

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 2. Общие характеристики процессов хранения и управления
данными**

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 2. Общие характеристики процессов хранения и управления данными

2.1. Основные технологии хранения данных.

В случае отдельного ПК под системой хранения данных можно понимать внутренний жесткий диск или систему дисков (RAID массив). Если же речь заходит о системах хранения данных масштаба предприятий и корпораций, то традиционно можно выделить три технологии организации систем хранения данных:

- Direct Attached Storage (DAS);
- Network Attach Storage (NAS);
- Storage Area Network (SAN).

Устройства DAS (Direct Attached Storage) – решение, когда устройство для хранения данных подключено непосредственно к серверу, или к рабочей станции, как правило, через интерфейс по протоколу SAS.

Архитектура системы хранения DAS представлена на рис.2.



Рис. 2. Архитектура СХД Direct Attached Storage

К основным преимуществам DAS систем можно отнести их низкую стоимость (в сравнении с другими решениями СХД), простоту развертывания и администрирования, а также высокую скорость обмена данными между системой хранения и сервером. Этим объясняется большая популярность их использования в сегменте малых офисов, хостинг-провайдеров и небольших корпоративных сетей.

По сути, данный тип СХД представляет собой дисковую полку (корзину) с жесткими дисками, вынесенную за пределы сервера.

Дисковая полка - это внешний программно-аппаратный массив жестких дисков с набором дополнительных функций, подключаемый к хосту через высокоскоростную магистраль передачи данных (SCSI U320 Мбит / сек или оптический интерфейс).

К недостаткам DAS-систем следует отнести неоптимальную утилизацию ресурсов, поскольку каждая DAS система требует подключения выделенного сервера и позволяет подключить максимум 2 сервера к дисковой полке в определенной конфигурации. Кроме того, СХД DAS обладает низкой надежностью, так как при выходе из строя сервера, к которому подключено хранилище данных, данные становятся недоступными.

Устройства NAS (Network Attached Storage) – отдельно стоящая интегрированная дисковая система, по-сути, NAS-сервер, со своей специализированной ОС и набором полезных функций быстрого запуска системы и обеспечения доступа к файлам.

Архитектура системы хранения NAS представлена на рис.3.



Рис. 3. Архитектура СХД Network Attached Storage

Технология NAS (сетевые подсистемы хранения данных) развивается как альтернатива универсальным серверам, реализующим множество функций (печати, приложений, факс сервер, электронная почта и т.п.). В отличие от них NAS-устройства исполняют только одну функцию - файловый сервер.

NAS подключаются к ЛВС и обеспечивают доступ к данным для неограниченного количества гетерогенных клиентов (клиентов с различными ОС) или других серверов. В настоящее время практически все NAS устройства ориентированы на использование в сетях Ethernet (Fast Ethernet, Gigabit Ethernet) на основе протоколов TCP/IP. Доступ к устройствам NAS производится с помощью специальных протоколов доступа к файлам. Наиболее распространенными протоколами файлового доступа являются протоколы CIFS, NFS и DAFS. Внутри подобных серверов стоят специализированные ОС, такие как MS Windows Storage Server.

К достоинствам NAS следует отнести:

- Дешевизну и доступность ресурсов СХД не только для отдельных серверов, но и для любых компьютеров предприятия или корпорации.
- Простоту коллективного использования ресурсов.
- Простоту развёртывания и администрирования
- Универсальность для клиентов (один сервер может обслуживать клиентов MS, Novell, Mac, Unix)

Недостатками NAS являются:

- Реализация доступа к данным через протоколы “сетевых файловых систем” что медленнее, чем доступ к локальному диску.
- Отсутствие возможности (для большинства недорогих NAS-серверов) обеспечить скоростной и гибкий метод доступа к данным на уровне блоков, а не на уровне файлов.

Storage Area Network (SAN) –это специальная выделенная сеть, объединяющая устройства хранения данных с серверами приложений, обычно строится на основе протокола Fibre Channel или протокола iSCSI.

