MESTRADO EM INFORMÁTICA

Engenharia de Aplicações

Frequência¹

11 de Fevereiro de 2014 Duração: 2h00min

- 1 Descreva na generalidade como é feita a previsão do número de itens produzidos por cada operador num plano de execução de uma interrogação relacional.
- 2 Diga qual o objetivo de um mecanismo de recuperação num servidor de bases de dados. Identifique quais os principais fatores que influenciam a sua escolha.
- 3 Compare sistemas transacionais oferecendo serializability e snapshot isolation. Utilize exemplos concretos nesta comparação.
- 4 Diga, justificando, a que se destina o sharding de bases de dados, como funciona, e quais são as suas limitações principais. Em que medida uma base de dados NoSQL como o HBase resolve, ou não, estes problemas?

II

Este grupo é opcional. Como alternativa será atribuída a nota do trabalho prático de ABD.

Considere um sistema pré-pago de inserção de anúncios em páginas Web. Neste sistema são registados anúncios que devem ser mostrados aleatoriamente um determinado número de vezes (coluna alvo). Logo que esse alvo tenha sido atingido, não é necessário inserir esse anúncio de novo. Cada página contém algumas dezenas de anúncios. Este sistema funciona num servidor PostgreSQL da seguinte forma: Durante a geração de uma página, é selecionado um conjunto de anúncios a inserir, entre os que têm saldo disponível, sendo atualizados os respetivos saldos. Este sistema está implementado da seguinte forma. Em primeiro lugar, para cada página são selecionados anúncios cujo saldo restante relativo é maior, usando a seguinte interrogação:

select id from anuncios order by (alvo-mostrado)/alvo limit 100;

Destes 100 são escolhidos aleatoriamente os artigos a mostrar efetivamente, sendo então, no contexto da mesma transação, atualizadas as contagens na coluna mostrado.

- 1 Admita que o servidor não tem capacidade para responder ao número de pedidos desejado. Como procederia para diagnosticar a origem do problema?
- 2 Avançando com as causas mais prováveis das limitações do sistema, como as resolveria sem mudar o servidor que o suporta?

I

¹Cotação — (10) + (10)

- 1 A equipa de infraestruturas dos jogos olímpicos de Inverno, de forma a garantir a disponibilidade dos serviços que tem de prestar montou dois centros de dados independentes e geograficamente isolados, o que lhe permite assegurar a continuidade de serviço em caso de falha. Durante o decorrer do evento verificaram que o datacenter principal não consegue servir todos os ficheiros de imagens que os utilizadores querem consultar e verificaram que têm capacidade disponível no datacenter de backup. Que solução propunha para poder fazer a distribuição da carga pelos dois centros de dados?
- 2 No trabalho prático que realizou, a monitorização foi relegada para segundo plano. No entanto ela é fundamental para garantir a correta operacionalidade da infraestrutura. Proponha um sistema de monitorização para essa infraestrutura. Como é que poderia evoluir essa infraestrutura para também permitir fazer gestão de capacidade, isto é, permitir-lhe tomar decisões sobre quando e como deve escalar/reduzir os recursos da infraestrutura.
- **3** Uma cadeia de lojas pediu-lhe para desenvolver uma arquitetura de servidores e storage para a sua loja online. Foi-lhe dado como informação que o número de artigos será na ordem das centenas de milhar ou milhões e que para cada artigo existem várias imagens (pode chegar às dezenas) com diferentes resoluções. Detalhe a solução que proporia.
- 4 A loja online foi um sucesso tão grande que voltaram a contactá-lo com o seguinte problema. A manter-se o ritmo atual dentro de pouco tempo o servidor de storage deixará de ter capacidade de responder a todos os pedidos que lhe são efetuados, atinge o número de operações de I/O. Propuseram-lhe o desafio de conseguir responder a este previsível aumento dos pedidos sem alterar a storage existente. Que soluções apresentaria para prolongar a vida deste equipamento, sabendo que os restantes servidores da infraestrutura estão relativamente folgados.