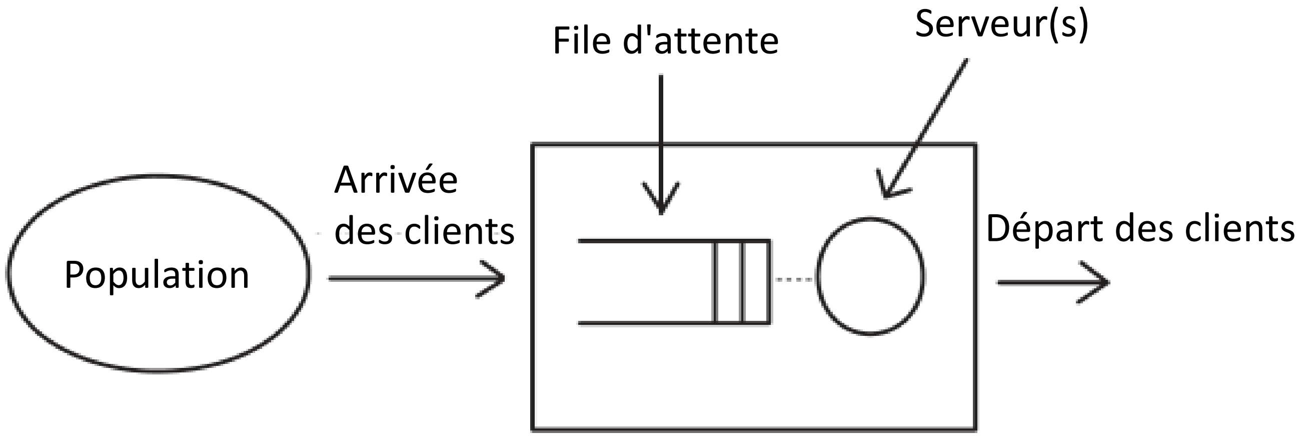
Centre Universitaire de AbdElhafid Boussouf, Mila
$2^{ème}$ Année Master Intelligence artificielle et ses applications
Année universitaire : 2022/2023
Matière: **Modélisation et simulation**

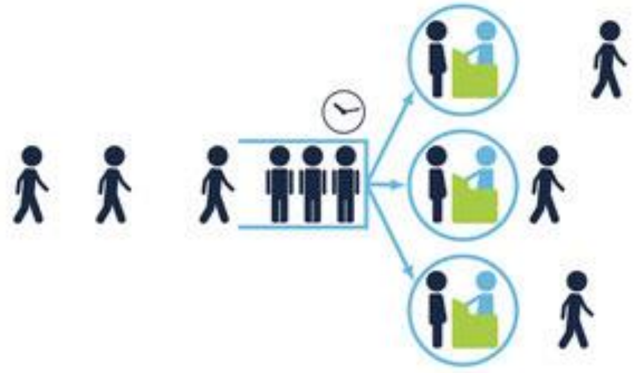
CHAPITRE III:
LES FILES D'ATTENTE

Responsable de la matière: DR. SADEK BENHAMMADA

1. Introduction

- Les files d’attentes 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**

Système / Client / Serveur	Objectifs
Système : Bureaux de services (banque, poste, etc.) **Client : client de la banque** **Serveur : Guichet**	• Minimiser les temps d'attente des clients. • Optimiser l'utilisation des ressources, notamment les guichets et le personnel.
Système : Centre d'appel **Client : Téléphoniste** **Serveur : Appelant**	• Réduire le temps d'attente des appelants. • Optimiser le nombre d'agents pour répondre aux appels.
Système : Hôpital **Client : Les patients** **Serveur : Médecin, lit, salle d'opération**	• Minimiser les temps d'attente pour les patients. • Optimiser l'utilisation des lits, des salles d'opération et du personnel médical.
Système : Trafic aérien **Client : Avions** **Serveur : Piste d'atterrissage**	• Optimiser l'utilisation des pistes d'atterissage et des portes d'embarquement.
Système : de production (Industrie) **Client : Tâche** **Serveur : Machine**	• Minimiser les temps d'attente des produits sur les lignes de production. • Optimiser l'utilisation des machines et des ressources.
Système : Trafic Routier **Client : Voitures** **Serveur : Feu de circulation**	• Réduire les embouteillages et les temps d'attente aux intersections. • Optimiser la gestion des feux de circulation.
Système : supermarché **Client : Acheteur** **Serveur : Poste de caisse**	• Minimiser les temps d'attente aux caisses. • Optimiser le nombre de caisses ouvertes en fonction de la charge.
Système : Ordinateur **Client : processus** **Serveur : CPU**	• Optimiser l'utilisation du processeur et des ressources système. • Réduire les temps d'attente pour l'exécution de tâches.
Système : Réseau informatique **Client : Paquets de données** **Serveur : Routeur**	• Optimiser la gestion du trafic sur le réseau. • Minimiser les temps d'attente pour la transmission de données.

