

ИНТЕЛЛЕКТУАЛЬНАЯ РАСПРЕДЕЛЕННАЯ СИСТЕМА ПАМЯТИ С ИЕРАРХИЧЕСКИМИ КОЛЬЦЕВЫМИ ШИНАМИ

*Институт кибернетики им. В.М. Глушкова НАН Украины, Киев, Украина

Анотація. Розглянуто архітектурно-структурну організацію інтелектуальної розподіленої системи пам'яті (ІРСП) з ієрархічними кільцевими шинами, застосування яких забезпечує розпаралелювання алгоритму рішення задачі користувача на двох і більше рівнях, що приводить у порівнянні з використанням однієї кільцевої шини до істотного підвищення продуктивності ІРСП в цілому. Приведені основні функції складових ІРСП вузлів і блоків і розглянуто функціонування системи в цілому. Дана попередня оцінка ефективності застосування ієрархічних кільцевих шин у складі ІРСП.

Ключові слова: інтелектуальна розподілена система пам'яті, ієрархічні кільцеві шини, підвищення продуктивності.

Аннотация. Рассмотрена архитектурно-структурная организация интеллектуальной распределенной системы памяти (ИРСП) с иерархическими кольцевыми шинами, применение которых обеспечивает распараллеливание алгоритма решения задачи пользователя на двух и более уровнях, что приводит по сравнению с использованием одной кольцевой шины к существенному повышению производительности ИРСП в целом. Приведены основные функции составляющих ИРСП узлов и блоков и рассмотрено функционирование системы в целом. Дана предварительная оценка эффективности применения иерархических кольцевых шин в составе ИРСП.

Ключевые слова: интеллектуальная распределенная система памяти, иерархические кольцевые шины, повышение производительности.

Abstract. The architectural and structural organization of intelligent distributed memory system (IDMS) with hierarchical ring-bus circuits, which application provides parallelization of the algorithm for solving user problems in two or more levels is considered. It results in essential increasing IDMS productivity as a whole in comparison with the use of one ring-bus circuit. The main functions of constituting IDMS units and blocks are resulted and system functioning as a whole is considered. Preliminary estimate of application effectiveness of hierarchical ring-bus circuits as a part of IDMS is given.

Keywords: intelligent distributed memory system, hierarchical ring-bus circuits, productivity increasing.

1. Введение

Рассмотренные системы с кольцевой шиной [1–6] имеют ограничение по быстродействию из-за временных задержек при прохождении сигналов до соответствующих модулей памяти при выполнении операций чтения или записи данных, а также из-за невозможности одновременного обращения к памяти за несколькими данными или для параллельной записи массива данных. При этом они также имеют ограничения наращивания емкости памяти и соответствующие технологические трудности, определяемые необходимостью размещения большого количества шин (сигнальных линий) внутри каждого модуля и системы памяти в целом.

Интеллектуальная система памяти, приведенная в [7], исключает некоторые из приведенных выше недостатков, однако имеет свои собственные, в том числе:

во-первых, в данной интеллектуальной системе памяти обеспечена возможность параллельной обработки информации секторными наборами только на одном уровне (например, отдельных фрагментов реализуемого пользователем алгоритма), что достигается при блокировании секторных разделителей от влияния процессов обработки в секторах слева и справа только одной кольцевой шины;

во-вторых, не использован один из основных принципов организации кольцевой шины, который предусматривает возможность уменьшить количество проводников кольцевой шины практически в два раза за счет передачи части данных по часовой стрелке, а другой части – против часовой стрелки до их состыковки при использовании соответствующей синхронизации. Это позволяет повысить помехозащищенность на линии передачи информации и тем самым повысить частоту передачи и соответственно производительность, а также уменьшить площадь, занимаемую кольцевой шиной на кристалле;

в-третьих, каждый разделитель кольцевой шины и блок управления разделителями являются достаточно сложными и содержат множество логических схем и схем памяти, соединенных большим количеством связей. Это приводит к значительным задержкам при прохождении информации от одного разделителя к другому и занимает достаточно большую площадь кристалла при размещении этих схем на одном кристалле.

Таким образом, требуется новый подход к созданию интеллектуальной системы памяти, которая исключала бы приведенные выше недостатки.

2. Актуальность

Ряд преимуществ кольцевых шин по сравнению со стандартной общей шиной, например, PCI, а также применение кольцевых шин в графических ускорителях, например, в сериях Radeon [1, 2] и в других вычислительных системах [3–6], доказывает их высокую актуальность и необходимость совершенствования с целью обеспечения высокой производительности распределенных вычислительных систем.

3. Цель работы

Создание архитектуры и структуры многоуровневой интеллектуальной распределенной памяти, содержащей иерархическую систему взаимосвязанных кольцевых шин в качестве коммутационной и передающей среды, обеспечивающей параллельную обработку информации не менее, чем на двух уровнях и отличающуюся простотой изготовления на кристалле (чипе). Такая система памяти может быть использована не только в качестве дополнительной внешней памяти к ЭВМ, но и в качестве ускорителя вычислительного процесса при решении с помощью ЭВМ сложных и трудоемких задач. При выполнении такой системы памяти на ПЛИС появляется возможность подстройки архитектуры и структуры системы под конкретную задачу, обеспечивая при этом дополнительный резерв повышения производительности системы в целом.

4. Основной материал

Схема интеллектуальной системы памяти с иерархическими кольцевыми шинами изображена на рис. 1 и содержит ПИП-1 – подсистему интеллектуальной памяти первого уровня; ПИП-2 – подсистему интеллектуальной памяти второго уровня, БУВ – блок управления взаимодействием интеллектуальных подсистем памяти различных уровней, СКП – системный контроллер памяти, БСлФ – блок служебных функций, КШ – кольцевую шину, КШД – кольцевую шину данных, КШУ – кольцевую шину управления, БУ(КШ) – блок управления кольцевой шиной, ЛШ – локальную шину хост-машины, ВП – ведущий процессор с КЭШ-памятью, Св/выв(Д) – системный ввод/вывод данных, Св/выв(У) – системный ввод/вывод управляющих сигналов, УСР – управляемые секторные разделители, СН – секторные наборы, УУ(УСР) – узел управления секторными разделителями, БУ(У) – блок управления уровнями, 1И(ПД) – первый интерфейс для передачи данных между уровнями, 2И(УС) – второй интерфейс для передачи управляющих сигналов между уровнями, Ин(вых) – интерфейс выхода данных и управляющих сигналов для других уровней, Уз_1(Сх) – первый узел синхронизации, Уз_2(Сх) – второй узел синхронизации.

