



Многопроцессорный вычислительный комплекс для задач «жёсткого» реального времени

Владимир Севбо, Анатолий Орлов, Андрей Лошаков

В статье рассматриваются принципы построения вычислительных систем «жёсткого» реального времени. Приводятся основные сведения о разработанном ЗАО НПП «Авиационная и Морская Электроника» многопроцессорном вычислительном комплексе и использовании данного вычислителя для построения кластерных высокопроизводительных вычислительных систем.

ТРЕБОВАНИЯ, ЗАДАЧИ, ПУТИ РЕШЕНИЯ

В настоящее время наблюдается неуклонный рост потребности в высокопроизводительных малогабаритных цифровых вычислительных комплексах (ЦВК), способных обеспечить обработку данных от большого числа источников информации в «жёстком» реальном времени. В первую очередь это относится к многоканальным системам обработки гидролокационной и радиолокационной информации, видеосигналов, к ультразвуковым и рентгеновским сканерам и др. Для создания таких комплексов в нашей стране используются многопроцессорные вычислительные системы или многомашинные вычислительные комплексы, построенные на основе одноплатных высокопроизводительных многопроцессорных ЭВМ (например «Эльбрус-90микро» или «Багет»).

Многомашинные вычислительные комплексы имеют ряд неоспоримых преимуществ:

- относительная простота реализации;
- возможность практически неограниченного увеличения производительности путём подключения дополнительных ЭВМ;
- относительная простота организации резервирования.

Однако достаточно эффективно эти комплексы функционируют только при реализации относительно малосвязанных задач, требующих процессоров универсального типа. При решении задач

многоканальной сигнальной обработки и схожих с ними по организации вычислительного процесса, в которых задействуется большое число процессоров с интенсивным межпроцессорным информационным обменом, использование многомашинных вычислительных комплексов нецелесообразно: как показано в [1], их реальная производительность составляет здесь лишь 5-10% от потенциально возможной.

Для обеспечения высокой производительности систем массового параллелизма, как правило, используются многопроцессорные векторно-конвейерные и массово-параллельные системы, включая системы кластерного типа [2]. Отдельным подклассом таких систем являются вычислительные системы, предназначенные для работы в «жёстком» реальном времени.

Организация вычислений в режиме «жёсткого» реального времени влечёт целый ряд специфических особенностей. В первую очередь это касается вопросов, от которых непосредственно зависит эффективность многопроцессорного вычислительного комплекса, — вопросов выбора способа построения систем межпроцессорного и внешнего обмена [1].

При построении системы межпроцессорного обмена должны быть обеспечены:

- равнодоступность и полносвязность процессоров;
- возможность простого и эффективного распараллеливания задач между процессорами;

- детерминированность задержек передачи данных;
- синхронизация параллельных процессов;
- выполнение групповых (синхронных) команд и многоадресной рассылки данных;
- исключение конфликтов и независимость работы процессоров от коммуникационной среды;
- возможность наращивания производительности;
- возможность реконфигурации в зависимости от характера решаемых задач.

Использование современных универсальных интерфейсов типа HyperTransport, RapidIO, PCI Express, Fibre Channel не в полной мере позволяет удовлетворить перечисленным требованиям, поэтому построение коммуникационных систем «жёсткого» реального времени, как правило, является специализированной задачей, обеспечивающей заданную эффективность межпроцессорного обмена в зависимости от решаемого класса задач, числа процессоров, выбранной архитектуры построения ЦВК и способа построения системы внешнего информационного обмена.

К основным требованиям, которые предъявляются к системам внешнего информационного взаимодействия, функционирующим в режиме «жёсткого» реального времени, следует отнести:

- обеспечение высокой скорости, помехоустойчивости и надёжности информационного обмена;

- обеспечение детерминизма времени доставки данных и передачи сигналов синхронизации и управления для большого числа источников информации;
- объединение в единый поток данных и команд управления для различных источников информации;
- наличие эффективных систем распределённого контроля, диагностики и резервирования.

Наиболее перспективным решением задач, связанных с выполнением перечисленных требований, является применение Ethernet-технологий. Одновременно использование сетевых технологий позволяет решить такие задачи, как

- возможность подключения к ЦВК разнообразного стандартного оборудования;
- значительное сокращение числа используемых каналов и количества используемых типов интерфейсов;
- простота реализации резервирования на основе использования стандартного коммутационного оборудования;
- полная гальваническая развязка взаимодействующих систем.

