

Федеральное государственное автономное образовательное учреждение
высшего образования «Национальный исследовательский университет
«Высшая школа экономики»

На правах рукописи

Лежнев Евгений Владимирович

**АВТОМАТИЗАЦИЯ НИЗКОУРОВНЕВОГО МОДЕЛИРОВАНИЯ
СЕТЕЙ НА КРИСТАЛЛЕ**

Резюме

диссертации на соискание ученой степени
кандидата технических наук

Научный руководитель:
кандидат технических наук,
доцент
Романов Александр Юрьевич

Москва – 2022

Актуальность работы

Постоянный рост сложности вычислительных задач, а также увеличение объема необходимых для вычислений данных становятся важными факторами, влияющими на индустрию разработки вычислительных систем. Отражением этого является переход от одноядерных к многоядерным процессорам, а также от однопроцессорных вычислительных комплексов к многопроцессорным. В последние несколько лет разработка процессоров и систем на кристалле (System-on-Chip, SoC, СнК), а также мультипроцессорных систем на кристалле (Multi Processor System-on-Chip, MPSoC) вышла на новый качественный уровень, представляющий собой объединение на одном чипе большого количества процессоров (от 100 и более) и соединение их в единую вычислительную сеть на кристалле (Network-on-Chip, NoC, СтнК). Это потребовало разработки и развития новых подходов к проектированию вычислительных систем.

В классическом процессе проектирования вычислительных систем можно выделить следующие этапы: разработка технического задания; построение высокоуровневых моделей на языках высокого уровня; разделение системы на функциональные блоки, и уточнение их характеристик; разработка поведенческих моделей на языках описания аппаратуры; прототипирование и верификация разрабатываемой системы; адаптация низкоуровневых моделей под требования производителя чипа; оценка физических характеристик реализуемой системы; производство чипа; корпусирование устройства. Появление новых подходов к проектированию вычислительных систем приводит к изменениям и самого процесса проектирования, в котором наиболее важным становится этап моделирования разрабатываемой системы.

На этапе моделирования при проектировании СтнК можно выделить два типа моделирования, которые позволяют в полной мере провести исследование: высокоуровневое и низкоуровневое моделирование. Благодаря им можно на раннем этапе оценить особенности работы разрабатываемой

системы. В большинстве случаев с помощью высокоуровневых моделей исследуется процесс передачи данных в СтнК в общем виде и невозможно учесть все особенности ее аппаратной реализации. При высокоуровневом моделировании изменения в описание модели СтнК, как правило, не вносят, так как модель является аппаратно-независимой. При низкоуровневом моделировании, когда вся система представлена в виде схемы уровня регистровых передач, исследование системы может происходить при различных архитектурных изменениях. К таким изменениям можно отнести разделение СтнК на структурные блоки, описывающие ее отдельные компоненты, изменение метода обмена информационными пакетами между компонентами СтнК, изменение самих описаний компонентов и т.д.

В процессе разработки СтнК отдельной важной задачей является разработка ее подсистемы связи. Процесс проектирования подсистемы связи СтнК в общем виде состоит в определении множества различных характеристик таких элементов, как: топология соединений маршрутизаторов, алгоритм маршрутизации, структура маршрутизаторов, методы управления и арбитража потоками данных в сети. В общий состав СтнК также входят следующие компоненты: вычислительные узлы и внешняя периферия. Но их влиянием на работу сети на первоначальном этапе можно пренебречь, т.к. они не влияют на подсистему связи, а только генерируют данные, которые подсистема связи должна передать. Соединение IP-блоков СтнК с ее подсистемой связи происходит только на уровне маршрутизаторов сети.

Для анализа влияния принятых архитектурных решений на производительность проектируемой СтнК, требуется проводить моделирование. Этап низкоуровневого моделирования является самым трудозатратным в проектировании, как в части описания модулей, так и в части получения результатов моделирования, что значительно больше по сравнению с высокоуровневым моделированием. В зависимости от вида данных, которые необходимо получить, на этом этапе могут применяться

различные виды моделирования: математическое, поведенческое. Подготовка модели обычно заключается в разработке большого количества однотипного программного кода на языках описания аппаратуры, в ходе которого происходит описание поведения компонентов СтнК, а также задаются различные параметры модели. Ввиду ограниченности ресурсов ПЛИС не всегда возможно прототипирование СтнК с требуемым количеством узлов. Разделение исходной модели на несколько моделей меньшей сложности, которые исследуют отдельные компоненты СтнК, позволяет существенно сократить время получения результатов моделирования. Использование специализированных САПР для проектирования и моделирования СтнК позволяют снизить трудозатраты на разработку и моделирование СтнК. При этом такие САПР могут автоматизировать как весь цикл разработки СтнК, так и отдельные его части.

Степень разработанности темы

Значительный вклад в развитие теории проектирования СтнК сделали такие известные зарубежные ученые, как W.J. Dally, D. Debb, I.H. Wang. Среди трудов отечественных исследователей можно выделить работы С.О. Быкова, Н.В. Вараскина, Е.А. Кичина, А.С. Кожина, В.В. Корнеева, В.О. Костенко, А.В. Лаврова, Э.А. Монаховой, Ю.А. Недбайло, Н.Ю. Полякова, Ю.Х. Сахина, В.В. Тихорского, С.Р. Тумковского, А.Ю. Романова, Д.И. Шпагилева и др.

Уже сейчас СтнК получили широкое распространение и применяются во многих проектах. Например, проекты таких компаний, как Intel, в которых применяются кольцевые соединения ядер в процессорах начиная с семейства Ivy Bridge, а также в процессорах Xeon 6-го поколения; компания МЦСТ использует для соединения вычислительных ядер и организации передачи данных между ними СтнК; фирма Tiler внедряет технологии СтнК в свои разработки больших многоядерных чипов. Также стоит выделить разработки фирмы Arteris, которая создает инструменты проектирования и внедрения СтнК для разработки устройств, что нашло применение в микросхемах

процессоров фирм Qualcomm и Samsung. Одной из последних тенденций в области разработки процессоров является создание крупноразмерных чипов, например, процессора Cerebras Wafer Scale Engine, который содержит более 850.000 вычислительных ядер, соединенных с помощью подсистемы связи, основанной на топологии mesh, где использование СтнК является единственным решением, позволяющим обеспечивать необходимую скорость передачи данных между вычислительными ядрами. Также стоит упомянуть проект Esperanto Technologies, в рамках которого реализовано 1088 энергоэффективных 64-битных RISC-V ET-Minion векторных/тензорных ядер в чипе ET-SoC-1.

Существуют также многочисленные академические научные группы, а количество исследований и публикаций по СтнК с каждым годом увеличивается. При этом следует отметить, что из-за новизны направления, большинство исследований разноплановые и не связанные между собой. Данная проблема поднимается в работах Ахмеда Бен Ачбаллаха, в которых проводится обзор низкоуровневых моделей по критериям синтеза аппаратного описания, по исследуемым компонентам СтнК и выводится утверждение, что результаты моделирования с использованием различных моделей не согласованы между собой по форматам представления данных. Наблюдается недостаток специализированных САПР для разработки СтнК и низкоуровневого моделирования в частности. Эта проблема продемонстрирована в работах Вим Мееуса, в которых проводится анализ инструментов моделирования СтнК по различным критериям: языку реализации модели, возможности добавления сторонних компонентов, возможности генерировать RTL-описание, сложности работы с инструментом моделирования. Автором описан ряд недостатков известных инструментов моделирования, которые усложняют их использование. Например, часто отсутствует стандартизация при вводе описания исследуемой системы, так что разработчикам требуется дополнительное обучение для разработки с

использованием конкретного инструмента моделирования СтнК. Часто для исследования проектируемой СтнК и оптимизации ее архитектуры требуется глубокая модификация исходного кода. Некоторые инструменты разработаны для одной области применения (например, только для описания потоков данных или логики управления). Это также является существенным недостатком, поскольку разработчикам необходимо использовать несколько инструментов моделирования и вручную адаптировать получаемые от них данные под свои задачи. Описанные выше проблемы определяют объект и предмет данного исследования.

