

Timely Redundant Service of Requests by a Sequence of Cluster

Vladimir Bogatyrev^{1,2,3}[0000-0003-0213-0223], Stanislav Bogatyrev¹[0000-0003-0836-8515], and Anatoly Bogatyrev¹[0000-0001-5447-7275]

¹ JSC NEO Saint Petersburg Competence Center, St. Petersburg, Russia
{info,stanislav,anatoly}@nspcc.ru

<https://www.nspcc.ru/en>

² ITMO University, Saint Petersburg, Russia
vabogatyrev@corp.ifmo.ru

³ Saint-Petersburg State University of Aerospace Instrumentation, Russia

Abstract. The possibilities of increasing the efficiency of the redundant service of latency-critical requests by a sequence of nodes of a multilevel cluster have been investigated. The analytic model is proposed, and the effectiveness of the option of redundant request servicing shown for which a copy of the request, executed first, is transferred for redundant service to the next cluster level, the remaining copies are destroyed.

The novelty of the proposed model is that it allows to taking into account the requirements of not exceeding the maximum permissible total accumulated waiting time for sequential redundant request servicing at all levels of the system.

A variant for which a certain number of copies are created during the formation of the request, for each of which a path is predefined as a sequence of nodes of different levels involved in maintenance, is considered as a prototype.

Keywords: Redundant service · Criticality of latency requests · Multi-level cluster · Distribution of requests.

1 Introduction

For multi-level cluster systems operating in real-time (especially in the cyber-physical systems), the reliability of operation is determined not only by the reliability of the system structure but also by the probability of timely execution of requests that are critical to service delay [1-4]. Providing timely execution of critical requests can be based on traffic prioritization [5-7] and load balancing [8-9].

Copyright ©2019 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

Transport coding [10, 11] reduces the average network delay of message transmissions, in which message fragments are transmitted along different routes, and as a result of encoding (information redundancy), even with the loss or delivery of part of the message with errors, it is possible to recover the entire message without retransmissions.

Reliability and timeliness of query execution can be improved as a result of redundant servicing of copies of queries with the issuance of the first serviced copy [12,14], the study of these capabilities in clusters, and multi-path transmission systems with redundant communication channels [15].

The redundant service of the request is considered successful if at least one copy of it is executed correctly without exceeding the maximum permissible total waiting time in the queues of the nodes that successively execute it. The effectiveness of redundant maintenance for multi-level systems is estimated by the probability of the timeliness of multi-stage maintenance (passing through all nodes making up the path) of at least one copy of the request [12-14].

By a path is meant a set of nodes whose operability ensures the fulfillment of a functional task of sequentially servicing a request. The path is active if the nodes included in it are involved in servicing the request or its copies. Potentially possible paths that are not used in executing generated copies of the request relate to backup paths.

The aim of the article is research the possibility of increasing the probability of timely servicing of redundant requests during their successful passage through the nodes of the cluster at all levels, taking into account the accumulation of the waiting time at each level. Reducing delays is achieved as a result of redundant servicing of copies of delay-critical requests.

The scientific novelty of this study is to build an analytic model of redundant query servicing by a sequence of clusters, with the requirement not to exceed the maximum permissible total waiting time in the queues of the nodes of the copies of the requests, the first in time performed at each level of the cluster.

The relevance of the research of the guaranteed stability of the functioning of distributed computer systems in real time is due to the active implementation of the technology of cyberphysical systems at the present stage of development, including those built on the basis of the Internet of Things technology [16-19].

2 Options for the formation of the path for redundant requests service in a multi-level cluster

As the object of research consider m levels computer cluster comprising at i -th level of n_i parallel-connected servers, each of them can be represented as single-channel queuing systems of the M / M / 1 type with infinite queues [20, 21].

In a multi-level cluster, the request is serviced sequentially by servers of all levels included in the request service path. With redundant maintenance at each level, multiple copies of the request are executed.

The following options for organizing redundant services in a multi-level cluster are researched:

Option S_1 : k copies of request (replicas) are created in the node of the request source, and for each copy, the service path (route) is specified with the servers that execute the request (copy of the request) at each level [14].

