Module 8: Lecture 3 of Process Simulation

Lecture 3: Queuing Theory & Markov Chain

dr. ir. Farid Vahdatikhaki

Image taken from: Dreamstime.com

What does it mean to improve this process?

Process Improvement

How is the efficiency of a process measured?

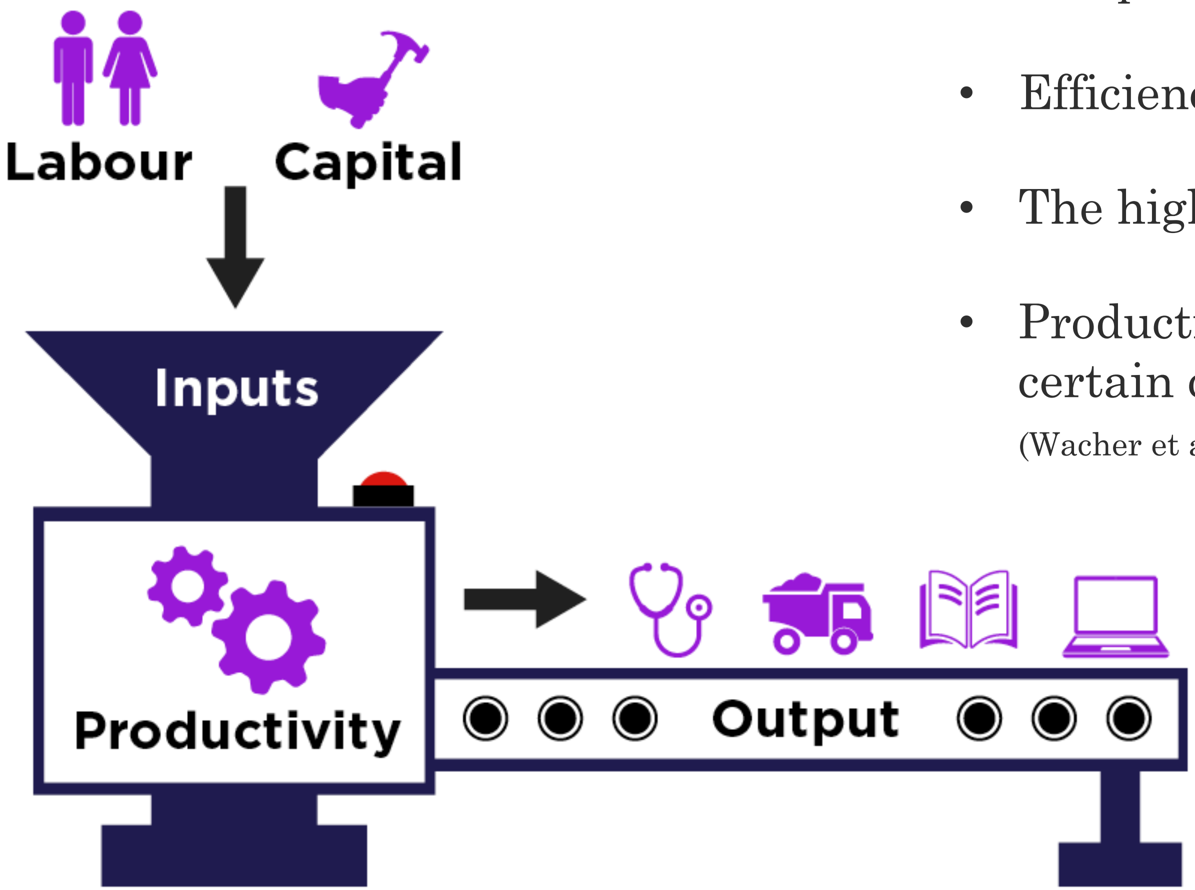
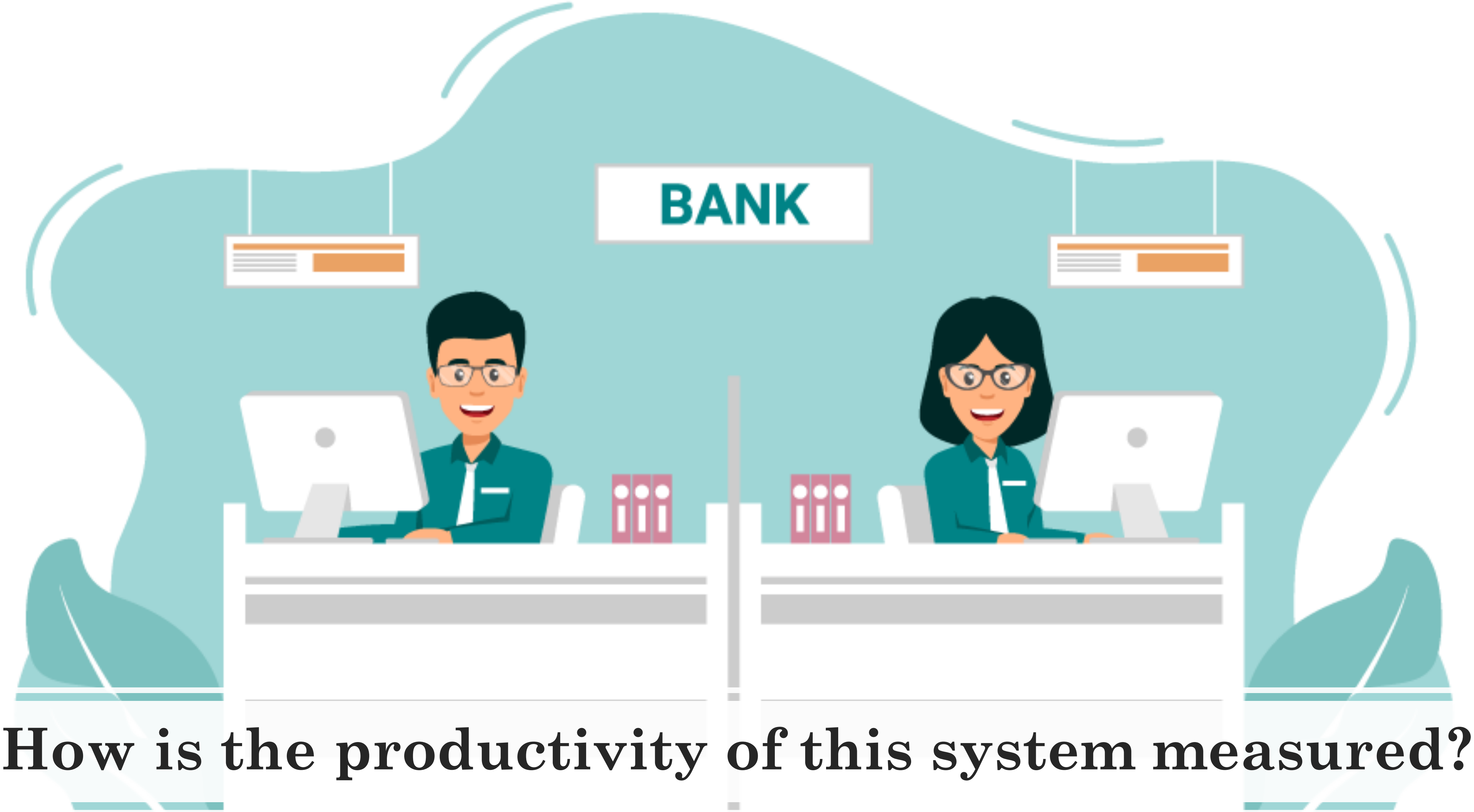


Image is taken from: westpac.com

- The process transforms certain inputs into outputs.
- Efficiency is defined in terms of the ratio of outputs to inputs
- The higher this ratio, the more efficient the process
- Productivity is defined as the amount of output produced with certain combinations of input resources (capital, labor, etc.) (Wacher et al. 2006).



How is the productivity of this system measured?

The image is taken from https://www.approvedcourse.com/

Let’s break down this process!

Clients

Image taken from: topsoftwarecompanies.co

Server

Image is taken from: www.kindpng.com

Customers arrive every **N** minutes (on average)

Productivity of this unit = **arrival rate** = $\lambda = \frac{1}{N}$

The teller can serve one client every **M** minutes (on average)

Productivity of this unit = **service rate** = $\mu = \frac{1}{M}$

What is the productivity of this system as a whole?

System Productivity

Server

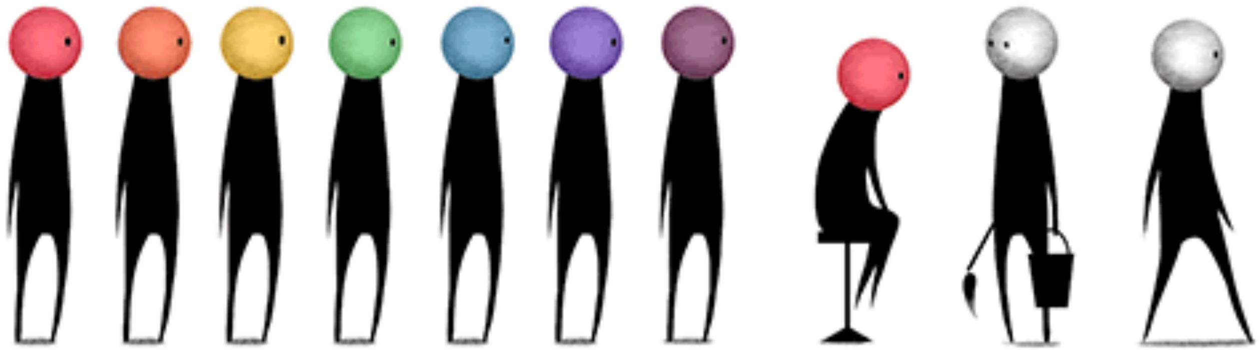
Image is taken from: www.kindpng.com

service rate $= \mu = \frac{1}{M}$

Clients

arrival rate $= \lambda = \frac{1}{N}$

Assuming that N and M are fixed (deterministic):



If $\mu < \lambda$ then **productivity** $= \mu$

If $\mu > \lambda$ then **productivity** $= \lambda$

What does this remind you of ?

