УДК 004.274:004.056

В.Ф. Евдокимов, Киев
А.Н. Давиденко, Киев
С.Я. Гильгурт, Киев

# ЦЕНТРАЛИЗОВАННЫЙ СИНТЕЗ ВЫЧИСЛИТЕЛЬНЫХ СТРУКТУР РЕКОНФИГУРИРУЕМЫХ УСКОРИТЕЛЕЙ ЗАДАЧ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ[1]

**Abstract.** The use of GRID or other HPC means as a platform for organizing remote centralized synthesizing of reconfigurable accelerators intended for information security protection is explored. Such approach migrates computation complexity from local LANs to supercomputer infrastructure. Besides the subsystem for generating bitstreams for FPGAs, a stage of synthesizing digital structure of reconfigurable accelerator is added.

## Введение

Высокие гибкость и быстродействие реконфигурируемых ускорителей на базе программируемых логических интегральных схем (ПЛИС) позволяют успешно решать вычислительно сложную задачу сигнатурного анализа признаков вредоносной активности в больших объемах данных. Такая задача возникает, в частности, в сетевых системах обнаружения вторжений (ССОВ), антивирусах и других средствах информационной защиты, обеспечивающих безопасность локальных компьютерных сетей различного уровня и назначения – от крупных информационных интранет-систем до смарт-сетей на электрических подстанциях электроэнергетических предприятий.

К сожалению, при практическом использовании программируемой логики возникают определенные сложности, связанные, прежде всего, с трудоемкостью процесса программирования ПЛИС. В этот процесс в качестве составляющих подзадач входят в общем случае: создание проекта (разработка цифровой схемы), ввод данных в специализированную САПР, отладка, компиляция проекта, его верификация, моделирование и тестирование, проверка работоспособности, оценка временных показателей создаваемой схемы, параметров энергопотребления и др. [1 – 4].

В некоторых случаях этот процесс можно упростить, сократив большую часть этапов. Например, в процессе функционирования систем обнаружения вторжений или антивирусных средств, построенных на базе реконфигурируемых ускорителей, время от времени возникает необходимость заново синтезировать некоторые аппаратные компоненты в связи с изменившимися условиями работы – обновлением базы данных сигнатур, изменением состава

[1] Исследование выполнено при частичном финансировании Целевой комплексной программой научных исследований НАН Украины «Грид-инфраструктура и грид-технологии для научных и научно-прикладных применений».

 

локальной сети, ее структуры, используемого программного обеспечения и т.п. В этом случае нет необходимости осуществлять полный цикл создания цифрового устройства, достаточно выполнить некоторые отдельные шаги, при этом остальные операции могут быть реализованы однократно заранее.

Другим направлением снижения трудоемкости процесса создания реконфигурируемых устройств является сосредоточение в едином центре вычислительной работы, которую необходимо выполнить для удовлетворения потребностей большого числа пользователей.

*Целью данной работы является* исследование процесса централизованного синтеза реконфигурируемых ускорителей задач информационной безопасности на предмет выделения из процесса создания цифровых устройств обязательных этапов, а также более подробное рассмотрение одного из них.

## 1. Предпосылки и преимущества централизации при разработке и сопровождении аппаратных средств информационной защиты

Концепция системы централизованного синтеза аппаратных ускорителей для решения задач информационной безопасности с использованием высокопроизводительной вычислительной инфраструктуры, предложенная в работе [5], предполагает следующую постановку задачи. Имеется большое число информационных объектов, для киберзащиты которых используются реконфигурируемые аппаратные устройства, построенные в общем случае на базе разных ПЛИС различной вычислительной мощности. Вследствие частого изменения условий функционирования защищаемых объектов (добавления в базу данных сигнатур вновь обнаруженных атак, изменение настроек либо системы защиты, либо параметров объектов) требуется оперативное реконфигурирование ПЛИС, то есть повторный синтез быстродействующих вычислительных структур, реализующих сложные в вычислительном смысле алгоритмы сигнатурного распознавания. При этом пользователи систем защиты (персонал, ответственный за информационную безопасность) не располагают ни достаточными вычислительными ресурсами, ни навыками синтеза сложных цифровых устройств на базе программируемой логики.

