

ПЛАНИРОВАНИЕ ИСПОЛНЕНИЯ НАБОРОВ КОМПОЗИТНЫХ ПРИЛОЖЕНИЙ С МЯГКИМИ СРОКАМИ ЗАВЕРШЕНИЯ ВО ВРЕМЕННЫХ ОКНАХ ГЕТЕРОГЕННЫХ ВЫЧИСЛИТЕЛЬНЫХ СРЕД

К. О. БОЧЕНИНА, Н. А. БУТАКОВ, А. В. БУХАНОВСКИЙ

*Университет ИТМО, 197101, Санкт-Петербург, Россия
E-mail: k.bochenina@gmail.com*

Повышение эффективности загрузки неоднородных вычислительных ресурсов при удовлетворении требований пользователей к срокам получения результатов расчета требует применения специализированных алгоритмов планирования исполнения задач. Такие алгоритмы должны учитывать как специфику распределенной среды (уровень утилизации, типы ресурсов), так и возможные взаимосвязи по данным между частями расчетного сценария. Предложены алгоритмы реализации позадачного, стадийного и кластеризационного подходов к планированию наборов композитных приложений (КП) с мягкими сроками завершения в условиях частичной доступности вычислительных ресурсов. Исследована сравнительная эффективность разработанных алгоритмов при планировании исполнения наборов КП на гетерогенных ресурсах. Результаты экспериментов демонстрируют, что: а) использование информации о доступности ресурсов и сроках завершения задач позволяет существенно повысить качество планов; б) кластеризационный подход превосходит стадийный и позадачный по интегральному критерию эффективности планирования.

Ключевые слова: алгоритмы планирования, облачные среды, композитные приложения, планирование загрузки ресурсов, временные окна

Введение. При проведении многоэтапных вычислительно сложных расчетов зачастую возникает необходимость использования различных программных сервисов в рамках единого численного эксперимента. Общепринятым формализмом логического сопряжения пакетов специализированного программного обеспечения (ПО) [1] является *композитное приложение* (КП, workflow, WF) — ориентированный ациклический граф задач. Как правило, КП выполняются в неоднородных (гетерогенных) распределенных средах, это позволяет использовать территориально разнесенные вычислительные мощности в рамках единого пула ресурсов, а также требует применения специальных алгоритмов планирования приложений. Проанализировав особенности практического использования КП, можно выделить класс задач планирования, характеризующихся доступностью информации о сроках завершения расчетов КП, а также о временных окнах доступности ресурсов. Например, при многопользовательском разделении вычислительной среды может возникнуть необходимость заполнения временных окон, оставшихся после планирования высокоприоритетных задач, менее приоритетными КП, для которых заданы желаемые сроки получения результатов.

Традиционно целью статического планирования является минимизация времени исполнения наборов КП в частично доступной аппаратной и программной гетерогенной вычислительной

среде при учете времени передачи данных и детерминированных оценках времени исполнения отдельных задач [2]. Среди подходов к ее достижению можно выделить [3]:

— представление каждого КП как одной задачи, которая может быть решена на произвольном числе процессоров [4]; подход применим, когда все подзадачи КП могут быть выполнены на одних и тех же ресурсах;

— поочередное планирование КП/частей КП (алгоритмы FPFT и FPCT [2]);

— объединение всех задач в одно КП (алгоритмы MWGS4 и MWGS2).

При работе с несколькими КП с разными критериями эффективности и ограничениями необходимо использовать интегральные метрики оценки качества формируемых планов, например, длительность выполнения набора КП, среднее замедление, среднее время ожидания [2, 3, 5]. Это позволяет минимизировать время выполнения набора КП без учета требований пользователей к качеству планов.

Целью настоящего исследования стала разработка алгоритмов, реализующих различные подходы к планированию исполнения наборов КП с мягкими сроками завершения и сравнительная оценка их эффективности; в алгоритме учитываются критерии оценки качества планов как степени нарушения сроков завершения и справедливости распределения ресурсов между КП.

