spark基于内存的计算方式

spark提供了python接口

hdfs分布式文件系统