Предлагаемое решение заключается в использовании централизованной системы (использующей вычислительные ресурсы современных высокопроизводительных сред), работа которой организована в виде удаленного сервиса и включает в себя выполнение следующих функций:

– оперативное снабжение клиентов актуальной информацией о вновь обнаруженных факторах злонамеренной активности (атаках, вирусах и т.п.);

– сбор данных о текущих параметрах настройки систем безопасности каждого из защищаемых объектов (клиентов);

– компиляция комплектов цифровых компонентов аппаратных ускорителей для клиентских систем, синтезированных для обслуживаемых

клиентов с учетом особенностей каждого из них;

– оперативная доставка потребителям готовых конфигураций для программирования ПЛИС реконфигурируемых ускорителей.

Можно заметить, что в отличие от известного механизма централизованной рассылки обновлений, используемого в широко распространенных программных антивирусных системах, вместо файлов обновлений антивирусной базы клиентам рассылаются загружаемые в ПЛИС конфигурации, а также учитываются индивидуальные требования каждой из обслуживаемой системы защиты.

В структурном плане система включает в себя на нижнем уровне набор реконфигурируемых устройств, обеспечивающих информационную защиту компьютерных объектов, а на верхнем – удаленный централизованный сервис, реализованный на базе современных высокопроизводительных компьютерных технологий.

Пользователи локальных систем (системные администраторы либо иные лица, ответственные за кибербезопасность) через соответствующий интерфейс передают сервису исходные данные (списки сигнатур, подлежащих распознаванию, параметры ССОВ и т.п.), а результаты обработки получают в виде конфигураций – двоичных файлов для программирования ПЛИС.

В работах [6, 7] описывается разработка, получившая название STRAGS (Security Tasks Reconfigurable Accelerators Grid-Service – грид-сервис для реконфигурируемых ускорителей задач информационной безопасности). С его помощью экспериментальным путем была подтверждена идея о целесообразности переноса ресурсоемкой операции синтеза аппаратных ускорителей с локальных реконфигурируемых устройств киберзащиты на высокопроизводительные компьютерные системы.

В результате дальнейшего исследования было установлено, что централизованный подход предоставляет ряд преимуществ по сравнению с локальным принципом создания реконфигурируемых средств киберзащиты. В частности, централизация позволяет следующее.

1. За счет использования суперкомпьютерной технологии повышается производительность системы в целом. Такие современные технологии как суперкомпьютинг, грид-системы и облачные вычисления предоставляют ресурсы, достаточные для быстрого выполнения вычислительно сложных задач оптимизации и синтеза параллельных распознающих структур на базе современных реконфигурируемых СБИС, последние семейства которых содержат миллионы эквивалентных логических элементов программируемых ресурсов, блоки внутренней памяти, аппаратные умножители и другие высокотехнологичные компоненты.

2. Благодаря разделению труда улучшаются технические характеристики локальных систем защиты информации. Централизация позволяет задействовать высококвалифицированных специалистов, результаты работы которых будут использоваться на каждой из локальных систем, что

невозможно достигнуть при индивидуальной разработке систем защиты по отдельности.

3. За счет снижения требований к квалификации персонала локальных систем также снижается совокупная стоимость владения.

4. Путем группирования схожих запросов сокращаются общие вычислительные затраты. Несмотря на то, что для аппаратных ускорителей каждой клиентской системы в общем случае требуется своя собственная, уникальная конфигурация, общая база сигнатур и ограниченность числа типов используемых ПЛИС позволяют оптимизировать вычислительный процесс таким образом, что многие из создаваемых комплектов реконфигурируемых компонентов могут (при незначительной избыточности) подходить одновременно нескольким клиентам. В итоге общая вычислительная сложность решения задачи снижается, причем тем существеннее, чем большее число информационных систем будет охвачено сервисом.

