

Обработка больших объемов сырых астрономических данных с помощью модели вычислений MapReduce¹

¹С.В. Герасимов < gerasimov@mlab.cs.msu.su >

²А.В. Мещеряков < mesch@iki.rssi.ru >

¹И.Ю. Колосов < zackwag32@gmail.com >

¹Е.С. Глотов < glot.unltd@gmail.com >

¹И.С. Попов < ivan@mlab.cs.msu.su >

¹ Московский государственный университет имени М.В. Ломоносова,
119991, Россия, Москва, Ленинские горы, д. 1., стр. 52, факультет ВМК

² Институт космических исследований РАН,
117342, Россия, г. Москва, Профсоюзная ул., 84/32

Аннотация. Экспоненциальный рост объемов, повышение качества данных в современных и будущих обзорах неба открывают перед астрофизиками новые горизонты, однако требуют применения новых подходов к их обработке, а именно технологий больших данных и облачных вычислений. В работе предлагается подход, основанный на модели MapReduce, для решения одной из самых масштабных и важных вычислительных задач астрофизики — задачи обработки сырых данных астрономических изображений.

Ключевые слова: MapReduce; Hadoop; небесный обзор; большие данные; облачные вычисления; обработка изображений

DOI: 10.15514/ISPRAS-2015-27(6)-20

Для цитирования: Герасимов С.В., Мещеряков А.В., Колосов И.Ю., Глотов Е.С., Попов И.С. Обработка больших объемов сырых астрономических данных с помощью модели вычислений MapReduce. Труды ИСП РАН, том 27, вып. 6, 2015 г., стр. 315-334. DOI: 10.15514/ISPRAS-2015-27(6)-20.

1. Введение

Развитие наблюдательной астрономии на современном этапе характеризуется взрывным ростом объема данных, получаемых телескопами в рамках программ *небесных обзоров* — больших наборов цифровых фотографий неба в заданном спектральном диапазоне (фильтре), покрывающих большую область

¹ Работа поддержана грантом РФФИ №15-29-07085 офи_м

неба. Крупнейший проект - Слоановский цифровой обзор неба (англ., SDSS) [1] — проводился в 1998-2009 годах и был завершен в январе 2011 года выпуском официального релиза и размещением всех данных обзора в публичном доступе. Бурное развитие цифровых приемников излучения (ПЗС-матриц) за последние 30 лет [2] к настоящему времени сделало возможным за одно наблюдение на современном оптическом телескопе с большим полем зрения получать широкоформатную фотографию в заданном фильтре участка ночного неба площадью несколько квадратных градусов (1 снимок неба $\sim 10^9$ пикселей имеет объем ~ 1 ГБ для современных обзорных телескопов Subaru-HSC [3], DES [4], PanSTARRS [5]). За одну ночь оптический телескоп может делать до тысячи фотографий неба, обеспечивая поток научных данных около 1ТБ/сутки. Скорость получения астрономических данных будет продолжать экспоненциально расти вместе с вводом в строй обзорных телескопов следующего поколения, таких как Большой синоптический обзорный телескоп (англ. LSST) [6], [7], который должен быть построен к 2019 году. При потоке научных данных ~ 15 ТБ/сутки ожидаемый объем всех астрономических изображений, полученных одним телескопом LSST за 10 лет работы, составляет 114 петабайт.

В настоящее время объем астрономических изображений, хранящихся в открытых архивах в центрах обработки данных обсерваторий по всему миру, составляет несколько петабайт, это сотни миллионов астрономических изображений.

Рост объемов данных наблюдений, повышение качества астрономических данных открывает перед астрофизиками новые горизонты, однако требует применения новых современных инженерных и математических подходов к их обработке, среди которых *технологии больших данных, облачные вычисления*.

Для извлечения данных о небесных объектах, содержащихся в больших массивах сырых изображений, полученных цифровыми камерами телескопов, используется цепочка преобразований, называемая *конвейер для обработки астрономических изображений* (далее просто *конвейер*, англ., *pipeline*). Финальной целью работы конвейера является получение *каталога небесных объектов* — таблицы чисел, содержащей значения свойств объектов, обнаруженных на изображениях.

Типичный конвейер по обработке сырых изображений, полученных телескопами, состоит из следующих этапов.

- *Первичная обработка изображений* с телескопа включает в себя вычитание шума считывания, коррекцию неравномерности чувствительности ПЗС-матрицы, удаление горячих пикселей и другие этапы предварительной обработки изображений, которые делаются в обсерватории перед тем, как данные попадут в архив.
- *Астрометрическая калибровка изображений*. На каждом изображении детектируются все яркие объекты, и карта этих объектов

сравнивается с картой той области неба, куда был направлен телескоп в момент наблюдения. В результате их сравнения строится система мировых координат на изображении (WCS, англ. World Coordinate System).

- *Фотометрическая калибровка изображения.* Величина яркости всех источников на изображении калибруется относительно яркости стандартных звезд.

