Criteria of reliability of computer networks and assessment quality of service

https://doi.org/10.31713/MCIT.2023.088

Besik Beridze
Department of Computer Engineering (Doctoral Study)
Georgian Technical University
Tbilisi, Georgia
b.beridze@bsu.edu.ge

Ibraim Didmanidze
Department of Languages and Information Technologies
Batumi Shota Rustaveli State University
Batumi, Georgia
ibraim.didmanidze@bsu.edu.ge

Mikheil Donadze
Department of Computer Sciences
Batumi Shota Rustaveli State University
Batumi, Georgia
mikheil.donadze@bsu.edu.ge

Vladimir Zaslavsky
Taras Shevchenko National University of Kyiv
Kyiv, Ukraine
zas@unicyb.kiev.ua

Abstract—**A modern computer network counts tens of millions of servers and hundreds of millions of workstations. Network technologies have found widespread use in industry and business, so network reliability has become one of the most pressing problems.**

In general, the reliability of all systems is determined by the reliability of their constituent elements. Different elements have different reliability. In many cases, reliability and reliability distributions are determined empirically.

If we define the reliability of the system as the possibility of communication between different nodes and workstations, then the reliability of the Internet turns out to be equal to the whole. Unless your workstation is among thousands of nodes, you cannot guarantee the absolute reliability of the network, so it is clear that both definitions are completely unacceptable.

The selection of criteria for determining the reliability and performance of information computer networks is one of the main problems of information systems.

Keywords—computer networks, effective performance criteria, network reliability.

I. Introduction

Today, the organization of information systems security regimes is a critically important strategic factor in the development of any foreign or local company. At the same time, as a rule, the main attention is paid to the requirements and recommendations of the relevant regulatory and methodological framework in the field of information protection. At the same time, many leading companies today, in order to maintain business continuity, use additional initiatives aimed at the sustainability and stability of corporate information systems.

In the process of organizing network resource protection, the concept of a systematic approach is often found in information sources. The concept of systematicity is not only the creation of appropriate protection mechanisms, it is a regular process that is carried out at all stages of the life cycle of an information system. At the same time, all means, methods and measures used to protect information are combined into a single, complete mechanism - the protection system. Unfortunately, the need for a systematic approach to information technology security issues has not yet found proper understanding among users of modern information systems [1].

II. Determine the degree of network reliability

The degree of reliability of a computer network is either a defined, random event probability or a random variation of expected values. This depends on the network structure, bandwidth, associated E members of the network graph, and corresponding event probabilities.

Performance quality evaluates network reliability based on various network performance criteria. As a result of the study, several performance criteria were considered. For example, networks use time-to-order delays (RTTs) for packet-switched packets. Such criteria can be considered random variables, so that they do not depend on the set of working nodes of the graph.

In order to determine the reliability criteria of computer networks and, accordingly, to increase the efficiency of local and global networks, it is desirable to solve the following tasks:

1. Establish criteria for effective network operation. Performance and reliability appear as such criteria in many cases. Which in turn need to select specific qualitative indicators. For example, such as reaction time and readiness ratio.
2. Determine the number of variable parameters. (plurality). Which directly or indirectly affect the performance criteria. All such parameters can be grouped in different ways. For example, parameters of a specific protocol (the maximum frame size of the Ethernet protocol or the frame size of unconfirmed packets in the TCP protocol) or device settings. (Addressing table size or bogie filtering speed, router's internal salt bandwidth).

3. Determine the sensitivity limit for the values of the efficiency criterion.

The network performance can be evaluated logically from "working" - "not working", then the optimization proceeds to diagnose the damage in the network, so that the network can accept any working state.

