

РОССИЙСКАЯ АКАДЕМИЯ НАУК
СИБИРСКОЕ ОТДЕЛЕНИЕ
А В Т О М Е Т Р И Я

№ 1

1993

ВЫЧИСЛИТЕЛЬНЫЕ СЕТИ И СИСТЕМЫ
В. Д. Бобко, В. Ф. Головин, Ю. Н. Золотухин

(Новосибирск)

ВЫСОКОПРОИЗВОДИТЕЛЬНЫЙ УЗЕЛ КОММУТАЦИИ
ДЛЯ ИНТЕГРАЛЬНОЙ ЦИФРОВОЙ СЕТИ

Предложена архитектура многопроцессорной вычислительной системы с общей памятью и многошинной структурой связей, использованная для реализации центра коммутации пакетов. Обсуждаются методы организации взаимодействия элементов системы.

1. Вычислительные сети — основные поставщики информационных ресурсов в современном обществе — обычно используют в качестве транспортной системы базовую сеть обмена данными [1]. Важнейшими элементами базовой сети являются каналы передачи информации и коммутационные узлы. К одной из ключевых характеристик, оказывающей влияние практически на все основные показатели сети, такие как время установления соединения, скорость доставки сообщений, достоверность передачи информации и другие, следует отнести используемый в сети метод коммутации.

Важнейшей задачей, решавшейся при проведении исследований в области интегральных цифровых сетей в Институте автоматики и электрометрии СО РАН [2], была разработка архитектуры, структуры связей и алгоритмов функционирования высокопроизводительного узла коммутации пакетов. Для получения оценок основных характеристик узла проведены исследования особенностей совместной цифровой пакетной передачи речи и данных на специализированном стенде [3].

Можно отметить, что требования к каналам как к среде передачи информации в достаточно слабой степени зависят от метода коммутации, в то время как требования к узлам при использовании, к примеру, метода коммутации каналов и метода коммутации пакетов совершенно различны. Метод коммутации каналов (коммутации цепей) основан на предоставлении абонентам, участвующим в сеансе связи, в монопольное распоряжение на время сеанса некоторого сетевого ресурса (пропускной способности, полосы частот и т. п.). Применительно к узлам это выливается в необходимость осуществлять коммутацию, обеспечивая абонентам непосредственное соединение через сеть на время сеанса и разъединение после его завершения; очевидно, что этот метод не предъявляет жестких требований к сложности и быстродействию коммутирующих устройств. Основной недостаток метода коммутации каналов — это, как правило, неэффективное использование сетевых коммуникационных ресурсов и значительное время установления соединения; тем не менее этот метод — исторически первый и широко используется в настоящее время (наиболее известным примером системы с коммутацией каналов является телефонная сеть).

В отличие от коммутации каналов метод коммутации пакетов не предполагает предоставления паре абонентов, участвующих в сеансе связи, сетевых ресурсов в монопольное владение на время сеанса. В сетях с коммутацией пакетов соседние узлы соединены высокоскоростными каналами, использующимися, по сути, в режиме асинхронного временного уплотнения. Подлежащая пересылке от абонента к абоненту информация разбивается на порции (пакеты), снабжается указателями принадлежности к некоторому сеансу связи и продвигается в нужном направлении от узла к узлу, подвергаясь в каждом из них воздействию коммуникационных процедур. При этом возможны: передача как с предварительным установлением соединения между абонентами, так и без него; следование всех пакетов, относящихся к одному соединению, по заранее «проложенному» маршруту либо индивидуальная маршрутизация каждого пакета; использование статических или динамических (например, адаптивных к состоянию сети) процедур маршрутизации и т. д. Повышение эффективности функционирования сети в целом, в том числе и улучшение использования наиболее дорогих элементов сети — каналов, связано с усложнением коммуникационных процедур в узлах сети.

2. Как правило, узлы в сети пакетной коммутации, соответствующей рекомендациям X.25 МККТТ [4], реализуют функции трех нижних уровней модели взаимодействия открытых систем ISO [5] (рис. 1):

1) физического (управление модемом или другим аналогичным устройством и обеспечение передачи и приема последовательностей цифровых сигналов);

2) канального (обмен информацией в виде кадров и обеспечение надежного канала доставки данных);

3) сетевого (организация нескольких логических каналов по одному физическому и взаимодействие не связанных непосредственно систем). Естественно также, что узел выполняет функции маршрутизации потоков информации.

