

Tổng hợp Kiến thức Kỹ thuật Dữ liệu (Data Engineering)

Học viên Cao học KHMT Bách Khoa TP.HCM (HCMUT)

Ngày 31 tháng 12 năm 2025

1 Nguyên lý Phân tích & Thiết kế CSDL

1.1 Tổng quan Các giai đoạn Thiết kế

- Mức Quan niệm (Conceptual):** Mục tiêu — nǎm bắt yêu cầu và ngữ nghĩa (độc lập với cài đặt); Mô hình/Công cụ — ER/EER; Đầu ra — lược đồ quan niệm (thực thể, thuộc tính, mối kết hợp, ràng buộc).
- Mức Logic:** Mục tiêu — ánh xạ từ mức quan niệm sang mô hình DBMS đích (ví dụ: quan hệ); Mô hình/Công cụ — ánh xạ ER-sang-quan hệ, chuẩn hóa (FDs); Đầu ra — lược đồ quan hệ (bảng, khóa, ràng buộc toàn vẹn).
- Mức Vật lý:** Mục tiêu — xác định cấu trúc lưu trữ và đường dẫn truy xuất để tối ưu hiệu năng; Mô hình/Công cụ — phân tích tải, chỉ mục, tổ chức tập tin, băm; Đầu ra — lược đồ trong (cấu trúc lưu trữ, chỉ mục, đường dẫn truy xuất).

1.2 Nguyên lý Thiết kế Mức Quan niệm

- Phân tích Yêu cầu:** Làm việc với người dùng/chuyên gia nghiệp vụ để nǎm bắt yêu cầu dữ liệu và yêu cầu chức năng (thao tác/giao dịch).
- Thành phần ER: Thực thể** (vd: NhanVien), **Thuộc tính** (đơn/phức hợp/đa trị/dẫn xuất), **Mối kết hợp** (sự liên kết giữa các thực thể).
- Ràng buộc Cấu trúc:** Tỷ số bản số (1:1, 1:N, M:N) và ràng buộc tham gia (toàn phần vs từng phần).
- Thực thể Yếu:** Được xác định thông qua **mối kết hợp xác định** với một thực thể chủ và một **khóa bộ phận**; thực thể yếu tham gia **toàn phần** vào mối kết hợp xác định.
- Tinh chỉnh Top-Down:** Tinh chỉnh lặp lại các thực thể tổng quát; áp dụng chuyên biệt hóa/tổng quát hóa (EER).

1.3 Nguyên lý Thiết kế Mức Logic

1.3.1 Cơ bản về Mô hình Quan hệ

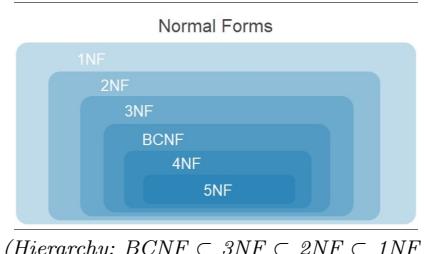
- Cấu trúc:** Lược đồ quan hệ $R(A_1, \dots, A_n)$; các bộ (tuple) không có thứ tự và không cho phép trùng lặp trong mô hình hình thức.

- Ràng buộc Toàn vẹn:** Ràng buộc miền giá trị (nguyên tố, có kiểu), ràng buộc khóa (siêu khóa/khóa ứng viên/khóa chính), toàn vẹn thực thể (khóa chính không NULL), toàn vẹn tham chiếu (giá trị khóa ngoại phải xuất hiện trong khóa chính được tham chiếu).

1.3.2 Lý thuyết Chuẩn hóa (Normalization Theory)

- Mục tiêu:** Giảm thiểu dư thừa, tránh dị thường (Thêm, Xóa, Sửa).
- Phụ thuộc Hàm (FD):** $X \rightarrow Y$ (Nếu $t_1[X] = t_2[X]$ thì $t_1[Y] = t_2[Y]$).
- Các loại Key:**
 - Superkey:* Xác định duy nhất một bộ.
 - Candidate Key:* Superkey tối thiểu.
 - Prime Attribute:* Thuộc tính nằm trong bất kỳ Candidate Key nào.

Các Dạng Chuẩn (Normal Forms)



1NF (Atomic): Miền giá trị nguyên tố.

→ *Vi phạm:* Thuộc tính đa trị, lồng nhau, lặp lại nhóm.

2NF (No Partial): Là 1NF + Thuộc tính *non-prime* phụ thuộc đầy đủ vào khóa.

→ *Vi phạm:* $\exists X \subsetneq Key$ sao cho $X \rightarrow NonPrime$.

(Chỉ xảy ra nếu Key là khóa phức hợp).

3NF (No Transitive): Là 2NF + Không có phụ thuộc bắc cầu giữa các *non-prime*.

→ *Định nghĩa:* Với mọi $X \rightarrow A$ (không tầm thường), phải thỏa: (a) X là Superkey **HOẶC** (b) A là Prime Attribute.

BCNF (Strict): Nghiêm ngặt hơn 3NF.

→ *Định nghĩa:* Với mọi $X \rightarrow A$, X **bắt buộc** là Superkey.

(Khác biệt: BCNF không chấp nhận ngoại lệ " A là Prime" như 3NF).

Tính chất Phân rã (Decomposition Properties)

- Kết nối bảo toàn thông tin (Lossless Join):** (*Bắt buộc*) Để phân rã R thành R_1, R_2 không bị mất dữ liệu, điều kiện là: $(R_1 \cap R_2) \rightarrow R_1$ HOẶC $(R_1 \cap R_2) \rightarrow R_2$.
(Giao của 2 bảng phải là khóa của ít nhất 1 bảng).
- Bảo toàn phụ thuộc (Dependency Preservation):** Các FD ban đầu có thể được kiểm tra riêng lẻ trên từng R_i mà không cần join lại. (BCNF có thể không bảo toàn phụ thuộc).

1.3.3 Phi chuẩn hóa (Denormalization)

- Mục tiêu:** Cải thiện hiệu suất đọc bằng cách đưa dư thừa vào lược đồ, ngược với chuẩn hóa.
- Động lực:** Tránh Join tốn kém; giảm độ phức tạp truy vấn; tăng locality dữ liệu.

Kỹ thuật Denormalization:

- Materialized Views:** Kết quả truy vấn được tính trước và lưu trữ; cập nhật khi dữ liệu thay đổi.
- Precomputed Aggregates:** Lưu giá trị tổng hợp (COUNT, SUM) trong bản ghi để tránh tính toán.
VD: Lưu số email chưa đọc trong bảng User thay vì đếm mỗi lần.
- Document Databases:** Nhúng (embedding) dữ liệu liên quan trong một document thay vì tham chiếu.
VD: MongoDB nhúng thông tin Worker vào document Project.
- Star Schema (DW):** Dimension tables được denormalize (VD: Brand, Category trong dim_product).
- Microservices:** Sao chép (replicate) dữ liệu giữa các service để tách biệt và giảm phụ thuộc.

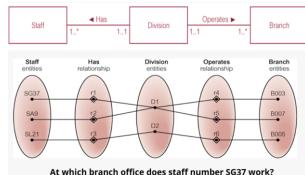
Dánh dối (Trade-offs):

- Write Overhead:** Mỗi cập nhật phải sửa tất cả bản sao dư thừa ⇒ ghi chậm hơn, phức tạp hơn.
- Data Inconsistency:** Rủi ro không nhất quán nếu một bản sao được cập nhật mà bản khác thì không.
- Storage Cost:** Lưu trữ dữ liệu trùng lặp tốn bộ nhớ hơn.

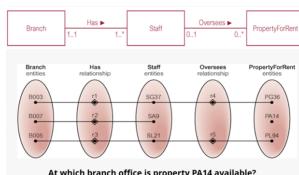
Khi nào dùng: OLAP/DW (đọc nhiều), NoSQL (thiếu join), microservices (tách biệt). Tránh: OLTP cần ACID chặt.

1.3.4 Bẫy Thiết kế CSDL (Design Traps)

- **Bẫy Kết nối trong ER:**



(Fan Trap)



(Chasm Trap)

- **Fan Trap:** Đường dẫn giữa các thực thể mơ hồ do nhiều quan hệ 1:N phân nhánh từ một thực thể. VD: NhanVien → PhongBan → ChiNhanh (không xác định được nhân viên làm ở chi nhánh nào). *Giải pháp:* Thêm quan hệ trực tiếp NhanVien-ChiNhanh.
- **Chasm Trap:** Đường dẫn không tồn tại do tham gia tùy chọn. VD: KhanhHang → TaiSan → ChiNhanh (nếu tài sản chưa niêm yết thì không liên kết được khách hàng với chi nhánh). *Giải pháp:* Đổi sang tham gia bắt buộc hoặc thêm quan hệ trực tiếp.
- **Dị thường Cập nhật (Update Anomalies):** Do thiết kế không chuẩn hóa:
 - *Insertion:* Không thể thêm phòng ban mới nếu chưa có nhân viên.
 - *Deletion:* Xóa nhân viên cuối cùng làm mất thông tin phòng ban.
 - *Modification:* Thay đổi tên phòng ban phải cập nhật nhiều bộ.
- **Bộ giả (Spurious Tuples):** Kết nối các quan hệ phân rã sai (không qua PK/FK hợp lệ) tạo ra bản ghi ảo. *Giải pháp:* Dùng phân rã bảo toàn thông tin (lossless join).
- **Bẫy NULL:** Quá nhiều thuộc tính NULL lảng phí bộ nhớ và gây khó khăn trong truy vấn tổng hợp.
- **Sai lầm phổ biến:** Dùng PK của thực thể này làm thuộc tính của thực thể khác thay vì mô hình hóa quan hệ; gán PK vào thuộc tính của quan hệ; dùng thuộc tính đơn trị khi cần đa trị.
- **Entity Trap (Kiến trúc):** Thiết kế component 1-1 với bảng DB (VD: CustomerManager cho bảng Customer) thay vì theo workflow nghiệp vụ → vi phạm tách biệt dữ liệu trong microservices.

