УДК 681.3

Л. П. Фельдман, д-р техн.наук, **Т. В. Михайлова**
Донецкий национальный технический университет
(Украина, 83000, Донецк, ул.Артема, 58,
тел.: (062) 3010757, E-mail: feldman@r5.dgtu.donetsk.ua, tanya@r5.dgtu.donetsk.ua)

Использование аналитических методов для оценки эффективности многопроцесорных вычислительных систем

Предложены модифицированные методы анализа и синтеза многопроцессорных вычислительных ресурсов различной топологии с помощью вероятностных моделей, позволяющие анализировать и проектировать более широкий класс параллельных вычислительных сред.

Запропоновано модифіковані методи аналізу і синтезу багатопроцесорних обчислювальних ресурсів різної топології за допомогою аналітичних моделей. Ці методи дозволяють аналізувати і проектувати більш широкий клас паралельних обчислювальних середовищ.

Ключевые слова: методы анализа и синтеза вычислительных систем, кластер, эффективность, дискретные и непрерывные марковские модели.

Актуальной проблемой в настоящее время является оценка эффективности многопроцессорных вычислительных систем. Один из способов решения этой проблемы — использование непрерывных [1—3] или дискретных аналитических моделей [4]. Дискретные модели Маркова, отображающие модель вычислительной среды наиболее точно, имеют большую размерность и эффективно распараллеливаются [5]. Непрерывные модели менее трудоемкие, поэтому более широко применяются для исследований.

В настоящее время широкое распространение получили кластерные системы. По критерию совместного использования дискового пространства их можно классифицировать следующим образом: с совместным использованием дискового пространства и без предоставления доступа к ресурсам [6].

Анализ кластерных систем. На основании методик, описанных в работах [7, 8], можно построить модели кластеров с совместным использованием дискового пространства (рис. 1 и 2) и без предоставления доступа к ресурсам.

Построим модель неоднородного кластера (рис. 1), которая содержит M рабочих станций пользователей и $N1$ серверов, выполняющих одинако-

