



CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
CURSO DE ENGENHARIA DE COMPUTAÇÃO

PERSISTÊNCIA POLIGLOTA

JOSÉ FRANCISCO CAMPOS LIMONGI

Orientador: Evandrino Barros
Centro Federal de Educação Tecnológica de Minas Gerais – CEFET-MG

BELO HORIZONTE
JULHO DE 2014

JOSÉ FRANCISCO CAMPOS LIMONGI

PERSISTÊNCIA POLIGLOTA

Modelo canônico de trabalho monográfico acadêmico
em conformidade com as normas ABNT apresentado à
comunidade de usuários L^AT_EX.

Orientador: Evandrino Barros
 Centro Federal de Educação Tecnológica
 de Minas Gerais – CEFET-MG

CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
CURSO DE ENGENHARIA DE COMPUTAÇÃO
BELO HORIZONTE
JULHO DE 2014

JOSÉ FRANCISCO CAMPOS LIMONGI

PERSISTÊNCIA POLIGLOTA

Modelo canônico de trabalho monográfico acadêmico
em conformidade com as normas ABNT apresentado à
comunidade de usuários \LaTeX .

Trabalho aprovado. Belo Horizonte, 24 de novembro de 2014

Evandrino Barros
Orientador

Professor
Convidado 1

Professor
Convidado 2

CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
CURSO DE ENGENHARIA DE COMPUTAÇÃO
BELO HORIZONTE
JULHO DE 2014

Espaço reservado para dedicatória. Inserir
seu texto aqui...

Agradecimentos

Inserir seu texto aqui... (esta página é opcional)

“Seja realista, exija o impossível.” (Roland Castro)

Lista de Figuras

Lista de Tabelas

Lista de Quadros

Lista de Algoritmos

SQL	Structure Query Language	1
Redis	Remote Dictionary Server	2

[siglas]

Sumário

1 – Introdução	1
1.1 Motivação	2
 Apêndices	 4
APÊNDICE A–Nome do Apêndice	5
APÊNDICE B–Nome do Apêndice	6
 Anexos	 7
ANEXO A–Nome do Anexo	8
ANEXO B–Nome do Anexo	9

1 Introdução

A necessidade de persistência de dados sempre esteve presente na computação. A medida que os sistemas evoluíram a complexidade da forma que os dados eram armazenados aumentou significativamente. Com isso, houve a necessidade da criação de um sistema computadorizado de manutenção de registros (??), o banco de dados.

O modelo relacional foi um dos primeiros gêneros de banco de dados, sua estrutura são tabelas de duas dimensões com linhas e colunas. Os dados armazenados são tipados podendo variar a quantidade de tipos de acordo com o banco utilizado. Para interagir com esse gênero é necessário realizar consultas com a linguagem Structure Query Language (SQL). Alguns exemplos de banco de dados relacional são MySQL ¹, Oracle ² e PostgreSQL ³.

Durante anos, o banco de dados relacional tem sido considerado a melhor opção para os problemas de escalabilidade, porém surgiram novas soluções com novas alternativas de estruturas, replicação simples, alta disponibilidade e novos métodos de consultas (??). Essas opções são conhecidas como NoSQL ou banco de dados não-relacional

Existem diversos gêneros de banco de dados não-relacional, entre eles chave-valor, orientado a documento, orientado a coluna e orientado a nó. Com o surgimento dessas novas soluções, o questionamento sobre qual banco de dados é melhor para resolver certo tipo de problema, vem à tona. A partir disso, o conhecimento e compreensão sobre os bancos de dados em geral se torna necessário para realizar uma boa escolha.

Entendendo que cada banco se destaca em determinados tipos de problema, é nítido perceber que sistemas que trabalham com mais de um banco de dados podem oferecer um melhor desempenho, dando origem à persistência poliglota.

