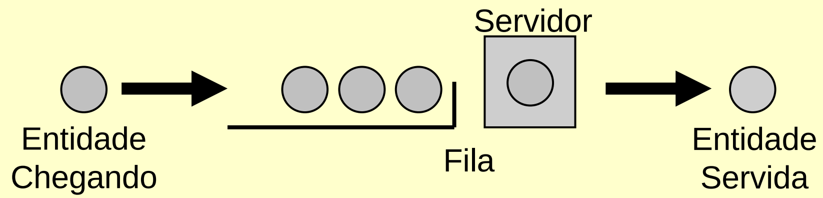
Teoria de Filas

Terminologia e Notação Básica

Referencia principal:

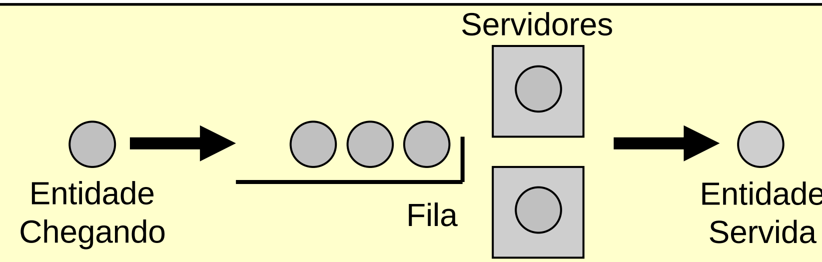
Freitas, P. J. <u>Introdução à Modelagem e Simulação de Sistemas</u>, 2ª Ed., Visual Books, 2008, Cap. 8.

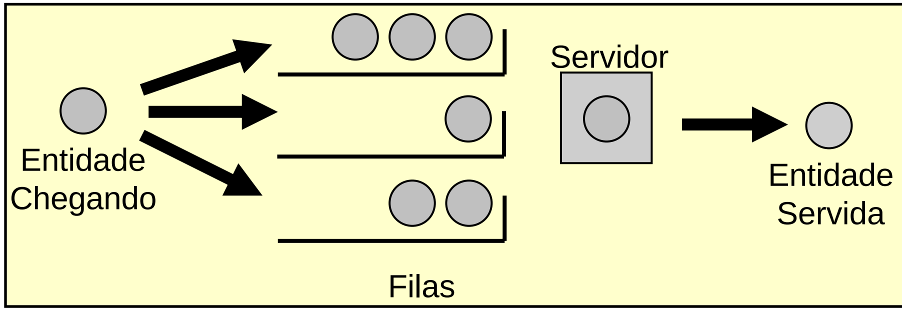
Exemplos de Sistemas de Filas



Fila Simples

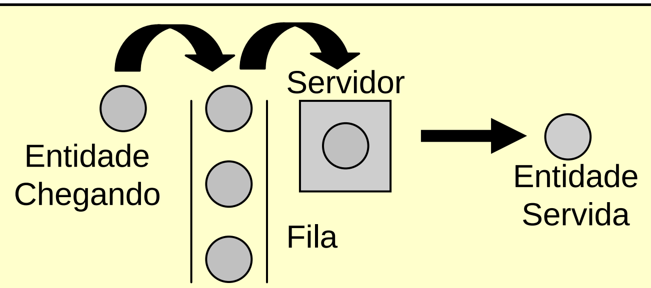
Centro com uma fila e múltiplos servidores





Centro com múltiplas filas e um servidor

Sistema de fila com regra de atendimento (LIFO)



Modelos de Filas e a Notação de Kendall

- A especificação de um modelo de fila normalmente requer que as seis características (parâmetros) sejam declarados.
- A notação mais conhecida é chamada de "notação de Kendall" que possui a forma A/S/m/B/K/SD, onde:
  - ✓ A: Distribuição do tempo de chegada;
  - ✓ S: Distribuição do tempo de serviço;
  - ✓ m: Número de servidores;
  - ✓ B: Capacidade do sistema (tamanho da fila);
  - ✓ K: Tamanho da população;
  - ✓ SD: Disciplina de serviço.
  - ✓ Se não especificado B e K são considerados infinitos e o tipo de disciplina de serviço é definido com FIFO.
  - ✓ As chegadas são individuais.

