

УПРАВЛЯЮЩАЯ ПРОГРАММА СИСТЕМЫ МИНИМАКС

В.Г.Кербель, В.Д.Корнеев, Н.Н.Миренков, Е.В.Щербаков

Описывается операционная часть программного обеспечения однородной вычислительной системы (OBC) МИНИМАКС [1,2]; главное внимание уделяется структуре и алгоритмам управляющей программы, функциям модулей и их взаимодействию между собой.

Основные задачи управляющей программы - обеспечение связи с внешним миром и рациональное использование ресурсов системы. Функционирование управляющей программы базируется на алгоритмах планирования [3,4], позволяющих синтезировать в системе различные подсистемы и специфицировать функции последних. Гибкое использование ресурсов и универсальность применения OBC достигаются: 1) введением специальных режимов работы подсистем: параллельной обработки, автономной работы машин, режима профилактики и режима диспетчера; 2) иерархической структурой управляющей программы, состоящей из Главного диспетчера, Старших диспетчеров режимов автономной работы и параллельной обработки и Диспетчеров элементарных машин (ЭМ) [1,5].

I. Главный диспетчер - программа, осуществляющая организацию подсистем, определяющая режимы работы для образованных подсистем, управляющая Старшими диспетчерами и операциями ввода/вывода для системных внешних устройств, присоединенных к машинам, реализующим режим диспетчера.

Организация подсистем и определение режимов работы осуществляется Главным диспетчером

1) по директиве оператора;

2) автоматически, при выполнении некоторых условий, связанных с характеристиками поступивших параллельных программ;

3) автоматически, при наличии информации о выходе машин из строя.

Директивы оператора позволяют задавать для подсистем любой режим, условия пункта 2 - режим параллельной обработки, пункта 3 - режим профилактики. Машины, не используемые подсистемами режимов диспетчера, параллельной обработки и профилактики, автоматически переводятся в режим автономной работы.

Для воздействия на организацию подсистем у оператора имеется три директивы: перевод ЭМ в режим профилактики, возвращение ЭМ в систему после профилактики и организация графика работы подсистем. Последняя директива, помимо прямого назначения режимов работы, может иметь условные списки ЭМ, работающих в режиме автономной работы, но предназначенных для динамического перевода в режим параллельной обработки при выполнении одного из следующих условий:

$$\sum_{i=1}^N T'_i R_i \geq k \cdot t \cdot L, \quad \max R_i > L,$$

где $T'_i = \begin{cases} T_i, & \text{если } T_i < t, \\ t, & \text{если } T_i \geq t; \end{cases}$

t - остаток времени действия директивы;

T_i - предполагаемое время, необходимое для завершения i -й задачи;

N - число задач, требующих режима параллельной обработки;

L - число ЭМ, работающих в режиме параллельной обработки;

k - некоторое положительное число, меньшее 1;

R_i - ранг i -й задачи (число ЭМ, требуемых для её решения).

Динамический перевод ЭМ из одного режима в другой используется Главным диспетчером при выходе из строя ЭМ, принадлежащих режиму параллельной обработки, а также для реализации эффективных алгоритмов загрузки подсистем параллельными задачами [4]. В основе организации подсистем режима параллельной обработки лежат запросы на параллельный счет, собираемые Главным диспетчером в очередях, которые создаются для задач каждого ран-

га. При построении очереди учитываются приоритеты поступивших запросов.

Организовать подсистему - значит выполнить следующие действия:

- проверить возможность реализации из заданных машин связной и линейной структур (см. [6]). Если такой возможности нет, прервать реализацию некоторых р-программ; из освободившихся и заданных машин получить новые подсистемы требуемых структур; возобновить счет прерванных р-программ;

- в таблицу подсистем, содержащуюся в Главном диспетчере, занести имя новой подсистемы, абсолютные номера ЭМ, включенных в нее, адрес настроичного массива^{*)} к этой подсистеме;

- в специальные ячейки оперативной памяти каждой ЭМ, включенной в подсистему, занести: относительный номер ℓ машин в подсистеме, ранг ℓ подсистемы, направление и длину кратчайшего пути (по связям(2)[1]) к подсистеме, работющей в режиме диспетчера, данные для регистра настройки $RH_{\ell-1}$ и $RH_{\ell+1}$, соответственно для связи с машинами $\ell-1$ и $\ell+1$ (для защиты подсистем друг от друга содержимое RH ℓ -й ЭМ равно $RH_{\ell-1}$, или $RH_{\ell+1}$; для первой и R -й ЭМ $RH_{\ell-1} = RH_{\ell+1}$);

- в каждой ЭМ подсистемы установить маски прерывания от индивидуальных внешних устройств.

