

Технология динамического блочного представления спутниковых данных системам распределённой обработки

А. А. Прошин, Е. А. Лупян, И. В. Балашов, А. В. Кашницкий,
А. М. Матвеев, Б. П. Руткевич

*Институт космических исследований РАН, Москва, 117997, Россия
E-mail: andry@smis.iki.rssi.ru*

Во многих задачах, связанных с обработкой и визуализацией пространственных данных, значительный выигрыш даёт использование блочного (тайлового) подхода, в рамках которого используются данные, представленные в виде блоков, на которые равномерно разбита вся зона интереса. Это позволяет эффективно организовать параллельную обработку данных на различных вычислительных узлах, использовать механизм кэширования и снизить нагрузку на централизованные ресурсы. При этом, как правило, используются уже подготовленные архивы данных с фиксированным пространственным разбиением. В то же время во многих случаях, в первую очередь когда необходимо совместно обрабатывать различные типы спутниковой информации, использование подготовленных архивов данных может быть нецелесообразным. Это связано с тем, что различные типы спутниковых данных существенно различаются как по пространственному разрешению, так и по схемам организации их хранения. В результате практически невозможно выбрать одно пространственное разбиение, которое было бы оптимальным для организации работы со всеми типами данных. Поэтому требуется разработка таких подходов к организации обработки спутниковых данных, которые могли бы максимально использовать преимущества блочного метода и оптимизации схемы хранения каждого конкретного типа данных. В статье представлен такой подход организации обработки спутниковых данных с использованием динамического блочного доступа к разнородным архивам спутниковых данных, в рамках которого необходимые для обработки данные формируются по запросу в таком пространственном разбиении и с такими характеристиками, которые будут оптимальны для проведения конкретного типа обработки. В работе приводится краткое описание разработанной технологии, а также рассматривается методика оценки оптимальных параметров разбиения для конкретной обработки и имеющейся инфраструктуры.

Ключевые слова: параллельная обработка спутниковых данных, архивы спутниковых данных, большие данные, ЦКП «ИКИ-Мониторинг»

Одобрена к печати: 10.12.2020
DOI: 10.21046/2070-7401-2020-17-7-79-93

Проблема и постановка задачи

Во многих задачах, связанных с обработкой и визуализацией пространственных, в первую очередь спутниковых, данных, значительный выигрыш в организации работы с ними даёт блочный (тайловый) подход. Он заключается в том, что информация передаётся компьютерам, на которых ведётся визуализация, анализ и/или обработка, в виде некоторых блоков (тайлов, гранул), на которые равномерно разбита вся зона интереса, покрытая данными. Такой подход в том числе позволяет:

- эффективно распараллеливать процесс обработки и анализа данных по значительным территориям;
- повысить эффективность кэширования данных в тех случаях, когда один и тот же фрагмент используется несколько раз, в том числе разными пользователями;
- понизить нагрузку на централизованные ресурсы за счёт того, что компьютеры, ведущие обработку и анализ данных, получают возможность кэширования данных в своих локальных архивах, а задачи решаются с использованием только собственных ресурсов.

Также блочный подход позволяет сокращать трафик между централизованными ресурсами и компьютерами, ведущими обработку и анализ данных, в первую очередь за счёт сокращения многократной передачи одной и той же информации.

Следует отметить, что в задачах визуализации пространственных данных блочный подход в настоящее время очень широко распространён, и его используют практически все системы, обеспечивающие массовый доступ пользователей к различным пространственным данным (Gorelick et al., 2017; Peterson, 2016).

Сегодня основным путём реализации блочного подхода является использование специально организованных архивов спутниковых данных, которые обеспечивают хранение различных данных в едином пространственном разбиении и проекции. Это позволяет реализовать унифицированные механизмы доступа к различным данным в архивах, существенно упростить и ускорить процедуры их получения (включая передачу и локальное кэширование), визуализации и обработки. Этот подход достаточно хорошо работает в случаях, когда нужно оперировать с относительно статическим и во многом однородным набором данных, для которых может быть выбрана единая схема организации хранения и представления. Поэтому подобные подходы успешно используются в системах, основной задачей которых является визуализация (представление) относительно ограниченного числа готовых (статических) информационных продуктов.

В то же время для проведения обработки и анализа данных по большим территориям, особенно за длительный период времени, с использованием информации различных систем наблюдения такой подход во многих случаях трудно реализуем. Это связано в том числе со следующими особенностями организации систем хранения спутниковых данных и результатов их обработки:

- оптимальные схемы организации хранения данных систем, обладающих различным разрешением и схемами наблюдения, могут существенно отличаться друг от друга. То есть, переходя в какую-то единую схему хранения данных, мы не можем обеспечить одинаково эффективную её работу для разных типов данных;
- значительное число архивов различных спутниковых систем имеют свою структуру и особенности, поэтому перевод их всех в какой-то регулярный блочный вид требует значительных ресурсов;
- практически невозможно выработать такое регулярное разбиение, которое было бы одинаково эффективно для разных типов данных и для разных задач по их обработке и анализу.

В то же время блочный подход организации обработки спутниковых данных, безусловно, имеет свои неоспоримые преимущества. Одним из главных является возможность эффективного распараллеливания процессов обработки и анализа данных благодаря задействованию в них значительного числа одновременно работающих вычислительных ресурсов. Однако, как уже обсуждалось выше, приведение всех имеющихся архивов в единый «блочный» вид имеет свои недостатки, а главное, является достаточно ресурсоёмкой задачей. Тем не менее в подавляющем числе случаев, когда стоит задача обработки и анализа информации по большим территориям, сам процесс обработки всё-таки приходится организовывать на основе блочного подхода. При этом на вычислительные узлы, которые ведут обработку данных, передаётся вся информация из архивов некими фрагментами хранения (сценами), которые используются в конкретных архивах. Выборка же информации, соответствующей конкретному обрабатываемому блоку, производится на вычислительном узле. Это приводит к тому, что на вычислительные узлы передаются значительные избыточные объёмы информации. Характерный пример, когда для проведения обработки блока данных необходима спутниковая информация из различных источников, для хранения которой используются разные географические проекции и разбиения, приведён на *рис. 1* (см. с. 81). На нём представлена ситуация, возникающая при подготовке данных для проведения радиометрической и геометрической коррекции данных прибора КМСС (комплекс многозональной спутниковой съёмки) (<http://russianspacesystems.ru/bussines/dzz/meteor-m>) с использованием данных приборов

MODIS (*англ.* Moderate Resolution Imaging Spectroradiometer) (<https://terra.nasa.gov/about/terra-instruments/modis>) и OLI-TIRS (*англ.* Operational Land Imager и Thermal Infrared Sensor) (<https://www.usgs.gov/core-science-systems/nli/landsat/landsat-8>). На рисунке различными цветами показаны контуры фрагментов разных типов спутниковых данных: чёрным — регулярная сетка разбиения на блоки, красным — одна из ячеек разбиения. Получается, что для обработки информации по выделенной ячейке необходимо использовать данные семи различных сцен спутниковых данных. Все эти сцены придётся передать на вычислительный узел, ведущий обработку, при этом значительный объём передаваемой информации не будет использован при проведении обработки.

