

ОПТИМИЗАЦИЯ ЭНДОГЕННЫХ ПАРАМЕТРОВ МОДУЛЕЙ ОБЩЕЙ ПАМЯТИ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

М. Л. Федорова (Новомосковск), Т. М. Леденева (Воронеж)

В данной работе исследована пропускная способность модулей общей памяти вычислительных систем в зависимости от объема буфера.

Процесс моделирования, т. е. описания функционирования системы во времени, в общем случае можно определить как исследование функциональной зависимости между экзогенными (независимыми) и эндогенными (зависимыми) переменными [1]. Т. е. на основе эксперимента с моделью можно спрогнозировать характер изменения выходных параметров при варьировании входными. Модель при этом используется в качестве «черного ящика». Для этих целей используются различные методы математического моделирования. Авторы данной статьи использовали имитационное моделирование для определения аппаратных характеристик компьютера, а также для исследования организации взаимодействия процессоров с общей памятью (ОП).

Данной теме посвящены многие публикации, в одной из них [2] рассматривается модель, состоящая из n процессоров и m n -портовых модулей общей памяти ОП _{j} ($j = 1, \dots, m$), представленная на рис. 1.

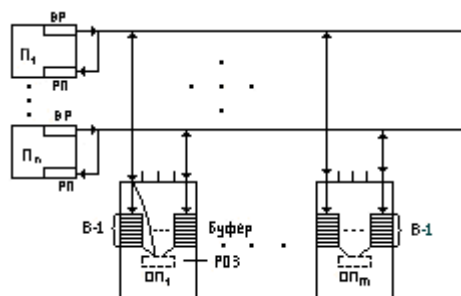


Рис. 1. Организация функционирования системы

Каждый процессор (П) системы с помощью собственной (неразделяемой) магистрали соединён с каждым модулем памяти ОП _{j} через предназначенный только для него порт. Процессоры независимо друг от друга с интенсивностью λ генерируют запросы в память, которые с равной вероятностью $q_j = 1/m$ направляются в один из модулей памяти ОП _{j} . Если модуль памяти свободен, то запрос сразу же поступает на обслуживание. В противном случае запрос записывается в соответствующий буфер модуля памяти ОП _{j} , недоступный другим процессорам системы. Процессор же после выдачи запроса продолжает работать и через некоторое время выдаёт новый запрос в модуль памяти.

Каждый порт модуля памяти имеет ограниченный буфер. В случае заполнения буфера выдаваемый процессором запрос сохраняется в блоке обмена информации на выходном регистре (ВР), а работа процессора блокируется до освобождения выходного регистра.

Приём процессором ответов из модулей памяти осуществляется на регистр приёма информации (РП), минуя ВР, который может быть занят блокирующим запросом. Синхронизация между выдачей и приёмом информации на магистрали обеспечивается соответствующим протоколом взаимодействия процессора с памятью.

При обслуживании запросов возможны два варианта организации модуля памяти: либо обслуживаемый запрос сохраняет за собой место в буфере на всё время обслуживания и только после передачи ответа в процессор освобождает его, либо место в

буфере освобождается сразу при приёме запроса на обслуживание. В последнем случае в модуле памяти должен существовать регистр обслуживаемого запроса (РОЗ), который, как и ВР процессора, увеличивает максимальное число запросов, находящихся в модуле памяти. Однако этот регистр является общим для всех процессоров и не относится к какому-либо отдельному буферу.