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.

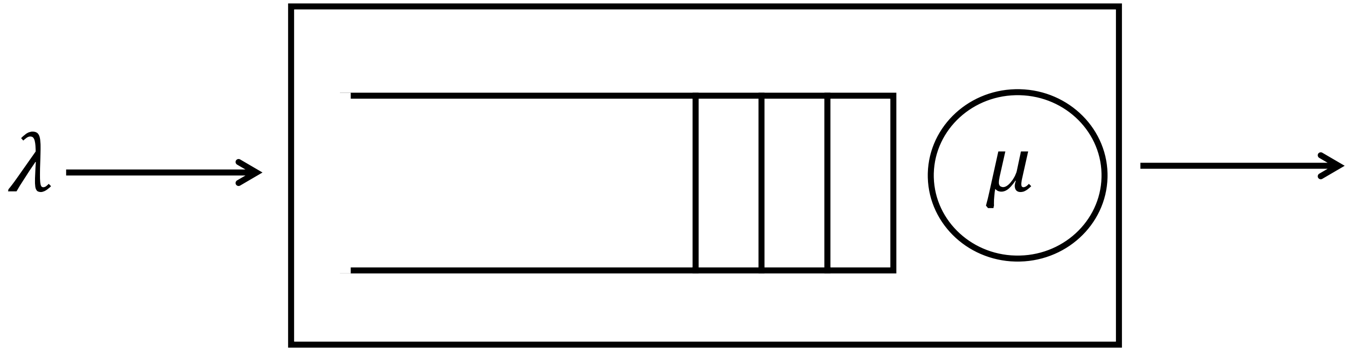
2.1. Notation de Kendell

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

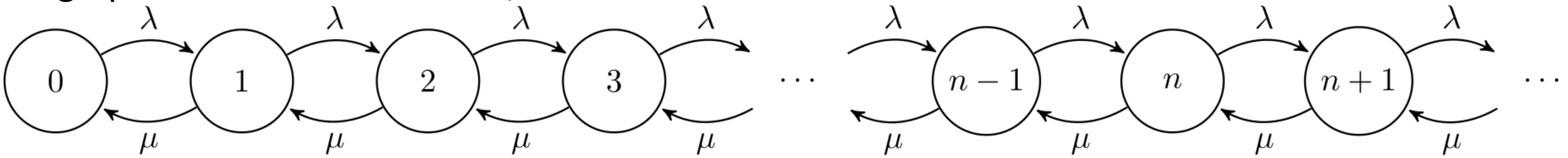
A/B/C/K/D

- **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'arrivé est général.
  - D (déterministe) si le processus d'arrivé et déterministe et déterministe constante.
  - etc.
- **B : La loi de la durée de service. Il peur être:** M,G, D, etc.
- **C :** Nombre de serveurs travaillant en parallèle (1 ou plus)
- **K :** Capacite 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 valeur correspondent aux 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:

$$\lambda_i = \lambda, \qquad i = 0,2,3, \dots$$
$$\mu_i = \mu, \qquad i = 1,2,3, \dots$$

- La matrice de taux de transition :

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

Arrivée avant un départ et départ avant une arrivée

- Temps pour qu’une nouvelle arrivée se produise :

$$A \sim Exp(\lambda)$$

- Temps pour qu’une nouvelle départ se produise :

$$D \sim Exp(\mu)$$

- 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 :

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

Les probabilités stationnaires:

