

Развитие инфраструктуры распределенных хранилищ спутниковых данных: интегрированная распределенная среда неоднородных информационных ресурсов исследования Земли из Космоса*

© Е.Б. Кудашев, А.Н. Филонов

Институт космических исследований РАН, Институт проблем информатики РАН
kudashev@iki.rssi.ru, filonov_a@inbox.ru

Аннотация

Статья представляет собой обзор текущего состояния технологий в области распределенных информационных систем, предназначенных для хранения и предоставления доступа к данным дистанционного зондирования Земли из космоса, а также возможных путей развития подобных систем в ближайшие годы. Рассматриваются отдельные технические аспекты реализации информационных систем трех поколений – от проекта первого поколения INTAS IRIS, до инициативы GMES НМА, которая сейчас находится на стадии разработки и тестирования.

1. От виртуальных центров научных данных к инфраструктуре Spatial Data Infrastructure (SDI, NSDI, GSDI)

Широкое признание научного сообщества в последние годы получает концепция «Виртуального управления информацией». Эта парадигма привела к идее построения «Виртуального Научного Центра данных» [1, 2]. В связи с бурным ростом количества наблюдательных данных в астрономии стремительно завоевывают популярность и интенсивно разрабатываются концептуальные основы «Виртуальной Обсерватории» как интегрированной системы распределенных гигантских астрономических архивов и баз данных.

В США с 2000 года разрабатывается проект National Virtual Observatory [3], немного позднее начался подобный проект в Европе [4]. В июне 2002 года создан международный консорциум IVOA (The International Virtual Observatory Alliance) [5].

Основной целью этой Программы является разработка «Виртуального Научного Центра астрономических данных» [6]. Исследуются проблемы управления и распространения астрономическими данными. Астрономические каталоги становятся взаимосвязанными: создается сеть интегрированных, интероперабельных баз данных в объемах, измеряемых в петабайтах [6, 7].

Концепция Виртуальной Обсерватории обеспечивает исследователю возможность использовать астрономические данные от многих различных инструментов, используя передовые ИТ и развивающуюся международную кооперацию научного сообщества. Общий уровень финансирования проектов за последние 3-5 лет приблизился к 20 миллионам долларов США.

Тенденцию лавинообразного увеличения данных можно отметить в области наук о Земле, в геофизике и в метеорологии. Интенсивные исследования глобального потепления и изменения климата приводят, например, к тому, что коллекции данных в наши дни увеличиваются в объеме на величины порядка нескольких петабайт в день.

Наряду с «Виртуальной обсерваторией» развивается программа «Виртуальный Научный Центр данных геофизики плазмы и солнечно-земной физики». Международный консорциум SPASE – Space Physics Archive Search and Extract [8, 9] создан научным сообществом Европы и США. Проект в настоящее время получил поддержку National Science Foundation USA и начал реализовываться; уровень финансирования достигает нескольких миллионов долларов США. Технологические достижения, объединяя распределенные по всему миру гигантские цифровые архивы и базы данных, фундаментально изменяют характер исследований космической плазмы, позволяют решить проблему доступа к данным и обмена информацией в Солнечно-Земной физике.

«Виртуальный Научный Центр» в области изучения планет создается NASA (National Aeronautics and Space Administration) – проект PDS (The Planetary Data System). Образован международный консорциум и сформирована система on-line архивов, состоя-

Труды 10-й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» – RCDL'2008, Дубна, Россия, 2008.

щая из нескольких географически распределенных центров, связанных Сетью Интернет [10, 11].

Естественным шагом в развитии концепции виртуальных обсерваторий является развитие геоинформационной инфраструктуры распределенных спутниковых данных. Проект выполняется в Институте космических исследований РАН и известен как Распределенная среда интегрированных информационных ресурсов в области космического мониторинга природной среды (Виртуальный Научный Центр спутниковых данных) [12, 28].

В США стратегия и концепция National Spatial Data Infrastructure (NSDI) начала разрабатываться с 1994 г. [13].

Основными целями программы National Spatial Data Infrastructure являются:

- Организация и проведение национальных дискуссионных форумов.
- Совершенствование механизма доступа к данным путем организации центров информационного обмена и создания баз метаданных.
- Создание баз пространственных данных (БПД).
- Создание тематических данных, критически важных для государства.
- Подготовка образовательных программ для обучения персонала.
- Развитие всестороннего партнерства и координация сбора и использования пространственных данных.

Концепция Spatial Data Infrastructure стремительно завоевала международную популярность. В 2003 г. Европейская Комиссия приняла решение о создании INSPIRE (Infrastructure for Spatial Information in Europe) – Глобальной геоинформационной инфраструктуры данных в Европе [14]. Аналогичный проект создается в Канаде – Canadian Geospatial Data Infrastructure (CGDI).

Основные задачи программы развития геоинформационной инфраструктуры Spatial Data Infrastructure (SDI) можно сформулировать следующим образом:

- Построение глобальной инфраструктуры ИТ и геоданных.
- Гармонизация спутниковой информации; использование согласованного набора стандартов, «понятных» всем участникам данной системы.
- Интероперабельность между независимо созданными приложениями и базами данных; Совместить между собой их интерфейсы, протоколы и форматы данных.
- Поддержка единой политики данных: а) доступ к данным и б) создание и поддержка спутниковой информации – для задач экологии, но открытой и для сельскохозяйственных и транспортных проблем.

