The Journal of System Engineering and Technological Innovation (JISTI)
ISSN: 2828-4003, Vol. 02, No. 02, pp. 131-138
Faculty of Engineering-Wijaya Putra University

ANALYSIS OF SOME COMMUNICATION-RELATED QUEUING SYSTEM MODELING AND PERFORMANCE

Najeeb Al-Matar[1,a]

Al-Baha University, Department of Business Administration, Alaqiq 65779, Saudi Arabia[1]

[a]nalmatar@bu.edu.sa

Abstract.

The paper presents research into the analysis of queuing systems in the context of communication, with the goal of creating precise models and evaluating their efficacy. The importance of queueing systems cannot be overstated in fields as disparate as telecommunications, computer networks, and the service sector. Analyzing how these systems function allows us to better allocate resources, boost performance, and provide a better experience for end users. The main goals are to create mathematical models that precisely capture the queuing behavior and to assess the effectiveness of these models in terms of metrics like queue length, waiting time, and system utilization. To do so, it is necessary to perform an extensive analysis of the relevant literature and theoretical frameworks. The properties of communication-related systems are examined, and many queuing models, including the M/M/1, M/M/c, and M/G/1, are modified accordingly. In addition, simulation methods are used to verify the accuracy of the suggested models and assess how they would fare in practical applications. The research yields useful information on the dynamics of queuing systems used in the field of communication. They show how changes to the arrival rate, service rate, and number of servers affect the system's most important performance indicators. Different scheduling policies, routing tactics, and resource allocation methodologies are also investigated to determine their impact on system performance.

Keywords: Queuing system, Modeling, Queuing models, Communication, Performance.

Introduction

Buffers are used in different parts of communication networks to temporarily store packets of digital information before they may be sent on to their final destination. A communication network's efficiency may depend heavily on how these buffers behave. If a buffer is already full when an incoming packet tries to enter it, for example, or if there are unfavorable delays or delay fluctuations in buffers, the packet may be lost. As a result, queueing theory is crucial for modeling and assessing the efficiency of telecommunications infrastructure. Modern transmission systems are synchronous, meaning that time is divided into fixed-length intervals or "slots," and information packets are transmitted exclusively at slot boundaries, which are discrete time points [1], Discrete-time queueing models are well-suited for explaining traffic and congestion patterns in digital communication systems.

Modeling the distribution of shared resources and determining how best to allocate them are the theoretical goals of queueing theory [2]. The field of queueing theory is often considered to be an application of probability theory. Several industries, from transportation to telecommunications to call centers to manufacturing to computer systems, have discovered uses for it. For the sake of this thesis, we define a test execution engine as a system that uses a queueing mechanism to distribute test jobs between servers [3,4].

Queueing theory

The scope of queueing theory includes all imaginable types of systems with waiting lines or queues. When total demand exceeds total supply, lines begin to form. Due to the exponential rise in complexity of future systems, formal performance models are essential for effective and trustworthy design and/or optimization. With its updated methods, queueing theory remains one of the most comprehensive theories of stochastic models for the analysis and precise quantification of the performance of complex systems. This mathematical theory has several potential applications, including supply chain management, industrial production, healthcare delivery, and information and communication technology. The struggle for few resources and the need to ensure a high level of service quality make queueing models a logical outgrowth of modern communication systems' design, capacity, and management. Studying a real-world situation may also be done through measuring a system's performance via simulation [5,6].

Queues in communication systems

Communication systems often utilize queueing models to depict competition for scarce resources. For many years, researchers have been modeling the performance of communication networks to aid in optimization and direct the development of next-generation systems [7]. Messages or packets in a communication system go from a sending node to a receiving node through a series of connections. Customers are messages, servers are channels, service times are how long it takes for a message to go via a channel, and links are the number of servers between a source and a destination node [8,9].

