



Л.Н. Демидов, В.Б. Терновсков  
С.М. Григорьев, Д.В. Крахмалев

# ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Рекомендовано  
Экспертным советом УМО в системе ВО и СПО  
в качестве **учебника** для студентов,  
обучающихся по направлениям подготовки  
«Информационные системы и технологии»,  
«Менеджмент», «Прикладная информатика»

УДК 004(075.8)  
ББК 32.973я73  
Д30

**Рецензенты:**

**В.А. Чобаня**, д-р техн. наук, заслуженный деятель науки и техники РФ, проф.,  
**В.А. Кадымов**, д-р физ.-мат. наук, проф.

**Д30** **Демидов, Лев Николаевич.** Информационные технологии : учебник / Л.Н. Демидов, В.Б. Терновсков, С.М. Григорьев, Д.В. Крахмалев. — Москва : КНОРУС, 2021. — 222 с. — (Бакалавриат).

**ISBN 978-5-406-09076-3**

В учебнике излагаются теоретические основы построения информационных технологий нового направления развития управления как комплексной научной и инженерной дисциплины. Рассмотрены основы информационной культуры, существующие информационные теории, математические основы информатики.

Соответствует ФГОС ВО последнего поколения.

*Для студентов, аспирантов и преподавателей различных специальностей, обучающихся по учебной дисциплине «Информационные технологии».*

**УДК 004(075.8)**  
**ББК 32.973я73**

Демидов Лев Николаевич  
Терновсков Владимир Борисович  
Григорьев Сергей Михайлович  
Крахмалев Дмитрий Владимирович

**ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ**

Изд. № 640711. Формат 60×90/16. Гарнитура «News GothicС».

Усл. печ. л. 14,0. Уч.-изд. л. 13,0. Тираж 500 экз.

ООО «Издательство «КноРус».

117218, г. Москва, ул. Кедрова, д. 14, корп. 2.

Тел.: +7 (495) 741-46-28.

E-mail: welcome@knorus.ru www.knorus.ru

Отпечатано в АО «Т8 Издательские Технологии».

109316, г. Москва, Волгоградский проспект, д. 42, корп. 5.

Тел.: +7 (495) 221-89-80.

**ISBN 978-5-406-09076-3**

© Демидов Л.Н., Терновсков В.Б.,  
Григорьев С.М., Крахмалев Д.В., 2021  
© ООО «Издательство «КноРус», 2021



# Оглавление

Предисловие . . . . .	6
Введение . . . . .	8
Список использованных сокращений . . . . .	12

## **Тема 1. Основные понятия информационных технологий (ИТ).**

### **Роль ИТ в развитии современного общества . . . . . 14**

1.1. Содержание новой информационной технологии как составной части информатики . . . . .	14
1.2. Типы и свойства современных ИТ . . . . .	16
1.2.1. Основные информационные процессы. . . . .	16
1.2.2. Взаимодействие информационных технологий и систем . . . . .	17
1.2.3. Обработка информации . . . . .	19
1.2.4. Классификация информационных технологий . . . . .	22
1.3. Этапы развития информационных технологий . . . . .	26
1.4. Общая классификация видов информационных технологий и их реализация в технических областях . . . . .	29
1.5. Глобальная, базовая и конкретные информационные технологии. . . . .	33
<i>Контрольные вопросы и задания . . . . .</i>	<i>34</i>

## **Тема 2. Информационные процессы в системах.**

### **Информационные процессы обработки и представления данных . . . 35**

2.1. Системный подход к решению функциональных задач и к организации информационных процессов в системах. . . . .	35
----------------------------------------------------------------------------------------------------------------------	----

2.2. Понятие технологического процесса обработки информации. Централизованная и распределенная обработка данных . . . . .	39
2.3. Основные этапы технологического процесса и составляющие их операции . . . . .	48
2.4. Технологии электронного офиса . . . . .	52
2.5. Технологии работы с документами . . . . .	54
2.6. Примеры реализации технологий электронного документооборота . . . . .	68
<i>Контрольные вопросы . . . . .</i>	<i>79</i>

### **Тема 3. Информационный процесс обмена данными.**

#### **Сетевые информационные технологии . . . . .**

3.1. Общие понятия о компьютерных сетях . . . . .	80
3.2. Телеобработка данных в компьютерных сетях . . . . .	92
3.3. Функциональный состав, структура и классификация сетей ЭВМ . . . . .	99
3.4. Открытая модель взаимодействия ISO/OSI . . . . .	105
3.5. Методы структуризации сетей ЭВМ . . . . .	110
3.6. Интернет как инструмент для ИТ . . . . .	117
<i>Контрольные вопросы . . . . .</i>	<i>140</i>

#### **Тема 4. Технологии поисковых систем . . . . .**

4.1. Документальные системы: информационно-поисковый язык, система индексирования, технология обработки данных, поисковый аппарат, критерии оценки документальных систем . . . . .	141
4.2. Компьютерные справочно-информационные системы, справочно-правовые системы . . . . .	156
<i>Контрольные вопросы и задания . . . . .</i>	<i>163</i>

### **Тема 5. Информационный процесс**

#### **накопления и хранения данных . . . . .**

5.1. Основные особенности построения ИТ . . . . .	165
5.2. Основные компоненты технологии . . . . .	167

5.3. Программное обеспечение технологии . . . . .	169
5.4. Техническое обеспечение технологии . . . . .	169
5.5. Хранилища данных . . . . .	172
<i>Контрольные вопросы.</i> . . . . .	189
<b>Тема 6. Технологии защиты данных</b> . . . . .	190
6.1. Информационная безопасность и защита информации . . . . .	190
6.2. Криптографические методы защиты данных. . . . .	191
6.3. Компьютерные вирусы и защита от них. . . . .	197
6.4. Защита информации в компьютерных сетях. . . . .	213
<i>Контрольные вопросы и задания</i> . . . . .	220
Список использованной литературы. . . . .	222



# ПРЕДИСЛОВИЕ

Роль информационных технологий сегодня трудно переоценить. Расширение области использования современных ИТ, ориентированных на базу в виде персональных компьютеров, компьютерных сетей и современных моделей взаимодействия, определяется технологическими возможностями по обработке экономической информации.

Любой человек, использующий такой мощный инструмент — ИТ, должен представлять, что же стоит под названием «информационные технологии». Вариантов построения систем на рынке сегодня масса. Специалисты в области информационных технологий в настоящее время должны обладать широким кругозором философского уровня и глубокими знаниями предметной области. А также достаточным объемом информации из пограничных областей.

В учебнике излагаются теоретические основы информационных технологий как комплексной научной и инженерной дисциплины. Рассмотрены основы информационной культуры, существующие информационные теории, математические основы информатики, криптографические основы защиты информации. А также представлены основные примеры реализации технологий для различных областей.

Учебник состоит из 6 тем, логически связанных друг с другом.

Учебник написан коллективом авторов в составе: Демидов Лев Николаевич, Терновсков Владимир Борисович, Григорьев Сергей Михайлович, Крахмалев Дмитрий Владимирович.

Демидов Лев Николаевич — кандидат технических наук, доцент. Круг научных интересов: компьютерная графика, телекоммуникации, информационные технологии, моделирование систем, бизнес-процессов. Доцент кафедры «Бизнес-информатика» Финансового университета при Правительстве РФ.

Терновсков Владимир Борисович — кандидат технических наук. Круг научных интересов: телекоммуникации, оценка рисков, информационные технологии. Доцент кафедры «Анализа рисков и экономической безопасности» Финансового университета при Правительстве РФ.

Григорьев Сергей Михайлович — кандидат военных наук, доцент. Круг научных интересов: моделирование безопасности систем, бизнес-процессов, оценка рисков, информационная безопасность. Доцент кафедры «Анализа рисков и экономической безопасности» Финансового университета при Правительстве РФ.

Крахмалев Дмитрий Владимирович — кандидат технических наук. Круг научных интересов: компьютерная графика, телекоммуникации, информационные технологии, моделирование систем, бизнес-процессов. Доцент кафедры «Бизнес-информатика» Финансового университета при Правительстве РФ.



# ВВЕДЕНИЕ

Современный мир не представим без применения информационных технологий и компьютерной техники. В особенности персональной и мобильной, использующей последние достижения в области разработки программного обеспечения (ПО). Применение компьютерной техники характеризуется следующими основными чертами:

- 1) доступностью и широким распространением персональных компьютеров (ПК), невероятно возросло число пользователей ЭВМ, в том числе не имеющих даже начальных знаний в области вычислительной техники;
- 2) значительным увеличением и числом пользователей ЭВМ, не имеющих соответствующей базовой подготовки, но мнящих себя чуть ли не гуру в области программирования и программной инженерии;
- 3) имеющимися международными, национальными и государственными стандартами в области разработки ПО, во многом отстающими от потребностей конечного пользователя.

Информация — это не откровение свыше, а достаточно привычное понятие, которое сегодня ассоциируется с еще одним значением, широко используемым как в СМИ, так и в обиходе, — ресурс. Да, информация — это стратегический ресурс каждого государства на земном шаре. Но это еще и ресурс для кропотливой, в какой-то степени рутинной обработки. Сегодня для обработки информации используют ЭВМ. Причем наибольшей эффективности такой обработки можно достичь, только с единым источником информации, работающим через объединяющую сеть с несколькими различными ЭВМ, управляемыми пользователями. Это определяет современный облик того, что мы называем информационные технологии — технологии обработки информации.

Учебник написан с избыточностью. На этот шаг авторы пошли обдуманно, поскольку вынуждены отметить тот факт, что современное по-

коление студентов не всегда охотно ищет ответы на вопросы в иных печатных источниках. Поэтому на страницах учебника приведены дополнительные объясняющие сноски, в большинстве своем взятые из «Википедии» — свободной Интернет-энциклопедии. Часть материала дана поверхностно, несмотря на значительный объем представленного материала. Это объясняется тем, что ряд вопросов (такие как телекоммуникации, сети, информационный менеджмент, управление проектами, реинжиниринг бизнес-процессов и т.п.) в ряде технических вузов изучается в течение двух и более семестров, что определяется сложностью и объемом рассматриваемого материала. В данном учебнике представлено лишь поверхностное их изложение. В результате изучения материала, представленного в данном учебнике, обучаемый должен

**знать:**

- современное состояние и тенденции развития ИТ;
- теоретические основы информационной технологии управления;
- принципы построения современных ИС;
- аппаратно-техническое и программное обеспечение глобальных компьютерных сетей и корпоративных ИС;
- содержание, стадии разработки и результаты выполнения этапов проектирования автоматизированных ИС (АИС);
- роль конечного пользователя в процессе проектирования АИС;
- содержание и цели применения интернет-технологий в деятельности менеджера;
- организацию системы поддержки принятия управленческих решений;
- технологию проектирования и создания баз данных;
- моделирование в рамках интегрированных пакетов;
- моделирование финансово-экономической деятельности предприятия;
- технологический процесс обработки и защиты данных;

**уметь:**

- пользоваться компьютерными системами и технологиями, применяемыми в менеджменте;
- использовать методы, способы и средства получения, хранения, переработки информации в ситуациях принятия решения;
- работать с компьютером как средством управления информацией в глобальных компьютерных сетях и корпоративных ИС;
- использовать программные средства анализа и количественного моделирования систем управления;
- использовать методы и программные средства обработки деловой информации.

**владеть:**

- основными методами, способами и средствами получения, хранения, переработки информации;
- навыками работы с компьютером как средством управления информацией;
- навыками работы в глобальных компьютерных сетях и корпоративных ИС;
- методами и программными средствами обработки деловой информации;
- навыками работы с универсальными пакетами прикладных программ для решения управленческих задач;
- навыками работы со специализированными пакетами программ для управленческих задач принятия решений технологиями систем поддержки принятия решений;
- основными методами хранения, поиска и обработки информации в системах управления базами данных.

Учебник состоит из шести разделов. В них излагаются теоретические основы построения и применения современных ИТ в менеджменте и различных уровнях их применения. В первом разделе рассматриваются главы, касающиеся истории возникновения и развития ИТ, их классификация, рассмотрены свойства ИТ. Во втором разделе рассматриваются взаимодействие ИТ и бизнес-процессов, процессный подход к управлению (менеджменту) и вопросы соотношения и взаимодействия организационных структур, бизнес-процессов и ИТ для использования в процессе управления, а также принципы распределения ИТ. В третьем разделе рассматриваются возможности взаимодействия информационных систем, реализующих ИТ, в том числе корпоративных. В четвертом разделе с достаточной степенью подробности рассмотрен инструментарий, позволяющий реализовать современные ИТ для их использования в менеджменте. В пятом разделе рассмотрены примеры современных ИТ, используемых на всех уровнях менеджмента современных компаний (предприятий, организаций). В шестом разделе рассматриваются начальные положения защиты информации.

Учебник рассчитан на широкую аудиторию и предназначается широкому кругу читателей. В том числе для студентов, аспирантов, проходящих подготовку по направлению обучения «Менеджмент». Книга, возможно, заинтересует «средний» управленческий аппарат современных компаний в различных областях деятельности, а также руководителей различного ранга.

Содержание учебника снабжено наборами вопросов для самостоятельной работы. Эти вопросы носят рекомендательный характер, по-

сколько подразумевается, что вопросы у обучаемых возникают по мере прочтения текста. А также каждый раздел снабжен дополнительными заданиями для самостоятельной отработки для более верного понимания того, что рассматривается в учебнике, в том числе реализованных в виде тестовых заданий. И в прилагаемом к каждой главе списке дополнительной литературы, рекомендуемой для самостоятельного изучения, они найдут ответы на все те вопросы, которые возникнут.



# СПИСОК ИСПОЛЬЗОВАННЫХ СОКРАЩЕНИЙ

АБД	— автоматизированный банк данных.
АВМ	— аналоговая вычислительная машина.
АИТ	— автоматизированная информационная технология.
АРМ	— автоматизированное рабочее место.
АСУ	— автоматизированная система управления.
БД	— база данных.
ВМ	— вычислительная машина.
ВЦ	— вычислительный центр.
ГВМ	— гибридная вычислительная машина.
ГИС	— географическая информационная система.
ИВС	— информационно-вычислительная сеть.
ИВЦ	— информационно-вычислительный центр.
ИПС	— информационно-поисковая система.
ИТ	— информационная технология.
КИС	— корпоративная информационная система.
ЛВС	— локальная вычислительная сеть.
МП	— микропроцессор.
НЖМД	— накопитель на жестком магнитном диске.
ОИТ	— обеспечивающая информационная технология.
ОС	— операционная система.
ПЗУ	— постоянное запоминающее устройство.
ПК	— персональный компьютер.
ПО	— программное обеспечение.
ПЭВМ	— персональная электронно-вычислительная машина.
РС	— рабочая станция.

СУБД	— система управлением базой данных.
ФИТ	— функциональная информационная технология.
ХД	— хранилище данных.
ЦВМ	— цифровая вычислительная машина.
ЭВМ	— электронно-вычислительная машина.
BBS	— Bulletin Board System (системы электронных досок).
CD-ROM	— Compact Disk Read-Only Memory (запоминающее устройство на компакт-диске в режиме «чтение»).
DME	— Document Management Extensions (система управления документами).
DVC	— Digital Video Conference (система видеоконференций).
DVD	— Digital Versalite/Video Disk (цифровой оптический видеодиск).
EDMS	— Electronic Document Management Systems (система управления электронными документами).
GAN	— Global Area Network (глобальная вычислительная сеть).
GIL	— Global Information Infrastructure (глобальная информационная инфраструктура).
GUI	— Graphical User Interface (графический интерфейс пользователя).
ISDN	— Integrated Services Digital Network (цифровая сеть с комплексными услугами).
ISO	— International Organization for Standardization (международная организация по стандартизации).
LAN	— Local Area Network (локальная вычислительная сеть).
MAN	— Metropolitan Area Network (региональная вычислительная сеть).
MIPS	— Million Instructions Per Second (миллион команд в секунду над числами с фиксированной запятой).
MIS	— Management Information System (информационные системы управления).
OLAP	— Online Analytical Processing (оперативная аналитическая обработка).
OLE	— Object Linking and Embedding (связывание и встраивание объектов).
QBE	— Query By Example (запрос по образцу).
SILK	— Speech Image Language Knowledge (интерфейс с речевыми командами).
SQL	— Structured Query Language (язык структурированных запросов).
SWIFT	— Society for Worldwide Interbank Financial Telecommunication (мировая банковская финансовая телекоммуникационная сеть).
UML	— Unified Modeling Language (универсальный язык моделирования).
WIMP	— Windows Image Pointer (графический интерфейс).

# ОСНОВНЫЕ ПОНЯТИЯ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ (ИТ). РОЛЬ ИТ В РАЗВИТИИ СОВРЕМЕННОГО ОБЩЕСТВА

## 1.1. Содержание новой информационной технологии как составной части информатики

Информационная технология является наиболее важной составляющей процесса использования информационных ресурсов общества. К настоящему времени она прошла несколько эволюционных этапов, смена которых определялась главным образом развитием научно-технического прогресса, появлением новых технических средств переработки информации. В современном обществе основным техническим средством технологии переработки информации служит персональный компьютер. Внедрение персонального компьютера в информационную сферу и применение телекоммуникационных средств связи определили новый этап развития информационной технологии и, как следствие, изменение ее названия за счет присоединения одного из синонимов: «новая», «компьютерная» или «современная».

Прилагательное «новая» подчеркивает новаторский, а не эволюционный характер этой технологии. Ее внедрение является новаторским актом в том смысле, что она существенно изменяет содержание раз-

личных видов деятельности в организациях. В понятие новой информационной технологии включены также коммуникационные технологии, которые обеспечивают передачу информации разными средствами, а именно: телефон, телеграф, телекоммуникации, факс и др.

Новая информационная технология — информационная технология с «дружественным» интерфейсом работы пользователя, использующая персональные компьютеры и телекоммуникационные средства. Прилагательное «компьютерная» подчеркивает, что основным техническим средством ее реализации является компьютер.

Три основных принципа новой (компьютерной) информационной технологии:

- интерактивный (диалоговый) режим работы с компьютером;
- интегрированность с другими программными продуктами;
- гибкость процесса изменения как данных, так и постановок задач.

Для эффективного взаимодействия конечных пользователей с вычислительной системой новые информационные технологии опираются на принципиально иную организацию интерфейса пользователей с вычислительной системой (так называемого дружественного интерфейса), который выражается прежде всего в следующем:

- в обеспечении права пользователя на ошибку благодаря защите информационно-вычислительных ресурсов системы от непрофессиональных действий на компьютере;
- наличию широкого набора иерархических меню, системы подсказок и обучения и т.п., облегчающих процесс взаимодействия пользователя с ПК;
- наличию системы «отката», позволяющей при выполнении регламентированного действия, последствия которого по каким-либо причинам не удовлетворили пользователя, вернуться к предыдущему состоянию системы.

По-видимому, более точным следует считать все же термин «новая», а не «компьютерная» информационная технология, поскольку он отражает в ее структуре не только технологии, основанные на использовании компьютеров, но и технологии, основанные на других технических средствах, особенно на средствах, обеспечивающих телекоммуникацию. Для информационных технологий является вполне естественным то, что они устаревают и заменяются новыми. Например, на смену технологии пакетной обработки программ на большой ЭВМ в вычислительном центре пришла технология работы на персональном компьютере на рабочем месте пользователя. Телеграф передал все свои функции телефону и т.д.

При внедрении новой информационной технологии в организации необходимо оценить риск отставания от конкурентов в результате ее неизбежного устаревания со временем, так как информационные продукты, как никакие другие виды материальных товаров, имеют чрезвычайно высокую скорость сменяемости новыми видами, версиями. Периоды сменяемости колеблются от нескольких месяцев до одного года. Если в процессе внедрения новой информационной технологии этому фактору не уделять должного внимания, возможно, что к моменту внедрения новой информационной технологии она уже устареет и придется принимать меры к ее модернизации. Основной причиной неудач является отсутствие или слабая проработанность методологии использования информационной технологии.

## 1.2. Типы и свойства современных ИТ

### 1.2.1. Основные информационные процессы

Процессы преобразования информации связаны с информационными технологиями. Технология в переводе с греческого — искусство, умение, а это не что иное, как процесс. Процесс — определенная совокупность действий, направленных на достижение поставленных целей.

Технология материального производства определяется как совокупность методов обработки, изготовления, изменения состояния, свойств, формы сырья, материала или полуфабриката в процессе производства (например, технология металлов, химическая технология, технология строительства).

**Цель технологии материального производства** — выпуск продукции, удовлетворяющей потребности человека или системы. Информационная технология — система взаимосвязанных методов и способов сбора, хранения, накопления, поиска, обработки информации на основе применения средств вычислительной техники. **Цель информационной технологии (ИТ)** — производство информации для анализа человеком и принятие на его основе решения по выполнению какого-либо действия (управленческого решения). Особенностью ИТ является то, что в ней и предметом и продуктом труда является информация, а орудиями труда — средства вычислительной техники и связи.

Основные принципы компьютерных (новых) информационных технологий:

- интерактивный режим (диалоговый или режим реального времени) работы с ПК;
- интегрированность (стыковка);

- гибкость процесса изменения как данных, так и постановок задач.
- Требования к информационным технологиям:
- малая стоимость, находящаяся в пределах доступности для индивидуального покупателя;
- автономность в эксплуатации без специальных требований к условиям окружающей среды;
- гибкость архитектуры, обеспечивающая ее адаптивность к разнообразным сферам применения: в управлении, науке, образовании, в быту;
- «дружественность» операционной системы и прочего программного обеспечения, обуславливающая работу с ней пользователя без специальной профессиональной подготовки;
- высокая надежность работы (более 8000 часов наработки на отказ).

### 1.2.2. Взаимодействие информационных технологий и систем

Успешное внедрение информационных технологий связано с возможностью их типизации. Конкретная информационная технология обладает комплексным составом компонентов, поэтому целесообразно определить ее структуру и состав.

Конкретная информационная технология определяется в результате компиляции и синтеза базовых технологических операций, специализированных технологий и средств реализации.

Технологический процесс — часть информационного процесса, содержащая действия (физические, механические и др.) по изменению состояния информации.

Информационная технология базируется на реализации информационных процессов, разнообразие которых требует выделения базовых, характерных для любой информационной технологии.

Базовый технологический процесс основан на использовании стандартных моделей и инструментальных средств и может быть использован в качестве составной части информационной технологии. К их числу можно отнести: операции извлечения, транспортировки, хранения, обработки и представления информации.

Среди базовых технологических процессов выделим:

- извлечение информации;
- транспортирование информации;
- обработку информации;

- хранение информации;
- представление<sup>1</sup> и использование информации.

Процесс извлечения информации связан с переходом от реального представления предметной области к его описанию в формальном виде и в виде данных, которые отражают это представление.

В процессе транспортирования осуществляют передачу информации на расстояние для ускоренного обмена и организации быстрого доступа к ней, используя при этом различные способы преобразования.

Процесс обработки информации состоит в получении одних «информационных объектов» из других «информационных объектов», путем выполнения некоторых алгоритмов; он является одной из основных операций, выполняемых над информацией и главным средством увеличения ее объема и разнообразия.

Процесс хранения связан с необходимостью накопления и длительного хранения данных, обеспечением их актуальности, целостности, безопасности, доступности. Процесс представления и использования информации направлен на решение задачи доступа к информации в удобной для пользователя форме.

Базовые информационные технологии строятся на основе базовых технологических операций, но, кроме этого, включают ряд специфических моделей и инструментальных средств. Этот вид технологий ориентирован на решение определенного класса задач и используется в конкретных технологиях в виде отдельной компоненты. Специфика конкретной предметной области находит отражение в специализированных информационных технологиях, например, организационное управление, управление технологическими процессами, автоматизированное проектирование, обучение и др. Среди них наиболее продвинутыми являются следующие информационные технологии:

- организационного управления (корпоративные информационные технологии);
- в промышленности и экономике;
- в образовании;
- автоматизированного проектирования.

Аналогом инструментальной базы (оборудование, станки, инструмент) являются средства реализации информационных технологий, которые можно разделить на методические, информационные, математические, алгоритмические, технические и программные.

---

<sup>1</sup> Данный термин не вполне корректен при применении его к технологиям, методам и способам обработки информации. Более корректно было бы сказать «отображение информации».

### 1.2.3. Обработка информации

Понятие обработки информации является весьма широким. Ведя речь об обработке информации, следует дать понятие инварианта обработки. Обычно им является смысл сообщения (смысл информации, заключенной в сообщении). При автоматизированной обработке информации объектом обработки служит сообщение, и здесь важно провести обработку таким образом, чтобы инварианты преобразований сообщения соответствовали инвариантам преобразования информации.

Цель обработки информации в целом определяется целью функционирования некоторой системы, с которой связан рассматриваемый информационный процесс. Однако для достижения цели всегда приходится решать ряд взаимосвязанных задач.

Обратимся теперь к вопросу о том, в чем сходство и различие процессов обработки информации, связанных с различными составляющими информационного процесса, используя при этом формализованную модель обработки. Прежде всего нельзя отрывать этот вопрос от потребителя информации (адресата), а также от семантического и прагматического аспектов информации. Наличие адресата, для которого предназначено сообщение (сигнал), определяет невозможность установления однозначного соответствия между сообщением и содержащейся в нем информацией. Совершенно очевидно, что одно и то же сообщение может иметь различный смысл для разных адресатов и различное прагматическое значение.

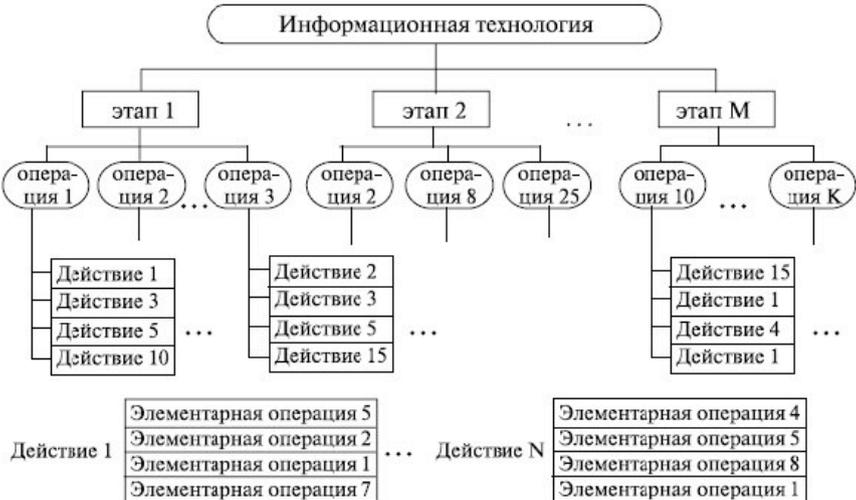
Используемые в производственной сфере такие технологические понятия, как технологический процесс, технологическая операция, метрика, норматив и т.п., могут применяться и в ИТ. Для этого нужно начинать с определения цели. Затем следует попытаться провести структурирование всех предполагаемых действий, приводящих к намеченной цели, и выбрать необходимый программный инструментарий (рис. 1.1).

1-й уровень — этапы, где реализуются базовые технологические процессы, состоящие из операций и действий последующих уровней.

2-й уровень — операции, в результате выполнения которых будет создан конкретный объект в выбранной на 1-м уровне программной среде.

3-й уровень — действия, совокупность стандартных для каждой программной среды приемов работы, приводящих к выполнению поставленной в соответствующей операции цели.

4-й уровень — элементарные операции по управлению элементарными действиями объектов.



**Рис. 1.1.** Технологический процесс переработки информации в виде иерархической структуры по уровням

ИТ, как и другие технологии, должны отвечать следующим требованиям:

- обеспечивать высокую степень расчленения всего процесса обработки информации на этапы, операции, действия;
- включать весь набор элементов, необходимых для достижения поставленной цели;
- иметь регулярный и масштабируемый характер;
- этапы, действия, операции технологического процесса должны быть стандартизированы и унифицированы, что позволит более эффективно осуществлять целенаправленное управление информационными процессами.

Традиционно в процессе обработки информации используются как измерительная аппаратура, обеспечивающая входные данные, так и собственно обрабатывающие (вычислительные) системы. И те, и другие прошли длинную дорогу развития вместе с человеческой цивилизацией. В следующем пункте будут перечислены основные вехи их истории.

Если раньше обрабатывающей системой был человек или какие-то механические приспособления, то для проведения процесса обработки было достаточно сформулировать набор правил (инструкций). Давно подметили, что повторяющиеся операции целесообразно автоматизиро-

вать в первую очередь и желательно порекомендовать машинам. При этом человек, задавая циклическое правило работы машине, колоссально выигрывает в трудозатратах.

Предположим, вам надо сложить 1000 последовательных данных измерений. Заводим специальный счетчик-сумматор и присваиваем ему значение 0. Для каждого из данных надо получить результат измерений и добавить его к счетчику, т.е. вам надо сделать 2001 операцию при «ручном» счете. Другой вариант — написать шесть инструкций для машины:

- 1) завести счетчик-сумматор и присвоить ему значение 0;
- 2) завести индекс (номер) текущей операции и присвоить ему значение 0;
- 3) получить новый результат измерений;
- 4) добавить его к счетчику-сумматору;
- 5) увеличить на 1 индекс текущей операции;
- 6) если он меньше 1000, то перейти к шагу 3.

За прошедшее время существенно усложнились задачи обработки информации, развились способы формулировки и записи правил работы машин (программ работы). Вычислительные устройства превратились в компьютеры, а правила работы — в компьютерные программы.

Программирование — процесс создания компьютерных программ с помощью языков программирования. Программирование сочетает в себе элементы искусства, науки, математики и инженерии.

В узком смысле слова программирование рассматривается как кодирование — реализация одного или нескольких взаимосвязанных алгоритмов на некотором языке программирования. Под программированием также может пониматься разработка логической схемы для интегральной микросхемы, а также процесс записи информации в микросхему ПЗУ (постоянного запоминающего устройства) некоторой электронной системы. В более широком смысле программирование — процесс создания программ, т.е. разработка программного обеспечения.

Составителями программ являются программисты. Большая часть работы программиста связана с написанием и отладкой исходного кода на одном из языков программирования.

Различные языки программирования поддерживают различные стили программирования (или парадигмы программирования). Отчасти искусство программирования состоит в том, чтобы на одном из языков эффективно реализовать алгоритм, наиболее полно подходящий для решения имеющейся задачи. Разные языки требуют от программиста различного уровня внимания к деталям при реализации алгоритма, результатом чего часто бывает компромисс между простотой и произ-

водительностью (или между временем программиста и временем пользователя).

Единственный язык, напрямую выполняемый процессором, — это машинный язык (также называемый машинный код). Изначально все программисты прорабатывали весь алгоритм в машинном коде, но сейчас эта трудная работа уже не делается. Вместо этого программисты пишут исходный код на языке высокого уровня (например, C, C++, C#, Java), а компьютер, используя компилятор или интерпретатор и уточняя все детали, транслирует его за один или несколько этапов в машинный код, готовый к исполнению на целевом процессоре. Если требуется полный низкоуровневый контроль над системой, программисты пишут программу на языке ассемблера, мнемонические инструкции которого преобразуются один к одному в соответствующие инструкции машинного языка целевого процессора.

В некоторых языках вместо машинного кода генерируется интерпретируемый двоичный код «виртуальной машины», также называемый байт-кодом (byte-code). Такой подход применяется в языке Forth, некоторых реализациях языков Lisp, Java, Perl, Python, а также в языках платформы Microsoft.NET.

Типичный процесс разработки программ состоит, в общем, из семи этапов:

- постановка задачи;
- формализация и специфицирование;
- выбор или составление алгоритма;
- программирование;
- компиляция (трансляция);
- отладка и тестирование;
- запуск в эксплуатацию.

Эксплуатируемая программа имеет дело с данными различных типов, предназначенных для решения конкретных задач.

### 1.2.4. Классификация информационных технологий

В настоящее время в большинстве источников информации классификацию информационных технологий (ИТ) принято проводить по следующим признакам:

- способу реализации в автоматизированных информационных системах (АИС);
- степени охвата задач управления;
- классам реализуемых технологических операций;
- типу пользовательского интерфейса;

УДК 004(075.8)  
ББК 32.973я73  
Д30

**Рецензенты:**

**В.А. Чобанян**, д-р техн. наук, заслуженный деятель науки и техники РФ, проф.,  
**В.А. Кадымов**, д-р физ.-мат. наук, проф.

**Демидов, Лев Николаевич.**

**Д30** Информационные технологии : учебник / Л.Н. Демидов, В.Б. Терновсков, С.М. Григорьев, Д.В. Крахмалев. — Москва : КНОРУС, 2020. — 222 с. — (Бакалавриат).

**ISBN 978-5-406-07568-5**

В учебнике излагаются теоретические основы построения информационных технологий нового направления развития управления как комплексной научной и инженерной дисциплины. Рассмотрены основы информационной культуры, существующие информационные теории, математические основы информатики.

Соответствует ФГОС ВО последнего поколения.

*Для студентов, аспирантов и преподавателей различных специальностей, обучающихся по учебной дисциплине «Информационные технологии».*

**УДК 004(075.8)**

**ББК 32.973я73**

Демидов Лев Николаевич  
Терновсков Владимир Борисович  
Григорьев Сергей Михайлович  
Крахмалев Дмитрий Владимирович

**ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ**

Изд. № 518207. Формат 60×90/16. Гарнитура «News GothicС».  
Усл. печ. л. 14,0. Уч.-изд. л. 13,0.

ООО «Издательство «КноРус»,  
117218, г. Москва, ул. Кедрова, д. 14, корп. 2.  
Тел.: 8-495-741-46-28.

E-mail: office@klorus.ru <http://www.klorus.ru>

Отпечатано в АО «Т8 Издательские Технологии»,  
109316, г. Москва, Волгоградский проспект, д. 42, корп. 5.  
Тел.: 8-495-221-89-80.

© Демидов Л.Н., Терновсков В.Б.,  
Григорьев С.М., Крахмалев Д.В., 2020  
© ООО «Издательство «КноРус», 2020

**ISBN 978-5-406-07568-5**



# Оглавление

Предисловие .....	6
Введение .....	8
Список использованных сокращений .....	12

## **Тема 1. Основные понятия информационных технологий (ИТ).**

### **Роль ИТ в развитии современного общества .....**

1.1. Содержание новой информационной технологии как составной части информатики .....	14
1.2. Типы и свойства современных ИТ .....	16
1.2.1. Основные информационные процессы .....	16
1.2.2. Взаимодействие информационных технологий и систем ..	17
1.2.3. Обработка информации .....	19
1.2.4. Классификация информационных технологий .....	22
1.3. Этапы развития информационных технологий .....	26
1.4. Общая классификация видов информационных технологий и их реализация в технических областях .....	29
1.5. Глобальная, базовая и конкретные информационные технологии .....	33
<i>Контрольные вопросы и задания .....</i>	<i>34</i>

## **Тема 2. Информационные процессы в системах.**

### **Информационные процессы обработки и представления данных .....**

2.1. Системный подход к решению функциональных задач и к организации информационных процессов в системах .....	35
-------------------------------------------------------------------------------------------------------------------	----

2.2. Понятие технологического процесса обработки информации. Централизованная и распределенная обработка данных . . . . .	39
2.3. Основные этапы технологического процесса и составляющие их операции. . . . .	48
2.4. Технологии электронного офиса . . . . .	52
2.5. Технологии работы с документами . . . . .	54
2.6. Примеры реализации технологий электронного документооборота. . . . .	68
<i>Контрольные вопросы.</i> . . . . .	79

### **Тема 3. Информационный процесс обмена данными.**

<b>Сетевые информационные технологии</b> . . . . .	80
3.1. Общие понятия о компьютерных сетях . . . . .	80
3.2. Телеобработка данных в компьютерных сетях . . . . .	92
3.3. Функциональный состав, структура и классификация сетей ЭВМ . . . . .	99
3.4. Открытая модель взаимодействия ISO/OSI . . . . .	105
3.5. Методы структуризации сетей ЭВМ . . . . .	110
3.6. Интернет как инструмент для ИТ. . . . .	117
<i>Контрольные вопросы.</i> . . . . .	140

### **Тема 4. Технологии поисковых систем** . . . . . 141

4.1. Документальные системы: информационно-поисковый язык, система индексирования, технология обработки данных, поисковый аппарат, критерии оценки документальных систем. . . . .	141
4.2. Компьютерные справочно-информационные системы, справочно-правовые системы. . . . .	156
<i>Контрольные вопросы и задания</i> . . . . .	163

### **Тема 5. Информационный процесс накопления и хранения данных.** . . . . . 164

5.1. Основные особенности построения ИТ. . . . .	165
5.2. Основные компоненты технологии . . . . .	167

СИО

ИЗДАНИЕ

Учебное пособие  
для студентов

В.А. Брунцов

# БАЗЫ ДАННЫХ

ИЗДАТЕЛЬСТВО

ИЗДАТЕЛЬСТВО

## ПРЕДИСЛОВИЕ

Роль информационных технологий сегодня трудно переоценить. Расширение области использования современных ИТ, ориентированных на базу в виде персональных компьютеров, компьютерных сетей и современных моделей взаимодействия, определяется технологическими возможностями по обработке экономической информации.

Любой человек, использующий такой мощный инструмент — ИТ, должен представлять, что же стоит под названием «информационные технологии». Вариантов построения систем на рынке сегодня масса. Специалисты в области информационных технологий в настоящее время должны обладать широким кругозором философского уровня и глубокими знаниями предметной области. А также достаточным объемом информации из пограничных областей.

В учебнике излагаются теоретические основы информационных технологий как комплексной научной и инженерной дисциплины. Рассмотрены основы информационной культуры, существующие информационные теории, математические основы информатики, криптографические основы защиты информации. А также представлены основные примеры реализации технологий для различных областей.

Учебник состоит из 6 тем, логически связанных друг с другом.

Учебник написан коллективом авторов в составе: Демидов Лев Николаевич, Терновсков Владимир Борисович, Григорьев Сергей Михайлович, Крахмалев Дмитрий Владимирович.

Демидов Лев Николаевич — кандидат технических наук, доцент. Круг научных интересов: компьютерная графика, телекоммуникации, информационные технологии, моделирование систем, бизнес-процессов. Доцент кафедры «Бизнес-информатика» Финансового университета при Правительстве РФ.

Терновсков Владимир Борисович — кандидат технических наук. Круг научных интересов: телекоммуникации, оценка рисков, информационные технологии. Доцент кафедры «Анализа рисков и экономической безопасности» Финансового университета при Правительстве РФ.

Григорьев Сергей Михайлович — кандидат военных наук, доцент. Круг научных интересов: моделирование безопасности систем, бизнес-процессов, оценка рисков, информационная безопасность. Доцент кафедры «Анализа рисков и экономической безопасности» Финансового университета при Правительстве РФ.

Крахмалев Дмитрий Владимирович — кандидат технических наук. Круг научных интересов: компьютерная графика, телекоммуникации, информационные технологии, моделирование систем, бизнес-процессов. Доцент кафедры «Бизнес-информатика» Финансового университета при Правительстве РФ.

# ВВЕДЕНИЕ

Современный мир не представим без применения информационных технологий и компьютерной техники. В особенности персональной и мобильной, использующей последние достижения в области разработки программного обеспечения (ПО). Применение компьютерной техники характеризуется следующими основными чертами:

- 1) доступностью и широким распространением персональных компьютеров (ПК), невероятно возросло число пользователей ЭВМ, в том числе не имеющих даже начальных знаний в области вычислительной техники;
- 2) значительным увеличением и числом пользователей ЭВМ, не имеющих соответствующей базовой подготовки, но мнящих себя чуть ли не гуру в области программирования и программной инженерии;
- 3) имеющимися международными, национальными и государственными стандартами в области разработки ПО, во многом отстающими от потребностей конечного пользователя.

Информация — это не откровение свыше, а достаточно привычное понятие, которое сегодня ассоциируется с еще одним значением, широко используемым как в СМИ, так и в обиходе, — ресурс. Да, информация — это стратегический ресурс каждого государства на земном шаре. Но это еще и ресурс для кропотливой, в какой-то степени рутинной обработки. Сегодня для обработки информации используют ЭВМ. Причем наибольшей эффективности такой обработки можно достичь, только с единым источником информации, работающим через объединяющую сеть с несколькими различными ЭВМ, управляемыми пользователями. Это определяет современный облик того, что мы называем информационные технологии — технологии обработки информации.

Учебник написан с избыточностью. На этот шаг авторы пошли обдуманно, поскольку вынуждены отметить тот факт, что современное по-

коление студентов не всегда охотно ищет ответы на вопросы в иных печатных источниках. Поэтому на страницах учебника приведены дополнительные объясняющие сноски, в большинстве своем взятые из «Википедии» — свободной Интернет-энциклопедии. Часть материала дана поверхностно, несмотря на значительный объем представленного материала. Это объясняется тем, что ряд вопросов (такие как телекоммуникации, сети, информационный менеджмент, управление проектами, реинжиниринг бизнес-процессов и т.п.) в ряде технических вузов изучается в течение двух и более семестров, что определяется сложностью и объемом рассматриваемого материала. В данном учебнике представлено лишь поверхностное их изложение. В результате изучения материала, представленного в данном учебнике, обучаемый должен

**знать:**

- современное состояние и тенденции развития ИТ;
- теоретические основы информационной технологии управления;
- принципы построения современных ИС;
- аппаратно-техническое и программное обеспечение глобальных компьютерных сетей и корпоративных ИС;
- содержание, стадии разработки и результаты выполнения этапов проектирования автоматизированных ИС (АИС);
- роль конечного пользователя в процессе проектирования АИС;
- содержание и цели применения интернет-технологий в деятельности менеджера;
- организацию системы поддержки принятия управленческих решений;
- технологию проектирования и создания баз данных;
- моделирование в рамках интегрированных пакетов;
- моделирование финансово-экономической деятельности предприятия;
- технологический процесс обработки и защиты данных;

**уметь:**

- пользоваться компьютерными системами и технологиями, применяемыми в менеджменте;
- использовать методы, способы и средства получения, хранения, переработки информации в ситуациях принятия решения;
- работать с компьютером как средством управления информацией в глобальных компьютерных сетях и корпоративных ИС;
- использовать программные средства анализа и количественного моделирования систем управления;
- использовать методы и программные средства обработки деловой информации.

**владеть:**

- основными методами, способами и средствами получения, хранения, переработки информации;
- навыками работы с компьютером как средством управления информацией;
- навыками работы в глобальных компьютерных сетях и корпоративных ИС;
- методами и программными средствами обработки деловой информации;
- навыками работы с универсальными пакетами прикладных программ для решения управленческих задач;
- навыками работы со специализированными пакетами программ для управленческих задач принятия решений технологиями систем поддержки принятия решений;
- основными методами хранения, поиска и обработки информации в системах управления базами данных.

Учебник состоит из шести разделов. В них излагаются теоретические основы построения и применения современных ИТ в менеджменте и различных уровнях их применения. В первом разделе рассматриваются главы, касающиеся истории возникновения и развития ИТ, их классификация, рассмотрены свойства ИТ. Во втором разделе рассматриваются взаимодействие ИТ и бизнес-процессов, процессный подход к управлению (менеджменту) и вопросы соотношения и взаимодействия организационных структур, бизнес-процессов и ИТ для использования в процессе управления, а также принципы распределения ИТ. В третьем разделе рассматриваются возможности взаимодействия информационных систем, реализующих ИТ, в том числе корпоративных. В четвертом разделе с достаточной степенью подробности рассмотрен инструментарий, позволяющий реализовать современные ИТ для их использования в менеджменте. В пятом разделе рассмотрены примеры современных ИТ, используемых на всех уровнях менеджмента современных компаний (предприятий, организаций). В шестом разделе рассматриваются начальные положения защиты информации.

Учебник рассчитан на широкую аудиторию и предназначается широкому кругу читателей. В том числе для студентов, аспирантов, проходящих подготовку по направлению обучения «Менеджмент». Книга, возможно, заинтересует «средний» управленческий аппарат современных компаний в различных областях деятельности, а также руководителей различного ранга.

Содержание учебника снабжено наборами вопросов для самостоятельной работы. Эти вопросы носят рекомендательный характер, по-

**владеть:**

- основными методами, способами и средствами получения, хранения, переработки информации;
- навыками работы с компьютером как средством управления информацией;
- навыками работы в глобальных компьютерных сетях и корпоративных ИС;
- методами и программными средствами обработки деловой информации;
- навыками работы с универсальными пакетами прикладных программ для решения управленческих задач;
- навыками работы со специализированными пакетами программ для управленческих задач принятия решений технологиями систем поддержки принятия решений;
- основными методами хранения, поиска и обработки информации в системах управления базами данных.

Учебник состоит из шести разделов. В них излагаются теоретические основы построения и применения современных ИТ в менеджменте и различных уровнях их применения. В первом разделе рассматриваются главы, касающиеся истории возникновения и развития ИТ, их классификация, рассмотрены свойства ИТ. Во втором разделе рассматриваются взаимодействие ИТ и бизнес-процессов, процессный подход к управлению (менеджменту) и вопросы соотношения и взаимодействия организационных структур, бизнес-процессов и ИТ для использования в процессе управления, а также принципы распределения ИТ. В третьем разделе рассматриваются возможности взаимодействия информационных систем, реализующих ИТ, в том числе корпоративных. В четвертом разделе с достаточной степенью подробности рассмотрен инструментарий, позволяющий реализовать современные ИТ для их использования в менеджменте. В пятом разделе рассмотрены примеры современных ИТ, используемых на всех уровнях менеджмента современных компаний (предприятий, организаций). В шестом разделе рассматриваются начальные положения защиты информации.

Учебник рассчитан на широкую аудиторию и предназначается широкому кругу читателей. В том числе для студентов, аспирантов, проходящих подготовку по направлению обучения «Менеджмент». Книга, возможно, заинтересует «средний» управленческий аппарат современных компаний в различных областях деятельности, а также руководителей различного ранга.

Содержание учебника снабжено наборами вопросов для самостоятельной работы. Эти вопросы носят рекомендательный характер, по-



## СПИСОК ИСПОЛЬЗОВАННЫХ СОКРАЩЕНИЙ

АБД	— автоматизированный банк данных.
АВМ	— аналоговая вычислительная машина.
АИТ	— автоматизированная информационная технология.
АРМ	— автоматизированное рабочее место.
АСУ	— автоматизированная система управления.
БД	— база данных.
ВМ	— вычислительная машина.
ВЦ	— вычислительный центр.
ГВМ	— гибридная вычислительная машина.
ГИС	— географическая информационная система.
ИВС	— информационно-вычислительная сеть.
ИВЦ	— информационно-вычислительный центр.
ИПС	— информационно-поисковая система.
ИТ	— информационная технология.
КИС	— корпоративная информационная система.
ЛВС	— локальная вычислительная сеть.
МП	— микропроцессор.
НЖМД	— накопитель на жестком магнитном диске.
ОИТ	— обеспечивающая информационная технология.
ОС	— операционная система.
ПЗУ	— постоянное запоминающее устройство.
ПК	— персональный компьютер.
ПО	— программное обеспечение.
ПЭВМ	— персональная электронно-вычислительная машина.
РС	— рабочая станция.

СУБД	— система управления базой данных.
ФИТ	— функциональная информационная технология.
ХД	— хранилище данных.
ЦВМ	— цифровая вычислительная машина.
ЭВМ	— электронно-вычислительная машина.
BBS	— Bulletin Board System (системы электронных досок).
CD-ROM	— Compact Disk Read-Only Memory (запоминающее устройство на компакт-диске в режиме «чтение»).
DME	— Document Management Extensions (система управления документами).
DVC	— Digital Video Conference (система видеоконференций).
DVD	— Digital Versalite/Video Disk (цифровой оптический видеодиск).
EDMS	— Electronic Document Management Systems (система управления электронными документами).
GAN	— Global Area Network (глобальная вычислительная сеть).
GIL	— Global Information Infrastructure (глобальная информационная инфраструктура).
GUI	— Graphical User Interface (графический интерфейс пользователя).
ISDN	— Integrated Services Digital Network (цифровая сеть с комплексными услугами).
ISO	— International Organization for Standardization (международная организация по стандартизации).
LAN	— Local Area Network (локальная вычислительная сеть).
MAN	— Metropolitan Area Network (региональная вычислительная сеть).
MIPS	— Million Instructions Per Second (миллион команд в секунду над числами с фиксированной запятой).
MIS	— Management Information System (информационные системы управления).
OLAP	— Online Analytical Processing (оперативная аналитическая обработка).
OLE	— Object Linking and Embedding (связывание и встраивание объектов).
QBE	— Query By Example (запрос по образцу).
SILK	— Speech Image Language Knowledge (интерфейс с речевыми командами).
SQL	— Structured Query Language (язык структурированных запросов).
SWIFT	— Society for Worldwide Interbank Financial Telecommunication (мировая банковская финансовая телекоммуникационная сеть).
UML	— Unified Modeling Language (универсальный язык моделирования).
WIMP	— Windows Image Pointer (графический интерфейс).

# ОСНОВНЫЕ ПОНЯТИЯ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ (ИТ). РОЛЬ ИТ В РАЗВИТИИ СОВРЕМЕННОГО ОБЩЕСТВА

## 1.1. Содержание новой информационной технологии как составной части информатики

Информационная технология является наиболее важной составляющей процесса использования информационных ресурсов общества. К настоящему времени она прошла несколько эволюционных этапов, смена которых определялась главным образом развитием научно-технического прогресса, появлением новых технических средств переработки информации. В современном обществе основным техническим средством технологии переработки информации служит персональный компьютер. Внедрение персонального компьютера в информационную сферу и применение телекоммуникационных средств связи определили новый этап развития информационной технологии и, как следствие, изменение ее названия за счет присоединения одного из синонимов: «новая», «компьютерная» или «современная».

Прилагательное «новая» подчеркивает новаторский, а не эволюционный характер этой технологии. Ее внедрение является новаторским актом в том смысле, что она существенно изменяет содержание раз-

личных видов деятельности в организациях. В понятие новой информационной технологии включены также коммуникационные технологии, которые обеспечивают передачу информации разными средствами, а именно: телефон, телеграф, телекоммуникации, факс и др.

Новая информационная технология — информационная технология с «дружественным» интерфейсом работы пользователя, использующая персональные компьютеры и телекоммуникационные средства. Прилагательное «компьютерная» подчеркивает, что основным техническим средством ее реализации является компьютер.

Три основных принципа новой (компьютерной) информационной технологии:

- интерактивный (диалоговый) режим работы с компьютером;
- интегрированность с другими программными продуктами;
- гибкость процесса изменения как данных, так и постановок задач.

Для эффективного взаимодействия конечных пользователей с вычислительной системой новые информационные технологии опираются на принципиально иную организацию интерфейса пользователей с вычислительной системой (так называемого дружественного интерфейса), который выражается прежде всего в следующем:

- в обеспечении права пользователя на ошибку благодаря защите информационно-вычислительных ресурсов системы от непрофессиональных действий на компьютере;
- наличию широкого набора иерархических меню, системы подсказок и обучения и т.п., облегчающих процесс взаимодействия пользователя с ПК;
- наличию системы «отката», позволяющей при выполнении регламентированного действия, последствия которого по каким-либо причинам не удовлетворили пользователя, вернуться к предыдущему состоянию системы.

По-видимому, более точным следует считать все же термин «новая», а не «компьютерная» информационная технология, поскольку он отражает в ее структуре не только технологии, основанные на использовании компьютеров, но и технологии, основанные на других технических средствах, особенно на средствах, обеспечивающих телекоммуникацию. Для информационных технологий является вполне естественным то, что они устаревают и заменяются новыми. Например, на смену технологии пакетной обработки программ на большой ЭВМ в вычислительном центре пришла технология работы на персональном компьютере на рабочем месте пользователя. Телеграф передал все свои функции телефону и т.д.

При внедрении новой информационной технологии в организации необходимо оценить риск отставания от конкурентов в результате ее неизбежного устаревания со временем, так как информационные продукты, как никакие другие виды материальных товаров, имеют чрезвычайно высокую скорость сменяемости новыми видами, версиями. Периоды сменяемости колеблются от нескольких месяцев до одного года. Если в процессе внедрения новой информационной технологии этому фактору не уделять должного внимания, возможно, что к моменту внедрения новой информационной технологии она уже устареет и придется принимать меры к ее модернизации. Основной причиной неудач является отсутствие или слабая проработанность методологии использования информационной технологии.

## 1.2. Типы и свойства современных ИТ

### 1.2.1. Основные информационные процессы

Процессы преобразования информации связаны с информационными технологиями. Технология в переводе с греческого — искусство, умение, а это не что иное, как процесс. Процесс — определенная совокупность действий, направленных на достижение поставленных целей.

Технология материального производства определяется как совокупность методов обработки, изготовления, изменения состояния, свойств, формы сырья, материала или полуфабриката в процессе производства (например, технология металлов, химическая технология, технология строительства).

**Цель технологии материального производства** — выпуск продукции, удовлетворяющей потребности человека или системы. Информационная технология — система взаимосвязанных методов и способов сбора, хранения, накопления, поиска, обработки информации на основе применения средств вычислительной техники. **Цель информационной технологии** (ИТ) — производство информации для анализа человеком и принятие на его основе решения по выполнению какого-либо действия (управленческого решения). Особенностью ИТ является то, что в ней и предметом и продуктом труда является информация, а орудиями труда — средства вычислительной техники и связи.

Основные принципы компьютерных (новых) информационных технологий:

- интерактивный режим (диалоговый или режим реального времени) работы с ПК;
- интегрированность (стыковка);

- гибкость процесса изменения как данных, так и постановок задач.
- Требования к информационным технологиям:
- малая стоимость, находящаяся в пределах доступности для индивидуального покупателя;
- автономность в эксплуатации без специальных требований к условиям окружающей среды;
- гибкость архитектуры, обеспечивающая ее адаптивность к разнообразным сферам применения: в управлении, науке, образовании, в быту;
- «дружелюбность» операционной системы и прочего программного обеспечения, обуславливающая работу с ней пользователя без специальной профессиональной подготовки;
- высокая надежность работы (более 8000 часов наработки на отказ).

### 1.2.2. Взаимодействие информационных технологий и систем

Успешное внедрение информационных технологий связано с возможностью их типизации. Конкретная информационная технология обладает комплексным составом компонентов, поэтому целесообразно определить ее структуру и состав.

Конкретная информационная технология определяется в результате компиляции и синтеза базовых технологических операций, специализированных технологий и средств реализации.

Технологический процесс — часть информационного процесса, содержащая действия (физические, механические и др.) по изменению состояния информации.

Информационная технология базируется на реализации информационных процессов, разнообразие которых требует выделения базовых, характерных для любой информационной технологии.

Базовый технологический процесс основан на использовании стандартных моделей и инструментальных средств и может быть использован в качестве составной части информационной технологии. К их числу можно отнести: операции извлечения, транспортировки, хранения, обработки и представления информации.

Среди базовых технологических процессов выделим:

- извлечение информации;
- транспортирование информации;
- обработку информации;

- хранение информации;
- представление<sup>1</sup> и использование информации.

Процесс извлечения информации связан с переходом от реального представления предметной области к его описанию в формальном виде и в виде данных, которые отражают это представление.

В процессе транспортирования осуществляют передачу информации на расстояние для ускоренного обмена и организации быстрого доступа к ней, используя при этом различные способы преобразования.

Процесс обработки информации состоит в получении одних «информационных объектов» из других «информационных объектов», путем выполнения некоторых алгоритмов; он является одной из основных операций, выполняемых над информацией и главным средством увеличения ее объема и разнообразия.

Процесс хранения связан с необходимостью накопления и длительного хранения данных, обеспечением их актуальности, целостности, безопасности, доступности. Процесс представления и использования информации направлен на решение задачи доступа к информации в удобной для пользователя форме.

Базовые информационные технологии строятся на основе базовых технологических операций, но, кроме этого, включают ряд специфических моделей и инструментальных средств. Этот вид технологий ориентирован на решение определенного класса задач и используется в конкретных технологиях в виде отдельной компоненты. Специфика конкретной предметной области находит отражение в специализированных информационных технологиях, например, организационное управление, управление технологическими процессами, автоматизированное проектирование, обучение и др. Среди них наиболее продвинутыми являются следующие информационные технологии:

- организационного управления (корпоративные информационные технологии);
- в промышленности и экономике;
- в образовании;
- автоматизированного проектирования.

Аналогом инструментальной базы (оборудование, станки, инструмент) являются средства реализации информационных технологий, которые можно разделить на методические, информационные, математические, алгоритмические, технические и программные.

<sup>1</sup> Данный термин не вполне корректен при применении его к технологиям, методам и способам обработки информации. Более корректно было бы сказать «отображение информации».

### 1.2.3. Обработка информации

Понятие обработки информации является весьма широким. Если речь об обработке информации, следует дать понятие инварианта обработки. Обычно им является смысл сообщения (смысл информации, заключенной в сообщении). При автоматизированной обработке информации объектом обработки служит сообщение, и здесь важно провести обработку таким образом, чтобы инварианты преобразований сообщения соответствовали инвариантам преобразования информации.

Цель обработки информации в целом определяется целью функционирования некоторой системы, с которой связан рассматриваемый информационный процесс. Однако для достижения цели всегда приходится решать ряд взаимосвязанных задач.

Обратимся теперь к вопросу о том, в чем сходство и различие процессов обработки информации, связанных с различными составляющими информационного процесса, используя при этом формализованную модель обработки. Прежде всего нельзя отрывать этот вопрос от потребителя информации (адресата), а также от семантического и прагматического аспектов информации. Наличие адресата, для которого предназначено сообщение (сигнал), определяет невозможность установления однозначного соответствия между сообщением и содержащейся в нем информацией. Совершенно очевидно, что одно и то же сообщение может иметь различный смысл для разных адресатов и различное прагматическое значение.

Используемые в производственной сфере такие технологические понятия, как технологический процесс, технологическая операция, метрика, норматив и т.п., могут применяться и в ИТ. Для этого нужно начинать с определения цели. Затем следует попытаться провести структурирование всех предполагаемых действий, приводящих к намеченной цели, и выбрать необходимый программный инструментарий (рис. 1.1).

