**CÔNG TY CỔ PHẦN MISA JSC**

**---🙠**🕮**🙢---**



**NGÀY HỘI SẢN XUẤT**

**NGHIÊN CỨU GIẢI PHÁP ĐỒNG BỘ DỮ LIỆU REAL-TIME GIỮA SQL VÀ ELASTICSEARCH ĐÁP ỨNG TÌM KIẾM DỮ LIỆU LỚN**

**CHO CÁC SẢN PHẨM CỦA MISA**

BÁO CÁO NGHIÊN CỨU

Nhóm: Dream Chaser From CRM

**HÀ NỘI, 08-2022**

**CÔNG TY CỔ PHẦN MISA JSC**

**---🙠**🕮**🙢---**



**NGÀY HỘI SẢN XUẤT**

**NGHIÊN CỨU GIẢI PHÁP ĐỒNG BỘ DỮ LIỆU REAL-TIME GIỮA SQL VÀ ELASTICSEARCH ĐÁP ỨNG TÌM KIẾM DỮ LIỆU LỚN**

**CHO CÁC SẢN PHẨM CỦA MISA**

BÁO CÁO NGHIÊN CỨU

Nhóm: Dream Chaser From CRM

**Cán bộ hướng dẫn: Nguyễn Quang Hoàng**

**HÀ NỘI, 08-2022**

# **NHẬN XÉT CỦA BAN GIÁM KHẢO**

**……………………………………….………………………………………………………….**

**……………………………………….………………………………………………………….**

**……………………………………….………………………………………………………….**

**……………………………………….………………………………………………………….**

**……………………………………….………………………………………………………….**

**……………………………………….………………………………………………………….**

**……………………………………….………………………………………………………….**

**……………………………………….………………………………………………………….**

**……………………………………….………………………………………………………….**

**……………………………………….………………………………………………………….**

**……………………………………….………………………………………………………….**

**……………………………………….………………………………………………………….**

**……………………………………….………………………………………………………….**

**……………………………………….………………………………………………………….**

**……………………………………….………………………………………………………….**

**……………………………………….………………………………………………………….**

**……………………………………….………………………………………………………….**

**……………………………………….………………………………………………………….**

**……………………………………….……………………………………………………………………………………………………………………………………………………………**

|  |
| --- |
| Hà Nội, ngày …. Tháng …. Năm 2022 |
| Người nhận xét |
|  |

**Tóm Tắt:** Với sự phát triển mạnh mẽ của công ty cổ phầm MISA JSC, lượng sản phầm của MISA ngày càng phát triển cả về số lượng cũng như chất lượng, cùng với đó là tập dữ liệu khách hàng khổng lồ từ các phần mềm. MISA luôn mong muốn đem lại cho người dung trải nghiệm dịch vụ tốt nhất nhằm đem lại hiệu quả về năng suất cũng như chất lượng công việc lên hàng đầu, luôn đề cao tinh thần phụ sự xã hội. Tuy nhiên, cùng với việc phát triển lớn mạnh của các phần mềm nhà MISA thì dữ liệu từ khách hàng là vô cùng lớn, việc dữ liệu lớn được lưu trữ trên cơ sở dữ liệu quan hệ cho ra một hiệu suất truy vấn không được tốt, MISA cũng đã có những sự cô liên quan đến việc tím kiếm dữ liệu lớn. Các công cụ tìm kiếm ra đời nhằm giải quyết tình trạng này như: Apache Solr, MeiliSearch, Sphinx, Manticore và Elasticsearch là một trong những công cụ tìm kiếm giúp tôi ưu việc tìm kiếm nhanh dữ liệu lớn. Tuy nhiên, việc phát sinh vấn đề đồng bộ dữ liệu giữa cơ sở dữ liệu quan hệ sang cơ sở dữ liệu tìm kiếm là vấn đề lớn nhất, Apache Kafka là một nền tảng theo kiến trúc phân tán cho phép lưu trữ sự kiện và xử lý dữ liệu luồng mã nguồn mở được phát triển bởi Apache Software, với khả năng chịu lỗi cao và là hệ thống nhắn tin nhanh, kafka đang dần được thay thế cho hệ thống nhắn tin truyền thống. Nó được sử dụng cho các hệ thống nhắn tin thông thường trong các ngữ cảnh khác nhau. Trong nghiên cứu này, việc sử dụng Elasticsearch và Apache Kafka giải quyết bài toán tìm kiếm dữ liệu lớn và đồng bộ dữ liệu real-time giữa cơ sở dữ liệu quan hệ và cơ sở dữ liệu tìm kiếm sẽ được chúng tôi ứng dụng thực tế tại dự án CRM.

