МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ УКРАИНЫ

НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
”ХАРЬКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ ИНСТИТУТ”

ФАКУЛЬТЕТ ИНФОРМАТИКИ И УПРАВЛЕНИЯ

КАФЕДРА ЭКОНОМИЧЕСКОЙ КИБЕРНЕТИКИ И МАРКЕТИНГОВОГО МЕНЕДЖМЕНТА

РЕФЕРАТ

По курсу ” Компьютерные информационные системы”

Тема «Сети Internet и Intranet»

ХАРЬКОВ 2007

**1. Общие сведения об Internet/Intranet**

Идея создания универсальной базы данных прорабатывалась в течение длительного периода, и только недавно были получены средства создания подобных баз данных. Многие рассматривают Internet и Всемирную паутину (WorldWideWeb, WWW) как экспериментальный образец такой базы данных. Технология, разработанная для WWW, воплощает идею глобальной информационной базы данных, реализованную в пределах современных возможностей.

**1.1. Основы технологии Internet/Intranet**

Мир Internet весьма разнообразен, мозаичен по своей структуре, включает много составляющих, каждая из которых имеет специфику применения. Разработчику приложений прежде всего необходимо разобраться в сопутствующей терминологии, уяснить принципы организации и основные элементы WWW, понять особенности корпоративных Intranet-сетей, а также преимущества и недостатки технологии Internet/Intranet.

**1.1.1. Сети Internet и WWW**

Internet в настоящее время является самым большим и популярным межсетевым объединением в мире. Оно объединяет десятки тысяч компьютерных сетей и миллионы пользователей во всем мире. При этом объединены компьютеры тысяч различных типов, оснащенных самым разным программным обеспечением. Пользователи Internet могут не обращать внимания на все эти различия.

В настоящее время существуют два созвучных термина - internet и Internet. Термин *internet* относится к технологии обмена данными, основанной на использовании семейства протоколов TCP/IP, а *Internet* - это глобальное сообщество мировых сетей, которые используют internet для обмена данными. Как правило, термин *"TCP/IP"* это то же самое, что и *"набор протоколов TCP/IP",* или *"набор протоколов internet",* или *"технология internet".*

Сеть Internet с самого начала не была централизованной и состояла из отдельных сегментов. Каждый узел сети независим от остальных узлов и может самостоятельно отвечать за прием и передачу сообщений. В основу информационного обмена положен принцип *коммутации пакетов*. Любое информационное сообщение делится на части, которые называются *пакетами*. Пакеты передаются по сети и собираются в сообщение в узле-получателе.

**1.1.2. Принципы организации WWW**

В марте 1989 года Тим Бернерс-Ли предложил руководству исследовательского центра CERN концепцию новой распределенной информационной системы, которую он назвал WorldWideWeb. Тим Бернерс-Ли считал, что информационная система, построенная на принципах гипертекста, должна объединить все множество информационных ресурсов CERN, которое состояло из базы данных отчетов, компьютерной документации, списков почтовых адресов, наборов данных экспериментов, информационной реферативной системы и т.п. Основная метафора гипертекста - это "электронная книга" с автоматически поддерживаемыми переходами по ссылкам. Гипертекстовая технология должна была позволить легко переходить с одного документа на другой с помощью гипертекстовых ссылок.

Проект был успешно реализован, в частности, к 1991 году был создан первый броузер (программа просмотра гипертекста), получивший название "www" и работавший в режиме командной строки. С этого момента основными элементами технологии WWW являются:

* язык гипертекстовой разметки документов HTML;
* универсальный способ адресации ресурсов в сети (URI и URL);
* протокол обмена гипертекстовой информацией HTTP;
* универсальный интерфейс шлюзов CGI, добавленный позже сотрудниками Национального Центра Суперкомпьютерных приложений (NSCA).

