

Петров А.Б.

ПРИМЕНЕНИЕ ПРИНЦИПОВ И ПОДХОДОВ ОТКРЫТЫХ СИСТЕМ ПРИ ИНТЕГРАЦИИ РАСПРЕДЕЛЕННЫХ ИНФОРМАЦИОННЫХ РЕСУРСОВ

*Московский государственный институт радиотехники, электроники и автоматики
(технический университет)*

Введение

Статья обобщает результаты работы по грантам РФФИ 05-07-90251-в, 06-07-89297-а и по ведущей научной школе НШ-6452.2006.9.

Актуальность решения задачи интеграции распределенных информационных ресурсов (ИР) неоднократно отмечалась в ряде публикаций [1-3]. К настоящему времени эта задача полностью перешла в плоскость практического решения.

Существует два пути решения стоящей задачи на основе модернизации существующей инфраструктуры ИР:

- с переводом всех интегрируемых ИР на использование одной выбранной коммерчески доступной технологии построения распределенных ИР,
- с объединением гетерогенных по архитектуре, технологиям построения и структурами данных ИР на основе создаваемой технологии интеграции.

Первый путь сопряжен с существенными затратами на модернизацию в момент ее осуществления, возможны потери информации при переходе на новую платформу и требуется значительное время на его осуществление. Вместе с тем, при использовании технологии построения ИР, отвечающей требованиям открытости [3-4] можно получить эффективно работающее решение, так как при построении ИР будет использоваться одна технология одного производителя с единой архитектурой и, при необходимости, с едиными словарями и классификаторами. Этот путь эффективен при создании «с нуля» ИР.

Второй путь обеспечивает постепенную интеграцию в процессе функционирования ИР по мере того, как будет доработана их организация на основе единых принципов интеграции, классификаторов и словарей. Постепенность процесса и возможность наращивания количества интегрируемых ИР является несомненным достоинством второго пути. Недостатками являются наличие дополнительных аппаратно-программных средств интеграции, необходимость модернизации каждого ИР при интеграции, а также затратность процесса. В общем случае, временные и ресурсные затраты могут превосходить затраты по первому пути, а достижение интеграции всех потребных ИР может быть не обеспечено. Тем не менее, интеграция распределенных ИР, построенных на основе однотипных принципов организации, а также ИР, использующих разные внутренние архитектуры, но имеющие встроенные WEB-службы, с учетом обеспечения непрерывности использования ИР является эффективной.

В работе будет рассмотрен второй путь интеграции распределенных ИР, реализуемый на основе принципов и подходов открытых систем [3-4].

Принципы и подходы открытых систем

Практическая реализация интеграции распределенных ИР основана на выборе решений архитектурных и технологических решений, отвечающих базовым принципам открытых систем [1]:

- масштабируемость - способность системы к расширению функций и предоставляемых услуг, внедрения новых технологий, увеличению количества источников информации и пользователей; замены отдельных компонентов без перестройки всей системы;
- переносимость - освобождение от зависимости от одного поставщика аппаратных или программных средств, возможность выбора продукции из имеющихся на рынке при условии соблюдения поставщиками соответствующих стандартов открытых систем;
- интероперабельность - способность к взаимодействию с другими автоматизированными информационными системами (ИС) (в том числе использующих разную аппаратную платформу и различные операционные системы) и использования информационных ресурсов, имеющихся в других ИС;
- дружелюбность к пользователю – способность сформировать комфортную, легко управляемую среду, в которой работает пользователь.

Соблюдение перечисленных принципов экономит инвестиции сейчас и при будущих модернизациях архитектурных и технологических решений при разработке и внедрении практических систем интеграции ИР.

Рассмотрим подробнее реализацию принципов и подходов открытых систем при создании систем интеграции распределенных ИР.

