

УДК 681.142.4

НЕКОТОРЫЕ ВОПРОСЫ ОРГАНИЗАЦИИ  
 ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ ВЫСОКОЙ ПРОИЗВОДИТЕЛЬНОСТИ

А.П. Заморин

Рассматриваются проблемы организации многопроцессорных систем на базе высокопроизводительных моделей электронно-вычислительных машин ЭВМ третьего поколения. Приводятся моделирование процессов обмена на уровне общего поля оперативной памяти.

I. Общая проблематика

Понятие "высокопроизводительная" ЭВМ является в достаточной степени условным. Очевидно, имеет смысл использовать это понятие в том случае, когда возникает качественно новый подход к разработке структур ЭВМ. Такой границей может быть условие

$$T_{np} < T_{on},$$

т.е. средний период повторения запросов со стороны процессора  $T_{np}$  меньше цикла обращения оперативной памяти, многоуровневое совмещение операций, промежуточная память и т.п.

При создании вычислительных систем на основе подобных ЭВМ необходимо рассматривать достаточно широкий круг проблем с тем, чтобы вновь созданная система была достаточно эффективной. Оцен-

ку можно вести учетом ряда следующих факторов: надежности системы, степени использования оборудования, себестоимости вычислений, реакции системы, адаптивности и т.д.

Надежность вычислительной системы определяется, во-первых, наличием дублирования каждого из устройств системы и, во-вторых, возможностью быстрой локализации неисправности работающим устройством. Надежность системы, в которой возможен частичный отказ устройств, определяется выражением

$$P_c = 1 - R \cdot h^{-s}, \quad (1)$$

где  $P_c$  - вероятность безотказной работы системы;

$R$  - вероятность отказа системы, в которой отсутствует дублирование;

$h$  - число уровней, на которых возможно дублирование;

$s$  - допустимое число одновременно отказавших устройств на каждом уровне.

Таким образом, требуемая надежность может быть достигнута увеличением числа уровней дублирования  $h$ , наличием "нагруженного" резервирования  $s$ , а также увеличением интенсивности восстановления  $\nu$ .

Практически полное дублирование на каждом из уровней по-лучить трудно, однако это условие должно быть положено в основу структуры системы.

Для большой системы, включающей множество различных устройств, важным фактором эффективности является интенсивность их использования. При этом в сравнении с отдельной ЭВМ может быть принята большая степень загрузки устройств.

Коэффициент возможного "сжатия" определяется соотношением:

$$\gamma = \left( \sum_{i=1}^k g_i \right)^{-2} \cdot \sum_{i=1}^k g_i^2, \quad (2)$$

где  $g_i$  - число однородных устройств в отдельной ЭВМ;

$k$  - степень комплексирования, количество объединяемых ЭВМ.

Что касается снижения себестоимости вычислений, то в качестве оценки можно использовать известный закон Гроша: увеличение производительности системы должно быть пропорционально квадрату увеличения ее стоимости.

Малая реактивность системы означает высокую скорость обработки прерываний, возможность реализации широкого диапазона приоритетов. Практически это определяется наличием многих параллельно работающих устройств обработки.

Под адаптивностью системы понимается возможность программной или аппаратной перестройки системы в соответствии со спецификой решаемой задачи.

Все указанные выше требования могут быть разрешены при организации системы по принципу полноты связей на уровне каждого из устройств отдельной ЭВМ, входящей в данный комплекс.

Полнота связей предполагает выполнение таких условий, как:  
о д н о р о д н о с т ь - функциональная идентичность устройств одного вида; возможность выполнения заданной процедуры любым из устройств данного вида;

а в т о н о м н о с т ь - взаимная независимость функционирования; возможность удаления устройства без нарушения режима функционирования системы;

д о с т у п н о с т ь - каждое устройство одного вида доступно, либо имеет доступ к любому из устройств другого вида.

Очевидно, что полная реализация данного принципа возможна только в полностью однородной системе. Тем не менее эффективность системы будет во многом определяться достигнутой степенью приближения к данному принципу.

## 2. Структура вычислительной системы

Основным условием при разработке структуры системы было требование использования типового набора устройств, имеющих в архитектуре ЭВМ третьего поколения, без их существенного изменения, сохранение принятой идеологии, системы команд, стандартного сопряжения и т.п.

Для организации связи на уровне внешних устройств могут быть использованы имеющиеся средства связи с несколькими каналами.

