# Solutions for the exercises in Chapter 4 (M/M/c-systems)

1. For an M/M/1-system with parameters $\lambda$ and $\mu$, the stationary probability distribution is given by

$$p_k = \left(\frac{\lambda}{\mu}\right)^k \left(1 - \frac{\lambda}{\mu}\right)$$

In this case we set $\rho = \lambda/\mu$, which means that $p_k = \rho^k(1-\rho)$.

(a) $\rho$=0.9
*Alternative 1*:
The average number of jobs in the system is given by

$$\bar{N} = \sum_{k=1}^{\infty} kp_k = \sum_{k=1}^{\infty} k \cdot 0,9^k \cdot 0,1 = 0,09 \cdot \sum_{k=1}^{\infty} k \cdot 0,9^{k-1}$$
$$= 0,09 \cdot \frac{1}{(1-0,9)^2} = 9$$

*Alternative 2:*

The average number of jobs in the system is given by $\bar{N} = \lim_{z \to 1} \frac{d}{dz}P(z)$ where

$$P(z) = \sum_{k=0}^{\infty} p_k z^k = \sum_{k=0}^{\infty} 0,9^k \cdot 0,1 \cdot z^k = 0,1 \cdot \frac{1}{1-0,9z}.$$

This means that

$$\bar{N} = \lim_{z \to 1} \frac{0,1 \cdot 0,9}{(1-0,9z)^2} = \frac{0,09}{0,01} = 9.$$

(b) $\bar{N} = \bar{N}_q + \bar{N}_s$ where $\bar{N}_q$ is the average number of jobs in the queue and $\bar{N}_s$ is the average number of jobs in the server.
*Alternative 1:*
$\bar{N}_s$ is determined as,

$$\bar{N}_s = 0 \cdot p_0 + 1 \cdot \underbrace{\sum_{k=1}^{\infty} p_k} = 1 - p_0 = \rho = 0,9$$

which means that $\bar{N}_q = \bar{N} - \bar{N}_s = 9 - 0,9 = 8,1$.

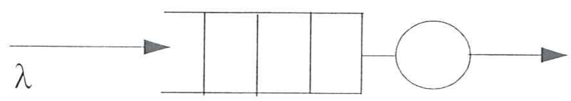
*Alternative 2*: $\bar{N}_q$ can be determined as:

$$\bar{N}_q = 0 \cdot p_0 + 0 \cdot p_1 + \sum_{k=2}^{\infty} (k-1)p_k = \sum_{k=2}^{\infty} kp_k - \sum_{k=2}^{\infty} p_k$$
$$= \left(\sum_{k=1}^{\infty} kp_k\right) - p_1 - \left(\sum_{k=0}^{\infty} p_k\right) + p_0 + p_1 = 9 - 1 + 0,1 = 8,1$$

(c) Since $p_k = \rho^k \cdot p_0$, where $\rho < 1$, will the most probably state be state 0.

(d) When $\rho \to \infty$, the queue will almost always be full, which means that $\bar{N}_q \to L$. Everytime a job departure occurs, a new job will enter the system. This means that $\lambda_{eff} \to \mu$. Using Little's theorem we can determine the average waiting time in queue, $W$: $W = \bar{N}_q/\lambda_{eff} \to L/\mu$.

2. Consider a normal M/M/1-system:




The stationary probability distribution is given by $p_k = \left(\frac{\lambda}{\mu}\right)^k \left(1 - \frac{\lambda}{\mu}\right)$.

The average number of jobs in the system is determined as in exercise 1.

Result: $E[N] = \bar{N} = \frac{\lambda/\mu}{1 - \lambda/\mu}$.

Let $\rho = \lambda/\mu$.

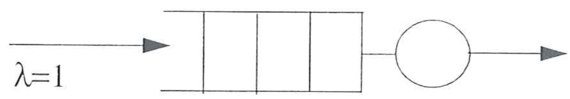
The second moment of $N$ can be determined from the generating function $P(z)$, which was determined in exercise 1:

$$E[N^2] = \lim_{z \to 1} \frac{d^2}{dz^2} P(z) + E[N] = \lim_{z \to 1} \frac{d^2}{dz^2}\left(\frac{1-\rho}{1-\rho z}\right) + E[N]$$

$$= \lim_{z \to 1} \frac{2(1-\rho)\rho^2}{(1-\rho z)^3} + E[N] = \frac{2(1-\rho)\rho^2}{(1-\rho)^3} + \frac{\rho}{1-\rho} = \ldots = \frac{\rho^2 + \rho}{(1-\rho)^2}$$

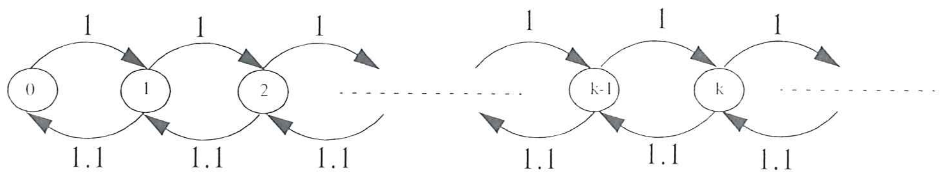
This means that: $V[N] = E[N^2] - E[N]^2 = \ldots = \frac{\rho}{(1-\rho)^2}$.

3. The tent is modelled as an M/M/1-system:




(a) All customers will be served which means that the total number of served customers during one day is $\lambda \cdot 60 \cdot 12 = 720$.

(b) State diagram of the system:




With the "cut-method" we can derive the following balance equations:

$$\begin{cases} p_0 = 1,1 p_1 \\ p_{k-1} = 1,1 p_k \end{cases}$$

Solve the equations. Result: $p_k = \left(\frac{1}{1,1}\right)^k \cdot p_0$

Use the normalisation condition $\sum_{k=0}^{\infty} p_k = 1$ to find $p_0$. Result: $p_0 = 1 - \frac{1}{1,1}$.

Therefore, the stationary probability distribution is $p_k = \left(\frac{1}{1,1}\right)^k \cdot \left(1 - \frac{1}{1,1}\right)$

(c) Let $\overline{N}_q$ =average number of customers in the queue.

