PCS-2039
Modelagem e Simulação de Sistemas Computacionais

Graça Bressan

gbressan@larc.usp.br

© *Copyright* LARC 2008

Redes de Filas e Leis Operacionais

Graça Bressan
LARC-PCS/EPUSP

© *Copyright* LARC 2008

Introdução

- O objetivo é apresentar a solução de sistemas que envolvem múltiplas filas.
- O enfoque dado é o de aplicar as soluções encontradas em sistemas reais;
- As leis operacionais de sistemas de filas são introduzidas;
- Se não existe soluções exatas, soluções numéricas aproximadas são fornecidas.

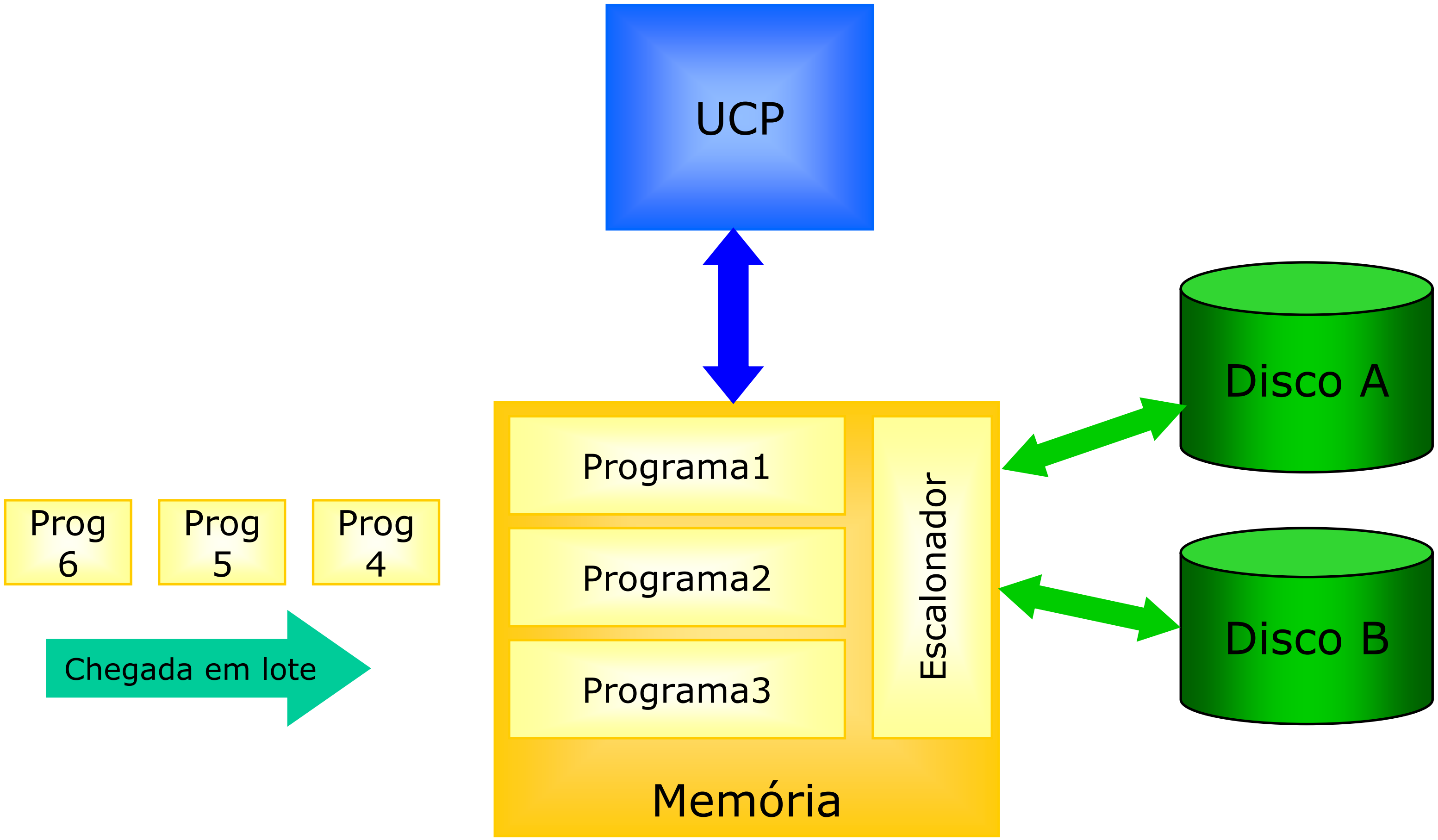
Tópicos a serem cobertos:

- Inicialmente são apresentadas as condições para se ter a solução na forma de produto para as redes de filas;
- A solução de sistemas de filas é apresentada;
- A aplicação de tais soluções em sistemas reais é discutida.