Что касается общего поля оперативной памяти, то проблема разрешения конфликтов при обращении многих процессоров и каналов остается достаточно сложной.

Целью настоящей работы был поиск оптимальной структуры, позволяющей разрешить данную проблему.

Очевидно, что данный поиск возможен как в направлении расширения пропускной способности системы обмена с оперативной памятью (увеличение степени расслоения, выборка нескольких слов по одному запросу и т.п.), так и более оптимальной организации процесса обмена.

В качестве основы был выбран распределительный коммутатор, дающий более простое аппаратное решение. Последующий анализ и результаты моделирования показали возможность построения коммутатора, распределенного по источникам запросов.

К разрабатываемому коммутатору предъявлялись следующие требования:

м о д у л ь н о с т ь - возможность разбиения на функционально идентичные узлы, автономно сопрягаемые с каждым источником запросов;

ф у н к ц и о н а л ь н о с т ь - полное разрешение конфликтов обращения к оперативной памяти; при минимальной доработке устройств;

у н и в е р с а л ь н о с т ь - возможность сопряжения с несколькими видами устройств.

Общий вид системы представлен на рис.1. Каждый процессор, канал или группа каналов связаны со своим модулем коммутатора. Каждый модуль имеет свой позиционный приоритет, определяемый приоритетом связанного с ним устройства. Соседние модули коммутатора соединены шинами разрешения обращения и шинами запросов.

Устройство сопряжения стандартной системой внешнего прерывания, прямого управления и управления каналами ввода-вывода имеет также модульное исполнение.

На рис.2. представлена функциональная схема коммутатора (один модуль). Она содержит ряд схем сравнения (по числу независимых блоков памяти) и такое же количество схем анализа запросов.

Схема сравнения может быть одноразрядной, если имеются приоритеты только двух уровней либо ее разрядность соответствует разрядности кода запроса. на схему поступает код запроса от источника и от устройства более низкого приоритета. Если при-

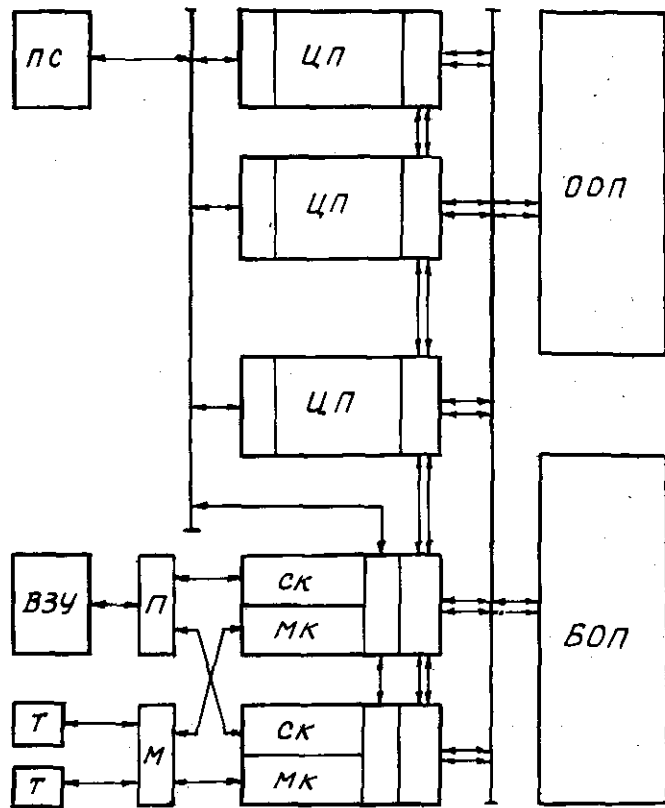


Рис. 1. Структура системы с полными связями

оритет первого выше или равен второму, то запрос поступает далее на схему анализа. Код запроса со старшим приоритетом поступает со схемы сравнения на устройство более высокого позиционного приоритета.