Modelos de Filas e a Notação de Kendall

- Os tipos de distribuição para tempos de chegada e de serviço, costumam ser representados pela seguinte nomenclatura:
  - M: *Exponencial* (M de *memoryless*);
  - E_k: *Erlang* com parâmetro k;
  - H_k: *Hiperexponencial* com parâmetro k;
  - D: Determinística;
  - G: *Geral*..

Exemplo 1

- Considere a especificação M/M/1/∞/∞/FIFO:
  - ✓ (M) tempos entre chegadas com distribuição exponencial;
  - ✓ (M) tempo de serviço com distribuição exponencial;
  - ✓ (1) um único servidor;
  - ✓ (∞) capacidade do sistema (tamanho da fila) sem limitações;
  - ✓ (∞) população também ilimitada
  - ✓ (FIFO) disciplina de serviço
  - ✓ Forma reduzida. Se os 3 últimos parâmetros são ∞/∞/FIFO então a notação ficaria M/M/1.

Exemplo 2

- Um sistema de fila que foi retratado com o emprego da notação de Kendall como M/G/4/50/2000/LIFO.
  - ✓ M: tempos entre chegadas com distribuição exponencial;
  - ✓ G: tempos de serviço distribuídos de forma arbitrária;
  - ✓ 4: quatro servidores;
  - ✓ 50: capacidade do sistema (tamanho da fila) limitada a 50 clientes (4 sendo servidos e no máximo 46 aguardando na fila);
  - ✓ 2000: a fonte de origem dos clientes com capacidade também limitada para 2000 clientes.
  - ✓ LIFO: a disciplina de serviço é do tipo último a chegar é o primeiro a ser atendido.

Análise de um Sistema de Fila Simples

- As principais variáveis associadas a estes modelos são as seguintes:
  - ✓ τ = tempo entre chegadas, isto é, tempo decorrido entre duas chegadas sucessivas;
  - ✓ λ = taxa média de chegadas = $1/E(\tau)$;
  - ✓ s = tempo de serviço por cliente;
  - ✓ μ = taxa média de serviço do servidor = $1/E(s)$;

Análise de um Sistema de Fila Simples

- Relacionamentos entre λ e μ que se aplicam em modelos de fila do tipo G/G/m (MM1 é um caso especial do G/G/m).

  1. *Condição de Estabilidade* ➔ $\lambda < m\mu$. Isto é, que taxa média de chegadas (λ) seja menor que a taxa média de serviço (μ), considerando m servidores, onde $m = 1, 2, 3....$
  2. *Número no Sistema versus Número na Fila* ➔ o número de clientes no sistema (n) = à soma do número de clientes na fila (nq) + o número de clientes sendo servidos (ns).

Análise de um Sistema de Fila Simples . . .

3. *Número versus Tempo* ➔ duas relações são importantes para a compreensão do comportamento geral de um sistema de filas. A condição fundamental é que se considere que clientes não são perdidos (deixam o sistema) por falta de capacidade da fila de espera.

> *Nº médio de clientes no sistema =*
>
> *= taxa de chegadas* X *tempo médio no sistema*

> *Nº médio de clientes na fila =*
>
> *= taxa de chegadas* X *tempo médio de espera*

Análise de um Sistema de Fila Simples . . .

4. *Tempo no Sistema versus Tempo na Fila* ➔ o tempo despendido por um cliente (ou o tempo de residência de um cliente) num sistema de fila (r) é igual à soma do tempo de espera (w) com o tempo recebendo serviço (s): $r = w + s$.

Modelo M/M/1

- É o modelo de fila mais empregado;
- Processos ou sistemas com um único servidor podem ser representados por este modelo.
- TEC e TS bem descritos por uma Exponencial (processos de Poisson).
- Sem limitações para o tamanho da fila
- Disciplina de gerenciamento é do tipo FIFO.
- Parâmetros necessários para sua análise são:
  - ✓ Taxa de chegada de clientes por unidade de tempo λ,
  - ✓ Taxa de demanda de serviço do servidor por unidade de tempo μ.