Объектом исследования являются сети на кристалле.

Предметом исследования является автоматизация низкоуровневого моделирования сети на кристалле.

Цель и задачи исследования

Целью исследования является уменьшение трудозатрат на подготовку низкоуровневой модели СтнК, сокращение времени моделирования, а также увеличение размеров исследуемой СтнК, за счет автоматизации процессов синтеза модели и анализа полученных результатов, путем разработки САПР для низкоуровневого моделирования СтнК.

Для достижения поставленной цели в работе решаются следующие задачи:

- Анализ принципов организации и работы низкоуровневых моделей СтнК, их типовых структур и решаемых задач;
- Разработка метода компьютерного моделирования, включающего в себя создание специализированных, предварительно настроенных на решение конкретных задач, низкоуровневых моделей СтнК, позволяющего выполнять раздельное моделирование и оценку ресурсов, занимаемых как подсистемой связи отдельно от остальных компонентов СтнК, так и сети в целом;
- Реализация универсального, т.е. требующего минимальной настройки, интерфейса для подключения компонентов СтнК к ее подсистеме связи для

сокращения трудозатрат по модификации подключаемых компонентов, а также для согласования формата данных между подключаемым компонентом и подсистемой связи СтнК;

- Разработка метода редуцирования модели СтнК для исследования сетей с большим количеством узлов;
- Разработка специализированного транслятора кода алгоритмов маршрутизации СтнК с языка C# на HDL;
- Разработка методики автоматизированного сквозного проектирования СтнК, заключающейся в автоматической параметризации низкоуровневой модели, с целью согласования между собой результатов моделирования СтнК на всех этапах проектирования и автоматизации процесса моделирования;
- Разработка новой архитектуры САПР, которая отличается от существующих объединением процессов синтеза и анализа, применительно к области проектирования СтнК, за счет возможности автоматизированной параметрической модификации ядра модели и дальнейшего анализа полученных результатов в сравнении с описанием модели на ЯВУ;
- Разработка алгоритмического обеспечения САПР для синтеза комплексных низкоуровневых моделей подсистем связи СтнК под различные задачи проектирования, с помощью разработки специализированного транслятора, а также автоматизации процесса генерации таких моделей;
- Разработка прикладного программного обеспечения, реализующего САПР, включающего информационное (база данных результатов моделирования), математическое (аппроксимационные модели предсказания результатов моделирования), лингвистическое (низкоуровневые модели на языке Verilog), методическое (разработанные методики моделирования), техническое (действующий прототип на ПЛИС) и программное обеспечение (прикладное ПО САПР и вспомогательные утилиты) для разработки низкоуровневых моделей СтнК;

– Апробация предложенной методики автоматизированного сквозного проектирования СтнК на примере решения задач оценки влияния топологии, алгоритма маршрутизации и других параметров на работу СтнК в целом для проведения ее сравнительного анализа с классическим циклом низкоуровневого моделирования.

Методология и методы исследования

В работе использованы методы моделирования на языках описания аппаратуры и имитационного моделирования для получения результатов работы модели СтнК и выявления зависимостей при ее параметрическом изменении; методы структурного и функционального программирования при описании моделей на языке Verilog, а также при разработке программного обеспечения САПР; методы статистической обработки, анализа и интерполяции данных.

Личный вклад автора

Все результаты и положения, выносимые на защиту, получены автором лично. На основе проведенного обзора и анализа предметной области, изучения источников литературы и особенностей создания современных низкоуровневых моделей СтнК, автором выявлена проблема параметрической генерации низкоуровневых моделей СтнК для множественного моделирования в части отсутствия систем автоматизации процесса создания однотипных низкоуровневых моделей.

Автором лично сформулирован объект исследования как процесс, в котором выявлена проблема и раскрыт предмет исследования как разработка средств для решения научной задачи автоматизации создания низкоуровневых моделей подсистемы связи СтнК. Поставлена цель работы и определен логически связанный комплекс задач для ее достижения, при решении которых автором лично получены новые научные результаты, имеющие

важное прикладное и практическое значение для области проектирования СтнК.

Личный вклад автора также отражен в достаточном количестве публикаций в рецензируемых и индексируемых изданиях, в части которых соискатель является основным автором.

Основные результаты работы:

1. Новый подход к низкоуровневому моделированию подсистемы связи СтнК, заключающийся в раздельном моделировании подсистемы связи и вычислительных IP-блоков, а также в редуцировании вычислительных IP-блоков при комплексном моделировании СтнК, который позволил проводить исследование СтнК с количеством узлов до 200 с использованием любых топологий и алгоритмов маршрутизации;

2. Новая методика синтеза низкоуровневых моделей СтнК на основе прототипа ядра низкоуровневой модели подсистемы связи СтнК, которая позволила автоматически синтезировать параметрические модели СтнК;

3. Разработанный метод редуцирования модели СтнК, дал возможность проводить моделирование подсистемы связи СтнК без необходимости использования вычислительных IP-блоков для генерации пакетов данных;

4. Разработанная САПР позволила автоматизировать параметрический синтез низкоуровневых моделей подсистемы связи СтнК, облегчить подключение компонентов СтнК к подсистеме связи, обеспечить унифицированное использование специализированных алгоритмов, как в высокоуровневых, так и в низкоуровневых моделях, обработку и хранение данных моделирования, а также обеспечить непрерывность процесса проектирования СтнК;

5. Разработанные новые подходы и методы автоматизации обеспечивают единую парадигму сквозного проектирования СтнК от параметрического описания до прототипа на ПЛИС, что позволило ускорить

процесс разработки подсистемы связи СтнК до 15 раз и обеспечить согласованность результатов моделирования на всех этапах проектирования.

Достоверность и обоснованность полученных результатов подтверждается корректностью постановки задачи и применяемых методов исследования, согласованностью результатов экспериментов с применением современных и широко апробированных средств проектирования и математического моделирования.

Достоверность также подтверждается апробацией основных результатов работы на протяжении ряда лет на многих всероссийских и международных конференциях и публикациями в общедоступных рецензируемых изданиях, индексируемых в международных и отечественных базах цитирования WoS, Scopus и РИНЦ.

Научная новизна исследования заключается в том, что:

- Разработан новый подход к низкоуровневому моделированию подсистемы связи СтнК, отличающийся от известных тем, что производится раздельное моделирование компонентов СтнК, а также редуцирование вычислительных IP-узлов при комплексном моделировании СтнК, и который позволяет проводить оценку аппаратных затрат на реализацию СтнК на основе любых топологий и алгоритмов маршрутизации с количеством узлов до 200;
- Предложена новая методика синтеза низкоуровневых моделей СтнК, отличающаяся от известных тем, что в ней применяется автоматизированный параметрический синтез моделей СтнК под заданные требования на основе использования прототипа ядра низкоуровневой модели подсистемы связи СтнК, которая дала возможность ускорить процесс отладки работы подсистемы связи СтнК до 15 раз;
- Разработан новый метод редуцирования низкоуровневой модели СтнК за счет замещения вычислительных IP-блоков модулем генерации

пакетов данных, что дало возможность проводить моделирование подсистемы связи СтнК с увеличенным количеством узлов в сети в 2,5 раза до 200 узлов.

Теоретическая значимость и практическая полезность

Теоретическая значимость данного исследования состоит в развитии теории автоматизации проектирования СтнК, а также средств моделирования подсистемы связи и методов сквозного проектирования СтнК.

