

Программные потоки распределённой обработки изображений на основе микросервисов

Попов Сергей Борисович

д.т.н., доцент

Институт систем обработки изображений РАН –
филиал ФНИЦ «Кристаллография и фотоника» РАН, тел. +7(846)3322764
s.b.popov@gmail.com

Abstract. The service-oriented distributed data processing technology of remote sensing is proposed, making use of metaprogramming along with multimodal approach, where users could refer to some integrating stages as a substitute for specific processing operation, set required processing goals and shape their implementation methods. Image processing metaprogramming is based on the following project solutions: the system basic structural storage unit is a distributed image, available for the user in the form of storage interacting services set, that is, each distributed image segment takes the form of a separate service with its unique name in the image service hierarchy; frame services apply the principle in-memory data storage and processing; standard processing operations are services, containing meta data, necessary to generate a processing graph, such as implemented operation class, processing pattern type and its parameters. The possibility to determine some roundup class instead of specific processing operation at the solution graph generation provides system multimodality. Graphic Charts, based on image processing operation master set, allow implementing efficient methods of image in-memory distributed processing by means of dataflow networks within the framework of computing process inside of the frame-service.

Аннотация. Предлагается сервис-ориентированная распределённая технология обработки данных дистанционного зондирования с использованием метапрограммирования и мультимодального подхода, при котором пользователь может указывать не конкретные операции обработки, а некоторые обобщающие этапы, формулировать требуемые цели обработки и намечать пути их реализации. Реализация метапрограммирования обработки изображений основывается на следующих проектных решениях: основной структурной единицей хранения в системе является распределённое изображение, которое доступно пользователю как набор взаимодействующих сервисов хранения, т.е. каждый фрагмент распределённого изображения представляется в виде отдельного сервиса (фрейм-сервиса) со своим уникальным именем в иерархии изображений-сервисов; фрейм-сервисы используют принцип обработки и хранения данных в оперативной памяти; типовые операции обработки являются сервисами, содержат необходимую для формирования графа обработки метаинформацию: класс реализуемой операции, тип паттерна обработки и его параметры. Возможность указать не конкретную операцию обработки, а некоторый обобщающий класс при формировании графа решения обеспечивает свой-

ство мультимодальности системы. Применение графов решений, заданных на множестве типовых операций обработки изображений, позволяет реализовать эффективные методы организации распределенной in-memory обработки изображений за счёт организации потоковых сетей в рамках вычислительного процесса "внутри" фрейм-сервиса.

В настоящее время использование облачной инфраструктуры для размещения данных и их обработки становится важным технологическим трендом. При размещении, хранении и обработке в распределённой информационной среде изображений большого размера и/или наборов связанных фрагментов изображений следует учитывать, помимо большого объёма данных, их структурные особенности, а именно, пространственную зависимость данных, необходимость согласования многомерной структуры данных и последовательного представления их при хранении.

Именно поэтому достаточно популярный путь перехода в «облако» на основе использования технологий «больших данных» (Big Data), в частности, программной платформы Hadoop, для организации обработки и хранения данных дистанционного зондирования наталкивается на значительное число проблем, связанных с тем, что технологии Big Data в первую очередь ориентированы на обработку текстовой информации. Отдельные проекты использования платформы Hadoop для хранения и обработки набора изображений только подтверждают это. Однако, общие принципы и методологию распределённого хранения и параллельной обработки данных полезно использовать при разработке систем обработки и хранения изображений.

Технологии обработки Big Data эксплуатируют две основных идеи: 1) тотальное распараллеливание данных и вычислений с обеспечением отказоустойчивости как на уровне хранения, так и при обработке, и 2) наличие базовой инструментальной среды, позволяющей исследователям сосредоточиться на написании содержательной части обработки, обеспечивая распределённое отказоустойчивое хранение данных, снимая проблемы организации распределённой обработки «на месте хранения», доставки исполняемого кода на узлы хранения для проведения вычислений.

