Modern problems of information systems and data networks: choice of network equipment, monitoring and detecting deviations and faults

N R Storozhenko, A I Goleva, D A Tunkov, V I Potapov
Omsk State Technical University, 11 Mira ave., Omsk, Russia

E-mail: frybkf07.93@mail.ru

Abstract. The rapid development, complexity of modern information systems and the growth of qualitative and quantitative characteristics of network anomalies and failures makes the task of network reliability more relevant. So, the aim of the article is to research modern problems of information systems and develop methods for its solution. For resolving of the considered problems in order to achieve the required level of network characteristics, the article contains developed the conditions for the optimal choice of network equipment, the methods for determining threshold values for detecting deviations from the normal operation of the network, the schemes of the algorithm for monitoring faults and problems.

1. Introduction
The rapid development of information systems and the increase in cost of software, server platforms and network components in recent years have particularly increased interest in the problems of reliability, efficient use of network resources and operational access to them. These problems are especially acute in large computer systems of research centers, industrial enterprises and defensive complex objects, when performance, reliability, query processing time and stability of network operation at peak loads do not satisfy the requirements of business processes, and do not justify large investments in information technology (IT).

Modern information systems should meet the requirements of reliability and accessibility for each registered user [1,2]. However, in the conditions of continuous dynamic due to changes in network structures, occurs a problem of network expansion, obsolescence of methods for protection and fault tolerance. So, there is a need to improve the network and methods for control its state, as well as the speed of reaction to changes in the network.

Operational, timely troubleshooting of the information system and the implementation of rational recovery procedures is a major task from a practical and economic point of view. In addition, the effective management troubleshooting is scientific interest due to the necessity of creating a plurality of models describing the functioning of information system components, development of algorithms for detection and localization of faults, designing a structure of the subsystem for management troubleshooting.

1.1. Literature Review
Research and analysis of problems in the field of information technology and data networks allows to identify priority direction for their resolution. A large number of works are devoted to the problems of security and fault tolerance of network resources.
Works [3-6] contain a description for calculation of basic reliability arguments (without taking into account modern problems in the field of information technologies) or include only a description and decision of some specific problem [1,7-10]. Works [11-15] describe the need for the application systems of monitoring, control and diagnostics for information systems, and contain the basic principles of their work without mathematical description.

The work [16] deals with monitoring problems, but only simple threshold schemes without hysteresis are involved in decision-making. The decision-making schemes proposed in [17] and tested in [18] use hysteresis mechanisms which allow to reduce the number of the decision schemes triggering, but the task of determining the values of the established thresholds is not considered.

A literature review on the topic of improving the reliability and regularity of computer networks and their resources showed that today, despite the great attention and increasing costs to ensure the reliability of information systems, there are a number of problems in this field of research.

According to Positive Technologies research, the number of incidents (related to attacks and network equipment failures) increased by 11% in the first half of 2019 compared to 2018 [19]. Based on this, the information systems of organizations in 2019 did not become less vulnerable to failures.

Statistics of network equipment failures show that the reliability of software and hardware does not meet the required level and does not respond the needs of modern information systems.

For this reason, this article based on an attempt to eliminate these shortcomings by conducting research of contemporary problems in this field of study in order to develop effective methods to solve them. The purpose of the article is to improve the efficiency of troubleshooting in network systems, including the development of methods and means for fault detection; in particular, by making recommendations for the optimal choice of network equipment, determining threshold values for detection deviations from the normal operation of the network, developing the algorithm for detection of modern faults and problems.

2. Peculiar properties of interaction modern networks resources

Figure 1 shows the order of network resources interaction during processing of user request in typical modern information system.

At receipt a request to access a network resource, access rights are checked on the security server. Then the domain controller will provide resources in accordance with the accounts and group policy.

Then, depending on the type of request, control is passed either to the mail server (if necessary for forwarding messages) or file server (if you want to work with files) or application server (if necessary, working with engineering files).
For the effective operation of these network services, it is necessary to ensure the reliability and uninterrupted functioning, because on the configuration and functioning of the network depends the performance of many quality-critical services of the enterprise, such as electronic commerce, videoconferencing, business-process management.

3. Modern problems of information systems

Analysis of the literature sources and author's experience allows to make a conclusion that today there are the following problems in this field of research.

