

10

DỰ ÁN CRM (MISA) 10 TRIỆU BẢN GHI HOẠT ĐỘNG

NGHIÊN CỨU GIẢI PHÁP ĐỒNG BỘ DỮ LIỆU PHỤC VỤ TÌM KIẾM DỮ LIỆU LỚN CHO CÁC SẢN PHẨM CỦA MISA

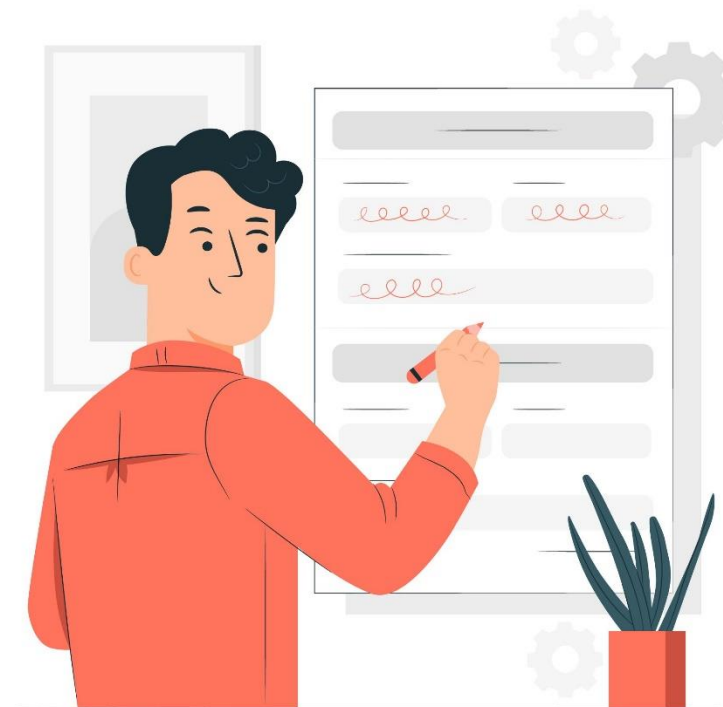
Dream Chaser From CRM

NỘI DUNG

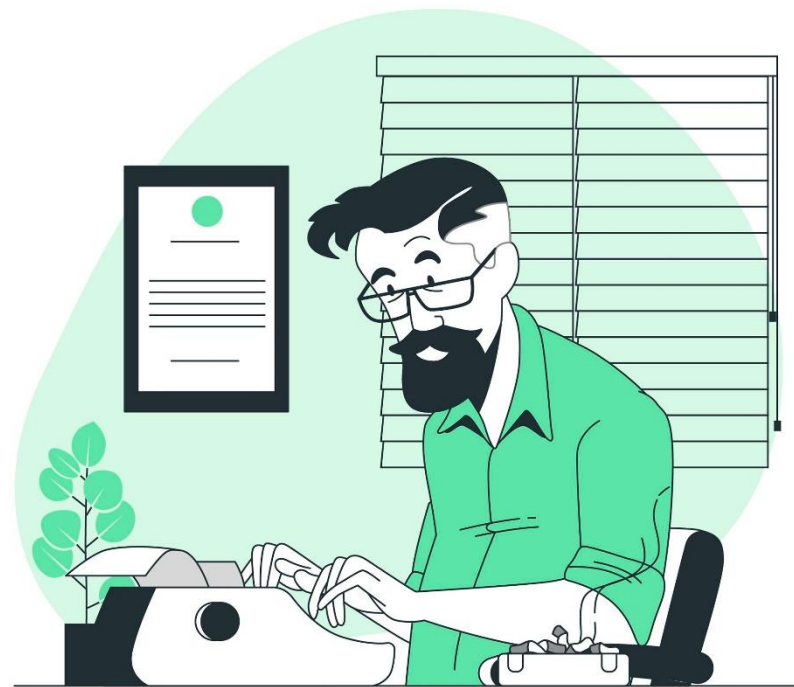
1 NÊU VẤN ĐỀ

2 GIẢI PHÁP ĐỒNG BỘ DỮ LIỆU REAL-TIME
GIỮA CSDL GỐC VÀ CSDL TÌM KIẾM

3 CASE STUDY CỦA TIKI VÀ ALIBABA



NÊU VẤN ĐỀ



NÊU VẤN ĐỀ

70.000+

ĐƠN VỊ HÀNH CHÍNH SỰ NGHIỆP
tin dùng sản phẩm MISA

