

Интеграция поставщиков спутниковых данных и услуг

А. И. Алексанин, М. Г. Алексанина, П. В. Бабяк, В. С. Ерёменко

*Институт автоматики и процессов управления ДВО РАН
Владивосток, 690041, Россия
E-mail: aleks@iacp.dvo.ru*

Рассматриваются проблемы создания информационной системы (ИС) спутникового центра с целью наиболее полного удовлетворения запросов потребителей. Необходимое условие решения поставленной задачи заключается во взаимодействии ИС с другими аналогичными системами. Глобальными информационными системами, объединяющими различных поставщиков данных, являются EOSDIS NASA и сервис-ориентированная платформа SSE (Service Support Environment) Европейского космического агентства. Расширение функциональных возможностей этих систем является актуальной задачей. С целью её решения «Роскосмосом» разрабатывается Государственная информационная система оперативной поставки данных дистанционного зондирования Земли (ГИС ОПД ДЗЗ). Эта система предназначена для интеграции различных поставщиков данных и сервисов их обработки. Ключевой вопрос при проектировании государственной системы состоит в минимизации затрат на её создание как со стороны «Роскосмоса», так и со стороны владельцев интегрируемых внешних информационных систем. Спутниковый центр ДВО РАН участвовал в проекте SSE и получил соответствующий опыт. Рассматриваются проблемы организации обмена данными и услугами между поставщиками спутниковой информации и возможные пути их решения.

Ключевые слова: информационная система, спутниковые данные, поставка данных и услуг

Одобрена к печати: 17.05.2019
DOI: 10.21046/2070-7401-2019-16-3-288-300

Введение

Современное развитие методов и средств дистанционного зондирования Земли из космоса делает возможным создание эффективной системы мониторинга различных процессов и явлений окружающей среды в режиме реального времени. Вместе с тем наблюдается низкий уровень первичной и тематической обработки данных дистанционного зондирования (ДДЗ), сложность доступа к имеющимся технологиям и накопленным спутниковым данным, а также существуют проблемы обработки этих данных и их верификации.

Разнообразие и сложность технологий обработки данных часто не позволяют пользователю освоить их в полном объёме в рамках одного спутникового центра. Даже крупные центры не располагают всем спектром необходимых данных и продуктов их обработки. Требуется кооперация и взаимодействие с другими центрами. Как правило, в зависимости от специфики задачи пользователь имеет собственные технологии обработки спутниковых данных, которые требуется встраивать в существующие схемы. Намечившаяся тенденция перехода к распределённой обработке данных через глобальные информационные системы (Лупян и др., 2004; Шокин и др., 2007) видится логичным решением этих проблем.

Сложившимся подходом к предоставлению спутниковых данных потребителю является создание каталогов и их интеграция в единую систему. В результате у пользователя появляется возможность получать данные различных организаций через единую точку доступа. Крупнейшими независимо развивающимися глобальными информационными системами, предоставляющими доступ к каталогам спутников данных, являются EOSDIS (Earth Observing System Data Information System) NASA (Kobler, Berbert, 1991) и eoPortal ESA Европейского космического агентства (SSE eoPortal, <https://wiki.services.eoportal.org>). В последние годы происходит смена основного направления развития этих систем от постав-

ки данных к поставке услуг в форме веб-служб. Так, система ECHO (Earth Observing System ClearingHouse) (Mitchell et al., 2009) и её модернизированная версия CMR (Common Metadata Repository, <https://earthdata.nasa.gov/about/science-system-description/eosdis-components/common-metadata-repository>) являются частью проекта NASA EOSDIS и предоставляют основу, позволяющую системам поставки данных наблюдения за Землёй и сервисов работать совместно. Современным этапом развития Европейской распределённой инфраструктуры доступа к спутниковой информации является среда SSE (Service Support Environment, <http://wiki.services.eoportal.org/tiki-index.php?page=Service+Support+Environment>) и основанный на ней проект НМА (Heterogeneous Mission Accessibility), обладающие аналогичным назначением. Последний проект был хорошо проработан теоретически и охватывал все ступени обработки и обмена данными различных поставщиков информации с последующей поставкой информации в ГИС-системы, работающие по OGC-стандартам. Тем не менее после остановки работы серверов проект и его информационная система были закрыты.

Данные ДЗЗ являются составной частью ресурсов многих ГИС-систем, но по проблеме более тесного взаимодействия этих двух типов систем нет достаточных проработок ни в теоретическом, ни в практическом плане (Шокин, Потапов, 2015). В этой же работе показаны современные достижения в области гибкого взаимодействия ГИС-систем по организации распределённой обработки с возможностью обмена данными и использования общего пула алгоритмов обработки. Системы приёма, обработки и поставки спутниковых данных такими возможностями не обладают. Начало работ по созданию «Роскосмосом» Государственной информационной системы оперативной поставки данных дистанционного зондирования Земли (ГИС ОПД ДЗЗ) поднимает вопросы её организации и расширения функций, отсутствующие сейчас в системах США и Европы. Встаёт также вопрос о требованиях к системам поставщиков и обработчиков спутниковой информации, выполнение которых необходимо для интеграции в ГИС ОПД ДЗЗ. В работе рассматриваются варианты решения этой проблемы с минимизацией затрат как со стороны поставщиков, так и со стороны разработчиков системы. Возможные решения показаны на примере организации информационной системы Центра коллективного пользования Регионального спутникового мониторинга окружающей среды ДВО РАН, который был интегрирован в европейскую систему, разрабатываемую в рамках проекта НМА.

