НИЯУ МИФИ
Национальный исследовательский ядерный университет МИФИ
Кафедра 75 «Кафедра финансового мониторинга»

Специальные технологии баз данных и информационных систем

Лекция 2. Введение в Hadoop

Авторы:

Мелконян С.Е.
Айрапетян С.В.

История развития Hadoop

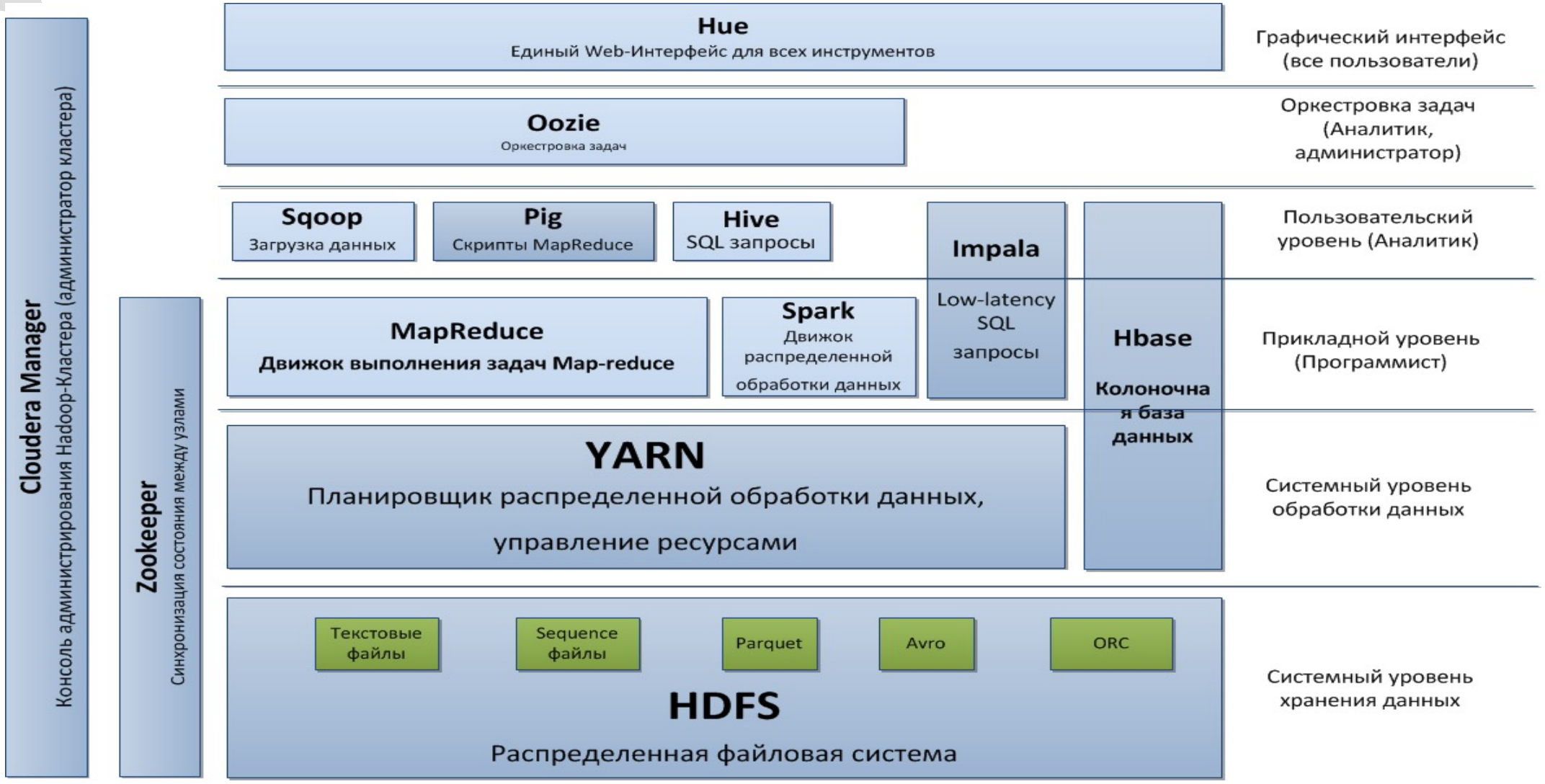
2004 – компания Google опубликовало две статьи, в которых описывалась файловая система Google File System (GFS) и концепция MapReduce.

2006 – корпорация Yahoo пригласила Каттинга возглавить специально выделенную команду разработки инфраструктуры распределённых вычислений.

2008 – Yahoo запустила кластерную поисковую машину на 10 тысяч процессорных ядер под управлением Hadoop.

2010 – корпорация Google предоставила Apache Software Foundation права на использование технологии MapReduce.