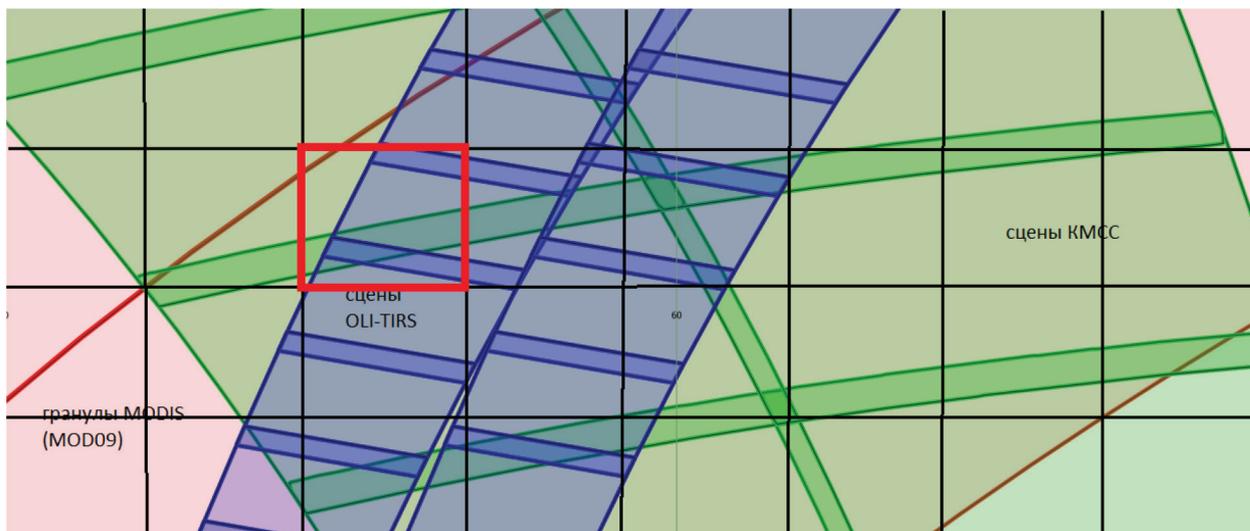


Рис. 1. Пример использования различных типов спутниковых данных в обработке

Альтернативой представленному методу может быть динамический блочный подход, который обеспечит представление на конкретный вычислительный узел только той информации по блоку данных, которую на нём планируется обработать. Это, с одной стороны, позволяет избежать предобработки архивов, а с другой — обеспечить предоставление данных для конкретной обработки в оптимальном именно для неё блочном виде. Для реализации динамического блочного доступа необходимо реализовать следующие механизмы:

- разбиения региона на блоки, по которому выполняется конкретная обработка (с учётом выбора оптимальных для обработки параметров);
- подготовки наборов данных для обработки (выбор необходимой информации по блоку);
- распределения блоков (или кластеров блоков) по имеющимся вычислительным ресурсам;
- кэширования блоков в случаях, когда они могут быть использованы неоднократно.

Фактически такой механизм динамического блочного доступа позволит оптимизировать представление данных для различных задач обработки и избежать необходимости переструктурирования имеющихся архивов данных под конкретные задачи.

В настоящей работе представлена технология динамического блочного доступа к архивам спутниковых данных Центра коллективного пользования (ЦКП) «ИКИ-Мониторинг» (Лупян и др., 2015, 2019) (<http://ckp.geosmis.ru/>). Особое внимание уделено особенностям задачи обработки данных по большим территориям, а также методике выбора оптимального разбиения зоны обработки на блоки и кластеры данных, обрабатываемые на отдельных вычислительных узлах.

Реализация предлагаемой технологии

Реализация представляемой в настоящей работе технологии динамического блочного доступа к данным базируется на использовании технической и программной инфраструктуры ЦКП «ИКИ-Мониторинг». Архивы центра разработаны и поддерживаются на основе технологии построения сверхбольших распределённых архивов разнородных спутниковых данных UNISAT (Прошин и др., 2016). Подсистемы обработки данных реализованы на базе технологии распределённой многопоточковой обработки спутниковых данных (Кобец и др., 2015, 2018), которая позволяет в полностью автоматическом режиме вести обработку на практически неограниченном числе вычислительных узлов. В рамках этой технологии за управление процессом обработки отвечают выделенные серверы диспетчеризации, на которых формируются задания для обработки на вычислительных узлах. При этом на самих вычислительных узлах устанавливается стандартное программное обеспечение (ПО), разработанное в Институте космических исследований РАН (ИКИ РАН) (Егоров и др., 2004; Кобец и др., 2015) или являющееся свободно распространяемым, которое позволяет выполнять широкий спектр различных типов обработки. Каждое задание содержит описание действий, которые необходимо произвести над набором данных, и информацию о местонахождении обрабатываемых файлов данных. Освободившиеся вычислительные узлы в соответствии с централизованно определяемыми приоритетами получают на сервере диспетчеризации очередное задание на обработку, выполняют его, а результаты передают на серверы архивации спутниковых данных, указанные в задании на обработку. Для управления и контроля за всеми автоматически выполняемыми процедурами обработки используются подходы, описанные в работах (Балашов и др., 2011; Кобец и др., 2018), которые, в частности, позволяют оперативно детектировать и диагностировать широкий спектр различных ошибок, оповещать о них персонал и отслеживать процесс их устранения.

Следует отметить, что обработка больших массивов спутниковых данных может требовать весьма значительных вычислительных ресурсов, высокопроизводительных систем хранения и каналов передачи данных. Поэтому задачи минимизации накладных расходов и повышения эффективности при реализации механизма динамического блочного доступа выходят на первый план. Это следует учитывать при построении эффективных сервисов доступа (предоставления) данных. Поэтому основным принципом построения таких сервисов является реализация процедур выборки конкретных данных с помощью вычислительных ресурсов, «максимально близких» (в сетевом смысле) к конкретным системам хранения данных. При этом для минимизации сетевого трафика может также быть предусмотрена возможность работы с различными буферами, обеспечивающими кеширование информации. Также при подготовке данных по блокам предпочтительнее использовать прямой доступ к файлам, позволяющий вычитывать только необходимое подмножество данных из них.

На *рис. 2* (см. с. 83) приведена принципиальная схема организации обработки с использованием динамического блочного доступа к данным. Серверы диспетчеризации отвечают за формирование заданий на обработку данных. В рамках этой операции сначала в соответствии с конфигурацией конкретной процедуры обработки запрашивается информация о данных (в общем случае различных типов), которые должны быть обработаны. Затем формируются запросы к сервису подготовки данных по блокам заданного разбиения, в которых содержится вся необходимая информация о местонахождении данных. После этого сами запросы в многопоточном режиме направляются на сервер переадресации запросов и балансировки нагрузки. На этом сервере устанавливается специально разработанный демон, позволяющий переадресовать запросы на соответствующие локальные или удалённые серверы подготовки данных.