В качестве примера рассмотрим многопроцессорный вычислительный комплекс (МПВК), разработанный ЗАО НПП «Авиационная и Морская Электроника» для корабельных систем обработки информации.

МПВК является многопроцессорной суперЭВМ, обладающей производительностью более 25 GFLOPS и предназначенной для многоканальной цифровой обработки больших потоков информации в «жестком» реальном масштабе времени. МПВК может быть использован как в качестве автономной вычислительной системы, так и для построения кластерных суперЭВМ с широким диапазоном производительности.

Внешний вид вычислительного блока МПВК показан на рис. 1.

Состав, особенности конструкции и параметры надёжности МПВК

В состав разработанного вычислительного комплекса входят:

- до 7-8 модулей цифровых процессоров сигналов (МЦПС);
- модуль центрального процессора (МЦП) на базе одноплатного универсального компьютера CPC-502;
- объединительная кросс-плата межмодульных соединений (КПМС-1);
- модуль блока питания (МБП).

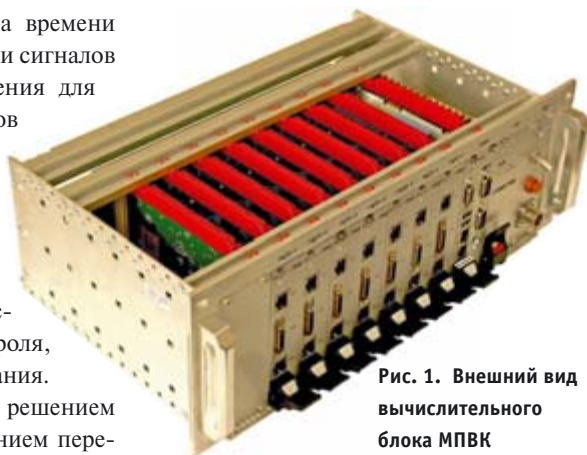


Рис. 1. Внешний вид вычислительного блока МПВК

Каждый МЦПС представляет собой двухпроцессорную систему. Один процессор установлен на основной плате, а второй — на дополнительном мезонинном модуле (ДЦПС). Для реализации в МПВК возможности непосредственного ввода композитного сигнала или радиолокационного видеосигнала с соответствующими сигналами синхронизации и их цифровой обработки в модули МЦПС вместо ДЦПС можно устанавливать модули ввода специальных сигналов, в том числе с высокоскоростными и многоканальными АЦП, модули цифрового ввода-вывода и др.

В основу архитектуры МПВК заложен принцип мультимодульности, что позволяет в зависимости от требований по производительности использовать от 1 до 8 (если не используется МЦП, в противном случае — до 7) вычислительных модулей МЦПС. В случае если МПВК не требует использования МЦП, вместо него устанавливается дополнительный МЦПС.

Для организации внешнего взаимодействия в МПВК применяются следующие интерфейсы:

- Ethernet 10/100/1000Base-T — до 28 плюс 2 канала модуля МЦП (Ethernet 10/100Base-T и 10/100/1000Base-T) для межсистемного взаимодействия или 4 канала восьмого модуля МЦПС (Ethernet 10/100/1000Base-T);



Рис. 2. Внешний вид платы КПМС-1

- RS-232, RS-422 (МЦП) — 2;
- USB 2.0 (МЦП) — 2.

Для организации внутрисистемного взаимодействия используются следующие интерфейсы:

- стандартный — системная шина CompactPCI (PICMG 2.0 D3.0) с пропускной способностью до 133 Мбайт/с;
- специализированные (собственной разработки):
 - магистраль транспортировки данных (МТД) с пропускной способностью от 400 до 1200 Мбайт/с,
 - канал межпроцессорного обмена (в каждом МЦПС) с пропускной способностью до 400 Мбайт/с,
 - канал глобальной синхронизации работы модулей МЦПС (7 линий, в случае установки 8 модулей МЦПС — 8 линий) для организации синхронизации многопроцессорных вычислений.

Конструктивное исполнение:

- блок стандарта Евромеханика (корпус EuroRack PRO фирмы Schroff с защитой от ЭМИ [3]) высотой 3U с габаритными размерами 435×253×149,5 мм;
- модули типоразмера 3U стандарта Евромеханика с габаритными размерами 178×100×11 мм;
- внешнее охлаждение не требуется;
- вес не более 8 кг.

