Modeling and Performance Analysis of Priority Queuing Systems

Dariusz Strzęciwilk and Włodek M. Zuberk

1 Department of Applied Informatics, University of Life Sciences, Nowoursynowska Street 159, 02-787 Warsaw, Poland
dariusz_strzeciwicki@agw.pl

2 Department of Computer Science, Memorial University, St. John’s, NL A1B 3X5, Canada
wlodek@mun.ca

Abstract. The paper presents the results of modeling and analysis of data performance on systems that support QoS (Quality of Service). In order to evaluate the performance of the modeled systems used were TPN (Timed Petri Nets). Studied were mechanisms of traffic shaping systems based on PQS (Priority Queuing System). Tested was the impact of the mechanism of generating traffic using TPN. Moreover, discussed were the basic mechanisms and queuing systems occurring in QoS structures. It is shown that models can be effectively used in the modeling and analysis of the performance of computer systems.

Keywords: Priority queuing system · Petri nets · Performance analysis
Modeling · QoS data

1 Introduction

Queueing theory is one of the heavily used tools in modeling and testing the quality of transmission in computer networks [1–3]. This theory uses a mathematical apparatus associated with the theory of stochastic processes, in particular Markov processes [4]. By queueing system herein is meant a system that on the one hand receives notification requiring maintenance, on the other hand, there are the so-called maintenance devices, designed to meet the needs of these applications. If the process of receiving applications exceeds the capabilities of their immediate service, a queue is created. A queueing system may be characterized by regulations of queues, i.e., the way one determines the order of service applications in the system [5]. The most common queueing systems are FIFO (First In First Out), LIFO (Last In First Out), SIRO (Select In Random Order), PQ (Priority Queuing). The basic mechanism that supports the transfer of packages is FIFO scheduling that is easy to implement and treats all packets equally. FIFO scheduling is not suitable to provide for a good quality of service transmission, as when the packets come from different traffic flows, one of them can easily disrupt the flow of the other remaining streams. Packet processing in the order of flue means that an aggressive stream can appropriate the higher capacity of a router queue. This can result in poor transmission causing, for example, sudden increase in delays of transmitting packets. Developed was a lot of packet scheduling algorithms that have better insulation between

© Springer International Publishing AG, part of Springer Nature 2019
https://doi.org/10.1007/978-3-319-91186-1_31
the streams [6]. In the case of the priority scheduling algorithms, some application can be handled before others, regardless of when they occurred in the system. Priority queuing systems form a large class of queuing systems where the incoming requests are to be distinguished by their importance [7]. A typical example of the use of such algorithms are routers, to which are flowing subsequent packets. Core routers classify incoming packets to the specified classes of traffic, and then handle packets belonging to the aggregated streams. Packets are handled in accordance with an implemented queuing mechanism and specific support and traffic shaping policies in order to provide services with the agreed QoS [8]. QoS is one of the most important challenges arising during the design and maintenance of both modern computer networks and next generation networks [9]. Guaranteeing adequate quality of service is of particular importance in the case of real-time applications such as Voice over IP [10] and video - IPTV [11]. These services are particularly sensitive to delay and require a guaranteed bandwidth [12]. To provide the desired QoS packages for the entire route from the sender to the recipient has been the subject of research for many years [13–15]. Research in this area can be divided into two groups. In the analytical methods the authors sought solutions of algebraic or differential equations which bind together the probability of events in the system. In the simulation methods were used most often implantation queuing algorithms, which were then subjected to statistical analysis. Based on the analysis of available research, it can be stated that the analytical methods most commonly include relatively simple queuing systems, which require implementation of many of the assumptions of the stochastic nature of the traffic flow. Complex systems are very difficult in the analysis and their functioning can be effectively examined by simulation methods. It should be noted that, to construct a sufficiently accurate model is not simple, and the waiting time for results could be discouragingly long. Hence, the aim of this study was to use models of Petri nets to assess the efficiency and to study the effectiveness of queuing mechanisms of PQS. Such an assessment may also be useful in the design and analysis of data in computer networks, distributed systems and multiprocessor systems. Constructed queuing models allowed estimation of significant features and parameters of the system under test.

2 Petri Nets and Network Models

Petri nets are a graphical and mathematical tool used in many fields. They are seen as a mathematical tool for modeling of concurrent systems [16, 17]. Although there are many varieties of Petri nets [18, 19] their common feature is the structure based on a bipartite directed graph, i.e. graph with two types of vertices, alternately connected by facing edges (or arcs). These two types of vertices represent, in general terms, conditions and events occurring in a modeled system, but each event can occur only when fulfilled are all the conditions associated with it. Formally Petri net is defined as a system \( N = (P, T, A) \) composed of a finite set of \( p \)-elements \( P \) (representing conditions), a finite set of \( t \)-element \( T \) (representing events) and set \( A \) of arcs connecting the \( p \)-components with \( T \)-elements and \( t \)-elements with \( p \)-elements, \( A \subseteq p \times T \cup T \times p \). set \( A \) is called the relation of parity. \( P \)-elements are connect by arcs directed to \( t \)-element and are called its input elements, while