

## ПРИМЕНЕНИЕ СТАНДАРТНЫХ ИНТЕРФЕЙСОВ ПЕРСОНАЛЬНЫХ КОМПЬЮТЕРОВ ДЛЯ ИНТЕГРАЦИИ ВЫЧИСЛИТЕЛЬНЫХ УСТРОЙСТВ В ПРОГРАММНО-АППАРАТНЫХ КОМПЛЕКСАХ

Тарасов И.Е., Люлява Д.В., Дуксин Н.А., Дуксина И.И.

*«МИРЭА - Российский технологический университет», 119454, Россия, г. Москва, проспект Вернадского, 78, e-mail: tarasov\_i@mirea.ru, lyulyava@mirea.ru, duksin@mirea.ru, duksina@mirea.ru*

---

В статье рассматриваются вопросы проектирования программно-аппаратных комплексов на базе персональных компьютеров, использующих локально подключенные аппаратные ускорители на базе графических процессоров (GPU) или программируемых логических интегральных схем (ПЛИС) с архитектурой FPGA или APSOC. Поскольку набор периферийных устройств в составе персонального компьютера ограничен, он определяет достижимую скорость передачи данных. Таким образом, производительность вычислений добавляемых в комплекс аппаратных ускорителей должна быть согласована с технически достижимой скоростью передачи входных данных в ускоритель и возврата вычисленных значений. Сопряжение с персональным компьютером также требует наличия контроллера соответствующего интерфейса в подключаемом ускорителе. Рассмотрены варианты подключения ускорителей на базе ПЛИС к персональному компьютеру с помощью интерфейса PCI Express для формирования программно-аппаратного комплекса, предназначенного для таких задач, как цифровая обработка сигналов, защита информации и распознавание регулярных выражений, реализуемые на базе ускорителя с параллельной архитектурой.

---

Ключевые слова: процессор, интерфейс, система на кристалле, аппаратная архитектура.

## APPLICATION OF STANDARD PERSONAL COMPUTER INTERFACES FOR INTEGRATION OF COMPUTING DEVICES IN SOFTWARE&HARDWARE COMPLEXES

Tarasov I.E., Lulyava D.V., Duksin N.A., Duksina I.I.

*«MIREA - Russian Technological University», 119454, Moscow, 78 Vernadskogo Avenue, Russia, e-mail: tarasov\_i@mirea.ru, lyulyava@mirea.ru, duksin@mirea.ru, duksina@mirea.ru*

---

The article discusses the design of software and hardware systems based on personal computers using locally connected hardware accelerators based on graphics processors (GPUs) or programmable logic integrated circuits (FPGAs) with FPGA or APSOC architecture. Since the set of peripheral devices included in a personal computer is limited, it determines the achievable data transfer rate. Thus, the computing performance of hardware accelerators added to the complex must be consistent with the technically achievable speed of transmitting input data to the accelerator and returning the calculated values. Interfacing with a personal computer also requires the presence of a corresponding interface controller in the connected accelerator. Options for connecting FPGA-based accelerators to a personal computer using the PCI Express interface to form a software and hardware complex designed for tasks such as digital signal processing, information security and regular expression recognition, implemented on the basis of an accelerator with parallel architecture, are considered.

---

Keywords: processor, interface, system on chip, hardware architecture.

### Введение

Проектирование высокопроизводительных вычислительных комплексов требует решения ряда проблем, связанных как с методическими подходами к разработке сложных систем, так и с технологическими проблемами микроэлектроники. Поэтому архитектурное проектирование программно-аппаратных комплексов необходимо выполнять с учетом ограничений, накладываемых существующими технологическими ограничениями. Также следует учитывать характеристики стандартных интерфейсов обмена данными, которые существенно ограничивают возможности масштабирования вычислительных систем.

