

## Teoria de Filas e Rede de Petri Estocástica: Um tutorial

Queueing Theory and Stochastic Petri Net: A tutorial

Teoría de las Colas y Red de Petri Estocástica: Un tutorial

Recebido: 10/02/2021 | Revisado: 20/02/2021 | Aceito: 24/02/2021 | Publicado: 03/03/2021

**John Wesley Soares de Lima**

ORCID: <https://orcid.org/0000-0003-1606-6517>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: john.lima@ufrpe.br

**Gustavo Callou**

ORCID: <https://orcid.org/0000-0002-7997-374X>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: gustavo.callou@ufrpe.br

**Ermeson Andrade**

ORCID: <https://orcid.org/0000-0002-9614-4492>

Universidade Federal Rural de Pernambuco, Brasil

E-mail: ermeson.andrade@ufrpe.br

### Resumo

A teoria das filas estuda sistemas que exigem que os clientes ou objetos esperem. Por outro lado, as redes de Petri são um formalismo matemático usado para modelar, analisar e otimizar sistemas complexos. Embora a teoria das filas tenha sido amplamente usada na literatura, ela não pode ser usada para modelar certos comportamentos de sistemas complexos como dependências entre componentes ou comportamento de falha e reparo. Assim, este tutorial tem como objetivo apresentar ao leitor o vocabulário e as construções de modelos de redes de Petri e ilustrar a equivalência desses modelos a sistemas de filas por meio de um conjunto de exemplos. Os resultados estatísticos obtidos via simulação revelaram que os modelos de rede de Petri podem representar sistemas de filas. Além disso, esperamos que este tutorial sirva de complemento às limitações da teoria das filas e ajude pesquisadores ou profissionais, tanto da academia quanto da indústria, a modelarem sistemas complexos com base nos exemplos apresentados neste trabalho.

**Palavras-chave:** Redes de Petri; Análise de filas; Modelagem de sistemas.

### Abstract

Queuing theory studies systems that require customers or objects to wait. On the other hand, Petri nets are a mathematical formalism used to model, analyze and optimize complex systems. Although queuing theory has been widely used in the literature, it cannot be used to model complex behavior like dependencies between system components or failure and repair behavior. Therefore, this tutorial is meant to introduce the reader to the vocabulary and constructions of Petri net models and illustrate the equivalence of these models to queuing systems by means of a set of examples. The statistical results obtained through simulation revealed that Petri net models can represent queuing systems. In addition, we hope that this tutorial will complement the limitations of queuing theory and help researchers or professionals, both from academia and industry, to model complex systems based on the examples presented in this work.

**Keywords:** Petri nets; Queueing analysis; Systems modeling.

### Resumen

La teoría de colas estudia sistemas que requieren que los clientes u objetos esperen. Por otro lado, las redes de Petri son un formalismo matemático utilizado para modelar, analizar y optimizar sistemas complejos. Aunque la teoría de las colas se ha utilizado ampliamente en la literatura, no se puede utilizar para modelar determinados comportamientos de sistemas complejos, como dependencias de componentes o comportamiento de fallos y reparaciones. Por lo tanto, este tutorial tiene como objetivo presentar al lector el vocabulario y las construcciones de los modelos de red de Petri e ilustrar la equivalencia de estos modelos con los sistemas de cola a través de un conjunto de ejemplos. Los resultados estadísticos obtenidos mediante simulación revelaron que los modelos de red de Petri pueden representar sistemas de colas. Además, esperamos que este tutorial complemente las limitaciones de la teoría de las colas y ayude a los investigadores o profesionales, tanto del mundo académico como de la industria, a modelar sistemas complejos basados en los ejemplos presentados en este trabajo.

**Palabras clave:** Redes de Petri; Análisis de colas; Modelado de sistemas.

## 1. Introdução

Na vida diária, no trabalho ou na universidade, nos deparamos com problemas complexos e abrangentes. Para que possamos resolver tais problemas com facilidade, modelos podem ser usados. A vantagem de se usar modelos é que eles nos permitem abstrair um problema complexo de modo a torná-lo mais simples. Isso consiste em desconsiderar detalhes sem importância para se concentrar nas características essenciais do problema (Dubois G. 2018; Doo-Kwon Baik, 2005; Havelange & Rêgo, 2015).

De maneira mais específica, a modelagem de sistemas é um estudo multidisciplinar que visa representar os sistemas através de modelos, de modo que os mesmos possam ser analisados sem a implementação real do sistema. Em se tratando de sistemas físicos ou computacionais, geralmente, usamos modelos matemáticos para simular e analisar numericamente tais sistemas. Modelagem matemática é uma representação matemática idealizada de um dado fenômeno do mundo real (Bassanezi, 2002; Doo-Kwon Baik, 2005; Hermann, et al., 2020). Por sua vez, um fenômeno pode ser classificado como determinístico ou não determinístico (também chamado de estocástico ou probabilístico). Um fenômeno determinístico é aquele que, quando ocorre sob condições idênticas, é possível se prever o resultado que irá ocorrer. O fenômeno não determinístico é aquele que, sob condições idênticas, não é possível prever o resultado particular que irá ocorrer, e sim, o conjunto dos possíveis resultados (Fogliatti & Mattos, 2007; Taha, 2006; Viali, 2006). Um fenômeno estocástico que iremos estudar e modelar neste artigo são as filas.

A teoria de filas descreve a modelagem de sistemas ou processos que resultam em espera, isto é, filas. As filas estão presentes em diversos setores de nossas vidas, seja numa fila de compras num mercado ou numa requisição computacional num servidor. O objetivo da teoria de filas é otimizar o desempenho de um sistema, reduzindo seus custos operacionais. Para isso, o sistema deve satisfazer alguns critérios de desempenho mínimo (Fogliatti & Mattos, 2007; Taha, 2006; Hillier & Lieberman, 2006). Embora a teoria de filas seja amplamente utilizada na literatura, ela não pode ser usada para modelar comportamentos complexos, como dependências entre componentes ou comportamentos de falha e reparo (Kounev & Buchmann, 2003). As redes de Petri, por outro lado, são um formalismo matemático que permite suprir tais limitações. Assim, advogamos o uso das redes de Petri em tais situações.

Existem na literatura alguns tutoriais com foco em redes de Petri. Através de exemplos, Zurawski e Zhou (1995) apresentam os conceitos fundamentais de uma rede de Petri e aplicam as redes de Petri em sistemas industriais. Blätke, et al. (2011) apresentam um tutorial mostrando como modelar o comportamento de um sistema biológico em rede de Petri. Zimmermann (2014) mostrou um tutorial com técnicas de simulação e aplicações em redes de Petri para avaliar a confiabilidade de sistemas. Diferentemente dos trabalhos citados, neste tutorial iremos apresentar ao leitor as construções dos modelos de redes de Petri e ilustrar a equivalência desses modelos aos sistemas de filas por meio de um conjunto de exemplos. Além disso, os resultados revelaram que os modelos de rede de Petri podem de fato representar sistemas de filas. Esperamos que este tutorial sirva de base para pesquisadores ou profissionais (da academia ou da indústria), de modo que, caso eles venham a enfrentar as limitações da teoria de filas, possam usar os modelos apresentados neste trabalho para modelagem, análises e otimização de sistemas complexos.

Este artigo está organizado da seguinte forma. A Seção 2 apresenta conceitos fundamentais para um melhor entendimento deste trabalho. A Seção 3 apresenta a abordagem adotada na pesquisa. A Seção 4 detalha os modelos e os resultados e apresenta uma discussão para os mesmos. Por fim, a Seção 5 conclui este trabalho e descreve os trabalhos futuros.

## 2. Fundamentação Teórica

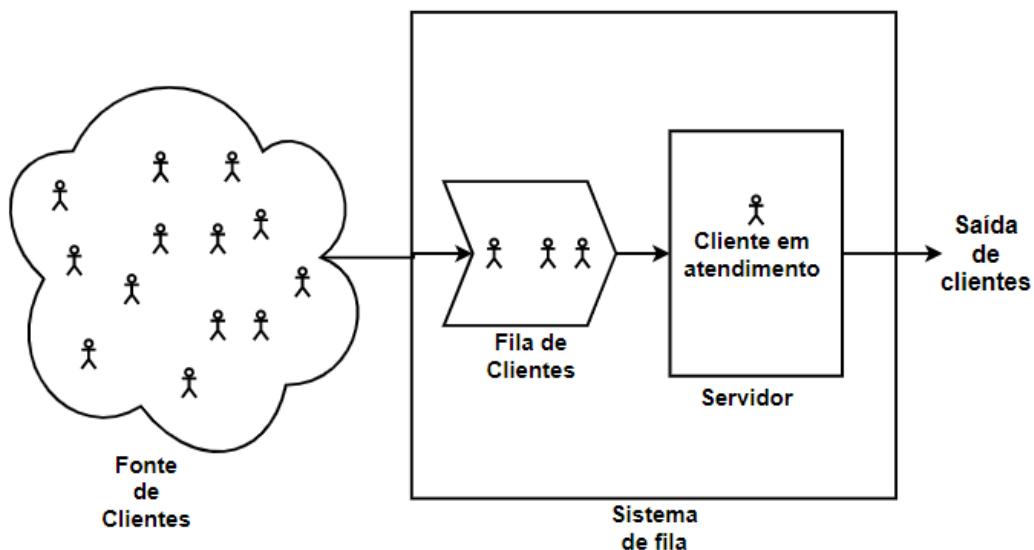
Nessa seção vamos apresentar os principais conceitos teóricos sobre teoria de filas, rede de filas e rede de Petri.

## 2.1 Teoria de Filas

Os elementos que compõe a estrutura de um sistema de filas são: fonte (população de clientes), cliente, fila de clientes e servidor. A fonte é o lugar de onde os clientes se originam e entram num sistema de filas de modo que esse sistema é composto pelos servidores e pela fila de clientes em espera para o atendimento (Hillier & Lieberman, 2006). A Figura 1 mostra a estrutura de um sistema de filas. Note que um sistema de filas é descrito pela teoria de filas. Tal teoria toma alguns critérios para modelar o sistema, esses critérios são: número de servidores, distribuição de probabilidade que representa a chegada e saída de clientes no sistema, limitação da quantidade de clientes no sistema, disciplina da fila e a fonte que pode ser finita ou infinita. A disciplina de uma fila é o modo como será gerenciado a entrada e saída de clientes da fila. A disciplina de fila mais usada é a FCFS (ou FIFO) que, do inglês, significa o primeiro a entrar é o primeiro a sair. A quantidade de clientes na fonte pode ser considerada, para a maioria dos casos, como infinita (Fogliatti & Mattos, 2007; Taha, 2006; Hillier & Lieberman, 2006).

A forma mais comum de se representar um sistema de fila é pela a notação de Kendall-Lee (Fogliatti & Mattos, 2007). Tal notação tem sua representação como a/b/c/d/e/f. Cada letra tem o seguinte significado: a - distribuição de chegada, b - distribuição de saída, c - número de servidores em paralelo, d - disciplina da fila, e - número máximo de clientes no sistema, f - quantidade de clientes na fonte (população) (Fogliatti & Mattos, 2007; Taha, 2006;). Nesse artigo, iremos considerar os modelos de fila que tenham população infinita, que seguem processos markovianos homogêneo de tempo contínuo (os tempos entre os eventos seguem uma distribuição exponencial), representados pela letra M (Taha, 2006; Bolch, *et al.*, 2006), e disciplina de fila FIFO. Por exemplo, suponha que um sistema de fila siga um processo markoviano homogêneo de tempo contínuo na chegada e saída de clientes, tenha um servidor para atendimento, não haja limites no número de clientes no sistema, tenha fonte de clientes infinita e que a disciplina da fila seja FIFO. Portanto, pela notação de Kendall-Lee, esse sistema seria representado como: M/M/1/FIFO/inf/inf.

