

## Система работы с распределенными архивами результатов обработки спутниковых данных центров приема НИЦ Планета

М.А. Бурцев<sup>1</sup>, В.Н. Антонов<sup>2</sup>, В.Ю. Ефремов<sup>1</sup>, А.В. Кашницкий<sup>1</sup>,  
Л.С. Крамарева<sup>2</sup>, Е.А. Лупян<sup>1</sup>, А.А. Мазуров<sup>1</sup>, А.М. Матвеев<sup>1</sup>,  
О.Е. Милехин<sup>2</sup>, А.А. Прошин<sup>1</sup>, В.И. Соловьев<sup>2</sup>

<sup>1</sup> Институт космических исследований Российской академии наук

117997 Москва, Профсоюзная 84/32

E-mail: evgeny@iki.rssi.ru;

<sup>2</sup> ФГБУ «НИЦ «Планета»

123242 Москва, Б.Предтеченский пер., д.7

E-mail: astmus@planet.iitp.ru

В настоящей работе описывается система работы с распределёнными архивами результатов обработки спутниковых данных центров ФГБУ «НИЦ «Планета». Приводятся основные задачи, решаемые системой, и предъявляемые к ней требования. Рассматривается общая архитектура системы и её основные составные блоки. Особое внимание в работе уделяется механизмам построения распределённых архивов спутниковых данных и способов доступа к ним. Также подробно рассматриваются структура и возможности картографического веб-интерфейса доступа к данным и их анализа. В заключение работы приводятся основные направления развития системы.

**Ключевые слова:** спутниковые данные, распределённые архивы, информационные системы, системы мониторинга, технологии построения автоматизированных систем, интерфейсы доступа к данным.

### Введение

Научно-исследовательский центр космической гидрометеорологии «Планета» (далее НИЦ «Планета») является ведущей организацией Российского агентства по гидрометеорологии и мониторингу окружающей среды (Росгидромет) по эксплуатации и развитию национальных космических систем гидрометеорологического, океанографического, гелиогеофизического мониторинга и мониторинга окружающей среды, а также по приему и обработке информации, поступающей с зарубежных спутниковых систем. НИЦ «Планета» обладает распределенной системой приема и обработки спутниковых данных, в состав которой входят Европейский (Москва, Обнинск, Долгопрудный), Сибирский (Новосибирск) и Дальневосточный (Хабаровск) центры. В этих центрах ежедневно принимается и обрабатывается информация, поступающая с российских и зарубежных космических аппаратов дистанционного зондирования (КА ДЗЗ), в том числе КА серии NOAA, TERRA, AQUA, «Метеор-М» №1, ЭЛЕКТРО-Л, Канопус-В и др. На базе полученной информации в центрах оперативно создается несколько сотен информационных продуктов как в интересах различных организаций Росгидромета, так и в интересах различных систем мониторинга, использующихся различными ведомствами и регионами.

Эффективная работа с большим числом информационных продуктов, постоянно использующихся сотнями потребителей, возможна только при обеспечении высокого уровня автоматизации процессов обработки, архивации и представления информации. Поэтому

в последние годы данному вопросу в НИЦ «Планета» уделяется очень большое внимание. В последние несколько лет в Центре также активно ведутся работы, направленные на создание единой унифицированной системы, обеспечивающей однотипную работу с распределенными архивами результатов обработки спутниковых данных, автоматически формируемых в центрах приема и обработки НИЦ «Планета» (Асмус и др., 2006, Прошин и др., 2004). Эти работы ведутся Центром совместно с Институтом космических исследований Российской академии наук (ИКИ РАН) при поддержке НИИ «Фотон» Рязанского государственного радиотехнического университета (РГРТУ) и других организаций

Настоящая работа посвящена описанию подходов к построению данной системы и ее основных возможностей. Система рассчитана на организацию максимальной автоматизации и унификации всех основных этапов жизненного цикла данных от их приёма, первичной и тематической обработки до предоставления конечным пользователям возможностей удобного доступа и анализа данных и результатов их обработки. В состав системы входят комплексы автоматической первичной и тематической обработки данных, распределённые архивы спутниковых данных и результатов их обработки, различные системы интерфейсов, обеспечивающих работу как с каталогами исходных данных, так и с различными информационными продуктами, полученными в результате обработки спутниковой информации. Основным инструментом, обеспечивающим доступ и анализ информации, полученной на основе спутниковых данных, является специализированный картографический web-интерфейс, обеспечивающий как работу с различными информационными продуктами, так и с различной сопутствующей информацией (метеоданные, различные картографические слои и т.д.). Система также обеспечивает оперативное представление информации в различные федеральные и региональные системы мониторинга

В работе рассматриваются общая архитектура системы и использованные для ее построения основные технологические решения. Описываются особенности реализации компонент системы. В работе также представлены основные возможности интерфейсов, обеспечивающих работу со спутниковыми данными. Особое внимание в докладе уделено вопросам обеспечения автоматизированного ведения распределённых архивов данных.

### **Основные задачи и возможности центров приёма «НИЦ «Планета»**

Территориально-распределенная система приема и обработки «НИЦ «Планета» по состоянию на ноябрь 2012 года обладала следующими характеристиками:

- система принимает и обрабатывает более 280 Гбайт/сутки;
- ежедневно получает данные с 16 зарубежных и отечественных КА ДЗЗ;
- ежедневно формирует более 100 информационных продуктов;
- обслуживает около 500 федеральных и региональных корпоративных потребителей (несколько тысяч пользователей);
- ежедневно формируются информационные продукты, покрывающие как отдельные районы России, так и крупные регионы Земли (территории Европы, Северной Евразии, полярные и экваториальные области и т.д.).

Блоки и комплексы приема и обработки данных работают в Европейском, Сибирском и Дальневосточном центрах, расположенных в Москве, Долгопрудном, Обнинске, Новосибирске и Хабаровске. Общая схема наземного комплекса приёма и обработки (НКПОР) данных НИЦ «Планета» приведена на рис. 1. (Асмус и др., 2012). Из неё видно, что на те-

кущий момент технические возможности Европейского, Сибирского и Дальневосточного центров НИЦ «Планета» во многом идентичны и обеспечивают приём данных с большого количества различных высокоинформационных зарубежных и отечественных космических аппаратов (КА), в том числе NOAA, TERRA, AQUA, MetOp, FY-1, Метеор-М, Канопус-В, Meteosat, MTSAT-2, Электро-Л и других.

Данные, получаемые с этих КА, используются для решения большого числа задач оперативного мониторинга, в том числе глобального и федерального масштаба. Поэтому необходимо, чтобы во всех этих центрах приема использовались стандартные механизмы работы с данными, включающие в себя обработку данных, их архивацию и предоставление пользователям, т.е. должна быть организована единая унифицированная система работы с данными.

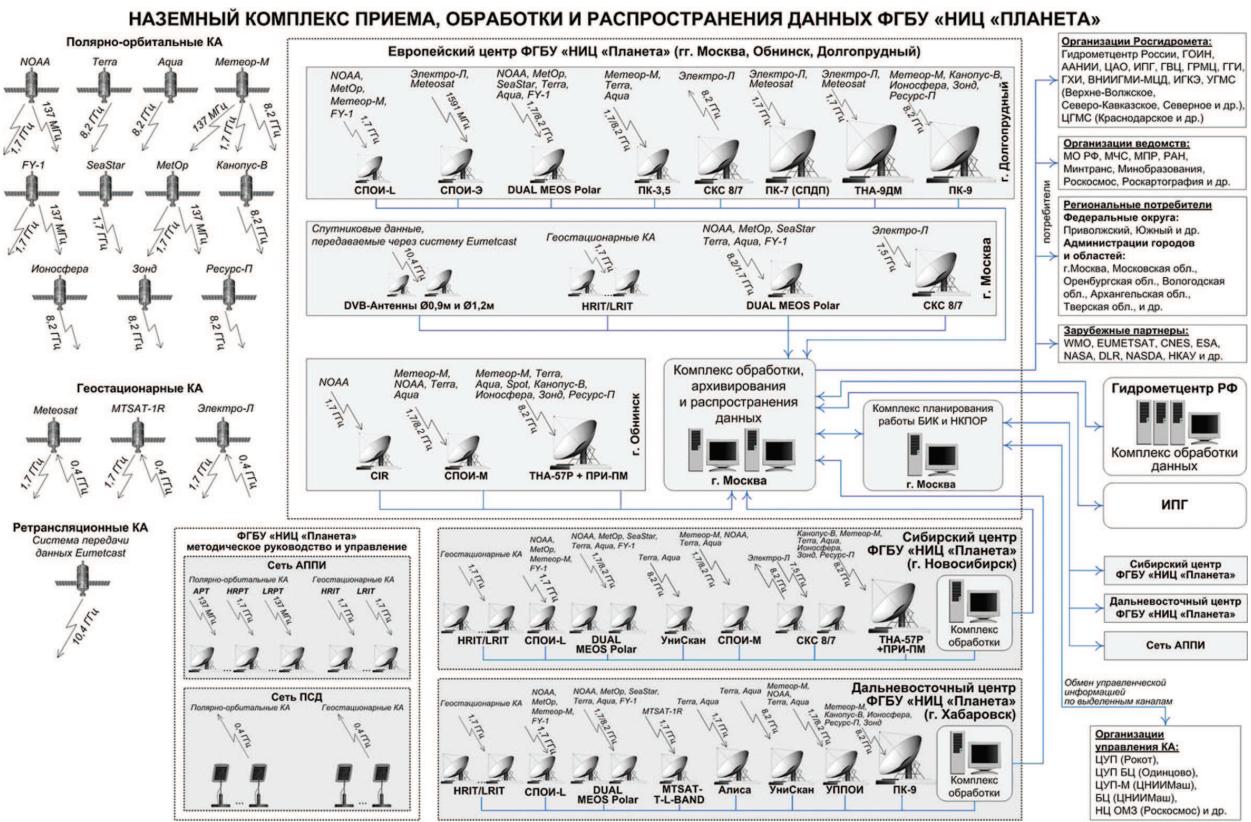


Рис. 1. Схема НКПОР данных ФГБУ «НИЦ «Планета»

### Основные требования к унифицированной системе работы с данными

Для решения задач крупномасштабного гидрометеорологического мониторинга требуется максимально оперативное покрытие единообразными данными больших территорий, например, всей России. Решения данной задачи крайне затруднительно добиться силами одного центра или центрами, в каждом из которых функционирует своя, уникальная система обработки и предоставления данных. Не создав в значительной мере унифицированную систему работы с данными в различных центрах, практически невозможно обеспечить их скоординированную работу и управление.