Кросс-плата КПМС-1 (рис. 2) установлена на задней стенке корпуса и обеспечивает электрическую коммутацию модулей МПВК и их механическое крепление. На кросс-плате реализованы следующие интерфейсы:

- системная шина CompactPCI (PICMG 2.0 D3.0), занимающая нижнюю часть разъёмов, — P1;
- магистраль транспортировки данных, занимающая верхнюю часть разъёмов, — P2.

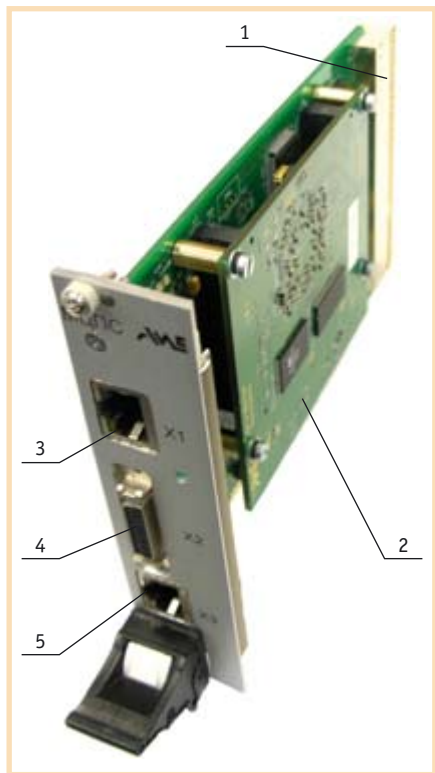
Параметры электропитания:

- 220 В (50 Гц) или 19-36 В;
- потребляемая мощность до 80 Вт.

Комплекс МПВК характеризуется следующими показателями надёжности:

- среднее время наработки на отказ — 5000 часов;
- ресурс до заводского ремонта — 40000 часов;
- назначенный срок службы — 15 лет;
- назначенный срок сохраняемости — 10 лет;
- срок действия гарантийных обязательств — 3 года.

МПВК имеет группу исполнения 2.1.2. по ГОСТ РВ 20.39-304-98 и ГОСТ РВ 20.39.305-98. Он полностью отвечает



Условные обозначения:
 1 — стандартный соединитель CompactPCI 3U для подключения МЦПС к плате КПМС-1;
 2 — съёмный модуль ДЦПС;
 3, 4, 5 — соединители Ethernet (X1, X2, X3).

Рис. 3. Внешний вид модуля МЦПС

требованиям, предъявляемым к корабельной аппаратуре, в частности, выдерживает одиночный механический удар с пиковым ускорением 5g, имеет защиту от соляного тумана и т.п.

ПРОЦЕССОРНЫЕ МОДУЛИ МЦПС И МЦП

Специализированный модуль МЦПС (рис. 3) представляет собой двух-процессорную систему, построенную на базе сигнального процессора TMS320C6713. Второй процессор установлен на мезонинном модуле. Общая производительность модуля составляет до 3600 MFLOPS.

Модуль МЦПС имеет стандартные размеры 3U, занимает в корпусе одну позицию (6НР) по ширине.

Основные характеристики МЦПС представлены в табл. 1.

В настоящий момент готовится к производству версия МЦПС на базе отечественного микропроцес-

Основные характеристики МЦПС

Объём оперативной памяти (на каждый процессор)	Не менее 256 Мбайт
Объём флэш-памяти (на каждый процессор)	Не менее 2 Мбайт
Число каналов JTAG для работы с процессорами и программирования ПЛИС	2
Каналы Ethernet 100/1000Base-TX	4
Протокол передачи	UDP, TCP/IP
Дальность передачи	До 100 м
Скорость передачи	До 1000 Мбит/с
Дополнительные функции использования каналов Ethernet	Произвольная коммутация на любой процессор Перекоммутация и сортировка данных, распараллеливание потока данных на разные МЦПС
Специализированный канал межпроцессорного обмена данными (с ДЦПС)	До 400 Мбайт/с
Два канала МТД по 32 разряда («верхний» и «нижний» порты)	400...1200 Мбайт/с

сора 1892ВМ5 производства ГУП НПП «Элвис».