Modelo M/M/1

- Estado do sistema determinado pelo número de clientes que se encontram no mesmo em determinado momento (na fila e no serviço).

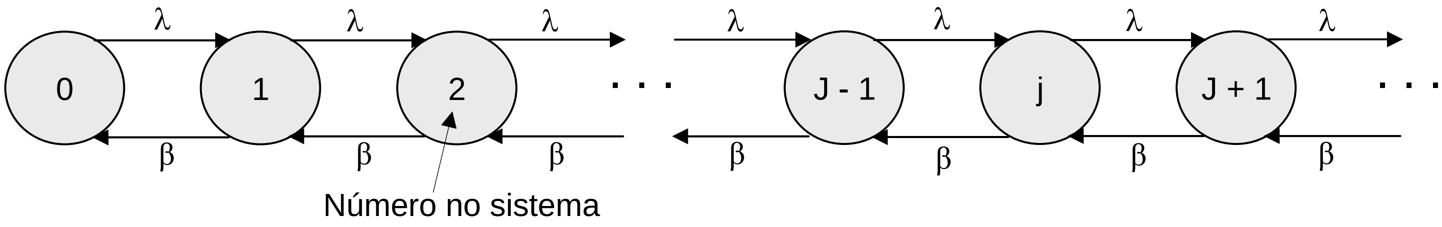


Diagrama de estados modelo M/M/1

Modelo M/M/1 - Fórmulas

1	Intensidade de tráfego	$\rho = \lambda / \mu$
2	Condição de estabilidade (requisito necessário para que a fila não cresça indefinidamente)	$\rho < 1$, isto é, $\lambda < \mu$
3	Probabilidade de *zero* clientes no sistema	$p_0 = 1 - \rho$
4	Probabilidade de *n* clientes no sistema	$p_n = (1 - \rho).\rho^n,\ n = 0,1,\ldots,\infty$
5	Probabilidade de *n ou mais* clientes no sistema	$P_{(n\ ou\ mais)} = \rho^n$
6	Número médio de clientes no sistema	$E[n] = \rho / (1 - \rho)$
7	Variância do número de clientes no sistema	$Var[n] = \rho / (1 - \rho)^2$
8	Probabilidade de *k* clientes na fila	$P(n_q = k) = \begin{cases} 1 - \rho^2, & k = 0 \\ (1 - \rho).\rho^{k+1}, & k > 0 \end{cases}$
9	Número médio de clientes na fila:	$E[n_q] = \rho^2 / (1 - \rho)$
10	Variância do número de clientes na fila	$Var[n_q] = \rho^2.(1 + \rho - \rho^2) / (1 - \rho)^2$
11	Tempo médio no sistema (ou de resposta)	$E[r] = (1/\mu) / (1 - \rho)$
12	Variância do tempo no sistema (ou de resposta)	$Var[r] = 1/\mu^2 \big/ (1 - \rho)^2$
13	Tempo médio de espera	$E[w] = \rho.(1/\mu) \big/ (1 - \rho)$
14	Variância do tempo de espera	$Var[w] = (2 - \rho).\rho / (\mu^2.(1 - \rho)^2)$

Exemplo 8.1

- Medições feitas em um roteador mostram que pacotes de dados chegam para serem roteados com uma taxa média de 1250 pacotes por segundo (pps) e são encaminhados (processados) com uma média de um pacote a cada 0,5 ms.
  - ✓ Qual é a probabilidade de haver uma sobrecarga no buffer do roteador se este dispõe de espaço para apenas 35 pacotes?
  - ✓ Qual o tamanho de buffer necessário para que a taxa de perda de pacotes seja inferior a 1 pacote por milhão.