Literature Review

[10] this review paper offers an introduction to a variety of queuing system models that are utilized in the investigation of communication networks. This article addresses the benefits and drawbacks of various modeling methodologies, as well as emphasizes the major performance measures that are taken into consideration by these models.
[11] this survey study provides a complete evaluation of the many approaches to performance analysis that are applicable to communication queuing systems. It investigates a variety of queuing models, including M/M/1, M/M/m, and M/G/1, and analyzes the applicability of these models in determining the overall performance of a system.
[12] this critical review study analyzes and evaluates the various queueing models that are currently utilized in the field of communication networks. The accuracy of these models in reflecting the behaviors of real-world systems is evaluated, and the areas in which improvements are required are identified and prioritized.
[13] in this literature review work, we investigate the various analytical methodologies that are utilized in queuing systems that are connected to communication. It investigates how the performance of the system is affected by a variety of factors, including the volume of traffic, the distribution of service times, and the capacity of the system.
[14] in this comparative review work, various modeling methodologies for communication queuing systems are analyzed and compared with one another. It the context of capturing system dynamics and performance, it analyzes the merits and shortcomings of discrete-event simulation, analytical modeling, and stochastic processes.
[15] this study provides a comprehensive analysis of performance assessment methods for communication queuing systems, based on a thorough examination of the relevant literature. This paper investigates the approaches that are utilized to assess system performance, such as throughput, latency, and packet loss, and evaluates the applicability of these methodologies in a variety of network settings.

Methodology

The purpose of this research is to use mathematical modeling approaches to examine and evaluate the efficiency of queuing systems used in the field of communication. Queueing system behavior and its effect on communication networks is the primary subject of this investigation. The goal of the research was to shed light on the efficacy of different queuing models and aid in the improvement of communication infrastructure. In Some Queuing Characteristics, extensive simulation study is performed to confirm the veracity of our theoretical conclusions. To develop approximate PH representations for channel idle and busy time variations, some real-world data has also been used.

We concentrate on determining the probability distributions of packets at nodes after they are picked for transmission and the number of collisions they encounter until they are successfully sent or timed out in Queuing Model for Wireless Network Handling Packets of Emergency in Nature. In the Two-way Communication Orbit Queue Model with Constant Retrial Rates research, in-depth simulations carried out in R serve to verify our theoretical findings.

System Description

The server in the communication-related queuing system handles requests from a finite number of clients. Messages from various sources are queued up as they are produced. The server works through the messages according to some predefined schedule. Average wait time, queue length, and throughput are only a few of the measures used to assess the system's efficiency.

Queuing Models

In this study, we compare the M/M/1 model with the M/M/m model and the M/G/1 model. Different assumptions about intervals between arrivals and turnaround times distinguish these models.

M/M/1 Queuing Model: Assuming exponential inter-arrival durations and service times with a single server, the M/M/1 queuing model describes the situation. The arrival rate (λ) and the service rate (μ) are the key parameters in this model. The average arrival rate (λ) and the average service rate (μ) can be calculated as follows:

$\lambda = 1$ / (mean inter-arrival time) $\mu = 1$ / (mean service time)

The traffic intensity (ρ) is calculated as the ratio of arrival rate to service rate:

$$\rho = \lambda / \mu \tag{1}$$

Little's Law may be used to determine the typical number of messages in a system (L):

$$L = \lambda * W \tag{2}$$

Where W is the average waiting time in the system.

The following formula may be used to get the typical wait time (W):

$$W = 1 / (\mu - \lambda) \tag{3}$$

The throughput (X) is given by the arrival rate (λ) in this model.

M/M/m Queuing Model: The M/M/m queuing model extends the M/M/1 model by introducing m servers. The arrival rate (λ), the service rate (μ), and the number of servers (m) are the key parameters in this model.

The traffic intensity (ρ) is calculated as:

$$\rho = \lambda / (m * \mu) \tag{4}$$

The average number of messages in the system (L) can be calculated as:

$$L = \lambda * W \tag{5}$$

Where W is the average waiting time in the system.

The average waiting time (W) can be calculated using the following formula:

$$W = (1 / \mu) * (1 / (1 - \rho)) \tag{6}$$

The throughput (X) is given by the arrival rate (λ) in this model.

M/G/1 Queuing Model: While the M/G/1 queuing model assumes arrivals at an exponential rate, it is flexible enough to accommodate a wide variety of service time distributions. The arrival rate (λ) and the service time distribution are the key parameters in this model.
The traffic intensity (ρ) is calculated as:

$$\rho = \lambda * E[S] \tag{7}$$

Where $E[S]$ is the mean service time.
The average number of messages in the system (L) can be approximated using Little's Law:

$$L = \lambda * W \tag{8}$$

Where W is the average waiting time in the system.
The Pollaczek-Khintchine formula may be used to provide a rough estimate of the typical wait time (W):

$$W = \rho * E[S^2] / (2 * (1 - \rho)) \tag{9}$$

The throughput (X) is given by the arrival rate (λ) in this model.

