

УДК 681.142.2

ОПЕРАТИВНОЕ ПЛАНИРОВАНИЕ ДЛЯ НЕОДНОРОДНЫХ СПЕЦИАЛИЗИРОВАННЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

С.Д. Панкеев, Р.И. Миняев

Рассматриваются алгоритмы (блок-схема приведена на рисунке) оперативного планирования для совместных систем с разделением времени, предназначенный для получения оптимального, в смысле приведенного ниже критерия, графика реализации пакетов задач.

Планирование предусматривается для неоднородной специализированной вычислительной системы (СВС) следующей структуры.

1. СВС состоит из M подсистем вычислительных машин (ВМ).
2. Каждая подсистема состоит из P типов ВМ, объединенных через общую оперативную память, и автономного устройства обмена (УО) для связи с внешними абонентами.
3. Все подсистемы объединены через внешнее поле памяти с помощью своих УО.
4. Машин подсистем могут быть специализированы по типам решаемых задач.

Исходными данными о пакетах задач для алгоритма планирования являются:

- 1) матрица $\{C_{ik}\}$ ($i, k = 1, \bar{m}$) отношений графа пакета задач;
- 2) матрица $\{\tau_{i\mu}\}$ времен реализации i -х задач на ВМ μ -го типа;
- 3) матрица $\{\tau_{okl}\}$ времен, затрачиваемых на обмен инфор-

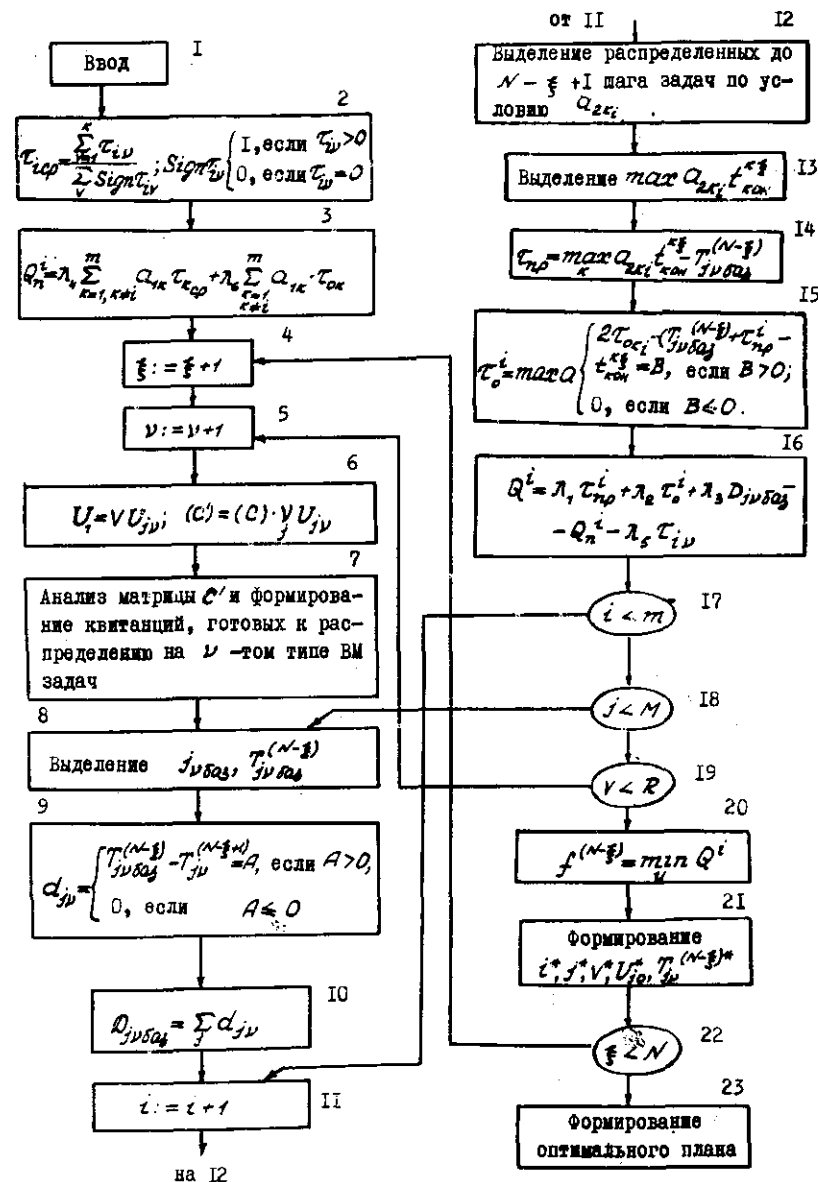


Рис. I

мащей при реализации связанных задач;

4) предварительная функциональная связность $i \rightarrow \nu$ задач с типом ВМ.

