

Predictive maintenance

Лекция 3

Распределённая файловая система HDFS

Власов Кирилл Вячеславович



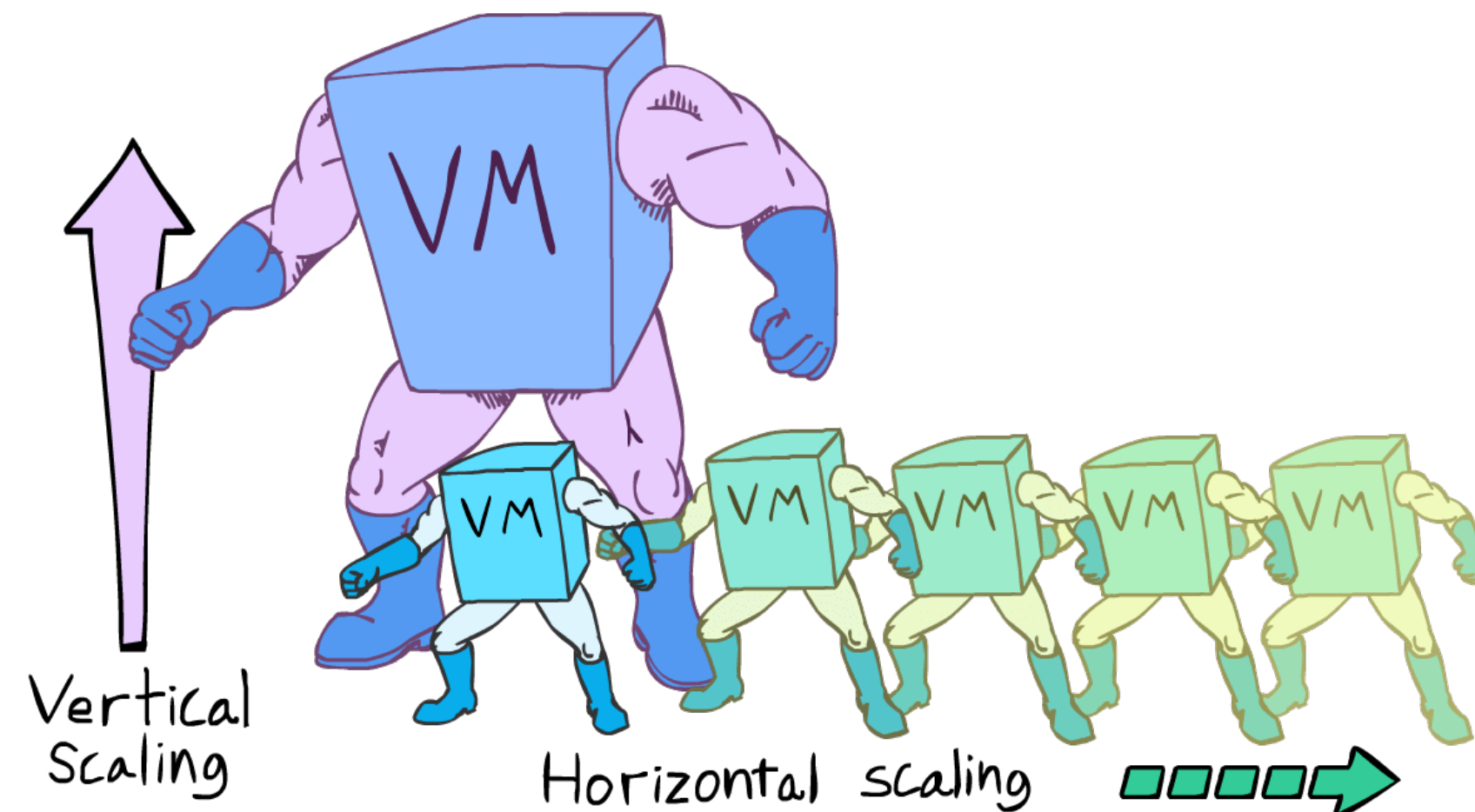
Google file system

Общие принципы построения распределенной файловой системы (GFS - Google file system) были опубликованы в 2003 году
<https://static.googleusercontent.com/media/research.google.com/en//archive/gfs-sosp2003.pdf>



Цели

1



Scale Up

Scale Out

2



Fault tolerance

Google file system

Общие принципы построения распределенной файловой системы (GFS - Google file system) были опубликованы в 2003 году
<https://static.googleusercontent.com/media/research.google.com/en//archive/gfs-sosp2003.pdf>



Критерии

Система должна хранить много больших файлов.

Несколько миллионов файлов, каждый от 100 Мб и больше. Маленькие файлы должны храниться, но для них не оптимизируется работа системы.

Google file system

Общие принципы построения распределенной файловой системы (GFS - Google file system) были опубликованы в 2003 году <https://static.googleusercontent.com/media/research.google.com/en//archive/gfs-sosp2003.pdf>



Критерии

Система должна хранить много больших файлов.

Несколько миллионов файлов, каждый от 100 Мб и больше. Маленькие файлы должны храниться, но для них не оптимизируется работа системы.

Система строится из большого количества недорогого оборудования. Должны существовать мониторинг сбоев, и возможность в случае отказа какого-либо оборудования восстановить функционирование системы.

Google file system

Общие принципы построения распределенной файловой системы (GFS - Google file system) были опубликованы в 2003 году <https://static.googleusercontent.com/media/research.google.com/en//archive/gfs-sosp2003.pdf>



Критерии

Система должна хранить много больших файлов.
Несколько миллионов файлов, каждый от 100 Мб и больше. Маленькие файлы должны храниться, но для них не оптимизируется работа системы.

Высокая пропускная способность является более предпочтительной, чем маленькая задержка.

Система строится из большого количества недорогого оборудования. Должны существовать мониторинг сбоев, и возможность в случае отказа какого-либо оборудования восстановить функционирование системы.

Google file system

Общие принципы построения распределенной файловой системы (GFS - Google file system) были опубликованы в 2003 году <https://static.googleusercontent.com/media/research.google.com/en//archive/gfs-sosp2003.pdf>



Критерии

Система должна хранить много больших файлов.

Несколько миллионов файлов, каждый от 100 Мб и больше. Маленькие файлы должны храниться, но для них не оптимизируется работа системы.

Система строится из большого количества недорогого оборудования. Должны существовать мониторинг сбоев, и возможность в случае отказа какого-либо оборудования восстановить функционирование системы.

Высокая пропускная способность является более предпочтительной, чем маленькая задержка.

Чтение большого последовательного фрагмента данных и чтение маленького объема произвольных данных.

