

# Динамические и статические методы планирования пакетов заданий в распределенных вычислительных системах

С.В. Минухин

Харьковский национальный экономический университет, 9-а, пр. Ленина, Харьков, Украина

ms\_vl@mail.ru

**Аннотация.** Рассмотрены вопросы планирования заданий в распределенных вычислительных системах на основе метода о покрытии. Разработан сценарий проведения экспериментальных исследований и сравнительного анализа различных методов решения задачи о покрытии и метода FCFS. Проведены исследования оценки эффективности исследуемых методов на основе времени выполнения и коэффициента использования ресурсов. Приведен сравнительный анализ полученных результатов экспериментов для динамического и статического методов пакетного планирования. Сформулированы требования к выбору параметров компонент используемого механизма планирования, позволяющие повысить производительность распределенных вычислительных систем и сформулировать предпочтения использования динамического и статического методов планирования.

## Ключевые слова

Распределенная вычислительная система, задание, планирование, метод покрытия, время выполнения, коэффициент использования, динамический метод, статический метод.

## 1 Введение

Одной из важнейших задач повышения производительности и эффективности работы распределенных вычислительных систем (мультикластерных, вычислительных Грид-систем) является выбор метода планирования, режима работы с заданиями входной очереди (пакетный, интерактивный), а также оценка эффективности на основе определенных критериев (целевых функций). В качестве таких критериев используются такие, как время разрешения очереди (время завершения выполнения последнего задания очереди, makespan), время ответа (или среднее время ответа – среднее время выполнения одного задания с учетом времени ожидания до обработки, времени, затрачиваемого на его планирование (совокупное время, называемое ожиданием), и времени решения задания на ресурсе), а также коэффициент использования ресурсов (коэффициент загрузки). При этом основной задачей является определение таких параметров планирования и параметров инфраструктуры распределенной системы, при которых при обработке потока входных заданий, оптимизируются приведенные целевые функции, а также обеспечивается требуемый (допустимый) уровень качества обслуживания пользователей QoS (Quality of Service).

Целевые функции отражают различные стратегии для оценки эффективности работы распределенной системы: коэффициент использования является системным (системно ориентированным) показателем, характеризующим загрузку ресурсов и основным для владельцев ресурсов; временные критерии отражают стратегии, ориентированные на пользователей (как одиночных, так и в составе виртуальных организаций). Таким образом, состав пользователей и решаемые ими задания (заявки, запросы) определяют и выбор режима работы с заданиями входной очереди, характеризующие приоритеты одиночных пользователей или участников виртуальных организаций. Для первой группы более предпочтительным является, как правило, интерактивный режим, для второй группы – пакетные режимы обработки заданий, получившие в настоящее время значительное развитие [1–6]. В данном исследовании развивается метод, в котором используется математическая модель задачи о наименьшем покрытии, основные результаты использования которого получены в работах [1–3, 5, 6].

Целью настоящего исследования является исследование и анализ основанных на методе покрытия методов планирования пакетов заданий с использованием статического и динамического режимов, определяемых изменением параметров системы планирования и инфраструктуры распределенных вычислений в зависимости от изменения параметров потоков входных заданий и распределенной вычислительной системы, а также оценка эффективности их использования на основе выбранных критериев функционирования распределенной системы.

## 2 Организация проведения исследования методов планирования пакетов заданий

В исследовании проведен анализ эффективности методов планирования на основе плана экспериментов, который представлен на рис. 1. Для проведения анализа были выбраны 3 метода, относящиеся к методам решения задачи о покрытии – точный метод решения задачи о покрытии, жадный алгоритм решения задачи о покрытии, эвристический метод МС (minimal cover) [1], а также, с целью проведения сравнительного анализа, метод FCFS.

Предварительные исследования [2, 4, 5] показали, что для всех этих методов наиболее важным, с точки зрения повышения эффективности их использования, является величина средней относительной сложности заданий (среднего времени выполнения заданий входной очереди на ресурсе), рассчитываемая как отношение средней сложности заданий (среднего времени выполнения) к средней производительности ресурсов системы. Для оценки влияния фактора сложности задания на критерии эффективности функционирования системы в данной работе используется следующая классификация: в качестве малой относительной сложности заданий было выбрано значение 10, в качестве большой – 1000 (как правило, дифференциация заданий по сложности (времени выполнения) определяется для такой классификации несколькими порядками). Таким образом, характеристика задания – «относительная сложность задания» – выбрана в качестве корневого узла предлагаемого дерева последовательности проведения экспериментов (см. рис. 1), а основные узлы определяют направления исследования, соответствующие заданиям с малой (левая ветвь) и большой (правая ветвь) относительной сложностью (рис. 1).