Option S_2 : When a request is generated by a source, it's k_1 copies are created, distributed for services in k_1 first-level servers. When one of the copies is serviced at the first level, k_2 copies of requests are created, which are transferred to services in selected k_2 next-level servers, and so on, until the request is serviced all m levels of the cluster [14].

For option S_1 : k paths of request execution are generated from the source, the change of which, as well as the multiplicity of backup copies, does not occur during sequential servicing of the request. The number of copies of requests (redundancy ratio) at different levels is the same. If the number of redundant service paths generated when a request arrives in the system is greater than the multiplicity of structural reservation of nodes at any level, then several copies of the request are sequentially serviced.

For option S_2 : request servicing paths are dynamically formed during the transfer of a request between cluster levels. The multiplicity of backup copies at different levels of the cluster may vary [14].

From [14], an analytic model of redundant request servicing by a sequence of clusters is known, when the maximum allowable waiting time is set for nodes at each level. Copies of requests whose waiting time is longer than the maximum set for server nodes at each level are destroyed.

In this paper, we set the task of constructing a new model that allows us to take into account the requirements of not exceeding the maximum allowable accumulated waiting time for sequential redundant request servicing at all levels of the system, including for service option S_2 , when generating replica requests distributed for servicing to cluster servers ($i + 1$) level is implemented by the first completed copy of the request at the i -th level.

3 Redundant service in a cluster of n parallel-connected servers

The probability that the queue request is shorter than the time t is calculated as [16, 17] in a cluster of n parallel connected servers with an input stream intensity Λ and average query execution time v :

$$P = 1 - \frac{\Lambda v}{n} e^{(\frac{\Lambda}{n} - \frac{1}{v})t},$$

The probability of the first completion of waiting in the queue of one of the copies in a time not exceeding t is calculated as [14] for redundant servicing of requests with multiplicity k (formations of k copies of requests sent to different servers in the cluster):

$$P = 1 - \left[\frac{\Lambda k v}{n} e^{(\frac{\Lambda k}{n} - \frac{1}{v})t} \right]^k. (1)$$

and the average time of the first time to complete waiting for a copy of a request in at least one of the k cluster servers that receive copies of the request in the queue is defined as [14]:

$$T_0 = \int_0^{\infty} \left[\frac{\Lambda v k}{n} e^{(\frac{\Lambda k_1}{n} - \frac{1}{v})t} \right]^k dt,$$

and the average residence time of the request in the cluster until the first copy of the request arrives from the queue of one of the nodes is calculated as:

$$T = v + \int_0^{\infty} \left[\frac{\Lambda v k}{n} e^{(\frac{\Lambda k_1}{n} - \frac{1}{v})t} \right]^k dt,$$

it should be noted that the average residence time of a selected (specific) copy of the request in the server node is as [20, 21]

$$T = \frac{v}{(1 - \Lambda k v / n)}.$$

The probability of a server failure or failure during the waiting time t is estimated as $r = e^{-(\lambda + \lambda_c)t}$, where λ, λ_c are intensity of server failures and errors. Then the probability of timely execution of an unreserved request, taking into account the impact of failures and errors of nodes we calculate as:

$$P = \left(1 - \frac{\Lambda v}{n} e^{(\frac{\Lambda}{n} - \frac{1}{v})t}\right) e^{-(\lambda + \lambda_c)t},$$

The probability of timely execution of an unreserved request, taking into account the possibility of failures, node errors, and calculation errors, is calculated as:

$$P = \left(1 - \frac{\Lambda v}{n} e^{(\frac{\Lambda}{n} - \frac{1}{v})t}\right) e^{-(\lambda + \lambda_c)t} e^{-(\lambda + \lambda_c + \lambda_o)v},$$

where λ_o is the error rate over time v the server executes the requests.

