

ПОДХОД К РАСПРЕДЕЛЕНИЮ ПАРАЛЛЕЛЬНЫХ ВЕТВЕЙ ЗАДАНИЯ В ВЫЧИСЛИТЕЛЬНОЙ СИСТЕМЕ

Рассмотрена проблема распределения заданий в вычислительной системе для случая, когда длительность выполнения параллельных ветвей задания, которые необходимо распределить, различна. Для повышения эффективности распределения разработан метод. Показано, что предложенный метод является эффективным, поскольку, используя его, время выполнения заданий заметно сокращается.

Ключевые слова: вычислительная система, grid, параллельные вычисления, ветви задания, распределение заданий, время выполнения задания, время работы системы, время ожидания.

ВВЕДЕНИЕ

Сегодня направление параллельных вычислений в GRID имеет большой интерес со стороны научных исследований, поскольку увеличение скорости вычислений является весьма актуальным для всех, чья деятельность связана с большим объемом вычислительных работ [1]. Использование параллельных вычислений позволяет решать задачи, требующие больших временных затрат, за меньшее время [2, 3]. Однако при параллельном вычислении бывают случаи, когда некоторые ветви задания вычисляются раньше и ожидают окончания работы остальных. При этом вычислительные узлы простаивают, что сказывается на общем времени выполнения заданий, а, следовательно, и на загрузке всей системы. Таким образом, сокращение времени простоя позволит увеличить скорость вычислений.

В системах, в которых приоритет заданий играет важную роль, диспетчеризация нагрузки происходит поочередно для каждого задания и большее внимание уделяется распределению ветвей параллельного задания между компьютерами. В таком случае зачастую количество ветвей любого из заданий меньше количества компьютеров.

ПОСТАНОВКА ЗАДАЧИ

Рассмотрим вычислительную систему, в которую в определенный момент времени поступают параллельные задания, т. е. состоящие из параллельно выполняемых ветвей. Для сокращения общего времени выполнения всех поступающих для решения заданий и при этом минимизации нагрузки на вычислительные ресурсы следует определить, какая ветвь выполняемого параллельного задания будет выполняться на каком компьютере. Таким образом, необходимо решить проблему эффективного распределения параллельных ветвей задания между вычислительными узлами системы.

Пусть каждый из компьютеров имеет различную производительность (у некоторых компьютеров производительность может совпадать, но не у всех), длительность

выполнения каждой из ветвей одного задания также различна, в результате чего время выполнения задания в целом определяется наибольшей длительностью выполнения одной из ветвей этого задания. Также следует отметить, что количество ветвей не превышает количества компьютеров.

Пусть данная вычислительная система состоит из n компьютеров, а очередь заданий состоит из q заданий (каждое задание содержит l_i ветвей, где $i = \overline{1, q}$). Тогда математическая модель этой задачи может быть записана следующим образом.

Мощность эталонного компьютера – P^e .

Мощность k -го компьютера – P_k , где $k = \overline{1, n}$.

Тогда коэффициент мощности k -го компьютера –

$$c_k = \frac{P^e}{P_k}, \quad (1)$$

т. е. чем больше мощность k -го компьютера, тем меньше его коэффициент мощности.

Во время распределения i -го задания свободно только a_i компьютеров. Учитывая (1), соответственно, коэффициент мощности m -го компьютера –

$$c_m = \frac{P^e}{P_m}, \quad (2)$$

где $m = \overline{1, a_i}$.

Тогда время выполнения j -ой ветви i -го задания на m -м компьютере t_{ijm} :

$$t_{ijm} = t_{ij}^e \times c_m, \quad (3)$$

где $j = \overline{1, l_i}$, t_{ij}^e – время выполнения j -ой ветви i -го задания на эталонном компьютере.