Схема анализа сравнивает поступивший запрос с сигналом разрешения (незанятости) с устройства более высокого позиционного приоритета. При возможности запроса организуется обращение и одновременно выдается сигнал занятости данного блока на

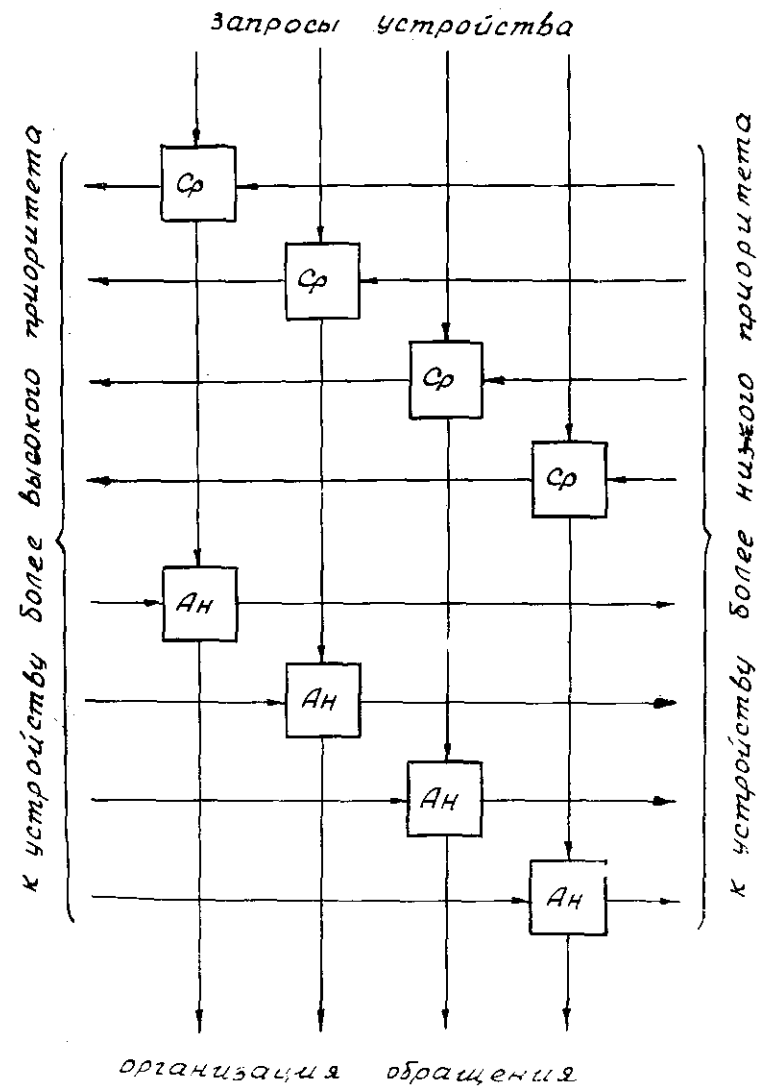


Рис. 2. Коммутатор запросов

устройство с более низким позиционным приоритетом.

Доработка устройств, посылающих запросы, сводится в этом случае к обеспечению "раскраски" запросов номером блока памяти и приоритетом запроса.

### 3. Модель общего поля памяти

Для вычисления системы высокой производительности характерно, что средняя частота запросов процессора к оперативной памяти превышает частоту выборки слов из отдельного блока памяти. С целью выравнивания потенциала оперативной памяти и процессора применяется деление памяти на отдельные блоки с возможностью независимого обращения (расслоение памяти). Запрос на обращение к памяти идет с опережением. Необходимое число предварительно выбираемых машинных слов с учетом конфликтных ситуаций при произвольном порядке выборки определяется соотношением

$$q = -N \cdot M^{-1} \cdot \ln(1 - K_{gon}), \quad (3)$$

где  $q$  - число предварительных запросов (глубина предвыборки);

$N$  - число независимых блоков памяти;

$M$  - количество устройств, посылающих запросы;

$K_{gon}$  - допустимый коэффициент снижения производительности памяти.

Производительность памяти при этом определится соотношением

$$W_{on} = T_{on}^{-1} \cdot K_{gon} \cdot N = T_{on}^{-1} \cdot N(1 - e^{-\frac{q}{N}}), \quad (4)$$

где  $T_{on}$  - цикл оперативной памяти.

Количество независимых блоков памяти определяется из условия динамического равновесия: потенциал памяти должен соответствовать интенсивности запросов со стороны процессора и каналов ввода-вывода.

$$N > \frac{M \cdot T_{on}}{T_{np} \cdot K_{gon}} \left(1 + \frac{1}{\alpha}\right), \quad (5)$$