Dados*:*

Taxa de chegada $\lambda = 1.250\ pps$

Taxa de serviço $\mu = 1/0{,}0005 = 2.000\ pps$

Exemplo 8.1 - Solução

Utilização do roteador (intensidade do tráfego): $\rho = \lambda / \mu$ = 1.250/2.000 =0,625

Numero médio de pacotes no roteador: E[n] = $\rho / (1- \rho)$ = 0,625 / 0,375 = 1,67

Tempo médio de resposta do roteador:

E[r] = (1/μ) / (1- ρ) = (1/2.000)/(1- 0,625) = (0,0005 / 0,375) = 0,0013 seg.

Probabilidade de estouro do *buffer*:

P (n ≥ 35) = ρ^{35} = 0.625^{35} = 7,2 x 10^{-8} ≈ 72 pacotes para cada bilhão de pacotes roteados.

Tamanho do buffer para que a probabilidade deste estourar seja ≤ 10^{-6}.

O que se deseja é que $\rho^n \leq 10^{-6}$.

$\log(\rho^n) \leq \log(10^{-6}) \therefore n.\log(\rho) \leq \log(10^{-6}) \therefore n > \log(10^{-6})/\log(\rho)$

$n > \log(10^{-6}) / \log(0.625) \therefore -5/-0{,}204 > 24{,}5$. Ou $n \geq 25$.

Com um buffer de tamanho 25, a taxa de perda menor que 1 pacote por milhão de pacotes roteados.

Exemplo 8.2

- Verifica-se que o tempo entre requisições a um servidor Web pode ser modelado por meio de uma distribuição exponencial com parâmetro para sua média igual a 8 ms. O tempo médio para o servidor processar cada requisição também pode ser descrito como uma exponencial com tempo médio de aproximadamente 5 ms. Com base nestas medições solicita-se:
  - ✓ Qual o tempo médio de resposta experimentado pelos usuários?
  - ✓ Quanto mais rápido deveria ser um novo servidor para que o tempo de resposta caia para a metade do tempo atual?
  - ✓ Qual deve ser o tamanho do buffer de entrada do servidor para que a perda de requisições seja de, no máximo, 1 a cada bilhão de requisições?

Exemplo 8.2 - Solução

- Dados:

  O tempo médio entre chegadas de requisições (req) é de uma a cada 8 ms.

  Logo a taxa de requisições λ = 1 req/8 ms = 0,125 req/ms.

  O tempo de processamento é de 5 ms/req.

  Logo a taxa de serviço μ = 1 req/5 ms = 0,2 req/ms

- Solução:

  Taxa de utilização do servidor $\rho = \lambda/\mu = 0,125/0,2 = 0,625$ ou 62,5 % de sua capacidade máxima.

  (a) Tempo médio de resposta (ou valor esperado para o tempo de resposta):

  $$E[r] = (1/\mu)/(1- \rho) = (1/0,2)/(1 – 0,625) = 13,333 \text{ ms};$$

Exemplo 8.2 - Solução

Qual deve ser o novo valor de μ para que r = 13,33/2 = 6,67 ms?

Considerando a mesma taxa de requisições λ e

$r = (1/\mu)/(1- \rho) = 1/(\mu - \lambda)$,

pode-se isolar $\mu = (1/r) + \lambda)$ = (1/6,67) + 0,125 = 0,275 req/ms.

Isto significa que o novo servidor deve ser ≈ ((0,275 – 0,2)/0,2)) x 100% = 37,5% mais rápido do que o anterior para que o tempo médio de resposta seja a metade do atual.

Exemplo 8.2 - Solução

É preciso encontrar um valor limite n, acima do qual haveria perda de requisições, caso a fila (buffer) tenha que ser limitada (como de fato acontece nos sistemas do mundo real).

Deseja-se que a probabilidade de uma requisição chegar ao sistema e encontrar ***k*** ou mais requisições seja menor ou igual a 10^{-9}. Empregando-se a relação 5 da Tabela 8.1 ($P(n \text{ ou mais}) = \rho^n$), obtém-se: $\rho^n = 10^{-9}$.