Тогда время выполнения i -го задания в данной системе T_i :

$$T_i = \max \{t_{ijm}\}, \quad (4)$$

Время ожидания окончания выполнения всего задания m -м компьютером, закончившим выполнение j -ой ветви i -го задания, обозначим как Δt_{ijm} :

$$\Delta t_{ijm} = T_i - t_{ijm}. \quad (5)$$

Суммарное время ожидания окончания всех ветвей i -го задания Δt_i :

$$\Delta t_i = \sum_{j=1}^{l_i} \Delta t_{ijm}. \quad (6)$$

Таким образом, для сокращения общего времени выполнения всех поступающих для решения заданий необходимо, чтобы

$$\begin{cases} T_i \rightarrow \min \\ \Delta t_i \rightarrow \min \end{cases} \quad (7)$$

Введем понятие времени ожидания k -м компьютером начала выполнения $(i+r)$ -го задания после окончания выполнения i -го задания, где $r = \overline{1, (q-1)}$. Обозначим это время как Δt_{ki} .

Тогда время работы системы, потраченное на выполнение всех заданий, рассчитывается как максимальное время работы одного из n компьютеров этой системы и равняется сумме длительностей ветвей заданий, выполненных на конкретном k -м компьютере плюс сумма всех Δt_{ki} для этого же компьютера.

А суммарное время ожидания всей системы Δt :

$$\Delta t = \sum_{k=1}^n \sum_{i=1}^q \Delta t_{ki} + \sum_{i=1}^q \Delta t_i.$$

Таким образом, необходимо, чтобы, учитывая (7), время работы системы стремилось к минимуму и $\Delta t \rightarrow \min$.

СУЩЕСТВУЮЩИЕ РЕШЕНИЯ

Для сравнения рассмотрим простейший случай, когда в каждый момент времени распределяется одно задание и самая длительная ветвь задания распределяется на наиболее мощный компьютер, а далее распределение происходит в порядке возрастания длительностей ветвей заданий и убывания мощностей компьютеров.

Согласно такому подходу, порядковые номера компьютеров и ветвей совпадают, т.е. $m = j$.

Такой подход обеспечивает выполнение первого из условий (7), но не в полной мере обеспечивает второе.

Для выполнения обоих условий в данной работе предлагается метод, в котором распределяются ветви заданий таким образом, чтобы были заняты также и менее мощные компьютеры. При этом вновь рассчитанное время выполнения j -й ветви i -го задания, t_{ijm} , не превышало уже рассчитанное время выполнения i -го задания, T_i . Таким образом, получаем, что выполняются условия (5) и для распределения $(i+1)$ -го задания, выполняемого одновременно с i -м на других компьютерах, остаются вычислительные узлы с большей средней мощностью. За счет этого общее время выполнения Q заданий уменьшится.

РАЗРАБОТАННЫЙ МЕТОД

Рассмотрим работу разработанного метода на примере. На рис. 1 изображено задание, содержащее 6 ветвей вычислений, которое выполняется в системе, состоящей из 6 компьютеров с одинаковой мощностью, – здесь все компьютеры можно считать эталонными. Длительности каждой ветви задания соответственно равны: 100, 80, 70, 60, 30, 20 условных единиц времени.

На рис. 1 система выглядит в виде 6 строк, каждая из которых представляет собой компьютер. Первая ячейка указывает номера компьютера; вторая – коэффициент мощности компьютера (здесь он равен 1.00, т. к. любой из компьютеров можно считать эталонным); третья – непосредственно графически представленную распределенную ветвь задания. Распределенная ветвь задания представлена в виде двух частей (или одной части для самой длительной ветви): собственно время выполнения с номером ветви задания и время ожидания окончания выполнения всех ветвей задания (заштрихованная область). При этом первое число номера ветви задания, обозначает номер задания, второе – номер ветви. К примеру, номер 1.2 указывает, что это распределена вторая ветвь первого задания.

Под изображением системы на рис. 1 представлены краткие статистические данные о работе системы с распределенными заданиями: время работы T и суммарное время ожидания системы Δt .

| | | | |
|---------------------------|------|-----|---|
| 1 | 1.00 | 1.1 | |
| 2 | 1.00 | 1.2 | ■ |
| 3 | 1.00 | 1.3 | ■ |
| 4 | 1.00 | 1.4 | ■ |
| 5 | 1.00 | 1.5 | ■ |
| 6 | 1.00 | 1.6 | ■ |
| Статистика: | | | |
| Время работы: | | 100 | |
| Суммарное время ожидания: | | 240 | |

Рис. 1. Задание, выполняемое в системе, состоящей из компьютеров одинаковой мощности