For the initial state of readiness of all n nodes of the cluster, the probability of waiting for at least one of k reserved copies of the request in the server queue for a time less than t will be found as:

$$P = 1 - \left[1 - \left(1 - \frac{\Lambda v}{n} e^{(\frac{\Lambda}{n} - \frac{1}{v})t}\right) e^{-(\lambda + \lambda_c)t}\right]^k.$$

The probability of timely and error-free redundant service in a cluster of at least one of k formed copies of the request is calculated as:

$$P = 1 - \left[1 - \left(1 - \frac{\Lambda v}{n} e^{(\frac{\Lambda}{n} - \frac{1}{v})t}\right) e^{-(\lambda + \lambda_c)t} e^{-(\lambda + \lambda_c + \lambda_o)v}\right]^k.$$

The example of calculating the probability of timely execution of reserved requests and the average time for their waiting is presented below. In the calculations, we assume that the average time of request execution $v = 4 \cdot 10^{-4} s^{-1}$, $n = 5$, and the maximum allowable waiting time is $8 \cdot 10^{-4} s^{-1}$. The dependence of the probability of timely execution P and average time waiting until the first

execution of at least one of k copies of requests from the input stream intensity Λ is presented in Fig. 5 and Fig. 6, respectively. In the graphs below, curves 1, 2, 3 correspond to the multiplicity of reservation requests $k = 1, 2, 3$.

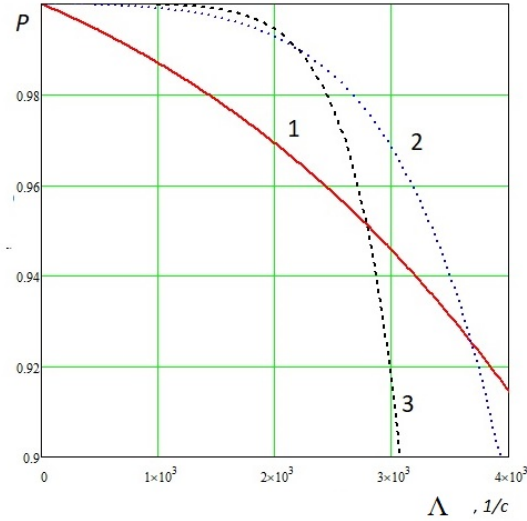


Fig. 1. The dependence of the probability of the timely execution of P on the intensity of the input stream Λ .

The dependence of the probability of the timely execution P and the average wait time until the first execution of at least one of the k copies of the request from the multiplicity of reservation requests is shown in Fig. 3 and Fig. 4 are presented in Fig. 1 and Fig. 2, respectively. In the presented graphs, curves 1,2,3 correspond to the input flow intensity $\Lambda = 1500, 2000, 2200$ 1/s

The presented dependencies allow us to conclude that there is an area of expediency of redundant servicing of requests, depending on the intensity of their receipt.

4 Redundant service with the preliminary formation of service paths according to option S_1

The servers will be presented in the form of single-channel queuing systems of the $M / M / 1$ type with endless queues [20, 21]. The probability of a request in a multi-level cluster is determined in accordance with [12] for the option of redundant service S_1 . Maximum allowable total time of waiting requests in the queues of nodes at all levels of the system t_0 is divided into N intervals, the

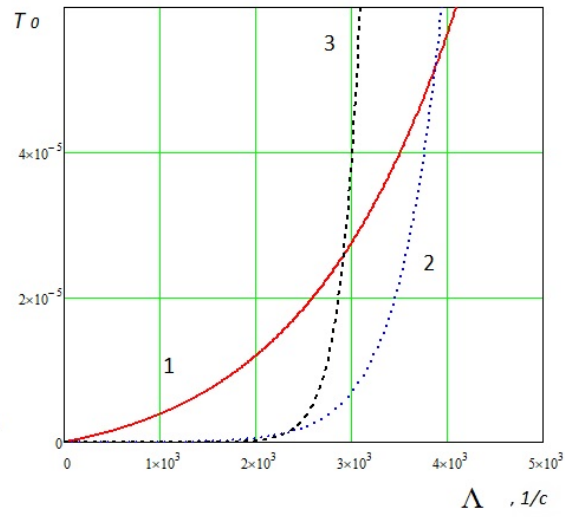


Fig. 2. Dependence of the average waiting time before the first execution of at least one of the k copies of the request on the input stream intensity Λ .