170.000+

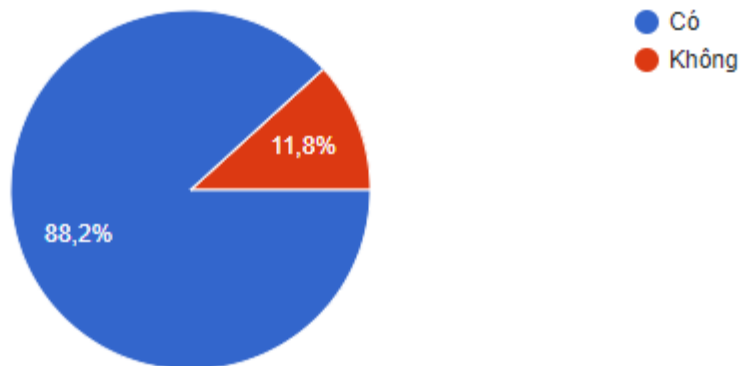
DOANH NGHIỆP
tin dùng sản phẩm MISA

2.500.000+

HỘ CÁ THỂ VÀ CÁ NHÂN
tin dùng sản phẩm MISA

2. Ứng dụng mà dự án của anh/chị đang phát triển có gặp vấn đề về dữ liệu lớn không?

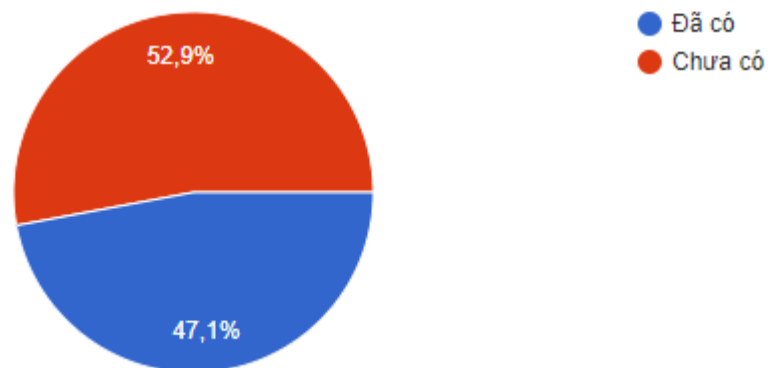
17 câu trả lời



NÊU VẤN ĐỀ

3. Nếu dự án của anh/chị đang gặp vấn đề về lượng dữ liệu lớn, gây ảnh hưởng đến hiệu năng của sản phẩm, thì anh/chị đã có giải pháp nào để giải quyết vấn đề đó chưa?

