**TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM**

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**THUYẾT TRÌNH**

**CƠ SỞ DỮ LIỆU PHÂN TÁN**

**PRINCIPLES OF DISTRIBUTED DATABASE SYSTEMS - CHƯƠNG 2**

*Người hướng dẫn*: **Ths. Nguyễn Thanh Khương**

*Người thực hiện:* **Nguyễn Khánh An – 51800166**

**Võ Hoàng Long – 51800898**

**Đậu Minh Lượng – 51403417**

**Lâm Phúc Nghi – 51403239**

**Đinh Đông Thức - 51800935**

THÀNH PHỐ HỒ CHÍ MINH, NĂM 2021

**TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM**

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**THUYẾT TRÌNH**

**CƠ SỞ DỮ LIỆU PHÂN TÁN**

**PRINCIPLES OF DISTRIBUTED DATABASE SYSTEMS - CHƯƠNG 2**

*Người hướng dẫn*: **Ths. Nguyễn Thanh Khương**

*Người thực hiện:* **Nguyễn Khánh An – 51800166**

**Võ Hoàng Long – 51800898**

**Đậu Minh Lượng – 51403417**

**Lâm Phúc Nghi – 51403239**

**Đinh Đông Thức - 51800935**

THÀNH PHỐ HỒ CHÍ MINH, NĂM 2021

LỜI CẢM ƠN

Qua nghiên cứu và thực hiện, nhóm chúng em đã hoàn thành bài thuyết trình môn Cơ sở dữ liệu phân tán.

Nhóm chúng em xin cảm ơn Ths. Nguyễn Thanh Khương đã hướng dẫn nhiệt tình những kiến thức cơ bản và hướng thực hiện bài thuyết trình này giúp chúng em có được nền tảng vững chắc để hoàn thành đề tài, thầy đã tận tình giải đáp những thắc mắc và hướng dẫn cách phát triển về đề tài này trong việc nghiên cứu.

Do kiến thức còn hạn hẹp, nên phần báo cáo của nhóm chúng em có thể còn nhiều sai sót. Em mong thầy/cô thông cảm và góp ý, nhận xét để nhóm em có thể tiếp tục khắc phục được những sai sót.

Nhóm em xin chân thành cảm ơn.

**BÀI THUYẾT TRÌNH ĐƯỢC HOÀN THÀNH**

**TẠI TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

Chúng tôi xin cam đoan đây là sản phẩm thuyết trình của riêng nhóm chúng tôi và được sự hướng dẫn của Ths. Nguyễn Thanh Khương. Các nội dung nghiên cứu, kết quả trong đề tài này là trung thực và chưa công bố dưới bất kỳ hình thức nào trước đây. Những số liệu trong các bảng biểu phục vụ cho việc phân tích, nhận xét, đánh giá được chính tác giả thu thập từ các nguồn khác nhau có ghi rõ trong phần tài liệu tham khảo.

**Nếu phát hiện có bất kỳ sự gian lận nào chúng tôi xin hoàn toàn chịu trách nhiệm về nội dung tiểu luận của mình.** Trường đại học Tôn Đức Thắng không liên quan đến những vi phạm tác quyền, bản quyền do chúng tôi gây ra trong quá trình thực hiện (nếu có).

*TP. Hồ Chí Minh, ngày tháng năm*

*Nhóm tác giả*

*(ký tên và ghi rõ họ tên)*

*Nguyễn Khánh An*

*Võ Hoàng Long*

*Đậu Minh Lượng*

*Lâm Phúc Nghi*

*Đinh Đông Thức*

PHẦN XÁC NHẬN VÀ ĐÁNH GIÁ CỦA GIẢNG VIÊN

Phần xác nhận của GV hướng dẫn

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Tp. Hồ Chí Minh, ngày tháng năm

(kí và ghi họ tên)

Phần đánh giá của GV chấm bài

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Tp. Hồ Chí Minh, ngày tháng năm

(kí và ghi họ tên)

**TÓM TẮT**

Trình bày những nội dung trong chương 2 sách “Principles of distributed database systems” cụ thể:

* Xây dựng lược đồ cơ sở dữ liệu, chuẩn hóa cơ sở dữ liệu, các phép toán đại số quan hệ, phép tính quan hệ.
* Tổng quan về mạng máy tính: khái niệm, các loại mạng, các hình thức giao tiếp, giao tiếp dữ liệu, cấu trúc frame, sơ lược chuyển mạch kênh và chuyển mạch gói, giao thức TCP/IP.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **PHÂN CÔNG CÔNG VIỆC** | | | | |
| **STT** | **TÊN** | **MSSV** | **CÔNG VIỆC** | **ĐÁNH GIÁ** |
| 1 | Nguyễn Khánh An (Nhóm trưởng) | 51800166 | 2.2.3 Data Communication Concepts (*đề mục theo sách*).  - Tìm hiểu nội dung.  - Tổng hợp thông tin cả nhóm, trình bày slide báo cáo. | 20% |
| 2 | Võ Hoàng Long | 51800898 | 2.1.1 Relational Database Concepts  2.1.2 Normalization  - Tìm hiểu nội dung. Tổng hợp thông tin, hình ảnh cho nhóm trưởng. | 20% |
| 3 | Đậu Minh Lượng | 51403417 | 2.2.4 Communication Protocols  - Tìm hiểu nội dung. Tổng hợp thông tin, hình ảnh cho nhóm trưởng. | 20% |
| 4 | Lâm Phúc Nghi | 51403239 | 2.1.3 Relational Data Languages  - Tìm hiểu nội dung. Tổng hợp thông tin, hình ảnh cho nhóm trưởng.  - Trình bày báo cáo (word). | 20% |
| 5 | Đinh Đông Thức | 51800935 | 2.2.1 Types of Networks.  2.2.2 Communication Schemes  - Tìm hiểu nội dung. Tổng hợp thông tin, hình ảnh cho nhóm trưởng. | 20% |

MỤC LỤC

[Chương 1. TỔNG QUAN VỀ CƠ SỞ DỮ LIỆU QUAN HỆ. 1](#_Toc68633136)

[1.1. Khái niệm về cơ sở dữ liệu quan hệ. 1](#_Toc68633137)

[1.2. Chuẩn hóa. 2](#_Toc68633138)

[1.2.1. Lặp lại dị thường. 2](#_Toc68633139)

[1.2.2. Cập nhật dị thường. 3](#_Toc68633140)

[1.2.3. Chèn dữ liệu dị thường. 3](#_Toc68633141)

[1.2.4. Xóa bất thường. 4](#_Toc68633142)

[1.2.5. Các dạng chuẩn. 5](#_Toc68633143)