Este trabalho consiste na comparação de dois sistemas, o primeiro que utilizará apenas um banco de dados, que será do gênero orientado a documento, e o segundo que utilizará dois bancos de dados, sendo um do gênero orientado a documento e um outro banco do gênero chave-valor. O segundo sistema por utilizar dois bancos de dados caracteriza a persistência poliglota, que é alvo desse trabalho. Os bancos escolhidos para fazer esses sistemas foram o MongoDB e o Remote Dictionary Server (Redis). O autor escolheu esses bancos de dados por ter experiência na linguagem *Ruby on Rails* que oferece excelentes bibliotecas para esses bancos. O intuito desse trabalho é comprovar

¹ Sítio oficial <<http://www.mysql.com>>

² Sítio oficial <<http://www.oracle.com/technetwork/oem/db-mgmt/db-mgmt-093445.html>>

³ Sítio oficial <<http://www.postgresql.org/>>

que o uso da persistência poliglota melhora o desempenho da aplicação.

MongoDB é um banco de dados do gênero orientado a documento e foi desenhado para ser gigante. O próprio nome é uma derivação da palavra inglesa *humongous* que significa gigantesco. O diferencial desse gênero é a maneira que os registros são armazenados. Cada registro fica armazenado em um documento que é análogo à tupla no modelo relacional. O documento é composto por um identificador único e um conjunto de valores de tipos e estruturas aninhadas. Esse gênero é bem flexível, pois não tem catálogo, ou seja, não existe tabela e permite valores multivalorados. Ao criar a arquitetura do sistema temos que identificar se as entidades criadas são expressivas como um documento (??). Além disso, tem soluções para tratar concorrência e foi desenhado para trabalhar em *clusters*. A linguagem utilizada para fazer consulta no MongoDB é JavaScript. Nesse trabalho iremos utilizar o banco MongoDB nos dois sistemas que serão criados. Esse banco está sendo utilizado em grandes empresas, como Cisco, eBay, Codecademy, Microsoft, Craigslist, The Guardian e outras, conforme sítio oficial do MongoDB ⁴.

O segundo banco que iremos utilizar se chama [Redis](#) do gênero chave-valor. Esse tipo de armazenamento é mais simples, como próprio nome indica, é armazenado um valor para determinada chave. O valor armazenado pode ter uma estrutura variável. Essa escolha foi feita devido ao cache que esse sistema realiza antes de efetivar a operação no disco. Esse cache tem um ganho muito alto em desempenho, porém poderá ocorrer perda de dados, caso ocorra uma falha de hardware (??). A forma de como estruturar esse banco é muito parecida com um tipo estruturado chamado *hash* que são implementadas em algumas linguagens de computação, como Java e Ruby. Esse banco será utilizado no segundo sistema a ser desenvolvido. O [Redis](#) está sendo utilizado no Twitter, Github, Craigslist e outros conforme referência do sítio oficial ⁵.

1.1 Motivação

A persistência poliglota é uma alternativa para melhorar o desempenho de uma aplicação. Utilizando-a conseguimos adaptar cada tipo de problema com um gênero de banco de dados.

Existem duas variáveis opostas no ambiente de persistência de dados, consistência e disponibilidade. Quanto mais consistente um dado, menos disponível ele será e quanto mais disponível um dado menos consistente ele estará. Em aplicações é comum termos um conjunto de dados que devem ser sempre consistentes e um outro conjunto de dados que devem estar sempre disponíveis. Logo, para atender a esses conjuntos de

⁴ <<http://www.mongodb.com/customers>>

⁵ <<http://redis.io/topics/whos-using-redis>>

dados devem ser utilizados dois banco de dados, um que garante disponibilidade e outro que garante consistência. Com esse ambiente misto é esperado que haja um ganho de desempenho.

Atualmente poucos trabalhos apresentam uma comparação entre sistemas que utilizam apenas um banco, sistema monoglota, e sistemas que utilizam persistência poliglota.

Apêndices

APÊNDICE A – Nome do Apêndice

Inserir seu texto aqui...

APÊNDICE B – Nome do Apêndice

Inserir seu texto aqui...

Anexos

ANEXO A – Nome do Anexo

Inserir seu texto aqui...

ANEXO B – Nome do Anexo

Inserir seu texto aqui...