Tutorial #3

M/M/1

## Rules

### Recall

- $\lambda \rightarrow$ arrival rate
- $T_s \rightarrow$ service time
- $T_Q \rightarrow$ response time
- $L_Q \rightarrow$ Avg # in system
- $T_W \rightarrow$ Avg waiting time
- $\rho \rightarrow$ utilization

$$\boxed{\text{Speed} = \frac{1}{T_s} = \mu} \quad (5)$$

$$\boxed{\rho = \lambda T_s = \frac{\lambda}{\mu}} \quad (6)$$

$$\boxed{L_Q = \frac{\rho}{1-\rho}} \quad (20)$$

From Little's law

$$\boxed{T_Q = \frac{L_Q}{\lambda} = \frac{T_s}{1-\rho}} \quad (21)$$

$$\boxed{T_Q = T_s + T_W} \quad (22)$$

$$\boxed{T_W = \rho T_Q} \quad (23)$$

From Little's

$$\boxed{L_W = \lambda T_W} \quad (24)$$

Avg # of jobs in a queue

- Jobs of non-zero waiting time, Distrib.

$$\boxed{T_{WD} = T_Q} \quad (25)$$

# of jobs in sys (R.V)

$$\boxed{P(n_Q = x) = \rho^x(1-\rho)} \quad (26)$$

$$\boxed{P(n_Q \geq x) = \rho^x} \quad (27)$$

$T_Q$ r W

$$\boxed{P(r \leq t) = 1 - e^{\frac{-t}{T_Q}}} \quad (28)$$

* The CDF of the response time r (exp distrib.)

## Examples

34

1) on a network gateway, measurements show packets arrive @ mean rate of 125 packets/sec & the gateway takes about 2 ms ($T_s$) to forward them. Assuming an M/M/1 model, determine -

1) mean time spent in the gateway (system)

$$T_Q = \frac{L_Q}{\lambda} = \frac{T_s}{1-\rho} = \frac{T_s}{1-\lambda T_s}$$

$$= \frac{2 \times 10^{-3}}{1-125(0.002)} = 2.67 \text{ ms}$$

2) The mean # of packets in the gateway buffers

$$L_Q = \frac{\rho}{1-\rho} = \frac{\lambda T_s}{1-\lambda T_s} = \frac{0.25}{1-0.25}$$

$$= 0.3\dot{3} \text{ Packets}$$

3) P(buffer of overflow) if gateway has only 13 buffers

$$P(n_Q \geq 13) = \rho^x = 0.25^{13}$$

I'm blocked [illegible]

4) # of buffer needed to keep 1 packet loss / million

$$P(n_Q \geq x) = 10^{-6}$$

$$\rho^x = 10^{-6}$$

$$x \lg(\rho) = \lg(10^{-6})$$

$$x = \frac{-6}{\lg(0.25)} = 9.966 \simeq 10 \text{ buffers}$$

*Dr & TA have diff sol & I just follow the dr's (for 2)

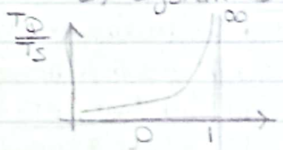
From (23) & (28)

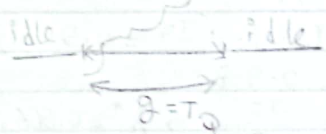
* The CDF of the waiting time

$$\boxed{P(W \leq t) = 1 - \rho e^{\frac{-t}{T_{WD}}}} \quad (29)$$

waiting time for those who must wait; $T_{WD} \rightarrow T_Q$

* As long as $\rho$ is well less than 1, system is stable




* The duration of the busy period

$$\boxed{g = T_Q} \quad (30)$$

* The # of jobs in this period

$$\boxed{h = \frac{g}{T_s} = \frac{T_Q}{T_s} = \frac{1}{1-\rho}} \quad (31)$$

* $\rho \uparrow \; T_Q \uparrow$

$$\boxed{\mu_{old} - \lambda_{old} = \mu_{new} - \lambda_{new}} \quad (32)$$

(2) The avg response time ($T_Q$) on a db system is 3s. During 1-min ($T$) observation, the idle time on the sys was 10 s (idle time). Using an M/M/1

