

知 在Hadoop系统中，HDFS将文件进行分块，将块作为基本的存储单元。这样可以设计的优势是？

厉梦如 2019-05-19 发表

问题描述

Q：在Hadoop系统中，HDFS将文件进行分块，将块作为基本的存储单元。这样可以设计的优势是？

解决方法

A：

- 1、HDFS可以存储任意的文件而不受到网络中任意单个磁盘大小的限制。理论上文件可以占满集群中所有的磁盘。
- 2、使用块作为存储单元：简化存储子系统，在HDFS中块的大小固定，可以简化存储系统的管理，将块和元数据信息分开存储，也消除了分布式管理文件元数据的复杂性。
- 3、方便容错，有利于数据复制。在HDFS中为了应对损坏的块以及磁盘、机器的故障，数据块会在不同机器上进行复制（一般副本数为3）。