

KIẾN TRÚC PHẦN MỀM

ĐỒ ÁN CUỐI KỲ



Bộ môn Công nghệ phần mềm
Khoa Công nghệ thông tin
Đại học Khoa học tự nhiên TP HCM

MỤC LỤC

1	Thông tin nhóm	3
2	Thiết kế kiến trúc	4
2.1	Sơ đồ kiến trúc	4
2.2	Trình bày kiến trúc	4
2.2.1	Mp3 Http Webserver	5
2.2.2	Mp3 Monitor Websocket	5
2.2.3	Caching System	5
2.2.4	Mp3 Data Middleware	6
2.2.5	Streaming System	6
2.2.6	Search Middleware System	7
2.2.7	Identification Middleware System	7
2.2.8	Crawler Worker	7
2.2.9	MongoDB	7
2.2.10	KyotoCabinet	7
3	Demo	8

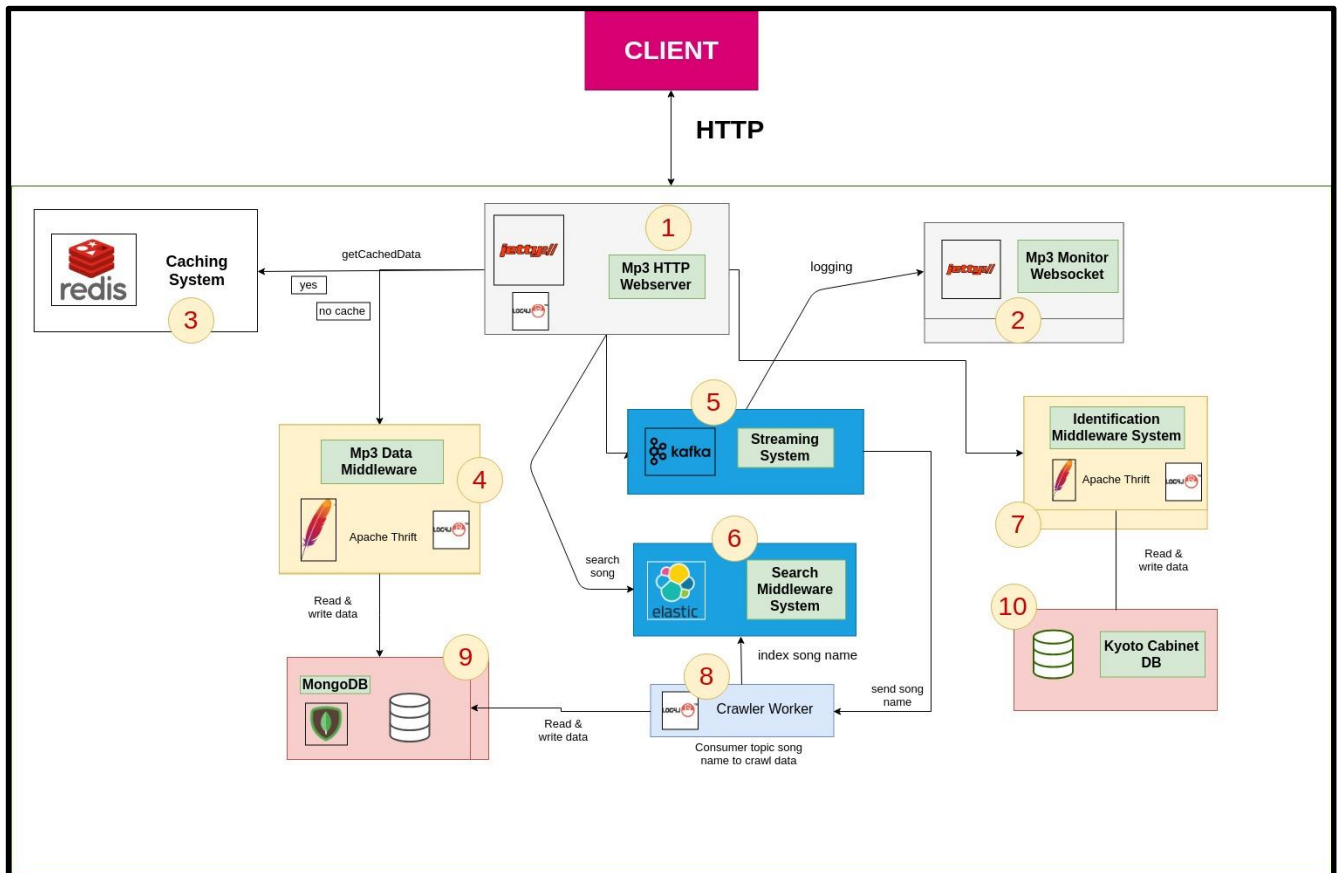
1

Thông tin nhóm

MSSV	Họ Tên	Email	Điện thoại
1512042	Nguyễn Thành Chung	1512042@student.hcmus.edu.vn	0329009486
1512063	Nguyễn Du Du	1512063@student.hcmus.edu.vn	0996027432

