

6. Zhilin L. E., Gorbachev A. N., Brumshteyn Yu. M., Vaskovskiy Ye. Yu. Analiz nomenklatury programmnykh sredstv massovogo ispolzovaniya, primenyaemykh v rossiyskikh vuzakh (na primere Astrakhanskogo gosudarstvennogo universiteta) [Analysis software for mass use of the nomenclature used in Russian universities (for example, the Astrakhan State University)]. *Prikaspiyskiy zhurnal: upravleniye i vysokiye tekhnologii* [Caspian Journal: Control and High Technologies], 2015, no. 2, pp. 42–51.
7. Zaripova V. M., Petrova I. Yu., Tsyrlunikov Ye. S. Klassifikatsiya avtomatizirovannykh sistem podderzhki innovatsionnykh protsessov na predpriyatii (Computer aided innovation – CAI) [Classification of automated systems to support innovative processes at the enterprise (Computer aided innovation – CAI)]. *Prikaspiyskiy zhurnal: upravleniye i vysokiye tekhnologii* [Caspian Journal: Control and High Technologies], 2012, no. 1, pp. 32–40.
8. Kazakov Yu. B., Shelikalov Yu. Ya. Issledovanie magnitnogo polya v vozdushnom zareze startera ST230B [Investigation of the magnetic field in the air gap starter ST230B]. *Tezisy doklada nauchno-tehnicheskoy konferentsii* [Theses of Report Scientific and Technical Conference], Ivanov, 2008, 129 p.
9. Karanchuk V. P. *Osnovy primeneniya EVM* [Fundamentals of computer application], Moscow, Radio i Kommunikatsii, 2008. 288 p.
10. Karpasyuk V. K., Smirnov A. M., Badelin A. G. Osobennosti postroyeniya datchikov magnitnogo polya na osnove efekta kolossalnogo magnitosoprotivleniya [Features of construction of the magnetic field sensors based on the colossal magnetoresistance effect]. *Prikaspiyskiy zhurnal: upravleniye i vysokiye tekhnologii* [Caspian Journal: Control and High Technologies], 2015, no. 4, pp. 291–297.
11. Morozov Yu. V. *Novaya tekhnologiya dlya izgotovleniya magnitoprovodov. Komponenty i tekhnologii* [New technology for the production of magnetic cores. Components and Technology], Omsk, Omsk State Technical University Publ. House, 2006. 95 p.
12. Nasedkin A. V. *Konechno-elementnoe modelirovanie na osnove ANSYS. Programmy resheniya staticheskikh zadach soprotivleniya materialov s variantami individualnykh zadaniy* [Finite-element simulation based on ANSYS. Program solutions of static problems of strength of materials with options for individual tasks], Rostov-on-Don, Rostov State University Publ. House, 2008. 44 p.
13. Chigarev A. V., Kravchuk A. S., Smalyuk A. F. *ANSYS dlya inzhenerov* [ANSYS for engineer], Moscow, Mashinostroyeniye-1, 2004. 512 p.
14. Aykeeva A. A., Zhautikov B. A., Zhautikov F. B., Mukhtarova P. A. The research loads on the skip of mine and quarry electromagnetic lifting installation. *Eurasian Physical Technical Journal*, 2015, no. 1 (23), pp. 59–64.
15. Aykeeva A. A., Zhautikov B. A., Rogovaya K. S., Zhautikov F. B., Mukhtarova P. A. 3-D modeling of elements of skip-electromagnet system. *Eurasian Physical Technical Journal*, 2015, vol. 12, no. 2 (24), pp. 28–32.
16. Bleuler H. A survey of magnetic levitation and magnetic bearing types. *International Journal of Japan Society of Mechanical Engineers*, 2012, no. 32 (5), pp. 335–342.
17. Demerdash N. A., Nehl T. W. An Evaluation of the Methods of Finite Elements and finite differences in the Solution of Nonlinear Electromagnetic Fields in Electrical Machines. *IEEE Transactions on Power Apparatus and Systems*, 2009, vol. 98 (1), pp. 74–87.
18. Hajjaji E., Ouladsine M. Modeling and nonlinear control of magnetic levitation systems. *IEEE Transactions on Industrial Electronics*, 2011, no. 48 (4), pp. 831–838.
19. Hyung-Suk Han, Dong-Sung Kim. *Magnetic Levitation. Maglev Technology and Applications*. Springer Science & Business Media Dordrecht, 2016, pp. 184–187.
20. Munaro C. J., et al. Modeling and observer-based nonlinear control of a magnetic levitation system. *Control Applications, 2002. Proceedings of the 2002 International Conference*, 2002, pp. 121–124.

УДК 004.942, 519.213, 519.872.4

СОКРАЩЕНИЕ ВРЕМЕНИ АППРОКСИМАЦИИ ЛОГОВ ВЫЧИСЛИТЕЛЬНОГО КЛАСТЕРА С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МОМЕНТОВ НА ГИПЕРЭКСПОНЕНЦИАЛЬНОМ РАСПРЕДЕЛЕНИИ

Статья поступила в редакцию 24.02.2017, в окончательном варианте – 22.03.2017.

