

ĐỀ CƯƠNG MÔN HỌC

IS211 – CƠ SỞ DỮ LIỆU PHÂN TÁN

1. THÔNG TIN CHUNG

Tên môn học (tiếng Việt):	Cơ sở dữ liệu phân tán
Tên môn học (tiếng Anh):	Distributed Databases
Mã môn học:	IS211
Thuộc khối kiến thức:	Chuyên ngành.....
Khoa/Bộ môn phụ trách:	Hệ thống thông tin.....
Giảng viên biên soạn:	ThS. Thái Bảo Trân, TS. Ngô Thanh Hùng Email: tranbt@uit.edu.vn , hungnt@uit.edu.vn
Số tín chỉ:	4
Lý thuyết:	3
Thực hành:	1
Môn học trước:	Cơ sở dữ liệu, Hệ quản trị Cơ sở dữ liệu, Nhập môn mạng máy tính

2. MÔ TẢ MÔN HỌC

Sinh viên được trang bị các 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, giải quyết được vấn đề về quản lý giao dịch, đặc trưng và các tính chất giao dịch. Cũng như, hiểu được các thuật toán điều khiển tương tranh, phục hồi dữ liệu nhằm ứng dụng vào thực tế và nghiên cứu. Đồng thời vận dụng được kỹ thuật xử lý phân tán và cách triển khai CSDL phân tán bằng Oracle/MS SQL Server.

3. MỤC TIÊU MÔN HỌC

Bảng 1

Mục tiêu (Gx)	Mục tiêu môn học [1]	CĐR trong CTĐT [2]
G1	Mô tả được các khái niệm cơ bản về CSDL phân tán.	1.3.5, 2.1.1, 2.3.1, 2.3.2, 2.3.3, 2.5.4
G2	Nắm được phương pháp thiết kế các hệ cơ sở dữ liệu phân tán.	1.3.5, 2.1.1, 2.1.3, 2.1.4

G3	Hiểu rõ được các vấn đề về giao dịch, quản lý giao dịch đồng thời phân tán, vấn đề tương tranh và hiệu năng xử lý phân tán.	1.3.5, 2.1.1,
G4	Sử dụng được phương pháp thiết kế các hệ cơ sở dữ liệu phân tán vào bài toán thực tế.	4.3.1, 4.3.2
G5	Trình bày về thu thập các thông tin cho bài toán phân mảnh.	4.4.1
G6	Thiết kế được kiến trúc phân mảnh dữ liệu.	4.4.3
G7	Áp dụng được kỹ thuật xử lý phân tán và cách triển khai CSDL phân tán bằng Oracle/SQL Server.	5.4.5
G8	Trình bày tối ưu hóa một truy vấn phân tán.	4.6.1
G9	Biết cách xử lý dữ liệu ứng với các mức trong suốt khác nhau	4.4.1

4. CHUẨN ĐẦU RA MÔN HỌC

Bảng 2

CĐRMH [1]	Mô tả CĐRMH (Mục tiêu cụ thể) [2]	Mức độ giảng dạy [3]
G1.1 (1.3.5.1)	Mô tả được các khái niệm cơ bản về CSDL phân tán: - Xử lý dữ liệu phân tán; - CSDL phân tán, các đặc điểm của CSDL phân tán, kiến trúc hệ quản trị cơ CSDL phân tán, lợi ích của CSDL phân tán; - Các khái niệm cơ bản trong một hệ quản trị CSDL phân tán (như Oracle hoặc SQL Server).	T
G1.2 (2.1.1.3)	Xác định và mô tả được bài toán phân tán, chọn lựa kiến trúc phân tán, chọn lựa hệ quản trị CSDL phân tán;	U
G1.3 (2.3.1.2, 2.3.1.3)	Xác định các thành phần chính của hệ quản trị CSDL phân tán.	U
G1.4 (2.3.2.2, 2.3.2.3)	- Phân tích tác động giữa các thành phần của hệ quản trị CSDL phân tán; - Sử dụng tư duy hệ thống để giải thích thành phần và cơ chế hoạt động của hệ thống.	TU
G1.5 (2.3.3.1)	Xác định công cụ giúp tư duy về hệ thống quản trị CSDL phân tán.	U
G1.6 (2.5.4.1)	Tìm hiểu, phân tích một hệ quản trị CSDL phân tán	U

