BIGDATA TEAM

X5 Course: Big Data

Hadoop MapReduce

Алексей Драль, aadral@bigdatateam.org
CEO at BigData Team, http://bigdatateam.org
https://www.facebook.com/bigdatateam

25.02.2021, Moscow, Russia

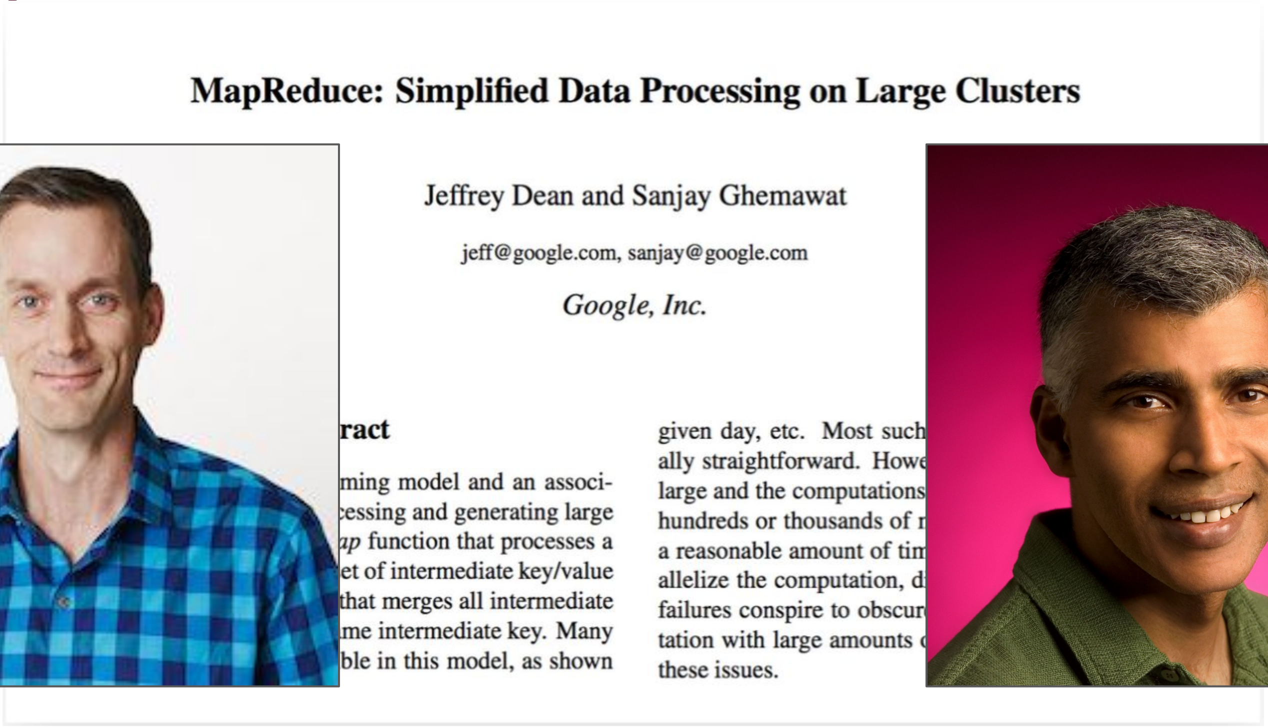
BIGDATA TEAM

План

- Проверка знаний (quiz)
- MapReduce (MR)
- Распределенные консольные утилиты
- Fault Tolerance
- -- (перерыв)
- MapReduce Streaming + Workshop

MapReduce (MR)

MapReduce



MapReduce: Simplified Data Processing on Large Clusters

Jeffrey Dean and Sanjay Ghemawat

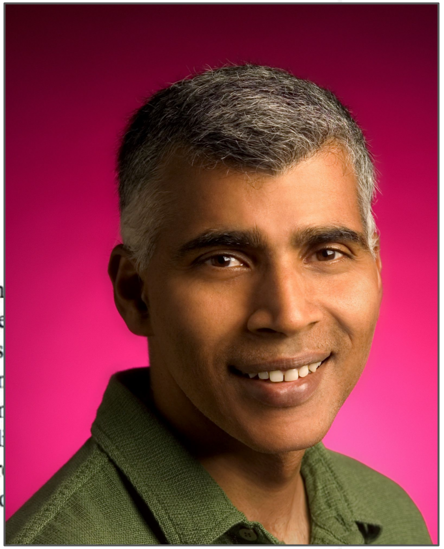
jeff@google.com, sanjay@google.com

Google, Inc.

ract

ming model and an associ-
cessing and generating large
ap function that processes a
et of intermediate key/value
that merges all intermediate
me intermediate key. Many
ble in this model, as shown

given day, etc. Most such
ally straightforward. Howe
large and the computations
hundreds or thousands of r
a reasonable amount of tim
allelize the computation, d
failures conspire to obscur
tation with large amounts o
these issues.



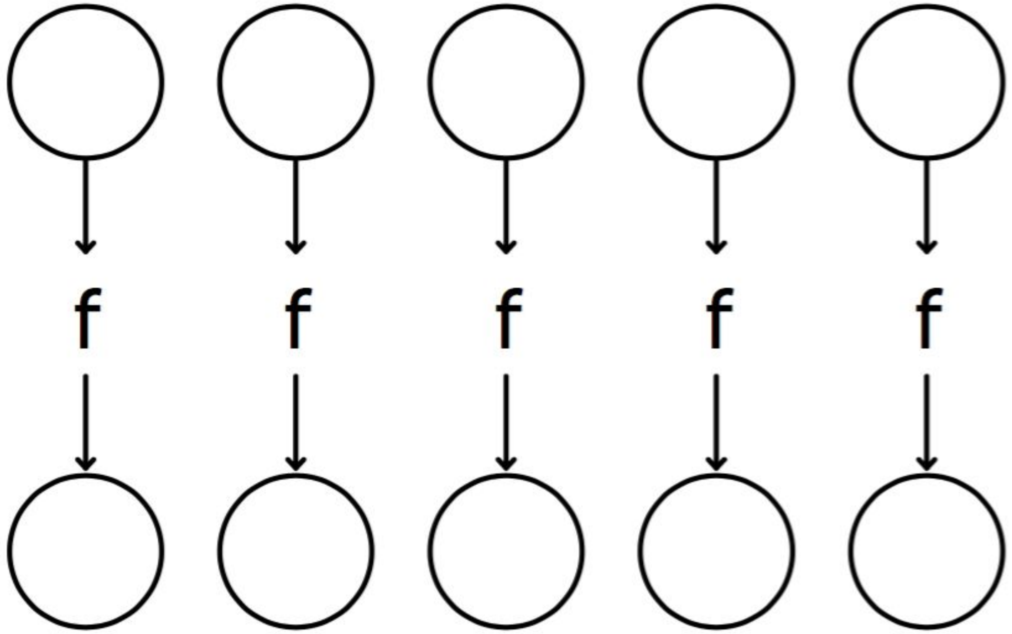
MapReduce: Simplified Data Processing on Large Clusters, Symposium on Operating Systems Design and Implementation (OSDI, 2004)

Jeffrey Dean

- Когда Jeff Dean разрабатывает ПО, он сначала создает бинарник, а потом пишет исходный код как документацию.
- Однажды Jeff Dean не прошел тест Тьюринга, потому что корректно посчитал 203 число Фибоначчи менее чем за 1 секунду.
- Скорость, с которой Jeff Dean разрабатывает ПО выросла в 40 раз в конце 2000, когда он обновил свою клавиатуру до USB2.0.
- Вы используете только 10% мозга. Остальные 90% используются под запуск MapReduce задач Джефа.

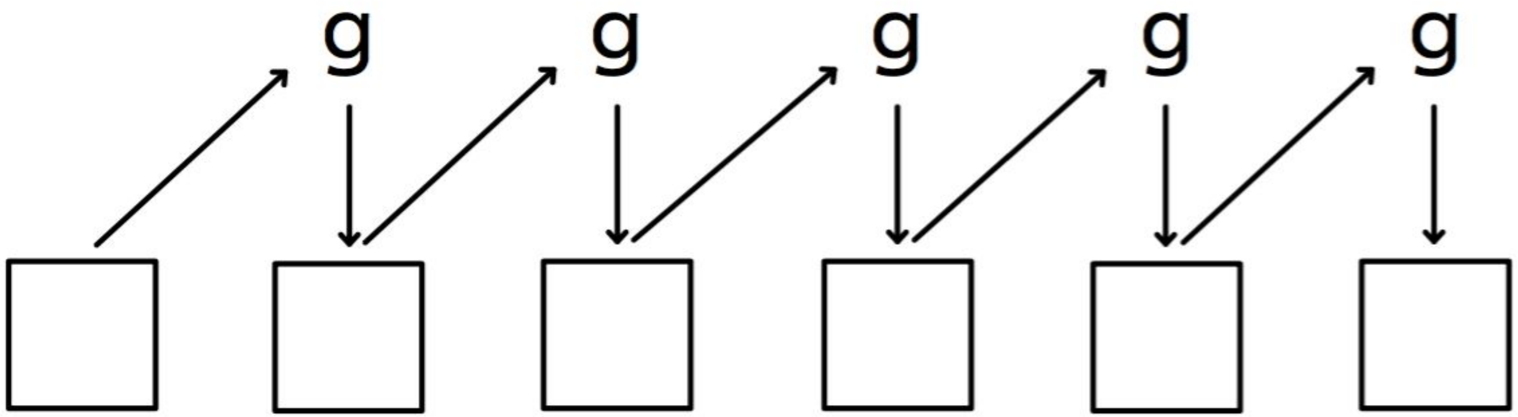
https://www.quora.com/What-are-all-the-Jeff-Dean-facts

Map



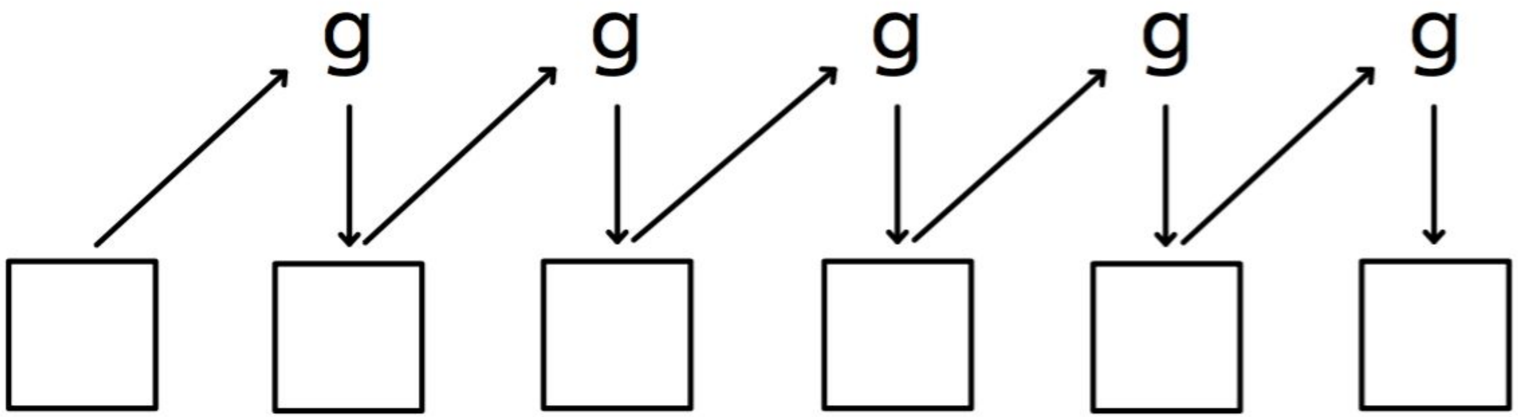
```
>>> map(lambda x: x*x, [1,2,3,4])
[1,4,9,16]
```

Fold / Reduce / Aggregate



```
>>> reduce(operator.sum, [1, 4, 9, 16])
>>> reduce(operator.sum, [5, 9, 16])
>>> reduce(operator.sum, [14, 16])
30
```

Fold / Reduce / Aggregate

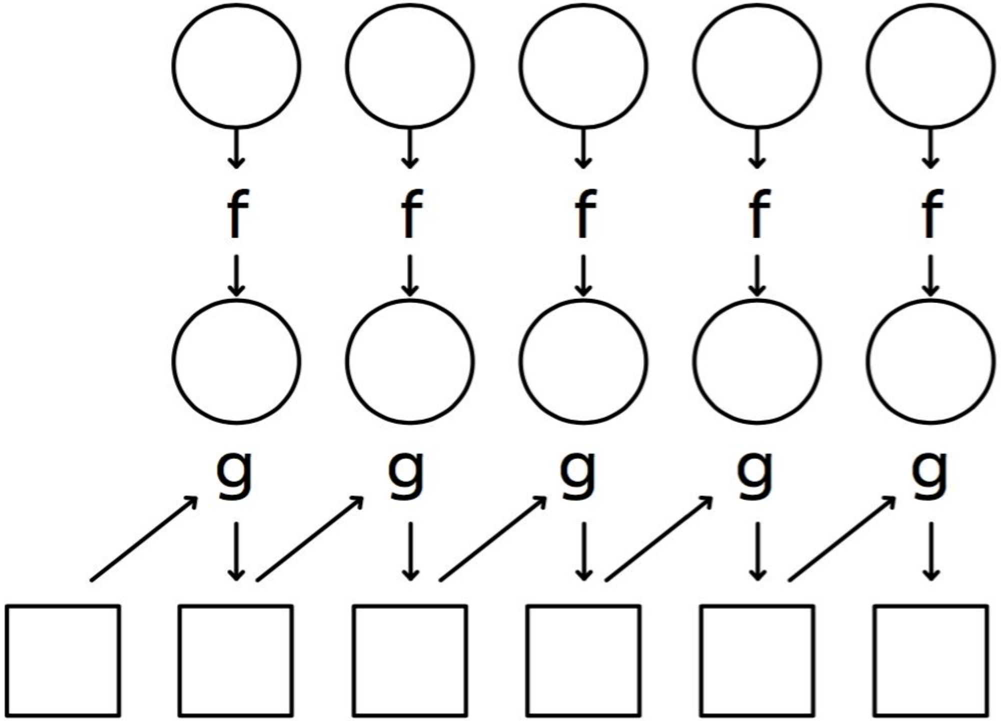


```
>>> average = lambda x, y: (x + y) / 2.
```

```
>>> reduce(average, [1, 2, 3])
2.25
```

```
>>> reduce(average, [3, 2, 1])
1.75
```

