Hadoop Ecosystem and Hadoop Core : HDFS

Boyu Diao

2016.05.27

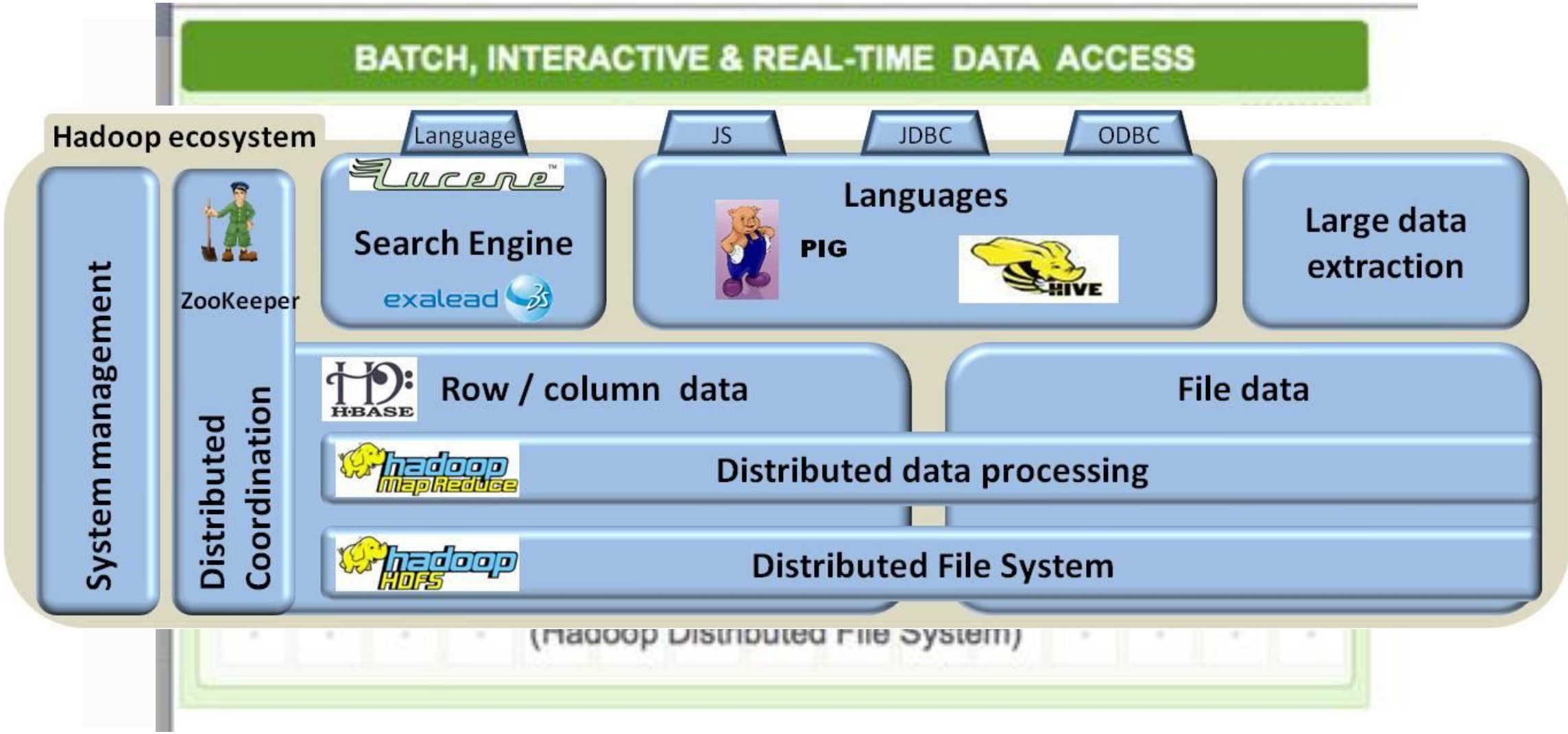
Outlines

- Recall Hadoop Ecosystem and Hadoop Core
- Some Concepts in Distributed System
- HDFS Architecture
- HDFS Read and Write
- HDFS Shell and API

Outlines

- Recall Hadoop Ecosystem and Hadoop Core
- Some Concepts in Distributed System
- HDFS Architecture
- HDFS Read and Write
- HDFS Shell and API

Recall Hadoop Ecosystem



Recall Hadoop Ecosystem

- **Hadoop Ecosystem Table** https://hadoopecosystemtable.github.io/

Categories	Examples
Distributed Filesystem	HDFS, GlusterFS
Distributed Programming	MapReduce / Spark /Storm
SQL-On-Hadoop	HIVE / Drill
NoSQL Databases	Hbase / Cassandra
Data Ingestion	Sqoop / Flume
Service Programming	Thrift / Avor / Zookeeper
Machine Learning	Mahout

Recall Hadoop Core

Apache Hadoop (pronunciation: /həˈduːp/) is an open-source software framework for distributed storage and distributed processing of very large data sets on computer clusters built from commodity hardware. All the

Components	Introduction
HDFS	designed to reliably store very large files across machines in a large cluster.
YARN	The fundamental idea of YARN is to split up the functionalities of resource management and job scheduling/monitoring into separate daemons.
MapReduce	MapReduce is a programming model and an associated implementation for processing and generating large data sets with a parallel, distributed algorithm on a cluster.

