

УДК 681.142.019.322

МНОГОМАШИННАЯ ВЫЧИСЛИТЕЛЬНАЯ СИСТЕМА С ОБЩИМ ПОЛЕМ ПАМЯТИ

А.А.Папернов, О.И.Гаранина

В статье рассматривается структура мультипрограммной вычислительной системы (ВС), предназначенной для решения набора задач, часть которых информационно связана друг с другом и с обработкой информации, поступающей от разнородных источников. Вычислительная система содержит несколько вычислителей и общее информационное поле достаточно большого объема, позволяющее вычислителям оперативно обмениваться информацией в процессе выполнения программ. В связи с мультипрограммной структурой ВС в общем информационном поле необходимо хранить большое количество программ и исходных данных к ним. По экономическим соображениям нецелесообразно хранить в оперативной памяти (ОЗУ) программы и данные по всем задачам в полном объеме. Для хранения информации по программам или участкам программ, непосредственно не выполняемым в данный момент времени, можно использовать более дешевые накопители на магнитных лентах (МЛ) и барабанах (МБ) и организовать обмен между ними и ОЗУ в процессе работы ВС. Имеется возможность включать программы, вводимые для решения в ВС, в несколько приоритетных списков с тем, чтобы программа из высшего приоритетного списка в случае возможности ее выполнения могла прерывать решение задач из низшего приоритет-

ного списка. С целью максимального использования вычислительных мощностей ВС предусмотрена возможность продолжения выполнения ранее прерванной программы не обязательно на том вычислителе, на котором она выполнялась до прерывания, а на любом освободившемся. Это возможно только тогда, когда вся записанная в памяти информация доступна любому вычислителю, т.е. вся имеющаяся в ВС оперативная и внешняя память входит в состав общего информационного поля.

Для организации работы ВС, а также для организации обмена между МБ и МЛ и ОЗУ используется программа-диспетчер, включаемая на любом вычислителе ВС.

Блок-схема показана на рис.1. В состав ВС входят:

- 1) вычислители,
- 2) устройства обмена,
- 3) общее информационное поле, состоящее из:
 - а) секционированного оперативного ЗУ,
 - б) группы МБ,
 - в) группы МЛ,
 - г) кодовых шин для связи оперативного ЗУ с вычислителями, устройствами обмена, МБ и МЛ,
- 4) устройство управления ВС.

Вычислителями предназначены для непосредственной обработки информации. В них не содержится никакой другой памяти, кроме буферной. Введение буферной памяти в состав вычислителей дает возможность уменьшить поток обращений к общему полю, усреднить этот поток по времени, а также за счет существенной разницы по времени обращения к буферной памяти и ОЗУ повысить производительность ВС.

В составе вычислителей имеется:

- 1) устройство подготовки команд,
- 2) буферное запоминающее устройство команд,
- 3) арифметическое устройство,
- 4) буферное запоминающее устройство для хранения промежуточных результатов и предварительной подготовки операндов,
- 5) регистр внутренних прерываний.

Основной функцией устройств обмена (УО) является обеспечение связи ВС со всеми источниками и потребителями информации. Прием информации от внешних источников, а также выдача инфор-

