УДК 004.274:004.056

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

ОРГАНИЗАЦИЯ ЦЕНТРАЛИЗОВАННОЙ ГЕНЕРАЦИИ ФАЙЛОВ КОНФИГУРАЦИЙ ДЛЯ АППАРАТНЫХ УСКОРИТЕЛЕЙ ЗАДАЧ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ[1]

Abstract. The use of GRID infrastructure as a platform for organizing remote centralized synthesizing of reconfigurable accelerators intended for solving computer security problems is explored. Such approach allows to migrate computation complexity from local devices to supercomputer network. Subsystem for generating bitstreams for FPGAs of reconfigurable accelerators is proposed. A service to link users to resources of GRID is developed. The solution is based on the RAINBOW technology, popular in the Ukrainian national grid.

Введение

Одной из сложных в вычислительном плане задач в области информационной безопасности является распознавание признаков вредоносной активности в интенсивном потоке данных, которая решается, в частности, в процессе функционирования сетевых систем обнаружения вторжений (ССОВ) и антивирусных средств.

Увеличение числа и сложности компьютерных атак, а также прекращение роста частоты универсальных процессоров заставляют разработчиков переходить от программных решений к аппаратным, в частности – к использованию реконфигурируемых устройств на базе ПЛИС, которые обладают высокой гибкостью, одновременно приближаясь по быстродействию к специализированным вычислительным устройствам. Применение программируемой логики, однако, обладает рядом особенностей, сдерживающих ее массовое применение. Одной из сложностей является ресурсоемкая процедура создания *конфигураций*, т.е. последовательностей битов (соответствующий англоязычный термин – bitstream), определяющих внутреннюю структуру межсоединений компонентов ПЛИС, которые необходимо загружать в реконфигурируемые ускорители перед их использованием.

В работе [1] была предложена концепция системы централизованного синтеза аппаратных ускорителей для решения задач информационной безопасности в энергетической отрасли с использованием высокопроизводительной вычислительной инфраструктуры. Процесс создания загружаемых в ПЛИС конфигураций является неотъемлемой функцией такой системы.

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

Целью данной работы является исследование и разработка принципов организации процесса генерации конфигураций при централизованном синтезе реконфигурируемых вычислителей в грид-среде.

1. Генерация файлов конфигураций для реконфигурируемых устройств

Разработка аппаратных ускорителей на базе ПЛИС, как любых сложных цифровых устройств, является комплексной задачей, состоящей из ряда трудоемких процедур. Сюда входят, в общем случае, операция создания проекта, ввод данных в специализированную САПР, отладка, компиляция и тестирование проекта, его верификация, моделирование и тестирование, проверка работоспособности, а также оценка временных показателей создаваемой схемы, параметров энергопотребления и др. [2 – 5].

В случае сужения функциональности до конкретной технической задачи либо класса задач, например, при создании аппаратных компонентов систем обнаружения вторжений или антивирусных средств, этот процесс можно упростить, выполнив ряд операций заранее. Выбор конкретного алгоритма распознавания и стандартизация формы представления входных данных приводят к еще большей унификации и сокращению числа операций. В итоге вся технологическая цепочка синтеза цифровой схемы в ПЛИС может быть сведена к двум этапам, первый из которых зависит от входных данных (размера и состава базы данных сигнатур), а второй может быть выполнен полностью автоматически средствами САПР. При этом в качестве ключевого разделительного признака выступает представление синтезируемой цифровой схемы на одном из языков описания аппаратуры (например, VHDL). Под понятием *генерация файлов конфигураций* в настоящей работе подразумевается выполнение второго этапа из указанных двух.

Данный этап синтеза цифровой схемы кроме собственно процедуры создания загружаемых в ПЛИС последовательностей битов (Bitstream Generating), включает предшествующую ему операцию компиляции проекта, которая в свою очередь состоит из ряда вычислительно емких процедур, таких как:

– синтез (Synthesize);

– трансляция (Translate);

– отображение (Map);

– размещение и трассировка (Place & Route).

Каждое из перечисленных действий в свою очередь может включать разное количество подзадач. На рис. 1. приведен пример процесса разработки цифровой схемы, отображаемого графическим интерфейсом САПР WebPack ISE фирмы Xilinx. (Процедуры Translate, Map и Place & Route структурно сгруппированы в единую операцию реализации проекта Implement Design). Здесь все стадии процесса, начиная со строки " Synthesize – XST" и по "Programming File Generation Report" выполняются в автоматическом режиме

без участия разработчика (в случае отсутствия синтаксических и логических ошибок в проекте).

