hadoop的一大基本原则是移动计算的开销要比移动数据的开销小，因此Hadoop通常尽量移动计算到拥有数据的节点上，这使得DFSClient和提供数据的datanode经常在一个节点上，这样就产生很多Local Reads。

最初设计的时候，Local Reads和remote Reads处理方式相同，先由datanode读取数据，然后再通过RPC把数据传给DFSclient，需要datanode在中间做一次中转。尽然DFSClient和数据都在一个机器上面，很自然的想法是让DFSClient绕开Datanode自己去读取数据。

在Linux中， Unix Domain Socket技术（进程间的通信方式），使得同一机器两个进程能以Socket的方式通信。HDFS-347，实现了DFSClient和Datanode两个进程以Socket方式通信，DFSClient可以直接读取Datanode数据目录中的某个文件，这样提高了Hadoop的运行速度。

使用short-circuit的配置如下：

<property>

<name>dfs.client.read.shortcircuit</name>

<value>true</value>

</property>

<property>

<name>dfs.domain.socket.path</name>

<value>/var/lib/hadoop-hdfs/dn\_socket</value>

</property>

short-circuit的实现是在libhadoop.so动态库中，源码位置为：../hadoop-com

mon/src/../net/unix/domainsocket.c中

注意：https://www.zybuluo.com/jewes/note/37713中有更详细的介绍。