[1.3. Đại số quan hệ (Relational Algebra). 5](#_Toc68633144)

[1.3.1. Phép chọn (Selection σ). 6](#_Toc68633145)

[1.3.2. Phép chiếu (Projection Π). 6](#_Toc68633146)

[1.3.3. Phép hội (Union ). 7](#_Toc68633147)

[1.3.4. Phép hiệu (Set Difference - ). 9](#_Toc68633148)

[1.3.5. Phép tích Descarte (Cartesian product x). 10](#_Toc68633149)

[1.3.6. Phép giao (Interesection ). 12](#_Toc68633150)

[1.3.7. Phép Theta kết (). 14](#_Toc68633151)

[1.3.8. Phép kết ngoài (Outer-join). 16](#_Toc68633152)

[1.3.9. Phép kết nửa (semijoin). 19](#_Toc68633153)

[1.3.10. Phép chia (Division ÷). 20](#_Toc68633154)

[1.4. Phép tính quan hệ (Relational Calculus). 21](#_Toc68633155)

[1.4.1. Phép tính quan hệ trên bộ (tuple relational calculus). 21](#_Toc68633156)

[1.4.2. Phép tính quan hệ trên miền (domain relational calculus). 22](#_Toc68633157)

[Chương 2. TỔNG QUAN VỀ MẠNG MÁY TÍNH. 23](#_Toc68633158)

[2.1. Khái niệm. 23](#_Toc68633159)

[2.2. Các loại mạng. 23](#_Toc68633160)

[2.3. Các hình thức giao tiếp. 24](#_Toc68633161)

[2.3.1. Point-to-point (unicast) 24](#_Toc68633162)

[2.3.2. Broadcast (đa điểm) 25](#_Toc68633163)

[2.3.3. Ngoài ra, còn các hình thức khác. 25](#_Toc68633164)

[2.4. Các khái niệm về giao tiếp dữ liệu. 25](#_Toc68633165)

[2.5. Cấu trúc Frame. 26](#_Toc68633166)

[2.6. Sơ lược chuyển mạch kênh và chuyển mạch gói. 27](#_Toc68633167)

[2.7. Các giao thức giao tiếp. 27](#_Toc68633168)

**DANH MỤC CÁC BẢNG BIỂU, HÌNH VẼ, ĐỒ THỊ**

**DANH MỤC HÌNH ẢNH**

[Hình 1‑1 Bảng EMP (nhân viên) và PROJ (dự án). 1](#_Toc68633169)

[Hình 1‑2 Bảng EMP và PROJ. 2](#_Toc68633170)

[Hình 1‑3 Bảng EMP và PROJ. 3](#_Toc68633171)

[Hình 1‑4 Bảng EMP và PROJ. 4](#_Toc68633172)

[Hình 1‑5 X xác định Y. 5](#_Toc68633173)

[Hình 1‑6 Bảng khách hàng. 6](#_Toc68633174)

[Hình 1‑7 Chọn khách hàng có doanh số lớn hơn 300.000. 6](#_Toc68633175)

[Hình 1‑8 Chọn khách hàng có địa chỉ ở quận 5. 6](#_Toc68633176)

[Hình 1‑9 Bảng khách hàng. 7](#_Toc68633177)

[Hình 1‑10 Lấy họ tên và số điện thoại khách hàng. 7](#_Toc68633178)

[Hình 1‑11 Minh họa phép hội. 7](#_Toc68633179)

[Hình 1‑12 Bảng khách hàng. 8](#_Toc68633180)

[Hình 1‑13 Bảng cung cấp. 8](#_Toc68633181)

[Hình 1‑14 Kết quả phép hội lấy tên Thành phố. 9](#_Toc68633182)

[Hình 1‑15 Minh họa phép hiệu. 9](#_Toc68633183)

[Hình 1‑16 Kết quả phép hiệu. 9](#_Toc68633184)

[Hình 1‑17 Bảng CTHD (chi tiết hóa đơn) 10](file:///D:\Workspace\HK2_2020_2021\CSDLPhanTan\ThuyetTrinh20\NhomB.docx#_Toc68633185)

[Hình 1‑18 Bảng sản phẩm. 10](#_Toc68633186)

[Hình 1‑19 Bảng hóa đơn. 10](#_Toc68633187)

[Hình 1‑20 Bảng khách hàng. 11](#_Toc68633188)

[Hình 1‑21 Kết quả phép tích. 11](#_Toc68633189)

[Hình 1‑22 Kết quả phép tích sau khi thêm điều kiện chọn. 12](#_Toc68633190)

[Hình 1‑23 Minh họa phép giao. 12](#_Toc68633191)

[Hình 1‑24 Bảng CTHD. 13](file:///D:\Workspace\HK2_2020_2021\CSDLPhanTan\ThuyetTrinh20\NhomB.docx#_Toc68633192)

[Hình 1‑25 Bảng sản phẩm. 13](file:///D:\Workspace\HK2_2020_2021\CSDLPhanTan\ThuyetTrinh20\NhomB.docx#_Toc68633193)

[Hình 1‑26 Kết quả phép giao. 13](#_Toc68633194)

[Hình 1‑27 Bảng HOADON. 14](#_Toc68633195)

[Hình 1‑28 Bảng KHACHHANG. 14](#_Toc68633196)

[Hình 1‑29 Kết quả phép tích sau khi thêm điều kiện chọn. 15](#_Toc68633197)

[Hình 1‑30 Bảng CTHD. 15](file:///D:\Workspace\HK2_2020_2021\CSDLPhanTan\ThuyetTrinh20\NhomB.docx#_Toc68633198)

[Hình 1‑31 Bảng sản phẩm. 15](file:///D:\Workspace\HK2_2020_2021\CSDLPhanTan\ThuyetTrinh20\NhomB.docx#_Toc68633199)

[Hình 1‑32 Bảng kết quả equijoin. 16](#_Toc68633200)

[Hình 1‑33 Bảng kết quả natural-join. 16](#_Toc68633201)

[Hình 1‑34 Bảng nhân viên. 17](#_Toc68633202)

[Hình 1‑35 Bảng hóa đơn. 17](#_Toc68633203)

[Hình 1‑36 Kết quả left outer join. 17](#_Toc68633204)

[Hình 1‑37 Kết quả right outer join. 18](#_Toc68633205)

[Hình 1‑38 Bảng khách hàng. 18](#_Toc68633206)

[Hình 1‑39 Bảng đơn hàng. 19](#_Toc68633207)

[Hình 1‑40 Bảng kết quả full outer join. 19](#_Toc68633208)

[Hình 1‑41 Bảng EMP. 20](file:///D:\Workspace\HK2_2020_2021\CSDLPhanTan\ThuyetTrinh20\NhomB.docx#_Toc68633209)