Время выполнения такого задания в данной системе:
 $T = \max\{100, 80, 70, 60, 30, 20\} = 100$, а суммарное время ожидания окончания всех ветвей задания:

$$\Delta t = (100 - 100) + (100 - 80) + (100 - 70) + (100 - 60) + (100 - 30) + (100 - 20) = 240.$$

Рассмотрим более сложный случай. Используем вычислительную систему из 12 компьютеров с коэффициентами мощности соответственно: 0,5; 0,75; 0,8; 1; 1,2; 1,25; 1,4; 1,5; 2; 2,1; 2,1; 2,5. Запустим для выполнения в ней два задания, таких же, как и рассмотренное выше, и проведем расчеты времени работы и суммарного времени ожидания.

Согласно рассмотренному простейшему подходу, время выполнения каждой ветви первого задания рассчитывается следующим образом:

$$\begin{aligned} t_{111} &= t_{11}^e \times c_1 = 100 \times 0,5 = 50, \\ t_{122} &= t_{12}^e \times c_2 = 80 \times 0,75 = 60, \\ t_{133} &= t_{13}^e \times c_3 = 70 \times 0,8 = 56, \\ t_{144} &= t_{14}^e \times c_4 = 60 \times 1 = 60, \\ t_{155} &= t_{15}^e \times c_5 = 30 \times 1,2 = 36, \\ t_{166} &= t_{16}^e \times c_6 = 20 \times 1,25 = 25. \end{aligned}$$

Тогда время выполнения первого задания в данной системе:

$T_1 = \max\{50, 60, 56, 60, 36, 25\} = 60$, а суммарное время ожидания окончания всех ветвей первого задания:

$$\Delta t_1 = (60 - 50) + (60 - 60) + (60 - 56) + (60 - 60) + (60 - 36) + (60 - 25) = 73.$$

Для второго задания время выполнения каждой ветви имеем:

$$\begin{aligned} t_{211} &= t_{21}^e \times c_7 = 100 \times 1,4 = 140, \\ t_{222} &= t_{22}^e \times c_8 = 80 \times 1,5 = 120, \\ t_{233} &= t_{23}^e \times c_9 = 70 \times 2 = 140, \\ t_{244} &= t_{24}^e \times c_{10} = 60 \times 2,1 = 126, \\ t_{255} &= t_{25}^e \times c_{11} = 30 \times 2,1 = 63, \\ t_{266} &= t_{26}^e \times c_{12} = 20 \times 2,5 = 50. \end{aligned}$$

А время выполнения второго задания в данной системе и суммарное время ожидания окончания всех ветвей второго задания соответственно:

$$T_2 = \max\{140, 120, 140, 126, 63, 50\} = 140,$$

$$\Delta t_1 = (140 - 140) + (140 - 120) + (140 - 140) + (140 - 126) + (140 - 63) + (140 - 50) = 201.$$

В таком случае время работы системы равно 140 условных единиц времени, а суммарное время ожидания системы – 274.

Согласно разработанному методу, время выполнения первого задания в данной системе рассчитывается аналогично и так же равняется $T_1 = 60$. Время выполнения каждой j -ой ветви первого задания на каждом m -м компьютере t_{1jm} представлено в табл. 1.

Из этой таблицы выбираются значения, которые максимально приближены к T_1 , но не более его. Эти значения в таблице выделены серым цветом. Выбор производится по очереди для каждого столбца, начиная с первого. При этом никакие два значения не должны принадлежать одному и тому же столбцу или строке.

Тогда суммарное время ожидания окончания всех ветвей первого задания:

$$\Delta t_1 = (60 - 50) + (60 - 60) + (60 - 56) + (60 - 60) + (60 - 60) + (60 - 50) = 24.$$

Время выполнения каждой j -й ветви второго задания на каждом m -м компьютере t_{2jm} и выбранные значения представлены в табл. 2.