Solução para $\rho = 0{,}667$ é dado por:

$$n \geq \log(10^{-9})/\log(0{,}625) = 44{,}09.$$

Logo, em se tratando de um sistema com a fila limitada, haveria a necessidade de espaço para comportar pelo menos 45 requisições para que a de descarte destas junto ao servidor ser menor do que 10^{-9}.

Modelo M/M/m

- É utilizado para a modelagem de sistemas que apresentam uma única fila diante de vários dispositivos (servidores) idênticos, ou sistemas multiprocessados.
- Como o M/M/1, TEC e TS são Exponenciais (processos de Poisson).
- Sem limitações para o tamanho da fila
- Disciplina de gerenciamento é do tipo FIFO.
- Parâmetros necessários para sua análise são:
  - ✓ Taxa de chegada de clientes por unidade de tempo λ,
  - ✓ Taxa de serviço de cada um dos servidores por unidade de tempo μ
  - ✓ Quantidade de servidores disponíveis m.

Modelo M/M/m

- Assim como no modelo anterior, o estado do sistema também é determinado pelo número *n* de clientes que ali se encontrem em determinado momento (na fila e nos servidores).

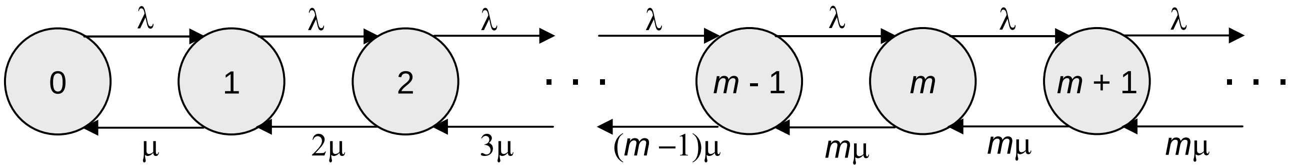


Diagrama de estados modelo M/M/m

Fórmulas Modelo M/M/m

1	Intensidade de tráfego	$\rho = \lambda / (m.\mu)$
2	Condição de estabilidade	$\rho < 1$, isto é, $\lambda < m.\mu$
3	Probabilidade de zero clientes no sistema	$p_0 = \left[1 + \frac{(m\rho)^m}{m!(1-\rho)} + \sum_{n=1}^{m-1} \frac{(m\rho)^n}{n!} \right]^{-1}$
4	Probabilidade de n clientes no sistema	$p_n = \begin{cases} \frac{(m\rho)^n}{n!} p_0, & n = 1,2,\ldots, m-1 \\ \frac{\rho^n m^m}{m!} p_0, & n = m, m+1, \ldots, \infty \end{cases}$
5	Probabilidade de enfileiramento (∂ = probabilidade de m ou mais clientes no sistema)	$\partial = P(\geq m \ \ clientes) = \frac{(m\rho)^m}{m!(1-\rho)} p_0$
6	Número médio de clientes no sistema	$E[n] = m\rho + \rho.\partial / (1 - \rho)$
7	Variância do número de clientes no sistema	$Var[n] = m\rho + \rho.\partial \left[\frac{1+\rho-\rho.\partial}{(1-\rho)^2} + m \right]$
8	Número médio de clientes na fila:	$E[n_q] = \rho.\partial / (1 - \rho)$
9	Variância do número de clientes na fila	$Var[n_q] = \partial\rho.(1 + \rho - \partial\rho) / (1 - \rho)^2$
10	Tempo médio no sistema (ou de resposta)	$E[r] = \frac{1}{\mu} \left(1 + \frac{\partial}{m(1-\rho)} \right)$
11	Variância do tempo no sistema (ou de resposta)	$Var[r] = \frac{1}{\mu^2} \left[1 + \frac{\partial(2-\rho)}{m^2(1-\rho)^2} \right]$
12	Tempo médio de espera	$E[w] = E(n_q) / \lambda = \partial / [m\mu(1-\rho)]$
13	Variância do tempo de espera	$Var[w] = \partial(2-\partial) / [m^2\mu^2(1-\rho)^2]$
14	Taxa de utilização média de cada servidor;	$U = \lambda / m\mu = \rho$

