

Г. С. Елизаров, В. С. Горбунов, А. Г. Титов

Аппаратно-программная платформа моделирующей гетерогенной ВС (МГВС)

АННОТАЦИЯ. В работе представлена аппаратно-программная реконфигурируемая платформа на базе плат, содержащих программируемые логические интегральные схемы (ПЛИС).

Подробно описаны основные компоненты, приведены технические характеристики. На основе предложенных компонент будет создана моделирующая гетерогенная вычислительная система (МГВС), предназначенная для моделирования работы суперкомпьютера экзафлопсного уровня производительности.

Ключевые слова и фразы: Реконфигурируемая вычислительная система, программируемая логическая интегральная схема, моделирующая гетерогенная вычислительная система, шина и коммутатор PCI Express.

Введение

Моделирующая гетерогенная вычислительная система (МГВС) предназначена для моделирования работы суперкомпьютера, объединяющего до нескольких миллионов процессорных ядер. МГВС содержит наиболее важные технические и программные решения моделируемой системы. МГВС предполагается строить из вычислительных узлов (ВУ) содержащих платы ускорителей с ПЛИС. ВУ объединяются с использованием сети InfiniBand.

1. Базовые аппаратные компоненты плат ускорителей с ПЛИС

Базовые аппаратные компоненты, реализованные на ПЛИС семейства Virtex-6 фирмы Xilinx, рассмотрены в [1]. В настоящее время выпускаются ПЛИС новых семейств Virtex-7 и Kintex-7. Семейство Kintex-7 характеризуется наилучшим соотношением про-

изводительность/стоимость, а Virtex-7 ориентировано для достижения максимальной производительности. Для обеспечения процесса моделирования в ПЛИС реализуются контроллеры памяти, набор процессорных ядер и контроллер PCI Express. Функциональность контроллера памяти ПЛИС расширена для обеспечения работы процессорных ядер “по готовности данных” на аппаратном уровне.

На ПЛИС семейства Virtex-7 будет реализован многоядерный многопоточковый процессор с общей памятью, содержащий порядка тысячи предельно компактных 32-х разрядных RISC-ядер. Микропроцессорные ядра не будут содержать аппаратные блоки для выполнения операций над числами с плавающей точкой [2].

На ПЛИС семейства Kintex-7 также будет реализован многофункциональный многопоточковый ускоритель. Ускоритель будет построен с использованием архитектуры HSA – гетерогенных систем с общей для разнородных вычислительных элементов памятью, поддержкой управления очередью задач. Здесь будут аппаратные блоки выполнения операций над числами с плавающей точкой. С учетом этих особенностей ФГУП «НИИ «Квант» и ООО «НПО «Роста» разработали новые изделия на основе семейств ПЛИС Kintex-7 и Virtex-7.

Ниже описаны базовые аппаратные компоненты, которые предлагается использовать при построении ВУ МГВС.

1.1. Реконфигурируемая вычислительная плата на базе Xilinx Kintex-7 KV8-K7-410-01 «Топаз-2»

Плата «Топаз-2» разработана в ФГУП «НИИ «Квант». Основные технические характеристики платы:

- размер платы: 375x150 mm;
- сервисная ПЛИС: Xilinx Virtex-6 (LX75/130/195/240T) FF(G)784;
- сервисная ПЛИС поддерживает PCI Express Gen1 x4 или Gen2 x4;

- восемь рабочих ПЛИС: Xilinx Kintex-7 (XC7K325/410) FF(G)900;
- все рабочие ПЛИС поддерживают PCI Express Gen1 x4 или Gen2 x4;
- сеть PCI Express: одна СБИС PEX8648 коммутатора PCI Express Gen2 x4 (используются 11 портов), все ПЛИС подключены к PEX8648;
- микросхема FLASH памяти: одна конфигурационная Flash 512 Мб;
- память SDRAM: 1Gb DDR3 SDRAM DDR-800;
- сеть Ethernet: Fast Ethernet Base-T.

Внешний вид платы показан на Рис.1.