MapReduce



```
>>> reduce(operator.add, map(lambda x: x*x, [1, 2, 3, 4]))

30
```

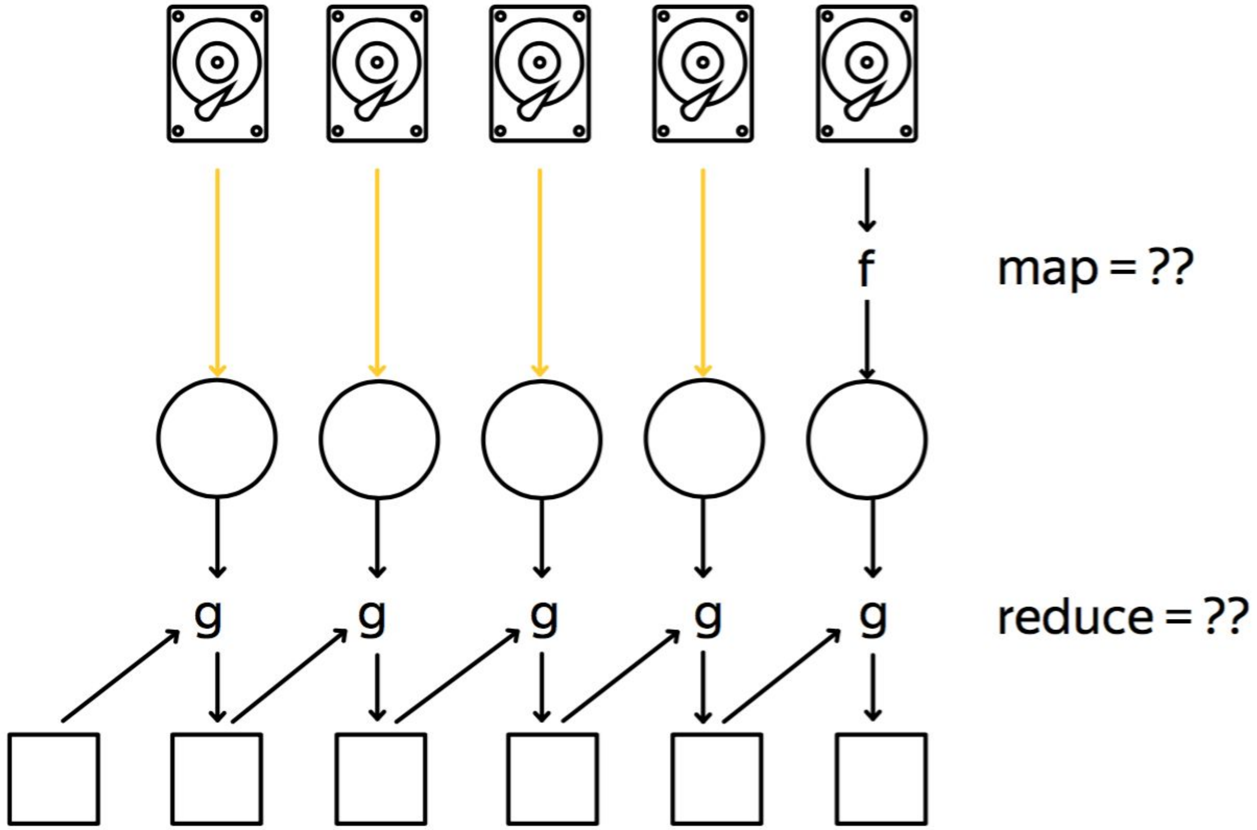
распределенные консольные утилиты

grep

```
$ grep <pattern> <file>
$ grep "hadoop" A.txt
Repository git-wip-us.apache.org/repos/asf/hadoop.git
Website  hadoop.apache.org
$ grep -i "hadoop" A.txt
Apache Hadoop
Apache Hadoop
Hadoop Logo
Repository git-wip-us.apache.org/repos/asf/hadoop.git
Website  hadoop.apache.org
Apache Hadoop ( /hə`du:p/) is
$ man grep
```

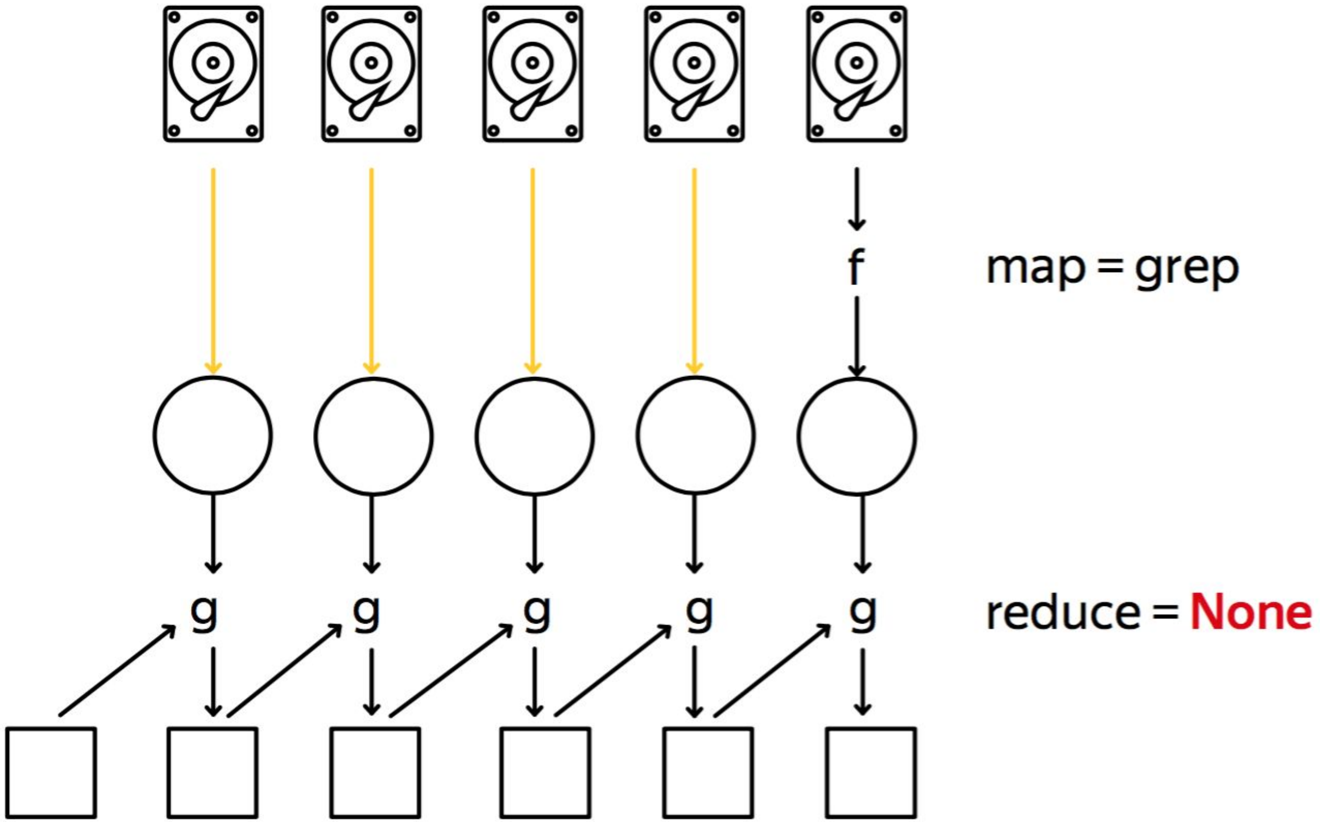
BIGDATA TEAM

Distributed Shell: grep



BIGDATA TEAM

Distributed Shell: grep



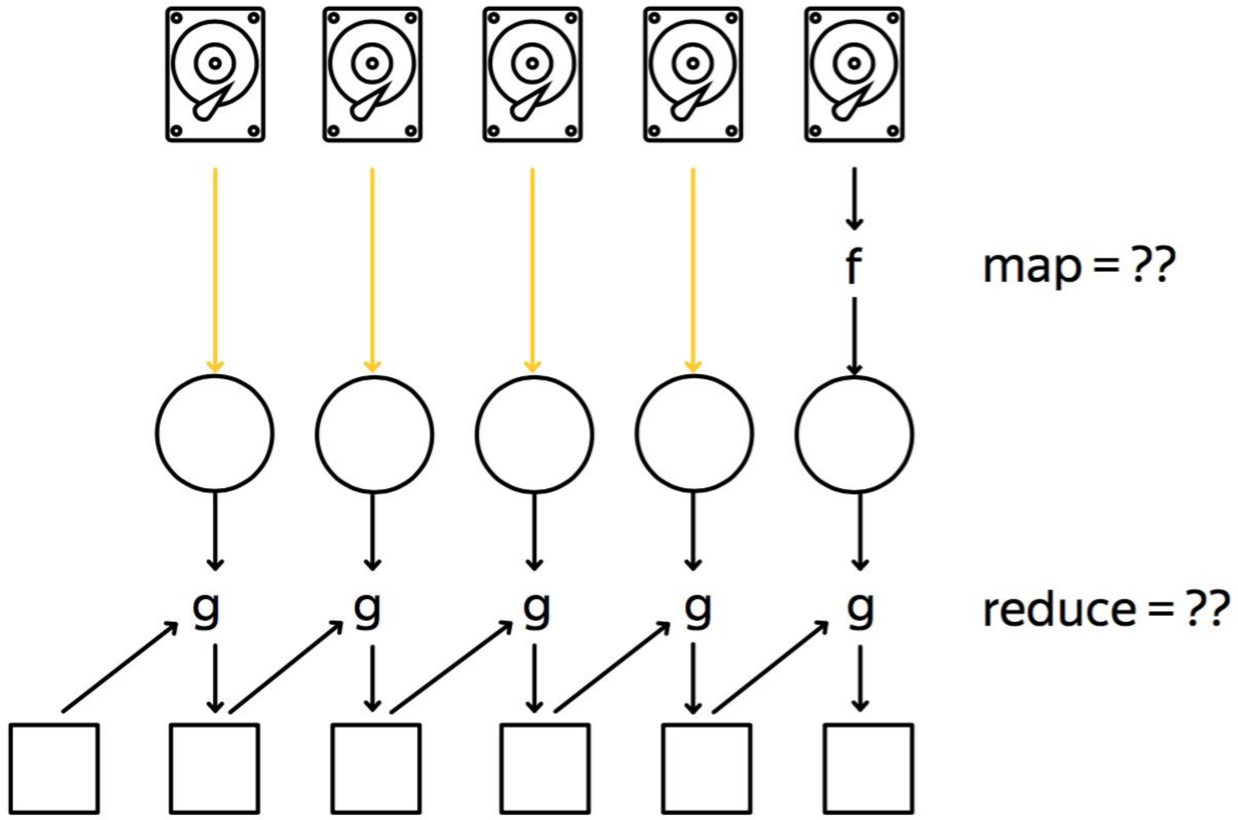
head

```
$ head <file>
$ head A.txt
```

