# ERP 420

## Research Project

## Single Station Queuing Systems

Dr A. De Freitas

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

# Performance Measures

**Probability of the Number of Jobs in the System $\pi_K$:**

It is often possible to describe the behavior of a queueing system by means of the probability vector of the number of jobs in the system $\pi_k$ The mean values of most of the other interesting performance measures can be deduced from $\pi_K$:

$\pi_K$ = P[there are ***k*** jobs in the system].

# Performance Measures

**Utilization** $\rho$: If the queueing system consists of a single server, then the utilization $\rho$ is the fraction of the time in which the server is busy, i.e., occupied. In case there is no limit on the number of jobs in the single server queue, the server utilization is given by

$$\rho = \frac{mean\,service\,time}{mean\,interarrival\,time} = \frac{arrival\ rate}{service\ rate} = \frac{\lambda}{\mu}$$

$\rho$ *is also called as* the traffic intensity (sometimes called occupancy). It is defined as the average arrival rate (lambda) divided by the average service rate (mu).

The utilization of a service station with multiple servers is the mean fraction of active servers. Since $m\mu$ is the overall service rate, we have:

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

# Performance Measures

**Throughput** $\lambda$: The throughput of an elementary queueing system is defined as the mean number of jobs whose processing is completed in a single unit of time, i.e., the departure rate. Since the departure rate is equal to the arrival rate $\lambda$ for a queueing system in statistical equilibrium, the throughput is given by

$$\lambda = m \cdot \rho \cdot \mu$$

**Note:** In the case of finite buffer queueing system, throughput can be different from the external arrival rate.

# Performance Measures

**Number of Jobs in the System** $K$ : The number of jobs in the queueing system is represented by $K$ . Then

$$\bar{K} = \sum_{k=1}^{\infty} k.\pi_k$$

The mean number of jobs in the queueing system $\bar{K}$ (average number of packets under transmission) and the mean queue length $\bar{Q}$ (Average time spent by a packet waiting in queue (average packet delay) can be calculated using one of the most important theorems of queueing theory, *Little's theorem.*

$$\bar{K} = \lambda \bar{T}$$

$$\bar{Q} = \lambda \bar{W}$$

# Performance Measures

**Waiting Time** $W$ **:** The waiting time is the time that a job spends in a queue waiting to be serviced. Therefore, we have

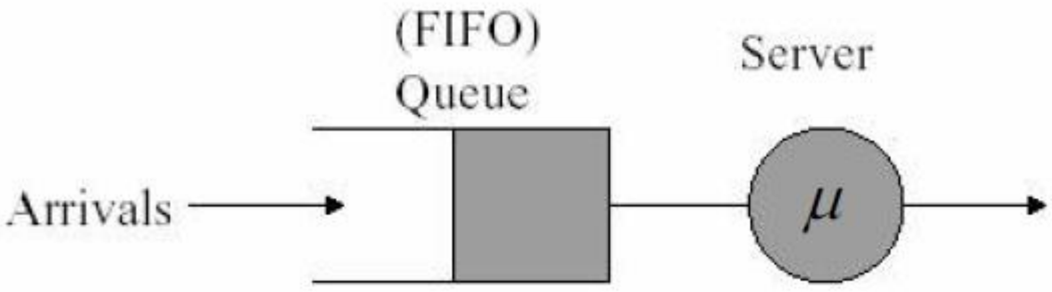
Response time = waiting time + service time ( $\frac{1}{\mu}$ ).

