

Лукин Николай Алексеевич – Институт машиноведения Уральского отделения РАН; e-mail: nicklookin@mail.ru; 620049, г. Екатеринбург, ул. Комсомольская, 34; тел.: +73433788903; с.н.с.

Lookin Nickolay Alexeevich – Institute of Engineering Science, Urals Department of RAS; e-mail: nicklookin@mail.ru; 34, Komsomol street, Yekaterinburg, 620049, Russia; phone: +73433788903; senior scientist.

УДК 681.324, 519.21

В.А. Павский, К.В. Павский

СТОХАСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ И ОЦЕНКИ РАЗМЕРА СТРУКТУРНОЙ ИЗБЫТОЧНОСТИ МАСШТАБИРУЕМЫХ РАСПРЕДЕЛЕННЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ*

В рамках теории массового обслуживания построена математическая модель для оценки надежности функционирования вычислительных систем со структурной избыточностью. Модель формализована системой дифференциальных уравнений. Анализ статистики отказов вычислительных систем показывает, что поток отказов вычислительных узлов (элементарных машин) описывается распределением Вейбулла с параметром формы 0,73 и 0,78. Однако для этих форм математическая модель оказывается громоздка и не допускает аналитического решения. В то же время при форме 1 (экспоненциальное распределение) аналитическое решение может существовать. Получено аналитическое решение, позволяющее рассчитать показатели надежности. Найдена вероятность нахождения вычислительных систем в состоянии низкой производительности в зависимости от размера структурной избыточности. Для этой вероятности приведена оценка и ее погрешность. Предложен расчет математического ожидания и дисперсии числа отказавших машин в зависимости от времени. Обосновано использование экспоненциального распределения. Для расчета моментов использован подход, позволяющий записать систему дифференциальных уравнений непосредственно, минуя нахождение вероятности состояний системы. Полученные формулы и их оценки просты и могут быть использованы в инженерных расчетах. Результаты подтверждены имитационным моделированием.

Распределенные вычислительные системы; структурная избыточность; математические модели; надежность; оценки показателей; распределение Вейбулла; анализ.

V.A. Pavsky, K.V. Pavsky

STOCHASTIC SIMULATION AND INDICES ESTIMATIONS OF STRUCTURAL REDUNDANCY OF LARGE-SCALE COMPUTER SYSTEMS

The mathematical model for estimation of reliability of distributed computer systems (CS) functioning with reserve is constructed by using methods of queuing theory. The model is formalized with system of differential equations. Based on the statistics of failure for cluster CSs, it is preferable to assume that the time between failures is Weibull distributed with a shape parameter value 0.73 and 0.78. But the mathematical model with these parameters is laborious and doesn't have analytical solution. But the analytical solution for shape parameter values of 1 (exponential distribution) is possible. The analytical solution allowing to calculate reliability indices is obtained. The functional dependency of the probability of computer system's low performance on the reserve size is found. The estimations for this probability are offered. The calculation of mathe-

* Работа выполнена при поддержке РФФИ (грант №13-07-00160).

mathematical expectation and dispersion of refusal machines numbers is proposed. Formulas derivation is based on methods allowing to get system of equations for moments without finding probabilities states. The formulas and they estimations are suitable to reverse engineering. The results of analytical modeling are confirmed by simulation modeling.

Distributed computer systems; reserve; mathematical models; reliability; indices estimations; Weibull distribution; analysis.