Другим ключевым компонентом механизма планирования для реализации исследуемых методов планирования является величина пула [3–5]. Исследование влияния динамики изменения величины пула позволяет определить те условия, при которых эвристические алгоритмы решения задачи о покрытии (жадный алгоритм и метод МС [4–6] повышают эффективность результатов планирования, а именно: меньшее среднее время (математическое ожидание) выполнения заданий, а также большее среднее значение (математическое ожидание) коэффициента использования ресурсов. Таким образом, второй задачей данного исследования является обоснование целесообразности использования процедуры выбора величины пула для заданий с различной сложностью и универсальностью в условиях фиксированной периодичности планирования.

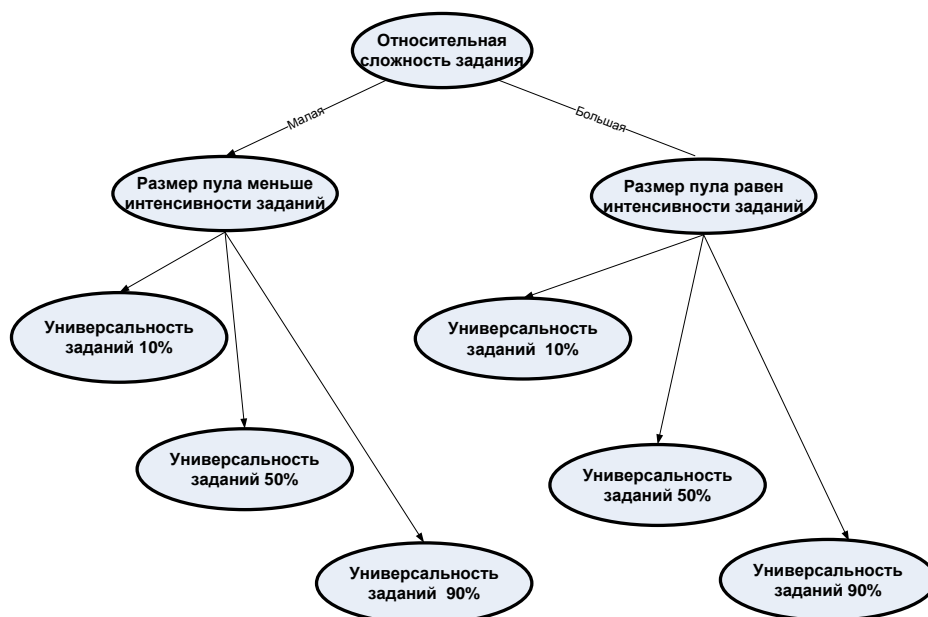


Рис. 1. Дерево сценария исследования и оценки эффективности методов планирования для заданий с различными значениями параметров их характеристик

Среднее время освобождения ресурсов системы [4, 5] является ключевым показателем для выбора величины периодичности планирования, так как он определяет среднюю продолжительность работы ресурса для решения заданий соответствующей сложности, и, следовательно, позволяет определить периодичность планирования (интервал планирования) в качестве некоторой границы, при превышении которой ресурсы, с большой вероятностью, будут простаивать, т.е. не будут загружены заданиями вследствие того, что планирование не проводилось. Распределение относительной сложности заданий определяется отношением используемых в моделировании двух нормально распределенных величин, и в случаях, когда среднее значение времени выполнения заданий превышает среднюю величину производительности ресурсов системы, оно также имеет нормальное распределение.

В исследовании использованы следующие значения периодичности: для заданий малой относительной сложности периодичность планирования равна 1 такту [2] (минимальной периодичности в используемой имитационной модели механизма планирования); для заданий большой относительной сложности равна  $(-3\sigma)$  относительно среднего времени освобождения ресурса. В данном случае использование правила  $3\sigma$  гарантирует, что периодичность планирования будет определяться минимально возможным временем, требуемым для освобождения ресурса. Так, например, если выбрать величину периодичности равной среднему времени освобождения ресурса  $+2\sigma$ , это приведет к тому, что задания, длительность которых равна среднему времени  $-2\sigma$  будут простаивать, так как периодичность планирования будет больше времени, требуемого на выполнение этого задания. Таким образом, выбор экстремальных значений для определения периодичности планирования, распределенной по случайному закону, позволяет определить условия достижимости максимальной эффективности планирования для заданных параметров выполняемых заданий и имеющихся ресурсов.

