Министерство образования и науки РФ

Государственное образовательное учреждение высшего профессионального образования

**«Санкт-Петербургский государственный электротехнический университет “ЛЭТИ” имени В.И. Ульянова (Ленина)» (СПбГЭТУ)**

РАБОЧАЯ ПРОГРАММА

дисциплины

*«Технологии распределенных баз данных и знаний»*

Для подготовки магистров по направлению

230100.68 *«Информатика и вычислительная техника»*

Санкт-Петербург

2009

Санкт-Петербургский государственный электротехнический

университет “ЛЭТИ”

"УТВЕРЖДАЮ"

проректор учебной работе

Лысенко Н.В.

"\_\_\_\_\_\_"\_\_\_\_\_\_\_\_\_\_\_\_\_2009 г.

РАБОЧАЯ ПРОГРАММА

дисциплины

*«Технологии распределенных баз данных и знаний»*

Для подготовки магистров по направлению

230100.68 *«Информатика и вычислительная техника»*

Уч. план №№ 231, 233, 234, 235

Факультет компьютерных технологий и информатики

Кафедра автоматизированных систем обработки информации и управления

Курс – 1

Семестр – 1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Лекции | 54 ч. |  | Экзамен | 1 семестр |
|  |  |  |  |  |
| Практические занятия | 18 ч. |  |  |  |
|  |  |  |  |  |
| Курсовое проектирование | ч. |  | Курсовое проектирование | 1 семестр |

|  |  |
| --- | --- |
| Аудиторные занятия | 72 ч. |
| Самостоятельные занятия | 72 ч. |
| Всего часов | 144 ч. |

2009 г.

Рабочая программа обсуждена на заседании кафедры автоматизированных систем обработки информации и управления “\_\_\_\_”\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_2009 г., протокол №\_\_\_\_\_\_.

Рабочая программа составлена в соответствии с государственным образовательным стандартом для подготовки магистров по направлению

230100.68 — «Информатика и вычислительная техника»

Дисциплина «Технологии распределенных БД и БЗ» преподается **на основе ранее изученных дисциплин программы подготовки бакалавров** и **является фундаментом для изучения последующих дисциплин:**

1) Управление информационными ресурсами предприятия.

Рабочая программа одобрена методической комиссией факультета КТИ “\_\_\_\_\_\_\_”\_\_\_\_\_\_\_\_\_\_\_\_\_2009 г.

АННОТАЦИЯ ДИСЦИПЛИНЫ

Дисциплина обеспечивает теоретическую и практическую подготовку в области принципов организации и функционирования распределенных баз данных и баз знаний.

Данная дисциплина служит фундаментом для изучения ряда специальных дисциплин, посвященных функционированию и проектированию информационных систем.

**DISCIPLINE ANNOTATION**

The discipline gives background theoretical knowledge and practical user technologies of date-bases and bases of knowledge.

This discipline is a prerequisite for interdiscipline project and master dissertation devoted for information system development.

Цели и задачи дисциплины

1. Освоение методики разработки математических и физических моделей исследуемых объектов, процессов и явлений, относящихся к области технологий распределенных БД и БЗ.
2. Освоение методов проектирования распределенных БД и БЗ, развитие способности к восприятию и разработке новых способов проектирования распределенных БД и БЗ.
3. Получение навыков организации информационных взаимодействий в распределенных БД и БЗ.

Требования к результатам освоения дисциплины

В результате освоения дисциплины студенты должны:

1. знать: классификацию и общие принципы построения систем хранения информации; теоретические основы организации и проектирования распределенных БД; принципы построения приложений БД, ориентированных на аналитическую обработку информации. принципы построения систем, основанных на использовании знаний; технологии доступа к БД в приложениях.
2. уметь: проектировать и применять CASE-средства и программные среды, использующие для разработки приложений БД и БЗ, управлять распределенными СУБД и обеспечивать управление репликацией данных в распределенных средах;
3. владеть: навыками применения SQL при разработке приложений БД; навыками проектирования БД на основе принципов нормализации и использования семантических диаграммных моделей.