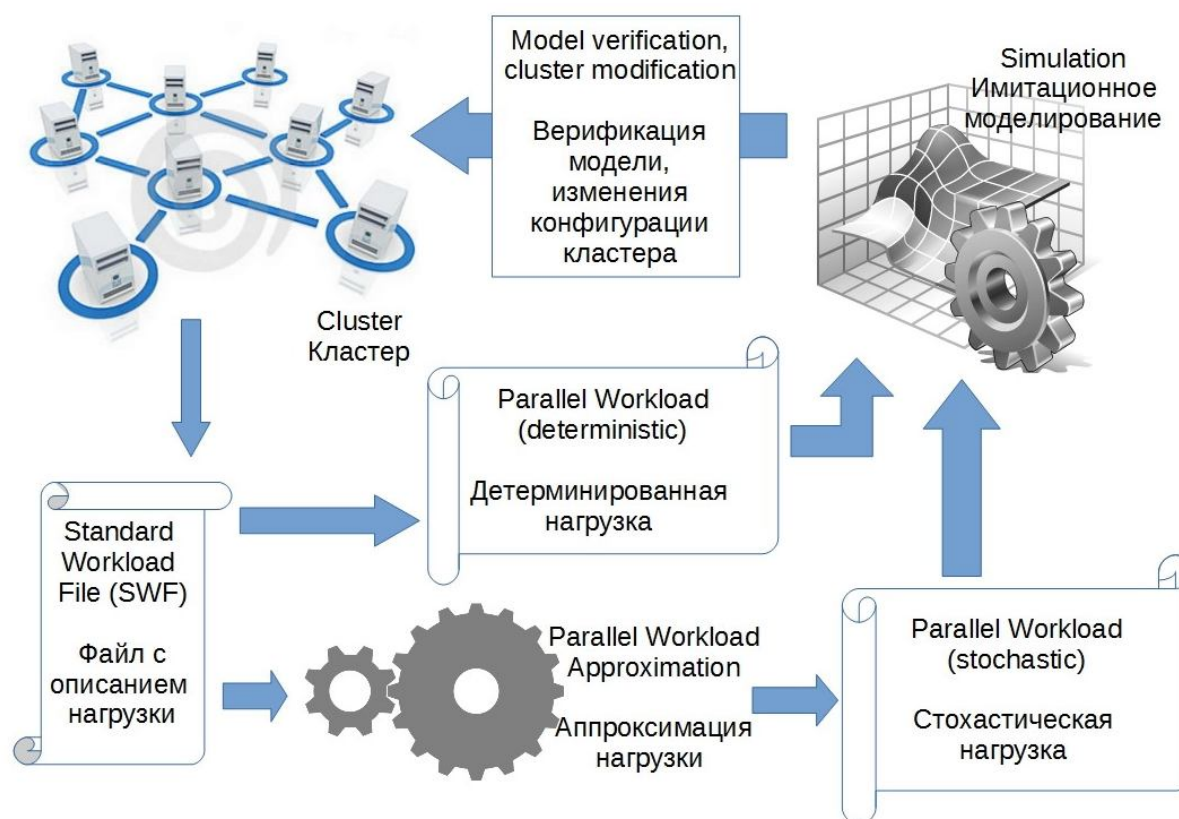
Гаевой Сергей Владимирович, кандидат технических наук, Волгоградский государственный технический университет, 400005, Российская Федерация, г. Волгоград, пр. Ленина, 28, e-mail: gaevserge@mail.ru
Ахмед Весам Мохаммед Абдо, аспирант, Волгоградский государственный технический университет, 400005, Российская Федерация, г. Волгоград, пр. Ленина, 28, e-mail: wesamalsofi@gmail.com
Быков Дмитрий Владимирович, кандидат технических наук, доцент, Волгоградский государственный технический университет, 400005, Российская Федерация, г. Волгоград, пр. Ленина, 28, e-mail: mitril@list.ru
Фоменков Сергей Алексеевич, доктор технических наук, профессор, Волгоградский государственный технический университет, 400005, Российская Федерация, г. Волгоград, пр. Ленина, 28, e-mail: saf@vstu.ru

Одним из типов вычислительных систем являются вычислительные кластеры. Они используются для обслуживания входящих заданий. Важным способом анализа нагрузки вычислительных кластеров является моделирование их работы с использованием модели входящей нагрузки. В работе в рамках такого моделирования предлагается использовать метод моментов с целью определения параметров гиперэкспоненциального распределения с двумя ветками

для модели входящей нагрузки. Это позволяет заметно сократить время аппроксимации модели входящей нагрузки по сравнению с методом наибольшего правдоподобия, но снижает качество результатов. Для проверки этого качества в данной статье используется метод имитационного моделирования аппроксимированной нагрузки и сравнение результатов вычислительных экспериментов с оригинальной (фактической) нагрузкой, взятой из лога вычислительной системы. Результаты для ранее предложенных моделей входящей нагрузки сравниваются с теми расчетами, которые приводятся в данной работе. Обоснована целесообразность использования выбора между двумя методиками для решения задач аппроксимации.

Ключевые слова: метод моментов, интегральная функция распределения, нагрузки вычислительных систем, не масштабируемые задачи, время выполнения заданий, имитационное моделирование, стохастическая аппроксимация

Графическая аннотация (Graphic annotation)



REDUCING THE APPROXIMATION TIME OF CLUSTER WORKLOAD BY USING METHOD OF MOMENTS ON HYPEREXPONENTIAL DISTRIBUTION

The article has been received by editorial board 21.06.2016, in the final version – 17.01.2017.

Gaevoy Sergey V., Ph.D. (Engineering), Volgograd State Technical University, 28 Lenina Ave., Volgograd, 400005, Russian Federation, e-mail: gaevserge@mail.ru

Ahmed Wesam M. A., postgraduate student, Volgograd State Technical University, 28 Lenin Ave., Volgograd, 400005, Russian Federation, e-mail: wesamalsofi@gmail.com

Bykov Dmitriy V., Ph.D. (Engineering), Associate Professor, Volgograd State Technical University, 28 Lenin Ave., Volgograd, 400005, Russian Federation, e-mail: mitril@list.ru

Fomenkov Sergey A., D.Sc. (Engineering), Professor, Volgograd State Technical University, 28 Lenin Ave., Volgograd, 400005, Russian Federation, e-mail: saf@vstu.ru

Computing clusters are one of the computing systems. They are used to execute incoming jobs. An important method to analyze parallel workloads is modeling execution of those systems by using parallel workload models. In this paper it is proposed to use method of moments to compute parameters of Hyperexponential distribution with two branches and get a parallel workload model. This allows us to drastically reduce the approximation time of the parallel workload model in comparison to maximum likelihood method, but it reduces the quality too. To validate the result quality we use the simulation of this approximation and compare the results with the original workload (from the log) in this paper. The results of the formerly proposed parallel workload models are compared with the results from this paper. The reasonability to select an appropriate approximation method for solving approximation tasks is justified.

Keywords: method of moments, cumulative distribution function, parallel workloads, rigid jobs, job length, simulation, stochastic approximation

Введение. Проблема рационального выполнения параллельных и высокопроизводительных вычислений сейчас достаточно актуальна [8]. В частности, стоит вопрос оптимального балансирования нагрузки и подбора оптимальной производительности [7] вычислительных кластеров (ВК). Одним из возможных путей решения этой задачи является моделирование работы ВК, в том числе и имитационное [11]. Последнее требует построения математической модели ВК и поступающей на нее (т.е. входной) нагрузки [16].

В данной работе мы рассмотрим один из аспектов моделирования работы ВК – аппроксимацию конкретных непрерывных случайных величин в процессе создания модели входной нагрузки (МВН).

