

Módulo 5 Análise do Valor Médio (AVM)

Graça Bressan LARC-PCS/EPUSP



AVM para Redes Abertas

- Hipóteses:
 - Assumindo Redes Abertas de Filas;
 - Supondo que todas as estações de serviço possuem servidores com capacidade fixa (CF) de processamento ou servidores que não apresentam espera em filas (infinitos servidores) (IS);
- Objetivo: determinar o tempo de resposta da rede de filas



 Para todos os servidores de capacidade fixa, tem-se:

$$R_i = S_i(1+Q_i)$$

• Um usuário ao chegar na estação i, em média, irá encontrar Q_i usuários à sua frente. Portanto, ele deverá esperar o atendimento de Q_i usuários, que levam em média S_i, antes de começar o seu próprio atendimento que levará também S_i.



 Assumindo que o fluxo é balanceado, temse:

$$X = \lambda$$

A vazão da i-ésima estação é dada por:

$$X_i = X V_i$$



A utilização da estação i é:

$$U_i = X_i S_i = X V_i S_i = \lambda D_i$$

O número médio de usuários na estação i é:

$$Q_i = X_i R_i = X_i S_i (1+Q_i) = U_i (1+Q_i)$$
 ou $Q_i = U_i / (1-U_i)$

note que esta equação é igual à do sistema M/M/1.



 Substituindo-se este na equação inicial de R_i tem-se:

$$R_i = S_i/(1-U_i).$$

No caso de infinitos servidores tem-se:

$$R_i = S_i e$$

 $Q_i = R_i X_i = S_i X V_i = X D_i = U_i$



- Algoritmo AVM para redes abertas:
 - Entradas:
 - X: Taxa externa de chegada, Vazão do sistema;
 - S_i: Tempo de serviço por visita à estação i;
 - V_i: número de visitas à estação i;
 - M: número de estações na rede;



- Saídas:
 - Q_i: número médio de usuários na estação i;
 - R_i: Tempo de resposta na estação i;
 - R: Tempo de resposta do Sistema;
 - U_i: Utilização da estação i;
 - N: Número médio de usuários no Sistema.



- Cálculos:
 - Demanda total de serviço: D_i = S_iV_i
 - Utilização dos dispositivos: U_i = XD_i
 - Vazão dos Dispositivos: X_i = XV_i;
 - Tempo de resposta dos Dispositivos:

$$R_i = S_i/(1-U_i)$$
 Servidores de capacidade fixa;

$$R_i = S_i$$
 Infinitos servidores



- Cálculos (continuação):
 - · Número médio de usuários nos dispositivos:

$$Q_i = U_i/(1-U_i)$$
 Servidores de capacidade fixa;

$$Q_i = U_i$$
 Infinitos Servidores



- Cálculos (continuação):
 - · Tempo médio de Resposta do Sistema:

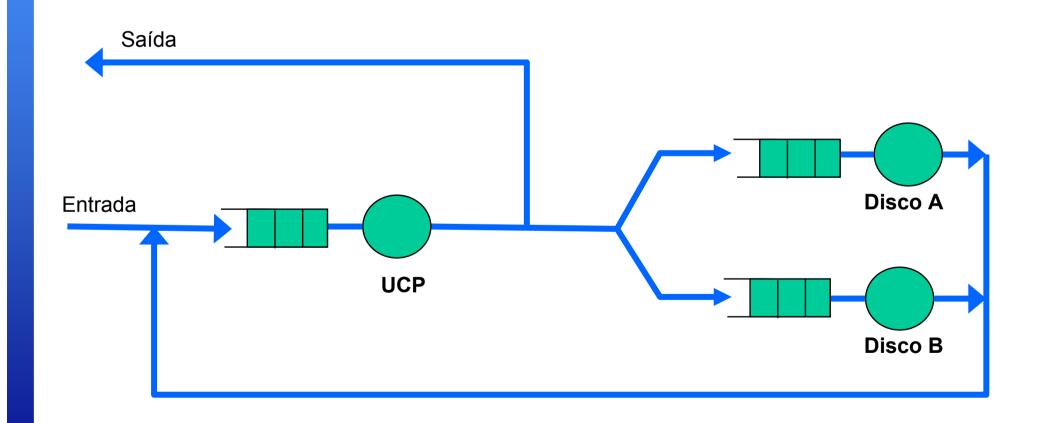
$$R = \sum_{i=1}^{M} R_i V_i$$

· Número médio de usuários no Sistema:

$$N = \sum_{i=1}^{M} Q_{i}$$



AVM Exemplo: Modelo de um Servidor de Arquivos





- O modelo representa um sistema distribuído com 6 clientes fazendo requisições ao servidor de arquivos;
- O servidor de arquivos possui 1 UCP e 2 discos;
- Medidas realizadas produziram os seguintes dados:



Medidas:

- Intervalo de Observação: 3 600 segundos;
- Número de requisições dos clientes: 10 800;
- Tempo de UCP ocupada: 1 728 segundos;
- Tempo de Disco A ocupado: 1 512 segundos;
- Tempo de Disco B ocupado: 2 592 segundos;
- Número de visitas ao disco A: 75 600 acessos;
- Número de visitas ao disco B: 86 400 acessos.



- Cálculos: Baseado nas medidas feitas pode-se calcular as seguintes grandezas:
 - Vazão do Sistema:
 X = 10800/3600 = 3 requisições por cliente;
 - V_A = 75600/10800 = 7 visitas por cliente ao disco A;
 - V_B = 86400/10800 = 8 visitas por cliente ao disco B;



- Cálculos (continuação):
 - V_{UCP} = 1 + 7 + 8 = 16 visitas por cliente à UCP;
 - D_{UCP} = 1728/10800 = 0,16 segundos de UCP por cliente;
 - D_A = 1512/10800 = 0,14 segundos de disco A por cliente;
 - D_B = 2592/10800 = 0,24 segundos de disco B por cliente (Gargalo do sistema);



- Cálculos (continuação):
 - S_{UCP} = 0,16/16 = 0,01 segundos por visita à UCP;
 - S_A = 0,14/7 = 0,02 segundos por visita ao disco A;
 - S_B = 0,24/8 = 0,03 segundos por visita ao disco B.



- Com estas grandezas calculadas temos todos os parâmetros necessários para resolver o modelo do servidor de arquivos:
 - Cálculo do fator de utilização de cada dispositivo do sistema:

•
$$U_{UCP} = XD_{UCP} = 3x0,16 = 0,48$$

•
$$U_{\Delta} = XD_{\Delta} = 3x0,14 = 0,42$$

•
$$U_B = XD_B = 3x0,24 = 0,72$$



Cálculo do tempo de resposta de cada dispositivo:

•
$$R_{UCP} = S_{UCP}/(1-U_{UCP}) = 0.01/(1-0.48) = 0.0192$$
 segundos;

•
$$\mathbf{R_A} = \mathbf{S_A}/(1 - \mathbf{U_A}) = 0.02/(1 - 0.42) = 0.0345$$
 segundos;

•
$$R_B = S_B/(1 - U_B) = 0.03/(1 - 0.72) = 0.107$$
 segundos;

Cálculo do tempo de resposta do Sistema:

$$\mathbf{R} = \Sigma \mathbf{V_i} \mathbf{R_i} = 16x0,0192+7x0,0345+8x0,107 = 1,406$$
 segundos.



- Este modelo pode ser utilizado para responder diversas questões, tais como:
 - Qual é o novo desempenho do sistema se:
 - a) tivermos 8 clientes?
 - b) usarmos um Cache para o disco B com uma taxa de acerto de 50%, embora ele aumente a sobrecarga da UCP em 30% e o tempo de serviço do disco B em 10%?
 - c) utilizarmos um servidor com um disco (disco A) e dirigirmos todos os acessos a ele?



- a) Sistema com 8 clientes:
 - Supondo que o cliente médio possua as mesmas características, pode-se dizer que a taxa de chegada vai crescer na proporção de 8:6;

X = 3 * (8/6) = 4 requisições/segundo;



 Os outros parâmetros do sistema podem ser obtidos:

•
$$U_{UCP} = XD_{UCP} = 4*0,16 = 0,64$$

•
$$U_A = XD_A = 4*0,14 = 0,56$$

•
$$U_B = XD_B = 4*0,24 = 0,96$$

•
$$R_{UCP} = S_{UCP}/(1 - U_{UCP}) = 0.01/(1 - 0.64) = 0.0278$$
 segundos;

 \cdot **R**_A = **S**_A/(1 - **U**_A) = 0.02/(1 - 0.56) = 0.0455 segundos;



- Outros parâmetros (cont.):
 - $R_B = S_B/(1 U_B) = 0.03/(1 0.96) = 0.75$ segundos;
 - Tempo de Resposta do Sistema:
 R = 16*0,0278+7*0,0455+8*0,75 = 6,76 segundos;
- Se o número de clientes vai de 6 para 8 o tempo de resposta do sistema degrada 6,76/1,406 = 4,8 vezes.