**Figura 1:** Estrutura de um sistema de filas.



Fonte: Autores.

Para que um sistema de fila seja eficiente, isto é, prestar um serviço ao menor custo possível e a um tempo tolerável para o cliente, é necessário que alguns parâmetros de desempenho sejam atendidos. Esses parâmetros de desempenho estão listados na Tabela 1 com suas respectivas equações, considere que  $n$  representa o número de servidores em paralelo,  $\mu$  a taxa média de atendimento dos servidores e  $\lambda$  representa a taxa de chegada ao sistema (Fogliatti & Mattos, 2007; Taha, 2006). Sobre os parâmetros de desempenho vale ressaltar que, quando  $n*\mu > \lambda$ , o sistema estará em regime estável. No entanto, quando  $n*\mu \leq \lambda$ , o sistema poderá entrar em colapso e a fila poderá aumentar sem limites (Fogliatti & Mattos, 2007; Taha, 2006; Sztrik, 2012; Hillier & Lieberman, 2006). Note que  $(1/\lambda)$  e  $(1/\mu)$ , são, respectivamente, o tempo médio de chegada de um cliente e o tempo médio de atendimento de um cliente. Dependendo do sistema de fila  $\lambda \neq \lambda_{ef}$ , por exemplo, um sistema de fila em que existe um limite no número de clientes que podem entrar (Fogliatti & Mattos, 2007; Taha, 2006;). Isso ocorre, pois, em um sistema com limitador e cheio, nem todo cliente que deseja entrar no sistema consegue entrar (Taha, 2006).

Nesse artigo, usaremos o conhecimento estabelecido na literatura acadêmica sobre os modelos matemáticos da teoria de fila para calcular os parâmetros de desempenho (Fogliatti & Mattos, 2007; Taha, 2006; Trivedi, 2008; Bolch, *et al.*, 2006; Sztrik, 2012; Hillier & Lieberman, 2006). Sobre os parâmetros  $W$  e  $W_q$ , eles têm as seguintes relações com  $L$  e  $L_q$ , respectivamente:  $W = L/\lambda_{ef}$  e  $W_q = L_q/\lambda_{ef}$ . Essas equações são chamadas de Lei de Little (Fogliatti & Mattos, 2007; Taha, 2006; Sztrik, 2012). O valor de  $W$  é chamado de tempo de resposta. O tempo de resposta ( $W$ ) é o tempo médio gasto pelo cliente desde que entra no sistema até finalizar o seu atendimento, isto é, como já mostrado na Tabela 1, é o tempo médio que um cliente permanece no sistema. Note que  $W = W_q + (1/\mu)$ , isto é, o tempo de resposta é a soma do tempo que um cliente permanece em espera na fila para iniciar o atendimento mais o tempo médio de atendimento (Bolch, *et al.*, 2006; Trivedi, 2008).

**Tabela 1:** Parâmetros de desempenho de um sistema de fila (sendo  $n$  o número de servidores em paralelo e com mesma taxa  $\mu$ ).

Parâmetro	Equação
$\lambda$ – taxa média nominal de chegada de clientes.	Nº de clientes chegando no sistema/unidade de tempo
$\lambda_{ef}$ – taxa média efetiva de chegada de clientes (vasão).	Nº efetivo de clientes chegando no sistema/unidade de tempo
$\mu$ – taxa média de atendimento de um servidor no sistema.	Nº de clientes atendidos/unidade de tempo
$\rho$ – fator de utilização (taxa de utilização).	$\lambda_{ef}/(n*\mu)$
$L$ – número médio de clientes no sistema.	$W * \lambda_{ef}$
$L_q$ – número médio de clientes esperando na fila.	$W_q * \lambda_{ef}$
$W$ – tempo médio que um cliente permanece no sistema (tempo de resposta).	$L / \lambda_{ef}$
$W_q$ – tempo médio que um cliente permanece em espera na fila.	$L_q / \lambda_{ef}$

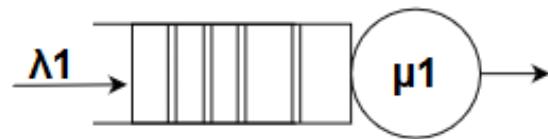
Fonte: Autores.

## 2.2 Rede de filas

Uma rede de filas é um sistema de filas complexo compostos por sistemas de filas simples. Um sistema de fila simples, por sua vez, é chamado de nó ou fila de estação única e é composto por uma fila de clientes e um serviço que pode ter um único servidor ou servidores em paralelo (Figura 1) (Trivedi, 2008; Bolch, *et al.*, 2006). Na literatura, é comum representar

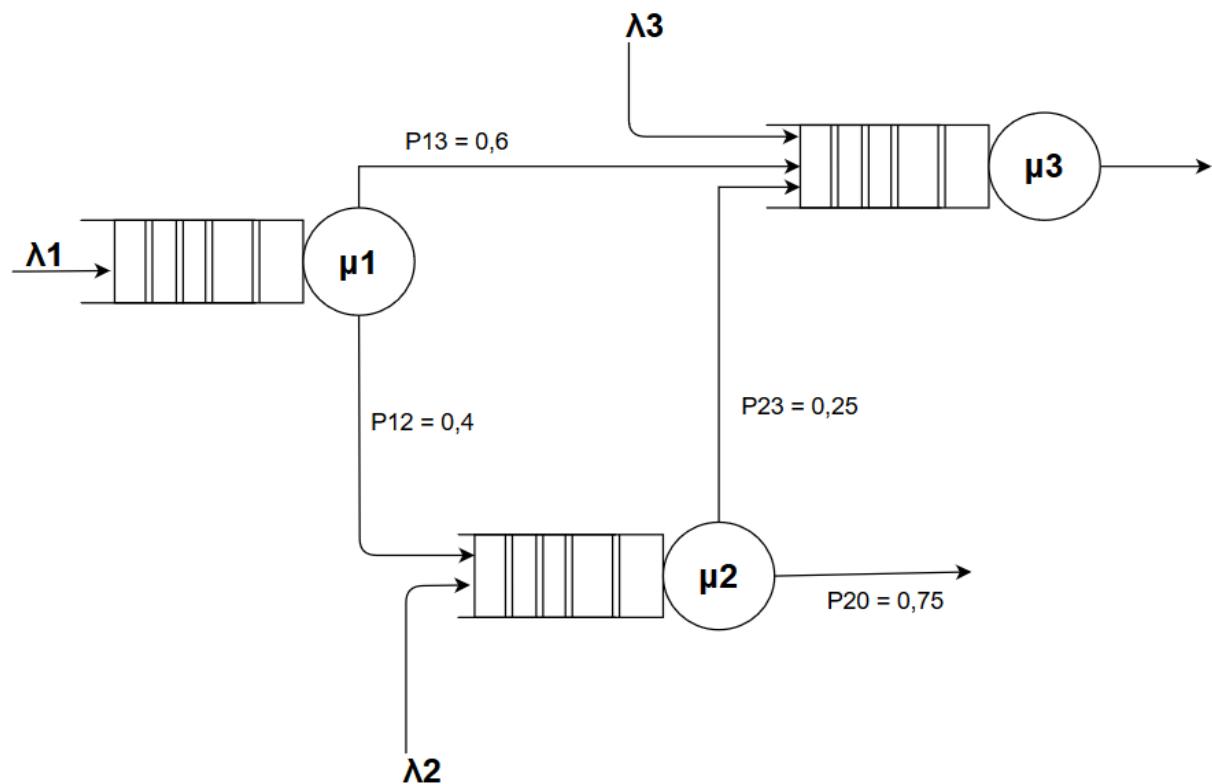
uma fila de estação única com taxas  $\lambda_1$  e  $\mu_1$ , como mostrado na Figura 2 (Fogliatti & Mattos, 2007; Taha, 2006; Trivedi, 2008).

**Figura 2:** Fila de estação única.



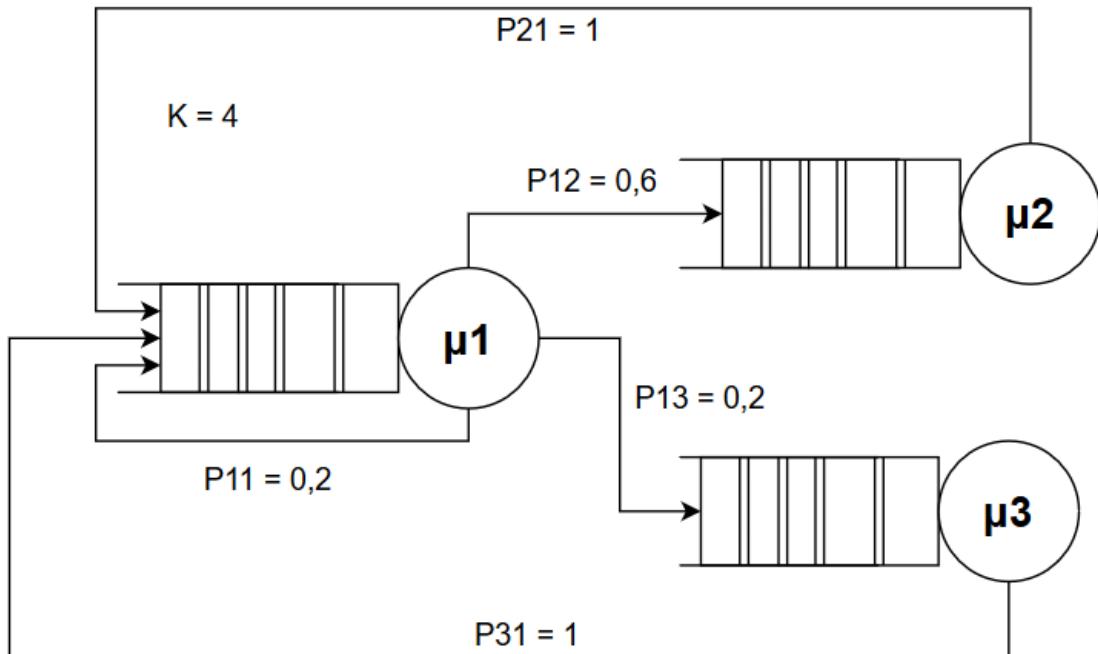
Fonte: Autores.

**Figura 3:** Rede de filas aberta com três nós com um único servidor cada.



Fonte: Autores.

**Figura 4:** Rede de filas fechada com  $K = 4$  (número de clientes) e com três nós com um único servidor cada.



Fonte: Autores.