*Alternative 1*:

$$\overline{N}_q = 0 \cdot p_0 + 0 \cdot p_1 + \sum_{k=2}^{\infty} (k-1)p_k = \sum_{k=2}^{\infty} (k-1)\left(\frac{1}{1,1}\right)^k \cdot \left(1 - \frac{1}{1,1}\right)$$

$$= \sum_{k=1}^{\infty} k\left(\frac{1}{1,1}\right)^{k+1} \cdot \left(1 - \frac{1}{1,1}\right) = \left(\frac{1}{1,1}\right)^2 \cdot \left(1 - \frac{1}{1,1}\right)\sum_{k=1}^{\infty} k\left(\frac{1}{1,1}\right)^{k-1} = \frac{\left(\frac{1}{1,1}\right)^2}{\left(1 - \frac{1}{1,1}\right)}$$

which means that $\overline{N}_q \approx 9,1$

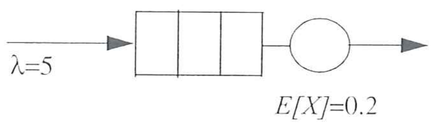
*Alternative 2*: Låt $\overline{N}$ =average number of customers in the whole system och $\overline{N}_s$= average number of .customers in the tent.

For an M/M/1-system: $\overline{N} = \left(\frac{1}{1,1}\right) / \left(1 - \frac{1}{1,1}\right) = 10$

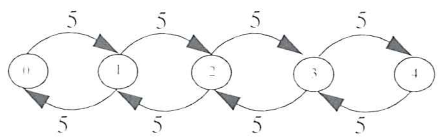
and $\overline{N}_s = 1 - p_0 = \frac{1}{1,1} \approx 0,91$ , which means that $\overline{N}_q = \overline{N} - \overline{N}_s \approx 9,1$

(d) P(a customer can enter the tent directly)=P(the tent is empty)= $p_0 = \left(1 - \frac{1}{1,1}\right) \approx 0,091$

4. The computer system is modelled as an M/M/1-system with a finite queue:




(a) The system has five states. The state diagram is:




By using the "cut-method" the following balance equations can be derived:

$$\begin{cases} 5p_0 = 5p_1 \\ 5p_1 = 5p_2 \\ 5p_2 = 5p_3 \\ 5p_3 = 5p_4 \end{cases} \Rightarrow p_0 = p_1 = p_2 = p_3 = p_4$$

Also, $\sum_{k=0}^{5} p_k = 1 \Rightarrow p_0 = p_1 = p_2 = p_3 = p_4 = \frac{1}{5}$

(b) The average number of jobs in the system is given by:

$$\bar{N} = 0 \cdot p_0 + 1 \cdot p_1 + 2 \cdot p_2 + 3 \cdot p_3 + 4 \cdot p_4 = \frac{10}{5} = 2$$

(c) The average time a job spends in the system, $T$, is determined with Little's theorem: $T = \frac{\bar{N}}{\lambda_{eff}}$.

The effective arrival rate is:

$$\lambda_{eff} = \lambda \cdot p_0 + \lambda \cdot p_1 + \lambda \cdot p_2 + \lambda \cdot p_3 + 0 \cdot p_4 = \lambda(1 - p_4) = 4$$

which means that $T = \frac{2}{4} = 0,5$ seconds.

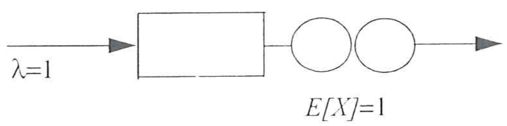
(d) *Alternative 1*: Use the definition of carried load:

$$\rho_c = \lambda_{eff} \cdot E[X] = 4 \cdot \frac{1}{5} = \frac{4}{5}$$

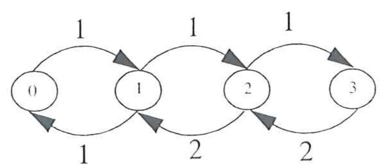
*Alternative 2*: The carried load, $\rho_c$, is the same as the average number of jobs in the server:

$$\rho_c = \bar{N}_s = 0 \cdot p_0 + 1 \cdot (p_1 + p_2 + p_3 + p_4) = \frac{4}{5}$$

5. The web site is modelled as an M/M/2-system with 1 queuing place:




(a) There are four states, where state $k$ = $k$ requests in the system. State diagram:




(b) First, we must know the stationary probability distribution
Let $p_k$=$P$($k$ requests in the system). The "cut-method" gives the following balance equations:

$$\begin{cases} p_1 = p_0 \\ 2p_2 = p_1 \Rightarrow p_2 = \frac{1}{2}p_0 \\ 2p_3 = p_2 \Rightarrow p_3 = \frac{1}{4}p_0 \end{cases}$$

Use the normalisation condition $\sum p_k = 1$ to determine $p_0$.

Result: $p_0 = \frac{4}{11}$, $p_1 = \frac{4}{11}$, $p_2 = \frac{2}{11}$ och $p_3 = \frac{1}{11}$

P(the system is full) = $p_3 \approx 0,09$.

P(a request is blocked)=$P_b = \frac{\lambda_b}{\lambda} = \frac{1p_3}{1p_0 + 1p_1 + 1p_2 + 1p_3} = p_3$

Note: Only when the arrival process is Poissonian will these two probabilities be equal!

(c) Offered load: $\rho_o = \lambda \cdot E[X] = (1p_0 + 1p_1 + 1p_2 + 1p_3) \cdot 1 = 1$ Erlangs

Carried load: $\rho_c = \lambda_{eff} \cdot E[X] = (1p_0 + 1p_1 + 1p_2 + 0p_3) \cdot 1 = \frac{10}{11} \approx 0,91$

(d) Utilisation of a server = fraction of time that the server is busy = P(server busy).
*Alternative 1*: Utilisation of a server = Carried load for this server.
Since the servers are chosen randomly, they will be equally loaded, which means that the utilisation of *one* server is half the total carried load= $\frac{1}{2} \cdot \frac{10}{11} \approx 0,45$ .

*Alternative 2:* Utilisation of one server = average number of requests in that server =

$0 \cdot p_0 + 1 \cdot p_1 \cdot \frac{1}{2} + 1 \cdot p_2 + 1 \cdot p_3 = \frac{5}{11} \approx 0,45$ (in state 1 a specific server is busy with probability 0.5)

(e) The average waiting time for a request, $W$, is determined with Little's theorem: $W = \frac{\overline{N}_q}{\lambda_{eff}}$.

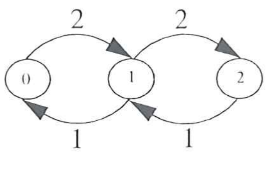
In this case:

$\lambda_{eff} = 1p_0 + 1p_1 + 1p_2 + 0p_3 = \frac{10}{11}$ and $\overline{N}_q = 0 \cdot (p_0 + p_1 + p_2) + 1 \cdot p_3 = \frac{1}{11}$

which means that $W = \frac{1/11}{10/11} = \frac{1}{10}$ seconds.