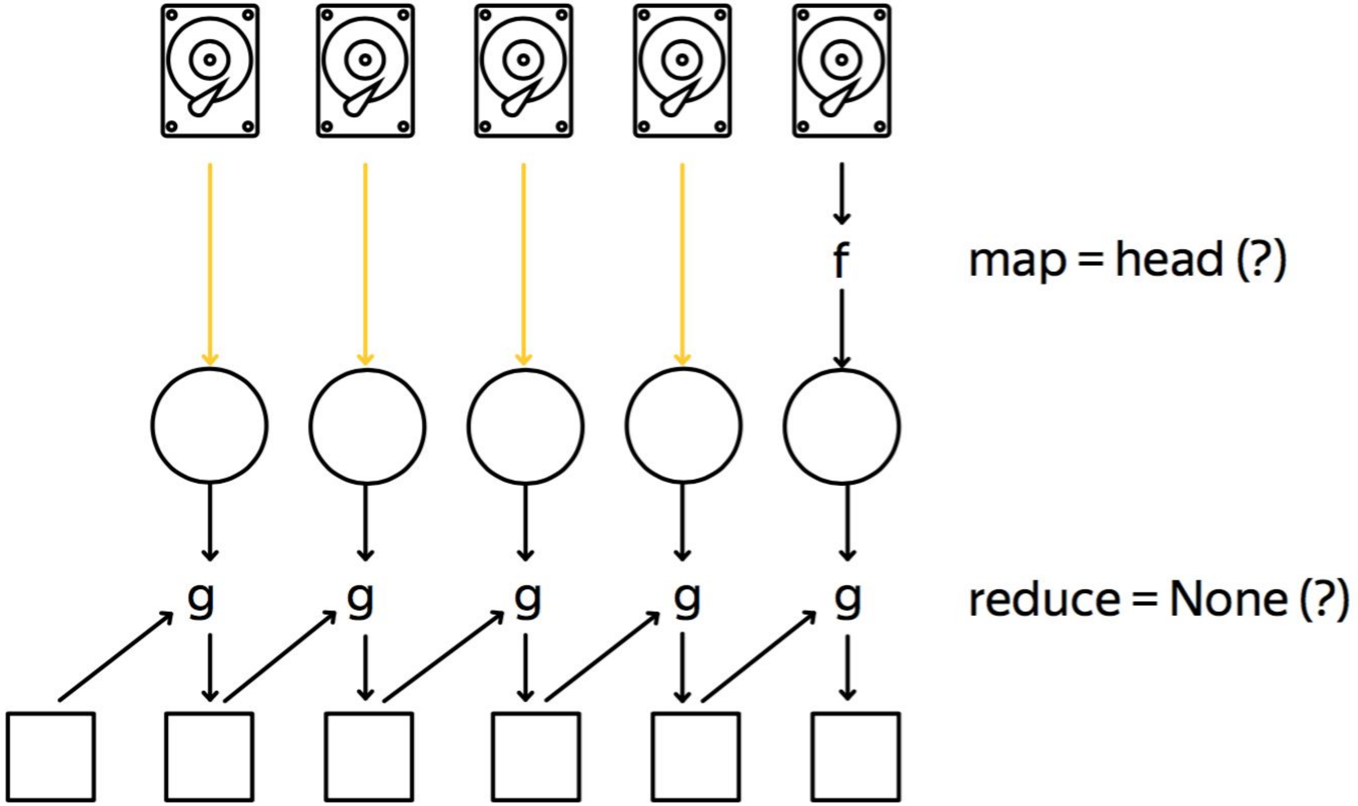
```
Apache Hadoop
From Wikipedia, the free encyclopedia
[hide]This article has multiple issues. Please help improve it or discuss these is-
sues on the talk page. (Learn how and when to remove these template messages)
This article contains content that is written like an advertisement. (October 2013)
This article appears to contain a large number of buzzwords. (October 2013)
This article may be too technical for most readers to understand. (May 2017)
Apache Hadoop
Hadoop Logo
Developer(s)Apache Software Foundation
```

BIGDATA TEAM

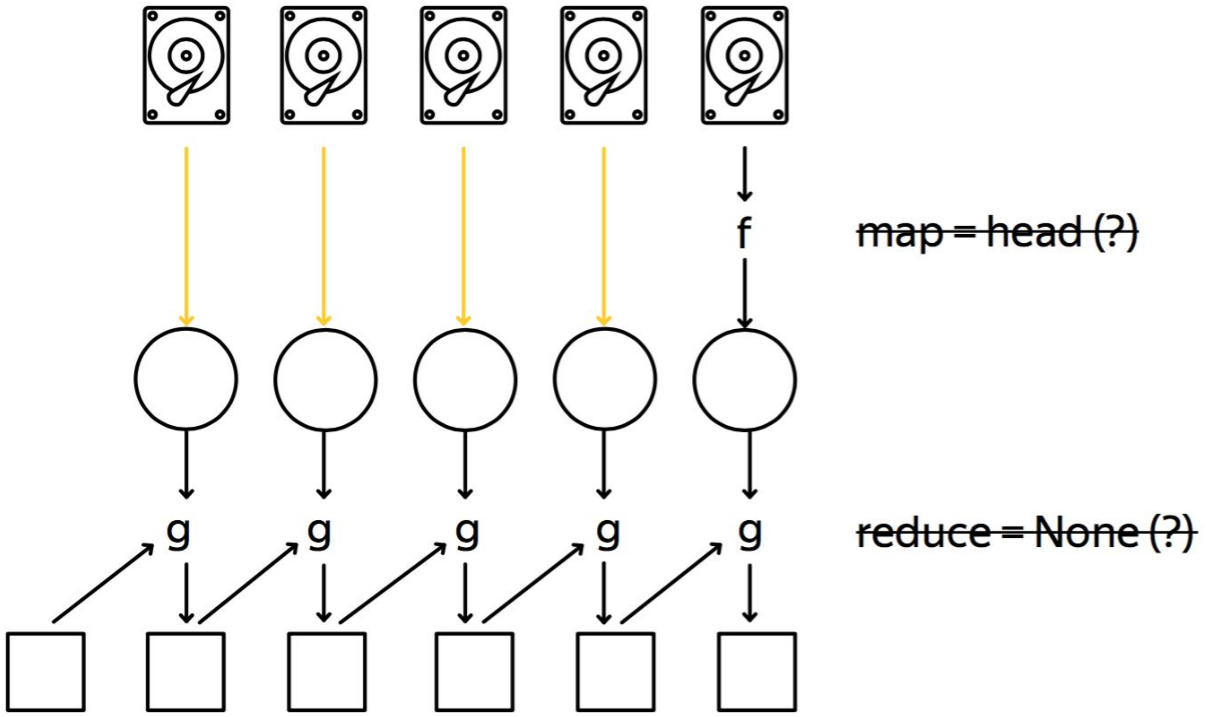
Distributed Shell: head



Distributed Shell: head

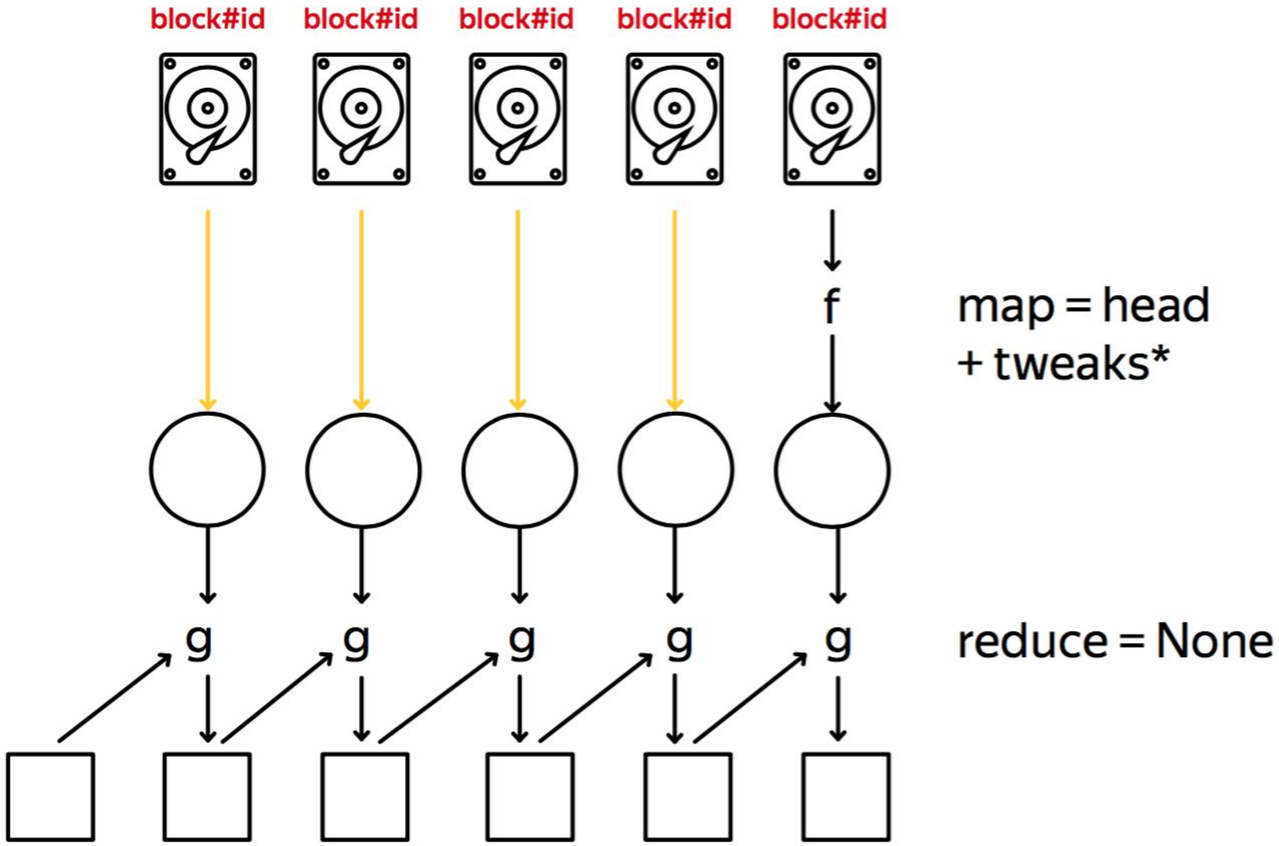


Distributed Shell: head



```
HDFS v.2*: hdfs dfs -text distributed_A.txt | head
HDFS v.3+: hdfs dfs -head distributed_A.txt
```

