

Результаты разработки распределённой Web-ориентированной системы управления библиографическими базами данных

Design of the Distributed Web-Oriented Bibliographic Database Management System

Результати розробки розподіленої Web-орієнтованої системи управління бібліографічними базами даних

Баженов С. Р., Федотов В. Б., Баженов И. С.

*Государственная публичная научно-техническая библиотека СО РАН,
Новосибирск, Россия*

S. R. Bazhenov, V. B. Fedotov, and I. S. Bazhenov

*State Public Library for Science and Technology, Siberian Division
of the Russian Academy of Sciences, Novosibirsk, Russia*

Баженов С. Р., Федотов В. Б., Баженов И. С.

*Державна публічна науково-технічна бібліотека СВ РАН,
Новосибірськ, Росія*

Описывается разработанная в ГПНТБ СО РАН распределённая Web-ориентированная система управления библиографическими базами данных CDS/ISIS, включающая в себя доступ к базам данных, предоставляемым по протоколу Z39.50. Рассматриваются особенности реализованной системы: возможности организации доступа к удалённым информационным источникам, гибкое распределение вычислительной нагрузки между серверами БД.

This is a description of the web-oriented bibliographic database management system (CDS/ISIS) that has been designed by the State Public Library for Science and Technology. The system offers Z39.50 access to the databases. Specific features of this system are studied, i. e., access to remote information resources, flexible distribution of computer load between database servers, etc.

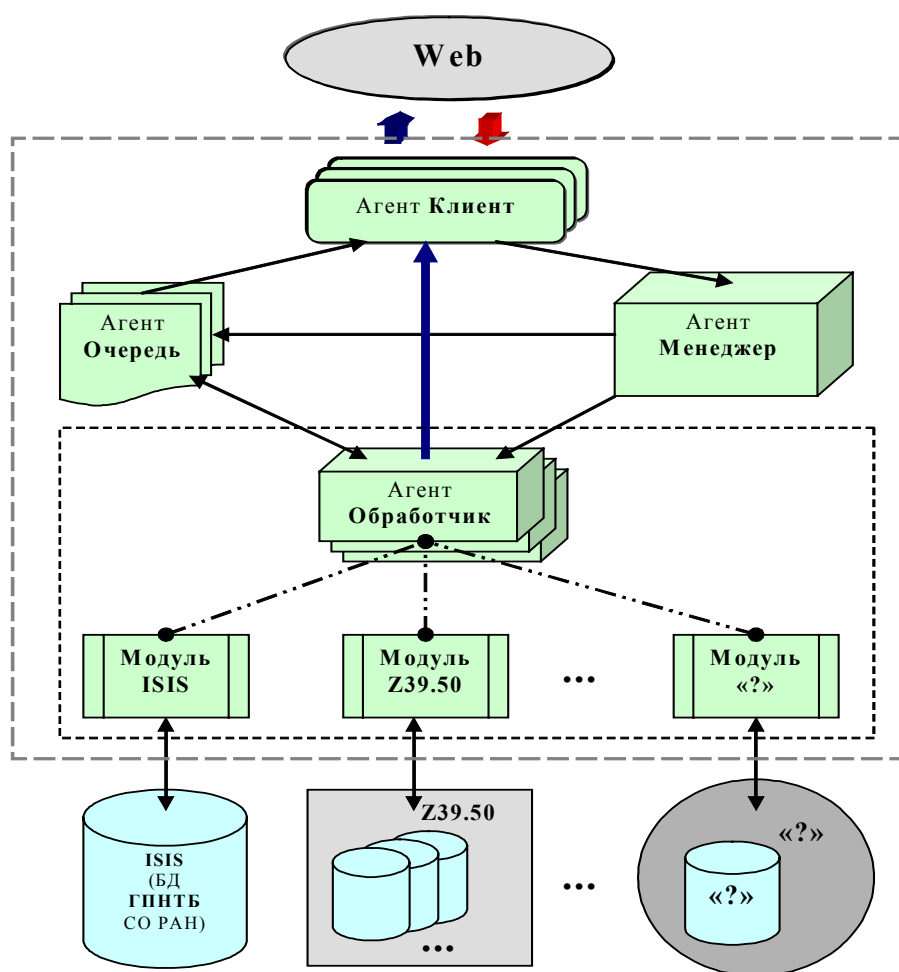
Описується розроблена в ДПНТБ СВ РАН розподілена Web-орієнтована система управління бібліографічними базами даних CDS/ISIS, яка включає в себе доступ до баз даних, що надаються за протоколом Z39.50. Розглядаються особливості реалізованої системи: можливості організації доступу до віддалених інформаційних джерел, гнучке розподілення обчислювального навантаження між серверами БД.

По мере надобности в течение эксплуатации WEB-ориентированной информационно-поисковой системы для доступа к базам данных CDS/ISIS [1-3] в неё было успешно внесено большое количество модификаций. Система, архитектура которой изначально проектировалась для обеспечения Web-доступа к локальным базам данных, многолетней эксплуатацией доказала свою эффективность.