Перечисленные выше этапы обработки, как правило, уже сделаны для изображений, входящих в состав официальных релизов астрономических данных небесных обзоров (в частности, обзора SDSS, с данными которого мы работаем в данной статье). Далее любой астрономический конвейер должен реализовывать два основных этапа, это:

- *Объединение изображений.* Каждый небесный обзор состоит из множества пересекающихся изображений — проекций небольших участков неба на плоскость ПЗС-матрицы телескопа в моменты наблюдений. Из-за движения Земли и телескопа, меняется ориентация ПЗС-матрицы в пространстве, и каждый новый кадр, полученный телескопом, имеет свои параметры проекции (которые определяются на этапе астрометрической калибровки изображения). В процессе объединения изображений все доступные кадры в заданном фильтре и для данного участка неба (по-разному ориентированные друг относительно друга в системе мировых координат) объединяются в одно большое изображение в заданной проекции. Основные подэтапы объединения изображений: (i) проецирование всех кадров в общую систему координат, (ii) удаление из изображений фоновой компоненты и (iii) сложение кадров для получения объединенной картинки заданного участка неба. Объединенное изображение обладает важными преимуществами по сравнению с каждым отдельным кадром: (а) на нем видно больше слабых объектов (сложение нескольких кадров увеличивает глубину астрономического изображения), (б) свойства объектов могут быть измерены более детально за счет увеличения динамического диапазона объединенного изображения, (в) свойства протяженных объектов (сравнимых с размером кадра) могут быть точно измерены на большом изображении, составленном из нескольких кадров и, наконец, (г) объединение изображений заметно сокращает размер астрономических данных без потери информации, что уменьшает затраты на их дальнейшую обработку.
- *Создание астрономического каталога.* Каталог представляет собой таблицу всех небесных объектов, обнаруженных на астрономических изображениях с фиксированным набором свойств, “измеренных” для каждого объекта. Основные подэтапы создания каталога для каждого изображения: (i) вычитание фоновой компоненты из изображения, (ii)

создание маски “плохих” областей на изображении (например, области вокруг ярких звезд), где данные по объектам сильно искажены, и исключение этих областей из дальнейшего анализа, (iii) поиск (детектирование) групп объектов на изображении, (iv) удаление отдельных артефактов (“искусственных” объектов, таких как следы от космических лучей, самолетов и спутников, блики от звезд и т.д.) из общего списка групп объектов, (v) пространственное разделение объектов внутри каждой из групп (деблендирование), (vi) экстракция разнообразных свойств (характеристик яркости, размера, формы, морфологии) для каждого объекта на изображении, (vii) классификация объектов на точечные/протяженные (звезды/галактики) на основе их свойств, (viii) измерение профилей ярких звезд на изображении и построение эмпирической модели функции отклика на точечный источник (англ. PSF), (ix) “измерение” свойств объектов, скорректированных моделью PSF², (x) сохранение всех свойств объектов в виде таблицы-каталога и объединение каталогов по всем изображениям.

Последовательность, набор этапов, качество реализации каждого из этапов, перечень “измеряемых” свойств могут варьироваться в зависимости от реализации конвейера. Как правило, конвейеры реализуются в рамках проектов небесных обзоров, и астрофизики имеют возможность либо скачать и самостоятельно обработать сырые данные (изображения), либо воспользоваться готовым каталогом соответствующего обзора и никоим образом не могут повлиять на алгоритмы конвейера, например, оптимизировать один из его этапов или “измерить” новую характеристику небесных объектов на всех изображениях и добавить ее в каталог.

Если построение своего конвейера для обработки изображений небольшого объема (несколько ГБ) на персональном компьютере еще представляется возможным, например, с помощью распространенных астрономических пакетов SWarp, SExtractor и PSFEx [8], то настраиваемая обработка больших подвыборок (от нескольких ТБ) сырых изображений современных небесных обзоров и архивов астрономических обсерваторий практически невозможна. В связи со взрывным ростом объемов данных в небесных обзорах следующего поколения (таких как PanSTARRS, LSST) ситуация будет резко усугубляться.

Необходимо отметить, что от алгоритмической базы конвейера напрямую зависит *качество данных в каталогах*, с которыми большинство астрофизиков

² Земная атмосфера и оптика телескопа искажают изображения небесных объектов — точечный источник на картинке с телескопа имеет конечный размер и сложную форму, которая меняется даже внутри одного изображения. Цель построения пространственной модели PSF на каждом изображении — свести к минимуму влияние атмосферы и телескопа на “измеряемые” свойства объектов.

будет работать в своих моделях и на которых будут проверяться научные гипотезы. Группы астрофизиков, работающие в таких предметных областях, как наблюдательная космология, физика галактик, звездная астрономия, исследование астероидов, могут быть заинтересованы в максимальной оптимизации тех (зачастую различных) элементов конвейера, которые наиболее критичны для решения их задач.

Цель настоящей работы — исследование и разработка горизонтально масштабируемого конвейера для обработки астрономических изображений, доступного для применения коллективами астрофизиков для пакетной обработки больших объемов сырых данных современных и будущих небесных обзоров, а также данных изображений архивов обсерваторий. Архитектура конвейера должна предоставлять широкие возможности по настройке и модификации алгоритмов, используемых на всех его этапах, включая возможность добавления возможности “измерения” новых свойств небесных объектов.

Одним из ключевых требований является легкость в развертывании на недорогих вычислительных мощностях астрономических групп разного масштаба или в облаке, отсутствие необходимости в привлечении высококвалифицированных программистов и системных администраторов. Как следствие, разработанный конвейер обладает рядом уникальных функциональных и нефункциональных возможностей:

- настраиваемая алгоритмическая база этапов конвейера;
- горизонтальная масштабируемость;
- легкость развертывания на ”дешевом” кластере и в облаке;
- ряд оптимизированных шагов обработки, в т.ч. корректная обработка протяженных объектов (часто отсутствует в современных конвейерах).