Возможные схемы построения ГИС ОПД ДЗЗ и требования к поставщикам данных и услуг

Базовые функции ГИС ОПД ДЗЗ

В создаваемой в рамках ОКР «Потребитель» информационной системе ГИС ОПД ДЗЗ желательно иметь основные функции существующих глобальных систем и некоторые дополнительные, позволяющие расширить возможности таких систем и организовать взаимодействие с прикладными геоинформационными системами. Наиболее детально идеология ИС ДЗЗ была развита при разработке портала EoPortal Европейского космического агентства и создании средств интеграции различных поставщиков данных и услуг. К последним относится SSE-Toolbox, позволявший при его установке единообразно генерировать и поставлять метаданные и данные по запросу. Среда SSE обеспечивала:

- создание инфраструктуры, позволяющей организовать взаимодействие между поставщиками и потребителями;
- сохранение базовых и сквозных служб на стороне поставщика;
- лёгкое подключение/отключение сервисов к/от среды;
- сочленение простых сервисов в составные;
- поддержку сервисов с подпиской;
- простой поиск и доступ к требуемым сервисам с возможностью наблюдения процесса вплоть до его завершения;
- интеграцию сервисов из различных областей знаний.

Центральным узлом среды был сервер SSE Portal (<http://services.eoportal.org>), предоставлявший пользователю веб-интерфейс для работы с системой. Пользователь работал с инструментами портала, не задумываясь о расположении поставщика, ресурсы которого он использует. Для выдачи результатов могли быть использованы сервисы WCS и WFS, протокол HTTP, FTP-доступ к файлам данных или SOAP-сообщение. На форматы файлов, доступных по протоколам FTP и HTTP, специальных ограничений не накладывалось.

Основной особенностью, препятствующей глубокой интеграции ИС спутниковых центров, является технология обработки поступающей информации. Получение тематического продукта обычно проходит в три этапа:

1. Предварительно обработанные данные помещаются в контейнер формата HDF5. Формат позволяет хранить флажки качества обработки каждого пикселя мультиспектрального изображения. Например, если мы поставляем данные для построения карт температуры поверхности океана, то флажки говорят о точности расчёта радиационной температуры, наличии облачности, солнечного зайчика, сбоев и др.
2. Для построения конкретных геофизических полей собирается сопутствующая информация и помещается в контейнер формата netCDF. Например, радиационная температура океана используется для построения четырёх типов полей температуры: температуры поверхностной плёнки, подповерхностной температуры, температуры на глубине 1 м, температуры на больших глубинах. Финальные температуры могут отличаться вплоть до 10 °С. Для их расчёта для каждого пикселя помещается информация о приводном ветре, аэрозоле над морем, озоне и др.
3. На основе данных формата netCDF строятся конкретные геофизические поля и переводятся в привычный для ГИС формат, например GeoTiff.

Приведённые базовые функции ИС ДЗЗ были хорошо разобраны с точки зрения стандартов обработки данных. Предпринимались попытки расширить функционал ИС в направлении создания единого пула методов тематической обработки данных. Была проведена организация сбора универсальных сервисов обработки данных от различных поставщиков сервисов, не привязанных к специфике исходных данных и информационным системам поставщика данных. Таким примером является GPOD (Grid Processing on Demand, <http://gpod.eo.esa.int>) — это европейский проект сбора сервисов обработки в единой информационной системе, использующий архив Европейского Космического Агентства в качестве источника данных для сервисов обработки. Разработчики этой системы предлагают владельцам сервисов обработки данных ДЗЗ размещать свои алгоритмы обработки на предоставляемых проектом виртуальных машинах, после чего в течение нескольких недель в результате взаимодействия разработчиков системы и поставщика сервиса обеспечивается доступ к размещённому на виртуальной машине сервису через интерфейс информационной системы. Для реализации этой схемы требуется переделка предоставляемых сервисов, и обычно это легко осуществить только для простых сервисов. Сама система часто недоступна по техническим причинам. Наиболее близкой по функциональности системой NASA EOSDIS является проект Giovanni (GES-DISC Interactive Online Visualization and analysis Infrastructure) (Berrick et al., 2009), предназначенный для анализа и визуализации данных каталога NASA в онлайн-режиме без необходимости скачивания исходных данных на компьютер пользователя. Однако этот проект не подразумевает возможности загрузки собственных пользовательских алгоритмов обработки и анализа данных для дальнейшего их использования. Приведённые решения не позволяют легко наращивать пул алгоритмов обработки для произвольных данных, которые можно получить в глобальной ИС.

В рамках глобальной ИС содержится огромное количество данных (BIG DATA), с которыми желательно организовать эффективную работу. В силу специфики обработки данных в ИС это сделать сложно. Особенно при условии, что алгоритмы первичной и тематической обработки данных непрерывно улучшаются, а пересчёт такого объёма данных требует огромных вычислительных затрат. Оригинальный способ решения этой проблемы создан в ИКИ РАН (Прошин и др., 2016; Kashnitskii et al., 2017). Разработана архитектура хранения и алго-

ритмы доступа к данным, обеспечивающие возможность формирования информационных продуктов «на лету», что даёт возможность в онлайн-режиме работать с продуктами, которые в системе представлены лишь виртуально. В распределённых архивах в основном хранятся данные, прошедшие первичную обработку. При изменении алгоритмов первичной обработки строящиеся по хранящимся данным продукты проходят коррекцию, обычно сводящуюся к преобразованию данных на основе простых функций пересчёта. Это снимает проблему постоянного пересчёта архивов накопленной информации. Опыт ИКИ РАН показывает, что требуется разработка стандартов для поставщиков данных, которые могли бы использовать возможности работы с большими данными. Это может стать одним из вариантов решения проблемы.

