

Engenharia Informática

Sistemas Distribuídos em Grande Escala

Teste¹

24 de maio de 2022

Duração: 1h30

I

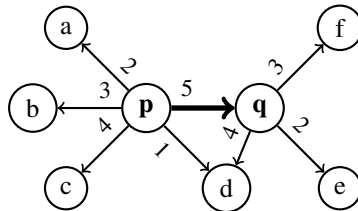
1 Descreva brevemente as garantias de sessão *monotonic reads*, *monotonic writes* e *read your writes*. Explique porque as duas primeiras não implicam (em geral) a terceira, e descreva o que pode ser feito, em termos da interação entre clientes e servidores, para obter a terceira, tendo as duas primeiras.

2 Explique porque relógios escalares (de Lamport) apenas são compatíveis com a causalidade, mas não a caracterizam, e porque não é suficiente a associação de um simples relógio escalar a mensagens num algoritmo de *causal broadcast*.

3 Compare as abordagens de implementação de CRDTs baseadas em estado com as baseadas em operações em termos da sua adequação aos seguintes objectivos ou cenários: “frescura” dos resultados (baixa latência de visibilidade), garantias exigidas em termos de envio de mensagens, tolerância a partições temporárias de rede, saída (terminação) abrupta de participantes (réplicas).

II

1 Considere um fragmento centrado nos nós p e q de uma rede sobreposta construída e mantida pelo algoritmo CYCLON com os parâmetros $c = 5$ (tamanho da *cache*) e $l = 2$ (tamanho da amostra). A idade de cada nó na cache é indicado sobre o arco. O nó p vai iniciar uma ronda de troca com q .



1.1 Qual foi o último nó que concluiu uma troca (*shuffle*) com p ? Justifique.

1.2 Descreva o passo de troca (*shuffle*) iniciada por p com q e indique o seu resultado final. Quando precisar de escolher nós aleatoriamente, admita que são escolhidos os que têm o nome alfabeticamente mais baixo.

2 Explique a utilidade de *consistent hashing* para sistemas distribuídos em grande escala.

3 Diga o que entende por transações de leitura num sistema geo-replicado com coerência *causal+*. Justifique a necessidade de guardar múltiplas versões de cada item para a conseguir.