

КООРДИНИРОВАННЫЕ СТРАТЕГИИ ПЛАНИРОВАНИЯ ЗАДАНИЙ НА ОСНОВЕ ОБОБЩЕННОГО СТАТИСТИЧЕСКОГО АНАЛИЗА ИХ ХАРАКТЕРИСТИК В РАСПРЕДЕЛЕННЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМАХ

В статье рассмотрены вопросы разработки и оценки стратегий планирования заданий в распределенных вычислительных системах на основе обобщенного статистического анализа характеристик заданий. Показано, что в условиях динамически изменяющейся нагрузки на вычислительную систему представляется эффективным для оценки производительности системы использование синтетических (сгенерированных для проведенных расчетов по экспоненциальному закону распределения) интенсивности и длительности заданий, с одинаковым средним значением, но различным соотношением между короткими и длинными заданиями. Приведены результаты расчетов метрик работы вычислительного кластера на реальных и синтетических данных, которые показали, что эффективность стратегий планирования можно повысить на основе дополнительной информации о распределении длительности входных заданий.

1. Введение

Одной из важнейших задач повышения производительности и эффективности работы распределенных вычислительных систем является оценка их текущей загрузки и выбор таких стратегий планирования входных потоков заданий, которые позволили бы в соответствии с выбранными критериями эффективности оптимизировать их работу [1–5]. В литературе решению этих вопросов уделяется значительное внимание. В связи с этим отметим и выделим следующие направления исследований. В работах [6–10], посвященных статистическому анализу находящихся в вычислительной системе заданий (workload), анализируются их следующие основные характеристики – закон распределения поступающих заданий (их интенсивности), закон распределения длительности заданий и времени между прибытиями заданий (interarrival time). Основной задачей является определение условий, при которых вычислительная система обрабатывает потоки заданий таким образом, что при этом оптимизируются такие показатели ее функционирования, как время ответа системы, коэффициент использования (загрузки) ресурсов, время завершения последнего задания (makespan), а также обеспечивается определенный уровень качества обслуживания пользователей QoS (Quality of Service). В работе [6] для повышения эффективности работы вычислительной системы предложено использовать координированную по отношению к характеристикам поступающих заданий стратегию: для этого предложено использовать вари-

ант агрессивного backfilling (Job Rank-backfilling policy), смысл которого заключается в том, что в первую очередь продвигаются те задания, которые имеют малую длительность (short или small jobs). Другой подход заключается в том, что с целью избегания «зависания» длинных заданий наряду с методом backfilling используется стратегия Shortest Job First (SJF) (Shortest Job Backfilled First, SJBF) [7, 8]. В работе [9] для исследования влияния стратегии работы с очередью был предложен «пороговый» подход: задание считается коротким, если ее длительность (рассчитанная или ожидаемая) меньше принятого порогового значения; работа считается длинной, если ее длительность (рассчитанная или ожидаемая) больше принятого порогового значения; при этом определяется процент тех заданий, длительность которых принимается по расчету как длинные, а они являются фактически короткими, и наоборот. Таким образом, данный подход позволяет определить целесообразность использования стратегий SJF или LJF вместе с алгоритмом backfilling, используемом для выполнения заданий на ресурсе [10].

Анализ приведенных методов координированных стратегий показывает, что координация осуществляется на основе 2-х основных подходов: анализа плотности распределения оценки (предварительно рассчитываемой) длительности поступающих на ресурс заданий и времени прибытия (поступления) заданий (стохастический вариант) или с использованием некоторых пороговых значений (детерминированный ва-

риант). В обоих случаях имеется неточность: в первом случае – в силу стохастичной природы оценивания имеются короткие и длинные задания, определяемые математическим ожиданием; во втором случае – неточность определяется самим подходом: пороговое значение, а, следовательно, и оценки значений для коротких и длинных заданий имеют погрешность. Это может привести к изменению (или повлиять на выбор) стратегий планирования заданий.

Необходимо отметить, что приведенные показатели функционирования отображают различные подходы для оценки эффективности работы распределенной системы: коэффициент использования является системным (системно ориентированным) показателем, характеризует уровень загрузки ресурсов и является ключевым для их владельцев; временные критерии отражают стратегии, ориентированные на пользователей (как одиночных, так и в составе виртуальных организаций). Это приводит к тому, что одновременное выполнение критериев работы указанных двух участников виртуальных организаций представляется затруднительным и, вместе с тем, определяет направленность исследований для нахождения компромиссного решения на основе реальных данных функционирования распределенных систем и определения границ, в которых такой компромисс может быть достигнут.

В работе [11] приведены результаты исследований для оценки степени влияния стратегий работы с входным потоком заданий на производительность вычислительных систем, для определения которого (и основных факторов) использовались такие методы планирования заданий как Shortest Job First (SJF), Longest Job First (LJF) и наиболее распространенный в промышленных системах метод FCFS. Данные о выполняемых заданиях разбивались на периоды (например, сутки, недели) и моделировались ситуации, в которых выбираемые задания сортировались по убыванию и по возрастанию их длительностей. После этого запускался механизм работы с очередью FCFS и задания обрабатывались в соответствии с методами пакета GridSim и интегрированным в программное обеспечение симулятора планировщика, основанного на методе наименьшего покрытия [5]. Необходимо отметить, что при проведении эксперимента не учитывались значения параметров плотности распределения такого **показателя** как математическое ожидание длительности

заданий, которое, как уже отмечалось, может повлиять на выбор стратегии работы с входной очередью. Аналогичные вопросы рассматривались при исследовании эффективности пакетных режимов, в которых преобладание коротких или длинных заданий непосредственно влияет на выбор метода планирования заданий [11].

Целью настоящего исследования является определение влияния характеристик входных заданий на выбор стратегий планирования на основе статистического анализа длительностей и интенсивностей реальных заданий и синтетических заданий с заданными статистическими характеристиками для распределенной вычислительной системы.

2. Статистический анализ входных потоков заданий

Для статистического анализа входного потока заданий использованы данные о выполненных заданиях на кластере ИСМА НАН Украины за период 2012–2013 гг.

Обобщенный статистический анализ предусматривает:

статистический анализ интенсивности заданий по генеральной выборке – в целом, и по этой же выборке – по дням недели;

статистический анализ длительности заданий;

построение временного ряда для определения динамики интенсивности входных заданий за анализируемый период.

Статистический анализ характеристик входных заданий проводился на основе следующих этапов.

Этап 1. Статистический анализ количества заданий (плотность распределения количества) за весь период показал, что они распределены по экспоненциальному закону (рис. 1).

Этап 2. Статистический анализ количества поступающих заданий по дням недели – определения наиболее загруженных дней недели или дней, имеющих определенные закономерности поступления заданий на выполнение. Результаты определения плотности распределения приведены на рис. 2.

Полученные результаты (рис. 2) показывают сохранение экспоненциального закона распределения количества заданий и по дням недели, при этом наибольшее количество заданий выполняется в среду (в целом – в первой половине недели).