1-й уровень — этапы, где реализуются базовые технологические процессы, состоящие из операций и действий последующих уровней.

2-й уровень — операции, в результате выполнения которых будет создан конкретный объект в выбранной на 1-м уровне программной среде.

3-й уровень — действия, совокупность стандартных для каждой программной среды приемов работы, приводящих к выполнению поставленной в соответствующей операции цели.

4-й уровень — элементарные операции по управлению элементарными действиями объектов.



Рис. 1.1. Технологический процесс переработки информации в виде иерархической структуры по уровням

ИТ, как и другие технологии, должны отвечать следующим требованиям:

- обеспечивать высокую степень расчленения всего процесса обработки информации на этапы, операции, действия;
- включать весь набор элементов, необходимых для достижения поставленной цели;
- иметь регулярный и масштабируемый характер;
- этапы, действия, операции технологического процесса должны быть стандартизированы и унифицированы, что позволит более эффективно осуществлять целенаправленное управление информационными процессами.

Традиционно в процессе обработки информации используются как измерительная аппаратура, обеспечивающая входные данные, так и собственно обрабатывающие (вычислительные) системы. И те, и другие прошли длинную дорогу развития вместе с человеческой цивилизацией. В следующем пункте будут перечислены основные вехи их истории.

Если раньше обрабатывающей системой был человек или какие-то механические приспособления, то для проведения процесса обработки было достаточно сформулировать набор правил (инструкций). Давно подметили, что повторяющиеся операции целесообразно автоматизиро-

вать в первую очередь и желателно перепоручить машинам. При этом человек, задавая циклическое правило работы машине, колоссально выигрывает в трудозатратах.

Предположим, вам надо сложить 1000 последовательных данных измерений. Заводим специальный счетчик-сумматор и присваиваем ему значение 0. Для каждого из данных надо получить результат измерений и добавить его к счетчику, т.е. вам надо сделать 2001 операцию при «ручном» счете. Другой вариант — написать шесть инструкций для машины:

- 1) завести счетчик-сумматор и присвоить ему значение 0;
- 2) завести индекс (номер) текущей операции и присвоить ему значение 0;
- 3) получить новый результат измерений;
- 4) добавить его к счетчику-сумматору;
- 5) увеличить на 1 индекс текущей операции;
- 6) если он меньше 1000, то перейти к шагу 3.

За прошедшее время существенно усложнились задачи обработки информации, развились способы формулировки и записи правил работы машин (программ работы). Вычислительные устройства превратились в компьютеры, а правила работы — в компьютерные программы.

Программирование — процесс создания компьютерных программ с помощью языков программирования. Программирование сочетает в себе элементы искусства, науки, математики и инженерии.

В узком смысле слова программирование рассматривается как кодирование — реализация одного или нескольких взаимосвязанных алгоритмов на некотором языке программирования. Под программированием также может пониматься разработка логической схемы для интегральной микросхемы, а также процесс записи информации в микросхему ПЗУ (постоянного запоминающего устройства) некоторой электронной системы. В более широком смысле программирование — процесс создания программ, т.е. разработка программного обеспечения.

Составителями программ являются программисты. Большая часть работы программиста связана с написанием и отладкой исходного кода на одном из языков программирования.

Различные языки программирования поддерживают различные стили программирования (или парадигмы программирования). Отчасти искусство программирования состоит в том, чтобы на одном из языков эффективно реализовать алгоритм, наиболее полно подходящий для решения имеющейся задачи. Разные языки требуют от программиста различного уровня внимания к деталям при реализации алгоритма, результатом чего часто бывает компромисс между простотой и произ-

водительностью (или между временем программиста и временем пользователя).

Единственный язык, напрямую выполняемый процессором, — это машинный язык (также называемый машинный код). Изначально все программисты прорабатывали весь алгоритм в машинном коде, но сейчас эта трудная работа уже не делается. Вместо этого программисты пишут исходный код на языке высокого уровня (например, C, C++, C#, Java), а компьютер, используя компилятор или интерпретатор и уточняя все детали, транслирует его за один или несколько этапов в машинный код, готовый к исполнению на целевом процессоре. Если требуется полный низкоуровневый контроль над системой, программисты пишут программу на языке ассемблера, мнемонические инструкции которого преобразуются один к одному в соответствующие инструкции машинного языка целевого процессора.

В некоторых языках вместо машинного кода генерируется интерпретируемый двоичный код «виртуальной машины», также называемый байт-кодом (byte-code). Такой подход применяется в языке Forth, некоторых реализациях языков Lisp, Java, Perl, Python, а также в языках платформы Microsoft.NET.

Типичный процесс разработки программ состоит, в общем, из семи этапов:

- постановка задачи;
- формализация и специфицирование;
- выбор или составление алгоритма;
- программирование;
- компиляция (трансляция);
- отладка и тестирование;
- запуск в эксплуатацию.

Эксплуатируемая программа имеет дело с данными различных типов, предназначенных для решения конкретных задач.

#### 1.2.4. Классификация информационных технологий

В настоящее время в большинстве источников информации классификацию информационных технологий (ИТ) принято проводить по следующим признакам:

- способу реализации в автоматизированных информационных системах (АИС);
- степени охвата задач управления;
- классам реализуемых технологических операций;
- типу пользовательского интерфейса;

- вариантам использования сети ЭВМ;
- обслуживаемой предметной области и др.

1. По способу реализации ИТ делятся на традиционные и современные. Традиционные ИТ существовали в условиях централизованной обработки данных до периода массового использования ПЭВМ. Они были ориентированы главным образом на снижение трудоемкости пользователя (например, инженерные и научные расчеты, формирование регулярной отчетности на предприятиях). Современные (новые<sup>1</sup>) ИТ связаны в первую очередь с информационным обеспечением процесса управления в режиме реального времени.

2. По степени охвата информационными технологиями задач управления выделяют: электронную обработку данных, автоматизацию функций управления, поддержку принятия решений, электронный офис, экспертную поддержку.

В первом случае электронная обработка данных выполняется с использованием ЭВМ без пересмотра методологии и организации процессов управления при решении локальных математических и экономических задач.

Во втором случае при автоматизации управленческой деятельности вычислительные средства используются для комплексного решения функциональных задач, формирования регулярной отчетности и работы в информационно-справочном режиме для подготовки управленческих решений. К этой же группе относятся ИТ поддержки принятия решений, которые предусматривают широкое использование экономико-математических методов и моделей, пакеты прикладных программ (ППП) для аналитической работы и формирования прогнозов, составления бизнес-планов, обоснованных оценок и выводов по процессам и явлениям производственно-хозяйственной деятельности.

К названной группе относятся и широко внедряемые в настоящее время ИТ, получившие название электронного офиса и экспертной поддержки принятия решений. Эти два варианта ИТ ориентированы на использование достижений в области новейших подходов к автоматизации работы специалистов и руководителей, создание для них наиболее благоприятных условий выполнения профессиональных функций, качественного и современного информационного обслуживания за счет автоматизированного набора управленческих процедур, реализуемых в условиях конкретного рабочего места и офиса в целом.

<sup>1</sup> Новые технологии — это те технические нововведения, которые представляют компаниям и предприятиям прогрессивные изменения в рамках области конкурентного преимущества. На сегодняшний день не существует однозначного мнения о целесообразности, значимости, статусу и экономической жизнеспособности различных новых технологий.

Электронный офис предусматривает наличие интегрированных ППП, которые обеспечивают комплексную реализацию задач предметной области. В настоящее время все большее распространение приобретают электронные офисы, сотрудники и оборудование которых могут находиться в разных помещениях. Необходимость работы с документами, материалами и базами данных (БД) конкретного предприятия или учреждения в гостинице, транспорте, дома привела к появлению электронных офисов, включенных в соответствующие сети ЭВМ.

ИТ экспертной поддержки принятия решений, составляют основу автоматизации труда специалистов-аналитиков. Эти работники кроме аналитических методов и моделей для исследования складывающихся ситуаций вынуждены использовать накопленный опыт в оценке ситуаций, т.е. сведения, составляющие базу знаний в конкретной предметной области.

3. По классу реализуемых технологических операций ИТ подразделяются: на работу с текстовым и табличным процессорами, графическими объектами, системы управления БД, гипертекстовые и мультимедийные системы.

Технология формирования видеозображения получила название компьютерной графики.

**Компьютерная графика** — это создание, хранение и обработка моделей объектов и их изображений с помощью ЭВМ. Эта технология проникла в область моделирования различных конструкций (машиностроение, авиационная техника, автомобилестроение, строительная техника и др.), экономического анализа, проникает в рекламную деятельность, делает занимательным досуг. Формируемые и обрабатываемые с помощью цифрового процессора изображения могут быть демонстрационными и анимационными. К демонстрационным изображениям относят, как правило, коммерческую (деловую) и иллюстрационную графику. Ко второй группе — анимационной графике — принадлежит инженерная и научная графика, а также графика, связанная с рекламой, искусством, играми, когда на экран выводятся не только одиночные изображения, но и последовательность кадров в виде фильма (интерактивный вариант). Интерактивная графика является одним из наиболее прогрессивных направлений среди современных ИТ. Это направление переживает бурное развитие в области появления новых графических станций и в области специализированных программных средств, позволяющих создавать реалистические объемные движущиеся изображения, сравнимые по качеству с кадрами видеофильма.

В классическом понимании система управления БД (СУБД<sup>1</sup>) представляет собой набор программ, позволяющих создавать и поддерживать БД в актуальном состоянии.

Обычно любой текст представляется как одна длинная строка символов, которая читается в одном направлении. Гипертекстовая технология — организация текста в виде иерархической структуры. Материал текста делится на фрагменты. Каждый видимый на экране ЭВМ фрагмент, дополненный многочисленными связями с другими фрагментами, позволяет уточнить информацию об изучаемом объекте и двигаться в одном или нескольких направлениях по выбранной связи.

Мультимедиа-технология — программно-техническая организация обмена с компьютером текстовой, графической, аудио- и видеoinформацией.

4. По типу пользовательского интерфейса можно рассматривать ИТ с точки зрения возможностей доступа пользователя к информационным и вычислительным ресурсам. Так, пакетная ИТ исключает возможность пользователя влиять на обработку информации, пока она проводится в автоматическом режиме. В отличие от пакетной диалоговая ИТ предоставляет пользователю неограниченную возможность взаимодействовать с хранящимися в системе информационными ресурсами в реальном масштабе времени, получая при этом всю необходимую информацию для решения функциональных задач и принятия решений.

Интерфейс сетевой ИТ предоставляет пользователю средства доступа к территориально распределенным информационным и вычислительным ресурсам благодаря развитым средствам связи.

В настоящее время наблюдается тенденция к объединению различных типов ИТ в единый компьютерно-технологический комплекс, который носит название интегрированного. Особое место в нем принадлежит средствам коммуникации, обеспечивающим не только чрезвычайно широкие технологические возможности автоматизации управленческой деятельности, но и являющимся основой создания самых разнообразных сетевых вариантов ИТ: локальных, многоуровневых, распределенных и глобальных информационно-вычислительных сетей.

5. По обслуживаемым предметным областям ИТ подразделяются разнообразно. Например, только в экономике ими являются бухгалтерский учет, банковская, налоговая и страховая деятельность и др.

<sup>1</sup> Система управления базами данных (СУБД) — совокупность программных и лингвистических средств общего или специального назначения, обеспечивающих управление созданием и использованием баз данных. Это более точно отражающее суть определение.

### 1.3. Этапы развития информационных технологий

Как и всякое развитие деятельности в рамках нашей цивилизации, история развития информационных технологий также имеет место быть. Длительный путь от осознания древним человеком необходимости сохранить увиденное в виде наскальных рисунков до современных технологий многоуровневого анализа данных с использованием удаленных распределенных компьютерных систем — это именно тот путь, который прошли ИТ. Но на этом пути возможно выделить ряд определяющих этапов, характерных для технических прорывов нашей цивилизации.

В истории развития цивилизации произошло несколько информационных революций — преобразований общественных отношений из-за кардинальных изменений в сфере обработки информации. Следствием подобных преобразований являлось приобретение человеческим обществом нового качества.

Первая революция связана с изобретением письменности, что привело к гигантскому качественному и количественному скачку. Появилась возможность передачи знаний от поколения к поколению.

Вторая (середина XVI в.) вызвана изобретением книгопечатания, которое радикально изменило индустриальное общество, культуру, организацию деятельности.

Третья (конец XIX в.) обусловлена изобретением электричества, благодаря которому появились телеграф, телефон, радио, позволяющие оперативно передавать и накапливать информацию в любом объеме.

Четвертая (70-е гг. XX в.) связана с изобретением микропроцессорной технологии и появлением персонального компьютера. На микропроцессорах и интегральных схемах создаются компьютеры, компьютерные сети, системы передачи данных (информационные коммуникации). Этот период характеризуют три фундаментальные инновации:

- переход от механических и электрических средств преобразования информации к электронным;
- миниатюризация всех узлов, устройств, приборов, машин;
- создание программно управляемых устройств и процессов.

Последняя информационная революция выдвигает на передний план новую отрасль — информационную индустрию, связанную с производством технических средств, методов, технологий для производства новых знаний. Важнейшая составляющая информационной индустрии — информационная технология.

Современные ИТ опираются на достижения в области компьютерной техники и средств связи, поэтому в этой связи историю развития ИТ

можно рассмотреть на фоне появления и развития компьютерных средств. В результате мы можем выделить этапы:

**1 этап** (до второй половины XIX в.) — «ручная» информационная технология, инструментарий которой составляли: перо, чернильница, книга. Коммуникации осуществлялись ручным способом путем переправки через почту писем, пакетов, депеш. Основная цель технологии — представление информации в нужной форме.

**2 этап** (с конца XIX в.) — «механическая» технология, инструментарий которой составляли: пишущая машинка, телефон, диктофон, оснащенная более совершенными средствами доставки почта. Основная цель технологии — представление информации в нужной форме более удобными средствами.

**3 этап** (40—60-е гг. XX в.) — «электрическая» технология, инструментарий которой составляли: большие ЭВМ и соответствующее программное обеспечение, электрические пишущие машинки, копировальные аппараты, портативные диктофоны. Изменяется цель технологии — акцент начинает перемещаться с формы представления информации на формирование ее содержания.

**4 этап** (с начала 70-х гг. XX в.) — «электронная» технология, инструментарием которой становятся большие ЭВМ и создаваемые на их базе автоматизированные системы управления (АСУ) и информационно-поисковые системы (ИПС), оснащенные широким спектром базовых и специализированных программных комплексов. Акцент смещается в сторону формирования более содержательной информации.

**5 этап** (с середины 80-х гг. XX в.) — «компьютерная» технология, основным инструментарием которой является персональный компьютер с широким спектром стандартных программных продуктов разного назначения. На этом этапе происходит процесс персонализации АСУ (создание систем поддержки принятия решения для разных специалистов).

В связи с переходом на микропроцессорную технологию существенным изменениям подвергается бытовая техника, приборы связи и коммуникации, оргтехника. Начинают широко развиваться компьютерные сети (локальные и глобальные).

**1-й этап** (до второй половины XIX в.) — «ручные» технологии. В качестве инструментов таких технологий использовались перо, чернильница, книга, элементарные ручные средства счета. Коммуникации осуществлялись путем доставки конной почтой писем, пакетов, депеш, в европейских странах применялся механический телеграф. Основная цель технологий — представление и передача информации в нужной форме.

**2-й этап** (конец XIX в. — 40-е гг. XX в.) — «механические» технологии. В качестве инструментов таких технологий использовались пишущая машинка, арифмометр, телеграф, телефон, диктофон<sup>1</sup>, оснащенная более совершенными средствами доставки почта. Основная цель технологий — представление информации в нужной форме более удобными средствами, сокращение затрат на исправление потерь и искажений.

**3-й этап** (40-е — 60-е гг. XX в.) — «электрические» технологии. В качестве инструментов таких технологий использовались первые ламповые ЭВМ и соответствующее программное обеспечение, электрические пишущие машинки, телетайпы (телексы), ксероксы, портативные диктофоны. Организация доставки информации в заданное время. Акцент в ИТ начинает перемещаться с формы представления информации на формирование ее содержания. Одной из первых технологий, активно разрабатывающихся в то время, была технология расчета заработной платы.

**4-й этап** (70-е гг. — середина 80-х гг. XX в.) — «электронные» технологии. Основной инструментарий — большие ЭВМ и создаваемые на их базе автоматизированные системы управления (АСУ) и информационно-поисковые системы, оснащенные широким спектром базовых и специализированных программных комплексов. Центр тяжести технологий смещается на формирование содержательной стороны информации для управленческой среды различных сфер общественной жизни, особенно на организацию аналитической работы.

**5-й этап** (с середины 80-х гг. XX в.) — «компьютерные» («новые») технологии, персональный компьютер с широким спектром стандартных и заказных программных продуктов широкого назначения. Появление первых ПК<sup>2</sup> (1983) и принятие стандартов «открытой архитектуры ПК»<sup>3</sup> оказало значительное влияние на облик ИТ, поскольку каждый пользователь получал возможность работать с информацией на собственной ЭВМ. А возникновение стандартов в области компьютерных сетей<sup>4</sup> (начиная с 1984 г.) оказало кардинальное влияние на облик ИТ, поскольку

<sup>1</sup> В этот момент времени диктофон занимал размеры комнаты, запись информации осуществлялась на специальную проволоку. Качество записи и воспроизведения было достаточно низким, но все же это был инструмент, дающий возможность зафиксировать аудиоинформацию на долговременный носитель и потом многократно (при необходимости) ее воспроизводить.

<sup>2</sup> Первым ПК, появившимся на рынке, был персональный компьютер «Альдар», созданный Полом Алленом, впоследствии совместно со Стивом Джобсом создавшим корпорацию Apple.

<sup>3</sup> Открытая архитектура ПК — термин подразумевающий, что архитектура компьютера, периферийного устройства или же программного обеспечения имеет опубликованные в открытой литературе спецификации, что позволяет другим производителям разрабатывать дополнительные устройства и системы с такой архитектурой.

<sup>4</sup> Первым в этой области стало принятие в качестве стандарта так называемой Модели взаимодействия открытых систем, имеющей ранг международного стандарта и известной под аббревиатурой ISO/OSI, а дальнейшие американские национальные стандарты группы 802 позволили стандартизировать область локальных компьютерных сетей, давших возможность ориентировать ИТ на малые компании и на группы одновременно работающих пользователей.

ку появилась возможность объединения в сети различных систем поддержки принятия решений на различных уровнях управления. Системы имеют встроенные элементы анализа и искусственного интеллекта, реализуются на персональном компьютере и используют сетевые технологии и телекоммуникации для работы в сети.

**6-й этап** (с середины 90-х гг. XX в.) — Internet (Intranet) («новейшие») технологии. Реализация архитектуры «Клиент — сервер». Широко используются в различных областях науки, техники и бизнеса распределенные системы, глобальные, региональные и локальные компьютерные сети. Развивается электронная коммерция. Увеличение объемов информации привело к созданию технологии Data Mining.

## 1.4. Общая классификация видов информационных технологий и их реализация в технических областях

ИТ в настоящее время можно классифицировать по ряду признаков, в частности (рис. 1.2):

- способу реализации в автоматизированных информационных системах (АИС);
- степени охвата АИТ задач управления;
- классам реализуемых технологических операций;
- типу пользовательского интерфейса;
- вариантам использования сети ЭВМ;
- обслуживаемой предметной области.

По способу реализации АИТ в автоматизированных информационных системах выделяют традиционно сложившиеся и новые информационные технологии.

Если традиционные АИТ прежде всего существовали в условиях централизованной обработки данных, до массового использования ПЭВМ были ориентированы главным образом на снижение трудоемкости при формировании регулярной отчетности, то новые информационные технологии связаны с информационным обеспечением процесса управления в режиме реального времени.

Новая информационная технология — это технология, которая основывается на применении компьютеров, активном участии пользователей (непрофессионалов в области программирования) в информационном процессе, высоком уровне дружественного пользовательского интерфейса, широком использовании пакетов прикладных программ общего и проблемного назначения, доступе пользователя к удаленным базам данных и программам благодаря вычислительным сетям ЭВМ.

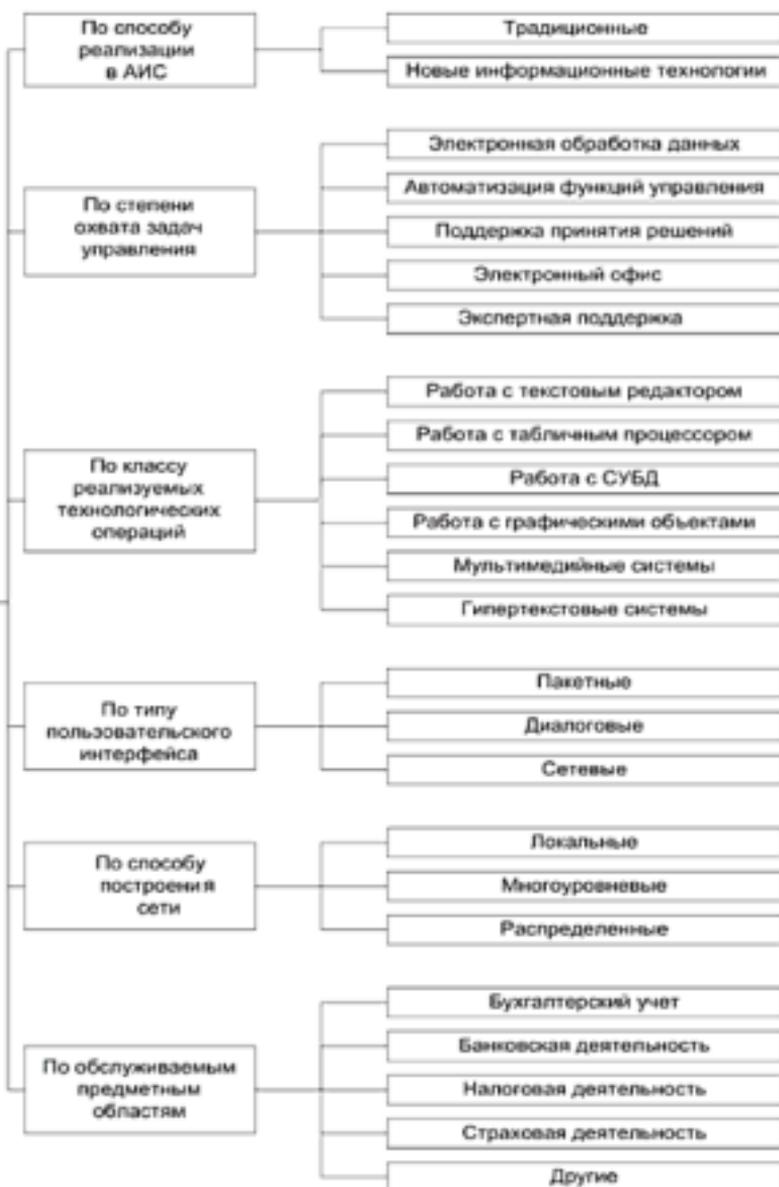


Рис. 1.2. Классификация автоматизированных информационных технологий

По степени охвата АИТ задач управления выделяют электронную обработку данных, когда с использованием ЭВМ без пересмотра методологии и организации процессов управления ведется обработка данных с решением отдельных экономических задач, и автоматизацию управленческой деятельности.

Во втором случае вычислительные средства, включая суперЭВМ и ПЭВМ, используются для комплексного решения функциональных задач, формирования регулярной отчетности и работы в информационно-справочном режиме для подготовки управленческих решений.

К этой же группе могут быть отнесены АИТ поддержки принятия решений, которые предусматривают широкое использование экономико-математических методов, моделей и ППП для аналитической работы и формирования прогнозов, составления бизнес-планов, обоснованных оценок и выводов по изучаемым процессам, явлениям производственно-хозяйственной практики.

К названной группе относятся и широко внедряемые в настоящее время АИТ, получившие название электронного офиса и экспертной поддержки решений. Эти два варианта АИТ ориентированы на использование последних достижений в области интеграции новейших подходов к автоматизации работы специалистов и руководителей, создание для них наиболее благоприятных условий выполнения профессиональных функций, качественного и своевременного информационного обслуживания за счет полного автоматизированного набора управленческих процедур, реализуемых в условиях конкретного рабочего места и офиса в целом.

Электронный офис предусматривает наличие интегрированных пакетов прикладных программ, включающих специализированные программы и информационные технологии, которые обеспечивают комплексную реализацию задач предметной области. В настоящее время все большее распространение приобретают электронные офисы, оборудование и сотрудники которых могут находиться в разных помещениях. Необходимость работы с документами, материалами, базами данных конкретной организации или учреждения в домашних условиях, в гостинице, транспортных средствах привела к появлению АИТ виртуальных офисов. Такие АИТ основываются на работе локальной сети, соединенной с территориальной или глобальной сетью. Благодаря этому абонентские системы сотрудников учреждения независимо от того, где они находятся, оказываются включенными в общую для них сеть.

Автоматизированные информационные технологии экспертной поддержки составляют основу автоматизации труда специалистов-аналитиков. Эти работники, кроме аналитических методов и моделей для ис-

следования складывающихся в рыночных условиях ситуаций по сбыту продукции, услуг, финансового положения предприятия, фирмы, финансово-кредитной организации, вынуждены использовать накопленный и сохраняемый в системе опыт оценки ситуаций, т.е. сведения, составляющие базу знаний в конкретной предметной области. Обработанные по определенным правилам такие сведения позволяют подготавливать обоснованные решения для поведения на финансовых и товарных рынках, выработать стратегию в областях менеджмента и маркетинга.

По классам реализуемых технологических операций АИТ рассматриваются по существу в программном аспекте и включают:

- текстовую обработку;
- электронные таблицы;
- автоматизированные банки данных;
- обработку графической и звуковой информации;
- мультимедийные и другие системы.

Перспективным направлением развития компьютерной технологии является создание программных средств для вывода высококачественного звука и видеоизображения. Технология формирования видеоизображения получила название компьютерной графики. Компьютерная графика — это создание, хранение и обработка моделей объектов и их изображений с помощью ЭВМ. Эта технология проникла в область экономического анализа, моделирования различного рода конструкций, она незаменима в производстве, проникает в рекламную деятельность, делает занимательным досуг. Формируемые и обрабатываемые с помощью цифрового процессора изображения могут быть демонстрационными и анимационными. К первой группе, как правило, относят коммерческую (деловую) и иллюстративную графику, ко второй — инженерную и научную, а также связанную с рекламой, искусством, играми, когда выводятся не только одиночные изображения, но и последовательность кадров в виде фильма (интерактивный вариант). Интерактивная машинная графика является одним из наиболее прогрессивных направлений среди новых информационных технологий. Это направление переживает бурное развитие в области появления новых графических станций и в области специализированных программных средств, позволяющих создавать реалистические объемные движущиеся изображения, сравнимые по качеству с кадрами видеофильма.

Программно-техническая организация обмена с компьютером текстовой, графической, аудио- и видеоинформацией получила название мультимедиа-технологии. Такую технологию реализуют специальные программные средства, имеющие встроенную поддержку мультимедиа и позволяющие использовать ее в профессиональной деятельности, учебно-

образовательных, научно-популярных и игровых областях. При применении этой технологии в экономической работе открываются реальные перспективы использовать компьютер для озвучивания изображений, а также понимания им человеческой речи, ведения компьютером диалога со специалистом на родном для специалиста языке. Способность компьютера с голоса воспринимать несложные команды управления программами, открытием файлов, выводом информации на печать и другими операциями в ближайшем будущем создаст самые благоприятные условия пользователю для взаимодействия с ним в процессе профессиональной деятельности.

По типу пользовательского интерфейса можно рассматривать АИТ с точки зрения возможностей доступа пользователя к информационным и вычислительным ресурсам. Так, пакетная АИТ исключает возможность пользователя влиять на обработку информации, пока она производится в автоматическом режиме. Это объясняется организацией обработки, которая основана на выполнении программно-заданной последовательности операций над заранее накопленными в системе и объединенными в пакет данными. В отличие от пакетной диалоговая АИТ предоставляет пользователю неограниченную возможность взаимодействовать с хранящимися в системе информационными ресурсами в реальном масштабе времени, получая при этом всю необходимую информацию для решения функциональных задач и принятия решений.

Интерфейс сетевой АИТ предоставляет пользователю средства теледоступа к территориально распределенным информационным и вычислительным ресурсам благодаря развитым средствам связи, что делает такие АИТ широко используемыми и многофункциональными.

## 1.5. Глобальная, базовая и конкретные информационные технологии

Уровни рассмотрения ИТ:

- *концептуальный* — рассматриваются общее содержание информационных процессов и структуры предметной области;
- *логический* — происходит формализация информационных процессов и разработка модели предметной области и систем обработки информации;
- *физический* — рассматриваются способы реализации информационных процессов, аппаратные программные средства, средства связи, которые реализуют процессы, методы и модели, выделенные на концептуальном и логическом уровне.

В общем случае ИТ можно разделить на глобальные, базовые и конкретные (частные).

Глобальные ИТ включают методы, модели и средства, позволяющие использовать информационные ресурсы всего общества и решающие задачи, связанные с изменением информационного ресурса всего общества.

Базовые ИТ — ИТ, которые основаны на одном из базовых процессов, выделяемых в рамках общей ИТ (процесс сбора и ввода информации, процесс отображения информации, процесс хранения информации, процесс обмена информацией, процесс получения знаний).

Конкретная ИТ — ИТ, характерная для какого-либо производства или науки (технологической подготовки производства).

Для каждого типа технологии можно выделить каждый из уровней рассмотрения.

## Контрольные вопросы и задания

1. Назовите основные основания классификации ИТ.
2. Перечислите основные вехи в развитии ИТ.
3. В чем, по вашему мнению, заключаются особенности так называемых новых ИТ?
4. Назовите три основных принципа новой (компьютерной) информационной технологии.
5. На какие достижения в области компьютерной техники и средств связи опираются современные ИТ?
6. На что опираются современные новые ИТ для эффективного взаимодействия конечных пользователей с вычислительной системой?

# ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ В СИСТЕМАХ. ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ ОБРАБОТКИ И ПРЕДСТАВЛЕНИЯ ДАННЫХ

## 2.1. Системный подход к решению функциональных задач и к организации информационных процессов в системах

В настоящее время концепция базы данных подразумевает разумный компромисс между сокращением до минимума необходимого дублирования информации и эффективностью процессов выборки и обновления требуемых данных. Действительное обеспечение такого решения возможно только при условии системного анализа всего комплекса задач, подлежащих автоматизации (описание целей системы и ее функций, а также состава и специфики информационных потоков и отдельных программных модулей). Системный подход базируется на положениях общих теорий систем и наиболее эффективен при решении сложных задач анализа и синтеза требующих одновременного использования ряда научных дисциплин. Объективное требование системного подхода

к разработке программных средств решения задач при автоматизации систем управления вызвало необходимость дифференциации специалистов-разработчиков: системные аналитики, системотехники, прикладные и системные программисты. Системный аналитик, исходя из общих целей — назначения, технических характеристик, состава и описания требований пользователей к прикладным задачам и системе в целом, — формулирует общие формальные требования к программному обеспечению системы. Специалист системотехник преобразует общие формальные требования, детальные спецификации на отдельные программы, участвует в разработке логической структуры базы данных, т.е. определяет общую информационно-структурную и программную часть проекта. Прикладной программист преобразует спецификации в логическую структуру программных модулей, а затем в программный код. Системный программист обеспечивает сопряжение программных модулей с программной средой, в рамках которой предстоит функционировать прикладным программам или задачам.

Системная разработка прикладных программ ориентирована также на использование интегрированных и распределенных баз данных. По этой же причине в качестве инструментальных средств разработки программного продукта наряду с языками программирования применяют языковые средства СУБД.

Применение компьютерных информационных технологий позволяет в ряде случаев при сравнительно небольших затратах получать ценные управленческие решения. Составление экономико-математических моделей и проведение расчетов с помощью вычислительных машин позволяют быстро и относительно недорого проводить разработку и сравнение многочисленных вариантов планов и управленческих решений. В настоящее время имеются отработанные методы решения ряда типовых задач по организации и планированию производства, для которых могут быть применены компьютерные информационные технологии. Все эти задачи могут быть классифицированы следующим образом.

Задачи в области производства, его организации (организация проектирования, задачи управления качеством, расчета потребностей в ресурсах с распределением во времени на основе календарного плана производства и т.п.).

Планирование производства (задачи планирования производства товарной продукции, технического развития производства, заработной платы, материально-технического обеспечения).

Такие отработанные решения определенных типовых задач базируются на методах имитационного моделирования, линейного программирования, вероятностного моделирования и других методах. Возможность

практического решения указанных задач в настоящее время расширяется в связи с компьютеризацией всех звеньев управленческого аппарата, созданием локальных и объединенных вычислительных сетей, организацией локальных и централизованных информационных баз данных и обеспечением к ним оперативного доступа.

На современном этапе автоматизированная обработка данных в организационных системах характеризуется переходом от централизованной обработки информации к децентрализованной на основе широкого применения ПК. Основными элементами современного электронного учреждения являются автоматизированные рабочие места пользователей, а также системы редактирования текстов БД И СУБД, информационно-вычислительные сети, электронная почта и т.п. Автоматизированное рабочее место — вычислительная система, предназначенная для автоматизации профессиональной деятельности. Примерами функций пользователей, реализуемых на АРМе, могут быть следующие: подготовка документов, содержащих текстовые, графические элементы на основе анализа доступной информации, хранение и поиск информации, прием или передача документов внутри учреждения и за его пределы, обеспечение режима использования и надежного хранения документов. Необходимость в обмене информацией в различных сферах управленческой деятельности, получении новых сведений в результате коллективного обсуждения привели к таким формам общения, как конференции, семинары и совещания. Эта идея с точки зрения информационной технологии была реализована в виде телеконференций.

**Автоматизированная обучающая система** — комплекс программных, технических и учебно-методических средств, предназначенных для активного индивидуального обучения человека на основе программного управления этим обучением.

Типы обучающих программ:

- тренировочные и контролирующие;
- наставнические;
- имитационные и моделирующие;
- развивающие.

Тренировочные программы предназначены для закрепления умений и навыков. Эти программы в случайной последовательности предлагают пользователю вопросы и задачи в случайной последовательности и подсчитывают количество правильно или неправильно решенных задач.

Наставнические программы предлагают теоретический материал для изучения, задачи и вопросы служат в этих программах для организации человеко-машинного диалога для управления ходом обучения.

Моделирующие программы основаны с одной стороны на графических иллюстративных возможностях ПК и вычислительных с другой и позволяют осуществлять компьютерный эксперимент. Такие программы предоставляют возможность наблюдать на экране некоторый процесс, влияя на его ход подачей команды с клавиатуры, меняющей значения параметров.

Развивающие игры предоставляют в распоряжение пользователя некоторую воображаемую среду, существующую только в компьютерном мире, набор каких-то возможностей и средств их реализации.

**Автоматизированные системы научных исследований (АСНИ)** представляют собой программно-аппаратные комплексы, обрабатывающие данные, поступающие от различного рода экспериментальных установок и измерительных приборов, и на основе их анализа облегчающие обнаружение новых эффектов и закономерностей.

**Системы автоматизированного проектирования (САПР).** Это комплекс программных и аппаратных средств, предназначенных для автоматизации процесса проектирования человеком технических изделий или продуктов интеллектуальной деятельности. Проектирование новых изделий является основной задачей изобретателей и конструкторов и протекает в несколько этапов, таких как:

- нормирование замысла;
- поиск физических принципов;
- поиск конструктивных решений;
- создание опытного образца;
- разработка технологий промышленного изготовления.

Этапы формирования замысла и поиска физических принципов являются творческими, не поддающимися автоматизации. При конструировании и расчетах применяются САПР.

В настоящее время САПР являются обязательным атрибутом конструкторских бюро и проектных организаций. Широко применяются в архитектуре, электротехнике, машиностроении, авиакосмической технике и т.д.

**Геоинформационные системы и технологии (ГИС).** ГИС объединяют компьютерную картографию и СУБД. Концепция технологии состоит в создании многослойной электронной карты, опорный слой которой описывает географию территории, а каждый из остальных слоев — один из аспектов состояния территории. Технология ГИС применима там, где необходимо учитывать, обрабатывать и демонстрировать территориально распределенную информацию. Пользователями ГИС могут быть как организации, чья деятельность целиком базируется на земле (нефтегазовые предприятия, экологические службы, ЖКХ), так и многочисленные ком-

мерческие организации (банки, страховые компании, торговые и строительные фирмы). В основе любой ГИС лежит информация о каком-либо участке земной поверхности. База данных организуется в виде набора слоев информации. Основой является географически привязанная карта местности (топооснова). В процессе создания и наложения слоев друг на друга между ними устанавливаются необходимые связи, что позволяет выполнить пространственные операции с объектами посредством моделирования и интеллектуальной обработки данных. Типовая схема организации ГИС складывается из следующих основных компонентов.

1. Приобретение и предварительная подготовка данных. Включается в себя манипуляции с исходными данными как результатами полевых исследований, текстовыми или табличными материалами, с архивными материалами.

2. Ввод и размещение данных (пространственных и непространственных). По функциональному назначению эта компонента конвертирует информацию во внутренние форматы системы и обеспечивает структурную и логическую совместимость всех имеющихся.

3. Управление данными. Предполагает наличие средств оптимальной внутренней организации данных, обеспечивающих эффективный доступ к ним.

4. Манипуляция данными и их анализ. Эти функции представлены средствами, предназначенными для содержательной обработки данных в целях их обработки и реорганизации.

5. Производство конечного продукта. Включает в себя вывод полученных результатов для конечных потребителей ГИС. Эти продукты могут представлять карты, статистические отчеты, различные графики, стандартные формы определенных документов.

Подавляющее большинство ГИС различают геометрическую и атрибутивную компоненты баз данных. Геометрическая или картографическая информация представляется точками, кривыми и площадными объектами. Атрибутивная информация содержит текстовые, числовые и логические данные о географических объектах. Она хранится в виде отдельных табличных файлов, как правило, в форматах реляционных баз данных.

## **2.2. Понятие технологического процесса обработки информации. Централизованная и распределенная обработка данных**

Понятие «информационная технология» опирается на понятие «технология». Более обширное по содержанию толкование данного термина

было дано польским философом и писателем Станиславом Лемом, определившим его как «...обусловленные состоянием знаний и общественной эффективностью способы достижения целей, поставленных обществом...».

А наиболее распространенным считается определение, приведенное в разнообразных словарях и энциклопедиях: «Технология (от греч. *techne* — искусство, умение, мастерство) определяется совокупностью методов обработки, изготовления, изменения состояния, формы, свойств материала, сырья или полуфабриката, осуществляемое в процессе изготовления продукции...».

Данное определение, бесспорно, является более определенным, чем сформулированное С. Лемом, так как ограничивает его применение сферой промышленного производства, но оно позволяет вычленить важнейшие его составляющие (рис. 2.1):

- объект технологии определяет направленность действия, осуществляемого в рамках технологии (материалы, сырье, полуфабрикаты);
- цель технологии определяется конечным результатом действий, выполняемых в рамках технологии (изготовление, обработка, изменение состояния, свойства или формы);
- средства технологии и методы их применения определяются способами реализации действий над объектом технологии для достижения цели или целей технологии.

Технология		
Объект	Цель	Средства и методы

Рис. 2.1. Составляющие понятия «технология»

В связи с тем что в соответствии с определением С. Лема технологии не ограничиваются сферой промышленного производства, а всецело определяются потребностями общества во всей их разнообразности, то различные области деятельности человека требуют и многообразных технологий.

Отличия технологий нашли проявление в том, на что направлена деятельность людей в определенной сфере, т.е. в объектах технологий. Для промышленного производства, как уже говорилось, это материалы, сырье или полуфабрикаты, составляющие материально-вещественные ресурсы производственного процесса.

Если в качестве объекта деятельности, а также надлежащих способов ее осуществления выступают энергетические ресурсы (например, элек-

трическая или тепловая энергия), то получают энергетические технологии (производство, преобразование, передача, распределение, потребление энергии).

Финансовые ресурсы, являясь объектом деятельности, порождают финансовые технологии (бухгалтерские и банковские технологии, технологии работы на рынке облигаций и ценных бумаг, технологии экономические и финансовые и т.д.).

Информация как общественный ресурс тоже может являться объектом деятельности и, значит, связана с соответствующими технологиями — информационными технологиями.

Основываясь на рассмотренном выше понятии «технология», можно сформулировать следующее определение понятия «информационная технология».

Информационная технология является совокупностью средств и методов их применения для целенаправленного изменения свойств информации, полностью определяемого содержанием решаемой задачи или проблемы.

Результатом, или целью, технологического процесса обработки информации является целенаправленная модификация свойств информации, определяемая постановкой решаемой проблемы или задачи.

Содержание информационного преобразования при обработке информации обуславливается определенным набором изменяемых свойств информации, и с данной точки зрения выделяют следующие преобразования исходной информации:

- сбор информации;
- накопление информации;
- регистрация информации;
- передача информации;
- копирование информации;
- упорядочение информации;
- хранение информации;
- поиск информации;
- представление информации;
- выдача информации;
- защита информации.

Первым способом решения проблемы совместного использования информации был централизованный принцип обработки данных с применением многопользовательского режима (70-е гг. XX в.). В ту эпоху ЭВМ отличались значительной дороговизной и большими габаритами, поэтому экономически обоснованным было приобретение пользователями одной мощной универсальной ЭВМ, на которой решались практиче-

ски все классы задач. Для обеспечения доступа к ЭВМ как можно большего числа пользователей к ней подключалось несколько терминалов, обычно территориально расположенных в одном большом зале или в пределах одного здания.

Преимуществом подобной системы является простота внешних устройств, а также концентрация данных и ресурсов общего пользования в одном месте (что, кстати, облегчало обслуживание ЭВМ).

Основными недостатками централизованной обработки данных были:

- низкая надежность — даже кратковременный выход из строя центральной ЭВМ приводил к роковым последствиям для системы в целом;
- низкая производительность — при диалоговой обработке данных в многопользовательском режиме ЭВМ не успевала обслуживать всех пользователей в нужном темпе;
- ограниченная возможность решения разнотипных прикладных задач, так как это требовало слишком высокой многофункциональности и универсальности ЭВМ и ОС;
- концентрация большой вычислительной мощности ЭВМ делала компьютеры слишком дорогими и недоступными для одиночных пользователей и небольших предприятий.

С появлением малых ЭВМ, микроЭВМ и, наконец, персональных компьютеров возникло логически обоснованное требование перехода к иным информационным технологиям — от использования крупных центральных ЭВМ к распределенной обработке данных. Распределенная обработка данных — обработка данных, выполняемая на независимых, но связанных между собой компьютерах, представляющих собой распределенную систему (т.е. не сосредоточенную в каком-то малом пространстве). Таким образом, имеет место централизация хранения данных и децентрализация вычислительных мощностей по их обработке. В этом заключается важнейшее отличие данных систем от классических многотерминальных систем.

Существуют различные способы объединения нескольких компьютеров в единую систему, отвечающую требованию распределенной обработки. Высшей формой такого объединения является компьютерная сеть.

Под сетью ПЭВМ (или компьютерной сетью) понимают коммуникационную систему, состоящую из двух или более компьютеров и включающую в себя специальные программы и аппаратное обеспечение, используемое для обмена информацией между компьютерами и совместного использования ресурсов. Таким образом, компьютерная сеть представляет собой совокупность трех компонент:

- сети передачи данных (включающей в себя каналы передачи данных и средства коммутации);
- компьютеров, взаимосвязанных сетью передачи данных;
- сетевого программного обеспечения.

Ресурсы сети — это данные, приложения и периферийные устройства, такие как диск, принтер, модем и т.д. В такой системе любое из подключенных устройств может использовать сеть для передачи или получения информации.

Основными отличиями компьютерной сети от прочих многомашинных вычислительных комплексов являются:

- размерность — сеть может состоять из десятков и даже сотен ЭВМ, расположенных друг от друга на расстоянии от нескольких метров до тысяч километров;
- разделение функций между ЭВМ — в сети одни машины могут выполнять функции обработки или передачи данных, другие — управление системой, третьи — обслуживанием периферийных устройств и т.д.;
- необходимость решения в сети задачи маршрутизации, т.е. определение оптимального маршрута передачи сообщений от одного компьютера другому; состояние каналов связи в крупной сети постоянно изменяется и в каждый момент времени может возникнуть необходимость искать новый маршрут.

Отметим преимущества, получаемые после объединения отдельных ПК в сеть:

- возрастает оперативность работы;
- появляется возможность организовать доступ всех пользователей к единому информационному ресурсу (например, базе данных), расположенному на одном компьютере;
- снижаются затраты на аппаратное обеспечение в расчете на одного пользователя; это достигается за счет совместного использования дискового пространства, дорогих высококачественных внешних устройств (лазерных принтеров, сканеров, плоттеров);
- повышается надежность системы в целом, поскольку при поломке одного устройства исполнение его функций может взять на себя другое.

С появлением компьютерных сетей возникла необходимость формирования новой специальной терминологии. Перечислим основные понятия, использующиеся в сетевых технологиях.

Абоненты сети — объекты, генерирующие или потребляющие информацию в сети. Абонентами сети могут быть отдельные ЭВМ и их ком-

плексы, терминалы, промышленные роботы, станки с числовым управлением и т.д.

Любой абонент сети подключается к станции. Станция — аппаратура, которая выполняет функции передачи или приема информации. Совокупность абонента и станции, к которой он подключен, называют абонентской системой. Для организации взаимодействия абонентов необходима физическая передающая среда — линия связи или пространство, в котором распространяются сигналы, и аппаратура передачи данных. На базе физической передающей среды строится коммуникационная сеть, которая обеспечивает передачу информации между абонентами сети. Такой подход позволяет рассматривать любую компьютерную сеть как совокупность абонентских систем и коммуникационной сети.

Компьютерные сети обладают многими новыми возможностями, недоступными для вычислительных систем на базе одной ЭВМ. Например:

- организация параллельного решения крупной задачи за счет одновременной обработки различных фрагментов данных на разных ЭВМ;
- создание распределенной базы данных, объединяющей информационные ресурсы многих ЭВМ;
- специализация отдельных ЭВМ для эффективного решения определенных классов задач;
- резервирование вычислительных мощностей и средств передачи данных на случай выхода из строя отдельных ЭВМ и устройств;
- перераспределение вычислительных мощностей между пользователями при изменении их потребностей;
- стабилизация уровня загрузки ЭВМ и дорогостоящего периферийного оборудования.

Информационная технология обработки данных предназначена для решения хорошо структурированных задач, по которым имеются необходимые входные данные и известны алгоритмы и другие стандартные процедуры их обработки. Эта технология применяется на уровне операционной (исполнительской) деятельности персонала невысокой квалификации в целях автоматизации некоторых рутинных постоянно повторяющихся операций управленческого труда. Поэтому внедрение информационных технологий и систем на этом уровне существенно повысит производительность труда персонала, освободит его от рутинных операций, возможно, даже приведет к необходимости сокращения численности работников.

На уровне операционной деятельности решаются следующие задачи:

- обработка данных об операциях, производимых фирмой;
- создание периодических контрольных отчетов о состоянии дел в фирме;
- получение ответов на всевозможные текущие запросы и оформление их в виде бумажных документов или отчетов.

При проектировании технологических процессов ориентируются на режимы их реализации. Режим реализации технологии зависит от объемно-временных особенностей решаемых задач: периодичности и срочности, требований к скорости обработки сообщений, а также от режимных возможностей технических средств и в первую очередь ЭВМ. Существуют:

- пакетный режим;
- режим реального масштаба времени;
- режим разделения времени;
- регламентный;
- запросный;
- диалоговый;
- телеобработки;
- интерактивный;
- однопрограммный;
- многопрограммный (мультиобработка).

Для пользователей финансово-кредитной системы наиболее актуальны следующие режимы: пакетный, диалоговый и режим реального времени.

**Пакетный режим.** При использовании этого режима пользователь не имеет непосредственного общения с ЭВМ. Сбор и регистрация информации, ввод и обработка не совпадают по времени. Вначале пользователь собирает информацию, формируя ее в пакеты в соответствии с видом задач или каким-то другим признакам (как правило, это задачи неоперативного характера, с долговременным сроком действия результатов решения). После завершения приема информации производятся ее ввод и обработка, таким образом, происходит задержка обработки. Этот режим используется, как правило, при централизованном способе обработки информации.

**Диалоговый режим (запросный) режим,** при котором существует возможность пользователя непосредственно взаимодействовать с вычислительной системой в процессе работы пользователя. Программы обработки данных находятся в памяти ЭВМ постоянно, если ЭВМ доступна в любое время, или в течение определенного промежутка времени, когда ЭВМ доступна пользователю. Взаимодействие пользователя с вычислительной системой в виде диалога может быть многоаспектным и определяться различными факторами: языком обще-

ния, активной или пассивной ролью пользователя; кто является инициатором диалога — пользователь или ЭВМ; временем ответа; структурой диалога и т.д. Если инициатором диалога является пользователь, то он должен обладать знаниями по работе с процедурами, форматами данных и т.п. Если инициатор — ЭВМ, то машина сама сообщает на каждом шаге, что нужно делать с разнообразными возможностями выбора. Этот метод работы называется «выбором меню». Он обеспечивает поддержку действий пользователя и предписывает их последовательность. При этом от пользователя требуется меньшая подготовленность.

Диалоговый режим требует определенного уровня технической оснащенности пользователя, т.е. наличие терминала или ПЭВМ, связанных с центральной вычислительной системой каналами связи. Этот режим используется для доступа к информации, вычислительным или программным ресурсам. Возможность работы в диалоговом режиме может быть ограничена во времени начала и конца работы, а может быть и неограниченной.

Режим реального масштаба времени означает способность вычислительной системы взаимодействовать с контролируемыми или управляемыми процессами в темпе протекания этих процессов. Время реакции ЭВМ должно удовлетворять темпу контролируемого процесса или требованиям пользователей и иметь минимальную задержку. Как правило, этот режим используется при децентрализованной и распределенной обработке данных. Пример: на рабочем столе операциониста установлен ПК, через который вся информация по операциям вводится в ЭВМ по мере ее поступления. Различаются следующие способы обработки данных: централизованная, децентрализованная, распределенная и интегрированная.

Централизованная предполагает наличие ВЦ. При этом способе пользователь доставляет на ВЦ исходную информацию и получает результаты обработки в виде результативных документов. Особенностью такого способа обработки являются сложность и трудоемкость налаживания быстрой, бесперебойной связи, большая загруженность ВЦ информацией (так как велик ее объем), регламентацией сроков выполнения операций, организация безопасности системы от возможного несанкционированного доступа.

Децентрализованная обработка. Этот способ связан с появлением ПЭВМ, дающих возможность автоматизировать конкретное рабочее место. В настоящее время существуют три вида технологий децентрализованной обработки данных.

Первая основывается на персональных компьютерах, не объединенных в локальную сеть (данные хранятся в отдельных файлах и на от-

дельных дисках). Для получения показателей производится перезапись информации на компьютер. Недостатки: отсутствие взаимоувязки задач, невозможность обработки больших объемов информации, низкая защита от несанкционированного доступа.

Вторая. ПК, объединенные в локальную сеть, что ведет к созданию единых файлов данных (но он не рассчитан на большие объемы информации).

Третья. ПК, объединенные в локальную сеть, в которую включаются специальные серверы (с режимом «клиент — сервер»).

Распределенный способ обработки данных основан на распределении функций обработки между различными ЭВМ, включенными в сеть. Этот способ может быть реализован двумя путями: первый предполагает установку ЭВМ в каждом узле сети (или на каждом уровне системы), при этом обработка данных осуществляется одной или несколькими ЭВМ в зависимости от реальных возможностей системы и ее потребностей на текущий момент времени. Второй путь — размещение большого числа различных процессоров внутри одной системы. Такой путь применяется в системах обработки банковской и финансовой информации, там, где необходима сеть обработки данных (филиалы, отделения и т.д.).

Преимущества распределенного способа: возможность обрабатывать в заданные сроки любой объем данных; высокая степень надежности, так как при отказе одного технического средства есть возможность моментальной замены его на другой; сокращение времени и затрат на передачу данных; повышение гибкости систем, упрощение разработки и эксплуатации программного обеспечения и т.д.

Распределенный способ основывается на комплексе специализированных процессоров, т.е. каждая ЭВМ предназначена для решения определенных задач или задач своего уровня.

Следующий способ обработки данных — интегрированный. Он предусматривает создание информационной модели управляемого объекта, т.е. создание распределенной базы данных. Такой способ обеспечивает максимальное удобство для пользователя. С одной стороны, базы данных предусматривают коллективное пользование и централизованное управление. С другой стороны, объем информации, разнообразие решаемых задач требуют распределения базы данных. Технология интегрированной обработки информации позволяет улучшить качество, достоверность и скорость обработки, так как обработка производится на основе единого информационного массива, однократно введенного в ЭВМ. Особенностью этого способа является отделение технологически и по времени процедуры обработки от процедур сбора, подготовки и ввода данных.

## 2.3. Основные этапы технологического процесса и составляющие их операции

Технологический процесс обработки информации можно разделить на процессы сбора и ввода в систему исходных данных, процессы размещения и хранения данных в памяти вычислительной системы, процессы обработки данных с целью получения определенных результатов и процессы выдачи данных в виде, удобном для восприятия человеком.

Технологический процесс можно разбить на четыре укрупненных этапа.

1. Начальный, или первичный, представляет сбор исходных данных, их регистрацию и передачу на устройство обработки информации.

2. Подготовительный этап предназначен для приема, контроля, регистрации входной информации и переноса ее на машинный носитель.

3. Основной заключается в непосредственной обработке информации.

4. Заключительный этап предназначен для контроля, выпуска и передачи результатной информации, ее размножения и хранения.

В зависимости от применяемых технических средств и требований к технологиям обработки информации видоизменяется и состав операций технологического процесса. Например, исходная информация на устройство обработки может быть предоставлена на машинных носителях информации, подготовленных для ввода в персональный компьютер, или быть передана по каналам связи в месте ее возникновения.

Операции сбора и регистрации данных осуществляются с помощью разнообразных средств. Различают следующие способы сбора и регистрации данных:

- механизированный;
- автоматизированный;
- автоматический.

Механизированный способ представляет собой сбор и регистрацию информации, осуществляемую человеком непосредственно с использованием простейших приборов (весов, счетчиков, мерной тары, приборов учета времени и т.д.).

Автоматизированный способ использует машиночитаемые документы, регистрирующие автоматы, универсальные системы сбора и регистрации, обеспечивающие совмещение операций формирования первичных документов и получения машинных носителей.

Автоматический способ применяется, как правило, при обработке информации в режиме реального времени. Информация с датчиков, учитывающая ход производственных процессов — выпуск продукции, затраты сырья, простой оборудования и т.д., — поступает непосредственно в компьютер.

Технические средства передачи информации включают:

- аппаратуру передачи информации (АПИ), объединяющую средства обработки и подготовки данных с телефонными, телеграфными и широкополосными каналами связи;
- устройства сопряжения персонального компьютера с аппаратурой передачи, управляющей обменом информацией, — мультиплексорами передачи информации.

Запись и передача информации по каналам связи на устройство обработки обладает рядом преимуществ:

- упрощается процесс формирования и контроля информации;
- соблюдается принцип однократной регистрации информации в первичном документе и машинном носителе;
- обеспечивается значительная достоверность информации, поступающей в обработку.

Дистанционная передача данных, основанная на применении каналов связи, представляет собой передачу информации в виде электрических сигналов, имеющих непрерывный или дискретный характер. Наиболее широко применяются телефонные и телеграфные каналы связи. Сигналы, передаваемые по телеграфному каналу связи, являются дискретными, а по телефонному — непрерывными. При выборе оптимального способа передачи информации должны быть учтены временные и объемные параметры доставки, требования к качеству передаваемой информации, стоимостные и трудовые затраты, связанные с передачей информации.

Говоря о технологических операциях, связанных со сбором, регистрацией и передачей информации при помощи разнообразных технических средств, необходимо рассказать и о сканирующих устройствах.

Ввод информации, в особенности графической, при помощи клавиатуры чрезвычайно трудоемок. В настоящее время наметились тенденции использования деловой графики, являющейся одним из основных видов информации, требующих оперативности ввода в компьютер и предоставления пользователям возможности создания гибридных документов и баз данных, связывающих графику с текстовым содержанием. Данные функции в компьютере возложены на сканирующие устройства, реализующие оптический ввод исходной информации и преобразование ее в цифровую форму для возможности последующей обработки.

Сканирующие устройства на базе персонального компьютера нашли применение не только для ввода графической или текстовой информации, но и в системах контроля, обработки писем, а также выполнения многообразных учетных функций.

Для указанных задач наибольшее применение нашли способы кодирования информации штриховыми кодами. Сканирование штриховых кодов для ввода информации в ЭВМ производится с помощью миниатюрных сканеров, напоминающих карандаш. Сканер перемещается пользователем перпендикулярно группе штрихов, внутренний источник света освещает область этого набора непосредственно около наконечника сканера. Штриховые коды нашли широкое применение и в сфере торговли, и на предприятиях (в системе табельного учета: при считывании с карточки работника фактически отработанного времени, регистрация времени, даты и т.д.).

В последнее время все большее внимание уделяется устройствам тактильного ввода — сенсорному экрану («сенсорный» — чувствительный). Устройства тактильного ввода широко применяются как информационно-справочные системы общего пользования и системы автоматизированного обучения. Сенсорный экран широко применяется для фондовых бирж (например, по нему передаются сведения о последних продажных ценах на акции).

На практике существует множество вариантов (организационных форм) технологических процессов обработки данных. Это зависит от использования различных средств вычислительной и организационной техники на отдельных операциях технологического процесса.

Построение технологического процесса зависит от характера решаемых задач, круга пользователей, от используемых технических средств, от систем контроля данных и т.д.

Технологический процесс обработки информации с использованием ЭВМ включает в себя следующие операции.

