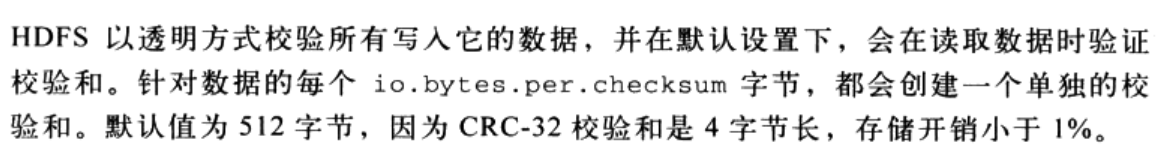
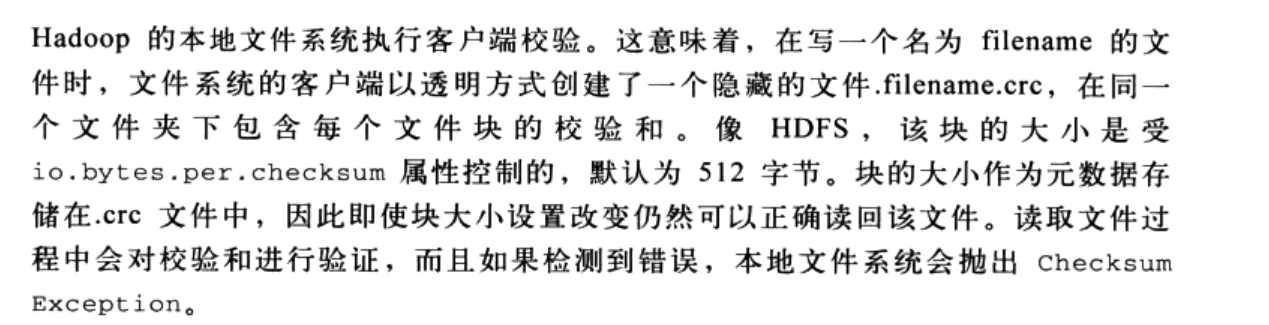
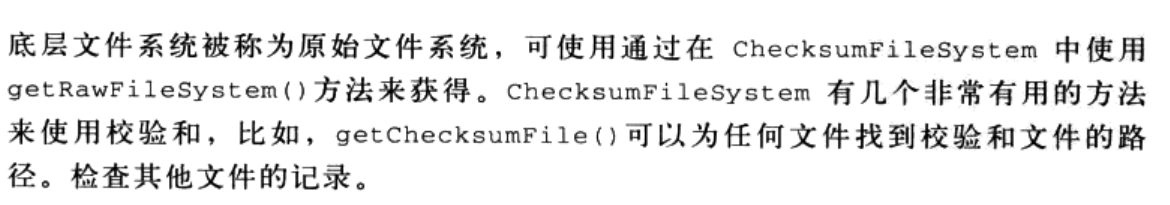
1. hadoop的IO
2. 数据完整性
3. HDFS的数据完整性