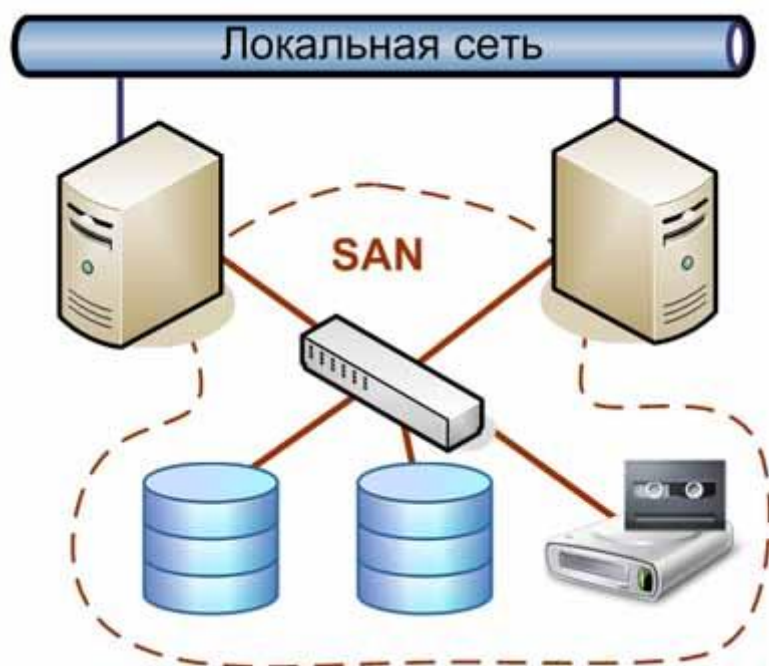


Рис. 4. Архитектура СХД Storage Area Network

В отличие от NAS, SAN не имеет понятия о файлах: файловые операции выполняются на подключенных к SAN серверах. SAN оперирует блоками, как некий большой жесткий диск. Идеальный результат работы SAN - возможность доступа любого сервера под любой операционной системой к любой части дисковой емкости, находящейся в SAN. Оконечные элементы SAN - это серверы приложений и системы хранения данных (дисковые массивы, ленточные библиотеки и т. п.). А между ними, как и в обычной сети, находятся адаптеры, коммутаторы, мосты, концентраторы.

iSCSI является более «дружелюбным» протоколом, поскольку он основан на использовании стандартной инфраструктуры Ethernet – сетевых карт, коммутаторов, кабелей. Более того, именно системы хранения данных на базе iSCSI являются наиболее популярными для виртуализированных серверов, в силу простоты настройки протокола. Архитектура системы хранения SAN представлена на рис.4.

К достоинствам SAN относятся:

- Высокая надёжность доступа к данным, находящимся на внешних системах хранения. Независимость топологии SAN от используемых СХД и серверов.

- Централизованное хранение данных (надёжность, безопасность).

- Удобное централизованное управление коммутацией и данными.

- Перенос интенсивного трафика ввода-вывода в отдельную сеть (разгружается LAN).

- Высокое быстродействие (обеспечивается скоростной и гибкий метод доступа к данным на уровне блоков).

- Масштабируемость и гибкость логической структуры SAN

- Возможность организации резервных, удаленных СХД и удаленной системы бэкапа и восстановления данных.

- Возможность строить отказоустойчивые кластерные решения без дополнительных затрат на базе имеющейся SAN.

К недостаткам SAN относятся:

- Более высокая стоимость

- Сложность в настройке FC-систем

- Необходимость сертификации специалистов по FC-сетям (iSCSI является более простым протоколом)

- Более жесткие требования к совместимости компонентов SAN.

2.2. Структурированные и неструктурированные данные.

Структурированные и неструктурированные данные — два пересекающихся подмножества данных информационного пространства объекта, и анализ каждого подмножества в отдельности позволяет выделить закономерности, присущие только ему. Например, используя структурированные данные о котировках акций компании, можно оценить динамику их роста или падения на фондовой бирже, а при обработке неструктурированных источников, таких как публикации в СМИ, открывается возможность исследовать эмоциональный фон вокруг бренда этой компании и сформулировать экспертную оценку влияния той или иной информации на котировки ее акций. Работая со всей информацией о конкретном событии или процессе, можно научиться осознанно и целенаправленно управлять этим процессом.

Различают структурированные данные, в которых отражаются отдельные факты предметной области (**это основная форма представления данных в СУБД**), и

неструктурированные, произвольные по форме, включающие и тексты, и графику, и прочие данные. Эта форма представления данных широко используется, например, в **Интернет-технологиях**, а сами данные предоставляются пользователю в виде отклика поисковыми системами.

Организация того или иного вида хранения данных (структурированных или неструктурированных) **связана с обеспечением доступа к самим данным.**

Под **доступом** понимается возможность выделения элемента данных (или множества элементов) среди других элементов по каким-либо признакам с целью выполнения некоторых действий над элементом.