При построении глобальной системы необходимо обеспечить максимальную по функциям интеграцию сторонних поставщиков данных и услуг при минимизации затрат на создание требуемого программного обеспечения как со стороны поставщиков, так и со стороны «Роскосмоса». Также должно обеспечиваться взаимодействие на уровне обмена данными и метаданными с крупными существующими внешними системами, по возможности с обходом их пользовательского интерфейса. Для обеспечения взаимодействия между участниками ГИС ОПД ДЗЗ наиболее целесообразно применение одного из двух подходов. Выбор подхода определяет трудозатраты на создание ИС. Существующее в настоящее время программное обеспечение (ПО), предоставляющее сервисы, которые основаны на международных стандартах работы с пространственными данными, разработанных Открытым геопространственным консорциумом (Open Geospatial Consortium — OGC), позволяет упростить реализацию взаимодействия поставщиков данных и сервисов обработки с глобальной информационной системой.

Первый подход предполагает взаимодействие между всеми участниками напрямую через ГИС «Потребитель» (рис. 1).

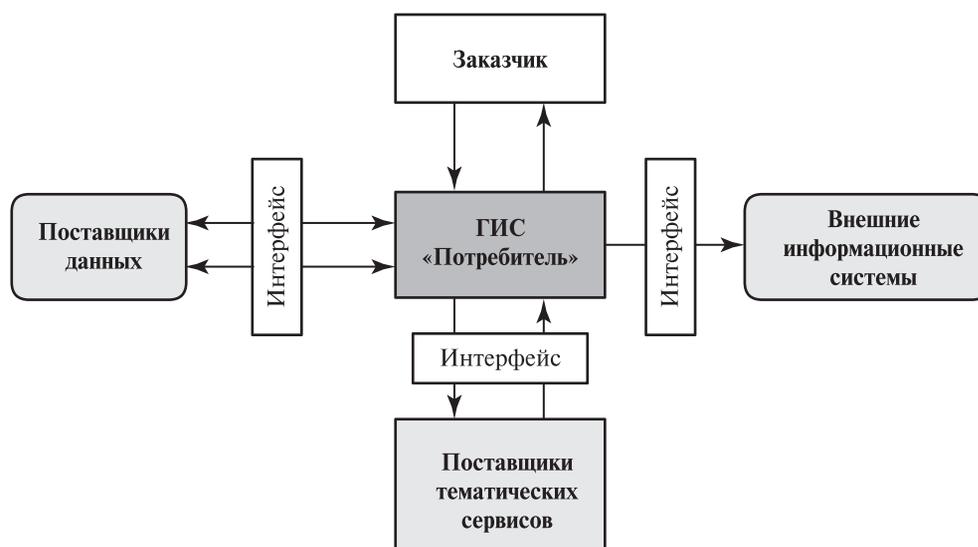


Рис. 1. Схема взаимодействия участников ГИС ОПД ДЗЗ по первому подходу

Полное описание работы предложенного подхода выглядит следующим образом:

1. Заказчик обращается в ГИС «Потребитель» и оставляет заявку на предоставление услуги.
2. ГИС «Потребитель» обращается к поставщикам тематических сервисов и вызывает соответствующую процедуру для запуска сервиса.
3. Поставщики тематических сервисов обращаются к ГИС «Потребитель» с запросом на поиск и получение необходимых для работы сервиса данных.

4. ГИС «Потребитель» производит поиск необходимых данных по всем доступным поставщикам данных и в случае, если формат данных не соответствует требуемому, производит преобразование формата (если это возможно), после чего данные передаются поставщикам тематических сервисов.
5. После получения результата работы сервиса поставщики тематических сервисов предоставляют доступ к полученной информации с использованием сервиса каталогизации пространственных данных OGC CSW (Catalogue Service for the Web), фактически представляя собой поставщика данных в рамках ГИС «Потребитель».
6. Для предоставления доступа к результатам работы тематических сервисов сторонним системам необходимо обеспечить ряд наиболее востребованных интерфейсов и форматов предоставления данных, таких как веб-сервисы OGC WMS (Web Map Service), OGC WFS (Web Feature Service), OGC WCS (Web Coverage Service) и общепринятые форматы для векторных и растровых данных (Shapefile, GeoTIFF, netCDF, HDF), а также доступ к сопроводительной информации по протоколам HTTP или HTTPS.

Для реализации данного подхода необходимо:

1. Разработать программное решение, предоставляющее в ГИС «Потребитель» унифицированный интерфейс поиска метаданных у различных поставщиков данных. Для реализации этой задачи требуется написать отдельный модуль поиска метаданных для каждого поставщика данных либо организовать доступ к метаданным на стороне поставщика в рамках каталога пространственных данных (OGC CSW 2.0.2) в формате ISO 19115. Для этого поставщик данных должен развернуть на своей стороне программные средства, реализующие сервис каталога, например программный пакет с открытым исходным кодом GeoNetwork.
2. Разработать сервис преобразования типов и форматов пространственных данных на стороне ГИС «Потребитель», а также систему хранения и предоставления заказанных данных заказчику/поставщику услуг.
3. Организовать систему хранения и передачи результатов работы тематических сервисов в виде каталога пространственных данных на базе интерфейса OGC CSW 2.0.2 с метаданными в формате ISO 19115 на стороне поставщиков тематических сервисов, аналогично поставщику данных.
4. Разработать интерфейс, позволяющий пользователям оставлять заявки на выполнение задач для поставщиков тематических сервисов.

Второй подход предполагает перенос функций обработки и поставки данных на сторону поставщика услуг/данных, минимизирует перекачки и уменьшает нагрузки на ГИС «Потребитель» (рис. 2, см. с. 293).