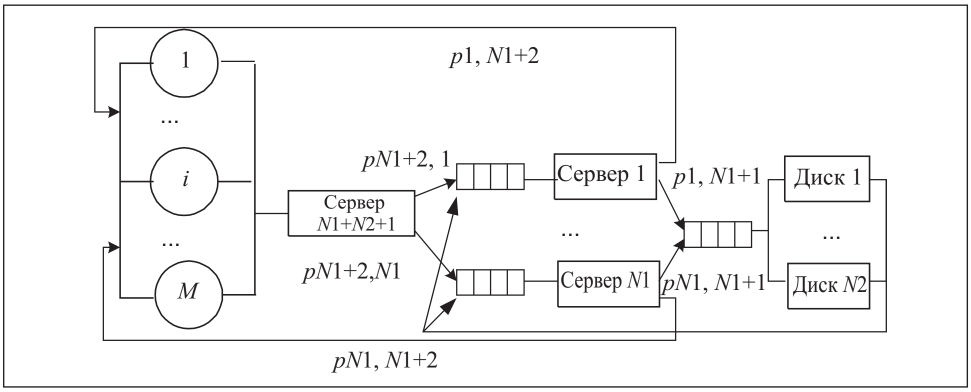

Рис. 1. Структура неоднородного кластера

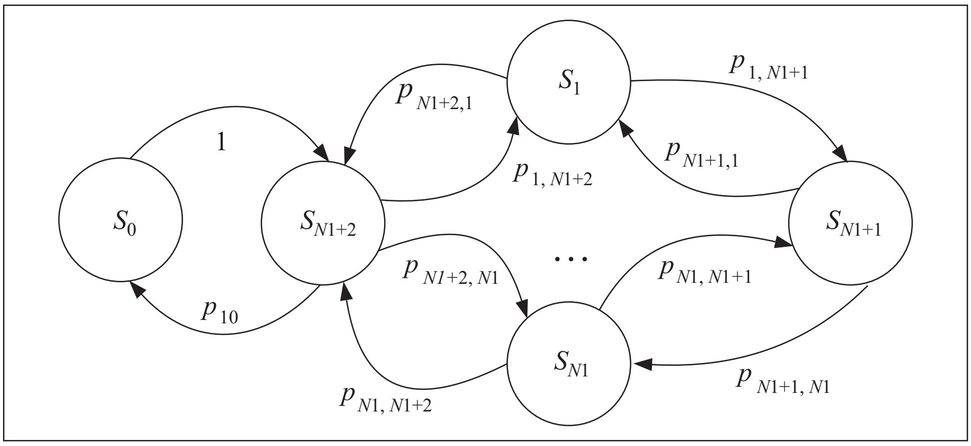

Рис. 2. Граф передач неоднородного кластера

вые приложения. База данных расположена на одном диске, при необходимости может иметь «зеркальное» отображение на нескольких дисках.

В каждом из кластеров M рабочих станций пользователей, $N1$ серверов, $N2$ дисковых массивов.

Представим серверы и диски многоканальными устройствами, время обслуживания которых имеет экспоненциальное распределение со средним параметром μ_i, $i = 1, \ldots, N1 + N2 + 1$. Требования, поступающие на обслуживание, становятся в очередь, из которой выбираются на обслуживание по правилу «первый пришел первый обслужен».

Каждый из M пользовательских запросов с вероятностью $p_{N1+2,i}$ обращается к i-му серверу, который, в свою очередь, обрабатывая этот запрос, обращается к одному из $N2$ дисков с вероятностью $p_{i,N1+1}$. Вероятности $p_{N1+1,i}$ $(i=\overline{1,N1})$ вычисляются по формуле

$$p_{N1+1,i}=\frac{p_{i,N1+1}}{\sum_{l=1}^{N1}p_{l,N1+1}}.$$

Функционирование рассматриваемой системы можно представить замкнутой стохастической сетью, содержащей $N1+2$ систем массового обслуживания (СМО), в которой циркулирует M заявок. Граф передач этой сети изображен на рис. 2, на котором обозначено: $S_{N1+2,1}$ — СМО, соответствующая главному серверу, $S_1, \ldots, S_{N1}$ — СМО, соответствующие серверам; S_{N1+1} — СМО, соответствующая дисковым пространствам.

Введем вектор $\mathbf{k}=(k_1,k_2,k_3)$, определяющий число устройств в узле $(k_1=1,\ k_2=N1,\ k_3=N2)$. По графу передач можно определить коэффициенты посещений каждой СМО, решив следующую систему уравнений:

$$\begin{aligned}
&\alpha_{N1+2}=1+p_{1,N1+2}\alpha_1+\ldots+p_{N1,N1+2}\alpha_{N1},\\
&\alpha_{N1+1}=p_{1,N1+1}\alpha_1+\ldots+p_{N1,N1+1}\alpha_{N1},\\
&\alpha_1=p_{N1+2,1}\alpha_{N1+2}+p_{N1+1,1}\alpha_{N1+1},\\
&.\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\\
&\alpha_i=p_{N1+2,i}\alpha_{N1+2}+p_{N1+1,i}\alpha_{N1+1},\\
&.\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\ .\\
&\alpha_{N1}=p_{N1+2,N1}\alpha_{N1+2}+p_{N1+1,N1}\alpha_{N1+1}.
\end{aligned}$$

Состоянием системы будем считать распределение заявок по СМО:

$$\mathbf{m}=(m_{N1+2},m_1,\ldots,m_{N1},m_{N1+1}),$$

где $m_{N1+2}+m_1+\ldots+m_{N1}+m_{N1+1}=M$. По теореме Джексона вычисляем стационарные вероятности, позволяющие определить характеристики вычислительной среды:

$$\pi(\mathbf{m})=\frac{1}{Q(M,N)}\prod_{j=1}^{N}R_j(m_j)(\alpha_j v_j)^{m_j}.$$

Здесь

$$Q(M,N)=\sum_{\Omega}\prod_{j=1}^{N}R_j(m_j)(\alpha_j v_j)^{m_j},$$

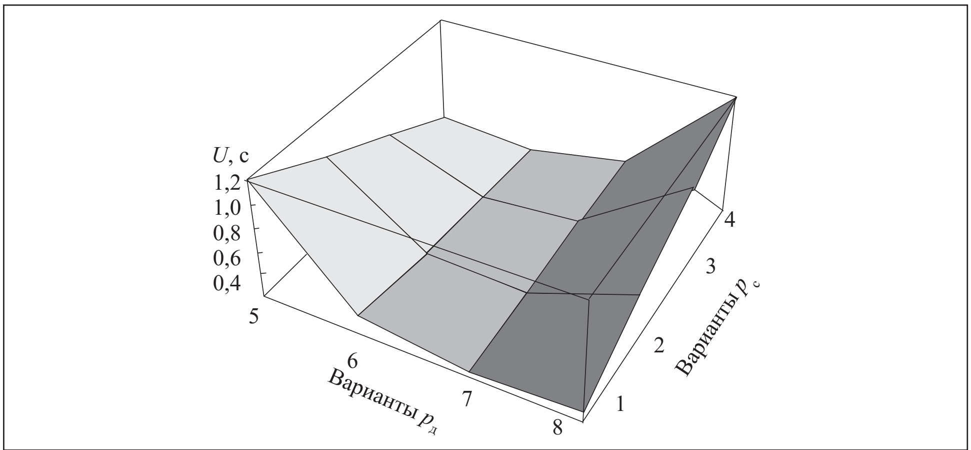

Рис. 3. Графики зависимости времени отклика от вероятностей обращения к диску и к серверу

где

$$R_j(m_j) = \begin{cases} 1/m_j!, m_j \le k_j, \\ 1/k_j! k_j^{m_j - k_j}, m_j > k_j, \end{cases} \qquad \nu_i = \begin{cases} m_i/\mu_i!, m_i \le k_i, \\ k_i/\mu_i, m_i > k_i, \end{cases}$$

Ω — все возможные состояния.

Используя стационарные вероятности, находим основные характеристики вычислительной среды: загрузка устройств, среднее число занятых устройств в s-м узле, среднее число задач, находящихся в s-м узле, среднее число задач, находящихся в очереди к s-му узлу, средние времена пребывания и ожидания в s-м узле, средние времена пребывания и ожидания в системе [4, 8].

Проанализируем этот кластер при $N2 = 2$, $N1 = 2$. Классы задач, решаемых в вычислительной среде, представлены в табл. 1 и табл. 2. Интенсивность обслуживания ν_i, $\nu_i = \theta_i / V_i$, $i = \overline{1, N}$, каждого этапа задачи в

Таблица 1

Номер варианта	Вероятность обращения к серверам, $p_с$	
	$p(5,1)$	$p(5,2)$
1	0,83125	0,11875
2	0,59375	0,35625
3	0,35625	0,59375
4	0,11875	0,83125

Таблица 2

Номер варианта	Вероятность обращения к дискам, $p_д$	
	$p(2,3)$	$p(1,3)$
5	0,125	0,875
6	0,375	0,625
7	0,625	0,375
8	0,875	0,125