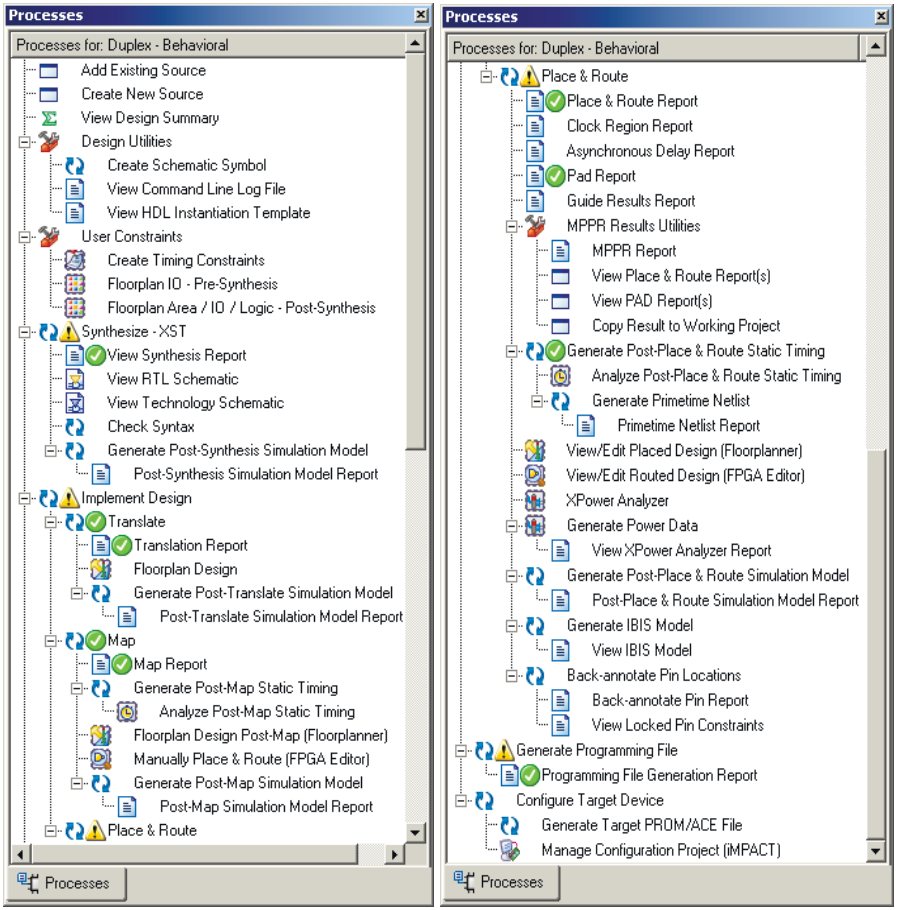

Рис. 1. Процесс генерации загружаемой в ПЛИС конфигурации

В зависимости от сложности синтезируемой схемы и типа ПЛИС процесс генерации файлов конфигураций может занимать от десятков минут до нескольких часов. Особенно критичной трудоемкость компиляции становится в случаях, когда задействованные ресурсы стремятся почти полностью занять площадь кристалла программируемой СБИС. В этом случае процедура размещения и трассировки превращается в комбинаторно сложную задачу перебора астрономически большого числа вариантов.

2. Особенности синтеза реконфигурируемых ускорителей задач информационной безопасности

Прикладная область киберзащиты накладывает определенную специфику на процесс разработки реконфигурируемых аппаратных ускорителей. Рассмотрим подробнее, к чему приводят данные особенности.

Угрозы информационной безопасности в компьютерных системах характеризуются высокой степенью динамичности, которая в свою очередь обусловлена стабильно нарастающей активностью злоумышленников – числом и изощренностью кибератак. Так, в современных системах антивирусной защиты обновление баз данных сигнатур в связи с появлением новых вредоносных программ, осуществляется в среднем несколько раз в сутки. С другой стороны, процессу защиты информации в компьютерных системах присуще разнообразие многочисленных настроек и режимов работы программного обеспечения, что делает каждый защищаемый объект (компьютер, локальную или корпоративную сеть), уникальным, не похожим на другие. Так, база данных сигнатур одной из самых популярных свободно распространяемых систем обнаружения вторжений Snort [6] насчитывает уже несколько десятков тысяч правил. Системный администратор в процессе работы с ССОВ обязан включать или выключать правила в зависимости от назначения и свойств защищаемого объекта. Например, если в защищаемой локальной сети отсутствуют компьютеры под управлением операционных систем семейства Windows, то все известные атаки на данный класс ОС не представляют угрозы, и соответствующие записи могут быть исключены из базы данных сигнатур.

