UDC 519.872

ANALYSIS OF DIFFERENCES IN THE CHARACTERISTICS OF QUEUING SYSTEMS WITH THE DYNAMICS OF INPUT STREAMS SELF-SIMILARITY

Leonid A. Uryvsky, Anastasiia V. Kryklyva

Institute of Telecommunication Systems
Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, Ukraine

Background. Modern queuing theory is used in many fields and in telecommunications it plays an important role. Today, the modern way to implement queuing systems is a cloud environment. And the main problem of queuing system is to ensure the quality of service.

Objective. The purpose of the paper is to study the indicators of service quality of queuing systems with different initial characteristics, to analyse the service quality characteristics when comparing quantitative characteristics of different parameters of queuing systems to identify features of queuing system with self-similarity.

Methods. Analyse the theory of QoS, types of distribution, such as Poisson and Weibull, as well as the properties of self-similar traffic. Compare the values of the main parameters that provide the desired quality of service at different inputs and different Hearst parameters.

Results. The importance of taking into account the self-similarity factor for the accuracy of the values of the parameters in the calculation that ensure the quality of service in the QMS.

Conclusions. Research of the influence of self-similar traffic on the quality of service in queuing systems.

Keywords: *queuing system; quality of service; Poisson distribution; Weibull distribution; self-similarity.*

Introduction

Queuing system - a system that performs the maintenance of the requirements that come to it. Maintenance requirements in the queuing system are performed by service devices. Classic queuing systems contain from one to an infinite number of devices. It is usually assumed that the input flow is Poisson [1].

Today, the rapid growth of traffic users, changes in its nature and structure, a significant increase in bandwidth can contribute to possible congestion of network objects, their buffer devices and, consequently, lead to delays and packet loss. Therefore, when servicing packet traffic, special attention is paid to maintaining the characteristics of the quality of service (QoS) [2], [3].

To obtain functional dependencies and assess the performance of the network, the cloud environment is considered as a set of queuing networks consisting of queuing systems.

A new trend in the modelling of queuing systems in recent years has been the study of systems not with independent flow of applications (as in the traditional Poisson model), but with the assumption of correlation of flows in total traffic. This case is classified as traffic self-similarity.

Self-similar processes are characterized by the presence of consequences: the probability of occurrence of another event depends not only on time but also on previous events. This means that the number of current events may depend on the number of previous events in the remote time.

Despite the long period of studying the self-similarity of teletraffic, a significant class of tasks remains unsolved:

1) in fact, there is no strict theoretical basis that would replace the classical theory of teletraffic in the design of modern mobile networks that use self-similar traffic;
2) there is no single generally accepted model of self-similar traffic;
3) there is no reliable and recognized method of calculating the quality characteristics of QoS for systems and networks serving self-similar traffic;
4) there are no mechanisms and algorithms that ensure the quality of service in terms of self-similar traffic [4].

Problem statement

To analyse the parameters of service in a cloud environment, a queuing system with Poisson distribution and Weibull distribution, which describes self-similarity, is used. To determine the degree of self-similarity the Hurst parameter is used, which may be in the range $0 < H < 1$.

The closer the Hurst parameter is to one, the more the properties of self-similarity are manifested, but limited strictly in the range of $0.5 < H < 1$. With a Hurst parameter equal to $H = 0.5$, this corresponds to the case of lack of self-similarity. So each queuing system is described by a given type of request flow, service duration, number of service channels and service discipline.

An important feature of the Poisson (simplest) flow of events is that the time between two consecutive events is a random variable distributed according to the exponential law [5], [6].

$$F(t) = 1 - e^{-\lambda t} \quad (1)$$

where λ> 0 is the flow intensity.

And the functional characteristic is the load factor ρ:

$$\rho = \lambda \cdot t \quad (2)$$

where λ - the intensity of applications in queuing systems,
t - duration of service applications in the queuing system.

Let's define indicators of quality of service for a case of Poisson's stream of applications and their service.