Введение. За последние 10 лет производительность лидеров суперкомпьютеров TOP 500 в среднем каждые 3 года повышалась в 10 раз. Количество ядер в таких ВС может достигать нескольких миллионов (например, в системе IBM BlueGene/Q Sequoia с количеством узлов 98 304, занимающей 3-е место 43-й редакции списка суперкомпьютеров Top500, число ядер равно 1 572 864) [1]. Sinisa Nolic – директор IBM – отметил, что за 10 лет эксплуатации суперкомпьютеров в среднем каждые 5 минут фиксировались различного рода сбои [2]. Следовательно, с развитием вычислительных средств высокой производительности актуальность проблемы производительности, надежности, живучести не только сохраняется, но и усиливается. Отказы ЭМ, принимающих участие в решении трудоемких задач, приводят к потере промежуточных результатов вычислений и необходимости перезапуска всей параллельной программы. Последнее снижает технико-экономическую эффективность ВС и осуществимость решения трудоемких задач. Более того, отказ или восстановление элементарной машины должен приводить лишь к изменению производительности ВС и не приводить к аварийному завершению задач пользователей.

Введение в состав распределенных вычислительных систем структурной избыточности позволяет повысить их живучее функционирование.

Итак, объектом исследования являются масштабируемые распределенные вычислительные системы (ВС) со структурной избыточностью [3]. Предложена стохастическая модель функционирования вычислительных систем со структурной избыточностью. Показано, что относительно быстрая замена вышедших из строя ЭМ машинами структурной избыточности позволяет поддерживать необходимую производительность в течение длительного промежутка времени.

1. Модель. Имеется ВС, состоящая из N элементарных машин (ЭМ), n – структурная избыточность (резерв), $N - n$ – основная подсистема (рис. 1). Время работы каждой ЭМ является случайной величиной, подчиненной экспоненциальному закону, с параметром λ – интенсивностью выхода ЭМ из строя. Вышедшая из строя ЭМ заменяется на ЭМ из структурной избыточности, а сама попадает в восстанавливающую систему (ВУ) и ждет восстановления. Время восстановления является случайной величиной, подчиненной экспоненциальному закону с параметром μ – интенсивностью восстановления. Предполагается, что независимо от числа ЭМ, находящихся в ВУ, среднее время восстановления любого числа $k \leq n$ ЭМ, находящихся на восстановлении, $t_{cp} = 1/\mu$.

2. Вероятности состояний восстанавливающей системы. Требуется вычислить $P_k(t)$ – вероятность того, что в момент времени t в системе массового обслуживания (СМО) находится k требований, $t \in [0, \infty)$, $k = 0, 1, \dots, n$, и провести анализ функционирования ВС. В отличие от модели, представленной в работе [4] (где поток отказов $\lambda(t) = N\lambda$), предполагаем, что для состояния k СМО поток отказов имеет интенсивность $(N - k)\lambda$. Итак, получаем

$$\begin{cases} P_0'(t) = -N \cdot \lambda P_0(t) + \mu \sum_{k=1}^n P_k(t), \\ P_k'(t) = -((N-k) \cdot \lambda + \mu) \cdot P_k(t) + (N-(k-1))\lambda \cdot P_{k-1}(t), \\ P_n'(t) = -\mu \cdot P_n(t) + (N-(n-1))\lambda \cdot P_{n-1}(t) \end{cases} \quad (1)$$