Дополнительно к вышеперечисленным обязательным функциям узел коммутации может производить сбор статистики и управление собственными устройствами, вести учет переданной информации, а также осуществлять взаимодействие с административной подсистемой сети.

Поскольку метод коммутации пакетов предполагает лишь кратковременное, промежуточное хранение пакетов в узлах (только на время обработки), а допустимое время задержки в узле обычно невелико, вычислительная система, реализующая функции узла коммутации, может быть отнесена к классу систем реального времени.

Известны различные подходы к построению узлов коммутации [6]. В простейших центрах коммутации пакетов (ЦКП), созданных на базе серий-

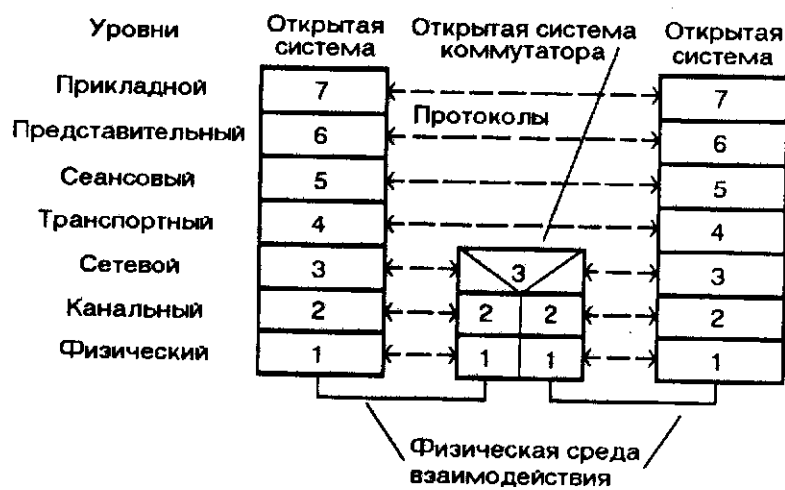


Рис. 1

ных ЭВМ, большинство функций сетевого узла коммутации реализуется программным путем с использованием вычислительных ресурсов универсальной ЭВМ, лишь подключение к каналам связи осуществляется с помощью специализированных устройств ввода/вывода — адаптеров линий связи (АЛС).

Такой подход вполне оправдан на начальных этапах разработок и во время опытной эксплуатации сетей, а также при реализации ЦКП небольшой коммутационной производительности. Так, при создании первой сети с коммутацией пакетов ARPANET [7] в качестве узлов коммутации использованы ЭВМ H516 фирмы "Honeywell"; производительность такого узла оценивалась в 750 Кбит/с. Центры коммутации в «Академсети» [8], реализованные на мини-ЭВМ СМ-4, имеют производительность порядка 30 пакетов/с. Повышение производительности ЦКП при таком подходе может быть достигнуто путем использования более мощных ЭВМ и разработки более совершенных линейных адаптеров.

Особенности метода коммутации пакетов (взаимная независимость пакетов, использование многоуровневой архитектуры процедур их обработки в узлах коммутации), а также необходимость одновременной работы с многими межузловыми и абонентскими линиями делают возможным применение также и другого подхода к построению ЦКП с высокими техническими характеристиками (такими как канальная емкость, производительность, максимальная скорость обмена в каналах связи) — использования архитектуры многопроцессорных вычислительных систем (МВС).

Как правило, определяющую роль при организации МВС играет структура связей между модулями, входящими в ее состав: процессорными блоками, блоками памяти и блоками ввода/вывода. Если исходить из задачи реализации МВС, в которой отсутствует функциональная специализация процессоров (что может предоставить ряд дополнительных полезных свойств, таких как однородность аппаратных и программных средств, упрощение эксплуатации, потенциальная отказоустойчивость и ремонтпригодность), то, очевидно, что каждому процессорному блоку необходимо обеспечить доступ ко всему пространству общей памяти, а также возможность управления работой всех устройств ввода/вывода; блокам ввода/вывода обеспечить доступ ко всем блокам общей памяти.