Outlines

- Recall Hadoop Ecosystem and Hadoop Core
- Some Concepts in Distributed System
- HDFS Architecture
- HDFS Read and Write
- HDFS Shell and API

Some Concepts in Distributed System

- Copy (副本)
  - replica

- Copy consistency(一致性)
  - strong consistency
  - week consistency
  - eventual consistency
  - Monotonic(单调) consistency

Prototype	Copy consistency
HDFS	Eventual Consistency
Hbase	Strong Consistency

Some Concepts in Distributed System

- Data distribution
  - Hashing
  - Data Range
  - Data Chunk(块)
  - Consistency Hashing
- Copy Control Protocol
  - Centralized
    - primary-secondary
  - Decentralized
    - Paxos

Prototype	Data distribution
HDFS	Data Chunk
Cassandra	Consistence Hashing
HBase	Data Range
Big Pipe	Hashing

Prototype	CC Protocol
HDFS	Primary-Secondary
Megastore	Paxos

Some Concepts in Distributed System

- Copy consistency(一致性)
  - strong consistency
    - 当更新操作完成之后，任何多个后续进程或者线程的访问都会返回最新的更新过的值。
  - week consistency
    - 系统并不保证续进程或者线程的访问都会返回最新的更新过的值。但保证每个副本都有相同的读写时序。
  - eventual consistency
    - 弱一致性的特定形式。系统保证在没有后续更新的前提下，系统最终返回上一次更新操作的值。
  - Monotonic(单调) consistency

Outlines

- Recall Hadoop Ecosystem and Hadoop Core
- Some Concepts in Distributed System
- HDFS Architecture
- HDFS Read and Write
- HDFS Shell and API

HDFS Architecture

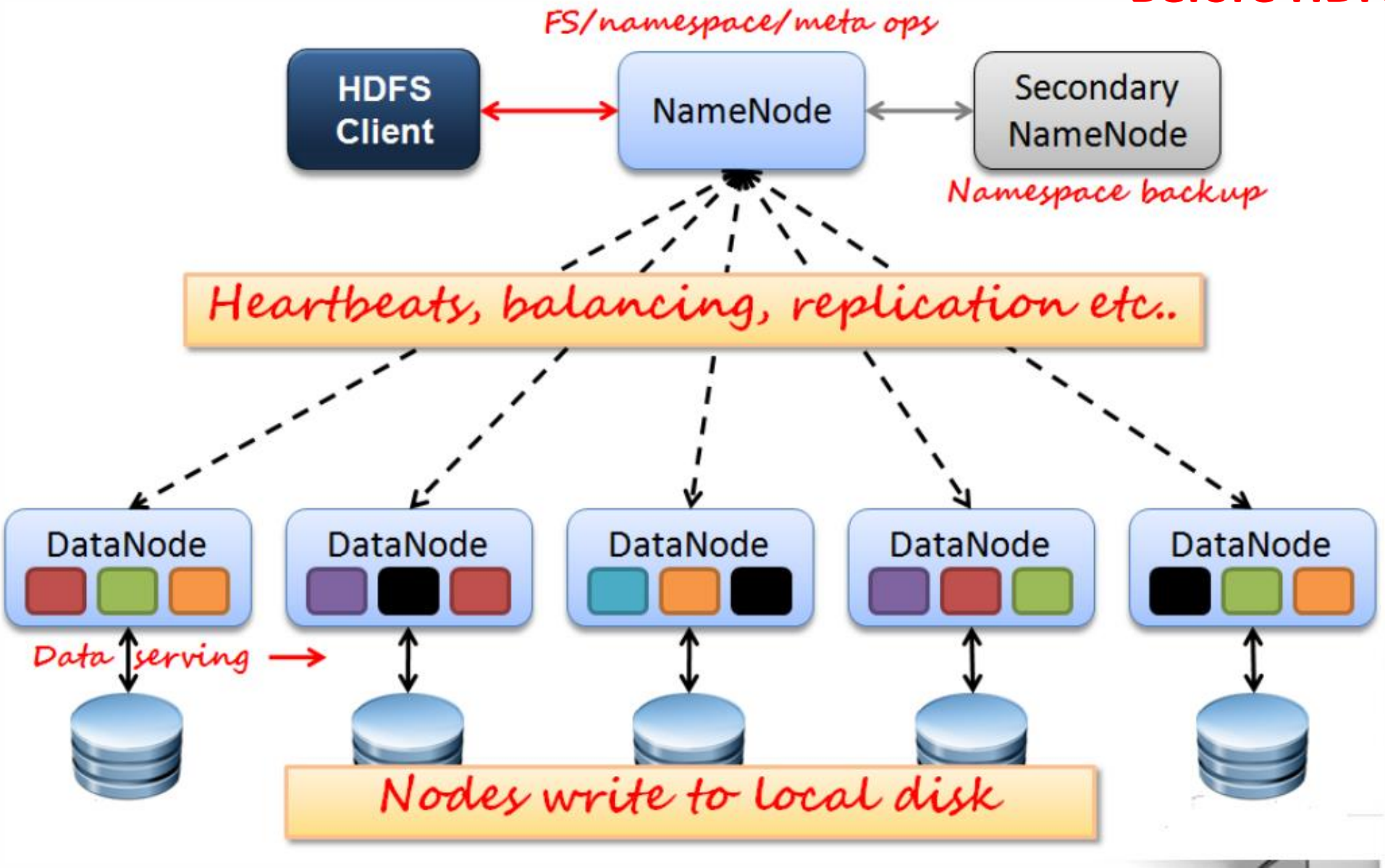
- HDFS is designed to reliably store very large files across machines in a large cluster.