Содержание рабочей программы

«Технологии распределенных БД и БЗ»

**Тема 1. Введение.**

Классификация и общие принципы построения систем хранения информации.

**Тема 2. Использование БД и БЗ в информационных системах.**

Корпоративные информационные системы. Производственные системы. Хранилища данных. Системы оперативной аналитической обработки данных. Системы поддержки принятия решений. Информационные системы, ориентированные на знания.

**Тема 3. Архитектурный подход к построению БД и БЗ в информационных системах.**

Обзор методологий разработки систем БД и БЗ, в том числе в распределенных средах.

**Тема 4. Проектирования БД на основе принципов нормализации и использования семантических диаграммных моделей.**

Классический подход к проектированию реляционных баз данных. Этапы разработки базы данных. Критерии оценки качества модели данных. Понятие функциональных и прочих зависимостей. Проектирование реляционных баз данных на основе принципов нормализации. Анализ критериев оценки для нормализованных и ненормализованных моделей данных.

Концептуальные модели и схемы баз данных. Диаграммное представление.

**Тема 5. CASE-средства и программные среды, использующие для разработки приложений БД и БЗ.**

CASE-системы для проектирования информационных систем. Назначение и разновидности CASE-систем. Примеры и различия распространенных CASE-систем.

**Тема 6. Технологии манипулирования данными в системах построенных на базе реляционных БД. Применения SQL.**

Язык SQL. Средства определения схемы – DDL (Data Definition Language). Операторы манипулирования данными DML – (Data Manipulation Language). Операторы защиты и управления данными.

**Тема 7. Технологии, обеспечивающие организацию и построение ИС РБД и БЗ**

Общая классификация архитектур построения приложений распределенных БД и БЗ. Технологические средства построения ИС в архитектуре «клиент-сервер», многозвенные приложения. Системы управления базами данных.

**Тема 8. Теоретические основы организации и проектирования распределенных БД.**

Подходы к построению распределенных БД. Безопасность данных. Целостность данных. Понятие транзакции.

**Тема 9. Архитектура систем поддержки принятия решений.**

Концепция хранилищ данных. Отличия хранилищ данных от систем OLTP. Витрины данных. «Добыча» данных.

**Тема 10. Архитектура приложений БД, ориентированных на аналитическую обработку информации.**

Концепция OLAP-систем. Связь OLAP и ХД. Структура информационно-аналитической системы и место OLAP в ней. Многомерная модель данных.

**Тема 11. Архитектура систем, основанных на использовании знаний.**

Логический вывод. Продукционные правила.

**Тема 12. Параллелизм и распределенные среды в приложениях БД**

Понятие распределенной среды применительно к построению систем БД и БЗ. Организация параллельных запросов, операций и команд.

**Тема 13. Управление параллельным сервером.**

Преимущества использования технологии Parallel Server.Применение технологий Parallel Server в бизнес-задачах.

**Тема 14.Управление распределенными СУБД.**

Типы распределенных СУБД. Использование распределенных СУБД. Распределенные транзакции.

**Тема 15.Управление репликацией данных в распределенных средах.**

Понятие репликации. Типы распространения данных. Типы репликации. Разрешение конфликтов. Инструменты организации репликации.

#### **Перечень практических занятий**

|  |  |  |
| --- | --- | --- |
| № | Наименование темы занятия | Номер темы программы |
| 1 | Проектирования БД на основе принципов нормализации и использования семантических диаграммных моделей. | 4 |
| 2 | CASE-средства и программные среды, использующие для разработки приложений БД и БЗ. | 5 |
| 3 | Технологии манипулирования данными в системах построенных на базе реляционных БД. Применения SQL. | 6 |
| 4 | Технологии, обеспечивающие организацию и построение ИС РБД и БЗ | 7 |
| 5 | Теоретические основы организации и проектирования распределенных БД. | 8 |
| 6 | Архитектура приложений БД, ориентированных на аналитическую обработку информации. | 10 |

**Цели и содержание курсового проекта**

и его ориентировочная трудоемкость

1. Освоение подходов к проектированию распределенных баз данных с использованием диаграммных моделей и средств автоматизации проектирования.
2. Овладение практическими навыками применения SQL при разработке приложений в архитектуре клиент-сервер.
3. Знакомство с теоретическими и практическими аспектами современных технологий разработки приложений баз данных с архитектурой клиент-сервер.