Начальные реализации платформы Hadoop основаны на парадигме MapReduce. Основная проблема заключается в сложности представления произвольного алгоритма обработки (получения результата) в виде двухэтапной процедуры, где на первом этапе (этапе Map) происходит формирование частичных промежуточных результатов обработки отдельных частей единого массива данных, а на втором этапе (этапе Reduce) на их основе собирается окончательный результат. Дополнительной проблемой такой организации вычислений является необходимость сохранять все промежуточные результаты в локальной дисковой памяти узлов распределённого вычислительного кластера, поэтому для обеспечения приемлемой эффективности процесса такой обработки объём данных после первого этапа должен уменьшиться на несколько порядков. Именно поэтому, организация распределённой обработки изображений на базе платформы Hadoop зачастую реализуется в вырожденном виде (только этап Map), по сути, используя только встроенные в данную инфраструктуру средства организации

распределённой обработки «на месте хранения». Поскольку не все методы обработки можно реализовать в рамках парадигмы MapReduce, то появились потоковые методы обработки больших данных, которые также основаны на идее разбиения исходного набора данных на независимые блоки с последующей потоковой схемой их преобразования в результат с использованием распределённой отказоустойчивой организации вычислений. Возросшая при этом эффективность обработки связана, в первую очередь, с реализованной возможностью не сохранять промежуточные результаты вычислений на диск.

При творческом развитии и адаптации изложенных идей применительно к задачам разработки распределённых систем обработки и хранения изображений необходимо учитывать основную особенность изображений как данных, а именно, пространственную зависимость данных. Простое разбиение крупноформатного изображения на отдельные фрагменты, которое реализуется в системах на базе платформы Hadoop, в большинстве случаев не применимо, оно должно быть программно-контролируемым, особенно при использовании в процессе обработки локальных операций на основе скользящей окрестности.

Особенностью вычислительных задач обработки изображений является тот факт, что значительное число методов сложной обработки могут быть реализованы как последовательное применение некоторых законченных типовых операций над изображениями. Именно это обстоятельство определяет эффективность применения разнообразных программных систем обработки изображений к решению широкого спектра исследовательских и прикладных задач обработки и анализа изображений.

Для технологий обработки изображений (в первую очередь, крупноформатных) характерно наличие предварительной обработки в виде многоэтапного процесса, при котором промежуточные данные существенно не редуцируются в объёме. Распределённая организация вычислений при этом связана с использованием декомпозиции на перекрывающиеся фрагменты, то есть совместно с основными отсчётами, которые будут изменяться в процессе обработки, хранятся пограничные отсчёты из соседних фрагментов, которые используются только в процессе обработки и при многоэтапной обработке должны корректироваться по результатам завершения очередной операции. Решением в данном случае является использование концепции Data-as-a-Service.

Основные проектные решения при этом развивают и адаптируют принципы технологии Big Data применительно к задачам разработки систем обработки и хранения изображений, ориентированных на использование в облачных вычислительных средах: (1) все компоненты системы хранения и обработки изображений реализуются как сервисы, размещенные в облачной вычислительной среде; (2) в качестве интеграционного решения используется децентрализованное объединение равноправных взаимодействующих друг с другом сервисов, т.е. федерация сервисов; (3) основной структурной единицей хранения в системе является распределённое изображение, которое доступно пользователю как набор взаимодействующих сервисов хранения, т.е. каждый фрагмент распределённого изображений представляется в виде отдельного сервиса (фрейм-сервиса) со своим уникальным именем в иерархии изображений-сервисов; (4) фрейм-сервисы используют принцип обработки и хранения данных в оперативной памяти; (5) сервисы регистрируются в распре-