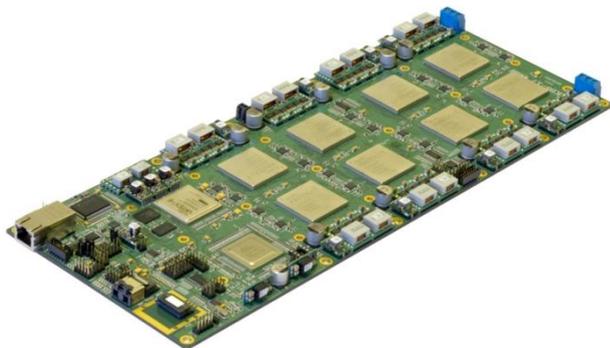


Рис.1. Внешний вид платы «Топаз-2»

Системная ПЛИС и 8 рабочих ПЛИС подключены к коммутатору PCI Express, реализованному на микросхеме PEX8648 (12 портов x4) фирмы PLX Technology. Объединение ПЛИС реализовано по схеме «звезда». Дополнительно все ПЛИС объединены в кольцо (кольцо RocketIOx4), что позволяет организовать высокоскоростной канал обмена данными. Кроме этого, все ПЛИС подключены к общей шине для организации управления. Загрузку конфигурационных данных в рабочие ПЛИС осуществляет сервисная ПЛИС через 8 независимых каналов в режиме Slave Serial. Для организации мониторинга основных компонентов на плате реализо-

ван внешний сетевой интерфейс Ethernet, а для связи с внешними PCI Express устройствами – два канала PCIe4.

1.2. Реконфигурируемая вычислительная плата на базе Xilinx Virtex-7 RC-47

В ООО «НПО «Роста» разработана реконфигурируемая вычислительная плата RC-47, предполагающая установку 4 рабочих ПЛИС Xilinx семейства Virtex-7.

Основные технические характеристики платы:

- размер платы: 280x150 mm;
- сервисная ПЛИС: Xilinx Spartan-6;
- четыре рабочих ПЛИС: Xilinx Virtex-7 (XC7V585/1500/2000T, XC7VX330/485/690T) FF(G) 1761;
- все рабочие ПЛИС поддерживают PCI Express Gen2 (для серии XC7V) или Gen3 (для серии XC7VX);
- сеть PCI Express: одна СБИС PEX8732 коммутатора PCI Express Gen3 x4 (8 портов);
- микроконтроллер STM32F, который обеспечивает контроль напряжений и температур, загрузку конфигураций в системный и рабочие ПЛИС, настройку и мониторинг коммутатора PCI Express по шине I2C;
- ножевой разъем, предназначенный для установки платы в стандартный слот расширения PCI Express;
- внешний разъем PCIe iPass 1, обеспечивающий подключение платы к серверу управления;
- внешний разъем PCIe iPass2, обеспечивающий возможность коммутации двух плат RC-47 между собой;
- опционально возможна установка на плату 2 слотов EM-4 для подключения внешней памяти – по одному на два рабочих ПЛИС;
- внешние интерфейсы JTAG, I2C, USB.

Внешний вид платы показан на Рис. 2.



Рис. 2. Внешний вид платы RC-47

Электропитание платы осуществляется через два 8-контактных разъема типа MiniFit от источника питания +12В.

Плата RC-47 предусматривает установку двух мезонинных модулей памяти EM4-DDR3, каждый из которых имеет 4 банка памяти объемом 512 Мбайт (128×32). При установке модулей EM4-DDR3 у каждого ПЛИС на плате появляются по два банка памяти.

1.3. Интерфейсная плата PCI Express КВА-P8619-PCI-E-01 «Янтарь-1»

Плата «Янтарь-1» разработана в ФГУП «НИИ «Квант». Основные технические характеристики платы:

- размер платы: 167,65x111,15 mm;
- микросхема коммутатора: PCI Express PEX8619;
- два внешних порта: PCI Express Gen2 x4 (до 40 Гбит/с);
- интерфейс с ПЭВМ: PCI Express Gen2 x8 (до 80 Гбит/с).

Плата «Янтарь-1» предназначена для подключения внешних устройств по интерфейсу PCI Express с использованием специализированного кабеля. Плата является компьютерной платой расширения в формате PCI Express, устанавливается в разъем PCI Express 16x или 8x материнской платы и выводит шину PCI Express на внешние порты 4x. «Янтарь-1» разработана в соответствии со спецификацией PCI Express 2.0. Подключение внешних устройств осуществляется по медному или оптическому кабелю.

