



# **Hadoop Sizing: оценка вычислительных мощностей для хранения данных**

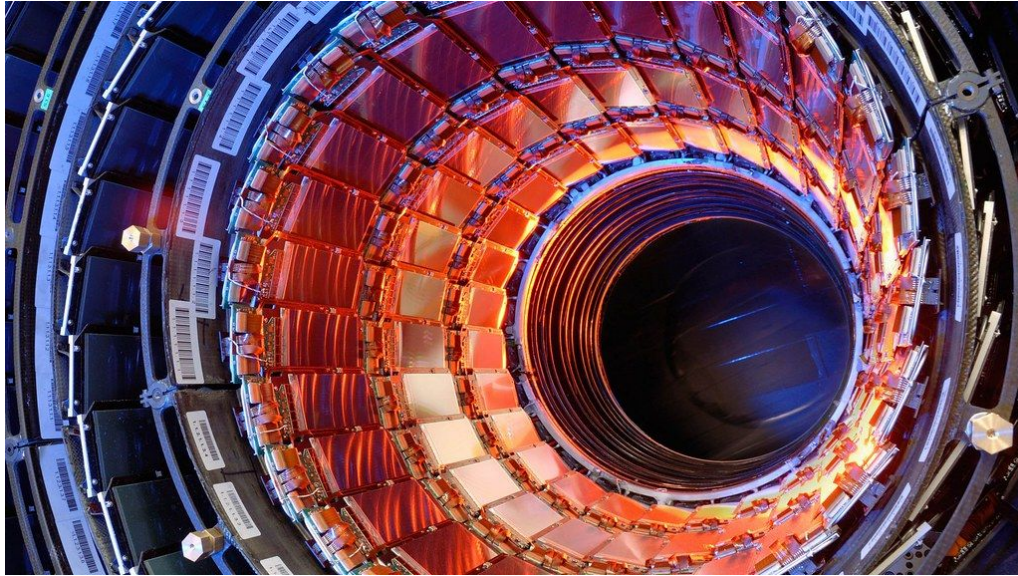
**Драль Алексей**, [study@bigdatateam.org](mailto:study@bigdatateam.org)

CEO at BigData Team, <https://bigdatateam.org>

<https://www.facebook.com/bigdatateam>



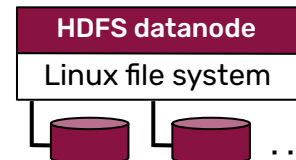
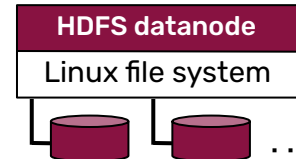
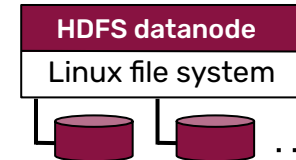
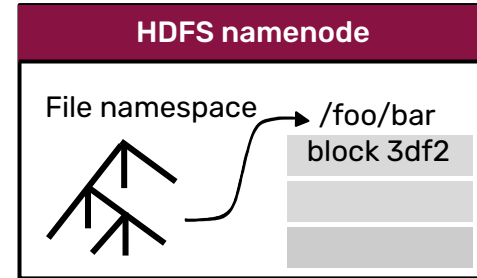
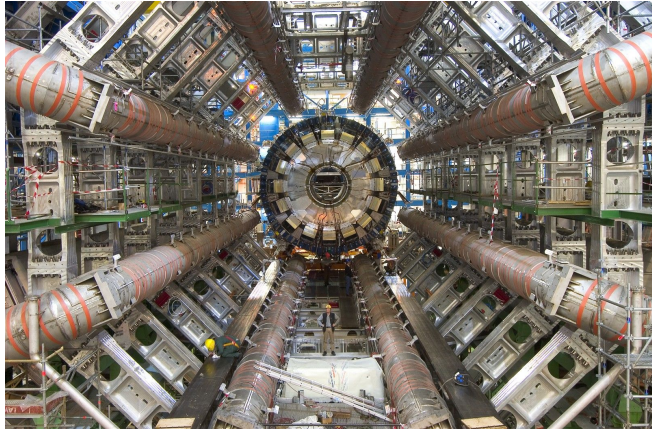
# Запрос на консультацию



