



# Девятая Международная Конференция «Крым 2002»

*Конференция проводится в рамках мероприятий ИФЛА 2002 г.*

***Библиотеки и ассоциации  
в меняющемся мире: новые технологии  
и новые формы сотрудничества***

***Тема 2002 года:***

***Электронные информационные  
ресурсы и социальная значимость  
библиотек будущего***

***Труды конференции.***

**Том 1**

**СУДАК**

**(основная программа)**

**Алушта, Бахчисарай, Белогорск, Феодосия, Керчь, Коктебель, Старый Крым**

**(выездные заседания)**

**Автономная Республика Крым, Украина**

**8–16 июня 2002**

Москва  
Издательство ГПНТБ России  
2002

УДК 02(063)  
ДКД 025.06047  
ББК 78я5  
Б 592

Б 592      **Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества. Тема 2002 года: Электронные информационные ресурсы и социальная значимость библиотек будущего: Тр. конф. / 9-я междунар. конф. «Крым 2002».** – М.: ГПНТБ России, 2002.  
Т. 1. – 520 с. – ISBN 5-85638-075-4.

Тексты докладов воспроизводятся  
с полным сохранением содержания,  
орфографии и синтаксиса,  
представленного авторами текста

**Опыт построения распределенных информационных систем  
на основе протокола Z39.50**

**Experience in Designing Distributed Information Systems  
on the basis of Z39.50 Protocol**

**Досвід побудови розподілених інформаційних систем  
на основі протоколу Z39.50**

*Мазов Н. А., Жижимов О. Л., Скибин С. В.*

*Объединенный институт геологии, геофизики и минералогии  
Сибирского Отделения РАН, Новосибирск, Россия*

*nikolai A. Mazov, Oleg L. Zhizhimov, Sergei V. Skibin  
Russian Academy of Science Siberian Branch Joint Institute of Geology,  
Geophysics and Mineralogy, Novosibirsk, Russia*

*Мазов Н. А., Жижимов О. Л., Скибин С. В.*

*Об'єднаний інститут геології, геофізики і мінералогії Сибірського відділення РАН, Новосибірськ, Росія*

Опыт построения распределенных информационно-библиотечных систем. В докладе рассмотрены современные информационно-библиотечные системы для хранения и обработки больших информационных массивов на основе различных СУБД, функционирующих в настоящее время в России, определены требования, предъявляемые к таким системам для построения распределенных информационно-библиотечных систем. Проведен анализ существующих решений в информационно-библиотечной практике России на основе различных СУБД

– решений с использованием Web-сервера совместно с CGI-приложением (Common Gateway Interface) и возможности решения «клиент-сервер» на основе использования протокола Z39.50 (ISO-23950). Вкратце описаны программные средства Z39.50 для обеспечения вышеперечисленных режимов при построении распределенных систем, разрабатываемые в Объединенном институте геологии, геофизики и минералогии СО РАН.

The experience of designing distributed library information systems is discussed. Modern information technologies to store and process vast information arrays on the basis of various DBMS popular in Russia are examined. The DBMS specifications essential for building distributed information library systems are determined. The solutions for various DBMS – in particular, the solution implicating the use of a Web-server together with the CGI-appendix (Common Gateway Interface) and potential of the «client-server» solution on the basis of the Z39.50 (ISO-23950) protocol – are examined. The software Z39.50 instruments to support the above mentioned modes being designed at the Joint Institute of Geology, Geophysics and Mineralogy of the Siberian Branch of the Russian Academy of Science are described in brief.

У доповіді розглянуто сучасні інформаційно-бібліотечні системи для збереження та обробки великих інформаційних масивів на основі різних СУБД, що функціонують на даний час у Росії; визначено вимоги, що пред'являються до побудови розподілених інформаційно-бібліотечних систем. Проведено аналіз існуючих рішень в інформаційно-бібліотечній практиці Росії на основі різних СУБД: з використанням Web-сервера спільно з CGI-додатком (Common Gateway Interface) і можливостей рішення «клієнт-сервер» на основі використання протоколу Z39.50 (ISO-23950). Коротко представлено програмні засоби Z39.50 для забезпечення перерахованих вище режимів при побудові розподілених систем, які розробляються в Об'єднаному інституті геології, геофізики і мінералогії СВ РАН.

Настоящий доклад посвящен изложению опыта построения распределенных информационно-библиотечных систем на основе протокола Z39.50. В докладе рассмотрены современные информационно-библиотечные системы для хранения и обработки больших информационных массивов на основе различных СУБД, функционирующих в настоящее время в России, определены требования, предъявляемые к таким системам для построения распределенных информационно-библиотечных систем. Проведен анализ существующих решений в информационно-библиотечной практике России на основе различных СУБД – решений с использованием Web-сервера совместно с CGI-приложением (Common Gateway Interface) и возможности решения «клиент-сервер» на основе использования протокола Z39.50 (ISO-23950). Вкратце описаны программные средства Z39.50 для обеспечения вышеперечисленных режимов при построении распределенных систем, разрабатываемые в Объединенном институте геологии, геофизики и минералогии СО РАН.