Язык гипертекстовой разметки HTML был создан на опыте использования редактора TeX и стандартного языка разметки SGML. Основная идея гипертекста заключается в присутствии внутри ASCII-текста форматирующих полей и ссылок как на части внутри документа, так и на другие документы. Поля и ссылки также являются фрагментами ASCII-текста, но, подобно программе, следуют строгим синтаксическим правилам. Благодаря этому, пользователь имеет возможность просматривать документы в том порядке, какой ему больше нравится, а не последовательно, как при чтении книг. Help-файлы, с которыми сталкивался любой пользователь ЭВМ, дают хорошее представление о гипертекстовой организации информации, позволяющей пользователю переходить от темы к теме, используя выделенные слова или поля текста.

Чтобы получить файл из Internet, броузер (browser, программа для просмотра Web, клиент) должен знать, где находится файл и как общаться с компьютером, на котором этот файл находится. Поэтому требуется, чтобы программа-клиент WWW передала имя определенного файла, его местоположение в Internet (адрес хоста) и метод доступа (обычно протокол типа HTTP или FTP). Комбинация этих элементов формирует универсальный идентификатор ресурса (UniversalResourceIdentifier, URI). URI определяет способ записи адресов различных информационных ресурсов. В основу URI были заложены идеи расширяемости, полноты и читаемости. Реализация URI для WWW называется URL (UniversalResourceLocator).

Общий формат ссылки URL:

протокол://узел/путь/файл[#метка]

протокол (или метод доступа) определяет способ взаимодействия с информационным ресурсом;
узел - имя или IP-адрес узла (сервера определенного типа), где расположена информация;
путь - имя каталога (возможно виртуального) или цепочки вложенных каталогов Web-сервера или файловой системы;
файл - простое имя файла с расширением, содержащее гипертекст, графический образ, прикладную программу или другую информацию;
метка- имя закладки в гипертекстовом файле, позволяет осуществлять внутренние переходы к разным фрагментам одного документа.

Вот некоторые примеры URL:

http://www.citmgu.ru/glossary.htm#P

http://citnt/text/docs/intro.htm

http://190.248.27.124/scripts/proc1.exe

В первом случае вызывается фрагмент HTML-файла с указанием доменного имени Web-сервера. Во втором примере используется простое имя узла. Третий пример содержит вызов процедуры при помощи IP-адреса Web-сервера.

Ссылки URL могут быть относительными. Например, из документа http://citnt/text/docs/intro.htm и для сервера citnt допустимы следующие обращения:

books/book1.doc

/images/pic24.gif

http://citnt/text/

http://citnt/

Первая ссылка выполняется относительно текущего каталога, вторая - по отношению к корню Web-сервера. Последние две ссылки правильнее называть неполными. При этом в указанных каталогах отыскивается стандартный индексный файл index.html (index.htm) или default.htm. В некоторых броузерах при отсутствии такого файла на экран выдается оглавление каталога.

При записи всегда используется нотация Unix (слэши, а не обратные слэши), для локальных дисков - идентификатору диска предшествует слэш, а вместо двоеточия ставится вертикальная черта, например: file:///C intranet/lab1/test1.htm

Глобальная сеть Internet, объединяющая гигантские информационные ресурсы, оставалась бы колоссальным банком данных, малоизвестных и потому малодоступных. Конечно, начальная страница по интересующей тематике и дальнейшие гипертекстовые ссылки рано или поздно выведут на необходимую информацию, но как найти нужные страницы среди сотен миллионов, даже если известно, что интересующая информация организована в гипертекстовые ссылки. Мощные поисковые службы Yahoo, Lycos и AltaVista обеспечивают быстрый поиск и доступ практически ко всей информации, находящейся в Internet.

Поисковый робот AltaVista пробегает Всемирную паутину со скоростью 2.5 млн. страниц в день и уже нашел и проиндексировал больше информации, чем любой другой программный робот или поисковая служба. Конечной целью является обнаружение всех страниц и индексация каждого слова в текстовом массиве Web. Когда эта работа будет завершена, пользователи впервые получат возможность точно установить, сколько ссылок на их титульные страницы содержится на других страницах всемирной паутины.

**1.1.3. Корпоративные Intranet-сети**

Применение internet- и WWW-технологий в корпоративной сети, изолированной от Internet, называется Intranet-технологией. Intranet-сети, получающие все более широкое распространение, состоят из внутрикорпоративных Web-серверов, доступ персонала к которым организован через ЛВС или глобальные сети.

