

УДК 519.872:004

ИССЛЕДОВАНИЕ МОДЕЛИ ДИСПЕТЧЕРА С РАСПРЕДЕЛЕННОЙ ОЧЕРЕДЬЮ ЗАДАЧ РЕКОНФИГУРИРУЕМОЙ ВЫЧИСЛИТЕЛЬНОЙ СИСТЕМЫ

Мартышкин А.И.*ФГБОУ ВО «Пензенский государственный технологический университет»,
Пенза, e-mail: Alexey314@yandex.ru*

Современные архитектуры реконфигурируемых вычислительных систем соответствуют тенденциям наращивания мощностей высокопроизводительных вычислительных систем. Применение реконфигурируемых вычислительных систем сегодня весьма актуально и перспективно для разработчиков высокопроизводительных вычислительных систем. Для обеспечения еще большей производительности реконфигурируемых вычислительных систем, возможно, имеет смысл отказаться от программной реализации функциональных узлов высокопроизводительных вычислительных систем, что как раз и обусловлено гибкостью и современными возможностями элементной базы. В статье описана идея и методы реализации диспетчера задач, а именно предложен вариант реализации модели с распределенной очередью задач. Рассмотрены модели реконфигурируемых вычислительных систем с подсистемой диспетчеризации задач. Проведен анализ основных характеристик моделей диспетчеров задач. Результаты моделирования показали, что реконфигурируемая вычислительная система, включающая диспетчеры задач с распределенными очередями, работает эффективней и быстрее, чем аналогичная система с единым диспетчером задач и глобальной очередью. Этот вывод подтверждается согласно результатам полученных исследований, представленных на графиках. По итогам проведенных исследований предложены упрощающие разработку и исследование вычислительных систем модели диспетчеров задач, построенные на базе теории массового обслуживания. Предложенные и исследуемые модели можно применять для анализа процесса диспетчеризации и назначения ресурсов внутри реконфигурируемых вычислительных систем при проектировании и построении новых структур высокопроизводительных вычислительных систем. Достоверность полученных результатов подтверждается проведенными экспериментами и сопоставлением полученных результатов с известными на текущий момент. В заключение приводятся основные выводы по проведенному исследованию.

Ключевые слова: аналитическое моделирование, диспетчер задач, имитационное моделирование, модель, планирование процессов, реконфигурируемая вычислительная система, система массового обслуживания

INVESTIGATION OF A TASK MANAGER MODEL WITH A DISTRIBUTED QUEUE OF A RECONFIGURABLE COMPUTING SYSTEM

Martyshkin A.I.*Penza State Technological University, Penza, e-mail: Alexey314@yandex.ru*

Modern architectures of reconfigurable computing systems correspond to the trends of increasing the capacity of high-performance computing systems. The use of reconfigurable computing systems today is very relevant and promising for developers of high-performance computing systems. To ensure even greater performance of reconfigurable computing systems, it may make sense to abandon the software implementation of functional nodes of high-performance computing systems, which is precisely due to the flexibility and modern capabilities of the element base. The article describes the idea and methods of implementing a task manager, namely, a variant of implementing a model with a distributed task queue is proposed. Models of reconfigurable computing systems with a task dispatching subsystem are considered. The analysis of the main characteristics of the task dispatcher models is carried out. The simulation results showed that a reconfigurable computing system, including task managers with distributed queues, works more efficiently and faster than a similar system with a single task manager and a global queue. This conclusion is confirmed according to the results of the studies presented in the figures and graphs. Based on the results of the conducted research, the models of task controllers that simplify the development and research of computing systems, built on the basis of queuing theory, are proposed. The proposed and investigated models can be used to analyze the process of dispatching and assigning resources within reconfigurable computing systems when designing and building new structures of high-performance computing systems. The reliability of the results obtained is confirmed by the experiments carried out and the comparison of the results obtained with the currently known ones. At the article concludes, the conclusions of the study are presented.

