AN ANALYSIS OF QUEUING MODEL

Raghavendra V
Research Scholar
Department of Mathematics
NIILM University, Kaithal, Haryana
raghuwnp9582@gmail.com

Dr. Ankush
Research Guide
Department of Mathematics
NIILM University, Kaithal, Haryana

Abstract

The mathematical analysis of waiting lines is known as queuing theory, and it is particularly helpful to define Modern information technologies necessitate innovations based on modeling, analyzing, and designing transactions as well as the process of traffic control in everyday human activities like telecommunications, reservation counters, super markets, big bazaars, and theater ticket windows, as well as figuring out the order of computer operations, computer performance, health services, airport traffic, and airline ticket sales. In the world of computers, Queue models are also the foundation of Parallel and Distributed systems. In this essay, we'll talk about the queueing model and theory's methodology.

Keywords: *Mathematics Study, Queueing System, traffic control, airline ticket*

Introduction to Queueing Theory

A.K. Erlang, a Danish mathematician who investigated telephone traffic congestion issues in the first decade of the 20th century, introduced the queueing theory. The queueing theory is highly helpful in a wide variety of real-world applications, including telephone exchange, traffic control, manufacturing, inventory, and communication systems, as well as supermarkets, gas stations, computer systems, etc. A collection of mathematical techniques known as queuing theory is used to analyze probabilistic systems of servers and clients. Operational research's queueing theory, commonly referred to as the theory of overcrowding, examines the connection between demand on a service system and the delays experienced by its customers.

Because the findings are often employed when making business choices regarding the resources required to deliver service, queueing theory is typically regarded as a subfield of operations research. The theory has a wide range of beneficial applications, the majority of which have been well-researched in the literature on probability, operations research, management science, and industrial engineering. Traffic flow (cars, planes, people, communications), scheduling (patients in hospitals, tasks performed by machines, and computer programs), and facility design (banks, post offices, theme parks, and fast-food restaurants) are a few examples.

Basic elements of Queue

Building a mathematical model to reflect how items enter the queue, the criteria by which they are put into service, and the processing time is the foundation of queue analysis. The whole range of these models, which include all perceivable systems that contain queue characteristics, is embodied in queueing theory. Whether it is a person or another kind of unit, we identify it.

A line or file of people is a queue. The word "queue" refers to the formation of a line while awaiting something, or a line of people waiting to be serviced at a location that offers the service they need.

The manner in which users enter the system; the sequence in which users enter the service facility after joining the line.

The way services are provided and leaving the system.

The term "arrival" describes the typical number of clients that need assistance within a certain amount of time.

Customers may be people, items that are still being produced, raw materials, incoming digital communications, or any other modeled entities that need to wait for a process to finish. The size of the queue can be unlimited or finite.

A server may be a person, a computer, or any other kind of processor who carries out a procedure for clients who are waiting.

The priority system by which the next client to receive service is chosen from a group of clients who are waiting is known as queue discipline. First-in-first-out, or FIFO, is a queue discipline that is frequently used.

The term "Service Rate" (also known as "Service Capacity") describes the average number of customers that a system can serve at any given time.

Systems of events called stochastic processes have random variables for the intervals between the events. Customers' patterns of arrival and service are modeled in queueing models as stochastic processes based on probability distributions.

Utilization is the percentage of time a server (or network of servers) is actively serving clients.

Refer to the server's adoption of scheduling for information on server work. The following rules are observed by servers.Customers are welcomed by the server for service. In this context, the criteria such as “first-come, firstserved” (FCFS), “last-come, first- served” (LCFS), and random selection for service” (RS) are self-explanatory. Other terms like "round robin" and "shortest processing time" might require some clarification. Customers in certain classes frequently receive preferential treatment when receiving services. For the effective operation of computers and communication systems, numerous other queue disciplines have been introduced. Other aspects of customer behavior, such as reneging, jockeying, and balking, must also be taken into account.

