

Centre Universitaire de AbdElhafid Boussouf, Mila
2^{ème} Année Master Intelligence artificielle et ses applications
Année universitaire : 2022/2023
Matière: **Modélisation et simulation**

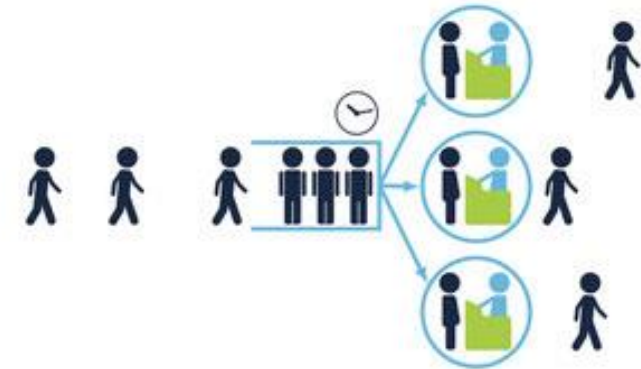
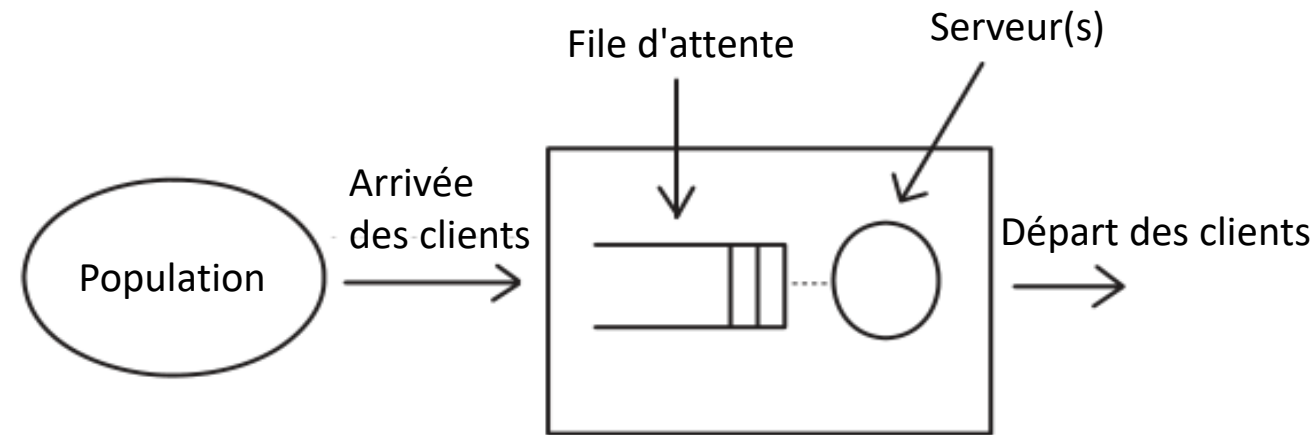
CHAPITRE III:

LES FILES D'ATTENTE

Responsable de la matière: DR. SADEK BENHAMMADA

1. Introduction

- Les files d'attente sont des systèmes où des clients se présentent à un dispositif de service, appelé serveur.



- Un client occupe le serveur pendant un certain temps, les autres clients doivent attendre avant d'être servis, formant ainsi une file d'attente.

1. Introduction

- **Exemples de files d'attentes et des objectifs de modélisations**

Systèmes	Clients	Serveur(s)	Objectif
Centres d'appels / services clients	Appels téléphoniques entrants	Agents / Opérateurs	Réduire le temps d'attente moyen et optimiser le nombre d'opérateurs
Réseaux informatiques / Internet	Paquets de données	Routeurs, serveurs	Minimiser la latence et éviter la congestion du réseau
Télécommunications	Appels téléphoniques	Commutateurs / Canaux radio	Minimiser le temps d'attente et le blocage
Systèmes bancaires	Clients physiques	Guichets ou conseillers bancaires	Diminuer le temps d'attente des clients et améliorer le service
Supermarchés / Caisses	Clients en file	Caissiers ou bornes automatiques	Réduire la longueur et la durée d'attente en caisse
Transports (aéroports, péages,)	Passagers, véhicules	Comptoirs d'enregistrement, guichets	Fluidifier le trafic et réduire les files d'attente
Systèmes hospitaliers / urgences	Patients	Médecins, infirmiers, salles de soins	Optimiser l'affectation du personnel et réduire le temps d'attente des patients
Usines / production industrielle	Produits ou pièces à traiter	Machines de production	Maximiser le débit et minimiser les temps d'arrêt
Cloud computing / serveurs web	Requêtes utilisateurs	Serveurs virtuels	Équilibrer la charge et minimiser le temps de réponse
Transport public (bus, métro)	Passagers	Véhicules (bus, trains)	Ajuster la fréquence pour réduire l'attente moyenne

1. Introduction

- La théorie des files d'attente (Queueing theory) consiste en la modélisation et l'étude des files d'attente
- Un modèle de file d'attente est construit de sorte à donner des réponses à des questions sur des caractéristiques des file d'attente telles que :
 - La longueur moyenne de la file d'attente,
 - Le temps d'attente d'un client,
 - Le taux d'utilisation du serveur,
 - ...
- Les modèles de files d'attente fournissent des outils pour évaluer les performances des systèmes de files d'attente, et la prise de décisions concernant les ressources nécessaires pour fournir un service.

- La notation de Kendall est le système standard utilisé pour décrire les modèles de file d'attente:

A/B/C/K/D

- **A : La loi d'arrivée des clients. Il peut être:**

- M (memoryless): Processus de Poisson (ou processus d'arrivée aléatoire) (c'est-à-dire, temps exponentiels entre les arrivées).
- G (général) si le processus d'arrivée est général.
- D (déterministe) si le processus d'arrivée est déterministe et déterministe constante.
- etc.

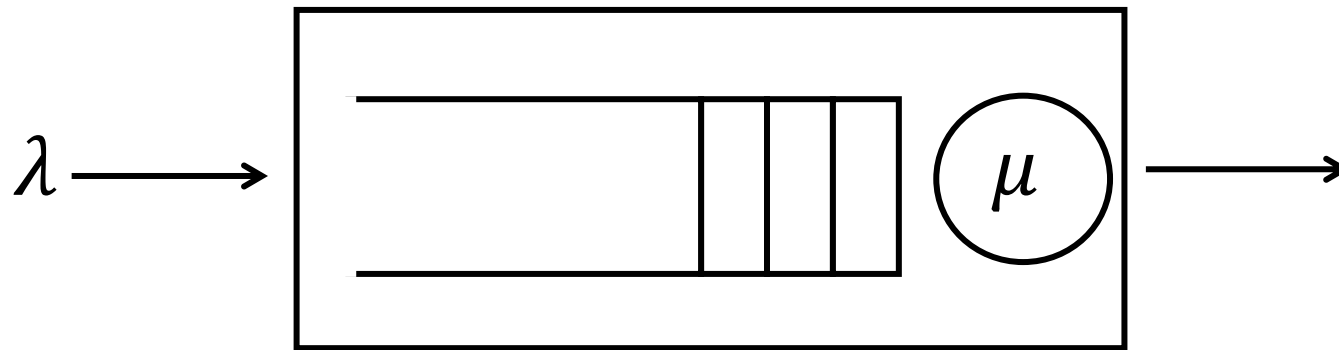
- **B : La loi de la durée de service. Il peut être:** M, G, D, etc.

