

## CHƯƠNG TRÌNH HỌC PHẦN

### 1. Thông tin chung về học phần

Tên học phần:	Cơ sở dữ liệu phân tán
Mã học phần:	
Số tín chỉ:	3
Học phần tiên quyết:	Cơ sở dữ liệu, Mạng máy tính
Đào tạo trình độ:	Đại học
Giảng dạy cho ngành:	Công nghệ Thông tin
Bộ môn quản lý:	Bộ môn Hệ thống Thông tin
Phân bổ tiết giảng của học phần:	
- Nghe giảng lý thuyết:	30 tiết
- Làm bài tập trên lớp:	9 tiết
- Thảo luận:	6 tiết
- Thực hành, thực tập:	0
- Tự nghiên cứu:	90 tiết.

### 2. Mô tả tóm tắt học phần

Cung cấp kiến thức về nguyên lý thiết kế cơ sở dữ liệu phân tán, quản lý giao tác, điều khiển tương tranh và phục hồi dữ liệu... Trên cơ sở này, người học có thể nắm vững phương pháp thiết kế cơ sở dữ liệu phân tán, các thuật toán điều khiển tương tranh, phục hồi dữ liệu nhằm ứng dụng vào thực tế và nghiên cứu.

### 3. Chủ đề và chuẩn đầu ra của học phần

#### 3.1. Danh mục chủ đề của học phần

1. Phân biệt cơ sở dữ liệu tập trung và cơ sở dữ liệu phân tán
2. Nguyên lý thiết kế CSDL phân tán
3. Tối ưu hóa truy vấn CSDL phân tán
4. Quản lý giao tác
5. Điều khiển tương tranh
6. Khôi phục sự cố

#### 3.2. Chuẩn đầu ra của quá trình dạy - học từng chủ đề của học phần

**Chủ đề 1: Phân biệt cơ sở dữ liệu tập trung và cơ sở dữ liệu phân tán**

<b>Nội dung</b>	<b>Mức độ</b>
<p><b>Kiến thức</b></p> <ol style="list-style-type: none"> <li>1. Ưu nhược điểm của cơ sở dữ liệu trung.</li> <li>2. Đặc trưng của CSDL phân tán.</li> <li>3. Hệ quản trị CSDL phân tán.</li> <li>4. Kiến trúc của một Hệ CSDLPT</li> <li>5. Các mức trong suốt của CSDLPT</li> </ol>	<p>1</p> <p>1</p> <p>1</p> <p>1</p> <p>2</p>
<p><b>Thái độ</b></p> <ol style="list-style-type: none"> <li>1. Công nghệ mạng máy tính ngày càng được triển khai ứng dụng rộng khắp. Vì vậy, tiếp cận xử lý dữ liệu được phân tán trên mạng là một nhu cầu bức thiết.</li> <li>2. Tiếp cận xử lý CSDLPT giúp có thể giải quyết các bài toán lớn theo nguyên lý “chia để trị”.</li> <li>3. Tùy đối tượng tiếp cận CSDLPT, hệ thống sẽ cung cấp các mức trong suốt khác nhau nhằm đảm bảo tính thân thiện và bảo mật.</li> </ol>	
<p><b>Kỹ năng</b></p> <ol style="list-style-type: none"> <li>1. Phân biệt CSDL phân tán, CSDL tập trung.</li> <li>2. Nhận biết tính chất và đặc điểm của hệ CSDL phân tán:</li> <li>3. Xử lý dữ liệu ứng với các mức trong suốt khác nhau</li> </ol>	<p>2</p> <p>2</p> <p>3</p>

**Chủ đề 2: Nguyên lý thiết kế cơ sở dữ liệu phân tán**

<b>Nội dung</b>	<b>Mức độ</b>
<p><b>Kiến thức</b></p> <ol style="list-style-type: none"> <li>1. Tính trong suốt của CSDL phân tán</li> <li>2. Phân mảnh dữ liệu - Tính đúng của phân mảnh dữ liệu</li> <li>3. Phương pháp phân mảnh ngang : Logic hội vị từ</li> <li>4. Phương pháp phân mảnh dọc: Thuật toán BEA</li> <li>5. Định vị dữ liệu</li> </ol>	<p>2</p> <p>3</p> <p>3</p> <p>3</p> <p>2</p>
<p><b>Thái độ</b></p> <ol style="list-style-type: none"> <li>1. Thiết kế CSDLPT phụ thuộc nhiều yếu tố: tổ chức logic của CSDL, hệ thống mạng, nhu cầu ứng dụng .. Xác định kiến trúc, mục tiêu, tổ chức hệ CSDLPT đảm bảo tính đúng là rất cần thiết</li> <li>2. Các giải pháp thiết kế CSDLPT hiện nay có độ phức tạp NP-khó. Trên cơ sở các giải thuật cổ điển, cần kết hợp với các kỹ thuật mới nhằm giảm bớt thời gian xử lý.</li> </ol>	

<b>Kỹ năng</b>	
1. Kiểm tra tính đúng của phân mảnh dữ liệu	3
2. Phân mảnh dữ liệu	3
3. Định vị dữ liệu	3