Приложения, критические по времени исполнения, должны накопить определенное количество таких запросов и отсортировать их по смещению от начала файла. Это позволит избежать при чтении блужданий вида назад-вперед

Google file system

Общие принципы построения распределенной файловой системы (GFS - Google file system) были опубликованы в 2003 году <https://static.googleusercontent.com/media/research.google.com/en//archive/gfs-sosp2003.pdf>



Критерии

Система должна хранить много больших файлов.
Несколько миллионов файлов, каждый от 100 Мб и больше. Маленькие файлы должны храниться, но для них не оптимизируется работа системы.

Система строится из большого количества недорогого оборудования. Должны существовать мониторинг сбоев, и возможность в случае отказа какого-либо оборудования восстановить функционирование системы.

Высокая пропускная способность является более предпочтительной, чем маленькая задержка.

Чтение большого последовательного фрагмента данных и чтение маленького объема произвольных данных.
Приложения, критические по времени исполнения, должны накопить определенное количество таких запросов и отсортировать их по смещению от начала файла. Это позволит избежать при чтении блужданий вида назад-вперед

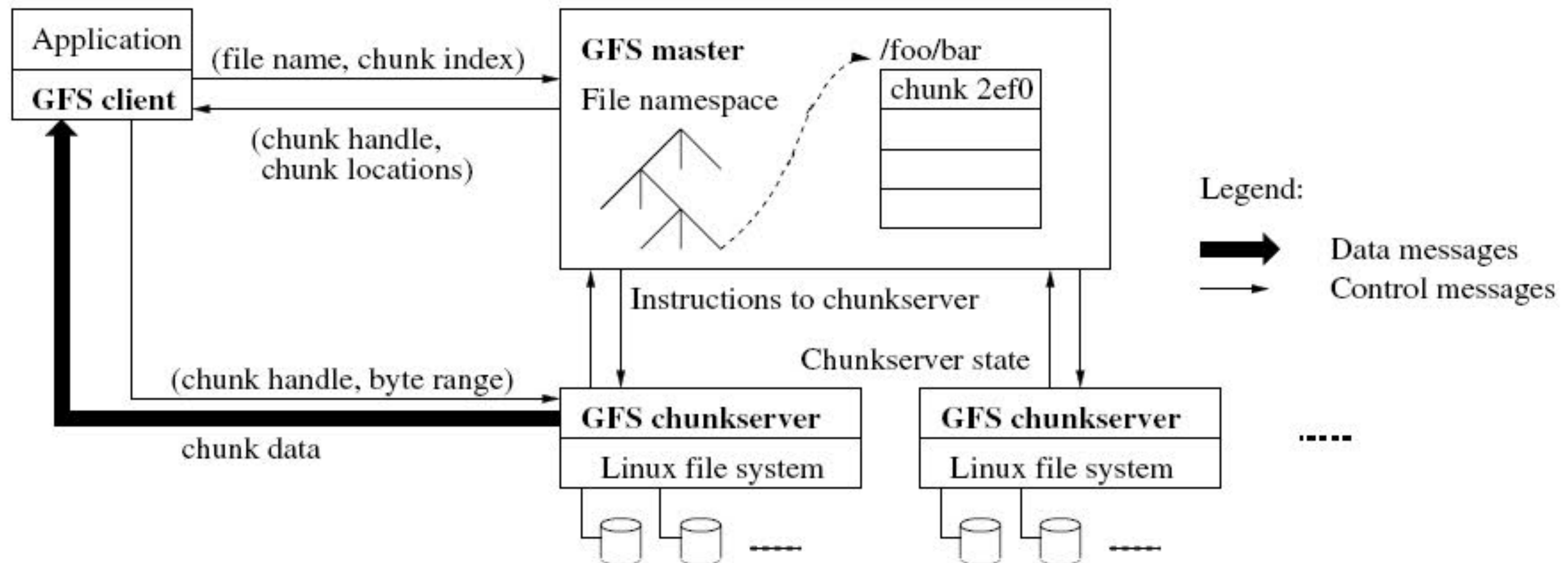
Система должна реализовывать строго очерченную семантику параллельной работы нескольких клиентов, в случае если они одновременно пытаются дописать данные в один и тот же файл. При этом может случиться так, что поступят одновременно сотни запросов на запись в один файл. Для того чтобы справиться с этим, используется атомарность операций добавления данных в файл, с некоторой синхронизацией. То есть если поступит операция на чтение, то она будет выполняться, либо до очередной операции записи, либо после.

Google file system

Общие принципы построения распределенной файловой системы (GFS - Google file system) были опубликованы в 2003 году
<https://static.googleusercontent.com/media/research.google.com/en//archive/gfs-sosp2003.pdf>



Архитектура



Google file system

Общие принципы построения распределенной файловой системы (GFS - Google file system) были опубликованы в 2003 году
<https://static.googleusercontent.com/media/research.google.com/en//archive/gfs-sosp2003.pdf>