- **C : Nombre de serveurs travaillant en parallèle (1 ou plus)**

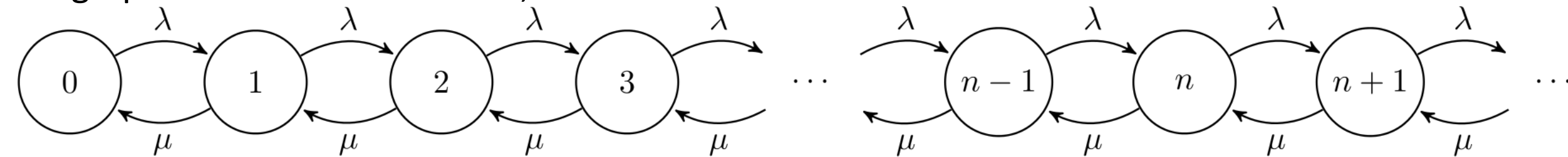
- **K : Capacité du système, c.à.d, le nombre maximum de places (clients) autorisés dans le système (file + serveurs). Lorsque le nombre est à ce maximum, d'autres arrivées sont refusées. Par défaut, la capacité est supposée illimitée.**

- **D : discipline de service. Par défaut FIFO/FCFP (First In First Out/First Come First Served).** Mais aussi: SIRO: Service In Random Order, PQ: Priority Queuing, etc.

- Le modèle M/M/1 est l'équivalent de $M/M/1/\infty /FIFO$:
 - Le processus des arrivées suit une loi de poisson de paramètre λ . ($\frac{1}{\lambda}$ est le temps moyen pour qu'une nouvelle arrivée se produise)
 - La durée de service suit une loi exponentielle de paramètre μ . ($\frac{1}{\mu}$ est la durée moyenne de service)
 - La capacité de la file est illimitée,
 - Un seul serveur.



- Une file d'attente M/M/1 est un processus de naissance-mort à espace d'états $E = \{0, 1, 2, 3, \dots\}$, où les valeurs correspondent au nombre de clients dans le système,
- Le processus naissance-mort modélisant un système M/M/1 peut être représenté par le graphe de transition suivant, donc:



Donc:

$$\begin{aligned} \lambda_i &= \lambda, & i &= 0, 2, 3, \dots \\ \mu_i &= \mu, & i &= 1, 2, 3, \dots \end{aligned}$$

- La matrice de taux de transition :

$$Q = \begin{pmatrix} -\lambda & \lambda & & & \\ \mu & -(\mu + \lambda) & \lambda & & \\ & \mu & -(\mu + \lambda) & \lambda & \\ & & \mu & -(\mu + \lambda) & \lambda \\ & & & \ddots & \ddots \end{pmatrix}$$

Arrivée avant un départ et départ avant une arrivée

- Temps pour qu'une nouvelle arrivée se produise :

$$A \sim \text{Exp}(\lambda)$$

- Temps pour qu'une nouvelle départ se produise :

$$D \sim \text{Exp}(\mu)$$

- Probabilité qu'une arrivée se produise avant un départ :

$$P(A < D) = \frac{\lambda}{\lambda + \mu}$$

- Probabilité qu'un départ se produise avant une arrivée :

$$P(D < A) = \frac{\mu}{\lambda + \mu}$$

2. Modélisation des files d'attente

2.2. Modèle M/M/1

Les probabilités stationnaires:

- Un processus M/M/1 est un processus de naissance-mort, donc:

La processus de Markov est ergodique si : $\lambda < \mu$

$$P(N=0) = P_0 = \left(1 + \sum_{n=1}^{\infty} \prod_{i=1}^n \frac{\lambda}{\mu}\right)^{-1} = \left(1 + \sum_{n=1}^{\infty} \left(\frac{\lambda}{\mu}\right)^n\right)^{-1}$$

N: Le nombre de clients au temps t dans le système

Par application de la somme géométrique: $\sum_{n=1}^{\infty} ar^n = \frac{a}{1-r}$ (ssi $r < 1$).

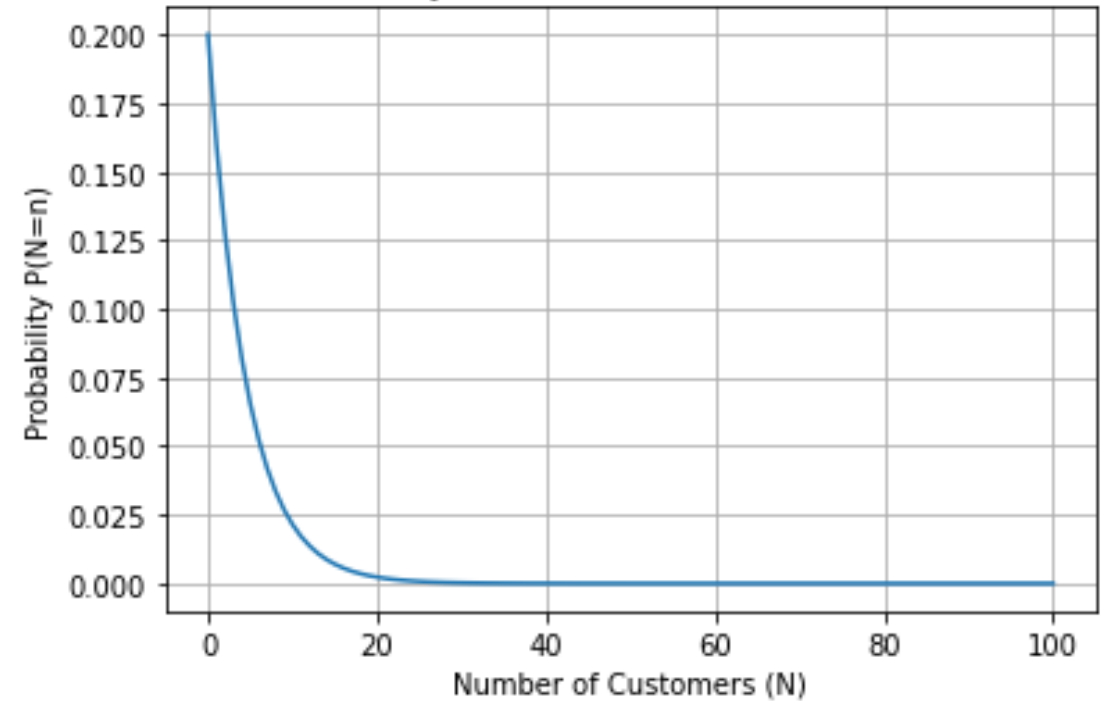
$\frac{\lambda}{\mu} < 1$, donc:

$$P_0 = 1 - \frac{\lambda}{\mu}$$

$$P(N=n) = P_n = P_0 \prod_{i=1}^n \frac{\lambda}{\mu} = P_0 \left(\frac{\lambda}{\mu}\right)^n, \text{ pour } n = 0, 1, 2, 3, \dots$$

$\rho = \frac{\lambda}{\mu}$ est appelé le coefficient d'utilisation du système ou l'intensité du trafic.