с начальными условиями

$$P_0(0) = 1, \quad P_k(0) = 0, \quad k \neq 0.$$

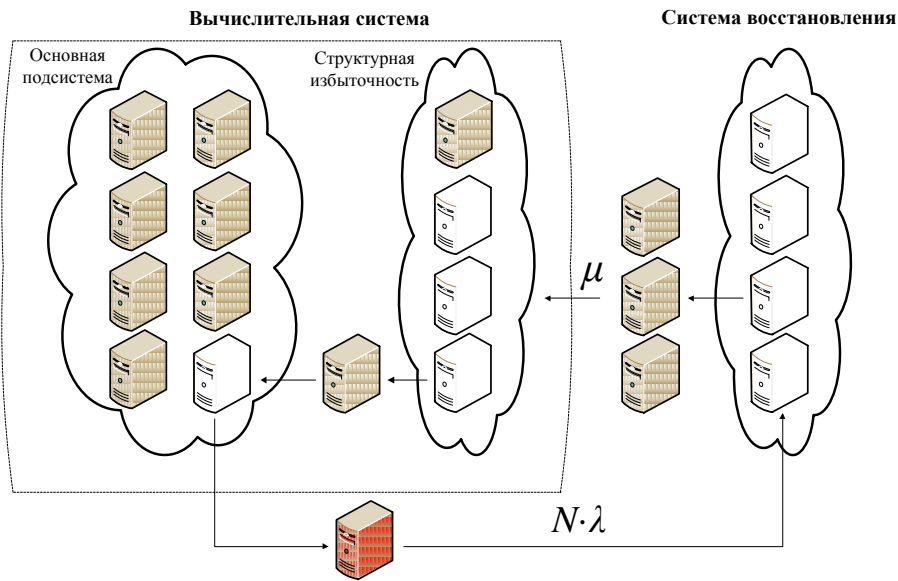


Рис. 1. Модель функционирования вычислительной системы со структурной избыточностью

Условие нормировки

$$\sum_{k=0}^n P_k(t) = 1, \quad t \in [0, \infty).$$

Решение (1) для стационарного режима $p_k = \lim_{t \rightarrow \infty} P_k(t)$, $k = 0, 1, \dots, n$, представляется в следующем виде:

$$\begin{cases} p_0 = \frac{\mu}{(N \cdot \lambda + \mu)}, \\ p_k = \left(\frac{\lambda}{\mu}\right)^k \frac{1}{((N-k) \cdot \lambda / \mu + 1)} \prod_{i=0}^{k-1} \frac{(N-i)}{((N-i) \cdot \lambda / \mu + 1)}, \\ p_n = \left(\frac{\lambda}{\mu}\right)^n \prod_{i=0}^{n-1} \frac{(N-i)}{((N-i) \cdot \lambda / \mu + 1)}. \end{cases} \quad (2)$$

Замечаем, что в (2) P_n – это вероятность пустоты резерва в стационарном режиме. Соответственно, на основе (2) считаем, что вероятность отказа ВС равна

$$P_{отк} = P_n = \left(\frac{\lambda}{\mu}\right)^n \cdot \prod_{i=0}^{n-1} \frac{N-i}{(N-i) \cdot \lambda / \mu + 1}. \quad (3)$$

Зависимость вероятности $P_{отк}$ от размера структурной избыточности (резерва) приведена на рис. 2. Очевидно, что

$$\left(\frac{(N-n)\lambda}{(N-n)\lambda + \mu}\right)^n < P_{отк} < \left(\frac{N \cdot \lambda}{N \cdot \lambda + \mu}\right)^n.$$

Пусть $\Phi_n(N) = \left(\frac{N \cdot \lambda}{N \cdot \lambda + \mu}\right)^n$, тогда

$$\left(1 - \frac{n}{N}\right) \left(1 + \frac{n \cdot \lambda}{(N-n) \cdot \lambda + \mu}\right)^n \Phi_n(N) < P_{отк} < \Phi_n(N). \quad (4)$$

Из (4) видно, что $P_{отк} \approx \Phi_n(N)$, $N \gg n$, что удобнее использовать для анализа. Например, для нахождения размера структурной избыточности по заданной доверительной вероятности того, что ВС находится в состоянии высокой производительности [4].

Оценим погрешность такого приближения. Имеем

$$\Phi_n(N) - P_{отк} < \delta(n) \Phi_n(N), \quad (5)$$

где

$$\delta(n) = 1 - \left(1 - \frac{n}{N}\right) \left(1 + \frac{n \cdot \lambda}{(N-n) \cdot \lambda + \mu}\right)^n. \quad (6)$$