При этом под элементом понимается как запись файла (в случае структурированных данных), так и сам файл (в случае неструктурированных данных).

Для данных любого вида доступ осуществляется с помощью специальных данных, которые называются **ключевыми (ключами).**

Для **структурированных** данных такие ключи входят в состав записей файлов в качестве отдельных полей записей.

Для **неструктурированных** поисковые слова или выражения входят, как правило, в искомый текст.

С помощью ключей выполняется идентификация требуемых элементов в информационном массиве (массиве хранения данных).

Дальнейшее изложение фазы хранения информации относится к структурированным данным.

Модели структурированных данных и технологии их обработки основаны на одном из **трех способов организации хранения данных:**

- в виде линейного списка (или табличном),
- иерархическом (или древовидном),
- сетевом.

Линейный список (далее – список) – это множество элементов хранения (далее – элементов) с заданным отношением строгого порядка, определяющим следование элементов в множестве.

Требование строгого порядка вызвано тем, что с каждым элементом линейного списка при хранении связан конкретный физический адрес. Примером линейного списка может быть состав студентов учебной группы или расписание экзаменов во время сессии.

В зависимости от места нахождения списка (**внутренняя или внешняя память компьютера**) он может быть организован как одномерный массив (в первом случае) или как файл (во втором случае). Тогда элементом списка может быть элемент массива или запись файла.

Элемент списка может иметь структуру. В таком случае он состоит из отдельных полей (это составляющие элемента). Например (рис.2.1), элемент списка, соответствующего составу учебной группы, может состоять из следующих полей: фамилия, имя, отчество студента, номер зачетной книжки, домашний адрес:

№ п/п	Фамилия	Имя	Отчество	Номер зачетной книжки	Домашний адрес
1	Строков	Иван	Иванович	02343	ул. Красная, 9 - 2
2	Скворцов	Олег	Иванович	02544	пр. Мира, 45 - 3
3	Соколов	Юрий	Кузьмич	01278	ул. Леонова, 23 - 98

Рис.2.1. Структура элементов списка

Иерархическая структура — многоуровневая форма организации объектов со строгой соотнесенностью объектов нижнего уровня определенному объекту верхнего уровня.

«У подчиненного может быть только один руководитель».

Графически представляется в виде дерева.

Сеть (или сетевая структура) – это два множества T и R , между которыми задано отображение $\Gamma: T \rightarrow R$, где T – множество элементов сети, R – множество бинарных отношений между ними, Γ – отображение, показывающее, какие элементы какими отношениями связаны.

Понятие “**информация**” тесно связано с такими терминами, как “**данные**”, “**сообщения**”, “**сведения**” и “**знания**”.

Данные - это формальные факты или идеи, которые можно хранить, обрабатывать и передавать на расстояние.

Данные также определяются, как числа, символы или буквы, которые используют при описании личностей, объектов, ситуаций, а также для их анализа, обсуждения или принятия соответствующих решений.

Другой формой представления информации является сообщение.

Сообщение - это текст, цифровые данные, изображения, звук, графика, таблицы и др.

Сообщения содержат информацию тогда, когда могут быть приняты и поняты любым живым существом или приёмником информации.

Сведения – практически синоним понятия “Сообщения”. Они чаще всего носят бытовой характер.

Можно считать, что сведения, сообщения и данные являются составляющими (компонентами) информации, особенно когда говорят, что они используются в вычислительной технике в виде электронных (машиночитаемых) данных.

Важной составляющей информации являются знания.

Знания – это:

- а) вид информации, отражающий опыт и восприятие человеком окружающего мира;
- б) понимание определённой информации с целью лучшего её использования при решении конкретных задач;
- в) факты и правила, сохраняющиеся в памяти людей и влияющие на их убеждения;
- г) способность получать информацию и отношение к полученным данным и др.

Знание - способность человека получать необходимые ему данные, обдумывать (осмысливать) и преобразовывать их в информацию.

Информация не всегда превращается в знания. Она может быть динамична, когда речь идёт о распространении и функционировании знаний потому, что одни и те же данные могут представлять разную информацию.

Получив какие-либо данные, человек усваивает (воспринимает и понимает), а затем превращает их (информационно-когнитивный процесс) в новую (по крайней мере для себя) информацию. Так происходит воссоздание (обновление) знаний, получение новых личных и общественных знаний. Этот процесс изображён на Рис. 2.2.



