

Análise e Modelagem de Desempenho de Sistemas de Computação

Departamento de Ciência da Computação – UFMG

Lista de Exercícios

Data de Entrega: 03/10

Questão 1

Considere um sistema fechado com 4 discos. Dois discos têm tempo médio de serviço igual a 15 milissegundos enquanto que os outros dois têm um tempo de serviço mais rápido, de 10 milissegundos. A carga deste sistema é composta de requisições de clientes que têm uma demanda média por CPU de 0.9 segundos. Cada requisição gera em média 6 operações de I/O para cada um dos dois discos mais lentos e 9 operações de I/O para cada um dos outros dois discos. Sabendo que o think time médio dos usuários é estimado como 10 segundos, responda:

- O sistema está balanceado? Por quê?
- Qual o ponto ótimo de operação do sistema?
- Qual o impacto no ponto ótimo de operação se a CPU é trocada por uma duas vezes mais rápida?

Questão 2

Em um período de 30 minutos, um disco particular de um servidor está ocupado por 12 minutos. Se é sabido que as requisições que chegam a este servidor realizam, em média, 320 acessos àquele disco, e que o tempo médio de serviço por acesso é de 25 milissegundos, qual o throughput do sistema?

Questão 3

Você é contratado pela Companhia Xulambis para realizar um estudo do desempenho do seu sistema de armazenamento. Você começa o estudo reduzindo o escopo para o cache e os cinco discos *idênticos*, que, segundo o gerente da companhia, são os principais gargalos de desempenho identificados previamente. Cada operação de I/O (requisição para um bloco de dado) submetida ao sistema passa primeiro pelo cache, uma memória mais rápida que armazena os blocos mais recentemente acessados. Caso o bloco requisitado não seja encontrado no cache, o pedido é enviado para um dos 5 discos do sistema. Sabe-se que cada acesso ao cache é servido em 2 milissegundos, independente do bloco ser ou não encontrado. Além disto, sabe-se que os blocos de dados são distribuídos aleatoriamente entre os 5 discos, isto é, a probabilidade de um bloco requisitado estar em um determinado disco é igual para todos os discos. A especificação do hardware estima um tempo médio de serviço no disco de 10 ms. Sabendo que a probabilidade de um bloco ser encontrado no cache é de 0.6 e que a taxa de chegada de requisições para blocos é de 50 requisições por segundo, faça:

- Mostre o modelo de filas correspondente ao sistema analisado, identificando cada centro de serviço modelado
- Calcule o tempo médio de resposta do sistema.
- Calcule a utilização de cada disco.
- Calcule o número médio de requisições na fila de cada disco (não inclui quem está em serviço).
- Calcule o tempo médio de espera de cada requisição nas filas do sistema.

Questão 4

Um amigo seu, iniciante em modelagem de sistemas, lhe apresenta os dados da tabela abaixo, obtidos a partir de medições em um sistema real e também de um modelo MVA de filas separáveis, fechado, que ele construiu. O sistema (real e modelado) tem 4 centros de serviço, uma carga única e 4 clientes. O seu amigo lhe pede ajuda pois, apesar da concordância entre modelo e dados medidos ser excelente para várias métricas de desempenho, ela é muito ruim para algumas outras. Ele lhe pergunta: as discrepâncias são provavelmente causadas por erro nas medições realizadas no sistema ou no código de solução do modelo MVA? Justifique sua resposta apresentando resultados quantitativos que a suportem.