1. Прием и комплектовка первичных документов — проверка полноты и качества их заполнения, комплектовки и т.д.
2. Подготовка МН и контроль.
3. Ввод данных в ЭВМ.
4. Контроль, результаты которого выдаются на ПУ, терминал. Различают визуальный и программный контроль, позволяющий отслеживать информацию на полноту ввода, нарушение структуры исходных данных, ошибки кодирования. При обнаружении ошибки производится исправление вводимых данных, корректировка и их повторный ввод.
5. Запись входной информации в исходные массивы.

6. Сортировка (если в этом есть необходимость).
7. Обработка данных.
8. Контроль и выдача результатной информации.

Перечисляя операции технологического процесса, хотелось бы несколько слов сказать об операции хранения информации. Еще совсем недавно информация хранилась на таких машинных носителях, как перфокарты, перфоленты, магнитные ленты, магнитные диски. С развитием средств вычислительной техники изменились и носители информации. Дискета (гибкий магнитный диск), которая подвергалась постоянно изменениям как внешне, так и объемом записываемой информации, в настоящее время уже не отвечает требованиям пользователей. Это касается не только технической надежности носителей информации, но и объема хранимой информации. Современные экономические информационные системы с мощными процессорами, оснащенными съемными винчестерами, CD-ROM'ами с лазерными дисками, обеспечивают более высокую скорость обработки информации и предоставляют пользователю работать с большими объемами данных, обеспечивая удобство в работе и надежность в сохранности информации.

Проектирование рациональных технологических процессов обработки данных является довольно сложной задачей. Эта сложность обуславливается тем, что сама автоматизированная система обработки информации относится к классу сложных систем, и при ее разработке должны учитываться многие параметры, среди которых не только чисто технические, но и параметры, учитывающие различные человеческие факторы, вопросы повышения сроков эксплуатации и использования инструментальных средств, уменьшения сроков разработки, ряд экономических соображений и т.д.

Технология проектирования автоматизированной обработки информации при решении любой задачи подразделяется на четыре этапа: начальный, подготовительный, основной, заключительный.

Состав и структура операций каждого из этапов технологического процесса могут быть различными в зависимости от используемых средств ВТ, средств оргсвязи и требований к технологии преобразования информации. По своему назначению технологические операции бывают вспомогательными, основными и контрольными. Вторые составляют основу и относятся к операциям внутримашинной технологии обработки данных. Это операции упорядочения, корректировки, накопления и собственно обработки.

Упорядочение — произвольно расположенные данные размещаются в определенной последовательности значений ключевых слов.

Корректировка — процесс внесения изменений в уже сформированные файлы данных, позволяющий поддержать их в актуальном для обработки состоянии.

Накопление — процесс периодического добавления данных в существующие файлы с целью формирования исходных данных за определенный интервал времени.

Обработка — выполнение всех арифметических и логических операций по преобразованию исходной информации в результатную.

Существуют различные формы внутримашинной технологии обработки информации. Наиболее распространенными формами являются обработка данных в пакетном и диалоговом режимах.

Иногда автоматизированное решение задач должно согласовываться по времени с ходом управляемых процессов. Соответственно организация обработки информации для этих нужд получила название технологии обработки данных в режиме реального времени. Важной характеристикой, определяющей область применения режима реального времени, является скорость реакции системы управления на изменение состояний объекта управления.

В настоящее время прослеживается тенденция к максимальному приближению информационных и программных ресурсов к пользователю. ЭВМ, работающие в сети, имеют существенное преимущество перед АРМ, работающими в режиме разделения времени. А главное, средства интеллектуального интерфейса обеспечивают пользователя простыми и надежными способами решения своих профессиональных задач.

Возвращаясь к вопросу об этапах разработки технологических процессов, необходимо сказать, что на заключительном этапе производится контроль и выпуск результатных документов.

Известно, что все этапы разработки технологических процессов (предпроектная стадия, техническое проектирование, стадия рабочего проектирования, ввода в действие, функционирование, сопровождение, модернизация) документируются.

Документирование — оформление описания выбранных вариантов построения информационной технологии с комментариями, обеспечивающими их использование в процессе эксплуатации системы.

Наличие документального обоснования позволяет проверить правильность варианта.

## 2.4. Технологии электронного офиса

Электронный офис организуется на предприятиях, в организациях или учреждениях в управленческих структурах.

Основными функциями электронного офиса являются:

- автоматизация рутинных работ с документами;
- организация электронного документооборота.

Для реализации перечисленных выше функций в электронном офисе выполняется ряд стандартных типовых процедур обработки документов.

В состав электронного офиса входят аппаратные и программные средства реализации типовых офисных процедур обработки информации.

Программные средства электронного офиса — это пакеты прикладных программ, которые позволяют автоматизировать основные процедуры обработки информации в процессе управления.

Программные средства электронного офиса делятся на три группы:

1) программы, предназначенные для создания и редактирования документов. Основой функционирования любого офиса является документ. В электронном офисе циркулируют документы, содержащие данные разных типов (текстовую информацию, табличные данные, графические изображения и т.д.). Офисные программы для создания и редактирования документов различных типов, как правило, интегрируют в единый пакет, который позволяет работать с одним документом, включая в него различные виды данных. Существуют и самостоятельные программные продукты, ориентированные на выполнение только одного вида работ: разработку электронных таблиц, редактирование текстов, электронную верстку документов и т.д.;

2) организаторы работ — это пакеты программ, предназначенные для автоматизации процедур планирования использования различных ресурсов (времени, денег, материалов) как отдельного человека, так и всей фирмы или ее структурных подразделений. Выделяют две разновидности пакетов данного класса:

- программы управления проектами предназначены для сетевого планирования и управления проектами. Эти программные средства позволяют быстро спланировать проект любой величины и сложности, эффективно распределить людские, финансовые и материальные ресурсы, составить оптимальный график работ и проконтролировать его исполнение,
- программы организации деятельности человека выполняют функции электронных секретарей и предназначены для эффективного управления деловыми контактами. В них предусмотрены следующие основные функции:
  - формирование графика деловой активности с автоматическим контролем со стороны программы,
  - ведение электронной картотеки, полностью настраиваемой под конкретные нужды пользователя,

- обеспечение безопасности и конфиденциальности данных,
- автоматический набор телефонных номеров (с автодозвоном) и занесение даты и времени звонка в соответствующую карточку,
- работа с электронной почтой и факсом и т.д.;

3) программы организации электронного документооборота позволяют решать следующие задачи:

- обеспечение более эффективного управления за счет автоматического контроля выполнения, прозрачности деятельности всей организации на всех уровнях,
- поддержка эффективного накопления, управления и доступа к информации и знаниям, обеспечение кадровой гибкости за счет большей формализации деятельности каждого сотрудника и возможности хранения всей предыстории его деятельности,
- оптимизация бизнес-процессов и автоматизация механизма их выполнения и контроля,
- исключение бумажных документов из внутреннего оборота предприятия,
- исключение необходимости или существенное упрощение и удешевление хранения бумажных документов за счет наличия оперативного электронного архива.

Аппаратные средства электронного офиса — это устройства, обеспечивающие техническую реализацию офисных процедур обработки данных. Они подразделяются на основные и дополнительные:

1) основными техническими средствами электронного офиса являются один или несколько ПК, объединенные в вычислительную сеть и имеющие широкий набор периферийных устройств, таких как устройства вывода информации на печать (принтеры), устройства автоматического ввода информации (сканеры, дигитайзеры) и др.;

2) дополнительные аппаратные средства электронного офиса служат для реализации операций копирования, размножения, оформления и уничтожения документов.

## 2.5. Технологии работы с документами

Любое управленческое решение оформляется в виде документа, поскольку только документ, отпечатанный на бумаге, имеет юридическую силу. Но что делать, когда документов масса, когда приходится отслеживать по документам все этапы процесса или окончательное решение требует коллективной работы над документом? В этом случае сегодня

имеется реально работающая технология, упрощающая рутинную работу с потоками документов, — технология электронного документооборота.

Каждый вид деятельности вызывает появление своего класса информации, поэтому для обеспечения решения задач управления, в частности в большинстве видов менеджмента, рассмотренных в темах 1—5, можно выделить несколько видов информации:

- управленческая;
- экономико-статистическая;
- научно-техническая;
- юридическая и др.

Ведущее место занимает управленческая информация, содержащая цели и подцели функционирования предприятия (компании), определяемые на базе использования всех остальных классов информации и, в первую очередь, экономической информации. Наиболее массовой и удобной для восприятия является документальная форма существования такой информации.

**Документ** — это материальный носитель, на который наносятся некоторые сведения, отражающие состояние системы, или принятое решение строго установленного содержания по строго регламентированной форме. Он обладает двумя отличительными свойствами: полифункциональностью и наличием юридической силы. К числу функций, которые реализуются с помощью документа, относятся: регистрация первичной информации или принятого решения, передача, обработка и хранение информации. Наличие юридической силы обеспечивается реквизитом — подписью лица, ответственного за достоверность сведений, содержащихся в документе.

Совокупность взаимосвязанных документов, систематически используемых для процессов управления объектом, называется системой документации (СД). В состав системы документации предприятия входит несколько сот форм различных документов, которые можно классифицировать по разным признакам.

В системе документации, используемой в любой экономической системе, по содержанию можно выделить класс управленческих или организационно-распорядительских документов (ОРД), используемых для административного управления; экономико-статистических документов, предназначенных для экономического управления предприятием; научно-технических документов, возникающих при проектировании и производстве новых видов продукции; юридических документов, регламентирующих отношения коллектива как внутри предприятия, так и его отношения с внешней средой, включая вышестоящие и правительствен-

ные организации, предприятия-потребители готовой продукции или услуг, предприятия-поставщики сырья и материалов и др.

Каждый выделенный класс в свою очередь по виду может делиться на подмножества документов. Так, среди экономико-статистических документов выделяют плановые, финансовые, статистические, бухгалтерские документы и др. К числу ОРД относят организационные, распорядительские и информационно-справочные документы. Научно-технические документы включают в себя стандарты, чертежи, техническую и технологическую документацию по изготовлению продукции. Юридические документы можно разделить по видам права: гражданского, уголовного, налогового, арбитражные документы и пр. Каждый вид документов можно подразделить по степени официальности на два класса: официальные и личные. По месту составления все документы делятся на внешние (входящие) и внутренние, которые могут оставаться как на предприятии, так и отправляться другим организациям, т.е. быть исходящими.

По уровню управления все документы можно разделить на четыре группы:

- государственного уровня;
- министерств;
- объединений — корпораций;
- предприятий, фирм или организаций.

По количеству вопросов, содержащихся в документе, можно выделить простые (например, докладная записка) и сложные документы (например, бизнес-план, договор). По степени подлинности документы подразделяются на подлинники, черновики и копии. По времени хранения выделяют документы постоянного и временного хранения. По методу обработки документов и их заполнения можно выделить четыре группы документов:

- ручного заполнения;
- машинно-ориентированные документы, предназначенные для ввода с клавиатуры содержащейся в них информации в ЭВМ;
- машинно-читаемые документы, информация которых автоматически вводится в базу данных;
- автоматического заполнения и выдачи информации на печать.

По степени гласности выделяют обычные документы, служебного пользования и секретные документы. По назначению существуют индивидуальные документы произвольной формы и документы с типовой формой. По языку отображения информации выделяют текстовые документы, табличные, графические, документы с аудиоинформацией и мультимедийные документы.

**Документопоток** — это процесс передвижения документов одного типа от источника возникновения или пункта обработки к потребителю. Документопотоки связывают все подразделения экономической системы в единую информационную систему. Документопотоки экономической системы делятся:

- на внешние, входящие в систему;
- внутренние документопотоки, имеющие обращение в системе и предназначенные для удовлетворения внутренних информационных потребностей;
- исходящие, связывающие данную систему с другими организациями и предприятиями.

Для организации эффективного управления экономической системой и рациональной автоматизации информационных процессов необходимо выявлять потоки информации, производить их оценку и оптимизацию.

Поскольку документы возникают и передвигаются в системе по мере выполнения функций управления или каких-либо деловых процессов, то на каждом предприятии, организации и фирме создается свой документооборот. Под документооборотом понимается регламентированная совокупность взаимосвязанных операций, выполняемых над документом в строго установленном порядке, на определенном рабочем месте с использованием определенных методов и средств, т.е. технологии обработки, начиная от момента возникновения документа и заканчивая сдачей его в архив. Документооборот на предприятии может быть двух типов.

1. **Операционный** — ориентированный на обработку первичных и получение сводных и аналитических документов, содержащих операционную атрибутику и используемых на первой и второй стадиях управления.

2. **Универсальный** — отражающий информационные потоки слабоструктурированной информации, используемой на третьей стадии управления, т.е. при исполнении управленческих решений или деловых процессов и процедур. Поскольку состав документов и процедур их обработки, состав потоков документов и структура документооборота в значительной степени определяется особенностями организационной структуры конкретного предприятия, а также происходящими на предприятии процессами управления и деловыми процессами, поэтому далее остановимся на рассмотрении этих вопросов.

Концепция электронного документа появилась в 80-х гг. XX в. с появлением на рынке персональных компьютеров и первого графического интерфейса пользователя (GUI) (рис. 2.2). Как утверждают сотрудники компании Delphi Consulting Group, уже сегодня в некоторых областях деятельности электронным способом обрабатывается до 90% информации.

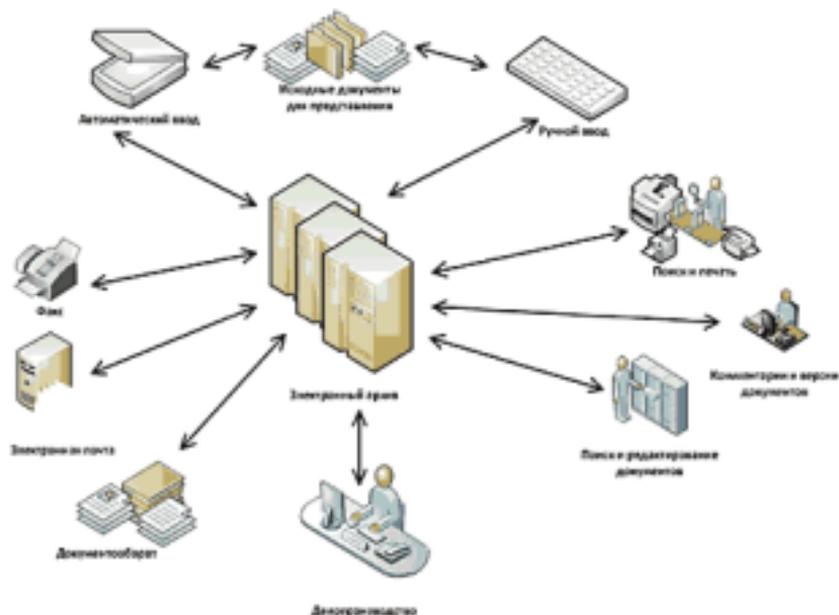


Рис. 2.2. Система управления электронными документами (СУД)

В отличие от документов на бумажных носителях с их жесткими рамками, статичной формой и ограниченными возможностями переход к динамичным цифровым электронным документам обеспечивает особые преимущества при создании, совместном использовании, распространении и хранении информации. Они могут увеличить производительность множества приложений, используемых в бизнесе, уменьшить требования к размерам накопителей, сократить или полностью решить проблему потерь и неверного размещения документов. Электронные документы могут одновременно использоваться сотрудниками в рамках одной рабочей группы, отдела или всего предприятия. Доступ к ним осуществляется за несколько секунд, а не минут, часов, дней, а иногда и недель, что случается при использовании документов на бумажных носителях. Ускоренный доступ к стратегической информации наряду со значительной экономией средств может обеспечить и важные конкурентные преимущества.

Помимо перечисленного электронные документы позволяют переместить центр тяжести компьютерной технологии с традиционных структурированных алфавитно-цифровых данных на потоки данных, допол-

ненные большими объемами неструктурированного текста, изображений, звука, видео и графики. Такие документы смогут также включать гипертекстовые связи, переработанные OLE-объекты, текстовые объекты и реляционные данные. Электронный документ будет ограничен такими параметрами, как его содержимое, структура данных, форматы и стандарты режима передачи и, самое важное, характер его использования. При изменении любого из этих параметров соответственно будет меняться документ. Он будет открытым, гибким, адаптируемым, многомерным.

За несколько лет концепция электронного документа получила свое развитие от обычного графического образа документа до идеи управления документами. Сегодня электронный документ — это форма знакомого вида, обработка которой происходит с помощью последовательного применения тесно взаимосвязанных технологий в рамках так называемых Систем управления электронными документами (СУД) или Electronic Document Management Systems (EDMS). Мощные системы класса EDMS, работающие чаще всего на распределенных архитектурах, основаны на комбинации технологий сбора, хранения, поиска и распространения электронных документов. Значительное повышение производительности при использовании технологии EDMS отмечено во многих отраслях: правительство, судопроизводство, юриспруденция, фармацевтика, страхование, финансовые услуги, здравоохранение, производство и система высшего образования.

Можно выделить следующие основные задачи, которые призвана решать система управления документами: уметь переводить бумажные документы в электронный вид и уметь перехватывать все создаваемые электронные документы из всех источников поступления:

- из приложений, которые создаются текстовыми процессорами, электронными таблицами, графическими процессорами, пакетами прикладных программ для настольных издательств;
- факсовые сообщения, пришедшие через факс-сервер предприятия;
- входящие, исходящие и внутренние сообщения, поступающие из электронной почты;
- отчеты и другие документы, которые готовят функциональные информационные системы, идущие на печать или отправляемые по факсу, но их надо перехватить и отправить в хранилище электронных документов.

Многофункциональные системы управления документами развиваются под влиянием перехода от вычислительных сред на базе мейнфреймов к системам на базе микропроцессоров, от монолитных закры-

тых приложений к открытым и интегрируемым программным компонентам.

Содержанием третьей концепции является разработка Системы электронного документооборота (СЭДО), призванной для интеграции всех информационных приложений в единую информационную среду, обеспечивающую оперативное взаимодействие всех пользователей при выполнении ими деловых процедур и функций управления необходимой информацией. Система электронного документооборота должна позволять планировать и составлять маршруты передвижения документов, контролировать это передвижение, уметь управлять документооборотом и регулировать его. Эта концепция имеет две стадии в своем развитии от воплощения идеи использования технологии docflow<sup>1</sup>, связанной с планированием и управлением документопотоками, до технологии workflow<sup>2</sup>, предназначенной для управления деловыми процессами и процедурами, при выполнении которых используются и создаются документы.

Системы, основанные на применении технологии workflow, получили название Систем автоматизации деловых процессов (САДП) и предназначены для моделирования деловых процессов и процедур, выполняемых на предприятии, контроля их исполнения, оценки и анализа эффективности их выполнения, реструктуризации системы управления фирмы.

В состав каждого предприятия включается самостоятельное структурное подразделение, основной целью функционирования которого является своевременное обеспечение документами процессов управления и деловых процедур. Таким подразделением является Система документального обеспечения управления (СДОУ), которая предназначена для решения следующих задач:

- документирования принимаемых организационных, плановых и административных управленческих решений;

<sup>1</sup> Дословный перевод означает «поток документов», где фактически раскрывается смысловое понимание термина, когда при разработке системы в качестве моделей — исходных элементов проектирования моделируются не только маршруты проведения документов, но и работа смысловой информации, заключенной в документе. Аналогом может служить диаграмма процессов, рассматриваемая при решении задач автоматизации управления.

<sup>2</sup> Поток работ (от англ. Workflow) — графическое представление потока задач в процессе и связанных с ним подпроцессов, включая специфические работы, информационные зависимости и последовательность решений и работ. Для изображения потока работ используют блок-схему или граф, который состоит из операций (работ), символов логики, стрелок. Разветвления блок-схемы имеют логические символы «и», «или». Стрелки используют для обозначения последовательности выполнения операций или потока объектов (документы, ресурсы). Кроме того модель потока работ может отображать исполнителей, используемое оборудование, программные средства и т.п. Поток работ в информационном смысле — способ поступления информации к различным объектам, участвующим в процессе. В частности, способ поступления документов к работникам.

- документационного обеспечения управления (ДОУ);
- получения, фильтрации и распределения потоков внешней и внутренней информации руководителям и исполнителям подразделений в соответствии с их полномочиями;
- контроля их исполнения, анализа качества и сроков исполнения;
- организации хранения, поиска документов и выдачи документов, необходимых для выполнения функций управления или деловых процессов и процедур.

Документирование представляет собой процедуру создания документов, отражающих факты, события или показатели, получаемые при выполнении функций управления или деловых процессов, т.е. их составление, оформление, согласование и изготовление (рис. 2.3).

Документационное обеспечение управления (ДОУ) осуществляется при выполнении следующих видов деятельности:

- делопроизводство, т.е. организация работы с документами;
- формирование и хранение архива документов.



Рис. 2.3. Документационное обеспечение управления

Существуют системы документооборота, которые можно настроить на необходимые правила делопроизводства. Но есть системы, которые изначально ставили своей целью поддержку именно этих правил и в этом

смысле не обладают более общими функциями в достаточной мере для того, чтобы их можно было назвать системами документооборота. С другой стороны, свод правил делопроизводства, который принят в России, довольно обширен, и настройка системы документооборота на эти правила — задача сложная и нетривиальная. Отсюда вытекает вывод: чтобы автоматизировать делопроизводство и документооборот на предприятии, необходимо предварительно хорошо изучить технологию выполнения основных процедур, составляющих содержание делопроизводства и документооборота на предприятиях и в организациях России.

В зависимости от объема информации, которую служба СДОУ получает и обрабатывает, эти системы делятся на четыре класса:

- к первому классу относятся СДОУ, обрабатывающие более 100 тыс. документов в год;
- ко второму классу относятся системы, обрабатывающие от 25 тыс. до 100 тыс. документов в год;
- в третий класс входят системы с обработкой от 10 тыс. до 25 тыс. документов в год;
- в четвертый — менее 10 тыс. документов в год.

Выделенным классам в свою очередь соответствуют следующие организационные формы СДОУ:

- управление делами (например, Управление делами при Президенте РФ);
- общий отдел (в министерствах);
- отдел делопроизводства (на предприятиях и в организациях);
- секретарь руководителя (в небольших фирмах).

Наиболее массовой формой организации работы СДОУ является отдел делопроизводства, структура которого представлена на рис. 2.4.

На территории РФ еще в XX в. приняты «Единые правила составления и оформления документов по РФ». Эти правила определяют:

- рациональные принципы документооборота;
- единые правила создания учетно-справочной системы;
- формы и методы контроля исполнения документов;
- правила составления номенклатуры дел;
- принципы формирования дел;
- правила подготовки дел к последующему хранению и использованию;
- рекомендации по повышению эффективности труда в делопроизводстве и автоматизации делопроизводственных процессов.

А с 1993 года действует «Типовая инструкция по делопроизводству для министерств и ведомств», устанавливающая общие требования к документированию управленческой деятельности и технологиям работы



Рис. 2.4. Структура подразделения «Отдел делопроизводства»

с документами в упомянутых организациях. Инструкция впервые закрепила не только требования к оформлению реквизитов, но и к тексту конкретных видов документов (приказу, распоряжению, протоколу, акту).

Целью разработки технологии документооборота является повышение эффективности управления экономическими системами на основе автоматизации управления документооборотом и деловыми процессами, всех видов работ с документами, обеспечивающими и координирующими совместную деятельность всех участников процесса управления. Системы документооборота обычно внедряются, чтобы решать определенные задачи, стоящие перед организацией, к таким задачам можно отнести следующие:

- обеспечение более эффективного управления за счет автоматического контроля выполнения, прозрачности деятельности всей организации на всех уровнях;
- поддержка системы контроля качества, соответствующей международным нормам;

- поддержка эффективного накопления, управления и доступа к информации и знаниям. Обеспечение кадровой гибкости за счет большей формализации деятельности каждого сотрудника и возможности хранения всей предыстории его деятельности;
- протоколирование деятельности предприятия в целом (внутренние служебные расследования, анализ деятельности подразделений, выявление «горячих точек» в деятельности);
- оптимизация бизнес-процессов и автоматизация механизма их выполнения и контроля;
- исключение бумажных документов из внутреннего оборота предприятия. Экономия ресурсов за счет сокращения издержек на управление потоками документов в организации;
- исключение необходимости или существенное упрощение и удешевление хранения бумажных документов за счет наличия оперативного электронного архива.

Внедрение электронного документооборота позволяет достичь реализации следующих целей:

- интеграции профессиональной и информационной деятельности специалистов;
- интеграции всех информационных потоков и создания единого информационного ресурса, используемого для принятия управленческих решений;
- коллективного использования информационного ресурса, упрощение процесса обмена информацией;
- оптимизации работ сотрудников и сокращения затрат труда и времени на администрирование их совместной деятельности.

Построение технологии (системы) электронного документооборота, как правило, осуществляется на основании сформировавшихся в течение последних 30 лет принципах. Первый принцип — ориентация на работу с документами. Из этого принципа вытекает как следствие тот факт, что информационное обеспечение ЭСУД должно предоставлять возможность интегрированной обработки всех видов информации, циркулирующей в организации, в том числе документов, порожденных электронным и бумажным документооборотом: внешней и внутренней переписки, осуществляемой как в электронной, так и в бумажной форме. Основной частью такой системы должна быть база данных электронных документов (БДД), которая становится элементом централизованной базы данных всей организации или предприятия и формируется как централизованный электронный архив документов (включаящий в том числе и бумажные оригиналы, и электронные копии оригиналов бумажных документов).

Система управления базой данных документов должна обеспечивать выполнение следующих функций:

- централизованную регистрацию всех документов, которые циркулируют в организации;
- хранение документов в электронном виде в различных форматах;
- ведение централизованного каталога документов организации, обеспечивающего возможность их поиска (по ключевым атрибутам, с использованием полнотекстового поиска и т.д.);
- хранение полной истории работы с документами (кто, когда и как работал с документом), а также различных версий документов;
- надежную систему защиты документов, регламентацию доступа персонала к документам различного назначения;
- возможность поддержки архивов документов на всех видах внешних устройств, включая магнитооптические и библиотеки CD-ROM.

Прикладное ПО должно включать следующие ключевые компоненты:

- систему управления хранением документов — программное обеспечение, реализующее функции управления единым документарным фондом организации (централизованным архивом);
- систему управления документооборотом — программное обеспечение, реализующее администрирование документооборота, управление маршрутизацией и движением документов, координацию документопотоков, контроль за передвижением документов, за своевременной их обработкой и т.д.;
- набор стандартных бизнес-приложений, используемых сотрудниками организации для подготовки документов — текстовых процессоров, электронных таблиц и т.п., набор специализированных функциональных приложений, предназначенных для подготовки документов (в отличие от стандартных бизнес-приложений они взаимодействуют с базой данных, поддерживающей структурированную информацию);
- систему экспорта (импорта) документов.

В качестве центрального управляющего блока программного обеспечения ЭСУД выступает система управления полномочиями пользователей, которая призвана осуществлять разграничение доступа пользователей к информации (в том числе к документам различной степени секретности) и регламентацию доступа пользователей к функциям, предоставляемым системой.

**Первый принцип** разработки технологии документооборота позволяет определить состав основных функциональных подсистем, к числу которых можно отнести (рис. 2.5):

<p style="text-align: center;"><b>СИСТЕМА СОЗДАНИЯ ЭЛЕКТРОННЫХ ДОКУМЕНТОВ</b></p>	<p style="text-align: center;"><b>СИСТЕМА ПРЕОБРАЗОВАНИЯ БУМАЖНЫХ ДОКУМЕНТОВ В ЭЛЕКТРОННЫЙ ВИД</b></p>
<p style="text-align: center;"><b>СИСТЕМА ХРАНЕНИЯ И ПОИСКА ЭЛЕКТРОННЫХ ДОКУМЕНТОВ</b></p>	<p style="text-align: center;"><b>СИСТЕМА ЭЛЕКТРОННОГО ДОКУМЕНТООБОРОТА И УПРАВЛЕНИЯ БИЗНЕС-ПРОЦЕССОМ</b></p>

Рис. 2.5. Схема функциональной структуры ЭСУД

- подсистему составления электронных документов (Система создания ЭД);
- подсистему организации массового ввода, преобразования бумажных документов в электронный вид и загрузки их в электронное хранилище (Система массового ввода бумажных документов — СМВ);
- подсистему осуществления хранения, поиска и выдачи электронных документов по запросам пользователей (Система управления электронными документами — СУД);
- подсистему планирования маршрутизации и контроля исполнения документов и выполнения деловых процессов с целью своевременного документационного обеспечения управления (Система электронного документооборота — СЭДО).

**Второй принцип** создания — разработка «активного» программного обеспечения для реализуемой системы. Работа с документами является основным способом выполнения функций для персонала организации. Цель автоматизации процессов движения документов и связанных с ними заданий — получение целостного механизма управления деятельностью предприятия, поэтому ПО должно выполнять следующие функции:

- обеспечивать возможность руководству организации распределять работы и контролировать процесс их выполнения исполнителями;

- до каждого исполнителя доводить информацию о том, какие работы, в какие сроки и в какой последовательности он должен выполнять;
- доставлять на рабочие места исполнителей документы и поручения;
- выстраивать список работ, которые должен выполнить исполнитель;
- обеспечивать исполнителю возможность после выбора из списка соответствующей работы, «погружения» его именно в то функциональное приложение, которое связано с соответствующим типом работы (или задания).

Согласно **третьему принципу** разработки, программное обеспечение должно легко адаптироваться к изменениям конкретных условий эксплуатации. Отсюда вытекает необходимость разрабатывать и внедрять ПО ЭСУД двумя методами:

- методом компонентного проектирования,
- с использованием средств CASE-технологии<sup>1</sup> проектирования ИС.

Создание и обработка формы электронного документа обычно включает в себя выполнение нескольких операций (рис. 2.6):



Рис. 2.6. Схема составления и обработки ЭД

Первая операция — проектирование структуры формы электронного документа (ЭД) заключается в рисовании линий, создании графических элементов (например, логотипов), т.е. подготовке внешнего вида ЭД с помощью графических средств. Поскольку переменные поля могут быть заполнены вручную или посредством выбора значений из какой-либо базы данных (в последнем случае должна быть установлена связь формы с базой данных), то второй операцией будет проектирование содержания формы ЭД.

<sup>1</sup> CASE (от англ. Computer-Aided Software Engineering) — набор инструментов и методов программной инженерии для проектирования программного обеспечения, который помогает обеспечить высокое качество программ, отсутствие ошибок и простоту в обслуживании программных продуктов. Также под CASE понимают совокупность методов и средств проектирования информационных систем с использованием CASE-инструментов.

Содержанием этой операции является привязка переменных полей формы, которые будут заполняться, к инструментальным средствам заполнения, или вычисления, или автоматического контроля, т.е. на этой операции проектирования осуществляется придание форме свойства интеллектуальности. Почти все продукты обеспечивают удобные средства установления простых связей с соответствующими средствами, часть из них предоставляет высокоуровневые языки скриптов или макросы. Дизайнер форм также может указать, что при заполнении поля будут выполняться определенные задачи, такие как, например, вычисление суммы, проверка типов и т.д.

На третьей операции с целью повышения достоверности информации при заполнении форм ЭД устанавливаются связи между формами, что позволяет автоматизировать процессы актуализации сведений, заключенных в связанных формах.

На четвертом этапе осуществляется работа с готовой формой путем ее заполнения из различных источников, автоматическая проверка правильности заполнения переменных полей, в которые информация вводится вручную, проставление отметок о просмотре документа, выполнение дальнейшей пересылки его на рабочий стол другого специалиста или передача его по электронной почте или факсу.

Наиболее проблематичным моментом при работе с ЭД является проставление электронной подписи. Очень часто при внедрении систем электронного документооборота встает вопрос о том, чем заменить обычную подпись на бумажном документе в его электронном аналоге. Существует несколько способов реализации технологии электронной подписи:

- криптографическая методика открытых ключей защиты форм от последующих изменений;
- использование биометрических характеристик специалистов;
- применение отпечатков пальцев и изображений лиц и др.

Наиболее часто используется криптографическая методика открытых ключей защиты форм от последующих изменений, основным недостатком которой является недостаточная степень защищенности от взлома или подделки.

## 2.6. Примеры реализации технологий электронного документооборота

Мировому рынку СЭД скоро будет уже 20 лет. Он очень сильно фрагментирован, так как на нем присутствуют как всемирно известные многопрофильные ИТ-компании, так и относительно малоизвестные

(или известные только в своей рыночной нише) фирмы. По различным оценкам в мире сейчас существует несколько сотен программных приложений (которые можно классифицировать как СЭД), отличающихся друг от друга как по функциональным возможностям, так и по технологическим решениям. Разработкой приложений в области электронного документооборота в мире занимаются сотни компаний, к наиболее известным из которых относятся (по алфавиту): ACS Software, Action Technologies, Adobe, Artesia, AXS-One, BroadVision, Cyco, Cypress, Datamax Technologies, Datawatch, Divine, Documentum, Dynamic Imaging, Eastman Software, Excalibur, FileNet, Hyland Software, HP/Dazel, Hummingbird, Gauss Enterprise, IBM, Ideal, Identitech, iManage, Interlucent Internet Solutions, Interwoven, InterTech, Ixos Software, Jetform, Keyfile, Kofax, Lotus Development, Microsoft, Mobius Management Systems, Novell, OIT, OpenText, Optio Software, Optika, Oracle, OTG, Plexus, Radnet, RedDot Solutions, Siemens Nixdorf, SER Macrosoft, SER Solutions, Saperion, Saros, Staffware plc, Stellent, Symantec, Tower Software, Tower Technology, TrueArc, TSP; Unisys, Vignette, Westbrook Technologies и др.

Аналитики IDC считают перспективы мирового рынка технологий управления документами и содержимым (document and content technologies — DCT) достаточно благоприятными (отчет «Document and Content Technologies Applications Forecast and Analysis, 2000—2004») вследствие продолжающегося роста потребности корпоративных пользователей в повышении эффективности их коллективной работы с корпоративными документами (по оценке GartnerGroup, к концу 2001 г. в мире насчитывалось около 40 млн пользователей СЭД). В данном отчете IDC выделяет следующие сегменты DCT-рынка: собственно СЭД; системы управления содержимым на порталах предприятий и системы управления содержимым для электронной коммерции. По прогнозу IDC, объем мирового DCT-рынка должен увеличиться с 1,1 млрд дол. в 1999 г. до почти 4,4 млрд дол. в 2004 г. со средними ежегодными темпами роста в 32% (для сравнения: в соответствии с отчетом IDC «Document Management Market Review and Forecast: 1998—2003» в 1998 г. объем мирового рынка СЭД составлял около 750 млн дол., включая 200 млн дол. — для рынка Западной Европы). Развитию DCT-рынка способствуют также дальнейшее распространение электронной коммерции и увеличение потребности предприятий в Web-совместимых интегрированных инструментальных средствах доступа к информации. При этом особенно быстро растет спрос потребителей на инструментальные средства сбора, поиска и анализа информации, с помощью которых возможно более оперативно обрабатывать разнородные собрания текстовых файлов, графических файлов, видео- и ау-

диофайлов. По оценке IDC, в настоящее время продажи СЗД формируют основу доходов на DCT-рынке. Однако сегмент приложений для электронной коммерции развивается все же быстрее. В 1998—1999 годах объем данного сегмента вырос на 143,1%. Для сравнения, объем сегмента СЗД в рассматриваемый период увеличился только на 19,5%, а объем сегмента приложений для порталов предприятий — на 64,6%.

Более поздний прогноз IDC (отчет «Document and Content Technologies Market Forecast and Analysis Summary, 2001—2005») оценивает перспективы мирового рынка систем управления документами и содержимым в период экономического спада (и с переоценкой перспектив его развития после событий 11 сентября). И хотя мировой рынок этих систем не повторил свой экстраординарный рост в 89% (как в 2000 г.), IDC прогнозирует для него хорошие перспективы развития. Начальный прогноз IDC основан на данных, собранных в 2000 г. и I квартале 2001 г. Спрогнозирован рост данного рынка с ежегодными темпами в 47,2% (с 2 млрд дол. в 2000 г. до более чем 14 млрд дол. в 2005 г.). После трагедии 11 сентября прогноз был пересмотрен в сторону уменьшения. Аналитики IDC считают, что планы корпоративных пользователей по приобретению систем управления документами и содержимым будут пока отложены в течение первых трех кварталов 2002 г. Предполагается, что рынок восстановится к концу 2002 г. и в течение 2003—2005 гг. По мнению Gartner Group, спрос на системы управления корпоративными документами продолжится, однако их разработчики находятся под давлением рыночных условий (требующих дальнейшего снижения цен), продолжающейся дифференциации продуктов и необходимости обеспечения VAR-возможностей. Популярность порталов продолжит расти. По прогнозу IDC (отчет «Worldwide Enterprise Information Portal Software Forecast and Analysis, 2001—2006»), объем мирового рынка ПО для создания корпоративных информационных порталов (enterprise information portal — EIP) увеличится с 550,4 млн дол. в 2001 г. до 3,1 млрд дол. в 2006 г. В свою очередь, по прогнозу Gartner Group, средние ежегодные темпы развития рыночного сегмента порталов в течение ближайших 5 лет составят 30% (в то время как аналитики Ovum считают, что при существующей непростой ситуации в мировой экономике многие «портальные» инициативы будут просто «положены на полку»). Также значительно растет интерес к системам управления образами.

Не обошли стороной аналитики IDC и рынок услуг в области управления содержимым и документами (отчет «Content and Document Management Service Market Forecast, 2001—2006»). По прогнозу IDC, объем данного рынка будет увеличиваться ежегодно со средними тем-

пами роста в 44%, чтобы достигнуть к 2006 г. уровня 24,4 млрд дол. В перечень таких услуг IDC включает сервисы планирования и проектирования, а также сервисы внедрения, обучения и поддержки, предоставляемые клиентам для оказания им помощи в эффективном управлении корпоративным содержанием.

По общему мнению аналитиков, важность внедрения современных СЭД для обеспечения успешного ведения бизнеса сохранилась, и в обозримом будущем их значимость будет только усиливаться.

Потребность российских предприятий и организаций в оптимизации своего документооборота остается высокой и продолжает расти. За последние годы в России появились десятки компаний, занимающихся разработкой и поставкой СЭД, как зарубежных, так и собственной разработки. Первые масштабные внедрения СЭД (пусть в большинстве своем пилотные) в России уже есть. Можно считать, что основы российского рынка ПО управления документооборотом сформированы. По данным IDC, объем российского рынка СЭД (СЭД отечественной разработки вместе с зарубежными СЭД) в 1999 г. составил около 2 млн дол. По мнению аналитиков, с 1999 г. ежегодные темпы роста российского рынка СЭД составляют не менее 30%. Есть и такое мнение, что, начиная с 1998 г., наблюдается почти двукратное ежегодное увеличение объема российского рынка СЭД. Перспективы российского рынка СЭД улучшает принятие закона об электронной цифровой подписи, создающего правовую основу для распространения СЭД при межкорпоративном взаимодействии. Существует мнение ряда отраслевых аналитиков, что потенциальный объем российского рынка СЭД составляет сотни миллионов долларов (при успешном развитии российской экономики).

Следует отметить, что наблюдаемые сейчас небольшие размеры российского рынка СЭД связаны не в последнюю очередь с относительной незначительностью доли электронного документооборота в общем документообороте российских предприятий и организаций (которые могут позволить себе приобрести СЭД стоимостью от нескольких десятков до нескольких сотен тысяч долларов). В подавляющем большинстве случаев на российских предприятиях доминирует бумажный документооборот. Это обстоятельство объясняется не только традициями и определенным консерватизмом, но и непростым финансовым и техническим состоянием большинства российских предприятий и организаций. Следует сказать, что на российском рынке присутствует и ряд известных зарубежных систем (Documentum, DOCS Open/Fusion, Staffware, Panagon, DocuLive, Lotus Notes и др.). Из ПО отечественной разработки наибольшую известность в России получили следующие программные системы

и их поставщики: БОСС-Референт (АйТи); Кодекс: Документооборот (Консорциум «Кодекс»); Гран-док (Гранит), Евфрат (Cognitive Technologies); Дело (ЭОС); LanDocs (Ланит); Крон (Анней); OfficeMedia (InterTrust); Effect Office (Гарант Интернэйшнл); N.System (Центр Компьютерных Технологий), LS Flow (Люция-Софт), Оптима (Optima Workflow), ЭСКАДО (ИнтерпрокомЛан), 1С:Документооборот и 1С:Архив (1С), Циркуляр и VisualDOC (ЦентрИнвест Софт), Документ-2000 (TelcomService), Ирида (IBS), RS-Document (R-Style Software Lab) и ряд других.

Следует отметить, что некоторые отечественные СЭД созданы в среде Lotus Domino/Notes (по различным объективным и субъективным причинам получившей достаточно широкое распространение в России): БОСС-Референт (АйТи), семейство продуктов «Золушка» и DIS-Assistant (Институт развития Москвы), CompanyMedia и OfficemMedia (ИнтерТраст), N.System (Центр Компьютерных Технологий), Делопроизводство (КСК) и др.

По классификации IDC большинство отечественных СЭД относится к классу систем, ориентированных на бизнес-процессы (нередко с элементами управления потоками работ). В различных источниках отмечено, что в большинстве отечественных СЭД реализованы следующие функции:

- обработка (хранение) документов;
- управление потоками работ (передача документов между исполнителями);
- контроль исполнения документов;
- поиск документов по атрибутам и полнотекстовый поиск;
- работа со взаимосвязанными документами;
- регламентация прав доступа;
- списание документов;
- интеграция с внешними системами электронной почты и др.

Основным достоинством отечественных СЭД является определенный учет российской специфики и традиций работы с документами (заложенный изначально в их бизнес-логику).

Техническую составляющую такой технологии рассмотрим на примере СЭД «Канцлер». «Канцлер» — система электронного документооборота (СЭД) на платформе IBM Lotus Domino/Notes, разработкой и внедрением которой занимается компания ИВА начиная с 2005 г.

В настоящее время ИВА предлагает семейство программных продуктов «Канцлер»:

- пакет прикладных программ (ППП) «Канцлер» предназначен для создания систем электронного документооборота: автоматизации документооборота и архивной обработки документов предприя-

тий разного профиля и формы собственности, в том числе с территориально-распределенной структурой и холдингового типа, а также органов государственного управления;

- пакет прикладных программ (ППП) «Канцлер» под ОС Linux предназначен для создания системы электронного документооборота предприятий различного типа. ППП «Канцлер» — серверная и клиентская части, функционал основного пакета модулей, ранее реализованные под управлением ОС Windows, теперь доступны и в версии «Канцлер» под управлением ОС Linux;
- программный продукт «iBoss Канцлер» предназначен для обеспечения условий автономной работы руководителей с документами СЭД «Канцлер» с последующей их синхронизацией в системе. Для работы могут быть использованы мобильные устройства Apple iPad, Apple iPad2 и Apple iPad3;
- система электронного документооборота «Облачный Канцлер» предназначена для предоставления услуг по автоматизации документооборота и архивному хранению документов, при этом СЭД развертывается на арендуемой инфраструктуре и ресурсах;
- система электронного документооборота «Канцлер Экспресс» — программный продукт, который совмещает в себе полнофункциональные версии IBM Lotus Domino/Notes и ППП «Канцлер»;
- программный комплекс «Канцлер Мини» предназначен для автоматизации документооборота и организации архивного делопроизводства небольших предприятий.

Система, построенная на ППП «Канцлер», является системой электронного документооборота, которая способна комплексно решить задачи по автоматизации делопроизводства, архивной обработке документов, а также по созданию, движению, маршрутизации документов в реальном масштабе времени, их централизованному хранению. Система на базе ППП «Канцлер» реализуется по модульному принципу и может включать в себя различные приложения-модули, связанные между собой «Базовой системой». Сегодня ИВА предлагает 16 готовых приложений, которые могут быть подключены к «Базовой системе» в зависимости от потребностей пользователя:

- «Мобильный Канцлер»;
- «Делопроизводство»;
- «Управление персоналом»;
- «Договоры»;
- «Архивное дело»;
- «Обращения граждан»;
- «Гербовые бланки»;

- «Потоковое сканирование»;
- «Электронная цифровая подпись»;
- «Система менеджмента качества»;
- «Задания»;
- «Нормативные акты»;
- «Командировки»;
- «Опубликование документов»;
- «Доска объявлений»;
- Система создания сайтов и управления контентом Content Management System.

ППП «Канцлер» объединяет ряд технологических решений: модульная архитектура ППП; централизованное администрирование при помощи «Базовой системы»; подключение ЭЦП любого криптопровайдера; создание распределенной схемы на основе XML-формата. Он поставляется с учетом индивидуальных потребностей заказчика. Пакет поставки может включать в себя любое количество приложений для автоматизации ключевых участков документооборота предприятия (организации). При расширении функционала системы докупаются необходимые приложения. Возможна адаптация и настройка под структуру и особенности организации, а также создание территориально-распределенных систем для организаций со сложной и территориально-распределенной структурой, для организаций холдингового типа. ППП «Канцлер» создает маршруты движения документов; заранее определенные маршруты, созданные в графическом виде, осуществляют управление жизненным циклом документов и обеспечивают удобство работы пользователя: быстрый доступ к документам как с помощью модуля «Рабочее место», с рабочего стола Lotus Notes, из почтового ящика по ссылке из письма-уведомления, так и через формирование графических маршрутов и интуитивно понятный дизайн приложений. Структура системы представлена на рис. 2.7.

Современная СЭД должна обеспечивать выполнение таких основных функций, как:

- регистрация входящей, исходящей и внутренней документации;
- направление документов исполнителям вместе с поручениями по ним и контроль за ходом исполнения поручений;
- списание документов в дело;
- передача дел в архив предприятия и ведение архива;
- рассылка исходящей корреспонденции;
- подготовка проектов документов и поручений и проведение процедур их согласования и утверждения;
- поиск по любому сочетанию параметров и выдача разнообразных сводок и отчетов.

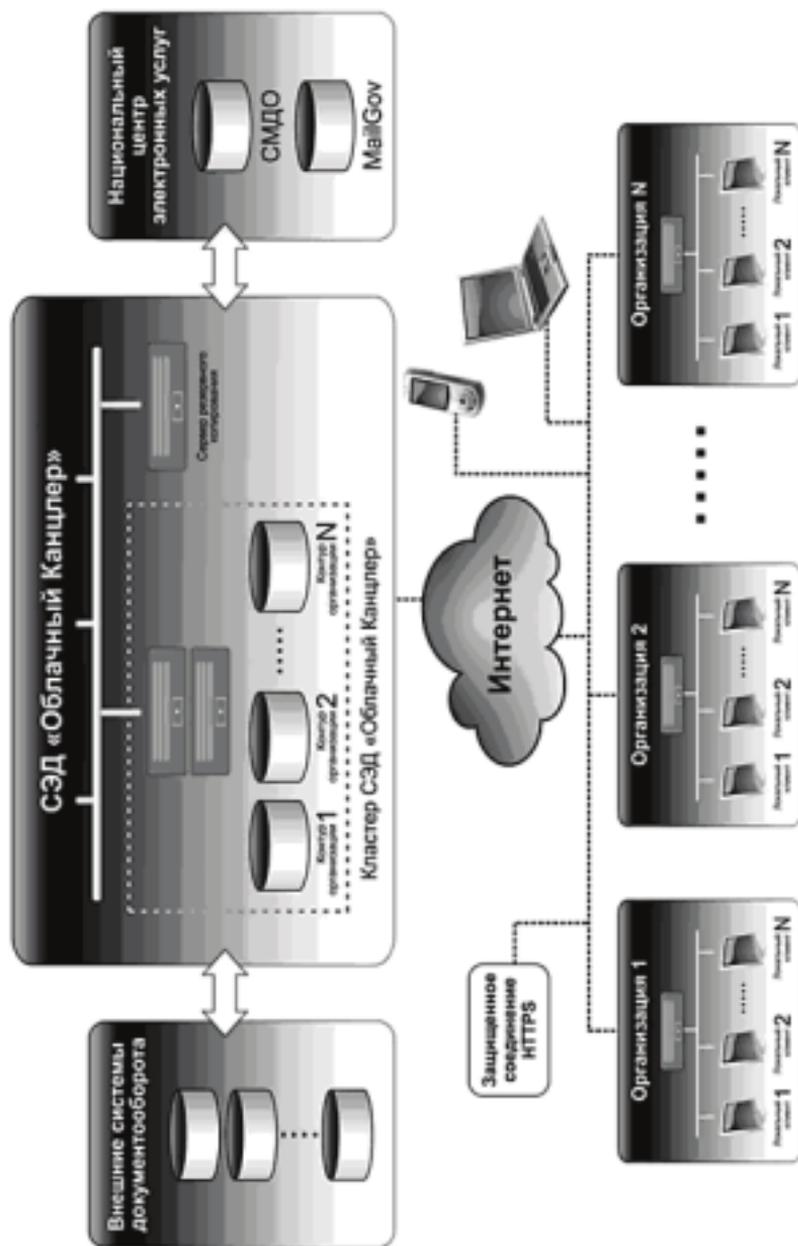


Рис. 2.7. Структура системы «Канцлер»

Состав функций и их реализация в каждой конкретной СЭД могут существенно различаться, и при выборе продукта следует детально и тщательно рассматривать возможности каждого из них.

Поскольку требования к свойствам системы при решении сходных задач в тех или иных подразделениях одной и той же организации могут значительно различаться, важным качеством СЭД является ее функциональная полнота. Структура примера системы представлена на рис. 2.8. При анализе функциональности надо отделить первостепенные потребности будущих пользователей системы от непринципиальных и порожденных в основном боязнью изменений и другими психологическими причинами. Практика показывает, что после периода привыкания к работе в новых условиях пользователи обычно успешно адаптируются и далее уже рассматривают систему как необходимый инструмент своей повседневной деятельности.

Эксплуатационные характеристики СЭД являются одним из важнейших критериев ее выбора, и, оценивая их, целесообразно обратить внимание на следующие качества системы:

- простоту ввода в действие;
- масштабируемость;
- разграничение доступа пользователей и обеспечение защиты от несанкционированного обращения к документам и функциям системы;
- важность протоколирования работы;
- адаптируемость;
- наличие механизмов администрирования и обеспечения надежности.

Ввод системы в действие. При выборе системы надо учесть статистические данные по срокам ее ввода в эксплуатацию у других потребителей и уточнить состав необходимых мероприятий по ее освоению. Одним из таких мероприятий является обучение пользователей, которое, как правило, совмещается ими с выполнением текущих служебных обязанностей. Поэтому крайне желательно, чтобы СЭД помимо работы с оперативной базой документов организации допускала бы и работу с их учебной базой, используемой для тренировки всех заинтересованных сотрудников.

Разумеется, современная СЭД должна иметь дружелюбный пользовательский интерфейс, максимально приближенный к интерфейсам наиболее популярных средств создания и редактирования документов (в частности, Microsoft Office). Следует обратить внимание на правила номерообразования различных групп документов при их регистрации. Поскольку в период ввода системы в действие отдельные участки об-

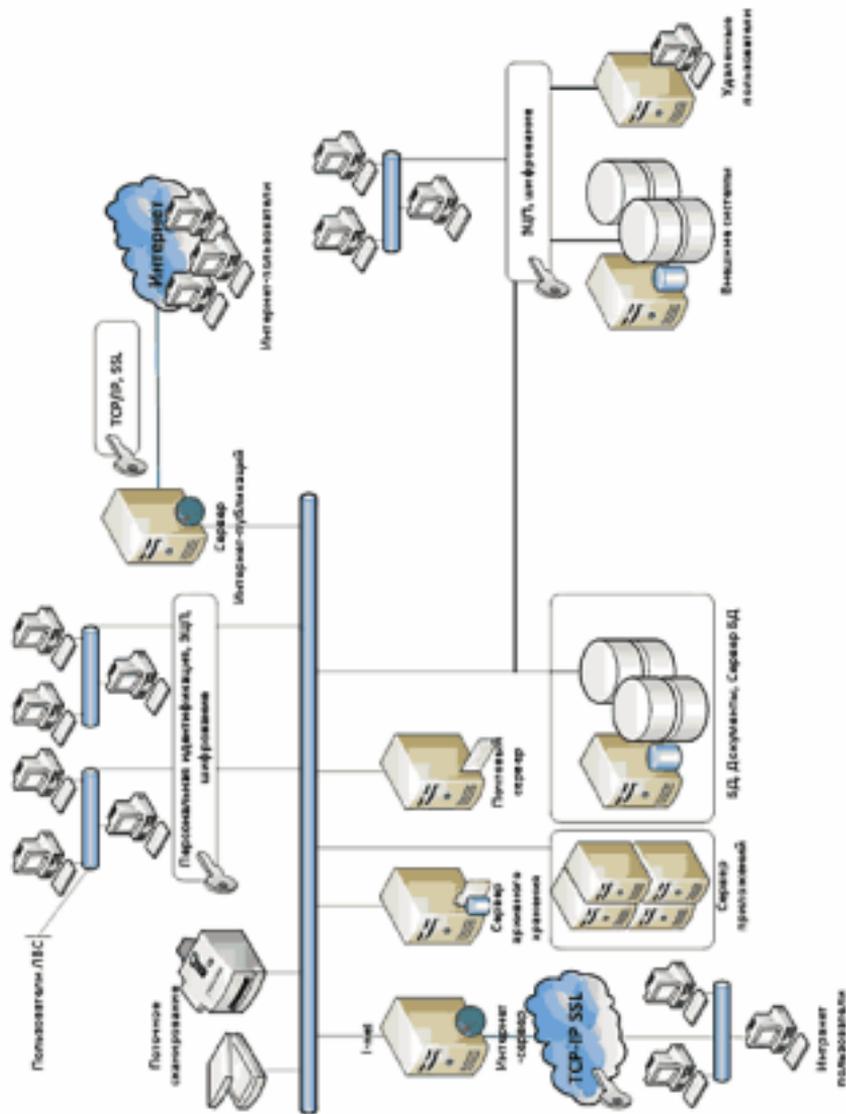


Рис. 2.8. Общая техническая структура СЗД

работки документов могут быть не охвачены автоматизацией, желательно, чтобы система, хотя бы на первых порах, сохраняла наряду с новым и прежний порядок нумерации документов. Так как во многих случаях внедряемая СЭД заменяет имеющуюся систему учета документов, актуальной задачей становится и импорт в нее ретроспективных данных.

**Масштабируемость.** Со временем количество рабочих мест в системе постепенно увеличивается. Обычно они устанавливаются сначала в местах регистрации документов, а затем поэтапно автоматизируются различные подразделения. Поэтому выбираемая СЭД должна обеспечивать возможность наращивания в широких пределах количества рабочих мест.

**Разграничение доступа пользователей и обеспечение защиты от несанкционированного обращения к документам и функциям системы.** СЭД должна обеспечивать разграничение доступа работников отдельных подразделений как непосредственно к текстам документов или их карточкам, так и ко всем этапам документооборота, включая наложение резолюций руководителями и подготовку отчетов исполнителями. Помимо аутентификации пользователей требуется поддержка нескольких уровней доступа (разрешающих, например, просмотр, редактирование, создание, удаление, печать и прочих), определяемых в отношении каждого документа (и даже полей в регистрационной карточке), а также ролевых групп с различными полномочиями (допустим, группы начальников отделов).

Если СЭД является территориально распределенной системой, желательно, чтобы для обмена документами по открытым каналам связи она имела встроенные и сертифицированные ФАПСИ средства цифровой подписи и шифрования.

**Протоколирование работы пользователей.** Протоколирование позволяет отследить всю историю документа (кто и когда его создал, редактировал, просматривал, печатал и т.д.) и дает возможность разрешать конфликтные ситуации при совместной работе над ним нескольких пользователей.

**Адаптируемость.** В силу разных причин на предприятии то и дело изменяются номенклатура дел, структура и состав подразделений. Система должна быть способна поддерживать эти изменения, сохраняя в неприкосновенности информацию о документах, обработанных ранее.

**Администрирование и обеспечение надежности.** Средства администрирования СЭД, позволяющие назначать и модифицировать ролевые полномочия пользователей по отношению к документам и функциям системы, должны уметь настраивать ее в соответствии с постоянно меняющейся организационной структурой. Желательно, чтобы действие

таких настроек распространялось не только на отдельные документы, но и на их группы, имеющие иногда сложную иерархическую структуру. Для аналитической работы с документами не обойтись без инструментов, позволяющих относить документы к различным темам и классифицировать по тем или иным признакам.

Важнейшими характеристиками устойчивости СЭД к неизбежным сбоям и отказам являются: независимость функционирования рабочих мест друг от друга, время восстановления сеанса после его аварийного завершения, наличие средств обеспечения целостности данных и их резервного копирования.

### **Контрольные вопросы**

1. Как моделируется информация для обеспечения процесса управления?
2. Какие этапы формирования информационного обеспечения менеджмента следует выделить?
3. Как следует классифицировать информацию по выполняемым функциям в процессе управления?
4. Какие основные черты вы могли бы выделить в модели информации в управлении экономическими объектами?
5. Зачем внедрять ИТ в управленческую деятельность?
6. Через что, по вашему мнению, следует вводить средства для обработки информации в управленческом процессе?
7. Каковы возможности системы управления по переработке этого потока информации?
8. Что такое «первый информационный барьер»?
9. Что такое «второй информационный барьер»?
10. При помощи чего следует преодолевать «второй информационный барьер»?

