

Relatório do Projeto de Big Data

Introdução

O projeto visa uma aplicação robusta para análise de dados musicais utilizando ferramentas de BigData e NoSQL, especialmente MongoDB para armazenamento dos dados, Redis para cache de consultas e Jupyter Notebook para análise e visualização dos dados. Tudo isso através do uso do Docker, garantindo o isolamento e facilidade de configuração.

MongoDB

MongoDB é utilizado como banco de dados principal do projeto, responsável pelo armazenamento do dataset musical. Ele permite consultas flexíveis e armazenamento de dados em forma de documentos, ideal para a manipulação de grandes volumes de dados não estruturados.

Redis

O Redis atua como uma camada de cache para o MongoDB, armazenando resultados de consultas frequentemente acessadas. Isso melhora significativamente a performance da aplicação, reduzindo o tempo de resposta para consultas repetidas.

Jupyter Notebook

O Jupyter é utilizado para análise e visualização de dados. Ele permite a execução de Python em ambiente interativo, facilitando a exploração de dados, visualizações e execução de análises complexas em tempo real.

Docker

O Docker é utilizado para containerização e gerenciamento de todos os serviços mencionados acima, garantindo que funcionem de maneira isolada e controlada. Docker Compose é utilizado para definir como esses containers interagem, suas dependências, e como são configurados em um ambiente unificado.

Integração dos serviços

A integração entre MongoDB, Redis e Jupyter é crucial para o sucesso da aplicação. O MongoDB fornece os dados, o Redis cacheia os resultados das consultas para aumentar a eficiência, e o Jupyter realiza as análises e visualizações dos dados. A rede 'mybridge' permite que esses serviços se comuniquem de maneira segura e eficaz dentro do ambiente Docker.

Segurança e acesso

O sistema usa senhas e configurações de rede para garantir que apenas serviços autorizados possam acessar o MongoDB e Redis. O Jupyter é acessado através de um token gerado automaticamente ou uma senha definida pelo usuário, proporcionando uma camada adicional de segurança.

Guia de Uso da aplicação

Após fazer a instalação do repositório certifique-se de estar dentro da pasta do mesmo

Utilize os comandos:

“docker-compose build”

“docker-compose up -d”

dentro de cada repositório para sua inicialização,

o acesso do Jupyter se encontra disponível no navegador
“localhost:8888”, sua senha foi definida como “pass”

o acesso do Mongodb se encontra na porta 8081, através do Express,
onde login e senha seguem como padrão “admin” “pass”

no Jupyter deve-se verificar o número e ip que foi gerado para o docker
“mongo service” para que possa ser feita a conexão com o Mongodb.