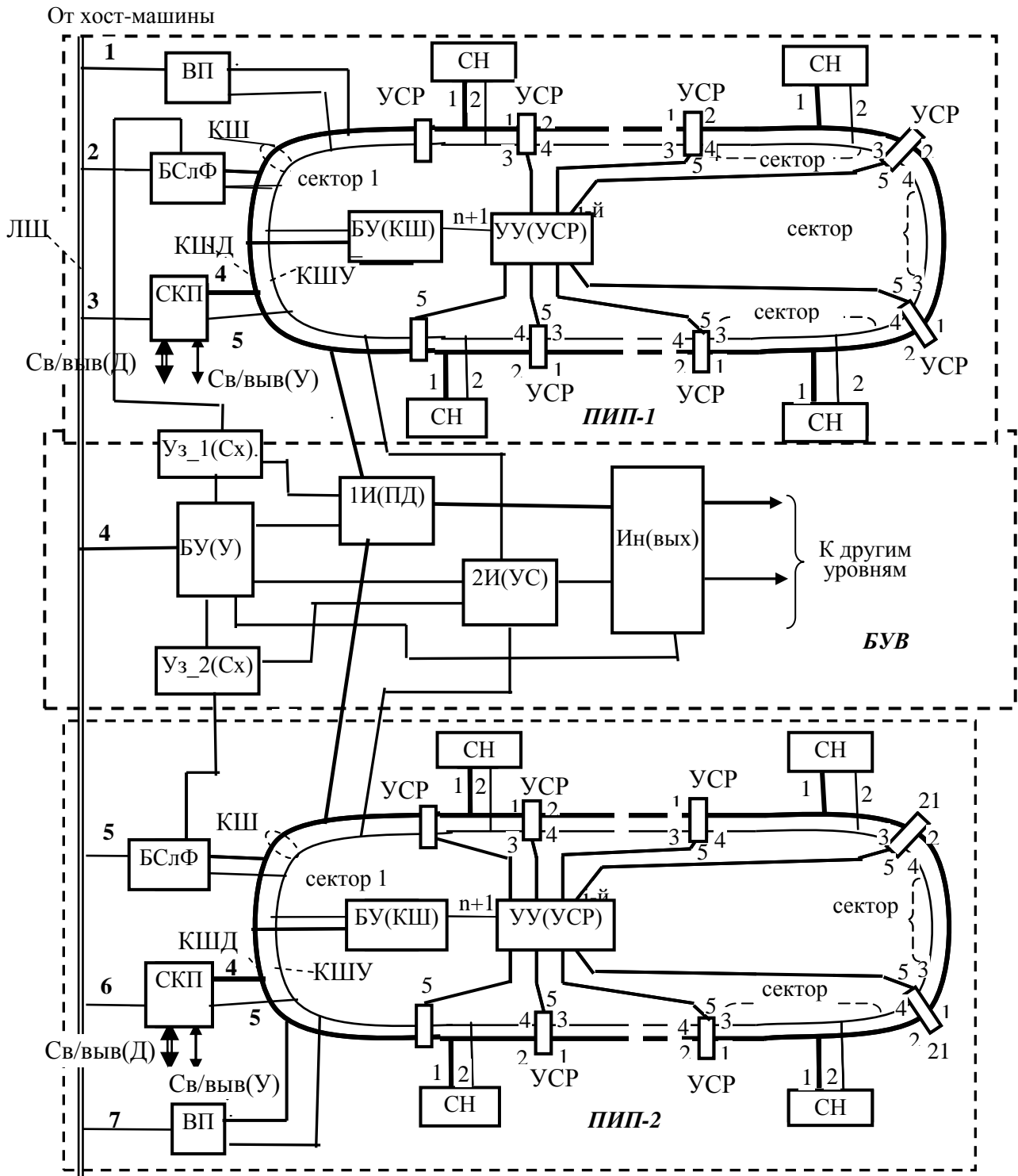


Рис. 1. Структурная схема интеллектуальной распределенной системы памяти с иерархическими кольцевыми шинами

Далее более подробно изложим организацию структуры и основные положения реализации функций наиболее важных узлов и блоков системы.

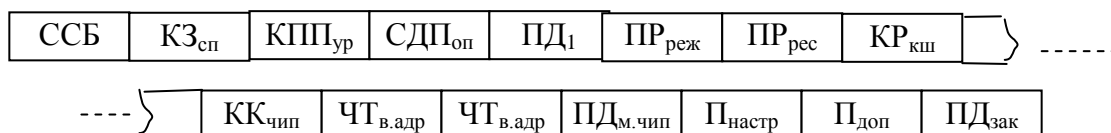
Системный контроллер памяти (СКП) управляет работой всей интеллектуальной распределенной системой памяти, обеспечивая её применение в качестве:

– основной или дополнительной памяти любой компьютерной системы с реализацией классических функций памяти – запись, хранение и чтение информации;

– средства обработки типа “Процессор-в-памяти” (Processor-in-memory”) вместе с имеющимися внутри системы устройствами памяти с реализацией функций параллельной обработки секторными наборами блоков интеллектуальной памяти и блоками интеллектуальной памяти внутри каждого секторного набора.

В соответствии с функциями, которые определены, а также методами использования интеллектуальной распределенной системы памяти (ИРСП) и режимами её работы, системный контроллер памяти содержит микропрограммный автомат (или микропроцессор) с локальной памятью, набор широких регистров для временного хранения и анализа управляющего пакета, схему для анализа поступающего кода адреса для определения “своего” обращения или обращения за пределы памяти ИРСП, а также набор интерфейсных схем для подключения системного контроллера памяти.

В основу функционирования системного контроллера положен системный пакет, который поступает от хост-машины на соответствующие его входы/выходы. При этом системный контроллер выбирает из состава пакета нужные коды полей, расшифровывает их и вырабатывает необходимые сигналы управления, которые он выдает вместе с данными, поступающими от хост-машины на кольцевую шину данных и кольцевую шину управления. Пример структуры пакета, содержащего необходимые коды полей, имеет следующий вид:



Назначение и функции каждого поля пакета приведены в табл. 1.