The CERN Data Centre passed a major milestone on 29 June 2017 with more than **200 petabytes** of data now archived on tape

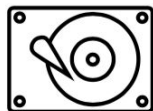
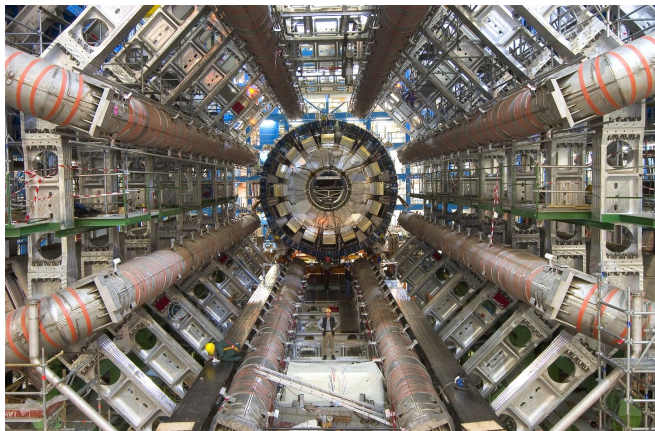


# Hadoop Sizing

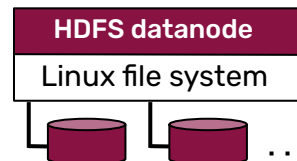
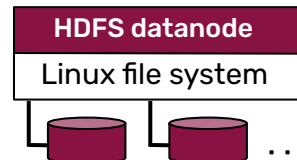
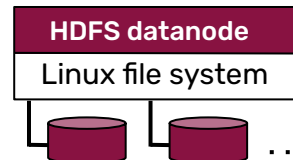
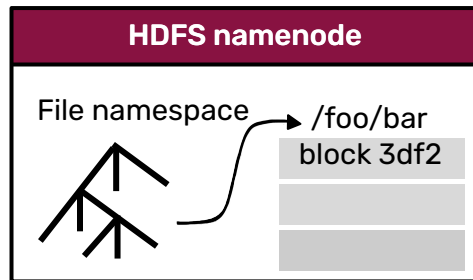




# Hadoop Sizing



10 PB / 2 TB \* 3 ~ 15 k





**BIGDATA  
TEAM**

Считаем сбои

 **BACKBLAZE**  
**DRIVE  
STATS  
2020**







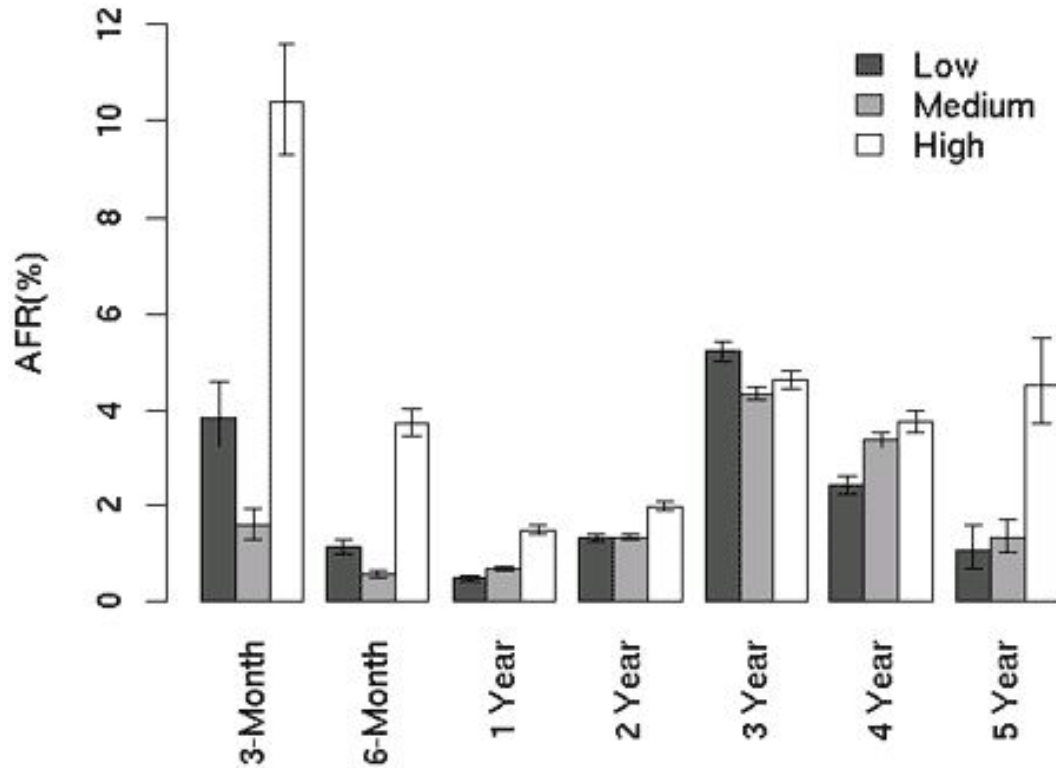
## Backblaze Hard Drive Failure Rates for 2020