6. M/M/1-system with one queueing place means maximum 2 jobs at the same time.

(a) State $k$ = $k$ jobs in the system. State diagram:




The "cut"-method gives the following balance equations:

$$\begin{cases} 2p_0 = p_1 \\ 2p_1 = p_2 \Rightarrow p_2 = 4p_0 \end{cases}$$

Use the normalisation condition $\sum_{k=0}^{2} p_k = 1$ to determine $p_0$.

Result: $p_0 = \frac{1}{7}$, $p_1 = \frac{2}{7}$, $p_2 = \frac{4}{7}$

(b) Let $N$ = number of customers at a random time instant ($N$ is a random variable)
This means that the average of $N$ = $E[N]$ is given by

$$E[N] = \sum_{k=0}^{2} k \cdot p_k = 1 \cdot p_1 + 2 \cdot p_2 = \frac{10}{7} \approx 1,43$$

The variance is $V[N] = E[N^2] - E[N]^2$

where $E[N^2] = \sum_{k=0}^{2} k^2 \cdot p_k = 1 \cdot p_1 + 4 \cdot p_2 = \frac{18}{7}$

which means that $V[N] = \frac{18}{7} - \left(\frac{10}{7}\right)^2 = \frac{26}{49} \approx 0,53$

(c) $T$=Time in the system for a job <u>that is served</u> (random variable)
Let the density function for $T$ be $f_T(t)$.

The distribution function is given by $P(T \le t) = \int_0^t f_t(v)dv$

First, determine the Laplace transform for $f_T(t)$, $F^*_T(s)$.
An arriving job that is served either finds 0 or 1 jobs in the system (if there are 2 jobs in the system, the arriving job is blocked)
If the job arrives at an empty system, the total time in the system is the job's service time, which means that $F^*_T(s|k=0) = \frac{\mu}{\mu + s} = \frac{1}{1+s}$, i.e. the Laplace transform for the service time.
If an arriving job finds 1 job in the system, this job's total time in the system consists of the other job's remaining service time plus the arriving job's own service time. Due to the memoryless property of the exponential distribution, will the remaining service time for the job in the server be exponentially distributed with mean $1/\mu$.

This means that $F^*_T(s|k=1) = \left(\frac{\mu}{\mu+s}\right)^2 = \left(\frac{1}{1+s}\right)^2$

Also, P(an arriving job finds 0 jobs (of those jobs that find 0 or 1 jobs)) = $\frac{\lambda p_0}{\lambda p_0 + \lambda p_1} = \frac{1}{3}$

and P(an arriving job finds 1 job) = $\frac{\lambda p_1}{\lambda p_0 + \lambda p_1} = \frac{2}{3}$.

The theorem of total probability then gives: $F^*_T(s) = \frac{1}{3} \cdot \frac{1}{1+s} + \frac{2}{3} \cdot \left(\frac{1}{1+s}\right)^2$.

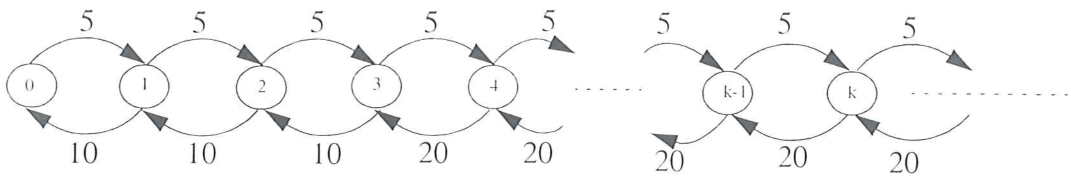
Inverse Laplace transform of $F^*_T(s)$ gives $L^{-1}(F^*_T(s)) = f_T(t) = \frac{1}{3} \cdot e^{-t} + \frac{2}{3} \cdot te^{-t}$,

which means that

$$P(T \le t) = \int_0^t f_t(v)dv = \int_0^t \left(\frac{1}{3} \cdot e^{-v} + \frac{2}{3} \cdot ve^{-v}\right)dv = \ldots = 1 - e^{-t} - \frac{2}{3} \cdot e^{-t}$$

7. This is a system where the number of servers depend on the state.

(a) State diagram:




Balance equations:

$$\begin{cases} 5p_0 = 10p_1 \Rightarrow p_1 = \frac{1}{2}p_0 \\ 5p_1 = 10p_2 \Rightarrow p_2 = \frac{1}{4}p_0 \\ 5p_2 = 10p_3 \Rightarrow p_3 = \frac{1}{8}p_0 \\ 5p_3 = 20p_4 \Rightarrow p_4 = \left(\frac{1}{8}\right) \cdot \frac{1}{4}p_0 \\ 5p_{k-1} = 20p_k \Rightarrow p_k = \left(\frac{1}{8}\right) \cdot \left(\frac{1}{4}\right)^{k-3} p_0, \quad k \geq 3 \end{cases}$$

$p_0$ is as always found with the normalisation condition:

$$\sum_{k=0}^{\infty} p_k = 1 \Rightarrow p_0 + \frac{1}{2}p_0 + \frac{1}{4}p_0 + \sum_{k=3}^{\infty} \left(\frac{1}{8}\right) \cdot \left(\frac{1}{4}\right)^{k-3} p_0 = 1 \Rightarrow p_0 = \ldots = \frac{12}{23}$$

which means that

$$p_1 = \frac{6}{23}$$

$$p_2 = \frac{3}{23}$$

$$p_k = \left(\frac{1}{4}\right)^{k-3} \cdot \frac{3}{46} \qquad k \geq 3$$

(b) *Alternative 1:*

Calculate the average number of jobs in the servers:

$$\begin{aligned} \bar{N}_s &= 0 \cdot p_0 + 1 \cdot (p_1 + p_2 + p_3) + 2 \cdot \sum_{k=4}^{\infty} p_k \\ &= (p_1 + p_2 + p_3) + 2 \cdot (1 - (p_0 + p_1 + p_2 + p_3)) = \ldots = \frac{1}{2} \end{aligned}$$

*Alternative 2:*

The average number of jobs in the servers is always equal to the carried load:

$$\bar{N}_s = \rho_c = \lambda_{eff} \cdot E[X] = \lambda \cdot \frac{1}{\mu} = \frac{1}{2}$$

$\lambda_{eff} = \lambda$ since the queue is unlimited.