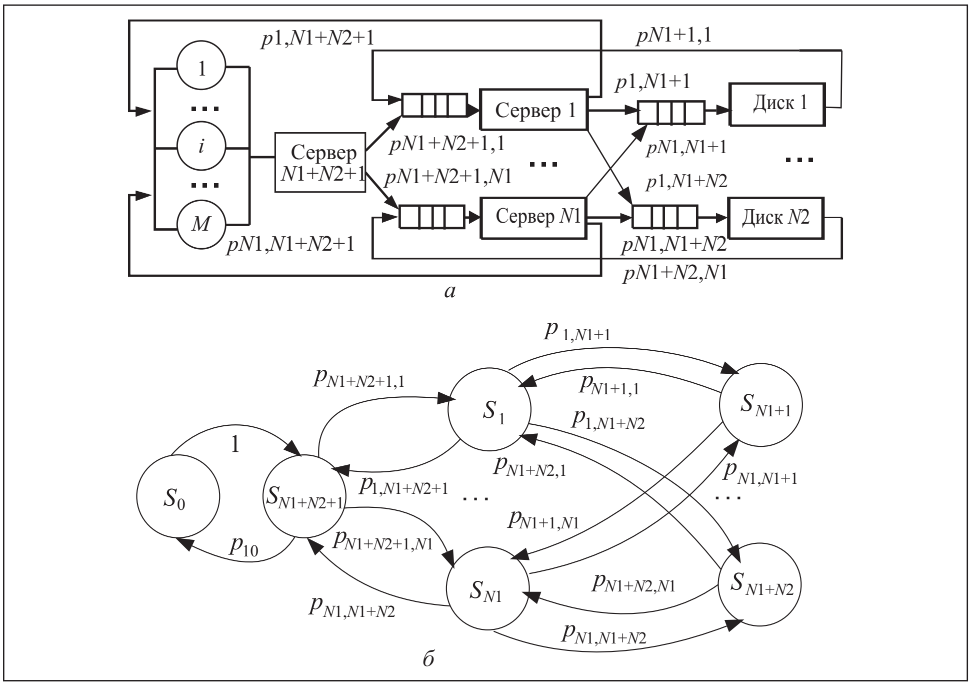

Рис. 4. Структура (*а*) и граф передач (*б*) кластера топологии $N{\times}N$

каждом из устройств изменялась следующим образом: на серверах $\nu_1 = \nu_2 = \{6, 12, 24\}\cdot 10^{-4}$ с, на дисках $\nu_3 = \nu_4 = \{0{,}35; 0{,}7; 1{,}5; 3\}\cdot 10^{-2}$ с. Число задач M =10.

Зависимости времени решения задачи U от вероятности обращения к серверам и дискам показаны на рис. 3. Эффективны классы задач, соответствующие вариантам (1, 7) и (4, 6), а также вариантам (1, 8), (2, 7), (4, 5), (3, 6), у которых время решения задачи наименьшее.

Аналогично можно построить модель кластера топологии $N \times N$ (рис. 4, *а*). По графу передач, изображенному на рис. 4, *б*, определяем коэффициенты посещений каждой СМО, решая систему уравнений

$$\alpha_{N1+N2+1} = 1 + \sum_{i=1}^{N1} p_{i,N1+N2+1}\alpha_i,$$

$$\alpha_i = p_{N1+N2+1,i}\,\alpha_{N1+N2+1} + \sum_{j=N1+1}^{N1+N2} p_{j,i}\,\alpha_j,\ i=\overline{1,N1},$$

