

# Multipath Redundant Real-Time Transmission with the Destruction of Expired Packets in Intermediate Nodes

Vladimir Bogatyrev<sup>1,2</sup>, Stanislav Bogatyrev<sup>1,3</sup> and Anatoly Bogatyrev<sup>3</sup>

<sup>1</sup> ITMO University, Kronverksky Pr. 49, bldg. A, Saint-Petersburg, 197101, Russia

<sup>2</sup> Saint-Petersburg State University of Aerospace Instrumentation, 67, Bolshaya Morskaya str. St Petersburg, Russia

<sup>3</sup> JSC NEO Saint Petersburg Competence Center, 1-Ya Sovetskaya, house 6 str. St. Petersburg, Russia

## Abstract

The possibilities of increasing the probability of timely service and reducing the average waiting time for requests for machine-to-machine exchange in distributed computer systems are investigated. Improving the reliability, timeliness and error-free transmission in automated distributed control systems focused on intelligent and cognitive methods of data and image analysis is fundamental in their real-time operation. The effect is achieved as a result of the reserved multipath transfers of packets critical to delays, at which their replication is provided with a task for each replica of the path (route) of sequential passage of network nodes. An analytical model is proposed for estimating the probability of timely delivery and the average total waiting time in the queues of route nodes with reserved and non-reserved packet transmission. The communication nodes that make up the data transmission route are represented by single-channel queuing systems with an infinite queue. The influence of the multiplicity of redundancy (replication) of transmissions on the probability of their timely maintenance is analyzed. The condition for the success of reserved transfers is that the accumulated total waiting in the queues of nodes that make up the path for at least one of the replicas of the packet should not exceed the given maximum allowed time. The efficiency of destroying expired packets in the intermediate nodes that make up the data transmission paths is shown. The existence of an optimal redundancy multiplicity critical to the total delay in the queues of packets is shown.

## Keywords

Fault tolerance, duplicate computer, reliability, redundancy, real time, probability of timely.

## 1. Introduction

The intellectualization of real-time application solutions in automated distributed control systems requires the support of high reliability and fault tolerance [1,2] with low service delays in data transmission and processing systems. The concept of Ultrareliable and Low-Latency Wireless Communication is aimed at solving these problems [3-4]. Improving the reliability, timeliness and error-free transmission in distributed systems focused on the implementation of intelligent and cognitive methods of data and image analysis is fundamental when they operate in real time.

A special feature of real-time image transmission and processing is the limitation of acceptable network delays and transmission delay spread (jitter), which significantly reduce the stability of communication and the quality of image perception, including video streams. Various methods are used to deal with transmission errors, the main ones being noise-tolerant coding (including transport coding) and retransmissions. Retransmissions lead to increased delays, which makes it advisable to use real-time protocols without delivery confirmations. Solutions to these problems are particularly important for real-time cyber-physical systems [5-9].

---

GraphiCon 2021: 31st International Conference on Computer Graphics and Vision, September 27-30, 2021, Nizhny Novgorod, Russia

EMAIL: vabogatyrev@corp.ifmo.ru (V. Bogatyrev); stanislav@nspcc.ru (S. Bogatyrev); anatoly@nspcc.ru (A. Bogatyrev)

ORCID: 0000-0003-0213-0223 (V. Bogatyrev); 0000-0003-0836-8515 (S. Bogatyrev); 0000-0001-5447-7275 (A. Bogatyrev)



© 2021 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

Transport coding [10, 11] allows to reduce the average network latency, which implements message fragmentation with the transmission of encoded fragments in different ways. If a part of the fragments is returned or not delivered, the message can be restored as a result of encoding. At the same time, transport encoding, allowing to reduce retransmissions, complicates the processing of packets on the receiving and transmitting sides, in connection with decoding and restoring the message.

The Real-Time Transport Protocol (RTP) and the Resource Reservation Protocol (RSVP) jointly used with it are aimed at ensuring the timeliness of transfers without delivery confirmations and retransmissions. RSVP allows you to reserve network resources to implement the required quality of real-time transmissions over the RTP protocol [12, 13].