Управление Старшими диспетчерами осуществляется Главным диспетчером на основе информационных таблиц, правила поочередного обслуживания или правила абсолютных приоритетов.

Правило поочередного обслуживания определяет передачу управления одному из Старших диспетчеров с частотой, пропорциональной приоритету, присвоенному соответствующему режиму; правило абсолютных приоритетов всегда передает управление старшему диспетчеру, имеющему наибольший приоритет (выбор правил поочередного обслуживания и абсолютного приоритета определяется при генерировании управляющей программы).

*) Настроичный массив (НМ) - массив из $\lceil \frac{\ell+2}{2} \rceil + 1$ ячеек, содержащий информацию для регистров настройки ЭМ ($[x]$ - целая часть x); ℓ - число ЭМ между настраивающей машиной и самой удаленной от неё [1] (хранится НМ, как правило, в упакованном виде).

Информационная таблица содержит информацию о завершении Главным диспетчером операций ввода/вывода на системных внешних устройствах, данные службы времени, приказы на перевод ЭМ из одного режима работы в другой, информацию о состоянии системы (наличие подсистем, графика их работы, статус системных внешних устройств и т.п.). Кроме того, в информационной таблице диспетчера режима автономной работы указаны настроочные массивы (информация для прокладывания путей до каждой машины, работающей в режиме автономной работы), общее число таких машин, управляющие данные для связи с пользователями, а для Старшего диспетчера режима параллельной обработки – настроочные массивы до каждой подсистемы, число подсистем, управляющие данные о загрузке подсистем параллельными программами.

Управление операциями ввода/вывода на системных внешних устройствах Главный диспетчер осуществляет по запросам, поступающим от Старших диспетчеров. Эти запросы содержат вид операции, тип внешнего устройства, объем передаваемых данных. Выделяются два вида буферной памяти: быстрый буфер (для внешних устройств типа диск) и медленный буфер (для устройства типа АЦПУ). Последний, кроме резидентной части, расположенной в оперативной памяти, имеет продолжение на диске.

Кроме того, Главный диспетчер распознает директивы оператора и выдает ему сообщения о состоянии системы, наличии запросов на параллельный счет, об окончании времени действия предыдущей директивы и т.п.

П. Старший диспетчер режима автономной работы, взаимодействуя с Главным диспетчером, обеспечивает доступ каждой машине, включенной в его подсистему, к Главному диспетчеру и системным внешним устройствам. Старший диспетчер автономной работы реализуется в каждой машине подсистемы, работающей в режиме диспетчера.

Основные функции диспетчера:

- 1) управление очередностью обслуживания ЭМ на основе заявленного приоритета или вида выполняемой работы;
- 2) предоставление рабочих каналов для доступа к подсистеме, работающей в режиме диспетчера, и системным внешним устройствам;

3) прием и анализ запросов на ввод/вывод для системных внешних устройств от программ и директив пользователей;

4) формирование заданий Главному диспетчеру;

5) реализация обменов с ЭМ, затребовавшей ввод/вывод для системных внешних устройств;

6) реализация "почтовой" связи между пользователями.