	(như Oracle hoặc SQL Server).	
G2.1 (1.3.5.2)	Hiểu được lý thuyết phân mảnh, giải thuật phân mảnh ngang, phân mảnh ngang dọc.	T
G2.2 (2.1.1.3, 2.1.3.1, 2.1.3.2, 2.1.4.3)	Thực hiện được việc: - Mô tả bài toán phân mảnh; - Liệt kê các giải pháp phân mảnh; - Ước lượng hiệu quả của giải pháp; - Tổng hợp, khuyến nghị kiến trúc phân mảnh CSDL;	U
G3.1 (1.3.5.3)	Nắm rõ về khái niệm giao dịch, xử lý tương tranh trong môi trường phân tán và lý thuyết giải quyết tương tranh;	T
G3.2 (2.1.1.3)	Xác định và mô tả được các bài toán giải quyết tương tranh trong truy vấn phân tán.	U
G4.1 (4.3.1.1, 4.3.1.2)	- Xác định nhu cầu phân tán CSDL cho tổ chức, doanh nghiệp (khách hàng); - Đề xuất kiến trúc phân tán CSDL và hệ quản trị CSDL phân tán phù hợp nhu cầu khách hàng	U
G4.2 (4.3.2.1, 4.3.2.2, 4.3.2.3, 4.3.2.3)	- Xác định các chức năng cần thiết của hệ thống CSDL phân tán; - Chọn lựa các kỹ thuật, công nghệ phân tán CSDL, xử lý phân tán, xử lý tương tranh, replicate dữ liệu; - Xác định mức độ phù hợp của công nghệ; - Xác định kiến trúc hệ thống CSDL phân tán.	U
G5.1 (4.4.1.1)	- Biết lựa chọn những yêu cầu đối với hệ thống CSDL phân tán từ các thông tin đã khảo sát	TU
G6.1 (4.4.3.1, 4.4.3.2, 4.4.3.3)	- Lựa chọn kỹ thuật phân mảnh thích hợp; - Thực hiện phân tích định lượng các phương án thiết kế; - Ước lượng phương án thiết kế phù hợp.	TU
G7.1 (4.5.3.1, 4.5.3.2)	- Cài đặt demo một hệ thống CSDL phân tán; Cài đặt demo một số xử lý truy vấn phân tán; Cài đặt demo một số cơ chế giải quyết tương tranh; - Mô tả việc kiểm tra các yêu cầu của hệ thống đã cài đặt	TU

G8.1 (1.3.5.3)	Hiểu được các khái niệm liên quan đến tối ưu hoá câu truy vấn	T
G8.1 (2.1.1.3)	Đề xuất giải pháp tối ưu hoá câu truy vấn	T
G8.3 (4.6.1)	Hiện thực yêu cầu tối ưu hoá câu truy vấn	TU
G9.1 (4.4.1)	Biết cách xử lý dữ liệu ứng với các mức trong suốt khác nhau	TU