Redes de Filas

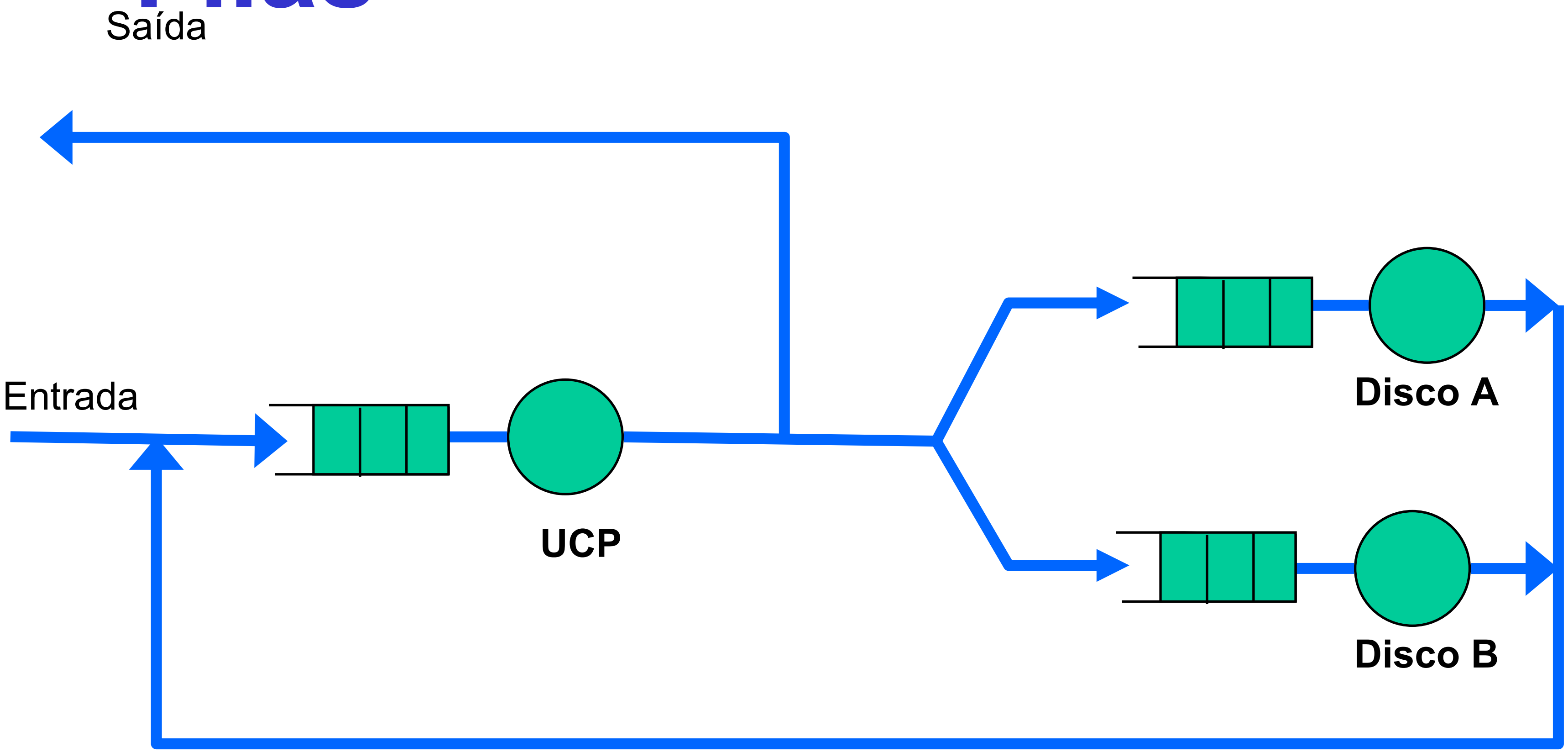
- Os sistemas de filas são classificados em:
  - **redes abertas**;
  - **redes fechadas**;
  - **redes mistas**.