Таблица 1. Время выполнения ветвей первого задания

| Номер ветви / Номер компьютера | 1 | 2 | 3 | 4 | 5 | 6 |
|--------------------------------|-----|-----|------|-----|------|----|
| 1 | 50 | 40 | 35 | 30 | 15 | 10 |
| 2 | 75 | 60 | 52,5 | 45 | 22,5 | 15 |
| 3 | 80 | 64 | 56 | 48 | 24 | 16 |
| 4 | 100 | 80 | 70 | 60 | 30 | 20 |
| 5 | 120 | 96 | 84 | 72 | 36 | 24 |
| 6 | 125 | 100 | 87,5 | 75 | 37,5 | 25 |
| 7 | 140 | 112 | 98 | 84 | 42 | 28 |
| 8 | 150 | 120 | 105 | 90 | 45 | 30 |
| 9 | 200 | 160 | 140 | 120 | 60 | 40 |
| 10 | 210 | 168 | 147 | 126 | 63 | 42 |
| 11 | 210 | 168 | 147 | 126 | 63 | 42 |
| 12 | 250 | 200 | 175 | 150 | 75 | 50 |

Таблица 2. Время выполнения ветвей второго задания

| Номер ветви / Номер компьютера | 1 | 2 | 3 | 4 | 5 | 6 |
|--------------------------------|-----|-----|------|-----|------|----|
| 5 | 120 | 96 | 84 | 72 | 36 | 24 |
| 6 | 125 | 100 | 87,5 | 75 | 37,5 | 25 |
| 7 | 140 | 112 | 98 | 84 | 42 | 28 |
| 8 | 150 | 120 | 105 | 90 | 45 | 30 |
| 10 | 210 | 168 | 147 | 126 | 63 | 42 |
| 11 | 210 | 168 | 147 | 126 | 63 | 42 |

Тогда время выполнения второго задания в данной системе:

$T_2 = \max\{120, 100, 98, 90, 63, 42\} = 120$, а суммарное время ожидания окончания всех ветвей второго задания:

$$\Delta t_2 = (120 - 120) + (120 - 120) + (120 - 98) + (120 - 75) + (120 - 63) + (120 - 42) = 202.$$

РЕЗУЛЬТАТЫ РАБОТЫ

В результате распределения заданий в данной системе согласно разработанному методу, время работы сис-

темы равно 120 условных единиц времени, а суммарное время ожидания системы – 226.

На рис. 2, 3 очевидны преимущества предлагаемого метода:

- распределив первое задание по компьютерам системы, не смотря на то, что время выполнения одинаково, время ожидания окончания первого задания гораздо меньше, а средняя мощность компьютеров, оставшихся в распоряжении второго задания, выше;
- распределив второе задание, имеем меньшее общее время выполнения и ожидания, что и требовалось получить.

| | | | | |
|---------------------------|------|-----|---|----|
| 1 | 0.50 | 1.1 | ▨ | |
| 2 | 0.75 | 1.2 | | |
| 3 | 0.80 | 1.3 | ▨ | |
| 4 | 1.00 | 1.4 | | |
| 5 | 1.20 | 1.5 | ▨ | |
| 6 | 1.25 | 1.6 | ▨ | |
| 7 | 1.40 | | | |
| 8 | 1.50 | | | |
| 9 | 2.00 | | | |
| 10 | 2.10 | | | |
| 11 | 2.10 | | | |
| 12 | 2.50 | | | |
| Статистика: | | | | |
| Время работы: | | | | 60 |
| Суммарное время ожидания: | | | | 73 |

| | | | | |
|---------------------------|------|-----|---|-----|
| 1 | 0.50 | 1.1 | ▨ | |
| 2 | 0.75 | 1.2 | | |
| 3 | 0.80 | 1.3 | ▨ | |
| 4 | 1.00 | 1.4 | | |
| 5 | 1.20 | 1.5 | ▨ | |
| 6 | 1.25 | 1.6 | ▨ | |
| 7 | 1.40 | 2.1 | | |
| 8 | 1.50 | 2.2 | | ▨ |
| 9 | 2.00 | 2.3 | | |
| 10 | 2.10 | 2.4 | | ▨ |
| 11 | 2.10 | 2.5 | | ▨ |
| 12 | 2.50 | 2.6 | | ▨ |
| Статистика: | | | | |
| Время работы: | | | | 140 |
| Суммарное время ожидания: | | | | 274 |