# ИНФОРМАЦИОННЫЙ ПРОЦЕСС ОБМЕНА ДАНЫМИ. СЕТЕВЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

## 3.1. Общие понятия о компьютерных сетях

Основным инструментом, используемым как база для построения современных ИТ, используемых в менеджменте, являются компьютерные сети. Компьютерные сети являются результатом компьютерных и телекоммуникационных технологий. С одной стороны, сети передачи данных представляют собой частный случай распределенных вычислительных систем, в которых группа компьютеров согласованно выполняет набор взаимосвязанных задач, обмениваясь данными в автоматическом режиме; с другой — компьютерные сети могут рассматриваться как средство передачи информации на большие расстояния, для чего в них применяются методы кодирования и мультиплексирования данных, получившие развитие в различных телекоммуникационных системах и технологиях передачи данных. Поэтому следует четко понимать:

- компьютерная сеть — это набор компьютеров, связанных коммуникационной системой и снабженных соответствующим программным обеспечением, которое предоставляет пользователям сети доступ к ресурсам этого набора компьютеров;
- компьютерную сеть образуют компьютеры разных типов — небольшие ЭВМ на базе микропроцессоров, рабочие станции,

мини-компьютеры, персональные компьютеры или суперкомпьютеры;

- передачу сообщений между любой парой компьютеров сети обеспечивает среда распространения информации, которая представляет собой воздух (беспроводные сети), металл (кабельные сети) или оптическое волокно (волоконно-оптические сети);
- компьютерная сеть позволяет пользователю работать со своим компьютером как с автономным и добавляет к этому возможность доступа к информационным и аппаратным ресурсам других компьютеров сети, а также сетевым службам, позволяющим работать двум и более пользователям одновременно.

Считается, что рождение именно цифровых вычислительных машин произошло вскоре после окончания Второй мировой войны. В 40-х годах XX в. были созданы первые ламповые вычислительные устройства. Для этого периода характерно следующее:

- компьютер представлял собой скорее предмет исследования, а не инструмент для решения каких-либо практических задач из других областей;
- программирование прикладных задач осуществлялось исключительно на машинном языке — кодах центрального процессора;
- не было никакого системного программного обеспечения, кроме библиотек математических и служебных подпрограмм;
- операционных систем не было в помине, все задачи организации вычислительного процесса решались вручную каждым программистом с пульта управления.

С середины 50-х гг. XX в. начался следующий период в развитии вычислительной техники, связанный с появлением новой технической базы — полупроводниковых элементов. В то время развитие ЭВМ определялось направлениями:

- в связи с внедрением новой элементной базы увеличилось быстродействие процессоров и, как следствие, выросла производительная мощность вычислительных машин, увеличились объемы оперативной и внешней памяти;
- появились первые алгоритмические языки, и, таким образом, выросла сложность программ, и соответственно это позволило решать более сложные задачи, а сами программы стали выглядеть как наборы кодов, написанных в виде текстовых конструкций и хранящихся в виде текстового файла;
- к библиотекам математических и служебных подпрограмм добавился новый тип системного программного обеспечения —

трансляторы, перевод исходного кода программ в коды центрального процессора;

- разработаны первые системные управляющие программы — мониторы, которые автоматизировали всю последовательность действий оператора по организации вычислительного процесса.

Программные мониторы явились прообразом современных операционных систем, они стали первыми системными программами, предназначенными не для обработки данных, а для управления вычислительным процессом. В ходе реализации мониторов был разработан формализованный язык управления заданиями, с помощью которого программист сообщал системе и оператору, какие действия и в какой последовательности он хотел бы выполнить на вычислительной машине. Типовой набор директив обычно включал признак начала отдельной работы, вызов транслятора, вызов загрузчика, признаки начала и конца исходных данных.

Оператор составлял пакет заданий, которые в дальнейшем без его участия последовательно запускались на выполнение монитором. Кроме того монитор был способен самостоятельно обрабатывать наиболее распространенные аварийные ситуации, возникающие при работе пользовательских программ, такие как отсутствие исходных данных, переполнение регистров, деление на ноль, обращение к несуществующей области памяти и т.д.

Следующий важный период развития операционных систем относится к 1965—1975 гг. XX в. В это время в технической базе вычислительных машин произошел переход от отдельных полупроводниковых элементов типа транзисторов к интегральным микросхемам, что открыло путь к появлению следующего поколения компьютеров, представителем которого является, например, IBM/360<sup>1</sup>.

В этот период были реализованы практически все основные механизмы, присущие современным операционным системам<sup>2</sup> (ОС):

<sup>1</sup> IBM System/360 (S/360) — семейство компьютеров класса мэйнфреймов, которое было анонсировано 7 апреля 1964 г. Это был первый ряд компьютеров, в котором проводилось четкое различие между архитектурой и реализацией. В отличие от предыдущих серий IBM создала линейку компьютеров от малых к большим, от низкой к высокой производительности, все модели которой использовали один и тот же набор команд (с двумя исключениями из правила — для специфичных рынков). Эта особенность позволяла заказчику использовать недорогую модель, после чего обновиться до более крупной системы с ростом компании — без необходимости переписывать программное обеспечение. Для обеспечения совместности IBM впервые применила технологию микрокода, который применялся во всех моделях серии, кроме самых старших.

<sup>2</sup> Операционная система (ОС) (от англ. operating system, OS) — комплекс управляющих и обрабатывающих программ, которые, с одной стороны, выступают как интерфейс между устройствами вычислительной системы и прикладными программами, а с другой стороны, предназначены для управления устройствами, управления вычислительными процессами, эффективного распределения вычислительных ресурсов между вычислительными процессами и организации надежных вычислений. Это определение применимо к большинству современных операционных систем общего назначения. В логической структуре типичной вычислительной системы операционная система занимает положение между устройствами с их микроархитектурой, машинным языком и, возможно, собственными (встроенными) микропрограммами с одной стороны и прикладными программами с другой.

- **мультипрограммирование** (способ организации выполнения нескольких программ на одном компьютере, обычно разделяют мультипрограммирование в пакетных системах, системах реального времени и мультипрограммирование в системах разделения времени);
- **мультипроцессирование** (синонимами еще являются многопроцессорность, мультипроцессорность, многопроцессорная обработка, образовано от английского *Multiprocessing* — использование пары или большего количества физических процессоров в одной компьютерной системе. Термин также относится к способности системы поддерживать больше чем один процессор и (или) способность распределить задачи между ними. Существует много вариантов данного понятия, и определение многопроцессорности может меняться в зависимости от контекста, главным образом в зависимости от того, как определены процессоры (много ядер в одном кристалле, множество чипов в одном корпусе, множество корпусов в одном системном модуле и т.д.). Многопроцессорностью иногда называют выполнение множественных параллельных программных процессов в системе в противоположность выполнению одного процесса в любой момент времени. Однако термины «многозадачность» или «мультипрограммирование» являются более подходящими для описания этого понятия, которое осуществлено главным образом в программном обеспечении, тогда как многопроцессорная обработка является более соответствующей, чтобы описать использование множественных аппаратных процессоров. Система не может быть и многопроцессорной и мультипрограммированной, только одной из двух, или ни той и ни другой);
- **поддержка многотерминального и (или) многопользовательского режима** (т.е. возможность работать одновременно за числом терминалов ЭВМ большим двум пользователям с различными правами доступа, учетными записями и т.д.);
- **виртуальная память** (виртуальная память происходит от английского *Virtual memory* — технология управления памятью ЭВМ, разработанная для многозадачных операционных систем. При использовании данной технологии для каждой программы используются независимые схемы адресации памяти, отображающиеся тем или иным способом на физические адреса в памяти ЭВМ. Технология позволяет увеличить эффективность использования памяти несколькими одновременно работающими программами, организовав множество независимых адресных про-

странств, и обеспечить защиту памяти между различными приложениями. Также позволяет программисту использовать больше памяти, чем установлено в компьютере, за счет отначки неиспользуемых страниц на вторичное хранилище);

- **файловые системы** (файловая система от английского *file system* — порядок, определяющий способ организации, хранения и именования данных на носителях информации в компьютерах, а также в другом электронном оборудовании: цифровых фотоаппаратах, мобильных телефонах и т.п. Файловая система определяет формат содержимого и способ физического хранения информации, которую принято группировать в виде файлов. Конкретная файловая система определяет размер имени файла (папки), максимальный возможный размер файла и раздела, набор атрибутов файла. Некоторые файловые системы предоставляют сервисные возможности, например, разграничение доступа или шифрование файлов.);
- **разграничение доступа** (совокупность правил, регламентирующих порядок и условия доступа субъекта к объектам информационной системы (информации, ее носителям, процессам и другим ресурсам), установленным правовыми документами или собственником, владельцем информации. Права доступа определяют набор действий (например, чтение, запись, выполнение), разрешенных для выполнения субъектам (например, пользователям системы) над объектами данных. Для этого требуется некая система Рис. 3.1. для предоставления субъектам различных прав доступа к объектам. Это система разграничения доступа субъектов к объектам, которая рассматривается в качестве главного средства защиты от несанкционированного доступа к информации).

В эти годы начинается расцвет системного программирования<sup>1</sup>. Из направления прикладной математики, представляющего интерес для узкого круга специалистов, системное программирование превращается в отрасль индустрии, оказывающую непосредственное влияние на практическую деятельность миллионов людей.

В условиях резко возросших возможностей компьютера, связанных с обработкой и хранением данных, выполнение только одной программы в каждый момент времени оказалось крайне неэффективным. На-

<sup>1</sup> Системное программирование (или программирование систем) — подраздел программирования, заключающийся в работе над системным программным обеспечением. Определение «системное» подчеркивает тот факт, что результаты этого вида программирования существенно меняют свойства и возможности вычислительной системы. В определенной степени этот результат имеет место при применении любых программ, выполняемых в вычислительной системе.



Рис. 3.1. Модель пакетной обработки заданий ЭВМ

чались разработки в области мультипрограммирования. Как уже было ранее отмечено, мультипрограммирование — способ организации вычислительного процесса, при котором в памяти компьютера находится одновременно несколько программ, попеременно выполняющихся на одном процессоре. Мультипрограммирование обычно реализовывалось в двух вариантах:

- пакетная обработка;
- разделение времени.

Системы пакетной обработки предназначались для решения задач в основном вычислительного характера, не требующих быстрого получения результатов. Главной целью и критерием эффективности систем пакетной обработки является максимальная пропускная способность, т.е. решение максимального числа задач в единицу времени.

Для достижения этой цели в системах пакетной обработки используется следующая схема функционирования (рис. 3.2): в начале работы формируется пакет заданий, каждое задание содержит требование к системным ресурсам; из этого пакета заданий формируется мультипрограммный набор, т.е. множество одновременно выполняемых задач. Для одновременного выполнения выбираются задачи, предъявляющие к ресурсам различные требования, так, чтобы обеспечивалась сбалансированная загрузка всех устройств вычислительной машины. Например, в мультипрограммном наборе желательно присутствие и вычислитель-

ных задач, и задач с интенсивным вводом-выводом. Таким образом, выбор нового задания из пакета заданий зависит от внутренней ситуации, складывающейся в системе, т.е. выбирается «выгодное» задание. Следовательно, в вычислительных системах, работающих под управлением пакетных ОС, невозможно гарантировать выполнение того или иного задания в течение определенного периода времени.

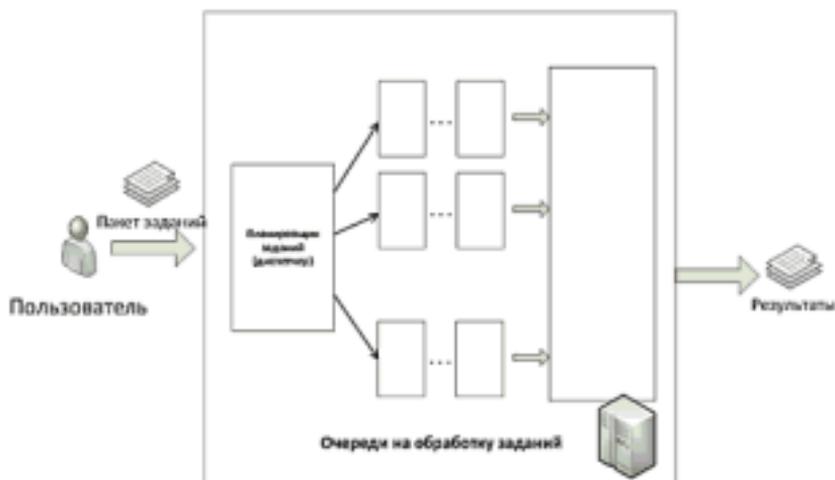


Рис. 3.2. Механизм пакетной обработки заданий ЭВМ

В системах пакетной обработки переключение процессора с одной задачи на другую происходит по инициативе самой активной задачи, например, когда она «отказывается» от процессора из-за необходимости выполнить операцию ввода-вывода. Поэтому существует высокая вероятность того, что одна задача может надолго занять процессор, и выполнение интерактивных задач станет невозможным. Взаимодействие пользователя с вычислительной машиной, на которой установлена система пакетной обработки, сводится к тому, что пользователь приносит задание, отдает его диспетчеру-оператору, а в конце дня после выполнения всего пакета заданий получает результат. Очевидно, что такой порядок повышает эффективность функционирования аппаратуры, но снижает эффективность работы пользователя.

В системах разделения времени пользователям (или одному пользователю) предоставляется возможность интерактивной работы сразу с несколькими приложениями. Для этого каждое приложение должно регу-

лярно взаимодействовать с пользователем. Понятно, что в пакетных системах возможности диалога пользователя с приложением ограничены.

В системах разделения времени эта проблема решается за счет того, что ОС принудительно периодически приостанавливает приложения, не дожидаясь, когда они сами освободят процессор. Всем приложениям попеременно выделяется квант процессорного времени  $q$ . Таким образом, пользователи, запустившие программы на выполнение, получают возможность поддерживать с ними диалог.

Системы разделения времени Рис. 3.3. призваны исправить основной недостаток систем пакетной обработки — изоляцию пользователя-программиста от процесса выполнения задач. Каждому пользователю в этом случае предоставляется терминал, с которого он может вести диалог со своей программой. Так как в системах разделения времени каждой задаче выделяется только квант процессорного времени, ни одна задача не занимает процессор надолго, и время ответа оказывается приемлемым. Если квант небольшой, то у всех пользователей, одновременно работающих на одной и той же машине, складывается впечатление, что каждый из них использует машину единолично.

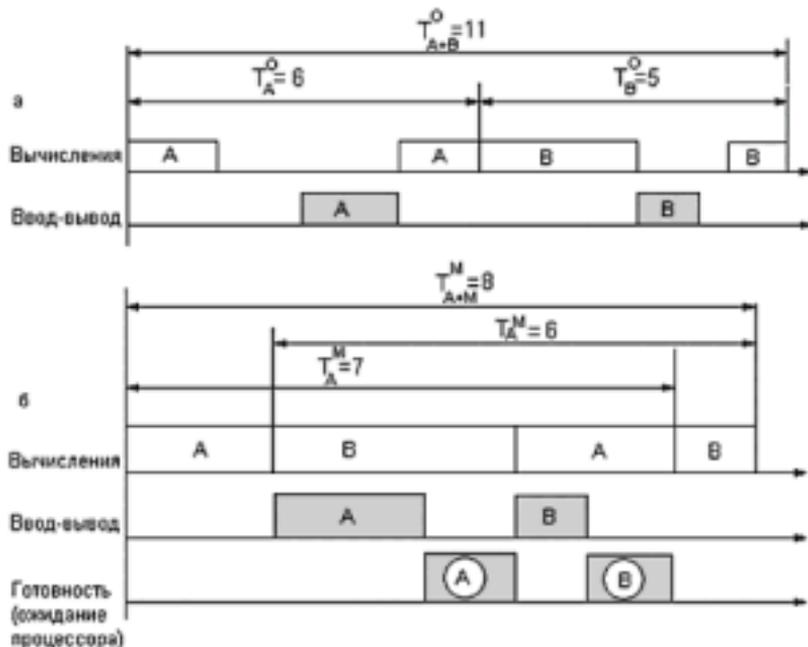


Рис. 3.3. Принцип работы системы с разделением времени

Все эти системы не зависимо от выбранного способа выполнения заданий взаимодействовали с пользователем либо через пульт ввода-вывода информации, либо через специализированное устройство ввода-вывода информации, называемой терминал<sup>1</sup>. Используемый в системах разделения времени (возможен еще английский термин *time sharing*), благодаря которым несколько пользователей могли одновременно работать с одной системой, каждый со своего терминала, подобное устройство стало основным средством взаимодействия с ОС и ЭВМ. Изначально в качестве терминалов применялись электромеханические телетайпы, уже использовавшиеся в телеграфии. Позже были разработаны специальные печатающие терминалы, такие как, например, DECwriter. Но скорость взаимодействия с машиной была ограничена невысокой скоростью печати, и печатная копия для процесса непосредственного общения не была необходимой. Позже для отображения информации также использовались дисплеи на основе знакопечатающих трубок. В таких трубках отображение символа на экране осуществлялась путем пропускания электронного луча через трафарет. Основным интерфейсом взаимодействия с пользователем использовался режим командной строки, где для управления всеми процессами использовались команды в виде текстовых сообщений вводимых пользователем с клавиатуры или выводимые на дисплей от системы. Поэтому такие терминалы называли текстовыми. Текстовый терминал (или чаще просто терминал, иногда также текстовая консоль) — это интерфейс компьютера для последовательной передачи данных — ввода и изображения текста. Информация представляется в виде массива предопределенных знаков. Первоначально текстовые терминалы были электронными устройствами, подключаемыми к последовательному порту, но позже в компьютеры стали встраивать текстовую консоль. В графическом интерфейсе пользователя используются программы — эмуляторы терминала.

Поскольку принцип разделения времени дал возможность работать одновременно нескольким пользователям с одной ОС и комплектом оборудования, называемый ЭВМ и периферия, то размещение пользователей в различных местах здания или зданий обеспечило необходимость

<sup>1</sup> Терминал (от англ. *terminal* — конечный) — конечная часть некой системы, которая обеспечивает связь системы с внешней средой. Компьютерный терминал — электронное или электромеханическое устройство, используемое для взаимодействия пользователя с компьютером или компьютерной системой. С появлением видео-дисплеев терминалы вытеснили другие способы ввода-вывода. Поэтому терминал — это устройство ввода-вывода, его основные функции заключаются в отображении и вводе данных. Устройство со значительным объемом обработки данных называют *smart terminal*, или «толстый клиент» (от англ. *fat client*). Терминал, сильно зависящий от своей *host-машины*, на которой выполняются основные вычисления, называют тонким клиентом (от англ. *thin client*).

решения проблемы в обеспечении таких пользователей средствами работы. Терминалы, выйдя за пределы вычислительного центра, рассредоточились по всему предприятию.

Многотерминальный режим использовался не только в системах разделения времени, но и в системах пакетной обработки. При этом не только оператор, но и все пользователи получали возможность формировать свои задания и управлять их выполнением со своего терминала. Такие операционные системы получили название систем удаленного ввода заданий.

Терминальные комплексы могли располагаться на большом расстоянии от процессорных стоек, соединяясь с ними с помощью различных глобальных связей — модемных соединений телефонных сетей или выделенных каналов. Для поддержки удаленной работы терминалов в операционных системах появились специальные программные модули, реализующие различные (в то время, как правило, не стандартные) протоколы связи. Такие вычислительные системы с удаленными терминалами, сохраняя централизованный характер обработки данных, в какой-то степени являлись прообразом современных компьютерных сетей (рис. 3.4), а соответствующее системное программное обеспечение — прообразом сетевых операционных систем. Многотерминальные централизованные системы уже имели все внешние признаки локальных вычислительных сетей, однако по существу ими не являлись, так как сохраняли сущность централизованной обработки данных автономно работающего компьютера.

Действительно, рядовой пользователь работу за терминалом мейнфрейма<sup>1</sup> воспринимал примерно так же, как сейчас воспринимает работу за подключенным к сети персональным компьютером. Пользователь мог получить доступ к общим файлам и периферийным устройствам, при этом у него создавалась полная иллюзия единоличного владения компьютером, так как он мог запустить нужную ему программу в любой момент и почти сразу же получить результат (некоторые далекие от вычислительной техники пользователи даже были уверены, что все вычисления выполняются внутри их дисплея).

<sup>1</sup> Мейнфрейм (также мэйнфрейм, оба произношения и варианта написания от англ. mainframe) — данный термин имеет три основных значения.

1. Большая универсальная ЭВМ — высокопроизводительный компьютер со значительным объемом оперативной и внешней памяти, предназначенный для организации централизованных хранилищ данных большой емкости и выполнения интенсивных вычислительных работ.

2. Компьютер с архитектурой IBM System/360, 370, 390, zSeries.

3. Наиболее мощный компьютер (например, удовлетворяющий признакам значения 1), используемый в качестве главного или центрального компьютера (например, в качестве главного сервера).

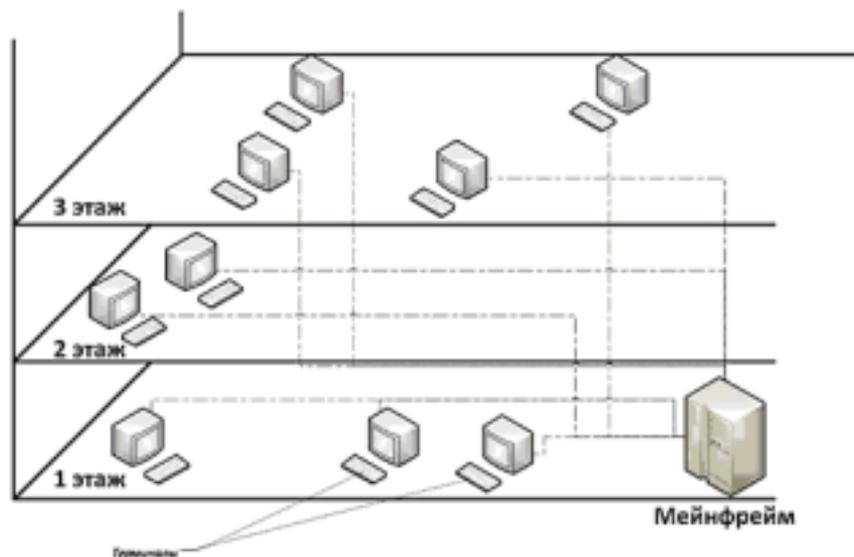


Рис. 3.4. Схема построения многотерминального вычислительного центра

Подобные идеи, такие как «... много терминалов разом» или «... А почему так долго считаем? Если объединить несколько ЭВМ в одну, будет же быстрее...» возникали одновременно с появлением новых и, главное, более производительных машин. До реализации связей «компьютер — компьютер» была решена более простая задача — организация связи «удаленный терминал — компьютер». Терминалы, находящиеся от компьютера на расстоянии многих сотен, а то и тысяч километров, соединялись с компьютерами через телефонные сети с помощью модемов. Такие сети позволяли многочисленным пользователям получать удаленный доступ к разделяемым ресурсам нескольких мощных компьютеров класса суперЭВМ. И только потом были разработаны средства обмена данными между компьютерами в автоматическом режиме. На основе этого механизма в первых сетях были реализованы службы обмена файлами, синхронизации баз данных, электронной почты и другие, ставшие теперь традиционными, сетевые службы.

В 1969 году Министерство обороны США инициировало работы по объединению в общую сеть суперкомпьютеров оборонных и научно-исследовательских центров. Эта сеть, получившая название ARPAnet<sup>1</sup>, по-

<sup>1</sup> ARPAnet (от англ. Advanced Research Projects Agency Network) — компьютерная сеть, созданная в 1969 г. в США Агентством Министерства обороны США по перспективным исследованиям и являвш-

служила отправной точкой для создания первой и самой известной ныне глобальной сети — Internet (Интернет). Сеть ARPAnet объединяла компьютеры разных типов, работавшие под управлением различных ОС с дополнительными модулями, реализующими коммуникационные протоколы, общие для всех компьютеров сети. Такие ОС можно считать первыми сетевыми операционными системами.

Сетевые ОС в отличие от многотерминальных позволяли не только рассредоточить пользователей, но и организовать распределенное хранение и обработку данных между несколькими компьютерами, связанными электрическими связями. Любая сетевая операционная система, с одной стороны, выполняет все функции локальной операционной системы, а с другой — обладает некоторыми дополнительными средствами, позволяющими ей взаимодействовать по сети с операционными системами других компьютеров. Программные модули, реализующие сетевые функции, появлялись в операционных системах постепенно, по мере развития сетевых технологий, аппаратной базы компьютеров и возникновения новых задач, требующих сетевой обработки. В 1974 году компания IBM объявила о создании собственной сетевой архитектуры для своих мэйнфреймов, получившей название SNA (System Network Architecture, системная сетевая архитектура). В это же время в Европе активно велись работы по созданию и стандартизации сетей X.25<sup>1</sup>.

---

ял основным полигоном для отработки и проведения экспериментов по тематике телекоммуникации и фактически прототипом сети «Интернет». 1 января 1983 г. она стала первой в мире сетью, прошедшей на маршрутизацию пакетов данных. В качестве маршрутизируемого протокола использовался TCP/IP, который и по сей день является основным протоколом передачи данных в сети «Интернет». ARPANET исключительно как сеть прекратила свое существование в июне 1990 г. Этот «полигон» фактически входит в состав созданного с 1958 г. в США агентства DARPA, DARPA (от англ. Defense Advanced Research Projects Agency — агентство передовых оборонных исследовательских проектов) — агентства Министерства обороны США, отвечающее за разработку новых технологий для использования в вооруженных силах. Миссией DARPA является сохранение технологического превосходства вооруженных сил США, предотвращение внезапного для США появления новых технических средств вооруженной борьбы, поддержка прорывных исследований, преодоление разрыва между фундаментальными исследованиями и их применением в военной сфере.

<sup>1</sup> X.25 — семейство протоколов канального уровня сетевой модели OSI. Предназначалось для организации WAN на основе телефонных сетей с линиями с достаточно высокой частотой ошибок, поэтому содержит развитые механизмы коррекции ошибок. Ориентирован на работу с установлением соединений. Исторически является предшественником протокола Frame Relay. X.25 обеспечивает множество независимых виртуальных каналов (Permanent Virtual Circuits, PVC и Switched Virtual Circuits, SVC) в одной линии связи, идентифицируемых в X.25-сети по идентификатором подключений к соединению идентификаторы логического канала (Logical Channel Identifier, LCI) или номера логического канала (Logical Channel Number, LCN). Благодаря надежности протокола и его работе поверх телефонных сетей общего пользования X.25 широко использовался как в корпоративных сетях, так и во всемирных специализированных сетях предоставления услуг, таких как SWIFT (банковская платёжная система) и SITA (от фр. Société Internationale de Télécommunications Aéronautiques — система информационного обслуживания воздушного транспорта), однако в настоящее время X.25 вытесняется другими технологиями канального уровня (Frame Relay, ISDN, ATM) и протоколом IP, оставаясь, однако, достаточно распространённым в странах и территориях с неразвитой телекоммуникационной инфраструктурой.

Таким образом, хронологически первыми появились глобальные сети (Wide Area Networks, WAN), то есть сети, объединяющие территориально рассредоточенные компьютеры, возможно, находящиеся в различных городах и странах. Именно при построении глобальных сетей были впервые предложены и отработаны многие основные идеи и концепции современных вычислительных сетей, такие, например, как многоуровневое построение коммуникационных протоколов, технология коммутации пакетов и маршрутизация пакетов в составных сетях.

Глобальные компьютерные сети очень многое унаследовали от других, гораздо более старых и глобальных сетей — телефонных.

## 3.2. Телеобработка данных в компьютерных сетях

Под телеобработкой понимается обработка данных (прием данных от источника, их преобразование вычислительными средствами и выдача результатов потребителю), передаваемых по каналам связи. Различают системную и сетевую телеобработку. Системная телеобработка основана на принципе централизованной обработки данных, когда удаленным пользователям, как правило, не имеющим своих вычислительных ресурсов, обеспечивается доступ к ресурсам одной высокопроизводительной ЭВМ (мэинфрейма) или вычислительной системы по каналам связи. Сетевая телеобработка основана на принципе распределенной обработки данных, реализуемой совокупностью ЭВМ, объединенных в сеть и взаимодействующих между собой с помощью каналов связи и специального сетевого оборудования. В сетях ЭВМ обеспечивается доступ локальных и удаленных пользователей к распределенным в сети информационно-вычислительным ресурсам и базам данных.

Системная телеобработка данных получила наибольшее распространение в 70-х — 80-х гг. XX в., в настоящее время продолжает широко использоваться в различных отраслях экономики, промышленности и в военном деле. Реализация данного вида обработки информации осуществляется на основе использования систем телеобработки данных (СТОД).

Система телеобработки данных представляет собой совокупность технических и программных средств, обеспечивающих одновременный и независимый удаленный доступ большого количества абонентов (пользователей, объектов управления) к централизованным информационно-вычислительным ресурсам. Структурная схема типовой системы телеобработки данных приведена на рис. 3.5.

Технические средства СТОД, как правило, включают в себя:

- универсальную ЭВМ (ПЭВМ, вычислительный комплекс или систему);
- устройства сопряжения (УС) ЭВМ с аппаратурой передачи данных (АПД);
- аппаратуру передачи данных с линиями связи, образующими в совокупности каналы связи (КС);
- абонентские пункты (АП).

ЭВМ является основным элементом и обеспечивает решение задач по обработке данных и управлению всей системой в целом. Устройства сопряжения предназначены для обеспечения физического и логического согласования ЭВМ и аппаратуры передачи данных. В качестве УС применяются линейные адаптеры, мультиплексоры передачи данных, связные процессоры и т.п.



Рис. 3.5. Структура системы телеобработки данных

Абоненты системы (пользователи, технические объекты) подключаются к ЭВМ с помощью каналов связи. Каждый канал связи состоит из

линии связи, по которой передаются сигналы, и аппаратуры передачи данных, преобразующей дискретные данные в сигналы, соответствующие конкретному типу линии связи, и наоборот. Канал связи может обслуживать единственного абонента, образуя двухточечное соединение (линия связи 1 на рис. 3.5), или одновременно нескольких абонентов, образуя многоточечное соединение (линия связи 2 на рис. 3.5). В этом случае абоненты разделяют между собой канал во времени, принимая адресованные им данные и снабжая передаваемые данные адресом (номером) источника.

Для передачи информации могут использоваться некоммутируемые (постоянно выделенные абоненту) и коммутируемые каналы. В последнем случае для подключения удаленных абонентов могут использоваться телефонные линии связи и автоматические телефонные станции (АТС) (линия связи 3 на рис. 3.5).

Абоненты взаимодействуют с ЭВМ через абонентские пункты. Абонентский пункт содержит в своем составе АПД, обслуживающую канал связи, набор ПУ, используемых для ввода-вывода данных, и обеспечивает обмен данными между каналом связи и периферийными устройствами. В качестве ПУ абонентских пунктов наиболее широко используются различные клавиатуры, дисплеи и печатающие устройства.

Функционирование технических средств систем телеобработки поддерживается программными средствами. Программные средства СТОД включают в себя специальные модули операционной системы ЭВМ, прикладные или пользовательские программы, реализующие телекоммуникационные методы доступа к информации и обеспечивающие решение следующих задач:

- управление работой ЭВМ в различных режимах телеобработки;
- прием данных от абонентов и их редактирование;
- управление очередями на прием и передачу данных;
- организация соединений с требуемыми абонентами;
- передача результатов обработки данных абонентам;
- обработка ошибок и восстановление работоспособности системы.

Системы телеобработки данных обеспечивают решение следующих задач:

- дистанционные вычисления, при выполнении которых с АП по каналам связи в ЭВМ вводятся исходные данные, а обратно выдаются результаты их обработки;
- дистанционный информационно-справочный режим, при котором по запросам АП из баз данных ЭВМ им предоставляется необходимая информация;

- дистанционный режим сбора данных, формируемых абонентами системы;
- коллективный доступ к ЭВМ абонентов с удаленных АП.

Таким образом, основной целью создания и основным достоинством систем телеобработки является повышение эффективности обработки данных за счет оперативного приема их непосредственно от источников информации и выдачи результатов обработки к местам их использования. Кроме того телеобработка позволяет эффективно использовать мощные ЭВМ за счет высокого уровня их загрузки и возможности создания на их основе больших баз данных. Наряду с указанными достоинствами СТОД имеют и ряд существенных недостатков:

- неравномерная интенсивность запросов от абонентских пунктов может привести к частичному простоям оборудования ЭВМ или возникновению пиковых нагрузок;
- отказ или сбой функционирования центральной ЭВМ приводит к нарушению работоспособности всей системы телеобработки;
- отказ отдельных каналов связи делает полностью невозможным взаимодействие соответствующих абонентских пунктов с ЭВМ;
- невозможность информационного взаимодействия между различными СТОД.

Недостатки централизованных систем телеобработки данных и технологический прорыв в области производства электронных компонентов привели к возникновению концепции объединения удаленных друг от друга ЭВМ в единую систему. Реализация данной концепции привела к созданию сетей ЭВМ и технологии распределенной обработки данных. Наиболее широкое внедрение сетей ЭВМ во все сферы человеческой деятельности началось в середине 70-х гг. XX в. и продолжается по настоящее время.

Сеть ЭВМ (вычислительная сеть, компьютерная сеть) — это сеть обмена и распределенной обработки информации, образуемая множеством абонентских систем, взаимодействующих между собой посредством телекоммуникационной сети. Абонентская система (АС) — это совокупность аппаратных и программных средств ЭВМ (ПЭВМ, рабочей станции, вычислительного комплекса и т.п.), периферийного оборудования и средств связи с телекоммуникационной сетью, реализующих прикладные процессы. Телекоммуникационная сеть (ТКС) — это совокупность физических линий связи, аппаратных и программных средств, обеспечивающих информационное взаимодействие абонентских систем. Прикладные процессы — это различные процедуры ввода, хранения, обработки и выдачи информации, выполняемые абонентскими системами.



Рис. 3.6. Обобщенная схема сети ЭВМ

ми по запросам и в интересах пользователей вычислительной сети. Обобщенная схема сети ЭВМ представлена на рис. 3.6.

Целесообразность создания сетей ЭВМ обуславливается двумя основными факторами:

- возможностью использования территориально рассредоточенными пользователями программного обеспечения и информационных ресурсов, размещенных в различных абонентских системах сети;
- возможностью организации распределенной обработки данных вычислительными ресурсами нескольких абонентских систем сети для решения особо сложных задач.

В общем случае сети ЭВМ позволяют:

- обеспечить широкий дистанционный доступ пользователей к аппаратным, программным и информационным ресурсам сети;
- повысить уровень загрузки и эффективность использования оборудования абонентских систем сети;
- оперативно перераспределять нагрузку между вычислительными средствами абонентских систем с целью недопущения ее пиковых значений;
- создавать распределенные по сети и централизованные базы данных;
- увеличить надежность обработки и передачи данных за счет избыточности и возможности резервирования отдельных технических компонентов сети.

Сети ЭВМ могут характеризоваться совокупностью показателей качества, к основным из которых относятся следующие:

- функциональные возможности сети — перечень основных информационно-вычислительных услуг, предоставляемых пользователям сети;
- производительность сети — среднее количество запросов пользователей сети, обслуживаемых за единицу времени;
- пропускная способность сети — объем данных, передаваемых по сети или отдельному ее сегменту за единицу времени;
- надежность сети — среднее время наработки на отказ основных компонентов сети;
- информационная безопасность сети — вероятность несанкционированного доступа к обрабатываемой и передаваемой по каналам сети информации;
- масштабируемость сети — возможность расширения сети без заметного снижения ее производительности.

Телекоммуникационная сеть — это совокупность физических линий связи, аппаратных и программных средств, обеспечивающих информационное взаимодействие абонентских систем.

В общем случае телекоммуникационную сеть образуют каналы связи и узлы коммутации (УК).

Обобщенная схема телекоммуникационной сети представлена на рис. 3.7.

Узлы коммутации и абонентские системы обобщенно называются конечным оборудованием данных (ООД).

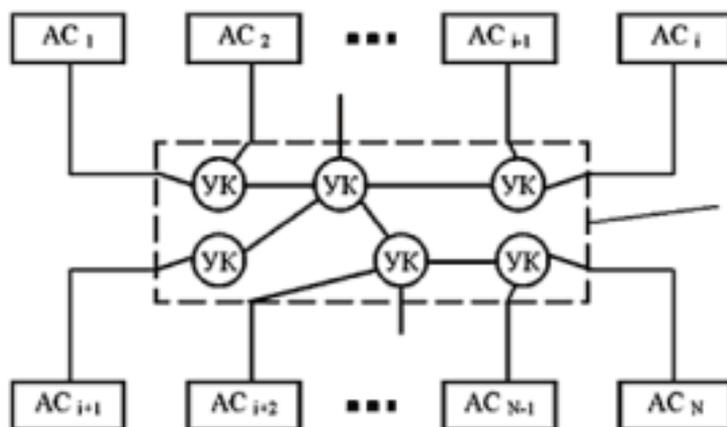


Рис. 3.7. Обобщенная схема телекоммуникационной сети

Каналом связи называют физическую среду и аппаратуру передачи данных (АПД), осуществляющих передачу информации от одного узла коммутации к другому либо между узлом коммутации и абонентской системой.

Физическая среда передачи данных — это пространство или материал, обеспечивающие распространение информационных сигналов.

В качестве физической среды передачи данных в телекоммуникационных сетях могут использоваться:

- проводные (воздушные) линии связи — неизолированные провода, натянутые между телеграфными столбами и висящие в воздухе;
- кабельные линии связи — кабели на основе витых пар медных проводов, коаксиальные кабели, волоконно-оптические кабели и т.п.;
- радиоканалы наземной и спутниковой связи.

Аппаратура передачи данных — совокупность технических средств, предназначенных для преобразования дискретных сигналов, формируемых передающим ООД, в сигналы, передаваемые по физическим линиям связи, и обратного преобразования сигналов, поступающих по линиям связи в принимающее ООД.

В качестве АПД используются модемы, сетевые адаптеры (сетевые карты), оптические модемы, устройства подключения к цифровым каналам и т.п. При большой протяженности линий связи АПД может включать промежуточное оборудование в виде различных усилителей — формирователей передаваемых сигналов. Состав канала связи телекоммуникационной сети представлен на рис. 3.8.

К основным характеристикам телекоммуникационных сетей относятся:

- пропускная способность сети или отдельного канала связи;
- достоверность передачи данных;
- надежность работы.



Рис. 3.8. Состав канала связи

Пропускная способность сети (канала) — максимально возможное количество информации, которое может быть передано по сети (по каналу) за единицу времени. Пропускная способность измеряется в битах в секунду (бит/с), в килобитах в секунду (Кбит/с), в мегабитах в секунду (Мбит/с), в гигабитах в секунду (Гбит/с) и т.д. Достоверность передачи данных — вероятность искажения каждого передаваемого бита информации. Для проводных линий связи вероятность искажения бита информации составляет  $10^{-3}$ — $10^{-4}$ , для кабельных —  $10^{-5}$ — $10^{-6}$ , для оптоволоконных линий —  $10^{-8}$ — $10^{-9}$ . Надежность — свойство телекоммуникационной сети сохранять во времени в установленных пределах значения всех параметров, характеризующих способность сети выполнять требуемые функции в заданных условиях применения.

### 3.3. Функциональный состав, структура и классификация сетей ЭВМ

Полный перечень функций, реализуемых любой сетью ЭВМ, можно представить двумя компонентами — обработка и передача данных.

Вычислительные средства (ЭВМ, вычислительные комплексы и системы) абонентских систем и их программное обеспечение (сетевые операционные системы и приложения) являются основными функциональными элементами сетей ЭВМ, выполняющих обработку данных. Их главная задача состоит в реализации функций предоставления, потребления и распределения ресурсов сети. Вычислительные средства, реализующие весь комплекс перечисленных функций, относятся к универсальным и составляют основу универсальных абонентских систем (УАС). Вычислительные средства, специализированные на предоставлении ресурсов, называются серверами и составляют основу сервисных абонентских систем (САС). Специализированные на потреблении сетевых ресурсов называются клиентами и составляют основу клиентских абонентских систем (КАС). Специализированные на управлении вычислительной сетью называются административными и составляют основу административных абонентских систем (ААС). Классификация сетевых абонентских систем по функциональному признаку представлена на рис. 3.9.

Универсальные абонентские системы используются для построения «одноранговых»<sup>1</sup> сетей ЭВМ. Остальные виды абонентских систем ис-

<sup>1</sup> Одноранговая, децентрализованная или пиринговая (от англ. peer-to-peer, P2P — равный к равному) сеть — это сверхсетевая компьютерная сеть, основанная на равноправии участников. В такой сети отсутствуют выделенные серверы, а каждый узел (peer) является как клиентом, так и сервером. В отличие от архитектуры клиент-сервера такая организация позволяет сохранять работоспособность сети при любом количестве и любом сочетании доступных узлов. Участниками сети являются пиры.



Рис. 3.9. Классификация сетевых абонентских систем

пользуются для построения сетей типа «клиент — сервер» или так называемых сетей с выделенным сервером<sup>1</sup>.

Вычислительные средства узлов коммутации совместно с каналами связи образуют телекоммуникационную сеть с определенной топологической структурой и реализуют функции передачи данных между всеми абонентскими пунктами.

Таким образом, в составе любой сети ЭВМ можно выделить следующие основные функциональные компоненты:

- абонентские системы различного назначения (УАС, САС, КАС, ААС), в совокупности образующие абонентскую сеть;
- узлы коммутации и каналы связи, образующие телекоммуникационную сеть.

В общем виде структура сети ЭВМ представлена на рис. 3.10.

Отдельные сети ЭВМ посредством специального межсетевого оборудования (МСО) могут объединяться между собой, образуя одноуровневые или многоуровневые иерархические структуры (рис. 3.11).

По такому принципу могут объединяться локальные, региональные и глобальные сети ЭВМ.

Сети ЭВМ относятся к разряду сложных вычислительных систем, поэтому для их классификации используется не один, а целый ряд признаков, наиболее характерные из которых представлены на рис. 3.12.

По функциональному назначению сети ЭВМ подразделяются:

- на информационные сети;
- вычислительные сети;
- информационно-вычислительные сети.

Информационные сети предоставляют пользователям в основном информационные услуги. К таким сетям относятся сети научно-технической и справочной информации, резервирования и продажи билетов

<sup>1</sup> Сеть с выделенным сервером (от англ. Client/Server network) — как правило, это локальная вычислительная сеть (LAN), в которой сетевые устройства централизованы и управляются одним или несколькими серверами. Индивидуальные рабочие станции или клиенты (такие как ПК) должны обращаться к ресурсам сети через сервер(ы). Подобное выделение реализуется, в первую очередь, на уровне логики взаимодействия элементов сетевой операционной системы и служит для увеличения управляемости, безопасности и других свойств сети. К понятию QoS имеет опосредованное отношение.

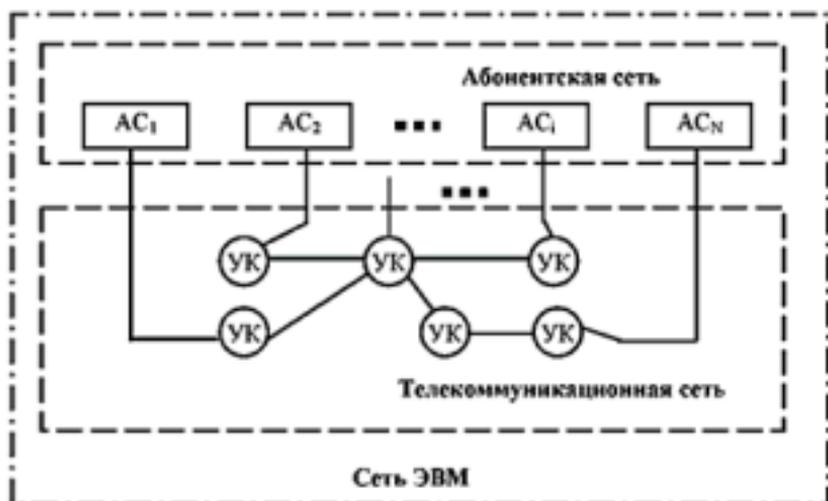


Рис. 3.10. Обобщенная структура сети ЭВМ

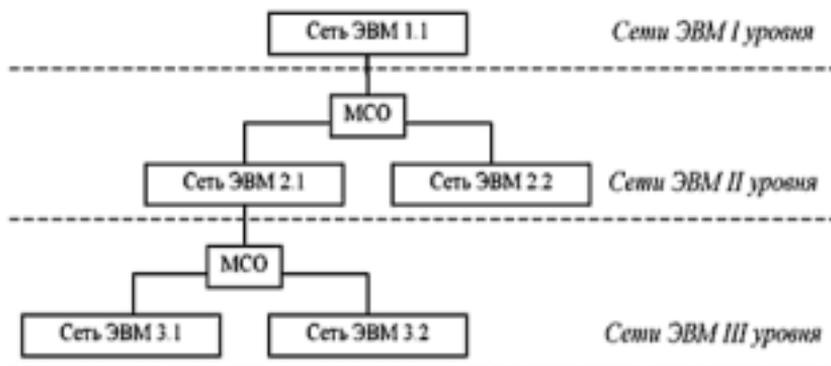


Рис. 3.11. Объединение сетей ЭВМ

на транспорте, сети оперативной информации служб специального назначения и т.д.

Вычислительные сети отличаются наличием в своем составе более мощных вычислительных средств, запоминающих устройств повышенной емкости для хранения прикладных программ, банков данных и знаний, доступных для пользователей, возможностью оперативного перераспределения ресурсов между задачами.



Рис. 3.12. Классификация сетей ЭВМ

На практике наибольшее распространение получили смешанные информационно-вычислительные сети, в которых осуществляются хранение и передача данных, а также решение различных задач по обработке информации. По размещению основных информационных массивов (банков данных) сети подразделяются на следующие типы:

- сети с централизованным размещением информационных массивов;
- сети с локальным (абонентским) размещением информационных массивов.

В сетях с централизованным размещением информационные массивы формируются и хранятся на главном файловом сервере сети. В сетях с локальным размещением информационные массивы могут находиться на различных файловых серверах.

По степени территориальной рассредоточенности компонентов различают:

- глобальные сети, охватывающие территорию страны или нескольких стран;
- сети с расстояниями между отдельными узлами сети в несколько тысяч километров;
- региональные сети, расположенные в пределах определенного территориального региона (города, района, области и т.п.);
- локальные вычислительные сети, охватывающие сравнительно небольшую территорию (в радиусе до 10 км).

По типу используемых вычислительных средств сети могут быть:

- однородными (ЭВМ всех абонентских систем сети аппаратно и программно совместимы);
- неоднородными (ЭВМ абонентских систем сети аппаратно и программно несовместимы).

Локальные сети ЭВМ обычно являются однородными, а региональные и глобальные — неоднородными.

По методу передачи данных различают сети:

- с коммутацией каналов;
- с коммутацией сообщений;
- с коммутацией пакетов;
- со смешанной коммутацией.

Для современных сетей ЭВМ наиболее характерным является использование метода коммутации пакетов. Особенности каждого из методов передачи данных более подробно будут рассмотрены в дальнейшем.

Важным признаком классификации сетей ЭВМ является их топология, т.е. структура связей между элементами сети. Топология оказывает существенное влияние на пропускную способность, на устойчивость сети к отказам ее оборудования, на качество обслуживания запросов пользователей, на логические возможности и стоимость сети.

Для построения сетей ЭВМ используются следующие топологические структуры (рис. 3.13):

- радиальная (звездообразная);
- кольцевая;
- шинная;
- полносвязная;
- древовидная (иерархическая).

Основу сетей с радиальной (звездообразной) топологией (рис. 3.13, а) составляет главный центр, который может быть как активным (выполняется обработка информации), так и пассивным (выполняется только ретрансляция информации). Такие сети довольно просты по своей структуре и организации управления. К недостаткам сетей с радиальной топологией можно отнести: нарушение связи при выходе из строя центрального узла коммутации, отсутствие свободы выбора различных маршрутов для установления связи между АС, увеличение задержек в обслуживании запросов при перегрузке центра обработки, значительное возрастание общей протяженности линий связи при размещении АС на большой территории.

В сетях с кольцевой топологией (рис. 3.13, б) информация между абонентскими станциями передается только в одном направлении. Кольцевая структура обеспечивает широкие функциональные возможности сети при высокой эффективности использования моноканала, низкой стоимости, простоте методов управления, возможности контроля работоспособности моноканала. К недостаткам сетей с кольцевой топологией можно отнести нарушение связи при выходе из строя хотя бы одного сегмента канала передачи данных.

В сетях с шинной топологией (рис. 3.13, в) используется моноканал передачи данных, к которому подсоединяются абонентские системы. Данные от передающей АС распространяются по каналу в обе стороны. Информация поступает на все АС, но принимает сообщение только та АС, которой оно адресовано. Шинная топология — одна из наиболее простых. Она позволяет легко наращивать и управлять сетью ЭВМ, является наиболее устойчивой к возможным неисправностям отдельных абонентских систем. Недостатком шинной топологии является полный выход из строя сети при нарушении целостности моноканала.

В полносвязной сети (рис. 3.13, г) информация может передаваться между всеми АС по собственным каналам связи. Такое построение сети требует большого числа соединительных линий связи. Оно эффективно для малых сетей с небольшим количеством центров обработки, работающих с полной загрузкой каналов связи.

В сетях с древовидной топологией (рис. 3.13, д) реализуется объединение нескольких более простых сетей с шинной топологией. Каждая

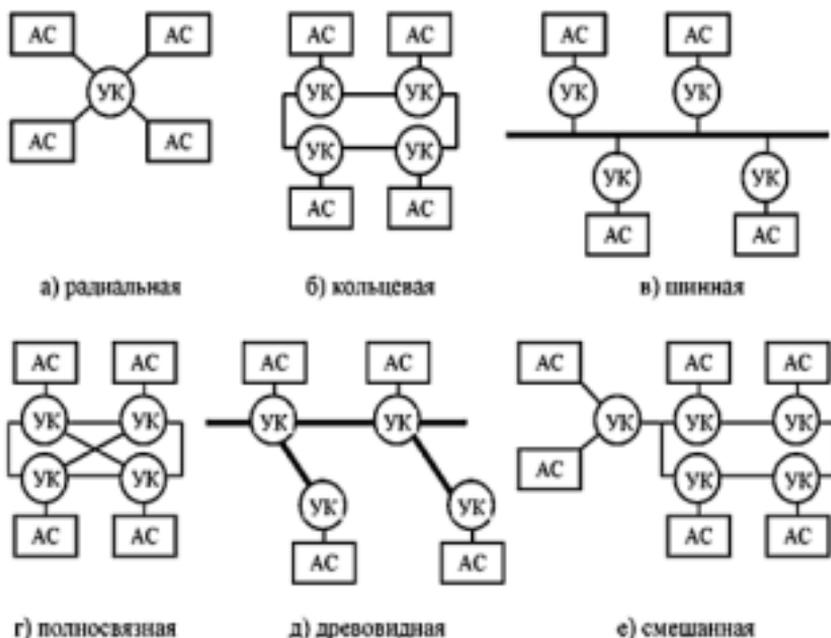


Рис. 3.13. Топологические структуры сетей ЭВМ

ветвь дерева представляет собой сегмент. Отказ одного сегмента не приводит к выходу из строя остальных сегментов.

Топология крупных сетей обычно представляет собой комбинации нескольких топологических решений. Примером такой сети может служить сеть со смешанной радиально-кольцевой топологией, представленная на рис. 3.13, е.

Правильный и рациональный выбор основных функциональных, технических и программных компонентов сетей ЭВМ, их топологической структуры оказывает непосредственное влияние на все технические характеристики и общую эффективность функционирования сетей ЭВМ в целом. Это особенно важно для вычислительных сетей военного назначения, предназначенных для обработки и передачи больших информационных массивов данных в условиях жесткого лимита времени и высоких требований к достоверности информации.

## 3.4. Открытая модель взаимодействия ISO/OSI

Организация взаимодействия между узлами в компьютерной сети (и, как следствие, одновременная работа пользователей в сети) является сложной проблемой, она включает много аспектов, начиная с согласования уровней электрических сигналов, формирования кадров, проверки контрольных сумм и кончая вопросами аутентификации приложений. Для ее решения используется универсальный прием — разбиение одной сложной задачи на несколько частных, более простых задач. Средства решения отдельных задач упорядочены в виде иерархии уровней. Для решения задачи некоторого уровня могут быть использованы средства непосредственно примыкающего нижележащего уровня. С другой стороны, результаты работы средств некоторого уровня могут быть переданы только средствам соседнего вышележащего уровня.

Многоуровневое представление средств сетевого взаимодействия имеет свою специфику, связанную с тем, что в процессе обмена сообщениями участвуют две машины, т.е. в данном случае необходимо организовать согласованную работу двух «иерархий». При передаче сообщений оба участника сетевого обмена должны принять множество соглашений. Например, они должны согласовать способ кодирования электрических сигналов, правило определения длины сообщений, договориться о методах контроля достоверности и т.п. Другими словами, соглашения должны быть приняты для всех уровней, начиная от самого низкого уровня передачи битов до самого высокого уровня, предоставляющего услуги пользователям сети.

Формализованные правила, определяющие последовательность и формат сообщений, которыми обмениваются сетевые компоненты, лежащие на одном уровне, но в разных узлах, называются протоколом.

Модули, реализующие протоколы соседних уровней и находящиеся в одном узле, также взаимодействуют друг с другом в соответствии с четко определенными правилами и с помощью стандартизованных форматов сообщений. Эти правила принято называть интерфейсом. Интерфейс определяет услуги, предоставляемые данным уровнем соседнему уровню.

В сущности, протокол и интерфейс выражают одно и то же понятие, но традиционно в сетях за ними закрепили разные области действия: протоколы определяют правила взаимодействия модулей одного уровня в разных узлах, а интерфейсы — модулей соседних уровней в одном узле.

Средства каждого уровня должны обрабатывать, во-первых, свой собственный протокол, а во-вторых, интерфейсы с соседними уровнями. Иерархически организованный набор протоколов, достаточный для организации взаимодействия узлов в сети, называется стекком коммуникационных протоколов.

Коммуникационные протоколы могут быть реализованы как программно, так и аппаратно. Протоколы нижних уровней часто реализуются комбинацией программных и аппаратных средств, протоколы верхних уровней, как правило, чисто программными средствами.

В начале 80-х гг. XX в. ряд международных организаций по стандартизации — ISO, ITU-T и некоторые другие — разработали модель, которая сыграла значительную роль в развитии сетей. Эта модель называется моделью взаимодействия открытых систем (Open System interconnection, OSI), или моделью OSI<sup>1</sup>. Модель OSI определяет различные уровни взаимодействия систем, дает им стандартные имена и указывает, какие функции должен выполнять каждый уровень. В модели ISO/OSI (рис. 3.14) средства взаимодействия делятся на семь уровней: прикладной, представительный, сеансовый, транспортный, сетевой, каналный и физический. Каждый уровень имеет дело с одним определенным аспектом взаимодействия сетевых устройств.

Функции всех уровней модели OSI могут быть отнесены к одной из двух групп: либо к функциям, зависящим от конкретной технической

<sup>1</sup> Также еще принято называть ее как сетевая модель OSI (for engl. Open Systems Interconnection Basic Reference Model) — базовая эталонная модель взаимодействия открытых систем. В советской научно-технической литературе XX в. также использовалось сокращение ЗМВОС, которое впервые было опубликовано в 1978 г. — Электронная Модель Взаимодействия Открытых Систем. Также ее понимают как сетевую модель стека сетевых протоколов OSI/ISO (ГОСТ Р ИСО/МЭК 7498-1—99).

реализации сети, либо к функциям, ориентированным на работу с приложениями.

Любой протокол модели OSI должен взаимодействовать либо с протоколами своего уровня, либо с протоколами на единицу выше и (или) ниже своего уровня. Взаимодействия с протоколами своего уровня называются горизонтальными, а с уровнями на единицу выше или ниже — вертикальными. Любой протокол модели OSI может выполнять

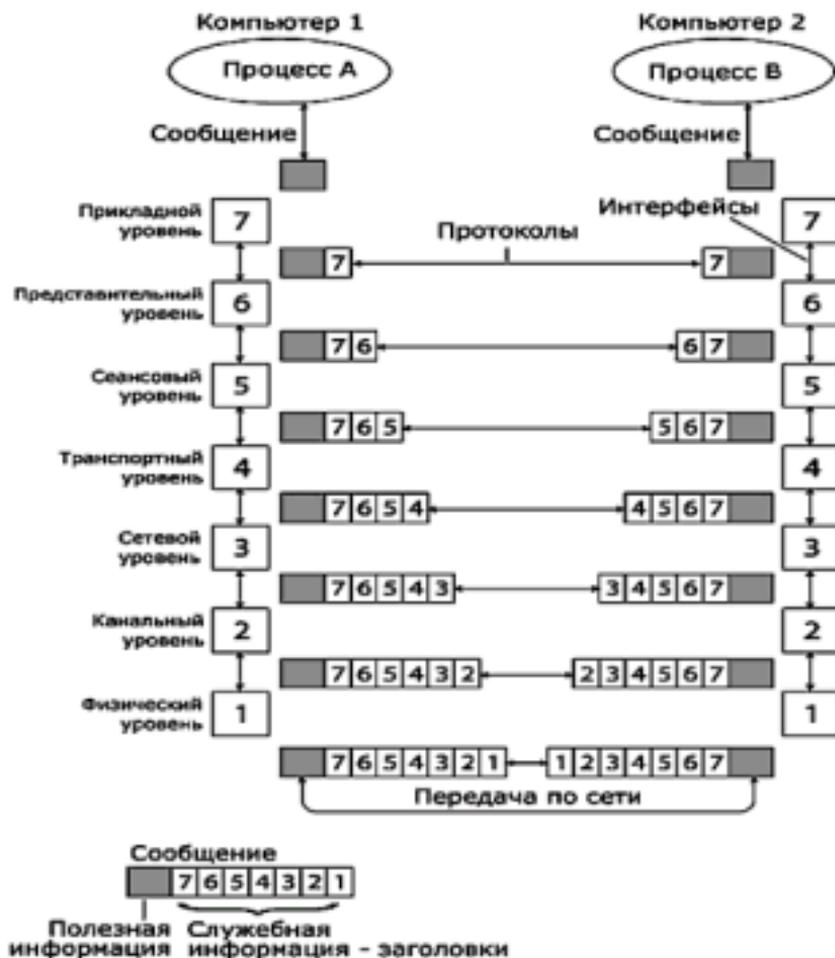


Рис. 3.14. Открытая модель взаимодействия ISO/OSI

функции только своего уровня и не может выполнять функций другого уровня, что не выполняется в протоколах альтернативных моделей.