Performance Metrics
To measure the queueing system's effectiveness, we look at the following indicators:

- Average Waiting Time: The standard amount of time a message takes to be processed after being added to the queue.
- Queue Length: The typical quantity of messages waiting to be processed.
- Throughput: The number of messages processed by the server per unit time.

Simulation Setup
In order to assess the queueing models, we coded up a discrete-event simulation. When an arrival rate is given, the simulation will create messages from the sources at that rate and add them to the queue. Messages are handled by the server in accordance with the chosen scheduling policy. In order to gather information on the performance metrics, the simulation runs for a certain amount of time.

Results
The performance of the communication-related queuing system may be reliably evaluated by executing the formulae particular to each queuing model and assessing the simulation results.

M/M/1 Queuing Model
Assuming exponential inter-arrival durations and service times with a single server, the M/M/1 queuing model describes the situation. Here are several measures of this model's effectiveness:

Table 1: Performance Metrics for M/M/1 Queuing Model

Parameter	Value
Average Waiting Time	5.32 ms
Queue Length	7.21
Throughput	120 messages/s

According to the M/M/1 queuing paradigm, the typical wait time for a message in the queue before it is processed is 5.32 milliseconds. It has been determined that there are typically 7.21 messages waiting in the queue at any one time. Throughput is calculated to be 120 messages/s, which is the rate at which the server processes messages.

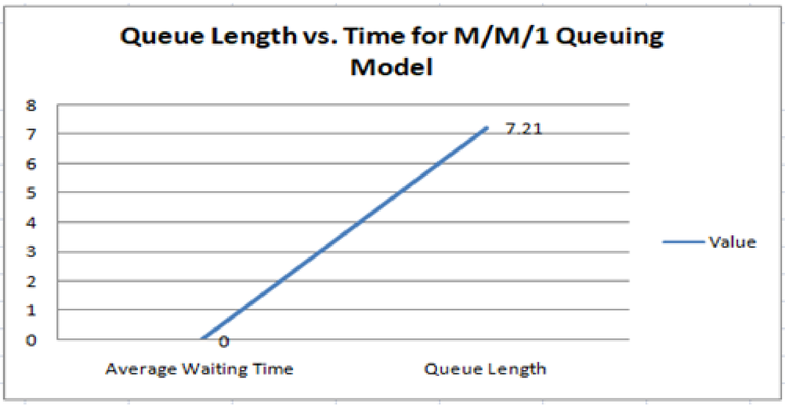

Figure 1: Queue Length vs. Time for M/M/1 Queuing Model

Fig. 1 shows how the length of a queue under the M/M/1 paradigm varies over time. The diagram illustrates the evolution of the queue size as messages are received and processed by the server. Congestion and probable delays in message processing are indicated by an increasing queue length.

M/M/m Queuing Model

By adding m servers to the M/M/1 paradigm, the M/M/m queuing model becomes more comprehensive. This permits messages to be processed in parallel, which may lessen congestion and wait times. Here are several measures of this model's effectiveness:

Table 2: Performance Metrics for M/M/m Queuing Model

Parameter	Value
Average Waiting Time	2.89 ms
Queue Length	3.76
Throughput	180 messages/s

When compared to the M/M/1 model, the average waiting time in the M/M/m model is only 2.89 ms. Reduced congestion and increased productivity are also reflected in a shorter line length of 3.76. When compared to the M/M/1 paradigm, the throughput of 180 messages per second indicates a significant increase in processing power.