Another form of evaluation is fine tuning of the network, where the characteristic parameters of the working network are changed to increase the performance of the network by at least a few percent. As a rule, network optimization can be imagined as a kind of transitional option, during the selection of which it is necessary to find such a value for the parameters that will significantly improve the performance indicators. For example, a request to a server (without data) should take 3 seconds instead of 10 seconds, and a file transfer to a remote computer should take 2 minutes instead of 2 minutes. But 30 seconds. In this way, we can choose 3 different forms of optimization tasks:

1. Bring the network to working condition. It involves searching for corrupted elements in the network. Wires, addresses, adapters, computers, device and software compatibility is checked here. Selection of correct values for the main parameters, which ensure message delivery to all nodes of the network, stacking of frame types and protocols, etc.
2. Coarse configuration – configuration of parameters that dramatically affect network characteristics. If the network is functional, but the data exchange is very slow, i.e. the delay is tens of seconds or minutes, or the connection between each other is often interrupted for an unknown reason, such a network can be called conditionally functional. It requires rough stacking. At this stage, it is necessary to find the reason for the delay in packets moving in the network.
3. Failure of one of the network elements or incorrect parameter setting is considered to be the cause of serious latency or variable network performance. If the network is large, it takes a long time to eliminate it, because the number of possible options is quite large. During normal network (busy) operation, the server's reaction to the user's request should not exceed 5 seconds.
4. Accurate setting of network parameters (optimization). If the network is working satisfactorily, then its performance and reliability cannot be improved by changing any parameter alone. In this case, in order to improve the quality of the network, it is necessary to find a successful agreement of its characteristics with various parameters.

It is not possible to achieve an optimal agreement between the parameters during detailed network stacking (in a pure mathematical sense), nor is it necessary to spend colossal efforts to achieve strict optimization, it is enough to find an agreement close to the optimal one, and the task of network optimization can be considered solved [1, 4, 9].

Such decisions are called rational options, and in practice, finding this decision is the main condition for network administrators.

Network fault detection is a trade-off between the two tasks of analysis (measurement, diagnosis, and localization of faults) and synthesis (deciding what changes need to be made to eliminate the fault).

Analysis – determining the value of the efficiency criterion. During the given agreement of the system parameters. At this stage, a monitoring sub-stage is allocated, which represents the procedure of collecting the primary data of network operation. Such data can be considered the statistics of various protocols and the number of frames circulating in the network. Port status of switches, routers and switches. The next step in analysis is the comparison of the data obtained as a result of monitoring with the previously obtained data. The task of analysis requires the active participation of a person, his highly intelligent data, as well as the use of such complex tools as expert systems, which are compiled as a result of the practical experience of network specialists.

Synthesis – choosing values for variable parameters in a network under the condition that efficiency is the best. If a threshold value is given as an indicator of efficiency, then the synthesis result should be a network variant that exceeds the given threshold. Bringing the network to a normal state is also one of the synthesis options. Often, the choice of device or model type is not technical in nature. It is determined either by commercial or general network development policies, so in some cases it is impossible to formulate an optimization task.

Effective network performance criteria are divided into two groups. They characterize network performance and reliability. Performance, in turn, is measured by two types of indicators, time - the time delay of data exchange in the network, or the duration of the delay, and the bandwidth indicator, which determines the amount of information in a unit of time. These two indicators are interdependent, so if we know one, the other can be calculated.

An important parameter of a computer network is its reliability, its ability to function correctly for a long period of time. This feature includes three components: reliability, availability and convenient service. Reliability enhancement is the detection of damage, fault and failure in the network, ensuring normal thermal regimes for electronic circuits. Reliability is measured by fault intensity and the average time spent on fault elimination. The readiness criterion is evaluated by the readiness coefficient, which is equal to the time the system is in working condition. It is calculated by the ratio of the time spent on the elimination of errors to the number of errors. If we consider the network as a transport system, then this problem can remain on any highway in the network, so in case of fallibility organization, all elements of the network through which this route will pass must be reserved. Transition from primary to reserved can be done both automatically and artificially. A high degree of network reliability can be achieved when performance testing procedures and failover elements are built into protocols. An example of this is the FDDI protocol, where physical connections

between network nodes and switches are always tested [2, 6].