Постановка задачи. Гетерогенная вычислительная среда представляет собой набор ресурсов — узлов, объединяющих на логическом уровне два типа компонентов: аппаратные (CPU/GPU, RAM, HDD) и программные (системное и прикладное ПО). Предполагается, что доступен набор ресурсов R различных типов, задана скорость передачи данных между ними, а также известны границы зарезервированных временных окон.

Пусть необходимо спланировать набор из N КП. i -е КП набора представляет собой ориентированный ациклический граф из Nt_i задач ($i=1, \dots, N$). Предполагается, что известны объем данных, передаваемых между задачами КП, оценки времени выполнения задач КП на ресурсах различных типов. Для каждого КП задан мягкий срок завершения d_i . Целью планирования является поиск отображения множества задач на множество ресурсов с указанием времени запуска каждой задачи s_{ij} , $0 \leq s_{ij} \leq \max\{d_i\} = T$.

Штраф для i -го КП будем рассчитывать как $F_i = \max\{e_i - d_i, 0\}$, где e_i — время окончания выполнения этого КП в плане. Так как предполагается, что задачи не могут запускаться по истечении периода планирования T , процедура планирования может предоставлять неполные планы — планы, содержащие назначения не для всех имеющихся задач. При расчете штрафов для неполных планов предполагается, что задачи, не вошедшие в план, выполняются в соответствии со структурой КП на единственном доступном ресурсе с усредненными оценками времени выполнения и передачи данных. Максимальный штраф рассчитывается исходя из тех же соображений в предположении, что ни одна задача не получила назначения в плане.

Средний штраф для плана будем рассчитывать как:

$$F_{\text{avg}} = \frac{1}{N} \sum_{i=1}^N \frac{F_i}{F_{\max i}}. \quad (1)$$

Штрафы позволяют оценить степень соблюдения мягких сроков завершения работ. Для учета справедливости распределения ресурсов между разными КП введем следующую метрику:

$$J = 1 - \frac{\max_{i,j} |F_i - F_j|}{F_{\max i}}. \quad (2)$$

Интегральный критерий качества плана рассчитывается как:

$$U = \rho_F (1 - F_{\text{avg}}) + \rho_J J. \quad (3)$$

Здесь $\rho_F, \rho_J \in [0,1]$, $\rho_F + \rho_J = 1$ — весовые коэффициенты, определяющие сравнительную значимость критериев.

Алгоритмы планирования наборов КП. Суть процесса планирования заключается в определении порядка доступа кластеров (задач) к ресурсам, т.е. в приоритизации задач. Возможно применение одного из видов планирования:

I) позадачное — все КП объединяются, после чего план формируется с использованием некоторого алгоритма планирования для одиночного КП;

II) на основе приоритизации КП — КП набора упорядочиваются, после чего планируется их поочередное исполнение;

III) на основе кластеризации — подход, объединяющий оба вида планирования. Основная идея заключается в разбиении всего множества задач на подмножества, которые затем последовательно исполняются.

Определим способы реализации указанных подходов для планирования исполнения наборов КП с мягкими ограничениями по срокам завершения задач.

I. *Позадачное планирование.* Предлагаемый алгоритм состоит из трех этапов: определение крайних сроков завершения отдельных задач; построение приоритетной очереди задач с учетом крайних сроков завершения и взаимосвязей задач; последовательное исполнение задач из очереди.

Крайние сроки завершения задач d_{ij} определяются следующим образом. Пусть \bar{t}_{ij} — усредненная оценка времени выполнения j -й задачи i -го КП. Накопленное время выполнения w_{ij} будем определять для стартовых задач КП как $w_{ij} = \bar{t}_{ij}$, а для остальных $w_{ij} = \bar{t}_{ij} + \max_{k \in P_{ij}} [w_{ik} + \bar{c}_{ikj}]$, где \bar{c}_{ikj} — среднее время передачи данных между k -й и j -й задачами i -го КП, P_{ij} — набор индексов задач-родителей j -й задачи i -го КП. Считаем, что задаче j^* с максимальным значением $w_{ij^*} = \max_j w_{ij}$ соответствует срок завершения d_i . Для остальных задач сроки завершения задач определяются как $d_{ij} = (w_{ij} / w_{ij^*}) d_i, j \neq j^*$.