Для того, чтобы глобальная геоинформационная инфраструктура стала реальностью в России, возникла необходимость определить те проблемы и особенно-

сти, с которыми сталкиваются национальные научные сообщества исследователей Земли из космоса при поиске оптимального доступа к спутниковым данным. Эти исследования на начальном этапе были выполнены авторами в проекте INTAS IRIS Project: Integration of Russian Satellite Data Information Resources with the Global Network of Earth Observation Information Systems [15, 16]. Проблема интеграции спутниковых архивов является ярким примером актуальности инфраструктурных решений при организации удаленного доступа и распределенной информационной среды. В проекте исследуются новые задачи аэрокосмического дистанционного зондирования, связанные с проблемой виртуального управления информацией в географически удаленных спутниковых цифровых архивах. Важнейшие научные задачи, охватывающие область исследования Земли из космоса, глобальный экологический мониторинг и аэрокосмическое дистанционное зондирование природной среды, актуальны также для процессов интеграции данных солнечно-земной физики, для программ интеграции Цифровых Архивов данных планетарных исследований. Проект IRIS представил концепцию Виртуального Научного Центра: было показано, что условием успешной интеграции служит использование технологий открытых систем и единых стандартов; решение проблемы гармонизации спутниковых данных для создания распределенных ресурсов. Основное внимание при решении задач поиска ресурсов, взаимодействия с ресурсами, извлечения ресурсов, совместного использования ресурсов необходимо уделять проблемам разработки масштабной информационной инфраструктуры интеграции ресурсов, удаленного доступа и управления гетерогенными ресурсами спутниковых архивов.

Главной целью проекта было улучшить доступ к российским спутниковым данным. Проект поддерживает объединение данных регионального спутникового экологического мониторинга. Обеспечивается доступ российских пользователей к европейской информационной системе INFEO (метаданные / on-line каталоги). При запросе в единой точке системы поиск данных осуществляется одновременно в ресурсах распределенных мировых каталогов. Запрос на метаданные оформляется единым образом для всех спутниковых центров, и информация ищется сразу по всем центрам, доступным из глобальных поисковых систем (в первую очередь, это система EOSDIS/ USA и система INFEO). Запрос на получение собственно спутниковых данных также выполняется автоматически и единым образом, а данные приходят в удобном формате для обработки в рамках GIS-систем.

Формирование геоинформационной инфраструктуры спутниковых данных определяется решением следующих задач: использование GIS и Web технологий для разработки новых средств поиска и обнаружения данных аэрокосмического зондирования; включение российских спутниковых каталогов в международную систему INFEO; создание описания российских спутниковых коллекций и регистрации их в IDN (International Directory Network).

Основной задачей проекта является интеграция географически удаленных российских спутниковых архивов. Требуется обеспечить работу пользователей с данными, которые получают из различных источников и имеют разную структуру хранения. Создается единая информационная система, основанная на концепции интероперабельных распределенных архивов. Необходимо обеспечить требования действующих международных стандартов: ISO/TC 211, FGDC USA (Federal Geographic Data Committee), CEOS (Committee on Earth Observation Satellites).

2. Информационная система первого поколения: проект INTAS IRIS

Международная Информационная система INFEO (Information about Earth Observation) – инициатива Европейского Космического Агентства (ESA), призванная решить проблему доступа к распределенным данным дистанционного зондирования Земли (ДЗЗ). основополагающим принципом системы является обеспечение интероперабельности – возможности взаимодействия гетерогенных сервисов и каталогов данных в рамках единой информационной системы.

В проекте IRIS были исследованы информационные задачи в области аэрокосмического дистанционного зондирования Земли, связанные с проблемой распределенного доступа и управлением информацией в географически удаленных спутниковых цифровых архивах. Разработка методов и программных средств сбора, тематической обработки, систематизации больших потоков аэрокосмической информации и хранения экстремальных объемов изображений особенно актуальна для космического мониторинга земной поверхности, водной среды, атмосферы, источников антропогенных воздействий на них, опасных природных процессов и объектов техносферы. Основные задачи Активного Архива данных спутниковых исследований, следуя работе [24], представим в такой последовательности: *Публикация* – процесс представления ресурса одним пользователем системы, в результате которого другие пользователи могут

найти этот ресурс и обратиться к нему. *Резюмирование* – процесс извлечения представительного множества характеристических сведений (метаданных) из ресурса или его окружения. *Индексирование* – процесс перевода информации, полученной после резюмирования, в форму, обеспечивающую эффективный распределенный поиск. *Поиск* – процесс обработки «индекса», имеющий целью сформировать ответ на запрос пользователя по распределенным метаданным. *Просмотр* – контролируемая пользователем деятельность, направленная на обследование распределенного информационного пространства, формируемого системой на основе ресурсов и их связей. *Извлечение* – деятельность пользователей, вызванная желанием получить (извлечь из распределенного информационного пространства) обнаруженную информацию. *Взаимодействие* – процесс, ведомый пользователем и находящийся в рамках, ведомых ресурсом, в течение которого ресурс предоставляет пользователю информационный или функциональный сервис. *Распределение* – деятельность системы, имеющая целью обеспечение работы с данными, расположенными на разных физических серверах, различных аппаратно-программных платформах. *Преобразование* – деятельность системы, направленная на перевод метаданных других систем, хранящихся в различных внутренних форматах, во внутренний формат системы для обеспечения интеграции информации и эффективного поиска. *Сотрудничество* – деятельность системы, ориентированная на совместную деятельность с другими информационными системами по обслуживанию поисковых запросов пользователей, в основном заключающихся в преобразовании поисковых запросов для этих систем и интеграции результатов их ответов. *Контроль* – деятельность системы по представлению различных уровней привилегий доступа к информации, её защите. *Управление* – деятельность по сопровождению системы, её ресурсов, соблюдению актуальности, целостности и сохранности информации в распределенном информационном пространстве. Схему функционирования Архива спутниковых данных иллюстрирует рис. 1 [25].