Каждому уровню с некоторой долей условности соответствует свой операнд — логически неделимый элемент данных, которым на отдельном уровне можно оперировать в рамках модели и используемых протоколов: на физическом уровне мельчайшая единица — бит, на канальном уровне информация объединена в кадры, на сетевом — в пакеты (датаграммы), на транспортном — в сегменты. Любой фрагмент данных, логически объединенных для передачи, — кадр, пакет, датаграмма — считается сообщением. Именно сообщения в общем виде являются операндами сеансового, представительского и прикладного уровней. Распределение вышеупомянутых вариантов взаимодействия представлено в табл. 3.1.

Таблица 3.1

#### Распределение вариантов взаимодействия в модели ISO/OSI

Тип данных	Уровень (layer) в модели ISO/OSI	Функции
Данные	7. Прикладной (application)	Доступ к сетевым службам
	6. Представительский (presentation)	Представление и шифрование данных
	5. Сеансовый (session)	Управление сеансом связи
Сегменты	4. Транспортный (transport)	Прямая связь между конечными пунктами и надежность
Пакеты	3. Сетевой (network)	Определение маршрута и логическая адресация
Кадры	2. Канальный (data link)	Физическая адресация
Биты	1. Физический (physical)	Работа со средой передачи, сигналами и двоичными данными

К базовым сетевым технологиям относятся физический и канальный уровни. Три нижних уровня — физический, канальный и сетевой — являются сетезависимыми, т.е. протоколы этих уровней тесно связаны с технической реализацией сети, с используемым коммуникационным оборудованием. Например, переход на оборудование FDDI означает полную смену протоколов физического и канального уровня во всех узлах сети.

Три верхних уровня — сеансовый, уровень представления и прикладной — ориентированы на приложения и мало зависят от технических особенностей построения сети. На протоколы этих уровней не вли-

яют никакие изменения в топологии сети, замена оборудования или переход на другую сетевую технологию. Так, переход от Ethernet на высокоскоростную технологию ATM не потребует никаких изменений в программных средствах, реализующих функции прикладного, представительного и сеансового уровней.

Транспортный уровень является промежуточным, он скрывает все детали функционирования нижних уровней от верхних уровней. Это позволяет разрабатывать приложения, независимые от технических средств, непосредственно занимающихся транспортировкой сообщений.

Рисунок 3.15 показывает уровни модели ISO/OSI, на которых работают различные элементы сети. Компьютер с установленной на нем сетевой ОС взаимодействует с другим компьютером с помощью протоколов всех семи уровней. Это взаимодействие компьютеры осуществляют

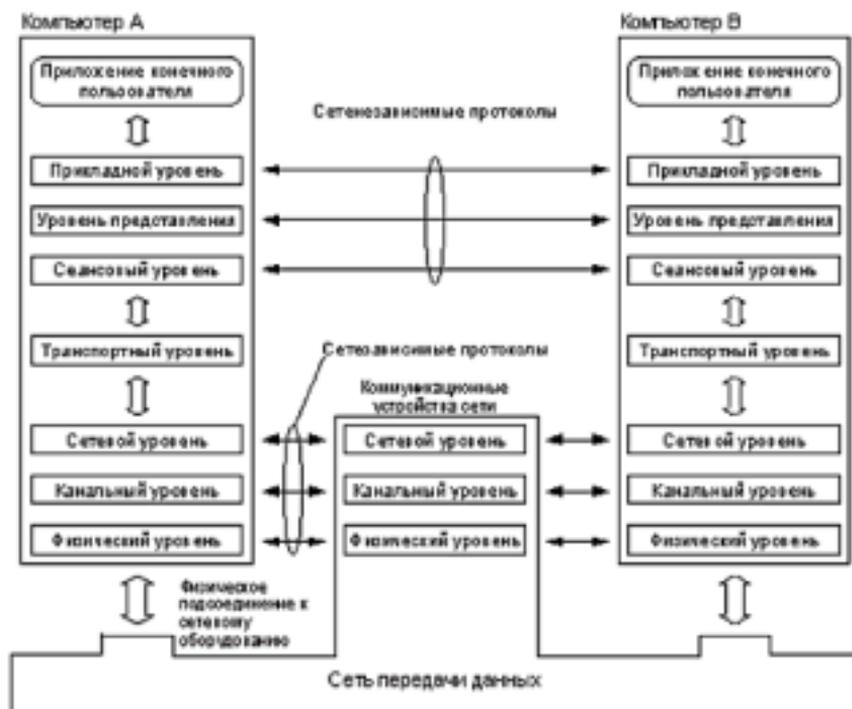


Рис. 3.15. Сетезависимые и сетезависимые протоколы ISO/OSI

ют через различные коммуникационные устройства: концентраторы, модемы, мосты, коммутаторы, маршрутизаторы, мультиплексоры.

В зависимости от типа коммуникационное устройство может работать либо только на физическом уровне (повторитель), либо на физическом и канальном (мост и коммутатор), либо на физическом, канальном и сетевом, иногда захватывая и транспортный уровень (маршрутизатор).

Рассмотренная ранее открытая модель взаимодействия применима к любому классу компьютерных сетей. Поэтому следует разделить все известные сегодня компьютерные сети в соответствии с известными классификационными признаками.

Подобные варианты взаимодействия, описанные в рассмотренной модели ISO/OSI (ЭМБОК), позволили создать то, что сегодня носит термин «конвергенция»<sup>1</sup>, или более просто — возможность объединения сетей в единое целое.

### 3.5. Методы структуризации сетей ЭВМ

Построение компьютерных сетей (или еще можно сказать сетей ЭВМ) с небольшим<sup>2</sup> (10—30) количеством абонентских систем чаще всего осуществляется на основе одной из типовых топологий: общая шина, кольцо, звезда или полносвязная сеть. Все перечисленные топологии обладают свойством однородности, т.е. все компьютеры абонентских систем в такой сети имеют одинаковые права в отношении информационного взаимодействия друг с другом (за исключением центрального компьютера при соединении звезда). Такая однородность структуры значительно упрощает процедуру наращивания общего числа абонентских систем, облегчает обслуживание и эксплуатацию сети ЭВМ. Однако увеличение количества абонентских систем в сети сверх указанного количества приводит к существенному снижению пропускной способности каналов связи и общей эффективности функционирования сети.

<sup>1</sup> Конвергенция (от лат. *convergo* — сближаю) — процесс сближения, схождения (в разном смысле), компромиссов; противоположна дивергенции. Конвергенция телекоммуникационных сетей — объединение несвязных, бывших ранее раздельными, услуг в рамках одной услуги. Например, услуга современных провайдеров интернет Triple Play — это объединение телефонии, интернета, телевидения в одном кабельном Интернет-подключении.

<sup>2</sup> Термин «небольшое число» достаточно относителен, поскольку такое число узлов сети обеспечивает незначительное по мнению большинства авторов загрузку основной среды распространения информации. Но такое положение характерно для обособленной сети, как правило, относящейся к классу «локальная». Как только такая сеть получает возможность связи на уровне обмена данными с другими сетями, такое число узлов сети не является «незначительным» и начинают работать совсем другие законы расчета загруженности, безопасности и т.д.

Одним из основных направлений разрешения указанной проблемы является применение методов структуризации больших сетей ЭВМ.

Построение больших сетей ЭВМ, объединяющих более 30 абонентских систем, на основе унифицированных типовых топологических структур порождает различные ограничения, наиболее существенными из которых являются:

- ограничения на длину связи между узлами;
- ограничения на количество узлов в сети;
- ограничения на интенсивность трафика, порождаемого узлами сети.

Например, технология Ethernet на тонком коаксиальном кабеле позволяет использовать кабель длиной не более 185 м, к которому можно подключить не более 30 сетевых ЭВМ (рис. 3.16). Однако если абонентские системы интенсивно обмениваются информацией между собой, то приходится снижать число подключенных к каналу компьютеров до 10—20, чтобы каждой абонентской системе доставалась приемлемая доля общей пропускной способности сети.

Для снятия этих ограничений используются специальные методы структуризации сети и специальное структурообразующее оборудование — повторители, концентраторы, мосты, коммутаторы, маршрутизаторы. Такое оборудование также называют коммуникационным.

Простейшим из коммуникационных устройств является повторитель (repeater). Повторители используются для физического соединения различных сегментов кабеля локальной сети ЭВМ с целью увеличения общей длины сети. Повторитель передает сигналы, приходящие из одного сегмента сети в другие ее сегменты (см. рис. 3.16). Повторитель позволяет преодолеть ограничения на длину линий связи за счет улучшения качества передаваемого сигнала — восстановления его мощности и амплитуды, улучшения фронтов и т.п.

Повторитель, который имеет несколько портов и соединяет несколько физических сегментов, называется концентратором, или хабом.

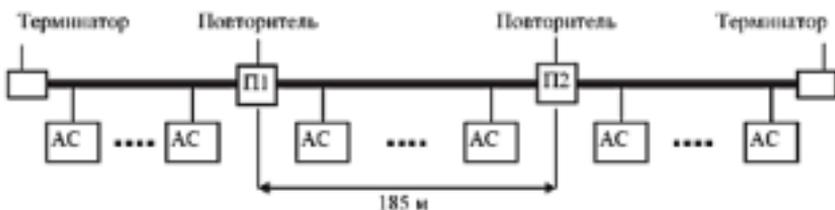


Рис. 3.16. Увеличение сети ЭВМ на основе повторителей

В данном устройстве сосредотачиваются все связи между сегментами сети. Концентраторы характерны практически для всех базовых технологий локальных сетей — Ethernet, ArcNet, Token Ring, FDDI и т.п. В работе концентраторов различных типов и технологий много общего — они повторяют сигналы, пришедшие с одного из своих портов, на других своих портах. Разница состоит в том, на каких именно портах повторяются входные сигналы. Так, концентратор Ethernet повторяет входные сигналы на всех своих портах, кроме того, с которого сигналы поступают (рис. 3.17, а). А концентратор Token Ring (рис. 3.17, б) повторяет входные сигналы, поступающие с некоторого порта, только на одном порту — на том, к которому подключена следующая в кольце АС.

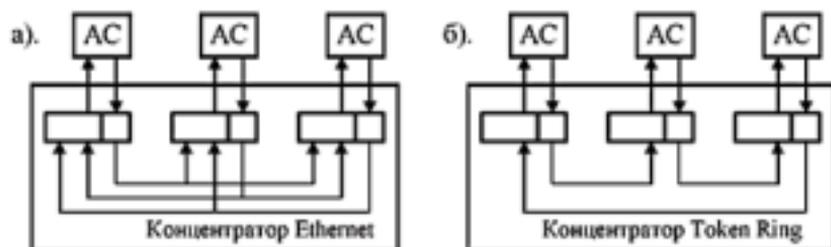


Рис. 3.17. Концентраторы технологий: а) — Ethernet; б) — Token Ring

Концентратор всегда изменяет физическую топологию сети, но при этом оставляет без изменения ее логическую топологию.

Под физической топологией понимается конфигурация связей, образованных отдельными частями кабеля, а под логической — конфигурация информационных потоков между компьютерами сети. Во многих случаях физическая и логическая топологии сети совпадают (рис. 3.18, а). Однако это выполняется не всегда. Сеть на рис. 3.18, б демонстрирует пример несовпадения физической и логической топологии. Физически компьютеры соединены по топологии «общая шина», а логически — по кольцевой топологии.

Физическая структуризация сети с помощью концентраторов целесообразна не только для увеличения расстояния между узлами сети, но и для повышения ее надежности. Например, если какая-либо абонентская система сети Ethernet с физической общей шиной из-за сбоя начинает непрерывно передавать данные по общему кабелю, то вся сеть выходит из строя, и для решения этой проблемы остается только один выход — вручную отсоединить сетевой адаптер этой абонентской системы от кабеля.

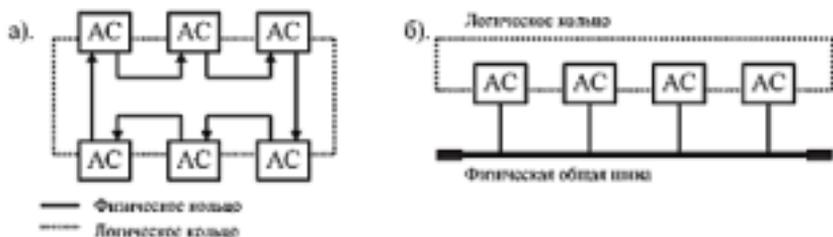


Рис. 3.18. Физическая и логическая топологии сети

В сети Ethernet, построенной с использованием концентратора, эта проблема может быть решена автоматически — концентратор отключает свой порт, если обнаруживает, что присоединенный к нему узел слишком долго монопольно занимает сеть. Концентратор может блокировать некорректно работающий узел и в других случаях, выполняя роль некоторого управляющего узла.

Физическая структуризация не позволяет решать в полной мере проблемы, связанные с построением и функционированием больших многоуровневых и составных сетей ЭВМ. Наиболее важной проблемой, нерешаемой путем физической структуризации, остается проблема перераспределения передаваемого трафика между различными физическими сегментами сети.

На рисунке 3.19, а приведена структурная схема составной сети, построенной на основе концентраторов (физической структуризации). В такой сети при информационном обмене между любой парой абонентских систем (предположим между абонентскими системами А и В сегмента 1) концентраторы распространяют любой информационный кадр по всем ее сегментам. Поэтому кадр, посылаемый абонентской системой А абонентской системе В, хотя и не нужен абонентским системам сегментов 2 и 3, в соответствии с логикой работы концентраторов поступает на эти сегменты тоже. И до тех пор, пока АС В не получит адресованный ей кадр, ни одна из абонентских систем этой сети не сможет передавать данные.

Такая ситуация возникает из-за того, что логическая структура сети после ее физической структуризации остается неизменной (однородной) — она никак не учитывает интенсивность трафика внутри сегментов и предоставляет всем парам абонентских систем равные возможности по обмену информацией (рис. 3.19, б). Данная проблема может быть решена путем логической структуризации сети.

Логическая структуризация сети — это разбиение сети на сегменты с локализованным трафиком. Под локализацией трафика понимается распространение трафика, предназначенного для абонентских систем конкретного сегмента сети, только в пределах этого сегмента.

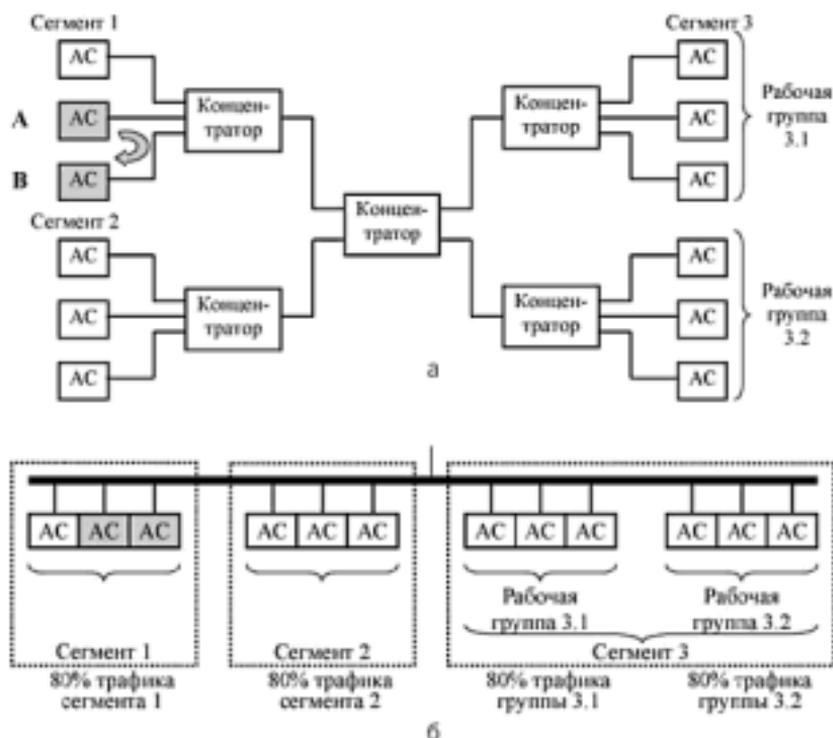


Рис. 3.19. Противоречие между логической структурой сети и информационными потоками: а — физическая структуризация с помощью концентраторов разделяемый канал; б — логическая структура сети

Локализация трафика позволяет значительно повысить общую эффективность сети, так как позволяет организовать информационный обмен одновременно внутри всех сегментов сети. Для логической структуризации сети используются специальные коммуникационные устройства, такие как мосты, коммутаторы, маршрутизаторы и шлюзы.

Мост<sup>1</sup> (Bridge) делит сеть на логические сегменты. Адреса абонентских систем внутри одного сегмента принадлежат заданному диапазону. Передача информации из одного сегмента сети в другой возможна лишь при условии, если адресуемая абонентская система принадлежит другому логическому сегменту, т.е. мост изолирует трафики логических сегментов друг от друга, повышая общую производительность передачи данных в сети. Одновременно мост может обеспечивать информационный обмен только между двумя сегментами сети. Локализация трафика также снижает вероятность несанкционированного доступа к передаваемым по сети данным, так как кадры не выходят за пределы своего сегмента и их сложнее перехватить злоумышленнику.

На рисунке 3.20 показана логическая структуризация сети с помощью моста. Абонентская система А (сегмент 1) передает информационный кадр абонентской системе В (сегмент 2).

Коммутатор (Switch) по принципу обработки информационных кадров полностью аналогичен мосту. Основное его отличие от моста состоит в том, что он способен осуществлять информационный обмен одновременно между несколькими парами логических сегментов сети, так как каждый его порт оснащен специализированным процессором, который обрабатывает кадры по алгоритму моста независимо от процессоров других портов. За счет этого общая производительность коммутатора обычно намного выше производительности традиционного моста, имеющего один процессорный блок.

Маршрутизатор (Router) — коммуникационное устройство с расширенными интеллектуальными возможностями по сравнению с мостами и коммутаторами. Маршрутизаторы реализуют более эффективные методы разграничения трафика отдельных логических сегментов сети. Это

<sup>1</sup> Мост — это устройство, обеспечивающее объединение двух (реже нескольких) локальных сетей посредством передачи кадров из одной сети в другую с их промежуточной буферизацией. Мост в отличие от регенераторов не обеспечивает побитовой синхронизации объединяемых сетей. По отношению к каждой из них он оконечный узел. Мост принимает кадр, буферизирует его, анализирует адрес назначения кадра и только в том случае, когда адресуемый узел действительно принадлежит другой сети, передает его туда. Для передачи кадра в другую сеть мост должен получить доступ к ее физической среде в соответствии с теми же правилами, что и обычный узел. Таким образом, мост, производя фильтрацию кадров, изолирует трафик одного сегмента от трафика другого. Наиболее распространенный тип мостов — прозрачные. Для них сеть представляется набором MAC-адресов устройств, используемых на канальном уровне, причем каждый набор связан с определенным портом моста. Мосты используют MAC-адреса, принимая решение о ретрансляции кадра, когда последний записывается во внутренний буфер моста из какого-либо его порта. Мосты не имеют доступа к информации об адресах, относящихся к более высокому (сетевому) уровню, и ничего не знают о топологии связей сегментов или сетей между собой. Таким образом, мосты совершенно прозрачны для протоколов сетевого уровня и выше. Это позволяет им передавать пакеты различных протоколов высших уровней без искажений. Мосты регенерируют кадры, передаваемые с порта на порт. Одно из преимуществ использования мостов — увеличение расстояния, покрываемого сетью, так как количество пересеканий мостов не оказывает влияния на качество сигнала.

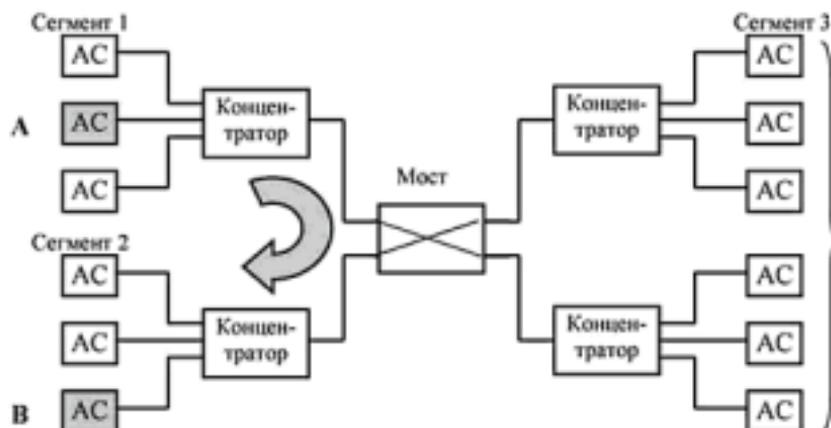


Рис. 3.20. Логическая структуризация сети с помощью моста

достигается за счет использования составных числовых адресов и явной адресации логических сегментов сети (рис. 3.21).

Кроме локализации трафика маршрутизаторы способны выполнять еще ряд задач, наиболее важными из которых являются выбор из нескольких возможных наиболее рационального маршрута передачи информационного кадра и организация взаимодействия сегментов сети, реализованных на основе различных сетевых технологий.

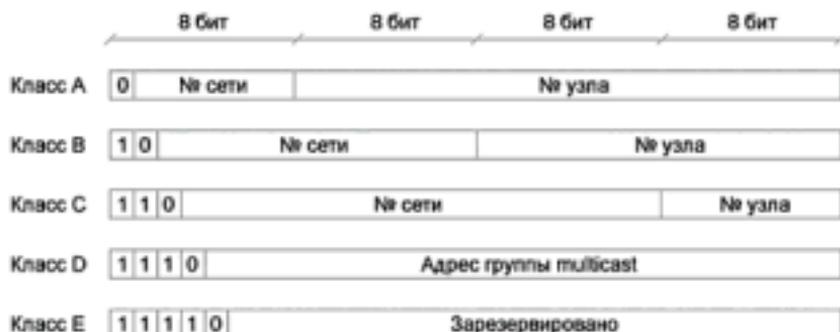


Рис. 3.21. Структура составного адреса (пример)

Логическая структуризация сети с помощью маршрутизаторов представлена на рис. 3.22. Особенностью данной сети является наличие дополнительной связи между сегментами 1 и 2, которая может благодаря наличию маршрутизаторов использоваться как для повышения производительности сети, так и для повышения ее надежности. В данной сети информационный обмен осуществляется одновременно между двумя парами абонентских систем А и В, С и D.

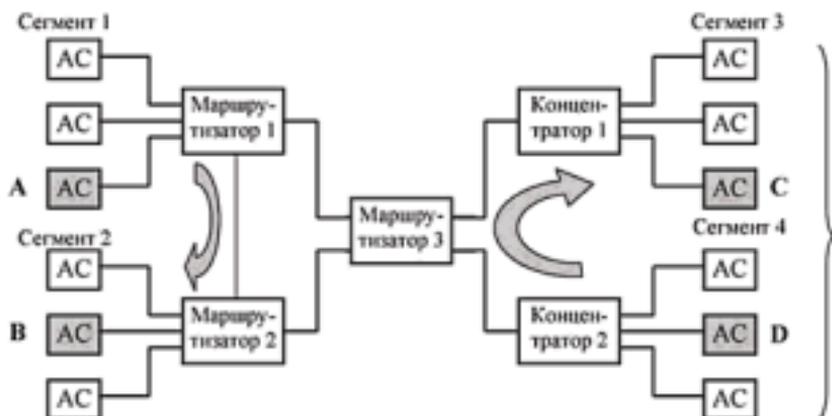


Рис. 3.22. Логическая структуризация сети с помощью маршрутизаторов

Такие системы способны локализовывать трафик и логически структурировать сети, однако основное их предназначение — объединять сети с разными типами системного и прикладного программного обеспечения. Крупные сети практически никогда не строятся без физической и логической структуризации. Для отдельных сегментов и подсетей характерны типовые однородные топологии базовых технологий, и для их объединения всегда используется оборудование, обеспечивающее локализацию трафика, — мосты, коммутаторы, маршрутизаторы, шлюзы и т.п.

Подобные подходы, которые исследовались еще на базе описанного ранее «полигона» — сети ARPAnet, позволили создать самый мощный инструмент для информационных технологий — сети «Интернет».

### 3.6. Интернет как инструмент для ИТ

Протокол IP был специально создан агностическим<sup>1</sup> в отношении физических каналов связи. То есть любая система (сеть) передачи цифровых

<sup>1</sup> Данный термин происходит от философского понятия «агностицизм». Агностицизм (от др.-греч.

данных, проводная или беспроводная, для которой существует стандарт инкапсуляции в нее IP-пакетов, может передавать и трафик Интернета. Агностицизм протокола IP, в частности, означает, что компьютер или маршрутизатор должен знать тип сетей, к которым он непосредственно присоединен, и уметь работать с этими сетями; но не обязан (и в большинстве случаев не может) знать, какие сети находятся за маршрутизаторами.

На стыках сетей специальные маршрутизаторы (программные или аппаратные) занимаются автоматической сортировкой и перенаправлением пакетов данных, исходя из IP-адресов получателей этих пакетов. Протокол IP образует единое адресное пространство в масштабах всего мира, но в каждой отдельной сети может существовать и собственное адресное подпространство, которое выбирается исходя из класса сети. Такая организация IP-адресов позволяет маршрутизаторам однозначно определять дальнейшее направление для каждого пакета данных. В результате между отдельными сетями Интернета не возникает конфликтов, и данные беспрепятственно и точно передаются из сети в сеть по всей планете и ближнему космосу.

Сам протокол IP был рожден в дискуссиях внутри организации IETF (от англ. Internet Engineering Task Force; Task force — группа специалистов для решения конкретной задачи), чье название можно вольно перевести как «Группа по решению задач проектирования Интернета». IETF и ее рабочие группы по сей день занимаются развитием протоколов Всемирной сети. IETF открыта для публичного участия и обсуждения. Комитеты организации публикуют так называемые документы RFC. В этих документах даются технические спецификации и точные объяснения по многим вопросам. Некоторые документы RFC возводятся организацией IAB (от англ. Internet Architecture Board — Совет по архитектуре Интернета) в статус стандартов Интернета (от англ. Internet Standard). С 1992 года IETF, IAB и ряд других Интернет-организаций входят в Общество Интернета (от англ. Internet Society, ISOC). Общество Интернета предоставляет организационную основу для разных исследовательских и консультативных групп, занимающихся развитием Интернета.

В стечке TCP/IP используются три типа адресов:

- локальные адреса (называемые также аппаратными);
- IP-адреса;
- символьные доменные имена.

---

αγνωστος — непознаваемый, непознанный) — позиция, существующая в философии, теории познания и теологии, полагающая принципиально невозможным познание объективной действительности только через субъективный опыт и невозможным познание любых предельных и абсолютных основ реальности. Также отрицается возможность доказательства или опровержения идей и утверждений, основанных полностью на субъективных посылах. Иногда агностицизм определяется как философское учение, утверждающее принципиальную непознаваемость мира.

В терминологии TCP/IP под локальным адресом понимается такой тип адреса, который используется средствами базовой технологии для доставки данных в пределах подсети, являющейся элементом составной интeрcети. В разных подсетях допустимы разные сетевые технологии, разные стеки протоколов, поэтому при создании стека TCP/IP предполагалось наличие разных типов локальных адресов. Если подсеть интeрcети является локальная сеть, то локальный адрес — это MAC-адрес. MAC-адрес назначается сетевым адаптерам и сетевым интерфейсам маршрутизаторов. MAC-адреса назначаются производителями оборудования и являются уникальными, так как управляются централизованно. Для всех существующих технологий локальных сетей MAC-адрес имеет формат 6 байт, например, 11-A0-17-3D-BC-01. Однако протокол IP может работать и над протоколами более высокого уровня, например, над протоколом IPX или X.25. В этом случае локальными адресами для протокола IP соответственно будут адреса IPX и X.25. Следует учесть, что компьютер в локальной сети может иметь несколько локальных адресов даже при одном сетевом адаптере. Некоторые сетевые устройства не имеют локальных адресов. Например, к таким устройствам относятся глобальные порты маршрутизаторов, предназначенные для соединений типа «точка-точка».

IP-адреса представляют собой основной тип адресов, на основании которых сетевой уровень передает пакеты между сетями. Эти адреса состоят из 4 байт, например, 109.26.17.100. IP-адрес назначается администратором во время конфигурирования компьютеров и маршрутизаторов. IP-адрес состоит из двух частей: номера сети и номера узла. Номер сети может быть выбран администратором произвольно либо назначен по рекомендации специального подразделения Internet (Internet Network Information Center, InterNIC), если сеть должна работать как составная часть Интернета. Обычно поставщики услуг Интернет получают диапазоны адресов у подразделений InterNIC, а затем распределяют их между своими абонентами. Номер узла в протоколе IP назначается независимо от локального адреса узла. Маршрутизатор по определению входит сразу в несколько сетей. Поэтому каждый порт маршрутизатора имеет собственный IP-адрес. Конечный узел также может входить в несколько IP-сетей. В этом случае компьютер должен иметь несколько IP-адресов, по числу сетевых связей. Таким образом, IP-адрес характеризует не отдельный компьютер или маршрутизатор, а одно сетевое соединение.

Символьные доменные имена. Символьные имена в IP-сетях называются доменными и строятся по иерархическому признаку. Составляющие полного символьного имени в IP-сетях разделяются точкой и перечис-

ляются в следующем порядке: сначала простое имя конечного узла, затем имя группы узлов (например, имя организации), затем имя более крупной группы (поддомена) и так до имени домена самого высокого уровня (например, домена объединяющего организации по географическому принципу: RU — Россия, UK — Великобритания, SU — США). Примером доменного имени может служить имя `base2.sales.zil.ru`. Между доменным именем и IP-адресом узла нет никакого алгоритмического соответствия, поэтому необходимо использовать какие-то дополнительные таблицы или службы, чтобы узел сети однозначно определялся как по доменному имени, так и по IP-адресу. В сетях TCP/IP используется специальная распределенная служба Domain Name System (DNS), которая устанавливает это соответствие на основании создаваемых администраторами сети таблиц соответствия. Поэтому доменные имена называют также DNS-именами.

Для идентификации компьютеров аппаратное и программное обеспечение в сетях TCP/IP полагается на IP-адреса, поэтому для доступа к сетевому ресурсу в параметрах программы вполне достаточно указать IP-адрес, чтобы программа правильно поняла, к какому хосту ей нужно обратиться. Например, команда `ftp://192.45.66.17` будет устанавливать сеанс связи с нужным ftp-сервером, а команда `http://203.23.106.33` откроет начальную страницу на корпоративном Web-сервере. Однако пользователи обычно предпочитают работать с символьными именами компьютеров, и операционные системы локальных сетей приучили их к этому удобному способу. Следовательно, в сетях TCP/IP должны существовать символьные имена хостов и механизм для установления соответствия между символьными именами и IP-адресами.

В операционных системах, которые первоначально разрабатывались для работы в локальных сетях, таких как Novell NetWare, Microsoft Windows или IBM OS/2, пользователи всегда работали с символьными именами компьютеров. Так как локальные сети состояли из небольшого числа компьютеров, то использовались так называемые плоские имена, состоящие из последовательности символов, не разделенных на части. Примерами таких имен являются: `NW1_1`, `mail2`, `MOSCOW_SALES_2`. Для установления соответствия между символьными именами и MAC-адресами в этих операционных системах применялся механизм широковещательных запросов, подобный механизму запросов протокола ARP. Так, широковещательный способ разрешения имен реализован в протоколе NetBIOS, на котором были построены многие локальные ОС. Так называемые NetBIOS-имена стали на долгие годы одним из основных типов плоских имен в локальных сетях.

Для стека TCP/IP, рассчитанного в общем случае на работу в больших территориально распределенных сетях, подобный подход оказывается неэффективным по нескольким причинам.

Плоские имена не дают возможности разработать единый алгоритм обеспечения уникальности имен в пределах большой сети. В небольших сетях уникальность имен компьютеров обеспечивает администратор сети, записывая несколько десятков имен в журнале или файле. При росте сети задачу решают уже несколько администраторов, согласовывая имена между собой неформальным способом. Однако если сеть расположена в разных городах или странах, то администраторам каждой части сети нужно придумать способ именования, который позволил бы им давать имена новым компьютерам независимо от других администраторов, обеспечивая в то же время уникальность имен для всей сети. Самый надежный способ решения этой задачи — отказ от плоских имен в принципе.

Широковещательный способ установления соответствия между символьными именами и локальными адресами хорошо работает только в небольшой локальной сети, неразделенной на подсети. В крупных сетях, где общая широковещательность не поддерживается, нужен другой способ разрешения символьных имен. Обычно хорошей альтернативой широковещательности является применение централизованной службы, поддерживающей соответствие между различными типами адресов всех компьютеров сети. Компания Microsoft для своей корпоративной операционной системы Windows NT разработала централизованную службу WINS, которая поддерживает базу данных NetBIOS-имен и соответствующих им IP-адресов.

Для эффективной организации именования компьютеров в больших сетях естественным является применение иерархических составных имен.

В стеке TCP/IP применяется доменная система имен, которая имеет иерархическую древовидную структуру, допускающую использование в имени произвольного количества составных частей (рис. 3.23).

Иерархия доменных имен аналогична иерархии имен файлов, принятой во многих популярных файловых системах. Дерево имен начинается с корня, обозначаемого здесь точкой (.). Затем следует старшая символьная часть имени, вторая по старшинству символьная часть имени и т.д. Младшая часть имени соответствует конечному узлу сети. В отличие от имен файлов, при записи которых сначала указывается самая старшая составляющая, затем составляющая более низкого уровня и т.д., запись доменного имени начинается с самой младшей составляющей, а заканчивается самой старшей. Составные части домен-

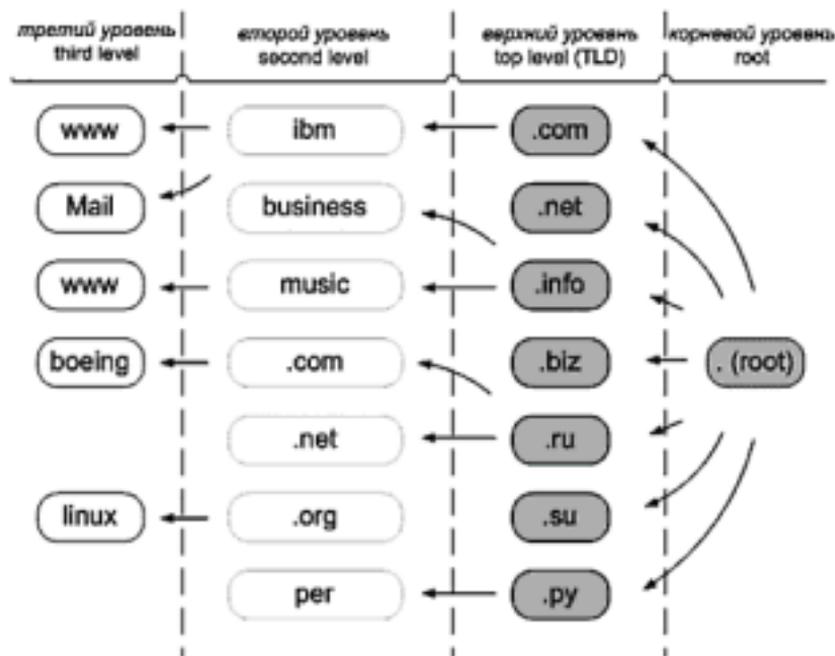


Рис. 3.23. Иерархия доменных имен

ного имени отделяется друг от друга точкой. Например, в имени `partnering.microsoft.com` составляющая `partnering` является именем одного из компьютеров в домене `microsoft.com`.

Разделение имени на части позволяет разделить административную ответственность за назначение уникальных имен между различными людьми или организациями в пределах своего уровня иерархии. Так, для примера, приведенного на рис. 3.23, один человек может нести ответственность за то, чтобы все имена, которые имеют окончание «ш», имели уникальную следующую вниз по иерархии часть. Если этот человек справляется со своими обязанностями, то все имена типа `www.ru`, `mail.mmt.ru` или `m2.zil.mmt.ru` будут отличаться второй по старшинству частью.

Разделение административной ответственности позволяет решить проблему образования уникальных имен без взаимных консультаций между организациями, отвечающими за имена одного уровня иерархии.

Очевидно, что должна существовать одна организация, отвечающая за назначение имен верхнего уровня иерархии.

Совокупность имен, у которых несколько старших составных частей совпадают, образуют домен имен (domain). Например, имена `www.zil.mmt.ru`, `ftp.zil.mmt.ru`, `yandex.ru` и `sl.mgu.ru` входят в домен `ru`, так как все эти имена имеют одну общую старшую часть — имя `ru`. Другим примером является домен `mgu.ru`. Из представленных на рис. 3.23 имен в него входят имена `sl.mgu.ru`, `s2.mgu.ru` и `m.mgu.ru`. Этот домен образуют имена, у которых две старшие части всегда равны `mgu.ru`. Имя `www.mmt.ru` в домен `mgu.ru` не входит, так как имеет отличающуюся составляющую `mmt`.

Термин «домен» очень многозначен, поэтому его нужно трактовать в рамках определенного контекста. Кроме доменов имен стека TCP/IP в компьютерной литературе также часто упоминаются домены Windows NT, домены коллизий и некоторые другие. Общим у всех этих терминов является то, что они описывают некоторое множество компьютеров, обладающее каким-либо определенным свойством.

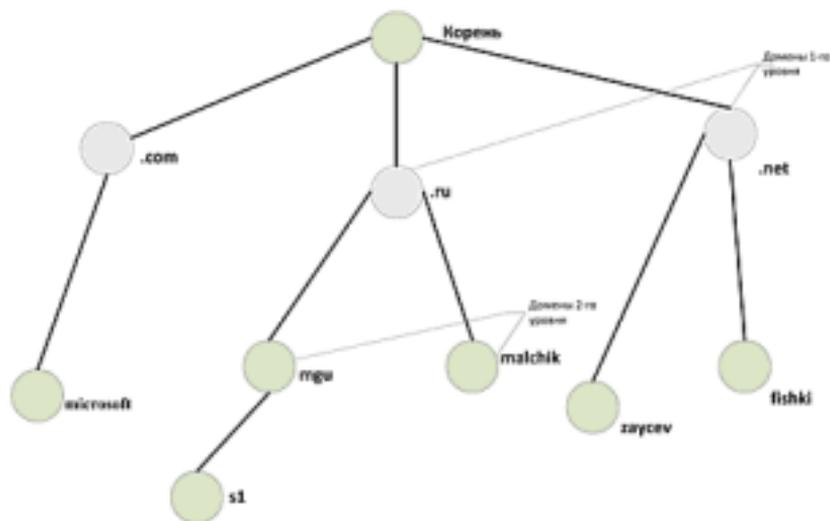


Рис. 3.23. Пространство доменных имен (пример)

Если один домен входит в другой домен как его составная часть, то такой домен могут называть поддоменом (subdomain), хотя название домен за ним также остается. Обычно поддомен называют по имени

той его старшей составляющей, которая отличает его от других поддоменов. Например, поддомен `mmt.ru` обычно называют поддоменом (или доменом) `mmt`. Имя поддомену назначает администратор вышестоящего домена. Хорошей аналогией домена является каталог файловой системы.

Если в каждом домене и поддомене обеспечивается уникальность имен следующего уровня иерархии, то и вся система имен будет состоять из уникальных имен.

По аналогии с файловой системой, в доменной системе имен различают краткие имена, относительные имена и полные доменные имена. Краткое имя — это имя конечного узла сети: хоста или порта маршрутизатора. Краткое имя — это лист дерева имен. Относительное имя — это составное имя, начинающееся с некоторого уровня иерархии, но не самого верхнего. Например, `www.zil` — это относительное имя. Полное доменное имя (fully qualified domain name, FQDN) включает составляющие всех уровней иерархии, начиная от краткого имени и кончая корневой точкой: `www.zu.mmt.ru`.

Необходимо подчеркнуть, что компьютеры входят в домен в соответствии со своими составными именами, при этом они могут иметь совершенно различные IP-адреса, принадлежащие к различным сетям и подсетям. Например, в домен `mgu.ru` могут входить хосты с адресами `132.13.34.15`, `201.22.100.33`, `14.0.0.6`. Доменная система имен реализована в сети «Интернет», но она может работать и как автономная система имен в крупной корпоративной сети, использующей стек TCP/IP, но не связанной с Интернетом.

В Интернете корневой домен управляется центром InterNIC. Домены верхнего уровня назначаются для каждой страны, а также на организационной основе. Имена этих доменов должны следовать международному стандарту ISO 3166. Для обозначения стран используются трехбуквенные и двухбуквенные аббревиатуры, а для различных типов организаций — следующие обозначения:

- com** — коммерческие организации (например, `http://www.nokia.com`);
- edu** — образовательные (например, `mit.edu`);
- gov** — правительственные организации (например, `nsf.gov`);
- org** — некоммерческие организации (например, `fidonet.org`);
- net** — организации, поддерживающие сети (например, `nsf.net`).

Каждый домен администрируется отдельной организацией, которая обычно разбивает свой домен на поддомены и передает функции администрирования этих поддоменов другим организациям. Чтобы получить доменное имя, необходимо зарегистрироваться в какой-либо организации, которой InterNIC делегировал свои полномочия по

распределению имен доменов. В России такой организацией является РосНИИРОС<sup>1</sup>, которая отвечает за делегирование имен поддоменов в домене.

Соответствие между доменными именами и IP-адресами может устанавливаться как средствами локального хоста, так и средствами централизованной службы. На раннем этапе развития Интернета на каждом хосте вручную создавался текстовый файл с известным именем `hosts`. Этот файл состоял из некоторого количества строк, каждая из которых содержала одну пару «IP-адрес — доменное имя», например, `102.54.94.97 — rhino.acme.com`.

По мере роста Интернета файлы `hosts` также росли, и создание масштабируемого решения для разрешения имен стало необходимостью.

Таким решением стала специальная служба — система доменных имен (Domain Name System, DNS). DNS — это централизованная служба, основанная на распределенной базе отображений «доменное имя — IP-адрес». Служба DNS использует в своей работе протокол типа «клиент-сервер». В нем определены DNS-серверы и DNS-клиенты. DNS-серверы поддерживают распределенную базу отображений, а DNS-клиенты обращаются к серверам с запросами о разрешении доменного имени в IP-адрес.

Служба DNS использует текстовые файлы почти такого формата, как и файл `hosts`, и эти файлы администратор также подготавливает вручную. Однако служба DNS опирается на иерархию доменов, и каждый сервер службы DNS хранит только часть имен сети, а не все имена, как это происходит при использовании файлов `hosts`. При росте количества узлов в сети проблема масштабирования решается созданием новых доменов и поддоменов имен и добавлением в службу DNS новых серверов.

Для каждого домена имен создается свой DNS-сервер. Этот сервер может хранить отображения «доменное имя — IP-адрес» для всего домена, включая все его поддомены. Однако при этом решение оказывается плохо масштабируемым, так как при добавлении новых поддоменов нагрузка на этот сервер может превысить его возможности. Чаще сервер домена хранит только имена, которые заканчиваются на следующем ниже уровне иерархии по сравнению с именем домена (аналогично каталогу файловой системы, который содержит записи о файлах и подкаталогах, непосредственно в него «входящих»). Именно при такой организации службы DNS нагрузка по разрешению имен распределяется

<sup>1</sup> Российский научно-исследовательский институт развития общественных сетей (РосНИИРОС) — российский институт, осуществляющий свою деятельность в рамках двух основных направлений:

- развитие компьютерных сетей для организаций науки и образования;
- развитие базовых элементов инфраструктуры российского сегмента сети «Интернет».

более-менее равномерно между всеми DNS-серверами сети. Например, в первом случае DNS-сервер домена mmt.ru будет хранить отображения для всех имен, заканчивающихся на mmt.ru: www.zil.mmt.ru, ftp.zil.mmt.ru, mail.mmt.ru и т.д. Во втором случае этот сервер хранит отображения только имен типа mail.mmt.ru, www.mmt.ru, а все остальные отображения должны храниться на DNS-сервере поддомена zil.

Каждый DNS-сервер кроме таблицы отображений имен содержит ссылки на DNS-серверы своих поддоменов. Эти ссылки связывают отдельные DNS-серверы в единую службу DNS. Ссылки представляют собой IP-адреса соответствующих серверов. Для обслуживания корневого домена выделено несколько дублирующих друг друга DNS-серверов, IP-адреса которых являются широко известными (их можно узнать, например, в InterNIC).

Процедура разрешения DNS-имени во многом аналогична процедуре поиска файловой системой адреса файла по его символьному имени. Действительно, в обоих случаях составное имя отражает иерархическую структуру организации соответствующих справочников — каталогов файлов или таблиц DNS. Здесь домен и доменный DNS-сервер являются аналогом каталога файловой системы. Для доменных имен так же, как и для символьных имен файлов, характерна независимость именования от физического местоположения.

Процедура поиска адреса файла по символьному имени заключается в последовательном просмотре каталогов, начиная с корневого. При этом предварительно проверяется кэш и текущий каталог. Для определения IP-адреса по доменному имени также необходимо просмотреть все DNS-серверы, обслуживающие цепочку поддоменов, входящих в имя хоста, начиная с корневого домена. Существенным же отличием является то, что файловая система расположена на одном компьютере, а служба DNS по своей природе является распределенной.

Существуют две основные схемы разрешения DNS-имен. В первом варианте работу по поиску IP-адреса координирует DNS-клиент:

- DNS-клиент обращается к корневому DNS-серверу с указанием полного доменного имени;
- DNS-сервер отвечает, указывая адрес следующего DNS-сервера, обслуживающего домен верхнего уровня, заданный в старшей части запрошенного имени;
- DNS-клиент делает запрос следующего DNS-сервера, который отсылает его к DNS-серверу нужного поддомена, и т.д., пока не будет найден DNS-сервер, в котором хранится соответствие запрошенного имени IP-адресу. Этот сервер дает окончательный ответ клиенту.

Такая схема взаимодействия называется нерекурсивной, или итеративной, когда клиент сам итеративно выполняет последовательность запросов к разным серверам имен. Так как эта схема загружает клиента достаточно сложной работой, то она применяется редко.

Во втором варианте реализуется рекурсивная процедура:

- DNS-клиент запрашивает локальный DNS-сервер, т.е. тот сервер, который обслуживает поддомен, к которому принадлежит имя клиента;
- если локальный DNS-сервер знает ответ, то он сразу же возвращает его клиенту; это может соответствовать случаю, когда запрошенное имя входит в тот же поддомен, что и имя клиента, а также может соответствовать случаю, когда сервер уже узнавал данное соответствие для другого клиента и сохранил его в своем кэше;
- если же локальный сервер не знает ответ, то он выполняет итеративные запросы к корневому серверу и т.д. точно так же, как это делал клиент в первом варианте; получив ответ, он передает его клиенту, который все это время просто ждал его от своего локального DNS-сервера.

В этой схеме клиент перепоручает работу своему серверу, поэтому схема называется косвенной, или рекурсивной. Практически все DNS-клиенты используют рекурсивную процедуру.

Для ускорения поиска IP-адресов DNS-серверы широко применяют процедуру кэширования проходящих через них ответов. Чтобы служба DNS могла оперативно обрабатывать изменения, происходящие в сети, ответы кэшируются на определенное время — обычно от нескольких часов до нескольких дней.

Интернет 2.0 — это эволюция существующего Интернета, его продолжение, дарующее пользователям Интернета возможности быть свободными, ответственными за свою свободу и самостоятельно управлять своим временем. Кроме того, Интернет 2.0 позволяет активно защищать свои права на свободу и управление временем, а также соблюдать права других людей. Одной из таких частных можно назвать авторское право.

Цели и задачи Интернета 2.0.

1. Свобода и безопасность пользователей Интернета, защита личной жизни и личной информации от посягательств других лиц как средствами программного обеспечения, так и личным участием (например, защита своего авторского права в кругу знакомых лиц).

2. Управление своим временем самостоятельно: вы следуете своим жизненным интересам — другие люди не нарушают их, и вы не на-

ланных аппаратных комплексах — суперкомпьютерах (не путать с ЦОД и ДЦ). Суперкомпьютеры (СК) — являются однотипным, блочно-модульным решением; набор модулей в соответствии с предназначением СК и предоставляемым сервисам. СК рассчитаны на онлайн-нагрузку в несколько млн человек (например, СК города, области, страны), но могут быть и меньше, например, тематические сообщества: спортивные (футбол, хоккей и т.п.); медиа (театры онлайн, цифровое ТВ в дом; авторское ТВ); игровые (обмен модов, форумы сообществ на едином СК, сайты профильных организаций, совпадающих с тематикой сообществ); транспортно-пассажирские, туристические (система управления автотранспортом городского, ж/д и авиасообщений); и т.д. Данные и все другие тематические сообщества и сервисы являются инструментами социально-экономических преобразований, основанных на открытости, ясности и свободе волеизъявлений людей, с возможностью активно защищать свои права и свободу и своевременно сигнализировать о нарушении их прав и свобод всему тематическому сообществу.

Виртуальная ОС объединяет сервисы и распределяет нагрузку на аппаратную часть: в случае аварии задействует аварийные модули, а в случае простоя отключает невостребованные; пользователи при этом продолжают получать сервисы в онлайн режиме. При необходимости универсальная виртуальная ОС может задействовать СК другого города-региона, с согласия пользователей (например, режим день-ночь: два города Токио в Японии и Марсель во Франции, когда одни спят, другие используют их СК, потом поменялись). Приложения для виртуальной ОС являются виртуализированными и содержатся на одном сервере — вирторепозитории. Автору приложения достаточно 1 раз обновить версию и она сразу станет доступна на всех СК и всех виртуальных ОС. Скорость распространения (важно потребителям), доступность моментальной публикации во всем Интернете (важно авторам), т.е. не надо никакой рекламы, ваши приложения сразу будут доступны всем потребителям и вы будете находиться на равных со всеми производителями ПО. Так сказать, Интернет 2.0 одинаков для всех и все перед ним равны и друг перед другом тоже. А также понадобится сообщество защиты авторских прав. Личные данные пользователей виртуальной ОС будут надежно защищены и зашифрованы в базах данных особым образом. При этом для пользователя его информация будет видна, а всем остальным — только набор зашифрованных данных.

Личное пространство (типа учетной записи) будет уникальным и в нем будет прописана полная информация о человеке: паспортные данные, где живет и т.д. Свое личное пространство пользователи будут сами определять, на каком СК хранить; их нужно хранить на разных

СК. При этом СК могут перемещать личное пространство пользователя по его первому требованию. Свои медиафайлы и файлы больших размеров пользователи могут хранить надежно и отдельно, также в зашифрованном виде, не допуская туда любых посторонних глаз. В случае аморального поведения человека на форуме или сообществе можно привлечь его к ответственности и при необходимости «забанить» по личному пространству, а не по IP (применимо и к вопросам защиты авторских прав).

Ответственность и восстановление порядка — не единственная функция личного пространства и взаимодействия. Изюминкой Интернета 2.0 является то, что все ссылки указывают на их автора, где бы он ни был и где бы он ни участвовал. И эта особенность без труда поможет вам отыскать автора и связаться с ним: например, отправить ему сообщение, и оно обязательно дойдет до адресата. Кстати, электронная почта, скорее всего, перестанет быть востребованной: общаться станет проще с помощью прямых взаимосвязей. Теперь при заходе на сайт не придется создавать аккаунты — эту ситуацию уладят между собой задействованные СК и виртуальные ОС, пользователь же выберет шаблон участия для таких сайтов и определит зал участия в своем личном пространстве. Также в личном пространстве пользователь будет хранить свою личную информацию, недоступную для просмотра другими лицами, в том числе техперсоналом, и соответственно силовым структурам, даже если захотят, — полное невмешательство в личную жизнь граждан. В зале участия будет находиться изолированная текущая онлайн сессия участия в сообществе, по выходу из нее ОЗУ СК будет высвобождено: нет большой необходимости в антивирусных средствах, хотя антивирусную защиту СК осуществлять будут именно средствами антивирусной защиты внутри себя и сертификатами безопасности между собой и (или) действиями пользователя. Залы участия будут разными: некоторые для просмотра сайтов общения, некоторые для личного творчества-работы; некоторые для работы со своим медиахранилищем; другие предназначения.

Для доступа в Интернет люди будут использовать как свой ПК, так и универсальные точки доступа (УТД). УТД чем-то похожа на тонкого клиента, только возможности гораздо шире. УТД будет поставлять сервисы: телевидение, телефонию (в том числе видеотелефон, сотовая связь через IP), доступ в интернет, подключение периферийных устройств и устройств хранения информации, Wi-Fi; содержать в себе тонкого клиента и его функции. Вероятно, будет иметь модульное исполнение (способность расширять возможности УТД вышеназванными модулями).

Использование УТД совместно с СК (суперкомпьютерами) позволит избежать информационно-технологического удара в сознание рядовых пользователей (технологический прогресс изобилует техническими терминами, употребляемыми специалистами и которые совсем не нужны обычным людям, уже владеющими другими, своими спецтерминами) и позволит технологическому прогрессу идти на стороне СК, а не на стороне ПК или ЦОД (в том числе ДЦ, которые, скорее всего, будут преобразованы в СК) и не «нагружать» людей ненужной им технологической информацией. А также это возможность не покупать более скоростной компьютер и долго не выбрасывать на свалку свой домашний компьютер, тем самым сокращая мусорные отходы общества потребления и экономя на электричестве (не нужно будет строить еще одну атомную станцию), повышая эффективность и гибкость применения информационных технологий.

Бесплатный доступ в Интернет 2.0, использование пропускной способности инфраструктуры Интернет-провайдеров на 100%; бесплатный доступ к услугам сотовой и наземной связи. Но вот за предоставляемые услуги плата все-таки будет взиматься и вся статистика доходов и расходов этих компаний должна быть в свободном открытом доступе, чтобы человек знал, что за услуги ему предоставляют, и самостоятельно осуществлял их выбор в соответствии со своими жизненными интересами, а также имел возможность «забанить» компанию, нарушающую его права и свободу, известив об этом все тематическое сообщество и при необходимости инициализировал электронное голосование на изъятие лицензии на право деятельности этой компании. То есть это возможность регулировать Интернет-сервисы свободно и самостоятельно, без участия государства, повышая оперативность реагирования и отсекая постороннюю заинтересованность, делая свободный выбор чистым и быстрым.

В Интернете 2.0 трафик<sup>1</sup> и биллинг<sup>2</sup> становятся не актуальны: любой человек может подключиться в сеть «Интернет» и просматривать ее со-

<sup>1</sup> Трафик (от англ. traffic — движение, транспорт, торговля). В данном случае речь идет о так называемом **сетевом трафике**. Сетевой трафик (от англ. Traffic — движение, грузосборот) — объем информации, передаваемой через компьютерную сеть за определенный период времени. Количество трафика измеряется как в гигабитах, так и в битах, байтах и их производных: килобайт (КБ), мегабайт (МБ) и т.д.

<sup>2</sup> Биллинг (понятие применительно исключительно для электросвязи) — комплекс процессов и решений на предприятиях связи, ответственных за сбор информации об использовании телекоммуникационных услуг, их тарификацию, выставление счетов абонентам, обработку платежей. Биллинговая система — прикладное программное обеспечение поддержки бизнес-процессов биллинга. Биллинг — важнейший компонент деятельности любого коммерческого оператора связи, вне зависимости от вида телекоммуникаций: операторы фиксированной и мобильной связи, Интернет-телефония, виртуальные операторы, Интернет-провайдеры, операторы транзитного цифрового трафика, провайдеры цифрового телевидения — не могут существовать без биллинга, благодаря которому выставляются счета потребителям их услуг и обеспечивается экономическая составляющая их деятельности.

вершено бесплатно, платными будут являться только сервисы. Конечно, это можно и нужно было бы правильнее назвать условно-бесплатным Интернетом, но об этом поподробнее. Дело в том, что, платя за сервис, вы еще оплачиваете аппаратно-программную часть суперкомпьютера (СК), реализующего тот или иной сервис, доступный пользователю через сеть, а также содержание собственно сети (техобслуживание, электричество, хозяйственные расходы компаний, предоставляющих сервисы). Любой пользователь сети Интернет 2.0 выбирает тот или иной сервис сети, «заказывает музыку», а компания, ее реализующая (так называемая сервис-компания), «исполняет». При этом всю ситуацию пользователь совершенно честно оплачивает и делает это совместно с другими пользователями однотипных сервисов, и совершенно обязательно это пользователи города или страны.

Пользователь выбирает нужный сетевой сервис<sup>1</sup> в любой точке планеты, потребляет и оплачивает его. Далее сервис-компания, предоставившая сервис, расходует средства на содержание СК и отчитывается перед ним о произведенных расходах. Помимо этого сервис-компания обязана оплатить и отчитаться по еще одной статье затрат: содержание сети Интернет 2.0. Для этого им в помощь идут компании-провайдеры. Компании-провайдеры осуществляют свою деятельность по прокладке Интернета 2.0 и техническому обслуживанию, а также бесперебойному функционированию сети в ее определенной зоне ответственности. Компании-провайдеры обобщают свои затраты и предлагают принять их к оплате сервис-компаниям. При этом пользователь лично указывает сервис-компаниям, какую зону ответственности они должны принять к оплате и оплатить от его имени. В принципе, как вариант, сервис-компания может предоставить пользователю всю статистику Интернет-сетииспользования, и он сам сможет оплатить компаниям провайдерам напрямую. Технические возможности для определения суммы оплаты также будут.