5. NỘI DUNG, KẾ HOẠCH GIẢNG DẠY

a. Lý thuyết

Bảng 3

Buổi (3 tiết) [1]	Nội dung [2]	CDR MH [3]	Hoạt động dạy và học [4]	Thành phần đánh giá [5]
Buổi 1	Chương 1. Tổng quan về CSDL phân tán - Xử lý phân tán - Định nghĩa CSDL phân tán. - Các đặc điểm của CSDL phân tán so với CSDL tập trung. - Tại sao sử dụng CSDL phân tán. - Hệ quản trị CSDL phân tán. - Triển vọng của các hệ cơ sở dữ liệu phân tán. - Phân mảnh dữ liệu (Fragmentation) - Các loại truy xuất CSDL phân tán	G1.1 G1.2	Dạy: cho các ví dụ, đặt câu hỏi, thuyết giảng Học ở lớp: sử dụng tư duy và dựa trên các kiến thức đã học để trả lời câu hỏi, tham gia xây dựng bài học Học ở nhà: Ôn tập và tự tìm hiểu trước các khái niệm đã học	
Buổi 2	Chương 2. Kiến trúc của hệ quản trị CSDLPT - Kiến trúc tham khảo cho cơ sở dữ liệu phân tán - Các thành phần của hệ quản trị cơ sở dữ liệu phân tán - Kiến trúc của hệ quản trị cơ sở dữ liệu phân tán	G1.3 G1.4 G1.5 G1.6	Dạy: thuyết giảng, phân tích Học ở lớp: nghe giảng, phát biểu ý kiến Học ở nhà: ôn tập, đối chiếu với một hệ quản trị CSDL phân tán tiên tiến.	
Buổi 3	Chương 3. Thiết kế các hệ cơ sở dữ liệu phân tán - Giới thiệu - Mục tiêu của thiết kế CSDL phân tán - Các bước thiết kế CSDL phân tán - Các chiến lược phân tán dữ liệu - Các phương pháp thiết kế CSDL phân tán	G2.1 G2.2	Dạy: thuyết giảng, phân tích. Học ở lớp: nghe giảng, phát biểu ý kiến Học ở nhà: ôn bài	
Buổi 4	- Khái niệm phân mảnh và các quy tắc phân mảnh. - Lý thuyết phân mảnh ngang trong mô hình quan hệ	G2.1 G2.2 G4.1 G4.2	Dạy: thuyết giảng, phân tích Học ở lớp: nghe giảng, phát biểu ý kiến	

	- Lý thuyết phân mảnh dọc không tồn thất thông tin.		Học ở nhà: ôn bài, đọc trước phần thiết kế phân mảnh	
Buổi 5	- Thiết kế phân mảnh	G2.1 G2.2 G4.1 G4.2 G5.1 G6.1	Dạy: thuyết giảng, phân tích, làm bài tập mẫu Học ở lớp: nghe giảng, phát biểu ý kiến, làm bài tập. Học ở nhà: ôn bài, làm bài tập cá nhân (BTVN1)	BTVN1
Buổi 6	- Thiết kế phân mảnh (tt)	G2.1 G2.2 G4.1 G4.2 G5.1 G6.1	Dạy: thuyết giảng, phân tích, làm bài tập mẫu Học ở lớp: nghe giảng, phát biểu ý kiến, làm bài tập. Học ở nhà: ôn bài, làm bài tập cá nhân (BTVN1)	BTVN1
Buổi 7	- Thiết kế định vị dữ liệu - Cấp phát tài nguyên trong hệ phân tán	G2.1 G2.2 G4.1 G4.2 G5.1 G6.1	Dạy: thuyết giảng, phân tích, làm bài tập mẫu Học ở lớp: nghe giảng, phát biểu ý kiến, làm bài tập. Học ở nhà: ôn bài, làm bài tập cá nhân (BTVN1)	BTVN1
Buổi 8	Chương 4. Tính trong suốt phân tán - Tính trong suốt phân tán của ứng dụng chỉ đọc - Tính trong suốt phân tán đối với các ứng dụng cập nhật - Các nguyên tắc truy xuất cơ sở dữ liệu phân tán	G9.1	Dạy: thuyết giảng Học ở lớp: nghe giảng, phát biểu ý kiến Học ở nhà: ôn bài, làm bài tập cá nhân (BTVN2)	BTVN2
Buổi 9	Chương 5. Tối ưu hóa truy vấn phân tán - Khái niệm về truy vấn - Mục tiêu của truy vấn - Các phép toán đại số quan hệ - Đặc trưng của xử lý truy vấn - Phân lớp xử lý truy vấn	G8.1	Dạy: thuyết giảng, phân tích, làm bài tập mẫu Học ở lớp: nghe giảng, phát biểu ý kiến Học ở nhà: ôn bài	
Buổi 10	- Phân rã truy vấn - Cục bộ hóa dữ liệu phân tán - Tối ưu hóa truy vấn phân tán	G8.2	Dạy: thuyết giảng, phân tích, làm bài tập mẫu Học ở lớp: nghe giảng, phát biểu ý kiến, làm bài tập. Học ở nhà: ôn bài, làm bài tập cá nhân (BTVN3)	BTVN3