2. Обзор существующих решений по параллельным конвейерам

В работе [9] представлен способ параллельного выполнения одного из шагов конвейера — объединения изображений. Объединение производится для небольшого (градус или его доли) целевого фрагмента неба и заданного цветового фильтра с использованием инфраструктуры Apache Hadoop [10]. Алгоритмы объединения реализованы авторами непосредственно в процедурах map и reduce. В качестве источника изображений используется обзор SDSS. Основное внимание в работе уделено повышению общей производительности обработки за счет эффективной фильтрации изображений, входящих в небольшой по размеру искомый участок неба с помощью предварительной фильтрации имен FITS-файлов (Flexible Image Transport System - формат, совмещающий растровое астрономическое изображение и мета-информацию о нем) регулярными выражениями либо

использовании информации из СУБД SDSS, позволяющей сформировать набор изображений, соответствующих целевому фрагменту неба.

Проект Montage [11] представляет собой известную open-source реализацию конвейера для обработки больших объемов сырых изображений. Архитектура Montage основывается на MPI [12] и Pegasus [13] — фреймворке для отображения графа научных задач на вычислительные ресурсы HTC (англ. High-Throughput Computing). Montage реализует следующие этапы параллельной обработки изображений:

- проецирование входных изображений;
- удаление фоновой компоненты;
- создание мозаики (объединенного изображения).

Причем для самого вычислительно сложного этапа обработки — проецирования изображений — дано несколько оптимизированных версий.

В работе [14] реализован конвейер на основе MapReduce [15] по получению каталога из изображений обзора SDSS. В качестве базового астрофизического ПО использовался SExtractor. Вычисления проводились на локальном кластере из 54 узлов (432 ядер).

В цепочку обработки для каждого изображения вошли следующие этапы:

- вычитание фона;
- поиск источников;
- деблендирование;
- измерение свойств объектов без коррекции на PSF;
- таблицы объектов объединялись в общий каталог.

В работе [16] реализован конвейер по получению каталога из изображений в архиве Национальной оптической астрономической обсерватории (англ., NOAO) из обзора DES (30ТБ данных изображений). Вычисления проводились на кластере Darwin Кембриджского университета и на локальном кластере, состоящем из 8 узлов / 96 ядер. В качестве базового астрофизического ПО использовались: SExtractor и PSFEx. В цепочку обработки для каждого изображения вошли следующие этапы:

- вычитание фона;
- поиск источников;
- деблендирование;
- отбор звезд и построение модели PSF;
- измерение свойств объектов с коррекцией на PSF;
- этап фотометрической калибровки в настоящее время отсутствует для данных обзора DES, поэтому авторы провели его самостоятельно.

Обработка 30ТБ данных изображений заняла у авторов около 15000 процессорных часов.

3. Предложенное решение

3.1 Базовые технологии распределенного хранения и пакетной обработки

В качестве основы архитектуры разработанного параллельного конвейера используются *распределенная файловая система HDFS* и *модель вычислений MapReduce*, реализованная в *Apache Hadoop*. Этот выбор обусловлен несколькими факторами.

1. Концепция планирования вычислений на основе *локальности данных* хорошо “ложится” на разновидности задач классического конвейера по обработке сырых данных небесных обзоров:

- обработка отдельного изображения, результатом которого является новое изображение (проецирование, удаление фоновой компоненты);
- обработка нескольких изображений для получения объединенного изображения;
- извлечение набора объектов и их свойств из отдельного изображения.

2. Apache Hadoop обеспечивает горизонтальное масштабирование, восстановление процесса вычислений после сбоев, надежность хранения данных, что необходимо для хранения и пакетной обработки больших массивов данных.

3. Apache Hadoop прост в администрировании и использовании, поддерживается популярными *облачными сервисами*, например, Microsoft Azure HDInsight, Amazon EMR. Существуют сторонние дистрибутивы Apache Hadoop, например, Cloudera и Hortonworks, предоставляющие дружественные пользовательские интерфейсы для разворачивания и мониторинга кластера Hadoop.

4. стек Hadoop включает в себя набор технологий, например, индексированные хранилища данных Apache HBase[17] и Hive[18], которые могут быть использованы на этапах обработки данных, следующих за обработкой изображений: систематизация данных небесных каталогов больших объемов, объединение данных из нескольких небесных обзоров, обработка запросов пользователей к данным каталогов.

5. Выбранные базовые технологии позволяют осуществить интеграцию с *интеграцией MapReduce* платформами (например, Apache Spark[19]) для эффективной итеративной обработки данных больших объемов алгоритмами машинного обучения.

3.2 Реализация базовых алгоритмов работы конвейера

Для реализации базовых алгоритмов работы конвейера используются де-факто признанные в мире астрофизиков пакеты ПО:

- SWarp (проецирование, измерение и удаление фоновой компоненты из изображений, объединение изображений);
- SExtractor (измерение и удаление фоновой компоненты, детектирование объектов на изображениях, классификация на точечные/протяженные и последующее “измерение” свойств объектов, в том числе исправленных на PSF, — при наличии готовых моделей функции отклика, построенных в PSFEx);
- PSFEx (определение модели инструментальной функции отклика на основе свойств точечных объектов, “измеренных” SExtractor).