Keywords: analytical modeling, task manager, simulation modeling, model, process planning, reconfigurable computing system, queuing system

На сегодняшний день методы и средства повышения производительности вычислительных систем (ВС), такие как более плотное размещение транзисторов на кристалле, увеличение числа одновременно выполняемых команд, расширение объемов кэш-памяти процессоров (ЦП), постепенно

исчерпывают свои возможности. Для того чтобы сохранить темпы роста производительности высокопроизводительных ВС (ВВС), недостаточно совершенствовать технологические решения, а также схемотехнику ЦП. В данной ситуации следует основательно рассмотреть способы пере-

стройки самой архитектуры BBC, например, представленной в виде реконфигурируемой вычислительной системы (PBC).

Синхронизация процессов в ВС с одним ЦП реализована программно в пространстве пользователя или ядра операционной системы (ОС), при этом не влияя на производительность BBC, тогда уже в PBC будет присутствовать достаточно серьезное увеличение времени выполнения при синхронизации процессов. Можно допустить, что программа, которая выполняет свои процессы в однопроцессорном и многопроцессорных режимах, имеет одинаковые затраты по времени на синхронизацию процессов. Затраты значительно могут отличаться в тот момент, когда будет наблюдаться уменьшение времени выполнения параллельных потоков. Упомянутый процесс обусловлен последовательностью связанного с синхронизацией процессов элемента программы и согласно закону Амдала [1] – это фактор, который снижает производительность PBC. Эффективное решение обозначенной проблемы – аппаратная реализация функций планирования и диспетчеризации задач, которая в значительной степени снимает проблему временных потерь при синхронизации.

Цель настоящей работы – проанализировать известные PBC и возможные способы диспетчеризации задач, а также разработать и исследовать математические модели таких систем с использованием широко применяемого аппарата теории массового обслуживания (ТМО). Описать возможности практического применения предложенных математических моделей. Достижению поставленной в работе цели способствует решение следующих задач: исследование, расчет и анализ параметров моделей диспетчеров задач (ДЗ) в составе PBC; подтверждение достоверности работы алгоритма ДЗ экспериментальным путем в режиме реального времени.

Материалы и методы исследования

К настоящему времени сложилась определенная классификация архитектур вычислительных систем, так называемая классификация по Флинну [2], которая анализирует архитектуру по способу обрабатываемых данных и взаимодействия потоков выполнения команд. К классу SISD относятся классические фон-неймановские машины. Архитектура SIMD формирует поток векторных команд, что позволяет производить арифметическую операцию над множеством данных. Векторная обработка может быть выполнена с помощью,

конвейера процессорной матрицей. Класс MISD предполагает архитектуру, в которой несколько ЦП осуществляют обработку одного и того же потока данных. Но данная архитектура так и не нашла практического применения. Широкое применение в сфере высокопроизводительных вычислений получила архитектура MIMD, на основе данной архитектуры строится большинство многопроцессорных параллельных ВС. MIMD разбивается на подклассы: мультипроцессоры и мультикомпьютеры. В свою очередь, мультипроцессоры делятся на UMA и NUMA-системы [1].

Система с общей памятью представляет собой структуру (рис. 1, а), в которой оперативная память является общей для всех ЦП. Достоинство такой архитектуры заключается в экономии времени на пересылку данных, так как эти данные, помещенные в общую память, доступны сразу всем ЦП с помощью общей шины. ЦП, имеющие одинаковый доступ к устройствам ввода-вывода и ко всем ячейкам оперативной памяти, к которой процессы обращаются единообразно и обращение происходит с одинаковым временем – такие системы называются симметричными (symmetric multiprocessor – SMP). Принцип действия подобной системы в том, что при помощи схемы приоритетов определенный ЦП получает данные из общей памяти, шина занята в течение времени передачи данных, другие ЦП отключаются от общей шины и выполняют задачи, не связанные с доступом в оперативную память.

