



Распределенные файловые системы на примере GFS и HDFS

Драль Алексей, study@bigdatateam.org

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

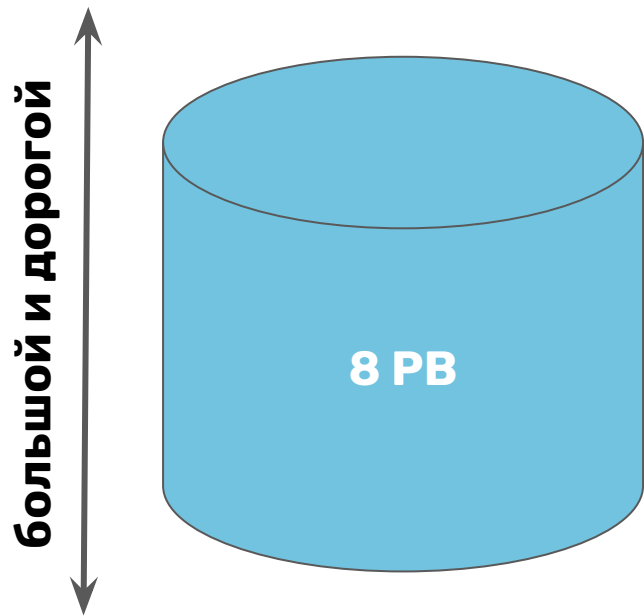
<https://www.facebook.com/bigdatateam>

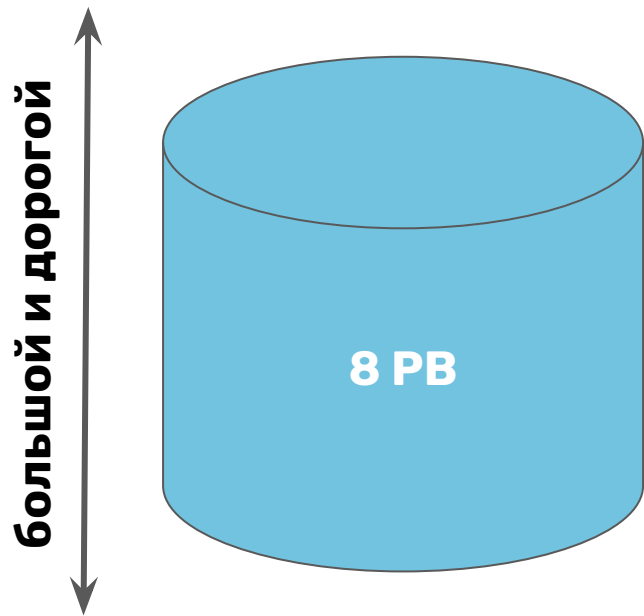


**BIGDATA
TEAM**

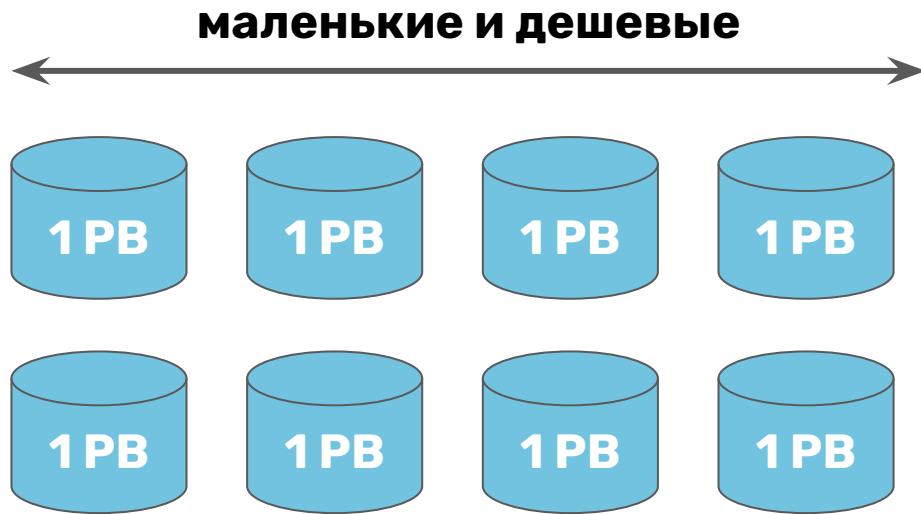
большой и дорогой



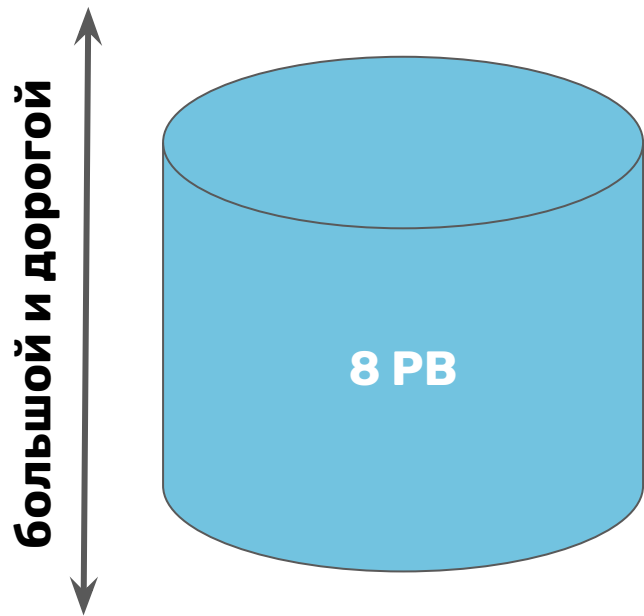




**вертикальное масштабирование
(scale-up)**



**горизонтальное масштабирование
(scale-out)**



**вертикальное масштабирование
(scale-up)**



**горизонтальное масштабирование
(scale-out)**



The Google File System

Sanjay Ghemawat, Howard Gobioff, and Shun-Tak Leung

Google*

ABSTRACT

We have designed and implemented the Google File System, a scalable distributed file system for large distributed data-intensive applications. It provides fault tolerance while running on inexpensive commodity hardware, and it delivers high aggregate performance to a large number of clients.

While sharing many of the same goals as previous distributed file systems, our design has been driven by observations of our application workloads and technological environment, both current and anticipated, that reflect a marked departure from some earlier file system assumptions. This has led us to reexamine traditional choices and explore rad-

1. INTRODUCTION

We have designed and implemented the Google File System (GFS) to meet the rapidly growing demands of Google's data processing needs. GFS shares many of the same goals as previous distributed file systems such as performance, scalability, reliability, and availability. However, its design has been driven by key observations of our application workloads and technological environment, both current and anticipated, that reflect a marked departure from some earlier file system design assumptions. We have reexamined traditional choices and explored radically different points in the design space.

The Google File System, Symposium on Operating Systems Principles (SOSP, 2003)



Q&A

Как обеспечить гарантии хранения
данных в рамках fail-recovery?