- b) A segunda alteração implica numa mudança do número de visitas ao disco B:
 - V_B = 0,5*8 = 4 visitas ao disco B com Cache e taxa de acerto de 50%;
 - $S_{UCP} = 1.3*0.01 = 0.013 ==> D_{UCP} = 0.208$ segundos (Sobrecarga de 30% na UCP);
 - $S_B = 1.1*0.03 = 0.033 ==> D_B = 4*0.033 = 0.132 segundos (Aumento de 10%).$



- A análise do sistema modificado é a seguinte:
 - $U_{UCP} = XD_{UCP} = 3*0,208 = 0,624$
 - $U_A = XD_A = 3*0,14 = 0,42$
 - $U_B = XD_B = 3*0,132 = 0,396$
 - $R_{UCP} = S_{UCP}/(1 U_{UCP}) = 0.013/(1 0.624) = 0.0346 \text{ segundos};$



- Análise do sistema modificado(cont.):
 - $\mathbf{R}_{A} = \mathbf{S}_{A}/(1 \mathbf{U}_{A}) = 0.02/(1 0.42) = 0.0345$ segundos;
 - $R_B = S_B/(1 U_B) = 0.033/(1 0.396) = 0.0546$ segundos;
- Tempo de Resposta do Sistema:
 - $\mathbf{R} = 16*0,0346+7*0,0345+4*0,0546 = 1,013$ segundos.



AVM

- Análise do sistema modificado (conclusão):
 - Se utilizarmos Cache para o disco B o tempo de resposta do sistema melhorará em (1,406 -1,013)/1,406 = 28%.



- c) Terceira modificação: implica em ajustes em **V**_A e **V**_B:
 - $V_{B} = 0$;
 - $V_{A} = 7 + 8 = 15$
 - D_{UCP} = 0,16 segundos (como antes);
 - $D_A = 15*0,02 = 0,3 \text{ segundos};$
 - $U_{UCP} = XD_{UCP} = 3*0,16 = 0,48$
 - $U_A = XD_A = 3*0.3 = 0.90$



c) Cont.

- $R_{UCP} = S_{UCP}/(1 U_{UCP}) = 0.01/(1 0.48) = 0.0192 segundos;$
- $R_A = S_A/(1 U_A) = 0.02/(1 0.90) = 0.2$ segundos;
- $\mathbf{R} = 16*0,0192+15*0,2 = 3,31 \text{ segundos};$
- Neste caso o tempo de resposta degrada por um fator de 3,31/1,406 = 2,35.



Estudo de Caso 2: Loja de Computadores On-line

 Voltando ao estudo de caso do módulo anterior, iremos agora calcular o tempo de resposta correspondente à vazão Xmax=70,92 transações/seg

 Utilizando o algoritmo de Análise do Valor Médio para sistemas abertos obtemos o tempo de resposta total R=0,2414 segundos como mostra a tabela a seguir.



Algoritmo de Análise do Valor Médio para Redes Abertas

		I		Γ	ı	
Dispositivos	Demandas do Cluster	Demanda de cada dispositivo	Utilização de cada dispositivo	Número médio de acessos ao dispositivo	Tempo médio de resposta de cada dispositivo	Tempo médio de resposta do cluster
	D _i (miliseg/ trans)	Di/numdisp (seg/ trans)	Ui = XD _i	Qi= Ui/(1-Ui) ou Q _i = U _i	Ri* Vi= Di/(1-Ui) ouRi* Vi= D _i	Ri* Vi= Di/(1- Ui) ou Ri* Vi= D _i
UCP do Servidor WEB	5,2	0,00173333	0,1124	0,1267	0,0020	0,0059
Disco do Servidor Web	9,5	0,00316667	0,2054	0,2585	0,0040	0,0120
UCP do Servidor de Aplicações	14	0,007	0,4540	0,8316	0,0128	0,0256
Disco do Servidor de Aplicações	10	0,005	0,3243	0,4799	0,0074	0,0148
UCP do Servidor de Base de Dados	14,1	0,0141	0,9145	10,6995	0,1650	0,1650
Disco do Servidor de Base de Dados	7,75	0,00775	0,5027	1,0107	0,0156	0,0156
LAN 1	0,49	0,00049	0,0318	0,0318	0,0005	0,0005
LAN 2	0,53	0,00053	0,0344	0,0344	0,0005	0,0005
LAN 3	0,38	0,00038	0,0246	0,0246	0,0004	0,0004
Link T3	1,2	0,0012	0,0778	0,0778	0,0012	0,0012
Tempo de Resposta					R =	0,2414