Probability Distribution $P(N=n)$ for $\rho=0.8$



$$P(N=0) = P_0 = 1 - \rho$$

$$P(N=n) = P_n = (1 - \rho)\rho^n$$

- $\bar{N}_Q$: Le nombre moyen de clients dans la file d'attente
- $\bar{N}_S$: Le nombre moyen de clients en train d'être servis.
- $\bar{N} = E(N) = \bar{N}_Q + \bar{N}_S$: Le nombre moyen de clients dans le système (attente + service)
- $\bar{T}_Q$: Le temps moyen d'attente dans la file.
- $\bar{T}_S$: Le temps moyen de service.
- $\bar{T} = \bar{T}_Q + \bar{T}_S$: temps moyen qu'un client passe dans le système (attente + service). (temps moyen de séjour d'un client dans le système)

Caractéristiques du système M/M/1

$\bar{N}$: Le nombre moyen de clients dans le système (attente + service)

$$\begin{aligned}\bar{N} &= E(N) = \sum_{i=1}^{\infty} i \times P(N = i) = \sum_{i=1}^{\infty} iP_i \\ &= \sum_{i=1}^{\infty} i(1 - \rho)\rho^i = (1 - \rho) \sum_{i=1}^{\infty} i\rho^i = (1 - \rho) \frac{\rho}{(1 - \rho)^2} = \frac{\rho}{(1 - \rho)}\end{aligned}$$

$$\bar{N} = \frac{\rho}{(1 - \rho)}$$

• **$\bar{N}_S$: Le nombre moyen de clients en train d'être servis: $\bar{N}_S = \frac{\lambda}{\mu} = \rho = 1 - P_0$**

$$\bar{N}_S = \rho = 1 - P_0$$

• **$\bar{N}_Q$: Le nombre moyen de clients dans la file d'attente :**

Soit N_Q la variable aléatoire qui donne le nombre de clients se trouvant dans la file d'attente

$$\begin{aligned}\bar{N}_Q &= E(N_Q) = \sum_{i=1}^{\infty} (i - 1)P_i = \sum_{i=1}^{\infty} iP_i - \sum_{i=1}^{\infty} P_i = \bar{N} - (1 - P_0) = \bar{N} - \rho = \bar{N} - \bar{N}_S \\ &= \frac{\rho}{(1 - \rho)} - \rho = \frac{\rho^2}{(1 - \rho)}\end{aligned}$$

$$\bar{N}_Q = \frac{\rho^2}{(1 - \rho)}$$

Caractéristiques du système M/M/1

- $\bar{T}$: Temps moyen de séjour dans le système (attente + service)

Loi de Little:

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

Donc:

$$\bar{T} = \frac{\bar{N}}{\lambda} = \frac{1}{\mu - \lambda}$$

$\bar{T}_S$: Le temps moyen de service: $\frac{1}{\mu}$

$$\bar{T}_S = \frac{1}{\mu}$$

$\bar{T}_Q$: Le temps moyen d'attente dans la file:

$$\bar{T}_Q = \bar{T} - \bar{T}_S = \frac{1}{\mu - \lambda} - \frac{1}{\mu} = \frac{\lambda}{\mu(\mu - \lambda)}$$

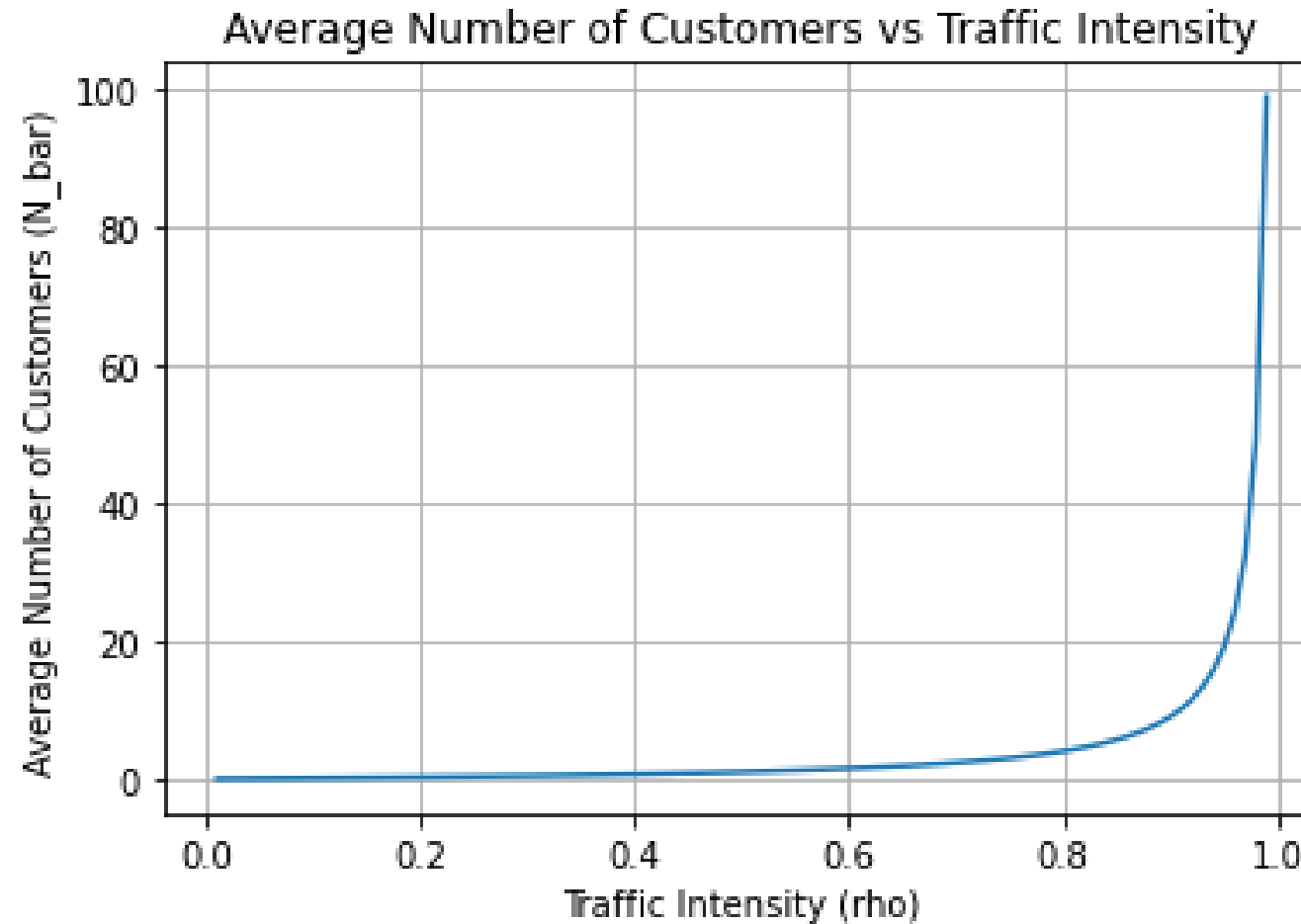
$$\bar{T}_Q = \frac{\lambda}{\mu(\mu - \lambda)}$$