$$\alpha_i = \sum_{j=1}^{N1} p_{j,i}\,\alpha_j,\ i=\overline{N1+1,N1+N2}.$$

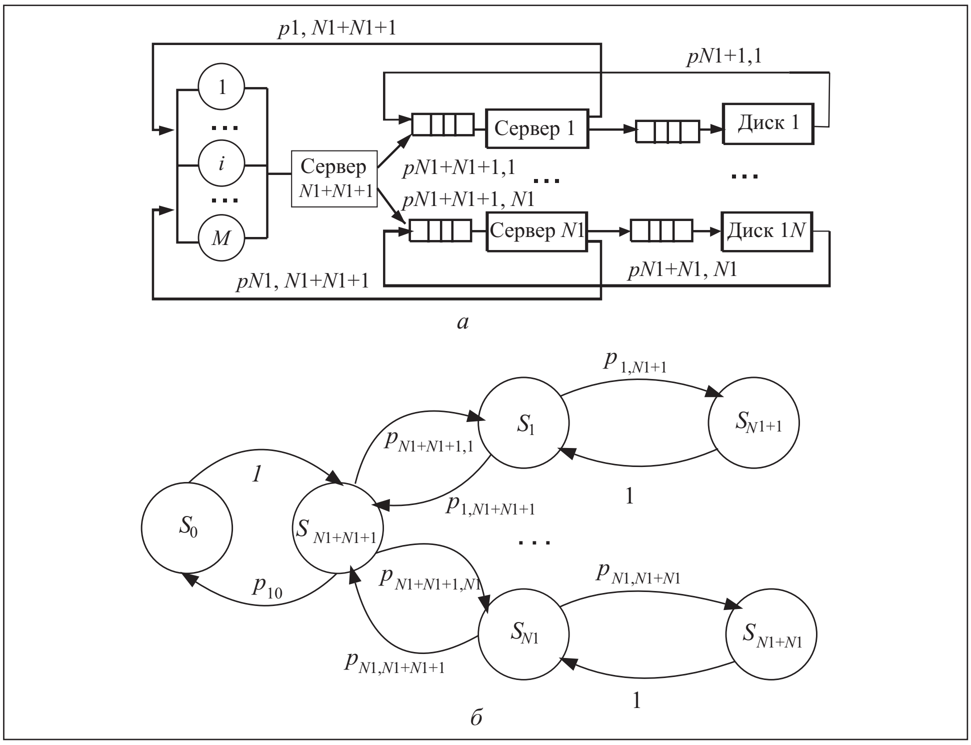

Рис. 5. Структура (*а*) и граф передач (*б*) кластера без разделения ресурсов

Состояние модели определяется распределением заявок по всем СМО:

$$\mathbf{m} = (m_{N1+N2+1}, m_1, ..., m_{N1}, m_{N1+1}, ..., m_{N1+N2}),$$

где $m_{N1+N2+1} + m_1 + ... + m_{N1} + m_{N1+1} + ... + m_{N1+N2} = M$.

Для модели кластера без предоставления доступа к ресурсам (рис. 5, *а*) граф передач представлен на рис. 5, *б*. Коэффициенты посещений вычисляем из системы уравнений

$$\alpha_{N1+N1+1} = p_{N1+N1+1,0} + p_{1,N1+N1+1}\alpha_1 + ... + p_{N1,N1+N1+1}\alpha_{N1},$$
$$\alpha_1 = \alpha_{N1+1} + p_{N1+N1+1,1}\alpha_{N1+N1+1},$$
$$. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad .$$
$$\alpha_{N1} = \alpha_{N1+N1} + p_{N1+N1+1,N1}\alpha_{N1+N1+1},$$
$$\alpha_{N1+1} = p_{1,N1+1}\alpha_1,$$
$$. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad .$$
$$\alpha_{N1+N1} = p_{N1,N1+N1}\alpha_{N1}.$$

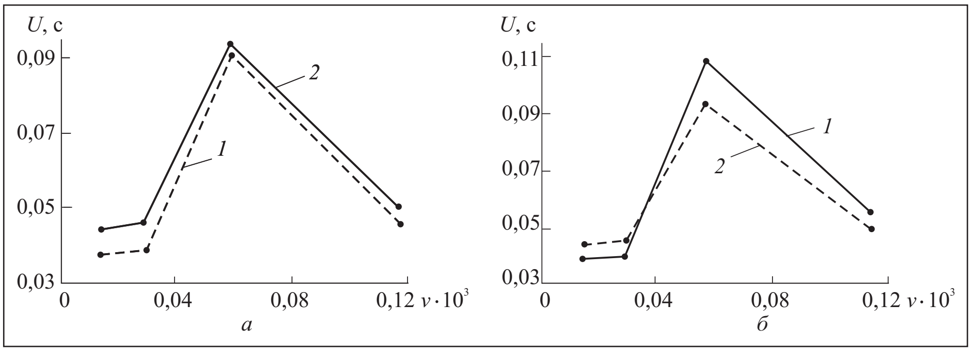

Рис. 6. Зависимость времени отклика от средней длительности обслуживания заявки на сервере при решении задач с равномерным (*а*) и неравномерным (*б*) обращением к серверам

Состояние модели определяется распределением заявок по всем СМО:

$$\mathbf{m} = (m_{N\,1+N\,1+1}, m_1, ..., m_{N\,1}, m_{N\,1+1}, ..., m_{N\,1+N\,1}),$$

где $m_{N\,1+N\,2+1} + m_1 + ... + m_{N\,1} + m_{N\,1+1} + ... + m_{N\,1+N\,2} = M$.

