2.31)$$

Определим реализацию МП $\xi(t)$ в некий момент времени t , как $\bar{\mathbf{q}}^T = (q_1, q_2, q_3, q_4)$, $\bar{\mathbf{q}} \in X$ а стационарную вероятность пребывания СМО в состоянии $\bar{\mathbf{q}}$ как $p(\bar{\mathbf{q}}) = P\{\bar{\mathbf{q}}\}, \forall \bar{\mathbf{q}} \in X$.

Для некоторых фиксированных $\bar{\mathbf{q}}$ и $\hat{\mathbf{q}}^T = (\hat{q}_1, \hat{q}_2, \hat{q}_3, \hat{q}_4)$, обозначим переход $\bar{\mathbf{s}} = (\bar{\mathbf{q}}; \hat{\mathbf{q}})$, $\bar{\mathbf{s}} \in H$, тогда матрица интенсивностей переходов над H будет иметь вид:

$$\mathbf{Q} = \{\gamma(\bar{\mathbf{s}}) : \bar{\mathbf{s}} \in H\}, \quad (2.32)$$

где $\gamma(\bar{\mathbf{s}})$ – интенсивность перехода $\bar{\mathbf{s}}$. Положение элемента $\gamma(\bar{\mathbf{s}})$ в матрице \mathbf{Q} определим согласно лексикографическому порядку, представленному в табл. 2.2. Алгоритм нумерации состояний системы представлен на рис. 2.11, фрагмент диаграммы интенсивностей переходов для $0 < q_1 < r_1, 0 < q_4 < r_4 + m_2; 0 < q_i < m_i, i = 2, 3$. представлен на рис. 2.12.

На рис. 2.11 и далее $q_k^{(i)}$ будет обозначать k -ю координату $\bar{\mathbf{q}}_i^T$, $i = \overline{1, I}$, $\bar{\mathbf{q}}_i^T \in X$. Для описания перехода $\bar{\mathbf{s}}$ между состояниями $\bar{\mathbf{q}}^T$ и $\hat{\mathbf{q}}^T$

будем использовать сокращенный формат записи, введенный в п. 2.2 диссертации.

Табл. 2.2. Лексикографический порядок состояний $\bar{q}_i^T, \bar{q}_i^T \in X$

Порядковый номер, i	q_1	q_2	q_3	q_4
1	0	0	0	0
2	0	0	0	1
3	0	0	0	2
...
I	r_1	m_1	$m - m_1$	r_2

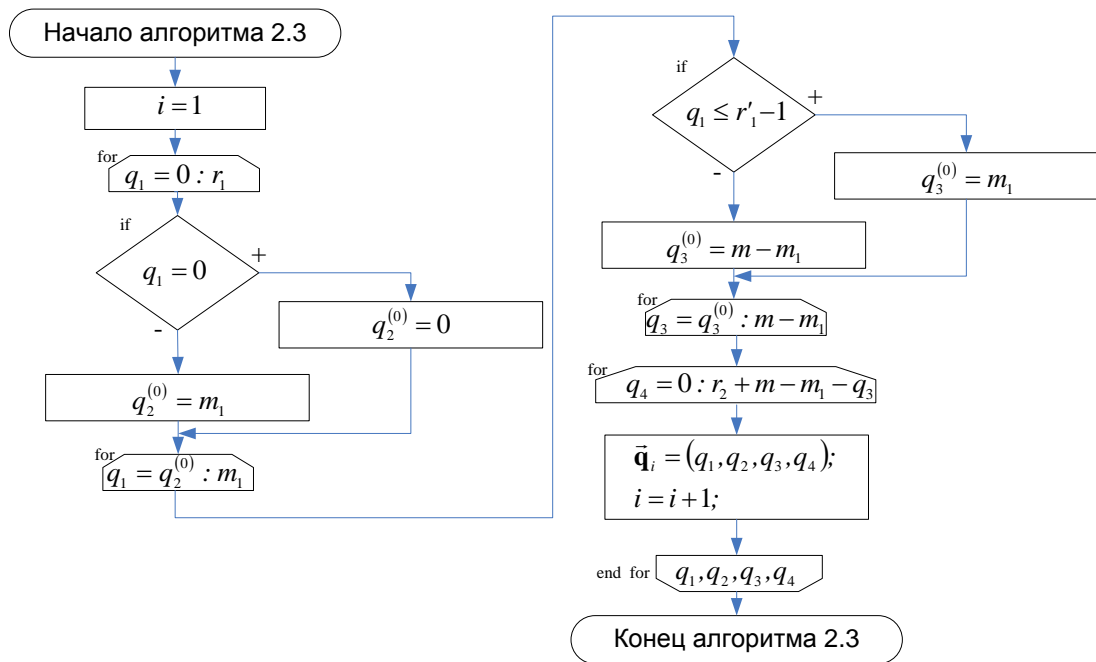


Рис.2.11. Алгоритм нумерации состояний системы

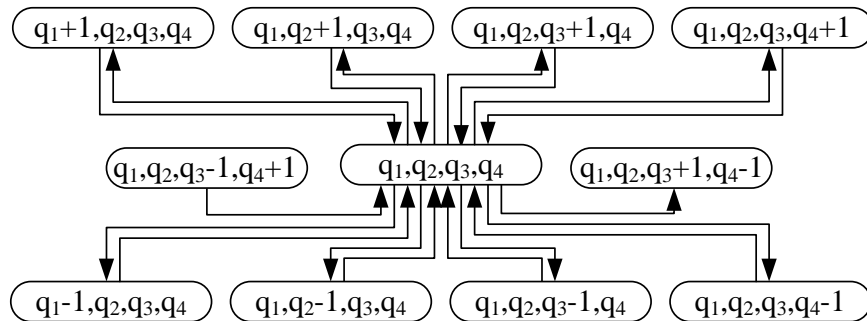


Рис. 2.12. Фрагмент диаграммы переходов

Ниже предложен пошаговый алгоритм заполнения матрицы интенсивностей переходов с учетом введенных обозначений:

Алгоритм 2.3. Пошаговый алгоритм заполнения матрицы **Q**

Шаг 1. Вычислить значение булевой переменной $b = (q_2 = m_1 \cap q_1 = r'_1 - 1 \cap q_4 \geq m - m_1 - q_3 \cap q_3 \neq m - m_1)$;

Переменная b принимает значение *истина* в случае, если СМО находится в состоянии, в котором возможно вытеснение НЗ.

Шаг 2. Пусть $\bar{s} = (q_1; q_1 - 1 / q_2, q_3, q_4)$, $\bar{s} \in H$, тогда

– если $q_1 \leq r'_1 - 1$: $\gamma(\bar{s}) = m_1 \mu$;

– если $q_1 \geq r'_1$: $\gamma(\bar{s}) = m \mu$.

Шаг 3. Пусть $\bar{s} = (q_2; q_2 - 1 / q_1, q_3, q_4)$, $\bar{s} \in H$, тогда $\gamma(\bar{s}) = q_2 \mu$.

Шаг 4. Пусть $\bar{s} = (q_3; q_3 - 1 / q_1, q_2, q_4)$, $\bar{s} \in H$, тогда $\gamma(\bar{s}) = q_3 \mu$.

Шаг 5. Пусть $\bar{s} = (q_4; q_4 - 1 / q_1, q_2, q_3)$, $\bar{s} \in H$, тогда

– если $\bar{b} \cap q_4 \geq m - m_1 - q_3$: $\gamma(\bar{s}) = (m - m_1 - q_3) \mu$;

– если $b = \text{истина}$: $\gamma(\bar{s}) = \max \{0, (m - m_1 - q_3) \mu - \lambda_1\}$;

– если $m - m_1 - q_3 > 0 \cap q_4 < m - m_1 - q_3$: $\gamma(\bar{s}) = q_4 \mu$.

Шаг 6. Пусть $\bar{s} = (q_1; q_1 + 1 / q_2, q_3, q_4)$, $\bar{s} \in H$, тогда если $((q_2 = m_1 \cap q_1 < r'_1 - 1) \cup (q_2 = m_1 \cap q_3 = m - m_1)) \cap q \neq r_1 + m$: $\gamma(\bar{s}) = \lambda_1$.

Шаг 7. Пусть $\bar{s} = (q_2; q_2 + 1 / q_1, q_3, q_4)$, $\bar{s} \in H$, тогда если $q_2 < m_1$: $\gamma(\bar{s}) = \lambda_1$.

Шаг 8. Пусть $\bar{s} = (q_3; q_3 + 1 / q_1, q_2, q_4)$, $\bar{s} \in H$, тогда если $q_3 = m_1 \cap q_1 = r'_1 - 1 \cap q_4 < m - m_1 - q_3$: $\gamma(\bar{s}) = \lambda_1$.

Шаг 9. Пусть $\bar{s} = (q_4; q_4 + 1 / q_1, q_2, q_3)$, $\bar{s} \in H$, тогда если $q_4 \neq r_2 + m - m_1 - q_3$: $\gamma(\bar{s}) = \lambda_2$.

Шаг 10. Пусть $\bar{s} = (q_3, q_4; q_3 + 1, q_4 - 1 / q_1, q_2)$, $\bar{s} \in H$, тогда если $b = \text{истина}$: $\gamma(\bar{s}) = \lambda_1$.

Шаг 11. Пусть $\bar{s} = (q_1, q_2, q_3, q_4; \hat{q}_1, \hat{q}_2, \hat{q}_3, \hat{q}_4)$, $\bar{s} \in H$ (главная диагональ), тогда: $\gamma(\bar{s}) = q(q_1, q_2, q_3, q_4; \hat{q}_1, \hat{q}_2, \hat{q}_3, \hat{q}_4)$.

Шаг 12. Все остальные элементы матрицы \mathbf{Q} являются нулевыми.

Используя нормировочное условие (2.10) запишем выражение (2.2) следующим образом:

$$\bar{\mathbf{p}}^T \hat{\mathbf{Q}} = (\bar{\mathbf{0}}^T, 1), \quad (2.33)$$

где $\hat{\mathbf{Q}}$ – матрица \mathbf{Q} , последний столбец которой заменен на единичный, в силу условия (2.9).

Отметим, что в соответствии с алгоритмом 2.3 в \mathbf{Q} имеется шесть ненулевых частичных полос диагонального типа. В связи с этим, при приведении матрицы $\hat{\mathbf{Q}}$ к верхнетреугольному виду с применением прямого хода метода Гаусса, количество операций умножения на константу и сложения существенно сокращается по сравнению с общим случаем. Пусть $X' = \{\bar{\mathbf{q}} \in X : q_1 = 0, q_2 = 0\}$, $\chi' = |X'|$, тогда количество необходимых операций умножения и сложения при прямом ходе Гаусса равно $\chi' \left(\sum_{j=0}^l (l-j) - \sum_{j=0}^{l-\chi'} (l-\chi'-j) \right)$.

2.6. Анализ ВВХ СМО Р-Л-CSC

Дополнительно к аналитическому решению СМО узла Р-Л-CSC проведем анализ характеристик его работы.

Вероятности потерь на БН₁ и БН₂ находятся по формулам:

$$\pi_1 = p.(\bar{\mathbf{q}}_i : q_1 = r_1 + m), \quad (2.34)$$

$$\pi_2 = p.(\bar{\mathbf{q}}_i : q_4 = r_2 + m - m_1 - q_3). \quad (2.35)$$

Вероятность потери неприоритетной заявки на приборе в результате вытеснения вычисляется по формуле:

$$\pi'_2 = p.(\bar{\mathbf{q}}_i : b = \text{истина}). \quad (2.36)$$

Среднее количество заявок в СМО:

$$N_1 = \sum_{i=1}^l p(\bar{\mathbf{q}}_i) \sum_{k=1,3} q_k^{(i)}, \quad (2.37)$$

$$N_2 = \sum_{i=1}^l p(\bar{\mathbf{q}}_i) q_4^{(i)}, \quad (2.38)$$

$$N = N_1 + N_2. \quad (2.39)$$

Среднее время пребывания ПЗ и НЗ соответственно, вычисляется по формуле:

$$T_i = N_i / \lambda_i, \quad i = 1, 2 \quad (2.40)$$

Среднее количество заявок (приоритетных и неприоритетных соответственно) на приборах есть:

$$M_1 = \sum_{i=1}^I p(\bar{q}_i) \sum_{k=2,3} q_k^{(i)}, \quad (2.41)$$

$$M_2 = \sum_{i=1}^I p(\bar{q}_i) q_4^{(i)} \max\{0, q_4^{(i)} - q_3^{(i)} - r_2\}. \quad (2.42)$$

Среднее время ожидания в БН₁ и БН₂ есть:

$$W_1 = \lambda_1^{-1} \sum_{i=1}^I p(\bar{q}_i) q_1^{(i)}, \quad (2.43)$$

$$W_2 = \lambda_2^{-1} \sum_{i=1}^I p(\bar{q}_i) \max\{0, q_4^{(i)} - (m - m_1 - q_3^{(i)})\}. \quad (2.44)$$

Графики расчетов средних характеристик СМО согласно выражениям (2.36) – (2.44) при $m_1 = 6$, $m_2 = 2$, $r_1 = 20$, $r_2 = 20$, $\lambda_1 = 1000$ заявок/с, $\lambda_2 = 300$ заявок/с представлены на рис. 2.13. Исходные параметры взяты из практических соображений: магистральные узлы Р-/I-CSC организовываются на основе мультипроцессорных ПАК. Решения с восьмью процессорами (приборами) широко распространены. Средняя вычислительная способность одного процессора при полной загрузке позволяет обслуживать от 3 до 5 тысяч одновременных SIP соединений, что соответствует интенсивности обработки сообщения и равной около 200 сообщений в секунду (при ориентировочной длительности вызова в 120 с).

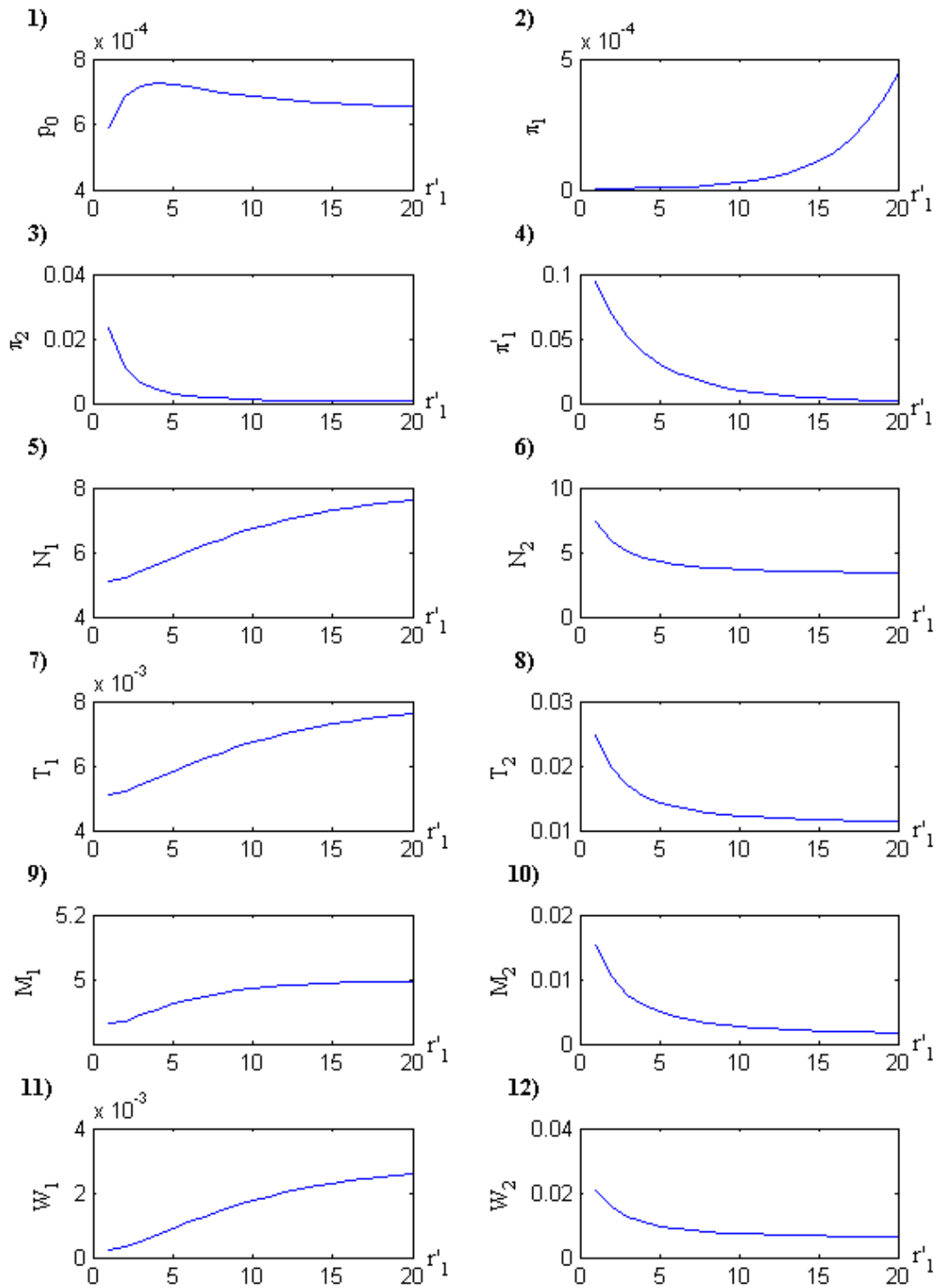


Рис. 2.13. Графики ВВХ функционирования СМО

Рассмотрим подробнее результаты анализа, представленные на рис. 2.13. Для вероятности простоя СМО $p_0(r'_1)$, график 1) на рис. 2.13, характерен максимум при $r'_1=4$, т.к. при данном значении порога

среднее количество заявок обоих типов в СМО минимально (рис. 2.13, 5 и 6). При малых значениях r'_1 вероятность простоя СМО меньше из-за накопления заявок в БН₂ (рис. 2.13, 12). При увеличении r'_1 наблюдается менее агрессивное поведение приоритетных заявок, что способствует резкому снижению времени ожидания в БН₂ у неприоритетных заявок, при незначительном росте времени ожидания приоритетных. Суммарный эффект выражается в наблюдаемом максимуме на рис. 2.13, 1 – для графика $p_0(r'_1)$.

При дальнейшем увеличении параметра r'_1 наблюдается плавное уменьшение $p_0(r'_1)$ вследствие того, что среднее время ожидания в БН₁ растет при практически неизменном параметре ожидания НЗ в БН₂.

Для численного решения уравнения (2.33) и нахождения значений вектора $\bar{\mathbf{p}}^T$ использовался прямой и обратный ход метода Гаусса, в котором учитывалась структура матрицы \mathbf{Q} (сокращение количества операций за счет учета верхней и нижней треугольной нулевой части). Вычисления данным способом при обозначенных параметрах СМО заняли менее минуты (на персональном компьютере на базе Intel Core i5, 8 Гб ОЗУ под управлением операционной системы (ОС) Linux Debian 6.3, Matlab R2011b 7.13).

Расчеты для небольших значений емкостей БН₁, БН₂ позволяют определить характер поведения СМО и составить прогноз поведения для больших значений емкостей накопителей. В случае, если необходим точный расчет подобных СМО с большим количеством состояний, оправдано использование подхода, описанного в п. 2 настоящей главы и, в частности, алгоритм 2.2 из п.2.2. Такой метод вычисления вектора стационарных вероятностей состояний СМО несколько усложняет программный код, но позволяет существенно повысить возможности метода для исследования систем с большими значениями емкостей БН и числом приборов.

2.7. Выводы

1. Предложена и исследована обобщенная аналитическая модель фрагмента конвергентной сети MSC/VLR/HLR, определены и решены системы уравнений глобального баланса для модели. Получено решение, которое позволило определить вероятности потери сигнального сообщения в сети на элементах фрагмента. На основе полученных зависимостей вероятности потерь сообщений входящего потока определяется максимально допустимый поток при заданном значении вероятности сброса.

2. Предложена и исследована обобщенная аналитическая модель фрагмента конвергентной сети на базе устройств P-I-CSC с двумя аппаратными БН конечной емкости, наличием заявок с абсолютным приоритетом и пороговым механизмом регулирования процесса обслуживания. Найдена система уравнений глобального баланса для предложенной аналитической модели и получено ее решение. На основании полученных результатов определено значение порогового параметра r'_1 , при котором вероятность простоя СМО максимальна при небольшом возрастании вероятности потерь.

3. Предложенная модель P-I-CSC является агрессивной по отношению к неприоритетным заявкам при малых значениях порогового параметра системы r'_1 , что соответствует реальным режимам работы устройства P-I-CSC в условиях обработки запросов от сетей сторонних операторов, являющихся потенциальными источниками угроз.

ГЛАВА 3

ИССЛЕДОВАНИЕ И РАЗРАБОТКА МОДЕЛИ СИГНАЛЬНОГО ОБМЕНА В КОНВЕРГЕНТНОЙ СЕТИ

3.1. Разработка процедуры построения математической модели предоставления услуг в конвергентной сети

Информация, передаваемая в конвергентных сетях, делится на два типа: служебная или сигнальная информация и пользовательские данные. Схемы обработки информации также принципиально отличаются в зависимости от ее типа.

Продолжая исследование конвергентных сетей с использованием методов ТМО, в данной главе перейдем от изучения фрагментов сети к рассмотрению обобщенной модели функционирования основных элементов конвергентной сети в их взаимодействии. Основной целью построения моделей функционирования конвергентных сетей можно считать оценку параметров работы сети как единой системы или оценку параметров работы каждого узла сети.

При оценке параметров качества функционирования сети как единого объекта наиболее показательными являются следующие параметры [126]:

- задержка от узла-источника до узла назначения при выполнении служебных процедур, обеспечивающих доступ к ресурсам сети, услугам;
- задержка от узла-источника до узла назначения при передаче пользовательских данных в рамках уже установленного соединения;
- вероятность потерь сигнального сообщения в сети;
- вероятность потерь при передаче пользовательских данных в уже установленном соединении;
- вероятность неуспешного завершения служебной процедуры.

При более детальном анализе работы сети, в котором возможны оценки характеристик функционирования каждого из узлов сети, наиболее важным с точки зрения работы сети представляется моделирование процессов обмена и обработки сигнальной информации. Приоритет построения моделей обработки именно сигнальной информации имеет несколько причин:

– процессы обмена сигнальными сообщениями и их обработки, зачастую являются более сложными, чем процессы обработки пользовательских данных.

– в случае проблем с обменом сигнальной информацией дальнейший анализ и исследование механизмов работы сети не имеет смысла в силу ее неработоспособности.

Применение СеМО к моделированию процессов обмена сигнальной информацией в конвергентной сети

При моделировании процессов обмена сигнальной информации будем использовать сеть ВСМР [70] для моделирования нестандартных типов узлов. Моделирование процессов сигнального обмена с помощью сетей Джексона [74, 122] и Гордона – Ньюела [90] не представляется возможным ввиду отсутствия в них описания заявок с помощью различных классов.

В сетях ВСМР [70] предполагается наличие входящих пуассоновских потоков заявок, классы заявок, смена классов заявок, возможность использования в качестве узлов СеМО ВСМР четырех типов СМО:

- $\bullet/M/m$ с экспоненциальным законом обслуживания и дисциплиной обслуживания FCFS;
- $\bullet/G/1$ с разделением ресурсов (PS, Processor Sharing);
- $\bullet/G/\infty$ с бесконечным количеством приборов (IS, Infinite Servers);
- $\bullet/G/1$ с дисциплиной «пришедший последним, обслуживается первым» (LCFS, Last Come – First Served).

Обозначив множество классов заявок M , определим согласно некому взаимнооднозначному соответствию нумерацию классов с диапазоном номеров $k = \overline{1, |M|}$. Для каждого из типов СМО ВСМР приведем выражение для вероятности нахождения узла в состоянии $(n_1, n_2, \dots, n_k, \dots, n_m)$, $k = \overline{1, |M|}$, где n_k – количество заявок k -класса в узле. Согласно [9], для всех типов узлов сети ВСМР существуют решения:

$$- P(n_1, n_2, \dots, n_l) = (1 - \rho) m! \prod_{k=1}^{|M|} \frac{\rho_k^{n_k}}{n_k!} \text{ для СМО } \bullet / M / m, \text{ FCFS}; \quad (3.1a)$$

$$- P(n_1, n_2, \dots, n_l) = (1 - \rho) m! \prod_{k=1}^{|M|} \frac{\rho_k^{n_k}}{n_k!} \text{ для СМО } \bullet / G / 1, \text{ PS}; \quad (3.1б)$$

$$- P(n_1, n_2, \dots, n_l) = e^{-\rho} \prod_{k=1}^{|M|} \frac{\rho_k^{n_k}}{n_k!} \text{ для СМО } \bullet / G / \infty, \text{ IS}; \quad (3.1в)$$

$$- P(n_1, n_2, \dots, n_l) = (1 - \rho) m! \prod_{k=1}^{|M|} \frac{\rho_k^{n_k}}{n_k!} \text{ для СМО } \bullet / G / 1, \text{ LCFS}, \quad (3.1г)$$

где $\rho_k = \frac{\lambda_k}{\mu_k}$.

Для СеМО, каждый из узлов которой описывается одним из четырех представленных в (3.1) выражений, справедливо уравнение потоков. Обозначим за N множество узлов сети и поставим в соответствие каждому узлу его номер из диапазона $\overline{1, |N|}$, тогда:

$$\lambda_{i,k} = \lambda'_{i,k} + \sum_{j=1}^{|N|} \sum_{l=1}^{|M|} q_{ji}^{lk} \lambda_{j,l}, \quad i = \overline{1, |N|}; \quad k = \overline{1, |M|}, \quad (3.2)$$

где $\lambda_{i,k}$ – общий входящий поток заявок k -класса в узел i , $\lambda'_{i,k}$ – общий внешний входящий поток заявок k -класса в узел i , q_{ji}^{lk} – вероятность перехода заявки класса l , находящейся на узле j , в заявку класса k на узел i .

Важным этапом моделирования процессов сигнального обмена в конвергентной сети является анализ диаграмм процессов сигнального обмена (рис. 3.1) и составление на их основе модели СеМО.

Опишем процедуру анализа диаграмм обмена сигнальными сообщениями и процесс составления СеМО. Для сохранения общности опишем подход к формализации задачи моделирования процессов сигнального обмена, абстрагируясь от конкретной задачи. На рис 3.1 используются следующие обозначения:

- узлы сети обозначаются A, B, C, D;
- на сети используются два вида протокола передачи сообщений – P1 и P2;
- для организации обмена используются несколько сообщений: M1–M3 для P1 и M1–M2 для P2.

Для перехода от инженерных диаграмм сигнального обмена к СеМО возможно использование подхода, который оказался удобным для решения сложных задач моделирования данной главы диссертации. Идеи формализации задачи моделирования сигнального обмена в конвергентных сетях представлены в ряде работ отечественных и зарубежных авторов [62, 131, 132].

Далее предложен подход к процессу формализации задачи моделирования СеМО конвергентной сети при изучении процессов сигнального обмена. Опишем последовательность шагов в рамках перехода к аналитической модели СеМО.

1) В существующей диаграмме сигнального обмена выделяем главную последовательность (ГП) сообщений. ГП должна начинаться с первого значащего для процесса сигнального сообщения. Окончание ГП должно означать завершение описываемого процесса (например, установление, модификацию или завершение сеанса). Последовательность должна быть конечна, не являться циклически замкнутой и иметь только однозначные переходы без разветвлений (что может наблюдаться в реальной ситуации, когда одно поступившее сообщение порождает несколько). Ситуация, в которой при поступлении на узел сети одного сообщения, генерируется несколько, будет описана при рассмотрении вторичных последовательностей (ВП). Согласно

рис. 3.1 конкретная реализация ГП разыгрывается с некоторыми вероятностями $p_1, p_2, p_1 + p_2 = 1$.

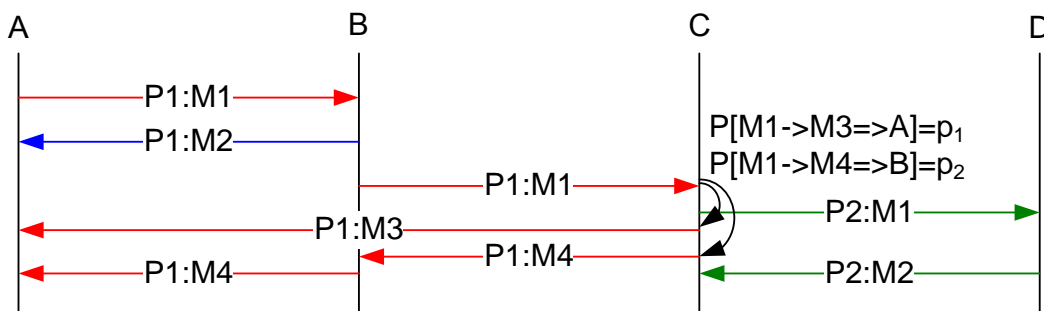


Рис.3.1. Типовая диаграмма обмена сигнальной информацией в конвергентной сети

В типовом случае, представленном на рис. 3.1, ГП будет состоять из сообщений:

$$A:P1:M1:B \rightarrow B:P1:M1:C \rightarrow C:P1:M3:A, \quad (3.3a)$$

$$A:P1:M1:B \rightarrow B:P1:M1:C \rightarrow C:P1:M4:B \rightarrow B:P1:M4:A, \quad (3.3б)$$

где используется следующий порядок записи: узел-отправитель сообщения, протокол, тип сообщения, узел-получатель. В типовом примере процесс начинается с сообщения A:P1:M1:B и заканчивается C:P1:M3:A. На рис. 3.1 ГП выделена красным цветом, ВП – зеленым и синим. Вероятность возникновения процесса, описываемого ГП согласно (3.3a) равна p_1 , (3.3б) – p_2 .

2) Любые последовательности, отличные от ГП, будем называть вторичными. Они, также как и ГП, должны быть конечны, не являться циклическими и иметь однозначные переходы. Данные последовательности – необходимые элементы для описания ситуации, при которой поступление одного сообщения на узел вызывает несколько ответных. На рис. 3.1 ВП являются следующие последовательности:

$$B:P1:M2:A \quad (3.4)$$

$$C:P2:M1:D \rightarrow D:P2:M2:C$$

3) Процедура предполагает приведение к открытой сети ВСМР. Дополним диаграмму на рис. 3.1 виртуальными узлами: S (Source node) –

источник заявок, и Т (Termination node) – узел-сток для заявок, покидающих сеть. С учетом виртуальных узлов, введенных обозначений и правил записи, ГП и ВП имеют вид:

$$\begin{aligned}
 &S:P1:M1:A \rightarrow A:P1:M1:B \rightarrow B:P1:M1:C \rightarrow C:P1:M3:A \rightarrow A:P1:M3:T \\
 &S:P1:M1:A \rightarrow A:P1:M1:B \rightarrow B:P1:M1:C \rightarrow C:P1:M4:B \rightarrow \dots \\
 &\dots \rightarrow B:P1:M4:A \rightarrow A:P1:M4:T \\
 &S:P1:M2:A \rightarrow A:P1:M2:T \\
 &S:P2:M1:D \rightarrow D:P2:M2:C \rightarrow C:P2:M2:T
 \end{aligned}
 \tag{3.5}$$

С учетом виртуальных узлов и последовательностей (3.5) диаграмма (рис. 2.1) может быть записана в виде, представленном на рис. 3.2. Отметим, что при введении расширенного класса узлов сети с включением S и T можно учитывать эти два виртуальных узла как один без потери общности.

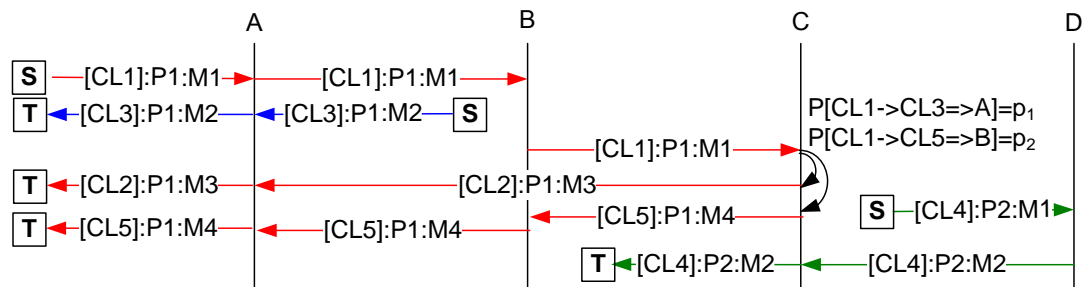


Рис.3.2. Определение потоков заявок и их классов

4) Для каждого типа сообщений назначается соответствующий ему класс заявок. Из рис. 3.2 (классы указаны в квадратных скобках) определено соответствие типов сообщения классам заявок (табл. 3.1).