Рис. 2.2. Соотношение понятий "информация", "данные", "знания"

Если рассматривать информацию как некоторый физический объект, то ее можно:

- 1) создавать (генерировать),
- 2) передавать (транслировать),
- 3) хранить и сохранять,
- 4) обрабатывать (перерабатывать).

Поскольку информация представляет интерес для различных категорий пользователей, то основным назначением информации является её использование. При этом выделяют такие её свойства, как: адресность, актуальность, возможность кодирования, высокая скорость сбора, обработки и

передачи, достаточность, достоверность, многократность использования, правовая корректность, полнота, своевременность.

Существенными составляющими информации являются её потребительские свойства, то есть те из них, которые наиболее важны для её потребителей. Основные *потребительскими свойствами информации* представлены на Рис. 2.3.



Рис. 2.3. Потребительские свойства информации.

2.3. Общие характеристики процессов хранения и управления данными

В информационных системах можно выделить отдельные этапы (фазы) обработки информации (данных), каждый из которых характеризуется определенными действиями.

Последовательность действий, выполняемых с информацией, называют информационным процессом.

Основными составляющими информационных процессов являются:

- сбор (восприятие) информации;
- подготовка (преобразование) информации;
- передача информации;
- обработка (преобразование) информации;
- хранение информации;
- отображение (воспроизведение) информации.

Основные этапы преобразования данных в составе информационных систем показаны на рис.2.4.



Рис. 2.4. Основные этапы преобразования данных в составе информационных систем

На этапе восприятия информации осуществляется целенаправленное извлечение и анализ информации о каком-либо объекте (процессе), в результате чего формируется образ объекта, проводятся его опознание и оценка. Главная задача на этом этапе – отделить полезную информацию от помех (шумов), что в ряде случаев связано со значительными трудностями.

На этапе подготовки информации осуществляется ее первичное преобразование. На этом этапе проводятся такие операции, как нормализация, преобразование, шифрование. Иногда этап подготовки рассматривается как вспомогательный на этапе восприятия. В результате восприятия и подготовки получается сигнал в форме, удобной для передачи, хранения или обработки.

На этапе передачи информация пересылается из одного места в другое (от отправителя получателю – адресату). Передача осуществляется по каналам различной физической природы, самыми распространенными из которых являются электрические, электромагнитные и оптические. Извлечение сигнала на выходе канала, подверженного действию шумов, носит характер вторичного восприятия.

На этапах обработки информации выявляются ее общие и существенные взаимозависимости, представляющие интерес для системы. Преобразование информации на этапе обработки (как и на других этапах) осуществляется либо средствами информационной техники, либо человеком.

Под обработкой информации понимается любое ее преобразование, проводимое по законам логики, математики, а также неформальным правилам, основанным на «здоровом смысле», интуиции, обобщенном опыте, сложившихся взглядах и нормах поведения. Результатом обработки является тоже информация, но либо представленная в иных формах (например, упорядоченная по каким-то признакам), либо содержащая ответы на поставленные вопросы (например, решение некоторой задачи). Если процесс обработки формализуем, он может выполняться техническими средствами.

Данными называют факты, сведения, представленные в формализованном виде (закодированные), занесенные на те или иные

носители и допускающие обработку с помощью специальных технических средств (в первую очередь ЭВМ).

Обработка данных предполагает производство различных операций над ними, в первую очередь арифметических и логических, для получения новых данных, которые объективно необходимы (например, при подготовке ответственных решений).

На этапе хранения информацию записывают в запоминающее устройство для последующего использования. Для хранения информации используются в основном полупроводниковые и магнитные носители.

Этап отображения информации должен предшествовать этапам, связанным с участием человека. Цель этого этапа – предоставить человеку нужную ему информацию с помощью устройств, способных воздействовать на его органы чувств.

Любая информация обладает рядом свойств, которые в совокупности определяют степень ее соответствия потребностям пользователя (качество информации). Можно привести немало разнообразных свойств информации, так как каждая научная дисциплина рассматривает те свойства, которые ей наиболее важны. С точки зрения информатики наиболее важными представляются следующие:

Актуальность информации – свойство информации сохранять ценность для потребителя в течение времени, т. е. не подвергаться «моральному» старению.

Полнота информации – свойство информации, характеризуемое мерой достаточности для решения определенных задач. Полнота информации означает, что она обеспечивает принятие правильного (оптимального) решения. Оценивается относительно вполне определенной задачи или группы задач.

Адекватность информации – свойство, заключающееся в соответствии содержательной информации состоянию объекта. Нарушение идентичности связано с техническим старением информации, при котором происходит расхождение реальных признаков объектов и тех же признаков, отображенных в информации.