Thực hành tốt: Dùng BCNF/3NF; phân rã bảo toàn thông tin; đảm bảo đường dẫn ER rõ ràng; thiết kế theo hành vi nghiệp vụ, không theo thực thể.

1.4 Nguyên lý Thiết kế Mức Vật lý

- **Kiến trúc Lưu trữ:** Dữ liệu bền vững trên đĩa/SSD trong các **khối (block)** kích thước cố định.
- **Phân tích Tài (Job Mix):** Xác định các quan hệ/tập tin thường truy cập, điều kiện chọn (bằng/khác/khoảng), và tần suất cập nhật so với truy vấn.
- **Cấu trúc Chỉ mục:** Chỉ mục có thứ tự (B+-Trees) và chỉ mục băm; chỉ mục **chính/phân cụm** (quy định thứ tự vật lý; tối đa một trên mỗi tập tin) so với chỉ mục **phụ**.
- **Tối ưu hóa Truy vấn:** Dựa trên chi phí (thống kê) và các quy tắc kinh nghiệm (dãy phép chọn/chiếu xuồng sớm) để chọn kế hoạch hiệu quả.

1.5 Các bước vẽ ERD (Steps to Create ERD)

1. Xác định Thực thể (Entities):

Tìm các **danh từ** (Nouns) quan trọng trong yêu cầu (Vd: Employee, Student). Tránh nhầm lẫn thuộc tính là thực thể.

2. Xác định Mối kết hợp (Relationships):

Tìm các **động từ** (Verbs) kết nối các thực thể (Vd: Works_for, Teaches).

3. Xác định Thuộc tính (Attributes):

Xác định thông tin chi tiết cho mỗi thực thể. Xác định thuộc tính đa trị, dãy xuất, phức hợp.

4. Xác định Khóa chính (Primary Keys):

Chọn thuộc tính định danh duy nhất cho mỗi thực thể và gạch chân nó.

5. Xác định Bản số (Cardinality Ratio):

Phân tích số lượng tham gia: 1:1, 1:N, hay M:N.

6. Xác định Ràng buộc tham gia (Participation):

Có bắt buộc không? (Total - Nét đôi) hay Tùy chọn? (Partial - Nét đơn).

Hỏi: "Thực thể A có thể tồn tại mà không cần B không?"

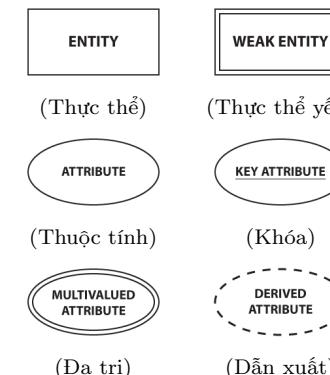
7. Vẽ phác thảo & Tinh chỉnh:

Vẽ sơ đồ, loại bỏ các thuộc tính dư thừa. Chuyển quan hệ M:N thành thực thể liên kết nếu cần thiết.

1.6 Chen Notation: ER & EER

1.6.1 Thực thể & Thuộc tính (Entities & Attributes)

Khi nào dùng: Thực thể cho đối tượng độc lập (NhanVien, SanPham, KhanhHang); Thực thể yếu cho đối tượng phụ thuộc (NguoiPhuThuoc của nhân viên, ChiTietDonHang); **Khóa** là định danh duy nhất (MSNV, CCCD); **Địa chỉ** cho thuộc tính nhiều giá trị (số điện thoại, email); **Dãy xuất** cho giá trị tính toán (tuổi từ ngày sinh, tổng tiền).



1.6.2 Mối kết hợp (Relationships)

Khi nào dùng: Quan hệ thường cho liên kết độc lập (NhanVien làm việc cho PhongBan, KhanhHang mua SanPham); Quan hệ xác định

khi thực thể yếu phụ thuộc vào thực thể chủ (NguoiPhuThuoc thuộc về NhanVien với khóa bộ phận là tên người phụ thuộc).



RELATIONSHIP



WEAK RELATIONSHIP

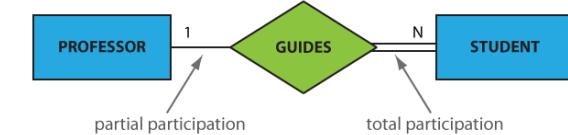
Bao gồm: Quan hệ (Hình thoi), Quan hệ xác định (Thoát đội)

1.6.3 Ràng buộc (Constraints)

Khi nào dùng: Xác định quy tắc nghiệp vụ giữa các thực thể.

Bản số (Cardinality): 1:1 (NhanVien quản lý PhongBan - mỗi phòng có 1 trưởng), 1:N (PhongBan có NhanVien - nhiều nhân viên/phòng), M:N (NhanVien tham gia DUAN - nhiều-nhiều).

Tham gia (Participation):

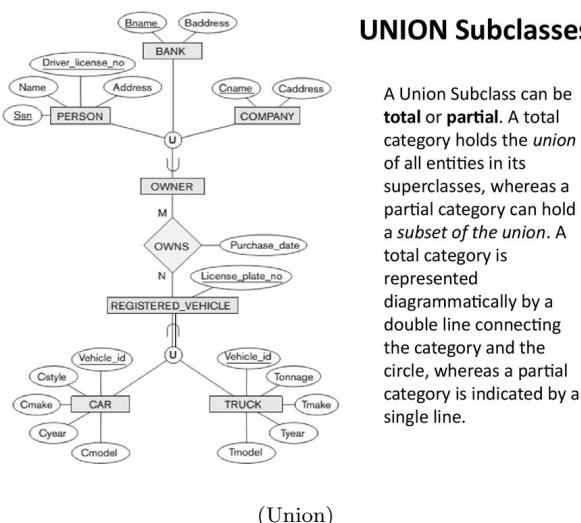
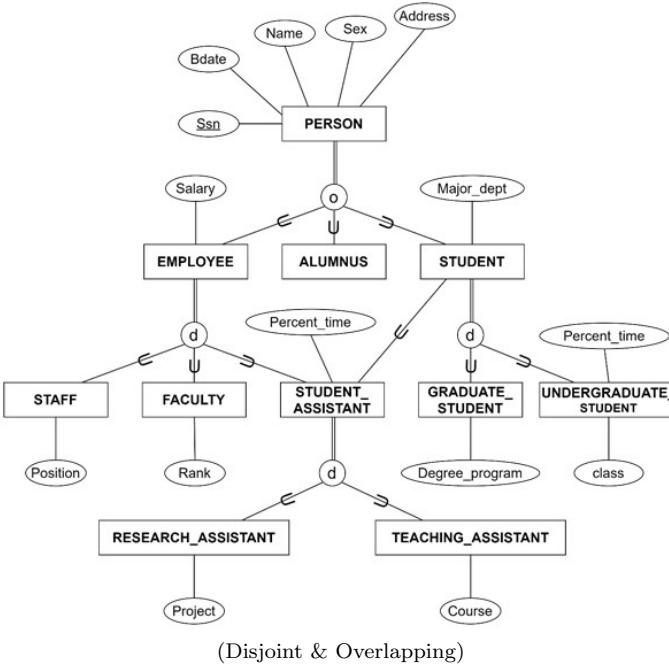


(Partial Participation (Từng phần): Nét đơn
Total Participation (Toàn phần): Nét đôi)

Min-Max (min, max): Ghi cặp số trên cạnh. Ví dụ: NhanVien (1,1) làm việc cho (0, N) PhongBan nghĩa là mỗi nhân viên bắt buộc làm việc cho đúng 1 phòng ban, mỗi phòng ban có thể có 0 đến nhiều nhân viên. (0, 1): tùy chọn, tối đa 1; (1, N): bắt buộc, có thể nhiều. (An extension of Participation và Cardinality)

1.6.4 EER Chuyên biệt hóa & Tổng quát hóa

Khi nào dùng: Disjoint khi lớp con không chồng lấp (NhanVien là KỶ SU hoặc QUẢN LÝ, không đồng thời); Overlapping khi có thể thuộc nhiều lớp (NGƯỜI là SINH VIÊN và/hoặc NHÂN VIÊN); Union khi lớp con kế thừa từ nhiều lớp cha (CHỦ SỞ HỮU có thể là NGƯỜI hoặc CÔNG TY hoặc NGÂN HÀNG). Total khi mọi thực thể cha phải thuộc ít nhất 1 lớp con; Partial khi không bắt buộc.



Ký hiệu: Hình tròn (d: disjoint, o: overlapping, U: union),
Nét dài (Total), Nét đơn (Partial).

