

Концептуальные подходы к построению распределенных систем*

А.М. Федотов[†]

Аннотация. Доклад посвящен обсуждению проблем, связанных с созданием и поддержкой распределенных информационно-вычислительных ресурсов в Сибирском отделении РАН. Излагаются основные технологические подходы, структура метаданных и использование GRID-технологий.

Введение

В настоящее время наиболее развитые страны мира находятся на завершающей стадии индустриального этапа развития общества и перехода к следующему – информационному этапу развития общества. Широкое использование современных средств доступа к информационно-вычислительным ресурсам открыло принципиально иные возможности использования информационных технологий и вычислительной техники. В условиях современного динамического развития общества, усложнения технической и социальной инфраструктуры информация становится таким же стратегическим ресурсом, как традиционные материальные и энергетические ресурсы. Современные информационные технологии, позволяющие создавать, хранить, перерабатывать и обеспечивать эффективные способы представления информационных ресурсов потребителю, стали важным фактором жизни общества и средством повышения эффективности управления всеми сферами общественной деятельности. Уровень использования информации становится одним из существенных факторов успешного экономического развития и конкурентоспособности региона как на внутреннем, так и на внешнем рынке. Информация является важнейшим стратегическим ресурсом и наибольший экономический и социальный успех сегодня сопутствует тем странам, которые активно используют современные средства компьютерных коммуникаций и сетей, информационных технологий и систем управления информационными ресурсами. Перенесенные на электронные носители информационные ресурсы приобретают качественно новое состояние и становятся активными. Доступная для оперативного воспроизводства средствами компьютерной обработки информация является важнейшим фактором социального развития общества.

Осознание мировым сообществом роли информации как стратегического ресурса стимулировало разработки новых технологий как для получения и переработки больших объемов информации, так и для ее хранения и предоставления пользователям. Стремительное развитие глобальных информаци-

*Работа выполнена при частичной поддержке РФФИ, проекты 03-07-90352, 03-07-90423, и интеграционного проекта СО РАН № 132.

[†]Институт вычислительных технологий СО РАН, Новосибирск.

онных и вычислительных сетей ведет к изменению фундаментальных парадигм обработки данных, которые можно охарактеризовать как переход к распределенным ресурсам и создание инфраструктуры для свободного доступа к ним. На первое место в современном информационном обществе выходят технологии использования распределенных информационно-вычислительных ресурсов [1]. В западной литературе несколько лет назад появился даже новый термин “GRID-технологии” – технологии создания и использования распределенных информационно-вычислительных ресурсов.

На самом деле идея использования распределенных ресурсов далеко не нова. Она была сформулирована еще на заре компьютерной эры – здесь можно вспомнить Вэннивера Буша¹, который в своей книге “As We May Think” (1945) описал концепцию гипертекста. В дальнейшем идея распределенной обработки информации стала активно развиваться с появлением первых “супер” ЭВМ второго поколения – это советская (российская) ЭВМ БЭСМ-6 и американская ILLIAC-IV. Именно на этих машинах впервые были поставлены эксперименты и создано соответствующее программное обеспечение по использованию распределенных вычислительных ресурсов. В дальнейшем технология массивованного счета была перенесена и на “мэйнфреймы” третьего поколения. Недаром одним из лозунгов создателей технологий GRID является следующий: “Вперед (назад) к мэйнфреймам”.

Второй этап развития технологий использования распределенных ресурсов можно связать с появлением сети Интернет. В 1989 году в CERN’e перед сотрудниками лаборатории ECP (Electronics & Computing for Physics) была поставлена задача по разработки системы, позволяющей обеспечить доступ к любым сведениям, содержащимся в сети, и объединения всех, имеющиеся в ней источники информации, для сотрудников, приезжающих из различных уголков мира – создать универсальную технологию доступа к распределенной, разнородной информации. Таким образом, родилась технология “World Wide Web” (“Всемирная паутина”). Создаваемая система должна не только поддерживать произвольные гипертекстовые ссылки, но предоставлять поиск по распределенным в сети базам данным, но что самое главное быть многопользовательской, платформенно-независимой системой. Дальнейшее развитие технологии WWW немного подкорректировало начальные установки и из системы доступа к разнородным ресурсам он превратилась во всемирную “презентационную” систему.