Используя полученные при моделировании стационарные вероятности, можно вычислить основные характеристики кластерных систем [4, 8]. Если анализируется конкретный вид кластера для решения определенного класса задач, то с помощью этих характеристик определяется эффективность вычислительной среды (в зависимости от критерия: равномерная загрузка всех узлов, минимальное время отклика и др.) [8].

Для выбора оптимального коэффициента мультипрограммирования используется методика, описанная в работе [9]. В ней предлагается критерий сбалансированности, составляющие которого — цена простоя оборудования и штраф за задержку выполнения запроса.

Выбор вида кластера для определенного класса решаемых задач осуществляется в результате сравнения получаемых при моделировании характеристик.

Модели, соответствующие структурам, изображенным на рис. 1, 3 и 4, *б*, обозначим *1*, *2*, *3*. При решении класса задач с равномерным обращением к серверам и равномерным обращением серверов к дискам для модели *1* требуется меньшее время отклика, чем для модели *2* при любой средней длительности обслуживания заявки (рис. 6, *а*). При решении задач с неравномерным обращением к серверам время отклика для модели *3* меньше при средней длительности обслуживания заявки меньшей 0,04 с. Если средняя длительность обслуживания заявки больше, то время отклика меньше для модели *2* (рис. 6, *б*).

Вероятностные модели используются для решения задач синтеза вычислительных структур.

Синтез кластерных систем. Для оптимизации состава и структуры вычислительных систем используем методы [10], позволяющие определить структуру вычислительной среды минимальной стоимости при заданном времени отклика или, наоборот, с минимальным временем отклика при заданной стоимости.

Пусть кластер состоит из M рабочих станций пользователей и N групп серверов ($k_1, \ldots, k_N$). В зависимости от вида ресурса сервера определяется тип сервера [11]. Предположим, что все рабочие станции заняты пользователями, каждый из которых может послать только один запрос на один из серверов.

Функционирование рассматриваемой системы представим в виде замкнутой стохастической сети, содержащей N СМО, в которой циркулирует M заявок.

Ведем класс задач, обслуживаемых в вычислительной среде: p_{ij} — вероятности поступлений заявок из i-й СМО в j-ю; p_{10} — вероятность завершения задачи.

По графу передач [10] определяем соотношения интенсивностей потоков заявок λ_i, поступающих в каждую из систем.

Задача оптимизации состава и структуры высокопроизводительных вычислительных систем. Рассмотрим решение следующей задачи: определить быстродействие рабочих станций V_1 и серверов V_2, …, V_N, обеспечивающих минимальное время решения задачи U таким образом, чтобы стоимость s системы с M рабочими станциями не превышала заданного значения. Следовательно, необходимо найти минимальное значение функции U, удовлетворяя ограничениям

$$c_1 M V_1 + c_2 V_2 k_2 + \ldots + c_N V_N k_N \le s,\ V_i > 0,\ i = 1, \ldots, N, \tag{1}$$

где c_i — стоимость единицы производительности i-го устройства; k_i — число серверов i-го вида.

Для решения рассматриваемой задачи применим метод множителей Лагранжа, т. е. найдем минимум функции

$$G = U + \gamma\,(c_1 M V_1 + c_2 V_2 k_2 + \ldots + C_N V_N k_N - s),$$

где γ — неопределенный постоянный множитель. В этом случае V_k ($k = 1, \ldots$ ..., N) и γ определяются как решение системы нелинейных уравнений

$$\frac{\partial G}{\partial V_1} = 0, \ldots, \frac{\partial G}{\partial V_k} = 0, \ldots, \frac{\partial G}{\partial V_N} = 0,$$

$$c_1 M V_1 + c_2 V_2 k_2 + \ldots + C_N V_N k_N \le s. \tag{2}$$

ISSN 0204–3572. Electronic Modeling. 2007. V. 29. № 2

Общее время решения задачи U вычисляется как сумма произведений средних времен пребывания в i-й СМО на коэффициент посещений: $u_i = (m_i \nu_i)/\rho_i$.