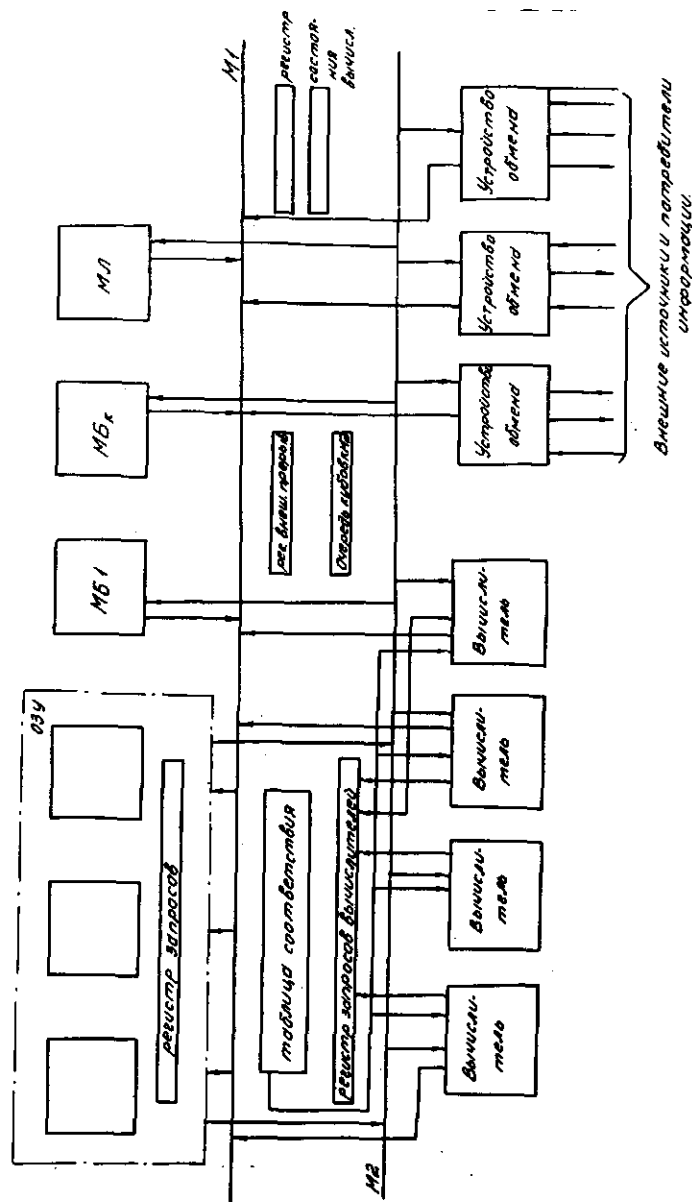


Рис. 1. Блок-схема вычислительной системы

мации потребителям производится параллельно с работой вычислителей по основной программе. Всю информацию, которую нужно выдать внешним потребителям, вычислители записывают в ОЗУ и сообщают устройствам обмена о ее готовности. Информацию, вводимую от внешних источников УО, также записывают в ОЗУ. Основными узлами УО являются:

- 1) регистр обмена,
- 2) регистр готовностей программ к обмену с внешними источниками,
- 3) регистр готовности внешних источников к обмену,
- 4) устройство управления.

Общее информационное поле в ВС состоит из группы МЛ, группы МБ и ОЗУ. В памяти на МЛ записаны все программы задач, константы и необходимые массивы информации. При вводе задачи для решения ее информационные массивы также вызываются во внутреннюю память. Внутренней памятью ВС являются ОЗУ, разделенное на отдельные секции, и группа МБ. В процессе работы ВС осуществляется динамическое распределение оперативной памяти.

Кодовые шины, предназначенные для связи вычислителей и устройств обмена с общим информационным полем, позволяют:

а) сократить оборудование по сравнению с необходимостью иметь самостоятельные цепи для связи каждой секции памяти с каждым вычислителем,

б) легко изменять состав и количество оборудования в ВС.

В рассматриваемой ВС используются две кодовые магистрали: одна для передачи информационного потока и адресов от вычислителей, МБ, МЛ и УО к секциям ОЗУ (М1), другая — для передачи информации из оперативной памяти запрашивающим устройствам (М2).

Устройство управления ВС содержит:

- 1) устройство управления кодовыми магистралями М1 и М2,
- 2) регистр состояния вычислителей,
- 3) таблицу распределения оперативной памяти между решаемыми задачами,
- 4) регистр прерывания ВС.

Устройство управления кодовой магистралью М1 является одновременно приоритетным регистром запросов к ОЗУ от всех устройств, которые могут обращаться к ОЗУ.

Управлением магистралью М2 является приоритетная очередь

секций памяти на передачу информации запрашивающим устройствам.

Вычислители, входящие в ВС, могут находиться в процессе работы ВС в следующих состояниях:

- 1) вычислитель свободен,
- 2) вычислитель выполняет программу и допускает прерывания,
- 3) вычислитель выполняет диспетчерскую программу и допускает только внутренние прерывания.

