

СОЗДАНИЕ РАСПРЕДЕЛЕННОЙ СИСТЕМЫ РАЗМЕЩЕНИЯ И УПРАВЛЕНИЯ ИНФОРМАЦИЕЙ НА КОРПОРАТИВНОМ ПОРТАЛЕ «АНАЛИТИЧЕСКАЯ ХИМИЯ В РОССИИ»

Колотов В.П., Широкова В.И., Аленина М.В.

Институт геохимии и аналитической химии им. В.И.Вернадского РАН
119991, ГСП-1, Москва В-334, ул. Косыгина, 19
shirokova@geokhi.ru

DEVELOPMENT OF DISTRIBUTED SYSTEM FOR INFORMATION ALLOCATION AND CONTROL ON THE CORPORATE WEB PORTAL “ANALYTICAL CHEMISTRY IN RUSSIA”

Shirokova V.I., Kolotov V.P., Alenina M.V

Institute of Geochemistry and Analytical Chemistry of Russian Academy of Sciences,
119991, GSP-1, Moscow B-334, Kosygin Str., 19, Russia
shirokova@geokhi.ru

A new Internet portal developed by community of Russian analysts has been launched in 2001 (<http://www.geokhi.ru/~rusanalytchem>, <http://www.rusanalytchem.org>). Now the portal contains a large amount of information, great part of it is stored in the form of SQL data base (MS SQL). The information retrieval is made by means of ASP pages, containing VB Scripts.

The obtained experience of work with such topical portal has detected some weak points, related with its centralized administration and updating.

It has been found that urgent supporting of all requests from different persons/organizations on information allocation on the portal's server takes a lot of efforts and time. That is why, the further development of portal we relate with development of a distributed system for information allocation and control, under preserving of centralized administration for ensuring of security and stable working of the portal. Analysis and testing of some available technologies lead us to conclusion to apply MS Share Point technologies. A MS Share Point Team Services (SPTS) has been selected as a technology supporting relatively small groups, where MS SQL is used for storage data and metadata. The last feature was considered as decisive one for SPTS selection, allowing easy integration with data base of the whole portal.

SPTS was launched as an independent Internet site accessible from home page of the portal. It serves as a root site to exit to dozens of subsites serving different bodies of Russian Scientific Council on analytical chemistry and external organizations located over the whole Russia. The secure functioning of such hierarchical system, which includes a lot of remote information suppliers, based

on use of roles to manage user rights independently for each subsite. The root site is controlled by portal administrator, whereas the daughter subsite has additional remote administrator representing the correspondent body. The remote administrator is responsible for arrangement of roles for contributors and information representation on the site.

For synchronization of data in the frames of the portal special MS SQL stored procedures have been developed. Now the developed system is under testing, but few first months of its functioning demonstrated high efficiency of such solution for Internet corporate portal.

The work is supported by Russian Foundation for Basic Research (RFBR), grant N01-03-32233

С апреля 2001 г. функционирует корпоративный Интернет-портал [1, 2], созданный сообществом российских специалистов в области аналитической химии. К настоящему времени портал содержит большой объем информации, которая частично загружена в базу данных (MS SQL). Выборка проводится с помощью ASP-страниц, содержащих программы на Visual Basic Script. Информационные характеристики портала: объем ~50 МБ, более 1300 html и ASP файлов, объем 3 файлов баз данных составляет >25 МБ). Для выборки и систематизации новых ресурсов Интернета в области аналитической химии разрабатывается поисковая машина второго уровня. Найденные ресурсы загружаются в базу данных для обеспечения on-line тематических запросов. Для поддержания целостности базы данных система периодически проверяет работоспособность всех Интернет-гиперссылок, находящихся в базе данных. Ссылка и вся связанная информация перегружается в архивные таблицы, если доступ по ссылке невозможен по каким-то причинам в течении месяца.

Вместе с тем, опыт эксплуатации тематического портала показал и некоторые ограничения, свойственные организации портала в традиционном варианте, когда портал администрируется и пополняется информацией централизованно.