The Notation use for Scheduling discipline is

FCFS: first-come, first-served LCFS: last-come, first-served RS: Random selection for Service RRS: Round Robin Faison

SPT: shortest processing time

Queueing Systems

a network with queues is Systems with an arbitrary but limited number of queues are called networks of queues. Through the network, sometimes in various queues, customers are serviced at the node. Other queuing systems within the same network may serve as the user sources for some of the queuing systems in the network.

Networks of service centers that require clients to wait in line for service at some or all of the centers. As a result, studying the complete network is required to gather data like the anticipated total waiting time, total projected consumer volume, and so forth. The relevance of queueing networks has led to a flurry of study in this vast domain of parallel and distributed computing networks.

Additional details on the connections between the queuing systems, how they interact, and how users are allocated to the waiting systems must be included in order to define a queuing network.

Service delays during busy periods are inevitable when responding to erratic requests whose timing and location are determined by probabilistic rules. It would be hard to afford to have enough capacity at all times to prevent delays. Designing service systems that strike a reasonable balance between system running expenses and the delays experienced by system users is the function of analysis.

Queue Model

Queueing Model are used to estimate desired performance measures of the system

Provide rough estimate of a performance measure

Typical measures

Server utilization

Length of waiting lines

Delays of customers

Applications

Determine the minimum number of servers needed at a service center

Detection of performance bottleneck or congestion

Evaluate alternative system designs

Common Distributions

M : Exponential

Ek :Erlang with parameter k

Hk :Hyperexponential with parameter k(mixture of k exponentials)

D : Deterministic(constant)

G : General(all)

Birth-Death Processes

The foundation of many of the most commonly used queuing models

Birth – equivalent to the arrival of a customer or job

Death – equivalent to the departure of a served customer or job

The discrete space Markov processes in which the transitions are restricted to neighboring states

Process in state n can change only to state n+1 or n-1

Example

The number of jobs in a queue with a single server and individual arrivals(not bulk arrivals)

Must trade off benefits of high utilization levels with benefits of flexibility and service

Conclusions

The usage of queueing systems is widespread throughout society. The effectiveness of these systems may significantly impact both the productivity of the process and the quality of human life. By creating mathematical models of how queueing systems work and utilizing those models to generate performance metrics, queueing theory analyzes queueing systems. This research gives key insights for effectively building queueing systems that strike a good balance between the price of providing a service and the price of using that service. Queueing models may be used to assess how well networking performs in parallel and distributed system models.

References

1. *Cooper, R. B.: Introduction to Queueing Theory, 2d ed., Elsevier North-Holland, New York, 1981.*

2. *I.G. Boldur-LăŃescu, I. Suciu, E. łigănescu, Operational Research with Applications in Economy, A.S.E., 1990 a. The art of computer systems performance analysis. Raj Jain Gh. Dodescu, Operating Systems, ASE 1997 [1]. Gh. Dodescu, B. Oancea, M. Raceanu, Parallel Processing, Ed. Economica, Bucharest, 2002 b. D. Gross,*

3. *C. M. Harris, Fundamentals of Queuing Theory, Wiley, New York, 2003 J. Joseph, C. Fellenstein, Grid Computing, Prentice Hall, 2003*

4. *Karlin, S., and Taylor, H. An introduction to stochastic modeling. Harcourt, New York, 1998. a. Khintchine, A. Mathematical methods in the theory of queueing. Hafner, New [1]. Kleinrock, L. Queueing systems. Vol. I.*

5. *Theory. John Wiley & Sons, New York, 1975*

6. *Queue Network,- Customers Signals and Product form Solution, by X. Chao, M. Miyazawa, and M. Pincdo Journal of Applied Mathematics and Stochastic Analysis, 14:4 (2001), 421-426*

[1]. Basic Queueing Theory:
7. Debrecen, 2011 by :Dr. JánosSztrik University of Debrecen, Faculty of Informatics [2]. Karlin, S., and Taylor, H. Stochastic process (in Hungarian). GondolatKiadó, Budapest, 1985.
8. Walrand, J.: An Introduction to Queueing Networks, Prentice-Hall, Englewood Cliffs, NJ, 1988.
9. Wolff, R. W.: Stochastic Modeling and the Theory of Queues, Prentice-Hall, Englewood Cliffs,NJ, 1989.