Однако, в связи с внедрением в технологию системы WinIRBIS произошли изменения в структурах, поддерживаемых ISIS-совместимых баз данных. Кроме того, возникла потребность в обеспечения доступа к удалённым библиографическим базам данных корпоративной библиотечной системы по протоколу Z39.50 [4-6]. Однако, при близком рассмотрении, расширение функциональности оказалось затруднительным. Можно сказать, что постоянная модификация и доделка системы отрицательно сказались на её структуре, удобстве работы администратора поисковой системы и интерфейсе пользователя, и постепенно привели систему к состоянию, когда дальнейшее расширение функциональности стало мало возможным.

Вывод, сделанный из этой ситуации, заключался в том, что система не только нуждается в кардинальном совершенствовании, но и сама концепция организации доступа нуждается в пересмотре. Вследствие чего была поставлена и успешно решена задача проектирования и реализации, концептуально новой версии поисковой системы, включающей в себя как все наработки с момента запуска предыдущей системы в эксплуатацию, так и многочисленные изменения и дополнения.

Архитектура многоагентной системы доступа к распределённым информационным ресурсам



В описываемой поисковой системе многоагентная система реализована следующим образом:

- через Web-интерфейс вызывается небольшой исполняемый модуль (Клиент), который принимает запрос с Web-страницы и передаёт его в очередь запросов, после чего ждет ответа от агентов через «сокет» (socket);
- автономные агенты на удаленных серверах (преимущественно там, где находятся базы данных) получают запрос из очереди (их может быть несколько, на разных серверах), выполняют его и посылают результат соответствующему агенту Клиент, который в свою очередь возвращает результат на Web-страницу.

Общие элементы программной архитектуры агентов

Все агенты в системе имеют общий программный костяк, базирующийся на передаче/приёме сообщений посредством коммуникации через транспортную среду. Транспортной средой служит сеть Internet. Каждый агент занимает на компьютере ровно один порт, через который он принимает и отправляет данные. Разработанная транспортная библиотека использует протокол передачи датаграмм и обеспечивает надёжную доставку сообщений. Для каждого агента, создаётся отдельный программный поток приёма сообщений, обеспечивающий прослушивание порта вне зависимости от занятости агента. Принимаемые агентом сообщения буферизуются в очередь для последовательной обработки.

Основной процесс

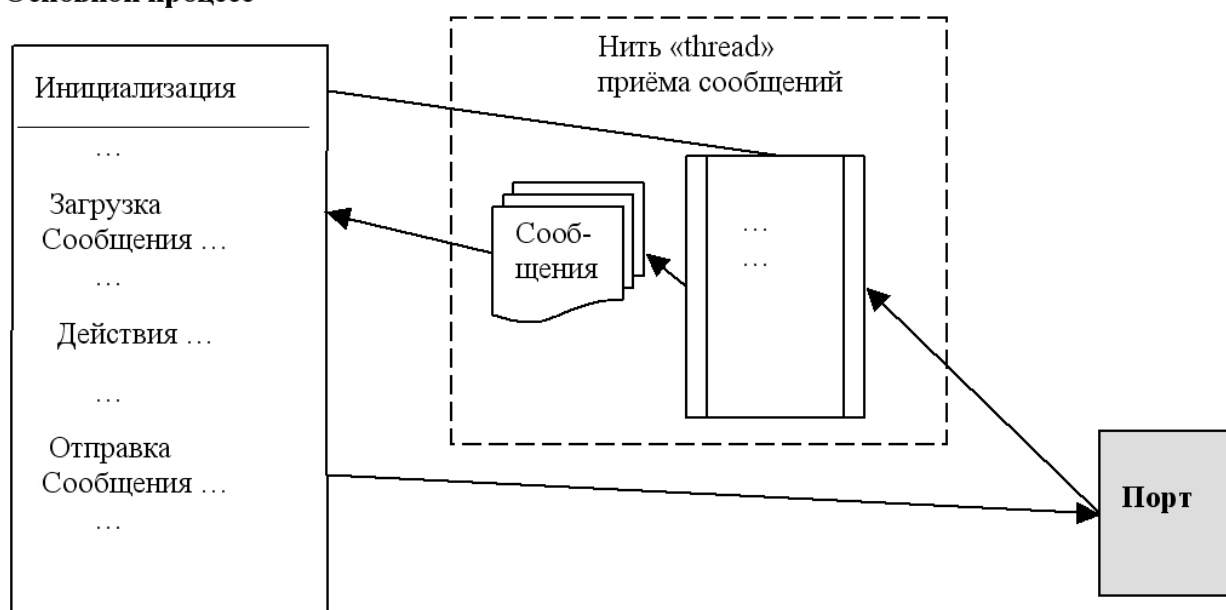


Рис. Общие элементы программной архитектуры агентов распределённой информационной системы

Основными преимуществами разработанной информационной системы являются:

- информационная распределенность: нет необходимости собирать все базы данных на одном сервере, достаточно установить соответствующего формата баз данных агента на удаленном сервере. Данная функция очень полезна для групп организаций, которые не могут установить свою поисковую систему, но могут предоставить сетевой доступ;
- распределение вычислительной нагрузки между серверами: обработка запросов к базам данных может быть распределена между серверами, что позволяет решать проблему высокой нагрузки на сервер БД не за счёт приобретения нового высоко-производительного оборудования, а за счёт использования двух и более серверов меньшей производительности. Использование более одного сервера для обработки запросов не только позволяет экономить средства на приобретении нового оборудования, но и позволяет достичь более высокой скорости работы системы (примеры графиков распределения нагрузки на систему приведены в докладе);
- повышенная отказоустойчивость: можно запустить несколько одинаковых агентов (как на одном, так и на разных серверах) и, если один из агентов перестает выполнять запросы, другой обрабатывает запросы вместо него;
- обеспечение одновременного поиска изданий во множестве библиографических БД, как основанных на системе ISIS, так и в любых других предоставляемых по протоколу Z39.50;
- высокая скорость работы информационно-поисковой системы во множестве различных информационных источников.

Полученная информационная система легко расширяема

- для добавления в систему удалённой БД достаточно, если это БД ISIS:
 - Чтобы она находилась в поле видимости хотя бы одного агента «Обработчик»;
 - Необходимо зарегистрировать базу данных в общем каталоге;
 - если нужно добавить удалённую БД по протоколу Z39.50 то для этого достаточно просто зарегистрировать её в каталоге
- для добавления новой поисковой функциональности, например доступ к реляционным БД. Нужно добавить к агенту «Обработчик» плагины новых типов и зарегистрировать в системе ещё один тип обрабатываемых запросов
- система сама естественным образом может распределять нагрузку обработки запросов к БД между разными рабочими станциями. Агенты, выполняющие запросы общих типов, конкурируют друг с другом за право выполнения запроса.

Архитектура информационной системы обладает высокой гибкостью

- при условии сохранения коммуникационного стандарта между агентами, любой отдельный узел (агент) можно заменить, не переделывая при этом всю программу и не меняя архитектуру в целом
- добавление в систему новых БД не требует модификации программного кода или даже остановки работы системы и производится «на лету»
- разработанная система является полностью распределённой — любая часть (агент) системы может быть установлена на удалённом компьютере с коммуникацией основанной на одном из базовых протоколов сети Internet — протоколе UDP

Как результат работы над унификацией программной архитектуры агентов разработана унифицированная модульная программная архитектура агента, воплощённая в программную реализацию комплекса WSDS («Web Search Distributed System»). При выборе программных модулей ядра комплекса WSDS в качестве основы разработки многоагентных распределённых систем, время программной реализации может быть существенно сокращено, давая разработчику агентов возможность не беспокоиться об общих для агентов функциях и концентрируя внимание только на реализации основной функциональности распределённой системы.

Программный комплекс WSDS позиционируется как программное обеспечение промежуточного слоя выступающее связующим звеном между интерфейсами доступа (в первую очередь Web-интерфейсами) и конкретными системами управления данными. Программный комплекс WSDS может быть использован для организации широкого круга библиотечных распределённых информационных систем, организации WEB-шлюзов сводных библиографических каталогов, обеспечению интерфейсов доступа к полнотекстовым и другим распределённым информационным ресурсам. В сочетании с Web-интерфейсом, реализованный программный комплекс является удобным и эффективным средством доступа к информационным ресурсам региональной библиотечной системы.

Литература

1. Баженов С. Р., Мазов Н. А., Малицкий Н. А., Баженов И. С. Создание программного комплекса доступа из Интернет к базам данных на основе WWW-ISIS // Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества: Материалы конф. «Крым 98». — М, 1998. — Т.2. — С.581 — 584.
2. Мазов Н. А., Малицкий Н. А., Баженов С. Р., Баженов И. С., Жижимов О. Л. WEB-ориентированная информационно-поисковая система для доступа к базам данных CDS/ISIS // Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества: Материалы 6 Междунар. конф. «Крым 99». — Судак, 1999. — Т.1. — С.113 — 115.
3. Баженов С. Р., Мазов Н. А., Малицкий Н. А., Баженов И. С. Создание программного комплекса доступа из Интернет к базам данных на основе WWW-ISIS// Научные и технические библиотеки. — 1999. — N2. — С.47 — 52.
4. Жижимов О. Л. Введение в Z39.50. Новосибирск, Изд-во НГОНБ, 2000, 196 с., ISBN 5-88742-037-5.
5. Мазов Н. А., Жижимов О. Л. CDS/ISIS и Z39.50: проблемы интеграции и решения // IV рабочее совещание по электронным публикациям — EL-PUB-99, Новосибирск, Академгородок, ИВТ СО РАН, 21-23 апреля 1999 года.: Тез. докл., Новосибирск. — 1999 (<http://www.ict.nsc.ru/ws/elpub99/40/>).
6. Баженов С. Р. Формирование корпоративной распределенной информационно-библиотечной системы г. Новосибирска // Научные библиотеки в новом тысячелетии: проблемы взаимного использования ресурсов. Новосибирск, Изд-во ГПНТБ СО РАН. — 2002. — С.109-114.