Модуль МЦП представляет собой отечественный универсальный одно-платный компьютер СРС502, выпускаемый ЗАО НПП «Доломант» по документации и под авторским надзором фирмы «Фаствел». Модуль выполнен в формате CompactPCI 3U. В зависимости от решаемых задач в составе МПВК он может выполнять следующие функции:

- начальная загрузка программы во все процессоры МЦПС;
- управление вычислительным процессом;
- приём и обработка итоговых данных от МЦПС;
- отображение информации;
- обмен данными с другими вычислителями и т.п.

Одним из наиболее важных преимуществ СРС502 является то, что данный одноплатный компьютер является полностью отечественной разработкой, поставляется с полным комплектом документации и с приёмкой заказчика. Опыт применения СРС502 во многих изделиях, выпускаемых ЗАО НПП «Авиационная и Морская Электроника», показал его высокую надёжность, эффективность в использовании и высокий уровень поддержки со стороны производителя.

Внешний вид модуля МЦП показан на рис. 4.

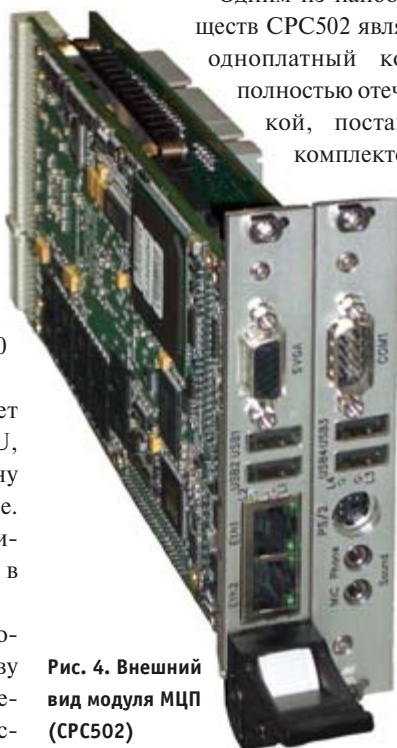


Рис. 4. Внешний вид модуля МЦП (СРС502)

Более подробную информацию о технических характеристиках СРС502 и других аналогичных по назначению изделий, выпускаемых НПП «Доломант», можно получить на официальном сайте фирмы www.dolomant.ru.

ОБЩИЕ ДАННЫЕ О СИСТЕМНОЙ АРХИТЕКТУРЕ МПВК

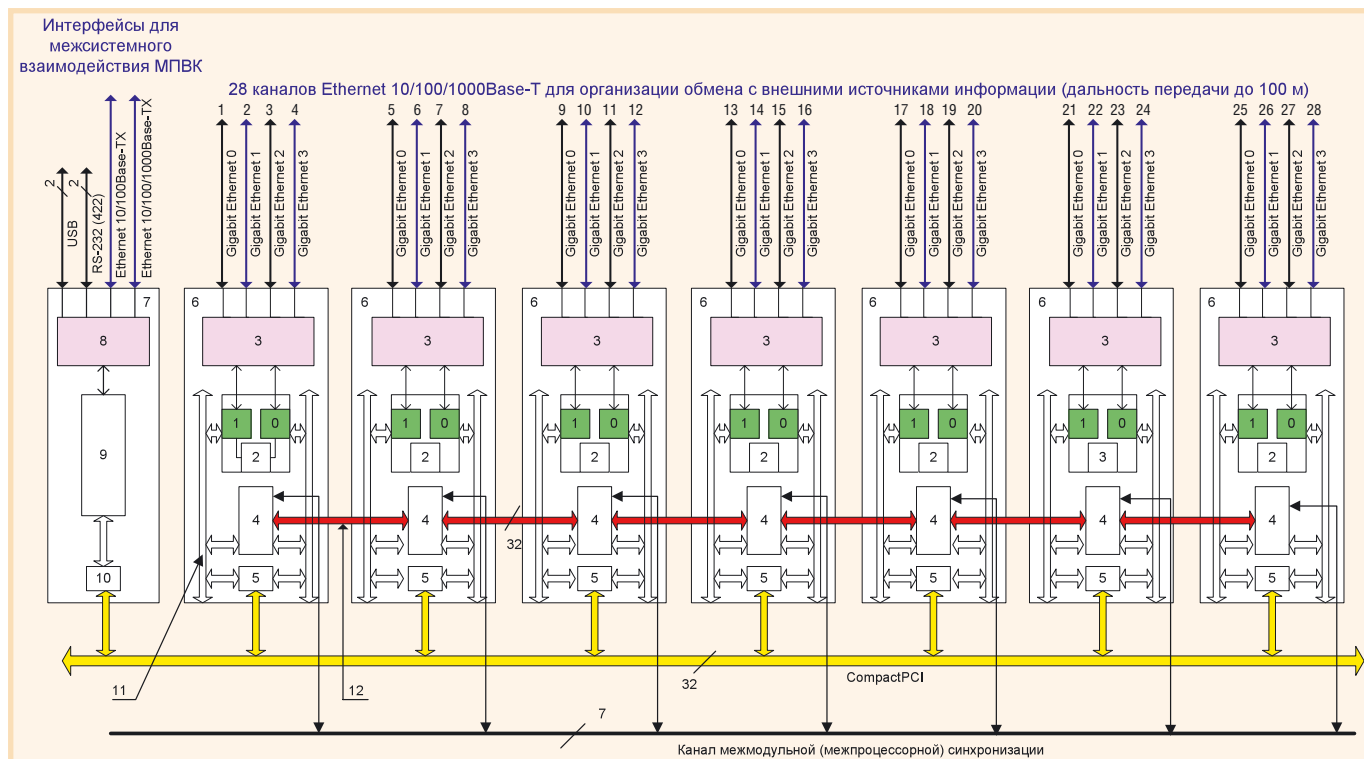
Архитектура МПВК разработана на основании рассмотренных ранее принципов построения многопроцессорных вычислительных систем «жёсткого» реального времени.