17 câu trả lời



NÊU VẤN ĐỀ



elasticsearch

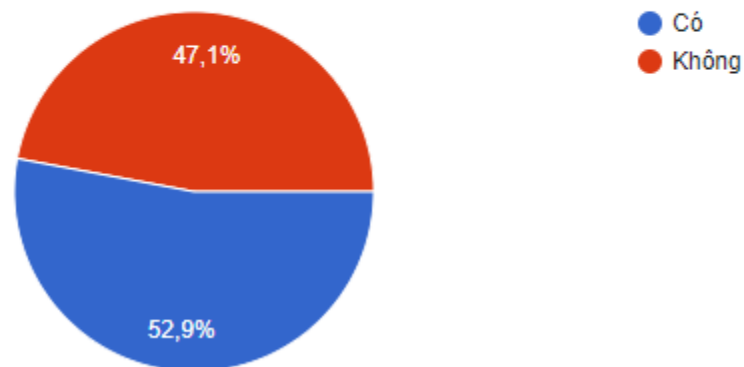


Open Source Search Server



6. Anh/chị có nghĩ Elasticsearch là câu trả lời cho vấn đề dữ liệu lớn của các ứng dụng của MISA không?

17 câu trả lời



NÊU VẤN ĐỀ



GIẢI PHÁP ĐỒNG BỘ DỮ LIỆU REAL-TIME GIỮA CSDL GỐC VÀ CSDL TÌM KIẾM

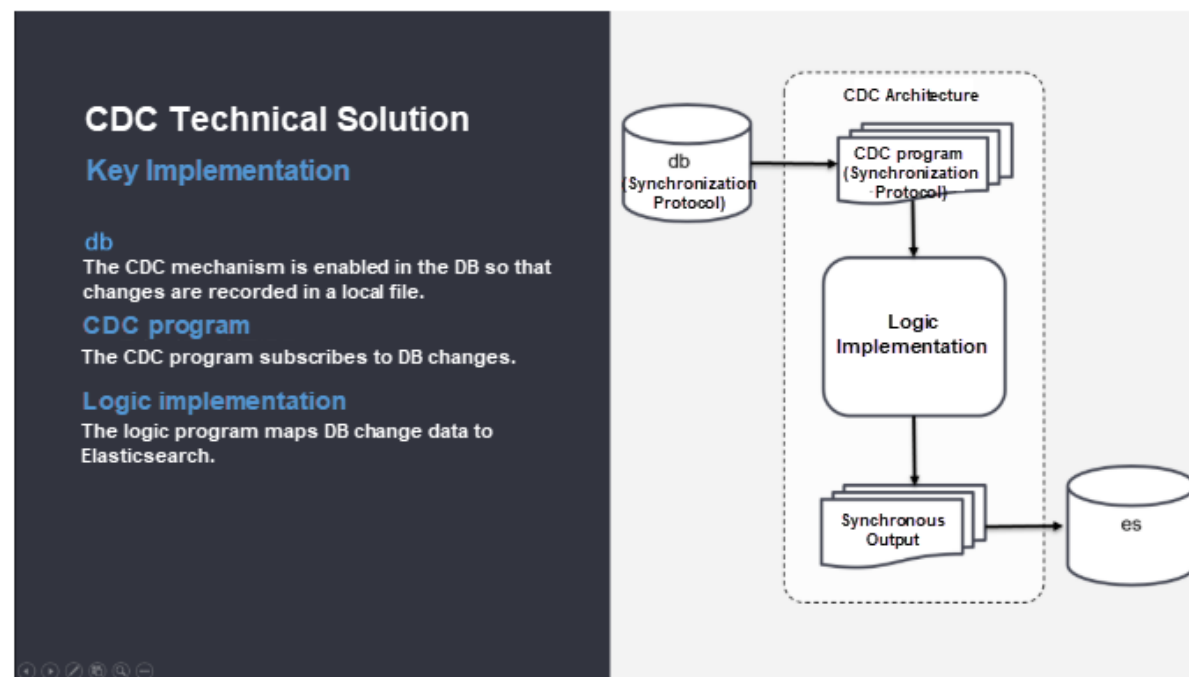


GIẢI PHÁP ĐỒNG BỘ DỮ LIỆU REAL-TIME GIỮA CSDL GỐC VÀ CSDL TÌM KIẾM

LƯU DỮ LIỆU ĐỒNG BỘ

LƯU DỮ LIỆU KHÔNG ĐỒNG BỘ

CHANGE DATA CAPTURE (CDC)



