Ramiz H. Shikhaliyev
Institute of Information Technology of ANAS, Baku, Azerbaijan
ramiz@science.az

DOI: 10.25045/jpit.v08.i2.03

THE CONCEPTUAL MODEL FOR THE INTELLECTUAL MONITORING SYSTEM OF COMPUTER NETWORKS

The effective management of computer networks (CN) is impossible without data about their status and function, which is provided by network monitoring. It is necessary to use a systematic approach for reliable and effective monitoring of the CN. The paper proposes a conceptual model of intelligent monitoring system of the CN, which seeks to create an effective infrastructure for collecting, storing and analyzing of monitoring data, as well as making decisions on network management.

Keywords*: computer networks, network monitoring, monitoring data, structure of the intellectual monitoring system.*

1. Introduction

Network monitoring is one of the components of the process of managing the computer networks (CN) and a source of obtaining objective data about their status and functioning. Without these data, it is impossible to make reasonable decisions on the management of the CN, especially if the scope of the CN is very large. At the same time, without conducting network monitoring, it is difficult to objectively assess the configuration of the hardware and software of the CN, as well as the changes occurred therein. Traffic management, network reconfiguration, fault identification, etc., which are based on the results of network monitoring, are implemented to ensure optimal performance and reliability of the CN.

In order for the management of the CN to be effective, it is necessary to constantly monitor them. However, constant monitoring of the CN may lead to some problems related to the collection, storage, processing and analysis of a larger volume of network traffic transmitted over the network. Since the scale and complexity of the CN is constantly changing, the network traffic of the CN is a complex and dynamic process consisting of different traffic flows. These traffic flows have many interconnected characteristics, and are generated by various devices, services, applications, protocols, etc. First of all, traffic flows are associated with CN management, for example, client initialization traffic, server traffic, and etc., which are periodically generated. Other traffic flows are the traffic of network services, applications and protocols, for instance, web service, DNS (Domain Name System), FTP (File Transfer Protocol), Windows Internet Naming Service (WINS) requests, ARP (Address Resolution Protocol) NetBIOS, HTTP (Hypertext Transfer Protocol), P2P (Peer-to-Peer), SMTP (Simple Mail Transfer Protocol), POP3 (Post Office Protocol v. 3), Telnet, etc., which constitute the bulk of network traffic of CN [1].

In terms of the ever-changing scope and complexity of the CN and the requirements for the CN itself (for example, performance, bandwidth, security, etc.), a systematic approach must be used to implement reliable and effective monitoring of the CN. A systematic approach to monitoring of the CN includes that the monitoring process is examined as a set of related elements, which are referred to collection (registration), storage, analysis of monitoring data and network management decision support, as well as feedback with the CN executed by the administrator, who performs specific management actions. At the same time, monitoring system model should be based on the use of intelligent technologies, since they minimize a human factor in the course of monitoring, reduce the loss of necessary information, minimize the impact of the monitoring system on the normal operation of the CN [2].

Nowadays, various monitoring systems for the CN are available in the information technology market [3]. Comparative analysis of these systems has shown that they differ depending on various criteria, such as the used technologies (for example, using the Simple

Network Management Protocol (SNMP) or agents), ways of data collection and storage, monitoring architectures, conditions for using these systems, be it commercial, conditionally free of charge or public. .

It is known that the main shortcomings of commercial monitoring systems are the high cost and closed-source software. These shortcomings are particularly important when the monitoring systems are used to monitor large CNs such as national CNs. Therefore, the development of the models for constructing a monitoring system for the CN is a topical matter.

This paper proposes a conceptual model of the intelligent monitoring system of the CN which seeks to create an effective infrastructure for collecting, storing and analyzing monitoring data and making network management decisions. At the same time, the intelligence of monitoring system of the CN mainly implies using intelligent technologies for analyzing the collected monitoring data and making decisions on network management.