2 Chỉ mục & Tối ưu hóa Truy vấn

2.1 Cơ bản về Chỉ mục

2.1.1 Chỉ mục là gì?

- Mục đích:** Tăng tốc độ truy xuất dữ liệu bằng cách tạo đường dẫn phụ trợ đến các bản ghi.
- Khóa tìm kiếm:** Các thuộc tính dùng để tìm bản ghi; không nhất thiết là khóa chính; có thể là khóa phức hợp (nhiều cột).
- Đánh đổi:** Đọc nhanh hơn so với ghi chậm hơn (chi phí bảo trì chỉ mục khi INSERT/UPDATE/DELETE).

2.1.2 Phân loại Chỉ mục

- Theo Cấu trúc:

- Có thứ tự (Ordered):** Các mục được sắp xếp (vd: B+-tree); hỗ trợ truy vấn khoảng.
- Băm (Hash):** Khóa được băm vào bucket; chỉ nhanh khi tìm kiếm chính xác (dấu bằng).

- Theo Mật độ:

- Đặc (Dense):** Một mục chỉ mục cho mỗi giá trị khóa tìm kiếm riêng biệt.
- Thưa (Sparse):** Một mục chỉ mục cho mỗi khối (hoặc mỗi giá trị phân cụm); nhỏ hơn, chi phí bảo trì thấp hơn.

- Theo Thứ tự Vật lý:

- Chỉ mục Chính/Phân cụm:** Khóa tìm kiếm quyết định thứ tự vật lý của tập tin (tối đa một per bảng); thường là chỉ mục thưa.
- Chỉ mục Phụ:** Đường dẫn truy cập thay thế độc lập với thứ tự vật lý; thường là chỉ mục đặc hoặc dùng gián tiếp cho các khóa không duy nhất.

2.2 B-Trees & B+-Trees

2.2.1 Tại sao dùng B+-Trees?

- Mục tiêu:** Giảm thiểu I/O đĩa tốn kém bằng cách giữ chiều cao cây thấp thông qua hệ số rẽ nhánh (fan-out) cao.
- Cấu trúc:** Cây cân bằng; nút trong chứa khóa + con trỏ con; nút lá chứa khóa + con trỏ dữ liệu (hoặc ID bản ghi) + liên kết đến lá kế tiếp.
- Thao tác:** Tìm/thêm/xóa trong $O(\log_{f_o} n)$ lần truy cập khối; tách/gộp nút để duy trì cân bằng.
- B+ so với B-Tree:** B+- chỉ lưu con trỏ dữ liệu ở lá \Rightarrow hệ số rẽ nhánh nút trong cao hơn, quét tuân tự hiệu quả qua chuỗi liên kết lá.

2.2.2 Ví dụ Tính Dung lượng

Tham số: Kích thước khối $B = 512$ bytes; kích thước khóa $V = 9$ bytes; con trỏ dữ liệu $Pr = 7$ bytes; con trỏ cây $P = 6$ bytes.

Bước 1: Tính Bậc (Số con trỏ tối đa mỗi nút)

- Nút trong B-Tree:** Chứa p con trỏ cây + $(p - 1)$ khóa + $(p - 1)$ con trỏ dữ liệu.

$$(p \times 6) + ((p - 1) \times (7 + 9)) \leq 512$$

$$6p + 16p - 16 \leq 512 \Rightarrow 22p \leq 528 \Rightarrow p = 23$$

- Nút trong B+-Tree:** Chứa p con trỏ cây + $(p - 1)$ khóa (không có con trỏ dữ liệu).

$$(p \times 6) + ((p - 1) \times 9) \leq 512$$

$$6p + 9p - 9 \leq 512 \Rightarrow 15p \leq 521 \Rightarrow p = 34$$

- Nút lá B+-Tree:** Chứa p_{leaf} cặp khóa/con trỏ dữ liệu + 1 con trỏ kế tiếp.

$$(p_{leaf} \times (7 + 9)) + 6 \leq 512$$

$$16 \times p_{leaf} \leq 506 \Rightarrow p_{leaf} = 31$$

Bước 2: Ước lượng Tổng dung lượng (Đầy 69%)

- B-Tree (3 mức):** Hệ số rẽ nhánh trung bình $f_o = 23 \times 0.69 \approx 16$.

$$- \text{Mức 0 (gốc): } 15 \text{ mục, } 16 \text{ con trỏ}$$

$$- \text{Mức 1: } 16 \times 15 = 240 \text{ mục}$$

$$- \text{Mức 2: } 256 \times 15 = 3,840 \text{ mục}$$

$$- \text{Tổng: } 15 + 240 + 3,840 \approx 4,095 \text{ mục}$$

- B+-Tree (3 mức):** Nút trong $f_o = 34 \times 0.69 \approx 23$; Dung lượng lá = $31 \times 0.69 \approx 21$.

$$- \text{Mức 0: } 22 \text{ mục, } 23 \text{ con trỏ}$$

$$- \text{Mức 1: } 23 \times 22 = 506 \text{ mục, } 529 \text{ con trỏ}$$

$$- \text{Mức lá: } 12,167 \times 21 \approx 255,507 \text{ con trỏ dữ liệu}$$

- Nhận xét:** B+- chứa được $\sim 4 \times$ số mục ở cùng chiều cao nhờ nút trong nhẹ hơn.

2.2.3 Các loại Chỉ mục Nâng cao

- Chỉ mục Phức hợp:** Khóa nhiều cột (vd: (City, LastName)); hỗ trợ truy vấn tiền tố trái nhất (City), (City, LastName); sắp xếp cột theo độ chọn lọc.

- Chỉ mục Dựa trên Hàm:** Chỉ mục trên biểu thức (vd: LOWER(email)); truy vấn phải dùng đúng hàm đó mới tận dụng được.

2.3 Chỉ mục Băm & Bitmap

2.3.1 Chỉ mục Băm (Hash Indexes)

- Sử dụng:** Tìm kiếm chính xác cực nhanh (truy vấn điểm); không hỗ trợ khoảng.

- Xử lý đúng đắn:** Dùng danh sách liên kết (chaining) với bucket tràn.

- Biến thể động:** Băm mở rộng/tuyến tính (Extendible/Linear hashing) tăng trưởng dần mà không cần xây lại toàn bộ.

2.3.2 Chỉ mục Bitmap

Tối ưu cho thuộc tính có độ chọn lọc thấp (*ít giá trị riêng biệt*).

- Cấu trúc:** Dánh số thứ tự bản ghi (0, 1, 2, ...); mỗi giá trị riêng biệt có một bitmap; bit $i = 1$ nếu bản ghi i có giá trị đó.

- Ví dụ (bảng 5 dòng):

– gender='m': 10010 gender='f': 01101

– income='L1': 11000 income='L2': 00100

- Truy vấn:** gender='f' AND income='L2'
01101 AND 00100 = 00100 \Rightarrow bản ghi 2

- **Ưu điểm:** Gọn nhẹ (1 triệu dòng = 125 KB mỗi bitmap); thao tác bitwise nhanh; hiệu quả cho bộ lọc nhiều điều kiện; hỗ trợ COUNT qua đếm bit.

2.4 Tối ưu hóa Truy vấn & Phân tích Chi phí

2.4.1 Ký hiệu Chi phí (Cost Function Notation)

Ước tính tài nguyên (disk I/O) để xử lý truy vấn: số block transfers (b) và số disk seeks (S).

- b (hoặc b_r): Tổng số blocks trong file.
- r (hoặc n_r): Tổng số records (tuples) trong relation.
- s (hoặc s_A): Selection cardinality - số records trung bình thỏa điều kiện.
- bfr : Blocking factor (số records trên mỗi block).
- x (hoặc h_i): Số levels trong index (chiều cao B+-tree).
- t_S : Average time cho disk seek.
- t_T : Average time để transfer một block.
- **Công thức tổng:** Cost = $b \times t_T + S \times t_S$ (transfer time + seek time).

2.4.2 Hàm Chi phí cho Phép Chọn (SELECT Cost Functions)

Phép chọn (σ) có thể dùng quét file hoặc truy cập index. Chi phí chưa tính ghi kết quả.

S1: Linear Search (File Scan)

Quét mọi block và test tất cả records. Áp dụng cho mọi cấu trúc file.

- **Trường hợp chung:** $C_{S1a} = b$ blocks.
Time: $t_S + b \times t_T$ (1 seek ban đầu + transfer all blocks).
- **Equality on Key:** $C_{S1b} = b/2$ blocks (trung bình, dùng khi tìm thấy).
Time: $t_S + (b/2) \times t_T$.

S2: Binary Search

File đã sắp xếp theo search attribute. $C_{S2} = \lceil \log_2 b \rceil + \lceil s/bfr \rceil - 1$ blocks.

Giải thích: $\log_2 b$ để locate first block + $\lceil s/bfr \rceil$ để retrieve matching blocks contiguous.

S3a: Primary Index (Equality on Unique Key)

File được sắp xếp vật lý theo search key. $C_{S3a} = x + 1$ blocks.

Time: $(h_i + 1) \times (t_T + t_S)$ (traverse index tree + 1 data block access).
VD: B+-tree height 3 → 3 index blocks + 1 data block = 4 I/Os.

S3b: Hash Index (Equality)

$C_{S3b} = 1$ (static/linear hashing) hoặc 2 (extendible hashing).