Экосистема Hadoop

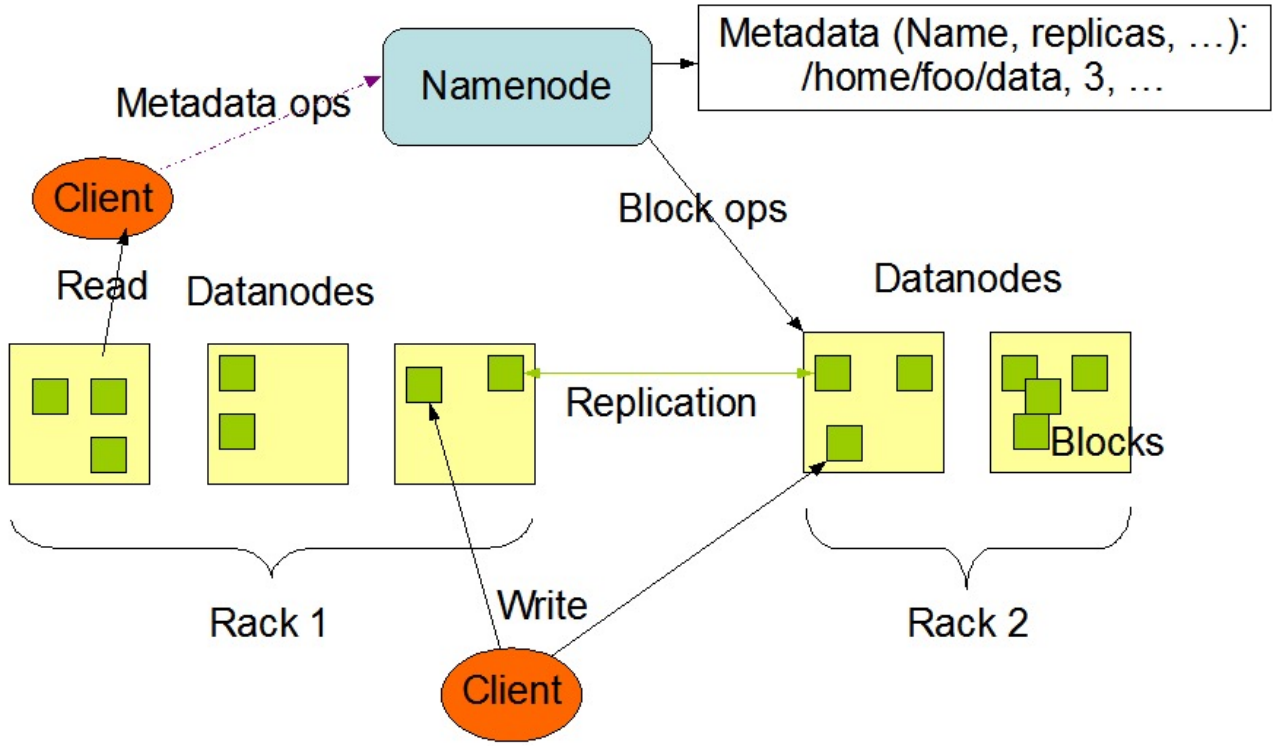


Особенности Hadoop

- ✓ Горизонтальное масштабирование
- ✓ Пары ключ/значение вместо реляционных таблиц
- ✓ Функциональное программирование (MapReduce) вместо декларативных запросов (SQL)
- ✓ Автономная пакетная обработка вместо оперативных транзакций

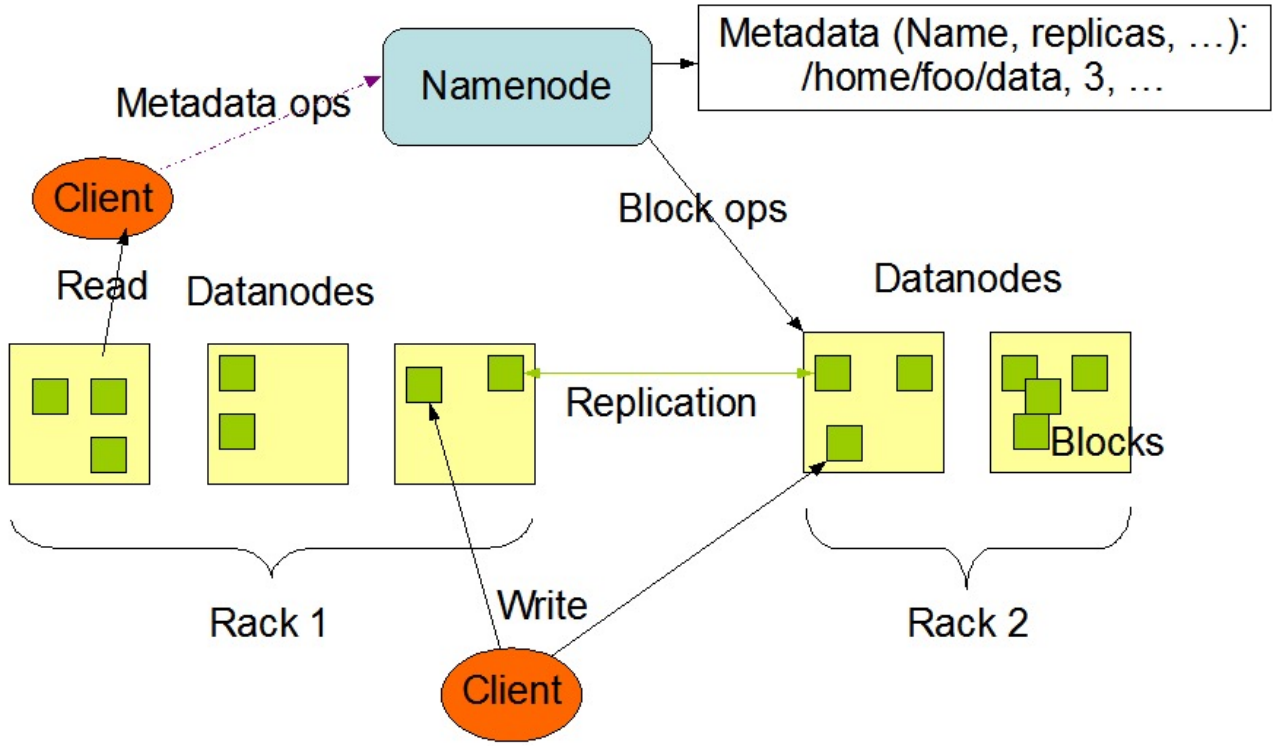
Архитектура HDFS

- **NameNode** — управляющий узел, узел имен
- **DataNode** — узел или сервер данных
- **Client** — клиент
- **Secondary NameNode** — вторичный узел имен