The obtained value of the load factor allows determining the main functional characteristics of the queuing system, using the known formulas of Little [7]:

- average number of Q applications in queuing systems (for service and in the queue):

$$Q = \frac{\rho}{1-\rho} \quad (3)$$

- the average length of the queue, ie applications pending service:

$$L = \frac{\rho^2}{1-\rho} \quad (4)$$

- the average length of stay W_{system} application in the system:

$$W_{system} = \frac{1}{\mu \cdot (1-\rho)} \quad (5)$$

- the average waiting time in the queue $V_{waiting}$, which is determined by the delay of the application in the queue and depends on the number of applications in the queue:

$$W_{waiting} = \frac{\rho}{\mu \cdot (1-\rho)} \quad (6)$$

In the case when time-dependent (correlated) applications arrive in the cloud environment, it is advisable to use the appropriate model with input flow. Corresponding to the Weibull distribution.

Consider a queuing system of the form Wb / M / 1 / ∞ (queuing system serves the flow of applications, which is described by the Weibull distribution (Wb), the service time has an exponential distribution (M), queuing system single-line with infinite queue), i.e. Weibull distribution, given by the differential distribution function [5, 8]:

$$f(x) = \begin{cases} \alpha \cdot \beta^{-\alpha} \cdot x^{\alpha-1} \cdot e^{-\beta \cdot x^{\alpha}}, & x \geq 0 \\ 0, & x \leq 0 \end{cases} \quad (7)$$

where α – is the shape parameter of the distribution curve, 0 < α < 1;

$$\alpha = 2 - 2 \cdot H \quad (8)$$

H – Hurst parameter, 0,5 < H < 1,

$$\beta = \left[\lambda \cdot \Gamma\left(1 + \frac{1}{\alpha}\right)\right]^{\alpha} \quad (9)$$

β – distribution parameter, β > 0,
λ – the intensity of service packets in the queuing system,
Γ(k) – Euler's gamma function.

It is known that for the queuing system $W_B/M/1/\infty$ the probability that the application received in the queuing system will find n service requests is defined as:

$$r_n = (1-\sigma) \cdot \sigma^n, \quad (10)$$

where σ – the root of the equation $0 \leq \sigma < 1$:

$$\sigma = F(\mu - \mu \cdot \sigma) \quad (11)$$

where F – Laplace-Stieltjes transformation,
μ– intensity of service of packages in queuing systems [applications / hour, packages / second] [8].

- the average waiting time of the application in the queue:

$$W_{waiting} = \frac{\sigma}{\mu(1-\sigma)} \quad (12)$$

- the average time spent in the application system:

$$W_{system} = \frac{\rho \cdot \sigma}{\mu \cdot \lambda(1-\sigma)} \quad (13)$$

- average number of applications Q:

$$Q = \frac{\rho \cdot \sigma}{\mu(1-\sigma)} \quad (14)$$

where ρ – load factor of queuing systems;
- the average queue length L of applications is:

$$L = \frac{\rho \cdot \sigma}{1-\sigma} \quad (15)$$

Proving the values of the parameters that at H = 0.5 it will be the simplest flow of events or Poisson's distribution, we will calculate the indicator σ according to formulas (7) - (11):

$$\alpha = 2 - 2 \cdot 0{,}5 = 1$$

$$\beta = \left[\lambda \cdot \Gamma\left(1 + \frac{1}{\alpha}\right)\right]^{-\alpha} = \beta = \left[\lambda \cdot \Gamma\left(1 + \frac{1}{1}\right)\right]^{-1} = \frac{1}{\lambda}$$

$\sigma = \alpha \cdot \beta \cdot \int_0^{+\infty} e^{-(\mu-\mu\cdot\sigma)\cdot t} \cdot t^{\alpha-1} \cdot e^{-\beta\cdot t^{\alpha}} \cdot dt = 1 \cdot \frac{1}{\lambda} \cdot \int_0^{+\infty} e^{-(\mu-\mu\cdot\sigma)\cdot t} \cdot t^{1-1} \cdot e^{-\lambda\cdot t^{1}} \cdot dt = \mu \cdot (1 - \sigma)$.