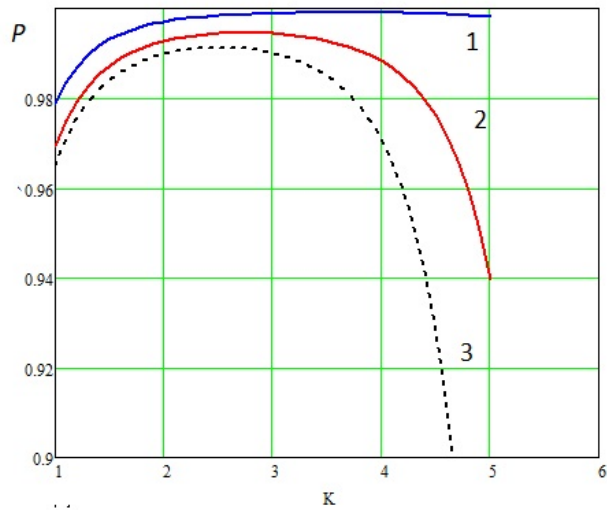


Fig. 3. Dependence of the probability of timely execution of the request P on its reservation ratio.

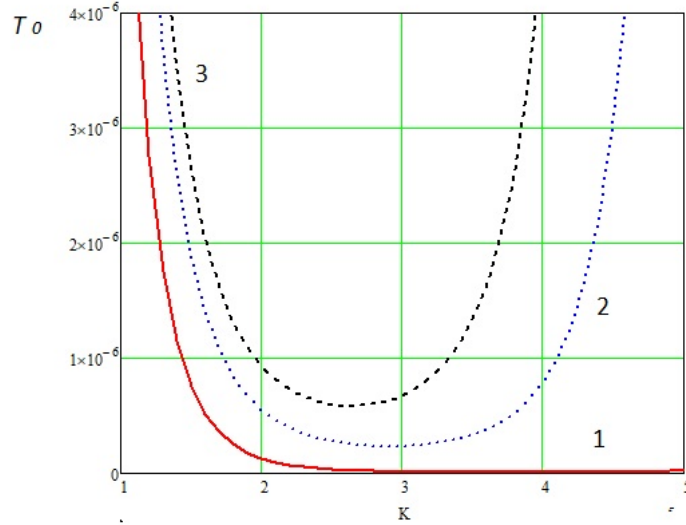


Fig. 4. Dependence of the average waiting time before the first execution of at least one of the k copies of the request on the multiplicity of reservation requests.

duration of which is equal to $t = t_0/N$. Let us single out the moments of the beginning of the countdown of the waiting intervals $i = 0, 1, \dots, N-1$. For the first level, the intervals $i = 0, 1, \dots, N-1$ are counted, for the second level, the intervals $j = 0, 1, \dots, N-1-i$ are counted (taking into account the already accumulated expectations), for the third level - intervals $d = 0, 1, \dots, N-1-i_1-i_2$ and so on [12]. Thus, the remaining acceptable wait time in the queue of the node serving the request will be $t_h = t_0 - (i_0 + i_1 + \dots + i_{h-1})t_0/N$. For a two-level system, we assume that given are the average query execution time as v_1, v_2 in nodes n_1, n_2 of the corresponding levels.

The probability of not exceeding the maximum permissible total waiting time in queues of all levels t_0 for each of k assigned paths of redundant execution of one copy of the request for a two-level cluster is calculated as:

$$\begin{aligned}
 p = & \sum_{i=1}^{N-1} \{ [1 - \{(k\Lambda/n_1)v_1 \exp(-i \frac{t_0}{N}(v_1^{-1} - (k\Lambda/n_1)))\}] - \\
 & - b_i [1 - \{(k\Lambda/n_1)v_1 \exp(-\frac{t_0}{N}(i-1)(v_1^{-1} - (k\Lambda/n_1)))\}] \times \\
 & \times [1 - \{(k\Lambda/n_2)v_2 \exp([t_0 - i \frac{t_0}{N}](v_2^{-1} - (k\Lambda/n_2)))\}] \},
 \end{aligned}$$

where

$$b_i = \begin{cases} 1, & \text{if } i > 1, \\ 0, & \text{if } i = 1. \end{cases}$$