Сохранность информации – свойство информации, характеризуемое степенью готовности определенных информационных массивов к целевому применению и определяемое способностью контроля и защиты информации обеспечить постоянное наличие и своевременное предоставление информационного массива, необходимых для автоматизированного решения целевых и функциональных задач системы.

Достоверность информации – свойство информации, характеризуемое степенью соответствия реальных информационных единиц их истинному значению. Требуемый уровень достоверности информации достигается путем внедрения методов контроля и защиты информации на всех стадиях ее переработки, повышения надежности комплекса технических и программных

средств информационной системы, а также административно-организационными мерами.

2.4. Общая структура информационных центров

Информационный центр представляет собой взаимосвязанную совокупность средств, методов и персонала, используемых для хранения, обработки и выдачи информации в интересах достижения поставленной цели

Типы обеспечивающих подсистем

Структуру информационной системы составляет совокупность отдельных ее частей, называемых подсистемами.

Подсистема - это часть системы, выделенная по какому-либо признаку.

Общую структуру информационной системы можно рассматривать как совокупность подсистем (рис.2.5) независимо от сферы применения. В этом случае говорят о *структурном признаке* классификации, а подсистемы называют обеспечивающими. Таким образом, структура любой информационной системы может быть представлена совокупностью обеспечивающих подсистем.



Рис. 2.5. Структура информационной системы как совокупность обеспечивающих подсистем

Среди обеспечивающих подсистем обычно выделяют информационное, техническое, математическое, программное, организационное и правовое обеспечение.

Назначение подсистемы информационного обеспечения состоит в современном формировании и выдаче достоверной информации для принятия управленческих решений.

Информационное обеспечение - совокупность единой системы классификации и кодирования информации, унифицированных систем документации, схем информационных потоков, циркулирующих в организации, а также методология построения баз данных.

Унифицированные системы документации создаются на государственном, республиканском, отраслевом и региональном уровнях. Главная цель - это обеспечение сопоставимости показателей различных сфер общественного производства. Разработаны стандарты, где устанавливаются требования:

- к унифицированным системам документации;
- к унифицированным формам документов различных уровней управления;
- к составу и структуре реквизитов и показателей;
- к порядку внедрения, ведения и регистрации унифицированных форм документов.

Однако, несмотря на существование унифицированной системы документации, при обследовании большинства организаций постоянно выявляется целый комплекс типичных недостатков:

- чрезвычайно большой объем документов для ручной обработки;
- одни и те же показатели часто дублируются в разных документах;
- работа с большим количеством документов отвлекает специалистов от решения непосредственных задач;
- имеются показатели, которые создаются, но не используются, и др.

Поэтому устранение указанных недостатков является одной из задач, стоящих при создании информационного обеспечения.

Схемы информационных потоков отражают маршруты движения информации и ее объемы, места возникновения первичной информации и использования результатной информации. За счет анализа структуры подобных схем можно выработать меры по совершенствованию всей системы управления.

Пример. В качестве примера простейшей схемы потоков данных можно привести схему, где отражены все этапы прохождения служебной записки или записи в базе данных о приеме на работу сотрудника - от момента ее создания до выхода приказа о его зачислении на работу.

Построение схем информационных потоков, позволяющих выявить объемы информации и провести ее детальный анализ, обеспечивает:

- исключение дублирующей и неиспользуемой информации;
- классификацию и рациональное представление информации.

При этом подробно должны рассматриваться вопросы взаимосвязи движения информации по уровням управления. Следует выявить, какие показатели необходимы для принятия управленческих решений, а какие нет. К

каждому исполнителю должна поступать только та информация, которая используется.

Методология построения баз данных базируется на теоретических основах их проектирования. Для понимания концепции методологии приведем основные ее идеи в виде двух последовательно реализуемых на практике этапов:

1-й этап - обследование всех функциональных подразделений фирмы с целью:

- понять специфику и структуру ее деятельности;
- построить схему информационных потоков;
- проанализировать существующую систему документооборота;
- определить информационные объекты и соответствующий состав реквизитов (параметров, характеристик), описывающих их свойства и назначение.

2-й этап - построение концептуальной информационно-логической модели данных для обследованной на 1-м этапе сферы деятельности. В этой модели должны быть установлены и оптимизированы все связи между объектами и их реквизитами. Информационно-логическая модель является фундаментом, на котором будет создана база данных.