Таблица 1. Назначение и функции полей управляющего пакета

№	Поля пакета	Функциональное назначение поля пакета
1	2	3
1	ССБ	Поле системного и других типов сигнала сброса триггеров
2	КЗ _{сп}	Код поля запуска сервисных программ для выполнения функций: распределения памяти и размещения данных, разбиения задачи на параллельные фрагменты, формирования нового управляющего пакета, запуска библиотеки стандартных подпрограмм, загрузки данных и др.
3	КПП _{ур}	Код поля, определяющий передачи информации между первым и вторым уровнями и обратно
4	СДП _{оп}	Спецификатор действия пакета, определяющий типы логических и арифметических операций, которые должны быть выполнены объектом-адресатом после получения пакета
5	ПД ₁	Поле значения параметров (данных), которые могут использоваться при выполнении текущего действия или могут быть частью процесса, который переносит значения в следующие местоположения
6	ПР _{реж}	Поле признака, отражающего характер и последовательность действий в различных режимах работы системы памяти, в том числе в качестве обычной памяти, а также в качестве “Процессора-в-памяти”
7	ПР _{рес}	Поле признака использования ресурсов обработки информации собственной системы или дополнительных ресурсов за счет других систем, подключенных через соответствующие интерфейсы

8	КР _{кш}	Код для управления режимами работы кольцевых шин (выделение секторных наборов БИП и соответственно секторов кольцевых шин для параллельной обработки, широковещательной передачи по кольцевым шинам данных или команд, передачи информации по кольцевым шинам по часовой стрелке и против часовой стрелки и др.)
9	КК _{чип}	Код для управления коммутацией внутри чипа (блока интеллектуальной памяти)
10	ЧТ _{в.адр}	Поле адреса для чтения отдаленного, виртуального адресного обращения за пределы системы
11	ПД _{м.чип}	Поле адреса для передачи блоков данных между чипами различных секторных наборов БИП
12	П _{настр}	Код поля для предварительной и в процессе работы настройки системы
13	П _{доп}	Дополнительные (вспомогательные) поля, необходимые для надежной транспортировки, обнаружения ошибок, маршрутизации и управления контекстом
14	ПД _{зак}	Заключительный элемент пакета – поле, определяющее последующие действия после окончания выполнения данного пакета, например, могут быть созданы один или более дочерних пакетов

Пакет может изменяться по длине, обеспечивая, таким образом, эффективную обработку простых операций с маленькими пакетами и эффективным использованием полосы пропускания для перемещения больших блоков данных.

Интерфейсы пакета (последовательно-параллельные) представлены каналами для активного управляемого сообщением вычисления между блоками интеллектуальной памяти (БИП) по кольцевым шинам.

Пакеты также могут использоваться для выполнения операций типа отдаленной загрузки или хранения, а также для вызова методов обработки на другой аналогичный БИП системы памяти, перемещая тем самым с помощью пакетов методы обработки ближе к данным за меньшее время, чем данные (как обычно) передаются к ним. Пакеты позволяют осуществлять разбиение или расщепление операций, обеспечивая тем самым допустимое время ожидания для всей системы. Применительно к предлагаемой системе интеллектуальной памяти (рис. 1) управляющий пакет, помимо указанных выше полей, содержит поля и признаки, отражающие характер и последовательность действий, обеспечивающих настройку ресурсов системы (П_{настр}) перед запуском её на решение конкретной задачи, а также при необходимости перестройку ресурсов в процессе работы.

Следует отметить, что состав полей управляющего пакета и разрядность каждого поля определяются типом и функциональными возможностями интеллектуальной распределенной системы памяти, а также особенностями решаемых с помощью ИРСП задач. При этом длина (количество разрядов) управляющего пакета может быть равна длине строк, хранимых в БИП данных (например, 256 бит), и поэтому пакеты могут быть сохранены в строке широких регистров и обработаны непосредственно арифметико-логическом устройством секторного набора.

Секторный набор (СН) предназначен для записи, чтения и хранения данных, поступающих через вход/выход блока интеллектуальной памяти от/в соответствующего сектора кольцевой шины при наличии управляющих сигналов, поступающих через вход/выход БИП от шины управления идентичного по номеру сектора кольцевой шины согласно режиму работы классической памяти, а также режиму работы типа “Процессор-в-памяти”. При этом каждый секторный набор (рис. 1) содержит $N / (n - 1)$ блоков интеллектуальной памяти, где n – количество секторов, а N – общее количество БИП. Блок ин-

теллектуальной памяти, входящий в состав каждого секторного набора (СН), предназначен для записи и хранения двух частей данных, поступающих через интерфейс (рис. 2) и мультиплексор по часовой и против часовой стрелки с регистров А и В на соответствующие входы узла памяти, реализуя тем самым режим 1 под управлением контроллера сектора.

Результат операции поступает либо в узел памяти сектора, либо непосредственно на шину данных сектора двумя частями с разделением по времени совместно с соответствующими сигналами управления, которые поступают на соответствующие входы/выходы процессора с шины управляющих сигналов сектора. При этом с выхода узла памяти считанные через демультимплексор части кодов данных поступают на регистры РгС и РгD, с выходов которых эти части кодов через интерфейс сектора поступают на шину данных сектора для последующих передач по кольцевой шине данных по часовой и против часовой стрелки.



Рис. 2. Структурная схема секторного набора первого и второго уровней

В режиме 2 для обработки информации по месту её хранения, согласно принципам организации РИМ-системы, процессорный элемент по команде секторного контроллера СКП выполняет операции над полноразрядным кодом. Основные принципы построения и работы РИМ-систем отражены в публикации [8]. Результат операции поступает либо в узел памяти, либо непосредственно на шину данных сектора двумя частями с разделением по времени совместно с соответствующими сигналами управления, которые поступают на соответствующие входы/выходы процессора с шины управляющих сигналов сектора. При этом с соответствующего выхода узла памяти сектора считанные через демультимплексор части кодов данных поступают на регистры РгС и РгD, с выходов которых эти части кодов через интерфейс сектора поступают на шину данных сектора для последующих передач по кольцевой шине данных по часовой и против часовой стрелки.