Execução de programas em Lotes

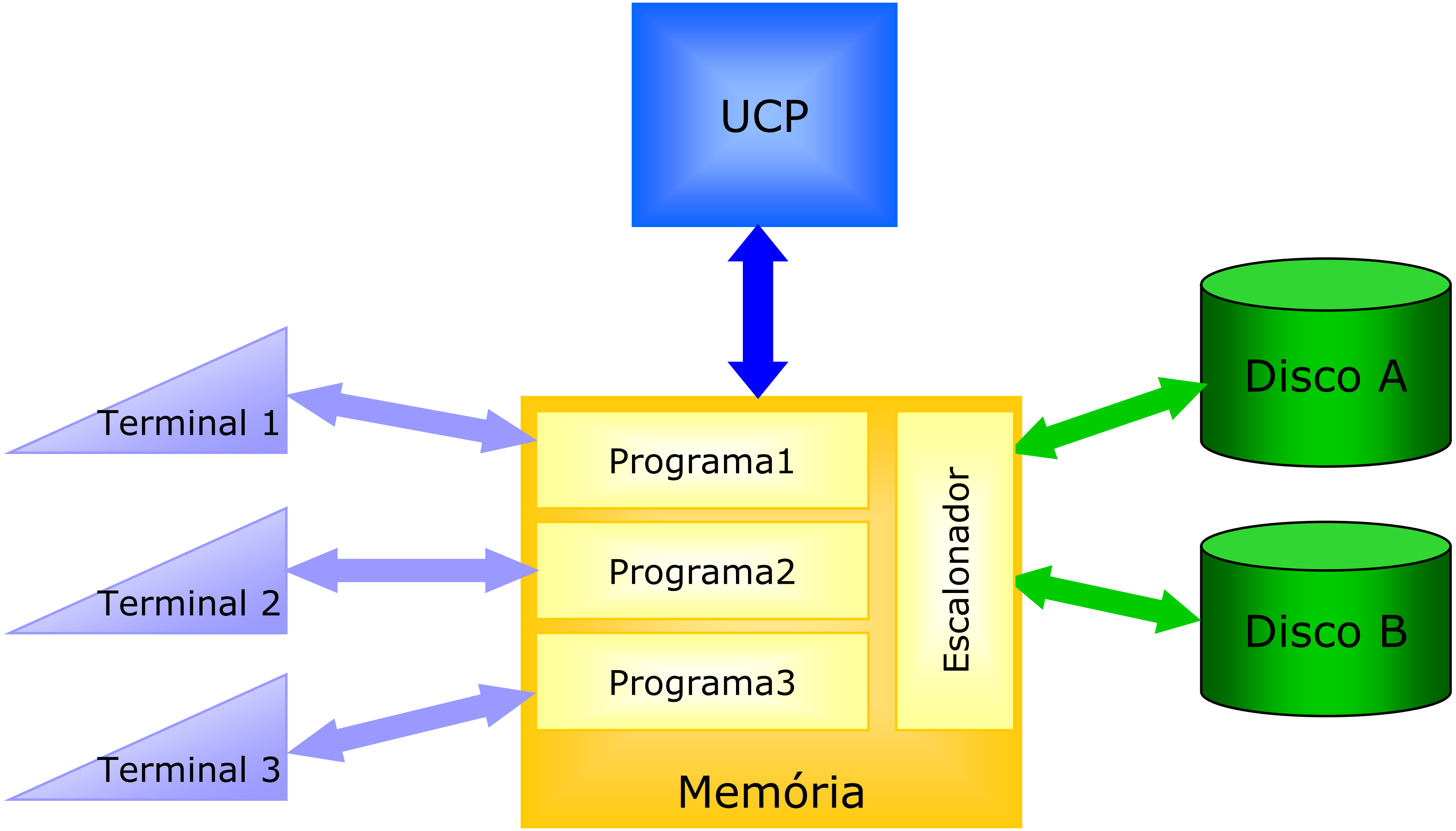


© *Copyright* LARC 2008

Exemplo de Rede Aberta de Filas

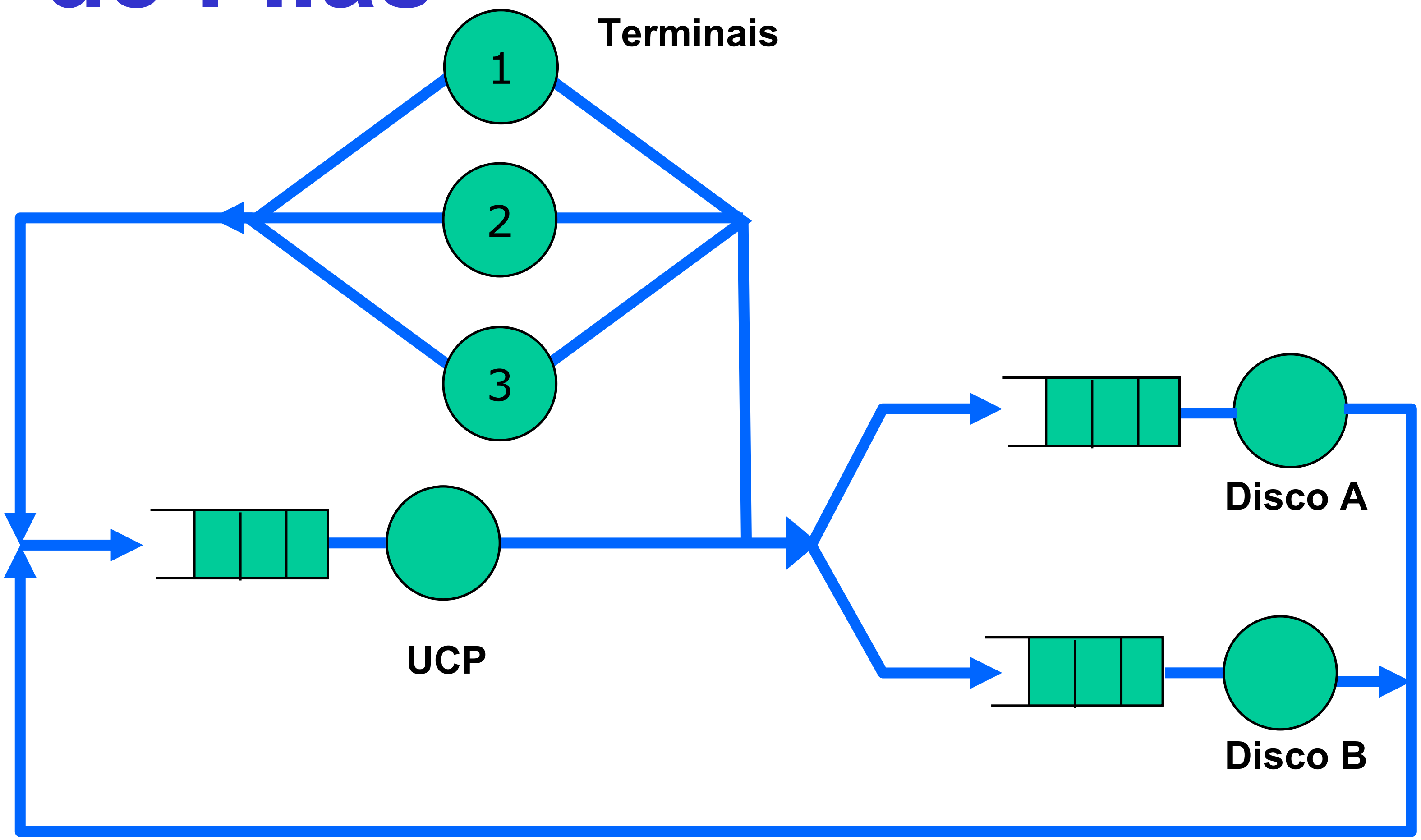


Execução de programas através de terminais

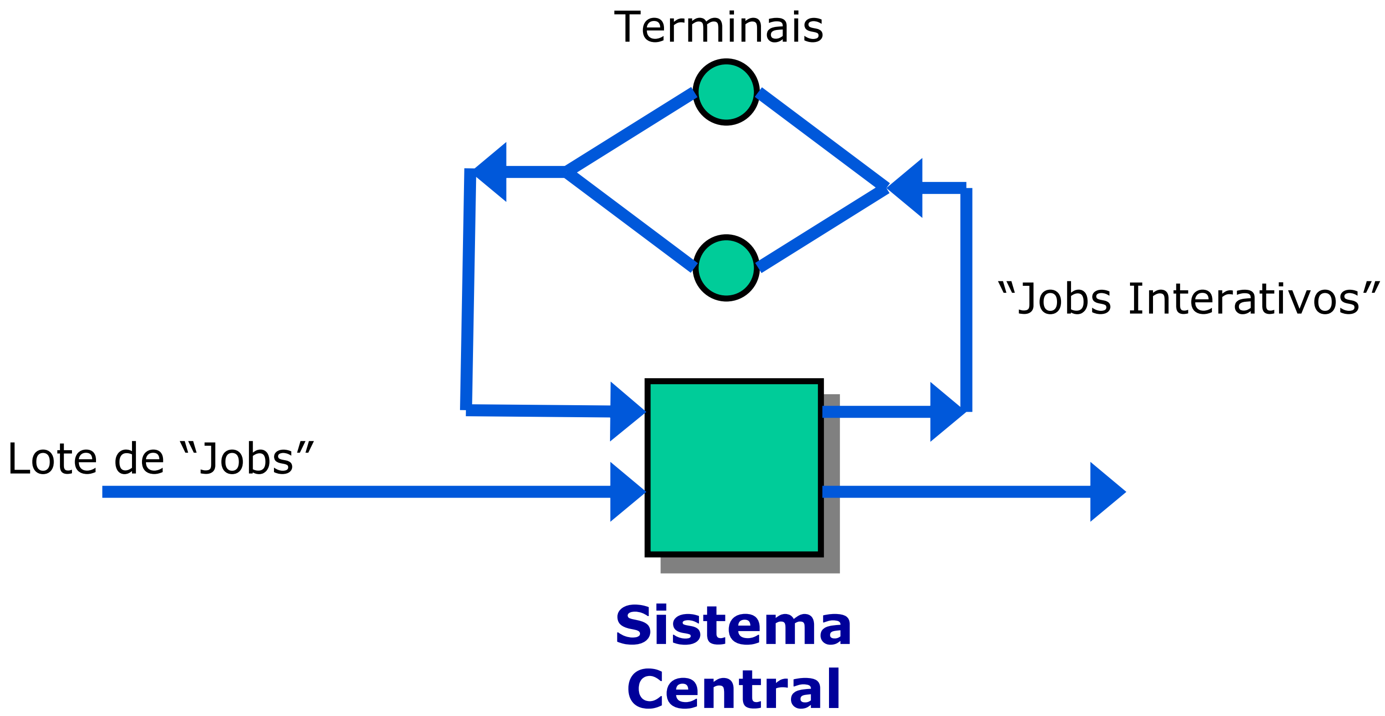


© *Copyright* LARC 2008

Exemplo de Rede Fechada de Filas

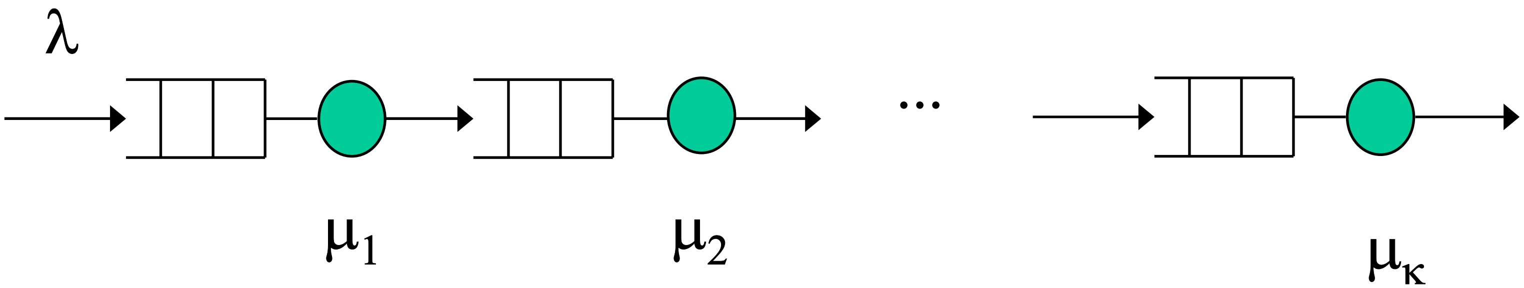


Exemplo de Rede Mista de Filas



Associação série de filas:

- A forma mais simples de uma associação de filas é uma associação em série.



Associação série de filas:

- Supondo que as filas estão em equilíbrio, tem-se:
  a taxa de chegada é igual a taxa de saída em todas as filas (λ).

- Neste caso o fator de utilização da i-ésima fila é:

$$\rho_i = \lambda/\mu_i$$

Associação série de filas:

- A probabilidade de se ter n_i usuários na fila i é:

  $$p_i(n_i) = (1 - \rho_i)*\rho_i^{ni}$$

- A probabilidade conjunta de se ter n_i usuários na fila i, para i = 1,2,...,M é dada por:

  $$P(n_1,n_2,...,n_M) = p_1(n_1)p_2(n_2)...p_M(n_M)$$

- Exemplo: Probabilidade de ter 2 usuários na fila 1, 4 usuários na fila 2, 3 usuários na fila 3 é $P(2,4,3) = p_1(2)p_2(4)\ p_M(3)$

Leis Operacionais

- Leis operacionais em redes de filas são relações entre as grandezas *diretamente mensuráveis* destes sistemas.
- Algumas grandezas diretamente mensuráveis:
  - $\mathbf{A}_i$: Número de chegadas;
  - $\mathbf{C}_i$: Número de partidas;
  - $\mathbf{B}_i$: Tempo ocupado.

