

MÃ ĐỀ

01

Tập đoàn Điện lực Việt Nam hiện muốn quản lý thông tin các nhà máy điện trực tiếp tham gia thị trường phát điện cạnh tranh. Tùy theo nguồn năng lượng được biến đổi thành điện năng mà các nhà máy được phân loại thành hai nhóm là Thủy điện (PhanLoai = “Thủy điện”) và Tuabin khí (PhanLoai = “Tuabin khí”). Ngoài ra, mỗi nhà máy sẽ được một trong hai cơ quan trực tiếp quản lý, đó là PVPower (CQQL = “PVPower”) và GENCO 3 (CQQL = “GENCO3”). Thông tin của các nhà máy điện được thể hiện qua lược đồ quan hệ sau:

NHAMAYDIEN (MaNMĐ, TenNMĐ, PhanLoai, CQQL, CongSuat)

Tên từ: Mỗi nhà máy điện có các thông tin bao gồm: Mã nhà máy điện (MaNMĐ), Tên nhà máy điện (TenNMĐ), phân loại nhà máy (PhanLoai), cơ quan quản lý (CQQL), công suất phát điện (CongSuat) được tính theo đơn vị MW.

1. (5 điểm) Cho thể hiện của quan hệ **NHAMAYDIEN** như sau:

MaNMĐ	TenNMĐ	PhanLoai	CQQL	CongSuat
NM01	Đak Đrinh	Thủy điện	PVPower	125
NM02	Hủa Na	Thủy điện	PVPower	180
NM03	Nhơn Trạch 2	Tuabin khí	PVPower	750
NM04	Sông Hình	Thủy điện	GENCO3	70
NM05	Phú Mỹ 1	Tuabin khí	GENCO3	1059
NM06	Buôn Kuốp	Thủy điện	GENCO3	280
NM07	Nhơn Trạch 1	Tuabin khí	PVPower	450
NM08	Thác Bà	Thủy điện	GENCO3	120

Giả sử có ba ứng dụng truy xuất đến **NHAMAYDIEN**:

Q1: SELECT COUNT(*) FROM **NHAMAYDIEN** WHERE **PhanLoai** = value

Q2: SELECT **MaNMĐ**, **TenNMĐ** FROM **NHAMAYDIEN** WHERE **CQQL** = value

Q3: SELECT * FROM **NHAMAYDIEN** WHERE **CongSuat** > 400

- Dùng giải thuật *COM_MIN*, tính Pr' thỏa tối tiểu và đầy đủ? (2 điểm)
- Sử dụng thuật toán *PHORIZONTAL*, thiết kế phân mảnh ngang chính cho quan hệ **NHAMAYDIEN**? Lưu ý, sinh viên không cần liệt kê dữ liệu trong các mảnh. (2 điểm)
- Hãy vẽ cây phân mảnh của lược đồ quan hệ **NHAMAYDIEN** đã làm. Chứng minh điều kiện **tái tạo** của phân mảnh này là đúng đắn. (1 điểm)

2. (5 điểm) Cho tập $Q = \{q1, q2, q3, q4, q5\}$ các truy vấn, tập $A = \{A1, A2, A3, A4, A5\}$ lần lượt là các thuộc tính MaNMD, TenNMD, PhanLoai, CQQL, CongSuat của quan hệ:

NHAMAYDIEN (MaNMD, TenNMD, PhanLoai, CQQL, CongSuat)

Tập $S = \{S1, S2, S3, S4\}$ các vị trí (sites) trong hệ cơ sở dữ liệu phân tán. Giả sử số truy xuất đến các cặp thuộc tính tại các vị trí được cho theo hàm:

$$ref_i(q_j) = \begin{cases} 1, \forall i \in [1,4], j \in \{1; 3; 5\} \\ 2, \forall i \in [1,4], j \in \{2; 4\} \end{cases}, \text{ với } i \text{ là số chỉ vị trí (site), } j \text{ là số chỉ của câu truy vấn.}$$

Ma trận truy vấn sử dụng thuộc tính (use) và ma trận tần số sử dụng truy vấn (acc) tại các vị trí như sau:

	A1	A2	A3	A4	A5
q1	0	1	1	0	1
q2	0	0	1	1	0
q3	1	0	1	0	1
q4	1	0	0	1	0
q5	0	1	0	0	1

Ma trận USE

	S1	S2	S3	S4
q1	3	0	5	0
q2	0	2	0	6
q3	1	0	3	0
q4	0	3	0	3
q5	2	0	1	3

Ma trận ACC

- Tính ma trận AA. (1 điểm)
- Sử dụng thuật toán BEA tính ma trận CA. (2 điểm)
- Sử dụng thuật toán PARTITION để tìm ra hai phân mảnh dọc của quan hệ **NHAMAYDIEN**. (2 điểm)

(----- Hết -----)

XÁC NHẬN CỦA KHOA

GIẢNG VIÊN RA ĐỀ

Nguyễn Hồ Duy Tri

Ghi chú: Đề thi này gồm có 2 trang. Cán bộ coi thi không giải thích gì thêm.