деленном репозитории, который обеспечивает группировку всех сервисов по передаваемой от них информации: сервисов-обработчиков по типам операций, фрейм-сервисов по принадлежности к исходному распределённому изображению, пространственным координатам и другим статистическим характеристикам; (6) организация вычислений в процессе обработки изображений использует ориентированный на данные подход, при котором фрагменты заранее распределены по виртуальным машинам (ВМ) облачной инфраструктуры, а в процессе обработки на ВМ распределенной системы вместо данных рассылаются процедуры их обработки; (7) в процессе обработки данные не изменяются, а создаётся новое распределённое изображение, фрагменты которого могут быть не согласованы между собой в части обеспечения последующей сбалансированности распределенной обработки, но наличие "интеллекта" у фрейм-сервиса и связей с соседями позволяет выполнить все необходимые для такого согласования действия автономно в фоновом режиме, предоставляя пользователю необходимые ему функции (чаще всего это визуализация) и данные непосредственно сразу после создания.

Декомпозиция данных при формировании федерации фрейм-сервисов одного изображения реализована как дизъюнктивное объединение подмассивов отсчетов, сформированных как результат применения к исходному изображению двумерной пространственной декомпозиции с полным перекрытием фрагментов. Данный вариант обеспечивает восстановление данных при отказе до трех сервисов хранения/обработки, позволяет использовать оригинальный алгоритм балансировки загрузки процессоров в процессе обработки всего массива данных крупноформатных распределённых изображений.

Предлагаемые решения опираются на следующие базовые принципы инфраструктуры распределенной обработки больших данных Hadoop:

- Обрабатываем там, где храним («Move Code to Data!») – данные в процессе обработки не перемещаются, наоборот, процедуры обработки доставляются к данным и запускаются на вычислительных ресурсах распределённых систем хранения.
- Данные навсегда (No UPDATE or «Data Are Forever!») – данные в процессе обработки не изменяются и не удаляются, результаты просто сохраняются здесь же в виде нового набора.
- In-Memory Processing – обработка и промежуточное хранение данных в оперативной памяти (добавлено в рамках современной технологии распределенной потоковой обработки Apache Spark).

В соответствии с этими принципами фрейм-сервис, помимо интерфейса доступа к данным, имеет интерфейс, обеспечивающий получение и выполнение задания на обработку контролируемых данных.

Типовые операции обработки также оформлены в виде сервисов. Сервисы-обработчики могут функционировать в двух режимах – как фильтр и как ресурс. В первом случае сервис получает и перерабатывает потоковые данные, во втором – по запросу передает специальным образом оформленную задачу, которая выполняется в виртуальной машине фрейм-сервиса. Все сервисы регистрируются в репозитории, который реализован на базе сервиса координации распределенных приложений Zookeeper, обеспечивающего синхронизацию и единое

иерархическое пространство имен для изображений-сервисов и их составных частей в виде фрейм-сервисов, группировку всех сервисов по передаваемой от них информации: сервисов-обработчиков по типам операций, фрейм-сервисов по принадлежности к исходному распределённому изображению, пространственным координатам и другим статистическим характеристикам.

Базовый вариант формирования задачи обработки изображений заключается в задании последовательности выполняемых операций обработки, т.е. графа обработки. Оформление типовых операций обработки в виде сервисов позволяет хранить совместно с собственно программами необходимую для формирования графа обработки метаинформацию: класс реализуемой операции, тип паттерна обработки и его параметры.

При формировании графа обработки используются операции следующих типов: операции поэлементной обработки (PO – point operators); операции обработки в локальной окрестности (LNO – local neighborhood operators) или локальной обработки скользящим окном; глобальные операции оценки характеристик или операции вычисления параметров изображения (RO), в т.ч. статистических; операции геометрических преобразований (реализованы практически важные операции масштабирования и поворота на небольшой угол, которые по своей структуре близки к локальным). Таким образом, в процессе формирования задачи обработки изображения возможно указать практически все типы операций низкоуровневой обработки изображений за исключением глобальных операций обработки, использование которых всё равно предполагает выход за рамки парадигмы параллельной обработки распределённо хранящихся фрагментов изображения.

