

Ciência da Computação e Sistemas de Informação PESQUISA OPERACIONAL (PO)

TEORIA DAS FILAS

Prof. Arthur arthur.battaglia@docente.unip.br 2021/2

Fonte: Introdução à Pesquisa Operacional – 5a edição – Eduardo Leopoldino de Andrade

INTRODUÇÃO

A Teoria das Filas trata de problemas de congestionamento de sistemas, cuja característica principal é a presença de "clientes" solicitando "serviços" de alguma maneira.

Um sistema de filas é composto de elementos que querem ser atendidos em um posto de serviço e que, eventualmente, devem esperar até que o posto esteja disponível.

Observação:

 Nesse contexto os "clientes" não são necessariamente pessoas, mas algum tipo de demanda que precisa ser atendida. Por exemplo, máquinas danificadas aguardando para serem consertadas.

INTRODUÇÃO

Quando há clientes solicitando serviços o atendimento é limitado por restrições próprias do sistema.

Dessa forma há a possibilidade de que esses clientes venham a formar uma fila, até que o serviço solicitado possa ser prestado.

Essas demandas ocorrem de forma irregular, isto é, num momento pode não haver demanda alguma e, no outro momento, surgem mais demandas do que a capacidade de atendimento do sistema.

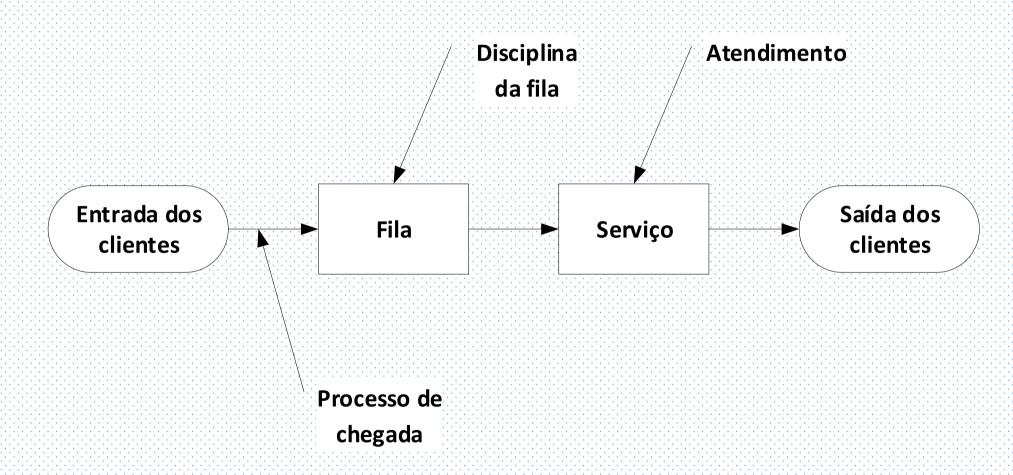
Essa irregularidade de comportamento é expressa em termos probabilísticos na Teoria das Filas.

INTRODUÇÃO

Quando o tamanho da fila ultrapassa o valor esperado (ou considerado normal) o sistema está entrando em congestionamento e, nessa situação, a qualidade e a produtividade do sistema caem e o custo total de operação tende a crescer sem controle.

Esse é o principal motivo que justifica o estudo dos problemas de congestionamentos.

COMPONENTES DE UM SISTEMA DE FILAS



Diversos fatores condicionam a operação de um sistema. Podem interferir tanto que o desempenho do sistema passa a ser função deles.

Esses fatores podem ser classificados em 4 categorias:

- Forma do atendimento
- Modo de chegada
- Disciplina da fila
- Estrutura do sistema

Vamos analisar cada um deles.

Formas de atendimento

Postos de atendimento, geralmente, são formados por pessoas, instalações e equipamentos. Por isso, existem diversos elementos passíveis de atuação por parte do administrador com o objetivo de melhorar o desempenho do sistema:

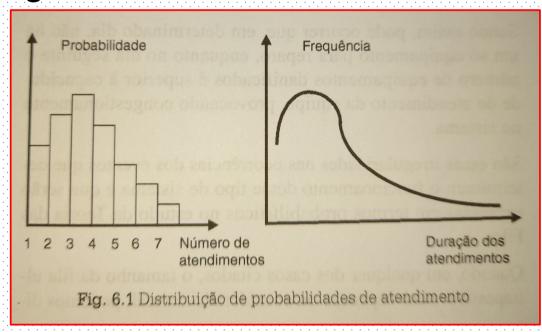
- Dimensionamento da capacidade
- Treinamento dos atendentes
- Rotinas administrativas
- Sistemas de informações
- Etc.

O resultado da interação desses fatores aparece, para o cliente, como o tempo gasto em cada atendimento ou como o número de atendimentos realizados.