Система с распределенной (сосредоточенной) памятью (рис. 1, б) содержит в себе множество ЦП, каждый из которых имеет свою собственную оперативную память. Чтобы организовать обмен данных между ЦП, существует канал связи, подобные системы делятся на два типа: с универсальной коммутацией, где любой ЦП может быть связан с любым другим, и с жесткой коммутацией, в которой связь определенного ЦП существует только с определенным количеством других ЦП. Достоинство данной архитектуры в эффективном решении задач, так как ЦП не требуется ожидать освобождения оперативной памяти. Архитектура NUMA обеспечивает обращение к данным при физическом распределении памяти, поэтому длительность доступа является неодинаковой для всех элементов памяти. Системы с распределенной памятью делятся еще на два типа, с локальной кэш-памятью (cache only memory architecture – COMA) ЦП и системы с когерентностью локальной кэш-памяти ЦП (cache coherent NUMA – CC-NUMA).

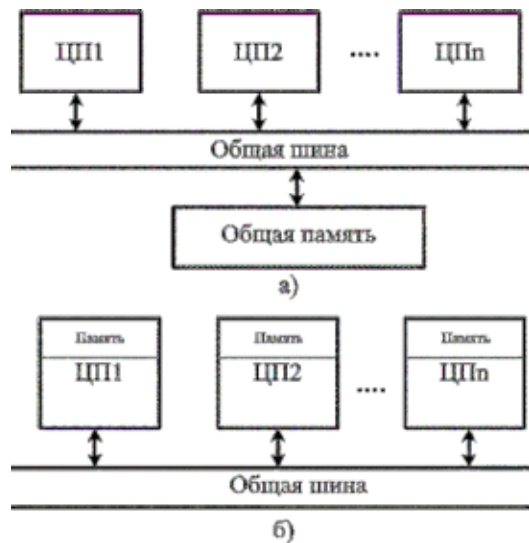


Рис. 1. Структурная схема ВС с общей (разделяемой) (а) и распределенной (б) памятью

В России известны и применяются следующие РВС – это, прежде всего, разработки ведущих научно-исследовательских институтов и крупных университетов, специализирующихся на РВС: РВС-1, РВС-0.2-РС, РУПК-50, БВР-01, KB8-K7-410-01 «Топаз-2» и др. [3, 4]. Разработчики компании Fastwel создали вычислитель БВР-01. Устройство направлено на решение таких задач, как цифровая обработка сигналов, применение методов линейной алгебры, символьной обработки, особенности – гибкая структура и возможности объединения с подобными устройствами для создания кластерных систем. Архитектура вычислителя БВР-01 (рис. 2, а) состоит из управляющего компьютера (УК) CPC1301, 8 ПЛИС Xilinx Virtex-6, модуль памяти EM-DDR3-2GB. ПЛИС соединены между собой высокоскоростным интерфейсом PCI Express Gen 2.0, что позволяет обмен данных между УК и 8 ПЛИС на скорости до 20 Гбит/с. РВС на базе Xilinx Kintex-7 KB8-K7-410-01 «Топаз-2» (рис. 2, б) создана в ФГУП НИИ «Квант». Разработанное устройство состоит из одной системной и восьми рабочих ПЛИС, которые соединены к коммутатору PCI Express, данный коммутатор представляет собой микросхему PEX8648 на 12 портов. Для соединения ПЛИС используются схемы «звезда» и кольцо, что обеспечивает высокоскоростной канал обмена информации. Разработчики ООО НПО «Роста» создали РВС RC-47 (рис. 2, в), на которой установлено четыре рабочих ПЛИС Xilinx семейства Virtex-7, также на плате установлены два мезонинных модуля памяти EM4-DD, где в каждом модуле по 4 банка памяти объемом 512 Мбайт. Основные технические характе-

ристики платы следующие, сервисная ПЛИС: Xilinx Spartan-6, четыре рабочих ПЛИС: Xilinx Virtex-7 (XC7V585/1500/2000T, XC7VX330/485/690T) FF(G) 1761, все рабочие ПЛИС поддерживают PCI Express Gen2 (для серии XC7V) или Gen3 (для серии XC7VX), сеть PCI Express: одна СБИС PEX8732 коммутатора PCI Express Gen3 x4 (8 портов), микроконтроллер STM32F, обеспечивающий контроль напряжений и температур, загрузку конфигураций в системный и рабочие ПЛИС, настройку и мониторинг коммутатора PCI Express по шине I2C [4].