The economic crisis forced organizations to limit the budget and tightly control costs for information technology. Now IT-department should be guided by economic factors, not technical rationality. It makes fundamentally changes the overall pattern of IT-management.

Poor level of mindfulness of the management segment of the company in the need to use tools to ensure not only the confidentiality of data transmitted in the network, but also the efficiency and continuity of uninterrupted its operations.

The choice from today's variety of diagnostic and protection tools for a particular information system should be based on the distinctive features and characteristics of this system. But this point is often not taken into account in building a system to ensure the security and reliability of data transmission. For this reason many solutions are not effective.

Now there are many users with different access rights, the dynamic nature of their presence in an information system, the necessity for expand the system and add new nodes, division of shared resources, the use of a variety of software and hardware in modern networks. Over the years, IT-organizations have accumulated large nomenclatures of complex applications that need to be preserved. Typically, these applications use outdated protocols and platforms. Administrators are faced with a difficult question: it is necessary either to implement the software into the updated network structure on a new platform (however, not always its may be realized due to limitations by software engineers) or to purchase/develop a new software (but may be a problem in training users to work with the functions of the new application).

The increasing number of IT-services for business technology automation, complication of distributed applications, the increasing number of network infrastructure components, the number and complexity of information and communication technologies that need to be supported, – lead to a decrease in the efficiency of IT-organizations and increase the economic and human costs of serving IT.

The rapid increase in the number of users who have access to network resources and data arrays requires not only a flexible system of administration, but also active, universal tools to control software, hardware and user actions in a network.

To date, vulnerabilities in software and network platforms make it possible to implement threats to information system resources, which, in turn, leads to failures of network equipment.

Many specialists do not understand the fundamental difference between incident management and problem management. Incidents are a manifestation of all sorts of non-standard situations, and problems can be caused by fundamental infrastructure deficiencies, which can lead to mass registration of incidents.
Therefore, a thorough search for the reasons of incidents can help to eliminate defects of designing or integration IT-components.

In order to eliminate existing vulnerabilities and possible network failures, it is extremely important to decision the above problems in the aim to improve the efficiency, reliability, security and fault tolerance of data network nodes and recourses.

4. Ways of solving problems in data networks

Modern information systems and networks should have the following characteristics (Figure 3).

4.1. Reasonable choice of communication equipment

In order to ensure the ability of network scaling and adding new nodes, it is necessary to make a well-founded choice of network equipment.

The performance of the communication device should not be less than aggregate intensity of the transmitted traffic:

\[ B \geq \sum P_{ij} \]  \hspace{1cm} (1)

where \( B \) – it is a performance of the switch,
\( P_{ij} \) – is the average traffic intensity from the i-th port to the j-th; the sum is taken for all ports of the device, from 1 to n.

If condition (1) is not met, the device will not cope with incoming traffic without loss and delays due to slow frame processing.
The total performance of the switch is provided by a sufficiently high performance of each of its individual elements: the port process, the switching matrix, the common bus connecting the modules, etc.

The maximum performance of each port must exceed the average intensity of the aggregate traffic passing through the port:

\[ C_k \geq \sum_j P_{kj} + \sum_i P_{ik} \]  

where \( C_k \) is a maximal performance of the k-th port (if the k-th port supports Ethernet, then \( C_k = 10 \) Mb/sec), \( \sum_j P_{kj} \) is an intensity of traffic leaving the port, \( \sum_i P_{ik} \) is an intensity of incoming traffic.

Condition (2) expects that the device operates in half-duplex mode, in the case of full-duplex mode, the parameter must be doubled.

So, these conditions are necessary for the switch to cope with the task on average and not lose frames constantly. If one of the above conditions is not met, the frame loss becomes not an episodic phenomenon at peak traffic values, but a constant manifestation, because even the average traffic values exceed the capabilities of the switch.

4.2 Development of methods and means for fault detection

Further in order to ensure fault-tolerant, reliable and uninterrupted operation, modern information systems need continuous monitoring and diagnostics [6], whereas the work of researchers, production halls, designers and technicians is impossible without a well-functioning communication system, without the possibility of instant data exchange, without diagnostic systems and fault management.