Since $W$ and $T$ are usually random numbers, their mean is calculated. Then

$$\bar{T} = \bar{W} + \frac{1}{\mu}$$

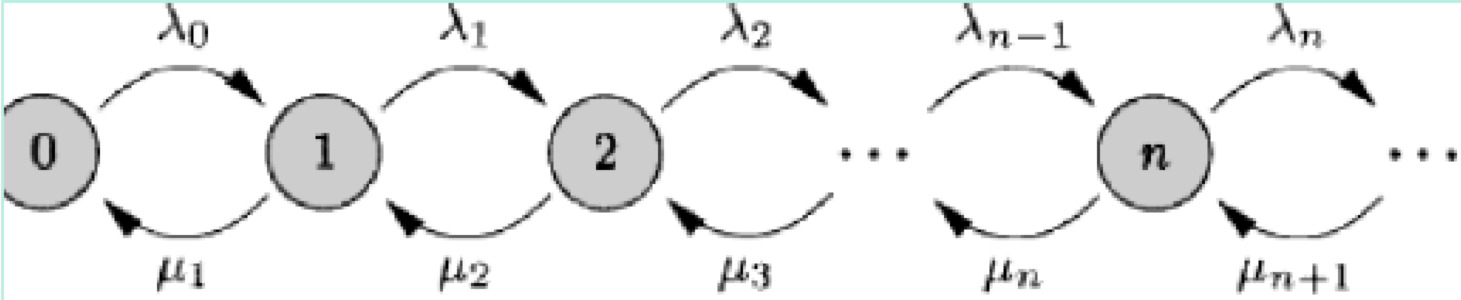
# Markov Queues – M/M/1

## M/M/1 Queue




- Notation                A/B/C
  - Arrivals process
  - Service Distribution
  - Number of Servers
- M ≡ Markov process. Transition probabilities depend only on current state. Memoryless.
- Poisson/Exponential is special case of Markov.

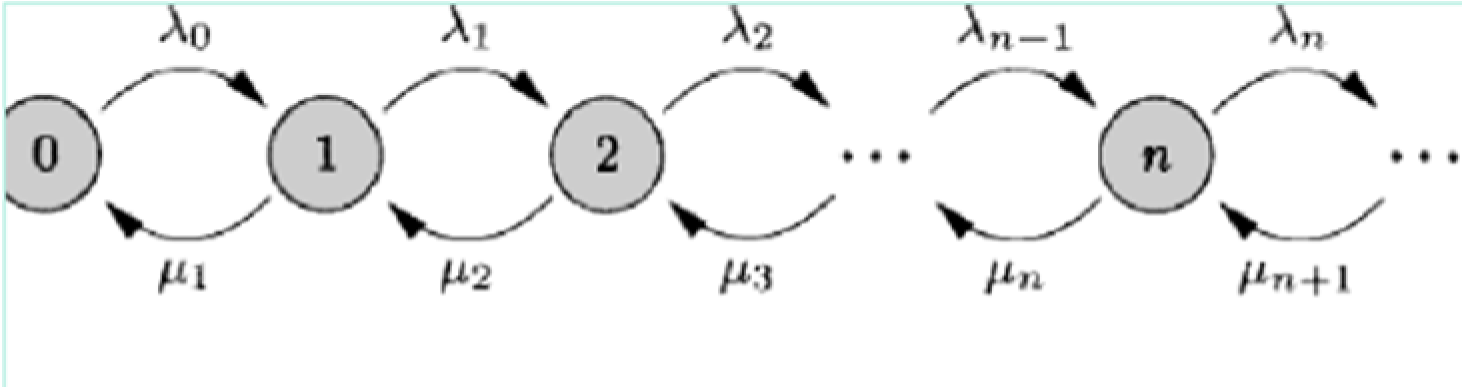
# Markov Queues – M/M/1




Birth–death process.

Birth-death processes are Markov chains where transitions are allowed only between neighboring states.