Полное описание работы предложенного подхода выглядит следующим образом:

1. Заказчик обращается в ГИС «Потребитель» и оставляет заявку на предоставление услуги.
2. ГИС «Потребитель» производит поиск данных, необходимых для работы сервиса, и передаёт ссылки на данные вместе с дополнительной информацией из заявки поставщикам сервиса.
3. Поставщики тематических сервисов загружают себе необходимые данные и преобразуют их в нужный формат в случае необходимости.
4. После получения результата работы сервиса поставщики тематических сервисов предоставляют доступ к полученной информации с использованием сервиса каталогизации пространственных данных (OGC CSW).
5. Для предоставления доступа к результатам работы тематических сервисов сторонним системам необходимо обеспечить ряд наиболее востребованных интерфейсов и форматов предоставления данных, таких как веб-сервисы (OGC WMS, OGC WFS, OGC WCS) и общепринятые форматы для векторных и растровых данных (Shapefile, GeoTIFF, netCDF, HDF), а также доступ к сопроводительной информации по протоколам HTTP или HTTPS.

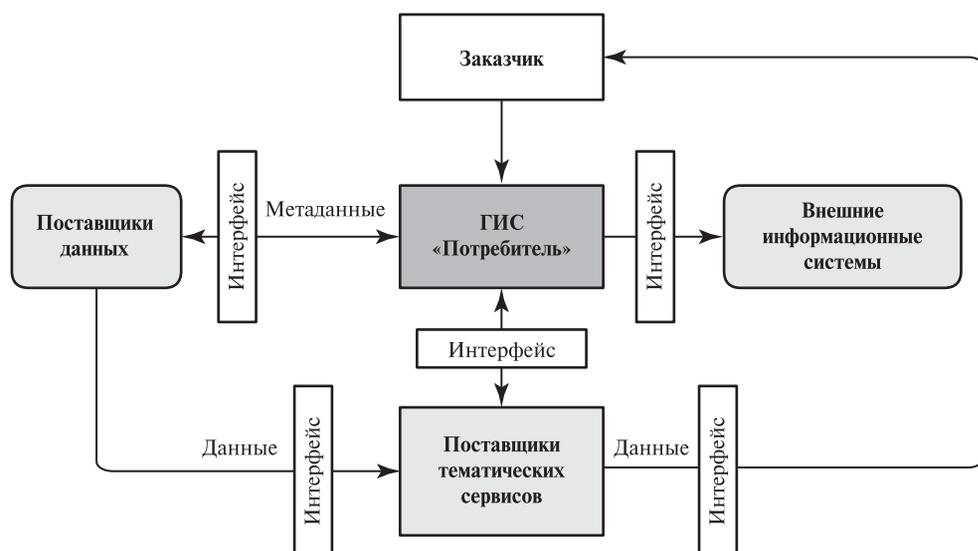


Рис. 2. Схема взаимодействия участников ГИС ОПД ДЗЗ по второму подходу

Для реализации данного подхода необходимо:

1. Разработать программное решение, предоставляющее в ГИС «Потребитель» унифицированный интерфейс поиска метаданных у различных поставщиков данных. Для реализации этой задачи требуется написать отдельный модуль поиска метаданных для каждого поставщика данных либо организовать доступ к метаданным на стороне поставщика в формате каталога пространственных данных (OGC CSW 2.0.2) с метаданными в формате ISO 19115.
2. Разработать сервис преобразования типов и форматов пространственных данных на стороне поставщиков.
3. Организовать систему хранения и передачи результатов работы тематических сервисов в виде каталога пространственных данных на базе интерфейса OGC CSW 2.0.2 с метаданными в формате ISO 19115 на стороне поставщиков тематических сервисов.
4. Разработать интерфейс, позволяющий пользователям оставлять заявки на выполнение задач для поставщиков тематических сервисов.

Преимущество второго подхода заключается в отсутствии необходимости организации системы хранения и передачи данных, полученных в процессе преобразования типов и форматов данных для передачи их поставщикам тематических сервисов, на стороне ГИС «Потребитель». Тем самым ГИС «Потребитель» будет «связывать» между собой поставщиков данных и поставщиков тематических сервисов напрямую, а не выступать посредником между ними, что существенно повысит скорость работы системы и уменьшит трудозатраты на разработку модулей для неё. Но потребуются создание унифицированных средств, передающихся поставщикам и работающих на их стороне.

Средства и возможности Спутникового центра ДВО РАН в области интеграции во внешние системы

Информационная система Спутникового центра ДВО РАН развивалась на принципах возможной интеграции во внешние глобальные системы. Ресурсы Центра были интегрированы в течение нескольких лет в информационную систему SSE Portal через SSE-Toolbox. Отрабатывались технологии не только обмена данными и метаданными. Сервис SSE-Toolbox позволял выбирать алгоритмы обработки данных, хотя в списке официальных функций это не значилось. Разработанное в Центре программное обеспечение позволяло интегрировать

в систему SSE Portal произвольного поставщика данных и сервисов обработки данных в течение недели вместе с его вычислительными ресурсами. Для этого была создана Платформа интеграции гетерогенных ресурсов, обеспечивающая возможность поиска, получения и обработки данных по технологии, управляемой пользователем (рис. 3). Данная платформа построена с использованием принципов слабосвязанной архитектуры, где каждый модуль представляет собой отдельную программную реализацию, независимую от других модулей платформы и поддерживающую определённый интерфейс взаимодействия. Платформа также позволяет производить поиск и получать данные из внешних информационных систем, основанных на стандарте OGC CSW. Для обеспечения доступа к сервисам платформы разрабатывается возможность их вызова посредством интерфейса OGC WPS (Web Processing Service). Использование стандартного интерфейса также позволяет интегрировать удалённые сервисы в глобальные информационные системы.