При проведении экспериментов использовано небольшое количество ресурсов (в качестве их могут выступать процессоры или ядра процессора) – в диапазоне от 10 до 30. Это связано с тем, что точный метод решения задачи о покрытии не позволяет проводить моделирование с требуемым количеством тестов (например, 40 наблюдений, позволяющих получать статистически достоверные результаты) и количества ресурсов более 30 вследствие экспоненциальной сложности точного метода решения задачи о покрытии. Выбираемая универсальность задания (отношение количества ресурсов, на которых может быть решено задание входной очереди, к общему количеству ресурсов распределенной системы) определяет граничные ситуации для планирования заданий на ресурсы: 10% - задания уникальные, 50% - задания могут быть решены на 50% от имеющихся ресурсов, 90 % - задания могут быть решены на 90 % от имеющихся ресурсов. Как показали исследования [4, 5], универсальность заданий также является важным фактором, влияющим на эффективность работы методов планирования.

Далее рассмотрен динамический метод пакетного планирования заданий, при котором исследуется влияние интенсивности заданий входного потока и количества ресурсов системы на величину пула и периодичность планирования для повышения эффективности функционирования распределенной вычислительной системы.

### **3 Исследование и анализ динамических методов пакетного планирования заданий**

В соответствии со сценарием проведения исследований рассмотрим влияние заданий малой сложности на результаты работы алгоритмов.

В качестве настроек динамической системы использованы следующие: средняя сложность заданий – 100, СКО – 10 (нормальный закон распределения); средняя интенсивность заданий входного потока – 50, СКО – 5 (нормальный закон распределения); универсальность заданий – 10, 50 и 90 %; размер пакета заданий – 1; величина пула – 10, 50; производительность ресурса – 10, СКО – 1 (нормальный закон распределения). Количество ресурсов изменялось в диапазоне от 10 до 30 с шагом 10; периодичность планирования – 1 такт [2, 4, 5].

На рис. 2 и 3 показаны результаты планирования заданий с заданными параметрами для точного метода, жадного алгоритма, метода МС и FCFS для пула 10 и универсальности заданий 10 %. Следует отметить, что результаты планирования для точного метода, жадного алгоритма и МС практически идентичны, как для времени выполнения, так и для коэффициента использования ресурсов. При увеличении количества ресурсов имеет место нелинейное уменьшение среднего времени выполнения заданий (рис. 2): при увеличении количества ресурсов до 20 (прирост – 10 ресурсов) уменьшение составляет 1100 тактов, а при увеличении количества ресурсов до 30 (прирост – 10 ресурсов) уменьшение составило только 300 тактов. Такой эффект проявляется вследствие уменьшения очередей заданий на соответствующие ресурсы. Увеличение количества ресурсов повышает вероятность нахождения подходящего ресурса для заданий низкой универсальности, и, таким образом, меньше времени затрачивается на нахождение задания в очереди на ресурс. Увеличение количества ресурсов в данном случае приводит к увеличению их общей загрузки и положительно влияет на балансировку загрузки ресурсов (рис. 3). Использование метода FCFS приводит к значительно худшим результатам, чем точный метод, жадный алгоритм и эвристический метод МС, в первую очередь, потому, что

FCFS загружает первый доступный ресурс в списке свободных и подходящих (First Fit), а в условиях низкой универсальности заданий эта процедура становится неэффективной в случае нескольких конкурирующих между собой заданий, претендующих на один и тот же ресурс.

На рис. 4 и 5 показаны результаты планирования для точного метода, жадного алгоритма, метода MC и FCFS для величины пула 10 и универсальности заданий равной 50 %. Особенностью этих результатов является практически линейный тренд уменьшения среднего времени выполнения для точного метода, жадного алгоритма и MC при увеличении количества ресурсов. Среднее время выполнения для метода FCFS, точного и эвристических методов (жадного алгоритма и MC) для количества ресурсов равного 10 становится минимальным, а коэффициент использования для количества ресурсов 10 одинаков для всех методов. Увеличение количества ресурсов практически не улучшает результаты планирования при использовании метода FCFS, более того, при увеличении количества ресурсов до 20 наблюдается резкое уменьшение коэффициента использования (рис. 5).