Архитектура

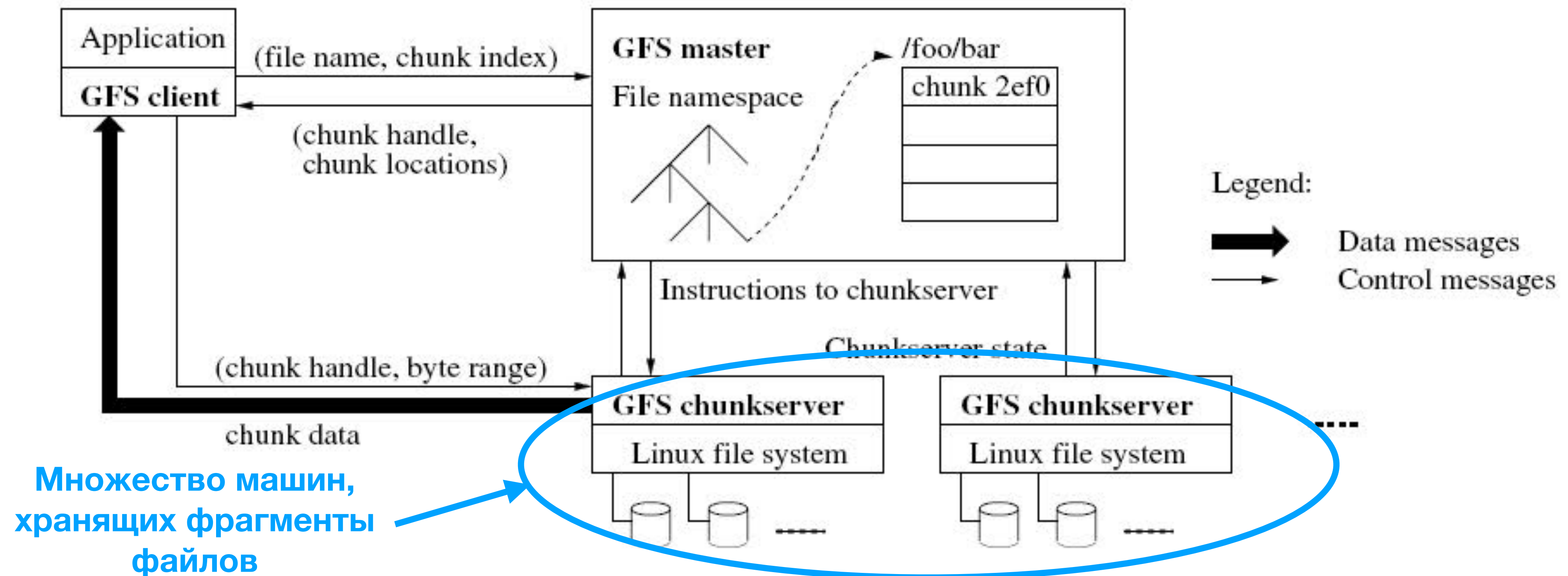


Google file system

Общие принципы построения распределенной файловой системы (GFS - Google file system) были опубликованы в 2003 году
<https://static.googleusercontent.com/media/research.google.com/en//archive/gfs-sosp2003.pdf>



Архитектура

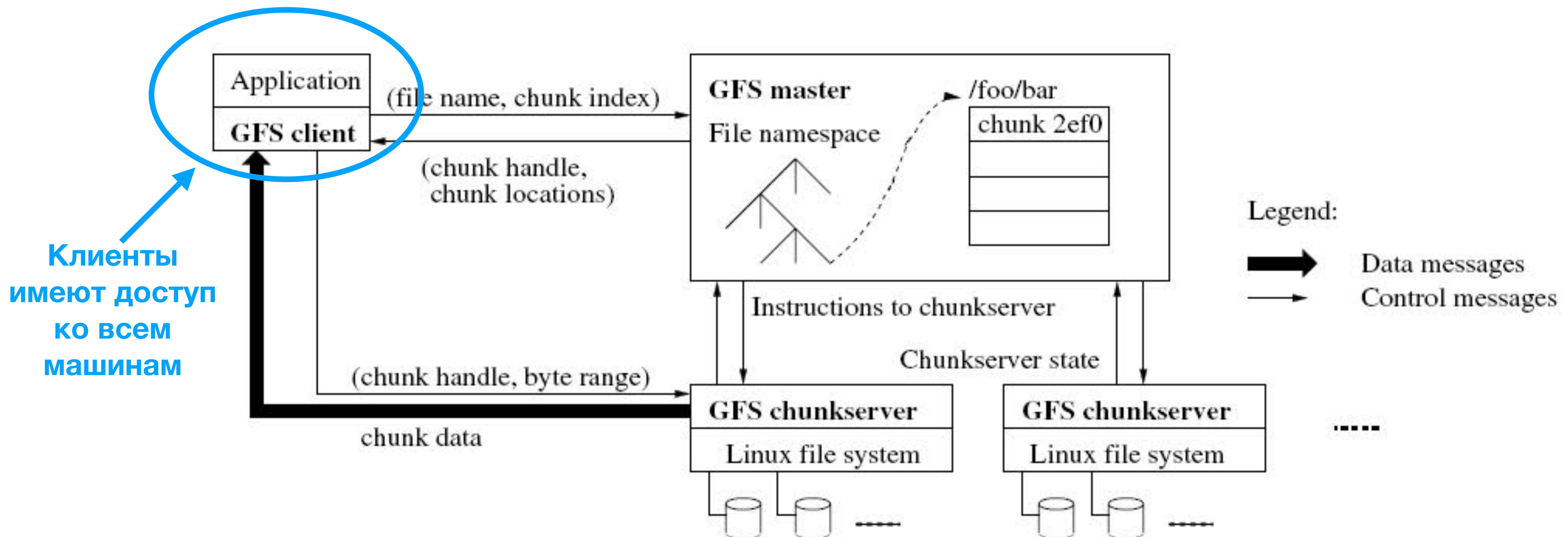


Google file system

Общие принципы построения распределенной файловой системы (GFS - Google file system) были опубликованы в 2003 году
<https://static.googleusercontent.com/media/research.google.com/en//archive/gfs-sosp2003.pdf>