Развитие концепций предоставления удобного и разветвленного поиска ресурсов, распределенных по сети Интернет, привело к созданию протокола WAIS, замыкающего триаду технологических средств, которую возглавляют Gopher и WWW. Данный протокол обеспечивает функционально полный набор информационных инструментов для Internet. Если Gopher (или сейчас более модный его собрат LDAP) – это виртуальная файловая система, где в качестве элементов системы могут использоваться не только файлы различных форматов и каталоги, но и виртуальные объекты в виде поисковых критериев, а WWW – распределенная мультимедийная система, то WAIS реализует кон-

¹Вэннивер Буш (Vannevar Bush) (1890–1974) – основатель Национального Научного Фонда (NSF) США, создатель “дифференциального анализатора” (1930) – релейной машины, способной решать дифференциальные уравнения, руководитель работ по созданию первых ЭВМ.

цепцию распределенной информационно-поисковой системы и распределенную систему доступа к ресурсам. Система базируется на универсальном протоколе WAIS, в основе которого лежит американский стандарт Z39.50 [2].

Стандарт Z39.50 – это один из протоколов семейства OSI², который описывает прикладной уровень взаимодействия распределенных информационно-поисковых систем (ИПС)³. Протокол определяет механизм информационного обмена в процессе обработки поисковых запросов и протокол обмена данными в системах, осуществляющих поиск. Основная область применения протокола в настоящее время – это библиотечные системы и системы научно-технической информации. Однако область применения протокола значительно шире перечисленных приложений – он может использоваться в информационно-поисковых системах общего назначения [3]. При разработке протокола подразумевалось, что он будет описывать порядок обмена информацией между пользователями информационной системы и ее ядром через сеть передачи данных. При этом сами системы могут управлять данными, используя разные модели и различные языки манипулирования этими данными. Таким образом, информационно-поисковую систему можно построить на основе любой системы управления данными, будь это обычная файловая система или объектно-ориентированная СУБД.

Для разработки распределенных информационных систем сегодня используются различные технологии (RPC, DCOM, RMI, ODBC, JDBC, CORBA, SOAP, eXML, WSDL, WSFL, UDDI и др.). Большая часть этих технологий по своей сути ориентирована не на работу с информационными ресурсами, а на сетевое взаимодействие программ и распределенные вычисления в гетерогенных средах. Однако ни одна из упомянутых технологий не обеспечивает универсальных способов работы с информационными ресурсами, т.е. необходимого уровня абстрагирования от конкретных систем и платформ при сохранении высокой степени функциональности и предоставления механизма доступа к данным. Эти требования присущи технологии, в основе которой лежит протокол ANSI Z39.50 [2] (стандарт ISO-23950).

В основе протоколов семейства OSI находится поддержка метаданных – формализованных знаний о внутренней структуре и поведении информационных и вычислительных ресурсов (или документов): структурированные сведения о ресурсе, представляющие его свойства (атрибуты) и функции – информации, предназначенной для анализа, дизайна, развития, обработки и использования Документов⁴. Понятие метаданных уже давно и успешно используют

²OSI – Open System Interconnection.

³В отечественной литературе за термином “Information Retrieval System” закреплено понятие “Информационно-поисковая система”. Стандарт Z39.50 посвящен организации механизма поиска в компьютерных сетях распределенных информационных ресурсов с использованием ключевых слов.

⁴Под понятием Документ в контексте данного доклада мы понимаем следующее: В информационном пространстве все информационные ресурсы: события, факты и любые другие сущности реального или виртуального мира существуют только в форме некоторых информационных объектов. В этом смысле информационный ресурс снабженный метаописанием превращается в Документ, который является основным “объектом”, с которым оперирует любая информационная система. Таким образом, Документ – это структурированное описание реальной сущности (объекта, факта или понятия, включая алгоритмы и программы), которые составляют информационное наполнение системы.