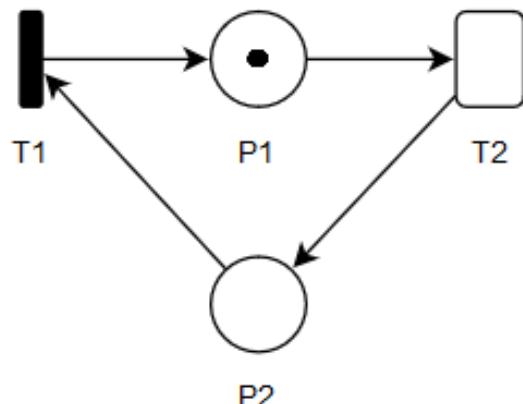
As redes de filas podem ser abertas (Figura 3) ou fechadas (Figura 4) (Bolch, *et al.*, 2006). Ambas as redes das Figuras 3 e 4 são compostas por três nós com um único servidor. O teorema de Jackson e o teorema de Gordon–Newell são bem conhecidos na literatura e, respectivamente, estabelecem suposições a respeito das redes de filas abertas e fechadas (Bolch, *et al.*, 2006). O teorema de Jackson considera que, em uma rede de filas aberta, cada nó fornece um serviço independente, o tempo de serviço segue uma distribuição exponencial, a disciplina de todos os nós é FIFO, o cliente entra no sistema a partir de uma fonte infinita e a taxa de chegada segue uma distribuição de Poisson. Além disso, qualquer cliente que sai do nó  $i$ , vai imediatamente para o nó  $j$ , com uma probabilidade  $P_{ij}$ , ou sai imediatamente do sistema com probabilidade  $P_{i0}$ . As probabilidades  $P_{ij}$  e  $P_{i0}$  são chamadas de probabilidades de roteamento. A soma das probabilidades de roteamento relativas à saída de clientes de um determinado nó  $i$  é igual a 1. Note que na Figura 3, por exemplo,  $P_{23} = 0,25$  (25%) significa a probabilidade de um cliente sair do nó 2 e entrar no nó 3 e  $P_{20} = 0,75$  (75%) significa a probabilidade de um cliente sair do sistema a partir do nó 2. Note também que a soma das probabilidades de roteamento relativas à saída de clientes do nó 2 é igual a 1, ou seja,  $P_{23} + P_{20} = 1$  (100%). O teorema de Gordon–Newell considera as mesmas suposições do teorema de Jackson, no entanto, difere a respeito de que a quantidade de clientes no sistema é finita (igual a  $K$ ), os clientes não podem sair do sistema e nenhum cliente externo ao sistema pode entrar (Bolch, *et al.*, 2006). As redes mostradas nas Figuras 3 e 4 seguem, respectivamente, os pressupostos dos teoremas de Jackson e de Gordon–Newell. Assim, as redes mostradas nas Figuras 3 e 4 são chamadas, respectivamente, de rede de Jackson e rede de Gordon–Newell (Bolch, *et al.*, 2006).

### 2.3 Rede de Petri Estocástica

As Redes de Petri são um formalismo matemático muito usado na modelagem e simulação de sistemas, os quais podem ser concorrentes, assíncronos, paralelos e distribuídos. Existem diversas extensões das redes de Petri. Neste trabalho, mais especificamente, iremos utilizar a rede de Petri estocástica (ou Stochastic Petri Net – SPN) (Murata, 1989; Trivedi, 2008). A rede de Petri estocástica, graficamente, é representada por arcos (setas) que podem ter pesos, transições temporizadas,

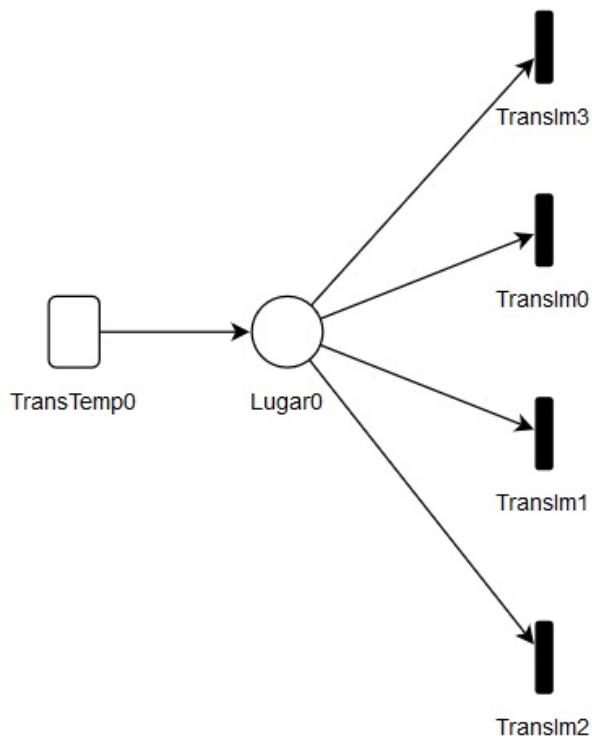
transições imediatas, lugares e tokens (também chamado de marcação). A Figura 5 ilustra um exemplo de uma rede SPN que é composta pelos lugares  $P1$  e  $P2$ , transição imediata  $T1$  e transição temporizada  $T2$ . No lugar  $P1$  existe um token associado (Murata, 1989; Trivedi, 2008).

**Figura 5:** Exemplo de uma GSPN.



Fonte: Autores.

**Figura 6:** Exemplo de uma GSPN.



Fonte: Autores.

De forma mais específica, uma rede SPN que contém uma transição imediata ( $T1$ ) é chamada de Generalized Stochastic Petri Net (GSPN) (Murata, 1989). A Figura 5 ilustra uma GSPN. A diferença entre as transições  $T2$  e  $T1$ , é que a transição  $T2$  demanda um tempo (ou delay) para ser executada, já a transição imediata  $T1$  não demanda tempo e ocorre de

modo prioritário em relação à  $T_2$ , isto é, ocorre de modo imediato ou com tempo equivalente a zero (0). Os tempos associados às transições temporizadas (ex.  $T_2$ ) são distribuídos exponencialmente, que é a principal característica de uma SPN ou GSPN (Murata, 1989). Além disso, a transição imediata tem um peso associado. Ele indica qual a probabilidade de uma determinada transição imediata disparar. Observe a Figura 6, existe um lugar  $Lugar0$  cuja origem de quatro arcos inicia nele e conecta quatro transições imediatas. Na modelagem de sistemas com RdP, quando um lugar cuja a origem dos arcos inicia nele e se conecta a transições imediatas, é necessário indicar qual a probabilidade de disparo dessas transições (peso). Se um lugar se conecta a apenas uma transição imediata o peso dessa transição será 1. De outro modo, quando um lugar se conecta a mais de uma transição imediata o peso dessas transições devem ser entre 0 e 1 (Balbo, 2001). No entanto, como o peso é uma probabilidade (varia de 0% a 100% ou 0 a 1), a soma dos pesos das transições imediatas conectadas a um mesmo lugar deve ser igual a 1. Na Figura 6, cada transição imediata tem peso igual a 0.25, cuja soma desses pesos é 1. Nesse artigo, iremos referenciar GSPN como SPN.

### 3. Metodologia

No processo de equiparar os modelos de teoria de filas com as SPNs, primeiramente, fizemos um estudo dos modelos matemáticos derivados da teoria de filas que são bem conhecidos e definidos na literatura acadêmica (Fogliatti & Mattos, 2007; Taha, 2006; Trivedi, 2008; Bolch, *et al.*, 2006; Sztrik, 2012; Hillier & Lieberman, 2006). Posteriormente, utilizamos a ferramenta Mercury (versão 4.6.5) para modelar os sistemas de fila em modelos SPNs equivalentes e realizar cálculos dos parâmetros de desempenho das filas (Silva, *et al.*, 2015). Os cálculos desses parâmetros de desempenho, tomando por base os modelos em SPN, foram feitos via o uso de simulação estacionária (Silva, *et al.*, 2015; Maciel, *et al.*, 2017; Chung, 2019; Kleijnen, 2007). Na simulação estacionária usamos um intervalo de confiança de 95% e consideramos um erro relativo máximo de 5%. Finalmente, comparamos os resultados entre os modelos matemáticos da teoria de filas e as SPNs, onde foi analisado se os valores médios tanto da simulação quanto do modelo teórico de fila estão dentro do intervalo de confiança.

### 4. Modelos, Resultados e Discussão

Nessa seção, serão mostrados os principais sistemas de filas e seus respectivos modelos em rede de Petri. Também comparamos os resultados dos cálculos teóricos, obtidos pelo estudo da teoria de filas, com os resultados de cálculos obtidos via simulação estacionária, oriundos dos modelos em SPN.

#### 4.1 Modelo M/M/1/FIFO/inf/inf

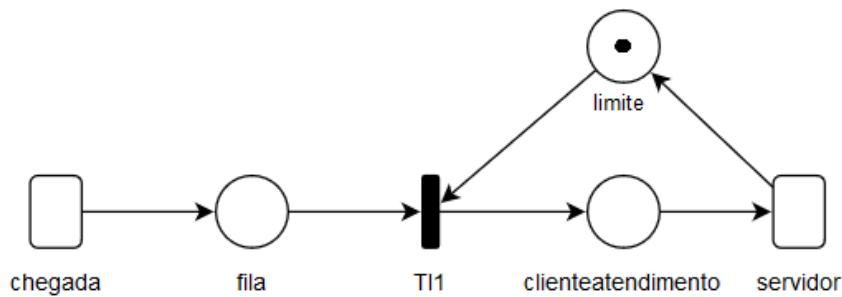
De acordo com a notação de Kendall-Lee, o modelo de sistema de filas M/M/1/FIFO/inf/inf apresenta uma distribuição markoviana dos tempos de chegada e saída de clientes no sistema; possui um servidor que atende um cliente por vez; a disciplina da fila é do tipo FIFO; o número máximo de clientes no sistema é infinito, isto é, não existem limites; e tem população infinita. Para exemplificar, suponha que a taxa média de chegada de clientes seja  $\lambda=(1/20)$  clientes/hora, e que a taxa média de atendimento do servidor no sistema seja  $\mu=(1/10)$  clientes/hora. Como não há limites de clientes nesse modelo, então  $\lambda=\lambda_{\text{ref}}$ .

O modelo de fila apresentado anteriormente pode ser mapeado na SPN mostrada na Figura 7. Neste modelo, as transições *chegada* e *servidor* representam, respectivamente, o tempo entre chegadas de clientes no sistema, e o tempo em que o cliente finaliza seu atendimento no servidor. Os valores de tempo das transições temporizadas são o inverso das taxas médias ( $\lambda$  e  $\mu$ ): *chegada* = 20 horas e *servidor* = 10 horas. Os lugares *fila* e *clienteatendimento*, representam, respectivamente, a fila de espera de clientes, e o cliente sendo atendido pelo servidor. O lugar *limite* indica que o servidor apenas atende um cliente por

vez. A transição imediata  $TII$ , com peso igual a 1, representa a ação do cliente iniciar o atendimento, desde de que o servidor estiver disponível.

A Tabela 2 lista as métricas usadas para calcular os parâmetros de desempenho via simulação estacionária, do modelo SPN. Por exemplo, a expressão  $(E\{\#fila\})$  calcula o número médio de tokens no lugar chamado *fila*. No modelo de SPN, o lugar *fila* representa a fila do sistema, e os tokens no lugar *fila* representam os clientes na fila esperando ser atendidos. Logo,  $Lq=(E\{\#fila\})$  representa o número médio de clientes que estão esperando na fila para serem atendidos.

**Figura 7:** Modelo M/M/1/FIFO/inf/inf.



Fonte: Autores.

**Tabela 2:** Métricas usadas na simulação estacionária do modelo spn (m/m/1/fifo/inf/inf).

Parâmetro	Métrica
<b>L</b>	$(E\{\#fila\})+(E\{\#clienteatendimento\})$
<b>Lq</b>	$(E\{\#fila\})$
<b>W</b>	$((E\{\#fila\})+(E\{\#clienteatendimento\}))/\lambda_{ef}$
<b>Wq</b>	$(E\{\#fila\})/\lambda_{ef}$

Fonte: Autores.