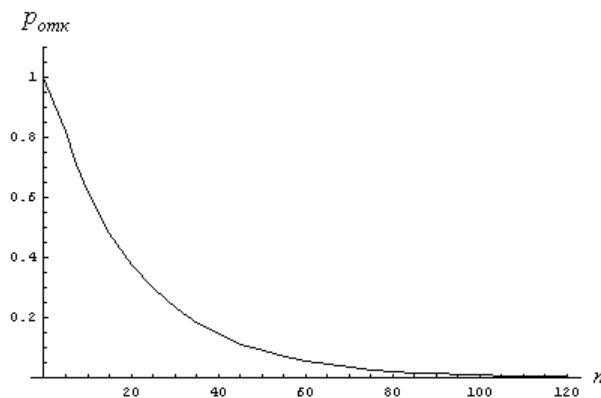


Рис. 2. Зависимость вероятности сохранения максимальной производительности от размера структурной избыточности:

$$\mu = 0,1 \text{ ч}^{-1}; \quad \lambda = 10^{-4} \text{ ч}^{-1}, \quad N = 2 \cdot 10^4 \text{ ЭМ}$$

Из рис. 3 видно, что для $\mu = 0,1 \text{ ч}^{-1}$; $\lambda = 10^{-4} \text{ ч}^{-1}$, $N = 2 \cdot 10^4 \text{ ЭМ}$ оценка погрешности (см. (5) и (6)) составляет примерно 1 % от $\Phi_n(N)$ при $n = 60$ и примерно 10 % при $n = 210$.

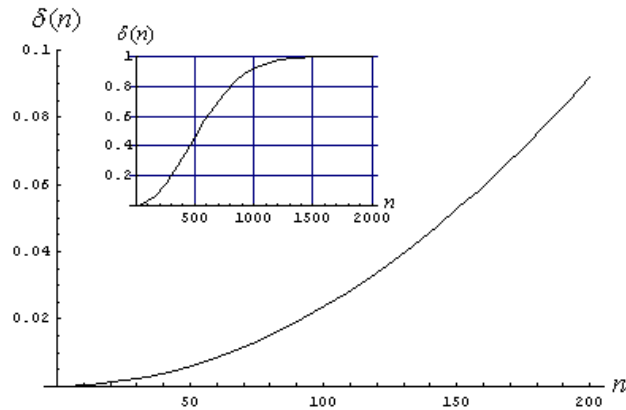


Рис. 3. Расчет $\delta(n)$ $\mu = 0,1 \text{ ч}^{-1}$; $\lambda = 10^{-4} \text{ ч}^{-1}$, $N = 2 \cdot 10^4 \text{ ЭМ}$

На рис. 4,а,б представлен пример расчета разницы $\Phi_n(N) - p_{отк}$ и $p_{отк} - \Phi_n(N-n)$ соответственно. Видно, что графики практически совпадают. Этот пример показывает, что оценка погрешности (5) несколько завышена (примерно в 2 раза).

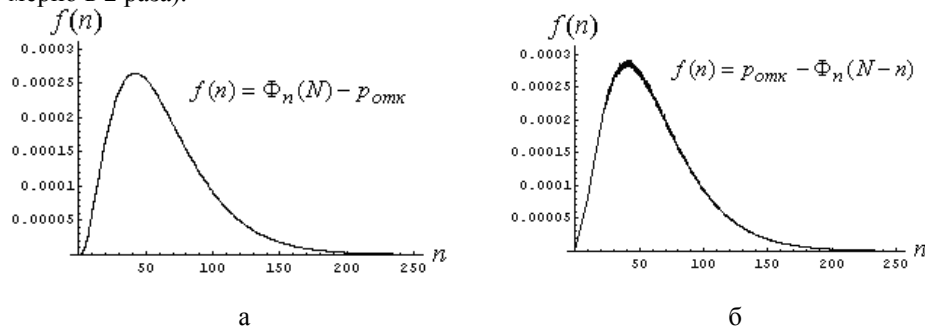


Рис. 4. Оценка погрешности $\mu = 0,1 \text{ ч}^{-1}$; $\lambda = 10^{-4} \text{ ч}^{-1}$, $N = 2 \cdot 10^4 \text{ ЭМ}$