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

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

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

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

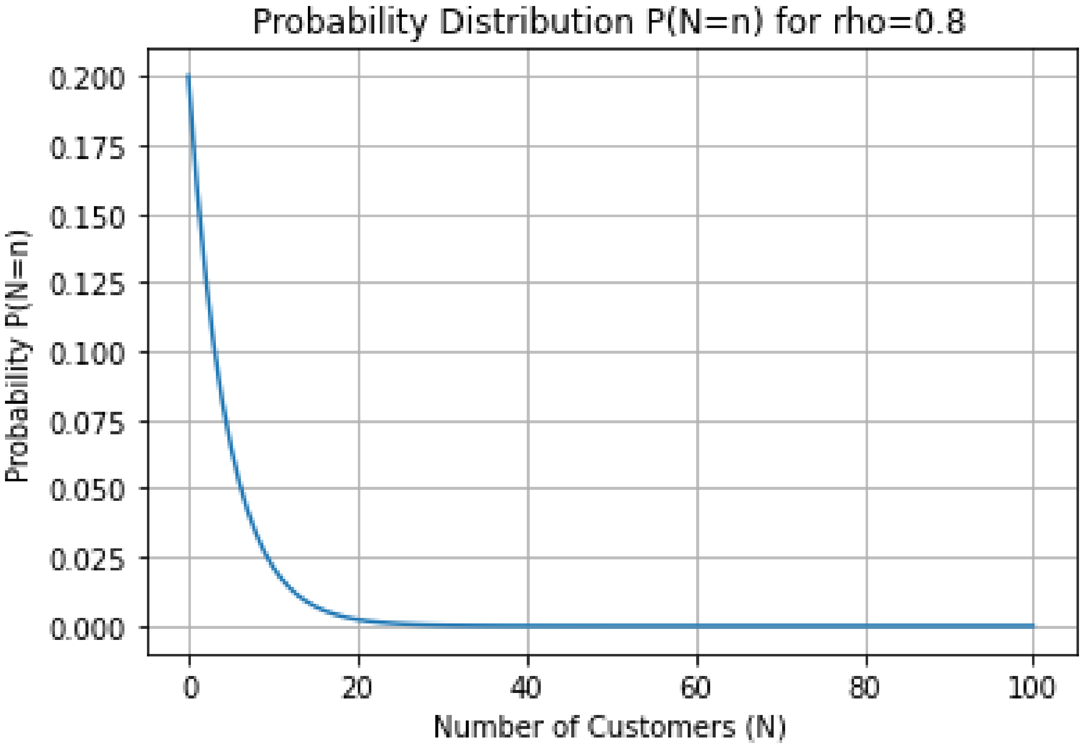
Par application de la somme géométrique: $\sum_{n=1}^{\infty} ar^n = \frac{a}{1-r}$ (ssi $r < 1$).

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

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

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

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



$$\boxed{\boldsymbol{P(N=0) = P_0 = 1 - \rho}}$$

$$\boxed{\boldsymbol{P(N=n) = P_n = (1-\rho)\rho^n}}$$

Caractéristiques du système M/M/1

- $\overline{N}_Q$: Le nombre moyen de clients dans la file d'attente
- $\overline{N}_S$: Le nombre moyen de clients en train d'être servis.
- $\overline{N} = E(N) = \overline{N}_Q + \overline{N}_S$: 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

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

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

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

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

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

- **$\boldsymbol{\bar{N}_Q}$: 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

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

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

Caractéristiques du système M/M/1

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

Loi de Little:

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

Donc:

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

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

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

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

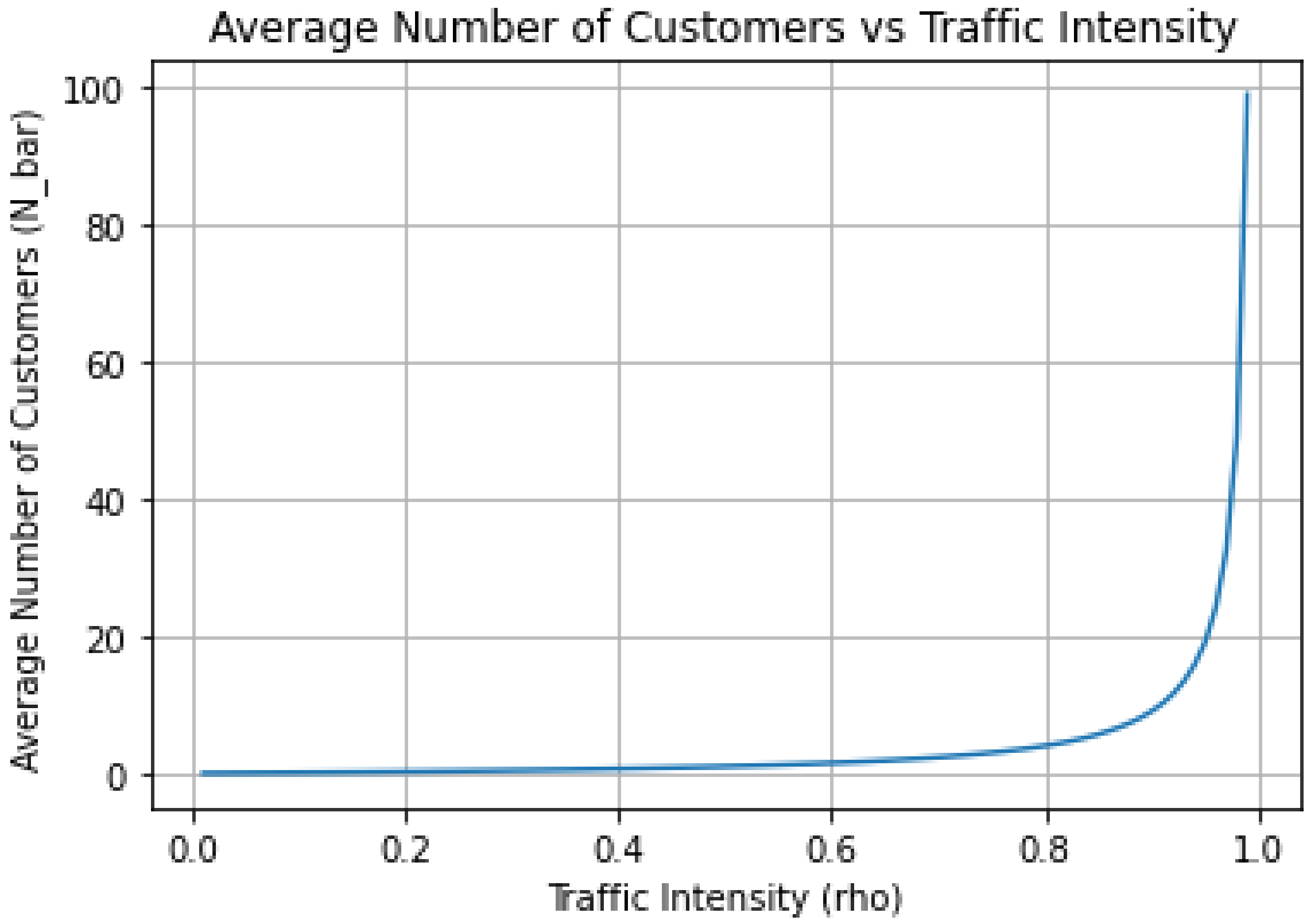
$$\bar{T}_Q = \bar{T} - \bar{T}_S = \frac{\mathbf{1}}{\boldsymbol{\mu - \lambda}} - \frac{\mathbf{1}}{\boldsymbol{\mu}} = \frac{\lambda}{\boldsymbol{\mu}(\boldsymbol{\mu} - \lambda)}$$

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

Caractéristiques du système M/M/1

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

Caractéristiques du système M/M/1



Caractéristiques du système M/M/1

Exemple

On considère une file d'attente M/M/1 de taux λ = 3/2 et μ = 2. Calculer :

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

Solution:

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

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

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

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

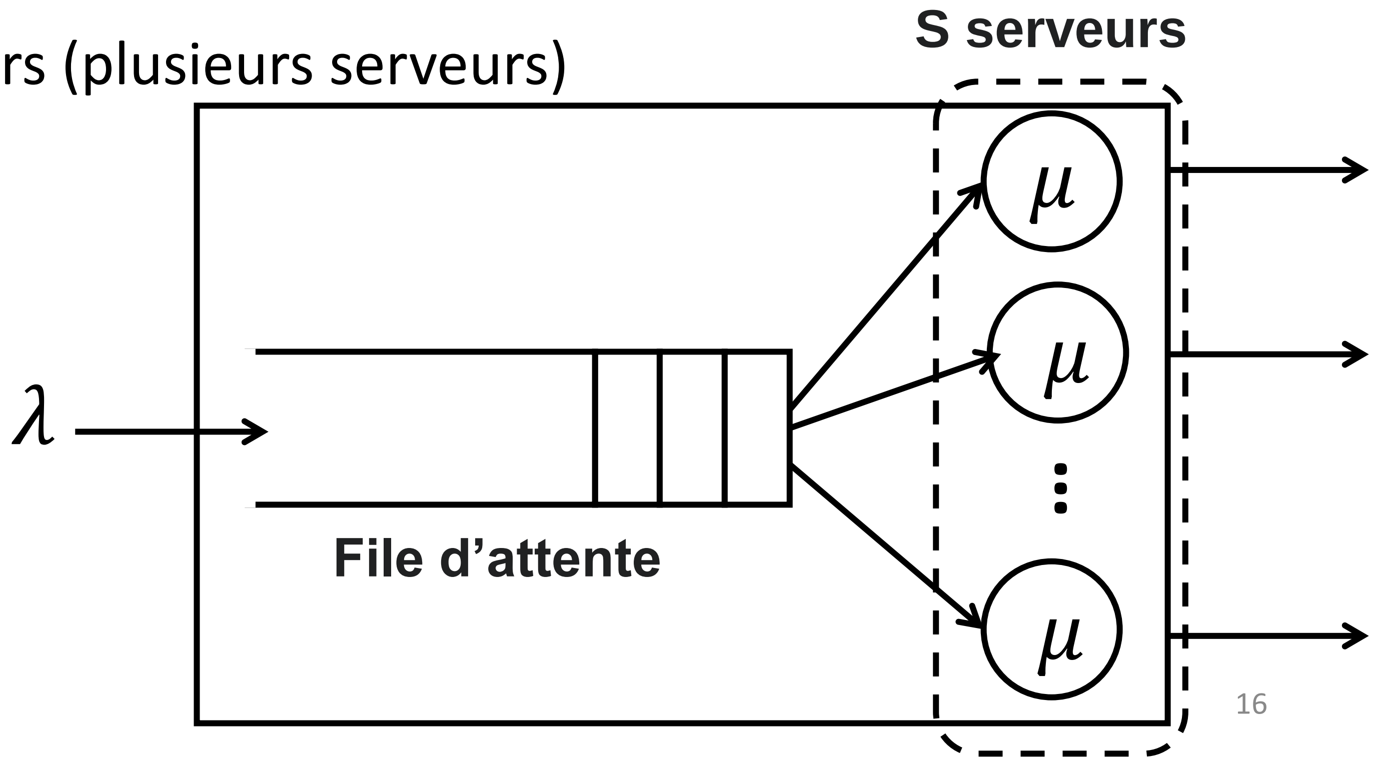
$$\boldsymbol{\overline{T} = \frac{\overline{N}}{\lambda} = \frac{3}{3/2} = 2}$$

$$\boldsymbol{\overline{T}_S = \frac{1}{\mu} = \frac{1}{2}}$$

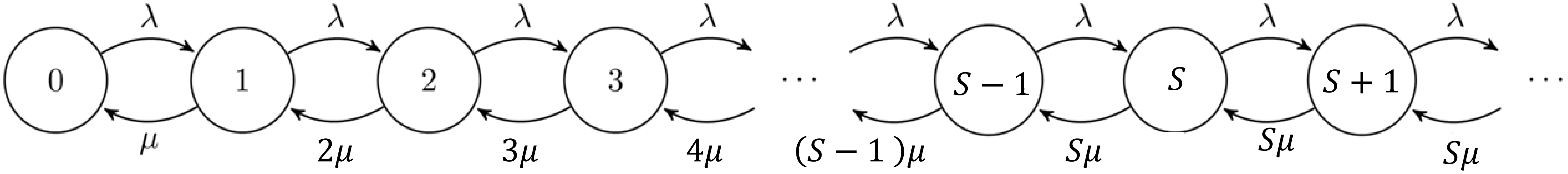
$$\boldsymbol{\overline{T}_Q = \frac{\lambda}{\mu(\mu-\lambda)} = \overline{T} - \overline{T}_S = \frac{3}{2}}$$

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:

$$\lambda_k = \lambda, \qquad k = 0,2,3, \ldots$$

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

Les probabilités stationnaires:

Le processus est ergodique et admet un 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:

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

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

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

$\boldsymbol{N}$: Le nombre de clients au temps t dans le système

$\boldsymbol{P(N} = n) = \mathrm{P}_n$: La probabilité d'avoir $\boldsymbol{n}$ clients dans le système

Formule Erlang C

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

$$P_{waiting} = P(n \geq S) = \sum_{n=S}^{\infty} P_n$$

$$= 1 - P(n < S) = \sum_{n=0}^{S-1} P_n = \frac{\frac{S}{S-\rho}}{\sum_{n=0}^{S-1} \frac{\rho^n}{n!} + \frac{\rho^S}{S!} \frac{1}{1-a}} = P_0 . \frac{\rho^S}{S!} \frac{S}{S-\rho}$$

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

$\mathbf{C(S, \rho)}$ est appelée **La formule Erlang C,** et exprime la probabilité qu'un client arrivant doive faire la queue (au lieu d'être immédiatement servi).

