Chapter 6 – M/M/m/m Loss Systems

Ioannis Glaropoulos

November 14, 2012

Abstract

The importance for this tutorial is to introduce the students in the loss systems, and, to underline the difference between the concepts of call-blocking and time-blocking probabilities, and to understand under which conditions these probabilities are identical. The selected exercises introduce, also, the Erlang tables, which is an important tool for easy calculations for the blocking probabilities.

1 Exercise 6.5

A telephone switch has 10 output lines and a large number of incoming lines. Upon arrival a call on the input line is assigned an output line if such line is available – otherwise the call is blocked and lost. The output line remains assigned to the call for its entire duration which is of exponentially distributed length. Assume that 180 calls / hour arrive in Poisson fashion whereas the mean call duration is 110 seconds.

1. *Determine the blocking probability.*
2. *How many calls are rejected per hour?*
3. *What is the average load per server (in Erlang)?*
4. *What is the maximum arrival rate at which a blocking probability of (at most) 2% can be guaranteed?*

Solution: This is an introductory problem for the Markovian Loss Systems. We start with the system mode and the Kendall Notation:

- Poisson arrivals with rate $\lambda = 180$ calls per hour
- Exponential service times with rate $\mu = \frac{1}{E[T]} = \frac{3600}{110}$
- 10 Servers
- No buffer

The Kendall Notation is: **M/M/10/10**.

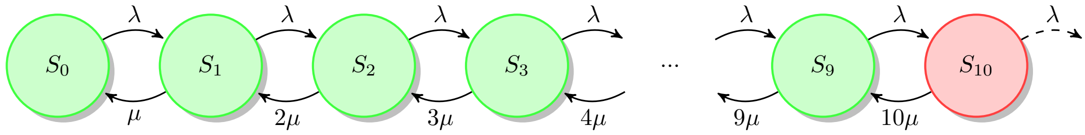


Figure 1: System diagram for exercise 6.4

We draw the system diagram (Fig. 1) and, based on that, we build the balance equations to compute the state probabilities. A state, S, defines the number of active calls in the system. We define

$$\rho = \frac{\lambda}{\mu}$$

which, here, is also the *offered load* of the system; of course this is NOT the actual load, since the system may drop calls when it is blocked. We have:

$$\lambda P_0 = \mu P_1 \rightarrow P_1 = \rho P_0$$

$$\lambda P_1 = 2\mu P_2 \rightarrow P_2 = \tfrac{\rho}{2} P_1 = \tfrac{\rho^2}{2} P_0$$

$$\lambda P_2 = 3\mu P_3 \rightarrow P_3 = \tfrac{\rho}{3} P_2 = \tfrac{\rho^3}{3\cdot 2} P_0$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\lambda P_{k-1} = k\mu P_k \rightarrow P_k = \tfrac{\rho}{k} P_{k-1} = \tfrac{\rho^k}{k!} P_0$$

We use the normalization equation to compute the P_0:

$$\sum_{k=0}^{n} P_k = 1 \rightarrow \sum_{k=0}^{n} \frac{\rho^k}{k!} P_0 = 1 \rightarrow P_0 = \frac{1}{\sum_{k=0}^{n} \frac{\rho^k}{k!}}.$$

where, here, $n = 10$. The blocking probability is the probability that the system can not accept new calls, and this is equal to the probability of the system being in state S_n (S_{10}). This probability is, also, the *TIME BLOCKING*, i.e. the percentage of time the system can not accept new calls, and it is the one seen by an independent observer.

$$P_{block} = P_n = \frac{\frac{\rho^n}{n!}}{\sum_{k=0}^{n} \frac{\rho^k}{k!}}$$

General rule: In Markov chains where the arrival rates do not depend on the system state, the time blocking is also **equal** to the probability of a random event being blocked. The latter is defined as *CALL BLOCKING* probability.

You can compute the P_{block} from the equation above, but you could also do it by using the Erlang tables.

- Servers: $n = 10$
- Offered Load: $\rho = \frac{\lambda}{\mu} = \frac{180}{\frac{3600}{110}} = 5.5$

We search in the tables for the solution. $E_{10}(5.5) \approx 0.029$.