2. Modélisation des files d'attente

2.2. Modèle M/M/1

Caractéristiques du système M/M/1

	M/M/1
Condition de stationnarité	$\rho = \frac{\lambda}{\mu} < 1$
P_0 : Probabilité que le système est vide (Aucun client dans le système)	$P_0 = 1 - \rho$
P_w : Probabilité d'attente (la probabilité qu'un client doive faire la queue (au lieu d'être immédiatement servi))	$P_w = \rho$
$\bar{N}$: Nombre moyen de clients dans le système	$\bar{N} = \frac{\rho}{1 - \rho} = \frac{\lambda}{\mu - \lambda} = \bar{N}_q + \bar{N}_s$
$\bar{N}_Q$: Nombre moyen de clients dans la file d'attente	$\bar{N}_Q = \frac{\rho^2}{1 - \rho} = \frac{\lambda^2}{\mu(\mu - \lambda)} = \bar{N} - \bar{N}_s$
$\bar{N}_S$: Nombre moyen de clients en service (au guichet)	$\bar{N}_S = \rho = 1 - P_0 = \bar{N} - \bar{N}_Q$
$\bar{T}$: Temps moyen qu'un client passe dans le système (attente + service). (temps moyen de séjour d'un client dans le système)	$\bar{T} = \frac{\bar{N}}{\lambda} = \frac{\rho}{\lambda(1 - \rho)} = \frac{1}{\mu - \lambda}$
$\bar{T}_Q$: Temps moyen d'attente dans la file	$\bar{T}_Q = \frac{\rho}{\mu(1 - \rho)} = \frac{\lambda}{\mu(\mu - \lambda)} = \bar{T} - \bar{T}_s$
$\bar{T}_S$: Temps moyen de service	$\bar{T}_S = \frac{1}{\mu}$



Caractéristiques du système M/M/1

Exemple

On considère une file d'attente M/M/1 de taux $\lambda = 3/2$ et $\mu = 2$. Calculer :

1. Le nombre moyen de clients dans le système, $\bar{N}$.
2. Le nombre moyen de clients en service, $\bar{N}_S$.
3. Le nombre moyen de clients dans la file d'attente, $\bar{N}_Q$.
4. Le temps moyen de séjour d'un client dans le système, $\bar{T}$.
5. Le temps moyen d'attente d'un client dans la file, $\bar{T}_Q$.
6. Le temps moyen de service d'un client, $\bar{T}_S$.

Solution:

$$\bar{N}_S = \rho = \frac{\lambda}{\mu} = \frac{3}{4}$$

$$\bar{N} = \frac{\rho}{(1 - \rho)} = \frac{\frac{3}{4}}{\left(1 - \frac{3}{4}\right)} = 3$$

$$\bar{N}_Q = \bar{N} - \bar{N}_S = 3 - \frac{3}{4} = \frac{9}{4}$$

$$\bar{N}_Q = \frac{\rho^2}{(1 - \rho)} = \frac{\left(\frac{3}{4}\right)^2}{\left(1 - \frac{3}{4}\right)} = \frac{9}{4}$$

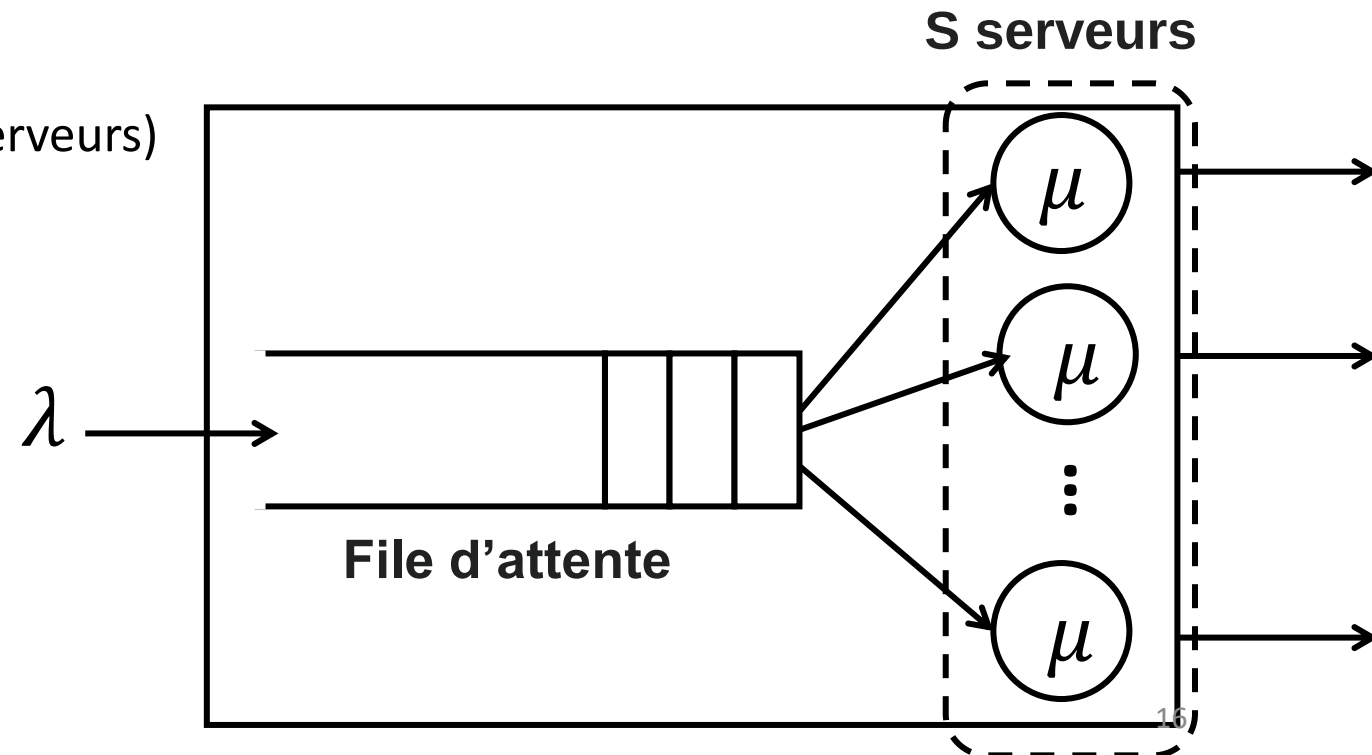
$$\bar{T} = \frac{\bar{N}}{\lambda} = \frac{3}{3/2} = 2$$

$$\bar{T}_S = \frac{1}{\mu} = \frac{1}{2}$$

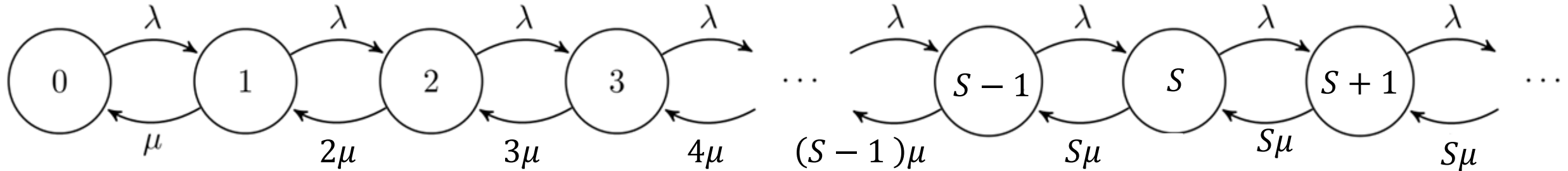
$$\bar{T}_Q = \frac{\lambda}{\mu(\mu - \lambda)} = \bar{T} - \bar{T}_S = \frac{3}{2}$$