Caractéristiques du système M/M/S

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

$$\overline{N} = E(N) = \sum_{\boldsymbol{k}=1}^{\infty} \boldsymbol{k} \times P(N = \boldsymbol{k}) = \sum_{\boldsymbol{k}=1}^{\infty} \boldsymbol{k} P_{\boldsymbol{k}}$$

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

$\boldsymbol{\overline{N}_Q}$**: 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

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

$$\boxed{\boldsymbol{\overline{N}_Q} = \rho \frac{C(S,\rho)}{S-\rho}}$$

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

$$\boxed{\boldsymbol{\overline{N}_S} = \rho = \frac{\boldsymbol{\lambda}}{\mu}}$$

Caractéristiques du système M/M/S

- $\boldsymbol{\overline{T}}$ **: Temps moyen d’attente dans le système** (attente + service)

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

Donc:

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

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

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

$$\boxed{\boldsymbol{\overline{T}_S = \frac{1}{\mu}}}$$

$\overline{T}_Q$: Le temps moyen d’attente dans la file:

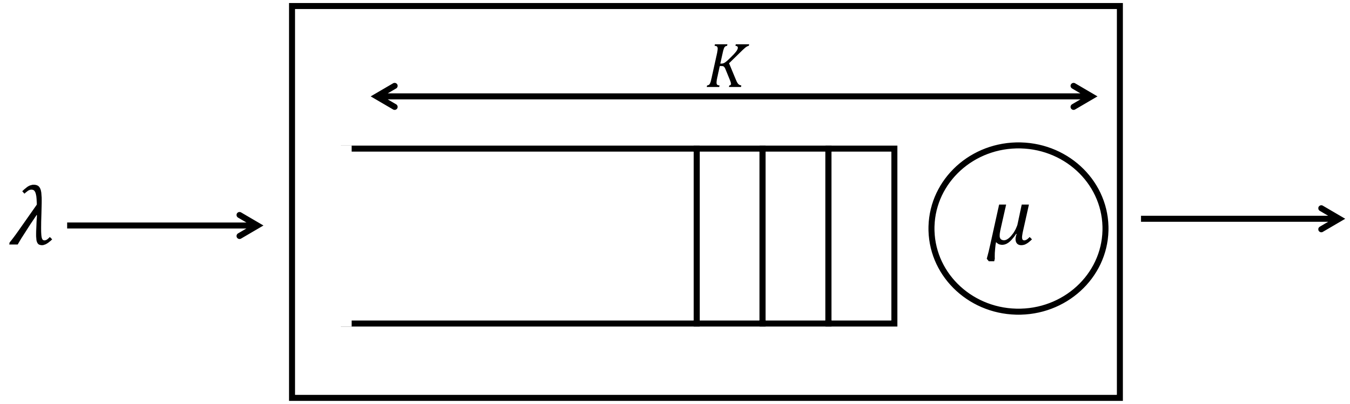
$$\overline{T}_Q = \overline{T} - \overline{T}_S = \boldsymbol{\frac{1}{\mu}\left(1 + \frac{C(S,\rho)}{(S-\rho)}\right) - \frac{1}{\mu}}$$

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

2.3. Modèle M/M/S

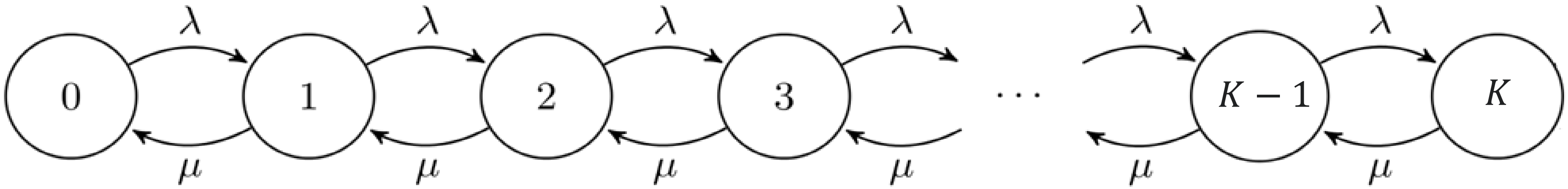
	M/M/1	**M/M/S**
Condition de stationnarité	$\rho = \frac{\lambda}{\mu} < 1$	$a = \frac{\lambda}{S\mu} < 1$ (a: le taux d'utilisation du système)
$\boldsymbol{P_0}$**: Probabilité que le système est vide (Aucun client dans le système)**	$P_0 = 1 - \rho$	$P_0 = \left(\sum_{k=0}^{S-1} \frac{\rho^k}{k!}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

- **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 $\boldsymbol{K}$ (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 $E = \{0,1,\ldots,K-1,K\}$, avec les taux de transitions $\lambda_i = \lambda, i = 0,\ldots,K-1$ et $\mu_i = \mu, i = 1,\ldots,K$.