## Используемое программное обеспечение

* FireBird – Сервер БД
* CA ERWin 4.0 - инструментальная среда проектирования БД
* Rational Rose – инструментальная среда объектно-ориентированного анализа и проектирования программных систем на основе языка UML.
* IBExpert – среда администрирования, управления и разработки БД для InterBase.
* Borland Delphi – интегрированная среда разработки приложений.

## Входные данные для курсовой работы

* БД в формате InterBase (файл с расширением gdb), содержащая набор таблиц.
* Краткое описание структуры и задач предметной области, для решения которых необходимо использовать предлагаемую БД.
* Дополнительная информация о предметной области, полученная в результате интервью с «заказчиком» (преподавателем).

## Порядок выполнения

### Проектирование модели предметной области с использование UML.

(2 часа)

* Изучить возможности анализа и проектирования информационных систем в среде Rational Rose с использованием UML.
* Построить семантическую модель предметной области по заданию на проектирование с использованием диаграммы Rational Rose.
* **Контрольная точка № 1.**

### Проектирование модели БД с использованием ERWin.

(2 часа)

* Разработать логическую и физическую модели данных в соответствии с задачами предметной области. В модель добавляются связи, ключи, другие ограничения целостности. Модель должна обеспечивать поддержку целостности, неизбыточности и непротиворечивости данных.
* Всем элементам модели дать описания в среде ERWin.
* Сформировать отчет по логической модели средствами ERWin.
* Получить SQL (DDL)-скрипт структуры БД с использованием ERWin.
* С помощью комментариев описать все значимые элементы полученного скрипта.
* **Контрольная точка № 2.**

### Проектирование хранилища данных

(2 часа)

* Разработать хранилище данных на основе данных OLTP.

### Моделирование архитектуры системы с учетом ее использования в распределенной среде

(2 часа)

* С использованием UML (диаграмма компонентов, диаграмма размещения и др.) смоделировать узлы распределенной БД и способы репликации данных.
* **Контрольная точка № 3.**

### Создание БД по физической модели с использованием IBExpert.

(2 часа)

* Создать БД по скрипту в IBExpert.
* Добавить в БД генераторы и триггеры на автоинкрементные поля.
* Добавить в БД хранимые процедуры и триггеры для подержки бизнес-логики предметной области.
* Наполнить БД тестовыми данными.
* Разработать набор типовых SQL-запросов к БД. Рассмотреть запросы на получение, редактирование и удаление данных.
* Проанализировать производительность выполнения запроса средствами IBExpert. Рассмотреть необходимость создания дополнительных индексов для таблиц БД. Создать индексы.
* На основе полученных SQL-запросов создать «Представление (View)».
* **Контрольная точка № 4.**

### Доступ к БД из приложений. Организация репликации данных

1. часа)

* С использованием Delphi создать простое приложение, обеспечивающее работу с БД. Должные быть представлены возможности по просмотру, изменению, удалению данных в БД для нескольких связанных таблиц БД.
* С использованием средств сервера БД организовать процесс репликации данных.

### Тестирование и отладка системы. Оформление работы. Защита работы.

(4 часа.)