[Hình 1‑42 Bảng PAY. 20](file:///D:\Workspace\HK2_2020_2021\CSDLPhanTan\ThuyetTrinh20\NhomB.docx#_Toc68633210)

[Hình 1‑43 Kết quả phép semi join. 20](#_Toc68633211)

[Hình 2‑1 Đinh dạng frame điển hình. 26](file:///D:\Workspace\HK2_2020_2021\CSDLPhanTan\ThuyetTrinh20\NhomB.docx#_Toc68633212)

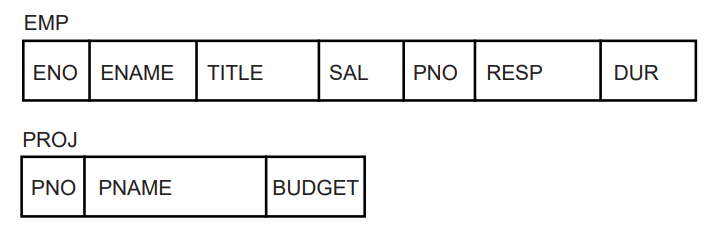
[Hình 2‑2 Truyền tin bằng TCP/IP. 28](#_Toc68633213)

# TỔNG QUAN VỀ CƠ SỞ DỮ LIỆU QUAN HỆ.

## Khái niệm về cơ sở dữ liệu quan hệ.

Một CSDL là một tập cấu trúc dữ liệu liên quan đến các hiện tượng ngoài đời thực mà chúng ta đang mô phỏng. Một dữ liệu quan hệ là một cấu trúc cơ dữ liệu ở dạng bảng.

*Ví dụ*: CSDL mô phỏng một công ty kỹ thuật.



Hình 1‑1 Bảng EMP (nhân viên) và PROJ (dự án).

Các thực thể sẽ là nhân viên (EMP) và dự án (PROJ). Với mỗi nhân viên, sẽ có số điện thoại nhân viên (ENO), tên nhân viên (ENAME), chức danh trong công ty (TITLE), lương (SAL), mã số dự án của nhân viên đang tham gia(PNO), vai trò trong dự án (RESP), và thời gian được giao để hoàn thành dự án (DUR).

Lược đồ quan hệ cho CSDL trên:

* EMP (ENO, ENAME, TITLE, SAL, PNO, RESP, DUR)
* PROJ (PNO, PNAME, BUDGET)

Trong lược đồ quan hệ EMP có 7 thuộc tính: ENO, ENAME, TITLE, SAL, PNO, RESP, DUR. Giá trị của ENO là miền dữ liệu của tất cả mã số nhân viên hợp lệ, chẳng hạn như D1, giá trị của ENAME là miền dữ liệu đến của tất cả tên hợp lệ, chẳng hạn như D2 và tiếp tục như thế.

Khóa của lược đồ quan hệ là một tập con khác rỗng sao cho các giá trị của các thuộc tính gồm cả khóa sẽ là duy nhất từng bộ của quan hệ. Những thuộc tính dùng để làm khóa được gọi là thuộc tính chính. Một tập các khóa thường được gọi là một siêu khóa. Như vậy đối với ví dụ khóa của PROJ là PNO, và của EMP là tập (ENO, PNO).

Một giá trị thuộc tính có thể không được xác định.

Có thể hiểu theo nhiều cách khác nhau, phổ biến nhất là “không xác định” hoặc “không áp dụng”. Giá trị đặc biệt của thuộc tính này được gọi là null.

Biểu diễn giá trị null phải khác với bất kỳ giá trị miền nào khác và cần đặc biệt chú ý phân biệt nó với giá trị 0.

## Chuẩn hóa.

Mục tiêu của chuẩn hóa là loại bỏ các điểm bất thường (hoặc các khía cạnh không mong muốn) của một quan hệ để có được các quan hệ tốt hơn.

Bốn vấn đề có thể tồn tại trong một lược đồ quan hệ.

### Lặp lại dị thường.

Một số dữ liệu có thể bị lặp lại không cần thiết.



Hình 1‑2 Bảng EMP và PROJ.

*Ví dụ*: xem xét mối quan hệ EMP trong hình: Tên, chức danh và mức lương của một nhân viên được lặp lại cho mỗi dự án mà người này tham gia vào.

### Cập nhật dị thường.

Do dữ liệu bị lặp lại, tốc độ cập nhật có thể bị chậm đi.



Hình 1‑3 Bảng EMP và PROJ.

Do dữ liệu bị lặp lại, tốc độ cập nhật có thể bị chậm.

*Ví dụ:* nếu lương của một nhân viên thay đổi, nhiều dòng sẽ phải thay đổi theo.

### Chèn dữ liệu dị thường.

Không thể thêm dữ liệu mới vào hệ cơ sở dữ liệu.



Hình 1‑4 Bảng EMP và PROJ.

*Ví dụ*: khi một nhân viên được nhận vào công ty, không thể thêm thông tin cá nhân (tên, chức danh, lương) vào EMP cho đến khi người đó tham gia vào dự án. Đó là bởi vì khóa của EMP bao gồm thuộc tính PNO, và giá trị null không được phép thêm vào khóa.

### Xóa bất thường.

Trái ngược với chèn dữ liệu dị thường. Nếu một nhân viên chỉ tham gia vào một dự án, và dự án đó bị hủy, thì không thể xóa dữ liệu dự án đó trong EMP.

Làm như vậy sẽ dẫn đến xóa tuple duy nhất về nhân viên đó, do đó dẫn đến mất thông tin cá nhân mà chúng ta muốn lưu trữ.

Chuẩn hóa sẽ giúp chuyển đổi các lược đồ quan hệ tùy ý thành các lược đồ không có những vấn đề trên.

Một quan hệ với một hoặc nhiều những vấn đề trên sẽ được tách thành 2 hoặc nhiều quan hệ có dạng chuẩn cao hơn.

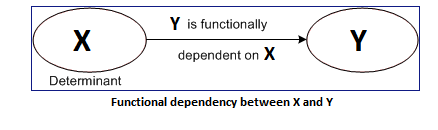
Một quan hệ được cho là ở dạng chuẩn nếu nó thỏa mãn các điều kiện liên quan đến dạng chuẩn đó.

### Các dạng chuẩn.

* Dạng chuẩn 1 (1NF)
* Dạng chuẩn 2 (2NF)
* Dạng chuẩn 3 (3NF)
* Dạng chuẩn Boyce-Codd (BCNF)
* Dạng chuẩn 4 (4NF)
* Dạng chuẩn 5 (5NF)

BCNF và các dạng bình thường thấp hơn dựa trên các phụ thuộc hàm (FD), 4NF dựa trên các phụ thuộc nhiều giá trị và 5NF dựa trên các phụ thuộc kết hợp phép chiếu.