Análise do Valor Médio Rede Fechada

- Considere uma rede fechada com N usuários, o tempo de resposta da i-ésima estação é dado por:
 - Ri(N) = Si[1 + Qi(N-1)] onde Qi(N1) é o número médio de usuários na estação i quando existem N-1 usuários no sistema;
 - O desempenho do sistema para N usuários pode ser calculado em função do desempenho com N-1 usuários.



Análise do Valor Médio Rede Fechada

- Considere uma rede fechada com N usuários, o tempo de resposta da i-ésima estação é dado por:
 - R_i(N) = S_i[1 + Q_i(N-1)] onde Q_i(N1) é o número médio de usuários na estação i quando existem N-1 usuários no sistema;
 - O desempenho do sistema para N usuários pode ser calculado em função do desempenho com N-1 usuários.



AVM Rede Fechada

- Em função das leis operacionais pode-se calcular os parâmetros do sistema;
- Conhecendo-se o tempo de resposta de cada estação pode-se calcular o tempo de resposta do Sistema por:

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



AVM Rede Fechada

A vazão do Sistema é dada por:

$$X(N) = N/(R(N) + Z)$$

 A vazão individual de cada estação é dada por:

$$X_i(N) = X(N)^*V_i$$

 O número médio de usuários em cada estação é dada por:

$$Q_{i}(N) = X_{i}(N)*R_{i}(N) = X(N)*V_{i}(N)*R_{i}(N)$$



AVM Rede Fechada

- As equações acima são válidas para servidores de capacidade fixa;
- Para estações com infinitos servidores (SI), o tempo de resposta da estação é dado por:

$$R_i(N) = S_i$$

as outras equações da vazão e do número de usuários são válidas;



AVM Rede Fechada

- Algoritmo AVM:
 - Entrada:
 - N: número de usuários;
 - Z: tempo de pensamento;
 - M: número de estações na rede;
 - S_i: tempo médio de serviço por visita na estação i;
 - V_i: número de visitas à estação i.



AVM Rede Fechada

- Saídas:
 - X: Vazão do sistema;
 - Q_i: Número médio de usuários na estação i;
 - R_i: Tempo de resposta da estação i;
 - R: Tempo de resposta do Sistema;
 - U_i: Utilização da estação i.



Algoritmo AVM para Rede Fechada

```
Inicialização:
    De i = 1 até M faça Q_i = 0;
Iterações:
    De K = 1 até N Faça
          De i = 1 até M Faça
                      R_i = S_i^*(1+Q_i)
                                                 {CF}
                  ou R_i = S_i
                                                 {SI}
             R = 0:
             De i = 1 até M faça R = R + V_i * R_i
             X = K/(R+Z);
             De i = 1 até M faça
                      Q_i = X^*V_i^*R_i
                      X_i = X^*V_i
                      U_i = X^*S_i^*V_i
```



- Considere a rede fechada do modelo do servidor central. Cada requisição do usuário faz 10 acessos de E/S ao disco A e 5 ao disco B.
- O tempo de serviço por acesso ao disco A e B é respectivamente 300 e 200 milisegundos.