For distributed real-time systems, including image transmission, transmission efficiency is characterized by the average delivery time and the variance of this time. For real-time transmissions, the key quality indicator is the probability of error-free delivery within the maximum allowed time.

The reliability and timeliness of request execution in structurally redundant infocommunication systems can be improved as a result of redundant servicing of packet transfer requests. In the case of redundant service, in the case of a request for the transmission of a packet, its copies (replicas) are formed, delivered to the addressee along different routes (paths). A request to transmit a packet is considered successful if at least one of the generated replicas is delivered to the addressee in a timely and error-free manner. When multiple replicas are delivered, all but one is destroyed and not issued to the addressable subscriber [15,16].

Multi-path redundant service, provides for the replication of requests with the assignment for each of them of the path (route) of the sequential passage of nodes, networks [15-18]. With redundant transmissions, a step-by-step route formation is possible, when for each replica, when it passes through the next communication node, the addresses of the next node that makes up the path are assigned.

The condition for the success of reserved transfers is that the accumulated total waiting in the queues of nodes that make up the path, for at least one of the copies, should not exceed the specified maximum allowed time [15-18].

The organization and models for estimating the probability of timely error-free delivery of packets to the addressee with multipath transmission redundancy during the passage of replicas of transmitted packets of the sequence of switching nodes that make up the path are proposed in [17,18].

Software models [15-18] are focused on transfers in which the intermediate communication nodes do not provide for the elimination of expired replicas of the transmitted packet. It should be noted that non-destroyed replicas that have lost their relevance create unproductive loading on switching nodes, and this leads to additional delays in the delivery of transmitted data.

The destruction of out-of-date replicas, implemented in intermediate switching links along the replica delivery path, can potentially increase the efficiency of multipath canned transmissions. This article is devoted to the study of this possibility.

The purpose of the article is to study the possibilities of increasing the reliability of timely delivery of packets to the addressee as a result of the elimination of irrelevant packets in the intermediate communication nodes that make up the packet delivery paths.

The study of the effectiveness of the destruction of expired replicas includes the construction of models and estimates of the probability of timely delivery of packets to the addressee through the sequence of communication nodes that make up the path, for cases with and without replication of transmitted packets with and without destruction of expired packets.

## **2. Assessing the probability of timely delivery without reserving transfers**

Let's consider a system that provides for the formation of a single path for the sequential passage of the communication nodes of the network when requesting a packet to be sent. In this section, we will assume that no transfer reservations are made.

The communication nodes that make up the data transmission route are represented by single-channel mass service systems with an infinite queue of the M/M/1 type [19, 20].

For the transmission option without destroying expired packets, the intensity of sequential requests to all nodes that make up the path will not change. In this case, expired packets will create an unproductive load on the nodes of the path.

When a packet of nodes composing a path passes sequentially, its waiting time accumulates in the outlines of the passed nodes. As a result, the margin of acceptable waiting time  $t_0$  is reduced on average by the total average waiting time in the queues of passed nodes.

If there are  $n$  possible non-overlapping routes in the network, then with a uniform distribution of the flow of intensity  $\Lambda_s$  requests between routes, the intensity of requests sent along each route will be  $\Lambda = \Lambda_s/n$ .

The probability of not exceeding the allowed waiting time of the packet when passing the  $i$ -th node, taking into account the delays in the queues of the nodes involved in the transmission, is defined as

$$P_i = \left( 1 - \Lambda v_i e^{\left( \Lambda - \frac{1}{v_i} \right) \left( t_0 - a_i \sum_{j=1}^{i-1} w_j \right)} \right) \quad (1)$$

$$a_i = \begin{cases} 1, & \text{if } i > 1, \\ 0, & \text{if } i = 1, \end{cases} \quad (2)$$

where  $v_i$  is the average time of packet transmission through the  $i$ -th node of the path without taking into account waiting in its queue,  $\Lambda$  is the intensity of the input stream of requests for packet transmission, and  $w_j$  is the average time of packet waiting in the queue of the  $j$ -th node

$$w_j = \frac{\Lambda v_j^2}{1 - \Lambda v_j}. \quad (3)$$

The probability of timely delivery when passing all  $m$  nodes that make up the path is defined as

$$P = \prod_{i=1}^m P_i, \quad (4)$$

where  $P_i$  is the conditional probability that the delay in the queue at the  $i$ -th node, taking into account the accumulation of waiting at the previous nodes, will not exceed the total allowable delay in the queues  $t_0$ .