Целью данной работы является решение задачи сокращения времени аппроксимации непрерывной случайной величины гипер-распределением за счет перехода от метода моментов (ММ) к методу наибольшего правдоподобия (МНП).

Общая характеристика проблематики работы. Ранее нами уже была построена модель [2, 4] рассматриваемого далее ВК и предложены способы моделирования для нее входящей нагрузки [1, 3].

Реальный ВК построен из вычислительных машин, которые обслуживают поступающие задания [17]. В данной работе используется представление такого ВК в виде обслуживающего блока с единой не приоритетной, не ограниченной по размеру очередью. Это обеспечивает обслуживание всех входящих заданий.

Каждое задание может исполняться параллельно на нескольких машинах (каналах обслуживания). Количество вычислительных машин, на которых исполняется задание, называется его шириной. Будем считать, что ширина задания определяется в момент его создания, что является довольно частым допущением [10, 13–16].

Введем следующие определения [6]. Длиной задания назовем время его выполнения. Площадью задания назовем произведение длины на ширину. Очевидно, что площадь – это сложность задания. Она же представляет собой суммарное машинное время обслуживания. Отметим, что авторы существующих публикаций по теме данной статьи используют и иную терминологию.

При завершении исполнения очередного задания очередь просматривается вся от начала и до конца для выбора (извлечения) заданий на выполнение. При этом извлекаемых из нее заданий может быть несколько – ведь ушедшее из очереди (т.е. исполненное) задание может освободить не только один, но и несколько каналов. Задание, для выполнения которого имеется достаточное количество свободных машин, ставится на исполнение. Затем просмотр очереди продолжается со следующего за ним задания.

Пример распределения заданий по узлам ВК дан на рисунке 1.

Узел 0 (задание 8, часть 1 из 5)	Узел 1 (свободен)	Узел 2 (задание 8, часть 3 из 5)
Узел 3 (задание 5, часть 1 из 1)	Узел 4 (задание 11, часть 1 из 2)	Узел 5 (задание 13, часть 4 из 8)
Узел 6 (свободен)	Узел 7 (задание 8, часть 2 из 5)	Узел 8 (свободен)
Узел 9 (задание 13, часть 7 из 8)	Узел 10 (задание 13, часть 1 из 8)	Узел 11 (задание 13, часть 5 из 8)
Узел 12 (задание 8, часть 4 из 5)	Узел 13 (свободен)	Узел 14 (задание 11, часть 2 из 2)
Узел 15 (задание 13, часть 3 из 8)	Узел 16 (задание 13, часть 2 из 8)	Узел 17 (задание 8, часть 5 из 5)
Узел 18 (задание 13, часть 6 из 8)	Узел 19 (свободен)	Узел 20 (задание 13, часть 8 из 8)

Рисунок 1 – Пример распределения заданий по узлам ВК

На нем представлено параллельное исполнение заданий с ширинами 1, 2, 5 и 8. При этом каждое из заданий может занимать произвольные узлы. Даже при наличии заданий в очереди, часть узлов может простаивать. В нашем примере, в очереди могут быть задания шириной шесть и более, которым для момента, показанного на рисунке, нет места на исполнение, т.е. нет достаточного количества узлов ВК. Более узкое задание, которому такого места хватает, должно быть извлечено из очереди, несмотря на то, что оно пришло позднее стоящих впереди него. Поскольку порядок обслуживания заданий может не соответствовать порядку их прихода (поступления в очередь), то номера заданий на рисунке идут не подряд.

Длиной очереди назовем количество, находящихся в ней заданий. Шириной (сложностью, площадью) очереди — сумму ширин (сложностей, площадей) входящих в нее заданий. Аналогично определим понятия длины, ширины и площади всей вычислительной системы, т.е. ВК.

Для генерации случайных нагрузок были предложены [3] различные МВН. В основе их лежат идеи разделения входного потока на несколько потоков по ширине заданий и/или искажения временной шкалы для моделирования нестационарной интенсивности приходящих заданий.

При использовании любой из этих моделей возникает необходимость аппроксимации следующих видов случайных величин [10, 13, 15, 16].

(1) Интервалов времени между моментами прихода заданий. (2) Ширин заданий. (3) Длин или площадей заданий.

Ширина является дискретной случайной величиной и представлена конечным числом значений. Поэтому ее можно аппроксимировать просто в виде массива вероятностей для набора ширин.

Для аппроксимации интервалов времени между моментами приходами заданий, длины или площади необходимо использовать непрерывную случайную величину. Эту аппроксимацию можно осуществить как с помощью ММ, так и МНП.

Согласно [10, 13–16] и [1, 3] аппроксимируемые параметры являются случайными величинами с большим коэффициентом вариации, поэтому гипер-распределения подходят для целей аппроксимации. Предыдущие работы [1, 3] показали, что одни из самых точных аппроксимаций дает МНП для гиперэкспоненциального и гипергамма-распределения. Наши аппроксимации МНП требуют значительных вычислительных ресурсов, при этом ММ для этих целей не применялся. В работе [13] показано, что возможно использование ММ на гиперэрланговом распределении, но для этого пришлось искусственным путем сократить число определяемых параметров.

В своей работе мы проводим моделирование нагрузок ВК систем (генерацию) и моделирование обслуживания этих нагрузок. В качестве исходного материала мы используем реальные нагрузки ВК, взятые из логов. Для проверки качества аппроксимаций мы сравниваем результаты моделирования сгенерированной нагрузки с результатами моделирования исходной нагрузки (детерминированная имитационная модель [2, 4]).

Введем следующие обозначения: $E(X)$ – математическое ожидание величины X , $VAR(X)$ – ее дисперсия, $stDev(X)$ – среднее квадратичное отклонение, $cov(X) = \frac{stDev(X)}{E(X)}$ – коэффициент вариации.

Если мы имеем дело с оценкой момента, то будем писать над ней горизонтальную черту.

Оценки этих параметров можно дать по формулам:

$$\overline{E(X^k)} = \frac{1}{N} \sum_{i=1}^N X_i^k, \quad (1)$$

$$\overline{VAR(X)} = \frac{1}{N-1} \sum_{i=1}^N (X_i - \overline{E(X)})^2 = \frac{N}{N-1} (\overline{E(X^2)} - \overline{E(X)}^2), \quad (2)$$

где N – число наблюдений, X_i – конкретное i -ое наблюдение.

Нам ТАКЖЕ потребуются $pdf(x)$ и $cdf(x)$ – дифференциальная и интегральная функции распределения.