Server

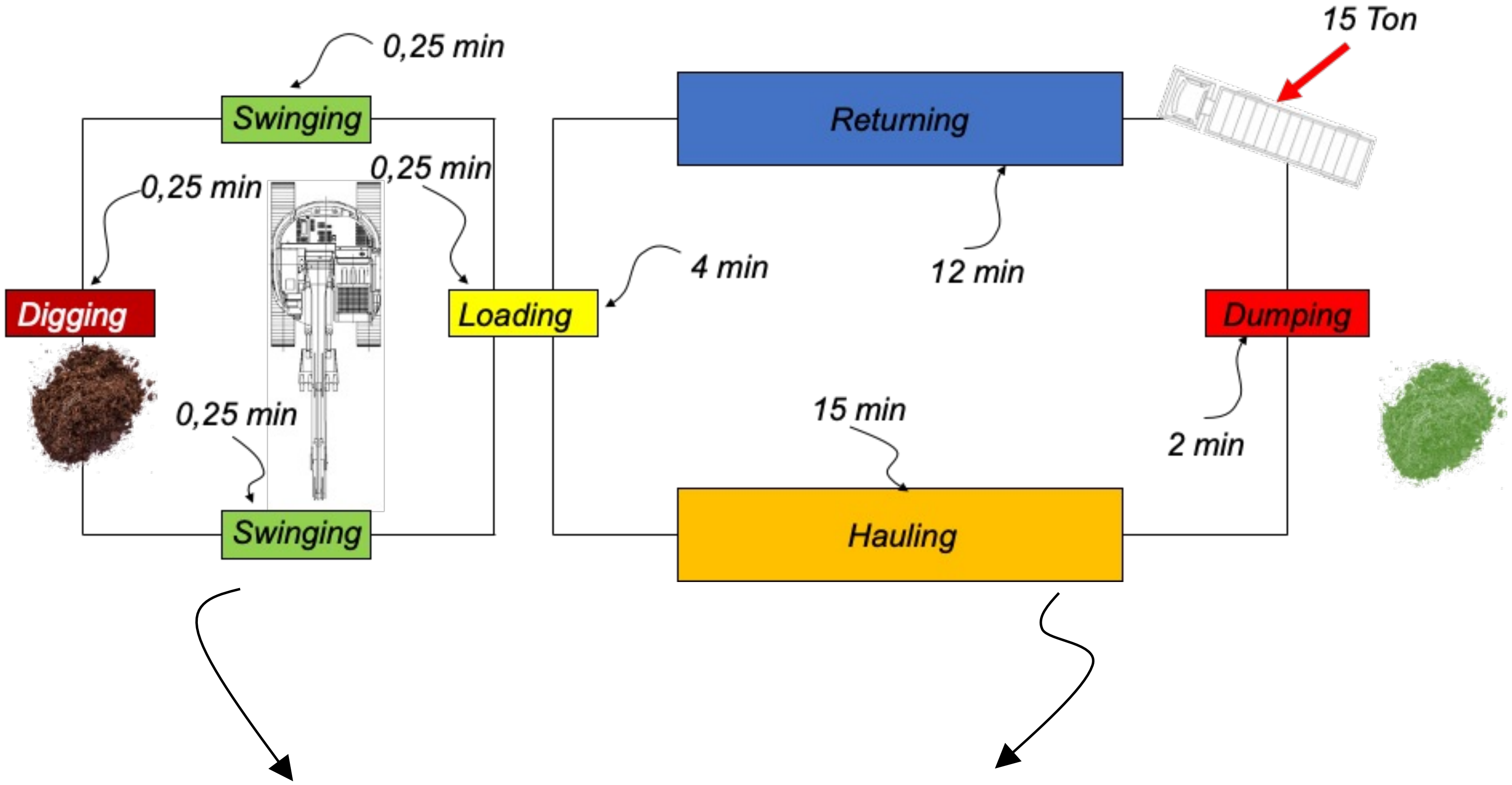
Image is taken from: www.kindpng.com

service rate = $\mu = \dfrac{1}{M}$

Clients

arrival rate = $\lambda = \dfrac{1}{N}$

=



Server

Clients

When is the system optimum?

Server

Image is taken from: www.kindpng.com

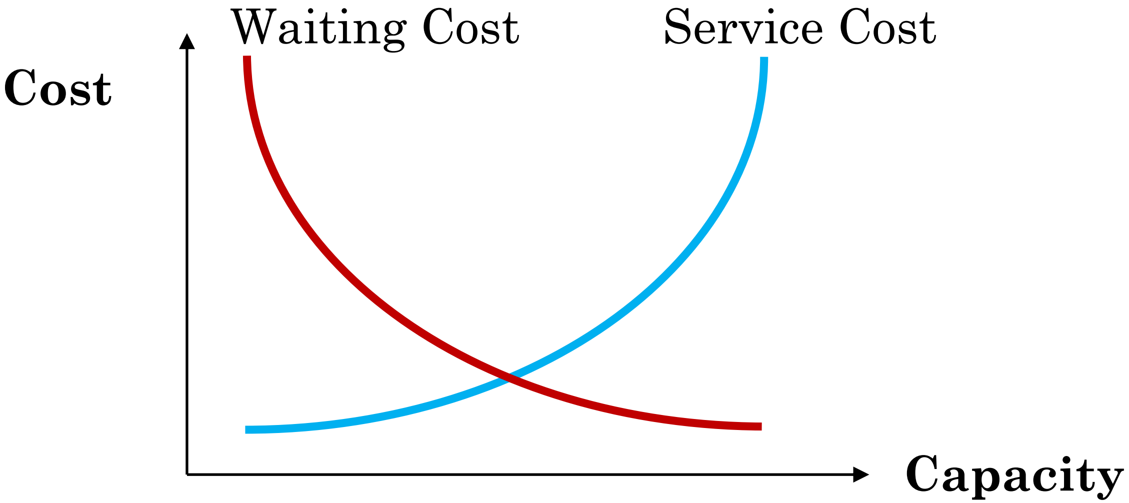
service rate = $\mu = \frac{1}{M}$

Clients

arrival rate = $\lambda = \frac{1}{N}$

If $\mu < \lambda$ $\longrightarrow$ Queues

If $\mu > \lambda$ $\longrightarrow$ Idle server



The System is optimum when $\mu = \lambda$

But, this almost never happens, right?

Again, you remember this from before!

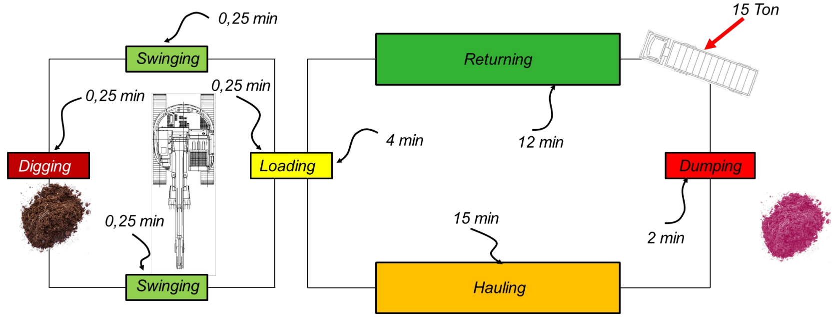
Server

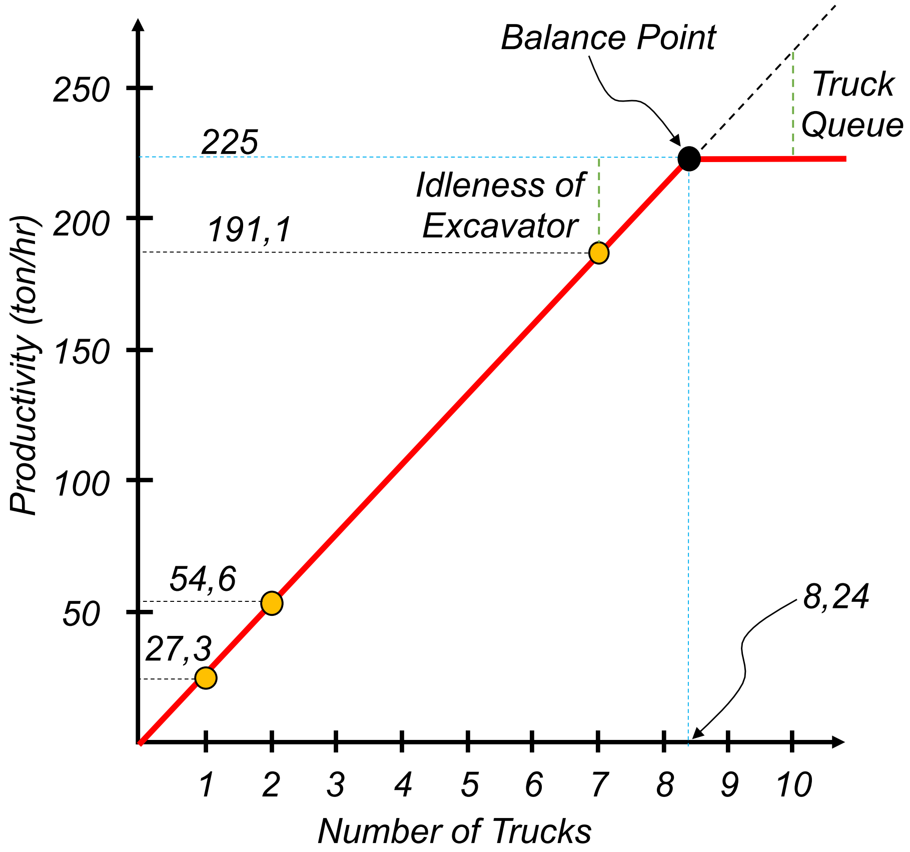
Image is taken from: www.kindpng.com

$$\text{service rate} = \mu = \frac{1}{M}$$

Clients

$$\text{arrival rate} = \lambda = \frac{1}{N}$$





When is the system optimum?

Server

Image is taken from: www.kindpng.com

service rate = $\mu = \frac{1}{M}$

Clients

arrival rate = $\lambda = \frac{1}{N}$

The arrival and service rates are almost never deterministic!!

Productivity of a stochastic queue model

Arriving Clients

Customers arrive every **N** minutes (on average)

Queue

Server

The teller can serve one client every **M** minutes (on average)

What are the possible states of this system?

S_0	S_1	S_2	S_3	S_4	• • • •	S_n
No client in the queue	1 client in the queue	2 clients in the queue	3 clients in the queue	4 clients in the queue		n clients in the queue
⬇	⬇	⬇	⬇	⬇		⬇
P_0	P_1	P_2	P_3	P_4		P_n

$$\underbrace{P_0 \quad P_1 \quad P_2 \quad P_3 \quad P_4 \quad \cdots \quad P_n}_{\text{Probabilities of occurrence}}$$

Probabilities of occurrence

Productivity of a stochastic queue model

What are the possible states of this system?

S_0	S_1	S_2	S_3	S_4	• • • •	S_n
No client in the queue	1 client in the queue	2 clients in the queue	3 clients in the queue	4 clients in the queue		n clients in the queue
P_0	P_1	P_2	P_3	P_4		P_n

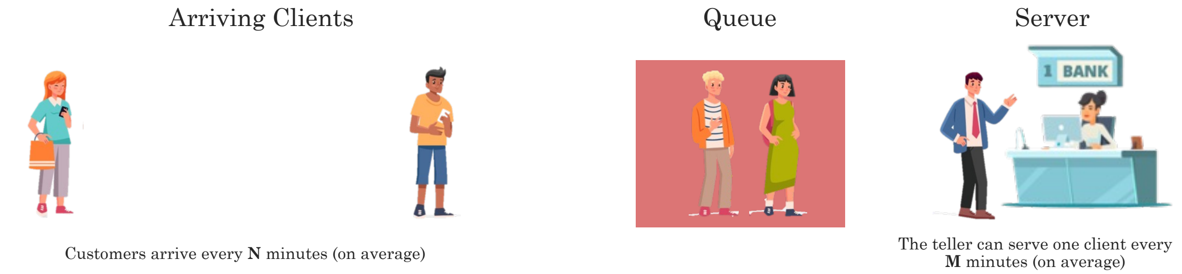
$$\underbrace{P_0 \quad P_1 \quad P_2 \quad P_3 \quad P_4 \quad \cdots \quad P_n}_{\text{Probabilities of occurrence}}$$

Thinking in a probabilistic way, we can state that Productivity (P) is service rate (μ) multiplied by the probability of having at least one unit in the service area (1-P_0).

