

Организация распределенных файловых хранилищ для ведения сверхбольших архивов спутниковых данных ДЗЗ

Прошин А.А., Бурцев М.А., Радченко М.В.

Институт космических исследований Российской академии наук, Москва

Современные проблемы дистанционного зондирования Земли из космоса, 14-18 ноября 2022г

Развитие систем спутникового дистанционного зондирования Земли в последние десятилетия привело к взрывному росту объемов спутниковых данных, получаемых с большого числа различных приборов наблюдения. Расширяется область применения данных дистанционного зондирования Земли, которые в настоящее время широко используются для решения самых разных исследовательских и прикладных задач, связанных с мониторингом природной среды и антропогенных объектов. Это привело к существенному возрастанию требований к функционалу информационных систем доступа к спутниковым данным, а в первую очередь, к возможностям, которые они предоставляют по анализу и обработке спутниковых наблюдений.

В соответствии с этими новыми требованиями в 2012 году был создан центр коллективного пользования системами архивации, обработки и анализа данных спутниковых наблюдений ИКИ РАН для решения задач изучения и мониторинга окружающей среды «ЦКП ИКИ-Мониторинг» (<http://ckp.geosmis.ru/>). Ключевым элементом этого сложного программно-аппаратного комплекса является система ведения сверхбольших распределенных архивов спутниковых данных, которая обеспечивает не только архивацию и хранение больших массивов разнородных спутниковых данных, но и поддержку сервисов для их обработки и анализа. На текущий момент эта система обеспечивает доступ к архивам спутниковых данных, суммарный объем которых приближается к 6 петабайтам, при этом в сутки в архивы поступает более 3 терабайт новых данных.

Технология UNISAT. Организация хранения файлов данных

Система ведения архивов спутниковых данных ЦКП «ИКИ-Мониторинг» реализована на базе унифицированной технологии ведения сверхбольших распределенных архивов разнородных спутниковых данных UNISAT.

В рамках использования этой технологии метаданные, т.е. информация об имеющихся в архивах данных и их атрибуты, сохраняются в базу данных **unisat**, а сами файлы данных помещаются в файловое хранилище.

Структура БД **unisat** включает в себя две основные таблицы:

- **fragments** – описание фрагментов спутниковых продуктов;
- **fragment_products** – описание информационных продуктов, относящихся к фрагментам, которым соответствуют файлы данных, которые в основном представляют из себя спутниковые изображения в формате GeoTIFF.

Хранение файлов спутниковых данных в архивах ЦКП «ИКИ-Мониторинг» в каждом из центров реализуется на основе использования распределенных файловых хранилищ, состоящих из UNIX серверов и специализированных систем хранения (NAS), объединенных в единое логическое пространство при помощи протокола NFS. На текущий момент в ИКИ РАН задействовано более 30 серверов хранения

Файлы архива хранятся поддиректориях, в каждой из которых по умолчанию может быть до ста тысяч файлов. Имена поддиректорий и самих файлов формируются на основании идентификатора соответствующей записи в таблице **product_fragments**. К примеру, записи с идентификатором 12345678, описывающей GeoTIFF файл, соответствует относительный путь к файлу “00123/45678.tif”

Реализация доступа к файлам в архиве

За пополнение и изменение данных каждого из архивов отвечает выделенный сервер архивации. На нем располагается базовая директория хранилища файлов архива, в которую линкуются поддиректории с данными, находящимися на различных сетевых разделах. При этом сами разделы должны быть присоединены к этому серверу в режиме чтения и записи по протоколу NFS.

На серверах, которым необходим только доступ к файлам в архиве, присоединяется по NFS базовая директория архива, а требуемые сетевые разделы присоединяются по протоколу NFS в режиме только чтения. К таким серверам, в первую очередь, относятся сервера, используемые для обеспечения доступа к данным пользователей, а также сервера, на которых реализована обработка данных, находящихся в этом архиве.

На серверах, которые обеспечивают формирование данных для занесения их в архив, присоединяется по NFS специальная директория с сервера архивации, которая всегда указывает на текущий раздел пополнения данными, а также сам этот раздел и следующие за ним. Это позволяет избежать лишнего копирования файлов и помещать их сразу на требуемый сетевой раздел, после чего в рамках процедур архивации они перемещаются внутри этого раздела в поддиректории с данными, что является нетребовательной к ресурсам операцией.