2. Conceptual model of intelligent network monitoring system for computer networks

Monitoring of the CN, as a whole, can be conditionally divided into the network infrastructure monitoring and network traffic monitoring. Network infrastructure monitoring is used to improve the performance and reliability of network infrastructure. The tasks of the network infrastructure monitoring include troubleshooting (failure), monitoring of the performance of network nodes (gates), and also controlling the load of communication channels, identifying bottlenecks, optimal distribution and using network resources, etc. The monitoring of network traffic, in its way, refers to the communication infrastructure, the task of which is to manage network traffic, to provide the necessary parameters of quality of service, such as bandwidth, minimum latency and packet loss, and user behavior monitoring, etc. According to the above mentioned, we can conclude, to what extent, the tasks of the network infrastructure monitoring and network traffic monitoring intersect. Since the purpose of traffic management is maximizing the performance of communication channels. It can be applied to improve the efficiency of using communication channel resources and to avoid congestion in the network, and also to change the routing configuration in the network in the case of failures or problems connected with overloads.

A conceptual model of the intelligent monitoring system of the CN, which is an information model for the interaction of the components of the monitoring system, is proposed, in order to solve the monitoring tasks in general. These components can be a part of any monitoring system, regardless of its architecture. The proposed model is a modular thus consists of several subsystems (Fig. 1).it also enables to include new subsystems easily on specific aspects of the monitoring process and to reorganize the interaction of subsystems. Thus, if necessary, the system can be scaled and any task of monitoring of the CN can be solved, which will allow adapting the system to any functional and infrastructural change that occurs in the CN. Furthermore, one of the most important aspects of the modularity of the model of the intelligent monitoring system of CN is the ability to distribute and integrate subsystems on the scale of the CN, which is especially important for very large scales of CN.

The proposed model of the intellectual monitoring system includes the following subsystems: collection (registration) of monitoring data; storage of collected monitoring data; monitoring data mining and decision support for network management. At the same time, the administrator (or the administrator's console) of the network is also an integral part of the intellectual monitoring structure of the CN. Because, based on the decisions made by the decision support subsystem for network management, the administrator provides feedback to the CN and performs specific actions. In this case, the first three subsystems solve the problems of systematic long-term accumulation and CN data analysis. The subsystems do not include any reaction to the data obtained.