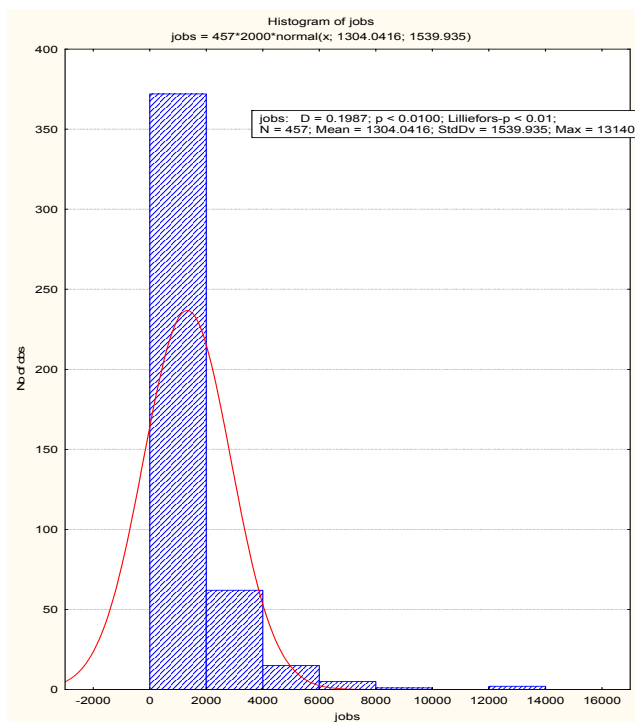


Рис. 1. Плотность распределения количества заданий за анализируемый период

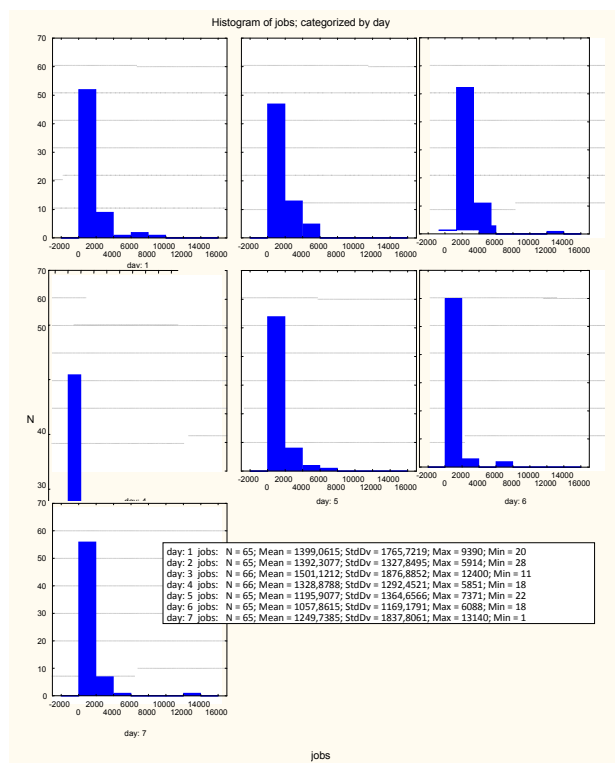


Рис. 2. Экспоненциальная плотность распределения количества поступающих заданий по дням недели

Таким образом, без потери общности анализа полученных результатов, в дальнейшем исследование проводилось для выборки заданий, поступивших за 1 неделю.

Этап 3. Анализ стационарности входных потоков заданий по дням.

Данный анализ проводился с целью определения стационарности потоков заданий (загрузки) кластера за весь анализируемый период (см. рис. 3). Из графика на рис. 3 следует, что имеется явный тренд увеличения количества заданий, поступающих на кластер на обработку, причем имеются изменения в политиках (стратегиях) использования ресурсов (240–250 дни исходной выборки заданий), при которых средняя загрузка на кластер увеличилась на 10%.

Последний результат свидетельствует о том, что интенсивность входного потока заданий возросла, что, в общем, может привести к изменению параметров функционирования кластера: в случае, если количество ресурсов не изменилось, коэффициент использования может уменьшиться; анализ динамики среднего времени ответа требует дополнительного исследования данных о выполненных заданиях.

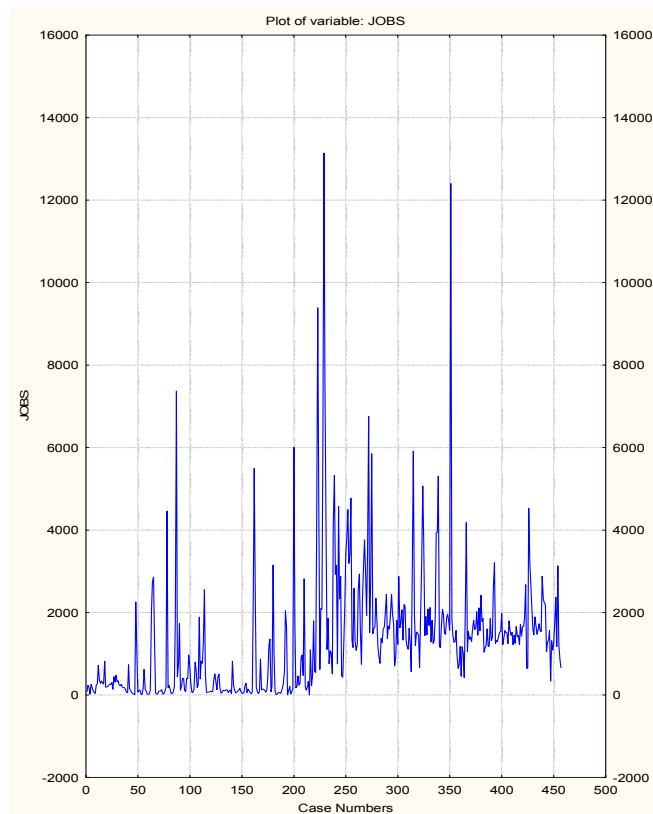


Рис. 3. Временной ряд количества поступивших заданий в течение 460 дней

Предварительный этап статистического анализа потоков заданий завершается определением закона распределения длительностей выполнения заданий (рис. 4).

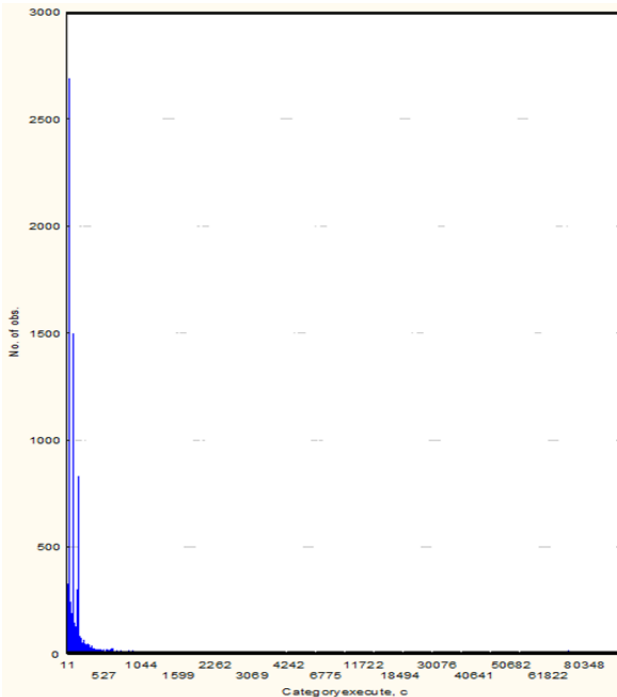


Рис. 4. Экспоненциальный закон распределения длительности заданий

Проведенный анализ показал, что длительности заданий подчиняются экспоненциальному закону распределения, среднее значение составляет 1308 сек. (22 мин.), т.е. большинство заданий – короткие.

