

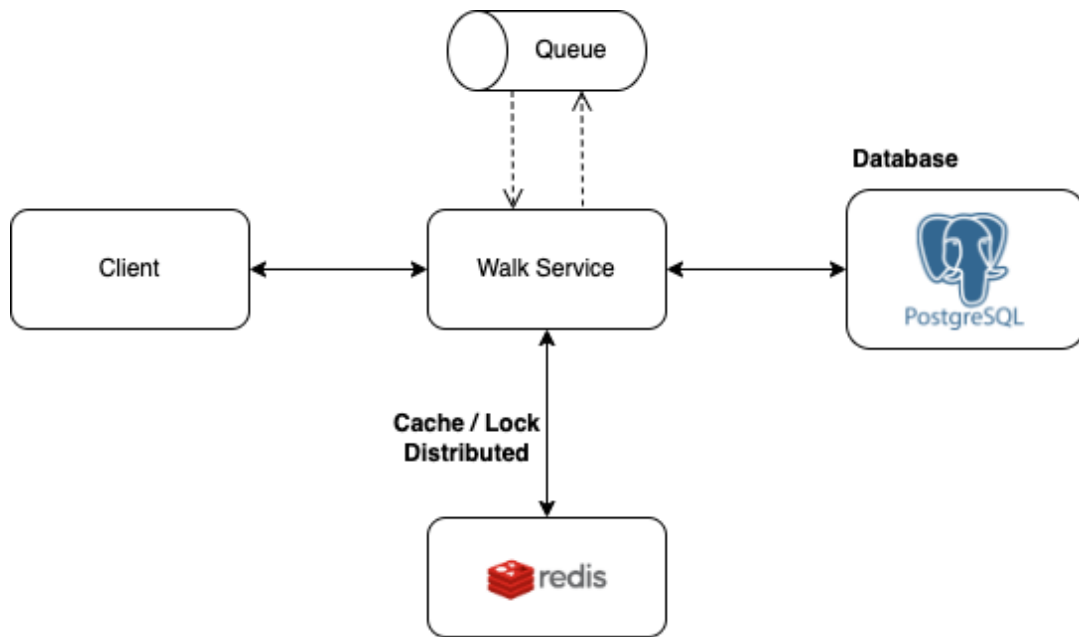
Tài Liệu

Tóm Tắt Kiến Trúc Hệ Thống Và Thiết Kế Cơ Sở Dữ Liệu
Tính Năng Di Bộ

Mục lục

1.	Tổng quan hệ thống	3
2.	Chọn cơ sở dữ liệu	4
3.	Diagram cơ sở dữ liệu	5
4.	Danh sách các bảng	6
5.	Mô tả chi tiết các bảng	7
	5.1 Bảng steps	7
	5.2 Bảng step_archives	8
	5.3 Bảng schedules	9
	5.4 Bảng daily_ranking	10

1. Tổng quan hệ thống



Tổng quan hệ thống

Hệ thống sẽ bao gồm các thành phần chính sau:

1. **Client:** Ứng dụng MoMo trên thiết bị của người dùng để ghi nhận số bước chân.
2. **Walk Service (Spring Boot):** Dịch vụ xử lý logic, ghi nhận và truy xuất dữ liệu.
3. **Redis:** Sử dụng redis để cache dữ liệu tạm thời như bảng xếp hạng, tổng số bước chân của người dùng theo tuần, tháng. Redis được sử dụng để quản lý lock distributed cho việc xử lý dữ liệu và tránh xung đột khi scale hệ thống chạy nhiều instance.
4. **Queue:** Hệ thống sử dụng redis làm queue. Khi client gửi yêu cầu ghi nhận bước chân, yêu cầu sẽ được đưa vào queue. Các yêu cầu sẽ được xử lý theo batch từ đó giảm tải cho database, không bị miss request khi số lượng người dùng đồng thời tăng đột biến.
5. **Database (PostgreSQL):** Lưu trữ dữ liệu về số bước chân của người dùng.

Hệ thống có thể scale dễ dàng bằng cách:

- Sử dụng Kubernetes để quản lý việc scale pods, tăng số lượng instances của microservices dựa trên nhu cầu
- Sử dụng kỹ thuật Shard hoặc Partition cho cơ sở dữ liệu PostgreSQL để phân tán dữ liệu người dùng.