Для того, чтобы обеспечить унифицированную работу с данными в различных распределенных центрах, необходимо, как минимум, выполнить следующие основные функциональные требования:

– все центры должны оперировать единым архивом исходных и обработанных данных и его каталогом;

– все центры должны быть идентичны и равнозначны по составу производимых продуктов, для этого во всех центрах должны быть созданы и внедрены однотипные комплексы обработки данных (это не исключает того, что в каждом центре дополнительно к стандартизованным продуктам могут производиться и специальные продукты, ориентированные на регион обслуживания центра);

– система должна располагать интерфейсами, обеспечивающими прозрачный доступ как внутренних, так и внешних пользователей к различным информационным продуктам вне зависимости от того, в каком из центров находится данный продукт и/или его фрагмент (в том числе обеспечивая и возможность работы с информацией по территории, заданной пользователями);

– должна иметься возможность организации обмена данными между различными центрами для обеспечения их совместной обработки для производства продуктов, покрывающих зоны работы различных центров.

Кроме того, для обеспечения работоспособности системы и возможности ее эксплуатации и поддержки должны быть выполнены следующие технологические требования:

– все блоки, входящие в состав системы, должны обеспечивать возможность удаленного управления и контроля работоспособности.

– должен иметься блок автоматизированного мониторинга состояния всех распределенных процессов приема, обработки, архивации и представления данных, работающих в системе;

– поскольку в системе задействовано значительное число процессов обработки данных, должен иметься блок, обеспечивающий автоматизированное выявление и диагностику сбоиных ситуаций;

– должна быть реализована развитая система интерфейсов, обеспечивающих возможность управления блоками обработки, архивации и представления данных в различных центрах. Интерфейсы должны обеспечивать возможность управления системой посредством территориально распределенной диспетчерской службы. Это позволит участвовать в работе по контролю системой и ее управлением специалистам всех центров НИЦ «Планета»;

– система должна быть масштабируема. Все блоки системы должны иметь возможность простого расширения и включения новых функций;

Исходя из этих функциональных и технологических требований и разрабатывалась система работы с данными объединённых архивов НИЦ «Планета», элементы которой будут описаны в настоящей работе.

## **Архитектура системы и основные технические решения**

Для разработки подобных систем, обеспечивающих работу со спутниковыми данными, может использоваться несколько различных подходов. Например, традиционный подход, предполагающий построение строго заданного набора тематических продуктов с фиксированным географическим покрытием и метаинформацией, и предоставление их пользователю посредством различных способов передачи данных (e-mail, ftp, http и т.д.). Данный подход, безусловно, обладает рядом достоинств:

– Простота организации хранения данных. Хранение каждого продукта происходит независимо, и организация его хранения не требует согласования с организацией хранения других продуктов, например, произведенных в других центра.

– Простота организации представления информации пользователям. Пользователи просто получают по запросу и/или в виде рассылки уже готовый продукт (чаще всего графический файл). При этом на момент представления данных пользователям уже не требуется выполнения никаких дополнительных операций по обработке данных.

– Простота учета и планирования ресурсов на обработку данных. В процессе работы системы обычно формируется определенный состав продуктов, которые, как уже говорилось, в дальнейшем не требуют никакой дополнительной обработки.

В то же время, традиционный подход имеет целый ряд существенных недостатков, которые не позволяют эффективно использовать его в достаточно сложных распределенных системах, работающих с большим объемом данных. К ним, в частности, относятся:

– Создаваемые на основе данного подхода системы очень негибки и тяжело наращиваются. Необходимость получения дополнительной информации (не содержащейся в формируемых продуктах или содержащейся в различных продуктах) фактически приводит к необходимости создания нового продукта, формирования для него процедуры обработки, хранения и представления данных.

– По мере увеличения числа производимых продуктов лавинообразно растет число процессов. Это приводит к такому же быстрому росту потребности в вычислительных ресурсах, необходимых для работы системы.

– Практически неизбежно возникает частичное дублирование информации в различных продуктах, что, в свою очередь, приводит к трудноконтролируемому росту объемов архивов производных продуктов;

– Достаточно сложно организовать взаимодействие различных распределенных систем обработки данных, функционирующих в разных центрах;

– Ввод в систему дополнительных продуктов обычно требует существенных временных затрат.

Вместо описанного традиционного подхода может быть предложен альтернативный подход, суть которого заключается в построении базовых информационных продуктов в форматах, позволяющих доступ к произвольным географическим областям, покрытым данными. Возможность реализации такого подхода обсуждались, в частности, в работах (Ефремов и др., 2007, Лупян, Саворский, 2012). В данном подходе для того, чтобы обеспечить пользователям быстрый доступ к информации по произвольным регионам, используются специально структурированные архивы данных. Созданные таким образом данные предоставляются пользователям посредством специализированных картографических web-интерфейсов, обеспечивающих как возможность просмотра и анализа данных, так и возможность создания собственных тематических продуктов на основе уже имеющихся базовых продуктов. У данного подхода можно выделить следующие преимущества:

– Имеется возможность доступа к данным различных КА по произвольной области в необходимом масштабе.

– Возможно построить достаточно универсальный инструментарий работы с данными, обеспечивающий возможности генерации и анализа данных в режиме онлайн.

– Во многих случаях не требуется наличия у пользователей сложных специализированных программных комплексов для проведения обработки спутниковых данных для получения необходимых ему специализированных информационных продуктов.

– Формируемые продукты практически не дублируются, что позволяет существенно сократить объемы хранимых данных.

– Пользователи имеют возможность самостоятельного формирования различных информационных продуктов по мере их необходимости.

– Пользователи имеют возможность одновременно работать с данными, расположеными в различных центрах приема и обработки информации.

– Возможно максимально унифицировать работу по подготовке, хранению и представлению данных во всех центрах приёма и обработки данных, что крайне облегчает построение, управление и расширение распределённой системы работы с данными.

В тоже время такой подход также имеет некоторые недостатки. Наиболее существенными из них являются:

– Необходимость создания достаточно сложных систем хранения и представления информации.

– Необходимость выделения специальных вычислительных ресурсов, производящих обработку данных по запросу пользователей. При этом следует учитывать, что может происходить сокращение мощностей на производство фиксированных информационных продуктов.

– Часть продуктов, которые формируются по запросам пользователей, могут производиться по несколько раз, при запросе их различными пользователями.

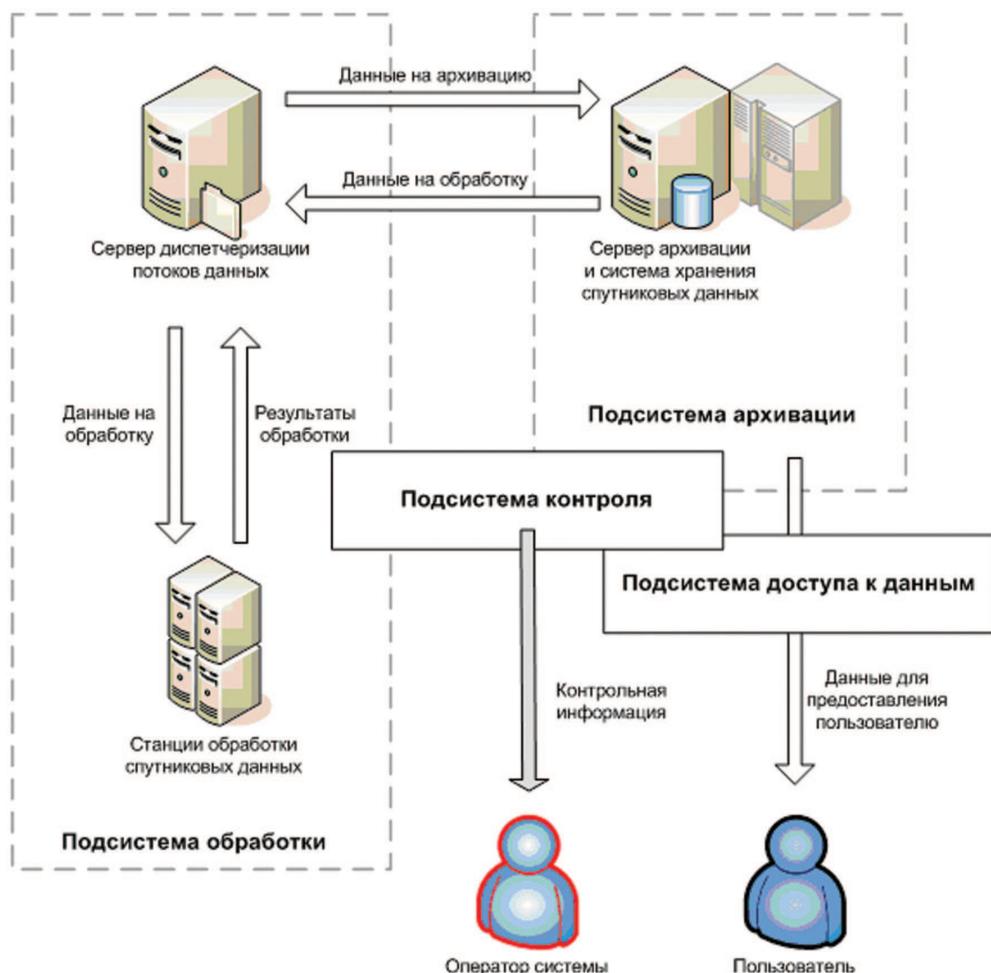
– Поскольку значительная часть продуктов получается и используется пользователями в режиме онлайн, резко возрастают требования к сетям передачи данных, использующихся в центрах приема. В тоже время следует учитывать, что за счет того, что во многих случаях в данной схеме для производства продуктов по большим территориям нет необходимости концентрировать все данные в одном центре, реальные объемы передаваемых данных могут быть и сокращены.