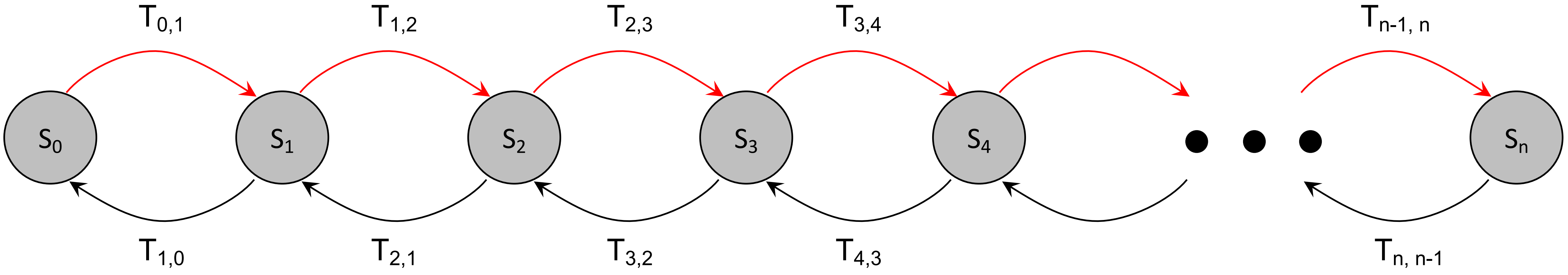
$$\textbf{Productivity} = \mu \times (1 - P_0)$$

What is the **unknown** in this equation?

Productivity of a stochastic queue model



What is the dynamic of this system?



Where $\mathbf{T}_{ij}$ is the probability of shifting from state S_i to state S_j

This is what is known as Markov Chain

Productivity of a stochastic queue model

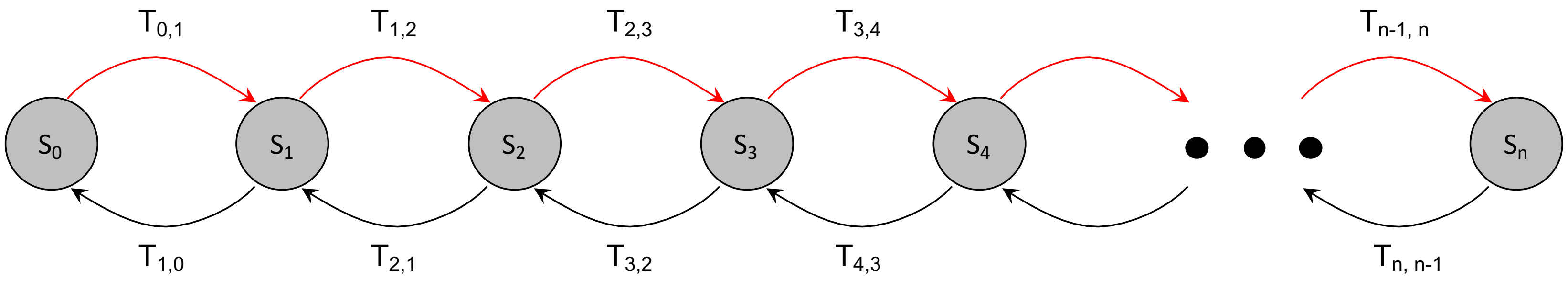
Server

Image is taken from: www.kindpng.com

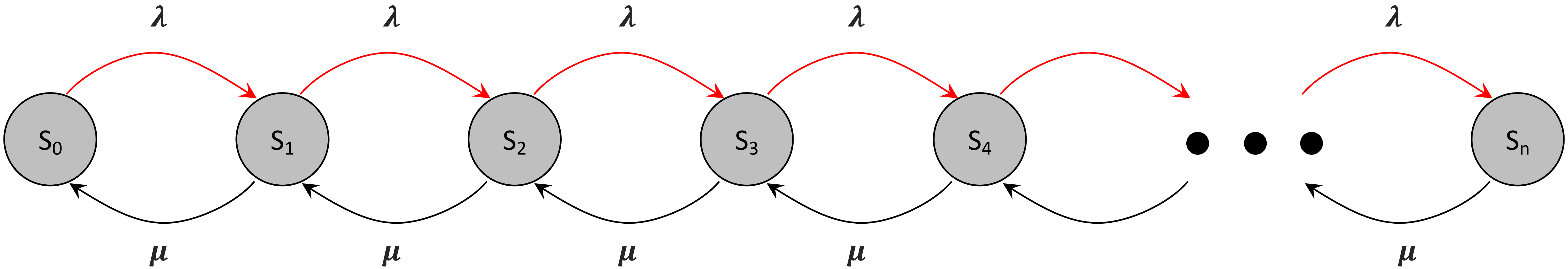
service rate = $\mu = \dfrac{1}{M}$

Clients

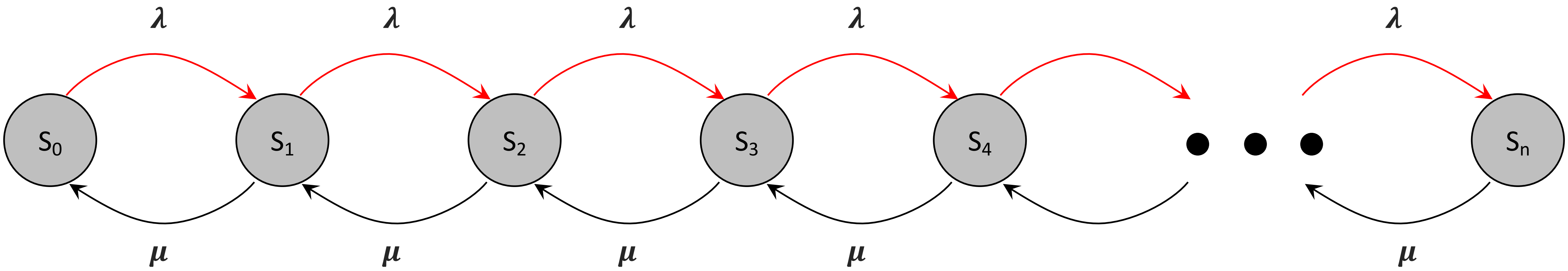
arrival rate = $\lambda = \dfrac{1}{N}$



What are $T_{i,i+1}$ and $T_{i,i-1}$ in this queue?

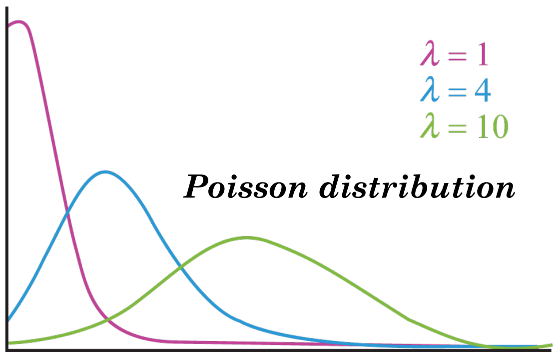


Productivity of a stochastic queue model

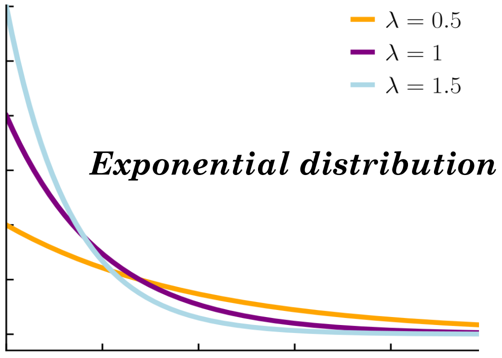


Important:

Again, the above statement holds only with the assumption that the arrival rate follows a Poisson distribution and service time follows an Exponential distribution



Poisson distribution assumes arrival of clients is random and independent



Exponential distribution measures the expected time for an event to happen in a continuous space

Productivity of a stochastic queue model

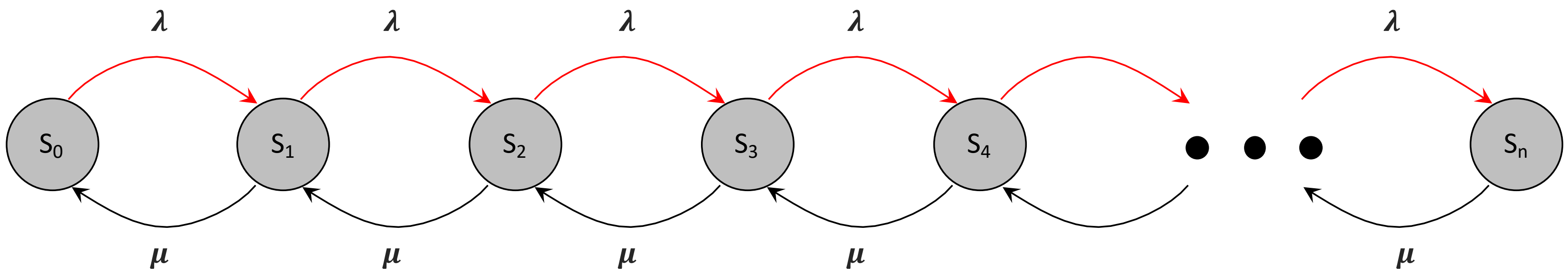
Server

Image is taken from: www.kindpng.com

service rate = $\mu = \frac{1}{M}$

Clients

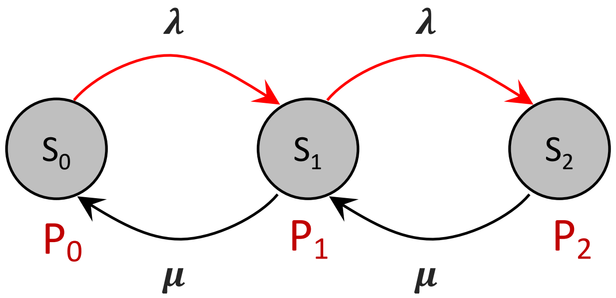
arrival rate = $\lambda = \frac{1}{N}$



What do we know about this system?