When expired packets are destroyed in intermediate nodes, the intensity of requests to the nodes will decrease as they continue along the route. Destroying expired packets reduces the unproductive load of nodes that make up the packet delivery path to the destination.

The probability of not exceeding the allowed waiting time of the packet during the passage of the  $i$ -th node, taking into account the reduction in the load of nodes on the meta of the packet promotion and the margin of the allowed total waiting time  $t_0$ , is defined as

$$P_i = 1 - \Lambda_i v_i e^{\left( \Lambda_i - \frac{1}{v_i} \right) \left( t_0 - a_i \sum_{j=1}^{i-1} w_j \right)}, \quad (5)$$

where  $a_i$  is calculated by (2).

For  $i=1$ ,  $\Lambda_1 = \Lambda$  and for  $i>1$

$$\Lambda_i = \Lambda \prod_{j=1}^{i-1} P_j, \quad (6)$$

$$w_j = \frac{\Lambda_j v_j^2}{1 - \Lambda_j v_j}. \quad (7)$$

The probability of timely delivery when passing all  $m$  nodes that make up the path is determined by (4) when calculating  $P_i$  by (5).

As an additional indicator of the effectiveness of the systems under study, the average delay  $W$  in queues when passing  $m$  nodes that make up the route can be used. In the case when the destruction of expired packets in intermediate nodes does not occur