Табл. 3.1 Классификация сообщений и их типы

Обозначение класса	CL1	CL2	CL3	CL4	CL5
Типы сообщений	S: P1:M1:A A:P1:M1:B B:P1:M1:C	C:P1:M3:A A:P1:M3:T	S:P1:M2:A A:P1:M2:T	S:P2:M1:D D:P2:M2:C C:P2:M2:T	C:P1:M4:B B:P1:M4:A A:P1:M4:T

На рис. 3.2 показаны межклассовые переходы и их вероятности, в случае если вероятность перехода не равна единице. Для обозначения интенсивностей внешних потоков заявок k -класса будем использовать

переменные λ'_k . Оперирова назначенными классами заявок, можно легко перейти к СеМО (рис. 3.3) и определить матрицы переходов для каждого из узлов. На основе данных о логике работы функциональных элементов сети (A, B, C, D), необходимо определить тип соответствующей СМО и ее характеристики.

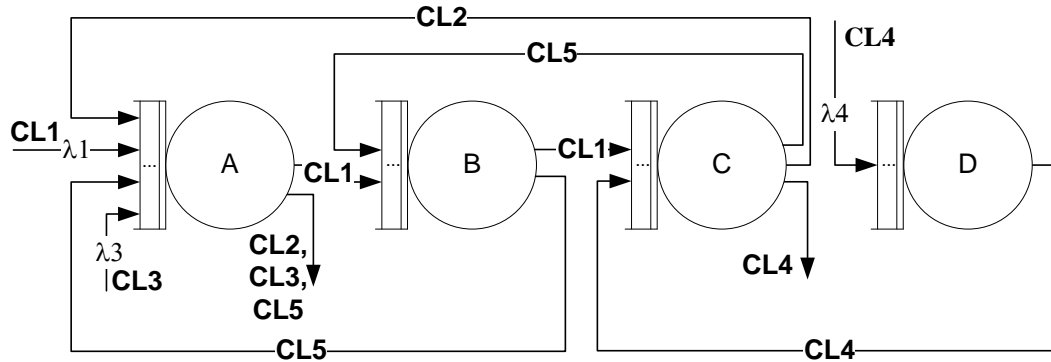


Рис.3.3 СеМО моделируемого процесса

Решение задачи моделирования процессов обмена сигнальной информацией в конвергентной сети с использованием сетей Джексона

Если для задачи моделирования возможны допущения, разрешающие использование в качестве моделей узлов сети СМО $M/M/t$, возможно упрощение процесса моделирования за счет использования открытой сети Джексона [74, 90, 122] вместо ВСМР [70].

В случае, если использование открытой сети Джексона допустимо, то задача существенно упрощается. Вместо одной многоклассовой СеМО типа ВСМР сеть разбивается на несколько независимых сетей Джексона.

Чтобы осуществить переход к сетям Джексона, на этапе разбиения на ГП и ВП необходимо рассмотреть независимую СеМО Джексона для каждой из последовательностей.

Получив для каждой из последовательностей аналитическое решение, обобщение полученных результатов до первоначальной задачи не составляет труда.

3.2. Исследование процессов сигнально обмена в конвергентных сетях на базе СПС 2/3/4-го поколений

Сети подвижной связи четвертого поколения во многом соответствуют принципам построения сетей NGN и подразумевают под собой единую IP архитектуру и соответствующий набор функций. Важно, чтобы процесс перехода на сети четвертого поколения был осуществлен плавно и незаметно для конечного потребителя услуг. В связи с этим, международные организации разработали поэтапный план перехода от мобильных сетей 2-го и 3-го поколений к 4-му. Поскольку в Российской Федерации переход к 4-му поколению будет осуществляться на основе стандарта LTE, будем рассматривать архитектуру сетей LTE и LTE-Advanced [41, 68]. Весь процесс перехода неразрывно связан с конвергенцией мобильных сетей различных поколений, что обусловлено необходимостью сделать процесс перехода максимально комфортным для конечного потребителя услуг.

Процесс построения сети LTE может быть разбит на два этапа:

– первый этап – построение сети 4-го поколения для целей оказания услуги передачи данных. Согласно архитектуре сети на этом этапе (рис. 3.4), в случае необходимости оказания голосовых и других услуг в режиме реального времени, мобильная станция, зарегистрированная в сети LTE использует ресурсы СПС сетей предыдущих поколений с помощью технологии CSFB [84] (см. гл. 1). В случае использования технологии CSFB наличие на сети оператора платформы IMS [87] не является необходимым условием для осуществления голосового вызова.

– второй этап – дооборудование сети платформой IMS и обеспечение возможности организации услуг реального времени без использования ресурсов сетей предыдущих поколений [138].

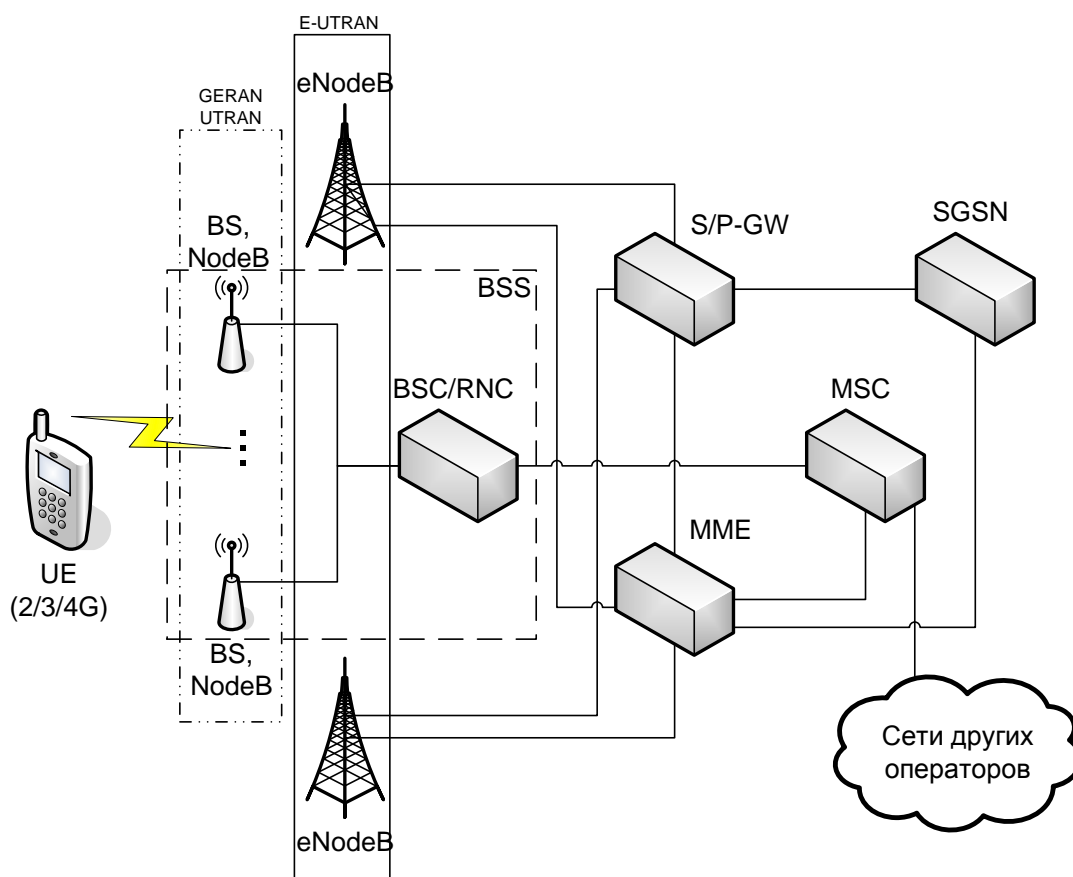


Рис.3.4 Архитектура конвергентной сети 2/3/4-го поколений на первом этапе перехода

На первом этапе одним из ключевых процессов, обеспечивающих возможность осуществления голосовых вызовов для мобильных станций, зарегистрированных в сети 4-го поколения, является процесс, возникающий при использовании CSFB. Оперативность выполнения процедур возврата мобильной станции под управление оборудования сетей предыдущих поколений в случае необходимости оказания услуг реального времени напрямую влияет на время доступа к ним. Для сетей 2-го поколения GSM нормальное время установления голосового соединения (задержка между набором номера абонентом и получения сигнала контроля отправки вызова (КПВ)) составляет 3.58 секунд [61]. На первом этапе внедрения LTE при CSFB, задержка может достигать 6-7.5 с [134], а в случае роуминга и в других сложных случаях – до 15-25 с.

Таким образом, при проектировании мобильных сетей LTE проблема временных задержек, вносимых процедурой CSFB, а также

дополнительная нагрузка, создаваемая сигнальным обменом в момент процедуры, является актуальной, требующей исследования.

Вопросам моделирования процессов в сетях 2-го и 3-го поколений, возникающих на радиоподсистеме, посвящено, например, подробное исследование [9]. В настоящем разделе рассмотрим процессы, возникающие в сети при CSFB [23, 53].

Рассмотрим согласно документам ETSI [84], диаграмму сигнального обмена для процедуры исходящего вызова с учетом CSFB в сеть 2-го поколения, в случае если обеспечение хэндовера пакетного соединения (PS HO, Packet Switched Hand Over) [86] не поддерживается сетью или мобильной станцией. Опуская промежуточные шаги, рассмотрим результат определения классов заявок (рис. 3.5) и таблицу соответствия сообщений классам заявок (Приложение 5). Обозначения и нумерация сигнальных сообщений на рис. 3.5 сохранены согласно рекомендации ETSI [84, ch. 6.3, fig 6.3–1]. Отметим, что рассматриваемая CeMO состоит из N' узлов, по которой циркулируют заявки M' классов, $N' = \overline{1,8}$, $M' = \overline{1,19}$.

Введем также вероятность ошибки при обработке любой заявки. Для учета ошибок введем дополнительный класс CL20, в результате чего множество классов $M = M' \cup \{20\} = \{1, \dots, 19\} \cup \{20\} = \{1, \dots, 20\}$, $|M| = 20$. Учитывая виртуальные узлы S и T полное множество узлов есть $N = N' \cup \{9\} = \{1, \dots, 9\}$, $|N| = 9$ (согласно п.3.1 полное множество узлов содержит только один узел, одновременно описывающий S и T). Отметим, что для описания процессов на радиоподсистеме сети GSM (BSS, Base Station Sybssystem) будем использовать единый узел, отмеченный на рис. 3.5 как BSS без разделения на базовую станцию (БС) и их контроллер.

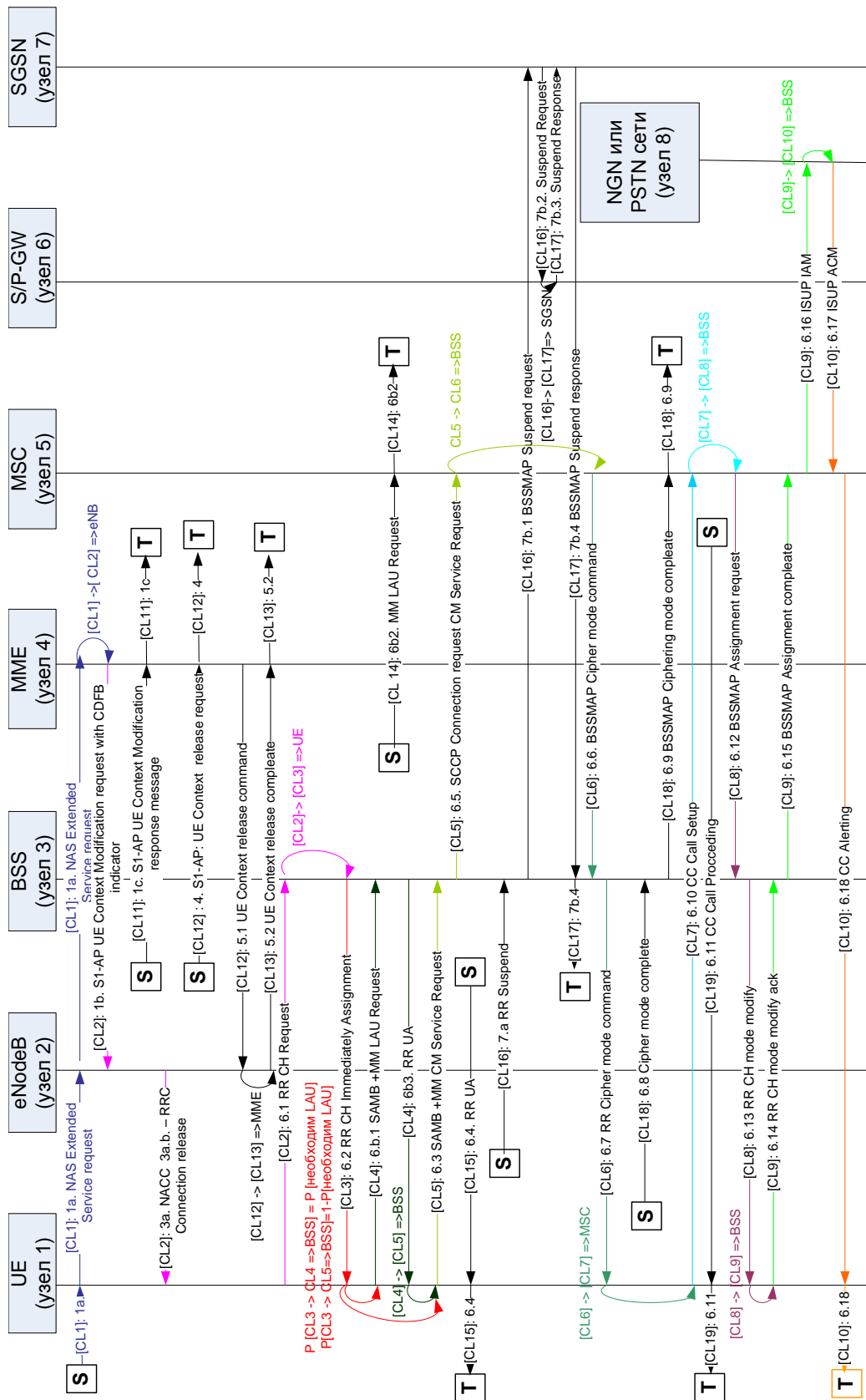


Рис.3.5 Определение потоков заявок и их классов для исходящего вызова абонента сети LTE с использованием CSFB в сеть СПС 2-го поколения GSM без поддержки PS HO

Для перехода к аналитическому решению и нахождению параметров СеМО пронумеруем каждый из узлов, как указано на рис. 3.5. На основе логики рассматриваемого процесса и схем рис. 3.5 определим СеМО (рис. 3.6), типы узлов исследуемой СеМО и интенсивности входящих потоков. Будем считать суммарный входящий поток CL1 заявок пуассоновским с интенсивностью λ_{in} .

С практической точки зрения наиболее актуальным представляется оценка параметров, связанных с процессом установления соединения. Процесс установления соединения здесь характеризуется последовательностью переходов заявок из класса CL1 в CL10, что показано на рис. 3.6. В результате прохождения сети заявка 10 раз меняет свой класс и выходит из сети, имея класс CL20. Согласно рис. 3.6 будем считать началом установления соединения момент, когда абонентское устройство начинает формирование сообщения запроса расширенной услуги (NAS Extended Service request). Окончанием установления соединения является момент, когда абонентское устройство получает и корректно обрабатывает сообщение Alerting.

Рассмотрим следующие варианты СеМО:

1) Вариант 1 показывает случай, когда ресурсы радиоподсистемы ограничены одной БС GSM (BS) и одной БС LTE (eNodeB), зоны радиопокрытия которых совпадают для возможности процедуры CSFB. Данный вариант позволяет оценить параметры работы узлов радиоподсистемы сетей GSM, LTE и отражает ситуацию, связанную с протеканием процессов на оборудовании базовой подсистемы сети GSM и оборудовании eNodeB LTE в случае возрастания нагрузки в пределах одной БС. Здесь и далее предполагается, что на одной базовой станции организуется одна сота независимо от поколения сети, чтобы исключить вопросы, связанные с хэндовером внутри между сотами на одной БС.

Поскольку элементы коммутационного ядра сети, такие как MSC, SGSN, MME, S/P-GW, проектируются для обслуживания запросов от нескольких сотен до нескольких тысяч БС, то, очевидно, что нагрузка,

создаваемая одной БС на данные элементы мала. Основываясь на этом положении, логичным выбором СМО для моделирования БС сетей GSM и LTE (узлы $i = \{2,3\}$ СеМО) является система $\bullet/G/1$ с разделением процессора PS. Остальные узлы СеМО есть бесконечно-линейные СМО $\bullet/G/\infty$ в силу малой нагрузки.

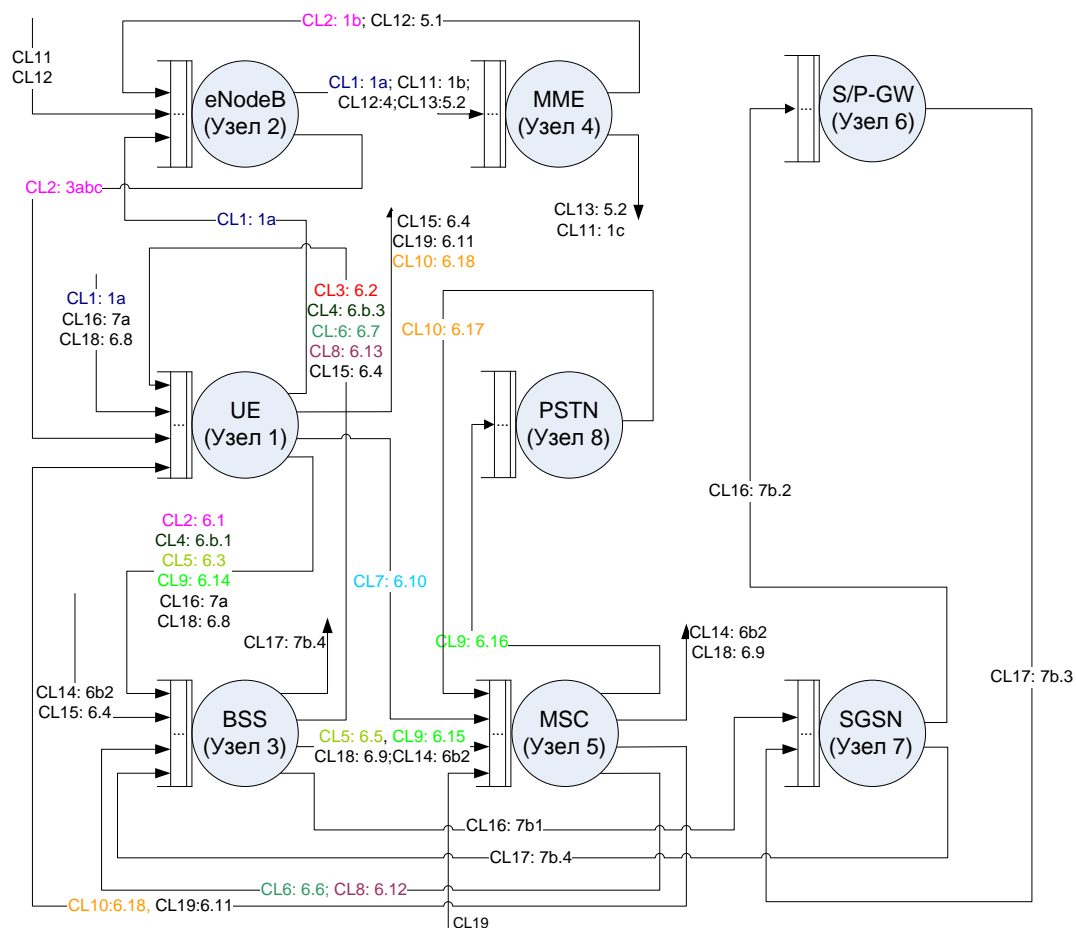


Рис.3.6 СеМО для конвергентной сети 2/4-го поколения и процесс исходящего вызова абонента LTE с использованием CSFB в сеть GSM без поддержки PS HO

2) Вариант 2 описывает ситуацию, в которой на узлы коммутационного ядра сети ($i = \overline{4,7}$) поступают запросы на установление соединения по процедуре CSFB от большого количества БС сетей GSM, LTE. Рассматривается случай, в котором количество БС не ограничено и, соответственно, нагрузка, поступающая от абонентов может возрастать бесконечно и не будет блокирована из-за недостатка ресурсов на

радиоподсистеме. Данный вариант отражает процессы работы элементов ядра сети, в случае, если радиоподсистема не является узким местом. В данном варианте логичным выбором СМО для элементов радиоподсистемы и пользовательского оборудования будет модель бесконечно-линейного сервера $\bullet/G/\infty$. Для моделирования процессов в узлах коммутационного ядра $i = \overline{4,7}$ в силу отсутствия нетерпеливых заявок логичным выбором является СМО $\bullet/G/1$, PS либо $\bullet/M/m$, FCFS. Остановимся на СМО $\bullet/G/1$, PS, т.к. часть ресурсов узлов, такие как входные и выходные сетевые интерфейсы, оперативная память, системная шина, очевидно, являются общими, а их ресурсы будут распределяться для обслуживания входящего потока заявок.

3) Вариант 3 рассматривает условия, аналогичные условиям варианта 2, но с предположением, что все элементы коммутационного ядра есть устройства с модульной архитектурой и некоторым количеством приборов. В данном случае как наиболее естественный выбор в качестве узлов $i = \overline{4,7}$ взяты экспоненциальные FCFS СМО $\bullet/M/m$; для остальных узлов – СМО $\bullet/G/\infty$, IS.

4) Вариант 4 рассматривает случай, когда все элементы сети работают в сильно недогруженном режиме. Данный вариант необходим для оценки минимальных значений задержки при выполнении рассматриваемого процесса. В качестве всех узлов СеМО в этом случае рассматриваются СМО типа $\bullet/G/\infty$, IS.

Обозначим через N_{IS} множество узлов типа IS, через N_{PS} – множество узлов типа PS. Положим, что q_{ij}^{kl} есть вероятность перехода заявки от узла i в узел j с изменением ее класса с k на l , $i, j \in N; k, l \in M$. Матрица $\mathbf{Q} = \{q_{ij}^{kl}\}, i, j \in N; k, l \in M$ будет четырехмерной, размерности $|M|^2 \times |N|^2$. Большинство элементов \mathbf{Q} не являются нулевыми.

При рассмотрении процессов передачи заявки с узла i на узел j все аспекты передачи, связанные с переходом от узла i к узлу j (например, задержки, потери, ошибки при передаче) относить к параметрам узла i .

Описание возможных переходов в условиях рассматриваемого процесса сигнального обмена представлено в Приложении 4.

Запишем уравнения баланса потоков [30] для каждого из узлов сети. Пусть $\lambda_{j,l}$ есть интенсивность входящего потока заявок класса l в узел с номером j . Внешние потоки заявок СеМО (рис. 3.7) определяются формулами:

$$\begin{aligned}
 \lambda'_{1,1} &= \lambda_{in} \\
 \lambda'_{1,16} &= \lambda_{in} \cdot q_{1,2}^{1,1} \cdot q_{2,4}^{1,1} \cdot q_{4,2}^{1,2} \cdot q_{2,1}^{2,2} \cdot q_{1,3}^{2,2} \cdot q_{3,1}^{2,3} \cdot (q_{1,3}^{3,4} \cdot q_{3,1}^{4,4} \cdot q_{1,3}^{4,5} + q_{1,3}^{3,5}) \\
 \lambda'_{1,18} &= \lambda_{in} \cdot q_{1,2}^{1,1} \cdot q_{2,4}^{1,1} \cdot q_{4,2}^{1,2} \cdot q_{2,1}^{2,2} \cdot q_{1,3}^{2,2} \cdot q_{3,1}^{2,3} \cdot (q_{1,3}^{3,4} \cdot q_{3,1}^{4,4} \cdot q_{1,3}^{4,5} + q_{1,3}^{3,5}) \cdot q_{3,5}^{5,5} \cdot q_{5,3}^{5,6} \cdot q_{3,1}^{6,6} \\
 \lambda'_{2,l} &= \lambda_{in} \cdot q_{1,2}^{1,1} \cdot q_{2,4}^{1,1} \cdot q_{4,2}^{1,2}, l = 11, 12 \\
 \lambda'_{3,14} &= \lambda_{in} \cdot q_{1,2}^{1,1} \cdot q_{2,4}^{1,1} \cdot q_{4,2}^{1,2} \cdot q_{2,1}^{2,2} \cdot q_{1,3}^{2,2} \cdot q_{3,1}^{2,3} \cdot q_{1,3}^{3,4} \\
 \lambda'_{3,15} &= \lambda_{in} \cdot q_{1,2}^{1,1} \cdot q_{2,4}^{1,1} \cdot q_{4,2}^{1,2} \cdot q_{2,1}^{2,2} \cdot q_{1,3}^{2,2} \cdot q_{3,1}^{2,3} \cdot (q_{1,3}^{3,4} \cdot q_{3,1}^{4,4} \cdot q_{1,3}^{4,5} + q_{1,3}^{3,5}) \\
 \lambda'_{5,19} &= \lambda_{in} \cdot q_{1,2}^{1,1} \cdot q_{2,4}^{1,1} \cdot q_{4,2}^{1,2} \cdot q_{2,1}^{2,2} \cdot q_{1,3}^{2,2} \cdot q_{3,1}^{2,3} \cdot (q_{1,3}^{3,4} \cdot q_{3,1}^{4,4} \cdot q_{1,3}^{4,5} + q_{1,3}^{3,5}) \cdot q_{3,5}^{5,5} \cdot q_{5,3}^{5,6} \cdot q_{3,1}^{6,6} \cdot q_{1,5}^{6,7}
 \end{aligned} \tag{3.6}$$

Тогда, с учетом выражений (3.2) и (3.6), система уравнений для входящих потоков выглядит следующим образом:

$$\begin{cases} \lambda_{j,l} = \sum_{i,k} \lambda_{i,k} \cdot q_{i,j}^{k,l}, \\ (j,l) \in N \times M \setminus \{(1,1), (1,16), (1,18), (2,11), (2,12), (3,14), (3,15), (5,19)\} \\ \lambda_{j,l} = \lambda'_{j,l} + \sum_{i,k} \lambda_{i,k} \cdot q_{i,j}^{k,l}, \\ (j,l) \in \{(1,1), (1,16), (1,18), (2,11), (2,12), (3,14), (3,15), (5,19)\} \end{cases} \tag{3.7}$$

Введем следующие обозначения:

- $\mu_{i,k}$ – интенсивность обслуживания заявок класса k на узле i ;
- $S = \{n_{ik}\}$ – матрица размерности $|N| \times |M|$ состояния СеМО, где n_{ik}

обозначает количество заявок типа k на узле i , $i = \overline{1, |N|}, k = \overline{1, |M|}$.

Поскольку все типы СМО в составе исследуемой СеМО являются совершенными [5, 27], то для узлов справедливо выражение

$$P(S) = G^{-1} \prod_{i=1}^{|N|} P(S(i, \bullet)),$$

где $P(S(i, \bullet))$ – вероятность нахождения узла i в

состоянии, описываемом i вектором столбцом матрицы S . Тогда, с учетом выражений (3.1), получим для исследуемой СеМО:

$$P(S) = G^{-1} \prod_{i \in N_{IS}} P_{PS}(S(i, \bullet)) \cdot \prod_{i \in N_{IS}} P_{IS}(S(i, \bullet)), \quad (3.8)$$

где $P_{PS}(S(i, \bullet)) = (1 - \rho_i) m_i! \prod_{l \in M} \frac{\rho_{il}^{n_{il}}}{n_{il}!}$, $P_{IS}(S(i, \bullet)) = e^{\rho_i} \prod_{l \in M} \frac{\rho_{il}^{n_{il}}}{n_{il}!}$, $\rho_{il} = \frac{\lambda_{il}}{\mu_{il}}$.

Оценка среднего времени установления соединения и времени ожидания

Для вычисления среднего времени ожидания для каждого узла, используем методику, описанную в [5], где сеть разбивается на отдельные изолированные узлы автономной системы. Используем также методы оценки средних значений, описанные в [25, 72].

Введем обозначения:

– w_{ik} – среднее время ожидания заявки класса k при обработке на узле i ;

– v_{ik} – среднее время отклика узла i при обработке заявок класса k ;

– μ_{ik}^{-1} – среднее время обслуживания заявки класса k при обработке на узле i ;

– μ_i – средняя интенсивность обслуживания заявок (без учета класса) при обработке на узле i , т.е.

$$\mu_i = \frac{\sum_{k \in M} \mu_{ik}}{|M|};$$

– B_i – вероятность того, что вновь прибывшая заявка будет помещена в очередь;

– V_i – суммарная временная задержка в процессе организации исходящего голосового вызова с использованием технологии CSFB, а также с учетом сценария с использованием процедуры LAU;

– V_{II} – суммарная временная задержка в процессе организации исходящего голосового вызова с использованием технологии CSFB в случае, когда LAU не требуется;

– m_i – количество обслуживающих элементов на узле i .

Тогда, для среднего времени отклика справедливо:

$$v_{ik} = \begin{cases} \mu_{ik}^{-1}, & \forall i \in N_{IS}, \forall k \in M, \\ w_{ik} + \mu_{ik}^{-1}, & \forall i \in N \setminus N_{IS}, \forall k \in M. \end{cases} \quad (3.9)$$

Согласно рис. 3.6 запишем выражения для оценки средней длительности процедуры установления соединения (процесс, описываемый главной последовательностью):

$$\begin{aligned} V_I &= v_{1,1} + v_{2,1} + v_{4,1} + v_{2,2} + v_{1,2} + v_{3,2} + v_{1,3} + v_{3,4} + v_{1,4} + v_{3,5} + \dots \\ &+ v_{5,5} + v_{3,6} + v_{1,6} + v_{5,7} + v_{3,8} + v_{1,8} + v_{3,9} + v_{5,9} + v_{8,9} + v_{5,10} + v_{1,10}, \\ V_{II} &= v_{1,1} + v_{2,1} + v_{4,1} + v_{2,2} + v_{1,2} + v_{3,2} + v_{1,3} + v_{3,5} + v_{5,5} + \dots \\ &+ v_{3,6} + v_{1,6} + v_{5,7} + v_{3,8} + v_{1,8} + v_{3,9} + v_{5,9} + v_{8,9} + v_{5,10} + v_{1,10}. \end{aligned} \quad (3.10)$$

Согласно [72], время ожидания прибывающей заявки класса k определяется формулой:

$$w_{ik} = B_{ik} \frac{1}{m_i \mu_{ik}} + \sum_{j=1}^M \overline{n_{ik}} \frac{1}{m_i \mu_{ik}}, \quad (3.11)$$

где $\overline{n_{ik}}$ – среднее число заявок класса k на узле i .

Учитывая закон Литтла, $\overline{n_{ik}} = \lambda_{jk} w_{jk} = \lambda_{jk} w_{ik}$, откуда получаем

$$w_{ik} = B_i \frac{1}{(1 - \rho_i) m_i \mu_{ik}}, \quad \rho_i = \frac{\sum_{k \in M} \lambda_{ik}}{m_i \mu_{ik}}. \quad (3.12)$$

Для узла типа IS, очевидно, время ожидания в очереди равно нулю, что следует из $B_i = 0, \forall i \in N_{IS}$.