В качестве серверов подготовки данных по ячейкам разбиения могут выступать как выделенные серверы с высокой скоростью доступа к архивам, так и непосредственно серверы, на которых располагаются данные. На каждом из них устанавливается стандартный сервис подготовки данных по заданной области (блоку или кластеру блоков). Реализованный демон позволяет гибко настраивать правила балансировки нагрузки на серверы подготовки данных.

В частности, при распределении запросов они в приоритетном порядке направляются на серверы, на которых непосредственно располагаются требуемые данные. Также для каждого из серверов может быть указан его вес, в конечном итоге определяющий количество направляемых на него запросов. После получения всех требуемых данных в дисковый буфер производится формирование заданий для серверов обработки. Вычислительные узлы получают сформированные задания от сервера диспетчеризации, скачивают данные по заданным блокам разбиения из дискового буфера и приступают к выполнению процедуры обработки. После завершения обработки результаты и описывающие их метаданные передаются на серверы архивации для занесения их в архивы центра.

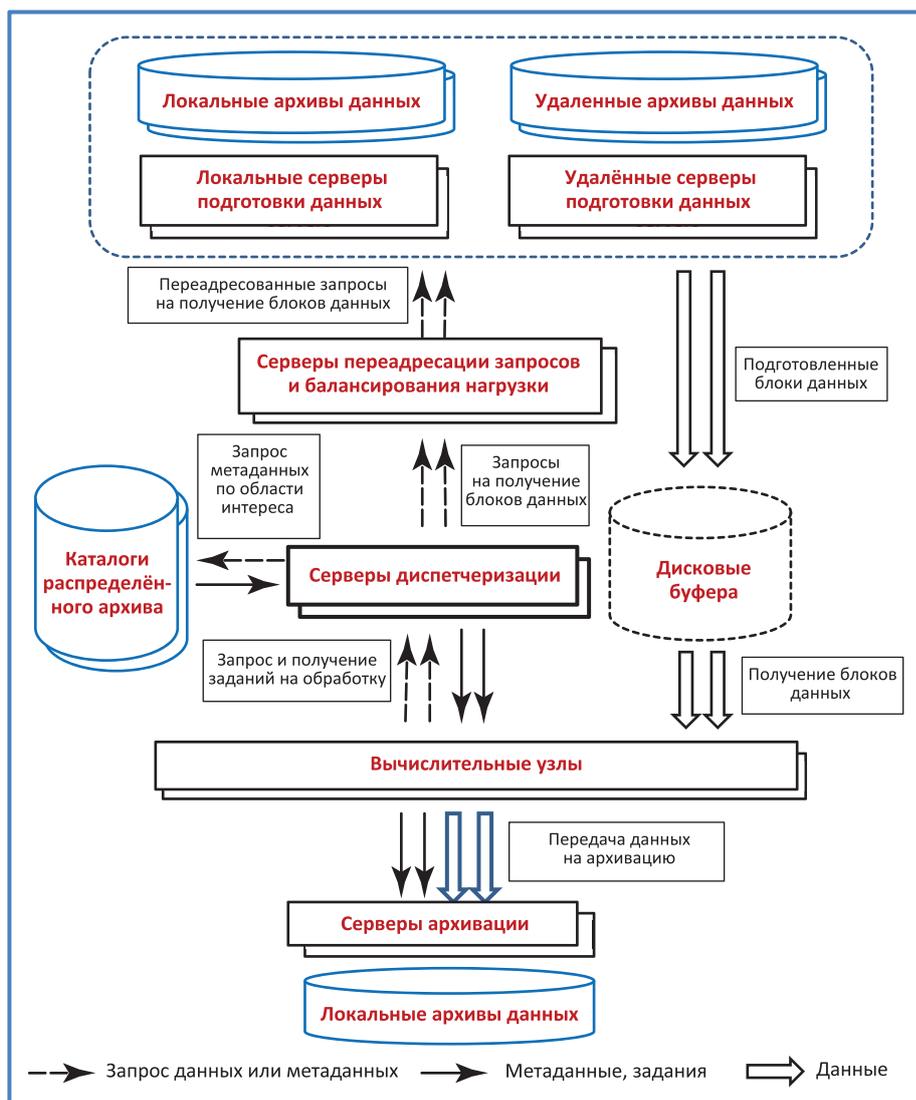


Рис. 2. Принципиальная схема организации обработки на основе динамического блочного доступа к данным

Реализованная технология позволяет предоставлять для выполнения обработки в блочном виде любые имеющиеся в архивах ЦКП «ИКИ-Мониторинг» спутниковые данные, для которых могут быть заданы следующие основные параметры: разбиение на блоки, проекция, разрешение, набор каналов и т.д. Это, как уже было отмечено выше, позволяет обеспечить динамический блочный доступ к данным в таком представлении, которое может быть максимально эффективным для решения конкретной задачи по обработке данных. Гибкая система конфигурации позволяет легко задать набор необходимых для проведения обработки данных

по заданной области интереса. С помощью реализованного механизма кэширования можно при необходимости использовать подготовленные по блокам данные многократно.

Непосредственно для реализации рассматриваемого в настоящей статье подхода были разработаны следующие программные компоненты:

- CGI-утилита для формирования данных по заданной ячейке разбиения, базирующаяся на использовании возможностей ПО GDAL (<https://gdal.org/>);
- демон для переадресации запросов и балансировки нагрузки на серверы подготовки данных;
- программа формирования запросов на получение данных по ячейкам разбиения для заданных наборов различных типов спутниковых данных;
- программа для многопоточкового скачивания данных по ячейкам на базе ПО Aria2c (<https://aria2.github.io/>);
- адаптированная для использования блочного доступа к данным процедура формирования заданий на обработку.

В настоящее время предложенная схема реализована в ЦКП «ИКИ-Мониторинг» и показала свою работоспособность.

Особенности организации блочной обработки данных по большим территориям

Как уже было отмечено выше, анализ данных по большим территориям может быть эффективно реализован только при организации параллельной обработки с использованием множества вычислительных узлов. В простейшем варианте, когда процедура обработки использует только операции с отдельными пикселями спутниковых изображений, для её распараллеливания может быть использовано разбиение области интереса на непересекающиеся блоки данных. При этом оптимальное число блоков в их кластерах (набор блоков, передающихся для обработки на вычислительный узел) определяется в конечном итоге числом блоков заданного размера, которое конкретная процедура может обрабатывать в параллельном режиме на характерном вычислительном узле.

