10选择，10判断，10填空，6个简答，2个综合

大数据特征选择填空，占比不大

完全分布式在选择填空，这么配置，配置什么，与阶段类似

HDFS考察不多，大分值题，算blocksize，算。。。

Mapreduce重点，框架，基本组成，每个干什么的，作用是什么。不考过程

Conbiner，切片，怎么切，inputformat

排序没有，shuffle过程没有

Yarn基本架构，压缩的种类