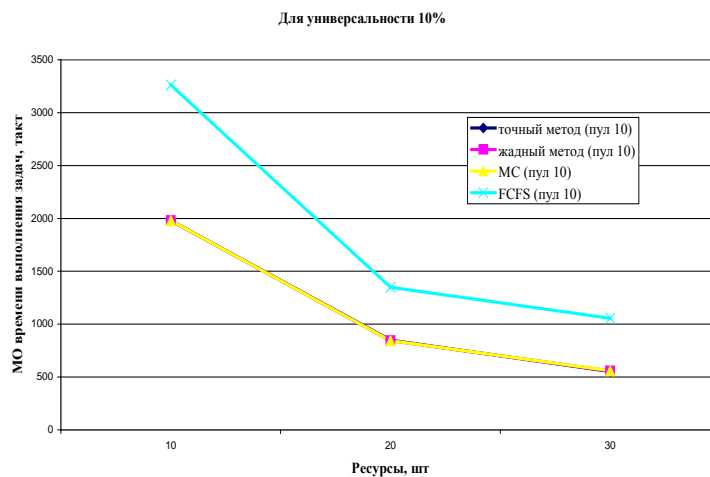


Рис. 2. Зависимость среднего времени выполнения заданий для точного метода, жадного метода, MC и FCFS для величины пула 10, значения универсальности, равной 10 % от количества ресурсов

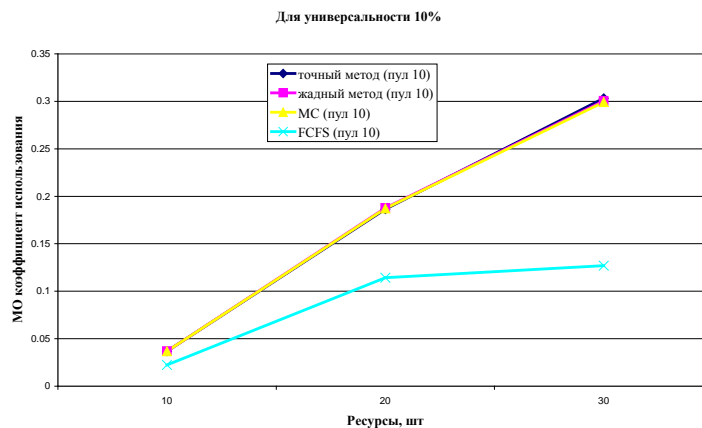
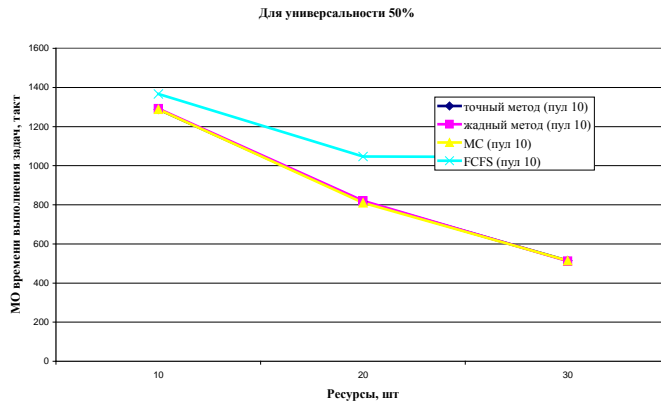
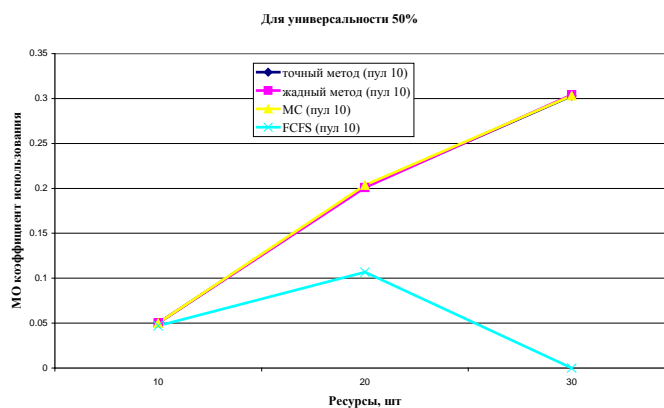