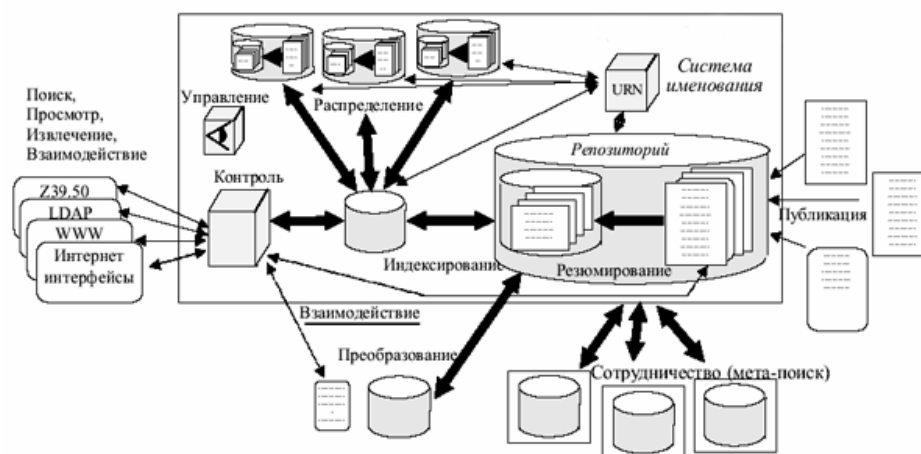


Рис. 1. Структура Активного Архива спутниковых данных

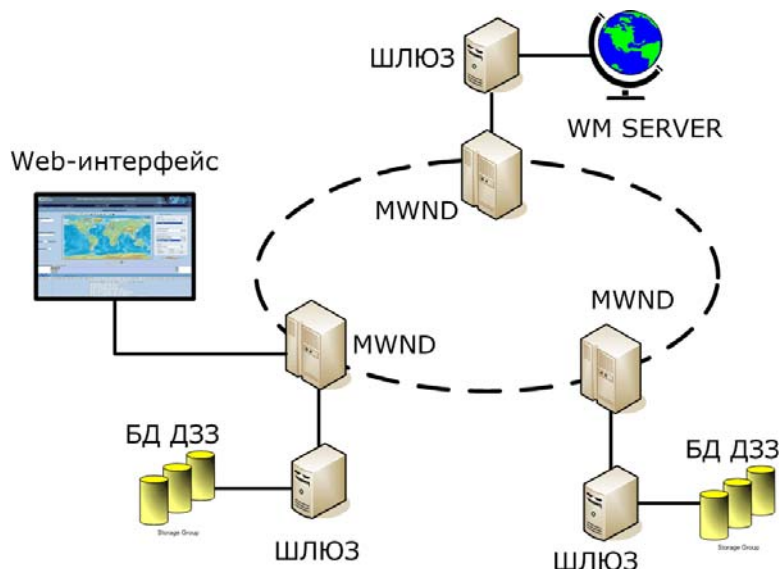


Рис. 2. Структурная схема распределенной системы первого поколения

Основная функция такой системы – создание единой информационной среды, предоставление единой точки доступа для поиска и получения информации по всем архивам и каталогам, зарегистрированным в системе, независимо от их географического расположения и внутреннего формата данных. Распределенная информационная система первого поколения представляет собой сеть узлов, так называемых Middleware Nodes (MWNDs) (рис. 2).

Каждый из узлов содержит метаданные коллекций, описывающие данные, доступ к которым предоставляется через Шлюзы (Gateways). Пользователь для поиска информации обращается к системе через WWW-сервер на одном из узлов MWND, где формируется распределенный запрос, рассылаемый релевантным шлюзам. Результаты выполнения запроса объединяются и возвращаются пользователю.

Шлюз представляет собой промежуточное звено между узлом MWND и каталогом или архивом поставщика данных. Шлюз обменивается сообщениями по протоколу CIP (Catalogue Interoperability Protocol) с узлом MWND, получая от него поисковые запросы и возвращая ему результаты поиска. Запросы передаются транслятору RDBMS, который преобразует их в запросы для непосредственного поиска в базе данных, либо формирует запрос для обращения к предоставляемому сервису (Web Map Server и т.п.).

В основе распределенной информационной системы лежит протокол CIP [17], регламентирующий правила взаимодействия пользователей и каталогов данных ДЗЗ. Для поддержки одновременного доступа пользователя к множеству каталогов используется трехуровневый принцип распределения запросов. Пользователь через web-интерфейс задает поисковый запрос и посылает его к узлу MWND, который, в свою очередь, перенаправляет запрос множеству серверов каталогов данных. Сервера, имеющие данные, удовлетворяющие критериям поиска, возвращают ответы на узел MWND, через

который пользователь получает объединенный результат поиска.