Problems	Methods
Copy / Replica	Multiple
Copy Consistency	Eventual Consistency
Data Distribution	Data Chunk
Copy Control Protocol	Primary-Secondary

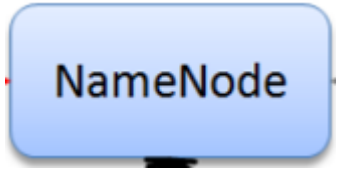
Ghemawat, S.; Gobioff, H.; Leung, S. T. (2003). "The Google file system". *Proceedings of the nineteenth ACM Symposium on Operating Systems Principles - SOSP '03*

HDFS Architecture

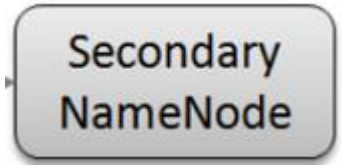
Before HDFS 2.0



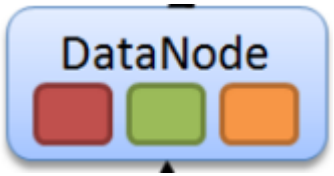
HDFS Architecture



- Master
- 管理命名空间镜像文件（fsimage）和操作日志（fsedits）
- 管理数据块映射信息
- 管理副本策略
- 处理Client的请求



- NameNode的热备
- 定期同步fsimage, fsedits
- 当NameNode出现故障时，帮助NameNode快速恢复



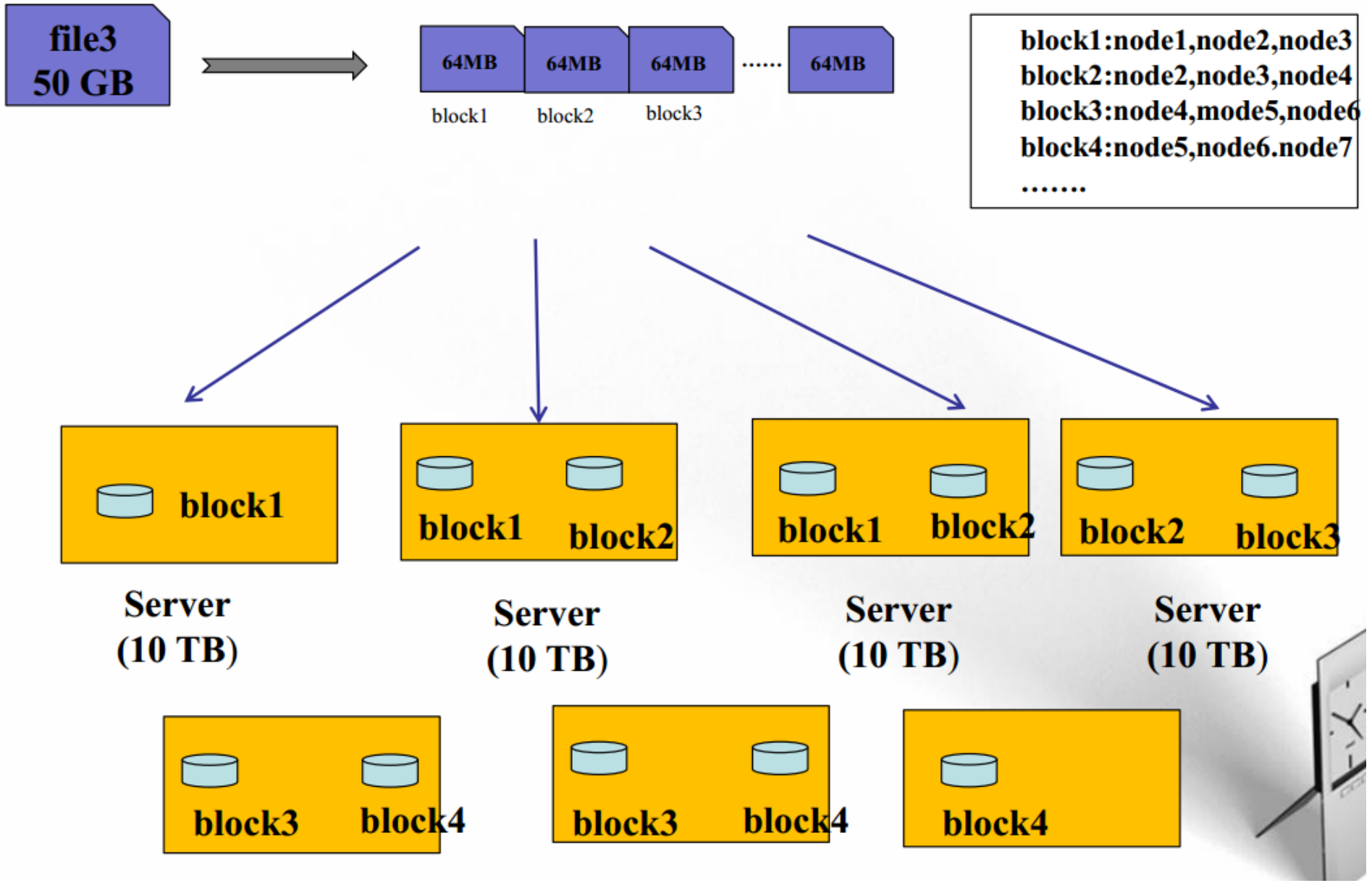
- Slave
- 负责数据在本地的读写，存储

- 文件切分
- 与NameNode交互获取副本策略
- 与DataNode交互，进行文件读写
- 管理HDFS集群

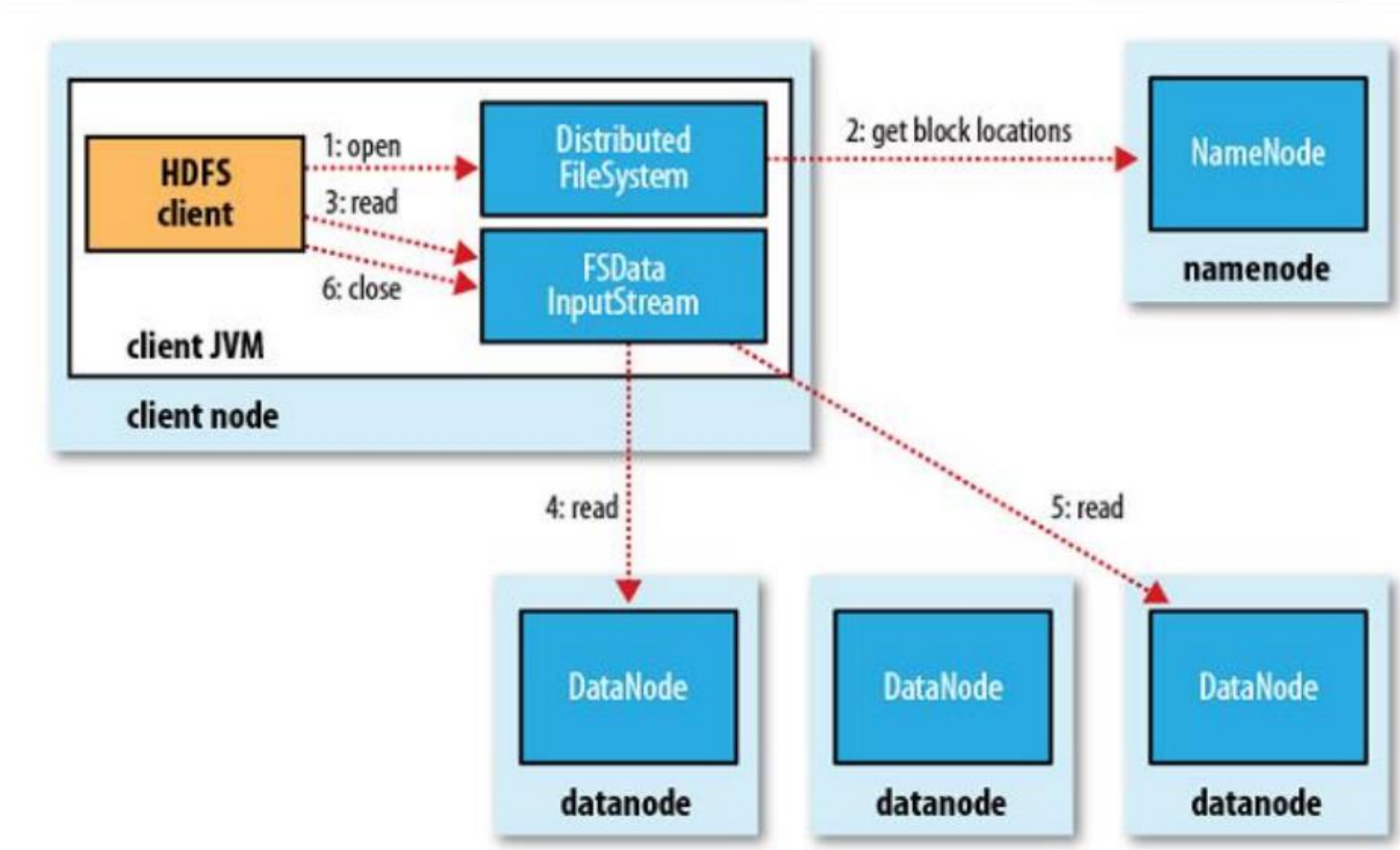
Outlines

- Recall Hadoop Ecosystem and Hadoop Core
- Some Concepts in Distributed System
- HDFS Architecture
- HDFS Read and Write
- HDFS Shell and API