3. В настоящей работе приведено описание предложенной авторами структуры связей, базирующейся на многошинной архитектуре. Основу коммуникационных средств МВС составляет разработанный специализиро-

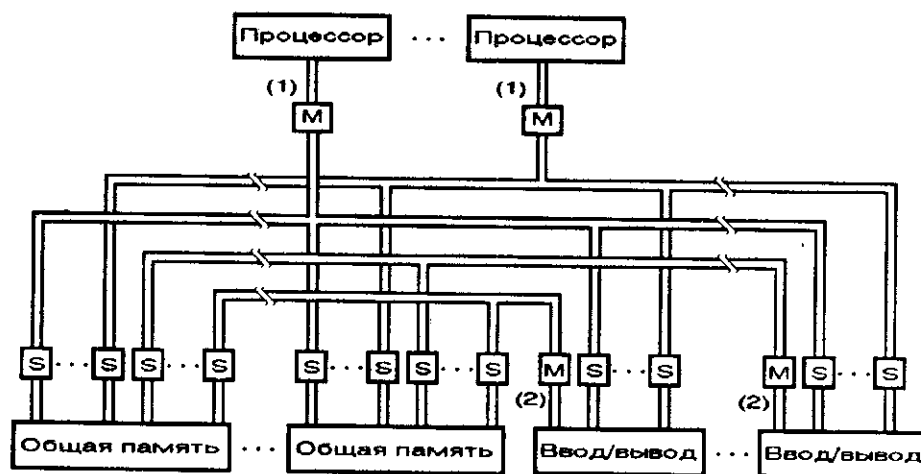


Рис. 2

ванный комплект соединителей шин, который обеспечивает процессорам возможность обращения к общей памяти в режиме прямого доступа с использованием механизма виртуальной адресации в 22-разрядном адресном пространстве. Схема использования соединителей для организации взаимодействия между блоками МВС приведена на рис. 2. Для обеспечения обмена информацией между блоками, находящимися на различных системных уровнях, разработаны межблочные соединители двух типов: «процессор — память — ввод/вывод» (1) и «ввод/вывод — память» (2). В комплект оборудования любого межблочного соединителя входят одно ведущее и несколько ведомых устройств. Для связи процессора с общей памятью и устройствами ввода/вывода в процессорном блоке устанавливается ведущее устройство соединителя (М), а в блоках памяти и ввода/вывода — ведомые (S); устройства соединителя объединяются кабелями, которые образуют общую шину. При обращении процессора к общей памяти или блоку ввода/вывода ведущее устройство соединителя преобразует 16-разрядный виртуальный адрес шины процессора в 22-разрядный системный адрес и передает его всем ведомым устройствам. Ведомое устройство блока памяти (ввода/вывода), дешифровав обращение по адресу, относящемуся к «своему» диапазону адресов, выдает запрос на доступ к внутренней шине блока. Передача данных при этом производится по тракту, образуемому шиной процессора, общей шиной соединителя и шиной выбранного блока. Соединители типа «ввод/вывод — память» работают аналогичным образом за исключением преобразования адресов: ведущее устройство (М) блока ввода/вывода «выдает» полный системный адрес; для обмена данными образуется тракт: шина блока ввода/вывода — общая шина соединителя — шина выбранного блока общей памяти.

Исходя из требований отказоустойчивости предусмотрена возможность процессорам отключать и вновь включать соединители других процессоров и устройств ввода/вывода; у отключенных соединителей запрещается работа дешифратора ведомого устройства.

В известной [9] многопроцессорной вычислительной системе для каждой связи «процессор — общая память» используется отдельный соединитель; аналогично организуются связи «процессор — ввод/вывод» и «ввод/вывод — общая память». Предложенная многосиная структура позволяет существенно сократить объем коммуникационного оборудования без снижения показателей производительности системы, что очевидно, так как в обеих структурах в каждый момент времени любой блок может осуществлять лишь одно взаимодействие через соединители с другими блоками.