Для создания информационного обеспечения необходимо:

- ясное понимание целей, задач, функций всей системы управления организацией;
- выявление движения информации от момента возникновения и до ее использования на различных уровнях управления, представленной для анализа в виде схем информационных потоков,
- совершенствование системы документооборота;
- наличие и использование системы классификации и кодирования;
- владение методологией создания концептуальных информационно-логических моделей, отражающих взаимосвязь информации;
- создание массивов информации на машинных носителях, что требует наличия современного технического обеспечения.

Техническое обеспечение - комплекс технических средств, предназначенных для работы информационной системы, а также соответствующая документация на эти средства и технологические процессы

Комплекс технических средств составляют:

- компьютеры любых моделей;

- устройства сбора, накопления, обработки, передачи и вывода информации;
- устройства передачи данных и линий связи;
- оргтехника и устройства автоматического съема информации;
- эксплуатационные материалы и др.

Документацией оформляются предварительный выбор технических средств, организация их эксплуатации, технологический процесс обработки данных, технологическое оснащение. Документацию можно условно разделить на три группы:

- общесистемную, включающую государственные и отраслевые стандарты по техническому обеспечению;
- специализированную, содержащую комплекс методик по всем этапам разработки технического обеспечения;
- нормативно-справочную, используемую при выполнении расчетов по техническому обеспечению.

К настоящему времени сложились две основные формы организации технического обеспечения (формы использования технических средств): централизованная и частично или полностью децентрализованная.

Централизованное техническое обеспечение базируется на использовании в информационной системе больших ЭВМ и вычислительных центров.

Децентрализация технических средств предполагает реализацию функциональных подсистем на персональных компьютерах непосредственно на рабочих местах.

Перспективным подходом следует считать, по-видимому, частично децентрализованный подход - организацию технического обеспечения на базе распределенных сетей, состоящих из персональных компьютеров и большой ЭВМ для хранения баз данных, общих для любых функциональных подсистем.

Математическое и программное обеспечение - совокупность математических методов, моделей, алгоритмов и программ для реализации целей и задач информационной системы, а также нормального функционирования комплекса технических средств.

К средствам *математического обеспечения* относятся:

- средства моделирования процессов управления;
- типовые задачи управления;
- методы математического программирования, математической статистики, теории массового обслуживания и др.

В состав *программного обеспечения* входят общесистемные и специальные программные продукты, а также *техническая документация*.

К *общесистемному программному обеспечению* относятся комплексы программ, ориентированных на пользователей и предназначенных для решения типовых задач обработки информации. Они служат для расширения функциональных возможностей компьютеров, контроля и управления процессом обработки данных.

Специальное программное обеспечение представляет собой совокупность программ, разработанных при создании конкретной информационной системы. В его состав входят пакеты прикладных программ (ППП), реализующие разработанные модели разной степени адекватности, отражающие функционирование реального объекта.

Техническая документация на разработку программных средств должна содержать описание задач, задание на алгоритмизацию, экономико-математическую модель задачи, контрольные примеры.

Организационное обеспечение - совокупность методов и средств, регламентирующих взаимодействие работников с техническими средствами и между собой в процессе разработки и эксплуатации информационной системы.

Организационное обеспечение реализует следующие функции:

- анализ существующей системы управления организацией, где будет использоваться ИС, и выявление задач, подлежащих автоматизации;
- подготовку задач к решению на компьютере, включая техническое задание на проектирование ИС и технико-экономическое обоснование ее эффективности;
- разработку управленческих решений по составу и структуре организации, методологии решения задач, направленных на повышение эффективности системы управления.

Организационное обеспечение создается по результатам предпроектного обследования на 1-м этапе построения баз данных, с целями которого вы познакомились при рассмотрении информационного обеспечения.

Правовое обеспечение - совокупность правовых норм, определяющих создание, юридический статус и функционирование информационных систем, регламентирующих порядок получения, преобразования и использования информации.

Главной целью правового обеспечения является укрепление законности.

В состав правового обеспечения входят законы, указы, постановления государственных органов власти, приказы, инструкции и другие нормативные документы министерств, ведомств, организаций, местных органов власти. В правовом обеспечении можно выделить общую часть, регулирующую функционирование любой информационной системы, и локальную часть, регулирующую функционирование конкретной системы.

Правовое обеспечение этапов разработки информационной системы включает нормативные акты, связанные с договорными отношениями разработчика и заказчика и правовым регулированием отклонений от договора.