(c) Little's theorem determines that the average time in the system for a job is $T = \frac{\bar{N}}{\lambda_{eff}}$.

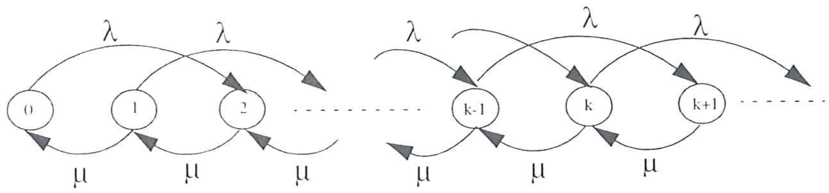
The average number of jobs in the system is calculated as:

$$\bar{N} = \sum_{k=1}^{\infty} kp_k = p_1 + 2p_2 + \sum_{k=3}^{\infty} k\left(\frac{1}{4}\right)^{k-3} \cdot \frac{3}{46} = \frac{12}{23} + \frac{24}{23} \cdot \sum_{k=3}^{\infty} k\left(\frac{1}{4}\right)^{k-1}$$
$$= \frac{12}{23} + \frac{24}{23} \cdot \left[\sum_{k=1}^{\infty} k\left(\frac{1}{4}\right)^{k-1} - 1 - \frac{2}{4}\right] = \ldots = \frac{12}{23} + \frac{24}{23} \cdot \frac{5}{18} \approx 0,81$$

which means that $T = \frac{\bar{N}}{\lambda_{eff}} = \frac{\bar{N}}{\lambda} \approx 0,16$

8. M/M/1-system with so-called batch arrivals.

(a) State diagram:




Balance equations: $\begin{cases} \lambda p_0 = \mu p_1 \\ \lambda p_0 + \lambda p_1 = \mu p_2 \\ \lambda p_{k-2} + \lambda p_{k-1} = \mu p_k & k \geq 2 \end{cases}$

(b) Multiplicate the balance equations above with $z^k$:

$$\begin{cases} \lambda p_0 z = \mu p_1 z \\ \lambda p_0 z^2 + \lambda p_1 z^2 = \mu p_2 z^2 \\ \lambda p_{k-2} z^k + \lambda p_{k-1} z^k = \mu p_k z^k & k \geq 2 \end{cases}$$

Summate all equations for $k \geq 2$:

$$\sum_{k=2}^{\infty} \lambda p_{k-2} z^k + \sum_{k=2}^{\infty} \lambda p_{k-1} z^k = \sum_{k=2}^{\infty} \mu p_k z^k$$

Identify P(z)!
Start with the right side:

$$\sum_{k=2}^{\infty} \mu p_k z^k = \mu \sum_{k=0}^{\infty} p_k z^k - \mu p_0 - \mu p_1 = \mu P(z) - \mu p_0 - \mu p_1$$

Then the left side:

$$\sum_{k=2}^{\infty} \lambda p_{k-2} z^k + \sum_{k=2}^{\infty} \lambda p_{k-1} z^k = \lambda z^2 \sum_{k=2}^{\infty} p_{k-2} z^{k-2} + \lambda z \sum_{k=2}^{\infty} p_{k-1} z^{k-1}$$
$$= \lambda z^2 P(z) + \lambda z (P(z) - p_0)$$

Derive an expression for $P(z)$ from the equation above, use that $\lambda p_0 z = \mu p_1 z$ :

$$\mu P(z) - \mu p_0 - \mu p_1 = \lambda z^2 P(z) + \lambda z (P(z) - p_0) \Rightarrow \ldots \Rightarrow P(z) = \frac{\mu p_0}{\mu - \lambda z - \lambda z^2}$$

$p_0$ is found by using the normalisation condition (in Laplace transform): $P(1)=1$:

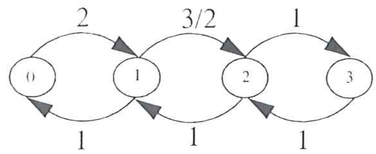
$$P(1) = \frac{\mu p_0}{\mu - \lambda - \lambda} = 1 \Rightarrow p_0 = 1 - \frac{2\lambda}{\mu}$$

(c) The average number of customers in the system. $\bar{N} = \lim_{z \to 1} \frac{d}{dz} P(z)$.

$$\frac{d}{dz} P(z) = \ldots = \frac{\mu p_0 (\lambda + 2\lambda z)}{(\mu - \lambda z - \lambda z^2)^2} \Rightarrow \bar{N} = \frac{3\lambda}{\mu - 2\lambda}$$

9. Use the M/M/1-model with 2 queueing places and 4 clients.

(a) $E[X] = 1 \Rightarrow \mu = 1, \beta = 1/2$.
State diagram:




(b) First, determine the stationary probability distribution!
The "cut"-method gives the following balance equations:

$$\begin{cases} 2p_0 = p_1 \\ \frac{3}{2} p_1 = p_2 \\ p_2 = p_3 \end{cases} \Rightarrow \begin{cases} p_1 = 2p_0 \\ p_2 = 3p_0 \\ p_3 = 3p_0 \end{cases}$$

Use the normalisation condition to find $p_0$.

Result: $p_0 = \frac{1}{9} \cdot p_1 = \frac{2}{9} \cdot p_2 = \frac{3}{9} \cdot p_3 = \frac{3}{9}$.

Use the definition of offered load:

$$\rho_o = \lambda \cdot E[X] \Rightarrow \rho_0 = \left(2p_0 + \frac{3}{2} p_1 + p_2 + \frac{1}{2} p_3\right) \cdot 1 = \frac{19}{18} \text{ Erlangs.}$$

In the same way, use the definition of carried load:

$$\rho_c = \lambda_{eff} \cdot E[X] \Rightarrow \rho_0 = \left(2p_0 + \frac{3}{2} p_1 + p_2\right) \cdot 1 = \frac{8}{9} \text{ Erlangs.}$$

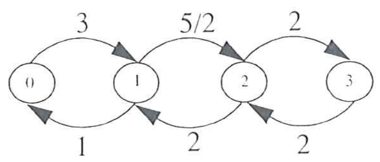
(c) P(system is full) = $p_3 = \frac{1}{3}$

$$\text{P(a job is blocked)}=\frac{\lambda_b}{\lambda} = \frac{\frac{1}{2}p_3}{2p_0 + \frac{3}{2}p_1 + p_2 + \frac{1}{2}p_3} = \frac{3}{19} \approx 0,16\,.$$

10. We use the M/M/2-model with 1 queueing place and 6 clients.

(a) Let state $k$ = $k$ jobs in the system.