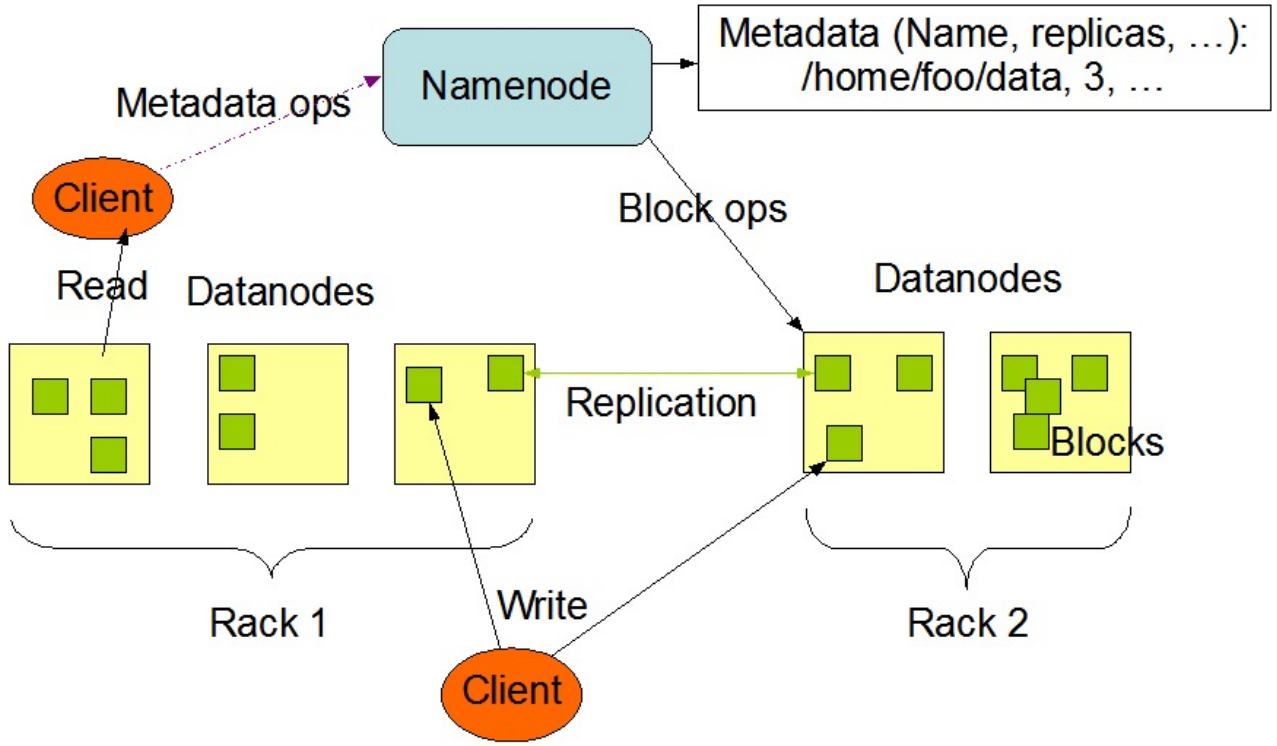
Архитектура HDFS: NameNode

- ведет учет разбиению файлов на блоки, хранит информацию о том, на каких узлах эти блоки находятся, и следит за общим состоянием распределенной файловой системы
- отвечает за открытие и закрытие файлов, создание и удаление каталогов, управление доступом со стороны внешних клиентов и соответствие между файлами и блоками, дублированными (реплицированными) на узлах данных

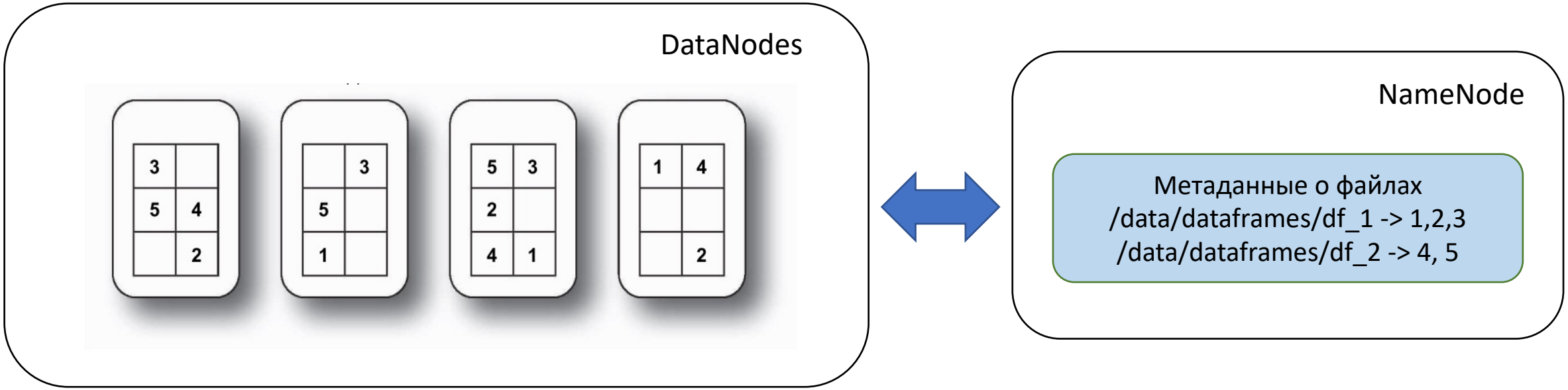


Архитектура HDFS: DataNode

- ✓ отвечает за запись и чтение данных
- ✓ данные хранятся в блоках, объем которых можно задавать при настройке кластера
- ✓ выполняет команды от узла NameNode по созданию, удалению и репликации блоков
- ✓ периодически отправляет сообщения о состоянии (heartbeats)
- ✓ обрабатывает запросы на чтение и запись, поступающих от клиентов файловой системы HDFS.