3. Математическое ожидание числа отказавших машин и соответствующая дисперсия. Для нахождения среднего числа машин $M(t)$, ожидающих восстановления в ВУ, и дисперсии $D(t)$ воспользуемся аппаратом производящих функций [5], тогда

$$\frac{\partial}{\partial t} F(z, t) = \mu(1 - F(z, t)) - N\lambda(1 - z)(F(z, t) - P_n(t)z^n) + \lambda z(1 - z)\left(\frac{\partial}{\partial z} F(z, t) - nP_n(t)z^{n-1}\right),$$

где

$$F(z, t) = \sum_{k=0}^n z^k \cdot P_k(t), \quad 0 < z \leq 1.$$

Откуда, используя метод нахождения моментов [6], после соответствующих преобразований получаем

$$\begin{cases} \frac{d}{dt} M(t) + (\mu + \lambda)M(t) = N\lambda - (N - n)\lambda P_n(t), \\ \frac{d}{dt} Q(t) + (\mu + 2\lambda)Q(t) = 2\lambda(N - 1)M(t) - 2\lambda n P_n(t)(N - n), \\ Q(t) = D(t) - M(t) + M^2(t). \end{cases} \quad (7)$$

Решение по (7), например, можно найти численными методами (см. рис. 5), где использовались результаты [4, 7].

В работах [8, 9] проводится исследование анализ отказов в распределенных вычислительных системах. На основе статистики отказов в 20-ти кластерных ВС показано, что предпочтительнее считать, что время между отказами распределено по закону Вейбулла с параметром формы $\delta = 0,78$.

Для предложенной модели (см. п. 1) была разработана имитационная модель функционирования ВС со структурной избыточностью. На рис. 5–8 приведены графики математического ожидания и дисперсии числа отказавших машин, находящихся на восстановлении и построенных по имитационной модели (при 10 000 экспериментов), в сравнении с решением численными методами для (7). На этих рисунках $M(t)$ и $M_w(t)$ – математические ожидания среднего числа отказавших машин, для которых время работы до отказа любой ЭМ распределено по закону Вейбулла с параметрами формы $\delta = 1$ и $\delta = 0,78$ соответственно. Моделирование восстановлений выполнено согласно экспоненциальному закону $N = 1,5 \cdot 10^4$, $\lambda = 10^{-4}$ 1/ч, $\mu = 0,1$ 1/ч.

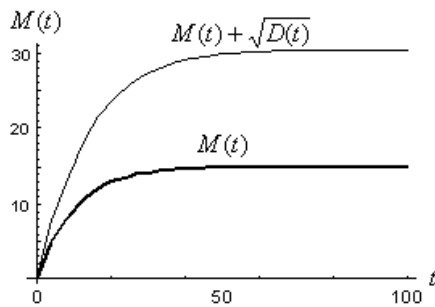


Рис. 5. Зависимость среднего числа отказавших ЭМ, находящихся в ВУ, от времени (расчет выполнен согласно формулам (7). Рассматриваются простейшие потоки)

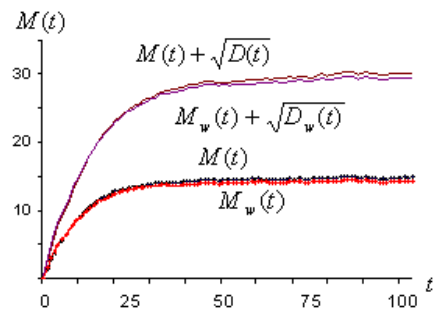


Рис. 6. Зависимость среднего числа отказавших ЭМ, находящихся в ВУ, от времени (расчет выполнен согласно имитационной модели)

Рис. 5, 6 показывают, что результаты моделирования по формулам (7) практически совпадают с имитационным. Рис. 6–8 показывают, как влияет $N \cdot \lambda$ на значение разницы между $M(t)$ и $M_w(t)$. При $N \cdot \lambda \approx 1,5$ значения $M(t)$ и $M_w(t)$ практически совпадают (рис. 6).