Leis Operacionais

- Valores derivados das grandezas mensuráveis:
  - $\lambda_\mathbf{i}$: Taxa de chegada = A_i/T;
  - $\mathbf{X_i}$: Vazão = C_i/T;
  - $\mathbf{U_i}$: Fator de Utilização = B_i/T;
  - $\mathbf{S_i}$: Tempo médio de serviço = B_i/C_i.

Leis Operacionais

- Observe que estas grandezas podem assumir diferentes valores em diferentes períodos de observação. Porém, existem certas relações que permanecem válidas para cada período de observação.
- Estas relações são as **Leis Operacionais** dos sistemas de filas:
  - Lei de utilização;
  - Lei de fluxo;
  - Lei de Little;
  - Lei to tempo de resposta;
  - Lei to tempo de resposta interativo;
  - Lei do gargalo.

© *Copyright* LARC 2008

Lei da Utilização

- Dado um número de partidas $\mathbf{C_i}$, um tempo de ocupação $\mathbf{B_i}$, de um sistema de filas **i** durante um intervalo de observação **T**, a seguinte relação é válida:

  $U_i = (B_i/T) = (C_i/T)*(B_i/C_i)$ ou

  $U_i = X_i*S_i$.

Lei de Fluxo

- Esta lei correlaciona a vazão global do sistema com as vazões de seus sub-sistemas;
- Numa rede aberta de filas, o número de usuários partindo da rede na unidade de tempo define a sua vazão;
- Numa rede fechada, a taxa com que se cicla no sistema define a sua vazão.

Lei de Fluxo

- Se num dado período **T** de observação, o número de usuários que entraram é igual ao número de usuários que saíram do sistema, isto é:

  $$\mathbf{A_i = C_i}$$

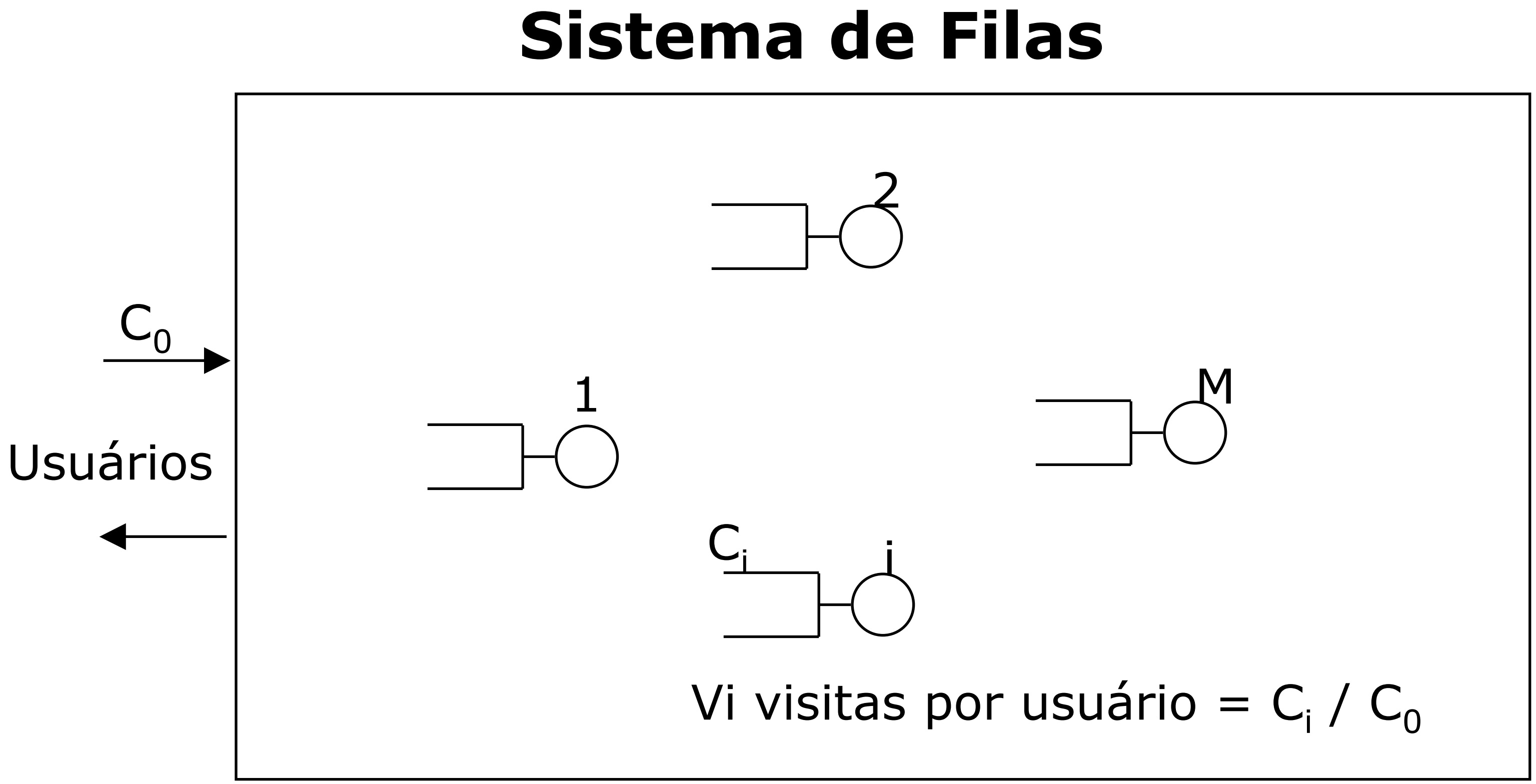
  pode-se dizer que este sistema satisfaz a hipótese de **fluxo balanceado**.
- Se o intervalo de observação é grande, C_i tende a se aproximar de A_i.