<sup>1</sup> Сетевой сервис — т.е. сервис (служба), предоставляемый сетью. Взаимодействие компьютеров между собой, а также с другим активным сетевым оборудованием в TCP/IP-сетях организовано на основе использования сетевых служб, которые обеспечиваются специальными процессами сетевой операционной системы (ОС) — демонами в UNIX-подобных ОС, службами в ОС семейства Windows и т.п. Службы ОС Windows (от англ. Windows Service, службы) — приложения, автоматически (если настроено) запускаемые системой при запуске Windows и выполняющиеся вне зависимости от статуса пользователя. Имеет общие черты с концепцией демонов в Unix. Демон (от англ. daemon) — в системах класса UNIX — служба, работающая в фоновом режиме без прямого общения с пользователем. Демоны обычно запускаются во время загрузки системы. Примеры сетевых сервисов: серверы сетевых протоколов (HTTP, FTP, электронная почта и др.), управление оборудованием, поддержка очереди печати, управление выполнением заданий по расписанию и т.д. В техническом смысле демоном считается процесс, который не имеет управляющего терминала. Чаще всего (но не обязательно) предком демона является init — жерновой процесс UNIX.

Варианты оплаты будут определять сами пользователи посредством голосования (всегда гласного). Например, правила «общего котла» — когда все платят одинаково или с использованием индивидуального маршрута; или и то и другое, дополняя друг друга. При этом право свободного выбора всегда остается за пользователем, и правила устанавливает именно он, а не любые компании. Кроме того развитие СК, к которому пользователь приписывает свой канал участия (типа аккаунта, учетной записи), также происходит с его подачи: если участников становится больше, то инициализируется голосование на добавление соответствующего модуля, а сервис-компания просто исполняет коллективные указания. В том случае, если сервис-компания по каким-либо причинам проигнорировала коллективные решения (а «музыку заказывают» пользователи сервиса), то любой пользователь вправе отозвать лицензию (гласное голосование) у этой компании или не пользоваться ее услугами совсем.

И самое интересное, что Интернет-компании будут работать без прибыли, т.е. не закладывать прибыль в свою деятельность, а будут содержаться целиком и полностью на средства Интернет-пользователей, реально оказывая Интернет-сервисы.

Одной из основных проблем современного Интернета считается малое число возможных IP-адресов: всего их может быть около 4 млрд. На первый взгляд этого более чем достаточно. Но на самом деле ситуация, когда IP-адресов будет не хватать на всех, может наступить в самом ближайшем времени — слишком много электронных устройств имеют сегодня доступ к сети. Другой важной проблемой является низкая производительность узлов, занимающихся передачей информации, например, маршрутизаторов. Дело даже не в самих устройствах, а в их предназначении. То есть вычислительные процессы, заложенные в основу Интернета, можно оптимизировать.

Следующим недостатком является неподготовленность глобальной сети к передаче по ее каналам больших объемов данных в реальном времени, например, видео- и аудиоинформации. Кроме того сюда же можно отнести невозможность широкоэвещательной передачи. Это очень сильно тормозит использование Интернета в качестве канала телевидения. Ну и, наконец, последний серьезный недостаток кроется в отсутствии «защитных в протоколы» механизмов безопасности.

Попытки решения этих задач и привели как раз к возникновению проекта Интернет 2.0. В основу сети Интернет 2.0 положен новый протокол IPv6. Его разработки начались еще в 1992 г. В 1996 году ряд крупных телекоммуникационных корпораций совместно с правительством США организовали проект Интернет 2.0, в задачи которого вхо-

дила практическая реализация сети нового поколения. И уже в 1998 г. первый «отрезок» этой сети был запущен в эксплуатацию. С его помощью были соединены более 100 самых крупных университетов, разбросанных по всей территории США. С тех пор Интернет 2.0 постоянно развивается. Причем в 2003 г. многие компании, занимающиеся производством телекоммуникационного оборудования, начали массовый выпуск устройств с поддержкой протокола IPv6.

Первоначально Интернет 2.0 использовался для проведения видеоконференций, совещаний и научных симпозиумов. Следующим шагом стали интерактивные лекции, т.е. телемосты между аудиториями различных университетов, в одном из которых находился преподаватель, а в другом — студенты. Фактически это позволяет профессорам проводить занятия в различных учебных заведениях. Причем может быть даже такое, что преподаватель читает лекцию сразу же для нескольких учебных групп из разных концов страны. Сегодня в список потенциальных возможностей Интернет 2.0 входит удаленное управление телескопами обсерватории и обратная передача полученной ими информации, дистанционное проведение различных экспериментов и т.п. Не за горами тот момент, когда с помощью сети Интернет 2.0 будут проводиться удаленные операции, в которых пациент лежит на столе в одном городе, а хирург находится в другом. Кроме того область возможного применения очень сильно вырастет с коммерциализацией Интернет 2.0. Здесь наверняка появятся виртуальные кинотеатры и альтернативные каналы телевидения, различные онлайн-игры и интерактивные шоу, уникальные торговые витрины и многое другое.

Протокол IPv6 существенно отличается от своего предшественника. Самым заметным различием стала система адресации. Теперь IP-адреса состоят не из 4, а из 16 байт. Кроме того многие корпоративные локальные сети могут отказаться от принципа выхода в глобальную сеть через один шлюзовой компьютер, что позволит каждому сотруднику иметь канал с более высокой скоростью связи. Другие изменения коснулись формата передаваемых пакетов, а также алгоритмов их обработки. Так, в IPv6 удалось избавиться от фрагментации пакетов. Для этого два компьютера, устанавливающие связь друг с другом, перед передачей данных определяют максимально допустимый размер пакета (он зависит от маршрутизаторов на пути между ними) и в дальнейшем используют именно его. Кроме того в IPv6 было внесено еще несколько подобных изменений, оптимизирующих скорость передачи информации.

При создании Интернет 2.0 особое внимание было оказано мультимедиа-технологии. В частности, в IPv6 была реализована поддержка мультикастинга, т.е. широковещательной передачи данных. Это можно

назвать знаковым отличием сети Интернет 2.0 от обычного Интернета. В привычной нам глобальной сети были радиостанции и телевизионные каналы. Однако они никогда не оказывались достаточно массовыми, чтобы сравниться с обычными, традиционными СМИ. Причем проблема заключалась именно в технической реализации радио- и телевидения. Ведь для отправки, например, 128-битного потока сервер должен был установить канал связи с каждым клиентом и передавать всем подключившимся абсолютно одинаковые данные. В сети же Интернет 2.0 данные могут по одному каналу передаваться сразу целой группе абонентов. Так что мультикастинг вместе с некоторыми другими особенностями протокола IPv6 позволяет вести речь о полноценном качественном теле- и радиовещании, полноценных телеконференциях, нормальной телефонии (без пропадания слов, «вканалья» и прочих прелестей VoIP).

Но Интернет 2.0 отличается от простого Интернета не только новым протоколом IPv6. Это совершенно другая сеть с другими физическими каналами связи. И основой этой сети послужила специально проложенная магистраль Abilence (рис. 3.24) с пропускной способностью 10 Гбит/сек. Причем эта магистраль практически не имеет «узких» мест, поскольку все сетевое оборудование, обслуживающее ее, действительно способно работать с такими объемами информации. Подключение к Интернет 2.0 осуществляется через специальные точки присутствия — *gigapops*. Стоит отметить, что минимальная скорость подключения составляет 100 Мбит/с<sup>1</sup>.

Первой компанией в России, которая заговорила о практической реализации Интернет 2.0, стал универсальный оператор связи «Корбина Телеком»<sup>2</sup>. В конце прошлого года представители этой фирмы провели впечатляющую демонстрацию возможностей нового поколения глобальной сети. В частности, специалисты компании смогли передать 12 Гб информации по новому каналу связи за 19 мин, причем по тому же каналу все это время транслировалось качественная видеопрограмма.

Сетевую основу Интернета 2.0 на текущий момент составляют две высокопроизводительные магистрали. Первая из них, vBNS (*very high speed Backbone Network Service*), создана по контракту с корпорацией MCI WorldCom. Оптические каналы, используемые для этой сети, обеспечивают пропускную способность в 622 Мбит/с. Второй магистралью проекта стала сеть под названием Abilene, созданная на основе национальной оптической сети Qwest с использованием технологий Cisco Systems и Nortel (Northern Telecom). Первоначальная пропускная способ-

<sup>1</sup> Для сети Интернет этот показатель составлял значение 14,4 Кбит/с.

<sup>2</sup> «Корбина Телеком» — российская телекоммуникационная компания. Штаб-квартира — в Москве. В настоящий момент полностью принадлежит компании «ВымпелКом». Образована в конце 1995 г. Данное упоминание не является рекламой.

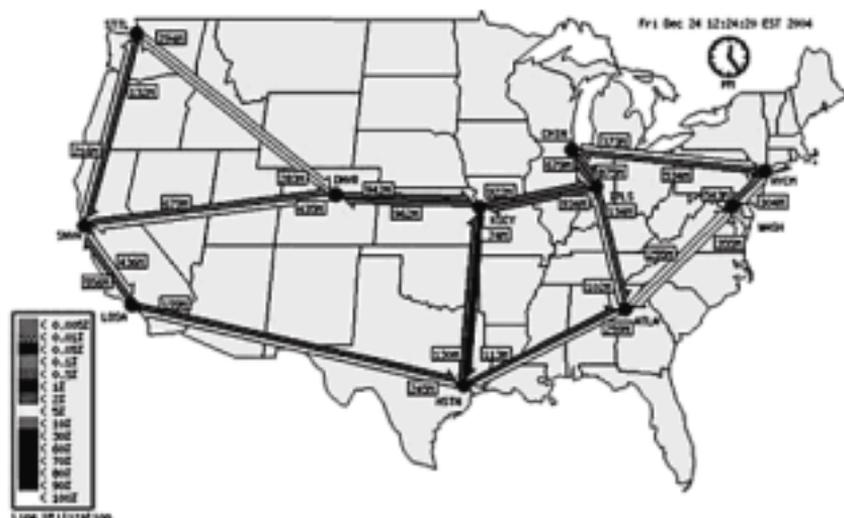


Рис. 3.24. Первая сеть, реализовавшая технологию и правила Интернет 2.0 (напоминает всем известную сеть ARPANET<sup>1</sup>, а по сути ею и является)

ность этой магистрали составляла 2,5 Гбит/с. С прошлого года ее пропускная способность доведена до 10 Гбит/с. А к концу 2006 г. планируется обеспечить каждый узел полосой 100 Гбит. Но развитие магистралей на этом не заканчивается. Консорциум Интернета 2.0 принял участие в не менее амбициозном проекте — строительстве National LambdaRail, первой трансконтинентальной Ethernet-сети. National LambdaRail — это оптическая сеть, построенная по технологии DWDM (Dense Wavelength Division Multiplexing). По одному волокну одновременно можно передавать от 32 до 40 каналов, каждому из которых соответствует определенная длина волны (отсюда в названии появилась Lambda). Пропускная способность каждого канала — 10 Гбит/с.

Для обеспечения такой пропускной способности необходимо и соответствующее оборудование. Этими проблемами занимаются такие известные компании, как Cisco, Novell и др. Оптическое волокно пока

<sup>1</sup> ARPANET (от англ. Advanced Research Projects Agency Network) — компьютерная сеть, созданная в 1969 г. в США Агентством Министерства обороны США по перспективным исследованиям (ARPA) и являвшаяся прототипом сети Интернет. 1 января 1983 г. она стала первой в мире сетью, перешедшей на маршрутизацию пакетов данных. В качестве маршрутизируемого протокола использовался TCP/IP, который и по сей день является основным протоколом передачи данных в сети Интернет. ARPANET прекратила свое существование в июне 1990 г. Более подробно рассматривалось в предыдущем параграфе.

единственное из имеющихся средств распространения информации по средствам связи, обеспечивая необходимые скорости<sup>1</sup>. В отличие от всех предыдущих вариантов Ethernet-сетей 10-гигабитовая технология полностью основана на оптической среде передачи данных (до недавнего времени решения, основанного на медном кабеле, не было). Для каналов доступа и протяженных сетей наилучшим выбором считается одномодовое волокно<sup>2</sup>. Для относительно коротких каналов связи решения могут быть различными.

IPv6 (от англ. Internet Protocol version 6) — новая версия протокола IP, призванная решить проблемы, с которыми столкнулась предыдущая версия (IPv4) при ее использовании в Интернете, за счет использования длины адреса 128 бит вместо 32. В настоящее время протокол IPv6 уже используется в нескольких тысячах сетей по всему миру (более 9000 сетей на май 2012 г.), но пока еще не получил столь широкого распространения в Интернете, как IPv4. В России используется почти исключительно в тестовом режиме некоторыми операторами связи, а также регистраторами доменов для работы DNS-серверов. Протокол был разработан IETF.

После того как адресное пространство в IPv4 закончится<sup>3</sup>, два стека протоколов — IPv6 и IPv4 — будут использоваться параллельно (от англ. dual stack), с постепенным увеличением доли трафика IPv6 по сравнению с IPv4. Такая ситуация станет возможной из-за наличия

<sup>1</sup> Потенциальные ресурсы оптического волокна (т.е. реально достижимая скорость передачи информации) на длине волны источника излучения  $\lambda = 1310$  нм соответствует  $34 \times 10^{18}$  бит/с.

<sup>2</sup> Оптическое волокно — нить из оптически прозрачного материала (стекла, пластика), используемая для переноса света внутри себя посредством полного внутреннего отражения. Стекловолоконные оптические волокна делаются из кварцевого стекла и используются для задачи передачи данных. Оптическое волокно, как правило, имеет круглое сечение и состоит из двух частей — сердцевины и оболочки. Для обеспечения полного внутреннего отражения абсолютный показатель преломления сердцевины несколько выше показателя преломления оболочки. Луч света, направленный в сердцевину, будет распространяться по ней. Оптические волокна, используемые в телекоммуникациях, как правило, имеют диаметр  $125 \pm 1$  микрон. Диаметр сердцевины может отличаться в зависимости от типа волокна и национальных стандартов. Диаметр сердцевины одномодовых волокон составляет от 7 до 10 микрон. Благодаря малому диаметру достигается передача по волокну лишь одной моды электромагнитного излучения, за счет чего исключается влияние дисперсионных искажений. В настоящее время практически все производимые волокна являются одномодовыми. Такие волокна обеспечивают передачу высокоскоростных данных на очень большие расстояния 100 км и выше без использования усилительных элементов в составе линии связи.

<sup>3</sup> Истощение IPv4-адресов — истощение запаса нераспределенных адресов протокола IPv4. Адресное пространство глобально управляется американской некоммерческой организацией IANA, а также пятью региональными Интернет-регистраторами, ответственными за назначение IP-адресов конечным пользователям на определенных территориях, и локальными Интернет-регистраторами, такими как Интернет-провайдеры. IPv4 позволяет использовать около 4,23 млрд адресов, и часть из них была распределена IANA региональным Интернет-регистраторам блоками примерно по 16,8 млн адресов (с учетом использования CIDR). В феврале 2011 г. IANA выделила региональным Интернет-регистраторам последние пять оставшихся блоков (8 из своего адресного пространства). APNIC является первым региональным Интернет-регистратором, выделенные IP-адреса которого закончились. Это произошло 15 апреля 2011 г.

огромного количества устройств, в том числе устаревших, не поддерживающих IPv6 и требующих специального преобразования для работы с устройствами, использующими только IPv6.

В конце 80-х гг. XX в. стала очевидна необходимость разработки способов сохранения адресного пространства Интернета. В начале 1990-х, несмотря на внедрение бесклассовой адресации, стало ясно, что этого недостаточно для предотвращения исчерпания адресов, и необходимы дальнейшие изменения инфраструктуры Интернета. К началу 1992 г. появилось несколько предложений, и к концу 1992 г. IETF<sup>1</sup> объявила конкурс для рабочих групп на создание Интернет-протокола следующего поколения (от англ. IP Next Generation — IPng). 25 июля 1994 г. IETF утвердила модель IPng, с образованием нескольких рабочих групп IPng. К 1996 году была выпущена серия RFC, определяющих Интернет-протокол версии 6, начиная с RFC 1883. IETF назначила новому протоколу версию 6, так как версия 5 была ранее назначена экспериментальному протоколу, предназначенному для передачи видео и аудио.

Оценки времени полного исчерпания IPv4 адресов различались в 2000-х гг., но в настоящее время все оценки сходятся на 2012 г. (рис. 3.25). В 2003 году директор APNIC Пол Уилсон (от англ. Paul Wilson) заявил, что, основываясь на темпах развертывания сети Интернет того времени, свободного адресного пространства хватит на одно — два десятилетия. В сентябре 2005 г. Cisco Systems предположила, что пула доступных адресов хватит на 4—5 лет. 3 февраля 2011 г. агентство IANA<sup>2</sup> распределило последние 5 блоков IPv4 региональным Ин-

<sup>1</sup> Инженерный совет Интернета (от англ. Internet Engineering Task Force, IETF) — открытое международное сообщество проектировщиков, ученых, сетевых операторов и провайдеров, созданное IAB в 1986 г. и занимающееся развитием протоколов и архитектуры Интернета. Вся техническая работа осуществляется в рабочих группах IETF, занимающихся конкретной тематикой (например, вопросами маршрутизации, транспорта данных, безопасности и т.д.). Работа в основном ведется через почтовые рассылки, но трижды в году проводятся собрания IETF. Результаты деятельности рабочих групп оформляются в виде рабочих проектов (от англ. Internet drafts), которые затем используются ISOC для кодификации новых стандартов. Задачи IETF (в соответствии с RFC 4677):

- идентификация проблем и предложение решений в технических аспектах организации Интернета;
- разработка спецификаций, стандартов и соглашений по общим архитектурным принципам протоколов Интернет;
- вынесение рекомендаций относительно стандартизации протоколов на рассмотрение Internet Engineering Steering Group (IESG);
- содействие широкому распространению технологий и стандартов, разрабатываемых в Internet Research Task Force (IRTF);
- организация дискуссии для обмена информацией в сообществе Интернета между учеными, разработчиками, пользователями, производителями оборудования и услуг, сетевыми администраторами и т. д.

<sup>2</sup> IANA (от англ. Internet Assigned Numbers Authority — Администрация адресного пространства Интернет) — американская некоммерческая организация, управляющая пространствами IP-адресов, доменов верхнего уровня, а также регистрирующая типы данных MIME и параметры прочих протоколов Интернета. Находится под контролем ICANN.

тернет-регистраторам. Выделение диапазонов адресов региональными службами RIR продолжается, однако, по данным исследований, остатки адресов закончились в августе 2012 г. 14 сентября 2012 г. организация RIPE NCC разослала информационное сообщение, где говорится о начале распределения последнего свободного блока. При этом новые аллокации не могут быть более (1024 адреса), и LIR уже должен иметь аллокацию в IPv6.

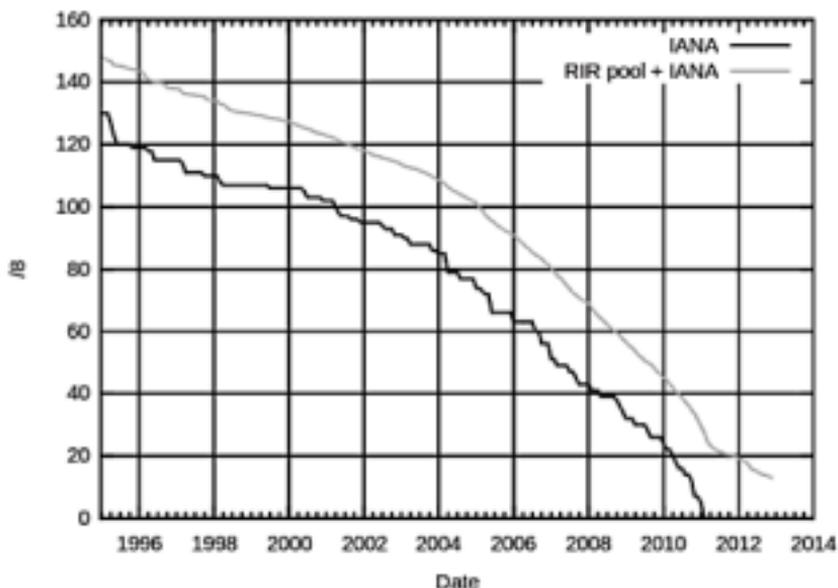


Рис. 3.25. Кривые «исчерпания» IP-адресов

Иногда утверждается, что новый протокол может обеспечить по  $5 \times 10^{28}$  адресов на каждого жителя Земли. Однако такое огромное адресное пространство IPv6 было введено ради иерархичности адресов (это упрощает маршрутизацию), и большая его часть в принципе не будет задействована никогда. Тем не менее увеличенное пространство адресов сделает NAT необязательным. Классическое применение IPv6 (по сети (64 на абонента); используется только unicast-адресация) обеспечит возможность использования более 300 млн IP-адресов на каждого жителя Земли. Из IPv6 убраны функции, усложняющие работу маршрутизаторов.

Что это дает рядовому пользователю помимо 100 Мбит/с минимальной скорости? Пожалуй, что не только это. Если задуматься, какого уровня достигла электронная коммерция сегодня, используя технологии 30-летней давности? Сегодняшний Интернет заполнен огромным количеством ресурсов и сервисов. А теперь стоит представить — чего можно достичь, если **конечный пользователь** получит в свое распоряжение скорость в 100 Мбит/с?

## Контрольные вопросы

1. Что такое компьютерная сеть?
2. Какое оборудование применяется в сетях ЭВМ, и каково его назначение?
3. Какие типы кабельных систем бывают?
4. Какая технология сетей нашла самое широкое применение в ЛВС?
5. Что дает предприятию использование сетей ЭВМ?
6. Что такое рабочая станция?
7. Что такое сервер?
8. Как осуществляется управление доступом к сетевым ресурсам?
9. Что такое протокол?
10. Каково место протокола TCP/IP в ЭМВОС (OSI)?

# ТЕХНОЛОГИИ ПОИСКОВЫХ СИСТЕМ

Решение задачи поиска документов в СЕД и технологиях электронного документооборота является крайне важной задачей. Помимо этого поиск информации в сети «Интернет», используемой в процессе управления, также является необходимой задачей, поэтому далее рассмотрим особенности построения поисковых машин.

## 4.1. Документальные системы: информационно-поисковый язык, система индексирования, технология обработки данных, поисковый аппарат, критерии оценки документальных систем

Функционирование современных ИПС основано на двух предположениях:

- 1) документы, необходимые пользователю, объединены наличием некоторого признака или комбинации признаков;
- 2) пользователь способен указать этот признак.

Оба эти предположения на практике не выполняются, и можно говорить только о вероятности их выполнения. Поэтому процесс поиска информации обычно представляет собой последовательность шагов, приводящих при посредстве системы к некоторому результату и позволяющих оценить его полноту. При этом поведение пользователя, как организующее начало управления процессом поиска, мотивируется не только информационной потребностью, но и разнообразием стратегий, технологий и средств, предоставляемых системой.

Пользователь обычно не имеет исчерпывающих знаний об информационном содержании ресурса, в котором проводит поиск. Оценить адекватность выражения запроса, как и полноту получаемого результата, он может, отыскав дополнительные сведения или так организовав процесс, чтобы часть результатов поиска могла использоваться для подтверждения или отрицания адекватности другой части. В то же время для пользователей-профессионалов характерна устойчивость тематического профиля. Когда они являются «информационно ориентированными», то им свойственно желание и способность организовать информационное пространство проблемы. Это означает, что пользователь создает по существу новый, «самостоятельный» проблемно ориентированный, индивидуально обновляемый и пополняемый ИР, включающий помимо подборок документов также и метаинформацию, например, словари специальной терминологии, классификаторы предметных областей, описания ресурсов и т.д.

Особенность работы пользователя в режиме «самообслуживания», в контексте задачи автоматизации совокупной деятельности означает, что система должна представлять среду, обеспечивающую поддержку функций потребителя по обработке найденной информации, а также традиционно относящихся к функциям информационного посредника (интерпретация запроса, его перевод на информационно-поисковый язык, выбор ИР, автоматизированный поиск и ручной отбор материалов), но также и такие «обеспечивающие» функции, как: структурирование информационной потребности, лексическая адаптация запроса, оценка, систематизация и обработка результатов поиска, причем на уровне как отдельного документа, так и информационных ресурсов в целом. Технические возможности, которыми располагает пользователь, позволяют ему создавать информационный ресурс — формировать массивы, систематизировать и создавать внешние представления их содержания для собственного или внешнего использования.

Автоматизированные ИПС (АИПС) используют компьютерные программно-технические средства и технологии и предназначаются для нахождения и выдачи пользователям информации по заданным критериям. Определяющими для понимания методов автоматизации поиска являются два следующих фактора:

- 1) сравниваются не сами объекты, а описания — так называемые поисковые образы;
- 2) сам процесс является сложным (составным и не одноактным) и обычно реализуется последовательностью операций.

Данные в АИПС вводятся на основе специально разрабатываемых форматов ввода. Все сведения об одном объекте в ИПС представляют-

ся в виде систематизированных данных, образующих одну строку таблицы и называются записью. При этом если ИПС представляет электронный каталог библиотеки, то любое библиографическое описание (БО) документа в нем — это одна запись, состоящая из полей, равных количеству элементов БО. Совокупность записей образует БД, которая, как правило, хранится в одном файле. Совокупность БД, объединенных одной СУБД, образует банк данных.

Поскольку АИПС инструмент, используемый человеком при поиске (а не интеллектуальный автомат для поиска информации — готовых решений задач основной деятельности), эффективность ее использования зависит от того, насколько хорошо человек знает природу операционных объектов и свойства инструмента, посредством которого он работает с этими объектами.

Информационный поиск подразумевает использование определенных стратегий, методов, механизмов и средств. Поведение пользователя, осуществляющего управление процессом поиска, определяется не только информационной потребностью, но и инструментальным разнообразием системы — технологиями и средствами, предоставляемыми системой.

Стратегия поиска — общий план (концепция, предпочтение, установка) поведения системы или пользователя для выражения и удовлетворения информационной потребности пользователя, обусловленный как характером цели и видом поиска, так и системными «стратегическими» решениями — архитектурой БД, методами и средствами поиска в конкретной АИПС. Выбор стратегии в общем случае является оптимизационной задачей. На практике в значительной степени он определяется искусством достижения компромисса между практическими потребностями и возможностями имеющихся средств.

Метод поиска — совокупность моделей и алгоритмов реализации отдельных технологических этапов: построения поискового образа запроса (ПОЗ), отбора документов (сопоставление поисковых образов запросов и документов), расширения и реформулирования запроса, локализации и оценки выдачи. Поисковый образ запроса — записанный на ИПЯ текст, выражающий смысловое содержание информационного запроса и содержащий указания, необходимые для наиболее эффективного осуществления информационного поиска. Методы поиска, т.е. выделение подмножества документов, потенциально содержащих описание решения задачи отбора документов (ОД), являются отражением процесса нахождения решения и зависят от характера задачи и предметной области. Сегодня активно используются и разрабатываются три модели поиска.

1. Векторная модель. Самой первой моделью, которая присутствует сегодня во всех системах, является векторная модель поиска. Ее изобрел Дж. Сэлтон в 60-х гг. Большинство машин работают по принципу наличия в релевантном документе всех терминов запроса, учета их встречаемости в документах и их средней языковой частотности. Эта модель используется при обработке запросов на естественном языке, особенно на поисковых страницах сайтов; она же применяется для поиска похожих документов.

2. Булева модель. Активно используется булева модель поиска, которая позволяет вводить в запрос логические операторы, контекстные ограничения на расстояние между словами, строить разветвленные мощные запросы, использовать стоп-словарь и лексические шаблоны аналогично регулярным выражениям в скриптовых языках. Профессиональные системы в дополнение к перечисленным базовым моделям предоставляют поиск с использованием нечеткой булевой модели поиска, позволяющей поисковой машине доставлять документы, которые она считает релевантными, даже если некоторые «слабые» элементы запроса в них не встречаются.

3. Поиск с использованием тезаурусов. Для семантического поиска широко используются тезаурусы, за счет которых происходит расширение запроса. Например, при поиске документов по автотранспортным происшествиям запрос «ДТП» имеет фактор расширения 1:150, т.е. из одной лексемы системой фактически генерируется 150 лексем для сервера поиска. Активное использование тезаурусов русского языка сдерживается сегодня отсутствием актуальных словарей синонимов.

Рассматривая поиск как итеративный процесс, методы сокращения пространства перебора (просматриваемого подмножества) образуют по существу методологическую основу стратегии поиска и могут быть разделены на следующие классы — методы поиска в:

- 1) одном пространстве (обычно тематическом);
- 2) иерархически упорядоченном пространстве;
- 3) альтернативных пространствах;
- 4) динамическом (изменяющемся в процессе поиска) пространстве.

Реализуемый метод построения ПОЗа должен обеспечивать эффективные способы построения запроса для достижения целей различного типа. Механизмы поиска — совокупность реализованных в системе моделей и алгоритмов процесса формирования выдачи документов в ответ на поисковый запрос. Средства поиска, с одной стороны, взаимозависимый комплекс информационно-поисковых языков (ИПЯ) и языков определения (управления) данными, обеспечивающий структурные и семантические преобразования объектов обработки (доку-

ментов, словарей, совокупностей результатов поиска), а с другой — объекты пользовательского интерфейса, обеспечивающие управление последовательностью выбора операционных объектов конкретной АИПС.

Поисковые технологии — унифицированные (оптимизированные в рамках конкретной АИПС) последовательности эффективного использования отдельных средств поиска в процессе взаимодействия пользователя с системой для устойчивого получения конечного и промежуточных результатов. Навигация как реализация процесса поиска по запросу в выбранной БД — целенаправленная, определяемая стратегией последовательность использования методов, средств и технологий конкретной АИПС для получения и оценки результата.

Средства навигации позволяют пользователю осуществлять управление процессом поиска. Они предоставляются пользователю в виде интерфейса, позволяющего организовать более или менее эффективный процесс взаимодействия с БД. При этом «дружественность» интерфейса характеризуется не только эргономичностью и понятностью, но и вариантноностью выбора операционных объектов.

Процесс поиска информации представляет последовательность шагов, приводящих при посредстве системы к некоторому результату и позволяющих оценить его полноту. Так как пользователь обычно не имеет исчерпывающих знаний об информационном содержании ресурса, в котором проводит поиск, то оценить адекватность выражения запроса, равно как и полноту получаемого результата, он может, основываясь лишь на внешних оценках или на промежуточных результатах и обобщениях, сопоставляя их, например, с предыдущими.

Процесс поиска можно представить в виде следующих основных компонент (рис. 4.1):

- формулирование запроса на естественном языке, выбор поисковых системы и сервисов, формализация запроса на соответствующем ИПЯ;
- проведение поиска в одной или нескольких поисковых системах;
- обзор полученных результатов (ссылок);
- предварительная обработка полученных результатов: просмотр содержания ссылок, извлечение и сохранение релевантных и pertinentных данных;
- при необходимости модификация запроса и проведение повторного (уточняющего) поиска с последующей обработкой полученных результатов.

Для уменьшения объема отобранных материалов осуществляют фильтрацию результатов поиска по типу источников (сайтов, порталов), те-



Рис. 4.1. Компоненты поисковых систем

матике и другим основаниям. По используемым поисковым технологиям ИС можно разбить на 4 категории.

1. Тематические каталоги.
2. Специализированные каталоги (онлайновые справочники).
3. Поисковые машины (полнотекстовый поиск).
4. Средства метапоиска.

В Интернете ИПС размещается на одном или нескольких серверах. В ИПС собирается, индексируется и регистрируется информация о документах, имеющихся в обслуживаемой системой группе веб-серверов. В документах индексируются все значащие слова или только слова из заголовков. Тематические каталоги предусматривают обработку документов и отнесение их к одной из нескольких категорий, перечень которых заранее задан. Фактически это индексирование на основе классификации. Индексирование может проводиться автоматически или вручную с помощью специалистов, просматривающих популярные веб-узлы и составляющих краткое описание документов-резюме (ключевые слова, аннотация, реферат).

Специализированные каталоги или справочники создаются по отдельным отраслям и темам, по новостям, по городам, по адресам электронной почты и т.п. Поисковые машины (самое развитое средство поиска в Интернете) реализуют технологию полнотекстового поиска. Индексируются тексты, расположенные на опрашиваемых серверах. Индекс может содержать информацию о нескольких миллионах документов. Например, в индексе популярной ИПС AltaVista более 56 млн URL-адресов.

При использовании средств метапоиска запрос осуществляется одновременно несколькими поисковыми системами. Результат поиска объединяется в общий, упорядоченный по степени релевантности список. Каждая система обрабатывает только часть узлов сети, что позволяет расширить базу поиска. К подобному классу можно отнести и «персональные программы поиска», позволяющие формировать свои собственные инструменты метапоиска (например, автоматически опрашивать часто посещаемые узлы).

Базы информационных данных могут содержать практически любые виды информации, в том числе в любой комбинации. Информационный поиск осуществляется как по существующим в полнотекстовых ЭИР терминам, так и по специальным элементам, входящим в состав ИПЯ. Для формирования запросов используются специальные информационно-поисковые языки.

ИПС внутри найденной выборки обычно пытаются расположить документы в порядке их «релевантности», т.е. близости к введенному пользователем запросу. Критериев такой близости много и выявление близких «по смыслу» к запросу документов не решает проблемы получения информации при отсутствии релевантного документа. Подобная ситуация достаточно тривиальна, в том числе и потому, что пользователь зачастую ищет документ, который сам собирается написать. Следует отметить, что в результате проведенного поиска пользователь может получить как релевантные, pertinentные, так и нерелевантные и непертинентные подмассивы данных.

ИПС фактически являются системами информационного обеспечения и представляют собой базы и банки данных. В качестве объекта в них выступает индивид, организация, отрасль, регион и т.п. Субъектом информационного обеспечения является специалист-информатик, любой потребитель информации.

Предлагается процедуру поиска необходимой информации разделить на девять основных этапов.

1. Определение области знаний.
2. Выбор типа и источников данных.
3. Сбор материалов, необходимых для наполнения информационной модели.
4. Отбор наиболее полезной информации.
5. Выбор метода обработки информации (классификация, кластеризация, регрессионный анализ и т.д.).
6. Выбор алгоритма поиска закономерностей.
7. Поиск закономерностей, формальных правил и структурных связей в собранной информации.

8. Творческая интерпретация полученных результатов.

9. Интеграция извлеченных «знаний».

Для проведения поиска первоначально на компьютере пользователя загружается интерфейс работы с соответствующей БД. Это может быть локальная или удаленная БД. Первоначально следует определиться с видом поиска (простой, расширенный и т.д.). Затем с набором предлагаемых для поиска полей. ИПС могут предложить для ввода одно или несколько полей. В последнем случае это обычно поля: автора, заглавия (названия), временного периода, вида документа, ключевых слов, рубрик и др. При формировании запроса практически все системы позволяют использовать логические элементы «И», «ИЛИ», «НЕТ».

Поисковые средства и технологии, используемые для реализации информационных потребностей, определяются типом и состоянием решаемой пользователем задачи основной деятельности: соотношением его знания и незнания об исследуемом объекте. Кроме того процесс взаимодействия пользователя с системой определяется уровнем знания пользователем содержания ресурса (полноты представления, достоверности источника и т.д.) и функциональных возможностей системы как инструмента. В целом эти факторы обычно сводятся к понятию «профессионализма» — информационного (подготовленный (неподготовленный) пользователь) и предметного (профессионал (непрофессионал) «профессионализма».

Процесс поиска информации обычно носит эмпирический характер. Он представляет последовательность шагов, приводящих при посредстве системы к некоторому результату, позволяющих оценить его полноту. При этом поведение пользователя, как организующее начало управления процессом поиска, мотивируется не только информационной потребностью, но и разнообразием стратегий, технологий и средств, предоставляемых системой.

Обычно пользователь не имеет исчерпывающих знаний об информационном содержании ресурса, в котором проводит поиск, поэтому оценить адекватность выражения запроса, как и полноту получаемого результата, он может, отыскав дополнительные сведения или организовав процесс так, чтобы часть результатов поиска могла использоваться для подтверждения или отрицания адекватности другой части.

Операционными объектами, непосредственно участвующими во взаимодействии пользователей с поисковой системой, являются поисковый образ документа (ПОД) и ПОЗ, соответствие которых устанавливается поисковым механизмом АИПС на формальном уровне. Адекватность образа действительному содержанию документа определяется качеством процесса свертки информации и уровнем знания субъектом средств

отражения — концептуальной схемы предметной области и возможностей ИПЯ. Поисковый образ документа — описание документа, выраженное средствами ИПЯ и характеризующее основное смысловое содержание или какие-либо другие признаки этого документа, необходимые для его поиска по запросу. Большинство ПС изначально предлагают пользователям либо БО, либо ссылки на полные или частичные документы, их описание и другое, хранящиеся в различных АИПС. Современные ПС позволяют определиться и указать, какой и в каком виде источник информации интересует пользователя.

По характеру преобразований (в контексте дальнейшего использования результатов обработки) методы обработки результатов поиска можно условно разделить на две группы.

1. Структурно-форматные преобразования.

2. Структурно-семантические преобразования (информационно-аналитические, логико-семантические).

Что обычно ищут в Интернете: персональные данные об индивидах и организациях; различные адресные данные; конкретные материалы (статьи, книги, фотографии, справочные данные, программное обеспечение и др.), в том числе место их хранения; где и сколько стоят те или иные материалы, услуги, продукты и т.п.; информационные сайты и порталы и др. Общепринята организация поиска по начальным фрагментам слова (поиск с усечением справа), например, вместо слова «библиотечный» можно ввести его фрагмент «библиоте\*». При этом будут найдены документы, в которых содержится не только слово «библиотечный», но и «библиотека», «библиотекарь», «библиотековедение» и др. В каждом случае пользователь должен представлять, что именно он хочет найти, так как в предложенном ему варианте будет найдено гораздо большее количество документов, чем при задании данного слова полностью (без усечения). В подобном случае возможно в полученном массиве информации провести уточняющий поиск и в результате получить более релевантные и pertinentные данные.

С точки зрения ИПС результат поиска в ней есть совокупность (подмножество) найденных документов или ссылок на них. Обычно он представляется пользователю в виде списка. То есть простейшей выходной формой в данном случае будет список ссылок в виде полных или частичных БО, найденных ИП. Такой список может быть тут же распечатан или послан на какой-либо адрес электронной почты, если такая возможность предоставляется ИПС и пользователь подключен к Интернету.

Графические и полнотекстовые ЭИР могут предлагаться пользователю только для просмотра, для копирования в различных форматах

и масштабах, причем полностью или частично. Графические ИР обычно существуют в общепринятых форматах типа: JPG, GIFF, TIFF, BMP и др., а для текстовых материалов обычно используют текстовые форматы TXT, DOC и др., HTML и PDF — фактически графический формат, в котором могут сохраняться как текстовые, так и графические данные. Полученные в результате поиска документы сохраняют.

Критерием результата поиска является получение пользователем списка документов, одного документа или их частей, максимально удовлетворяющего его потребностям, сформулированным в поисковом запросе. В ИПС принято формировать список полученных в результате поиска документов по их релевантности. Различают критерии смыслового и формального соответствия между поисковым предписанием и выдаваемым документом. Полнота и точность поиска являются взаимосвязанными показателями. Увеличение одного из них ведет к снижению другого. В современных ИПС при сбалансированном поиске их значения составляет примерно 70%. Следует учитывать ситуацию, при которой список выданных поисковой системой ссылок содержит несколько, а порой и десятки разных адресов с одним и тем же текстом. Подобные ссылки характеризуются как дубликаты. Из них при подсчете коэффициентов учитывается только один документ.

Учитывая, что идеальный результат поиска должен удовлетворять требованиям единственности, полноты и непротиворечивости, получаем, что различные виды поиска определяют различные требования к функциональным возможностям системы в части оценивания результата. Однако для случая предметного поиска доказательство полноты является тривиальным: непустой результат поиска подтверждает факт существования (или отсутствия) объекта, обладающего искомыми свойствами. При этом результат тематического поиска множественен и требует последующей систематизации — еще одного процедурного шага для упорядочения полученного множества объектов по значениям не определенного явно основания. В свою очередь проблемный поиск предлагает уже двухуровневую систематизацию.

Развитие процесса поиска осуществляется путем модификации выражения ПОЗ, путем реформулирования запроса и проведения повторного поиска в том же массиве данных или в подмассиве, полученном в результате осуществления первоначального поиска. Интерфейсные средства обработки результата и развития поиска используют два типа операционных объектов — отдельные документы или коллекции документов.

Для получения информации в среде Интернета создаются специальные поисковые системы. Как правило, они общедоступны и обслу-

живают пользователей в любой точке планеты, где имеется возможность работы с Интернетом. Непосредственно для поиска используются поисковые машины, число которых в мире исчисляется несколькими сотнями. Они ориентируются на определенные типы запросов или их сочетание (библиографический, адресный, фактографический, тематический и др.). Кроме того бывают полнотекстовые, смешанные и другие поисковые машины. Для проведения поиска в Интернете (в WWW) функционирует множество сайтов и поисковых систем, поэтому необходимо не только ориентироваться в таких системах, но и уметь осуществлять в них эффективный поиск, т.е. использовать соответствующие технологии.

Технология поиска (от англ. Search Technology) означает совокупность правил и процедур, в результате выполнения которых пользователь получает ИР. При поиске в Интернете рекомендуется обращать внимание на две составляющие: полноту (ничего не потеряно) и точность (не найдено ничего лишнего). Обычно соответствие найденных материалов этим критериям называют релевантностью, т.е. соответствием ответа вопросу (запросу). Поисковые системы характеризуются также временем выполнения поиска, интерфейсом, предоставляемым пользователю и видом отображаемых результатов. При выборе поисковых систем обращают внимание на такие их параметры, как охват и глубина. Под охватом понимается объем базы поисковой машины, измеряемый тремя показателями: общим объемом проиндексированной информации, количеством уникальных серверов и количеством уникальных документов. Под глубиной понимается, существует ли ограничение на количество страниц или на глубину вложенности директорий на одном сервере.

Каждая поисковая машина имеет свои алгоритмы сортировки результатов поиска. Чем ближе к началу списка, полученного в результате проведения поиска, оказывается нужный документ, тем выше релевантность и лучше работает поисковая машина. Поисковые машины используют общие принципы работы, ориентированные на выполнение двух основных функций. Первая функция реализуется программой-роботом, автоматически просматривающей различные сервера в Интернете. Находя новые или изменившиеся документы, она осуществляет их индексацию и передает на базовый компьютер поисковой машины. Робот — автоматизированный браузер, загружающий веб-страницу, изучающий ее и при необходимости переходящий к одной из ее гиперсвязей. Когда ему попадается страница, не содержащая связей, робот возвращается на одну — две ступени назад и переходит по адресу, указанному в одной из обнаруженных ранее связей. Запущенный робот проходит

огромные расстояния в среде Интернета (киберпространстве), ориентируясь на развитие веб-сети и изменяя в соответствии с этим свои маршруты. Индексирующие роботы обрабатывают лишь HTML-файлы, игнорируя изображения и другие мультимедийные файлы. Они могут: обнаруживать связи с уже несуществующими страницами; устанавливать связь с наиболее популярными узлами, подсчитывая количество ссылок на них в других веб-страницах; регистрировать веб-страницы для оценки роста системы и др. Чаще всего роботы просматривают сервера самостоятельно, находя новые внешние ссылки в уже обследованных документах. Вторая функция заключается в обработке выявленных документов. При этом учитывается все содержание страниц (не только полный текст, но и наличие иллюстраций, аудио- и видеофайлов, Java-приложений). Индексации подвергаются все слова в документе, что дает возможность использовать поисковые системы для детального поиска по самой узкой тематике. Образующие гигантские индексные файлы, хранящие информацию о том, какое слово, сколько раз, в каком документе и на каком сервере употребляется, составляют БД, к которой собственно и обращаются пользователи, вводя в поисковую строку ПОЗ (сочетания ключевых слов). Выдача результатов осуществляется с помощью специальной подсистемы, производящей интеллектуальное ранжирование результатов. В своих расчетах она опирается на местоположение термина, частоту его повторения в тексте, процентное соотношение данного термина с остальным текстом на данной странице и другие параметры, характеризующие возможности конкретной поисковой машины.

Роботы имеют ряд разновидностей, одной из которых является «паук» (от англ. spider). Он непрерывно «ползает по сети», переходя с одной веб-страницы к другой с целью сбора статистических данных о самой «паутине» (Web) и (или) формирования некоторой БД с индексами содержимого веб. Автоматизированные агенты «спайдеры» регулярно сканируют веб-страницы и актуализируют БД адресов (гиперссылки), средства индексирования информации, расположенные по указанным адресам. Полученные индексы используются для быстрого и эффективного поиска по набору терминов, задаваемых пользователем. В разных системах эта цель достигается различным образом. Одни посылают «агентов» на каждую попадающуюся веб-страницу, индексируя все встречающиеся слова. Другие сначала анализируют БД адресов, определяя наиболее популярные (обычно подсчитывается число имеющихся ссылок на них). Именно эти веб-страницы в различной степени индексируются (только заголовки веб-страниц и ссылки, включая автоматическое аннотирование документов или весь текст).

Все чаще применяются «интеллектуальные агенты» — небольшие программы, обладающие способностью самообучаться и действовать самостоятельно от имени своего владельца. Имея связь с компьютером пользователя, они выступают в роли персональных помощников, выполняющих ряд задач с применением знаний о потребностях и интересах пользователя. Интеллектуальные роботы-агенты ведут самостоятельный поиск в сети по собственным уникальным алгоритмам. Некоторые из них не только просматривают ключевые слова, но и осуществляют в Интернете семантический анализ информации, выявляя степень ее смыслового соответствия поставленной задаче.

Эффективный доступ к информации в Интернете обеспечивают такие зарубежные поисковые системы (машины), как Альта-Виста (AltaVista), Lycos, Yahoo, Google, OpenText, Wais, WebCrawler и др. Их адреса в Интернете: [www.altavista.com](http://www.altavista.com), [www.yahoo.com](http://www.yahoo.com), [www.gogle.com](http://www.gogle.com), [www.opentext.com](http://www.opentext.com).

К отечественным поисковым машинам относятся: Нигма (Научный проект «Интеллектуальная поисковая система Nigma.ru» МГУ им. М.В. Ломоносова, факультетов ВМиК и психологии), Апорт (Aport АО Агама), Rambler (фирма Stack Ltd.), Яндекс (Yandex фирма CompTek Int), «Русская машина поиска», «Новый русский поиск», и др. Их адреса в Интернете: [www.nigma.ru](http://www.nigma.ru) (нигма.пф), [www.aport.ru](http://www.aport.ru), [www.rambler.ru](http://www.rambler.ru), [www.yandex.ru](http://www.yandex.ru), [search.interrussia.com](http://search.interrussia.com), [www.openweb.ru](http://www.openweb.ru) соответственно и др.

Все эти поисковые машины позволяют по ключевым словам, тематическим рубрикам и даже отдельным буквам оперативно находить в сети, например, все или почти все тексты, где эти слова присутствуют. При этом пользователю сообщаются адреса сайтов, где найденные IP постоянно присутствуют. Однако ни одна из них не имеет подавляющих преимуществ перед другими. Для проведения надежного поиска по сложным запросам специалисты рекомендуют использовать последовательно или параллельно (одновременно) различные ИПС.

Полнотекстовая поисковая машина индексирует все слова видимого пользователю текста. Наличие морфологии дает возможность находить искомые слова во всех склонениях или спряжениях. Кроме этого в языке HTML существуют тэги, которые также могут обрабатываться поисковой машиной (заголовки, ссылки, подписи к картинкам и т.д.). Некоторые машины умеют искать словосочетания или слова на заданном расстоянии, что часто бывает важно для получения разумного результата.

Google — единственная компания, сосредоточившая свои усилия на разработке «идеальной поисковой системы», которая, по словам соучредителя компании Лэрри Пейджа, «точно определит, что подразумевает пользователь, и покажет именно те результаты, которые ему нуж-

ны». С этой целью Google непрестанно ищет новые методы и отказывается смиряться с ограничениями существующих технологий. В итоге Google разработала собственную инфраструктуру и революционную технологию PageRank, которая изменила подход к выполнению поиска. С самого начала разработчики Google понимали, что для более быстрого получения наиболее точных результатов необходим новый способ настройки сервера. Большинство поисковых систем использовали несколько больших серверов, которые часто работали медленно при пиковых нагрузках. Компания Google задействовала связанные ПК, позволяющие быстро находить ответы на все запросы. Внедрение этой инновационной технологии привело к сокращению времени отклика, повышению масштабируемости и снижению расходов. С тех пор все остальные компании копируют эту идею, в то время как Google продолжает постоянно улучшать внутреннюю технологию с целью повышения ее эффективности.

Программное обеспечение, используемое для реализации технологии поиска Google, проводит ряд одновременных вычислений, которые занимают не больше доли секунды. Традиционные поисковые системы в большей степени основываются на том, как часто слово появляется на веб-странице. Google же изучает всю структуру веб-ссылок и определяет, какие страницы наиболее важны, с помощью PageRank. Затем проводится анализ соответствия гипертекста и выбор страниц, наиболее подходящих для конкретного поиска. На основании общей значимости и соответствия запросу Google отображает в первую очередь наиболее релевантные и достоверные результаты.

Технология PageRank: PageRank объективно оценивает значимость веб-страниц, основываясь на уравнении, включающем более 500 млн переменных и 2 млрд терминов. Вместо того чтобы подсчитывать прямые ссылки, PageRank рассматривает ссылку со страницы А на страницу Б как голос в пользу страницы Б от страницы А. Затем по количеству полученных голосов PageRank определяет значимость данной страницы. PageRank также оценивает важность каждой страницы, принимающей участие в голосовании. При получении голосов от страниц с большей значимостью ссылка становится более ценной. Значимые страницы получают более высокий рейтинг PageRank и отображаются в начале результатов поиска. Технология Google использует совокупные интеллектуальные веб-средства, чтобы определить значимость страницы. Человеческий фактор или подтасовка результатов невозможны, и именно поэтому пользователи доверяют Google как источнику объективной информации, в результатах поиска которого отсутствуют оплаченные рекламные объявления.

Анализ соответствия гипертекста: поисковая система Google, как и другие системы, также анализирует содержание страницы. Однако вместо простого сканирования текста страницы (который может выполнить веб-издатель с помощью метатегов — HTML- или XHTML-теги, предназначенные для предоставления структурированных метаданных о веб-странице.) технология Google анализирует все содержание страницы, особенности шрифтов, разбивки текста и точное расположение каждого слова. Google также анализирует содержание соседних веб-страниц, чтобы убедиться в том, что полученные результаты наиболее точно соответствуют запросу пользователя (рис. 4.2).

Инновационные технологии Google распространяются не только на настольные ПК. Компания Google поставила перед собой задачу донести результаты своего точного и быстрого поиска до пользователей, работающих в Интернете через мобильные устройства. С этой целью Google разработала первую в своем роде технологию беспроводного поиска для мгновенного преобразования HTML в форматы, оптимизированные

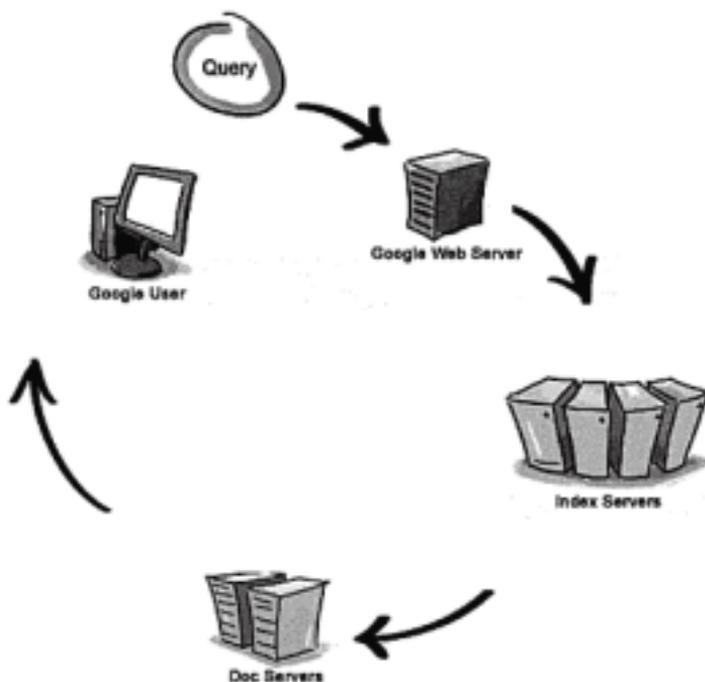


Рис. 4.2. Работа Google

для WAP, i-mode, J-SKY и EZWeb. В настоящее время беспроводная технология Google используется множеством лидирующих на рынке компаний, включая AT & T Wireless, Sprint PCS, Nextel, Palm, Handspring и Vodafone.

Время поиска Google обычно занимает менее чем полсекунды, однако эта процедура включает в себя ряд действий, без которых результаты поиска не будут доставлены пользователю. Этапы запроса Google:

1) веб-сервер отправляет запрос на серверы индексации. Содержимое серверов индексации соответствует содержанию в конце книги — в нем сообщается, на каких страницах представлены слова, соответствующие запросу;

2) запрос отправляется на серверы, которые осуществляют поиск хранящихся документов. Для описания каждого результата поиска создаются фрагменты;

3) всего лишь доля секунды требуется для получения результатов поиска.

«Идеальная поисковая система, — говорит один из основателей компании Google Лэрри Пейдж (Larry Page), — точно определит, что подразумевает пользователь, и покажет именно те результаты, которые ему нужны». Учитывая текущий уровень развития технологии поиска, для этого нужно смотреть далеко вперед, проводить исследования, разработку и внедрение новейших методов работы. Google прилагает все усилия к тому, чтобы прокладывать новые пути в этой области. Несмотря на то, что Google уже признана ведущей компанией мира, занимающейся технологией поиска.

## 4.2. Компьютерные справочно-информационные системы, справочно-правовые системы

Машины веб-поиска — это сервера с огромной базой данных URL-адресов, которые автоматически обращаются к страницам WWW по всем этим адресам, изучают содержимое этих страниц, формируют и прописывают ключевые слова со страниц в свою базу данных (индексируют страницы). Более того, роботы поисковых систем переходят по встречаемым на страницах ссылкам и переиндексируют их. Так как почти любая страница WWW имеет множество ссылок на другие страницы, то при подобной работе поисковая машина в конечном результате теоретически может обойти все сайты в Интернет. Именно этот вид поисковых инструментов является наиболее известным и популярным среди всех пользователей сети «Интернет». У каждого на слуху названия известных машин веб-поиска (поисковых систем) — Yandex, Rambler, Aport.

Чтобы воспользоваться данным видом поискового инструмента, необходимо зайти на него и набрать в строке поиска интересующее вас ключевое слово. Далее вы получите выдачу из ссылок, хранящихся в базе поисковой системы, которые наиболее близки вашему запросу. Чтобы поиск был наиболее эффективен, заранее обратите внимание на следующие моменты.

1. Определитесь с темой запроса. Что именно в конечном итоге вы хотите найти?

2. Обращайте внимание на язык, грамматику, использование различных небуквенных символов, морфологию. Важно также правильно сформулировать и вписать ключевые слова. Каждая поисковая система имеет свою форму составления запроса — принцип один, но могут различаться используемые символы или операторы. Требуемые формы запроса различаются также в зависимости от сложности программного обеспечения поисковых систем и предоставляемых ими услуг. Так или иначе, каждая поисковая система имеет раздел Help («Помощь»), где все синтаксические правила, а также рекомендации и советы по поиску доступно объясняются (скриншот страничек поисковиков).

3. Используйте возможности разных поисковых систем. Если не нашли на Yandex, попробуйте на Google. Пользуйтесь услугами расширенного поиска.

4. Чтобы исключить документы, содержащие определенные термины, используйте знак «-» перед каждым таким словом. Например, если вам нужна информация о работах Шекспира за исключением «Гамлета», то введите запрос в виде: «Шекспир — Гамлет». И для того чтобы наоборот в результаты поиска обязательно включались определенные ссылки, используйте символ «+». Так, чтобы найти ссылки о продаже именно автомобилей, вам нужен запрос «продажа + автомобиль». Для увеличения эффективности и точности поиска используйте комбинации этих символов.

5. Каждая ссылка в списке результатов поиска содержит сниппет — несколько строчек из найденного документа, среди которых встречаются ваши ключевые слова. Прежде чем переходить по ссылке, оцените соответствие сниппета теме запроса. Перейдя по ссылке на определенный сайт, внимательно окиньте взглядом главную страничку. Как правило, первой страницы достаточно, чтобы понять, по адресу вы пришли или нет. Если да, то дальнейшие поиски нужной информации ведите на выбранном сайте (в разделах сайта), если нет — возвращайтесь к результатам поиска и пробуйте очередную ссылку.

6. Помните, что поисковые системы не производят самостоятельную информацию (за исключением разъяснений о самих себе). Поисковая



Кроме того некоторые тематические каталоги позволяют искать по ключевым словам. Пользователь вводит необходимое ключевое слово в строку поиска и получает список ссылок с описаниями сайтов, которые наиболее полно соответствуют его запросу. Стоит отметить, что этот поиск происходит не в содержимом WWW-серверов, а в их кратком описании, хранящемся в каталоге.

Подборки ссылок (рис. 4.4) — это отсортированные по темам ссылки. Они достаточно сильно отличаются друг от друга по наполнению, поэтому, чтобы найти подборку, наиболее полно отвечающую вашим интересам, необходимо ходить по ним самостоятельно, дабы составить собственное мнение. В качестве примера приведем Подборку ссылок «Сокровища Интернет» АО «Релком».



Рис. 4.4. Рабочее окно системы подборки ссылок

Пользователь, нажимая на любую из заинтересовавших его рубрик:

- Содержание
- Автомобилистам
- Астрономия и астрология
- Ваш дом
- Ваши питомцы
- Дети — цветы жизни
- Досуг
- Города в сети «Интернет»
- Здоровье и медицина
- Информационные агентства и службы
- Краеведческий музей и т.д.,

попадает на подборку со ссылками на полезные Интернет-ресурсы.

Преимуществом такого вида поисковых инструментов является их целенаправленность, обычно подборка включает в себя редкие Интернет-ресурсы, подобранные конкретным веб-мастером или хозяином Интернет-странички.