в таких контекстах, как информационные хранилища и системы аналитической обработки данных, электронный документооборот и управление потоками работ, управление знаниями. Этот стандарт позволяет описывать классы объектов, представляющие типы информационных документов, атрибуты и отношения между различными типами документов и ресурсов.

В моделях распределенных вычислений метаданные являются основным структурным и функциональным описанием ресурса, использование которого и обеспечивает унифицированный доступ. В рамках стандартов OSI уже разработан целый набор схем метаданных для описания Документов, о которых мы расскажем в следующем разделе.

Основная задача, поставленная разработчиками GRID – это работа по объединению всех ресурсов сети Интернет в единую интегрированную среду распределенных ресурсов (GRID, ИРИС⁵), которая составит информационно-вычислительную инфраструктуру будущего [4]. GRID можно определить как исходно распределенную систему, которая сводит воедино данные, вычислительные мощности и ресурсы для представления данных. Единый интерфейс должен предоставлять доступ ко всем необходимым ресурсам так, словно мы имеем дело с одним огромным “метакомпьютером”. Все задачи, как традиционные для обычных компьютеров (управление процессами, памятью, файловой системой, вводом/выводом и пр.), так и принципиально новые/старые (учет, контроль, способ доступа и распределение ресурсов, обеспечение безопасности, совместная работа над набором данных в реальном масштабе времени и пр.) должен решать специализированный комплекс программного обеспечения на базе соответствующей аппаратной инфраструктуры.

Разработчики технологии GRID (термин не является аббревиатурой, это английское слово “сеть, решетка”⁶) составляют отдельное сообщество с собственной терминологией⁷. Уже из названия можно увидеть, что они делают заявку на “второй Интернет” – подход концепции GRID претендует на глобальность.

1. ИРИС

Создание распределенной информационно-вычислительной сети опирается на идею электронных (цифровых) библиотек. Основная задача – это формирование в ресурсах “Сети” единого, математически однородного поля компьютерной информации, способного стать универсальным и машинезависимым носителем данных, программ и глобально распределенных вычислительных ресурсов. В рамках этого подхода цифровые библиотеки рассматриваются как отдельная конкретная технология работы с информацией.

Помимо задач модификации первичных ресурсов организаций и институтов в направлении их частичного открытия внешним пользователям, стоит про-

⁵ИРИС – Интегрированная Распределенная Информационная Система Сибирского отделения РАН, см. [3].

⁶Правда некоторые авторы расшифровывают этот термин следующим образом: “Global Resource Internet Distributed”.

⁷Это составляет некоторые трудности для сопряжения с концепцией OSI, работающий на открытых международных стандартах.

блема разработки принципов организации ресурсов внутреннего пользования и самих метаданных для их эффективного использования⁸.

Создание единой технологической системы хранения, поиска и использования информации является одной из важнейших задач интеграции научных исследований, проводимых различными группами исследователей. Стремительное развитие глобальных информационных и вычислительных сетей ведет к изменению фундаментальных парадигм обработки данных, которые можно охарактеризовать как переход к распределенным ресурсам и создание инфраструктуры для свободного доступа к ним. Системы поиска информации эволюционируют от синтаксического к семантическому поиску, технологии обработки информации совершенствуются в направлении все более изощренных и эффективных алгоритмов анализа всего спектра данных.

В последние годы в СО РАН начаты работы по созданию Интегрированной Распределенной Информационной Системы (ИРИС) [5]. При создании единой системы доступа к ресурсам должны быть решены следующие принципиальные задачи:

- обеспечение интероперабельности в глобальной программно-аппаратной инфраструктуре;
- диспетчеризация, включая идентификацию доступных ресурсов, статистика использования и загрузки ресурсов и пр.;
- система безопасности и контроля доступа, в т.ч. гибкое регулирование объема прав и привилегий пользователей;
- обращение к наборам данных в удаленных архивах (включая протоколы, которые необходимо использовать для работы с гетерогенными источниками данных, и библиотеки программных комплексов) и др.

Организация данных в ИРИС основана на взаимодействии соответствующих служб:

- публикации данных, поддержка и их аутентичности и качества;
- поиска и представления информации;
- анализа распределенных данных.