Таким образом, исходя их сказанного выше, можно отметить, что первый подход оптимален для локальных систем мониторинга с ограниченным числом тематической информационной продукции, а второй подход гораздо лучше приспособлен для глобальных систем мониторинга, включающих в себя большое количество разнообразных данных. Кроме того, при планировании распределенной системы работы с данными в центрах НИЦ «Планета» следовало учесть, что на протяжении значительного времени в центрах использовался традиционный подход к работе со спутниковыми данными. Поэтому для создаваемой распределенной системы работы с данными был выбран гибридный вариант ее построения. Основой для него стал описанный выше альтернативный подход, в рамках которого была реализована распределенная система хранения данных в виде, обеспечивающем возможность доступа к информации по произвольной пространственной области и возможность прозрачного получения данных из различных центров. В то же время в создаваемой системе была реализована возможность работы с фиксированными тематическими продуктами, традиционно пользующимися спросом со стороны потребителей, наряду с базовыми информационными продуктами, покрывающими всю территорию наблюдений.

Для реализации данного подхода была разработана общая архитектура системы, функциональная схема которой приведена на рис 2. Системы включает в себя четыре основных

компоненты, особенности построения которых будут подробно рассмотрены в настоящей работе:

- подсистема архивации данных;
- подсистему обработки данных;
- подсистему доступа к данным;
- подсистему контроля.



*Рис. 2. Функциональная схема системы работы с распределенными архивами результатов обработки спутниковых данных центров приема НИЦ «Планета»*

### Подсистема архивации

Архивы и блоки их ведения являются одним из ключевых элементов всей системы работы с данными. В настоящий момент в центрах НИЦ «Планета» принимаются данные с девяти полярно-орбитальных КА Д3З и с шести геостационарных КА Д3З. За сутки в одном центре может быть принято около 200 сеансов. Все принятые в центре данные проходят через тематическую обработку, порождающую от одного до нескольких информационных продуктов. Подсистема архивации данных должна обеспечить возможность хранения как исходных данных, так и различных информационных продуктов. При этом суточный поток подлежащих архивации сеансов и информационных продуктов зачастую превышает 500 единиц. Поэтому главное требование, предъявляемое к системе архивации, — это максимальная автоматизация процедур, связанных с ведением архива, т.е. подготовки данных

к занесению в архив, погружение данных в архив, миграции данных между узлами архива и физическими системами хранения и т.д.

Одной из основных проблем, связанных с организацией архивов в центрах приема НИЦ «Планета», являются большие объемы получаемой и обрабатываемой информации. С учетом того, что уже сейчас в центрах принимается и обрабатывается около 300 Гбт данных в день, скорость пополнения архивов данных может превышать 1 Тбт в сутки. Для поддержки архивов, пополняющихся с такой скоростью, требуется использовать специальные подходы. При этом следует иметь в виду, что, даже при достаточно быстром развитии дисковых систем хранения, организация хранения на них всех потоков информации, получаемых в центрах, достаточно дорога. Поэтому следует разделить данные на как минимум две категории, для одной из которых должны использоваться системы хранения, поддерживающие быстрый (непосредственный) доступ ко всей информации, а для другой часть информации может быть перенесена на сменные (долговременные) носители. Ко второй категории для центров НИЦ «Планета» может быть отнесена телеметрическая (сырая) информация, объем которой достаточно велик, а востребованность после прохождения процедур обработки и построения базовых продуктов достаточно ограничена.

Для информации этой категории архивы данных могут быть разделены на циклические оперативные, обеспечивающие доступ к текущим данным в режиме реального времени или близком к нему, и долговременные, доступ к данным которых может производиться с некоторой задержкой по мере их запросов. В оперативных архивах обычно хранятся данные за заданный промежуток времени, по его истечении они переносятся в долговременные архивы, а освободившееся пространство замещается свежими данными. По мере поступления запросов в оперативных архивах могут восстанавливаться данные из долговременных архивов. Оперативные архивы строятся на дисковых системах хранения, а для построения долговременных архивов могут использоваться системы хранения со сменными носителями (оптические диски, магнитные ленты и т.д.). Естественно, что система ведения таких архивов должна поддерживать единообразное ведение обоих типов архивов для обеспечения прозрачной миграции данных из оперативных архивов и обратно, а также для минимизации процедур занесения данных в архивы.

Архивы центров, к которым должен быть обеспечен постоянный оперативный доступ, например, архивы результатов обработки спутниковых данных, также имеют достаточно большой объем и постоянно растут. Поэтому система их ведения должна обеспечивать поддержку хранения данных на различных, в том числе и территориально распределенных серверах и дисковых системах. Таким образом, в таких архивах должна поддерживаться возможность работы с множеством томов хранения данных. При этом, к данным архивам должен быть обеспечен прозрачный доступ вне зависимости от того, на какой физической системе и в каком центре осуществляется их хранение.

Исходя из выше перечисленных особенностей, были сформулированы следующие основные требования к организации и подсистеме ведения архивов данных центров НИЦ «Планета». Подсистема должна обеспечивать:

- Максимальную автоматизацию всех процедур ведения архива.
- Поддержку работы распределенных архивов данных.
- Ведение как оперативных и долговременных архивов.
- Унифицированное ведение архивов в различных центрах.

Только совместное выполнение этих требований может позволить обеспечить максимальную оперативность доступа к данным за счёт минимизации задержек на всех этапах жизненного цикла данных. Также следует отметить, что автоматизированная подсистема ведения архивов данных позволяет минимизировать ошибки при ведении архивов за счёт устранения человеческого фактора.

Рассмотрев общие требования к организации архивов спутниковых данных и результатов их обработки в НИЦ «Планета», остановимся на некоторых деталях основных типов архивов, использующихся в центрах. К таким типам относятся: архивы исходных данных (телеметрии) и архивы результатов обработки данных.

Архивы исходных данных строятся, как правило, для однотипных данных, получаемых с КА одной серии, как, например, КА серии NOAA, либо данных одинаковых измерительных приборов, установленных на разных КА (КА TERRA, AQUA, прибор MODIS). С точки зрения реализации они представляют собой выделенную файловую БД, в которую заносятся метаданные о каждом сеансе данных, а также ссылки на сами файлы данных, помещаемые в специализированное файловое хранилище. Системы архивации или ведения таких архивов являются типовыми, отличаясь друг от друга только настройками, отражающими специфику и набор хранимых данных и метаданных (фактически – структуру БД архива). В состав метаданных, как правило, входят дата и время сеанса приёма данных, КА и прибор, с которого данные получены, номер витка КА и идентификатор станции приёма. Основной задачей типовой системы ведения архива исходных данных является пополнение БД. Кроме этого на неё возложены также различные сопутствующие задачи, связанные с ведением базы данных: экспорт и импорт данных, удаление неактуальных данных, перенос части данных на автономные носители информации, резервное копирование баз данных и др.

Архивы результатов обработки спутниковых данных строятся таким образом, чтобы обеспечить максимально быстрый выбор данных. В том числе они должны быть рассчитаны на возможность выбора информации по произвольной территории. Для организации таких архивов должны быть использованы специальные подходы, которые рассматриваются, в частности, в работах (Ефремов и др., 2007, Балашов и др., 2008).

Один из подходов к построению таких архивов основан на том, что изображения в архиве хранятся в виде фиксированного разбиения на гранулы небольшого размера, общего для всех продуктов. При этом для организации доступа к изображениям в произвольном масштабе используются дополнительные разбиения на гранулы, формирующие пирамиду разрешений. Необходимое пользователю изображение в этом случае динамически формируется на сервере на основе соответствующих гранул нужного масштаба. Для этого готовые спутниковые изображения разбиваются на гранулы базового масштаба и уже в таком виде поступают на архивацию. На этапе архивации на основе гранул базового масштаба получаются также гранулы дополнительных масштабов, формирующие пирамиду разрешений. Подобный подход наиболее эффективен, когда продукты обработки спутниковых данных представляют собой покрытия больших территорий, полученные на основе различных сеансов наблюдений (например, безоблачные композитные изображения)

В случае, когда результаты обработки формируются для одного сеанса (сцены) или их фрагмента, может быть использован вариант индивидуального разбиения на гранулы для каждого сеанса (сцены). Этот подход целесообразно использовать в том случае, когда разбиение на гранулы реализуется в рамках графического формата, применяемого для хранения спутниковых изображений. В настоящее время в качестве такого формата реально

может быть использован формат GeoTIFF. Разбиение изображения на гранулы, формирующие пирамиду разрешений, возможно в рамках стандартной реализации этого формата. Использование формата GeoTIFF позволяет упростить реализацию архива, но накладывает определенные ограничения на область его применения. В частности, подход перестает быть эффективным при работе со сценами большого размера.

Особо следует отметить, что для ведения больших распределенных архивов данных целесообразно использовать специальные технологии, позволяющие организовывать ведение файловых архивов под управлением специализированной базы данных, в которой хранятся все метаданные об отдельных изображениях, помещаемых в архив, и адреса их физического хранения. Это позволяет ускорить все процессы поиска данных в архивах, а также организовывать распределенное хранение данных с возможностью централизованного ведения базы метаданных. Для организации архивов данных в НИЦ «Планета» была использована подобная технология, разработанная в ИКИ РАН и описанная в работах (Лупян и др., 2000, Ефремов и др., 2004)

### **Взаимодействие архивов центров НИЦ «Планета»**

Для построения единого распределённого архива на основе архивов различных центров можно выделить два подхода. Первый предполагает постоянный обмен метаданными между центрами по мере пополнения их архивов с целью постоянно иметь информацию о полном содержимом объединённого архива во всех его узлах. Второй предполагает независимое ведение архивов в каждом центре и объединение их не на уровне метаданных, а на уровне пользовательских интерфейсов доступа к данным.