Exemplo 8.4

- Para o mesmo sistema citado no exemplo 8.2, considere uma nova configuração no servidor Web. Agora, o número de processadores é igual a quatro. Com esta nova configuração, até quatro requisições podem ser simultaneamente atendidas. Neste caso, considerando a mesma demanda, qual seria o tempo médio de resposta experimentado pelos usuários?
- Dados:
  - ✓O tempo médio entre chegadas de requisições (*req*) é de uma a cada 8 ms.
  - ✓Logo a taxa de requisições λ = 1 req/8 ms = 0,125 *req*/ms.
  - ✓O tempo de processamento é de 5 ms/req.
  - ✓Logo, Taxa de serviço μ = 1 req/5 ms = 0,2 *req*/ms
  - ✓Número de servidores m = 4.

Exemplo 8.4 - Solução

Taxa de utilização do servidor $\rho = \lambda / m.\mu = 0,125/4.0,2 = 0,156$ ou 15,6 % de sua capacidade máxima.

Tempo médio de resposta: $E[r] = \dfrac{1}{\mu}\left(1 + \dfrac{\partial}{m(1-\rho)}\right)$, onde $\partial =$ probabilidade de m ou mais clientes no sistema.

$$\partial = P(\geq m \ \ clientes) = \frac{(m\rho)^m}{m!(1-\rho)} p_0 \text{, onde } p_0 = \left[1 + \frac{(m\rho)^m}{m!\,(1-\rho)} + \sum_{n=1}^{m-1} \frac{(m\rho)^n}{n!}\right]^{-1}.$$

Assim, $p_0 = 0{,}536$ e $\partial = 0{,}004$.

O novo valor esperado para o Tempo de Resposta será:

$E[r] = (1/0{,}2) + (0{,}004/(4.0{,}2).(1 - 0{,}156)) = 5{,}01\ ms.$

Portanto, pode-se esperar que um aumento de quatro vezes na capacidade nominal de processamento do servidor provoque uma redução de aproximadamente 60% no tempo de resposta, isto é, de 13,33 ms para 5,01 ms.

Exemplo 8.5

- Uma companhia aérea pretende disponibilizar 5 terminais de auto-atendimento em um aeroporto.
- A gerencia de atendimento da companhia está em dúvida entre duas alternativas:
  - ✓(a) colocar os 5 terminais juntos, com acesso por fila única;
  - ✓(b) distribuí-los individualmente em vários locais do aeroporto, considerando, neste caso, que a demanda se distribua igualmente entre os locais onde se encontrarem os 5 terminais.
- Para apoiar futura decisão, a gerencia solicita ao departamento técnico uma análise para prever o comportamento das duas alternativas, principalmente em relação a expectativa do tempo médio despendido pelos clientes no sistema. Os dados fornecidos pelo gerente são os seguintes:

Exemplo 8.5 – Dados e Solução

- Dados:

  ✓Estima-se a chegada de 100 clientes por hora nos terminais;

  ✓Cada cliente deverá permanecer, em média, 2 minutos junto ao terminal;

- Solução

  O sistema poderá ser avaliado considerando

  (a) Emprego de um modelo M/M/5 ou

  (b) Empregando-se um modelo M/M/1 (neste caso com a demanda reduzida a 1/5 da demanda original).

Exemplo 8.5 – Solução

a) Modelo M/M/5

Taxa de chegadas 100 clientes por hora. Logo, $\lambda = 1{,}667$ clientes/min;

Taxa de serviço $\mu = 1$ cliente/2 min $= 0{,}5$ *clientes*/min

Número de servidores $m = 5$.