Все эти механизмы должны составлять основу системы “усвоения данных (документов)” – системы превращения информации в систему библиотек, оперирующих с “документами”. Сама по себе информация, хранящаяся в репозиториях (архивах) является только лишь набором битов, комбинацией данных и метаданных, выполненных с использованием адекватного языка описания или разметки. Как именно конкретный пользователь (приложение) будет использовать эту информацию, определяется пользователем в соответствии с метаописанием. Создаваемые технологии должны предоставлять возможности для точного и адекватного удовлетворения потребностей пользователей, формально обращающихся к одной и той же информации (документу из архива).

⁸По оценке многих экспертов, уже сейчас Интернет представляет собой мировую “свалку” информации.

ИРИС СО РАН можно рассматривать как пример виртуальной организации унифицирующей механизмы работы с документами – большая часть компонентов системы может находиться в разных местах, при выполнении прикладных задач эти компоненты могут функционировать независимо, а интероперабельность и унифицированность достигается использованием согласованного набора протоколов и служб, действующих на основе метаинформации. Единство достигается, прежде всего, за счет унифицированного интерфейса и централизованного администрирования.

Любая электронная библиотека (корпоративная или общедоступная) опирается на метаинформацию, содержащую онтологию, описывающую принципы организации информации. Онтология конкретизируется в схеме данных (атрибутов), описывающих информацию в конкретных метаданных.

2. Метаданные

Согласно определению, данному Meta Data Coalition (<http://www.mdcinfo.com>) в документе “Open Information Model”: “метаданные – это описательная информация о структуре и смысле данных, а также приложений и процессов, которые манипулируют данными”. Согласно нашим определениям: в распределенной системе обработки информации все данные представлены в виде Документов.

Под Документом понимается целостный информационный объект (как это принято в ООП – объектно-ориентированном программировании⁹) фиксированного Класса, помещенный в информационное пространство, который описывает, представляет, отображает или моделирует некоторую сущность реального (или субъективного) мира.

Принадлежность документа к некоторому Классу определяет смысловую структуру документа, атрибуты и функции, а так же методы доступа к нему. В Классе задается функциональность: множество Документов, принадлежащих одному Классу, выполняют одинаковые функции. Таким образом, при введении в рассмотрение документов в их описании наряду с объектами появляются функции. Исходя из объектной модели представления информации, в основе которой лежат “метаданные”. “МЕТАДААННЫЕ – это описательная информация о структуре и смысле данных, а также приложений и процессов, которые оперируют данными и программами”. На основе метаданных осуществляется поиск документов (ресурсов), вывод результатов поиска, управление ресурсами, взаимодействие с ними. Формальное определение смысловой структуры Документа дается метаописанием Класса документов, в котором каждый тип документов представляется в виде набора объектов со своими характеристиками и атрибутами.

⁹На сегодняшний день в программировании не существует точной формулировки понятия объект! Точнее сказать, их столько много, что нет никакой возможности выбрать правильную формулировку. Диапазон широк, от общепризнанной “Все есть объект” (значит и стул, на котором сидит программист, тоже объект программирования), до слишком ограничивающих формулировок, например, в ООП объект наделяют обязательным свойством наследования. Некоторые формулировки просто устарели, хотя находятся в постоянном употреблении, например, утверждение о том, что объект должен находиться в оперативной памяти компьютера (тогда как быть с объектами, расположенными в БД?).

В информационном пространстве события, сущности, факты существуют в виде информационных ресурсов. Под *информационным ресурсом* понимается целостный информационный объект, имеющий имя. Ресурсом может быть может быть научная статья, визитная карточка, файл с изображением и т. п. С функциональной точки зрения метаданные разбиваются на два типа:

- *Структурные* метаданные – определяют свойства класса ресурсов, в соответствии с которым осуществляется их обработка (типы, связи, форматы представления, ограничения на управление доступом и т. п.).
- *Описательные* метаданные – описывают содержательную часть ресурса или некоторую реальную сущность. Описательные метаданные, характеризующие ресурс, могут являться частью ресурса и, в то же время, могут содержать в соответствии с выбранной *схемой данных* сведения о ресурсе (авторы, название, дата создания и т. д.).