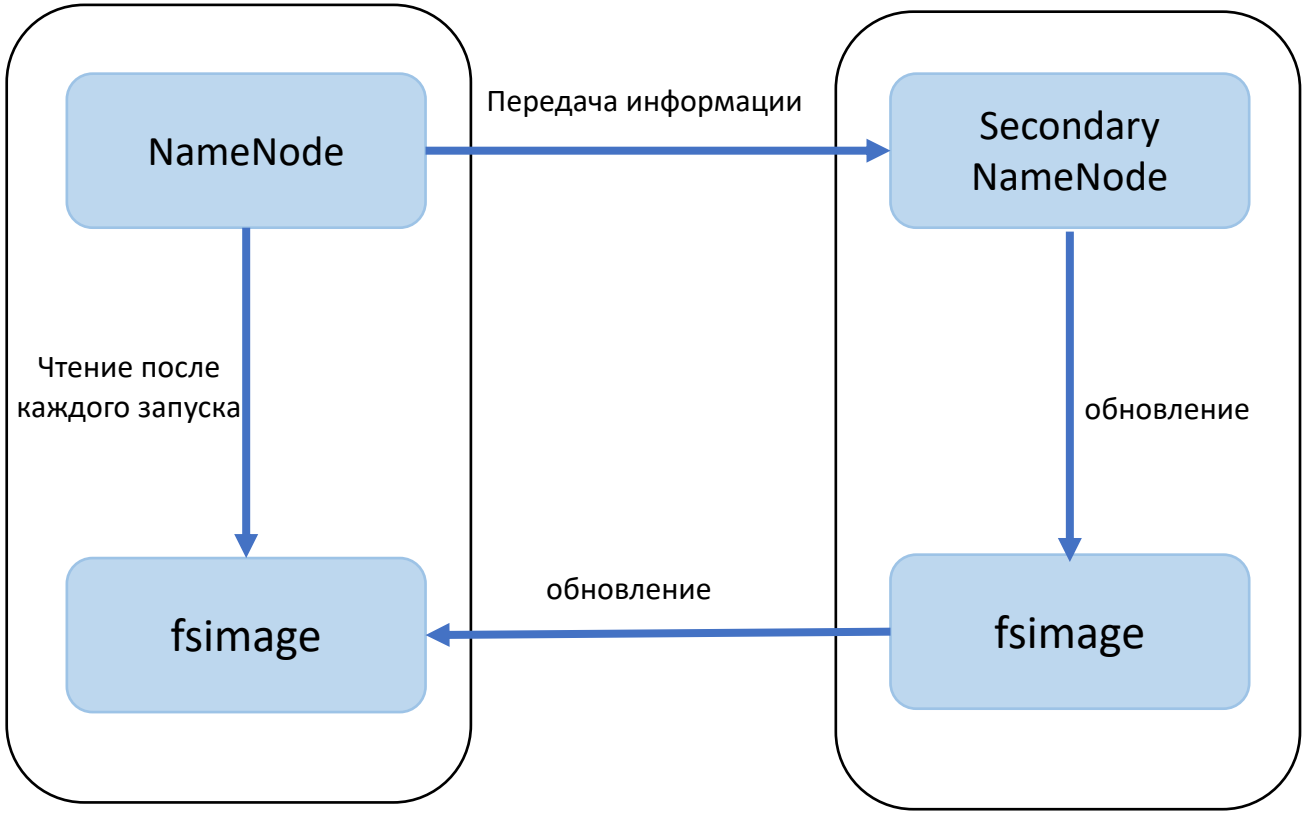


Пример хранения файлов в HDFS



Архитектура HDFS: Secondary NameNode

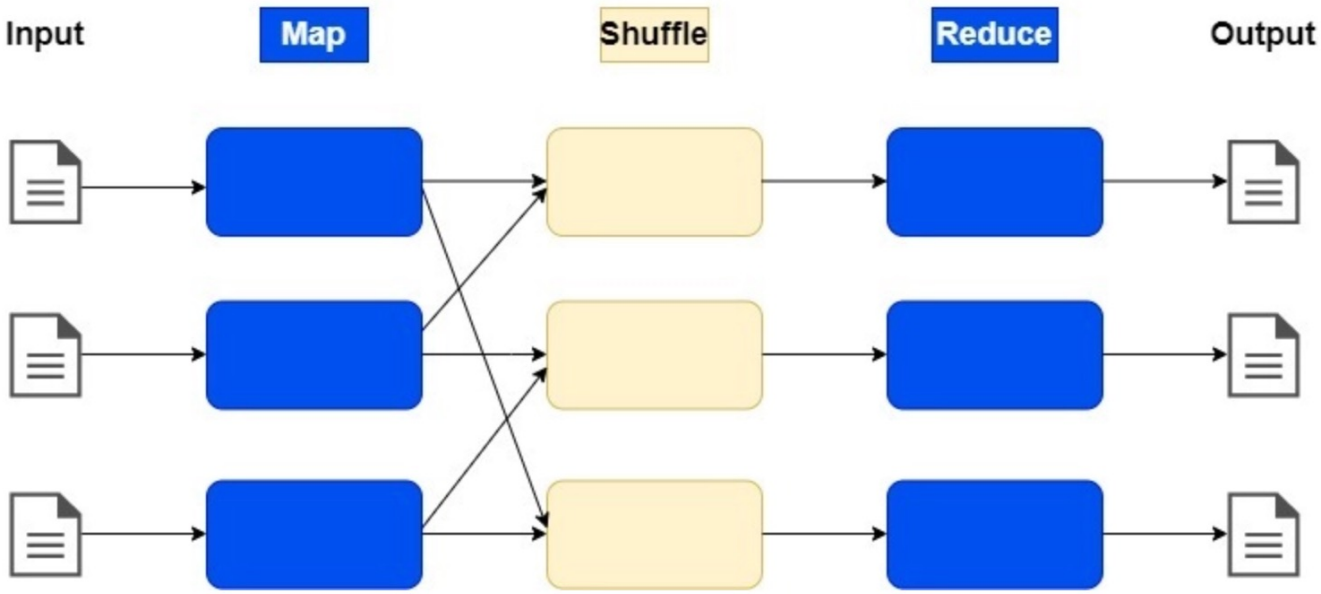
Secondary NameNode — вторичный узел имен, отдельный сервер, единственный в кластере, который копирует образ HDFS и лог транзакций операций с файловыми блоками во временную папку, применяет изменения, накопленные в логе транзакций к образу HDFS, а также записывает его на узел NameNode и очищает лог транзакций. Secondary NameNode необходим для быстрого ручного восстановление NameNode в случае его выхода из строя.



Знакомство с MapReduce

MapReduce - модель распределённых вычислений для параллельной обработки больших объёмов информации.