Resultados do modelo:

Taxa média de utilização dos servidores: $\rho = \lambda/m.\mu = 1{,}667/5.0{,}5 = 0{,}668$ ou 66,8 % de sua capacidade máxima

Tempo médio no sistema: $E[r] = \dfrac{1}{\mu}\left(1 + \dfrac{\partial}{m(1-\rho)}\right)$, onde

$$\partial = P(\geq m \;\; clientes) = \frac{(m\rho)^m}{m!(1-\rho)} p_0 \text{ e } p_0 = \left[1 + \frac{(m\rho)^m}{m!(1-\rho)} + \sum_{n=1}^{m-1} \frac{(m\rho)^n}{n!}\right]^{-1}.$$

Calculando-se $p_0 = 0{,}128$ e $\partial = 1{,}314$. Assim, o tempo esperado no sistema:

$E[r] = (1/0{,}5).[1+(1{,}314/5.(1-0{,}668))] =$ 3,58 min, com $Var[r] = 4{,}1$ min.

Exemplo 8.5 – Solução

b) Modelo M/M/1

Taxa de chegadas será de 1/5 da taxa original. Logo, chegam 20 clientes/hora ou 1 cliente a cada 3 min. Assim, $\lambda = 0{,}333$ clientes/min;

Taxa de serviço permanece a mesma: $\mu = 1$ cliente/2 min = 0,5 *clientes*/min

O número de servidores é igual a um servidor por local $m = 1$.

Resultados do modelo:

Taxa média de utilização dos servidores: $\rho = \lambda/\mu = 0{,}333/0{,}5 = 0{,}666$ ou 66,6 % de sua capacidade máxima.

Tempo médio no sistema: $E[r] = (1/\mu)/(1 - \rho) = (1/0{,}5)/(1-0{,}666) = \underline{5{,}99 \text{ min}}$

A variância do tempo no sistema será: Var[r] = (1 / µ2) / (1 - ρ)2 = 35,86 e o desvio-padrão será 5,98 min.

Modelo M/M/m/B

- O modelo M/M/m/B apresenta fila com múltiplos servidores e tamanho limitado para área de espera.
- Se a capacidade da área de espera ser alcançada, todos os novos clientes são “perdidos”.
- B deve ser maior ou igual a m.
- Como o M/M/1, TEC e TS são Exponenciais (processos de Poisson) e o gerenciamento da fila é do tipo FIFO.
- Parâmetros necessários para sua análise são:
  - ✓ Taxa de chegada de clientes por unidade de tempo λ,
  - ✓ Taxa de serviço de cada um dos servidores por unidade de tempo μ
  - ✓ Quantidade de servidores disponíveis m.
  - ✓ O tamanho máximo da fila B

Modelo M/M/m/B

- Assim como no modelo anterior, o estado do sistema também é determinado pelo número *n* de clientes que ali se encontrem em determinado momento (na fila e nos servidores). Há agora uma limitação em B.

