Stochastic Modelling of Large-Scale Distributed Computer Systems Functioning with Group Restorations

Valery Pavsky [1], and Kirill Pavsky [2]

[1] Kemerovo Institute of Food Science and Technology (University), Kemerovo, Russia
pavva46@mail.ru

[2] Rzhanov Institute of Semiconductor Physics Siberian Branch of Russian Academy of Sciences, pr. Lavrentieva, 13, Novosibirsk, 630090, Russia
pkv@isp.nsc.ru

Abstract.
The model of functioning of distributed computer systems with group restorations of failed machines is considered. The model is formalized in the form of a system of differential equations, in which are unknown probabilities of the states. The paper proposes solutions for calculating the mathematical expectation of the number of efficient machines and variances, which are the basis for creating indices of potential robustness.
The investigation of the functioning of the CS under the assumption of the validity of the exponential law of failure of computers makes it possible, due to a well-developed theory, to obtain profound results, in contrast to the use of other distribution laws. And the obtained analytical solutions can be used for rapid analysis of the functioning of the CS.

Keywords: distributed computer systems, mathematical model, group restoration, number of working machines, analytical solutions.

1 Introduction

The problem of the reliability of high-performance scalable computing devices, including supercomputers and computer systems [1], increases with the number of elementary machines (EM) [2], the number of which in such systems ranges from several tens to hundreds of thousands [3]. Practice shows that in scalable computing systems (CS) the time between different types of failures can be measured by hours [4,5]. The preservation of the efficiency of the CS in the conditions of failures [6-8], the analysis of functioning with regard to robustness, is an urgent task.
This work is devoted to the development of means for analyzing the efficiency of the operation of larger-scale distributed CS [9, 10]. The queuing theory apparatus is used as an analysis tool.
The paper proposes formulas for calculating the mathematical expectation and variance of the number of working machines, which are the basis for creating indicators of potential robustness in group recovery [2]. The indices of potential robustness of

the CS take into account the fact that in the solution of problems all the working EMs are used, the number of which is actually instant. This assumes that parallel programs of complex tasks, when implemented on survivable CS, are capable of using the total performance of all working EM systems. Assuming mathematical idealization, CS can be regarded as a stochastic object.

2 Model of the CS functioning

A stochastic model of the distributed CS operation is proposed, in which the EMs are not absolutely reliable [2]. Each of them fails with λ intensity. Out of order EM gets into the restoration system and is waiting for resoration. At random moments of time, the recovery is exercised by groups in r EM with μ intensity (Fig.1). We believe that at the initial time the system contains n EMs. As performance indices (evaluation of potential viability), we use numerical characteristics – the mathematical expectation of the number of effective EMs and its variance [2].

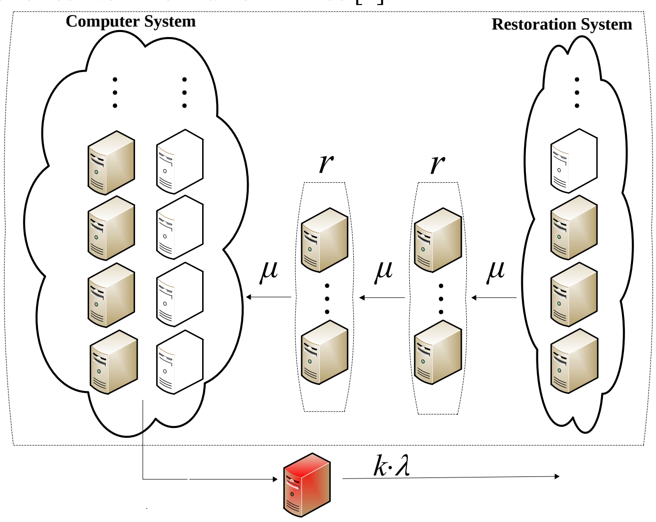

Fig. 1. Model of the functioning of a scalable CS in the group recovery of failed EM.

In constructing the model of the functioning of the CS, we use the methods of queuing theory, in which models of this class are formulated according to the tried-and-tested method – a system of differential equations is compiled, and the probability distribution is considered as unknown functions. The analytical solution of such systems is far from always available, even stationary [2]. Assuming further mathematical idealization of the model, we believe that the number of EMs in the CS is potentially infinite, which is permissible because of its scalability. From the formalized system of differential equations of the model, we find an analytical solution, directly for the moments, bypassing the probability distribution.