If there are $k$ jobs in the system, the total arrival rate is $(6-k)\cdot\frac{1}{2}$ jobs per second.




(b) The "cut"-method gives the following balance equations:

$$\begin{cases} p_1 = 3p_0 \\ 2p_2 = \frac{5}{2}\cdot p_1 \Rightarrow p_2 = \frac{15}{4}\cdot p_0 \\ 2p_3 = 2p_2 \Rightarrow p_3 = \frac{15}{4}\cdot p_0 \end{cases}$$

Use the normalisation condition to determine $p_0$.

Result: $p_0 = \frac{2}{23}$, $p_1 = \frac{6}{23}$, $p_2 = \frac{15}{46}$ och $p_3 = \frac{15}{46}$.

(c) Average number of blocked jobs per second: $\lambda_b = \frac{3}{2}\cdot p_3 = \frac{45}{92} \approx 0,49$

(d) P(the system is full) = $p_3 \approx 0,33$.

$$\text{P(a job is blocked)}=P_b = \frac{\lambda_b}{\lambda} = \frac{\frac{3}{2}\cdot p_3}{3p_0 + \frac{5}{2}\cdot p_1 + 2p_2 + \frac{3}{2}\cdot p_3} = \frac{45}{189} \approx 0,24\,.$$

(e) *Alternative 1*:

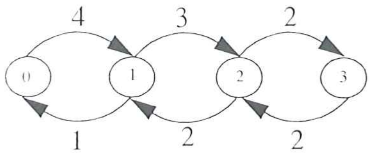
Use the definition of carried load: $\rho_c = \lambda_{eff}\cdot\bar{x} = \left(3p_0 + \frac{5}{2}\cdot p_1 + 2p_2\right)\cdot 1 = \frac{72}{46} \approx 1,57$.

*Alternative 2:*

Carried load = Average number of jobs in the servers: $\rho_c = \bar{N}_s = 1p_1 + 2p_2 + 2p_3 = \frac{72}{46}$.

11. We use the M/M/2-model with 1 queueing place and 4 clients.

(a) State diagram:




The "cut"-method gives the following balance equations: $\begin{cases} 4p_0 = p_1 \\ 3p_1 = 2p_2 \Rightarrow p_2 = 6p_0 \\ 2p_2 = 2p_3 \Rightarrow p_3 = 6p_0 \end{cases}$

$p_0$ is determined with the normalisation condition: $\sum_{k=0}^{3} p_k = 1 \Rightarrow p_0 = \frac{1}{17}$,

which means that $p_1 = \frac{4}{17}$, $p_2 = \frac{6}{17}$, and $p_3 = \frac{6}{17}$.

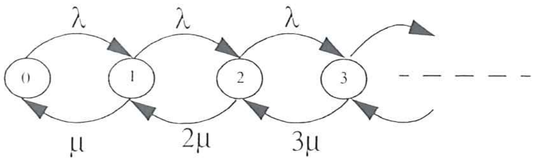
(b) The average number of jobs in the queue: $\bar{N}_q = 0 \cdot (p_0 + p_1 + p_2) + 1 \cdot p_3 = \frac{6}{17}$

(c) P(the system is full)=$p_3 \approx 0,35$

P(a job is blocked)=$\frac{\lambda_b}{\lambda} = \frac{1 \cdot p_3}{4p_0 + 3p_1 + 2p_2 + p_3} = \frac{3}{17} \approx 0,18$

(d) Throughput = Effective arrival rate = $\lambda_{eff} = 4p_0 + 3p_1 + 2p_2 = \frac{28}{17} \approx 1,65$

12. State diagram:




(a) Since no jobs are blocked, the carried load = offered load $= \rho_o = \lambda \cdot E[X] = \lambda/\mu$

(b) There is no queue, which means that the average waiting time is zero.

(c) The "cut"-method gives the following balance equations:

$$\lambda p_0 = \mu p_1$$
$$\lambda p_1 = 2\mu p_2$$
$$\vdots$$
$$\lambda p_k = (k+1)\mu p_{k+1}$$

that can be expressed as: $p_k = \left(\frac{\lambda}{\mu}\right)^k \cdot \frac{1}{k!} \cdot p_0$ for $k \geq 0$.

Let $\lambda/\mu = \rho$.

$p_0$ is determined with the normalisation condition:

$$\sum_{k=0}^{\infty} p_k = 1 \Rightarrow \sum_{k=0}^{\infty} \rho^k \cdot \frac{1}{k!} \cdot p_0 = 1 \Rightarrow p_0 = e^{-\rho}.$$

which means that $p_k = \frac{\rho^k}{k!} e^{-\rho}$.

(d) *Alternative 1:*
The average number of jobs in the system:

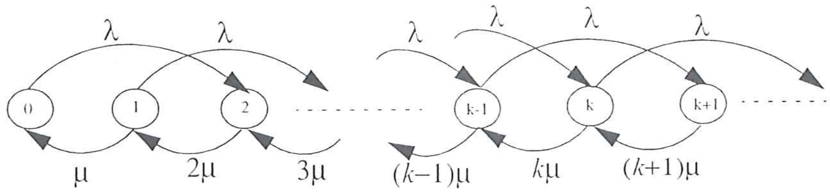
$$\bar{N} = \sum_{k=1}^{\infty} k p_k = \sum_{k=1}^{\infty} k \frac{\rho^k}{k!} e^{-\rho} = \rho e^{-\rho} \sum_{k=1}^{\infty} \frac{\rho^{k-1}}{(k-1)!} = \rho e^{-\rho} \cdot e^{\rho} = \rho$$

*Alternative 2:*

No queue means that $\bar{N} = \bar{N}_s = \rho_c = \rho_o = \rho$. (The average number of jobs in the system is equal to the number of jobs in the servers, which is the same as the carried load that in this case is equal to the offered load since there is no blocking).

13. $M/M/\infty$-system with two jobs in each arrival.