Рис. 3. Зависимость среднего значения коэффициента использования ресурсов для точного метода, жадного, MC и FCFS для величины пула 10, универсальности, равной 10 % от количества ресурсов

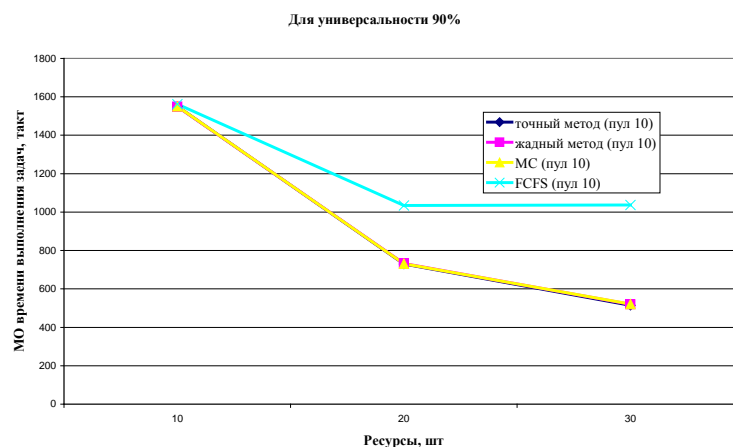


**Рис. 4.** Зависимость среднего времени выполнения заданий для точного метода, жадного алгоритма, MC и FCFS для величины пула 10, универсальности, равной 50 % от количества ресурсов

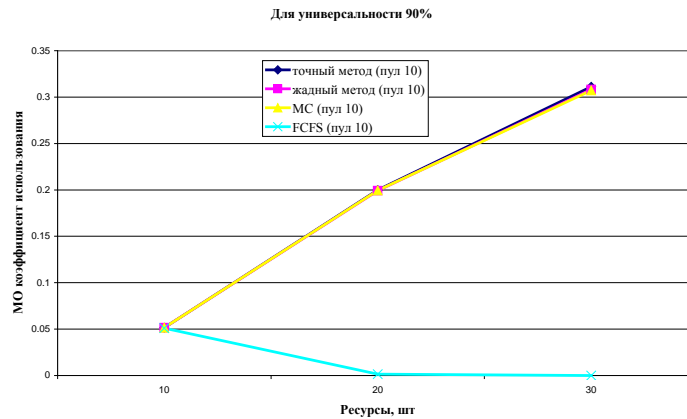


**Рис. 5.** Зависимость среднего значения коэффициента использования ресурсов для точного метода, жадного, MC и FCFS для величины пула 10 и универсальности заданий, равной 50 % от количества ресурсов

Анализ результатов планирования для точного метода, жадного алгоритма, MC и FCFS при величине пула 10 и универсальности заданий равной 90% показан на рис. 6, 7. При количестве ресурсов равным 10 среднее время выполнения и коэффициент использования ресурсов у всех методов одинаковы, а при увеличении количества ресурсов до 20 эффективность метода FCFS снижается: уменьшается среднее значение коэффициента использования (рис. 6). Это свидетельствует о неэффективности планирования методом FCFS в условиях тенденции увеличения (или большого количества) ресурсов для заданий с высокой универсальностью. В случае практической реализации это соответствует ситуации, при которой вычислительная архитектура имеет низкую гетерогенность (фактически является гомогенной).



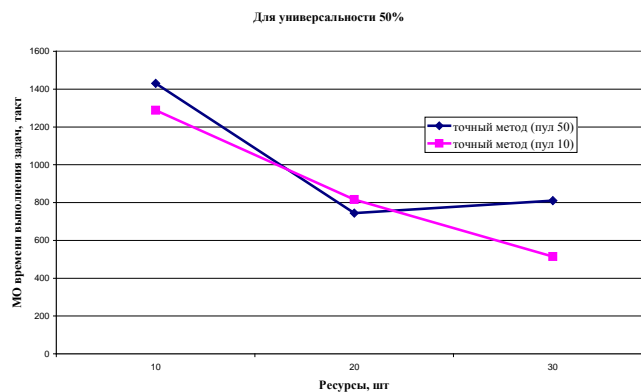
**Рис. 6.** Зависимость среднего времени выполнения задач для точного метода, жадного алгоритма, MC и FCFS для величины пула 10 и универсальности заданий 90 % от количества ресурсов