Перечисленные пакеты программ могут быть эффективно использованы на отдельных персональных компьютерах с многоядерными процессорами для обработки небольших объемов изображений. К примеру, SWarp может совместить на современном персональном компьютере с использованием двух рабочих нитей 17 изображений общим размером 180 МБ за 4 минуты 41 секунду. Указанные пакеты ПО свободно лицензируются (GPL), написаны на языке Си и обладают продуманной модульной структурой, упрощающей их доработку. Все перечисленные утилиты имеют файловый интерфейс входных/выходных данных, настройки обработки задаются в виде конфигурационных файлов. Следует отметить, что имеется возможность *ограничивать данные пакеты ПО по использованию ресурсов*: количеству нитей, объему ОЗУ и жесткого диска.

3.3 Реализация параллельного конвейера

Целевой участок неба, для которого выполняется обработка, разделяется на прямоугольные *клетки* с фиксированными сторонами. Клетки нумеруются двумя индексами, соответствующими строке и столбцу (рис. 1).



Рис.1. Покрывтие исходных изображений “клетками”.

На 1-м этапе работы предложенного алгоритма параллельной конвейеризации сырых изображений (рис. 2), хранящихся в виде файлов в формате FITS, производятся операции фильтрации изображений по попаданию в целевой участок неба, а также удаление фона, проецирование, объединение изображений с помощью пакета SWarp. Все перечисленные операции кроме объединения проводятся независимо над каждым исходным изображением и могут быть распараллелены в рамках map задач с задействованием локальности данных. Функция map:

- пропускает входное изображение, если оно не попадает в целевую область;
- в случае попадания модифицирует изображение и возвращает номер клетки, к которой принадлежит изображение и само модифицированное изображение в качестве пары ключ-значение.

Следует отметить, что изображение может пересекаться сразу с несколькими клетками, в этом случае map вернет соответствующее число пар: номер клетки, модифицированное изображение.

В reduce попадают изображения, принадлежащие одной клетке. Над ними производится операция объединения.

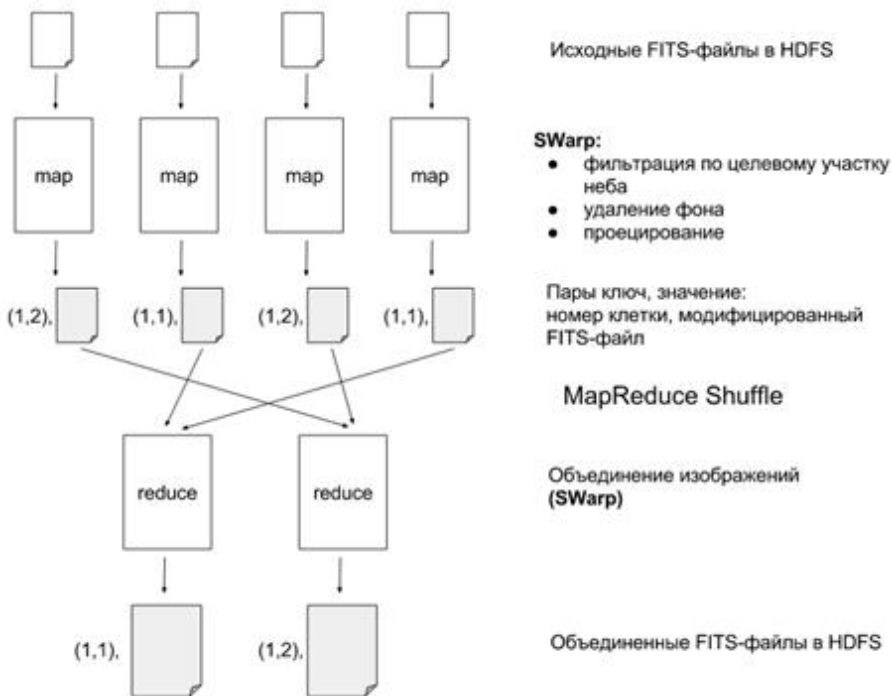


Рис.2. 1-й этап работы параллельного конвейера.

На 2-м этапе параллельной обработки производится извлечение небесных объектов и их свойств из изображений-клеток. Для этого в каждой map функции:

- осуществляется обнаружение объектов на изображениях, классификация точечных источников и извлечение свойств с помощью SExtractor;
- на основе свойств точечных объектов с помощью пакета PSFEx строится пространственная модель PSF;
- SExtractor осуществляет “измерение” свойств каждого объекта с учетом PSF-модели.

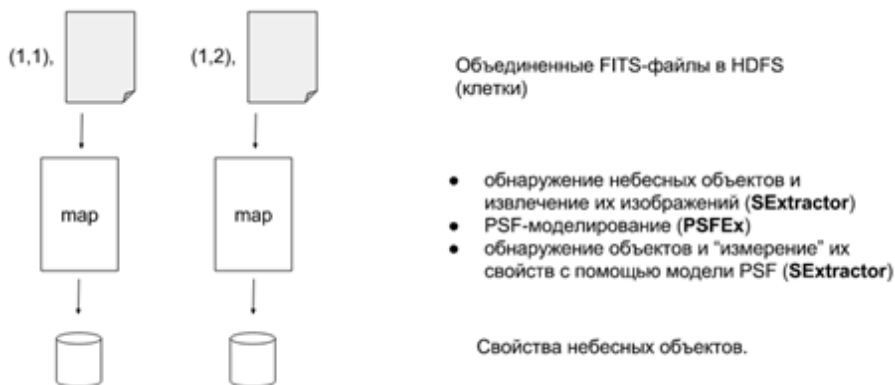


Рис.3. 2-й этап работы параллельного конвейера.