Многоуровневая система организации информационного обмена МПВК, построенная по иерархическому принципу, обеспечивает выполнение требований, предъявляемых к организации внутрисистемного и внешнего информационного обмена.

Уровень внутрисистемного взаимодействия представлен:

- системным уровнем, который обеспечивает организацию информационного взаимодействия МЦПС с управляющей ЭВМ (модуль МЦП) по стандартному магистрально-шинному интерфейсу CompactPCI;
- уровнем информационного взаимодействия МЦПС, который обеспечивает высокоскоростной обмен данными между модулями МЦПС по специализированному интерфейсу МТД;
- уровнем межпроцессорного взаимодействия, который обеспечивает обмен данными между основным и мезонинным процессорами в модуле МЦПС (обмен организован по методу «почтового ящика»).

Структурно-функциональная схема вычислительного блока МПВК приведена на рис. 5.



Условные обозначения:

0 — основной процессор МЦПС; 1 — мезонинный процессор МЦПС (ДЦПС); 2 — контроллер межпроцессорного обмена МЦПС; 3 — контроллер внешних интерфейсов МЦПС; 4 — контроллер МТД МЦПС; 5 — контроллер CompactPCI МЦПС; 6 — модуль МЦПС; 7 — модуль МЦПС; 8 — контроллер внешних интерфейсов МЦПС; 9 — процессорный блок МЦПС; 10 — контроллер CompactPCI МЦПС; 11 — шина EMIF процессора; 12 — шина МТД.

Рис. 5. Структурно-функциональная схема вычислительного блока МПВК

Организация информационного обмена по МТД

Для обеспечения требуемой эффективности работы многопроцессорной системы в режиме «жесткого» реального времени сотрудниками ЗАО НПП «Авиационная и Морская Электроника» была разработана магистраль МТД, отличающаяся высокой скоростью передачи данных, присущей системам, построенным на основе Link-каналов, которая обеспечивает равнодоступность и полносвязность процессоров, а также гибкость в управлении информационным обменом, свойственную системам с архитектурой гиперкуб. Реализация синхронного протокола передачи данных позволила исключить возникновение конфликтных ситуаций и обеспечить детерминированность времени доставки данных. С помощью канала глобальной синхронизации обеспечивается синхронизация параллельных процессов.

Структурно-функциональная схема МТД приведена на рис. 6.

Каждый МЦПС содержит коммутационный узел, в состав которого входят два независимых двунаправленных 32-разрядных канала, представленных «нижним» портом (НП) и «верхним» портом (ВП). МТД организована так,

что все коммутационные узлы МЦПС последовательно соединены друг с другом, причём НП предыдущего узла МЦПС соединён с ВП последующего узла. Первый МЦПС в данном случае содержит верхний начальный узел (ВНУ) МТД, а последний — нижний конечный узел (НКУ) МТД. Определение количества МЦПС и их нумерация осуществляются при начальном включении МПВК. Теоретически количество узлов в МТД может быть произвольным и определяться только требованиями к производительности многопроцессорной вычислительной системы.