Архитектура

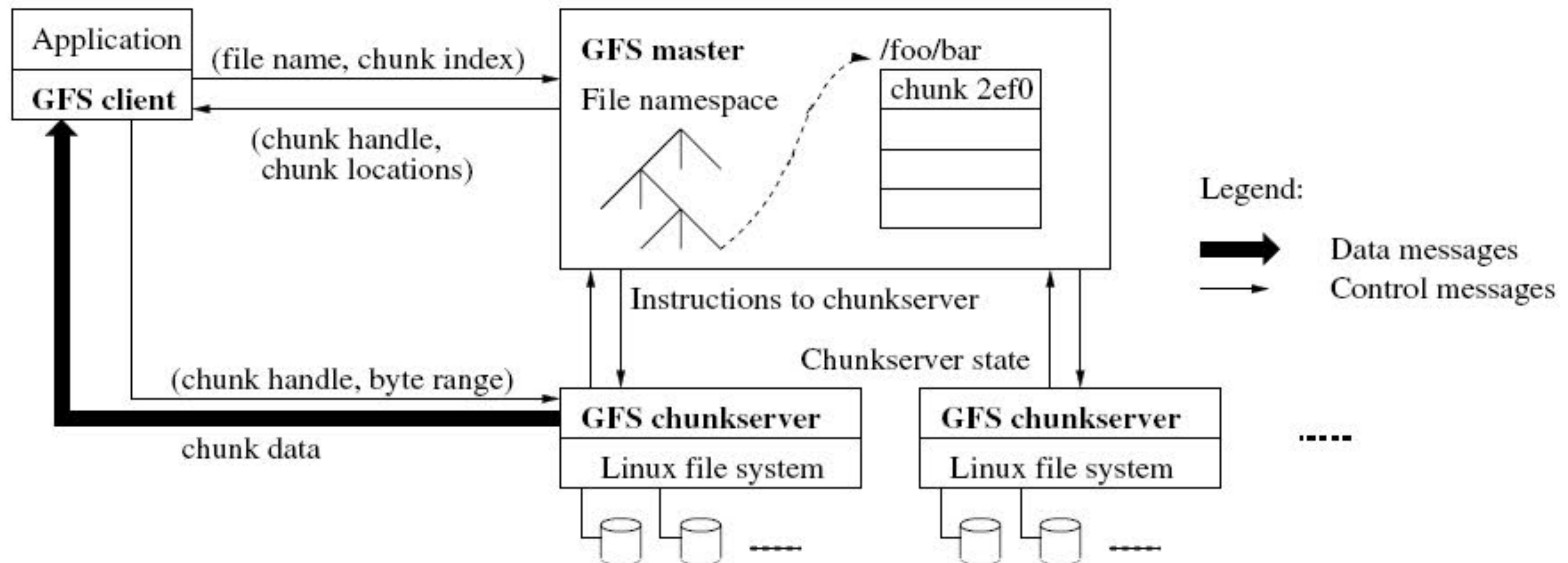


Google file system

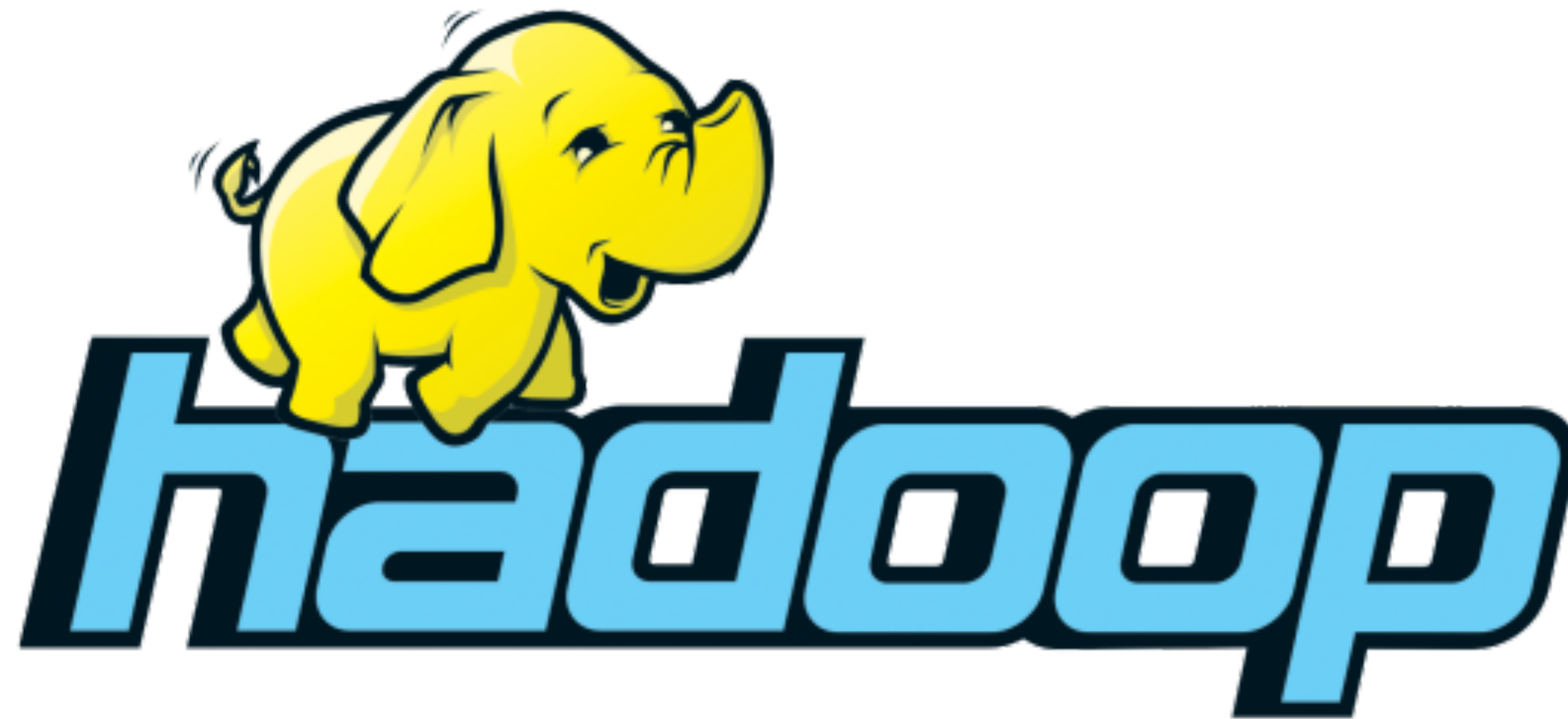
Общие принципы построения распределенной файловой системы (GFS - Google file system) были опубликованы в 2003 году
<https://static.googleusercontent.com/media/research.google.com/en//archive/gfs-sosp2003.pdf>