Блоки служебных функций (БСлФ) системы памяти предназначены для распределения ресурсов интеллектуальной системы памяти в целом, включая распределение всего адресного пространства памяти по суммарной памяти ПИП-1 и ПИП-2 соответствующего размещения данных, разделения программ пользователя по секторным наборам на параллельно выполняемые части и загрузки средств обработки секторных наборов.

При этом блоки служебных функций, входящие в состав ПИП-1 и ПИП-2, могут работать в двух основных режимах: режим 1 и режим 2. В режиме 1 этот блок осуществляет исходную настройку системных ресурсов для оптимального решения пользовательской задачи. В режиме 2 он осуществляет непосредственно распределение ресурсов для работы, а также при необходимости перераспределение ресурсов в процессе работы системы. При этом алгоритмы реализации перечисленных выше функций распределения ресурсов с помощью блока служебных функций в общем случае основываются на известных способах, например, способах распределения памяти РІМ-систем, которые приведены в публикациях [9–11].

Блоки управления кольцевой шиной (БУ(КШ)) для ПИП-1 и для ПИП-2 предназначены для управления соответствующими кольцевыми шинами при передачах информации, поступающей на кольцевую шину из блоков служебных функций, системных контроллеров памяти в БИП при выполнении операции записи, либо на кольцевую шину из БИП любого секторного набора при выполнении операции чтения и передачи полученных данных на первый сектор кольцевой шины и далее на соответствующие входы/выходы системного контроллера (СКП) и системного блока служебных функций (БСлФ). При этом каждый блок БУ(КШ) выполняет следующие функции:

- фиксация (например, в виде соответствующих индексов или табличного описания) компонентов и связей кольцевой шины с указанием количества секторных разделителей и секторов с выделением первого сектора кольцевой шины для связи с системным контроллером памяти и системным блоком служебных функций, а остальных секторов – с соответствующими секторными наборами БИП;

- управление широкоэвентальным режимом передачи, когда все секторные разделители УСР открыты для передачи по кольцевой шине информации от каждого предыдущего сектора к каждому последующему, и наоборот;

- управление адресным режимом передачи информации, когда открыты секторные разделители только по пути передачи информации либо из одного секторного набора БИП в другой, либо из секторного набора БИП в первый сектор кольцевой шины и далее на соответствующие входы блоков;

- управление режимом секторной блокировки, когда один или несколько секторов кольцевой шины в соответствии с адресной зоной секторных наборов БИП блокируются секторными разделителями от других секторов шины, предоставляя тем самым возможность выделенным (заблокированным) секторным наборам БИП параллельно выполнять приписанные им участки алгоритмов;

- вырабатывает сигнал опорной синхронизации, который поступает с соответствующего выхода БСлФ ПИП-1 на второй вход узла синхронизации Уз_1(Сх) блока БУВ, а также со второго выхода Уз_2(Сх) уровня ПИП-2 на соответствующий вход узла синхронизации Уз_2(Сх) при передаче информации из ПИП-1 на ПИП-2 и обратно.

При реализации указанных функций блок управления кольцевой шиной ПИП-1 и ПИП-2 использует следующие поля управляющего пакета, выделенные системными контроллерами памяти СКП, переданные на его вторые входы/выходы через первый сектор кольцевой шины управления: $KP_{кш}$, $PR_{реж}$, $KZ_{сп}$, $KK_{м/чип}$, $PD_{м.чип}$, $P_{доп}$ и их комбинации.

Таким образом, исходя из набора функций, подлежащих реализации, блок управления кольцевой шиной содержит следующие компоненты, соединенные соответствующими связями:

- таблицу параметров кольцевой шины, идентифицирующих структуру кольцевой шины с указанием разрядности шины данных, шины управления, количества секторов, на которые разделена кольцевая шина, количества секторных разделителей, количества секторных наборов БИП и количества БИП в каждом секторе с указанием адресного пространства емкости памяти каждого БИП и сектора в целом;

– набор регистров КЕШ-памяти для хранения кодов полей $KP_{кш}$, $PR_{реж}$, $KЗ_{сп}$, $KK_{м/чип}$, $ПД_{м.чип}$, $П_{доп}$ и их комбинаций;

– блок анализа и обработки кодов полей регистров КЕШ-памяти, а также входных управляющих сигналов, поступающих на вторые входы/выходы системного блока управления с кольцевой шины управления первого сектора;

– блок формирования кодов управляющих сигналов по кодам адреса и обработанным кодам полей управляющего пакета для реализации перечисленных выше режимов работы кольцевой шины. При этом коды управляющих сигналов поступают с третьего входа/выхода системного блока управления кольцевой шиной БУ(КШ) на $(n + 1)$ -ые входы/выходы узла управления УУ(УСР) секторными разделителями УСР, а коды адреса – с первого сектора кольцевой шины данных КШ на первые входы/выходы системного блока управления кольцевой шиной БУ(КШ);

– интерфейсные схемы для связи системного блока управления БУ(КШ) с кольцевой шиной данных КШД и кольцевой шиной управления КШУ первого сектора кольцевой шины.

Узлы управления секторными разделителями УУ(УСР) (рис. 1) предназначены для передачи данных и управляющих сигналов предыдущего сектора кольцевой шины через контакты 1 и 3 секторного разделителя на последующий сектор с входов/выходов 2 и 4 секторного разделителя и в обратном направлении в зависимости от двоичного кода управляющего сигнала (01 или 10) на пятом входе разделителя, поступающего с соответствующего выхода узла управления секторными разделителями. При значении кода управляющего сигнала 00 на пятом входе секторного разделителя передача информации по кольцевой шине через секторный разделитель УСР запрещена, тем самым имеется возможность блокировки выбранного сектора кольцевой шины КШ с двух сторон, обеспечивая при этом обработку данных соответствующим секторным набором СН независимо от других секторных наборов.