2 Thiết kế kiến trúc

2.1 Sơ đồ kiến trúc



2.2 Trình bày kiến trúc

Ứng dụng "Nghe nhạc trực tuyến mp3" cung cấp các tính năng cơ bản của một trang web nghe nhạc như: Nghe nhạc online, xem lyric bài hát, xem thông tin album, thể loại của bài hát; xem thông tin của ca sỹ; tìm kiếm tên bài hát. Trang website hỗ trợ chức năng đăng nhập cho admin, hỗ trợ xem monitor hệ thống. Khi người dùng tìm kiếm tên bài hát, nếu bài hát chưa tồn tại trong hệ thống thì hệ thống sẽ tự động đi crawl data từ trang mp3.zing.vn.

Hệ thống được xây dựng theo kiến trúc microservice, hướng chức năng. Mỗi một system trong hệ thống sẽ có một chức năng riêng biệt đảm bảo scale hệ thống nhanh, phát triển độc lập, một hệ thống bị lỗi không gây ảnh hưởng tới các hệ thống còn lại. Các phần trên hình vẽ sẽ được trình bày dưới đây.

2.2.1 Mp3 Http Webserver

Server Http phục vụ các chức năng của hệ thống đối với người dùng từ web browser qua giao thức HTTP. Server được cài đặt bằng mã nguồn mở Jetty. Jetty là một framework giúp xây dựng các trang web standalone hoạt động riêng biệt mà không phải deploy server vào một server container như Tomcat.

Server Http hỗ trợ các handler sau:

LoginHandler: Chức năng đăng nhập vào hệ thống.

LogoutHandler: Chức năng đăng xuất khỏi hệ thống.

HomeHandler: Xem trang chủ của trang web.

AdminHandler: Chức năng monitor, quản lý cho admin của hệ thống.

SongHandler: Chức năng nghe nhạc, xem thông tin kèm theo bài hát như ca sỹ, album, thể loại, lời bài hát.

SingerHandler: Chức năng xem thông tin chi tiết cho ca sỹ.

LyricHandler: Chức năng xem lyric.

SearchHandler: Chức năng search tên bài hát.

2.2.2 Mp3 Monitor Websocket

Websocket hỗ trợ realtime xem monitor cho admin của hệ thống. Websocket được xây dựng bằng open source Jetty như ở Mp3 Http Server.

Việc monitor trên hệ thống rất quan trọng để lập trình viên và quản trị hệ thống có thể nâng performan, khắc phục sự cố cho service cần thiết.

2.2.3 Caching System

Caching là một việc rất quan trọng mà bất kỳ một hệ thống nào cũng sử dụng để tăng performance cho service. Truy xuất nhanh tức thời, hạn chế tiêu tốn tài nguyên của server.

Trong hệ thống hiện tại sử dụng Redis như một service riêng biệt phục vụ việc caching cho hệ thống. Redis là một mã nguồn mở hỗ trợ nosqlDB dạng key-value giúp truy xuất

data nhanh phù hợp cho việc cache và lưu session. Redis hỗ trợ lưu các loại kiểu dữ liệu như String, List, Set, SortedSet, Hash.

Hiện tại hệ thống đang config Redis sử dụng chiến lược LRU để tối ưu việc lưu cache cần thiết cho hệ thống, không gây tiêu tốn tài nguyên quá nhiều. Redis hỗ trợ việc dump và backup data theo định kỳ xuống ổ cứng để khắc phục tình trạng khi server xảy ra vấn đề trực trực như mất điện, server chết thì có thể khôi phục lại dữ liệu đã cache. Mọi request từ Mp3HttpServer đến các Middleware đều được cache lại. Trước khi gọi request tới Middleware luôn được kiểm tra trong cache trước để truy xuất dữ liệu được nhanh chóng.