Практическая значимость полученных результатов заключается в том, что:

- На основе обзора и анализа литературных источников, обзора способов моделирования СтнК, типов моделей, их структуры и исследуемых ими характеристик, проведен анализ предметной области, что позволило сформулировать научную задачу диссертационного исследования;
- Разработана низкоуровневая модель подсистемы связи СтнК, которая дала возможность проводить моделирование передачи данных в СтнК с увеличенным количеством узлов до 200;
- Разработан универсальный интерфейс связи, который дает возможность подключать любые компоненты СтнК, упростить процесс согласования данных между ними и подсистемой связи СтнК за счет устранения необходимости модификации описания этих элементов;
- С учетом особенностей структуры описания компонентов подсистемы связи СтнК разработан транслятор для перевода высокоуровневых описаний специализированных алгоритмов в низкоуровневое представление (на HDL), что позволило проводить сравнительный анализ влияния различных топологий алгоритмов маршрутизации, описанных на языках высокого уровня, на функционирование подсистемы связи СтнК, описанной низкоуровневой моделью;
- На основе разработанного подхода к разделению моделированию компонентов СтнК в рамках созданной САПР реализована возможность прототипирования подсистемы связи отдельно от остальных компонентов

СтнК, за счет чего удалось ускорить процесс отладки работы подсистемы связи до 15 раз.

Положения, выносимые на защиту:

- Новый подход к низкоуровневому моделированию подсистемы связи СтнК позволяет проводить исследование СтнК с количеством узлов до 200 с использованием любых топологий и алгоритмов маршрутизации;
- Новая методика синтеза низкоуровневых моделей СтнК на основе прототипа ядра низкоуровневой модели подсистемы связи СтнК позволяет автоматически синтезировать параметрические модели СтнК, за счет чего процесс отладки работы подсистемы связи СтнК ускорился до 15 раз;
- Метод редуцирования модели СтнК позволяет проводить моделирование подсистемы связи СтнК без необходимости использования реальных вычислительных IP-блоков для генерации пакетов данных;
- Разработанная САПР позволяет автоматизировать параметрический синтез низкоуровневых моделей подсистемы связи СтнК, облегчает подключение компонентов СтнК к подсистеме связи, обеспечивает одинаковое поведение специализированных алгоритмов, как в высокоуровневых, так и в низкоуровневых моделях, обработку и хранение данных моделирования, а также обеспечивает непрерывность процесса проектирования СтнК;
- Новые подходы и методы автоматизации сквозного проектирования СтнК от параметрического описания до прототипа на ПЛИС позволяют ускорить процесс разработки подсистемы связи СтнК до 15 раз и обеспечить согласованность результатов моделирования на всех этапах проектирования.

Апробация результатов

Предлагаемые в диссертации новые результаты внедрены в учебный процесс МИЭМ НИУ ВШЭ, нашли применение в проектных работах МИЭМ

НИУ ВШЭ, в исследовательских проектах ЦФИ НИУ ВШЭ, а также в гранте РФФИ, что подтверждается соответствующими актами:

– Акт внедрения в учебный процесс департамента компьютерной инженерии Московского института электроники и математики Национального исследовательского университета «Высшая школа экономики». Внедрение результатов диссертационной работы позволило использовать в учебной и научной деятельности студентов департамента компьютерной инженерии современные теоретические и практические разработки в области моделирования сетей на кристалле (СтнК). Разработанные низкоуровневые модели дали возможность студентам проводить изучение СтнК различных конфигураций, применять теоретические знания в области теории графов в практике исследования и разработки алгоритмов маршрутизации для различных топологий, использовать инструменты анализа RTL-схем и скорости работы устройств в САПР Quartus Prime. Разработанная в рамках диссертационной работы модель подсистемы связи обеспечивает возможность студентам изучать методы ко-симуляции, работы с TCL-скриптами, программной генерации кода на Verilog и работы с САПР Quartus Prime без использования графического интерфейса;

– Акт об использовании результатов диссертации в проектной работе «Аппаратно-программный комплекс для обучения в режиме удаленного доступа к лабораторному оборудованию». Результаты диссертационной работы используются в части лабораторных работ, адаптированных для выполнения на оборудовании лаборатории САПР в удаленном режиме и без использования отладочных плат за счет использования виртуализации, что открывает новые возможности для изучения сетей на кристалле.

– Исследовательский проект ЦФИ НИУ ВШЭ «Синтез циркулянтных топологий для применения в сетях на кристалле», рег. № НИОКТР АААА-А18-118051690145-1, 01.02.2018–29.12.2018;

- Исследовательский проект ЦФИ НИУ ВШЭ «Моделирование сетей на кристалле с подсистемой связи на основе циркулянтных топологий», рег. № НИОКТР АААА-А19-119061490099-1, 01.02.2019–31.12.2019;
- Исследовательский проект ЦФИ НИУ ВШЭ «Разработка гибридной модели для проектирования и симуляции сетей на кристалле», рег. № НИОКТР АААА-А20-120070390136-2, 03.02.2020-31–31.12.2020;
- Исследовательский проект ЦФИ НИУ ВШЭ «Разработка алгоритмов маршрутизации в сетях на кристалле», рег. № НИОКТР 121051100322-4, 01.02.2021–31.12.2021;
- Грант РФФИ «Самоорганизация в сетях на кристалле: принципы, модели, алгоритмы маршрутизации, программы, производственные технологии», соглашение № 22-29-00979, 29.12.2021 – 29.12.2023.

Основные положения диссертационной работы и ее отдельные разделы были изложены в виде докладов и обсуждались на 8 международных конференциях:

- Научно-технической конференции студентов, аспирантов и молодых специалистов НИУ ВШЭ им. Е.В. Арменского, г. Москва, НИУ ВШЭ, 2019 г. (тема доклада: «Разработка HDL модели сети на кристалле на основе маршрутизатора с одной очередью»);
- Международной конференции «Актуальные проблемы системной и программной инженерии (АПСПИ)», г. Москва, НИУ ВШЭ, 2019 г. (тема доклада: «Modification of the BookSim simulator for modeling networks-on-chip based on two dimensional circulant topologies»);
- Международной конференции «Компьютерное моделирование в физике и не только (CSP)», г. Москва, НИУ ВШЭ, 2018 г. (тема доклада: «Routing in Networks on Chip with Multiplicative Circulant Topology»);
- Международной конференции «Компьютерное моделирование в физике и не только (CSP)», г. Москва, НИУ ВШЭ, 2019 г. (тема доклада: «Development of multiprocessor system-on-chip based on soft processor cores schoolMIPS»);

- Московском семинаре по электронным и сетевым технологиям (MWENT), г. Москва, НИУ ВШЭ, 2020 г. (тема доклада: «Analytical Routing Algorithm for Networks-on-Chip with the Three-dimensional Circulant Topology»);
- Международной научно-технической конференции «Автоматизация» (RusAutoCon), г. Сочи, 2019 г. (тема доклада: «Routing in Networks-on-Chip with Circulant Topology with Three Generatrices of Type $C(N;S_1,S_2,S_3)$ »);
- Международной научно-технической конференции «Автоматизация» (RusAutoCon), г. Сочи, 2021 г. (тема доклада: «Development of Automation System for HDL Modeling of the Communication Subsystem for Networks-on-Chip»);
- Всероссийской научно-технической конференции «Проблемы разработки перспективных микро- и наноэлектронных систем (МЭС)», г. Зеленоград, ИПИМ РАН, 2021 г. (тема доклада: «Автоматизация низкоуровневого моделирования сетей на кристалле»).

Список опубликованных статей, отражающих основные результаты диссертации

Результаты диссертационной работы отражены в 14 печатных работах, из них 10 статей – в изданиях, включенных в международные наукометрические базы (WoS, Scopus).

Работы, опубликованные автором в изданиях, включенных в международные наукометрические базы Web of Science и Scopus:

1. Monakhov, O.G. Adaptive Dynamic Shortest Path Search Algorithm in Networks-on-Chip Based on Circulant Topologies / O.G. Monakhov, E.A. Monakhova, A.Yu. Romanov, A.M. Sukhov, E.V. Lezhnev // IEEE Access. – IEEE, 2021. – Vol. 9. – P. 160836–160846. (Q1, WoS).
2. Monakhova, E.A. Shortest Path Search Algorithm in Optimal Two-Dimensional Circulant Networks: Implementation for Networks-on-Chip /

E.A. Monakhova, A.Yu. Romanov, E.V. Lezhnev // IEEE Access. – IEEE, 2020. – Vol. 8. – P. 215010–215019. (Q1, WoS).