Formas de atendimento

O primeiro passo no estudo de um sistema de filas é o levantamento estatístico do número de clientes atendidos por unidade de tempo, ou do tempo gasto em cada atendimento. Esse tempo pode ser regular ou aleatório.

A finalidade do levantamento estatístico é determinar a distribuição de probabilidades do número de atendimentos, ou da duração de cada atendimento:



Formas de atendimento

Além disso, mais dois fatores devem ser analisados na definição do regime de atendimento:

- Disponibilidade do serviço, pois nem todos os serviços estão disponíveis a qualquer tempo.
- Capacidade de atendimento simultâneo do sistema

Modo de chegada

A chegada de clientes ao sistema que mais interessa à administração ocorre, na maioria das vezes, de modo aleatório.

Por isso, é importante realizar um levantamento estatístico com a finalidade de descobrir se o processo de chegada pode ser caracterizado por uma distribuição de probabilidades.

Para que essa caracterização possa ser feita, o processo de chegadas tem que estar no "estado estacionário", ou seja, a distribuição de probabilidades que identifica o processo hoje será a mesma de amanhã.

Modo de chegada

Quando a distribuição de probabilidades de um evento varia com o tempo, o sistema está em "estado não estacionário" ou "transitório".

Exemplo:

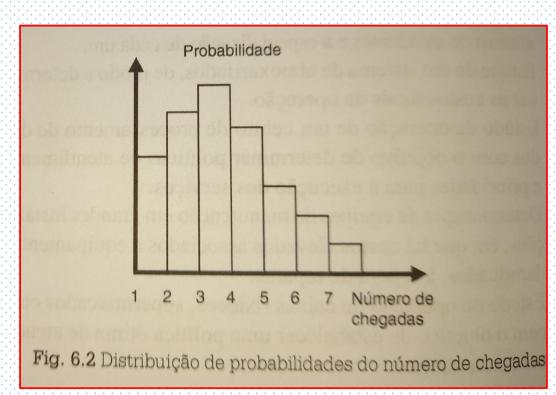
- Estacionário: fila de pessoas num banco num dia normal.
- Não estacionário: fila de pessoas num banco às vésperas de uma greve bancária.

Modo de chegada

A distribuição de probabilidades de um evento estacionário é diferente daquela

de um evento não estacionário.

Distribuição de probabilidades do número de chegadas:



Disciplina da fila

É um conjunto de regras que determinam a ordem em que os clientes serão atendidos.

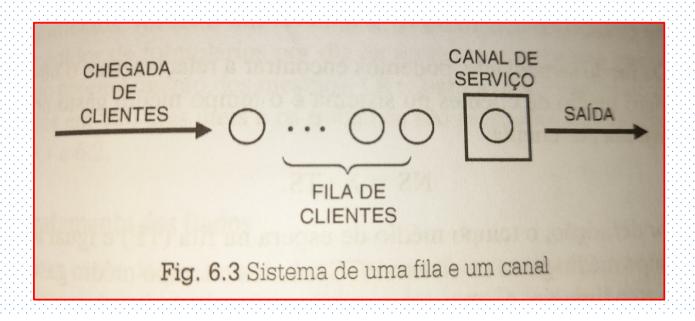
Esse atendimento pode ser feito:

- Por ordem de chegada (o 1º a chegar é o 1º a ser atendido).
- Por ordem inversa de chegada (o último a chegar é o 1º a ser atendido).
- Por prioridade de certas classes.

Estrutura do sistema

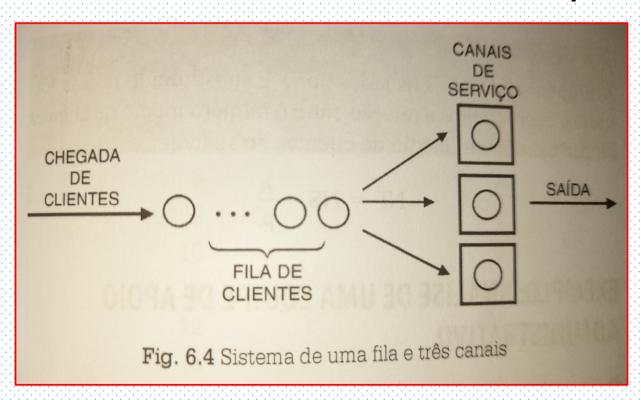
Os sistemas de filas podem ter estruturas variadas, e cada uma exige um estudo analítico diferente.