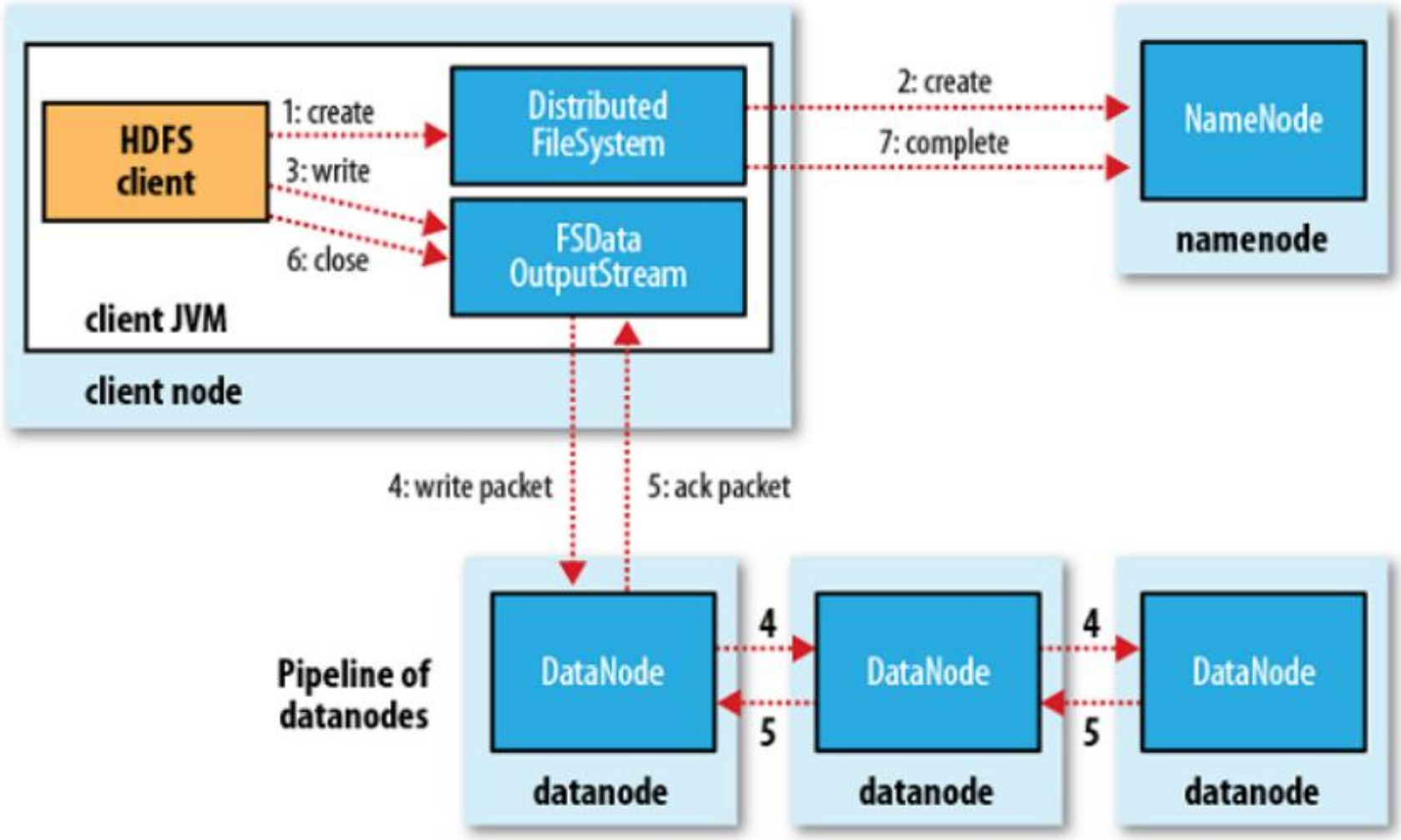
HDFS Read and Write



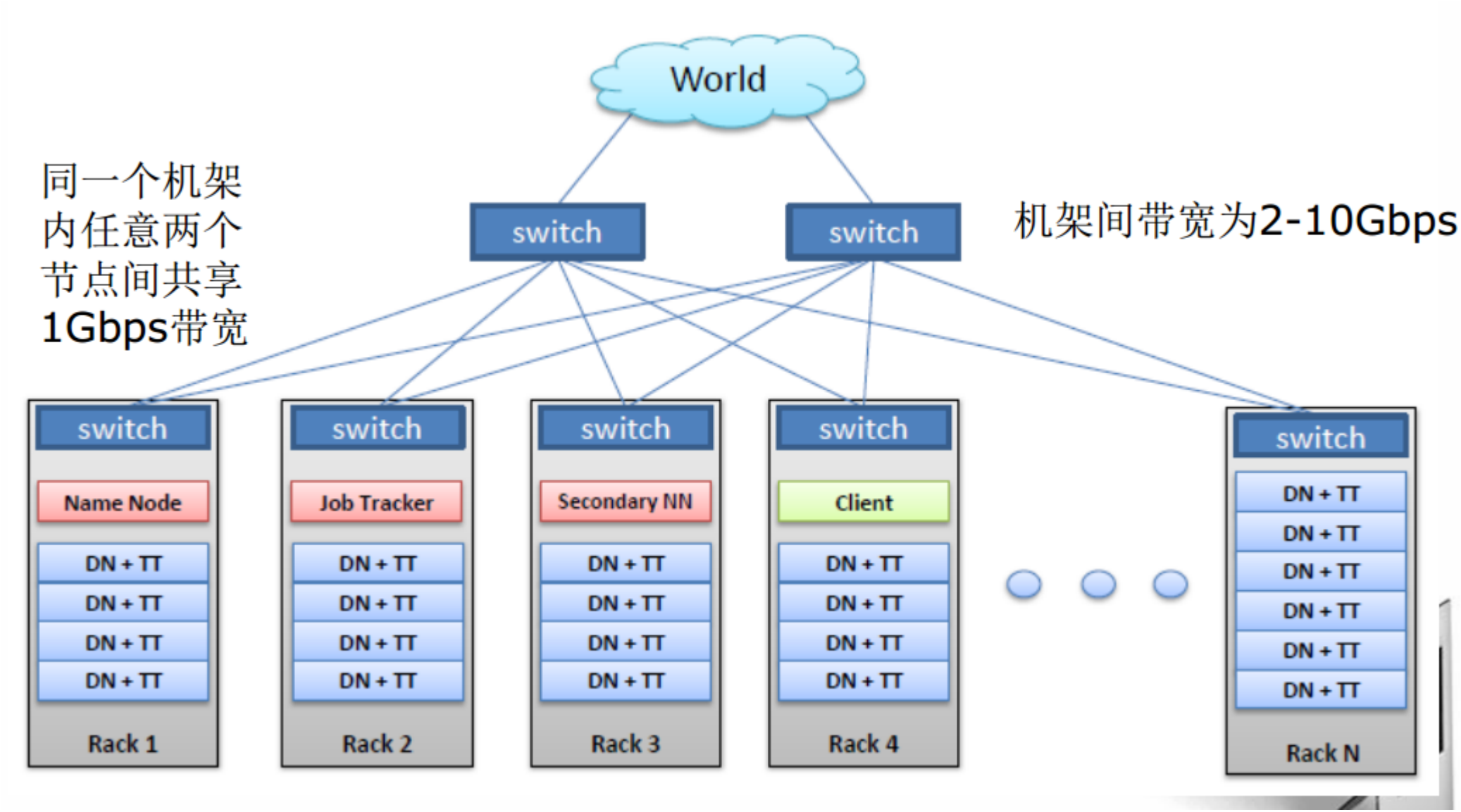
HDFS Read and Write



