**十、海量数据并行处理系统**  
1、MapReduce基础概念、运行框架、调度模型等；  
2、HDFS基础概念、命令行接口、数据序列化、数据流等；  
3、Hadoop伪分布式及完全分布式模型的构建、测试及应用；  
4、生产环境中Hadoop多种常用属性的调优及其注意事项等；  
5、Hadoop生态系统及常用组件说明；  
6、Hadoop调度算法及其选用；  
7、Hadoop硬件选型要点；  
8、NoSQL数据库特性、技术流派及代表性项目；Zookeeper功能、特性、安装配置及应用；列式数据库HBase基础原理、安装配置及其应用；  
9、数据仓库Hive基础原理、安装配置及应用；  
10、讲解flume/scribe/chukwa/kafka日志收集系统其中之一；  
11、使用工具完成Hadoop自动化部署；