Как правило, информация внутри отдельной организации распределена по множеству компьютеров и хранится в виде разнообразных файлов, отчетов и сообщений электронной почты. Как правило, способ доступа ко всей информации не организован. Поэтому многие организации создают intranet-сети с внутренними Web-серверами, чтобы облегчить своим сотрудникам доступ к разнообразной информации. Благодаря связям с корпоративными базами данных, файл-серверами и хранилищами документов Web-серверы предоставляют сотрудникам компании самые различные виды информации через единый интерфейс - хорошо знакомый Web-броузер. Несколько начальных страниц служат гипертекстовыми связями со всеми видами документов и данных.

При работе в корпоративной сети прежде всего необходимо знать имя файла, содержащего необходимую информацию. Для опытных в обращении с компьютерами сотрудников поиск и совместное использование файлов в сети становится чем-то вроде второй профессии, но даже для них задача оказывается трудновыполнимой, если кто-то переименует файл, изменит его содержимое или переместит в другое место. Для неискушенных же пользователей попытки найти такую "потерянную" информацию вызывают раздражение и чаще всего оканчиваются неудачей.

Что бы их устроило на самом деле - это страничка текста на экране, описывающая, что делается в системе, и подсказывающая в процессе перемещения из одного пункта сети в другой местонахождение и содержимое файлов. Именно такой текстовый путеводитель, показывающий, как организованы данные, и предоставляет intranet-приложение. Если на экране появилась ссылка на интересующий информационный объект, чтобы добраться до него, достаточно щелкнуть кнопкой мыши.

*"Меня убедил опыт собственной компании,* - говорит Билл Гейтс. - *В Microsoft мы опубликовали в своей сети intranet тысячи документов для использования внутри корпорации. Я с удивлением обнаружил, что с вводом у нас сети intranet число обращений к данным возросло в несколько раз (прим.: по данным другой статьи - в 3 раза). Это весьма примечательный факт, если учесть, что и раньше ко многим из таких файлов добраться было нетрудно, а пользующиеся ими сотрудники неплохо владеют компьютером и весьма заинтересованы в изучении хранящейся в них информации"*. Это свидетельство особенно показательно, так как именно Microsoft выпускает и рекламирует пакет коллективной работы Exchange как средство публикации общедоступной информации через систему папок.

С точки зрения менеджера, по мнению Билла Гейтса, наиболее привлекательным моментом концепции intranet является то, что для ее организации не требуется затрачивать больших усилий или вкладывать значительные денежные средства. Предприятия, обладающие связанными в IP-сеть компьютерами, практически подготовлены к ее внедрению. Единственное, на что следует потратиться, это нанять специалиста, который составит описательные гипертекстовые страницы и разработает систему ссылок.

Чтобы облегчить поиск информации по ключевым словам, не вынесенным в гипертекстовые ссылки, фирма DigitalEquipment выпустила версию программы AltaVista для персональных компьютеров, что позволяет применять этот мощный и эффективный поисковый механизм в Intranet.

**1.1.4. Преимущества и недостатки Internet/Intranet-технологии**

Достоинства и недостатки Intranet-технологии проистекают из особенностей, заложенных при создании WWW-технологий:

* использование гипертекста обеспечивает связность разнородной информации;
* броузер (клиентская программа WWW-сервера) предоставляет единый, более простой интерфейс пользователя;
* унаследованные преимущества и недостатки архитектуры клиент-сервер приводят к сходным проблемам, возникающим при использовании баз данных.

**Преимущества**

Корпоративная сеть Intranet - прекрасная платформа для публикации информации внутри предприятия. Web-броузер доступен практически для любой клиентской системы - в отличие от фирменных клиентов ПО групповой работы, которых для части платформ может не быть. Web-серверы не требуют аппаратных средств такой большой мощности (быстродействия процессора и емкости жесткого диска), как, например, классические пакеты для групповой работы типа LotusNotes. Рынок ПО для Web-серверов характеризуется высоким уровнем конкуренции - пользователи не привязаны к одному поставщику. Тем не менее программы прекрасно взаимодействуют. Технология Web обладает свойством наращиваемости и может использоваться в региональных вычислительных сетях. Новые средства авторской разработки в комплексе прикладных программ для настольных ПК облегчают новичкам создание HTML-страниц для Web-серверов.