As a result of the study of heterogeneous computer networks of different levels, it can be said that there are different grades of computer network fallibility, for example:

High availability - characterizes systems implemented with conventional computer technologies, where redundant hardware and software resources are used, and the system recovery time interval ranges from 2 to 20 minutes.

Fault-resistant - characterizes such systems, which have additional (reserve) equipment in reserve for all functional blocks. Recovery time does not exceed 1 second.

Continuous standby - a feature of such systems where system recovery takes place within 1 second. Such systems provide continuous functionality and performance regardless of network problems [12].

Performance and reliability characteristics are closely related. These two characteristics are mutually exclusive. An increase in one leads to a significant decrease in the other. This is explained by the fact that the delay and fallibility of the connection channels cause some packets to be distorted and lost, therefore communication protocols are forced to retransmit data, and because local networks are busy with data recovery at the transport and application levels, which time-out In the mode they work for tens of seconds, the performance of the network drops significantly in case of low reliability.

III. Network service quality assessment

Based on the discussed criteria, modern corporate networks should meet certain strict requirements, one of the main requirements of which is to ensure its main function, i.e. to ensure the cooperation of computers and network devices connected to the network. All other requirements - performance, reliability, compatibility, manageability, security, extensibility and scalability - are related to the performance of the basic function. In network technologies, the term "quality of service" QoS (Quality of Service) was introduced, which includes only two main features - performance and reliability [1].

Two approaches have been adopted for assessing the quality of services. The first of these, from the user's point of view, implies that network service personnel guarantee certain quantitative indicators will be observed. For example, the packets sent by the subscriber of the network will be delayed for a time no more than 150 milliseconds, or the average bandwidth between the subscribers of the network A and B will be no less than 5 Mbit/s.

The second approach consists in the fact that the network will serve the subscribers according to their priorities, that is, only the given priority is guaranteed in the service. Such a service is called best effort. The network tries to serve subscribers with maximum quality, but does not make any guarantees. For example, local networks built on switches with frame priorities work according to this principle [1].

Performance - Network performance has several characteristics:

- Reaction time;
- Bandwidth;
- Transmission delay.

Reaction time - is defined as the time interval from the moment of generating an order for any service of the network until receiving a response to it. Of course, it is clear that this time depends on the type of service, which user applies, and which service he applies. In addition, it also depends on the state of network elements - the load on network segments, routers and switches, server load, etc. From the user, these characteristics cause the reply "The network is working slowly today".

Network response time has many components. Their knowledge is not of interest to the user. Network specialists should analyze these times, calculate the productivity of network elements, identify bottlenecks and, if necessary, modernize the network.

Bandwidth - reflects the amount of data transmitted by the network or its parts in a unit of time.

Bandwidth characterizes the speed of performance of internal network operations - the transfer of data packets between network nodes and through communication devices. Thus, this is one of the main characteristics of a network. Network throughput is measured either by the number of bits transmitted per second or by the number of packets transmitted per second. Instantaneous, peak and average bandwidth can be measured for the network.

Instantaneous bandwidth is calculated by dividing the total volume of data transferred by the time spent, while taking a fairly large period of time - hours, days or weeks. Average bandwidth characterizes a network or its elements over a long period of time.

Instantaneous throughput differs from average throughput in that it takes a very small portion of time for analysis, e.g. 1 millisecond or 1 second.

Maximum throughput This is the maximum value of the instantaneous throughput during the observation period. Maximum bandwidth characterizes the network's ability to handle peak network loads.

Bandwidth can be measured between any two nodes in a network. In some cases, it may be useful to calculate the total network bandwidth, which is defined as the amount of data transmitted by all nodes in the network per unit of time. This feature characterizes the quality of the entire network, without dividing it into segments and devices.

Transmission delay - is defined for any node or element of the network and represents the delay from the appearance of a message at the input of a device or node to the appearance of the same message at its output. This parameter is essentially close to reaction time.

