

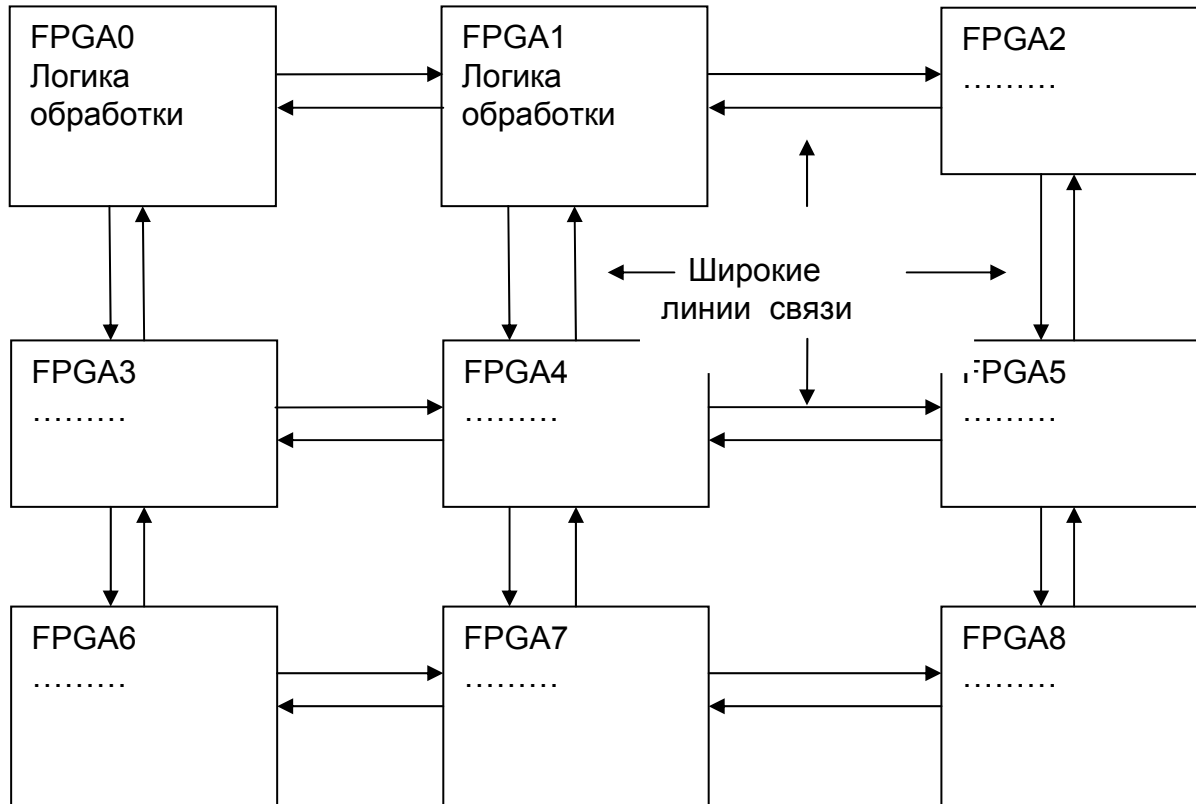
**Андреев С. С.,
Лацис А. О.,
Плоткина Е. А.**

ООО НПО «Роста», ИПМ им. М. В. Келдыша РАН.

**ИССЛЕДОВАНИЕ И СОЗДАНИЕ
СИСТЕМНОГО
АППАРАТНО-ПРОГРАММНОГО
ОБЕСПЕЧЕНИЯ ПЕРЕДАЧИ ДАННЫХ ПО
ЛИНИЯМ РОСКЕТИО ДЛЯ СИСТЕМ НА
КРИСТАЛЛАХ ПЛИС ФИРМЫ XILINX.**