Caractéristiques du système M/M/1

- Si dans un file d'attente M/M/1 le taux d'arrivée moyen λ est supérieur au taux de service μ le système ne sera pas stable, c'est pourquoi il faut augmenter le nombre de serveurs.
- **Le modèle M/M/S est l'équivalent de $M/M/S/\infty$ /FIFO:**
 - Le processus des arrivées suit une loi de poisson de paramètre λ .
 - La durée de service suit une loi exponentielle de paramètre μ .
 - la capacité de la file est illimitée,
 - Le service est fourni par **S** serveurs (plusieurs serveurs)



- Une file d'attente M/M/S est un processus naissance-mort
- Le processus naissance-mort modélisant un système M/M/S peut être représenté par le graphe de transitions suivant:



Donc:

$$\lambda_k = \lambda, \quad k = 0, 2, 3, \dots$$

$$\mu_k = \min(k, S)\mu = \begin{cases} k\mu & \text{pour } k < S \\ S\mu & \text{pour } k \geq S \end{cases} \quad k = 1, 2, 3, \dots$$

Les probabilités stationnaires:

Le processus est ergodique et admet une distribution stationnaire si: $\frac{\lambda}{S\mu} < 1$, (ou bien $\lambda < S\mu$)

Supposons que cette condition est vérifiée, donc, les probabilités stationnaires du système:

$$P(N = 0) = P_0 = \left(\sum_{n=0}^{S-1} \frac{\rho^n}{n!} + \left(\frac{\rho^S}{S!} \frac{1}{1-a} \right) \right)^{-1}$$

$$P(N = n) = P_n = \begin{cases} P_0 \frac{\rho^n}{n!} & \text{pour } n < S \\ P_0 \frac{a^n S^S}{S!} & \text{pour } n \geq S \end{cases}$$

où $a = \frac{\lambda}{S\mu} = \frac{\rho}{S}$ (a est le taux d'utilisation du système)

N : Le nombre de clients au temps t dans le système

$P(N = n) = P_n$: La probabilité d'avoir n clients dans le système

Formule Erlang C

La probabilité qu'un client qui arrive doit attendre est:

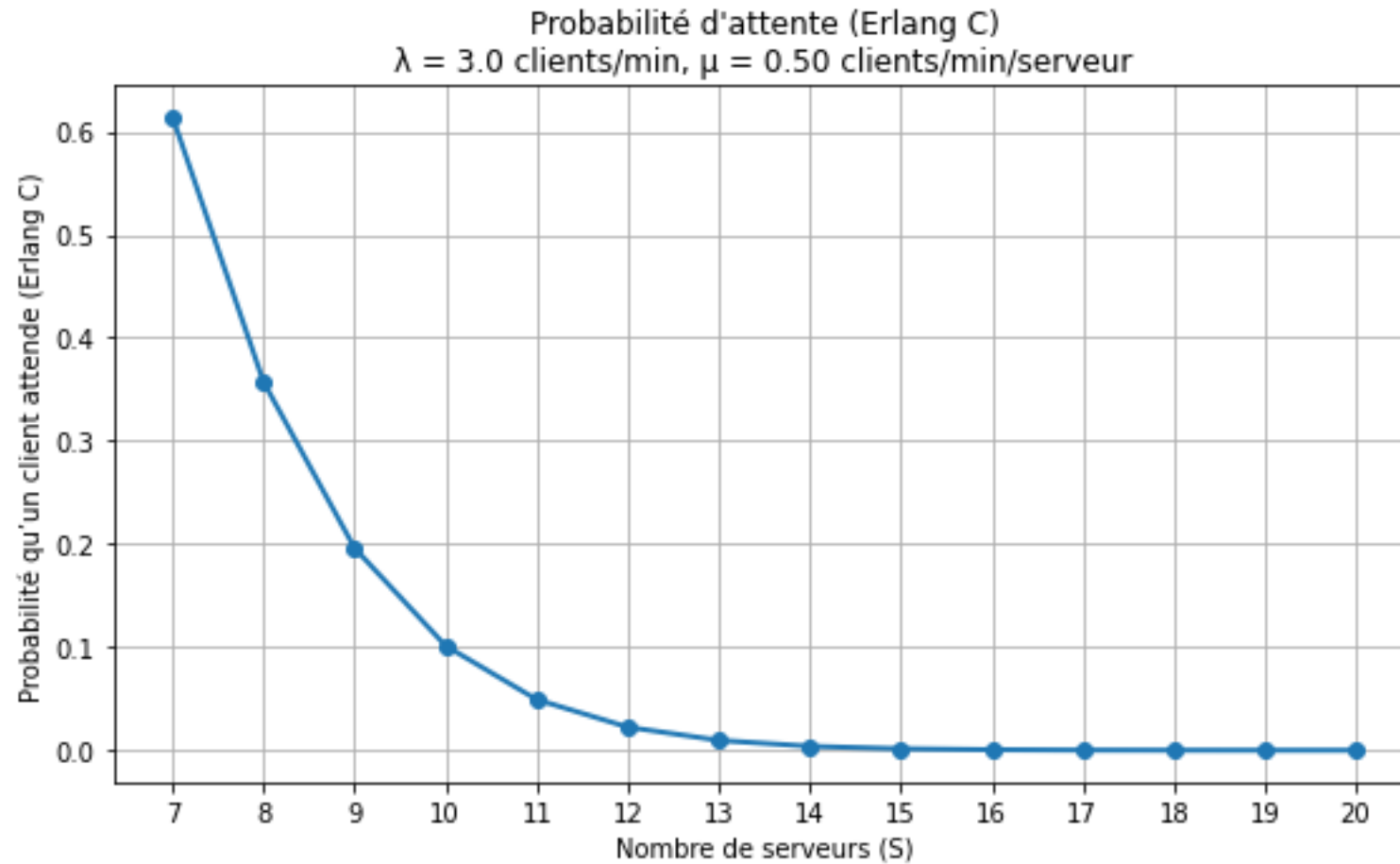
$$\begin{aligned} P_{\text{waiting}} &= P(n \geq S) = \sum_{n=S}^{\infty} P_n \\ &= 1 - P(n < S) = \sum_{n=0}^{S-1} P_n = \frac{\frac{S}{S-\rho}}{\sum_{n=0}^{S-1} \frac{\rho^n}{n!} + \frac{\rho^S}{S!} \frac{1}{1-\rho}} = P_0 \cdot \frac{\rho^S}{S!} \frac{S}{S-\rho} \end{aligned}$$

$$P_{\text{waiting}} = C(S, \rho) = P_0 \frac{\rho^S}{(S-1)!(S-\rho)}$$

C(S, ρ) est appelée **La formule Erlang C**, et exprime la probabilité qu'un client arrivant doit faire la queue (au lieu d'être immédiatement servi).