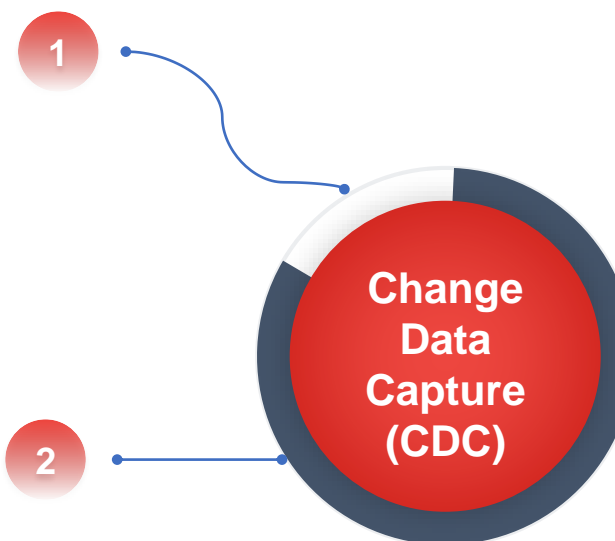
CHANGE DATA CAPTURE (CDC)

BinLog

Là một tệp nhị phân trên đĩa, chứa tất cả các sự kiện thay đổi nội dung hoặc cấu trúc của cơ sở dữ liệu MySQL

Replicate binary log

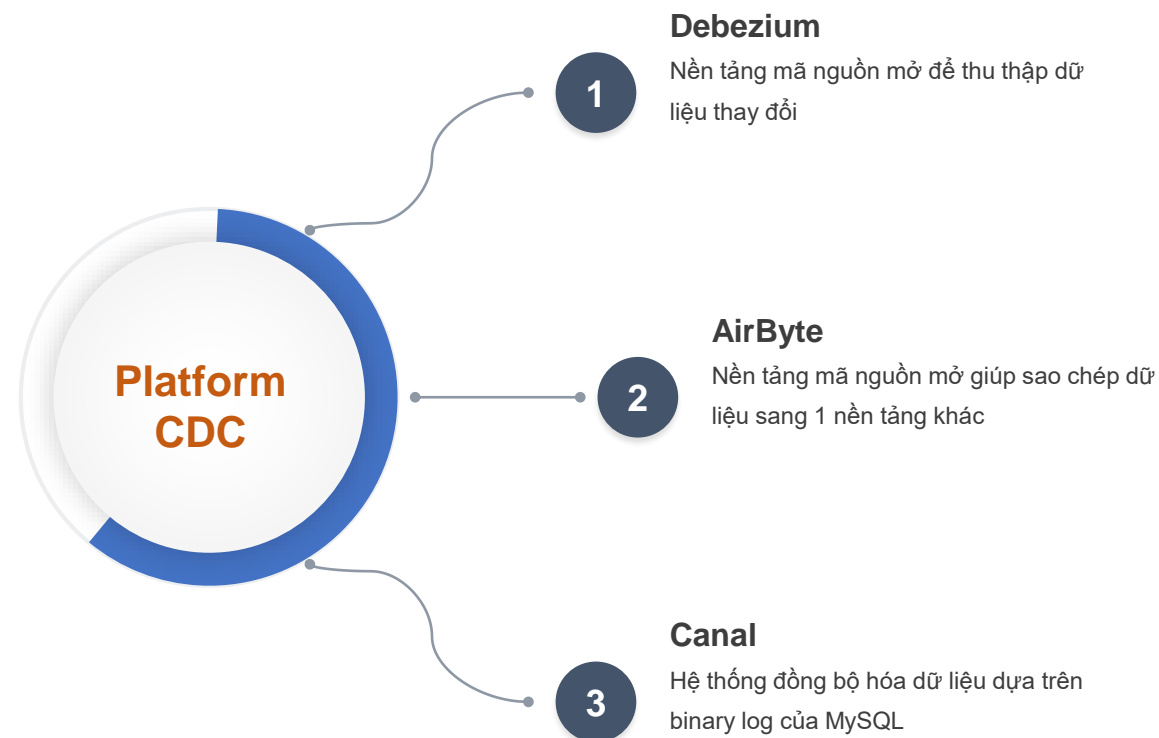
streamming tất cả các thay đổi của database từ master tới slave



CHANGE DATA CAPTURE (CDC)