Reporting period 1/1/2020 - 12/31/2020 inclusive

MFG	Model	Drive Size	Drive Count	Avg Age (months)	Drive Days	Drive Failures	AFR
HGST	HMS5C4040ALE640	4TB	3,100	56.65	1,083,774	8	0.27%
HGST	HMS5C4040BLE640	4TB	12,744	50.43	4,663,049	34	0.27%
HGST	HUH728080ALE600	8TB	1,075	34.85	372,000	3	0.29%
Seagate	ST12000NM001G	12TB	1,100	3.00	1,200,170	30	0.04%
Seagate	ST14000NM001G	14TB	5,987	2.89	454,090	13	1.04%
Seagate	ST14000NM0138	14TB	360	1.56	5,784	0	0.00%
Seagate	ST16000NM001G	16TB	59	12.93	21,323	1	1.71%
Seagate	ST18000NM000J	18TB	60	3.27	5,820	2	12.54%
Toshiba	MD04ADA400V	4TB	60	67.00	20,000	0	0.00%

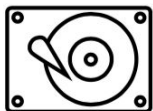
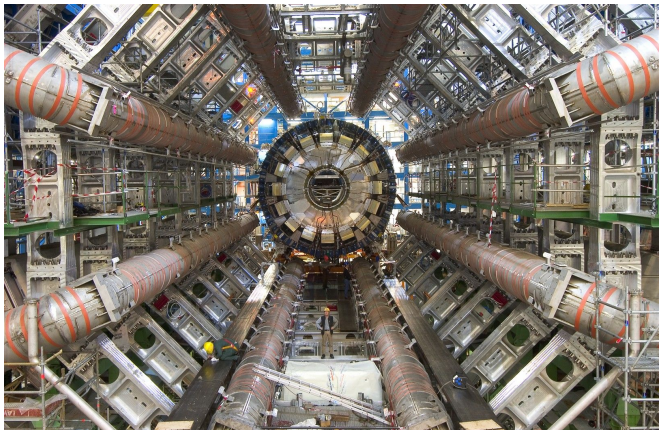


# Вероятность сбоя от перегрузки





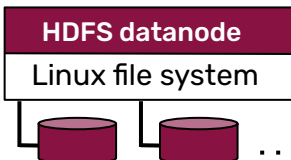
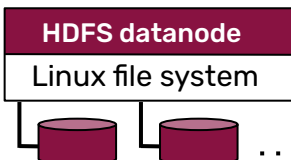
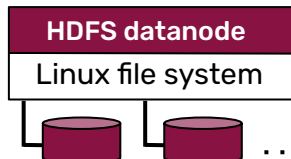
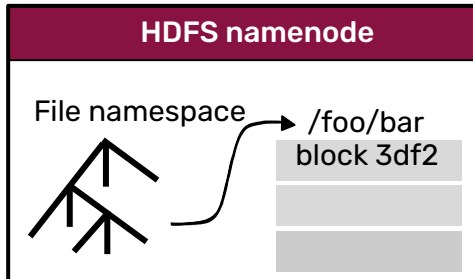
## Hadoop Sizing



