Queuing Theory in today's world-An Overview

Vatsal Mishra[1], Harshit Jain[1], Ashish Nagra[1], Naresh K[2], Jagadeesh Kannan R[3]

[1]*Student, BTech Computer Science and Engineering, VIT University, Vellore, TN, India*
[2]*Assistant Professor, School of Computer Science and Engineering, VIT University, Vellore, TN, India*
[3] *Professor, School of Computer Science and Engineering, VIT University, Chennai,* TN, India

---***---

Abstract - *When all is said in done we don't prefer to hold up. Be that as it may, decrease of the holding up time for the most part requires additional speculations. To choose whether or not to contribute, it is essential to know the impact of the speculation on the holding up time. So we require models and strategies to investigate such circumstances. In this course we treat various basic queueing models. Consideration is paid to techniques for the examination of these models, furthermore to uses of queueing models. Essential application regions of queueing models are creation frameworks, transportation and stocking frameworks, correspondence frameworks and data handling frameworks. Queueing models are especially helpful for the outline of these framework regarding design, limits and control.*

Keywords: Queuing hypothesis, Queuing models, Distributions, Probability Theory, Queuing Networks

1. INTRODUCTION

A lining framework comprises of at least one servers that give administration or something to that affect to arriving clients. Clients who touch base to discover all servers occupied for the most part go along with at least one (lines) before the servers, consequently the name lining frameworks. There are a few regular cases that can be depicted as lining frameworks, for example, bank employee benefit, PC frameworks, producing frameworks, upkeep frameworks, correspondences frameworks et cetera.

Lining hypothesis is the scientific investigation of holding up lines, or lines. In lining hypothesis, a model is developed so that line lengths and holding up time can be anticipated. Lining hypothesis is for the most part considered a branch of operations research in light of the fact that the outcomes are regularly utilized when settling on business choices about the assets expected to give an administration.

Queueing hypothesis has its sources in research by Agner Krarup Erlang when he made models to depict the Copenhagen phone trade. The thoughts have since seen applications including media transmission, activity building, processing and the outline of industrial facilities, shops, workplaces and healing centers.

2. USE OF QUEUING

Cases Below we quickly portray a few circumstances in which queueing is imperative.

2.1 Example: Supermarket.

To what extent do clients need to hold up at the checkouts? What happens with the holding up 7 time amid pinnacle hours? Are there enough checkouts?

2.2 Example: Production framework.

A machine produces distinctive sorts of items. What is the generation lead time of a request? What is the decrease ahead of the pack time when we have an additional machine? Would it be advisable for us to dole out needs to the requests?

2.3 Example: Data correspondence.

In PC correspondence systems, for example, the Internet information bundles are transmitted over connections starting with one switch then onto the next. In every switch approaching parcels can be cushioned when the approaching interest surpasses the connection limit. Once the cushion is full, approaching bundles will be lost. What is the parcel delay at the switches? What is the division of parcels that will be lost? What is a decent size of the cradle?

2.4 Example: Parking parcel.

They will make another parking area before a grocery store. How expansive would it be a good idea for it to be?

2.5 Example: Assembly of printed circuit sheets.

Mounting vertical parts on printed circuit sheets is done in a gathering focus comprising of various parallel addition machines. Every machine has a magazine to store parts. What is the creation lead time of the printed circuit loads up? By what means ought to the parts vital for the get together of printed circuit sheets be separated among the machines?

3. SINGLE QUEUEING NODES

Single queueing hubs are generally portrayed utilizing Kendall's documentation as a part of the frame A/S/C where A depicts the time between landings to the line, S the span of occupations and C the quantity of servers at the hub. Numerous hypotheses in lining hypothesis can be demonstrated by lessening lines to numerical frameworks known as Markov chains, initially depicted by Andrey Markov in his 1906 paper.