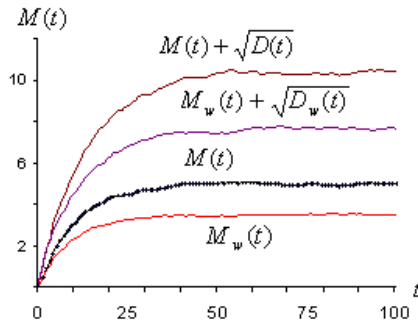


Рис. 7. Зависимость среднего числа отказавших ЭМ, находящихся в ВУ, от времени (расчет выполнен согласно имитационной модели. $N = 5 \cdot 10^3$, $\lambda = 10^{-4} 1/\text{ч}$, $\mu = 0,1 1/\text{ч}$)

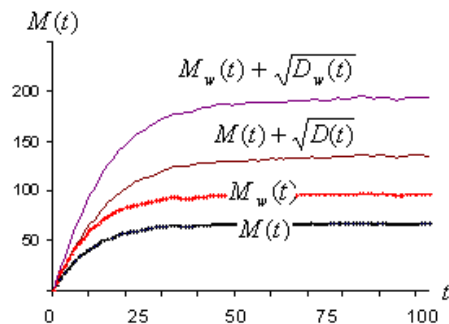


Рис. 8. Зависимость среднего числа отказавших ЭМ, находящихся в ВУ, от времени (расчет выполнен согласно имитационной модели. $N = 7,5 \cdot 10^4$, $\lambda = 10^{-4} 1/\text{ч}$, $\mu = 0,1 1/\text{ч}$)

Заключение. В работе предложена математическая модель функционирования масштабируемых вычислительных систем со структурной избыточностью. Найдена вероятность нахождения ВС в состоянии низкой производительности в зависимости от размера структурной избыточности. Предложена ее оценка и погрешность, которые более удобны для анализа. Предложен расчет среднего числа отказавших машин в зависимости от времени. Приведены результаты расчета среднего числа отказавших ЭМ в восстанавливающей системе, построенной по формулам (7) и имитационной модели, где время работы до отказа любой ЭМ является случайной величиной, распределенной по закону Вейбулла, форм 0,78 и 1. Показано, что с точностью до одного стандартного отклонения эти формы взаимозаменяемы.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Top500 Supercomputer sites// <http://www.top500.org> (дата обращения 10.11.2014).
2. *Nikolic S.* High Performance Computing Directions: The Drive to ExaScale Computing // Труды Международной научной конференции “Параллельные вычислительные технологии (ПаВТ’2012). – Новосибирск, 2012, URL: <http://pavt.susu.ru/2012/talks/Nikolic.pdf> (дата обращения 10.11.2014).
3. *Хорошевский В.Г.* Архитектура вычислительных систем. – М.: МГТУ им. Баумана, 2008. – 520 с.
4. *Павский В.А., Павский К.В., Хорошевский В.Г.* Математическая модель и расчет показателей функционирования вычислительных систем со структурной избыточностью // Известия ЮФУ. Технические науки. – 2012. – № 5 (130). – С. 37-41.
5. *Саати Т.Л.* Элементы теории массового обслуживания и ее приложения. –3-е изд. – М.: Книжный дом «ЛИБРОКОМ», 2010. – 520 с.
6. *Хорошевский В.Г., Павский В.А., Павский К.В.* Расчет показателей живучести распределенных вычислительных систем // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2011. – № 2 (15). – С. 81-88.
7. *Овчаров Л.А.* Прикладные задачи теории массового обслуживания. – М.: Машиностроение, 1969. – 324 с.
8. *Schroeder B., Gibson Garth A.* A large-scale study of failures in high-performance computing systems // Proceedings of the International Conference on Dependable Systems and Networks (DSN2006), Philadelphia, PA, USA, June 25-28. – 2006. – 10 p.
9. Analyzing failure data. URL: <http://www.pdl.cmu.edu/FailureData/> (дата обращения 10.11.2014).