(i) sys utilization

$$\rho = U = \frac{B}{T} = \frac{60-10}{60} = 0.8\dot{3}$$

(ii) Avg service time per query

$$T_Q = \frac{T_s}{1-\rho} = \frac{T_s}{1-0.83} = 3 \text{ sec}$$

$$T_s = 0.5 \text{ sec} \qquad \lambda = \frac{\rho}{T_s} = 1.66 \text{ sec}$$

(iii) Avg # of jobs in system

$$L_Q = \frac{\rho}{1-\rho} = \frac{0.8\dot{3}}{1-0.83} = 4.88$$

(iv) P(# jobs > 10)

$$P(n_Q \geq 11) = \rho^{11} = (0.8\dot{3})^{11}$$

(v) 90-percentile response time

$$P(r \leq t) = 1 - e^{\frac{-t}{T_Q}} = 0.9$$

$$1 - e^{\frac{-t}{3}} = 0.9$$

$$t = 6.91 \text{ sec}$$

(vi) 90-percentile waiting time

$$1 - \rho e^{\frac{-t}{T_{WD}}} = 0.9$$

$$1 - 0.83 e^{\frac{-t}{3}} = 0.9$$

$$t = 6.349 \text{ sec}$$

(vii) mean length of a busy period

$$g = T_Q = 3 \text{ sec}$$

# Tutorial #3

## Rules

### Recall

- * $\lambda$ → arrival rate
- * $T_S$ → service time
- * $T_Q$ → response time
- * $L_Q$ → Avg # in system
- * $T_W$ → Avg waiting time
- $\rho$ → utilization

$$\boxed{\text{Speed} = \frac{1}{T_S} = \mu} \quad (5)$$

$$\boxed{\rho = \lambda T_S = \frac{\lambda}{\mu}} \quad (6)$$

$$\boxed{L_Q = \frac{\rho}{1-\rho}} \quad (20)$$

From little's law

$$\boxed{T_Q = \underbrace{\frac{L_Q}{\lambda}}_{\text{general}} = \frac{T_S}{1-\rho}} \quad (21)$$

$$\boxed{T_Q = T_S + T_W} \quad (22)$$

$$\boxed{T_W = \rho T_Q} \quad (23)$$

From little's

$$\boxed{L_W = \lambda T_W} \quad (24)$$

Avg # of jobs in a queue

* Jobs of non-zero waiting time,

$$\boxed{T_{WD} = T_Q} \quad (25)$$

Distrib.
$n_Q$
→ $r$
→ $W$

→ # of jobs in sys (R.V)

$$\boxed{P(n_Q = X) = \rho^X (1-\rho)} \quad (26)$$

$$\boxed{P(n_Q \geq X) = \rho^X} \quad (27)$$

$n_Q$

$$\boxed{P(r \leq t) = 1 - e^{\frac{-t}{T_Q}}} \quad (28)$$

* The CDF of the response time $r$ (exp distrib.)

## Examples

34

1) on a network gateway, measurements show packets arrive @ mean rate of 125 packets/sec ($\lambda$) & the gateway takes about 2 ms ($T_s$) to forward them. Assuming an M/M/1 model, determine..

1) mean time spent in the gateway (system)

$$T_Q = \frac{L_Q}{\lambda} = \frac{T_S}{1-\rho} = \frac{T_S}{1-\lambda T_S}$$

$$= \frac{2\times10^{-3}}{1-125(0.002)} = 2.67 \text{ ms}$$

* 2) The mean # of packets in the gateway buffers

$$L_Q = \frac{\rho}{1-\rho} = \frac{\lambda T_S}{1-\lambda T_S} = \frac{0.25}{1-0.25}$$

$$= 0.3\dot{3} \text{ Packets}$$

* 3) P(buffer of) if gateway has only 13 buffers → overflow

$$P(n_Q \geq 13) = \rho^X \doteq 0.25^{13}$$

(انا داخل الـ sys وفيه 13 او اكثر فـ I'm blocked)