Lei de Fluxo

- Suponha que cada usuário faça $\mathbf{V_i}$ visitas ao i-ésimo sub-sistema. Se o fluxo deste sistema é balanceado, o número de usuários $\mathbf{C_0}$ que entram ou saem, e o número de visitas ao i-ésimo sub-sistema estão relacionados pela seguinte expressão:

$$\mathbf{C_i = C_0 * V_i \quad \text{ou} \quad V_i = C_i / C_0}$$

Lei de Fluxo


Sistema de Filas
C_0
Usuários
1
2
M
C_i
i
Vi visitas por usuário = C_i / C_0

Lei de Fluxo

- A variável **Vi** representa a taxa de visitas ao sub-sistema **i** para cada usuário.
- A vazão global do sistema durante este período de observação é dada por:

X : Vazão do Sistema = C_0/T

Lei de Fluxo

- A vazão do i-ésimo sub-sistema é dada por:

  $\mathbf{X_i}$: Vazão do sub-sistema i = C_i/T = $(C_i/C_0)*(C_0/T)$

  isto é,

  $\mathbf{X_i = X*V_i}$ **Esta é a lei de Fluxo.**

Lei de Fluxo

- Combinando a lei de utilização com a lei de Fluxo tem-se:

  $U_i = X_i{*}S_i = X{*}V_i{*}S_i$

  ou

  $\mathbf{U_i = X{*}D_i}$ onde $\mathbf{D_i = V_i{*}S_i}$, e é chamado de demanda total sobre o i-ésimo sub-sistema.

- O sub-sistema que possuir o maior $\mathbf{D_i}$ será o gargalo do sistema.

Lei de Fluxo

- A taxa de visitas é uma das maneiras de se especificar o roteamento dos usuários numa rede de filas.
- Uma outra forma é se especificar as probabilidades de transição $\mathbf{p_{ij}}$ de um usuário ao terminar o serviço em **i** se mover para **j**.

Lei de Fluxo

- Num sistema com o fluxo balanceado tem-se:

$$C_j = \sum_{i=0}^{M} C_i p_{ij}$$

  - $\mathbf{p_{i0}}$ é a probabilidade do usuário deixar o sistema tendo terminado o serviço em i;
  - $\mathbf{C_0}$ representa o número de usuários que entraram ou sairam do sistema;

Lei de Fluxo

- Dividindo ambos os lados da relação por $\mathbf{C_0}$ tem-se:

$$V_j = \sum_{i=0}^{M} V_i p_{ij}$$

Como a tarefa de um usuário termina ao sair do sistema, então $\mathbf{V_0 = 1}$.

- As duas equações anteriores permitem que se obtenha as relações entre $\mathbf{V_i}$ e $\mathbf{p_{ij}}$.

Lei de Little

- A lei de Little já foi vista e é expressa por:

  $Q_i = \lambda_i * R_i$, onde Q_i é o número de usuários em **i** e **Ri** é o tempo gasto em **i**;

- Para o caso de sistemas com fluxo balanceado pode-se escrever:

  $Q_i = X_i * R_i$, **onde** X_i **é a vazão em i.**

Lei do Tempo de Resposta

- Todo sistema de **"Time-Sharing"** pode ser dividido em dois sub-sistemas: os **Terminais** e o **Sistema Central**;
- A lei de Little pode ser aplicada para qualquer destes sub-sistemas desde que ele possua fluxo balanceado:

Q = X*R, para o sistema Central.

Lei do Tempo de Resposta

- Conhecendo-se o número de usuários em cada um dos sub-sistemas do Sistema Central, pode-se escrever:

  $Q = Q_1 + Q_2 + \ldots + Q_{M,}$, como

  $Q_i = X_i * R_i$, tem-se:

Lei do Tempo de Resposta

- $XR = X_1R_1 + X_2R_2 + \ldots + X_MR_M$, dividindo-se ambos os lados por X e usando a lei do fluxo, tem-se:

$$R = \sum_{i=1}^{M} R_i V_i$$

- Esta é a **lei do Tempo de Resposta**.

Lei do Tempo de Resposta Interativo

- Num sistema interativo o tempo em que um usuário gasta pensando antes de fornecer uma nova requisição ao sistema é **Z**;
- Se o tempo de resposta do sistema é **R**, então o tempo de um ciclo completo pelo sistema é :

  (R + Z).

Lei do Tempo de Resposta Interativo

- Cada usuário produz **T/(R+Z)** requisições ao sistema num intervalo de tempo **T**;
- Num sistema com **N** usuários a vazão do sistema será dada por:

  $X = \{N[T/(R+Z)]/T\} = N/(R+Z)$ ou

  R = (N/X) - Z, esta é a **lei do Tempo de Resposta Interativo**.

Análise de Gargalo

- Num sistema o dispositivo gargalo é aquele que possui a maior demanda de serviço $\mathbf{D_i}$, ou equivalentemente, o maior fator de utilização $\mathbf{U_i}$;
- Suponha que o elemento gargalo seja **b**. Isto implica em $\mathbf{D_b = D_{max}}$, onde D_{max} é o maior valor entre $D_1, D_2, \ldots, D_M$;