Поскольку базы данных системы шибер защиты постоянно обновляются, ресурсоемкую процедуру синтеза конфигураций, строго говоря, надо повторять при возникновении и добавлении в базу сигнатур каждой новой атаки. Как следствие, при синтезе вычислительной структуры блока распознавания системы обнаружения вторжений каждому пользователю в общем случае требуется своя собственная, уникальная конфигурация для ПЛИС ускорителя. И эту конфигурацию необходимо синтезировать заново как при изменении настроек ССОВ, так и при добавлении в базу данных новых сигнатур недавно обнаруженных атак, несущих угрозу данному объекту.

Упомянутые выше особенности приводят к тому, что разработать завершенное универсальное решение, одинаково подходящее для работы в любой ССОВ или антивирусной системе на практике не представляется возможным. С другой стороны, необходимость частого повторного выполнения сложной в вычислительном плане задачи создания конфигурации для реконфигурируемых модулей распознавания каждой конкретной системы обнаружения вторжений, сводит на нет преимущества от использования программируемой логики.

С целью устранения указанных противоречий в данном исследовании предлагается организация вычислительного процесса таким образом, что

уникальная для каждого пользователя ресурсоемкая процедура синтеза цифровой схемы выполняется не локально на каждой отдельной реконфигурируемой системе киберзащиты, а централизованно с помощью грид-инфраструктуры. Результатом проведения вычислений в гриде являются файлы конфигураций для модификации клиентских ССОВ либо антивирусных средств, построенных на базе реконфигурируемых ускорителей. Такой подход позволяет существенно сократить общие вычислительные и эксплуатационные затраты на создание конфигураций, повысить удобство использования (usability), гибкость и масштабируемость средств информационной защиты.

3. Принципы организации процесса централизованной генерации конфигураций

Техническое решение сформулированного выше подхода представляет собой двухуровневую структуру, состоящую из большого числа реконфигурируемых устройств информационной безопасности на низшем уровне и централизованной системы синтеза конфигураций для них – на высшем.

Низший уровень составляют аппаратные ускорители на базе ПЛИС, предназначенные для защиты объектов обработки информации (различных как по масштабу, так и по присутствующим рискам). Реконфигурируемые ускорители работают по схожим алгоритмам, но построены в общем случае на базе ПЛИС разных типов и различной вычислительной мощности.

Высший уровень централизованной системы синтеза конфигураций реализует следующие функции:

- оперативное пополнение имеющихся баз сигнатур актуальной информацией о недавно выявленных факторах злонамеренной активности (атаки, вирусы и т.п.);
- сбор исходных данных о текущих параметрах безопасности с каждого из защищаемых объектов;
- максимально быстрая генерация файлов конфигураций с учетом особенностей каждой клиентской ССОВ / антивирусной системы;
- оперативная доставка конфигураций потребителям.

Сформулированная выше функциональность может быть реализована с применением современных информационных технологий в виде интерактивного сервиса наподобие механизма централизованной рассылки обновлений в современных антивирусных приложениях. Следует отметить, что в качестве платформы для реализации подобного сервиса помимо грид-сети могут бать задействованы и другие высокопроизводительные и / или распределенные компьютерные технологии, в частности, облачные вычисления, центры обработки данных и т.п.

В настоящем исследовании в качестве высшего уровня выступает среда Украинского национального грида (УНГ).

Пользователи средств информационной безопасности в энергетике (системные администраторы, обслуживающий персонал, ответственные лица по киберзащите эксплуатирующих организаций и служб в энергетической отрасли и в смежных областях) не являются специалистами ни в области разработки реконфигурируемых аппаратных средств, ни в технологиях распределенных вычислений. Поэтому предложенную двухуровневую структуру необходимо дополнить промежуточной прослойкой в виде грид-сервиса, который, обладая дружественным интерфейсом, должен позволять в удобной форме передавать на верхний уровень первичные данные (перечень распознаваемых сигнатур, параметры ССОВ, характеристики реконфигурируемого ускорителя и т.п.), а также получать обратно и пересылать пользователям результаты вычислений в виде бинарных файлов для программирования ПЛИС. Изнутри данный грид-сервис от имени сервисного грид-сертификата (hostcert.pem) должен формировать и направлять грид-задания в вычислительную инфраструктуру, контролировать процесс их выполнения и автоматизировать процесс агрегации полученных файлов конфигураций.