Задачей планирования является выработка оптимального графика реализации пакета задач на СВС в смысле критерия

$$T_{BC} = \min \{ \max_j T_j \}, \quad (I)$$

где T_j - время занятости j -й ВМ при реализации пакета задач.

Общий алгоритм оптимального планирования для СВС [1], основанный на аппарате динамического программирования, позволяет получать точное решение задачи планирования. Однако реализация его на планирующей ВМ требует весьма значительных затрат производительности и особенно объема памяти.

Предлагаемый алгоритм оперативного планирования позволяет решать задачу планирования за значительно меньшее время и при весьма малых объемах памяти планирующей ВМ.

Идея оперативного планирования состоит в выборе на каждом шаге многошагового процесса возвратного динамического программирования единственного решения о распределении i -й задачи на j -ю машину ν -го типа ($j = 1, M$; $\nu = 1, R$), оптимального в смысле критерия Q^i . Так как значительное сокращение перебора вариантов планирования по сравнению с методом оптимального планирования может ухудшить оперативный план, то предлагается применять более сложный критерий оптимальности оперативного планирования вида (2), учитывающий не только характеристики процесса планирования на каждом шаге $N-\xi$, но и будущие его свойства на основе характеристик графа пакета задач.

$$Q^i = \lambda_1 \tau_{np}^i + \lambda_2 \tau_o^i + \lambda_3 \cdot D_{j\nu\delta\alpha_3} - \lambda_4 \sum_{k=i, k \neq i}^m Q_{ik} \tau_{kcp} - \lambda_5 \tau_{i\nu} - \lambda_6 \sum_{k=i, k \neq i}^m Q_{ik} \tau_{ok i} \quad (2)$$

Первое слагаемое функции критерия Q^i вычисляется операторами (13 - 16) алгоритма и представляет собой время τ_{np}^i простоя ВМ, на которую назначена на данном шаге $N-\xi$ i -я задача, информационно связанная с задачами, распределенными на предыдущем $N-\xi+1$ шаге планирования.

$$\tau_{np}^i = \max_{k \neq i} Q_{k i} t_{kon}^{k \xi} - T_{j\nu\delta\alpha_3}^{(N-\xi)}, \quad (3)$$

где $t_{kon}^{k \xi}$ - планируемое инверсное время окончания реализации k -й задачи;

$T_{j\nu\delta\alpha_3}$ - время занятости j -й ВМ ν -го типа, на которую назначена i -я задача;

$$a_{k i} = \begin{cases} 1, & \text{если } i \text{ -я задача информационно связана с } k \text{ -й задачей, распределенной до } N-\xi+1 \text{ шага,} \\ 0, & \text{в противном случае.} \end{cases}$$

Второе слагаемое функции Q^i вычисляется оператором (15) и представляет собой время τ_o^i , необходимое для обмена информацией через внешнее поле памяти при реализации информационно связанных задач на различных подсистемах СВС:

$$\tau_o^i = \max_k a_{k i} \begin{cases} 2\tau_{ok i} - (T_{j\nu\delta\alpha_3} + \tau_{np}^i - t_{kon}^{i \xi}) = B, & \text{если } B > 0, \\ 0, & \text{если } B \leq 0, \end{cases} \quad (4)$$

где $\tau_{ok i}$ - входная мощность k -й задачи, информационно связанной с рассматриваемой i -й задачей.