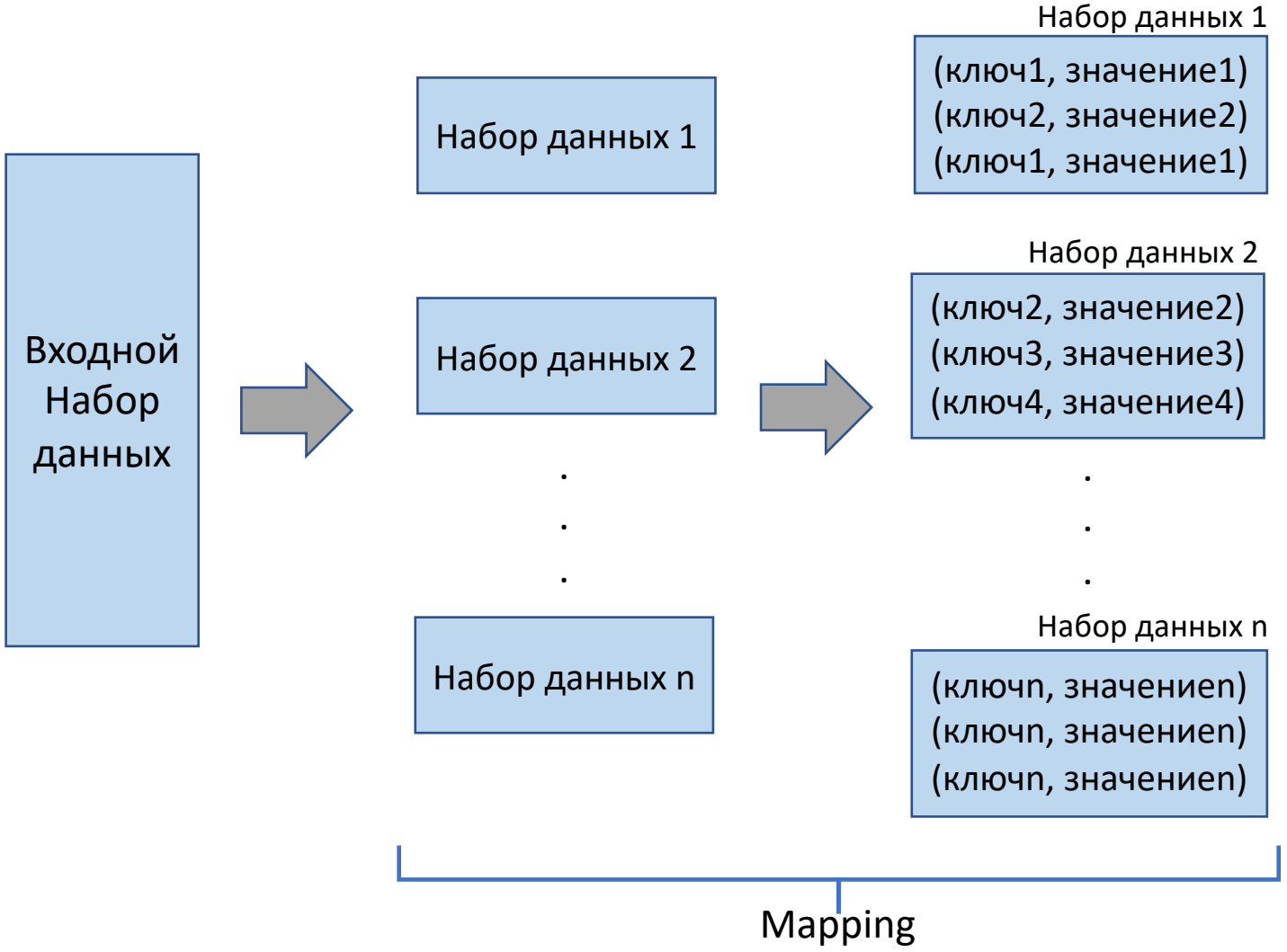
- ✓ **Map** – распределение
- ✓ **Shuffle** – перераспределение
- ✓ **Reduce** – редукция