3 Mathematical Model

The queuing system (QS) with an infinite number of channels receives a Poisson stream of packets with μ intensity [11, 12]. Each packet consists of r requirements. At any fixed time $t \in [0,\infty)$, the QS is in one of a number of incompatible states C_k - where k is the number of requirements in the QS, including undeserved requirements. The service time of each requirement is subject to an exponential distribution with a parameter λ. If the system is in the C_k state, then one of the k requirements leaves the system with the $k\lambda$ intensity. It is required to find the mathematical expectation of the $M(t)$ state number, in which the system is located when servicing the requirements and the corresponding variance $D(t)$, provided that $M(0)=n, D(0)=0$. Figure 2 shows a graph-scheme of QS states that allows us to better understand the relationship between the formulation of the model and its formalization by a system of differential equations.

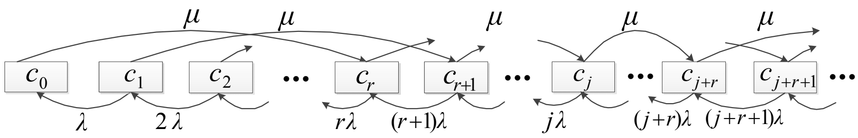

Fig. 2. Graph-scheme of QS states, for the scalable CS operation model.

$P_k(t)$ denotes the probability that at the instant t the QS is in the state C_k, $k = 0,1,\ldots$ The system of differential equations has the form:

$$\begin{cases} P_0'(t) = -\mu P_0(t) + \lambda P_1(t), \\ P_k'(t) = -(\mu + k\lambda)P_k(t) + (k+1)\lambda P_{k+1}(t), \ 0 < k < r, (*) \\ P_k'(t) = -(\mu + k\lambda)P_k(t) + \mu P_{k-r}(t) + (k+1)\lambda P_{k+1}(t), \quad k \geq r. \end{cases} \tag{1}$$

To find the mathematical expectation and variance, we apply the method of generating functions. We transform the system of equations (1) so that in it the middle equation (*), with sliding parameter k, would be obtained from the last equation, at 0<*k*<*r*. For this, we set $P_{k-r}(t) \equiv 0$, $0 \leq k < r$, then the system (1) will look as follows:

$$\begin{cases} P_0'(t) = -\mu P_0(t) + \lambda P_1(t), \\ P_k'(t) = -(\mu + k\lambda)P_k(t) + \mu P_{k-r}(t) + (k+1)\lambda P_{k+1}(t), \end{cases} \tag{2}$$

In accordance with the formulation of the model, we give the initial conditions

$$P_n(0) = 1; \ P_k(0) = 0, \ k \neq n \tag{3}$$

and the normalization condition, which is a consequence of the formulation of the model,

$$\sum_{k=0}^{\infty} P_k(t) = 1. \tag{4}$$

To solve the system (2), taking into account (4), we introduce the generating function

$$F(z,t) = \sum_{k=0}^{\infty} z^k P_k(t)\,. \tag{5}$$

Multiplying the corresponding equation k of system (2) by z^k and summing, we obtain

$$\sum_{k=0}^{\infty} z^k P_k'(t) = -\mu \sum_{k=0}^{\infty} z^k P_k(t) - \lambda \sum_{k=1}^{\infty} kz^k P_k(t) + \mu \sum_{k=0}^{\infty} z^k P_{k-r}(t) + \lambda \sum_{k=0}^{\infty} (k+1) z^k P_{k+1}(t)$$

Expressing each term of the resulting equation in terms of the generating function and reducing such terms, we obtain a linear equation for the generating function

$$\frac{\partial}{\partial t} F(z,t) = -\mu(1-z^r) F(z,t) + \lambda(1-z) \frac{\partial}{\partial z} F(z,t)\,. \tag{6}$$

To find the mathematical expectation $M(t)$ and the corresponding variance $D(t)$, we use the method of moments [13].

We take the derivative with respect to the variable z from the right and left sides of (6), then

$$\frac{\partial^2}{\partial z \partial t} F(z,t) = \mu r z^{r-1} F(z,t) - \mu(1-z^r) \frac{\partial}{\partial z} F(z,t) - \lambda \frac{\partial}{\partial z} F(z,t) + \lambda(1-z) \frac{\partial^2}{\partial z^2} F(z,t)\,. \tag{7}$$

Assume ξ is a random variable characterizing the number of requirements in the QS. From (5) and the properties of the generating functions [13, 14] it follows that

$$M\xi = \frac{\partial}{\partial z} F(1), \text{ a } D\xi = \frac{\partial^2}{\partial^2 z} F(1) + \frac{\partial}{\partial z} F(1) - \left(\frac{\partial}{\partial z} F(1) \right)^2 .$$

After the corresponding transformations over (7) we obtain

$$\begin{cases} \dfrac{\partial}{\partial t} M(t) + \lambda M(t) = r\mu, \\ \dfrac{\partial}{\partial t} Q(t) + 2\lambda Q(t) = 2r\mu M(t) + r(r-1)\mu, \\ D(t) = Q(t) + M(t) - M^2(t), \end{cases} \tag{8}$$

where, on the basis of (3), $M(0) = n, \quad D(0) = 0$.