II. *Планирование на основе приоритизации КП.* В настоящей работе в качестве критерия приоритизации КП предлагается использовать резервное время $RT = d_i - te_i$ и определять порядок следования кластеров с использованием стадийной схемы, включающей N стадий планирования. На первой стадии определяется набор планов независимо для всех КП в условиях монопольного доступа к ресурсам. После этого приоритетный доступ к ресурсам получает КП с минимальным значением резервного времени. Затем процесс повторяется итеративно для оставшихся КП.

Подходы I и II были впервые предложены в работе [6], где они использовались для обеспечения максимальной эффективности использования совокупности ресурсов набором КП с ограничениями по срокам завершения задач, причем эксперименты на синтетических данных показали преимущество подхода II по среднему резервному времени, эффективности, доли нарушенных крайних сроков завершения.

III. *Кластеризационный подход* в общем случае состоит из двух этапов.

— Кластеризация

1) каждая задача выделяется в отдельный кластер. Для каждого кластера, как и при позадачном планировании, определяются вес и срок завершения. После этого кластеры упорядочиваются по возрастанию сроков завершения. Процесс кластеризации основан на нахожде-

нии баланса между средней разностью весовых коэффициентов кластеров dw_{avg} и суммой длин кластеров l (под длиной кластера понимается разность между сроком завершения и сроком старта кластера);

2) на каждой итерации (начиная с кластера 2 и заканчивая $N-1$) производится попытка слияния текущего кластера с последующим или с предыдущим. Кластеры сливаются, если по итогам слияния уменьшается значение метрики $f = dw_{\text{avg}}/dw_{\text{max}} + l_{\text{avg}}/l_{\text{max}}$. Если f уменьшается, выбирается вариант слияния, обеспечивающий минимальное значение f . Здесь l_{max} — сумма длин кластеров в случае, когда каждая задача представляет собой отдельный кластер, dw_{max} — разница весовых коэффициентов кластера, объединяющего все задачи КП, и кластера нулевого веса. Процесс итеративного слияния кластеров продолжается до тех пор, пока при проходе от кластера 2 до кластера $N-1$ не будет произведено ни одного слияния.

— Упорядочивание и планирование кластеров:

1) построение орграфа зависимостей кластеров для каждого КП;

2) определение набора C_i , из которых выбираются кластеры на этом шаге планирования. Сначала в данный набор помещаются кластеры, не имеющие родителей в ориентированных графах соответствующих КП, далее — кластеры, родители которых уже были спланированы;

3) определение номера КП i^* , кластер из которого будет спланирован на текущем шаге, в соответствии с минимальным резервным временем;

4) выбор кластера для планирования в соответствии с максимальным значением следующей метрики:

$$M(C_{i^*}) = \rho_{WL} \frac{1}{wl_{\text{max}}} \frac{w(C_{i^*})}{l(C_{i^*})} + \rho_D \frac{d(C_{i^*})}{d_{\text{min}}}, \quad (4)$$

где wl_{max} — максимальное отношение веса кластера к его длине для кластеров набора C_{i^*} , d_{min} — минимальный срок завершения для кластеров того же набора; ρ_{WL} и ρ_D — весовые коэффициенты;

5) планирование задач выбранного кластера, фиксация занятых временных интервалов, обновление срока завершения текущего кластера и сроков старта и завершения всех следующих кластеров, удаление спланированного кластера из набора C_{i^*} .

Шаги 2—5 повторяются до тех пор, пока все кластеры не будут спланированы.