Если рассматривать процессоры как источники, а запросы, выдаваемые процессором, как транзакты, модули общей памяти и регистры – как каналы, то буфер будет многоканальным устройством. В этом случае данную задачу можно рассматривать как задачу массового обслуживания. Рассмотрим частный случай, в котором два процессора взаимодействуют с двумя модулями общей памяти. Согласно классификации Кендалла, это задача типа $M_1 | M_2 | CP(S) | N | FCFS$, т. е. многоканальная модель с ограниченной очередью, с полной разделенностью по каналам, блокировкой источника и с беспriorитетной дисциплиной обслуживания. В работе [2] данная система рассмотрена как цепь Маркова. Составляется система уравнений равновесия (СУР) и далее показано, что даже если рассматривать систему всего лишь из двух процессоров и двух модулей общей памяти, то СУР будет состоять из 28 уравнений. Естественно, если количество моделей памяти увеличить, то и количество уравнений неизмеримо возрастет. В публикациях по этой теме подробно описаны системы с показательным временем обслуживания. По этой причине в статье [2] используется экспоненциальное распределение, хотя данное допущение несколько занижает показатель эффективности системы. Авторы настоящей статьи предлагают решить задачу с помощью языка имитационного моделирования, а именно: GPSS WORLD. Использование данного средства позволяет исследовать модель с детерминированным обслуживанием и без использования численных методов, искажающих результат. За единицу времени взят один такт работы модуля памяти. Можно также найти не только загрузку системы, но и коэффициенты загрузки и эффективности каждого из компонентов: модулей памяти, процессоров и регистров. Для решения задачи были построены две модели: одна с РОЗ, вторая – без него. На рис. 2 приводится функциональная схема симуляции модели.

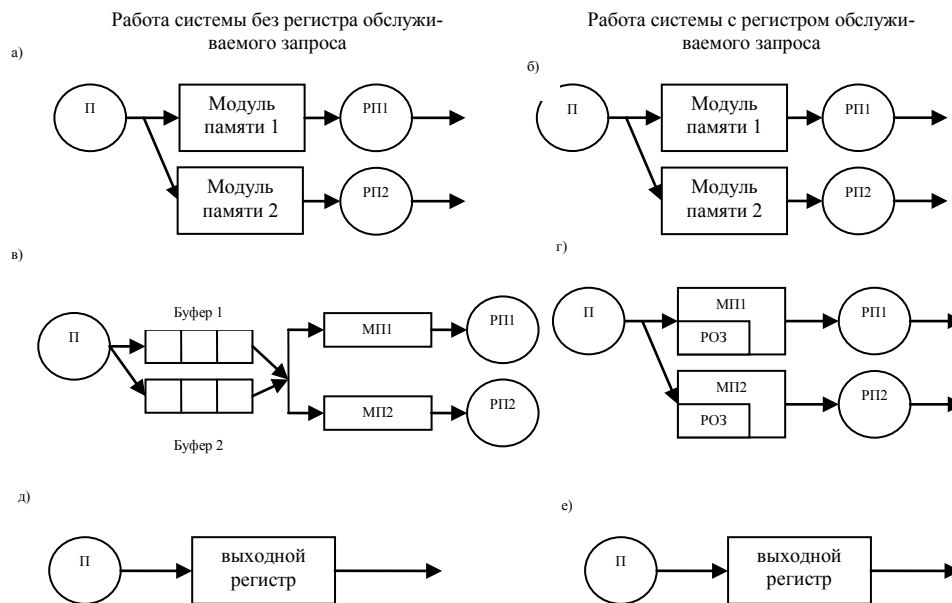


Рис. 2. Функциональная схема модели

Общим элементом для всех схем является запрос. Если это возможно, то запрос обрабатывается одним из свободных модулей памяти с равной вероятностью и затем передается на регистр приема (см. рис. 2, а, б), в противном случае запрос с равной вероятностью поступает в один из буферов моделей памяти (МП), причем в случае, показанном на рис. 2, в, запрос сохраняется в буфере все время его обслуживания. В случае, приведенном на рис. 2, з, запрос в буфере не сохраняется, а хранится в регистре обслуживаемого запроса, освобождая при этом место в буфере. На рис. 2, д и е изображен случай, когда поступающий запрос не может быть помещен в буфер из-за его заполненности. В этом случае запрос поступает в выходной регистр, а все остальные запросы будут потеряны, пока один из буферов не освободится.