- Cada requisição à UCP leva 2 segundos e o usuário pensa 4 segundos.
- Os parâmetros do sistema são:

$$S_A = 0.3$$
 $V_A = 10 ==> D_A = 3$
 $S_B = 0.2$ $V_B = 5 ==> D_B = 1$
 $D_{UCP} = 2$ $V_{UCP} = V_A + V_B + 1 = 16 ==> S_{UCP} = 0.125$
 $Z = 4$ $N = 20$

Inicialização:

Número de usuários : N = 0;

Número de usuários nas estações:

$$Q_{UCP} = 0, Q_A = 0, Q_B = 0$$



Iteração 1:

Número de usuários: **N** = 1;

Tempo de resposta dos dispositivos:

$$R_{UCP} = S_{UCP} (1+Q_{UCP}) = 0.125(1+0) = 0.125$$

Tempo de resposta dos dispositivos(cont.):

$$\mathbf{R}_{\Delta} = \mathbf{S}_{\Delta} (1 + \mathbf{Q}_{\Delta}) = 0.3(1 + 0) = 0.3 \text{ segundos};$$

$$R_B = S_B (1+Q_B) = 0.2(1+0) = 0.2 \text{ segundos};$$

Tempo de Resposta do Sistema:

$$\mathbf{R} = \mathbf{R}_{UCP} \mathbf{V}_{UCP} + \mathbf{R}_{A} \mathbf{V}_{A} + \mathbf{R}_{B} \mathbf{V}_{B} = 0,125 \times 16 + 0,3 \times 10 + 0,2 \times 5 = 6$$
 segundos;

Vazão do Sistema:

$$X = N/(R+Z) = 1/(6+4) = 0,1$$



Número de usuários nas estações:

$$\mathbf{Q_{UCP}} = \mathbf{XR_{UCP}V_{UCP}} = 0.1x0.125x16 = 0.2$$

$$\mathbf{Q}_{\Delta} = \mathbf{X} \mathbf{R}_{\Delta} \mathbf{V}_{\Delta} = 0,1 \times 0,3 \times 10 = 0,3$$

$$Q_B = XR_BV_B = 0.1x0.2x5 = 0.1$$



Iteração 2:

Número de usuários: **N** = 2;

Tempo de resposta das estações:

$$R_{UCP} = S_{UCP}(1+Q_{UCP}) = 0.125(1+0.2) = 0.15 \text{ segundos};$$

Tempo de resposta das estações (cont.):

$$R_A = S_A(1+Q_A) = 0.3(1+0.3) = 0.39$$
 segundos;

$$\mathbf{R}_{B} = \mathbf{S}_{B}(1+\mathbf{Q}_{B}) = 0.2(1+0.1) = 0.22$$
 segundos;

Tempo de Resposta do Sistema:

$$R = R_{UCP}V_{UCP} + R_AV_A + R_BV_B = 0.15x16 + 0.39x10 + 0.22x5 = 7.4$$
 segundos;

Vazão do Sistema:

$$X = N/(R+Z) = 2/(7,4+4) = 0,175;$$



Número de usuários nas estações:

$$\mathbf{Q_{UCP}} = \mathbf{XR_{UCP}V_{UCP}} = 0.175 \times 0.15 \times 16 = 0.421$$

$$\mathbf{Q}_{\Delta} = \mathbf{X} \mathbf{R}_{\Delta} \mathbf{V}_{\Delta} = 0,175 \times 0,39 \times 10 = 0,684$$

$$\mathbf{Q_B} = \mathbf{XR_BV_B} = 0.175 \times 0.22 \times 5 = 0.193$$

A iterações seguintes produziram os resultados mostrados a seguir:



Iter no.	R _{UCP}	R _A	R_{B}	R	X	Q _{UCP}	Q _A	Q _B
1	0,125	0,3	0,2	6,0	0,1	0,2	0,3	0,1
2	0,15	0,39	0,22	7,4	0,175	0,421	0,684	0,193
		•••	•••	•••	•••	•••	•••	•••
20	0,373	4,854	0,3	56,016	0,333	1,991	16,177	0,5



AVM Resumo

- O método AVM é aplicável somente se a rede possuir solução na forma de produto, ou seja, satisfizer as condições de:
 - fluxo balanceado;
 - eventos únicos;
 - dispositivos homogêneos.



AVM Resumo

- Adicionalmente, deve-se ter também
 - Servidores de capacidade fixa (CF), ou
 - Infinitos servidores (IS), e
 - Em ambos os casos, com tempo de serviço distribuído exponencialmente.



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, o tempo de serviço dos discos A e B é 30 e 25 ms, e o tempo médio de serviço por visita à UCP é 40 mseg.

