

Big data

Подходы, инструменты и применение в ФП

Наблюдения

- 1.8 зетабайт данных создано в 2011
- Объем данных удваивается каждые 40 месяцев
- Общий объем данных на серверах Google ~10 эксабайт
- Интерес к big data вырос в 10 раз за последние 3 года

Big data?

Данные объемом > 1 ТБ (вплоть до пета-, экса-)

- сохранение
- передача
- поиск
- анализ
- визуализация

About 2,300,000,000 results (0.23 seconds)

Ads related to **big data** ⓘ

[**Big Data** in 2013 - tableausoftware.com](http://www.tableausoftware.com)

www.tableausoftware.com/big-data ▾

7 Things you Need to Do About **Big Data** in 2013. Get the Free Article!

[**Big Data** Graph Database - Objectivity.com](http://www.objectivity.com)

www.objectivity.com/bigdata ▾ 1 (855) 601 4951

Solutions for Complex **Big Data** The **Big Data** Connection Platform

InfiniteGraph Features - Objectivity News - Download

[**Real-Time Big Data** - SolaceSystems.com](http://www.solacesystems.com)

www.solacesystems.com/ ▾

High throughput **Big Data** distribution on LAN, WAN and web.

[**Big data** - Wikipedia, the free encyclopedia](http://en.wikipedia.org/wiki/Big_data)

en.wikipedia.org/wiki/Big_data ▾

Big data is the term for a collection of data sets so large and complex that it becomes difficult to process using on-hand database management tools or traditional ...

[Statistical inference](#) - [Unstructured data](#) - [Connectomics](#) - [Descriptive statistics](#)

Где встречается big data

В науке:

- Метеорология
- Физ моделирование
- Геномика
- Экономика

В производстве:

- High load
- Поиск в Интернете

Big data

Подходы

Подходы big data

- MapReduce
- Bulk Synchronous Parallel
- “Сделай сам”

MapReduce

MapReduce

- Разделяем большие данные на маленькие куски
- Работаем с каждым куском независимо
- Агрегируем результаты

MapReduce

Map(k_1, v_1) \rightarrow list(k_2, v_2)

Reduce($k_2, \text{list}(v_2)$) \rightarrow list(k_3, v_3)

MapReduce: подсчет числа слов

```
function map(String name, String documentPart):
```

```
  for each word w in documentPart:
```

```
    emit (w, 1)
```

```
function reduce(String word, List<Int> partialCounts):
```

```
  sum = 0
```

```
  for each pc in partialCounts:
```

```
    sum += pc
```

```
  emit (word, sum)
```

MapReduce: область применения

Хорошо:

- Distributed sorting
- Pattern matching
- Indexing
- Machine learning

Плохо:

- Global shared state
- Small data

Bulk Synchronous Parallel

Bulk Synchronous Parallel (BSP)

Each iteration (superstep):

1. Concurrent computation
2. Communication
3. Barrier synchronisation

BSP: Область применения

- Итерационные алгоритмы

Graph Computing

1. Вход - граф
2. Для каждой вершины вычисляем её функцию (итеративно)
3. Выход - граф

Graph Computing

Вершина:

- Получает сообщения с пред. итерации
- Изменяет своё состояние
- Создаёт/удаляет исходящие рёбра
- Изменяет состояние исходящих ребёр
- Отправляет сообщения по исходящим рёбрам

Graph Computing: shortest path

```
function compute(Vertex me, List<Int> messages):  
  minDist = me.isStartVertex() ? 0 : INF  
  for each otherDist in messages:  
    if otherDist < minDist: minDist = otherDist  
  if minDist < me.value:  
    me.setValue(minDist)  
    for each edge in me.edges:  
      me.sendMessage(edge, minDist + edge.value)  
  else:  
    me.voteToHalt()
```

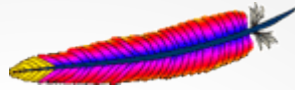
Big data

Инструменты

Требования к инструментам big data

- Масштабируемость
- Отказоустойчивость

Apache Hadoop



- Де-факто платформа для big data
- Реализует модель MapReduce

Apache Hadoop: модули

- **Common** - общие библиотеки и утилиты
- **Distributed File System** - файловая система
- **YARN** - менеджмент ресурсов
- **MapReduce** - реализация модели MapReduce