Ряд организаций, располагающих библиографическими базами данных, работает в настоящее время над тем, чтобы обеспечить унифицированный сетевой доступ к имеющимся ресурсам по стандартным протоколам для более широкого обращения к ним ученых, исследователей или других пользователей. Цель этих работ – построение единого информационного пространства и интеграция различных источников данных, т. е. создание распределенной информационной системы.

Однако, в гетерогенной информационной среде, где используются серверы различных производителей, плохо взаимодействующих друг с другом, например, Oracle, Informix, MS SQL Server, IBM DB2 и др., возникают проблемы формирования единого информационного пространства. Все перечисленные серверы работают на фирменных протоколах обмена данными, а эти протоколы, как правило, несовместимы между собой. Поэтому на клиентском компьютере необходимо держать клиентское программное обеспечение для каждого из этих серверов.

Проблемы несовместимости протоколов обмена данными становятся неразрешимыми, когда в качестве единого информационного пространства рассматривается вся сеть Интернет или ее большая часть. Заранее неизвестно, какие серверы могут предоставить данные, и какое клиентское программное обеспечение потребуется для доступа к ним.

Решение проблемы может быть лишь в унификации доступа к серверам любых производителей и базам данных. Именно для этой цели в мировом сообществе принят набор правил под общим названием Z39.50 [1, 2].

В последнее время в информационно-библиотечном сообществе России отмечается тенденция использования в качестве базового программного обеспечения небольшого числа отечественных АБИС целевого назначения для построения информационно-библиотечных систем.

Их репертуар весьма скромнен. По оценкам специалистов ГПНТБ России более 95% потребностей АБИС покрывают пять-шесть сравнительно недорогих российских разработок [3]. Такое положение дел обуславливается еще и тем, что при внедрении больших СУБД с целью создания информационно-библиотечных систем возникает ряд трудностей, а именно:

- имеющиеся в распоряжении у пользователей СУБД представляют собой достаточно сложные программные комплексы, для эксплуатации которых требуется наличие высококвалифицированных специалистов и значительных машинных ресурсов;

- пользователи, имеющие опыт создания и эксплуатации систем, разработанных в соответствии с принципами «табличной» организации данных, часто бывают морально неподготовлены к восприятию принципов организации и ведения библиографических баз данных.

Таким образом, сложность существующих коммерческих СУБД, с одной стороны, предполагает наличие достаточно мощного машинного парка, а с другой стороны, требует значительных усилий на их освоение и предъявляет повышенные требования к квалификации пользователей.

В целях обеспечения коллективного использования информационных ресурсов был проведен анализ существующих информационно-поисковых систем для обработки больших информационных массивов, который показал, что системы подобного рода должны обладать следующими основными компонентами:

- собственно данные;
- программное обеспечение работы с данными;
- среда передачи информации;
- интерфейс пользователя.

На основании предшествующего опыта эксплуатации локальных и сетевых информационных систем, а также исходя из основных концепций протокола Z39.50 было установлено, что распределенная информационная система должна отвечать следующим требованиям:

- Возможность работы с распределенными данными – информационная система должна допускать возможность работы с данными, расположенными на разных физических серверах, различных аппаратно-программных платформах и хранящихся в разнообразных внутренних форматах, в контексте одного клиентского сетевого соединения.
- Логическая группировка данных – система должна позволять обрабатывать все запросы на логических группах баз данных, полностью скрывая тем самым физическое расположение последних.
- Абстрактная модель данных – информационная система должна строиться на основе абстрактной схемы данных, на которую должны быть отображены конкретные базы данных. Это позволяет объединять данные из разнородных систем в одной логической группе.
- Абстрактная система запросов – система должна оперировать не конкретным синтаксисом запросов, а его логической сутью на основе абстрактных атрибутов.
- Метаинформация – система должна предоставлять полную информацию о себе и обо всех своих ресурсах.
- Разграничение доступа – система должна быть способна предоставлять различные уровни привилегий для пользователей по доступу к информации.
- Учет и контроль – система должна уметь собирать статистические данные по запросам пользователей и вести их бюджеты.
- Открытость – система должна допускать расширение и быть основана на открытых стандартах и протоколах.
- Связь с другими системами – возможность интегрировать свои ресурсы с ресурсами других информационных систем.
- Демократичность в общении – система должна предоставлять как простые и понятные для неподготовленного пользователя интерфейсы, так и профессиональные интерфейсы для доступа к информации.
- Связь с WWW – система должна иметь шлюз для доступа к ней из WWW.

Реализация этих требований в реальной информационной системе позволила бы максимально удовлетворить сегодняшние потребности сетевого информационного сервиса для доступа к распределенным базам данных.