A Tabela 3 mostra a comparação dos cálculos dos parâmetros de desempenho do modelo matemático teórico (teoria de filas) em relação ao modelo SPN (via simulação estacionária). Conforme mostrado na tabela, o valor médio e o valor teórico estão dentro do intervalo de confiança indicando que o modelo SPN é estatisticamente equivalente ao modelo matemático da teoria de filas.

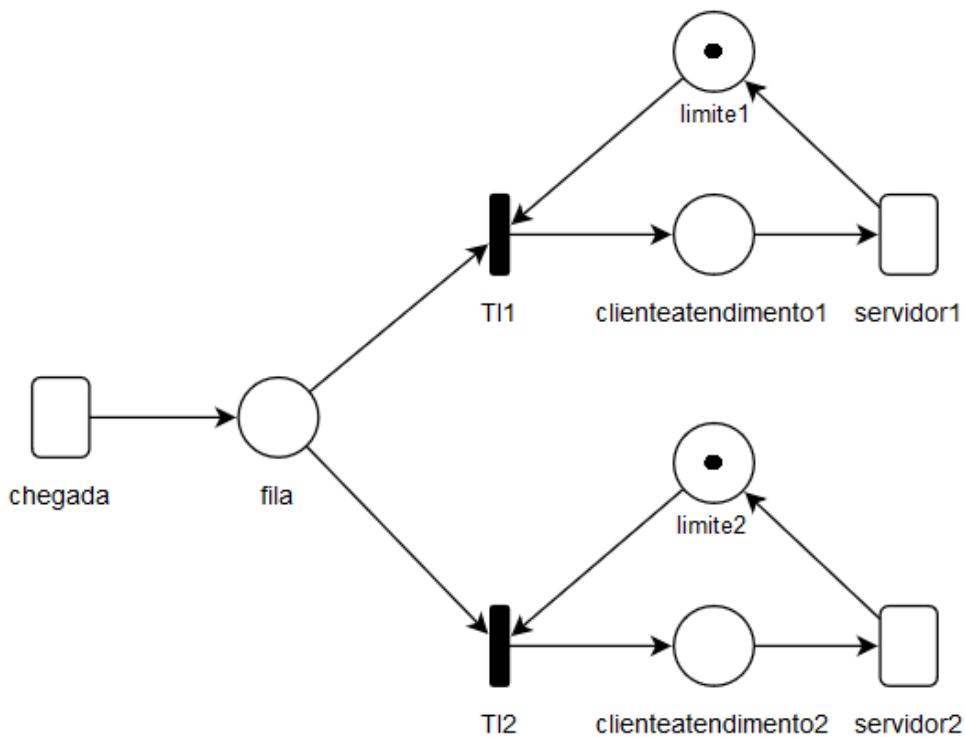
**Tabela 3:** Comparação do modelo teórico da teoria de filas com a spn (M/M/1/FIFO/inf/inf).

Parâmetros	Teórico	SPN	SPN
		(média)	(intervalo de confiança)
L	1,0	1,005188	[ 0,9966547 ; 1,013722 ]
Lq	0,5	0,5001593	[ 0,4927012 ; 0,5076174 ]
W	20,0	19,91019	[ 19,77336 ; 20,04703 ]
Wq	10,0	10,00336	[ 9,847298 ; 10,15942 ]

Fonte: Autores.

#### 4.2 Modelo M/M/n/FIFO/inf/inf e M/M/inf/FIFO/inf/inf

**Figura 8:** Modelo M/M/2/FIFO/inf/inf.

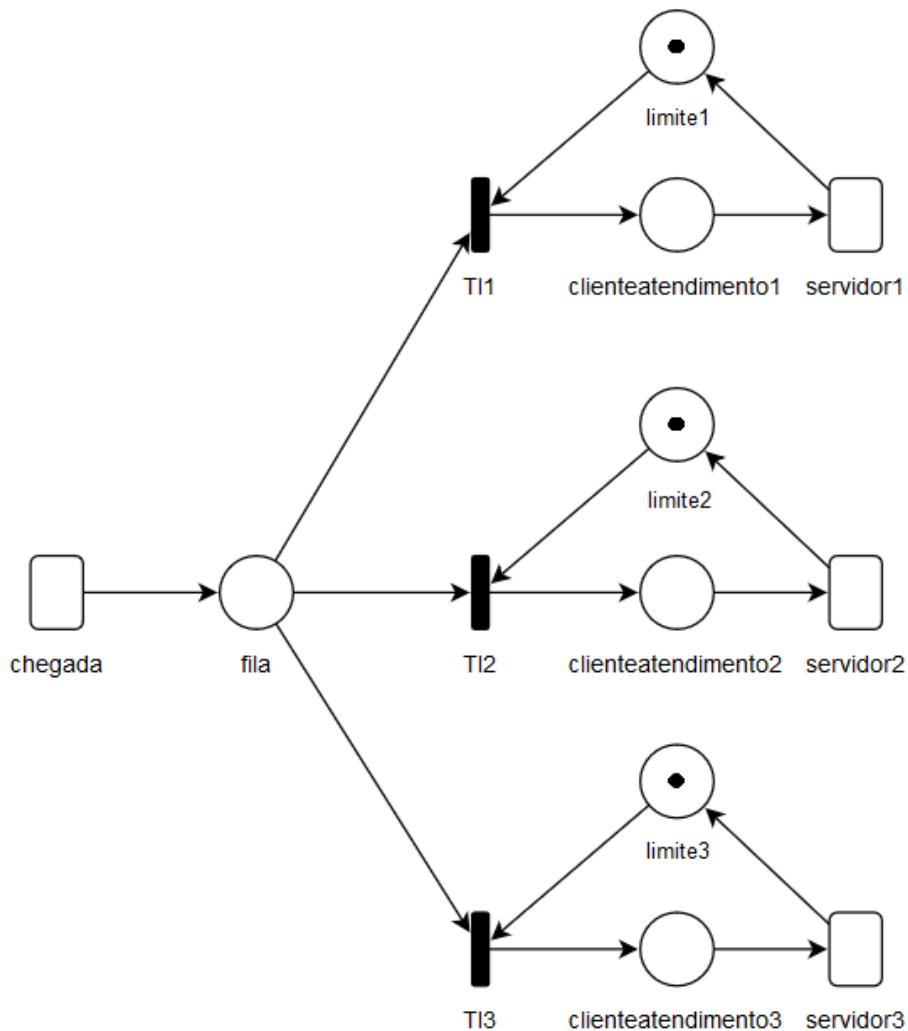


Fonte: Autores.

O modelo M/M/n/FIFO/inf/inf tem as mesmas características do modelo anteriormente apresentado para o M/M/1/FIFO/inf/inf, a diferença é que no modelo M/M/n/FIFO/inf/inf temos  $n$  servidores em paralelo e cada um com taxa média  $\mu$ . Aqui, mostraremos três modelos com 2, 3 e 4 servidores em paralelo. Também, mostraremos que os modelos SPNs correspondentes são equivalentes aos modelos teóricos de filas. Além disso, mostraremos que, quando se tem as taxas  $\lambda$  e  $\mu$  fixas, e que quando se aumenta a quantidade de servidores em paralelos, os valores dos parâmetros da fila tendem a um limite. Isso também ocorre tanto para os modelos SPN quanto para os modelos teóricos de filas. Esse limite acontece quando aumentamos a quantidade de servidores em paralelo até tal ponto que os valores dos parâmetros tendem para um valor fixo. Neste caso, dizemos que é um sistema de fila com infinitos servidores, e este sistema é representado pela notação de Kendall-Lee como M/M/inf/FIFO/inf/inf (Fogliatti & Mattos, 2007; Taha, 2006).

Nas Figuras 8, 9 e 10 são mostrados modelos de sistemas de fila com 2, 3 e 4 servidores em paralelo, respectivamente. Suponha que para esses modelos as taxas médias sejam fixas e que seus valores sejam:  $\lambda=(1/20)$  clientes/hora e  $\mu=(1/10)$  clientes/hora. Como não há limites de clientes nesse modelo, então  $\lambda=\lambda_{cf}$ . Neste caso, os valores de tempo colocados nas transições temporizadas dos modelos são: *chegada*=20 horas e *servidor1=servidor2=servidor3=servidor4*=10 horas. Além disso, os valores dos pesos das transições imediatas são: para o modelo mostrado na Figura 8, *TII=TI2=0,5*; para o modelo mostrado na Figura 9, *TII=TI2=TI3=0,333*; e para o modelo mostrado na Figura 10, *TII=TI2=TI3=TI4=0,25*.

**Figura 9:** Modelo M/M/3/FIFO/inf/inf.

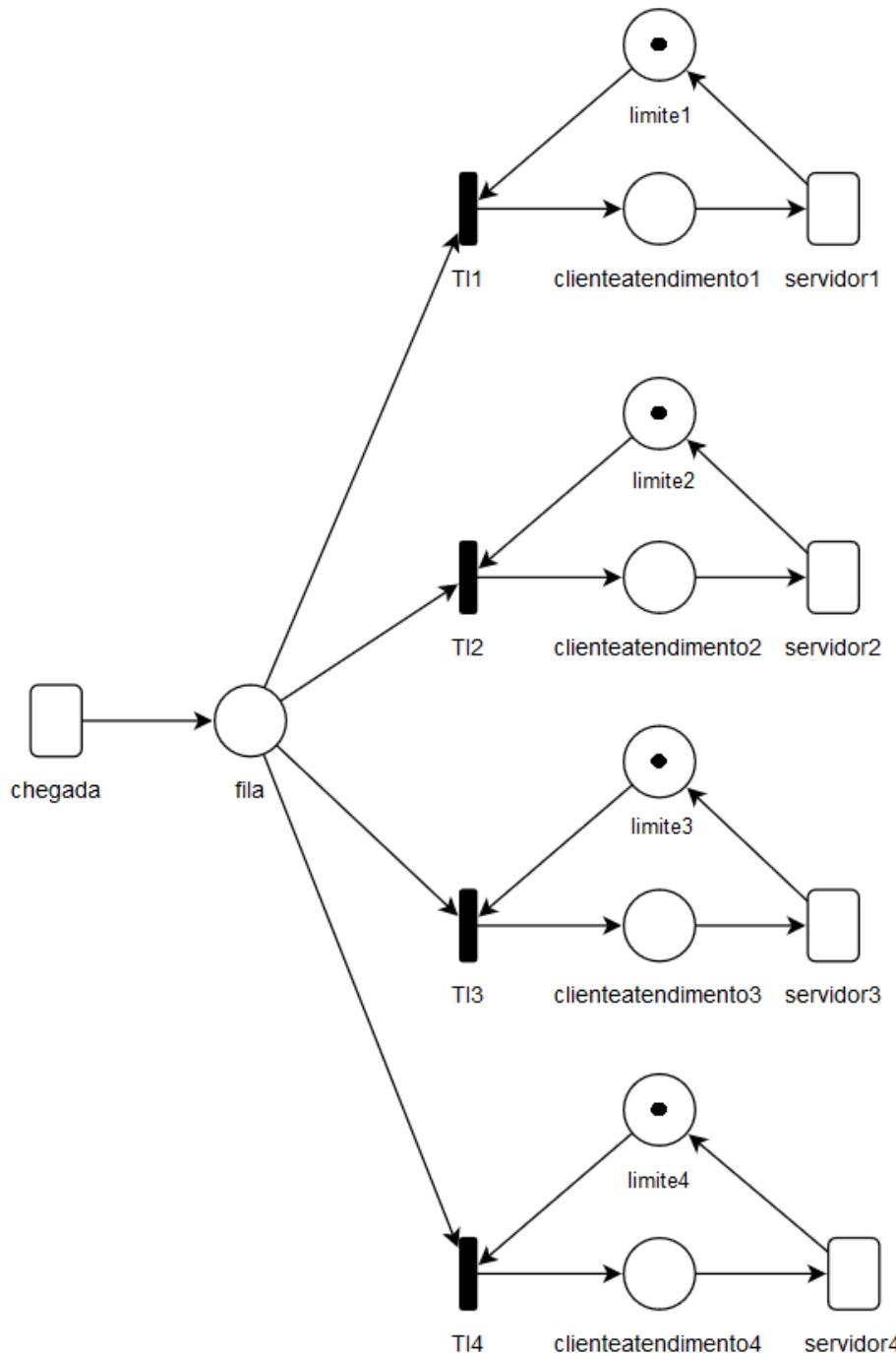


Fonte: Autores.

