PCS-2039
Modelagem e Simulação de Sistemas Computacionais

Graça Bressan

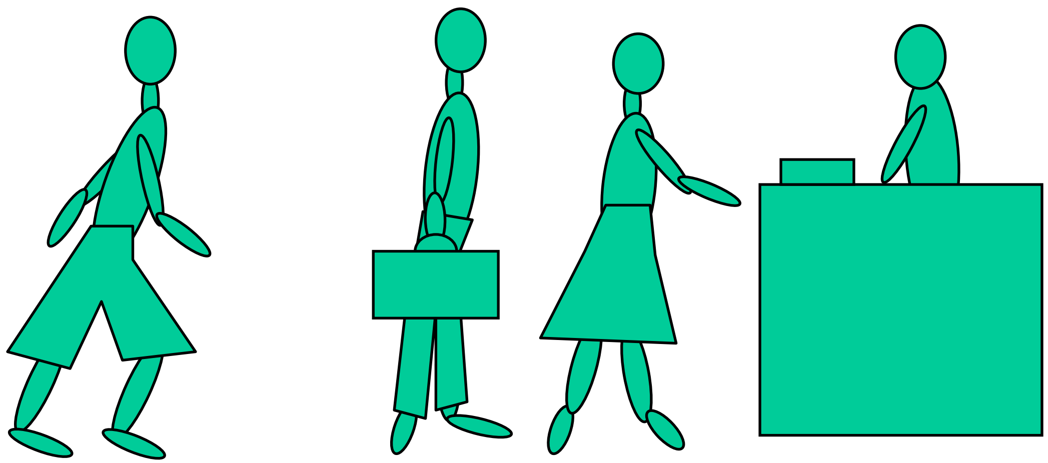
gbressan@larc.usp.br

Sistemas de Filas Simples

Graça Bressan
LARC-PCS/EPUSP

Sistemas de Filas

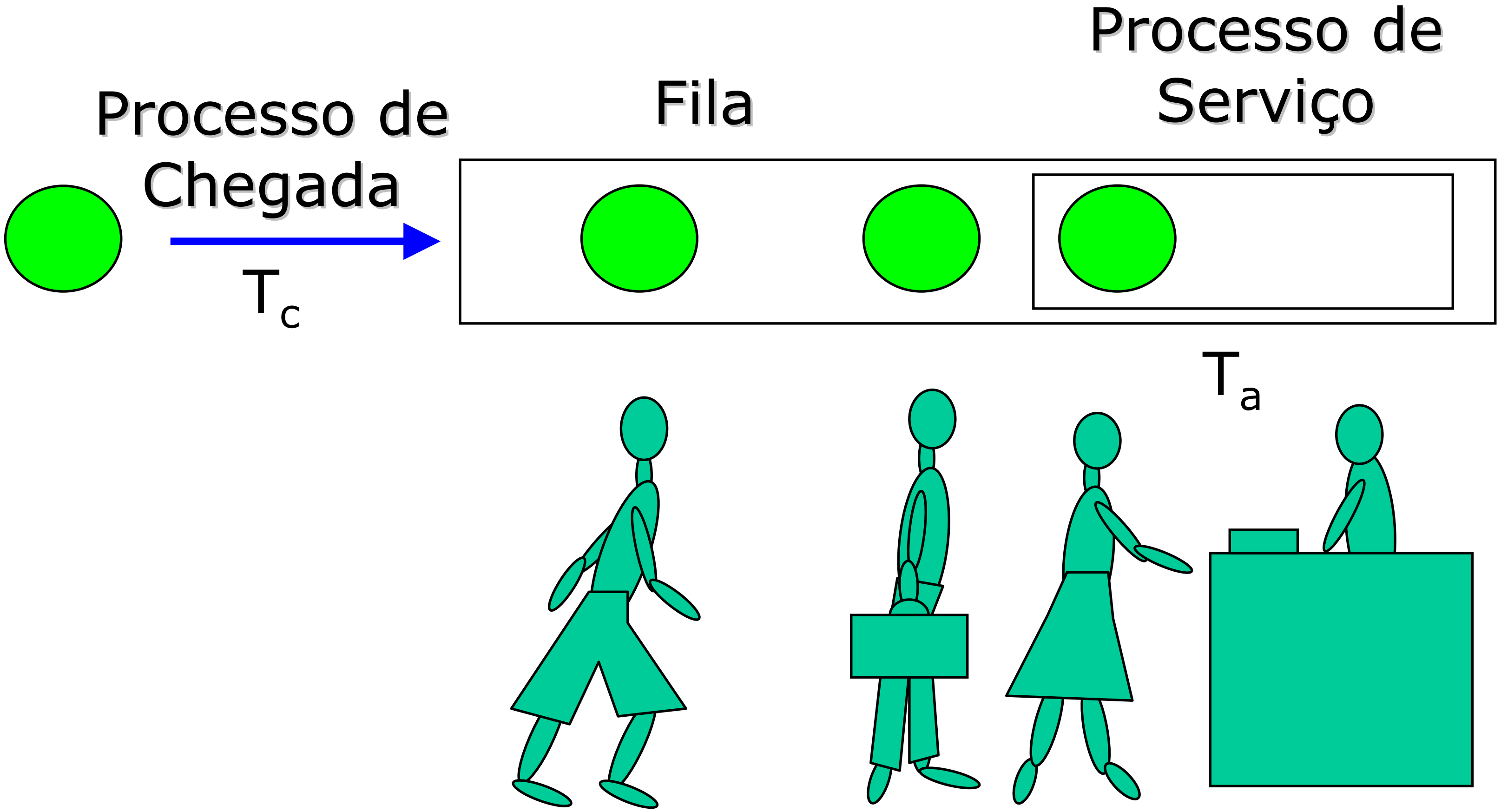
- A maior parte dos sistemas de serviços tais como bancos, supermercados, postos de gasolina, lanchonetes, sistemas de comunicação e sistemas computacionais podem ser modelados por sistemas de filas.



Teoria de Filas

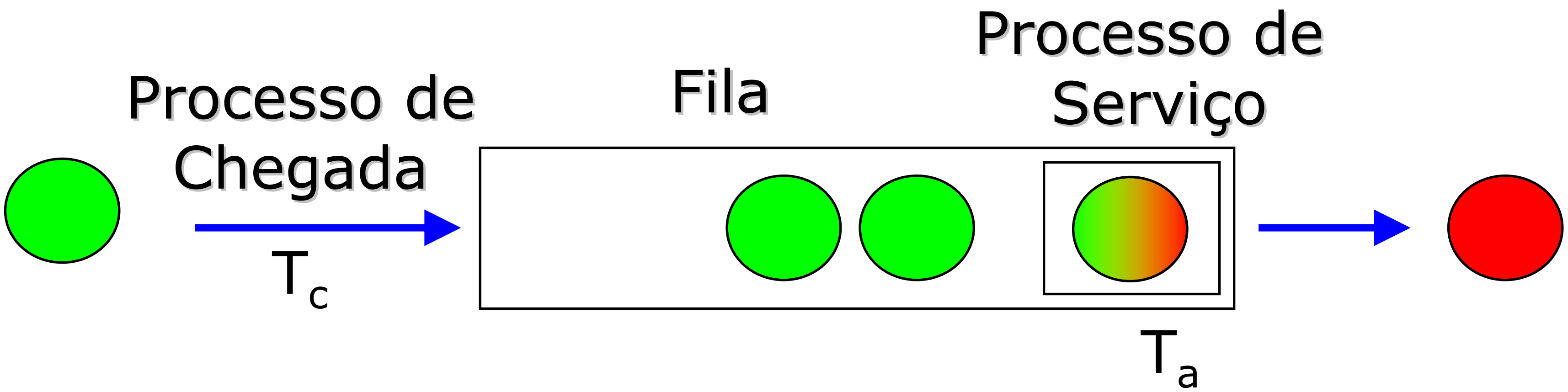
- Neste módulo será apresentada a teoria de filas que é um ramo da área de processos estocásticos.
- Esta teoria permitirá a obtenção métodos analíticos para avaliação desempenho de sistemas de filas simples como alternativa aos modelos de simulação.

Sistemas de Filas Simples



Sistemas de Filas Simples

- As filas simples são definidas por dois parâmetros principais:
  - T_c - Intervalo entre chegadas que é parâmetro de carga
  - T_a - Tempo de atendimento que é parâmetro de sistema
- Os parâmetros T_c e T_a em geral são variáveis aleatórias descritas por distribuições de probabilidade.



Sistemas de Filas Simples

- **Processo de chegada:** é a seqüência de variáveis aleatórias que definem os intervalos entre chegadas;
- **Processo de atendimento ou serviço:** é a seqüência de variáveis aleatórias que definem os tempos de atendimento.
- As variáveis aleatórias formam uma seqüência de valores aleatórios Independentes e Identicamente Distribuídos (IID).

Parâmetros de Sistemas de Filas Simples

- **Número de servidores**: define o número de servidores disponíveis no sistema de fila. Normalmente, estes servidores são idênticos.
- **Capacidade do sistema**: é o número máximo de usuários que o sistema de fila pode apresentar. Este número considera tanto os usuários na fila como aqueles em serviço. Quando este parâmetro não representa nenhuma grande limitação é comum utilizar-se um valor infinito para a sua capacidade.

Parâmetros de Sistemas de Filas Simples

- **Tamanho da população:** representa o número total de usuários em potencial, que podem chegar no sistema de fila. Na maioria dos sistemas reais a população é finita, porém se esse número é suficientemente grande, pode-se utilizar o valor infinito como tamanho da população.

Parâmetros de Sistemas de Filas Simples

- **Disciplina de Serviço:** a ordem com que os usuários do sistema são tratados define a disciplina de serviço ou atendimento. A disciplina mais comum é a **FCFS** (Primeiro a Chegar- Primeiro a ser atendido). Outras disciplinas podem ser aplicadas, tais como: **LCFS**, **LCFS-preemptivo**, **RR** (Round Robin) e **PS** (Processor Sharing).

Parâmetros de Sistemas de Filas Simples

- A **disciplina RR** com um quantum suficientemente pequeno comparado com o tempo médio de serviço é chamado de **PS**, pois ele reparte o processador em **n** partes iguais.
- Sistemas de fila que não possuem tempo de espera são chamados de **Centros de Atraso** (Delay Center). Em geral, sistemas com infinitos servidores possuem esta característica;

Notação

- Indica-se os sistemas de fila simples na notação:

  M/M/1 –chegada e atendimento exponenciais e 1 servidor.

  M/M/m - chegada e atendimento exponenciais e m servidores.

  M/M/m/k -chegada e atendimento exponenciais, m servidores e k espaços na fila.

  M/G/1 - chegada exponencial e atendimento genérico e 1 servidor.
- As distribuições dos intervalos entre chegadas e dos tempos de serviço são representadas em geral por uma letra:

M	Exponencial;
D	Determinística
G	Genérica

Determinística e Genérica

- Uma **distribuição determinística** é aquela que define tempos constantes, portanto não existe nenhuma variabilidade.
- Uma **distribuição genérica** significa uma distribuição não especificada. Os resultados assim obtidos são válidos para qualquer distribuição.