Архитектура



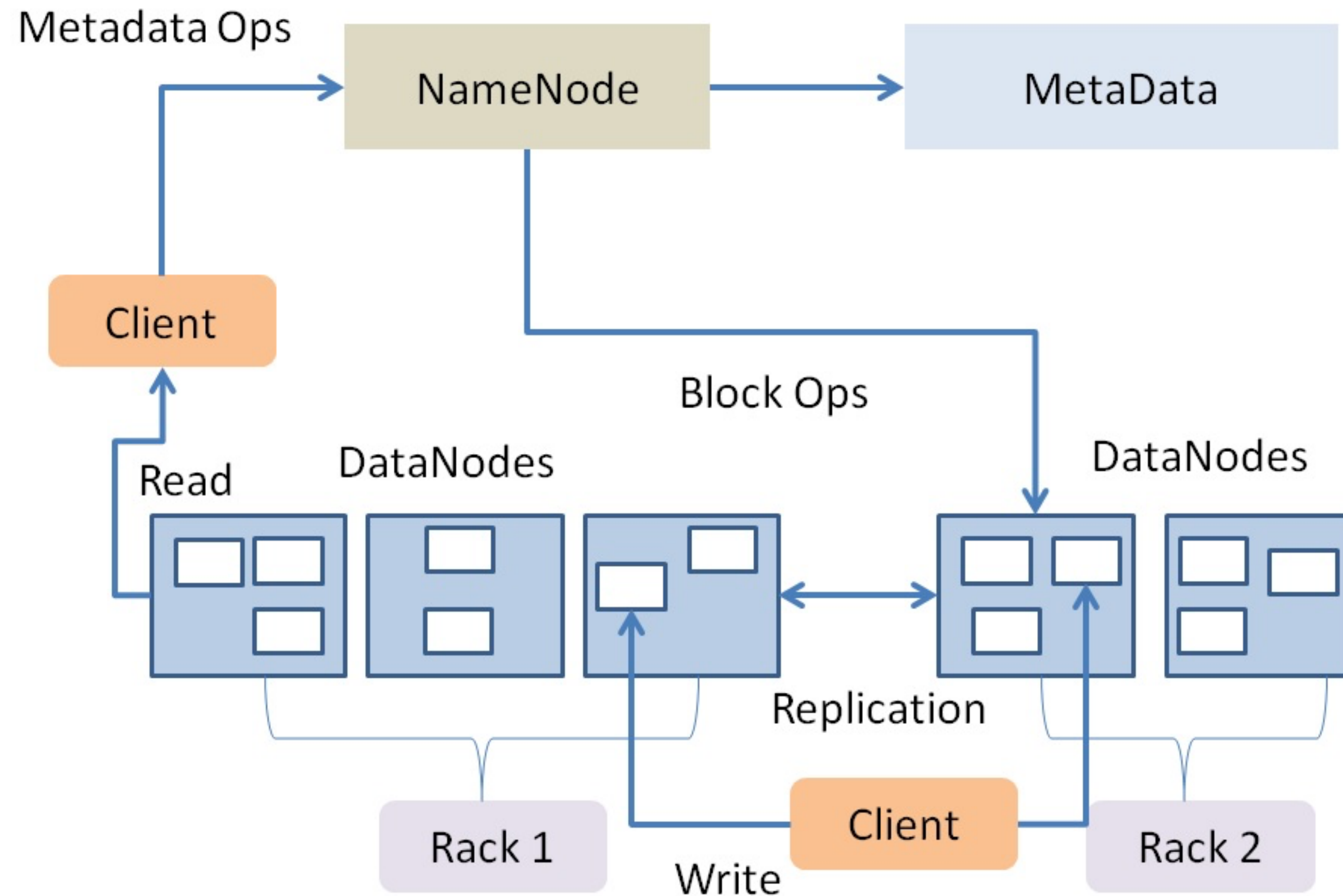
Hadoop Distributed File System



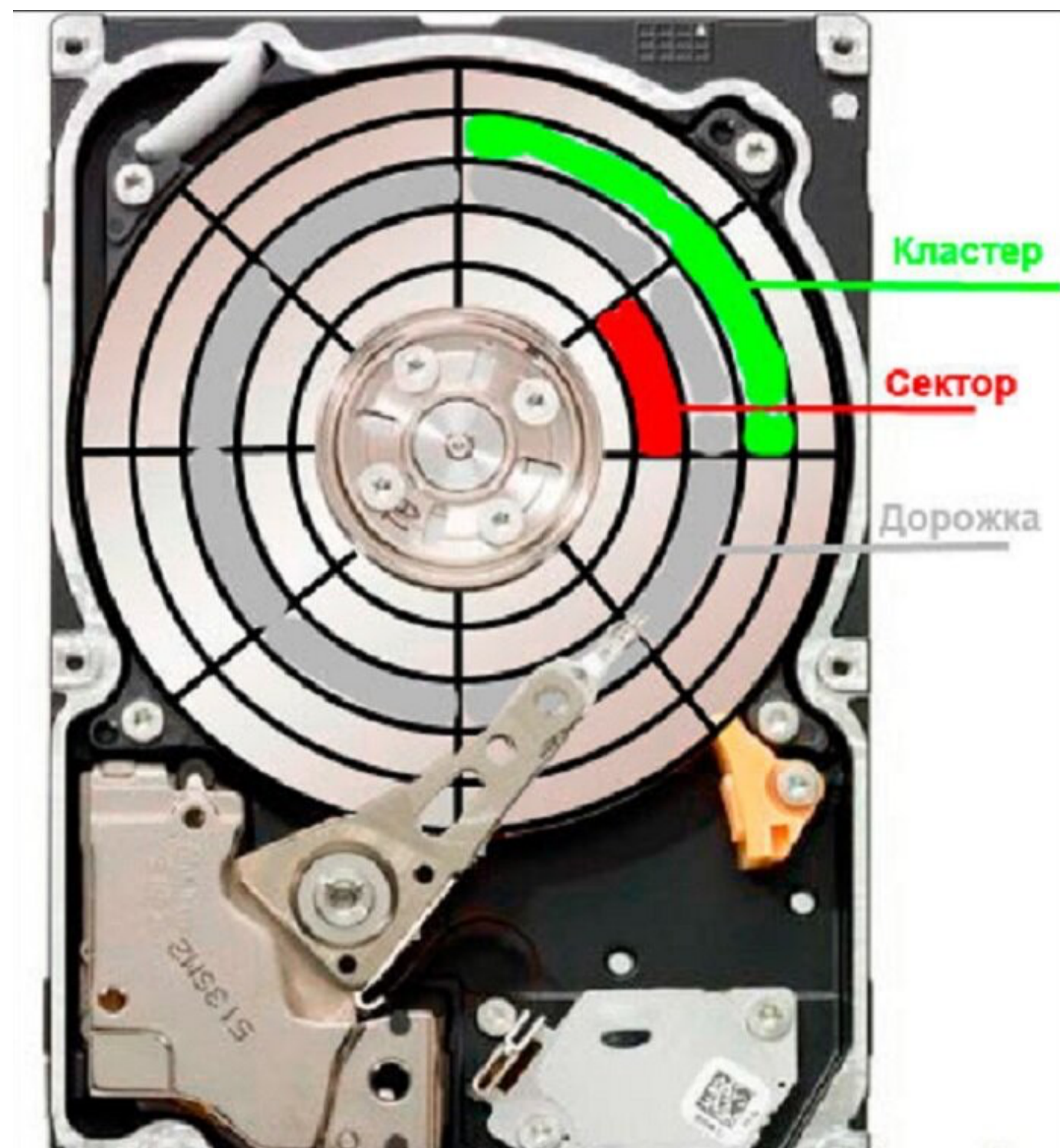
Hadoop Distributed File System



Архитектура



Файлы и блоки



В HDFS файлы хранятся в блоках. При этом стандартный размер блока ***64 Mb*** или ***128 Mb***

Rack Awareness Algorithim

Block A: 

Block B: 

Block C: 