Однако во многих процедурах обработки используются не только данные в отдельных точках, но также данные в соседних к ним областях. Такие данные необходимы, например, для отсекающей теней от облаков, улучшения калибровки, интерполяции данных и дополнительной привязки данных по наблюдениям более высокого пространственного разрешения. В результате возникает необходимость в предоставлении для обработки не только данных по области, по которой будет получен результат обработки конкретного кластера, но и данных по требуемой буферной зоне. Это приводит к тому, что часть данных используется многократно. Иллюстрация этого приведена на *рис. 3* (см. с. 85). На нём красными контурами показаны области, по которым получается результат конкретного блока, чёрными — граница области, по которой для этого должны быть представлены данные (включая буфер). Штриховкой показаны участки данных, которые при обработке полной области интереса используются два и более раз.

Наиболее простым способом предоставления буферных данных является использование расширенных пересекающихся друг с другом блоков данных, т. е. блоков, включающих в себя необходимые «соседние» данные. Однако такой вариант приводит к дублированию подготовки и передачи по сети одних и тех же областей данных. В тех случаях, когда необходимо получить данные из других информационных центров по каналам с ограниченной пропускной способностью, это может быть неэффективным. В случае, когда обработка на вычислительных узлах ведётся кластерами блоков, наиболее простым вариантом организации буферной зоны данных, на наш взгляд, является просто использование в составе кластера обработки дополнительных блоков, включающих требуемую буферную зону. Это, с одной стороны, не требует проведения дополнительной выборки данных из архивов, а с другой — до некоторой степени сокращает объёмы передаваемой на вычислительные узлы информации. Далее

мы будем рассматривать именно такой вариант. Пример формирования кластеров для обработки в таком случае приведён на *рис. 4*. В то же время следует учитывать, что в тех случаях, когда число блоков, которые покрывают буферную область, составляет значительную часть от числа блоков, входящих в кластер, передаваемый для обработки на вычислительный узел, возникают существенные накладные расходы, связанные с дублированием передачи данных из дискового буфера вычислительным узлам.

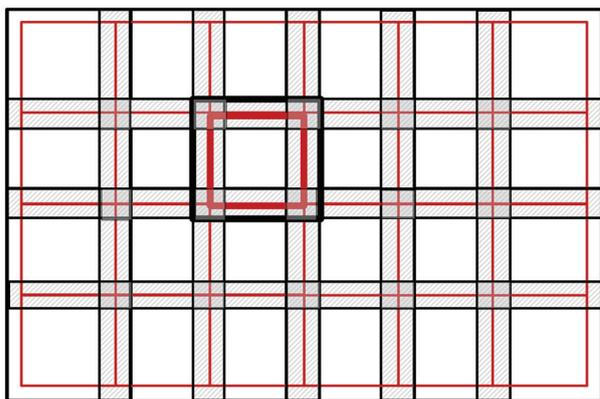


Рис. 3. Схема предоставления «буферных» данных для обработки

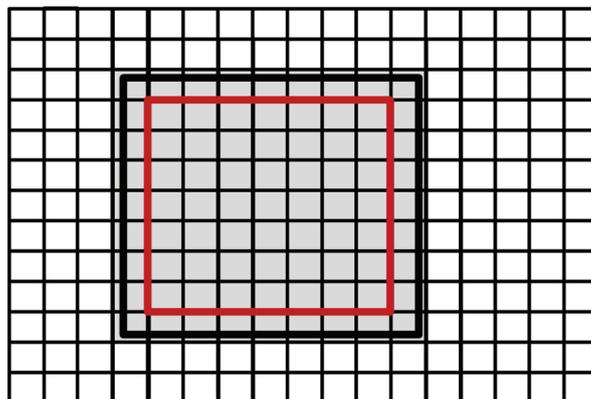


Рис. 4. Организация буферной зоны с использованием набора дополнительных блоков

При использовании такой схемы формирования наборов данных для обработки на одном вычислительном узле встаёт вопрос оптимизации размеров как блоков, так и кластеров. Это должно позволить выбрать оптимальное число используемых вычислительных узлов с учётом того, что при увеличении их числа происходит уменьшение размера кластера и, соответственно, рост числа неоднократно обрабатываемых блоков. Отметим, что такую оптимизацию необходимо проводить для каждой вычислительной процедуры с учётом особенностей обрабатываемых данных, имеющихся в наличии вычислительных ресурсов, требований ко времени, за которое должна быть проведена обработка, и т. д. Подход к решению данной задачи будет рассмотрен в следующем разделе.

Выбор оптимальных разбиений на блоки и кластеры обработки

Для выбора оптимальных параметров разбиения обрабатываемой области на блоки и кластеры необходимо оценить зависимости временных затрат на проведение различных операций с данными от значения этих параметров. Сразу отметим, что размеры области, по которой ведётся обработка, а также блоков удобнее всего определять в единицах географических координат, например в градусах (как мы будем делать в дальнейшем). Это особенно важно, когда при обработке используются данные различного пространственного разрешения. Выбор таких единиц измерения размеров позволяет избегать необходимости поиска оптимальных размеров для каждого типа данных различного пространственного разрешения. Необходимость в этом возникала бы в том случае, когда размеры областей и блоков определялись, например, в пикселях. Весь процесс обработки можно поделить на два основных этапа: этап подготовки данных и этап обработки.

Этап подготовки данных

На этапе подготовки данных серверы диспетчеризации производят запросы к каталогам распределённых архивов на получение требуемых метаданных. На основании этих метаданных формируются запросы на получение данных по блокам, после чего они направляются на сервер

переадресации, который перенаправляет их на серверы подготовки данных. По мере получения всех требуемых блоков данных в дисковый буфер на сервере диспетчеризации формируются задания для обработки на отдельных вычислительных узлах. Среди перечисленных операций существенного времени и ресурсов, как правило, требует только сам процесс подготовки (выборки) данных и передачи их по сети в дисковый буфер, поэтому временными затратами на остальные операции в первом приближении можно пренебречь. Суммарное время получения заданного числа блоков по используемому в обработке набору разных типов спутниковых данных можно описать формулой:

$$T_{prep}(a, n) = n \left(\sum_{i=1}^k T_{fi}(a) + \sum_{i=1}^k T_{di}(a) \right),$$

где a — размер блока; n — количество блоков; i — номер типа данных, использующихся в обработке; k — количество разных типов данных; $T_{fi}(a)$ — среднее время формирования блока i -го типа спутниковых данных, эта функция может быть получена экспериментально для каждого типа данных; $T_{di}(a)$ — среднее время получения блока i -го типа данных по сети в дисковый буфер, которое также вычисляется экспериментальным путём.

Серверы подготовки данных, как локальные, так и удалённые, одновременно используются для обеспечения данными разных процессов обработки. При этом скорость отдачи ими данных напрямую зависит от конкретных каналов связи и средней загруженности как самих серверов, так и канала связи. Для упрощения в рамках рассматриваемой методики сервисы подготовки данных считаются фиксированным внешним фактором, поэтому зависимость времени подготовки блока от его размера $T_{prep}(a)$ для конкретной обработки может определяться либо на реальной конфигурации вычислительной системы, использующей для её реализации, либо на имитационном стенде, учитывающем особенности использующей вычислительной среды.