2. Chọn cơ sở dữ liệu

Việc chọn cơ sở dữ liệu phù hợp là một yếu tố quan trọng trong việc phát triển một ứng dụng. Với yêu cầu của bài toán cả PostgreSQL và MongoDB đều đáng được cân nhắc. PostgreSQL hỗ trợ ACID, truy vấn phức tạp phù hợp với các ứng dụng có tính chất quan hệ, như hệ thống bảng xếp hạng và tính toán tổng số bước chân của người dùng theo tuần/tháng. Nhược điểm cần biết trước cấu trúc dữ liệu, khó scale theo chiều ngang. MongoDB được thiết kế để xử lý dữ liệu lớn và có hiệu năng tốt khi scale ngang, cấu trúc dữ liệu linh hoạt. Nhược điểm là không hỗ trợ quan hệ dữ liệu mạnh mẽ và không đảm bảo ACID, không phù hợp cho các truy vấn phức tạp.

Tóm lại, với tính năng Đi bộ trong ứng dụng MoMo, việc thu thập, xử lý và tính toán số bước chân của nhiều người dùng theo ngày là rất quan trọng. PostgreSQL với tính năng tính toán vẹn dữ liệu và độ tin cậy cao là lựa chọn phù hợp để đảm bảo tính chính xác và hiệu năng của hệ thống, đặc biệt khi số lượng người dùng và dữ liệu ngày càng tăng lên.

Khi lưu số bước chân của người dùng sẽ lưu vào 2 bảng **'steps'** và **'step_archives'**. Bảng **'steps'** chỉ lưu bước chân của tháng hiện tại vì vậy sẽ có bộ lập lịch để dọn dẹp dữ liệu của tháng trước đó vào đầu tháng. Để phục vụ việc xem bảng xếp hạng nhanh nhất và đảm bảo hiệu năng của hệ thống, hệ thống sẽ tạo bảng **'daily_ranking'** là **materialized view** được làm mới 30 phút mỗi lần.