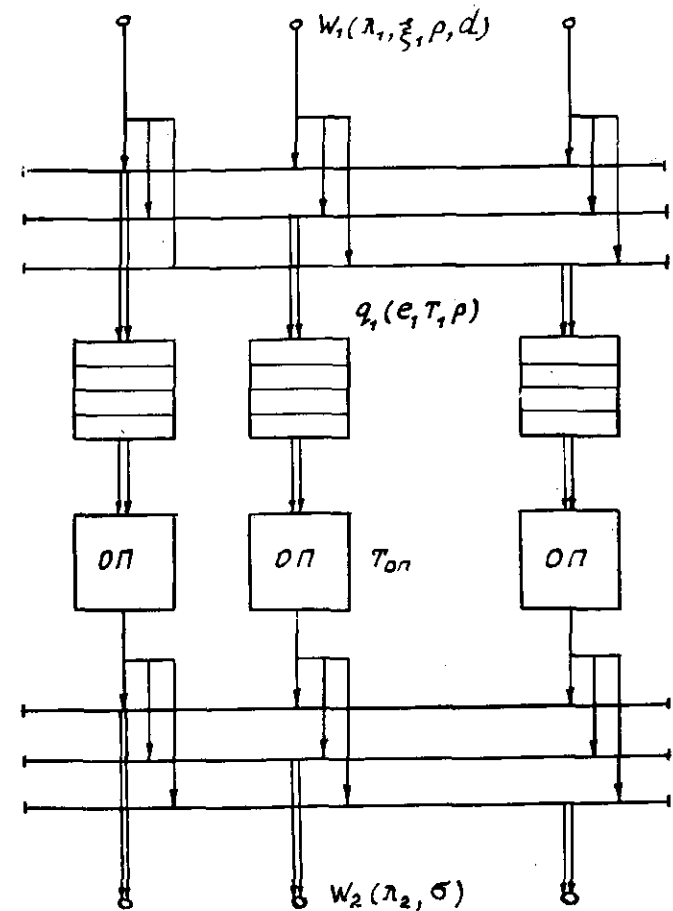


Рис. 3. I вариант общего поля памяти

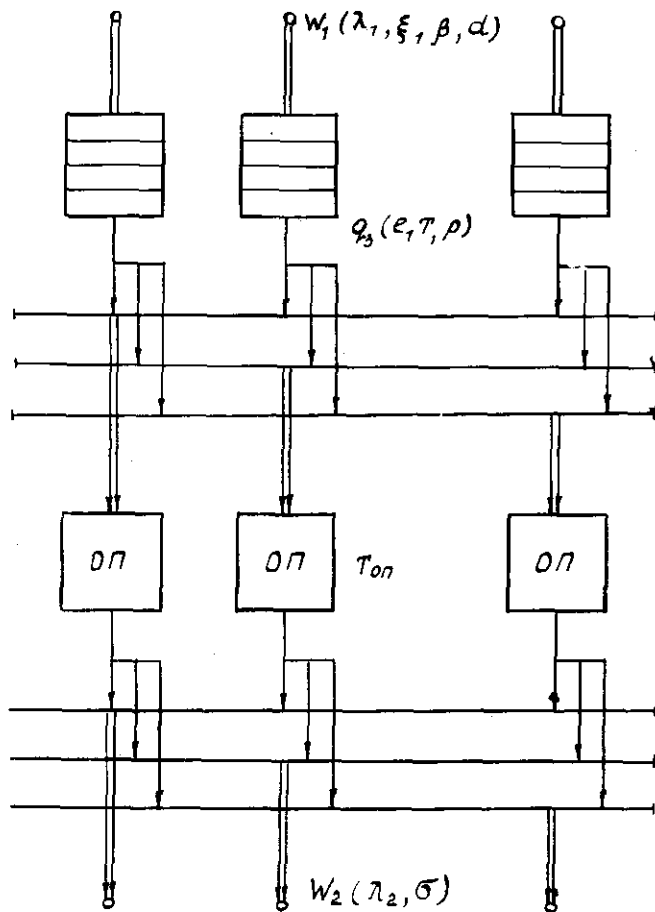


Рис. 4. II вариант общего поля памяти

где  $\alpha$  - коэффициент активности информации - среднее число обращения памяти, приходящееся на одно слово, вводимое в оперативную память.

Использование предвыборки предполагает наличие буферных устройств (магазинов) для хранения очереди запросов.

Возможны две схемы организации буферных устройств:

I вариант (рис.3) - очереди запросов организуются на входе каждого блока оперативной памяти;

II вариант (рис.4) - очереди запросов организуются на выходе каждого из источников запросов.

Схема варианта I имеет некоторые, вполне очевидные достоинства: простоту организации асинхронной работы, возможность реализации широкого диапазона приоритетов. Из наиболее явных недостатков можно отметить большой объем оборудования, связанный с необходимостью запоминания полноразрядных посылок, невозможность управления очередью запросов, уязвимость схемы.