## 2. Два этапа синтеза реконфигурируемых модулей распознавания сигнатур

Как упоминалось выше, разработка аппаратных ускорителей задач информационной безопасности на базе ПЛИС, как и любых сложных цифровых устройств, является комплексной задачей, состоящей из ряда последовательно выполняемых процедур. Однако, за счет сужения функциональности синтезируемых устройств до конкретной технической задачи этот процесс можно упростить, выполнив ряд операций заранее. Конкретизация в выборе алгоритмов распознавания, а также стандартизация формы представления входных данных приводят к еще большей унификации и сокращению числа операций. В результате вся технологическая цепочка в случае повторного синтеза цифровых схем сигнатурного распознавания сводится фактически к двум этапам, рассмотренным ниже.

Этап 1. Синтез вычислительной структуры. Данный этап не может быть сокращен, поскольку зависит от изменившихся исходных данных – размера и состава базы данных сигнатур, подлежащих распознаванию. Результатом его выполнения являются спецификации разработанных схем, представленные, как правило, на одном из языков описания аппаратуры (Hardware Description Language – HDL), таком как VHDL, Verilog и т.п. Данный этап может быть выполнен как в автоматизированном, так и автоматическом режиме (хотя второй вариант представляет собой нетривиальную научную задачу). В следующем разделе рассмотрен именно этот этап синтеза цифровых схем.

Этап 2. Автоматический синтез загружаемой в ПЛИС конфигурации. На данном шаге все операции выполняются полностью автоматически с применением фирменной САПР или специализированного программного продукта. Этап включает в себя такие процедуры, как синтез (Synthesize), трансляция (Translate), отображение (Map), размещение/трассировка (Place & Route) и собственно генерация битовой последовательности (Bitstream

Generating). Данный этап был реализован в рамках упомянутого выше грид-сервиса STRAGS на базе грид-узла PIMEE Института проблем моделирования в энергетике им. Г.Е. Пухова НАН Украины [7]. Результатом его работы являются файлы конфигурации, которые загружаются в ПЛИС реконфигурируемых ускорителей.

Оба рассмотренных выше этапа являются необходимыми при каждом повторном синтезе реконфигурируемых структур. В качестве разделителя между ними выступает представление синтезируемой цифровой схемы на языке описания аппаратуры.

### 3. Синтез аппаратной структуры системы обнаружения вторжений

Исторически первыми разработками в сфере защиты информации с применением ПЛИС типа FPGA (Field Programmable Gate Array) были сетевые системы обнаружения вторжений. В публикации [8] проведено исследование и обобщение основных принципов построения ССОВ на базе программируемой логики по результатам анализа накопленного в мире опыта применения ПЛИС типа FPGA. Рассмотрены основные параметры подобных систем и предъявляемые к ним требования. Сформулированы в общем виде состав и структура аппаратной реализации на ПЛИС системы обнаружения вторжений (рис. 1).

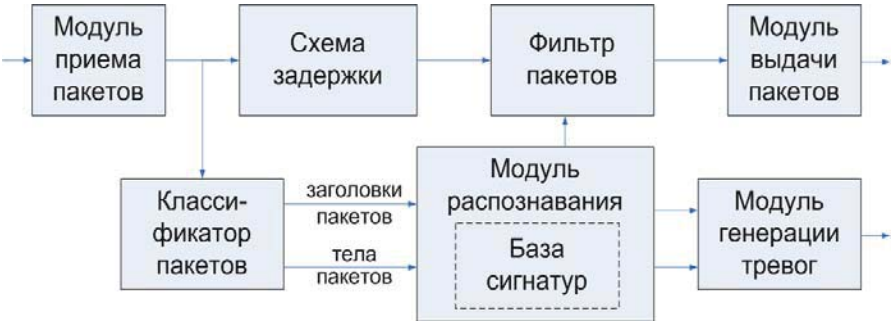


Рис. 1. Обобщенная структурная схема ССОВ на базе ПЛИС

В состав структуры входят:

- модуль приема пакетов;
- классификатор пакетов;
- модуль распознавания;
- модуль генерации тревог;
- схема задержки;
- фильтр пакетов;
- модуль выдачи пакетов.