Состояния всех вычислителей в системе в каждый момент времени отражаются на регистре состояний вычислителей и используются при выборе вычислителя для выполнения прерывания ВС. Для реакции на прерывание назначается либо любой свободный вычислитель, либо вычислитель, допускающий прерывание.

С целью сокращения разрядов в команде и тем самым объема памяти, необходимой для хранения программ, в разных задачах используются повторяющиеся математические адреса страниц. Информация, вводимая в ОЗУ в процессе выполнения программы, может размещаться на любых свободных страницах. Для учета размещения информации в ОЗУ и преобразования математических адресов в физические в состав ВС вводится таблица соответствия между математическими адресами программ, выполняемых в настоящее время на вычислителях, и соответствующими физическими адресами их в ОЗУ. Полная таблица размещения программных и информационных массивов всех включенных задач в ОЗУ и на МБ доступна диспетчерской программе (ДП) и помещается в ОЗУ.

Блок-схема управления ВС приведена на рис.2.

Основными функциями диспетчерской программы в мультипрограммной ВС с памятью нескольких уровней является:

- организация обмена массивами информации между разными уровнями памяти,
- учет использования страниц ОЗУ и организация динамического распределения памяти,
- организация решения задач,
- решение конфликтных ситуаций при обмене с внешними источниками информации,
- организация реакции на сбой и неисправности различных устройств ВС.

Рассмотренная структура ВС является наиболее гибкой с точки зрения связи вычислителей друг с другом. Такая структура

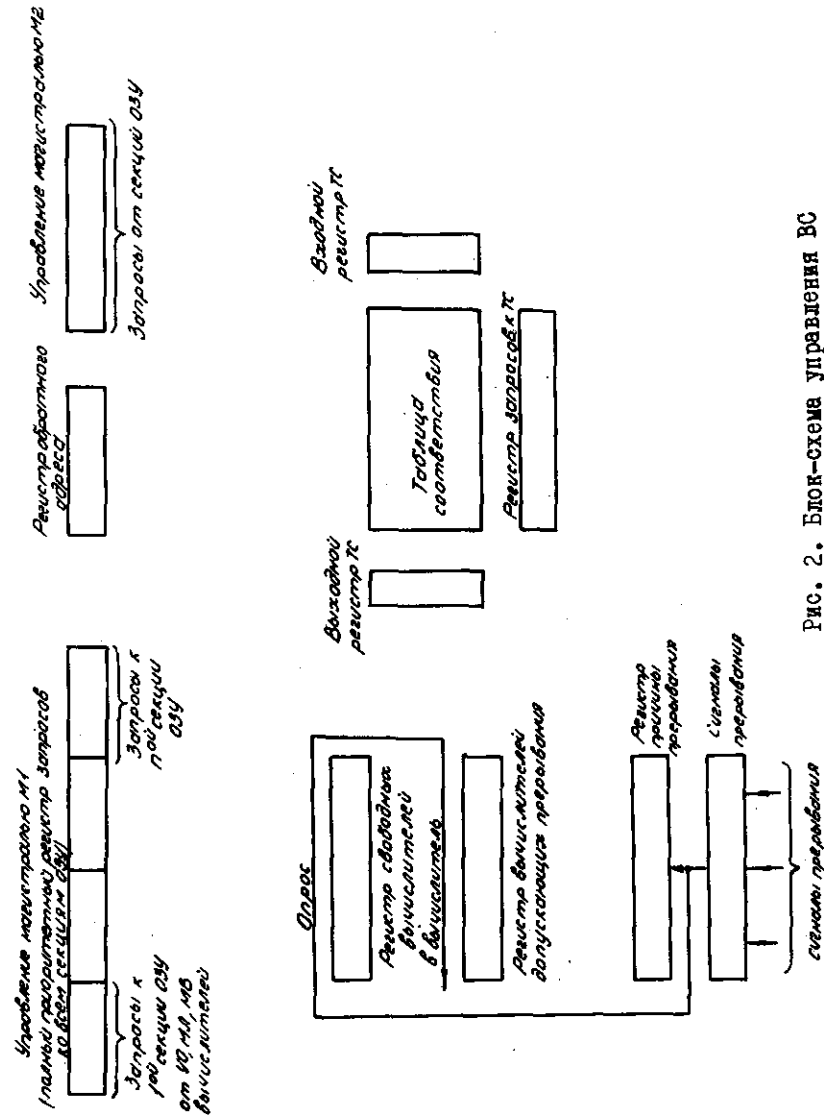


Рис. 2. Блок-схема управления ВС

позволяет легко распределять задачи между вычислителями, перераспределять их после прерывания, совмещать операции обмена с внешними источниками с решением задач.