**Недостатки**

Прикладные системы совместной работы для intranet-сетей не обладают мощными возможностями традиционного ПО для групповой работы. Существует ограниченный набор средств привязки Web-серверов к базам данных и другим вспомогательным прикладным системам. Необходимо организовывать и поддерживать функционирование отдельных прикладных систем, в частности, таких, как электронная почта и Web-серверы, вместо того, чтобы пользоваться одной унифицированной системой, как в пакетах для групповой работы. Реализация Intranet требует наличия сети TCP/IP в отличие от других пакетов для групп, которые будут работать с существующими протоколами передачи данных по ЛВС. В intranet-сетях нет встроенных средств тиражирования для удаленных пользователей, в то время как в пакетах типа LotusNotes они есть. Язык HTML и протокол HTTP не обладают достаточной мощностью для разработки прикладных систем клиент-сервер. Стандарты программирования для Web, например, такие, как Java, являются относительно новыми и недостаточно освоенными.

**1.2. Архитектура Internet/Intranet-приложений**

Для анализа архитектуры Internet/Intranet-приложений необходимо рассмотреть основные информационные ресурсы и потоки, классификацию средств для разработки, эксплуатации и сопровождения подобных приложений, а также основные схемы взаимодействия компонентов системы.

**1.2.1. Основные информационные ресурсы и потоки**

Системы на основе гипертекста позволяют создать информационную инфраструктуру корпорации, объединив различные информационные ресурсы и предоставив к ним единообразный доступ.

В информационную инфраструктуру корпорации могут входить следующие информационные ресурсы:

* гипертексты и гипермедиа;
* офисные документы;
* графическая информация;
* архивные файлы;
* сообщения электронной почты;
* новости;
* базы данных;
* хранилище данных;
* прикладное программное обеспечение.

Направление и содержание информационных потоков зависит от природы ресурсов, средств их создания, механизмов и прав доступа к информации.

Гипертекст и гипермедиа выполняют роль своеобразного клея, который соединяет различные информационные ресурсы. Гипертекстовая информация не только отображается и служит для навигации, но и поддерживает диалог с пользователем, при необходимости - ввод данных по электронной форме с передачей их на сервер. Пользователь может даже пересылать на сервер произвольные файлы.

Гипертекстовые страницы могут быть сгенерированы динамически с использованием данных из других ресурсов. Гипермедиа является динамической по своей природе, но поток информации, как правило, однонаправленный, связан с воспроизведением аудио, видео и других мультимедиа-файлов. Обратный поток ограничен управлением навигацией и воспроизведением.

Офисные документы представляют собой тексты, электронные документы, планы и т.п., подготовленные пакетами офисной автоматизации или групповой работы. Доступ к этой информации может быть обеспечен либо в режиме чтения (с использованием программ просмотра), либо полный доступ с возможностью редактирования в той среде, где документ был создан. Помимо навигации с помощью гипертекста возможно организовать контекстный поиск документов.

Графическая информация отображается на клиенте в виде статических иллюстраций, анимации или трехмерных сцен виртуальной реальности. Для выполнения навигации с изображением или его частями связывают гиперссылки, кроме того, имеется возможность изменять виртуальные сцены.

Архивные файлы служат основой электронных архивов документов и программ, доступных на серверах FTP и Gopher для распространения. Пользователь может выбрать нужную информацию и получить ее с сервера по запросу. Загрузка файлов возможна и с Web-серверов.

Сообщения электронной почты представляют собой информационный ресурс, хранимый в почтовых ящиках и общих папках. Они образуют двунаправленные потоки, сообщения можно посылать и получать, не выходя из броузера. Однако больше распространен сбор данных с использованием почты. В общих папках обычно организуются дискуссии, часто задаваемые вопросы FAQ, планы встреч и другие групповые работы.

