

ЭФФЕКТИВНОСТЬ РАСПРЕДЕЛЕНИЯ ЗАДАНИЙ В КОМПЬЮТЕРНЫХ СИСТЕМАХ

В данной статье приведены результаты сравнения времени работы разработанного гибридного алгоритма с существующими алгоритмами FCFS и Backfill. Выявлена зависимость между соотношением числа заданий к числу компьютеров и временем работы алгоритмов. Приведена формула расчета и рассчитана эффективность разработанного гибридного метода для конкретного примера.

The article describes a comparison of the running time of developed hybrid algorithm with the existing ones (FCFS and Backfill). Revealed a relationship between the ratio of the number of jobs to the number of computers and running time of the algorithms. The calculation formula and computation of the efficiency of the developed hybrid algorithm for a specific example are presented.

Введение

В настоящее время применение распределенных и параллельных компьютерных систем и вычислений становится одним из основных условий успешного решения сложных научных и практических задач, основой создания распределенных и виртуальных организаций, средством для решения глобальных задач человечества и фундаментом во всеобщей информатизации общества. Сегодня участие в международных программах становится стратегическим вопросом информационного развития государств. Украина также стоит на пути научно-технической интеграции в мировое и европейское пространство, о чем говорят ряд проектов, в которых Украина является полноправным участником. Примером тому является создание национальной GRID-инфраструктуры – UkrGRID.

Для построения распределенных и параллельных вычислений и их эффективного использования необходимо иметь мощную программную и аппаратную базу, позволяющую управлять распределением вычислительных заданий. Решение этой задачи возложено на планировщиков заданий, работа которых базируется на некотором алгоритме распределения заданий. От выбранного алгоритма зависит эффективность работы того или иного планировщика, а следовательно и эффективность управления вычислительными ресурсами.

1. Постановка задачи

Проведенный в работе [1] анализ систем, методов и алгоритмов показал, что известные методы диспетчеризации не учитывают конкретную длительность выполнения задания на определенном компьютере, и качество оптимизации во многом зависит от порядка возникновения задания в системе для распределения. Поэтому для повышения эффективности диспетчеризации необходимо разработать методы распределения заданий, позволяющие учитывать длительность выполнения заданий и при использовании которых результат распределения не склонен к зависимости от порядка поступления заданий хотя бы в пределах некоторого интервала времени. Такой метод – гибридный метод диспетчеризации заданий – был разработан и подробно описан в работе [2]. Однако, не был проведен анализ, на сколько данный метод эффективнее по сравнению с существующими.

Современные планировщики заданий в составе промежуточного программного обеспечения (middleware), к примеру, gLite, чаще всего используют для работы такие алгоритмы распределения как FCFS и Backfill.

Поскольку данные алгоритмы являются наиболее популярными, то была поставлена задача сравнить результаты работы разработанного гибридного метода с данными алгоритмами, и таким образом, определить, какой из алгоритмов является наиболее эффективным.

2. Проведение экспериментов

Для сравнения алгоритмов FCFS, Backfill и гибридного, реализующего гибридный метод, были проведены следующие эксперименты.

Для этого были взяты различные комбинации числа заданий и числа

компьютеров. Так, к примеру, в начале количество компьютеров было постоянным и равным пяти, а количество заданий менялось от 10 до 200.

Таблица 1. Результаты сравнения реализованных программно алгоритмов FCFS, Backfill и гибридного

