



Чтение и запись данных в HDFS

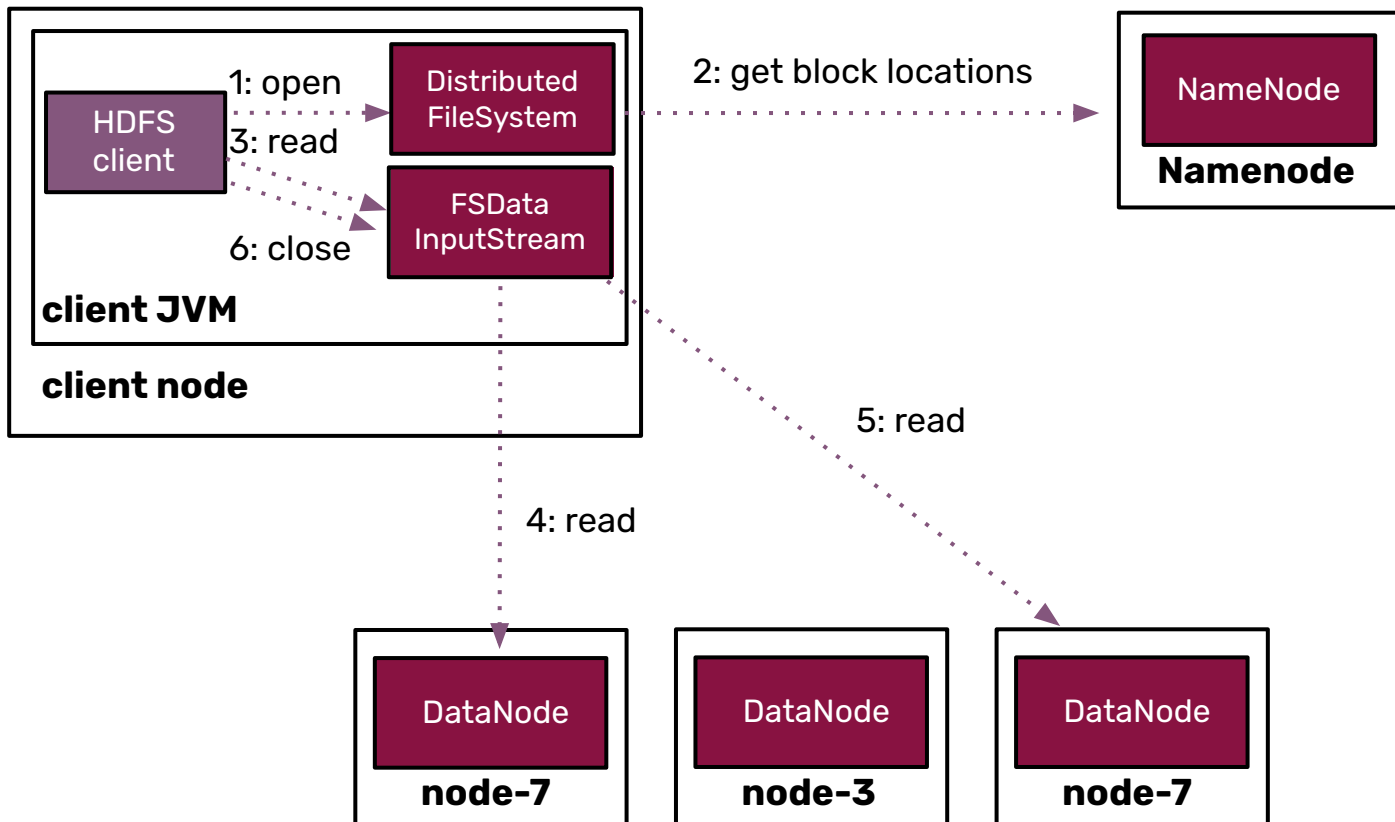
Драль Алексей, study@bigdatateam.org