The solution for (8), taking (3) into account, will be

$$\begin{cases} M(t) = \dfrac{r\mu}{\lambda} + \left(n - \dfrac{r\mu}{\lambda} \right) e^{-\lambda t}, \\ D(t) = (n^2 - n - \dfrac{r\mu}{\lambda}(2n - \dfrac{r\mu}{\lambda} + \dfrac{(r-1)}{2}))e^{-2\lambda t} + \dfrac{r\mu}{\lambda}(\dfrac{r\mu}{\lambda} + 2\left(n - \dfrac{r\mu}{\lambda} \right) e^{-\lambda t}) + \\ + \dfrac{r(r-1)\mu}{2\lambda} + M(t) - M^2(t) \end{cases} \tag{9}$$

For the stationary regime we have

$$\begin{cases} M = \dfrac{r\mu}{\lambda}, \\ D = \dfrac{r^2\mu + r\mu}{2\lambda}, \\ \sigma = \sqrt{\dfrac{r^2\mu + r\mu}{2\lambda}}. \end{cases} \qquad (10)$$

Figure 3 shows the calculation of the mathematical expectation with allowance for dispersion by formulas (9) at $\mu = 10^{-1}$ 1/hr, $\lambda = 10^{-3}$ 1/hr, $r = 20$, $n = 10^4$ EM.

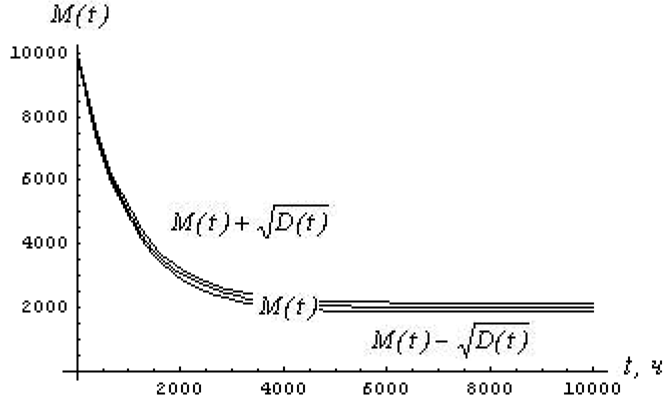

Fig. 3. The average number of working machines with variance.

From formulas (9) and their graphical implementation (Fig. 3), it is also evident that the robustness of the CS and high performance is provided by two thousand EM, while reliability should be calculated on the basis of ten thousand EM. Therefore, in order to achieve sufficient robustness, given the volume of the aircraft, it is necessary to increase the reliability of the element base and other parameters associated with the process of executing parallel programs.

In Fig. 4 shows an example where the system is initially in a state of equilibrium between failure and recovery ((9) and (10)) for $\mu = 5 \cdot 10^{-2}$ 1/hr, $\lambda = 10^{-4}$ 1/hr, $r = 20$, $n = 10^4$ EM. In Fig. 5 also shows the calculation of the root-mean-square deviation for this case. It can be seen how quickly the system enters a stationary mode and it is possible to estimate the limits of its potential performance.

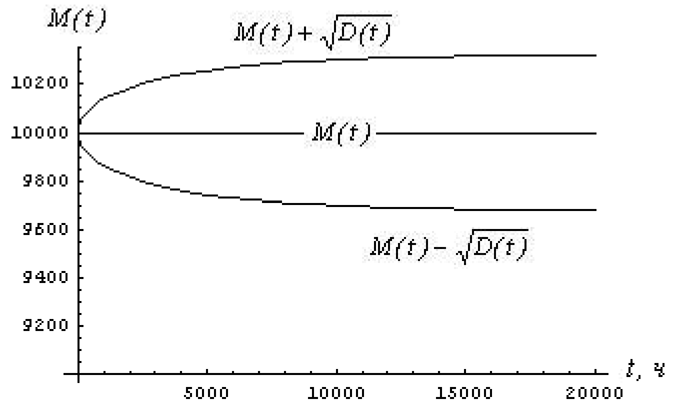

Fig. 4. The average number of working machines with variance.

