STUDY OF A NETWORK OF THREE QUEUE BY ANALYTICAL METHOD

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 a study of a three-queue network by analytical method. Its stability parameters will be highlighted in order to arrive at a more stable network model.

Keyword : Queue networks, product form, routing network, analytical method

---

1. INTRODUCTION

For a large network like the Internet, modeling will take a long time. The purpose of this chapter is to model a network of queues, and to predict the stability and the maximum load supported by the system. This automation will be done by load test or by increment.

Performing this work involved the use of an algorithm to make modeling a network of queues easy, that is, without much manual intervention from a human administrator. The detection of bottlenecks remains to be specified in future work.

2. CONSIDERED ARCHITECTURE

Fig1 shows a network with three queues and its. Queues each have a single server, but in practice the number of servers in a queue is at least one[1].

This network has three entries Ei to queue i, for i varying from 1 to 3. For each entry Ei, the external arrival rate is γi and the average arrival rate is λi.

Likewise, there are three exits Si with the probability ri0 of going from queue i to the exit.

From the node point of view, the network has three loops on the same queue (on queue i, the probability of arrival is rii), normal interconnections (from queue i to queue j with the probability rij), interconnections back (from queue j to queue i with the probability rji where j is less than i).

Fig.1 shows the general case of a three-queue network with its parameters.

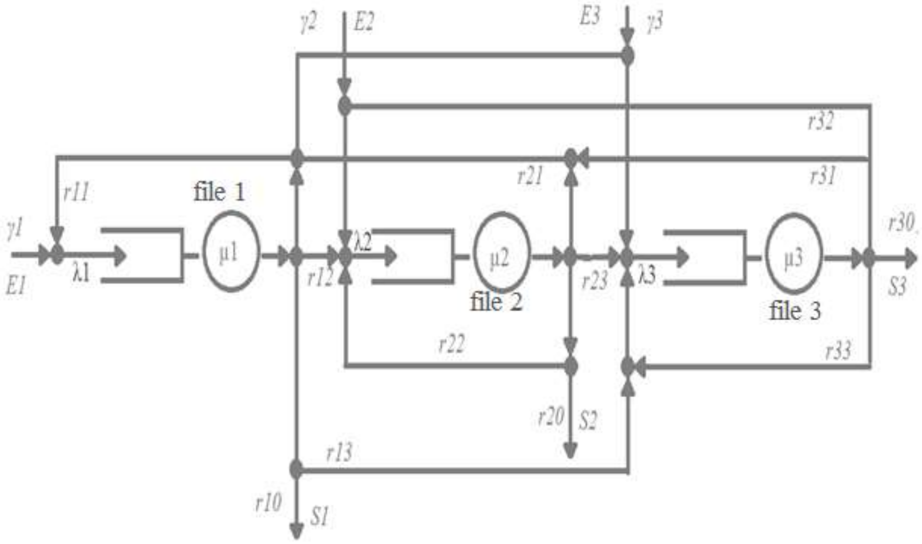

Fig.1: General case of a three-queue network.

The service rates are μi for queue i [2].
And the equilibrium probabilities for the outputs are such as:

$$\begin{cases} r_{10} + r_{11} + r_{12} + r_{13} = 1 \\ r_{20} + r_{12} + r_{22} + r_{23} = 1 \\ r_{30} + r_{13} + r_{32} + r_{33} = 1 \end{cases}$$

where ri0 are the probabilities of exiting the network from queue i. If there is only one S3 exit (from queue 3) then r10 and r20 will be zero.

Note :
For a closed network, the γi as well as the probabilities ri0 are all zero. That is, there is no external input and there is no output.

3. SIMPLIFYING HYPOTHESIS

a) ARCHITECTURE

The consideration of all possible routing for this study can influence the computation time and the system resources. Arbitrary simplifications have been imposed.

Indeed, there is only one input E1 and one output S3. Also, some routings have been eliminated, reducing some probabilities to zero, namely: r11, r13, r33, as well as the output probabilities r10 and r20.