Объединение архивов на уровне метаданных позволяет пользователю в любой момент времени получить информацию о наличии данных по любым интересующим его критериям при обращении к любому из узлов архива, независимо от того, принадлежат ли требуемые данные к архиву этого центра или не принадлежат. Обмен метаданными осуществляется асинхронно посредством стандартных протоколов обмена данными. При каждом пополнении или изменении архива в одном центре метаданные добавленных и изменившихся данных пересыпаются во все остальные центры, где усваиваются в локальные каталоги архивов со ссылками на источник данных. При необходимости доступа к самим данным они запрашиваются из соответствующих центров.

Объединение архивов на уровне интерфейсов доступа также позволяет пользователю в любой момент времени получить информацию о наличии данных по интересующим его критериям при обращении к любому из узлов архива. Основное отличие от предыдущего метода заключается в том, что информация о наполнении локальных архивов на постоянной основе не передаётся, а для получения каждой требуемой выборки данных во все центры единого архива направляются запросы, и по их результатам формируется ответ пользователю. Функциональные схемы методов взаимодействия архивов центров между собой представлены на рис. 3 и 4.

У каждого из подходов есть свои преимущества и недостатки. Объединение на уровне метаданных предполагает асинхронный обмен данными. С одной стороны, это может привести к тому, что за счёт временных задержек на пересылку и усвоение метаданных пользователь может получить информацию о требуемых ему данных с некоторым запозданием. С другой стороны, такая схема в сочетании с хранением полного комплекта

метаданных архива в каждом его узле позволит пользователю иметь максимально полную информацию об архиве в любой момент времени независимо от тех или иных проблем со связью между узлами. Кроме того, все сложные запросы на выборку данных осуществляются в рамках одного центра, что существенно увеличивает быстродействие системы доступа к данным. Объединение на уровне интерфейсов предполагает синхронный обмен данными, что позволяет пользователю всегда получать информацию на текущий момент времени. Вместе с тем, эта схема очень требовательна к каналам связи и постоянной работе всего набора оборудования распределенной системы хранения во всех центрах, так как в случае задержек или сбоев связи или отказе серверов, ведущих базы метаданных в одном из центров, его архивы становятся полностью недоступны системе доступа к данным. Необходимость частого обмена достаточно сложными запросами между центрами способна также понизить быстродействие системы доступа к данным. В то же время необходимо отметить, что минусом объединения на уровне метаданных является достаточно большая сложность реализации и учета всех нюансов, которые могут возникнуть при синхронизации баз данных.

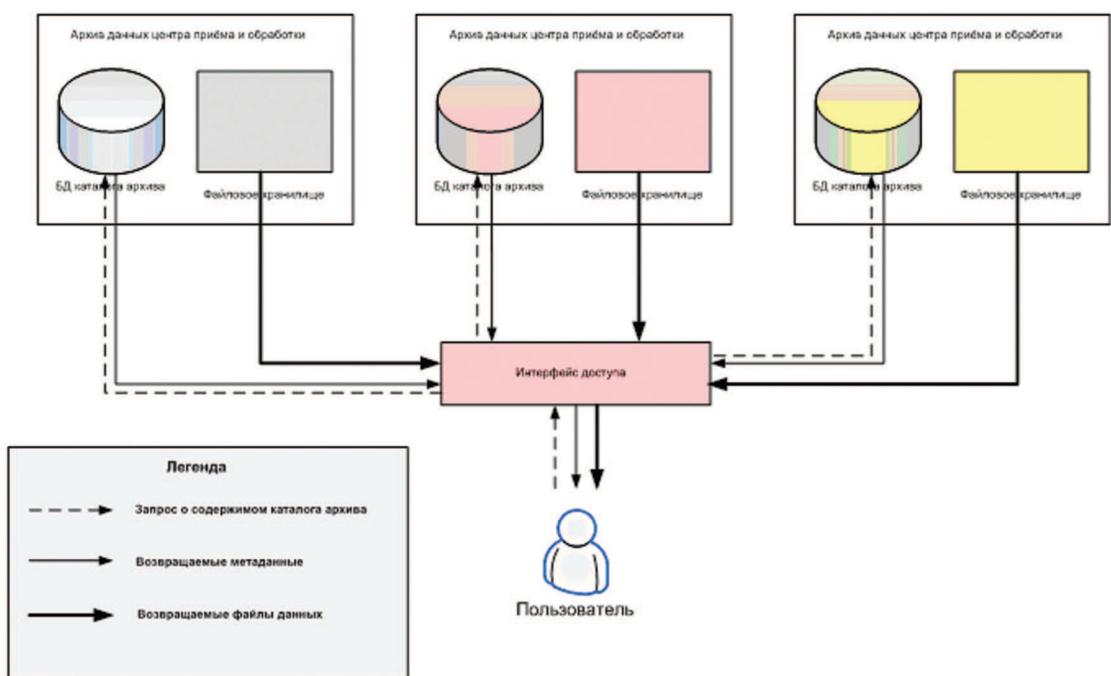


Рис. 3. Взаимодействие архивов исходных данных

Рассмотренные подходы оптимально использовать для решения работы с архивами разного класса. Подход, основанный на объединении на уровне метаданных, оптимален для сложных архивов с большим количеством данных, в том числе единовременно предоставляемых пользователю. Сложность его реализации полностью компенсируется скоростью и надёжностью работы системы доступа к данным. Подход с объединением на уровне интерфейсов доступа оптимален для простых архивов, к которым не идёт сложных запросов и не так критична надёжность и скорость доступа. Для реализации распределенной системы работы с данными в центрах НИЦ «Планета» было выбрано комплексное решение, при котором единый архив результатов тематической обработки данных построен на уровне объединения метаданных, а архивы исходных данных – на уровне объединения интерфейсов доступа.

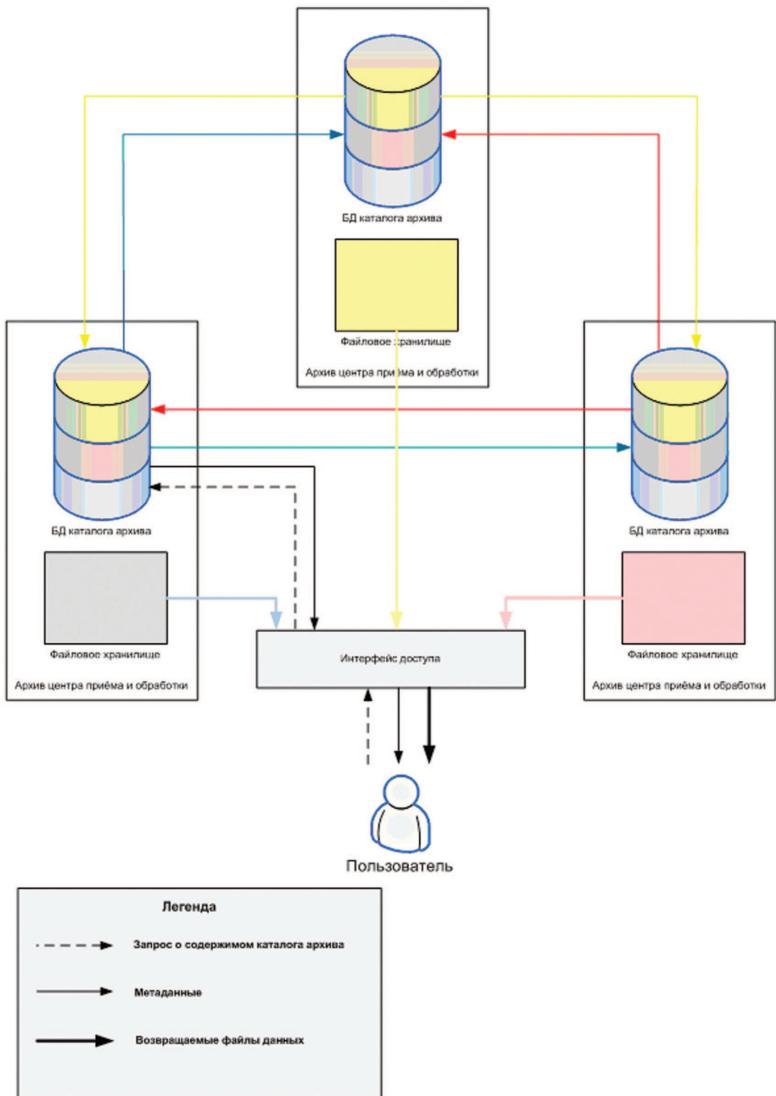


Рис. 4. Взаимодействие архивов результатов обработки

### Построение системы доступа к данным

Для организации доступа к распределенным архивам данных НИЦ «Планета» была разработана система web-интерфейсов, позволяющих работать с информацией, хранящейся в архивах Европейского, Сибирского и Дальневосточного центра. Основной задачей данных интерфейсов является предоставление пользователям возможности получения доступа к данным вне зависимости от того, в каком из центров НИЦ «Планета» они физически располагаются. В настоящее время в системе реализовано два типа интерфейсов:

- табличный – рассчитанный на получение файлов исходных данных;
- картографический – обеспечивающий работу с результатами обработки спутниковых данных.

Оба типа интерфейсов созданы на основе технологий и базового программного обеспечения, разработанного в ИКИ РАН (Андреев и др., 2004, Толпин и др., 2011).

Табличный интерфейс разработан для обеспечения доступа к содержимому распределённого архива исходных данных. Интерфейс позволяет просматривать информацию о на-

полнении каталога, количестве сеансов данных за требуемые даты, просматривать текстовые аннотации к сеансам данных, обзорные изображения к ним же, и скачивать сами файлы данных. Помимо этого, интерфейс обеспечивает возможность фильтрации данных по станциям приёма и аппаратам, а также возможность перехода в картографический web-интерфейс для просмотра тематических продуктов по данным выбранного сеанса. Интерфейс позволяет работать с исходными данными, полученными в центрах приема НИЦ «Планета» вне зависимости от того, в архивах каких центров физически расположены данные. Основной задачей интерфейса является обеспечение возможности удобного поиска файлов данных и организация их получения.

