

ОРГАНИЗАЦИЯ И ПРОПУСКНАЯ СПОСОБНОСТЬ ИНТЕРФЕЙСА PCI-E УСКОРИТЕЛЬНОЙ ПЛАТЫ ALVEO U250 С УПРАВЛЯЮЩИМ КОМПЬЮТЕРОМ

Качинский М. В., Петровский Н. А., Станкевич А. В.

Кафедра электронных вычислительных средств,

Белорусский государственный университет информатики и радиоэлектроники

Минск, Республика Беларусь

E-mail: {kachinsky, nick.petrovsky, stankevich}@bsuir.by

Рассматриваются вопросы построения интерфейса для обмена данными между Alveo U250 и управляющим компьютером по стандарту PCI Express Gen3. Получена фактическая пропускная способность между ускорительной платой Alveo U250 и хост-компьютером под управлением ОС Linux по шине PCI-e с использованием регистрового AXI4-Lite и отображаемого на память AXI4-MM интерфейсов.

ВВЕДЕНИЕ

Последние несколько десятилетий экспоненциально возрос объем обрабатываемых данных в области машинного обучения, специализированных алгоритмов поиска в базах данных, криптографических системах и т.п. Таким же образом изменился и спрос на высокопроизводительные энергоэффективные вычислительные системы для специализированных задач обработки данных.

Ускорение обработки больших объемов информации только за счет эксплуатации параллелизма процессоров общего назначения и графических карт (CPU/GPU) не всегда эффективно из-за фиксированной архитектуры таких систем [1]. Одним из решений данной проблемы является использование адаптируемых реконфигурируемых ускорителей (ускорительных карт для компьютера) на базе ПЛИС, которые в составе вычислительных систем общего назначения (совместно с CPU/GPU) позволят обеспечить требуемый уровень производительности для решения задач с интенсивными вычислениями.

Фирма Xilinx выпускает семейство адаптируемых карт ускорителей Xilinx Alveo Data Center в форм-факторе PCIe, с широким спектром готовых приложений, которые обеспечивают быстрое развертывание адаптивных ускорителей, что приводит к значительному увеличению вычислительных возможностей и экономии эксплуатационных расходов (TCO — от англ. Total Cost of Ownership) до 65%. Один из таких ускорителей Xilinx Alveo U250 Data Center (далее Alveo U250) представляет собой совместимую с PCI Express Gen3 x16 карту [2]. Карта предназначена для ускорения приложений со сложными вычислительными операциями, таких как машинное обучение, обработка медиаданных.

I. ВЗАИМОСВЯЗЬ ПОДСИСТЕМЫ PCIe И АППАРАТНЫХ ИНТЕРФЕЙСОВ

В настоящем докладе рассматриваются вопросы построения интерфейса для обмена дан-

ными между Alveo U250 и управляющим компьютером.

Alveo U250 выполнена на базе FPGA XCU250, имеющей в своем составе интегрированный блок PCI Express. Для доступа к этому интерфейсу необходимо использовать IP-ядро фирмы Xilinx. В качестве такого ядра может использоваться DMA/Bridge Subsystem for PCI Express v4.1 [3] (далее PCIe-DMA). Данная версия IP-ядра поддерживает, в частности, XCU250.

PCIe-DMA может быть сконфигурирована как высокопроизводительное устройство прямого доступа в память (модуль DMA) для обмена данными между PCI Express и пользовательским устройством, либо как мост между PCI Express и пространством AXI memory. Как модуль DMA, ядро можно использовать для перемещения блочных данных между адресным пространством PCIe и адресным пространством AXI, используя специальный драйвер. При настройке в качестве моста PCIe полученные пакеты PCIe преобразуются в трафик AXI, а принятый трафик AXI преобразуется в трафик PCIe. Функциональность моста идеально подходит для периферийных устройств AXI, которым требуется быстрый и простой способ доступа к PCI Express. В этом режиме ядро может использоваться в качестве конечной точки или корневого порта. Для реализации внутрисистемного интерфейса аппаратной платформы IP-ядро PCIe-DMA целесообразно использовать в режиме DMA, т.к. в этом режиме ядро может быть дополнительно сконфигурировано таким образом, чтобы обеспечить прямое подключение к пользовательской логике (RTL logic). AXI является частью ARM AMBA (Advanced Microcontroller Bus Architecture). В настоящее время используется 4 версия интерфейса.

Существует три типа интерфейсов AXI4 [4]:

- AXI4 memory-mapped (AXI4-MM) — для высокопроизводительных приложений;
- AXI4-Lite — для простых приложений с низкой пропускной способностью (напри-

мер, для чтения/записи регистров управления и состояния);

- AXI4-Stream — для высокоскоростной потоковой передачи данных.

Протоколы с отображением в память AXI4-MM и AXI4-Lite во всех транзакциях используют концепцию целевого адреса в пространстве системной памяти для передаваемых данных. Протокол AXI4-Stream используется для приложений, которые обычно ориентированы на высокоскоростной обмен данными. В этом протоколе понятие адреса не используется. Каждый AXI4-Stream интерфейс действует как один однонаправленный канал, использующий при передаче данных механизм «рукопожатия» (handshake).