Организация обмена между ОЗУ и памятью на МБ

При организации обмена между ОЗУ и памятью на МБ прежде всего следует исходить из условия: влияние невысокого быстродействия запоминающих устройств на МБ на время выполнения программ должно быть сведено к минимуму. По этой причине для хранения на МБ назначается информация, не используемая в процессе выполнения программ в течение сравнительно длительных отрезков времени. Групповой характер обмена между ОЗУ и МБ позволяет снизить удельные потери времени на ожидание, приходящиеся на каждое слово, по которому производится обмен. Для организации обмена необходимо иметь следующую служебную информацию:

- 1) таблицу размещения информации в ОЗУ,
- 2) таблицу свободных страниц в ОЗУ,
- 3) таблицу размещения информации в памяти на МБ.

Первой таблицей приходится пользоваться при исполнении каждой команды, поэтому она реализована как схемная таблица соответствия (ТС). Обращение ко 2-й и 3-й таблицам производится при организации обмена между ОЗУ и МБ, т.е. значительно реже, поэтому эти таблицы хранятся в ОЗУ. Обмен между ОЗУ и МБ выполняется массивами стандартной длины — страницами. Постраничная структура памяти позволяет за счет некоторого схемного усложнения осуществить такую организацию обменов, когда необходимость обмена выявляется автоматически в процессе исполнения программы. Автоматическое выявление окончания использования страниц, упрощая программирование и в некоторой степени облегчая организацию мультипрограммной работы, в то же время снижает быстродействие ВС за счет частого включения диспетчерской программы и ухудшает использование объемов ОЗУ. Ухудшение использования объемов ОЗУ связано с тем, что каждая страница должна находиться в ОЗУ не только в течение времени, которое требуется для использования ее программой, но и в течение времени, необходимого для выявления окончания ее использования. Это время в сред-

нем равно полутора интервалам проверки использования страниц. Во избежание излишних задержек при исполнении программы желательно, чтобы к моменту запроса на ввод страницы была хотя бы одна свободная страница в ОЗУ, готовая к приему информации. Для этого после заполнения информацией последней свободной страницы включается подпрограмма диспетчерской программы, организующая отсылку выбранной заранее страницы на МБ. Для выбора страницы, предназначенной для отсылки из ОЗУ на МБ, используется одна из модификаций аналогичного алгоритма, примененного в английской машине "Атлас". Алгоритм основан на изучении характера обращения к страницам в процессе исполнения программы и прогнозирования на основе этой информации предполагаемого характера обращения к тем же страницам в будущем.

Выбор интервала проверки использования страниц должен производиться с учетом следующих противоречащих друг другу факторов:

увеличение интервала проверки приводит к относительному снижению потерь быстродействия ВС, связанному с необходимостью после окончания каждого интервала включать программу учета характера использования страниц,

с другой стороны, увеличение интервала проверки приводит к ухудшению использования объемов ОЗУ.

Триггеры использования страниц в ТС в начале каждого интервала проверки обрасываются в 0. Страница, к которой на этом интервале было обращение, считается используемой, ее триггер использования устанавливается в 1. Совокупность триггеров, относящихся ко всем страницам, образует регистр, на котором зафиксирован код, характеризующий использование всех страниц в течение одного интервала проверки. В конце каждого интервала проверки этот код поразрядно сопоставляется с хранящимся в ОЗУ аналогичным кодом, выработанным в конце предыдущего интервала проверки, после чего сам переписывается в ОЗУ.