Следует отметить, что для корректной обработки на 2-м этапе больших объектов на краю клетки, уже на 1-м этапе клетки формируются внахлест за счет дополнительной *рамки*, ширина которой больше, чем характерный размер объектов, свойства которых мы хотим надежно "измерять". При этом обработка и измерение свойств объектов будут производиться корректно (объект на краю клетки не будет обработан дважды и "по частям"): каждый объект будет обработан в пределах единственной клетки. Объект, находящийся на границе, будет обрабатываться в той клетке, на стороне которой он находится относительно границы клеток без учета рамок (рис.4).

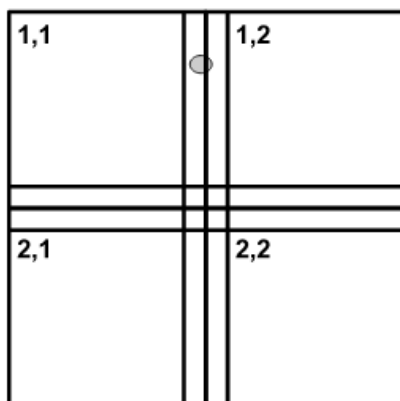


Рис.4. Обработка объекта на границе клеток. Данный объект попадет в обрамленную клетку 1,1.

Для встраивания пакетов SWarp, SExtractor и PSFEx использовался интерфейс локальных файлов, создаваемых во временных директориях узлов. В будущем планируется адаптировать интерфейсы пакетов для использования стандартных потоков ввода, вывода. Следует отметить, что благодаря встроенной возможности ограничения ресурсов, используемых перечисленными пакетами, в особенности используемого объема ОЗУ, стала возможной мотивированная настройка параметров контейнеров YARN, необходимых map и reduce задачам (см. следующий раздел).

4. Эксперименты

В качестве данных для экспериментов использовалось подмножество сырых изображений красного фильтра набора *Stripe82* небесного обзора SDSS DR12. SDSS камера (рис.5) представлена 6 вертикальными рядами ПЗС-матриц, фиксирующих излучение в одном из 5 фильтров видимого спектра волн (условно названных *r*, *i*, *u*, *z*, *g*). Вертикальные ряды ПЗС-матриц имеют “слепые” зоны между собой. Снятие изображений непрерывных полос неба обеспечивается за счет перемещения небесных объектов “по вертикали” (вдоль рядов ПЗС-матриц) благодаря движению Земли. Каждый сеанс снятия изображений называется *проходом* (англ., *run*). Для устранения “слепых” вертикальных зон обычно выполняется два прохода, второй со смещением камеры по “горизонтали” (эти проходы называются северный и южный).

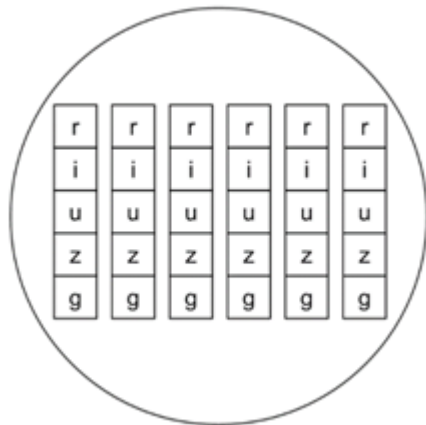


Рис. 5. Камера SDSS

Эксперименты проводилась в облачной инфраструктуре Azure³ на сборке Hadoop 2.6.0 HDInsight. Сырые изображения, полученные с сервера архива проекта SDSS, были предварительно преобразованы в формат SequenceFile.

³ Предоставлена компанией Microsoft в виде гранта по программе “Azure for Research”

Дело в том, что модель MapReduce эффективна в том случае, когда фрагмент данных, обрабатываемый одним Mapper (англ. split), по объему не сильно меньше размера блока HDFS. На сегодняшний момент с учетом характеристик существующих жестких дисков наиболее распространенным размером блока HDFS является 128МБ. А каждый исходный FITS-файл занимает около 12 МБ. Поэтому для эффективной обработки исходные файлы были преобразованы в файлы в формате SequenceFile, позволяющие разбивать их на логические части, близкие к размеру блока.

Эксперименты проводились на кластере HDInsight, построенном на узлах D12 из линейки типов узлов Azure, обладающих следующими характеристиками:

- число ядер: 4
- объем ОЗУ: 28 ГБ

Разработанный экспериментальный образец написан на Java и запускается с помощью командной строки, поддерживающей следующие опции:

- RA, DEC — координаты центра запрашиваемого участка в экваториальной системе координат выраженные в градусах. RA — аббревиатура от англ. Right Ascension (прямое восхождение), DEC — аббревиатура от англ. Declination (склонение). Для этих координат также используются обозначения α , δ .
- RA_SPAN, DEC_SPAN — ширина запрашиваемого участка по обеим координатам.
- RA_OVERLAP, DEC_OVERLAP — ширина перекрытия соседних клеток, на которые делится участок;
- RA_SUBREGIONS, DEC_SUBREGIONS — число клеток, на которые будет разбиваться участок вдоль обеих координат. Эти параметры позволяют задавать размеры одной клетки.
- FILTER — диапазон спектра, изображения в котором будут совмещаться. Один из пяти диапазонов, присутствующих в небесном обзоре SDSS DR12, обозначаемых буквами u, g, r, i и z.
- PIXEL_SCALE — размер одного пикселя выходного изображения в арксекундах. Например, в изображениях небесного обзора SDSS размер пикселя составляет 0.396 арксекунд.