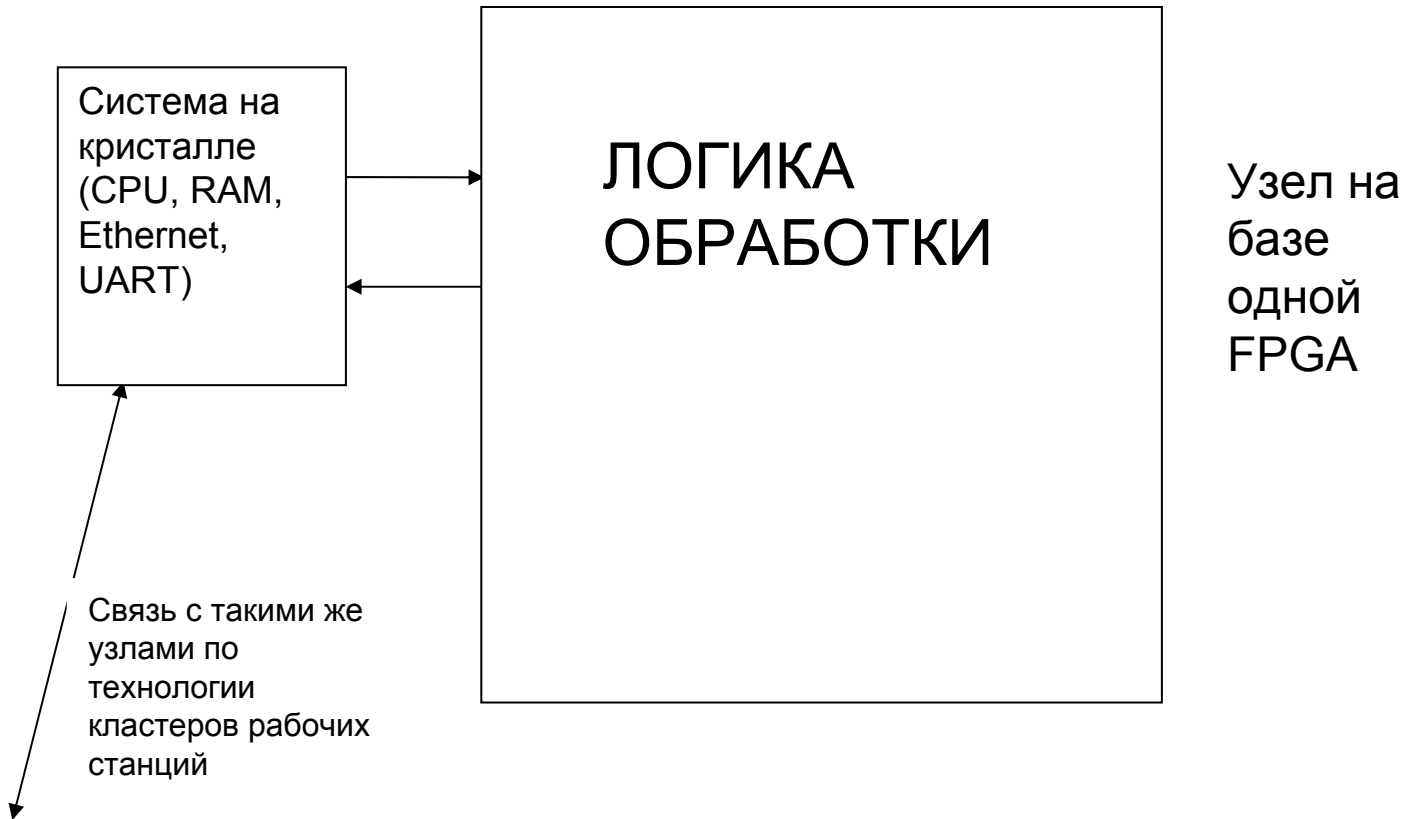
Как интегрировать комплексы из значительного (десятки, сотни) числа FPGA? Два способа:

1. «Силовой»: создаем конструкцию для объединения требуемого числа кристаллов в единую схему, потом эту схему реализуем.



Специально разработанная конструкция (плата или набор плат)

С использованием возможностей систем на кристалле (СнК).



Достоинства второго способа известны и очевидны, недостатки – в слабости коммуникационной среды. Кластеры из таких СнК настоятельно нуждаются в высокоскоростных сетях типа Myrinet или Infiniband.

- Специализированные аппаратные ядра, пригодные для реализации этих коммуникационных технологий, в Virtex-2pro присутствуют, называются RocketIO.
- В общей структуре программного обеспечения вычислительных кластеров ПО для высокоскоростных сетей традиционно стоит особняком. Оно не очень хорошо стандартизовано и интегрировано в общую структуру ПО. Это позволяет нам не стремиться реализовать в точности Myrinet или Infiniband, а ограничиться функциональным аналогом, простым, малым по объему оборудования и эффективным.
- Такое решение было разработано и получило название ***устройства rocketmem***.

Что получает программист.