Число компьютеров, n	Число заданий, m	$\frac{m}{n}$	суммарное время			максимальное время		
			гибридный алгоритм	fcfs, backfill	Эффективность, %	гибридный алгоритм	fcfs, backfill	Эффективность, %
5	10	2.00	138.27	148.54	7.43	29.04	36.30	25.00
5	15	3.00	248.46	309.28	24.48	50.74	89.44	76.27
5	30	6.00	499.69	537.36	7.54	100.62	128.34	27.55
5	30	6.00	499.69	549.22	9.91	100.62	135.00	34.17
5	40	8.00	637.96	652.34	2.25	129.00	140.00	8.53
5	50	10.00	799.74	853.30	6.70	161.82	190.06	17.45
5	60	12.00	1002.15	1042.53	4.03	201.24	228.78	13.69
5	70	14.00	1140.42	1147.92	0.66	229.19	234.78	2.44
5	80	16.00	1302.20	1337.46	2.71	262.26	279.00	6.38
5	90	18.00	1499.88	1534.09	2.28	301.32	329.22	9.26
5	100	20.00	1640.11	1683.71	2.66	329.81	363.35	10.17
5	110	22.00	1801.89	1838.74	2.05	362.70	379.44	4.62
5	120	24.00	1999.57	2037.55	1.90	401.76	429.66	6.94
5	130	26.00	2137.84	2170.77	1.54	430.00	450.00	4.65
5	140	28.00	2299.63	2361.50	2.69	462.68	491.92	6.32
5	150	30.00	2502.03	2523.85	0.87	502.20	525.00	4.54
5	160	32.00	2639.72	2658.09	0.70	530.00	547.82	3.36
5	170	34.00	2802.09	2836.22	1.22	562.65	577.17	2.58
5	180	36.00	2999.77	3028.95	0.97	602.58	620.49	2.97
5	190	38.00	3139.41	3183.36	1.40	630.38	665.21	5.53
5	200	40.00	3301.78	3344.51	1.29	663.06	691.92	4.35
10	200	20.00	3301.78	3349.07	1.43	331.96	355.00	6.94
20	200	10.00	3275.42	3483.84	6.36	166.98	206.83	23.87
30	200	6.67	3263.72	3531.22	8.20	111.80	140.00	25.22
40	200	5.00	3276.55	3505.26	6.98	83.85	108.90	29.87
50	200	4.00	3269.63	3663.73	12.05	67.08	111.80	66.67
60	200	3.33	3272.96	3686.35	12.63	55.90	95.03	70.00
70	200	2.86	3193.65	3673.19	15.02	49.02	85.00	73.40
80	200	2.50	3090.53	3753.38	21.45	43.86	83.70	90.83
90	200	2.22	3135.52	3768.70	20.19	38.70	75.00	93.80
100	200	2.00	3109.88	3742.58	20.34	35.26	75.00	112.71
57	567	9.95	4313348.76	4456768.53	3.33	76188.44	99468.81	30.56
57	567	9.95	4320363.65	4456768.53	3.16	75845.00	99468.81	31.15
50	1000	20.00	16609.75	16998.06	2.34	335.00	373.86	11.60
50	100	2.00	1530.24	1893.86	23.76	35.26	65.00	84.34
50	400	8.00	6597.91	6999.87	6.09	134.31	167.70	24.86
50	2000	40.00	33289.56	33596.41	0.92	669.08	714.24	6.75
100	2000	20.00	33287.73	33956.43	2.01	335.00	374.53	11.80
234	2345	10.02	7647598.14	7859909.15	2.78	32698.52	44806.25	37.03
Среднее значение эффективности					6.52	28.41		

Далее, константным было взято количество заданий, равное 200, а количество компьютеров менялось от 5 до 100. После этого, эксперименты проводились с большим количеством заданий и компьютеров, но по такому же принципу (с константным числом компьютеров и переменным числом заданий или переменным числом компьютеров и числом константным заданий).

3. Результаты сравнения алгоритмов

Основные результаты сравнения реализованных программно алгоритмов FCFS, Backfill и разработанного гибридного алгоритма для однопроцессорных заданий приведены в табл. 1.

В таблице приведены следующие обозначения:

n – число компьютеров системы;

m – число поступающих для распределения заданий.

Стоит отметить, что при данном эксперименте результаты работы алгоритмов FCFS и Backfill совпали, поэтому были занесены в одну таблицу. Такое совпадение объясняет условием задачи, что каждое задание для решения требует только одного процессора.