Обозначим через $H(n)$ гиперэкспоненциальное распределение:

$$pdf(x) = \sum_{i=1}^n \alpha_i \lambda_i e^{-\lambda_i x}, \quad (3)$$

$$cdf(x) = 1 - \sum_{i=1}^n \alpha_i e^{-\lambda_i x}, \quad (4)$$

где n – количество веток распределения (задается перед аппроксимацией как один из параметров распределения), α_i – вероятности использования веток ($1 \geq \alpha_i \geq 0$, $\sum_{i=1}^n \alpha_i = 1$), λ_i – интенсивность составляющего его экспоненциального распределения на ветке № i ($\lambda_i \geq 0$). Стоит также отметить, что $cov(X) \geq 1$.

Начальные моменты имеют вид:

$$E(X^k) = k! \sum_{i=1}^n \frac{\alpha_i}{\lambda_i^k}. \quad (5)$$

В частности, математическое ожидание есть:

$$E(X) = \sum_{i=1}^n \frac{\alpha_i}{\lambda_i}, \quad (6)$$

Из центральных моментов нам потребуется только дисперсия:

$$VAR(X) = E^2(X) + \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j (\lambda_i^{-1} - \lambda_j^{-1})^2. \quad (7)$$

Из условия $\sum_{i=1}^n \alpha_i = 1$ следует, что одна из α_i задается по остаточному принципу. Поэтому число параметров этого распределения равно $2n - 1$.

В [1, 3] использованы такие подходы к аппроксимации как ММ и МНП. Однако для гиперэкспоненциального распределения авторами [1, 3] был применен только МНП, который требует использования больших вычислительных ресурсов.

Использовать ММ можно лишь для распределений, у которых не более четырех параметров, так как на практике определение моментов выше четвертого порядка затруднено из-за большой доли случайности.

Число параметров гиперэкспоненциального распределения равняется $2n - 1$, поэтому для распределения с двумя ветками ММ все-таки применим. Число параметров гипегамма-распределения равняется $3n - 1$, поэтому даже для распределения с двумя ветками ММ не применим. Рассматривать вариант с одной веткой нет смысла, так как это обычно будет экспоненциальное или гамма-распределение, которые уже рассмотрены в [1, 3].

Конкретная цель – необходимо получить аппроксимацию случайной величины с большим коэффициентом вариации ММ для двухветочного гиперэкспоненциального распределения.

Теоретическое рассмотрение вопросов. В данной работе используется ММ для аппроксимации двухветочным гиперэкспоненциальным распределением. Для этого введем упрощенные обозначения:

$$m_i = \lambda_i^{-1}, \alpha_1 = \alpha, \alpha_2 = 1 - \alpha, cov(X) = v, \mu_i = \frac{m_i}{E(X)}.$$

Преобразуем выражение для математического ожидания на основе формулы (6) из

$$\alpha_1 m_1 + \alpha_2 m_2 = E(X) \quad (8)$$

в

$$\alpha \mu_1 + (1 - \alpha) \mu_2 = 1. \quad (9)$$

Преобразуем выражение для дисперсии (7) из

$$VAR(X) = E^2(X) + 2\alpha_1 \alpha_2 (\lambda_1^{-1} - \lambda_2^{-1})^2 \quad (10)$$

в

$$\frac{v^2 - 1}{2} = \alpha(1 - \alpha)(\mu_1 - \mu_2)^2. \quad (11)$$

Обозначим:

$$\beta = \sqrt{\frac{v^2 - 1}{2}}. \quad (12)$$

Под корнем гарантировано находится неотрицательное число, так как $v = cov(X) \geq 1$.

В (9, 11) у нас три неизвестных – μ_1 , μ_2 и α . Выразим из этих соотношений μ_1 и μ_2 . В силу того, что уравнение (11) является квадратным, получаем два возможных решения:

$$\mu_1 = 1 - \sqrt{\frac{1-\alpha}{\alpha}}\beta \quad (13)$$

$$\mu_2 = 1 + \sqrt{\frac{\alpha}{1-\alpha}}\beta \quad (14)$$

и

$$\mu_1 = 1 + \sqrt{\frac{1-\alpha}{\alpha}}\beta \quad (15)$$

$$\mu_2 = 1 - \sqrt{\frac{\alpha}{1-\alpha}}\beta. \quad (16)$$

Очевидно, что при замене α на $1-\alpha$ ветви просто меняются местами. Поэтому достаточно рассмотреть (проанализировать) только одно решение. Остановимся на первом из них, т.е. на формулах (13) и (14).

Значения μ_1 и μ_2 не могут быть отрицательными. При этом μ_2 гарантировано положительное. Для μ_1 с целью обеспечения его не отрицательности необходимо выполнение условия:

$$\sqrt{\frac{1-\alpha}{\alpha}}\beta \leq 1. \quad (17)$$

Отсюда

$$\alpha \geq \frac{\beta^2}{1+\beta^2} \quad (18)$$

Теперь нам надо получить третье уравнение. Возьмем третий начальный момент (5)

$$\frac{E(X^3)}{3!} = \alpha_1 m_1^3 + \alpha_2 m_2^3 \quad (19)$$

и преобразуем его в форму:

$$\frac{E(X^3)}{6E^3(X)} = \alpha\mu_1^3 + (1-\alpha)\mu_2^3. \quad (20)$$

Подставив значения μ_1 и μ_2 и раскрыв скобки, получим:

$$\frac{E(X^3)}{6E^3(X)} = 1 + 3\beta^2 + \frac{2\alpha-1}{\sqrt{\alpha(1-\alpha)}}\beta^3. \quad (21)$$

Преобразуем его в форму:

$$\frac{2\alpha-1}{\sqrt{\alpha(1-\alpha)}} = \gamma, \quad (22)$$

где

$$\gamma = \frac{E(X^3)}{6E^3(X)\beta^3} - \frac{1+3\beta^2}{\beta^3}. \quad (23)$$

Для решения уравнения (22) необходимо возвести обе части в квадрат. В результате мы получим два решения при различных γ :

$$\alpha = \frac{1}{2} \left(1 \pm \frac{|\gamma|}{\sqrt{\gamma^2 + 4}} \right) \quad (24)$$

Однако одно из этих решений лишнее, так как то же самое получится и при решении уравнения $\frac{2\alpha-1}{\sqrt{\alpha(1-\alpha)}} = -\gamma$.