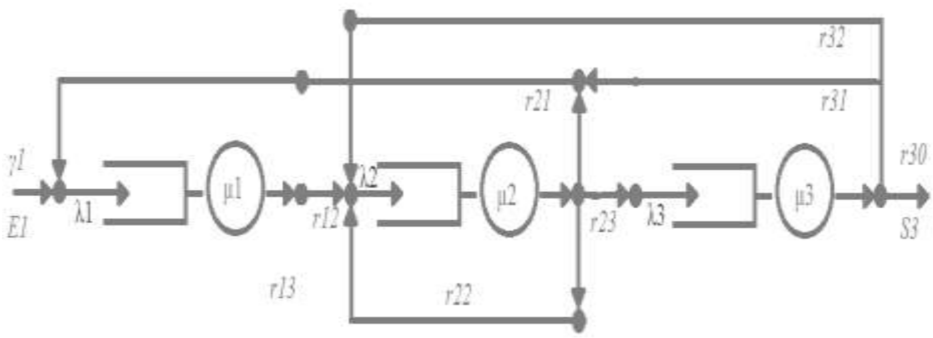

Fig.2: Three-queue network studied

The probability of going from lane 1 to lane 2 is one because that is the only possible direction from lane 1. The sum of the probabilities of going from lane 2 to all lanes is one. The sum of the probabilities of going from lane 3 to the exit and all lanes is one.

On the other hand, the external arrival rates $\gamma 1$ and $\gamma 2$ are zero since there is no external arrival in queues 2 and 3.

b) FORM OF THE NETWORK

This approximation can be summed up by replacing a network with a non-produced form by a network with a produced form while keeping the same topology, changing each general service law by exponential service laws and a rate μi (n) depending on the charge.

Next, it is necessary to determine the arrival rates depending on the load λi (n), by short-circuiting station i and replacing all the other stations with a station C. It is also necessary to analyze each node of the network and determine the rate λi (n) of station i with n clients and the stationary probabilities pi (n).

Note :

All the results will be average values of the performance indices of the network and its stability conditions. Indeed, the state considered is the stationary state.

c) STABILITY OF EACH QUEUE

For a queue network to be considered stable, every node in the network must be stable. On the other hand, stability is only defined in a steady state. It will be considered that each node of the system to be designed has been manually verified and that they are all stable. Its stability in the system will then depend on the parameters of arrival from the previous node.

d) TYPE OF NETWORK

The network to be modeled is an open queue network. Indeed, a closed network is only considered in the case of network troubleshooting but in our case we are in the phase of modeling a network for use with several clients. In addition, the number of clients in a closed network is constant at all times since there are no additional clients, and the clients do not leave the network.

4. STAGES OF THE STUDY

To model a network of queues and have the result by simulation, it is necessary to have an algorithm which can extract the number of servers required, as well as the saturation load which is necessary to estimate the period of stability of the system. . But there are still some simplifications to be made. Indeed, if we consider all network cases, there should be several types of algorithm, which would increase the response time of the simulation.

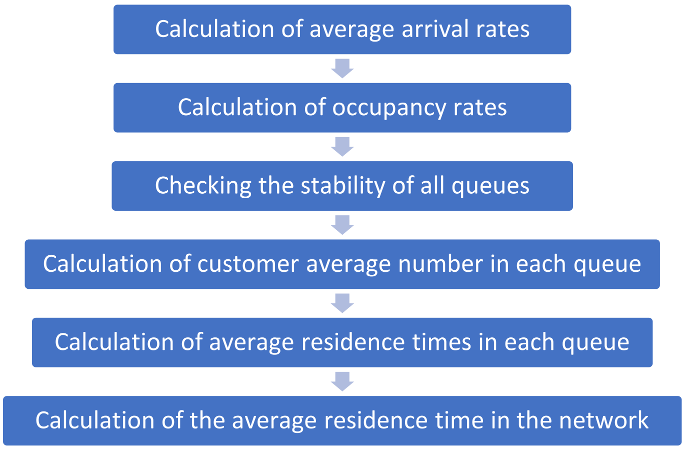

Fig.3: Functional diagram of the test process.

Consequently, this presents a lot of approximation in our case. But if the number of clients is large, the solution obtained will be approximate but the response time small [3].