Схема варианта II не содержит недостатков, свойственных схеме варианта I, однако и лишена достоинств упомянутой схемы. Практически сложно организовать асинхронное обращение и реализовать весь диапазон приоритетов.

Представляет интерес сравнение обеих схем с точки зрения пропускной способности и объема оборудования.

Входной поток  $W_1$  определяется средней интенсивностью  $\lambda_1$ , видом запроса  $\xi$ , приоритетом исполняемой программы  $\beta$ , характером последовательности запросов  $\alpha$  (случайной или детерминированной).

Очередь запросов  $Q$  определяется средней длиной  $\ell$ , средним ожиданием  $T_{ож}$ , дисциплиной обслуживания  $\rho$ . Порядок обслуживания устанавливается по приоритету запроса, который в общем случае определяется соотношением

$$\rho = b_1 \cdot i + b_2 \beta + b_3 \Delta t + b_4 \Delta \ell + b_5 \xi \quad (6)$$

где  $i$  - номер устройства (позиционный приоритет);

$\beta$  - программный приоритет;

$\Delta t$  - динамический приоритет;

$\Delta \ell$  - порядковый приоритет (зависит от места запроса в очереди данного источника);

$\xi$  - вид запроса (видовой приоритет);

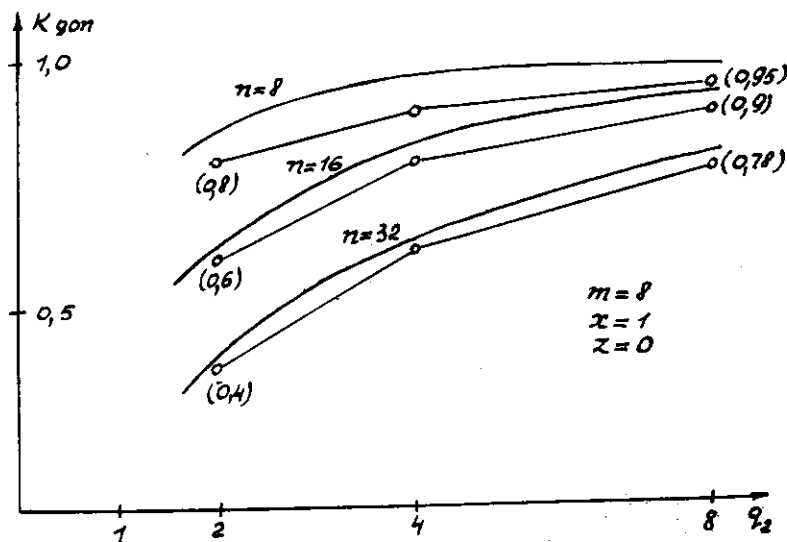
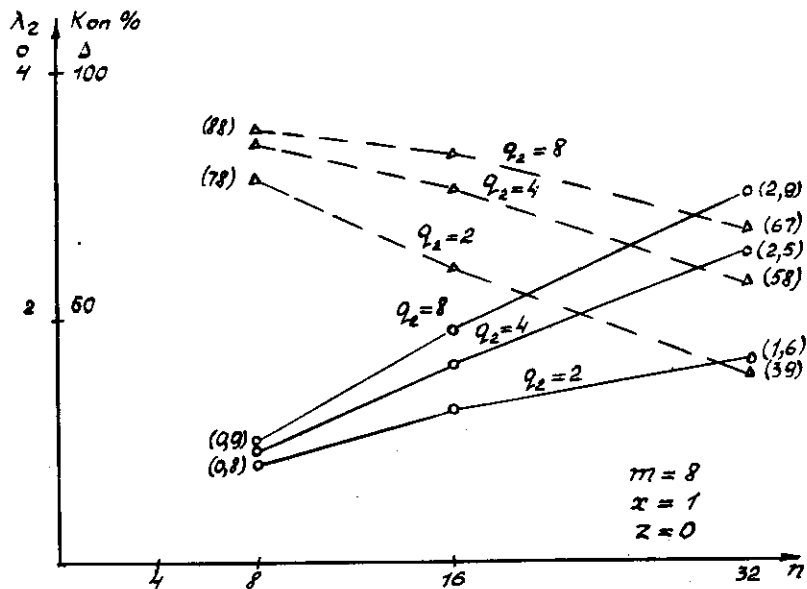


Рис. 5. Влияние глубины предвыборки ( $q_2$ ).