Из всех указанных параметров на время работы прототипа непосредственно влияют те, которые задают *размер целевого участка неба* и *размеры клеток*. Число клеток выбиралось таким образом, чтобы размер одной клетки составлял примерно 0.7 градуса. В соответствии с этим был выбран участок неба с центром в точке с координатами $\alpha=51.5912$ град., $\delta=0.0131$ град. и размером 30 градусов по α и 2.5 градуса по δ . Размер по δ выбран так, чтобы участок вмещал в себя изображения, полученные всеми 6 столбцами ПЗС-камер телескопа (см. рис. 5). Участок разбивался на 60 частей по α и на 5 частей по δ , а перекрытие клеток (рамка) было выбрано равным 0.2 градусам,

таким образом, каждая клетка имела размеры 0.7 х 0.7 градусов и перекрывалась с соседними клетками.

Помимо параметров самого прототипа на время работы также влияют настройки MapReduce, в частности, размер фрагмента данных, подаваемого на вход одной задаче отображения (англ., map), и число задач свертки (англ., reduce). В ходе экспериментов число задач свертки выбиралось равным числу клеток. Размер фрагмента данных составил 128 МБ.

Пакет SWarp, лежащий в основе реализации, также имеет свои настройки, влияющие на производительность. Так, он позволяет задавать число рабочих нитей, которое бралось равным 2, и размер буфера для совмещения изображений в памяти. Использовался буфер размером в 1 ГБ.

Настройки кластера, на котором производились эксперименты, предусматривали 3 ГБ памяти для задач отображения и 5 ГБ памяти для задач свертки. Настройки, задающие максимальный размер кучи, выделяемой виртуальной машине Java (JVM), выбирались так, чтобы соответствовать этим значениям.

В приведенных ниже таблицах указан объем данных, занимаемых входными изображениями, общее время выполнения, а также время выполнения стадий шага свертки. Этих стадий три. Первая стадия, стадия перемешивания (англ., shuffle), включает в себя передачу вывода шага отображения задачам свертки. Стадия перемешивания включает в себя передачу данных по сети. Вторая стадия, стадия сортировки (англ., sort) сортирует ключи и значения, переданные задаче свертки. Благодаря сортировке значения разбиваются на группы. Наконец, следует стадия свертки (англ., reduce), на которой значения каждой группы сворачиваются и вывод записывается в распределенную файловую систему. Свертка в Hadoop может начинаться до того, как завершится отображение, поэтому общее время выполнения, вообще говоря, не является суммой времени выполнения шагов отображения и свертки. Временные показатели работы 1-го этапа обработки представлены в табл.1 и 2.

Табл.1. Результаты 1-го этапа обработки на 6 рабочих узлах

Объем данных	Время выполнения в минутах				
	Общее	Отображение	Перемешивание	Сортировка	Свертка
14 ГБ	34	19	10	1	14
21 ГБ	60	33	17	1	24
33 ГБ	79	45	24	1	31

Табл.2. Результаты 1-го этапа обработки на 12 рабочих узлах

Объем данных	Время выполнения в минутах				
	Общее	Отображение	Перемешивание	Сортировка	Свертка
14 ГБ	27	16	17	1	10
21 ГБ	33	17	17	1	14
33 ГБ	49	23	27	1	21

Рассматривая полученные результаты, стоит отметить, что затраты времени на шаге свертки в основном складываются из затрат на перемешивание результатов отображения и затрат на собственно свертку, то есть совмещение изображений. Время, требуемое на перемешивание, ожидаемо растет с увеличением объема данных. При этом затраты на перемешивание увеличиваются при использовании двенадцати рабочих узлов против изначальных шести, что можно объяснить издержками на передачу данных по сети.

В экспериментах по 2-му этапу обработки получены следующие результаты по производительности (табл.3, табл.4).

Табл.3. Результаты 2-го этапа обработки на 6 рабочих узлах

Объём данных	Время выполнения
14 ГБ	3 минуты 27 секунд
21 ГБ	3 минуты 47 секунд
33 ГБ	6 минут 6 секунд

Табл.4. Результаты 2-го этапа обработки на 12 рабочих узлах

Объём данных	Время выполнения
14 ГБ	1 минута 54 секунды
21 ГБ	2 минуты 55 секунд
33 ГБ	4 минуты 11 секунд

Графики на рис.6 иллюстрируют масштабируемость работы экспериментального образца.

По 1-му этапу обработки были проведены эксперименты на больших объемах данных (для 12 узлов кластера), результаты которых показаны в виде графика на рис. 7.