Formule Erlang C

Exemple:



Caractéristiques du système M/M/S

$\bar{N}$: Le nombre moyen de clients dans le système (attente + service)

$$\bar{N} = E(N) = \sum_{k=1}^{\infty} k \times P(N = k) = \sum_{k=1}^{\infty} k P_k$$

$$\bar{N} = \rho + \frac{\rho}{S - \rho} C(S, \rho) = \rho \left(1 + \frac{C(S, \rho)}{(S - \rho)} \right)$$

$\bar{N}_Q$: Le nombre moyen de clients dans la file d'attente

Soit N_Q la variable aléatoire qui donne le nombre de clients se trouvant dans la file d'attente

$$\bar{N}_Q = E(N_Q) = \sum_{k=S+1}^{\infty} (k - S) \times P_k$$

$$\bar{N}_Q = \rho \frac{C(S, \rho)}{S - \rho}$$

$\bar{N}_S$: Le nombre moyen de clients en train d'être servis: $\bar{N}_S = \bar{N} - \bar{N}_Q = \rho$

$$\bar{N}_S = \rho = \frac{\lambda}{\mu}$$

Caractéristiques du système M/M/S

- **$\bar{T}$: Temps moyen d'attente dans le système** (attente + service)

Loi de Little: $\bar{N} = \lambda \bar{T}$

Donc:

$$\bar{T} = \frac{\bar{N}}{\lambda} = \frac{1}{\lambda} \rho \left(1 + \frac{C(S, \rho)}{(S - \rho)} \right) = \frac{1}{\lambda} \times \frac{\lambda}{\mu} \left(1 + \frac{C(S, \rho)}{(S - \rho)} \right)$$

$$\bar{T} = \frac{1}{\mu} \left(1 + \frac{C(S, \rho)}{(S - \rho)} \right)$$

$\bar{T}_S$: Le temps moyen de service: $\frac{1}{\mu}$

$$\bar{T}_S = \frac{1}{\mu}$$

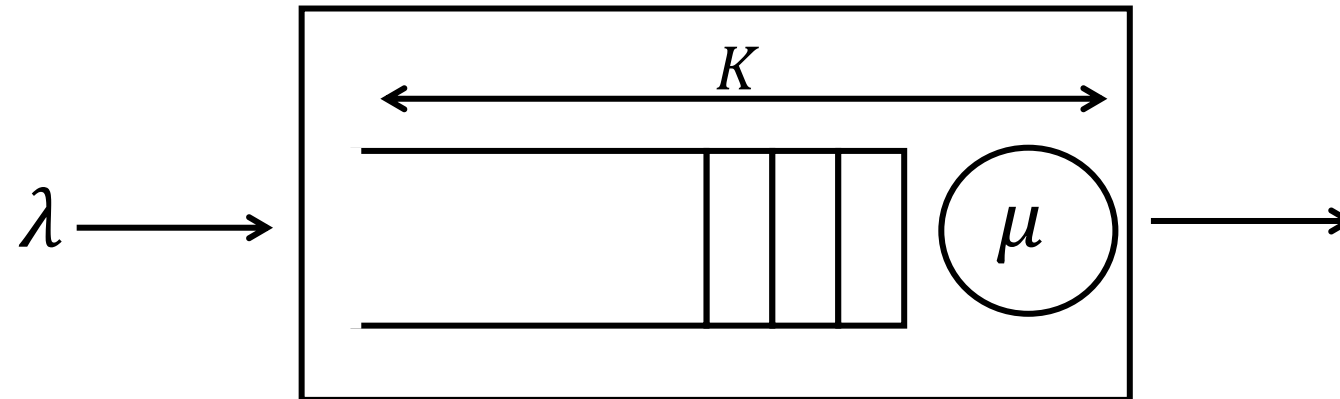
$\bar{T}_Q$: Le temps moyen d'attente dans la file:

$$\bar{T}_Q = \bar{T} - \bar{T}_S = \frac{1}{\mu} \left(1 + \frac{C(S, \rho)}{(S - \rho)} \right) - \frac{1}{\mu}$$

$$\bar{T}_Q = \frac{C(S, \rho)}{\mu(S - \rho)}$$

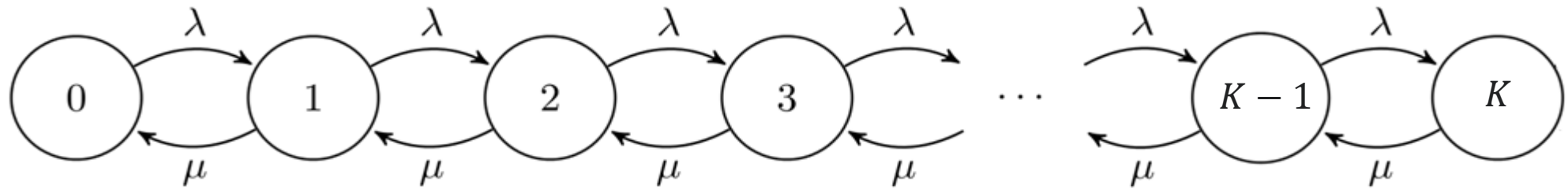
	M/M/1	M/M/S
Condition de stationnarité	$\rho = \frac{\lambda}{\mu} < 1$	$a = \frac{\lambda}{S\mu} < 1$ (a : le taux d'utilisation du système)
P_0 : Probabilité que le système est vide (Aucun client dans le système)	$P_0 = 1 - \rho$	$P_0 = \left(\sum_{k=0}^{S-1} \frac{\rho^k}{k!} + \left(\frac{\rho^S}{S!} \frac{1}{1-a} \right) \right)^{-1}$
P_w : Probabilité d'attente (la probabilité qu'un client doive faire la queue (au lieu d'être immédiatement servi))	$P_w = \rho$	$C(S, \rho) = P_0 \times \frac{\rho^S}{(S-1)!(S-\rho)}$
$\bar{N}$: Nombre moyen de clients dans le système	$\bar{N} = \frac{\rho}{1-\rho}$	$\bar{N} = \rho \left(1 + \frac{C(S, \rho)}{S-\rho} \right)$
$\bar{N}_Q$: Nombre moyen de clients dans la file d'attente	$\bar{N}_Q = \frac{\rho^2}{1-\rho}$	$\bar{N}_Q = \rho \frac{C(S, \rho)}{S-\rho}$
$\bar{N}_S$: Nombre moyen de clients en service (au guichet)	$\bar{N}_S = \rho$	$\bar{N}_S = \rho$
$\bar{T}$: Temps moyen qu'un client passe dans le système (attente + service). (temps moyen de séjour d'un client dans le système)	$\bar{T} = \frac{\rho}{\lambda(1-\rho)}$	$\bar{T} = \frac{1}{\mu} \left(1 + \frac{C(S, \rho)}{S-\rho} \right)$
$\bar{T}_Q$: Temps moyen d'attente dans la file	$\bar{T}_Q = \frac{\rho}{\mu(1-\rho)}$	$\bar{T}_Q = \frac{C(S, \rho)}{\mu(S-\rho)}$
$\bar{T}_S$: Temps moyen de service	$\bar{T}_S = \frac{1}{\mu}$	$\bar{T}_S = \frac{1}{\mu}$

- Le système **M/M/1/K** est identique au système M/M/1 exceptée que sa capacité est finie (le nombre maximum de clients).
- Le processus des arrivées suit une loi de poisson de paramètre λ .
- La durée de service suit une loi exponentielle de paramètre μ .
- La capacité de du système est limitée à nombre maximum de clients **K** (y compris celui en service).
- Le service est assuré par **1 seul serveur**



Probabilités stationnaires M/M/1/K

- Le système **M/M/1/K** est un processus de naissance-mort à espace d'états fini $E = \{0, 1, \dots, K - 1, K\}$, avec les taux de transitions $\lambda_i = \lambda, i = 0, \dots, K - 1$ et $\mu_i = \mu, i = 1, \dots, K$.