Исследование влияния учета информации о доступности ресурсов на качество планов. Для оценки того, насколько учет временных окон влияет на улучшение качества планов, алгоритм позадачного планирования сравнивался с эвристикой Min-Min, основанной на тех же принципах, но не берущей в расчет информацию о зарезервированных окнах при планировании. Эксперименты проводились на синтетических данных, полученных с помощью генератора КП DAGGEN и генератора ресурсов, разработанного авторами настоящей статьи [6]. Размер набора КП менялся от 20 до 100 с шагом 20. В каждый набор входили случайно выбранные КП с $Nt=5, 10, 20$ или 50 задач. Базовая длина периода планирования T_0 была установлена равной 24 часам. Для временных окон варьировалась доля времени, занятого на ресурсах ($p=0,25, 0,5, 0,75$) и число окон на ресурс ($w=1-10$). Сроки завершения всех КП были установлены равными $T = T_0 + pT_0$. Для каждой комбинации этих трех параметров (N, Nt, p) было произведено пять запусков процедуры планирования с временными окнами, случайно назначенными в диапазоне $[0, T]$. Результаты (интегральный критерий, рис. 1, а) иллюстрируют, что использование информации о временных окнах определяет преимущество по-

задачного подхода (Task) в среднем на 18 %, что, прежде всего, обусловлено неравномерностью распределения штрафов (рис. 1, б) при использовании Min-Min.

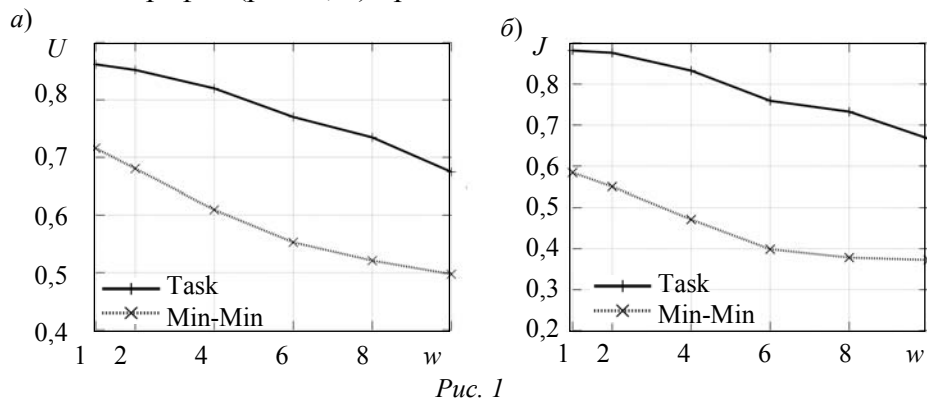


Рис. 1

Исследование эффективности алгоритмов планирования наборов КП с мягкими ограничениями по срокам завершения. Эффективность различных подходов к планированию определялась по значениям интегрального критерия (3) и среднего резервного времени. Эксперименты проводились для тестовых наборов КП Cybershake (рис. 2, а, б) и Montage (рис. 2, в, г) размером от 50 до 1000 задач (на рис. 2 приведены результаты для 20 задач). Минимальный мягкий срок завершения для каждого КП набора варьировался для шагов $n=1—12$ по формуле $d = l_{\min} + (n-1)cl_{\min}$, где l_{\min} — срок, достаточный для выполнения одного КП на заданном наборе ресурсов без срыва сроков, $c = 0,125$ — коэффициент, определяющий рост d с увеличением шага. Усредненные результаты (рис. 2) демонстрируют, что кластеризационный алгоритм (III) позволяет достичь наилучших значений как интегрального критерия, так и резервного времени для всех тестовых значений, начиная с $n = 4$. Наихудшие значения получены с использованием алгоритма позадачного планирования (I), что подтверждает результаты работы [6].