Картографический интерфейс создавался для решения следующих основных задач:

- Обеспечение работы с многомерными архивами данных (которые содержат различные информационные продукты (характеристики), являющиеся пространственными покрытиями, полученными в различные моменты времени, т.е. обеспечивать работу как с пространственными данными, так и с временными рядами);
  - Обеспечение возможности одновременного отображения и анализа различных видов информационных продуктов;
  - Обеспечение динамического формирования комплексных информационных продуктов на основе базовых продуктов по запросу пользователей;
  - Предоставление различных инструментов для анализа данных (например, контрастирования отдельных изображений, сравнения разновременных изображений, расчета различных характеристик изображений, выделения на изображениях различных объектов и т.д.);
  - обеспечение прозрачной работы с данными, находящимися в архивах различных центров.
  - обеспечение онлайн-импорта информации из различных сторонних информационных систем и экспорта данных в такие системы;
  - и т.д.

Эти задачи определяют основные требования, которым должен удовлетворять картографический интерфейс. В том числе:

- удобное и интуитивно понятное управление данными (выбор, поиск, включение для отображения нужных характеристик и т.д.);
  - быстрое отображение выбранных наборов данных;
  - возможность простого расширения функциональности интерфейсов (включение в него дополнительных функциональных возможностей);
  - возможность «прозрачной» работы с распределенными архивами данных;
  - поддержка гибкой системы авторизации пользователей для разграничения доступа к различным функциям и информационным продуктам, используемым в системе;

Для создания картографического интерфейса была использована технология GEOSMIS (Толпин и др., 2011), разработанная в ИКИ РАН, предназначенная для построения сложных картографических интерфейсов для работы с данными систем дистанционного мониторинга.

Центральным элементом интерфейса является его программное ядро, обеспечивающее взаимодействие всех модулей получения данных с одной стороны и сервисов предоставления данных с другой. Оно представляет собой набор базовых функций и методов, позволяющих получать запросы от интерфейса на метаданные и данные, передавать их соответствующим модулям получения данных, принимать от них ответ и передавать его соответствующим сервисам интерфейса для отображения пользователю. Также в рамках

ядра реализованы функции отображения растровой и векторной информации и основные базовые элементы управления. Важным компонентом ядра являются файлы настроек, где хранятся адреса и параметры доступа к архивам данных, в т.ч. удалённых. Отдельным пунктом в составе ядра интерфейса следует выделить систему авторизации и разграничения доступа пользователей.

За работу с отдельными типами данных отвечают специализированные модули, обеспечивающие получение стандартизованных запросов от ядра интерфейса, возврат результатов в стандартизированном же виде и набор элементов интерфейса для работы с этими данными. Каждый модуль состоит из двух компонентов – серверного и клиентского. Серверный компонент обеспечивает доступ к данным в соответствии со стандартом WMS. Клиентский компонент обеспечивает работу элементов управления, предназначенных для формирования WMS-запросов на получение данных. Также при необходимости в модуль добавляются функции взаимодействия с другими модулями системы и функции работы с данными и их анализа. Общий вид интерфейса представлен на рис. 5. Принципиальная схема архитектуры интерфейса приведена на рис. 6.

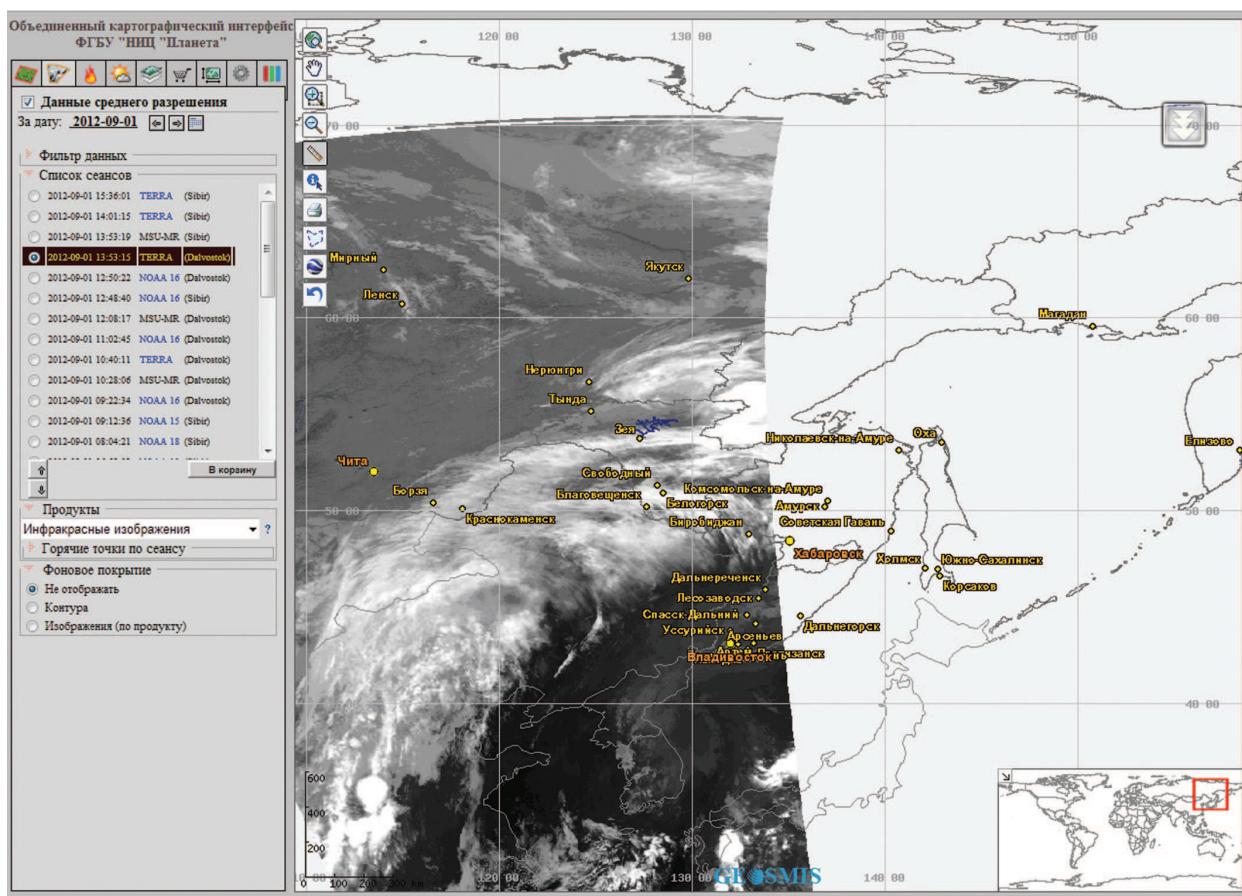


Рис. 5. Общий вид интерфейса

Следует отметить, что интерфейс позволяет организовать работу с данными, физически находящимися в системах хранения различных центров. В случае запросов данных из разных центров их обработка фактически идет по описанной выше схеме. Сервера, обеспечивающие работу интерфейсов, установлены во всех центрах НИЦ «Планета». При этом, как уже обсуждалось в разделе, посвященном организации архивов, в каждом центре хранится копия всех метаданных распределённого каталога. Это позволяет системе интерфейсов различных центров использовать для работы с метаданными только локальные БД.

С точки зрения интерфейса работа с данными, находящимися в локальных или удаленных архивах, отличается только адресами, по которым идёт обращение к модулю получения данных. С точки зрения пользователя, работающего с интерфейсом, при этом нет никаких различий, из каких архивов данных реально получен тот или иной продукт.

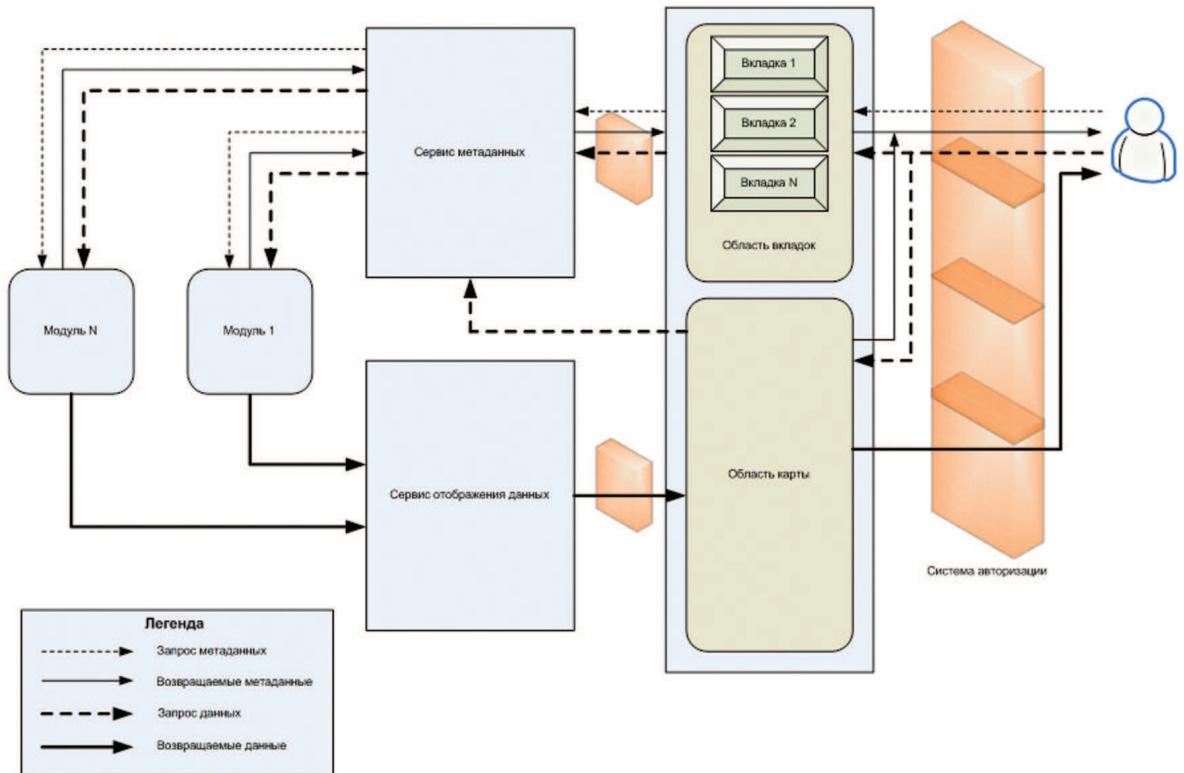


Рис. 6. Принципиальная схема интерфейса

При работе с интерфейсом пользователю доступны две его основные области – блок управления («вкладок») и область карты. Область вкладок обеспечивает управление составом отображаемых данных и позволяет включать/отключать отображение данных по типам, осуществлять фильтрацию данных в рамках одного типа по временным, пространственным и качественным характеристикам. В области карты производится отображение выбранных данных, а также в ней присутствуют инструменты работы с картой (сдвиг, масштабирование), инструменты работы с точками, линиями и полигонами и инструменты экспорта данных. Подробно базовые функциональные возможности подобных интерфейсов описаны в (Толпин и др., 2011).