3. Romanov, A.Yu. Development of routing algorithms in networks-on-chip based on two-dimensional optimal circulant topologies / A.Yu. Romanov, E.V. Lezhnev, A.Yu. Glukhikh, A.A. Amerikanov // Heliyon. – Elsevier, 2020. – Vol. 6. – No. 1. – P. 1–8. (Q1, Scopus, WoS).

4. Romanov, A.Yu. Analysis of Approaches for Synthesis of Networks-on-chip by Using Circulant Topologies / A.Yu. Romanov, A.A. Amerikanov, E.V. Leghnev // Journal of Physics: Conference Series. – IOP Publishing, 2018. – Vol. 1050. – No. 1. – P. 1–12. (Q4, Scopus).

5. Romanov, A.Yu. Modification of the BookSim simulator for modeling networks-on-chip based on two dimensional circulant topologies / A.Yu. Romanov, E.V. Lezhnev, A.A. Amerikanov // Proceedings of the 6th International Conference Actual Problems of System and Software Engineering (APSSE). – Moscow: HSE University, 2019. – Vol. 2514. – Ch. 107. – P. 182–192. (Scopus).

6. Romanov, A.Yu. Routing in Networks-on-Chip with Circulant Topology with Three Generatrices of Type $C(N;S1,S2,S3)$ / A.Yu. Romanov, M.V. Sidorenko, E.V. Lezhnev // International Russian Automation Conference (RusAutoCon). – IEEE, 2019. – P. 1–6. (Scopus).

7. Schegoleva, M.A. Routing in Networks on Chip with Multiplicative Circulant Topology / M.A. Schegoleva, A.Yu. Romanov, E.V. Lezhnev, A.A. Amerikanov // Journal of Physics: Conference Series. – IOP Publishing, 2019. – Vol. 1163. – No. 1. – P. 1–7. (Q4, Scopus).

8. Ryazanova, A.E. Development of multiprocessor system-on-chip based on soft processor cores schoolMIPS / A.E. Ryazanova, A.A. Amerikanov, E.V. Lezhnev // Journal of Physics: Conference Series. – IOP Publishing, 2019. – Vol. 1163. – No. 1. – P. 1–7. (Q4, Scopus).

9. Lezhnev, E.V. Development of Automation System for HDL Modeling of the Communication Subsystem for Networks-on-Chip / E.V. Lezhnev //

International Russian Automation Conference (RusAutoCon). – IEEE, 2021. – P. 780–784. (Scopus).

10. Monakhova, E.A. Analytical Routing Algorithm for Networks-on-Chip with the Three-dimensional Circulant Topology / E.A. Monakhova, O.G. Monakhov, A.Yu. Romanov, E.V. Lezhnev // Moscow Workshop on Electronic and Networking Technologies (MWENT). – Moscow: 2020. – No. 9067418. – P. 1–6. (Scopus).

Работы, опубликованные автором в других научных изданиях:

11. Прилепко, П.М. Модификация высокоуровневой модели NoCModel 2.0 для моделирования сетей на кристалле с циркулянтными топологиями / П.М. Прилепко, А.Ю. Романов, Е.В. Лежнев // Проблемы разработки перспективных микро- и наноэлектронных систем (МЭС-2020). – Москва: ИППМ РАН, 2020. – № 4. – С. 23–30.

12. Лежнев, Е.В. Автоматизация низкоуровневого моделирования сетей на кристалле / Е.В. Лежнев // Проблемы разработки перспективных микро- и наноэлектронных систем (МЭС-2021). – Москва: ИППМ РАН, 2021. – № 1. – С. 46–50.

13. Завьялов, А.Н. Разработка HDL модели сети на кристалле на основе маршрутизатора с одной очередью / А.Н. Завьялов, Е.В. Лежнев // Межвузовская научно-техническая конференция студентов, аспирантов и молодых специалистов им. Е.В. Арменского. – Москва, НИУ ВШЭ. – 2019. – С. 95–96. (РИНЦ).

Также есть

14. Свидетельство о государственной регистрации программы для ЭВМ (2021616623 РФ). Генератор Verilog кода подсистемы связи сетей на кристалле (Verilog Code Generator of Communication Subsystem for Networks-on-Chip, HDLNoCGen) / А.Ю. Романов, Е.В. Лежнев; заявитель и правообладатель Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет

«Высшая школа экономики» (RU). – № 2021615675; заявл. 20.04.21; опубл. 23.04.21, Реестр программ для ЭВМ.

Структура и содержание работы

Диссертация состоит из введения, пяти глав, заключения, списка используемой литературы и приложения с актами внедрения результатов исследования.

Во **введении** обоснована актуальность темы диссертации, сформулированы цель и задачи исследования, определены научная новизна и практическая ценность полученных результатов, приведены данные об их апробации.

В **первой главе** проведен анализ низкоуровневых моделей СтнК, а также основных подходов к моделированию СтнК, заключающихся в выборе уровня абстракции модели. В зависимости от требуемых задач модель может быть реализована на языках описания аппаратуры (HDL) или на языках высокого уровня (HLL). Определены характеристики основных видов моделирования в зависимости от уровня абстракции модели, обрабатываемых моделью данных и целей моделирования. Одними из следствий выбора способа реализации модели является скорость моделирования, а также точность результатов. Несмотря на то, что модели, разработанные на языках описания аппаратуры, по природе своей являются более точными, они обладают недостатком в виде длительного по времени моделирования. У высокоуровневых моделей среднее время моделирования значительно меньше, но точность может быть хуже.

В главе приведено описание структуры и основных компонентов моделей СтнК. Можно выделить следующие компоненты СтнК, используемые в моделях: маршрутизаторы, структура связей между ними, вычислительные IP-блоки.

Было проведено сопоставление структуры СтнК и решаемых ее компонентами задач с сетевой моделью OSI (рис. 1).

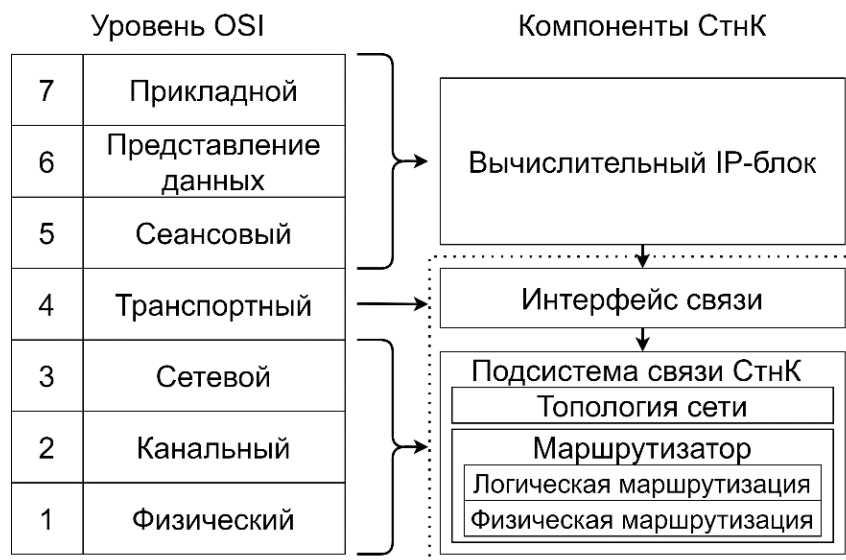


Рис. 1. Модель OSI в СтнК.

Прикладной уровень, уровень представления данных и сеансовый уровень реализуются вычислительным IP-блоком, который формирует данные для передачи по сети и указывает IP-блок назначения, куда необходимо передать эти данные. Оставшиеся уровни реализуются подсистемой связи СтнК, которая принимает от IP-блока данные, преобразует их в формат, пригодный для передачи между маршрутизаторами в сети с заранее определенной топологией, передает по рассчитанному маршрутизатором пути и, по достижении целевого узла, производит обратное преобразование для передачи в IP-блок.