**Chủ đề 3:** Tối ưu hóa truy vấn CSDL phân tán.

<b>Nội dung</b>	<b>Mức độ</b>
<b>Kiến thức</b>	
1. Cơ sở toán học của tối ưu hóa	2
2. Các qui tắc tối ưu hóa	3
<b>Thái độ</b>	
Cần có một chiến lược thích hợp đáp ứng các truy vấn nhanh nhất có thể trên môi trường đa người dùng ứng với một hệ CSDLPT. Việc tối ưu hóa có thể dựa trên cơ sở tối ưu hóa các biểu thức đại số quan hệ.	
<b>Kỹ năng</b>	
Tối ưu hóa một truy vấn	3

**Chủ đề 4:** Quản lý giao tác

<b>Nội dung</b>	<b>Mức độ</b>
<b>Kiến thức</b>	
1. Khái niệm giao tác	2
2. Lịch giao tác, giao tác tuần tự, khả tuần tự	2
3. Khả tuần tự xung đột.	3
4. Khả tuần tự khung nhìn.	3
<b>Thái độ</b>	
Giao tác là một đơn vị xử lý nguyên tố đảm bảo tính nhất quán dữ liệu, một sự thực hiện tuần tự của các giao tác sẽ đảm bảo tính nhất quán được bảo toàn	
<b>Kỹ năng</b>	
1. Xác định giao tác	2
2. Kiểm tra tính khả tuần tự của lịch	3

**Chủ đề 5: Điều khiển tương tranh.**

<b>Nội dung</b>	<b>Mức độ</b>
<b>Kiến thức</b> <ol style="list-style-type: none"> <li>1. Các Chủ đề khi truy xuất đồng thời</li> <li>2. Kỹ thuật khóa</li> <li>3. Kỹ thuật nhãn thời gian</li> <li>4. Kỹ thuật xác nhận tính hợp lệ</li> </ol>	<p>1</p> <p>3</p> <p>3</p> <p>3</p>
<b>Thái độ</b> <ol style="list-style-type: none"> <li>1. Tính khả tuần tự của lịch thực hiện được sinh ra bởi các giao tác thực hiện đồng thời có thể được đảm bảo thông qua một trong những cơ chế được gọi là điều khiển tương tranh.</li> <li>2. Các kỹ thuật điều khiển tương tranh ngày càng được bổ sung và khắc phục nhược điểm và tích hợp vào các hệ quản trị CSDL.</li> </ol>	
<b>Kỹ năng</b> <ol style="list-style-type: none"> <li>1. Nhận biết các tình huống tương tranh.</li> <li>2. Điều khiển tương tranh bằng các kỹ thuật: khóa, nhãn thời gian, xác nhận tính hợp lệ</li> </ol>	<p>2</p> <p>3</p>

**Chủ đề 6: Khôi phục sự cố**

<b>Nội dung</b>	<b>Mức độ</b>
<b>Kiến thức</b> <ol style="list-style-type: none"> <li>1. Phân loại sự cố</li> <li>2. Mục tiêu khôi phục sự cố</li> <li>3. Nhật ký giao tác</li> <li>4. Phương pháp khôi phục: Undo, Redo, Undo_Redo logging</li> </ol>	<p>1</p> <p>1</p> <p>1</p> <p>3</p>
<b>Thái độ</b> <p>Việc các giao tác cùng thực hiện trên cùng dữ liệu trong cùng khoảng thời gian, trên môi trường mạng có thể gây ra sự không nhất quán dữ liệu do sự cố hệ thống hay sự cố giao tác. Mục tiêu của khôi phục sự cố là đưa CSDL về trạng thái nhất quán sau cùng nhất trước khi gặp sự cố đảm bảo các giao tác nguyên tố, bền vững.</p>	
<b>Kỹ năng</b> <ol style="list-style-type: none"> <li>1. Xác định sự cố</li> <li>2. Khôi phục dữ liệu</li> </ol>	<p>2</p> <p>3</p>

#### 4. Phân bổ thời gian chi tiết

Chủ đề	Phân bổ số tiết cho hình thức dạy - học					Tổng
	Lên lớp			Thực hành, thực tập	Tự nghiên cứu	
	Lý thuyết	Bài tập	Thảo luận			
Chủ đề 1	2				5	
Chủ đề 2	4	1			10	
Chủ đề 3	2	1			5	
Chủ đề 4	8	2	2		20	
Chủ đề 5	7	3	2		30	
Chủ đề 6	7	2	2		20	

#### 5. Tài liệu

TT	Tên tác giả	Tên tài liệu	Năm xuất bản	Nhà xuất bản	Địa chỉ khai thác tài liệu
1	M. Tamer Ozsu and Patrick Valduriez	"Principles of Distributed Database Systems," Second Edition	1999	Prentice Hall	GV cung cấp
2	Ng Bá Tường	Nhập môn CSDL phân tán	2004	NXBKH&KT	Thư viện
3	Ng Kim Anh	Nguyên lý của các hệ CSDLPT	2004	NXB ĐHQG Hà Nội	Thư viện
4	Anand Rajaraman	Slide bài giảng)		<a href="http://www.stanford.edu/class/cs347">http://www.stanford.edu/class/cs347</a>	Web

#### 6. Đánh giá kết quả học tập

TT	Các chỉ tiêu đánh giá	Phương pháp đánh giá	Trọng số (%)
1	Tham gia trình bày	Chấm bài	20
2	Tham gia đóng góp bài, làm bài tập	Chấm bài.	10
3	Kiểm tra giữa kỳ	Viết	20
4	Thi kết thúc học phần	Viết	50

**TRƯỞNG KHOA**

**TRƯỞNG BỘ MÔN**  
(Ký và ghi họ tên)