TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**ĐỒ ÁN CUỐI KÌ MÔN CƠ SỞ DỮ LIỆU PHÂN TÁN**

**THI TRẮC NGHIỆM**

*Người hướng dẫn*: **GV Trần Thị Hồng Nhung**

*Người thực hiện*: **NGUYỄN TRUNG TÍNH – 51603330**

**NGUYỄN TUẤN HUY – 51603144**

Lớp **: 16050302-16050301**

Khoá  **: 20**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2019**

TỔNG LIÊN ĐOÀN LAO ĐỘNG VIỆT NAM

**TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

**KHOA CÔNG NGHỆ THÔNG TIN**



**ĐỒ ÁN CUỐI KÌ MÔN CƠ SỞ DỮ LIỆU PHÂN TÁN**

**THI TRẮC NGHIỆM**

*Người hướng dẫn*: **GV Trần Thị Hồng Nhung**

*Người thực hiện*: **NGUYỄN TRUNG TÍNH – 51603330**

**NGUYỄN TUẤN HUY – 51603144**

Lớp **: 16050302-16050303**

Khoá  **: 20**

**THÀNH PHỐ HỒ CHÍ MINH, NĂM 2019**

LỜI CẢM ƠN

Trước tiên chúng em xin gửi lời cảm ơn cô Trần Thị Hồng Nhung đã tận tình giúp đỡ chúng em hoàn thành đề tài này.

Với vốn kiến thức hạn hẹp nên chúng em không tránh khỏi những sai sót. Chúng em rất mong nhận được ý kiến đóng góp, phê bình của cô. Chúng em sẽ cố gắng hoàn thiện kiến thức của mình, trau dồi thêm nhiều kiến thức mới hơn nữa. Đó sẽ là hành trang quý giá của chúng em sau này.

**ĐỒ ÁN ĐƯỢC HOÀN THÀNH**

**TẠI TRƯỜNG ĐẠI HỌC TÔN ĐỨC THẮNG**

Tôi xin cam đoan đây là sản phẩm đồ án của riêng chúng tôi và được sự hướng dẫn của GV Trần Thị Hồng Nhung;. 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.

Ngoài ra, trong đồ án còn sử dụng một số nhận xét, đánh giá cũng như số liệu của các tác giả khác, cơ quan tổ chức khác đều có trích dẫn và chú thích nguồn gốc.

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

*TP. Hồ Chí Minh, ngày tháng năm*

*Tác giả*

*(ký tên và ghi rõ họ tên)*

*Nguyễn Tuấn Huy*

*Nguyễn Trung Tính*

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

MỤC LỤC

[LỜI CẢM ƠN i](#_Toc8176026)

[PHẦN XÁC NHẬN VÀ ĐÁNH GIÁ CỦA GIẢNG VIÊN iii](#_Toc8176027)

[TÓM TẮT iv](#_Toc8176028)

[MỤC LỤC 1](#_Toc8176029)

[DANH MỤC CÁC BẢNG BIỂU, HÌNH VẼ, ĐỒ THỊ 3](#_Toc8176030)

[CHƯƠNG 1 – TỔNG QUAN VỀ CƠ SỞ DỮ LIỆU PHÂN TÁN 4](#_Toc8176031)

[1.1 Khái niệm 4](#_Toc8176032)

[1.2 Tính chất 4](#_Toc8176033)

[1.3 Thành phần hệ thống 4](#_Toc8176034)

[1.4 Ưu điểm và nhược điểm 5](#_Toc8176035)

[1.5 So sánh với CSDL tập trung: 6](#_Toc8176036)

[1.6 Phân loại 6](#_Toc8176037)

[1.7 Kiến trúc của một hệ CSDL phân tán 8](#_Toc8176038)

[1.7.1 Phương pháp phân mảnh 9](#_Toc8176039)

[CHƯƠNG 2 – GIỚI THIỆU BÀI TOÁN 10](#_Toc8176040)

[2. Bài toán: Thi trắc nghiệm 10](#_Toc8176041)

[CHƯƠNG 3 – THIẾT KẾ PHÂN TÁN DỮ LIỆU 10](#_Toc8176042)

[CHƯƠNG 4 – ỨNG DỤNG 10](#_Toc8176043)

[CHƯƠNG 5 – HƯỚNG DẪN CÁCH CHẠY DEMO 10](#_Toc8176044)

