**Mục lục**

[**1.** **Vấn đề** 2](#_Toc83787001)

[**2.** **Thiết kế từ trên xuống (top-down)** 2](#_Toc83787002)

[*Lưu ý làm BTL* 2](#_Toc83787003)

[*Lưu ý làm BTL* 3](#_Toc83787004)

[**3.** **Thiết kế từ dưới lên (Bottom-Up)** 4](#_Toc83787005)

[**4.** **Phân mảnh dữ liệu** 4](#_Toc83787006)

[*Lưu ý làm BTL* 4](#_Toc83787007)

[**5.** **Cấp phát** 5](#_Toc83787008)

***- by Team 8 CSDL Phân tán PTIT***

# **Vấn đề**

* Để quyết định bố trí dữ liệu, chương trình như thế nào. Thiết kế mạng ra sao.
* Bố trí liên gồm có:
  + Bố trí các phần mềm DBMS (Database Management System)
  + Bố trí các ứng dụng chạy trên CSDL
* Chiến lược thiết kế:
  + Tiếp cận từ trên xuống (top-down)
    - Thiết kế hệ thống từ đầu
    - Các hệ thống đồng nhất
  + Tiếp cận từ dưới lên (Bottom-up)
    - Các CSDL đã tồn tại ở một số vị trí
    - Các CSDL cần được kết nối để giải quyết các nhiệm vụ chung

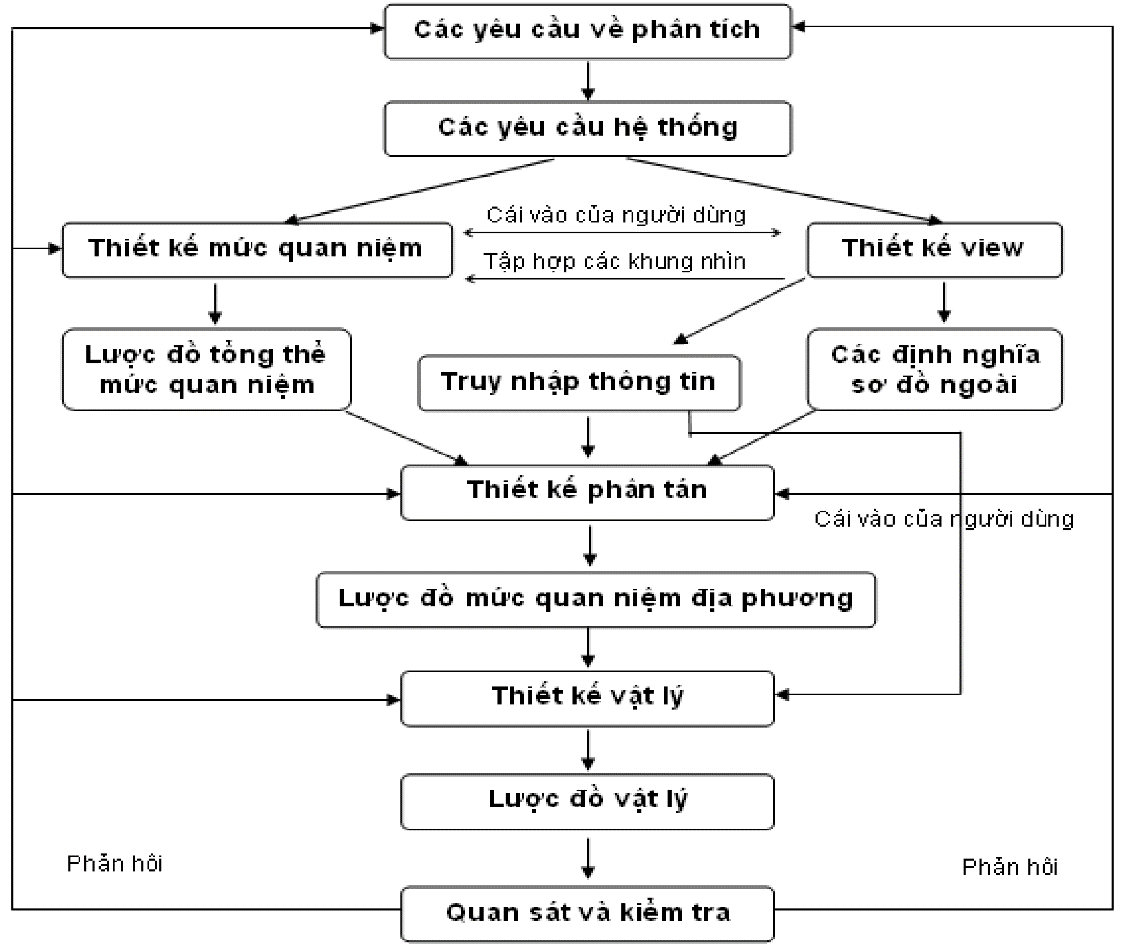
# **Thiết kế từ trên xuống (top-down)**

* Thiết kế từ tổng thể đến riêng biệt. Thiết kế từ một khối ban đầu, chia nhỏ thành các khối con

## *Lưu ý làm BTL*

*\*Trong chương trình học thầy yêu cầu thiết kế BTL theo kiểu này. Thực hiện thiết kế khối tổng thể như môn CSDL. Sau khi thiết kế khối tổng thể đó rồi mới phân chia ra các mảnh. Qúa trình thiết kế từ bước 1-2-3-4 là trên giấy, gần như giống nhau. Từ B5 là bước thiết kế vật lý thì mới phân mảnh dữ liệu rồi phân bổ nó đi đến các site.*

* Phân rã một hệ thống lớn thành các hệ thống con
  + Phân tích các yêu cầu nhằm định nghĩa môi trường hệ thống
  + Thu thập các yêu cầu về dữ liệu và nhu cầu xử lý của các trạm có sử dụng CSDL
* Mô hình:



* + **Thiết kế view**: xây dựng khung nhìn dữ liệu cho người sử dụng ở các trạm

## *Lưu ý làm BTL*

*\*Trong chương trình này thầy không yêu cầu thiết kế giao diện. Thầy chỉ yêu cầu chạy câu truy vấn ngay trên hệ quản trị CSDL, mỗi người khác nhau sẽ chạy các câu truy vấn khác nhau*

* + **Thiết kế mức quan niệm:** là một tiến trình kiểm tra và xác định rõ hai nhóm quan hệ: phân tích thực thể và phân tích chức năng
    - Phân tích thực thể: xác định các tập thực thể, các thuộc tính và các mối quan hệ giữa chúng
    - Phân tích chức năng: xác định các chức năng của hệ thống và đưa ra các chức năng cơ sở
  + **Thiết kế phân phân tán:** Bằng cách tách các quan hệ thành các quan hệ nhỏ (mảnh), đặt tại các vị trí (site), bao gồm hai phần:
    - Thiết kế phân đoạn (Fragmentation)
    - Thiết kế định vị và nhân bản (Allocation and Replication)
  + **Thiết kế lược đồ quan niệm địa phương**: tạo ra các lược đồ mức quan niệm tại các địa phương
  + **Thiết kế vật lý**: thực hiện ánh xạ lược đồ mức quan niệm tại các địa phương ra các đơn vị lưu trữ vật lý
  + **Quan sát và kiểm tra**: kiểm tra các giai đoạn của quá trình thiết kế cơ sở dữ liệu

# **Thiết kế từ dưới lên (Bottom-Up)**

* Có hiệu quả khi xây dựng một hệ thống mới. Nghĩa là bắt đầu từ những mảnh có sẵn, những dữ liệu có sẵn rồi kết nối lại thành một hệ thống tổng thể, thống nhất.
* Cách thiết kế:
  + Chọn một mô hình dữ liệu chung để mô tả lược đồ tổng thể
  + Chuyển mỗi lược đồ cục bộ theo mô hình dữ liệu chung đã chọn
  + Tích hợp các lược đồ cục bộ vào lược đồ tổng thể

# **Phân mảnh dữ liệu**

## *Lưu ý làm BTL*

*\*Trong BTL thì phải có “ít nhất” 1 bảng phân mảnh ngang và 1 bảng phân mảnh dọc*

* Là quá trình chia một quan hệ toàn cục thành nhiều mảnh có mối quan hệ logic với nhau
* Các loại phân mảnh  
  Một quan hệ thường được biểu diễn dưới dạng bảng. Phân mảnh một bảng thành nhiều bảng con
  + ***Phân mảnh dọc:*** chia các quan hệ theo chiều dọc. Nghĩa là thiết lập một tập các quan hệ mới từ 1 quan hệ gốc, mỗi quan hệ mới chỉ có một số thuộc tính từ quan hệ gốc. Thực chất đây là phép chiếu trên tập con các thuộc tính của quan hệ
  + ***Phân mảnh ngang:*** Quan hệ được chia theo chiều ngang. Thực chất là phép chọn quan hệ. Chọn những bộ của quan hệ thỏa mãn một biểu thức điều kiện cho trước
  + ***Phân mảnh hỗn hợp***
* Các quy tắc phân mảnh:
  + ***Tính đầy đủ***
    - R được phân rã thành các mảnh , …
    - Phân mảnh ngang: sao cho   
      Mục dữ liệu là các
    - Phân mảnh dọc: sao cho   
      Mục dữ liệu là các thuộc tính
    - Quy tắc này đảm bảo cho các mục dữ liệu trong R được ánh xạ hoàn toàn vào các mảnh và không bị mất
  + ***Tính phục hồi (Tái cấu trúc)***
    - Phân mảnh ngang: Quan hệ toàn cục phục hồi lại bằng cách hợp các quan hệ mảnh con
    - Phân mảnh dọc: Quan hệ toàn cục phục hồi bằng kết nối tự nhiên các quan hệ mảnh
    - Tính phục hồi đảm bảo quan hệ toàn cục phân rã không tổn thất thông tin
  + ***Tính tách biệt:***
    - Mỗi một của quan hệ toàn cục được chứa duy nhất trong một quan hệ con
    - Các thuộc tính của quan hệ con chỉ chung nhau thuộc tính khóa

# **Cấp phát**

* Cấp phát không nhân bản: mỗi mảnh nằm ở một vị trí
* Cấp phát có nhân bản:
  + Nhân bản toàn phần: Mỗi mảnh ở mọi vị trí
  + Nhân bản bán phần: Mỗi mảnh ở 1 số vị trí
* Luật cấp phát:
  + Nếu sẽ có lợi ích nhân bản
    - Nếu như tần suất sử dụng câu truy vấn Update > tần suất sử dụng câu truy vấn Select (C < 1) trên một mảnh nào đó, thì mảnh đấy không nên được nhân bản (số lượng nơi lưu trữ nhỏ nhất có thể),
    - Còn ngược lại Câu Update nhỏ hơn so với sử dụng câu Select thì mảnh đó có thể/nên được nhân bản
  + Trường hợp khác, nhân bản sẽ gây ra nhiều vấn đề
* Các yêu cầu thông tin cho cấp phát:
  + Thông tin CSDL (Thiết kế tổng thể, giúp chúng ta thiết kế từ B1->B4)
  + Thông tin ứng dụng (Giúp chúng ta phân mảnh, phân bổ phù hợp)
  + Thông tin kết nối (Kết nối mạng, truyền thông,,..)
  + Thông tin hệ thống máy tính (Hệ thống, dung lượng,…)