# Markov Queues – M/M/1




Birth–death process.

$$\mathbf{Q} = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & \cdots \\ 0 & 0 & \mu_3 & -(\lambda_3+\mu_3) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

# Markov Queues – M/M/1

$$\mathbf{0} = \boldsymbol{\pi}\mathbf{Q}$$

$$\boldsymbol{\pi}\mathbf{1} = \sum_{i \in S} \pi_i = 1$$

By means of the recursion, all the state probabilities can be expressed in terms of that of the state 0, $\pi_0$,

$$\boxed{\pi_k = \frac{\lambda_{k-1}\lambda_{k-2}\cdots\lambda_0}{\mu_k\mu_{k-1}\cdots\mu_1}\pi_0 = \prod_{i=0}^{k-1}\frac{\lambda_i}{\mu_{i+1}}\pi_0}$$

The probability $\pi_0$ is determined by the normalization condition $\pi_0$

$$\boxed{\pi_0 = \frac{1}{1 + \frac{\lambda_0}{\mu_1} + \frac{\lambda_0\lambda_1}{\mu_1\mu_2} + \cdots} = \frac{1}{1 + \sum_{k=1}^{\infty}\prod_{i=0}^{k-1}\frac{\lambda_i}{\mu_{i+1}}}}$$

# Markov Queues – M/M/1

- How we can derive a general formula for the probability of having ***n*** packet in the system ($\pi_k$) or $\pi_0$ ?

$$\lambda_k = \lambda, \quad k \geq 0$$

Then,

$$\mu_i = \mu, \quad i \geq 1$$

# Markov Queues – M/M/1

- for the steady-state probability that there are ***k*** jobs in the system we get:

$$\pi_k = \prod_{k=0}^{k-1}(\frac{\lambda}{\mu})\pi_0 = \rho^k \pi_0$$

- How to find $\pi_0$ ? $\sum_{k=0}^{\infty}\pi_k = 1 \ =>$

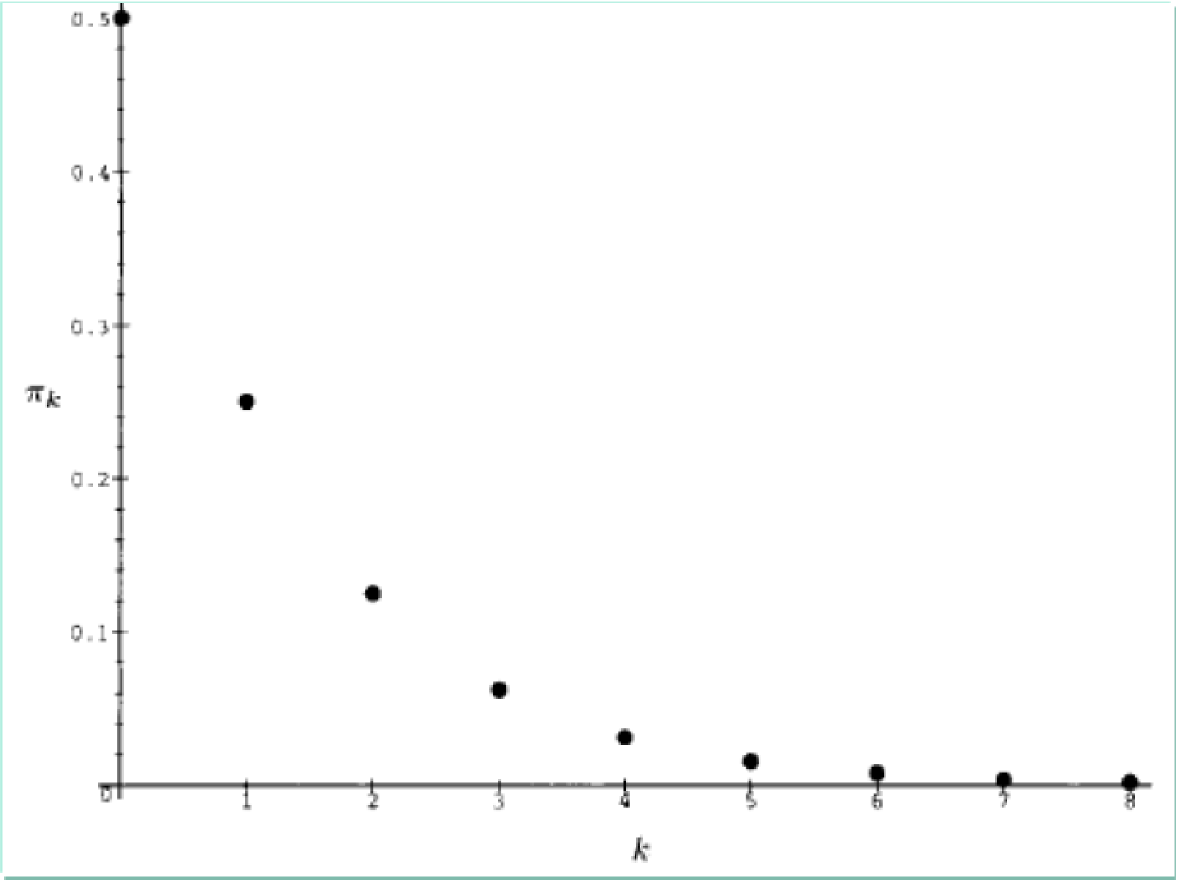
$$\sum_{k=0}^{\infty}\rho^k\pi_0 = 1 => \pi_0\sum_{k=0}^{\infty}\rho^k = 1 =>$$