Фрагмент программы с регистром обслуживаемого запроса выглядит следующим образом:

```

GENERATE                               (Exponential(1,0,2))
TRANSFER                               0.5,KANAL1,KANAL2
KANAL1 GATE NU REGOBCL1,QUE1; общая память свободна
SEIZE   REGOBCL1
ADVANCE 1
RELEASE REGOBCL1
SEIZE   REGPR1
ADVANCE (Exponential(1,0,0.35))
RELEASE REGPR1
SAVEVALUE NOBR+,1
TERMINATE                               1
KANAL2 GATE NU REGOBCL2,QUE2; общая память свободна
SEIZE   REGOBCL2
ADVANCE 1
RELEASE REGOBCL2
SEIZE   REGPR2
ADVANCE (Exponential(1,0,0.35))
RELEASE REGPR2
SAVEVALUE NOBR+,1
TERMINATE                               1
QUE1   GATE SNF BUF1,BLOKIR1
BB1    ENTER   BUF1
PER    TRANSFER 0.5,KANAL3,KANAL4
KANAL3 SEIZE   REGOBCL1
LEAVE  BUF1
ADVANCE 1
RELEASE REGOBCL1
SEIZE   REGPR1
ADVANCE (Exponential(1,0,0.35))
RELEASE REGPR1
SAVEVALUE NOBR+,1
TERMINATE                               1
KANAL4 SEIZE   REGOBCL2
LEAVE  BUF1
ADVANCE 1
RELEASE REGOBCL2
SEIZE   REGPR2

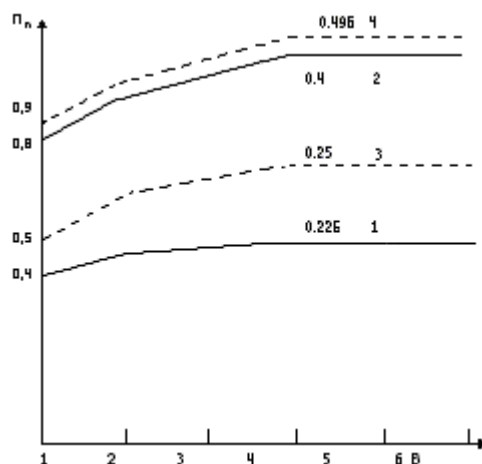
```

```

ADVANCE (Exponential(1,0,0.35))
RELEASE REGPR2
SAVEVALUE NOBR+,1
TERMINATE                                1
QUE2   GATE SNF BUF2,BLOKIR2
BB2    ENTER   BUF2
PER1   TRANSFER 0.5,KANAL5,KANAL6
KANAL5 SEIZE    REGOBCL1
LEAVE  BUF2
ADVANCE 1
RELEASE REGOBCL1
SEIZE    REGPR1
ADVANCE (Exponential(1,0,0.35))
RELEASE REGPR1
SAVEVALUE NOBR+,1
TERMINATE                                1
KANAL6 SEIZE    REGOBCL2
LEAVE  BUF2
ADVANCE 1
RELEASE REGOBCL2
SEIZE    REGPR2
ADVANCE (Exponential(1,0,0.35))
RELEASE REGPR2
SAVEVALUE NOBR+,1
KON    TERMINATE                                1
BLOKIR1 TEST E   F$BR1,0,POT
SEIZE   BR1
TEST E SF$BUF1,0
RELEASE BR1
TRANSFER ,BB1
BLOKIR2 TEST E F$BR2,0,KON
SEIZE   BR2
TEST E SF$BUF2,0
RELEASE BR2
TRANSFER ,BB2

```

Модель с введением РОЗ отличается тем, что запрос при обслуживании его памятью уже не находится в буфере, поэтому количество запросов, посылаемых в буфер, будет увеличено. Это иллюстрируется графиком (рис. 3) зависимости пропускной способности системы Π_m , определяемой как среднее число запросов, обслуженных в единицу времени, от количества мест в буфере. При увеличении среднего коэффициента использования общей памяти пропускная способность памяти монотонно возрастает.



**Рис. 3. График зависимости пропускной способности системы от количества мест в буфере (цифры над графиками показывают средний коэффициент использования общей памяти):
1, 2 – модель без РОЗ; 3, 4 – модель с РОЗ**

Выводы

Используя средства имитационного моделирования, можно точно определить характеристики полносвязной многопроцессорной системы с секционированной многопортовой памятью до ее введения в эксплуатацию.

Литература

1. Советов Б. Я., Яковлев С. А. Моделирование систем. М.: Высш. школа, 1985. 271 с.
2. Сокол Ю. М. Анализ пропускной способности буферизированной многопортовой памяти//Автоматика и телемеханика. 1992. № 3. С. 153–163.