Agner Krarup Erlang, a Danish specialist who worked for the Copenhagen Telephone Exchange, distributed the main paper on what might now be called lining hypothesis in

1909. He demonstrated the quantity of phone calls touching base at a trade by a Poisson procedure and settled the M/D/1 line in 1917 and M/D/k lining model in 1920. In Kendall's documentation:

- M remains for Markov or memory less and implies landings happen as indicated by a Poisson procedure
- D remains for deterministic and means occupations touching base at the line require a settled measure of administration
- k depicts the quantity of servers at the queueing hub (k = 1, 2,...). On the off chance that there are a larger number of occupations at the hub than there are servers then employments will line and sit tight for administration

The M/M/1 line is a basic model where a solitary server serves employments that land as indicated by a Poisson procedure and have exponentially dispersed administration prerequisites. In a M/G/1 line the G remains for general and shows a discretionary likelihood appropriation. The M/G/1 model was fathomed by Felix Pollaczek in 1930, an answer later recast in probabilistic terms by Aleksandr Khinchin and now known as the Pollaczek–Khinchine recipe.

After the 1940s queueing hypothesis turned into a region of research enthusiasm to mathematicians. In 1953 David George Kendall comprehended the GI/M/k line and presented the current documentation for lines, now known as Kendall's documentation. In 1957 Pollaczek concentrated on the GI/G/1 utilizing a basic condition. John Kingman gave an equation for the mean holding up time in a G/G/1 line: Kingman's recipe.

The network geometric strategy and framework systematic strategies have permitted lines with stage sort appropriated between entry and administration time dispersions to be considered.

Issues, for example, execution measurements for the M/G/k line remain an open issue.

3. SERVER DISCIPLINES

Different planning strategies can be utilized at lining hubs:

3.1 First in first out

This standard expresses that clients are served each one in turn and that the client that has been holding up the longest is served first.

3.2 Last in first out

This standard likewise serves clients each one in turn, yet the client with the most brief holding up time will be served first. Otherwise called a stack.

3.3 Processor sharing

Benefit limit is shared similarly between clients.

3.4 Priority

Clients with high need are served first. Need lines can be of two sorts, non-pre-emptive (where an occupation in administration can't be interfered) and pre-emptive (where an occupation in administration can be hindered by a higher-need work). No work is lost in either show.

3.5 Shortest employment first

The following employment to be served is the one with the littlest size

3.6 Pre-emptive briefest occupation first

The following occupation to be served is the one with the first littlest size

3.7 Shortest outstanding handling time

The following employment to serve is the one with the littlest outstanding handling necessity.

3.8 Service office

- Single server: clients line up and there is stand out server
- Parallel servers: clients line up and there are a few servers
- Tandem line: there are numerous counters and clients can choose going where to line

3.9 Customer's conduct of holding up

- Balking: clients choosing not to join the line in the event that it is too long
- Jockeying: clients switch between lines in the event that they think they will get served speedier by so doing
- Reneging: clients leave the line on the off chance that they have sat tight too ache for administration

4. QUEUING NETWORKS

Systems of lines are frameworks in which various lines are associated by client directing. At the point when a client is adjusted at one hub it can join another hub and line for administration, or leave the system. For a system of m the condition of the framework can be depicted by a m–dimensional vector (x1,x2,...,xm) where xi speaks to the quantity of clients at every hub.

The principal huge results around there were Jackson systems, for which a proficient item shape stationary dispersion exists and the mean esteem analysis[24] which permits normal measurements, for example, throughput and visit times to be figured. In the event that the aggregate number of clients in the system stays consistent the system is known as a shut system and has likewise been appeared to have a product–form stationary dispersion in the Gordon–Newell hypothesis. This outcome was reached out to the BCMP arrange, where a system with exceptionally broad administration time, administrations and client directing is appeared to likewise show an item frame stationary appropriation. The normalizing steady can be computed with the Buzen's calculation, proposed in 1973.

Systems of clients have additionally been researched, Kelly systems where clients of various classes encounter diverse need levels at various administration hubs. Another kind of system are G-organizes initially proposed by Erol Gelenbe in 1993: these systems don't expect exponential time appropriations like the exemplary Jackson Network.