3 Технология интеграции ресурсов и сервисов второго поколения: среда SSE

В проекте INTAS IRIS авторы выполнили исследование инфраструктуры системы международного обмена космической информацией на основе технологий INFEO. Система INFEO, являясь активным многоцелевым архивом космических данных, предоставляет пользователям специализированные интерфейсы доступа к данным, обеспечивает долговременное хранение данных и позволяет производить поиск спутниковых снимков и коллекций данных.

Логическим развитием системы INFEO стала разработка и внедрение технологий интеграции второго поколения. Это было вызвано стремлением расширить участие ESA в глобальном экологическом мониторинге, улучшить в задачах охраны окружающей среды взаимодействие с мировыми информационными системами ДЗЗ [13] и необходимостью поддержать требования Директивы ЕС – INSPIRE [14] о совместимости с другими информационными системами. Развитие технологий интеграции второго поколения направлено на миграцию от подхода, основанного на технологии Z39.50/Gateways, к использованию веб-сервисов (XML над SOAP/HTTP) для связи внешних каталогов с каталогами INFEO. Технологии интеграции второго поколения, получившие название EOLI-XML, разрабатываются ESA в сотрудничестве с национальными космическими Агентствами: Франции (Centre National d'Etudes Spatiales – CNES), Италии (Italian Space Agency – ASI), Германии (German Space Agency – DLR), European Union Satellite Centre (EUCS), Joint Research Centre (JRC) при участии ряда партнеров в промышленности: SPOT Image, Siemens, Spacebel и др. EOLI-XML

представляет собой внешний интерфейс для обмена сообщениями между сервером и клиентом на основе протокола SOAP [18].

Технология EOLI-XML лежит в основе системы SSE (Service Support Environment) [19], разрабатываемой для решения указанных задач с помощью создания открытой, сервис-ориентированной, распределенной среды. Система SSE является инфраструктурой, единой средой для потребителей и провайдеров данных и сервисов. SSE объединяет своих пользователей на основе стандарта XML с использованием протоколов обмена сообщениями SOAP и языка описания веб-сервисов WSDL (Web Service Definition Language) [20]. Таким образом, любой сервис, интегрируемый в среду SSE, должен предоставить SOAP-интерфейс, в соответствии с SSE ICD (Interface Control Document) [21].

Итак, основные цели создания инфраструктуры SSE:

- Предоставить среду, максимально упрощающую взаимодействие типа «сервис-провайдер – пользователь» и «сервис-провайдер – сервис-провайдер».
- Упростить интеграцию уже существующих сервисов, предоставляя универсальный интерфейс на основе XML, позволяющий не изменять их структуры.
- Предоставить пользователям единую точку входа – Интернет-портал системы.

Система SSE состоит из двух основных компонентов: SSE Portal Server и AOI Server, образующих вместе Интернет-портал, с которым взаимодействует конечный пользователь.

SSE Portal Server предоставляет web-интерфейс для доступа пользователей к portalу. Сервер построен с использованием промышленной технологии J2EE, которая удовлетворяет требованиям SSE: хорошая масштабируемость, интеграция с существующими информационными системами, гибкая политика безопасности, поддержка стандартных протоколов и языков, переносимость компонент без необходимости перекомпиляции. Выполнение этих требований является обязательным для успешного построения информационной инфраструктуры. В качестве сервера приложений выбран сервер JBoss с открытым исходным кодом.

AOI Server работает в связке с SSE Portal Server и служит для предоставления для сервис-провайдеров функций визуального выделения области на карте (area of interest, AOI) при задании критериев поиска, а также визуализации результатов поиска.

Так как большинство сервисов в качестве одного из входных параметров требуют указание географической области, SSE Portal предоставляет специальную поддержку этой возможности. При обращении к сервису пользователь должен указать географическую область и сделать это он может разными способами: выбрать из списка, указать область на карте, загрузить файл, описывающий AOI, задать AOI с помощью указания координат.

При обращении пользователя к сервису, SSE Portal использует апплет, загружаемый с сервера AOI для визуального отображения выделенной области на карте. SSE позволяет использовать для визуализации как локально хранимую информацию, так и подключаться к OGC WMS серверам для загрузки дополнительных слоев карты [22]. Выделенная таким образом на карте область затем преобразуется в описание в формате GML (Geography Markup Language) [23] описание вместе с запросом отправляется сервис-провайдеру в виде SOAP сообщения.

Типичный сценарий работы выглядит следующим образом. Пользователь через интерфейс SSE портала формирует запрос к нужному сервису, выделив на карте область AOI и указав характеристики искомой информации. После этого сформированный в формате SOAP-сообщения запрос пересылается на URL сервиса, указанный провайдером при регистрации. Здесь сообщение поступает на вход сервера веб-приложений, где в зависимости от поступившего запроса запускается соответствующий сервлет, который занимается дальнейшей обработкой запроса и передачей его на внутренний интерфейс сервиса. Для выполнения сервлетов сервис-провайдеру вместе с Toolbox необходимо установить web-сервер с поддержкой сервлетов (рекомендуется к установке веб-сервер Tomcat).

Обработка входного запроса и формирование ответа на запрос выполняется на основе WSDL-документа в соответствии с XML-схемой сервиса. При установке Toolbox уже имеется предопределенная схема, тем не менее, некоторые сервисы могут иметь специфические параметры входных и выходных сообщений, что можно легко исправить, отредактировав соответствующие файлы XML-схем.