Сформированные в определенной последовательности сигналы управления поступают на пятые входы соответствующих разделителей УСР, обеспечивая при этом необходимый набор сигналов для реализации конкретных режимов обработки и передачи кодов адреса, данных и управляющих сигналов. Примеры отдельных типов управляющих сигналов приведены в табл. 2. При этом, помимо управляющих сигналов, вырабатываемых контроллерами интеллектуальной распределенной системы памяти, используются следующие основные системные сигналы (табл. 2).

Ведущий процессор ВП, входящий в состав ПИП-1 с памятью, а также в состав ПИП-2 с памятью, предназначен для хранения и предварительного анализа данных, находящихся в первом секторе кольцевых шин ПИП-1 и ПИП-2, а также для хранения методов распределения адресного пространства, распределения данных по соответствующим адресам и метода распределения приложений по процессорам, которые при поступлении соответствующих команд пересылают в соответствующие блоки служебных функций ПИП-1 и ПИП-2 для реализации этих функций. Кроме того, эти процессоры используются при необходимости чтения длинных слов (например, 256 бит), считанных из памяти различных секторных наборов. При требованиях к большой емкости памяти для этих процессоров они могут быть вынесены за пределы чипа в виде самостоятельного изделия.

Таблица 2. Основные системные управляющие сигналы, используемые при работе ИРСП

№ сигнала	Функциональное назначение сигнала	Откуда поступает сигнал
1	2	3
1	Запуск ведущего процессора ВП в составе ПИП-1	От хост-машины
2	Запуск блока служебных функций БСлФ в составе ПИП-1	От хост-машины
3	Запуск системного контроллера СКП в составе ПИП-1	От хост-машины
4	Запуск блока управления взаимодействием (БУВ) подсистем памяти ПИП-1 и ПИП-2	От хост-машины
5	Запуск блока служебных функций БСлФ в составе ПИП-2	От хост-машины
6	Запуск системного контроллера СКП в составе ПИП-2	От хост-машины
7	Запуск ведущего процессора ВП в составе ПИП-2	От хост-машины
8	Разрешение чтение / запись для доп. строки триггеров разделителя	От узлов УУ(УСР)
9	Входной/выходной сигнал данных на первом вводе/выводе системного контроллера в составе ПИП-1	От хост-машины
10	Входной/выходной управляющий сигнал на втором вводе/выводе системного контроллера в составе ПИП-1	От хост-машины
11	Сигнал синхронизации разделителей УСР шин ПИП-1 и ПИП-2	От узлов УУ(УСР)
12	Передача информации по кольцевой шине против часовой стрелки (налево)	От узлов УУ(УСР)
13	Команда "Чтение"	От узлов УУ(УСР)
14	Разрешение / блокировка передачи информации по кольцевой шине	От узлов УУ(УСР)
15	Передача информации по кольцевой шине по часовой стрелке (направо)	От узлов УУ(УСР)
16	Сигнал сброса Тг ₃ столбца триггеров разделителя ПИП-1 и ПИП-2	От узлов УУ(УСР)
17	Сигнал сброса Тг строки дополнительной строки триггеров	От узлов УУ(УСР)
18	Сигнал разрешения на запись управляющих сигналов в столбец триггеров разделителя	От узлов УУ(УСР)
19	Сигнал разрешения на запись данных в столбец триггеров разделителя	От узлов УУ(УСР)
20	Сигнал запуска секторных процессоров ПИП-1 и ПИП-2	От узлов УУ(УСР)
21	Входной/выходной сигнал данных на втором вводе/выводе системного контроллера СКП в составе ПИП-2	От хост-машины
22	Входной/выходной управляющий сигнал на первом вводе/выводе системного контроллера СКП в составе ПИП-2	От хост-машины

Секторный разделитель для ПИП-1 и ПИП-2. Структурная схема управляемого секторного разделителя приведена на рис. 3. Она представлена в виде матрицы элементов схем, содержащей первый столбец элементов "2И", второй столбец элементов "3И", третий столбец – схем триггеров "Тг₃", четвертый столбец элементов "3И", пятый столбец элементов "2И", а также дополнительную строку схем триггеров с элементами "2И".

Интеллектуальная распределенная система памяти с иерархическими кольцевыми шинами работает следующим образом. При этом она может работать в четырех режимах:

Режим 1. Это режим параллельной реализации программы пользователя одновременно на двух уровнях с использованием ПИП-1 и ПИП-2. Далее несколько подробнее рассмотрим процессы, протекающие при реализации режима 1. Так как остальные режимы

являются в основном частными случаями первого режима, то ограничимся только их краткими комментариями.