Плата «Янтарь-1» построена на микросхеме коммутатора PCI Express PEX8619 фирмы PLX Technology. Микросхема позволяет коммутировать внутреннюю шину PCI Express x8 на один или два внешних разъема PCI Express x4.

Внешний вид платы показан на Рис 3.



Рис 3. Внешний вид платы «Янтарь-1»

Микросхема PEX8619 представляет собой не блокируемый коммутатор каналов PCI Express с низкой задержкой и поддержкой DMA (прямой доступ к памяти). PEX8619 имеет аппаратные и программные возможности для конфигурации режимов работы.

При подаче питания на коммутатор PEX8619 считывается начальная конфигурация, которая задается аппаратно резисторами, переключателями или перемычками на плате. При наличии PROM памяти, подключенной к PEX8619, чтение конфигурации произво-

дится из нее и заносится в регистры PEX8619. Наличие энергонезависимой памяти значительно ускоряет процесс конфигурации и инициализации коммутатора.

Доступ к PROM памяти для изменения значений осуществляется через PEX8619 по шине PCI Express с помощью программных средств PLX SDK или программ сторонних производителей. Также во время работы коммутатора по шине PCI Express можно изменять конфигурационные параметры, записывая значения непосредственно в конфигурационные регистры PEX8619.

2. Компоненты программного обеспечения

Ниже описаны компоненты программного обеспечения, которые предлагается использовать при построении ВУ МГВС.

2.1. Программное обеспечение ПЛИС, интерфейс пользователя

Программное обеспечение ПЛИС предназначено для создания проекта пользователя с использованием систем автоматизированного проектирования (САПР) ПЛИС. Для обеспечения возможности пользователю создавать свой собственный проект, разработан базовый проект для рабочих ПЛИС. Базовый проект рабочих ПЛИС и программные компоненты, входящие в его состав, необходимы для создания пользовательских проектов в САПР ISE компании Xilinx версии 13.3 и выше. Верхний уровень базового проекта рабочих ПЛИС описан на языке VHDL. На верхнем уровне базового проекта подключены следующие компоненты второго уровня:

- HOST – интерфейс сопряжения с шиной PCI Express, реализует аппаратную часть обмена с универсальным микропроцессором в командном (по одному слову) и блочном (пачками слов) режимах;

- BLOCK_IO – блок ввода-вывода, реализующий исполнительные сигналы пользовательского интерфейса и интерфейса HOST;

- BRAM1 – память произвольного доступа, предназначена для организации записи данных в исполнительные устройства пользователя;

- FIFO – память последовательного доступа для организации обмена данными с исполнительными устройствами пользователя;

- BRAM2 – память произвольного доступа, предназначена для организации чтения данных из исполнительных устройств пользователя;

- USER – компонент, объединяющий исполнительные устройства пользователя.

С использованием предложенных компонентов пользователь может самостоятельно спроектировать свои исполнительные устройства.

2.2. Комплект программного обеспечения для взаимодействия с ПЛИС

Комплект программного обеспечения для взаимодействия с ПЛИС реконфигурируемых плат разработан ФГУП «НИИ «Квант».

Комплект содержит:

- драйвер;
- коммуникационную библиотеку;
- программы и утилиты конфигурирования, управления и диагностики ПЛИС;
- примеры и тесты.

Драйвер выполняет следующие функции:

- верификацию PCI Express интерфейса ПЛИС;
- исправление конфликтов PCI Express интерфейса ПЛИС с BIOS и операционной системой;
- создание и регистрацию ПЛИС-устройств;
- выделение и освобождение участков памяти в PCI-области для каждой ПЛИС.
- отображение областей памяти PCI-устройств в область виртуальных адресов хост-программы;
- поддержку горячего отключения/подключения (горячая замена) без перезагрузки драйвера;

- отработку запросов коммуникационной библиотеки по низкоуровневому управлению ПЛИС;
- программную поддержку взаимодействия с подсистемой мониторинга по температурным состояниям ПЛИС;
- маскирование деградирующих ПЛИС.