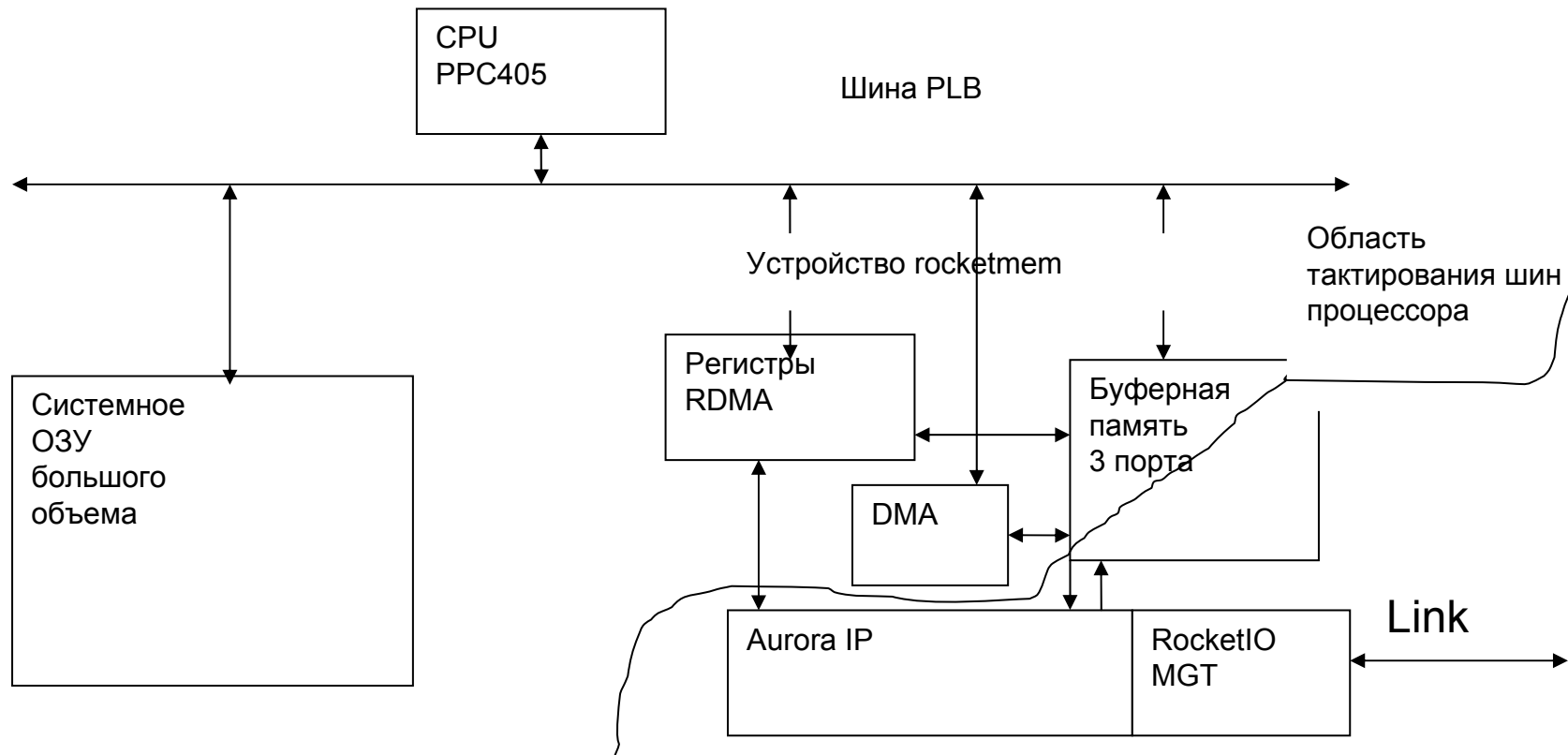
- Реализовано устройство для обмена данными «точка-точка»:



- Можно строить транспьютероподобные системы, поскольку в состав СнК может входить не одно устройство.
- С точки зрения программиста, в устройство входят:
 - - блок регистров управления,
 - - блок буферной памяти (64КБ),
 - - логика DMA.
- Обмены происходят между буферными памятьями устройств, соединенных линком.
- Режим обмена – RDMA Write.
- Сигнализация завершения обмена:
 - - на передающей стороне – бит готовности в регистре состояния,
 - - на приемной – прерывание.
- Логика DMA служит для быстрых обменов буферной памяти с основной памятью СнК.
- Для реализованного в настоящее время одноканального линка получены:
 - **скорость около 260 МБ/с и латентность 0.4мкс**

Как устройство включено в СнК.

- Устройство подключено к шине PLB. Блок регистров и блок буферной памяти адресуются на PLB. Логика DMA взята готовой из PLB IPIF. Скорость и латентность DMA примерно равны скорости и латентности линка.
- Объем необходимого оборудования: в 20-м кристалле минимальная тестовая система (PLB+OPB+OCM BRAM+UART) и два устройства rocketmem занимают 56% кристалла.



Какие были сложности в реализации.

- - несколько областей тактирования,
- - необходимость использования трехпортовой памяти.
- В обоих случаях применялись стандартные решения, рекомендованные Xilinx.
- Для управления самими RocketIO использовался предоставляемый в исходных текстах протокол Aurora от Xilinx. Это позволило сделать схему устройства не зависимой от числа каналов в линке.

Выводы.

- - в современных протоколах высокоскоростных сетей есть заметный резерв упрощения с целью повышения эффективности,
- - при использовании коммуникационных технологий RocketIO шинная инфраструктура СнК перестает быть адекватной.

Спасибо за внимание.