Базы данных адресов (рис. 4.5) — это специальные поисковые сервера, которые обычно используют классификации по роду деятельности, по выпускаемой продукции и оказываемым услугам, по географическому признаку. Иногда они дополнены поиском по алфавиту. В записях базы данных хранится информация о сайтах, которые предоставляют информацию об электронном адресе, организации и почтовом адресе за определенную плату. Крупнейшей англоязычной базой данных адресов можно назвать <http://www.lookup.com/> — представляет собой рубрику с множеством поддиректорий, таких как: Find Person, Phone Numbers, Name Search и др. Попадая в данные поддиректории, пользователь обнаруживает ссылки на сайты, которые и предлагают интересующую его информацию.

Gopher — это взаимосвязанная система серверов (Gopher-пространство), распределенная по Интернет. В пространстве Gopher собрана богатейшая литературная библиотека, однако материалы недоступны для просмотра в удаленном режиме: пользователь может только просматривать иерархически организованное оглавление и выбирать файл по названию. С помощью специальной программы (Veronica) такой поиск можно сделать и автоматически, используя запросы, построенные на ключевых словах.

До 1995 г. Gopher являлся самой динамичной технологией Интернет: темпы роста числа соответствующих серверов опережали темпы роста серверов всех других типов Интернет. В сети EUnet/Relcom активного



Рис. 4.5. Рабочее окно базы данных адресов

развития серверы Gopher не получили, и сегодня о них практически никто не вспоминает.

Система поиска FTP-файлов — это особый тип средств поиска в Интернете, который позволяет находить файлы, доступные на «анонимных» FTP-серверах. Протокол FTP предназначен для передачи по сети файлов, и в этом смысле он функционально является своеобразным аналогом Gopher. Основным критерием поиска является название файла, задаваемое разными способами (точное соответствие, подстрока, регулярное выражение и т.д.). Данный тип поиска, конечно же, не может соперничать по возможностям с поисковыми машинами, так как содержимое файлов никак не учитывается при поиске, а файлам, как известно, можно давать произвольные имена. Тем не менее если вам

требуется найти какую-нибудь известную программу или описание стандарта, то с большой долей вероятности файл, его содержащий, будет иметь соответствующее имя, и вы сможете найти его при помощи одного из серверов FTP Search.

FileSearch (рис. 4.6) ищет файлы на FTP-серверах по именам самих файлов и каталогов. Если вы ищете какую-либо программу или еще что-то, то на WWW-серверах вы, скорее, найдете их описание, а с FTP-серверов вы сможете перекачать их к себе.

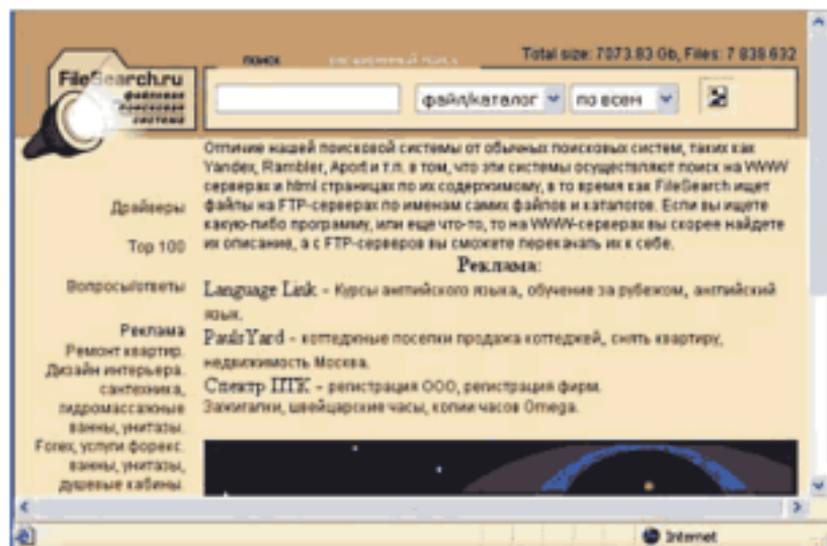


Рис. 4.6. Рабочее окно FileSearch

USENET NEWS — это система телеконференций сообщества сетей «Интернет». На Западе этот сервис принято называть новостями. Ближним аналогом телеконференций являются и так называемые эхи в сети FIDO.

С точки зрения абонента телеконференции, USENET представляет из себя доску объявлений, в которой есть разделы, где можно найти статьи на любую тему — от политики до садоводства. Эта доска объявлений доступна через компьютер, подобно электронной почте. Не отходя от компьютера, можно читать или помещать статьи в ту или иную конференцию, найти полезный совет или вступать в дискуссии. Естественно, статьи занимают место на компьютерах, поэтому не хранятся вечно,

а периодически уничтожаются, освобождая место для новых. Во всем мире лучшим сервисом для поиска информации в конференциях USENET является сервер Google Groups (Google Inc.). Группы Google — это бесплатное интерактивное сообщество и служба групп обсуждений, которая предлагает самый обширный в Интернете архив сообщений сети USENET (более млрд сообщений).

### **Контрольные вопросы и задания**

1. Дайте определение понятия «информационные ресурсы».
2. Перечислите основные виды ИР.
3. Что такое «информационный поиск»?
4. Назовите виды информационного поиска и дайте их определения.
5. Перечислите средства информационного поиска для каждого из его видов.
6. Дайте типологию справочных изданий.
7. Перечислите и охарактеризуйте основные виды свертывания информации.
8. Каковы сущность, цели, задачи и условия применения формализованного анализа документов?
9. Дайте определения понятий «информационная технология» и «новая информационная технология».
10. Приведите пример реализации технологии.

## ИНФОРМАЦИОННЫЙ ПРОЦЕСС НАКОПЛЕНИЯ И ХРАНЕНИЯ ДАННЫХ

Работа в ИТ должна быть хорошо скоординирована между выполняющими ее людьми и программным и техническим обеспечением; должны быть обеспечены тесные связи, позволяющие обмениваться информацией в кратчайшие сроки. Анализ бизнес-процессов, выполняемых работниками, использующими ИТ, позволяет классифицировать в общем виде как задачи, решаемые предприятием (организацией), так и исполнителями этих задач. Такая классификация задач основана на степени их интеллектуальности и сложности.

1. Наиболее простые задачи, состоящие из полностью формализуемых процедур, выполнение которых не представляет особых трудностей для исполнителей. Эти задачи легко стандартизируются и программируются. К ним относятся контроль и учет, оформление документов, их тиражирование, рассылка и т.п. Подобного рода задачи в настоящее время решаются практически всеми автоматизированными информационными системами (например, «Бухгалтерский учет», «Подготовка производства», «Кадры», «Складской учет»). Задачи этого класса, если они используются для принятия решений, называются задачами принятия решений в условиях полной определенности. При этом случайные и неопределенные факторы отсутствуют. Такие задачи часто решаются путем разработки различного вида информационных систем с использованием средств языка системы управления базами данных (СУБД).

2. Более сложные задачи — задачи принятия решений в условиях риска, т.е. в том случае, когда имеются случайные факторы, для которых известны законы их распределения. Постановка и решение таких задач возможны на основе методов теории вероятностей, аналитического и имитационного моделирования.

3. Слабоструктурированные задачи, содержащие неизвестные или не измеряемые компоненты (количественно не оцениваемые). Для таких задач характерно отсутствие методов решения на основе непосредственных преобразований данных. Постановка задач базируется на принятии решений в условиях неполной информации. В ряде случаев на основе теории нечетких множеств и приложений этой теории удается построить формальные схемы решения.

4. Задачи принятия решений в условиях противодействия или конфликта (например, необходимо учитывать наличие активно действующих конкурентов, преследующих собственные интересы). В задачах этого класса могут иметься случайные факторы, для которых известны или не известны законы их поведения. Постановка и решение таких задач возможны (правда, не всегда) методами теории вероятностей, теории нечетких множеств и теории игр.

5. Наиболее сложные задачи принятия решений при отсутствии возможности формализации из-за высокой степени неопределенности. К таким задачам относится большинство проблем прогнозирования, перспективного планирования и т.п. Основой решения этого класса задач являются творческий потенциал исполнителя, особенности его личности, а также атрибуты его деятельности (информированность, квалификация, талант, интуиция, образование и т.п.). Решение таких задач возможно с применением экспертных систем.

Именно те задачи, которые мы рассмотрели в данной классификации под номером 1, и являются основой для разработки и внедрения технологий обработки данных.

## 5.1. Основные особенности построения ИТ

Повторимся, информационная технология обработки данных предназначена для решения хорошо структурированных задач, по которым имеются необходимые входные данные (причем эти данные заведомо были известны, их форматы, также была заранее известна и структура, размещающая эти данные) и известны алгоритмы и другие стандартные процедуры их обработки. Эта технология применяется на уровне операционной (исполнительской) деятельности персонала невысокой квалификации в целях автоматизации некоторых рутинных постоянно повто-

ряющихся операций управленческого труда. Поэтому внедрение информационных технологий и систем на этом уровне существенно повысит производительность труда персонала, освободит его от рутинных операций, возможно, даже приведет к необходимости сокращения численности работников.

На уровне операционной деятельности решаются следующие задачи:

- обработка данных об операциях, производимых фирмой;
- создание периодических контрольных отчетов о состоянии дел в фирме;
- получение ответов на всевозможные текущие запросы и оформление их в виде бумажных документов или отчетов.

Примером может послужить ежедневный отчет о поступлениях и выдачах наличных средств банком, формируемый в целях контроля баланса наличных средств, или же запрос к базе данных по кадрам, который позволит получить данные о требованиях, предъявляемых к кандидатам на занятие определенной должности.

Существует несколько особенностей, связанных с обработкой данных, отличающих данную технологию от всех прочих: выполнение необходимых фирме задач по обработке данных. Каждой фирме предписано законом иметь и хранить данные о своей деятельности, которые можно использовать как средство обеспечения и поддержания контроля на фирме. Поэтому в любой фирме обязательно должна быть:

- информационная система обработки данных и разработана соответствующая информационная технология,
- решение только хорошо структурированных задач, для которых можно разработать алгоритм,
- выполнение стандартных процедур обработки. Существующие стандарты определяют типовые процедуры обработки данных и предписывают их соблюдение организациями всех видов,
- выполнение основного объема работ в автоматическом режиме с минимальным участием человека,
- использование детализированных данных. Записи о деятельности фирмы имеют детальный (подробный) характер, допускающий проведение ревизий. В процессе ревизии деятельность фирмы проверяется хронологически от начала периода к его концу и от конца к началу,
- акцент на хронологию событий,
- требование минимальной помощи в решении проблем со стороны специалистов других уровней.

## 5.2. Основные компоненты технологии

Представим основные компоненты информационной технологии обработки данных и приведем их характеристики, основные компоненты представлены на рис. 5.1.

Сбор данных. По мере того как фирма производит продукцию или услуги, каждое ее действие сопровождается соответствующими записями данных. Обычно действия фирмы, затрагивающие внешнее окружение, выделяются особо как операции, производимые фирмой.



Рис. 5.1. Основные типовые операции ИТ

Обработка данных. Для создания из поступающих данных информации, отражающей деятельность фирмы, используются следующие типовые операции:

- классификация или группировка;
- сортировка, с помощью которой упорядочивается последовательность записей;
- вычисления, включающие арифметические и логические операции (эти операции, выполняемые над данными, дают возможность получать новые данные);

- Укрупнение, или агрегирование, данных, служащее для уменьшения количества данных и реализуемое в форме расчетов итоговых или средних значений.

Первичные данные обычно имеют вид кодов, состоящих из одного или нескольких символов (эти коды, выражающие определенные признаки объектов, используются для идентификации и группировки записей).



Рис. 5.2. Компоненты ИТ для работы с документами

**Пример.** При расчете заработной платы каждая запись включает в себя код (табельный номер) работника, код подразделения, в котором он работает, занимаемую должность и т.п. В соответствии с этими кодами можно произвести разные группировки.

**Хранение данных.** Многие данные на уровне операционной деятельности необходимо сохранять для последующего использования либо здесь же, либо на другом уровне. Для их хранения создаются базы данных.

**Создание отчетов (документов).** В информационной технологии обработки данных необходимо создавать документы для руководства и работников фирмы, а также для внешних партнеров. При этом документы могут создаваться как по запросу или в связи с проведенной фирмой операцией, так и периодически в конце каждого месяца, квартала или года. Пример использования технологии при работе с документами представлен на рис. 5.2.

### 5.3. Программное обеспечение технологии

Для реализации данной технологии в качестве основного ПО, способного ее реализовать, следует рассматривать нижеперечисленные компоненты:

- 1) операционные системы (их мы подробно рассматривали ранее, причем нет обязательной ориентации на конкретный тип, чистая кроссплатформенность и обязательная ориентация на работу с сетью);
- 2) офисный пакет (любой имеющийся в наличии офисный пакет дает возможность реализовать такую технологию);
- 3) СУБД, способную интегрироваться с имеющимся офисным пакетом на уровне обмена данных и способную реализовать интересующие пользователей отчеты на основе заранее спроектированных алгоритмов);
- 4) электронная почта (компонент технологий, рассмотренный детально в предыдущих темах).

### 5.4. Техническое обеспечение технологии

С точки зрения реализации на основе современных достижений вычислительной техники выделяют следующие виды обработки информации:

- последовательная обработка, применяемая в традиционной фонеймановской архитектуре ЭВМ, располагающей одним процессором;
- параллельная обработка, применяемая при наличии нескольких процессоров в ЭВМ;
- конвейерная обработка, связанная с использованием в архитектуре ЭВМ одних и тех же ресурсов для решения разных задач.

Поэтому обработка данных сегодня — это не только возможность просто запустить исходные данные в прикладной процесс на ЭВМ, а еще и возможность обработать эти данные так, что результат будет превосходить имеющиеся ожидания и обеспечить одновременную работу с этими данными множества пользователей.

Помимо выполнения расчетов для получения конечных данных интересна еще и операция корректировки данных. Корректировка — модификация ранее сформированных данных, в результате которой их состояние соответствует реально существующим условиям. При корректировке могут выполняться следующие действия: добавление, исключение, изменение записей существующих файлов данных. Создание файлов данных можно интерпретировать как частный случай реализации действий добавления. Объектами корректировки могут быть записи файла или отдельные поля записей. Одним из основных условий выполнения корректировки является поиск местоположения данных, который, как правило, осуществляется по ключам.

В зависимости от того, сколько записей одновременно подвергается модификации, принято разделять ее на индивидуальную и групповую. При индивидуальной корректировке одна корректирующая запись вызывает модификацию одной записи файла, а при групповой — нескольких записей. Так, примером групповой корректировки могут служить изменения значения некоторого реквизита во всех записях файла.

Необходимо различать корректировку автономных файлов и базы данных. В первом случае модификации подвергаются только записи соответствующего файла, а во втором — записи файлов и соответствующие связи.

Корректировка проводится в основном по принципу «отец — сын». Сущность его заключается в том, что для выполнения модификации необходим исходный файл (отец) и файл корректуры, в результате получается откорректированный файл (сын).

Большое значение при корректировке придается контролю достоверности информации и ее защите от несанкционированного доступа. Это обеспечивается путем сохранения исходного файла и файла корректуры, а также введением системы паролей и ключей защиты.

Выдача информации — отражение (представление) результатов решения задач. По способу отражения выходной (результатной) информации различают вывод данных на бумагу, на машинные носители (в основном магнитные диски) и на видеотерминальные устройства (дисплей).

Традиционно результаты обработки отображаются в виде бумажных документов. При этом они проверяются на комплектность и логическую непротиворечивость, а затем оформляются юридически. С развитием технического, программного и иного обеспечения стала преобладать

тенденция перехода на безбумажную технологию. Но полный отказ от бумажных документов пока невозможен, в частности, из-за нерешенности правовых вопросов оформления информации. Вывод данных может осуществляться как непосредственно в месте обработки, так и по каналам связи для удаленных абонентов.

Основной задачей операции представления информации пользователю является создание эффективного интерфейса в системе «человек — компьютер». Среди существующих вариантов интерфейса в системе «человек — компьютер» можно выделить два основных типа: на основе меню и на основе языка команд.

Интерфейсы типа «меню» облегчают взаимодействие пользователя с компьютером, так как не требуют предварительного изучения языка общения с системой. На каждом шаге диалога пользователю предъявляются все возможные в данный момент команды в виде наборов пунктов меню, из которого пользователь должен выбрать нужный. Такой способ общения удобен для начинающих и непрофессиональных пользователей. Интерфейс на основе языка команд требует знания пользователем синтаксиса языка общения с компьютером. Достоинством командного языка является его гибкость и мощность. Технология представления информации должна давать дополнительные возможности для понимания данных пользователями, поэтому целесообразно использование графики, диаграмм, карт.

Пассивные пользователи, называемые иногда потребителями, обладают рядом специфических качеств, связанных с отсутствием времени, желания и квалификации для более глубокого изучения используемых инструментальных средств (СУБД, ЭТ, электронной почты и т.д.). В этом случае алгоритм общения с системой должен быть предельно простым, а кроме того, пользователь должен использовать единственный универсальный интерфейс, позволяющий единообразно работать с подготовленной информацией. Другая часть пользователей требует предоставления достаточно широкого круга средств активного влияния на выполняемые информационные процессы. Этим требованиям удовлетворяет Web-технология, которая имеет следующие особенности:

- 1) информация предоставляется потребителю в виде публикаций;
- 2) публикация может объединять информационные источники различной природы и географического расположения;
- 3) изменения в информационных источниках мгновенно отражаются в публикациях;
- 4) в публикациях могут содержаться ссылки на другие публикации без ограничения на местоположение и источники последних (гипертекстовые ссылки);

5) потребительские качества публикаций соответствуют современным стандартам мультимедиа (доступны текст, графика, звук, видео, анимация);

6) число потенциальных потребителей информации практически не ограничено;

7) информация легко усваивается потребителем благодаря широкому спектру изобразительных возможностей, предоставляемых Web-технологией;

8) технология не предъявляет особых требований к типам и источникам информации.

Любая система создается на довольно длительный период эксплуатации, за время которого накапливается достаточно большое количество данных. Эти данные составляют основу архивного хранения и используются при решении задач, связанных с ретроспективными данными, анализом хозяйственной деятельности. В настоящее время определяющим направлением реализации этой операции является концепция базы данных и хранилища данных. База данных может быть определена как совокупность данных, организованных по определенным правилам, предусматривающим общие принципы описания, хранения и манипулирования данными, независимая от прикладных программ.

## 5.5. Хранилища данных

Дальнейшим развитием технологии обработки данных является технология «хранилищ данных». Хранилище данных<sup>1</sup> (ХД) — это база, хранящая данные, агрегированные по многим измерениям. Подобная многомерность в структуре данных в отличие от баз данных, к примеру, реляционной модели, имеющей исключительно плоскую двумерную структуру, дает более высокие возможности по проведению анализа данных.

Основные отличия БД и ХД:

- агрегирование данных;
- данные из ХД никогда не удаляются;
- наполнение ХД происходит на периодической основе;

<sup>1</sup> Хранилище данных (от англ. Data Warehouse) — предметно ориентированная информационная база данных, специально разработанная и предназначенная для подготовки отчетов и бизнес-анализа с целью поддержки принятия решений в организации. Строится на базе систем управления базами данных и систем поддержки принятия решений. Данные, поступающие в хранилище данных, как правило, доступны только для чтения. Данные из OLTP-системы копируются в хранилище данных таким образом, чтобы построение отчетов и OLAP-анализ не использовали ресурсы транзакционной системы и не нарушали ее стабильность. Как правило, данные загружаются в хранилище с определенной периодичностью, поэтому актуальность данных может несколько отставать от OLTP-системы.

- формирование новых агрегатов данных, зависящих от старых, автоматическое;
- доступ к ХД осуществляется на основе многомерного куба.

Альтернативой хранилищу данных является концепция витрины данных (*Data Mart*<sup>1</sup>). Витрины данных — множество тематических БД, содержащих информацию, относящуюся к отдельным информационным аспектам предметной области.

Концепция витрин данных была предложена Forrester Research еще в 1991 г. По мысли авторов, витрины данных — множество тематических баз данных (БД), содержащих информацию, относящуюся к отдельным аспектам деятельности организации.

Концепция имеет ряд несомненных достоинств:

- аналитики видят и работают только с теми данными, которые им реально нужны;
- целевая БД максимально приближена к конечному пользователю;
- витрины данных обычно содержат тематические подмножества заранее агрегированных данных, их проще проектировать и настраивать;
- для реализации витрин данных не требуется высокомоощная вычислительная техника.

Но концепция витрин данных имеет и очень серьезные пробелы. По существу, здесь предполагается реализация территориально распределенной информационной системы с мало контролируемой избыточностью, но не предлагается способов, как обеспечить целостность и непротиворечивость хранимых в ней данных. Идея соединить две концепции — хранилищ данных и витрин данных, по видимому, принадлежит М. Демаресту (М. Demarest), который в 1994 г. предложил объединить две концепции и использовать хранилище данных в качестве единого интегрированного источника данных для витрин данных. На сегодня имеется такое многоуровневое решение:

- первый уровень — общекорпоративная БД на основе реляционной СУБД с нормализованной или слабо денормализованной схемой (детализированные данные);
- второй уровень — БД уровня подразделения (или конечного пользователя), реализуемые на основе многомерной СУБД (агрегированные данные);

<sup>1</sup> Витрина данных, другие варианты: хранилище данных специализированное, киоск данных, рынок данных (от английского *Data Mart*) — чрез хранилища данных, представляющий собой массив тематической, узконаправленной информации, ориентированный, например, на пользователей одной рабочей группы или департамента.

- третий уровень — рабочие места конечных пользователей, на которых непосредственно установлен аналитический инструментарий;
- постепенно становится стандартом де-факто, позволяя наиболее полно реализовать и использовать достоинства каждого из подходов;
- компактное хранение детализированных данных и поддержка очень больших БД, обеспечиваемые реляционными СУБД;
- простота настройки и хорошие времена отклика, при работе с агрегированными данными, обеспечиваемые многомерными СУБД.

Реляционная форма представления данных, используемая в центральной общекорпоративной БД, обеспечивает наиболее компактный способ хранения данных. Современные реляционные СУБД уже умеют работать с БД, имеющими размер порядка нескольких терабайт. Хотя такая центральная система обычно не сможет обеспечить оперативного режима обработки аналитических запросов при использовании новых способов индексации и хранения данных, а также частичной денормализации таблиц, время обработки заранее регламентированных запросов (а в качестве таких можно рассматривать и регламентированные процедуры выгрузки данных в многомерные БД) оказывается вполне приемлемым.

В свою очередь использование многомерных СУБД в узлах нижнего уровня обеспечивает минимальные времена обработки и ответа на нерегламентированные запросы пользователя. Кроме того в некоторых многомерных СУБД имеется возможность хранить данные как на постоянной основе (непосредственно в многомерной БД), так и динамически (на время сеанса) загрузить данные из реляционных БД (на основе регламентированных запросов).

Таким образом, имеется возможность хранить на постоянной основе только те данные, которые наиболее часто запрашиваются в данном узле. Для всех остальных хранятся только описания их структуры и программы их выгрузки из центральной БД. И хотя при первичном обращении к таким виртуальным данным время отклика может оказаться достаточно продолжительным, такое решение обеспечивает высокую гибкость и требует менее дорогих аппаратных средств.

Попытки строить системы принятия решений, которые обращались бы непосредственно к базам данных **систем оперативной обработки транзакций (OLTP-систем)**, оказываются в большинстве случаев неудачными. Во-первых, аналитические запросы «конкурируют» с оперативными транзакциями, блокируя данные и вызывая нехватку ресурсов.

Во-вторых, структура оперативных данных предназначена для эффективной поддержки коротких и частых транзакций и в силу этого слишком сложна для понимания конечными пользователями и, кроме того, не обеспечивает необходимой скорости выполнения аналитических запросов. В-третьих, в организации, как правило, функционирует несколько оперативных систем; каждая со своей базой данных. В этих базах используются различные структуры данных, единицы измерения, способы кодирования и т.д. Для конечного пользователя (аналитика) задача построения какого-либо сводного запроса по нескольким подобным базам данных практически неразрешима.

Для того чтобы обеспечить возможность анализа накопленных данных, организации стали создавать **хранилища данных**, которые представляют собой интегрированные коллекции данных, которые собраны из различных систем оперативного доступа к данным. Хранилища данных становятся основой для построения систем принятия решений. Несмотря на различия в подходах и реализациях, всем хранилищам данных свойственны следующие общие черты<sup>1</sup>:

- **предметная ориентированность.** Информация в хранилище данных организована в соответствии с основными аспектами деятельности предприятия (заказчики, продажи, склад и т.п.); это отличает хранилище данных от оперативной БД, где данные организованы в соответствии с процессами (выписка счетов, отгрузка товара и т.п.). Предметная организация данных в хранилище способствует как значительному упрощению анализа, так и повышению скорости выполнения аналитических запросов. Выражается она, в частности, в использовании иных, чем в оперативных системах, схемах организации данных. В случае хранения данных в реляционной СУБД применяется схема «звезды» (star) или «снежинки» (snowflake). Кроме того данные могут храниться в специальной многомерной СУБД в n-мерных кубах;
- **интегрированность.** Исходные данные извлекаются из оперативных БД, проверяются, очищаются, приводятся к единому виду, в нужной степени агрегируются (т.е. вычисляются суммарные показатели) и загружаются в хранилище. Такие интегрированные данные намного проще анализировать;
- **привязка ко времени.** Данные в хранилище всегда напрямую связаны с определенным периодом времени. Данные, выбранные из оперативных БД, накапливаются в хранилище в виде

<sup>1</sup> Все приведенные ниже особенности принципов построения ХД были впервые сформулированы в 1992 г. «опорно-основателем» хранилищ данных Биллом Инмонем (Bill Inmon) в его книге «Building the Data Warehouse».

«исторических слоев», каждый из которых относится к конкретному периоду времени. Это позволяет анализировать тенденции в развитии бизнеса;

- **неизменяемость.** Попав в определенный «исторический слой» хранилища, данные уже никогда не будут изменены. Это также отличает хранилище от оперативной БД, в которой данные все время меняются, «дышат», и один и тот же запрос, выполненный дважды с интервалом в 10 мин, может дать разные результаты. Стабильность данных также облегчает их анализ.

Англоязычный термин Data Warehousing, который сложно перевести лаконично на русский язык, означает «создание, поддержку, управление и использование хранилища данных» и хорошо подтверждает тот факт, что речь идет о процессе. Цель этого процесса — непрерывная поставка необходимой информации нужным сотрудникам организации. Этот процесс подразумевает постоянное развитие, совершенствование, решение все новых задач и практически никогда не кончается, поэтому его нельзя уместить в более или менее четкие временные рамки, как это можно сделать для разработки традиционных систем оперативного доступа к данным.

Хранилища данных могут быть разбиты на два типа: корпоративные хранилища данных (enterprise data warehouses) и киоски данных (data marts).

**Корпоративные хранилища данных** содержат информацию, относящуюся ко всей корпорации и собранную из множества оперативных источников для консолидированного анализа. Обычно такие хранилища охватывают целый ряд аспектов деятельности корпорации и используются для принятия как тактических, так и стратегических решений. Корпоративное хранилище содержит детальную и обобщающую информацию; его объем может достигать от 50 Гбайт до одного или нескольких терабайт. Стоимость создания и поддержки корпоративных хранилищ может быть очень высокой. Обычно их созданием занимаются централизованные отделы информационных технологий, причем создаются они сверху вниз, т.е. сначала проектируется общая схема, и только затем начинается заполнение данными. Такой процесс может занимать несколько лет.

**Киоски данных** содержат подмножество корпоративных данных и строятся для отделов или подразделений внутри организации. Киоски данных часто строятся силами самого отдела и охватывают конкретный аспект, интересующий сотрудников данного отдела. Киоск данных может получать данные из корпоративного хранилища (зависимый киоск) или, что более распространено, данные могут поступать непосредственно из оперативных источников (независимый киоск).

Киоски и хранилища данных строятся по сходным принципам и используют практически одни и те же технологии.

Основными компонентами хранилища данных являются следующие:

- оперативные источники данных;
- средства проектирования (разработки);
- средства переноса и трансформации данных;
- СУБД;
- средства доступа и анализа данных;
- средства администрирования.

Существуют два архитектурных направления — нормализованные хранилища данных и хранилища с измерениями.

В нормализованных хранилищах, данные находятся в предметно ориентированных таблицах третьей нормальной формы. Нормализованные хранилища характеризуются как простые в создании и управлении, недостатки нормализованных хранилищ — большое количество таблиц как следствие нормализации, из-за чего для получения какой-либо информации нужно делать выборку из многих таблиц одновременно, что приводит к ухудшению производительности системы.

Хранилища с измерениями используют схему «звезда» или схему «снежинка». При этом в центре «звезды» находятся данные (таблица фактов), а измерения образуют лучи звезды. Различные таблицы фактов совместно используют таблицы измерений, что значительно облегчает операции объединения данных из нескольких предметных таблиц фактов (пример: факты продаж и поставок товара). Таблицы данных и соответствующие измерения образуют архитектуру «шина». Измерения часто создаются в третьей нормальной форме, в том числе для протоколирования изменения в измерениях. Основным достоинством хранилищ с измерениями является простота и понятность для разработчиков и пользователей, также благодаря более эффективному хранению данных и формализованным измерениям облегчается и ускоряется доступ к данным, особенно при сложных анализах. Основным недостатком являются более сложные процедуры подготовки и загрузки данных, а также управление и изменение измерений данных.

**Схема «звезды»** (или еще пользуются терминами «схема звездного соединения», «звездоподобная схема», «звездная схема» от английского термина *star schema*) — специальная организация реляционных таблиц, удобная для хранения многомерных показателей (рис. 5.3). Лежит в основе реляционного OLAP. Модель данных состоит из двух типов таблиц: одной таблицы фактов (*fact table*) — центр «звезды» — и нескольких таблиц измерений (*dimension table*) по числу измерений в модели данных — лучи «звезды».

Таблица фактов обычно содержит одну или несколько колонок типа DECIMAL, дающих числовую характеристику какому-то аспекту предметной области (например, объем продаж для торговой компании или сумма платежей для банка), и несколько целочисленных колонок-ключей для доступа к таблицам измерений.

Таблицы измерений расшифровывают ключи, на которые ссылается таблица фактов; например, таблица products измерения «товары» базы данных торговой компании может содержать сведения о названии товара, его производителе, типе товара. За счет использования специальной структуры таблицы измерений реализуется иерархия измерений, в том числе ветвящаяся.

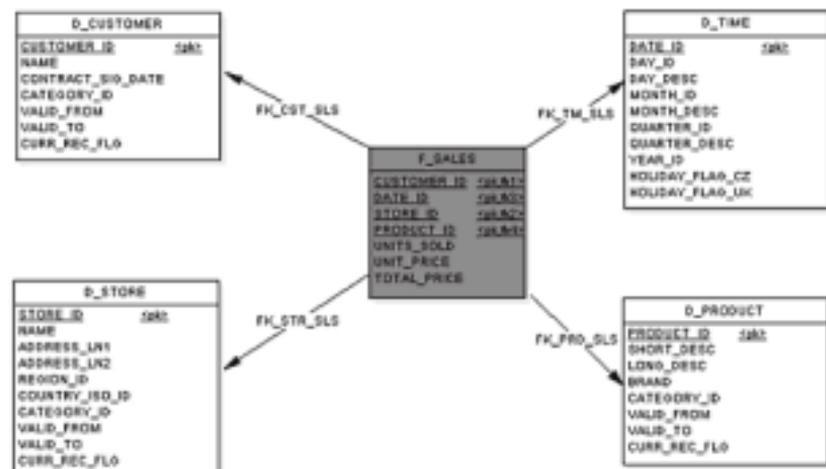


Рис. 5.3. Схема «звезды»

Обычно данные в таблицах-измерениях денормализованы: ценой несколько неэффективного использования дискового пространства удается уменьшить число участвующих в операции соединения таблиц, что обычно приводит к сильному уменьшению времени выполнения запроса. SQL-запрос к схеме «звезда» обычно содержит в себе:

- одно или несколько соединений таблицы фактов с таблицами измерений;
- несколько фильтров (SQL-оператор WHERE), применяемых к таблице фактов или таблицам измерений;
- группировку и агрегирование по требуемым элементам иерархии измерений (dimension elements).

Например:

```
SELECT
d_product.brand,
d_store.country_iso_id,
SUM (f_sales.units_sold) AS summa
FROM
f_sales, d_customer, d_time, d_store, d_product
WHERE
f_sales.customer_id = d_customer.customer_id AND
f_sales.date_id = d_time.date_id AND
f_sales.store_id = d_store.store_id AND
f_sales.product_id = d_product.product_id AND
d_time.year_id = 1997 AND
d_product.category_id = «tv»
GROUP BY
d_product.brand, d_store.country_iso_id
```

Иногда тем не менее требуется произвести нормализацию таблиц-измерений; такая схема носит название «снежинка» (*snowflake schema*). **Схема снежинки** получила свое название за свою форму, в виде которой отображается логическая схема таблиц в многомерной базе данных. Так же, как и в схеме звезды, схема снежинки представлена централизованной таблицей фактов, соединенной с таблицами измерений.

Таблица фактов является основной таблицей хранилища данных. Как правило, она содержит сведения об объектах или событиях, совокупность которых будет в дальнейшем анализироваться. Обычно говорят о четырех наиболее часто встречающихся типах фактов. К ним относятся:

- факты, связанные с транзакциями (от англ. Transaction facts). Они основаны на отдельных событиях (типичными примерами которых являются телефонный звонок или снятие денег со счета с помощью банкомата);
- факты, связанные с «моментальными снимками» (от англ. Snapshot facts). Основаны на состоянии объекта (например, банковского счета) в определенные моменты времени, например на конец дня или месяца. Типичными примерами таких фактов являются объем продаж за день или дневная выручка;
- факты, связанные с элементами документа (от англ. Line-item facts). Основаны на том или ином документе (например, счете за товар или услуги) и содержат подробную информацию об элементах этого документа (например, количестве, цене, проценте скидки);

- факты, связанные с событиями или состоянием объекта (от англ. Event or state facts). Представляют возникновение события без подробностей о нем (например, просто факт продажи или факт отсутствия таковой без иных подробностей).

Таблица фактов, как правило, содержит уникальный составной ключ, объединяющий первичные ключи таблиц измерений. Чаще всего это целочисленные значения либо значения типа «дата/время» в целочисленном формате — ведь таблица фактов может содержать сотни тысяч или даже миллионы записей, и хранить в ней повторяющиеся текстовые описания, как правило, невыгодно — лучше поместить их в меньшие по объему таблицы измерений. При этом как ключевые, так и некоторые неключевые поля должны соответствовать будущим измерениям OLAP-куба. Помимо этого таблица фактов содержит одно или несколько числовых полей, на основании которых в дальнейшем будут получены агрегатные данные.

Отличием является то, что здесь таблицы измерений нормализованы с рядом других связанных измерительных таблиц, в то время как в схеме звезды таблицы измерений полностью денормализованы с каждым измерением, представленным в виде единой таблицы, без соединений на связанные таблицы в схеме снежинки. Чем больше степень нормализации таблиц измерений, тем сложнее выглядит структура схемы снежинки. Создаваемый «эффект снежинки» затрагивает только таблицы измерений и не применим к таблицам фактов. Схема построения «снежинка» представлена на рис. 5.4.

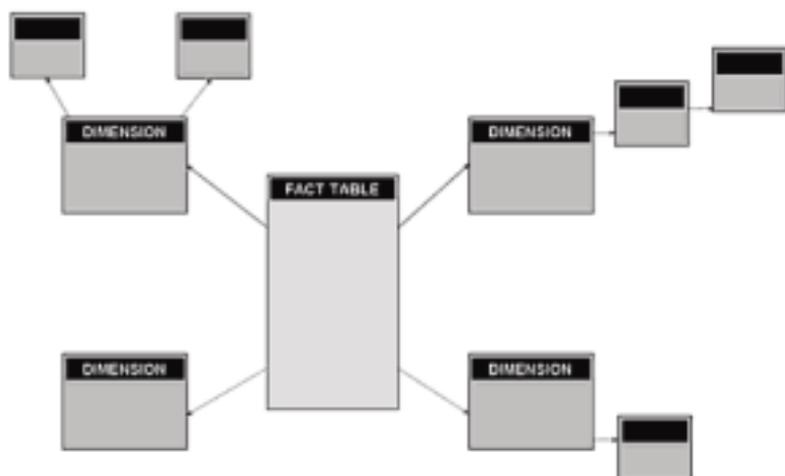


Рис. 5.4. Схема снежинки

Таблица измерений — таблица в структуре многомерной базы данных, которая содержит атрибуты событий, сохраненных в таблице фактов. Атрибуты представляют собой текстовые или иные описания, логически объединенные в одно целое. Например, имя покупателя может являться атрибутом в таблице измерений покупателей, а наименование товара — в таблице измерений товаров. В то время как сумма транзакции является величиной аддитивной, и ее значение должно храниться в таблице фактов.

Таблица фактов связана с таблицами измерений с помощью внешнего ключа. Схема снежинки так же, как и схема звезды, наиболее часто встречается в таких хранилищах данных, для которых скорость получения данных более важна, чем эффективность их манипуляции. Следовательно, таблицы должны быть нормализованы в малой степени, и зачастую разрабатываются с применением не выше третьего уровня нормализации. Схема снежинки образуется из схемы звезды, в случае нормализации таблиц измерений последней.

Решение в сторону использования схемы звезды или же схемы снежинки обуславливается относительной мощностью платформы БД и инструментария для реализации запросов. Схема звезды подходит окружению, в котором инструментарий реализации запросов предоставляет пользователям широкий доступ к структуре таблиц, а также в окружениях, где большинство запросов просты по своей природе. Схема снежинки более подходит для случаев применения более сложного инструментария для реализации запросов, который в большей степени изолирует пользователей от детальной структуры таблиц, а также для среды с множеством запросов сложной структуры. Процессы создания, поддержки и использования хранилищ данных традиционно требовали значительных затрат, что в первую очередь было вызвано высокой стоимостью доступных на рынке специализированных инструментов. Эти инструменты практически не интегрировались между собой, так как были основаны не на открытых и стандартных, а на частных и закрытых протоколах, интерфейсах и т.д. Сложность и дороговизна делали практически невозможным построение хранилищ данных в небольших и средних фирмах, в то время как потребность в анализе данных испытывает любая фирма, независимо от масштаба.

Корпорация Microsoft давно осознала важность направления, связанного с хранилищами данных, и необходимость принятия мер по созданию инструментальной и технологической среды, которая позволила бы минимизировать затраты на создание хранилищ данных и сделала бы этот процесс доступным для массового пользователя. Это привело к созданию **Microsoft Data Warehousing Framework** (рис. 5.5) — спец-

ификации среды создания и использования хранилищ данных. Данная спецификация определяет развитие не только новой линии продуктов Microsoft (например, Microsoft SQL Server), но и технологий, обеспечивающих интеграцию продуктов различных производителей. Открытость среды Microsoft Data Warehousing Framework обеспечила ее поддержку многими производителями ПО, что в свою очередь дает возможность конечным пользователям выбирать наиболее понравившиеся им инструменты для построения своих решений.



Рис. 5.5. Microsoft Data Warehousing Framework

Цель Microsoft Data Warehousing Framework — упростить разработку, внедрение и администрирование решений на основе хранилищ данных. Эта спецификация призвана обеспечить:

- открытую архитектуру, которая легко интегрируется и расширяется третьими фирмами;
- экспорт и импорт гетерогенных данных наряду с их проверкой, очисткой и возможным ведением истории накопления;
- доступ к разделяемым метаданным со стороны процессов разработки хранилища, извлечения и трансформации данных, управления сервером и анализа данных конечными пользователями;

- встроенные службы планирования задач, управление дисковой памятью, мониторинг производительности, оповещение и реакция на события.

**OLE DB — стандарт обмена данными.** Построение хранилищ данных требует, с одной стороны, взаимодействия с различными оперативными БД для извлечения данных, а с другой — обмена данными и метаданными между различными компонентами. И та и другая задача решается крайне сложно — при отсутствии единого интерфейса для доступа к разнородным данным. Но такой интерфейс существует — это **OLE DB**. OLE DB целиком основан на открытой модели **COM** (Component Object Model) и представляет собой набор интерфейсов, которые могут быть использованы, например, в приложениях на Visual C++. Для упрощения использования OLE DB создан набор ActiveX-компонентов — **ActiveX Data Objects (ADO)**. Эти компоненты могут вызываться из приложений на Visual Basic, Access, Excel, встраиваться в активные Web-страницы и т.п.

Практически все компоненты, которые мы будем обсуждать ниже, используют OLE DB для доступа к данным.

OLE DB обеспечивает доступ не только к реляционным данным, но и к таким ресурсам, как почтовые сообщения, файловые каталоги, полнотекстовые индексы и т.п.

**Метаданные.** Одна из наиболее важных задач при построении хранилища данных — интеграция различных компонентов и инструментов, используемых для проектирования, хранения данных, переноса и трансформации, а также анализа данных. Ключевым моментом при такой интеграции является возможность использования разделяемых **метаданных** (т.е. данных о данных).

Центральным компонентом Data Warehousing Framework, представленным на рис. 5.6, является хранилище метаданных (репозиторий) — Microsoft Repository, представляемое как один из компонентов Microsoft SQL Server Microsoft Repository — это база данных, которая хранит описательную информацию о компонентах программного обеспечения и об их отношениях. Microsoft Repository состоит из набора **открытых информационных моделей (Open Information Model, OIM)**, а также набора опубликованных COM-интерфейсов.

Открытые информационные модели — это объектные модели определенного типа информации, при этом они достаточно гибки, чтобы обеспечить поддержку новых типов информации. Корпорация Microsoft, опираясь на сотрудничество с представителями отрасли, уже разработала модели (OIM) для схемы баз данных (Database Schema), преобразования данных (Data Transformations) и OLAP. Последующие модели

будут поддерживать репликацию, планирование задач, семантические модели и информационный справочник, который предназначен для обеспечения метаданными конечного пользователя.

Коалиция метаданных (The Metadata Coalition), отраслевой консорциум 53 производителей, заявила о поддержке Microsoft Repository. От-

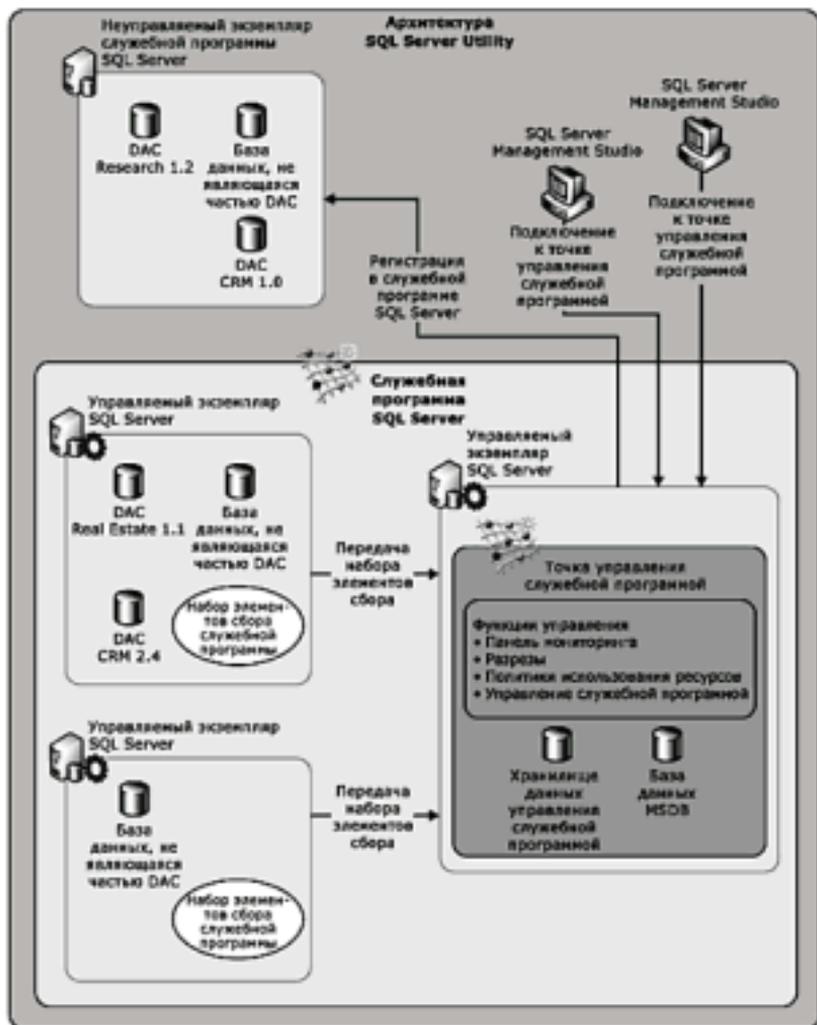


Рис. 5.6. Архитектура MS SQL Server

крытые информационные модели получили широкую поддержку у независимых разработчиков ПО.

Сердцем хранилища данных является, безусловно, СУБД, обеспечивающая надежное и производительное хранение и обработку данных. Как правило, данные из оперативных БД перемещаются в реляционное хранилище, где они становятся доступными для анализа. В дальнейшем при использовании OLAP-средств они могут быть перемещены в многомерную СУБД либо будут выбираться процессором многомерных запросов прямо из реляционных таблиц. Microsoft SQL Server обеспечивает как реляционный, так и многомерный вид хранения. Подробную информацию о Microsoft SQL Server можно найти в разделе «Microsoft SQL Server». Ниже кратко перечислены его основные характеристики: сначала возможности реляционной СУБД, а затем — многомерной.

Microsoft SQL Server обладает целым рядом свойств, делающих его превосходной платформой для построения хранилищ данных:

- поддержка баз данных, размер которых исчисляется терабайтами;
- масштабируемость как «вверх» — в сторону мощнейших современных аппаратных платформ для поддержки очень больших баз данных, так и «вниз» — в сторону серверов небольших рабочих групп и даже настольных и мобильных компьютеров (при этом обеспечивается полная совместимость);
- улучшенная обработка запросов, обеспечивающая оптимизацию и эффективное выполнение сложных запросов, типичных для хранилищ данных, в частности, запросов по схеме типа «звезда»;
- средства параллельного выполнения сложных запросов;
- эффективные средства настройки производительности, загрузки данных и построения индексов;
- распределенные запросы, позволяющие выбирать связанные данные из различных OLE DB-источников;
- надежные и эффективные средства тиражирования данных, незаменимые при поддержке нескольких связанных хранилищ или киосков данных.

Кроме того средства тиражирования по-прежнему остаются одним из механизмов перемещения данных из оперативной БД в хранилище. Ниже рассматривается ряд механизмов, входящих в состав SQL Server.

**OLAP (Online Analytical Processing, оперативная аналитическая обработка)** — все более популярная технология, которая может коренным образом усовершенствовать анализ данных. Microsoft SQL Server OLAP Services — это новый, полнофункциональный OLAP-сервер, поставляемый в составе SQL Server 7.0. OLAP Services включает в себя собствен-

но сервер, доступный по протоколу OLE DB for OLAP, а также клиентский компонент, являющийся поставщиком протокола OLE DB for OLAP и обеспечивающий эффективное изширование и возможность локального сохранения многомерных выборок для их дальнейшего анализа без подключения к OLAP-серверу.

Традиционно OLAP характеризовался дорогим инструментарием и сложным процессом реализации. Включение OLAP-функциональности в Microsoft SQL Server (рис. 5.7) сделает многомерный анализ значительно более приемлемым с точки зрения затрат для небольших и средних организаций. Кроме того небольшие группы или отделы в крупных организациях также смогут в полной мере воспользоваться новыми возможностями анализа — от сложной отчетности до продвинутых систем принятия решений.

Организация извлечения данных из оперативных БД, их очистки, интеграции и помещения в хранилище может потребовать значительных усилий и затрат, если не пользоваться встроенной в Microsoft SQL Server службой — Data Transformation Services (DTS). DTS обладает следующими свойствами:

- он на 100% использует OLE DB для доступа как к источнику, так и к приемнику данных. Благодаря этому DTS может извлекать и преобразовывать данные практически из любых источников (и соответственно помещать их в любые приемники данных);
- для переноса и трансформации данных используется расширяемый набор ActiveX-объектов, которыми легко управлять при помощи языка сценариев, например, VBScript или JavaScript. Таким образом, имеются практически неограниченные возможности управления переносом и преобразованием данных;
- DTS способен интегрироваться с Microsoft Repository для использования метаданных об источнике, приемнике и схеме преобразования данных;
- задания по переносу и преобразованию данных, которые могут включать в себя множество последовательных шагов, оформляются в виде пакетов (DTS Package), которые могут быть сохранены в хранилище метаданных (Repository), в базе SQL Server или в файле. Пакеты могут затем автоматически выполняться по расписанию при помощи сервиса SQL Server Agent.

Именно в сфере средств представления и анализа данных следует ожидать (и уже можно видеть) наибольшее количество продуктов, предлагаемых третьими фирмами, хотя и Microsoft предлагает здесь не только базовые технологии, но и средства для конечного пользователя. К ним относятся компоненты нового поколения Microsoft Office — пре-

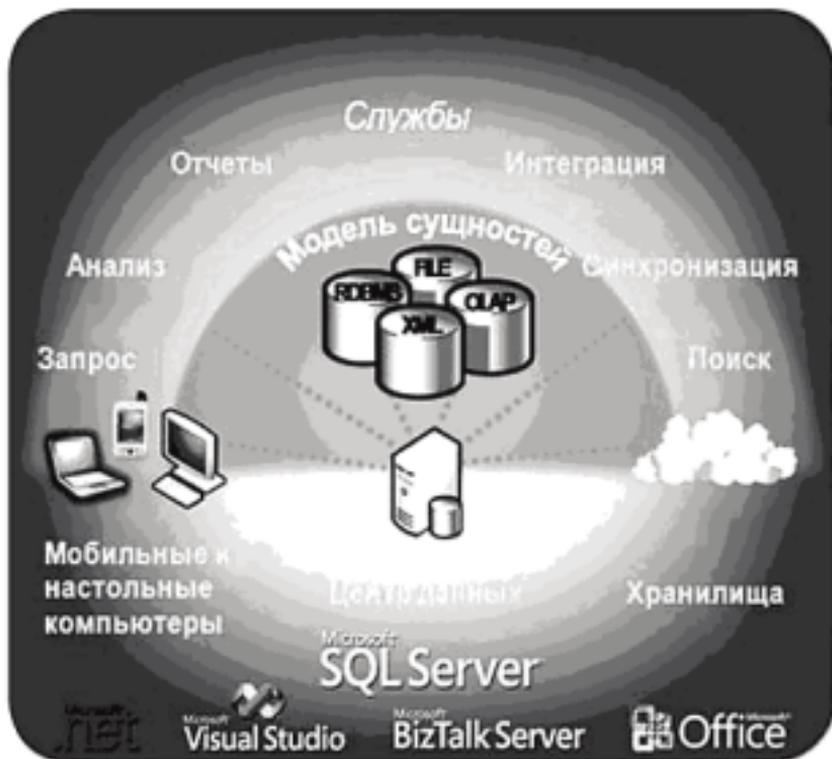


Рис. 5.7. Архитектура OLAP для MS SQL Server

где всего Microsoft Excel. Его популярное средство анализа данных PivotTable® теперь сможет задействовать всю мощь OLAP-сервера, подключаясь к нему через упоминавшийся выше клиентский компонент PivotTable Services.

Еще одно средство от Microsoft — English Query — позволяет строить запросы к SQL Server на естественном языке (по-английски).

Существенной составляющей расходов на внедрение хранилища данных являются расходы на текущее сопровождение и администрирование хранилища. Средства администрирования, в том числе средства автоматизации выполнения административных задач, предусмотренные Data Warehousing Framework и включенные в состав Microsoft SQL Server, позволяют значительно сократить эти расходы.

Единой средой администрирования различных компонентов является Microsoft Management Console (рис. 5.8). Средства управления каждым конкретным компонентом (например, SQL Server или OLAP Services) представляют собой так называемый snap-in, т.е. модуль администрирования, использующий единые средства пользовательского интерфейса.



Рис. 5.8. Установка оснастки MMC для Microsoft SQL Server

Средство управления SQL Server — SQL Enterprise Manager (рис. 5.9) включает в себя более 25 программ-мастеров (Wizards), помогающих не слишком искусственному администратору решать самые важные задачи, в том числе создавать и копировать базы данных, производить настройку тиранирования, импорт (экспорт) данных, управлять правами пользователей и т.п.

Кроме того в SQL Enterprise Manager входят средства создания и редактирования графических диаграмм баз данных, значительно облегчающих создание и модификацию структуры хранилища. Средства автоматизации администрирования позволяют создавать многоступенчатые задания, состоящие как из команд языка Transact-SQL, так и из сценариев на языках VBScript или JavaScript. При этом выполнение последующих шагов может быть поставлено в зависимость от результатов выполнения предыдущих. Эти задания могут охватывать множество серверов и выполняются по заданному расписанию.

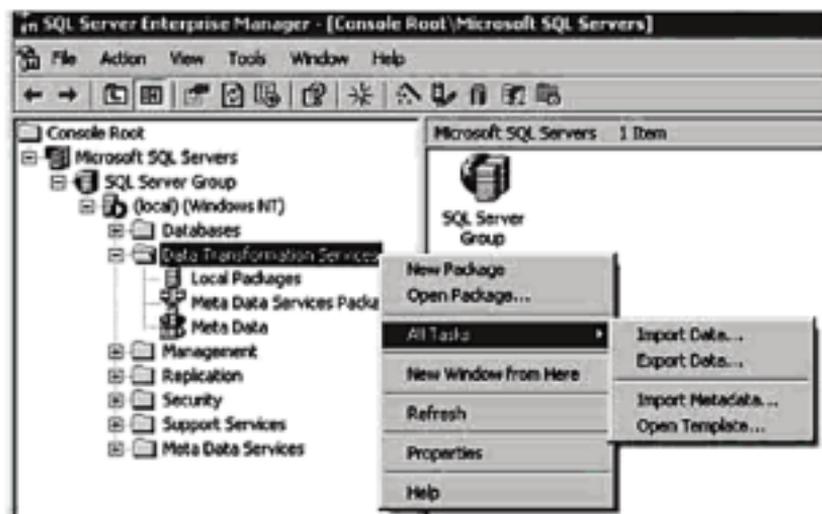


Рис. 5.9. SQL Enterprise Manager

## Контрольные вопросы

1. Каковы преимущества электронной почты?
2. Что собой представляет электронный почтовый ящик?
3. Из каких частей состоит адрес электронной почты? Приведите пример.
4. Что не рекомендуется указывать в пароле?
5. В каком режиме чаще осуществляется работа с электронной почтой?
6. Как работать с адресной книгой?
7. Как переложить письмо в другую папку?
8. Как прикрепить файл к письму, какие ограничения накладываются на прикрепляемые файлы?
9. Какие сервисные услуги предлагает выбранный почтовый сервер по обработке почтового сообщения?
10. Какие вы знаете бесплатные почтовые сервисы в сети «Интернет»?

# ТЕХНОЛОГИИ ЗАЩИТЫ ДАННЫХ

Новые ИТ активно внедряются во все сферы деятельности человека. Появление локальных и глобальных сетей передачи данных предоставило пользователям компьютеров новые возможности для оперативного обмена информацией. Развитие Интернета привело к использованию глобальных сетей передачи данных в повседневной жизни практически каждого человека. По мере развития и усложнения средств, методов и форм автоматизации процессов обработки информации повышается зависимость общества от степени безопасности используемых им ИТ.

## 6.1. Информационная безопасность и защита информации

Современные методы обработки, передачи и накопления информации способствовали появлению угроз, связанных с возможностью потери, искажения и раскрытия данных, адресованных или принадлежащих конечным пользователям. Поэтому обеспечение информационной безопасности компьютерных систем и сетей является одним из ведущих направлений развития ИТ. Информационная безопасность компьютерных систем достигается обеспечением конфиденциальности, целостности и достоверности обрабатываемых данных, а также доступности и целостности информационных компонентов и ресурсов системы.

Корпоративные сети относятся к распределенным автоматизированным системам (АС), осуществляющим обработку информации. Обеспечение безопасности АС предполагает организацию противодействия любому несанкционированному вторжению в процесс функционирова-

ния АС, а также попыткам модификации, хищения, выведения из строя или разрушения ее компонентов, т.е. защиту всех компонентов АС — аппаратных средств, программного обеспечения (ПО), данных и персонала. Конкретный подход к проблеме обеспечения безопасности основан на разработанной для АС политике безопасности.

## 6.2. Криптографические методы защиты данных

Безопасность данных означает их конфиденциальность, целостность и подлинность. Критерии безопасности данных могут быть определены следующим образом.

Конфиденциальность данных предполагает их доступность только для тех лиц, которые имеют на это соответствующие полномочия. Под обеспечением конфиденциальности информации понимается создание таких условий, при которых понять содержание передаваемых данных может только законный получатель, которому данная информация предназначена.

Целостность информации предполагает ее неизменность в процессе передачи от отправителя к получателю. Под обеспечением целостности информации понимается достижение идентичности отправляемых и принимаемых данных.

Подлинность информации предполагает соответствие этой информации ее явному описанию и содержанию, в частности, соответствие действительным характеристикам указанных: отправителя, времени отправления и содержания. Обеспечение подлинности информации, реализуемое на основе аутентификации, состоит в достоверном установлении отправителя, а также защите информации от изменения при ее передаче от отправителя к получателю.

Своевременно обнаруженное нарушение подлинности и целостности полученного сообщения позволяет предотвратить отрицательные последствия, связанные с дальнейшим использованием такого искаженного сообщения.

Криптография является методологической основой современных систем обеспечения безопасности информации в компьютерных системах и сетях. Исторически криптография (в переводе с греческого этот термин означает «тайнопись») зародилась как способ скрытой передачи сообщений. Криптография представляет собой совокупность методов преобразования данных, направленных на то, чтобы защитить эти данные, сделав их бесполезными для незаконных пользователей. Такие преобразования обеспечивают решение трех главных проблем защиты

данных: обеспечение конфиденциальности, целостности и подлинности передаваемых или сохраняемых данных.