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

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

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

Remarques.

- Si $\rho = 1$ $(\lambda = \mu)$, alors: $P_n = \frac{1}{K+1}$, $n = 0,1,\ldots,K$
- La probabilité de rejet d'un client est la probabilité que le nombre maximum de clients ***K*** est atteint:

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

Caractéristiques du système M/M/1/K

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

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

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

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

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

$\boldsymbol{\bar{N}_Q}$: 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

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

Caractéristiques du système M/M/1/K

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

Loi de Little:

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

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).

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

Donc:

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

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

$$\boxed{\overline{T}_S = \frac{1}{\mu}}$$

$\overline{T}_Q$: Le temps moyen d’attente dans la file:

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

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

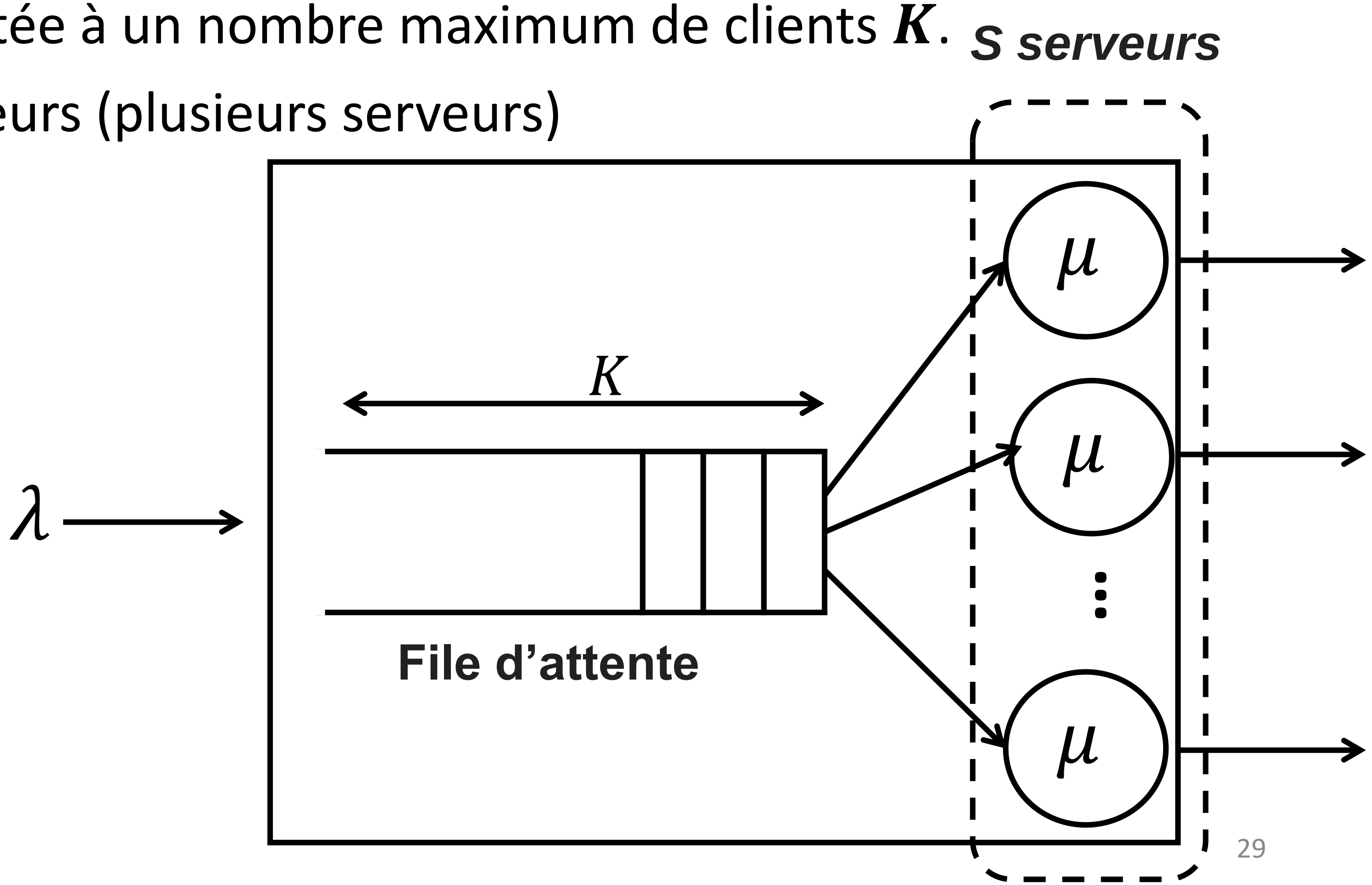
Caractéristiques du système M/M/1/K

	M/M/1	**M/M/1/K**
Condition de stationnarité	$\frac{\lambda}{\mu} < 1$	Le système est stationnaire
$\boldsymbol{P_0}$**: 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**	/	$P_K = \frac{(1-\rho)\rho^K}{1-\rho^{K+1}}$