Bảng 1: So sánh các tool đồng bộ theo cơ chế CDC

	Debezium	Airbyte	Canal
Khả năng kết nối nhiều hệ quản trị CSDL	SQLServer MySQL PostgreSQL ...	SQLServer MySQL PostgreSQL ...	Chỉ hỗ trợ MySQL
Kết nối không giới hạn cơ sở dữ liệu	Có	Có	Có
Trà phí	Miễn phí	Có phí	Miễn phí



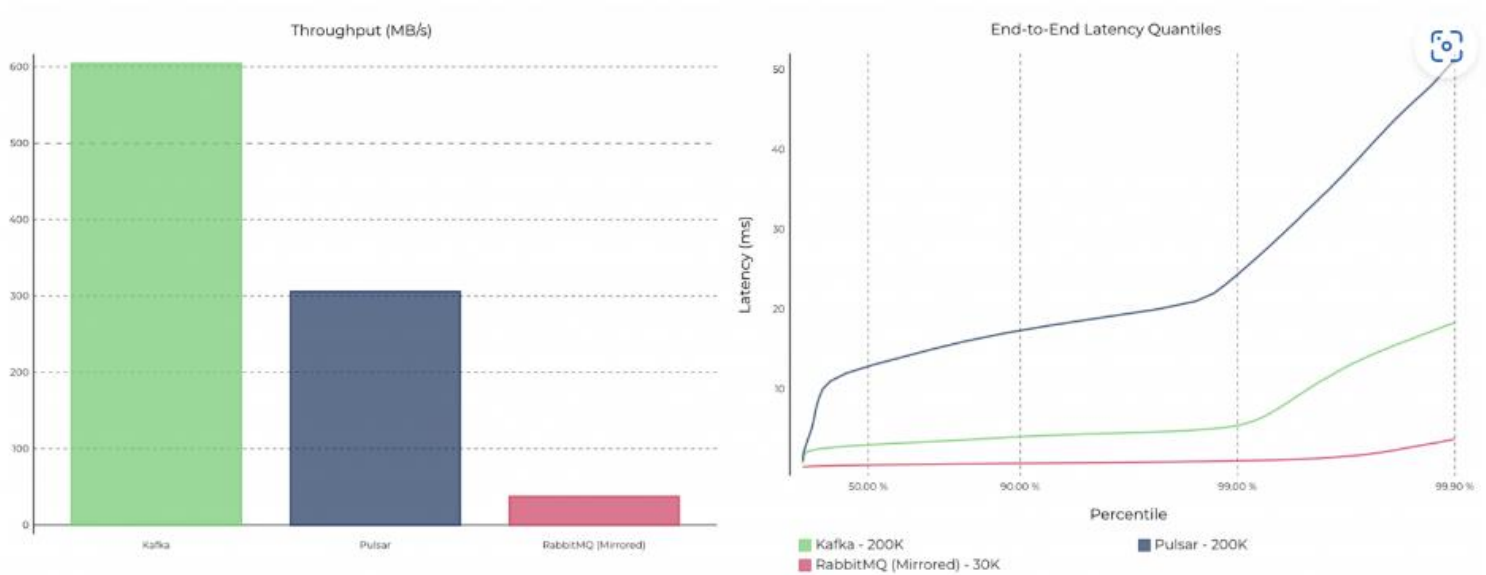
MESSAGE QUEUE



MESSAGE QUEUE

Kafka cung cấp thông lượng ghi nhanh hơn 15 lần so với RabbitMQ

Kafka có trễ thấp với thông lượng cao, còn RabbitMQ có thể đạt được độ trễ đầu cuối thấp hơn Kafka nhưng chỉ ở mức thông lượng thấp hơn đáng kể



	Kafka	Pulsar	RabbitMQ (Mirrored)
Peak Throughput (MB/s)	605 MB/s	305 MB/s	38 MB/s
p99 Latency (ms)	5 ms (200 MB/s load)	25 ms (200 MB/s load)	1 ms* (reduced 30 MB/s load)