Обоснована необходимость низкоуровневого моделирования при разработке СтнК. Приведено сравнение низкоуровневого моделирования с высокоуровневым и поведенческим моделированием, показано, что только низкоуровневые модели могут быть синтезированы в реальную СтнК и обеспечивают наибольшую точность моделирования, поскольку исследуется прототип сети, наиболее приближенный к реальной реализации СтнК. Высокоуровневые модели реализуют СтнК в общем виде и не могут учесть все особенности ее реализации.

В главе представлена классификация низкоуровневых моделей СтнК. На рисунке 2 приведена общая структура низкоуровневых моделей СтнК.

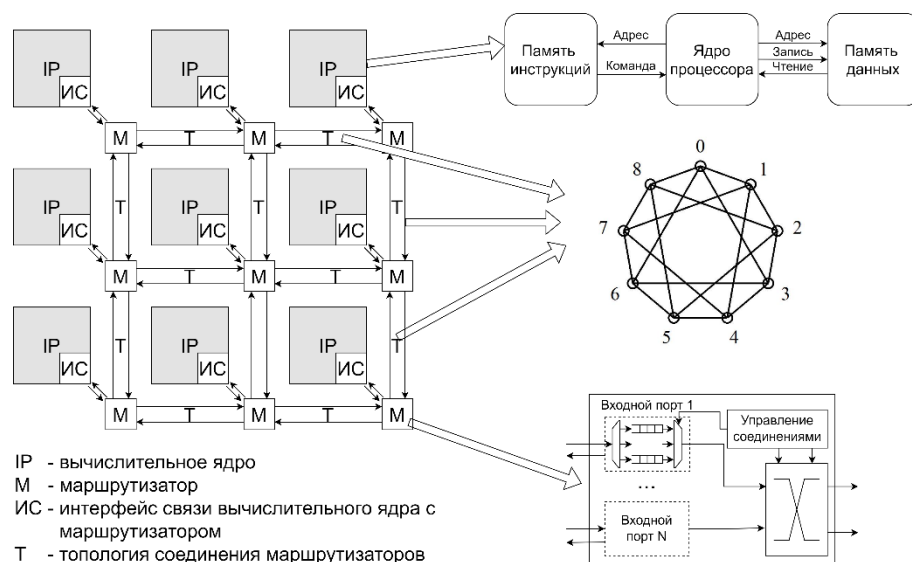


Рис. 2. Структура низкоуровневой модели СтнК.

Также в этой главе проведен обзор низкоуровневых моделей СтнК, согласно которому большинство моделей позволяют исследовать отдельные характеристики СтнК и реализуют ее отдельные компоненты. В результате обзора показано, что результаты моделирования различных моделей невозможно совместно использовать, т.к. они представляются в различных форматах. Кроме того, модификация моделей обычно невозможна, т.к. в большинстве реализаций отсутствует параметрическая настройка моделей.

Во **второй главе** описан процесс разработки моделей СтнК. Показано, что не во всех случаях требуется проводить моделирование всей СтнК. Зачастую необходимо проводить моделирование определенных ее частей, таких как подсистема связи, алгоритм маршрутизации, система управления трафиком. Проведение моделирования отдельных компонентов СтнК позволяет выполнить оценку их влияния на сеть по отдельности, а также на основе полученных данных, произвести параметрический выбор оптимальных настроек для работы выбранного компонента СтнК. Также раздельное моделирование компонентов СтнК позволяет ускорить процесс моделирования как сети в целом, так и исследуемых компонентов в отдельности.

В главе обоснована важность топологического подхода при разработке и моделировании СтнК. Структурно СтнК можно разделить на 2 основные части: вычислительный IP-блок и подсистему связи. Эти части СтнК взаимодействуют между собой только при передаче данных между IP-блоком и подсистемой связи и фактически не влияют на работу друг друга, так как являются отдельными структурно независимыми компонентами сети. Следствием этого является возможность раздельного моделирования этих частей с сохранением корректности получаемых результатов. В зависимости от выбранной топологии соединения маршрутизаторов, а также выбранного алгоритма маршрутизации, изменяются характеристики работы сети. Поэтому при проектировании СтнК моделирование подсистемы связи является важнейшим этапом.

В **третьей главе** представлена разработанная низкоуровневая модель подсистемы связи СтнК (рис. 3, блок «Ядро модели»), на основе которой разработана САПР для моделирования СтнК HDLNoC Gen. Модель состоит из 4 компонентов: ядро модели – сгенерированные файлы на языке Verilog, которые реализуют подсистему связи исследуемой СтнК; модуль настройки модели – получает необходимые данные для параметрической генерации подсистемы связи СтнК; модуль обработки данных – необходим для получения результатов работы модели и генерации тестовых данных для подачи в модель в реальном режиме времени; инфраструктура тестирования – содержит файлы для проведения автоматизированного событийного моделирования работы модели на заранее подготовленных данных.

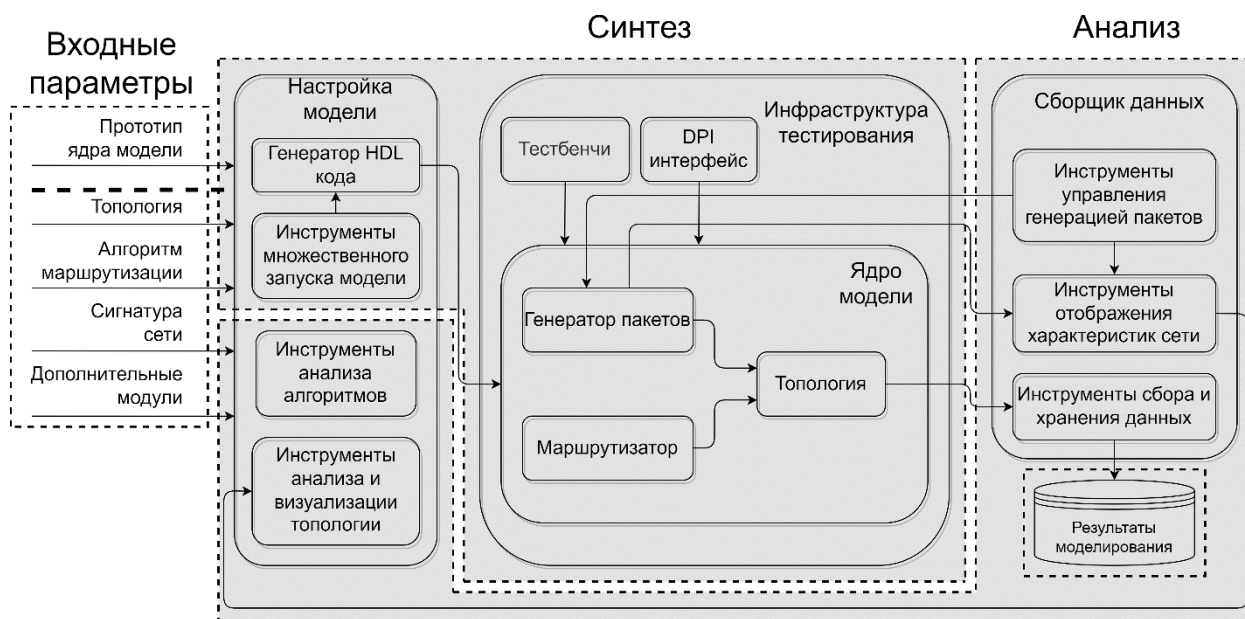


Рис. 3. Архитектура САПР для низкоуровневого моделирования СтК.

Ядро модели состоит из трех основных файлов – модуля маршрутизатора, модуля топологии и вспомогательного файла генератора пакетов тестовых данных. Модуль маршрутизатора реализует исследуемый алгоритм маршрутизации и фактически представляет собой конечный автомат, на основе которого работает маршрутизатор в сети. С помощью хранимой в маршрутизаторе информации о топологии подсистемы связи, а также сервисной информации, хранящейся в пересылаемых пакетах данных, производится выбор порта, в который нужно отправить пакет, чтобы он достиг узла назначения. Модуль топологии реализует исследуемую топологию подсистемы связи. В ней производится соединение всех маршрутизаторов между собой, на основе заданной в настройках топологии, не предполагающей структурных ограничений на ее выбор, выполняется передача данных между ними и вывод информации в вычислительный узел. Генератор пакетов является вспомогательным и фактически не является структурным элементом подсистемы связи в разрабатываемой системе; он нужен для редуцирования модели СтК – замены вычислительных IP-блоков. Он производит эмуляцию генерации пакетов данных при тестировании работы сети. Все порты маршрутизаторов для связи с вычислительными узлами соединены с модулем

генератора пакетов. Это дает возможность проводить исследование СтнК с большим количеством узлов, чем при использовании реальных вычислительных IP-блоков. Для проведения тестирования была разработана надстройка с файлами тестбенчей и для использования инструментов DPI.