В процессе разработки описанной ВС было произведено моделирование работы ДП, в частности работы ее подпрограмм, связанных с организацией динамического распределения ОЗУ с автоматической организацией обмена между ОЗУ и МБ. Моделировалась работа ВС, состоящей из 3 и 4 вычислителей с общим быстродействием 300 тыс. и 400 тыс. коротких операций в сек. Частота про-

верки использования страниц была выбрана 0,02 сек, объем ОЗУ — 16 тыс. ячеек. Моделирование выполнялось на задачах с различной цикличностью решения и использующих разное количество массивов. В результате моделирования выявились значительные трудности, возникающие при решении задач, требующих довольно высоких частот повторения. По результатам моделирования оказалось, что на работу ДП тратится 13% быстродействия ВС из 4-х вычислителей и 17% при 3-х вычислителях, причем основное время затрачивается на выбор страницы для вытеснения на МБ и учет использования страниц. Можно добиться снижения потерь быстродействия в случае программной организации отсылки страниц из ОЗУ на МБ. При этом с помощью команд, включаемых в программу программистом, организуется отсылка страниц, которые больше не понадобятся, и обращение к подпрограммам, заполняющим списки страниц, в использовании которых предполагается перерыв. Информация о предполагаемой длительности интервалов должна сообщаться программистом. Эти списки используются диспетчерской программой, если при необходимости ввести в ОЗУ не окажется свободных страниц. Возможны комбинированные способы, сочетающие программную и автоматическую организацию обмена.

Логическая структура таблицы соответствия

Записываемые в командах программы математические адреса данных, а также математические адреса самих программ не отражают фактического расположения в памяти. Связь между математическими и физическими адресами страниц устанавливается ДП в процессе динамического распределения памяти. Поскольку математические адреса страниц нужно преобразовывать в физические при исполнении каждой команды, это преобразование надо выполнять самым быстрым способом, т.е. аппаратным путем. С целью экономии объема программ в задачах используются повторяющиеся математические адреса. Для страниц, в которых имеется информация, используемая несколькими программами, отводятся номера, не повторяющиеся ни в одной из задач. Для различения страниц разных задач, имеющих одинаковые математические номера, используется номер вычислителя, на котором в данный момент решается задача.

Таблица соответствия отражает распределение ОЗУ по вычислителям в каждый момент времени. Логическая структура таблицы соответствия, выполненной на триггерных регистрах, показана на рис.3. Число строк в ТС равно числу страниц в ОЗУ. Регистры с 1 по 72 содержат математические адреса страниц, записанных в ОЗУ. Физический адрес страницы, т.е. фактический номер секции и страницы в ОЗУ, формируется на шифратор и выдается на выходной регистр ТС. Запросы на обращение к ТС собираются на приоритетном регистре запросов. Математический адрес ТС, по которому находится соответствующий физический адрес, принимается на входной регистр ТС. Обращение к ТС по своему характеру ассоциативное. Для реализации ассоциативного поиска на каждом регистре математического адреса имеется многоадресная схема совпадения адреса, записанного в данной строке ТС с математическим адресом на входном регистре ТС. В каждой строке ТС также указан номер вычислителя (или вычислителей), которому разрешено обращение к данной странице. Разряды в регистрах математических номеров, указывающие номер вычислителя, объединены в регистры-ключи вычислителей. При включении новой задачи для решения ДП засылает на регистр ключа этого вычислителя новое значение. По окончании решения задачи на вычислителе его регистр ключа устанавливается в 0. Ключи одновременно выполняют функцию защиты памяти, так как никакой другой вычислитель не может обратиться в страницы, разрешенные этому вычислителю. Динамическое распределение памяти в процессе выполнения программ требует учета использования страниц, записанных в ОЗУ. С этой целью у каждой страницы ОЗУ есть триггер использования, присланный к каждой строке ТС. Содержимое триггеров использования анализируется ДП.