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

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

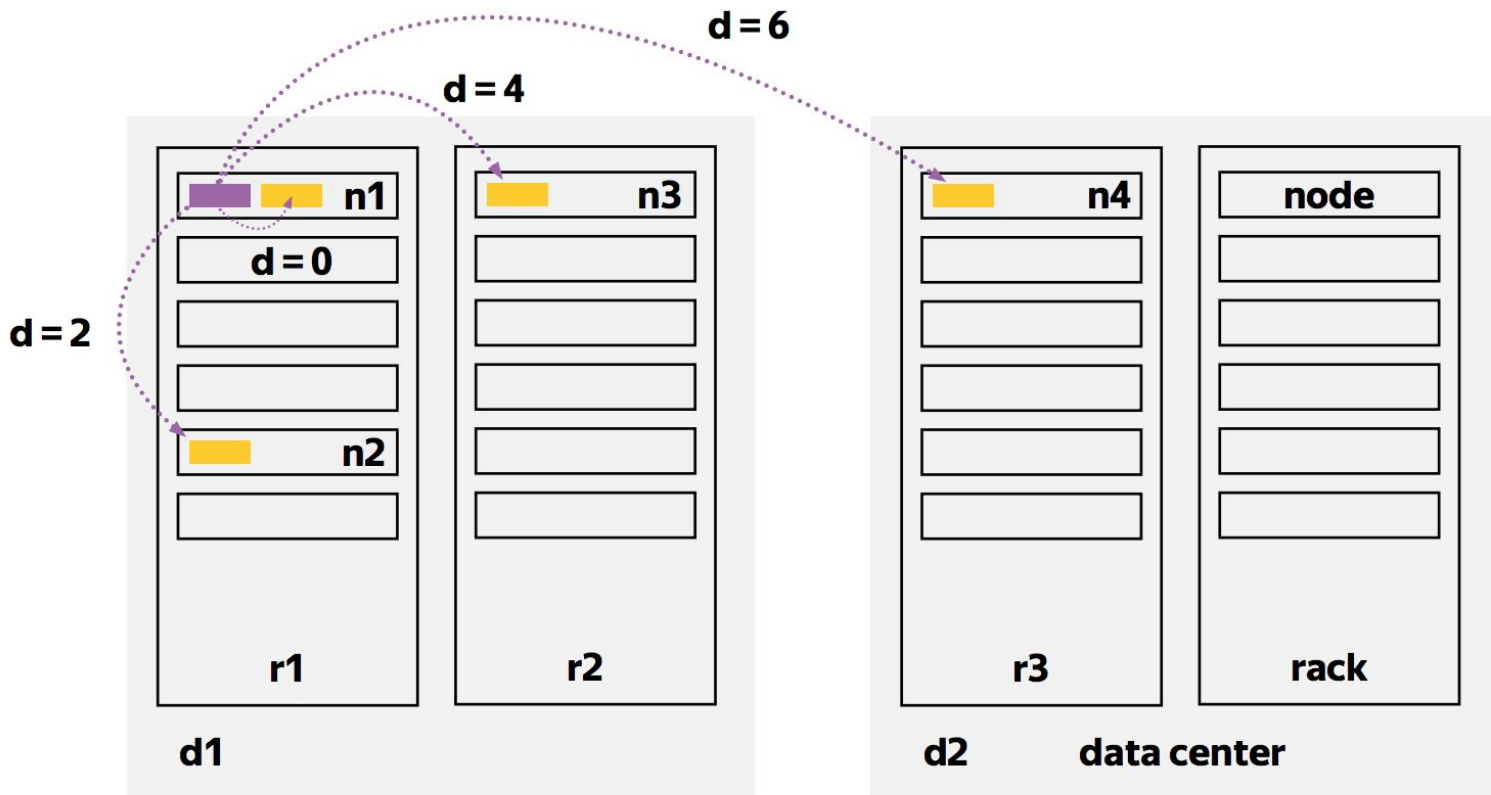