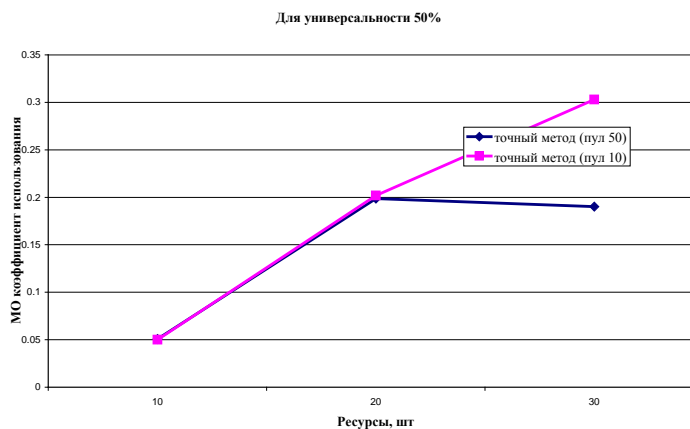
**Рис. 7.** Зависимость среднего значения коэффициента использования ресурсов для точного метода, жадного, MC и FCFS для величины пула 10 и универсальности заданий, равной 90 % от количества ресурсов

Таким образом, при небольшом количестве ресурсов и с увеличением универсальности поступающих на обработку в систему заданий все методы, основанные на использовании метода покрытия, позволяют получить существенно лучшие результаты по сравнению с наиболее используемым в настоящее время в промышленных системах методом FCFS.

Сравнительный анализ особенности влияния величины пула на эффективность использования методов планирования для универсальности заданий 50 % приведен на рис. 8–13. На рис. 8, 9 показаны результаты планирования для точного метода, которые подтверждают влияние величины пула на эффективность планирования. В данном случае небольшой пул обеспечивает лучшие результаты планирования – времени выполнения и коэффициента использования ресурсов.



**Рис. 8.** Зависимость среднего времени выполнения заданий для точного метода для величины пула 10 и 50 при универсальности заданий, равной 50 % от количества ресурсов



**Рис. 9.** Зависимость среднего значения коэффициента использования ресурсов для точного метода для величины пула 10, 50 при универсальности заданий, равной 50 % от количества ресурсов

Таким образом, величина пула, даже при небольшом количестве ресурсов, влияет на производительность распределенной вычислительной системы, и оптимизация ее работы может быть осуществлена на основе выбора его величины.

В работе приведены результаты, обосновывающие влияние периодичности планирования на эффективность функционирования распределенной вычислительной системы.

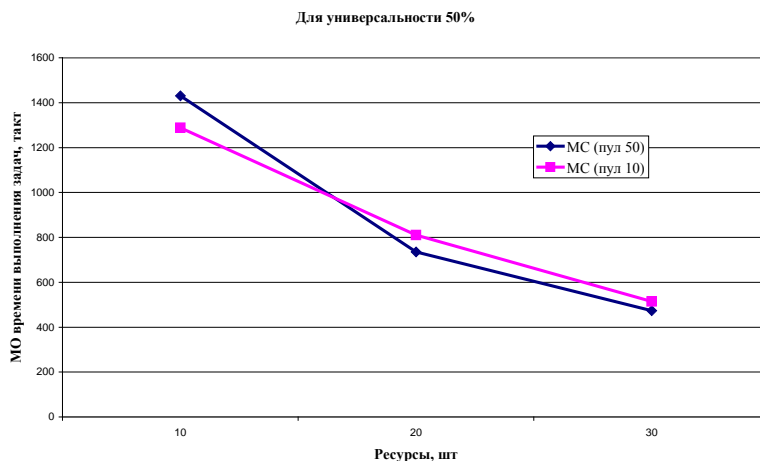


Рис. 10. Зависимость среднего времени выполнения для метода МС для величины пула 10, 50 при универсальности заданий, равной 50 % от количества ресурсов