A request is considered to be completed in a timely manner if the total waiting time in queues during the sequential passage of all servers for at least one assigned path does not exceed the maximum permissible time t_0 for redundant request servicing with reservation ratio k according to service option S_1 . Thus, the probability of timely redundant service according to option S_1 is calculated as [12]

$$P = 1 - (1 - p)^k.$$

5 Redundant service with the formation of copies of requests at each cluster level according to option S_2

The probability of fulfilling the request in a time not exceeding the specified t_0 , taking into account the variation in the accumulation of service time at different levels of the cluster when organizing redundant service according to option S_2 , will be estimated below.

For redundant service according to option S_2 , in contrast to S_1 , service replication is carried out at each level and according to the first result obtained, the specified number of copies of the request for execution is generated and distributed to the servers of the next cluster level. Considering that the probability of tinning for a time not exceeding t for at least one of k replicas of a request is calculated according to (1), we find the probability of timely execution of requests for a two-level cluster as:

$$\begin{aligned} P = & \sum_{i=1}^{N-1} \{ [1 - \{(k_1 \Lambda / n_1) v_1 \exp(-i \frac{t_0}{N} (v_1^{-1} - (k_i \Lambda / n_1)))\}^{k_1}] - \\ & - b_i [1 - \{(k_1 \Lambda / n_1) v_1 \exp(-\frac{t_0}{N} (i-1) (v_1^{-1} - (k_1 \Lambda / n_1)))\}^{k_1}] \times \\ & \times [1 - \{(k_2 \Lambda / n_2) v_2 \exp([t_0 - i \frac{t_0}{N}] (v_2^{-1} - (k_2 \Lambda / n_2)))\}^{k_1}] \}. \end{aligned}$$

For option S_2 of the redundant service of the m level cluster, the servers of which are structurally redundant at the i -th level ($i = 1, 2, \dots, m$) with multiplicity n_i , the average residence time of the request when creating k_i copies of the request when serving it on the i -th as:

$$T = \sum_{i=1}^m \{ v_i + \int_0^\infty [\frac{\Lambda v_i k_i}{n_i} e^{(\frac{\Lambda k_i}{n_i} - \frac{1}{v_i}) t}]^{k_i} dt \}.$$

6 The results of calculations and comparison of redundant service options

In the calculations, we assume that the number of reserved servers at each level is the same and equal to $n = 8$, the average request execution times for servers of the first and second level are $v_1 = v_2 = 0.4s$, and the total allowable wait time is $t_0 = 0.2s$.

In Fig. 5 shows the dependences of the probabilities of timely service in a two-level cluster on the intensity of the input request stream Λ . In Fig. 1, curve 1 corresponds to non-redundant service $k = 1$, curves 2-3 correspond to service option S_1 with a request reservation ratio of $k = 2, 3$, and curves 4, 5 correspond to option S_2 with a reservation ratio of $k = 2, 3$.

Fig. 6 shows the dependences of the probability of timely service of requests on the multiplicity of their reservation (number of generated copies) k . In Fig. 2, curves 1-3 represent service option S_1 , and curves 4-6 option S_2 for request flow intensities corresponding to $\Lambda = 3.3; 3; 3.5$ 1/s.