В области стандартов хранения и обмена метаданными единого комплекса взаимодополняющих стандартов к сожалению пока не существует. На сегодня пока существует целый ряд разрозненных стандартов:

- CDIF (CASE Data Interchange Format);
- IRDS (Information Resource Dictionary System);
- AD/Cycle CASE Platform;
- MOF (Meta Object Facility) – определяет набор интерфейсов CORBA и объектов для взаимодействия с метаданными (разработанная Object Management Group);
- XMI (XML Metadata Interchange) – формат транспорта метаданных. Существует также XMI-схема для представления MOF-моделей;
- серия стандартов, направленных на построение инфраструктуры взаимодействия приложений через Internet: SOAP, eXML, WSDL, WSFL, UDDI;
- международный стандарт ISO-23950 (Z39.50).

В основу манипулирования метаданными, определяющими функциональные связи между документами, положена схема RDF¹⁰, которая определяет правила представления структур данных.

В модели RDF документ рассматривается как частично-упорядоченный набор абстрактных объектов (элементов), обладающие свойствами (атрибутами) и имеющими уникальный идентификатор. Любой объект при своем создании получает генерируемый системой уникальный идентификатор, который связан с объектом во все время его существования и не меняется при изменении состояния объекта. RDF позволяет определять произвольные объекты в документе. Атрибуты (имена и значения) должны выбираться из словарей, связанных с теми или иными предметными областями. Формально RDF не накладывает никаких ограничений на значения атрибутов объектов, перекладывая создание соответствующих словарей на заинтересованные организации. Основной

¹⁰RDF – Resource Description Framework, RDF-Schema.

словарь имен объектов системы создан на основе словарей стандартных схем метаданных.

3. GRID

Использование распределенных информационно-вычислительных ресурсов становятся магистральным направлением развития современной компьютерной индустрии. На смену отдельно стоящим, независимым компьютерам и суперкомпьютерам должны прийти группы высокопроизводительных серверов, объединенных либо в кластеры, либо в виртуальные системы управления вычислительными ресурсами. Развитие сетевых технологий в настоящий момент сделало возможным объединение распределенных по сети компьютеров в мощный территориально распределенный “суперкомпьютер”.

Часто вычисления с использованием больших объемов данных выделяются в отдельную категорию. К общим характеристикам потребностей, которые делают оправданной организацию вычислительных архитектур типа GRID, можно отнести следующие:

- большие объемы данных, распределенных по различным научным центрам, странам и континентам;
- участие большого количества специалистов в обработке данных из разных институтов и университетов;
- информация, которую следует проанализировать, имеет сложную структуру;
- алгоритмы обработки информации имеют нетривиальный характер (объем программ составляет миллионы строк текста);
- наконец, масштабируемость базового программного обеспечения (фактически, всего того, что лежит ниже прикладного уровня), которое должно устойчиво работать как на настольной машине, так и на суперкомпьютере.

К настоящему времени уже имеются практические примеры реализации различных систем типа GRID. Следует также обратить внимание на то, что мощные вычислительные установки не могут функционировать сами по себе без квалифицированного персонала, имеющего практический опыт организации крупномасштабных вычислений. Ведь при работе на удаленном вычислителе потребитель пользуется также результатами труда персонала, который сопровождает данный вычислитель. Зачастую это обстоятельство оказывается таким же важным, как и наличие доступа к мощным вычислительным ресурсам – подготовка новых специалистов и создание продуктивного коллектива для нового суперкомпьютерного центра обойдется дороже любых суперкомпьютеров. Таким образом, создание средств для объединения вычислительных мощностей одновременно ведет и к появлению средств для объединения, через национальные границы и административные барьеры, усилий людей.

В рамках создания ИРИС в СО РАН разрабатывается проект распределенной вычислительной среды Meta-DataGrid.

Цели проекта:

- Интеграция информационных и вычислительных ресурсов и создание распределенной системы доступа к ним в рамках (система DataGrid).
- Разработка уникальной системы доступа к международным банкам данных, моделей и программ анализа данных (в том числе и в области “биологической информатики”).

