

Atividade Complementar 2 com solução

Curso: Especialização em Arquitetura de Software Distribuído

Disciplina: Análise, Projeto e Avaliação de Arquitetura de Software

Docente: Maria Augusta Vieira Nelson

Contexto: para a realização das atividade 2 é fornecida a descrição de um problema de um cliente do domínio de entretenimento que descreve o Sistema de Aluguel e Streaming de Filmes (SASF).

Fonte: A descrição do Sistema de Aluguel e Streaming de Filmes (SASF) foi adaptada de:

BARBOSA, Guilherme Mauro Germóglio. Um livro-texto para o ensino de projeto de arquitetura de software, Campina Grande, 2009. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de Campina Grande, Centro de Engenharia Elétrica e Informática.

Atividade

Objetivo da atividade: exercitar o conceito de estimativa de uso de memória secundária para um servidor de dados e o conceito de planejamento da capacidade da arquitetura de sistemas. A capacidade de um sistema em geral é um requisito de qualidade que precisamos atender em diversos sistemas. Esta atividade trabalha com o conceito de capacidade de armazenamento de memória secundária.

Descrição do problema: Sistema de Aluguel e Streaming de Filmes (SASF).

O SASF é um sistema de informação com dois grandes objetivos: (1) gerenciar o processo de locação via web de vídeos e (2) proporcionar infraestrutura de software para realizar streaming de vídeos também via web.

O SASF é dividido em cinco grandes módulos funcionais. Cada módulo é responsável por prover um conjunto de funcionalidades relacionadas entre si. Os grandes módulos do SASF são:

- a) Locadora de Filmes;
- b) Transmissor de Filmes;
- c) Motor de Sugestões;
- d) Cadastro de Usuários;
- e) Cadastro de Filmes.

O armazenamento das informações dos módulos será realizado usando um Sistema Gerenciador de Banco de Dados Relacional (SGBDR) de modo a permitir criação, edição, obtenção e remoção das entradas. As informações guardadas no SGBDR são os atributos dos Usuários, Filmes, Locações, Produtoras, e são essencialmente textuais ou metainformações para localização de arquivos de mídia (fotos ou vídeos). Além disso, o armazenamento de arquivos textuais que não são atributos dos Usuários, Filmes ou Locações, por exemplo, mensagens para usuários, recomendações de filmes ou comentários sobre filmes também devem ser armazenados no SGBDR. O armazenamento de arquivos de mídia (fotos de Usuários, fotos de Filmes e arquivos de vídeo) serão armazenados usando uma Rede de Fornecimento de Conteúdo (Content Delivery Network ou CDN) e não serão tratados aqui. Não serão consideradas aqui também as informações sobre pagamento das mensalidades de usuários.

Volume de dados inicial

10.000.000 usuários cadastrados ao todo (estima-se chegar a 100.000.000) (20% usam o sistema por dia)

100.000 filmes cadastrados (5.000.000 de DVDs disponíveis)

175 produtoras de filmes

Frequência das atividades:

10.000 novos usuários são cadastrados por dia

20.000 pedidos de locações recebidos por dia

São cerca de 150 mil vídeos transmitidos por dia, ou seja, um stream de vídeo sendo iniciado a cada 0.57 segundos caso as transmissões fossem distribuídas uniformemente ao longo do dia. Se considerarmos que a transmissão de um vídeo dura em média uma hora, isso gera uma carga de pouco mais de seis mil usuários simultâneos fazendo *streaming*.

Usuários trocam em média 500 mensagens por dia e em média cada arquivo de mensagem ocupa 400 bytes

Usuários escrevem 10.000 comentários por mês e em média cada comentário ocupa 200 bytes

Recomendações de filmes de um usuário para outro acontecem em média 500 por dia e em média cada arquivo de recomendação possui 100 bytes.

Em média 1500 novos filmes passam a fazer parte do acervo de filmes do sistema a cada ano.

Critérios de esvaziamento:

Locações realizadas há mais de um mês passam para um outro banco secundário fora do servidor principal de dados

Usuários sem solicitação de locações ou transmissão de vídeo há mais de um ano passam para um outro banco secundário (em média são 10.000 por ano)

Mensagens entre usuários enviadas há mais de um ano são removidas do sistema.