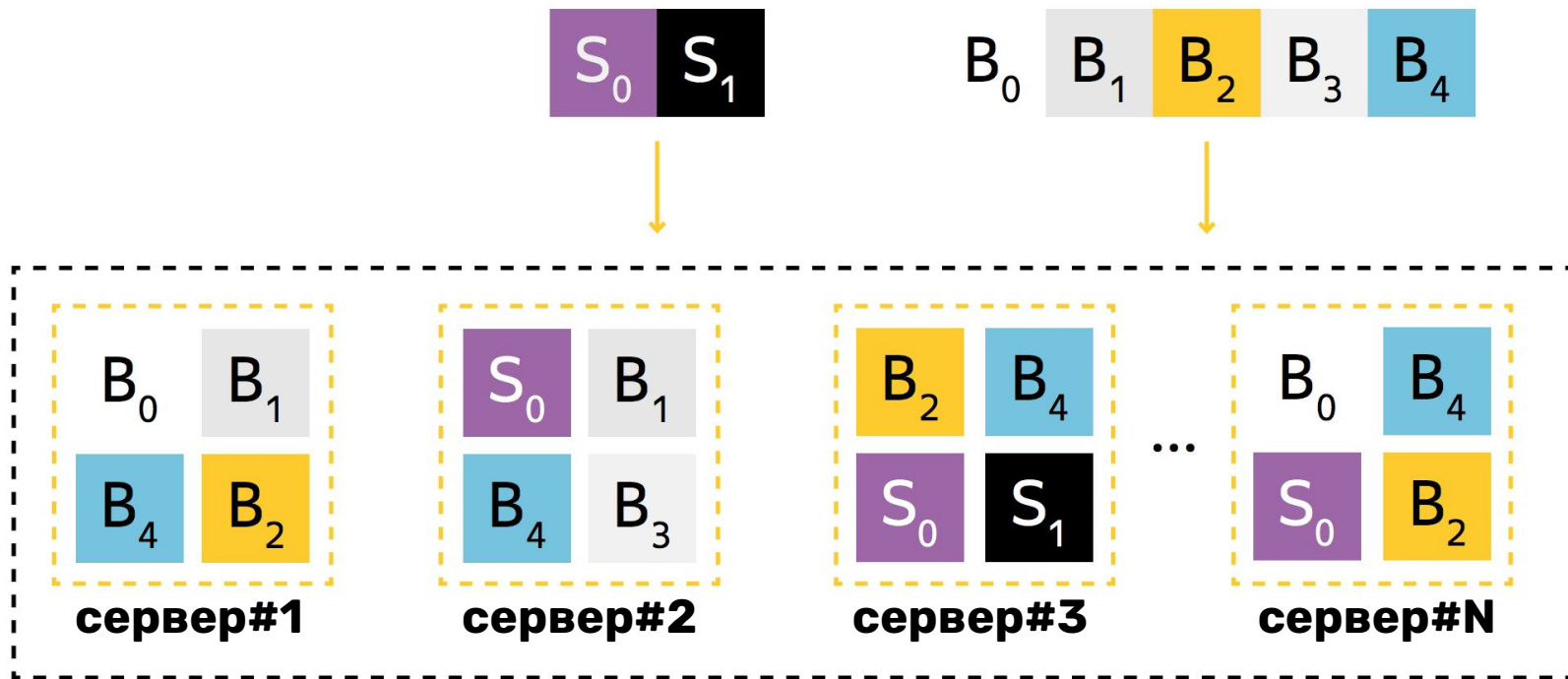


Q&A

Как храним файлы размером 2 TB и
10 GB в кластере?



Распределенная ФС





BIGDATA
TEAM

Принципы GFS



- ▶ Поломки компонент - это норма (используем репликацию)



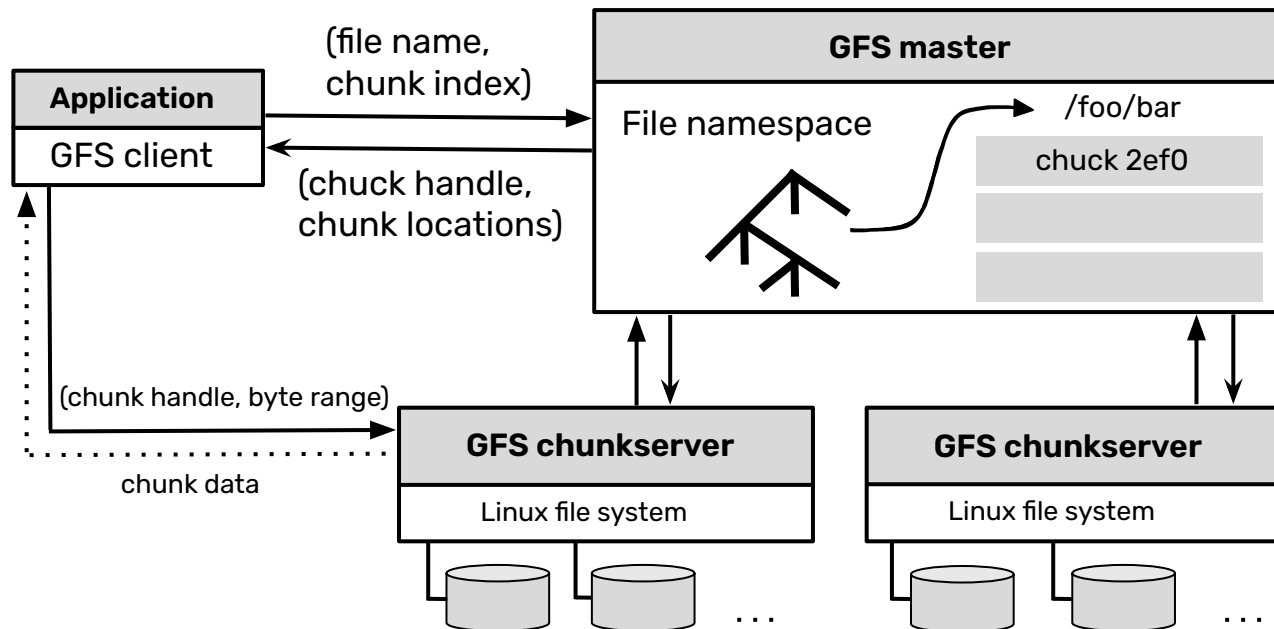
- ▶ Поломки компонент - это норма (используем репликацию)
- ▶ Равномерная утилизация компонент кластера