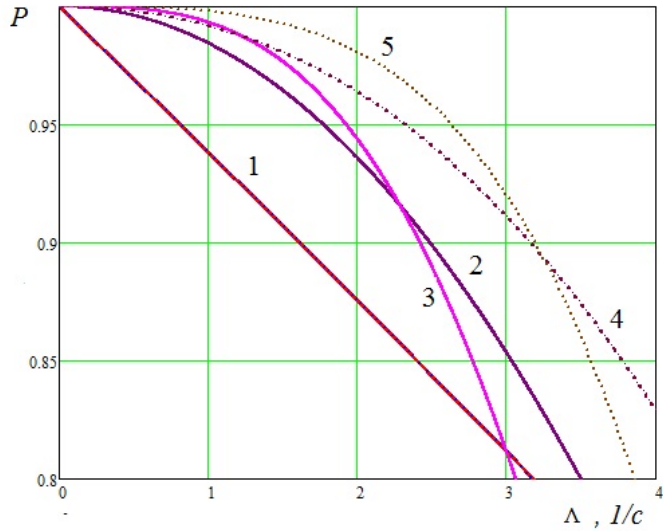


Fig. 5. Dependences of the probabilities of timely service in a two-level cluster on the intensity of the input request stream Λ .

The above graphs allow concluding that there is an optimal multiplicity of redundant service, and the smaller the system load and the permissible total waiting time, the greater the redundancy rate at which the maximum probability of multi-stage service timeliness is achieved. The calculations show that the option of redundant service S_2 allows increasing the probability of timely service requests.

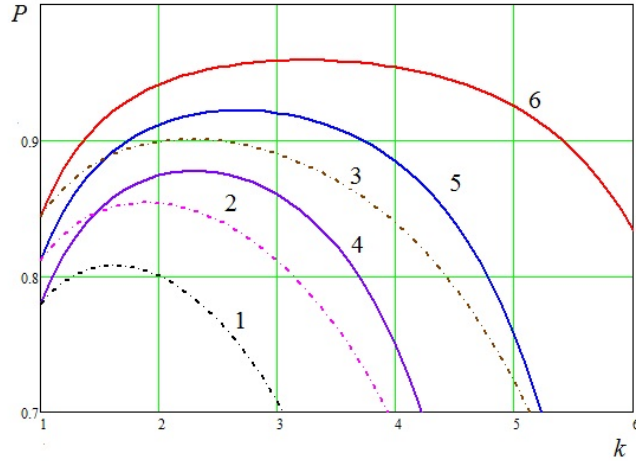


Fig. 6. shows the dependences of the probability of timely service requests on the frequency of their reservation.

7 Conclusion

An analytic model is proposed for multi-level cluster systems involving the execution of queries by a sequence of redundant nodes, and the effectiveness of the options for redundant query servicing critical for queue latency is determined.

The novelty of the proposed model is that it allows you to take into account the requirements of not exceeding the maximum allowable accumulated waiting time for sequential redundant request servicing at all levels of the system, including for the service option when generating replicas of the request distributed to the redundant service in the cluster servers each level, is implemented by a copy of the request made by the servers of the previous level first in time.

The influence of the redundant service multiplicity on the probability of timely query execution, taking into account sequential redundant execution in servers at all levels of the cluster, is analyzed.

The existence of the efficiency domain of redundant request servicing and the optimal frequency of their reservation depending on the system load and restrictions on the allowable waiting time in queues is shown. The efficiency of redundant maintenance with the formation of copies of requests at each cluster level is shown.

The efficiency of redundant maintenance with the formation of a given number of copies of requests and their distribution in the server queue at each level of the cluster is shown.