Поддержка структуры поддиректорий архива

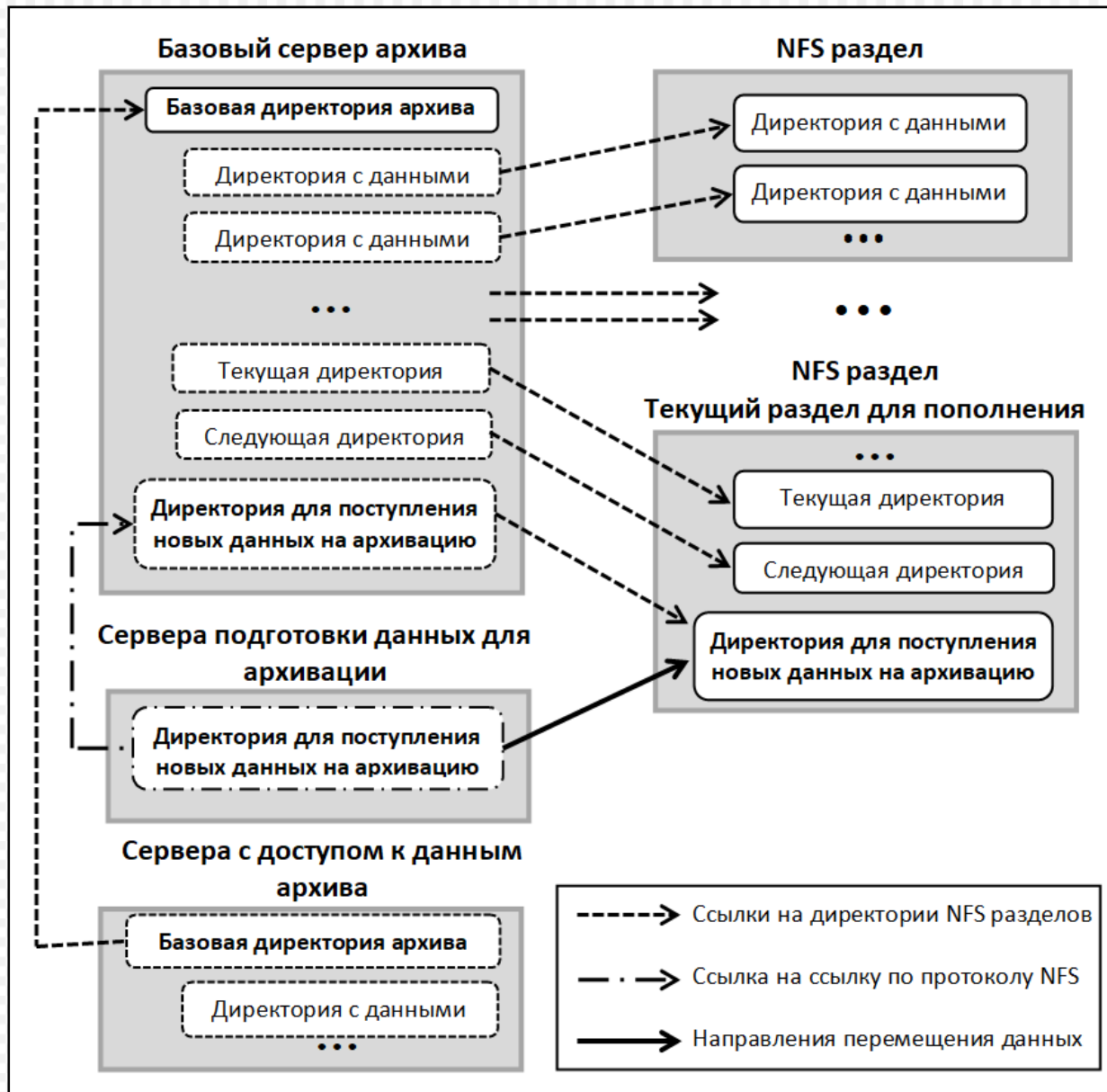
За создание поддиректорий на серверах хранения и ссылок на них отвечает специальная программа **manage_store_subdirs**, устанавливаемая на выделенных серверах архивации. Она запускается перед каждым запуском процедур пополнения данными или импорта данных из других информационных центров.

Программа обеспечивает наличие как минимум одной следующей поддиректории, в которую может осуществляться пополнение файлов. Число таких директорий определяется соответствующими настройками.

При заполнении текущего сервера хранения поддиректории создаются на следующем сервере, на котором имеется достаточное свободное пространство.

В случае достижения минимального лимита на свободное пространство на текущем сетевом разделе, указанного в настройках программы, форсируется переход на поддиректорию, располагающуюся на другом сервере хранения. Это достигается путем изменения текущего значения идентификатора в таблице **fragment_products**, используемого для формирования пути к файлу данных в хранилище.

Схема организации распределенного файлового хранилища



Организация автоматического монтирования требуемых сетевых разделов

Для автоматического монтирования NFS разделов отвечает специальная процедура **archmount**, периодически выполняемая на каждом из серверов, которым необходим доступ к данным .

В рамках централизованно рассылаемых настроек указывается состав разделов, используемых для хранения данных каждого из архивов, определяются списки архивов, доступ к которым необходим тому или иному серверу, а также режим их монтирования (запись или только чтение).

В случае неполадок в работе сервера хранения во избежание неполадок производится отключение соответствующих ему разделов. При этом после возобновления работы сервера хранения монтирование автоматически возобновляется. Это позволяет избежать потенциально проблемных ситуаций, когда подсоединенный по сети раздел не функционирует штатно, что может приводить к зависанию системных процессов на сервере.

Заключение

Работы по развитию технологий ведения сверхбольших архивов спутниковых данных с 2019 года выполняются в рамках темы Минобрнауки РФ «Большие данные в космических исследованиях: астрофизика, солнечная система, геосфера» (№122042500019-6) с использованием возможностей ЦКП «ИКИ-Мониторинг» (<http://ckp.geosmis.ru/>).