To the most effective method of network analysis can be attributed the system of monitoring, control and diagnostics the parameters of information systems.

The developed scheme (Figure 4) describes the process of monitoring and diagnosing network parameters.

The packet capture module receives data from network equipment and the filter module allows to discard packets that should be ignored by the system [20]. Then monitored packets get into the anomaly detection module. In general, the operation of this module is based on the control of exceeding the threshold value. If the local characteristics differ significantly from the threshold, it is abnormal behaviour of the packet flow. And the response module provides timely notification of the behaviour of network devices.

![Figure 4. Figure shows the operation of the network monitoring and diagnostics system](image-url)
Complication of modern networks and a large number of IT-services require the control of a large number of different parameters. Moreover, the network is used not only to transmit the values of these parameters, but also to transmit user information. In this case, the transfer of a large number of various parameters creates a large load on the IT-system. Therefore, reducing the amount of control information transmitted is an important task.

4.3 Determining threshold values for detection network problems
All types of fault detection systems determine the state of functioning of the network elements $S_i(t)$, then these values are compared with the threshold values $L_i$, $i$ – the number of parameters for determining the state of the elements.

When the value $S_i(t)$ exceeds the threshold $L_i$, the value of the discrete signal $D_i(t)$ changes at the output of the comparator (which performs the comparison of values), and a failure message is transmitted to the fault management subsystem.

When using thresholds without hysteresis (Figure 3(a)) fluctuations of the $S_i(t)$ in the zone of the threshold $L_i$ cause a large number of comparator trips and the values of $D_i(t)$ can change very often. Each change of $S_i(t)$ causes the need to transmit information about the new state for generation the control decisions. It can lead to software and network overload during transmission of service traffic.

![Figure 5. Using threshold levels without hysteresis (a) and with hysteresis (b)](image)

In the case of three-threshold schemes, the value of each threshold $L_i$ is converted into two values $L_i^{(+)}$ and $L_i^{(-)}$. And the solving scheme works in such a way that the signal $D_i(t)$ takes the value of a logical 1 when the value $S_i(t)$ is exceeded of the threshold $L_i^{(+)}$ and becomes equal to logical 0 when the value $S_i(t)$ is less than the threshold $L_i^{(-)}$. If the condition $L_i^{(-)} \leq S_i(t) \leq L_i^{(+)}$ is satisfied, the signal $D_i(t)$ retains the previous value.

Therefore, to make a decision on the state of a network element it is reasonable to use thresholds with the hysteresis property (Figure 3(b)), that allows to significantly reduce the number of trips of the solving scheme. Because the signal of the change in the state of the element is generated only after the change in the state is reliably recorded.

When using thresholds with the hysteresis property, the task of determining the values of $L_i^{(+)}$ and $L_i^{(-)}$ arises.

Despite the fact that the main purpose of use the three-threshold circuits for decision-making is to minimize the triggering of deciding circuits, nevertheless, a brief performance degradation should not be ignored. It is in practice difficult to implement. To fix any (not always informative and necessary) deviations, it is easier and more reliable to use single-threshold schemes, which are used in most network monitoring systems.
Given the fact that the system for managing faults solves a large number of diverse tasks to detect and eliminate different faults in various hardware and software network elements, it is difficult to create a universal method for determining the values of the $L_i$, $L_i^{(+)}$ and $L_i^{(-)}$ thresholds.

However, to assess the quality of the application, it is possible to do the following. Control tasks are periodically run on workstations and the time of their execution is estimated. After accumulation of statistics, the parameters of normal operation and values of thresholds $L_i$, $L_i^{(+)}$ and $L_i^{(-)}$ are defined.

The authors propose several variants for constructing solving schemes.

For demonstration, the change of $S_i$ values during the execution of control tasks on the server, the graph in Figure 6(a) is presented.

![Figure 6](image)

**Figure 6.** Changing the $S_i$ values during the execution tasks on the server (a), applying a threshold level with hysteresis and the output signal (b)

In the first case, for a long time $T$ with a period $\Delta t$, $M=T/\Delta t$ test tasks are performed, and the threshold value of the $i$-th parameter is determined as follows:

$$L_i = \frac{\sum_{m=1}^{M} S_{i,m}}{M}$$  \hspace{1cm} (3)

where $M$ – is a number of test tasks during the accumulation of statistics for the parameter $i$,

$S_{i,m}$ – is a result of the test task execution to control of the $i$-th parameter at a time $t_m = m\Delta t$.