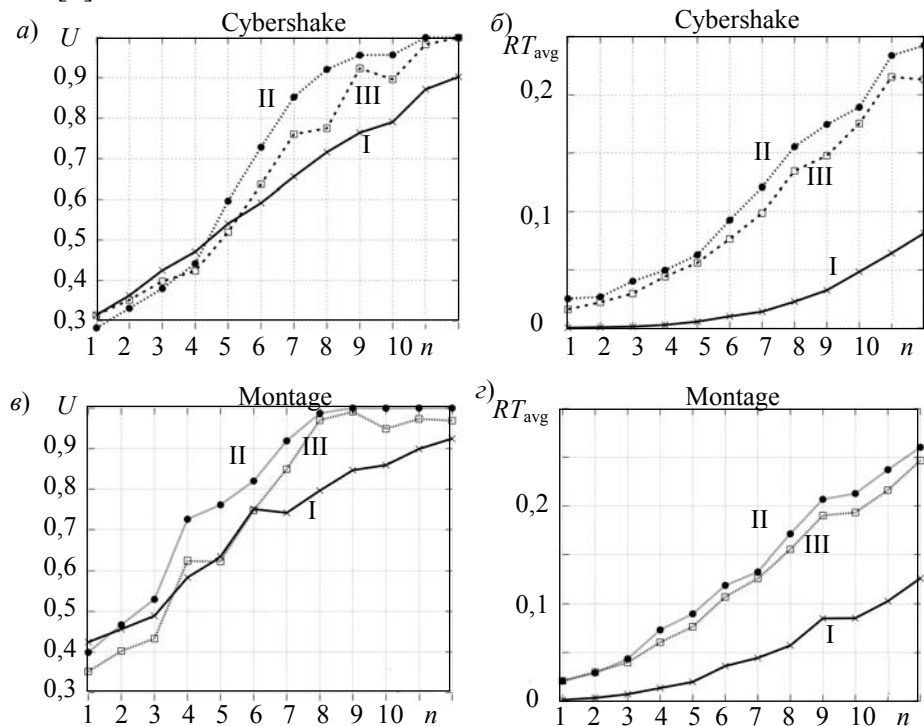


Рис. 2

Для набора Cybershake максимальное преимущество алгоритма III по сравнению с алгоритмом II составляет 11 %, по сравнению с I — 20 %, среднее преимущество (для $n > 4$) — 7 и 17 % соответственно. Алгоритмы III и II дают близкие значения среднего резервного

времени, максимальное преимущество по сравнению с алгоритмом I составляет 6 и 5 % соответственно. Для набора Montage сравнительное преимущество алгоритма III составляет по интегральному критерию — 13 % по сравнению с II и 9 % по сравнению с I, по среднему резервному времени — 11 и 20 % соответственно.

Заключение. В настоящей работе сравнивается эффективность позадачного, стадийного и кластеризационного подходов к статическому планированию наборов композитных приложений с мягкими сроками завершения во временных окнах вычислительных гетерогенных сред. Экспериментальные исследования продемонстрировали целесообразность учета информации о периодах доступности ресурсов и преимущество кластеризационного алгоритма над стадийной схемой и алгоритмом позадачного планирования. Так как вычислительная сложность стадийного алгоритма составляет $O(N^2)$, его использование целесообразно для наборов ограниченного размера, содержащих КП с большим числом задач, в то время как кластеризационный алгоритм, имеющий сложность $O(NNt^2)$, необходимо применять для объемных наборов КП со сравнительно малым числом задач.

Работа выполнена при поддержке проекта Российского научного фонда, соглашение № 14-21-00137.