Для исследования результатов применения стратегий планирования SJF (или LJF) в данной работе использованы синтетические тестовые данные для интенсивностей поступления заданий и их длительностей, сгенерированных по экспоненциальному закону распределения.

При генерации синтетических заданий приняты следующие допущения:

интенсивность и закон распределения сгенерированных входных заданий должен соответствовать интенсивности и закону распределения заданий на кластере ИСМА НАН Украины за анализируемый период (см. рис. 1, 2, 4);

для оценки влияния длительности заданий на стратегии планирования предполагается осуществить следующие генерации распределений синтетических заданий по длительности:

а) генерация длительностей заданий на основе среднего значения, но с меньшей долей коротких заданий в тестовой выборке, т.е. увеличение в выборке доли средних и длинных заданий на 10 %;

б) генерация длительности на основе среднего значения и закона распределения длительностей заданий кластера (рис. 4);

время завершения задания рассчитывается как время его старта, увеличенное на его длительность.

Предварительный анализ реальных данных интенсивности поступления заданий на кластер показал, что среднее значение количества заданий по часам (группировка выполнена по часам суток для генеральной выборки) составляет 113 заданий (2 задания в мин.). Таким образом, при среднем времени решения заданий на узлах кластера 22 мин. в очереди на все узлы в среднем может находиться 44 задания.

Генерация синтетических длительностей заданий и их интенсивностей осуществлена в пакете STATISTICA 7.1 на основе экспоненциального закона распределения.

При этом экспоненциальное распределение имеет вид:

$$f(x) = \frac{1}{\lambda} \cdot e^{-\frac{x}{\lambda}}, \quad x \in [0; \infty), \quad \lambda > 0,$$

где среднее значение ряда определяется как отношение x к λ (в нашем случае – 113).

В проведенных экспериментах для генерации интенсивности и интервалов времени поступления заданий на кластер были использованы следующие характеристики $\lambda=2$, $\lambda=0,017544$, которые позволили получить среднее значение синтетического ряда равным 113.

Для генерации синтетических последовательностей поступления заданий в пакете STATISTICA 7.1 использована формула $Vexpon(\text{rnd}(2); 0,017544)$. Статистический анализ интенсивности поступления заданий на кластер в различные интервалы времени показал, что она имеет экспоненциальный закон распределения (рис. 5, 6).

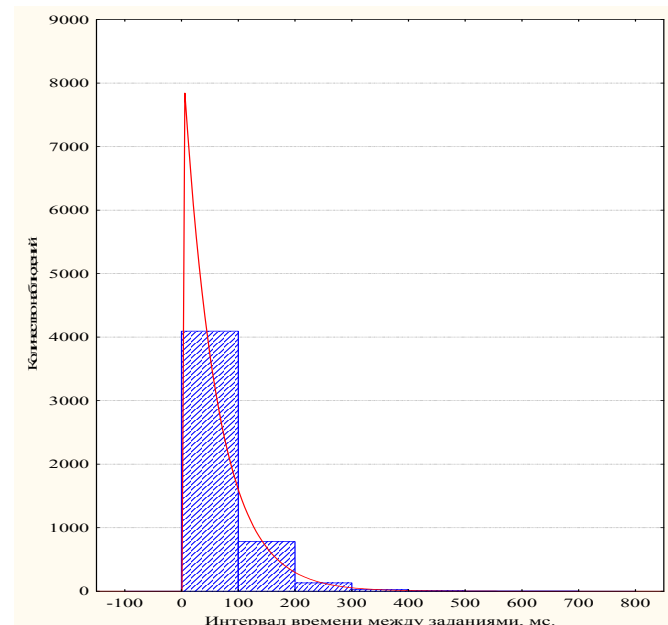


Рис. 5. Экспоненциальный закон распределения интенсивности поступления заданий в различные интервалы времени

| Category | K-S d=,16159, p<,01 ; Lilliefors p<,01 | | | | | |
|----------------------|--|------------------|------------------|------------------|----------------|---------------------|
| | Count | Cumulative Count | Percent of Valid | Cumul % of Valid | % of all Cases | Cumulative % of All |
| -100,000<x<=0,000000 | 0 | 0 | 0,00000 | 0,0000 | 0,00000 | 0,0000 |
| 0,000000<x<=100,000 | 4094 | 4094 | 81,26241 | 81,2624 | 40,94000 | 40,94000 |
| 100,0000<x<=200,000 | 781 | 4875 | 15,50218 | 96,7646 | 7,81000 | 48,75000 |
| 200,0000<x<=300,000 | 129 | 5004 | 2,56054 | 99,3251 | 1,29000 | 50,04000 |
| 300,0000<x<=400,000 | 26 | 5030 | 0,51608 | 99,8412 | 0,26000 | 50,30000 |
| 400,0000<x<=500,000 | 6 | 5036 | 0,11909 | 99,9603 | 0,06000 | 50,36000 |
| 500,0000<x<=600,000 | 1 | 5037 | 0,01985 | 99,9802 | 0,01000 | 50,37000 |
| 600,0000<x<=700,000 | 1 | 5038 | 0,01985 | 100,0000 | 0,01000 | 50,38000 |
| Missing | 4962 | 10000 | 98,49146 | | 49,62000 | 100,00000 |

Рис.6. Частотная таблица распределения интенсивности заданий в различные интервалы времени

Для генерации случайных последовательностей для длительностей заданий (времени выполнения) в пакете STATISTICA 7.1 со средним значением 1400 сек. были использованы следующие параметры: для варианта а) $x=0,9$, $\lambda=0,55$; варианта б) – $x=1$, $\lambda=0,000714$.

Вариант а) был сгенерирован при помощи функции $1000 * Vexron(rnd(0,9); 0,55)$ (см. рис.

7). Вариант б) был сгенерирован при помощи функции $Vexron(rnd(1); 0,000714)$ (см. рис. 8).

Варианты генерации синтетических данных отличаются тем, что вариант а) использует константу равную 1000 для увеличения доли средних и длинных заданий, а вариант б) использует генерацию стандартного экспоненциального распределения (без константы).

Для проверки отличия синтетических данных, сгенерированных в соответствии с вариантами а) и б), построены частотные таблицы, в которых отражены частоты поступления заданий в различные интервалы времени (см. рис. 9, 10). Таким образом, в отличие от работы [10], где генерируется одна генеральная тестовая выборка для получения двух различных тестовых последовательностей с одинаковыми характеристиками, в данном исследовании использован более обоснованный, со статистической точки зрения, подход к выбору стратегии планирования заданий входной очереди.

