BASIC CHARACTERISTIC OF A PRODUCT FORM QUEUING NETWORK

RANDRIAMAMPIANINA Sitraka Sedera Nandrasana[1], HERINANTENAINA Edmond Fils[2], RAKOTOMIRAHO Soloniaina[3]

[1] *PhD student, SE-I-MSDE, ED-STII, Antananarivo, Madagascar*
[2] *Thesis co-director, SE-I-MSDE, ED-STII, Antananarivo, Madagascar*
[3] *Thesis director, SE-I-MSDE, ED-STII, Antananarivo, Madagascar*

ABSTRACT

This paper describes the basic parameters of a product form queue networks, its performance and stability parameters, and queue network categories it aims to develop a reliable model of queuing network.

Keywords : Queue networks, product form, routing network, Kendhal notation

---

1. INTRODUCTION

Queues are visible everywhere, in everyday life. Queue networks, on the other hand, are not directly visible. But by studying its performance parameters, it can result in a lot of simplification in several areas, especially in the use of the internet.

Currently, networks extend to several thousand or even millions of nodes. What poses a problem during modeling is surely the calculation time that the administrator has to do as well as several tests to evaluate the period of stability as well as the performance of the network to be tested. This work consists of studying a network of queues by knowing its basic input parameters. But it is still necessary to know each queue that makes up the network and to assess its performance and stability conditions. It is only after this that the administrator can design a model of the network.

This document describes the generality of queuing networks. And before using a queuing network, it is necessary to know its types and its characteristic parameters in order to prevent possible problems linked to the random behavior of the system to better adapt to the need for the realization of a model.

2. CHARACTERISTICS OF A QUEUE

A queue is defined as a system where the "client" arrives from outside, processed in the system by one or more servers, and then exits the system. This action can be represented by Fig-1 which is the universal form of a queue [1].

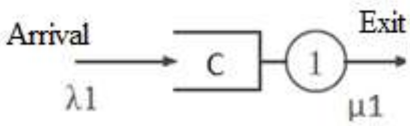

Fig-1: Representation of a queue.

λ_1 represents the arrival rate, μ_1 the service rate and C the server capacity. The circle with a number inside means the number of servers in the queue.

A queue differs from another queue, but they are all characterized by special parameters: the arrival process as well as the average arrival number, the service process as well as the service rate d. 'a queue, the number of servers, the capacity of the queue. By convention, if the capacity of the queue is omitted, it is considered to be infinite.

The Poisson law with parameter λ is the law that governs the arrival of customers in the queue, it is a discrete random variable X with positive values defined by its state probabilities:

$$P(X = n) = (\frac{\lambda^n}{n!})e^{-\lambda}$$

for n=0, 1, 2,…

It is the law of small probabilities or the law of rare and memoryless events in a given time interval.

A continuous time chain N is a Poisson process with parameter λ if and only if N is a counting process, N is a stationary process with independent increments. N (t) is distributed according to a Poisson law with parameter λt and has as probability [2].

$$P(N(t) = k) = (\frac{\lambda t^k}{k!})e^{-\lambda t}$$

for k=0, 1, 2,…

2.1. Kendhal notation

Each queue can be characterized according to the notation A/B/C/K/m/Z which is called "Kendhal notation".

Table 1 : Summary of meanings of these letters.

LETTER	MEANINGS
A	Law of the customer arrival process, or the law of time between two successive arrivals. It is described using a counting process.
B	Law that governs the service or law of the length of service of the queue
C	Number of servers in the queue, the minimum number of servers is one
K	Capacity, buffer or buffer of the queue. It is infinite if it is not defined. If the queue is at finite capacity, any arriving customers will be lost if the queue is full.
m	Population of users, clients. It is infinite if it is not indicated
Z	Discipline of service. The possible disciplines are FIFO, LIFO, Random,…

By default, K and m are infinite and Z is FIFO discipline. This notation is universal for describing a queue with its characteristics.

The arrival of customers is also known as the "arrival process". It is characterized by a probability density function of the time between arrivals. The possible parameters A and B are summarized in Table 2.

Table 2 : Types of arrival process

TYPES OF ARRIVAL	DESCRIPTIONS
M (Markov)	Exponential or Poissonian distribution
D (Deterministic)	Uniform distribution, i.e. customers arrive at regular intervals
G (General)	Other type of distribution
GI(General ndependent)	Other type of distribution with independent inter-arrival
H	Geometric distribution
Ek	Distribution according to an Erlang-k law

The residence time R of N packets arriving in a system with an arrival rate λ is given by the formula:

λ N= R

This formula has a general character and remains valid for average values of R, λ and N.