**MỤC LỤC**

[NHẬN XÉT CỦA BAN GIÁM KHẢO 3](#_Toc111843267)

[MỤC LỤC 5](#_Toc111843268)

[DANH MỤC HÌNH ẢNH 7](#_Toc111843269)

[DANH MỤC BẢNG BIỂU 8](#_Toc111843270)

[CHƯƠNG 1: GIỚI THIỆU CHUNG 9](#_Toc111843271)

[1.1 Lý do chọn đề tài 9](#_Toc111843272)

[1.2 Giới thiệu 9](#_Toc111843273)

[1.1.1 Bài toán tìm kiếm nhanh với dữ liệu lớn 9](#_Toc111843274)

[1.1.2 Tìm kiếm trên cơ sở dữ liệu quan hệ 9](#_Toc111843275)

[1.1.3 Tìm kiếm trên công cụ tìm kiếm elasticsearch 9](#_Toc111843276)

[1.1.4 Đồng bộ dữ liệu real-time dựa trên Apache Kafka 9](#_Toc111843277)

[1.3 Tiềm năng ứng dụng của bài toán tìm kiếm dữ liệu lớn và đồng bộ dữ liệu real-time tại công ty cổ phần MISA JSC 9](#_Toc111843278)

[1.4 Mục tiêu của nghiên cứ 9](#_Toc111843279)

[1.5 Cấu trúc bài nghiên cứu 9](#_Toc111843280)

[CHƯƠNG 2. A 10](#_Toc111843281)

[**2.1 A** 10](#_Toc111843282)

[2.1.1 A 10](#_Toc111843283)

[2.2 Kết luận 10](#_Toc111843284)

[CHƯƠNG 3. B 11](#_Toc111843285)

[**3.1 B** 11](#_Toc111843286)

[**3.2 B** 11](#_Toc111843287)

[**3.3 Kết luận** 11](#_Toc111843288)

[CHƯƠNG 4. C 11](#_Toc111843289)

[**4.1 C** 11](#_Toc111843290)

[**4.5 Kết quả** 11](#_Toc111843291)

[CHƯƠNG 5. KẾT LUẬN 11](#_Toc111843292)

[TÀI LIỆU THAM KHẢO 11](#_Toc111843293)

# **DANH MỤC HÌNH ẢNH**

**DANH MỤC BẢNG BIỂU**

# **CHƯƠNG 1: GIỚI THIỆU CHUNG**

## Lý do chọn đề tài

## 1.2 Giới thiệu

### **1.1.1 Bài toán tìm kiếm nhanh với dữ liệu lớn**

### **1.1.2 Tìm kiếm trên cơ sở dữ liệu quan hệ**

### **1.1.3 Tìm kiếm trên công cụ tìm kiếm elasticsearch**

### **1.1.4 Đồng bộ dữ liệu real-time dựa trên Apache Kafka**

## 1.3 Tiềm năng ứng dụng của bài toán tìm kiếm dữ liệu lớn và đồng bộ dữ liệu real-time tại công ty cổ phần MISA JSC

## 1.4 Mục tiêu của nghiên cứ

## 1.5 Cấu trúc bài nghiên cứu

Dựa trên mục tiêu cụ thể đã trình bày, bài nghiên cứu được chia ra làm năm chương với nội dung cụ thể như sau:

Chương 1: Giới thiệu chung

Chương 2: Tổng quan về tìm kiếm trên các loại cơ sở dữ liệu

Chương 3: Ứng dụng công cụ tìm kiếm nhanh

Chương 3: Đồng bộ dữ liệu real-time giữa CSDL gốc và CSDL tìm kiếm

Chương 4: Kết quả thực nghiệm

Chương 5: Kết luận

# **CHƯƠNG 2. A**

**2.1 A**

### **2.1.1 A**

## Kết luận

**CHƯƠNG 3. B**

**3.1 B**

**3.2 B**

**3.3 Kết luận**

**CHƯƠNG 4. C**

**4.1 C**

**4.5 Kết quả**

# **CHƯƠNG 5. KẾT LUẬN**

# **TÀI LIỆU THAM KHẢO**