So we get:

$$\frac{\sigma}{1-\sigma} = \mu$$

Solution

To compare the parameters of the Poisson version of traffic (H = 0.5) and self-similar input flow (H = 0.6 and H = 0.7), the calculated values according to relations (3) - (6) and (12) - (15) are summarized to Table 1.

Table 1. Estimated values for comparing parameters.

The load factor of the queuing system, ρ	0.8	0.7	0.64	0,6	0.5	0.49	0.48	0.47	0.46	0.45
The average time the application is in the system, W_{system}										
H = 0,5	6.67	3.33	2.22	1.67	1	0.87	0.769	0.686	0.617	0.56
H = 0,6	9.403	3.131	1.544	0.91	0.348	0.265	0.207	0.165	0.134	0.111
H = 0,7	14.99	5.024	2.481	1.479	0.575	0.436	0.341	0.273	0.224	0.185
Average delay time, $W_{waiting}$										
H = 0,5	5.33	2.33	1.42	1	0.5	0.43	0.37	0.33	0.228	0.225
H = 0,6	7.052	3.132	1.930	1.366	0.696	0.594	0.517	0.457	0.409	0.372
H = 0,7	11.245	5.024	3.102	2.219	1.150	0.980	0.854	0.755	0.679	0.616
Average number of applications, Q										
H = 0,5	4	2.33	1.78	1.5	1	0.96	0.92	0.9	0.87	0.86
H = 0,6	5.642	2.192	1.235	0.819	0.348	0.291	0.248	0.214	0.188	0.167
H = 0,7	8.996	3.517	1.985	1.331	0.575	0.48	0.409	0.355	0.313	0.277
The length of the application queue, L										
H = 0,5	3.2	1.63	1.137	0.9	0.5	0.47	0.44	0.42	0.39	0.37
H = 0,6	4.231	2.193	1.544	1.229	0.696	0.655	0.621	0.591	0.565	0.543
H = 0,7	6.747	3.517	2.482	1.997	1.15	1.08	1.023	0.976	0.938	0.901

According to the data from table 1, graphs were constructed to compare the average residence time of the application in the system and the average number of applications at different values of the Hearst parameter.

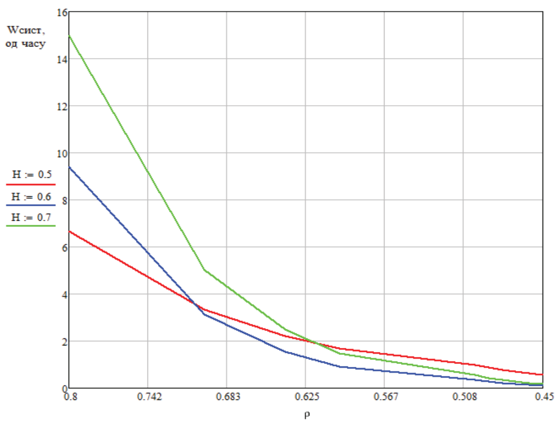

Fig. 1. The average residence time of the application in the system at different values of the Hearst parameter

Consider the first of the dependences - the average residence time of the application in the system Wsystem, as a function of the intensity of the input flow (load factor of the queuing system) ρ, depending on the degree of self-similarity of traffic.

The nature of the dependence is such that the discrepancy in the indicators increases with increasing intensity of the input flow and, at the same time, the self-similarity factor.

When the value of ρ = 0.8 and H = 0.7, the discrepancy in the value of W_{system} reaches 230%.

Quantitatively, this is equivalent to errors in determining the performance of the queuing system using traditional, Poisson models.

Therefore, in the case of using traditional models of queuing systems to predict the quality of service in queuing systems with signs of self-similarity of traffic, the error can reach significant values!