Коммуникационная библиотека выполняет следующие функции:

- настройку и предоставление программного интерфейса ввода/вывода между программой пользователя и элементами памяти ПЛИС (регистры, BRAM, FIFO, внешняя подключаемая память) в режимах пословной и блочной передачи данных;
- поддержку сквозной логической нумерации ПЛИС, независимой от их аппаратного расположения (наличия) на реконфигурируемых платах;
- корректную работу интерфейса ввода/вывода в многопользовательском режиме;
- корректную работу интерфейса ввода/вывода в параллельном режиме;
- корректную работу интерфейса ввода/вывода при одновременной работе с разными сериями ПЛИС;
- корректную работу интерфейса ввода/вывода с выделенными в пользовательской области элементами памяти ПЛИС (BRAM, FIFO) разного размера;
- корректное завершение и освобождение ресурсов при возникновении событий, связанных с нажатием «Ctrl-C» (экстренное завершение программы), поступлением сигнала «Kill» и возникновении ошибок сегментации оперативной памяти;
- предоставление программного интерфейса ввода/вывода для функций интерфейса Mitrion Mithal;
- предоставление программного интерфейса ввода/вывода для функций интерфейса Mentor Catapult C.

Программы и утилиты конфигурирования, управления и диагностики ПЛИС выполняют следующие функции:

- загрузку битовых последовательностей с проверкой контрольной суммы;
- частичную загрузку битовых последовательностей с сохранением работоспособности интерфейса ввода/вывода;
- поддержку идентификации используемого оборудования;

- выдачу информации о состоянии PCI Express интерфейса ПЛИС;
- расширенную диагностику состояния PCI Express устройств (ПЛИС и PCI Express мостов).

Наборы примеров и тестов обеспечивают проверку работоспособности ПЛИС и её памяти, содержат последовательные и параллельные тесты ввода/вывода, работающие в пословном и блочном (DMA) режимах передачи данных, для различных типов памяти (регистры, BRAM, FIFO), содержат программы-примеры на Mittrion и CatapultC.

Комплект программного обеспечения для взаимодействия с ПЛИС реконфигурируемых плат работает под управлением ОС Linux с ядром версии не ниже 3.0.

2.3. Программное обеспечение САПР ПЛИС

Для разработки пользовательских проектов ПЛИС используются продукты трех ведущих в этой области компаний: Mentor Graphics, Xilinx и Synopsys. В принципе, для целей проектирования, минимально необходимым комплектом является комплект САПР фирмы Xilinx. Продукты Mentor Graphics расширяют этот комплект средствами удобного графического ввода данных проекта и мощной системой моделирования. Синтезатор фирмы Synopsys позволяет существенно повысить эффективность проектов.

2.3.1. САПР компании Mentor Graphics

Для разработки пользовательских проектов ПЛИС используются следующие продукты компании Mentor Graphics: HDL-редактор HDL Designer Series, симулятор ModelSim, средство синтеза Precision Synthesis, транслятор Catapult C. Эти САПР работают под управлением операционных систем Windows, HP-UX, Solaris, Linux.

САПР HDL Designer Series предоставляет набор текстовых, графических и табличных инструментов для формирования HDL-описания, а также средства визуализации и отладки. В частности,

HDL Designer включает текстовый редактор, редактор блок-схем, редактор конечных автоматов, редактор таблиц истинности логических функций, редактор алгоритмов (flow chart), редактор диаграмм соединений (IBD). В последних версиях САПР имеется поддержка работы с функциональными блоками, описанными на языке С (при помощи программы Catapult C).

САПР Precision Synthesis представляет собой семейство средств синтеза, позволяющее синтезировать HDL-описания на языках VHDL, Verilog (всех модификаций), EDIF. Precision позволяет выполнять как логический, так и физический синтез с привязкой к архитектуре конкретного типа ПЛИС и выполнением соответствующей оптимизации, ориентированной, в зависимости от настроек, на оптимизацию по объёму проекта или по быстродействию.

В Precision Synthesis предусмотрена возможность сокращения времени повторного синтеза проекта после внесения в него изменений, за счёт ретрансляции только изменённой части (это может быть выполнено в автоматическом режиме, либо можно явно указать блок, который должен подвергнуться ретрансляции).