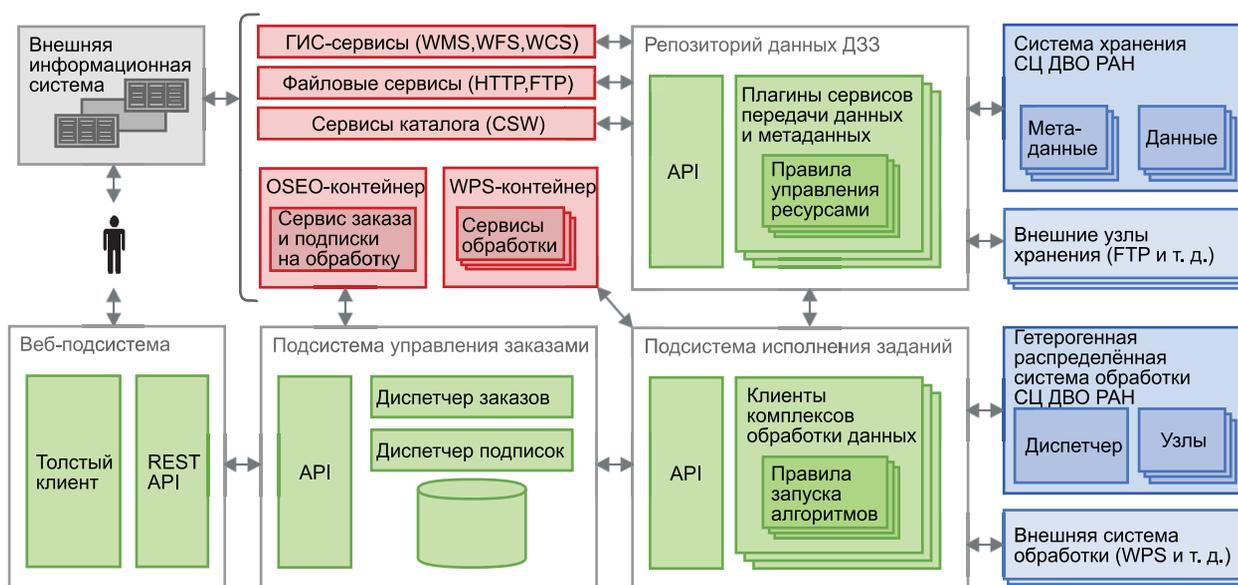


Рис. 3. Схема платформы для интеграции гетерогенных ресурсов, обеспечивающая возможность поиска, получения и обработки данных по технологии, управляемой пользователем

В связи со значительным ростом числа вычислительных алгоритмов и областей применения продуктов обработки спутниковых данных в последние годы на первый план выходит решение задач, связанных с долговременным хранением данных (за весь период приёма), интеграцией спутниковых центров в единые территориально-распределённые информационные системы, увеличением потока данных за счёт расширения списка принимаемых спутников. Обмен данными в распределённой среде, наряду с обеспечением информационно-поисковых систем (каталогов) различных сервисов по управлению обработкой данных и разграничением доступа, является одним из важнейших вопросов при построении ИС.

Для обеспечения функционирования ИС и передачи больших объёмов данных между различными частями распределённой системы Центра было решено использовать iRODS — интегрированную систему данных, основанную на правилах. iRODS реализует такое перспективное направление в области распределённой поставки и обмена данными, как архитектура data grid, которая предоставляет пользователям возможность доступа, изменения и передачи больших объёмов географически распределённых данных для различных прикладных задач. Обеспечивается это посредством множества промежуточных приложений (middleware) и сервисов, которые объединяют данные и ресурсы из нескольких административных доменов, а затем предоставляют их пользователям по запросу. При этом данные в data grid могут располагаться как на одном ресурсе, так и на множестве разделённых ресурсов, где каждый из них может быть участником административного домена, регулируемого своим собственным

набором правил разграничения доступа. Кроме того, несколько реплик данных могут быть распределены по сети за пределами их первоначального административного домена, но ограничения безопасности на исходные данные должны одинаково применяться и к репликам.

Система управления алгоритмами обработки данных

Для реализации обработки принимаемых и внешних данных используется распределённая система обработки данных (Бабяк и др., 2011, 2012). Вся система построена на grid-вычислениях, где каждый вычислительный узел (в том числе задействованный для вычислений многопроцессорный кластер) является независимым компьютером, связанным с другими посредством сети. Запуск и контроль выполнения задач на вычислительных узлах осуществляется выделенным управляющим сервером, содержащим формализованное описание узлов и выполняемых на них задач. Все соединения с сервером управления обработкой, а также обрабатывающими машинами происходят по протоколу SSH, что обеспечивает простоту взаимодействия и высокий уровень безопасности устанавливаемых соединений.

Структура управляющего сервера представлена на *рис. 4*. В базе «БД объектов системы» хранятся: формализованное описание задач для запуска в распределённой системе обработки и счётных машин, на которых эти задачи запускаются; триггеры, контролирующие последовательность выполнения задач; разграничения прав доступа к запускаемым задачам. «Брокер сообщений» отвечает за приём и отправку сообщений от запуска и до завершения различных задач внутри системы. «Диспетчер» отвечает за запуск задач на «Обрабатывающих компьютерах».