Чтение данных в HDFS



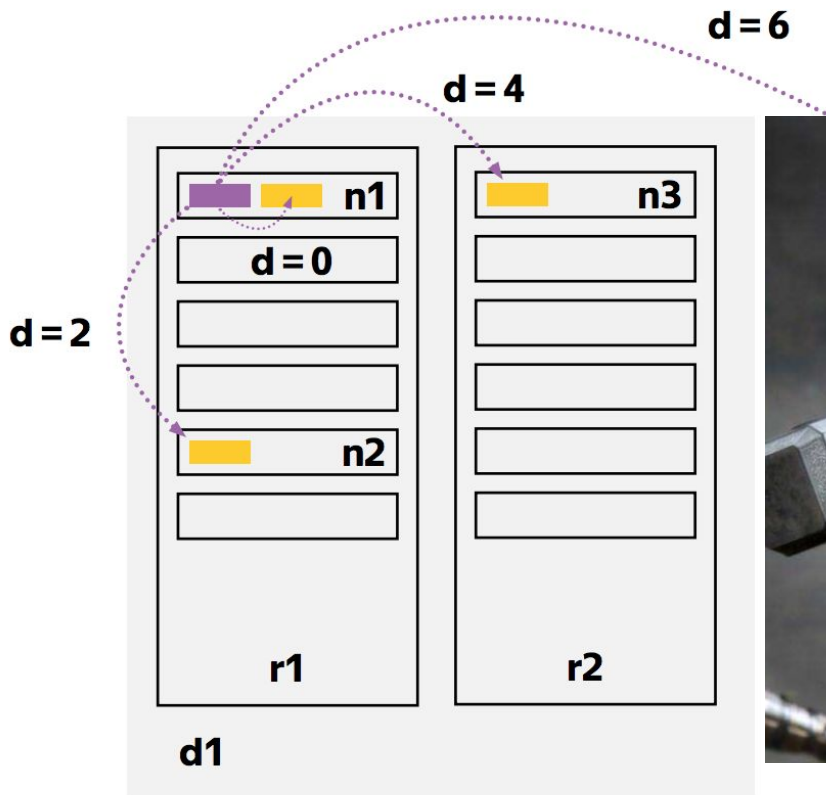


метрика близости HDFS