Этап обработки данных

Оценку временных затрат на получение заданием обработки блоков данных из буфера, их обработку и передачу результатов на сервер архивации можно представить в виде следующей формулы:

$$T_{proc}(a, p) = (p + 2)^2 T_{get}(a) + p^2 (T_{proc}(a) + T_{upl}(a)), \quad (1)$$

где a — размер блока; p — число блоков, входящих в кластер, без учёта буферной зоны; $T_{get}(a)$ — среднее время скачивания блока по всем используемым типам данных на сервер обработки; $T_{proc}(a)$ — среднее время обработки данных блока, включая чтение и запись данных; $T_{upl}(a)$ — среднее время передачи блока на сервер архивации.

Зависимость функций $T_{get}(a)$, $T_{proc}(a)$, $T_{upl}(a)$ от размера блока a в общем случае не может быть описана аналитически, поэтому должна вычисляться экспериментальным путём для конкретной обработки.

Следует отметить, что в общем случае также необходимо оценивать и учитывать нагрузку на сетевую инфраструктуру, обеспечивающую работу вычислительных узлов. Но в первом приближении для решения задачи выбора именно оптимальных размеров используемых блоков и их кластеров данными особенностями процессов обработки мы пренебрежём.

В то же время для оценки ограничений количества используемых вычислительных узлов необходимо оценить нагрузку на канал доступа к дисковому буферу и канал передачи данных на архивацию, которая создаётся при выполнении одного задания на обработку. Такая нагрузка может быть оценена следующим образом:

$$V_{iget}(a, p) = \frac{SourceUnitSize(a) \cdot (p + 2)^2}{T_{iproc}(a, p)}, \quad (2)$$

$$V_{upl}(a, p) = \frac{ResultUnitSize(a) \cdot p^2}{T_{proc}(a, p)}, \quad (3)$$

где $V_{get}(a, p)$ — средний трафик, создаваемый одним заданием на канал получения данных из дискового буфера; $V_{upl}(a, p)$ — средний трафик, создаваемый заданием на канал отдачи результатов обработки; $SourceUnitSize(a)$ — средний объём всех типов исходных данных для блока размера a ; $ResultUnitSize(a)$ — средний объём результатов, получаемых при обработке блока с размером a .

Важно отметить, что для равномерной нагрузки на сетевую инфраструктуру при проведении параллельной обработки на множестве вычислительных узлов необходимо тем или иным образом разводить по времени выполнение операций по получению данных и возврату получаемых результатов различными вычислительными узлами.

Восстановление экспериментальных зависимостей времени различных процессов от размера блока данных

При экспериментальном восстановлении обсуждаемых зависимостей для сокращения времени их оценки целесообразно определить разумные для конкретной реализации процедуры обработки размеры блоков. Например, для процедуры обработки, требующей формирования буферной области, минимальный размер блока может быть связан с её размером. Максимальный размер блока зависит от конкретной реализации процедуры обработки и возможности использующихся вычислительных узлов, в первую очередь имеющейся на них оперативной памяти. Для получения требуемых зависимостей в пределах выбранного диапазона обычно используется ряд значений размера блока с фиксированным шагом, от которого в конечном итоге зависит точность выбора оптимальных параметров, обеспечивающих минимизацию суммарного времени обработки.

Для того чтобы избежать влияния на экспериментально восстанавливаемые зависимости неравномерности загрузки реальной вычислительной среды, в которой производится обработка, замеры скорости выполнения различных операций и её зависимости от размеров блоков целесообразно проводить на специально созданном стенде. Это, с одной стороны, не должно изменить вида получаемых зависимостей, а с другой — позволяет отчасти сократить время тестирования, в частности за счёт отсутствия необходимости проводить долгое накопление информации для исключения эффектов неоднородной во времени нагрузки реальной вычислительной среды. Поэтому тестирование скорости выполнения обработки целесообразно полностью осуществлять на специализированном стенде в режиме, в котором одновременно проводится тестирование процедуры обработки только одного типа и отсутствует другая вычислительная нагрузка. Такой стенд был создан в рамках ЦКП «ИКИ-Мониторинг». В его состав вошли следующие серверы:

- сервер диспетчеризации, на котором реализуется дисковый буфер;
- сервер обработки, производительность которого по возможности должна примерно соответствовать серверам, которые предполагается использовать в реальной конфигурации.

При проведении тестирования для каждого из анализируемых значений размера блока на сервере обработки производятся тестовые запуски различных процедур, в результате которых определяются следующие зависимости: $T_{get}(a)$, $T_{proc}(a)$, $T_{upl}(a)$. При этом оценивается не время выполнения отдельных заданий в рамках отдельных вычислительных потоков (с учётом, что практически все вычислительные узлы центра являются многопроцессорными), а эффективные значения при выполнении обработки в многопоточковом режиме. В базовом варианте сначала в одном потоке закачиваются все необходимые блоки данных, затем они обрабатываются в параллельном режиме, после чего результаты обработки уже одним потоком передаются на сервер архивации. В этом случае среднее значение для этих операций может быть получено как суммарное время их выполнения, делённое на количество

блоков, а время обработки одного блока — как время многопоточковой обработки, делённое на число блоков. Также при проведении тестирования оцениваются характерные значения $SourceUnitSize(a)$ и $ResultUnitSize(a)$, которые в дальнейшем используются для оценки среднего трафика, возникающего при обработке в используемой сетевой инфраструктуре с помощью выражений (2) и (3).

Оценку времени подготовки данных следует проводить с использованием реальных систем хранения. Поэтому зависимость времени подготовки данных $T_{prep}(a)$ мы исследовали на реальной инфраструктуре ведения сверхбольших распределённых архивов спутниковых данных ЦКП «ИКИ-Мониторинг». Так как нагрузка на эту инфраструктуру может достаточно сильно меняться во времени, то для получения $T_{prep}(a)$ необходимо многократно провести запросы данных и получить средние значения для конкретных a . Для автоматизированного проведения таких запросов в течение продолжительного времени также использовался созданный стенд. При этом следует учитывать, что количество блоков, запрашиваемых в каждом отдельном тесте, должно быть достаточным, чтобы в полной мере активировать параллельный режим подготовки данных на группах серверов подготовки данных.

Выбор оптимальных значений размеров блоков

В настоящем разделе мы рассмотрим вопрос выбора оптимального значения размера блока при обработке области данных площадью S . Фактически мы должны минимизировать $T_{area}(S, a)$, которое в общем случае представляется выражением:

$$T_{area}(S, a) = \frac{S}{a^2} \left[T_{prep}(a) + \frac{1}{p^2 N} \left((p+2)^2 T_{get}(a) + p^2 (T_{proc}(a) + T_{upl}(a)) \right) \right], \quad (4)$$

где кроме описанных выше параметров ещё фигурирует параметр N — число используемых вычислительных узлов.