Gọi R là một quan hệ được xác định trên tập các thuộc tính A = {A1, A2, ..., An} và cho X ⊂ A, Y ⊂ A. Nếu với mỗi giá trị của X trong R, chỉ có một giá trị Y liên kết, thì “X xác định Y” hoặc “Y phụ thuộc X” ký hiệu X → Y.



Hình 1‑5 X xác định Y.

## Đại số quan hệ (Relational Algebra).

Là ngôn ngữ lý thuyết (hình thức) gồm các phép toán thực hiện trên một hay hai quan hệ nhằm định nghĩa một quan hệ mới mà không làm thay đổi quan hệ ban đầu. Toán tử: operator, toán hạng: operand.

*Ví dụ*: Trong đại số, toán tử là: +, -, \*, /; toán hạng là các biến: a, b,.. ; biểu thức: (a + b)/(c – d).

Đại số quan hệ có 5 phép toán cơ sở.

### Phép chọn (Selection σ).

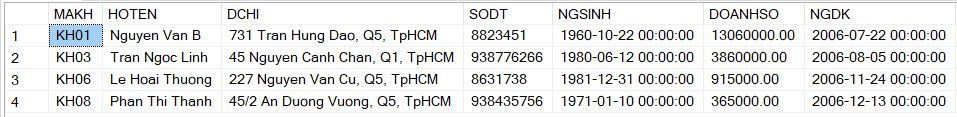
Chọn các dòng từ một quan hệ (chiều ngang).

: quan hệ, : điều kiện.



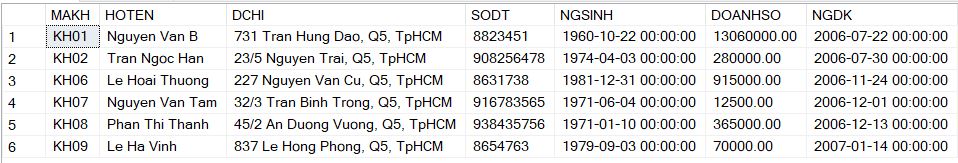
Hình 1‑6 Bảng khách hàng.

*Ví dụ 1*: Chọn khách hàng có doanh số lớn hơn 300.000.



Hình 1‑7 Chọn khách hàng có doanh số lớn hơn 300.000.

*Ví dụ 2*: Chọn khách hàng có địa chỉ Quận 5.



Hình 1‑8 Chọn khách hàng có địa chỉ ở quận 5.

### Phép chiếu (Projection Π).

Chọn các cột từ một quan hệ (chiều dọc), các dòng trùng nhau chỉ lấy 1 dòng.

: quan hệ, : các thuộc tính (các cột).

*Ví dụ*: Lấy Họ tên và số điện thoại của các khách hàng.

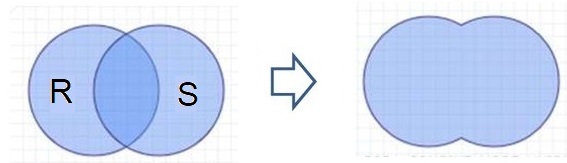


Hình 1‑9 Bảng khách hàng.



Hình 1‑10 Lấy họ tên và số điện thoại khách hàng.

### Phép hội (Union ).

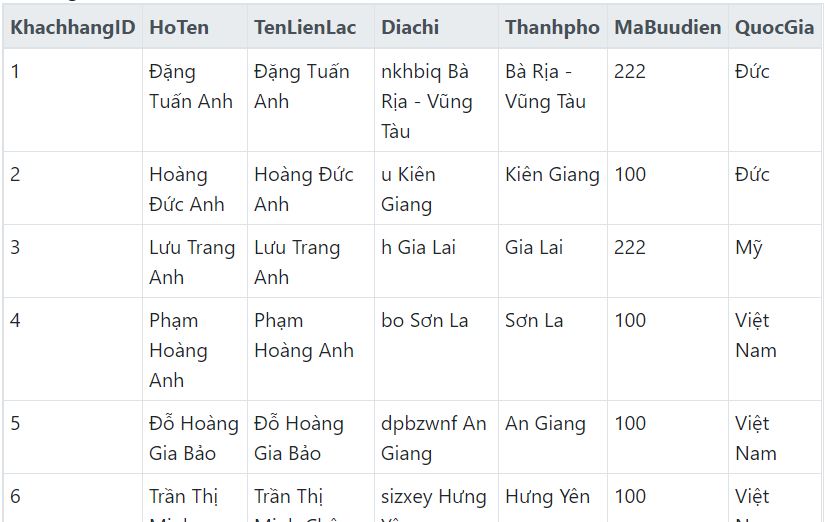


Hình 1‑11 Minh họa phép hội.

Hội của hai quan hệ *R* và *S* tạo một quan hệ mới chứa tất cả các bộ của *R*, hoặc của *S* hoặc của cả 2, loại bộ trùng.

*Ví dụ*: Có 2 bảng: Khách hàng, Cung cấp.

Bảng Khachhang (Khách hàng):



Hình 1‑12 Bảng khách hàng.

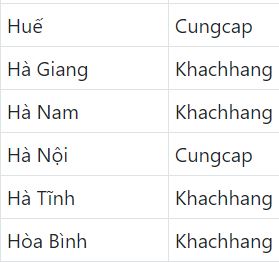
Bảng Cungcap (Cung cấp):

Hình 1‑13 Bảng cung cấp.

Giả sử: R: Lấy Thanhpho (tên thành phố) từ bảng Khachhang (khách hàng).

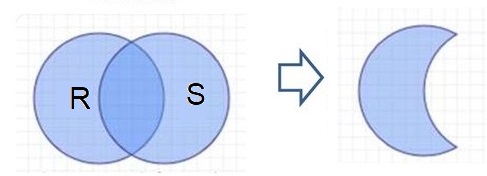
S: Lấy Thanhpho từ bảng Cungcap (cung cấp).

Phép hội của R và S là:



Hình 1‑14 Kết quả phép hội lấy tên Thành phố.

### Phép hiệu (Set Difference - ).



Hình 1‑15 Minh họa phép hiệu.

Hình thành một quan hệ chứa các bộ trong *R* nhưng không thuộc *S*.