4. Грид-сервис централизованного синтеза конфигураций

Начиная с 2015 года частично за счет финансовой поддержки со стороны целевой комплексной программы научных исследований НАН Украины «Грид-инфраструктура и грид-технологии для научных и научно-прикладных применений» по инициативе и при непосредственном участии специалистов ИПМЭ им. Г.Е. Пухова НАНУ ведутся работы по созданию и исследованию в грид-инфраструктуре УНГ сервиса централизованного синтеза конфигураций для аппаратных устройств кибербезопасности [1].

Данная разработка получила название STRAGS (Security Tasks Reconfigurable Accelerators Grid-Service – грид-сервис для реконфигурируемых ускорителей задач информационной безопасности).

На основании сформулированных в предыдущем разделе требований был разработан функционал, который должен быть реализован для успешной работы создаваемого грид-сервиса:

- на удаленных вычислительных узлах грид-инфраструктуры (CE – Computing Element) необходимо обеспечить наличие и возможность запуска специализированного программного обеспечения для синтеза реконфигурируемых устройств (аналогичного либо функционально эквивалентного фирменным САПР разработки конфигураций для ПЛИС);
- обеспечить прием от пользователей в унифицированной форме входных данных (VHDL-описаний аппаратных компонент, ucf-файлов временных/топологических ограничений и других вспомогательных файлов проекта) и их передачу на удаленный CE;
- прозрачный для пользователя запуск автоматического этапа создания

конфигураций с возможностью выявления ошибок на ранних стадиях;
- оперативный мониторинг процессов компиляции и генерации файлов конфигураций.

Из перечисленных функций в плане технической реализации наиболее сложной задачей представляется обеспечение функционирования фирменной САПР на узлах грид-инфраструктуры. При ее традиционном решении (когда каждый компонент программного обеспечения (ПО), необходимого для выполнения удаленно запускаемой задачи, инсталлируется на каждом грид-узле перед ее запуском) возникает ряд проблем, в частности:

- объем специализированного программного обеспечения синтеза цифровых схем, которое должно быть предустановленно на удаленном грид-узле, даже в случае использования усеченного комплекта пакетов составляет порядка 10 Гб, и его передача одновременно с файлами грид-задания при каждым запуске представляется нецелесообразной;
- упомянутое программное обеспечение для своего функционирования требует наличия ряда системных библиотек (в т.ч. 32-битных), которые в общем случае отсутствуют на вычислительных узлах СЕ;
- в случае использования фирменного ПО лицензионные ограничения не позволяют установить его на каждый вычислительный узел каждого кластера грид-инфраструктуры, которая в случае УНГ насчитывает несколько тысяч процессоров [7].

В качестве оригинального решения данной проблемы было предложено использовать новейшую отечественную грид-технологию Rainbow ("ARC in the Cloud") [8]. Данная система изначально разрабатывалась для запуска специализированного ПО moldyngrid в рамках виртуальной организации medgrid, однако, впоследствии оказалась удачной разработкой и получила более широкое распространение. В настоящее время число грид-узлов Украинского национального грида, поддерживающих данную технологию, стабильно растет.

Суть подхода заключается в запуске на удаленных узлах грид-инфраструктуры в качестве грид-задачи виртуальной машины со всем необходимым предустановленным и настроенным программным обеспечением. Кроме того, что важно отметить, в рамках создания технологии отлажены и отработаны механизмы интерактивного взаимодействия пользователя с запущенными на виртуальных машинах программными пакетами в реальном масштабе времени.

При использовании такого подхода решение рассматриваемой в данном исследовании задачи сводится к созданию унифицированного образа виртуальной машины, включающего все необходимое для генерации конфигураций инструментальное ПО. Такой образ может быть успешно запущен на любом из узлов грид-инфраструктуры УНГ, поддерживающем технологию Rainbow.

С целью упрощения и унификации процедуры передачи пользователями входных данных грид-сервису STRAGS, все необходимые для компиляции конкретного проекта компоненты – VHDL-описания узлов цифровой схемы, ucf-файлы и другая сопутствующая информация – помещаются в единый архивный файл, который распаковывается в среде виртуальной машины после запуска задания на целевом грид-узле. Удобство такого приема заключается в том, что пользователь выполняет первый этап разработки цифровой схемы локально на своем компьютере в привычной для него графической среде на фирменной САПР, а все необходимые для архива файлы копирует из директории текущего проекта, не вникая в их состав и назначение. После чего второй, более ресурсоемкий по затратам процессорного времени этап синтеза выполняется на свободном в данный момент высокопроизводительном сервере в распределенной суперкомпьютерной сети.

