

ПРИМЕНЕНИЕ СОВРЕМЕННЫХ WEB-ТЕХНОЛОГИЙ ПРИ СОЗДАНИИ ЭЛЕКТРОННЫХ ЭНЦИКЛОПЕДИЙ

Марусин Д.В., Шаркевич И.В.

(Волгоград)

В работе описываются современные подходы к построению веб-ориентированных информационных систем, и предлагается новый способ построения таких систем. Способ заключается в применении возможностей операционной системы Windows, в частности встраиваемых протоколов, для создания программного продукта, способного без изменения функционировать в нескольких операционных средах. В работе приводится описание технологии и примеры ее использования для создания электронной энциклопедии.

APPLICATION OF MODERN WEB-TECHNOLOGIES FOR CREATION OF ELECTRONIC ENCYCLOPEDIAS

Marusin D.V., Sharkevich I.V.

(Volgograd)

Modern approaches for construction web-based information systems are described and newly developed way for creating such systems is offered. The approach is based on application of features of Windows operation system, in particular technologies of asynchronous pluggable protocol handlers, for creation a software product, capable to function in different environments without any changes. Description of this technology, way it works and usage examples included.

В последнее время интенсификация информационных обменов в различных областях знаний приводит к тому, что значительно увеличивается объем информации, требующей ее оперативного анализа. Решить эту проблему позволяют современные информационные технологии, предоставляющие эффективные

средства поддержки принятия решений, представления и анализа данных. Одним из перспективных направлений накопления и структурирования знаний являются электронные энциклопедии в определенных предметных областях. Традиционно энциклопедия представляется в виде совокупности статей, объединенных в информационные блоки с помощью перекрестных ссылок. Такая концепция организации знаний в электронной форме может быть построена с помощью гипертекстовой технологии WWW.

Разработка энциклопедических информационных систем, обладающих web-интерфейсом, прежде всего подразумевает их применение в сети Интернет. В этом случае система обладает необходимым инструментарием для накопления, обработки и анализа информации, является делокализованной, и имеет максимально широкий круг пользователей. При этом к архитектуре информационной системы предъявляется ряд серьезных требований, в частности, обеспечение многопользовательской работы в режиме клиент-сервер, что обуславливает специфический способ ее организации (рис. 1).

При такой организации информационная система представляет собой совокупность:

- веб-страниц, образующих пользовательский интерфейс;
- клиентских скриптов, управляющих поведением клиентских мест;
- серверных скриптов, реализующих бизнес-логику системы (в основном это технологии ASP, JSP, CGI, PHP, SSI и т.п.);
- серверов приложений, обеспечивающих базовые функции системы;
- веб-сервера, являющегося «центром» всей системы и связывающего все элементы между собой.

Это приводит к тому, что система перестает быть мобильной и переносимой; она создается в единственном варианте, жестко привязана к конкретной среде исполнения и т.п. Существует большое количество способов преодолеть эти трудности, например, оформление системы как единого клиент-серверного проекта в рамках одной архитектуры.

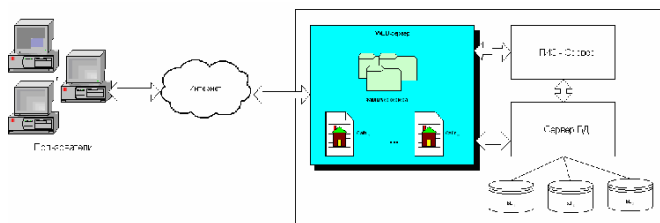


Рис. 1. Традиционная архитектура веб – ориентированной информационной системы.

Помимо разработки таких Интернет-ориентированных информационных систем, представляется актуальным разработка информационных систем того же уровня удобства и функциональности, с применением тех же технологий, но способных работать вне сети Internet. В этой связи становится важным вопрос поиска технологий позволяющих проектировать и реализовывать информационные системы, способные функционировать в двух вариантах использования:

- *локальном*, при котором система работает на изолированном рабочем месте, не имеющем доступа к сетевым ресурсам;
- *сетевом*, когда система работает как web-сайт с использованием Интернет или Интранет.

Целью данного исследования являлось разработка комплекса инструментов и технологий, позволяющих создать систему обработки и анализа данных, способную работать в обоих вариантах использования без нарушения ее функциональности. Основными задачами являлось разработка приемлемого по сложности технического решения и его реализация для ОС Windows с использованием WEB-интерфейса.