2.2. M/M/1 queue

It is the most frequent queue in the modeling of a computer system with a single server, a Poissonian arrival process of rate λ and an exponential service law. The capacity of the queue is endless.

Fig-2 shows an M/M/1 queue.

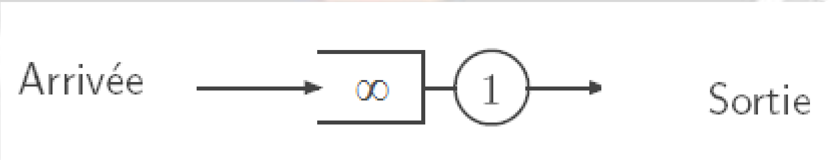

Fig-2: Representation of an M/M/1 type queue.

Its parameters are summarized in Table3

Table 3: **Parameters of a queue type M/M/1.**

PARAMETERS	**FORMULAS**
Average flow	$D = [1 - P(0)]\mu = \rho\mu = \lambda$
Use rate	$U = [1 - P(0)] = \rho$
Average number of customers	$N = \frac{\rho}{(1-\rho)}$
Average crossing time	$R = \frac{N}{D} = \frac{1}{\mu(1-\rho)} = \frac{1}{\mu} + \frac{\rho}{\mu(1-\rho)}$
Number of server	1

2.3. M/M/K queue

This queue is characterized by its Poissonian arrival process and exponential service law, but the number of servers is K and the capacity of the queue is still infinite.

Fig-3 shows a M/M/K file.

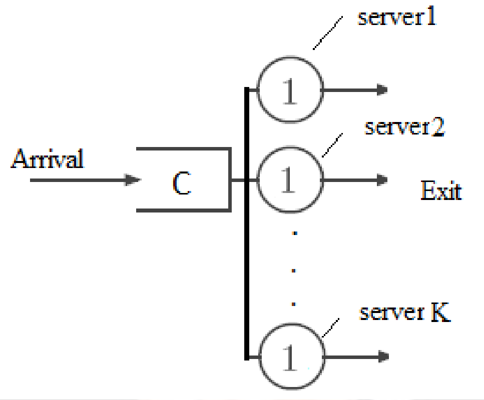

Fig-3: Representation of an M/M/K type queue.

Its parameters are summarized in Table4.

Table 4: **Parameters of an M / M / K type queue.**

PARAMETERS	**FORMULAS**
Average flow	D=λ
Average number of customers	N=R*D=R* λ
Average crossing time	$R = P(0)\frac{\rho^k}{\mu(k-\rho)^2(k-1)!} + \frac{1}{\mu}$
Number of server	K

2.4. G/G/1 queue

Consider the following general case:

The G/G/1 queue is characterized by a General arrival and service process, with a single server and infinite capacity. However, for a general arrival process, the parameters of the queue can only be calculated approximately, they are however framed by two inequalities.

The average waiting time in the queue is denoted W with:

$$\frac{\lambda\sigma_X^2 - X(2-\rho)}{2(1-\rho)} \leq W \leq \frac{\lambda(\sigma_T^2 + \sigma_X^2}{2(1-\rho)}$$

where X is the average service time.

σ_X: the standard deviation of the random variable that describes the service time.

σ_T: the standard deviation of the random variable that describes the inter-arrivals.

By setting the upper bound W_{sup} of W

$$W_{sup} \leq \frac{\lambda(\sigma_T^2 + \sigma_X^2}{2(1-\rho)}$$

The previous equation can be written :

$$W_{sup} - \frac{1+\rho}{2\lambda} \leq W \leq W_{sup}$$

Where W_{sup} is the upper bound of W.

By applying Little's formula, the average number of N_{buffer} clients waiting in the buffer with $0 \leq \rho \leq 1$ is given by

$$\lambda W_{sup} - \frac{1+\rho}{2} \leq N_{buffer} \leq \lambda W_{sup}$$

3. OVERVIEW QUEUING NETWORK

A system that has more than one queue is said to be a network of queues. This is the case with the Internet, which is the largest computer network nowadays [3]. Fig-4 shows a network of three queues with the three service rates μ_1, μ_2 et μ_3.