S5: Clustering Index (Equality on Non-Unique Key)

File sắp xếp, matching records lùu contiguous. $C_{S5} = x + \lceil s/bfr \rceil$ blocks.

Time: $h_i \times (t_T + t_S) + t_S + b_{matches} \times t_T$ (index + 1 seek to data +

transfer).

VD: Tìm Dno=5 với 1000 matching records, bfr=10 → $x + 100$ blocks.

S6a: Secondary Index (Equality on Non-Unique Key)

File khong sắp xếp theo attribute này → matching records phân tán. $C_{S6a} = x + s + 1$ blocks (worst case: mỗi record cần 1 random I/O). Time: $(h_i + n) \times (t_T + t_S)$ với n = số records.

Nguy cơ: Nếu s lớn, secondary index đắt hơn linear scan do nhiều random seeks.

S6b: Secondary Index (Range Query)

$C_{S6b} = x + (b_{I1}/2) + (r/2)$ (giả sử half file thỏa điều kiện).

Performance Implication: Nếu số records lớn, optimizer thích linear scan hơn secondary index.

2.4.3 Selections với Comparisons (Range Queries)

- **Clustering Index:** Hiệu quả - records contiguous.
Cost: $x + b_{range}$ (find first + sequential scan).
- **Secondary Index:** Đắt - index leaves scan + random I/O cho mỗi record.
Nếu range lớn: Prefer linear scan.

2.4.4 Complex Selections

- **Conjunctive (AND):** Chọn single most restrictive access path (smallest selectivity) → check conditions còn lại trong RAM.
Hoặc: Dùng nhiều secondary indices → retrieve pointers → compute intersection.
- **Disjunctive (OR):** Nếu 1 condition cần linear scan → toàn bộ query cần linear scan.
Nếu indices cho tất cả: Retrieve pointers → compute union.

2.4.5 Bảng Tóm tắt Chi phí SELECT

Algorithm	Mô tả	Block Accesses
S1	Linear Search	b
S2	Binary Search	$\lceil \log_2 b \rceil + \lceil s/bfr \rceil - 1$
S3a	Primary Index (Unique)	$x + 1$
S5	Clustering Index (Equality)	$x + \lceil s/bfr \rceil$
S6a	Secondary Index (Equality)	$x + s + 1$
S6b	Secondary Index (Range)	$x + (b_{I1}/2) + (r/2)$

Lưu ý: Chi phí thời gian = $b \times t_T + S \times t_S$ (tách biệt transfer và seek).

2.4.6 Ví dụ: Phép chọn trên EMPLOYEE

Kịch bản: Quan hệ EMPLOYEE có thông số:

- $r_E = 10,000$ records, $b_E = 2,000$ blocks, $bfr_E = 5$ records/block.
- **Available Access Paths:**
 - Ssn (Secondary, Unique Key): $x = 4, s = 1$.
 - Dno (Secondary, Non-Unique): $x = 2, s = 80$ (assumption: 125 depts, so $10,000/125 = 80$).

- $Salary$ (Clustering, Non-Unique): $x = 3, s = 20$.

- Sex (Secondary, Non-Unique): $x = 1, s = 5,000$ (assumption: 2 values, so $10,000/2 = 5,000$).

Example 1: Equality on Unique Key

Query: $\sigma_{Ssn='123456789'}(\text{EMPLOYEE})$.

- **S1b (Linear Search):** $C = b_E/2 = 2,000/2 = 1,000$ blocks.

Explanation: Scan half file on average until record found.

- **S6a (Secondary Index on Key):** $C = x + 1 = 4 + 1 = 5$ blocks.

Explanation: Traverse 4-level index + retrieve 1 data block.

- **Decision:** Use S6a (5 blocks ≪ 1,000 blocks). Speedup: 200x.

Example 2: Equality on Non-Unique Key

Query: $\sigma_{Dno=5}(\text{EMPLOYEE})$.

- **S1a (Linear Search):** $C = b_E = 2,000$ blocks.

- **S6a (Secondary Index):** $C = x + s = 2 + 80 = 82$ blocks.

Explanation: 2 index blocks + 80 random I/Os (records scattered across blocks).

- **S5 (If Clustering Index existed):** $C = x + \lceil s/bfr \rceil = 3 + \lceil 80/5 \rceil = 3 + 16 = 19$ blocks.

Explanation: 3 index blocks + 16 contiguous data blocks (records stored together).

- **Decision:** Use S6a (82 blocks ≪ 2,000). Note: Clustering index would be 4.3x better (19 vs 82).

Example 3: Conjunctive Selection (Multiple Conditions)

Query: $\sigma_{Dno=5 \wedge Salary > 30,000 \wedge Sex='F'}(\text{EMPLOYEE})$.

Strategy: Choose most selective access path to retrieve candidate set, then filter remaining conditions in memory.

- **Via Dno (S6a):** $C = x + s = 2 + 80 = 82$ blocks. Retrieves 80 records.

- **Via Salary Range (Clustering):** $C \approx x + (b_E/2) = 3 + 1,000 = 1,003$ blocks.

Assumption: Half employees earn > 30,000, need to scan half file.

- **Via Sex (S6a):** $C = x + s = 1 + 5,000 = 5,001$ blocks. 5,000 random I/Os (very expensive).

- **Linear Scan (S1a):** $C = 2,000$ blocks.

- **Decision:** Use Dno index (82 blocks) to retrieve 80 candidate records → filter Salary > 30,000 and Sex='F' in memory.

Final result size estimated: $80 \times 0.5 \times 0.5 = 20$ records (assuming 50% pass each filter).

Example 4: Secondary Index vs Linear Scan Trade-off

Query: $\sigma_{Sex='F'}(\text{EMPLOYEE})$ (highly non-selective).

- **S6a (Secondary Index):** $C = x + s = 1 + 5,000 = 5,001$ blocks.

Problem: 5,000 random seeks → extremely expensive. At $t_S = 10\text{ms}$, $t_T = 1\text{ms}$: $(1 + 5,000) \times 11 = 55,011\text{ms} \approx 55$ seconds.

- **S1a (Linear Scan):** $C = 2,000$ blocks.

Time: $1 \times 10 + 2,000 \times 1 = 2,010\text{ms} \approx 2$ seconds.

- **Decision:** Use linear scan (2,000 blocks ≪ 5,001 blocks). 27x faster despite having index!

Key Insight: Secondary index only beneficial when selectivity is high ($s \ll b$).

2.4.7 Cost Functions for JOIN Operations

Notation

- b_r, b_s : Number of blocks (số khối) in relations r and s .
- n_r, n_s (or $|R|, |S|$): Number of tuples (số bộ) in relations r and s .
- js : Join selectivity (độ chọn lọc kết nối) - fraction of tuple pairs matching. For equi-join: $js \approx 1/\max(NDV(A), NDV(B))$.
- jc : Join cardinality (lực lượng kết nối) - number of result tuples = $js \times |R| \times |S|$.
- bfr_{result} : Blocking factor (hệ số khối) of result relation.
- n_B (or M): Number of available memory buffer blocks (số khối bộ đệm).
- hi (or x): Height of index tree (chiều cao B+-tree).
- s_B : Selection cardinality - average matching records in index lookup.

J1: Nested-Loop Join Variants

- Basic Nested-Loop:** For each record in outer relation (quan hệ ngoài) r , scan entire inner relation (quan hệ trong) s .
Cost: $b_r + n_r \times b_s$ block transfers + $(n_r + b_r)$ seeks.
Problem: Cực kỳ đắt nếu n_r lớn.
- Block Nested-Loop:** Process block-by-block thay vì row-by-row.
Cost (Worst): $b_r + (b_r \times b_s)$ transfers + $2 \times b_r$ seeks.
Cost (With n_B buffers): $b_r + \lceil b_r / (n_B - 2) \rceil \times b_s$ transfers.
Strategy: Dùng $n_B - 2$ buffers cho outer, 1 cho inner, 1 cho output. Dùng quan hệ nhỏ hơn làm outer.

J2: Indexed Nested-Loop Join

Index tồn tại trên join attribute của inner relation s . Thay file scans bằng index lookups (tra cứu chỉ mục).

- General Formula:** $b_r + (n_r \times \text{cost of selection on } s)$.
Explanation: Đọc outer r (b_r) + thực hiện n_r index lookups trên s .
- Secondary Index (Chỉ mục phụ):** $b_r + n_r \times (x + s_B + 1)$ blocks.
 s_B : Average matching records. Mỗi record cần random I/O nếu non-unique.
- Primary/Clustering Index (Chỉ mục chính/phân cụm):** $b_r + n_r \times (x + 1)$ blocks.
Assumption: Unique keys hoặc matching records lưu contiguous (liền kề).

J3: Sort-Merge Join

Cả hai quan hệ phải được sắp xếp theo join attributes. Hiệu quả cao nếu đã pre-sorted.

- If already sorted:** $b_r + b_s$ block transfers (đọc cả hai một lần).
- If not sorted:** Cost(Sort r) + Cost(Sort s) + $b_r + b_s$.
Sort Cost: Thường $b \times (2 \lceil \log_{n_B-1}(b/n_B) \rceil + 1)$ cho external merge sort.
- Best Use Case:** Dữ liệu đã sorted, hoặc khi kết quả cần sorted.

J4: Hash Join

Partition both relations by hash of join key. Matching tuples fall into same partition.