Для узлов ВСМР $\bullet/G/1$ с дисциплиной PS:

$$B_i = 1 - P(0) = 1 - (1 - \rho_i) 0! \prod_{k \in M} \frac{\rho_{ik}^0}{0!} = \rho_i. \quad (3.13)$$

С учетом выражения (3.11) и $m_i = 1$ получаем

$$w_{ik} = \frac{\rho_i}{(1 - \rho_i)\mu_{ik}}. \quad (3.14)$$

Проведя аналогичные выкладки для системы $\bullet/M/m$ FCFS, согласно [72], вероятность постановки прибывшей заявки в очередь имеет выражение

$$B_i = \frac{(m_i \rho_i)^{m_i}}{m_i!} \left((1 - \rho_i) \sum_{p=0}^{m_i-1} \frac{(m_i \rho_i)^p}{p!} + \frac{(m_i \rho_i)^{m_i}}{m_i!} \right)^{-1}. \quad (3.15)$$

Учитывая выражение (3.12), получаем для среднего времени ожидания в очереди:

$$w_{ik} = \frac{m_i^{m_i-1} \rho_i^{m_i}}{(1 - \rho_i)\mu_{ik} m_i!} \left((1 - \rho_i) \sum_{p=0}^{m_i-1} \frac{(m_i \rho_i)^p}{p!} + \frac{(m_i \rho_i)^{m_i}}{m_i!} \right)^{-1}. \quad (3.16)$$

Подставляя полученные для w_{ik} выражения (3.16) в (3.10) с учетом (3.9), не сложно получить расчетные формулы для потерь и общей задержки с учетом типа заявок.

Если допустить что для зафиксированного $\forall i \in N$ справедливо $\forall k, l \in M : \mu_{ik} = \mu_{il} = \mu_i$, т.е. интенсивность обслуживания заявок на узле не зависит от их класса, то, с учетом (3.9–3.10), получаем оценку для среднего времени установления соединения:

– для варианта 1 (очевидно, что имеет смысл только значение времени процедуры осуществления исходящего вызова без использования LAU):

$$V_{II} = 6\mu_1^{-1} + 2(\mu_2^{-1} + w_2) + 5(\mu_3^{-1} + w_3) + \mu_4^{-1} + 4\mu_5^{-1} + \mu_8^{-1}, \quad (3.17)$$

– для варианта 2 и 3:

$$\begin{aligned} V_I &= 7\mu_1^{-1} + 2\mu_2^{-1} + 6\mu_3^{-1} + (\mu_4^{-1} + w_4) + 4(\mu_5^{-1} + w_5) + \mu_8^{-1}, \\ V_{II} &= 6\mu_1^{-1} + 2\mu_2^{-1} + 5\mu_3^{-1} + (\mu_4^{-1} + w_4) + 4(\mu_5^{-1} + w_5) + \mu_8^{-1}, \end{aligned} \quad (3.18)$$

– для варианта 4:

$$\begin{aligned} V_I &= 7\mu_1^{-1} + 2\mu_2^{-1} + 6\mu_3^{-1} + \mu_4^{-1} + 4\mu_5^{-1} + \mu_8^{-1}, \\ V_{II} &= 6\mu_1^{-1} + 2\mu_2^{-1} + 5\mu_3^{-1} + \mu_4^{-1} + 4\mu_5^{-1} + \mu_8^{-1}. \end{aligned} \quad (3.19)$$

Основываясь на приведенных выкладках настоящего раздела,

перейдем к численной оценке результатов моделирования и проверки возможности использования предложенных вариантов.

3.3. Численные результаты моделирования и выводы

Пусть $\forall i \in N, \forall k, l \in M : \mu_{ik} = \mu_{il} = \mu_i$. Предположим также, что уровень ошибок на узле не зависит от класса заявки и зависит только от узла сети. Обозначим вероятность успешной обработки заявки на узле i как p_i . Для целей численного анализа построенных СеМО будем рассматривать значения характеристик функционирования узлов согласно табл. 3.2, при выборе значений использовались международные стандарты МСЭ-Т, ETSI, технические отчеты производителей оборудования, опубликованные результаты измерений [67, 79, 134].

Данные о значениях параметров работы узлов $i = \{3,5\}$ в табл. 3.2 взяты из [61, 78, 134], а для узлов $i = \{2,4\}$ – на основе анализа технической документации некоторых производителей телекоммуникационного оборудования. Для узла 1 введена дополнительная задержка, отражающая время отклика программного обеспечения оборудования UE. За среднее время обработки сообщений на узле 8 взято среднее время между моментом посылки SS7 ISUP IAM (либо SIP INVITE) и регистрацией ответного SS7 ISUP ACM (либо SIP 183 Ringing).

Вероятность LAU для СеМО вариантов 2-4 взята $q = 0.2$. Для СеМО варианта 1 вероятность LAU – нулевая (в силу условий). Результаты численных расчетов для параметров w_i, V_i и V_{II} СеМО вариантов 1, 4 представлены на рис. 3.7.

Задержки установления соединения для СеМО варианта 4 $V_i = 6.89$ с и $V_{II} = 5.99$ с (рис. 3.7, при учете выражения (3.19) и табл. 3.2) согласуются с характеристиками реальных сетей [134].

Возможности современных eNodeB и базовых станций 2G GSM позволяют обслуживать до 2 тысяч пользователей (из расчета полосы в 5МГц на соту и нагрузки, создаваемой одним мобильным абонентом в

0.017 Эрл). При средней продолжительности вызова в 120 с, интенсивность вызовов на одну БС есть $\lambda_{in} \approx 0.284$ вызова в секунду. Из рис. 3.7 видно что СеМО варианта 1 неплохо отражает реальную работу узлов $i = \{2,3\}$, показывая переполнение очереди при $\lambda_{in} > 0.3$. Возможности eNodeB по сравнению с пропускной способностью базовой станции сети GSM выше, что также согласуется с численными результатами.

Табл. 3.2 Параметры работы узлов конвергентной сети

i	1	2	3	4	5	6	7	8
μ_i^{-1}, c	0.5	0.25	0.4	0.03	0.04	0.03	0.03	0.3
P_i	0.95	0.99	0.99	0.999	0.999	0.999	0.999	0.95

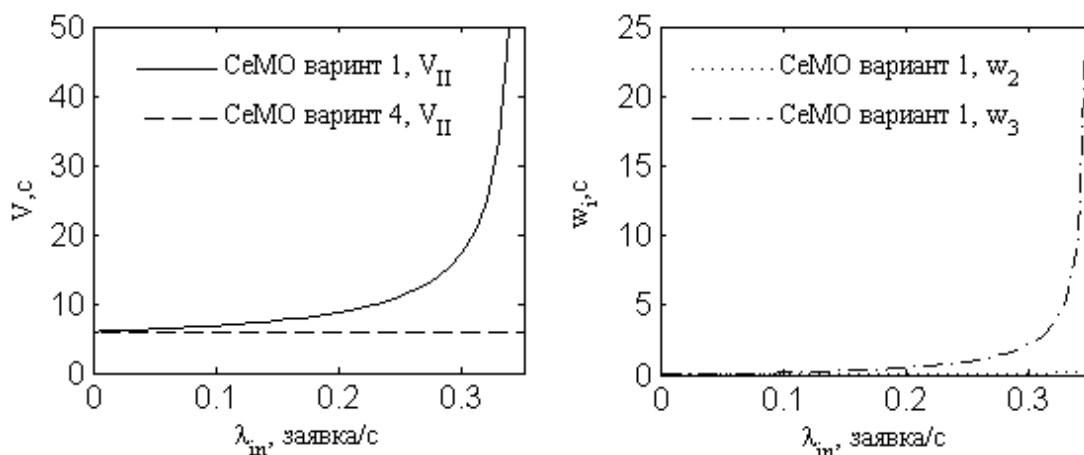


Рис. 3.7 Результаты численного анализа СеМО вариантов 1, 4

Анализируя СеМО варианта 2 (рис. 3.8), можно судить лишь о том, что способ, при котором узлы коммутационного ядра сети моделируются СМО $\bullet/G/1, PS$ плохо соответствует реальным режимам работы узлов. Результаты, полученные на СеМО варианта 2 корректны только лишь при незначительных нагрузках.

Узлы конвергентной сети $i = \overline{4,7}$ в настоящее время делаются из расчета на бoльшую нагрузку, чем показывает рис. 3.8, но при этом демонстрируют характеристики согласно табл. 3.2. Данный факт, а также

принцип модульной архитектуры современных узлов, позволяет предполагать, что в качестве узлов $i = \overline{4,7}$ необходимо брать СМО с более чем одним обслуживающим элементом. Для численного анализа примем значения $m = \{5,10,15\}$. Семейство кривых для каждого из значений количества обслуживающих элементов представлены на рис. 3.9. С увеличением m наблюдается ожидаемое увеличение производительности СеМО и сдвиг начала переполнения очередей в сторону большей нагрузки пропорционально количеству обслуживающих элементов.

Численный анализ СеМО варианта 3 подтверждает, что тип СМО $\bullet/M/m$ для узлов $i = \overline{4,7}$ дает лучшие результаты для более широкого диапазона значений λ_{in} , по сравнению с СеМО варианта 2.

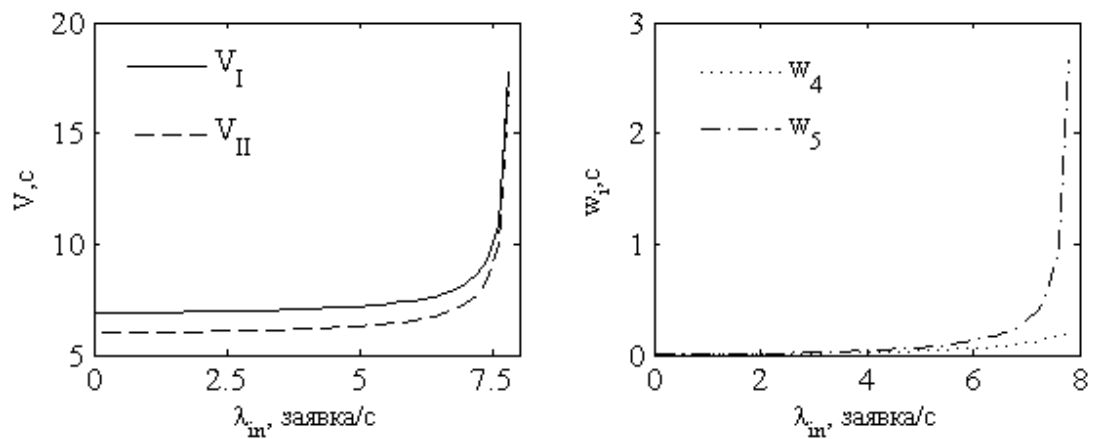


Рис. 3.8 Результаты численного анализа СеМО варианта 2

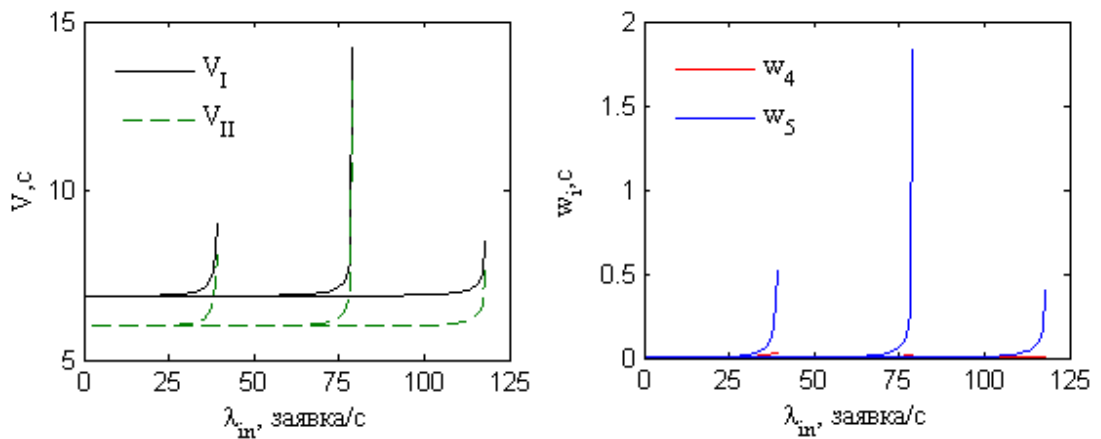


Рис. 3.9 Результаты численного анализа СеМО варианта 3, $m = \{5,10,15\}$

3.4. Решение СеМО для узлов сети типа GI|G|m

Предположение об экспоненциальном распределении времени обслуживания будучи серьезным допущением [2, 66, 77], позволяет проводить инженерные расчеты и получать оценки для определения тенденций поведения и характеристик функционирования реальных объектов. То же относится к предположению, что узлы $i = 2,3$ являются бесконечными СМО типа IS. Реальная сеть может быть спроектирована так, что в одном случае узким местом будет оборудование коммутационного ядра ($i = \overline{4,7}$), а в другом сегменте будут наблюдаться перегрузки на узлах базовой подсистемы ($i = 2,3$). Для учета реальных свойств сети необходимо рассмотреть для $i \in N$ СМО типа $GI/G/m$. Для построения аналитической модели используем подходы построения СеМО [30]. Применяя подход [70] и методику [62], в совокупности с аппроксимационными подходами оценки средних значений характеристик СМО типа $GI/G/m$ [139], используем комбинированный подход для построения СеМО и оценки ВВХ сети.

Как и раньше будем считать суммарный входящий поток CL1 заявок пуассоновским с интенсивностью λ_{in} . Исключим ситуации размножения сообщений (одно сообщение порождает два и более), а также поглощения или объединения сообщений [139]. Выберем последовательности (рис. 3.6) таким образом, чтобы интенсивность суммарного входящего потока заявок была равна суммарной исходящей интенсивности:

$$\sum \lambda_{\text{входящий}} \equiv \sum \lambda_{\text{исходящий}}. \quad (3.20)$$

В рассматриваемой СеМО общего вида предположим для $i = \overline{2,7}$ есть СМО типа $GI/G/m$, а узлы $i = \{1,8\}$ – $GI/G/\infty$. В дополнение к вариантам 1-4 СеМО (п. 3.2) приведем расчеты параметров для СеМО общего вида, а затем подробно рассмотрим частные варианты общего вида СеМО:

1) Вариант I описывает сеть в условиях варианта 1 (п. 3.2), однако в качестве узлов $i = \{2,3\}$ взяты СМО типа $GI/G/1$ с дисциплиной FCFS; остальные узлы есть бесконечно-линейные СМО $GI/G/\infty$.

2) Вариант II описывает сеть в условиях варианта 2 (п. 3.2), но в качестве узлов $i = \overline{4,7}$ взяты СМО $GI/G/m_i$; для остальных узлов – СМО $GI/G/\infty$.

3) Вариант III отражает работу сети в сильно недогруженном режиме – в качестве всех узлов СеМО рассматриваются СМО типа $GI/G/\infty$.

Уравнения баланса в данной модификации СеМО не будут отличаться от (3.6), (3.7) в силу условия (3.20).

Учитывая общий характер распределения времени поступления заявок и их обработки, в дополнение к системе (3.7) необходимо определить коэффициенты вариации для потоков. Применим технику декомпозиции [5], аналогично п. 3.3 настоящей диссертации. Также будем предполагать равнозначность классов.

Введем новые и переопределим уже введенные в п. 3.2 обозначения:

– $c_{a,i}^2$ – квадрат коэффициента вариации интервалов времени между поступлением заявок на узел i ;

– $c_{a,ji}^2$ – квадрат коэффициента вариации интервалов времени между поступлением заявок с узла j на узел i ;

– $c_{s,i}^2$ – квадрат коэффициента вариации времени обслуживания заявок в узле i ;

– $c_{d,ij}^2$ – квадрат коэффициента вариации временных интервалов между заявками потока на выходе узла i по направлению к узлу j ;

– $c_{d,i}^2$ – квадрат коэффициента вариации временных интервалов между заявками потока на выходе узла i ;

– N_{IS} – множество узлов, моделируемых СМО типа $GI/G/\infty$;

– N_m – множество узлов, моделируемых СМО типа $GI/G/m_i$, где натуральное m_i такое, что $m_i < \infty$.

Тогда, с учетом предположения равнозначности классов сети, примем для $\forall i \in N$ и $\forall k, l \in M$: $\mu_{ik} = \mu_{il} = \mu_i$. Обозначим матрицу переходов для случая отсутствия классов как $Q' = \{q_{ij}\}$, где $q_{ik} = \sum_k \sum_l q_{ij}^{kl}$.

Будем предполагать справедливость следующих утверждений:

- для $\forall i \in N$: $\rho_i = \frac{\lambda_i}{m_i \mu_i} < 1$ – отсутствие переполнения узлов СеМО;
- $\sum_i \lambda_{ij} = \sum_i \lambda_{ji}, \forall i, j \in N \setminus \{9\}$ – условие (3.20);
- $\lambda_{ij} = \lambda_i q_{ik}$.

Учитывая общий характер входящих потоков заявок и распределений времени обслуживания на узлах СеМО, для определения потоков необходимо использовать не только средние значения интенсивностей входящих заявок и времени обслуживания, но коэффициенты вариации. Для описания задачи, необходимо определять коэффициенты вариации для [139, п. 4.3–4.5]

– общего потока, состоящего из нескольких

$$c_{a,i}^2 = \sum_{j \in N} \left(\frac{\lambda_{ji}}{\sum_{r \in N} \lambda_{ri}} c_{a,ji}^2 \right) = \frac{1}{\lambda_i} \left(\sum_{j \in N} \lambda_{ji} c_{a,ji}^2 \right), \forall i \in N; \quad (3.21)$$

– исходящего потока на выходе из узла СеМО

$$c_{d,i}^2 = \begin{cases} c_{a,i}^2, & i \in N_{is} \\ 1 + (1 - \rho_i^2)(c_{a,i}^2 - 1) + \frac{\rho_i^2}{\sqrt{m_i}}(c_{s,i}^2 - 1), & i \in N_m \end{cases}; \quad (3.22)$$

– известной части общего исходящего потока, определяемой вероятностью q_{ij}

$$c_{d,ij}^2 = q_{ij} c_{d,i}^2 + 1 - q_{ij}, \quad \forall i, j \in \overline{1,8}. \quad (3.23)$$

Отметим, что $\forall i, j \in N$: $c_{d,ij}^2 = c_{a,ij}^2$. Подставляя выражения (3.22–

3.23) в (3.21), можно записать систему линейных уравнений для общего вида СеМО

$$\begin{aligned} c_{a,i}^2 \lambda_i - \sum_{j \in N_m} \lambda_{ji} q_{ji} (1 - \rho_j^2) (c_{a,j}^2 - 1) = \\ = \sum_{j \in N_m} \lambda_{ji} \left\{ \frac{q_{ji} \rho_j^2}{\sqrt{m_j}} (c_{s,j}^2 - 1) + 1 \right\} + \sum_{j \in N_{IS}} \lambda_{ji} \{ q_{ji} (c_{a,j}^2 - 1) + 1 \}. \end{aligned} \quad (3.24)$$

Решение систем (3.7) и (3.24) позволит полностью описать потоки в исследуемой СеМО. Определив параметры потоков, перейдем к ВВХ СеМО.

Аппроксимация временных характеристик

В соответствии с рис. 3.5, 3.6 и с учетом (3.10), запишем выражения для оценки средней длительности процедуры установления соединения в условиях предположения о равноправности классов:

$$\begin{aligned} V_I &= 7v_1 + 2v_2 + 6v_3 + v_4 + 4v_5 + v_8, \\ V_{II} &= 6v_1 + 2v_2 + 5v_3 + v_4 + 4v_5 + v_8. \end{aligned} \quad (3.25)$$

Для определения средней длительности процедуры установления соединения необходимо определить w_i . Согласно [139, п. 5.2] для СМО типа $GI / G / m$ при $\rho \rightarrow 1$ ожидаемое среднее w_i может быть аппроксимировано ожидаемым средним w_i СМО с входящими пуассоновскими потоками $\bullet / M / m$ с нормировочным коэффициентом

$$w_i(c_{a,i}^2, c_{s,i}^2, m_i / GI / G / m_i) = \left(\frac{c_{a,i}^2 + c_{s,i}^2}{2} \right) w_i(M / M / m_i). \quad (3.26)$$

Для оценки параметров СМО типа $GI / G / 1$ используем подход, предложенный в [62] и применим аппроксимацию Крамера и Лангенбаха-Белтца [124]:

$$w_i = \frac{\rho_i (c_{a,i}^2 + c_{s,i}^2)}{2\mu_i (1 - \rho_i)} g(\rho_i, c_{a,i}, c_{s,i}), \quad (3.27)$$

где

$$g(\rho_i, c_{a,i}, c_{s,i}) = \begin{cases} \exp\left\{-\frac{2(1-\rho_i)(1-c_{a,i}^2)^2}{3\rho_i(c_{a,i}^2+c_{s,i}^2)}\right\}, c_{a,i}^2 < 1, \\ \exp\left\{-\frac{(1-\rho_i)(c_{a,i}^2-1)^2}{c_{a,i}^2+4c_{s,i}^2}\right\}, c_{a,i}^2 \geq 1. \end{cases}$$

Оценку для w_i , с учетом (3.16), получим в следующем виде:

$$w_i = \frac{1}{\lambda_i} \sum_{k \in M} \lambda_{ik} w_{ik}. \quad (3.28)$$

Используя параметры потоков, полученные при решении систем (3.9) и (3.24), выражение (3.27) и выражение (3.26) с учетом (3.16), можно получить приближенные оценки средних значений времени ожидания w_i , $i \in N$.

Используя выражение (3.25) с учетом (3.9) для различных вариантов СеМО запишем выражения для среднего времени установления соединения:

– общий вид СеМО (3.29):

$$\begin{aligned} V_I &= 7\mu_1^{-1} + 2(\mu_2^{-1} + w_2) + 6(\mu_3^{-1} + w_3) + (\mu_4^{-1} + w_4) + 4(\mu_5^{-1} + w_5) + \mu_8^{-1}, \\ V_{II} &= 6\mu_1^{-1} + 2(\mu_2^{-1} + w_2) + 5(\mu_3^{-1} + w_3) + (\mu_4^{-1} + w_4) + 4(\mu_5^{-1} + w_5) + \mu_8^{-1}; \end{aligned}$$

– вариант I (очевидно, что имеет смысл только значение времени процедуры осуществления исходящего вызова без использования LAU):

$$V_{II} = 6\mu_1^{-1} + 2(\mu_2^{-1} + w_2) + 5(\mu_3^{-1} + w_3) + \mu_4^{-1} + 4\mu_5^{-1} + \mu_8^{-1}; \quad (3.30)$$

– вариант II:

$$\begin{aligned} V_I &= 7\mu_1^{-1} + 2\mu_2^{-1} + 6\mu_3^{-1} + (\mu_4^{-1} + w_4) + 4(\mu_5^{-1} + w_5) + \mu_8^{-1}, \\ V_{II} &= 6\mu_1^{-1} + 2\mu_2^{-1} + 5\mu_3^{-1} + (\mu_4^{-1} + w_4) + 4(\mu_5^{-1} + w_5) + \mu_8^{-1}; \end{aligned} \quad (3.31)$$

– для варианта III:

$$\begin{aligned} V_I &= 7\mu_1^{-1} + 2\mu_2^{-1} + 6\mu_3^{-1} + \mu_4^{-1} + 4\mu_5^{-1} + \mu_8^{-1}, \\ V_{II} &= 6\mu_1^{-1} + 2\mu_2^{-1} + 5\mu_3^{-1} + \mu_4^{-1} + 4\mu_5^{-1} + \mu_8^{-1}. \end{aligned} \quad (3.32)$$

Основываясь на приведенных выкладках настоящего раздела, перейдем к численной оценке результатов моделирования и проверке применимости предложенных вариантов.

Численные результаты моделирования

Для целей численного анализа построенных СеМО примем характеристики функционирования узлов согласно табл. 3.2. Примем также что $c_{s,i}^2 = c_{s,j}^2 = c_s^2$ для $\forall i, j \in N$.

Для СеМО варианта I вероятность LAU нулевая (в силу предположений варианта I). Результаты численных расчетов для параметров V_{II} , w_i , $i = 2,3$ представлены на рис. 3.10.

Учитывая (3.26–3.27) заметим, что значения V_{II} при $c_s^2 = 1$, не отличаются от значений, полученных в п. 3.3 при оценке характеристик процесса установления соединения для экспоненциальных вариантов СеМО. В случае, если $c_s^2 \neq 1$, очевидно, значения V_I , V_{II} будут отличны от полученных в п. 3.3. Можно наблюдать, как при уменьшающихся коэффициентах вариации форма кривых зависимости V_{II} и w_i , $i = 2,3$ начинает стремиться к сети, узлы которой описываются детерминированными СМО $D/D/1$. При возрастании коэффициентов вариации разброс значений временных промежутков между моментами регистрации входящих сообщений и времени их обработки увеличивается, что сглаживает зависимости на рис. 3.10.

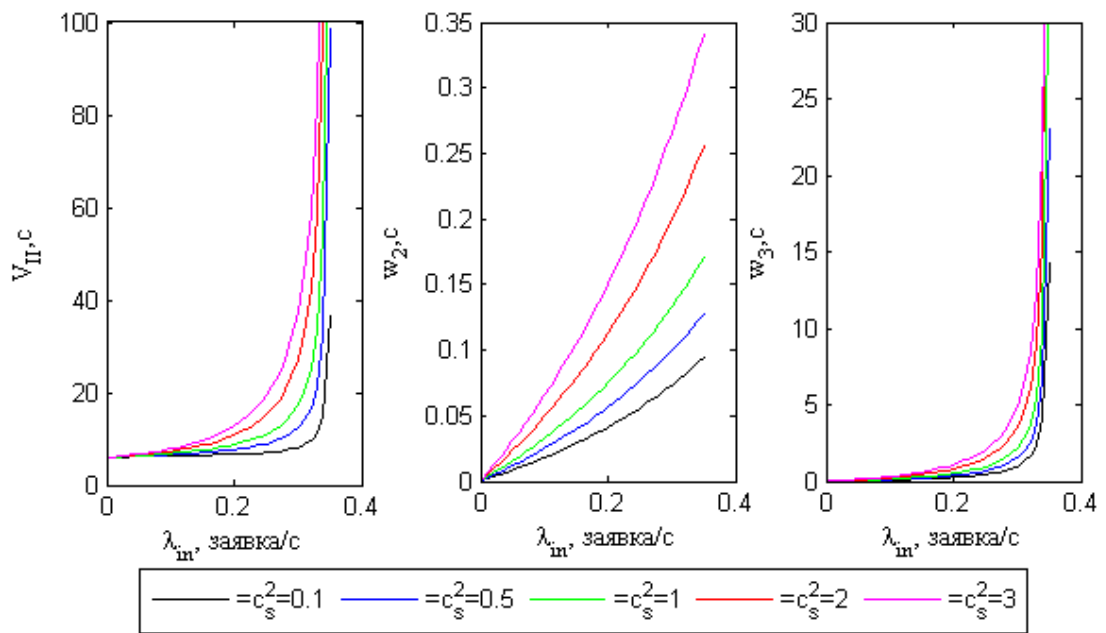


Рис. 3.10 Зависимости V_{II} и $w_i, i = \{2,3\}$ от интенсивности входящей нагрузки λ_{in} для СеМО варианта I

Характер зависимости задержки при установлении соединения и времени ожидания при высоких показателях c_s , говорит о негативном влиянии роста c_s на показатели работы сети. Таким образом, при конструировании сетевых элементов, по возможности, необходимо внедрять механизмы снижения значений c_s и c_a .

Результаты, представленные на рис. 3.10, хорошо согласуются с оценками п. 3.3 и отражают изменения при различных вариантах коэффициента вариации времени обслуживания.

Для варианта II СеМО приведем результаты расчетов, аналогично условиям п. 3.3, взяв $\forall i, j = \overline{4,7}: m_i = m_j$, и $m_i = \{5,10,15\}$. На рис. 3.11 хорошо видны семейства кривых, соответствующих различным числам m на узле.

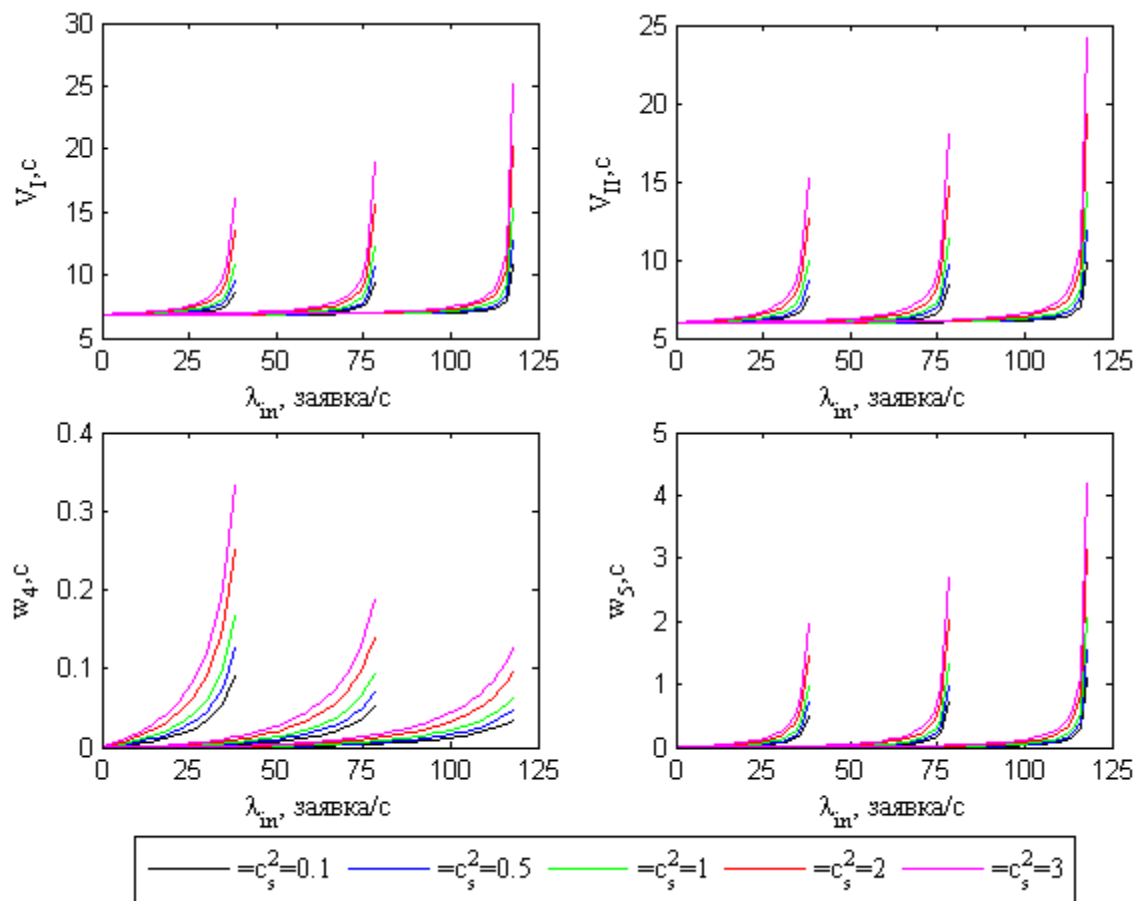


Рис. 3.11 Результаты численного моделирования СеМО вариант II

Как и в СеМО ВСМР, отметим, что для описания многопроцессорных элементов конвергентной сети, таких как MSC и MME целесообразно использовать СМО с несколькими приборами, чтобы обеспечить необходимую производительности [23, 53].

3.5. Выводы

1. Формализована процедура моделирования процессов обмена сигнальной информацией в конвергентной сети. Описан процесс перехода от диаграмм сигнального обмена к СеМО на основе аппарата ТМО сетей ВСМР.

2. Предложена и исследована модель СеМО конвергентной сети 2G/4G и проведен анализ временных характеристик процессов установления исходящих голосовых вызовов с использованием технологии CSFB.

3. Разработаны несколько вариантов выбора СМО ВСМР сети для моделирования узлов конвергентной сети. Для моделирования процессов на радиоподсистеме с целью анализа характеристик базовой станции 2G/4G рекомендуется использовать модели узлов $\bullet/G/1$, для остальных элементов сети – СМО IS. При моделировании процессов на коммутационных элементах конвергентной сети (MSC, MME, S-/P-GW, SGSN) рекомендуется использовать модель СМО $\bullet/M/m$ FCFS.

4. В рамках усложнения модели конвергентной сети и перехода от пуассоновского потока к потоку общего вида приведены результаты средних временных характеристик работы сети с узлами $GI/G/m$. Предложена схема расчета временных параметров системы на основе приближений Крамера и Лангенбаха-Белтца – для узлов радиоподсистемы, моделируемых СМО $GI/G/1$, и Вайтга – для коммутационных элементов конвергентной сети, моделируемых СМО $GI/G/m$. Аналитическая модель СеМО может быть использована для расчетов и оптимизации архитектуры реальных конвергентных сетей при наличии потоков сигнальных сообщений не пуассоновского характера.