Análise de Gargalo

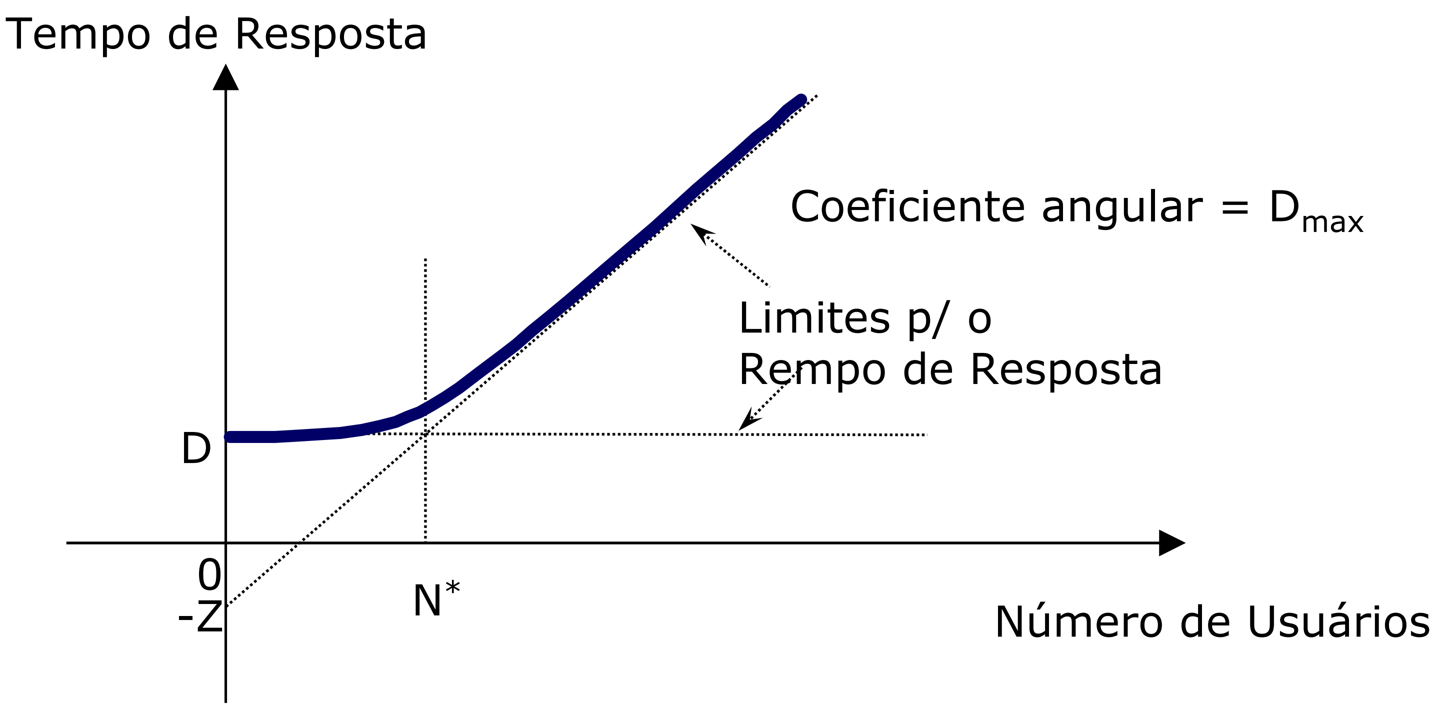
- A vazão e o tempo de resposta do sistema são limitados pelos seguintes valores:

X(N) <= min { (1/Dmax),(N/(D+Z)) }

R(N) >= max{ D,(NDmax - Z) }

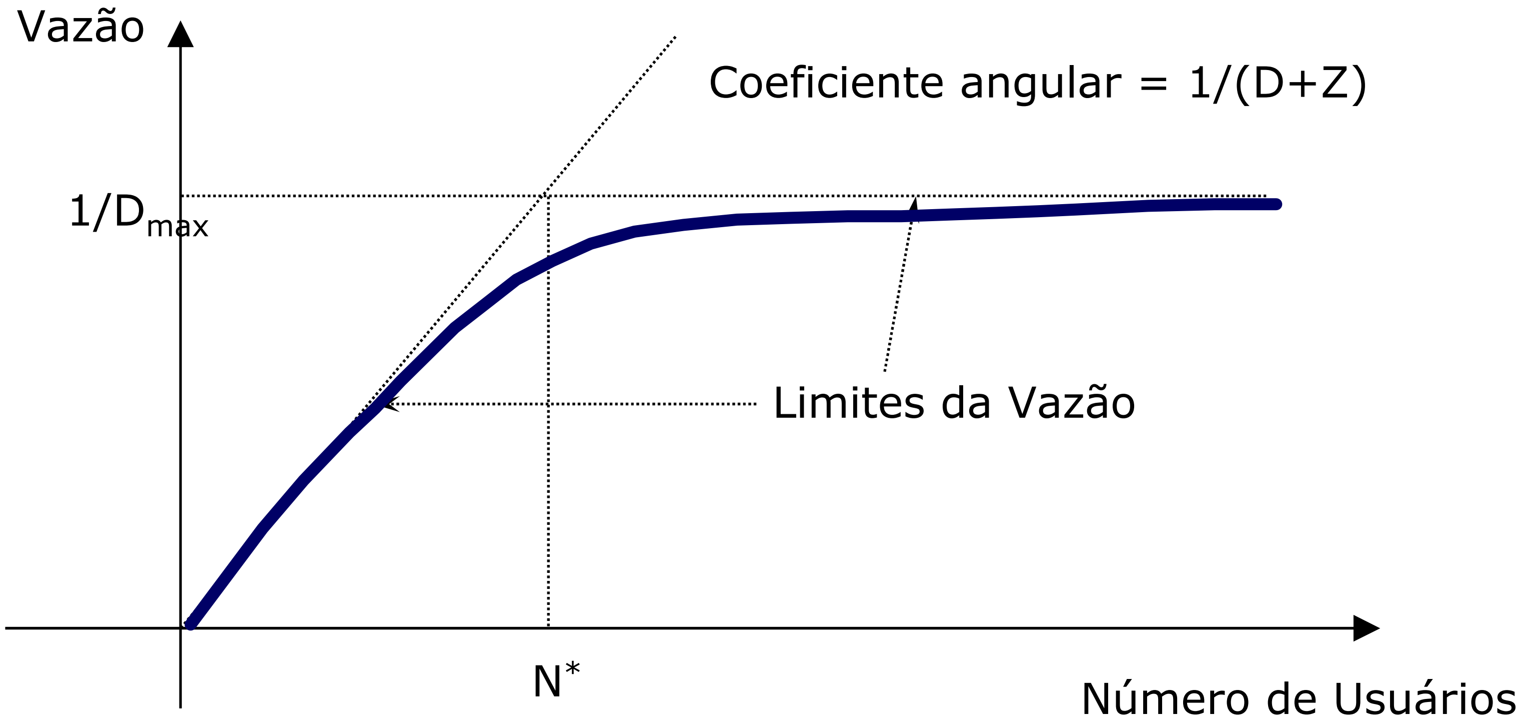
onde $D = \Sigma D_i$ é a soma da demanda de serviço de todos os sub-sistemas exceto os terminais. Estas inequações são chamadas de limites assintóticos.