При построении МВС одной из центральных является проблема разрешения конфликтов при попытке одновременного обращения к общим ресурсам (блокам памяти и ввода/вывода). Для систем с жестко установленным приоритетом в случае интенсивного обращения возможны ситуации, когда источники заявок, находящиеся на низших приоритетных уровнях, в течение длительного времени не получают доступа к шине, что может оказаться недопустимым. Из этого следует, что необходимо создать такой механизм арбитража, когда для каждого источника заявок обеспечивается гарантированное время доступа. Адекватным методом решения этой проблемы является дисциплина обслуживания, при которой приоритет источника циклически меняется после каждого обслуживания. Такая же циклическая дисциплина может быть использована и для разрешения конфликтов между заявками на обслуживание в циклах прямого доступа и предоставления прерывания. Для осуществления арбитражных и иных управляющих функций в каждом блоке предусмотрен арбитр шины.

Еще одна проблема, требующая решения при разработке многопроцессорных систем, — это распределение заданий между процессорами и взаимодействие процессоров между собой. Иногда для этого выделяется специальный процессор-диспетчер, в некоторых случаях — распределение заданий производится программным способом. В рамках принятого подхода «равноправности» процессоров реализована известная (см., например, [10]) модель системы обслуживания с относительным приоритетом, когда заявки образуют упорядоченную в соответствии с установленным приоритетом очередь, а каж-

дый завершивший текущее обслуживание заявки процессор выбирает для обработки наиболее приоритетную заявку из очереди.

Описанные выше функции распределения заданий в системе реализованы 4. Как уже упоминалось, в состав мультипроцессорной системы входят блоки трех основных типов: процессорные, общей памяти и ввода/вывода.

Вычислительная мощность системы сосредоточена в процессорных блоках. Процессорный блок функционирует как самостоятельная ЭВМ в многозадачном режиме реального времени, все взаимодействия его с другими процессорами системы осуществляются через общую память. Наличие локальной памяти в каждом процессорном блоке позволяет размещать в общей памяти, наряду с буферами для транзитных пакетов, лишь необходимые для взаимодействия системные компоненты, вследствие чего снижается интенсивность обращений к общей памяти.

Для уменьшения количества конфликтов при одновременном обращении к общей памяти нескольких процессоров предусмотрено секционирование памяти, обеспечивающее возможность одновременного независимого доступа к различным блокам памяти. В случае одновременного обращения нескольких процессоров к одному блоку общей памяти возникают задержки вычислительного процесса, связанные с образованием очереди за предоставлением доступа к памяти. Как уже отмечено, в арбитраже каждого из блоков общей памяти реализован механизм циклического обслуживания запросов, что обеспечивает «выравнивание» средних параметров обслуживания для различных запросов. При этом уровень приоритета регенерации модулей динамической памяти всегда самый высокий.

Арбитр шины — это многофункциональное устройство, реализующее, с одной стороны, специальные, описанные выше функции для организации МВС, а с другой — необходимые управляющие функции шины внутри блока, к которым можно отнести: обмен данными в цикле предоставления прямого доступа к шинам в блоках памяти и ввода/вывода; стандартное взаимодействие в цикле предоставления прерывания в блоке ввода/вывода; регенерацию информации, записанной в ячейки динамической памяти, в блоках общей памяти, причем узел регенерации работает в соответствии с дисциплиной прямого доступа; выдачу сигнала начальной установки (сброса) при включении/восстановлении питания.

Для повышения надежности функционирования системы в целом в арбитраже предусмотрены: возможность отключения неисправных соединителей шин; сброс шины блока программным способом или по ситуации, когда шина неактивна в течение определенного времени.

Состав блока ввода/вывода в МВС несет на себе отпечаток функций, которые должны реализовываться системой. В данной работе, направленной на решение проблемы интеграции услуг, предоставляемых цифровой сетью, состав устройств, входящих в блок ввода/вывода, определяется задачей создания высокопроизводительного центра коммутации пакетов. Каждый блок ввода/вывода содержит адаптеры линий связи — основные функциональные устройства, арбитр шины, интерфейсы общесистемных периферийных устройств, плату диспетчера очереди заданий, общесистемное ПЗУ, одно ведущее и несколько ведомых устройств соединителей.

Многопроцессорная архитектура, высокая производительность, работа с высокоскоростными линиями определили основные требования к АЛС, используемым в многопроцессорном ЦКП. Каждый адаптер линии связи может непосредственно обращаться к общей памяти в режиме прямого доступа для записи или чтения информационных пакетов. Это дает возможность замкнуть основной информационный поток непосредственно через общую память ЦКП, минуя процессоры, на которые при этом возлагаются лишь задачи

управления информационными потоками и распределения ресурсов общей памяти.