Modelo de fila simples

- Uma distribuição muito utilizada nos modelos de fila é a **exponencial** cujo parâmetro **1/λ** (ou **1/μ**) pode representar o intervalo médio de chegada (ou tempo médio de atendimento).
- Sistemas em que a chegada e o tempo de serviço ambos têm distribuição exponencial são denominados **M/M/1** (M de Markoviano).

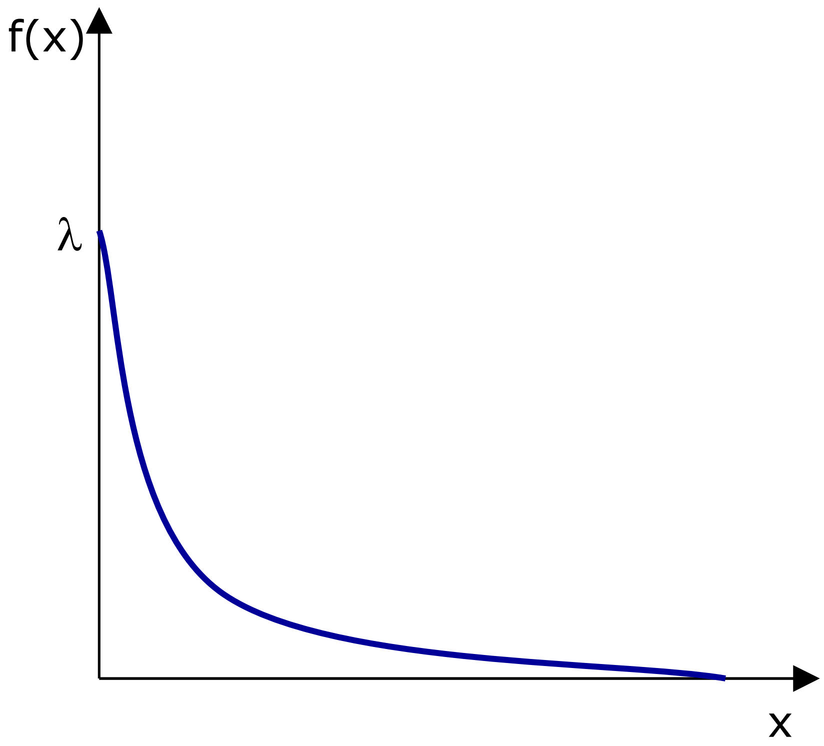
Distribuição Exponencial

Função densidade:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases}$$

Função distribuição:

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases}$$

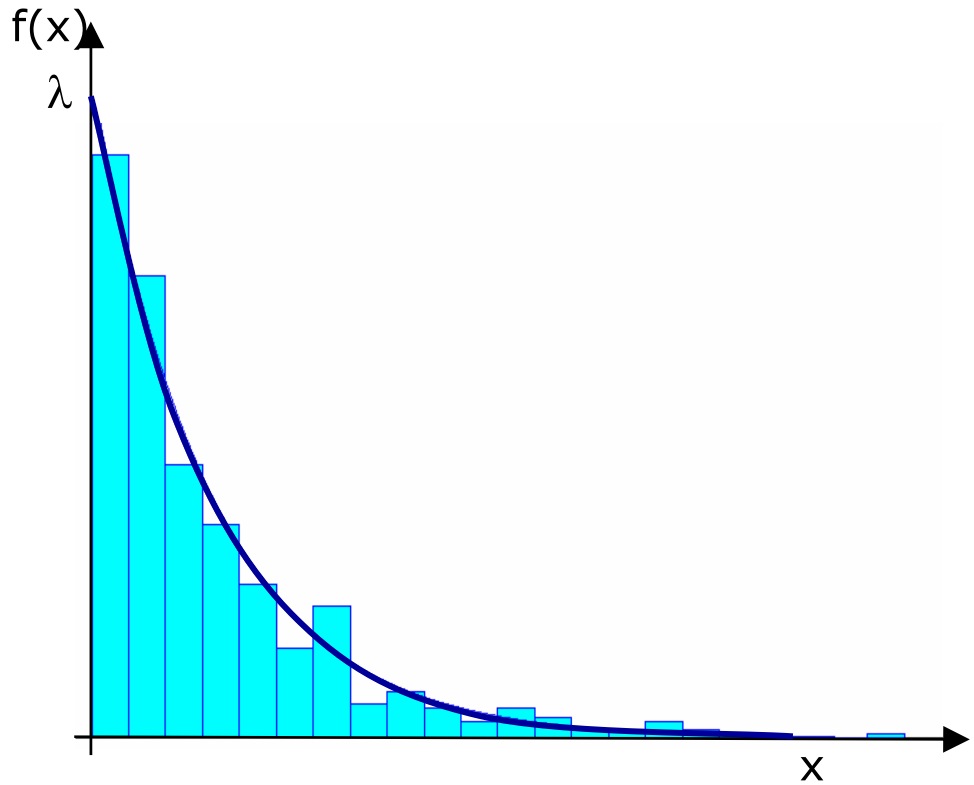


Média: $1/\lambda$

Variância: $1/\lambda^2$

Distribuição Exponencial - Histograma

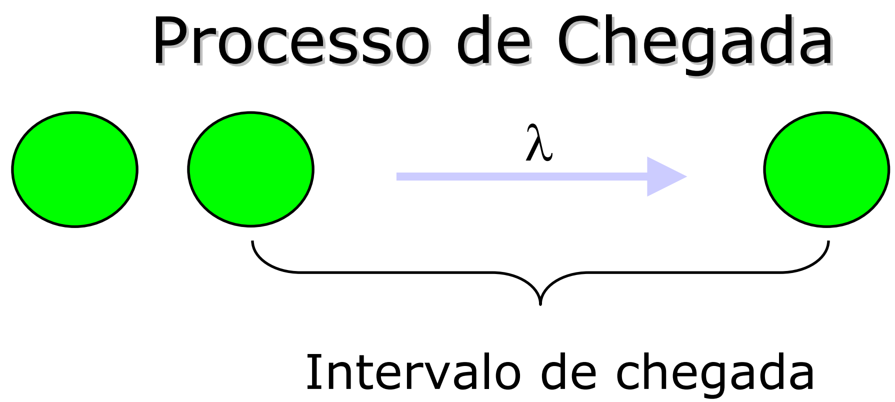
- A distribuição pode ser obtida de dados amostrados em situações reais, através da construção de histogramas de freqüências e ajuste de curvas ao histograma (Best-fit).



Distribuição Exponencial

- O parâmetro λ é a taxa média de chegada (chegadas por unidade de tempo) e $1/\lambda$ é o intervalo médio de tempo entre chegadas.
- A **distribuição exponencial** possui a propriedade de não apresentar memória (“memoryless”), isto é, sendo $1/\lambda$ o tempo médio entre chegadas, o tempo esperado para a próxima chegada é sempre $1/\lambda$, independente do tempo que já transcorreu desde a última chegada.

Processo de Poisson x Distribuição Exponencial



- Seja um processo de chegada em que os intervalos de chegada são variáveis aleatórias IID (Independentes e Identicamente Distribuídas) que obedecem distribuição exponencial com intervalo médio de chegada $1/\lambda$.
- Neste caso, o número de elementos que chegam por unidade de tempo tem distribuição (discreta) de Poisson com taxa de chegada λ.

Parâmetros do sistema de filas:

Parâmetro	Significado
λ	Taxa média de chegada
$1/\lambda$	Intervalo médio de chegada
μ	Taxa média de serviço por servidor
$1/\mu$	Tempo médio de serviço de um usuário
m	Número de servidores

Variáveis do sistema de filas

Variável	Significado
n	Número de usuários no sistema. Inclui os que estão na fila e sendo atendidos no servidor.
nq	Número de usuários esperando para serem atendidos. É sempre menor que n, pois não inclui os usuários em serviço.
ns	Número de usuários em serviço
E[n]	Número médio de usuários no sistema
E[nq]	Número médio de usuários na fila sem serem atendidos
E[ns]	Número médio de usuários sendo atendidos ($\leq$m)

Variáveis do sistema de filas

Variável	Significado
r	Tempo de resposta ou simplesmente tempo no sistema. Inclui tanto o tempo de espera como o tempo de serviço.
w	Tempo de espera, isto é, intervalo de tempo entre o instante de chegada e o início do serviço
s	Tempo de serviço
E[r]	Tempo médio de resposta
E[w]	Tempo médio dos usuários na fila (sem serem atendidos)
$E[s] = 1/\mu$	Tempo médio de serviço

Sistema Balanceado

- Diz-se que um sistema é **balanceado** se usuários não são perdidos no sistema, isto é, todos que entraram no sistema saíram em algum momento.
- Este conceito não depende da distribuição da chegada ou do atendimento.
- De forma aproximada, o sistema será balanceado se dado um intervalo de tempo **T** suficientemente grande, o número de chegadas se aproxima do número de partidas.
- Sistemas **balanceados** equivalem a sistemas **estáveis**.

Fórmula de Little

- A **Fórmula de Little** exige que o sistema seja **balanceado** e é válida independentemente das distribuições de chegada e atendimento.
- Dados:

  T - um intervalo de tempo suficientemente grande;

  λ - a taxa média de chegada neste período;

  E[n] - número médio de usuários no sistema neste período;

  E[r] - tempo médio de resposta (entre a chegada e a saída) neste período.
- A **Fórmula de Little** relaciona o número de usuários com o tempo de resposta.

$$E[n] = \lambda * E[r]$$

Fórmula de Little

- Além disso, sendo, neste período:

  E[nq] - número médio de usuários no na fila (esperando para serem atendidos neste período;

  E[w] - tempo médio de espera neste período.

- A **Fórmula de Little** pode ser aplicada resultando em:

$$E[nq] = \lambda * E[w]$$

Exemplo 1: Atendimento em caixa automático de banco (ATM)

- Em um período de 30 minutos chegaram 6 clientes para utilizar um caixa automático de um banco (ATM). Observou-se que entre a chegada à fila e a saída, cada cliente levou em média 2 minutos.
- Pergunta: Quantos clientes existem em média na fila do ATM, incluindo o que está sendo atendido?