4) # of buffer needed to keep 1 Packet loss / million

$$P(n_Q \geq X) = 10^{-6}$$

$$\rho^X = 10^{-6}$$

$$X \lg(\rho) = \lg(10^{-6})$$

$$X = \frac{-6}{\lg(0.25)} = 9.966 \simeq 10 \text{ buffers}$$

* Dr & TA have diff sol & I just follow the dr's (for 2)

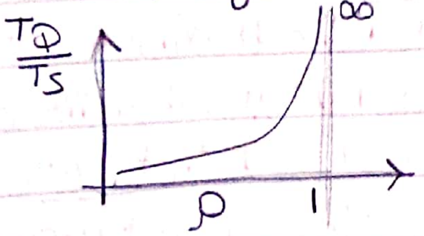
From (23) & (28)

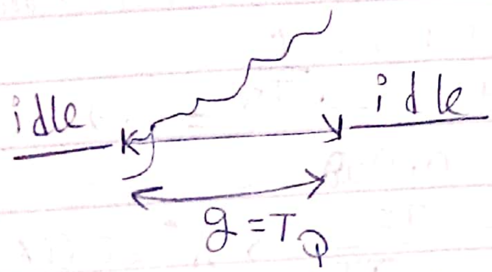
* The CDF of the waiting time

$$\boxed{P(W \leq t) = 1 - \rho e^{\frac{-t}{T_{WD}}}}$$ (29)

waiting time for those who must wait $T_{WD} \rightarrow T_Q$

* As long as $\rho$ is well less than 1, system is stable




* The duration of the busy period

$$\boxed{g = T_Q}$$ (30)

* The # of jobs in this period

$$\boxed{h = \frac{g}{T_S} = \frac{T_Q}{T_S} = \frac{1}{1-\rho}}$$ (31)

* $\rho \uparrow \; T_Q \uparrow$

$$\boxed{\mu_{old} - \lambda_{old} = \mu_{new} - \lambda_{new}}$$ (32)

(2) The avg response time on a db system is 3s ($T_Q$). During 1-min ($T$) observation, the idle time on the sys was 10 s (idle time). using an M/M/1..

(i) sys utilization

$$\rho = U = \frac{B}{T} = \frac{60-10}{60} = 0.8\dot{3}$$

(ii) Avg service time per query

$$T_Q = \frac{T_S}{1-\rho} = \frac{T_S}{1-0.83} = 3 \text{ sec}$$

$$T_S = 0.5 \text{ sec} \qquad \lambda = \frac{\rho}{T_S} = 1.66 \text{ sec}$$

iii) Avg # of jobs in system

$$L_Q = \frac{\rho}{1-\rho} = \frac{0.8\dot{3}}{1-0.8\dot{3}} = 4.88$$

iv) $P(\# \text{jobs} > 10)$

$$P(n_Q \geq 11) = \rho^{11} = (0.8\dot{3})^{11}$$

v) 90-percentile response time

$$P(r \leq t) = 1 - e^{\frac{-t}{T_Q}} = 0.9$$

$$1 - e^{\frac{-t}{3}} = 0.9$$

$$t = 6.91 \text{ sec}$$

vi) 90-percentile waiting time

$$1 - \rho e^{\frac{-t}{T_{WD}}} = 0.9$$

$$1 - 0.83 e^{\frac{-t}{3}} = 0.9$$

$$t = 6.349 \text{ sec}$$

vii) mean length of a busy period

$$g = T_Q = 3 \text{ sec}$$

(3) In a busy city street with crowds, there is a single public telephone. From time to time, a passer by dicides to make a phone call. The avg duration of a phone call is $T_s$ 2 min. & avg people arrive to make calls once every $1/\lambda$ 5 min. Assuming an M/M/1 model

i) How long on avg does each person have to wait?