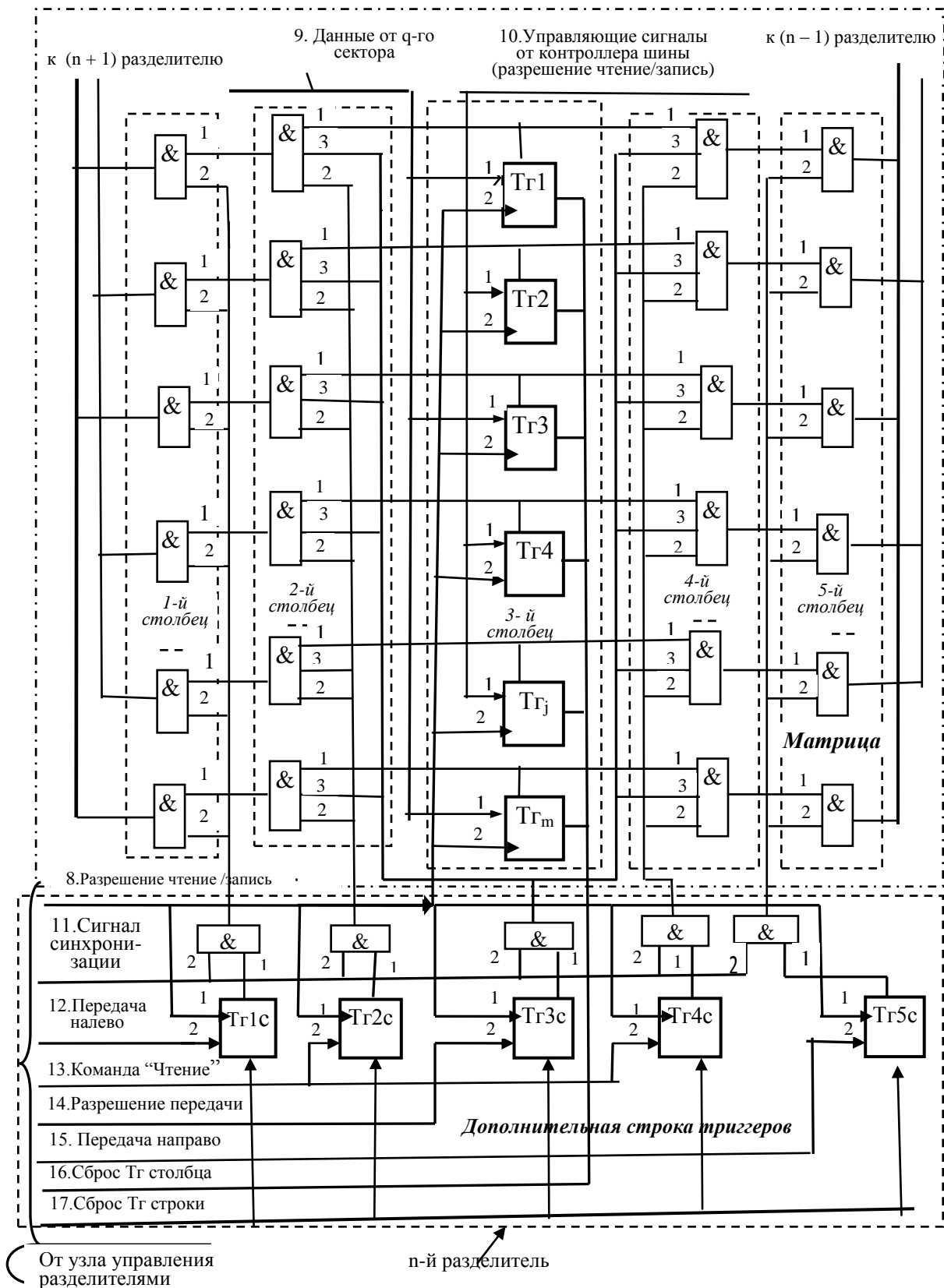


Рис. 3. Структурная схема управляемого секторного разделителя

Режим 2. Первый этап заведомо реализован без участия блоков служебных функций ПИП-1 и ПИП-2, например, хост-машиной программным способом, и все атрибуты реализации первого этапа размещены в памяти соответствующих секторных наборов обоих уровней. При этом для достижения баланса загрузки процессоров возможны пересылки информации через блок управления уровнями БУВ не только от ПИП-1 к ПИП-2, но и в обратном направлении – от ПИП-2 к ПИП-1. Выдача результатов параллельной обработки информации аналогична выдаче соответствующих результатов в режиме 1, который описан выше.

Режим 3. Одновременная реализация двух независимых программ пользователя, одна из которых реализуется на ПИП-1, а вторая одновременно и независимо от первой на ПИП-2. При этом исключаются функции БУВ, процесс выдачи результатов аналогичен процессу выдачи в режиме 1.

Режим 4. Применение интеллектуальной распределенной системы памяти с кольцевыми шинами или в виде дополнительной памяти хост-машины, или в виде дополнительной распределенной памяти процессорного элемента ВП, принадлежащего ПИП-1, или в виде дополнительной распределенной памяти процессорного элемента ВП, принадлежащего ПИП-2. В этом режиме возможно также выполнение команд чтения и записи слов большой разрядности, например, 256 бит и более с последующим их выводом по частям, например, по 8 или 16 бит.

Описание работы ИРСИ в режиме 1. На этапе подготовки с помощью блока служебных функций происходит распределение адресного пространства для всей системы по двум уровням, размещение данных в область адресного пространства ПИП-1, а также размещение подпрограмм (или крупных фрагментов программы) по памяти секторных наборов с учетом баланса загрузки процессорных элементов каждого сектора ПИП-1. Далее каждую подпрограмму с первого сектора кольцевой шины ПИП-1 передают через блок управления уровнями БУВ в определенной последовательности, зафиксированной в общем алгоритме реализации программы, на первый сектор кольцевой шины ПИП-2, где с помощью блока служебных функций происходит разделение адресного пространства по секторам, размещение данных и распределение фрагментов каждой подпрограммы между памятью секторных наборов ПИП-2 с учетом баланса загрузки процессорных элементов каждого сектора ПИП-2.

На втором этапе первого режима работы происходит непосредственная реализация программы пользователя с применением параллелизма на двух уровнях ПИП-1 и ПИП-2. При этом результаты обработки поступают через первый сектор кольцевой шины ПИП-1 и далее через соответствующие входы/выходы системного контроллера, входящего в состав ПИП-1, а также через первый сектор кольцевой шины ПИП-2 и далее через входы/выходы системного контроллера СКП.

В исходном состоянии при выполнении режима 1 системный пакет, сформированный хост-машиной, данные, соответствующие системные наборы управляющих сигналов, а также программы, реализующие методы распределения адресного пространства, размещения данных и распределения подпрограмм (крупных фрагментов программ), размещены либо в памяти хост-машины, либо в памяти (при наличии достаточной емкости памяти) ведущих процессоров ВП, расположенных в ПИП-1 и ПИП-2. Кроме того, все схемы триггеров ИРСИ должны находиться в сброшенном (нулевом) состоянии. Предпочтительным является второй вариант, так как интеллектуальная распределенная система памяти в целом может использоваться как приставка к ЭВМ, расширяющая её функциональные возможности и повышающая производительность при решении задач пользователя.