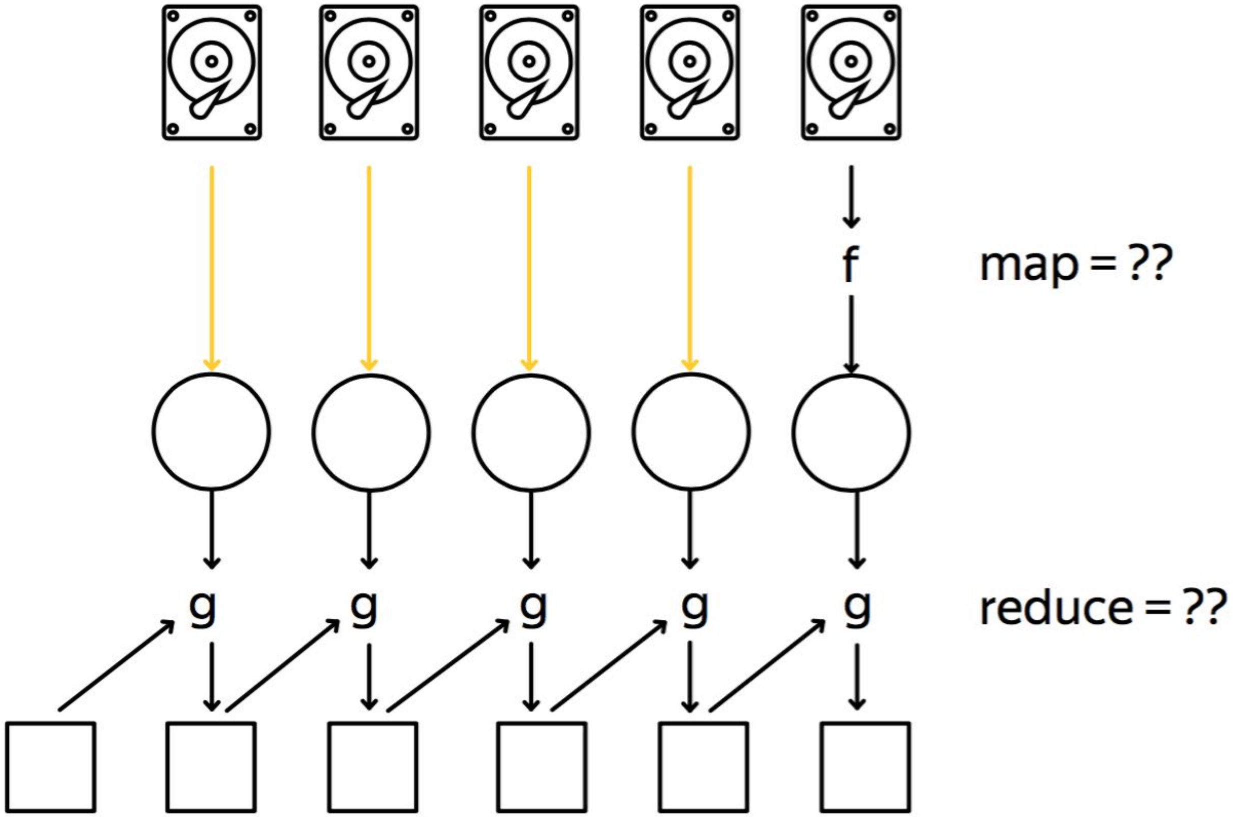
Distributed Shell: head



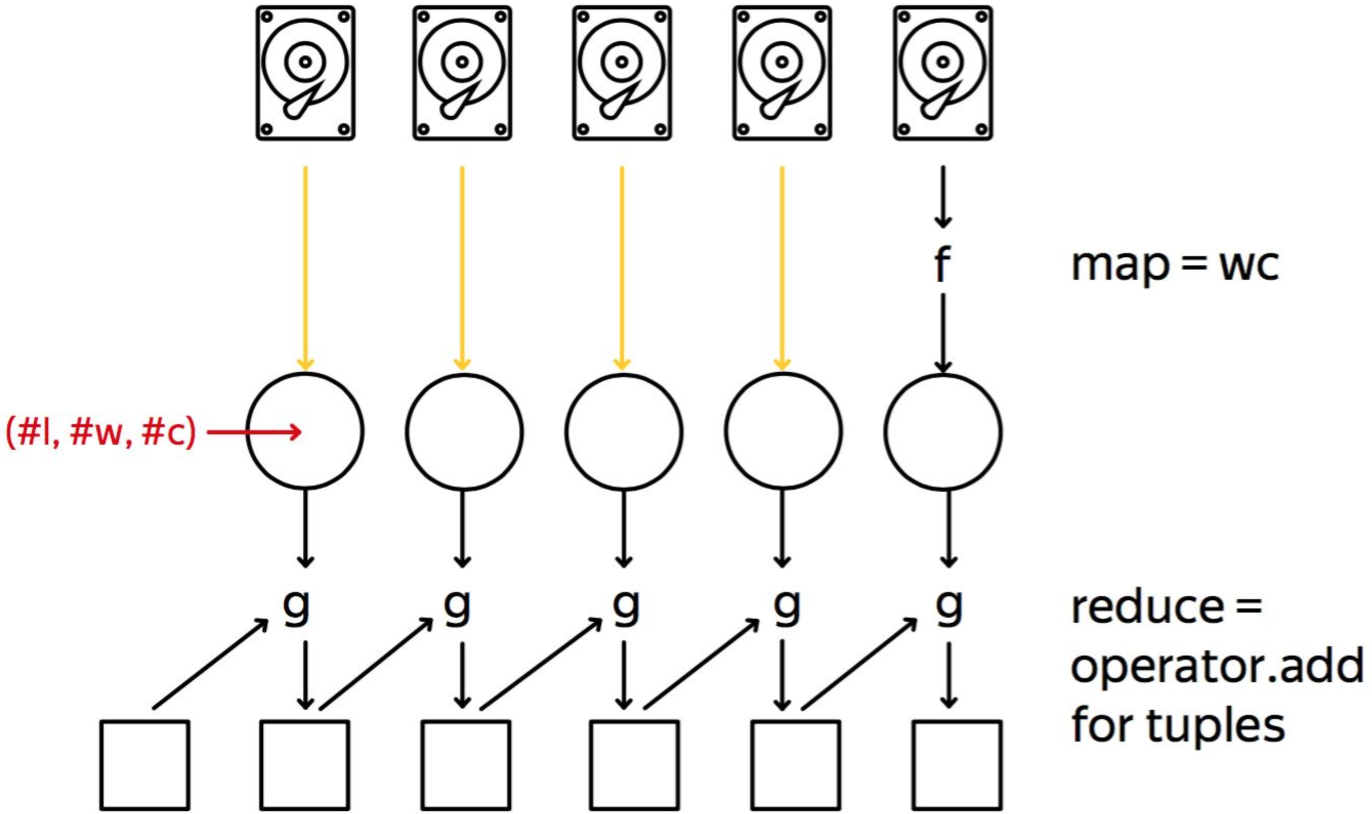
wc

```
$ wc <file>
$ wc A.txt
269  4319  28001  A.txt
```

Distributed Shell: wc



Distributed Shell: wc



Word Count

Apache Hadoop (/həˋdu:p/) is an open-source software framework used for distributed storage and processing of dataset of big data using the MapReduce programming model. It consists of computer clusters built from commodity hardware.

All the modules in Hadoop are designed with a fundamental assumption that hardware failures are common occurrences and should be automatically handled by the framework…

↓

```
'the': 3, 'of': 3, 'hadoop': 2, …
```

Word Count на одном компьютере

```
$ cat dataset.txt
```

```
   Apache Hadoop is a collection of open-source
software utilities that facilitates using a
network of many computers to solve problems
involving massive amounts of data and computation.
It provides a software framework for distributed
storage and processing of big data using the
MapReduce programming model...
```

BIGDATA TEAM

Word Count на одном компьютере

```
$ cat dataset.txt | tr ' ' '\n'
```

```
   Apache
   Hadoop
   is
   a
   collection
   of
   ...
```

Word Count на одном компьютере

```
$ cat dataset.txt | tr ' ' '\n' | sort
```

```
   All
   Apache
   Hadoop
   Hadoop
   Hadoop
   It
   ...
```

BIGDATA TEAM

Word Count на одном компьютере

```
$ cat dataset.txt | tr ' ' '\n' | sort | uniq -c
```

```
   1 All
   1 Apache
   3 Hadoop
   2 It
   1 MapReduce
   4 a
   ...
```

Что делаем с “sort”?

Распределенный Word Count

```
$ cat dataset.txt | tr ' ' '\n' | sort | uniq -c
```

❌ Фаза Map (нужна агрегация)
❌ Фаза Reduce (не хватит RAM / HDD)

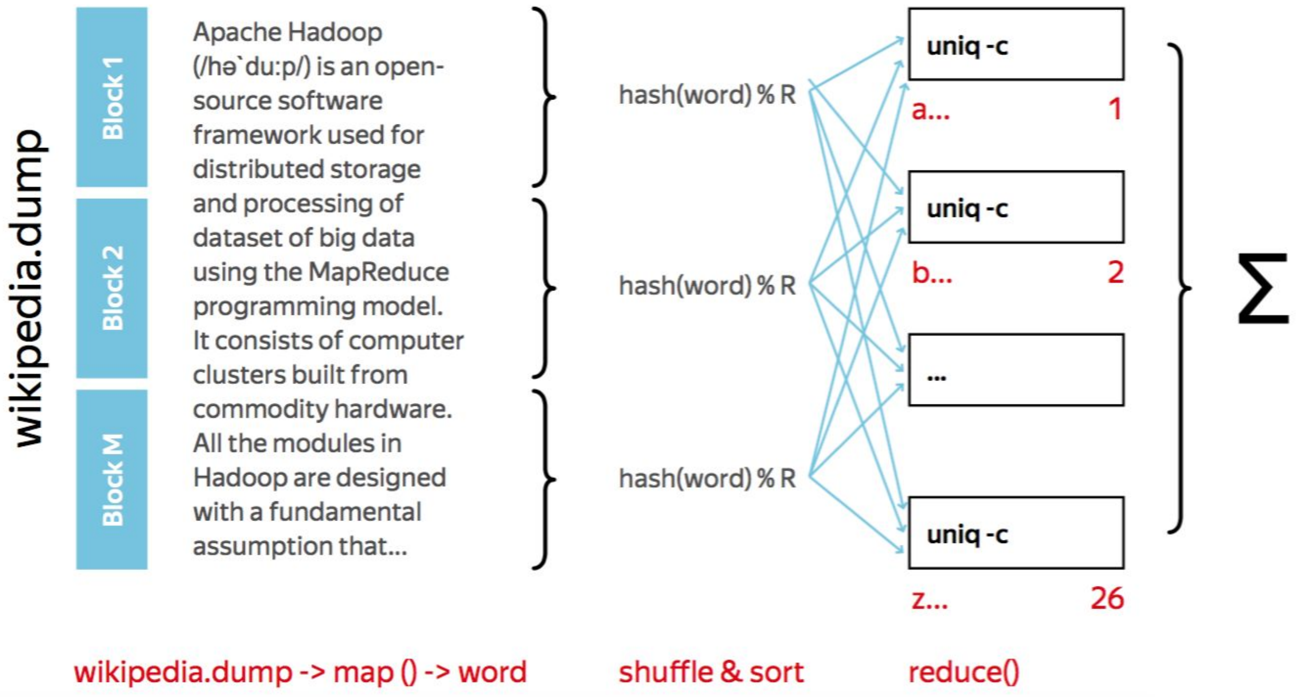
MapReduce =

Map → Shuffle & Sort → Reduce

BIGDATA TEAM

MapReduce (example)

```
wikipedia.dump | tr ' ' '\n'  |    sort     |      uniq -c
```


wikipedia.dump
Block 1
Block 2
Block M
Apache Hadoop (/həˈdu:p/) is an open-source software framework used for distributed storage and processing of dataset of big data using the MapReduce programming model. It consists of computer clusters built from commodity hardware. All the modules in Hadoop are designed with a fundamental assumption that...
hash(word) % R
hash(word) % R
hash(word) % R
uniq -c
a... 1
uniq -c
b... 2
...
uniq -c
z... 26
Σ
wikipedia.dump -> map () -> word
shuffle & sort
reduce()

Формальная модель MapReduce

Фазы:

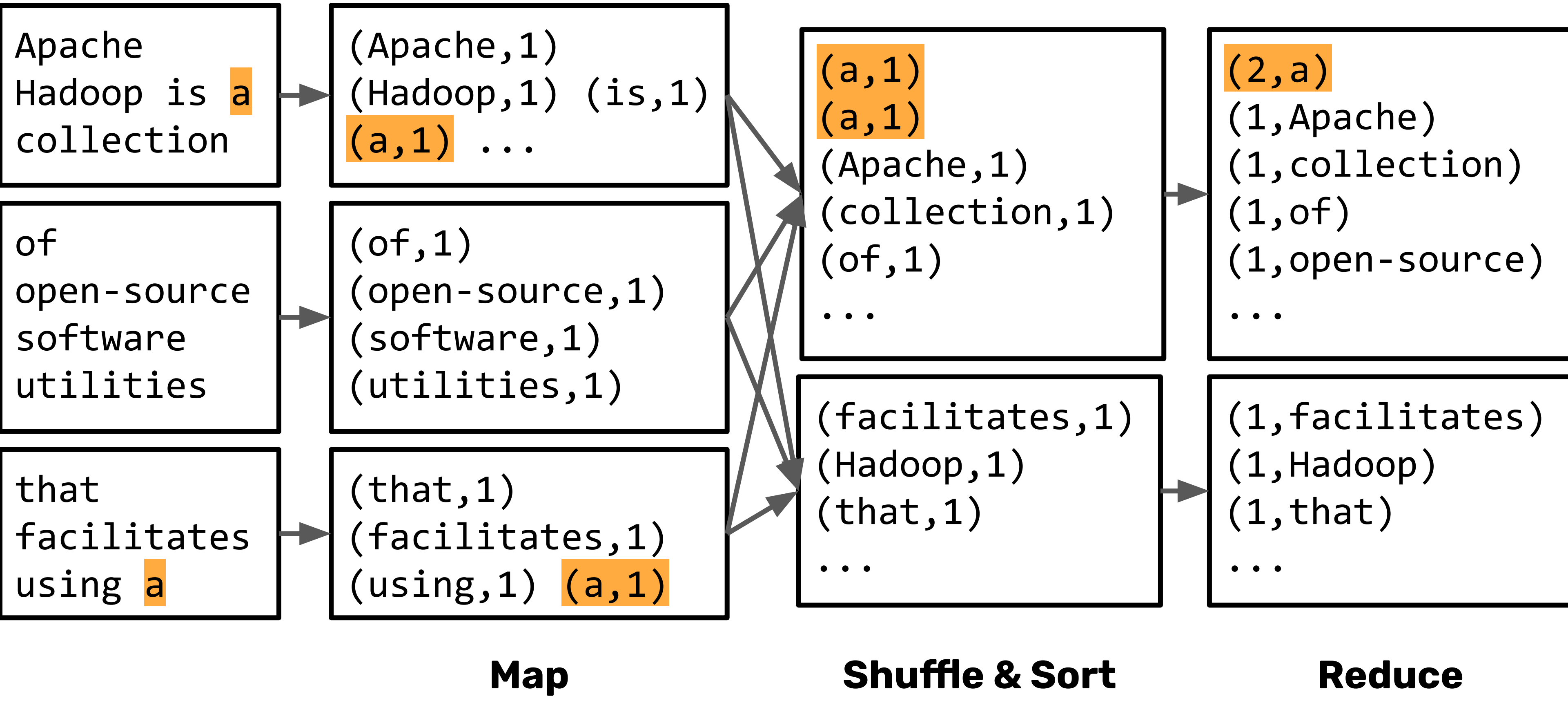
1. Map
2. Shuffle & Sort
3. Reduce

Worker'ы (контейнеры):

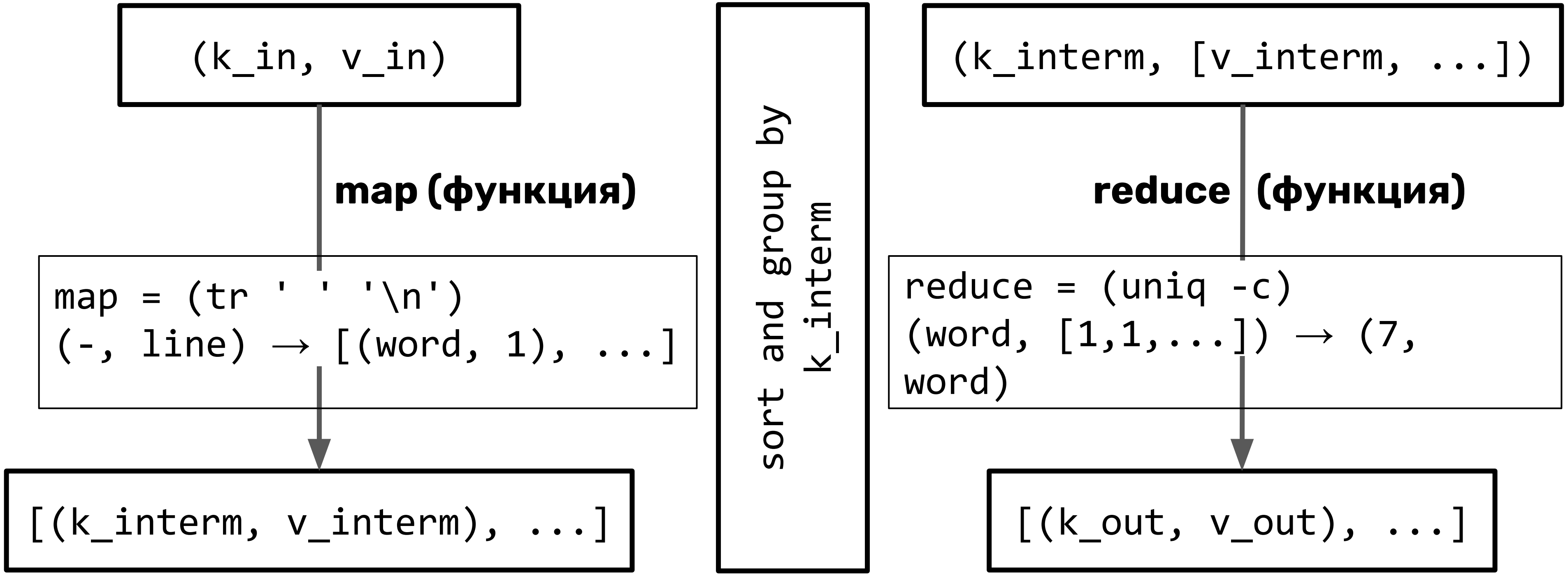
- Фаза Map → Mapper (использует функцию map)
- Фаза Reduce → Reducer (использует функцию reduce)