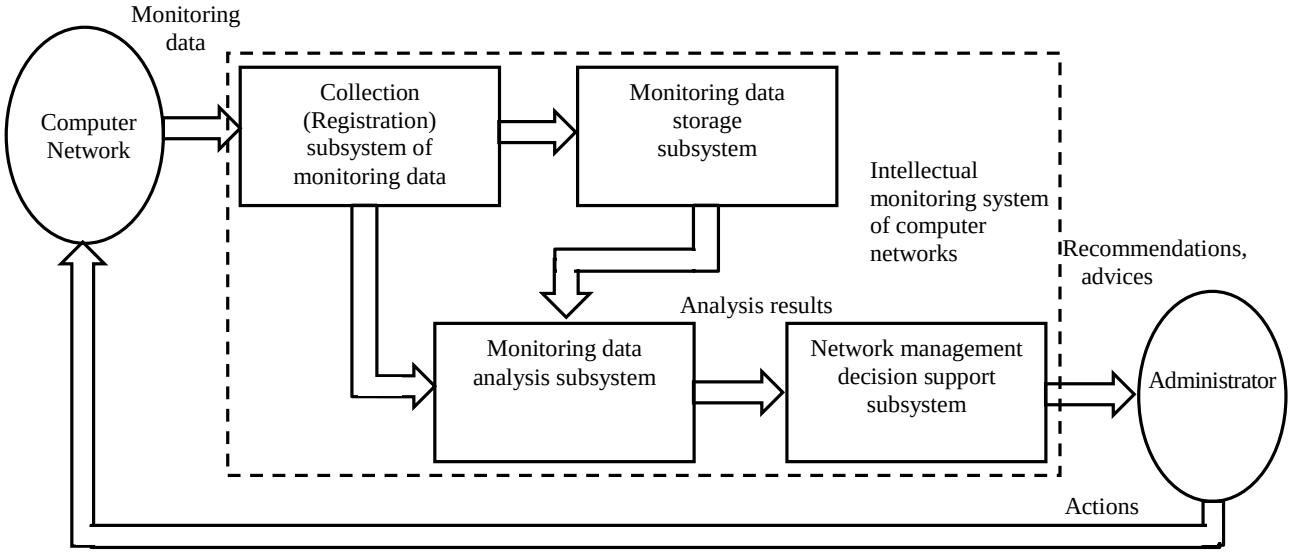

Fig. 1 Conceptual model of the intelligent monitoring system of the CN

3. Interaction and functions of subsystems

In the above proposed model of the intelligent monitoring system of CN, the interaction of the subsystems is as follows. First, network traffic (monitoring data) is collected in the monitoring data collection (registration) subsystem using network sensors. It can also carry out a preliminary analysis of these data. At the same time, most of the network traffic collection tools record the received data in files (log files) and usually have their own file formats. Therefore, it is very important to know the format of the stored file to easily organize data transfer between collection applications and data analysis, since most of them support certain file formats [4]. However, there are some common formats (for example, pcap) that are supported by most collection applications and data analysis, but the formats of the generated files can determine the necessary volumes for storing files. At the same time, as it was said above, constant monitoring of the CN can lead to some problems related to the collection, storage, processing and analysis of a larger volume of network traffic transmitted over the network. To solve the problems associated with the collection and storage of a larger amount of monitoring data, in [5], the authors proposed a method that allows converting the format of the collected monitoring data. This method minimizes the loss of veracity of monitoring data and reduces the space required to store them.

Depending on the monitoring tasks, various methods can be used: collection of all packages; collection of the network stream and the so-called collection of the extended flow [4, 6, 7]. Moreover, these methods have different requirements to the amount of memory required to store the collected data.

The purpose of packet collection is to collect all network traffic generated by computers and devices of the CN, in which the header of each data packet and the information transmitted in packets are collected and stored. In other words, a copy of each traffic packet is collected, processed and stored for further analysis. These collected data provide analysts with complete information about the traffic: packet header information and packet information. Therefore, this method of collecting monitoring data can be the most versatile, since a large amount of information can be intensively stored and processed.

A network stream is defined as a set of IP packets that pass through an observation point in the network for a certain time interval. All packages belonging to a particular flow have a set of common properties. The requirements to IP packet flows are defined in RFC 3917 [8].

The collection of an extended stream involves the collection of all packets and a network stream. At the same time, information directly taken from the packet headers or from the

information transmitted in the packets is added to the flow information. However, the extended stream may also contain additional information about some external source, for example, the geographical location of IP addresses of source and destination.

After the collection, monitoring data should be transferred to the storage subsystem of monitoring data, which accumulates, stores and archives data. At the same time, monitoring data should be kept long enough and reliable, if necessary, to be submitted for the analysis to the monitoring data analysis subsystem or it can be directly transferred to the monitoring data analysis subsystem.

Such an approach may enable offline and online monitoring of the CN.

Depending on the place and method of data storage, the required amount of storage for storage can significantly change, as well as problems related to administration and maintenance may occur, etc. Therefore, this subsystem should include data compression tools, working with databases of various types, and so on. However, data can be stored locally in a cloud or other external storage, and various data storage methods are used, such as: files (e.g., log files); databases and their hybrids [4, 6, 7]. Moreover, technologies, such as Cloud Computing [9], MapReduce, Hadoop [10], can be used for the large scale storage and processing of data.

The next stage in the monitoring data processing includes their analysis, which allows extracting information from the collected monitoring data and creating new knowledge about the operation of the CN. In order to achieve these goals, the basic tasks of data analysis, such as classification and clustering, finding associative rules, statistical regularities, finding a correlation etc., should be solved. For this purpose, various methods of data mining can be used, including methods of data merging [11], as well as methods of extracting knowledge from data (data mining) [12]. At the same time, network traffic analysis can be performed on several abstract levels: at port number level, packet content, stream, packet header and bit level (i.e. traffic volume). In point of fact, at each level, the analyzed characteristics of network traffic will differ, for example, at the packet level, the network traffic is characterized by the packet size and the time interval between packets. Moreover, bit-level analysis mainly refers to the quantitative characteristics of the network, such as transmission intensity and throughput of exchange in the communication channels of the network. At the packet level, the procedure of the arrival of IP packets is considered, i.e. the intensity of their delay and packet loss. As a result, a comprehensive and effective analysis of network traffic will allow solving such basic monitoring tasks as identifying the real state of the CN, identifying faults, determining the priorities for the generation of bandwidth for separate traffic, provision of the CN security, and so on.