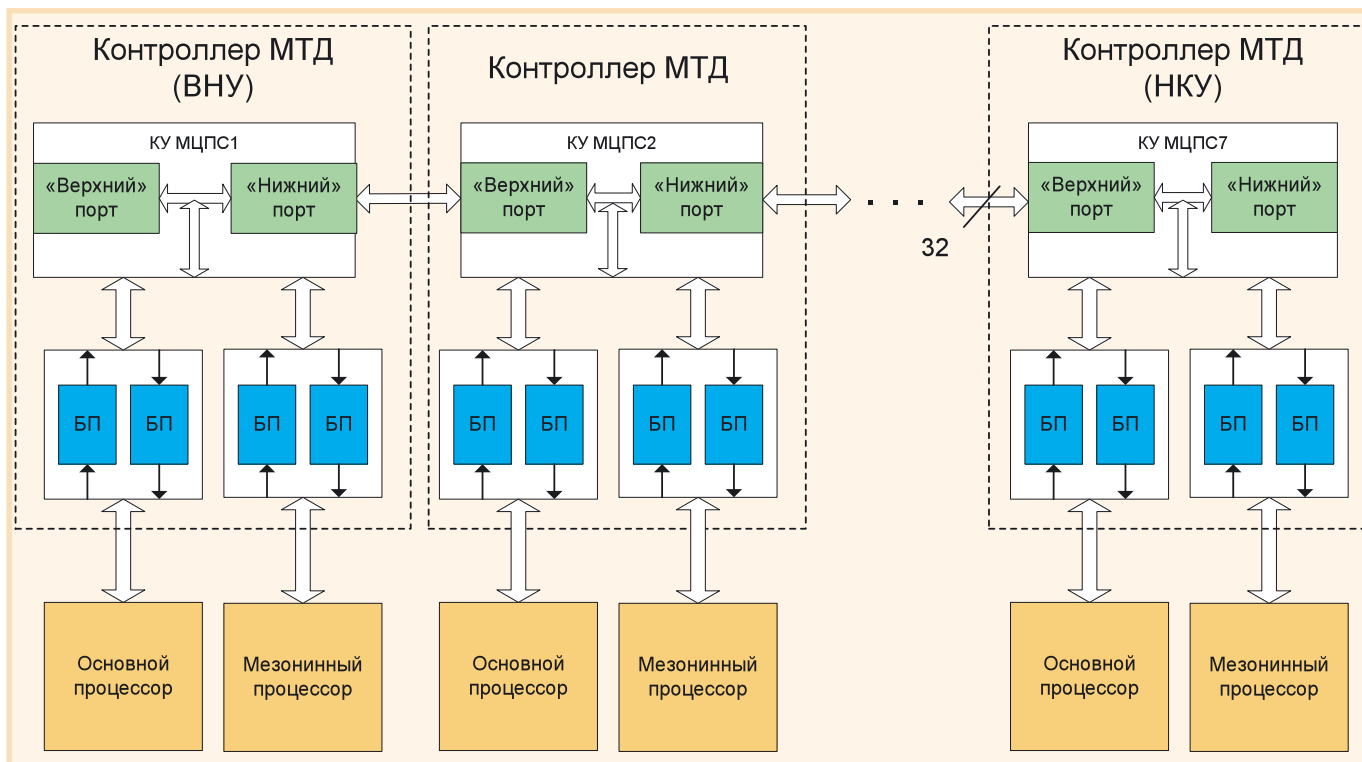
Для обеспечения возможности исключения процессоров в ходе информационного обмена передача данных по МТД основана на использовании механизма буферизации. Каждый процессор помещает в буферную память данные с адресом получателя, которые необходимо передать, и считывает поступившие для него данные из другой буферной памяти контроллера МТД. Запись и чтение данных осуществляются с помощью каналов EDMA в режиме Burst Mode со скоростью 400 Мбайт/с. Для организации взаимодействия процессора с контроллером МТД используется механизм прерываний или механизм работы с флагом.

Передача данных по МТД осуществляется пакетами. Пакет состоит из заголовка и блока данных. Заголовок пакета содержит адрес источника, адрес приёмника информации (адреса приёмников при групповой рассылке) и размер передаваемого блока данных. Блок данных должен состоять не менее чем из восьми 32-разрядных слов, но не превышать размер в 1024 (2048) 32-разрядных слова.