(a) State diagram:




Global balance equations: $\begin{cases} \lambda p_0 = \mu p_1 \\ \lambda p_{k-1} + \lambda p_{k-2} = k\mu p_k & k \geq 2 \end{cases}$

(b) Multiplicate the balance equations with $z^k$ and then summate them:

$$\lambda \sum_{k=2}^{\infty} p_{k-1} z^k + \lambda \sum_{k=2}^{\infty} p_{k-2} z^k = \sum_{k=2}^{\infty} k\mu p_k z^k$$

(c) Solve each part individually.
Right part:

$$\sum_{k=2}^{\infty} k\mu p_k z^k = \mu z \frac{\partial}{\partial z} \sum_{k=2}^{\infty} p_k z^k = \mu z \frac{\partial}{\partial z}(P(z) - p_0 - p_1 z) = \mu z(P'(z) - p_1)$$

First sum in the left part:

$$\lambda \sum_{k=2}^{\infty} p_{k-1} z^k = \lambda z \sum_{k=1}^{\infty} p_k z^k = \lambda z(P(z) - p_0)$$

Second sum in the left part:

$$\lambda \sum_{k=2}^{\infty} p_{k-2} z^k = \lambda z^2 \sum_{k=0}^{\infty} p_k z^k = \lambda z^2 P(z)$$

By using $p_1 = \frac{\lambda}{\mu}p_0$ the following differential equation can be derived: :

$$P'(z) = P(z)(1+z)\frac{\lambda}{\mu}$$

(d) Just insert the proposed solution in the differential equation above....

(e) Use $P(1)=1$: $P(1) = Ce^{\frac{3\lambda}{2\mu}} = 1 \Rightarrow C = e^{-\frac{3\lambda}{2\mu}}$

(f) The average number of jobs in the system: $\overline{N} = \lim_{z \to 1} \frac{d}{dz}P(z) = \frac{2\lambda}{\mu}$

14. M/M/1-system with $\lambda=1s^{-1}$ and $\mu=2s^{-1}$.

(a) Let $N(t)$=number of arrived jobs in a time interval $[T, T+t]$.

$P(N(t) = k) = \frac{(\lambda t)^k}{k!} \cdot e^{-\lambda t}$ since $N(t)$ is Poissonian distributed (The arrival process is Poissonian).

This means that P(at least one job has arrived in interval [20,22])=

$P(N(2) \geq 1) = 1 - P(N(2) = 0) = 1 - e^{-2} \approx 0,865$

(b) Since $N(t)$ is Poissonian distributed, the time between two arrivals is exponentially distributed with mean $1/\lambda$. The memoryless property of the exponential distribution gives that when looking at the system at time $t_0$, the time to the next arrival is always exponentially distributed with mean $1/\lambda$, independent of $t_0$. Here, $t_0$=the time instant when the system becomes empty.

(c) Let $X_i$=service time for job $i$, $i$=1,2 ($X_i$ is a random variable).

M/M/1-system means that $X_i \in \exp(1/\mu) \Rightarrow F_{X_i}(s) = \frac{\mu}{\mu+s}$.

Let $Z$ be the total remaining time in the system for the two jobs. ($Z$ is a random variable).

In this case, $Z = X_1 + X_2$ since the remaining time for the job in the server at time 10 has the same distribution as this job's total service time (due to the memoryless property of the exponential distribution).

This means that $F_Z(s) = \left(\frac{\mu}{\mu+s}\right)^2 \Rightarrow f_Z(t) = \mu^2 te^{-\mu t}$

P(the two jobs have left the system at time $12 |$ 2 jobs in the system at time 10)=

$P(Z \leq 2) = \int_0^2 f_Z(t)dt = \ldots = 1 - (1+2\mu) \cdot e^{-2\mu} = 0,908$

15. The blocking probability in an M/M/c-system without queue can be found in the Erlang tables.

$\rho_o = 5$ Erlangs : $E_c(5) < 0,01 \Rightarrow c \geq 11$.

$\rho_o = 10$ Erlangs: $E_c(10) < 0,01 \Rightarrow c \geq 18$.

$\rho_o = 20$ Erlangs: $E_c(20) < 0,01 \Rightarrow c \geq 30$

Let $U_s$ be the utilisation per server. All servers are chosen randomly, which means that they are evenly utilised. Also, the utilisation per server is the same as the carried load per server. In this system, $\rho_c = \rho_o(1 - E_c(\rho_o))$, which means that $U_s = \frac{\rho_c}{c} = \frac{\rho_o(1-E_c(\rho_o))}{c}$.

$\rho_o = 5$ Erlangs: $U_s = \dfrac{5(1-E_{11}(5))}{11} = 0,45$.

$\rho_o = 10$ Erlangs: $U_s = 0,55$.

$\rho_o = 20$ Erlangs: $U_s = 0,66$

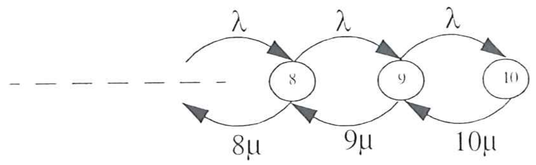
20% increase in offered load means that $\rho_{new} = 1,2 \cdot \rho_{old}$.

$\rho_{old} = 5$ Erlangs: $P_b = E_{11}(6) = 0,023$, $U_s = 0,53$ (for the new system).

$\rho_{old} = 10$ Erlangs: $P_b = E_{18}(12) = 0,027$, $U_s = 0,65$

$\rho_{old} = 20$ Erlangs: $P_b = E_{30}(24) = 0,040$, $U_s = 0,77$

16. M/M/10-system without queue.
The offered load is 5 Erlangs, i.e. $\rho_o = \lambda \cdot E[X] = \lambda/\mu = 5$.
Consider the end of the state diagram:

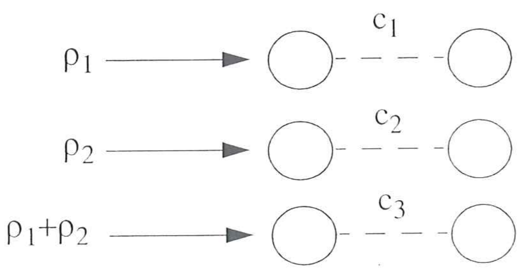



We know from Erlang's loss formula that $p_{10} = E_{10}(\rho_o) = E_{10}(5) \approx 0,018385$
The "cut"-method gives the following balance equations:

$$\lambda p_9 = 10\mu p_{10} \Rightarrow p_9 = 10 \cdot \frac{\mu}{\lambda} p_{10} = \frac{10}{\rho_o} p_{10} = 2 \cdot E_{10}(5) \approx 0,03677$$

$$\lambda p_8 = 9\mu p_9 \Rightarrow p_8 = \frac{9}{\rho_o} p_9 \approx 0,066186$$

17. We consider three loss systems:




where $\rho_1 = 3$ Erlangs and $\rho_2 = 6$ Erlangs.