Network bandwidth and transmission delay time are independent parameters. A network may have high throughput, but its nodes may have long transmission delays. An example of such a situation is the

geostationary satellite communication channel. The bandwidth of such a connection channel is quite high - 2 Mbps, but the transmission delay time is also high - 0.24 s, which is caused by the long transmission distance (72,000 km).

Reliability and security - one of the main goals of creating distributed systems is to achieve higher reliability compared to the connection of individual computers in the network. It is important to distinguish several aspects of reliability. For technical devices, reliability indicators such as the average time of fault-free operation, the probability of faults, and the intensity of faults are used. These characteristics are useful for such simple elements and devices, which are characterized by only two states, working and non-working. In complex systems, the situation is more complicated. They are characterized by intermediate states, which cannot be taken into account by these characteristics. Therefore, for complex systems such as computer networks, other characteristics are used.

Readiness or availability coefficient - (availability) refers to the part of time during which the system can be used.

It is possible to increase readiness in systems by adding spare elements to them, so that in case of failure of one element, other elements ensure the functioning of the system.

For a system to be highly reliable, it must have a high availability rate. This is a necessary condition, but not sufficient. Data protection and *consistency* must also be ensured, that is, if the same data is stored on several servers in order to increase reliability, their identity must also be ensured.

An important feature in networks is also the probability of forwarding a packet without distortion to the addressee node. Along with this characteristic, other characteristics can be used: the probability of packet loss (for any reason), the probability of distortion of individual bits in the transmitted data, and the ratio of lost to transmitted packets.

To characterize the overall reliability of the system, *security* is also used, that is, the ability of systems to protect data from unauthorized access. This is more difficult to achieve in distributed systems than in centralized systems.

Another feature of the network is ***fault tolerance***. In networks, falsification means the ability of the system to hide from users the fact of falsity of its individual elements, which the user may not even discover. Failure of any of their elements in fire-resistant systems leads only to a decrease in work quality, and not to a complete stop.

Extensibility and Scalability - The terms extensibility and scalability are sometimes used interchangeably in the literature, but this is incorrect. Both have a completely defined meaning.

Extension - (extensibility) refers to the network's ability to relatively easily add individual elements to the network, to complicate individual segments of the network, to replace existing equipment with more complex ones. Ease of expansion is often provided, but within very limited limits. For example, in an Ethernet network built with a large coaxial cable, it is easy to add new subscribers, but their number should not exceed 30-40. However, up to 100 subscribers can be physically connected to the segment, but at this time the performance of the network deteriorates dramatically. The presence of such a limitation indicates poor scalability of the system under conditions of maximum expansion.

Scalability - refers to the ability of the network to increase the number of nodes in the network and the connection distances over a very wide range without degrading its performance. In order to ensure the scalability of the network, it is necessary to use additional communication devices and it is often necessary to change the structure of the network. E.g. A multi-segment hierarchical structure network built on the basis of routers and switches is characterized by good scalability. Several thousand computers can be included in such networks so that each user is guaranteed the required level of service.

Transparency - The transparency of the network is achieved when the network is not a set of interconnected computers, but a unified computing environment in the traditional sense. E.g. Sun Microsystems' famous slogan "The network is the computer" indicates the need for good network transparency.

Hardness can be achieved at two different levels - the user level and the program level. User-level persistence means that it will use the same commands and procedures to work with remote resources as it does with local resources. Transparency at the program level means that the application program needs the same calls and references to access a remote resource as it does with local resources. Difficulty is easier to achieve at the user level than at the software level.

Networks must cover the peculiarities of operating systems and different types of computers. The term hardship can be used in different aspects. Statelessness means that the user is not required to know the location of software or hardware resources. Portability means that resources should be freely moved from one computer to another without changing their names. The difficulty of parallelism lies in the fact that the parallelization of the computing process takes place automatically, without the intervention of the programmer.

It is difficult to say about modern networks that they are all poor. The difficulty is more about the goal, namely how networks should be developed.

