

КОНСОЛИДАЦИЯ ЭЛЕКТРОННЫХ БИБЛИОТЕЧНЫХ И ИНТЕРНЕТ РЕСУРСОВ ДЛЯ ОБРАЗОВАТЕЛЬНЫХ И НАУЧНЫХ ЦЕЛЕЙ НА ОСНОВЕ GRID-ТЕХНОЛОГИИ

Б.В. Олейников, А.И. Шалабай

*ФГАОУ ВПО «Сибирский Федеральный Университет», г. Красноярск, Россия
oleynik48@mail.ru, andrsh@gmail.com*

Одним из критериев доступности образования и проведения качественных научных исследований является возможность получения необходимой учебной и научной литературы. Если еще несколько лет назад основным поставщиком литературы являлись библиотеки, то сегодня в силу определенных причин наблюдается стремительно нарастающая тенденция размещения требуемой для обучения и проведения научных исследований литературы в сети Интернет. Давая очевидные преимущества, связанные с доступом ко все увеличивающемуся объему литературы, данная тенденция в тоже время порождает и множество проблем, главные из которых связаны со специфическим сопровождением размещаемых в сети ресурсов, мало способствующим их консолидации с целью повышения эффективности поиска.

Одним из путей решения указанных проблем является консолидация электронной литературы, находящейся как в классических библиотеках, так и в сети Интернет, на библиотечных принципах. Основой является наличие библиографического описания любого ресурса (хранящегося в библиотеке или найденного в Интернет), в соответствии с общепринятым форматом семейства MARC, используемым в современных автоматизированных библиотечно-информационных системах (АБИС). При этом описание ресурса, найденного в Интернет, большей частью должно осуществляться автоматически (включая и определение тематики до определенного уровня библиотечной классификации).

В силу огромного количества электронных документов, находящихся в сети Интернет, ни одна отдельная организация не способна описать все ресурсы. Решением проблемы может стать использование Grid-технологий для создания распределенной научно-образовательной библиотечной системы. Её концепция строится на основе следующих положений.

Полные тексты документов хранятся в узлах Grid-сети, являющихся территориально распределенными клиентскими рабочими местами, имеющими выход в Интернет. На узлах обеспечивается порождение требуемых библиографических описаний, индексирование документов, а также их резервное долговременное хранение в виде полных текстов.

Полученные библиографические описания, а также список ссылок на полные тексты документов направляются в централизованное хранилище, позволяющее осуществлять поиск без последовательного опроса узлов, каталогизировать поступающую информацию, а также обмениваться библиографической информацией с существующими АБИС.

Внешний пользователь (получатель информации) после осуществления поиска в хранилище, выбирает определенные документы для скачивания, которые предоставляются ему с ближайшего доступного в данный момент узла Grid-сети. При доступности источников полнотекстовых данных скачивание происходит по технологии р2р, в противном случае специальная серверная служба постоянно

отслеживает их подключение, самостоятельно скачивает и отправляет получателю требуемые данные.

Узлам сети присваиваются определенные полномочия, например, добавление новой информации определенных тематик, её модификация, возможность блокировки определенных узлов с изъятием полученных от них данных и т.д. Также целесообразно для каждого узла разделять полномочия для документов разной области знаний.

С внедрением библиотечной Grid-системы и объединением баз данных описательной информации с классическими библиотеками появится возможность одновременного поиска требуемых документов в Интернет, в классических и электронных библиотеках (реализация принципа «одного окна»). Авторами получены определенные результаты, связанные с реализацией предложенного подхода.

CONSOLIDATION OF ELECTRONIC LIBRARY AND INTERNET RESOURCES FOR EDUCATIONAL AND SCIENTIFIC PURPOSES BASED ON GRID-TECHNOLOGY

B.V. Oleynikov, A.I. Shalabay

*FSAEI HPE Siberian Federal University, Krasnoyarsk, Russia
oleynik48@mail.ru, andrsh@gmail.com*

One of the criteria of access to education and scientific researches is the possibility of obtaining the necessary educational and scientific literature. If a few years ago the main supplier of literature were libraries, then today in view of the known causes there is a rapidly growing trend of placing the required training and research literature in the Internet. Giving the obvious advantages associated with access to ever-increasing volume of the literature, the tendency at the same time initiates many problems, the main of them are connected with a certain support of the resource which is placed in a network. It does not assist consolidation of resources and, as a consequence, to their effective search.

One of solutions of these problems is the consolidation of electronic literature, located both in the classical libraries and the Internet on library principles. The basis of this approach is the presence of bibliographic description of any resource (stored in the library or found on the Internet), in accordance with a generally accepted format of a family of MARC, used in the modern automated library information system (ALIS). A description of the resource found on the Internet for the most part should be done automatically (including identification of subjects to a certain level of library classification).

Considering the huge number of electronic documents in the Internet, no single organization can describe all of the resources. The solution is to use Grid-technologies for creation of distributed scientific and educational library system. The concept of such a system is based on the following assumptions.

The full texts of documents are stored in the nodes of the Grid-network, which are client workstations with Internet access and geographically distributed. Bibliographic descriptions, indexing resources and their long-term backup storage in the form of the full text are provided at the nodes.

The resulting bibliographic descriptions and a list of links to full texts of documents are sent to a centralized repository that allows one to search a desired document without serial poll sites, to catalog incoming information and exchange bibliographic information with existing ALIS.

The external user (recipient of the information) after the search in the repository selects certain documents for download. These documents become available to him from the nearest and available Grid-node. If a Grid-node is available, then at downloading p2p technology is used. Otherwise, a special service server continuously monitors connectivity of Grid-nodes, independently downloads and sends the requested data to the recipient.

The Grid-nodes are supplied with a certain powers, for example, addition of new information of certain subjects, its modification, a possibility of blocking certain nodes with withdrawal of the data received from them, etc. Also it is advisable for each node to share its responsibility for the processing of documents relating to various areas of knowledge.

The introduction of the Grid-library system and the uniting of a descriptive information databases with the classic libraries provide a way for simultaneous search for the required documents in the Internet, in the classical and digital libraries (the principle of "one window"). The authors have obtained some results relating to the implementation of the proposed approach.