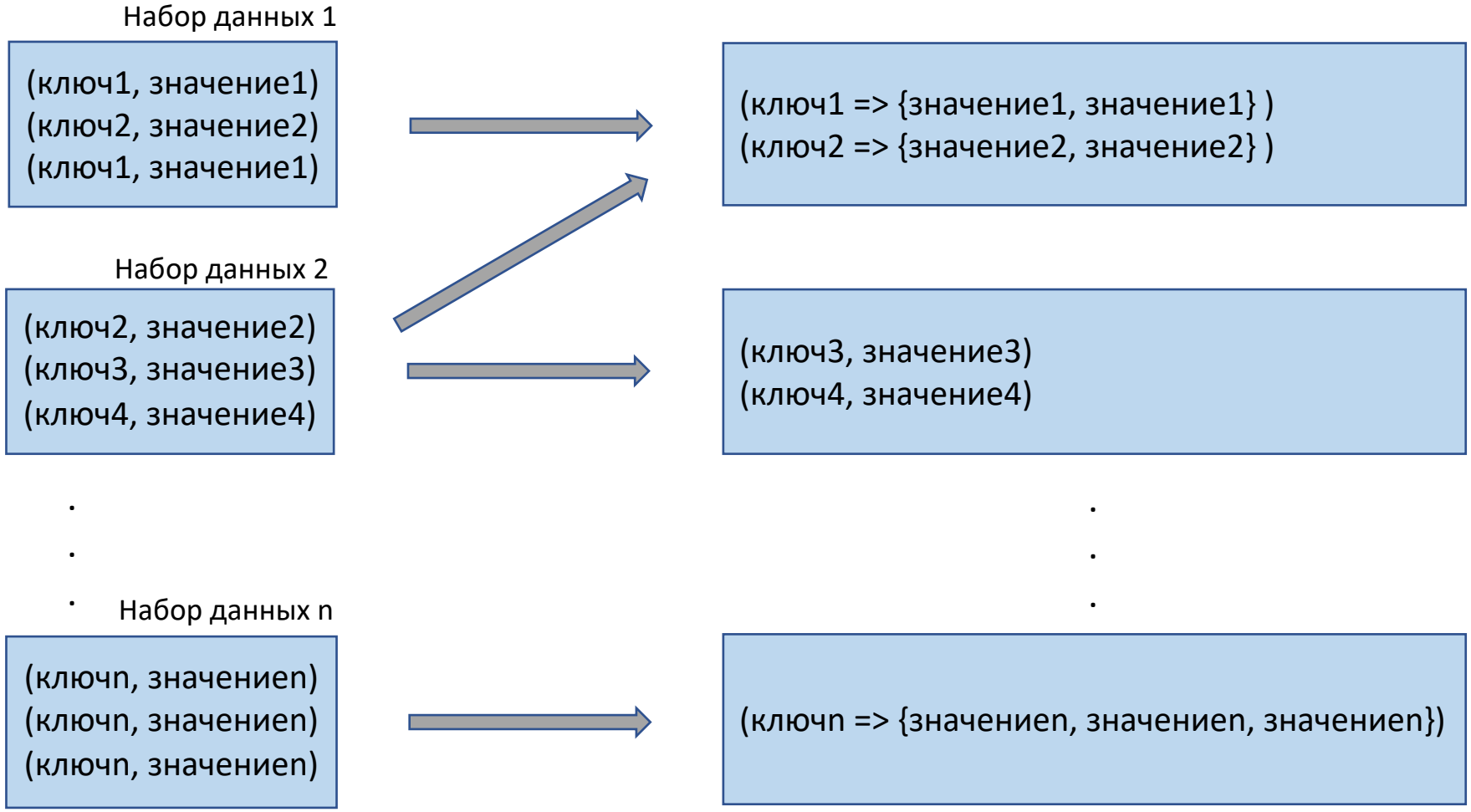
MapReduce: шаг Map

На шаге Map происходит предварительная обработка данных - главный узел кластера (master node) получает набор данных, делит его на части и передает рабочим узлам (worker node).

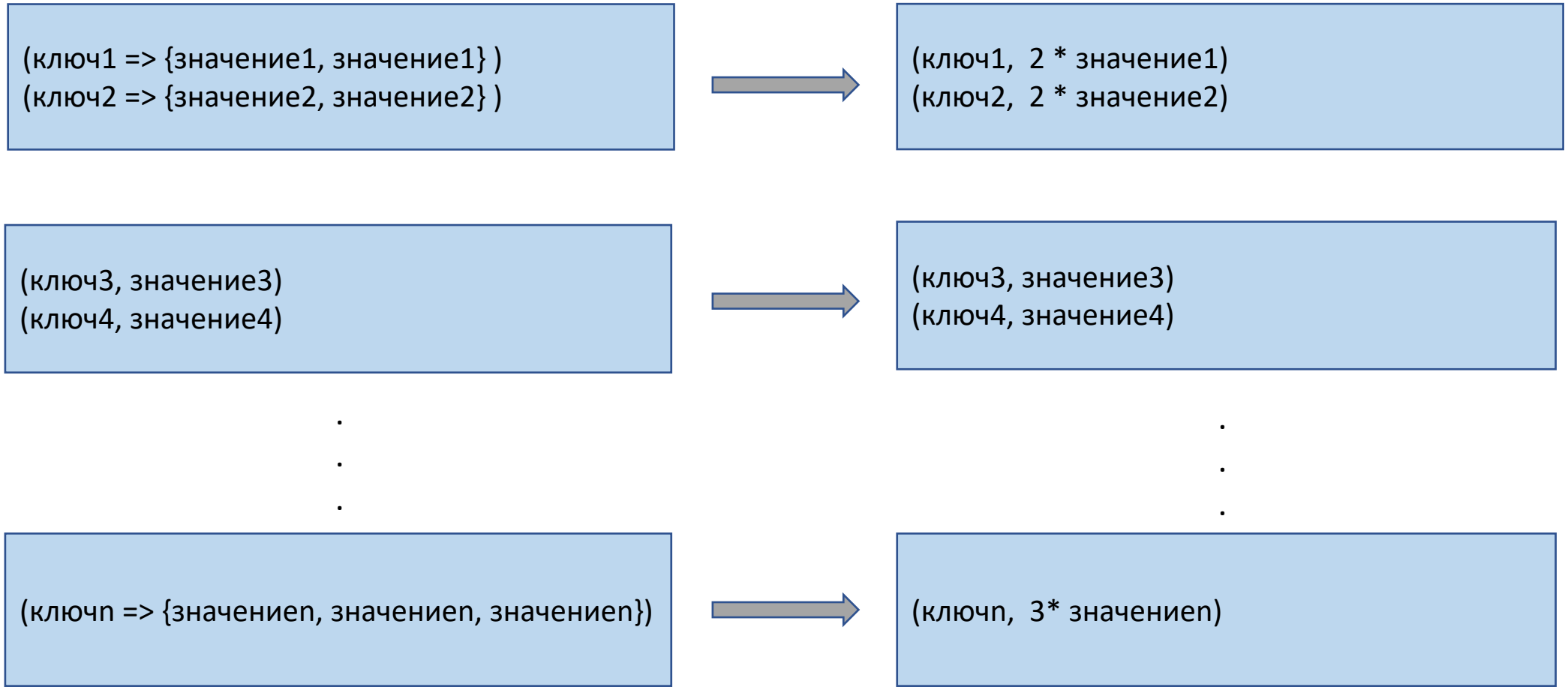
Каждый рабочий узел применяет функцию Map к локальным данным и записывает результат в формате «ключ-значение» во временное хранилище.