Этап подготовки ИРСИ в режиме-1 начинается с подачи с 1-го, 2-го и 3-го выходов локальной шины хост-машины сигналов запуска на соответствующие входы ведущего процессора ВП, блока служебных функций БСлФ и системного контроллера СКП, а также

одновременно системного пакета от хост-машины на входы/выходы СКП системного контроллера ПИП-1 с полным набором кодов его полей, который запоминается в “широких” регистрах системного контроллера СКП (табл.1), а также данных на вторые входы/выходы СКП. Системный контроллер выделяет из системного пакета коды следующих полей КЗ_{СП}, ПД₁, ПР_{реж.}, ПР_{рес.}, КР_{кш.}, расшифровывает их и вырабатывает соответствующие управляющие сигналы, которые через кольцевую шину управляющих сигналов КШУ поступают на соответствующие входы/выходы блока служебных функций БСлФ. В качестве управляющих сигналов системный контроллер формирует сигналы для выполнения следующих функций: КЗ_{СП}=(разделение адресного пространства памяти, размещение данных, размещение подпрограмм по секторным наборам); ПД₁=(значения данных); ПР_{реж.}=(режим: “Процессор-в-памяти”); ПР_{рес.}=(используются только внутренние ресурсы ИРСР); КР_{кш.}=передача по кольцевым шинам данных и команд. Эти сигналы с кольцевой шины управления КШУ первого сектора поступают на соответствующие входы блока управления кольцевыми шинами БУ(КШ), который анализирует их и вырабатывает в определенной последовательности набор управляющих сигналов, которые поступают на входы 5 секторных разделителей, обеспечивая тем самым прохождение информации к каждому сектору от блока БСлФ на кольцевые шины ПИП-1 после реализации своих служебных функций, и при необходимости происходит её запоминание в секторной памяти.

Таким образом, по окончании этапа подготовки в режиме 1 должны быть представлены:

- разделенные по секторным наборам ПИП-1 подпрограммы (крупные фрагменты программы пользователя) для их параллельной реализации (первый уровень);
- разделенные по секторным наборам ПИП-2 фрагменты распараллеливаемых подпрограмм (крупных фрагментов программ) для их параллельной реализации (второй уровень);
- способы и средства организации параллельного вычислительного процесса с использованием иерархических кольцевых шин;
- способы представления и выдачи полученных результатов при распараллеливании вычислительного процесса на двух иерархических уровнях.

На втором этапе режима 1 подаются сигналы запуска для каждого секторного процессора ПИП-1 и ПИП-2, который поступает на его входы/выходы непосредственно с кольцевой шины управления конкретного сектора (фиг. 2, рис. 1), и каждый секторный процессор СПр реализует назначенные ему функции блоками служебных функций БСлФ. При этом секторные разделители УСР уровня ПИП-1 и УСР уровня ПИП-2 должны быть заблокированы от передачи информации от соседних секторов, например, с помощью сигналов сброса дополнительной строки триггеров (фиг.1, фиг.3, рис. 1). Полученные результаты каждого секторного процессора запоминаются в соответствующей секторной памяти (рис. 2), с первых выходов которой этот результат при чтении данных проходит через соответствующий демультимплексор на регистры PгC и PгD и с них на четвертые и пятые входы интерфейса секторного набора (ИСН) под управлением контроллера секторного набора, обеспечивающего выдачу с секторного интерфейса разделенной на две части информации на кольцевые шины сектора в определенные кванты времени. При этом полные результаты со всех секторов также передаются по кольцевым шинам ПИП-1 и кольцевым шинам ПИП-2 в определенные кванты времени как по часовой, так и против часовой стрелки.

Для этого системные контроллеры СКП (рис. 1) анализируют коды полей системного пакета СДП_{оп.}, ПР_{реж.}, ПР_{рес.}, КР_{кш.}, выделяя при этом следующие сигналы: “Типы реализуемых логических и арифметических операций, которые должны быть выполнены при получении пакета”, “Применение режима типа процессор-в-памяти”, “Используются только собственные ресурсы системы”, “Применение кольцевых шин для параллельной обработки”. Эти сигналы поступают через шины управляющих сигналов первых секторов

ПИП-1 и ПИП-2 на вторые входы блоков управления кольцевыми шинами БУ(КШ), которые формируют управляющие сигналы для секторных разделителей УСР. Функциональные назначения наиболее важных сигналов отмечены в табл. 2 (позиции 14 и 24), а также показаны на рис. 3 (сигнал разрешение чтения/записи; сигнал синхронизации; сигнал команды чтения; разрешения передачи; передача данных по часовой стрелке; передача данных против часовой стрелки; сброс Тг столбца; сброс Тгс дополнительной строки триггеров. Для реализации процедуры выдачи результирующих данных прежде всего в соответствующий разделитель (например, УСР), в третий столбец триггеров записывают данные и управляющие сигналы при наличии соответствующего сигнала “Разрешение чтения/записи”, который поступает от узла управления секторными разделителями УУ(УСР), сформированными при наличии на их входах соответствующих сигналов от узла УУ(УСР) для ПИП-1 и узла УУ(УСР) для ПИП-2. Выходные сигналы данных и выходные управляющие сигналы с соответствующих триггеров третьего столбца поступают на первые входы схем “ЗИ” второго и четвертого столбцов (рис. 3), на вторые входы которых от узлов УУ(УСР) поступает сигнал “Разрешение передачи”, а на третьи входы – сигнал команды “Чтение”. При этом передача данных и управляющих сигналов по кольцевым шинам против часовой стрелки происходит при наличии на вторых входах схем “ЗИ” первого столбца соответствующего управляющего сигнала (табл. 2, позиция 12), а передача по часовой стрелке – при наличии на вторых входах схем “ЗИ” пятого столбца (рис. 3) соответствующего сигнала, отмеченного в табл.2 на позиции 15. Для синхронизации этих сигналов при их состыковке используется сигнал синхронизации (табл. 2, позиция 11), который поступает на вторые входы схем “ЗИ” дополнительной строки триггеров. Далее результаты обработки сигналов данных и управления через первые секторы кольцевых шин ПИП-1 и ПИП-2 поступают в хост-машину через соответствующие входы/выходы четвертого и пятого системных контроллеров СКП, а также на входы/выходы Св/выв(У), Св/выв(Д) для ПИП-1 и соответственно на такие же входы/выходы для ПИП-2.

4. Оценка эффективности применения иерархических кольцевых шин

Под оценкой эффективности будем понимать отношение времени реализации распараллеленных фрагментов алгоритма с помощью ИРСРП при наличии иерархических кольцевых шин к времени реализации ИРСРП при наличии одной кольцевой шины.