Utilizando o AVM, determine a vazão e o tempo de resposta do sistema para N=1,...,5 usuários interativos.



Anexo: Análise do Valor Médio Aproximada



Motivação

- O algoritmo de análise do valor médio é iterativo: para calcular o desempenho do sistema com N usuários deve-se conhecer o desempenho com (N – 1) usuários.
- Para simplificar o processo e reduzir o número de passos, utiliza-se uma aproximação: aumentando o número de usuários no sistema, o número de usuários em cada estação aumenta aproximadamente na mesma proporção, i.e. Q_i(N) ≈ α_iN.



AVM Aproximada

- Aproximação inicial: $Q_i(N) \approx \alpha_i N \Rightarrow Q_i(N-1)/(N-1)$ $\approx Q_i(N)/N \Rightarrow Q_i(N-1) \approx [(N-1)/N] Q_i(N).$
- Tempo de resposta na estação i: $R_i(N) \approx S_i \cdot \{1+[(N-1)/N] \ Q_i(N)\}$ (capacidade fixa), ou $R_i(N) = S_i \cdot (\infty \text{ servidores})$.
- Tempo de resposta do sistema: R(N) ≈ Σ_i V_i R_i(N).
- Vazão no sistema: X(N) ≈ N/[Z + R(N)].
- Nova aproximação: Q_i(N) ≈ X(N) V_i R_i(N).



AVM Aproximada

- Algoritmo de Schweitzer: a partir de uma aproximação inicial qualquer para Q_i(N), obter uma nova aproximação até que a diferença seja menor que uma margem de erro ε.
- A escolha da aproximação inicial só afeta o número de passos, não o valor final. Tipicamente escolhe-se Q_i(N) = N/M (isto é, ocupação uniforme das estações).



Algoritmo de Schweitzer

Entradas:

- N: número de usuários;
- M: número de estações na rede;
- Z: atraso na interação com um usuário;
- S_i: tempo de serviço por visita à estação i;
- V_i: número de visitas à estação i;
- ε: margem de erro da aproximação;

Saídas:

- R: tempo de resposta do sistema;
- Ri: tempo de resposta da estação i;
- X: vazão do sistema;
- X_i: vazão na estação i;
- Q_i: número médio de usuários na estação i;
- U_i: fator de utilização da estação i;

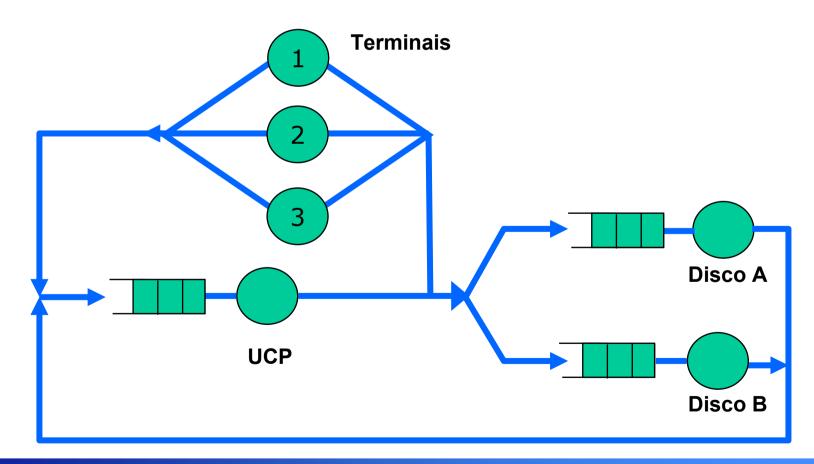


Algoritmo de Schweitzer

```
para i = 1 até M faça Q_i \leftarrow N/M fim
repita
   R \leftarrow 0
   para i = 1 até M faça
      R_i \leftarrow S_i(1+((N-1)/N)Q_i) /* ou S_i */; R \leftarrow R + V_iR_i
   fim
   X \leftarrow N/(R+Z); aprox \leftarrow 0
   para i = 1 até M faça
       aprox \leftarrow max(aprox, |Q_i - XV_iR_i|)
      Q_i \leftarrow XV_iR_i; X_i \leftarrow XV_i; U_i \leftarrow XS_iR_i
   fim
até aprox < ε
```



Sistema de time-sharing (1 UCP e 2 discos):