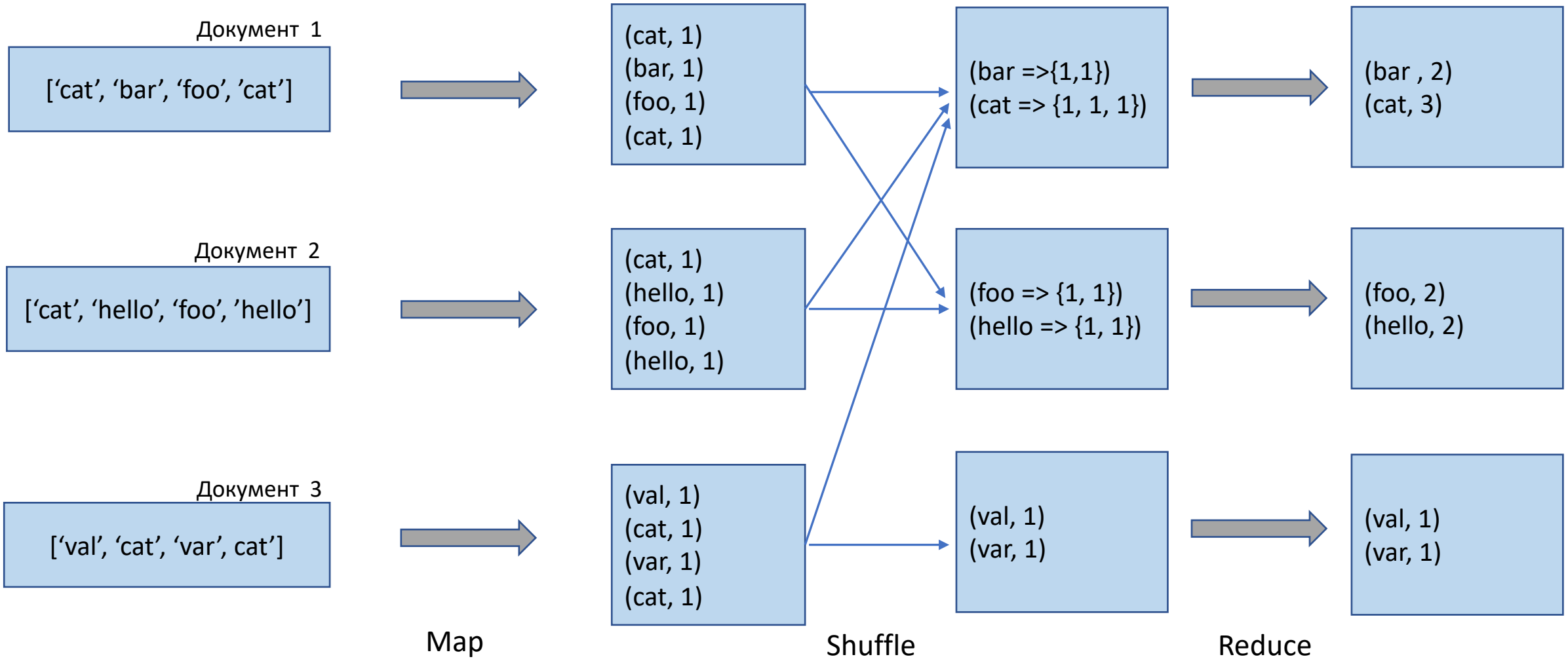
MapReduce: шаг Shuffle



MapReduce: шаг Reduce



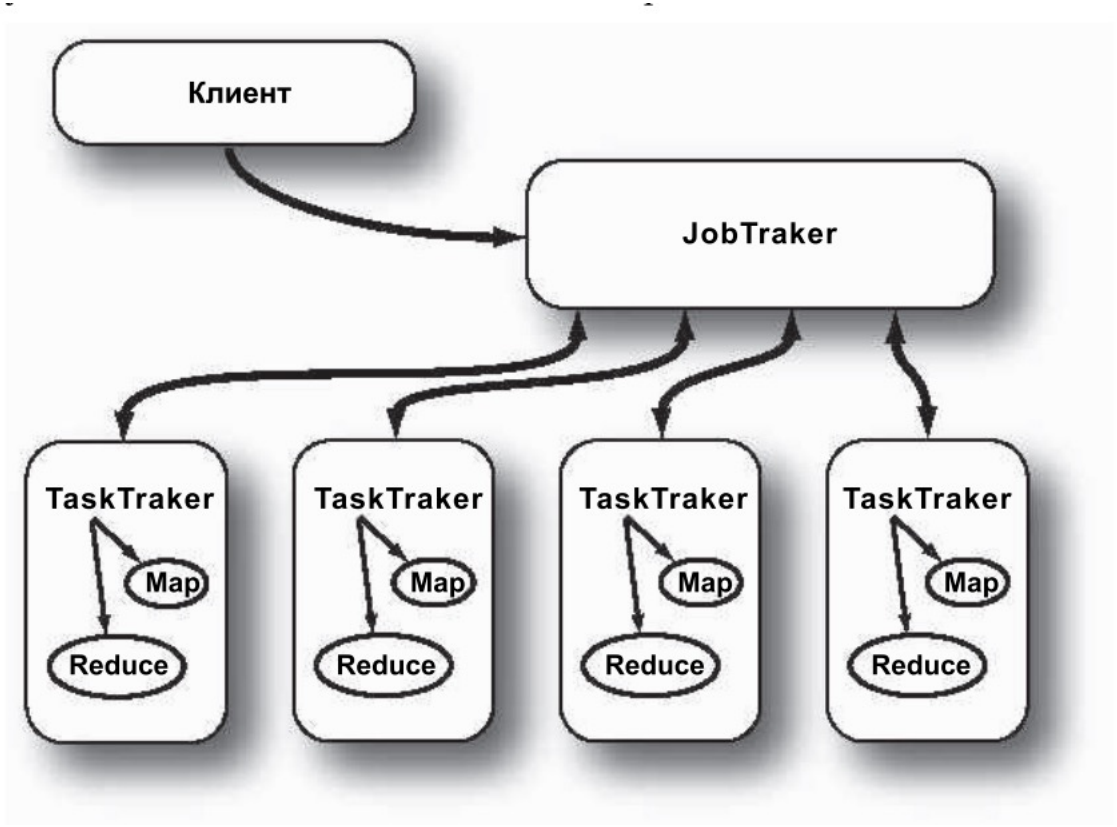
Пример использования



Реализация в Hadoop

JobTraker - строит план выполнения, то есть определяет, какие файлы обрабатывать, назначает узлы различным задачам и следит за ходом исполнения этих задач. В кластере Hadoop может быть только один демон JobTracker.