Диспетчеризация задач. Суть планирования и диспетчеризации процессов и задач в РВС заключается в назначении целевого ЦП для ее решения. Современные работы по планированию в операционных системах говорят о двух часто применяемых способах организации работы ДЗ: с глобальной и распределенной очередью задач [5].

Для исследования параметров ВВС применяют простой метод, основанный на различных системах массового обслуживания (СМО), – абстрактный объект, который включает в себя заявки (задачи), поступающие в систему, несколько обслуживающих каналов (приборов) заявок и накопитель, где организуется очередь из поступивших заявок, ожидающих обслуживания. СМО или базовые модели можно соотносить в соответствии со следующей классификацией: по числу классов заявок, поступающих в СМО; по количеству мест в накопителе; по количеству обслуживающих приборов. Основные параметры и характеристики СМО и построенных на их базе сетей подробно рассмотрены и описаны в работах [6].

В РВС с общей памятью программы обработки и данные, связанные с ними $P = \{P_1, \dots, P_M\}$, делятся на некоторые подмножества Q_1, \dots, Q_N , $Q_i = \{P_{\alpha_1}, \dots, P_{\alpha_{\omega_i}}\}$ $P(i = 1, \dots, N)$, и распределяются в памяти процессора $ЦП_1, \dots, ЦП_N$. По этой причине каждый из n -процессоров обрабатывает данные с размещенными в памяти ЦП программами обслуживания. При такой ситуации можно принять процесс работы РВС как функционирование совокупности n -одноканальных СМО, причем каждая из них включает в себя некоторый поступающий с интенсивностью λ_i поток заявок, очередь O_i и обслуживающее устройство (процессор) $ЦП_i$. Для того чтобы выявить наиболее эффективный метод реализации при проектировании РВС, требуется произвести оценку потери производительности исследуемой системы. В данной работе рассмотрен алгоритм диспетчеризации с распределенной очередью.