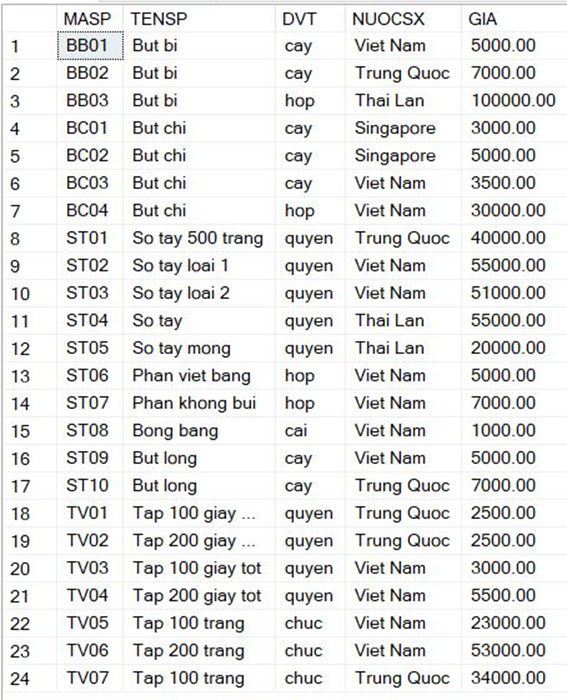
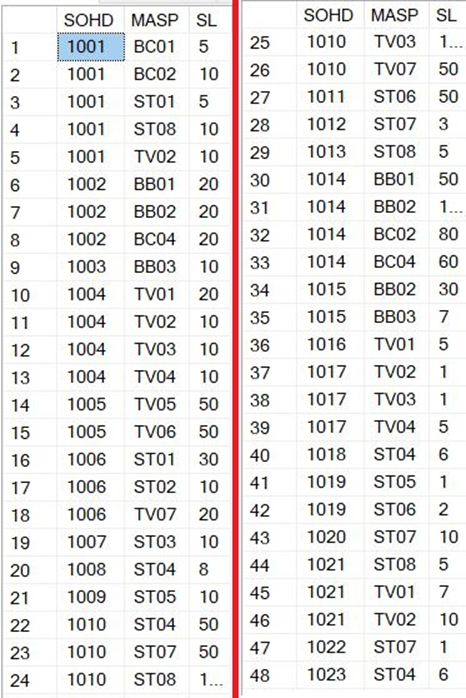
*Ví dụ*: R: Lấy MASP (mã sản phẩm) từ bảng SANPHAM (sản phẩm).

S: Lấy MASP từ bảng CTHD (chi tiết hóa đơn).

Phép hiệu giữa R và S là: (những sản phẩm chưa bán)



Hình 1‑16 Kết quả phép hiệu.



Hình 1‑18 Bảng sản phẩm.

Hình 1‑17 Bảng CTHD (chi tiết hóa đơn)

### Phép tích Descarte (Cartesian product x).

Hình thành một quan hệ nối các bộ của *R* với mọi bộ của *S.*

*Ví dụ:* Xét 2 bảng: HOADON và KHACHANG.

Bảng HOADON (hóa đơn):



Hình 1‑19 Bảng hóa đơn.

Bảng KHACHHANG (khách hàng):

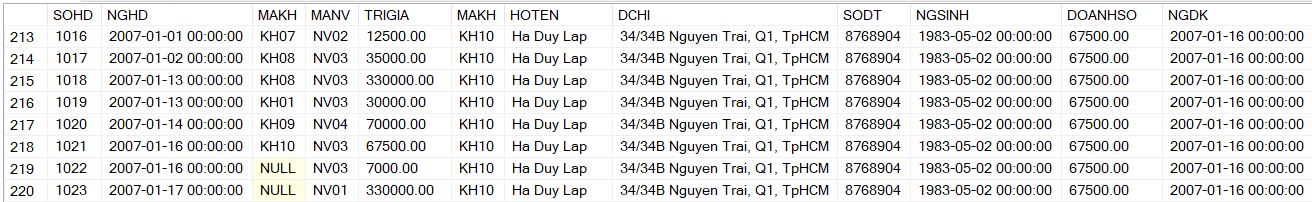


Hình 1‑20 Bảng khách hàng.

Giả sử: R: tất cả thuộc tính của bảng HOADON (hóa đơn).

S: tất cả thuộc tính của bảng KHACHHANG (khách hàng).

Tích của R và S là:



Hình 1‑21 Kết quả phép tích.

*Nhận xét:*

Bảng HOADON có 5 cột, 22 dòng.

Bảng KHACHHANG có 7 cột, 10 dòng.

Kết quả phép tích trên có 12 cột, 220 dòng.

*Nhận xét*: khi thực hiện phép tích, cần thêm điều kiện để giảm số dòng không có ý nghĩa. Cụ thể ví dụ: Thêm điều kiện khách hàng có mã khách hàng trùng với mã khách hàng trên hóa đơn.

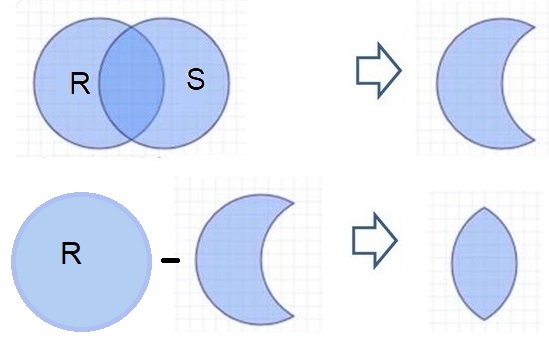


Hình 1‑22 Kết quả phép tích sau khi thêm điều kiện chọn.

Kết quả: mỗi dòng thể hiện thông tin khách hàng cùng với thông tin hóa đơn đúng với khách hàng.

Ngoài ra còn các phép toán mở rộng.

### Phép giao (Interesection ).



Hình 1‑23 Minh họa phép giao.

Hình thành một quan hệ mới chứa các bộ thuộc cả *R* và *S.*

Biểu diễn bằng các toán tử căn bản:

Bảng SANPHAM Bảng CTHD

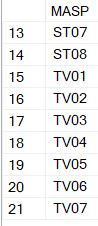
Hình 1‑24 Bảng CTHD.

Hình 1‑25 Bảng sản phẩm.

*Ví dụ*: R: Lấy MASP (mã sản phẩm) từ bảng SANPHAM (sản phẩm).

S: Lấy MASP từ bảng CTHD (chi tiết hóa đơn).

Phép giao giữa R và S là: (những sản phẩm chưa bán)



Hình 1‑26 Kết quả phép giao.

### **Phép Theta kết ().**

Hình thành một quan hệ mới chứa các bộ thỏa điều kiện *F* của tích descarte của hai quan hệ *R* và *S.*

*F* có dạng , là một trong các phép so sánh .

Nếu là dấu = thì phép trên được gọi là equi-join.

Một phép equi-join mà bỏ đi các cột trùng nhau thì gọi là natural-join.

Xét 3 ví dụ về 2 bảng HOADON và KHACHHANG.