Разработанная в рамках исследования технология *Virtual WEB (VWEB)* для решения поставленных задач использует возможности стандартных прикладных интерфейсов (API) ОС Windows и браузера Internet Explorer. Имеющийся в составе ОС Windows модуль URLMON, управляющий подключениями к Internet, позволяет применять в прикладных программах так называемые фильтры для расширения возможностей браузера. Описываемая технология VWEB использует возможности одно-

го из видов фильтров, так называемые встраиваемые протоколы (Asynchronous Pluggable Protocols), с помощью которых обеспечивается единообразный доступ к ресурсам информационной системы в обоих вариантах функционирования.

Технология встраиваемых протоколов позволяет любому приложению ОС Windows получить доступ к самым разным информационным ресурсам сети Интернет и прочих источников через стандартный интерфейс. Достигается это тем, что ОС Windows регистрирует для каждого протокола Интернет *обработчик*, который отвечает за интерпретацию запросов, их исполнение и возврат результатов. Никаких ограничений на внутреннее устройство обработчика не налагается, что позволяет реализовать любой механизм получения данных. Эта технология широко используется в приложениях ОС Windows, например, в системе гипертекстовой справки HTML Help, стандартные протоколы Интернет и шлюз моникеров Windows. Обработчики протоколов могут быть постоянными и временными. Первые регистрируются в реестре и доступны в каждый момент времени, а вторые устанавливаются только на время функционирования конкретного приложения и доступны только для него.

Основным требованием для встраиваемых протоколов является использование COM-технологии для взаимодействия с ОС Windows. В соответствии с этим обработчик протокола представляет собой совокупность классов, реализующих те или иные интерфейсы, главным из которых является *Internet*. Использование абстрактных объектов для доступа к Интернет-ресурсам позволяет использовать Интернет-технологии даже там, где нет подключения к локальной сети, не требуя при этом даже наличия сетевой поддержки в операционной системе.

В соответствии с этими требованиями, ядро технологии VWEB выполнено в виде COM-компонента, помещенного в модуль DLL, с применением технологий DCOM, ActiveX, WebSnap, DataSnap. Этот компонент состоит из нескольких элементов (см. рис. 2):

- блок обработки и распределения запросов (Обработчик запросов);
- блок исполнения (WebSnap и Внутренняя логика ИС);

- интерфейс доступа к данным (Диспетчер и Хранилище данных);
- локальные информационные ресурсы (Базы данных и Локальные файлы).

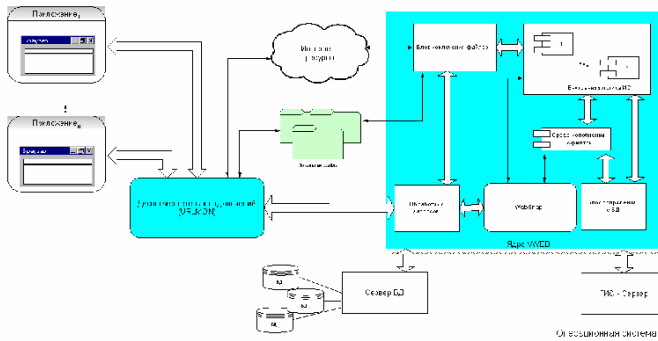


Рис. 2. Архитектура веб-ориентированной информационной системы с использованием технологии VWEB.

При использовании технологии VWEB технологический цикл проектирования и реализации информационной системы осуществляется только для построения Internet-варианта. Для того чтобы выпустить систему в локальном варианте, разрабатывается вспомогательное приложение-оболочка, к которому подключается ядро VWEB и архив файлов, из которых состоит web-сайт системы. К оболочке подключаются специализированные компоненты-заместители серверных технологий, использующиеся в сетевом варианте. При этом сама система (файлы, данные и метаданные web-сайта) не изменяется.

При работе в локальном варианте пользователь использует программу-оболочку для доступа к системе. Для него оболочка представляется как специализированная версия браузера Internet Explorer. С помощью этого браузера он генерирует http-запросы, которые передаются в модуль URLMON. После их анализа вызывается соответствующий обработчик протоколов, например, для адреса <http://www.yandex.ru> используется стандартный протокол HTTP. При указании протокола `vweb://` управление передается на блок обработки запросов (см. рис. 2). После интерпретации запроса, управление передается либо в

блок исполнения, либо в блок извлечения файлов.