- Standard Hash Join:** $3(b_r + b_s)$ block transfers.
Passes: 1. Read/write partitions for r . 2. Read/write partitions for s . 3. Join partitions.
Requirement: Partitions of smaller relation must fit in memory.

- Hybrid Hash Join:** $b_r + b_s$ (best case).

Optimization: If entire build relation fits in memory, skip disk writes for some partitions.

Writing the Result

Chi phí trên chỉ ước tính *processing*. Phải cộng thêm write cost (chi phí ghi) để có tổng chi phí.

- Write Cost:** jc/bfr_{result} blocks.
- Total Cost:** Processing Cost + $(js \times |R| \times |S|)/bfr_{result}$.

2.4.8 Example: EMPLOYEE \bowtie DEPARTMENT

Scenario: EMPLOYEE ($|E| = 10,000, b_E = 2,000$) $\bowtie_{Dno=Dnumber}$ DEPARTMENT ($|D| = 125, b_D = 13$).

Indices: Secondary on $E.Dno$ ($x = 2, s = 80$). Primary on $D.Dnumber$ ($x = 1$).

Parameters: $js = 1/125, jc = 10,000, bfr_{result} = 4 \rightarrow$ write cost = 2,500 blocks. $n_B = 3$ buffers.

- J1: Block Nested-Loop (D outer):**

$$C = 13 + \lceil 13/1 \rceil \times 2,000 + 2,500 = 13 + 26,000 + 2,500 = \mathbf{28,513} \text{ blocks.}$$

Explanation: Đọc D (13) + scan E 13 lần (26,000) + ghi kết quả (2,500).

- J2a: Indexed Nested-Loop (D outer \rightarrow E inner):**

$$C = 13 + 125 \times (2 + 80) + 2,500 = 13 + 10,250 + 2,500 = \mathbf{12,763} \text{ blocks.}$$

Explanation: Đọc D (13) + 125 lookups trên E (mỗi lần: 2 index + 80 scattered records) + ghi.

- J2b: Indexed Nested-Loop (E outer \rightarrow D inner):**

$$C = 2,000 + 10,000 \times (1+1) + 2,500 = 2,000 + 20,000 + 2,500 = \mathbf{24,500} \text{ blocks.}$$

Explanation: Đọc E (2,000) + 10,000 lookups trên D (mỗi lần: 1 index + 1 data block) + ghi.

- J3: Sort-Merge (if already sorted):**

$$C = 2,000 + 13 + 2,500 = \mathbf{4,513} \text{ blocks.}$$

Best case: Chỉ đọc cả hai quan hệ một lần + ghi kết quả.

If not sorted: Cộng thêm sorting cost (thường $\approx 2b \log b$ per relation).

- J4: Hash Join:**

$$C = 3 \times (2,000 + 13) + 2,500 = 6,039 + 2,500 = \mathbf{8,539} \text{ blocks.}$$

Explanation: 3 passes qua cả hai quan hệ (partition + build + probe) + ghi kết quả.

Decision Ranking

J3 (if sorted) < J4 < J2a < J2b < J1.

- Best:** Sort-Merge nếu đã sorted (4,513).
- Good:** Hash Join nếu chưa sorted (8,539) - không có sorting overhead.
- Moderate:** Indexed Nested-Loop với small outer (12,763).
- Avoid:** Block Nested-Loop (28,513) - chỉ dùng khi không có index/memory.

2.5 Quy trình Tối ưu hóa Truy vấn

2.5.1 Tổng quan Quy trình

- Phân tích & Kiểm tra (Parse):** Kiểm tra cú pháp, tuân thủ lược đồ.
- Dịch (Translate):** Chuyển SQL sang đại số quan hệ (cây truy vấn).
- Tối ưu hóa Kinh nghiệm (Heuristic):** Áp dụng các quy tắc biến đổi.
- Tối ưu hóa Dựa trên Chi phí:** Liệt kê các kế hoạch, ước tính chi phí, chọn chi phí thấp nhất.
- Thực thi:** Hiện thực hóa kết quả trung gian hoặc pipeline kết quả.

2.5.2 Quy tắc Heuristic

- Đẩy phép chọn (σ) xuống:** Lọc sớm để giảm kích thước trung gian.
- Đẩy phép chiểu (Π) xuống:** Giảm chiều rộng bộ dữ liệu sớm.
- Thay thế $\sigma + \text{Tích đề-các} (\times)$ bằng Kết (\bowtie):** Tránh tích đề-các tốn kém.
- Sắp xếp lại thứ tự kết:** Dùng tính giao hoán/kết hợp để tìm kết quả trung gian có lực lượng thấp.
- Cây nghiêng trái (Left-deep trees):** Con phải luôn là bảng cơ sở \Rightarrow cho phép tìm kiếm chỉ mục, giảm không gian tìm kiếm.

2.5.3 Quyết định Dựa trên Chi phí

- Chọn đường dẫn truy cập:** So sánh quét toàn bộ, chỉ mục phân cụm, chỉ mục phụ, bitmap cho mỗi vị trí.
- Thuật toán kết nối:** Nested-loop (index/block), hash join, sort-merge dựa trên số lượng bộ và bộ nhớ.
- Thứ tự kết nối:** Ước lượng độ chọn lọc kết nối $js \approx 1/\max(NDV(A), NDV(B))$; số lượng $jc = js \times |R| \times |S|$.

2.6 Kỹ thuật Chuyên biệt

2.6.1 Cấu trúc Tối ưu Ghi

- LSM-Tree:** Bộ đệm trong RAM (memtable) + các mức đĩa đã sắp xếp; ghi tuần tự; nén định kỳ; Bloom filter để bỏ qua các mức.
- Buffer Tree:** Biến thể B-tree với bộ đệm ghi tại mỗi nút; gom nhóm các thay đổi xuống cây; độ trễ đọc tốt hơn LSM.

2.6.2 Truy cập Không gian & Đa khóa

- Đa khóa:** Chỉ mục phức hợp cho truy vấn tiền tố; chỉ mục bao phủ (covering index) tránh tra cứu bảng.
- Không gian:** R-tree (hình chữ nhật bao), kd-tree, quadtree cho dữ liệu địa lý/khoảng; hỗ trợ truy vấn vùng + lân cận gần nhất.

3 Big Data & Data Engineering

3.1 Đặc trưng Big Data (The Vs)

- 5V Cốt lõi (Core):**
 - Volume:** Dung lượng khổng lồ (TB, PB, ZB).

- Velocity:** Tốc độ sinh ra & xử lý (Batch → Streaming).
- Variety:** Diverse định dạng (Structured, JSON, Video, Log).
- Veracity:** Độ tin cậy, tính xác thực (Messy/Noisy data).
- Value:** Giá trị chuyển hóa thành lợi ích kinh doanh.

- Các V Mở rộng (Extended):

- Variability:** Tính biến thiên (Ý nghĩa dữ liệu thay đổi theo ngữ cảnh/thời gian).
- Validity:** Tính hợp lệ (Dữ liệu có đúng định dạng/chuẩn để dùng khống).
- Vulnerability:** Tính bảo mật (Đã bị tấn công/rò rỉ).
- Volatility:** Độ bay hơi (Thời gian lưu trữ trước khi xóa/lưu trữ lâu dài).
- Visualization:** Khả năng trực quan hóa (Để con người hiểu được).

3.2 Paradigm: Batch vs Streaming

Hai mô hình xử lý dữ liệu cơ bản trong Big Data.

	Batch Processing	Streaming Processing
Đặc trưng	Xử lý dữ liệu tĩnh, lượng lớn theo lô.	Xử lý dữ liệu động, liên tục theo thời gian thực.
Độ trễ	Cao (minutes - hours).	Thấp (seconds - milliseconds).
Công cụ	Hadoop MapReduce, Apache Spark (Batch mode).	Apache Flink, Spark Streaming, Kafka Streams.
Use Case	ETL, báo cáo cuối ngày, ML training.	Real-time analytics, fraud detection, monitoring.
Ưu điểm	Xử lý hiệu quả khối lượng lớn, đơn giản.	Phản hồi nhanh, phát hiện sự kiện ngay lập tức.
Nhược điểm	Không real-time, lãng phí khi data nhỏ.	Phức tạp, khó debug, cần xử lý out-of-order.

Kiến trúc Lai (Hybrid):

- Lambda Architecture:** Batch layer (chính xác) + Speed layer (real-time) + Serving layer. Phức tạp, duy trì 2 code base.
- Kappa Architecture:** Chỉ dùng Streaming (đơn giản hóa). Mọi dữ liệu qua stream processor, replay từ Kafka khi cần.

3.3 Phân mảng & Sao chép (Partitioning & Replication)

3.3.1 Sao chép (Replication)

Mục đích: High Availability (HA) và giảm độ trễ đọc.

- Single-Leader (Master-Slave):** Mọi ghi vào Leader, Leader chép sang Followers. Đã nhất quán, nhưng Leader là nút cổ chai.
- Multi-Leader:** Nhiều node chấp nhận ghi. Tốt cho đa trung tâm dữ liệu, nhưng khó xử lý xung đột.
- Leaderless (Dynamo-style):** Ghi/Đọc gửi tới nhiều node. Dùng cơ chế Quorum để xác nhận.

