

ФЕДЕРАЛЬНОЕ АГЕНТСТВО СВЯЗИ
Федеральное государственное образовательное бюджетное
учреждение высшего профессионального образования
«САНКТ-ПЕТЕРБУРГСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ТЕЛЕКОММУНИКАЦИЙ
им. проф. М. А. БОНЧ-БРУЕВИЧА»

А.Н.Губин

Сети хранения данных

Краткий конспект лекций

Раздел 1. Содержание курса, цели и задачи дисциплины

САНКТ-ПЕТЕРБУРГ
2017

Содержание

Раздел 1. Содержание курса, цели и задачи дисциплины

- 1.1. *Общее содержание курса, цели и задачи изучения дисциплины.*
- 1.2. *Структура дисциплины и ее связь с другими курсами.*
- 1.3. *Хранение информации как основная задача информационных центров.*

Раздел 2. Общие характеристики процессов хранения и управления данными

- 2.1. *Основные технологии хранения данных.*
- 2.2. *Структурированные и неструктурированные данные.*
- 2.3. *Общие характеристики процессов хранения и управления данными.*
- 2.4. *Общая структура информационных центров.*
- 2.5. *Управление хранением данных. Многоуровневое хранение данных.*

Раздел 3. Основные компоненты среды хранения данных

- 3.1. *Физические компоненты. Хост, дисковое устройство.*
- 3.2. *Логические компоненты. Файловые системы.*
- 3.3. *Производительность дисковых устройств.*

Раздел 4. Защита данных. RAID-массивы

- 4.1. *Распределение данных. Зеркалирование данных.*
- 4.2. *Контроль четности.*
- 4.3. *Основные конфигурации RAID-массивов.*
- 4.4. *IOPS-операций и конфигурация дисков.*
- 4.5. *Влияние наличия RAID на производительность дисковых устройств.*

Раздел 5. Системы хранения данных. Интеллектуализация систем хранения данных

- 5.1. *Основные компоненты интеллектуальных систем хранения информации.*
- 5.2. *Операция чтения и записи данных с использованием КЭШ-памяти.*
Защита КЭШ- данных
- 5.3. *Основные компоненты системы хранения данных Symmetrix.*

Раздел 6. Сети хранения данных

- 6.1. *Архитектура сетей хранения данных. Особенности архитектуры СХД Fiber- Channel. Зонирование данных.*
- 6.2. *Контекстная адресация данных.*
- 6.3. *Виртуализация систем хранения данных.*

Раздел 7. Перспективные направления развития технологий хранения данных и управления информацией

- 7.1. *Внеполостная и внутриполостная виртуализация систем данных.*
- 7.2. *Основные проблемы виртуализации систем хранения данных.*

Раздел 1. Содержание курса, цели и задачи дисциплины

1.1. *Общее содержание курса, цели и задачи изучения дисциплины*

Общее содержание курса представлено в табл. 1.

Таблица 1.

№ п/п	Наименование раздела дисциплины	Содержание раздела
1	Введение. Содержание курса, цели и задачи дисциплины.	Общее содержание курса, цели и задачи изучения дисциплины. Структура дисциплины и ее связь с другими курсами. Хранение информации как основная задача информационных центров.
2	Общие характеристики процессов хранения и управления данными. Данные и информация. Основные технологии хранения данных.	Структурированные и неструктурированные данные. Общая структура информационных центров. Управление хранением данных. Многоуровневое хранение данных.
3	Основные компоненты среды хранения данных Физические компоненты. Хост, дисковое устройство. Логические компоненты. Файловые системы.	Хост, соединение устройств и хранение данных. Средства и устройства хранения данных. Производительность дисковых устройств. Файловые системы.
4	Защита данных. RAID-массивы. IOPS-операций и конфигурация дисков.	Распределение данных. Зеркалирование данных. Контроль четности. Основные конфигурации RAID-массивов. Влияние наличия RAID на производительность дисковых устройств.
5	Системы хранения данных. Интеллектуализация систем хранения данных. Особенности использования КЭШ-памяти. Основные компоненты системы хранения данных Symmetrix.	Основные компоненты интеллектуальных систем хранения информации. Структура КЭШ-памяти. Операция чтения и записи данных с использованием КЭШ-памяти. Защита КЭШ-данных. Высокопроизводительные системы хранения данных. Реализация интеллектуальных массивов данных в Symmetrix.
6	Сети хранения данных. Контекстная адресация данных. Виртуализация систем	Архитектура сетей хранения данных. Особенности архитектуры СХД Fiber-Channel. Зонирование данных. Система хранения данных с адресацией по содержанию. Хранение и извлечение объектов в системах с

	хранения данных.	контекстной адресацией данных. Виртуализация систем хранения данных на блочном и файловом уровнях.
7	Заключение. Перспективные направления развития технологий хранения данных и управления информацией.	Внеполостная и внутриполостная виртуализация систем данных. Основные проблемы виртуализации систем хранения данных.