Одним из актуальных эффектов является т.н. «стена памяти» или, в другой формулировке, «стена интерфейсов» [1]. Эффект заключается в опережающем росте производительности вычислений по сравнению с

производительностью передачи данных. Эффект усиливается по мере перехода к меньшим технологическим нормам, поэтому данная проблема требует прежде всего архитектурного решения. Дополнительно необходимо отметить повышенный уровень брака при изготовлении компонентов высокоскоростных интерфейсов, что является дополнительным фактором увеличения стоимости системы.

В то же время, при проектировании высокопроизводительных вычислительных комплексов применяется гетерогенная архитектура, содержащая CPU, GPU и ПЛИС с архитектурой FPGA, а также специализированные СБИС в различных сочетаниях. Наличие нескольких компонентов, требующих передачи данных между ними, формирует проблему выбора интерфейсов с соответствующей пропускной способностью, достаточной для организации совместной работы подсистем проектируемого комплекса. Поскольку CPU и GPU представляют собой массово выпускаемые компоненты, естественно ориентироваться на интерфейсные подсистемы, уже реализованные в устройствах на их базе – прежде всего персональных компьютерах. При этом специализация программно-аппаратных комплексов производится за счет добавления конфигурируемых компонентов – ПЛИС, а также, при необходимости, специализированных СБИС. В статье рассматриваются характеристики стандартных интерфейсных устройств, входящих в состав персональных компьютеров (ПК), и возможные архитектурные подходы к проектированию высокопроизводительных программно-аппаратных комплексов на их базе.

### **Характеристики высокопроизводительных интерфейсов, входящих в состав персональных компьютеров**

В современных персональных компьютерах наиболее распространенные интерфейсы, доступными без применения дополнительных контроллеров, обычно представлены набором PCI Express, Gigabit Ethernet и USB. Максимальная скорость передачи данных для версий стандарта USB приведены в таблице 1.

Таблица 1. Характеристики версий интерфейса USB

Название (год внедрения)	Максимальная скорость передачи данных
USB 4 2.0 (2022)	80 Гб/с
USB 4 (2019)	40 Гб/с
USB 3.2 (2017)	20 Гб/с
USB 3.1 (2013)	10 Гб/с
USB 3.0 (2008)	5 Гб/с
USB 2.0 (2000)	480 Мб/с
USB 1.1 (1998)	12 Мб/с, 1.5 Мб/с
USB 1.0 (1996)	12 Мб/с, 1.5 Мб/с

Ориентируясь на данные таблицы 1, можно предположить, что интерфейс USB, с учетом относительной простоты доступных микросхем контроллеров, является предпочтительным для подключения высокопроизводительных ускорителей к ПК. В то же время, указанные максимальные скорости передачи данных соответствуют пиковой скорости, измеряемой в течение короткого интервала времени, когда данные действительно физически передаются. Практическими ограничениями USB являются максимальный размер пакета данных и количество пакетов, передаваемых в секунду. Принципиальным ограничением, обеспечивающим совместимость широкого спектра устройств, оставалась величина 1000 пакетов в секунду, которая для USB 2.0 была увеличена до 8000 пакетов в секунду. При этом максимальный размер пользовательских данных (payload) в составе пакета также лимитируется практическими реализациями микросхем внешних контроллеров и для распространенных моделей FTDI, Infineon, WCH ограничено величинами 128-256 байт, что обеспечивает в итоге верхний предел практической скорости обмена данными 2-4 Мбит/с. Увеличение скорости за пределами этих величин предполагает применение периферийных контроллеров, поддерживающих прямой доступ к буферам приема и передачи, что также требует поддержки протокола USB в подключаемом устройстве (вместо того, чтобы использовать встроенную поддержку профилей SPP, HID или Mass Storage в готовых микросхемах контроллеров USB). Поэтому применение USB в варианте Serial Port Profile представляется приемлемым для низкоскоростного обмена данными, а также в случае, когда

подключаемый аппаратный ускоритель обладает высоким уровнем автономности и не требует постоянного обмена данными с ПК.