$$\pi_0(\frac{1}{1-\rho}) = 1 => \pi_0 = 1-\rho$$

$$\pi_k = \rho^k(1-\rho)$$

# Markov Queues – M/M/1

- In the following Fig, we plot $\pi_k = (1-\rho)\rho^k \quad \rho = \frac{1}{2}$



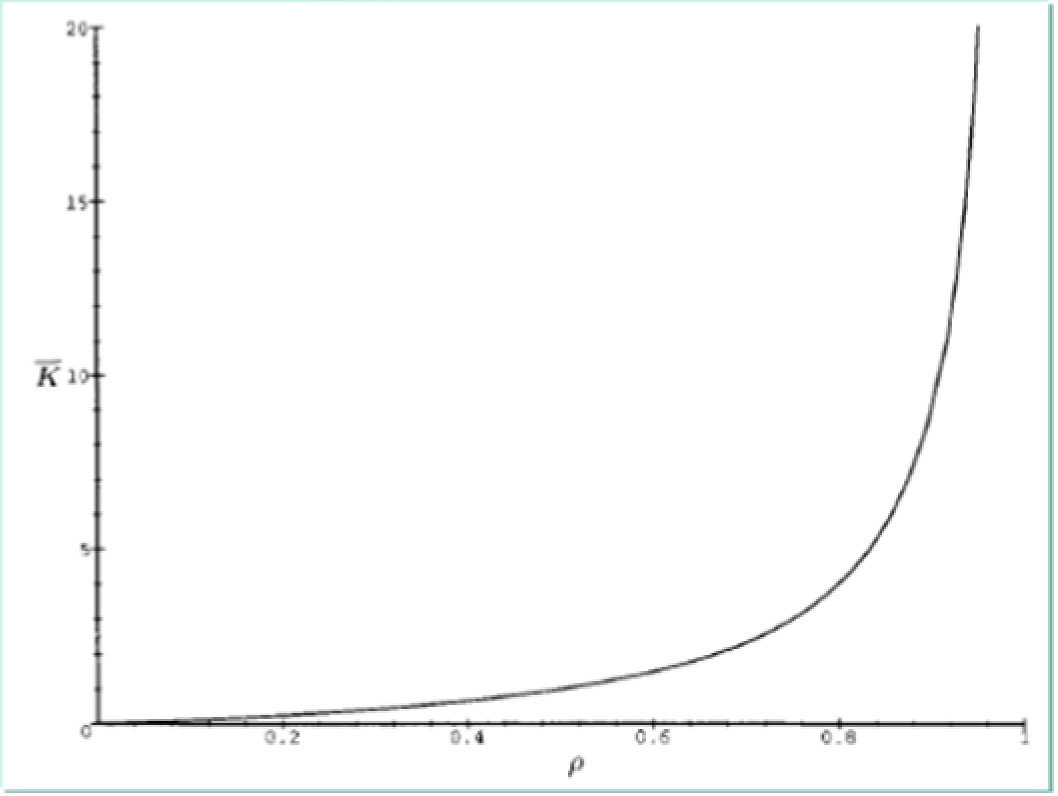
# Markov Queues – M/M/1

- The mean number of jobs is obtained using the following Equation:

$$\bar{K} = \sum_{k=0}^{\infty} k.\pi_k$$

- the mean number of jobs is plotted as a function of the utilization ($\rho$).

$$\bar{K} = \frac{\rho}{1-\rho}$$