В общем случае оптимальное число используемых вычислительных узлов также зависит от размера блока, и это следует учитывать при решении полной оптимизационной задачи, направленной на определение размера блока a , при котором суммарное время обработки области площадью S будет минимально.

В то же время при решении поставленной задачи могут быть сделаны некоторые упрощения, которые помогут проанализировать особенности её решения. Одно из таких упрощений связано с отказом от учёта ограничений возможностей вычислительной сети и использования предположения, что оптимальным числом вычислительных узлов для конкретного разбиения, которое определяется параметрами a и p , является такое, при котором имеется возможность одновременного запуска обработки всех кластеров блоков, на которые разбита область S . Для выполнения этого необходимо, чтобы выполнялось условие:

$$N = \frac{S}{(pa)^2}. \quad (5)$$

В этом случае выражение (4) может быть преобразовано в вид:

$$T_{area}(S, a) = \frac{S}{a^2} T_{prep}(a) + \left((p+2)^2 T_{get}(a) + p^2 (T_{proc}(a) + T_{upl}(a)) \right). \quad (6)$$

Конечно, это выражение имеет некоторые ограничения, связанные в том числе с тем, что физически N должно быть целым числом, однако, на наш взгляд, это несущественно понижает точность оценки. Используя эту зависимость, можно решить оптимизационную задачу и определить значения размера блока a_{opt} и размера кластера p_{opt} , при которых суммарное время обработки области площадью S будет минимальным.

При этом при проведении оптимизации мы должны также учитывать различные ограничения, связанные, в частности, с возможностями используемой сетевой инфраструктуры. Поэтому необходимо минимизировать $T_{area}(S, a)$ с учётом интегрального пикового

трафика, который может быть оценён путём умножения $V_{iget}(a, p(a))$, $V_{upl}(a, p(a))$ (см. выражения (2) и (3)) на число используемых вычислительных узлов. В принципе, эта задача численно может решаться и в самом общем виде, поскольку для конкретной обработки используемой сетевой и вычислительной инфраструктуры, а также системы хранения и доступа к архивам спутниковых данных, как было показано выше, могут быть экспериментально получены все необходимые зависимости.

Выражение (6) позволяет нам также проанализировать динамику временных затрат как в предельных случаях малых, так и больших a , при условии, что мы не рассматриваем ограничения, связанные с возможностями вычислительной сети. В этом случае при больших a (в пределе, когда a равно размеру всей обрабатываемой области) $p = 1$, поскольку в обрабатываемой области остаётся только один блок. Число вычислительных узлов, которые можно использовать в этом случае, становится равным единице. То есть время, которое в этом пределе необходимо для обработки всей области интереса, становится равным времени обработки на одном вычислительном узле. При больших размерах обрабатываемой области это время может быть очень значительным. В то же время при малых a , из-за того что величина слагаемых, отвечающих за обработку данных при уменьшении a фактически стремится к 0 (это связано с тем, что скорость обработки малого объёма данных на отдельном компьютере становится очень высокой), выражение (6) фактически приобретает вид:

$$T_{area}(S, a) = \frac{S}{a^2} T_{prep}(a). \quad (7)$$

Важно отметить, что удельная на единицу объёма скорость выполнения практически всех операций с данными катастрофически падает при сильном уменьшении размера блока a . В результате, несмотря на рост выделяемых вычислительных узлов, происходит рост интегрального времени обработки. Это отражает тот факт, что система архивации, обслуживающая обработку, не справляется с обеспечением данными вычислительных узлов. При проведении потоковой обработки множества областей заданного размера это приводит к тому, что увеличение числа вычислительных узлов целесообразно до тех пор, пока время подготовки данных не станет сопоставимым с временем их обработки. Дальнейшее увеличение числа вычислительных узлов приведёт к неоправданному расходу ресурсов, так как при уменьшении размера кластера растёт доля «буферных» блоков данных, передаваемых по сети и используемых в обработке. Таким образом, мы можем сделать предположение, что должны существовать некоторые оптимальные значения a и p , при которых интегральное время обработки заданной области S будет минимально. Этот минимум должен находиться в области значений a и p , в которой скорость обработки данных будет равна скорости подготовки данных. То есть первый и второй члены в выражении (6) будут равны. Это соображение можно учитывать для оптимизации численного поиска оптимальных значений a и p .

Следует также учитывать, что на практике мы в основном сталкиваемся с задачей, в которой требуется найти оптимальные значения a и p для решения задачи с помощью имеющихся ресурсов. То есть необходимо найти оптимальное значение размера блока a при обработке заданной области S при фиксированном значении числа вычислительных узлов N , и в формулу (4) нужно подставить значения размера кластера p , вычисляемые по формуле:

$$p = \sqrt[2]{\frac{S}{Na^2}}. \quad (8)$$

В этом случае выражение (4) принимает вид:

$$T_{area}(S, a) = \frac{S}{a^2} T_{prep}(a) + \left(\left(\sqrt[2]{\frac{S}{Na^2}} + 2 \right)^2 T_{iget}(a) + \frac{S}{Na^2} (T_{proc}(a) + T_{upl}(a)) \right). \quad (9)$$

Поскольку N в данном выражении является константой, то видно, что поиск минимального значения T_{area} фактически превращается в поиск минимума по единственному параметру a .

Это, безусловно, упрощает процедуру решения задачи. Отметим, что после нахождения оптимального a оптимальное значение p вычисляется по формуле (8).

Ниже приведён пример проведения такой оценки для реальной процедуры обработки.

Пример расчёта оптимального размера блока при выделении для обработки фиксированного числа вычислительных узлов

Для демонстрации данной ситуации была использована задача выбора оптимальных параметров пространственного разбиения на блоки и кластеры данных для задачи радиометрической и геометрической коррекции данных прибора КМСС (спутники серии «Метеор-М»). В качестве допустимых значений блока были выбраны значения от $0,5$ до 3° с шагом $0,25^\circ$. На рис. 5 в виде графиков представлены выявленные по описываемой предлагаемой методике зависимости времени выполнения различных операций от размера блока, необходимые для выбора оптимальных параметров разбиения.

