

**ĐẠI HỌC THÁI NGUYÊN  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG**

**NGUYỄN THỊ LAN ANH**

**THIẾT KẾ CƠ SỞ DỮ LIỆU PHÂN TÁN ĐỂ TỔ CHỨC  
KHAI THÁC THÔNG TIN VỀ GIAO THÔNG**

**LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH**

Thái Nguyên - 2014

**ĐẠI HỌC THÁI NGUYÊN  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG**

**NGUYỄN THỊ LAN ANH**

**THIẾT KẾ CƠ SỞ DỮ LIỆU PHÂN TÁN ĐỂ TỔ CHỨC  
KHAI THÁC THÔNG TIN VỀ GIAO THÔNG**

**CHUYÊN NGÀNH: KHOA HỌC MÁY TÍNH**

**Mã số: 60 48 01 01**

**LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH**

**NGƯỜI HƯỚNG DẪN KHOA HỌC  
PGS. TS ĐỖ TRUNG TUẤN**

**Thái Nguyên - 2014**

## Lời cảm ơn

Để hoàn thành chương trình cao học và viết luận văn này, em đã nhận được sự giúp đỡ và đóng góp nhiệt tình của các thầy cô trường Đại học Công nghệ Thông tin và Truyền Thông, Đại học Thái Nguyên.

Trước hết, em xin chân thành cảm ơn các thầy cô trong bộ phận Đào tạo sau đại học, Đại học Công nghệ thông tin và Truyền thông, trường Đại học Thái Nguyên đã tận tình giảng dạy, trang bị cho em những kiến thức quý báu trong suốt những năm học qua. Em xin gửi lời biết ơn sâu sắc tới PGS. TS Đỗ Trung Tuấn đã dành rất nhiều thời gian và tâm huyết hướng dẫn, chỉ bảo em trong suốt quá trình thực hiện đề tài.

Xin chân thành cảm ơn gia đình, bạn bè đã nhiệt tình ủng hộ, giúp đỡ, động viên cả về vật chất lẫn tinh thần trong thời gian học tập và nghiên cứu.

Trong quá trình thực hiện luận văn, mặc dù đã rất cố gắng nhưng cũng không tránh khỏi những thiếu sót. Kính mong nhận được sự cảm thông và tận tình chỉ bảo của các thầy cô và các bạn.

## **Lời cam đoan**

Luận văn thạc sỹ này tôi nghiên cứu và thực hiện dưới sự hướng dẫn của PGS.TS Đỗ Trung Tuấn. Để hoàn thành bản luận văn này, ngoài các tài liệu đã liệt kê, tôi cam đoan không sao chép các công trình hoặc đồ án tốt nghiệp của người khác.

Tác giả

**NGUYỄN THỊ LAN ANH**

## MỤC LỤC

Lời cảm ơn .....	i
Lời cam đoan .....	iv
MỤC LỤC .....	v
Danh mục hình vẽ.....	vii
Danh mục các từ viết tắt.....	viii
MỞ ĐẦU .....	1
CHƯƠNG 1. ....	3
CƠ SỞ DỮ LIỆU PHÂN TÁN .....	3
1.1. Cơ sở dữ liệu tập trung.....	3
1.2. Cơ sở dữ liệu phân tán.....	3
1.2.1. Khái niệm cơ sở dữ liệu phân tán .....	3
1.2.3. Các hình thức tổ chức hệ thống phân tán .....	4
1.2.4. Ưu nhược điểm của hệ phân tán .....	6
1.2.5. Kiến trúc cơ bản của CSDL phân tán.....	6
1.3. So sánh cơ sở dữ liệu phân tán và cơ sở dữ liệu tập trung .....	9
1.4. Sự cần thiết của cơ sở dữ liệu phân tán.....	12
1.5. Kết luận chương.....	14
CHƯƠNG 2. ....	15
THIẾT KẾ CƠ SỞ DỮ LIỆU PHÂN TÁN .....	15
2.1. Các vấn đề về phân mảnh dữ liệu .....	15
2.1.1. Lý do phân mảnh .....	15
2.1.2. Các kiểu phân mảnh .....	16
2.1.3. Mức độ phân mảnh.....	16
2.1.4. Các quy tắc phân mảnh.....	17
2.1.5. Các kiểu cấp phát.....	17
2.1.6. Các yêu cầu thông tin.....	18
2.2. Các phương pháp phân mảnh .....	19
2.2.1. Phân mảnh dữ liệu ngang .....	19
2.2.2. Phân mảnh dọc .....	24
2.2.3. Phân mảnh hỗn hợp.....	27
2.3. Cấp phát.....	27
2.3.1. Bài toán cấp phát .....	27
2.3.2. Yêu cầu về thông tin .....	28
2.3.3. Mô hình cấp phát .....	29
2.4. Truy vấn trong cơ sở dữ liệu phân tán .....	31
2.4.1. Mục đích của xử lý truy vấn.....	31
2.4.2. Các tầng của quá trình xử lý truy vấn .....	32
2.5. Kỹ thuật thực hiện phép nối và phép hợp của đại số quan hệ.....	33
2.5.1. Phép chọn .....	35
2.5.2. Phép nối .....	36
2.5.3. Phân mảnh .....	38