Интерфейс Ethernet, преимущественно в разновидности TriMode Ethernet (т.е. поддерживающий автоматическое определение скорости в 10, 100 или 1000 Мбит/с), повсеместно распространен в составе ПК и обеспечивает формально более высокие скорости по сравнению с практически достижимыми для простых модификаций контроллеров USB. В то же время, его использование совместно с другими устройствами, в первую очередь штатно подключаемой локальной сетью, увеличивает трафик в этой сети, что препятствует достижению высоких показателей скорости обмена. Также возможности, предоставляемые распространенными протоколами обмена данными по Ethernet, как UDP, так и, в особенности, TCP, представляются избыточными при локальном подключении устройств. Наконец, необходимость обеспечения программной поддержки стека протоколов TCP/IP или, по крайней мере, UDP/IP ужесточает требования к программной подсистеме контроллера подключаемого оборудования, который должен обеспечивать поддержку протоколов ARP, DHCP и ICMP в дополнение к UDP или TCP, на базе которых будет производиться основной обмен данными.

Преимущественным способом подключения локальных ускорителей является использование интерфейса PCI Express в варианте «конечной точки» (endpoint). При этом, наиболее простом, варианте от контроллера не требуется подключение нисходящих устройств, поэтому сложность и дополнительные накладные расходы относительно невысоки. Характеристики модификаций интерфейса PCI Express приведены в таблице 2.

Таблица 2. Характеристики интерфейса PCI Express: x2, x4, x8, x16

Название (год внедрения)	Максимальная скорость передачи данных (по одной линии)	Кодирование	Способ передачи сигнала
PCIe 7.0 (2025)	128.0 Гб/с	1b/1b (режим Flit)	PAM4
PCIe 6.0 (2022)	64.0 Гб/с	1b/1b (режим Flit)	PAM4
PCIe 5.0 (2019)	32.0 Гб/с	128b/130b	NRZ
PCIe 4.0 (2017)	16.0 Гб/с	128b/130b	NRZ
PCIe 3.0 (2010)	8.0 Гб/с	128b/130b	NRZ
PCIe 2.0 (2007)	5.0 Гб/с	8b/10b	NRZ
PCIe 1.0 (2003)	2.5 Гб/с	8b/10b	NRZ

При использовании PCI Express важной особенностью является использование высокоскоростных последовательных приемопередатчиков, которые представляют собой функционально сложные блоки (СФ-блоки), реализуемые в виде аналого-цифровой схемы. Это связано с жесткими требованиями к целостности сигналов на используемых скоростях обмена данными. Поэтому контроллер PCI Express должен присутствовать в составе СБИС ускорителя (или, по крайней мере, находиться в составе подсистемы ускорителя в виде отдельной микросхемы). В таблице 3 приведены варианты отладочных плат на базе ПЛИС AMD/Xilinx, которые могут выступать прототипами ускорителей для работы в составе программно-аппаратных комплексов.

Таблица 3. Модели отладочных плат на базе ПЛИС/СнК (FPGA/SoC) AMD/Xilinx с поддержкой PCIe

Название отладочной платы	Используемая ПЛИС	Скорость передачи данных	Интерфейс	Версия протокола PCIe
Xilinx Virtex UltraScale+ VU9P FPGA	Virtex UltraScale+	до 16 Гб/с	PCIe Gen3 x16	Gen3
Xilinx Virtex UltraScale+ VU13P FPGA	Virtex UltraScale+	до 16 Гб/с	PCIe Gen3 x16	Gen3
Xilinx Kintex UltraScale+ KU115 FPGA	Kintex UltraScale+	до 8 Гб/с	PCIe Gen3 x8	Gen3

AMD Alveo U250 Accelerator Card	XCU250	до 15.8 Гб/с	PCIe Gen3 x16	Gen3
AMD Alveo U280 Accelerator Card	XCU280	до 15.75 Гб/с	PCIe Gen4 x8 2xGen4 x8	Gen4
Xilinx Zynq UltraScale+ MPSoC ZCU102 Evaluation Kit	Zynq UltraScale+ MPSoC	до 2 Гб/с	PCIe Gen2 x4	Gen2
Xilinx Zynq-7000 ZC706 Evaluation Kit	Zynq-7000	до 8 Гб/с	PCIe Gen2 x16	Gen2