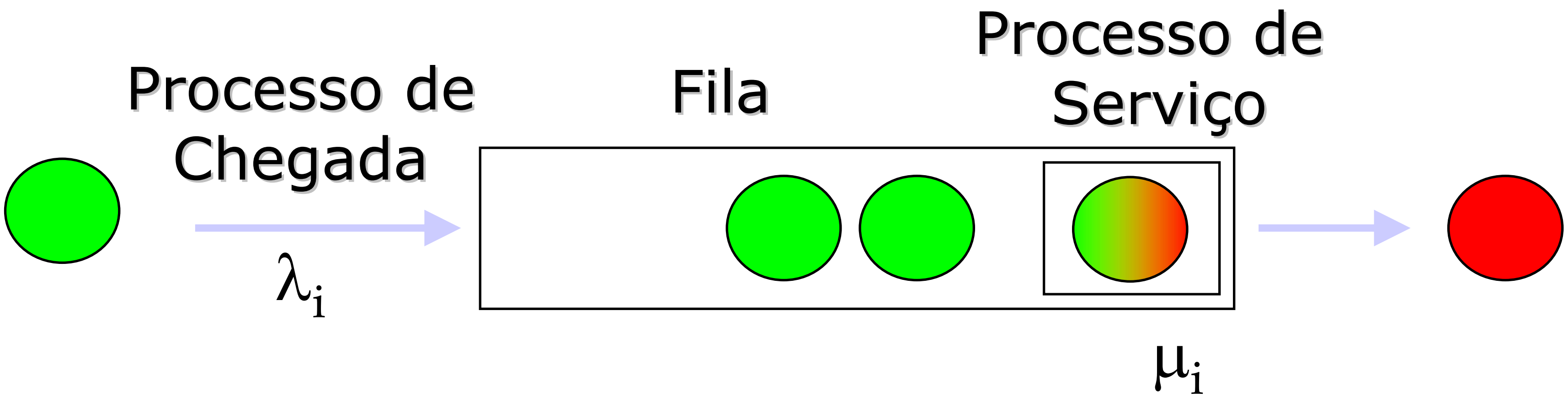
  Taxa de chegada λ = 6/30 clientes/minutos

  Tempo médio de resposta E[r] = 2 minutos
- Utilizando-se a fórmula de Little **E[n] = λ*E[r]** calcula-se o número médio de usuários na ATM:

  E[n] = (6/30) * 2 = 0,4 usuários

Sistemas de Fila Única

- Estes sistemas são modelados como **Processos de Markov de Nascimento e Morte.**



- Sendo:
  - λ_i - Taxa média de chegada quando o sistema possui i usuários.
  - μ_i - Taxa média de atendimento quando o sistema possui i usuários.
- Esta é a situação geral em que λ_i e μ_i podem variar conforme o estado do sistema.

Exemplo 2: Fila única de supermercado

- Um pequeno supermercado possui uma única caixa sendo que os clientes chegam à fila do caixa em intervalos de tempo cuja média é 15 minutos com distribuição exponencial. O tempo médio de atendimento de um cliente pelo caixa é de 10 minutos também com distribuição exponencial.

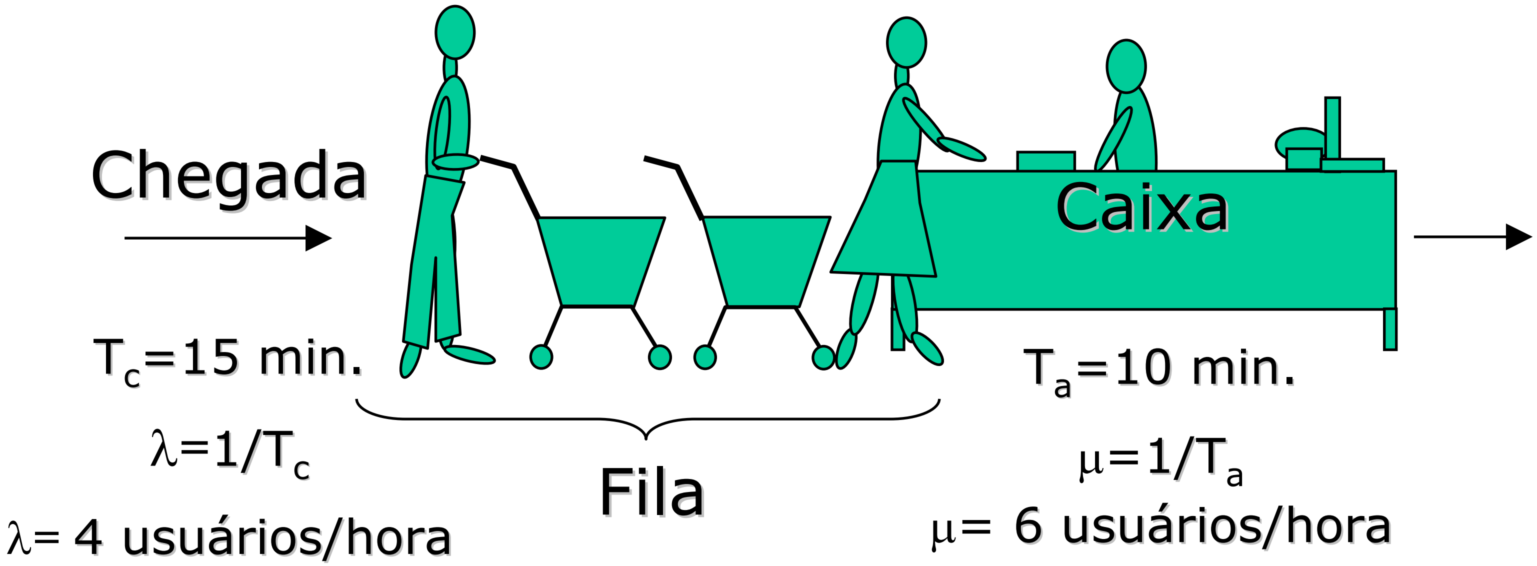


Diagrama de Transição de Estados

- Inicialmente o sistema está vazio. A cada chegada, o sistema muda para o próximo estado e quando for completado um serviço o sistema volta para o estado anterior.

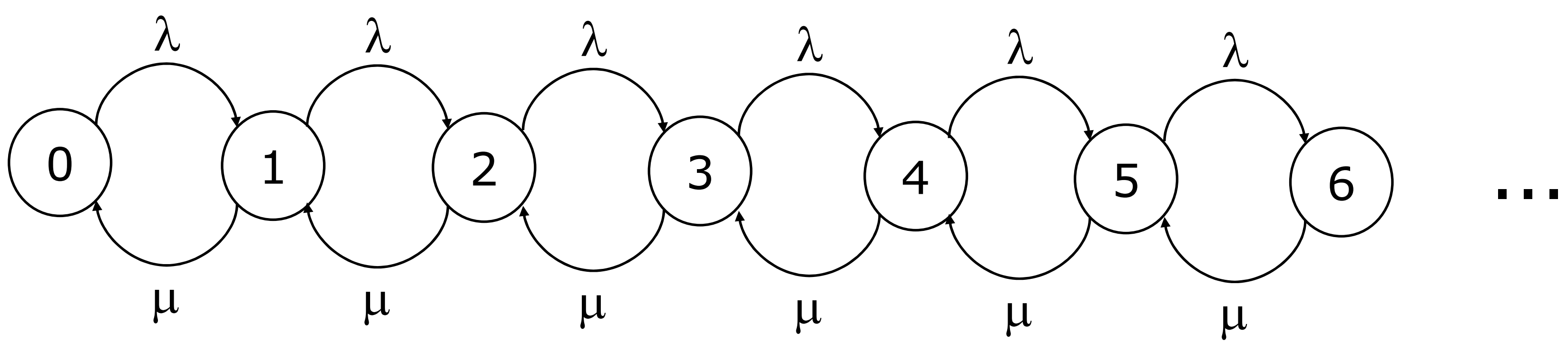


Diagrama de transição de estados

Sistemas de Fila Única M/M/1

- Nos **Processo nascimento e morte** o estado deste sistema indica o número de usuários no sistema.

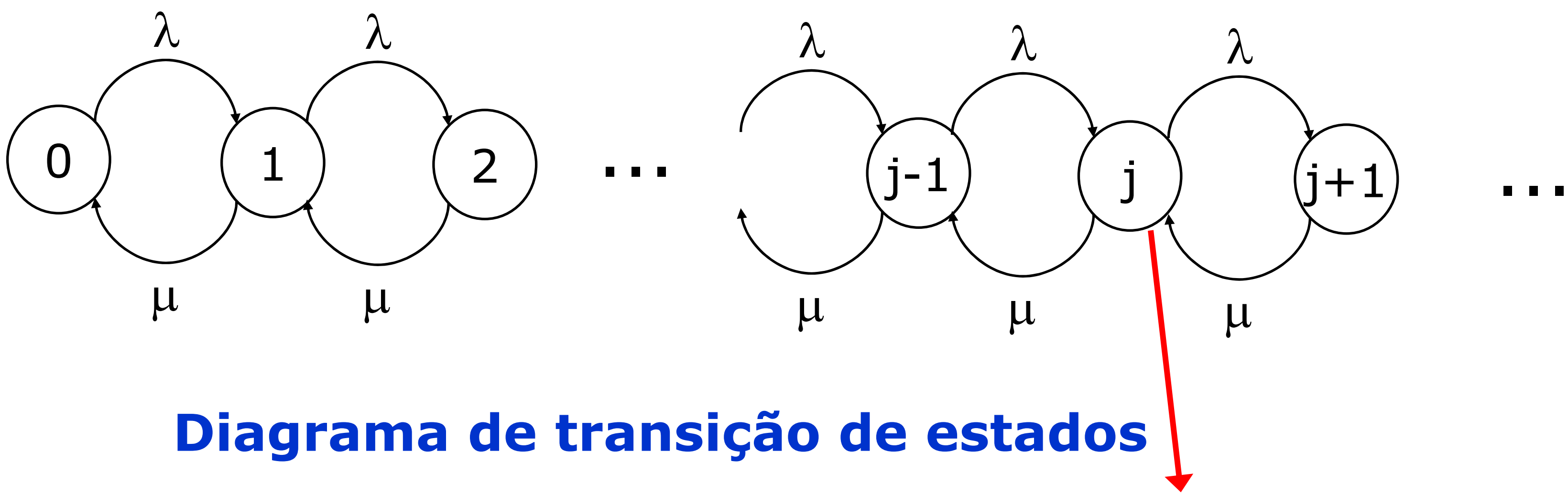


Diagrama de transição de estados

- Quando o sistema está no estado j, existem j usuários no sistema.

Exemplo 1: Fila única de supermercado

1. Para este sistema tem-se os parâmetros básicos:

$\lambda = 1/T_c$ onde T_c = 15 minutos, logo
λ = 1/15 = 4 clientes/hora

$\mu = 1/T_a$ onde T_a = 10 minutos, logo
μ = 1/10 = 6 clientes/hora

Define-se o parâmetro $\rho = \lambda / \mu$, sendo que ρ deve ser menor que 1 ($\rho < 1$) para que o sistema seja estável (balanceado).