Масштабируемость системы интеграции заключается в осуществлении следующего функционала:

- обеспечение возможности наращивания *количества* интегрируемых ИР;
- обеспечение возможности наращивания *типов построения (архитектур, моделей данных, семантик и т.д.)* интегрируемых ИР;
- обеспечение возможности наращивания *количества пользователей* интегрированными ИР;
- расширение и развитие *содержания* классификаторов и словарей;
- адаптация модели данных к требованиям потребителя информации;
- адаптация объема и формата запрашиваемых данных к объему и форматам запросов пользователя.

Переносимость и интероперабельность системы интеграции распределенных ИР заключаются в обеспечении следующих организационных, технических и функциональных возможностей:

- тиражируемость решения по интеграции ИР на существующие и перспективные ИР с разной архитектурой, разной организацией, выполненных в разных средах;

- использование единых архитектурных принципов организации системы интеграции распределенных ИР;
- применение единых интерфейсов информационного обмена между элементами системы интеграции;
- унификация набора используемых словарей и классификаторов, применение для их организации требований международных стандартов и спецификаций;
- применение типовых языков программирования и представления данных вне конкретного ИР;
- обеспечение возможности интеграции ИР имеющих различных владельцев без нарушения их прав собственности, выполненных по различным принципам, в различных средах, различными разработчиками.

Дружественность к пользователю реализуется посредством обеспечением доступа распределенного пользователя к ресурсам системы интеграции вне зависимости от технологий передачи информации, способа доступа, удаленности от ресурса и т.д. Важным является также единство применяемого интерфейса пользователя интегрированными ресурсами, его независимость от операционных систем, технологических платформ и принципов организации. Также, для обеспечения друженности к пользователю, в системе интеграции необходимо предусмотреть возможность самообучения системы, выполнения автоматических настроек и т.д.

Применение принципов и подходов открытых систем [3-4] необходимо использовать на всех этапах разработки: от выбора стандартов и спецификаций на конкретное решение до выбора аппаратных и программных платформ и сред программирования.

Реализация принципов и подходов открытых систем на разных уровнях интеграции

Анализ положений международных стандартов и публикаций в научной сфере [5-8] позволяет выделить следующие основные уровни интеграции:

- уровень физической связности, когда обеспечивается взаимодействие различных ИР на уровне реализации обмена информацией на уровне телекоммуникационных протоколов;
- уровень приложений, когда обеспечивается взаимодействие различных ИР на уровне обеспечения связности между ними, совместной работы и единого интерфейса пользователя,
- уровень данных, когда обеспечивается взаимодействие различных ИР на уровне обеспечения единого (единообразного) представления и восприятия данных.

Уровень физической связности в настоящее время достаточно хорошо разработан, имеются базовые технологии передачи информации, технологические и программные решения, стандартные протоколы. Принципы и подходов открытых систем на этом уровне реализуется в виде наличия открытых функциональных стандартов (профилей) на элементы базовых телекоммуни-

кационных технологий и на технологии в целом, на стандартные интерфейсы, на среды передачи данных.

Практические решения по интеграции приложений разрабатываются сейчас практически всеми крупными разработчиками аппаратных средств и программного обеспечения: IBM, Sun, Microsoft, IBM, Oracle, рядом европейских и российских компаний. Большинство из них основывается на использовании открытых стандартов и спецификаций, открытых архитектурных решений, стандартных интерфейсов.

При этом разработчики используют оба пути интеграции, перечисленные во введении: либо создавая универсальную оболочку для интеграции (например, WebSphere), либо используя решения на основе WEB-служб и сервис ориентированной архитектуры (SOA).

Интеграция приложений

Интеграция приложений является первым уровнем интеграции, который необходимо создать для обеспечения взаимодействия распределенных ИР и пользователя после обеспечения физической связности между ними. Для ее осуществления необходимо, чтобы в составе ИР была встроенная WEB-служба, реализующая функции адаптера [6-7] для преобразования данных из вида внутреннего хранения в единый вид, определяемой системой интеграции, и обратно. Наличие такой службы не затрагивает целостность ресурса, не меняет его архитектуру, не нарушает его информационное содержание.

Решения, предлагаемые сегодня, используют сервис-ориентированную архитектуру SOA (service oriented architecture).