При выборе версии PCI Express необходимо иметь в виду, что физические скорости передачи данных выше 6,5 Гбит/с обеспечиваются только при использовании более дорогих и сложных в производстве корпусов СБИС типа flip-chip. Это связано с тем, что в корпусах типа wire bond индуктивность проволоки, соединяющей контакт кристалла СБИС с выводом корпуса, слишком высока для скоростей обмена свыше указанной величины. В то же время в СБИС типа «перевернутый кристалл» (flip chip) индуктивность проводника, соединяющего контактную площадку кристалла СБИС с соответствующим шариковым выводом корпуса, существенно меньше, что является общим свойством для высокоскоростных последовательных приемопередатчиков и интерфейсов на их основе [7].

### Архитектура программно-аппаратного комплекса на базе персонального компьютера

С учетом ограниченной пропускной способности стандартных интерфейсов ПК и особенностей протоколов обмена, приводящим к неравномерной мгновенной скорости обмена данными, необходимо предусматривать архитектурные решения, приводящие поток данных, передаваемый в ПК, к значениям, допустимым для используемых интерфейсов. Ранее было отмечено, что для ПК, совместимых с x86-64, наибольшая производительность передачи данных достигается при использовании интерфейса PCI Express, однако в зависимости от версии интерфейса и количества задействованных линий пиковая производительность может существенно различаться. Пакетный характер передачи данных дополнительно корректирует действительные значения пропускной способности интерфейса, делая скорость передачи пользовательских данных, зависимой от алгоритмов обмена данными и усредненного размера пакета. Поэтому предпочтительной архитектурой для создания высокопроизводительных масштабируемых программно-аппаратных комплексов представляется гетерогенная система с частично автономными вычислительными подсистемами, подключаемыми к ПК с помощью PCI Express (в ряде случаев с помощью менее производительных интерфейсов Ethernet и USB). Архитектура вычислительной системы, использующей аппаратные ускорители, соединенного с ПК с помощью одного из стандартных интерфейсов, показана на рисунке 1.

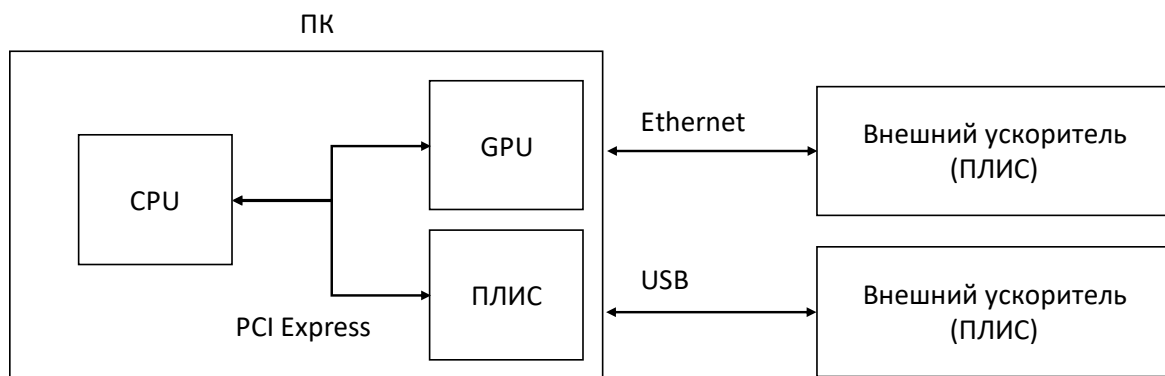


Рисунок 1. - Архитектура программно-аппаратного комплекса с подключением аппаратных ускорителей вычислений с помощью стандартных интерфейсов

Поскольку суммарная производительность вычислений ускорителей на базе ПЛИС или специализированных СБИС существенно выше, чем производительность передачи данных, предпочтительными классами реализуемых алгоритмов являются те, которые обеспечивают выполнение большого количества операций над переданными данными. Среди них можно перечислить вычисления, основанные на сверточных преобразованиях

данных (спектральный и вейвлет-анализ), цифровая фильтрация данных, вычисление хэш-функций, а также другие алгоритмы, реализуемые на базе конвейерных вычислительных структур. В этих случаях над относительно небольшим объемом передаваемых данных выполняетсякратно большее количество операций по их преобразованию.