Neste caso, ρ = 4/6 = 1/3

Probabilidades em Equilíbrio do Sistema M/M/1

- Probabilidade p_0 do sistema possuir zero usuários

$$p_0 = 1 - \lambda / \mu \qquad \text{ou} \qquad p_0 = 1 - \rho$$

- Probabilidade P_n de existirem n usuários no sistema (n = 1,2,3,...) é:

$$p_n = p_0 (\frac{\lambda}{\mu})^n \qquad \text{ou} \qquad p_n = p_0 \rho^n$$

- Soma de todas probabilidades :

$$\sum_{n=0}^{\infty} p_n = 1$$

Exemplo 1: Fila única de supermercado

2. Probabilidade de não existirem clientes no caixa, isto é, probabilidade de 0 clientes no sistema :

$$p_0 = 1 - \lambda / \mu$$

$$p_0 = 1 - 4/6 = 1 - 2/3 = 1/3 = 33{,}33\%$$

Exemplo 1: Fila única de supermercado

3. Probabilidade de n clientes no caixa, isto é, probabilidade de estar no estado n:

$$p_n = p_0 (\frac{\lambda}{\mu})^n$$

Por exemplo para saber a probabilidade de existirem 4 clientes no caixa basta calcular

$p_4 = p_0 * (\lambda / \mu)^4$

$p_4 = 1/3 * (2/3)^4 = 0{,}0658$ isto é, 6,58 %

Exemplo 1: Fila única de supermercado

4. O fator de utilização **U** do caixa é a porcentagem do tempo que o caixa ficou ocupado:

$$U = 1 - P_0$$

No exemplo,

$$U = 1 - P_0 = 1 - 1/3 = 0{,}6666$$

Logo a utilização do sistema é de 66,66 %

Exemplo 1: Fila única de supermercado

5. Número médio de clientes no sistema:

$$E[n] = \rho / (1-\rho)$$

$$E[n] = (2/3)/[1-(2/3)] = 2$$

6. Variância do número de clientes no sistema:

$$Var[n] = \rho / (1-\rho)^2$$

$$Var[n] = (2/3)/[1-(2/3)]^2 = 6$$

Exemplo 1: Fila única de supermercado

5. Probabilidade de se ter n ou mais usuários no sistema:

   $P_{\geq n} = \rho^{n} = (10/15)^{n}$

   No exemplo, a probabilidade de ter 4 ou mais clientes na fila é

   $p_{\geq 4} = (10/15)^{4} = 0{,}1975$ isto é, 19,75 %

6. Tempo médio de resposta: E[r]

   $E[n] = \lambda * E[r)$ Fórmula de Little

   então $E[r] = E[n]/\lambda$

   $E[r] = 2 / (1/15) = 30$ minutos

Sistemas de Fila Única

- No caso geral de **Processo nascimento e morte** o estado deste sistema indica o número de usuários no sistema.

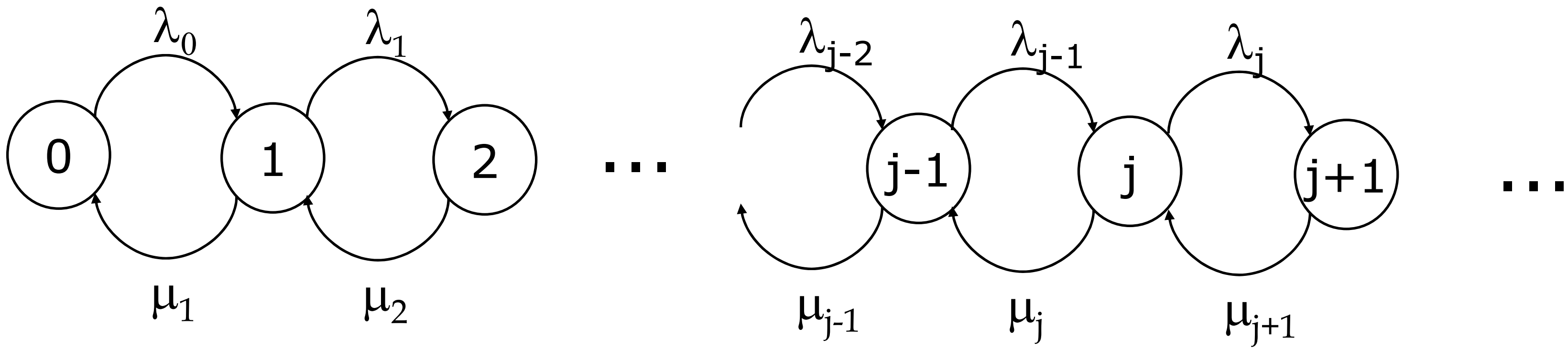


Diagrama de transição de estados

- Quando o sistema está no estado j, existem j usuários no sistema.

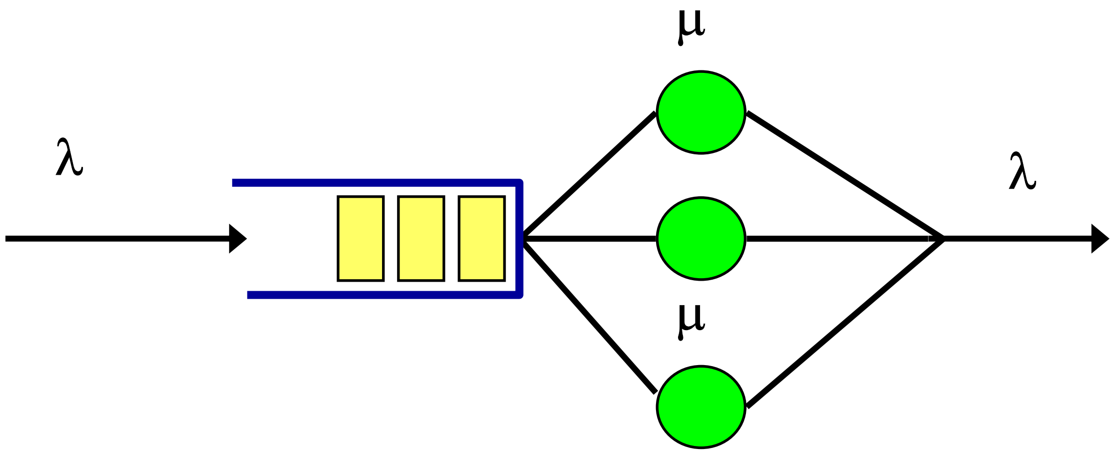
Probabilidades em Equilíbrio

- A solução geral de um processo nascimento e morte é dada por:

$$p_0 = \frac{1}{1 + \sum_{n=1}^{\infty} \prod_{j=0}^{n-1} \frac{\lambda_j}{\mu_{j+1}}}$$

$$p_n = p_0 * \prod_{j=0}^{n-1} \frac{\lambda_j}{\mu_{j+1}}$$

Fila Única M/M/m



m é o número de servidores

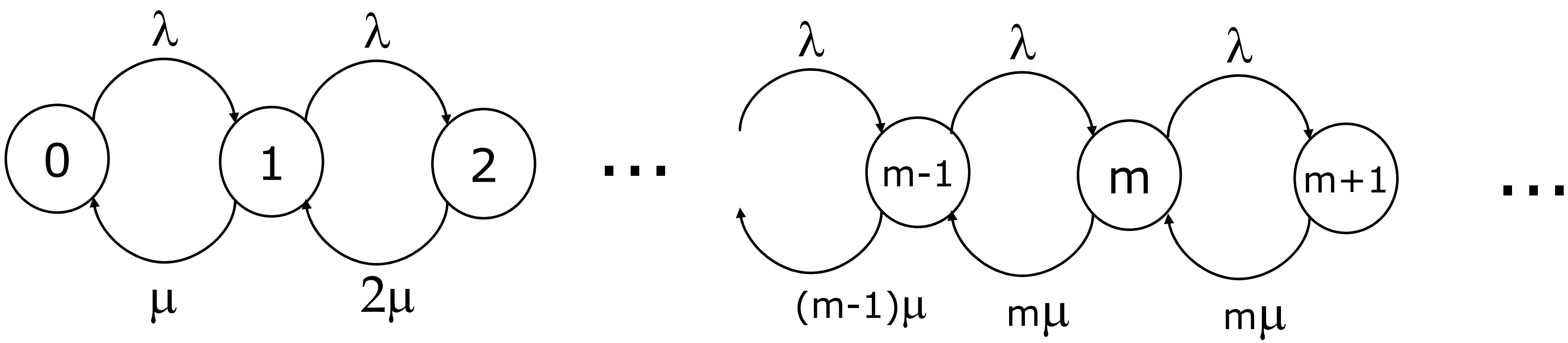
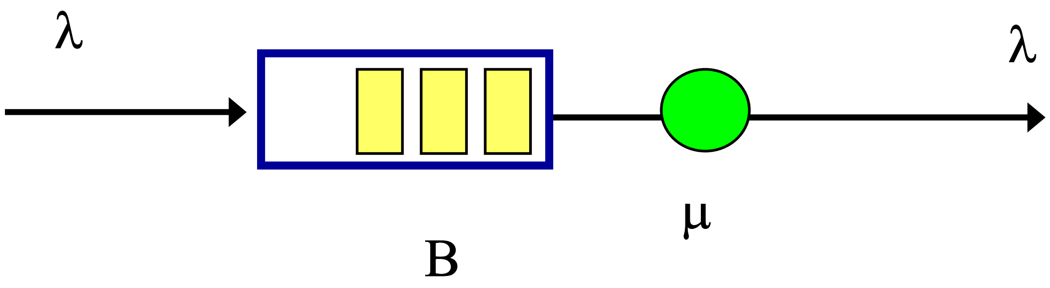


Diagrama de transição de estados

Fila Única M/M/1/B



Máximo de:

B no sistema
B-1 na fila

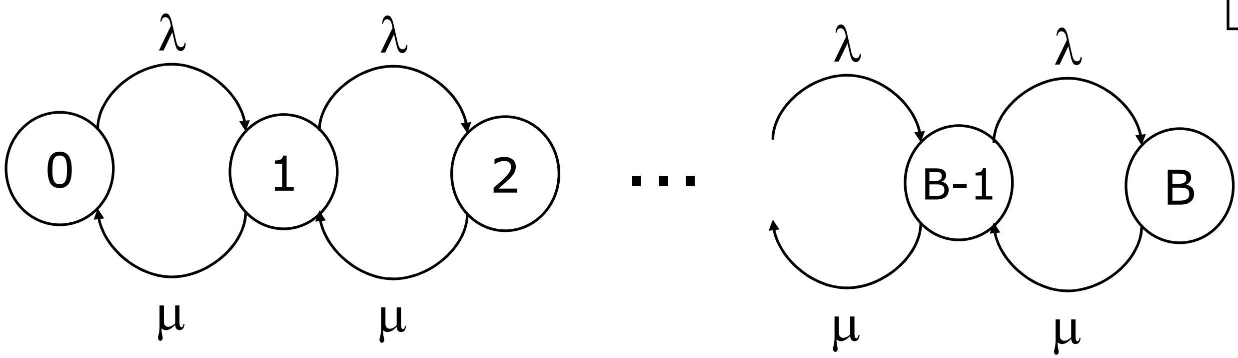
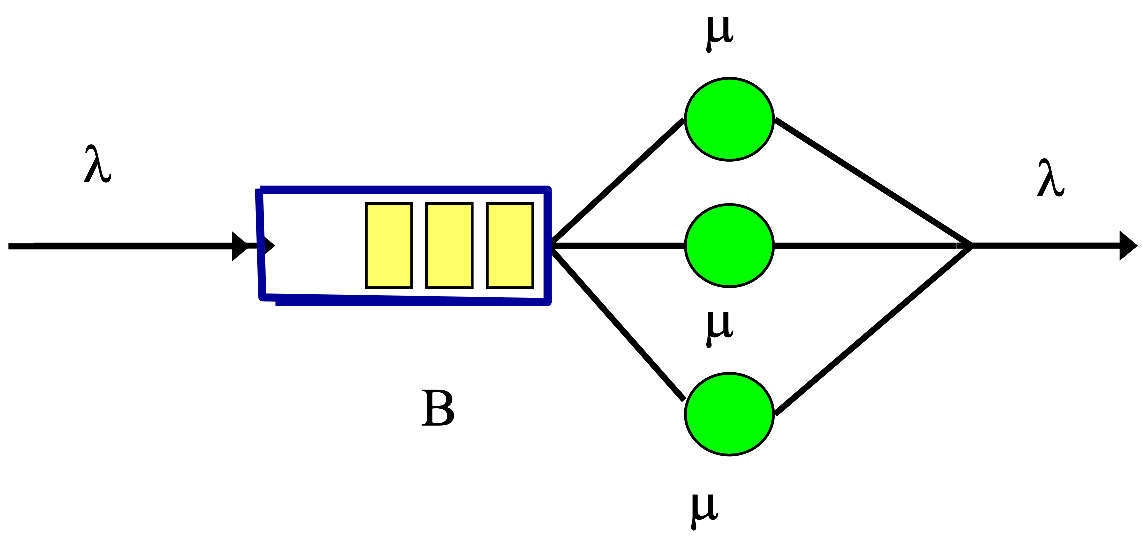


Diagrama de transição de estados

Fila Única M/M/m/B



Máximo de:

B no sistema
B-1 na fila

e B>m

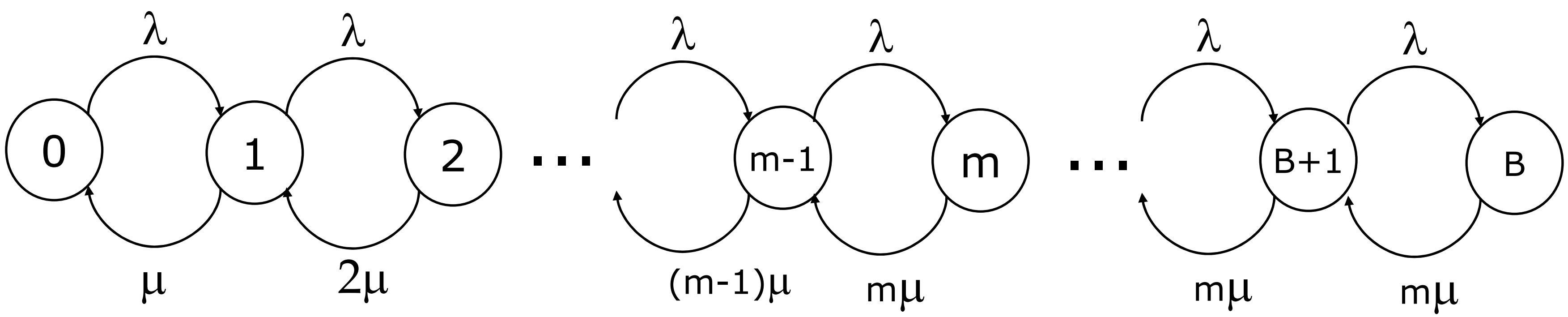


Diagrama de transição de estados

Exercícios

1) O tempo médio de resposta em um sistema de bases de dados do campus de uma universidade é 3 segundos. Durante um período de observação de 1 minuto, o tempo ocioso no sistema foi medido como 10 segundos. Usando o modelo M/M/1 para o sistema determine o seguinte:
   a) Utilização do sistema
   b) Tempo médio de serviço por consulta
   c) Número de consultas completadas durante o intervalo de observação
   d) Número médio de consultas no sistema
   e) Probabilidade do número de consultas no sistema ser maior que 10
   f) Tempo de resposta em 90%
   g) Tempo de espera em 90%.

   Resp.: a) 5/6; b) 0,5 s; c)100; d) 5; e) 0,135; f) 6,9 s; g) 6,36 s.

Exercícios

2) Considerando que o tempo médio de resposta no exercício anterior não está aceitável, a universidade está analisando uma das seguintes alternativas. Qual dessas alternativas garante um tempo médio de resposta menor?

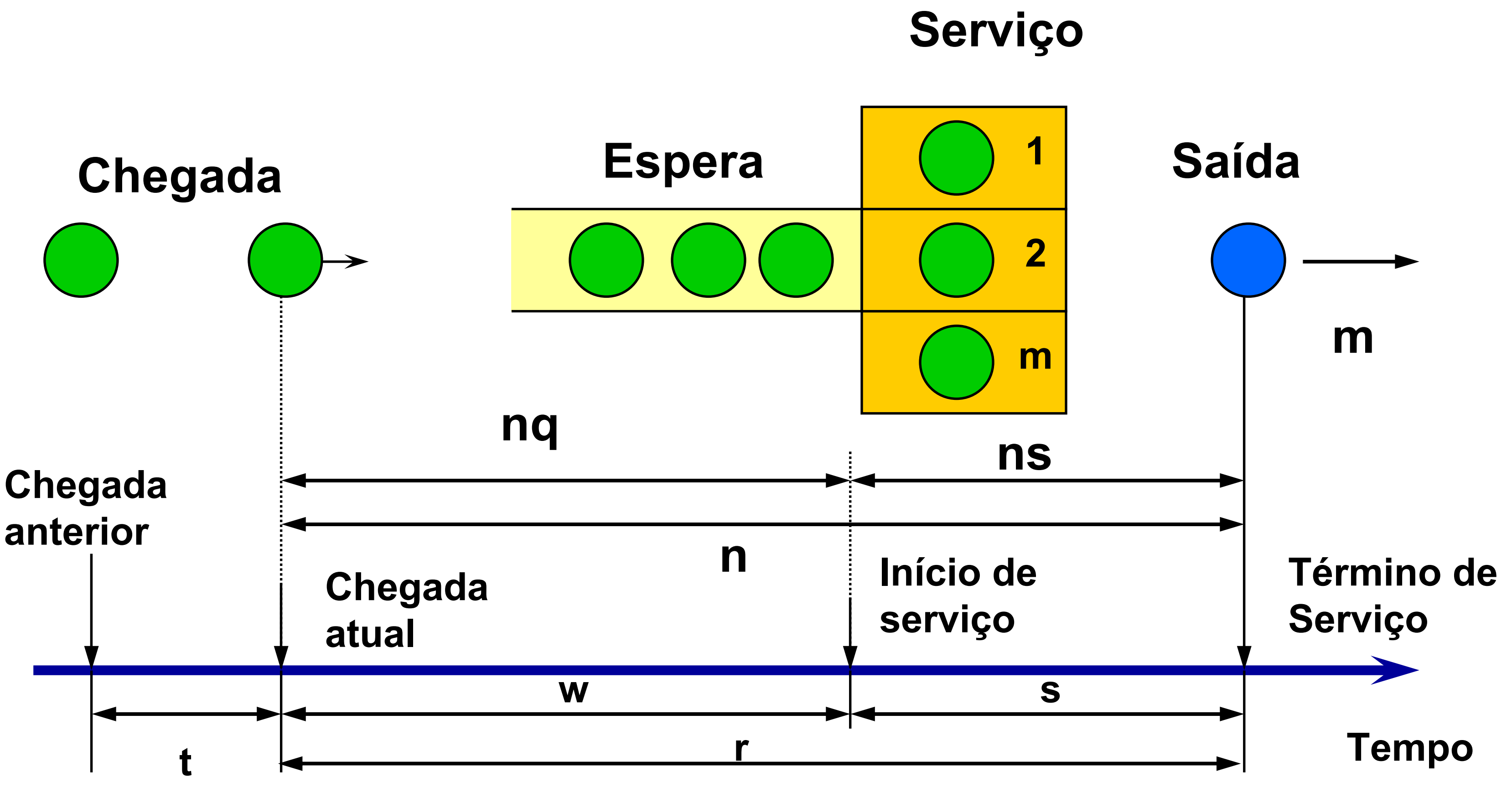
a) Substituir o computador por um que seja duas vezes mais rápido.

b) Colocar outro computador idêntico ao primeiro em outro lugar do campus recebendo a metade da carga.

c) Colocar outro computador idêntico ao primeiro com um balanceado de carga entre os dois.

Anexo: Fórmulário

Variáveis associadas a um sistema de filas:



Prâmetros do sistema de filas:

Parâmetro	Significado
λ	Taxa média de chegada
$1/\lambda$	Intervalo médio de chegada
μ	Taxa média de serviço por servidor
$1/\mu$	Tempo médio de serviço de um usuário
m	Número de servidores

Variáveis do sistema de filas

Variável	Significado
n	Número de usuários no sistema. Inclui os que estão na fila e sendo atendidos no servidor.
nq	Número de usuários esperando para serem atendidos. É sempre menor que n, pois não inclui os usuários em serviço.
ns	Número de usuários em serviço
E[n]	Número médio de usuários no sistema
E[nq]	Número médio de usuários na fila sem serem atendidos
E[ns]	Número médio de usuários sendo atendidos ($\leq$m)

Variáveis do sistema de filas

Variável	Significado
r	Tempo de resposta ou simplesmente tempo no sistema. Inclui tanto o tempo de espera como o tempo de serviço
w	Tempo de espera, isto é, intervalo de tempo entre o instante de chegada e o início do serviço
s	Tempo de serviço
E[r]	Tempo médio de resposta
E[w]	Tempo médio dos usuários na fila (sem serem atendidos)
E[s] = 1/μ	Tempo médio de serviço

Variáveis do sistema de filas

- Todas as variáveis anteriores com exceção de λ e μ são **variáveis aleatórias.**
- As variáveis λ e μ são parâmetros das distribuições.

Propriedades das Variáveis

- As propriedades a serem vistas a seguir valem para qualquer sistema de filas, não apenas aquelas com distribuições exponenciais).

Sistema Balanceado

- Diz-se que um sistema é **balanceado** se usuários não são perdidos no sistema, isto é, todos que entraram no sistema saíram em algum momento.
- Se o número de usuários cresce continuamente e se torna infinito, o sistema é dito **instável**.
- Para que um sistema seja **balanceado**, é necessário que a seguinte relação seja verdadeira

$$\lambda < m\mu$$

- De forma aproximada, o sistema será balanceado se dado um intervalo de tempo **T** suficientemente grande, o número de chegadas se aproxima do número de partidas.
- Sistemas **balanceados** equivalem a sistemas **estáveis** (em equilíbrio).