Bảng HOADON (hóa đơn):



Hình 1‑27 Bảng HOADON.

Bảng KHACHHANG (khách hàng):



Hình 1‑28 Bảng KHACHHANG.

*Ví dụ 1*: R: tất cả thuộc tính của bảng HOADON (hóa đơn).

S: tất cả thuộc tính của bảng KHACHHANG (khách hàng).

Theta kết của R và S là:

Hoặc



Hình 1‑29 Kết quả phép tích sau khi thêm điều kiện chọn.

*Ví dụ 2*: Phép kết equi-join:

Bảng SANPHAM Bảng CTHD

Hình 1‑30 Bảng CTHD.

Hình 1‑31 Bảng sản phẩm.



Hình 1‑32 Bảng kết quả equijoin.

*Ví dụ 3*: Phép kết natural-join:



Hình 1‑33 Bảng kết quả natural-join.

### Phép kết ngoài (Outer-join).

* Left outer join:

Là phép kết bảo toàn các bộ của quan hệ bên trái, bộ nào không kết được với quan hệ bên phải thì điền null.

*Ví dụ*: Xét 2 bảng NHANVIEN và HOADON.

Bảng NHANVIEN



Hình 1‑34 Bảng nhân viên.

Bảng HOADON



Hình 1‑35 Bảng hóa đơn.

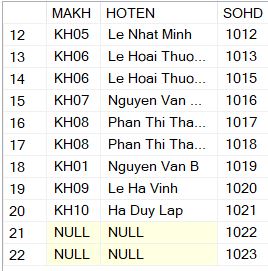


Hình 1‑36 Kết quả left outer join.

* Right outer join:

Là phép kết bảo toàn các bộ của quan hệ bên phải, bộ nào không kết được với quan hệ bên trái thì điền null.

*Ví dụ*:

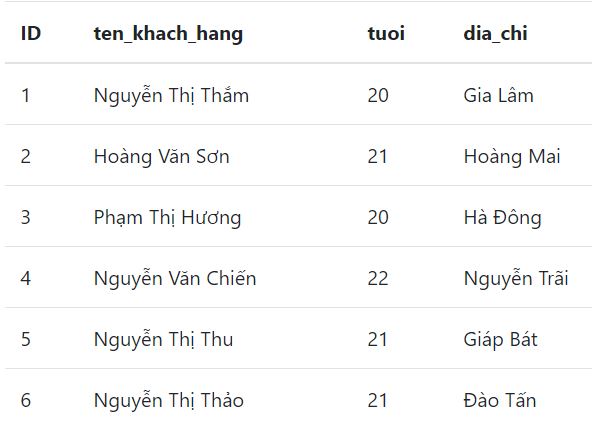


Hình 1‑37 Kết quả right outer join.

* Full outer join:

Là phép kết bảo toàn các bộ của hai quan hệ, bộ nào không kết được thì điền null.

*Ví dụ*: Bảng khách hàng.



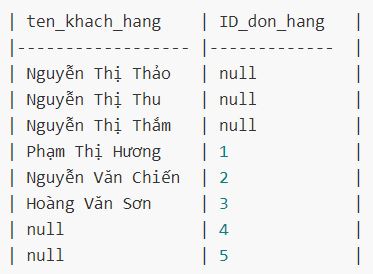
Hình 1‑38 Bảng khách hàng.

Bảng đơn hàng.



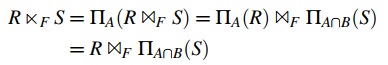
Hình 1‑39 Bảng đơn hàng.

Kết quả full outer join.

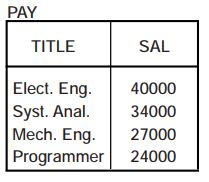


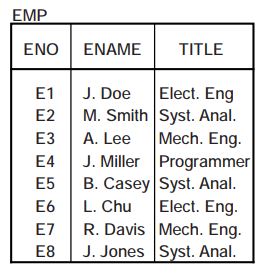
Hình 1‑40 Bảng kết quả full outer join.

### Phép kết nửa (semijoin).



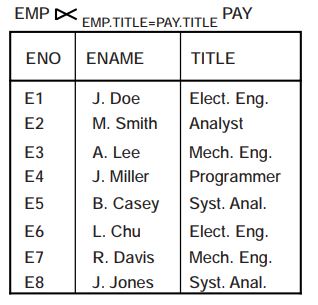
Hình thành quan hệ chứa các bộ của *R* có thể tham giá kết với quan hệ *S.*

*Ví dụ:*



Hình 1‑41 Bảng EMP.

Hình 1‑42 Bảng PAY.



Hình 1‑43 Kết quả phép semi join.

### Phép chia (Division ÷).

Tìm các bộ trong quan hệ *R* sao cho mỗi bộ đó có thể kết với tất cả các bộ trong quan hệ *S*.

*Z* là tập thuộc tính của *R*, *X* là tập thuộc tính của *S*. Để chia được thì . (cột và miền giá trị của X là con của Z).

## Phép tính quan hệ (Relational Calculus).

Ngôn ngữ đại số quan hệ (relational algebra languages) biểu diễn làm thế nào để có kết quả thì ngôn ngữ phép tính quan hệ (relational calculus languages) biểu diễn kết quả là những gì.

Ngôn ngữ phép tính quan hệ (relational calculus languages) chia thành 2 nhóm: phép tính quan hệ trên bộ (tuple relational calculus) và phép tính quan hệ trên miền (domain relational calculus).

### Phép tính quan hệ trên bộ (tuple relational calculus).

Biểu thức phép tính quan hệ trên bộ có dạng.

{ t | F(t) }

t là biến bộ: Biến nhận giá trị là một bộ của quan hệ trong CSDL.

F là công thức có liên quan đến t.

Phép tính quan hệ trên bộ là tìm những dòng t trong CSDL đúng với điều kiện F(t).

*Ví dụ 1*: Tìm các nhân viên có lương trên 30000.

*Ví dụ 2*: Cho biết mã và tên nhân viên có lương trên 30000.

*Ví dụ 3*: Cho biết các nhân viên (MANV) làm việc ở phòng ‘Nghien cuu’.

### Phép tính quan hệ trên miền (domain relational calculus).

: là các biến miền: nhận giá trị là một miền của thuộc tính.

là công thức theo .

*Ví dụ 1*: Cho biết mã và tên nhân viên có lương trên 30000.

NHANVIEN (HONV, TENLOT, TENNV, MANV, NGSINH, DCHI, PHAI, LUONG, MA\_NQL, PHG)

Đặt các thuộc tính lần lược là <p, q, r, s, t, u, v, x, y, z>.