СПИСОК ЛИТЕРАТУРЫ

1. Goble C., Roure D. De. The Impact of Workflow Tools on Data-centric Research // Fourth Paradig. Data-intensive Sci. Discov. / Ed. by T. Hey, S. Tansley, K. Tolle. Microsoft Research, 2009. P. 137—145.
2. Zhao H., Sakellariou R. Scheduling Multiple DAGs onto Heterogeneous Systems // Parallel Distrib. Process. Simp. IPDPS'06. Washington: IEEE Computer Society, 2006. P. 14.
3. Bittencourt L. F., Madeira E. R. M. Towards the Scheduling of Multiple Workflows on Computational Grids // J. Grid Comput. 2009. Vol. 8, N 3. P. 419—441.
4. Masko L. et al. Scheduling moldable tasks for dynamic SMP clusters in SoC technology, parallel processing and applied mathematics // PPAM—2005. Poznań, Poland, September 2005. Springer, Berlin, 2005. Vol. 3911. P. 879—887.
5. Hiraes-Carbajal A. et al. Multiple Workflow Scheduling Strategies with User Run Time Estimates on a Grid // J. Grid Comput. 2012. Vol. 10, N 2. P. 325—346.
6. Bochenina K. A Comparative Study of Scheduling Algorithms for the Multiple Deadline-constrained Workflows in Heterogeneous Computing Systems with Time Windows // Procedia Comput. Sci. Elsevier Masson SAS. 2014. Vol. 29. P. 509—522.

Сведения об авторах

Клавдия Олеговна Боченина

— канд. техн. наук; Университет ИТМО; НИИ наукоемких компьютерных технологий; младший научный сотрудник;
E-mail: k.bochenina@gmail.com

Николай Алексеевич Бутаков

— канд. техн. наук; Университет ИТМО; НИИ наукоемких компьютерных технологий; младший научный сотрудник;
E-mail: alipoov.nb@gmail.com

Александр Валерьевич Бухановский

— д-р техн. наук; Университет ИТМО; мегафакультет трансляционных информационных технологий; директор;
E-mail: avb_mail@mail.ru

Рекомендована НИИ НКТ

Поступила в редакцию
25.11.16 г.

Ссылка для цитирования: Боченина К. О., Бутаков Н. А., Бухановский А. В. Планирование исполнения наборов композитных приложений с мягкими сроками завершения во временных окнах гетерогенных вычислительных сред // Изв. вузов. Приборостроение. 2017. Т. 60, № 4. С. 295—301.

**SCHEDULING SETS OF COMPOSITE APPLICATIONS WITH SOFT DEADLINES
IN TIME WINDOWS OF HETEROGENEOUS COMPUTING ENVIRONMENTS****K. O. Bochenina, N. A. Butakov, A. V. Boukhanovsky***ITMO University, 197101, St. Petersburg, Russia**E-mail: k.bochenina@gmail.com*

The loading performance of heterogeneous computing resources under user deadline requirements to results of calculation necessitates the use of specialized planning algorithms of the tasks execution. Such algorithms should consider both the specifics of the distributed environment (the level of utilization, types of resources) and possible data relationship between parts of the design scenario. The proposed algorithms implement per-task, phasic and clusterization approaches to planning sets of composite applications with a soft deadline requirements in conditions of partial availability of computing resources. Comparative efficiency of the developed algorithms for scheduling sets of the composite applications on heterogeneous resources is investigated. Presented experimental results demonstrate that: a) the use of information about the availability of resources and the tasks completion timing can significantly improve the scheduling quality; b) clusterization approach is superior to stage and per-task according to integral criterion of scheduling efficiency.

Keywords: *scheduling algorithm, cloud media, composite applications, resource load scheduling, time windows*

Data on authors

- Klavdiya O. Bochenina** — PhD; ITMO University, eScience Research Institute; Junior Researcher; E-mail: k.bochenina@gmail.com
- Nikolay A. Butakov** — PhD; ITMO University, eScience Research Institute; Junior Researcher; E-mail: alipoov.nb@gmail.com
- Alexander V. Boukhanovsky** — PhD; ITMO University, School of Translational Information Technologies; Head of the School; E-mail: avb_mail@mail.ru

For citation: *Bochenina K. O., Butakov N. A., Boukhanovsky A. V. Scheduling sets of composite applications with soft deadlines in time windows of heterogeneous computing environments // Journal of Instrument Engineering. 2017. Vol. 60, N 4. P. 295—301 (in Russian).*

DOI: 10.17586/0021-3454-2017-60-4-295-301