As métricas usadas no cálculo dos parâmetros no modelo SPN são mostradas na Tabela 4. Na Tabela 5 são mostrados os resultados dos cálculos dos parâmetros de desempenho dos modelos teóricos e dos modelos SPNs. Nessa Tabela 5, o valor médio e o valor teórico estão dentro do intervalo de confiança indicando que o modelo SPN é estatisticamente equivalente ao modelo matemático da teoria de filas. Além disso, ainda podemos observar que o modelo teórico de filas, assim como o modelo SPN, tende a estabilizar os valores de seus parâmetros de desempenho na medida em que a quantidade de servidores em paralelo aumenta. À medida que se aumenta a quantidade de servidores, mais os valores dos parâmetros de desempenho se aproximam de um limite. Tal limite é mostrado na Tabela 6, onde se encontram os valores dos parâmetros para um sistema de

fila com infinitos servidores, quando  $\lambda=(1/20)$  clientes/hora e  $\mu=(1/10)$  clientes/hora M/M/inf/FIFO/inf/inf (Fogliatti, Mattos, 2007; Taha, 2006; Bolch, *et al.*, 2006; Sztrik, 2012). Também podemos concluir que, para M/M/n/FIFO/inf/inf, seja no modelo teórico de filas ou no modelo SPN, quando as taxas  $\lambda$  e  $\mu$  são fixas, os parâmetros de desempenho do sistema de fila tendem aos valores dos parâmetros do modelo M/M/inf/FIFO/inf/inf quando o número de servidores em paralelo aumenta.

**Figura 10:** Modelo M/M/4/FIFO/inf/inf.



Fonte: Autores.

**Tabela 4:** Métricas usadas na simulação estacionária dos modelos spns desta subseção (M/M/n/FIFO/inf/inf).

Arranjo	Parâmetro	Métrica
<b>2 servidores em paralelo</b>	<b>L</b>	$((E\{\#fila\}) + (E\{\#\text{clienteatendimento1}\}) + (E\{\#\text{clienteatendimento2}\}))$
<b>3 servidores em paralelo</b>	<b>Lq</b>	$(E\{\#fila\})$
	<b>W</b>	$((E\{\#fila\}) + (E\{\#\text{clienteatendimento1}\}) + (E\{\#\text{clienteatendimento2}\})) / \lambda_{ef}$
	<b>Wq</b>	$(E\{\#fila\}) / \lambda_{ef}$
Arranjo	Parâmetro	Métrica
<b>4 servidores em paralelo</b>	<b>L</b>	$((E\{\#fila\}) + (E\{\#\text{clienteatendimento1}\}) + (E\{\#\text{clienteatendimento2}\}) + (E\{\#\text{clienteatendimento3}\}))$
<b>Lq</b>		$(E\{\#fila\})$
	<b>W</b>	$((E\{\#fila\}) + (E\{\#\text{clienteatendimento1}\}) + (E\{\#\text{clienteatendimento2}\}) + (E\{\#\text{clienteatendimento3}\})) / \lambda_{ef}$
	<b>Wq</b>	$(E\{\#fila\}) / \lambda_{ef}$
Arranjo	Parâmetro	Métrica
<b>5 servidores em paralelo</b>	<b>L</b>	$((E\{\#fila\}) + (E\{\#\text{clienteatendimento1}\}) + (E\{\#\text{clienteatendimento2}\}) + (E\{\#\text{clienteatendimento3}\}) + (E\{\#\text{clienteatendimento4}\}))$
<b>Lq</b>		$(E\{\#fila\})$
	<b>W</b>	$((E\{\#fila\}) + (E\{\#\text{clienteatendimento1}\}) + (E\{\#\text{clienteatendimento2}\}) + (E\{\#\text{clienteatendimento3}\}) + (E\{\#\text{clienteatendimento4}\})) / \lambda_{ef}$
	<b>Wq</b>	$(E\{\#fila\}) / \lambda_{ef}$

Fonte: Autores.

**Tabela 5:** Comparação do modelo teórico de filas com spn para 2,3 e 4 servidores em paralelo (M/M/n/FIFO/inf/inf).

Arranjo	Parâmetro	Teórico	SPN (média)	SPN (intervalo de confiança)
<b>2 servidores em paralelo</b>	<b>L</b>	0,533	0,5341814	[ 0,5319212 ; 0,5364416 ]
	<b>Lq</b>	0,0333	0,03358946	[ 0,03290705 ; 0,03427187 ]
	<b>W</b>	10,6667	10,6447	[ 10,6063 ; 10,68309 ]
	<b>Wq</b>	0,6667	0,6637744	[ 0,6525564 ; 0,6749924 ]
Arranjo	Parâmetro	Teórico	SPN (média)	SPN (intervalo de confiança)
<b>3 servidores em paralelo</b>	<b>L</b>	0,503	0,5022509	[ 0,5003349 ; 0,504167 ]
	<b>Lq</b>	0,003	0,00302816	[ 0,002882657 ; 0,003173674 ]
	<b>W</b>	10,06	10,01989	[ 9,987837 ; 10,05194 ]
	<b>Wq</b>	0,0606	0,05796135	[ 0,05445931 ; 0,0614634 ]
Arranjo	Parâmetro	Teórico	SPN (média)	SPN (intervalo de confiança)
<b>4 servidores em paralelo</b>	<b>L</b>	0,5002	0,500031	[ 0,4992558 ; 0,5008071 ]
	<b>Lq</b>	0,00025786	0,00025654	[ 0,0002439 ; 0,00026918 ]
	<b>W</b>	10,0051	10,00063	[ 9,9851165 ; 10,0161423 ]
	<b>Wq</b>	0,00515	0,00513088	[ 0,004878 ; 0,0053837 ]

Fonte: Autores.

**Tabela 6:** Valores dos parâmetros no modelo teórico de filas para infinitos servidores (M/M/inf/FIFO/inf/inf).

Parâmetro	Teórico
<b>L</b>	0,5
<b>Lq</b>	0,0
<b>W</b>	10,0
<b>Wq</b>	0,0

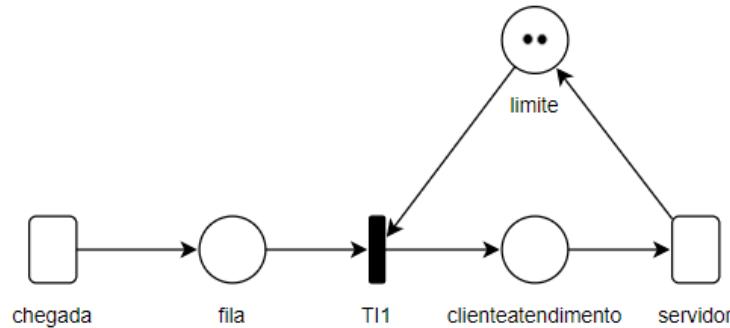
Fonte: Autores.

#### 4.3 Modelo M/M/n/FIFO/inf/inf (Equivalência)

As características e as taxas médias dos modelos de fila apresentados nesta subseção são os mesmos apresentados na Subseção 4.2, desse modo, os modelos apresentados nas Figuras 11, 12 e 13 são, respectivamente, equivalentes aos modelos apresentados nas Figuras 8, 9 e 10. Para entender como esses modelos equivalentes funcionam, vamos apresentar dois conceitos sobre as semânticas das transições temporizadas. Esses dois conceitos são o “single-server” e o “infinite-server”. Tais conceitos indicam o modo como uma transição temporizada será executada durante a simulação. O “single-server” simula a execução da transição temporizada de modo que ela não possa ser executada múltiplas vezes simultaneamente. Por outro lado, o “infinite-server” simula a execução de modo que a transição possa ser executada múltiplas vezes simultaneamente (Arteiro, *et al.*, 2007). Por exemplo, a transição *servidor* apresentada no modelo da Figura 11 tem semântica “infinite-server”.

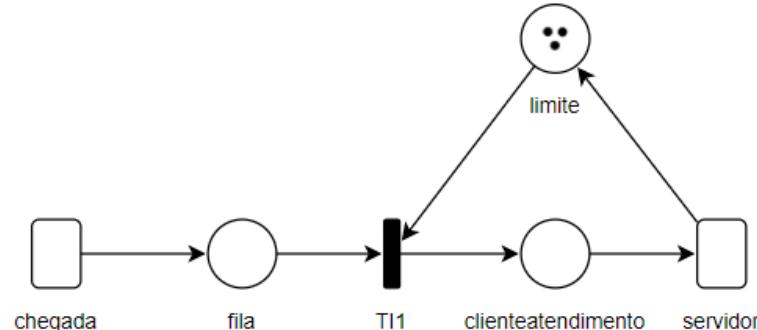
Logo, se durante a simulação, houver dois tokens no lugar *clienteatendimento*, esses tokens serão consumidos simultaneamente quando a transição *servidor* for executada. Por outro lado, se a transição *servidor* tivesse a semântica “single-server”, os tokens no lugar *clienteatendimento* não seriam consumidos simultaneamente, eles seriam consumidos um a um. Vale ser ressaltado que neste artigo apenas as transições *servidor* apresentadas nos modelos desta subseção apresentam a semântica “infinite-server”. As outras transições temporizadas dos modelos apresentados nas outras subseções usam a semântica “single-server”.

**Figura 11:** Modelo M/M/2/FIFO/inf/inf (equivalência).



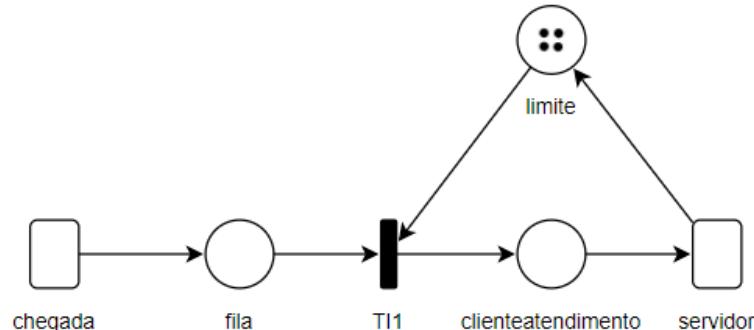
Fonte: Autores.

**Figura 12:** Modelo M/M/3/FIFO/inf/inf (equivalência).



Fonte: Autores.