Есть основание полагать, что адреса в основном будут идти последовательно и смена страниц в процессе исполнения программ будет происходить нечасто. Справедливость такого предположения очевидна для программ и в большом количестве случаев для данных. Для сокращения числа обращений в ТС и уменьшения времени простоя запроса в очереди для каждого абонента в вычислителе, обращающегося в ТС, вписана строка ТС в специально введенный для этого регистр. В связи с тем, что программа и операнды располагаются в разных страницах, вводятся разные табличные регистры для операндов и команд. Наличие в ВС схемной ТС

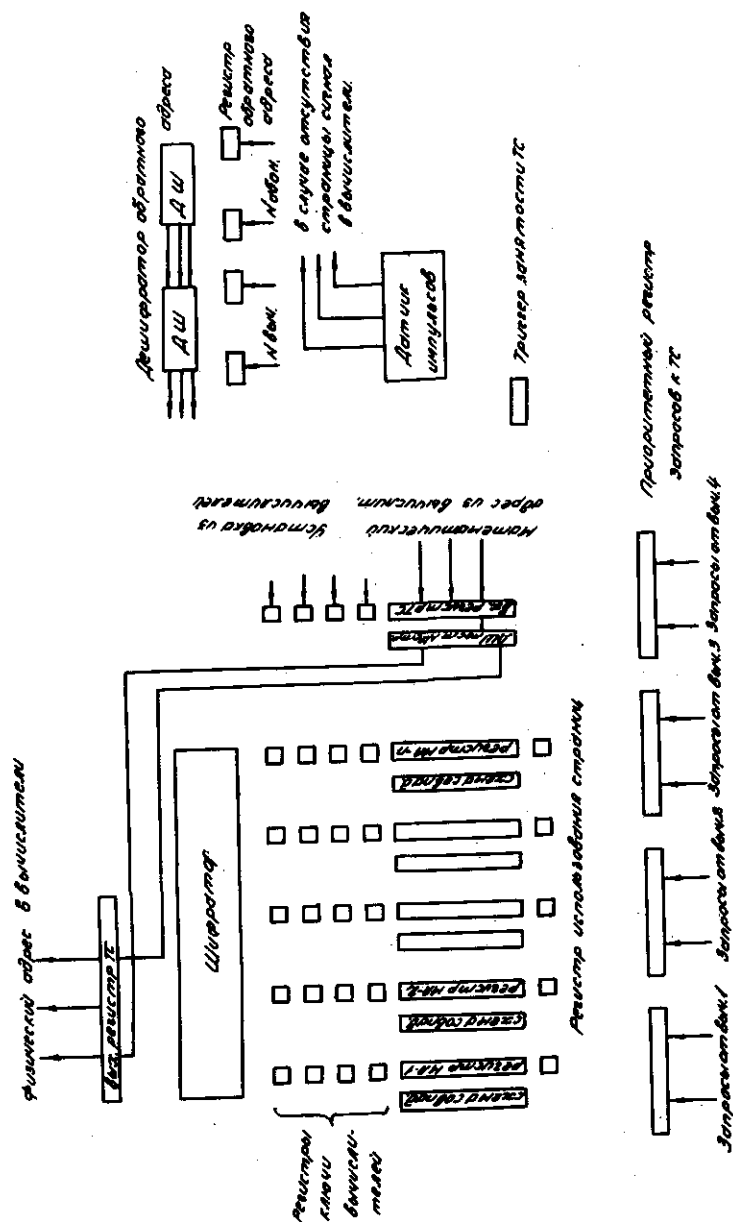


Рис. 3. Логическая структура таблицы соответствия

позволяет повысить ее быстродействие по сравнению со случаем выполнения ее функций программными средствами.

Расчет производительности ВС

Производительность ВС, состоящей из нескольких вычислителей, можно определить, используя методы моделирования и аналитического расчета ВС как системы массового обслуживания. Моделирование всей ВС в целом весьма громоздко и практически трудно осуществимо. Ввиду того, что потоки запросов от отдельных вычислителей к общесистемному оборудованию (линейные магистрали, ТС) независимы друг от друга, расчет задержки, вносимой общесистемным оборудованием, можно производить аналитически с использованием аппарата цепей Маркова, если потоки запросов от каждого вычислителя известны. Отдельные факторы, влияющие на поток запросов различных вычислителей, сильно коррелированы друг с другом и, следовательно, трудно поддаются аналитическому расчету. Поэтому для определения потоков запросов от вычислителей к общесистемному оборудованию используется метод моделирования.