TaskTraker - управляют исполнением отдельных заданий на подчиненных узлах.



Преимущества применения MapReduce

возможность распределенного выполнения операций предварительной обработки (map) и свертки (reduce) большого объема данных. При этом функции map работают независимо друг от друга и могут выполняться параллельно на разных узлах кластера.

быстрота обработки больших объёмов данных за счет распределения операций (сортировка петабайта данных при использовании MapReduce за пару часов)

отказоустойчивость и оперативное восстановления после сбоев: при отказе рабочего узла, производящего операцию map или reduce, его работа автоматически передается другому рабочему узлу в случае доступности входных данных для проводимой операции.

Недостатки MapReduce

недостаточно высокая производительность – классическая технология, в частности, реализованная в ядре Apache Hadoop, обрабатывает данные ациклично в пакетном режиме. При этом функции Reduce не запустятся до завершения всех процессов Map. Все операции проходят по циклу чтение-запись с жесткого диска, что влечет задержки (latency) в обработке информации.

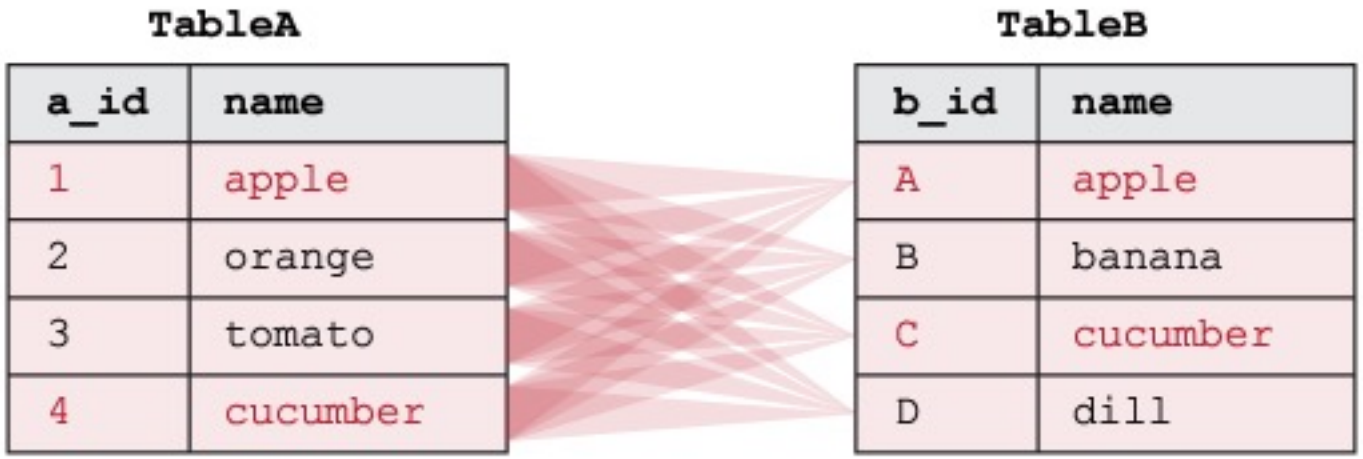
ограниченность применения – высокие задержки распределенных вычислений, приемлемые в пакетном режиме обработки, не позволяют использовать классический MapReduce для потоковой обработки в режиме реального времени, повторяющихся запросов и итеративных алгоритмов на одном и том же датасете, как в задачах машинного обучения.

Часть 2. Основы SQL

Соединение (JOIN) - операция для сопоставления строки одной таблицы строкам другой таблицы.

Синтаксис

```
SELECT * FROM
TableA a  INNER join TableB b
ON
a. name=b.name;
```



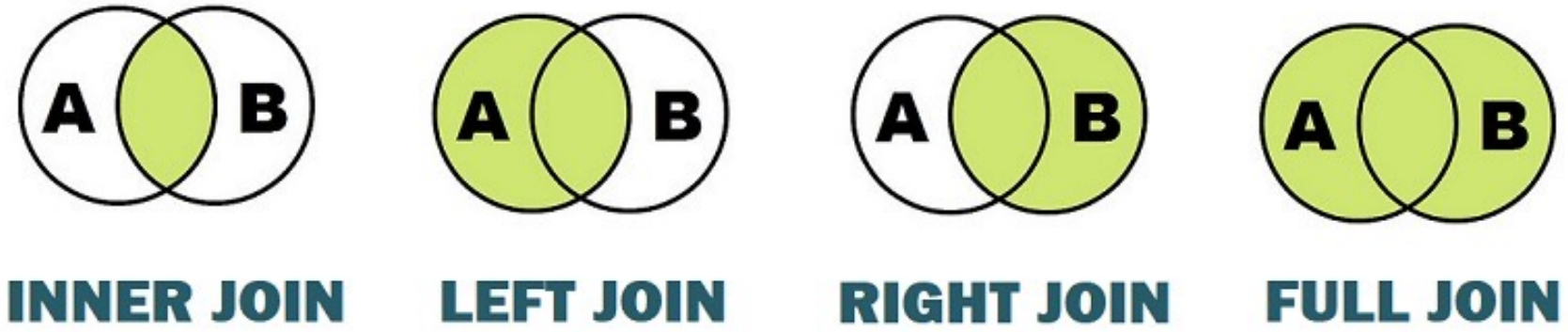
TableA

a_id	name
1	apple
2	orange
3	tomato
4	cucumber

TableB

b_id	name
A	apple
B	banana
C	cucumber
D	dill

Виды JOIN



TableA

a_id	name
1	apple
2	orange
3	tomato
4	cucumber

TableB

b_id	name
A	apple
B	banana
C	cucumber
D	dill