Dados Experimentais	Estimativas a partir do Modelo MVA
$S_1 = 70$	
$S_2 = 12$	
$S_3 = 25$	
$V_1 = 1$	
$V_2 = 2$	
$V_3 = 2$	
$X = 0.01$	$X = 0.01$
$R_1 = 198$	$R_1 = 198$
$U_1 = 70\%$	$U_1 = 70\%$
$R_2 = 18$	$R_2 = 14$
$U_2 = 30\%$	$U_2 = 12\%$
$R_3 = 100$	$R_3 = 87$
$U_3 = 50\%$	$U_3 = 40\%$

Questão 5

Você foi contratado para planejar um upgrade de um servidor Web de abrangência nacional. Como primeiro passo, você executa um estudo do desempenho atual do sistema. Por questões de segurança, o provedor não permite que você realize qualquer medição no sistema real e exige que toda e qualquer coleta seja feita por um de seus especialistas. Você então pede ao especialista que colete medidas sobre a utilização de CPU do servidor, o tempo de resposta, o número médio de requisições no sistema (incluindo as pendentes e em processamento). Além disso, sabendo que os arquivos disponibilizados pelo servidor têm, em média, 10 Kbytes você utiliza a seguinte expressão para estimar a demanda média por CPU, $D_{CPU} = 0.02 + 0.01 \cdot RequestSize$, onde $RequestSize$ é o número médio de blocos de 1 KByte processados. O especialista retornou as seguintes medidas $U_{CPU} = 0.48$, $R = 1s$ e $N = 10$. Você confia nestas medidas? Qual o seu próximo passo? Justifique ambas respostas.

Questão 6

Um servidor Web contém uma CPU e dois discos. Em um período de uma hora, o servidor atendeu 36,000 requisições, houve 360,000 acessos ao disco A e 180,000 acessos ao disco B, e a CPU ficou ocupada por 360 segundos. Os discos A e B são idênticos com um tempo médio de serviço igual a 12 milissegundos. Sabendo que o servidor mantém duas *threads* ativas para processamento das requisições (parâmetro de configuração) e que nenhuma *thread* ficou ociosa durante o período de medição, responda:

- Qual o throughput e o tempo de resposta médio do servidor?

- b) O número de *threads* utilizado é ótimo, se a carga medida for típica? Justifique. Em particular, se não for, apresente o número ótimo.
- c) A configuração de hardware do sistema *suportaria* um tempo de resposta médio inferior a 0.2 segundos?

Questão 7

Um servidor Web com uma CPU e um disco processa requisições para pequenos arquivos HTML e para arquivos de imagem, em média, 5 vezes maiores. Uma medição no sistema durante o período de 1 hora resultou nos seguintes dados:

$$C_{\text{imagem}} = 9000 \quad D_{\text{imagem, CPU}} = 0.1 \text{ segundos} \quad D_{\text{imagem, disk}} = 0.04 \text{ segundos}$$

$$C_{\text{HTML}} = 18000 \quad D_{\text{HTML, CPU}} = 0.01 \text{ segundos}$$

- a) Qual o tempo de residência total médio de cada tipo de requisição na CPU?
- b) O administrador do sistema esqueceu de medir a demanda média por disco das requisições HTML, $D_{\text{HTML, disk}}$ mas ele conhece o tempo médio de residência destas requisições no disco. Ele precisa refazer a medição a fim de obter $D_{\text{HTML, disk}}$? Se sim, diga como isto pode ser feito da forma mais eficiente possível. Se não, apresente uma fórmula que ele possa usar a fim de obter $D_{\text{HTML, disk}}$ a partir dos dados conhecidos.
- c) Se $D_{\text{HTML, disk}}$ é igual a 0.04 segundos, o sistema está balanceado ou há um centro de gargalo que poderia passar por um upgrade para melhorar o desempenho do servidor? Se há, qual é ele?
- d) Você fica sabendo que o mesmo servidor Web também roda, periodicamente e em background, um processo de geração de relatório de acompanhamento do desempenho do servidor. Este processo tem demandas médias por CPU e disco iguais a 1 e 0.05 segundos, respectivamente. Além disto, o início de cada geração de relatório ocorre exatamente 20 segundos após o término da geração anterior. Assumindo que $D_{\text{HTML, disk}}$ é igual a 0.04 segundos, quanto tempo em média dura a geração de cada relatório?
- e) Qual o impacto que a geração de relatórios causa no desempenho do servidor quanto ao tratamento das requisições de imagem e HTML? Quantifique este impacto avaliando o crescimento no tempo de resposta médio para cada tipo de requisição.