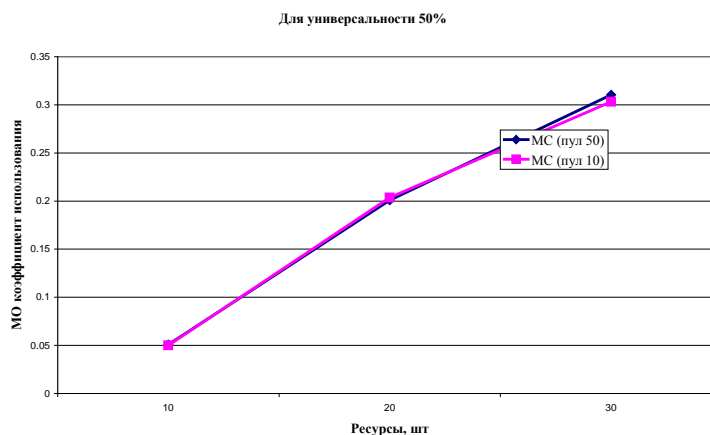


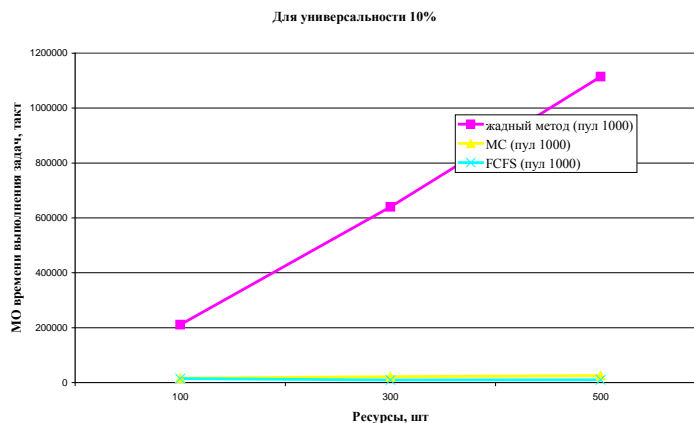
Рис. 11. Зависимость среднего значения коэффициента использования ресурсов для метода МС для величины пула 10, 50 при универсальности заданий, равной 50 % от количества ресурсов

## 4 Исследование и анализ статических методов пакетного планирования заданий

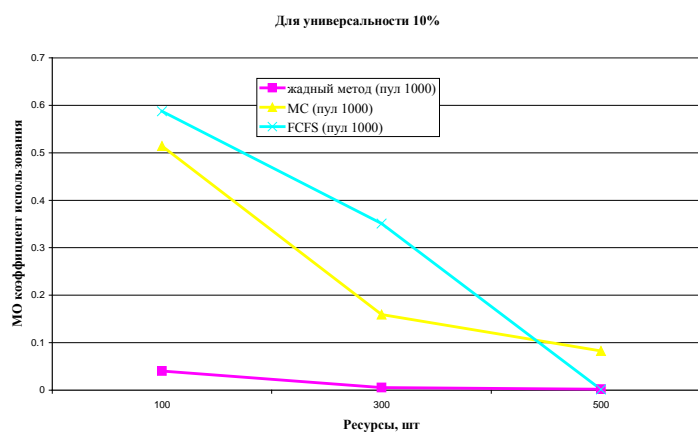
Использование статического метода планирования предполагает отсутствие динамически настраиваемых параметров модели, как было рассмотрено в п. 3 данного исследования, в случае динамически изменяющихся параметров заданий входного потока и распределенной вычислительной системы (количества ресурсов).

Для проведения экспериментов сгенерировано 1000 заданий большой относительной сложности при фиксированной интенсивности поступления заданий – 1000 и пуле, равным 1000 (т.е. все задания сразу же попадают в пул). Результаты планирования получены для универсальности заданий 10%, 50% и 90%.

Результаты планирования для различных критериев эффективности работы системы для универсальности заданий 10% показаны на рис. 12, 13.



**Рис. 12.** Зависимость среднего времени выполнения заданий для FCFS, жадного алгоритма и метода MC для величины пула 1000 при универсальности заданий, равной 10 % от количества ресурсов (большое количество)



**Рис. 13.** Зависимость среднего значения коэффициента использования для FCFS, жадного алгоритма и метода MC для величины пула 1000 при универсальности заданий, равной 10 % от количества ресурсов (большое количество)

Следует отметить, что при статическом планировании коэффициент использования FCFS для 100 ресурсов является лучшим среди исследуемых методов, однако для 500 ресурсов, в отличие от метода MC, коэффициент использования для метода FCFS стремится к 0 (см. рис. 13).