For example, the execution time of a control task, running on a server from a remote workstation.

The value of the upper threshold $L_i^{(+)}$ for each $i$-th parameter is defined as follows:

$$L_i^{(+)} = \frac{\sum_{m=1}^{M} B_{i,m}^{(+)} S_{i,m}}{\sum_{m=1}^{M} B_{i,m}^{(+)}},$$

$$B_{i,m}^{(+)} = \begin{cases} 1 & \text{if } S_{i,m} > L_i^{+} \\ 0 & \text{else} \end{cases}$$  \hspace{1cm} (4)

The lower threshold value is determined similarly:
\[ L_{i}^{\pm} = \frac{\sum_{n=1}^{m} B_{i,n}^{\pm} S_{i,m}}{\sum_{n=1}^{m} B_{i,n}^{\pm}}, \]
\[ B_{i,n}^{\pm} = \begin{cases} 1 & \text{if } S_{i,m} \leq L_{i}^{\pm} \\ 0 & \text{else} \end{cases} \]

For the figure 7(a) normalized threshold values are: \( L_{i} = 0.7, L_{i}^{(+) = 0.79, L_{i}^{(-) = 0.49}. \)

In this case, the signal at the output of the solving circuit will have the form shown in figure 6(b).

The second method of determining the threshold values is based on the approximation of the initial data by a linear function. For this purpose, the period of time during which the \( S_{i} \) parameter was measured is divided into intervals \( \Delta T = R \Delta t \). And \( R \) is selected so that the interval \( \Delta T \) corresponds to a period in tens of minutes or hours and can be used to display and account for the daily workload of the IT-system components. At each interval \( \Delta T \), a linear function is determined, for example, by the least squares method [4], when the coefficients \( b_{0} \) and \( b_{1} \) are first calculated:

\[ b_{0} = \frac{2(2R + 1) \sum r S_{i,r} - 6 \sum r S_{i,r}}{R(R - 1)}, \]
\[ b_{1} = \frac{12 \sum r S_{i,r} - 6(R + 1) \sum r S_{i,r}}{\Delta t(R(R - 1)(R + 1))}. \]

An example of converting the source data (Figure 6(a)) when dividing the time interval into nine intervals is shown in figure 7(a).

Further, the thresholds \( L_{i} = 0.7, L_{i}^{(+) = 0.79, L_{i}^{(-) = 0.49}. \) are determined by the formulas (3), (4), (5), with the difference that the minimum and maximum values are taken from the graph in figure 7(a), and not from the array of accumulated statistical data on the value of the parameter \( S_{i} \), as in the first case.

For the figure 7(a) the normalized threshold are \( L_{i} = 0.58, L_{i}^{(+) = 0.72, L_{i}^{(-) = 0.45}. \) The signal at the output of the solving circuit for this case is shown in figure 7(b). In this case, the number of the solving scheme triggering is much smaller, and the triggering occurred as a reaction to the apparent failure.
Selecting of the method for threshold values is determined by specific of the task. Thus, to identify more deviations from normal operation the first method of determining the values of three-threshold schemes can be used, to identify only significant and long deviations the second variant can be used.

Thus for the purpose of development the methods of control and diagnostics of network resources the algorithmic diagram of monitoring the network parameters (Figure 8).

The algorithm determines the intensity of polling network devices, polls equipment, analyses of received packets. Then, at the stage of collecting parameters, local characteristics of network nodes are determined, their accumulation and transfer to the analyser are carried out. And the obtained values are compared with the thresholds.

![Figure 8. Operation of a network monitoring and diagnostics system](image)

When errors are detected, the system must provide a variety of response options: from notifying the administrator and issuing an alarm to the network management console to automatically troubleshooting by restarting devices.

5. Conclusion

At present in the conditions of growth of qualitative and quantitative characteristics of network failures in the operation of information systems equipment, as well as various incidents, there is a need to consider the modern problems and methods to improve existing fault tolerance systems in this field of research.

Because despite the fact that the software market supplies a considerable number of network management tools, according to international statistics, the number of failures and levels of data loss and device performance are not reduced.