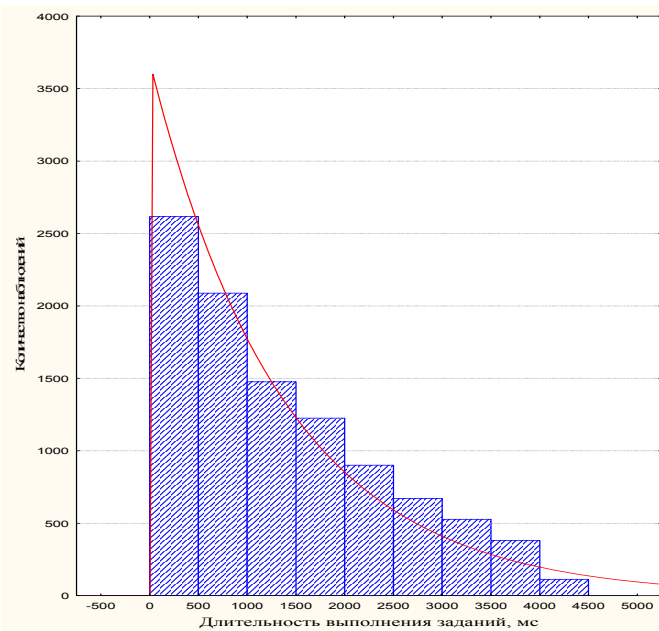


Рис. 7. вариант а) генерации длительностей

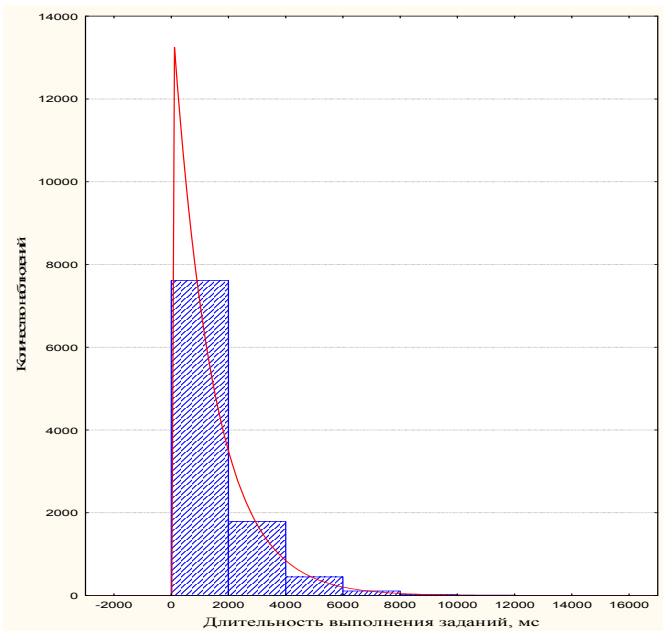


Рис. 8. Вариант б) генерации длительностей

| Category | Frequency table: K-S d=,10338, p<,01 ; Lilliefors p<,01 | | | | | |
|----------------------|--|------------------|------------------|------------------|----------------|---------------------|
| | Count | Cumulative Count | Percent of Valid | Cumul % of Valid | % of all Cases | Cumulative % of All |
| -1000,00<x<=0,000000 | 0 | 0 | 0,00000 | 0,0000 | 0,00000 | 0,0000 |
| 0,000000<x<=1000,00 | 4705 | 4705 | 47,05000 | 47,0500 | 47,05000 | 47,05000 |
| 1000,000<x<=2000,00 | 2702 | 7407 | 27,02000 | 74,0700 | 27,02000 | 74,07000 |
| 2000,000<x<=3000,00 | 1572 | 8979 | 15,72000 | 89,7900 | 15,72000 | 89,79000 |
| 3000,000<x<=4000,00 | 908 | 9887 | 9,08000 | 98,8700 | 9,08000 | 98,87000 |
| 4000,000<x<=5000,00 | 113 | 10000 | 1,13000 | 100,0000 | 1,13000 | 100,00000 |
| Missing | 0 | 10000 | 0,00000 | | 0,00000 | 100,00000 |

Рис.9. Частотная таблица распределения длительности выполнения заданий для варианта а)

| Category | K-S d=,16062, p<,01 ; Lilliefors p<,01 | | | | | |
|----------------------|--|------------------|------------------|------------------|----------------|---------------------|
| | Count | Cumulative Count | Percent of Valid | Cumul % of Valid | % of all Cases | Cumulative % of All |
| -2000,00<x<=0,000000 | 0 | 0 | 0,00000 | 0,0000 | 0,00000 | 0,0000 |
| 0,000000<x<=2000,00 | 7615 | 7615 | 76,15000 | 76,1500 | 76,15000 | 76,15000 |
| 2000,000<x<=4000,00 | 1790 | 9405 | 17,90000 | 94,0500 | 17,90000 | 94,05000 |
| 4000,000<x<=6000,00 | 449 | 9854 | 4,49000 | 98,5400 | 4,49000 | 98,54000 |
| 6000,000<x<=8000,00 | 111 | 9965 | 1,11000 | 99,6500 | 1,11000 | 99,65000 |
| 8000,000<x<=10000,00 | 23 | 9988 | 0,23000 | 99,8800 | 0,23000 | 99,88000 |
| 10000,00<x<=12000,00 | 10 | 9998 | 0,10000 | 99,9800 | 0,10000 | 99,98000 |
| 12000,00<x<=14000,00 | 2 | 10000 | 0,02000 | 100,0000 | 0,02000 | 100,00000 |
| Missing | 0 | 10000 | 0,00000 | | 0,00000 | 100,00000 |

Рис.10. Частотная таблица распределения длительности выполнения заданий для варианта б)

3. Сценарии проведения экспериментов и анализ полученных результатов

Для проведения экспериментов были использованы результаты предварительной статистической обработки данных о входных потоках заданий, приведенные в п. 2, и методы планирования заданий, имеющиеся в пакете GridSim – FCFS, Backfill и интегрированный в него и программно реализованный метод планирования, основанный на методе наименьшего покрытия МС [5]. Организация и проведение экспериментов осуществлялось следующим образом: для анализа была определена контрольная выборка заданий за весь период – 9200 заданий; для определения текущих показателей функционирования системы были определены периоды времени (интервалы планирования) таким образом, чтобы среднее количество поступающих заданий за эти периоды было равным 600.

Для оценки влияния исследуемых методов (стратегий) планирования заданий на функционирование кластера в соответствии с полученными п. 2 результатами реализованы следующие сценарии.

Сценарий 1. Планирование и расчет показателей функционирования системы для реальных данных.

Сценарий 2. Планирование и расчет показателей функционирования для сгенерированных синтетических данных для двух предложенных вариантов генерации длительностей заданий. В качестве метрик работы кластера выбраны следующие: коэффициент использования ресурсов, коэффициент балансировки загрузки, характеристика качества обслуживания, определяемая как отношение среднего времени ожидания заданий в очереди к среднему времени их выполнения $\frac{t_{wait}}{t_{run}}$. Для корректного сравнения по-

лученных результатов для всех методов планирования выбран один и тот же период. Далее на графиках использованы следующие обозначения: КИ – коэффициент использования ресурса; КБ – коэффициент балансировки загрузки; количество ресурсов (КР) и количество процессорных элементов (КПЭ) на ресурсе обозначается как КР КПЭ. На рис. 11, 12 приведены результаты моделирования в пакете GridSim за выбранный период для сортированных по методу LJF заданий и без сортировки для 50 ресурсов, на рис.13, 14 – для 100 ресурсов.

Анализ приведенных на рис. 11–14 результатов показывает, что сортировка заданий по методу LJF практически не влияет на величину КИ как для 50, так и для 100 ресурсов. При этом наилучший результат показывает метод МС (30–33 % – для 50 ресурсов и 19–20% – для 100), затем из наиболее используемых следует backfill, наихудший результат – у метода FCFS.

На рис. 15, 16 приведены результаты моделирования и полученные значения КБ за выбранный период, которые показывают, что этот показатель также практически не изменяется, при этом наибольшее значение имеет метод МС (92 %), наименьшее – метод FCFS (для сортированных заданий 2%, для несортированных – 20%). Это указывает на то, что данный метод может иметь низкий коэффициент использования в случае, когда большинство заданий – длинные, или они отсортированы по убыванию их длительности.