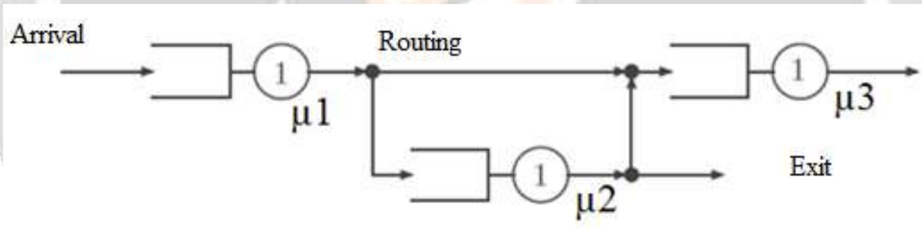

Fig-4: Example of a three-queue network.

This is only an example, but the type of routing may vary depending on the network. Indeed, the two most famous types of routing are: ring routing and star routing. But as with the Internet, these two types of routing can coexist.

Ring routing can be summed up as routing where clients that leave a queue do not return to it unless they have been served by all the other servers in each queue that make up the network, i.e. say that service only goes one way (Fig-5).

Star routing is more complex than the previous one, but it is the most general and widespread case. Indeed, in this configuration, there is always a probability that a customer already processed in a queue i will return to queue i-1. Routing is done in both directions (Fig-6).

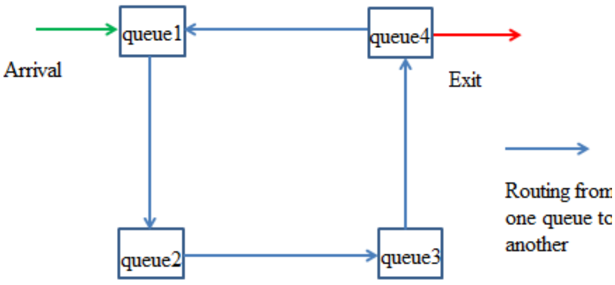

Fig-5 : Queue network ring routing

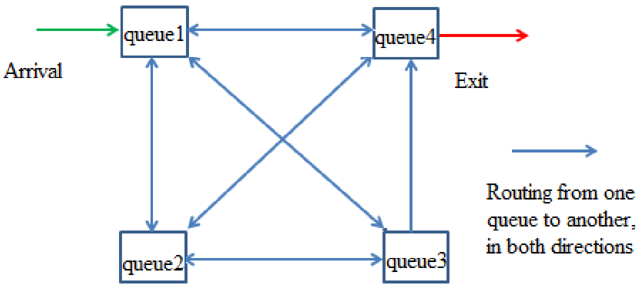

Fig-6 : Queue network star routing.

4. MAIN TYPES

There are several types of queue networks, depending on the structure and operation as well as the needs of the system in question. It may happen that these different types are mixed to form yet other types of network [4].

4.1. Open network

An **open network** is a network where customers come from outside, progress through the system to be processed by the nodes, and then leave it. (Fig-7).

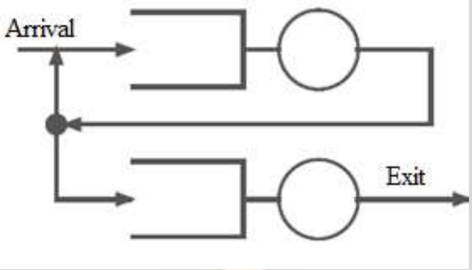

Fig-7: Open queue network.

This is the most common case since entry can be considered as an arrival and exit as a departure. In summary, a network where any client present or entering the system, able to leave the system is an open network.

4.2. Closed network

The closed network is a network where the number of clients is fixed at all times. This means that customers are caught in one or even loops and never get out. The exit becomes the entrance and the entrance becomes the exit. This type of queue network is used in network troubleshooting, by not letting out any packets or messages to see or estimate any hardware or software problems. (Fig-8).

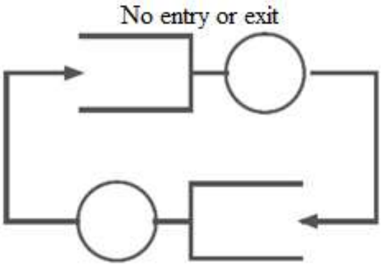

Fig-8: Closed network of queues.

4.3. Network with class

The **Network with class** is a network where the processing priorities of messages or packets in the system have the same processing priorities (single-class) or are of different priorities (multi-class). Classes define the order of priority for dealing with customers. The highest class is processed first, then other classes of clients will be processed. This type of class can exist in open networks or closed networks.

A **network** is said to be **mixed** if it is open for certain classes of customers and closed for others.