Стандарт IEEE P1471 [8] определяет, что архитектура программного обеспечения является основополагающим руководством для построения какой-либо системы, реализованной в виде ее компонентов, их отношений между собой и с внешней средой, а также принципов проектирования и эволюции системы.

Архитектура, ориентированная на SOA, обладает следующими свойствами: модульностью, возможностью инкапсуляции, документированным интерфейсом. Выделяют три уровня архитектуры SOA: системная архитектура (первый уровень), оркестровка сервисов (маршрутизация на базе контента), управление метаданными, мониторинг и администрирование, безопасность, надежность, готовность, масштабируемость (второй уровень) и функциональные элементы (третий уровень).

SOA позволяет разработчикам, провайдерам и потребителям сервисов взаимодействовать вне зависимости от использованных ими технических решений и с минимальным знанием друг о друге – достаточно уметь идентифицировать и использовать необходимый сервис. Взаимодействие с сервисами может быть прямое или опосредованное через ПС промежуточного слоя (интеграционные брокеры и очереди сообщений).

SOA поддерживает синхронный режим сервиса, то есть при отработке запроса обе стороны находятся в режиме взаимодействия. SOA использует

для интеграции приложений web-сервисы. Выделяют два поколения развития web-сервисов. Первое поколение WEB-сервисов обеспечивало открытие существующих систем для внешнего взаимодействия. Второе поколение WEB-сервисов основано на обмене документами, интерфейсах взаимодействия и многочисленных расширениях протокола SOAP.

Все продукты, входящие в семейство WebSphere, созданы на основе открытых стандартов. Инструменты среды поддерживают текущие версии таких стандартов, как XML, WEB -сервисов и UML.

Применение принципов и подходов открытых систем при осуществлении интеграции приложений осуществляется посредством применения:

- открытых стандартов и спецификаций, функциональных стандартов (профилей), определяющих основные требования и решения;
- открытой архитектуры, определяющей решение по интеграции;
- открытых языков программирования и открытых программных платформ;
- единых стандартных интерфейсов.

За счет этого обеспечивается:

- множественность доступных для интеграции типов ИР;
- интеграция распределенных ИР;
- вариативность архитектурных решений, аппаратно-программных сред реализации и разработчиков ИР.

Таким образом, технологии интеграции приложений, основанные на использовании SOA и WEB-сервисов, должны служить основой для решения интеграционных задач в области фундаментальных и прикладных научных исследований.

Интеграция данных

Необходимость интеграции данных возникает после осуществления интеграции приложений, когда уже обеспечена логическая связность ИР и пользователя и возникает необходимость в устранении следующих неоднородностей в восприятии данных, получаемых от распределенных ИР:

- различия в представлении одинаковых данных, включая написание, последовательность, синтаксис, формат и т.д.;
- различия в интерпретации однотипных данных;
- семантики данных, имеющих одинаковое написание и различный смысл;
- единообразия написания и толкования объектов.

Проблема интеграции данных возникает при решении ряда управленческих задач, связанных с необходимостью поиска конкретной информации по запросу, и обусловлена многообразием видов ресурсов, моделей и сред, подходов к организации хранения и представления информации, программных сред, способов доступа и т.д.

Среди основных решений по интеграции данных можно выделить проекты электронного правительства, электронного государства и электронного бизнеса [9-12].

Таким образом, интеграция данных является вторым уровнем интеграции, который необходимо создать для обеспечения взаимодействия распределенных ИР и пользователя после обеспечения логической связности между ними.

Реализация семантической связности требует, чтобы встроенная WEB-служба в составе ИР, реализующая функции адаптера, выполняла не только логическое преобразование данных, но и осуществляла семантическое преобразование к единому виду данных, определяемых содержанием словарей и классификаторов [1] из вида внутреннего хранения и обратно.

Возможны и другие решения, когда подобное преобразование осуществляется на стороне пользователя или в промежуточном между ИР и пользователем слое.