ГЛАВА 4

ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ СИГНАЛЬНОГО ОБМЕНА НА МАГИСТРАЛИ КОНВЕРГЕНТНОЙ СЕТИ

В настоящей главе изложены результаты исследования магистральных участков сети конвергентного оператора связи, использующих сигнальные протоколы SIPv2.0 и SIP-I [2, 19, 65-66]. Необходимость анализа потоков сигнальных сообщений протокола SIP и его подвидов (SIP-I/-T) а также экспериментальное определение параметров работы устройств P-I-CSC обусловлена необходимостью получения актуальных данных для проверки результатов аналитического моделирования, полученных в главах 2, 3 диссертации.

В разделе 4.1 рассматриваются процесс и возможные проблемы организации экспериментального изучения сетевого трафика с магистралей конвергентной сети. Исследуется вопросы организации сбора, хранения и обработки массивов сигнальной информации. Данное исследование является необходимым этапом в рамках организации экспериментального исследования свойств сетевого трафика конвергентной сети и режимов работы ее фрагментов.

В разделе 4.2 приведены результаты исследования потока сообщений протокола SIPv2.0 на магистрали между сетью агрегации вызовов и справочным контактном центром одного из клиентов оператора, использовавшего сеть конвергентного оператора связи в качестве сети агрегации вызовов от потенциальных клиентов.

4.1. Исследование проблем экспериментального изучения магистральных участков конвергентной сети

Данный раздел посвящен инженерно-техническим аспектам организации сбора данных о трафике на действующей сети. Затронуты

аспекты, связанные с идентификацией пакетов при прохождении их через сетевые устройства различных уровней, а также вопросы первичной обработки данных. Особое внимание уделено методам измерения задержки передачи информации.

Задержка передачи информации по сети связи – один из ключевых показателей, на основе которого можно делать выводы о качестве функционирования сети и, отчасти, о качестве предоставляемых на сети услуг. Особенно важным он является для потоковых видов трафика, возникающих при предоставлении аудио или видео услуг в режиме реального времени. Предоставление услуг связи в большинстве случаев сопряжено с передачей сигнальных и пользовательских данных в режиме реального времени. В отличие от эластичного трафика передачи данных, качество услуг связи сильно зависит от величины задержки при передаче. Пакеты сигнальной и пользовательской информации, потоки которых передаются в сети при организации услуг связи, должны соответствовать определенным требованиям, в частности, нормам на задержки, несоблюдение которых приводит как к искажению передаваемой аудио и/или видеoinформации, так и к снижению воспринимаемого пользователями качества услуг (QoE, Quality of Experience) [32, 38].

Как уже упоминалось выше, одним из важных показателей качества услуг телефонной связи в конвергентных сетях является время установления соединения. В случае возникновения больших задержек при передаче сигнального трафика по сети возможны проблемы с установлением сеансов (например, аварийное прерывание, «зависание»), что существенным образом влияет на QoE пользователя. Задержка передачи сигнального трафика складывается из задержек его передачи по каналам связи (propagation delay) и задержек обработки сигнальных сообщений в оборудовании (processing delay).

Для изучения реальных свойств пакетного трафика, режимов работы оборудования, сети, и вносимых ею задержек необходимы

системы сбора и обработки данных. Подобные системы должны производить анализ данных в режиме реального или отложенного времени и не оказывать влияния на работу сети. Далее кратко рассматриваются варианты организации пассивных и активных систем мониторинга сети, методы захвата и обработки данных. Предложены варианты решения типовых проблем в процессе организации подобных систем, таких как:

- синхронизация временных меток при захвате данных несколькими устройствами или от нескольких источников;
- переполнение очередей буферных накопителей на сетевых интерфейсных картах;
- сбор и хранение больших объемов данных в схеме с отложенной обработкой информации;
- методы идентификации сообщений после прохождения их через сетевой элемент либо сегмент сети.

Классификация систем сбора данных

Анализируя типовые решения по организации систем сбора данных в пакетных сетях, можно выделить два основных подхода к классификации [54]:

- по типу влияния системы на существующую сеть (активные и пассивные системы);
- по режиму обработки данных (реального времени и с отложенной обработкой данных).

Остановимся на характеристиках систем более подробно:

1) Активные – системы сбора данных, предполагающие изменение топологии сети и внедрение измерительных устройств или устройств сбора данных. Системы такого рода обладают весомым преимуществом – возможностью сбора данных непосредственно с изучаемого сегмента сети или узла, не искаженных коммутационным оборудованием, используемым для зеркалирования трафика (рис. 4.1). Однако, имеют место и ряд недостатков:

– для построения подобной системы требуется внедрение нового активного оборудования на сети;

– в случае изучения реального коммерческого трафика возникает вероятность отказа системы сбора данных, который может повлечь за собой отказ одного из сегментов сети. В силу обозначенных выше проблем с организацией и эксплуатацией подобного рода систем, внедрение их на сетях с коммерческим трафиком сильно затруднено и, в связи с этим, не использовалось для получения результатов, представленных далее в п.4.2–4.3 настоящей главы.

2) Пассивные – системы сбора данных, предполагающие установку дополнительных узлов и зеркалирование потоков данных с изучаемого участка сети. Пассивные схемы сбора данных более широко используются для изучения трафика в коммерческих сетях. Рассмотрим основные схемы организации пассивных систем:

2.1) Схема с организацией зеркалирования трафика на одном из коммутационных узлов сети представлена на рис. 4.2.

Данная схема организации позволяет анализировать трафик, проходящий через коммутационный узел, в том числе, оценивать его объемы и структуру. Данная схема часто используется в коммерческих сетях для организации систем мониторинга и оценки качества работы сети и ее узлов. Однако, она не позволяет оценить временные задержки, возникающие на сетевом элементе или участке сети.

2.2) Схема с организацией пассивной системы сбора данных с тестируемого узла (рис. 4.3).

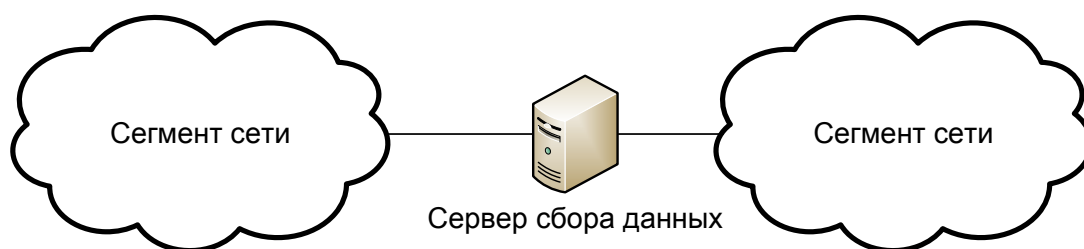


Рис. 4.1. Организация активной системы сбора данных

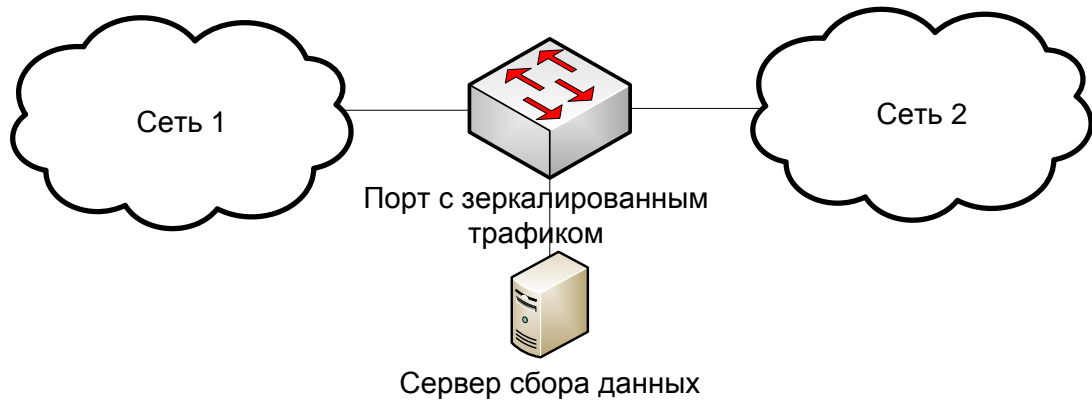


Рис. 4.2. Организация пассивной системы сбора данных с зеркалированием на основе одного коммутационного элемента

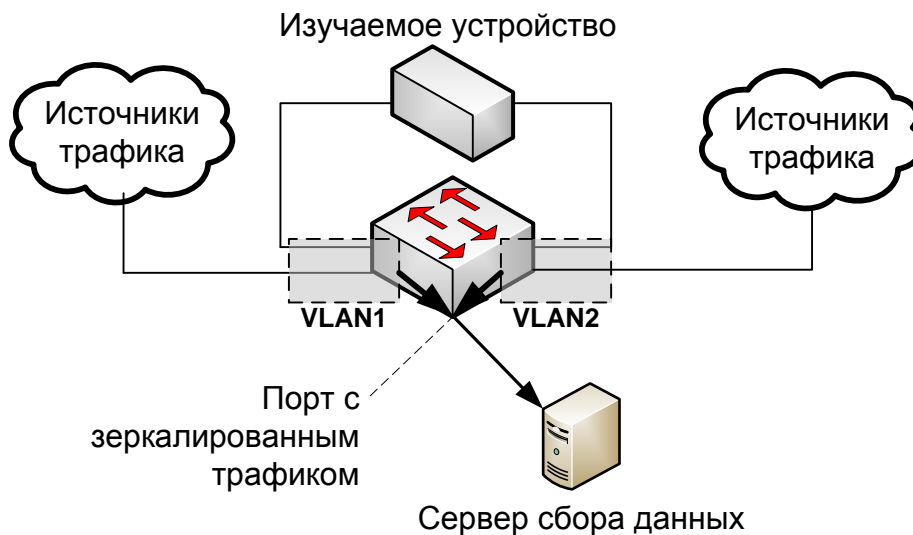


Рис. 4.3. Организация пассивной системы сбора данных с тестируемого узла

На основе данной схемы возможен анализ как входящих, так и исходящих потоков трафика, анализ особенностей работы тестируемого устройства в условиях реальной или тестовой нагрузки. В данной схеме на трафик оказывает влияние тестируемое устройство, но не система сбора данных. При организации подобной системы сбора данных возможен не только анализ качественных и количественных показателей проходящего через тестируемый узел или сегмент сети трафика, но и анализ влияния на проходящий трафик. Данная схема является оптимальной для изучения временных характеристик работы

тестируемого узла или сегмента сети, поскольку не возникает проблемы синхронизации временных меток.

2.3) Схема организации пассивной системы сбора данных с агрегированием заркалированного трафика на едином сервере сбора данных представлена на рис. 4.4. Данная схема позволяет собирать трафик с различных участков сети, подключенных к различным коммутационным устройствам. При создании подобных схем необходимо учитывать влияние на задержку коммутационного оборудования и линий связи на пути передачи трафика от места зеркалирования до сервера регистрации данных. Данная схема позволяет обойти проблемы синхронизации за счет использования единого сервера регистрации данных.

2.4) Схема пассивной распределенной системы сбора данных (рис. 4.5) – применяется, обычно, для исследования сети как единого целого, или ее разрозненных сегментов и представляет собой распределенную систему сбора данных на сети. В такой схеме данные собираются с нескольких узловых элементов сети и регистрируются на различных серверах. В связи с этим возникает проблема синхронизации на серверах при регистрации данных. Данная проблема может быть решена стандартными средствами синхронизации времени в IP-сетях.

Распределенные системы регистрации данных обычно применяются на больших пакетных сетях с целью оперативной оценки качества предоставляемых услуг и мониторинга возможных аварий. Распределенные системы часто представляют собой наложенную сеть, по которой на центральные серверы обработки данных приходит информация от удаленных серверов и центров регистрации данных.

Подобные системы сбора и анализа данных обычно работают в режиме квази-реального времени (обработка данных с небольшой задержкой). Величина этой задержки обычно определяется вычислительными мощностями центральных серверов обработки

данных, а также длиной временного интервала, за который вычисляются статистические показатели качества работы сети и сетевых элементов.

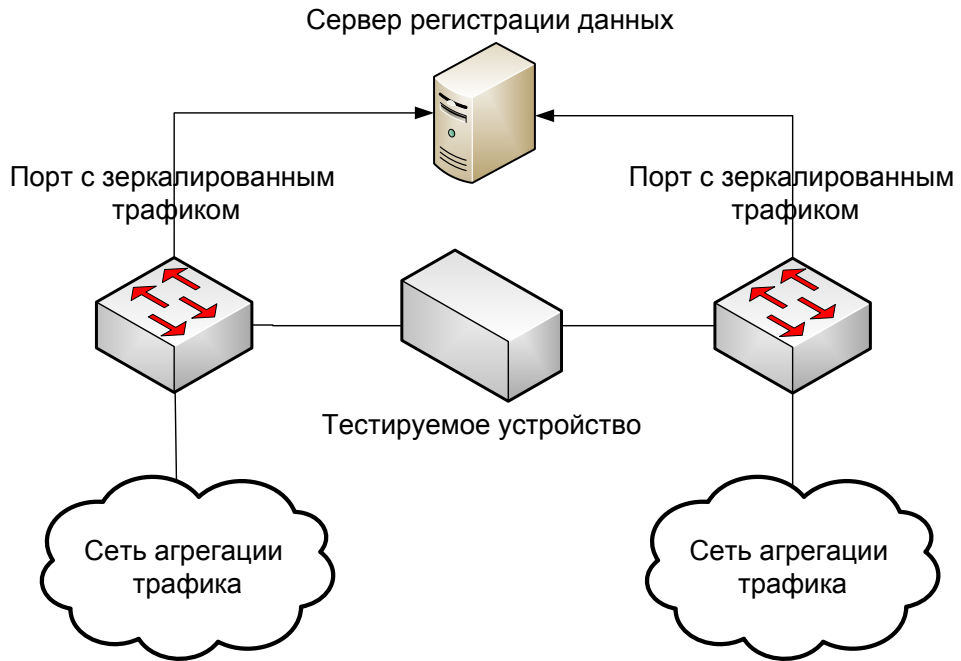


Рис. 4.4. Организация пассивной системы с агрегированием трафика на едином сервере сбора данных

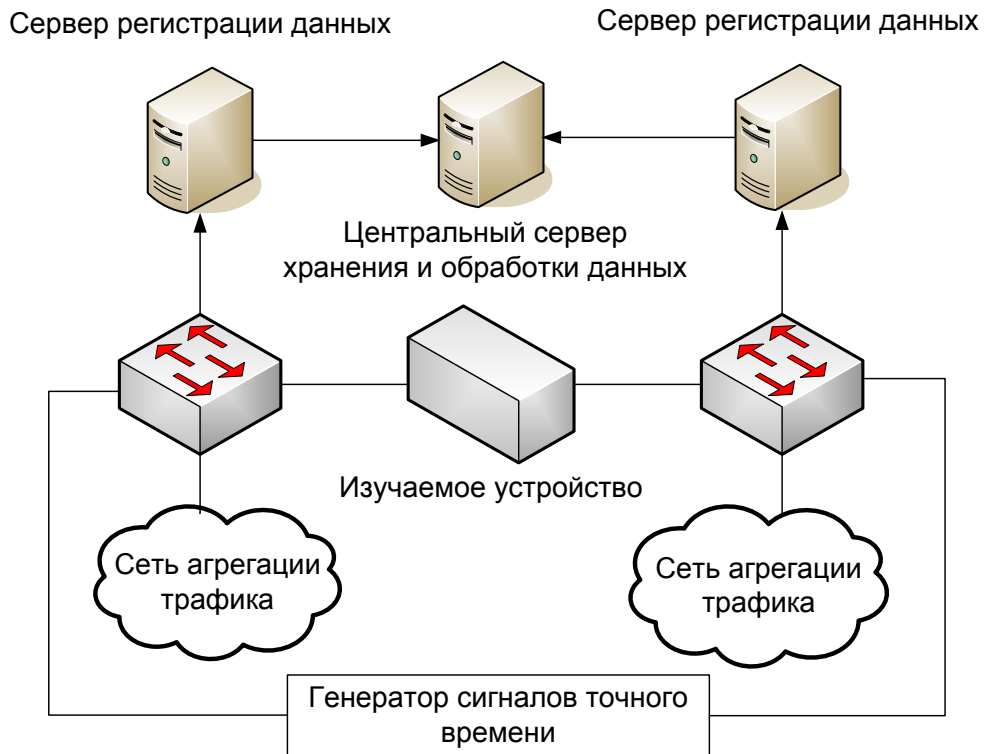


Рис. 4.5. Организация пассивной распределенной системы с агрегированием трафика на едином сервере сбора данных

Синхронизация времени на устройствах системы сбора данных

Одной из основных проблем при построении распределенных систем сбора данных является проблемы синхронизации времени на элементах системы с заданной точностью. Для большинства задач связанных с мониторингом качества предоставляемых услуг вполне достаточно точности, обеспечиваемой протоколом синхронизации сетевого времени (NTP, Network Time Protocol) [107]. В тех случаях, когда необходимо проводить измерения временных задержек, вносимых сегментом сети необходимы более точные способы синхронизации времени. В случае необходимости высокоточных измерений необходимо также учитывать, что точность большинства кварцевых генераторов импульсов, устанавливаемых на компьютерах и серверах, не велика и колеблется в пределах 10^{-3} - 10^{-4} с.

Рассмотрим основные способы и источники для обеспечения временной синхронизации:

- использование протокола NTP;
- использование спутниковых систем точного времени (например, ГЛОНАСС либо GPS (Global Positioning System));
- использование синхронизации от оптических сетей;
- использование первичного эталонного генератора.

По результатам тестирования систем, в которых синхронизацию времени выполняли средствами протокола NTP, абсолютная погрешность измерений составляла порядка 10^{-3} с. Такие значения погрешности при синхронизации через протокол NTP накладывают серьезные ограничения на круг измерений, которые можно проводить при помощи подобных систем. Проверка точности синхронизации по протоколу NTP проводилась на тестовом стенде, оборудованном источниками сигналов GPS с синхронизацией через ежесекундные импульсы (PPS, Pulse Per Second), обеспечивающей точность измерений до 10^{-6} с. Техническая реализация сервера точного времени с синхронизацией от GPS описана в [47].

В виду проблем с точностью при синхронизации по NTP, из доступных решений остаются следующие:

- организация систем сбора данных на основе одного сервера регистрации;

- организация распределенных систем сбора данных, каждый из серверов которой подключен к спутниковому источнику сигналов точного времени (через системы ГЛОНАС или GPS) либо к генератору сигналов точного времени.

Точность синхронизации спутниковых систем вполне достаточна для изучения процессов обработки пакетов или проведения другого рода исследований задержек при передаче информации в пакетных сетях на втором и более высоких уровнях модели OSI [116]. Предельную оценку снизу можно получить на основе данных о задержках при обработке данных устройством уровня L2 – они имеют порядок 10^{-5} с [75].

Кроме использования протокола NTP и спутниковых систем, источником сигнала точного времени может служить оптическая сеть связи (например, синхронная оптическая сеть SDH [38]) либо первичный эталонный генератор. Данные решения обеспечивают высокую точность синхронизации, но не всегда имеются на сети оператора.

Особенности выбора оборудования для организации систем сбора данных

При организации систем сбора данных и, в особенности, при использовании их как инструмента исследования задержек на сети необходимо учитывать следующее:

- влияние на задержку оборудования, расположенного на пути от точки зеркалирования трафика до его записи на сервере сбора данных;

- необходимо удостовериться в том, что оборудование, через которое проходит зеркалированный трафик, нагружено незначительно или же величина нагрузки не влияет существенным образом на вклад, вносимый в задержку, возникающую при обработке пакета этим устройством;

– необходимо удостовериться, что на всем пути следования зеркалированного трафика, пакеты этого трафика не отбрасываются ни одним из устройств;

– необходимо убедиться, что интерфейсные сетевые платы сервера регистрации данных обладают достаточной пропускной способностью для обработки входящего потока. При выборе серверного оборудования необходимо учитывать не только скорость входящего потока, но и его структуру – средний размер одного пакета, т.к. данный фактор существенно влияет на производительность интерфейсных плат [17];

– в случае необходимости организации системы обработки зеркалированного потока в режиме реального времени, необходимо убедиться в достаточности вычислительных ресурсов.

Методы идентификации пакетов

При прохождении пакета через сетевое устройство или несколько сетевых устройств, информация, содержащаяся в заголовках различных уровней может быть изменена. В процессе исследования вопросов, связанных с изучением задержек на оборудовании различных уровней всегда возникает вопрос об однозначной идентификации пакета после его прохождения через сегмент сети или сетевое устройство (рис. 4.6). Одним из методов решения задачи идентификации пакета является определение его уникального идентификатора на основе неизменяемой на исследуемом сегменте сети части. Пакеты регистрируются с обеих сторон от исследуемого сегмента или узла (исследуемый объект может иметь более двух точек регистрации трафика – предложенный метод легко обобщить для такого случая). На основе неизменяемой части зарегистрированного пакета вычисляется его уникальный идентификатор – хэш-функция, которая вместе со временем регистрации заносится в одну из двух таблиц в соответствии с тем, на какой стороне от исследуемого объекта был зарегистрирован пакет. После этого проводится поиск совпадений хэш-функций по таблицам, и для

совпавших значений вычисляются задержки путем определения абсолютной разницы значений соответствующих моментов регистрации.

Ситуация усложняется, если необходимо исследовать задержки на устройствах, работающих на уровне приложений и изменяющих не только всю служебную информацию, но и сами передаваемые данные. Примером такого рода устройств могут служить контроллер пограничных взаимодействий IBC и шлюзы TrGW. Устройство IBC функционирует как прокси-сервер SIP и может изменять не только заголовки L2-L4, но и сами данные, т.е. поля SIP-сообщений. Тот же эффект наблюдается и с TrGW, в функции которого входит обработка потоков пользовательских данных и, часто, их перекодирование.

В таких сложных ситуациях идентификация пакетов может производиться путем анализа неизменяемых частей пакета. Так, для исследования потоков SIP сообщений и режимов работы SIP-прокси устройств было разработано программное обеспечение [43].

В п.4.2-4.4 настоящей главы представлены результаты измерений, полученных с применением пассивных систем сбора данных. Для проведенных экспериментов каждый рассмотренный случай будет описываться более подробно.

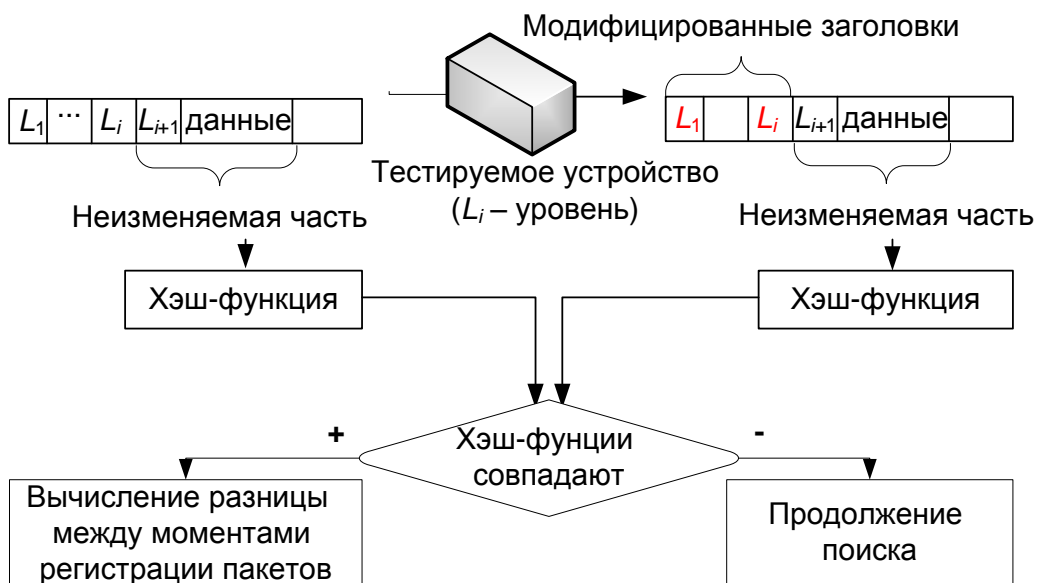


Рис. 4.6. Идентификация пакета по неизменяемой части

4.2. Исследование трафика и структуры потоков сигнальных SIP сообщений на магистральном канале между сетями агрегации и центром обработки вызов

В главе 3 рассматривалась СеМО с учетом общего вида потоков сообщений в ней. В данном разделе приведены результаты экспериментального исследования потоков сигнальных сообщений конвергентной сети, подтверждающего наличие в конвергентных сетях потоков не пуассоновского характера [2, 66]. В связи с растущей актуальностью услуг VoIP большой популярностью пользуются решения по организации контактных центров (КЦ) на основе пакетных сетей [128]. Наиболее распространенным и открытым с точки зрения международных стандартов решением, в части организации сигнального обмена для КЦ, является протокол SIP. Изучению параметров работы КЦ с применением аппарата ТМО посвящен ряд публикаций в отечественных и зарубежных изданиях [16, 24, 76], в то время как особенности потоков сообщений SIP изучены мало. В п. 4.3 представлены результаты исследований автора [2], целью которых являлось определение характеристик SIP трафика КЦ, имеющего распределенную архитектуру сети доступа, на основе данных измерений, проведенных на сети крупного оператора связи. Для сбора статистики и ее обработки применена методика, предложенная в работах [3, 17, 34, 75]. Результаты анализа [2], представленные в настоящем разделе, могут быть использованы для прогнозирования нагрузки на КЦ, аппроксимации роста нагрузки на оборудование, а также при создании генераторов сигнальной нагрузки. Далее приводится подробное описание методики сбора и получения данных измерений, затем на основе результатов обработки полученных данных рассматривается ряд характеристик потока SIP-сообщений.

Средства сбора информации и цели исследования

Рассмотрим схему распределенной сети КЦ оператора услуг связи, представленную на рис. 4.7. Исследовался поток сообщений SIP,

поступающий на один из транзитных узлов КЦ (транзитный SIP-прокси узел).

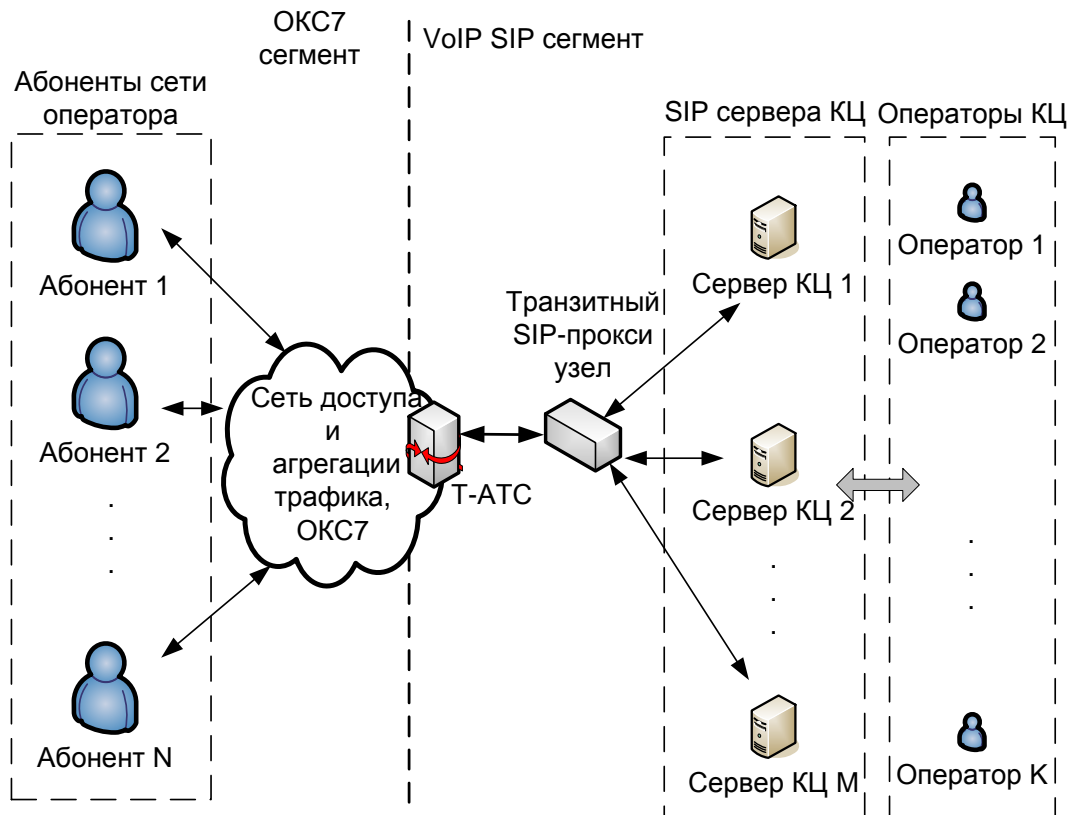


Рис. 4.7. Схема прохождения сигнального SIP-трафика между абонентами и КЦ

Сетью доступа и агрегации трафика является сеть оператора связи, используя ресурсы которой абоненты могут обратиться к операторам КЦ по единому многоканальному номеру или группе номеров. Согласно правилам маршрутизации вызовы, поступающие от абонентов, отправляются на транзитный SIP-прокси сервер, который является точкой агрегации сигнального трафика для КЦ. Транзитный SIP-прокси сервер, согласно предустановленным правилам, маршрутизирует звонки на сервера КЦ, которые обеспечивают дальнейший процесс обслуживания входящего вызова (например, проигрывание автоматического сообщения, постановка вызова в очередь, передача вызова на оператора).

Поток сигнального трафика между КЦ и сетью агрегации состоит из последовательностей сигнальных сообщений, сильно отличающихся от стандартной процедуры установления соединения (RFC 3665 [102]), за счет участия в ней нескольких элементов на стороне КЦ, между которыми происходят последовательные переключения вызовов. Автоматическая процедура обработки вызова и переключения речевых потоков между отдельными устройствами на стороне КЦ, а также возможные переводы вызовов между операторами могут увеличивать число сигнальных сообщений в рамках одного сеанса до 30 и более. Данный факт, не характерный для обычного SIP соединения «точка-точка», должен быть учтен при расчетах нагрузки и выборе оборудования при организации КЦ.

Сбор сигнальной информации на сети производится на основе пассивной схемы сбора и обработки информации № 2.1 (п. 4.1, рис. 4.2). Сервер сбора информации находился под управлением ОС Linux, сбор данных производился стандартной утилитой tcpdump в открытом формате libpcap.

В качестве сетевых интерфейсных модулей использовали стандартное оборудование операторского класса [13], позволяющее организовать запись пакетов без потерь на уровне сетевого интерфейса [31, 96].

Первичная обработка libpcap файлов с SIP трафиком производилось с помощью стандартных возможностей программы tshark. Автоматизация процесса обработки данных была произведена с помощью shell-скрипта консольной оболочки bash операционной системы Linux. В результате, после первичной обработки libpcap файлов с трафиком был получен массив данных, содержащий данные о пакетах SIP, необходимых для дальнейшего исследования. Последующая обработка производилась в пакете прикладных программ MATLAB.

В рассматриваемой сети размер SIP сообщений ограничен сверху максимально возможной длиной Ethernet кадра в 1542 байт [93].

Сегментированные SIP сообщения были исключены из рассмотрения, т.к. их доля не превышала 0,1%. Данные файлов libpcap, содержащие полную информацию обо всех SIP пакетах за период измерений, позволяли определить интенсивность поступления SIP сообщений, их тип, размер, для каждого зарегистрировать время поступления.

Объекты измерения

На рис. 4.7 изображена схема подключения IP сети КЦ, работающего по протоколу SIP, к сети агрегации трафика, построенной по технологии TDM и использующей протокол ОКС 7. На границе сетей TDM и IP находится транзитная автоматическая телефонная станция (Т-АТС), работающая на основе технологии программной коммутации и позволяющие пропускать трафик из сегмента ОКС 7 в SIP и обратно. На Т-АТС со стороны сети агрегации трафика поступают вызовы на номера КЦ, которые транслируются в SIP сегмент на транзитный SIP-прокси сервер. Сервер SIP-прокси в свою очередь обеспечивает агрегацию трафика от Т-АТС и маршрутизацию на один из серверов КЦ в соответствии с заранее заданными правилами маршрутизации.