BIGDATA TEAM

Word Count v.2


Apache Hadoop is a collection
of open-source software utilities
that facilitates using a
(Apache,1) (Hadoop,1) (is,1) (a,1) ...
(of,1) (open-source,1) (software,1) (utilities,1)
(that,1) (facilitates,1) (using,1) (a,1)
(a,1) (a,1) (Apache,1) (collection,1) (of,1) ...
(facilitates,1) (Hadoop,1) (that,1) ...
(2,a) (1,Apache) (1,collection) (1,of) (1,open-source) ...
(1,facilitates) (1,Hadoop) (1,that) ...
Map
Shuffle & Sort
Reduce

Формальная модель Word Count



(k_in, v_in)
map (функция)
[(k_interm, v_interm), ...]
sort and group by k_interm
(k_interm, [v_interm, ...])
reduce (функция)
[(k_out, v_out), ...]

```
map = (tr ' ' '\n')
(-, line) → [(word, 1), ...]
```

```
reduce = (uniq -c)
(word, [1,1,...]) → (7, word)
```

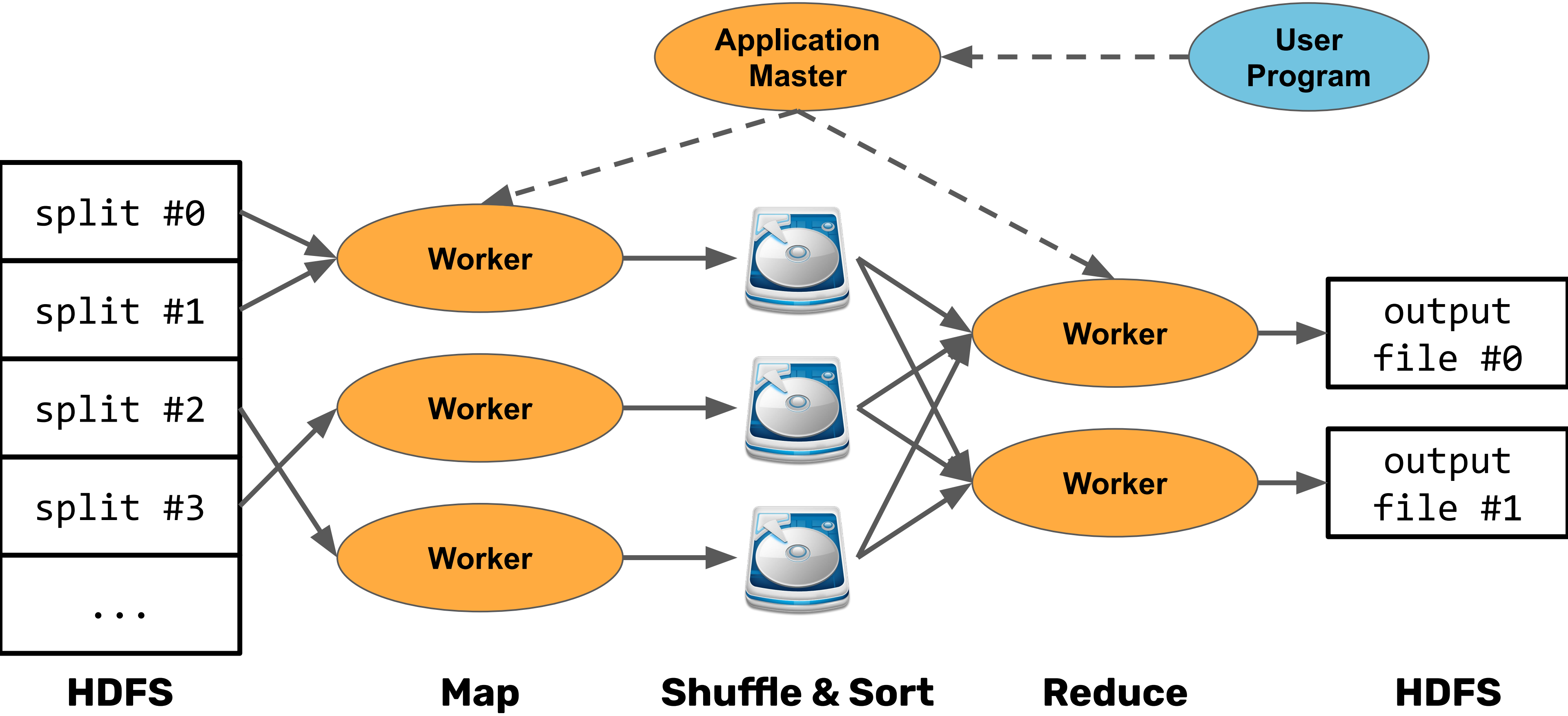
Map (фаза) **Shuffle & Sort** **Reduce (фаза)**

BIGDATA
TEAM

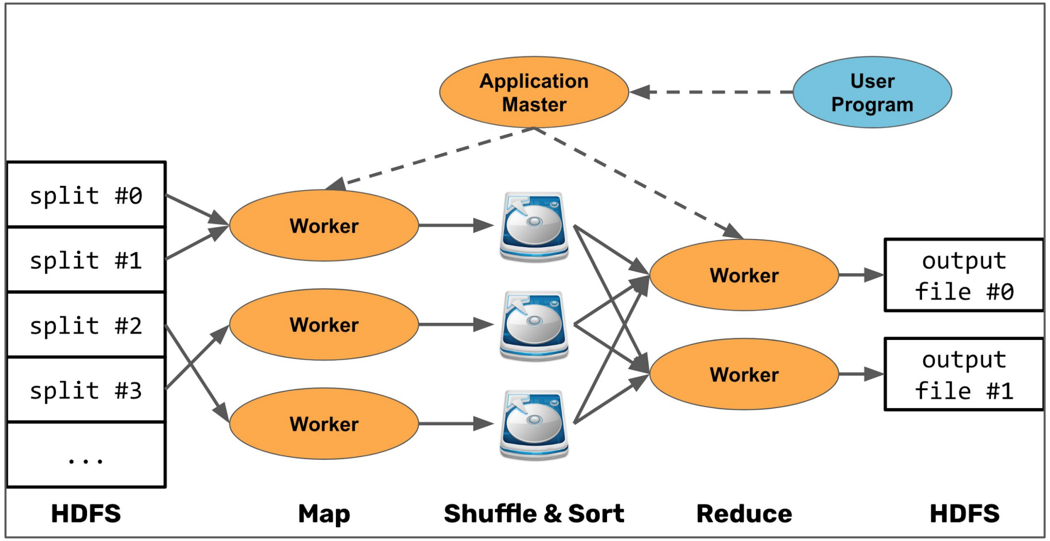
Fault Tolerance

BIGDATA TEAM

MapReduce



Где храним Shuffle & Sort?



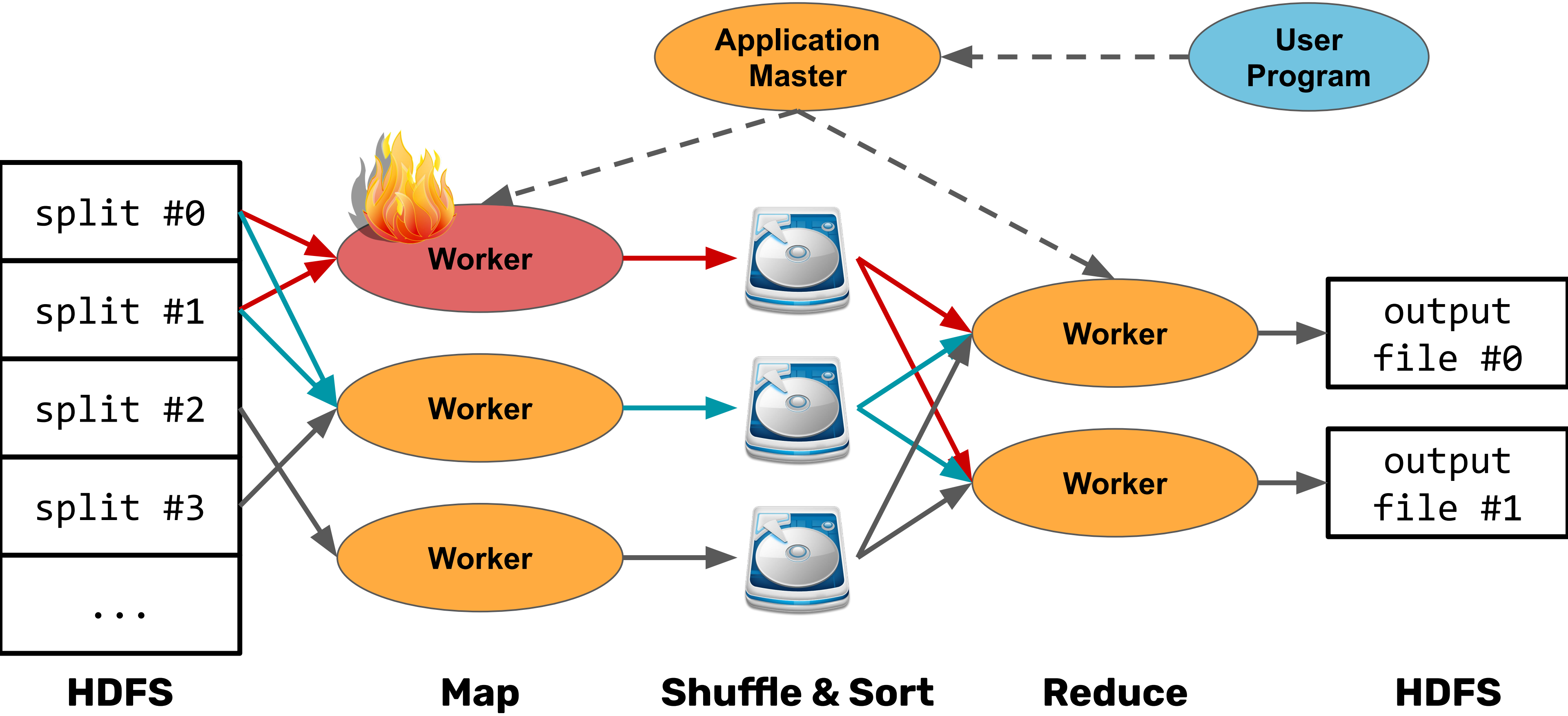
А. RAM

Б. HDFS

В. Local FS

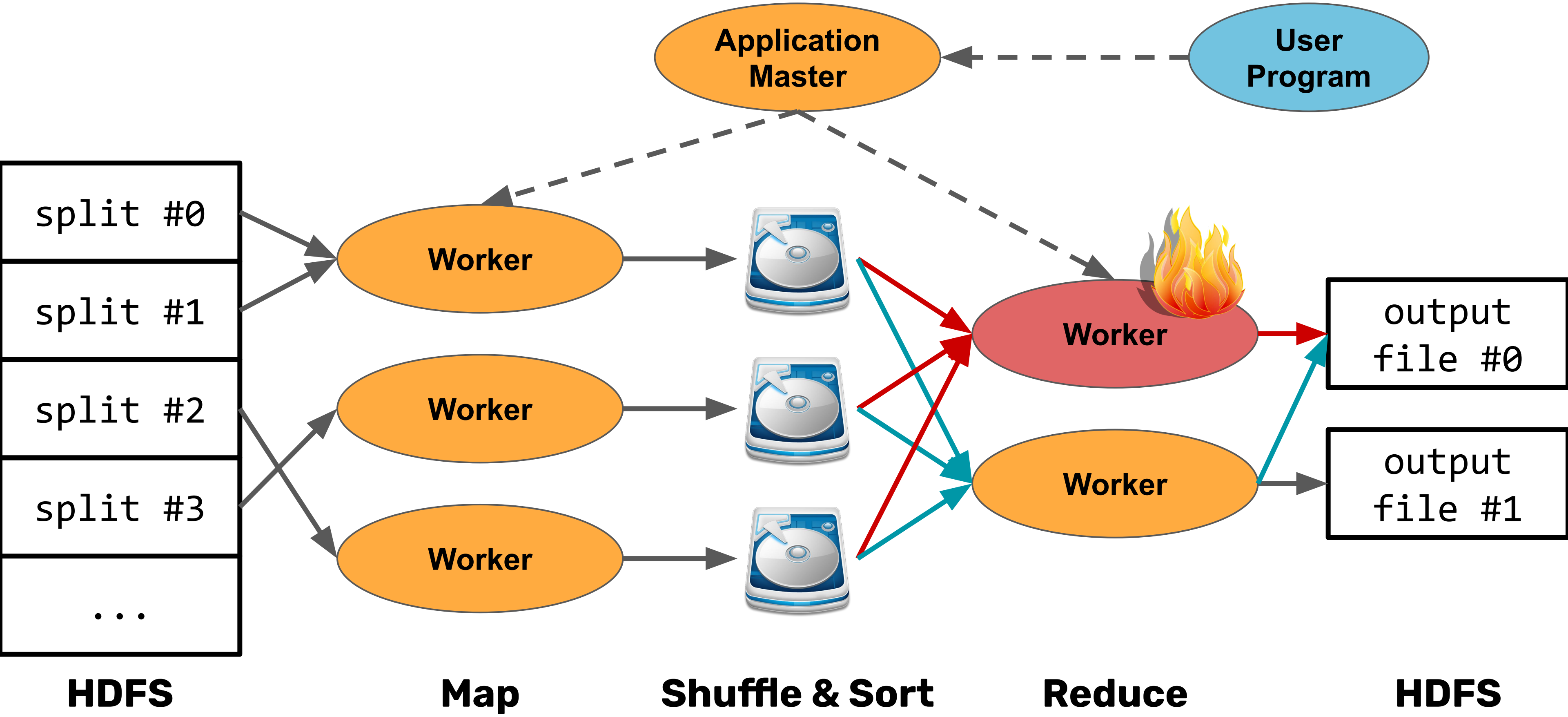
Г. Где-то там

Fault Tolerance (защита от сбоев)