**Figura 13:** Modelo M/M/4/FIFO/inf/inf (equivalência).



Fonte: Autores.

Os valores de tempo usados nas transições temporizadas dos modelos são: *chegada*=20 horas e *servidor*=10 horas. Além disso, o valor do peso da transição imediata *TII* é igual a 1. As métricas usadas na simulação são as mesmas mostradas na Tabela 2. Na Tabela 7 são mostrados os resultados dos cálculos dos parâmetros de desempenho dos modelos teóricos e dos modelos SPNs. Assim, podemos concluir que ambos os modelos teóricos e SPN são estatisticamente equivalentes.

**Tabela 7:** Comparação do modelo teórico de filas com spn para 2,3 e 4 servidores em paralelo (M/M/n/FIFO/inf/inf) – Equivalência.

Arranjo	Parâmetro	Teórico	SPN (média)	SPN (intervalo de confiança)
<b>2 servidores em paralelo</b>	<b>L</b>	0,533	0,536448	[ 0,53164317 ; 0,5412539 ]
	<b>Lq</b>	0,0333	0,0337	[ 0,0333132 ; 0,034087 ]
	<b>W</b>	10,6667	10,72897	[ 10,63286 ; 10,82507 ]
	<b>Wq</b>	0,6667	0,674002	[ 0,66626 ; 0,68174 ]
Arranjo	Parâmetro	Teórico	SPN (média)	SPN (intervalo de confiança)
<b>3 servidores em paralelo</b>	<b>L</b>	0,503	0,50298	[ 0,502118 ; 0,5038479 ]
	<b>Lq</b>	0,003	0,003093	[ 0,0029881 ; 0,003198 ]
	<b>W</b>	10,06	10,059659	[ 10,04236 ; 10,076958 ]
	<b>Wq</b>	0,0606	0,06187	[ 0,059763 ; 0,0639772 ]
Arranjo	Parâmetro	Teórico	SPN (média)	SPN (intervalo de confiança)
<b>4 servidores em paralelo</b>	<b>L</b>	0,5002	0,500937	[ 0,499741 ; 0,502133 ]
	<b>Lq</b>	0,00025786	0,00026181	[ 0,0002489 ; 0,0002747 ]
	<b>W</b>	10,0051	10,00331	[ 9,996461 ; 10,010177 ]
	<b>Wq</b>	0,00515	0,005281	[ 0,005026 ; 0,005536 ]

Fonte: Autores.

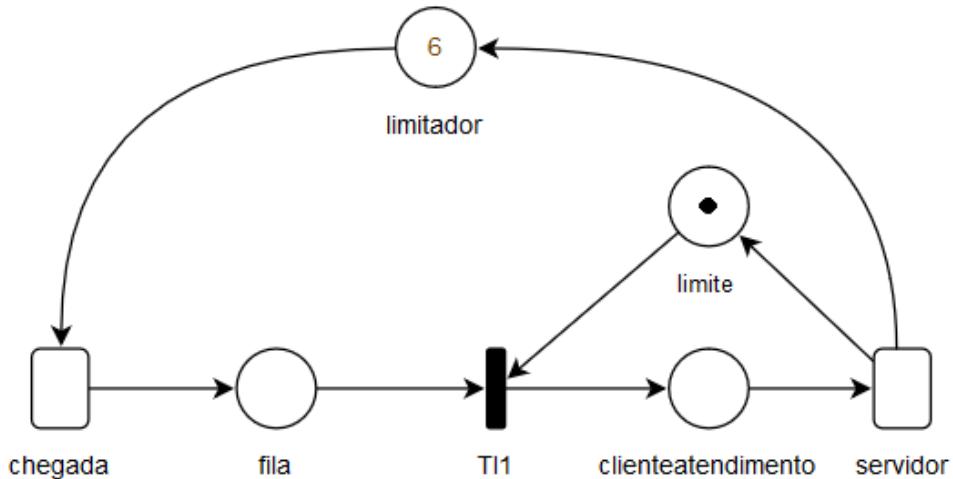
#### 4.4 Modelo M/M/1/FIFO/N/inf

Este modelo de sistema de fila apresenta uma distribuição markoviana dos tempos de chegada e saída de clientes no sistema; possui um servidor que atende um cliente por vez; a disciplina da fila é do tipo FIFO; e o número máximo de clientes no sistema é finito em N; tem população infinita (Taha, 2006). A Figura 14 apresenta o modelo de fila descrito anteriormente. Estamos supondo que existe um limitador para a quantidade de clientes no valor de N=6, e que os valores das taxas médias são:  $\lambda=0,3$  clientes/hora e  $\mu=0,5$  clientes/hora. Note que os valores de tempo associados às transições temporizadas foram o inverso das taxas médias ( $\lambda$  e  $\mu$ ): *chegada*=3,33 horas e *servidor*=2 horas. Além disso, o peso da transição *TII* é igual a 1.

No modelo de fila há um limite de 6 clientes no sistema, portanto,  $\lambda \neq \lambda_{ef}$ . O cálculo da taxa média efetiva de chegada de clientes é  $\lambda_{ef} = (P\{\#limitador > 0\}) * \lambda$ . A expressão  $P\{\#limitador > 0\}$  significa a probabilidade de haver pelo menos um token no lugar *limitador*. Isto é, é a probabilidade de que exista pelo menos uma vaga no sistema para o novo cliente (Taha, 2006; Silva, et al., 2015; Bolch, et al., 2006). Outro modo de calcular a taxa efetiva é usando a equação da taxa de utilização do servidor mostrada na Tabela 1, como  $\rho = \lambda_{ef}/(n*\mu)$ , então  $\lambda_{ef} = \rho*(n*\mu)$ . O cálculo taxa de utilização  $\rho$  é feita através da

métrica  $P\{(\#clienteatendimento > 0)\}$ . Desse modo, para o modelo M/M/1/FIFO/N/inf, como o número de servidores é  $n=1$ , então  $\lambda_{ef} = (P\{(\#clienteatendimento > 0)\}) * \mu$ .

**Figura 14:** Modelo M/M/1/FIFO/N/inf.



Fonte: Autores.

As métricas mostradas na Tabela 2 para o modelo M/M/1/FIFO/inf/inf são as mesmas usadas para o modelo desta subseção. As métricas apenas diferem no valor de  $\lambda_{ef}$ , de modo que  $\lambda_{ef} = (P\{\#limitador > 0\}) * \lambda$  (Taha, 2006; Silva, *et al.*, 2015; Bolch, *et al.*, 2006). Na Tabela 8 são mostrados os resultados dos cálculos dos parâmetros de desempenho dos modelos teóricos e na SPN. Novamente, o valor médio e o valor teórico estão dentro do intervalo de confiança indicando que o modelo SPN é estatisticamente equivalente ao modelo matemático da teoria de filas.

**Tabela 8:** Comparação do modelo teórico de filas com a spn (M/M/1/FIFO/N/inf).

Parâmetros	Teórico	SPN	
		(média)	(intervalo de confiança)
L	1,2984	1,2937	[ 1,2823755 ; 1,3051023 ]
Lq	0,7099	0,7071	[ 0,6916841 ; 0,722529 ]
W	4,4127	4,3973718	[ 4,3458209 ; 4,4489228 ]
Wq	2,4127	2,40343	[ 2,3439473 ; 2,4629186 ]

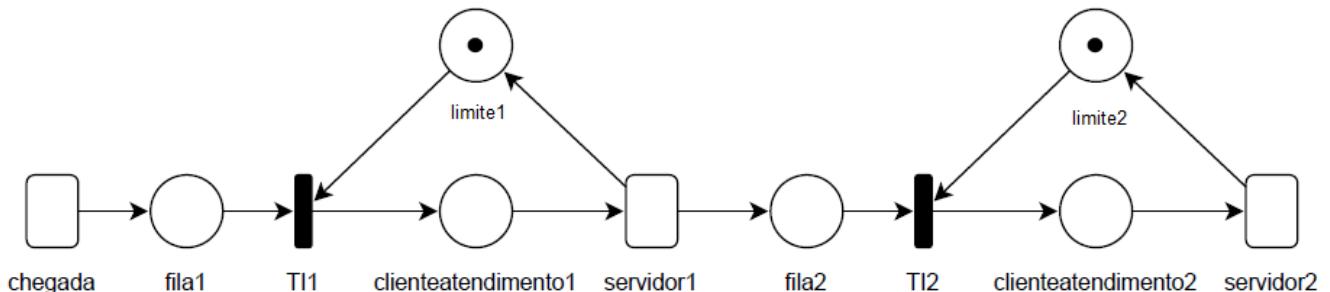
Fonte: Autores.

#### 4.5 Modelo M/M/1/FIFO/inf/inf (Série)

Sistemas de filas em série são aqueles em que um cliente é atendido numa primeira fila e, logo em seguida, ele entra em outra fila e segue sucessivamente até ser atendido na última fila (Fogliatti & Mattos, 2007; Bolch, *et al.*, 2006). A Figura 15 representa um modelo, SPN, de um sistema com duas filas M/M/1/FIFO/inf/inf em série. Assim, consideramos a taxa de chegada  $\lambda = \lambda_{ef} = (1/20)$  clientes/hora, a taxa de atendimento para o primeiro servidor  $\mu_1 = (1/10)$  clientes/hora e a taxa de atendimento para o segundo servidor  $\mu_2 = 2$  clientes/hora. Para as transições temporizadas consideramos os seguintes tempos:  $chegada = 20$  horas e  $servidor1 = 10$  horas e  $servidor2 = 0,5$  horas. O peso das transições  $TI1$  e  $TI2$  é igual a 1. As Tabelas 9 e 10 mostram, respectivamente, as métricas e a comparação dos cálculos dos parâmetros para o modelo apresentado na Figura

15. Observando os valores nas tabelas, podemos concluir que ambos os modelos teóricos e SPN são estatisticamente equivalentes.

**Figura 15:** Modelo M/M/1/FIFO/inf/inf (Série).



Fonte: Autores.

**Tabela 9:** Métricas usadas na simulação estacionária do modelo spn (M/M/1/FIFO/inf/inf) – Série.

Parâmetro	Métrica
<b>L</b>	$(E\{\#fila1\}) + (E\{\#clienteatendimento1\}) + (E\{\#fila2\}) + (E\{\#clienteatendimento2\})$
<b>Lq</b>	$(E\{\#fila1\}) + (E\{\#fila2\})$
<b>W</b>	$((E\{\#fila1\}) + (E\{\#clienteatendimento1\}) + (E\{\#fila2\}) + (E\{\#clienteatendimento2\})) / \lambda_{ef}$
<b>Wq</b>	$((E\{\#fila1\}) + (E\{\#fila2\})) / \lambda_{ef}$

Fonte: Autores.