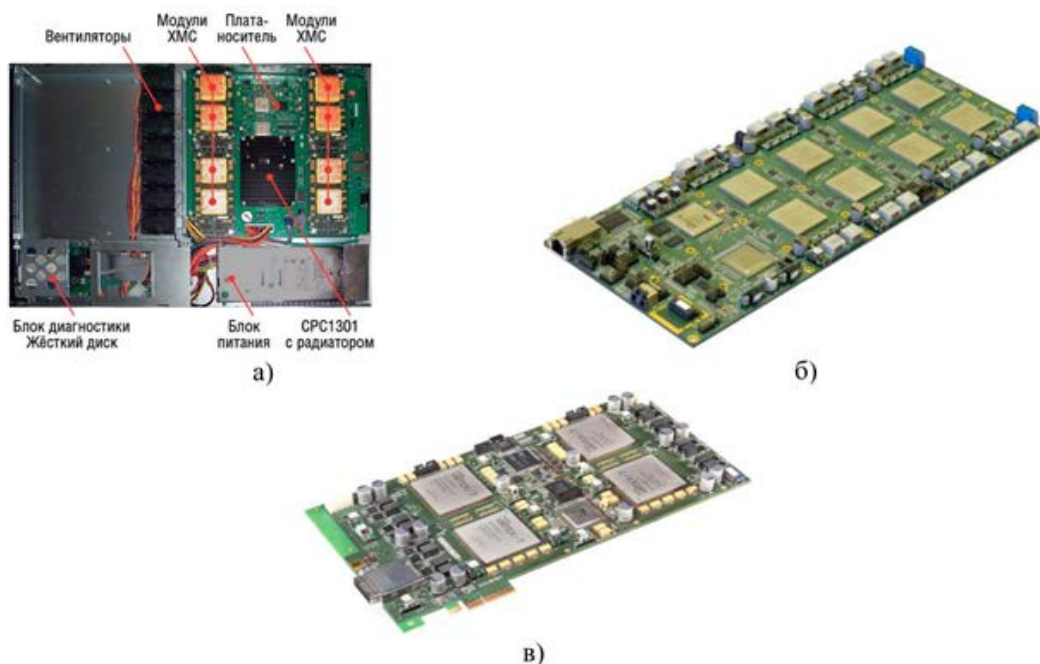


Рис. 2. БВР-01 (а); PBC KB8-K7-410-01 «Тоназ-2» (б); PBC RC-47 (в)

ДЗ с общей очередью [5] имеет один существенный недостаток в структуре организации, который понижает производительность всей РВС. Недостаток заключается в том, что возникают конфликтные ситуации при обращении к ДЗ, т.е. только определенный ЦП может обратиться к очереди задач. При этом ЦП требуется обращаться к ДЗ, чтобы получить новую заявку, последние два случая требуют дополнительных затрат по времени. Возможен еще случай, когда в РВС есть не занятые работой ЦП, но обслуживание ожидающих задач не осуществляется, поскольку ДЗ не справляется с потоком поступающих на обработку задач. Использование другого по архитектуре ДЗ с индивидуальными очередями процессов к ЦП – один из путей разрешения проблемной ситуации.

Математическая модель ДЗ с индивидуальными очередями состоит из n -одноканальных СМО (S_1, \dots, S_m) (рис. 3, а), которые моделируют обслуживание в ДЗ и ЦП (S_1, S_2, \dots, S_m) и граф передач. Структура и принцип взаимодействия блоков модели не отличается от ДЗ с общей очередью, единственное, что РВС с n ДЗ кроме формирования очередей и назначения задач на обслуживание в ЦП согласно некоторому алгоритму осуществляет балансировку наполненности очередей, когда заявки могут с определенной вероятностью (чем более загружена очередь, тем больше вероятность извлечения задачи из нее) перемещаться из i -й очереди в менее занятую очередь j -го процессора ЦП.

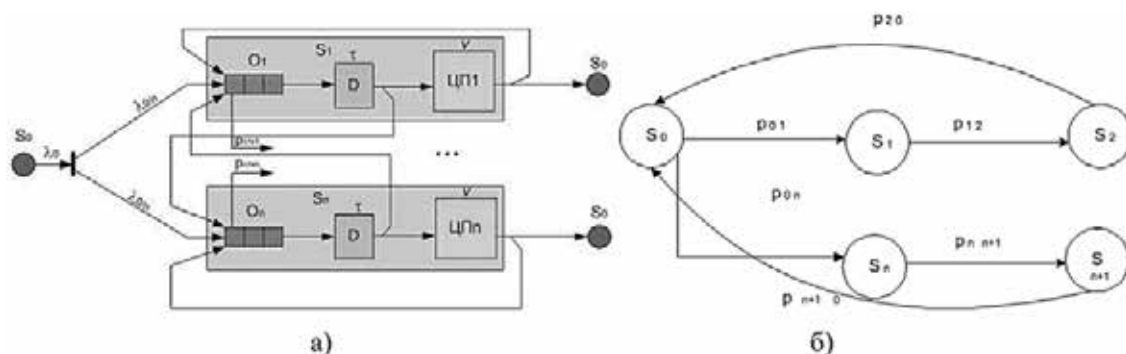


Рис. 3. Схема модели РВС с ДЗ с распределенным ДЗ (а) и ее графовая модель (б)

Параметры моделей

Трудоемкость	Интенсивность входного потока, задач/мс	Число ЦП	Средняя занятость ЦП обработкой заявки, мс
Высокая	9,3	4	0,009
	93	16	
Средняя	0,333	4	0,068
	3,33	16	
Низкая	0,07	4	0,15
	0,7	16	