Знаки $2\alpha-1$ и γ совпадают, поэтому имеем:

$$\alpha = \frac{1}{2} \left(1 + \frac{\gamma}{\sqrt{\gamma^2 + 4}} \right). \quad (25)$$

Может сложиться впечатление, что решение уравнения (22) есть при любом значении третьего момента. Но это не так, поскольку α ограничен снизу (18). Поэтому мы вынуждены отойти от равенства третьего момента предлагаемой оценке в тех случаях, когда необходимо обеспечить (18).

Используя минимальное α из (18) и монотонное строгое возрастание $\frac{2\alpha-1}{\sqrt{\alpha(1-\alpha)}}$ из (21), получим, что

$$\frac{E(X^3)}{6E^3(X)} \geq (1+\beta^2)^2. \quad (26)$$

В итоге получаем:

$$E(X) = \overline{E(X)}, \quad VAR(X) = \overline{VAR(X)}, \quad E(X^3) = \max\left(\overline{E(X^3)}, 6E^3(X)(1+\beta^2)^2\right), \quad (27)$$

$$v^2 = \frac{VAR(X)}{E^2(X)}, \quad \beta = \sqrt{\frac{v^2-1}{2}}, \quad \bar{\gamma} = \frac{E(X^3)}{6E^3(X)\beta^3} - \frac{1+3\beta^2}{\beta^3}, \quad (28)$$

$$\alpha = \max\left(\frac{1}{2}\left(1 + \frac{\bar{\gamma}}{\sqrt{\bar{\gamma}^2+4}}\right), \frac{\beta^2}{1+\beta^2}\right), \quad (29)$$

$$\lambda_1 = \left(E(X)\left(1 - \sqrt{\frac{1-\alpha}{\alpha}}\beta\right)\right)^{-1}, \quad \lambda_2 = \left(E(X)\left(1 + \sqrt{\frac{\alpha}{1-\alpha}}\beta\right)\right)^{-1}. \quad (30)$$

Реальное значение третьего момента может отличаться от аппроксимированного, а вот математическое ожидание и дисперсия всегда равны своим оценкам по (27). Итак, искомые параметры гиперэкспоненциального распределения представлены в (29–30). Обозначим такую аппроксимацию гиперэкспоненциального распределения с использованием ММ как H_μ .

Сравнение расчетных и реальных результатов. В [14] предоставлены логи работы реальных вычислительных систем (ВК). Описание потока заданий содержит время прихода задания, его ширину и длину. Для аппроксимации этих стохастических параметров используются различные вышеупомянутые методики [1, 3], а для моделирования обслуживания нагрузок – средства из [2, 4]. В рамках данной работы будет использован лог UniLu-Gaia-2014-2.swf, который принадлежит кластеру UniLu Gaia (The University of Luxemburg Gaia Cluster log [12]) с 2004 каналами обслуживания.

Сравним время работы и качество результата ММ и МНП. В силу универсальности МНП обозначим его аппроксимации так же, как и само распределение $H(n)$. Каждая аппроксимация в нашей программе [2, 4] обычно сопровождается еще и анализом выполнения самой процедуры аппроксимации – поэтому тратится дополнительное время на вычисления (табл. 1). Очевидно, что анализ занимает примерно 2 мин. 30 с. Примеры аппроксимаций эмпирической функции распределения (ЭФР) времени между моментами прихода заданий изображены на рисунке 2.

Таблица 1 – Скорости выполнения различных видов аппроксимации

Анализ	Время выполнения		
	H_μ	$H(2)$	$H(3)$
Не проводится	5 сек.	2 мин. 16 с	10 мин. 31 с
Проводится	2 мин. 34 сек.	4 мин. 45 с	12 мин. 59 с

Для оценки качества аппроксимации было проведено стохастическое имитационное моделирование предложенных моделей. С этой целью было усовершенствовано средство, описанное в [2, 4]. В качестве допустимой погрешности имитационного моделирования была взята величина 5 %. В соответствии с Центральной предельной теоремой [5] это требует более 40 испытаний на проверку каждого из вариантов.

Каждая модель нагрузки состоит из двух частей: модели времен приходов заданий и модели обслуживания (определение длины, ширины и площади). Обсуждение этих моделей, как уже говорилось, тема отдельной статьи [1, 3]. Поэтому мы рассмотрим их лишь кратко.

Начнем с самого простого варианта: A , B , $B^\wedge$, где A – аппроксимация некоторым законом распределения интервала между приходами заданий, B – аппроксимация законом распределения площади (при указании $^\wedge$ – длины) заданий. Ширина моделируется отдельно по вероятности ее появления.

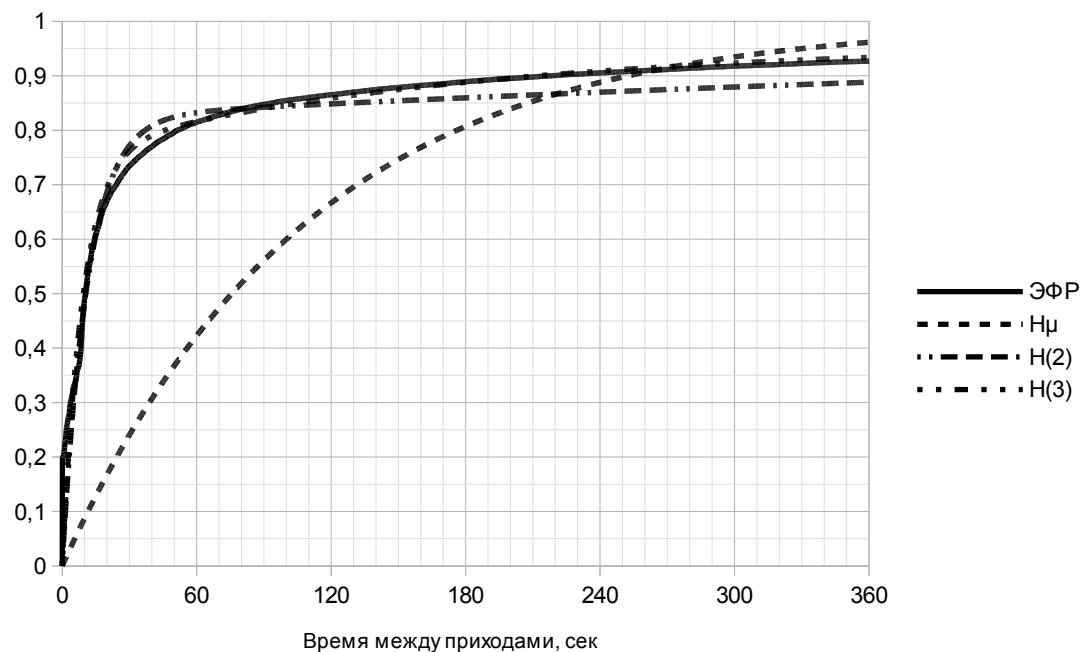


Рисунок 2 – Аппроксимация времен между приходами заданий

Задания различной ширины могут иметь весьма различные характеристики распределения длины/площади. Поэтому имеет смысл выделить отдельные законы для интервалов ширины [13, 16]. В самом простом случае мы выделяем для каждой ширины свое распределение длины/площади задания. Обозначим это значком $\$$ перед обозначением закона распределения площади: $\$B$. В силу того, что в каждом из входных потоков будет только одна ширина, длина будет пропорциональна площади и в отдельной аппроксимации не нуждается.