Fórmulas de Little

- As fórmulas de Little exigem que o sistema seja **balanceado** e são válidas independentemente das distribuições de chegada e atendimento.
- **Número x Tempo:** As fórmulas de Little relacionam o número de usuários com o tempo de resposta.

$$E[n] = \lambda * E[r]$$

e

$$E[nq] = \lambda * E[w]$$

Relações entre as variáveis

- Número no sistema, número em serviço e número na fila:

$$n = nq + ns$$

- Número médio de entidades no sistema: sendo n, nq e ns variáveis aleatórias então:

$$E[n] = E[nq] + E[ns]$$

- Se o número de usuários no servidor é independente do número de usuários na fila, isto é, $Cov(nq, ns) = 0$

então $Var[n] = Var[nq] + Var[ns]$

Relações entre as variáveis

- Tempo no sistema x Tempo na fila

$$r = w + s$$

- Tempo médio no sistema: sendo r,w e s variáveis aleatórias então:

$$E[r] = E[w] + E[s]$$

- Se o tempo de serviço é independente do tempo na fila, isto é,

$$Cov(w, s) = 0$$

então $$Var[r] = Var[w] + Var[s]$$

Sistemas de Fila Única

- No caso geral de **Processo nascimento e morte** o estado deste sistema indica o número de usuários no sistema.

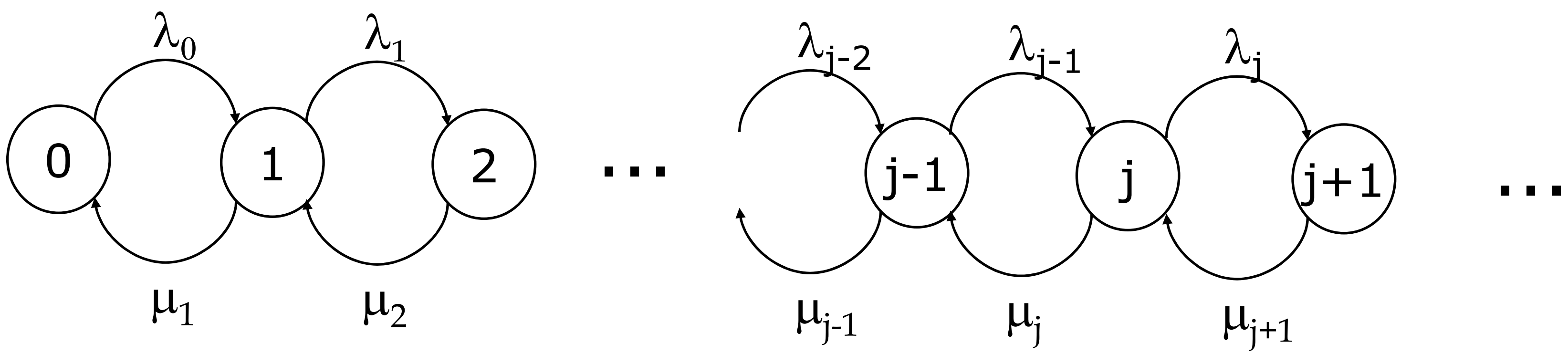


Diagrama de transição de estados

- Quando o sistema está no estado j, existem j usuários no sistema.

Probabilidades em Equilíbrio

- A probabilidade P_n de equilíbrio de um processo nascimento e morte se encontrar num determinado estado é:

$$p_n = p_0 \prod_{j=0}^{n-1} \frac{\lambda_{p_n j}}{\mu_{j+1}} \qquad n = 1,2,3,...$$

- **p_0** é a probabilidade do sistema possuir zero usuários.
- Considerando que $\sum_{n=0}^{\infty} p_n = 1$

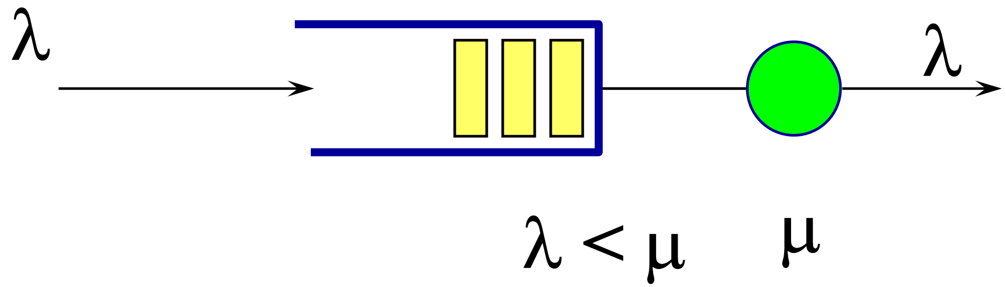
Probabilidades em Equilíbrio

- A solução geral de um processo nascimento e morte é dada por:

$$p_0 = \frac{1}{1 + \sum_{n=1}^{\infty} \prod_{j=0}^{n-1} \frac{\lambda_j}{\mu_{j+1}}}$$

$$p_n = p_0 * \prod_{j=0}^{n-1} \frac{\lambda_j}{\mu_{j+1}}$$

Fila Única M/M/1



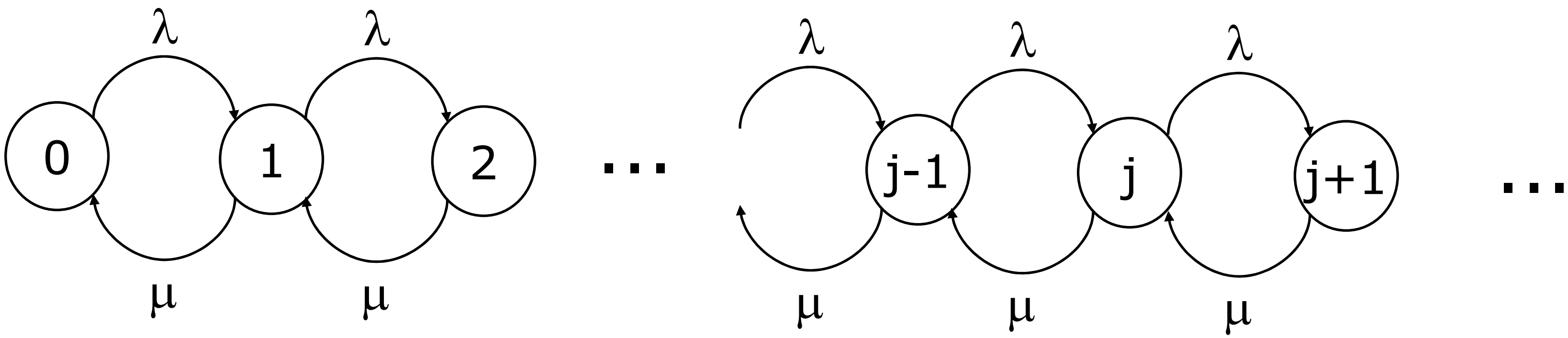


Diagrama de transição de estados

M/M/1

- Para este sistema tem-se:

  $\lambda_n = \lambda$

  $\mu_n = \mu$

1. Aplicando a forma geral de solução do processo nascimento e morte, chega-se:

$$p_n = \left(\frac{\lambda}{\mu}\right)^n * p_0$$

Denominando

$\rho = \lambda/\mu$, tem-se

$p_n = p_0 * \rho^n$, onde $p_0 = 1 - \rho$

M/M/1

2. Fator de utilização do servidor: **U**

   U = 1 - P_0 = ρ

3. Número médio de usuários no sistema:

$$E[n] = \sum_{n=1}^{\infty} n.P_n = \frac{\rho}{(1-\rho)}$$

4. Variância do número de usuários no sistema:

$$Var[n] = E[n^2] - E[n]^2 = \frac{\rho}{(1-\rho)^2}$$

M/M/1

5. Probabilidade de se ter n ou mais usuários no sistema:

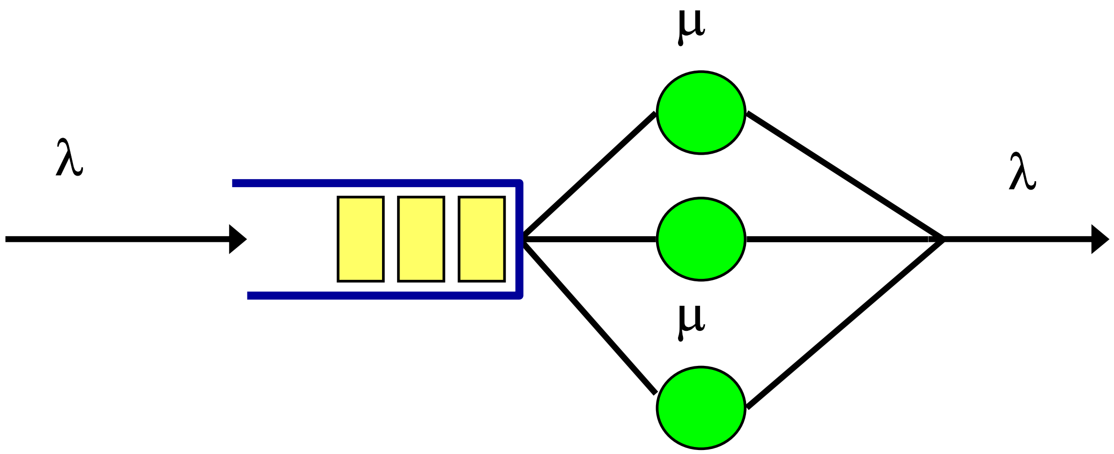
$$p_{\geq n} = \sum_{j=n}^{\infty} p_j = \sum_{j=n}^{\infty} \rho^j.(1-\rho) = \rho^n$$

6. Tempo médio de resposta: E[r]

E[n] = λ * E[r]

$$\mathrm{E[r]} = \frac{\mathrm{E[n]}}{\lambda} = \frac{\rho}{(1-\rho)} \cdot \frac{1}{\lambda} = \frac{\frac{1}{\mu}}{1-\rho}$$