При этом для предварительной оценки целесообразно принять следующие положения, которые очевидны при решении большого количества задач:

- ПИП-1 и ПИП-2 идентичны по составу узлов и блоков и их функциональному назначению и содержат n секторов, при этом активных секторных наборов, исключая первый сектор, будет $(n - 1)$;

- все $(n - 1)$ секторных наборов ПИП-1 одновременно (параллельно) реализуют m подпрограмм (крупных фрагментов алгоритма) задачи пользователя, выделенные блоком служебных функций (БСлФ), принадлежащем ПИП-1;

- все $(n - 1)$ секторных наборов ПИП-2 одновременно (параллельно) реализуют b фрагментов алгоритма задачи пользователя, выделенные блоком служебных функций (БСлФ), принадлежащем ПИП-2, при распараллеливании крупного фрагмента алгоритма, реализуемого секторным набором ПИП-1;

- разрядность обрабатываемых слов условно принимаем равной m битам, при этом примерно половина $0,5m$ передается по кольцевой шине по часовой стрелке, а вторая половина соответственно одновременно передается против часовой стрелки, что приводит к уменьшению времени передачи информации примерно в два раза.

Так как все процессы для ПИП-1 распараллелены, то каждый секторный набор заканчивает обработку информации примерно в одно и то же время. При этом целесообразно

в дальнейшем проводить оценку максимального времени реализации одной параллельной подпрограммы и соответственно оценку эффективности применительно к одному секторному набору ПИП-1. Аналогично также и для ПИП-2, где целесообразно проводить оценку максимального времени реализации одной параллельной ветви, но уже фрагмента подпрограммы.

Принимая во внимание приведенные выше положения, можно предварительно оценить значение коэффициента эффективности Ω применения иерархических кольцевых шин в составе интеллектуальной системы памяти, что является предметом рассмотрения отдельной статьи.

5. Выводы

Таким образом, распределенная интеллектуальная система памяти при работе системы обеспечивает широкие функциональные возможности, включая уменьшение проводников кольцевой шины практически в два раза и передачу информации по частям по этим шинам по часовой и против часовой стрелки, мультирежим уплотнения при передаче информации по кольцевой шине в определенные кванты времени, возможность параллельной обработки подпрограмм и фрагментов подпрограмм на двух и более уровнях, что привело к повышению помехоустойчивости системы и, следовательно, к повышению частоты её работы, а также в целом – к повышению производительности системы за счет реализации двухуровневого параллелизма при выполнении программы пользователя. Кроме того, уменьшение количества шин и упрощение схемотехнических решений узлов и блоков обеспечивает возможности создания высокопроизводительной системы на одном (либо на двух) кристалле [12].

Приведены основные функции составляющих ИРСП узлов и блоков и рассмотрено функционирование системы в целом. Предложенная архитектурно-структурная организация ИРСП имеет ряд преимуществ по сравнению с известными системами памяти, обеспечивая в режиме “Processor-in-Memory” (PIM-системы) существенное повышение производительности. При этом узлы и блоки такой системы отличаются простотой и технологичностью при их изготовлении на кристалле, например, на ПЛИС. Наличие 4-х режимов работы ИРСП обеспечивает эффективное её использование не только в качестве дополнительной памяти к различным вычислительным системам, но и в качестве интеллектуального средства обработки информации при решении ряда задач в области графики, систем алгебраических и дифференциальных уравнений, распознавания образов и других задач, связанных с искусственным интеллектом. Таким образом, применение иерархических кольцевых шин в составе ИРСП показывает высокую актуальность проведения исследований в этом направлении.

СПИСОК ЛИТЕРАТУРЫ

1. Пахомов С. Современные видеокарты для ПК [Электронный ресурс] / С. Пахомов // КомпьютерПресс. – 2005. – № 12. – Режим доступа: <http://www.compress.ru/article.aspx?id=15082&iid=713>.
2. Чеканов Д. Radeon X1300, Radeon X1600 и Radeon X1800: тест [Электронный ресурс] / Чеканов Д. – Режим доступа: <http://www.thg.ru/graphic/20051006/index.html>.
3. Pat. App. Pub. 2006031593 (A1) Ring bus structure and its use in flash memory systems / Sinclair A.W.; US, G06F15/16; 2006. – 38 p.
4. Pat. 6,728,206. Crossbar switch with communication ring bus / Carlson, J.R.; US, 370/222, 370/423; EU, H04L 12/56; H04L 12/427; H04L; 2004. – 19 p.
5. Pat. 6,253,292. Distributed shared memory multiprocessor system based on a unidirectional ring bus using a snooping scheme / Jhang S.T.J., Chu S.K., Hyung H.; US, 711/146; 709/218; 711/148; 711/E12.025, EU, G06F 12/08; G06F 012/00; 2001. – 22 p.

6. Pat. 6,658,509. Multi-tier point-to-point ring memory interface / Bonella; Randy M., Halbert; John B.; US, 710/100; 370/223; 710/300, EU G06F 13/42; G06F 013/00; G01R 031/08; 2003. – 19 p.
7. Пат. України № 99164. Інтелектуальна розподілена система пам'яті з кільцевою шиною / Палагін О.В., Яковлев Ю.С., Тихонов Б.М., Єлісеєва О.В. Клас UA: C2, Міжнародний клас: G06F 15/16, G06F 13/42.
8. Палагін А.В. Основные принципы построения вычислительных систем с архитектурой “Процессор-в-памяти” / А.В. Палагін, Ю.С. Яковлев, Б.М. Тихонов // Управляющие системы и машины. – 2004. – № 5. – С. 30 – 37.
9. Яковлев Ю.С. О распределении памяти в компьютерных системах / Ю.С. Яковлев, Б.М. Тихонов // Управляющие системы и машины. – 2006. – № 5. – С. 40 – 47.
10. Яковлев Ю.С. Об оптимизации размещения данных в РІМ-системе / Ю.С. Яковлев, Б.М. Тихонов // Математичні машини і системи. – 2006. – № 3. – С. 24 – 35.
11. Яковлев Ю.С. Математическая модель и стратегия распределения приложений для интеллектуальной памяти распределенных компьютерных систем / Ю.С. Яковлев, Е.В. Елисеєва // Математичні машини і системи. – 2009. – № 4. – С. 3 – 17.
12. Боюн В.П. Динамическая теория информации. Основы и приложения / Боюн В.П. – Киев: Ин-т кибернетики им. В.М. Глушкова НАН Украины, 2001. – 326 с.

Стаття надійшла до редакції 14.06.2016