The article describes current problems that can lead to negative effects up to the complete shutdown of critical services and programs, such as e-mail, IP-telephony, video conferencing, communication with banks.

Thus, to solve modern problems, the article includes the conditions for a choice of network equipment, the methods for determining thresholds for assessing the quality of equipment, as well as designed schemes for the development of control and diagnostics of equipment.
With further research, the algorithm can be realized, which will optimize the monitoring process and implement the ability to control the behaviour of the system in accordance with the condition of using the decisive three-threshold schemes.

6. References

[1] Potapov V I, Shafoeva O P, Doroshenko M S, Chervenchuk I V and Gritsay A S 2018 Numerically-analytical solution of problem gaming confrontation hardware-redundant dynamic system with the enemy operating in conditions of incomplete information about the behavior of participants in the game Journal of Physics: Conference Series 1546 p 012062

[2] Lukyanov V S, Gayevo V V, Al-Hadsha F A H 2013 Evaluation of reliability indicators of networks Bulletin of computer and information technologies 8, pp 47–52

[3] Ventzel E S 2018 Operations research: tasks, principles, methodology (Moscow: Justice) p 192

[4] Ushakov I A 1985 Reliability of technical systems (Moscow: Radio and communication) p 608

[5] Mozhaev A S 2003 Universal graphical-analytical method, algorithm and software module for the construction of monotonic and non-monotonic logical functions of system efficiency Proc. Int. Scientific. Schools (Saint-Petersburg: SUAI) pp 101–110

[6] Achilov R 2013 Building secure enterprise networks (Moscow: Science and technology) p 250

[7] Ge J, Zhang B, Fang Y 2016 Research on the Resource Monitoring Model Under Cloud Computing Environment Web Information Systems and Mining. WISM 2010. Lecture Notes in Computer Science (Berlin: Springer) 6318, pp 111–118

[8] Wolters D, Gerth C, Engels G 2017 Visual Requirements Modeling for Cross-Device Systems Computer Science and Information Systems vol 14(2), pp 517–536

[9] Storozenko N R, Goleva A I, Potapov V I, Pastuhova E I, Tunkov D A, Ralovec M L, Fakhrutdinov A R 2019 The mathematical model of monitoring information system with failure prediction Journal of Physics: Conference Series 1260, p 022010

[10] Montes J, Sanchez A, Memishi B, Perez M, Antonii G 2013 GMonE: A complete approach to cloud monitoring Future Generation Computer Systems 29(8), pp 2026–2040

[11] Fatema K, Emeakaroha V, Healy P, Morrison J, Lynn T 2014 A survey of cloud monitoring tools: taxonomy, capabilities and objectives Journal of Parallel and Distributed Computing 74 (10), pp 2918–2933

[12] Aceto G, Botta A, Donato W. Pescape A 2013 Cloud monitoring: a survey Computer Networks 57(9), pp 2093–2115

[13] Averchenkov V I 2011 Monitoring and system analysis of information on the Internet (Moscow: Flinta) p 160

[14] Baklanov A I 2014 Surveillance and monitoring systems (Moscow: Binom) p 234

[15] Lee S, Levanti K and Kim H S 2014 Network monitoring: Present and future Computer Networks 65 pp 84–98

[16] Cormode G, Muthukrishnan S, Yi K 2008 Algorithms for Distributed Functional Monitoring Symp. on Discrete algorithms (San Francisco, California). pp 1076–1085

[17] Wuhib F, Dam M, Stadler R, Clemm A 2005 Decentralized computation of threshold crossing alerts Ambient Networks DSOM 2005. Lecture Notes in Computer Science (Heidelberg: Springer) 3775, pp 220–232

[18] Wuhib F, Stadler R, Clemm C 2006 Decentralized service-level monitoring using network threshold crossing alerts IEEE Communications Magazine 44(10), pp 70–76

[19] Current cyber threats: II quarter 2019, https://www.ptsecurity.com/ru/ru/research/analytics/cybersecurity-threatscape-2019-q2/#id1, last accessed 2019/10/21.

[20] Kufel L 2013 Security Event Monitoring in a Distributed Systems Environment IEEE Security and Privacy 11 (Washington: IEEE Computer Society) pp 36–43