1- Summation of the probability of being in each state $\sum_{i=0}^{n} P_i = 1$

2- For the system to be stable, inflow to each state = outflow from that state



For S_1

$$P_0 \times \lambda + P_2 \times \mu = P_1 \times \mu + P_1 \times \lambda$$

Productivity of a stochastic queue model

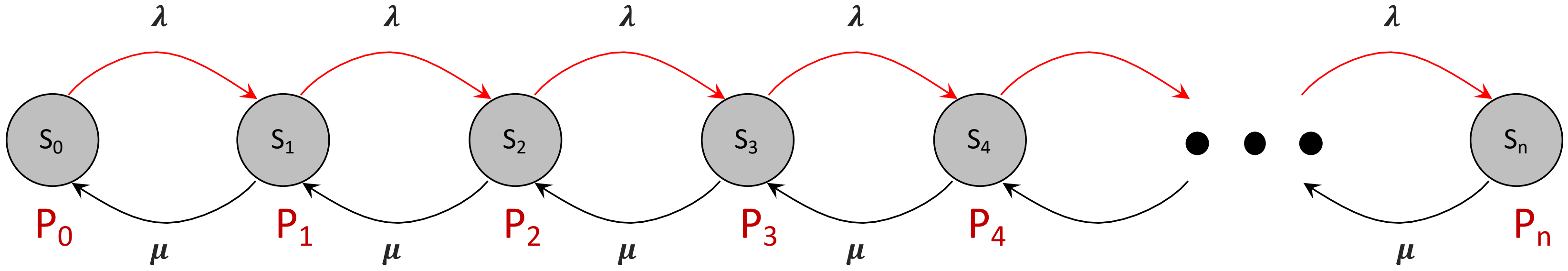
Server

Image is taken from: www.kindpng.com

service rate = $\mu = \frac{1}{M}$

Clients

arrival rate = $\lambda = \frac{1}{N}$



For S_0 ➡ $P_0 \times \lambda = P_1 \times \mu$ ➡ $P_1 = \dfrac{P_0 \times \lambda}{\mu}$

For S_1 ➡ $P_0 \times \lambda + P_2 \times \mu = P_1 \times \mu + P_1 \times \lambda$ ➡ $P_2 = P_0 \times (\dfrac{\lambda}{\mu})^2$

For S_2 ➡ $P_1 \times \lambda + P_3 \times \mu = P_2 \times \mu + P_2 \times \lambda$ ➡ $P_3 = P_0 \times (\dfrac{\lambda}{\mu})^3$

⋮

For S_n ➡ $P_n = P_0 \times (\dfrac{\lambda}{\mu})^n$

Productivity of a stochastic queue model

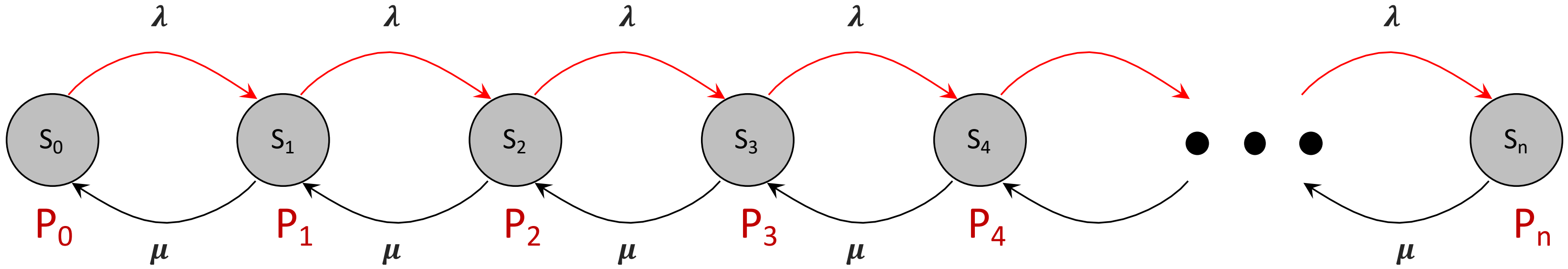
Server

Image is taken from: www.kindpng.com

service rate = $\mu = \frac{1}{M}$

Clients

arrival rate = $\lambda = \frac{1}{N}$



$$Since \sum_{i=0}^{n} P_i = 1 :$$

$$\Rightarrow P_0 + P_0 \times \left(\frac{\lambda}{\mu}\right) + P_0 \times \left(\frac{\lambda}{\mu}\right)^2 + \cdots + P_0 \times \left(\frac{\lambda}{\mu}\right)^n = 1$$

$$\Rightarrow P_0 + P_0 \times \left(\frac{\lambda}{\mu}\right) + P_0 \times \left(\frac{\lambda}{\mu}\right)^2 + \cdots + P_0 \times \left(\frac{\lambda}{\mu}\right)^n = 1$$

$$\Rightarrow P_0 \underbrace{\left(1 + \left(\frac{\lambda}{\mu}\right) + \left(\frac{\lambda}{\mu}\right)^2 + \cdots + \left(\frac{\lambda}{\mu}\right)^n\right)}_{\text{Finite Geometric Series}} = 1$$

$$\Rightarrow P_0 \left(\frac{1}{1-\frac{\lambda}{\mu}}\right) = 1 \Rightarrow P_0 = 1 - \frac{\lambda}{\mu}$$

Productivity of a stochastic queue model

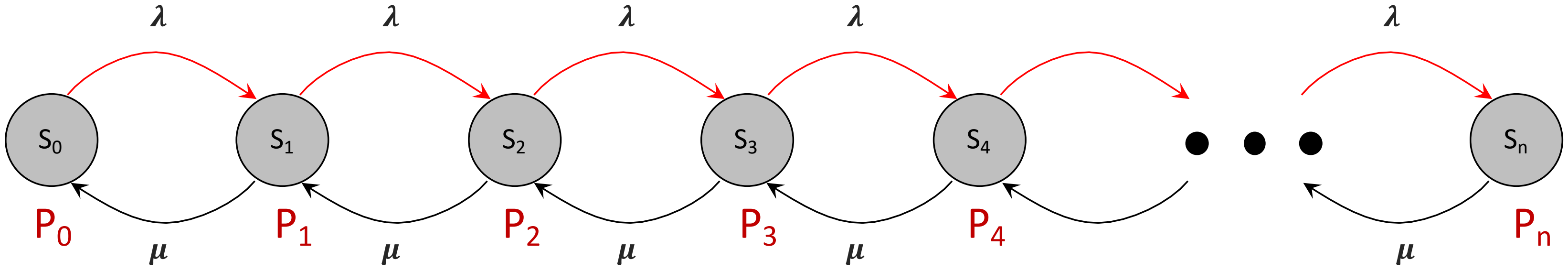
Server

Image is taken from: www.kindpng.com

service rate = $\mu = \dfrac{1}{M}$

Clients

arrival rate = $\lambda = \dfrac{1}{N}$



$\textit{Since}$ Productivity $= \boldsymbol{\mu} \times (\mathbf{1} - \mathbf{P_0})$ and $\boldsymbol{P_0} = \mathbf{1} - \dfrac{\boldsymbol{\lambda}}{\boldsymbol{\mu}}$

$$\textbf{Productivity} = \lambda$$

Productivity of a stochastic queue model

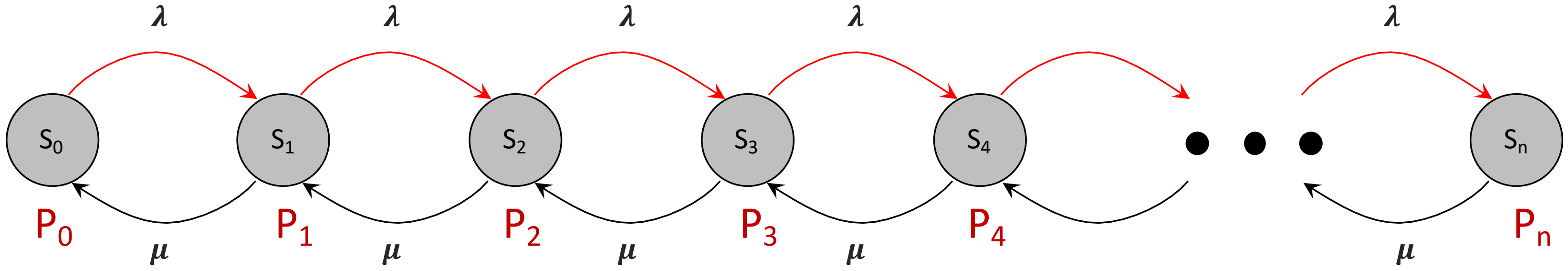
Server

Image is taken from: www.kindpng.com

service rate = $\mu = \frac{1}{M}$

Clients

arrival rate = $\lambda = \frac{1}{N}$



But, there is a twist here!
This system assumes an infinite arrival of clients!

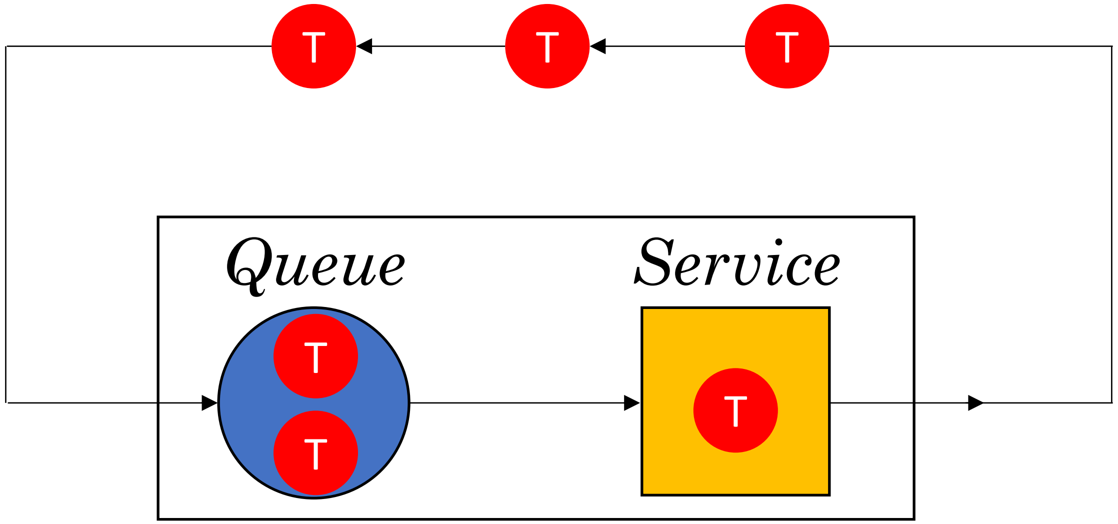
In the construction industry, we are mostly dealing with close-looped, finite systems, right?