На рис. 17, 18 приведены результаты расчетов для показателя качества обслуживания $\frac{t_{wait}}{t_{run}}$ для 50 и 100 ресурсов, соответственно. Из графиков следует, что, в целом, с увеличением количества ресурсов его величина уменьшается, а лучшие результаты – у методов МС и backfill.

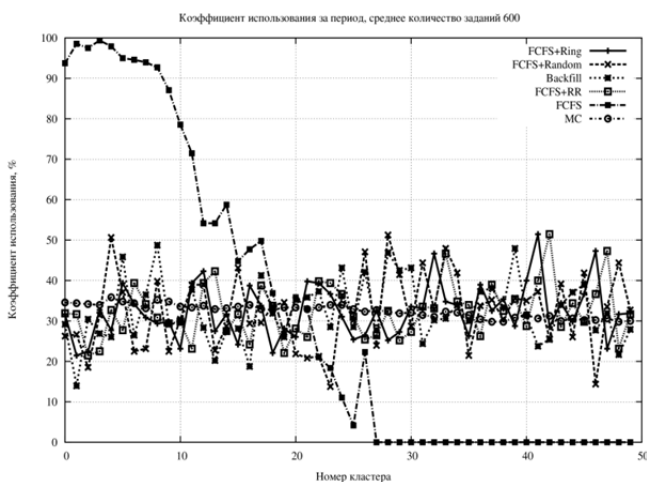


Рис. 11. Зависимость КИ для 50 ресурсов (с LJF)

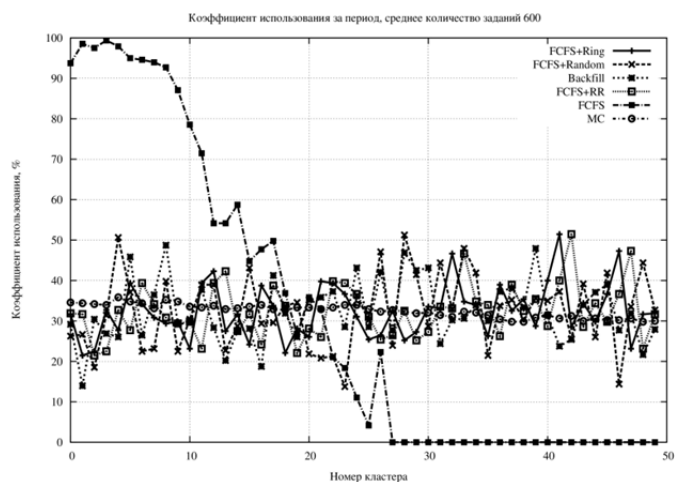


Рис. 12. Зависимость КИ для 50 ресурсов

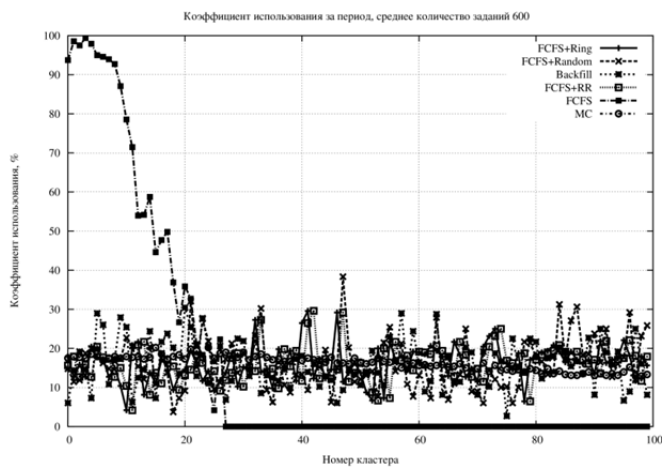


Рис. 13. Зависимость КИ 100 ресурсов (с LJF)

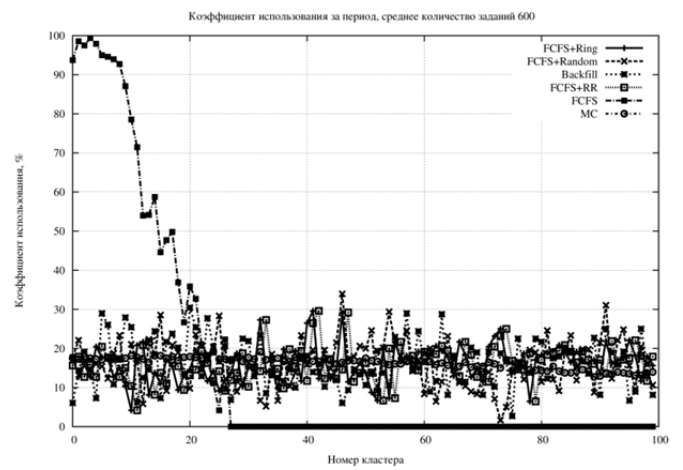


Рис. 14. Зависимость КИ для 100 ресурсов

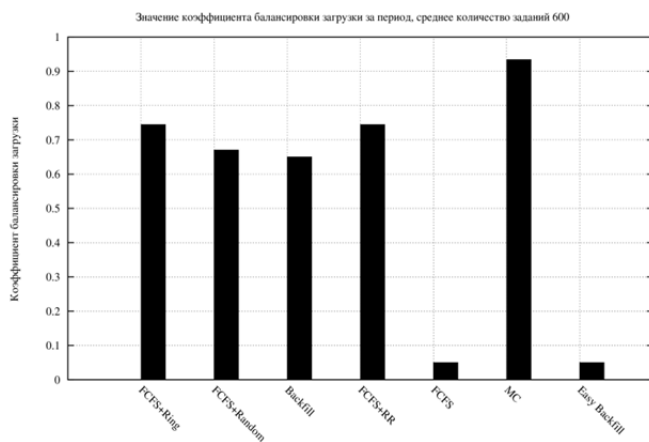


Рис. 15. Зависимость КБ для 70 ресурсов (с LJF)