Dicionário de dados:

ARQUIVO-EXTERNO = COD-ARQ + CAMINHO-ARQUIVO

CAMINHO-ARQUIVO = * caminho com a localização do arquivo no sistema operacional; tipo: alfanumérico; 100 bytes *

COD-ARQ = * código de identificação do arquivo; tipo: numérico; 4 bytes *

COD-DIRETOR = * código de identificação do diretor; tipo: numérico; 4 bytes *

COD-FILME = * código de identificação do filme; tipo: numérico; 4 bytes *

COD-PRODUTORA = * código de identificação da produtora do filme; tipo: numérico; 4 bytes *

COD-USUÁRIO = * código de identificação do usuário; tipo: numérico; 4 bytes *

COMENTÁRIO = COD-USUÁRIO + COD-FILME + COD-ARQ

DATA = * tipo: numérico; 4 bytes *

ENDEREÇO = * tipo: alfanumérico; 150 bytes *

FILME = COD-FILME + TÍTULO + COD-PRODUTORA + COD-DIRETOR + TRANSMISSÃO

LOCAÇÃO = COD-USUÁRIO + COD-FILME + DATA

LOGIN = * palavra que descreve o login do usuário no sistema; tipo: alfanumérico; 20 bytes *

MENSAGEM = COD-USUÁRIO *origem* + COD-USUÁRIO *destino* + COD-ARQ

NOME = * tipo: alfanumérico; 60 bytes *

RECOMENDAÇÃO = COD-USUÁRIO *origem* + COD-USUÁRIO *destino* + COD-FILME + COD-ARQ

SENHA = * senha do usuário no sistema; tipo: alfanumérico; 20 bytes *

TELEFONE = * telefone do usuário do sistema; tipo: alfanumérico; 15 bytes *

TÍTULO = * tipo: alfanumérico; 100 bytes *

TRANSMISSÃO = *endereço URL que armazena a transmissão via streaming do filme; tipo: alfanumérico; 50 bytes *

USUÁRIO = COD-USUÁRIO + NOME + ENDEREÇO + TELEFONE + LOGIN + SENHA

Estimativa de uso de memória secundária em um servidor:

Nome do servidor onde os dados serão armazenados

Tempo estimado para armazenagem dos dados neste servidor (TE)

Nome do arquivo/tabela do banco de dados

Tamanho do registro (TR)

Número inicial de registros (NIR)

Estimativa de crescimento (EC) do número de registros no Tempo de Estimado

Tamanho do arquivo/tabela $TA = TR \times (NIR + EC)$

Considere para a estimativa deste servidor que o tempo estimado é 1 ano. Calcule o tamanho de cada tabela separadamente e em seguida faça o somatório das tabelas.

Nome do Arquivo/Tabela	TR	NIR	EC	TA
Usuários				
Filmes				
Locações				
....				
Total (em Gigabytes)				

Solução

Para calcularmos o TA de USUÁRIO

Considere o registro de usuário descrito no dicionário de dados do SASF:

USUÁRIO = COD-USUÁRIO + NOME + ENDEREÇO + TELEFONE + LOGIN + SENHA

O campo COD-USUÁRIO = * código de identificação do usuário; tipo: numérico; 4 bytes *

O campo NOME ...

O campo ENDEREÇO ...

O campo TELEFONE ...

O campo LOGIN...

O campo SENHA...

TR usuários: 4 bytes + 60 bytes + 150 bytes + 15 bytes + 20 bytes + 20 bytes = 269 bytes

Para o cálculo do NIR voltamos ao texto e percebemos que na descrição do Volume inicial de dados, existem 10.000.000 de usuários cadastrados no sistema.

NIR usuários: 10.000.000

Para calcular o EC precisamos considerar tudo que vai entrar e sair na tabela de usuários do banco de dados dentro do período estimado de um ano. No texto do enunciado do problema percebemos que na descrição da frequência das atividades é dito que 10.000 novos usuários se cadastram no sistema diariamente. E nos critérios de esvaziamento está descrito que 10.000 usuários em média por ano deixam de usar o sistema.

EC usuários: $10.000 \times 365 - 10.000 = 3.640.000$

$TA = TR \times (NIR + EC)$

TA usuários = $269 \text{ bytes} \times (10.000.000 + 3.640.000) = 3.669.160.000 \text{ bytes}$

TA usuários = 3,4 Gigabytes

Completando a tabela temos:

Nome do Arquivo/Tabela	TR	NIR	EC	TA
Usuários	269	10.000.000	3.640.000	3,4Gb
Filmes				
Locações				
....				
Total (em Gigabytes)				

Agora faça o cálculo para as outras tabelas.