Выражение для функции U достаточно громоздкое. Поэтому необходимые для получения системы (2) преобразования выполняются с помощью системы Mathematica. Программа формирования и решения уравнений (2) подробно описана в [9].

Полученная таким образом система нелинейных уравнений (2) на современных ЭВМ решается при ограниченном значении M ($M < 10$), так как при вычислении значения U используются стационарные вероятности (число которых увеличивается с увеличением числа задач). В методе средних [1, 12] используются рекуррентные формулы для определения времени отклика вычислительной системы, что позволяет упростить решение поставленной задачи. Аналогично решается задача синтеза кластерных систем минимальной стоимости при ограниченном времени отклика [10].

Метод оптимизации состава и структуры высокопроизводительных вычислительных систем и оценка его трудоемкости. Алгоритм оптимизации состава и структуры высокопроизводительных вычислительных систем следующий.

1. Выбираем исходный вектор производительности устройств $\boldsymbol{V}_k$ ($k = 1, \ldots, N$) из условия (1). Задаем начальные значения величины шага h, времени отклика $U_{\text{opt}} = \infty$, признак определения лучшей точки $f_{\text{lag}} = 0$ ($f_{\text{lag}} = 1$, если определена лучшая точка).

2. Методом средних считаем основные характеристики: средние времена пребывания в i-й СМО, время отклика задачи $U(M)$, среднее число задач, находящихся в i-й СМО.

3. Если $U(M) < U_{\text{opt}}$, запоминаем новый вектор $\boldsymbol{V}_k$ ($k = 1, \ldots, N$), новый рекорд по времени отклика $U(M) = U_{\text{opt}}$ и $f_{\text{lag}} = 1$ (определена лучшая точка на данном шаге).

4. Если не найдена лучшая точка с шагом h ($f_{\text{lag}} = 0$), уменьшаем его: $h = h/2$.

5. Если $h > \varepsilon$, переходим к п. 6, иначе — конец алгоритма.

6. Строим новый вектор $\boldsymbol{V}_k (k = 1, \ldots, N) = \boldsymbol{V}_k (k = 1, \ldots, N) + h$ при условии (1). Переходим к п. 2.

Начальным приближением для этого метода является точка, полученная градиентным методом [6]. На границе с единичным шагом выбирается направление координаты, по которой функция улучшается и вычисляется новая точка. Если не определена лучшая точка, шаг уменьшается. Процесс продолжается до достижения заданной точности.

Для вычисления значения U необходимо, используя теорему Джексона, рассчитать стационарные вероятности и среднее время пребывания в каждой СМО, загрузку и среднее число решаемых ею задач. Для вычисления нормировочной константы следует выполнить $CNMC_{M+N+1}^{N-1}$ операций сложения и умножения (где C — константа), в том числе MC операций произведений сомножителей $R_j(m_j)\alpha_j v_j)^{m_j}$ (так как $\sum_{j=1}^{N} m_j = M$) и N операций сложения этих произведений, а для вычисления стационарных вероятностей — C_{M+N+1}^{N-1} операций деления. Чтобы вычислить необходимые для расчета времени отклика основные характеристики требуется MN операций сложения. Сложность этого алгоритма — комбинаторная.

Одна итерация с использованием теоремы о среднем требует LNM операций сложения и умножения (где L = const) и еще столько же операций для уточнения на границе. Сложность одной итерации этого алгоритма — полиномиальная.

Алгоритм с использованием теоремы Джексона имеет комбинаторный порядок, а алгоритмы с использованием теоремы о среднем [10] — полиномиальный, что позволяет решать задачи, которые вообще не решаются аналитическим методом на современных ЭВМ в течение реального времени.