2.7. Kết luận chương.....	43
CHƯƠNG 3. ....	44
THIẾT KẾ CƠ SỞ DỮ LIỆU VỀ CÁC CUNG ĐƯỜNG THÁI NGUYÊN.....	44
3.1. Phát biểu bài toán.....	44
3.2. Thiết kế cơ sở dữ liệu phân tán các cung đường.....	45
3.2.1. Thiết kế lược đồ quan hệ tổng thể .....	45
3.2.2. Thiết kế phân mảnh .....	52
3.2.3. Thiết kế định vị.....	54
3.2.4. Thiết kế sơ đồ ánh xạ địa phương.....	54
3.2.5. Thiết kế hệ thống mạng cho hệ thống.....	55
3.3. Cài đặt ứng dụng.....	56
3.3.1. Cấu hình các Server.....	56
3.3.2. Môi trường hoạt động .....	57
3.3.3. Công nghệ sử dụng.....	57
3.4. Giới thiệu chương trình.....	59
3.4.1. Mục tiêu mà hệ thống đạt được .....	59
3.4.2. Cấu trúc chương trình .....	59
3.4.3. Một số giao diện chính.....	60
3.5. Kết luận chương.....	61
KẾT LUẬN.....	62
Kết quả đạt được.....	62
Hướng nghiên cứu tiếp .....	62
Tài liệu tham khảo .....	63

## Danh mục hình vẽ

Hình 1.1. Mô hình CSDL phân tán .....	4
Hình 1.2. Mô hình Peer-to-peer .....	4
Hình 1.3. Mô hình File Server .....	5
Hình 1.4. Mô hình Client Server .....	5
Hình 1.5. Kiến trúc cơ bản của CSDL phân tán .....	6
Hình 1.6. Các phân mảnh và mô hình vật lý cho một quan hệ toàn cục .....	8
Hình 2.1 So sánh các phương pháp nhân bản .....	18
Hình 2.1. Lược đồ phân tầng tổng quát để xử lý truy vấn phân tán.....	33
Hình 2.2. Rút gọn phân mảnh ngang với phép chọn.....	35
Hình 2.3a. Cây đại số quan hệ truy vấn gốc .....	37
Hình 2.3b. Rút gọn phân mảnh ngang với phép kết nối .....	38
Hình 2.5. Rút gọn cho phân mảnh gián tiếp.....	42
Hình 2.6. Rút gọn phân mảnh hỗn hợp.....	43
Hình 3.1. Tình Thái Nguyên .....	44
Hình 3.2. Cần thiết an toàn giao thông .....	45
Bảng 1. Cấu trúc bảng HUYEN .....	46
Bảng 2. Cấu trúc bảng DUONG .....	47
Bảng 3. Cấu trúc bảng LOAIMADUONG .....	47
Bảng 4. Cấu trúc bảng KIEUDUONG .....	48
Bảng 5. Cấu trúc bảng TOCHUCGIAOTHONG .....	48
Bảng 6. Cấu trúc bảng MUCDOHUHONG.....	49
Bảng 7. Cấu trúc bảng LOAIBAOTRI .....	49
Bảng 8. Cấu trúc bảng DONVITHICONG .....	50
Bảng 9. Cấu trúc bảng THONGTINBAOTRI .....	51
Hình 3.3. Sơ đồ về mối quan hệ .....	51
Hình 3.3. Sơ đồ định vị của các mảnh tại các vị trí .....	54
Hình 3.3. Các mảnh và hình ảnh vật lý của một quan hệ tổng thể.....	55
Hình 3.3. Mô hình mạng của hệ thống quản lý các cung đường.....	56
Hình 3.7. Giao diện form quản lý thông tin cung đường .....	60
Hình 3.8. Giao diện form quản lý thông tin bảo trì.....	61
Hình 3.9. Giao diện form tìm kiếm cung đường .....	61

## Danh mục các từ viết tắt

DBMS	Database Management System
ADSL	Đường truyền bất đối xứng Asymmetric Digital Subscriber Line
C#	Ngôn ngữ lập trình
CPU	Central Processing Unit
CSDL	Cơ sở dữ liệu
CSDLPT	Cơ sở dữ liệu phân tán
DB2	Hệ quản trị cơ sở dữ liệu DB2 của IBM
DHM	Highway Development and Management System
ER	Entity-Relationship
ETL	<i>Extract Transform Load</i>
IMS	Information Management System
LAN	Local Area Network
Mainframe	Máy tính mainframe
MDX	Multidimensional eXpressions
MSIL	MicroSoft Intermediate Language
MSSQL	Microsoft SQL Server
NT	Network Technology
ORACLE	Hệ quản trị cơ sở dữ liệu ORACLE
Partition	Phân đoạn
PC	Máy vi tính, máy cá nhân
SQL	Structured Query Language
XML	Extensible Markup Language