The rate of rejected calls is ALWAYS given by: $\lambda \cdot P_{call\_block}$. However, as stated above, here,

$$P_{call\_block} = P_{block}$$

since the arrival process is independent of the system state[1]. We get:

$$\lambda_{rejected} = \lambda P_{block} \approx 180 \cdot 0.029 = 5.27.$$

To find the average load per server, we first need to find the total ACTUAL load. The actual load is the nominal (offered) load minus the rejected load:

$$\lambda_{eff} = \lambda - \lambda_{rejected} = \lambda(1 - P_10) = 5.5 \cdot (1 - 0.029)$$

For the average load per server, we divide the above with the number of severs.[2]

As we see the blocking probability is above the target of 2%. So we must decrease the offered load, keeping the number of servers to 10. We check in the tables the maximum offered load that guarantees the blocking probability below the target. Then, from that we find the nominal arrival rate.

2 Exercise 6.3

We consider two types of call arrivals to a cell in a mobile telephone network: new calls that originate in a cell and calls that are handed over from neighboring cells. It is desirable to give preference to handover calls over new calls. For this reason, some of the channels in the cell are reserved for handover calls, while the rest of the channels are available to both types of calls. For the questions below, assume the following: Channels in the cell are held for two minutes on average, with exponential distribution. All calls arrive according to a Poisson process with rate $\lambda_{nc} = 125$ calls per hour for new calls, and $\lambda_{ho} = 50$ calls per hour for handover calls. The cell has a capacity of 10 channels; each call occupies 1 channel.

1. *Draw a state diagram of the channel occupancy in a cell when 2 channels are used exclusively for handover calls.*
2. *Calculate the blocking probability in the cell if no channels are reserved for handover calls. What is the average number of channels used?*
3. *Find the minimum number of channels reserved for handover calls so that their blocking probability is below 1 percent. What is the blocking probability for the new calls in this case?*

Solution:

[1] or, as we say, alternatively, the arrival process is homogeneous.

[2] Check also exercise 4.5.

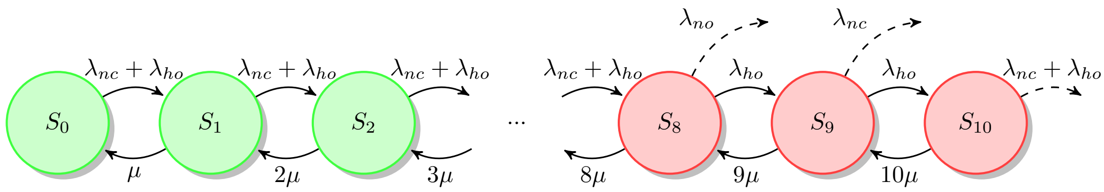


Figure 2: State diagram for exercise 6.3. Case of 2 reserved channels for handover calls. Notice that the system accepts only handover arrivals when only 2 channels remain vacant. In state 10, both handover and new calls are dropped.

This is an interesting exercise that models a Mobile network, like GSM, UTRAN etc. We have a system which prioritizes some calls (handover) over some others (new).

We draw the state transition diagram for the case of 2 reserved channels for handover calls (Fig. 2)

- Poisson arrivals − $\lambda = \lambda_{nc} + \lambda_{ho}$ for states 0,1,2,...,7, $\lambda = \lambda_{ho}$ for states 8,9. This is a nce application of the Poisson SPLIT property.
- Exponential Service times − $\mu = \frac{1}{E[t]} = \frac{1}{2/60} = 30$.
- 10 servers, no buffer

We can draw and solve the balance equations for this system, and calculate, for example, the average number of active calls in the cell. The offered load in the system is

$$\lambda_{total} \cdot E[T] = (\lambda_{nc} + \lambda_{ho}) \cdot E[T].$$

What is the actual load of this system?

We consider, now, the case where no channels are reserved for handover calls. The Kendall notation is then M/M/10/10. So, the system resembles a standard M/M/10/10 case!