Further, the information extracted from the collected monitoring data is transmitted and stored in the decision support subsystem for network management. However, this subsystem accumulates the experience of decision-making by network administrators to manage the CN. Based on the knowledge extracted from the monitoring data and the decisions support subsystem for network management, the network administrator makes decisions and performs certain specific actions.

The Network Management Decision Support Subsystem is an information system and is an interactive software-based system designed to support decision-makers. This subsystem will allow administrators of the CN to choose one from a variety of alternative network management solutions. The decision-making process should be automated so that a large amount of information can be promptly analyzed. This will enable administrators to respond quickly to critical situations occurring in the CN, for example, in the cases of nodes and communication channels failure, network security attacks, etc. Actually, for CN administrators, this subsystem is a means for the network monitoring and making operational decisions on network management.

The creation of decision support systems for the management of CN is a separate research topic and therefore, this paper does not explore it. However, different approaches to the creation of decision support systems in this or that field are available in various literatures. Five types of these systems are distinguished: decision support systems (managed by communications); data-

driven decision support systems; document-driven decision support systems; knowledge-based decision-making support systems, and decision-driven decision support systems [13].

Conclusion

The scale and complexity of the modern CN and the demand for CN, are constantly growing. Consequently their effective management becomes a daunting task. In order to manage the CN effectively, objective data on their status and functioning are required, which can be collected by network monitoring.

The paper proposes the modular conceptual structure of the intelligent monitoring system of the CN for the reliable and effective monitoring of the CN.

The proposed monitoring system will mitigate to include new subsystems on specific aspects of the monitoring process, to reorganize the interaction of subsystems, and to distribute and integrate subsystems on the CN scale. Thus, if necessary, the system can be scaled, and any task of the CN monitoring can be solved, which will adapt the system to any functional and infrastructural change that occurs in the CN.

References

1. Shikhaliyev R.H. The analysis and classification of the computer networks traffic // Problems of Information Technologies, 2010, No 2, pp. 15-23.
2. Shikhaliyev R.H. Application of Intelligent Technologies at the Network Monitoring of Computer Networks / / Artificial Intelligence, 2011, No 1, pp. 124-132.
3. Comparison of network monitoring systems. http://en.wikipedia.org/wiki/Comparison_of_network_monitoring_systems
4. Shikhaliyev R.H. Methods of collection, storage and analysis of large network traffic // Problems of Information Technologies, 2016, No 2, pp. 56-62.
5. Aceto G., Botta A., and Pescape A. Efficient storage and processing of high-volume network monitoring data // IEEE Transactions on Network and Service Management, 2013, vol. 10, no. 2, pp. 162−175.
6. Horneman A., Dell N. Smart collection and storage method for network traffic data. Technical Report CMU/SEI-2014-TR-011, 2014, p. 62.
7. Shikhaliyev R.H. On the methods of collecting and storing big network traffic / 10th IEEE International Conference on Application of Information and Communication Technologies (AICT2016), Azerbaijan, Baku, 12−14 October 2016, pp. 585−587.
8. Quittek J., Zseby T., Claise B., Zander S. RFC 3917: Requirements for IP Flow Information Export (IPFIX). Internet Engineering Task Force, 2004. http://tools.ietf.org/html/rfc3917.
9. Sivashakthi T., and Prabakaran N. A survey on storage techniques in cloud computing // International Journal of Emerging Technology and Advanced Engineering, 2013, vol. 3, no. 12, pp. 125−128.
10. White T. Hadoop: The Definitive Guide. O'Reilly Media, p. 768, 2015.
11. Bleiholder J., and Naumann F. Data fusion // ACM Computing Surveys, 2008, vol.41, no.1, pp. 1−41.
12. Han J., Kamber M., Data mining: concepts and techniques. Morgan Kaufmann, p. 743, 2006.
13. Marin G. Decision support systems // Journal of Information Systems & Operations Management, 2008, vol. 2, no. 2, pp. 513−520.