3. Diagram cơ sở dữ liệu

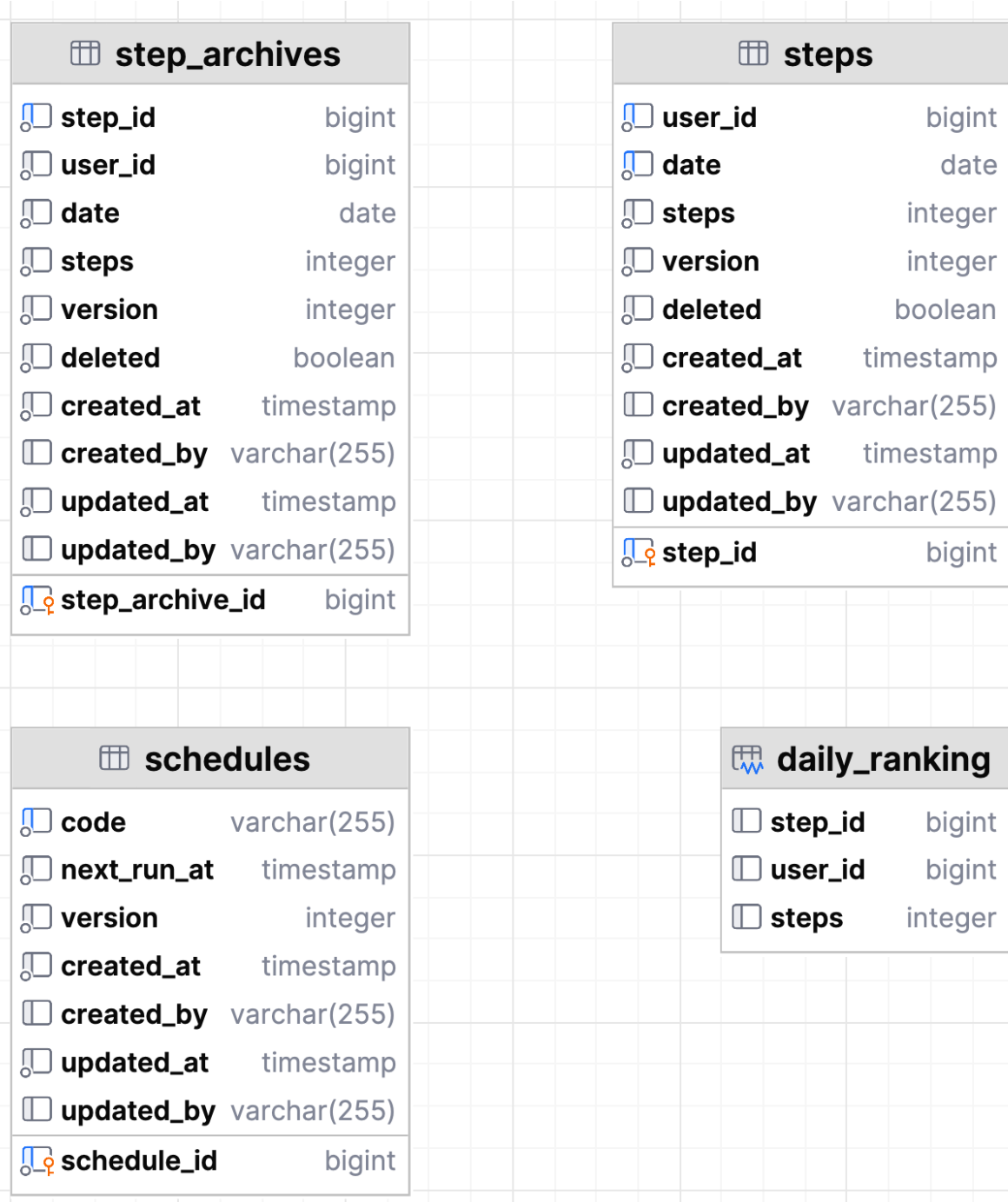


Diagram cơ sở dữ liệu

4. Danh sách các bảng

STT	Tên bảng	Mô tả
1	steps	Lưu dữ liệu số bước chân theo ngày của người dùng trong tháng hiện tại
2	step_archives	Lưu toàn bộ dữ liệu số bước chân theo ngày của người dùng
3	schedules	Lưu lịch trình để dọn dẹp bảng steps
4	daily_ranking	Bảng materialized view giúp xem bảng xếp hạng số bước chân theo ngày hiện tại

5. Mô tả chi tiết các bảng

5.1 Bảng steps

STT	Tên trường	Kiểu dữ liệu	Nullable	Ràng buộc	Mô tả
1	step_id	bigint	No	PK	Mã định danh step
2	user_id	bigint	No		Mã định danh người dùng
3	date	date	No		Ngày ghi nhận
4	steps	int	No		Số bước chân người dùng đi được
5	version	int	No		
6	deleted	bool	No		Đã xoá mềm hay chưa
7	created_at	timestampz	No		Tạo lúc nào
8	created_by	varchar(255)	Yes		Tạo bởi ai
9	updated_at	timestampz	No		Cập nhật lúc nào
10	updated_by	varchar(255)	Yes		Cập nhật bởi ai

5.2 Bảng step_archives

STT	Tên trường	Kiểu dữ liệu	Nullable	Ràng buộc	Mô tả
1	step_archive_id	bigint	No	PK	Mã định danh step archive
2	step_id	bigint	No	PK	Mã định danh step
3	user_id	bigint	No		Mã định danh người dùng
4	date	date	No		Ngày ghi nhận
5	steps	int	No		Số bước chân người dùng đi được
6	version	int	No		
7	deleted	bool	No		Đã xoá mềm hay chưa
8	created_at	timestamptz	No		Tạo lúc nào
9	created_by	varchar(255)	Yes		Tạo bởi ai
10	updated_at	timestamptz	No		Cập nhật lúc nào
11	updated_by	varchar(255)	Yes		Cập nhật bởi ai

5.3 Bảng schedules

STT	Tên trường	Kiểu dữ liệu	Nullable	Ràng buộc	Mô tả
1	schedule_id	bigint	No	PK	Mã định danh schedule
2	code	varchar(255)	No	Unique	Code của lịch trình
3	next_run_at	timestamptz	No		Lịch trình sẽ chạy tiếp lúc nào
4	version	int	No		
5	created_at	timestamptz	No		Tạo lúc nào
6	created_by	varchar(255)	Yes		Tạo bởi ai
7	updated_at	timestamptz	No		Cập nhật lúc nào
8	updated_by	varchar(255)	Yes		Cập nhật bởi ai

5.4 Bảng daily_ranking

STT	Tên trường	Kiểu dữ liệu	Nullable	Ràng buộc	Mô tả
1	step_id	bigint	No		Mã định danh step
2	user_id	bigint	No		Mã định danh người dùng
3	steps	bigint	No		Tổng số bước chân ngày hiện tại của người dùng