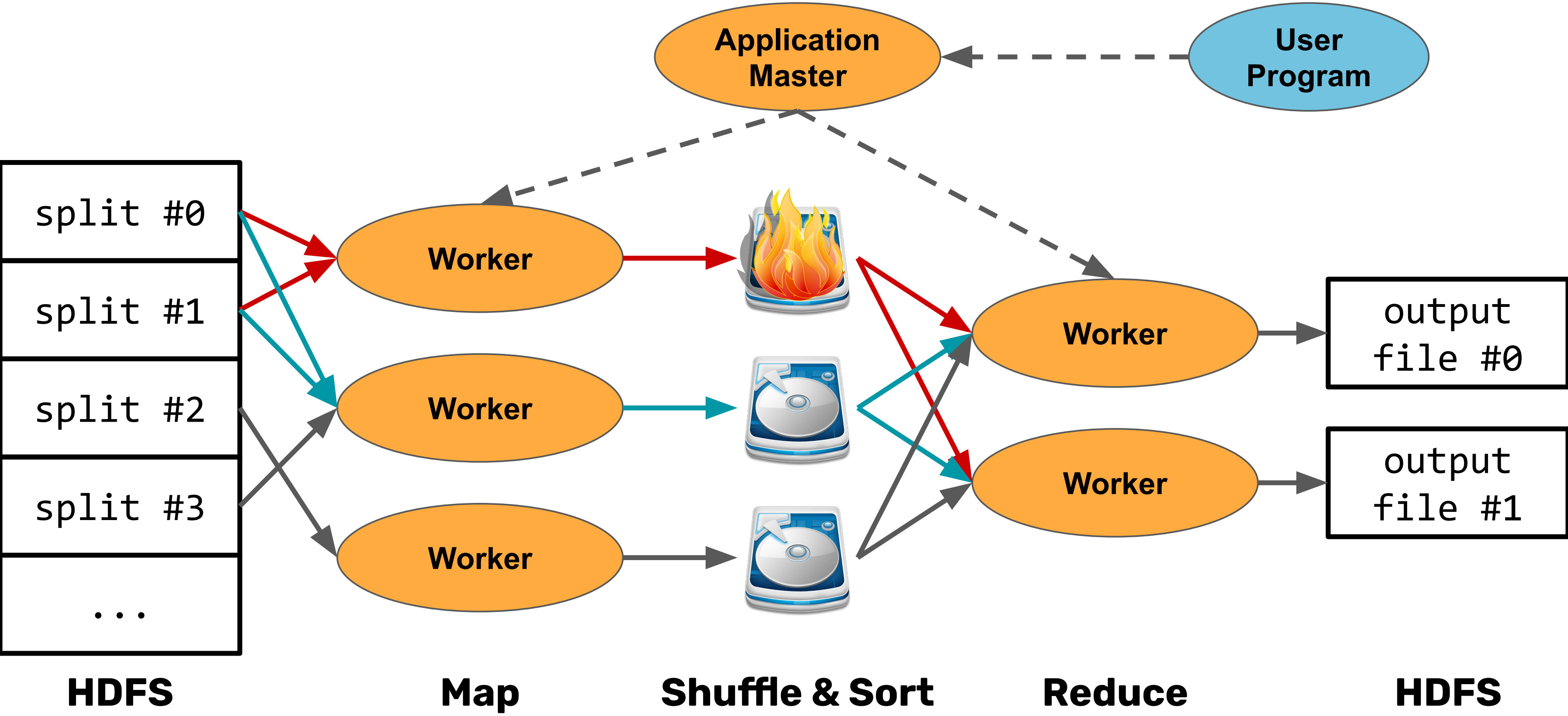


BIGDATA TEAM

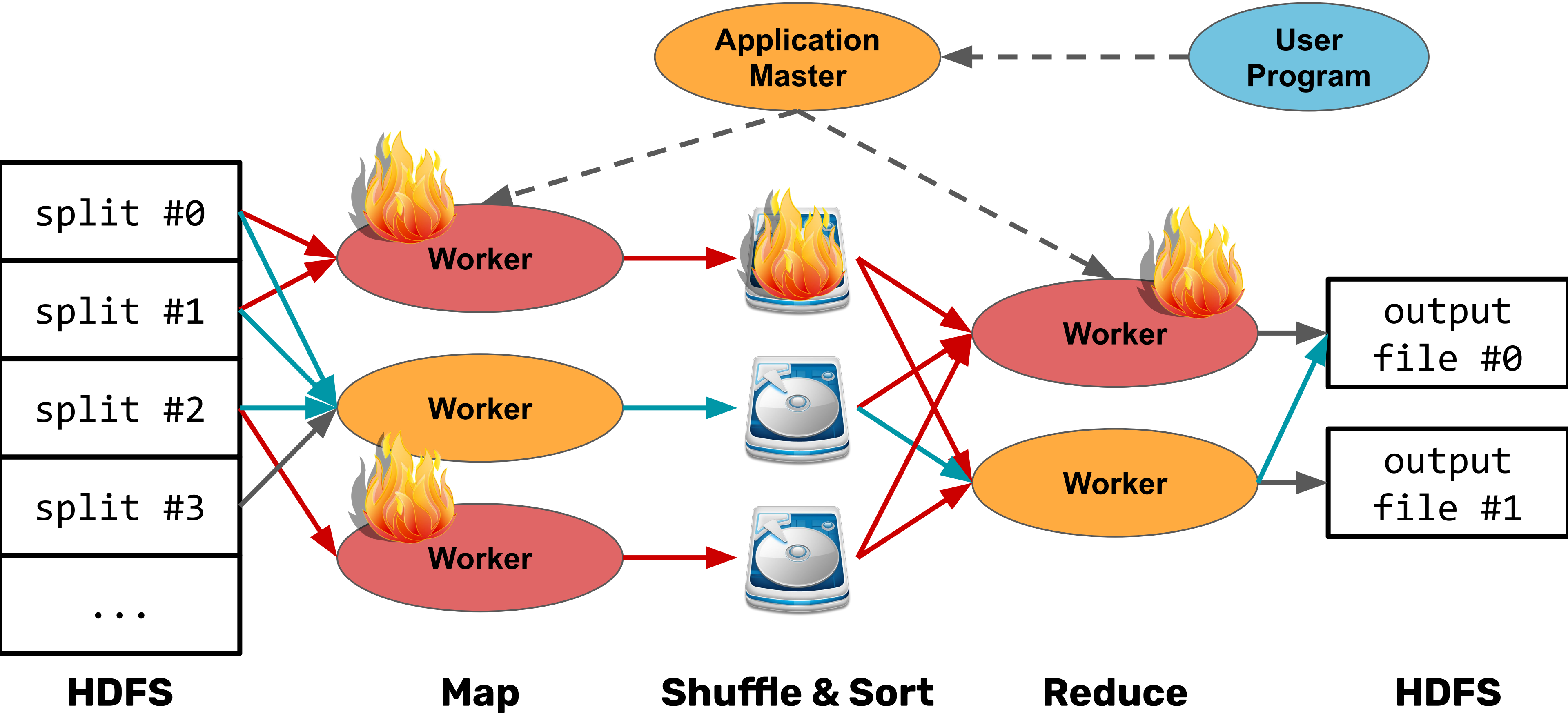
Fault Tolerance (защита от сбоев)

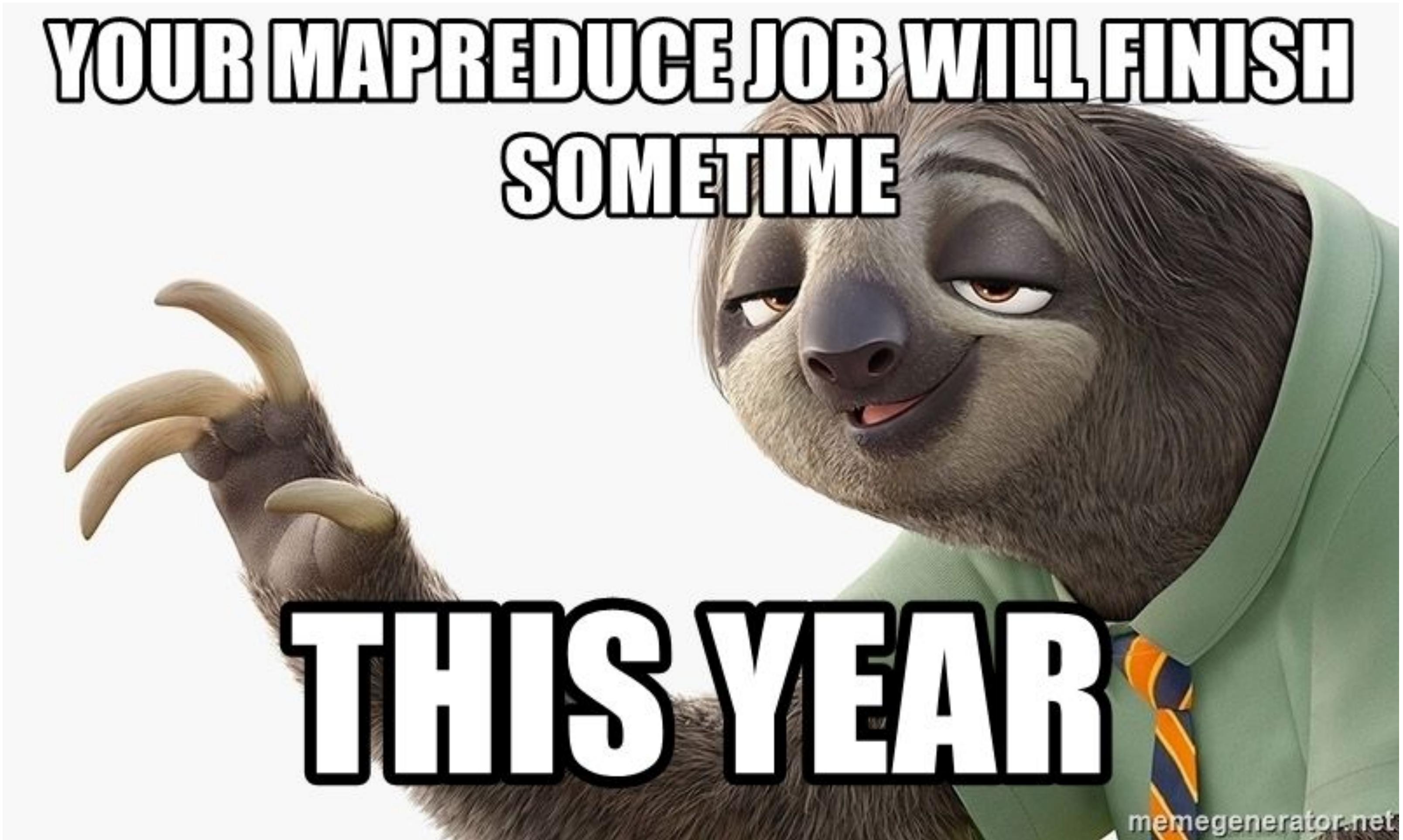


Fault Tolerance (защита от сбоев)