- Le processus est fini et irréductible, donc, il est ergodique.
- Les probabilités stationnaires du système ($\rho = \frac{\lambda}{\mu}$):

$$P_0 = \begin{cases} \frac{1}{K+1} & \rho = 1 \\ \frac{1-\rho}{1-\rho^{K+1}} & \rho \neq 1 \end{cases}$$

Remarques.

- Si $\rho = 1$ ($\lambda = \mu$), alors: $P_n = \frac{1}{K+1}$, $n = 0, 1, \dots, K$

$$P_n = \begin{cases} P_0 \cdot \rho^n & n \leq K \\ 0 & \text{sinon} \end{cases}$$

- La probabilité de rejet d'un client est la probabilité que le nombre maximum de clients K est atteint:

$$P_K = \frac{(1-\rho)\rho^K}{1-\rho^{K+1}}$$

$\bar{N}$: Le nombre moyen de clients dans le système (attente + service)

$$\bar{N} = E(N) = \sum_{n=1}^K nP_n$$

$$\bar{N} = \frac{\rho(1 - (K+1)\rho^K + K\rho^{K+1})}{(1-\rho)(1-\rho^{K+1})}$$

$\bar{N}_S$: Nombre moyen de clients en service (au guichet)

$$\bar{N}_S = 1 - P_0$$

$\bar{N}_Q$: Le nombre moyen de clients dans la file d'attente

Soit N_Q la variable aléatoire qui donne le nombre de clients se trouvant dans la file d'attente

$$\bar{N}_Q = E(N_Q) = \sum_{n=1}^K (n-1)P_n = \bar{N} - \bar{N}_S$$

Caractéristiques du système M/M/1/K

$\bar{T}$: Temps moyen d'attente dans le système (attente + service)

Loi de Little:

$$\bar{N} = \lambda_e \bar{T}$$

λ_e : taux d'entrée dans le système

Dans le système M/M/1/K, le taux d'entrée dans le système (λ_e) et le taux d'arrivée des clients (λ) sont différents: Un client qui arrive peut entrer dans le système ou peut être perdu parce que le système est plein (le nombre maximum de clients K est atteint).

$$\lambda_e = \lambda \times P(N < K) = \lambda \times (1 - P(N = K)) = \lambda(1 - P_K)$$

Donc:

$$\bar{T} = \frac{\bar{N}}{\lambda(1 - P_K)}$$

$\bar{T}_S$: Le temps moyen de service:

$$\bar{T}_S = \frac{1}{\mu}$$

$\bar{T}_Q$: Le temps moyen d'attente dans la file:

$$\bar{T}_Q = \bar{T} - \bar{T}_S = \frac{\bar{N}}{\lambda(1 - P_K)} - \frac{1}{\mu}$$

$$\bar{T}_Q = \frac{\bar{N}}{\lambda(1 - P_K)} - \frac{1}{\mu}$$

	M/M/1	M/M/1/K
Condition de stationnarité	$\frac{\lambda}{\mu} < 1$	Le système est stationnaire
P_0: Probabilité que le système est vide (Aucun client dans le système)	$P_0 = 1 - \rho$	$P_0 = \begin{cases} \frac{1}{K+1} & \rho = 1 \\ \frac{1-\rho}{1-\rho^{K+1}} & \rho \neq 1 \end{cases}$
P_K: Probabilité de rejet d'un client	/	$P_K = \frac{(1-\rho)\rho^K}{1-\rho^{K+1}}$