$w + r > n$ (Write nodes + Read nodes > Total replicas) → Đảm bảo đọc thấy dữ liệu mới nhất.

3.3.2 Phân mảng (Partitioning/Sharding)

Mục đích: Scalability (Mở rộng dung lượng/băng thông).

- Key Range Partitioning:** Chia theo khoảng khóa (A-C, D-F).
 - **Ưu:** Query theo khoảng (Range scan) hiệu quả.
 - **Nhược:** Dễ bị Hotspot (nếu user dồn vào vần A).
- Hash Partitioning:** Băm khóa để chia đều ($hash(key) \% N$).
 - **Ưu:** Phân phối đều, tránh Hotspot.
 - **Nhược:** Mất khả năng Range Query (phải quét tất cả).

3.4 Định dạng Lưu trữ (File Formats)

Lựa chọn định dạng ảnh hưởng trực tiếp đến hiệu năng đọc/ghi.

3.4.1 Row-based vs. Column-based

- Row-oriented (CSV, Avro):

- Lưu trữ tuần tự từng dòng.
- Ưu điểm: Ghi nhanh (append), tốt khi truy xuất toàn bộ thông tin của 1 entity (OLTP).
- Nhược điểm: Chậm khi tính toán tổng hợp (SUM, AVG) vì phải đọc cả dữ liệu không cần thiết.

- Column-oriented (Parquet, ORC):

- Lưu trữ riêng biệt từng cột.
- Ưu điểm: Nén cực tốt (do dữ liệu cùng kiểu), tối ưu cho OLAP (chỉ đọc cột cần thiết).
- Nhược điểm: Ghi chậm, update tốn kém.

3.4.2 So sánh Avro, Parquet, ORC

Đặc điểm	Avro	Parquet
Mô hình	Row-based	Column-based
Schema	JSON (lưu trong file)	Binary (footer)
Tối ưu cho	Ghi nhiều (Write heavy)	Đọc nhiều (Read heavy)
Schema Evo	Rất tốt (Thêm/bớt field)	Hạn chế
Ecosystem	Kafka, Hadoop	Spark, Impala, Presto

3.5 Hệ sinh thái Hadoop (Hadoop Ecosystem)

Khung làm việc mã nguồn mở cho lưu trữ và xử lý phân tán.

1. HDFS (Storage):

Hệ thống tệp phân tán.

- NameNode:** Quản lý metadata (vị trí block).
- DataNode:** Lưu trữ block dữ liệu thực tế.
- Cơ chế:** Chia file thành block (128MB), replicate (x3) để chịu lỗi.

2. YARN (Resource Management):

"Hệ điều hành" của cluster.

- Phân phối tài nguyên (RAM, CPU) cho các ứng dụng.

- Cho phép chạy nhiều engine (Spark, MapReduce) trên cùng 1 cụm.

3. MapReduce (Processing):

Mô hình xử lý Batch - chia để trị trên cụm phân tán.

- Map:** Chia nhỏ & Gán nhãn. Input → Split → **<Key, Value>**.
- Shuffle:** Xáo trộn & Gom nhóm. Chuyển dữ liệu qua mạng, gom cùng Key.
- Reduce:** Tổng hợp. Xử lý danh sách Value của mỗi Key.

3.6 Công nghệ NoSQL (Storage Tech)

Các mô hình NoSQL cho use case khác nhau.

MongoDB (Document Store)

- Mô hình:** Schema-on-read, lưu trữ JSON/BSON documents. Collections thay vì tables.
- Ưu điểm:** Linh hoạt schema (mỗi doc có cấu trúc khác nhau), dễ scale horizontal (sharding), query mạnh (aggregation pipeline).
- Architecture:** Replica Sets (HA), Sharding (scale-out), WiredTiger storage engine.
- Use Case:** CMS, Mobile apps, Catalog, Real-time analytics. VD: Forbes, eBay, Uber.
- Trade-offs:** Không ACID cross-document (trước v4.0), chiếm RAM nhiều.

Redis (Key-Value Store)

- Mô hình:** In-memory key-value, data structures (String, Hash, List, Set, Sorted Set).
- Ưu điểm:** Cực nhanh (< 1ms latency), atomic operations, support pub/sub, Lua scripting.
- Persistence:** RDB (snapshot) hoặc AOF (append-only log). Có thể dùng cả 2.
- Use Case:** Caching (session, query result), Message Queue (Celery), Leaderboard, Rate limiting. VD: Twitter, GitHub, Stack Overflow.
- Trade-offs:** Giới hạn RAM, single-threaded (1 core), không có query phức tạp.

Cassandra (Wide-Column Store)

- Mô hình:** Wide-column, mỗi row có thể có số cột khác nhau. Organize theo Column Family.
- Ưu điểm:** Ghi cực nhanh (LSM Tree), linear scalability, masterless (P2P), multi-datacenter replication.
- Architecture:** Consistent hashing (ring), tunable consistency (quorum), compaction strategies.
- Use Case:** Time-series data, IoT sensor logs, Event logging, Messaging. VD: Netflix, Apple, Instagram.
- Trade-offs:** Đọc chậm hơn (nhiều SSTable), không join, modeling phức tạp (query-first design).

Neo4j (Graph Database)

- Mô hình:** Nodes (entities) + Relationships (edges) + Properties. Native graph storage.
- Ưu điểm:** Traversal cực nhanh (follow pointers), query trực quan (Cypher), ACID transactions.

- Architecture:** Index-free adjacency (mỗi node chứa pointer đến neighbors).
- Use Case:** Social networks, Recommendation engines, Fraud detection, Knowledge graphs. *VD: LinkedIn, Walmart, eBay.*
- Trade-offs:** Scale khó hơn NoSQL khác, không tốt cho bulk data processing.

3.7 Xử lý Batch (Batch Processing)

3.7.1 Apache Spark (Unified Analytics Engine)

Thay thế MapReduce với tốc độ cao hơn 100x (*in-memory*).

- Core Concept:** RDD (Resilient Distributed Dataset) - immutable, partitioned, parallel.
 - Transformations:* Lazy (map, filter, join) - tạo DAG.
 - Actions:* Eager (collect, count, save) - trigger execution.
- Components:**
 - Spark SQL:* Query structured data (DataFrame/Dataset API).
 - Spark Streaming:* Micro-batch streaming (DStream).
 - Mlib:* Machine learning library (classification, clustering, etc.).
 - GraphX:* Graph processing (PageRank, connected components).

- Ưu điểm:** In-memory caching, lazy evaluation, DAG optimization, unified API (batch + streaming).
- Nhược điểm:** Tốn RAM, không true streaming (micro-batch), overhead cho job nhỏ.
- Use Case:** ETL, ML training, interactive analytics, log processing. *VD: Netflix, Uber, Airbnb.*

3.8 Xử lý Luồng (Streaming Processing)

Xử lý dữ liệu liên tục, độ trễ thấp (*Real-time*).

3.8.1 Các chiến lược xử lý (Strategies)

- Thời gian (Time Domain):**
 - Event Time:* Thời gian sự kiện xảy ra (quan trọng nhất).
 - Processing Time:* Thời gian hệ thống nhận được dữ liệu.
 - Watermark:* Cơ chế xử lý độ trễ (data đến muộn) trong Event Time.
- Cửa sổ (Windowing):**
 - Tumbling:* Cố định, không chồng (vd: mỗi 5p).
 - Hopping/Sliding:* Có chồng lấp (vd: 5p, trượt mỗi 1p).
 - Session:* Dựa trên hoạt động người dùng (hết timeout thì đóng).
- Đảm bảo (Guarantees):**
 - At-most-once:* Gửi 1 lần, chấp nhận mất (vd: Log).
 - At-least-once:* Không mất, chấp nhận trùng lặp.
 - Exactly-once:* Chính xác 1 lần (Khó nhất, cần Flink/Kafka).

3.9 Giao thức IoT (IoT Protocols)

3.9.1 MQTT vs HTTP

HTTP (HyperText Transfer)	MQTT (Message Queuing)
Mô hình: Request - Response (Client-Server).	Mô hình: Publish - Subscribe (qua Broker).
Kết nối: Ngắn, đóng sau khi xong.	Kết nối: Dài, <i>Keep-alive</i> , nhẹ.
Header: Lớn, cồng kềnh (Metadata).	Header: Rất nhỏ (2 byte), tiết kiệm băng thông.
Use Case: Web, API, truyền tải lớn.	Use Case: IoT, mạng chập chờn, pin yếu.

3.10 Pipelines & Orchestration

3.10.1 Apache Kafka vs Airflow

Apache Kafka	Apache Airflow
Loại: Event Streaming Platform (Message Broker).	Loại: Workflow Orchestration (Quản lý quy trình).
Đặc trưng: <ul style="list-style-type: none"> Log bền vững (Durable log). Decoupling (Tách rời). Replayable, High throughput. 	Đặc trưng: <ul style="list-style-type: none"> Code-as-infra (Python DAGs). Quản lý dependency phức tạp. Backfill (chạy lại quá khứ).
Vai trò: "Xương sống" vận chuyển dữ liệu <i>Real-time</i> .	Vai trò: "Nhạc trưởng" điều phối Job (<i>Batch/ETL</i>).

3.10.2 CDC (Change Data Capture)

Theo dõi và đồng bộ thay đổi từ database nguồn sang đích.