* **Защита.**

**Распределение учебных часов по темам, видам занятий**

**и видам самостоятельной работы**

| № темы | Название разделов и тем | **Объем учебных часов** | | | | | | Семестр |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Лекции | Лабор. занят. | Практ. занят. | Аудит. занят. | Самост. Работа | **Всего** | **Литература по темам** |
|
| 1 | Классификация и общие принципы построения систем хранения информации. | 2 |  |  | 2 |  | 2 | 1 | Л1,Л2, Д1,Э2, Э4, Э5 |
|
| 2 | Использование БД и БЗ в информационных системах. | 2 |  |  | 2 |  | 2 | 1 | Л1, Л2, Д1,Э2, Э4, Э5, Э8, Э9 |
|
| 3 | Архитектурный подход к построению БД и БЗ в информационных системах. | 4 |  |  | 4 |  | 4 | 1 | Д1, Э2, Э4 |
|
| 4 | Проектирования БД на основе принципов нормализации и использования семантических диаграммных моделей. | 2 |  | 4 | 6 |  | 6 | 1 | Л1, Л3, Л4, Л5, Л6, Э1, Э2, Э4 |
|
| 5 | CASE-средства и программные среды, использующие для разработки приложений БД и БЗ. | 2 |  | 4 | 6 |  | 6 | 1 | Л3, Л4, Л5. Д2,Э1, Э2, |
|
| 6 | Технологии манипулирования данными в системах построенных на базе реляционных БД. Применения SQL. | 4 |  | 4 | 8 |  | 8 | 1 | Л7, Л8, Л9, Э2, Э3 |
|
| 7 | Технологии, обеспечивающие организацию и построение ИС РБД и БЗ | 4 |  | 2 | 6 |  | 6 | 1 | Л7, Э2, Э4, Э7, Э8, Э9 |
|
| 8 | Теоретические основы организации и проектирования распределенных БД. | 6 |  | 2 | 8 |  | 8 | 1 | Л7, Э2 |
|
| 9 | Архитектура систем поддержки принятия решений. | 4 |  |  | 4 |  | 4 | 1 | Э2, Э5, Э6, Э7, Э8, Э9, Э10,Э11 |
|
| 10 | Архитектура приложений БД, ориентированных на аналитическую обработку информации. | 4 |  | 2 | 6 |  | 6 | 1 | Э2, Э5, Э6, Э7, Э9, Э10, Э11 |
|
| 11 | Архитектура систем, основанных на использовании знаний. | 4 |  |  | 4 |  | 4 | 1 | Э2, Э7 |
|
| 12 | Параллелизм и распределенные среды в приложениях БД. | 4 |  |  | 4 |  | 4 | 1 | Л9 |
|
| 13 | Управление параллельным сервером. | 4 |  |  | 4 |  | 4 | 1 | Л9 |
| 14 | Управление распределенными СУБД | 4 |  |  | 4 |  | 4 | 1 | Л9 |
| 15 | Управление репликацией данных в распределенных средах. | 4 |  |  | 4 |  | 4 | 1 | Л9 |
| Курсовое проектирование | |  |  |  |  | 36 | 36 | 1 |
| Подготовка к экзамену | |  |  |  |  | 36 | 36 | 1 |
| **ИТОГО:** | | 54 |  | 18 | 72 | 72 | 144 |  |

# Учебно-методическое обеспечение дисциплины

# Основная литература

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № | Название, библиографическое описание | Л | Лр | Пз (С) | Кп | Кр | К-во экз. в библ. (на каф.) | Гриф |
| Л1 | М.Р. Когаловский. Энциклопедия технологий баз данных. М.: Финансы и статистика, 2002 | 8 | 8 |  |  | 8 | ЧЗ1(2) Ф(2)  Б(1) |  |
| Л1 | Гектор Гарсиа-Молина, Джеффри Ульман, Дженифер Уидом. Системы баз данных. Полный курс. М., С.-Петербург, Киев: Вильямс, 2003. | 8 | 8 |  |  | 8 | У(16) |  |
| Л3 | Мартин Грабер. SQL. Справочное руководство, 2001 | 8 | 8 |  |  | 8 | нет |  |
| Л4 | Вильям Дж. Пейдж (мл.). Использование  Oracle8/8i, 2001 | 8 | 8 |  |  | 8 | нет |  |