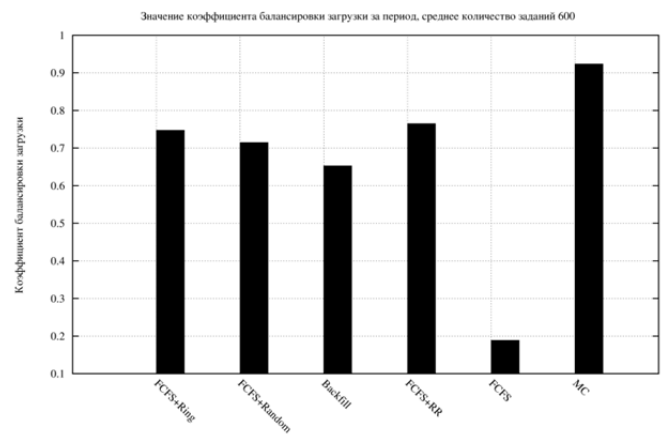


Рис. 16. Зависимость КБ для 70 ресурсов

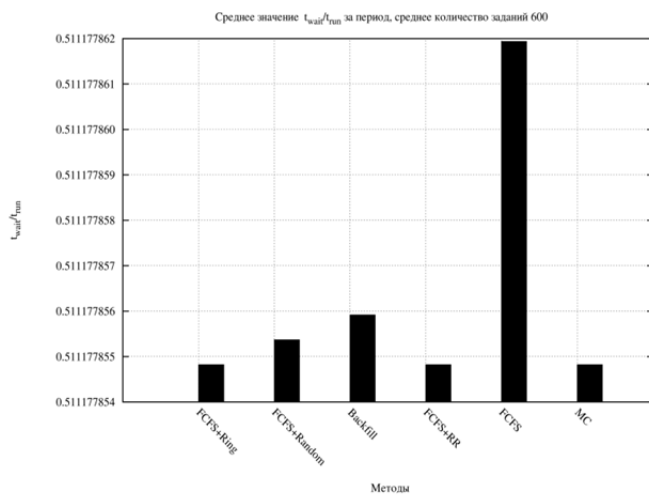


Рис. 16. Зависимость $\frac{t_{wait}}{t_{run}}$ для 50 ресурсов

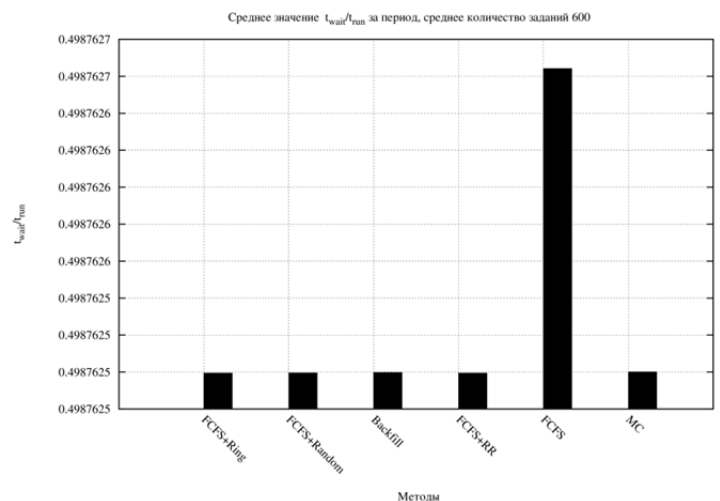


Рис. 17. Зависимость $\frac{t_{wait}}{t_{run}}$ для 100 ресурсов

Моделирование на сгенерированных синтетических данных проводилось в соответствии с описанными ранее сценариями.

Для обеспечения требуемого уровня достоверности полученных результатов сгенерированы 2 множества синтетических данных длительностью заданий – по 10 различных масси-

вов с приведенными ранее параметрами экспоненциального закона распределения, каждый из которых включает 10000 наблюдений. Полученные в результате расчетов значения метрик работы системы усредняются по количеству сгенерированных синтетических длительностей. На рис. 18–23

приведены результаты расчетов для различных показателей для данных, сгенерированных по вариантам а) и б), соответственно, для 50 ресурсов. Как следует из графиков, коэффициент использования с увеличением доли длинных заданий в генеральной тестовой выборке увеличился для всех методов, при этом лучшие результаты показывают методы МС (улучше-

ние составляет в среднем с 23% до 26%) и backfill (улучшение составляет в среднем с 22% до 25%); улучшение для метода FCFS составляет с 16% до 20%, но при этом ресурсы загружаются крайне неравномерно.

На рис. 24–29 приведены результаты моделирования для рассматриваемых вариантов для 100 ресурсов.

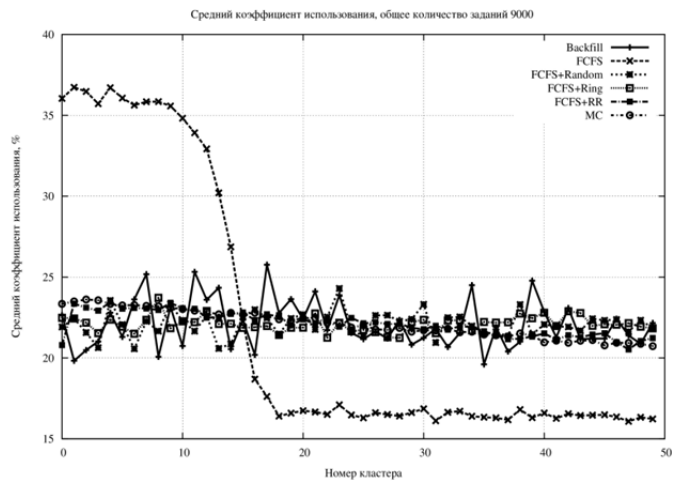
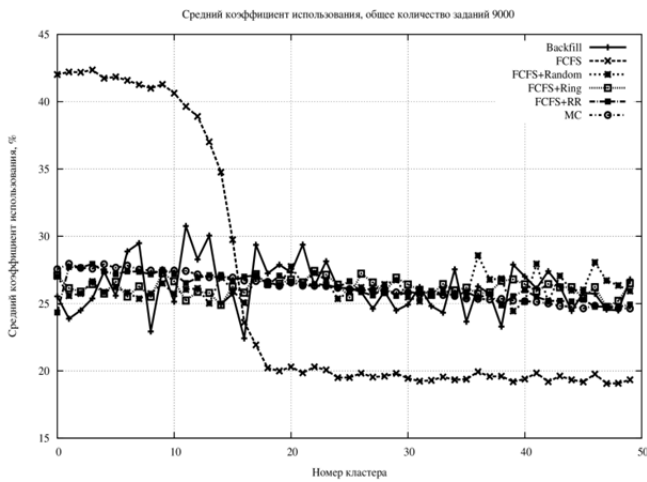


Рис. 18. Среднее значение КИ для варианта а) Рис. 19. Среднее значение КИ для варианта б)

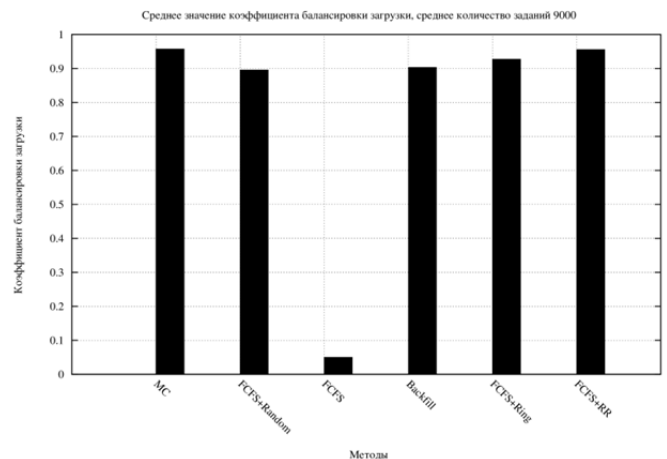
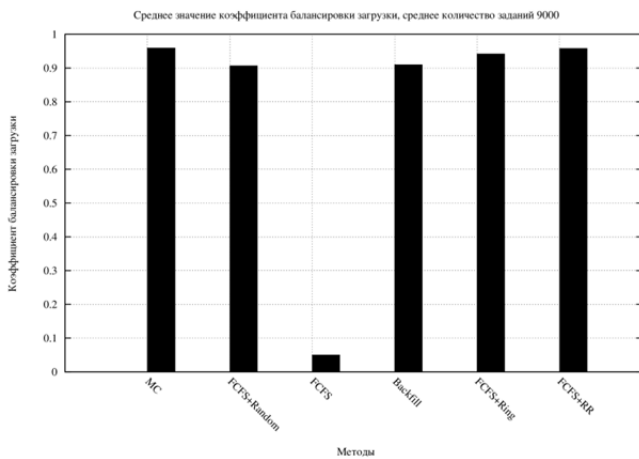