Hadoop Distributed File System (HDFS)

- Master/slave архитектура
- Репликация
- Кросс-платформенность

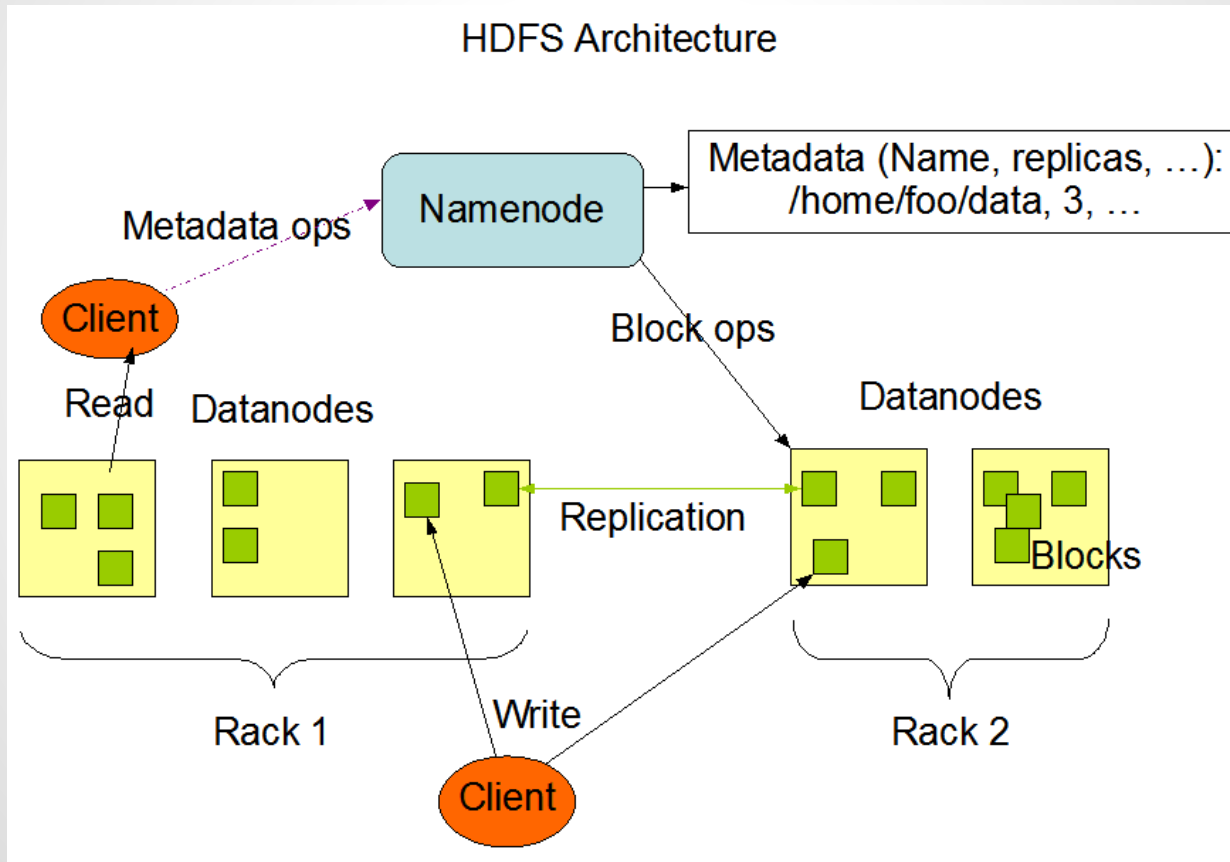
HDFS: Namenode

- Хранит метаданные о FS
- Производит журналирование
- Следит за репликацией
- Следит за состоянием узлов данных (Datanode) и блоков на них