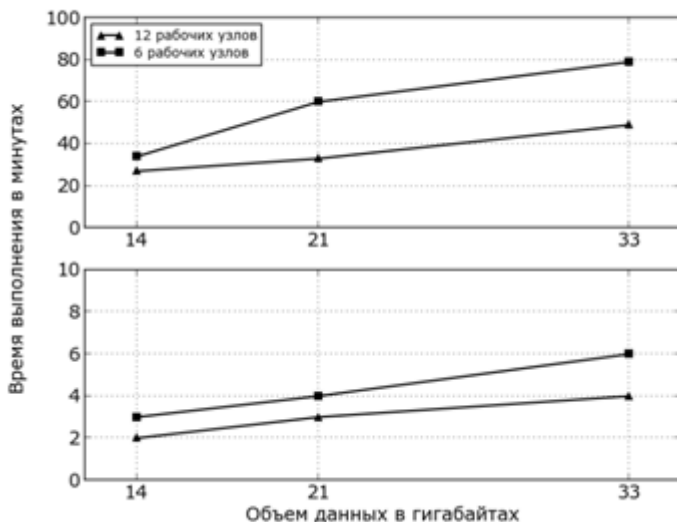


Рис 6. Время работы 1-го (сверху) и 2-го (снизу) этапов обработки в зависимости от количества узлов в кластере и объемов данных.

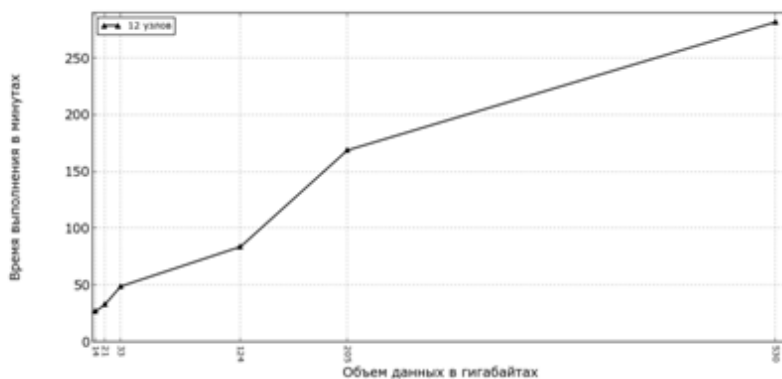


Рис. 7. Время работы 1-го этапа в зависимости от объема данных на 12 узлах

5. Вывод

Проведенные эксперименты подтвердили целесообразность применения предложенной архитектуры параллельного конвейера, основанной на MapReduce, для обработки наборов астрономических изображений, объемы которых находятся в пределах нескольких десятков ГБ. Легко доступная и дешевая инфраструктура Hadoop позволяет астрофизику заменить собой персональный компьютер в разы ускоряя вычисления (последовательная обработка 33ГБ данных из экспериментов заняла бы более 12 часов на персональном компьютере с 2 ядрами). Для апробации метода на больших объемах данных (5-20 ТБ) запланированы дополнительные эксперименты на обзорах неба SDSS и DES. Эксперименты позволят уточнить эффективность масштабирования и работоспособность решения на объемах изображений, соответствующих современным небесным обзорам.

Также планируется реализация ряда оптимизаций, в частности:

- оптимизация интерфейсов взаимодействия пакетов SWarp, SExtractor, PSFEx и инфраструктуры Hadoop;
- проработка механизма, позволяющего интегрировать в цепочку обработки свои реализации любых шагов конвейера и апробация механизма на оптимизированных версиях нескольких алгоритмов, в частности:
 - вычитание фоновой компоненты, основанное на нескольких соседних исходных изображениях (на 1-м этапе обработки);
- добавление возможности расширения набора “измеряемых” свойств небесных объектов (на 2-м этапе обработки).

Литература

- [1]. The Sloan Digital Sky Survey (SDSS) <http://www.sdss.org/>
- [2]. Burke B., Gregory J., Cooper M., Loomis A., Young D., Lind T., Doherty P., Daniels P., Landers D., Ciampi J., Johnson K., O'Brien P. CCD Imager Development for Astronomy. Lincoln Laboratory Journal, 2007, Volume 16, Number 2
- [3]. Subaru-HSC <http://www.naoj.org/Projects/HSC/>
- [4]. The Dark Energy Sky Survey (DES) <http://www.darkenergysurvey.org/>
- [5]. Pan-STARRS <http://pan-starrs.ifa.hawaii.edu/public/>
- [6]. The Large Synoptic Survey Telescope (LSST) <http://www.lsst.org/>
- [7]. Zhang Y., Zhao Y. Astronomy in the Big Data Era. Data Science Journal, 2015
- [8]. <http://www.astromatic.net/>
- [9]. Wiley K., Connolly A., Gardner J., Krughof S., Balazinska M., Howe B., Kwon Y., Bu Y. Astronomy in the Cloud: Using MapReduce for Image Coaddition. Publications of the Astronomical Society of the Pacific, 2011, Vol. 123, No. 901, pp. 366-380
- [10]. Apache Hadoop <http://hadoop.apache.org>
- [11]. Montage: an astronomical image mosaic engine <http://montage.ipac.caltech.edu/>
- [12]. Message Passing Interface Forum <http://www.mpi-forum.org/>
- [13]. Pegasus: workflow management system <http://pegasus.isi.edu/>