$$T_W = \rho T_Q = \rho \frac{T_s}{1-\rho} = \frac{\lambda T_s^2}{1-\lambda T_s}$$

$$= \frac{\frac{1}{5}(2)^2}{1-(\frac{1}{5})(2)} = 1.3\dot{3} \text{ min}$$

ii) what is the 90th percentile of waiting time?

$$P(W \le t) = 1-\rho e^{\frac{-t}{T_{WD}}} = 0.9$$

$$1-(0.4)e^{\frac{-t}{3.33}} = 0.9$$

$$t = 4.616 \text{ min}$$

iii) How many people on avg, are waiting to make a call?

$$L_W = \lambda T_W = \frac{1}{5}(1.33) = 0.\dot{2}66$$

iv) How long, on avg, does a busy period last?

$$g = T_Q = 3.3\dot{3} \text{ min}$$

(4) Consider a LAN with 100 personal computers & a servers which maintains a common db for a query application. The avg time for the server to respond to a query is 0.6 s $T_s$, and the stdev is estimated to be equal to the mean

The query rate reaches $\lambda$ 20 qu/min we would like to answer the following

(i) what is the avg response time ignoring transmission overhead?

$$T_Q = \frac{T_s}{1-\rho} = \frac{0.6}{1-\frac{20}{60}0.6} = 0.75 \text{ sec}$$

ii) if a $T_{Q\,max}$ 2.3 sec response time is considered the maximum acceptable, what percent growth in message load can occur such that 90% of all response times do not exceed this maximum. $t \to 2.3$

$$P(r \le t) = 1-e^{\frac{-t}{T_Q}} = 0.9$$

$$T_Q^* = 0.998 \text{ sec}$$

$$= \frac{T_s}{1-\lambda T_s} \Rightarrow \lambda^* = 0.666 \text{ /sec}$$

$$\%\text{load change} = \left|\frac{\lambda^*-\lambda}{\lambda}\right| \times 100$$

$$= \left|\frac{0.666-0.333}{0.666}\right| = 100\,\%$$

iii) If 20% more utilization is experienced, what's the corresponding percentage increase in mean response $T_Q$ time?

$$\rho^* = (0.2)(1.2) = 0.24$$

$$T_Q^* = \frac{T_s}{1-\rho^*} = \frac{0.6}{1-0.24} = 0.789 \text{ sec}$$

$$\%\text{ Change} = \left|\frac{0.789-0.75}{0.75}\right| \times 100$$

$$= 5.2\,\%$$

(5) Given an M/M/1 model with mean arrival rate $\lambda = 5$ jobs/hr & mean service time $T_s = 6$ min. If a faster server is used with $T_s^* = 3$ min, then compute the mean arrival rate that can be tolerated for the same service levels

$$\mu - \lambda = \mu^* - \lambda^*$$

$$\frac{1}{6} - \frac{5}{60} = \frac{1}{3} - \lambda^*$$

$\lambda^* = 0.25$ jobs/min

(6) In a data communications network, messages arrive to be transmitted over a particular link. The avg time required to transmit a message is $0.6$ sec, and msgs arrive @ an avg rate of 1 msg/s. Assume an M/M/1

(i) P(an arriving msg does not have to wait)

$P(T_W = 0) = P(\text{server is idle})$
$= 1 - \rho = 1 - \lambda T_s = 0.4$

ii) How long does each take, on avg, to traverse the link?

$$T_Q = \frac{T_s}{1-\rho} = \frac{0.6}{1-0.6} = 1.5\,s$$

iii) 90th percentile of the waiting time of a message $t$

$$P(W \leq t) = 1 - \rho e^{-t/T_Q} = 0.9$$

$t = 2.688$ s

(iv) If the time required to transmit a message is 0.4 sec, What the avg arrival reduced to rate that can be tolerated while providing the same service level mean time-in-sys?

$$\frac{1}{T_s} - \lambda = \frac{1}{T_s^*} - \lambda^*$$

$$\frac{1}{0.6} - 1 = \frac{1}{0.4} - \lambda^*$$

$\lambda^* = 1.833$ msg/sec

v) P(@least 5 msg in sys)

$$P(n_Q \geq 5) = \rho^5 = (0.6)^5$$