Правовое обеспечение этапов функционирования информационной системы включает:

- статус информационной системы;
- права, обязанности и ответственность персонала;
- правовые положения отдельных видов процесса управления;
- порядок создания и использования информации и др.

2.5. Управление хранением данных. Многоуровневое хранение данных

Объемы хранимых данных ежегодно увеличиваются на 50–80%, что заставляет разработчиков искать альтернативы сложным СХД с ограниченной масштабируемостью, создавать решения, более эффективно использующие ресурсы ЦОД и сокращающие время простоя, а кроме того, упрощать администрирование за счет автоматизации операций управления данными, ведь расходы на управление также быстро растут. Эта задача повышения эффективности хранения данных решается с помощью консолидации и многоуровневого хранения. Многоуровневое, иерархическое хранение информации (рис. 2.6) — один из подходов, которые приходят на смену экстенсивному наращиванию емкости хранения данных.

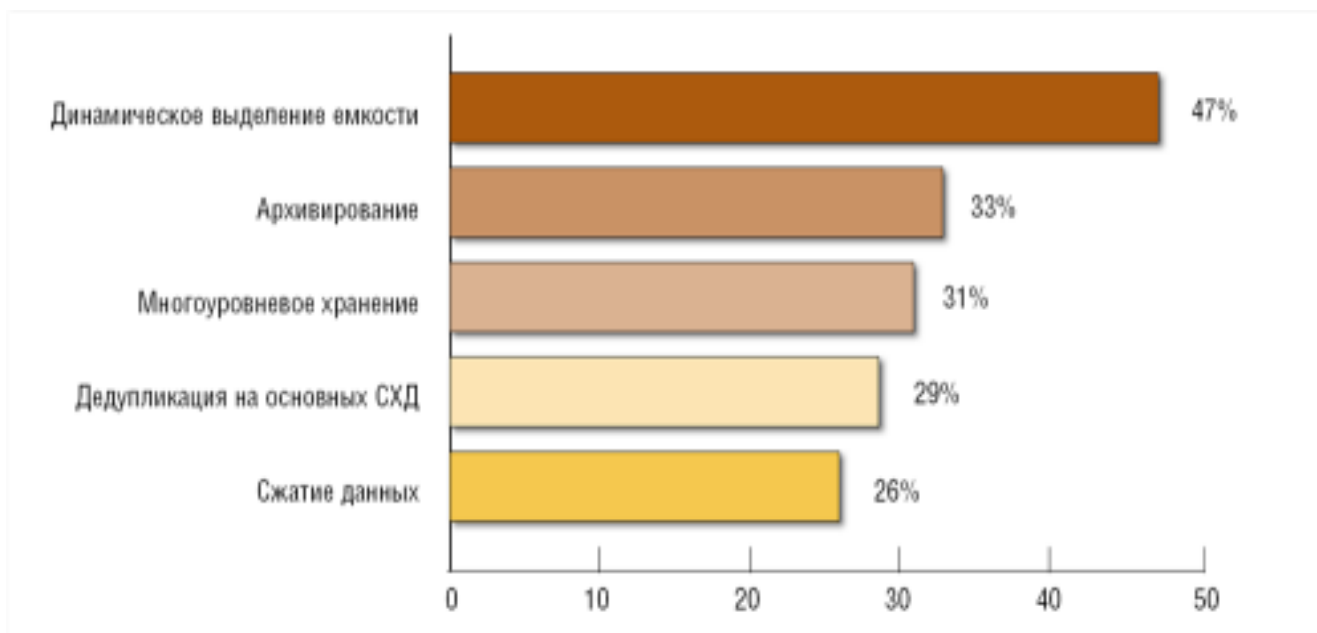


Рис.2.6. Многоуровневое хранение данных

Нередко дисковые массивы содержат большие объемы информации, доступ к которой осуществляется очень редко, если происходит вообще.

Согласно статистике, 90% данных после их создания не востребованы. Идея отделить «нужные» данные (частота обращения к ним считается лучшим критерием ценности) и поместить их на более дорогие, но быстрые носители, а все остальное хранить на дешевых и медленных или в СХД, — возникла уже давно. В частности, она нашла выражение в концепции управления жизненным циклом данных (ILM) — процедуре, при которой системный администратор должен постоянно менять приоритет данных в зависимости от их важности для бизнеса. Это сложная задача, отнимающая много времени, к тому же тома (LUN) перемещаются целиком, хотя могут содержать данные, запрашиваемые с разной степенью интенсивности.

Концепция ILM существует уже не один год, и все это время предпринимались более или менее успешные попытки реализовать ее на практике.