Новости - динамически изменяемый ресурс, организуемый на основе внешних или внутренних (корпоративных) каналов, последние пока не получили широкого распространения. Отображение новостей производится по запросу или в режиме бегущей строки с помощью тикера. Распространение информации осуществляется методом опроса каналов или широковещательной рассылки, нередко необходима "подписка" на требуемые новости.

Базы данных не имеют прямой поддержки в Internet/Intranet. Доступ и ведение БД производится через серверные или клиентские приложения. При этом запросы на поиск и ввод данных оформляются в виде HTML-форм.

Хранилище данных может иметь различную реализацию (реляционная многомерная БД, совокупность ODBC-источников данных), но предназначено для целей оперативной аналитической обработки данных. Доступ к хранилищу организуется подобно доступу к базе данных. Нужно отметить, что аналитическая обработка требует наличия средств для динамического построения и манипулирования плоскими и трехмерными объектами деловой и научной графики.

Программное обеспечение в виде клиентских и серверных приложений составляет важный ресурс. Гипертекст служит средством создания "меню", ссылающегося на разные общие и специальные прикладные программы.

Для ведения перечисленных информационных ресурсов и организации потоков между клиентами и серверами используются различные средства разработки, эксплуатации и сопровождения Internet/Intranet-приложений.

**1.2.2. Средства разработки, эксплуатации и сопровождения Internet/Intranet-приложений**

В состав средств, предназначенных для разработки Internet/Intranet-приложений, их эксплуатации и сопровождения, входят следующие группы:

* средства Run-time (времени выполнения);
* инструментальные средства разработки;
* средства администрирования.

Средства Run-time включают:

* программы просмотра и навигации (броузеры);
* клиентские приложения и расширения;
* программное обеспечение Web-серверов;
* серверные приложения и расширения;
* средства поиска информации;
* программное обеспечение Internet-серверов;
* средства безопасности.

Программы просмотра и навигации обеспечивают интерпретацию гипертекста или гипермедиа, организацию диалога, активацию и выполнение клиентских приложений и расширений, обращение к серверным приложениям с передачей входных данных. В броузер встроены интерпретаторы сценариев и мобильных кодов. Броузер может служить унифицированным интерфейсом к различным информационным ресурсам.

Клиентские приложения имеют различную природу. Это может быть, прежде всего, гипертекст или гипермедиа. Дополнительную функциональность обеспечивают сценарии на языках JavaScript или VBScript, мобильные Java-апплеты, helper-программы (редакторы текста, процессоры электронных таблиц и другие готовые программы, активируемые в зависимости от типа файла) и клиентские расширения (ActiveX и Plug-in).

Программное обеспечение Web-сервера осуществляет передачу гипертекста, гипермедиа и других файлов клиентам по запросам, активацию серверных приложений, связь с файл-серверами и серверами баз данных.

Серверные приложения представляют собой загрузочные модули или сценарии для интерпретации, выполняемые на Web-сервере. Серверные расширения - библиотеки времени выполнения, к которым обращаются серверные сценарии.

Средства поиска информации помогают найти в Internet/Intranet необходимые сведения, удовлетворяющие условиям поиска. Для этого используются как поисковые машины, так и средства управления документами с полнотекстовым индексированием.

Программное обеспечение Internet-серверов служит для поддержки электронной почты, FTP-сервиса для передачи файлов, возможностей доступа к новостям и др.

Средства безопасности могут быть встроены в ПО Internet-серверов или представлены в виде дополнительных компонентов: комплексов Firewall и Proxy-серверов, выполняющих фильтрацию на различных уровнях.

Инструментальные средства разработки Internet/Intranet-приложений очень разнообразны и включают:

* редакторы гипертекста;
* графические редакторы и конверторы изображений;
* средства разметки карт изображений;
* средства мультимедиа (аудио, анимация, видео);
* средства генерации виртуальной реальности;
* системы программирования клиентских приложений;
* средства программирования серверных приложений;
* системы программирования для создания клиентских и серверных расширений.