Информационный обмен в МТД организован методом циклической передачи прав на передачу данных каждому МЦПС (метод циклического опроса заявок на передачу). Каждый узел МТД имеет возможность передачи данных одному или нескольким (групповая рассылка) МЦПС.

Данные по МТД передаются в оба направления через «верхний» и «нижний» порты. Передача данных может производиться с подтверждением готовности контроллеров МТД к приёму информации или без подтверждения. Считается, что узел не готов к приёму информации, если в данный момент времени занята его приёмная буферная память. В случае получения сигнала неготовности узла к приёму передача пакета прекращается. Права на передачу переходят к следующему узлу, а попытка повторной передачи данного пакета будет осуществлена в следующем цикле. Каждый узел обеспечивает синхронизацию передачи данных до очередного узла.

После окончания передачи данных или при их отсутствии передающий узел МТД



Условные обозначения:

КУ — коммутационный узел; ВНУ — «верхний» начальный узел; НКУ — «нижний» конечный узел; БП — буферная память.

Рис. 6. Структурно-функциональная схема МТД

уступает права на передачу следующему узлу. Последний узел МТД по окончании передачи данных передаёт права на передачу в начало МТД (первому узлу).

Узлы МТД, не осуществляющие передачу, обеспечивают трансляцию принятых данных следующему узлу. За счёт реализации механизма передачи прав каждый узел МТД знает, где относительно него находится активный узел, и поэтому может однозначно определить направление трансляции данных. При ретрансляции осуществляется регенерация передаваемых сигналов, что обеспечивает возможность высокоскоростной передачи данных без ограничения дальности.

Для обеспечения слаженной работы процессоров по каналу глобальной синхронизации МТД передаются сигналы синхронизации и управления. Для организации синхронизации и управления используется один из процессоров (задаётся программным способом), который является ведущим, а все остальные процессоры — ведомыми.

В зависимости от характера решаемых задач предусмотрена возможность логического разделения МТД на отдельные участки, способные осуществлять независимую передачу данных в своём сегменте, что существенно повышает (до 1200 Мбайт/с) пропускную способность МТД. Сегментирование

МТД обеспечивается за счёт создания нескольких «верхних» и «нижних» портов, ограничивающих область сегмента. Реконфигурация МТД может осуществляться многократно без нарушения работоспособности системы.

Системная архитектура модуля МЦПС

Функциональная схема модуля МЦПС приведена на рис. 7.

Начальная установка и загрузка программы в основной и мезонинный процессоры TMS320C6713 осуществляется от управляющей ЭВМ (МЦП) по каналам CompactPCI через НРІ с помощью контроллеров НРІ. Возможность доступа МЦП к внутренним ресурсам процессоров (регистры, память) также реализуется по интерфейсу НРІ.

Оперативная память SDRAM и флэш-память подключены к процессорам по интерфейсу EMIF в адресном пространстве СЕ 2.

Для организации межпроцессорного обмена используется соответствующий контроллер, который подключён к каналам EMIF каждого процессора. В состав контроллера межпроцессорного обмена входят два независимых буфера обмена БП 0 (буферная память основного процессора) и БП 1 (буферная память мезонинного процессора) объёмом по

1 кбайт и общедоступные регистры управления. Скорость чтения/записи в буферы обмена — до 400 Мбайт/с. Для управления обменом данными предусмотрены процессорные прерывания или механизм работы с флагом.

Для организации информационного обмена по каналам CompactPCI используются контроллеры CompactPCI, реализующие системный интерфейс CompactPCI (стандарт PICMG 2.0 D3.0). Во входной буфер осуществляется запись данных для процессора, принятых по шине CompactPCI, а в выходной буфер — данных от процессора для передачи по шине CompactPCI. Взаимодействие буферов с процессором осуществляется по каналам EMIF. Объём каждого буфера составляет 1 кбайт, вместе они образуют буферную память контроллера CompactPCI. Скорость чтения/записи в буферную память по шине EMIF составляет до 400 Мбайт/с. Обмен данными с МЦП каждый МЦПС может осуществлять в режимах «Target» или «Master». В контроллере реализована возможность передачи прерываний со стороны процессоров на управляющую ЭВМ (МЦП) или приёма сигналов сброса процессоров со стороны МЦП.

Обмен данными между МЦПС организован с помощью магистрали МТД. Для управления передачей данных используется контроллер МТД, в состав которого входят коммутационный узел, реализующий интерфейс передачи данных по магистрали, и приёмопередающие блоки для каждого процессора. Ка-

