

Universidade Federal do ABC
Prova de MCTA025-13 — Sistemas Distribuídos

Turmas NA1MCTA025-13SA e NA2MCTA025-13SA
 Segundo Quadrimestre de 2018

Prof. Dr. Emilio Franceschini

Nome:
RA:

Questão	Pontos	Nota
1	2	
2	2½	
3	2½	
4	3	
5	1	
Total:	11	

- A prova tem duração de **110 minutos**
- Não esqueça de preencher a caneta **todos** os campos na primeira página da folha de prova e seu nome e RA nesta folha de questões.
- Comece a resposta para cada uma das questões em uma nova página.
- Antes da resposta da questão i escreva “Questão i” em letras garrafais.
- As respostas às questões podem ser deixadas a lápis.
- Ao final da prova entregue tanto a folha de questões quanto a folha de prova.
- Em caso de fraude, **TODOS** os envolvidos:
 - **Receberão conceito final F (reprovado) na disciplina**
 - Serão **denunciados** à Comissão de Transgressões Disciplinares Discentes da Graduação e à Comissão de Ética da UFABC cuja punição pode resultar em **advertência, suspensão ou desligamento**, de acordo com os artigos 78-82 do Regimento Geral da UFABC e do artigo 25 do Código de Ética da UFABC.

Boa Prova!

Questão 1. Sobre escalabilidade.

- (a) (1 ponto) Descreva, com exatidão, o que quer dizer *sistema escalável*.

Solução: [ST] Seção 1.2.4

A escalabilidade de um sistema pode ser medida em basicamente 3 dimensões. Escalabilidade em relação ao seu tamanho, que indica que o sistema pode ser facilmente adaptado a uma quantidade maior de usuários pela simples adição de recursos computacionais. Escalabilidade geográfica, que indica a capacidade de um sistema atender clientes que não se localizam perto geograficamente. Escalabilidade administrativa, que se refere a um sistema que é de fácil manutenção ainda que ele abranja diferentes organizações e que se estenda geograficamente.

- (b) ($\frac{1}{2}$ ponto) Como o uso de distribuição como técnica para alcançar escalabilidade pode ajudar a melhorar o desempenho de um sistema distribuído quando seus nós estão espalhados pelo mundo todo?

Solução: O uso da distribuição permite resolver problemas de escalabilidade de tamanho através do offloading de servidores.

- (c) ($\frac{1}{2}$ ponto) Qual novo problema de desempenho pode aparecer com o uso de distribuição como técnica de escalabilidade em tais sistemas?

Solução: Como os servidores estão espalhados geograficamente, e possivelmente por grandes distâncias, pode-se inadvertidamente causar um problema de escalabilidade geográfica.

Questão 2. Um dos mecanismos de comunicação mais usados em sistemas distribuídos é o de *chamada remota de procedimento* (RPC).

- (a) (1 ponto) Explique em detalhes o que é o *stub* cliente.

Solução: Um *stub* cliente fornece a mesma interface que é disponibilizada pelo servidor, contudo é capaz de esconder as possíveis diferenças entre as arquiteturas dos computadores envolvidos assim como os detalhes de comunicação. O *stub* do lado do cliente transforma chamadas locais em mensagens que são enviadas pela rede para o servidor, e o *stub* do servidor transforma as mensagens em chamadas locais e devolve as eventuais respostas pela rede para o *stub* cliente. A impressão do cliente é que ele efetuou uma chamada local ordinária enquanto que para o servidor é como se tivesse sido feita uma chamada local comum a um de seus procedimentos.

- (b) ($\frac{1}{2}$ ponto) Discuta como esse *stub* pode ser mais ou menos “inteligente” e qual a relação disso com os chamados clientes magros e clientes gordos.

Solução: Alguns *stubs* podem ser capazes de resolver algumas requisições na própria máquina do cliente. Por exemplo, é possível que códigos sejam baixados do servidor para serem executados no cliente para evitar uma chamada remota que seria mais custosa em termos de desempenho. Por exemplo,

um formulário poderia verificar se o CPF digitado é válido localmente, e só depois de fazer validações deste tipo enviar efetivamente a chamada para o servidor. Clientes gordos efetuam praticamente todo o processamento localmente, e enviam o resultado para o servidor. Clientes magros não efetuam nada no cliente e delegam todo o processamento para o servidor. Stubs “inteligentes” permitem que alcancemos um meio termo.