PCIe-DMA обеспечивает пересылку данных между памятью хоста и подсистемой прямого доступа к памяти (DMA). Передача данных может быть от хоста к ускорительной карте (H2C) либо в обратном направлении (C2H).

II. АРХИТЕКТУРА СИСТЕМЫ

Простейший вариант подключения PCIe-DMA к пользовательскому проекту заключается в использовании интерфейса AXI4-Lite Master Configuration port. В этом случае доступ хоста к регистрам управления и состояния в пользовательской логике обеспечивается через порт AXI4-Lite Master. Запись и чтение осуществляется 32-разрядными словами. Запрос на чтение или запись от хоста передается через PCIe к AXI4-Lite Master BAR0 (Base Address Register). При чтении данные, передаваемые от пользовательского проекта, возвращаются к хосту через внутреннюю шину CC PCIe-DMA (рисунок 1).

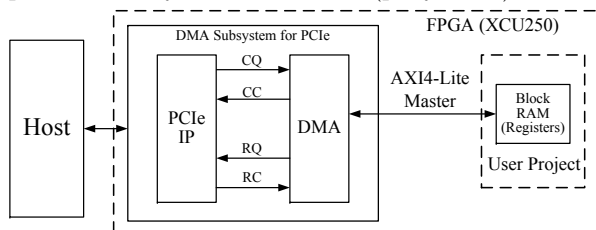


Рис. 1 – Использование интерфейса AXI4-Lite Master

Рассмотренный вариант подключения через интерфейс AXI4-Lite Master обладает относительно невысокой пропускной способностью, однако ее может быть достаточно для определенного класса задач.

Более высокую пропускную способность обеспечивает использование AXI4-MM интерфейса. При этом в пользовательском проекте должна использоваться блочная память с интерфейсом AXI4-MM (рисунок 2). Для H2C передачи PCIe-DMA читает данные от хоста и записывает их в блочную память на пользовательской стороне. Для обратной передачи (C2H) PCIe-DMA читает данные из блочной памяти на пользовательской стороне и записывает их в память хоста.

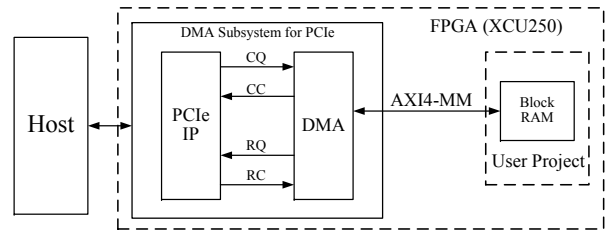


Рис. 2 – Использование AXI4-MM интерфейса

Для реализации дополнительных сигналов управления для доступа к хосту в дополнение к AXI4-MM интерфейсу можно использовать интерфейс AXI4-Lite Master.

Для взаимодействия с IP-ядром PCIe-DMA использовался стандартный драйвер от компании Xilinx — `xdma.ko`. Драйвер обеспечивает высокоскоростной доступ с использованием контроллера DMA к заданным конечными точками устройства (endpoint).

С использованием PCIe-DMA и AXI4-Lite Master реализован регистровый интерфейс. В интерфейсе имеется регистр данных и регистр управления/состояния. Также был реализован интерфейс на основе AXI4-MM.

III. ЭКСПЕРИМЕНТАЛЬНЫЕ РЕЗУЛЬТАТЫ

При исследовании пропускной способности интерфейсов использовались 2^{20} транзакций (запись/чтение) с последующим усреднением. При исследовании регистрового интерфейса обмен данными производился 32-разрядными словами. Было получено значение пропускной способности 56 Мбит/с. При исследовании интерфейса на основе AXI4-MM изменялся размер пакета (размер транзакции). Наиболее высокая пропускная способность была получена при максимальном размере блока (размере транзакции) равном 32768 байт (размер блока памяти на пользовательской стороне интерфейса AXI4-MM). При указанном размере блока и выключенных прерываниях (`poll_mode = 1`) получена максимальная пропускная способность порядка $\approx 6,8$ Гбит/с при чтении и $\approx 6,7$ Гбит/с при записи.

1. Breathe New Life into Your Data Center with Alveo Adaptable Accelerator Cards. WP499 (v1.0) November 19, 2018; [Электронный ресурс]. – Режим доступа: https://www.xilinx.com/support/documentation/white_papers/wp499-alveo-intro.pdf. Дата доступа: 19.11.2020.
2. ADM-PCIe-7V3; [Электронный ресурс]. – Режим доступа: <https://www.xilinx.com/products/boards-and-kits/1-4i8a6z.html>. Дата доступа: 19.11.2020.
3. DMA/Bridge Subsystem for PCI Express v4.1. Product Guide Vivado Design Suite PG195 (v4.1). June 20, 2019; [Электронный ресурс]. – Режим доступа: https://www.xilinx.com/support/documentation/ip_documentation/xdma/v4_1/pg195-pcie-dma.pdf. Дата доступа: 19.11.2020.
4. AXI Reference Guide UG761 (v14.3). November 15, 2012; [Электронный ресурс]. – Режим доступа: https://www.xilinx.com/support/documentation/ip_documentation/axi_ref_guide/latest/ug761_axi-reference_guide.pdf. Дата доступа: 19.11.2020.