- N = 20 usuários utilizam o sistema.
- Cada requisição do usuário faz V_A = 10 acessos ao disco A e V_B = 5 ao disco B. O tempo de serviço por acesso é S_A = 0,3 s para o disco A e S_B = 0,2 s para o disco B.
- Cada requisição tem uma demanda de D_{UCP} = 2 s da UCP e o usuário pensa por Z = 4 s.
- Valores derivados: $D_A = S_A V_A = 0.3*10 = 3$; $D_B = S_B V_B = 0.2*5 = 1$; $V_{UCP} = 1 + V_A + V_B = 16$; $S_{UCP} = D_{UCP}/V_{UCP} = 2/16 = 0.125$.



 Usando AVM para redes fechadas, obtém-se após N = 20 iterações:

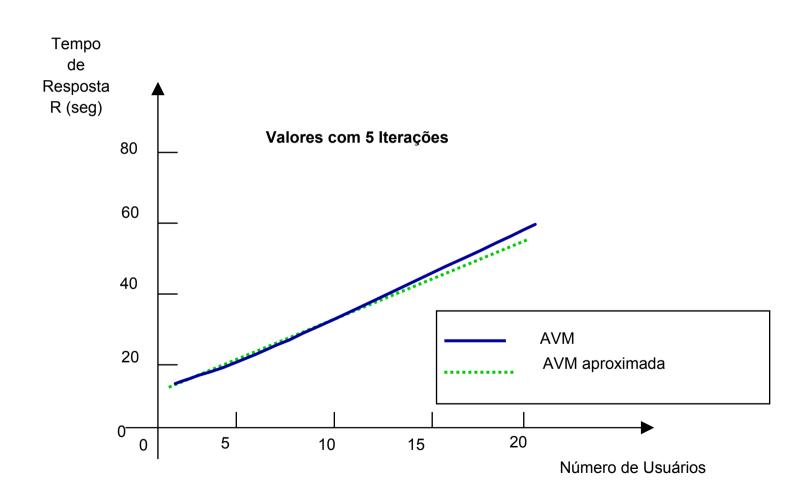
R _{UCP}	R_A	R_{B}	R	X	Q_{UCP}	Q_A	Q_{B}
0,37	4,85	0,30	56,02	0,33	1,99	16,18	0,50

 Usando AVM aproximada, obtém-se com ε = 0,01 (16 iterações):

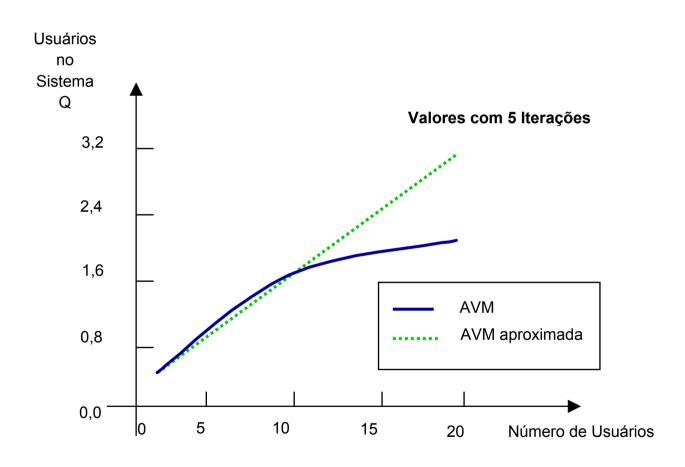
R_{UCP}	R_A	R_{B}	R	X	Q _{UCP}	Q_A	Q_{B}
0,34	4,98	0,29	56,63	0,33	1,78	16,42	0,48

 Mesmo com apenas 5 iterações, os resultados são bastante aceitáveis.











Resposta:

N U			R		Vazão X	Q		
	UCP	Disco A	Disco B	Sistema		UCP	Disco A	Disco B
1	0.040	0.030	0.025	1.700	0.149	0.149	0.090	0.015
2	0.046	0.033	0.025	1.904	0.290	0.333	0.189	0.029
3	0.053	0.036	0.026	2.149	0.420	0.559	0.299	0.043
4	0.062	0.039	0.026	2.443	0.537	0.838	0.419	0.056
5	0.074	0.043	0.026	2.795	0.641	1.179	0.546	0.068



Fim do Módulo Análise do Valor Médio (AVM)