Блок обработки запросов является связующим звеном между инфраструктурой встраиваемых протоколов и внутренними компонентами VWEB. Он выполнен как COM-объект и реализует ряд интерфейсов: стандартные интерфейсы (Internet, InternetProtocol, InternetProtocolSink, InternetBindInfo), интерфейсы протокола HTTP (IWinInetInfo, IWinInetHttpInfo), интерфейсы потокового доступа к данным. Реализация двух последних групп интерфейсов потребовалась вследствие того, что информационная система может использовать не только статические, цельные документы, но и средства гипермедиа. Последние по своей природе являются потоковыми, то есть их использование начинается еще до полного получения документа. Примерами гипермедиа средств могут служить звуковые клипы в форматах MP3, ASF, WMA, RealAudio, видео-клипы в форматах MPEG, RealVideo, WMV, ASF, анимационные клипы в форматах Flash, Shockwave и многие другие.

Блок исполнения активируется для веб-страниц, содержащих серверные скрипты технологии ASP (с расширением.asp). Этот блок выполнен на основе технологии WebSnap компании Inprise, которая применяется для создания специализированных веб-серверов. При активации блока исполнения для веб-страницы имитируется среда исполнения веб-сервера, что обеспечивает полиморфность информационной системы. Средства VWEB обеспечивают все механизмы обычного веб-сервера, в том числе компоненты доступа к данным, серверы приложений и т.п. Использование последних означает передачу управления в блок внутренней логики к соответствующему блоку-заместителю. Таким образом реализуется полиморфный характер информационной системы – при работе в Интернет информационная система обращается к реальным серверам, службам и компонентам; при работе в локальном – к их заместителям.

Блок извлечения отвечает за работу со статичными файлами системы и, при необходимости, с внешними Интернет-ресурсами. При получении запроса он извлекает требуемые файлы из архива системы или переадресует запрос стандартному обработчику HTTP. Запрос может поступить как из блока обработки запросов (от пользователя), так и из блока исполне-

ния или блока внутренней логики.

Блок внутренней логики состоит из компонентов-заместителей серверных технологий Интернет-варианта и набора СОМ-объектов, реализующих бизнес-логику системы. Конкретный состав этого блока определяется спецификой информационной системы.

В результате, ядро технологии VWEB представляет собой специализированный web-сервер, не требующий наличия сетевой инфраструктуры на рабочем месте пользователя.

Наибольшую сложность при использовании такого подхода представляет создание компонентов-заместителей стандартных серверных технологий. Потребность в этих компонентах вызвана тем, что для системы, разработанной для сети Интернет, в локальном варианте нецелесообразно применять те же решения, что и сетевом варианте. Прежде всего это потребует значительно больших аппаратных ресурсов и усложнит использование системы. В ходе исследования были найдены решения, заменяющие стандартные серверные технологии с сохранением требуемой функциональности. Рассмотрим, например, реализацию компонента доступа к данным. В сетевом варианте информационная система обязана использовать многопользовательскую СУБД, но в локальном варианте это не обязательно. Поэтому при создании локальной системы возможно применить настольную версию той же СУБД, снизив при этом сложность системы (для СУБД SQL Server применяется SQL Server Desktop Engine и т.п.). При использовании стандартных интерфейсов доступа к данным (ADO, ODBC, JDBC) функции заместителя сводятся к замене одного драйвера СУБД. Аналогичным образом можно поступить и с другими компонентами, например с доступом к функциям ГИС-сервера.

В настоящем исследовании технология VWEB была использована для создания локальной версии экономической Интернет-энциклопедии Волгоградской области. Данный проект разрабатывается в Волгоградском государственном университете в рамках построения единого информационного пространства нижневолжского региона. Энциклопедия представляет собой веб-сайт, построенный с применением технологии ASP, взаимодействующий с базой данных под управлением СУБД Interbase,

ГИС – сервером и набором СОМ-объектов бизнес-логики. При построении локальной версии системы в среде Inprise Delphi была создана оболочка, использующая технологию VWEB, к которой подключены адаптеры доступа к СУБД и ГИС, разработанные в рамках исследования. Блок поиска, использованный в сетевом варианте, заменен аналогичным компонентом VCL, полученным из сети Интернет.

В результате проведенного исследования показано, что применение технологии VWEB позволяет существенно снизить трудоемкость процесса разработки электронной энциклопедии, способной функционировать в двух средах – локальной и сетевой. Разработанная система была спроектирована и только в сетевом варианте, а локальный получен путем автоматической генерации с помощью мастера. Это позволило избежать создания и сопровождения двух идентичных систем, сконцентрировавшись на создании единого прототипа.