1、归档 tar -zcvf 文件.gz.tar 文件

解压缩tar -zxvf 文件名.gz.tar

2、Hadoop的核心组件包括: HDFS、Yarn、MapReduce

HDFS

目前大量采用的分布式文件系统，是整个大数据应用场景的基础通用文件存储组

Yarn

YARN,分布式资源调度，可以接收计算的任务把它分配到集群各节点处理，相当于大数据操作系统，通用性好，生态支持好；

MapReduce

分布式计算基本计算框架，编程难度高，执行效率低

1. 方法重载(Overload):一个类中定义多个相同名称的方法。  
    要求:  
    (1)、方法名相同  
    (2)、参数列表不同(类型、个数、顺序)  
    (3)、与访问修饰和返回值类型无关  
         
   方法覆盖(Override):父类方法无法满足子类需求，需要进行覆盖。  
    要求:  
    (1)、返回值类型、方法名、形参列表和父类完全一致  
    (2)、修饰符和父类相同或者比父类更宽
2. JDBC的操作步骤?
3. 加载驱动
4. 获取连接
5. 书写sql语句
6. 准备执行语句的对象Statement
7. 处理结果集
8. 关闭资源