Рис. 20. Среднее значение КБ для варианта а) Рис. 21. Среднее значение КБ для варианта б)

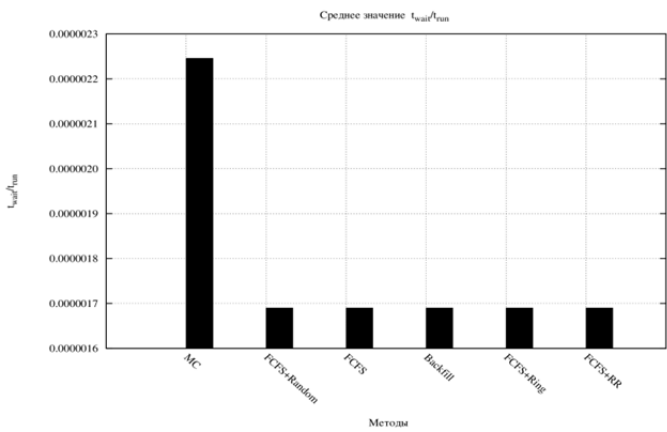
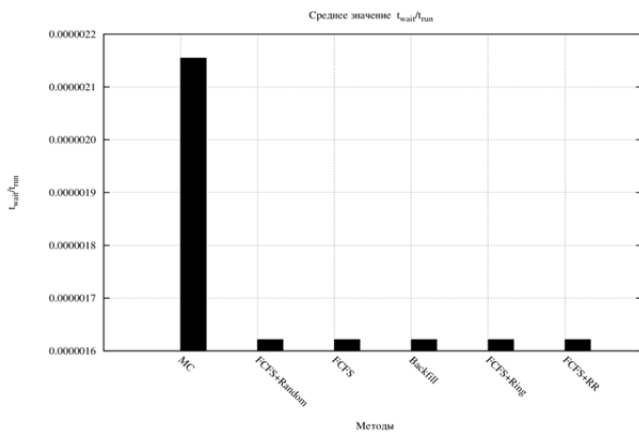


Рис. 22. Среднее значение $\frac{t_{wait}}{t_{run}}$ для варианта а)

Рис. 23. Среднее значение $\frac{t_{wait}}{t_{run}}$ для варианта б)

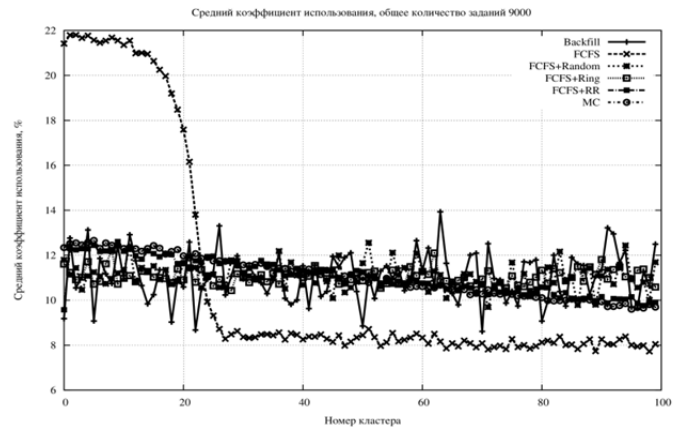
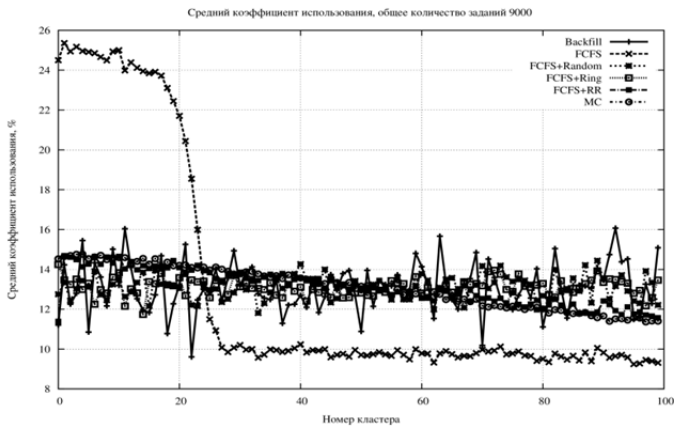


Рис. 24. Среднее значение КИ для варианта а)

Рис. 25. Среднее значение КИ для варианта б)

Приведенные на рис. 24, 25 результаты показывают, что для 100 ресурсов величина КИ для всех методов уменьшилась практически в 2 раза, при этом улучшение для отдельных методов снизилось до уровня: для MC – с 12% до 13%, для backfill – с 11% до 13%, для FCFS – 8% до 10%. Приведенные на рис. 26, 27 ре-

зультаты показывают, что среднее значение КБ для различных вариантов не изменилось, а среднее значение $\frac{t_{wait}}{t_{run}}$ несколько увеличилось, что предсказуемо для случая длинных заданий и постоянных интервалов планирования.

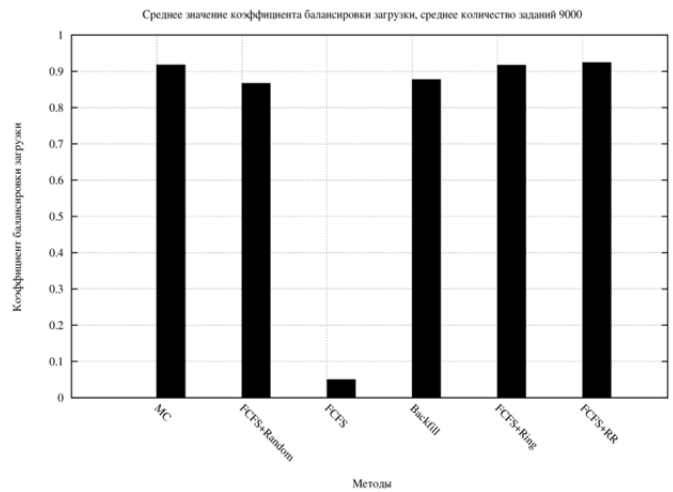
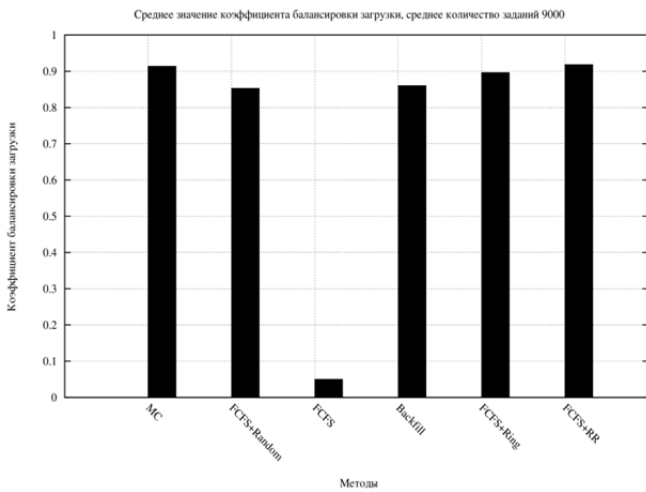