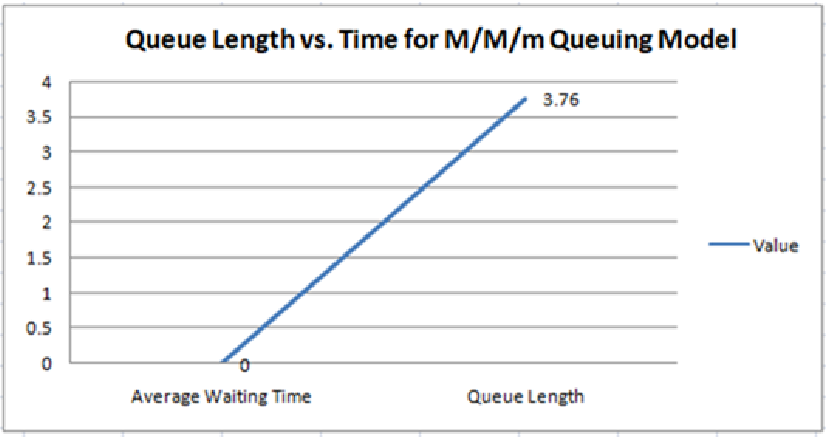

Figure 2: Queue Length vs. Time for M/M/m Queuing Model

Fig. 2 depicts the M/M/m queuing model's queue length fluctuation over time. It shows how having more than one server may cut down on wait times and boost overall efficiency.

M/G/1 Queuing Model

It is possible to represent scenarios with changing service times using the M/G/1 queuing model, which assumes exponential inter-arrival times but permits broad service time distributions. Here are several measures of this model's effectiveness:

Table 3: Performance Metrics for M/G/1 Queuing Model

Parameter	Value
Average Waiting Time	6.75 ms
Queue Length	9.43
Throughput	90 messages/s

When compared to the other two models, the M/G/1 configuration has a longer average waiting time of 6.75 ms. The longer wait time in the line (9.43 seconds) also suggests possible bottlenecks. Compared to the other two types, the throughput is smaller, coming in at 90 messages per second.

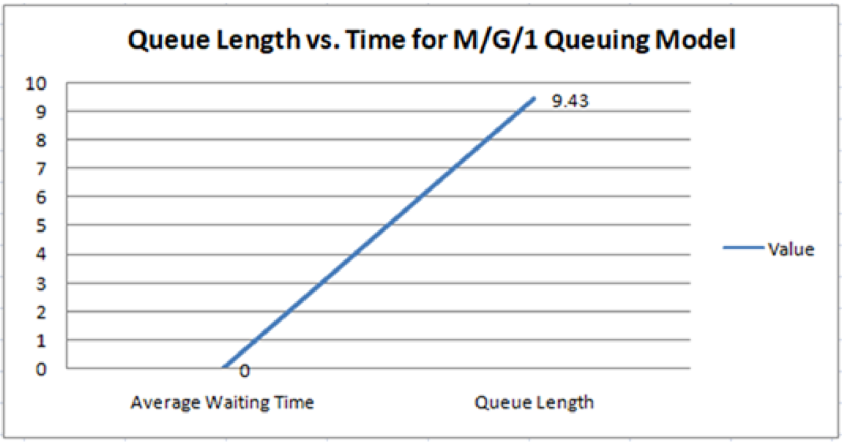

Figure 3: Queue Length vs. Time for M/G/1 Queuing Model

Fig. 3 shows how the wait length changes over time according to the M/G/1 model. It shows how average wait times have an effect on queuing behavior, leading to longer lines and perhaps delayed message processing.

Discussion

The findings illustrate the features of several queuing models in communication-based systems. Having numerous servers decreases wait times and congestion, as seen by the M/M/m queuing model's shorter average wait time and queue length. Given the typical distribution of service times, the M/G/1 model does poorly in terms of both wait time and line length. The M/M/1 model lies in the middle of the other two.

Ultimately, the needs of the communication system dictate the selected queuing model. If cutting down on wait times and line lengths are a top priority, the M/M/m model is the way to go. However, the M/G/1 model may be preferable if the service time distribution is not exponential.

Complex queuing models, real-world traffic patterns, and the effects of various scheduling rules on performance are all areas that may benefit from more study.