*Ví dụ 2:* Cho biết các nhân viên (MANV) làm việc ở phòng ‘Nghien cuu’.

NHANVIEN (HONV, TENLOT, TENNV, MANV, NGSINH, DCHI, PHAI, LUONG, MA\_NQL, PHG)

Đặt các thuộc tính bảng NHANVIEN lần lượt là <p, q, r, s, t, u, v, x, y, z>.

PHONGBAN(TENPHG, MAPHG, TRPHG, NG\_NHANCHUC)

Đặt các thuộc tính bảng PHONGBAN lần lượt là <a,b,c,d>.

# TỔNG QUAN VỀ MẠNG MÁY TÍNH.

## Khái niệm.

Mạng máy tính (computer network) là hệ thống bao gồm nhiều hệ máy tính đơn lẻ (nút mạng) được kết nối với nhau theo kiến trúc nào đó và có khả năng trao đổi thông tin.

Các thành phần:

* Máy chủ (nút, các thiết bị đầu cuối)
* Switch
* Liên kết giao tiếp

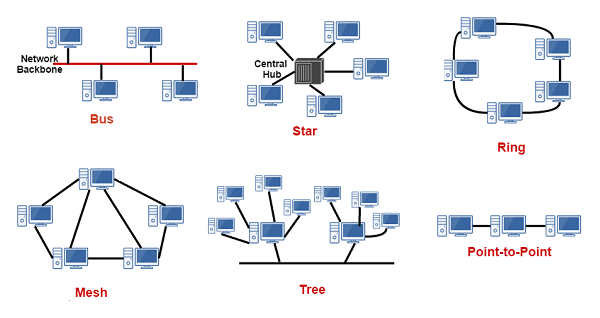
## Các loại mạng.

Theo quy mô (phân bố địa lý):

* Mạng rộng (WAN)
  + Khoảng cách giữa hai nút bất kỳ> 20km và có thể lên tới hàng nghìn km
  + Chậm trễ lâu do khoảng cách di chuyển
  + Tính không đồng nhất của phương tiện truyền dẫn
  + Tốc độ từ 150Mbps đến 10Gbps (OC192 trên đường trục)
* Mạng cục bộ (LAN)
* Giới hạn trong phạm vi địa lý (thường <2km)
* Tốc độ 10-1000 Mbps
* Độ trễ ngắn và tiếng ồn thấp
* Mạng khu vực đô thị (MAN)
* Giữa LAN và WAN

Theo cách bố trí phần tử của mạng:

* Bus
* Star
* Ring
* Mesh



Hình 2‑1 Topology.

## Các hình thức giao tiếp.

### Point-to-point (unicast)

* Một hoặc nhiều liên kết (trực tiếp hoặc gián tiếp) giữa mỗi cặp nút.
* Luôn luôn giao tiếp giữa hai nút.
* Người nhận và người gửi được xác định bằng địa chỉ của họ có trong tiêu đề thư.

Tin nhắn có thể đi theo một trong nhiều liên kết giữa người gửi và người nhận bằng cách sử dụng chuyển mạch hoặc định tuyến.

### Broadcast (đa điểm)

* Thông điệp được truyền qua một kênh chia sẻ và được tất cả các nút nhận.
* Mỗi nút kiểm tra địa chỉ và nếu nó không phải là người nhận dự định, hãy bỏ qua.
* Nhiều thiết bị: trường hợp đặc biệt.
* Thông báo được gửi đến một tập hợp con của các nút.

### Ngoài ra, còn các hình thức khác.

* Cặp xoắn
* Đồng trục
* Cáp quang
* Vệ tinh
* Microwave
* Không dây

## Các khái niệm về giao tiếp dữ liệu.

***Data communication:***

Truyền thông dữ liệu (Data communication) là 1 tập hợp các công nghệ hỗ trợ cho việc giao tiếp giữa 2 máy chủ (hosts).

Các máy chủ được kết nối bằng các liên kết (links), mỗi liên kết có thể có một hoặc nhiều kênh:

* Liên kết là một thực thể vật lý trong khi kênh là một thực thể logic.
* Các liên kết truyền thông có thể mang tín hiệu ở dạng số (digital form) hoặc dạng analog.