- [14]. Farivar R., Brunner R., Santucci R., Campbell R. Cloud Based Processing of Large Photometric Surveys. Astronomical Data Analysis Software and Systems XXII, 2013, p.91
- [15]. Dean J., Ghemawat S. MapReduce: Simplified Data Processing on Large Clusters. OSDI'04: Sixth Symposium on Operating System Design and Implementation, December 2004
- [16]. Koposov S., Belokurov V., Torrealba G., Wyn N. Evans Beasts of the Southern Wild: Discovery of nine Ultra Faint satellites in the vicinity of the Magellanic Clouds. The Astrophysical Journal, March 2015
- [17]. Apache HBase <http://hbase.apache.org>
- [18]. Apache Hive <http://hive.apache.org>
- [19]. Apache Spark <http://spark.apache.org>

Processing of Raw Astronomical Data of Large Volume by MapReduce Model⁴

¹S. Gerasimov < gerasimov@mlab.cs.msu.su >

²A. Mesheryakov < mesch@iki.rssi.ru >

¹I. Kolosov < zackwag32@gmail.com >

¹E. Glotov < glot.unltd@gmail.com >

¹I. Popov < ivan@mlab.cs.msu.su >

¹*Lomonosov Moscow State University Faculty CMC, 2nd Education Building,
GSP-1, Leninskie Gori, Moscow, 119991, Russian Federation*

²*Space Research Institute of the Russian Academy of Sciences,
84/32 Profsoyuznaya Str, Moscow, Russian Federation, 117997*

Abstract. Exponential grow of volume, increased quality of data in current (SDSS, DES, PanSTARRS) and incoming sky surveys (LSST) open new horizons for astrophysics but require new approaches to data processing especially big data technologies and cloud computing. This work presents a MapReduce-based approach to solve a major and important computational task in astrophysics - raw astronomical image data processing. We present architecture of Hadoop-based astrophysical pipeline which combines following steps of data processing: background removal, projection, co-addition, PSF-modelling, sky objects features extraction from images. The architecture uses modern implementations of astrophysical image processing algorithms from software packages SWarp, PSFEx, SExtractor. These tools are integrated in MapReduce procedures. The pipeline steps are joined in two phases. First phase - "raw" data processing - includes background removal, projection and images co-edition. Results of the first phase are preprocessed and co-added images into so called cells. Cells interleave by borders. Interleavings help us to process correctly on the second stage large sky objects on borders. The second stage includes steps of PSF-modelling and creation of the sky catalogue by extraction of sky object properties from cells. Experiments showed linear scalability of all processing steps and small impact of Hadoop infrastructure on entire

⁴ The project is supported by RFBR grant number 15-29-07085 ofi_m

performance costs. We used one filter data (red) from the Stripe82 dataset. All experiments are made inside cloud platform Microsoft Azure HDInsight.

Keywords: MapReduce; Hadoop; sky survey; big data; cloud computing; image processing

DOI: 10.15514/ISPRAS-2015-27(6)-20

For citation: Gerasimov S., Mesheryakov A., Kolosov I., Glotov E., Popov I. Processing of Raw Astronomical Data of Large Volume by MapReduce Model. Trudy ISP RAN/Proc. ISP RAS, vol. 27, issue 6, 2015, pp. 315-334 (in Russian). DOI: 10.15514/ISPRAS-2015-27(6)-20

References

- [1]. The Sloan Digital Sky Survey (SDSS) <http://www.sdss.org/>
- [2]. Burke B., Gregory J., Cooper M., Loomis A., Young D., Lind T., Doherty P., Daniels P., Landers D., Ciampi J., Johnson K., O'Brien P. CCD Imager Development for Astronomy. Lincoln Laboratory Journal, 2007, Volume 16, Number 2
- [3]. Subaru-HSC <http://www.naoj.org/Projects/HSC/>
- [4]. The Dark Energy Sky Survey (DES) <http://www.darkenergysurvey.org/>
- [5]. Pan-STARRS <http://pan-starrs.ifa.hawaii.edu/public/>
- [6]. The Large Synoptic Survey Telescope (LSST) <http://www.lsst.org/>
- [7]. Zhang Y., Zhao Y. Astronomy in the Big Data Era. Data Science Journal, 2015
- [8]. <http://www.astromatic.net/>
- [9]. Wiley K., Connolly A., Gardner J., Krughof S., Balazinska M., Howe B., Kwon Y., Bu Y. Astronomy in the Cloud: Using MapReduce for Image Coaddition. Publications of the Astronomical Society of the Pacific, 2011, Vol. 123, No. 901, pp. 366-380
- [10]. Apache Hadoop <http://hadoop.apache.org>
- [11]. Montage: an astronomical image mosaic engine <http://montage.ipac.caltech.edu/>
- [12]. Message Passing Interface Forum <http://www.mpi-forum.org/>
- [13]. Pegasus: workflow management system <http://pegasus.isi.edu/>
- [14]. Farivar R., Brunner R., Santucci R., Campbell R. Cloud Based Processing of Large Photometric Surveys. Astronomical Data Analysis Software and Systems XXII, 2013, p.91
- [15]. Dean J., Ghemawat S. MapReduce: Simplified Data Processing on Large Clusters. OSDI'04: Sixth Symposium on Operating System Design and Implementation, December 2004
- [16]. Koposov S., Belokurov V., Torrealba G., Wyn N. Evans Beasts of the Southern Wild: Discovery of nine Ultra Faint satellites in the vicinity of the Magellanic Clouds. The Astrophysical Journal, March 2015
- [17]. Apache HBase <http://hbase.apache.org>
- [18]. Apache Hive <http://hive.apache.org>
- [19]. Apache Spark <http://spark.apache.org>