Второй вариант разделения – выделить в отдельную группу каждую ширину, равную степени двойки. Это имеет смысл, так как согласно [15], в логах доминируют именно задания, ширина которых является степенью двойки – даже тогда, когда к этому нет технических предпосылок. Такие работы, как [10], выделяют еще и другие доминирующие (но слабее, чем степени двойки) ширины заданий, например, кратные десяти. В других работах [9] наоборот пытаются уйти от этой тенденции.

Интервалы ширин между степенями двойки выделим также в отдельные группы: по одной группе на каждый интервал. То есть будут группы 1, 2, 3, 4, 5–7, 8, 9–15, 16, 17–31, 32, 33–63, 64 и т.д. Обозначим такое выделение групп знаком $\&$: $\&B$ и $\&B^{\wedge}$.

Аналогичное разделение можно провести и для входных времен заданий: выделим несколько входных потоков, в каждом из которых приходят задания с ширинами, относящимися к определенному интервалу. Используем аналогичные принципы разбиения и обозначения. Получатся $\$A$, $\&A$.

В [15] предлагается анализировать входной поток как нестационарный. В данной работе будет рассмотрено изменение интенсивности потока заявок в течение недели. За начало недели мы взяли полночь с воскресенья на понедельник. Будем полагать, что интенсивность прихода заданий остается постоянной в течение получаса (как в [15]).

Под приведенной интенсивностью подразумевается:

$$\bar{\lambda}(t) = \lambda(t) / \bar{\lambda}, \quad (31)$$

где $\bar{\lambda}(t)$ – приведенная интенсивность прихода заданий, $\lambda(t)$ – интенсивность, $\bar{\lambda}$ – средняя интенсивность в течение недели.

Для генерации интервала между событиями нестационарного потока искадим временную шкалу. Интервалом времени между реальными точками t_0 и t_1 временной шкалы будем считать значение интеграла

$$L(t, t_0) = \int_{t_0}^t \bar{\lambda}(t) dt. \quad (32)$$

Назовем ее «случайным приведенным временем» между приходами заданий. Такой поток будет стационарным и его можно аппроксимировать обычным способом, а потом вернуться к изначальной шкале времени.

Обозначим такую модель знаком «~» перед обозначением входного потока, например, $\sim A$. Здесь сразу же надо оговорить одну важную особенность. Обозначения $\sim SA$ и $S \sim A$ не одинаковы. В первом случае мы подразумеваем введение единой интенсивности для всех входных потоков, а во втором – что каждый поток получает свою собственную интенсивность.

Таким образом, получаем следующие варианты приходов: A , $\sim A$, SA , $\& A$, $\sim SA$, $\sim \& A$, $S \sim A$, $\& \sim A$ и варианты обслуживания: B , SB , $\& B$, $B^\wedge$, $\& B^\wedge$. Сочетание этих двух моделей дает модель нагрузки. Будем обозначать его (сочетание) через слеш, например, $\& \sim A / \& B^\wedge$. Итого здесь мы имеем 40 комбинаций приходов заданий и параметров их обслуживания. С учетом трех возможных вариантов аппроксимации (H_μ , $H(2)$, $H(3)$) получаем 120 вариантов моделей.

В качестве эталонного результата моделирования возьмем результат детерминированного моделирования исходной нагрузки ВК.

Чтобы выбрать лучший вариант используется критерий отклонения:

$$Dev = \sqrt{\frac{1}{M} \sum_{i=1}^M \left(\frac{\bar{P}_i - P_i}{P_i} \right)^2}, \quad (33)$$

где M – число параметров, P_i – эталонное значение параметра, $\bar{P}_i$ – значение, полученное по стохастической модели.

Пять моделей с самым малым значением этого параметра представлены в таблице 2, а пять моделей, полученных по ММ (также с самым малым значением этого показателя) – в таблице 3.

В ней указаны два средних времени ожидания. Причина - из-за того, что в очередь не попадают все задания, а лишь их определенный процент, то среднее время ожидания можно рассчитывать двумя способами. (1) Для всех заданий, учитывая нулевое время ожидания не попавших в очередь (без штриха). (2) Только для попавших в очередь (со штрихом).

Таблица 2 – Самые хорошие аппроксимации входной нагрузки

	$\sim \& H(3)/H(3)^\wedge$	$S \sim H(2)/SH(2)$	$\sim H(2)/\& H(2)^\wedge$	$\sim \& H(2)/H(2)^\wedge$	$\& \sim H(3)/H(3)^\wedge$	Эталон
Среднее время выполнения заданий, сек	14230	14356	14237	14293	14231	14329
Среднее число выполняемых заданий	95,884	96,679	95,903	96,146	95,86	93,067
Среднее число занятых каналов	961,21	921,77	919,4	962,9	957,52	872,26
Среднее время ожидания задания в очереди, с	59,191	83,9	48,56	47,864	103,143	72,41
Среднее время ожидания задания в очереди', с	1752,4	2823,8	2393,7	2026,4	2395,1	2259,7
Доля заданий попавших в очередь	0,028911	0,023727	0,018048	0,020327	0,037374	0,032044
Средняя длина очереди	0,40144	0,56642	0,3279	0,32306	0,70042	0,4703
Средняя ширина очереди	12,375	17,711	13,829	11,628	16,377	15,31
Среднее время пребывания в системе, сек	14289	14440	14285	14341	14334	14402
Средняя длина системы	96,286	97,245	96,231	96,469	96,561	93,537
Средняя ширина системы	973,59	939,49	933,23	974,53	973,9	887,57
Отклонение от эталона	0,12547	0,14491	0,19336	0,19947	0,20817	0
Место по степени отклонения	1	2	3	4	5	