Авторами настоящего доклада был построен прототип распределенной информационной системы, центральное место в которой, несомненно, занимает модульный сервер баз данных ZooPARK, поддерживающий протокол Z39.50, допускающий работу с данными различных СУБД и удовлетворяющий следующим требованиям:

- поддержка протокола Z39.50-1995 (v. 3);
- работа с различными СУБД;
- переносимость на различные аппаратные платформы.

Немаловажную роль при этом играло желание максимально использовать существующие открытые разработки и библиотеки поддержки Z39.50, распространяемые другими разработчиками в исходных текстах. Это позволяет применить накопленный мировым сообществом опыт и избежать многих ошибок в реализации базовых функций Z39.50.

Включение в распределенную систему данных из новой СУБД для описываемого сервера ZooPARK сводится к разработке специального модуля с фиксированным программным интерфейсом (провайдер данных).

Следует отметить, что функциональные характеристики сервера существенно зависят от возможностей провайдеров данных. Все провайдеры данных представляют собой динамически загружаемые библиотеки.

«Распределенность» информационной системы достигается двумя способами:

1. Включение в комплектацию каждого сервера Z39.50 специального провайдера данных (Z-REMOTE для ZooPARK), способного открывать сеансы Z-связи с другими серверами, перенаправлять к ним запросы, получать ответы и возвращать последние ядру сервера.
2. Установка серверного программного обеспечения промежуточного (между клиентом и серверами Z39.50) слоя, способного поддерживать множественные сеансы Z-связи. В качестве такого программного обеспечения используется шлюз HTTP-Z39.50, построенный на основе Z-CGI (разработка ОИГГМ СО РАН).

В настоящее время в СО РАН создана распределенная информационная система, построенная на основе вышеописанного сервера Z39. 50, для обеспечения доступа научных сотрудников к библиографической информации (в первую очередь аналитического характера – электронные варианты баз данных ВИНТИ, каталогов библиотек и др.). Система состоит из ряда серверов Z39.50 ZooPARK, установленных в различных организациях и обеспечивающих доступ к базам данных под управлением CDS/ISIS (библиографическая аналитическая информация), к реляционным базам данных MS SQL, MS ADO, MySQL (справочная информация, тезаурусы и классификационные схемы) и текстовым данным в формате SGML. Распределенность системы реализована двумя указанными выше способами. Информационная система также интегрирована в WWW через шлюз Z39. 50-HTTP, функционирующий в среде серверов IIS для Windows или Apache для UNIX.

Следует отметить, что опыт эксплуатации серверов ZooPARK в составе информационной системы ОИГГМ СО РАН, распределенной информационной системы СО РАН, корпоративных распределенных библиотечных системах Новосибирска, Москвы, Омска и ряда других городов России показал его жизнеспособность и эффективность на больших информационных массивах библиографической информации. При этом особо следует отметить возможность перенаправления запросов (работа соответствующего провайдера Z39. 50), логическое объединение пространственно разнесенных и разнородных баз данных (например, данных CDS/ISIS и MS ADO), время реакции системы. Примечательно, что на сегодняшний день в России именно этот сервер успешно применяется при построении распределенных информационно-библиотечных систем и занимает по различным оценкам от 60 % до 70 % серверного программного обеспечения Z39. 50.

В ходе промышленной эксплуатации распределенной информационной системы были дополнительно построены и протестированы специальные средства, расширяющие функциональные возможности эксплуатируемой системы. В частности, был организован доступ к тезаурусам и различным классификаторам по схеме Zthes, к музейным коллекциям по схеме CIMI. При этом для однозначной идентификации словарных статей тезаурусов был применены оригинальные алгоритмы кодирования. В настоящее время ведутся работы по расширению возможностей сервера в части поддержки дополнительных функций Z39. 50.

Таким образом, описанная выше технология построения распределенной информационной системы на основе протокола Z39.50 с применением сервера ZooPARK позволяет снять ряд ограничений, присущих оригинальным СУБД и информационным системам, а также позволяет обеспечить расширение функциональных характеристик базовых СУБД, в частности:

- возможность работы с базами данных в режиме «клиент-сервер»;
- возможность реализации мультибазовой концепции работы с данными;
- возможность доступа к данным по стандартному открытому протоколу;
- возможность интеграции баз данных с другими базами данных в гетерогенной информационной системе.

## Литература

1. ANSI/NISO Z39.50-1995. Information Retrieval (Z39.50): Application Service Definition and Protocol Specification. Z39.50 Maintenance Agency Official Text for Z39.50-1995, July 1995.
2. Жижимов О.Л. Введение в Z39.50. – Новосибирск: Изд-во НГОНБ, 3-е изд., доп. и перераб. 2002. – 253 с.
3. Шрайберг Я.Л., Воройский Ф.С. Автоматизированные библиотечно-информационные системы. – М.: Либерия: ГПНТБ России, 1996. – 273 с.