### **Заключение**

Рассмотренная архитектура программно-аппаратного комплекса основана на совместном применении аппаратных платформ трех основных типов, используемых в настоящее время для создания высокопроизводительных систем – CPU, GPU и ПЛИС. В качестве наиболее производительного интерфейса передачи данных целесообразно использовать PCI Express, с учетом его практических ограничений по достигаемой скорости передачи пользовательских данных. Применение таких стандартных интерфейсов, как Ethernet и USB, несмотря на формально высокие показатели производительности передачи данных, следует использовать ограниченно, ориентируясь на практически достижимые показатели.

Работа выполнена в рамках государственного задания Министерства науки и высшего образования Российской Федерации (тема № FSFZ-2022-0004 Архитектуры специализированных вычислительных комплексов, методики, алгоритмы и инструменты проектирования цифровых вычислительных устройств).

### **Список литературы**

---

1. Hennessy J L, Patterson D A 2017 Computer Architecture. 6th Edition. A Quantitative Approach (The Morgan Kaufmann Series in Computer Architecture and Design) 936 p
2. <https://ftdichip.com/product-category/products/> (дата обращения 18.07.2024)
3. <https://www.infineon.com/cms/en/product/universal-serial-bus/usb-peripheral-controllers-for-superspeed/ez-usb-fx3-usb-5gbps-peripheral-controller/> (дата обращения 18.07.2024)
4. <https://wch-ic.com/products/CH340.html> (дата обращения 18.07.2024)
5. Y. A. Nedbailo et al., "Elbrus-2C3: a Dual-Core VLIW Processor with Integrated Graphics," 2021 International Conference Engineering and Telecommunication (En&T), Dolgoprudny, Russian Federation, 2021, pp. 1-5, doi: 10.1109/EnT50460.2021.9681771.
6. Тарасов, И. Е. Перспективы применения софт-процессоров в системах на кристалле на базе программируемых логических интегральных схем / И. Е. Тарасов, Д. С. Потехин, О. В. Платонова // Russian Technological Journal. – 2022. – Т. 10, № 3. – С. 24-33. – DOI 10.32362/2500-316X-2022-10-3-24-33. – EDN UVLVBA.
7. <https://www.xilinx.com/products/technology/high-speed-serial.html> (дата обращения 18.07.2024)

### **References**

---

1. Hennessy J L, Patterson D A 2017 Computer Architecture. 6th Edition. A Quantitative Approach (The Morgan Kaufmann Series in Computer Architecture and Design) 936 p
2. <https://ftdichip.com/product-category/products/> (access date 18.07.2024)
3. <https://www.infineon.com/cms/en/product/universal-serial-bus/usb-peripheral-controllers-for-superspeed/ez-usb-fx3-usb-5gbps-peripheral-controller/> (access date 18.07.2024)
4. <https://wch-ic.com/products/CH340.html> (access date 18.07.2024)
5. Y. A. Nedbailo et al., "Elbrus-2C3: a Dual-Core VLIW Processor with Integrated Graphics," 2021 International Conference Engineering and Telecommunication (En&T), Dolgoprudny, Russian Federation, 2021, pp. 1-5, doi: 10.1109/EnT50460.2021.9681771.
6. Тарасов, И. Е. Перспективы применения софт-процессоров в системах на кристалле на базе программируемых логических интегральных схем / И. Е. Тарасов, Д. С. Потехин, О. В. Платонова // Russian Technological Journal. – 2022. – Т. 10, № 3. – С. 24-33. – DOI 10.32362/2500-316X-2022-10-3-24-33. – EDN UVLVBA.
7. <https://www.xilinx.com/products/technology/high-speed-serial.html> (access date 18.07.2024)