Для поиска необходимой информации пользователь может воспользоваться поисковым интерфейсом по адресу <http://catalogues.eoportal.org/eoli.html>. Интерфейс выполнен в виде Java-апплета (рис. 3).

Существует возможность ограничивать поиск либо по поставщикам данных (панель Find Products), либо по коллекциям данных (Find Collections). В раздел Find Products входят крупные поставщики, имеющие несколько собственных обширных коллекций. Например, ESA (коллекции Radar Imagery, Radar Altimetry, Optical/Multispectral Imagery), DLR (коллекции Atmospheric Sensors, Thematic Maps, Optical Sensors, Digital Elevation Model, Synthetic Aperture Radar Data) и др. После выбора типа данных, по которым будет ограничен поиск, следует выбрать площадь земной поверхности. Для более точного поиска информации можно воспользоваться режимом расширенного поиска Advanced Query Mode. Этот режим позволяет дополнительно задавать ключевые слова для поиска, а также выбирать спутник и сенсор, с помощью которого были получены данные. Ключевые слова служат для отбора данных, поле описания которых содержит строку, заданную в поле Free Text.



Рис. 3. Поисковый интерфейс пользователя

4 Интеграция неоднородных распределенных информационных ресурсов спутниковых данных: инициатива GMES и проект НМА

Основным предназначением инициативы Global Monitoring for Environment Security (GMES) является поддержка задач Европейского Союза, касающихся устойчивого развития и глобального управления с помощью обеспечения доступа к высококачественным данным дистанционного зондирования Земли в режиме реального времени. Доступ к такой информации имеет стратегическое значение как с точки зрения современного развития отдельных регионов Европы, так и с точки зрения Европейского Союза в целом, как члена мирового сообщества.

Информация, получаемая с помощью проекта GMES, используется в трех основных направлениях деятельности: экология и защита окружающей среды, поддержка развития инфраструктуры ЕС, обеспечение безопасности граждан, как в повседневной жизни, так и в режиме чрезвычайных ситуаций.

Европейское космическое агентство (ESA) – главный партнер европейской комиссии по разработке и внедрению проекта GMES, работающий в тесном сотрудничестве с национальными космическими агентствами Европейских стран, а так же Канады и России.

В настоящий момент доступ к данным дистанционного зондирования Земли каждого спутникового оператора осуществляется через отдельную точку доступа («портал») этого поставщика данных. Поль-

зователь вынужден выполнять поиск нужных ему данных отдельно в каждой точке доступа. Особенно актуальной проблема своевременного получения данных становится в условиях режима чрезвычайных ситуаций, когда счет идет на часы и минуты.

Ключевой особенностью GMES в таком аспекте является гармонизация доступа к гетерогенным данным разных поставщиков данных, включая национальные космические агентства. С этой целью в рамках инициативы GMES европейское космическое агентство инициировало проект Heterogeneous Missions Accessibility (HMA). Это совместный проект с участием большого количества корпораций, поставщиков данных, научных и коммерческих организаций из Европы и Канады.

Проект НМА обеспечит гармонизацию сервисов и данных наземных сегментов поставщиков данных, с помощью сервис-ориентированной архитектуры, предоставляя единообразный доступ к данным и сервисам каждого из участников проекта. В данный момент проект находится на стадии активного внедрения и тестовой эксплуатации. Основные этапы разработки проекта НМА показаны на рис. 4 [26].

Не только базовые спецификации, лежащие в основе проекта НМА, но и большая часть архитектурных решений проекта находятся в свободном доступе, то есть принадлежат области всеобщего достояния (public domain), что значительно увеличивает возможности распространения технологий и программных продуктов проекта. Широко используются стандарты и протоколы ISO (International Organization for Standardization) и OGC (Open

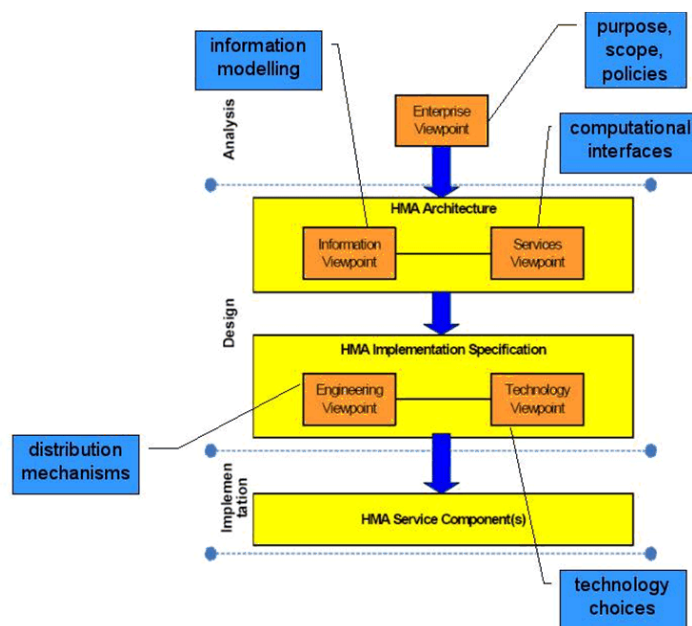


Рис. 4. Разработка проекта Heterogeneous Missions Accessibility (HMA)

Geospatial Consortium). Основополагающим стандартом системы является международный стандарт для построения открытых распределенных систем (RM-ODP ISO/IEC 10746-1:1998).