$b_1 - b_2$  - коэффициенты (веса) приоритетов.

Выходной поток  $W_2$  характеризуется средней интенсивностью  $\lambda_2$  и дисперсией рассеивания потока  $\sigma$ . Под "рассеиванием" потока понимается степень его неупорядоченности, вызванной произвольным порядком обслуживания запросов.

#### 4. Результаты моделирования

Целью моделирования была проверка эффективности расслоения при различной интенсивности запросов, влияния емкости буферных устройств, использования группового расслоения, режима работы устройств (синхронного и асинхронного), различной дисциплины обслуживания.

Результаты моделирования приведены на рис. 5 - 8. На всех рисунках используются общие обозначения, принятые в алгоритме моделирования:

$m$  - число источников запросов;

$n$  - число блоков памяти;

$q_1$  - емкость буферной памяти блока памяти;

$q_2$  - глубина предвыборки;

$q_3$  - емкость буферной памяти источника;

$x$  - характер последовательности адресов запросов источника (1 - случайная, 0 - детерминированная);

$z$  - режим работы устройств (1 - синхронный, 0 - асинхронный);

$\lambda$  - средняя интенсивность запросов (число запросов на цикл обращения блока памяти);

$K_{оп}$  - коэффициент использования блоков памяти;

$T_{ож}$  - среднее время ожидания;

$\sigma_{рас}$  - дисперсия рассеивания потока запросов от отдельного источника;

$z$  - число блоков памяти, используемых для расслоения.

На рис. 5 можно проследить влияние глубины предвыборки ( $q_2 = 2, 4, 8$ ) на результаты использования расслоения. Пунктирной линией показан процент использования блоков памяти.

Внизу дается сравнение значений  $K_{гон}$ , вычисленных согласно выражению (4) и полученных в результате моделирования.

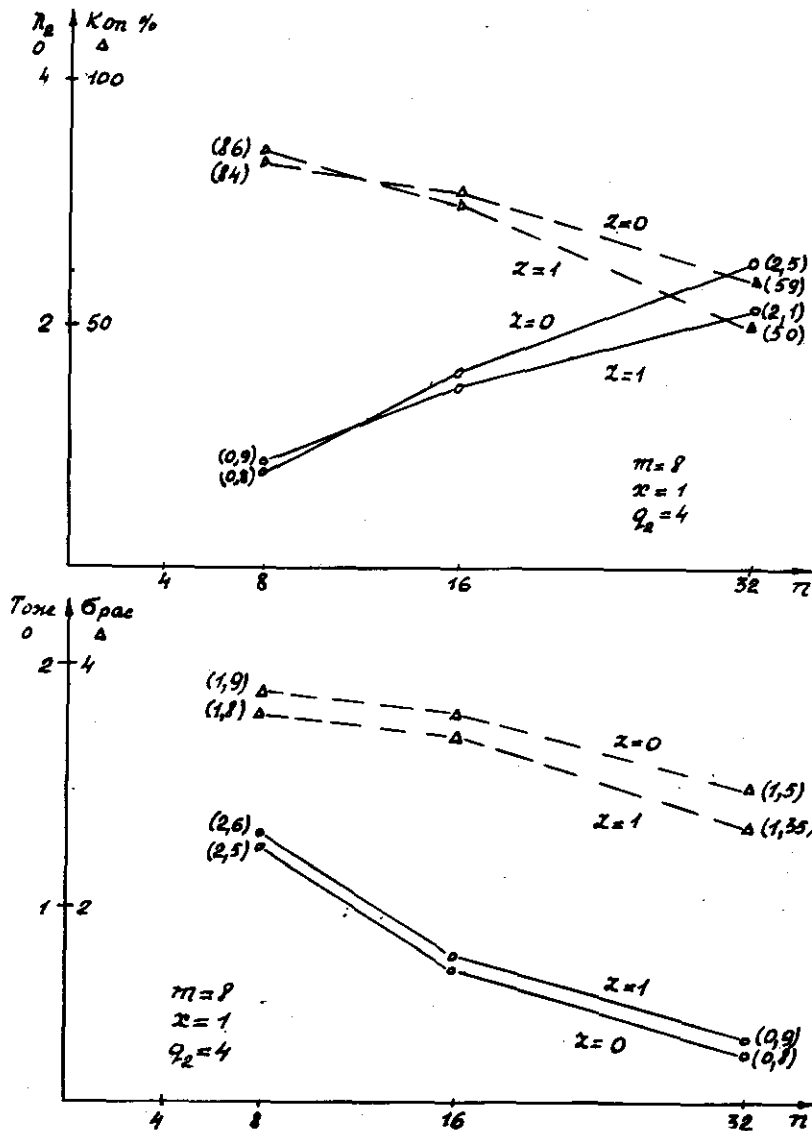


Рис. 6. Влияние режима работы ( $\alpha$ ).