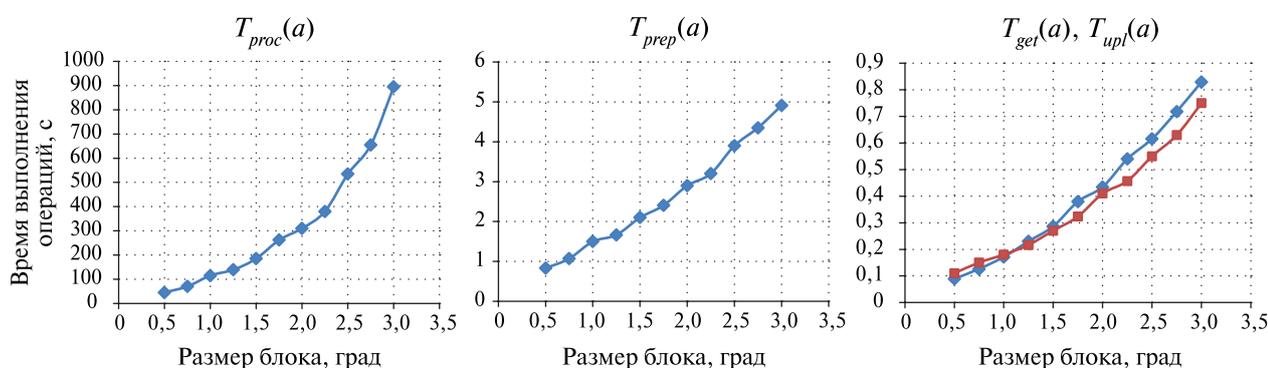


Рис. 5. Зависимости времени выполнения операций над блоком данных от размера блока. Красный цвет — $T_{upl}(a)$

Оценка оптимального размера блока проводилась для обработки участка площадью $40 \times 40^\circ$ с одновременным использованием 10 вычислительных узлов. Для каждого из допустимых значений размера блока был выбран размер кластера данных, позволяющий проводить обработку с требуемым уровнем распараллеливания, после чего по представленной выше формуле было оценено интегральное время выполнения обработки. На основании полученной зависимости (рис. 6) может быть сделан вывод, что для проведения обработки оптимальный размер блока a составляет 2° , которому соответствует размер кластера $p = 7$.

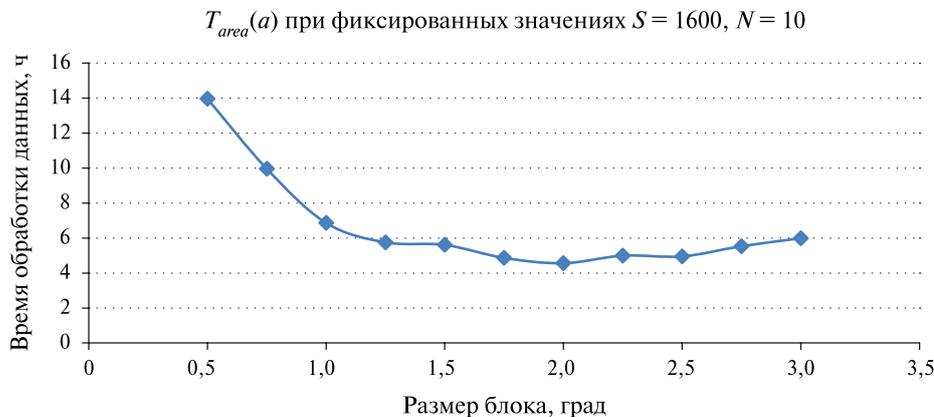


Рис. 6. Оценка времени обработки заданного объёма данных на 10 вычислительных узлах в зависимости от размера блока

Заключение

В статье представлен новый подход к организации обработки спутниковых данных с использованием динамического блочного доступа к разнородным архивам спутниковых данных. При этом оценены преимущества подобного подхода при организации работы с долговременными архивами спутниковых данных.

На базе ЦКП «ИКИ-Мониторинг» на основе предложенного подхода разработана и реализована схема организации работы системы обработки спутниковых данных. В настоящий момент эта схема уже опробована при создании различных процедур потоковой обработки данных и показала свою эффективность.

Также в работе представлен подход, позволяющий осуществлять выбор оптимальных параметров разбиения регионов, по которым проводится обработка, на блоки и их кластеры для достижения минимального суммарного времени, необходимого для выполнения конкретной обработки.

Таким образом, фактически в работе представлена реализованная на базе ЦКП «ИКИ-Мониторинг» новая технология организации обработки спутниковых данных с использованием разработанного динамического блочного подхода.

Представленная технология была успешно апробирована при реализации процедур проведения радиометрической и геометрической коррекции данных спутников серии «Метеор-М» с использованием данных спутников Terra, Aqua и Landsat-8. Это позволило существенно упростить реализацию самих процедур, а также в несколько раз сократить время этапа подготовки данных для обработки. В настоящее время ведутся работы по переводу на использование представленной технологии и других типов обработки данных, реализованных в ЦКП «ИКИ-Мониторинг».

Работа выполнена в рамках темы «Большие данные в космических исследованиях: астрофизика, солнечная система, геосфера» (государственная регистрация № 0024-2019-0014) с использованием возможностей ЦКП «ИКИ-Мониторинг» (<http://ckp.geosmis.ru/>).

Литература

1. *Балашов И. В., Ефремов В. Ю., Мазуров-мл. А. А., Мамаев А. С., Матвеев А. М., Прошин А. А.* Особенности организации контроля и управления распределенных систем дистанционного мониторинга // *Современные проблемы дистанционного зондирования Земли из космоса.* 2011. Т. 8. № 3. С. 161–166.
2. *Егоров В. А., Ильин В. О., Лупян Е. А., Мазуров А. А., Прошин А. А., Флитман Е. В.* Возможности построения автоматизированных систем обработки спутниковых данных на основе программного комплекса XV_SAT // *Современные проблемы дистанционного зондирования Земли из космоса.* 2004. Вып. 1. С. 431–436.
3. *Кобец Д. А., Матвеев А. М., Мазуров А. А., Прошин А. А.* Организация автоматизированной потоковой обработки спутниковой информации в системах дистанционного мониторинга // *Современные проблемы дистанционного зондирования Земли из космоса.* 2015. Т. 12. № 1. С. 145–155.
4. *Кобец Д. А., Матвеев А. М., Прошин А. А., Мазуров А. А.* Контроль работоспособности и управление распределенными комплексами автоматической потоковой обработки спутниковых данных // *Вопросы электромеханики.* Тр. ВНИИЭМ. Приложение за 2017 г.: Материалы 5-й международной научно-техн. конф. «Актуальные проблемы создания космических систем дистанционного зондирования Земли». 2018. С. 225–234. URL: http://www.vniiem.ru/ru/uploads/files/conferences/170525/materialy_2017.pdf.
5. *Лупян Е. А., Прошин А. А., Бурцев М. А., Балашов И. В., Барталев С. А., Ефремов В. Ю., Кашицкий А. В., Мазуров А. А., Матвеев А. М., Суднева О. А., Сычугов И. Г., Толпин В. А., Уваров И. А.* Центр коллективного пользования системами архивации, обработки и анализа спутниковых данных ИКИ РАН для решения задач изучения и мониторинга окружающей среды // *Современные проблемы дистанционного зондирования Земли из космоса.* 2015. Т. 12. № 5. С. 263–284.
6. *Лупян Е. А., Прошин А. А., Бурцев М. А., Кашицкий А. В., Балашов И. В., Барталев С. А., Константинова А. М., Кобец Д. А., Мазуров А. А., Марченков В. В., Матвеев А. М., Радченко М. В., Сычугов И. Г., Толпин В. А., Уваров И. А.* Опыт эксплуатации и развития центра коллективного пользования