Таблица 3 – Самые хорошие аппроксимации входной нагрузки по ММ

	$\&\sim H_{\mu}/\$H_{\mu}$	$\sim\&H_{\mu}/H_{\mu}^{\wedge}$	$\&\sim H_{\mu}/H_{\mu}^{\wedge}$	$\&\sim H_{\mu}/\&H_{\mu}^{\wedge}$	$\&\sim H_{\mu}/\&H_{\mu}$	Эталон
Среднее время выполнения заданий, с	14352	14196	14196	14375	14324	14329
Среднее число выполняемых заданий	96,511	96,2	96,105	97,161	96,816	93,067
Среднее число занятых каналов	905,98	966,42	959,97	920,9	910,3	872,26
Среднее время ожидания задания в очереди, с	38,28	35,247	33,043	28,605	25,24	72,41
Среднее время ожидания задания в очереди', с	1873,8	1391,7	1323,5	2327,7	1789,4	2259,7
Доля заданий попавших в очередь	0,016405	0,019913	0,020755	0,007809	0,009683	0,032044
Средняя длина очереди	0,26044	0,2413	0,2262	0,19438	0,17137	0,4703
Средняя ширина очереди	9,395	12,048	11,916	8,023	7,073	15,31
Среднее время пребывания в системе, с	14390	14231	14229	14404	14349	14402
Средняя длина системы	96,771	96,441	96,331	97,355	96,987	93,537
Средняя ширина системы	915,37	978,46	971,88	928,92	917,37	887,57
Отклонение от эталона	0,27683	0,27981	0,29096	0,37155	0,38778	0
Место по степени отклонения	15	16	17	21	22	

Очевидно, что качество результатов ММ ниже, чем у МНП. Однако с учетом разницы во времени аппроксимации, имеет смысл использование ММ. Также приведенные результаты показывают, что нет серьезных оснований использовать трехветочное распределение с МНП.

Заключение. Таким образом, была получена несколько грубая, но очень быстрая аппроксимация, которая дала не самые плохие результаты (см. табл. 1). Время расчета для нее составляет 5 сек. против 126 сек. по сравнению с двухветочным МНП. Это быстрее примерно в 25 раз при отклонении результата от эталона всего в два-три раза больше (см. табл. 2–3) чем у самой лучшей аппроксимации. Также было показано, что использование трехветочного распределения МНП не обеспечивает значительно лучшего результата: отклонение от эталона сокращается на 15 % по сравнению с двухветочным МНП при увеличении времени расчета в 4,6 раза.

Полученная оценка ММ может быть использована при необходимости аппроксимировать данные большого объема, что позволяет упростить моделирование обслуживания в ВК. Отметим, что ММ не применим к гиперэкспоненциальному распределению с большим числом веток.

Список литературы

1. Гаевой С. В. Аппроксимация времени выполнения заданий на примере вычислительного кластера LPC EGEE 2004 / С. В. Гаевой, Ф. А. Х. Аль-Хадша, С. А. Фоменков // Известия Волгоградского государственного технического университета. Серия «Актуальные проблемы управления, вычислительной техники и информатики в технических системах». – 2014. – Вып. 21. – № 12 (139). – С. 135–141.
2. Гаевой С. В. Детерминированная имитационная модель кластеров грид-системы, обслуживающих задания / С. В. Гаевой, Ф. А. Х. Аль-Хадша, В. С. Лукьянов // Вестник компьютерных и информационных технологий. – 2014. – № 6. – С. 39–43.
3. Гаевой С. В. Моделирование работы вычислительного кластера на примере LANL CM5 / С. В. Гаевой, Ф. А. Х. Аль-Хадша // SCI-ARTICLE.RU : электронный периодический научный журнал. – 2013. – № 3 (ноябрь). – С. 304–313. – Режим доступа: http://sci-article.ru/stat.php?i=modelirovanie_raboty_vychislitelnogo_klastera_na_primere_LANL_CM5, свободный. – Заглавие с экрана. – Яз. рус.

4. Детерминированная имитационная модель кластеров грид-системы для сравнения эффективности использования эвристик распределения заданий / С. В. Гаевой, Ф. А. Х. Аль-Хадша, С. А. Фоменков, В. С. Лукьянов // Прикаспийский журнал: управление и высокие технологии. – 2014. – № 2. – С. 148–157.
5. Фоменков С. А. Математическое моделирование системных объектов : учебное пособие / С. А. Фоменков, В. А. Камаев, Ю. А. Орлова. – Волгоград : Волгоградский государственный технический университет, 2014. – 335 с.
6. Эвристики распределения задач для брокера ресурсов Grid / А. И. Аветисян и другие. – 2009. – Режим доступа: <http://www.citforum.ru/nets/digest/grid/index.shtml>, свободный. – Заглавие с экрана. – Яз. рус.
7. An Evaluation of Parallel Job Scheduling for ASCI Blue-Pacific / H. Franke, et al. // Supercomputing. – 1999. – Nov.
8. Clúster de balanceo de carga y alta disponibilidad para servicios web y mail / M. M. Sinisterra, et al. – Режим доступа: <https://dialnet.unirioja.es/descarga/articulo/4364562.pdf>, свободный. – Заглавие с экрана. – Яз. рус.
9. Downey A. B. A Parallel Workload Model and Its Implications for Processor Allocation / A. B. Downey // The Rachel and Selim Benin School of Computer Science and Engineering. – 2013. – Режим доступа: <http://allendowney.com/research/allocation/>, свободный. – Заглавие с экрана. – Яз. рус.
10. The Feitelson 1996 Model // The Rachel and Selim Benin School of Computer Science and Engineering. – 2013. – Режим доступа: http://www.cs.huji.ac.il/labs/parallel/workload/m_feitelson96/, свободный. – Заглавие с экрана. – Яз. рус.
11. GridMe: Grid modeling environment // Google code. – 2014. – Режим доступа: <https://code.google.com/p/gridme/>, свободный. – Заглавие с экрана. – Яз. рус.
12. HPC @ Uni.lu. – 2017. – Режим доступа: <https://hpc.uni.lu/systems/gaia/>, свободный. – Заглавие с экрана. – Яз. рус.
13. The Jann et al 1997 Model // The Rachel and Selim Benin School of Computer Science and Engineering. – 2013. – Режим доступа: http://www.cs.huji.ac.il/labs/parallel/workload/m_jann97/, свободный. – Заглавие с экрана. – Яз. рус.
14. Logs of Real Parallel Workloads from Production Systems // The Rachel and Selim Benin School of Computer Science and Engineering. – 2013. – Режим доступа: <http://www.cs.huji.ac.il/labs/parallel/workload/logs.html>, свободный. – Заглавие с экрана. – Яз. рус.
15. Lublin U. The Workload on Parallel Supercomputers: Modeling the Characteristics of Rigid Jobs / U. Lublin, D. G. Feitelson // The Rachel and Selim Benin School of Computer Science and Engineering. – 2013. – Режим доступа: <http://www.cs.huji.ac.il/~feit/papers/Rigid01TR.pdf>, свободный. – Заглавие с экрана. – Яз. рус.
16. Modeling of Workload in MPPs / J. Jann, et al. // Job Scheduling Strategies for Parallel Processing : Lect. Notes Comput. Sci. / ed. by D. G. Feitelson, L. Rudolph. – Springer-Verlag, 1997. – Vol. 1291. – P. 95–116.
17. Den optimalen Rechnerverbund gibt es nicht einmal auf dem Papier // Computerwoche. – Режим доступа: <http://www.computerwoche.de/a/den-optimalen-rechnerverbund-gibt-es-nicht-einmal-auf-dem-papier,1087149>, свободный. – Заглавие с экрана. – Яз. рус.