Выводы. Предложенные методики анализа эффективности вычислительных систем позволяют получить характеристики их функционирования при решении ими различных классов задач. Рассмотренные способы оптимизации состава и структур высокопроизводительных вычислительных систем можно использовать при их проектировании.

Таким образом, использование вероятностных моделей при проектировании, эксплуатации и оптимизации вычислительных систем позволяет вырабатывать рекомендации по рациональному использованию ресурсов этой вычислительной среды.

Modified methods are proposed for the analysis and synthesis of multiprocessor computational resources of various topology by means of probabilistic models. They allow to analyze and design more wide class of parallel computational media.

1. *Авен О. И. и др.* Оценка качества и оптимизация вычислительных систем. — М. : Наука, 1982. — 464 с.
2. *Cremonesi P., Gennaro C.* Integrated Performance Models for SPMD Applications and MIMD Architectures //IEEE Transactions on Parallel and Distributed Systems. — 2002. — **13,** № 7, — P. 745—757.
3. *Varki E.* Response Time Analysis of Parallel Computer and Storage Systems //Ibid. — 2001. — **12**, № 11. — P. 1146—1161.
4. *Клейнрок Л.* Вычислительные системы с очередями. — М. :Мир, 1979. — 600 с.

5. *Фельдман Л. П., Михайлова Т. В.* Параллельный алгоритм построения дискретной марковской модели /Высокопроизводительные параллельные вычисления на кластерных системах: Материалы четвертого Междунар. науч.-практич. семинара и Всероссийской молодежной школы /Под ред. чл.-кор. РАН В. А. Сойфера. — Самара, 2004. — С. 249—255.
6. *Спортак М., Франк Ч., Паппас Ч. и др.* Высокопроизводительные сети. Энциклопедия пользователя. — Киев : ДиаСофт, 1998. — 432 с.
7. *Михайлова Т. В.* Анализ оценки эффективности кластерных систем с использованием вероятностных моделей //Системний аналіз та інформаційні технології: Тез. доп. Міжнар. науково-практичної конференції студентів, аспирантів та молодих вчених, 1—3 липня 2003р. —Київ : Изд. Киевского политехнического ун-та, 2003. — С. 83—85.
8. *Основы* теории вычислительных систем/С. А. Майоров, Г. И. Новиков, Т. И. Алиев и др. — М. : Высшая шк., 1978. — 408 с.
9. *Фельдман Л. П., Михайлова Т. В.* Оценка эффективности кластерных систем с использованием моделей Маркова //Изв. ТРТУ. Тем. вып.: Материалы Всерос. науч.-техн. конф. с международным участием «Компьютерные технологии в инженерной и управленческой деятельности». — Таганрог: ТРТУ, 2002. — № 2 (**25**). — С. 50—53.
10. Ф*ельдман Л. П., Михайлова Т. В.* Способы оптимизации состава и структуры высокопроизводительных вычислительных систем //Науч. тр. Донецкого государственного технического университета. Серия «Информатика, кибернетика и вычислительная техника»(ИКВТ-2001). — Донецк : ДонГТУ. — 2000. — С. 80—85.
11. *Шнитман В.* Современные высокопроизводительные компьютеры// Информационно-аналитические материалы центра информационных технологий. — 1996: http://hardware/app_kis
12. *Фельдман Л. П., Дедищев В. А.* Математическое обеспечение САПР. Моделирование вычислительных и управляющих систем. — Киев : УМК ВО, 1992. — 256 с.

Поступила 04.10.06

ФЕЛЬДМАН Лев Петрович, д-р техн. наук, профессор кафедры прикладной математики Донецкого национального технического университета. В 1951 г. окончил Московский госуниверситет. Область научных исследований — параллельные вычисления и моделирование вычислительных систем.

МИХАЙЛОВА Татьяна Васильевна, ст. преподаватель кафедры прикладной математики Донецкого национального технического университета. В 1984 г. окончила Донецкий политехнический ин-т. Область научных исследований — моделирование вычислительных систем.