Управление очередностью обслуживания ЭМ осуществляется согласно специальной последовательности, генерируемой на основе приоритетов $\rho_i(t)$, поставленных в соответствие (или заданных пользователем) некоторому виду работы, выполняемой на i -й машине в момент времени t . Эта последовательность состоит из $S = \sum_{i=1}^{L_1} \rho_i$ элементов, каждый из которых является номером ЭМ, работающей в режиме автономной работы (L_1 – число таких ЭМ); при этом одинаковые номера встречаются с частотой, пропорциональной ρ_i . Процесс генерирования состоит из восьми этапов:

1) упорядочивания ρ_i по убыванию;

2) вычисления длины последовательности $S = \sum_{i=1}^{L_1} \rho_i$;

3) $k := 0$;

4) $k := k + 1$;

5) вычисления $F_k = \left[\frac{S}{\rho_{i_k}} \right]$ и $f_k = S - F_k \cdot \rho_{i_k}$ для k -го элемента упорядоченной последовательности;

6) разбиения интервала слева на ρ_{i_k} подинтервалов, первые $\rho_{i_k} - f_k$ из которых имеют длину F_k , остальные $F_k + 1$; занесения в каждый подинтервал на самое левое свободное место номера i_k (если в подинтервале свободных мест нет, занесение производится на первое свободное место, встретившееся при движении к правому концу интервала),

8) передачи управления на п.4, если $k < L_1$.

Предоставление рабочих каналов к подсистеме, работающей в режиме диспетчера, и системным внешним устройствам Старший диспетчер режима автономной работы осуществляет на основе установленной очередности путем опроса ЭМ. Опрос осуществляется следующим образом.

С помощью настроочного массива в третий разряд регистра предварительной настройки машины, которой предполагается пре-

доставить каналы, заносится единица. После чего начинается процедура занесения нуля в тот же разряд рассматриваемого регистра предварительной настройки с помощью того же настроичного массива.

Машине, "желающей" выполнить операции ввода/вывода для системных внешних устройств, запускается на проверку содержимого регистра предварительной настройки. Обнаружение "1" в третьем разряде она воспринимает как информацию о предоставлении ей рабочих каналов и сообщает Старшему диспетчеру о "желании" воспользоваться указанной ситуацией. Сообщение воспринимается Старшим диспетчером в виде сигнала о конфликте, получаящемся в результате посылки встречной настройки из ЭМ в сторону Старшего диспетчера, или появления "1" в четвертом разряде регистра предварительной настройки машины, реализующей Старший диспетчера. Отсутствие конфликтного сигнала и единицы в четвертом разряде воспринимается как отказ от использования рабочих каналов. Получив сообщение о "желании" ЭМ воспользоваться рабочими каналами, Старший диспетчер прокладывает к ней путь по связям (I) и настраивается на прием от неё управляющего слова.

Прием и анализ запросов на ввод/вывод для системных устройств Старший диспетчер режима автономной работы выполняет после упомянутых выше взаимодействий с ЭМ. Запрос представляет собой управляющую последовательность вида:

0	C	n	- управляющее слово
1			
...			
n			- вызывающая последовательность

Управляющее слово содержит код C вида предполагаемой работы и число n слов вызывающей последовательности, специфицирующей эту работу. Выделяются следующие виды работ:

C = 0 - запрос системной программой (транслятором и т.п.) рабочей области на диске;

C = 1 - запрос, связанный с выполнением директивы пользователя или оператора;

C = 2 - запрос на ввод/вывод от программы пользователя;

C = 3 - запрос на ввод с диска от резидента диспетчера элементарной машины.

Вызывающая последовательность содержит данные об устройствах ввода/вывода, с которыми связана запрограммированная работа, объем и вид передаваемой информации, адреса буферных полей и т.п. Например, вызывающая последовательность для работы с диском может иметь следующий вид:

I	ДВС I (или 2)
2	ОСТ control
3	A
4	ДВС n, (или - 2n,)
5	ДВС F
6	ДВС Z

I - чтение; 2 - запись;
control - восьмеричное условное слово;
адрес буфера A;
число передаваемых слов (+n,),
символов (-2n,);
десятичный номер тракта рабочей зоны;
десятичный номер сектора.

Формирование заданий Главному диспетчеру Старший диспетчер осуществляет на основе управляющей последовательности. При этом адрес буферного поля ЭМ, содержащийся в вызывающей последовательности, заменяется на адрес стандартного буферного поля Главного диспетчера.

Реализация обменов между ЭМ и Старшим диспетчером выполняется частями массива, если его размер превышает длину стандартного буферного поля Главного диспетчера. Указанный размер выбирается из вызывающей последовательности и сообщается драйверам, реализующим "прием - передачу".