Аппаратная реализация в АЛС некоторых функций обработки кадров, ориентированная на принятую структуру разбиения адресного пространства общей памяти, позволяет повысить быстродействие системы и дополнительно высвободить вычислительные ресурсы процессоров. В упрощенном виде процедуры обмена информацией между процессором и адаптером линии связи выглядят следующим образом. Информация передается по линиям связи в виде кадров и хранится в общей памяти ЦКП в специальном формате, называемом буфером. Каждый буфер предназначается для хранения одного кадра, имеет фиксированную длину и состоит из двух частей: заголовка и остатка. Заголовок буфера состоит из 12 байтов: нулевой и первый байты заголовка содержат указатель на следующий буфер, второй — информацию о длине кадра, третий — статусную информацию о буфере; следующие шесть байтов заголовка содержат первые шесть байтов кадра, т. е. поля А и С кадра и четыре первых байта из информационного поля кадра; десятый и одиннадцатый — указатель на остаток буфера. Остаток буфера предназначен для хранения информационной части пакета.

Адресное пространство ЦКП условно разделено на сегменты по 128 байтов (32768 сегментов). Для хранения номера сегмента в системе отводится одно 16-разрядное слово (два байта указателя остатка в заголовке буфера). Области хранения заголовков и остатков буферов в общей памяти разделены.

В каждом АЛС имеется буферная память для хранения списков адресов заголовков восьми сегментов памяти для приема и восьми — для передачи кадров, что позволяет согласовать процессы приема, передачи и обработки последовательности кадров при максимальной уплотненности (только одна флаговая последовательность между двумя соседними кадрами) и высокой скорости передачи в линии связи. В начале работы операционная система производит распределение общей памяти, в частности, выделяются области заголовков и остатков, составляются и загружаются в буферную память АЛС списки адресов заголовков отведенных соответствующим адаптерам сегментов общей памяти. Для приема последовательности кадров из линии адаптер выбирает очередной адрес буфера из списка, принимает кадр, упаковывает принятые данные в пустой буфер, указывая длину принятого кадра, статусную информацию, через механизм ДОЗ передает системе адрес уже заполненного данными буфера и выбирает следующий адрес. Аналогичным образом осуществляется передача кадра: линейному адаптеру передается список последовательности адресов подлежащих передаче буферов с упакованными в них данными. Адаптер выбирает очередной буфер, распаковывает данные, передает их в линию в формате кадра, затем через ДОЗ информирует мультипроцессор о завершении передачи кадра и выбирает адрес следующего буфера. Таким образом, процессорам для поддержания непрерывного процесса приема и передачи кадров необходимо лишь своевременно пополнять буферные памяти заголовков в АЛС. Использование режима прямого доступа освобождает процессоры от осуществления пересылок данных между линейными адаптерами и памятью. Использование буферной памяти заголовков значительно снижает критичность требований ко времени реакции мультипроцессора при малых межкадровых промежутках.

5. В основу реализации многопроцессорной системы ЦКП положен метод максимальной унификации аппаратных средств наряду с широким применением принципа однородности. Базовым конструктивным элементом для построения МВС выбрана микроЭВМ с общей шиной МПИ [11], используемая во всех блоках системы, что позволило применить серийно изготавливаемые устройства, такие как процессоры, запоминающие устройства, интерфейсы периферийного оборудования. На рис. 3 приведен созданный в ИАиЭ СО РАН макет многопроцессорного ЦКП, а на рис. 4 — его структурная схема. Экспериментальный макет ЦКП содержит четыре процессорных блока, по одному блоку общей памяти и ввода/вывода. Каждый процессорный блок содержит процессор (ПР), локальную оперативную память (ЗУ), постоянное запоминающее устройство, интерфейсы для подключения локальных внеш-

них устройств (И), а также ведущее устройство соединителя шин (М). Блок общей памяти содержит арбитр шины (АШ), ведомые устройства соединителей (S) и собственно модули динамической памяти (ДЗУ), требующей периодического восстановления (регенерации) хранимой информации. Блок ввода/вывода содержит арбитр шины, адаптеры линий связи, интерфейсы для подключения общесистемных внешних устройств, диспетчер очереди заданий, ведущее М и ведомые S устройства соединителей шин.