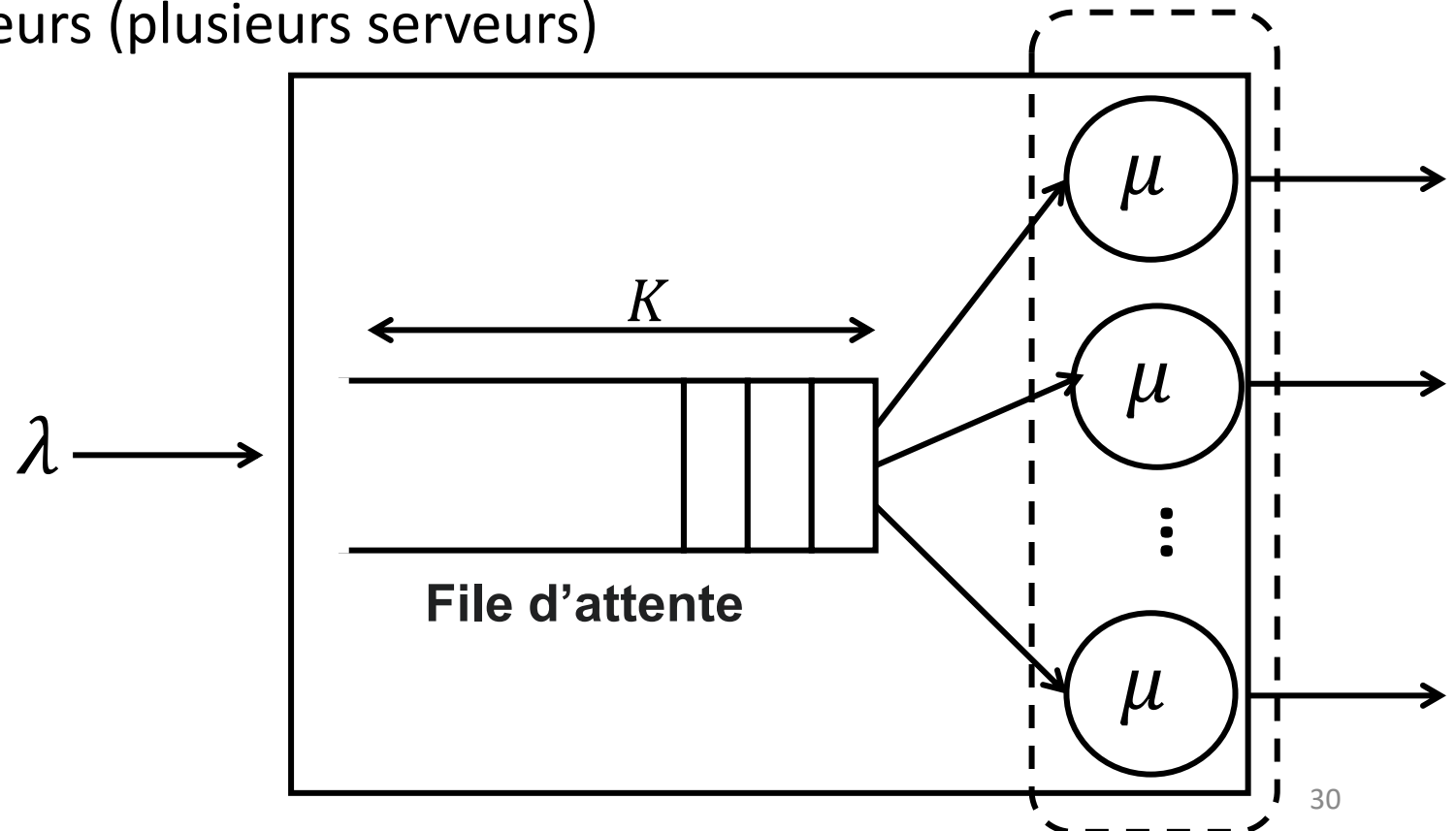
2. Modélisation des files d'attente

2.4. Modèle M/M/1/K

Caractéristiques du système M/M/1/K

	M/M/1	M/M/1/K
$\bar{N}$: Nombre moyen de clients dans le système	$\bar{N} = \frac{\rho}{1 - \rho}$	$\bar{N} = \frac{\rho(1 - (K + 1)\rho^K + K\rho^{K+1})}{(1 - \rho)(1 - \rho^{K+1})}$
$\bar{N}_Q$: Nombre moyen de clients dans la file d'attente	$\bar{N}_Q = \frac{\rho^2}{1 - \rho}$	$\bar{N}_Q = \sum_{n=1}^K (n - 1)P_n = \bar{N} - \bar{N}_S$
$\bar{N}_S$: Nombre moyen de clients en service (au guichet)	$\bar{N}_S = \rho$	$\bar{N}_S = 1 - P_0$
$\bar{T}$: Temps moyen qu'un client passe dans le système (attente + service). (temps moyen de séjour d'un client dans le système)	$\bar{T} = \frac{\bar{N}}{\lambda}$	$\bar{T} = \frac{\bar{N}}{\lambda(1 - P_K)}$
$\bar{T}_Q$: Temps moyen d'attente dans la file	$\bar{T}_Q = \frac{\bar{N}_Q}{\lambda}$	$\bar{T}_Q = \frac{\bar{N}_Q}{\lambda(1 - P_K)} = \frac{\bar{N}}{\lambda(1 - P_K)} - \frac{1}{\mu}$
$\bar{T}_S$: Temps moyen de service	$\bar{T}_S = \frac{1}{\mu}$	$\bar{T}_S = \frac{1}{\mu}$

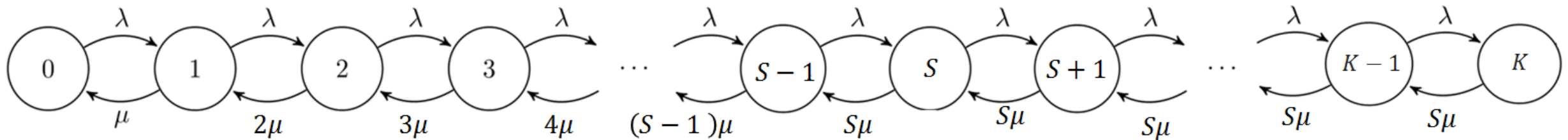
- Le modèle **M/M/S** est l'équivalent de **M/M/S/K /FIFO**:
 - Le processus des arrivées suit une loi de poisson de paramètre λ .
 - La durée de service suit une loi exponentielle de paramètre μ .
 - La capacité du système est limitée à un nombre maximum de clients **K**. *S serveurs*
 - Le service est fourni par **S** serveurs (plusieurs serveurs)



- Le système **M/M/S/K** est un processus de naissance-mort à espace d'états fini $E = \{0, 1, \dots, K - 1, K\}$, avec les taux de transitions:

$$\lambda_n = \lambda, n = 0, \dots, K - 1$$

$$\mu_n = \begin{cases} n\mu & 1 \leq n < S \\ S\mu & S \leq n \leq K \end{cases}$$



- Le processus est fini et irréductible, donc, il est ergodique.

- Les probabilités stationnaires du système **M/M/S/K**:

- Soit $\rho = \frac{\lambda}{\mu}$ et $a = \frac{\rho}{S}$

- P_K : Probabilité de rejet d'un client

$$P_0 = \left(\sum_{n=0}^{S-1} \frac{\lambda^n}{n! \mu^n} + \sum_{n=S}^K \frac{\lambda^n}{S^{n-S} S! \mu^n} \right)^{-1}$$

$$P_n = \begin{cases} \frac{\rho^n}{n!} P_0 & 1 \leq n \leq S \\ \frac{\rho^n}{S^{n-S} S!} P_0 & S \leq n \leq K \end{cases}$$

La probabilité qu'un client qui arrive doit attendre est:

$$P_{\text{waiting}} = P(S \leq n \leq K) = \sum_{n=S}^K P_n = P_0 \frac{\rho^S}{S!} \frac{1 - \left(\frac{\rho}{S}\right)^{K-S}}{1 - \frac{\rho}{S}}$$

$$P_K = \frac{\rho^K}{S^{K-S} S!} P_0$$

- Le taux d'entrée dans le système $\lambda_e = \lambda(1 - P_K)$

$$\bar{N} = \sum_{n=0}^K n P_n$$

$$\bar{N}_Q = \sum_{n=S+1}^K (n - S) P_n$$

$$\bar{N}_S = \rho(1 - P_K)$$

$$\bar{T} = \frac{\bar{N}}{\lambda_e} = \frac{\bar{N}}{\lambda(1 - P_K)}$$

$$\bar{T}_Q = \frac{\bar{N}_Q}{\lambda_e} = \frac{\bar{N}_Q}{\lambda(1 - P_K)}$$

	M/M/S/K	M/M/S
Condition de stationnarité	Le système est irréductible à espace d'états fini, donc, il est stationnaire.	$a = \frac{\lambda}{S\mu} < 1$ (a: le taux d'utilisation du système)
P_0 : Probabilité que le système est vide (Aucun client dans le système)	$P_0 = \left(\sum_{n=0}^{S-1} \frac{\lambda^n}{n! \mu^n} + \sum_{n=S}^K \frac{\lambda^n}{S^{n-S} S! \mu^n} \right)^{-1}$	$P_0 = \left(\sum_{k=0}^{S-1} \frac{\rho^k}{k!} + \left(\frac{\rho^S}{S!} \frac{1}{1 - \rho/S} \right) \right)^{-1}$
P_w : Probabilité d'attente (la probabilité qu'un client doive faire la queue (au lieu d'être immédiatement servi)	$P_w = P_0 \frac{\rho^S}{S!} \frac{1 - \left(\frac{\rho}{S}\right)^{K-S}}{1 - \frac{\rho}{S}}$	$C(S, \rho) = P_0 \times \frac{\rho^S}{(S-1)! (S - \rho)}$
$\bar{N}$: Nombre moyen de clients dans le système	$\bar{N} = \sum_{n=0}^K n P_n$	$\bar{N} = \rho \left(1 + \frac{C(S, \rho)}{S - \rho} \right)$
$\bar{N}_Q$: Nombre moyen de clients dans la file d'attente	$\bar{N}_Q = \sum_{n=S+1}^K (n - S) P_n$	$\bar{N}_Q = \rho \frac{C(S, \rho)}{S - \rho}$
$\bar{N}_S$: Nombre moyen de clients en service (au guichet)	$\bar{N}_S = \rho(1 - P_K)$	$\bar{N}_S = \rho$
$\bar{T}$: Temps moyen qu'un client passe dans le système (attente + service). (temps moyen de séjour d'un client dans le système)	$\bar{T} = \frac{\bar{N}}{\lambda_e} = \frac{\bar{N}}{\lambda(1 - P_K)}$	$\bar{T} = \frac{1}{\mu} \left(1 + \frac{C(S, \rho)}{(S - \rho)} \right)$
$\bar{T}_Q$: Temps moyen d'attente dans la file	$\bar{T}_Q = \frac{\bar{N}_Q}{\lambda_e} = \frac{\bar{N}_Q}{\lambda(1 - P_K)}$	$\bar{T}_Q = \frac{C(S, \rho)}{\mu(S - \rho)}$
$\bar{T}_S$: Temps moyen de service	$\bar{T}_S = \frac{1}{\mu}$	$\bar{T}_S = \frac{1}{\mu}$