Третье слагаемое (операторы 9, 10) представляет собой дебаланс загрузки машин в системе, вызванный распределением i -й задачи на машину $j\nu\delta\alpha_3$.

$$D_{j\nu\delta\alpha_3} = \sum a_{j\nu}, \quad a_{j\nu} = \begin{cases} T_{j\nu\delta\alpha_3}^{(N-\xi)} - T_{j\nu}^{(N-\xi+1)} = A, & \text{если } A > 0, \\ 0, & \text{если } A \leq 0. \end{cases} \quad (5)$$

Минимизация первых трех слагаемых функции Q^i позволяет получать план с минимальными простоями машин системы, временами обмена через внешнее поле памяти и с равномерной загрузкой машин системы. Эти слагаемые вычисляются на основе характеристик процесса планирования на каждом шаге $N-\xi$.

Следующие три слагаемых функции Q^i (операторы 2, 3, 16) формируются на основе характеристик графа-схемы пакета задач и сопоставляются каждой i -й задаче один раз перед началом решения задачи планирования.

Четвертое и шестое слагаемые (операторы 2, 3) функции Q^i представляют собой суммарные длины всех ветвей графа, оканчивающихся i -й вершиной. Здесь: $\tau_{i co}$ - среднее время реализации i -й задачи на различных типах машин системы

$$\tau_{i\sigma} = \frac{\sum \tau_{i\nu}}{\sum \text{sign} \tau_{i\nu}}, \quad \text{sign} \tau_{i\nu} = \begin{cases} 1, & \text{если } \tau_{i\nu} > 0, \\ 0, & \text{если } \tau_{i\nu} = 0, \end{cases} \quad (6)$$

$$Q_{ik} = \begin{cases} 1, & \text{если } k \text{ -я вершина принадлежит ветвям графа,} \\ & \text{оканчивающимся } i \text{ -й вершиной,} \\ 0, & \text{в противном случае.} \end{cases}$$

Применение сложного составного критерия позволяет с помощью выбора весовых коэффициентов λ_i адаптировать алгоритм к различным классам пакетов задач. Для выбора оптимальных значений весовых коэффициентов λ_i для различных классов пакетов задач можно легко построить поисковый алгоритм по критерию вида (1).

Алгоритм планирования представляет N - m -шаговую процедуру динамического программирования.

На каждом шаге N - ξ для каждого ν -го типа ВМ выделяются готовые к распределению задачи (операторы 6,7). В качестве базовых машин для распределения i -й готовой задачи поочередно выбираются все ВМ каждого типа (оператор 8). Формируются варианты планирования и вычисляются значения функции критерия Q^i (операторы 9 - 20).

По основному рекуррентному соотношению динамического программирования (оператор 21) выбирается единственный оптимальный вариант распределения для N - ξ этапа. Для этого варианта формируются и запоминаются следующие данные:

1. i^* - номер распределенной задачи;
2. $j_{\nu\delta Q_3}^*$ - номер и тип машины, на которую распределена i^* -я задача;
3. $U_{j\nu}^*$ - новое значение кода распределенных на $j\nu$ -ю ВМ задач;
4. $T_{j\nu}^{(N-\xi)}$ - новое значение времени занятости базовой ВМ.

$$T_{j\nu}^{(N-\xi)} = T_{j\nu\delta Q_3}^{(N-\xi)} + \tau_{i^*} + \tau_{no}^{i^*} + \tau_o^{i^*} \quad (7)$$

В конце N -го этапа распределения формируется оптимальный план реализации пакета задач.

План содержит следующие данные:

1. Матрицу отношений $\{C_{ik}^{n,l}\}$ планового графа пакета задач.

2. Ожидаемые времена $t_{\text{нач}}^i$ начала и $t_{\text{кон}}^i$ окончания реализации i -х задач на СВС.

3. Ожидаемое общее системное время T_{BC} реализации пакета задач.

4. Указание о графике работы УО.

Алгоритм реализован на ЭВМ "Минск-22". Программа состоит из 800 команд. Время реализации алгоритма для пакетов из нескольких десятков задач не превосходит в среднем несколько десятков секунд.

Л и т е р а т у р а

1. ПАНКЕЕВ С.Д. Методы оптимального диспетчирования мультипрограммной работы вычислительных систем с учетом случайных факторов. - "Доклады I-го Всесоюзного симпозиума по статистическим проблемам в технической кибернетике", М, 1967.