Решения HSM функционируют на достаточно высоком уровне в инфраструктуре систем хранения данных. Обычно с их помощью расширялся функционал файловой системы за счет присоединения к ней дополнительного, «внешнего» уровня хранения. На практике подобные продукты показали весьма ограниченную применимость к базам данных, успешно реализованные решения базировались в основном на платформе Windows, что сократило сферу использования продуктов HSM до файловых серверов.

Современное многоуровневое хранение реализовано на более низких уровнях, внутри СХД, прозрачно для файловых систем и приложений и, таким образом, может использоваться для очень широкого круга задач — от транзакционных СУБД до файловых серверов. Причем оно нисколько не противоречит ILM: эти подходы могут сочетаться в рамках одной крупной инфраструктуры».

В концепции иерархического, или ярусного, хранения (Tiered Storage или Storage Tiering) в СХД выделяются несколько уровней хранения с разными показателями производительности и стоимости: стандартные твердотельные накопители (Solid State Drive, SSD) или специально разрабатываемые вендорами для своих СХД флэш-накопители (Enterprise Flash Drive, EFD), жесткие диски SAS, FC и/или SATA с разной частотой вращения или разные уровни RAID, магнитные ленты и даже облако. Причем, по мнению специалистов IBM, Tiered Storage — скорее технический термин, а Storage Tiering — организационный.

Новые технологии позволяют динамически перемещать данные между разными уровнями. SSD можно применять для активных («горячих») блоков данных или в качестве кэш-памяти для чтения/записи. В HDS считают, что 80% информации, размещенной на рабочих СХД, следует перенести на менее дорогие и производительные системы хранения. При этом не только не теряется скорость доступа к часто используемым данным, но и значительно снижаются потребности в дорогостоящей дополнительной емкости хранения.

В системах с автоматическим или динамическим многоуровневым хранением (Automated Tiered Storage, ATS или Automated Storage Tiering, AST)

данные автоматически переносятся между разными уровнями в зависимости от интенсивности использования. AST предполагает их динамическую классификацию и миграцию. В результате самые быстрые и дорогие накопители выделяются для данных, поступающих от критически важных приложений (рис.2.7). Динамическое многоуровневое хранение позволяет увеличить производительность последних (например, существенно улучшить время отклика в СУБД) и сократить издержки на хранение данных и энергопотребление.

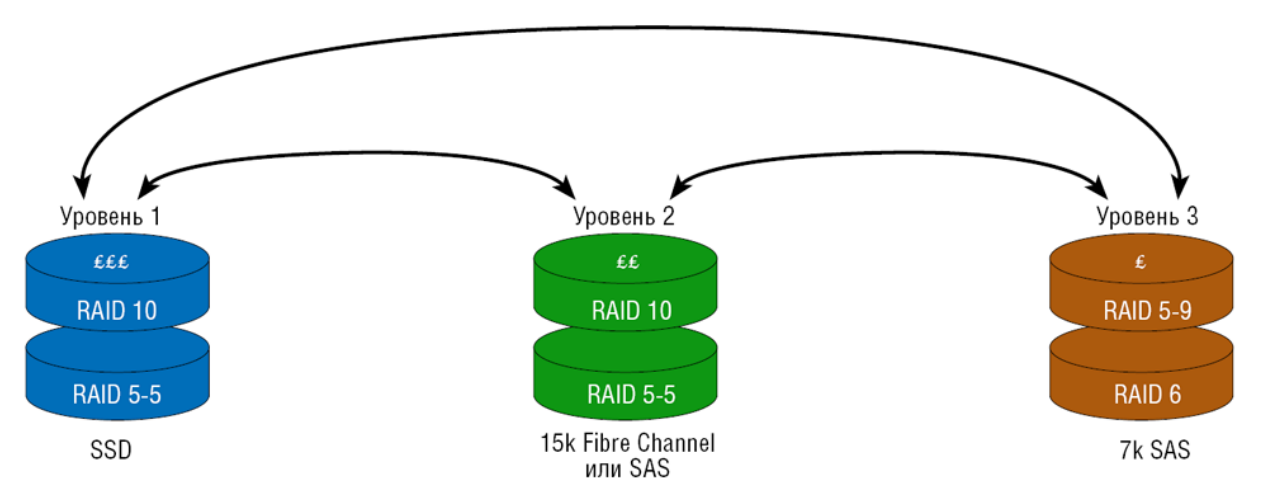


Рис. 2.7. Иерархическая система хранения данных