- Offered Load: $\rho = (\lambda_{nc} + \lambda_h o)E[T] = 175 \cdot \frac{2}{60} = \frac{35}{6}$.
- Servers: $n = 10$
- Blocking Probability: $E_{10}(\frac{35}{6}) = ...$

Consequently, the effective, or actual load of the system is $\lambda_{eff} = 175 \cdot (1 - E_{10}(35/6)) = ...$

We can apply the state probabilities to derive the average number of calls in the cell. However, since we know the effective load, and the average system time, we can also apply the LITTLE formula

$$\overline{N} = \lambda_{eff} \cdot E[T] = ...$$

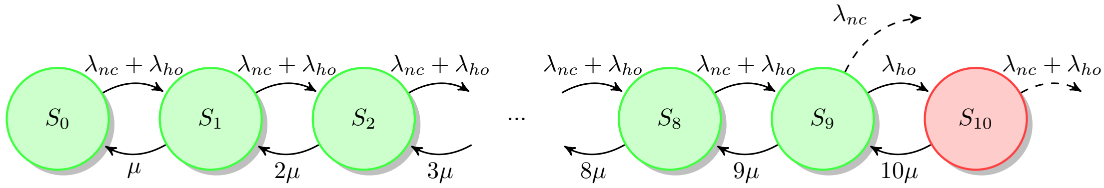


Figure 3: State diagram for exercise 6.3c. Case of 1 reserved channel for handover calls. Notice that the system accepts only handover arrivals when only one channel remains vacant. State S_9 is now green, to show that it is a non-blocked state for handover calls.

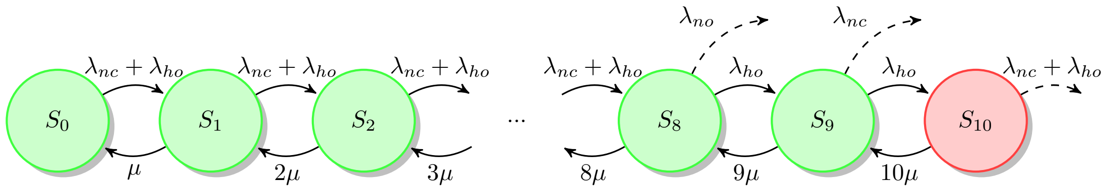


Figure 4: State diagram for exercise 6.3c. Case of 2 reserved channel for handover calls. Notice that the system accepts only handover arrivals when only one channel remains vacant. State S_9 is now green, to show that it is a non-blocked state for handover calls.

For the third question we need to go step by step. First, we do the calculations with one reserved channel. The diagram in shown in Fig. 3 We have the result:

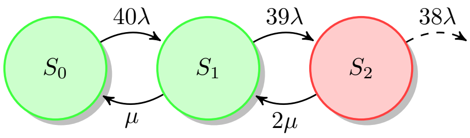
$$P_{block} = P_{10} = \frac{\frac{\lambda_{ho}}{10\mu}\frac{\lambda_{tot}^9}{9!\mu^9}}{1 + \frac{\lambda_{tot}}{\mu} + \frac{\lambda_{tot}^2}{2\mu^2} + ... + \frac{\lambda_{tot}^9}{9!\mu^9} + \frac{\lambda_{ho}\lambda_{tot}^9}{10!\mu^{10}}} = \frac{\frac{\lambda_{ho}}{10\mu}}{\frac{1}{E_9(\rho)} + \frac{\lambda_{ho}}{10\mu}}$$

If we replace the number we find a blocking probability of 1.1%, which is above the target. We now do the calculations for two reserved channels (Fig. 4):

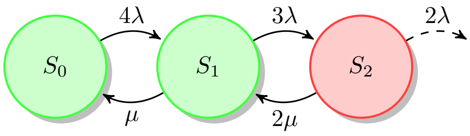
$$P_{block} = P_{10} = \frac{\frac{\lambda_{ho}^2}{10\mu 9\mu}\frac{\lambda_{tot}^8}{8!\mu^8}}{1 + \frac{\lambda_{tot}}{\mu} + \frac{\lambda_{tot}^2}{2\mu^2} + ... + \frac{\lambda_{tot}^8}{8!\mu^8} + \frac{\lambda_{ho}\lambda_{tot}^8}{9!\mu^9} + \frac{\lambda_{ho}^2\lambda_{tot}^8}{10!\mu^{10}}} = \frac{\frac{\lambda_{ho}^2}{10 \cdot 9\mu^2}}{\frac{1}{E_8(\rho)} + \frac{\lambda_{ho}}{9\mu} + \frac{\lambda_{ho}^2}{10 \cdot 9\mu^2}}$$

We do again the calculations and find a probability of 0.2%, which is acceptable.