Productivity of a stochastic finite queue model

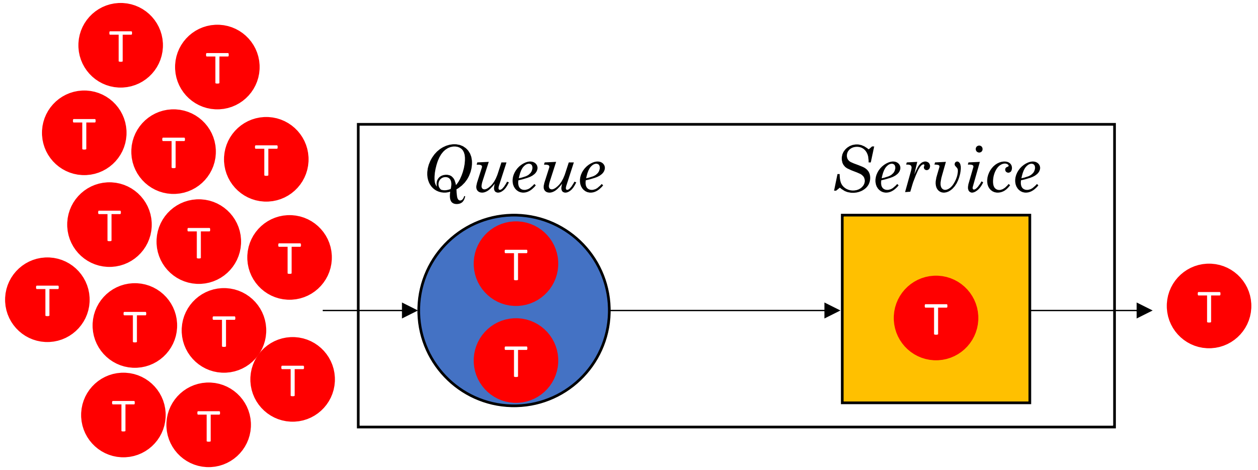
There are two systems:

- **Finite Systems:** Limited number of units arriving at the queue (truck-excavator)

- **Infinite Systems:** Large number of units arriving at the queue (Telephone calls)

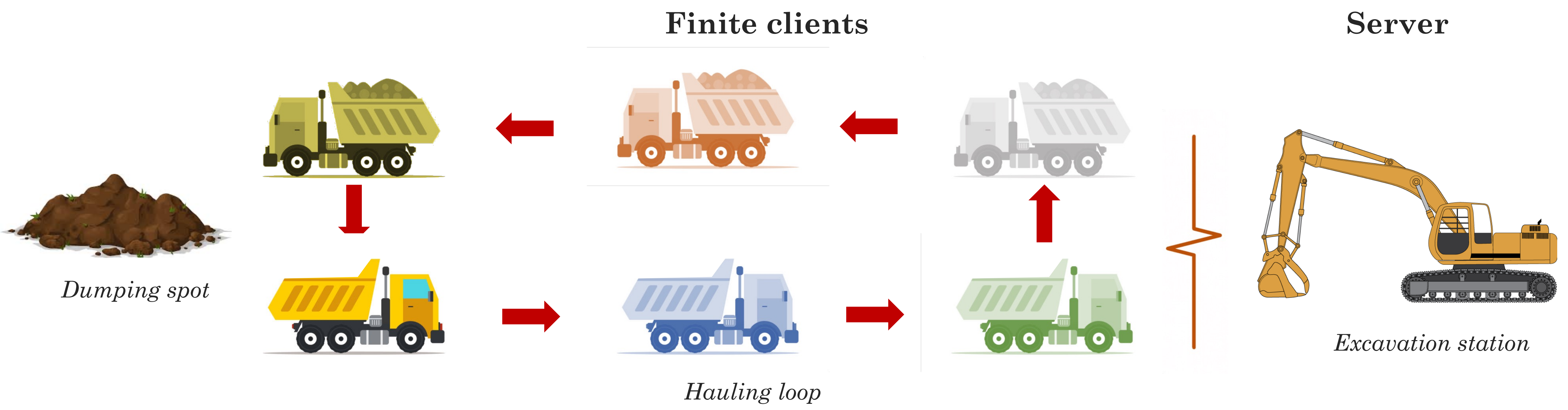


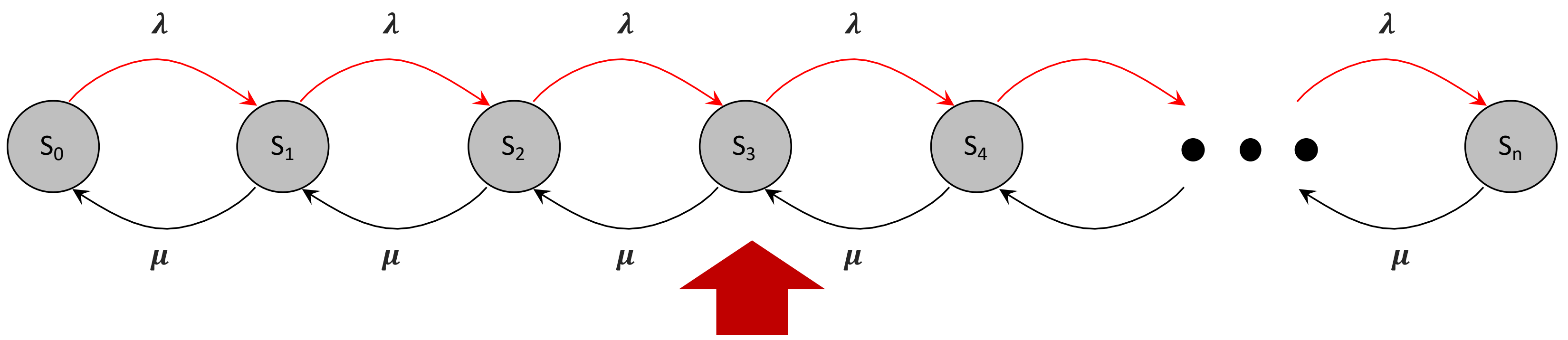
Finite Systems



Infinite Systems

Productivity of a stochastic finite queue model

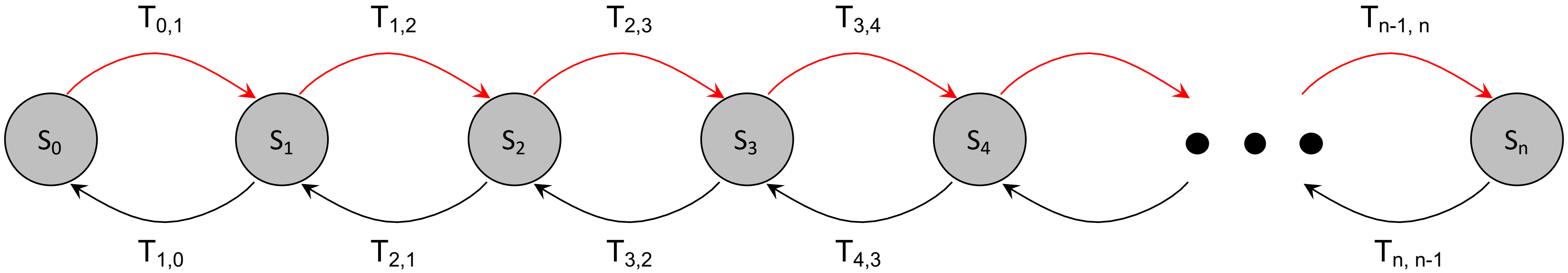




How do you think this model would change for the finite system ?

Productivity of a stochastic finite queue model

Let's go back to the generic model and see how it would change for a finite system



- Let's assume each truck takes t_b to move from A to B. This is called the Back Cycle.

- Also, let's assume it takes t_s for an excavator to fill a truck

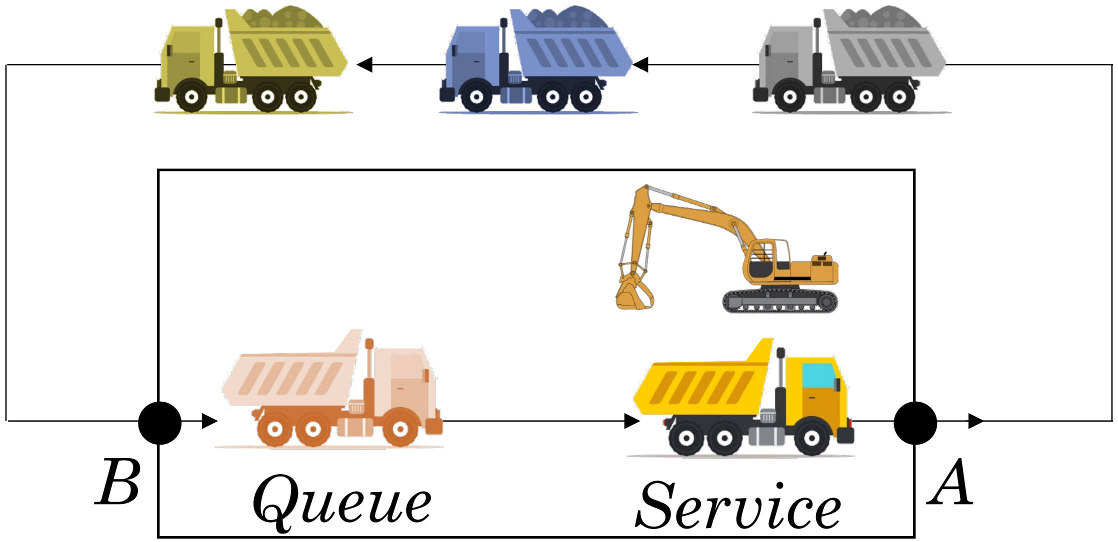
- Then:

Arrival rate $\boldsymbol{\lambda} = \frac{1}{t_b}$

Processing rate $\boldsymbol{\mu} = \frac{1}{t_s}$

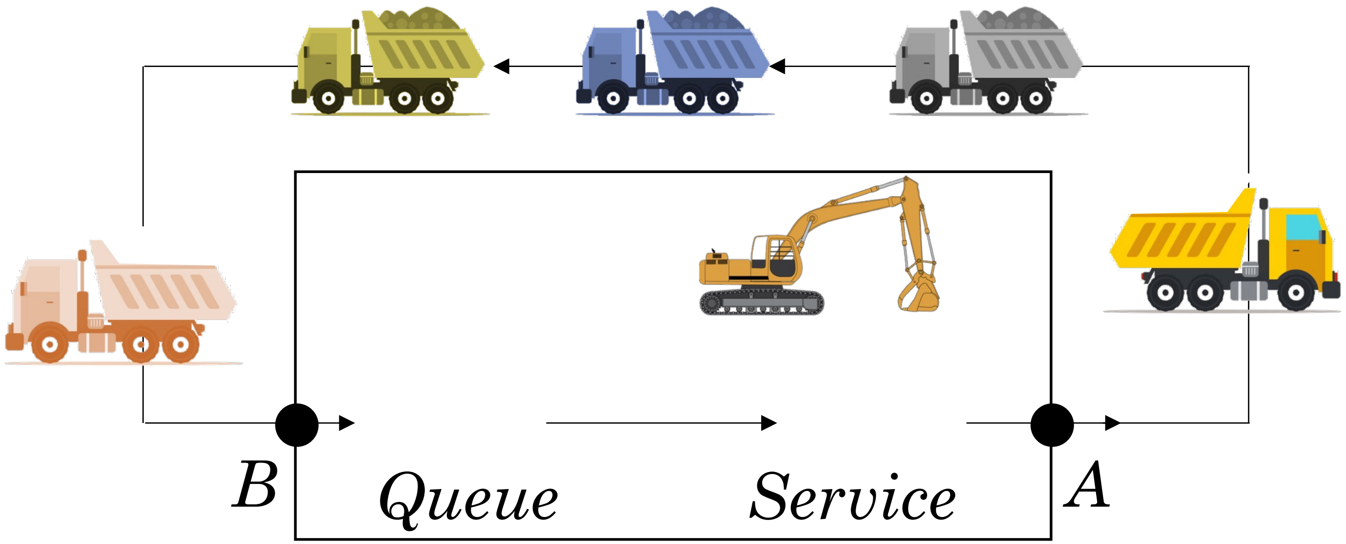
- **How many possible states are there?**

The No. of states = No. Clients + 1



Productivity of a stochastic finite queue model

- Now, if there are no trucks in the queue or service point, what is the probability of one truck arriving in the queue?