Модель построения распределенных систем RM-ODP была выбрана в качестве базовой для проекта HMA GMES, так как основные задачи проекта практически совпадают с задачами, для которых разрабатывалась модель RM-ODP:

- Поддержка распределенной обработки.
- Поддержка интероперабельности между гетерогенными системами.
- Скрытие механизмов распределенного взаимодействия от конечных пользователей и системных разработчиков.

Но есть и важное отличие от базовой модели – GMES-HMA будет действовать на основе слабо связанной сети систем и сервисов вместо распределенной вычислительной системы, базирующейся на взаимодействующих объектах.

Определены пять основных направлений развития проекта, определяющих минимальный набор требований и объектную модель, гарантирующую целостность и логическую связность системы:

- Организационное направление: определяет цели, область применения, ограничения и правила использования и развития системы.
- Информационное направление: определяет семантику информационных процессов.
- Вычислительное направление: декомпозиция всей системы на отдельные объекты, взаимодействующие между собой и с внешними интерфейсами с целью оптимизации вычислительных затрат.
- Направление инжиниринга: разработка механизмов и функций, необходимых для под-

держки распределенного взаимодействия между отдельными объектами системы.

- Технологическое направление: определяет выбор конкретных технологических решений, необходимых для выполнения поставленных задач.

Предложенная архитектура является сервис-ориентированной (Service Oriented Architecture, SOA), в которой поставщики данных получают возможность предоставлять доступ к своим сервисам через стандартизованные интерфейсы системы.

Ключевым компонентом системы HMA является интеграционный слой доступа к данным (Data Access and Interoperability Layer – DAIL), предоставляющий единый интерфейс доступа к каталогам и сервисам всех участников проекта GMES с помощью использования установленных стандартов интероперабельного доступа.

В структуре DAIL можно выделить основные компоненты:

- Среда выполнения сервисов – среда для выполнения основных компонентов DAIL, предоставляет средства выполнения сервисов и управления их взаимодействием с другими компонентами DAIL.
- База данных – служит для хранения данных, необходимых для функционирования служебных сервисов HMA.
- Рабочий процесс – основное приложение ядра архитектуры DAIL, служит для запуска и управления сервисами DAIL.
- Репозиторий профилей пользователей – база данных для хранения профилей пользователей системы и информации, необходимой для аутентификации пользователей в системе.
- Сервер коллекций – служит для хранения списка доступных в системе коллекций и отслеживания их статуса.

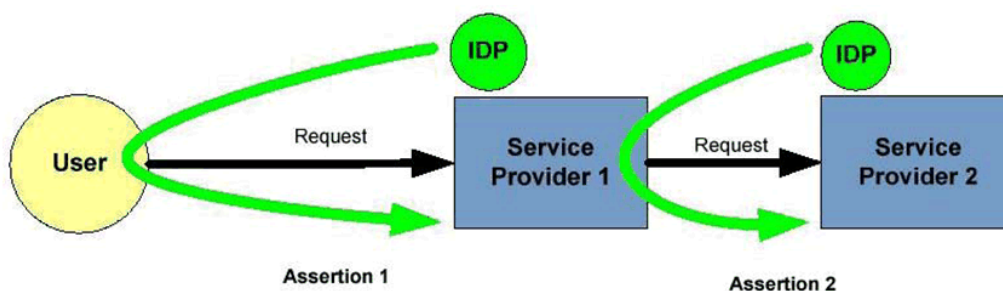


Рис. 5. Стратегия управления доступом к данным и сервисам в НМА

- Сервер сервисов – аналогично серверу коллекций, служит для хранения списка доступных в системе сервисов и отслеживания их статуса.
- Виртуальный FTP сервер – компонента для поддержки обмена данными между поставщиками и пользователями через систему НМА.

В процессе проектирования системы большое внимание было уделено вопросам безопасности, а при разработке учитывались такие стандарты, регламентирующие информационную безопасность, как ISO/IEC17799, ISO27001 и т.п. В качестве базовой для проекта была принята модель управления доступом «Distributor» (рис. 5). В данной модели пользователь авторизуется только у одного поставщика сервисов (SP1), который затем, в случае необходимости воспользоваться сервисами другого поставщика, обращается к нему под своими собственными правами; таким образом пользователю нет необходимости выполнять аутентификацию у каждого из поставщиков данных и сервисов, участвующих в цепочке обработки информации.

5 Заключение

Технологии и принципы построения распределенных систем, рассмотренные в данной статье имеют большой потенциал для дальнейшего развития в будущем и уже сегодня предоставляют широкие возможности для организации распределенных гетерогенных данных и сервисов в единую глобальную систему. Вектор развития каждого следующего поколения распределенных информационных систем состоит из трех основных слагаемых: уход от недостатков существующей системы, улучшение технических и информационных характеристик, решение новых задач, которых просто не существовало на момент разработки системы предыдущего поколения.

Коллектив авторов принимал активное участие в проектах всех трех поколений на этапе как разработки системы, так и практического ее внедрения. Так, на основе проектов INFEO IRIS и INSPIRE (технологии EOLI-XML и SSE) был построен и введен в действие российский сегмент распределенной информационной системы.

