



UNIVERSIDADE ESTADUAL DE MATO GROSSO DO SUL - UEMS
Curso de Ciência da Computação

Disciplina de Algoritmos Paralelos e Distribuídos

Professor: Dr. Rubens Barbosa Filho

Data: 29/11/2023

Ano: 4

Turma: U

Nome: _____ RGM: _____

Exame de PPD. ASSUNTOS: Todo o conteúdo de SD.



Instruções

1. Preencha o cabeçalho corretamente - nome e rgm.
2. Verifique se sua prova está completa - Não faltando folhas.
3. Algumas questões precisam de cálculos. Estes podem ser feitos a lápis. Porém a resposta final deve ser feita a caneta.
4. Não é permitida a consulta a livros, materiais eletrônicos e principalmente ao colega.
5. Leia com atenção as questões antes de responder.
6. Cada questão vem identificada com o seu valor e assunto a que se refere.
7. A interpretação faz parte da avaliação. Se o seu raciocínio estiver correto mas sua resposta estiver errada, uma parte da questão será considerada. Mas se seu raciocínio estiver errado e seu resposta estiver correta, sua questão será desconsiderada.
8. A fraude, a indisciplina e o desrespeito ao professor são passíveis de correção disciplinar.

Questão 1 (Introdução) (1.0)

Explique o que quer dizer **transparência de distribuição** e dê exemplos de diferentes tipos de transparência.

Resposta: A transparência da distribuição é o fenômeno pelo qual os aspectos de distribuição em um sistema ficam ocultos dos usuários e aplicativos. Os exemplos incluem transparência de acesso, transparência de localização, transparência de migração, transparência de relocação, transparência de replicação, transparência de simultaneidade, transparência de falhas e transparência de persistência.

Questão 2 (Arquitetura) (2.0)

Considere uma cadeia de processos P_1, P_2, \dots, P_n implementando uma arquitetura cliente-servidor multidivida. O processo P_i é cliente do processo P_{i+1} e P_i retornará uma resposta a P_{i-1} somente após receber uma resposta de P_{i+1} . Quais são os principais problemas dessa organização quando se examina o desempenho de requisição-resposta no processo P_1 ?

Resposta: Pode-se esperar que o desempenho seja ruim para n grandes. O problema é que cada comunicação entre duas camadas sucessivas ocorre, em princípio, entre duas máquinas diferentes. Consequentemente, o desempenho entre P_1 e P_2 também pode ser determinado por $n-2$ iterações solicitação-resposta entre as outras camadas. Outro problema é que se uma máquina na cadeia tiver um mau desempenho ou estiver temporariamente inacessível, isso irá degradar imediatamente o desempenho ao mais alto nível.

Questão 3 (Processos) (2.0)

Faça uma comparação entre ler um arquivo usando um servidor de arquivos monothread ou um servidor multithread. Lembre-se que para obter uma requisição para o trabalho, despachá-la e fazer o resto do processamento necessário demora 15 ms, considerando que os dados necessários estejam em uma cache de memória principal. Se for preciso uma operação de disco, como acontece em um terço das vezes, serão necessários mais 75 ms, durante os quais a thread dorme. Pergunta: Quantas requisições por segundo o servidor pode manipular se for monothread? E se for multithread?

Resposta: No caso de thread único, os acertos de cache levam 15 ms e as falhas de cache levam 90 ms. A média ponderada é $2/3 \times 15 + 1/3 \times 90$. Assim, a solicitação média leva 40 ms e o servidor pode fazer 25 por segundo. Para um servidor multithread, toda a espera pelo disco é sobreposta, de modo que cada solicitação leva 15 ms e o servidor pode lidar com 66 $2/3$ solicitações por segundo.

Questão 4 (Comunicação) (1.0)

Quando se trata de procurar arquivos em um sistema peer-to-peer não estruturado, pode ser útil restringir a busca a nós que tenham arquivos semelhantes aos seus. Explique como o Protocolo *Gossiping* pode ajudar a achar esses nós.

Resposta: Se, durante a execução do *Gossiping*, os nós trocarem informações de associação, cada nó eventualmente conhecerá todos os outros nós do sistema. Cada vez que descobre um novo nó, ele pode ser avaliado quanto à sua proximidade semântica, por exemplo, contando o número de arquivos incomuns. Os nós semanticamente mais próximos são então selecionados para enviar uma consulta de pesquisa.

Questão 6 (Nomeação) (1.0)

Qual é a maior desvantagem de consultas recursivas na resolução de uma chave em um sistema baseado em DHT?

Resposta: Um problema é que o cliente solicitante nunca será capaz de descobrir o que deu errado quando nenhuma resposta for retornada. Em algum lugar ao longo da rota que corresponde à resolução da chave, uma mensagem pode ter sido perdida ou um nó pode ter falhado. Por esse motivo, às vezes é preferível uma pesquisa iterativa: o cliente saberá exatamente qual parte da pesquisa não foi realizada e poderá escolher um nó alternativo para ajudar.

Questão 7 (Sincronização)(1.0)

Considere o comportamento de duas máquinas em um sistema distribuído. Ambas têm relógios que devem pulsar 1000 vezes por milissegundo. Um deles realmente pulsa a essa taxa, mas o outro pulsa somente 990 vezes por milissegundo. Se as atualizações UTC chegam uma vez por minuto, qual será a máxima defasagem entre os relógios?

Resposta: O segundo relógio bate 990.000 vezes por segundo, dando um erro de 10 ms por segundo. Em um minuto, esse erro cresceu para 600 ms. Outra maneira de resolver é: o segundo relógio está um por cento lento, então, depois de um minuto, ele está atrasado em $0,01 \times 60$ segundos, ou 600 ms.

Questão 8 (Consistência e Replicação) (2.0)

Vimos em sala que o protocolo Multicast totalmente ordenado utilizando relógios lógicos de Lamport não é escalável. Explique por quê.

Resposta: O multicasting totalmente ordenado de Lamport exige que todos os servidores estejam funcionando, prejudicando efetivamente o desempenho quando um deles fica lento ou travou. Isso terá que ser detectado por todos os outros servidores. À medida que o número de servidores cresce, esse problema se agrava.

BOA PROVA!