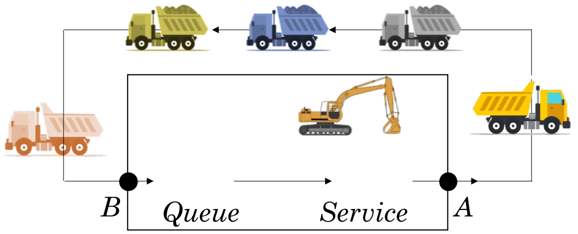


Because each truck has the arrival probability of λ, when 5 trucks are in the back cycle, the probability of one truck getting into the queue is 5λ

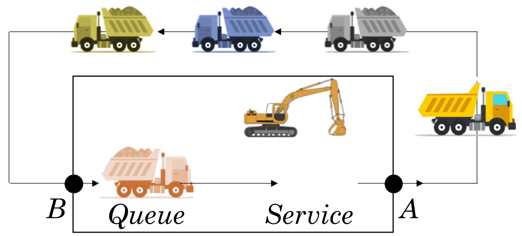
Important:

Again, the above statement holds only with the assumption that the arrival rate follows a Poisson distribution (arrivals are independent and random)

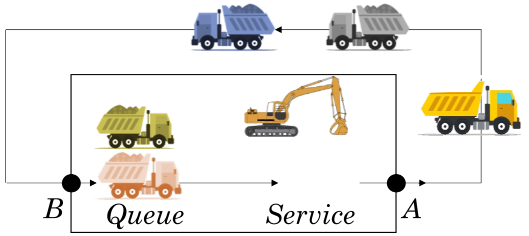
Productivity of a stochastic finite queue model



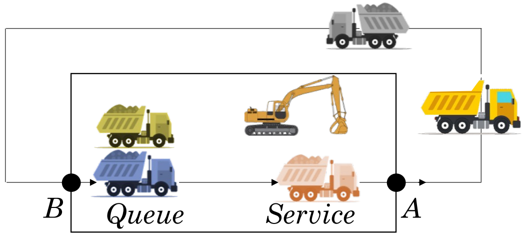
5 trucks in the back cycle → the arrival probability 5λ



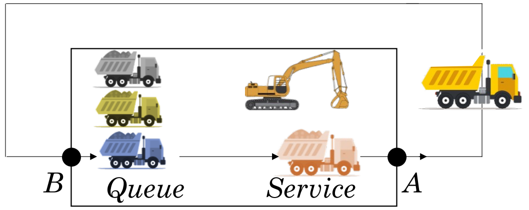
4 trucks in the back cycle → the arrival probability 4λ



3 trucks in the back cycle → the arrival probability 3λ

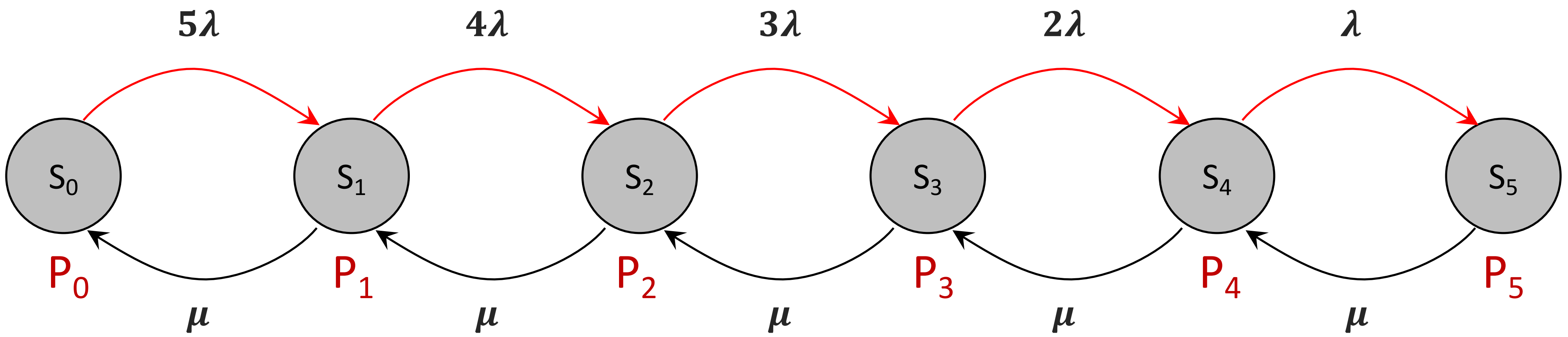


2 trucks in the back cycle → the arrival probability 2λ



1 truck in the back cycle → the arrival probability λ

Productivity of a stochastic finite queue model



For S_0 ➡ $P_0 \times 5\lambda = P_1 \times \mu$

For S_1 ➡ $P_0 \times 5\lambda + P_2 \times \mu = P_1 \times \mu + P_1 \times 4\lambda$

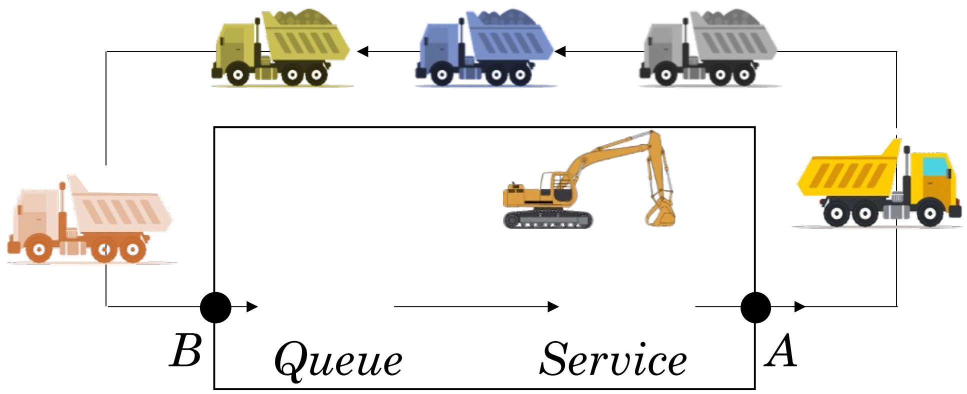
For S_2 ➡ $P_1 \times 4\lambda + P_3 \times \mu = P_2 \times \mu + P_2 \times 3\lambda$

For S_3 ➡ $P_2 \times 3\lambda + P_4 \times \mu = P_3 \times \mu + P_3 \times 2\lambda$

For S_4 ➡ $P_3 \times 2\lambda + P_5 \times \mu = P_4 \times \mu + P_4 \times \lambda$

For S_5 ➡ $P_4 \times \lambda = P_5 \times \mu$

Also, we know that : $\sum P_i = 1$



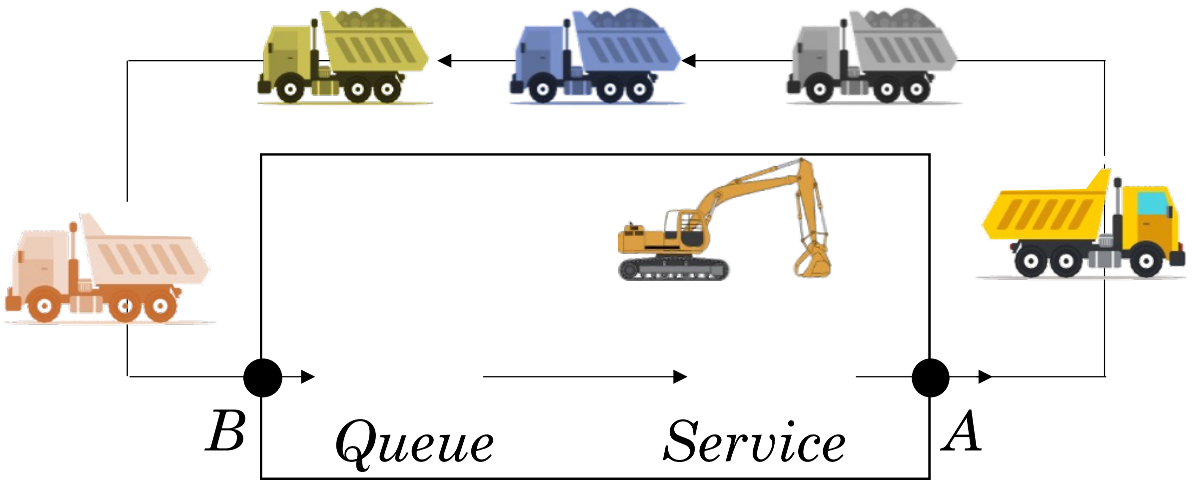
Through solving this system of equations, P_0 can be calculated

Productivity of a stochastic finite queue model

- Once P_0 is calculated, the below generic equation can be used to determine productivity for any given length of time:

$$P = L \times C \times \mu \times (1 - P_0)$$

Where:



L: the period of time considered
C: capacity of the unit
μ: the processor rate (e.g., excavator)
P_0: the probability of no units in the service area

Productivity of a stochastic finite queue model

- Additionally, using the following equations, we can easily calculate (1) the average number of units in the system [N], and (2) the average length of queue in the system [Q].