(a) The number of servers should be chosen so that P(call blocking) is less than 0.01%.
The call blocking probability is found with Erlang's loss formula.

System 1: $E_{c_1}(\rho_1) < 0,001 \Rightarrow c_1 \geq 10$ $\quad E_{10}(3) = 0,00081$

System 2: $E_{c_2}(\rho_2) < 0,001 \Rightarrow c_2 \geq 15$ $\quad E_{15}(6) = 0,00089$

System 3: $E_{c_3}(\rho_3) < 0,001 \Rightarrow c_3 \geq 20$ $\quad E_{20}(9) = 0,00062$

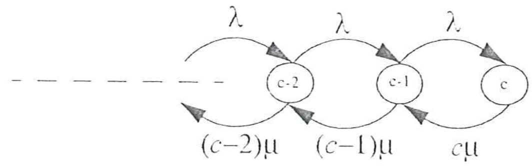
(b) Average utilisation per server = Carried load per server = $\dfrac{\rho_i \cdot (1 - E_{c_i}(\rho_i))}{c_i}$ for system $i$.

System 1: $\dfrac{3 \cdot (1 - E_{10}(3))}{10} = 0,03$

System 2: $\dfrac{6 \cdot (1 - E_{15}(6))}{15} = 0,4$

System 3: $\dfrac{9 \cdot (1 - E_{20}(9))}{20} = 0,45$

(c) P(no more than two servers are idle)=$p_c+p_{c-1}+p_{c-2}$




From the state diagram the following expression can be derived:

$$p_c = E_c(\rho_o)$$

$$p_{c-1} = \frac{c\mu}{\lambda} p_c$$

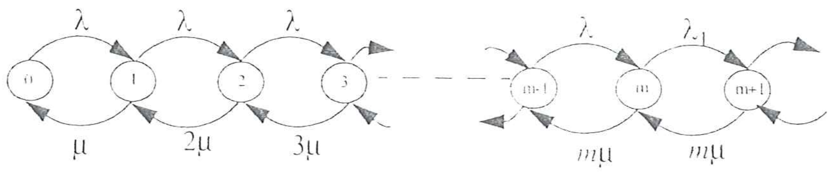
$$p_{c-2} = \frac{(c-1)\mu}{\lambda} p_{c-1}$$

$$\rho_o = \lambda/\mu$$

which means that we can calculate P(no more than two servers are idle) for the three systems above. Result:
System 1: P(no more than two servers are idle)=0.01161
System 2: P(no more than two servers are idle)=0.008325
System 3: P(no more than two servers are idle)=0.004883
This means that system 3, which has most servers, can be more efficiently used. This phenomenon is called "The bigger the better" and applies to all queuing systems. For example, a small radio cell in a mobile network is less efficient than a larger cell with more radio channels (and more traffic).

18. Let $\lambda = \lambda_1 + \lambda_2 \qquad \rho = \dfrac{\lambda}{\mu} \qquad \rho_1 = \dfrac{\lambda_1}{\mu} \qquad \rho_2 = \dfrac{\lambda_2}{\mu}$

(a) State diagram:




The "cut method" gives the following balance equations:

$$\lambda p_{k-1} = k\mu p_k \qquad k \le m$$

$$\lambda_1 p_{k-1} = m\mu p_k \qquad k > m$$

which means that $p_k$ can be derived as:

$$p_k = \left(\frac{\lambda}{\mu}\right)^k \cdot \frac{1}{k!} p_0 \qquad k \le m$$

$$p_k = \left(\frac{\lambda_1}{m\mu}\right)^{k-m} \cdot p_m = \left(\frac{\lambda_1}{m\mu}\right)^{k-m} \cdot \left(\frac{\lambda}{\mu}\right)^m \cdot \frac{1}{m!} p_0 \qquad k \ge m$$

$p_0$ is found by using the normalisation condition:

$$\sum_{k=0}^{\infty} p_k = 1 \Rightarrow p_0 = \frac{1}{\sum_{k=0}^{m} \left(\frac{\lambda}{\mu}\right)^k \cdot \frac{1}{k!} + \sum_{k=m+1}^{\infty} \left(\frac{\lambda_1}{m\mu}\right)^{k-m} \cdot \left(\frac{\lambda}{\mu}\right)^k \cdot \frac{1}{k!}}$$

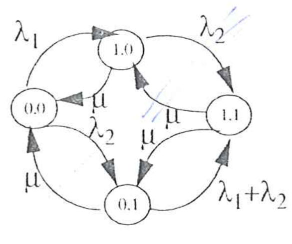
(b) The blocked load, $\rho_b$, is calculated as

$$\rho_b = \lambda_b \cdot E[X] = \left(\sum_{k=m}^{\infty} \lambda_2 p_k\right) \cdot \frac{1}{\mu} = \frac{\lambda_2}{\mu} \sum_{k=m}^{\infty} p_k = \rho_2 \sum_{k=m}^{\infty} p_k$$

$$= \rho_2 \sum_{k=m}^{\infty} \left(\frac{\lambda_1}{m\mu}\right)^{k-m} \cdot \left(\frac{\lambda}{\mu}\right)^m \cdot \frac{1}{m!} p_0 = \rho_2 p_0 \cdot \frac{\rho^m}{m!} \sum_{k=m}^{\infty} \left(\frac{\rho_1}{m}\right)^{k-m} = \rho_2 p_0 \frac{\rho^m}{m!} \left(\frac{m}{m - \rho_1}\right)$$

19. This is a special type of loss system, and we can therefore not use the theory of Erlang systems.

(a) The state $(i,j)$ corresponds to $i$ jobs in server $A$ and $j$ jobs in server $B$.
State diagram:




(b) Let: $p(0,0) = p_0 \qquad p(0,1) = p_1 \qquad p(1,0) = p_2 \qquad p(1,1) = p_3$

The "flow-in-flow-out method" gives the following balance equations:

$$(\lambda_1 + \lambda_2)p_0 = \mu p_1 + \mu p_2$$
$$(\lambda_1 + \lambda_2 + \mu)p_1 = \lambda_2 p_0 + \mu p_3$$
$$(\lambda_2 + \mu)p_2 = \lambda_1 p_0 + \mu p_3$$
$$2\mu p_3 = (\lambda_1 + \lambda_2)p_1 + \lambda_2 p_2$$

Let: $\lambda_1 = 1 \qquad \lambda_2 = 2 \qquad \mu = 1$

Solve the equations and use the normalisation condition: $\sum_{k=0}^{3} p_k = 1$.