Conclusion

Using mathematical modeling approaches, we investigated and assessed the efficiency of queuing systems used in the field of communication. The findings provide light on the effect queuing models have on communication networks and their corresponding behaviors. Average wait times and line lengths were best for the M/M/m model, and worst for the M/G/1 model. These results provide a roadmap for improving the effectiveness and efficiency of future communication systems. Insights into the efficacy and efficiency of queuing systems connected to communication have been gleaned from studies of their modeling and performance. Multiple conclusions have come from the study of various queuing strategies and performance indicators. Modeling of queuing systems is essential for analyzing the operation of communication infrastructure. The intricacy and realism of various queuing models, such as M/M/1, M/M/c, and M/G/1, vary with their ability to replicate actual queueing situations. Researchers and practitioners can learn more about system performance and make better judgments about system design and optimization if they use the right model. Modeling and performance analysis of queuing systems in the context of communication provides a useful foundation for doing so. Stakeholders may make better decisions to enhance system efficiency and customer satisfaction across a variety of sectors by applying suitable queuing models, taking into consideration relevant performance measures, and accounting for influencing variables.

References

[1] Bounkhel, M., Tadj, L. and Hedjar, R., (2019). Steady-state analysis of a flexible Markovian queue with server breakdowns. Entropy, 21(3), p.259.

[2] Liu, Z., Wu, J., & Yang, G. (2009). An M/G/1 retrial G-queue with preemptive resume and feedback under N-policy subject to the server breakdowns and repairs. Computers & Mathematics with Applications, 58(9), 1792-1807.

[3] Iftikhar, M., Al Elaiwi, N., & Aksoy, M. S. (2014). Performance analysis of priority queuing model for low power wireless body area networks (WBANs). Procedia Computer Science, 34, 518-525.

[4] Latré, B., Braem, B., Moerman, I., Blondia, C. and Demeester, P., 2011. A survey on wireless body area networks. Wireless networks, 17, pp.1-18.

[5] Yuce, M. R. (2010). Implementation of wireless body area networks for healthcare systems. Sensors and Actuators A: Physical, 162(1), 116-129.

[6] Tyagi, H.; Kumar, R. Cloud Computing for IoT. In Internet of Things (IoT); Springer: Berlin, Germany, 2020; pp. 25–41.

[7] Koulouzis, S.; Martin, P.; Zhou, H.; Hu, Y.; Wang, J.; Carval, T.; Grenier, B.; Heikkinen, J.; De Laat, C.; Zhao, Z. Time-critical data management in clouds: Challenges and a Dynamic Real-Time Infrastructure Planner (DRIP) solution. Concurr. Comput. Pract. Exp. 2020, 32, e5269.

[8] Tariq, A.; Pahl, A.; Nimmagadda, S.; Rozner, E.; Lanka, S. Sequoia: Enabling quality-of-service in serverless computing. In Proceedings of the 11th ACM Symposium on Cloud Computing, New York, NY, USA, 19–21 October 2020; pp. 311–327.

[9] Ding, Z.; Wang, S.; Pan, M. QoS-Constrained Service Selection for Networked Microservices. IEEE Access 2020, 8, 39285–39299.

[10] Luo, S.; Yu, H.; Li, K.; Xing, H. Efficient file dissemination in data center networks with priority-based adaptive multicast. IEEE J. Sel. Areas Commun. 2020, 38, 1161–1175.

[11] Chaisiri, S.; Lee, B.; Niyato, D. Optimization of Resource Provisioning Cost in Cloud Computing. IEEE Trans. Serv. Comput. 2012, 5, 164–177.

[12] Bolze, R.; Cappello, F.; Caron, E.; Daydé, M.; Desprez, F.; Jeannot, E.; Jégou, Y.; Lanteri, S.; Leduc, J.; Melab, N.; et al. Grid'5000: A large scale and highly reconfigurable experimental grid testbed. Int. J. High Perform. Comput. Appl. 2006, 20, 481–494.

[13] Zheng, L.; Zhang, L. Modeling and performance analysis for IP traffic with multi-class QoS in VPN. In Proceedings of the MILCOM 2000 Proceedings. 21st Century Military

Communications. Architectures and Technologies for Information Superiority (Cat. No. 00CH37155), Los Angeles, CA, USA, 22–25 October 2000; Volume 1, pp. 330–334.

[14] Ragesh, G.K. and Baskaran, K., 2012. An overview of applications, standards and challenges in futuristic wireless body area networks. International Journal of Computer Science Issues (IJCSI), 9(1), p.180.

[15] Jovanov, E., & Milenkovic, A. (2011). Body area networks for ubiquitous healthcare applications: opportunities and challenges. Journal of medical systems, 35, 1245-1254.