It has already been seen that

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

with j = 1, ..., 3 and n = 3 the number of queues.

By developing this formula with three queues, the result is:

$$\begin{cases} \lambda_1 = \gamma_1 + r_{11}\lambda_1 + r_{21}\lambda_2 + r_{31}\lambda_3 \\ \lambda_2 = \gamma_2 + r_{12}\lambda_1 + r_{21}\lambda_2 + r_{31}\lambda_3 \\ \lambda_3 = \gamma_3 + r_{13}\lambda_1 + r_{21}\lambda_2 + r_{31}\lambda_3 \end{cases} \qquad \begin{cases} (r_{11} - 1)\lambda_1 + r_{21}\lambda_2 + r_{31}\lambda_3 = -\gamma_1 \\ r_{12}\lambda_1 + (r_{22} - 1)\lambda_2 + r_{32}\lambda_3 = -\gamma_2 \\ r_{13}\lambda_1 + r_{23}\lambda_2 + (r_{33} - 1)\lambda_3 = -\gamma_3 \end{cases}$$

For known r_{11}, r_{21}, r_{31}, r_{21}, r_{22}, r_{23}, r_{31}, r_{32}, r_{33}, γ_1, γ_2 and γ_3.

Using the Gaussian n-unknown equation method, it is possible to find the λ_n.

For a system to be stable, all the queues that make it up must be stable. This leads to

$$\begin{cases} \rho_1 = \dfrac{\lambda_1}{m_1\mu_1} < 1 \\ \rho_2 = \dfrac{\lambda_2}{m_2\mu_2} < 1 \\ \dots\dots\dots\dots \\ \rho_n = \dfrac{\lambda_n}{m_n\mu_n} < 1 \end{cases} \qquad \begin{cases} N_1 = \dfrac{\rho_1}{(1-\rho_1)} \\ N_2 = \dfrac{\rho_2}{(1-\rho_2)} \\ N_3 = \dfrac{\rho_3}{(1-\rho_3)} \end{cases}$$

for a three-queue network[4].

with m1, m2 and m3 are respectively the number of servers in queue 1, queue 2 and queue 3 and that µ1, µ2 and µ3 are respectively the average arrival rates of queue 1, queue 2 and queue 3. These data are previously given.

Average residence time

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

Because γ_2 and γ_3 are zero then $\lambda = \gamma_1$

The steady-state probability is therefore[5]

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

Where $x_1 \geq 0$,$x_2 \geq 0$ et $x_3 \geq 0$ represent, respectively, the number of customers in the first, second and third rows.

For what follows, the most important parameters are: on input the external arrival rates to be able to extract the average arrival rates in each lane, as well as the utilization rates or stability conditions at the output to establish the period system stability.

a) EQUIVALENT NETWORK

The entire network can be thought of as a single queue with equivalent parameters[6].

The average arrival rate λ (since there is only one entry in the network) is still displayed. Finally, the probability of the stationary distribution is displayed last, because it is the product of all the steady-state probabilities of all the queues. Fig.4 illustrates the assimilation of the system into a single queue.

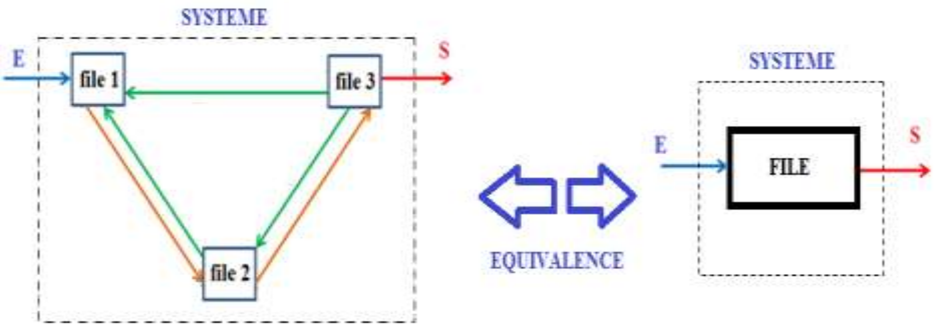

Fig.4 : Network equivalent to a queue.