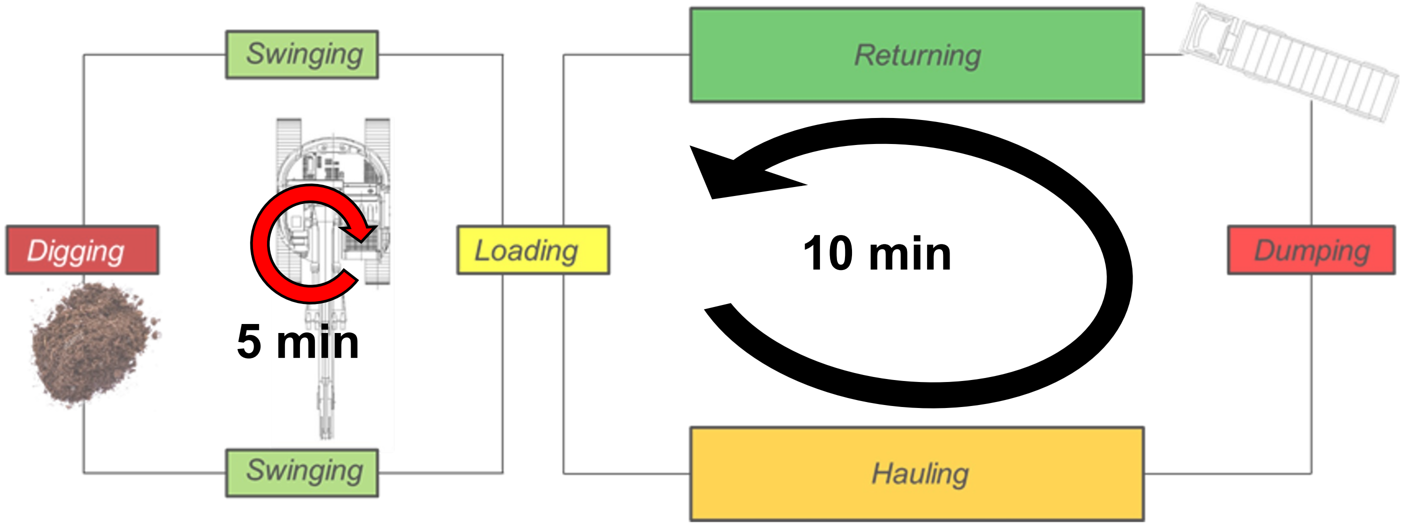
$$\mathrm{N} = \sum_{i=0}^{M} P_i \times X_i \quad \text{and} \quad Q = \sum_{i=0}^{M} P_i \times (X_i - 1)$$

Where:

N: the average number of units in the system
Q: the average length of queue
P_i: the probability of the system being in state i
M: number of units
X_i: number of units in the system at states i

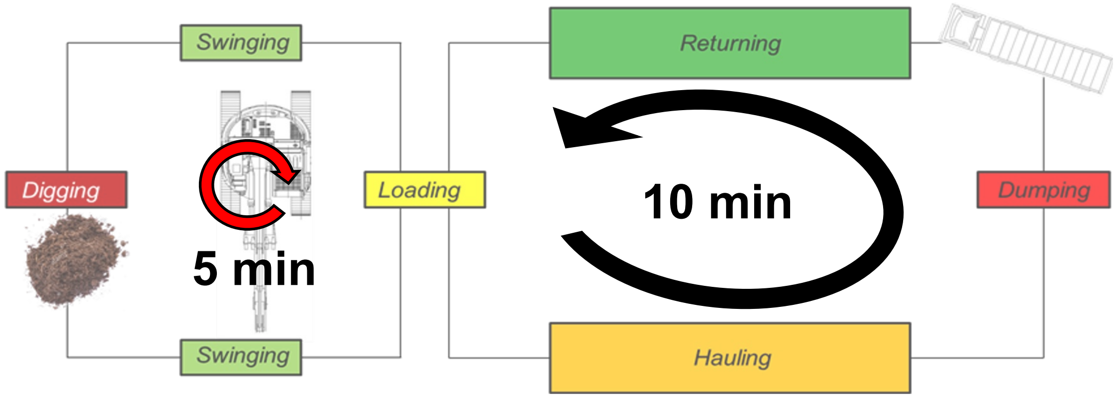
Example

In an excavation operation, one excavator and 6 trucks are moving the soil. It takes trucks 10 minutes to move from the loading point to the dumping point and return. The excavator takes 5 minutes to feel a truck. The capacity of each truck is 20 tons. (a) How much soil is removed after 1,5 if we assume a deterministic system?, (b) what if we assume stochasticity according to queuing theory?; and (c) What is the average length of the queue according to queuing theory?



Example

(a) Solution:



$$\boldsymbol{Truck\ cycle} = \lambda = \frac{\mathbf{1}}{\boldsymbol{T_b}} = \frac{\mathbf{1}}{(\frac{\mathbf{15}}{\mathbf{60}})} = \mathbf{5}\,\frac{\boldsymbol{Tr}}{\boldsymbol{Hr}}$$

For 6 trucks, productivity = 6× 5 Tr/Hr × 20 = 600 Ton/ Hr

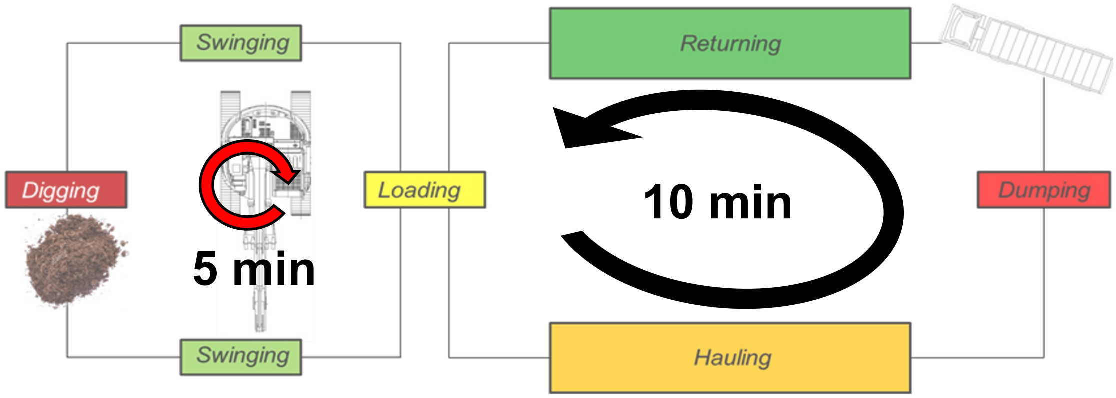
$$\boldsymbol{Excavator\ Cycle} = \mu = \frac{\mathbf{1}}{\boldsymbol{T_s}} = \frac{\mathbf{1}}{(\frac{\mathbf{5}}{\mathbf{60}})} = \mathbf{12}\,\boldsymbol{Tr/Hr}$$

For 1 excavator, productivity = 12 Tr/Hr × 20 = 240 T/ Hr

240 T/ Hr is dominant. So, in a deterministic system, this is the productivity of the system. After 1.5 hours, 360 tons of soil is removed

Example

(b) Solution:



$$arrival\ rate = \lambda = \frac{1}{T_b} = \frac{1}{(\frac{10}{60})} = 6\,\frac{Tr}{Hr}$$

$$Processing\ rate\ \mu = \frac{1}{T_s} = \frac{1}{(\frac{5}{60})} = 12\ Tr/Hr$$

$$\left\{\begin{array}{lll}
\mu\times P_1 = 6\lambda\times P_0 & \rightarrow & 12\times P_1 = 36\times P_0 \\
6\lambda\times P_0 + \mu\times P_2 = 5\lambda\times P_1 + \mu\times P_1 & \rightarrow & 36\times P_0 + 12\times P_2 = 30\times P_1 + 12\times P_1 \\
5\lambda\times P_1 + \mu\times P_3 = 4\lambda\times P_2 + \mu\times P_2 & \rightarrow & 30\times P_1 + 12\times P_3 = 24\times P_2 + 12\times P_2 \\
4\lambda\times P_2 + \mu\times P_4 = 3\lambda\times P_3 + \mu\times P_3 & \rightarrow & 24\times P_2 + 12\times P_4 = 18\times P_3 + 12\times P_3 \\
3\lambda\times P_3 + \mu\times P_5 = 2\lambda\times P_4 + \mu\times P_4 & \rightarrow & 18\times P_3 + 12\times P_5 = 12\times P_4 + 12\times P_4 \\
2\lambda\times P_4 + \mu\times P_6 = \lambda\times P_5 + \mu\times P_5 & \rightarrow & 12\times P_4 + 12\times P_6 = 6\times P_5 + 12\times P_5 \\
\lambda\times P_5 = \mu\times P_6 & \rightarrow & 6\times P_5 = 12\times P_6 \\
\sum_{i=0}^{6} P_i & \rightarrow & P_0 + P_1 + P_2 + P_3 + P_4 + P_5 + P_6 = 1
\end{array}\right\}$$

$$P_0 = 0,0121$$
$$P_1 = 0,0363$$
$$P_2 = 0,0906$$
$$P_3 = 0,1813$$
$$P_4 = 0,2719$$
$$P_5 = 0,2719$$
$$P_6 = 0,1359$$

Example

(b) Solution (cont.):

$P = C \times \mu \times (1 - P_0)$ ➡ $P = 20 \times 12 \times (1 - 0{,}0121) =$ **237,096 ton/hr**

$P = L \times C \times \mu \times (1 - P_0)$ ➡ $P = 1{,}5 \times 237{,}096 =$ **355,644 ton in 1,5 hours**

(c) Solution

$$Q = \sum_{i=0}^{M} P_i \times (X_i - 1)$$

X_i	P_i	$P_i \times (X_i - 1)$
1	0,0363	0
2	0,0906	0,0906
3	0,1813	0,3626
4	0,2719	0,8157
5	0,2719	1,0876
6	0,1359	0,6795
		Q = 3,036 trucks

Generic equations for stochastic finite queue model

The general form of the equation for this type of queuing system looks lite:

$$P_0 = \left[\sum_{i=0}^{M}\left(\frac{M!}{(M-i)!} \times \left(\frac{\lambda}{\mu}\right)^{i}\right)\right]^{-1}$$

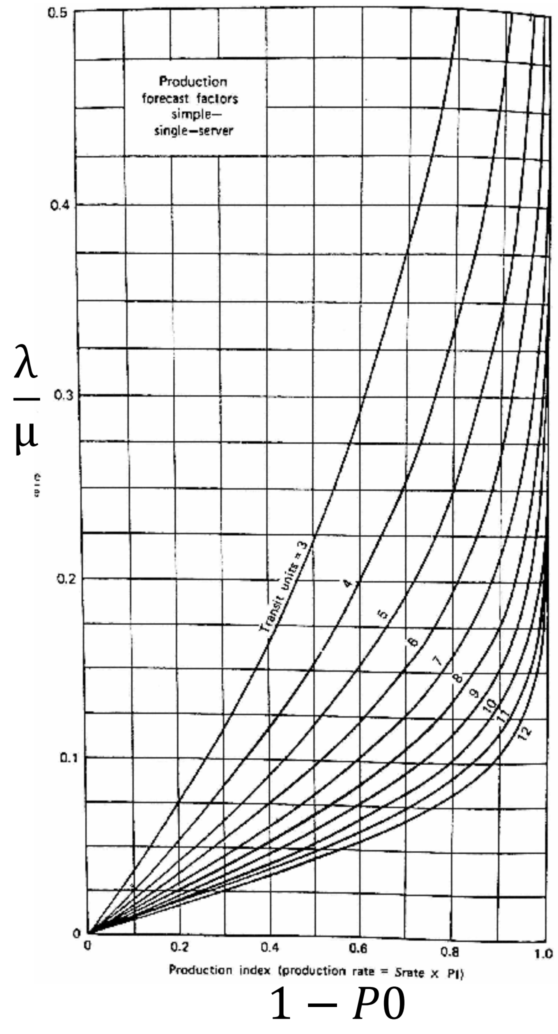
$$P_i = \left[\frac{M!}{(M-i)!} \times \left(\frac{\lambda}{\mu}\right)^{i}\right] \times P_0$$

Try these equations for the previous example!