Данный алгоритм также обеспечивает автоматическую передачу на верхний уровень системы сведений о семействе ПЛИС и ее конкретной модели, выбранной для реализации локального реконфигурируемого ускорителя информационной защиты.

Упомянутая выше возможность интерактивного взаимодействия с запущенным на виртуальной машине вычислительным процессом, позволяет реализовать в рамках грид-сервиса STRAGS возможность удаленного наблюдения за ходом выполнения стадий компиляции проекта. На рис. рис. 2 представлен один из экранов пользовательского интерфейса грид-сервиса, на котором отображается прогресс выполнения грид-задания, детализированный до конкретной процедуры компиляции (см. п. 1).

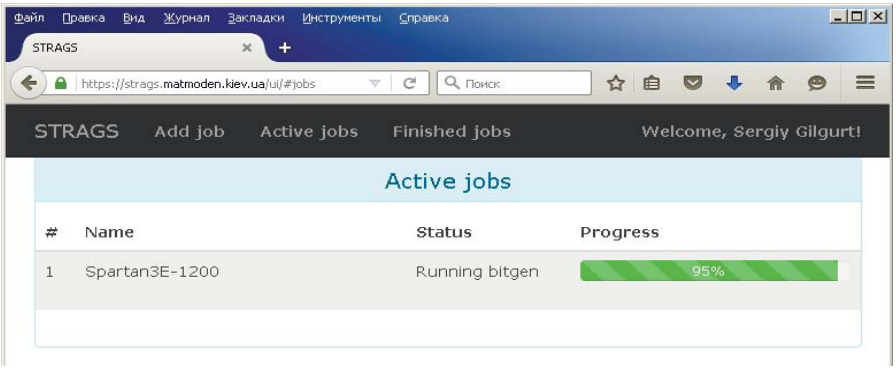

Рис. 2. Интерфейс пользователя грид-сервиса STRAGS. Экран выполнения задания

Использование стандартных возможностей грид-технологии не позволило бы достичь такой степени удобства взаимодействия с удаленным вычислительным процессом.

Проведенные в рамках данного исследования эксперименты позволили проверить и подтвердить эффективность идеи переноса ресурсоемких

операций с локальных систем киберзащиты в высокопроизводительную вычислительную среду с целью централизованного выполнения. В настоящее время проводятся работы по дальнейшему развитию и совершенствованию созданной централизованной системы синтеза реконфигурируемых аппаратных ускорителей для решения ресурсоемких задач киберзащиты.

Выводы

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

Предложены и исследованы принципы организации процесса генерации конфигураций – реконфигурационных последовательностей битов, загружаемых в ПЛИС аппаратных ускорителей. Разработана двухуровневая структура соответствующей подсистемы. Проанализированы возникающие проблемы. Предложено решение, основанное на использовании технологии Rainbow. Разработан грид-сервис STRAGS, связывающий нижний и верхний уровни системы централизованной генерации файлов конфигураций.

Экспериментальным путем подтверждена состоятельность идеи о переносе ресурсоемких операций с локальных реконфигурируемых устройств киберзащиты на высокопроизводительные компьютерные системы.

1. *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 НАН України. – Київ, 2017. – Вип. 79. – С.3-8.

2. *Грушвицкий Р.И., Мурсаев А.Х., Угрюмов Е.П.* Проектирование систем на микросхемах программируемой логики. СПб.: БХВ-Петербург, 2002. – 608 с.

3. *Зотов В.Ю.* Проектирование цифровых устройств на основе ПЛИС фирмы XILINX в САПР WebPACK ISE. М.: Горячая линия – Телеком, 2003. – 624 с.

4. Реконфигурируемые вычислительные системы: Основы и приложения. / А.В. Палагин, В.Н. Опанасенко. – К.: «Просвіта», 2006. – 280 с.

5. *Угрюмов Е.П.* Цифровая схемотехника: Учеб. пособие для вузов. – 3-е изд., перераб. и доп. – СПб.: БХВ-Петербург, 2010. – 816 с.

6. *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 НАН України. – Київ, 2012. – Вип. 65. – С.94-103.

7. UA-Grid: Украинская национальная грид-программа / А.Г. Загородний, С.Я. Свистунов, Л.Ф. Белоус, А.Л. Головинский // International Conference "Parallel and Distributed Computing Systems" PDCS 2013 (Ukraine, Kharkiv, March 13-14, 2013), pp.346-356.

8. *Сальников А.А., Вишневский В.В., Борецкий А.Ф.* «Платформа як сервіс» у грід для інтерактивного аналізу медичних даних // Математичні машини і системи. – 2015. – № 1. – С.53-64.

Поступила 2.10.2017р.