На языке C# было разработано программное обеспечение САПР, реализующее представленную архитектуру низкоуровневого моделирования подсистемы связи СтнК. В ее состав входят компоненты настройки генерации ядра модели, инструменты сбора данных от ядра модели, а также инструменты управления генерацией пакетов данных.

Важной частью САПР является модуль настройки модели, который позволяет генерировать код ядра модели, а также проводить множественный запуск модели с различными параметрами.

Оставшимися частями САПР являются инструменты анализа результатов. Часть из них (инструменты отображения характеристик сети) располагаются в модуле сборщика данных, т.к. предоставляют доступ к необработанным данным. Остальные части располагаются в блоке настройки модели, т.к. производят обработку данных от модели и позволяют их сравнить с результатами, полученными на этапе высокоуровневого моделирования.

На вход САПР подается следующий набор параметров: топология сети, алгоритм маршрутизации, сигнатура сети (в которой содержится информация о количестве узлов и топологии их соединения), дополнительные модули, которые необходимо подключить к сети. Опционально можно задать расположение прототипа ядра модели.

Разработанная САПР позволяет проводить моделирование параметризированной подсистемы связи СтнК для получения оценки количества занимаемых арифметико-логических блоков (ALM) и регистров (REG, блоков памяти), необходимых для прототипирования подсистемы связи.

Все компоненты подсистемы связи реализованы в виде отдельных модулей, за счет чего снижены трудозатраты на добавление необходимых компонентов в модель для исследования за счет отсутствия необходимости полной переработки кода модели. Это возможно за счет созданного универсального интерфейса, в который все подключенные компоненты передают данные в своем формате, после чего происходит их преобразование для дальнейшего использования другими компонентами.

В соответствии с пунктом 3.3.2 ГОСТа 25301.101-87 «Системы автоматизированного проектирования», в составе САПР представлен следующий набор обеспечений: информационное (база данных результатов моделирования), математическое (аппроксимационные модели предсказания результатов моделирования), лингвистическое (низкоуровневые модели на языке Verilog), методическое (разработанные методики моделирования), техническое (действующий прототип на ПЛИС) и программное обеспечение (прикладное ПО САПР и вспомогательные утилиты). К алгоритмическому обеспечению относится блок генератора HDL-кода, который представляет собой транслятор кода, обеспечивающий генерацию ядра модели. К информационному обеспечению относятся входные данные, на основе которых происходит параметрический синтез модели, а также результаты работы модели, такие как: занимаемые ресурсы чипа, частота работы синтезированного проекта сети, эффективность размещения компонентов. В состав технического обеспечения входят рабочая станция на операционной системе Windows, на которой работает разработанная САПР, САПР Intel Quartus Prime Lite, которая производит создание файла описание сети для чипа ПЛИС, а также сам чип ПЛИС в составе отладочной платы.

В главе 4 рассмотрены способы автоматизации процесса подготовки HDL-модели подсистемы связи СтнК. Для каждой анализируемой конфигурации подсистемы связи необходимо создавать свои файлы ядра модели, что является рутинной и ресурсоемкой по времени задачей. Для

автоматизации этого процесса на языке C# был разработан модуль для настройки модели, в состав которого входят инструменты множественного запуска, а также генератор HDL-кода. Инструменты множественного запуска позволяют задавать пользователю несколько наборов данных для синтеза модели и запускают все необходимые программные модули для синтеза и выполнения модели.

Основным компонентом данной подсистемы является генератор HDL-кода. С помощью него производится параметрический синтез ядра модели. Структурно генератор HDL-кода представляет собой транслятор, который может работать в двух режимах в зависимости от подаваемых в САПР данных. Режимы работы транслятора приведены на рис. 4.

В первом режиме работа транслятора основана на использовании прототипа ядра модели (рис. 4а). Разработчик один раз создает прототип ядра модели. При дальнейшей эксплуатации системы требуется изменять только файл, в котором реализован алгоритм маршрутизации. В этом файле разработчик отмечает строки кода, в которые нужно внести изменения для создания параметризированной модели.

Второй режим работы генератора представляет собой C#-to-Verilog трансляцию (рис. 4б). В этом режиме работы генератору не нужен прототип ядра модели. На вход транслятора подается файл с алгоритмом маршрутизации, разработанный на языке C#. Также подается файл с ограничениями для лексического анализатора. На основе этих входных данных генератор кода формирует файл с алгоритмом маршрутизации на языке Verilog, а также все остальные необходимые файлы.

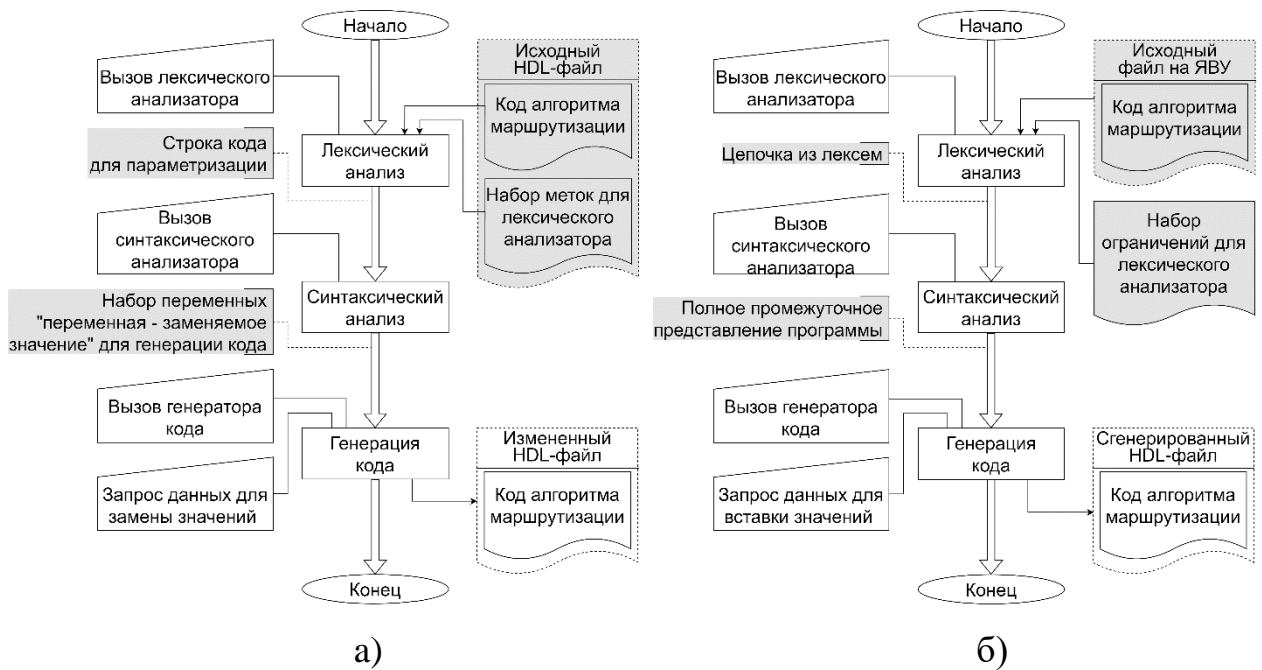


Рис. 4. Режимы работы генератора HDL-кода: а) на основе прототипа ядра модели; б) в режиме C#-to-Verilog транслятора.