Моделирование вычислителя проводилось на реальных задачах. Основной задачей моделирования было получение выходных характеристик вычислителя в ВС: потока обращения к ОЗУ, к магистралям М1 и М2 и ТС. В процессе моделирования были определены также быстродействие вычислителя, входящего в состав ВС, загрузка его основных устройств, простои, а также определены необходимые размеры буфера чисел и команд. В результате моделирования была определена эффективность буфера чисел и команд задач с точки зрения сокращения количества обращений в ОЗУ.

Источниками снижения производительности ВС являются столкновения запросов в общесистемном оборудовании:

- у магистралей М1 и М2,
- у секций ОЗУ,
- у ТС.

Систему "вычислители-магистраль" можно представить в виде, показанном на рис.4. Будем рассматривать состояние системы "вычислители-магистраль М1 (М2)", предполагая, что нахождение со-

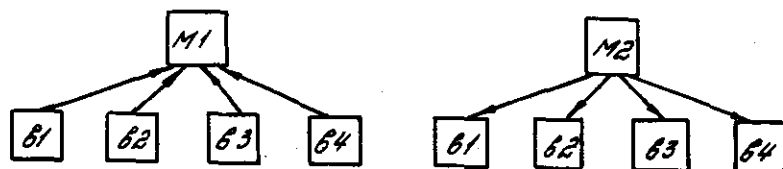


Рис.4. Система "вычислители-магистраль M1 (M2)"

остояния возможно с временными интервалами в один такт. Для расчета задержек, связанных с занятостью M1 (M2), применяется аппарат цепей Маркова. Состояния системы "вычислители-магистраль" описываются следующим образом:

"0" — в магистраль нет ни одного запроса, магистраль не работает. Вероятность этого состояния P_0 .

"1" — имеется один запрос от вычислителя (вычислитель не работает) — магистраль работает. Вероятность этого состояния P_1 .

"2" — два вычислителя послали запрос к M1, M1 обслуживает один запрос. Вероятность этого состояния P_2 .

"3" — три вычислителя послали запрос, магистраль работает. Вероятность этого состояния P_3 .

"4" — четыре вычислителя послали запрос, один из которых обслуживается магистралью. Вероятность этого состояния P_4 .

Причинами перехода системы из одного состояния в другое являются:

освобождение магистрали от выполняемых запросов,
поступление новых запросов.

Если магистраль занята исполнением запроса от некоторого вычислителя, то она с вероятностью, равной 1, освобождается от него в конце такта. Выполнение этого условия не зависит от времени поступления этого и других запросов, т.е. от предистории. Вероятность поступления новых запросов от каждого вычислителя фактически зависит от того, каким образом поступали запросы от этого вычислителя в течение нескольких последних тактов. Вслед-

ствие такой зависимости поток запросов к магистрали от каждого вычислителя, строго говоря, не является пуассоновским. Однако суммарный поток от четырех вычислителей вследствие независимости этих потоков друг от друга приближается к пуассоновскому. Это дает нам право без большой погрешности считать суммарный поток от четырех вычислителей пуассоновским, т.е. не зависящим от предистории. Следовательно, система, в которой состояния определены указанным выше образом, в случае принятия общего потока заявок от четырех вычислителей к магистрали пуассоновским, является цепью Маркова. Нас интересуют вероятности нахождения системы в указанных выше состояниях.

Пусть:

p — вероятность того, что от вычислителя в данном такте поступил запрос к M1,

$(1-p)$ — вероятность того, что в данном такте запрос не поступил.

По результатам моделирования для M1 $p = 0,93$, для M2 $p = 0,083$.

P_i — вероятность нахождения системы в i -м состоянии после окончания переходного процесса,

P_{ij} — вероятность перехода из i -го состояния в j -ое в течение данного такта.