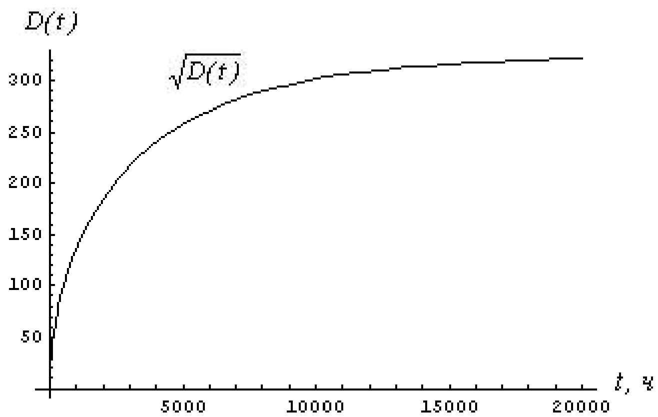

Fig. 5. Calculation of variance.

4 Conclusion

Calculations of the mathematical expectation and dispersion of the number of working machines in distributed scalable computing systems use a technique designed primarily to assess the effectiveness of the functioning of predictive and projected computing systems. Here the mathematical apparatus is used as a research method, demonstrating not only the result, but also the prospects for its development. This is the main advantage of analytical solutions - internal information meaningfulness formulas to numerical and algorithmic approaches, using this device as a research tool. The solution is obtained by the method of moments [13]. In the theory of queuing of this type, models are usually formalized by systems of differential equations with unknowns that form a probability distribution. As a rule, this is sufficient for constructing a probability space. Consequently, any probabilistic characteristics associated with the random value of this space can be obtained one way or another. In our case, it is impossible to find an exact solution of the probability distribution, since, even in the steady-state regime, complications arise that lead to an approximate solution. The method of moments makes it possible to find an exact solution for the moments of any order. In addition, finding an exact solution allows us to obtain additional information that is inaccessible to an approximate solution. For example, the number n, EM in the computer system, appears conditionally (in the initial conditions and in the formulas for the transitional regime), but is absent in the systems of differential equations (1), (2) and in the formulas of the stationary regime – this is the property of the Markov processes. In addition to the quantitative evaluation of the productivity of the QS, a qualitative assessment is obtained – it is impossible to achieve any desired pre-set performance by simply increasing the computing system by elementary machines, without improving their parameters.

References

1. Dongarra J. J., A. J. van der Steen High-performance computer systems: Status and outlook, Acta Numerica (2012), pp. 1–96.
2. Khoroshevsky V.G. Architecture of computer systems. Moscow: Bauman MSTU, 2008, p. 520.
3. TOP500 Supercomputers Official Site. TOP500 Lists [website] [Electronic resource] / URL: http://www.top500.org (16.05.2017).
4. Nikolic S. High Performance Computing Directions: The Drive to ExaScale Computing. // Proceedings of the International Scientific Conference "Parallel Computing Technologies (ПаВТ'2012). –Novosibirsk, 2012, URL: http://pavt.susu.ru/2012/talks/Nikolic.pdf (16.05.2017).
5. 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.
6. Kalyaev I.A., Korobkin V.V., Melnik E.V., Malakhov I.V. Fault-tolerant control computer complex of a VVER-type reactor of a rechargeable nuclear reactor // Mechatronics, Automation, Control. - 2003.-No. 3. P.143-146.
7. S. Di, M. S. Bouguerra, L. Bautista-Gomez, and F. Cappello. Optimization of multilevel checkpoint model for large scale hpc applications. In Parallel and Distributed Processing Symposium, 2014 IEEE 28th International, pages 1181–1190, May 2014.
8. Korneev V.V., Semenov D.V., Telegin P.N., Shabanov B.M. Fault-tolerant decentralized grid resource management // Izvestiya Vuzov. Electronics. 2015. No. 1, P. 83-89.
9. Vishnevsky, V.M. Theoretical Foundations of Computer Network Design. V.M. Vishnevsky. - Moscow: Technosphere, 2003. - 512 p.
10. Khoroshevsky V.G. Models of analysis and organization of large-scale distributed computer systems. // Electronic modeling. - Kiev, 2003. - Vol. 25, No. 6.
11. Saati T.L. Elements of queuing theory and its applications. Ed. 3rd. - Moscow: The Book House "LIBROKOM", 2010. - 520p.
12. Kleinrok L. Queuing theory. M .: Mechanical Engineering, 1979. - 432 p.
13. Khoroshevsky V.G., Pavsky V.A. Calculating the efficiency indices of distributed computer system functioning// Optoelectronics, Instrumentation and Data Processing. – 2008. – V. 44. – №2. – P. 95–104.

14. Gibson Garth A. [Electronic resources] // Analyzing failure data: [сайт]. URL: http://www.pdl.cmu.edu/FailureData/ (27.05.2017 г.).