Разработанный транслятор не только позволяет автоматически синтезировать параметрическую модель подсистемы связи, но и обеспечивает возможность согласовать исследование алгоритма маршрутизации с высокоуровневой моделью.

В модели СтнК реализована возможность подключения внешних вычислительных IP-блоков для проверки сети в целом. Для согласования данных и приведения их в необходимый формат, между IP-блоками и подсистемой связи реализован интерфейс связи, который преобразует данные. Интерфейс связи реализован в виде отдельного модуля, который разработчик может настраивать для подключения различных IP-блоков.

Используя реальные вычислительные IP-блоки для генерации входных данных для работы СтнК можно исследовать сети с небольшим количеством узлов, что было показано на расчете потребляемых ресурсов сетями на основе ядер schoolMIPS и ядер NIOS II. На основе отладочной платы ПЛИС DE1-SoC можно реализовать СтнК с 50 узлами NIOS II, и с 127 узлами schoolMIPS. Таким образом, при использовании реальных вычислительных IP-блоков,

исследовать СтнК с количеством узлов, превышающим сотню, является проблематичным.

Решением этой проблемы является редуцирование модели. Этой задачей занимается модуль генератора пакетов. Он создает тестовые пакеты данных для имитации работы подсистемы связи. Традиционное разделение на сообщения и данные в данном типе сетей не рассматривается. В разработанной низкоуровневой модели используется упрощенная структура пакета данных (рис. 5).

Флаг наличия пакета данных	Маршрутная информация	
	Адрес отправления	Адрес назначения

Рис. 5. Структура пакета данных, генерируемого модулем редуцирования модели генератора пакетов данных.

Используя генератор пакетов данных в качестве замены вычислительных IP-блоков, удалось при тех же ограничения отладочной платы, увеличить сеть по сравнению с сетью на основе ядер schoolMIPS на 37 %, а по сравнению с сетью на основе ядер NIOS II – на 248 %. Максимальный размер сети на основе редуцированной модели составил 174 узла. Используя другие алгоритмы маршрутизации, удалось увеличить размер СтнК до 200 узлов.

Представленная редуцированная модель имеет в своей структуре несколько дополнительных модулей, которые не являются архитектурными компонентами подсистемы связи и СтнК в целом, но также участвуют в моделировании и вносят искажения в получаемые результаты. Оценка влияния этих модулей также рассмотрена в **четвертой главе**. Представлены формулы, на основании которых можно заранее вычислить вносимые искажения в результаты моделирования с использованием дополнительных модулей; их можно использовать для корректировки результатов моделирования.

Структура модуля генератора пакетов постоянна и линейно зависит от параметров подсистем связи СтнК. Ее основу составляет вектор размерности

$L = N \cdot BIT$, где N – количество узлов, а BIT – длина пакета, поступающего в маршрутизатор. Далее этот вектор разделяется на N векторов длиной BIT , которые передаются в маршрутизаторы. Таким образом можно получить теоретическую оценку вносимой избыточности в модель подсистемы связи регистрах памяти (рис. 6).

$$U_{reg} = N \cdot BIT + 16 + 2 \cdot (BIT - 1),$$

где N – количество узлов в сети;

BIT – длина пакета.

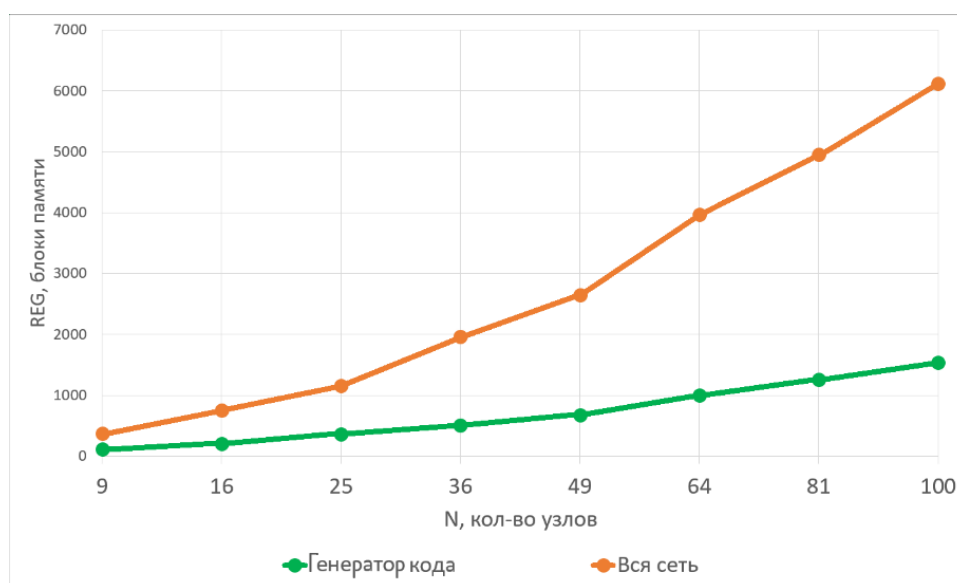


Рис. 6. Зависимость использования регистров ПЛИС для вспомогательных модулей от количества узлов в сети.

Оценка избыточности в логических блоках (ALM) (рис. 7):

$$U_{alm} = 0,0006 \cdot N^2 + 8,546 \cdot N + 25,101,$$

где N – количество узлов в сети.

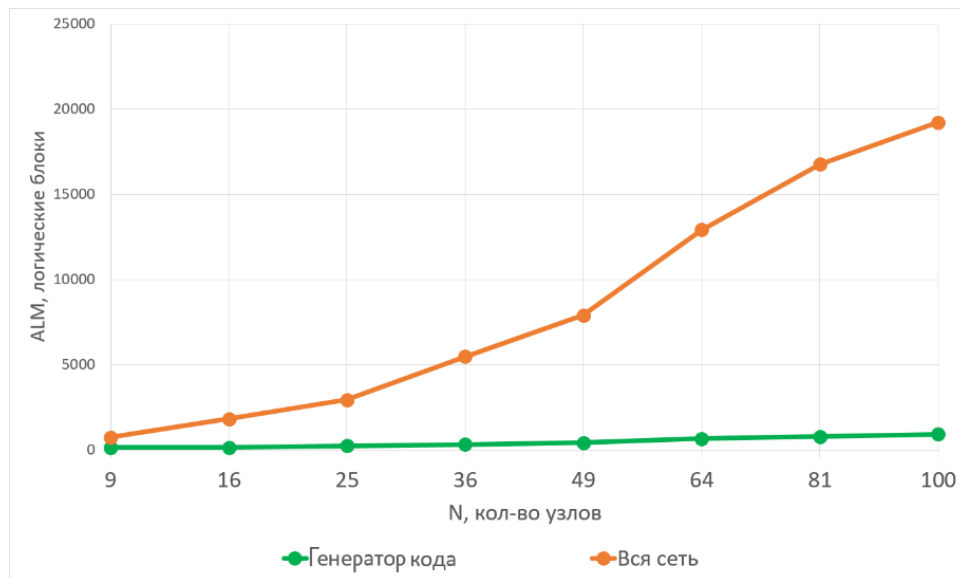


Рис. 7. Зависимость использования логических ресурсов ПЛИС для всей сети и вспомогательных модулей от количества узлов в сети.

Для регистров памяти оценка ресурсов точно совпадает с теоретически рассчитанными значениями. Оценка логических ресурсов может быть выражена полиномиальной функцией второй степени, представленной выше, с коэффициентов достоверности $R^2 = 0,991$. Также из полученных результатов следует, что по сравнению с подсистемой связи, модуль генератора пакетов занимает совсем мало ресурсов и для большего количество узлов этой избыточностью можно пренебречь.

На основании полученных формул оценки используемых неархитектурными компонентами СтК ресурсов был получен график расчета используемых ресурсов для большего количества узлов методом экстраполяции данных (рис. 8).

Полученные на основе полиномиальной функции данные для СтК с количеством узлов от 121 до 324 были сравнены с экспериментально полученными данными и доказано, что экстраполированные данные соответствуют измеренным.