$$W = \sum_{j=1}^m \frac{\Lambda v_j^2}{1 - \Lambda v_j}.$$

If the expired packets in the intermediate nodes are destroyed, then

$$W = \sum_{j=1}^m \frac{\Lambda_j v_j^2}{1 - \Lambda_j v_j}$$

An indicator can also be used as a comprehensive performance indicator

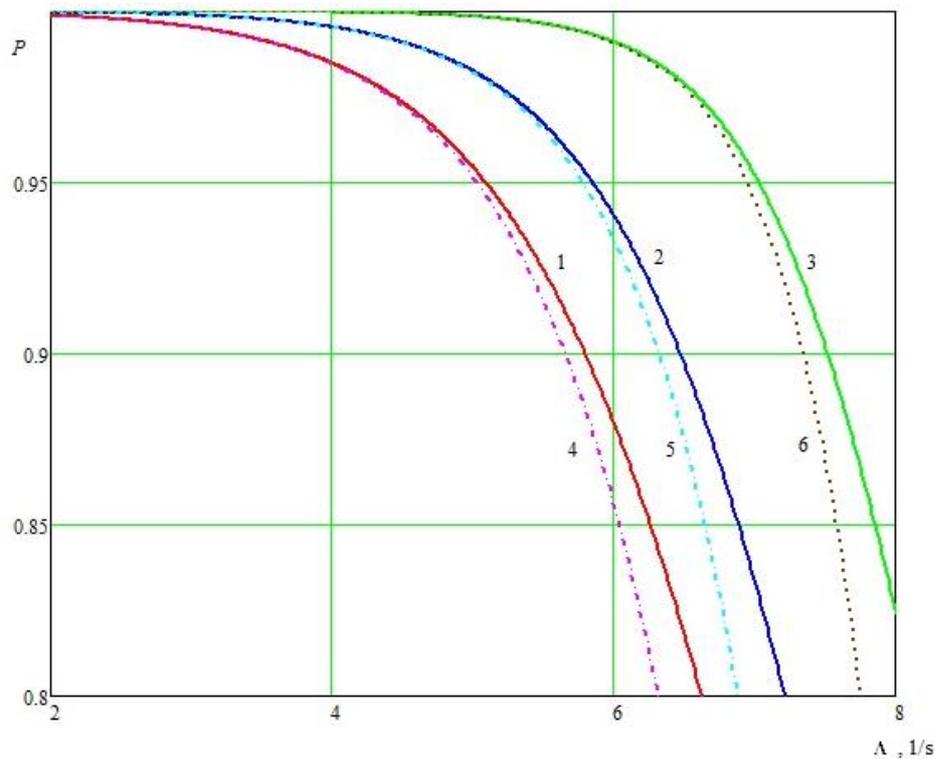
$$Z = P(t_0 - W),$$

expressing the mathematical expectation of the margin until the maximum allowable delay  $t_0$  is reached, the average time of total waiting in the queues of nodes that make up the route for randomly delivered packets.

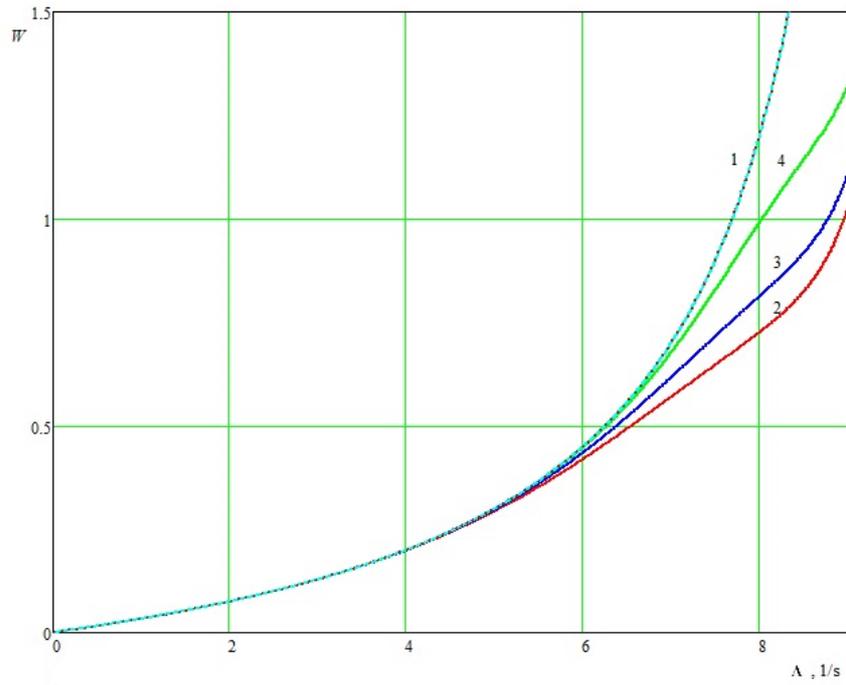
When calculating, we will assume that the path passes through  $m=3$  nodes, the average transmission time through which (without taking into account waiting in queues),  $v_1 = v_2 = v_3 = 0.1$  s, and the maximum allowable total waiting time in the queues of all nodes that make up the path is  $t_0 = 0.8$  s.

Figure 1 shows the dependence of the probability of timely service of requests for the delivery of packets to the destination, and Figure 2 shows the average waiting time for transmitted packets in the queues of the nodes of the path on the intensity of the input stream of requests  $\Lambda$ . In Fig.1, curves 1-3 correspond to the passage of the path with the destruction of expired packages, and curves 4-6 without their destruction at  $t_0=0.8, 1, 1.5$  s. In Fig.2, curves 1 correspond to the total average delay in points when passing the path without destroying expired packets, and curves 2-4 with their destruction at  $t_0=0.8, 1, 1.5$  s.