На рис. 6 приводятся сравнение результатов моделирования для синхронного и асинхронного режимов работы устройств. Синхронный режим подразумевает жесткую временную диаграмму обращения к различным блокам памяти. Как следует из рис.6 (вверху), различие между двумя режимами невелико при высокой интенсивности запросов. Внизу, где дается время ожидания и степень рас-сеивания, также не наблюдается значительного отличия двух ре-жимов работы.

На рис.7 показано влияние емкости буферной памяти устрой-ства (вверху) и источника (внизу). Результаты были получены для сравнительно небольшой глубины предвыборки ( $q_2 = 4$ ), по-этому увеличение емкости буферной памяти свыше 4 практически не оказывает влияния.

На рис. 8 (вверху) приводятся результаты, полученные при использовании группового расслоения. Результаты получены для случая, когда каждый из источников запросов обращается опреде-ленным образом в определенную группу блоков памяти. Результаты для случая произвольного обращения будут, очевидно, несколько хуже.

На рис. 8 (внизу) даются линейно сглаженные интенсивности обслуженных запросов ( $\lambda_2$ ) по источникам ( $i$ ), полученные для различных дисциплин обслуживания. Как следует из рисунка, нет существенного различия в обслуживании при использовании разных приоритетов. С другой стороны, введение любого вида приоритета приводит к дифференциации скорости обслуживания. Объясняется это тем, что позиционный принцип сохраняется всегда, благодаря наличию определенных схемных соединений в системе.

Результаты моделирования дадут возможность сделать следую-щие выводы:

- 1) Расслоение памяти может быть достаточно эффективным при увеличении глубины предвыборки,
- 2) Возможно использование синхронного режима работы уст-ройств; при незначительном снижении эффективности работы можно получить более простую схемную реализацию,
- 3) Возможно применение ограниченного расслоения, что мо-жет увеличить надежность системы и упростить ее организацию,
- 4) Результаты моделирования дают основание полагать, что предложенная организация обмена на уровне общего поля памяти

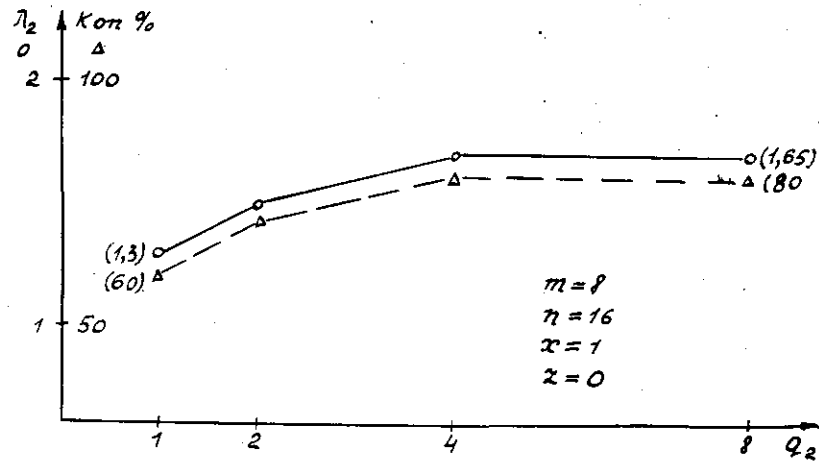
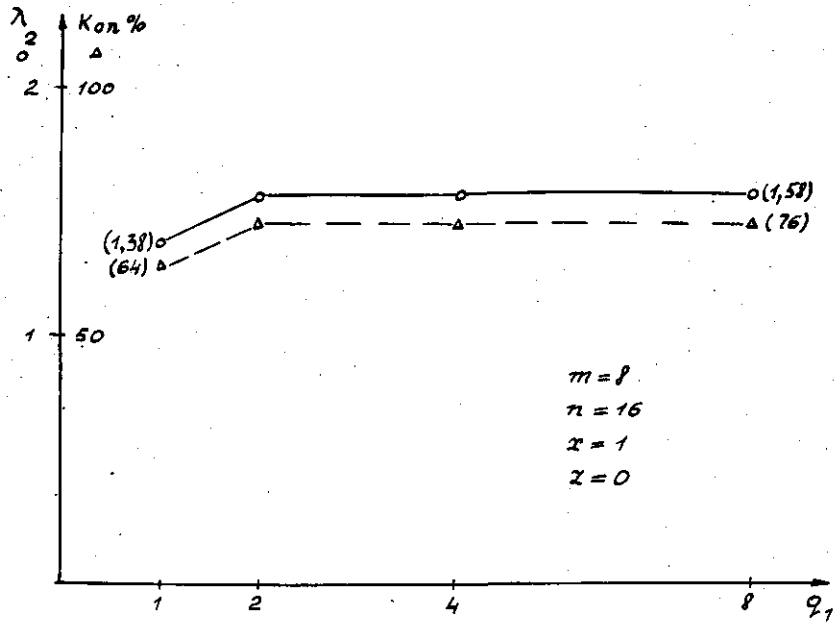