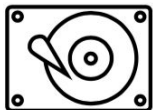
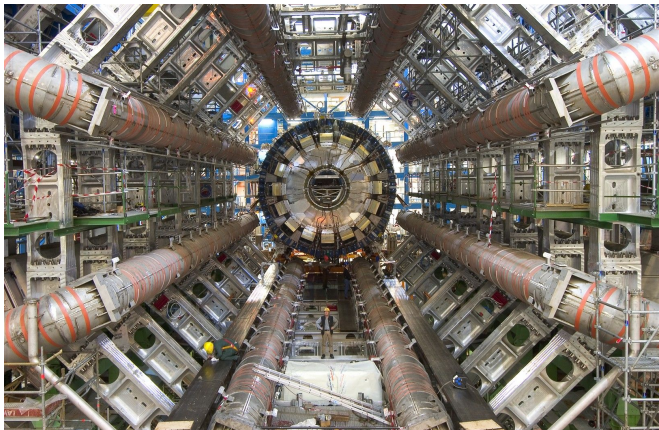
10 PB / 2 TB \* 3 ~ 15 k



150 B - average block size on Namenode





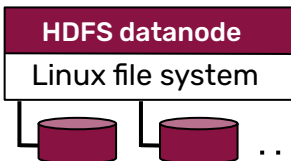
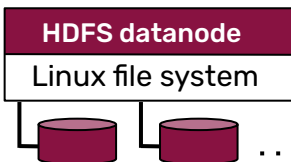
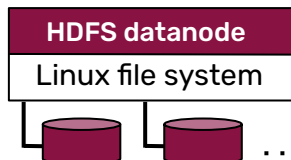
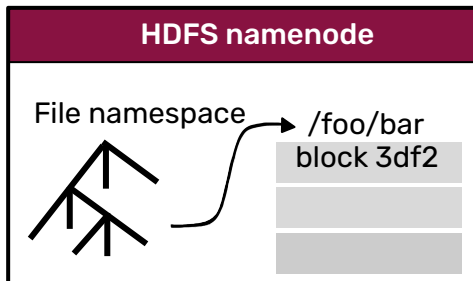


10 PB / 2 TB \* 3 ~ 15 k



150 B - average block size on Namenode

## Hadoop Sizing



Input interpretation:

$$\frac{10 \text{ PB (petabytes)}}{128 \text{ MB (megabytes)}} \times 150 \text{ bytes}$$

Result:

$$1.172 \times 10^{10} \text{ bytes}$$

Unit conversions:

$$11.72 \text{ GB (gigabytes)}$$



**BIGDATA  
TEAM**

# Hadoop “Small Files Problem”



**Когда залил много  
маленьких файлов в HDFS**



# Hadoop “Small Files Problem”



Когда залил много  
маленьких файлов в HDFS



- ▶ reading speed - 600 MB/sec
- ▶ 10 PB → **207** дней



- ▶ reading speed - 600 MB/sec
- ▶ 10 PB → **103.5** дней



RAM: 16GB, block size: **32MB**, replication factor: 3

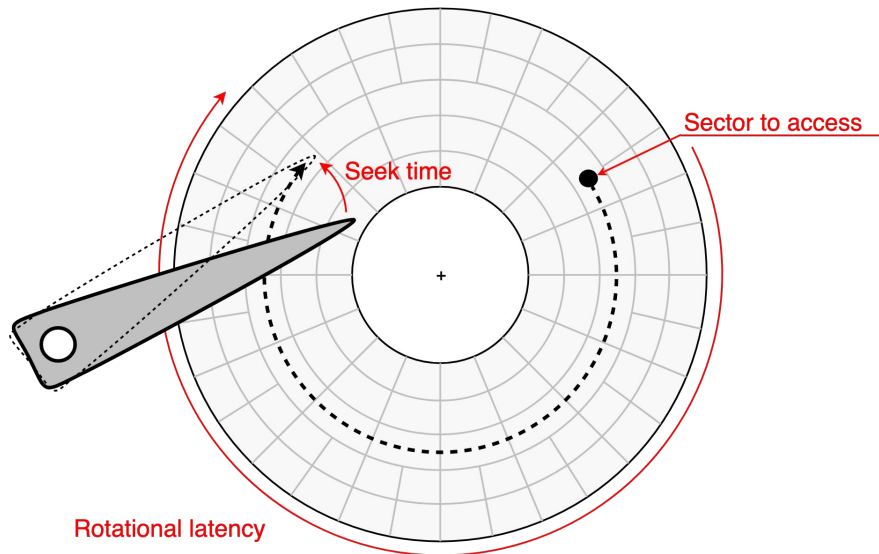
► максимальный размер хранилища:  $16\text{GB} / 150\text{B} \times \mathbf{32\text{MB}} / 3 = 1.138\text{PB}$



# Выбор Block Size

RAM: 16GB, block size: **32MB**, replication factor: 3

► максимальный размер хранилища:  $16\text{GB} / 150\text{B} \times \mathbf{32\text{MB}} / 3 = 1.138\text{PB}$



- HDD seek time: 0.2-0.8 мс
- SSD ~seek time: 0.08-0.16 мс

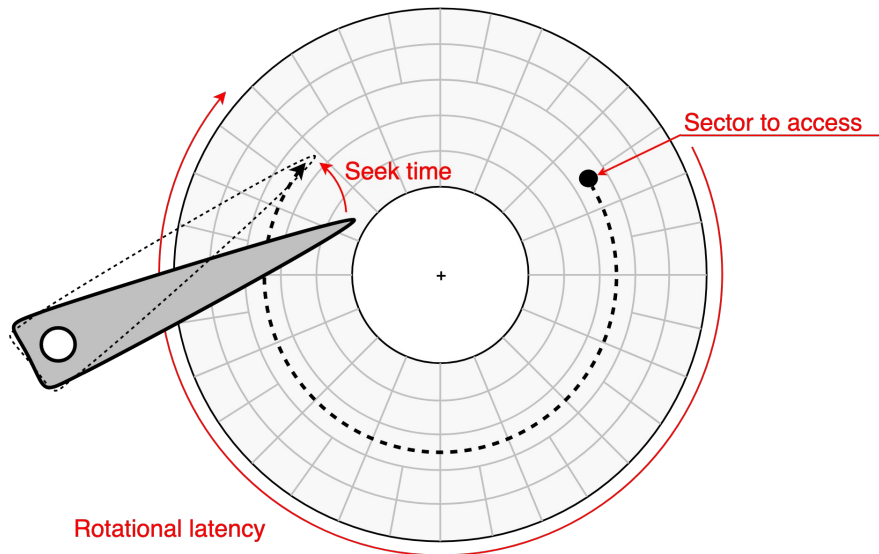




# Выбор Block Size

RAM: 16GB, block size: **32MB**, replication factor: 3

- ▶ максимальный размер хранилища:  $16\text{GB} / 150\text{B} \times \mathbf{32\text{MB}} / 3 = 1.138\text{PB}$



- ▶ HDD seek time: 0.2-0.8 мс
- ▶ SSD ~seek time: 0.08-0.16 мс
  
- ▶ reading speed - 600 MB/сек
- ▶ 32 MB ~ 50 мс



# Алгоритмы во внешней памяти



Алгоритмы во внешней памяти | Максим Бабенко

См. также: [https://en.wikipedia.org/wiki/External\\_memory\\_algorithm](https://en.wikipedia.org/wiki/External_memory_algorithm)



**BIGDATA  
TEAM**

Резюме



Вы должны уметь:

- ▶ Поговорить за Small Files Problem



Вы должны уметь:

- ▶ Поговорить за Small Files Problem
- ▶ Решить задачу Hadoop Sizing





Вы должны уметь:

- ▶ Поговорить за Small Files Problem
- ▶ Решить задачу Hadoop Sizing
- ▶ Оценить влияние Block Size на Namenode RAM