5. RESULTS

For 3 queues, here is the conservation equation:

$$\begin{cases} \lambda_1 = \gamma_1 + r_{11}\lambda_1 + r_{21}\lambda_2 + r_{31}\lambda_3 \\ \lambda_2 = \gamma_2 + r_{12}\lambda_1 + r_{21}\lambda_2 + r_{31}\lambda_3 \\ \lambda_3 = \gamma_3 + r_{13}\lambda_1 + r_{21}\lambda_2 + r_{31}\lambda_3 \end{cases} \qquad \begin{cases} (r_{11}-1)\lambda_1 + r_{21}\lambda_2 + r_{31}\lambda_3 = -\gamma_1 \\ r_{12}\lambda_1 + (r_{22}-1)\lambda_2 + r_{32}\lambda_3 = -\gamma_2 \; where \; \gamma_2 = 0 \\ r_{13}\lambda_1 + r_{23}\lambda_2 + (r_{33}-1)\lambda_3 = -\gamma_3 \; where \; \gamma_3 = 0 \end{cases}$$

The unknowns are λ_1, λ_2 et λ_3.

a) SYSTEM STABILITY

ρ_i with i varying from one to three, is the rate of use or condition of stability of each queue. If its value is greater than or equal to one, then the system is unstable [7].

By considering several tests, the system can be stable or unstable.

i) Unstable System

Numerical data is represented in Tab1:

Tab 1: **Input parameters for unstable network simulation.**

QUEUE 1	*QUEUE 2*	*QUEUE 3*
r_{11} =0	r_{21} =0.25	r_{31} =0.125
r_{12} =1	r_{22} =0.25	r_{32} =0.125
r_{13} =0	r_{23} =0.5	r_{33} =0.25
m_1 =4	m_2 =6	m_3 =5
g_1 =1500	g_2 =0	g_3 =0
μ_1 =800	μ_2 =600	μ_3 =600

So according to Tab8 and after calculation, the average arrival rates are given by
ρ1 = λ_1/ m_1 μ_1=3000/4x800

$$\begin{cases}\lambda_1 = 3000\\ \lambda_2 = 4500\\ \lambda_3 = 3000\end{cases} \quad \begin{cases}\rho_1 = 0,9375\\ \rho_2 = 1,25\\ \rho_3 = 1\end{cases} \quad \begin{cases}N_1 = 15\\ N_2 = -5\\ N_3 = \infty\end{cases}$$

So the system is unstable, the instability of the system can cause calculation errors which could give unpredictable results.

Et N = N_1 + N_2 + N_3 = ∞

R = T = N/ λ = ∞/ 1500 = ∞

i) Stable System

Numerical data is represented in Tab2:

<u>*Tab 2*</u>: **Input parameters for stable network simulation.**

QUEUE 1	***QUEUE 2***	***QUEUE 3***
r_{11} =0	r_{21} =0.25	r_{31} =0.125
r_{12} =1	r_{22} =0.25	r_{32} =0.125
r_{13} =0	r_{23} =0.5	r_{33} =0.25
m_1 =4	m_2 =6	m_3 =5
g_1 =900	g_2 =0	g_3 =0
μ_1 =800	μ_2 =800	μ_3 =700

For the calculation of the probability of the stationary distribution, the average values of the numbers of clients in each queue were taken. Which gives after calculation and according to Tab9, the external arrival rate, the utilization rate and the average number of customers in each queue:

$$\begin{cases}\lambda_1 = 1800\\ \lambda_2 = 2700\\ \lambda_3 = 1800\end{cases} \quad \begin{cases}\rho_1 = 0,5625\\ \rho_2 = 0,5625\\ \rho_3 = 0,5143\end{cases} \quad \begin{cases}N_1 = 1,2857\\ N_2 = 1,2857\\ N_3 = 1,0589\end{cases} \quad \begin{cases}\pi_1 = 0,20879\\ \pi_2 = 0,20879\\ \pi_3 = 0,2402\end{cases}$$

And N = N_1 + N_2 + N_3 = 3,6303

R = T = N/ λ = 3,6303/ 900 = 0,004

Either π = π_1 x π_2 x π3

Then π = 0,1003

The data as well as the results of the two simulations can be compared from Tab3 and Tab4 [8]:

<u>*Tab3*</u>: **Initial data for simulations and analytical calculations:**

STABLE SYSTEM			***UNSTABLE SYSTEM***		
QUEUE 1	***QUEUE 2***	***QUEUE 3***	***QUEUE 1***	***QUEUE 2***	***QUEUE 3***
r_{11} =0	r_{21} =0.25	r_{31} =0.125	r_{11} =0	r_{21} =0.25	r_{31} =0.125
r_{12} =1	r_{22} =0.25	r_{32} =0.125	r_{12} =1	r_{22} =0.25	r_{32} =0.125
r_{13} =0	r_{23} =0.5	r_{33} =0.25	r_{13} =0	r_{23} =0.5	r_{33} =0.25
m_1 =4	m_2 =6	m_3 =5	m_1 =4	m_2 =6	m_3 =5
g_1 =900	g_2 =0	g_3 =0	g_1 =1500	g_2 =0	g_3 =0
μ_1 =800	μ_2 =800	μ_3 =700	μ_1 =800	μ_2 =600	μ_3 =600

Tab4: **Results of the two analytical calculations.**

STABLE SYSTEM			***UNSTABLE SYSTEM***		
QUEUE 1	***QUEUE 2***	***QUEUE 3***	***QUEUE 1***	***QUEUE 2***	***QUEUE 3***
N1 = 4	N2 = 6	N3 = 5	N1 = 4	N2 = 6	N3 = 5
R1 = 0.00071	R2 = 0.00047	R3 = 0.00058	R1 = 0.005	R2 = 0.0033	R3 = ∞
λ_1 = 1800	λ_2 = 2700	λ_3 = 1800	λ_1 = 3000	λ_2 = 4500	λ_3 = 3000
ρ_1 = 0.5625	ρ_2 = 0.5625	ρ_3 = 0.5142	ρ_1 = 0.9375	ρ_2 = 0.9375	ρ_3 = 1
π1 = 0.2088	π2 = 0.2088	π3 = 0.2065	π1 = 0.0237	π2 = 0.0237	π3 = 0
N = 3.63	λ = 1800		N = ∞	λ = 3000	
R = 0.004	π = 0.009		R = ∞	π = 0	

After several tests, the following two curves show the evolution of the stability of the system.

Exceeding an arrival rate of 1100, the system is unstable but the first line is still stable. This is explained by the instability of the second or third queue of the network. The calculations are wrong, especially for the average number of customers in the network as well as the average transit time. From Figure 3.8, row # 2 tends to instability more quickly. This is after the # 1 row becomes unstable in turn, then the third row is the last to be unstable.

Note:

The number of clients N1 as well as the average crossing time R1 of the first queue have been demonstrated because only the first queue has external arrivals.

The number of average customers in the network grows exponentially to an arrival rate of 1,100. This means that the expectations in the system increase as the system can no longer maintain quality of service. If the arrival rate exceeds 1100, the calculations are no longer correct due to the instability of the system.

b) CROSSING TIME

The results are summarized in Tab5 and in Fig.5;