Российский сегмент распределенной информационной системы спутниковых архивов был реали-

зован при сотрудничестве ИКИ РАН и ИАПУ ДВО РАН. В качестве поставщика данных дистанционного зондирования Земли выступает Лаборатория спутникового мониторинга ИАПУ Дальневосточного Отделения РАН (<http://satellite.dvo.ru>). Программное обеспечение и серверное оборудование для функционирования шлюза сегмента установлено в Москве, в Институте Космических Исследований РАН (<http://iris.iki.rssi.ru>) [16]. Таким образом, мировое научное сообщество получило возможность доступа к данным одного из российских поставщиков данных ДЗЗ с помощью глобальной распределенной системы.

В заключении нельзя не отметить немаловажный аспект, выходящий за рамки данного обзора. К сожалению, в России до недавнего времени практически отсутствовала поддержка и заинтересованность государства в развитии аналогичных национальных информационных систем. Однако, в последнее время наблюдаются положительные сдвиги в этом направлении. Так, в 2006 году правительством РФ была одобрена и принята к исполнению «Концепция создания и развития инфраструктуры пространственных данных Российской Федерации» [27].

В 2004 году стартовал проект информатизации наук о Земле (геология полезных ископаемых, физика Земли, экология) «Электронная Земля: научные информационные ресурсы и информационно-коммуникационные технологии». В этом проекте исследован подход к решению задачи интеграции геопространственных (ГИС) данных на основе стандартов OpenGIS Consortium и ISO и специализированных методов их обработки с помощью технологий Grid и Интегрированной Системы Информационных Ресурсов (ИСИР). Этот проект является одним из пунктов Программы фундаментальных исследований Президиума РАН «Разработка фундаментальных основ создания научной распределенной информационно-вычислительной среды на основе технологий GRID». Проект имеет своей целью создание на основе технологий IT-инфраструктуры сетевой распределенной информационно-аналитической системы по наукам о Земле, обеспечивающей единообразный доступ к объединяемым в рамках проекта информационно-аналитическим и другим ресурсам. Создаваемая инфраструктура будет предоставлять возможность использования этих ресурсов для решения фундаментальных и прикладных задач, компьютерного моделирования и параллельных вычислений.

Участниками проекта являются свыше двадцати институтов РАН, как геофизической, химической, картографической направленности, так и решающих задачи в сфере распределенных вычислений и телекоммуникаций [29].

Литература

- [1] Кудашев Е.Б., Филонов А.Н. Технология и стандарты интеграции сервисов, каталогов и баз данных дистанционного исследования Земли из космоса // Труды Девятой Всероссийской научной конференции «Электронные Библиотеки»-RCDL'2007, с. 273–279. – Переславль-Залесский: Изд-во «Университет г. Переславля», 2007.
- [2] Antikidis J.P. Virtual Access to Information: an emerging concept // In Proceedings Workshop PV-2004. Ensuring the Long Term Preservation and Adding Value to the Scientific and Technical Data, pages 63–67, 2004.
- [3] Szalay A.S. The National Virtual Observatory // In Astronomical Data Analysis Software and System X. ASP Conf. Ser. Vol. 238, p. 3–12, 2001.
- [4] Virtual Observatory in Europe home page: <http://esavo.esac.esa.int/registry>
- [5] IVOA (International Virtual Observatory Alliance) home page: <http://www.ivoa.net/>
- [6] Arviset C, Dowson J, Hernandez J, Ortiz I, Osuna P., Salgado J, San Miguel G, Venet A. ESA RSSD Science Archives User Interfaces and Inter-Operability Systems. In Proceedings Workshop PV-2004 Ensuring the Long Term Preservation and Adding Value to the Scientific and Technical Data, pages 159–164, 2004.
- [7] Arviset C. ESA Scientific Archives and interoperable Virtual Observatory Systems // In Proceedings PV-2007 Conference. Ensuring the Long Term Preservation and Adding Value to the Scientific and Technical Data – Munich, DLR, 2007.
- [8] Thieman J. Et al. The SPASE Standard for Heliophysics Data Description, Location, and Acquisition // In Proceedings PV-2007 Conference. Ensuring the Long Term Preservation and Adding Value to the Scientific and Technical Data – DLR, Munich, 2007.
- [9] Harvey C. et al. SPASE –Space Physics Archive Search and Extract – <http://www.igpp.ucla.edu/spase/>
- [10] Hughes S. et al. A Planetary Data system for the 2006 Mars Reconnaissance Orbiter Era // In Proceedings PV-2004 Workshop. Ensuring the long-term Preservation and Value Adding, p. - ESA/ESRIN, Frascati, 2004.
- [11] Hughes S. et al. Defining the Core Archive Data Standards of the International Planetary Data Alliance (IPDA) // In Proc. PV-2007 Conference. Ensuring the Long Term Preservation and Adding Value to the Scientific and Technical Data – DLR, Munich, 2007.
- [12] Kudashev E.B. Digital Library: Improving the accessibility to Russian Satellite Data for mitigating Natural Disaster // In Proc. PV-2004 Workshop. Ensuring the long-term Preservation and Adding Value to the Scientific and Technical data, p. 233–240 – ESA/ESRIN, Frascati, Italy. 2004
- [13] National Spatial Data Infrastructure (NSDI) home page, 2007. <http://www.fgdc.gov/nsdi>
- [14] INSPIRE home page, 2007. <http://www.ec-gis.org/inspire>
- [15] Е.Б. Кудашев, А.Н. Филонов. Организация информационной распределенной среды и интеграция спутниковых архивов // Труды Седьмой Всероссийской научной конференции «Электронные Библиотеки» – RCDL'2005. – Ярославль: Ярославский ун-т, 2005.
- [16] Кудашев Е.Б., Филонов А.Н. Интегрированная распределенная информационная система спутниковых данных в программах исследования Земли из космоса // Труды Восьмой Всероссийской научной конференции «Электронные библиотеки» – RCDL'2006, Суздаль, с. 131–139. – Ярославль: Ярославский ун-т, 2006.
- [17] Catalog Interoperability Protocol (CIP) Specification – Release B-24, 2005. http://wgiss.ceos.org/ics/documents/cip2.4/cipspec-2_4_75_6.pdf
- [18] Earthnet On-line XML Front-End Interface Control Document, EOLI-XML-006-ICD, Issue 1.7, 2003. <http://earth.esa.int/rtd/Documents>
- [19] Service Support Environment Architecture, Model and Standards, 2007. http://earth.esa.int/services/esa_doc/doc_sse.html
- [20] Web Services Description Language (WSDL) 1.1, W3C Note, 2001. <http://www.w3.org/TR/wSDL>
- [21] Interface Control Document (ICD) for the Service Support Environment (SSE), 2007. <http://services.eoportal.org/massRef/documentation/icd.pdf>
- [22] The Open Geospatial Consortium, 2007. <http://www.opengeospatial.org/>
- [23] ISO/TC 211/WG 4/PT 19136 Geographic information – Geography Markup Language (GML), 2004. http://portal.opengeospatial.org/files/?artifact_id=4700
- [24] Бездушный А.Н., Жижченко А.Б., Кулагин М.В., Серебряков В.А. Интегрированная система информационных ресурсов РАН и технология разработки электронных библиотек // Программирование. 2000. Т. 26. № 4. С. 177–185.
- [25] Балашов А.Д., Каленова Н.И., Кудашев Е.Б. Интеграция электронных архивов спутниковых данных // VIII Международная конференция по электронным публикациям EL-Pub2003. <http://www.ict.nsc.ru/ws/elpub2003/rep6254.zip>
- [26] Lawrence B., Pritchard M., Woolf A. Review of the Heterogeneous Mission Accessibility Project // BNSC EO International Co-operation Programme. Year 4 (2006/07): Documents and Presentations. <http://www.pinkbox.space.qinetiq.com/icp2/docs>