Программное обеспечение ЦКП [12] поддерживает три нижних уровня модели ВОС в соответствии с требованиями рекомендации X.25 МККТТ в режимах коммутируемых и постоянных виртуальных каналов; кроме того, обеспечивается возможность совместной коммутации речевых пакетов и пакетов данных.

6. Для исследования производительности ЦКП разработана специальная тестовая программа, генерирующая поток запросов ко второму уровню X.25 и осуществляющая регистрацию количества переданных и принятых кадров. Тестовая программа также исполняется процессорами ЦКП, предусмотрена возможность вари-

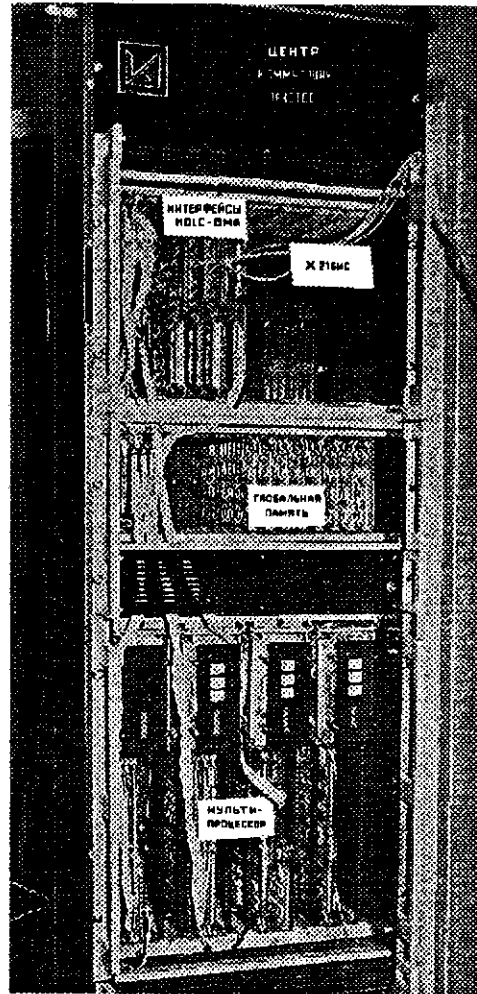


Рис. 3

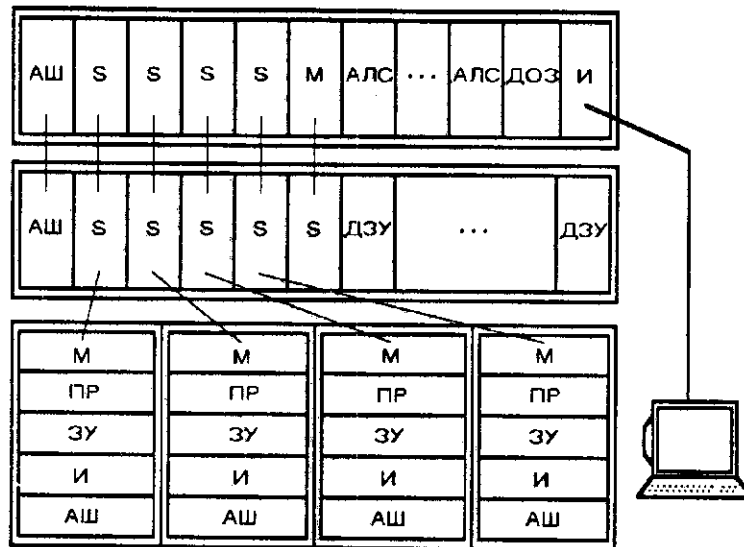


Рис. 4

Число процессоров	Число кадров/с
1	800
2	1900
3	2600
4	2800

рования числом процессоров в системе. Полученные результаты приведены в таблице.

Непропорционально большой рост числа обрабатываемых кадров для конфигурации с двумя и тремя процессорами является кажущимся и объясняется уменьшением суммарных накладных расходов от дополнительной загрузки системы исполнением тестовой программы. Рост производительности при трех и четырех процессорах замедляется из-за увеличения числа конфликтов при обращении к общим ресурсам.