References

1. Sorin D. Fault Tolerant Computer Architecture. Morgan & Claypool 2009. 103 p.

2. Koren I. Fault tolerant systems. Morgan Kaufmann publications, visit our San Francisco 2009. 378 p.
3. Kutuzov O. I., Tatarnikova T. M. Model of a self-similar traffic generator and evaluation of buffer storage for classical and fractal queuing system. In Moscow Workshop on Electronic and Networking Technologies, MWENT 2018 - Proceedings 1, pp. 1-3 (2018).
4. Tatarnikova, T. M., Dziubenko, I. N. Wireless Sensor Network Clustering Model. In 2018 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF), pp. 14 (2018).
5. Astakhova, T., Verzun, N., Kolbanev, M., Shamin, A. A model for estimating energy consumption seen when nodes of ubiquitous sensor networks communicate information to each other. In Proceedings of the 10th Majorov International Conference on Software Engineering and Computer Systems, Saint Petersburg, Russia, December 20-21 (2018).
6. Astakhova, T. N., Verzun, N. A., Kasatkin, V. V., Kolbanev, M. O., Shamin, A. A. Sensor network connectivity models. *Informatsionno-upravliaiushchie sistemy*, no. 5, pp. 3850 (2019). <https://doi.org/10.31799/1684-8853-2019-5-38-50>.
7. Aliev, T.I.: The synthesis of service discipline in systems with limits. *Communications in Computer and Information Science, IET*, vol. 601, pp. 151-156 (2016).
8. Dudin, A.N., Sun, B.: A multiserver MAP/PH/N system with controlled broadcasting by unreliable servers. *Automatic Control and Computer Sciences*, No. 5, pp. 3244 (2009).
9. Prasenjit Chanak, Tuhina Samanta, Indrajit Banerjee Fault-tolerant multipath routing scheme for energy efficient wireless sensor networks. *International Journal of Wireless & Mobile Networks (IJWMN)* Vol. 5, No.2, April 2013 p 33-45.
10. Kabatiansky G., Krouk E., Semenov S. Error Correcting Coding and Security for Data Networks. Analysis of the Superchannel Concept. Wiley, 2005. 288 .
11. Krouk E., Semenov S. Application of Coding at the Network Transport Level to Decrease the Message Delay // Proc. of 3rd Intern. Symp. on Communication Systems Networks and Digital Signal Processing. Staffordshire University, UK, 2002. P. 109112.
12. Bogatyrev A.V., Bogatyrev S.V., Bogatyrev V.A. Analysis of the Timeliness of Redundant Service in the System of the Parallel-Series Connection of Nodes with Unlimited Queues // 2018 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF) 2018.
13. A. V. Bogatyrev, V. A. Bogatyrev and S. V. Bogatyrev, Multipath Redundant Transmission with Packet Segmentation, 2019 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF), Saint-Petersburg, Russia, 2019, pp. 1-4. doi: 10.1109/WECONF.2019.8840643.
14. V. A. Bogatyrev, S. V. Bogatyrev and A. V. Bogatyrev, Model and Interaction Efficiency of Computer Nodes Based on Transfer Reservation at Multipath Routing, 2019 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF), Saint-Petersburg, Russia, 2019, pp. 1-4. doi: 10.1109/WE-CONF.2019.8840647.
15. Bogatyrev V.A., Parshutina S.A. Redundant Distribution of Requests Through the Network by Transferring Them Over Multiple Paths // *Communications in Computer and Information Science - 2016*, Vol. 601, pp. 199-207.
16. Zakoldaev, D.A., Korobeynikov, A.G., Shukalov, A.V., Zharinov, I.O. Workstations Industry 4.0 for instrument manufacturing // IOP Conf. Series: Materials Science and Engineering 665 (2019) 012015 IOP Publishing doi:10.1088/1757-899X/665/1/012015.

17. Zakoldaev D.A., Korobeynikov A.G., Shukalov A.V., Zharinov I.O. Cyber and physical systems technology classification for production activity of the Industry 4.0 smart factory//IOP Conference Series: Materials Science and Engineering, 2019, Vol. 582, No. 1, pp. 012007.
18. Zakoldaev D.A., Gurjanov A.V., Shukalov A.V., Zharinov I.O. Classification of cyber and physical systems of Industry 4.0//IOP Conference Series: Materials Science and Engineering, 2019, Vol. 582, No. 1, pp. 012008.
19. Bogatyrev V.A. Increasing the fault tolerance of a multi-trunk channel by means of inter-trunk packet forwarding (1999) Automatic Control and Computer Sciences, 33 (2) , pp. 70- 76.
20. Kleinrock, L). Queueing Systems: Volume I Theory. New York: Wiley Interscience. 1975 p. 417. ISBN 978-0471491101.
21. Kleinrock, L. Queueing Systems: Volume II Computer Applications. New York: Wiley Interscience. 1976 p. 576. ISBN 978-0471491118.