При обращении пользователя к функциям интерфейса, обеспечивающим выбор различных типов данных («вкладки управления продуктами»), посыпается стандартизованный запрос к сервису метаданных по выбранному активному типу данных с учётом временных, качественных и пространственных фильтров, определяемых заданными во вкладке параметрами и выбранным участком карты. Сервисом метаданных этот запрос перенаправляется в специализированный модуль работы с этим типом данных, который запрашивает БД каталога архива и возвращает результаты обратно в сервис метаданных. Полученные результаты запроса сервис направляет «вкладке», где они отображаются в виде списка доступных данных и тематических продуктов. Пользователь на основе этого списка запрашивает необходимые данные для отображения. Запрос, содержащий ме-

таданные с описанием требуемого, передаётся в специализированный модуль, который извлекает данные из непосредственно БД или связанного с ней файлового хранилища и наряду с необходимой для визуализации информацией передаёт сервису отображения данных. Сервис отображения данных, в свою очередь, передаёт полученные данные в область карты, где они становятся доступны пользователю для работы с ними. Также запрос на данные может быть инициирован из области карты, например, при сдвиге или масштабировании изображения.

### **Основные возможности картографического интерфейса обеспечивающего работу с распределенными архивами данных НИЦ «Планет»**

Картографический интерфейс в настоящее время обеспечивает возможность работы с различными типами продуктов, получаемых в центрах приема и обработки НИЦ «Планета», в том числе:

- продуктами, полученными на основе данных высокого пространственного разрешения (КА Landsat, Канопус-В);
- продуктами, полученными на основе данных среднего пространственного разрешения (прибор КМСС);
- Продуктами, получаемыми на основе информации, поступающей с низкоорбитальных спутниковых систем (NOAA, Terra, Aqua, «Метеор-М» №1 и других);

Он также позволяет оперировать результатами обработки спутниковых данных, предоставляемыми различными специализированными системами дистанционного мониторинга и системами ведения архивов спутниковых данных. Например, данными спутникового сервиса ВЕГА (Лупян, Савин и др.), геопортала Роскосмоса (Носенко, Лошкарев, 2010) и др.

В интерфейсе реализованы специализированные «вкладки», ориентированные на работу с различными типами продуктов, например:

- продуктами, полученными на основе отдельных сеансов и сцен;
- продуктами, полученными на основе серий наблюдений различных территорий (например, многодневными композитами);
- различными специализированными, эпизодически формируемыми продуктами для локальных районов (например, продуктами, ориентированными на анализ ледовой обстановки);
- специализированными продуктами, ориентированными на мониторинг природных пожаров (данными об активном горении и площадях, пройденных огнем).

Одновременно с различными продуктами, полученными на основе спутниковых данных, интерфейс позволяет оперировать оперативной и архивной информацией, необходимой для решения задач гидрометеорологического мониторинга, например:

- картографической информацией;
- фактическими метеоданными;
- прогностическими метеоданными.

Основной особенностью созданного интерфейса является то, что он ориентирован не только на поиск и просмотр отдельных продуктов обработки спутниковых данных из системы распределенных архивов НИЦ «Планета», но и на возможность проведения комплексного анализа информации и её обработки. Интерфейсы позволяют производить различные операции над получаемыми продуктами, например:

- осуществлять выбор произвольной территории и удобного масштаба для анализа данных;
- проводить различные контрастирования;
- проводить перевод данных в различные проекции;
- формировать цветостинтезированные изображения, в том числе и с использованием композитных съемок;
- выделять и производить измерения параметров различных объектов;
- осуществлять сравнение разновременных наблюдений различных районов и объектов;

Пример работы вкладки дополнительного контрастирования различных информационных продуктов приведен на рис. 7.

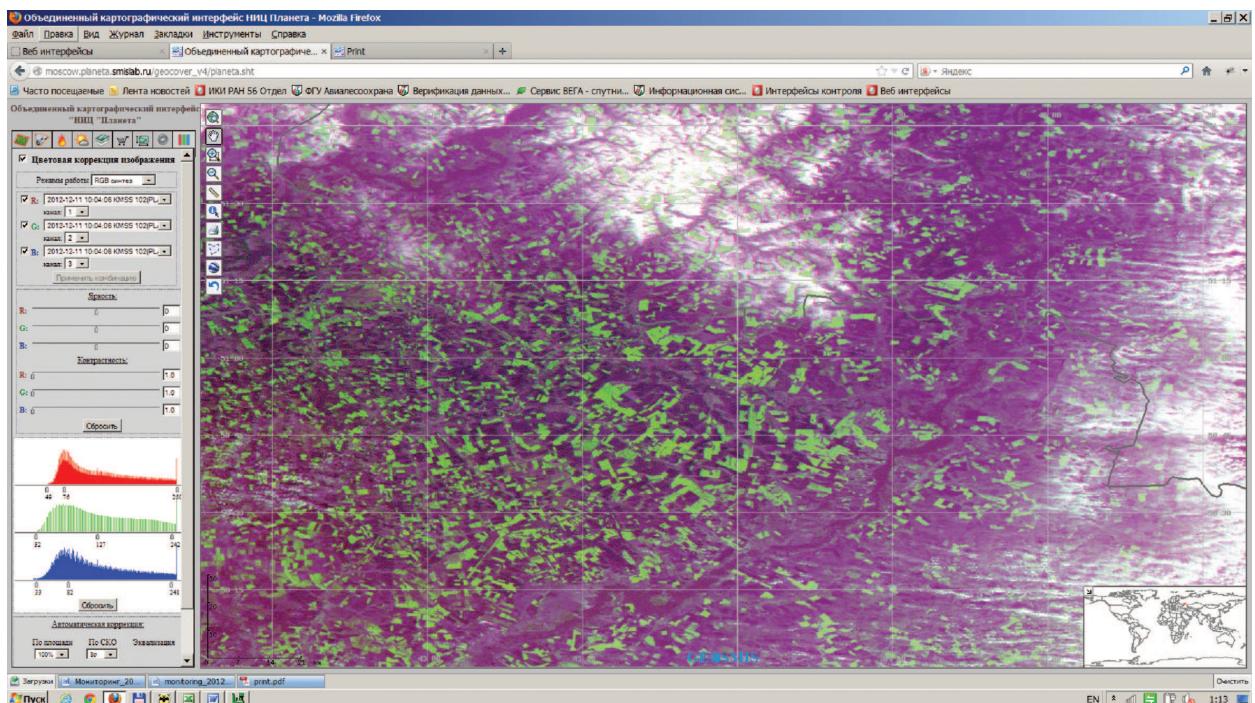


Рис. 7. Пример работы вкладки контрастирования изображений

Для сравнения продуктов (изображений) интерфейс позволяет использовать различные механизмы. Например, возможно, используя данные, полученные в различное время, синтезировать изображения, в разных каналах которых будут представлены данные, полученные в разное время. Имеются также механизмы типа «шторка», позволяющие проводить сравнение разновременных изображений или изображений из различных источников (см. рис 8.).

Одной из основных задач интерфейса является обеспечение возможности комплексного анализа различных ситуаций с использованием разнородных данных. Иллюстрация такой возможности приведена на рис. 9. На нём на примере ситуации, наблюдавшейся в Красноярском крае 01.08.2012, демонстрируется возможность совместного анализа информации о пожарах (красные контуры), облачном покрове (базовое изображение), распределения температуры (сетка цифр). При этом представленное распределение облачности получено по данным прибора MODIS (разрешение 250 м.), в то же время показано, что на некоторые участки представленного района за те же сутки имеются и данные LANDSAT (разрешение 30 м.) (синие контуры границ сцен и пример изображения одной из сцен).