Организацию "почтовой" связи между пользователями Старший диспетчер выполняет по их директивам. Имеются директивы передачи сообщения в "почтовый ящик" и директива запроса содержимого "почтового ящика". Содержимое "почтового ящика" выдается адресату автоматически при входе его в систему.

Ш. Старший диспетчер режима параллельной обработки, взаимодействуя с Главным диспетчером, обеспечивает каждой подсистеме доступ к системным внешним устройствам, инициирует загрузку, запуск, счет и окончание параллельных программ. Этот диспетчера реализуется в каждой машине подсистемы, работающей в режиме диспетчера. Часть функций Старшего диспетчера, связанных с реализацией взаимодействий между ЭМ во время счета параллельной программы, перенесена в языки программирования.

ния на уровне операторов. Эти операторы имеют обращения к драйверам обменных взаимодействий, которые включены в Диспетчер элементарной машины.

Основные функции диспетчера:

- 1) управление очередностью обслуживания подсистем;
- 2) предоставление рабочих каналов к подсистеме, работющей в режиме диспетчера, и внешним устройствам;
- 3) прием и анализ запросов (индивидуальная связь со Старшим диспетчером, ввод/вывод для системных внешних устройств) от параллельных программ, реализующихся на подсистемах;
- 4) формирование заданий Главному диспетчеру;
- 5) реализация обменов с подсистемой, потребовавшей ввод / вывод;
- 6) загрузка и запуск параллельных программ;
- 7) снятие параллельных программ.

Управление очередностью обслуживания подсистем Старший диспетчер выполняет на основе рангов параллельных задач и заявленных приоритетов. Приоритеты характеризуют интенсивность обращений параллельной программы к системным внешним устройствам. Алгоритм генерирования обслуживаний последовательности подобен описанному в Старшем диспетчере режима автономной работы. Отличие в том, что роль $P_i(t)$ здесь играет произведение ранга на заявленный приоритет.

Предоставление рабочих каналов Старшим диспетчером режима параллельной обработки осуществляется на основе установленной очередности путем опроса подсистем. Опрос осуществляется в несколько приемов аналогично опросу, реализуемому Старшим диспетчером режима автономной работы.

С помощью настроичного массива в третий разряд регистра предварительной настройки машины, с относительным номером I в подсистеме, заносится единица. После чего начинается процедура занесения нуля в тот же разряд рассматриваемого регистра предварительной настройки с помощью того же настроичного массива.

Машины подсистемы, "желания" выполнить операции ввода/вывода на системных внешних устройствах, защищиваются на проверке содержимого своих регистров предварительной настройки. Первой обнаруживает возможность использования рабочих каналов

ЭМ с относительным номером I. Она информирует Старший диспетчер режима параллельной обработки о "желании" использовать эти каналы путем посылки встречной настройки. После чего они вступают во взаимодействия. Старший диспетчер режима параллельной работы настраивает рабочие каналы, а ЭМ с номером I передает ему управляющее слово следующего вида:

C	K	n
---	---	---

,

где C - код вида взаимодействия; n - длина вызывающей последовательности;

$$K = \begin{cases} 0, & \text{если каждая ЭМ подсистемы требует связи со Старшим диспетчером;} \\ l, & \text{если только } l\text{-я ЭМ требует этой связи.} \end{cases}$$

Прием и анализ запросов Старший диспетчер режима параллельной обработки выполняет после выработки конфликтного сигнала, связанного с посылкой встречной настройки, или появления "I" в четвертом разряде регистра предварительной настройки ЭМ, реализующей режим диспетчера. Запрос представляет собой следующую управляющую последовательность:

0	C	K	n
1			
2			
...			
N			

- управляющее слово
- вызывающая последовательность

$C = 0$ - запрос на ввод/вывод от параллельной программы; $C = 1$ - запрос, связанный с нормальным завершением счета; $C = 2$ - запрос, связанный с аварийной ситуацией.

В случае $C = 0$ формат вызывающей последовательности аналогичен формату вызывающей последовательности Старшего диспетчера режима автономной работы. При $C = 1$ вызывающая последовательность отсутствует (предполагается, что пользователь нужную ему информацию запомнил в своих файлах). При $C = 2$ вызывающая последовательность содержит относительные номера l_i машин, в которых произошли аварийные ситуации, и типы ситуаций (Z_i):

Z_1	l_1
Z_2	l_2
...	...
Z_n	l_n

- вызывающая последовательность

Формирование заданий Главному диспетчеру Старший диспетчер режима параллельной обработки осуществляет на основе управляющей последовательности.