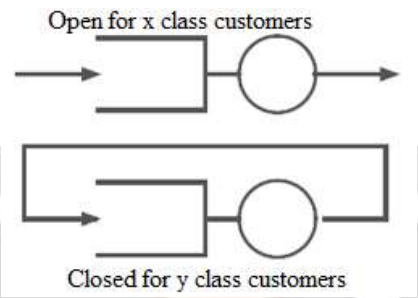

Fig-9: Mixed queue network.

Note

Mixtures may exist between these class networks and open and closed networks. For example, it is possible to have an open single-class network, a closed multi-class network, and so on.

5. RELEVANT NETWORK PARAMETERS

Queuing networks are characterized by:

a) ***Their queue management policy (FIFO, LIFO,…),***
b) ***The probability distributions of service times,***
c) ***The capacity of each queue***
d) ***The law of arrival of packets to the network.***

And to measure their performance, it is necessary to know:

- the rate of use (occupancy) of the server i which results in the formula

$$\rho_i = \frac{\lambda_i}{\mu_i}$$

- the average number of customers in station i which results in the formula

$$N_i = \frac{\lambda_i}{(\mu_i - \lambda_i)}$$

- the average stay time in a station i which is expressed by the formula

$$R_i = \frac{N_i}{\lambda_i} = \frac{1}{(\mu_i - \lambda_i)}$$

- Average residence time in the network which results in the formula

$$\mathrm{T} = \frac{N}{R}$$

6. DIFFERENT CATEGORIES

Queue networks fall into two broad categories: product form and non-product form queue networks [5].

6.1. Product form queue network

A product form queue network is a network whose equilibrium probability can be decomposed into the product of all the steady-state probabilities of all the nodes in the system.

6.2. Non-product form queue network

A **non-product form** queue network is a network whose steady-state probabilities cannot be expressed as a product to give the steady-state probability of the entire system.

Note :

For the simulation, the steady-state probability, which should be unique for a product-shaped queue network is not calculated with the real data since this probability requires knowing each number of clients present in each node in real time. The average number of customers in each queue will be taken to calculate this probability, which may be slightly different from the real value of this probability.

7. CONDITION OF STABILITY

It is interesting to model a network of queues, but knowing its stability limits is necessary in order not to quickly saturate the system and to collect non-erroneous data. A system is only stable in a steady state [6].

A network of queues is stable provided that all of the queues that make up the network are all stable. If λ_j is the average arrival rate in queue j, the conservation equations are written:

$$\lambda_j = \gamma_j + \sum_{i=1}^{n} \lambda_i r_{ij}$$

Where j varies from 1 to n (n is the number of queues).

So, if there are n queues, the average arrival rate is written:

$$\lambda_n = \gamma_n + r_{1n}\lambda_1 + r_{2n}\lambda_2 + r_{3n}\lambda_3 + \ldots + r_{nn}\lambda_n$$

In matrix form, these equations become:

$$(\lambda) = (\gamma) + (\lambda)(R) \text{ then } \lambda\,(1 - R) = \gamma$$

For an open network:

$$\lim_{n \to +\infty} R^n = 0$$

The solution of the previous equation is

$$\lambda = \frac{\gamma}{1 - R}$$

The stability condition of queue j is given by the following relation, if λ is known:

$$\rho_j = \frac{\lambda_j}{m_j \mu_j} < 1$$

ρ_j is also called the utilization rate of queue j.
Factorizing the λ is the method adopted to facilitate the resolution in this work.
Equation therefore becomes

$$\begin{cases} (r_{11}-1)\lambda_1 + r_{21}\lambda_2 + r_{31}\lambda_3 + \cdots + r_{j1}\lambda_j = -\gamma_1 \\ r_{12}\lambda_1 + (r_{22}-1)\lambda_2 + r_{32}\lambda_3 + \cdots + r_{j2}\lambda_j = -\gamma_2 \\ r_{13}\lambda_1 + r_{23}\lambda_2 + (r_{33}-1)\lambda_3 + \cdots + r_{j3}\lambda_j = -\gamma_3 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ r_{1j}\lambda_1 + r_{2j}\lambda_2 + r_{3j}\lambda_3 + \cdots + (r_{jj}-1)\lambda_j = -\gamma_j \end{cases}$$

