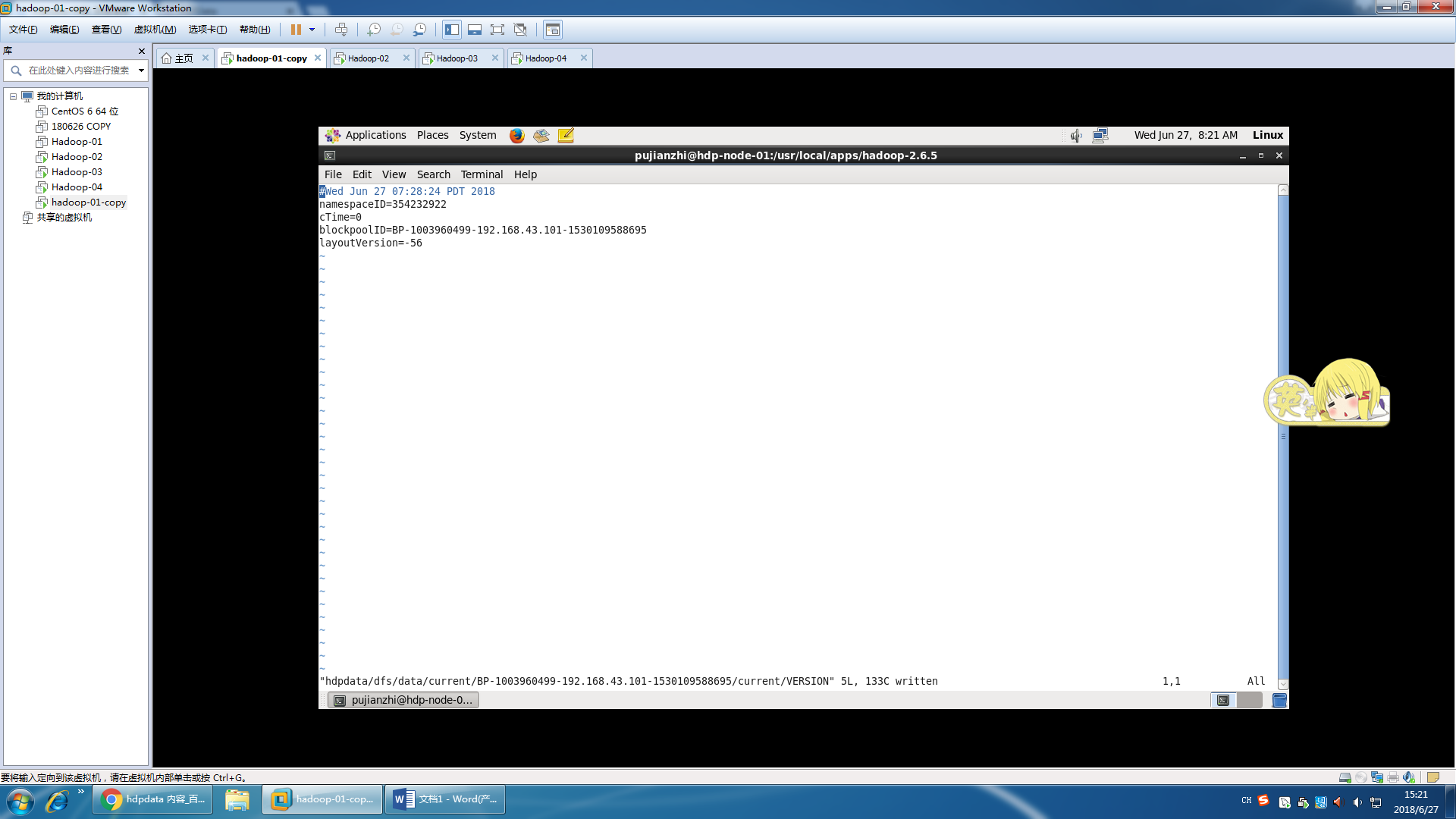
主节点下目录为



Date保存的为各个文件的数据，其后的数字为ID

Blk\_id保存的的是文件的各个块

VERSION保存的为集群ID等重要信息

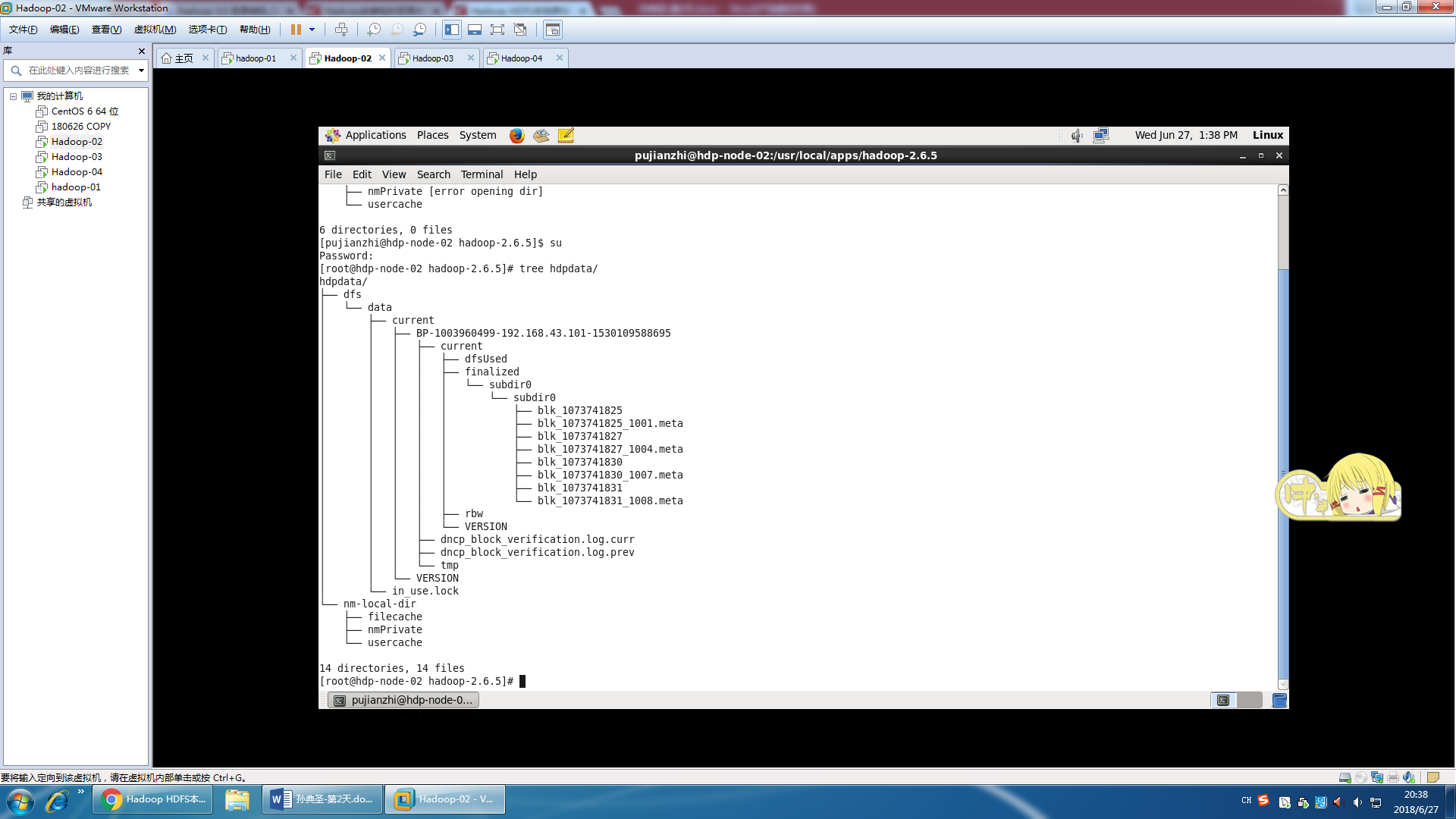


Name下有edits类文件，保存的是修改的记录

它们会被定期的合并到fsimage类文件当中

in\_use.lock防止一台机器同时启动多个Namenode进程导致目录数据不一致

datanode结点的信息为



BP-random integer-NameNode-IP address-creation time

BP代表BlockPool的意思，就是上面Namenode的VERSION中的集群唯一blockpoolID，如果是Federation HDFS，则该目录下有两个BP开头的目录，IP部分和时间戳代表创建该BP的NameNode的IP地址和创建时间戳

VERSION 与Namenode类似，其中storageType是DATA\_NODE

finalized/rbw 这两个目录都是用于实际存储HDFS BLOCK的数据，里面包含许多block\_xx文件以及相应的.meta文件，.meta文件包含了checksum信息。

rbw是“replica being written”的意思，该目录用于存储用户当前正在写入的数据。

dncp\_block\_verification.log 该文件用于追踪每个block最后修改后的checksum值，该文件会定期滚动，滚动后会移到.prev文件

in\_use.lock防止一台机器同时启动多个Datanode进程导致目录数据不一致