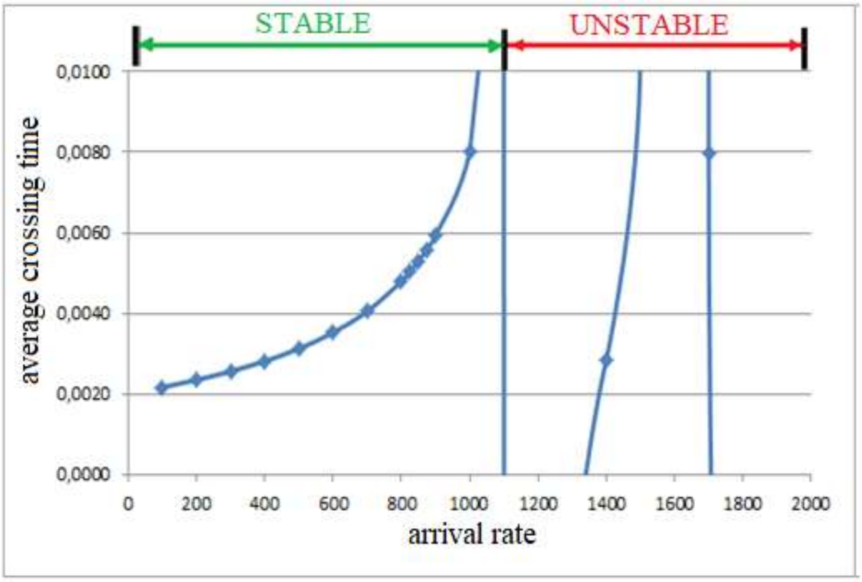

Fig.5 : Average crossing time depending on the arrival rate.

Tab5: **Miscellaneous results after variation in the arrival rate**

EA	***N***	***R***	***λ***	***STABILITY***
100	0,218	0,0022	200	STABLE
200	0,472	0,0024	400	STABLE
300	0,771	0,0026	600	STABLE
400	1,130	0,0028	800	STABLE
500	1,569	0,0031	1000	STABLE
600	2,122	0,0035	1200	STABLE
700	2,844	0,0041	1400	STABLE
800	3,842	0,0048	1600	STABLE
825	4,156	0,0050	1650	STABLE
850	4,506	0,0053	1700	STABLE
875	4,899	0,0056	1750	STABLE
900	5,345	0,0059	1800	STABLE
1000	8,000	0,0080	2000	STABLE
1100	14,892	0,0135	2200	STABLE
1200	-12496	-10,413	2400	UNSTABLE
1300	-5,766	-0,0044	2600	UNSTABLE
1400	4,000	0,0029	2800	UNSTABLE
1500	16,000	0,0107	3000	UNSTABLE
1600	10000	10000	3200	UNSTABLE
1700	13,541	0,0080	3400	UNSTABLE
1800	-48,063	-0,0267	3600	UNSTABLE
1900	-21,721	-0,0114	3800	UNSTABLE
2000	-15,503	-0,0077	4000	UNSTABLE

Average transit times increase exponentially up to an arrival rate of 1100. As the arrival rate increases and the service rate of the system remains unchanged, there is an increase in waiting time due. network overload. Which means the system is getting slow. If the arrival rate exceeds 1100, the first queue is stable, but the entire system is unstable.

6. RESIZING

Resizing a system is arbitrary. Indeed, it is possible to resize the network by increasing the number of queues, or by increasing the number of servers. This is to make the queue as well as the system more stable. Tab6 summarizes the stability conditions of the previous experiment.

Tab6 : **Récapitulation de la stabilité de chaque files d'attente.**

EA	ρ_1	ρ_2	ρ_3	STABILITY
100	0,063	0,083	0,057	STABLE
200	0,125	0,167	0,114	STABLE
300	0,188	0,250	0,171	STABLE
400	0,250	0,333	0,229	STABLE
500	0,313	0,417	0,286	STABLE
600	0,375	0,500	0,343	STABLE
700	0,438	0,583	0,400	STABLE
800	0,500	0,667	0,457	STABLE
825	0,516	0,688	0,471	STABLE
850	0,531	0,708	0,486	STABLE
875	0,547	0,729	0,500	STABLE
900	0,563	0,750	0,514	STABLE
1000	0,625	0,833	0,571	STABLE
1100	0,688	0,917	0,629	STABLE
1200	0,750	1,000	0,686	UNSTABLE
1300	0,813	1,083	0,743	UNSTABLE
1400	0,875	1,167	0,800	UNSTABLE
1500	0,938	1,250	0,857	UNSTABLE
1600	1,000	1,333	0,914	UNSTABLE
1700	1,063	1,417	0,971	UNSTABLE
1800	1,125	1,500	1,029	UNSTABLE
1900	1,188	1,583	1,086	UNSTABLE
2000	1,250	1,667	1,143	UNSTABLE

Where EA denotes the external arrival rate to the network, N1 the average number of customers in the first queue. R1 designates the average crossing time of the second row. N designates the number of average clients in the network. R is the average traverse time of the entire network. Lambda refers to the average arrival rate in the network. r01 is the stability condition of the first row. If it is less than one, the network is stable. But if the network is unstable while ro1 is less than one then one of the queues other than the first is unstable.