Análise de Gargalo



Limites do Tempo de Resposta

Análise de Gargalo



Limites para a Vazão do Sistema

© *Copyright* LARC 2008

Análise de Gargalo

- O ponto de interseção das duas retas limites é chamado **"joelho"** do sistema e é dado por:

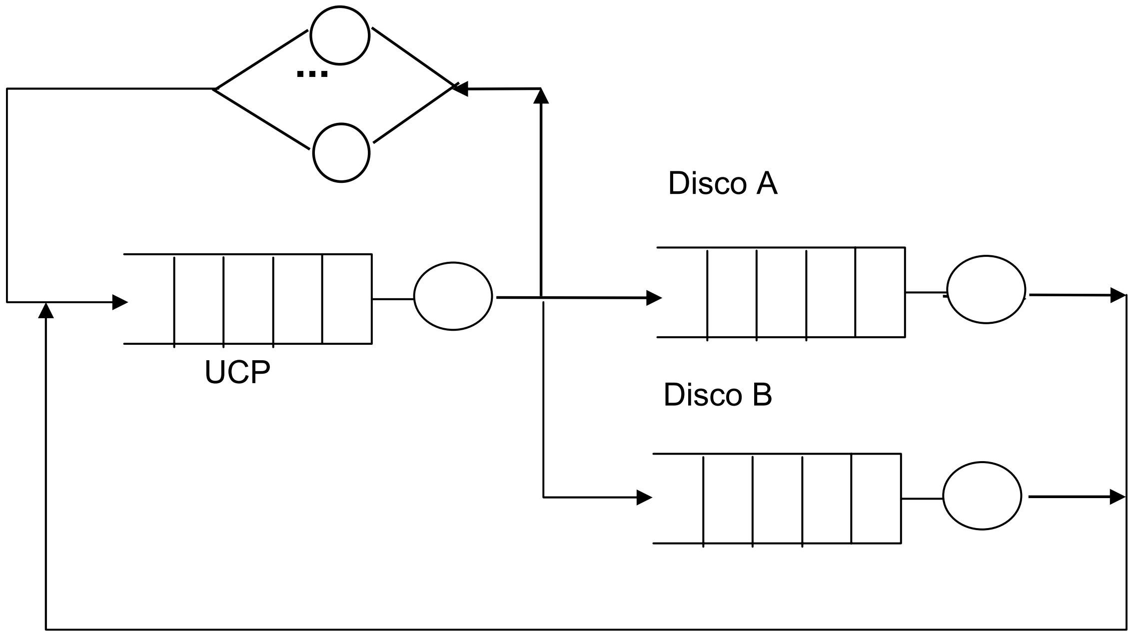
  $\mathbf{N^* = (D+Z)/D_{max}}$, onde N^* é o número de usuários no "joelho" do sistema.

- Se o número de usuários no sistema for maior que **N***, pode-se dizer com certeza que existirá espera em algum lugar do sistema.

Exercício

- Em um sistema de timesharing com 2 discos (para usuários e sistema), após o uso da UCP, a probabilidade de um programa utilizar o disco A é de 0,80, de utilizar o disco B é de 0,16 e de utilizar os terminais é de 0,04. O tempo que o usuário fica pensando é de 5 segundos, os tempos de serviço dos discos A e B são de 30 e 25 ms respectivamente, e o tempo médio de serviço por visita à UCP é 40 mseg. Considerando que com 20 usuários a utilização do disco A é 60%, realize a análise do sistema.

Exercício



Exercício

- Perguntas:
  a) Para cada programa, qual é a taxa de visitas à UCP, disco A e disco B?
  b) Para cada dispositivo, qual é a demanda total de serviço?Qual é a utilização da UCP e do disco B?
  c) Qual é o tempo médio de resposta?
  d) Qual dispositivo é o gargalo do sistema?
  e) Qual é o tempo de resposta mínimo do sistema (independente do número de usuários)?
  f) Qual é a utilização máxima do disco A (independente do número de usuários)?
  g) Qual é a vazão máxima deste sistema (independente do número de usuários)?

Exercício

h) Qual mudança na velocidade da UCP é recomendada para obter-se um tempo de resposta de 10 segundos com 25 usuários? Também serão necessários discos A e B mais rápidos?

i) Escreva as expressões para os limites assintóticos da vazão e do tempo de resposta e desenhe os gráficos correspondentes.

Resp.:a) 25; 20 e 4 b) 1; 0,6 e 0,1 c) 1 e 0,1 d) 15 s e) UCP f) 1,7 g) 0,60 h) 1 prog/s; i) R≥max{D, N Dmax-Z}=> Dmax≤ 0.6 e UCP 40% mais rápida j) $X \leq \min\{ N / 6.7 ; 1\}$; $R \geq \max\{1.7, N-5\}$.

Fim do Módulo
Leis Operacionais