Целью преподавания дисциплины является изучение современных технологий управления информацией и хранения данных используемых в информационных системах. Дисциплина «Сети хранения данных» должна обеспечивать формирование фундамента подготовки будущих специалистов в области информационных технологий, а также, создавать необходимую базу для успешного овладения последующими специальными дисциплинами учебного плана. Приобретенные студентами знания и навыки необходимы для осуществления профессиональной деятельности в области разработки, проектирования и эксплуатации современных информационно-вычислительных систем.

1.2. Структура дисциплины и ее связь с другими курсами

Общая структура содержания дисциплины представлена в табл. 2.

Таблица 2.

Вид учебной работы	Всего часов	Семестр 6
Общая трудоемкость	108	108
Аудиторные занятия (всего)	50	50
В том числе:		
Лекции	20	20
Практические занятия (ПЗ)	16	16
Лабораторные работы (ЛР)	14	14
Самостоятельная работа (всего)	58	58
В том числе:		
Подготовка к лабораторным работам	28	28
Подготовка к зачету	30	30
Вид промежуточной аттестации	Зачет	Зачет

Овладение предметом дисциплины «Сети хранения данных» является обязательным для изучения последующих дисциплин учебного плана:

«Технологии обработки информации», «Современные технологии проектирования информационных систем».

1.3. Хранение информации как основная задача информационных центров.

Хранение информации является важнейшей составляющей современных информационных технологий. Рост количества производящих контент устройств обуславливает существенный (экспоненциальный) рост объемов как структурированных (данные организованные в определенные структуры и строго форматированные по содержанию), так и неструктурированных данных (веб-страницы, изображения, аудио/видео потоки, тексты и др.). Возможность обработки данных предполагает их хранение и обеспечение доступности для приложений, использующих эти данные. Причем, в последнее время все большую актуальность приобретает возможность хранения и обработки неструктурированных данных.

Данные представляют собой некоторую совокупность фактов представленных в формализованном виде.

Традиционные базы данных, как правило, ориентированы на выполнение жестко определенных операций по обработке данных. Возможности обработки нерегламентированных запросов в таких базах данных существенно ограничены.

В то же время, необходимость выполнения по отношению к системе хранения данных таких операций как

- формирование запросов произвольной формы;
- интегрирование данных из различных генерирующих контент систем;
- обработка больших объемов данных,

вызвало появление новых технологий организации баз данных, а именно технологии хранилищ данных.

В основе концепции хранилища данных лежат две основные идеи:

- интеграция разьединенных детализированных данных в едином хранилище;

- разделение наборов данных и приложений, используемых для оперативной обработки и применяемых для решения задач анализа.

Определение понятия "хранилище данных"(Data Warehouse) первым дал Уильям Г. Инмон в 2002 году. Он определил хранилище данных как "предметно-ориентированную, интегрированную, содержащую исторические данные, не разрушаемую совокупность данных, предназначенную для поддержки принятия управленческих решений".

Концептуально модель хранилища данных можно представить в виде схемы, представленной на рис. 1.

Данные из различных источников помещаются в хранилище данных, а описания этих данных в репозиторий метаданных. Конечный пользователь, используя различные инструменты (средства визуализации, построения отчетов, статистической обработки и т.д.) и содержимое репозитория,

анализирует данные в хранилище. Результатом его деятельности обычно является информация в виде готовых отчетов, найденных скрытых закономерностей, каких-либо прогнозов. Так как средства работы конечного пользователя с хранилищем данных могут быть самыми разнообразными, то теоретически их выбор не должен влиять на структуру и функции поддержания в актуальном состоянии хранилища данных.

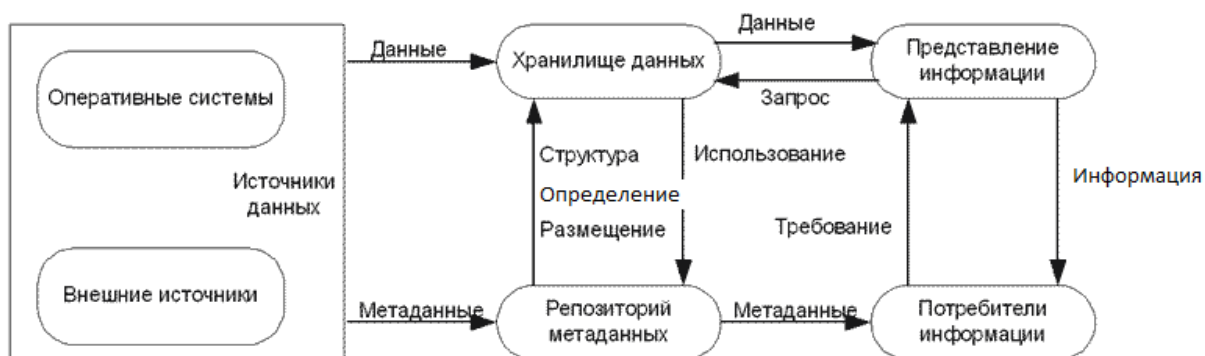


Рис. 1. Концептуальная модель хранилища данных