Rack - 1

1



2



3

4



Rack - 2

5

6



7



8



Rack - 3

9



10

11



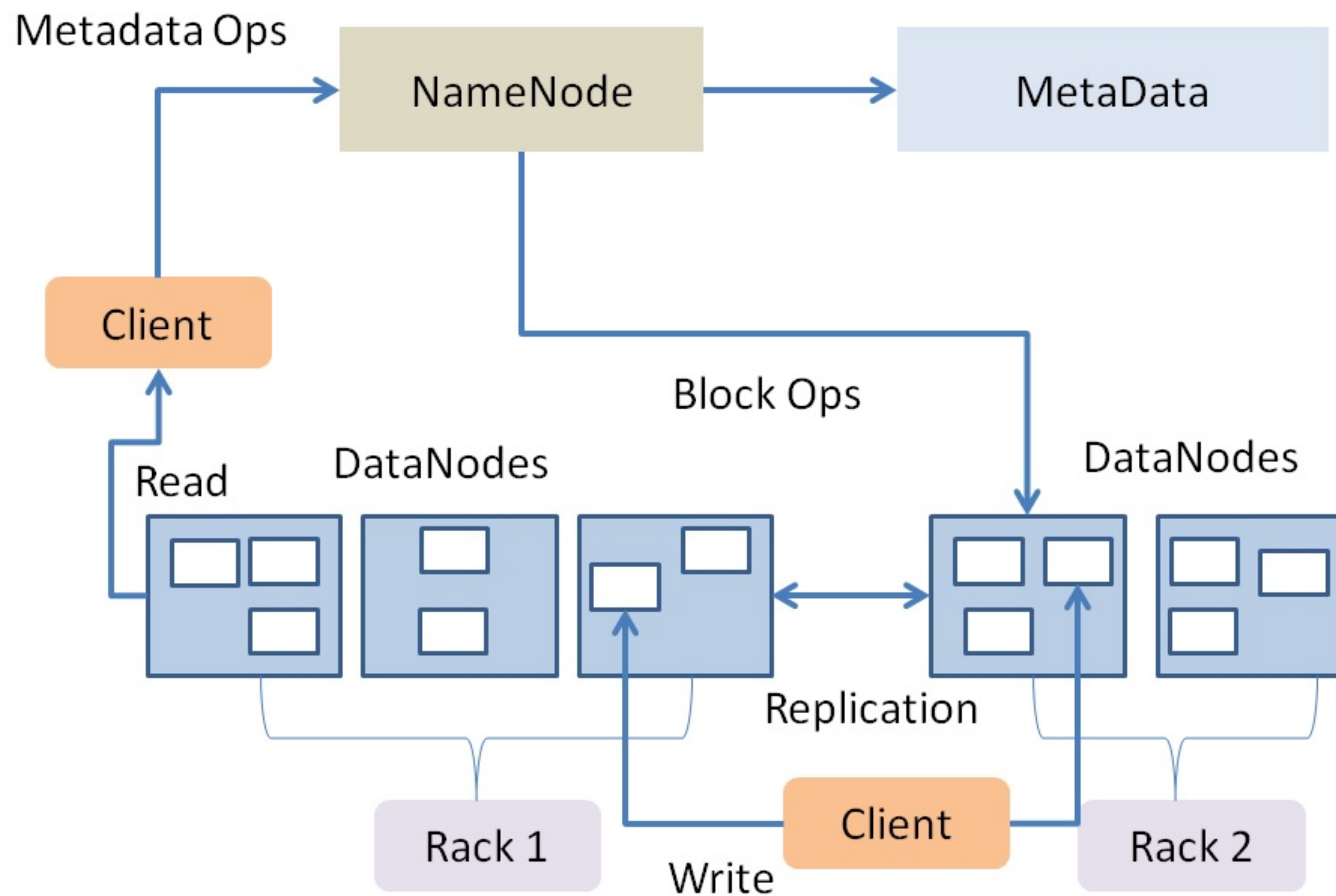
12



Hadoop Distributed File System



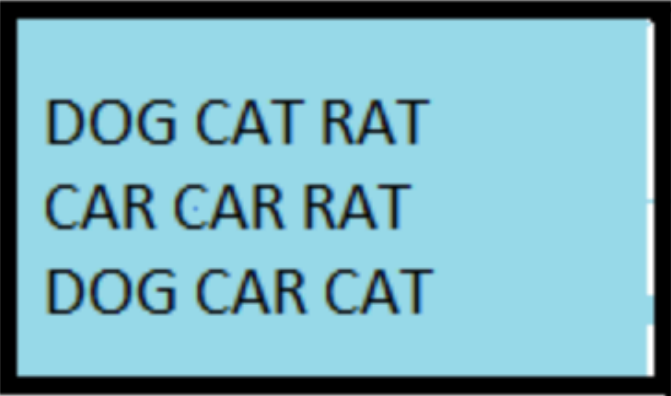
Архитектура



Парадигма MapReduce

Парадигма MapReduce

Input

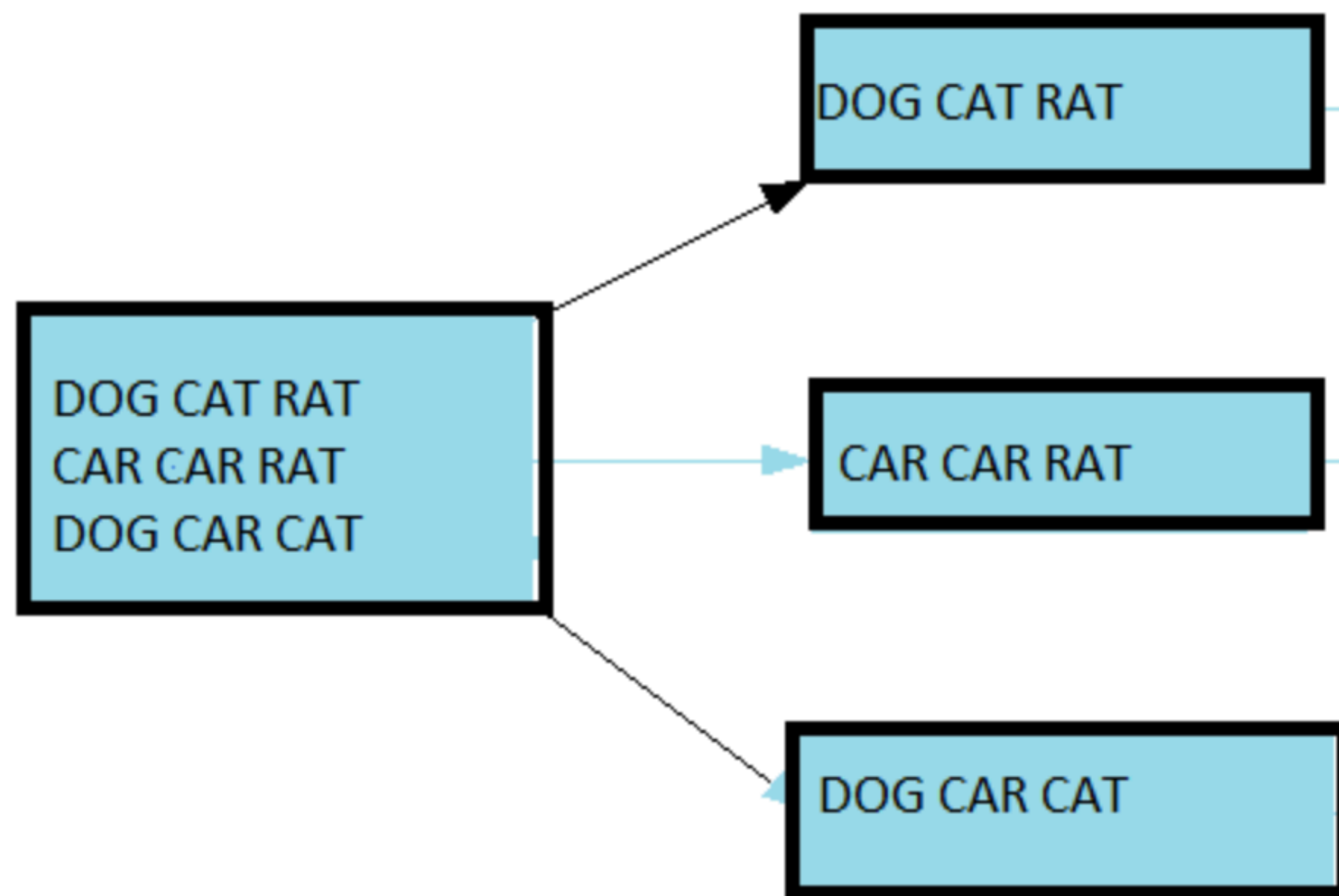


```
DOG CAT RAT  
CAR CAR RAT  
DOG CAR CAT
```