Редакторы гипертекста предназначены для формирования HTML-файлов в режимах программирования или WYSIWYG. Для создания гипертекста могут использоваться и обычные текстовые редакторы, а также средства, встроенные в броузеры. К этой же группе относятся конверторы офисных документов в гипертекст.

Графические редакторы служат для создания изображений, включаемых в гипертекст. Конверторы изображений обеспечивают преобразование форматов, размеров и цветов, создание специальных эффектов.

Средства разметки карт изображений позволяют разбить изображение на участки и связать с каждым из них гиперссылки. Такие средства могут быть встроены в графический редактор.

Средства мультимедиа предназначены для создания звукового и музыкального сопровождения, анимационных и видео роликов. Часто воспроизведение файлов мультимедиа осуществляется клиентскими расширениями или helper-программами.

Средства генерации виртуальной реальности позволяют запрограммировать трехмерные сцены и управление ими на языке VRML. Воспроизведение виртуальной реальности может потребовать дополнительных средств.

Системы программирования клиентских приложений предназначены для разработки и отладки сценариев (на языках JavaScript, VBScript) и мобильных приложений (на языке Java), выполняемых на стороне клиента, наибольшие удобство и производительность разработки дают средства визуального программирования.

В качестве средств программирования серверных приложений могут использоваться как обычные системы программирования (C/C++, VisualBasic, Java и др.), так и интерпретаторы команд (UNIX-shell, REXX и др.), интерпретаторы и компиляторы сценариев на JavaScript, VBScript и Perl.

Для создания клиентских и серверных расширений используются системы программирования, позволяющие создавать компоненты с использованием механизмов ActiveX или Plug-in, представленных в виде встроенных или дополнительных библиотек интерфейсов.

Средства администрирования, как правило, поставляются в составе ПО Web-сервера и служат для конфигурирования, активации и мониторинга Internet-сервисов, для настройки системы безопасности, для контроля связности гипертекстовой структуры и актуальности гиперссылок, для учета и протоколирования использования серверов.

Далее будем рассматривать в основном инструментальные средства разработки Intranet-приложений и только по необходимости касаться средств Run-time и администрирования.

**1.2.3. Традиционная схема публикации и просмотра документов**

Изначально технология Internet/Intranet/WWW предназначалась для облегчения доступа к информации и публикации документов (рисунок 1.1). Программа-клиент (броузер) выполняет функции интерфейса пользователя (диалоговый ввод-вывод PS и логика диалога PL) и обеспечивает доступ практически ко всем информационным ресурсам Internet/Intranet посредством HTTP-сервиса. База данных гипертекста - это часть файловой системы, которая содержит текстовые файлы в формате HTML и связанные с ними графику и другие ресурсы. Фактически, броузер является интерпретатором HTML-текста. И как типичный интерпретатор клиент в зависимости от команд разметки выполняет различные функции. В круг этих функций входит не только размещение текста на экране, но и обмен информацией с сервером по мере анализа полученного HTML-текста, что наиболее наглядно происходит при отображении встроенных графических образов. При анализе URL-спецификаций или по командам сервера клиент запускает дополнительные help-программы для работы с документами в форматах, отличных от HTML, например, GIF, JPEG, MPEG, Postscript и т.п. Кроме того, в гипертекст может быть встроен текст сценария на JavaScipt, VBScript и др., который расширяет логику диалога (PL) и обеспечивает простую прикладную обработку (BL). Броузер вызывает соответствующий интерпретатор для исполнения сценария.

Рис. 1.1. Традиционная схема публикации и просмотра документов

До недавнего времени сеть Internet была "улицей с односторонним движением" - информация с Web-страниц поступала к пользователю от Web-сервера при наличии запроса. С появлением в языке HTML диалоговых свойств пользователь получил обратную связь с Web-сервером. Для обеспечения специальной обработки (поиска, карт изображений, анкетных листов, включений и т.п.) на Web-сервере может быть запущена серверная программа, которая обменивается параметрами с броузером. Обмен параметров при этом осуществляется через интерфейс CGI. В последнее время все большее распространение получает механизм согласования запускаемых программ через MIME-типы.