Рис. 2. Пример распределения заданий с использованием существующего метода

| | | | | |
|---------------------------|------|-----|---|----|
| 1 | 0.50 | 1.1 | ▨ | |
| 2 | 0.75 | 1.2 | | |
| 3 | 0.80 | 1.3 | ▨ | |
| 4 | 1.00 | 1.4 | | |
| 5 | 1.20 | | | |
| 6 | 1.25 | | | |
| 7 | 1.40 | | | |
| 8 | 1.50 | | | |
| 9 | 2.00 | 1.5 | | |
| 10 | 2.10 | | | |
| 11 | 2.10 | | | |
| 12 | 2.50 | 1.6 | ▨ | |
| Статистика: | | | | |
| Время работы: | | | | 60 |
| Суммарное время ожидания: | | | | 24 |

| | | | | |
|---------------------------|------|-----|---|-----|
| 1 | 0.50 | 1.1 | ▨ | |
| 2 | 0.75 | 1.2 | | |
| 3 | 0.80 | 1.3 | ▨ | |
| 4 | 1.00 | 1.4 | | |
| 5 | 1.20 | 2.1 | | |
| 6 | 1.25 | 2.4 | | ▨ |
| 7 | 1.40 | 2.3 | | ▨ |
| 8 | 1.50 | 2.2 | | |
| 9 | 2.00 | 1.5 | | |
| 10 | 2.10 | 2.6 | | ▨ |
| 11 | 2.10 | 2.5 | | ▨ |
| 12 | 2.50 | 1.6 | ▨ | |
| Статистика: | | | | |
| Время работы: | | | | 120 |
| Суммарное время ожидания: | | | | 226 |

Рис. 3. Пример распределения заданий с использованием предлагаемого метода

ВЫВОДЫ

Таким образом, в данной статье был рассмотрен подход к распределению ветвей параллельного задания между компьютерами гетерогенной вычислительной среды. Был продемонстрирован пример оптимизации распределения, используя рассмотренный метод, эффективно решающий поставленную задачу.

СПИСОК ЛИТЕРАТУРЫ

1. Linux. Кластер. Практическое руководство по параллельным вычислениям [Электронный ресурс] / авт. курса Ю. Сбитнев – Электрон. текстовые дан. – [Волгоград?], [199–?] – Режим доступа: <http://cluster.linux-ekb.info>, свободный. – Загл. с экрана.
2. Шпаковский Г. И. Реализация параллельных вычислений: кластеры, многоядерные процессоры, грид, квантовые компьютеры / Г. И. Шпаковский. – Минск : БГУ.– 2010. – 155 с.
3. Технология запуска параллельных задач в различных распределенных средах / [В. М. Волохов, Д. А. Варламов, Н. Ф. Сурков, А. В. Пивушков, А. В. Волохов] // Распределенные вычисления и Грид-технологии в науке и образовании : труды 4-й междунар. конф. (Дубна, 28 июня – 3 июля, 2010 г.). – Дубна : ОИЯИ. – С. 329–334.

Стаття надійшла до редакції 20.05.2011.

Сметанін Р. І., Тягунова М. Ю.

ПІДХІД ДО РОЗПОДІЛУ ПАРАЛЕЛЬНИХ ГІЛОК ЗАВДАННЯ В ОБЧИСЛЮВАЛЬНІЙ СИСТЕМІ

Розглянуто проблему розподілу завдань в обчислювальній системі для випадку, коли тривалість виконання паралельних гілок завдання, які необхідно розподілити, різна. Розроблено метод для підвищення ефективності розподілу. Показано, що запропонований метод є ефективним, оскільки, використовуючи його, час виконання завдань помітно скорочується.

Ключові слова: обчислювальна система, grid, паралельні обчислення, гілки завдання, розподілення завдань, час виконання завдання, час роботи системи, час очікування.

Smetanin R. I., Tyagunova M. Yu.

APPROACH TO THE DISTRIBUTION OF PARALLEL TASK THREADS IN COMPUTING SYSTEMS

The problem of the tasks distribution in the computer system with different execution time of parallel threads, which must be distributed, is considered. Method to improve distribution efficiency is developed. Method is effective, because it much reduces execution time of task.

Keywords: computing system, grid, parallel computations, task threads, tasks distribution, task execution time, system work time, waiting time.