**DANH MỤC KÍ HIỆU VÀ CHỮ VIẾT TẮT**

**CÁC CHỮ VIẾT TẮT**

CSDL: Cơ sở dữ liệu

DANH MỤC CÁC BẢNG BIỂU, HÌNH VẼ, ĐỒ THỊ

**DANH MỤC HÌNH**

[Hình 1: Mô hình CSDL phân tán 4](file:///C:\Users\trung\Desktop\CSDL_Phan_Tan\CSDL_PHAN_TAN_DE_01.docx#_Toc8176020)

[Hình 2: Hệ quản trị CSDL phân tán 5](file:///C:\Users\trung\Desktop\CSDL_Phan_Tan\CSDL_PHAN_TAN_DE_01.docx#_Toc8176021)

[Hình 3: Mô hình hệ quản trị CSDL thuần nhất 7](file:///C:\Users\trung\Desktop\CSDL_Phan_Tan\CSDL_PHAN_TAN_DE_01.docx#_Toc8176022)

[Hình 4: Mô hình hệ quản trị CSDL không thuần nhất 8](file:///C:\Users\trung\Desktop\CSDL_Phan_Tan\CSDL_PHAN_TAN_DE_01.docx#_Toc8176023)

[Hình 5: Các mức trong suốt phân tán 8](file:///C:\Users\trung\Desktop\CSDL_Phan_Tan\CSDL_PHAN_TAN_DE_01.docx#_Toc8176024)

**DANH MỤC BẢNG**

[Bảng 3.1 Ví dụ cho chèn bảng 1](#_Toc387689363)

CHƯƠNG 1 – TỔNG QUAN VỀ CƠ SỞ DỮ LIỆU PHÂN TÁN

* 1. Khái niệm

Cơ sở dữ liệu phân tán: là một tập dữ liệu có quan hệ logic với nhau, được phân bổ trên các máy tính của một mạng máy tính.

Hệ quản trị cơ sở dữ liệu (CSDL) phân tán: là hệ thống phần mềm cho phép quản lý CSDL phân tán và đảm bảo tính trong suốt về sự phân tán đối với người dùng. Nói cách khác CSDL phân tán là một CSDL tập trung được phân tán vật lý nhưng được thống nhất là một CSDL duy nhất.

Hình 1: Mô hình CSDL phân tán

**Network**

**Trạm 3**

**Trạm 1**

**Trạm 2**

**Trạm 4**

* 1. Tính chất
* Tính chất phân tán: thực tế dữ liệu không lưu trữ trên cùng một trang (site) mà lưu trữ ở nhiều trạm thuộc một mạng máy tính, vì vậy có thể phân biệt CSDL phân tán với CSDL tập trung.
* Tương quan logic: từng loại dữ liệu đều mang một số tính chất ràng buộc lẫn nhau, như vậy ta có thể phân biệt được CSDL phân tán với CSDL địa phương (local) hoặc phân biệt những tệp lưu trữ trên nhiều trang khác nhau.
  1. Thành phần hệ thống

Để sự phân tán dữ liệu được trong suốt đối với người dùng, việc phân tán dữ liệu tại mỗi trang đòi hỏi phải có những thành phần sau:

* Thành phần quản trị CSDL (Database Management - DM)
* Thành phần truyền dữ liệu (Data Communication - DC)
* Từ điển dữ liệu (Data Dictionary - DD): thông tin về sự phân tán dữ liệu trên mạng.
* Thành phần CSDL phân tán (Distributed Database DDB)

Các dịch vụ cần có của hệ thống trên:

* Các ứng dụng truy cập CSDL từ xa
* Cung cấp các mức trong suốt phân tán
* Hỗ trợ điều khiển và quản trị CSDL
* Khả năng mở rộng giữa các hệ thống với nhau
* Khả năng điều khiển đồng thời và phục hồi các giao tác phân tán.