Эффективная организация параллельных вычислений для задач сложной обработки изображений, реализованных как последовательное применение некоторых законченных типовых операций над изображениями, в рамках выбранной сервис-ориентированной архитектуры системы требует минимизации создания промежуточных изображений в процессе обработки. Это обеспечивается за счёт передачи фрейм-сервисам не отдельной операции обработки контролируемых им данных, а последовательности выполнения операций, заданной в виде графа решения, необходимой для формирования нового распределённого изображения.

Таким образом, при решении задачи обработки изображений необходимо сформировать полный граф обработки, решающий задачу, разбить его на подграфы (возможно тривиальные, то есть состоящие из одной вершины-операции), допускающие выполнение параллельной обработки распределённого изображения в рамках отдельных фрейм-сервисов без формирования промежуточных распределённых изображений.

В практически важном случае для графов, реализующих последовательность операций PO и/или LNO, число таких операций, которые можно выполнять в рамках одного задания для фрейм-сервиса (т.е. без создания промежуточного распределённого изображения), определяется суммой радиусов окна обработки участвующих в наборе операций LNO, которая не должна превышать размер перекрытия фрагментов распределённого изображения.

Для подграфов, которые определяют одно задание для фрейм-сервиса, определены два варианта организации вычислений в рамках фрейм-сервиса: после-

довательный и потоковый. Для ациклических подграфов используется последовательный вариант организации вычислений, для подграфов, которые содержат один или несколько циклов (если рассматривать графы как неориентированные), необходимо реализовать потоковый вариант организации вычислений. Потоковая обработка изображений основана на последовательном характере вычислений "внутри" операции обработки. Тот факт, что определенная часть информации готова значительно ранее полного завершения текущего шага обработки, позволяет использовать схему с параллельным выполнением всех шагов алгоритма и непосредственной передачей по мере готовности строк изображения от одной операции к другой. Фактически формируется некоторый поток данных (потоковая сеть), элементы которого преобразуются на каждом шаге и передаются последующему для дальнейшей обработки. Для подграфов, которые содержат один цикл, т.е. имеют две параллельные ветви обработки, условием отсутствия тупика при потоковой обработке является равенство сумм радиусов окна обработки участвующих в отдельных ветках операций LNO. Если условие не выполняется, то корректность потоковой сети восстанавливается добавлением буферизации строк изображения в той ветке, которая имеет меньшую сумму радиусов окна обработки участвующих в ней операций LNO.

Выполнение вышеперечисленных условий означает возможность распределенной in-методу-обработки крупноформатных изображений в облачных вычислительных средах.

Предлагаемые решения позволяют организовать технологию обработки данных дистанционного зондирования Земли с использованием метапрограммирования и мультимодального подхода, при котором пользователь может указывать не конкретные операции обработки, а некоторые обобщающие этапы, формулировать требуемые цели обработки и намечать пути их реализации. Реализация метапрограммирования обработки изображений основывается на следующих проектных решениях: типовые операции обработки являются сервисами, содержат необходимую для формирования графа обработки метаинформацию: класс реализуемой операции, тип паттерна обработки и его параметры. Возможность указать не конкретную операцию обработки, а некоторый обобщающий класс при формировании графа решения обеспечивает свойство мультимодальности системы.

Использование потоковой модели вычислений в виде ациклического направленного графа применяется в современной технологии Apache Spark. Однако, в качестве обрабатываемой распределенной структуры в данном случае используется RDD. Предлагаемое решение использует специализированные фрейм-сервисы, точнее соответствующие пространственной природе изображений. Применение графов решений, заданных на множестве типовых операций обработки изображений, позволяет реализовать эффективные методы организации распределенной in-методу-обработки крупноформатных изображений за счёт использования потоковых сетей в рамках вычислительного процесса "внутри" фрейм-сервиса.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект № 1529-07077).