# Markov Queues – M/M/1

**Waiting Time** $\bar{W} = \dfrac{\rho}{\mu - \lambda}$

Response time = waiting time + service time $= \dfrac{1}{\mu - \lambda}$

If $\lambda$ is very small, then the queue is almost always empty when a customer arrives. In that case,the average time $T$ is $1/\mu$ and is equal to the average time the cashier spends serving a customer.

Another useful result is that the average number $L$ of customers in line or with the server is given by L=$\lambda$ .w

$$L = \frac{\rho^2}{1-\rho}$$

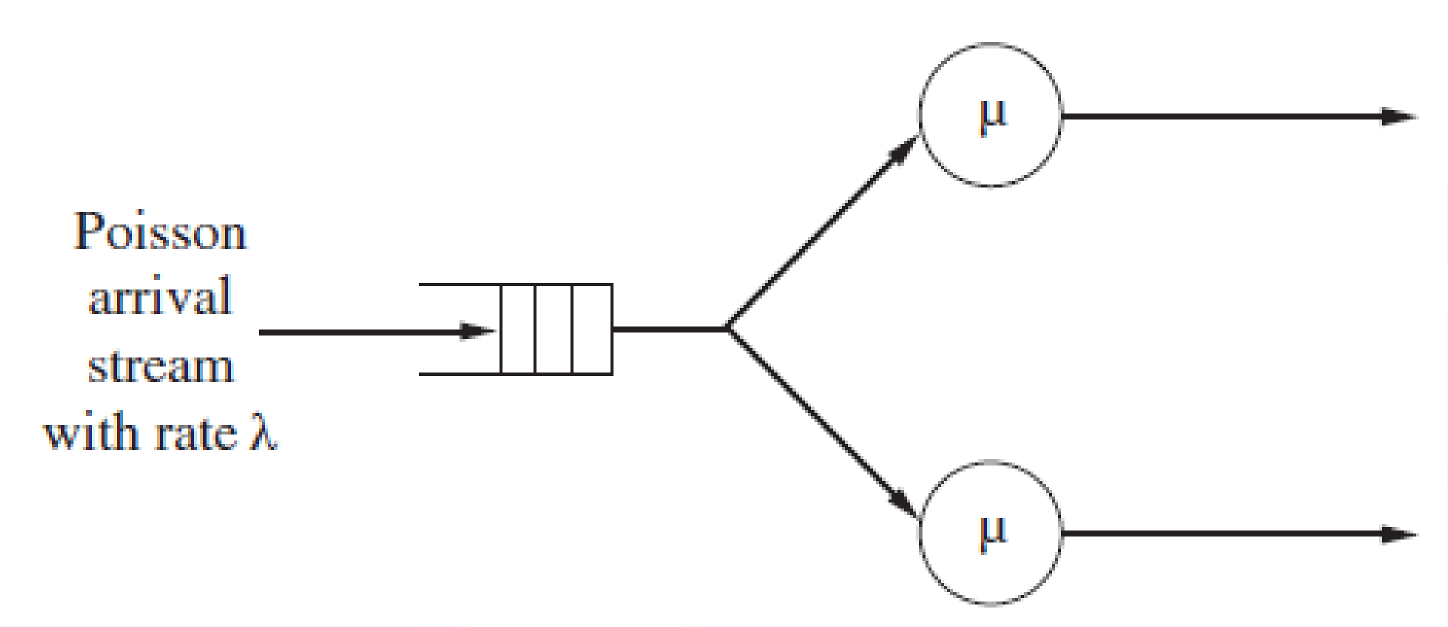
UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