Основное направление работ – “система распределенного управления ресурсами”, предназначенная для консолидации распределенных ресурсов при выполнении какой-либо задачи путем формирования метакомпьютерной среды из временно простаивающих рабочих станций, суперкомпьютеров, серверов и ПК.

Для реализации системы метавычислений решаются следующие задачи:

- Оптимальное распределение частей работы по вычислительным системам различной архитектуры и различной мощности.
- Однородный доступ к вычислительным ресурсам большого числа компьютеров в локальной или глобальной сети.
- Управление структурой МетаВычислителя, который может не иметь постоянной конфигурации – отдельные компоненты могут включаться в его конфигурацию или отключаться от нее; при этом технологии метавычислений должна обеспечить непрерывное функционирование системы в целом.

Реализация проекта обеспечит:

1. Интероперабельность БД (информационных хранилищ) и стандартизацию хранилища данных – интеграция БД.
2. Извлечение данных – построение знаний – интеллектуальные агенты.
3. Доступ к библиотекам алгоритмов для обработки данных и организация систем моделирования.
4. Настройка данных на алгоритмы.
5. Распределение задач по различным порталам (распределенная обработка данных).
6. Мультипроцессорную обработку ресурсов и данных.

Реализация распределенной информационной системы позволит перейти к построению интеллектуальной системы обработки запросов, например в идеологии популярной сегодня технологии GRID, главную роль в которой играет программное обеспечение, реализующее функцию управления моделями данных и метаданных (диспетчера). Полученный от приложения запрос направляется в систему обработки, которая посредством системы поиска информации разыскивает необходимые данные и после выполнения удаленных процедур (например, выделения требуемого подмножества из данного множества) возвращает затребованные данные приложению. Разработанная динамическая система формирования электронных коллекций предоставляет возможности для

точного и адекватного удовлетворения потребностей пользователей, формально обращающихся к одной и той же информации (цифровому объекту в репозитории). Эта методология была использована при создании “Электронного атласа биоразнообразия животного и растительного мира Сибири” [6].

Система обмена метаданными основывается на сервере метаданных СО РАН, который поддерживает следующий набор служб:

- публикация/регистрация новых наборов данных;
- база метаданных для поиска данных по атрибутам;
- доступ к гетерогенным ресурсам посредством брокера ресурсов;
- контроль аутентификации и доступа;
- мониторинг информационных ресурсов и ресурсов ввода/вывода;
- распределенное исполнение служб.

Соответствующие службы могут регистрироваться в объектно-ориентированной базе данных и вызываться для обработки любого набора данных, хранящегося в библиотеке (репозитории). Комбинация возможности доступа к данным через базу метаданных, служб каталогов, зарегистрированного набора методов обработки данных позволяет решать задачи для создания среды обработки данных. Главную роль играет программное обеспечение, реализующее функцию управления моделью данных (диспетчера). Полученный от приложения запрос направляется в систему обработки, которая посредством систем поиска разыскивает необходимые данные и после выполнения удаленных процедур (например, выделение требуемого подмножества из данного множества) возвращает затребованные данные приложению. Основная задача проекта состоит в разработке и использовании программных компонентов для получения доступа к различным ресурсам компьютерной сети с целью организации крупномасштабной обработки данных.

Одним из основных результатов работ по проекту является разработка промежуточного ПО, экранирующего сложность различных распределенных гетерогенных ресурсов при сохранении их возможностей и прозрачном позиционировании трех подсистем:

1. **Именованное.** Поскольку уникальные имена практически невозможны, используются метаданные. Система поиска информации, выполняя запросы по этим данным, осуществляет поиск в каталоге метаданных.
2. **Нахождение ресурса.** Описание расположения хранится в каталоге метаданных. Становится возможным разнесенное хранение данных и метаданных, а в каталоге должны храниться и данные о протоколе для доступа к конкретному набору.
3. **Конверсия протоколов,** которая должна осуществлять преобразование протоколов с целью обеспечения интероперабельности.