Generic equations for stochastic finite queue model

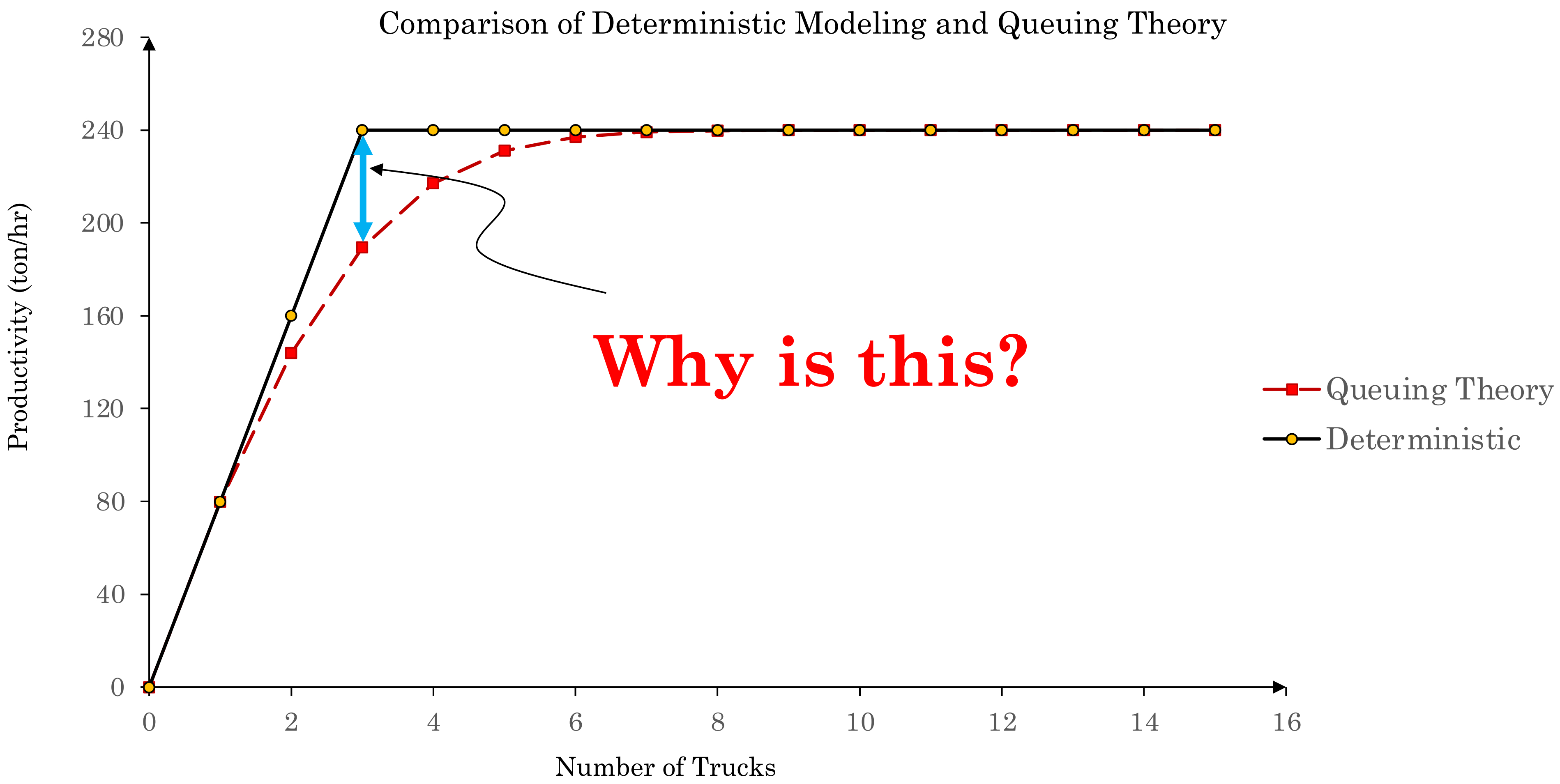
The general form of the equation for this type of queuing system looks lite:

Another way to solve the queuing problem is to use available nomographs.



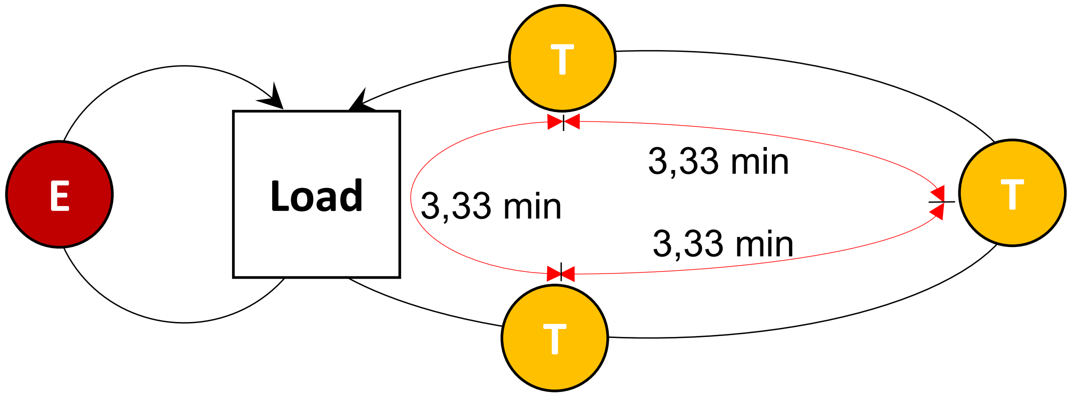
Queuing Theory Vs. Deterministic Model

Let's compare the difference between Queuing Theory and deterministic modeling

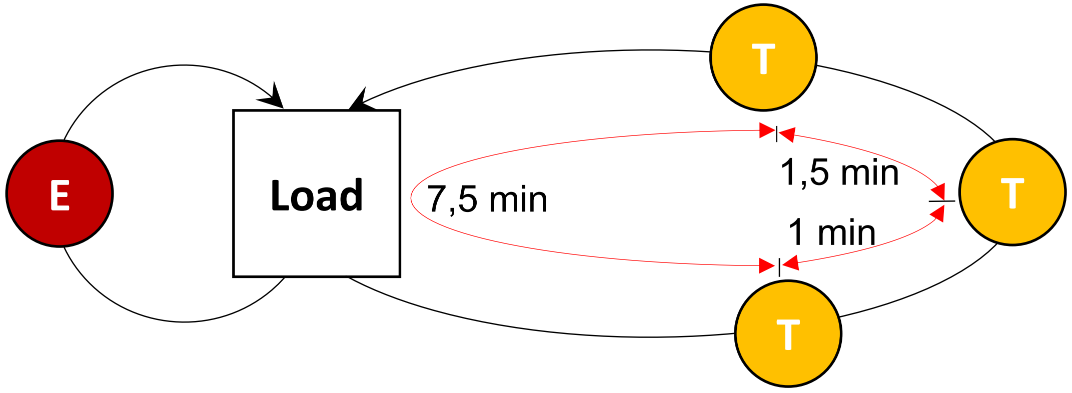


Queuing Theory Vs. Deterministic Model

- In a deterministic system, the assumption is that the distances between trucks are always the same. In a Queuing model, this does not hold. As a result of randomness, **"Bunching"** is always inevitable, which, in turn, causes more idleness.



Deterministic Model



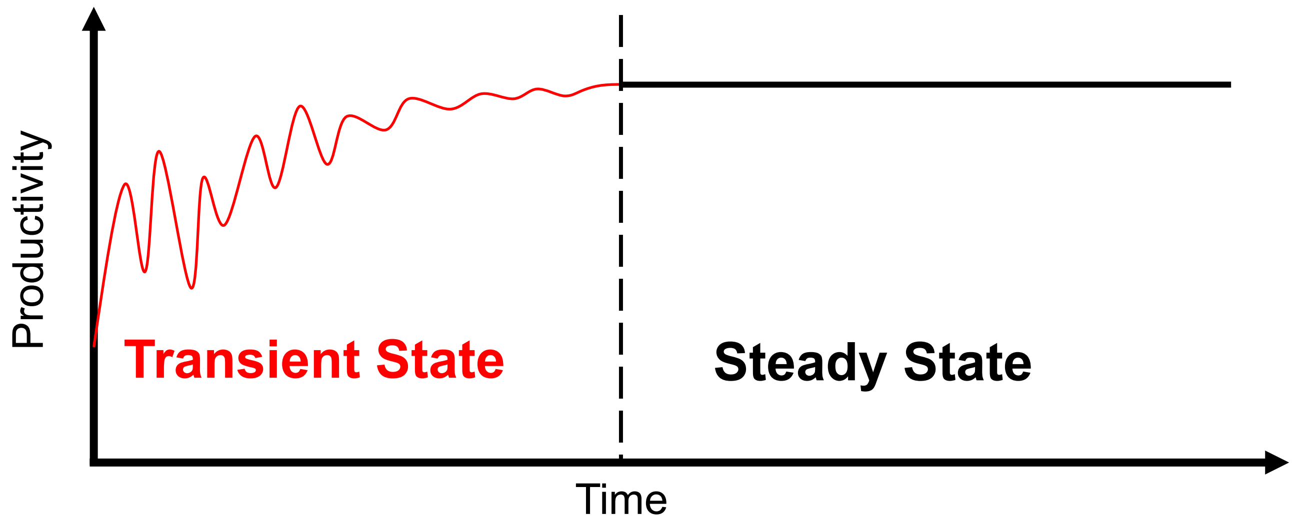
Queuing Model

The main limitations of Queuing Theory

- As a mathematical model, queuing theories **function on several assumptions** (e.g., specific distributions, rational behavior, complete randomness, time-invariant behavior, etc.). While helpful, this can have an impact on the accurate modeling of complex systems/processes.

- Again, given the mathematical basis, it **demands a certain level of affinity with statistics**. This may reduce its practical applicability. Especially, because experts may need to convert a complex problem into a simple queue representation.

- Finally, it **assumes that the system is in a steady state**. This ignores the time it takes for the system to reach equilibrium after changes are applied to the system. More about this in the next lecture!



Burning Questions

- How would the queue model change if we have **multiple serves**?
- **Is a queueing model stochastic**? Can we apply Monte Carlo to it?
- Do you think **Queuing theory is suitable for the construction industry**? Why?

What did we learn in this lecture?

- Importance of queues
- Queuing theory
- Markov chain
- Limitations of queuing theory