В случае операций ввода/вывода действия Старшего диспетчера режима параллельной обработки аналогичны действиям Старшего диспетчера автономной работы. Отличие состоит в том, что операция ввода/вывода повторяется λ раз (λ - ранг параллельной программы), при этом (кроме замены адреса буферного поля ЭМ, содержащегося в вызывающей последовательности, на адрес стандартного буферного поля Главного диспетчера) происходит λ -кратная модификация адреса дисковой памяти.

В случае нормального завершения счета Старший диспетчер режима параллельной обработки информирует Главный диспетчер об освобождении подсистемы. В случае аварийного завершения Старший диспетчер режима параллельной обработки обеспечивает вывод причины аварийной ситуации и информации, заказанной пользователем, на буфер вывода.

Реализация обменов между Старшим диспетчером и подсистемой, работающей в режиме параллельной обработки, выполняется:

- 1) за λ этапов (последовательно для всех ЭМ), если требуется ввод/вывод для каждой машины подсистемы;
- 2) за один этап, если требуется ввод/вывод для одной машины подсистемы.

Выполнение обменов основывается на управляющей последовательности, переданной Старшему диспетчеру режима параллельной обработки из первой машины.

Обмены реализуются частями массивов, если их размеры превышают длину стандартного буферного поля Главного диспетчера.

Загрузку и запуск параллельных программ Старший диспетчер режима параллельной обработки выполняет по заданию Главного диспетчера. Вид задания иллюстрируется схемой (см. стр. I39).

На приведенной схеме:

НЗ - номер записи в таблице подсистем. Запись содержит имя и ранг подсистемы, абсолютные номера ЭМ и настроочный массив для прокладывания пути к этой подсистеме.

$\gamma = 0$ соответствует загрузке и запуску параллельной про-

γ	НЗ	α
A_1		
0	ℓ_{11}	ℓ_{12}
1	ℓ_{13}	ℓ_{14}
A_2		
0	ℓ_{21}	ℓ_{22}
1	ℓ_{23}	0000
...		
A_λ		
1	$\ell_{\lambda 1}$	$\ell_{\lambda 2}$

граммы; $\gamma = 1$ - её снятие с сохранением результатов счета; $\gamma = 2$ - её снятие без сохранения результатов счета.

α - количество различных ветвей в параллельной программе ($\alpha = 1$ указывает на идентичность всех ветвей).

A_i - длина i -й ветви из общего числа различных ветвей.

ℓ_{ik} - относительные номера ЭМ подсистемы, в которые загружается эта i -я ветвь.

Единица в нулевом разряде указывает на конец списка относительных номеров ЭМ, в которые передается одна и та же ветвь (i -я из общего числа различных ветвей).

После получения задания от Главного диспетчера на загрузку подсистемы Старший диспетчер прокладывает путь к машинам, назначенным в эту подсистему, настраивает их на прием команды Обобщенного безусловного перехода (ОБП), приказывает принять управляющее слово, содержащее размер подготовленной к счету р-ветви и осуществляет загрузку последней. В случае идентичности р-ветвей загрузка осуществляется сразу для всех ЭМ подсистемы, иначе она реализуется за несколько приемов, число которых равно числу различных р-ветвей. Загруженная в ЭМ р-ветвь сразу же получает управление.

После загрузки всех р-ветвей Старший диспетчер настраивает ЭМ подсистемы для работы в режиме параллельной обработки и устанавливает в четвертом разряде регистра предварительной настройки каждой машины единицу, разрешающую автономное (не связанное с обращением к Старшему диспетчеру) использование рабочих каналов в пределах подсистемы.

Снятие параллельных программ осуществляется в следующих четырех случаях: при нормальном завершении счета, при аварийном завершении его, по истечении заданного интервала времени, при экстренном снятии задачи.