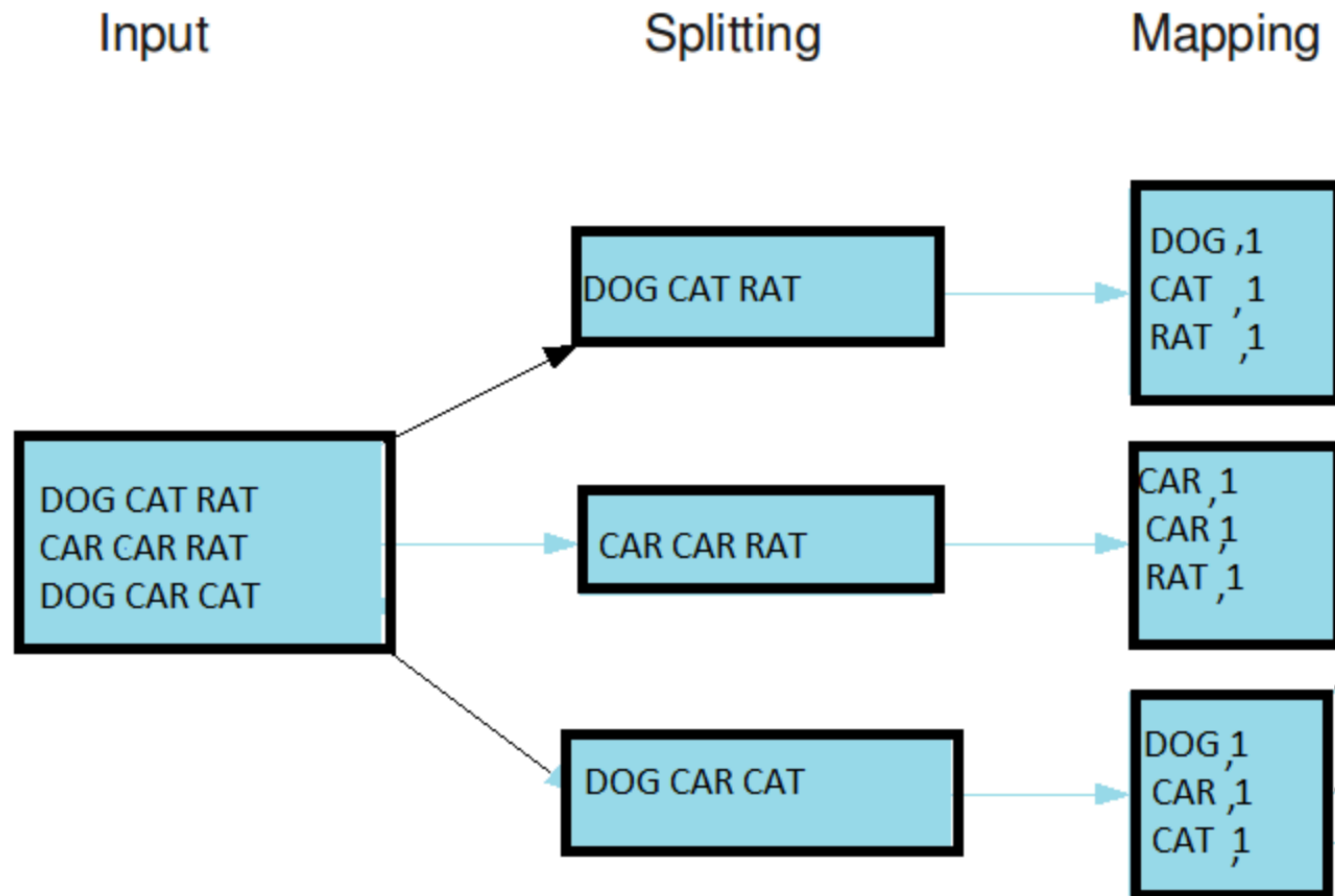
Парадигма MapReduce

Input

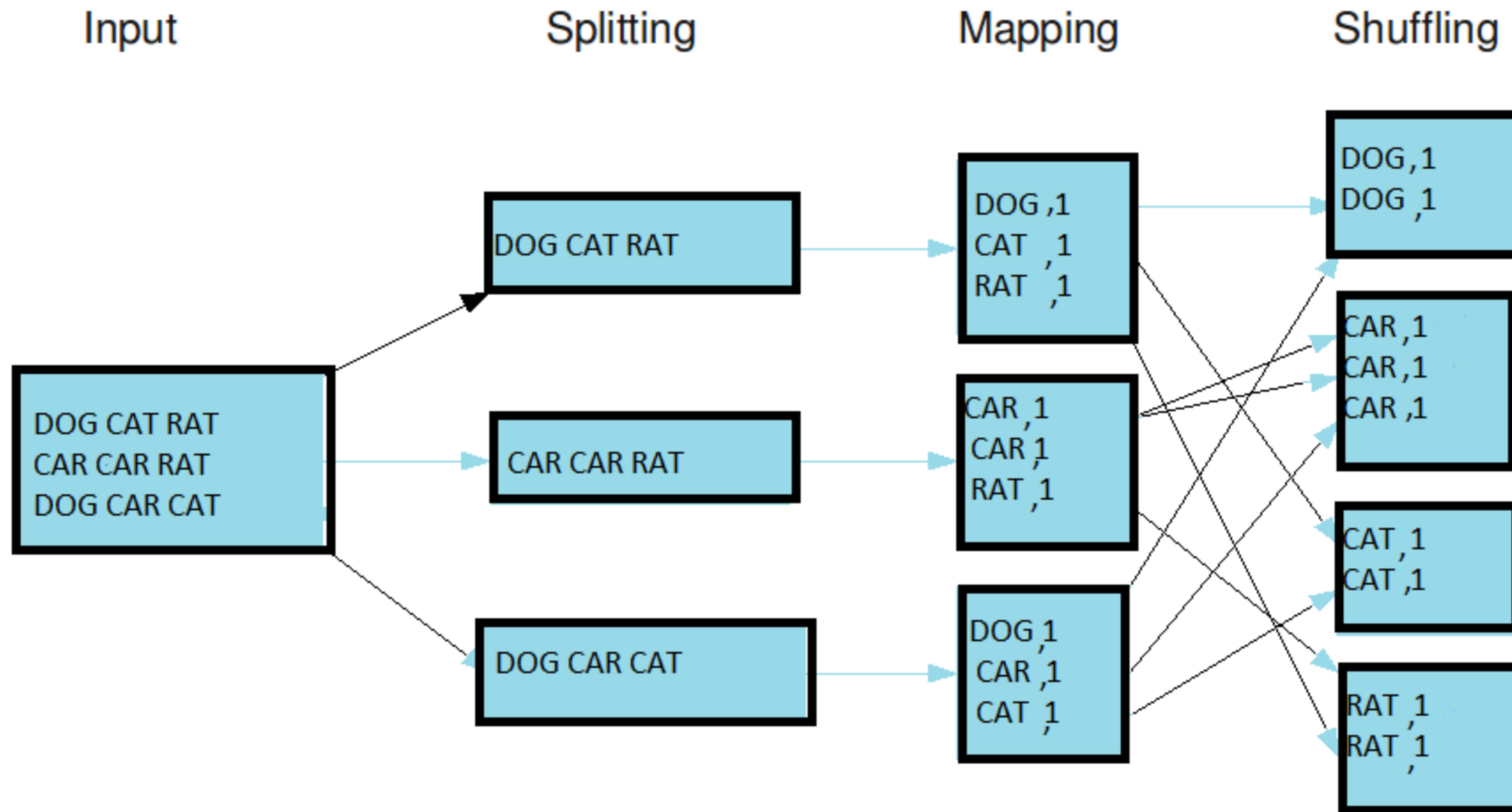
Splitting