Для обеспечения безопасности данных необходимо поддерживать три основные функции:

- защиту конфиденциальности передаваемых или хранимых в памяти данных;
- подтверждение целостности и подлинности данных;
- аутентификацию абонентов при входе в систему и при установлении соединения.

Для реализации указанных функций используются криптографические технологии шифрования, цифровой подписи и аутентификации.

Конфиденциальность обеспечивается с помощью алгоритмов и методов симметричного и асимметричного шифрования, а также путем взаимной аутентификации абонентов на основе многоразовых и однократных паролей, цифровых сертификатов, смарт-карт и т.п.

Целостность и подлинность передаваемых данных обычно достигается с помощью различных вариантов технологии электронной подписи, основанных на односторонних функциях и асимметричных методах шифрования.

Аутентификация позволяет устанавливать соединения только между легальными пользователями и предотвращает доступ к средствам сети нежелательных лиц. Абонентам, доказавшим свою легальность (аутентичность), предоставляются разрешенные виды сетевого обслуживания.

Обеспечение конфиденциальности, целостности и подлинности передаваемых и сохраняемых данных осуществляется прежде всего использованием криптографических способов и средств защиты информации. Основой большинства криптографических средств защиты информации является шифрование данных.

Преобразование шифрования может быть симметричным или асимметричным относительно преобразования расшифровывания. Соответственно различают два класса криптосистем:

- симметричные криптосистемы (с единым ключом);
- асимметричные криптосистемы (с двумя ключами).

Первыми появились симметричные криптографические системы. В симметричной криптосистеме шифрования используется один и тот же ключ для зашифровывания и расшифровывания информации. Это означает, что любой, кто имеет доступ к ключу шифрования, может расшифровать сообщение.

Соответственно с целью предотвращения несанкционированного раскрытия зашифрованной информации все ключи шифрования в симметричных криптосистемах должны держаться в секрете. Именно поэтому

симметричные криптосистемы называют *криптосистемами с секретным ключом* — ключ шифрования должен быть доступен только тем, кому предназначено сообщение. Симметричные криптосистемы называют еще *одноключевыми криптографическими системами*, или *криптосистемами с закрытым ключом*. Схема симметричной криптосистемы шифрования показана на рис. 6.1.



Рис. 6.1. Схема симметричной криптосистемы шифрования

Данные криптосистемы характеризуются наиболее высокой скоростью шифрования, и с их помощью обеспечиваются как конфиденциальность и подлинность, так и целостность передаваемой информации. Конфиденциальность передачи информации с помощью симметричной криптосистемы зависит от надежности шифра и обеспечения конфиденциальности ключа шифрования.

Обычно ключ шифрования представляет собой файл или массив данных и хранится на персональном ключевом носителе, например, флэш-накопителе или смарт-карте; обязательно принятие мер, обеспечивающих недоступность персонального ключевых носителя кому-либо, кроме его владельца.

Подлинность обеспечивается за счет того, что без предварительного расшифрования практически невозможно осуществить смысловую модификацию и подлог криптографически закрытого сообщения. Фальшивое сообщение не может быть правильно зашифровано без знания секретного ключа.

Целостность данных обеспечивается присоединением к передаваемым данным специального кода (имитовставки), вырабатываемой по секретному ключу. Имитовставка является разновидностью контрольной суммы, т.е. некоторой эталонной характеристикой сообщения, по которой осуществляется проверка целостности последнего. Алгоритм фор-

мирования имитовставки должен обеспечивать ее зависимость по некоторому сложному криптографическому закону от каждого бита сообщения. Проверка целостности сообщения выполняется получателем сообщения путем выработки по секретному ключу имитовставки, соответствующей полученному сообщению, и ее сравнения с полученным значением имитовставки. При совпадении делается вывод о том, что информация не была модифицирована на пути от отправителя к получателю.

Симметричное шифрование идеально подходит для шифрования информации «для себя», например, с целью предотвращения НСД к ней в отсутствие владельца. Это может быть как архивное шифрование выбранных файлов, так и прозрачное (автоматическое) шифрование целых логических или физических дисков.

Обладая высокой скоростью шифрования, одноключевые криптосистемы позволяют решать многие важные задачи защиты информации. Однако автономное использование симметричных криптосистем в компьютерных сетях порождает проблему распределения ключей шифрования между пользователями.

Асимметричные криптографические системы были разработаны в 1970-х гг. Принципиальное отличие асимметричной криптосистемы от криптосистемы симметричного шифрования состоит в том, что для шифрования информации и ее последующего расшифровывания используются различные ключи:

- открытый ключ  $K$  используется для шифрования информации, вычисляется из секретного ключа  $k$ ;
- секретный ключ  $k$  используется для расшифровывания информации, зашифрованной с помощью парного ему открытого ключа  $K$ .

Эти ключи различаются таким образом, что с помощью вычислений нельзя вывести секретный ключ  $k$  из открытого ключа  $K$ . Поэтому открытый ключ  $K$  может свободно передаваться по каналам связи.

Асимметричные системы называют также двухключевыми криптографическими системами, или криптосистемами с открытым ключом.

Обобщенная схема асимметричной криптосистемы шифрования с открытым ключом показана на рис. 6.2.

Для криптографического закрытия и последующего расшифровывания передаваемой информации используются открытый и секретный ключи получателя  $B$  сообщения.

В качестве ключа зашифровывания должен использоваться открытый ключ получателя, а в качестве ключа расшифровывания — его секретный ключ.

Секретный и открытый ключи генерируются попарно. Секретный ключ должен оставаться у его владельца и быть надежно защищен от НСД (аналогично ключу шифрования в симметричных алгоритмах). Копия открытого ключа должна находиться у каждого абонента криптографической сети, с которым обменивается информацией владелец секретного ключа.



Рис. 6.2. Обобщенная схема асимметричной криптосистемы шифрования

Совместное использование этих криптосистем позволяет эффективно реализовывать такую базовую функцию защиты, как криптографическое закрытие передаваемой информации с целью обеспечения ее конфиденциальности. Комбинированное применение симметричного и асимметричного шифрования устраняет основные недостатки, присущие обоим методам, и позволяет сочетать преимущества высокой секретности, предоставляемые асимметричными криптосистемами с открытым ключом, с преимуществами высокой скорости работы, присущими симметричным криптосистемам с секретным ключом.

Метод комбинированного использования симметричного и асимметричного шифрования заключается в следующем.

Симметричную криптосистему применяют для шифрования исходного открытого текста, а асимметричную криптосистему с открытым ключом применяют только для шифрования секретного ключа симметричной криптосистемы. В результате асимметричная криптосистема с открытым ключом не заменяет, а лишь дополняет симметричную криптосистему с секретным ключом, позволяя повысить в целом защищенность передаваемой информации. Такой подход иногда называют схемой электронного «цифрового конверта» (рис. 6.3).



Рис. 6.3. Схема шифрования сообщения комбинированным методом

Электронная цифровая подпись используется для аутентификации текстов, передаваемых по телекоммуникационным каналам. При таком обмене существенно снижаются затраты на обработку и хранение документов, упрощается их поиск. Но возникает проблема аутентификации автора электронного документа и самого документа, т.е. установления подлинности автора и отсутствия изменений в полученном электронном документе.

Целью аутентификации электронных документов является их защита от возможных видов злоумышленных действий, к которым относятся:

- активный перехват — нарушитель, подключившийся к сети, перехватывает документы (файлы) и изменяет их;
- маскарад — абонент С посылает документ абоненту В от имени абонента А;
- ренегатство — абонент А заявляет, что не посылал сообщения абоненту В, хотя на самом деле послал;
- подмена — абонент В изменяет или формирует новый документ и заявляет, что получил его от абонента А;
- повтор — абонент С повторяет ранее переданный документ, который абонент А посылал абоненту В.

Эти виды злоумышленных действий могут нанести существенный ущерб банковским и коммерческим структурам, государственным предприятиям и организациям, частным лицам, применяющим в своей деятельности компьютерные ИТ.

Проблему проверки целостности сообщения и подлинности автора сообщения позволяет эффективно решить методология электронной цифровой подписи.

### 6.3. Компьютерные вирусы и защита от них

На сегодняшний день известны десятки тысяч различных компьютерных вирусов. Несмотря на такое изобилие, число типов вирусов, отличающихся друг от друга механизмом распространения и принципом действия, достаточно ограничено. Существуют и комбинированные вирусы, которые можно отнести одновременно к нескольким типам. Вирусы можно разделить на классы:

- по среде обитания;
- операционной системе (ОС);
- особенностям алгоритма работы;
- деструктивным возможностям.

Основной и наиболее распространенной классификацией компьютерных вирусов является классификация по среде обитания, или по типам объектов компьютерной системы, в которые внедряются вирусы (рис. 6.4). По среде обитания компьютерные вирусы можно разделить:

- на файловые;
- загрузочные;
- макровирусы;
- сетевые.

Файловые вирусы либо внедряются в выполняемые файлы (наиболее распространенный тип вирусов) различными способами, либо создают файлы-двойники (компаньон-вирусы), либо используют особенности организации файловой системы (link-вирусы).

Загрузочные вирусы записывают себя либо в загрузочный сектор диска (boot-сектор), либо в сектор, содержащий системный загрузчик винчестера (Master Boot Record). Загрузочные вирусы замещают код программы, получающей управление при загрузке системы. В результате при перезагрузке управление передается вирусу. При этом оригинальный boot-сектор обычно переносится в какой-либо другой сектор диска. Иногда загрузочные вирусы называют бутовыми вирусами.

Макровирусы заражают макропрограммы и файлы документов современных систем обработки информации, в частности, файлы-документы и электронные таблицы популярных редакторов Microsoft Word, Microsoft Excel и др. Для размножения макровирусы используют возможности макроязыков и при их помощи переносят себя из одного

зараженного файла в другие. Вирусы этого типа получают управление при открытии зараженного файла и инфицируют файлы, к которым впоследствии идет обращение из соответствующего офисного приложения.

Сетевые вирусы используют для своего распространения протоколы или команды компьютерных сетей и электронной почты. Иногда сетевые вирусы называют программами типа «червь». Сетевые черви подразделяются на Интернет-черви (распространяются по Интернету), LAN-черви (распространяются по локальной сети), IRC-черви Internet Relay Chat (распространяются через чаты). Существуют также смешанные типы, которые совмещают в себе сразу несколько технологий.



Рис. 6.4. Классификация компьютерных вирусов по среде обитания

Существует много комбинированных типов компьютерных вирусов, например, известен сетевой макровирус, который заражает редактируемые документы, а также рассылает свои копии по электронной почте. В качестве другого примера вирусов комбинированного типа можно указать файлово-загрузочные вирусы, заражающие как файлы, так и за-

грузочные секторы дисков. Такие вирусы имеют усложненный алгоритм работы и применяют своеобразные методы проникновения в систему.

Другим признаком деления компьютерных вирусов на классы является операционная система, объекты которой подвергаются заражению. Каждый файловый или сетевой вирус заражает файлы какой-либо одной или нескольких ОС — DOS, Windows 95/98, Windows NT/2000 и т.д. Макровирусы заражают файлы форматов Word, Excel, Microsoft Office. На определенные форматы расположения системных данных в загрузочных секторах дисков также ориентированы загрузочные вирусы.

Естественно, эти схемы классификации не являются единственно возможными, существует много различных схем типизации вирусов. Однако ограничимся пока классификацией компьютерных вирусов по среде обитания, поскольку она является базовой, и перейдем к рассмотрению общих принципов функционирования вирусов. Анализ основных этапов «жизненного цикла» этих вредоносных программ позволяет выделить их различные признаки и особенности, которые могут быть положены в основу дополнительных классификаций.

Для защиты от компьютерных вирусов могут использоваться:

- общие методы и средства защиты информации;
- специализированные программы для защиты от вирусов;
- профилактические меры, позволяющие уменьшить вероятность заражения вирусами.

Общие средства защиты информации полезны не только для защиты от вирусов. Они используются так же, как страховка от физической порчи дисков, неправильно работающих программ или ошибочных действий пользователя. Существуют две основные разновидности этих средств:

- средства копирования информации (применяются для создания копий файлов и системных областей дисков);
- средства разграничения доступа (предотвращают несанкционированное использование информации, в частности, обеспечивают защиту от изменений программ и данных вирусами, неправильно работающими программами и ошибочными действиями пользователей).

При заражении компьютера вирусом важно его обнаружить. К внешним признакам проявления деятельности вирусов можно отнести следующие:

- вывод на экран непредусмотренных сообщений или изображений;
- подача непредусмотренных звуковых сигналов;
- изменение даты и времени модификации файлов;
- исчезновение файлов и каталогов или искажение их содержимого;

- частые зависания и сбои в работе компьютера;
- медленная работа компьютера;
- невозможность загрузки ОС;
- существенное уменьшение размера свободной оперативной памяти;
- прекращение работы или неправильная работа ранее успешно функционировавших программ;
- изменение размеров файлов;
- неожиданное значительное увеличение количества файлов на диске.

Однако следует заметить, что перечисленные выше явления необязательно вызываются действиями вируса, они могут быть следствием и других причин. Поэтому правильная диагностика состояния компьютера всегда затруднена и обычно требует привлечения специализированных программ.

Для обнаружения и защиты от компьютерных вирусов разработано несколько видов специальных программ, которые позволяют обнаруживать и уничтожать компьютерные вирусы. Такие программы называются *антивирусными*. Практически все антивирусные программы обеспечивают автоматическое восстановление зараженных программ и загрузочных секторов. Антивирусные программы используют различные методы обнаружения вирусов.

К основным методам обнаружения компьютерных вирусов можно отнести следующие:

- метод сравнения с эталоном;
- эвристический анализ;
- антивирусный мониторинг;
- метод обнаружения изменений;
- встраивание антивирусов в BIOS компьютера и др.

**Метод сравнения с эталоном.** Самый простой метод обнаружения заключается в том, что для поиска известных вирусов используются так называемые маски. Маской вируса является некоторая постоянная последовательность кода, специфичная для этого конкретного вируса. Антивирусная программа последовательно просматривает (сканирует) проверяемые файлы в поиске масок известных вирусов. Антивирусные сканеры способны найти только уже известные вирусы, для которых определена маска.

Если вирус не содержит постоянной маски или длина этой маски недостаточно велика, то используются другие методы. Применение простых сканеров не защищает компьютер от проникновения новых вирусов. Для шифрующихся и полиморфных вирусов, способных полностью

изменять свой код при заражении новой программой или загрузочного сектора, невозможно выделить маску, поэтому антивирусные сканеры их не обнаруживают.

**Эвристический анализ.** Для того чтобы размножаться, компьютерный вирус должен совершать какие-то конкретные действия: копирование в память, запись в секторы и т.д. Эвристический анализатор (который является частью антивирусного ядра) содержит список таких действий и проверяет программы и загрузочные секторы дисков, пытаясь обнаружить в них код, характерный для вирусов. Эвристический анализатор может обнаружить, например, что проверяемая программа устанавливает резидентный модуль в памяти или записывает данные в исполнимый файл программы. Обнаружив зараженный файл, анализатор обычно выводит сообщение на экране монитора и делает запись в собственном или системном журнале. В зависимости от настроек антивирус может также направлять сообщение об обнаруженном вирусе администратору сети. Эвристический анализ позволяет обнаруживать неизвестные ранее вирусы. Первый эвристический анализатор появился в начале 1990-х гг. Практически все современные антивирусные программы реализуют собственные методы эвристического анализа. В качестве примера такой программы можно указать сканер McAfee VirusScan.

**Антивирусный мониторинг.** Суть данного метода состоит в том, что в памяти компьютера постоянно находится антивирусная программа, осуществляющая мониторинг всех подозрительных действий, выполняемых другими программами. Антивирусный мониторинг позволяет проверять все запускаемые программы, создаваемые, открываемые и сохраняемые документы, файлы программ и документов, полученные через Интернет или скопированные на жесткий диск с флэш-накопителя либо компакт-диска. Антивирусный монитор сообщит пользователю, если какая-либо программа попытается выполнить потенциально опасное действие. Пример такой программы — сторож Spider Guard, который входит в комплект сканера Doctor Web и выполняет функции антивирусного монитора.

**Метод обнаружения изменений.** При реализации этого метода антивирусные программы, называемые ревизорами диска, запоминают предварительно характеристики всех областей диска, которые могут подвергнуться нападению, а затем периодически проверяют их. Заражая компьютер, вирус изменяет содержимое жесткого диска: например, дописывает свой код в файл программы или документа, добавляет вызов программы-вируса в файл AUTOEXEC.BAT, изменяет загрузочный сектор, создает файл-спутник. При сопоставлении значений характери-

стик областей диска антивирусная программа может обнаружить изменения, сделанные как известным, так и неизвестным вирусом.

**Встраивание антивирусов в BIOS компьютера.** В системные платы компьютеров встраивают простейшие средства защиты от вирусов. Эти средства позволяют контролировать все обращения к главной загрузочной записи жестких дисков, а также к загрузочным секторам дисков. Если какая-либо программа пытается изменить содержимое загрузочных секторов, срабатывает защита, и пользователь получает соответствующее предупреждение. Однако эта защита не очень надежна. Известны вирусы, которые пытаются отключить антивирусный контроль BIOS, изменяя некоторые ячейки в энергонезависимой памяти (CMOS-памяти) компьютера.

Виды антивирусных программ. Различают следующие виды антивирусных программ:

- программы-фаги (сканеры);
- программы-ревизоры (CRC-сканеры);
- программы-блокировщики;
- программы-иммунизаторы.

**Программы-фаги (сканеры)** используют для обнаружения вирусов метод сравнения с эталоном, метод эвристического анализа и некоторые другие методы. Программы-фаги осуществляют поиск характерной для конкретного вируса маски путем сканирования в оперативной памяти и в файлах и при обнаружении выдают соответствующее сообщение. Программы-фаги не только находят зараженные вирусами файлы, но и «лечат» их, т.е. удаляют из файла тело программы-вируса, возвращая файлы в исходное состояние. В начале работы программы-фаги сканируют оперативную память, обнаруживают вирусы и уничтожают их и только затем переходят к «лечению» файлов. Среди фагов выделяют полифаги — программы-фаги, предназначенные для поиска и уничтожения большого числа вирусов.

Программы-фаги можно разделить на две категории — универсальные и специализированные сканеры. *Универсальные сканеры* рассчитаны на поиск и обезвреживание всех типов вирусов вне зависимости от ОС, на работу в которой рассчитан сканер. *Специализированные сканеры* предназначены для обезвреживания ограниченного числа вирусов или только одного их класса, например, макровирусов. Специализированные сканеры, рассчитанные только на макровирусы, оказываются более удобным и надежным решением для защиты систем документооборота в средах MS Word и MS Excel.

Программы-фаги делятся также на *резидентные мониторы*, производящие сканирование «на лету», и *нерезидентные сканеры*, обеспечи-

вающие проверку системы только по запросу. Резидентные мониторы обеспечивают более надежную защиту системы, поскольку они немедленно реагируют на появление вируса, в то время как нерезидентный сканер способен опознать вирус только во время своего очередного запуска.

К достоинствам программ-фагов всех типов относится их универсальность. К недостаткам следует отнести относительно небольшую скорость поиска вирусов и относительно большие размеры антивирусных баз.

Наиболее известные программы-фаги: Aidstest, Scan, Norton Antivirus, Doctor Web. Учитывая, что постоянно появляются новые вирусы, программы-фаги быстро устаревают, и требуется регулярное обновление версий.

**Программы-ревизоры (CRC-сканеры)** используют для поиска вирусов метод обнаружения изменений. Принцип работы CRC-сканеров основан на подсчете CRC-сумм (кодов циклического контроля) для присутствующих на диске файлов (системных секторов). Эти CRC-суммы, а также некоторая другая информация (длины файлов, даты их последней модификации и др.) затем сохраняются в БД антивируса. При последующем запуске CRC-сканеры сверяют данные, содержащиеся в БД, с реально подсчитанными значениями. Если информация о файле, записанная в БД, не совпадает с реальными значениями, то CRC-сканеры сигнализируют о том, что файл был изменен или заражен вирусом. Как правило, сравнение состояний производят сразу после загрузки ОС.

CRC-сканеры, использующие алгоритмы анти-стелс, являются довольно мощным средством против вирусов: практически 100% вирусов оказываются обнаруженными почти сразу после их появления на компьютере. Однако у CRC-сканеров имеется недостаток, заметно снижающий их эффективность: они не могут определить вирус в новых файлах (в электронной почте, на переносных накопителях, в файлах, восстанавливаемых из backup или при распаковке файлов из архива), поскольку в их БД отсутствует информация об этих файлах.

К числу CRC-сканеров относится широко распространенная в России программа ADInf (*Advanced Diskinfoscope*) и ревизор AVP Inspector. Вместе с ADInf применяется лечащий модуль ADInf Cure Module (ADInfExt), который использует собранную ранее информацию о файлах для их восстановления после поражения неизвестными вирусами. В состав ревизора AVP Inspector также входит лечащий модуль, способный удалять вирусы.

**Программы-блокировщики** реализуют метод антивирусного мониторинга. Антивирусные блокировщики — это резидентные программы, перехватывающие «вирусо-опасные» ситуации и сообщающие об этом пользователю. К «вирусоопасным» ситуациям относятся вызовы, которые

характерны для вирусов в моменты их размножения (вызовы на открытие для записи в выполняемые файлы, запись в загрузочные секторы дисков или MBR винчестера, попытки программ остаться резидентно и т.п.).

При попытке какой-либо программы произвести указанные действия блокировщик посылает пользователю сообщение и предлагает запретить соответствующее действие. К достоинствам блокировщиков относится их способность обнаруживать и останавливать вирус на самой ранней стадии его размножения, что бывает особенно полезно в случаях, когда регулярно появляется давно известный вирус. Однако они не «лечат» файлы и диски. Для уничтожения вирусов требуется применять другие программы, например, фаги. К недостаткам блокировщиков можно отнести существование путей обхода их защиты и их «назойливость» (например, они постоянно выдают предупреждение о любой попытке копирования исполняемого файла).

Следует отметить, что созданы антивирусные блокировщики, выполненные в виде аппаратных компонентов компьютера. Наиболее распространенной является встроенная в BIOS защита от записи в MBR винчестера.

**Программы-иммунизаторы** — это программы, предотвращающие заражение файлов. Иммунизаторы делятся на два типа: иммунизаторы, сообщающие о заражении, и иммунизаторы, блокирующие заражение каким-либо типом вируса. Иммунизаторы первого типа обычно записываются в конец файлов и при запуске файла каждый раз проверяют его на изменение. У таких иммунизаторов имеется один серьезный недостаток — они не могут обнаружить заражение стелс-вирусом. Поэтому этот тип иммунизаторов практически не используются в настоящее время.

Иммунизатор второго типа защищает систему от поражения вирусом определенного вида. Он модифицирует программу или диск таким образом, чтобы это не отразилось на их работе, вирус при этом воспринимает их зараженными и поэтому не внедряется. Такой тип иммунизации не может быть универсальным, поскольку нельзя иммунизировать файлы от всех известных вирусов. Однако в качестве полумеры подобные иммунизаторы могут вполне надежно защитить компьютер от нового неизвестного вируса вплоть до того момента, когда он будет определяться антивирусными сканерами.

Качество антивирусной программы можно оценить по нескольким критериям:

- надежность и удобство работы — отсутствие «зависаний» антивируса и прочих технических проблем, требующих от пользователя специальной подготовки;

- качество обнаружения вирусов всех распространенных типов, сканирование внутри файлов-документов (таблиц) (MS Word, Excel, Office), упакованных и архивированных файлов; возможность лечения зараженных объектов;
- существование версий антивируса под все популярные платформы (DOS, Windows NT, Novell NetWare, OS/2, Alpha, Linux и т.д.); наличие режимов сканирования «по запросу» и «на лету», существование серверных версий с возможностью администрирования сети;
- скорость работы и другие полезные особенности.

Надежность работы антивируса является наиболее важным критерием, поскольку даже «абсолютный» антивирус может оказаться бесполезным, если он не в состоянии довести процесс сканирования до конца, т.е. «повиснет» и не проверит часть дисков и файлов, и в результате вирус останется незамеченным в системе.

Качество обнаружения вирусов стоит на следующем месте по вполне естественной причине. Главная обязанность антивирусных программ — обнаруживать 100% вирусов и лечить их. При этом антивирусная программа не должна иметь высокий уровень ложных срабатываний.

Следующий по важности критерий — многоплатформенность антивируса, поскольку только программа, рассчитанная на конкретную ОС, может полностью использовать функции этой системы. Моментальная и принудительная проверка приходящих на компьютер файлов и подключаемых накопителей — это практически 100%-ная гарантия от заражения вирусом. Если в серверном варианте антивируса присутствует возможность антивирусного администрирования сети, то его ценность еще более возрастает.

Скорость работы также является важным критерием качества антивирусной программы. В разных антивирусах используются различные алгоритмы поиска вирусов, один алгоритм может оказаться более быстрым и качественным, другой — медленным и менее качественным.

Профилактические меры защиты. Своевременное обнаружение зараженных вирусами файлов и дисков, полное уничтожение обнаруженных вирусов на каждом компьютере позволяют избежать распространения вирусной эпидемии на другие компьютеры. Абсолютно надежных программ, гарантирующих обнаружение и уничтожение любого вируса, не существует. Важным методом борьбы с компьютерными вирусами является своевременная профилактика. Чтобы существенно уменьшить вероятность заражения вирусом и обеспечить надежное хранение информации на дисках, необходимо выполнять следующие меры профилактики:

- применять только лицензионное ПО;
- оснастить компьютер современными антивирусными программами и постоянно возобновлять их версии;
- всегда проверять электронные носители информации на наличие вирусов (запуская антивирусные программы своего компьютера) перед считыванием с них информации, записанной на других компьютерах;
- при переносе на свой компьютер файлов в архивированном виде проверять их сразу же после разархивации на жестком диске, ограничивая область проверки только вновь записанными файлами;
- периодически проверять на наличие вирусов жесткие диски компьютера, запуская антивирусные программы для тестирования файлов, памяти и системных областей дисков с защищенного от записи носителя информации, предварительно загрузив ОС с защищенного от записи системного носителя информации;
- всегда защищать свои носители информации от записи при работе на других компьютерах, если на них не будет производиться запись информации;
- обязательно делать на электронных носителях архивные копии ценной для пользователя информации;
- использовать антивирусные программы для входного контроля всех исполняемых файлов, получаемых из компьютерных сетей.

У каждого типа антивирусных программ есть свои достоинства и недостатки. Только комплексное использование нескольких типов антивирусных программ может привести к приемлемому результату. Программные средства защиты представляют собой комплекс алгоритмов и программ, нацеленных на контроль и исключение проникновения несанкционированной информации.

Существует спектр программных комплексов, предназначенных для профилактики заражения вирусом, обнаружения и уничтожения вирусов. Они обладают универсальностью, гибкостью, адаптивностью и др.

Перечислим наиболее распространенные антивирусные программные комплексы:

- антивирус Касперского (AVP);
- антивирус Dr.Web;
- антивирус Symantec Antivirus;
- антивирус McAfee;
- антивирус AntiVir Personal Edition.

Как и у любой программы, у компьютерных вирусов можно выделить две основные стадии жизненного цикла — хранение и исполнение.

**Стадия хранения** соответствует периоду, когда вирус просто хранится на диске совместно с объектом, в который он внедрен. На этой стадии вирус является наиболее уязвимым со стороны антивирусного ПО, так как он не активен и не может контролировать работу ОС с целью самозащиты.

Некоторые вирусы на этой стадии используют механизмы защиты своего кода от обнаружения. Наиболее распространенным способом защиты является шифрование большей части тела вируса. Его использование совместно с механизмами мутации кода (об этом идет речь ниже) делает невозможным выделение сигнатур — устойчивых характеристических фрагментов кода вирусов.

Стадия исполнения компьютерных вирусов, как правило, включает пять этапов:

- загрузка вируса в память;
- поиск жертвы;
- заражение найденной жертвы;
- выполнение деструктивных функций;
- передача управления программе-носителю вируса.

Рассмотрим эти этапы подробнее.

**Загрузка вируса.** Загрузка вируса в память осуществляется ОС одновременно с загрузкой исполняемого объекта, в который вирус внедрен. Например, если пользователь запустил на исполнение программный файл, содержащий вирус, то, очевидно, вирусный код будет загружен в память как часть этого файла. В простейшем случае процесс загрузки вируса представляет собой не что иное, как копирование с диска в оперативную память, сопровождаемое иногда настройкой адресов, после чего происходит передача управления коду тела вируса. Эти действия выполняются ОС, а сам вирус находится в пассивном состоянии. В более сложных ситуациях вирус может после получения управления выполнять дополнительные действия, которые необходимы для его функционирования. В связи с этим рассматриваются два аспекта.

**Первый аспект** связан с максимальным усложнением процедуры обнаружения вирусов. Для обеспечения защиты на стадии хранения некоторые вирусы используют достаточно сложные алгоритмы. К таким усложнениям можно отнести шифрование основного тела вируса. Однако использование только шифрования является полумерой, так как в открытом виде должна храниться та часть вируса, которая обеспечивает расшифрование вируса на стадии загрузки. Для избегания подобной ситуации разработчики вирусов используют механизмы «мутаций» кода расшифровщика. Суть этого метода состоит в том, что при внедрении в объект копии вируса часть ее кода, относящаяся к расшифровщику,

модифицируется так, чтобы возникли текстуальные различия с оригиналом, но результаты работы остались неизменными. Обычно применяют следующие приемы модификации кода:

- изменение порядка независимых инструкций;
- замена некоторых инструкций на эквивалентные по результату работы;
- замена используемых в инструкциях регистров на другие;
- введение случайным образом зашумляющих инструкций.

Вирусы, использующие подобные механизмы мутации кода, получили название *поллиморфных вирусов*. При совместном использовании механизмов шифрования и мутации внедряемая копия вируса окажется отличной от оригинала, так как одна ее часть будет изменена, а другая окажется зашифрованной на ключе, сгенерированном специально для этой копии вируса. А это существенно осложняет выявление вируса в вычислительной системе.

*Полиморфные вирусы* (polymorphic) — это трудно обнаруживаемые вирусы, не имеющие сигнатур, т.е. не содержащие ни одного постоянного участка кода. В большинстве случаев два образца одного и того же полиморфного вируса не будут иметь ни одного совпадения. Полиморфизм встречается в вирусах всех типов — файловых, загрузочных и макровирусах.

Дополнительные действия, которые выполняют полиморфные вирусы на этапе загрузки, состоят в расшифровывании основного тела вируса.

При использовании стелс-алгоритмов вирусы могут полностью или частично скрыть себя в системе. Наиболее распространенный стелс-алгоритм осуществляет перехват системных запросов с целью контроля действий ОС. Вирусы, использующие стелс-алгоритмы, называются стелс-вирусами.

*Стелс-вирусы* (Stealth) способны скрывать свое присутствие в системе и избегать обнаружения антивирусными программами. Эти вирусы могут перехватывать запросы ОС на чтение (запись) зараженных файлов, при этом они либо временно лечат эти файлы, либо «подставляют» вместо себя незараженные участки информации, эмулируя «чистоту» зараженных файлов.

В случае макровирусов наиболее популярным способом является запрет вызовов меню просмотра макросов. Одним из первых файловых стелс-вирусов был вирус Frodo, первым загрузочным стелс-вирусом был вирус Brain.

Нередко в вирусах используются различные нестандартные приемы с целью глубже спрятаться в ядре ОС, либо защитить от обнаружения свою резидентную копию, либо затруднить лечение от вируса и т.п.

**Второй аспект** связан с так называемыми резидентными вирусами. Поскольку вирус и объект, в который он внедрен, являются для ОС единым целым, то после загрузки они располагаются, естественно, в едином адресном пространстве. После завершения работы объекта он выгружается из оперативной памяти, при этом одновременно выгружается и вирус, переходя в пассивную стадию хранения. Однако некоторые типы вирусов способны сохраняться в памяти и оставаться активными после окончания работы вирусоносителя. Эти вирусы получили название резидентных.

*Резидентные вирусы* при инфицировании компьютера оставляют в оперативной памяти свою резидентную часть, которая затем перехватывает обращения ОС к объектам заражения и внедряется в них. Резидентные вирусы находятся в памяти и являются активными вплоть до выключения компьютера или перезагрузки ОС.

Резидентными можно считать макровирусы, так как для большинства из них выполняются основные требования — постоянное присутствие в памяти компьютера на все время работы зараженного редактора и перехват функций, используемых при работе с документами. При этом роль ОС берет на себя редактор, а понятие «перезагрузка операционной системы» трактуется как выход из редактора.

Нерезидентные вирусы не заражают память компьютера и сохраняют активность ограниченное время. Некоторые вирусы оставляют в оперативной памяти небольшие резидентные программы, которые не распространяют вирус. Такие вирусы считаются нерезидентными.

Следует отметить, что деление вирусов на резидентные и нерезидентные справедливо в основном для файловых вирусов. Загрузочные вирусы, как и макровирусы, относятся к резидентным вирусам.

**Поиск жертвы.** По способу поиска жертвы вирусы можно разделить на два класса.

К **первому классу** относятся вирусы, осуществляющие «активный» поиск с использованием функций ОС. Примером являются файловые вирусы, использующие механизм поиска исполняемых файлов в текущем каталоге.

**Второй класс** составляют вирусы, реализующие «пассивный» механизм поиска, т.е. вирусы, расставляющие «ловушки» для программных файлов. Как правило, файловые вирусы устраивают такие ловушки путем перехвата функции Exec ОС, а макровирусы — с помощью перехвата команд типа Save as из меню File.

**Заражение жертвы.** В простейшем случае заражение представляет собой самокопирование кода вируса в выбранный в качестве жертвы объект. Классификация вирусов на этом этапе связана с анализом особенностей этого копирования и способов модификации заражаемых объектов.

**Особенности заражения файловыми вирусами.** По способу инфицирования жертвы вирусы можно разделить на два класса.

К **первому классу** относятся вирусы, которые не внедряют свой код непосредственно в программный файл, а изменяют имя файла и создают новый, содержащий тело вируса.

**Второй класс** составляют вирусы, внедряющиеся непосредственно в файлы-жертвы. Они характеризуются местом внедрения. Возможны следующие варианты.

Внедрение в начало файла. Этот способ является наиболее удобным для COM-файлов MS-DOS, так как данный формат не предусматривает наличие служебных заголовков. При внедрении этим способом вирусы могут либо производить конкатенацию собственного кода и кода программы-жертвы, либо переписывать начальный фрагмент файла в конец, освобождая место для себя.

Внедрение в конец файла. Это наиболее распространенный тип внедрения. Передача управления коду вирусов обеспечивается модификацией первых команд программы (COM) или заголовка файла (EXE).

Внедрение в середину файла. Как правило, этот способ используется вирусами применительно к файлам с заранее известной структурой (например, к файлу COMMAND.COM) или же к файлам, содержащим последовательность байтов с одинаковыми значениями, длина которой достаточна для размещения вируса. Во втором случае вирусы архивируют найденную последовательность и замещают собственным кодом. Помимо этого вирусы могут внедряться в середину файла, освобождая себе место путем переноса фрагментов кода программы в конец файла или же «раздвигая» файл.

**Особенности заражения загрузочными вирусами** определяются особенностями объектов, в которые они внедряются, — загрузочными секторами гибких и жестких дисков и главной загрузочной записью (MBR) жестких дисков. Основной проблемой является ограниченный размер этих объектов. В связи с этим вирусам необходимо сохранить на диске ту свою часть, которая не уместилась на месте жертвы, а также перенести оригинальный код инфицированного загрузчика. Существуют различные способы решения этой задачи.

Используются псевдосбойные секторы. Вирус переносит необходимый код в свободные секторы диска и помечает их как сбойные, защищая тем самым себя и загрузчик от перезаписи.

Используются редко применяемые секторы в конце раздела. Вирус переносит необходимый код в эти свободные секторы в конце диска. С точки зрения ОС эти секторы выглядят как свободные.

Используются зарезервированные области разделов. Вирус переносит необходимый код в области диска, зарезервированные под нужды ОС, а потому неиспользуемые.

Короткие вирусы могут уместиться в один сектор загрузчика и полностью взять на себя функции MBR или загрузочного сектора.

**Особенности заражения макровирусами.** Процесс заражения сводится к сохранению вирусного макрокода в выбранном документе-жертве. Для некоторых систем обработки информации это сделать не просто, так как формат файлов документов может не предусматривать возможность сохранения макропрограмм. В качестве примера приведем Microsoft Word 6.0. Сохранение макрокода для этой системы возможно только в файлах шаблонов (имеющих по умолчанию расширение .DOT). Поэтому для своего сохранения вирус должен контролировать обработку команды Save as из меню File, которая вызывается всякий раз, когда происходит первое сохранение документа на диск. Этот контроль необходим, чтобы в момент сохранения изменить тип файла-документа (имеющего по умолчанию расширение .DOC) на тип файла-шаблона. В этом случае на диске окажутся и макрокод вируса, и содержимое документа.

Помимо простого копирования кода вируса в заражаемый объект на этом этапе могут использоваться более сложные алгоритмы, обеспечивающие защиту вируса на стадии хранения. К числу таких вирусов относятся описанные выше полиморфные вирусы.

**Выполнение деструктивных функций.** Вирусы могут выполнять помимо самокопирования деструктивные функции.

По деструктивным возможностям вирусы можно разделить на безвредные, неопасные, опасные и очень опасные.

Безвредные вирусы — это вирусы, в которых реализован только механизм самораспространения. Они не наносят вред системе, за исключением расхода свободной памяти на диске в результате своего распространения.

Неопасные вирусы — это вирусы, присутствие которых в системе связано с различными эффектами (звуковыми, видео) и уменьшением свободной памяти на диске, но которые не наносят вред программам и данным.

Опасные вирусы — это вирусы, которые могут привести к серьезным сбоям в работе компьютера. Последствием сбоя может стать разрушение программ и данных.

Очень опасные вирусы — это вирусы, в алгоритм работы которых заведомо заложены процедуры, непосредственно приводящие к разру-

шениям программ и данных, а также к стиранию информации, записанной в системных областях памяти и необходимой для работы компьютера.

На «степень опасности» вирусов оказывает существенное влияние та среда, под управлением которой вирусы работают.

Так, вирусы, созданные для работы в MS-DOS, обладают практически неограниченными потенциальными возможностями.

Распространение вирусов под управлением Windows NT/2000 ограничивается развитой системой разграничения доступа.

Возможности макровирусов напрямую определяются возможностями макроязыков, на которых они написаны. В частности, язык Word Basic позволяет создать мощные макровирусы, способные доставить пользователям серьезные неприятности.

Дополняя эту классификацию, можно отметить также деление вирусов на вирусы, наносящие вред системе вообще, и вирусы, предназначенные для целенаправленных атак на определенные объекты.

**Передача управления программе-носителю вируса.** Здесь следует указать на деление вирусов на разрушающие и неразрушающие.

Разрушающие вирусы не заботятся о сохранении работоспособности инфицированных программ, поэтому для них этот этап функционирования отсутствует.

Для неразрушающих вирусов этот этап связан с восстановлением в памяти программы в том виде, в котором она должна корректно исполняться, и передачей управления программе-носителю вируса.

Кроме вирусов принято выделять еще несколько видов вредоносных программ. Это троянские программы, логические бомбы, хакерские утилиты скрытого администрирования удаленных компьютеров, программы, ворующие пароли доступа к ресурсам Интернет и прочую конфиденциальную информацию. Четкого разделения между ними не существует: троянские программы могут содержать вирусы, в вирусы могут быть встроены логические бомбы и т.д.

Троянские программы не размножаются и не рассылаются сами. Внешне они выглядят совершенно безобидно и даже предлагают полезные функции. Но когда пользователь загрузит такую программу в свой компьютер и запустит ее, она может незаметно выполнять вредоносные функции. Чаще всего троянские программы используются для первоначального распространения вирусов, для получения удаленного доступа к компьютеру через Интернет, кражи данных или их уничтожения.

Логической бомбой называется программа или ее отдельные модули, которые при определенных условиях выполняют вредоносные действия.

Логическая бомба может, например, сработать по достижении определенной даты или тогда, когда в БД появится или исчезнет запись, и т.п. Такая бомба может быть встроена в вирусы, троянские программы и даже в обычные программы.

## 6.4. Защита информации в компьютерных сетях

Система информационной безопасности должна оградить информационные ресурсы сети от наиболее распространенных внешних и внутренних атак, направленных на вывод из строя серверов и уничтожение данных, от нежелательного проникновения в локальные вычислительные сети через «дыры» в ОС, от целенаправленного вторжения в систему для получения конфиденциальной информации.

Для успешного использования современных ИТ необходимо надежное и эффективное управление не только самими сетями, но и средствами сетевой безопасности. И если раньше задача заключалась в управлении отдельными серверами, сетями и маршрутизаторами, то сейчас требуется обеспечить информационную безопасность корпоративных бизнес-процессов. Все это предъявляет жесткие требования к управлению средствами сетевой безопасности.

Важнейшим компонентом системы управления корпоративной сетью является система информационной безопасности. Эта система должна:

- централизованно и оперативно осуществлять управляющие воздействия на средства сетевой безопасности;
- проводить регулярный аудит и мониторинг, дающие объективную информацию о состоянии информационной безопасности для принятия оперативных решений.

Сформулируем основные задачи управления системой сетевой безопасности масштаба предприятия. Функционально система управления средствами защиты информации в распределенной сети масштаба предприятия должна решать следующие задачи:

- управление глобальной политикой безопасности (ГПБ) в рамках сети предприятия, формирование локальных политик безопасности (ЛПБ) отдельных устройств и доведения ЛПБ до всех устройств защиты информации;
- управление конфигурацией объектов и субъектов доступа; включает управление составом, версиями, компонентами устройств и ПО защиты, а также управление патчами (patch), которые служат для закрытия дыр, обнаруженных в поставленных продуктах обеспечения безопасности;

- предоставление сервисов защиты распределенным прикладным системам, а также регистрацию защищенных приложений и их ресурсов. Приложения этой группы должны обеспечивать прежде всего интерфейс (API) для обеспечения управления сервисами защиты со стороны прикладных систем;
- управление криптосредствами, в частности, ключевое управление (ключевая инфраструктура). Ключевая инфраструктура должна функционировать в составе инфраструктурных (системообразующих) служб;
- событийное протоколирование; включает настройку выдачи логов на разные устройства, управление уровнем детализации логов, управление составом событий, по которым ведется протоколирование;
- аудит безопасности ИС; обеспечивает получение и оценку объективных данных о текущем состоянии защищенности ИС, иногда под аудитом безопасности понимают анализ логов, поиск нарушителей и дыр в существующей системе, однако эти функции покрываются, скорее, задачами управления логами;
- мониторинг безопасности системы; обеспечивает получение информации в реальном времени о состоянии, активности устройств и о событиях с контекстом безопасности, происходящих в устройствах, например, о потенциальных атаках;
- обеспечение работы специальных защищенных приложений, например, нотариального надзора за операциями, поддержка регламентных мероприятий (смена ключей, паролей, устройств защиты, выпуск смарт-карт и др.);
- обеспечение работы проектно-инвентаризационной группы приложений; эта группа приложений должна осуществлять:
  - определение точек установки средств защиты в сети предприятия,
  - учет применяемых средств защиты,
  - контроль модульного состава средств защиты,
  - контроль состояния средств защиты и др.

Существует проблема комплексования и организации взаимодействия традиционных систем управления сетями и систем управления средствами защиты информации в сети. Для решения этой проблемы применяются два основных подхода.

Первый подход заключается в интеграции средств сетевого или системного управления с механизмами управления средств защиты. Средства сетевого и системного управления ориентированы, в первую очередь, на управление сетью или ИС, т.е. поддерживают традиционные

действия и услуги: управление учетными записями пользователей, управление ресурсами и событиями, маршрутизацию, производительность и т.п. Ряд компаний — Cisco Systems, Computer Associates, Hewlett Packard, Tivoli Systems — пошли по пути интеграции механизмов управления средств защиты в традиционные системы управления сетями. Однако такие комплексные системы управления часто отличаются высокой стоимостью и, кроме того, некоторые аспекты управления безопасностью остаются за пределами внимания этих систем.

Второй подход заключается в использовании средств, предназначенных для решения только задачи управления безопасностью. Например, Open Security Manager (OSM) от Check Point Software Technologies дает возможность централизованно управлять корпоративной политикой безопасности и устанавливать ее на сетевые устройства по всей компании. Продукт OSM является одним из основных компонентов технологии OPSEC (Open Platform for Secure Enterprise Connectivity), разработанной компанией Checkpoint, он создает интерфейс для управления устройствами сетевой безопасности различных производителей (например, Cisco, Bay, 3Com).

Для обеспечения безопасности информационных ресурсов предприятия средства защиты информации обычно размещаются непосредственно в корпоративной сети. МЭ контролируют доступ к корпоративным ресурсам, отражая атаки злоумышленников извне, а шлюзы виртуальных частных сетей (VPN) обеспечивают конфиденциальную передачу информации через открытые глобальные сети, в частности, Интернет. Для создания надежной шелонированной защиты в настоящее время применяются также такие средства безопасности, как системы обнаружения вторжений IDS (Intrusion Detection Systems), средства контроля доступа по содержанию информации, антивирусные системы и др.

Большинство КИС построено на основе программных и аппаратных средств, поставляемых различными производителями.

Каждое из этих средств требует тщательного и специфического конфигурирования, отражающего взаимосвязи между пользователями и доступными им ресурсами. Чтобы обеспечить в гетерогенной КИС надежную защиту информации, нужна рационально организованная система управления безопасностью КИС, которая обеспечила бы безопасность и правильную настройку каждого компонента КИС, постоянно отслеживала происходящие изменения, устанавливала «заплатки» на найденные в системе бреши, контролировала работу пользователей. Очевидно, что чем разнообразнее ИС, тем сложнее обеспечить управление ее безопасностью.

В основе централизованного управления безопасностью КИС лежит концепция глобального управления безопасностью GSM (Global Security

Management). Концепция GSM позволяет построить комплексную систему управления и защиты информационных ресурсов предприятия со следующими свойствами:

- управление всеми существующими средствами защиты на базе политики безопасности предприятия, обеспечивающее целостность, непротиворечивость и полноту набора правил защиты для всех ресурсов предприятия (объектов политики безопасности) и согласованное исполнение политики безопасности средствами защиты, поставляемыми разными производителями;
- определение всех информационных ресурсов предприятия через единый (распределенный) каталог среды предприятия, который может актуализироваться как за счет собственных средств описания ресурсов, так и посредством связи с другими каталогами предприятия (в том числе по протоколу LDAP);
- централизованное, основанное на политике безопасности (policy-based) управление локальными средствами защиты информации;
- строгая аутентификация объектов политики в среде предприятия с использованием PKCS#11 токенов и инфраструктуры открытых ключей PKI, включая возможность применения дополнительных локальных средств аутентификации LAS (по выбору потребителя);
- расширенные возможности администрирования доступа к определенным в каталоге ресурсам предприятия или частям всего каталога (с поддержкой понятий групп пользователей, доменов, департаментов предприятия), управление ролями как набором прав доступа к ресурсам предприятия, введение в политику безопасности элементов косвенного определения прав через атрибуты прав доступа (credentials);
- обеспечение подотчетности (регистрации всех операций взаимодействий распределенных объектов системы в масштабах корпоративной сети) и аудита, мониторинга безопасности, тревожной сигнализации;
- интеграция с системами общего управления, инфраструктурными системами безопасности (PKI, LAS, IDS).

В рамках данной концепции управление, основанное на политике безопасности — PBM (Policy based management) — определяется как реализация набора правил управления, сформулированных для бизнес-объектов предприятия, которая гарантирует полноту охвата бизнес-области объектами и непротиворечивость используемых правил управления.

Система управления GSM, ориентированная на управление безопасностью предприятия на принципах PBM, удовлетворяет следующим требованиям:

- политика безопасности предприятия представляет собой логически и семантически связанную, формируемую, редактируемую и анализируемую как единое целое структуру данных;
- политика безопасности предприятия определяется в едином контексте для всех уровней защиты как единое целое сетевой политики безопасности и политики безопасности информационных ресурсов предприятия;
- для облегчения администрирования ресурсов и политики безопасности предприятия число параметров политики минимизируется.

Для того чтобы минимизировать число параметров политики, используются следующие приемы:

- 1) групповые определения объектов безопасности;
- 2) косвенные определения, например, определения на основе верительных (credential) атрибутов;
- 3) мандатное управление доступом (в дополнение к фиксированному доступу), когда решение о доступе определяется на основе сопоставления уровня доступа, которым обладает субъект, и уровня конфиденциальности (критичности) ресурса, к которому осуществляется доступ.

Система управления GSM обеспечивает разнообразные механизмы анализа политики безопасности за счет средств многокритериальной проверки соответствия политики безопасности формальным моделям концепции безопасности предприятия.

Ниже приводится концепция определения ГПБ (GSP — Global Security Policy) сети предприятия и описание построенной на базе ГПБ системы управления безопасностью (policy based security management).

Глобальная политика безопасности корпоративной сети представляет собой конечное множество правил безопасности (security rules) (рис. 6.5), которые описывают параметры взаимодействия объектов корпоративной сети в контексте информационной безопасности:

- необходимый для соединения сервис безопасности (правила обработки, защиты и фильтрации трафика);
- направление предоставления сервиса безопасности;
- правила аутентификации объектов;
- правила обмена ключами;
- правила записи результатов событий безопасности в системный журнал;
- правила сигнализации о тревожных событиях и др.

При этом объектами ГПБ могут быть как отдельные рабочие станции и подсети, так и группы объектов, которые могут включать в себя целые структурные подразделения компании (например, отдел маркетинга или



Рис. 6.5. Структура правила глобальной политики безопасности

финансовый департамент) или даже отдельные компании (входящие, например, в холдинг). Политика безопасности для каждого объекта в группе автоматически реплицируется всем объектам группы.

Задачи защиты бизнес-объектов распределенной корпоративной системы можно сформулировать в терминах правил, поскольку сетевое взаимодействие можно представить как простую передачу информации между субъектом Subj и объектом Obj доступа на основе некоторого сетевого сервиса защиты SecSrv, настроенного при помощи параметров P. В результате глобальная политика безопасности предприятия представляется как набор правил вида Subj, Obj, SecSrv (P).

При этом отсутствие правила для объекта Obj означает запрет любого доступа к данному Obj.

Для простоты определения целей безопасности предприятия в GSM предусмотрено два типа объектов, выступающих в качестве Subj и Obj. Это пользователь (U) и ресурс (R).

Ресурс R может быть информационным (IR) или сетевым (NR).

Пользователь и ресурс могут выступать в любой из форм агрегации, поддерживаемых в системе: группы, домены, роли, департаменты, разделы каталога.

Пример: правило (U, IR, S1) представляет собой правило защиты S1, обеспечиваемое при доступе пользователя U к информационному ресурсу IR. Правило (IR1, IR2, S2) означает разрешение сетевого взаимодействия двух информационных модулей (программ) с необходимостью обеспечения свойств защиты S2.

Политика по умолчанию для доступа к любому защищаемому объекту корпоративной системы представляет собой запретительное правило: все, что не разрешено явно, запрещено. Такое правило обеспечивает полноту защиты информации в сети предприятия и априорное отсутствие «дыр» в безопасности.

Чтобы обеспечить взаимодействие устройств в сети, для них создается и доставляется (в общем случае не по каналам сети) стартовая конфигурация, содержащая необходимые правила настройки устройств только для их централизованного управления — стартовая политика безопасности устройства.

Правила ГПБ могут быть распространены как на сетевые взаимодействия, так и на функции контроля и управления самой системы.

Функционально правила ГПБ разбиты по группам:

- правила VPN. Правила данного типа реализуются при помощи протоколов IPSec; агентом исполнения правила является драйвер VPN в стеке клиентского устройства или шлюза безопасности (IPX, IP2, VPNRule);
- правила пакетной фильтрации. Они обеспечивают пакетную фильтрацию типа stateful и stateless; исполнение этих правил обеспечивают те же агенты, что исполняют VPN-правила (IP1, IP2, PacketRule);
- проxy-правила, включая антивирусную защиту «на лету». Эти правила отвечают за фильтрацию трафика, передаваемого под управлением заданных прикладных протоколов; их исполнительным агентом является проxy-агент, например, (User, Protocol, ProxyRule) или (Application, Protocol, Proxy-Rule);
- правила аутентифицированного (авторизованного) доступа, включая правила Single Sign-On. Управление доступом Single Sign-On обеспечивает данному пользователю работу на едином пароле или другой аутентификационной информации со многими информационными ресурсами; понятно, что символическая запись правила сетевого доступа легко распространяется на Single Sign-On (User, Application, Authentication Scheme). Правила этой группы могут комбинированно исполняться агентами различного уровня, от VPN-драйвера до проxy-агентов; кроме того, агентами исполнения таких правил могут быть системы аутентификации запрос-отклик и продукты третьих разработчиков;
- правила, отвечающие за сигнализацию и *событийное протоколирование*. Политика протоколирования может оперативно и централизованно управляться агентом протоколирования; исполнителями правил являются все компоненты системы.

Набор правил ГПБ является логически целостным и семантически полным описанием политики безопасности в масштабах сети, на основе которой может строиться локальная политика безопасности отдельных устройств.

**Локальная политика безопасности.** Любому средству защиты, реализующему какой-либо сервис информационной безопасности, необходима для выполнения его работы ЛПБ — точное описание настроек для корректной реализации правил аутентификации пользователей, управления доступом, защиты трафика и др. При традиционном подходе администратору приходится отдельно настраивать каждое средство защиты или реплицировать какие-то простейшие настройки на большое число узлов с последующей их корректировкой. Очевидно, что это неизбежно приводит к большому числу ошибок администрирования и, как следствие, существенному снижению уровня защищенности корпоративной сети.

После формирования администратором ГПБ Центр управления на основе интерпретации ГПБ автоматически вычисляет и, если это необходимо, корректирует отдельные ЛПБ для каждого средства защиты и автоматически загружает нужные настройки в управляющие модули соответствующих средств защиты.

В целом ЛПБ сетевого устройства включает в себя полный набор правил разрешенных соединений данного устройства, исполняемых для обеспечения какой-либо информационной услуги с требуемыми свойствами защиты информации.

Различие между правилами, реализующими ГПБ в сети, и правилами, реализующими ЛПБ конкретного устройства, заключается в том, что в правилах группы ГПБ объекты и субъекты доступа могут быть распределены произвольным образом в пределах сети, а правила группы ЛПБ, включая субъекты и объекты ЛПБ, предназначены и доступны только в пределах пространства одного из сетевых устройств.

## Контрольные вопросы и задания

1. Дать определение ИБ. Назвать основные направления и задачи обеспечения ИБ общества.
2. Назвать основные компоненты ИБ автоматизированных ИС.
3. Определить основные виды угроз информационным ресурсам.
4. Описать причины возникновения каналов несанкционированного доступа к информации. Классифицировать виды каналов несанкционированного доступа к информации.

5. Охарактеризовать технические каналы несанкционированного доступа к информации. Охарактеризовать легальные и нелегальные методы обеспечения действия каналов утечки информации.
6. Показать порядок защиты прав граждан на личную тайну и неприкосновенность частной жизни законодательством Российской Федерации о СМИ.
7. Определить объекты защиты авторских прав и интеллектуальной собственности, защищаемые патентным законодательством.
8. Дать определение государственной тайны и назвать грифы секретности и перечислить сведения, составляющие государственную тайну, и сведения, которые не могут относиться к государственной тайне.
9. Дать определение коммерческой тайны и перечислить сведения, которые не могут быть ее объектом.

## СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Информационные технологии (для экономиста) : учеб. пособие для студ. экон. спец. / А.К. Волков [и др.]. М. : ИНФРА-М, 2014. 305 с.
2. Уткин В.Б. Информационные системы и технологии в экономике : учебник. М. : ЮНИТИ-ДАНА, 2013. 355 с.
3. Буслова М.К. Новые технологии как фактор социальной динамики // Право и экономика. 2015. 48 с.
4. Демидов Л.Н., Звездарев Д.В. Информационный менеджмент : учебник. М. : Бестселлер. 2015.
5. Современные информационные технологии : учебник / Л.Н. Демидов, В.Б. Терновсков, В.В. Терновский, О.А. Смирнов. Межрегиональный институт развития территорий. 2016. 726 с.
6. Князев В.В., Демидов Л.Н. Информатика. Часть 1. Теоретические основы информатики. М. : МО РФ, ВА им. Петра Великого, 2013.
7. Введение в информационные технологии : учебник / Л.Н. Демидов, С.М. Григорьев, В.Б. Терновсков, Д.В. Крахмалев. Межрегиональный институт развития территорий, 2015. 266 с.
8. Системы передачи информации : учеб. пособие. Часть 1 / Д.С. Крапов, Л.Н. Демидов [и др.]. М. : ФГОУ ВПО РГУТиС, 2013.
9. Системы передачи информации : учеб. пособие. Часть 2 / Д.С. Крапов, Л.Н. Демидов [и др.]. М. : ФГОУ ВПО РГУТиС, 2013.
10. Демидов Л.Н., Терновский В.В., Терновсков В.Б., Тарасов Б.А. Модель представления информации для применения в экономике. «Экономика: вчера, сегодня, завтра» №3\2016
11. Демидов Л.Н., Калищук С.А. Введение в информационный менеджмент : учебник. 2013.
12. Демидов Л.Н., Звездарев Д.В. Введение в информационные технологии в менеджменте (управлении). М. : Бестселлер, 2014 .
13. Демидов Л.Н., Звездарев Д.В., Гладышев А.И. Информационные технологии. М. : Бестселлер, 2014.
14. Борисов Р.С., Карпов Д.С., Демидов Л.Н. Администрирование в информационных системах : практикум. М. : ФГОУ ВПО РГУТиС, 2013.
15. Менеджмент в малом и среднем бизнесе : учеб. пособие / Т.А. Платонова, Л.Н. Демидов [и др.]. М. : Институт Управления, бизнеса и технологий, 2014.