Данные для исследования были получены путем записи SIP трафика, циркулирующего между Т-АТС и SIP-прокси сервером в течение 14 суток (с 00:00 10.09 по 00:00 24.09). Поскольку время продолжительности соединения в сети было ограничено одним часом, при расчетах было сделано предположение, что трафик имеет зависимости на периоде, не превышающем максимально возможную длительность соединения. Также имеются колебания нагрузки в течение суток. Основываясь на данных предположениях и утверждениях, можно говорить о том, что период наблюдения и собранные данные достаточны для учета всех зависимостей и характеристик исследуемого потока сообщений.

На данном участке трафик является двунаправленным. Входящий трафик со стороны сети агрегации составил около 3700 Эрл в ЧНН, зависимость количества входящих вызовов от времени наблюдения

представлена на рис. 4.8. Из графика видно, что в течение суток нагрузка начиная с 8:00 резко возрастает и достигает максимального уровня примерно к 10:00. Падение нагрузки происходит резко, начиная с 20:00, и достигает уровня ночного минимума уже около 23:00.

В период с 10:00 до 20:00 наблюдаются слабые колебания нагрузки, в которых можно выделить два пика ЧНН. С 13 до 15 часов (обеденные часы) нагрузка находится немного ниже уровня ЧНН. График изменения сигнальной нагрузки в течение одного дня в период с 10:00 до 20:00 показан на рис. 4.8. Из графика видно, что в данный период изменения в характере поступающей нагрузки минимальны.

Модель сигнального трафика

На основе данных измерений, собранных с исследуемого участка сети, была сформирована табл. 4.1, содержащая информацию о встречающихся в SIP трафике типах сообщений и их процентном содержании в общем потоке. Отметим, что соотношение числа сообщений INVITE, проходящих в направлении от сети агрегации в сторону КЦ приблизительно в 7-8 раз меньше количества INVITE, проходящих в обратном направлении. Данный факт объясняется большим числом переключений точек терминации RTP-потоков, происходящих на стороне КЦ, в процессе которых генерируются сообщения re-INVITE с новыми параметрами сеанса.

На основе полученных статистических данных был проведен анализ возможных сценариев установления соединения, детальные результаты анализа представлены в работе [2, табл. 2], где приведены схемы сценариев (последовательность SIP сообщений, которые были переданы для установления сеанса), процентный вес которых в общей массе вызовов был не менее 0.1 %. Поскольку сценарии вызовов сложные и предполагают частые пересогласования параметров сеанса, то количество сообщений INVITE не связано на прямую с количеством вызовов.

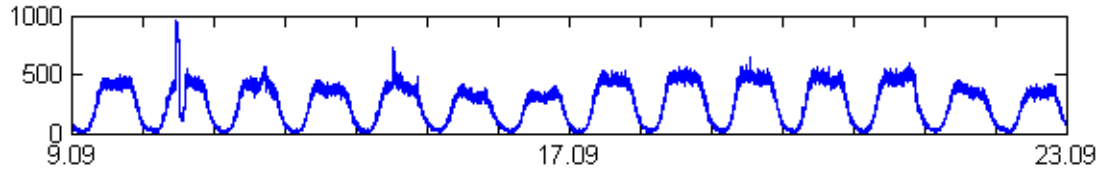


Рис. 4.8. Количество входящих вызовов (14 сут.)

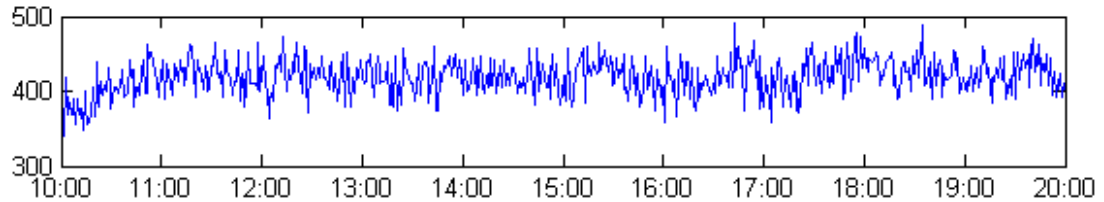


Рис. 4.9. Изменение нагрузки период с 10:00 до 20:00

Табл. 4.1. Процент SIP сообщений по типам и направлениям (за 14 сут)

Тип	Кол-во сообщений на направлении, шт.		Средний размер сообщений на направлении, байт		Процент от общего числа сообщений на направлении, %	
	к КЦ	от КЦ	к КЦ	от КЦ	к КЦ	от КЦ
ACK	251242	1696386	472.53	532.49	5.6564	38.5495
BYE	199561	87983	503.92	489.76	4.4928	1.9994
INVITE	251494	1697586	1216.20	991.75	5.6620	38.5767
CANCEL	829	3856	442.39	451.81	0.0187	0.0876
INFO	95	3815	504.39	599.77	0.0021	0.0867
NOTIFY	32	0	665.59		0.0007	0
REFER	0	16		749.69	0	0.0004
UPDATE	29	64	597.52	813.23	0.0007	0.0015
100	1697607	229508	408.52	384.86	38.2193	5.2154
180	17039	229040	637.06	585.68	0.3836	5.2048
181	28575	0	588.60		0.6433	0
183	202841	0	1393.65		4.5667	0
200	1635525	449919	1145.06	808.38	36.8216	10.2242
202	16	0	486.69		0.0004	0
400	64	63	418.63	385.06	0.0014	0.0014
404	1737	2	570.27	406.50	0.0391	0
408	0	474		439.00	0	0.0108
415	0	14		465.00	0	0.0003
480	98025	0	571.07		2.2069	0
481	1610	758	440.88	469.49	0.0362	0.0172
482	27	0	423.00		0.0006	0
483	1	1	394.00	503.00	0	0
484	69	0	580.06		0.0016	0
486	4898	0	565.64		0.1103	0
487	2377	657	522.56	549.83	0.0535	0.0149
488	661	319	491.14	470.19	0.0149	0.0072
491	40919	8	491.14	495.75	0.9212	0.0002
500	6238	63	516.44	455.84	0.1404	0.0014
502	241	0	513.56		0	0
503	0	11		473.27	0	0.0002
Всего	4441752	4400543				

На основе данных результатов статистического анализа, можно разбить сценарии вызовов на следующие группы [2]:

- успешное установление вызова с ответом, общий вес 58.24 % ;
- успешное установление вызова на занятого оператора, общий вес 1 %;
- успешное завершение соединения без установления разговорного сеанса, связанное с неверным набором номера абонентом, общий вес 22.64 %;
- завершение соединения со стороны абонента до установления соединения с оператором или автоматическим информатором, общий вес 0.4 %;
- завершение соединения, связанное с ошибками в процессе согласования параметров сеанса, неожиданной реакцией оборудования и иного рода ошибками, общий вес 0.85 %.
- другие сценарии установления соединения (каждый из которых в процентном соотношении занимает не более 0.1 %), остались за рамками рассмотрения исследования, общий вес 16.87 %.

На основе полученных данных были проведены исследования характеристик потока сигнальных SIP сообщений. Исследуемая выборка была составлена на основе наблюдений за моментами прихода SIP сообщений в направлении к КЦ и от КЦ в течение временного периода равного 8 часам (с 10 по 18 в течение одних суток). Объем выборки (общее число SIP сообщений) составил около 4.4 млн. в каждом из направлений. Оценки для среднего времени, среднеквадратического отклонения и коэффициента вариации времени между двумя последовательными SIP сообщениями и между двумя последовательными INVITE сообщениями представлены в табл. 4.2, доверительные интервалы $\alpha = 0,95$ вычислены bootstrap-методом [129], вследствие отсутствия свойства нормальности исследуемой выборки [2]. Для построения доверительных интервалов использовалось 1000 bootstrap-выборок, что является рекомендуемым значением [129], при

котором наблюдается тенденция к сходимости определяемой характеристики. В [66] показано, что аппроксимирующим распределением для потоков сигнальных сообщений рассмотренного выше типа может служить управляемый Марковским процессом поступающий поток.

Табл. 4.2. Характеристики потока сообщений

	Для всех типов SIP сообщений в направлении		Для сообщений INVITE в направлении	
	к КЦ	от КЦ	к КЦ	От КЦ
Общее число сообщений, шт.	4441752	4400543	251494	1697586
Граничные точки доверительного интервала $\alpha = 0.95$ для среднего время между сообщениями, с	0.00797-0.00826	0.00808-0.00829	0.1345-0.1498	0.0171-0.0292
Граничные точки доверительного интервала $\alpha = 0.95$ для среднеквадратического отклонения, с	0.0050-0.0063	0.0047-0.0062	0.1361-0.1504	0.0202-0.0231
Граничные точки доверительного интервала $\alpha = 0.95$ для коэффициента вариации	0.7201-0.7273	0.713-0.737	0.9982-1.0140	1.0004-1.0737

Результаты анализа могут быть использованы для прогнозирования нагрузки на КЦ, аппроксимации роста нагрузки на оборудование, при создании генераторов сигнальной нагрузки. В результате исследования были определен ряд характеристик потока SIP сообщений, основные сценарии установления вызовов и их особенности, что может быть использовано при создании аналитических и имитационных моделей, описывающих процессы на серверах SIP-проки и элементах КЦ.

4.3. Выводы

1. Рассмотрен процесс и возможные проблемы организации экспериментального изучения сетевого трафика с магистралей конвергентной сети. Предложены варианты решения типовых проблем в процессе организации систем сбора и обработки трафика. Исследованы вопросы организации сбора, хранения и обработки массивов сигнальной информации. Предложен алгоритм идентификации сообщения по его неизменяемой части при прохождении сетевого узла для целей определения времени его обработки. Результаты анализа могут быть

использованы при организации систем изучения потоков сигнальных сообщений, задержек при их обработке на узлах и фрагментах действующих конвергентных сетей.

2. Получены результаты исследования потока SIPv2.0 пакетов на магистрали между сетью агрегации вызовов и справочным контактном центром одного из клиентов оператора, использовавшего сеть конвергентного оператора связи в качестве сети агрегации вызовов от потенциальных клиентов, что позволило экспериментально подтвердить наличие в конвергентных сетях потоков высокой интенсивности не пуассоновского характера. Данные результаты наряду с другими исследованиями структуры потоков сигнальных сообщений подтверждают необходимость построения СеМО конвергентной сети для потоков заявок общего типа.

ЗАКЛЮЧЕНИЕ

В работе рассмотрены принципы функционирования конвергентных сетей, ее элементы и потоки сигнальной информации. Были определены основные элементы и фрагменты, участвующие в обработке потоков сигнальных сообщений, предложены их аналитические модели. Разработана СеМО конвергентной сети и процедура определения средних временных параметров установления соединения в ней. Статистический анализ экспериментально полученных данных о сигнальных потоках конвергентной сети позволил определить основные их характеристики для различного типа услуг, позволил оценить корректность построенных аналитических моделей.

В диссертации получены следующие результаты:

1. Предложена и исследована аналитическая модель фрагмента конвергентной сети MSC/VLR/HLR, получено решение, которое позволило определить вероятности потери сигнального сообщения в сети на элементах комбинированного узла. На основе полученных зависимостей вероятности потерь сообщений входящего потока определяется максимально допустимый поток при заданном значении вероятности сброса.

2. Предложена и исследована аналитическая модель узла на базе устройств P-I-CSC с двумя аппаратными БН конечной емкости, наличием заявок с абсолютным приоритетом и пороговым механизмом регулирования процесса обслуживания. На основании полученных результатов определено значение порогового параметра r'_1 , при котором вероятность простоя СМО максимальна при небольшом возрастании вероятности потерь.

3. Сформулирована процедура формализации задачи моделирования процессов сигнального обмена в конвергентной сети. С

помощью данного подхода построена СеМО ВСМР конвергентной сети и исследован процесс организации исходящего голосового вызова из сети LTE с применением технологии CSFB. Даны рекомендации по выбору СМО для моделирования основных узлов конвергентной сети. Проведена оценка средних значений времени установления соединения и времени ожидания в очереди для каждого узла конвергентной сети, участвующего в процессе организации голосового вызова. Проведено усложнение СеМО: в качестве СМО сети взяты системы общего вида $GI/G/1$, $GI/G/m$ и $GI/G/\infty$ и получены значения средних времени установления соединения и времени ожидания на узле. Для обоих случаев СеМО также получены результаты численных расчетов для типовых характеристик конвергентной сети.

4. Проведено исследование проблем, возникающих при организации систем записи трафика с конвергентной сети, находящихся в коммерческом использовании, определены основные схемы, отмечены их особенности и даны рекомендации к применению.

5. Исследованы потоки сигнальных SIP сообщений в точке их агрегации между абонентскими сетями и КЦ. Определены основные статистические характеристики потоков сообщений, установлены сценарии установления вызовов и их процентное содержание в общей массе, что позволило подтвердить наличие в конвергентных сетях потоков сигнальных сообщений высокой интенсивности не пуассоновского характера.

СПИСОК ИСТОЧНИКОВ

1. Абаев П.О., Гайдамака Ю.В., Самуйлов К.Е. Гистерезисное управление сигнальной нагрузкой в сети SIP-серверов // Вестник Российского университета дружбы народов. «Математика. Информатика. Физика». – М.:РУДН, 2011. – № 4. – С. 54-71.
2. Абаев П.О., Разумчик Р. В., Углов И.В. Анализ модели SIP-трафика контактного центра по результатам обработки данных сетевых измерений // Т-Comm – Телекоммуникации и Транспорт. – № 11. – 2013. – С. 4-10.
3. Андрианов Г.А., Самуйлов К.Е., Гайдамака Ю.В. Анализ модели трафика ОКС 7 по результатам обработки статистики измерений // Вестник связи. – №11. – 2007. – С. 17-23.
4. Бакланов И.Г., Сондак С.Ю. Концепция «мягкого TMN» и роль измерительной техники в системах управления // Вестник связи. - 2002, №8. – С. 46-48.
5. Башарин Г. П., Толмачев А. Л. Теория сетей массового обслуживания и ее приложения к анализу информационно-вычислительных систем / Итоги науки и техники, серия «Теория вероятностей. Математическая статистика. Теоретическая кибернетика». – М.: ВИНТИ, 1983. – №21 – С. 3-119.
6. Башарин Г.П. Лекции по математической теории телетрафика: Учеб. пособие. – Изд. 2-е испр. и доп. – М.:РУДН, 2007. – 268 с.
7. Башарин Г.П., Бочаров П.П., Коган Я.А. Анализ очередей в вычислительных сетях. Теория и методы расчета. – М.: Наука, 1989. – 336 с.
8. Башарин Г.П., Самуйлов К.Е. Об оптимальной структуре буферной памяти в сетях передачи данных с коммутацией пакетов. – М.: ВИНТИ Препринт НСК АН СССР, 1982. – 70 с.

9. Гайдамака Ю.В., Зарипова Э.Р., Самуйлов К.Е. Модели обслуживания вызовов в сети сотовой подвижной связи: Учебно-метод. пособие. – М.: РУДН, 2008. – 72 с.
10. Голдстейн М., Голдстейн И. Как мы познаем. Исследование процесса научного познания / Сокр. пер. с англ. А. Е. Петрова. — М.: Знание, 1984. — 256 с.
11. Гольдштейн Б.С. Сигнализация в сетях связи. – М.: Радио и связи, 1997. – 423 с.
12. Гольдштейн Б.С., Пинчук А.В., Суховицкий А.Л. IP-телефония. – М.: Радио и связь, 2001. – 336 с.
13. Гольдштейн Б. С., Соколов Н. А., Яновский Г.Г. Сети связи. – СПб.: БХВ – Петербург, 2010. – 400 с.
14. Горелов Г.В., Ромашкова О.Н., Чан Т. А. Качество управления речевым трафиком в телекоммуникационных сетях. – М.: Радио и связь, 2001. – 111 с.
15. Деарт В.Ю. Мультисервисные сети связи. Транспортные сети и сети доступа. – М.: Инсвязьиздат, 2007. – 166 с.
16. Дибби В.Н. Модели и методы расчета мультисервисных контакт-центров: Автореф. дис. канд. техн. наук. – СПб., 2011. –17 с.
17. Донелли С. Производительность плат DAG при захвате пакетов данных // [Электронный ресурс] // URL: <http://endace.ru/solutions/DAGPacketCapturePerformance/>.
18. Ефимушкин В.А. Особенности взаимоувязанного функционирования систем управления качеством услуг и соглашений об уровне обслуживания // В сб.: Материалы конференции «Услуги электросвязи. Инновационные решения, тенденции и проблемы». – 26 февраля 2010 г. – М.: Экспо-Телеком, 2010. – С. 3-4.
19. Ефимушкин В.А., Углов И.В. Анализ задержек обработки сообщений протокола SIP-I в оборудовании ИВС // Электросвязь. – М., 2013. – № 1. – С.37-43.

20. Ефимушкин В.А., Углов И.В. Анализ сценариев взаимодействия функциональных элементов конвергентной сети при предоставлении различных услуг // Материалы 4-й ежегодной конференции "Конвергенция в телекоммуникациях"; Тезисы. – 7 октября 2010 г. – М., Экспо-Телеком, 2010. – С. 23.
21. Ефимушкин В.А., Углов И.В. Архитектура QoS для конвергентных сетей и особенности ее применения // Т-Comm – Телекоммуникации и Транспорт. – № 7. – 2010. – С. 162-163.
22. Ефимушкин В.А., Углов И.В. Механизмы взаимодействия функциональных элементов конвергентной сети при предоставлении инфокоммуникационных услуг // Электросвязь. – М., 2010. – № 8. – С. 29-32.
23. Ефимушкин В.А., Углов И.В. Моделирование процессов управления исходящими голосовыми вызовами в сетях LTE с использованием технологии CSFB в GSM // XII Всероссийское совещание по проблемам управления. – 16-19 июня 2014. – С. 8674-8685. // [Электронный ресурс] // URL: <http://vspu2014.ipu.ru/proceedings/prcdngs/8674.pdf>.
24. Ефросинин Д.В., Петухова Н.В., Семенова О.В. Фардахов М.П. Моделирование гибридного центра связи с сервисами самообслуживания и пороговым управлением размещением заявок // Управление большими системами: сборник трудов. – № 30-1. – 2010. – С. 352-370.
25. Жожикашвили В.А., Вишневский В.М. Сети массового обслуживания. Теория и применение к сетям ЭВМ. – М.: Радио и связь, 1988 – 192 с.
26. Зарядов И.С., Королькова А.В., Разумчик Р.В. Математические модели расчета и анализа характеристик систем активного управления очередями с двумя входящими потоками и различными приоритетами // Т-Comm – Телекоммуникации и Транспорт. – № 7. – 2012. – С. 107-111.

27. Ивницкий В.А. Теория сетей массового обслуживания // Физико-математическая литература – 2004 – 772 с.
28. Ильин В.П., Кузнецов Ю.И. Трехдиагональные матрицы и их приложения. – М.: Наука, 1985. – 208 с.
29. Клейнрок Л. Вычислительные системы с очередями. – М.: Мир, 1979. – 600 с.
30. Клейнрок Л. Теория массового обслуживания. – М.: Машиностроение, 1979. – 432 с.
31. Корн Г., Корн Т. Справочник по математике для научных работников и инженеров; Перевод Арамановича И.Г., Березмана А.М., Вайнштейна И.А., Румшинского Л.З., Цлафа Л.Я.; Под общ. ред. Арамановича И.Г. – М.: Наука, 1973. – 832 с.
32. Лагутин В. С., Степанов С. Н. Телетрафик мультисервисных сетей связи. – М.: Радио и связь, 2000. – 320 с.
33. Лившиц Б. С., Пшеничников А. П., Харкевич А. Д. Теория телетрафика: Учебник для вузов. – 3-е изд., перераб. и доп. – М.: Связь, 1979. – 224 с.
34. Назаров Н.Г., Назаров А.Н. Теоретические основы формирования оптимальных планов измерений с использованием нецентральных статистик. – М.: ООО «ИПЦ “Маска”», 2011 – 272 с.
35. Наумов В.А., Самуйлов К.Е., Яркин Н.В. Теория телетрафика мультисервисных сетей. – М.: РУДН, 2007. – 191 с.
36. Нетес В.А. Качество обслуживания в сетях связи. Обзор рекомендаций МСЭ-Т // Сети и системы связи. – 1999.– №3. – С. 59-69.
37. О внесении изменений в приказы Министерства информационных технологий и связи Российской Федерации и Министерства связи и массовых коммуникаций Российской Федерации по вопросам применения оборудования и построения телефонной сети связи общего пользования / Минкомсвязь России, Приказ от 6 декабря 2012 г. № 284 г. // Российская газета. – М., 25.01.2013. – № 5991.

38. Олифер В.Г., Олифер Н.А. Компьютерные сети. Принципы, технологии, протоколы. Учебник для вузов: 3-е изд. СПб.: Питер, 2006. – 958 с.
39. Писсанецки С. Технология разреженных матриц. – М.: Мир, 1988. – 410 с.
40. Печинкин А.В. Разумчик Р.В. Стационарные характеристики системы с гистерезисной политикой управления интенсивностью входящего потока // Информационные процессы. – 2013. – Т.3., №3 – С. 125-140.
41. План мероприятий «Развитие конкуренции в сфере электросвязи» / Правительство Российской Федерации, Распоряжение № 130-р. – 03.02.2014.
42. Правила присоединения сетей электросвязи и их взаимодействия / России, Постановление Правительства Российской Федерации от 28 марта 2005 г. № 161 г. // Российская газета. – М., 01.04.2005. – №3735.
43. Программа обработки данных SIP трафика, определения последовательностей сообщений в рамках установления соединения и частоты их появления: пат. 2013661814 РФ: МПК-2014.01 Н 04 Q 1/10 / Разумчик Р.В., Углов И.В.; правообладатель: Федеральное государственное бюджетное учреждение науки Институт проблем информатики Российской академии наук. – № 2013619644; заявл. 24.10.2013; дата гос. рег. 16.12.2013; опубл. 20.01.2014.
44. Проектирование и техническая эксплуатация цифровых телекоммуникационных систем и сетей / Алексеев Е.Б., Гордиенко В.Н., Крухмалев В.В., Моченов А.Д., Тверецкий М.С. – М.: Горячая линия - Телеком, 2008. – 392 с.
45. Самуйлов К.Е., Зарипова Э.Р. Модель локального механизма контроля перегрузок SIP-сервера // Т-Comm – Телекоммуникации и Транспорт. – № 7. – 2012. – С. 185-187.

46. Семенов Ю.В. Проектирование сетей связи следующего поколения. – СПб.: Наука и Техника, 2005. – 240 с.
47. Сервер очень точного времени stratum 1 на Garmin GPS // [Электронный ресурс] // URL: <http://habrahabr.ru/blogs/sysadm/118266>.
48. Соколов Н.А. Процессы конвергенции, интеграции и консолидации в современной телекоммуникационной системе // Connect. – № 10. – 2007. – С. 2-6.
49. Степанов С.Н. Основы телетрафика мультисервисных сетей. – М.: Эко-Трендз, 2010. – 392 с.
50. Степанов С.Н., Иверсен В.Б. Способы уменьшения объема вычислений при расчете моделей систем связи с потерями, основанные на игнорировании маловероятных событий. // Проблемы передачи информации. – Т. 37, №3. – 2001. С. 82-95.
51. Требования к построению телефонной сети связи общего пользования / Мининформсвязи России, Приказ от 8 августа 2005 г. № 97. [Электронный ресурс] // <http://docs.pravo.ru/document/view/12863/34569612>.
52. Углов И.В. Исследование архитектур конвергентных сетей общего пользования // В сб.: Конференция МТУСИ. – Москва, 13-15 февраля 2008 г. – С. 90-93.
53. Углов И.В. Исследование вероятностно-временных характеристик организации вызовов в конвергентных сетях LTE/GSM с использованием технологии CSFB // T-Comm – Телекоммуникации и Транспорт. – Т.8, № 5. – 2014. – С. 56-62.
54. Углов И.В. Методология организации измерения временных характеристик VoIP сети. // T-Comm – Телекоммуникации и Транспорт. – № 7. – 2012. – С. 211-214.
55. Углов И.В. Модели функционирования узлов конвергентной сети при обработке сигнальной информации. // Научно-просветительский портал «Академия современных инфокоммуникационных

- технологий», ЭЛ № ФС 77-50669. [Электронный ресурс] // URL: http://www.acikt.ru/components/com_jshopping/files/demo_products/OP.16.pdf (дата обращения 04.11.2014). – 47 с.
56. Углов И.В. Оптимизация маршрутов прохождения голосового трафика при предоставлении NGN-услуг в распределенных сетях доступа // Т-Comm – Телекоммуникации и Транспорт. – № 11. – 2013. – С. 158-160.
 57. Углов И.В. Перспективы создания виртуальных конвергентных сетей // Т-Comm – Телекоммуникации и Транспорт. – № 7 – 2012. – С. 215-218.
 58. Углов И.В. Сравнительный анализ сценариев функционального взаимодействия элементов конвергентной сети // Т-Comm – Телекоммуникации и Транспорт. – Спецвыпуск, часть 1. – 06.2009. – С. 29-31.
 59. Шнепс-Шнеппе М.А. Системы распределения информации. Методы расчета. – М.: Связь, 1979. – 342 с.
 60. Яновский Г.Г. IP Multimedia Subsystem: принципы, стандарты и архитектура // Вестник связи. – №3. – 2006. – С. 71-76.
 61. 3GPP TSG GERAN Ad Hoc on Release 2000 and Beyond 2. A Comparison Between GERAN Packet-Switched Call Setup Using SIP and GSM Circuit-Switched Call Setup Using RIL3-CC, RIL3-MM, RIL3-RR, and DTAP. – Rev. 0.3. – Munich, German, 2000. – 27 p.
 62. Abaev P. A. On SIP Session Setup Delay Modeling in Next Generation Networks // ICUMT. 2010, – Pp. 1125-1131.
 63. Abaev P., Gaidamaka Yu., Samouylov K. Queuing Model for Loss-Based Overload Control in a SIP Server Using a Hysteretic Technique / 12th International Conference NEW2AN. – Heidelberg, Springer-Verlag, 2012, – № 7469, T. 2. – Pp. 371-378.
 64. Abaev P., Razumchik R. Queuing Model for SIP Server Hysteretic Overload Control with Bursty Traffic // Lecture Notes in Computer

- Science; Internet of Things, Smart Spaces, and Next Generation Networking, 2013. – vol. 8121. – Pp. 383-396.
65. Abaev P., Razumchik R., Uglov I. Some results of measurement based statistical analysis of SIP traffic // DCCN-2013. – 2013. – Pp. 226-230.
 66. Abaev P., Razumchik R., Uglov I. Statistical analysis and modeling of SIP traffic for parameter estimation of server hysteretic overload control // Journal of Telecommunications and Information Technology. – 2013. – № 4. – Pp. 22-31.
 67. Abaev P., Razumchik R., Uglov I. Statistical analysis of message delay in SIP proxy server // Journal of Telecommunications and Information Technology. – 2014. – № 4. – Pp. 1-9.
 68. Ali-Yahiya T. Understanding LTE and its Performance. –Orsay, France: Springer, 2011. – 250 p.
 69. Babu T.V.J.G., Hayes J.F. Modeling and Analysis of Telecommunications Networks. – USA, Hoboken, New Jersey: John Wiley & Sons, Ltd., 2004. – 416 p.
 70. Baskett F., Chandy K.M., Muntz R.R., Palacios F.G. Open, closed and mixed networks of queues with different classes of customers // Journal of the ACM. –vol. 22, № 2. –1975. – Pp. 248-260.
 71. Bestak R., Ulvan A., Ulvan M. IMS Signalling in LTE-based Femtocell Network // IARIA Journal / UBICOMM 2010. – 2010. – Pp. 157-163.
 72. Boucherie R.J., Dijk N.M. Queueing Networks: A Fundamental Approach. – USA, NY, New York: Springer, 2011 – 798 p.
 73. Broadcom NetXtremeI: Network Adapter User Guide. [Электронный ресурс] // URL: http://www.broadcom.com/docs/support/ethernet_nic/Broadcom_NetXtremeII_Server_T7.8.pdf
 74. Burke P.J. The Output of a Queuing System. // Operations Research. – vol.4, №6. – 1956. – Pp. 699-704.
 75. Diot C., Fraleigh C., Moon S., Papagiannaki K., Thiran P. Measurement and Analysis of Single-Hop Delay on an IP Backbone Network // IEEE

- Journal on Selected Areas in Communications. – № 6. – 2003. – Pp. 908-921.
76. Dombacher C. Queueing Models for Call Centres. [Электронный ресурс] // URL: http://www.telecomm.at/documents/Queueing_Models_CC.pdf.
 77. Erramilli A., Narayan O., Willinger W. Experimental queuing analysis with long-range dependent packet traffic // IEEE/ACM Transactions on Networking. – vol. 4, №. 2. – 1996. – Pp. 209-223.
 78. ETSI ES 202 504-1 v. 1.1.1. Open Service Access (OSA); Parlay X Web Services; Part 1: Common (Parlay X 3). Sophia Antipolis, 2008. – 33 p.
 79. ETSI TR 351. Digital cellular telecommunications system; Technical performance objectives. Sophia Antipolis, 1996. – 25 p.
 80. ETSI TS 101 043 v.6.1.0. Digital cellular telecommunications system (Phase 2+); Basic call handling; Technical realization. Sophia Antipolis, 1998. – 181 p.
 81. ETSI TS 101 252. v.5.1.0. Digital cellular telecommunications system (Phase 2+); Description for the use of a Shared Inter Working Function (SIWF) in a GSM PLMN; Stage 2. Sophia Antipolis, 1997. – 25 p.
 82. ETSI TS 123 002 v.11.6.0. Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Network architecture. Sophia Antipolis, 2013. – 107 p.
 83. ETSI TS 123 018 v.11.4.0. Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Basic call handling; Technical realization. Sophia Antipolis, 2013. – 299 p.
 84. ETSI TS 123 272 v. 10.6.0. Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Circuit Switched (CS) fallback in Evolved Packet System (EPS); Stage 2. Sophia Antipolis, 2013. – 98 p.

85. ETSI TS 123 231. v.8.2.0. Technical Specification Group Core Network and Terminals; SIP-I based circuit-switched core network; Stage 2. Sophia Antipolis, 2008. – 72 p.
86. ETSI TS 123 401 v.8.6.0. LTE; General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access. . Sophia Antipolis, 2009. – 287 p.
87. ETSI TS 123 228 v. 8.4.0. IP Multimedia Subsystem (IMS); Stage 2. Sophia Antipolis, 2008. – 236 p.
88. ETSI TS 187 003 v. 3.4.1. Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); NGN Security; Security Architecture. Sophia Antipolis, 2011. – 53 p.
89. ETSI TS Q.704. Signalling network functions and messages. Sophia Antipolis, 1996. – 217 p.
90. Gordon W.J.; Newell G. F. Closed Queuing Systems with Exponential Servers // Operations Research. – vol. 15, № 2. – 1967. – Pp. 254-265.
91. Heyde A., Stewart L. Technical Report 080507A. Using the Endace DAG 3.7GF Card With FreeBSD 7.0. – Australia, Melbourne: CAIA. – 05.2008. – 4 p.
92. IEEE 754. Standard for Binary Floating-Point Arithmetic // IEEE Xplore. – 2008. – 70 p. [Электронный ресурс] // URL: <http://ieeexplore.ieee.org/servlet/opac?punumber=4610933>.
93. IEEE 802.3. Ethernet // IEEE Xplore. – 2001. [Электронный ресурс] // URL: <http://standards.ieee.org/about/get/802/802.3.html>.
94. IETF RFC 2617. HTTP Authentication: Basic and Digest Access Authentication. / Franks, J., Hallam-Baker, P., Hostetler, J., [etc.]. [Электронный ресурс] // URL: <https://www.ietf.org/rfc/rfc2617.txt>.
95. IETF RFC 2719. Framework Architecture for Signaling Transport / Ong L., Rytina I., Garcia M., [etc.]. [Электронный ресурс] // URL: <http://www.ietf.org/rfc/rfc2719.txt>.