- системами архивации, обработки и анализа спутниковых данных (ЦКП «ИКИ-Мониторинг») // Современные проблемы дистанционного зондирования Земли из космоса. 2019. Т. 16. № 3. С. 151–170. DOI: 10.21046/2070-7401-2019-16-3-151-170.
7. Прошин А. А., Лупян Е. А., Балашов И. В., Кашицкий А. В., Бурцев 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. DOI: 10.21046/2070-7401-2016-13-3-9-27.
 8. Gorelick N., Hancher M., Dixon M., Flyushchenko S., Thau D., Moore R. Google Earth Engine: Planetaryscale geospatial analysis for everyone // Remote Sensing of Environment. 2017. V. 202. P. 18–27.
 9. Peterson M. P. Web-Mapping Services // Intern. Encyclopedia of Geography: People, the Earth, Environment and Technology. 2016. P. 1–19.

Technology of satellite data dynamic block provision to distributed processing systems

A. A. Proshin, E. A. Loupian, I. V. Balashov, A. V. Kashnitskii,
A. M. Matveev, B. P. Rutkevich

Space Research Institute RAS, Moscow 117997, Russia
E-mail: andry@smis.iki.rssi.ru

In many tasks related to the processing and visualization of spatial data, a significant benefit is provided by the use of a block (tile) approach, within which data are used, presented in the form of blocks, into which the entire area of interest is evenly divided. This makes it possible to effectively organize parallel data processing on various computational nodes, use the caching mechanism and reduce the load on centralized resources. In this case, as a rule, already prepared data archives with fixed spatial partitions are used. At the same time, in many cases, first of all, when it is necessary to jointly process various types of satellite information, the use of prepared data archives may be impractical. This is due to the fact that various types of satellite data differ significantly both in spatial resolution and in the organization of their storage. As a result, it is almost impossible to choose one spatial partitioning that would be optimal for organizing work with all types of data. Therefore, it is required to develop such approaches to organizing satellite data processing that could maximize the advantages of the block approach and optimize the storage scheme for each specific type of data. The article presents such an approach to organizing the processing of satellite data using dynamic block access to heterogeneous archives of satellite data, within which the data necessary for processing are generated upon request in such a spatial division and with such characteristics that will be optimal for a specific type of processing. The paper provides a brief description of the developed technology, and also considers a methodology for evaluating the optimal partitioning parameters for a specific processing and existing infrastructure.

Keywords: parallel processing of satellite data, satellite data archives, big data, IKI-Monitoring center

Accepted: 10.12.2020

DOI: 10.21046/2070-7401-2020-17-7-79-93

References

1. Balashov I. V., Efremov V. Yu., Mazurov A. A., Mamaev A. S., Matveev A. M., Proshin A. A., Osobennosti organizatsii kontrolya i upravleniya raspredelennykh sistem distantsionnogo monitoringa (Features of remote monitoring distributed systems control and monitoring), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2011, Vol. 8, No. 3, pp. 161–166.
2. Egorov V. A., Il'in V. O., Loupian E. A., Mazurov A. A., Flitman E. V., Vozможности postroeniya avtomatizirovannykh sistem obrabotki sputnikovykh dannykh na osnove programmnoho kompleksa XV_SAT (Possibilities of developing automated satellite data processing systems on the basis of XV_SAT software package), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2004, No. 1, pp. 431–436.

3. Kobets D. A., Matveev A. M., Mazurov A. A., Proshin A. A., Organizatsiya avtomatizirovannoi mnogopotokovoi obrabotki sputnikovoi informatsii v sistemakh distantsionnogo monitoringa (Organization of automated multithreaded processing of satellite information in remote monitoring systems), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2015, Vol. 12, No. 1, pp. 145–155.
4. Kobets D. A., Matveev A. M., Proshin A. A., Mazurov A. A., Kontrol' rabotosposobnosti i upravlenie raspredelennymi kompleksami avtomaticheskoi potokovoi obrabotki sputnikovykh dannykh (Operation control and management of distributed complexes of automatic streaming processing of satellite data), *Voprosy elektromekhaniki. Trudy VNIIEМ. Prilozhenie za 2017 g. Materialy 5-i mezhdunarodnoi nauchno-tekhnicheskoi konferentsii "Aktual'nye problemy sozdaniya kosmicheskikh sistem distantsionnogo zondirovaniya Zemli"* (Materials of the 5th international scientific and technical conf. "Actual problems of creation of space remote sensing systems of the Earth"), Electromechanical matters, VNIIEМ studies, 2018, pp. 225–234, available at: http://www.vniiem.ru/ru/uploads/files/conferences/170525/materialy_2017.pdf.
5. Loupian E. A., Proshin A. A., Burtsev M. A., Balashov I. V., Bartalev S. A., Efremov V. Yu., Kashnitskii A. V., Mazurov A. A., Matveev A. M., Sudneva O. A., Sychugov I. G., Tolpin V. A., Uvarov I. A., Tsentr kollektivnogo pol'zovaniya sistemami arkhivatsii, obrabotki i analiza sputnikovykh dannykh IKI RAN dlya resheniya zadach izucheniya i monitoringa okruzhayushchei sredy (IKI center for collective use of satellite data archiving, processing and analysis systems aimed at solving the problems of environmental study and monitoring), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2015, Vol. 12, No. 5, pp. 263–284.
6. Loupian E. A., Proshin A. A., Burtsev M. A., Kashnitskii A. V., Balashov I. V., Bartalev S. A., Konstantinova A. M., Kobets D. A., Mazurov A. A., Marchenkov V. V., Matveev A. M., Radchenko M. V., Sychugov I. G., Tolpin V. A., Uvarov I. A., Opyt ekspluatatsii i razvitiya tsentra kollektivnogo pol'zovaniya sistemami arkhivatsii, obrabotki i analiza sputnikovykh dannykh (TsKP "IKI-Monitoring:") (Experience of development and operation of the IKI-Monitoring center for collective use of systems for archiving, processing and analyzing satellite data), *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*, 2019, Vol. 16, No. 3, pp. 151–170, DOI: 10.21046/2070-7401-2019-16-3-151-170.
7. 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, DOI: 10.21046/2070-7401-2016-13-3-9-27.
8. Gorelick N., Hancher M., Dixon M., Ilyushchenko S., Thau D., Moore R., Google Earth Engine: Planetaryscale geospatial analysis for everyone, *Remote Sensing of Environment*, 2017, Vol. 202, pp. 18–27.
9. Peterson M. P., Web-Mapping Services, In: *Intern. Encyclopedia of Geography: People, the Earth, Environment and Technology*, 2016, pp. 1–19.