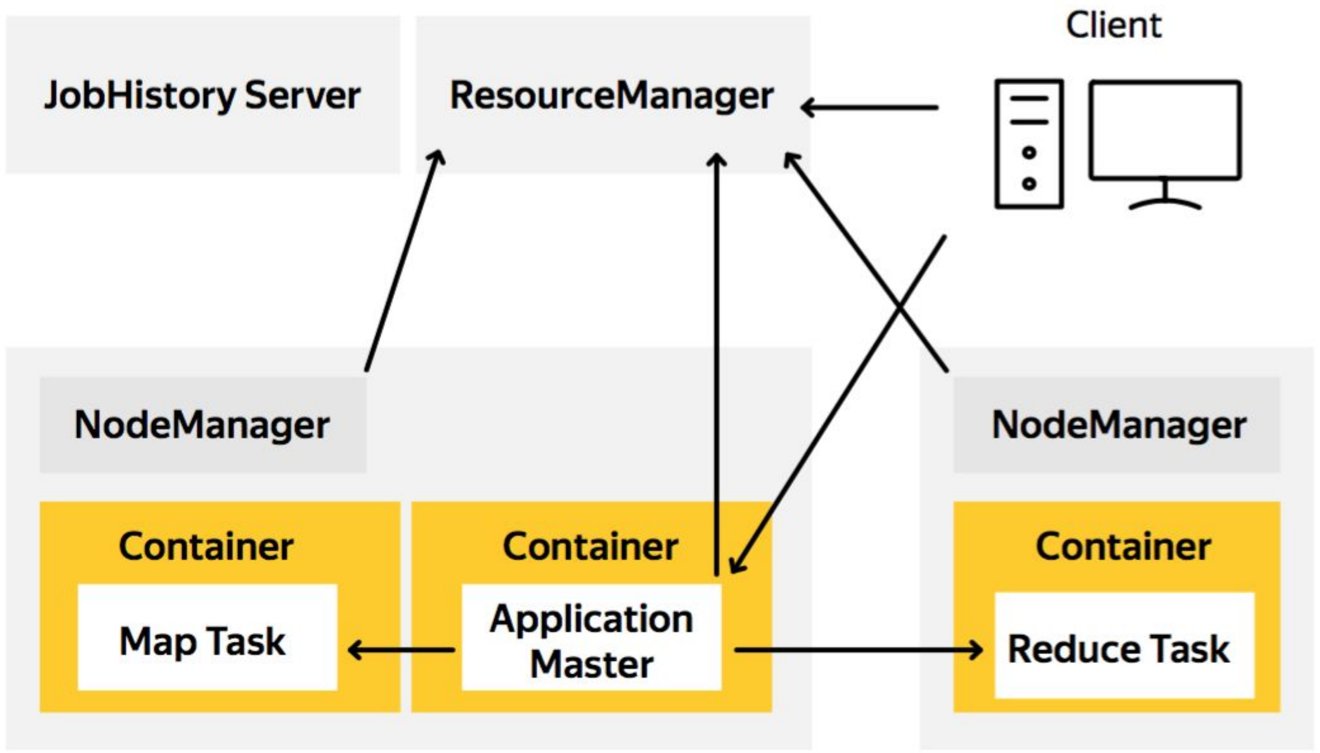


Fault Tolerance (защита от сбоев)




YOUR MAPREDUCE JOB WILL FINISH SOMETIME
THIS YEAR
memegenerator.net

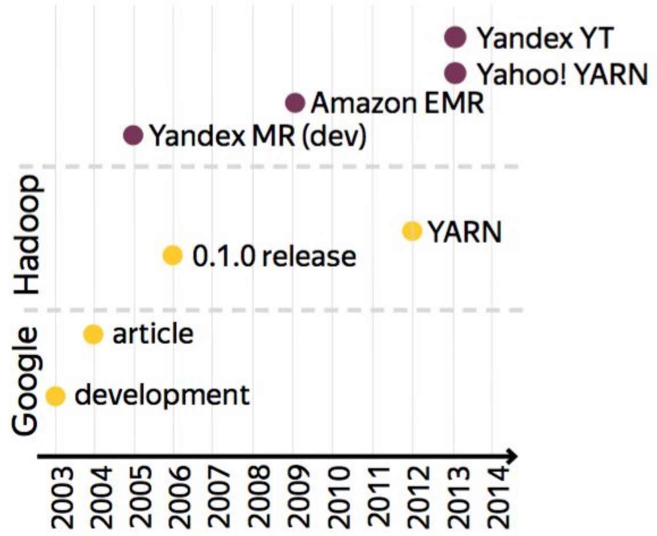
YARN



Client
JobHistory Server
ResourceManager
NodeManager
Container
Map Task
Container
Application Master
NodeManager
Container
Reduce Task

MapReduce Frameworks (Timeline)

- [2003] Google MapReduce (development)
- [2004] Google MapReduce (article)
- [2005] Yandex MapReduce (development)
- [2006] Hadoop 0.1.0 release
- [2009] Amazon EMR (Hadoop inside)
- [2012] MapReduce —> YARN
- [2013] Yahoo! YARN deployed in production
- [2013] Yandex YT…
- MapReduce in MongoDB, Riak, …



Q&A

Как зная топологию данных оптимизировать MapReduce?

BIGDATA TEAM

Data Locality

```
$ yarn jar map_reduce_example.jar
```

```
 ...
 INFO mapreduce.Job: Counters: 30
 ...
    Job Counters
        Launched map tasks=2
        Data-local map tasks=2
        Total time spent by all maps in occupied slots (ms)=27360
 ...
```