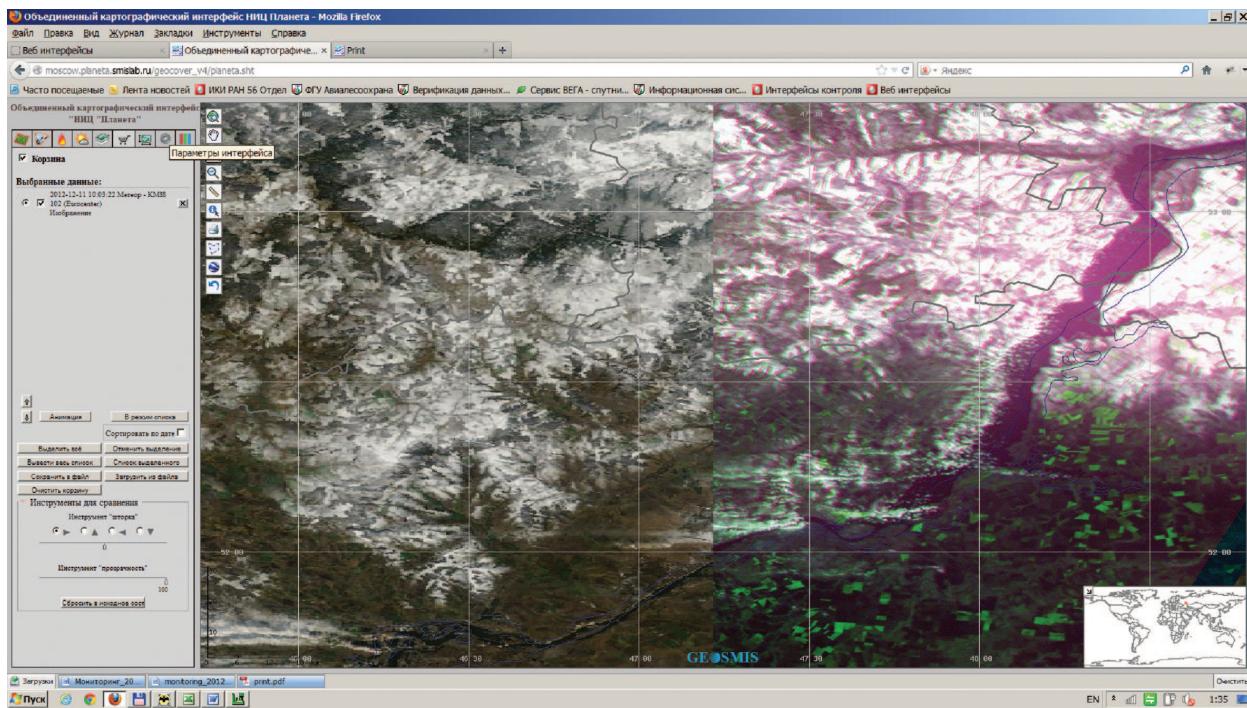


Рис. 8. Пример сравнения различных продуктов. Сравнения с использованием механизма «шторка»

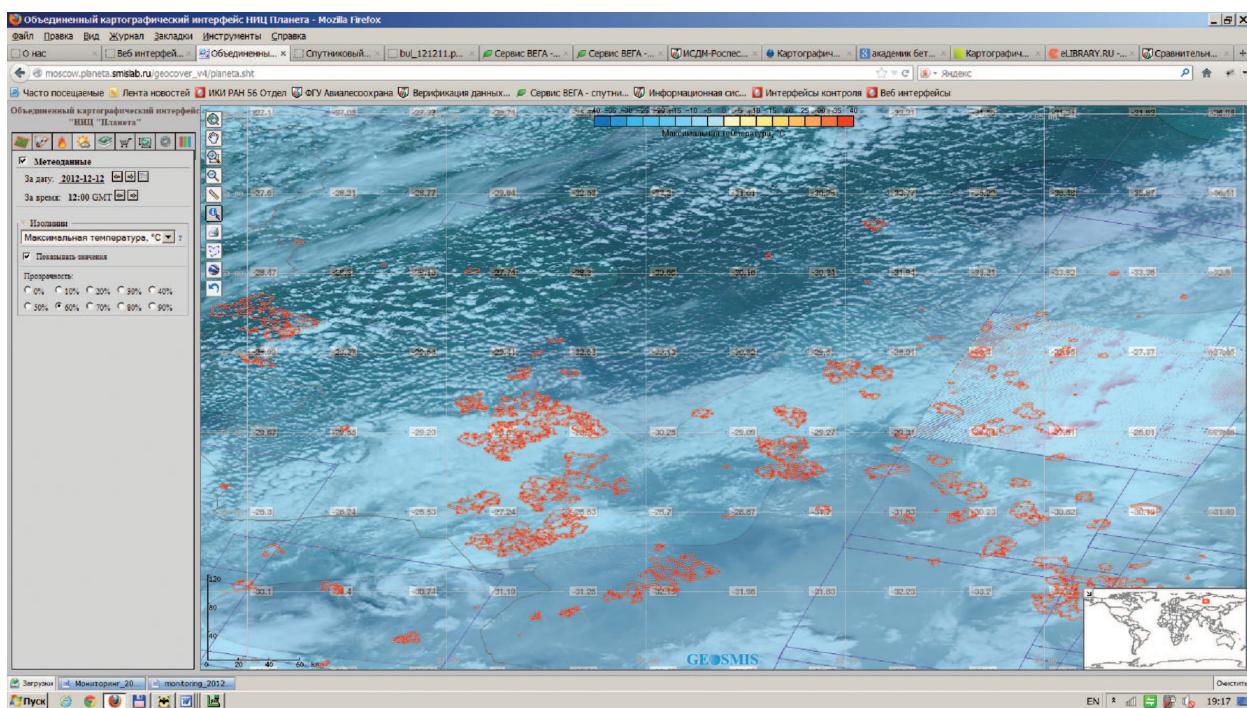


Рис. 9. Пример возможности анализа различных типов данных

Следует обратить внимание, что созданная система интерфейсов по факту позволяет динамически создавать новые информационные продукты на основе базовых продуктов, имеющихся в распределенных архивах данных НИЦ «Планета». Например, на основе данных, полученных в различных центрах, могут быть сформированы различные покрытия («склейки») для выбранных территорий наблюдений. Пример такой «склейки» для территории Центральной и Восточной Сибири, построенной на основе данных прибора MODIS,

установленного на спутниках TERRA и AQUA, полученных в Сибирском и Дальневосточном центрах приема, приведен на рис. 10. Отметим, что построение «склейки» осуществляется непосредственно в интерфейсе, при этом данные, входящие в её состав, автоматически получаются из архивов различных центров на момент запроса.

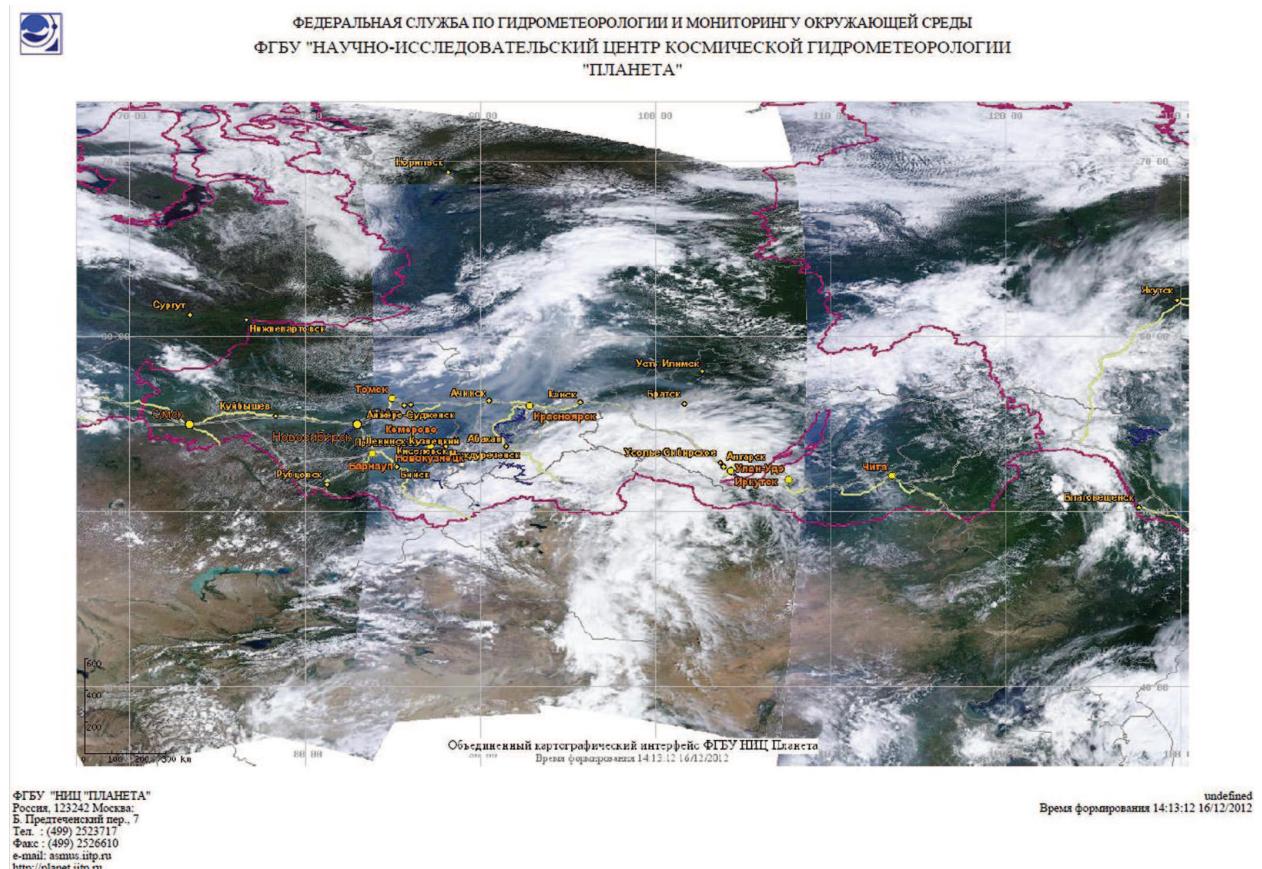


Рис. 10. Пример оперативной «склейки» цветосинтезированных изображений, полученных на основе данных прибора MODIS в различных центрах приема НИЦ «Планета»

В тоже время, интерфейс позволяет решать и традиционные задачи, связанные с поиском, выбором и заказом данных. Он позволяет пользователям найти интересующие его данные по заданным территориям и периодами времени. Он также может сформировать заказ необходимых ему данных, произведя их поиск и отбор. Заказанные данные будут автоматически выбраны из архивов различных центров, и пользователю будет направлена информация с параметрами доступа к ним.

Следует также отметить, что система, обеспечивающая доступ к данным, позволяет организовать специальные программные интерфейсы (API), обеспечивающие предоставление данных различным специализированным системам дистанционного мониторинга. Сегодня такие интерфейсы используются для организации оперативного поступления данных в различные системы, например:

- Информационная система дистанционного мониторинга лесных пожаров Федерального агентства лесного хозяйства (Коровин и др., 2010, Нестеренко и др., 2004);
- Отраслевая система мониторинга Федерального агентства по рыболовству (Андреев и др., 2007, Лупян и др., 2004);
- Спутниковый Сервис Вега (Лупян, Савин и др., 2011);
- Различные региональные информационные системы.