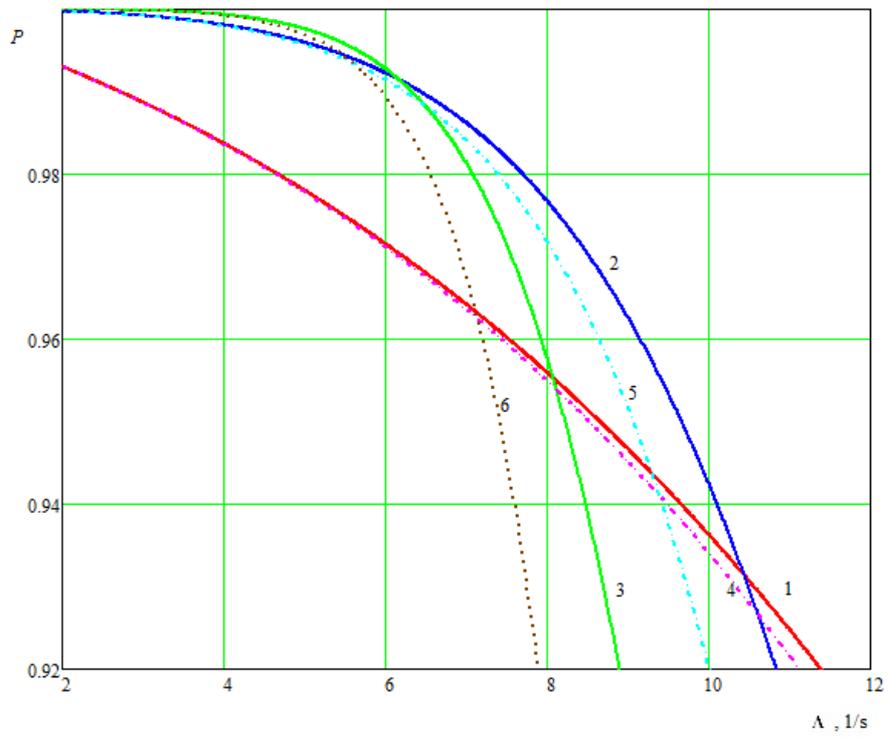
Figure 3 shows the dependence of the complex efficiency indicator  $Z$  on the intensity of the input stream of requests  $\Lambda$ . Curves 1-3 correspond to the passage of the path with the destruction of expired packages, and curves 4-6 without their destruction at  $t_0=0.8, 1, 1.5$  s.



**Figure 1:** The dependence of the probability of timely servicing of requests for packet delivery to the addressee on the intensity of the input stream of requests  $\Lambda$



**Figure 2:** The dependence of the average waiting time at the nodes of the path when delivering a packet to the destination on the intensity of the input stream of requests  $\Lambda$



**Figure 3:** The dependence of the complex efficiency indicator  $Z$  on the intensity of the input stream of requests  $\Lambda$

### 3. Estimates of the probability of timely delivery of packages with redundant transfers

When reserving transfers (with the creation of  $k$  replicas of each transmitted packet), we assume that all paths include  $m$  nodes. At the  $i$ -th stage (section) of the path, one of the  $n_i$  nodes with identical characteristics can be used in its formation, the load of which is assumed to be the same. Thus, the probability of timely delivery of a single replica of a packet when passing through all  $m$  nodes that make up one of the paths, while not destroying expired replicas of packets in intermediate nodes is calculated as

$$P = \prod_{i=1}^m \left( 1 - \frac{\Lambda_s k v_i}{n_i} e^{\left( \frac{\Lambda_s k - 1}{n_i} \right) \left( t_0 - a_i \sum_{j=1}^{i-1} w_j \right)} \right),$$

where

$$w_j = \frac{(\Lambda_s k / n_j) v_j^2}{1 - (\Lambda_s k / n_j) v_j},$$

the probability of timely delivery of a single replica of a packet when passing through all  $m$  nodes that make up one of the paths, when destroying expired replicas of packets in intermediate nodes, is calculated by (4) when finding  $P_i$  according to the formula (5) with substitution for  $i=1$ , and for  $i>1$

$$\Lambda_i = \frac{\Lambda_s k}{n_i} \prod_{j=1}^{i-1} P_j. \quad (8)$$

In the formula (5)  $w_j$  is calculated by the formula (7) in the calculation according to the formula (8).