Поступающий на входы n ДЗ РВС характеризуется интенсивностью, которая описывается выражением $\lambda_{00} = \lambda_0 / n$. ДЗ обрабатывает поступающие задачи с некоторой интенсивностью μ_D . Когда ДЗ занят обработкой текущей задачи, другие поступающие на обслуживание требования ожидают своего сервиса очереди. В случае переполнения очереди поступающие задачи перенаправляются в другую менее занятую очередь. После описания характеристик перейдем непосредственно к построению моделей в программах [7, 8]. С помощью графового представления (рис. 3, б) можно рассчитать вероятностно-временные характеристики выхода ($p_{20}, p_{n+1,0}$) задачи, полностью получившей обслуживание, и вероятностно-временные характеристики возвращения ($p_{12}, p_{n,n+1}$) на дообслуживание в ДЗ. Приведенные вероятностно-временные характеристики напрямую связаны с трудоемкостью поступившей задачи: чем задача более трудоемкая, тем дольше она обрабатывается в ЦП. При исследова-

нии приняты параметры, представленные в таблице. Данные параметры входной интенсивности поступающих задач на обслуживание обеспечивают среднюю загрузку ЦП примерно до уровня 65–70%. Принято, что длина очереди перед ДЗ составляет 128 задач. Среднее время активности ДЗ принято равным 0,002 мс [9].

Обзор методов математического описания характеристик РВС показывает, как можно простым и наглядным образом получить значения характеристик РВС, не затрачивая при этом больших ресурсов и средств. Для проведения эксперимента получены и определены параметры исследуемых моделей, которые нашли свое применение при построении аналитических и имитационных моделей РВС.

Результаты исследования и их обсуждение

В ходе проведенных экспериментов над моделями получены результаты, которые отражены на графиках на рис. 4–6.

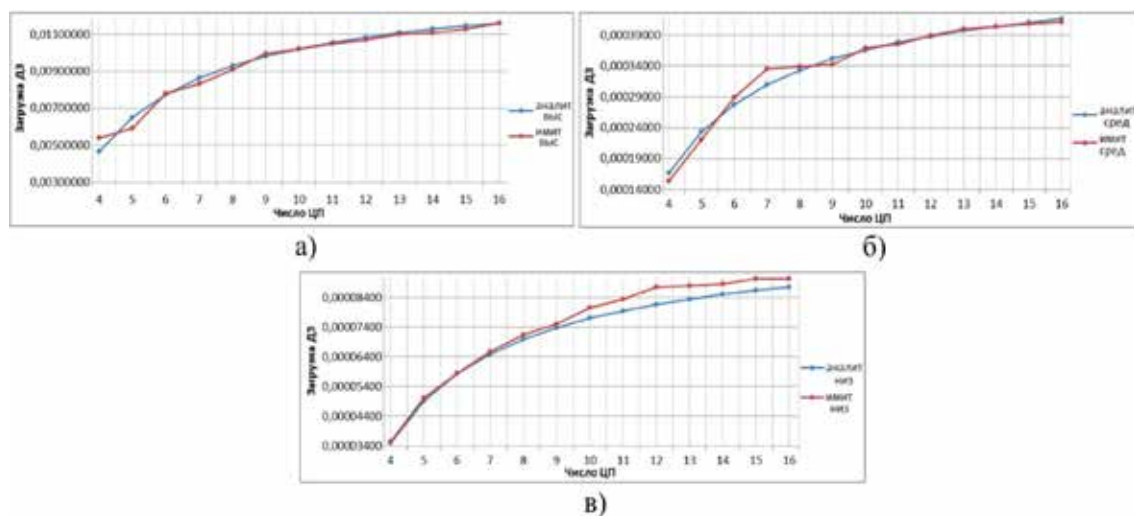


Рис. 4. Зависимость загрузки ДЗ с распределенной очередью от числа ЦП при различной трудоемкости: высокой (а) средней (б) и низкой (в)