Buổi 11	- Các thuật toán tối ưu hóa truy vấn phân tán	G8.3	Dạy: thuyết giảng, phân tích, làm bài tập mẫu Học ở lớp: nghe giảng, phát biểu ý kiến, làm bài tập. Học ở nhà: ôn bài, làm bài tập cá nhân (BTVN3)	BTVN3
Buổi 12	Chương 6. Quản lý giao dịch và điều khiển đồng thời phân tán - Định nghĩa giao dịch - Đặc trưng của giao dịch - Các tính chất giao dịch	G3.1	Dạy: thuyết giảng, phân tích, làm bài tập mẫu Học ở lớp: nghe giảng, phát biểu ý kiến Học ở nhà: ôn bài	
Buổi 13	- Các loại giao dịch	G3.1	Dạy: thuyết giảng, phân tích, làm bài tập mẫu Học ở lớp: nghe giảng, phát biểu ý kiến, làm bài tập Học ở nhà: ôn bài, làm bài tập cá nhân (BTVN4)	BTVN4
Buổi 14	- Điều khiển đồng thời phân tán	G3.1 G3.2	Dạy: thuyết giảng, phân tích, làm bài tập mẫu Học ở lớp: nghe giảng, phát biểu ý kiến, làm bài tập. Học ở nhà: ôn bài, làm bài tập cá nhân (BTVN4)	BTVN4
Buổi 15 (3 tiết)	Ôn tập			

b. Thực hành

Bảng 4

Buổi (4 tiết)	Nội dung	CĐR MH	Hoạt động dạy và học	Thành phần đánh giá [5]
Buổi 1	Giới thiệu nội dung thực hành Ôn tập kiến thức về Oracle: - Tạo bảng, quan hệ, ràng buộc, truy vấn; - Trigger, Function, Procedure	G1.2 G1.6 G2.2 G3.2 G4.1 G7.1	Dạy: ôn tập câu lệnh, demo thao tác Học ở lớp: theo dõi, xem tài liệu và thực hành trên máy Học ở nhà: đọc thêm tài liệu, 2 bài tập lớn	
Buổi 2	- Phân quyền - Tạo database link	G1.2 G1.6 G2.2 G3.2	Dạy: ôn tập câu lệnh, demo thao tác Học ở lớp: theo dõi, xem tài liệu và thực	

		G4.1 G7.1	hành trên máy Học ở nhà: ôn tập, đọc thêm tài liệu, suy nghĩ về 2 bài tập lớn	
Buổi 3	Bài tập lớn 1 (bài tập nhóm): Xây dựng CSDL phân tán và cài đặt xử lý các truy vấn phân tán - Xây dựng mô hình dữ liệu, phân mảnh CSDL phù hợp với bài toán đặt ra;	G1.2 G1.6 G2.2 G4.1 G7.1 G8.1	Dạy: Theo dõi công việc của các nhóm; nhắc lại lý thuyết; hỗ trợ SV khi cần Học ở lớp: thực hiện phân tích, cài đặt hệ thống theo bài toán đã cho Học ở nhà: tìm hiểu tài liệu bổ sung cho các nội dung phân tích, cài đặt.	Bài tập lớn 1
Buổi 5	Bài tập lớn 1 (bài tập nhóm): Xây dựng CSDL phân tán và cài đặt xử lý các truy vấn phân tán (tt) Xây dựng ứng dụng, cài đặt xử lý truy vấn phân tán trên hệ thống.	G1.2 G1.6 G2.2 G4.1 G7.1 G8.1 G8.2 G8.3	Dạy: Theo dõi công việc của các nhóm; nhắc lại lý thuyết; hỗ trợ SV khi cần Học ở lớp: thực hiện phân tích, cài đặt hệ thống theo bài toán đã cho Học ở nhà: tìm hiểu tài liệu bổ sung cho các nội dung phân tích, cài đặt. Hoàn thiện báo cáo về bài tập lớn 1.	Bài tập lớn 1
Buổi 6	Bài tập lớn 2 (bài tập nhóm): Xây dựng CSDL phân tán và cài đặt xử lý tương tranh giữa các giao tác - Xây dựng mô hình dữ liệu, phân mảnh CSDL phù hợp với bài toán đặt ra;	G1.2 G1.6 G2.2 G3.1 G4.1 G7.1	Dạy: Theo dõi công việc của các nhóm; nhắc lại lý thuyết; hỗ trợ SV khi cần Học ở lớp: thực hiện phân tích, cài đặt hệ thống theo bài toán đã cho Học ở nhà: tìm hiểu tài liệu bổ sung cho các nội dung phân tích.	Bài tập lớn 2
Buổi 7	Bài tập lớn 2 (bài tập nhóm): Xây dựng CSDL phân tán và cài đặt xử lý tương tranh giữa các giao tác (tt) - Xây dựng ứng dụng, cài đặt giả lập các giao tác gây tương tranh.	G1.2 G1.6 G2.2 G3.2 G4.1 G7.1	Dạy: Theo dõi công việc của các nhóm; nhắc lại lý thuyết; hỗ trợ SV Học ở lớp: thực hiện phân tích, cài đặt hệ thống theo bài toán đã cho Học ở nhà: tìm hiểu tài liệu bổ sung cho các nội dung phân tích, cài đặt.	Bài tập lớn 2
Buổi 8	Bài tập lớn 2 (bài tập nhóm): Xây	G1.2	Dạy: Theo dõi công việc	Bài tập lớn