The probability of not exceeding the allowable time  $t_0$  of walking all nodes in the path, at least one of the  $k$  replicas of the packages, provided that the number of nodes that make up all the way is equal to  $m$ , and their characteristics are the same, calculate as

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

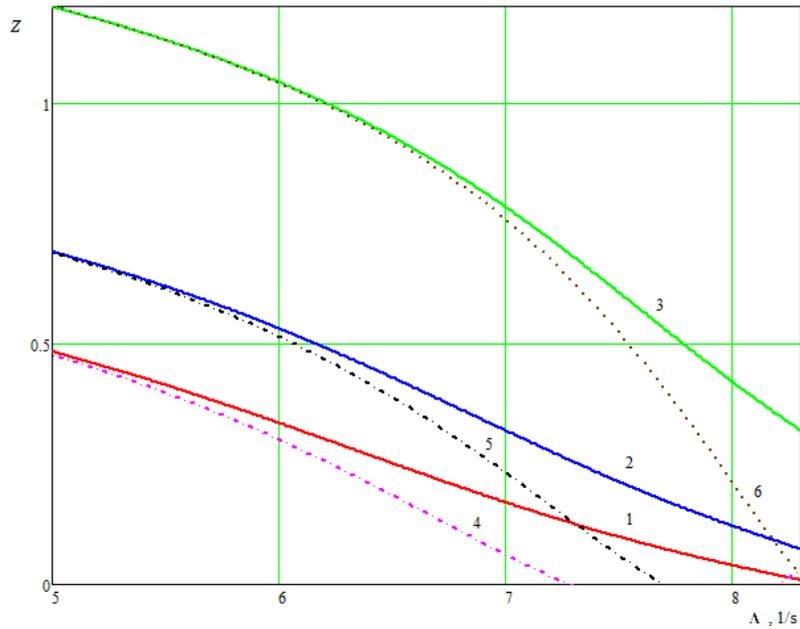
The dependence of the probability of timely reserved service on the intensity of the input stream of requests is shown in Figure 4. Curves 1-3 correspond to the option with the destruction of expired replicas in intermediate nodes, and curves 4-6 correspond to the option without their destruction, respectively, with a redundancy multiplicity of 1, 2, 3. The calculation is performed at  $t_0=0.3$  s. The dependence of the probability of timely reserved service on the multiplicity of reserved requests is shown in Figure 5. Curves 1-3 correspond to the variant with the destruction of expired replicas in intermediate nodes, and curves 4-6 correspond to the variant without their destruction, respectively, at  $t_0=0.8, 1, 1.4$  s. The calculation is performed at  $\Lambda=6$  1/s.

The presented graphs confirm the high efficiency of destroying expired packets in intermediate nodes and the existence of an optimal multiplicity of packet redundancy critical to the total delay in the queues of the transmitting nodes. With a decrease in the system load, the recommended redundancy rate for waiting-critical packets, at which the maximum probability of timely delivery of at least one of the replicas of packets is achieved, increases.