2.2.4 Mp3 Data Middleware

Hệ thống Mp3 Data Middleware có trách nhiệm trung gian trong việc lưu trữ và lấy dữ liệu từ Mp3 Http Server. Server Middleware được cài đặt bằng open source "Thrift". Thrift là một ngôn ngữ trung lập giúp định nghĩa ra các model và service hỗ trợ nhiều loại ngôn ngữ. Đây là điều đặc biệt khi mà server ta có thể viết bằng một ngôn ngữ khác, và client bằng một ngôn ngữ khác đều được thrift server. Thrift hỗ trợ nhiều loại server, protocol, transport có khả năng truyền tải dữ liệu tốc độ cao.

Các middleware giao tiếp với Database MongoDB và KyotoCabinet để lưu trữ dữ liệu.

2.2.5 Streaming System

Hệ thống có trách nhiệm nhận các message job từ các đầu server và bắn message vào các worker để thực hiện một công việc nào đó.

Hệ thống đang cài đặt với 3 loại message chính:

"song_lookup": Message Job đi crawl data.

"download_resource": Message Job download các resources như hình ảnh, âm thanh, lyric text.

"logs_server": Message Job bắn log tới Websocket Monitor.

Hệ thống được cài đặt bằng mã nguồn mở "Kafka".

Về cơ bản kafka là hệ thống message pub/sub phân tán mà có khả năng scale rất tốt.

Message của kafka được lưu trên đĩa cứng, đồng thời được replicate trong cluster giúp phòng tránh mất dữ liệu.

Kafka có thể hiểu là một hệ thống **logging**, nhằm lưu lại các **trạng thái** của hệ thống, nhằm phòng tránh mất thông tin.

2.2.6 Search Middleware System

Tính năng search là một tính năng quan trọng mà người dùng sử dụng rất nhiều trong hệ thống để tìm kiếm thông tin cần thiết. Hệ thống search hỗ trợ người dùng tìm kiếm tên bài hát cực kỳ nhanh chóng.

Hệ thống đang được cài đặt bằng mã nguồn mở "ElasticSearch".

ElasticSearch là một công cụ tìm kiếm cấp doanh nghiệp (enterprise-level search engine). Mục tiêu của nó là tạo thành một công cụ, nền tảng kỹ thuật tìm kiếm và phân tích thời gian thực (ở đây là nhanh trong và chính xác), cũng như cách để nó có thể áp dụng hay triển khai một cách dễ dàng vào các nguồn dữ liệu (data sources) khác nhau.

2.2.7 Identification Middleware System

Hệ thống có vài trò xác thực người dùng. Tạo và quản lý các session của người dùng.

Hệ thống cũng được cài đặt bằng mã nguồn mở "Thrift".

2.2.8 Crawler Worker

Worker có trách nhiệm nhận message job từ Streaming System để đi crawl data từ trang mp3.zing.vn như ca sỹ, bài hát, lời bài hát, album,...

Worker gồm một bể Executor gồm các tiến trình được khởi tạo sẵn để tăng hiệu năng crawl với lượng message lớn.

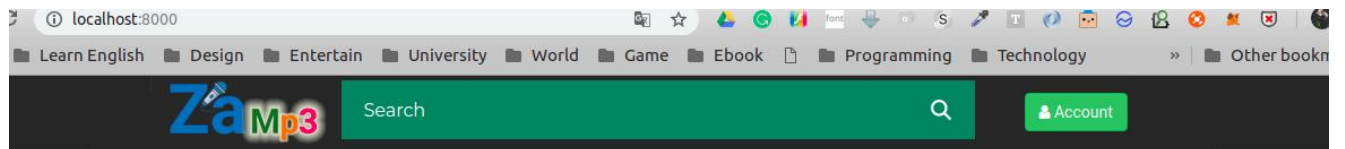
2.2.9 MongoDB

MongoDB là một chương trình cơ sở dữ liệu mã nguồn mở được thiết kế theo kiểu hướng đối tượng trong đó các bảng được cấu trúc một cách linh hoạt cho phép các dữ liệu lưu trên bảng không cần phải tuân theo một dạng cấu trúc nhất định nào. Chính do cấu trúc linh hoạt này nên MongoDB có thể được dùng để lưu trữ các dữ liệu có cấu trúc phức tạp và đa dạng và không cố định (hay còn gọi là Big Data).

2.2.10 KyotoCabinet

KyotoCabinet cũng là một loại database có performance rất cao trong việc lưu và đọc dữ liệu. KyotoCabinet là một nosql database hỗ trợ theo dạng key-value. Các object được serialize thành bytes rất tiện lợi.

3 Demo



Nhạc hot tháng 8



Website is developed by | NGUYEN THANH CHUNG