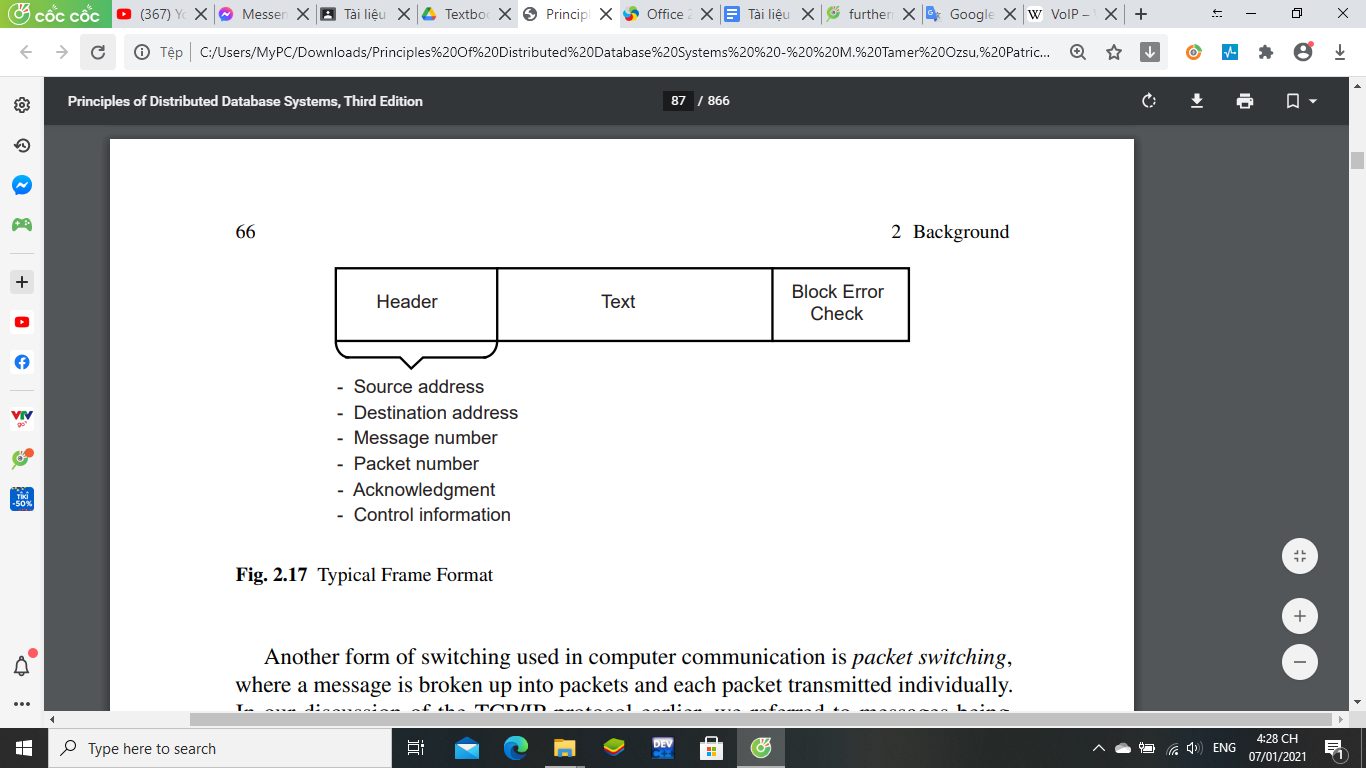
***Bandwidth:***

Mỗi kênh giao tiếp có 1 dung lượng nhất đinh, có thể được định nghĩa là lượng thông tin có thể được truyền qua kênh trong một đơn vị thời gian nhất định, thường được gọi là băng thông của kênh.

* Trong các kênh truyền dẫn analog, băng thông được định nghĩa là sự khác biệt (tính bằng hertz) giữa tần số thấp nhất và cao nhất có thể được truyền qua kênh mỗi giây.
* Trong các liên kết kỹ thuật số (digital), băng thông đề cập đến số lượng bit có thể được truyền mỗi giây (bps).

## Cấu trúc Frame.

Thông thường, giới hạn của kích thước của từng frame sẽ được thiết lập cho mỗi mạng, kèm theo những dữ liệu có vai trò kiểm soát việc truyền tải (the control informations) như địa chỉ nguồn (source addresses), địa chỉ đích (destinations), các mã kiểm tra lỗi khối (block error check codes),…



Hình 2‑1 Đinh dạng frame điển hình.

## Sơ lược chuyển mạch kênh và chuyển mạch gói.

1. ***Chuyển mạch kênh:***

Có nhiều hình thức chuyển mạch / định tuyến có thể được thực hiện trong mạng point-to-point. Cho phép tạo ra 1 kênh chuyên dụng riêng cho người gửi và người nhận. Được gọi là chuyển mạch gói (circuit switching), thường được sử dụng cho các trong các kết nối di động truyền thống.

1. ***Chuyển mạch gói:***

Một hình thức chuyển mạch khác được sử dụng trong computer communication là chuyển mạch gói (packet switching): Dữ liệu cần vận chuyển sẽ được chia thành các gói có kích thước (size) và định dạng (format) xác định, được truyền đi 1 cách riêng rẽ thông qua các đường truyền (route) khác nhau để đến được nơi nhận. Do đó, các gói tin này có thể đến nơi nhận tại các thời điểm khác nhau cũng như sẽ sắp xếp không theo thứ tự ban đầu.

Vì thế khi các gói tin đến đích, giao thức lớp transport (The transport layer protocol) sẽ chịu trách nhiệm đối chiếu và sắp xếp các gói tin theo đúng thứ tự để tạo nên dữ liệu ban đầu.

## Các giao thức giao tiếp.

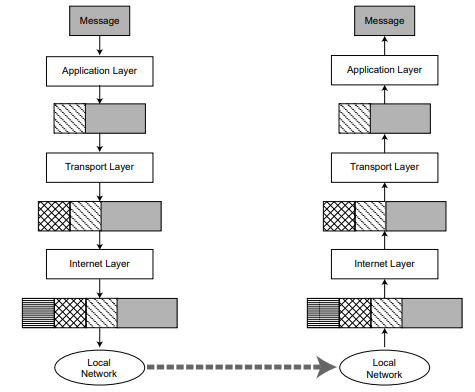
Thiết lập kết nối vật lý giữa hai máy chủ là không đủ vì không đảm bảo tính tin cậy, toàn vẹn dữ liệu, hiệu quả ,xảy ra lỗi… vì vậy yêu cầu xây dựng giao thức.

Giao thức mạng được phân lớp mỗi lớp một chức năng xác định bởi lớp dưới cung cấp cho lớp trên gọi là giao thức stack/ bộ giao thức.

Giao thức tiêu chuẩn đang thường dùng hiện nay là TCP/IP (Transport Control Protocol/Internet Protocol).

Lớp trên thêm 1 đoạn thông tin vào đầu gói tin rồi chuyển cho lớp dưới, lớp dưới lại thêm thông tin tiếp vào đầu gói tin để hỗ trợ cung cấp dịch vụ của nó.

Qua máy nhận sẽ làm ngược lại là lớp dưới đọc thông tin đầu gói tin đó và chuyển cho lớp trên.



Hình 2‑2 Truyền tin bằng TCP/IP.

1. ***Giao thức TCP/IP:***

TCP / IP gồm 2 bộ giao thức một bộ tầng vận chuyển một bộ tầng network.

Tầng vận chuyển giải quyết vấn đề mất dữ liệu ,băng thông, thời gian.

bộ giao thức tầng network cung cấp phương tiện đóng gói gói tin để truyền qua đường truyền vật lý.

Tầng vận chuyển cung cấp 2 giao thức là TCP và UDP.

* + Giao thức TCP hướng kết nối, đáng tin cậy, kiểm soát luồng nhưng không đảm bảo thời gian.
  + Giao thức UDP không hướng kết nối, không đáng tin cậy, không kiểm soát luồng, nhưng tốc độ truyền cao.

Bộ giao thức tầng network chứa thêm thông tin người nhận người gửi giúp xác định máy đích.

1. ***Các tầng giao tiếp khác.***

* Tầng ứng dụng cung cấp thông số kỹ thuật mà các ứng dụng phân tán phải tuân theo vd html cho web.
* Lớp dưới cùng đại diện cho mạng mỗi mạng có định dạng và giao thức truyền tải riêng.

Tiêu chuẩn hóa cho mạng LAN được dẫn đầu bởi Viện Kỹ sư Điện và Điện tử (IEEE), gọi là Tiêu chuẩn IEEE 802 gồm 3 lớp :

* + Lớp vật lý giải quyết các vấn đề truyền dữ liệu vật lý như tín hiệu.
  + Lớp kiểm soát truy cập phương tiện xác định ai có có quyền truy cập vào phương tiện truyền dẫn và khi nào.
  + Lớp điều khiển liên kết logic đảm bảo truyền gói tin giữa hai máy tính

Trong hầu hết các mạng LAN, giao thức lớp TCP và lớp IP được thực hiện trên đầu ba lớp này, cho phép mỗi máy tính có thể giao tiếp trực tiếp trên Internet.