Result: $p_0 = 0,12963 \qquad p_1 = 0,18519 \qquad p_2 = 0,2087 \qquad p_3 = 0,4815$

The utilisation of server A = P(server A is busy): $U_A = p_2 + p_3 = 0,6852$

(c) Utilisation of server B: $U_B = p_1 + p_3 = 0,6667$

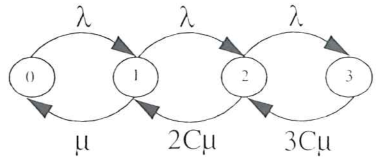
(d) *Alternative 1:*
Carried load is defined as:

$$\rho_c = \lambda_{eff} \cdot E[X] = [(\lambda_1 + \lambda_2)p_0 + (\lambda_1 + \lambda_2)p_1 + \lambda_2 p_2] \cdot \frac{1}{\mu} \approx 1,35186 \text{ Erlangs}$$

*Alternative 2:*
Carried load is the same as the average number of jobs in the servers, which in this case is: $\bar{N}_s = U_A + U_B \approx 1,35186$, since the utilisation of a specific server is the same as the average number of jobs in that server.

20. In this system the average service time for a job depend on the state of the system.

(a) State diagram:




The "cut method" gives the following balance equations:

$$\begin{cases} \lambda p_0 = \mu p_1 \\ \lambda p_1 = 2\mu p_1 \\ \lambda p_2 = 3\mu p_1 \end{cases}$$

which means that

$$\begin{cases} p_1 = \frac{\lambda}{\mu} p_0 \\ p_2 = \left(\frac{\lambda}{\mu}\right)^2 \frac{1}{2C} p_0 \\ p_3 = \left(\frac{\lambda}{\mu}\right)^3 \frac{1}{6C} p_0 \end{cases}$$

$p_0$ is as usual found with the normalisation condition:

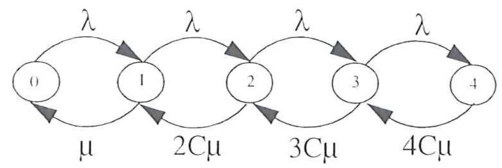
$$p_0 = \frac{1}{1 + \frac{\lambda}{\mu} + \left(\frac{\lambda}{\mu}\right)^2 \frac{1}{2C} + \left(\frac{\lambda}{\mu}\right)^3 \frac{1}{6C}}$$

(b) P(a call is blocked)=$P_b = \frac{\lambda_b}{\lambda} = \frac{\lambda p_3}{\lambda} = p_3 \approx 0,2195$

(c) In this case it is not easy to use the definition of carried load, since the average service time varies. Use instead the fact that the carried load is equal to the average number of jobs in the servers: $\rho_c = \bar{N}_s = 0 \cdot p_0 + 1 \cdot p_1 + 2 \cdot p_2 + 3 \cdot p_3 \approx 1,52$

(d) Alternative 1: One more processor.

New state diagram:




Derive new balance equations, and find the new probability distribution.

The blocking probability, $P_b = p_4 \approx 0,1385$

Alternative 2: A better memory.

Use $C$=0.95 in the calculations in (a) and (b).

The blocking probability, $P_b = p_3 \approx 0,1512$.

Conclusion: Alternative 1 gives the lowest blocking probability.

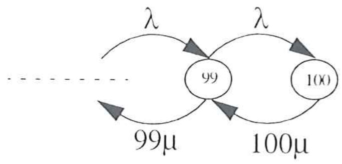
21.

(a) We have conducted measurements on an *M/G/m*-system without queue (Erlang system).

The measurements showed that $\begin{cases} p_{100} = \dfrac{72,3}{3600} \approx 0,020 \\ p_{99} = \dfrac{82,2}{3600} \approx 0,023 \end{cases}$

The offered load is defined as $\rho_o = \lambda \cdot E[X] = \dfrac{\lambda}{\mu}$

The end of the Markov chain for this system is as follows:




This means that we can derive the following equation: $\lambda p_{99} = 100\mu p_{100}$, from which we can find the offered load: $\rho_o = \dfrac{\lambda}{\mu} = \dfrac{100p_{100}}{p_{99}} \approx 87$

(b) Let $T$ be the interarrival time between två calls (random variable)

Exponentially distributed interarrival time means that: $f_T(t) = \lambda e^{-\lambda t}$,

where $\lambda = 400$ calls/hour, which is about 0.11 calls/second.

Therefore, $P(T \leq t) = F_T(t) = 1 - e^{-0,11t}$

which means that $P(T \leq 0,02) \approx 0,0022$.

In average $400 \cdot 0,0022 \approx 0,9$ calls/hour are lost.

22.M/M/10 loss system with $\lambda = N/2\, s^{-1}$ and $E[X] = 1\, s$ (Erlang system).

(a) The conditions for Erlang's loss formula are fulfilled which means that P(call blocked)=$E_{10}(\rho_o)$ where $\rho_o$ is the offered load.

$\rho_o = \lambda \cdot E[X] = \frac{N}{2}$ which means that P(call blocked)=$E_{10}(N/2) < 0,01$

By using the Erlang tables we find that $\frac{N}{2} \leq 4,4 \Rightarrow N \leq 8,8$ to achieve a low enough blocking probability.

Therefore, maximum 8 sources can send jobs to the system.

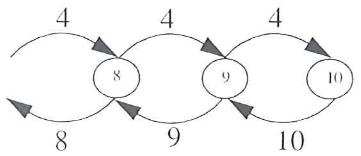
(b) If N=8 then $\rho_o$=4.

Carried load=

$\rho_c = \lambda_{eff} \cdot E[X] = \lambda \cdot (1 - E_{10}(\rho_o)) \cdot E[X] = \rho_o \cdot (1 - E_{10}(\rho_o)) \approx 3,98$

Carried load per server $= \frac{\rho_o}{10} \approx 0,4$ Erlang, since the load is evenly distributed.

(c) The end of the Markov chain:




P(at least 8 servers busy) = $p_8 + p_9 + p_{10}$. Also $p_{10} = E_{10}(4) \approx 0,005308$

The "cut method" gives the following equations:

$$\begin{cases} 4p_9 = 10p_{10} \Rightarrow p_9 = \frac{5}{2} \cdot p_{10} \approx 0,01327 \\ 4p_8 = 9p_9 \Rightarrow p_8 = \frac{9}{4} \cdot p_9 \approx 0,02986 \end{cases}$$

Therefore, P(at least 8 servers busy)=0.0484