**Дополнительная литература**

|  |  |  |
| --- | --- | --- |
| № | Название, библиографическое описание | К-во экз. в библ. (на каф.) |
| Д1 | Вендров А.М. Проектирование программного обеспечения экономических информационных систем. М.: Финансы и статистика, 2000 | нет |
| Д2 | Маклаков С.В. BPwin и Erwin. CASE-средства разработки информационных систем. – М.: ДИАЛОГ-МИФИ, 1999. | нет |
| Д3 | К. Дейт. Введение в системы баз данных. 7-е изд. М.: СПб.: Вильямс, 2001. | Ф(2) ЧЗ1(2) |
| Д4 | Д. Мейер Теория реляционных баз данных. М.: Мир, 1987 | Ф(3) |
| Д5 | Вендров А.М. CASE-технологии. Современные методы и средства проектирования информационных систем. М.: Финансы и статистика, 1998 | нет |
| Д6 | Фаулер М., Скотт К. UML в кратком изложении. Применение стандартного языка объектного моделирования. М.: Мир, 1999 | Ф(1) |
| Д7 | Буч Г., Рамбо Д., Джекобсон А. Язык UML: руководство пользователя. М.: ДМК, 2000 | У(11) |

|  |  |
| --- | --- |
| Зав. отделом учебной литературы *(для технических дисциплин)* | Киселева Т.В |

**Электронные информационные ресурсы**

|  |  |
| --- | --- |
| № | Название (адрес в Интернет) |
| Э1 | Кузнецов С.Д. Введение в реляционные базы данных. Интернет-университет информационных технологий - ИНТУИТ.ру, 2005. http://www.intuit.ru/department/database/rdbintro/ |
| Э2 | С. Д. Кузнецов. Проектирование и разработка корпоративных информационных систем. Центр Информационных Технологий, 1998. <http://citforum.ru/cfin/prcorpsys/> |
| Э3 | Кузнецов С.Д. Введение в модель данных SQL - Интернет-университет информационных технологий - ИНТУИТ.ру, 2005. <http://www.intuit.ru/department/database/sqlmdintro/> |
| Э4 | Баженова И. Ю. Основы проектирования приложений баз данных Интернет-университет информационных технологий - ИНТУИТ.ру, 2006. <http://www.intuit.ru/department/database/cdba2/> |
| Э5 | Щавелёв Л.В. Оперативная аналитическая обработка: концепции и технологии. - http://www.olap.ru/basic/olap\_and\_ida.htm |
| Э6 | Хрусталёв Е.М. Агрегация данных в OLAP-кубах. – <http://www.olap.ru/Home/mut.htm> |
| Э7 | Эйриэнн Х. Слотер Архитектуры OLAP. - http://www.olap.ru/Basic/olap\_arch.htm |
| Э8 | Альперович М. Введение в OLAP и многомерные базы данных. - <http://www.olap.ru/Basic/alpero2i.htm> |
| Э9 | What is OLAP? - http://[www.olapreport.com](http://www.olapreport.com)/fasmi.htm |
| Э10 | OLAP Architectures - http://[www.olapreport.com](http://www.olapreport.com)/Architectures.htm |

|  |  |
| --- | --- |
| Авторы: |  |
| (с к.т.н., доцент | Савосин С.В. |
| к.т.н., доцент | Егоров С.С. |
| ассистент | Пономарёв А.В. |
|  |  |
| Рецензент |  |
| к.т.н., доцент | Первицкий А.Ю. |
|  |  |
| Зав. кафедрой автоматизированных систем обработки информации управленияи и управления |  |
| д.т.н., профессор | Советов Б.Я. |
|  |  |
| Декан факультета компьютерных технологий и информатики |  |
| д.т.н., профессор | Куприянов М.С. |
|  |  |
| Программа согласована: |  |
|  |  |
| Председатель методической комиссии факультета *компьютерных*  *технологий и информатики* |  |
| к.т.н., доц. | Михалков В.А. |
|  |  |
| Руководитель методического отдела |  |
| к.т.н., доцент | Марасина Л.А. |
|  |  |