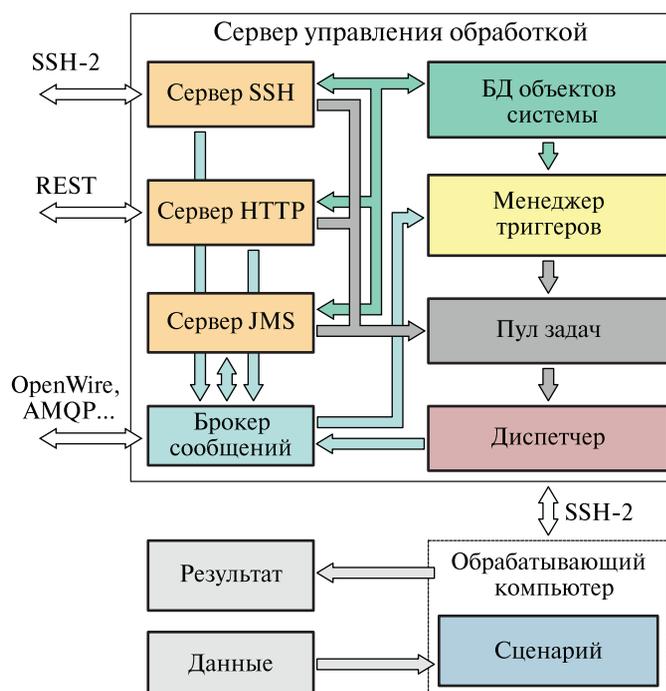


Рис. 4. Общая структура работы распределённой вычислительной системы

При срабатывании установленных триггеров (чаще всего — по появлению новых данных) происходит запуск схемы, обрабатывающей данную задачу. Схема является формальным описанием того, как и на каких обрабатывающих машинах должна запускаться та или иная задача. По завершении задачи «Диспетчер» посылает соответствующее сообщение, в котором присутствует некоторая описательная информация о задаче.

Такая автоматическая распределённая система позволяет достичь нескольких целей, а именно: полностью автоматический запуск необходимых задач с возможностью их ручного

запуска; оперативность обработки поступающих данных; автоматическое распределение вычислительных ресурсов между задачами; устойчивость к выходу из строя тех или иных обрабатывающих машин; все поставленные в очередь задачи в конечном итоге будут обработаны в порядке их поступления.

Апробация подходов на макете ГИС «Потребитель»

В рамках выполнения начального этапа ОКР «Потребитель» был разработан макет системы для отработки подходов к построению ГИС ОПД ДЗЗ. Он позволяет в простейшей форме осуществлять поиск и получение данных у различных поставщиков и запускать сервисы по обработке выбранных данных. Макет реализован на базе фреймворка Java Spring Boot и был апробирован на тестовых примерах, обеспечивающих связь с различными информационными системами. В экспериментах обеспечивалось взаимодействие ИС следующих организаций: АО НИИ ТП «Роскосмос», НЦ ОМЗ «Роскосмос», ООО Центр инновационных технологий, ИТЦ СканЭкс, ИАПУ ДВО РАН. На тестовых примерах обрабатывались процедуры:

- запрос на поиск и поставку нужных метаданных и данных;
- получение данных от одного поставщика и их обработка алгоритмами другого поставщика с использованием вычислительных средств обработчика;
- заказ на обработку данных выбранными алгоритмами с настройкой параметров этих алгоритмов.

Основные узлы макета:

- каталог метаданных с возможностью поиска адресов данных;
- средства заказа данных с автоматическим поиском и поставкой данных;
- средства авторизации доступа;
- репозиторий, позволяющий по идентификатору из метаданных находить реальные адреса данных или услуг;
- интерфейсы для интеграции в систему сторонних информационных ресурсов.

В качестве системы, обеспечивающей аутентификацию и авторизацию пользователей в системе ГИС ОПД ДЗЗ, использовалась система геопортала «Роскосмос» для осуществления доступа к данным и сервисам обработки. Здесь необходимо разработать модуль для осуществления взаимодействия системы разграничения доступа геопортала «Роскосмос» с ГИС «Потребитель».

Доступ к данным «Роскосмоса» и ведение каталога данных реализуются через ЕТРИС ДЗЗ (Носенко и др., 2010). Для обеспечения доступа к данным ведомств и организаций — поставщиков данных использовались сервисы для организации каталогов метаданных. Этот вид сервисов позволяет осуществлять быстрый поиск и получение информации о данных, а также сами данные, если информация о способе их получения представлена в метаданных. Для реализации указанного подхода предлагается применять сервисы каталога по стандартам OGC Catalogue Service for the Web (CSW 2.0.2, CSW 3.0) с использованием профилей метаданных в формате ISO 19115 / ISO 19139. Рекомендуются использовать существующее открытое программное обеспечение, поддерживающее работу по вышеописанным стандартам, а именно:

- GeoNetwork;
- ESRI GeoPortal Server.

Профили метаданных ЕТРИС ДЗЗ желательно поменять — перейти на латиницу при задании служебных полей описаний метаданных (иначе сложно будет единообразно интегрировать в ГИС ОПД ДЗЗ сторонние ресурсы, особенно международные). Для этого нужно создать программное обеспечение по отображению одного формата профиля метаданных в другой.

Метаданные должны содержать информацию о получении данных в виде прямой ссылки на них либо в виде ссылки на сервис, предоставляющий эти данные по запросу. Для дальней-

шего доступа к этим данным на стороне ГИС ОПД ДЗЗ необходимо будет создать ПО для поиска и получения списка метаданных от каждого поставщика данных.

Существует два подхода по дальнейшей организации работы с поставщиками данных в случае поискового запроса на данные:

- отправка запроса к внешнему сервису каталога поставщика данных;
- отправка запроса к локальному сервису каталога системы ГИС ОПД ДЗЗ.

В случае использования первого подхода ГИС ОПД ДЗЗ будет возвращать наиболее актуальную информацию от каждого поставщика данных. Этот подход не требует установки дополнительного ПО и пространства для хранения метаданных, однако скорость его работы зависит от каналов передачи данных между ГИС ОПД ДЗЗ и поставщиками данных.

При использовании второго подхода все метаданные от всех поставщиков данных периодически скачиваются в локальный каталог. Таким образом, скорость работы поиска будет выше, однако на стороне ГИС ОПД ДЗЗ требуется установка дополнительного ПО, реализующего сервис каталога с хранением метаданных и их синхронизацию, а актуальность данных в локальном каталоге будет зависеть от частоты их обновления.

Хранение данных

Для хранения данных на стороне поставщика рекомендуется использовать сервис репозитория. В качестве программной реализации сервиса может быть взято любое ПО, позволяющее получить доступ к данным с использованием идентификаторов по протоколам HTTP/HTTPS. В качестве репозитория можно использовать встроенные средства хранения данных GeoNetwork или стороннее ПО, например SKAN, GeoNode, DSpace, Omeka и др.