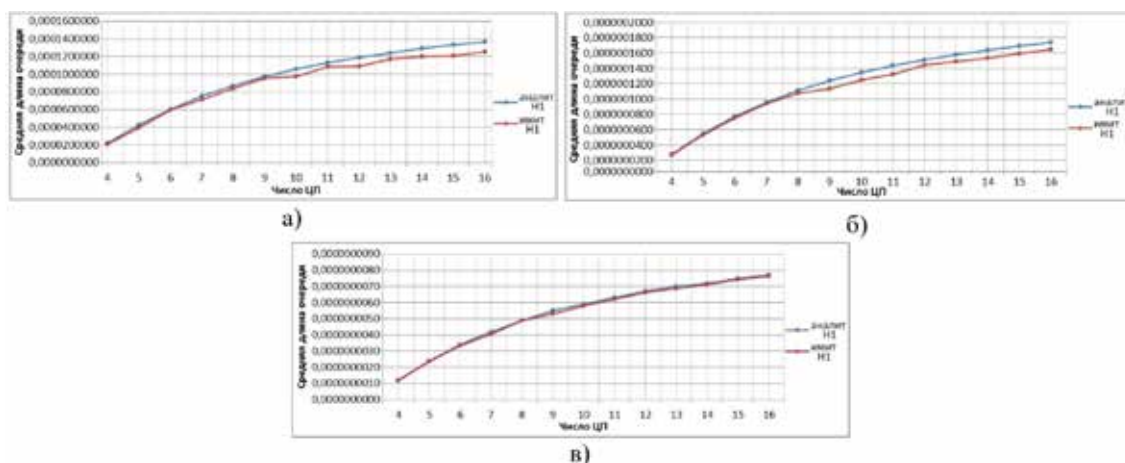


Рис. 5. Соотношение средней длины очереди H1 в PBC от числа ЦП при высокой (а) средней (б) и низкой (в) трудоемкости поступающих на обслуживание задач

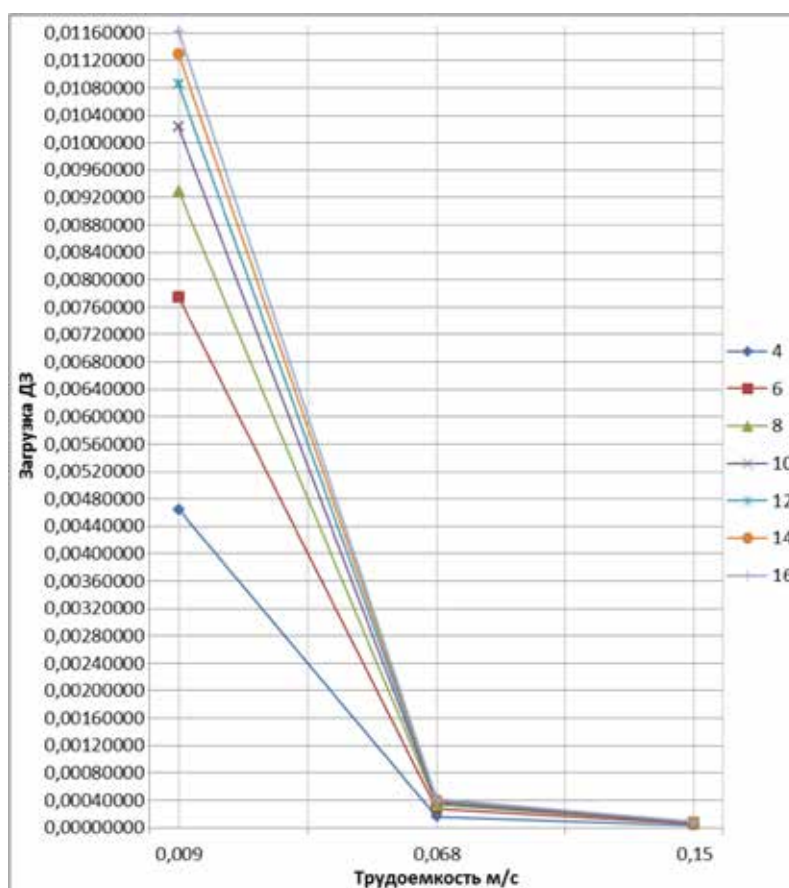


Рис. 6. Трудоемкость задач в PBC, влияющая на загрузку ДЗ с распределенной очередью