- (c) (1 ponto) Que vantagens o uso de chamada remota de procedimentos pode oferecer ao programador, se comparado com o uso direto de uma API de mensagens como a de sockets TCP/IP? Cite uma desvantagem do uso de RPC em relação ao uso de uma API de mensagens.

Solução: O uso de soluções de RPC facilitam o trabalho do programador que não precisa se preocupar com a representação dos dados, com o controle de conexão entre as máquinas, sobre as diferenças arquiteturais entre os computadores,... Por outro lado, podem trazer como desvantagem um controle menos fino da comunicação o que, em alguns casos, pode ser percebido na diferença de desempenho entre uma chamada RPC e uma simples troca de mensagens via TCP/IP.

Questão 3. Queremos comparar o desempenho de dois servidores hipotéticos para a leitura de um arquivo. O primeiro servidor utiliza apenas um thread (*single-threaded*, processamento sequencial) enquanto o segundo pode utilizar quantos threads desejar (*multi-threaded*). Assuma que 10 ms são necessários para fazer a requisição, leitura e processamento dos dados caso eles estejam na memória principal. Caso seja necessário ler os dados do disco (o que é verdade em $\frac{1}{4}$ dos casos), são necessários 65 ms adicionais para completar a requisição.

- (a) ($1\frac{1}{2}$ pontos) Quantas requisições por segundo o servidor com execução sequencial é capaz de atender?

Solução: Considerando-se o tempo de requisição médio como: $\frac{3}{4}10 + \frac{1}{4}75 = 26.25$ temos que o número de requisições por segundo é $\frac{1000}{26.25} = 38.09$.

- (b) (1 ponto) E o servidor *multi-threaded* utilizando 2 threads? (Assuma que só pode haver uma requisição ao disco ativa a cada instante).

Solução: Houve um problema com a elaboração desta questão. O enunciado deveria dizer “Assuma que podem haver duas requisições ao disco ativas a cada instante”. Da maneira como o enunciado está elaborado a resposta passa a ser, na verdade, uma faixa de valores dependendo do escalonamento das requisições (se há sobreposição de requisições ao disco ou não). Nesta questão será considerada como correta tanto a resposta que não leva em conta a sobreposição (dando portanto o upper-bound do intervalo) assim como tentativas de levar em conta a sobreposição (dando o lower-bound).

Caso a questão não tenha sido feita, a letra (a) passa a valer 2,5 pontos.

Questão 4. (3 pontos) Considere um ambiente empresarial cujo negócio dependa de um sistema distribuído que é composto por diversos outros subsistemas (que podem ser distribuídos ou não) fornecidos por diversas empresas de TI. Por exemplo, considere um Pet Shop que oferece como produto não apenas rações, xampus e todo tipo de produtos relacionados a animais de estimação como também oferece um serviço pago mensalmente para que uma pessoa venha buscar e passear todos os dias com o seu *pet*. Neste contexto, "O sistema" da empresa seria, então, um grande sistema distribuído que envolve diversos sistemas de cobrança (cartão de crédito, boleto, débito em conta, ...), sistemas de controle de estoque, sistemas para avisar tanto o dono do animal quanto os prestadores de serviço sobre horários e locais para buscar e deixar os animais, sistemas de controle de entregas, sistemas de controle de assinaturas, interfaces com sistemas de proteção ao crédito, ...

Discorra (detalhadamente) sobre as vantagens e desvantagens que um esquema de comunicação persistente assíncrono pode ter sobre as demais alternativas (combinações de transiente/persistente e síncrono/assíncrono) em um ambiente como o descrito.

Solução: Em um ambiente como o descrito onde há vários serviços e sistemas oferecidos por possivelmente diferentes fornecedores, cada um com diferentes critérios de qualidade e disponibilidade, o uso de alternativas de comunicação persistentes e assíncronas tem a vantagem de desacoplar os diversos sistemas em termos de desempenho (o desempenho de um sistema deixa de ser limitado por aquele mais lento) e também livra os desenvolvedores de um sistema específico de ter que lidar com as indisponibilidades e eventuais alterações na API dos sistemas com os quais se comunicam. O uso de uma solução síncrona e transiente acoplaria cada sistema ao desempenho dos demais e também à sua disponibilidade, sendo que a falha em um único sistema poderia, em um caso extremo, comprometer o funcionamento de todos os demais.

Questão 5. (1 ponto) Comente, critique, reclame, e/ou elogie a disciplina, o material, o projeto e/ou o professor. Caso deseje, sugira mudanças que você gostaria de ver na segunda metade do quadrimestre.