The ρ_2 rate is the highest among ρi, and the ρ_3 rate is the lowest. This means that the second queue is the most used among the three queues in the system.

When $\rho_2 = 1$, TA = 1200, which means the system has become unstable. So the lower the number of servers in queue # 2, the faster the system tends to instability.

The following experiment takes place in three phases namely cas1 with six servers, cas2 with seven servers and cas3 with eight servers. The results are summarized in Tab8. After several test with a service rate set at 800 but a variable arrival rate (incremented by 100 after each measurement), Tab7 summarizes the input parameters.

Tab7 : **Fixed entries for experimentation**

QUEUE 1	*QUEUE 2*	*QUEUE 3*
r_{11} =0	r_{21} =0.25	r_{31} =0.125
r_{12} =1	r_{22} =0.25	r_{32} =0.125
r_{13} =0	r_{23} =0.5	r_{33} =0.25
m_1 =4	m_2 =6	m_3 =5
μ_1 =800	μ_2 =600	μ_3 =700

All the parameters of the third chapter for the stable system are kept except the external arrival rate which will be variable.
After the ρ utilization rate (i varying from 1 to 3) is equal to 1, the transit time and the number of clients are infinite in the network.
For an experiment on the influence of the number of servers on the system.

Tab8 : **Variation of $\rho 2$ for the three cases**

EA	*case1 : m=6*	*case2 : m=7*	*case3 : m=8*
100	0,083	0,077	0,062
200	0,167	0,143	0,125
300	0,25	0,214	0,187
400	0,333	0,286	0,25
500	0,417	0,357	0,312
600	0,5	0,428	0,374
700	0,583	0,5	0,437
800	0,667	0,571	0,499
900	0,75	0,643	0,562
1000	0,833	0,714	0,625
1100	0,917	0,786	0,687
1200	1	0,857	0,75
1300	1,083	0,928	0,812
1400	1,167	1	0,87
1500	1,25	1,07	0,937
1600	1,333	1,143	0,999
1700	1,417	1,214	1,06
1800	1,5	1,285	1,125
1900	1,583	1,357	1,187
2000	1,667	1,429	1,25

The purpose of the tests is to see the behavior of the second row for a variation in the external arrival rate. But after several additional tests, additional conclusions could be interpreted. Indeed, apart from the instability of the system, it should be noted that some data distorts the calculations.

For the study of a network with three queues, having a bottleneck located on the second queue, and that the parameters of each queue have already been seen in the preceding tables, the proposal of a resizing in the second queue can be summed up by Tab9.

Tab9 : **Sizing of the second queue.**

Number of server	***EA FOR THE THRESHOLD***	***DIFFERENCE***
6	1200	0
7	1400	200
8	1700	500

Which means that the most appropriate resizing is to increase the number of servers in the second queue to eight servers. Indeed, the arrival rate to reach the stability threshold goes from 1200 to 1700.

7. CONCLUSION

This work consists of modeling a network of queues knowing its basic parameters as input. But it is still necessary to know each queue that makes up the network, to model them as a black box and to assess its performance and stability conditions. It was only after that there was the modeling of the entire network.

The modeling generally requires the knowledge of the expected type of entry, but in our case it was assumed that the entry has the Poissonnian arrival process due to the uncertainty of the arrivals which can increase considerably depending on the type of network and the independence of each arrival on different intervals.

This work had provided approximate solutions for large networks. A possible improvement too would be to research an exact-solution solving algorithm for multi-client systems to further improve simulations and final network implementation.

Maintaining the stability of each node in the network is important because a single unstable node can destabilize the network, which would require knowledge of the node having the bottleneck.

8. 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] « Aide-mémoire sur les réseaux de files d'attente markoviennes », Télécom 2A, INPG, 2005.

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

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

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

[6] Hakim Badis, « Modélisation des Réseaux », IGM, Université de Paris-Est Marne-la-Vallée

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

[8] Evaluation et sureté de fonctionnement, « Technique de modélisation: Méthodes analytique », Conservatoire National Des Arts et Métiers.