Средство моделирования ModelSim SE работает под управлением ОС Windows (32/64), Linux (32/64) и Solaris. В семействе ModelSim версия SE обладает полным набором функциональных возможностей, позволяющим моделировать высокоуровневые и низкоуровневые (уровень RTL или «вентильный» уровень) описания без ограничения на размер проекта. Modelsim SE поддерживает языки VHDL, Verilog и SystemC. Пакет обеспечивает достаточно высокий уровень наглядности за счет нескольких средств визуализации, из которых основное – Waveform, отображающее диаграммы изменения значений сигналов во времени. В Modelsim SE также реализована технология ускоренного повторного моделирования при внесении изменений – для блоков проекта, не подвергавшихся изменениям, перекомпиляция не производится. Modelsim SE предъявляет повышенные требования к объёму памяти ПК – для работы с большими ПЛИС желателен объём памяти не менее 12 Гбайт.

Пакет Catapult C выполняет трансляцию описания на адаптированном языке программирования C в HDL-описание, дальнейший синтез которого может быть выполнен в САПР Precision Synthesis.

2.3.2. САПР компании Xilinx

ПО компании Xilinx включает САПР ISE Design Suite и Vivado, которые работают под управлением операционных систем Windows (32/64), Red Hat Enterprise Linux, SUSE Linux Enterprise.

Пакет ISE Foundation (Integrated software environment) представляет собой систему сквозного проектирования, которая реализует полный цикл разработки цифровых устройств на базе ПЛИС, включая программирование кристалла. Управляющая оболочка Project Navigator предоставляет пользователю интерфейс для работы с проектом и управления всеми процессами проектирования и программирования ПЛИС. Запуск всех программных модулей пакета может осуществляться непосредственно в среде Project Navigator. Исходные описания проектируемых устройств могут быть представлены в текстовой форме с использованием языков HDL, в виде принципиальных схем или диаграмм состояний. В состав пакета включен схемотехнический редактор, редактор конечных автоматов и комплект библиотек. Кроме того, предусмотрена возможность импорта проектов, подготовленных в других системах проектирования, как описаний в формате EDIF.

HDL-редактор, входящий в состав пакета, позволяет формировать описания модулей проекта на языках VHDL, Verilog и ABEL HDL (последний – только для семейства микросхем CPLD) за счет использования шаблонов. Генератор логических ядер CORE Generator позволяет автоматически создавать блоки системного уровня с требуемыми характеристиками. HDL-синтез проектов может выполняться не только встроенными средствами Xilinx Synthesis Technology (XST), но и программами третьих фирм, например Synplify и Precision Synthesis, поддерживающими языки VHDL и Verilog. Встроенные приложения FloorPlanner (PlanAhead) и FPGA Editor предоставляют возможности по ручному управле-

нию размещением и трассировкой проекта в ПЛИС. Средство XPower позволяет произвести оценку потребляемой мощности при загруженной в ПЛИС конфигурации. Среда EDK (Embedded development kit) позволяет проектировать процессорные микросистемы с использованием встроенных в ПЛИС процессорных ядер или эмуляторов процессорных ядер Microblaze, реализованных на логических ресурсах кристалла.

Конфигурирование кристаллов семейств ПЛИС фирмы Xilinx, выполненных по различной технологии (CPLD и FPGA), осуществляется с помощью единого модуля Impact.

САПР ISE предъявляет высокие требования к быстродействию и объёму памяти ПК. Даже на процессорах с частотой 3 ГГц трансляция типового проекта для ПЛИС среднего размера (20 млн. условных вентиляей) может длиться более суток. Объём используемой динамической памяти при этом примерно пропорционален объёму необходимых для реализации проекта логических ресурсов и, например, для ПЛИС V6LX240T может превышать 6 Гбайт (требования к объёму памяти, конкретизированные для различных типов ПЛИС, представлены Xilinx в свободно распространяемой документации).

САПР Vivado выпускается с 2012 г. В состав Vivado входят программные модули, функционально аналогичные программным модулям ISE, и некоторые новые. Например, встроенное приложение Vivado HLS позволяет транслировать описания на языках C, C++, System C в HDL-описание. Указанные модули оптимизированы под выполнение на многоядерных процессорах. Первые пробы практического использования САПР Vivado для ПЛИС XC7K325T показали заметные преимущества этого средства над ISE как по времени трансляции проекта, так и по быстродействию реализованного в ПЛИС проекта.