Рис. 7. Влияние емкости буфера ( $Q_1$  и  $Q_2$ ).

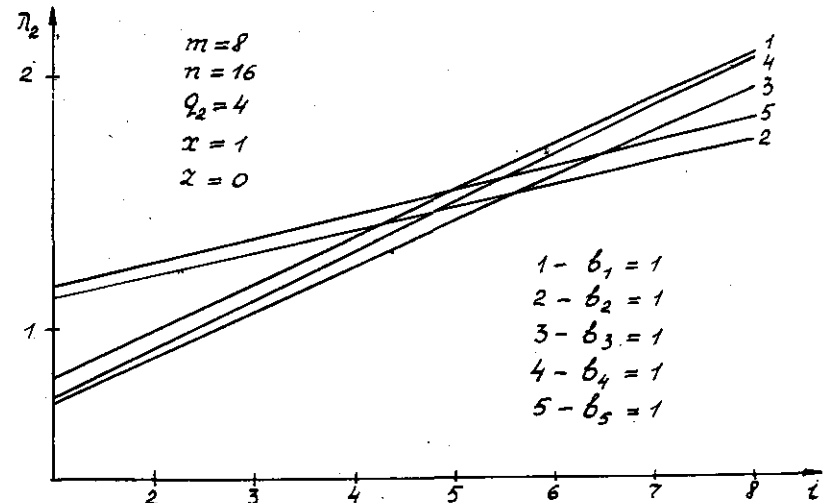
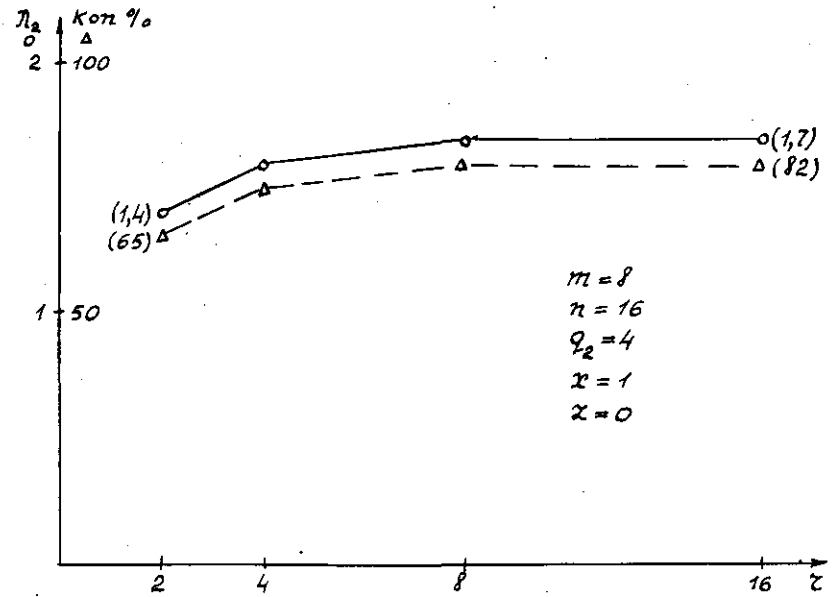


Рис. 8. Влияние глубины расслоения ( $z$ ) и дисциплины обслуживания (внизу).



окажется достаточно эффективной.

Автор выражает благодарность Сахнюк М.А. и коллективу ИК АН УССР, которым была поставлена данная задача моделирования.