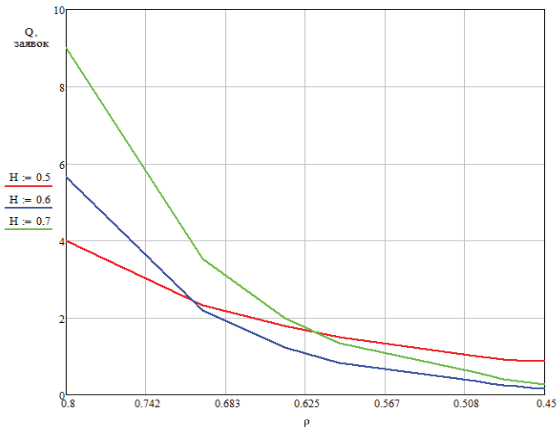

Fig. 2. The average number of applications for different values of the Hearst parameter

When the value of ρ = 0.8 and H = 0.7, the difference in the value of the average number of applications Q reaches 225%.

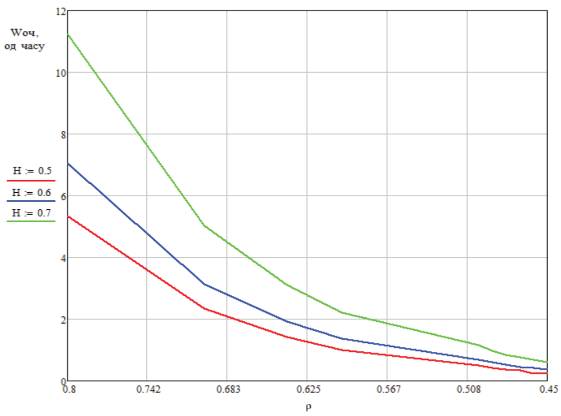

Fig. 3 - Average waiting time of the application at different values of the Hearst parameter

Analysing the obtained graph of the average waiting time of the application in the queue, we can observe a faster growth of the graphs of the average delay time with the property of self-similarity compared to the simplest flow.

When the value of ρ = 0.8 and H = 0.7, the difference in the value of the average waiting time for applications Wwaiting reaches 211%.

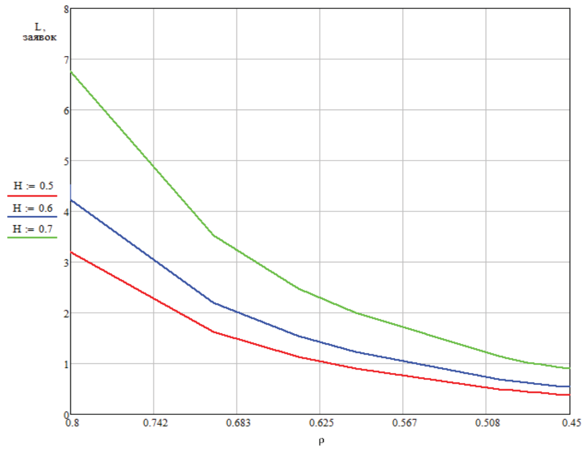

Fig. 4 - The length of the application queue at different values of the Hearst parameter

When the value of ρ = 0.8 and H = 0.7, the discrepancy in the value of the length of the queue of applications L reaches 211%.

Therefore, having received graphs, it is possible to draw the general conclusion that. if we do not take into account the quantitative characteristics of the degree of self-similarity of traffic, i.e. the Hirst H parameter, it is impossible to adequately represent the characteristics to meet the quality of service in a cloud environment [9].

Conclusions. In the report of the method of finding the values of service quality parameters in queuing systems is investigated using the formulas of Little, who is one of the founders of queuing theory.

An important point for the analysis of parameters to ensure the quality of service in queuing systems is the use of self-similarity, which determines more accurate values of indicators compared to the analysis of parameters in the simplest flow of events.

To analyse the parameters of service in a cloud environment, a queuing system with Poisson distribution and Weibull distribution is used, which describes self-similarity at different Hirst (H) parameters, as this parameter determines the degree of self-similarity.

The Hearst parameter is considered in the range of values from 0.5 to one. The closer H is to unity, the more the property of self-similarity is manifested.

The mathematical model with Weibull distribution asymptotically tends to the Poisson distribution with a Hearst coefficient H = 0.5.

As the Hearst parameter increases, the values of the considered service quality indicators increase, namely: the average time spent by applications in the system and the average number of applications in queuing systems.

