

## 7、HDFS 的文件写入过程

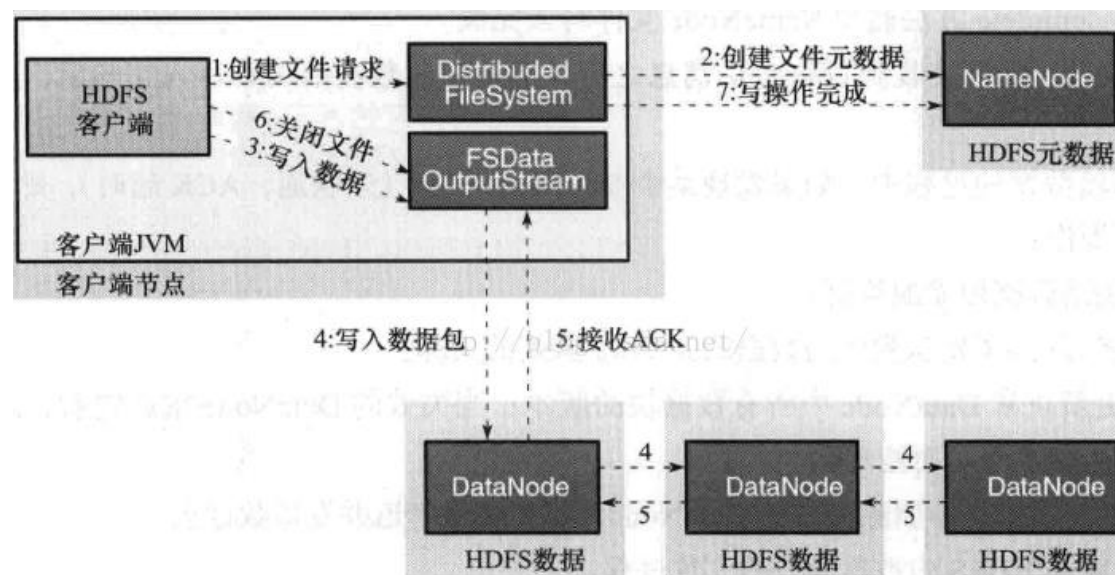


图 3-7 HDFS 数据写入过程

### 详细步骤解析：

- 1、 client发起文件上传请求，通过RPC与NameNode建立通讯，NameNode检查目标文件是否已存在，父目录是否存在，返回是否可以上传；
- 2、 client请求第一个block该传输到哪些DataNode服务器上；
- 3、 NameNode根据配置文件中指定的备份数量及机架感知原理进行文件分配，返回可用的DataNode的地址如：A，B，C；

注：Hadoop在设计时考虑到数据的安全与高效，数据文件默认在HDFS上存放三份，存储策略为本地一份，同机架内其它某一节点上一份，不同机架的某一节点上一份。

4、 client请求3台DataNode中的一台A上传数据（本质上是一个RPC调用，建立pipeline），A收到请求会继续调用B，然后B调用C，将整个pipeline建立完成，后逐级返回client；

5、 client开始往A上传第一个block（先从磁盘读取数据放到一个本地内存缓存），以packet为单位（默认64K），A收到一个packet就会传给B，B传给C；A每传一个packet会放入一个应答队列等待应答。

6、 数据被分割成一个个packet数据包在pipeline上依次传输，在pipeline反方向上，逐个发送ack（命令正确应答），最终由pipeline中第一个DataNode节点A将pipelineack发送给client；

7、 当一个block传输完成之后，client再次请求NameNode上传第二个block到服务器。