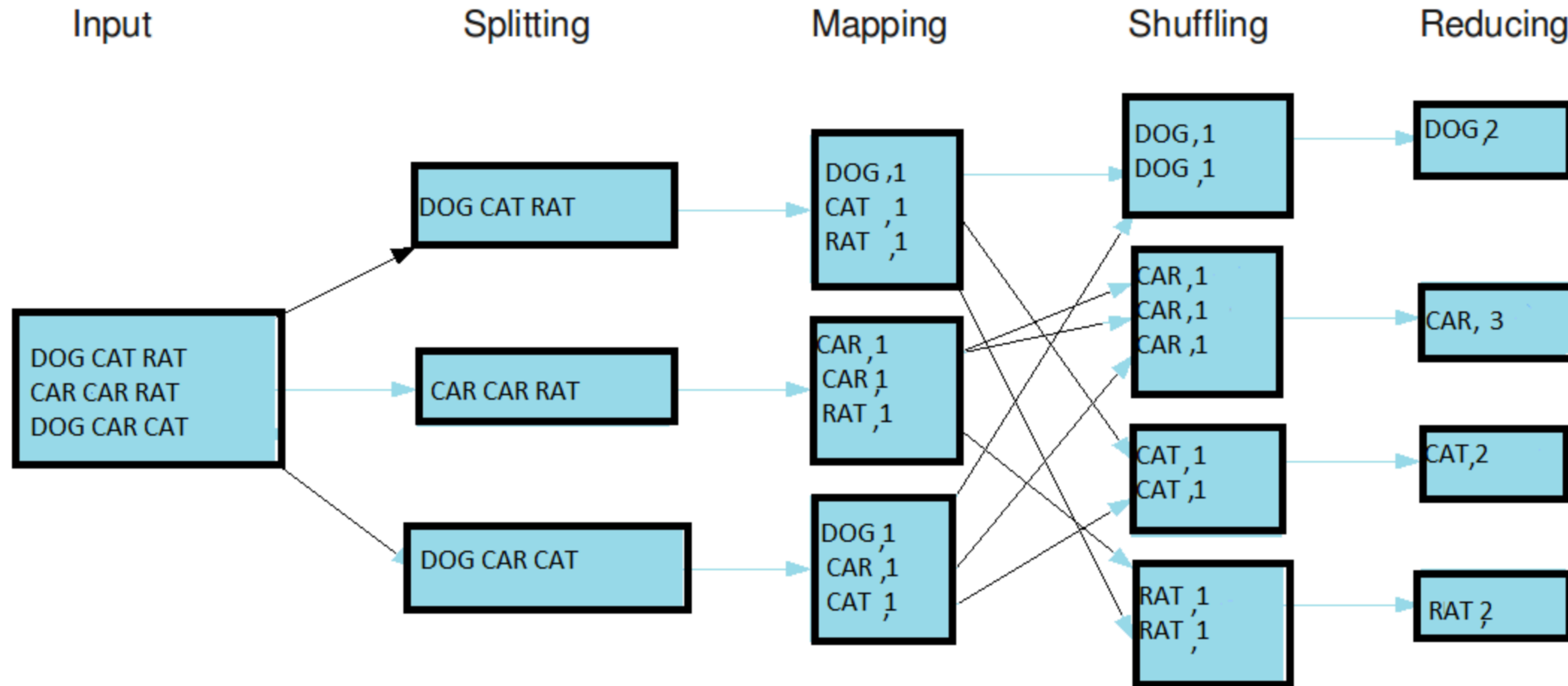
Парадигма MapReduce



Парадигма MapReduce



Парадигма MapReduce



Парадигма MapReduce

