

### *CP.3: Minimize explicit sharing of writable data*





ОНЛАЙН-ОБРАЗОВАНИЕ



# Хранилища Много данных

Дмитрий Шебордаев

Ведущий специалист



## Сегодня

- ▶ Опять день открытых дверей?
- ▶ Считаем слова
- ▶ Пишем опечаточник
- ▶ Балобол
- ▶ MapReduce
- ▶ Распределяем вычисления



## Опять день открытых дверей?

Patrick Lucy Ignacio Carlo Rudy Carly Christi Rachael Cletus  
Lesley Cheri Abraham Clement Gerald Aimee Janet Susie Eula  
William Dorian



Опять день открытых дверей?  
P L I C R C C R C L C A C G A J S E W D



Опять день открытых дверей?  
A A C C C C C C D E G I J L L P R R S W



## Опять день открытых дверей?

- ▶ 6 C
- ▶ 2 R
- ▶ 2 L
- ▶ 2 A





## Опять день открытых дверей?

- ▶ 2 Cl
- ▶ 2 Ch
- ▶ 2 Ca



## Опять день открытых дверей?

```
cut -c -5 names.txt | sort | uniq -d
```



## Считаем слова

I would have gotten the promotion, but my attendance wasn't good enough. The book is in front of the table. Wednesday is hump day, but has anyone asked the camel if he's happy about it? A glittering gem is not enough. A purple pig and a green donkey flew a kite in the middle of the night and ended up sunburnt. Two seats were vacant. Joe made the sugar cookies; Susan decorated them. The quick brown fox jumps over the lazy dog. How was the math test? I'd rather be a bird than a fish.



## Считаем слова

a a a a a about and and anyone asked attendance be bird  
book brown but but camel cookies day decorated dog donkey  
ended enough enough fish flew fox front gem glittering good  
gotten green happy has have hes how hump i id if in in is is is it  
joe jumps kite lazy made math middle my night not of of over  
pig promotion purple quick rather seats sugar sunburnt susan  
table test than the the the the the the the the the them two  
up vacant was wasnt wednesday were would



## Считаем слова

- ▶ 10 the
- ▶ 6 a
- ▶ 3 is
- ▶ 2 of
- ▶ 2 in
- ▶ 2 enough
- ▶ 2 but
- ▶ 2 and



## Пишем опечаточник

- I would have gotten the promotion, but my attendance wasn't good enough.



## Пишем опечаточник

- ▶ I would have gotten the promotion, but my attendance wasn't good enough.
- ▶ wou oul uld hav ave got ott tte ten the pro rom omo mot oti  
tio ion but att tte ten end nda dan anc nce was asn snt goo  
ood eno nou oug ugh



## Пишем опечаточник

- ▶ I would have gotten the promotion, but my attendance wasn't good enough.
- ▶ wou oul uld hav ave got ott tte ten the pro rom omo mot oti  
tio ion but att tte ten end nda dan anc nce was asn snt goo  
ood eno nou oug ugh
- ▶ te d n n r s





## Балобол

- ▶ A - B C D
- ▶ B - D
- ▶ B - A

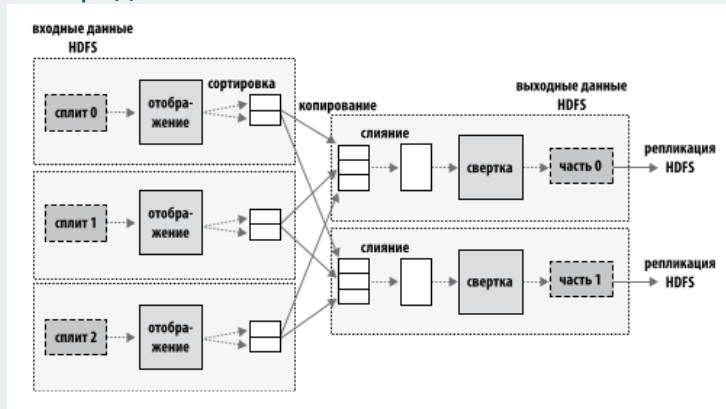


## MapReduce

map -> sort -> reduce



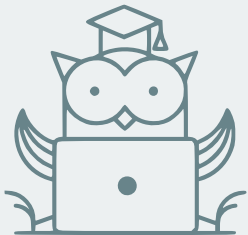
## Распределяем вычисления





# Hadoop: The Definitive Guide 3rd Edition 978-1449311520





Спасибо за  
внимание!