5. Programming projects FOR SIMULATION/ANALYSIS

The accompanying are the significant programming which are utilized for reenactment/examination

- Java Modeling Tools, a GPL suite of lining hypothesis apparatuses written in Java
- Queuing Package for GNU Octave
- Discrete Event Simulation for Python
- Queuing Process Models in the Wolfram Language
- PDQ programming bundle for R measurable processing
- SimEvents for MATLAB

6. CONCLUSIONS

With the information of likelihood hypothesis, information and yield models, and birth demise forms, it is conceivable to determine various lining models, including however not constrained to the ones we saw in this paper. Lining hypothesis can be pertinent in numerous true circumstances. For instance, seeing how to display a various server line could make it conceivable to decide what number of servers are really required and at what wage so as to expand money related proficiency. Then again maybe a lining model could be utilized to concentrate on the life expectancy of the globules in road lights keeping in mind the end goal to better see how habitually they should be supplanted. The utilizations of lining hypothesis amplify well past holding up in line at a bank. It might take some imaginative considering, yet in the event that there is any kind

REFERENCES

1. Sundarapandian, V. (2009). "7. Queueing Theory". Probability, Statistics and Queueing Theory. PHI Learning. ISBN 8120338448.
2. Lawrence W. Dowdy, Virgilio A.F. Almeida, Daniel A. Menasce. "Performance by Design: Computer Capacity Planning by Example".
3. Schlechter, Kira (March 2, 2009). "Hershey Medical Center to open redesigned emergency room". The Patriot-News.
4. Mayhew, Les; Smith, David (December 2006). Using queuing theory to analyse completion times in accident and emergency departments in the light of the Government 4-hour target. Cass Business School. ISBN 978-1-905752-06-5. Retrieved 2008-05-20.
5. Tijms, H.C, Algorithmic Analysis of Queues", Chapter 9 in A First Course in Stochastic Models, Wiley, Chichester, 2003
6. Kendall, D. G. (1953). "Stochastic Processes Occurring in the Theory of Queues and their Analysis by the Method of the Imbedded Markov Chain". The Annals of Mathematical Statistics. **24** (3): 338. doi:10.1214/aoms/1177728975. JSTOR 2236285.
7. A.A. Markov, Extension of the law of large numbers to dependent quantities, Izvestiia Fiz.-Matem. Obsch. Kazan Univ., (2nd Ser.), 15(1906), pp. 135–156 [Also [37], pp. 339–361].
8. "Agner Krarup Erlang (1878 - 1929) | plus.maths.org". Pass.maths.org.uk. Retrieved 2013-04-22.
9. Asmussen, S. R.; Boxma, O. J. (2009). "Editorial introduction". Queueing Systems. **63**: 1. doi:10.1007/s11134-009-9151-8.
10. Erlang, Agner Krarup (1909). "The theory of probabilities and telephone conversations" (PDF). Nyt Tidsskrift for Matematik B. **20**: 33–39. Archived from the original (PDF) on 2011-10-01.
11. Kingman, J. F. C. (2009). "The first Erlang century—and the next". Queueing Systems. **63**: 3–4. doi:10.1007/s11134-009-9147-4.
12. Pollaczek, F., Ueber eine Aufgabe der Wahrscheinlichkeitstheorie, Math. Z. 1930
13. Whittle, P. (2002). "Applied Probability in Great Britain". Operations Research. **50**: 227–177. doi:10.1287/opre.50.1.227.17792. JSTOR 3088474.
14. Kendall, D.G.:Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded Markov chain, Ann. Math. Stat. 1953
15. Pollaczek, F., Problèmes Stochastiques posés par le phénomène de formation d'une queue

16. Kingman, J. F. C.; Atiyah (October 1961). "The single server queue in heavy traffic". Mathematical Proceedings of the Cambridge Philosophical Society. **57** (4): 902. doi:10.1017/S0305004100036094. JSTOR 2984229.

17. Ramaswami, V. (1988). "A stable recursion for the steady state vector in markov chains of m/g/1 type". Communications in Statistics. Stochastic Models. **4**: 183–188. doi:10.1080/15326348808807077.

18. Penttinen A., Chapter 8 – Queueing Systems, Lecture Notes: S-38.145 - Introduction to Teletraffic Theory.

19. Harchol-Balter, M. (2012). "Scheduling: Non-Preemptive, Size-Based Policies". Performance Modeling and Design of Computer Systems. p. 499. doi:10.1017/CBO9781139226424.039. ISBN 9781139226424.