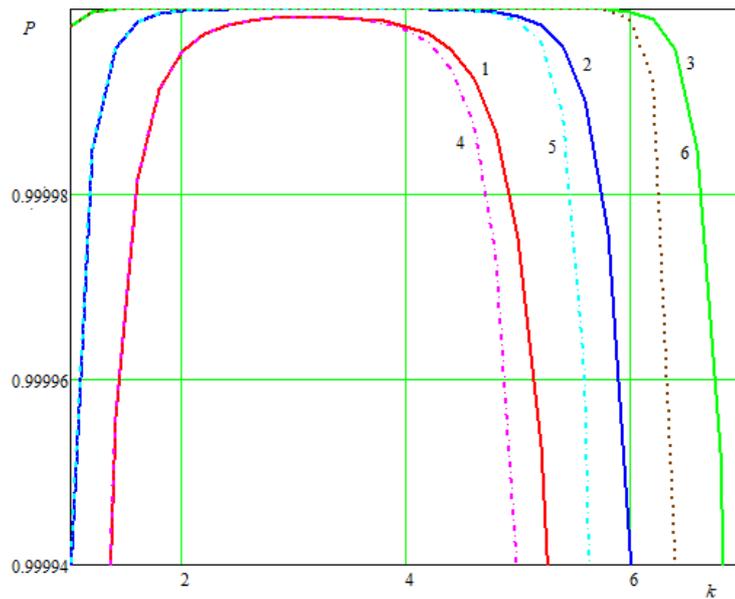
Figure 4 shows the feasibility of adaptive changes in the multiplicity of reservation requests as the intensity of the input stream increases. So, for the example under consideration, when the request intensity is low, the transmitted packets should be reserved with a multiplicity of three, as it increases, with a multiplicity of two, and finally, there is a limit on the request intensity above which it is not advisable to reserve packets.

Figure 5 shows the existence of an optimal redundancy multiplicity for requests critical to the total delay in the queues. Moreover, when the expired requests are destroyed, the effect of reserving transfers increases.

The proposed models and technical solutions for ensuring the reliability and timeliness of multipath transmission of packets critical to queue delays are supposed to be adapted for use within the framework of the concept of Ultrareliable and Low-Latency Wireless Communication [21-23].



**Figure 4:** The dependence of the probability of the timeliness of the reserved service on the intensity of the input stream of requests



**Figure 5:** Dependence of the probability of timely reserved service on the multiplicity of reserved requests

#### 4. Conclusion

For multipath redundant transmissions, an analytical model is proposed for estimating the probability of timely delivery and the average total waiting time in queues before the first replica is delivered along one of the paths for redundant and non-redundant packet transmission.

The efficiency of destroying expired packets in the intermediate nodes that make up the data transmission paths is shown.

For requests critical to delays in queues, the influence of the multiplicity of reservation (replication) requests on the probability of their timely service is analyzed.

It is shown that there is an optimal multiplicity of reserving requests critical to the total delay in queues, taking into account options with and without destroying expired packets in the intermediate nodes that make up the paths.