Support of different types of traffic - one of the main purposes of computer networks is to allow users to access different computers' resources. Traffic generated in computer networks is characterized by its own characteristics and is significantly different from telephone or cable TV traffic. Since the 90s of the last century, the creation of computer data traffic, which combines sound and video images in digital format, began. Computer networks began to be used for video conferencing, chat and entertainment. It is clear that dynamic transmission of multimedia traffic requires other algorithms, rules and devices. Today, the share of multimedia data in computer data traffic is quite large. It

has already infiltrated global and local networks and is an integral part of them.

In general, the strictest requirement for dynamic voice and image traffic is their synchronicity. Even small delays will distort sound and images.

Computer data traffic is characterized by completely different features. There are no strict requirements for synchronicity here, it is not a difficulty. But combining these two types of traffic is already a much more difficult task. Today, serious work is being done on networks to solve this problem. Networks based on ATM technology are closest to this goal. This technology was originally designed for the simultaneous presence of different types of traffic in the network.

Manageability - it implies the possibility of centralized control of its main elements, detection and elimination of problems in the work process, analysis of network performance and planning of its development. Ideally, network management tools provide a means of monitoring, controlling, and managing network elements, while treating the network as a whole rather than as a collection of individual computers.

An effective network management system monitors the network and, when a problem is detected, takes some action, corrects the situation, and notifies the administrator of what happened and what actions to take. Along with this, the management system should collect data on the basis of which the development of the network should take place.

That the management system is very useful is clearly visible in a large network. In corporate and global networks, through a good management system, network management personnel can be reduced.

There is still much to be done in the network management system. Modern management systems only monitor the current processes in the network and cannot take active actions.

Compatibility - Integrability, means that the network can include various types of software and hardware, different operating systems and regulations, different stacks. A network consisting of various elements is called non-homogeneous or heterogeneous. And if the heterogeneous network works without problems, it is called integrated [1, 10, 12].

CONCLUSION

The article discusses computer networks' functioning and reliability criteria and their determination methods. Network service quality assessment boundaries are presented. Because of the difficulty of direct calculations, many researchers are limited to only estimating confidence limits. In practice, the most productive computing systems are used, where it is possible to estimate the reliability of the network by limiting the number of network nodes. Modern approaches to the functioning of computer networks and the corresponding characteristics presented in the paper allow us to develop recommendations and criteria for the selection of network topology and structure, which will allow us to achieve higher reliability of computer networks.

REFERENCES

[1] J. Beridze, T. Burkadze, and A. Burkadze, “Management and Routing in Computer Networks,” Tb., 2009.

[2] A. Frangishvili, M. Kurdadze, and Z. Gasitashvili, “Fundamentals of computer network engineering,” Tb., 2009.

[3] G. Murjiknel, A. Robitashvili, T. Vekua et al, “Modern digital technologies of telecommunication,” Tb., 2006.

[4] O. Natroshvili, “Data reception-transmission management and diagnostic algorithms in computer networks,” GTU 2006.

[5] O. Shonia, G. Janelidze, and B. Mefarishvili, “Security of information and network resources GTU,” Tb., 2009.

[6] I. Gertsbakh, and Y. Shpungin, “Network Reliability A Lecture Course,” Springer, 2020.

[7] J.F. Kurose, and K. Ross, “Computer Networking A Top-Down Approach,” 7th. ed., 2017.

[8] S.K. Chaturvedi, Network Reliability: Measures and Evaluation, New York: John Wiley & Sons, 2016.

[9] M. O. Ball, C. J. Colbourn and J. S. Provan, “Network Reliability,” Handbook of Operations Research: Network Models, Elsevier North-Holland, vol. 7, 1995 , pp. 673–762.

[10] V. Olifer, “Computer Networks: Principles, Technologies and Protocols for Network Design,” Wiley, 2006.

[11] M. Whitman, and H. Mattord, “Principles of Information Security,” 5-th ed., 2014.

[12] P.B. Khorev, “Methods and means of protecting information in computer systems,” Academy, 2005.