## **Заключение**

Система работы с распределёнными архивами данных, созданная в интересах ФГБУ «НИЦ «Планета», предоставляет большой диапазон возможностей доступа к спутниковым данным и работы с ними. Система позволяет обеспечить доступ как к исходным, так и обработанным данным по всей территории России, полученным в Европейском, Сибирском и Дальневосточном центрах приема. Система предоставляет пользователю не только возможности поиска и заказа данных, но и достаточно большой набор инструментов для анализа данных.

За счёт гибкой модульной архитектуры построения созданная система легко наращивается и позволяет достаточно просто включать в себя дополнительные функции и наборы данных.

В настоящее время основными направлениями развития системы являются:

– развитие функциональности интерфейсов работы с данными, которая должна в перспективе обеспечить полнофункциональные возможности обработки спутниковых данных для решения задач гидрометеорологии и мониторинга окружающей среды. Это в перспективе должно позволить изменить традиционную технологию работы с данными и предоставить специалистам, работающим в данной области, не только удобные возможности поиска и выбора данных, но и достаточно развитые средства их анализа, фактически совмещенные с оперативно пополняющимися архивами данных;

– расширение возможностей работы с данными перспективных КА Д33.

В заключение отметим, что при создании системы использовались технологии и базовое программное обеспечение, разработанное в ИКИ РАН (Лупян и др., 2004, Егоров и др., 2004, Андреев и др., 2004, Ефремов и др., 2004, Лупян, Мазуров и др., 2011). Элементы системы были созданы при поддержке проектов Росгидромета, Роскосмоса, при ее построении были также использованы методы и элементы технологий, созданные в рамках проектов РАН (тема «Мониторинг») и РФФИ (проекты 11-07-12028-офи-м-2011, 11-07-12026-офи-м-2011, 11-07-12025-офи-м-2011).

## **Литература**

1. *Андреев М.В., Ефремов В.Ю., Лупян Е.А., Мазуров А.А., Наглин Ю.Ф., Прошин А.А., Флитман Е.В.* Построение интерфейсов для организации работы с архивами спутниковых данных удаленных пользователей // Современные проблемы дистанционного зондирования Земли из космоса, Выпуск 1. 2004. с. 514-520.
2. *Андреев М.В., Лаврова О.Ю., Лупян Е.А., Мазуров А.А., Митягина М.И., Наглин Ю.Ф., Соловьев А.В., Нестеренко А.А., Проценко И.Г., Прошин А.А., Пырков В.Н.* Использование данных спутниковых РЛС для решения задач контроля позиционирования промысловых судов // Электронный журнал «Исследовано в России», 2007. № 021. С. 211-222.
3. *Асмус В.В., Бурцев М.А., Воронин А.А., Кузнецов А.Е., Лупян Е.А., Мазуров А.А., Милехин О.Е., Прошин А.А., Соловьев В.А., Успенский А.Б., Флитман Е.В., Хоменок Н.И.* Разработка автоматизированного комплекса приема, обработки и архивации данных геостационарных спутников в НИЦ «Планета» // Современные проблемы дистанционного зондирования Земли из космоса, 2006. Выпуск 3. Т. I. С. 156-162.
4. *Асмус В.В., Кровотынцев В.А., Милехин О.Е., Соловьев В.И., Успенский А.Б., Фролов А.В.* Проблемы и перспективы развития отечественных космических наблюдательных систем гидрометеорологического назначения // Тезисы докладов десятой всероссийской открытой ежегодной конференции «Современные проблемы дистанционного зондирования Земли из космоса» (Физические основы, методы и технологии мониторинга окружающей среды, природных и антропогенных объектов) Москва, ИКИ РАН, 12-16 ноября 2012 г.

5. Балашов И.В., Бурцев М.А., Ефремов В.Ю., Лупян Е.А., Прошин А.А., Толгин В.А. Построение архивов результатов обработки спутниковых данных для систем динамического формирования производных информационных продуктов // Современные проблемы дистанционного зондирования Земли из космоса, 2008. Выпуск 5. Т. I. С. 26-32.
6. Бурцев М.А., Воронин А.А., Еремеев В.В., Злобин В.К., Кузнецов А.Е., Лупян Е.А., Милехин О.Е., Соловьев В.И. Комплекс оперативной обработки гидрометеорологической спутниковой информации // Исследование Земли из космоса, 2009. № 1. С.16-23.
7. В.А. Егоров, В.О. Ильин, Е.А. Лупян, А.А. Мазуров, Р.Р. Прошин, Е.В. Флитман Е.В. Возможности построения автоматизированных систем обработки спутниковых данных на основе программного комплекса XV\_SAT // Современные проблемы дистанционного зондирования Земли из космоса, Выпуск 1. 2004. с. 431-436.
8. Ефремов В.Ю., Крашенинникова Ю.С., Лупян Е.А., Мазуров А.А., Прошин А.А., Флитман Е.В. Оптимизированная система хранения и представления географически привязанных спутниковых данных // Современные проблемы дистанционного зондирования Земли из космоса, 2007. Выпуск 4. Т. I. С. 125-132
9. Ефремов В.Ю., Лупян Е.А., Мазуров А.А., Прошин А.А., Флитман Е.В. Технология построения автоматизированных систем хранения спутниковых данных // Современные проблемы дистанционного зондирования Земли из космоса, 2004. Выпуск 1. С. 437-443.
10. Ефремов В.Ю., Лупян Е.А., Мазуров А.А., Прошин А.А., Флитман Е.В. Управление и контроль работоспособности систем автоматизированной обработки спутниковых данных // Современные проблемы дистанционного зондирования Земли из космоса. 2004. № 1. С. 467-475.
11. Лупян Е.А., Мазуров А.А., Назиров Р.Р., Прошин А.А., Флитман Е.В. Универсальная технология построения систем хранения спутниковых данных. М.: ИКИ РАН. Препринт Пр-2024. 2000. 22 с.
12. Лупян 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, Выпуск 1. 2004. С. 81-89.
13. Лупян Е.А., Мазуров А.А., Назиров Р.Р., Прошин А.А., Флитман Е.В., Крашенинникова Ю.С. Технологии построения информационных систем дистанционного мониторинга // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8. № 1. С. 26-43.
14. Лупян Е.А., Савин И.Ю., Барталев С.А., Толгин В.А., Балашов И.В., Плотников Д.Е. Спутниковый сервис мониторинга состояния растительности («ВЕГА») // Современные проблемы дистанционного зондирования Земли из космоса. 2011. Т. 8. № 1. С. 190-198.
15. Лупян Е.А., Саворский В.П. Базовые продукты обработки данных дистанционного зондирования Земли // Современные проблемы дистанционного зондирования Земли из космоса, 2012. Т. 9. № 2. С.87-97.
16. Коровин Г.Н., Котельников Р.В., Лупян Е.А., Щетинский В.Е. Основные возможности и структура информационной системы дистанционного мониторинга лесных пожаров Федерального агентства лесного хозяйства (ИСДМ Рослесхоз) // Современные проблемы дистанционного зондирования Земли из космоса, 2010. Т.7. № 2. С.97-105.
17. Нестеренко А.А., Романов А.А., Андреев М.В., Лупян Е.А. Общесистемное обеспечение отраслевой системы мониторинга Госкомрыболовства // Современные проблемы дистанционного зондирования Земли из космоса, 2004. Выпуск 1. Т.1. С. 204-212.
18. Носенко Ю.И., Лошкарев П.А. Единая территориально-распределенная информационная система Д33 – проблемы, решения, перспективы (часть 1) // Геоматика 2010. № 3. С. 35-43.
19. Прошин А.А., Бурцева Т.Н. Т.Н., Ефремов В.Ю., Лупян Е.А., Милехин О.Е., Мазуров А.А., Флитман Е.В., Ковалев А.Ф., Кормашева Т.Л. Автоматизированная система сбора, обработки и представления спутниковых данных НИЦ «Планета» // Современные проблемы дистанционного зондирования Земли из космоса, Выпуск 1. 2004. С.317-323.
20. Толгин В.А., Балашов И.В., Ефремов В.Ю., Лупян Е.А., Прошин А.А., Уваров И.А., Флитман Е.В. Создание интерфейсов для работы с данными современных систем дистанционного мониторинга (система GEOSMIS) // Современные проблемы дистанционного зондирования Земли из космоса, 2011. Т.8. № 3. С. 93-108.

# **Distributed satellite data processing products archives operation system in the SRC “Planeta” centres**

**M.A. Burtsev<sup>1</sup>, V.N. Antonov<sup>2</sup>, V.Yu. Efremov<sup>1</sup>, A.V. Kashnitskiy<sup>1</sup>, L.S. Kramareva<sup>2</sup>,  
E.A. Loupiian<sup>1</sup>, A.A. Mazurov<sup>1</sup>, A.M. Matveev<sup>1</sup>, O.E. Milekhin<sup>2</sup>,  
A.A. Proshin<sup>1</sup>, V.I. Solovyev<sup>2</sup>**

*<sup>1</sup> Space Research Institute of Russian Academy of Sciences  
117997, Moscow, Profsoyuznaya str., 84/32  
E-mail: evgeny@iki.rssi.ru;*

*<sup>2</sup> Scientific Research Center of Space Hydrometeorology «Planeta» (SRC «Planeta»)  
123242 Moscow, Bolshoy Predtechensky st. 7  
E-mail: asmus@planet.iitp.ru*

This paper describes the system for distributed satellite data processing products archives operation implemented in the regional centers of SRC “Planeta”. Basic tasks and requirements for the system are given along with the general structure of the system and its basic blocks description. Particular attention is paid to the distributed satellite data archives creation and access methods. Another important topic described in the paper is the structure and the capabilities of the cartographic web interface for data access and analysis. The conclusion of the paper gives the ways of the system further development.

**Keywords:** satellite data, distributed archives, information systems, monitoring systems, automated systems development technologies, data access interfaces.