Графики, изображенные на рис. 4, показывают плавную зависимость загрузки ДЗ при увеличении числа ЦП. Отличим от системы с общей очередью также является то, что в системе с распределенными очередями ДЗ значения загруз-

ки примерно в 10 раз меньше при всех трех значениях трудоемкости. Данные результаты дают основание утверждать, что система ДЗ с распределенной очередью имеет большую производительность, чем аналогичная.

Графики, изображенные на рис. 5, показывают соотношение средней длины очереди перед ДЗ(Н1) от количества ЦП в РВС. Данное соотношение, в сравнении с зависимостью в системе ДЗ с общей очередью [9], имеет равномерный характер роста при росте числа ЦП, само же значение средней длины очереди имеет разницу в 100 раз.

Графики, приведенные на рис. 6, характеризуют зависимость загрузки ДЗ с распределенной очередью от трудоемкости задач в РВС с числом ЦП от 4 до 16 соответственно. Отметим, что в РВС с ДЗ с распределенными очередями загрузка примерно в 10 раз меньше, чем в аналогичной системе с другим типом ДЗ [9].

Сделаем выводы по проведенным исследованиям. При проведении экспериментов на исследуемых моделях значения основных характеристик (коэффициент загрузки ДЗ, время реакции РВС, средняя длина очереди перед ДЗ) имели небольшой разброс, что подчеркивает адекватность разработанных моделей и целесообразность их применения при проведении аналогичных исследований.

Заключение

В работе проанализированы современные РВС и подсистемы диспетчеризации задач, в том числе рассмотрены возможности практического применения моделей в виде реализации функциональных узлов-диспетчеров задач в составе РВС. Основными достигнутыми результатами являются пред-

ложенные модели ДЗ с целью исследования влияния различного вида производительности, верифицированные экспериментально.

Исследование выполнено за счет гранта Российского научного фонда № 21-71-00110.

Список литературы

1. Таненбаум Э., Остин Т. Архитектура компьютера. 6-е изд. СПб.: Питер, 2013. 816 с.
2. Орлов С.А., Цилькер Б.Я. Организация ЭВМ и систем: учебник для вузов. 3-е изд. Стандарт третьего поколения. СПб.: Питер, 2014. 688 с.
3. Каляев И.А., Левин И.И., Семерников Е.А., Дордопуло А.И. Реконфигурируемые вычислительные системы на основе ПЛИС семейства VIRTEX-6 // Вестник Уфимского государственного авиационного технического университета. 2011. Т. 15. № 5 (45). С. 148–154.
4. Левин И.И., Дордопуло А.И., Каляев И.А., Гудков В.А. Высокопроизводительные реконфигурируемые вычислительные системы на основе ПЛИС VIRTEX-7 // Параллельные вычислительные технологии (ПАВТ'2014): труды Международной научной конференции. 2014. С. 131–139.
5. Таненбаум Э., Бос Х. Современные операционные системы. СПб.: Питер, 2015. 1120 с.
6. Алиев Т.И. Основы моделирования дискретных систем. СПб.: СПбГУ ИТМО, 2009. 363 с.
7. Мартышкин А.И. Программа для расчета основных вероятностно-временных характеристик подсистемы планирования и диспетчеризации многопроцессорных систем // Свидетельство о регистрации программы для ЭВМ RU 2016619353, 18.08.2016. Заявка № 2016616801 от 27.06.2016.
8. Юданова В.В. Имитационное моделирование систем массового обслуживания // Отходы и ресурсы. 2019. № 4. URL: <https://resources.today/PDF/23INOR419.pdf> (дата обращения: 10.04.2022).
9. Мартышкин А.И. Разработка и анализ математической модели диспетчера задач реконфигурируемой вычислительной системы // Современные наукоемкие технологии. 2022. № 3. С. 73–79.