**Tabela 10:** Comparação do modelo teórico de filas com a spn (M/M/1/FIFO/inf/inf em série).

Parâmetros	Teórico	SPN	
		(média)	(intervalo de confiança)
<b>L</b>	1,02564	1,022579	[ 1,0015715 ; 1,0435861 ]
<b>Lq</b>	0,5006	0,4993658	[ 0,48250 ; 0,5162312 ]
<b>W</b>	20,5128	20,451576	[ 20,031429 ; 20,871723 ]
<b>Wq</b>	10,0128	9,9873162	[ 9,6500075 ; 10,3246249 ]

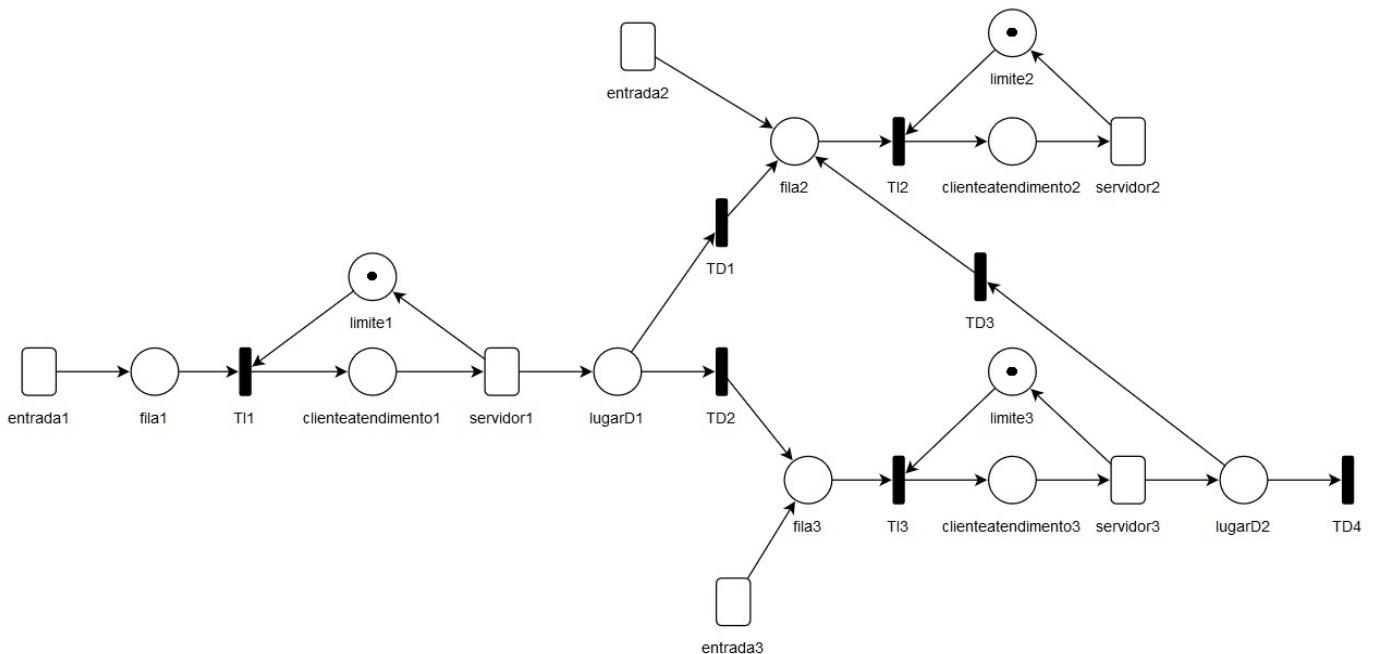
Fonte: Autores.

#### 4.6 Modelo de rede de filas aberta

Já abordamos a teoria a respeito de rede de filas na Seção 2. Agora vamos mostrar um modelo SPN de uma rede de filas, suas métricas e comparar com o modelo matemático da teoria de filas. A Figura 16 mostra uma SPN que representa a rede de filas aberta da Figura 3. Essa rede de filas é composta por três nós e cada nó tem um único servidor, além disso, ela tem três entradas com taxas  $\lambda_1 = 0,5$  clientes/hora,  $\lambda_2 = 0,25$  clientes/hora e  $\lambda_3 = 0,25$  clientes/hora. Tais entradas são representadas, respectivamente, por três transições com seus tempos associados  $entrada1 = 2$  horas,  $entrada2 = 4$  horas e  $entrada3 = 4$  horas. Os três nós da rede tem as seguintes taxas de serviços  $\mu_1 = 1$  cliente/hora,  $\mu_2 = 1$  cliente/hora e  $\mu_3 = 1$  cliente/hora. Tais serviços são representados, respectivamente, por três transições com seus tempos associados  $servidor1 = 1$

hora,  $servidor2 = 1$  hora e  $servidor3 = 1$  hora. As transições imediatas  $TI1$ ,  $TI2$  e  $TI3$  tem pesos iguais 1. Os lugares  $lugarD1$  e  $lugarD2$  são apenas um artifício usado na modelagem para representar o estado em que o cliente deve tomar uma decisão para qual centro de serviço ele deverá entrar em seguida. Note que, quando um cliente termina seu atendimento no primeiro nó, ele deverá entrar imediatamente no segundo ou no terceiro nó. Desse modo, o lugar  $lugarD1$  representa o estado em que ele irá tomar a decisão. O lugar  $lugarD1$  está conectado às transiões  $TD1$  e  $TD2$ , tais transições representam a ação da tomada de decisão por parte do cliente, seja ele ir para o segundo ou terceiro centro de serviço. A essas transições é atribuído um peso que representa a probabilidade do cliente entrar num ou em outro nó, ou seja, é a probabilidade de roteamento. Assim, conforme a Figura 3, os pesos das transições  $TD1$  e  $TD2$  devem ser, respectivamente, iguais as probabilidades de roteamento  $P13 = 0,6$  e  $P12 = 0,4$ . De modo semelhante, ocorre com as transições  $TD3$  e  $TD4$  conectadas ao lugar  $lugarD2$ . Neste caso, o cliente poderá ir para o segundo centro ou poderá sair da rede. Assim, os pesos das transições  $TD3$  e  $TD4$  são, respectivamente, iguais a  $P23 = 0,25$  e  $P20 = 0,75$ .

**Figura 16:** Modelo de rede de filas aberta.



Fonte: Autores.

Seja  $\lambda_{ef}$ ,  $L$  e  $W$ , respectivamente, a taxa média efetiva de entrada de clientes no sistema (rede de filas), a quantidade média de clientes na rede e o tempo de resposta na rede, sendo  $\lambda_{ef} = \lambda_1 + \lambda_2 + \lambda_3 = 1$  cliente/hora. Além disso, a taxa média efetiva, a quantidade média de clientes e o tempo de resposta para cada um dos três nós é, respectivamente,  $\lambda_{ef1}$ ,  $\lambda_{ef2}$  e  $\lambda_{ef3}$ ;  $L_1$ ,  $L_2$  e  $L_3$ ;  $W_1$ ,  $W_2$  e  $W_3$ . Note que, para o cálculo da taxa efetiva para cada nó, devemos usar a equação da taxa de utilização do servidor mostrada na Tabela 1, isto é,  $\rho = \lambda_{ef}/(\mu * n)$ , então  $\lambda_{ef} = \rho * (\mu * n)$ . Assim, os três nós têm as seguintes taxas efetivas, respectivamente,  $\lambda_{ef1} = (P\{(\#clienteatendimento1 > 0)\}) * \mu_1$ ,  $\lambda_{ef2} = (P\{(\#clienteatendimento2 > 0)\}) * \mu_2$  e  $\lambda_{ef3} = (P\{(\#clienteatendimento3 > 0)\}) * \mu_3$ .

As Tabelas 11 e 12 mostram, respectivamente, as métricas e a comparação dos cálculos dos parâmetros para o modelo apresentado na Figura 16. Observando os valores nas tabelas, podemos concluir que ambos os modelos teóricos e SPN são estatisticamente equivalentes.

**Tabela 11:** Métricas usadas na simulação estacionária do modelo spn (rede de filas aberta).

Parâmetro	Métrica
<b>L</b>	$(E\{\#fila1\}) + (E\{\#clienteatendimento1\}) + (E\{\#fila2\}) + (E\{\#clienteatendimento2\}) + (E\{\#fila3\}) + (E\{\#clienteatendimento3\}) + (E\{\#lugarD1\}) + (E\{\#lugarD2\})$
<b>W</b>	$((E\{\#fila1\}) + (E\{\#clienteatendimento1\}) + (E\{\#fila2\}) + (E\{\#clienteatendimento2\}) + (E\{\#fila3\}) + (E\{\#clienteatendimento3\}) + (E\{\#lugarD1\}) + (E\{\#lugarD2\})) / \lambda_{ef}$
<b>L1</b>	$(E\{\#fila1\}) + (E\{\#clienteatendimento1\})$
<b>L2</b>	$(E\{\#fila2\}) + (E\{\#clienteatendimento2\})$
<b>L3</b>	$(E\{\#fila3\}) + (E\{\#clienteatendimento3\})$
<b>W1</b>	$(E\{\#fila1\}) + (E\{\#clienteatendimento1\}) / \lambda_{ef1}$
<b>W2</b>	$(E\{\#fila2\}) + (E\{\#clienteatendimento2\}) / \lambda_{ef2}$
<b>W3</b>	$(E\{\#fila3\}) + (E\{\#clienteatendimento3\}) / \lambda_{ef3}$

Fonte: Autores.

**Tabela 12:** Comparação do modelo teórico de filas com spn (rede de filas aberta).

Parâmetros	Teórico	SPN	SPN
		(média)	(intervalo de confiança)
<b>L</b>	3,646	3,590263	[ 3,4504524 ; 3,7300736 ]
<b>W</b>	3,646	3,590263	[ 3,4504524 ; 3,7300736 ]
<b>L1</b>	1	0,9883164	[ 0,9400184 ; 1,0366144 ]
<b>L2</b>	1,424	1,398272	[ 1,3432284 ; 1,4533155 ]
<b>L3</b>	1,222	1,2036745	[ 1,1662209 ; 1,2411281 ]
<b>W1</b>	2	1,978284	[ 1,925251 ; 2,0313169 ]
<b>W2</b>	2,4238	2,4057203	[ 2,3342079 ; 2,4772326 ]
<b>W3</b>	2,221	2,22947	[ 2,165333 ; 2,2936082 ]

Fonte: Autores.