Модуль приема пакетов осуществляет низкоуровневый захват сетевых пакетов и их преобразование в более удобный для внутрисхемной обработки тип кодирования, например, из формата XAUI (10 Gigabit Attachment Unit Interface) в формат XGMII (10 Gigabit Media Independent Interface) [9].

Классификатор разбирает пакеты на основе анализа заголовков вплоть до требуемого уровня в зависимости от используемого метода обнаружения вторжений [10].

Модуль распознавания выполняет самую ресурсоемкую вычислительную операцию поиска сигнатур в соответствии с выбранным алгоритмом. От качества его реализации в значительной степени зависят главные характеристики всей системы обнаружения вторжений: производительность, ресурсоемкость и масштабируемость. В большинстве случаев база данных правил распознавания сигнатур, включая сами сигнатуры непосредственно "вшита" в распознающую вычислительную структуру [11].

Вопросы выбора алгоритма распознавания и технического решения для его реализации выходят за рамки данной работы и будут рассмотрены в последующих публикациях цикла.

Помимо аппаратуры, реализующей собственно алгоритм поиска вхождения шаблонов в содержимом сетевых пакетов, модуль распознавания в общем случае содержит также схему распознавания заголовков пакетов и детектор правил базы данных сигнатур [8].

Модуль генерации тревог служит для формирования сообщений об обнаруженных вторжениях. Он идентифицирует вредоносные пакеты и объединяет информацию о типе атаки, поступающую от узла распознавания с дополнительными сведениями из заголовков пакетов, позволяющими идентифицировать источник вторжения.

Схема задержки синхронизирует поток пакетов с работой модуля распознавания.

Фильтр пакетов служит для пресечения вредоносного трафика путем отбрасывания вредоносных пакетов.

Модуль выдачи пакетов реализует преобразование, обратное тому, что выполнялось в модуле приема пакетов.

Следует заметить, что последние три компонента из рассмотренных, а именно: схема задержки, фильтр пакетов и модуль выдачи пакетов, в общем случае не являются необходимыми компонентами ССОВ и присутствуют только в структурах сетевых систем предотвращения вторжений (ССПВ), соответствующий англоязычный термин – Network Intrusion Prevention System (NIPS). ССПВ предъявляют повышенные требования как по быстродействию, так и по достоверности процедуры распознавания, поскольку могут оказывать существенное, в том числе и негативное влияние на работу защищаемой вычислительной сети. Системы предотвращения вторжений в данном исследовании не рассматриваются, соответствующие им компоненты приведены для полноты изложения.

Анализ компонентов систем обнаружения вторжений и выполняемых

ими функций приводит к выводу, что при изменении исходных данных перекомпиляции подлежит только модуль распознавания, включающий в себя на аппаратном уровне базу данных сигнатур.

Автоматическая компиляция модуля распознавания является нетривиальной задачей. Подробное рассмотрение принципов ее решения планируется в последующих публикациях цикла. В качестве примера возможного решения можно привести описанную в работе [12] процедуру.

Завершая рассмотрение принципов построения аппаратной структуры ССОВ, заметим, что при реализации антивирусных систем используются похожие подходы. Анализ специфики, вносимой увеличенным размером антивирусной базы данных сигнатур, также выходит за рамки данной работы.

**4. Технические особенности реализации**

Как указывалось выше, в пробной версии грид-сервиса, реализованной ранее, выполнялся только второй из двух этапов, описанных в разделе 2. Рассмотрим технические особенности реализации системы централизованного синтеза реконфигурируемых модулей распознавания сигнатур при добавлении первого этапа.

Как указывалось в работе [7], в предыдущей версии системы в качестве исходных данных, передаваемых пользователями грид-сервису STRAGS, выступали все необходимые для автоматической компиляции конкретного проекта компоненты – VHDL-описания узлов цифровой схемы, ucf-файлы и другая сопутствующая информация в формате используемой САПР. С целью унификации и упрощения работы с сервисом вся эта информация упаковывалась в единый архивный файл в предопределенном формате.