20. Harchol-Balter, M. (2012). "Scheduling: Preemptive, Size-Based Policies". Performance Modeling and Design of Computer Systems. p. 508. doi:10.1017/CBO9781139226424.040. ISBN 9781139226424.

21. Harchol-Balter, M. (2012). "Scheduling: SRPT and Fairness". Performance Modeling and Design of Computer Systems. p. 518. doi:10.1017/CBO9781139226424.041. ISBN 9781139226424.

22. Jackson, J. R. (1957). "Networks of Waiting Lines". Operations Research. **5** (4): 518–521. doi:10.1287/opre.5.4.518. JSTOR 167249.

23. Jackson, James R. (Oct 1963). "Jobshop-like Queueing Systems". Management Science. **10** (1): 131–142. doi:10.1287/mnsc.1040.0268. JSTOR 2627213.

24. Reiser, M.; Lavenberg, S. S. (1980). "Mean-Value Analysis of Closed Multichain Queuing Networks". Journal of the ACM. **27** (2): 313. doi:10.1145/322186.322195.

25. Van Dijk, N. M. (1993). "On the arrival theorem for communication networks". Computer Networks and ISDN Systems. **25** (10): 1135–2013. doi:10.1016/0169-7552(93)90073-D.

26. Gordon, W. J.; Newell, G. F. (1967). "Closed Queuing Systems with Exponential Servers". Operations Research. **15** (2): 254. doi:10.1287/opre.15.2.254. JSTOR 168557.

27. Baskett, F.; Chandy, K. Mani; Muntz, R.R.; Palacios, F.G. (1975). "Open, closed and mixed networks of queues with different classes of customers". Journal of the ACM. **22** (2): 248–260. doi:10.1145/321879.321887.

28. Buzen, J. P. (1973). "Computational algorithms for closed queueing networks with exponential servers" (PDF). Communications of the ACM. **16** (9): 527. doi:10.1145/362342.362345.

29. Kelly, F. P. (1975). "Networks of Queues with Customers of Different Types". Journal of Applied Probability. **12** (3): 542–554. doi:10.2307/3212869. JSTOR 3212869.

30. Gelenbe, Erol (Sep 1993). "G-Networks with Triggered Customer Movement". Journal of Applied Probability. **30** (3): 742–748. doi:10.2307/3214781. JSTOR 3214781.

31. Bobbio, A.; Gribaudo, M.; Telek, M. S. (2008). "Analysis of Large Scale Interacting Systems by Mean Field Method". 2008 Fifth International Conference on Quantitative Evaluation of Systems. p. 215. doi:10.1109/QEST.2008.47. ISBN 978-0-7695-3360-5.

32. Bramson, M. (1999). "A stable queueing network with unstable fluid model". The Annals of Applied Probability. **9** (3): 818. doi:10.1214/aoap/1029962815. JSTOR 2667284.

33. Chen, H.; Whitt, W. (1993). "Diffusion approximations for open queueing networks with service interruptions". Queueing Systems. **13** (4): 335. doi:10.1007/BF01149260.

34. Yamada, K. (1995). "Diffusion Approximation for Open State-Dependent Queueing Networks in the Heavy Traffic Situation". The Annals of Applied Probability. **5** (4): 958. doi:10.1214/aoap/1177004602. JSTOR 2245101.

35. Bertoli, Marco; Casale, Giuliano; Serazzi, Giuseppe (2009). "JMT: performance engineering tools for system modeling". SIGMETRICS Perform. Eval. Rev. **36**: 10–15. doi:10.1145/1530873.1530877.

36. Marzolla, Moreno. "The queueing Package" (PDF). Retrieved 31 July 2014.

37. Müller, Klaus; Vignaux, Tony. "SimPy Acknowledgements". Retrieved 24 May 2016.

38. "Queueing Processes—Wolfram Language Documentation". reference.wolfram.com. Retrieved 2016-06-23.

39. "PDQ-R: {\em Pretty Damn Quick} for R Statistical Computing". www.perfdynamics.com. Retrieved 2016-06-23.

40. "Queuing - MATLAB & Simulink". www.mathworks.com. Retrieved 2016-06-23.