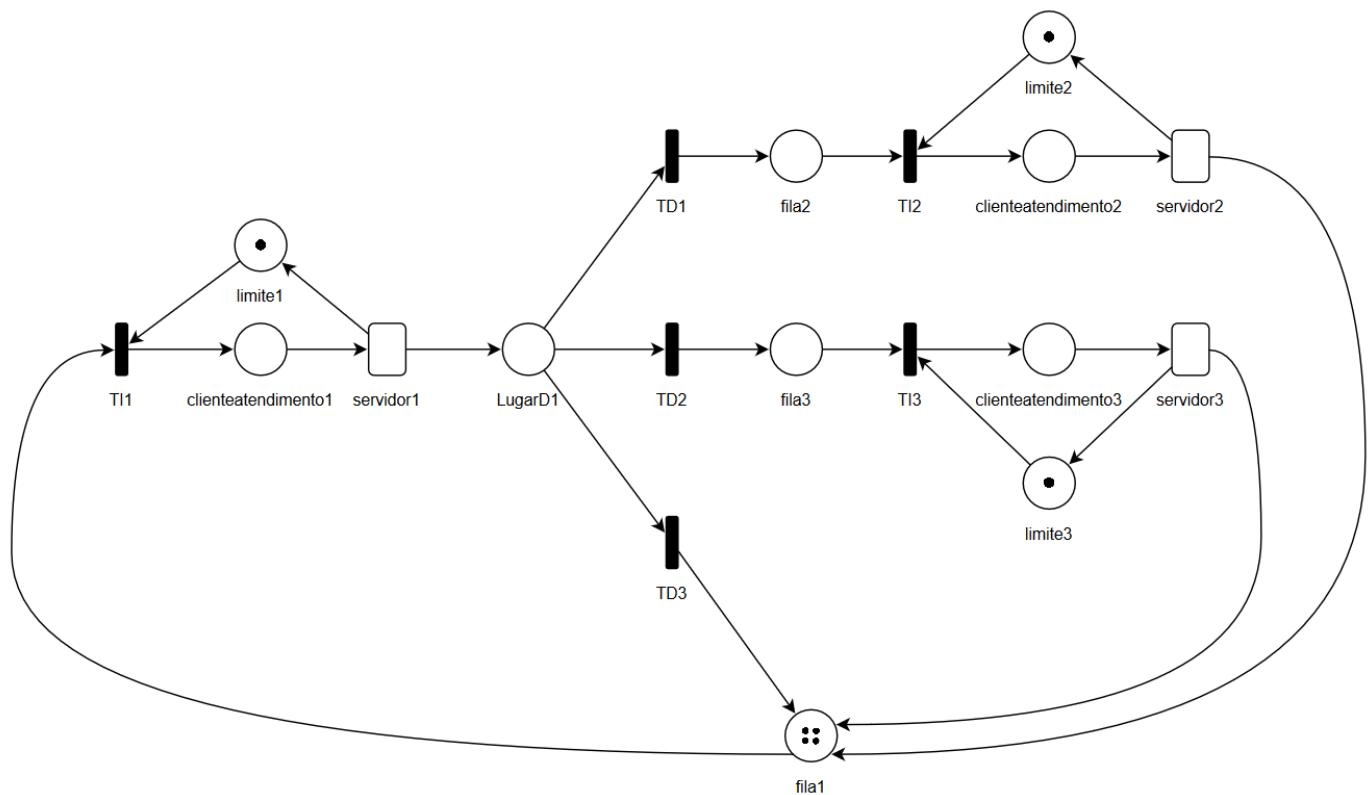
#### 4.7 Modelo de rede de filas fechada

A Figura 17 mostra uma SPN que representa a rede de filas fechada apresentada na Seção 2 (Figura 4). Essa é uma rede de Gordon-Newell e, portanto, apresenta quantidade de clientes no sistema fixo em  $K = 4$  e nenhum cliente externo ao sistema pode entrar nele. Assim, a quantidade média de clientes no sistema é  $L = 4$  e a taxa efetiva de entrada de clientes no sistema é  $\lambda_{ef} = 0$  clientes/hora, desse modo, o valor de  $W = L / \lambda_{ef}$  fica indefinido. Essa rede de filas é composta por três nós, e cada nó tem um único servidor com as taxas de serviços  $\mu_1 = 10$  cliente/hora,  $\mu_2 = 5$  clientes/hora e  $\mu_3 = 1$  cliente/hora. Esses serviços são representados, respectivamente, por três transições com seus tempos associados  $servidor1 = 0,1$  hora,  $servidor2 = 0,2$  hora e  $servidor3 = 1$  hora. As transições imediatas  $T11$ ,  $T12$  e  $T13$  tem pesos iguais 1 e os pesos das transições  $TD1$ ,  $TD2$  e  $TD3$  são, respectivamente, iguais a  $P12 = 0,6$ ,  $P13 = 0,2$  e  $P11 = 0,2$ . A taxa média efetiva, a quantidade média de

clientes e o tempo de resposta para cada um dos três nós é, respectivamente,  $\lambda_{ef1}$ ,  $\lambda_{ef2}$  e  $\lambda_{ef3}$ ; L1, L2 e L3; W1, W2 e W3. Note que, os cálculos das taxas efetivas para cada nó são:  $\lambda_{ef1} = (P\{(\#clienteatendimento1>0)\}) * \mu_1$ ;  $\lambda_{ef2} = (P\{(\#clienteatendimento2>0)\}) * \mu_2$ ;  $\lambda_{ef3} = (P\{(\#clienteatendimento3>0)\}) * \mu_3$ .

As Tabelas 13 e 14 mostram, respectivamente, as métricas e a comparação dos cálculos dos parâmetros para o modelo apresentado na Figura 17. Observando os valores nas tabelas, podemos concluir que ambos os modelos teóricos e SPN são estatisticamente equivalentes.

**Figura 17:** Modelo de rede de filas fechada.



Fonte: Autores.

**Tabela 13:** Métricas usadas na simulação estacionária do modelo spn (rede de filas fechada).

Parâmetro	Métrica
<b>L1</b>	$(E\{\#fila1\}) + (E\{\#clienteatendimento1\})$
<b>L2</b>	$(E\{\#fila2\}) + (E\{\#clienteatendimento2\})$
<b>L3</b>	$(E\{\#fila3\}) + (E\{\#clienteatendimento3\})$
<b>W1</b>	$(E\{\#fila1\}) + (E\{\#clienteatendimento1\}) / \lambda_{ef1}$
<b>W2</b>	$(E\{\#fila2\}) + (E\{\#clienteatendimento2\}) / \lambda_{ef2}$
<b>W3</b>	$(E\{\#fila3\}) + (E\{\#clienteatendimento3\}) / \lambda_{ef3}$

Fonte: Autores.

**Tabela 14:** Comparação do modelo teórico de filas com spn (rede de filas fechada).

<b>Parâmetros</b>	<b>Teórico</b>	<b>SPN</b>	<b>SPN</b>
		<b>(média)</b>	<b>(intervalo de confiança)</b>
<b>L1</b>	0,7038	0,6912616	[ 0,66935 ; 0,7131732 ]
<b>L2</b>	0,9347	0,9291622	[ 0,9083043 ; 0,9500201 ]
<b>L3</b>	2,3615	2,379576	[ 2,3368065 ; 2,4223455 ]
<b>W1</b>	0,1585	0,1580261	[ 0,1567419 ; 0,1593103 ]
<b>W2</b>	0,3509	0,3503456	[ 0,3493116 ; 0,3513797 ]
<b>W3</b>	2,6599	2,6746584	[ 2,6315147 ; 2,717802 ]

Fonte: Autores.

## 5. Conclusão

Sistemas de filas são bem comuns em nossa sociedade. Tais sistemas podem aparecer de modo direto como em uma fila de um banco, ou podem aparecer de modo indireto como em um sistema computacional de armazenamento de dados. A importância do estudo da teoria de filas reside no fato de nós necessitarmos prever demandas futuras e/ou analisarmos o desempenho de um determinado sistema de modo que possamos encontrar gargalos e propor melhorias. No entanto, a teoria de filas não pode ser usada para modelar certos comportamentos específicos, como dependências entre componentes ou comportamentos de falhas e reparos. As redes de Petri, por outro lado, são um formalismo matemático que permite suprir tais limitações. Consequentemente, entender as redes de Petri não só nos habilita a estudar os sistemas de filas como também sistemas mais complexos. Podemos exemplificar esses sistemas complexos como, sistema ciber-físicos, sistemas de telecomunicações, sistemas distribuídos e sistemas mecatrônicos.

Assim, neste trabalho mostramos, analisamos e comparamos modelos de sistemas de filas teóricos com os resultados obtidos através de modelos SPNs. Os resultados mostraram que a modelagem em SPN pode representar os sistemas de filas de forma acurada. Advogamos que as SPNs são um excelente meio de auxílio à modelagem de sistemas complexos. Adicionalmente, esperamos que este tutorial sirva como um guia para pesquisadores e profissionais modelarem sistemas complexos com base nos exemplos apresentados. Como trabalhos futuros, iremos detalhar as equações matemáticas dos modelos de fila e também iremos comparar os resultados dos modelos com medições em sistemas reais.

## Referências

- Arteiro, R. & Souza, F. & Rosa, N. & Maciel, P. (2007). Utilizando redes de Petri para modelagem de desempenho de middleware orientado a mensagem. *Anais do XXVII Congresso da SBC - SBC*, Sociedade Brasileira de Computação, 595-614.
- Bassanezi, R. C. (2002). *Ensino-aprendizagem com modelagem matemática: uma nova estratégia*. Contexto.
- Balbo, G. (2001). Introduction to Stochastic Petri Nets. *Lectures on Formal Methods and Performance Analysis*, 2090 of LNCS, 88-155.
- Blätke, M. A. & Heiner, M. & M, W. (2011). Tutorial - Petri Nets in Systems Biology. *Technical Report*. 10.13140/2.1.3796.6402.
- G. Bolch, S. Greiner, H. de M., & Trivedi. K. S. (2006). *Queueing Networks and Markov Chains*. (2a ed.), Wiley-Interscience.
- Chung, C. (2019). *Simulation Modeling Handbook: A Practical Approach*. CRC Press.
- Dubois G. (2018). *Modeling and Simulation*. CRC Press-Taylor & Francis.
- Doo-Kwon B. (2005). *Systems modeling and simulation: theory and applications - third Asian Simulation Conference, AsiaSim 2004, Jeju Island, Korea, October, 2004*. Springer.
- Fogliatti, M. C.; Mattos, N. M. C. (2007). *Teoria de Filas*. Editora Interciência. .
- Havelange, S.& Rêgo, R. (2015). *O Concreto e o Abstrato no Ensino de Matemática*. 10.13140/RG.2.1.1385.7760.

- Hermann, W., Juvanelli, C., & Coqueiro, V. dos S. (2020). Overview of publications on Mathematical Modeling in four journals of the teaching area. *Research, Society and Development*, 9(8), e73985139. <https://doi.org/10.33448/rsd-v9i8.5139>.
- Hillier, F. Lieberman, G. (2013). *Introdução à pesquisa operacional*, (9a ed.), AMGH.
- Kleijnen, Jack. (2007). *Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice*, 173-223. 10.1002/9780470172445.
- Kounev, S., & Buchmann, A. (2003). *Performance modelling of distributed e-business applications using Queuing Petri Nets*, 143- 155. 10.1109/ISPASS.2003.1190241.
- Maciel, P. et al. (2017). Mercury: Performance and Dependability Evaluation of Systems with Exponential, Expolynomial and General Distributions. In: *The 22nd IEEE Pacific Rim International Symposium on Dependable Computing (PRDC 2017)*. January 22-25. Christchurch, New Zealand.
- Murata, T. (1989). Petri nets: Properties, analysis and applications. *Proceedings of the IEEE*, 77(4), 541-580.
- Silva, B. et al. (2015). Mercury: An Integrated Environment for Performance and Dependability Evaluation of General Systems. In: *Proceedings of Industrial Track at 45th Dependable Systems and Networks Conference (DSN-2015)*.
- Sztrik, J. (2012). *Basic Queuing Theory*, University of Debrecen, Faculty of Informatics. Hungary.
- Taha, H. A. (2006). *Operations research: an introduction*, (8th ed.), Macmillan Publishing Co., Inc. Indianapolis, IN, USA.
- Trivedi, K. S. (2008). *Probability & Statistics with Reliability, Queuing and Computer Science Applications*. John Wiley & Sons.
- Viali, L. (2006). Contribuições para o ensino da distribuição normal ou curva de Gauss em cursos de Graduação. In: *III Seminário Internacional de Pesquisa em Educação Matemática*. São Paulo, SP, Brazil.
- Zimmermann, A. (2014). Reliability Modelling and Evaluation of Dynamic Systems with Stochastic Petri Nets (Tutorial). *VALUETOOLS 2013 - 7th International Conference on Performance Evaluation Methodologies and Tools*. 10.4108/icst.valuetools.2013.254370.
- Zurawski, R. & Zhou, M. (1995). Petri net and industrial application: A tutorial. *IEEE Transactions on Industrial Electronics*. 41. 567 - 583. 10.1109/41.334574.