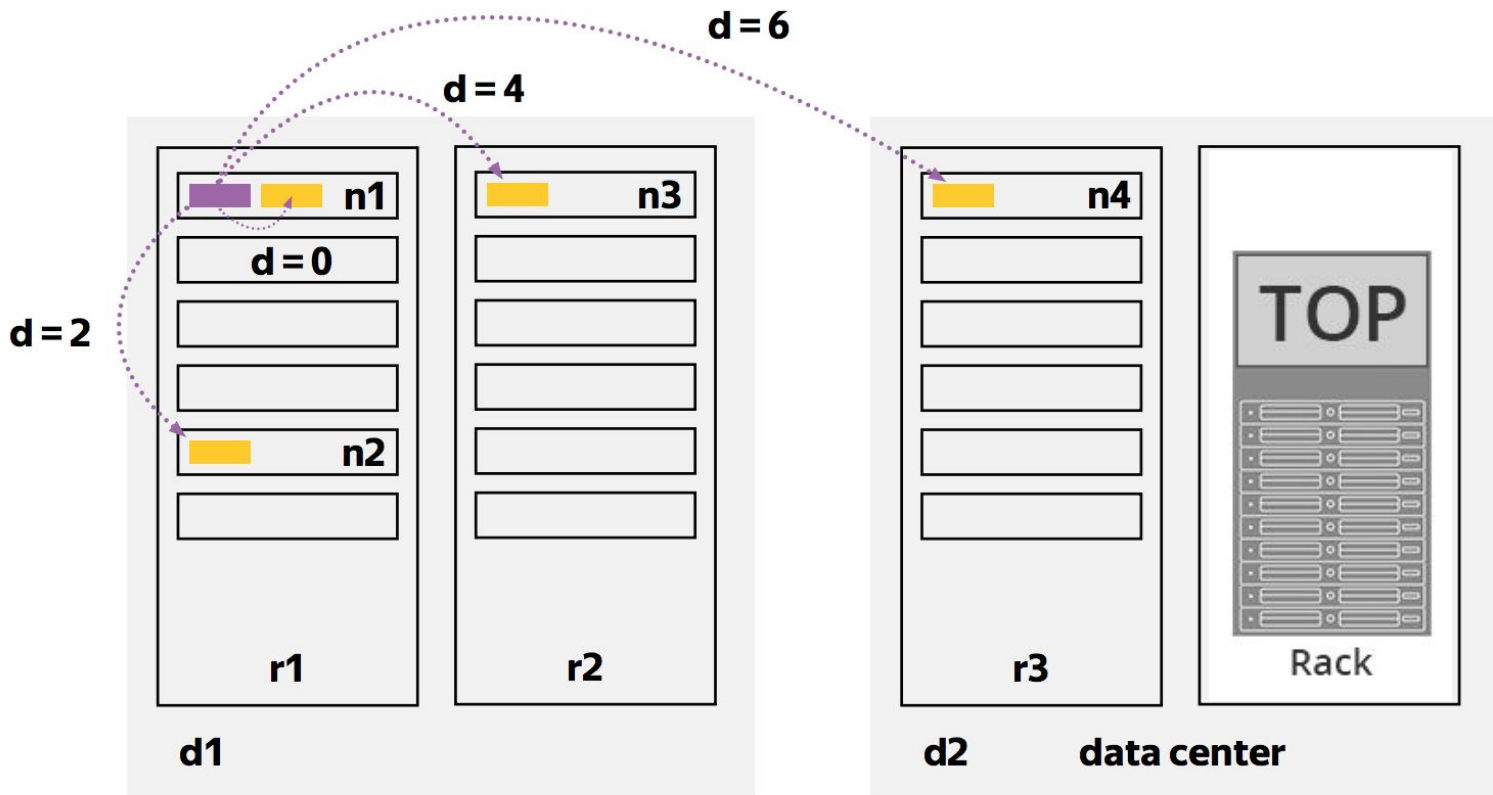


метрика близости HDFS



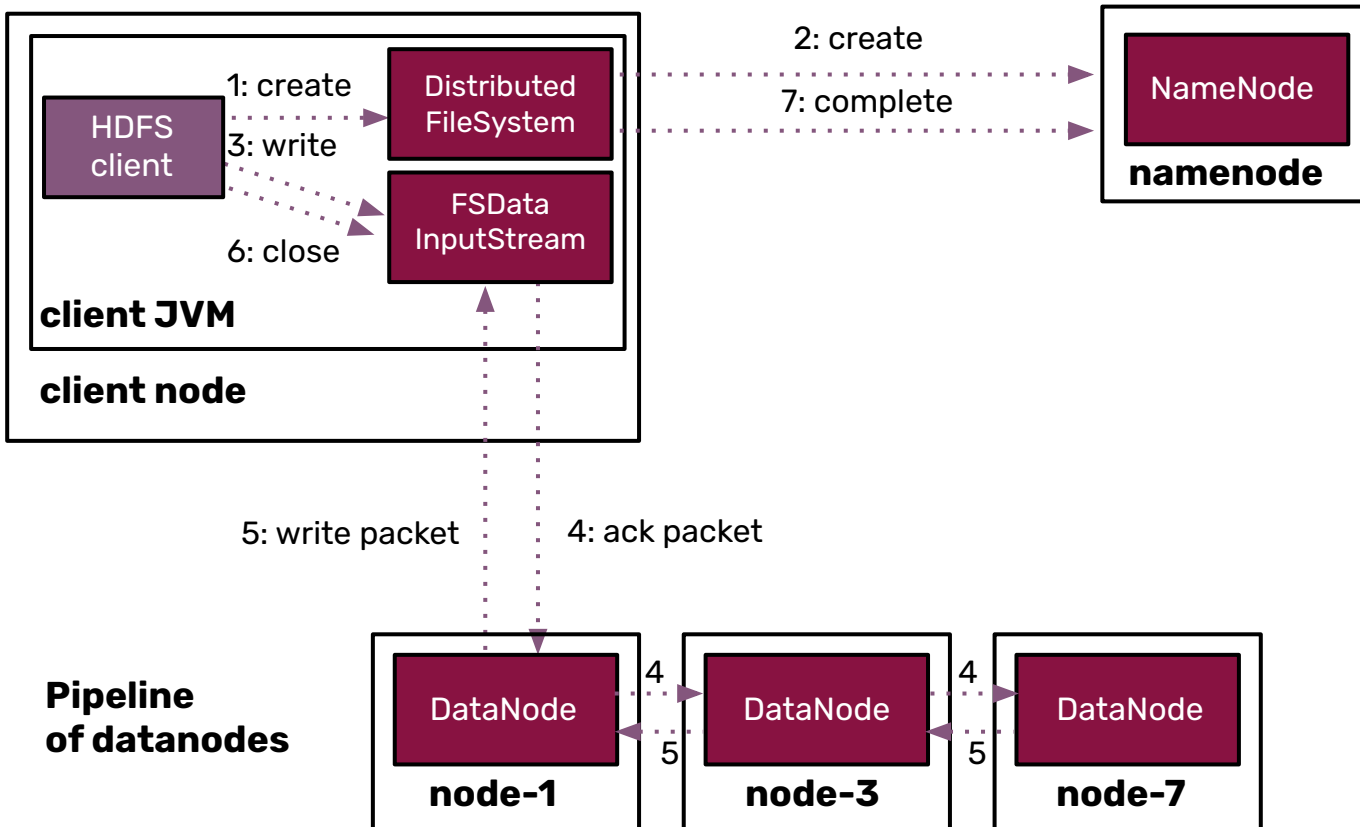


метрика близости HDFS