- Mục đích:** Real-time data replication, sync giữa OLTP và OLAP, event-driven architecture.
- Cơ chế:**
 - Log-based CDC:* Đọc database transaction log (binlog MySQL, WAL PostgreSQL). **Tốt nhất** - không ảnh hưởng source.
 - Trigger-based:* Trigger trên INSERT/UPDATE/DELETE.Ảnh hưởng performance.
 - Timestamp/Version-based:* Poll dựa trên `updated_at` column. Thiếu DELETE events.
- Tools:** Debezium (Kafka Connect), AWS DMS, Oracle GoldenGate, Airbyte.
- Use Case:** Sync OLTP → DW, Microservices data sharing, Cache invalidation, Audit logs.
- Pattern:** Source DB → CDC Tool → Kafka → Sink (DW/Cache/Search).

3.10.3 ETL vs ELT

- ETL (Extract-Transform-Load):** Transform *trước* khi vào kho. Schema-on-write. Dữ liệu sạch, bảo mật. (Truyền thống).

- ELT (Extract-Load-Transform):** Load raw vào kho *trước*, transform sau. Schema-on-read. Tận dụng sức mạnh Cloud DW (BigQuery, Snowflake). (Hiện đại).

3.11 Kho Dữ liệu (Data Warehousing)

3.11.1 OLTP vs OLAP

	OLTP (Transactional)	OLAP (Analytical)
Mục tiêu	Vận hành hàng ngày (Operational).	Ra quyết định (Decision support).
Dữ liệu	Hiện hành, chi tiết, cập nhật liên tục.	Lịch sử, tổng hợp, đa chiều.
Truy vấn	Đơn giản, trả về ít dòng (Lookup).	Phức tạp, join nhiều, quét băng lớn.
Thiết kế	Chuẩn hóa cao (3NF) để tránh dị thường.	Phi chuẩn hóa (Star/Snowflake) để đọc nhanh.
User	NV, App, Khách hàng.	Manager, Data Analyst.

3.11.2 4 Đặc trưng Chính (Inmon)

- Hướng chủ đề (Subject-oriented):** Tổ chức theo chủ đề chính (Khách hàng, Sản phẩm) thay vì theo ứng dụng (App Bán hàng, App Kho).
- Tích hợp (Integrated):** Dữ liệu từ nhiều nguồn được làm sạch, đồng nhất (đơn vị, format, encoding) trước khi nạp.
- Bất biến (Non-volatile):** Dữ liệu đã vào DW thì (thường) không bị sửa/xóa, chỉ đọc.
- Biến thiên theo thời gian (Time-variant):** Mọi dữ liệu đều gắn với mốc thời gian để phân tích xu hướng (Historical data).

3.11.3 Thách thức Xây dựng DW

- Data Quality:** "Garbage In, Garbage Out". Dữ liệu nguồn bẩn làm sai lệch báo cáo.
- ETL Complexity:** Tích hợp các hệ thống cũ (Legacy) rất phức tạp.
- Performance:** Truy vấn phân tích tốn tài nguyên, cần tối ưu index/partition.
- User Acceptance:** Người dùng không hiểu hoặc không tin tưởng dữ liệu.
- Cost:** Chi phí lưu trữ và duy trì hạ tầng cao.

3.11.4 Mô hình hóa (Modeling)

- Star Schema:** Fact ở giữa, Dimension xung quanh. Phi chuẩn hóa dimension. *Hiệu năng cao, dễ query.*
- Snowflake Schema:** Chuẩn hóa dimension (tách nhỏ). *Tiết kiệm không gian, join phức tạp.*

3.11.5 DW vs DL vs Lakehouse

- Data Warehouse:** Dữ liệu có cấu trúc, cho BI/Reporting.

- **Data Lake:** Dữ liệu thô (Raw), đa dạng, giá rẻ, cho ML/DS.
- **Lakehouse:** Kết hợp (Lưu trữ rẻ của Lake + Quản lý/ACID của Warehouse).

3.11.6 SCD (Slowly Changing Dimensions)

Tại sao cần? Để đảm bảo báo cáo lịch sử chính xác. Nếu KH chuyển từ HCM ra HN năm 2024, doanh số năm 2020 vẫn phải tính cho HCM.

Type	Chiến lược	Đặc điểm & Use Case
0	Retain Original: Giữ nguyên, không bao giờ sửa.	Dữ liệu gốc là chân lý. (VD: Ngày sinh).
1	Overwrite: Ghi đè giá trị mới lên cũ.	Không cần lịch sử. Sửa lỗi chính tả.
2	Add Row: Thêm dòng mới + <i>Effective Date</i> + <i>Current_Flag</i> .	Chuẩn mực nhất. Theo dõi toàn bộ lịch sử biến động.
3	Add Column: Thêm cột <i>Previous_Value</i> .	Chỉ cần biết giá trị liền trước. (Ít dùng).
4	Add History Table: Tách bảng lịch sử riêng (Mini-Dimension).	Tối ưu khi bảng chính quá lớn và chỉ một nhóm thuộc tính thay đổi nhanh.
5	Hybrid (4 + 1): Mini-dimension + tham chiếu "Current" ở bảng chính.	Tối ưu truy vấn khi cần cả lịch sử chi tiết và giá trị hiện tại nhanh chóng.
6	Hybrid (1 + 2 + 3): Type 2 row + cột chứa giá trị hiện tại (Type 1).	"Pure Type 6": Giúp truy vấn lịch sử nhưng vẫn group by theo giá trị hiện tại dễ dàng.
7	Hybrid (Dual Keys): Fact table chứa cả Surrogate Key (lịch sử) và Natural Key (hiện tại).	Linh hoạt nhất: Join theo Surrogate để xem lịch sử, join theo Natural để xem hiện tại.

Các kỹ thuật xử lý dữ liệu thay đổi theo thời gian (0 → 7).

3.12 Quản lý Dữ liệu

3.12.1 Vấn đề Tích hợp (Data Integration Issues)

- **Heterogeneous data sources:** Khác biệt về hệ quản trị cơ sở dữ liệu và định dạng tệp.
VD: MySQL, PostgreSQL, Oracle, MongoDB, CSV, JSON, XML, Parquet.
 - **Data Mapping:** Khác biệt về cấu trúc schema giữa các nguồn.
VD: Bảng 'Employee' (Full_Name, DOB) vs 'Emp' (First_Name, Last_Name, Birth_Date).
 - **Data Conflicts:** Xung đột về kiểu dữ liệu, giá trị, định dạng, đơn vị, định nghĩa.
- Kiểu: String vs Integer cho mã nhân viên. Định dạng: DD/MM/YYYY vs MM/DD/YYYY.

Đơn vị: USD vs VND, km vs miles. **Độ chính xác:** 2 vs 4 chữ số thập phân.

- **Data Redundancy:** Dữ liệu trùng lặp từ nhiều nguồn cần khử trùng (Deduplication).

VD: Cùng khách hàng xuất hiện trong CRM và ERP với ID khác nhau.

- **Entity Resolution:** Xác định 2 bản ghi từ 2 nguồn khác nhau là cùng 1 thực thể.

VD: 'Nguyen Van A' (DB1) và 'A Nguyen' (DB2), 'IBM' vs 'International Business Machines'.

Kỹ thuật: Fuzzy matching, similarity scores (Levenshtein distance), Master Data Management (MDM).

- **Constraints Violation:** Vi phạm ràng buộc khi tích hợp dữ liệu.
Primary Key: Trùng lặp khóa chính khi merge. **Foreign Key:** Tham chiếu không tồn tại.

Semantic: Giá trị không hợp lệ (tuổi âm, ngày trong tương lai).

- **Data Quality Issues:** Các vấn đề chất lượng dữ liệu trong quá trình tích hợp.

Accuracy: Sai lệch so với thực tế. **Completeness:** Thiếu trường bắt buộc (NULL).

Uniqueness: Trùng lặp. **Timeliness:** Dữ liệu cũ/lỗi thời. **Consistency:** Mâu thuẫn giữa các nguồn.

- **Communication Heterogeneity:** Khác biệt về giao diện và giao thức truyền thông.

VD: REST API vs SOAP, HTTP vs FTP, Batch files vs Real-time streams, GraphQL vs SQL.

3.12.2 Chiều Chất lượng Dữ liệu (Data Quality Dimensions)

Các tiêu chí đánh giá chất lượng dữ liệu trong hệ thống thông tin.

6 Chiều Cốt lõi (Core Dimensions)

1. **Accuracy (Chính xác):** Mức độ dữ liệu phản ánh đúng đối tượng/sự kiện thực tế.

Định nghĩa: Sự không chính xác có nghĩa là hệ thống biểu diễn một trạng thái thế giới thực khác với trạng thái đáng lẽ phải được biểu diễn.

VD: Địa chỉ sai, số điện thoại cũ, thông tin lỗi thời.

2. **Completeness (Đầy đủ):** Tỷ lệ dữ liệu được lưu trữ so với khả năng 100% hoàn chỉnh.

Định nghĩa: Khả năng của hệ thống biểu diễn mọi trạng thái có nghĩa của hệ thống thế giới thực.

VD: NULL values, missing fields, incomplete records (thiếu email, số điện thoại).

3. **Consistency (Nhất quán):** Sự vắng mặt của sự khác biệt khi so sánh hai hoặc nhiều biểu diễn của cùng một thứ.