KẾT LUẬN

MÔ HÌNH TRIỂN KHAI

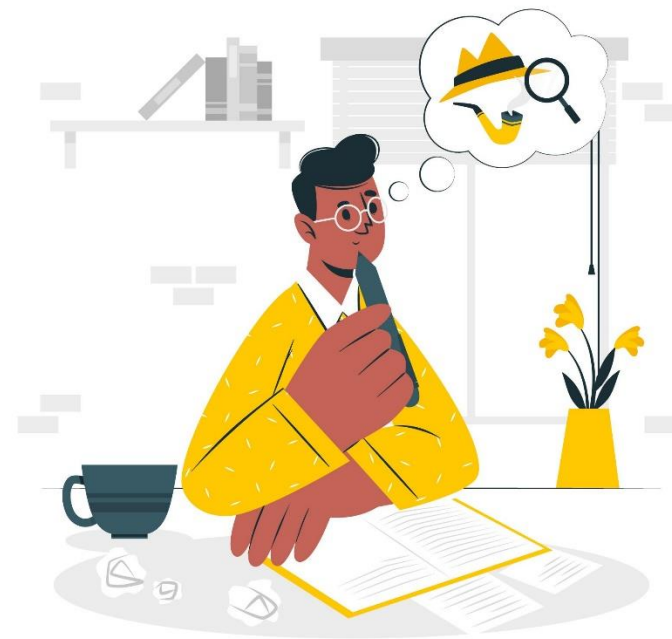
Database
nguồn



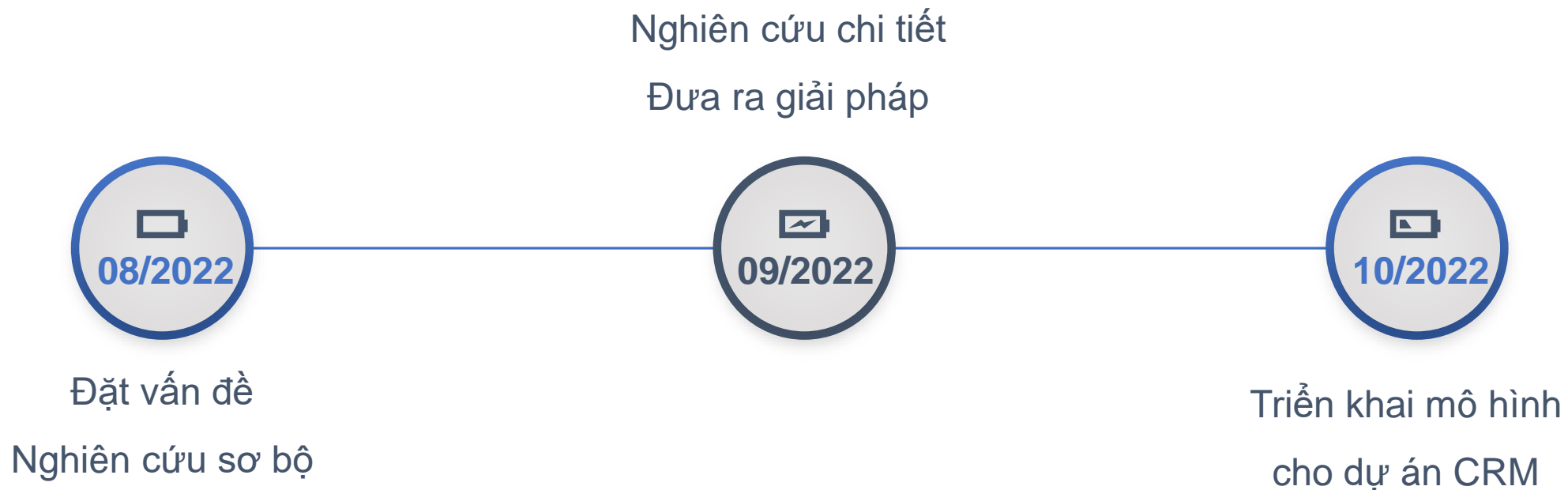
Elastic search



KẾ HOẠCH HÀNH ĐỘNG



KẾ HOẠCH HÀNH ĐỘNG



THANK YOU



Q & A