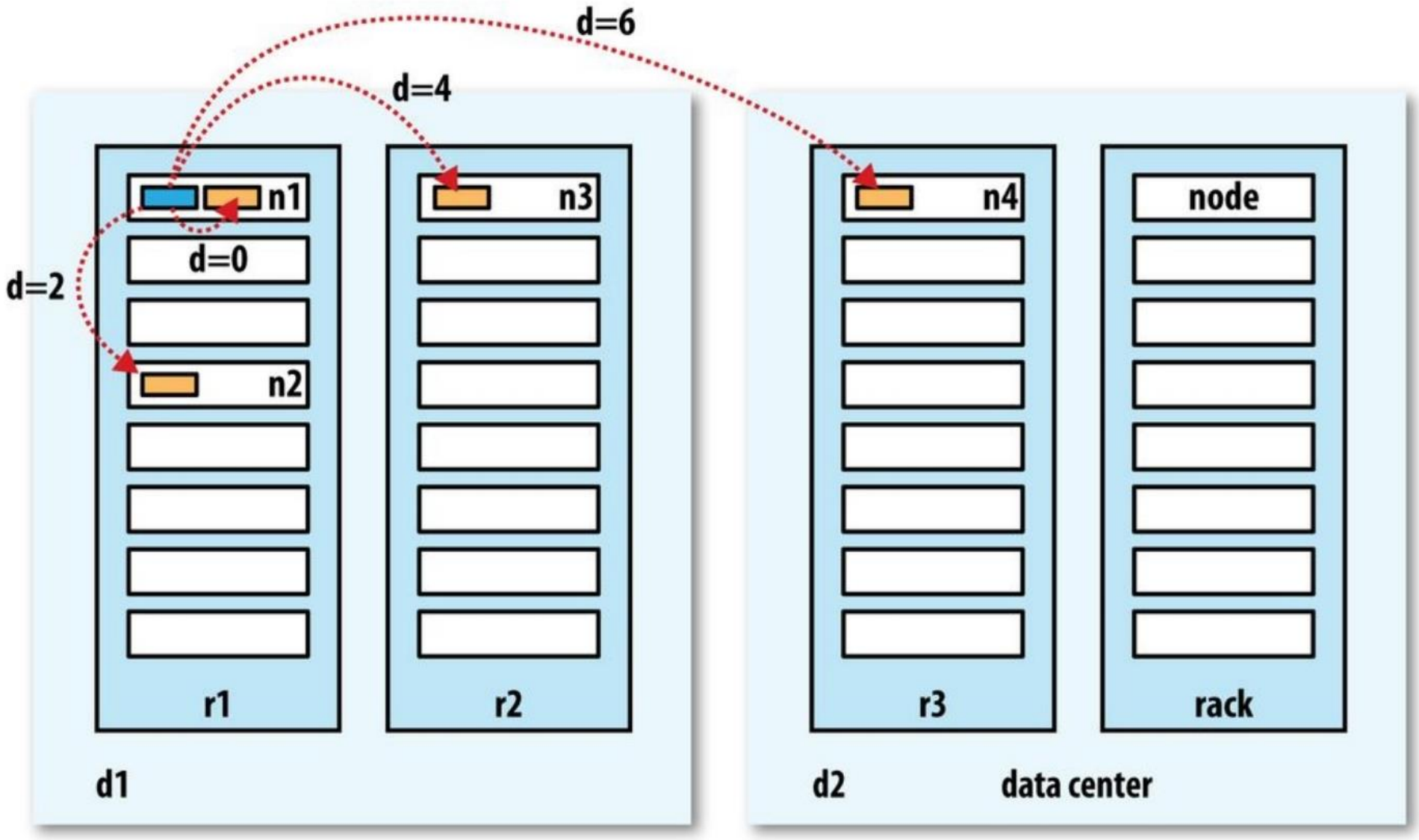
HDFS Read and Write



HDFS Read and Write



HDFS Read and Write

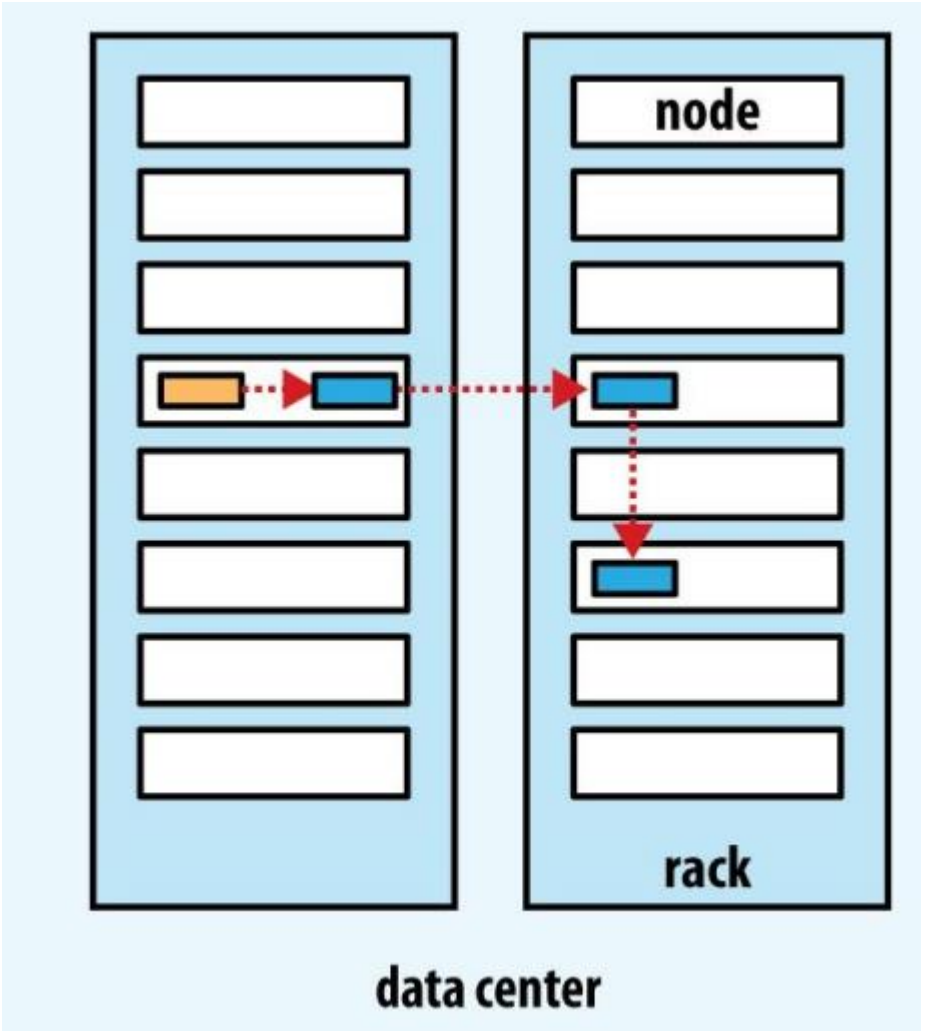


HDFS Read and Write

- REPLICA PLACEMENT
  - 1. Same node as Client
  - 2. Node in Different Rack
  - 3. Another node in this Rack