## 5. References

- [1] I. Koren, Fault tolerant systems. Morgan Kaufmann publications, San Francisco, 2009.
- [2] H. Aysan, Fault-tolerance strategies and probabilistic guarantees for real-time systems Mälardalen University, Västerås, Sweden. 2012. 190 p.
- [3] S. Kim, Y. Choi, Constraint-aware VM placement in heterogeneous computing clusters. *Cluster Comput* 23, 71–85 (2020). doi.org/10.1007/s10586-019-02966-6.
- [4] M. Siddiqi<sup>1</sup>, H. Yu, J. Joung, 5G Ultra-Reliable Low-Latency Communication Implementation Challenges and Operational Issues with IoT Devices *Electronics* 2019, 8, 981; www.mdpi.com/journal/electronics. doi:10.3390/electronics8090981
- [5] D.A. Zakoldaev, A.G. Korobeynikov, A.V. Shukalov, I.O. Zharinov, 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.
- [6] T. N. Astakhova, N. A. Verzun, V. V. Kasatkin, M. O. Kolbanev, A. A. Shamin, Sensor network connectivity models. *Informatsionno-upravliaiushchie sistemy [Information and Control Systems]*, 2019, no. 5, pp. 38–50. doi:10.31799/16848853-2019-5-38-50. <https://doi.org/10.31799/1684-8853-2019-5-38-50>.
- [7] B. Sovetov, T. Tatarnikova, V. Cehanovsky, Detection system for threats of the presence of hazardous substance in the environment. *Proceedings of 2019 22nd International Conference on Soft Computing and Measurements, SCM 2019* (2019) 121-124. doi: 10.1109/SCM.2019.8903771.
- [8] D.A. Zakoldaev, A.G. Korobeynikov, A.V. Shukalov, I.O. Zharinov, O.O. Zharinov, Industry 4.0 vs Industry 3.0: the role of personnel in production//*IOP Conference Series: Materials Science and Engineering*, 2020, Vol. 734, No. 1, pp. 012048.
- [9] S.B. Ya, T.M. Tatarnikova, E.D. Poymanova, Organization of multi-level data storage (2019) *Informatsionno-Upravliaiushchie Sistemy*, 2019 (2), pp. 68-75. doi: 10.31799/1684-8853-2019-2-68-75.
- [10] E. Krouk, S.Semenov, 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. pp. 109—112
- [11] G.Kabatiansky, E. Krouk, S.Semenov, Error Correcting Coding and Security for Data Networks. *Analysis of the Superchannel Concept*. Wiley, 2005. 288 p
- [12] C. Perkins, RTP—Addison- Wesley, 2003. — P. 414. — ISBN 9780672322495.
- [13] R. Zurawski,. RTP, RTCP and RTSP protocols // *The industrial information technology handbook* CRC Press, 2004. — P. 28—70. — ISBN 9780849319
- [14] V.A. Bogatyrev, A.V. Bogatyrev, S.V. Bogatyrev, Redundant Servicing of a Flow of Heterogeneous Requests Critical to the Total Waiting Time During the Multi-path Passage of a Sequence of Info-Communication Nodes. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2020. Vol. 12563. pp. 100-112. doi 10.1007/978-3-030-66471-89.
- [15] V.A. Bogatyrev, S.V. Bogatyrev, A.N. Derkach, Timeliness of the Reserved Maintenance by Duplicated Computers of Heterogeneous Delay-Critical Stream. *CEUR Workshop Proceedings*. 2019. Vol. 2522. pp. 26-36.
- [16] V.A. Bogatyrev, S.V. Bogatyrev, A.V. Bogatyrev, Redundant multi-path service of a flow heterogeneous in delay criticality with defined node passage paths. *Journal of Physics: Conference Series, Volume 1864, 13th Multiconference on Control Problems (MCCP 2020) 6-8 October 2020, Saint Petersburg, Russia 2021* J. Phys.: Conf. Ser. 1864 012094 - 2021, Vol. 1864, 012094, No. 1, pp. 012094. doi 10.1088/1742-6596/1864/1/012094.

- [17] V.A. Bogatyrev, S.V. Bogatyrev, A.V. Bogatyrev, Inter-machine exchange of real time in distributed computer systems. CEUR Workshop Proceedings. - 2020. - Vol. 2744.
- [18] V.A. Bogatyrev, A.V. Bogatyrev, S.V. Bogatyrev, The probability of timeliness of a fully connected exchange in a redundant real-time communication system. Wave Electronics and its Application in Information and Telecommunication Systems (WECONF 2020). - 2020. - № - Access mode: <https://ieeexplore.ieee.org/document/9131517>. doi: 10.1109/WECONF48837.2020.9131517.
- [19] L. Kleinrock, Queueing Systems: Volume I. Theory. New York: Wiley Interscience. 1975 p. 417. ISBN 978-0471491101.
- [20] L. Kleinrock, Queueing Systems: Volume II. Computer Applications. New York: Wiley Interscience. 1976 p. 576. ISBN 978-0471491118.
- [21] H.Ji, S. Park, J. Yeo, Y. Kim, J. Lee, B. Shim, Ultra-Reliable and Low-Latency Communications in 5G Downlink: Physical Layer Aspects. *IEEE Wirel. Commun.* 2018, 25, 124–130.
- [22] J. Sachs, G. Wikström, T. Dudda, R. Baldemair.; K. Kittichokechai, 5G Radio Network Design for Ultra-Reliable Low-Latency Communication. *IEEE Netw.* 2018, 32, 24–31. doi:10.1109/MNET.2018.1700232.
- [23] M. Bennis, M Debbah, Poor, H.V. Ultrareliable and Low-Latency Wireless Communication: Tail, Risk and Scale. *Proc. IEEE* 2018, 106, 1834–1853. doi: 10.1109/JPROC.2018.2867029.