Hình 2: Hệ quản trị CSDL phân tán

**Network**

**Database-1**

**PC**

**PC**

**Trình quản lý ứng dụng**

**Database-2**

**PC**

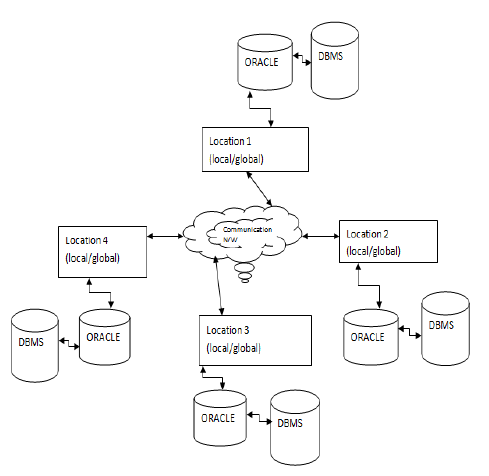
**PC**

**Trình quản lý ứng dụng**

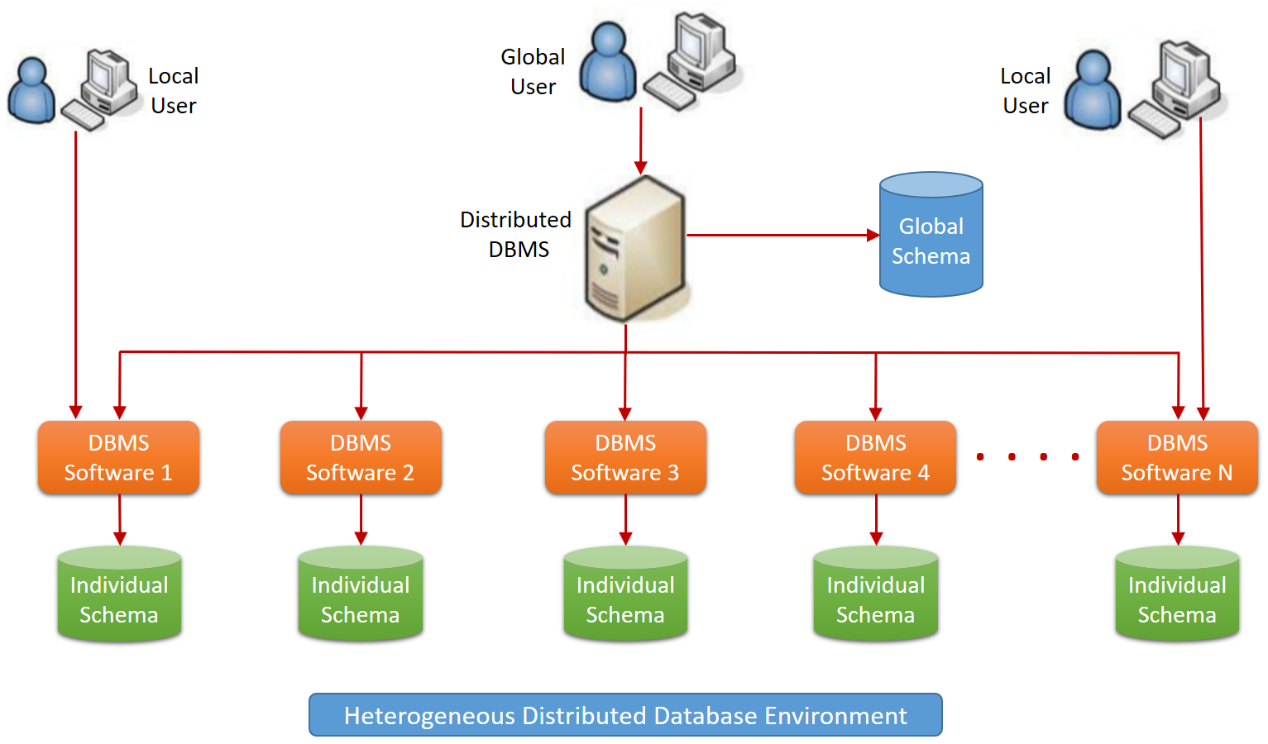
* 1. Ưu điểm và nhược điểm
* Ưu điểm:
* Phù hợp cấu trúc của một tổ chức
* Nâng cao khả năng chia sẻ tài nguyên
* Nâng cao tính sẵn sàng
* Nâng cao độ tin cậy
* Nâng cao hiệu năng
* Dễ mở rộng.
* Nhược điểm:
* Thiết kế CSDL phức tạp hơn
* Khó điều khiển được tính nhất quán của dữ liệu
* Khó phát hiện lỗi
* Giá thành cao
* Không có sẵn chuẩn
* Có vấn đề liên quan đến bảo mật.
  1. So sánh với CSDL tập trung:

|  |  |
| --- | --- |
| **Cơ sở dữ liệu tập trung** | **Cơ sở dữ liệu phân tán** |
| Dữ liệu tập trung tại DataCenter | Dữ liệu phân cấp trên các chi nhánh |
| Chi nhánh phải kết nối với DataCenter | Các chi nhánh kết nối với nhau |
| Dịch vụ trường lửa và cơ sở hạ tầng tại DataCenter phải đủ mạnh để phục vụ kết nối cho các chi nhánh | Dịch vụ tường lửa tự cung cấp tại mỗi chi nhánh |
| Các truy vấn bắt buộc phải thông qua DataCenter | Các thông tin không bắt buộc đi qua DataCenter |
| Chi phí thấp, dễ quản lý | Chi phí cao, khó quản lý |
| Xác xuất sự cố cao | Linh hoạt, tính sẵn sàng cao |