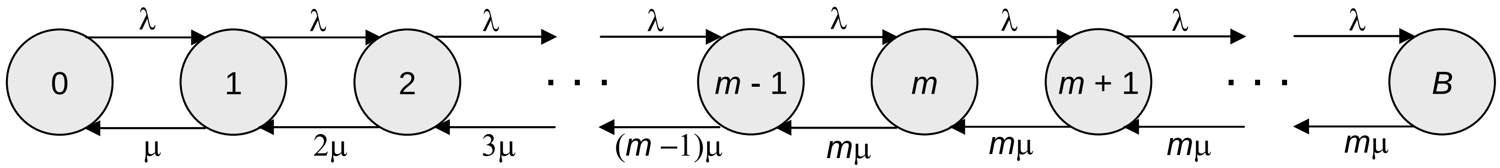


Diagrama de estados modelo M/M/m/B

Fórmulas Modelo M/M/m/B

1	Intensidade de tráfego	$\rho = \lambda/(m.\mu)$
2	Condição de estabilidade	$\rho < \infty$
3	Probabilidade de zero clientes no sistema	$p_0 = \left[1 + \frac{(1-\rho^{B-m+1})(m\rho)^m}{m!(1-\rho)} + \sum_{n=1}^{m-1} \frac{(m\rho)^n}{n!}\right]^{-1}$
4	Probabilidade de n clientes no sistema em termos da intensidade de tráfego	$p_n = \begin{cases} \frac{1}{n!}(m\rho)^n p_0, & 0 \le n < m \\ \frac{m^m \rho^n}{m!} p_0, & \mathrm{m} \le \mathrm{n} \le B \end{cases}$
5	Número médio de clientes no sistema	$E[n] = \sum_{n=1}^{B} n p_n$
6	Número médio de clientes na fila	$E[n_q] = \sum_{n=m+1}^{B} (n-m) p_n$
7 8 9	Taxa efetiva de chegada no sistema Tempo médio no sistema (ou de resposta) Tempo médio de espera	$\lambda' = \sum_{n=0}^{B-1} \lambda p_n = \lambda(1-p_B)$ $E[r] = E[n]/\lambda' = E[n]/[\lambda(1-P_B)]$ $E(w) = E(r) - 1/\mu = E[n_q]/[\lambda(1-p_B)]$
10	Taxa de perda por unidade de tempo	λp_B *clientes/unidade de tempo*
11	Taxa de utilização média de cada servidor;	$U = \lambda' / m\mu = \rho(1-p_B)$

Exemplo 8.6

- Para o roteador do exemplo 8.1, considere agora que o mesmo tenha seu *buffer* (área de espera) limitado a um máximo de 5 pacotes. Medições feitas no roteador mostram que pacotes de dados chegam para serem roteados com uma taxa média de 1250 pacotes por segundo (pps) e são encaminhados (processados) numa média de um pacote a cada 0,5 ms.
- Dados:
  - ✓ Taxa de chegada $\lambda = 1.250$ pps;
  - ✓ Taxa de serviço $\mu = 1/0{,}0005 = 2.000$ pps;
  - ✓ $m = 1$;
  - ✓ $B = 5$

Exemplo 8.6 - Solução

Intensidade de tráfego: $\rho = \lambda / \mu$ = 1.250/1.2.000 =0,625

Para n = 1, 2,...B, Pn será:

$p_1 = \rho . p_0 = 0{,}625 . p_0$

$p_2 = \rho^2 . p_0 = 0{,}391 . p_0$

$p_3 = \rho^3 . p_0 = 0{,}256 . p_0$

$p_4 = \rho^4 . p_0 = 0{,}163 . p_0$

$p_5 = \rho^5 . p_0 = 0{,}103 . p_0$

Pelo axioma da probabilidade total:

$p_0 + p_1 + p_2 + p_3 + p_4 + p_5 = 1$

$p_0 + 0{,}625 . p_0 + 0{,}391 . p_0 + 0{,}256 . p_0 + 0{,}163 . p_0 + 0{,}103 . p_0 = 1$

$p_0 = 1/(1 + 0{,}625 + 0{,}391 + 0{,}256 + 0{,}163 + 0{,}103) = 1/2{,}538 = 0{,}39$

Exemplo 8.6 - Solução

Logo,

p_0 = 1/(1 + 0,625 + 0,391 + 0,256 + 0,163 + 0,103) = 1/2,538 = 0,39

Substituindo p_0 nas equações de p_n, com n = 1, 2, ... 5 obtêm-se:

p_1= 0,625.0,39 = 0,247

p_2=0,391.0,39 = 0,152

...

p_5= 0,103.0,39 = 0,04

$$E[n] = \sum_{n=1}^{B} np_n = 1.0,247 + 2.\ 0,152 + ... + 5.0,04 = 1,31$$

Exemplo 8.6 - Solução

Como o número de buffers do sistema é limitado, os pacotes excedentes são descartados. Pode-se então calcular a taxa efetiva de chegada, isto é, os pacotes que efetivamente acessam o roteador como sendo:

$$\lambda' = \lambda.(1 - p_B) = 1.250.(1 - 0,103) = 1.121 \text{ pps}$$

Desta forma, a partir do cálculo de λ' pode-se calcular o número esperado de pacotes que serão perdidos com esta limitação do buffer:

$$\lambda - \lambda' = 1.250 - 1.121 = 129 \text{ pps.}$$

Exercícios

- Responda aos exercícios de teoria das filas do capítulo 8 do livro texto.