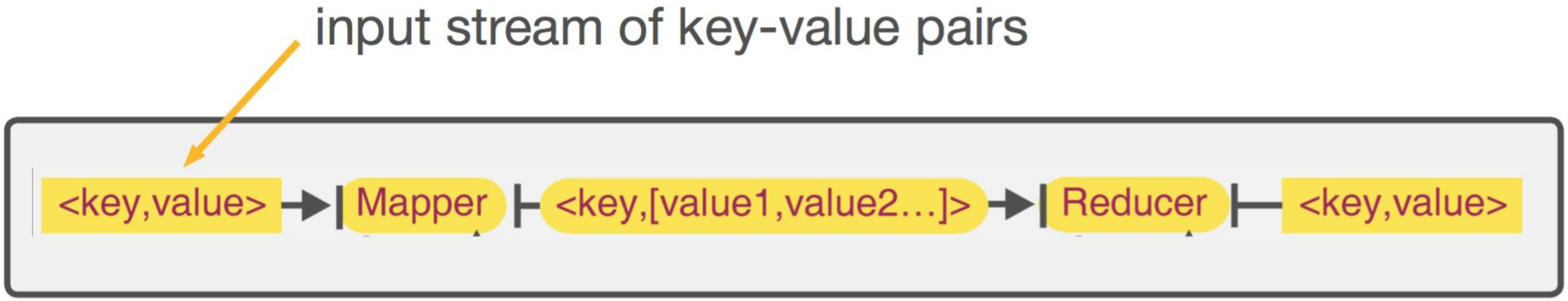
BIGDATA TEAM

Tea / Coffee Break

BIGDATA TEAM

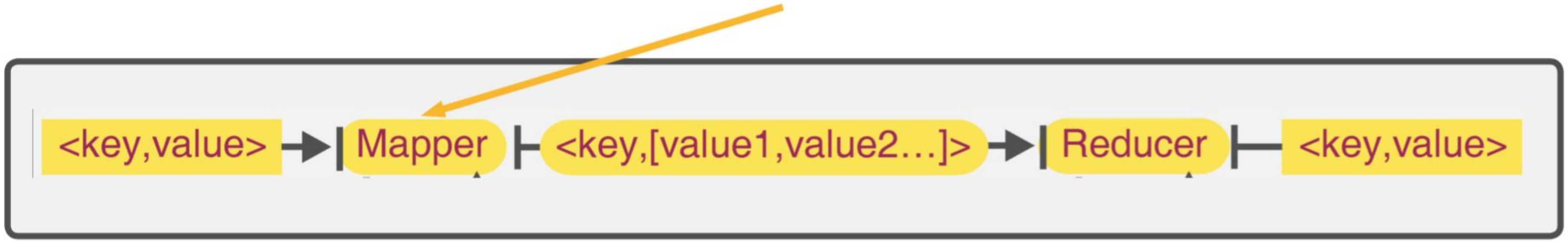
MapReduce Streaming

MapReduce Streaming

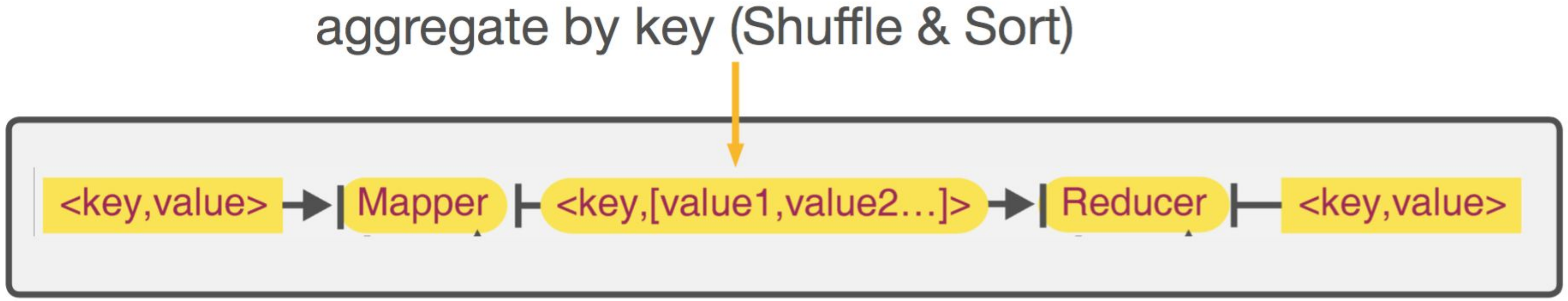


MapReduce Streaming

```
map: (k_in, v_in) --> [(k_interm, v_interm), …]
```

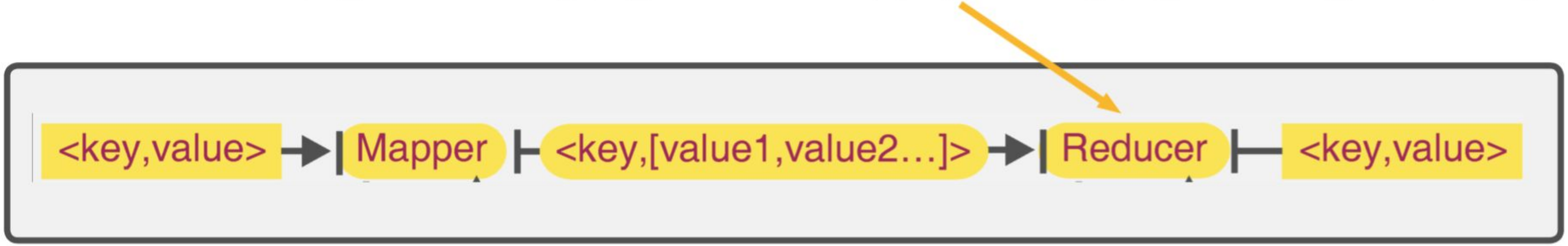


MapReduce Streaming

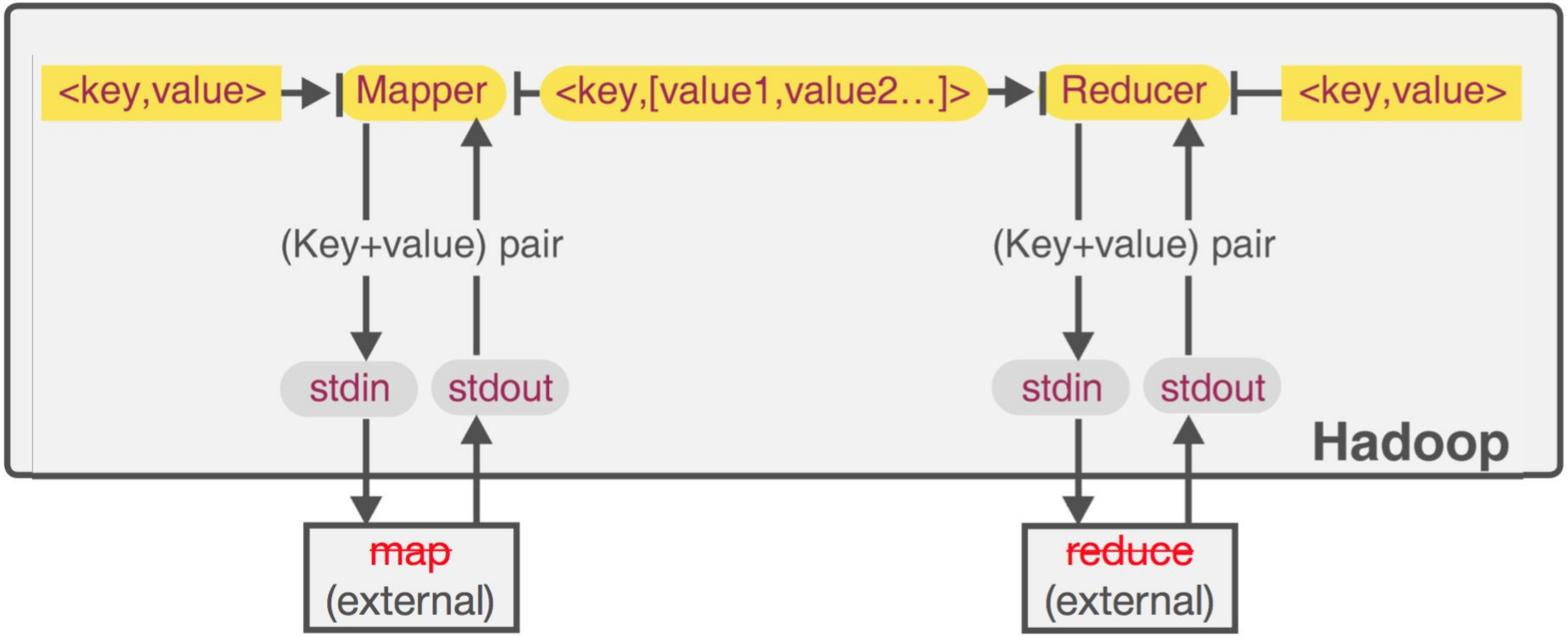


MapReduce Streaming

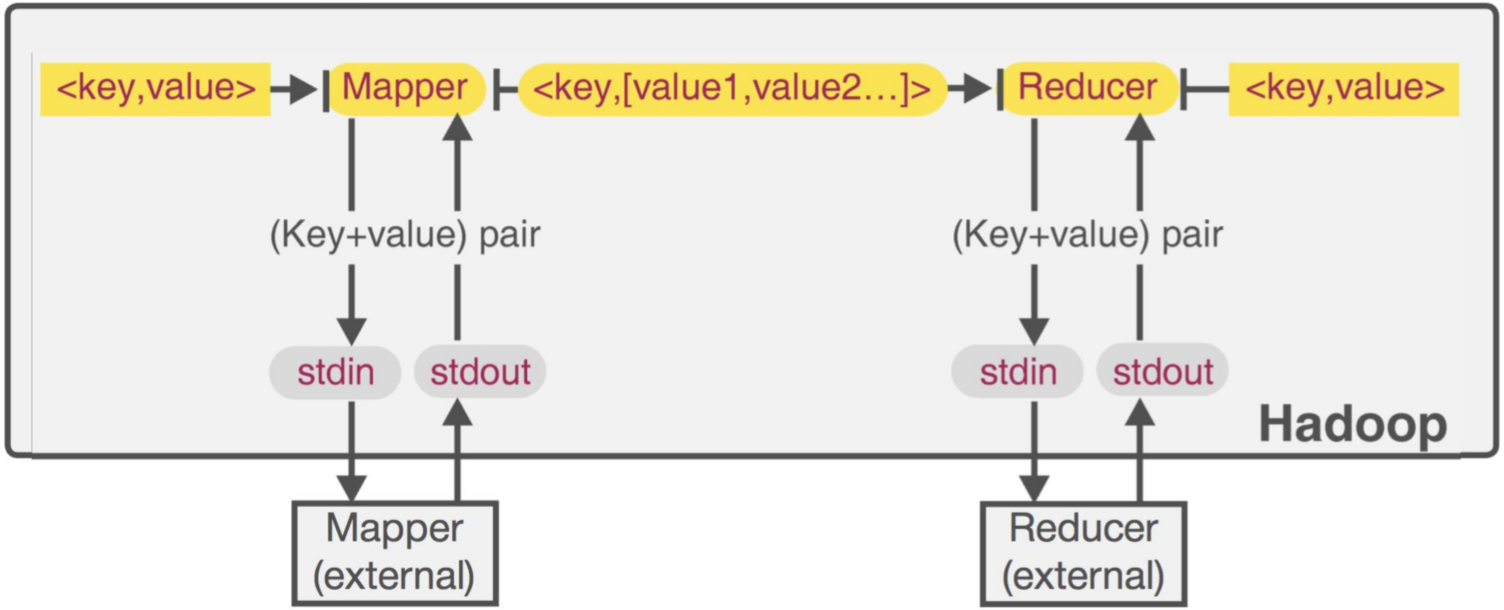
```
reduce: (k_interm, [(v_interm, …)] ) --> [(k_out, v_out), …]
```



MapReduce Streaming



MapReduce Streaming



Mapper:

- ▶ Как данные читаем (input format)
- ▶ Как данные обрабатываем
- ▶ Как данные выводим (output format)

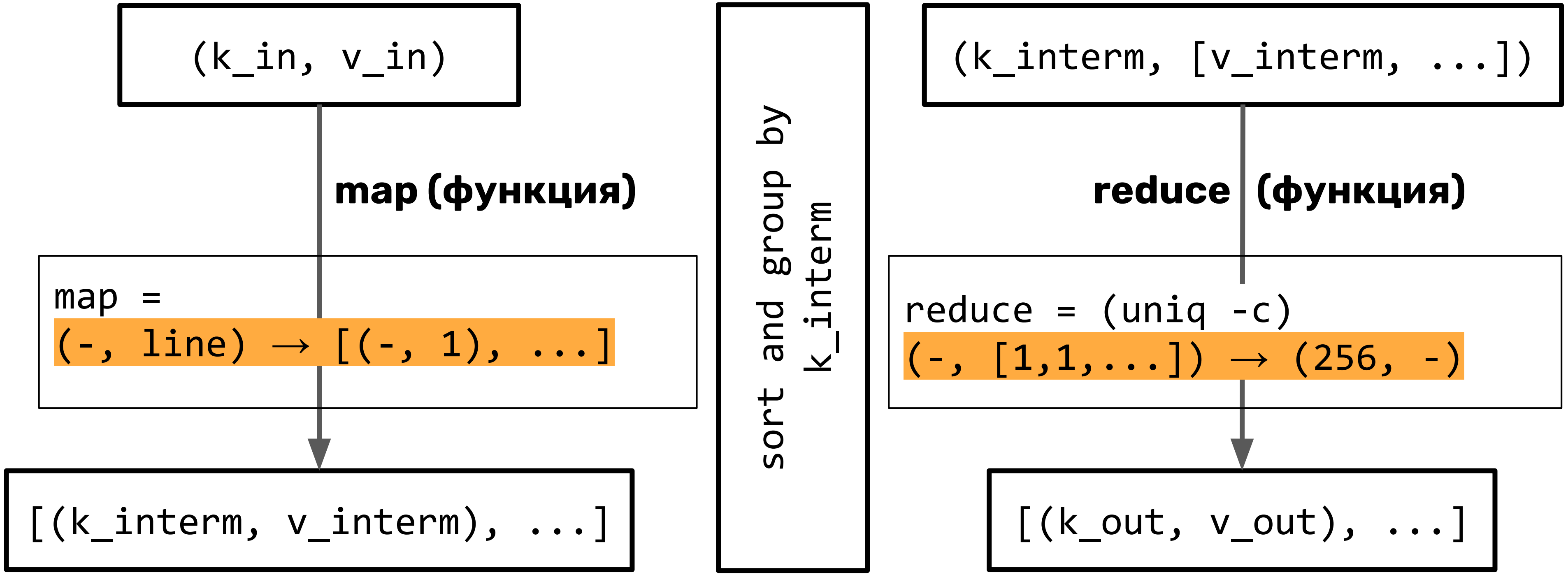
Тоже, что и Mapper, плюс:

- ▶ Как агрегируем по ключам отсортированные данные

Постановка задачи Line Count

`<article_id> <tab> <article_content>`

Line Count



Map (фаза) **Shuffle & Sort** **Reduce (фаза)**

run.sh

```
HADOOP_STREAMING_JAR=/path/to/hadoop-streaming.jar
OUT_DIR=my_hdfs_output

yarn jar $HADOOP_STREAMING_JAR \
    -mapper "wc -l" \
    -numReduceTasks 0 \
    -input /data/wiki/en_articles_part \
    -output $OUT_DIR
```

```
$ hdfs dfs -ls my_hdfs_output
```

```
Found 3 items
-rw-r--r--   3 aadral hdfs    0 2021-02-15 18:38 my_hdfs_output/_SUCCESS
-rw-r--r--   3 aadral hdfs    6 2021-02-15 18:38 my_hdfs_output/part-00000
-rw-r--r--   3 aadral hdfs    5 2021-02-15 18:38 my_hdfs_output/part-00001
```

```
$ hdfs dfs -text my_hdfs_output/*
```

```
3624
476
```

```
$ hdfs dfs -text my_hdfs_output/*
```

```
3624
476
```

```
$ hdfs dfs -ls -h /data/wiki/en_articles_part
```

```
Found 1 items
-rw-r--r--   3 hdfs hdfs      73.3 M 2020-03-12 21:03
/data/wiki/en_articles_part/articles-part
```

```
$ ./run.sh
```

```
...
ERROR streaming.StreamJob: Error Launching job : Output directory
hdfs://brain-master.bigdatateam.org:8020/user/aadral/my_hdfs_outpu
t already exists
Streaming Command Failed!
```

run.sh

```
HADOOP_STREAMING_JAR=/path/to/hadoop-streaming.jar
OUT_DIR=my_hdfs_output

hdfs dfs -rm -r $OUT_DIR

yarn jar $HADOOP_STREAMING_JAR \
    -mapper "wc -l" \

    -numReduceTasks 0 \
    -input /data/wiki/en_articles_part \
    -output $OUT_DIR
```

BIGDATA TEAM

run.sh

```
HADOOP_STREAMING_JAR=/path/to/hadoop-streaming.jar
OUT_DIR=my_hdfs_output

hdfs dfs -rm -r $OUT_DIR

yarn jar $HADOOP_STREAMING_JAR \
    -mapper "wc -l" \
    -reducer "awk '{line_count += \$1} END { print line_count }'" \
    -numReduceTasks 1 \
    -input /data/wiki/en_articles_part \
    -output $OUT_DIR
```

```
$ hdfs dfs -ls my_hdfs_output
```

```
Found 2 items
-rw-r--r--   3 aadral hdfs    0 2021-02-17 11:22 my_hdfs_output/_SUCCESS
-rw-r--r--   3 aadral hdfs    6 2021-02-17 11:22 my_hdfs_output/part-00000
```

```
$ hdfs dfs -text my_hdfs_output/*
```

```
4100
```

BIGDATA TEAM

reducer.sh

```
#!/usr/bin/env bash
awk '{line_count += $1} END { print line_count }'
```

BIGDATA TEAM

run.sh

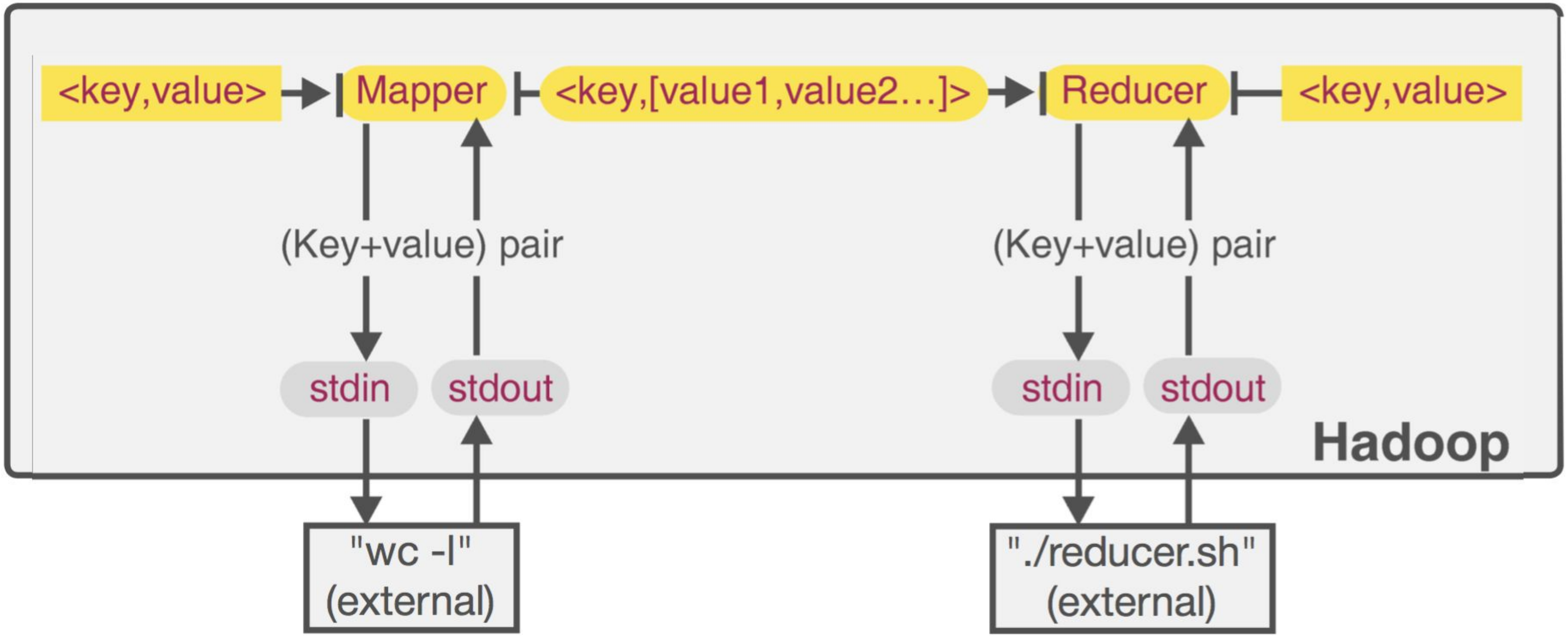
```
HADOOP_STREAMING_JAR=/path/to/hadoop-streaming.jar
OUT_DIR=my_hdfs_output

hdfs dfs -rm -r $OUT_DIR

yarn jar $HADOOP_STREAMING_JAR \
    -files reducer.sh \
    -mapper "wc -l" \
    -reducer "./reducer.sh" \
    -numReduceTasks 1 \
    -input /data/wiki/en_articles_part \
    -output $OUT_DIR
```

BIGDATA TEAM

Line Count



Резюме

- Вы можете объяснить, что происходит когда “умирает” Mapper или Reducer
- Вы знаете, за что отвечают ResourceManager и NodeManager в YARN
- Вы знаете 3 фазы MapReduce (Map, Shuffle & Sort, Reduce)
- Вы знаете, что такое MapReduce Streaming и как он работает (пример: Line Count)

KEEP
CALM
AND
TRY
CODING

BIGDATA TEAM

Спасибо! Вопросы?

Feedback: http://rebrand.ly/x5bd2021q1_feedback_02_mr

Алексей Драль, aadral@bigdatateam.org
CEO at BigData Team, http://bigdatateam.org
https://www.linkedin.com/in/alexey-dral
https://www.facebook.com/bigdatateam