  Random pick the first one

  If client is not a cluster node



Outlines

- Recall Hadoop Ecosystem and Hadoop Core
- Some Concepts in Distributed System
- HDFS Architecture
- HDFS Read and Write
- HDFS Shell and API

HDFS Shell and

- Shell命令
- Hadoop fs –xxx
- Hadoop dfsadmin –xx
- Hadoop fsck –xxx

```
dongxicheng@dongxicheng-laptop:~/hadoop/hadoop-0.20.2-cdh3u6$ bin/hadoop fs
Usage: java FsShell
           [-ls <path>]
           [-lsr <path>]
           [-df [<path>]]
           [-du <path>]
           [-dus <path>]
           [-count[-q] <path>]
           [-mv <src> <dst>]
           [-cp <src> <dst>]
           [-rm [-skipTrash] <path>]
           [-rmr [-skipTrash] <path>]
           [-expunge]
           [-put <localsrc> ... <dst>]
           [-copyFromLocal <localsrc> ... <dst>]
           [-moveFromLocal <localsrc> ... <dst>]
           [-get [-ignoreCrc] [-crc] <src> <localdst>]
           [-getmerge <src> <localdst> [addnl]]
           [-cat <src>]
           [-text <src>]
           [-copyToLocal [-ignoreCrc] [-crc] <src> <localdst>]
           [-moveToLocal [-crc] <src> <localdst>]
           [-mkdir <path>]
           [-setrep [-R] [-w] <rep> <path/file>]
           [-touchz <path>]
           [-test -[ezd] <path>]
           [-stat [format] <path>]
           [-tail [-f] <file>]
           [-chmod [-R] <MODE[,MODE]... | OCTALMODE> PATH...]
           [-chown [-R] [OWNER][:[GROUP]] PATH...]
```

HDFS Shell and API

- Hadoop dfsadmin –xxx

```
dongxicheng@dongxicheng-laptop:~/hadoop/hadoop-0.20.2-cdh3u6$ bin/hadoop dfsadmin
Usage: java DFSAdmin
           [-report]
           [-safemode enter | leave | get | wait]
           [-saveNamespace]
           [-refreshNodes]
           [-finalizeUpgrade]
           [-upgradeProgress status | details | force]
           [-metasave filename]
           [-refreshServiceAcl]
           [-refreshUserToGroupsMappings]
           [-refreshSuperUserGroupsConfiguration]
           [-setQuota <quota> <dirname>...<dirname>]
           [-clrQuota <dirname>...<dirname>]
           [-setSpaceQuota <quota> <dirname>...<dirname>]
           [-clrSpaceQuota <dirname>...<dirname>]
           [-setBalancerBandwidth <bandwidth in bytes per second>]
           [-help [cmd]]
```

HDFS Shell and API

- Hadoop fsck –xxx

```
dongxicheng@dongxicheng-laptop:~/hadoop/hadoop-0.20.2-cdh3u6$ bin/hadoop fsck /home/dongxicheng/README.txt -files -blocks -locations
FSCK started by dongxicheng (auth:SIMPLE) from /127.0.1.1 for path /home/dongxicheng/README.txt at Sat Sep 21 22:39:52 CST 2013
/home/dongxicheng/README.txt 1366 bytes, 1 block(s):  OK
0. blk_7247523855706538504_1001 len=1366 repl=1 [127.0.0.1:50010]

Status: HEALTHY
 Total size:    1366 B
 Total dirs:    0
 Total files:   1
 Total blocks (validated):      1 (avg. block size 1366 B)
 Minimally replicated blocks:   1 (100.0 %)
 Over-replicated blocks:        0 (0.0 %)
 Under-replicated blocks:       0 (0.0 %)
 Mis-replicated blocks:         0 (0.0 %)
 Default replication factor:    1
 Average block replication:     1.0
 Corrupt blocks:                0
 Missing replicas:              0 (0.0 %)
 Number of data-nodes:          1
 Number of racks:               1
FSCK ended at Sat Sep 21 22:39:52 CST 2013 in 1 milliseconds
```

HDFS Shell and API

- Hadoop API https://hadoop.apache.org/docs/stable/api/

- HTTP

Summary I

- Recall Hadoop Ecosystem and Hadoop Core
- Some Concepts in Distributed System
- HDFS Architecture
- HDFS Read and Write
- HDFS Shell and API

Source: Wikipedia (Japanese rock garden)