HDFS: Datanode

- Хранит блоки данных в виде файлов
- Отчитывается перед Namenode

HDFS: Architecture



Data Replication

- Доступность
- Быстрое чтение (локальность)
- Быстрая запись

HDFS: Replication

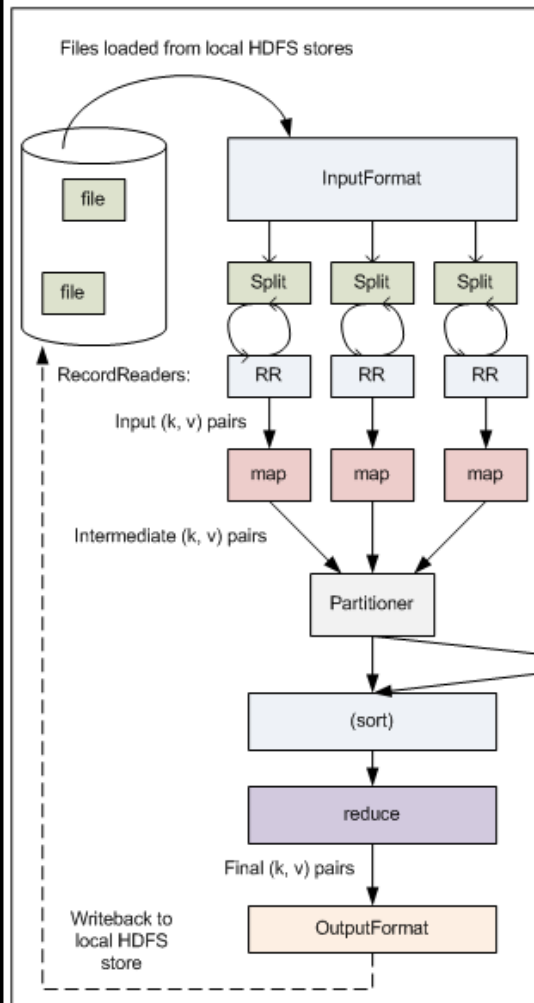
- “Rack awareness”
- Конфигурируется отдельно для отдельных файлов
- По умолчанию - 1 блок на локальной стойке, 2 блока на удаленной

Hadoop MapReduce

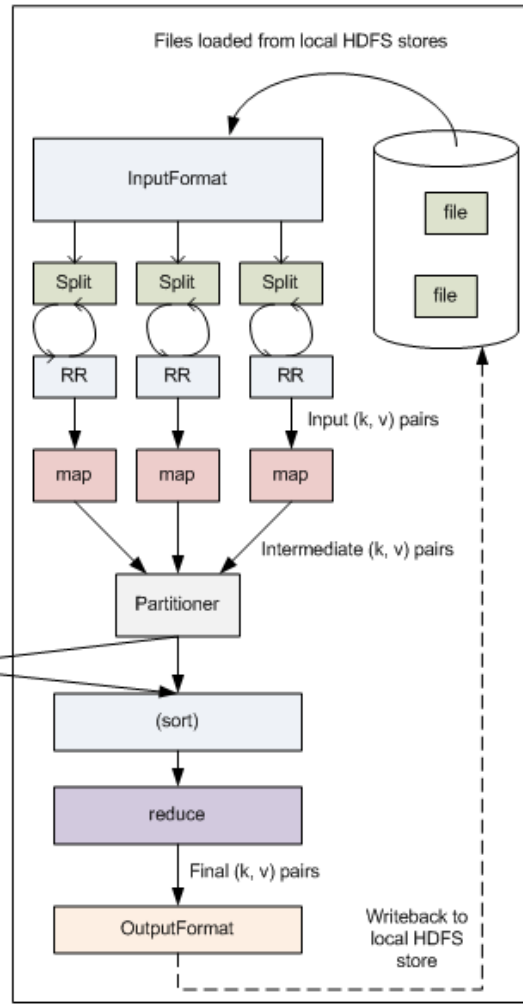
Пользователь определяет:

- Input reader
- Функцию Map
- Функцию partition
- Функцию compare
- Функцию Reduce
- Output writer