# Markov Queues – M/M/1

- In M/M/1 utilisation is equal to the traffic intensity and throughput is equal to arrival rate.

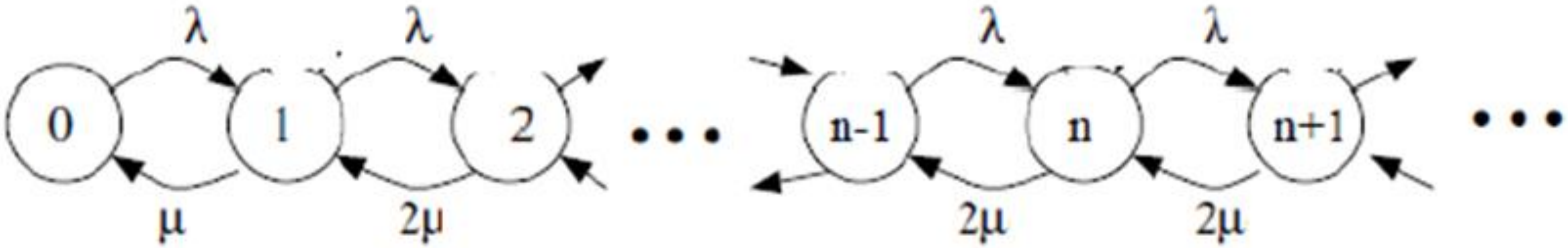
$$utilisation : 1-\pi_0 = 1-(1-\rho) = \rho(traffic\ intensity)$$
$$Throughput : (1-\pi_0)\mu = (1-(1-\rho))\mu = \lambda$$

# Markov Queues – M/M/2




- Draw the state diagram

# Markov Queues – M/M/2




- Find an expression for $\pi_k$ in terms of $\pi_0$

# Markov Queues – M/M/2

$$\Pi_k = 2\rho^k \Pi_0$$

- Find an expression for $\pi_0$ and substitute into $\pi_k$

# Markov Queues – M/M/2

$$\sum_{k=0}^{\infty} \pi_k = 1$$

$$\Pi_0 + \sum_{k=1}^{\infty} \pi_k = 1$$

$$\Pi_0 + \sum_{k=1}^{\infty} 2\rho^k \Pi_0 = 1$$

$$\Pi_0 + 2\Pi_0 \rho \sum_{k=1}^{\infty} \rho^{k-1} = 1$$

$$\Pi_0 + 2\Pi_0 \rho \sum_{k=0}^{\infty} \rho^k = 1$$

$$\Pi_0 + 2\Pi_0 \rho (\frac{1}{1-\rho}) = 1$$

$$\Pi_0 (1 + \frac{2\rho}{1-\rho}) = 1$$

$$\Pi_0 = \frac{1-\rho}{1+\rho}$$

$$\Pi_k = 2\rho^k \Pi_0 = 2\rho^k (\frac{1-\rho}{1+\rho})$$

# Markov Queues – M/M/2

- Find average the number of packets in the system

$$\bar{K} = \sum_{k=0}^{\infty} k.\pi_k = \frac{2\rho}{(1+\rho)(1-\rho)}$$

# Markov Queues – M/M/2

- Find average waiting time in the system

$$\bar{T} = \frac{1}{\mu(1+\rho)(1-\rho)}$$