Для обеспечения обратной совместимости с работающей версией при введении нового этапа предлагается сохранить основной принцип обмена с центральной частью. Отличительной особенностью нового режима будет наличие в составе архивного файла, передаваемого локальным клиентом центральному грид-сервису, текстового файла "rules.txt", который содержит набор сигнатур сетевых атак, подлежащих обнаружению в локальной сети данного клиента. Состав и назначение остальных файлов архива остается прежним. Кроме упрощения адаптации пользователей к новому режиму такой подход обеспечит универсальность центральной части системы и даст возможность ее использования как в одноэтапном (когда выполняется только Этап 2), так и в двухэтапном (выполняются оба этапа) режимах. Индикатором выбора нужного режима для грид-сервиса является наличие во входном архиве файла сигнатур "rules.txt".

Заметим, что при таком подходе передача на верхний уровень сведений о типе ПЛИС, водящей в состав аппаратного ускорителя конкретной системы киберзащиты, как и в предыдущей версии, обеспечивается автоматически за счет наличия всех необходимых для компиляции компонентов проекта системы САПР, которая используется грид-сервисом в качестве инструмента создания загружаемых в ПЛИС конфигураций.

Формат файла "rules.txt" зависит от используемой в качестве прототипа ССОВ и соответствующей ей структуры базы данных сигнатур. Во многих академических публикациях по системам обнаружения вторжений на реконфигурируемой элементной базе в качестве наборов сигнатур используются базы таких свободно распространяемых систем обнаружения вторжений, как Bro, Snort или Suricata [13 – 15]. Однако большинство исследователей используют в качестве устоявшегося стандарта де-факто набор сигнатур от Snort [16, 17], что дает возможность достаточно объективно сравнивать между собой количественные показатели быстродействия разработок разных авторских коллективов.

**Выводы**

В данной работе продолжены исследования, связанные с созданием на базе грид-инфраструктуры системы централизованного синтеза аппаратных ускорителей для решения задач информационной безопасности, которые могут быть использованы в энергетической отрасли.

Проанализированы предпосылки и преимущества подхода, заключающегося в централизованном выполнении трудоемких процессов синтеза реконфигурируемых средств информационной защиты на базе ПЛИС.

Также в работе в результате исследования полного процесса синтеза, в нем выделены два обязательных этапа, которые требуется выполнять всякий раз при изменении условий функционирования информационных объектов, подлежащих защите. Подробно рассмотрен один из них – этап синтеза вычислительной структуры, завершающийся формированием спецификации цифровых схем на языке описания аппаратуры.

В результате анализа обобщенной структуры аппаратной реализации на ПЛИС системы обнаружения вторжений выделены, с одной стороны, неизменные компоненты, которые могут быть сформированы заранее, с другой – модули, подлежащие повторному синтезу при изменении исходных данных.

Разработанный ранее (на базе грид-узла Института проблем моделирования в энергетике им. Г.Е. Пухова НАН Украины) грид-сервис STRAGS дополнен недостающим этапом синтеза вычислительной структуры. Приведены сведения по его технической реализации в рамках централизованной системы. Предложено решение, позволяющее сохранить функциональность предыдущей версии сервиса. Обновленная система способна по составу водных данных, присылаемых пользователем, автоматически распознавать в каком из режимов следует работать – урезанном или обновленном.

1. *Грушвицкий Р.И., Мурсаев А.Х., Угрюмов Е.П.* Проектирование систем на микросхемах программируемой логики. СПб.: БХВ-Петербург, 2002. – 608 с.
2. *Зотов В.Ю.* Проектирование цифровых устройств на основе ПЛИС фирмы XILINX