Запись данных в HDFS



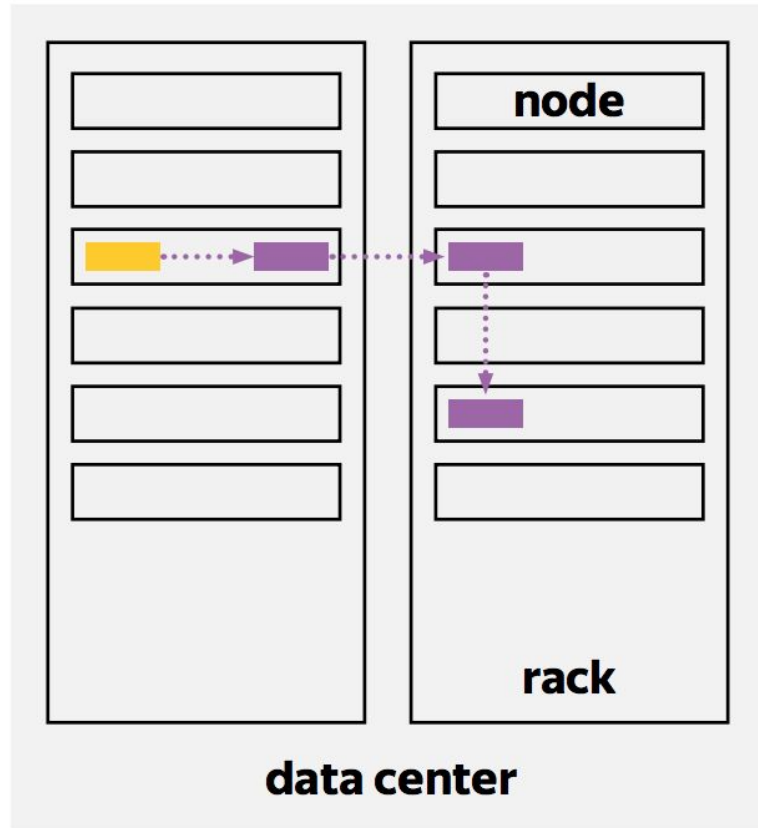


Q&A

Как расположить реплики с учетом сетевой топологии кластера?