Node 1



Node 2



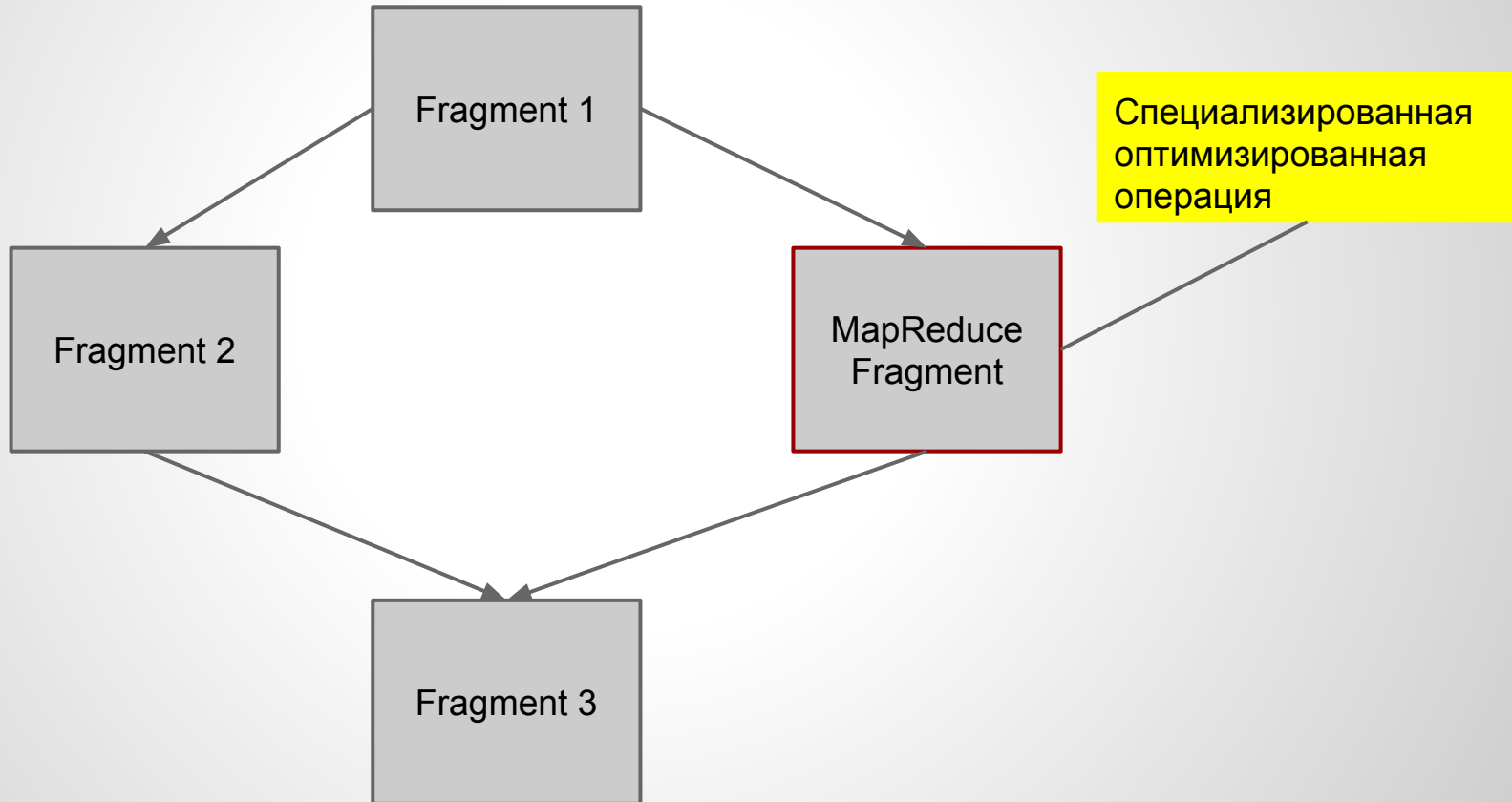
"Shuffling" process

Intermediate (k, v) pairs exchanged by all nodes

Big data

Применение в ФП

Big data в ФП



Big data в ФП: проблемы

- Работа с файлами = побочные эффекты
- Интеграция с библиотеками на языках, отличных от C/C++
- Необходимость поддержки модульности
- Необходимость поддержки функций высшего порядка

Big data в ФП: планы

1. Интеграция LuNA с библиотекой MapReduce-MPI
2. Перенос решения на Hadoop
3. Расширение для поддержки других языков (Java)