REFERENCES

1. Available at: Top500 Supercomputer sites// <http://www.top500.org> (Accessed 10 November 2014).
2. *Nikolic S.* High Performance Computing Directions: The Drive to ExaScale Computing, Trudy Mezhdunarodnoy nauchnoy konferentsii "Parallelnye vychislitel'nye tekhnologii (PaVT'2012) [Proceedings of International scientific conference "Parallel computing technologies (Pushchino'2012)]. Novosibirsk, 2012, Available at: <http://pavt.susu.ru/2012/talks/Nikolic.pdf> (Accessed 10 November 2014).
3. *Khoroshevskiy V.G.* Arhitektura vychislitel'nykh sistem [Architecture of computing systems]. Moscow: MGTU im. Baumana, 2008, 520 p.
4. *Pavskiy V.A., Pavskiy K.V., Khoroshevskiy V.G.* Matematicheskaya model' i raschet pokazateley funktsionirovaniya vychislitel'nykh sistem so strukturnoy izbytochnost'yu [Mathematical model and of indices calculation of computer systems functioning with structural redundancy], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2012, No. 5 (130), pp. 37-41.
5. *Saati T.L.* Elementy teorii massovogo obsluzhivaniya i ee prilozheniya [Elements of queueing theory and its applications]. 3rd ed. Moscow: Knizhnyy dom «LIBROKOM», 2010, 520 p.
6. *Khoroshevskiy V.G., Pavskiy V.A., Pavskiy K.V.* Raschet pokazateley zhivuchesti raspredelennykh vychislitel'nykh sistem [To calculate the survivability of distributed computing systems], *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika* [Vestnik of Tomsk state University. Control, computer engineering and Informatics], 2011, No. 2 (15), pp. 81-88.
7. *Ovcharov L.A.* Prikladnye zadachi teorii massovogo obsluzhivaniya [Applied problems in theory of mass service]. Moscow: Mashinostroenie, 1969, 324 p.
8. *Schroeder B., Gibson Garth A.* A large-scale study of failures in high-performance computing systems, *Proceedings of the International Conference on Dependable Systems and Networks (DSN2006), Philadelphia, PA, USA, June 25-28, 2006*, 10 p.
9. Analyzing failure data. Available at: <http://www.pdl.cmu.edu/FailureData/> (Accessed 10 November 2014).

Статью рекомендовал к опубликованию д.т.н., профессор А.М. Гудов.

Павский Валерий Алексеевич – Кемеровский технологический институт пищевой промышленности; e-mail: pavva46@mail.ru; 650056, г. Кемерово, б-р Строителей, 47; тел.: 83842734200; кафедра высшей математики; зав. кафедрой; д.т.н.; профессор.

Павский Кирилл Валерьевич – Федеральное государственное бюджетное учреждение науки Институт физики полупроводников им. А.В. Ржанова СО РАН; e-mail: pkv@isp.nsc.ru; 630090, г. Новосибирск, пр. Лаврентьева, 13; тел.: 8383332171, 83833305626; лаборатория ВС; д.т.н.; научный сотрудник.

Pavsky Valery Alexeevich – Kemerovo institute of technology of the food-processing industry; e-mail: pavva46@mail.ru; 47, Stroiteley bulvar, Kemerovo, 650056, Russia; phone: +73842734200; the department of mathematica chair; head of department; dr. of eng. sc.; professor.

Pavsky Kirill Valerievich – A.V. Rzhzanov institute of semiconductor physics of Siberian branch of the RAS; e-mail: pkv@isp.nsc.ru; 13, ak/ Lavrentieva av., Novosibirsk, 630090, Russia; phone: +7383332171, +73833305626; dr. of eng. sc.; researcher.