Эффективность вычисляется по формуле:

$$F = \frac{(T_e - T_h)}{T_h} * 100\%, \quad (4.1)$$

где F – эффективность применения гибридного метода;

T_e – время, максимальное или суммарное, полученное при использовании для диспетчеризации сравниваемого алгоритма;

T_h – время, максимальное или суммарное, полученное при использовании для диспетчеризации гибридного метода.

Зависимость эффективности применения гибридного метода от соотношения количества заданий к количеству компьютеров распределенной компьютерной системы отображено на графике (рис. 1).

На графике пунктирной линией обозначено аппроксимированное значение суммарного времени выполнения заданий, сплошной – аппроксимированное значение максимального времени выполнения заданий на одном из компьютеров распределенной системы.

Исходя из результатов, представленных в таблице 1 и на рис. 1, можно сделать вывод, что наиболее эффективно применение гибридного метода в случае, когда количество

заданий превышает количество компьютеров не более чем в 20 раз.

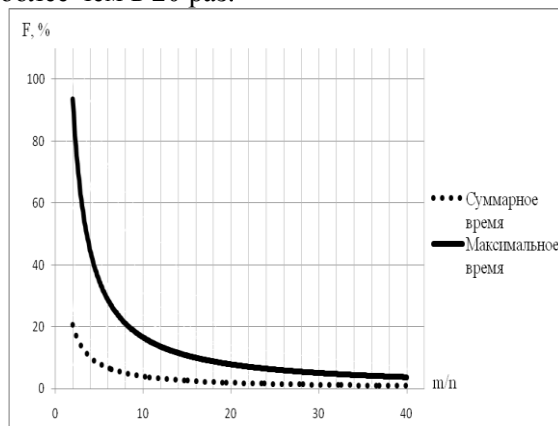


Рис. 1. Зависимость эффективности применения гибридного метода от соотношения количества заданий к количеству компьютеров распределенной компьютерной системы

4. Эффективность гибридного метода

Приведем ниже в более расширенном виде результаты работы сравниваемых трех алгоритмов (табл. 2 – 4) на следующем примере.

Пусть для распределения в компьютерную систему из 5 компьютеров поступило 15 заданий. Пусть массив коэффициентов мощностей компьютеров:

0,86 3,63 5,00 5,58 5,59.

Пусть массив длительностей выполнения заданий:

2	3	5	5	6	6
7	7	7	8	8	8
9	9	10.			

Результаты распределения заданий на компьютеры представим в виде таблиц 2 – 4.

В табл. 2 – 4 номер столбца соответствует номеру компьютера, а номер строки – номеру задания, светло-серым цветом отмечены номера заданий и мощности компьютеров; темно-серым – задания, распределенные на определенный компьютер; в нижней строке показана суммарная длительность выполнения заданий после процесса диспетчеризации по каждому из компьютеров и сумма выполнения всех поступивших заданий по всем компьютерам.

Как видно из результатов работы алгоритмов, порядок поступления заданий в случае использования алгоритмов FCFS и Backfill очень сильно влияет на эффективность распределения заданий. В гибридном алгоритме данный фактор не влияет на результаты, поскольку в алгоритме предусмотрена автоматическая сортировка

заданий и мощностей компьютеров до реализации непосредственной диспетчеризации заданий, собранных за определенный промежуток времени.

Эффективность для конкретно взятого примера составляет 6.26% по суммарному времени и 20.96% по максимальному времени выполнения заданий в лучшем случае реализации алгоритмов-аналогов и 24.48% по суммарному времени и 76.27% по максимальному времени выполнения заданий в худшем случае.