References

1. Gaevoy S. V., Al-Khadsha F. A. Kh., Fomenkov S. A. *Approximatsia vremeni vypolneniya zadaniy na primere vychislitel'nogo klustera LPC EGEE 2004* [Approximation of job execution time discovering computing cluster LPC EGEE 2004]. *Izvestiya Volgogradskogo gosudarstvennogo tekhnicheskogo universiteta, seriya «Aktualnye problemy upravleniya, vychislitel'noy tekhniki i informatiki v tekhnicheskikh sistemakh»* [Bulletin of the Volgograd State Technical University, Series "Actual Problems of Management, Computing Hardware and Informatics in Engineering Systems"], 2014, no. 12 (139), pp. 135–141.
2. Gaevoy S. V., Al-Khadsha F. A. Kh., Lukyanov V. S. *Determinirovannaya imitatsionnaya model klasterov grid-sistemy, obsluzhivayushchikh zadaniya* [Deterministic simulation model of clusters of a Grid-system executing jobs]. *Vestnik kompyuternykh i informatsionnykh tekhnologiy* [Bulletin of Computer and Information Technologies], 2014, no. 6, pp. 39–43.
3. Gaevoy S. V., Al-Khadsha F. A. Kh. *Modelirovanie raboty vychislitel'nogo klustera na primere LANL CM5* [Simulation of computing cluster discovering LANL CM5]. *SCI-ARTICLE.RU : elektronnyy periodicheskiy nauchnyy zhurnal* [SCI-ARTICLE.RU : Electronic Periodical Scientific Journal], 2013, no. 3, pp. 304–313. Available at : http://sci-article.ru/stat.php?i=modelirovanie_raboty_vychislitel'nogo_klustera_na_primere_LANL_CM5.
4. Gaevoy S. V., Al-Khadsha F. A. Kh., Fomenkov S. A., Lukyanov V. S. *Determinirovannaya imitatsionnaya model klasstrov grid-sistemy dlya sravneniya effektivnosti ispolsovaniya evristik raspredeleniya zadaniy* [Deterministic simulation model of clusters of a Grid-system for comparison of heuristics for task distribution]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian Journal: Control and High Technologies], 2014, no. 2, pp. 148–157.
5. Fomenkov S. A., Kamaev V. A., Orlova Yu. A. *Matematicheskoe modelirovanie sistemnykh obektov* [Mathematical modeling of system objects], Volgograd, Volgograd State Technical University Publ. House, 2014. 335 p.
6. Avetisyan A. I., et al. *Evrystiki raspredeleniya zadach dlya brokera resursov Grid* [Heuristics of job distribution for Grid resource broker]. Available at: <http://www.citforum.ru/nets/digest/grid/index.shtml>.
7. Franke H., et al. An Evaluation of Parallel Job Scheduling for ASCI Blue-Pacific. *Supercomputing*, 1999, Nov.
8. Sinisterra M. M., et al. *Clúster de balanceo de carga y alta disponibilidad para servicios web y mail* [Cluster of load balancing and high availability for web and mail services]. Available at: <https://dialnet.unirioja.es/descarga/articulo/4364562.pdf>.
9. Downey A. B. A Parallel Workload Model and Its Implications for Processor Allocation. *The Rachel and Selim Benin School of Computer Science and Engineering*, 2013. Available at: <http://allendowney.com/research/allocation/>.
10. The Feitelson 1996. *The Rachel and Selim Benin School of Computer Science and Engineering*, 2013. Available at: http://www.cs.huji.ac.il/labs/parallel/workload/m_feitelson96/.
11. GridMe: Grid modeling. *Google code*, 2014. Available at: <https://code.google.com/p/gridme/>.

-
12. *HPC @ Uni.lu*, 2017. Available at: <https://hpc.uni.lu/systems/gaia/>.
 13. The Jann et al 1997 Model. *The Rachel and Selim Benin School of Computer Science and Engineering*, 2013. Available at: http://www.cs.huji.ac.il/labs/parallel/workload/m_jann97/.
 14. Logs of Real Parallel Workloads from Production Systems. *The Rachel and Selim Benin School of Computer Science and Engineering*, 2013. Available at : <http://www.cs.huji.ac.il/labs/parallel/workload/logs.html>.
 15. Lublin U., Feitelson D. G. The Workload on Parallel Supercomputers: Modeling the Characteristics of Rigid Jobs. *The Rachel and Selim Benin School of Computer Science and Engineering*, 2013. Available at: <http://www.cs.huji.ac.il/~feit/papers/Rigid01TR.pdf>.
 16. Jann J., et al. Modeling of Workload in MPPs. *Job Scheduling Strategies for Parallel Processing : Lect. Notes Comput. Sci.*, Springer-Verlag, 1997, vol. 1291, pp. 95–116.
 17. Den optimalen Rechnerverbund gibt es nicht einmal auf dem Papier [There is the optimal computing cluster not only on the paper]. *Computerwoche*. Available at: <http://www.computerwoche.de/a/den-optimalen-rechnerverbund-gibt-es-nicht-einmal-auf-dem-papier,1087149>.