4. Реализация ИРИС

Реализация ИРИС идет в рамках работ по созданию электронной библиотеки Сибирского отделения РАН [7]. В основе реализации ИРИС лежит мета-

модель, исходящая из того, что документ (информационный ресурс: документ, алгоритм, программа, файл или вычислительный ресурс) характеризуется набором присущих ему атрибутов и методов, характеризующих связи с другими документами. Информация о документах системы, их атрибутах и методах поддерживается сервером метаданных, содержащий метаописания системы и метаописания отдельных коллекций документов. Сервер метаданных является отдельной частью системы, содержащей описание информационной модели предметной области, параметров настройки стандартных функций системы. По информации сервера метаданных осуществляется динамическая генерация схем базы данных системы и ведение программного служебных баз данных, в которых хранятся данные, обеспечивающие поддержку стандартных функций системы, динамически определяемые отношения между документами и динамическое распределение ресурсов системы.

В настоящее время в ИРИС выделяются следующие типы ресурсов – ДОКУМЕНТОВ:

- каталоги информационных ресурсов, информация об информационных ресурсах и метаинформация;
- электронные коллекции;
- классификационные ресурсы;
- вычислительные ресурсы и ресурсы обработки данных;
- хранилища программного обеспечения;
- административные информационные ресурсы, включающие полную информацию об актуальном состоянии системы и ее отдельных компонент.

В основу создания ИРИС положен принцип информационных хранилищ, с учетом поддержки уже функционирующих технологий, технологических решений используют принципы обработки метаописаний данных на основе серверов метаданных и протокола Z39.50. Разработана технология, которая обеспечивает виртуальную интеграцию разнородных информационных и вычислительных ресурсов, расположенных на серверах различных организаций, в единую систему с унифицированным доступом на основе открытых международных стандартов. Это позволило объединить в единое информационное пространство уже существующие многочисленные сервера организаций, входящих в научно-образовательную сеть, основанные на стандартных протоколах, организовать информационное обеспечение проведения исследований по фундаментальным и прикладным направлениям, проводимым в вузах, академических и институтах, ГНЦ «Вектор». Обеспечить доступ к фактографическим базам данных и уникальным коллекциям, накопленным в научных подразделениях СО РАН, с помощью современных средств доступа к данным. Таким образом, данные доступные в обычном режиме узкому кругу сотрудников той или иной организации станут доступными в пределах всего научно-образовательного сообщества Новосибирска.

Пользователю система (на первом этапе) предоставляет:

- Каталоги организаций и сотрудников.
- Интегрированный каталог основных разработок, проектов, публикаций и изданий, включая электронные.
- Каталоги документов и ресурсов, касающиеся всего научно-образовательного пространства.
- Каталоги предоставляемых программных, алгоритмических и вычислительных ресурсов.
- Доступ к электронной библиотеке – хранилищу ресурсов.

Список литературы

- [1] Шокин Ю.И., Федотов А.М. Распределенные информационные системы // Вычислительные технологии. – 1998. – Т. 3, № 5.
- [2] 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.
- [3] Шокин Ю.И., Федотов А.М., Жижимов О.Л., Мазов Н.А. Интегрированная распределенная информационная система (ИРИС) Сибирского отделения РАН // Выездное заседание научно-координационного совета по целевой программе “Информационно-телекоммуникационные ресурсы СО РАН”, Иркутск, 29-30 августа 2002. – Иркутск: ИДСТУ СО РАН, 2003.
- [4] The Grid: Blueprint for a New Computing Infrastructure / Eds. I. Foster and C. Kesselman. – San Francisco: Morgan Kaufmann Pub. CA., 1999.
- [5] Шокин Ю.И., Федотов А.М. Информационная система Сибирского Отделения РАН // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Вторая Всероссийская научная конференция, Протвино, 26-28 сентября 2000 г.: Сб. докл. – Протвино: ГНЦ ИФВЗ, 2000. – С. 6–15. – ISBN 5-88738-029-2.
- [6] Байков К.С., Коропачинский И.Ю., Шокин Ю.И., Шумный В.К., Ермаков Н.Б., Колчанов Н.А., Федотов А.М. Электронные коллекции и проблемы биоразнообразия // Там же. – С. 58–65.
- [7] Федотов А.М., Шокин Ю.И. Электронная библиотека Сибирского отделения РАН // Информационное общество. – 2000. – № 2.