Fila Única M/M/m



m é o número de servidores μ

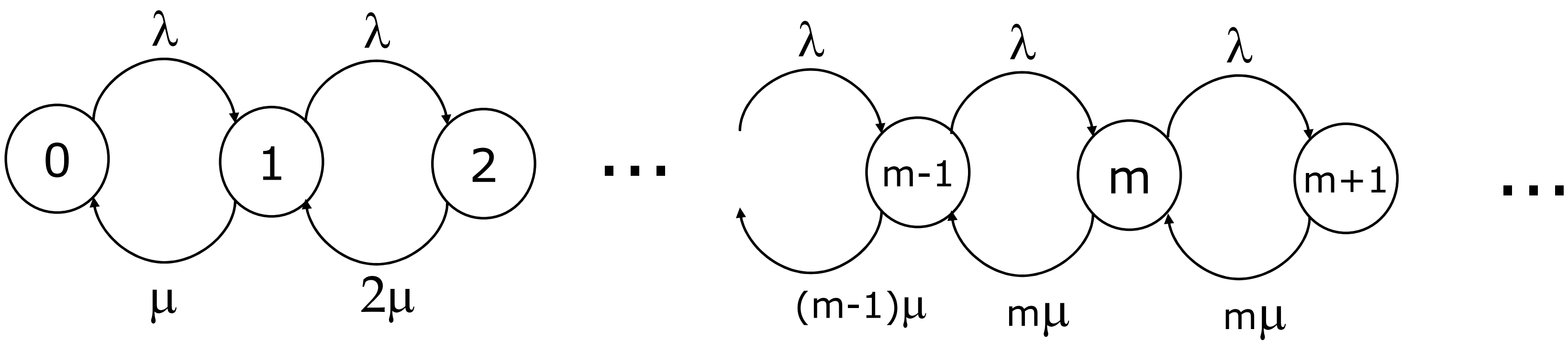


Diagrama de transição de estados

Fórmulas do M/M/1 e M/M/m

	M/M/1	**M/M/m**
• Taxa de chegada • Taxa de serviço • Número de Servidores	λ μ 1	λ μ m
• Fator de utilização	$U = \rho = \frac{\lambda}{\mu}$	$U = \rho = \frac{\lambda}{m \cdot \mu}$
• Probabilidade de zero usuários no sistema	$p_0 = 1 - \rho$	$p_0 = \left[1 + \frac{(m \cdot \rho)^m}{m!(1-\rho)} + \sum_{n=1}^{m-1} \frac{(m \cdot \rho)^n}{n!} \right]^{-1}$
• Probabilidade de n usuários no sistema n= 0,1,2,...,∞	$p_n = (1-\rho) \cdot \rho^n$	$p_n = p_0 \frac{(m \cdot \rho)^n}{n!}$ n<m $p_n = p_0 \frac{m^m \cdot \rho^n}{m!}$ n≥m
• Probabilidade de n ou mais usuários no sistema	$p_{\geq n} = \rho^n$	

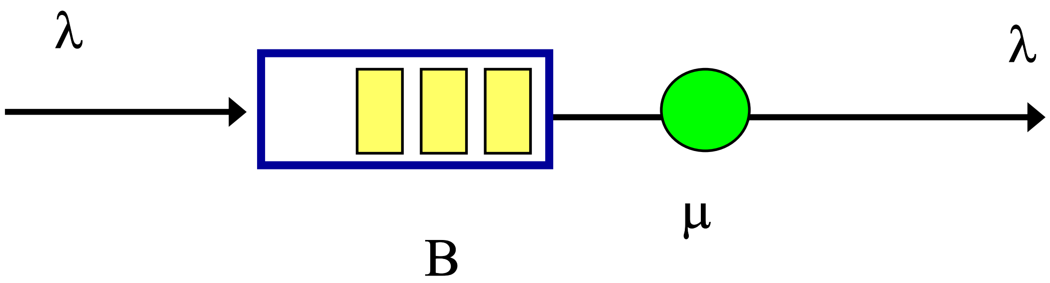
Fórmulas do M/M/1 e M/M/m

	M/M/1	M/M/m
• Probabilidade de esperar na fila	$p_{\geq 1} = \rho$	$\varsigma = P(\geq m \text{ usuários}) = \frac{(m\rho)^m}{m!(1-\rho)} P_0$
• Número médio de usuários no sistema	$E[n] = \rho/(1-\rho)$	$E[n] = m\rho + \rho\varsigma/(1-\rho)$
• Variância do número de usuários no sistema	$Var[n] = \rho/(1-\rho)^2$	$Var[n] = m\rho + \rho\varsigma\left[\frac{1+\rho-\rho\varsigma}{(1-\rho)^2} + m\right]$
• Tempo médio de resposta	$E[r] = (1/\mu)/(1-\rho)$	$E[r] = \frac{1}{\mu}\left[1 + \frac{\varsigma}{m(1-\rho)}\right]$
• Variância do tempo de resposta	$Var[r] = (1/\mu^2)/(1-\rho)^2$	$Var[r] = \frac{1}{\mu^2}\left[1 + \frac{\varsigma(2-\varsigma)}{m^2(1-\rho)^2}\right]$
• Número médio de usuários na fila	$E[n_q] = \rho^2/(1-\rho)$	$E[n_q] = \rho\varsigma/(1-\rho)$

Fórmulas do M/M/1 e M/M/m

	M/M/1	M/M/m
• Probabilidade de k usuários na fila	$p(n_q = k) = 1 - \rho^2$, $k=0$; $p(n_q = k) = (1-\rho)\cdot\rho^{k+1}$, $k>0$	
• Variância do número de usuários na fila	$Var[n_q] = \rho^2(1+\rho-\rho^2)/(1-\rho)^2$	$Var[n_q] = \rho\varsigma(1+\rho-\rho\varsigma)/(1-\rho)^2$
• Tempo médio de espera	$E[w] = \rho/[\mu(1-\rho)]$	$E[w] = \varsigma/[m\mu(1-\rho)]$
• Variância do tempo de espera	$Var[w] = (2-\rho)\rho/[\mu^2(1-\rho)^2]$	$Var[w] = \varsigma(2-\varsigma)/[m^2\mu^2(1-\rho)^2]$
• Tempo de resposta em q%	$r_{q\%} = E[r]\ln\left(\frac{100}{100-q}\right)$	
• Tempo de espera em q%	$w_{q\%} = \max\left[0, \frac{E[w]}{\rho}\ln\left(\frac{100\rho}{100-q}\right)\right]$	$w_{q\%} = \max\left(0, \frac{E[w]}{\varsigma}\ln\frac{100\varsigma}{100-q}\right)$

Fila Única M/M/1/B



Máximo de:

B no sistema
B-1 na fila

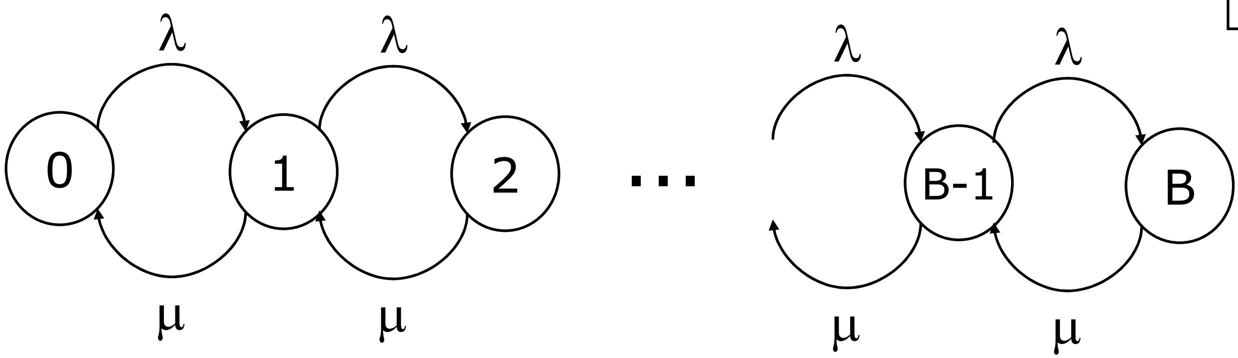
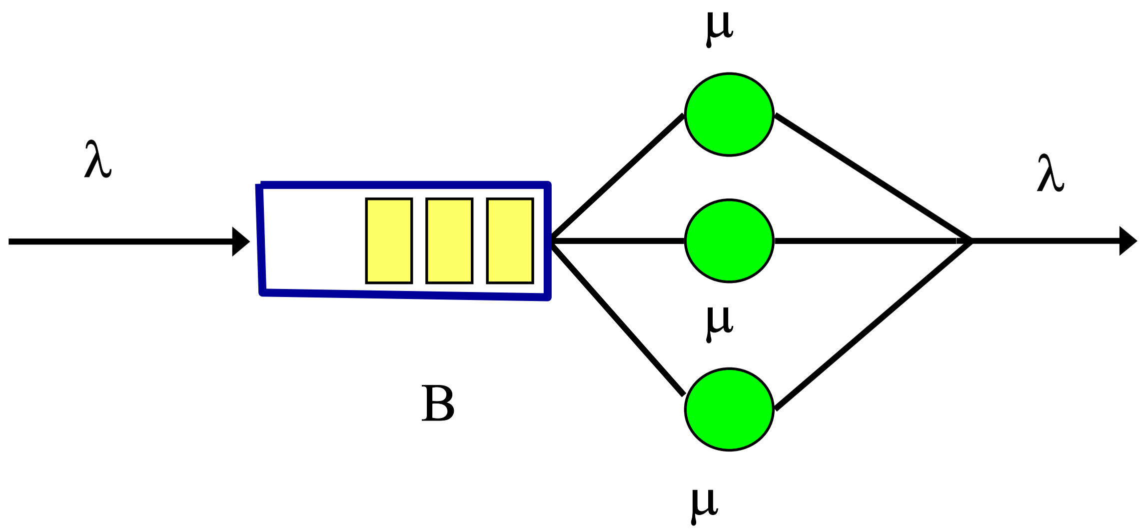


Diagrama de transição de estados

Fila Única M/M/m/B



Máximo de:

B no sistema
B-1 na fila

e B>m

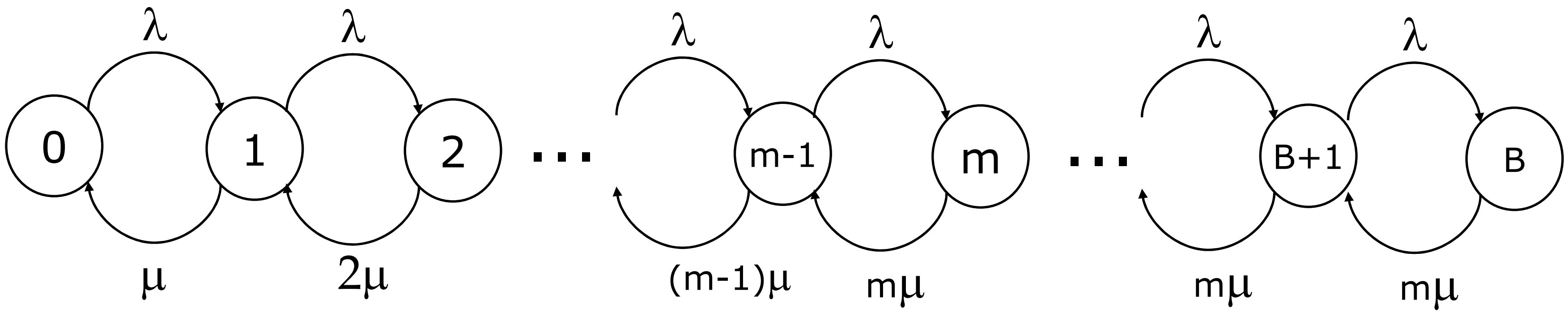


Diagrama de transição de estados

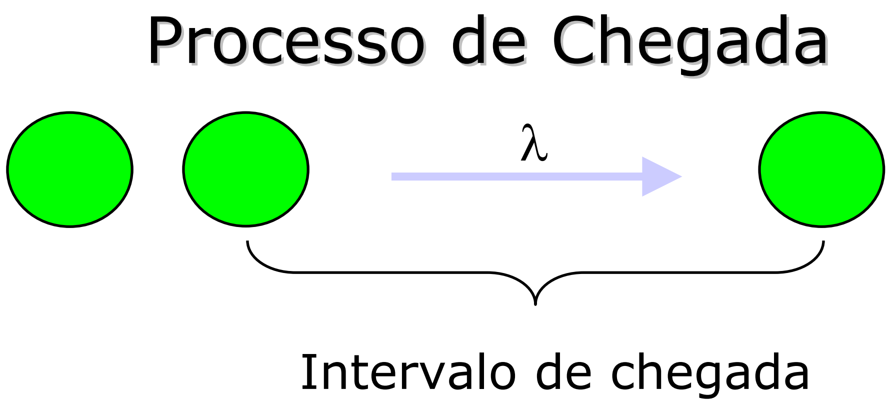
Fórmulas do M/M/1/B e M/M/m/B

	M/M/1/B	**M/M/m/B**
• Taxa de chegada • Taxa de serviço • Número de servidores • Número de buffers • Probabilidade de zero usuários no sistema	λ μ 1 B (B≥1) $p_0 = \frac{(1-\rho)}{(1-\rho^{B+1})}$ $\rho \neq 1$ $p_0 = \frac{1}{(B+1)}$ $\rho = 1$	λ μ m B (B≥1) $p_0 = \left[1 + \frac{(1-\rho^{B-m+1})(m \cdot \rho)^m}{m!(1-\rho)} + \sum_{n=1}^{m-1} \frac{(m \cdot \rho)^n}{n!}\right]^{-1}$
• Probabilidade de n usuários no sistema	$p_n = \frac{(1-\rho)}{(1-\rho^{B+1})}\rho^n$ $\rho \neq 1$ e $0 \leq n \leq B$ $p_n = \frac{1}{(B+1)}$ $\rho = 1$ e $0 \leq n \leq B$ $p_n = 0$ $n > B$	$p_n = p_0 \frac{(m \cdot \rho)^n}{n!}$ $0 \leq n \leq m$ $p_n = p_0 \frac{m^m \cdot \rho^n}{m!}$ $m \leq n \leq B$

Fórmulas do M/M/1/B e M/M/m/B

	M/M/1/B	**M/M/m/B**
• Taxa de chegada efetiva	$\lambda' = \lambda(1 - P_B)$	$\lambda' = \lambda(1 - P_B)$
• Taxa de perda	λP_B	λP_B
• Fator de utilização	$U = \rho(1 - P_B)$	$U = \rho(1 - P_B)$
• Número médio de usuários no sistema	$E[n] = \frac{\rho}{1-\rho} - \frac{(B+1)\rho^{B+1}}{1-\rho^{B+1}}$	
• Número médio de usuários na fila	$E[n_q] = \frac{\rho}{1-\rho} - \rho\frac{1+B\rho^{B}}{1-\rho^{B+1}}$	
• Tempo médio de resposta	$E[r] = E[n]/[\lambda(1 - P_B)]$	$E[r] = E[n]/[\lambda(1 - P_B)]$
• Tempo médio de espera	$E[w] = E[n_q]/[\lambda(1 - P_B)]$	$E[w] = E[n_q]/[\lambda(1 - P_B)]$

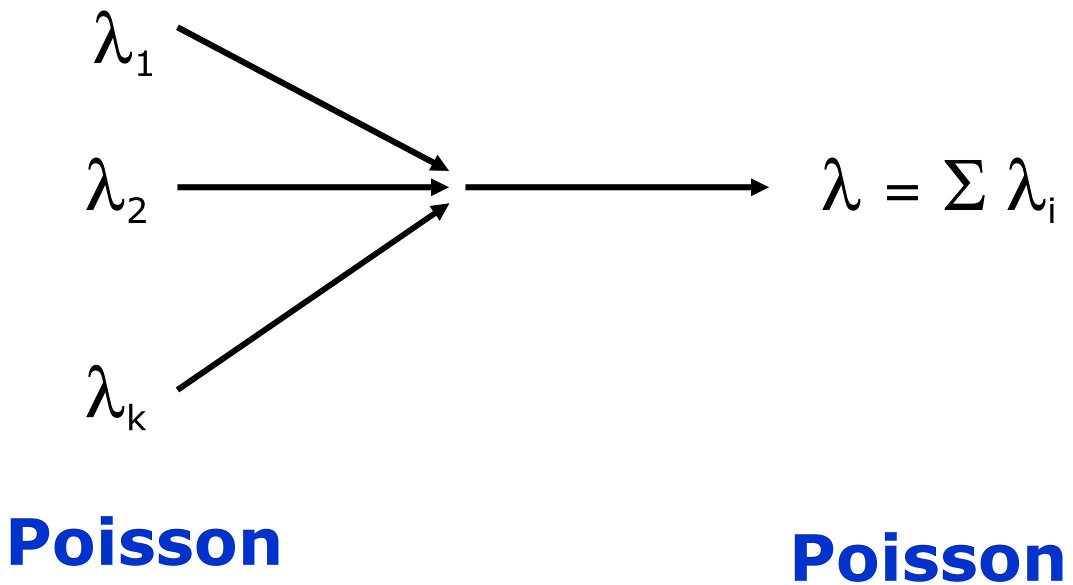
Processo de Poisson x Distribuição Exponencial



- Seja um processo de chegada em que os intervalos de chegada são variáveis aleatórias IID que obedecem distribuição exponencial com média $1/\lambda$.
- Neste caso, o número de elementos que chegam por unidade de tempo tem distribuição (discreta) de Poisson com taxa de chegada λ.

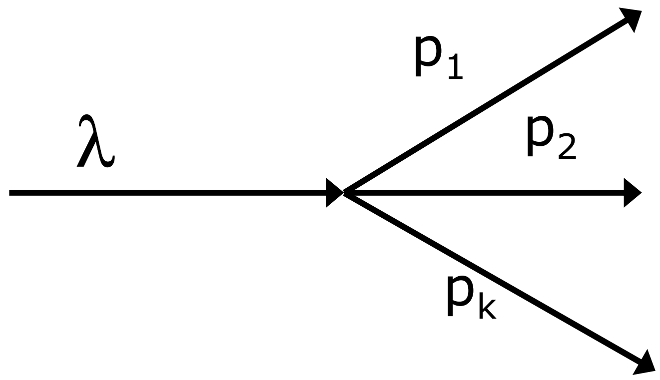
Processos de Poisson

- Junção de processos de Poisson:



Processos de Poisson

- Distribuição de processos de Poisson:



$\lambda_1 = p_1 * \lambda$

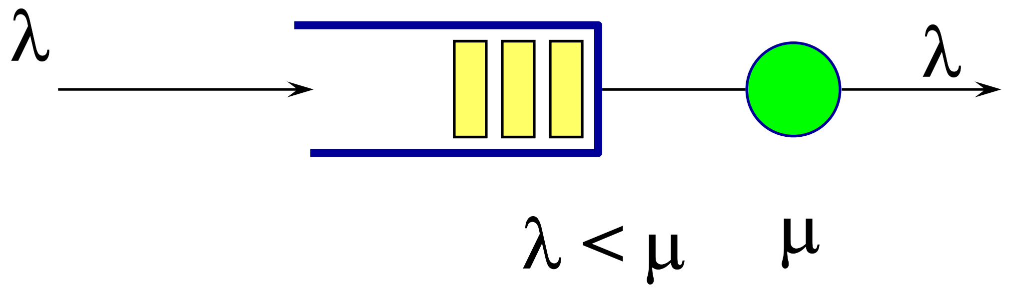
$\lambda_2 = p_2 * \lambda$

$\lambda_k = p_k * \lambda$

Poisson **Poisson**

Processos de Poisson

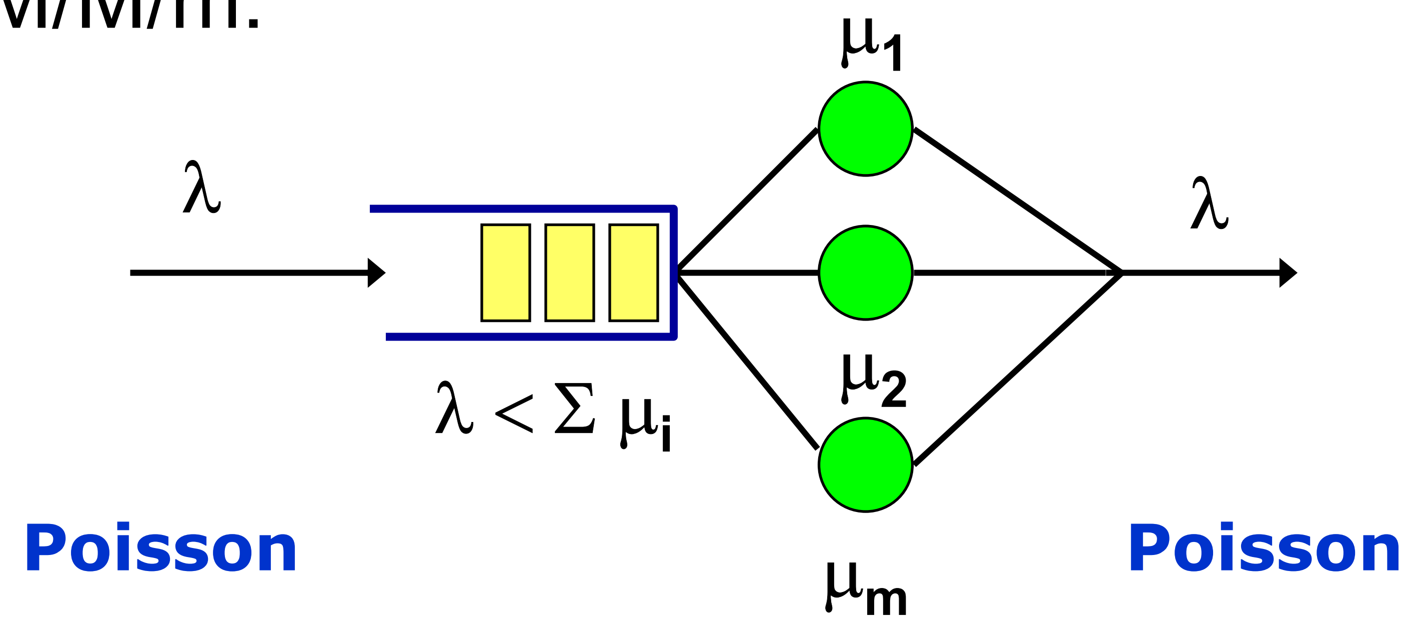
- Chegadas e partidas de um sistema M/M/1:



Poisson **Poisson**

Processos de Poisson

- Chegadas e partidas de um sistema M/M/m:



Fim do módulo
Sistemas de Filas Simples