Доступ к сервисам обработки

Для организации доступа к сервисам обработки всем поставщикам необходимо реализовать доступ к своим сервисам, используя единый интерфейс запуска обработки. В качестве универсального интерфейса для запуска и контроля процесса обработки рекомендуется использовать OGC Web Processing Service (OGC WPS 1.0.0). Данный интерфейс поддерживает возможность получения списка существующих у поставщика сервисов обработки, а также описание, аргументы и параметры запуска каждого сервиса. Для этого поставщикам предлагается использовать открытое программное обеспечение, предоставляющее возможность создания собственных WPS-сервисов на основе готовых программных решений, например:

- GeoServer,
- Deegree,
- 52 North,
- pyWPS.

Интерфейс OGC WPS реализован поверх протокола HTTP (HTTPS), таким образом, в нём предусмотрена возможность разграничения доступа и шифрования соединения через SSL. Со стороны ГИС ОПД ДЗЗ необходимо реализовать каталог сервисов и WPS-клиент, взаимодействующий со всеми поставщиками сервисов для обработки пространственно-привязанных данных.

Интерфейс доступа к сервисам заказа и поиску данных

В качестве интерфейса доступа к сервисам заказа системы ГИС ОПД ДЗЗ можно также использовать существующий интерфейс заказа геопортала «Роскосмоса». В связи с использованием единых стандартов для доступа к данным и сервисам обработки организация доступа

из внешних информационных систем потребует минимальных затрат. Таким образом, доступ к данным и сервисам можно будет организовать как непосредственным использованием интерфейса ГИС ОПД ДЗЗ, так и из любой другой внешней информационной системы, реализовав соответствующие программные компоненты.

Заключение

Интеграция информационных систем центров приёма и обработки спутниковых данных дистанционного зондирования Земли является основой для максимального удовлетворения запросов пользователей. Существующие подходы развиты только для обмена данными и метаданными. Попытки объединения методов обработки данных различных поставщиков информационных услуг в этой области не демонстрируют удобного универсального решения проблемы. Не разработаны также стандарты и средства работы с огромными накопленными архивами спутниковой информации. В то же время некоторые решения центров позволяют сделать заключение о возможности создания необходимых схем и средств. Начавшаяся разработка Государственной системы оперативной поставки данных дистанционного зондирования Земли (ГИС ОПД ДЗЗ) делает актуальной выработку схем и стандартов для интеграции информационных систем разнородных поставщиков спутниковых данных и услуг. Созданный в рамках проекта макет информационной системы продемонстрировал возможности взаимодействия ИС следующих организаций: АО НИИ ТП «Роскосмоса», НЦ ОМЗ «Роскосмоса», ООО Центр инновационных технологий, ИТЦ СканЭкс, ИАПУ ДВО РАН. На тестовых примерах обрабатывались процедуры поставки метаданных, данных, заказа на обработку данных по схеме, использующей одновременно данные, алгоритмы и вычислительные ресурсы нескольких поставщиков.

Работа поддержана РФФИ (проект № 17-05-41152РГО_a).

Литература

1. *Бабяк П. В., Недолужко И. В., Тарасов Г. В.* Инфраструктура приёма, распределённой обработки и поставки спутниковых данных ЦКП Регионального Спутникового Мониторинга ДВО РАН // 9-я Всерос. открытая конф. «Современные проблемы дистанционного зондирования Земли из космоса»: сб. тез. конф. Москва: ИКИ РАН, 2011. С. 73.
2. *Бабяк П. В., Недолужко И. В., Фомин Е. В.* Подход к предоставлению услуг по обработке спутниковых данных в Центре коллективного пользования регионального спутникового мониторинга окружающей среды ДВО РАН // Современные проблемы дистанционного зондирования Земли из космоса. 2012. Т. 3. № 9. С. 324–331.
3. *Лузян Е. А., Мазуров А. А., Назиров Р. Р., Прошин А. А., Флитман Е. В.* Технология построения автоматизированных информационных систем сбора, обработки, хранения и распространения спутниковых данных для решения научных и прикладных задач // Современные проблемы дистанционного зондирования Земли из космоса: Физические основы, методы и технологии мониторинга окружающей среды, потенциально опасных явлений и объектов. М.: Полиграф сервис, 2004. С. 81–88.
4. *Носенко Ю. И., Новиков М. В., Заичко В. А., Ромашкин В. В., Лошкарев П. А.* Единая территориально-распределённая информационная система ДЗЗ — проблемы, решения, перспективы (часть 2) // Геоматика. 2010. № 4. С. 31–36.
5. *Прошин А. А., Лузян Е. А., Балашов И. В., Кашицкий А. В., Буццев М. А.* Создание унифицированной системы ведения архивов спутниковых данных, предназначенной для построения современных систем дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2016. Т. 13. № 3. С. 9–27.
6. *Шокин Ю. И., Потапов В. П.* ГИС сегодня: состояние, перспективы, решения // Вычислительные технологии. 2015. Т. 20. № 5. С. 175–213.
7. *Шокин Ю. И., Жижимов О. Л., Пестунов И. А., Синявский Ю. Н., Смирнов В. В.* Распределённая информационно-аналитическая система для поиска, обработки и анализа пространственных данных // Вычислительные технологии. 2007. Т. 12. Спецвып. 3. С. 108–115

8. *Kashnitskii A. V., Loupian E. A., Balashov I. V., Konstantinova A. M.* Technology for Designing Tools for the Process and Analysis of Data from Very Large Scale Distributed Satellite Archives // *Atmospheric and Oceanic Optics*. 2017. V. 30. No. 1. P. 84–88.
9. *Kobler B., Berbert J.* NASA Earth Observing System Data Information System (EOSDIS) // *Digest of Papers Eleventh IEEE Symp. Mass Storage Systems*. Monterey, USA, 1991. P. 18–19.
10. *Mitchell A., Ramapriyan H., Lowe D.* Evolution of web services in EOSDIS — Search and order metadata registry (ECHO) // *2009 IEEE Intern. Geoscience and Remote Sensing Symp.* Cape Town, 2009. V. 5. P. 371–374.