В современных ПЛИС Xilinx (начиная с семейства Virtex-6) и, соответственно, в средствах САПР поддержан режим частичной реконфигурации (partial reconfiguration). Использование этого режима позволяет реконфигурировать часть ПЛИС (динамическую составляющую проекта), сохраняя другую часть (статическую со-

ставляющую) в рабочем состоянии. Сборка проекта, включающего статическую и динамическую составляющие (динамических может быть несколько) выполняется в среде ISE Plan Ahead.

Режим частичной реконфигурации даёт следующие преимущества. Достаточно один раз выполнить размещение в ПЛИС той части проекта, которая включает сложные интерфейсные IP-блоки, реализуемые на встроенных аппаратных ядрах ПЛИС (высокоскоростные трансиверы, ядра PCI-Express), и к которой предъявляются жёсткие требования по размещению и быстродействию, если она отнесена в статическую составляющую. Стабильность работы этой части не будет зависеть от особенности реализации других частей проекта и не будет нарушаться в ходе реконфигурации динамической составляющей ПЛИС. Кроме того, имеется возможность ускорения стартового конфигурирования ПЛИС за счёт загрузки проекта с «пустой» динамической составляющей. Это позволяет управляющей операционной системе быстрее инициализировать системные ресурсы, реализованные в ПЛИС (в ряде случаев от скорости зависит корректность инициализации).

2.3.3. САПР компании Synopsys

САПР Synplify компании Synopsys является средством логического синтеза и работает под управлением ОС Windows, Solaris, Linux. Пакет Synplify Pro выполняет оптимизацию исходного языкового описания и генерирует низкоуровневое описание (например, в формате EDIF). Кроме того, Synplify предоставляет набор средств графической визуализации: средство изображения блок-схем уровня RTL (register transfer language – язык регистровых пересылок), средство визуализации диаграмм автоматически выделенных из описания проекта конечных автоматов, а также средства контроля и управления временными характеристиками проекта. Пакет позволяет интегрировать IP-ядра системного уровня, в том числе микропроцессорные системы, реализованные в среде EDK

ISE.

Преобразование исходного описания, выполняемое пакетом Synplify Pro, носит универсальный характер: результат практически не зависит от типа микросхемы, принимается во внимание по существу только объём доступных логических ресурсов.

Заключение

Представленная аппаратно-программная платформа на базе плат, содержащих ПЛИС, может служить основой для построения моделирующей гетерогенной системы. МГВС, включающая несколько десятков тысяч ПЛИС, сможет моделировать работу многопроцессорной вычислительной системы, содержащей несколько миллионов ядер, что соответствует представлению о будущем экзафлопсном суперкомпьютере.

Список литературы

- [1] Горбунов В.С., Титов А.Г. *Аппаратная платформа суперкомпьютеров с использованием ПЛИС* // «Перспективные направления развития средств вычислительной техники». Научно-техническая конференция ОАО «НИЦЭВТ».
- [2] Елизаров С.Г., Лукьянченко Г.А., Корнеев В.В. *Создание вычислительной системы для моделирования суперкомпьютера с производительностью экзафлопсного уровня* // «НСКФ-2012».

Об авторах:

Георгий Сергеевич Елизаров

Кандидат физико-математических наук. Директор ФГУП «НИИ «Квант».

e-mail: elizarov@rdi-kvant.ru

Виктор Станиславович Горбунов

Заместитель директора ФГУП «НИИ «Квант» по научной работе.

e-mail: gorbunov@rdi-kvant.ru

Александр Георгиевич Титов

Кандидат технических наук. Главный инженер научно-исследовательского отделения ФГУП «НИИ «Квант».

e-mail: titov@rdi-kvant.ru

G.S. Elizarov, V.S. Gorbunov, A.G. Titov.

Hardware-software platform of modeling heterogeneous computing system (MHCS).

ABSTRACT. In work the hardware-software platform on the basis of the reconfigurable boards containing the FPGA is presented. The main components are in detail described, technical characteristics are given. On the basis of offered the component can be created the modeling heterogeneous computing system (MHCS) intended for modeling of a exaFLOPS supercomputer.

Key Words and Phrases: Reconfigurable computing board, Field-Programmable Gate Array, the modeling heterogeneous computing system, PCI Express switch.