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

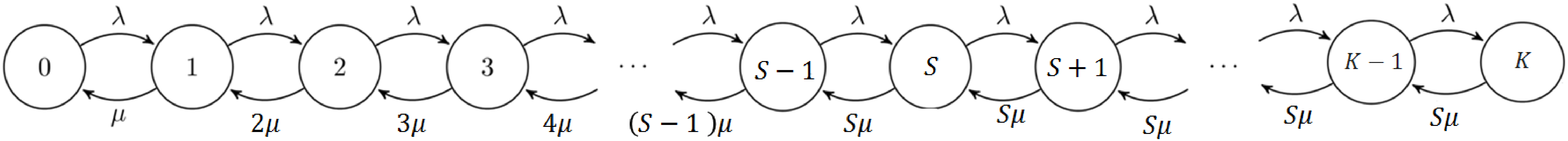
- **Le modèle M/M/S est l’équivalent de $M/M/S/K\ /FIFO$:**
  - Le processus des arrivées suit une loi de poisson de paramètre λ.
  - La durée de service suit une loi exponentielle de paramètre μ.
  - La capacité du système est limitée à un nombre maximum de clients $\boldsymbol{K}$.
  - Le service est 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:

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

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



- 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* $a = \frac{\rho}{S}$

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

$$\boldsymbol{P_n} = \begin{cases} \dfrac{\boldsymbol{\rho^n}}{\boldsymbol{n}!}\boldsymbol{P_0} & 1 \le n \le \boldsymbol{S} \\[2ex] \dfrac{\boldsymbol{\rho^n}}{S^{\boldsymbol{n-S}}\boldsymbol{S}!}\boldsymbol{P_0} & S \le n \le \boldsymbol{K} \end{cases}$$

- P_K**: Probabilité de rejet d'un client**

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

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

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

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

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

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

- Le taux d’entrée dans le système $\lambda_e = \lambda(1 - P_{\boldsymbol{K}})$

$$\bar{T} = \frac{\bar{N}}{\lambda_e} = \frac{\bar{N}}{\lambda(1 - P_{\boldsymbol{K}})}$$

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

	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)	$P_w = P_0 \frac{\rho^S}{S!} \frac{1-\left(\frac{\rho}{S}\right)^{K-S}}{1-\frac{\rho}{S}}$	$C(S,\rho) = P_0 \times \frac{\rho^S}{(S-1)!\,(S-\rho)}$
$\overline{N}$: Nombre moyen de clients dans le système	$\overline{N} = \sum_{n=0}^{K} n P_n$	$\overline{N} = \rho \left(1 + \frac{C(S,\rho)}{S-\rho} \right)$
$\overline{N}_Q$: Nombre moyen de clients dans la file d'attente	$\overline{N}_Q = \sum_{n=S+1}^{K} (n-s) P_n$	$\overline{N}_Q = \rho \frac{C(S,\rho)}{S-\rho}$
$\overline{N}_S$: Nombre moyen de clients en service (au guichet)	$\overline{N}_s = \rho(1-P_K)$	$\overline{N}_S = \rho$
$\overline{T}$: Temps moyen qu'un client passe dans le système (attente + service). (temps moyen de séjour d'un client dans le système)	$\overline{T} = \frac{\overline{N}}{\lambda_e} = \frac{\overline{N}}{\lambda(1-P_K)}$	$\overline{T} = \frac{1}{\mu} \left(1 + \frac{C(S,\rho)}{(S-\rho)} \right)$
$\overline{T}_Q$: Temps moyen d'attente dans la file	$\overline{T}_Q = \frac{\overline{N}_Q}{\lambda_e} = \frac{\overline{N}_Q}{\lambda(1-P_K)}$	$\overline{T}_Q = \frac{C(S,\rho)}{\mu(S-\rho)}$
$\overline{T}_S$: Temps moyen de service	$\overline{T}_S = \frac{1}{\mu}$	$\overline{T}_S = \frac{1}{\mu}$