Integration of satellite data and service providers

A. I. Aleksanin, M. G. Aleksanina, P. V. Babyak, V. S. Eremenko

Institute of Automation and Control Processes FEB RAS, Vladivostok 690041, Russia
E-mail: aleks@iacp.dvo.ru

The problems of creating an information system of a satellite centre that best satisfies users' requirements have been considered. Information system interaction with other similar ones is a necessary condition for the task solution. The global information systems that have united different service-providers are EOSDIS/NASA and the service oriented platform SSE (Service Support Environment) of European Space Agency. The enlargement of functional abilities of such systems is an urgent task. The development by Roscosmos of the State Information System for Real Time Delivery of Remote Sensing Data (SIS RTD RSD) is an attempt to solve the task. SIS RTD RSD is intended for integration of different service providers. Minimization of expense for both Roscosmos and the owners of external information systems under integration is the key aim of the system design. The Satellite Centre of Far Eastern Branch of the Russian Academy of Sciences took part in the SSE project and acquired appropriate experience. The problems of data and service exchange between providers, and possible solutions are considered.

Keywords: information system, satellite data, data and service delivery

Accepted: 17.05.2019

DOI: 10.21046/2070-7401-2019-16-3-288-300

References

1. Babyak P. V., Nedoluzhko I. V., Tarasov G. V., *Infrastruktura priema, raspredelennoi obrabotki i postavki sputnikovyx dannykh TsKP Regional'nogo Sputnikovogo Monitoringa DVO RAN (Infrastructure of satellite data reception, distributed processing and delivery at the Center for Regional Satellite Monitoring of the FEB RAS), 9-ya Vserossiiskaya otkrytaya konferentsiya "Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa" (9th All-Russia Open Conf. "Current Problems in Remote Sensing of the Earth from Space")*, Book of Abstracts, Moscow: IKI RAN, 2011, p. 73.
2. Babyak P. V., Nedoluzhko I. V., Fomin E. V., *Podkhod k predostavleniyu uslug po obrabotke sputnikovyx dannykh v Tsentre kollektivnogo pol'zovaniya regional'nogo sputnikovogo monitoringa okruzhayushchei sredy DVO RAN (Approach to the provision of satellite data processing services in the Center of regional satellite monitoring of environment of the FEB RAS), Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2012, Vol. 3, No. 9, pp. 324–331.
3. Loupian E. A., Mazurov A. A., Nazirov R. R., Proshin A. A., Flitman E. V., *Tekhnologiya postroeniya avtomatizirovannykh informatsionnykh sistem sbora, obrabotki, khraneniya i rasprostraneniya sputnikovyx dannykh dlya resheniya nauchnykh i prikladnykh zadach (Development of Automated Information Systems for Scientific and Application Tasks), Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa: Fizicheskie osnovy, metody i tekhnologii monitoringa okruzhayushchei sredy, potentsial'no opasnykh yavlenii i ob"ektov*, Moscow: Poligraf servis, 2004, pp. 81–88.

4. Nosenko Yu. I., Novikov M. V., Zaichko V. A., Romashkin V. V., Loshkarev P. A., Edinaya territorial'no-raspredeleonnaya informatsionnaya sistema DZZ — problemy, resheniya, perspektivy (chast' 2) (Unified geographically distributed information system of remote sensing - problems, solutions, prospects (part 2)), *Geomatika*, 2010, No. 4, pp. 31–36.
5. Proshin A. A., Loupian E. A., Balashov I. V., Kashnitskii A. V., Burtsev M. A., Sozdanie unifitsirovannoi sistemy vedeniya arkhivov sputnikovykh dannykh, prednaznachennoi dlya postroeniya sovremennykh sistem distantsionnogo monitoringa (Unified satellite data archive management platform for remote monitoring systems development), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2016, Vol. 13, No. 3, pp. 9–27.
6. Shokin Yu. I., Potapov V. P., GIS segodnya: sostoyanie, perspektivy, resheniya (GIS today: state, prospects, solutions), *Vychislitel'nye tekhnologii*, 2015, Vol. 20, No. 5, pp. 175–213.
7. Shokin Yu. I., Zhizhimov O. L., Pestunov I. A., Sinyavskii Yu. N., Smirnov V. V., Raspredeleonnaya informatsionno-analiticheskaya sistema dlya poiska, obrabotki i analiza prostranstvennykh dannykh (Distributed information and analytical system for searching, processing and analysis of spatial data), *Vychislitel'nye tekhnologii*, 2007, Vol. 12, No. S3, pp. 108–115.
8. Kashnitskii A. V., Loupian E. A., Balashov I. V., Konstantinova A. M., Technology for Designing Tools for the Process and Analysis of Data from Very Large Scale Distributed Satellite Archives, *Atmospheric and Oceanic Optics*, 2017, Vol. 30, No. 1, pp. 84–88.
9. Kobler B., Berbert J. B., Earth Observing System Data Information System (EOSDIS), *Digest of Papers Eleventh IEEE Symp. Mass Storage Systems*, Monterey, USA, 1991, pp. 18–19.
10. Mitchell A., Ramapriyan H., Lowe D., Evolution of web services in EOSDIS — Search and order metadata registry (ECHO), *2009 IEEE Intern. Geoscience and Remote Sensing Symp.*, Cape Town, 2009, Vol. 5, pp. 371–374.