Для вероятностей состояний системы справедливо:

$$P_0 [(1-p)^4 - 1] + P_1 (1-p^3) = 0,$$

$$P_0 [4p(1-p^3)] + P_1 [3p(1-p)^2 - 1] + P_2 (1-p)^2 = 0,$$

$$P_0 6p^2(1-p)^2 + P_1 3p^2(1-p) + P_2 [2p(1-p) - 1] + P_3 (1-p) = 0,$$

$$P_0 4p^3(1-p) + P_1 p^3 + P_2 p^2 + P_3 (p-1) + P_4 = 0,$$

$$P_0 p^4 - P_4 = 0,$$

$$P_0 + P_1 + P_2 + P_3 + P_4 = 1.$$

После подсчета вероятностей состояний рассчитывается за -

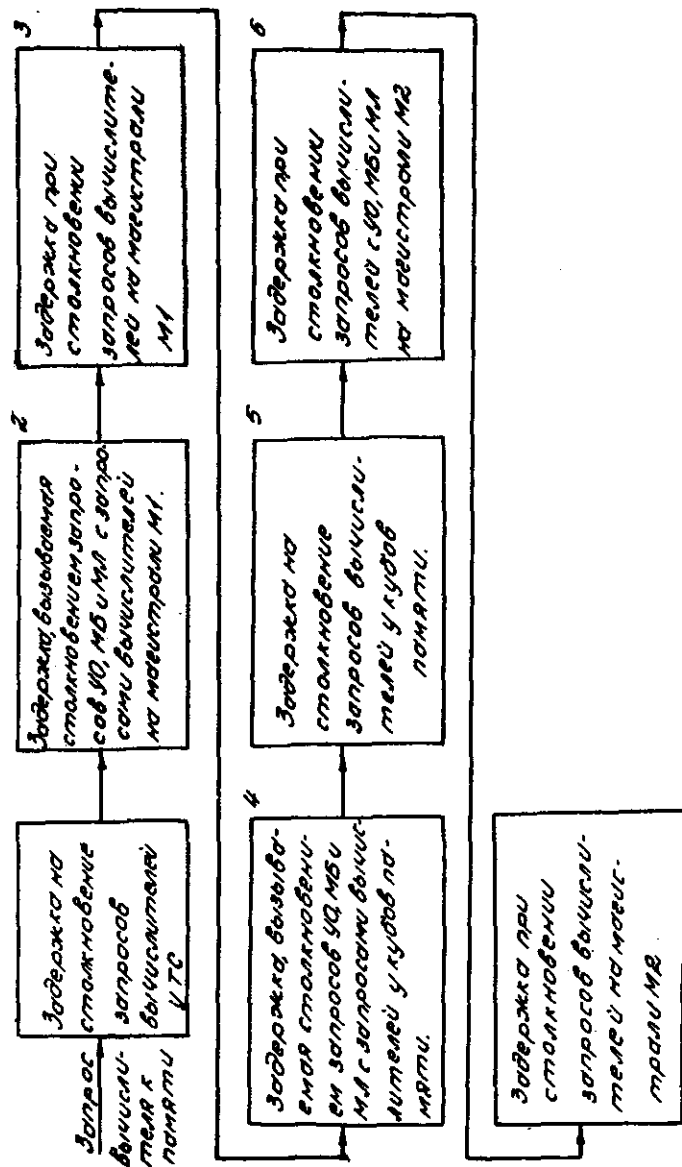


Рис. 5. Суммарная задержка запросов вычислителей в ВК с общим полем памяти

держка, вносимая столкновением запросов вычислителей на магистральных М1 и М2. При расчете задержек, вызываемых столкновением запросов вычислителей и ТС, используем тот же метод и той же матрицей переходов. Максимальное значение вероятности занятости ТС при обращении вычислителя по результатам моделирования равно $p = 0,0736$. Далее определяются вероятности состояний и подсчитывается задержка, вносимая столкновением запросов у ТС. Задержка, вносимая взаимными столкновениями вычислителей у кубов памяти, вычисляется также рассмотренным способом.

На рис.5 показана схема образования суммарной задержки запросов вычислителей в ВК при работе с общим полем памяти.

В результате всех задержек суммарная производительность ВК, состоящей из четырех вычислителей, уменьшается на 17,3%.