96. IETF RFC 2881. Network Access Server Requirements Next Generation (NASREQNG) NAS Model / Mitton D., Beadles M. [Электронный ресурс] // URL: <http://tools.ietf.org/html/rfc2881>.
97. IETF RFC 3261. SIP: Session Initiation Protocol / Rosenberg J., Schulzrinne H., Camarillo G., [etc.]. [Электронный ресурс] // URL: <http://www.ietf.org/rfc/rfc3261.txt>.
98. IETF RFC 3264. An Offer/Answer Model with the Session Description Protocol (SDP) / Rosenberg J., Schulzrinne H. [Электронный ресурс] // URL: <http://www.ietf.org/rfc/rfc3264.txt>.
99. IETF RFC 3435. Media Gateway Control Protocol (MGCP) / Andreassen F., Foster B. [Электронный ресурс] // URL: <http://tools.ietf.org/html/rfc3435>.
100. IETF RFC 3550. RTP: A Transport Protocol for Real-Time Applications / Schulzrinne H., Casner S., Frederick R., [etc.]. [Электронный ресурс] // URL: <http://tools.ietf.org/html/rfc3550>.
101. IETF RFC 3588. Diameter Base Protocol / Calhoun P., Loughney J., Guttman E., [etc.]. [Электронный ресурс] // URL: <http://tools.ietf.org/html/rfc3588>.
102. IETF RFC 3665. Session Initiation Protocol (SIP) Basic Call Flow Examples // Johnston A., Donovan S., Sparks R., [etc.]. [Электронный ресурс] // URL: <http://tools.ietf.org/html/rfc3665>.
103. IETF RFC 3864. Registration Procedures for Message Header Fields / Klyne G., Nottingham M., Mogul J. [Электронный ресурс] // URL: <http://tools.ietf.org/html/rfc3864>.
104. IETF RFC 4506. XDR: External Data Representation Standard / Eisler M. [Электронный ресурс] // URL: <http://tools.ietf.org/html/rfc4506>.
105. IETF RFC 4566. SDP: Session Description Protocol / Handley M., Jacobson V., Perkins C. [Электронный ресурс] // URL: <http://www.ietf.org/rfc/rfc4566.txt>.
106. IETF RFC 4960. Stream Control Transmission Protocol / Stewart R. R. [Электронный ресурс] // URL: <http://tools.ietf.org/html/rfc4960>.

107. IETF RFC 5905. Network Time Protocol Version 4: Protocol and Algorithms Specification / Burbank J., Delaware U., Kasch W., [etc.]. [Электронный ресурс] // URL: <http://tools.ietf.org/html/rfc5905>.
108. IETF RFC 768. User Datagram Protocol / Postel J. [Электронный ресурс] // URL: <http://tools.ietf.org/html/rfc768>.
109. IETF RFC 793. Transmission Control Protocol. Prepared for Defense Advanced Research Projects Agency Information Processing Techniques. 1400 Wilson Boulevard, Arlington, Virginia 22209, 1981. – 85 p. [Электронный ресурс] // URL: <http://www.ietf.org/rfc/rfc793.txt>.
110. ITU-T Recommendation E.800. Terms and definitions related to quality of service and network performance including dependability. Institute from 09.2008. – Geneva, 2009. – 30 p.
111. ITU-T Recommendation G.107. The E-model, a computational model for use in transmission planning. Institute from 1998. – Geneva, 1998. – 17 p.
112. ITU-T Recommendation H.248.1. Gateway control protocol: Version 3. Institute from 03.2013. – Geneva, 2013. – 242 p.
113. ITU-T Recommendation H.323. Packet-based multimedia communications systems. Institute from 12.2009. – Geneva, 2010. – 320 p.
114. ITU-T Recommendation Q.1912.5. Interworking between Session Initiation Protocol (SIP) and Bearer Independent Call Control protocol or ISDN User Part. Institute from 2004. – Geneva, 2004. – 100 p.
115. ITU-T Recommendation Q.700. Introduction to CCITT Signaling System No.7. Approved by the WTSC. Helsinki, March 1-12, 1993. – 24 p.
116. ITU-T Recommendation X.200. Information technology - Open Systems Interconnection - Basic Reference Model: The basic model. Institute from 07.1994. – Geneva, 1994. – 59 p.
117. ITU-T Recommendation Y.2001. General overview of NGN. Institute from 12.2004. – Geneva, 2005. – 18 p.

118. ITU-T Recommendation Y.2011. General principles and general reference model for Next Generation Networks. Institute from 10.2004. – Geneva, 2005. – 34 p.
119. ITU-T Recommendation Y.2013. Converged services framework functional requirements and architecture. Institute from 12.2006. – Geneva, 2006. – 39 p.
120. ITU-T Recommendation E.212. List of mobile country or geographical area codes. Institute from 11.2010. – Geneva, 2011. – 10 p.
121. ITU-T Technical Documentation TD 246 (WP 2/13). Working draft of Q.FMC-IMS: Fixed mobile convergence with a common IMS session control domain. – Geneva, 2006. – 17 p.
122. Jackson J.R. Networks of Waiting Lines // Operations Research. – vol.5, №4. – 1957. – Pp. 518-521.
123. Kitaev M.Yu., Rykov V.V. Controlled Queueing Systems. New York: CRC-Press, 1995.
124. Kramer W., Lagenbach-Belz M. Approximate Formulae for the Delay in the Queueing System GI|G|1 // International Teletraffic Congress 8. – 1976. – Pp. 235-1-235-8.
125. Kuthan J., Schafer G., Sisalem D. Denial of service attacks and sip infrastructure. Attack Scenarios and Prevention Mechanisms // 1st international conference on Principles, systems and applications of IP telecommunications. – USA, New York, NY, 2007. – Pp.57-66.
126. Laszlo L., Laszlo S., Miklos T. Introduction to Queueing Systems with Telecommunication Applications. – Springer, 2013. – 385 p.
127. Mayer G., Poikselka M. The IMS: IP multimedia concepts and services. – 3rd ed. – Kingdom, Chippenham, Wiltshire: John Wiley & Sons, Ltd., 2006. – 533 p.
128. Moulton N. The Evolution of the Contact Centre // XI Call Center World Forum 2012. [Электронный ресурс] // URL: <http://ccwf.ru/content/2012/>.

129. Navidi W. Statistics for engineers and scientists – 3rd ed. – New USA, New York: McGraw-Hill, 2011. – 908 p.
130. Nong G, Hamdi M. On the Provision of Quality-of-Service Guarantees for Input Queued Switches // IEEE Commun. Magazine. – 2000. – № 12. – Pp. 62-69.
131. Ohta M. Overload Control in a SIP Signaling Network // World Academy of Science, Engineering and Technology. – vol.1. № 12. – 2007. – Pp. 820-825.
132. On the self-similar nature of Ethernet traffic / Leland W., Taqqu M., Willinger W., [etc.]. // IEEE/ACM Transactions on Networking. – vol. 2, № 1, – 1994. – Pp. 1-15.
133. Park D., Perros H.G. m-MMBP Characterization of the Departure Process of an m-MMBP/Geo/1/K Queue // Proc. 14th Int. Teletraffic Cong., 6-10 June 1994. North-Holland Elsevier Science B.V. – 1994. – № 1. – Pp. 75-84.
134. Qualcomm Circuit-switched fallback. The first phase of voice evolution for mobile LTE devices // Qualcomm in collaboration with Ericsson. – 2012. – 10 p. [Электронный ресурс] // URL: <http://www.qualcomm.com/media/documents/circuit-switched-fallback-first-phase-voice-evolution-mobile-lte-devices>.
135. Riddington E.P., Saily M., Sebire G. GSM/EDGE: Evolution and Performance. - Kingdom, Chippenham, Wiltshire: John Wiley & Sons, Ltd., 2010. – 504 p.
136. Stamatelos G.M., Koukoulidis V.N. Reservation-based bandwidth allocation in a radio ATM network // IEEE Journal / ACM Trans. on Networking 5. – 1997. – Pp. 420-428.
137. TeleMagement Forum. Business Process Framework Overview. [Электронный ресурс] // URL: <http://www.tmforum.org/Overview/13763/home.html>

138. Voice over LTE (VoLTE) / Poikselka M, Holma H., Hongisto J., [etc.]. – United Kingdom, Chichester, West Sussex: John Wiley & Sons, Ltd., 2012. – 243 p.
139. Whitt W. The Queuing Network Analyzer // USA: The Bell System Technical Journal, 1983. – vol. 62, № 9. – Pp. 2779-2815.

ПРИЛОЖЕНИЕ 1. СПИСОК СОКРАЩЕНИЙ

3 GPP	Консорциум 3 GPP (3rd Generation Partnership Project)
AC	Контроллер доступа (Access Controller)
AN	Узел доступа (Access Node)
AP	Точка доступа (Access Point)
AS	Сервер приложений (Application Server)
ATM	Асинхронный режим переноса (Asynchronous Transfer Mode)
BCMP	Подход к моделированию СеМО, названный по первым буквам фамилий авторов: Baskett, Chandy, Muntz and Palacios
BG	Пограничный шлюз (Border Gateway)
BGC	Контроллер пограничных шлюзов (Border Gateway Controller)
BSC	Контроллер базовых станций (Base Station Controller)
BSS	Радиоподсистема в сетях 2G GSM (Base Station Subsystem)
BTS	Базовая приемопередающая станция (Base Transceiver Station)
CAMEL	Протокол CAMEL (Customised Applications for Mobile networks Enhanced Logic)
CDMA	Множественный доступ с кодовым разделением (Code Division Multiple Access)
CNG	Сетевой шлюз пользователя (Customer Network Gateway)
CNGCF	Конфигуратор сетевого шлюза пользователя (CNG Configuration Function)
CP	Схема полного разделения БП (Complete Partitioning)
CRC	Циклический избыточный код (Cyclic Redundancy Check)
CS	Полнодоступная схема разделения памяти (Complete Sharing)
CSC	Контроллер управления вызовами (Call-Session Controller)
CSFB	Технология использования ресурсов сетей КК для организации вызовов (Circuit Switched FallBack)
DDoS	Распределенная DoS атака (Distributed DoS)
DoS	Отказ в обслуживании (Denial of Service)
DSS1	Протокол цифровой абонентской сигнализации № 1 (DSS1 - Digital Subscriber Signaling № 1)
EDGE	Технология передачи данных EDGE (Enhanced Data rates for GSM Evolution)
eNodeB	Базовая станция сети стандарта LTE (Evolved Node B)
eTOM	Расширенная карта процессов оператора связи (enhanced Telecom Operations Map)
ETSI	Европейский институт по стандартизации в области телекоммуникаций (European Telecommunications

	Standards Institute)
ETTH	Технология “Ethernet к дому” (Ethernet To The Home)
FCFS	Пришедший первым обслуживается первым (First Come–First Served)
GERAN	Сеть радиодоступа GSM EDGE (GSM EDGE radio access network)
GGSN	Шлюз GPRS (GPRS Gateway Service Node)
GPRS	Пакетная радиосвязь общего пользования (General Packet Radio Service)
GSM	Глобальная система мобильной связи (Global System for Mobile Communications)
GW	Шлюз (Gateway)
H-CLF	Функция определения и хранения местоположения гостей пользователей (Home Connectivity session Location and repository Function)
HLR	Домашний регистр местоположения (Home Location Registry)
H-PDBF	Гостевая база данных профилей пользователя (Home Profile Database Function)
HSS	Сервер домашней подписки (Home Subscriber Server)
HTTP	Протокол передачи гипертекста (HyperText Transfer Protocol)
H-U AAF	Функция авторизации гостей пользователей (Home User Access Authorization Function)
IBC	Контроллер пограничных взаимодействий (Interconnection Border Controller)
I-CSCF	Функция управления доступом и управления вызовами (Interrogating-CSCF)
IM	Услуга «мгновенные сообщения» (Instant Messaging)
IMS	Мультимедийная IP подсистема (IP Multimedia Subsystem)
IMSI	Международный идентификатор мобильного абонента (International Mobile Subscriber Identity)
IP	Протокол межсетевое взаимодействие (Internet Protocol)
IS	Бесконечный сервер (Infinite Servers)
ISDN	Цифровая сеть с интеграцией услуг (Integrated Services Digital Network)
ISIM	Механизм индивидуальной идентификации абонентов (Individual Subscriber Identification Module)
ISUP	Прикладная часть ISDN (ISDN User Part)
ISUP IAM	Сообщение по системе ISUP ОКС 7 (ISUP Initial Address Message)
LAI	Идентификатор области местоположения абонента (Location Area Identity)
LAU	Обновление информации о местоположении (Location Information Update)
LCFS	Пришедший последним обслуживается первым (Last

	Come–First Served)
LTE	Семейство стандартов мобильной связи LTE (Long Term Evolution)
MAP	Подсистема MAP ОКС 7 (Mobile Application Part)
MG	Транспортный шлюз (Media Gateway)
MGC	Контроллер управления шлюзами (Media Gateway Controller)
MGCP	Протокол управления шлюзом (Media Gateway Control Protocol)
MMS	Услуги мультимедиа сообщений (Multimedia Message Service)
MNP	Услуга переносимости номера в мобильных сетях (Mobile Number Portability)
MPLS	Технология многопротокольной коммутации меток (Multiprotocol Label Switching)
MRC	Контроллер мультимедиа ресурсов (Multimedia Resource Controller)
MRFC	Функция контроля медиа ресурсов (Media Resource Function Control)
MSC	Центр мобильной коммутации (Mobile Switching Centre)
MTP	Подсистема переноса сообщений (Message Transfer Part)
NACF	Функция конфигурации доступа (Network Access Configuration Function);
NAS	Система присоединения к сети (Network Attachment Subsystem)
NGN	Сети последующего поколения (Next Generation Network)
NodeB	базовая станция сети стандарта UMTS (Node B)
NTP	Протокол сетевого времени (Network Time Protocol)
OFDMA	Множественный доступ с ортогонально-частотным разделением (Orthogonal Frequency-Division Multiple Access)
OSI	Модель взаимодействия открытых систем (Open Systems Interconnection Reference Model)
OSS	Средства диагностирования и поддержки системы (Operation Sussort System)
P-AAA	Прокси узел аутентификации, авторизации и аккаунтинга (Proxy-Authentication, Authorization, and Accounting)
PCRF	Функции доступа и расчетов (Policy and Charging Rules Function)
P-CSC	Прокси-контроллер управления вызовами (Proxy Call Session Controller)
P-CSCF	Функция соединения зон гостевой и домашней сети (Proxy-CSCF)
PCU	Устройство контроля пакетной передачи (Packet Controller Unit)
PDF	Функция стратегических решений (Policy Decision

	Function)
PDF	Функция плотности вероятности (Power Density Function)
PDN	Сеть передачи данных на основе КП (Packet Data Network)
PES	Посистема эмуляции ССОП (PSTN/ISDN Emulation Subsystem)
PPS	Ежесекундный импульс (Pulse Per Second)
PS	Разделение ресурсов (Processor Sharing)
PS HO	Процедура обеспечения хэндовера пакетного соединения (Packet Switched Hand Over)
QoE	Качество восприятия (Quality of Experience)
QoS	Качество обслуживания (Quality of Service)
RFC	Запрос предложений (Request For Comments)
RSVP	Протокол резервирования ресурсов (Resource ReSerVation Protocol)
RTP	Протокол реального времени (Real-Time Protocol)
S	Узел – источник заявок (Source node)
S-AAA	Защищенный узел аутентификации, авторизации и аккаунтинга (Serving Authentication, Authorization, and Accounting)
SBC	Контроллер пограничных сеансов (Session Border Controller)
S-CSC	Функциональный узел управления вызовами на уровне услуг (Serving-CSCF)
S-CSCF	Функция управления вызовами на уровне услуг (Serving-CSCF)
SCTP	Протокол передачи с управлением потоком (Stream Control Transmission Protocol)
SDH	Синхронная цифровая иерархия (Synchronous Digital Hierarchy)
SDP	Протокол описания сеансов (Session Description Protocol)
SEG	Защищенный шлюз (Security Gateway)
SG	Сигнальный шлюз (Signal Gateway)
SGSN	Узел сервисного обслуживания GPRS (Serving GPRS Support Node)
SIGTRAN	Семейство протоколов Signaling Transport
SIP	Протокол установления сеанса (Session Initiation Protocol)
SLA	Соглашение о качестве обслуживания (Service Level Agreement)
SPDF	Функция разделения доступа на базе уровня услуг (Service-based Policy Decision Function)
SRTP	Защищенный протокол реального времени (Secure Real-Time Protocol)
T	Точка выхода заявок из СеМО (Termination node)
TCP	Протокол управления передачей (Transmission Control Protocol)
TDM	Мультиплексирование на основе временного разделения

	(Time Division Multiplexing)
TE	Терминальное оборудование (Terminal Equipment)
TMN	Международная исследовательская организация ТМ Forum
TrGW	Шлюз транскодирования (Translation GateWay)
TV	Телевидение (TeleVision)
UAN	Услуга универсального номера (Universal Access Number)
UCS	Универсальная услуга связи (Universal Communication Service)
UDP	Протокол пользовательских датаграмм (User Datagram Protocol)
UE	Оборудование пользователя (User Equipment)
UMTS	Универсальная мобильная телекоммуникационная система (Universal Mobile Telecommunications System)
UPT	Услуга универсальной персональной связи (Universal Personal Telecommunications)
USIM	механизм универсальной идентификации абонентов (Universal Subscriber Identification Module).
UTRAN	Наземная радиосеть UMTS (UMTS Terrestrial Radio Access Network)
V-CLF	Функция определения и хранения местоположения гостей пользователей (Visitor Connectivity session Location and repository Function)
VLR	Гостевой регистр местоположения (Visitor Location Registry)
VoLTE	Технология реализации голосовых вызовов в сетях LTE (Voice over LTE)
V-PDBF	Гостевая база данных профилей пользователя (Visitor Profile Database Function)
V-U AAF	Функция авторизации гостей пользователей (Visitor User Access Authorization Function)
WAG	Шлюз доступа беспроводной зоны (WLAN Access Gateway)
WLAN	Беспроводная локальная сеть (Wireless LAN)
WTSC	Всемирная конференция по стандартизации электросвязи (The World Telecommunication Standardization Conference)
xDR	Расширенное представление данных (Extended Data Representation)
xDSL	Семейство технологий обеспечения цифровой абонентской линии (Digital Subscriber Line)
AAA	Аутентификация, авторизации и учет (Authentication, Authorization, Accounting)
ATC	Автоматическая телефонная станция
БН	Буферный накопитель
БП	Буферная память
БС	Базовая станция

ВАК	Высшая аттестационная комиссия
ВВХ	Вероятностно-временные характеристики
ВП	Вторичная последовательность
ГП	Главная последовательность
КК	Коммутация каналов
КП	Коммутация пакетов
КПВ	Контроль посылки вызова
КПК	Карманный персональный компьютер
КЦ	Контактный центр
МП	Марковский процесс
МСЭ-Т	Международный союз электросвязи, Сектор стандартизации
НЗ	Неприоритетная заявка
ОБ	Обслуживающий блок
ОКС 7	Общеканальная система Сигнализации № 7
ОС	Операционная система
ПАК	Программно-аппаратный комплекс
ПЗ	Приоритетная заявка
СеМО	Сеть массового обслуживания
СМО	Система массового обслуживания
СПД	Сеть передачи данных
СПС	Сеть подвижной связи
ССОП	Сеть связи общего пользования
Т-АТС	Транзитная АТС
ТМО	Теория массового обслуживания
ТРИК	Территориально-распределенные измерительные комплексы
ТЧ	Тональная частота
Ц-АТС	Цифровая АТС
ЦСИС	Цифровая сеть с интеграцией сервисов
ЧНН	Час наибольшей нагрузки

ПРИЛОЖЕНИЕ 2. СПИСОК ОСНОВНЫХ ОБОЗНАЧЕНИЙ

$c_{a,i}$	Коэффициент вариации интервалов времени между поступлением заявок на i -м элементе
$c_{a,ji}$	Коэффициент вариации интервалов времени между поступлением заявок с узла j на узел i
$c_{s,i}$	Коэффициент вариации времени обслуживания заявок в узле i
$c_{d,ij}$	Коэффициент вариации временных интервалов между заявками потока на выходе узла i по направлению к узлу j
$\delta(i, j)$	Символ Кронекера
$\xi(t)$	Случайная функция времени, описывающая многомерный Марковский процесс (МП)
H	Пространство возможных переходов над X
λ	Интенсивность входящего потока заявок
μ_i	Интенсивность обслуживания одного прибора i -го элемента
m_i	Количество приборов i -го элемента
M_i	Среднее количество заявок на приборах i -го элемента
N_i	Среднее количество заявок на i -м элементе
N_{IS}	Множество узлов, моделируемых бесконечно-линейными системами типа $GI / G / \infty$
N_m	Множество узлов, моделируемых системами типа $GI / G / m$, где натуральное m такое, что $m < \infty$.
π_i	Вероятность потерь на накопителе i -го элемента
$\bar{\mathbf{p}}$	Вектор стационарных вероятностей
q_i	Количество заявок на i -м элементе
$\bar{\mathbf{q}}$	Вектор, описывающий состояние системы
Q	Инфинитезимальный оператор
r_i	Емкость буферного накопителя i -го элемента
R_i	Максимально-возможное количество заявок на i -м элементе
$\ \bar{\mathbf{r}}\ $	Вектор невязки численного и точного результатов
\bar{s}	Переход \bar{s} между двумя состояниями системы
T_i	Среднее время обслуживания i -го элемента
$u(x)$	Функция Хэвисайда
X	Пространство состояний МП
V_I	Средняя длительность процедуры установления исходящего голосового соединения в конвергентной сети при использовании ресурсов сетей предыдущих поколений (CSFB, Circuit Switched FallBack), при необходимости

	обновления информации о местоположении абонента (LAU, Location Information Update)
V_{II}	Суммарная временная задержка в процессе организации исходящего голосового вызова с использованием технологии CSFB в случае, когда LAU не требуется
W_i	Среднее время ожидания на i -м элементе

ПРИЛОЖЕНИЕ 3. АНАЛИЗ АРХИТЕКТУР КОНВЕРГЕНТНЫХ СЕТЕЙ

3.1. Исследование технологий конвергентных сетей

Конвергенция сетей связи различных типов – это процесс их гармоничной эволюции, направленной на создание единой инфраструктуры для поддержки широкого спектра услуг и приложений.

Далее в разделе мы проанализируем основные технологии, используемых при организации конвергентных сетей.

Классификация и краткий обзор технологий конвергентных сетей

Конвергентная сеть предполагает использование широкого спектра технологий передачи информации для организации доступа к ресурсам сети и ее услугам, взаимную интеграцию услуг и технологий [48]. Такая сеть состоит из различных сегментов, которые можно классифицировать следующим образом:

1) сети связи общего пользования (ССОП), ориентированные в основном на передачу голоса, и использующие технологии коммутации каналов (обычно технологию временного разделения каналов (TDM, Time Division Multiplexing) или каналы тональной частоты (ТЧ)) [38];

2) сети передачи данных (СПД), использующие технологии, как коммутации каналов (КК), так и коммутации пакетов (КП). СПД могут использовать такие технологии передачи информации как Asynchronous Transfer Mode (ATM), Ethernet, FrameRelay, Wi-Fi, WiMAX, X.25, семейство технологий Digital Subscriber Line (xDSL), Ethernet To The Home (ETTH) [38];

3) сети подвижной связи (СПС) строятся на основе беспроводных технологий КП или КК, таких как Wi-Fi, Enhanced Data rates for GSM Evolution (EDGE), услуга пакетной радиосвязи (GPRS, General Packet Radio Service), глобальная система мобильной связи (GSM, Global System for Mobile Communications), технология множественного доступа с кодовым разделением (CDMA, Code Division Multiple Access),

широкополосный CDMA (WCDMA, Wideband CDMA), технология множественного доступа с ортогонально-частотным разделением (OFDMA, Orthogonal Frequency-Division Multiple Access) [38, 68, 135].

Очевидно, что такая классификация, предоставляя практическую наглядность, предусматривает некоторую неоднозначность (некоторые технологии могут обеспечивать информационный обмен в сетях, принадлежащих к различным классам). Поэтому в некоторых случаях может быть более удобным пользоваться классификацией по типу коммутации, используемой в сети:

- сети с КК;
- сети с КП.

В диссертации будут использоваться обе системы классификации, в зависимости от конкретной ситуации и смыслового контекста.

Анализ протоколов сигнализации в конвергентных сетях

Для организации передачи данных в конвергентных сетях может использоваться достаточно широкий спектр протоколов – начиная от закрытых (например, Skype) и заканчивая открытыми, например, H.323 [113] или протокол установления сеанса (SIP, Session Initiation Protocol) [97].

Ориентированность производителей на открытые стандарты международных организаций, таких как Международный союз электросвязи, Сектор стандартизации (МСЭ-Т) и Европейский институт по стандартизации в области телекоммуникаций (ETSI, European Telecommunications Standards Institute), позволяет совместно использовать сетевое оборудование различных производителей и упрощает процесс взаимодействия операторов сетей связи друг с другом. Использование открытых протоколов значительно облегчает процесс построения конвергентной сети, облегчая процесс централизации управления. В целом, по сравнению с использованием частных протоколов, достигается большая прозрачность конвергентной сети.

Рассмотрим подробнее открытые протоколы H.323 и SIP.

Протокол H.323 широко применяется многими производителями для организации управления передачей голоса поверх сетей КП, построенных на базе протокола межсетевого взаимодействия (IP, Internet Protocol) [38], многими производителями. Однако, в случае использования протокола H.323 для организации дополнительных мультимедийных услуг в сети возникает ряд трудностей:

- описание соединения в H.323 происходит в специальном формате, усложняющем адаптацию протокола для новых услуг;

- протокол оптимизирован для организации IP-телефонии, но не для мультимедийных услуг в целом, таких как видео-услуги, например.

Протокол SIP – относительно новый протокол (разработка его началась в 1996), на данный момент, являющийся одним из наиболее универсальных протоколов организации сеанса связи, который может включать обмен мультимедийным содержимым (видео- и аудио-конференция, мгновенные сообщения, онлайн игры). В эталонной модели (Reference Model) взаимодействия открытых систем (OSI, Open Systems Interconnection [116]) SIP относится к сетевым протоколам прикладного уровня. Для описания параметров устанавливаемого соединения может использоваться протокол описания параметров сеансов (SDP, Session Description Protocol) [105].

Протокол SIP описывает процедуры установления, сеансового управления и завершения связи между двумя или более сторонами, используя их уникальные имена. Также он определяет способы согласования между сторонами, устанавливающими соединение, открытие каналов информационного обмена на основе других протоколов, которые могут использоваться для непосредственной передачи пользовательской информации, например, на основе протокола реального времени (RTP, Real-Time Protocol), защищенного RTP (SRTP, Secure Real-Time Protocol) [100]. Протокол RTP может работать поверх протокола управления передачей (TCP, Transmission Control Protocol), протокола пользовательских датаграм (UDP, User Datagram Protocol)

либо протокола передачи информации управления потоком (SCTP, Stream Control Transmission Protocol) [106, 108, 109].

Кроме протоколов организации сеансов в конвергентных сетях используется большое количество служебных протоколов передачи сигнальной информации. К ним можно отнести такие протоколы как DIAMETER, MGCP, H.248 [99, 101, 112], семейство протоколов сигнализации Signaling Transport (SIGTRAN) [95].

Транспортировка информации сигнализации по технологии SIGTRAN предназначена для передачи сообщений протокола сигнализации сети КК через сеть КП и должна обеспечивать:

- передачу сообщений разнообразных протоколов сигнализации, обслуживающих соединения сетей КК, например, протоколов прикладных и пользовательских подсистем общеканальной системы сигнализации № 7 (ОКС 7) [115]. Семейство протоколов SIGTRAN обеспечивает передачу сообщений различных уровней, таких как МТР-3, ISUP [95, 115];

- средства идентификации конкретного транспортируемого протокола сигнализации сети с КК;

- общий базовый протокол, определяющий форматы заголовков, расширения в целях информационной безопасности и процедуры для транспортировки сигнальной информации, а также (при необходимости) расширения для введения конкретных индивидуальных протоколов сигнализации сети с коммутацией каналов;

- функциональные возможности (с использованием нижележащего сетевого протокола, например IP), соответствующие нижнему уровню конкретной сети с КК.

Поскольку сети КК и их сигнальные системы разработаны и используются давно [11], механизмы формирования задержек при установлении соединения детально изучены на настоящий момент [2, 5, 7].

Протокол DIAMETER обеспечивает передачу сигнальных сообщений для предоставления услуг авторизации, учета и аутентификации (AAA, Authentication, Authorization, Accounting) [38].

Протоколы H.248, MGCP используются для передачи сигнальной информации в процессе управления коммутацией информационных потоков на транспортном шлюзе (MG, Media Gateway) [46].

Начало использования протоколов H.248, MGCP, DIAMETER было обусловлено реализацией концепции перехода к программным коммутаторам Softswitch [12], пришедшим на замену цифровым автоматическим телефонным станциям (Ц-АТС). Вклад этих протоколов в задержку установления сеанса должен учитываться при исследовании SIP-сетей.

Исследование транспортных протоколов в конвергентных сетях

Согласно архитектурной модели конвергентной сети, ее можно логически разделить на ядро сети и сети доступа [15]. Отметим, что технологии и протоколы, используемые в телекоммуникационных сетях, и описанные в [13, 38] не являются предметом исследования настоящей диссертации. Кратко остановимся на структуре транспортной сети.

Ядро конвергентной сети обычно представляет собой совокупность магистральных каналов передачи данных и сетевого оборудования, обеспечивающих передачу информации и управляющего работой конвергентной сети в целом. Ядро конвергентной сети использует стек протоколов TCP/IP, а так же обширный диапазон специализированных протоколов уровней 4-7 модели OSI для транспортировки потоков трафика различных типов поверх IP-сети.

Сети доступа организуются на основе сетей КК и КП. Разнообразие технологий, использующихся в сетях доступа, обусловлено процессом естественной эволюции сетей связи. Таким образом, в сетях доступа используется все основные транспортные протоколы и технологии сетей КК и КП.

Анализ технологий обеспечения качества передачи информации

Качество обслуживания (QoS, Quality of Service) детально рассматривается многими организациями по стандартизации, в том числе МСЭ-Т. В Рекомендации МСЭ-Т E.800 [110] QoS определяется как «суммарный эффект показателей качества услуги, который определяет степень удовлетворенности пользователя услуги».

Механизмы и технологии обеспечения QoS в пакетных сетях – обширная тема [36]; в данном параграфе мы кратко остановимся на основных, применяемых для обеспечения и контроля качества предоставляемых услуг.

В сетях КП существуют механизмы, позволяющие предоставлять услуги с гарантированным качеством. К механизмам обеспечения сквозного качества услуг относятся технологии дифференцированного обслуживания очередей и приоритезации трафика (DiffServ), протокол резервирования ресурсов (RSVP, Resource ReSerVation Protocol) [38] и технология многопротокольной коммутации меток (MPLS, Multiprotocol Label Switching) [38]. Механизмы обеспечения QoS в IP сетях развиваются и структурируются в архитектуре QoS [21]. Отметим, что нормы, регламентирующие качество передачи данных могут различаться в зависимости от конкретных условий эксплуатации и назначения сети [44], также как и методы его оценки [111].

Для целей контроля качества существуют два подхода, связанные с применением методик выборочного или сплошного контроля. В эксплуатации эти два подхода используются в различных системах управления и мониторинга качества.

Сплошной контроль качества и формирование базы расширенного представления данных (xDR, Extended Data Representation) [104] основан на принципе полного знания обо всех элементарных компонентах сети. Применение этого принципа, например, к контролю качества телефонной сети приводит к необходимости формирования базы данных всех вызовов в сети.

В качестве средств сбора информации могут выступать:

– территориально-распределенные измерительные комплексы (ТРИК), представляющие собой систему измерительных средств, объединенных сетью, осуществляющей сбор и обработку информации [4];

– встроенные диагностические средства в составе оборудования на основе системы операционной поддержки (OSS, Operation Support System).

Подводя итог вышесказанному, выделяют четыре варианта построения современных систем контроля качества (рис. 1).

Существующие системы сбора информации позволяют обеспечить необходимый для эксплуатации сети функционал. При глубоком исследовании сети, исследователь зачастую будет вынужден отказаться от использования существующих коммерческих систем и прибегнуть к разработке узкоспециализированных алгоритмов сбора и обработки информации, отвечающих конкретной задаче [43, 91, 54].