- ▶ Поломки компонент - это норма (используем репликацию)
- ▶ Равномерная утилизация компонент кластера
- ▶ Семантика write-once-read-many



GFS с высоты птичьего полета

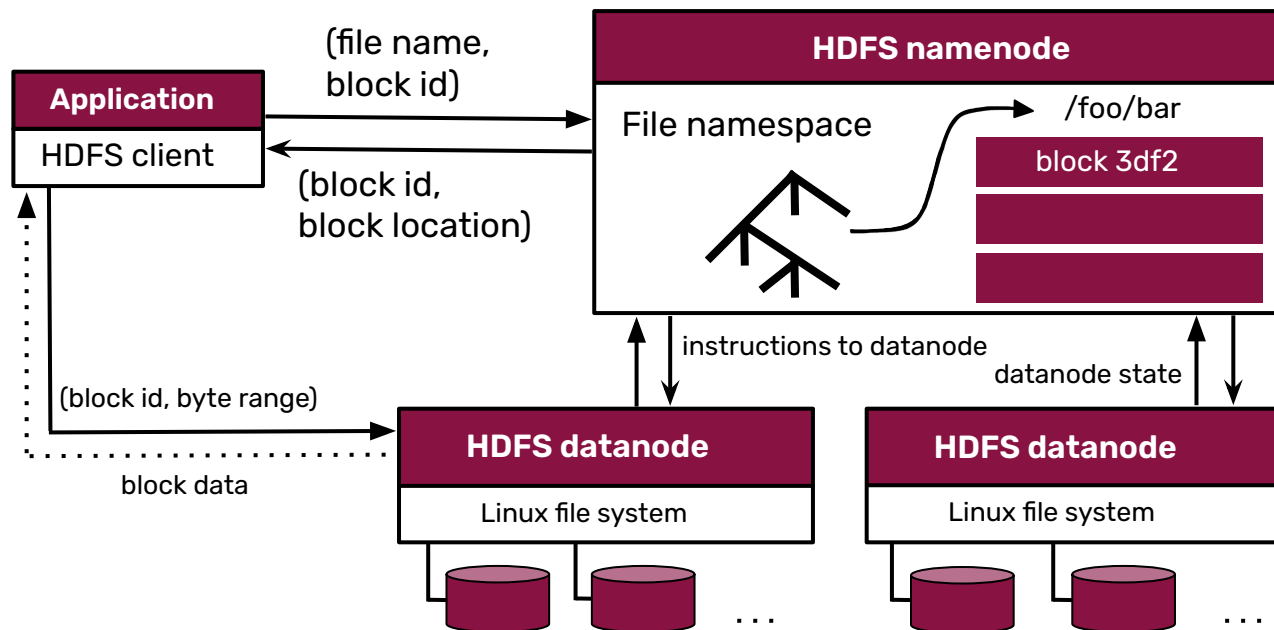


GFS

..... ➤ Data messages
————➤ Control messages



HDFS с высоты птичьего полета



HDFS



**BIGDATA
TEAM**

Резюме



Вы должны уметь:

- ▶ Донести разницу scale-up vs scale-out



Вы должны уметь:

- ▶ Донести разницу scale-up vs scale-out
- ▶ Перечислить три ключевых принципа построения распределенных файловых систем



Вы должны уметь:

- ▶ Донести разницу scale-up vs scale-out
- ▶ Перечислить три ключевых принципа построения распределенных файловых систем
- ▶ Рисовать архитектуру HDFS



Вы должны уметь:

- ▶ Донести разницу scale-up vs scale-out
- ▶ Перечислить три ключевых принципа построения распределенных файловых систем
- ▶ Рисовать архитектуру HDFS
- ▶ Любить Yahoo!



Yahoo!