Bảng 1: Bảng so sánh CSDL tập trung và CSDL phân tán

* 1. Phân loại
* Hệ quản trị CSDL phân tán thuần nhất:
* Tất cả các nút cùng sử dụng một loại hệ quản trị CSDL
* Thuận lợi cho việc tăng trưởng và nâng cao hiệu năng
* Có thể có được bằng cách chia một CSDL thành một tập CSDL cục bộ

Hình 3: Mô hình hệ quản trị CSDL thuần nhất

* Hệ quản trị CSDL phân tán không thuần nhất:
* Các nút có thể thực hiện trên các hệ quản trị CSDL khác nhau
* Diễn ra khi các nút thiết đặt từng CSDL riêng
* Có thể có được bằng cách tích hợp những CSDL cục bộ đã có
  1. Kiến trúc của một hệ CSDL phân tán

Hình 4: Mô hình hệ quản trị CSDL không thuần nhất

Hình 5: Các mức trong suốt phân tán

**Sơ đồ tổng thể**

**Sơ đồ phân mảnh**

**Sơ đồ định vị**

**Sơ đồ ánh xạ địa phương 1**

**Sơ đồ ánh xạ địa phương 2**

**DBMS vị trí 1**

**DBMS vị trí 2**

**CSDL-1**

**CSDL-2**

* Sơ đồ tổng thể: xác định tất cả dữ liệu lưu trong CSDL phân tán
* Sơ đồ phân mảnh: mỗi quan hệ tổng thể có thể chia thành một vài phần nhỏ hơn không giao nhau được gọi là mảnh. Có nhiều cách để thực hiện việc chia mảnh này. Ánh xạ này là một – nhiều.
* Sơ đồ định vị: Các mảnh là các phần logic của một quan hệ tổng thể được định vị trên một hoặc nhiều vị trí vật lý trên mạng. Sơ đồ định vị xác định mảnh nào ở trạm nào.
* Sơ đồ ánh xạ địa phương: Ánh xạ các ảnh vật lý tới các đối tượng được các hệ quản trị CSDL địa phương thao tác tại các trạm. Ánh xạ này phụ thuộc vào hệ quản trị CSDL địa phương.
  + 1. Phương pháp phân mảnh
* Phân mảnh ngang
* Phân mảnh dọc
* Phân mảnh hỗn hợp

CHƯƠNG 2 – GIỚI THIỆU BÀI TOÁN – THIẾT KẾ PHÂN TÁN DỮ LIỆU

2.1 Bài toán

* Nội dung: Thi trắc nghiệm các môn học theo các trình độ khác nhau.
* Yêu cầu: Giả sử trường có 2 cơ sở chính : cơ sở 1 (CS1), cơ sở 2 (CS2). Phân tán cơ sở dữ liệu THI\_TN ra làm hai mảnh với điều kiện sau:
* THI\_TN được đặt trên server-1: chứa thông tin của các khoa thuộc cơ sở 1.
* THI\_TN được đặt trên server-2: chứa thông tin của các khoa thuộc cơ sở 2.
* Viết ứng dụng thực hiện một số việc sau trên từng cơ sở:

Nhập dữ liệu: môn học, sinh viên, khoa, lớp, giáo viên, đề, chuẩn bị thi

Phân quyền: gồm 5 nhóm (TRUONG-COSO1-COSO2-GIANGVIEN-SINHVIEN).

Hiện thực một số chức năng: đăng nhập, thi, xem kết quả, bảng điểm môn học.

2.2 Thiết kế phân tán dữ liệu

CHƯƠNG 3 – ỨNG DỤNG

CHƯƠNG 4 – HƯỚNG DẪN CÁCH CHẠY DEMO

**TÀI LIỆU THAM KHẢO**

**Tiếng Việt**

**Tiếng Anh**