Два первых случая распознаются по управляющей последовательности, полученной из подсистемы; два вторых - по заданию

Главного диспетчера. Предполагается, что при нормальном завершении счета пользователь всю необходимую ему информацию записал во внешнюю память с помощью операторов вывода. По истечении заданного интервала времени реализуется запоминание оперативной памяти подсистемы, обеспечивающее дальнейшее продолжение счета. При аварийном завершении обеспечивается выдача причины аварийной ситуации и распечатка областей памяти, заданной пользователем. При экстренном снятии задачи результаты счета теряются.

Снятие задачи требует выполнения следующей последовательности действий:

- 1) занесение нуля в четвертый разряд регистра предварительной настройки каждой ЭМ;
- 2) настройка подсистемы на прием команды ОБП от Старшего диспетчера (пп. 1-2 не используются при нормальном и аварийном завершении);
- 3) передача управляющего слова, содержащего информацию о причине снятия задачи и областях памяти, подлежащей запоминанию;
- 4) выгрузка содержимого указанных областей памяти (пп. 3-4 не используются при нормальном завершении);
- 5) сообщение Главному диспетчеру об освобождении подсистемы.

IV. Диспетчер элементарной машины - программа, обеспечивающая вход в систему и распознавание директив пользователя, взаимодействие со Старшим диспетчером, обработку аварийных ситуаций и прерываний, управление вводом/выходом для индивидуальных внешних устройств, участие в реализации операторов обменных взаимодействий.

Основой Диспетчера элементарной машины может служить практически любая операционная система машины, на базе которой построена ОВС.

Для входа в систему пользователь должен включить телетайп одной из ЭМ, работающей в режиме автономной работы, и напечатать произвольную последовательность символов. В результате в память этой машины с диска будет загружена программа управления заданиями, сообщающая о своей готовности воспринимать ди-

рективы. На основе поступающих директив формируются задания Старшему диспетчеру.

При обработке аварийных ситуаций Диспетчер элементарной машины выдает сообщение на телетайп или вырабатывает соответствующий признак о ситуации в зависимости от функционирования ЭМ в режимах автономной работы или параллельной обработки. Анализ признака аварийной ситуации осуществляется при выходе машин подсистемы, работающей в режиме параллельной обработки, на операторы обменных взаимодействий.

Помимо обработки обычных прерываний (связанных, например, с индивидуальными внешними устройствами) Диспетчер элементарной машины обрабатывает прерывания, связанные с командой ОБП, посланной Старшим или Главным диспетчераами. ОБП используется для перевода ЭМ из одного режима работы в другой.

Ячейки прерывания
МОНИТОР
Драйверы для ИВУ ^{**}
Драйверы системных взаимодействий
Область программ пользователя

Л и т е р а т у р а

1. ВИНОКУРОВ В.Г., ДМИТРИЕВ Ю.К., ЕВРЕИНОВ Э.В., КОСТЕЛЯНСКИЙ В.М., ЛЕХНОВА Г.М., МИРЕНКОВ Н.Н., РЕЗАНОВ В.В., ХОРОШЕВСКИЙ В.Г. Однородная вычислительная система из мини-машин. -В кн.: Вычислительные системы. Вып.51. Новосибирск, 1972, стр. 127-146.
2. МИРЕНКОВ Н.Н. МИНИМАКС - вычислительная система коллективного пользования. -Настоящий сборник, стр.115-128.
3. МИРЕНКОВ Н.Н. Алгоритмы планирования для диспетчера однородной вычислительной системы. -В кн.: Вычислительные системы. Вып.42. Новосибирск, 1970, стр. 34-46.
4. КРЫЛОВ Э.Г., МИРЕНКОВ Н.Н. Алгоритм планирования функциональных состояний однородной вычислительной системы. Отчет ИМ СО АН СССР, 1973.
5. МИРЕНКОВ Н.Н. Диспетчирование в однородных вычислительных системах. -В кн.: Материалы 3 Всесоюзной конференции по однородным вычислительным системам и средам, Таганрог, 1972, стр. 51-54.
6. МИРЕНКОВ Н.Н., ФИНЕРМАН С.Б. Алгоритмы распознавания подсистем заданных структур в однородной вычислительной системе. Отчет ИМ СО АН СССР, 1973.

Поступила в ред.-изд.отд.

9 июля 1973 года