Рис.26. Среднее значение КБ для варианта а)

Рис.27. Среднее значение КБ для варианта б)

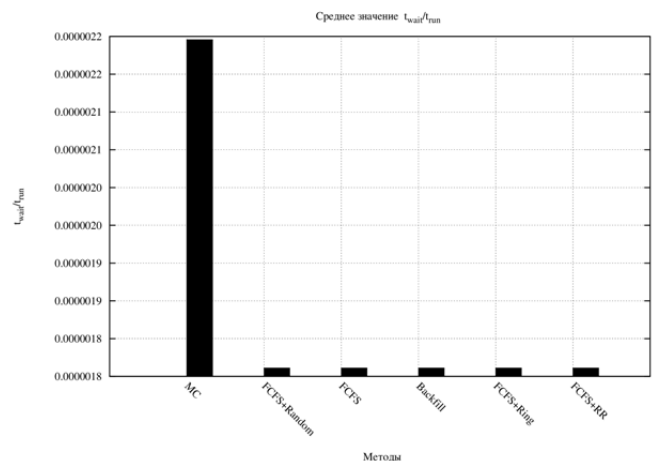
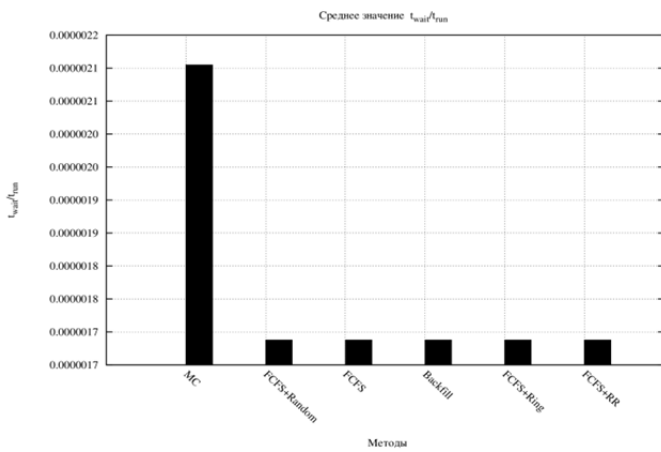


Рис. 28. Среднее значение $\frac{t_{wait}}{t_{run}}$ для варианта а)

Рис. 29. Среднее значение $\frac{t_{wait}}{t_{run}}$ для варианта б)

4. Выводы

В данном исследовании получены результаты, позволяющие сформулировать следующие выводы относительно оценки влияния характеристик поступающих на вычислительный кластер заданий на выбор стратегии планирования и производительность системы.

1. Для обоснованного выбора координированной стратегии планирования заданий необходимо проводить статистический анализ интенсивности поступающих заданий, их длительности, а также динамики интенсивности заданий, показывающей, как при этом может измениться производительность системы в целом.

2. Для выбора стратегии планирования заданий ключевым является, как показали проведенные исследования, соотношение между количеством длинных и коротких заданий, влияющего на оценки метрики производительности вычислительной системы. Для подтверждения этой гипотезы в работе предложено использовать генерацию синтетических (искусственных) данных об интенсивностях и длительностях заданий, имеющих одинаковые средние значения, соответствующие характеристикам реальных данных. При этом используется частотный анализ, в соответствии с которым можно достаточно точно определить долю длинных заданий

(их количество) в генерируемых случайных тестовых последовательностях.

3. Проведенные на синтетических данных исследования показали, что сортировка для массива с увеличенным количеством длинных заданий в соответствии со стратегией SJF практически не влияет на значения используемых метрик. Это, по-видимому, объясняется тем, что для определенных периодов (в исследовании они выбраны в среднем равным 600 заданиям), соотношение между длинными и короткими заданиями можно рассчитать, а для каких-то – нет, в силу неравномерного поступления заданий по периодам. Поэтому необходимо использовать динамическое планирование: например, с увеличением интенсивности заданий или увеличением доли коротких заданий уменьшать интервал планирования или определять его по среднему времени решения заданий на ресурсе данной вычислительной системы.

4. Качество обслуживания с увеличением доли длинных заданий падает (отношение $\frac{t_{wait}}{t_{run}}$ увеличивается), поэтому для обоснования правильности выбора координированной стратегии планирования необходимо определять приоритеты для критериев эффективности работы вычислительной системы.

Список литературы

1. Симоненко В.П. Математическая постановка задачи динамического распределения работ в grid системах и оценки качества решения. // Вісник НТУУ «КПІ» Інформатика, управління та обчислювальна техніка. – 2011. – №53. – С.37–41.
2. Симоненко В.П. Теоретические основы проектирования динамических пространственных планировщиков неоднородных GRID систем. // Электронное моделирование. – 2011. – Т. 33. – №5. – С. 57–71.
3. Симоненко А.В. Система пространственного распределения заданий в распределенных вычислительных системах. / А.В. Симоненко, С.В. Пих, Н.В. Слуцкий, В.В. Воробйов. // Вісник НТУУ «КПІ» Інформатика, управління та обчислювальна техніка. – 2012. – №56. – С.94–97.
4. Петренко А.І. Гібридний алгоритм брокера для Nordugrid ARC 2.0. /А.І Петренко, С.Я. Свистунов, П.В. Свірін.// Міжнародна конференція "Високопродуктивні обчислення" НРС-UA'2012 (Україна, Київ, 8-10 жовтня 2012 року. – С. 275–277.
5. Листровой С.В. Модель и подход к планированию распределения ресурсов в гетерогенных ГРИД-системах / С.В. Листровой, С.В. Минухин. // Проблемы управления и информатики. – 2012. – №5. – С.120–133.
6. Rodero I. Evaluation of Coordinated Grid Scheduling Strategies / I. Rodero, F. Guim, J. Corbalan // HPCC '09 Proceedings of the 2009 11th IEEE International Conference on High Performance Computing and Communications, 2009, p. 1–10.
7. Tsafir D. Backfilling Using Runtime Predictions Rather Than User Estimates / D.Tsafir, Y. Etsion, D. G. Feitelson. // IEEE Transactions on Parallel and Distributed Systems. – 2007. – V.18. – Iss. 6. – p.789–803.
8. Srinivasan S. Characterization of Backfilling Strategies for Parallel Job Scheduling / S. Srinivasan, R. Kettimuthu, V. Subramani, P. Sadayappan.// Proceedings of the ICPP Workshop on Scheduling and Resource Management for Cluster Computing, 2002.

9. Guim F. Prediction f based Models for Evaluating Backfilling Scheduling Policies / F. Guim, J. Corbalan, J. Labarta. // PDCAT '07 Proceedings of the Eighth International Conference on Parallel and Distributed Computing, Applications and Technologies, 2007, p. 9–17.
10. Guim F. A Job Self-Scheduling Policy for HPC Infrastructures / F. Guim, J. Corbalan. // JSSPP'07 Proceedings of the 13th international conference on Job scheduling strategies for parallel processing, 2007, p. 51–75.
11. Минухин С.В. Статистический анализ и исследование влияния методов работы с очередью входных заданий на функционирование распределенных вычислительных систем. / С.В. Минухин, А.В. Корвин. // Third International Conference "High Performance Computing" HPC-UA 2013 (Ukraine, Kyiv, October 7-11, 2013), 2013, С. 258–265.