Таблица 2. Результаты работы алгоритмов FCFS и Backfill (задания по возрастанию)

FCFS и Backfill (в порядке возрастания)					
1	1,72	7,25	10	11,16	11,18
2	2,58	10,89	15	16,74	16,76
3	4,3	18,13	25	27,90	27,93
4	4,3	18,13	25	27,90	27,93
5	5,16	21,78	30	33,47	33,54
6	5,16	21,78	30	33,47	33,54
7	6,02	25,41	35	39,04	39,1
8	6,02	25,41	35	39,06	39,1
9	6,02	25,41	35	39,06	39,1
10	6,88	29,01	40	44,64	44,72
11	6,88	29,01	40	44,64	44,72
12	6,88	29,01	40	44,64	44,72
13	7,74	32,67	45	50,22	50,31
14	7,74	32,67	45	50,2	50,31
15	8,6	36,26	50	55,78	55,86
	0,86	3,63	5	5,58	5,59
309,2	40,42	36,30	65	78,12	89,44

Таблица 3. Результаты работы алгоритмов FCFS и Backfill (задания по убыванию)

FCFS и Backfill (в порядке убывания)					
1	8,6	36,26	50	55,78	55,86
2	7,74	32,67	45	50,2	50,31
3	7,74	32,67	45	50,2	50,31
4	6,88	29,01	40	44,64	44,72
5	6,88	29,01	40	44,64	44,72
6	6,88	29,01	40	44,64	44,72
7	6,02	25,41	35	39,04	39,1
8	6,02	25,41	35	39,04	39,1
9	6,02	25,41	35	39,04	39,1
10	5,16	21,78	30	33,47	33,54
11	5,16	21,78	30	33,47	33,54
12	4,3	18,13	25	27,89	27,93
13	4,3	18,13	25	27,89	27,93
14	2,58	10,88	15	16,74	16,76
15	1,72	7,25	10	11,16	11,18
	0,86	3,63	5	5,58	5,59
264,033	47,3	54,45	45	61,38	55,90

Таблица 4. Результаты работы гибридного алгоритма

Гибридный алгоритм					
1	1,72	7,25	10	11,16	11,18
2	2,58	10,89	15	16,74	16,76
3	4,3	18,13	25	27,90	27,93
4	4,3	18,13	25	27,90	27,93
5	5,16	21,78	30	33,47	33,54
6	5,16	21,78	30	33,47	33,54
7	6,02	25,41	35	39,04	39,1
8	6,02	25,41	35	39,06	39,1
9	6,02	25,41	35	39,06	39,1
10	6,88	29,01	40	44,64	44,72
11	6,88	29,01	40	44,64	44,72
12	6,88	29,01	40	44,64	44,72
13	7,74	32,67	45	50,2	50,31
14	7,74	32,67	45	50,2	50,31
15	8,6	36,26	50	55,78	55,86
	0,86	3,63	5	5,58	5,59
248.46	50,74	47,19	50	50,22	50,31

Выводы

Проведенные эксперименты дают основание сделать вывод, что применение гибридного метода для однопроцессорных заданий является более эффективным по сравнению с алгоритмами FCFS и Backfill.

СПИСОК ЛИТЕРАТУРЫ

1. Юрич М.Ю. Анализ систем и методов балансировки нагрузки вычислительных систем /М.Ю. Юрич // Науковий вісник Чернівецького національного університету. Комп'ютерні системи та компоненти. – Випуск 1. – Т. 1. – 2010. – С. 14-19.
2. Тягунова М.Ю. Гибридный метод диспетчеризации заданий в распределенных компьютерных системах // Науковий вісник Чернівецького національного університету. Комп'ютерні системи та компоненти. – Випуск 1. – Т. 2. – 2011. – с. 27-30.

М.Ю. Tyagunova. Efficient of tasks distribution in computer systems

М.Ю. Тягунова. Ефективність розподілення завдань у комп'ютерних системах

У даній статті наведено результати порівняння часу роботи розробленого гібридного алгоритму з існуючими алгоритмами FCFS і Backfill. Виявлено залежність між співвідношенням числа завдань до числа комп'ютерів і часом роботи алгоритмів. Наведена формула розрахунку і розрахована ефективність розробленого гібридного методу для конкретного прикладу.

