**Tóm tắt Bộ Bài Tập Hệ Quản Trị Cơ Sở Dữ Liệu**

**I. Mục đích Cốt Lõi**

| **Mục tiêu** | **Nội dung chi tiết** |
| --- | --- |
| **1. Xây dựng nền tảng vững chắc** | Bắt đầu từ thiết kế lược đồ, xác định khóa chính, khóa ngoại, các ràng buộc toàn vẹn để đảm bảo cấu trúc dữ liệu tốt. |
| **2. Kết nối lý thuyết và thực hành** | Kết hợp truy vấn bằng đại số quan hệ và SQL để hiểu rõ bản chất toán học của truy vấn. |
| **3. Trang bị kỹ năng truy vấn đa tầng** | Bao gồm các kỹ thuật từ cơ bản (SELECT, WHERE) đến nâng cao (CTE, Subquery, Window Functions). |
| **4. Tối ưu hóa hiệu năng** | Sử dụng EXPLAIN ANALYZE, chỉ mục, view vật lý hóa để tối ưu cách truy vấn và thực thi. |
| **5. Hiểu hoạt động bên trong CSDL** | Tìm hiểu cách PostgreSQL lựa chọn chiến lược thực thi: Sequential Scan, Index Scan, JOIN Strategy. |

**II. Kỹ Thuật Sử Dụng**

**1. Thiết kế & Định nghĩa lược đồ (DDL)**

| **Kỹ thuật** | **Mô tả** |
| --- | --- |
| Tạo cơ sở dữ liệu & bảng | Sử dụng CREATE DATABASE, CREATE TABLE để khởi tạo lược đồ. |
| Kiểu dữ liệu phổ biến | TEXT, VARCHAR, INT, NUMERIC, DATE, TIMESTAMP, CHAR. |
| Ràng buộc toàn vẹn | PRIMARY KEY (có thể phức hợp), FOREIGN KEY, NOT NULL, UNIQUE, CHECK. |

**2. Thao tác và Sinh dữ liệu**

| **Kỹ thuật** | **Mô tả** |
| --- | --- |
| Chèn dữ liệu | Sử dụng INSERT INTO. |
| Sinh dữ liệu tự động | Dùng generate\_series, khối DO $$ ... $$. |
| Lập trình thủ tục | Dùng PROCEDURE, FUNCTION để xử lý logic phức tạp. |

**3. Truy vấn Dữ liệu (SQL & Đại số quan hệ)**

| **Nhóm kỹ thuật** | **Cụ thể** |
| --- | --- |
| **Đại số quan hệ** | Phép Chọn (σ), Phép Chiếu (π), Phép Kết nối (⨝). |
| **SQL cơ bản** | SELECT, FROM, WHERE. |
| **JOINs** | INNER JOIN. |
| **Lọc nâng cao** | BETWEEN, LIKE, IN, EXISTS. |
| **Tổng hợp & nhóm** | GROUP BY, HAVING với COUNT, SUM, AVG. |
| **Sắp xếp & phân trang** | ORDER BY, LIMIT, OFFSET. |
| **Truy vấn con** | Subquery trong SELECT, WHERE, FROM. |
| **CTEs** | WITH ... AS để viết truy vấn rõ ràng hơn. |
| **Window Functions** | RANK(), ROW\_NUMBER(), OVER (PARTITION BY ...). |

**4. Tối ưu hóa Hiệu năng**

**a. Chỉ mục (Indexes)**

| **Loại chỉ mục** | **Mô tả** |
| --- | --- |
| Đơn cột | CREATE INDEX ON table(column); |
| Chỉ mục phức hợp | Trên nhiều cột kết hợp. |
| Chỉ mục thứ tự giảm | Dùng DESC. |
| Chỉ mục một phần | Dùng điều kiện WHERE để chỉ tạo trên tập con. |

**b. Khung nhìn (Views)**

| **Loại View** | **Mô tả** |
| --- | --- |
| View logic | CREATE VIEW giúp tái sử dụng truy vấn phức tạp. |
| View vật lý | CREATE MATERIALIZED VIEW lưu kết quả để tăng tốc độ. |

**c. Phân tích Kế hoạch Truy vấn (Query Plan)**

| **Công cụ/Kỹ thuật** | **Mục đích** |
| --- | --- |
| EXPLAIN ANALYZE | Phân tích chi phí và chiến lược thực thi truy vấn. |
| Kiểu quét | Sequential Scan, Index Scan. |
| Chiến lược JOIN | Nested Loop Join, Hash Join, Merge Join. |
| Ép chiến lược | Dùng SET enable\_hashjoin = off để thử nghiệm. |

**d. Viết lại truy vấn tối ưu (Query Rewriting)**

| **Kỹ thuật** | **Mục đích** |
| --- | --- |
| Selection Pushdown | Đưa điều kiện WHERE vào sâu nhất. |
| Projection Pushdown | SELECT chỉ cột cần thiết trong subquery. |
| Join Reordering | JOIN bảng nhỏ trước để giảm bản ghi trung gian. |
| Subquery sang JOIN | So sánh IN/EXISTS với JOIN. |
| Loại bỏ JOIN thừa | Giảm chi phí truy vấn. |
| Tránh hàm trên cột có index | Tránh vô hiệu hóa index. |
| Tối ưu phân trang | OFFSET vs Keyset Pagination. |

**III. Tổng kết**

Bộ bài tập là công cụ học tập toàn diện, đào tạo sinh viên hoặc lập trình viên:

* **Có tư duy logic nền tảng** về dữ liệu
* **Viết truy vấn hiệu quả, rõ ràng, tối ưu**
* **Hiểu bản chất hoạt động** của hệ thống CSDL

Đây không chỉ là tài liệu luyện tập mà còn là nền tảng để xây dựng các hệ thống xử lý dữ liệu có quy mô và hiệu năng cao.