Định nghĩa: Sự không nhất quán có nghĩa là ánh xạ biểu diễn là một-nhiều (one-to-many).

VD: Tổng doanh thu trong bảng Orders ≠ tổng trong Report, tên KH khác nhau giữa CRM và ERP.

4. **Validity (Hợp lệ):** Dữ liệu hợp lệ nếu tuân thủ cú pháp của định nghĩa (format, type, range).

VD: Email đúng format (có @), tuổi trong khoảng [0, 150], enum đúng giá trị cho phép.

5. **Timeliness (Kịp thời):** Mức độ dữ liệu phản ánh thực tế từ mốc thời gian yêu cầu.

Định nghĩa: Độ trễ giữa thay đổi trạng thái thế giới thực và sửa đổi tương ứng trong hệ thống.

VD: Giá cổ phiếu cập nhật delay 10s, báo cáo tồn kho cách 1 ngày (stale data).

6. **Uniqueness (Duy nhất):** Không có thứ nào được ghi lại nhiều hơn một lần dựa trên cách xác định thứ đó.

VD: Không có duplicate records, một khách hàng chỉ có một ID duy nhất.

Các Chiều Mở rộng (Extended Dimensions)

- **Interpretability (Khả năng Giải thích):** Liên quan đến tài liệu và metadata có sẵn để diễn giải chính xác ý nghĩa và thuộc tính của dữ liệu.

VD: Data dictionary, schema documentation, column descriptions, business glossary.

- **Accessibility (Khả năng Truy cập):** Độ lưỡng khả năng người dùng truy cập dữ liệu từ văn hóa, trạng thái/chức năng vật lý và công nghệ có sẵn của họ.

VD: API có rate limit, quyền truy cập dựa trên role (RBAC), hỗ trợ đa ngôn ngữ.

- **Usability (Khả dụng):** Độ lưỡng hiệu quả, hiệu suất, sự hài lòng mà người dùng cụ thể cảm nhận và sử dụng dữ liệu.

VD: Dashboard dễ hiểu, query response time nhanh, format phù hợp với use case.

- **Trustworthiness (Độ Tin cậy):** Độ lưỡng mức độ đáng tin cậy của tổ chức trong việc cung cấp nguồn dữ liệu.

VD: Data lineage rõ ràng, audit logs đầy đủ, SLA đảm bảo, data provenance.

3.12.3 Quản lý Thông tin (Information Management)

Quản lý dữ liệu như một tài sản chiến lược của doanh nghiệp.

6 Khía cạnh Quản lý Thông tin

1. **Information Collection (Thu thập):** Xác định và thu thập dữ liệu từ các nguồn khác nhau.

Hoạt động: Data ingestion, ETL pipelines, API integration, web scraping, sensors/IoT.

VD: Crawl web data, stream từ Kafka, batch import từ CSV, CDC từ transactional DB.

2. **Information Organization (Tổ chức):** Cấu trúc hóa và phân loại dữ liệu để dễ quản lý và truy xuất.

Hoạt động: Schema design, data modeling (Star/Snowflake), taxonomy, metadata management.

VD: Thiết kế DW dimensions/facts, tạo data catalog, tag dữ liệu theo domain.

3. **Information Storage (Lưu trữ):** Chọn và triển khai hệ thống lưu trữ phù hợp với đặc điểm dữ liệu.

Hoạt động: Chọn DBMS (RDBMS, NoSQL, DW), partitioning, replication, backup strategy.

VD: OLTP trên PostgreSQL, OLAP trên Snowflake, unstructured data trên S3/Data Lake.

4. **Information Manipulation (Thao tác):** Thực hiện các thao tác CRUD và chuyển đổi dữ liệu.

Hoạt động: INSERT/UPDATE/DELETE, data transformation,

cleansing, enrichment, aggregation.

VD: Chuẩn hóa địa chỉ, deduplicate records, merge datasets, derive calculated fields.

5. Information Processing (Xử lý): Phân tích và tính toán để tạo insights từ dữ liệu thô.

Hoạt động: Query execution, analytics, ML training, reporting, dashboarding, data mining.

VD: SQL queries, Spark jobs, BI reports, predictive models, real-time analytics.

6. Information Protection (Bảo vệ): Đảm bảo an toàn, riêng tư và tuân thủ quy định.

Hoạt động: Access control (RBAC), encryption (at-rest/in-transit), audit logging, compliance (GDPR).

VD: Mã hóa PII, masking sensitive data, role-based permissions, backup/disaster recovery.

Tổng thể: Information Management bao gồm Data Governance (quản trị), Data Quality, Master Data Management (MDM), Security, và Lifecycle Management. Mục tiêu cuối cùng là đảm bảo dữ liệu **tin cậy, an toàn, dễ truy cập** để tạo giá trị kinh doanh.

3.13 Thách thức Hệ Phân tán (Distributed Systems Challenges)

Khác với hệ thống đơn, hệ phân tán đối mặt với *partial failures* - một số phần hỏng trong khi phần khác hoạt động.

3.13.1 8 Fallacies (Ngụy biện) của Distributed Computing

Những giả định sai lầm mà lập trình viên thường mắc phải (L. Peter Deutsch, Sun Microsystems).

1. The Network is Reliable: Mạng không đáng tin cậy - switches hỏng, cables ngắt, packets bị mất/reorder.

2. **Latency is Zero:** Local call ($\text{ns}/\mu\text{s}$) \neq Remote call (ms). Chain calls \rightarrow latency tích lũy.
3. **Bandwidth is Infinite:** Băng thông hữu hạn. Stamp coupling (truyền dữ liệu thừa) gây bottleneck.
4. **The Network is Secure:** Attack surface tăng. Mọi endpoint phải được bảo mật.
5. **Topology Never Changes:** Topology thay đổi liên tục (upgrades, failures, scaling) \rightarrow timeout/failures.
6. **Only One Administrator:** Nhiều admin quản lý các segment khác nhau (firewall, DB) \rightarrow khó phối hợp.
7. **Transport Cost is Zero:** Serialization, marshalling, network infra đều tốn chi phí.
8. **Network is Homogeneous:** Mạng gồm nhiều vendor (routers, switches) \rightarrow packet loss, anomalies.

3.13.2 Unreliability: Network & Time

- **Network Unreliability:**
 - *Timeouts:* Không có bounded time \rightarrow phải dùng timeout. Ngắn quá: false positive. Dài quá: delay detection.
 - *Network Partition (Netsplit):* Nhóm nodes bị cô lập \rightarrow "Split brain"(2 leaders cùng chấp nhận ghi xung đột).
- **Time Unreliability:** "Time is an illusion không có global clock.
 - *Clock Drift:* Đồng hồ quartz drift do nhiệt độ. Ngay cả NTP sync vẫn có sai lệch.
 - *Process Pauses:* GC hoặc VM suspension \rightarrow node pause lâu hơn lease timeout \rightarrow bị khai tử nhưng vẫn nghĩ mình sống \rightarrow data corruption.
 - *Ordering Issues:* Last Write Wins (LWW) dùng timestamp nguy hiểm - clock không sync \rightarrow ghi đè giá trị mới bằng cũ.

3.13.3 Consistency & Consensus Challenges

- **CAP Theorem:** Chỉ đạt 2/3: Consistency, Availability, Partition Tolerance. Partition không tránh được \rightarrow chọn CP hoặc AP.

\rightarrow NoSQL thường chọn AP (*Sẵn sàng + Chịu lỗi*) thay vì CP.

- **FLP Impossibility:** Hệ async không thể đảm bảo consensus termination nếu có 1 node crash \rightarrow phải dùng timeout.
- **Two Generals' Problem:** Không thể đạt common knowledge (chắc chắn đồng ý) qua kênh unreliable - ACK cuối cùng có thể bị mất.
- **BASE:** Basically Available, Soft state, Eventual consistency (Nhất quán cuối cùng - chấp nhận dữ liệu cũ tạm thời).

3.13.4 Failure Models

- **Crash Faults:** Process dừng hẳn, không bao giờ quay lại (đơn giản nhất).
- **Omission Faults:** Process bỏ qua bước hoặc không gửi/nhận message (buffer overflow, congestion).
- **Byzantine Faults:** Process hành xử tùy ý/dộc hại, gửi message sai (khó nhất, blockchain).

3.13.5 Managing "Truth" in Distributed Systems

- **Truth by Majority:** Không node nào tin được view của chính mình. *Sự thật = Quorum quyết định.*
VD: Node tưởng mình là leader, nhưng quorum khai tử (do GC pause) \rightarrow node đó "đã chết".
- **Fencing Tokens:** Số monotonic để storage từ chối ghi từ node cầm lock đã hết hạn (zombie nodes).

3.13.6 Thách thức Xử lý Dữ liệu Phân tán

- **Data Skew (Lệch dữ liệu):** Một partition chứa quá nhiều dữ liệu.
 - \rightarrow *Hậu quả:* Straggler problem - job đợi node chậm nhất.
 - \rightarrow *Giải pháp:* Salting (thêm prefix ngắn nhiên) để chia nhỏ hot key.
- **Shuffle:** Chuyển dữ liệu giữa nodes qua mạng (Map \rightarrow Reduce).
 - \rightarrow *Tối ưu:* Broadcast Join (chèp bảng nhỏ đến tất cả node) để tránh shuffle bảng lớn.