Finally, the blocking probability for the new calls is $P_8 + P_9 + P_{10}$, since the system is time homogeneous and the last two channels are reserved for the handovers.



(a)



(b)

Figure 5: State diagrams for the A and B systems of exercise 8.

3 Exercise 8 [Collection of Exam Problems]

Consider two M/M/2 loss systems, A and B. A has 40 subscribers; to B there are 4 subscribers connected. Each subscriber generates on average 1 call per minute. The mean service time is 6 seconds.

1. *Compare the call blocking probability for A and B*
2. *Compare the mean blocking time for A and B*
3. *Compare the mean time without blocking for A and B*
4. *How many additional servers do you have to provide to system A to achieve a time blocking that is not higher than that of system B?*

Solution: There is a clear difference between the two system, and that is, the user population. In Fig. 5 we depict the two system diagrams. Both systems have finite population, i.e. the user arrival rate depends on the system state. However, the first system (A) could be approximated with a typical M/M/2/2 system, since the population size (40) is high compared to the number of servers (2). In other words, $40\lambda \approx 39\lambda \approx 38\lambda$. We can not say the same for system B.

System A The offered load is

$$\rho = \frac{40\lambda}{\mu} = 40\lambda \cdot E[T] = 40 \cdot 1 \cdot \frac{6}{60} = 4 \text{ Erl.}$$

We can either draw and solve the balance equations, or use the Erlang Tables to compute the TIME blocking probability (P_2). In the second case, we obtain:

$$E_m(\rho) = E_2(4) = 0.615.$$

We notice that the TIME Blocking probability is equal to the CALL blocking probability since the arrival intensity is constant and does not depend on the state (we assumed typical M/M/2/2). So, we obtain:

$$P_{block} = P_2 = 0.615$$

The mean blocking time (T_b) is the average time the system CONTINUOUSLY remains in state S_2, i.e. from the time it arrives at S_2 until it departs to S_1. The transition rate ($S_2 \rightarrow S_1$) is 2μ, so the average time the system stays at S_2 is $E[T_b] = 1/2\mu = 3$sec.

The **mean time without blocking** (T_n) is the average time between a **departure** from state S_2 and an **arrival** at state S_2. It is **not** the percentage of time without blocking! That would be, simply, $P_0 + P_1$.

We must calculate $E[T_n]$ based on $P_b, E[T_b]$. We define a **cycle** of the system as a set of adjacent blocking and non-blocking periods. Consider that we observe the system for a large period of cycles, say M. Then it holds:

$$P_b = \lim_{M\to\infty} \frac{\sum_{i=1}^{M} T_b^i}{\sum_{i=1}^{M} T_b^i + T_n^i}, \tag{1}$$

where T_b^i is the duration of the blocking period at cycle i and T_n^i is the duration of the non-blocking period at cycle i. We rewrite the above as

$$P_b = \lim_{M\to\infty} \frac{\frac{1}{M}\sum_{i=1}^{M} T_b^i}{\frac{1}{M}\sum_{i=1}^{M} T_b^i + T_n^i} = \frac{\lim_{M\to\infty} \frac{1}{M}\sum_{i=1}^{M} T_b^i}{\lim_{M\to\infty} \frac{1}{M}\sum_{i=1}^{M} T_b^i + T_n^i} \tag{2}$$

and finally:

$$P_b = \frac{\lim_{M\to\infty} \frac{1}{M}\sum_{i=1}^{M} T_b^i}{\lim_{M\to\infty} \frac{1}{M}\sum_{i=1}^{M} T_b^i + \lim_{M\to\infty} \frac{1}{M}\sum_{i=1}^{M} T_n^i} = \frac{E[T_b]}{E[T_b] + E[T_n]}. \tag{3}$$

From the above we can calculate $E[T_n]$ with respect to P_b and $E[T_b]$.