Аналитическое исследование модели предложенной многопроцессорной вычислительной системы и анализ ее производительности проведены в [13]. Возможности повышения производительности многопроцессорной вычислительной системы в значительной степени определяются организацией использования общих ресурсов. Коэффициент загрузки общей памяти определяет асимптотическую оценку производительности при росте числа процессоров. В рамках предложенной архитектуры возможно увеличение производительности ЦКП путем увеличения числа блоков нужного типа; например, увеличение числа блоков общей памяти при обеспечении независимого к ним доступа приводит к уменьшению коэффициента загрузки и соответственно снижению потерь из-за конфликтов. При этом появляется возможность дальнейшего увеличения производительности при увеличении числа процессорных блоков и блоков ввода/вывода. Соответствующим подбором параметров можно добиться практически линейной зависимости производительности многопроцессорной вычислительной системы в достаточно широкой области изменения числа процессоров в системе.

7. Архитектура предложенной многопроцессорной вычислительной системы с общей памятью и многослойной структурой связей позволяет создавать комплексы различной конфигурации, отличающиеся по производительности, числу подключаемых внешних устройств и объему памяти.

Модульность и отсутствие функциональной специализации процессоров, положенные в основу системы, позволяют добиться потенциальной отказоустойчивости при введении минимальной избыточности и обеспечивают постепенную деградацию характеристик при отказах.

Разработанная многопроцессорная вычислительная система, а также аппаратные и программные методы организации взаимодействия ее элементов обеспечивают реализацию высокопроизводительных узлов сети с коммутацией пакетов.

СПИСОК ЛИТЕРАТУРЫ

1. Мизин И. А., Богатырев В. А., Кулешов А. П. Сети коммутации пакетов.—М.: Радио и связь, 1986.
2. Бобко В. Д., Головин В. Ф., Золотухин Ю. Н., Халбашкеев Ю. Ф. Высокопроизводительный центр коммутации пакетов для интегральной цифровой сети связи // X Всесоюз. шк.-сем. по вычислительным сетям: Тез. докл.—Ч. 3.—Москва — Тбилиси: ВИНТИ, 1985.
3. Ефимов В. М., Золотухин Ю. Н., Лившиц З. А., Якушев В. С. Исследование пакетной передачи речи в интегральных сетях связи // Автоматизация научных исследований: Тез. докл. XIX Всесоюз. шк.—Новосибирск: ИАиЭ СО АН СССР, 1985.
4. CCITT. The International Telegraph and Telephone Consultative Committee: VIIIth Plenary Assembly, Malaga-Torremolinos, 8-19 oct. 1984.—Geneva, ITU: Red Book, 1985.
5. ISO/IS 7498. Information Processing Systems. Open System Interconnection // Basic Reference Model, 1983.
6. Шигин Г. А. Сети и системы с коммутацией пакетов (Ч. 3) // Зарубеж. радиоэлектрон.—1981.—№ 5.
7. Шигин Г. А. Сети и системы с коммутацией пакетов (Ч. 1) // Там же.—№ 3.
8. Калтыгин М. И. Центр коммутации пакетов опытной зоны АКАДЕМСЕТИ // X Всесоюз. шк.-сем. по вычислительным сетям: Тез. докл.—Ч. 3.—Москва — Тбилиси: ВИНТИ, 1985.

9. Катсуки Д., Элсам Э. С., Манн У. Ф. и др. Pluribus — отказоустойчивый операционный мультипроцессор // ТИИЭР.—1978.—66, № 10.
10. Кофман А., Крюон Р. Массовое обслуживание: теория и применение.—М.: Мир, 1965.
11. ГОСТ 26765.51-86. Интерфейс магистральный параллельный МПИ системы электронных модулей.—М.: Изд-во стандартов, 1986.
12. Песляк П. М. Архитектура программного обеспечения центра коммутации пакетов // Автометрия.—1993.—№ 1.
13. Бобко В. Д., Головин В. Ф., Золотухин Ю. Н. Анализ производительности мультипроцессорной вычислительной системы // Автометрия.—1992.—№ 6.

Поступила в редакцию 27 ноября 1992 г.

23