It has already been mentioned that for a system to be stable, all the queues that constitute it must be stable. The stability check for each queue has already been seen in the first part of this chapter. Which give:

$$\begin{cases} \rho_1 = \dfrac{\lambda_1}{m_1\mu_1} < 1 \\ \rho_2 = \dfrac{\lambda_2}{m_2\mu_2} < 1 \\ \dots\dots\dots\dots \\ \rho_n = \dfrac{\lambda_n}{m_n\mu_n} < 1 \end{cases}$$

with n the number of queues in the network.
Now is the time to calculate the performance indices of the network which are N, R since
$\lambda = \gamma_1$ because there is only one external input.
For $N_1, N_2, N_3 \dots N_j$

$$\begin{cases} N_1 = \dfrac{\rho_1}{1-\rho_1} \\ N_2 = \dfrac{\rho_2}{1-\rho_2} \\ N_3 = \dfrac{\rho_3}{1-\rho_3} \\ \dots\dots\dots\dots\dots \\ N_J = \dfrac{\rho_J}{1-\rho_J} \end{cases}$$

And

$$N = N_1 + N_2 + N_3 + \dots\dots\dots\dots + N_J$$

Average residence time :

$$R = \frac{N}{\lambda} = \frac{N_1 + N_2 + N_3 + \cdots + N_J}{\gamma_1 + \gamma_2 + \gamma_3 + \cdots + \gamma_J}$$

The stationary distribution of the network is as follows:

$$\prod(x_1 + x_2 + x_3 + \cdots + x_j) = [(1-\rho_1)\rho_1{}^{x_1}][(1-\rho_2)\rho_2{}^{x_2}][(1-\rho_3)\rho_3{}^{x_3}] \dots [(1-\rho_j)\rho_j{}^{x_j}]$$

where x_j represent, respectively, the number of customers in the j queue.

8. SPECIAL CASE: JACKSON NETWORK

A Jackson's network is a special case of a queue network composed of n queues of the exponential type.

Each of the queues in the network has one or more identical servers (mj for queue j) as the case may be, providing exponential-type services (the service rate of queue j is denoted µj) and the type of arrival is Poissonian. The capacity of each queue, therefore of the network, is infinite. The service discipline is FIFO. This type of network is of the open mono-class type and the routings between the queues are probabilistic.

A customer who has finished his service at station j moves to station k with probability r_{jk} or leaves the system with probability r_{j0}.

Where

$$r_{j0} = 1 - \sum_{k=1}^{n} r_{jk}$$

Note:

A Jackson network is an open network in product form. Another case is the Gordon-Newell network that is similar to a Jackson closed network. This type of network has only one class of client. Each queue that composes it has only one server. The service law is exponential for all the queues that make up the network. the capacity of each queue, therefore of the network, is infinite, that is to say that there is no loss of customer [7]. The discipline which governs this network is FIFO type, the routings are probabilistic. So almost identical to the properties of the open Jackson lattice, except that the latter is a closed lattice.

9. CONCLUSION

This document has shown two main categories of network: product form and non-product form. The problem is the non-product form queue networks. Indeed, the total probability of the system in the stationary state cannot be decomposed into the product of all the probabilities of the nodes constituting the system. We will have to find a method to simplify the modeling of queuing networks as much as possible.

10. REFERENCES

[1] Randriamampianina Sitraka Sedera Nandrasana, « *Etude d'une file d'attente dans un réseau WAN par modèle analytique* », Mémoire de fin d'études- Département Electronique-EA, ESPA-Université d'Antananarivo, 2009.

[2] Franck Jedrzejewski, « *Modèles aléatoires et physique probabiliste* », Springer, Juillet 2009

[3] Henri Nussbaumer, « *Téléinformatique 2 : Conception des réseaux-Réseau-Transport* », Collection Informatique, Lausanne 1987

[4] Stephan Robert, « *Modélisation stochastique, Files d'attente* », Master of Science in Ingeneering, HEIG-Vd, 30 novembre 2009.

[5] G Pujolle, « *Réseau de files d'attente à forme produit* », Recherche Opérationnelle, Tome 14, N°4, p.317-330, AFCET, 1980

[6] Ahmed Harbaoui, «*Vers une modélisation et un dimensionnement automatiques des applications réparties* », THÈSE Pour obtenir le grade de DOCTEUR en Informatique DE L'UNIVERSITÉ DE GRENOBLE, 21 Octobre 2011.

[7] « *Architecture d'une boîte à outils d'algorithmes d'ingénierie de trafic et application au réseau GÉANT* », ftp://ftp.run.montefiore.ulg.ac.be/pub/RUN-PP05-02.pdf