Так, например, одной из задач портала, поставленной при его организации была электронная поддержка Научного Совета по аналитической химии РАН, объединяющего ведущих специалистов, работающих в различных регионах России и некоторых зарубежных странах. Структура Совета достаточно сложная и включает в себя многочисленные комиссии по направлениям, научные ассоциации и семинары, региональные бюро и др. Кроме того, под эгидой Совета часто проводятся различные научные конференции и другие мероприятия. Практика показала, что поддержка всех оперативных запросов с мест на размещение информации в портале является не простой задачей (главным образом в организационном плане). Кроме того, информация, поставляемая структурами Совета, подвергается

периодическому пересмотру, что ведет к необходимости ее синхронизации с информацией, отображаемой на страницах портала.

Таким образом, совершенно очевидно возникает потребность в разработке средств распределенного размещения и управления информацией при сохранении централизованного администрирования для обеспечения безопасности и устойчивого функционирования портала. Мы сознательно отказались от использования известных средств, основанных на программировании дистанционного доступа к базе данных портала в силу того, что учесть все потребности большого числа клиентов очень сложно и разработка такой системы потребует много времени. Внимание было обращено на готовые решения, такие как MS Share Point Team Services (SPTS) и Share Point Portal (SPP), предназначенные для организации корпоративной работы, что отвечает части задач портала. Названные решения достаточно похожи, но рекомендованы для различного масштаба применения. SPP предназначен для больших корпораций, использует развитые средства Dashboard сайта как интерфейса и специальное хранилище информации, в то время как SPTS предназначен для меньших групп и использует MS SQL для хранения данных и метаданных. Последнее обстоятельство оказалось решающим в пользу выбора SPTS, так как обеспечивало наиболее простой способ интеграции распределенной информации с центральной базой данных портала.

SPTS был развернут в виде сайта, доступ к которому имеется с домашней страницы портала. Этот сайт является головным для выхода в десятки субсайтов, предназначенных для организации работ подразделений Совета, в том числе и готовящихся конференций. Функционирование такой иерархической системы, включающей большое число удаленных поставщиков информации основано на ролевой системе безопасности, когда каждый сайт (субсайт) имеет свою независимую систему учетных записей, назначающую посетителям определенные права. При попытке изменить информацию сайта (базы данных) система запрашивает учетное имя и пароль и далее проверяет возможность выполнения планируемого изменения в рамках назначенной роли.

Корневой SPTS сайт управляется администратором портала, в то время как каждый субсайт имеет 2 администратора (один из которых представляет соответствующее удаленное подразделение, а второй является администратором портала). Удаленный администратор назначает роли авторам субсайта (также удаленным), а также определяет общую настройку представления информации на субсайте. На субсайтах размещены новости, объявления, общие файлы, проводятся обсуждения по различным темам, размещены различные базы данных. Посетители, которым назначена определенная роль, должны ввести пароль прежде, чем им будет разрешено менять содержание субсайта или размещать на нем файлы (анонимный доступ разрешен и позволяет только просматривать субсайт и документы).

Вся информация, включая метаданные, находится в общей базе данных SPTS сайта, что обеспечивает легкую интеграцию информации в пределах портала. Для синхронизации данных в пределах портала разработаны хранимые процедуры для MS SQL сервера, поддержка которых и является задачей администратора портала. Трудоемкость этой процедуры существенно меньше, чем непосредственное размещение информации, более того, децентрализованная загрузка информации (в данном случае на уровне портала), соответствует основной идеологии развития Интернета, доказавшей свою эффективность.

Система находится в процессе обкатки, но уже первые месяцы ее функционирования показали высокую эффективность такого решения.

Хотя SPTS и SPS позиционированы прежде всего для применения в корпоративной интрасети, тем не менее полученный опыт свидетельствует и в пользу Интернет-применения SPTS. К сожалению, миграция от SPTS к SPS (функционально более развитому) невозможна. Более того, эти 2 решения не могут сосуществовать на одном сервере.

Работа проводится при поддержке Российского фонда фундаментальных исследований (грант № 01-03-32233).

Литература

- [1] Корпоративный портал «Аналитическая химия в России»: <http://www.geokhi.ru/~rusanalytchem>, <http://www.rusanalytchem.org>
- [2] Колотов В.П., Широкова В.И., Аленина М.В. «Создание и развитие информационно-поискового портала “Аналитическая химия в России”» / III Всероссийская научная конференция «Электронные библиотеки: перспективные методы и технологии, электронные коллекции». RCDL'2001, Петрозаводск, 11-13 сентября 2001 г. (<http://rcdl2001.krc.karelia.ru/posters.ru.shtml>)