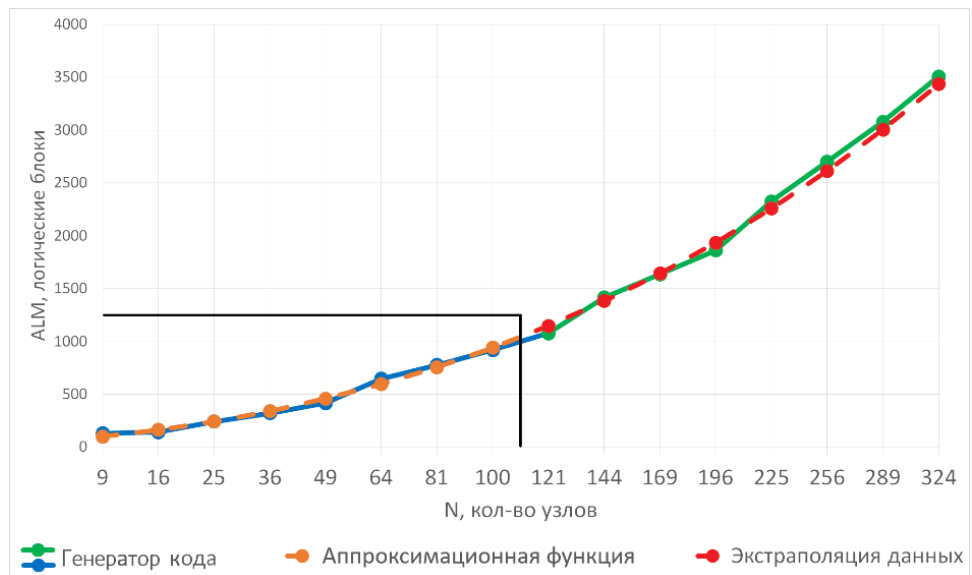


Рис. 8. Зависимость использования логических ресурсов ПЛИС для всей сети и вспомогательных модулей от количества узлов в сети.

Все созданные инструменты моделирования позволили ускорить процесс получения характеристик работы сети. На рис. 9 показана зависимость скорости получения данных работы сети от количества узлов с помощью разработанной САПР и с использованием низкоуровневого моделирования (в ModelSim).

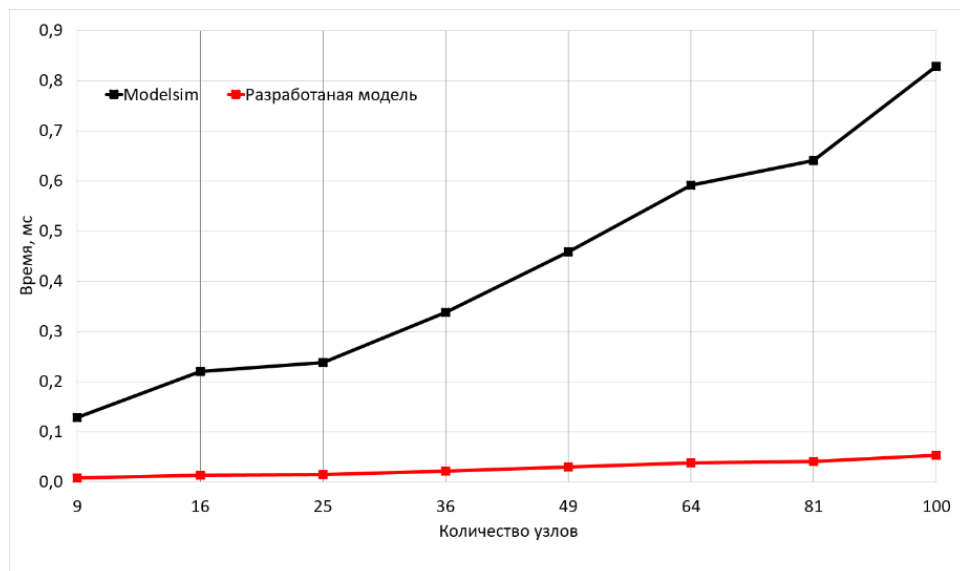


Рис. 9. Сравнение скорости получения результатов работы СтНК с использованием разработанных инструментов и с использованием инструментов ModelSim.

Для оценки скорости получения результатов проводилось моделирование прохождения одного пакета данных. По сравнению с ModelSim ускорение моделирования составило несколько порядков (до 15 раз).

Пятая глава посвящена апробации разработанной модели. Разработанная САПР и методы моделирования были применены в исследованиях работы алгоритмов и влияния топологии на работу СтнК, которые проводились в рамках исследовательских проектов ЦФИ НИУ ВШЭ и РФФ. Были получены данные о занимаемых ресурсах СтнК для сетей с различными параметрами. На основе этих результатов построены графики и зависимости.

Было проведено моделирование с использованием неархитектурных элементов в структуре подсистемы связи СтнК, в сравнении с такой же сетью, но в которой в качестве вычислительных узлов использовались ядра процессора schoolMIPS и NIOS II. На рис. 10 показано подключение в модель софт-процессора schoolMIPS.

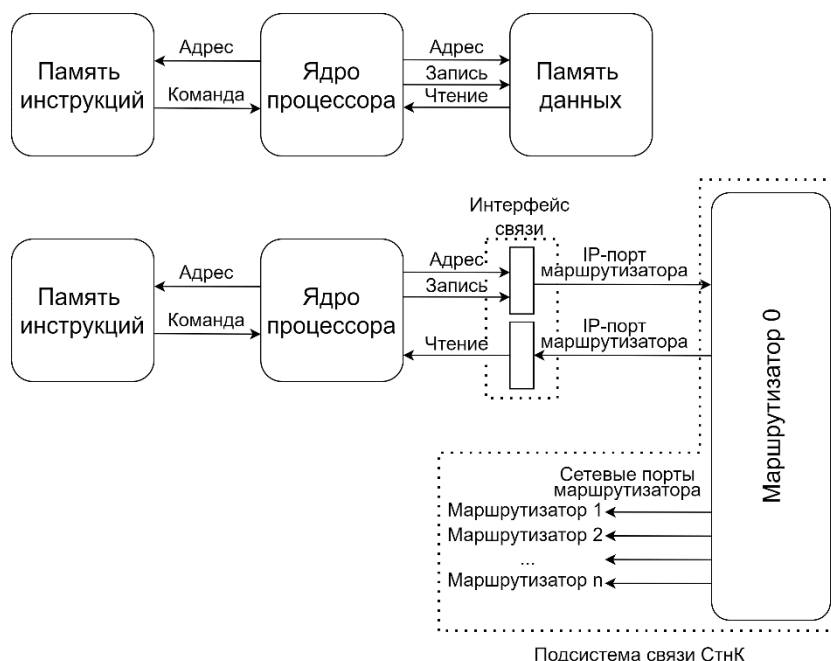


Рис. 10. Демонстрация реализации универсального интерфейса для подключения софт-процессора schoolMIPS в подсистему связи СтнК

В структуру процессора было добавлено 2 модуля для согласования формата данных между процессором и подсистемой связи. Первый модуль формирует пакет заданного формата для его передачи в маршрутизатор, а второй модуль занимается обратным преобразованием данных для передачи пакета в процессор schoolMIPS.

Таким образом, благодаря апробации на различных задачах, было показано, что разработанная САПР может применяться для практического применения для моделирования СтнК с количеством узлов, достигающим сотен, для исследования различных аспектов ее функционирования: от алгоритмов маршрутизации до метода контроля трафика, или их комбинаций.

В **выводах** перечисляются основные итоги диссертационной работы, важнейшим из которых является решение актуальной научно-прикладной задачи, связанной с автоматизацией процесса разработки низкоуровневых моделей подсистем связи СтнК, а также увеличением размеров моделируемой СтнК, путем создания новой низкоуровневой модели ее подсистемы связи, которая, за счет созданной САПР для разработки низкоуровневых моделей СтнК, включая универсальный интерфейс связи компонентов СтнК, а также модуль редуцирования модели, позволяет проводить моделирование как сети в целом, так и отдельных ее компонентов без необходимости модификации модели.