в САПР WebPACK ISE. М.: Горячая линия – Телеком, 2003. – 624 с.
3. Реконфигурируемые вычислительные системы: Основы и приложения. / А.В. Палагин, В.Н. Опанасенко. – К.: «Просвіта», 2006. – 280 с.
4. *Угрюмов Е.П.* Цифровая схемотехника: Учеб. пособие для вузов. – 3-е изд., перераб. и доп. – СПб.: БХВ-Петербург, 2010. – 816 с.
5. *Гильгурт С.Я.* Организация вычислительного процесса синтеза файлов конфигураций для аппаратных ускорителей при решении задач информационной безопасности // Моделювання та інформаційні технології. Зб. наук. пр. ІПМЕ ім Г.Є. Пухова НАН України. – Вип. 74. – К.: 2015. – С.29-33.
6. *Євдокимов В.Ф., Давиденко А.М., Гільгурт С.Я.* Створення на базі грід-сайту ІПМЕ ім. Г.Є. Пухова НАНУ системи централізованого синтезу апаратних прискорювачів для вирішення задач інформаційної безпеки в енергетичній галузі // Моделювання та інформаційні технології. Зб. наук. пр. ІПМЕ ім Г.Є. Пухова НАН України. – Вип. 79. – К.: 2017. – С.3-8.
7. *Евдокимов В.Ф., Давиденко А.Н., Гильгурт С.Я.* Организация централизованной генерации файлов конфигураций для аппаратных ускорителей задач информационной безопасности // Моделювання та інформаційні технології. Зб. наук. пр. ІПМЕ ім Г.Є. Пухова НАН України. – Вип. 81. – К.: 2017. – С.3-11.
8. *.Коростиль Ю.М., Гильгурт С.Я.* Принципы построения сетевых систем обнаружения вторжений на базе ПЛИС // Моделювання та інформаційні технології. Зб. наук. пр. ІПМЕ ім Г.Є. Пухова НАН України. – Вип. 57. – К.: 2010. – С.87-94.
9. *Katashita T., Yamaguchi Y., Maeda A., Toda K.* FPGA-Based Intrusion Detection System for 10 Gigabit Ethernet // IEICE – Transactions on Information and Systems, v.E90-D n.12, p.1923-1931, December 2007.
10. *Jiang W., Prasanna V.* Scalable Multi-Pipeline Architecture for High Performance Multi-Pattern String Matching // IEEE International Parallel and Distributed Processing Symposium (IPDPS '10), April 2010.
11. *Jiang W., Prasanna V.* A FPGA-based Parallel Architecture for Scalable High-Speed Packet Classification // Proceedings of the International Conference on Application-Specific Systems, Architectures and Processors, July 2009, pp.24-31.
12. *Mitra A., Najjar W., Bhuyan L.* Compiling PCRE to FPGA for accelerating SNORT IDS, Proceedings of the 3rd ACM/IEEE Symposium on Architecture for networking and communications systems // December 03-04, 2007, Orlando, Florida, USA.
13. The Bro Network Security Monitor [Электронный ресурс]. – Режим доступа: http://www.bro-ids.org. – Загл. с экрана. – (Дата обращения: 01.07.2018).
14. SNORT [Электронный ресурс]. – Режим доступа: http://snort.org. – Загл. с экрана. – (Дата обращения: 01.07.2018).
15. Suricata. Open Source IDS / IPS / NSM engine [Электронный ресурс]. – Режим доступа: http://suricata-ids.org. – Загл. с экрана. – (Дата обращения: 01.07.2018).
16. *Давиденко А.Н., Гильгурт С.Я., Сабат В.И.* Аппаратное ускорение алгоритмов сигнатурного обнаружения вторжений в открытой системе информационной безопасности Snort // Моделювання та інформаційні технології. Зб. наук. пр. ІПМЕ ім. Г.Є. Пухова НАН України. – Вип. 65. – К.: 2012. – С.94-103.
17. *.Коростиль Ю.М., Гильгурт С.Я., Назаренко О.М.* Анализ базы данных системы информационной безопасности Snort и вопросы быстродействия // Моделювання та інформаційні технології. Зб. наук. пр. ІПМЕ ім Г.Є. Пухова НАН України. – Вип. 66. – К.: 2012. – С.77-84.

*Поступила 15.01.2018р.*