A estrutura mais simples é um sistema de uma fila e um canal:



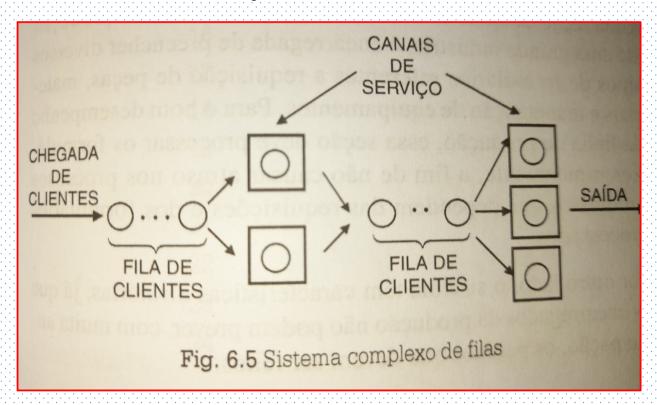
Estrutura do sistema

Outra estrutura é um sistema de uma fila e três canais em paralelo:



Estrutura do sistema

Outra estrutura é um sistema complexo de filas e canais em série:



A Notação de Kendall-Lee é uma maneira de descrever os sistemas para simplificação de sua análise, a partir da utilização de cinco características descritas da seguinte maneira: A/B/c/D/E. Onde:

- A Faz referência ao tipo de distribuição de probabilidade do processo de chegada ao sistema;
- B Refere-se ao tipo de distribuição de probabilidade do processo de atendimento;
- c Determina o número de postos de atendimento em paralelo;
- D Indica a capacidade física do sistema;
- E Faz menção à disciplina da fila

Fonte: http://www.abepro.org.br/biblioteca/TN_STP_243_410_31843.pdf

Os tipos de modelos de filas são definidos a partir da Notação de Kendall, que representa cada cadeia de filas pelos símbolos A, B, c, K, N, Z da seguinte forma:

A/B/c/K/N/Z

A: distribuição do tempo entre chegadas;-

B: disciplina de serviços;

c: número de servidores;

K: capacidade total de usuários no sistema;

N: número de usuários potenciais em uma população fonte;

Z: disciplina de atendimento.

Fonte: https://lamfo-unb.github.io/2018/07/06/Otimiza%C3%A7%C3%A3o-de-Filas/

Usualmente, quando não declarado a capacidade máxima do sistema (K), e o número de clientes a serem atendidos (N), estes serão considerados como ilimitados.

A disciplina de atendimento pode admitir duas politicas: First In First Out (primeiro que entra primeiro a sair), ou fila, é a estrutura de atendimento que bancos e hospitais utilizam pois sequem a ordem de chegada em que os primeiros a chegarem serão os primeiros a serem atendidos; Last In First Out (último que entra primeiro a sair) é a estrutura habitualmente chamada de pilha e amplamente utilizada na estocagem de produtos. Caso o modelo não declare qual a disciplina de atendimento, entende-se que será utilizada a disciplina FIFO.

Fonte: https://lamfo-unb.github.io/2018/07/06/Otimiza%C3%A7%C3%A3o-de-Filas/

De maneira geral:

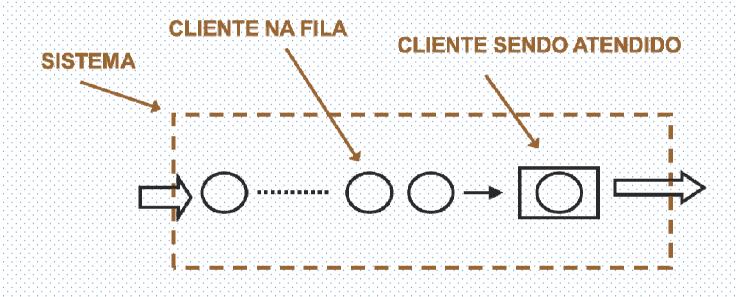
1/2/3/4/5/6

- 1 processo de chegada
- 2 processo de atendimento
- 3 número de atendentes
- 4 número máximo de clientes no Sistema
- 5 regra da fila
- 6 tamanho da população

Exemplo:

M/M/1

- 1 processo de chegada exponencial
- 2 processo de atendimento exponencial
- 3 número de atendentes = 1
- 4 regra da fila: geral
- 5 número máximo de clientes no sistema: ∞
- 6 tamanho da população:∞





λ - taxa de chegada

μ - taxa de

atendimento

IC – tempo médio entre chegadas $IC=1/\lambda$

TF - tempo médio na fila

NF - número médio de clientes na fila

TA - tempo médio de atendimento TA=1/μ

NA – número médio de clientes em atendimento

TS - tempo médio no sistema

NS – número médio de clientes no sistema



$$NS = NF + NA$$

$$TS = TF + TA$$

$$NA = \lambda / \mu = TA/IC$$

$$NS = NF + NA = NF + (\lambda / \mu) = NF + (TA/IC)$$

intensidade de tráfego $\rho = \lambda / \mu$

Fórmulas básicas:





Ciência da Computação e Sistemas de Informação PESQUISA OPERACIONAL (PO)

TEORIA DAS FILAS

Prof. Arthur arthur.battaglia@docente.unip.br 2021/2