Рис.1. Системы контроля качества

3.2. Анализ функциональной модели конвергентной сети

Являясь промежуточным этапом развития сетей связи следующего поколения, конвергентная сеть базируется на тех же принципах функционального разделения, что и NGN. В результате, в соответствии с рекомендацией МСЭ-Т Y.2001 [117], для конвергентной сети возможно

применение трехуровневой модели NGN, часть требований которой не будет выполнена. Детальный анализ стандартов и рекомендаций, описывающих архитектуру конвергентной сети [121, 127], позволяют выделить следующие функциональные зоны:

1) Зона пользовательского оборудования – здесь обеспечивается простой, интуитивно-понятный пользователю интерфейс, используемый для доступа к услугам конвергентной сети. В качестве аппаратной реализации могут выступать фиксированные и беспроводные терминалы доступа: мобильные телефоны, компьютеры с поддержкой соответствующих технологий, карманные персональные компьютеры (КПК), web-интерфейсы сайтов сети Интернет. В зоне пользовательского оборудования должны использоваться следующие технологии доступа:

– беспроводные технологии доступа (WiFi, WiMAX, GPRS, EDGE, GSM, CDMA);

– технологии передачи данных по фиксированным линиям, например, цифровая сеть с интеграцией служб (ЦСИС), Ethernet, ССОП, xDSL, FTTH, dial-up, FrameRelay.

2) Зона станционного доступа – здесь пользовательский трафик агрегируется от проводных и беспроводных терминалов и осуществляется организация передачи пользовательской информации на верхние архитектурные уровни конвергентной сети.

В качестве аппаратной реализации могут выступать беспроводные точки доступа (станции WLAN), точки доступа операторов сотовой связи (UTRAN, GERAN), автоматические телефонные станции (АТС), серверы приложений (в случае доступа через сеть компании или Интернет – сеть предприятия).

3) Зона конвергенции технологий доступа – в данной зоне трафик агрегируется в единый поток, поступающий от разнотехнологичных сетей доступа, и перенаправляется на транспортную сеть конвергентной сети.

В качестве аппаратной реализации могут выступать мосты, маршрутизаторы (с наличием соответствующих интерфейсов), программные коммутаторы (Softswitch).

4) Зона гостевой сети – здесь предоставляются основные услуги по передаче агрегированного пользовательского и сигнального трафика в гостевой сети с учетом информации, полученной из домашней сети абонента.

В качестве аппаратной реализации должны выступать серверы аутентификации, авторизации и аккаунтинга (AAA, Authentication, Authorization, Accounting), маршрутизаторы и платформа мультимедийной IP подсистемы (IMS, IP Multimedia Subsystem) в целом.

5) Зона домашней сети – в задачи данной зоны входит предоставление услуг по запросу абонента, обработка AAA-запросов, контроль над передачей и распределением пользовательского и сигнального трафика, управление профилями пользователей.

В качестве аппаратной реализации могут выступать серверы баз данных, серверы приложений, а также платформа IMS.

Сравнительная спецификация функциональных узлов конвергентной сети

В состав узлов конвергентной сети, на этапе ее организации и для обеспечения постепенного перехода к структуре сети нового поколения, могут входить элементы NGN, построенные на платформах Public switched telephone network Emulation Subsystem (PES) и IMS. На первых этапах построения конвергентной сети, построение телефонной сети осуществляется, обычно, на базе решения программного коммутатора.

Проведем краткую классификацию типовых узлов связи, используемых в конвергентной сети:

1) узел связи типа «А» (Softswitch) включает [46]:

а) контроллер управления шлюзами (MGC, Media Gateway Controller);

б) транспортный шлюз MG;

в) сигнальный шлюз (SG, Signal Gateway).

2) узел связи типа «Б» (IP-АТС на базе программного коммутатора), – телефонный концентратор (IP-АТС) [46]. Данный узел в основном является частным решением на базе программного коммутатора или решений на базе гибридных АТС, имеющих в своем составе функции MGC, MG и SG.

3) узел связи типа «В» (IP-АТС на базе контроллера управления соединениями (P-CSC, Proxy – Call-Session-Controller), выделенного узла IMS). Контроллер P-CSC является вынесенным узлом платформы IMS. На практике функциональность P-CSC часто выполняется в составе пограничного контроллера сеансов (SBC, Session Border Controller). Отметим, что в документах ETSI начиная с 2011 года, функционал SBC был разделен на два функциональных элемента:

– пограничный контроллер взаимодействий (IBC, Interconnection Border Controller), отвечающий за обработку передачу потоков сигнальных сообщений;

– шлюз транскодирования (TrGW, Translation Gateway).

4) узел связи типа «Г» (на базе платформы IMS). Реализация на сети узла связи Г-типа позволит в дальнейшем исключить применение на сети выделенных решений с использованием независимых центральных управляющих элементов.

Помимо узлов связи, которые можно отнести к элементам сети NGN, в конвергентную сеть, могут входить узлы связи транспортной сети и сетей доступа. Узлы этих классов различаются в зависимости от используемых технологий.

К функциональным узлам конвергентной сети можно отнести элементы сети, не выполняющие функции коммутации и, в соответствии с терминологией принятой Минкомсвязи России, не являющиеся узлами связи [37, 51]. К таким функциональным узлам относятся:

– терминальное оборудование (TE, Terminal equipment) или оборудование пользователя (UE, User equipment);

- повторители (Repeater) и шлюзы (GW, GateWay);
- базы данных: сервер домашней подписки (HSS, Home Subscriber Server), гостевой регистр местоположения (VLR, Visitors Location Register);
- серверы приложений (AS, Application server).

Исследование механизмов взаимодействия функциональных узлов конвергентной сети

В состав конвергентной сети может входить большое число различных функциональных узлов, характер и механизмы взаимодействия которых определяется типом узлов, участвующих в процессе [20, 22]. Остановимся на рассмотрении типичных механизмов взаимодействия функциональных узлов конвергентной сети. Механизмы взаимодействия функциональных узлов, наследованные конвергентными сетями из сетей СПС, СПД или ССОП, особенно на транспортном уровне, часто не претерпевают изменений, поэтому рассмотрим механизмы взаимодействия, характерные для конвергентных сетей, и в которых, в основном, принимают участие функциональные элементы сетей NGN.

Взаимодействие между составными частями узла связи (между MGC и SG), определенного как тип «А» в предыдущем подразделе, происходит по правилам, формализованным в рабочем предложении (RFC, Request for Comments) №2719 [95], и известного как стек протоколов передачи сигнализации SIGTRAN.

Взаимодействие между MGC и AS, а также между центральным коммутационным узлом платформы IMS – контроллером управления вызовами (CSC, Call Session Control) и AS осуществляется, преимущественно, двумя различными способами, в зависимости от специфики AS, с помощью протоколов:

- SIP [97];
- Parlay, Parlay-X [78].

Взаимодействие между AS, базами данных, например, HSS, и узлами платформы, например, CSC, осуществляется по протоколам MAP, RADIUS, DIAMETR, и используется, обычно, для организации процессов работы с профилями пользователя.

Взаимодействие терминального оборудования и конвергентной сети может быть организовано с помощью различных механизмов, которые рассмотрены ниже.

Исследование механизмов взаимодействия пользователя с конвергентной сетью

Взаимодействие пользователя с конвергентной сетью в большинстве случаев представляет собой использование ресурсов сети либо ее услуг. Процессы использования ресурсов или услуг сети можно, в свою очередь разбить на множество процессов (например, процессы установления соединения, аутентификации, авторизации пользователя). Чем ближе конвергентная сеть в своем развитии приближается к NGN, тем более универсальные механизмы взаимодействия пользователя с сетью используются в ней.

Такие условия, как обеспечение глобальной мобильности пользователя и непрерывности вызова при переходе через границы разнородных сетей накладывают на механизмы идентификации особые требования. В результате, сценарии взаимодействия пользователя в конвергентной сети могут сильно отличаться от аналогичных по функциональности сценариев, реализованных в обычных сетях (ССОП, СПС и СПД).

Процесс регистрации пользователя в сети доступа может быть осуществлен вручную, путем набора идентификационного кода через абонентский терминал, или автоматически, с использованием возможностей реализуемых в унифицированных терминалах с поддержкой GSM/UMTS/WLAN.

Для обеспечения мобильности пользователей необходимо разделять процесс идентификации пользователя и идентификации терминала

фиксированного доступа. Процесс идентификации не должен зависеть от технологии, используемой для осуществления доступа. Одновременно необходимо учитывать возможность перемещения пользователя из одного технологического сегмента конвергентной сети в другой.

Рекомендация МСЭ-Т E.212 [120] устанавливает требования к идентификации пользователей в мобильных сетях связи. Этот же механизм может быть использован и для идентификации пользователей в конвергентных сетях связи.

Кратко перечислим несколько возможных механизмов идентификации абонентов в конвергентных сетях [127]:

- механизм индивидуальной идентификации абонентов (ISIM, Individual Subscriber Identification Module);

- механизм универсальной идентификации абонентов (USIM, Universal Subscriber Identification Module).

Появляющиеся и используемые в настоящее время в России терминалы 2-го, 3-го и 4-го поколения, не поддерживающие ISIM и/или не поддерживающие технологию IPSec [38], которая является основным механизмом защиты в сетях на базе оборудования IMS, тоже могут подключаться к конвергентным сетям.

Наряду с используемыми в настоящее время мобильными терминалами, не имеющими механизмов ISIM, на фиксированных сетях связи также могут использоваться фиксированные SIP-терминалы, не обладающие данным механизмом. Такие аппараты также должны иметь возможность подключаться к конвергентным сетям. В этом случае, обеспечение регистрации, аутентификации и защиты осуществляется средствами сети доступа фиксированной сети связи [88].

В результате, на обычных терминалах как проводной, так и беспроводной связи могут применяться механизмы SIP HTTP аутентификации [94] и аутентификации на сетевом сервере (NAS, Network Access Server) [96].

Анализ принципов взаимодействия конвергентных сетей друг с другом и с сетями связи других типов

Исходя из рекомендаций МСЭ-Т и документов ETSI состав и архитектура конвергентной сети могут быть различными (п. 1.2). В качестве примера для анализа принципов взаимодействия рассмотрим ситуацию, в которой предполагается совместная работа конвергентной сети, построенной на базе оборудования программного коммутатора, фиксированной сети, СПС и СПД (рис.2) [46, 134]. Сценарии взаимодействия разнотехнологичных сегментов конвергентной сети будут существенно различаться [58]. На рис. 2 использованы следующие сокращения:

- BTS (Base Transceiver Station) – базовая станция;
- BSC (Base Station Controller) – контроллер базовых станций;
- PCU (Packet Controller Unit) – контроллер устройств, использующих пакетный способ передачи данных;
- MSC (Mobile Switching Centre) – Центр мобильной коммутации;
- VLR – Гостевой регистр местоположения;
- HLR (Home Location Registry) – Домашний регистр местоположения.

Взаимодействие сегментов сети связи на основе технологий КК и КП осуществляется через MG и SG. Шлюз MG обеспечивает преобразование формата передаваемых голосовых сообщений, поступающих из TDM-сегмента в IP-сегмент и обратно. Шлюз SG реализует функции передачи сигнальных сообщений ОКС 7, поступающих из TDM-сегмента в IP-сегмент.

Управление MG осуществляется контроллером MGC на основе обработки поступающей в него сигнальной информации. В зависимости от реализации оборудования программного коммутатора сигнальная информация поступает в MGC из сегмента КК через шлюзы SG по протоколам SIGTRAN, либо от других MGC через сеть КП.

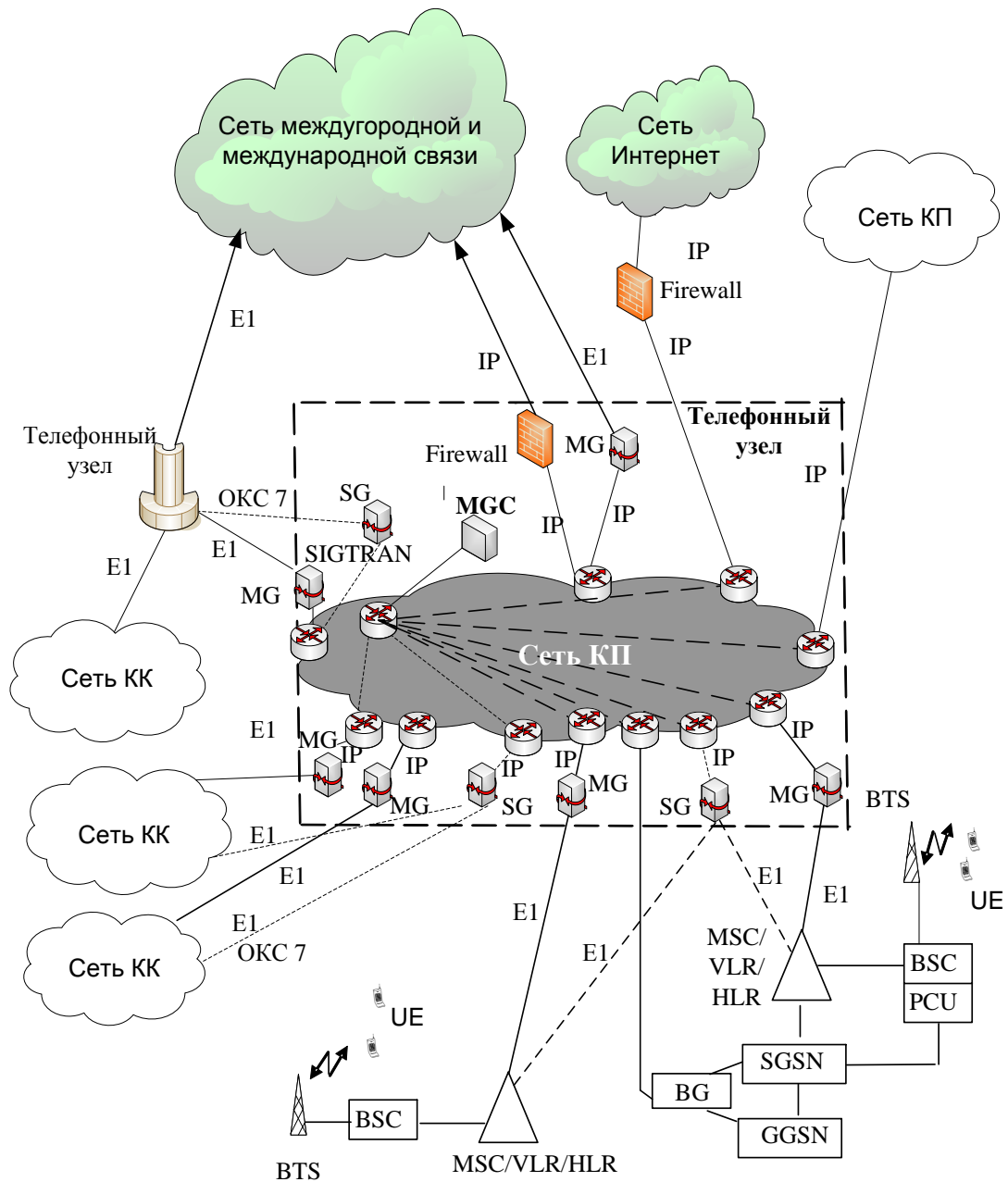


Рис. 2. Взаимодействие различных сегментов конвергентной сети на базе Softswitch с сетями фиксированной и мобильной связи

Взаимодействие пакетной сети фиксированной связи на базе оборудования программного коммутатора с сетью СПС, имеющей оборудование 2-го поколения, осуществляется между узлами сети фиксированной связи и узлом СПС через шлюз MG для голосового трафика и через шлюз SG для сигнального трафика.

Взаимодействие пакетной сети фиксированной связи на базе оборудования Softswitch с сетью СПС, которая имеет оборудование 2.5-

го поколения, осуществляется через узел обслуживания абонентов GPRS (SGSN, Serving GPRS Support Node), шлюз GPRS (GGSN, GPRS Gateway Service Node) и пограничный шлюз (BG, Border Gateway), входящий в состав СПС, а также маршрутизатор пакетной сети связи.

Далее рассмотрим взаимодействие фрагментов сети фиксированной связи, использующих разные технологии коммутации, и сегмента сети СПС посредством платформы IMS, рис.3.

Взаимодействие узла связи, реализованного на технологии КК, и узла связи, реализованного на технологии КП, осуществляется через платформу IMS [87]. В сопрягающее оборудование входят шлюзы SG и MG, а также MGC. MGC осуществляет преобразование протоколов сигнализации, поступающих от сети КК, в сообщения SIP, которые платформа IMS использует для управления сеансом связи между сетями.

Управление вызовами осуществляется элементом CSC [60, 86], на базе информации сигнализации, поступающей от MGC. Он же преобразует сообщения ISUP, поступающие от шлюза SG по протоколу SIGTRAN, в сообщения SIP. Шлюз SG производит преобразование нижних протокольных уровней сигнализации (обычно L2-L4) при взаимодействии сетей КП и КК.

Взаимодействие между конвергентной сетью на базе IMS и СПС, имеющей оборудование 2-ого поколения, осуществляется через шлюзы SG и MG IMS-сети, а также MGC.

Взаимодействие пакетной сети фиксированной телефонной связи на базе технологии IMS с сетью СПС, которая имеет оборудование поколения 2.5, осуществляется через оборудование SGSN, GGSN узла СПС, граничный шлюз BG, входящий в состав СПС, маршрутизатор пакетной сети связи. Элемент CSC платформы IMS, являющийся узлом управления вызовами, осуществляет маршрутизацию вызовов и контроль установленных соединений.

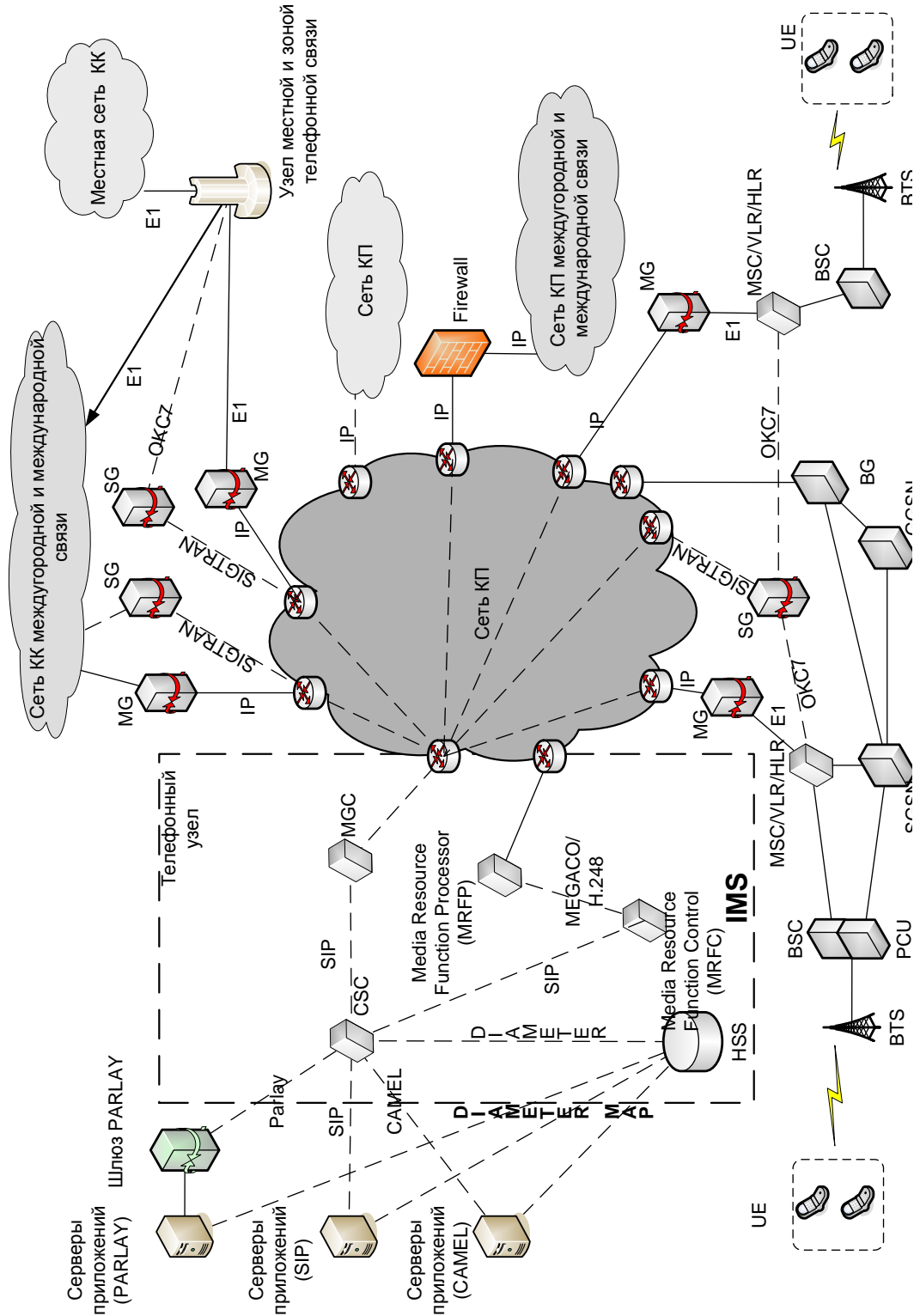


Рис. 3. Взаимодействие узлов конвергентной сети, построенной на базе платформы IMS

Элементы конвергентной сети могут быть разделены не только технологически, но и образовывать облачные структуры, давая возможность создавать виртуальные конвергентные сети [57].

В заключение рассмотрим развитие и конвергенцию сетей СПС разных поколений. В последние два десятилетия развитие технологий перешагнуло 2-е и 3-е поколение СПС и вплотную подошло к полноценной реализации 4-го поколения. В настоящее время все три поколения сосуществуют в сетях операторов связи в различных пропорциях, представляя собой сложную конвергентную структуру. Конвергенция в данном случае происходит на всех трех, определенных МСЭ-Т, уровнях [117,118].

3.3. Исследование архитектурных решений по организации связи в конвергентных сетях

Классификация архитектурных решений по организации связи конвергентных сетях

Конвергентные сети, в зависимости от стадии эволюции, могут быть организованы на базе узлов связи различных типов (п. 1.2), и, соответственно, могут обладать различной архитектурой [46].

При переходе к сетям NGN, можно выделить три этапа эволюции конвергентной сети. На каждом этапе архитектура существенно меняется.

При эволюции конвергентной сети большое внимание уделяется обеспечению комфортного взаимодействия пользователя с сетью, сохранению всего набора предоставляемых на ССОП услуг и обеспечению взаимодействия с ССОП с сохранением функциональности [117, 118, 127]. Плавный переход к новым технологическим решениям достигается за счет использования платформы PES на базе решения Softswitch.

Постепенно, переводя соединительные линии из действующего узла в область Softswitch, сеть преобразуется таким образом, что коммутационная мощность будет полностью переведена с

традиционного узла на узел NGN, построенный на решении Softswitch, для целей создания логики платформы PES. Назовем данный тип архитектуры конвергентной сети первым, рис. 4.

При дальнейшем развитии конвергентной сети требуется установка узлов связи типа «Б» и «В» для подключения абонентов через IP-телефоны. Установка узлов связи типа «Б» позволяет концентрировать трафик от абонентов сети КП и управлять ее абонентами (рис. 5). Появляются новые услуги (например, услуга инициализации и завершения вызовов на разнотехнологичных сетях связи). Назовем данный тип архитектуры конвергентной сети вторым.

На последнем этапе развития конвергентной сети (третий тип архитектуры (рис.6)) и последующем переходе к полноценной NGN предполагается установка узлов связи типа «Г». На этом этапе развития появляются новые абоненты — абоненты сети NGN.

Подключение пользователей телефонной сети происходит уже по новым технологиям с использованием для этого имеющихся возможностей сети передачи данных на базе протокола IP. Реализация платформы IMS позволяет построить сеть как полноценную NGN.

Сравнительный анализ типовых архитектурных решений по организации конвергентных сетей

При внедрении архитектурного решения первого типа, после установки шлюзов SG, MG и контроллера MGC, сеть связи обладает следующими характеристиками:

- центральная система управления;
- общая база данных пользователей;
- расширенные возможности конвергентной сети в части предоставления услуг (за счет установки AS);
- упрощение процедур подключения пользователей к сети;
- упрощение процедур подписки, активации, деактивации и т.д. пользователей сети на услуги.

В этом случае абонентам сети КК будет представлен широкий набор голосовых услуг сети КП.

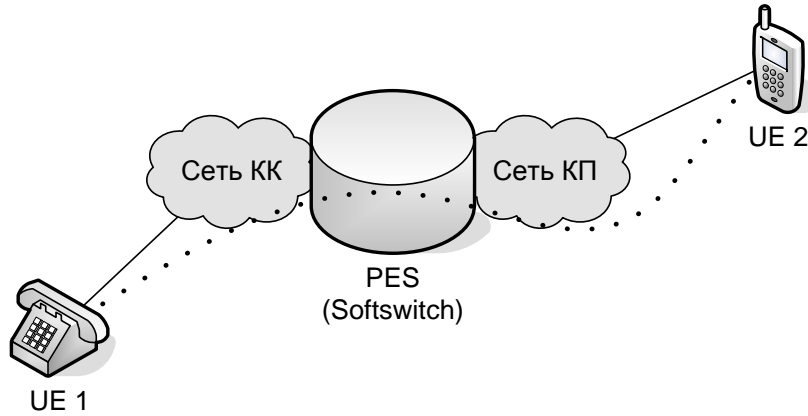


Рис.4. Схема взаимодействия сетей ССОП на базе платформы PES

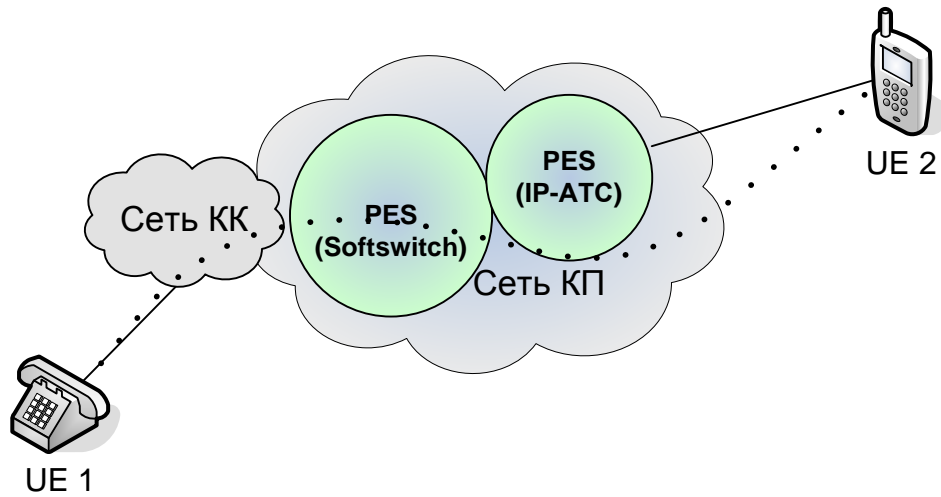


Рис. 5. Внедрение IP-АТС

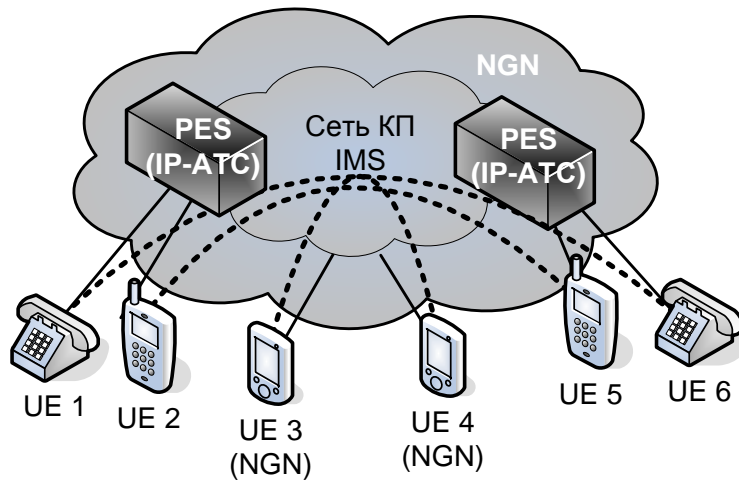


Рис.6. Взаимодействие основных узлов сети, построенной на платформе IMS

Архитектурное решение второго типа существенно отличается от первого. На этапе его реализации устраняется проблема масштабируемости сети (как КК, так и КП сегмента) и проблема расширения сегмента, пользователи которого подключаются при помощи IP-телефонов.

После реализации данного архитектурного решения возможно расширение спектра предоставляемых на сети услуг, например, за счет услуг инициализации и окончания вызовов на разнотехнологичных сетях связи.

Архитектурное решение третьего типа предполагает организацию конвергентной сети на базе платформы IMS. Возможности платформы IMS позволяют предоставлять пользователям сети конвергентные и NGN-услуги и решают проблемы масштабируемости и унификации сети. После внедрения на сети платформы IMS, сеть находится под единым управлением центрального коммутационного ядра платформы, имеет единую базу данных абонентов, легко масштабируется как в сегменте КП так и сегменте КК, предоставляет стандартные механизмы взаимодействия с серверами приложений на основе технологий Parley, SIP, CAMEL.

Определение критериев оптимальных способов построения конвергентных сетей

Построение конвергентной сети может осуществляться по различным сценариям, с использованием широкого спектра технологий и оборудования. Для каждого конкретного случая необходима разработка индивидуального сценария, учитывающего особенности объединяемых сегментов. Вместе с тем, следует отметить некоторые общие критерии, оптимизирующие процесс конвергенции и устанавливающие определенные требования к будущей сети [44]:

1) Обеспечение достаточной производительности. При этом необходимо учитывать, что значения пиковых нагрузок могут многократно превышать их нормальные повседневные значения. Также

необходимо обеспечить необходимый запас производительности, оценив потенциальный рост потребностей абонентов сети;

2) Обеспечение масштабируемости решения. В современных условиях структура сети общего пользования оперативно меняется, согласно растущим потребностям абонентов. Необходимо чтобы сетевая инфраструктура также быстро менялась без ущерба для бюджета и работы оператора сети;

3) Обеспечение высокой доступности. Необходимо, чтобы инфраструктура сети работала максимально непрерывно, а возможный отказ отдельных компонентов был либо незаметен, либо быстро устраним;

4) Обеспечение информационной безопасности. Сетевая инфраструктура должна соответствовать существующим на предприятии политикам безопасности по разграничению доступа, защите от внутренних и внешних атак;

5) Простота применения или управления. Оборудование и решения, внедряемые в сетевую инфраструктуру должны легко управляться, иметь возможность быстрого диагностирования и замены. Не требовать излишних административных расходов;

6) Обеспечение обслуживания различного типа трафика. Сетевая инфраструктура современной сети общего пользования должна обеспечивать возможность функционирования интегрированных приложений, обеспечивать передачу помимо обычных данных, передачу голоса и видео с требуемым качеством;

7) Экономичность. Уменьшение стоимости внедрения и владения сетевой инфраструктурой.

Исследование возможных способов организации сетевой инфраструктуры

Основой сетевой инфраструктуры конвергентной сети можно считать магистральную сеть, построенную на основе технологии КП. Магистральная сеть содержит центральный коммутационный узел или

узлы, отвечающий за управление вызовами в сети, точки взаимодействия с серверами приложений, сторонними операторами сетей, как конвергентных и NGN, так и традиционных КК и КП.

В зависимости от этапа развития сети и типов узлов (п. 1.2), используемых для ее организации, структура сети может существенно различаться. Глобально можно выделить два способа организации сетевой инфраструктуры:

- на базе оборудования программного коммутатора Softswitch (использование узлов типа «А», «Б», «В»);
- на базе платформы IMS (использование узла типа «Г»).

Организация сетевой инфраструктуры на базе оборудования типа Softswitch используется на данный момент достаточно широко для объединения сегментов сетей с КП и КК. Этот вариант рассмотрен в п. 1.2, на рис. 2 в п. 1.3 на рис. 4, 6. Второй же вариант пока не получил широкого распространения на сетях операторов, действующих на территории РФ и находится на стадии внедрения. Рассмотрим его более подробно.

Функциональность платформы IMS позволяет объединить передачу голоса и данных в рамках единой пакетной сети с разделением управления вызовами и голосового трафика, роумингом вызовов между мобильными и фиксированными сетями. Платформа предполагает использование SIP протокола и позволяет реализовать широкий спектр традиционных и NGN-услуг.