## MỞ ĐẦU

Trong hệ thống kết cấu hạ tầng kinh tế - xã hội, hạ tầng giao thông vận tải đóng vai trò rất quan trọng trong việc tạo điều kiện đi lại thuận lợi cho nhân dân, giao thương hàng hóa, thúc đẩy tăng trưởng kinh tế, giảm nhẹ thiên tai, phòng thủ quốc phòng, cải thiện đời sống cho nhân dân, do đó phải ưu tiên phát triển. Nhiều cung đường liên tỉnh, liên huyện, liên xã được bảo trì, nâng cấp, nhiều cung đường mới được mở rộng. Vì vậy khối lượng thông tin về giao thông ngày càng lớn, phong phú và đa dạng, cần phải quản lý thông tin này một cách khoa học.

Hiện tại, Sở Giao thông Thái Nguyên quản lý các thông tin này theo cách thức sau: các đơn vị quản lý giao thông tại các cung đường thường xuyên có báo cáo về các Sở giao thông như: lưu lượng tham gia giao thông, số vụ tai nạn xảy ra, chất lượng các cung đường, công tác bảo dưỡng duy tu, giám sát cung đường,... Với và cách thức tổ chức đó sẽ tốn thời gian và dữ liệu có thể không đồng bộ.

Để khắc phục những nhược điểm trên cần áp dụng công nghệ thông tin vào trong quá trình quản lý. Vì vậy tôi đã chọn đề tài “*Thiết kế cơ sở dữ liệu phân tán để tổ chức khai thác thông tin về giao thông*” làm đề tài luận văn tốt nghiệp của mình.

### **Đối tượng và phạm vi nghiên cứu.**

- Đối tượng nghiên cứu của đề tài gồm: Nghiên cứu về CSDL phân tán, thiết kế CSDL phân tán như: các vấn đề về thiết kế phân tán, các phương pháp thiết kế phân tán, phương pháp phân mảnh, cấp phát cho các mảnh, ... tìm hiểu Các thông tin về giao thông. Ứng dụng lý thuyết thiết kế CSDL phân tán vào hệ thống. Ứng dụng khả năng quản trị CSDL phân tán của SQL Server và Visual Studio vào hệ thống.
- Trong phạm vi thực hiện của đề tài sẽ phân tích, thiết kế cơ sở dữ liệu các cung đường trên địa bàn tỉnh Thái Nguyên.

**Hướng nghiên cứu của đề tài :** Nghiên cứu và ứng dụng CSDL phân tán, tìm hiểu, thu thập các thông tin về giao thông. Từ những thông tin thu thập được phân tích, thiết kế CSDL phân tán để quản lý các thông tin đó

**Phương pháp nghiên cứu :** Để thực hiện được mục tiêu và nhiệm vụ đặt ra trong đề tài, tôi áp dụng hai phương pháp nghiên cứu đó là: phương pháp nghiên cứu lý thuyết và phương pháp nghiên cứu thực nghiệm. Đối với phương pháp nghiên cứu lý thuyết: tôi tiến hành nghiên cứu và thu thập các tài liệu liên quan đến CSDL phân tán, thiết kế CSDL phân tán, các công cụ có thể triển khai quản trị CSDL phân tán. Tiếp đến tôi thu thập các tài liệu liên quan đến thông tin về hệ thống các cung đường. Đối với phương pháp thực nghiệm: tôi phân tích yêu cầu thực tế của hệ thống và để xác định được các chức năng, quy trình hoạt động của hệ thống. Tiếp theo vận dụng cơ sở lý thuyết liên quan như CSDL phân tán, thiết kế CSDL phân tán vào thiết kế hệ thống, sử dụng công cụ SQL Server và Visual Studio để quản trị CSDL phân tán và công cụ thiết kế giao diện và cuối cùng đánh giá kết quả đạt được.

Ý nghĩa khoa học của đề tài : Việc ứng dụng CSDL phân tán để quản lý dữ liệu về các cung đường hiện nay là vấn đề quan trọng và cần thiết cho những người làm công tác quản lý giao thông vận tải có cái nhìn chính xác, và đưa ra quyết định kịp thời.

Bố cục của luận văn

Luận văn chia thành các chương:

1. Chương 1: Tổng quan về cơ sở dữ liệu phân tán
2. Chương 2: Thiết kế cơ sở dữ liệu phân tán
3. Chương 3: Thiết kế, xây dựng cơ sở dữ liệu các cung đường tại Thái Nguyên

Cuối luận văn là phần kết luận và tài liệu tham khảo.