Полученные результаты планирования для универсальности заданий 50% и 90% свидетельствует о том, что при увеличении универсальности заданий при использовании метода MC среднее время выполнения заданий увеличивается. При универсальности заданий 90% и количестве ресурсов 100 время выполнения заданий при статическом планировании для MC значительно увеличивается (100 000 тактов), чем при динамическом планировании составил 0.52, а при статическом планировании – 0.08. Таким образом, при увеличении универсальности заданий эффективность статического планирования для метода MC уменьшается. Это объясняется тем, что при использовании только одной процедуры планирования количество решений задачи о покрытии в процессе перепланирования резко увеличивается вследствие того, что задания, находящиеся в очереди, с увеличением универсальности практически все решаются на одних и тех же ресурсах, что, в свою очередь, приводит к увеличению времени нахождения заданий в очередях на ресурсы.

## 5 Выводы

Проведенные исследования показали, что при выборе стратегий планирования выполнения заданий в распределенных вычислительных системах необходимо учитывать факторы, влияющие на компоненты механизма планирования в зависимости от изменения характеристик поступающих на обработку заданий и распределенной инфраструктуры. Для используемой модели планирования [1, 2] пакетов заданий такими компонентами являются величина пула, универсальность заданий и периодичность планирования.

В целом, полученные результаты позволяют сформулировать следующее:



для заданий большой относительной сложности ключевым фактором является периодичность планирования. Увеличение периодичности планирования приводит к ухудшению эффективности методов решения задачи о покрытии. Однако, в тех случаях, когда периодичность планирования меньше величины  $(-3\sigma)$  относительно среднего времени освобождения ресурса, эффективность планирования не изменяется;

универсальность заданий существенно влияет на результаты планирования. При увеличении универсальности заданий эффект использования методов планирования на основе метода покрытия уменьшается. При этом при универсальности заданий 90% для количества ресурсов 20, 30 результаты планирования при использовании этих алгоритмов практически идентичны;

для заданий большой сложности использование большого пула является нецелесообразным, так как за время решения заданий можно выполнить неоднократное перепланирование заданий на освободившиеся ресурсы, и, следовательно, принципиальным является не величина пула, а периодичность планирования. При этом пул рекомендуется выбирать равным (или меньшим) количеству ресурсов (в исследовании рекомендуется выбирать величину пула равной 10);

применение статического метода пакетного планирования приводит к ухудшению результатов по сравнению с динамическим. Вместе с тем, при статическом планировании при низкой универсальности заданий метод МС позволяет получить лучшие результаты по времени выполнения заданий по отношению к динамическому при сохранении величины коэффициента использования ресурсов. Статическое планирование целесообразно использовать для заданий с большой сложностью, высокой интенсивностью входного потока заданий и большого количества ресурсов.

## Литература

- [1] С.В. Листровой, С.В. Минухин. Модель и подход к планированию распределения ресурсов в гетерогенных Грид-системах. *Международный научно-технический журнал «Проблемы управления и информатики»*, 2012, 5, 120–133.
- [2] С.В. Минухин, С.В. Знахур. Имитационная модель и ее программная реализация планирования ресурсов Грид-системы. УкрПРОГ 2012. *Проблеми програмування*, 2012, №2–3, спеціальний випуск, 133–143.
- [3] С.В. Минухин, С.В. Знахур. Исследование эффективности методов планирования ресурсов для различных интенсивностей потоков заданий в Грид. *Радіоелектронні і комп'ютерні системи*, 2012, № 1, 165–171.
- [4] С.В. Минухин, С.В. Знахур. Исследование и анализ основанных на методе покрытия алгоритмов планирования для гетерогенной распределенной системы с использованием сценарного подхода. В монографии «Информационные технологии и системы в управлении, образовании, науке», 2013, Харьков, Цифрова друкарня, 39–54.
- [5] С.В. Минухин, С.В. Знахур. Разработка дерева сценариев исследования и анализа результатов процедуры планирования для гетерогенной распределенной системы. International Conference «Parallel and Distributed Computing Systems». PDCS 2013 (Ukraine, Kharkiv, March 13-14, 2013), 218 – 226.
- [6] С.В. Минухин. Моделирование и анализ влияния основанных на методе покрытия алгоритмов планирования заданий и распределения ресурсов на производительность гетерогенных распределенных систем. *Системи обробки інформації*, 2012, вип. 8 (106), 27–32.