Общая архитектура IMS, в соответствии со спецификациями ETSI [86], предполагает наличие следующих основных компонентов:

- элементов управления вызовами, например, варианты CSC;
- баз данных, например, HSS;
- серверов AS и элементов управления, например, функциональных элементов контроля медиа ресурсов (MRFC, Media resource function control);

- элементов, обеспечивающих межсетевое взаимодействие, например, MGC, MG, SG и пограничных контроллеров шлюзов (BGC, Border Gateway Controller). BGC организуется на базе MGC, однако выполняет функции пограничного узла;

- элементов обеспечения защиты информации и качества услуг (PDF, Policy Decision Function; SEG, Security Gateway);

- элементов системы управления и мониторинга OSS .

На рис. 7 показана схема включения составных элементов платформы IMS. Каждый элемент архитектуры IMS отвечает за определенную функциональность.

В качестве центрального элемента платформы IMS выступает контроллер управления вызовами CSC. CSC может быть реализован в нескольких вариантах: P-CSC, S-CSC и I-CSC.

Элемент P-CSC включает в себя функциональность прокси-сервера и обеспечивает подключение SIP абонентов к сети.

Элемент S-CSC обеспечивает подключение абонентов к ресурсам сети напрямую или посредством взаимодействия с P-CSC, получая от них транзитную информацию о вызовах между абонентами.

В состав основных функций элемента I-/S-/P-CSC входят:

- определение следующего элемента (S-CSCF или сервер приложений), которому необходимо передать вызов, на основе данных из HSS;

- управление взаимодействием со сторонними сетями;

- трансляция SIP сообщений без преобразований на I-CSC стороннего IMS решения;

- прием со стороны внешней сети SIP сообщений и их трансляция к S-CSC своей доменной зоны;

- присвоение S-CSCF на основе данных из HSS в процессе регистрации абонента в сети или использования абонентом услуг, не требующих регистрации в сети, например, voice mail;

- прямое взаимодействие с BGC.

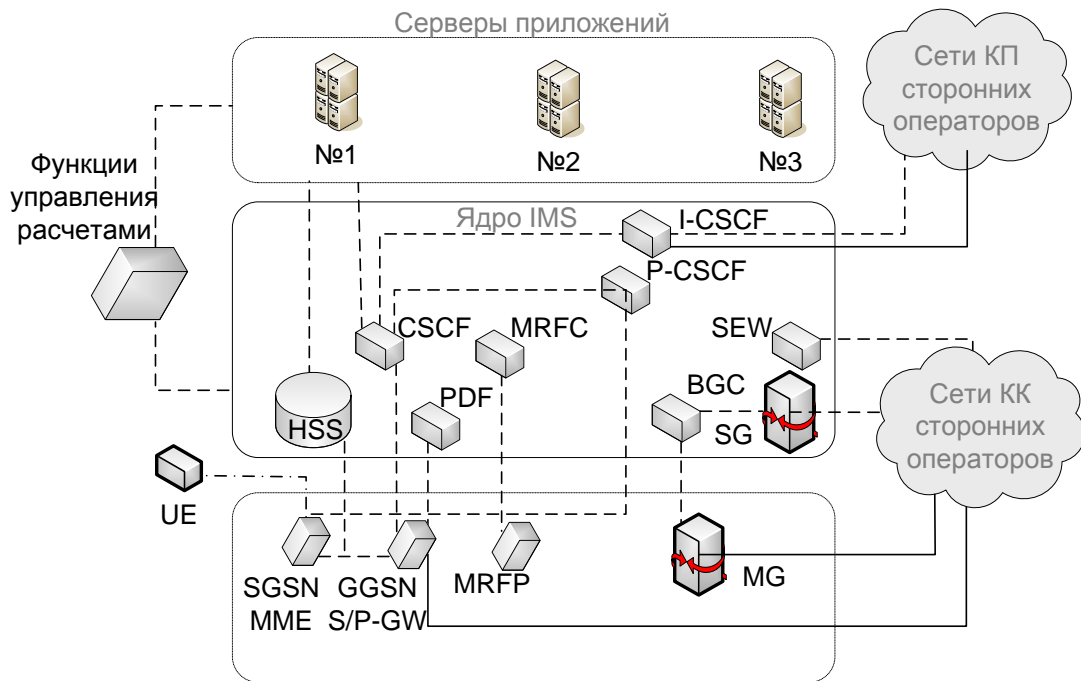


Рис.7. Общая схема взаимодействия на основе IMS

Таким образом, два из трех вариантов реализации центрального коммутационного элемента CSC (P-CSC, I-CSC) предполагают осуществление взаимодействия с потенциально опасными с точки зрения информационной безопасности участками сети (абонентские сети доступа и сети других операторов). Через подобные опасные участки сети возможны различные злонамеренные, либо случайные деструктивные воздействия на оборудование сети оператора. В связи с этим, функциональность P-CSC, I-CSC часто совмещают с обеспечением защиты от внешних угроз на уровнях L3-L7. Такие элементы обозначают как контроллеры SBC.

Элемент BGC (контроллер MGC, установленный на границе сети оператора) обеспечивает точку обмена сигнальной информацией между домашней сетью оператора и сетями других операторов, принимает решение о маршрутизации вызова в другую сеть исходя их информации, содержащейся в протоколе SIP, информации базы данных HSS и политик доступа, определяемых S-CSC данной доменной зоны.

Основными функциями BGC являются:

- определение маршрутизации вызовов на сторонние сети связи;

– выбор соответствующего MGC при маршрутизации вызова на ССОП;

– трансляция вызовов на сети сторонних операторов;

Контроллер мультимедиа ресурсов (MRC, Multimedia resource controller [86]) отвечает за выделение мультимедийных и сетевых ресурсов в сети.

Основными функциями контроллера MRC являются:

– управление потоками данных (управление транспортными и сигнальными шлюзами);

– преобразование информации от серверов AS и S-CSC в сообщения управления, используемые для конфигурации условий пропуска трафика для каждого конкретного вызова (например, выделение ресурсов сети для передачи данных пользователя).

3.4. Анализ основных сценариев предоставления услуг в конвергентных сетях

Классификация и краткий перечень услуг, предоставляемых в конвергентных сетях

В соответствии с рекомендацией МСЭ-Т Y.2013 [119] конвергентные услуги можно классифицировать по типам [58]:

– услуги на базе одной сети оператора. К их числу можно отнести такие услуги как: услуга «присутствия» (presence), Push-to-Talk on Cellular (для СПС, СПД), передача мультимедийных сообщений (MMS, Multimedia Message Service) (для СПС);

– услуги на базе множества разнотехнологичных сетей. К данному типу относятся такие услуги как: универсальный номер доступа (UAN, Universal Access Number), универсальная персональная связь (UPT, Universal personal telecommunications), универсальные услуги связи (UCS, Universal Communication Service). В данном случае, все услуги доступны всем типам сетей;

– услуги на базе взаимодействующих между собой конвергентных сетей связи. Примером услуги такого типа могут стать услуги видео или

аудио связи, услуга мгновенной передачи сообщений (IM, Instant Messaging), услуга переносимости номера мобильных сетях (MNP, Mobile Number Portability).

Определение бизнес-процессов основных сценариев предоставления конвергентных услуг

В сфере услуг связи за последние годы наблюдался их резкий рост, вызванный быстрым развитием технологий обработки и передачи информации. Вслед за технологиями под влиянием жесткой конкуренции на рынке телекоммуникационных услуг совершенствовались способы оптимизации сетей, интеграции технологий, различного телекоммуникационного оборудования, а так же алгоритмы создания конвергентный и мультисервисных сетей следующего поколения. В результате эволюции систем управления предприятий и совместной работы более четырехсот компаний по всему миру в рамках проекта TM Forum была разработана структурная модель бизнес-процессов компании поставщика телекоммуникационных услуг enhanced Telecom Operations Map (eTOM) [137]. Целью структурной модели является создание общего представления о бизнес процессах предприятия и их стандартизация. eTOM уже несколько лет негласно является стандартом в отрасли, поскольку в TMN входят практически все крупнейшие компании – операторы связи по всему миру, но официально статус мирового стандарта за eTOM не закреплён.

Структурная модель eTOM состоит из трех областей, классифицирующих бизнес-процессы компании по типам: «Стратегия, Инфраструктура, Продукт», «Операционные (производственные) процессы» и «Управление предприятием». Каждая из областей имеет свою многоуровневую структуру, позволяющую рассматривать бизнес-процессы компании как в общем (нулевой уровень модели), так и достаточно детализировано (третий уровень модели).

**ПРИЛОЖЕНИЕ 4. СООТВЕТСТВИЕ ПРОЦЕССОВ
СИГНАЛЬНОГО ОБМЕНА ПЕРЕХОДАМ И КЛАССАМ ЗАЯВОК
СеМО КС**

$(i, j), (k, l)$	Комментарий
Главная последовательность	
(1,2), (1,1)	Успешная отправка на E-UTRAN uU интерфейсе NAS Extended Service request, $q_{1,2}^{1,1} + q_{1,2}^{1,20} = 1$.
(1,2), (1,20)	Неуспешная отправка на E-UTRAN uU интерфейсе NAS Extended Service request, $q_{1,2}^{1,1} + q_{1,2}^{1,20} = 1$
(2,4), (1,1)	Успешная отправка на S1-MME интерфейсе NAS Extended Service request, $q_{2,4}^{1,1} + q_{2,4}^{1,20} = 1$.
(2,4), (1,20)	Неуспешная отправка на S1-MME интерфейсе NAS Extended Service request, $q_{2,4}^{1,1} + q_{2,4}^{1,20} = 1$.
(4,2), (1,2)	Успешная отправка на S1-MME интерфейсе S1-AP UE Context Modification request with CDFB indicator, $q_{4,2}^{1,2} + q_{4,2}^{1,20} = 1$.
(4,2), (1,20)	Неуспешная отправка на S1-MME интерфейсе S1-AP UE Context Modification request with CDFB indicator, $q_{4,2}^{1,2} + q_{4,2}^{1,20} = 1$.
(2,1), (2,2)	Успешная отправка NACC – RRC Connection release через Uu интерфейс E-UTRAN, $q_{2,1}^{2,2} + q_{2,1}^{2,20} = 1$.
(2,1), (2,20)	UE не получил и не обработал сообщение NACC – RRC Connection release через Uu интерфейс E-UTRAN, $q_{2,1}^{2,2} + q_{2,1}^{2,20} = 1$.
(1,3), (2,2)	Успешная отправка сообщения RR Channel Request через RAN сети 2G, $q_{1,3}^{2,2} + q_{1,3}^{1,20} = 1$.
(1,3), (2,20)	Неуспешная передача RR Channel Request через RAN сети 2G, $q_{1,3}^{2,2} + q_{1,3}^{1,20} = 1$.
(3,1), (2,3)	Успешная передача сообщений RR CH Immediately Assignment, $q_{3,1}^{2,3} + q_{3,1}^{2,20} = 1$.
(3,1), (2,20)	Неуспешная отправка сообщения RR CH Immediately Assignment, либо ошибки на стороне BSS, по причине которой команда о назначении канала не была сформирована, $q_{3,1}^{2,3} + q_{3,1}^{2,20} = 1$
(1,3), (3,4)	Успешная отправка сообщения о новой LA, в случае если процедура обновления местоположения необходима, $q_{1,3}^{3,4} + q_{1,3}^{3,20} + q_{1,3}^{3,5} = 1$.
(3,1), (4,4)	Успешное завершение процедуры Location Update и прием на стороне UE сообщения RR UA, $q_{3,1}^{4,4} + q_{3,1}^{4,20} = 1$.
(3,1), (4,20)	Не успешное завершение процедуры Location Update или ошибка на стороне UE при приеме сообщения RR UA,

	$q_{3,1}^{4,5} + q_{3,1}^{4,20} = 1.$
(1,3), (4,5)	Успешное завершение процедуры и приема на стороннее BSS сообщения SAMB +MM CM Service Request (После окончания процедуры Location Update), $q_{1,3}^{4,5} + q_{1,3}^{4,20} = 1.$
(1,3), (4,20)	Неуспешная отправка сообщения SAMB +MM CM Service Request (После окончания процедуры Location Update), $q_{1,3}^{4,5} + q_{1,3}^{4,20} = 1.$
(1,3), (3,5)	Успешная отсылка и приема на стороннее BSS сообщения SAMB + MM CM Service Request (в случае если обновление информации об LA не был необходим), $q_{1,3}^{3,4} + q_{1,3}^{3,20} + q_{1,3}^{3,5} = 1.$
(1,3), (3,20)	Суммарный процесс, описывающий ошибки на этапе обработки или генерации сообщения о новой LA, в случае если процедура обновления местоположения необходима и вероятности неуспешной транзакции SAMB +MM CM Service Request (в случае если обновление информации об LA не был необходим), $q_{1,3}^{3,4} + q_{1,3}^{3,20} + q_{1,3}^{3,5} = 1.$
(3,5), (5,5)	Успешная передача SCCP Connection request CM Service Request, $q_{3,5}^{5,5} + q_{3,5}^{5,20} = 1.$
(3,5), (5,20)	Неуспешная передача SCCP Connection request CM Service Request, $q_{3,5}^{5,5} + q_{3,5}^{5,20} = 1.$
(5,3), (5,6)	Успешная передача BSSMAP Cipher mode command, $q_{5,3}^{5,6} + q_{5,3}^{5,20} = 1.$
(5,3), (5,20)	Неуспешная передача BSSMAP Cipher mode command, $q_{5,3}^{5,6} + q_{5,3}^{5,20} = 1.$
(3,1), (6,6)	Успешная передача RR Cipher mode command, $q_{3,1}^{6,6} + q_{3,1}^{6,20} = 1.$
(3,1), (6,20)	Неуспешная передача RR Cipher mode command, $q_{3,1}^{6,6} + q_{3,1}^{6,20} = 1.$
(1,5), (6,7)	Успешная передача и обработка на стороне MSC сообщения Call Setup, $q_{1,5}^{6,7} + q_{1,5}^{6,20} = 1.$
(1,5), (6,20)	Неуспешная передача или обработка на стороне MSC сообщения Call Setup, $q_{1,5}^{6,7} + q_{1,5}^{6,20} = 1.$
(5,3), (7,8)	Успешная передача BSSMAP Assignment request, $q_{5,3}^{7,8} + q_{5,3}^{7,20} = 1.$
(5,3), (7,20)	Неуспешная передача BSSMAP Assignment request, $q_{5,3}^{7,8} + q_{5,3}^{7,20} = 1.$
(3,1), (8,8)	Успешная передача RR CH mode modify, $q_{3,1}^{8,8} + q_{3,1}^{8,20} = 1.$
(3,1), (8,20)	Неуспешная передача RR CH mode modify, $q_{3,1}^{8,8} + q_{3,1}^{8,20} = 1.$
(1,3), (8,9)	Успешная передача RR CH mode modify ack, $q_{1,3}^{8,9} + q_{1,3}^{8,20} = 1.$
(1,3), (8,20)	Неуспешная передача RR CH mode modify ack, $q_{1,3}^{8,9} + q_{1,3}^{8,20} = 1.$
(3,5), (9,9)	Успешная передача BSSMAP Assignment complete,

	$q_{3,5}^{9,9} + q_{3,5}^{9,20} = 1.$
(3,5), (9,20)	Неуспешная передача BSSMAP Assignment complete, $q_{3,5}^{9,9} + q_{3,5}^{9,20} = 1.$
(5,8), (9,9)	Успешная передача ISUP IAM, $q_{5,8}^{9,9} + q_{5,8}^{9,20} = 1.$
(5,8), (9,20)	Неуспешная передача ISUP IAM, $q_{5,8}^{9,9} + q_{5,8}^{9,20} = 1.$
(8,5), (9,10)	Ответ сети на IAM сообщением ISUP ACM и его корректная обработка, $q_{8,5}^{9,10} + q_{8,5}^{9,20} = 1.$
(8,5), (9,20)	Ошибка на сети, невозможности адресовать звонок вызываемому абоненту (либо любая иная причина, по которой не может быть выслан ACM), $q_{8,5}^{9,10} + q_{8,5}^{9,20} = 1.$
(5,1), (10,10)	Успешного формирования, доставки и последующей обработки на UE сообщения Alerting о приключении канала до вызываемого абонента и начале посылки сигнала КПВ, $q_{5,1}^{10,10} + q_{5,1}^{10,20} = 1.$
(5,1), (10,20)	Не успешное формирование, ошибка доставки или некорректная обработки на UE сообщения Alerting о приключении канала до вызываемого абонента и начале посылки сигнала КПВ, $q_{5,1}^{10,10} + q_{5,1}^{10,20} = 1.$
(1,9), (10,20)	Завершение процедуры установления соединения
Вторичная последовательность: Процесс окончания обработки сообщения	
(2,4), (11,11)	Успешная передача и генерация сообщения S1-AP UE Context Modification response message, $q_{2,4}^{11,11} + q_{2,4}^{11,20} = 1.$
(2,4), (11,20)	Неуспешная передача и генерация сообщения S1-AP UE Context Modification response message, $q_{2,4}^{11,11} + q_{2,4}^{11,20} = 1.$
(4,9), (11,20)	Процесс окончания обработки сообщения S1-AP UE Context Modification response message на MME
Вторичная последовательность: процедура освобождение контекста на eNodeB	
(2,4), (12,12)	Генерация на eNodeB и успешная доставка сообщения S1-AP: UE Context release request, $q_{2,4}^{12,12} + q_{2,4}^{12,20} = 1.$
(2,4), (12,20)	Неуспешная генерация на eNodeB или не успешная доставка сообщения S1-AP: UE Context release request, $q_{2,4}^{12,12} + q_{2,4}^{12,20} = 1.$
(4,2), (12,12)	Неуспешная доставка и последующая обработка сообщения UE Context release command на узле eNodeB, $q_{4,2}^{12,12} + q_{4,2}^{12,20} = 1.$
(4,2), (12,20)	Неуспешная доставка или последующая некорректная обработка сообщения UE Context release command на узле eNodeB, $q_{4,2}^{12,12} + q_{4,2}^{12,20} = 1.$
(2,4), (12,13)	Успешная доставка и обработка UE Context release complete сообщения на MME, $q_{2,4}^{12,13} + q_{2,4}^{12,20} = 1.$
(2,4), (12,20)	Неуспешная доставка или ошибка при обработке UE Context release complete сообщения на MME, $q_{2,4}^{12,13} + q_{2,4}^{12,20} = 1.$
(4,9), (13,20)	Окончание обработки последовательности

Вторичная последовательность: обновление информации о LA на MSC	
(3,5), (14,14)	Успешная передача сообщения MM LAU Request на MSC и завершение процедуры Location Update, $q_{3,5}^{14,14} + q_{3,5}^{14,20} = 1$.
(3,5), (14,20)	Неуспешная передача сообщения MM LAU Request на MSC, $q_{3,5}^{14,14} + q_{3,5}^{14,20} = 1$.
(5,9), (14,20)	Окончание обработки сообщения MM LAU Request на MSC и завершение процедуры Location Update
Вторичная последовательность: Отклик на на SAMB +MM CM Service Request	
(3,1), (15,15)	Успешная передача сообщения RR UA – отклика на SAMB +MM CM Service Request, $q_{3,1}^{15,15} + q_{3,1}^{15,20} = 1$.
(3,1), (15,20)	Неуспешная передача сообщения RR UA – отклика на SAMB +MM CM Service Request, $q_{3,1}^{15,15} + q_{3,1}^{15,20} = 1$.
(1,9), (15,20)	Окончание обработки сообщения RR UA – отклика на SAMB +MM CM Service Request.
Вторичная последовательность: процедура приостановления услуги пакетной передачи данных на время вызова	
(1,3), (16, 16)	Успешная передача сообщения Suspend request, $q_{1,3}^{16,16} + q_{1,3}^{16,20} = 1$.
(1,3), (16, 20)	Неуспешная передача сообщения Suspend request, $q_{1,3}^{16,16} + q_{1,3}^{16,20} = 1$.
(3,7), (16, 16)	Успешная передача и последующая обработка на SGSN сообщения BSSMAP Suspend request, $q_{3,7}^{16,16} + q_{3,7}^{16,20} = 1$.
(3,7), (16, 20)	Неуспешная передача или последующая некорректная обработка на SGSN сообщения BSSMAP Suspend request, $q_{3,7}^{16,16} + q_{3,7}^{16,20} = 1$.
(7,6), (16, 16)	Успешная передача и последующая обработка на S/P-GW сообщения Suspend request, $q_{7,6}^{16,16} + q_{7,6}^{16,20} = 1$.
(7,6), (16, 20)	Успешная передача или ошибочная последующая обработка на S/P-GW сообщения Suspend request, $q_{7,6}^{16,16} + q_{7,6}^{16,20} = 1$.
(6,7), (16, 17)	Успешная передача и последующая обработка на SGSN сообщения Suspend response, $q_{6,7}^{16,17} + q_{6,7}^{16,20} = 1$.
(6,7), (16, 20)	Неуспешная передача или последующая некорректная обработка на SGSN сообщения Suspend response, $q_{6,7}^{16,17} + q_{6,7}^{16,20} = 1$.
(7,3), (17, 17)	Успешная передача и последующая обработка на BSS сообщения BSSMAP Suspend response, $q_{7,3}^{17,17} + q_{7,3}^{17,20} = 1$.
(7,3), (17, 20)	Неуспешная передача или последующая некорректная обработка на SGSN сообщения BSSMAP Suspend response, $q_{7,3}^{17,17} + q_{7,3}^{17,20} = 1$.
(3,9), (17, 20)	Окончание обработки сообщения BSSMAP Suspend response на BSS, окончание последовательности приостановления услуг пакетной передачи данных для абонента в процессе голосового

	соединения.
Начало вторичной последовательности включения режима шифрования.	
(1,3), (18, 18)	Успешная передача сообщения RR Ciphering mode compleate, $q_{1,3}^{18,18} + q_{1,3}^{18,20} = 1$.
(1,3), (18, 20)	Неуспешная передача или последующая некорректная обработка на MSC сообщения RR Ciphering mode compleate, $q_{1,3}^{18,18} + q_{1,3}^{18,20} = 1$.
(3,5), (18, 18)	Успешная передача на MSC сообщения BSSMAP Ciphering mode compleate, $q_{3,5}^{18,18} + q_{3,5}^{18,20} = 1$.
(3,5), (18, 20)	Неуспешная передача на MSC сообщения BSSMAP Ciphering mode compleate, $q_{3,5}^{18,18} + q_{3,5}^{18,20} = 1$.
(5,9), (18,20)	Окончание последовательности.
Окончание обработки Call Proceeding	
(5,1), (19,19)	Успешная передача Call Proceeding, $q_{5,1}^{19,19} + q_{5,1}^{19,20} = 1$.
(5,1), (19,20)	Неуспешная передача Call Proceeding, $q_{5,1}^{19,19} + q_{5,1}^{19,20} = 1$.
(1,9), (19,20)	Окончание последовательности обработки Call Proceeding
Остальные переходы	
(i,9), (20,20), $\forall i \in \mathbb{N}$	Все заявки, помеченные как CL20 уходят на узел 9.

ПРИЛОЖЕНИЕ 5. СООТВЕТСТВИЕ КЛАССОВ И ТИПОВ СИГНАЛЬНЫХ СООБЩЕНИЙ ДЛЯ ПРОЦЕДУРЫ ИСХОДЯЩЕГО ВЫЗОВА В СЕТИ LTE С ИСПОЛЬЗОВАНИЕМ CSFB В СЕТЬ 2/3-ГО ПОКОЛЕНИЙ БЕЗ ПОДДЕРЖКИ PS HO

Класс	Тип сообщения
CL1	S : 1a. NAS Extended Service request : UE UE : 1a. NAS Extended Service request : eNodeB eNodeB: 1a. NAS Extended Service request : MME
CL2	MME : 1b. S1-AP UE Context Modification request with CDFB indicator: eNodeB eNodeB : 3a. NACC 3a.b. – RRC Connection release : UE UE : 6.1 RR CH Request : BSS
CL3	BSS : 6.2 RR CH Immediately Assignment : UE
CL4	UE: 6.b.1 SAMB +MM LAU Request: BSS BSS : 6b3. RR UA : UE
CL5	UE : 6.3 SAMB +MM CM Service Request : BSS BSS : 6.5. SCCP Connection request CM Service Request : MSC
CL6	MSC : 6.6. BSSMAP Cipher mode command : BSS BSS : 6.7 RR Cipher mode command : UE
CL7	UE : 6.10 CC Call Setup : MSC
CL8	MSC : 6.12 BSSMAP Assignment request : BSS BSS : 6.13 RR CH mode modify : UE
CL9	UE : 6.14 RR CH mode modify ack :BSS BSS : 6.15 BSSMAP Assignment compleate : MSC MSC : 6.16 ISUP IAM : PSTN
CL10	PSTN или NGN: 6.17 ISUP ACM : MSC MSC : 6.18 CC Alerting : UE UE : 6.18 CC Alerting : T
CL11	S : 1c. S1-AP UE Context Modification response message : MME MME : 1c. S1-AP UE Context Modification response message : T
CL12	S : 4. S1-AP: UE Context release request MME MME : 5.1 UE Context release command : eNodeB
CL13	eNodeB : 5.2 UE Context release compleate : MME MME: 5.2 UE Context release compleate : T
CL14	S : 6b2. MM LAU Request : MSC MSC : 6b2. MM LAU Request : T
CL15	S : 6.4. RR UA : UE UE : 6.4. RR UA : T
CL16	S : 7.a RR Suspend : BSS BSS : 7b.1 BSSMAP Suspend request : SGSN SGSN : 7b.2. Suspend Request : S/P-GW
CL17	S/P-GW : 7b.3. Suspend Response : SGSN SGSN : 7b.4 BSSMAP Suspend response : BSS BSS: 7b.4 BSSMAP Suspend response : T

CL18	S : 6.8 Cipher mode complete : BSS BSS : 6.9 BSSMAP Ciphering mode complete : MSC MSC : 6.9 BSSMAP Ciphering mode complete : T
CL19	S : 6.11 CC Call Proceeding : UE UE : 6.11 CC Call Proceeding : T
CL20	Выходящие из системы

**ПРИЛОЖЕНИЕ 6. ДОКУМЕНТЫ, ПОДТВЕРЖДАЮЩИЕ
ИСПОЛЬЗОВАНИЕ РЕЗУЛЬТАТОВ ДИССЕРТАЦИИ**

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2013661814

**«Программа обработки данных SIP трафика, определения
последовательностей сообщений в рамках установления
соединения и частоты их появления»**

Правообладатель: **Федеральное государственное бюджетное
учреждение науки Институт проблем информатики Российской
академии наук (ИПИ РАН) (RU)**

Авторы: **Разумчик Ростислав Валерьевич (RU),
Углов Иван Валерьевич (RU)**

Заявка № **2013619644**
Дата поступления **24 октября 2013 г.**
Дата государственной регистрации
в Реестре программ для ЭВМ **16 декабря 2013 г.**

Руководитель Федеральной службы
по интеллектуальной собственности

Б.П. Симонов



УТВЕРЖДАЮ

Директор Департамента
эксплуатации сети
Макро-региона «Москва»
ОАО «Мобильные ТелеСистемы»


_____ Медведев В.А.
12.01.2012 г.



А К Т

о внедрении результатов кандидатской диссертационной работы
Углова Ивана Валерьевича

Комиссия в составе:

председатель комиссии:

Евгеев А.В. – начальник отдела эксплуатации коммутационной подсистемы, департамент эксплуатации сети, макро-регион Москва;

члены комиссии:

Ромаченко Д.М. – руководитель группы модернизации и расширения, отдел эксплуатации коммутационной подсистемы, департамент эксплуатации сети, макро-регион Москва;

Жуков Ф.В. – ведущий технический специалист группы модернизации и расширения, отдел эксплуатации коммутационной подсистемы, департамент эксплуатации сети, макро-регион Москва;

составили настоящий акт, что результаты диссертационной работы Углова И.В. в части:

- организации процессов мониторинга трафика сети;
- методов автоматизации обработки данных с сетевых узлов;
- организации систем анализа и последующей оценки качества работы узлов сетей и ее сегментов;
- методов анализа качественных характеристик сетевого трафика;

были использованы при организации системы мониторинга и анализа качества сети ОАО «МТС». Решение зарегистрировано в рамках внутренней инновационной программы «Фабрика идей» под номером FI-899/11.

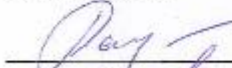
Использование указанных результатов позволяет: повысить качество предоставляемых услуг связи, сократить время и затраты на проведение мероприятий по поиску и устранению неполадок на сети.

Председатель комиссии



_____ Евгеев А.В.

Члены комиссии:



_____ Ромаченко Д.М.



_____ Жуков Ф.В.



УТВЕРЖДАЮ

Директор департамента сервисных и фиксированных сетей Технического блока Корпоративного центра Группы МТС

 Чекарёва Е.В.

« 15 » декабря 2014 г.

А К Т

**об использовании результатов диссертационной работы
Углова Ивана Валерьевича**

Настоящий акт составлен о том, что разработанные в кандидатской диссертации Углова И.В. на тему «Разработка обобщенных аналитических моделей процесса сигнального обмена в конвергентной сети» следующие материалы:

- аналитические модели работы узлов конвергентной сети в виде многолинейных систем массового обслуживания в непрерывном времени с поступающим пуассоновским потоком;
- методы и алгоритмы расчета стационарного распределения вероятностей числа заявок в системах массового обслуживания;
- формулы для расчета вероятностно-временных характеристик функционирования указанных моделей и результаты численного анализа;
- аналитические модели работы конвергентной сети LTE/UMTS/GSM с приоритизацией трафика в виде сети массового обслуживания с дисциплинами обработки заявок на узлах общего вида;
- методы изучения характеристик сигнальных потоков конвергентных сетей операторского класса;
- программные средства на языке Matlab, Unix Shell, C++, обеспечивающие обработку потоков сигнальной информации для определения свойств и характеристик исследуемых потоков, использованы для решения производственных задач в рамках развития сети ОАО «МТС», в целях анализа и обеспечения качества функционирования сети, при разработке внутрикорпоративных стандартов.

Использование указанных результатов позволило:

- построить модели функционирования фрагментов сетей для целей анализа качества их работы в рамках предоставления услуг на основе протоколов различных типов;
- провести анализ качественных характеристик функционирования этих фрагментов;
- разработать практические рекомендации по построению конвергентных сетей с распределенной архитектурой;
- осуществить оценку функционирования современного коммутационного оборудования.

Данные результаты имеют существенное значение для дальнейшего развития средств расчета и проектирования сети ОАО «МТС».

Начальник отдела МГ/МН и транзитной сигнальной сети департамента сервисных и фиксированных сетей Технического блока Корпоративного центра Группы МТС

 Хижняк Д.В.



МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
"Московский государственный технический университет
радиотехники, электроники и автоматики"
МГТУ МИРЭА

УТВЕРЖДАЮ
Проректор МГТУ МИРЭА
по учебной работе



Панков В.Л.

« 20 » декабря 2014 г.

АКТ

**о внедрении результатов диссертационной работы
Углова Ивана Валерьевича
на тему «Разработка обобщенных аналитических моделей процессов
сигнального обмена в конвергентной сети»**

Результаты диссертационной работы Углова И.В. на тему «Разработка обобщенных аналитических моделей процессов сигнального обмена в конвергентной сети», представленной на соискание ученой степени кандидата технических наук, были использованы в период с 2013 по 2014 г. в учебном процессе кафедры «Компьютерная и информационная безопасность» МГТУ МИРЭА, а именно:

1. По материалам диссертационной работы разработано четыре лабораторных работы по курсу «Конструктивно-технологическое обеспечение средств связи», в которых проводится исследование процессов

сигнального обмена трафиком в фиксированных и беспроводных сетях передачи данных.

2. Подготовлено методическое пособие «Основы построения конвергентных сетей» по курсу «Конструктивно-технологическое обеспечение средств связи». Указанное пособие содержит полученные соискателем результаты анализа архитектур фиксированных и подвижных сетей, механизмов взаимодействия их функциональных элементов и сценариев установления соединений в них

Разработанные Уловым И.В. лабораторные работы и программно-алгоритмическое обеспечение лабораторных работ являются существенным вкладом в работу учебной лаборатории и в учебный процесс в целом кафедры «Компьютерная и информационная безопасность» в соответствии с современными требованиями.

Зав. кафедрой «Компьютерная
и информационная безопасность»
МГТУ МИРЭА

к.т.н., доцент  С.М. Ярлыкова