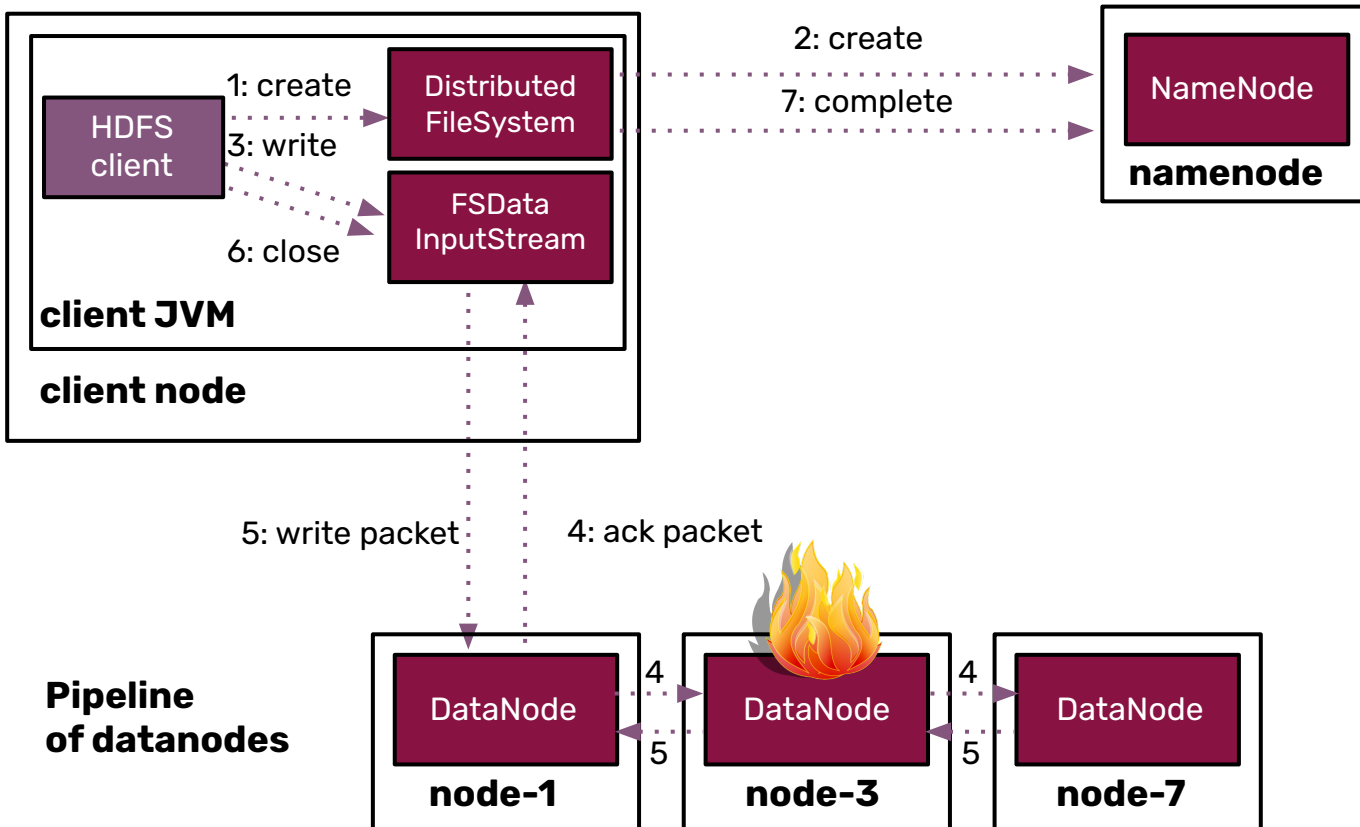


Выбор расположения реплик HDFS





Запись данных в HDFS v.2





**BIGDATA
TEAM**

А что если..?

WE NEED TO GO



DEEPER



Вы должны уметь:

- ▶ Считать сетевые соединения на диаграмме



Вы должны уметь:

- ▶ Считать сетевые соединения на диаграмме
- ▶ Дискутировать на тему стратегий выбора расположения реплик



Вы должны уметь:

- ▶ Считать сетевые соединения на диаграмме
- ▶ Дискутировать на тему стратегий выбора расположения реплик
- ▶ Считать расстояние в HDFS и организовывать Datanode в пайплайн



Вы должны уметь:

- ▶ Считать сетевые соединения на диаграмме
- ▶ Дискутировать на тему стратегий выбора расположения реплик
- ▶ Считать расстояние в HDFS и организовывать Datanode в пайплайн
- ▶ Минимизировать время простоя кластера из-за недоступности реплик (а также любить Yahoo! ❤️)