- [27] Концепция создания и развития инфраструктуры пространственных данных Российской Федерации // Распоряжение Правительства РФ №1157-р от 21 авг. 2006 г.
- [28] Kudashev E. Spatial Data Infrastructure development of Russian Satellite Data for Global Monitoring and Disaster Management support // 2008 IEEE International Geoscience & Remote Sensing Symposium (IGARSS 2008). July 6-11, 2008 | Boston, Massachusetts, U.S.A. Abstracts of Symposium Presentations. Abstract 1754.
- [29] Вершинин А.В., Бездушный А.Н., Серебряков В.А. «Электронная Земля»: интеграция ГИС-данных с использованием технологий Grid // Труды Восьмой Всероссийской научной конференции «Электронные библиотеки» – RCDL'2006, Суздаль. – Т. 1. Статья 94. – Ярославль : Ярославский ун-т, 2006.

Distributed Infrastructure Development of Satellite Data: Integrated Distributed Environment of Heterogeneous Information Resources for Earth Observation

E.B. Kudashev, A.N. Filonov

Our presentation will focus on distributed infrastructure development for the integration of Russian satellite data information resources with EO Portal/INFEO. Access to geo-data and services are organised via the geo-Portal. Geo-Portal provides an efficient means of distributing geo-data, and for users to search for and obtain information and gain access to data. Distributed Infrastructure of Spatial Data is based on standards, technology and architecture that will lead towards distributed GIS wherein updated data is accessed over the internet directly from data owners in the form of WMS and WFS. Distributed infrastructure for Spatial Information supports the easier access to geo-data and organize better products and services, which in turn will support increased use of geo-information, necessary to acquire process, store, distribute and improve utilisation of geospatial data. Direct integration with other datasets may be organised in an increasingly flexible way. It is intended to extend the current routine collection of digital data and create new information databases, to generate catalogues of Remote Sensing Data and filling in of the Digital Archive, browse images and ordering capability via the web, which are tailored to support environmental monitoring and natural disaster mitigation. Important here is interoperability of data and services. We are facing a transition: from data interoperability to service interoperability. Efficient organization of information resources and open access to spatially distributed experiment data are founded on the Web technology (Access to data including data search and request; Access distributed metadata catalogues, which may be: distributed catalogues themselves based on different application profiles or protocols). Appropriate metadata

management systems are built to provide for the collection and distribution of experiment data and thematic processing results; while the archive is linked to the regional centers of environmental monitoring via Internet. An important element is the elaboration of interface, archiving and network data exchange structures. This calls for the development of search engines and a remote interactive access regime for external users via Internet to catalogues of experiment data and processing results and the realization of the on-line access mode. The principle of distributed data processing gains an ever-growing importance for satellite monitoring. At the same time, remote users get access through Internet to satellite data obtained by the regional monitoring centers. Further development of Remote Sensing satellite data electronic archives implies introduction of automatic user-oriented information selection as a result of enhanced information models of electronic collections coupled to upgraded techniques of user interaction with the archive's information system.

* Работа проводится при поддержке РФФИ (гранты №№ 08-07-00002, 08-07-90401-Укр_а).