	dựng CSDL phân tán và cài đặt xử lý tương tranh giữa các giao tác (tt) - Xây dựng ứng dụng, cài đặt xử lý tương tranh giữa các giao tác.	G1.6 G2.2 G3.2 G4.1 G7.1	của các nhóm; nhắc lại lý thuyết; hỗ trợ SV Học ở lớp: thực hiện phân tích, cài đặt hệ thống theo bài toán đã cho Học ở nhà: tìm hiểu tài liệu bổ sung cho các nội dung phân tích, cài đặt. Hoàn thiện báo cáo về bài tập lớn 2.	2
--	----------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------

6. ĐÁNH GIÁ MÔN HỌC

Bảng 5

Thành phần đánh giá	CĐRMH	Tỷ lệ (%)
A1. Thực hành: Bài tập lớn 1,2 (bài tập nhóm) + bài tập cá nhân về nhà (BTVN1, BTVN2, BTVN3 và BTVN4)	G1, G2, G3, G5, G6, G7, G8, G9	30%
A2. Thi giữa kỳ (tự luận)	G1.4, G2.1, G9.1	20%
A4. Thi lý thuyết cuối kỳ (tự luận)	G2, G3, G8	50%

7. QUY ĐỊNH CỦA MÔN HỌC

- SV dành nhiều thời gian để chủ động trong việc tự học và tự tìm hiểu thêm các tài liệu liên quan dưới sự hướng dẫn của GV.
- Thực hiện các bài tập nhóm (nhóm tối đa 4 SV), bài tập nhóm để phát triển khả năng làm việc nhóm và trình bày.
- Sinh viên vắng quá 30% số buổi học trên lớp sẽ không được tham dự thi lý thuyết.
- SV chủ động ôn tập, làm bài tập về nhà, bài tập lớn và nộp đúng thời hạn yêu cầu. Nộp bài trễ hạn 1 ngày bị trừ 0.25 điểm, trễ hơn 1 ngày xem như không nộp bài tập.

8. TÀI LIỆU HỌC TẬP, THAM KHẢO

1. Hector Garcia-Mollina, Jeffrey D. Ullman, Jennifer Widom. *Database Systems: The Complete Book*. Prentice Hall.
2. Elmasri & Navathe. *Fundamentals of database systems*. Pearson Education, Inc, 2004.
3. Abraham, Silberschatz, Henry F.Korth, S. Sudarshan. *Database System Concepts*. McGraw-Hill, 2002.
4. Thomas Connolly, Carolyn Begg. *Database Systems*. Addison Wesley, 2002.

9. PHẦN MỀM HAY CÔNG CỤ HỖ TRỢ THỰC HÀNH

HQTCSDL thương mại (Microsoft SQL Server, Oracle, DB2, MySQL...).

Trưởng khoa/ bộ môn

Giảng viên

ThS. Thái Bảo Trân