Эффективность каждого из решений и их применимость в каждом конкретном случае можно определить и выработать соответствующие рекомендации. Главное же, что их объединяет, это осуществление преобразования семантики ИР в единую семантику пользователя.

Применение принципов и подходов открытых систем при интеграции данных заключается в использовании следующих функциональных решений:

- стандартных представлениях данных, базирующихся единых словарях и на общепризнанных классификаторах;
- модернизации языка представления данных для обеспечения семантической однородности;
- унификации описания и представления объектов;
- единых регламентов, описывающих правила создания ИР, представления данных в ИР, интерпретации данных в ИР.

Заключение

Многообразие решений, моделей и архитектур, разрабатывающихся в настоящее время, говорит о том, что сами процессы интеграции находятся только на начальном этапе своего развития. Вместе с тем, в ближайшее время следует ожидать появления стандартизованных (де-факто и де-юре) моделей интеграции и архитектурных решений, что позволяет выразить уверенность в том, что на рынке появится 3-5 общепризнанных эффективных решений, а дальнейшее развитие интеграционных процессов пойдет по пути их совершенствования.

На уровне физической связности мы такие решения имеем. Здесь возможно только появление более совершенных технологий передачи данных.

На уровне логической связности большинство производителей прошли начальный этап пути, выработали базовые подходы и приступили к их практической реализации.

На уровне семантической связности при полной ясности в целом того, что и как надо делать, практических работ по разработке, модернизации и актуализации классификаторов, создания унифицированных словарей, а, главное, работающих решений, крайне мало.

Таким образом, можно сделать вывод о том, что необходим перенос центра тяжести от работ, связанных с логической интеграцией на семантику. Применение принципов и подходов открытых систем для решения всего спектра задач интеграции даст экономию инвестиций и обеспечит эффективность самой интеграции.

Список использованных источников

1. Петров А.Б. Модель интеграции информационных ресурсов на основе стандартных словарей запросов, описания ресурсов и протоколов обмена и представления. - Современные проблемы информатизации в системах моделирования, программирования и телекоммуникациях. СПИ-МП-2004 - Сб. трудов IX Международной открытой научной конф. Выпуск 9. – Воронеж: «Научная книга», 2004. – с. 369-370.
2. Батоврин В.К., Петров А.Б. и др. Интеграция информационных ресурсов на основе методов функциональной стандартизации. Труды XI Всеросс. научно-методич. конф. «Телематика '2004» - С.П., 7-10 июня 2004 г. – с.47-48.
3. Батоврин В.К., Королев А.С., Петров А.Б. Технология интеграции распределенных информационных ресурсов. - Информационные технологии моделирования и управления №7(25), 2005 – с. 987-992.
4. Батоврин В.К., Гуляев Ю.В., Петров А.Б. и др./ Под ред. Олейникова А.Я. Технология открытых систем. – М., «Янус-К», 2004. – 288 с.
5. Концепция создания интегрированной автоматизированной информационной системы Минобразования России (проект). Телекоммуникации и информатизация образования. №1(8), январь-февраль 2002, стр. 5-17.
6. Когаловский М.Р. Интеграция данных в информационных системах./Сб. трудов III-й Всероссийской практической конференции «Стандарты в проектах современных информационных систем». М., 2003 - с.83-85.
7. М.Р. Когаловский. Энциклопедия технологий баз данных. – М.: Финансы и статистика, 2002. – 800 с.
8. IEEE P1471-2000 Recommended Practice for Architectural Descriptions.
9. Электронное правительство: Рекомендации по внедрению в Российской Федерации. Под редакцией Дрожжинова В.И., Зиндера Е.З. – М.: ЭКО-ТРЕНДЗ, 2004.
10. SAGA. Standards and Architecture for e-government Applications. Version 2.0. KBSt Publication Series ISSN 0179-7263, Vol. 59, December 2003.
11. E-Gov Enterprise Architecture Guidance (Common Reference Model). Draft – Version 2.0, FEA Working Group, 25 July 2002.
12. <http://www.govtalk.gov.uk>. e-Government Interoperability Framework. Version 6.0. 30 April 2004.