Therefore, if the self-similarity factor is ignored, the system will not meet the expected parameters, which can lead to a significant error in determining the quantitative values of the characteristics of queuing systems needed to ensure the desired quality of customer service.

References

1. Erlang, Agner K., The Theory of Probabilities and Telephone Conversations, in The Life and Works of A.K. Erlang, by Brockmeyer, E.; Halstrøm, H. L.; Jensen, Arne (eds.), Transactions of the Danish Academy of Technical Sciences, 2, Akademiet for de Tekniske Videnskaber, 1909, pp. 131-137.

2. A. Kryklyva, "Analysis of the influence of the properties of self-similar traffic on the quality of service", XIII International scientific and technical conference of students and graduate students "Prospects for the development of information and telecommunications technologies and systems" PRITS-2021: Proceedings of the conference – K .: Igor Sikorsky Kyiv Polytechnic Institute, p.390, 2021.

3. I. Strelkovskaya, I. Solovskaya, A.Makoganiuk, "Optimization of QoS characteristics of self-similar traffic ", Problems of Infocommunications 169 Science and Technology: conference proceedings of the 2017 4th International Scientific-Practical Conference Proceedings (PICS&T 2017), p.497-500, Kharkiv, Ukraine, Oct. 10-13, 2017. https://doi.org/10.1109/INFOCOMMST.2017.8246447. (Scopus)

4. I. Strelkovskaya, I. Solovskaya, A. Makoganiuk, A. Balyk, "Research of the quality characteristics of self-similar traffic of a mobile communication network on the basis of software release", International Research Journal. Information and telecommunication sciences: Volume 11 Number 2(21), p. 51-57, Jul.-Dec. 2020.

5. Kleinrock L., Queueing Systems, Vol. I, II, John Wiley & Sons, 1975

6. Giambene G., Queuing Theory and Telecommunications: Networks and Applications, Giovanni Giambene. –Springer, 2014.

7. Little, J.D.C., "A Proof for the Queueing Formula L = λW", Operations Research, Vol. 9, 1961, pp. 383–387.

8. I. Strelkovskaya, T. Grigorieva, I. Solovskaya, "Service of self-similar traffic in G / M / 1 queuing systems with Weibull distribution", Proceedings of universities. Radioelectronics, v.61, № 3, 2018, pp.173 – 178.

9. L. Uryvsky, A. Kryklyva, "Research of service indicators dynamics in queuing system with self-similar traffic", XVI International scientific and technical conference "Prospects for the development of information and telecommunications technologies and systems" PRITS-2021: Proceedings of the conference. – K .: Igor Sikorsky Kyiv Polytechnic Institute, pp.53-58, 2022.

Уривський Л.О., Криклива А.В.

Аналіз відмінностей в характеристиках систем обслуговування з динамікою самоподібності вхідних потоків

Проблематика. Сучасна теорія масового обслуговування використовується в багатьох сферах діяльності і в сфері телекомунікацій теорія масового обслуговування займає важливу роль. На сьогоднішній день сучасним способом реалізації систем масового обслуговування є хмарне середовище. І головною проблемою СМО є забезпечення якості обслуговування.

Мета дослідження. Дослідити показники якості обслуговування СМО з різними вихідними характеристиками, аналіз характеристик якості обслуговування при порівнянні кількісних характеристик при різних параметрах СМО задля виявлення особливостей СМО із властивістю самоподібності.

Методика реалізації. Проаналізувати теорію СМО, види розподілу, такі як Пуассона і Вейбулла, а також властивості самоподібного трафіку. Порівняти значення основних параметрів, що забезпечують потрібну якість обслуговування при різних вхідних показниках і різних параметрах Херста.

Результати дослідження. Важливість урахування фактору самоподібності для точності значень параметрів при розрахунку, що забезпечують якість обслуговування в СМО.

Висновки. Дослідження впливу самоподібного трафіку на якість обслуговування в системах масового обслуговування.

Ключові слова: *система масового обслуговування, якість обслуговування, розподіл Пуассона, розподіл Вейбулла, самоподібність.*