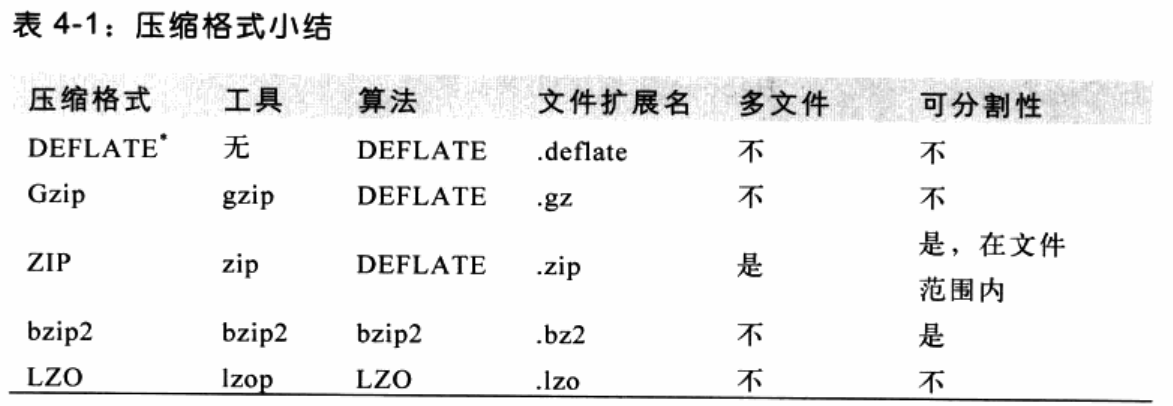
1. 本地文件系统



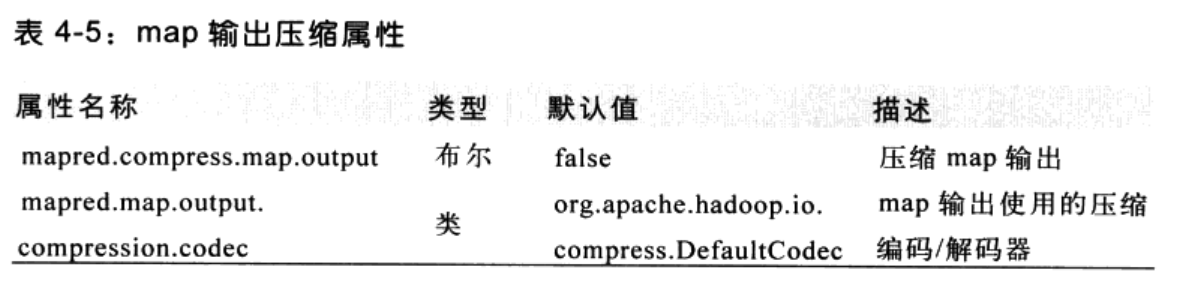
1. ChecksumFileSystem



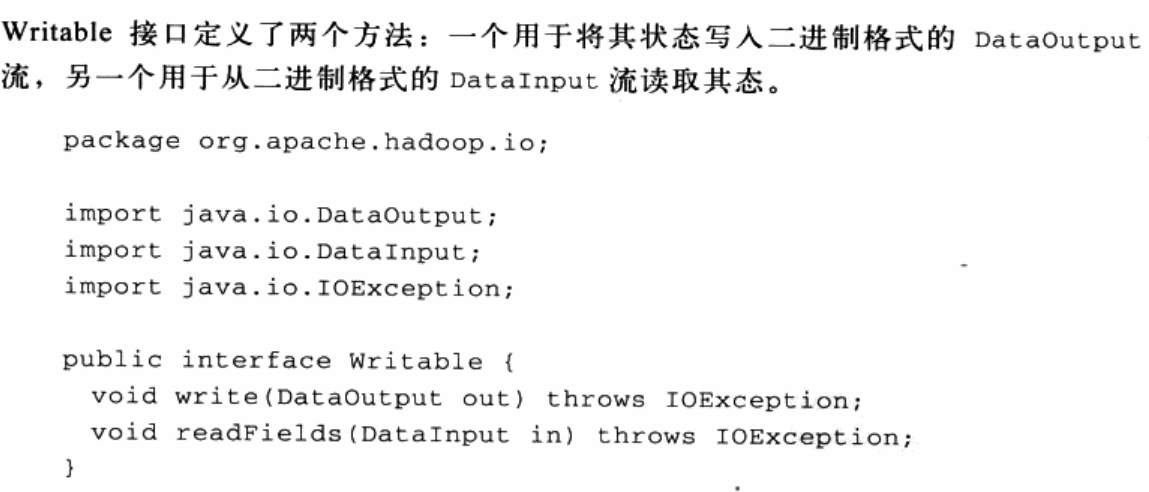
1. 压缩



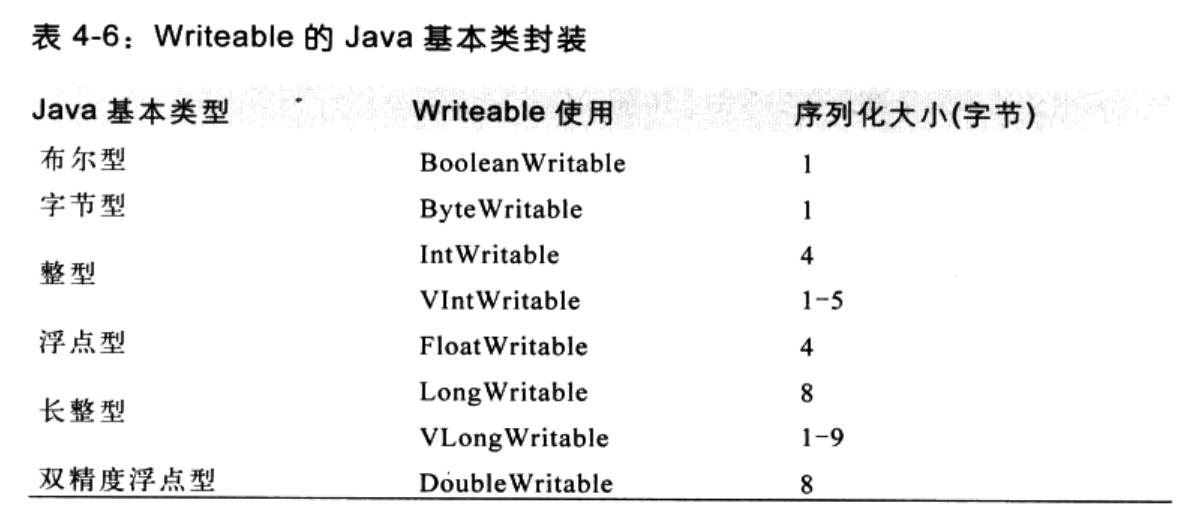
1. 编码/解码器
2. 压缩和输入分割
3. 在mapreduce使用压缩



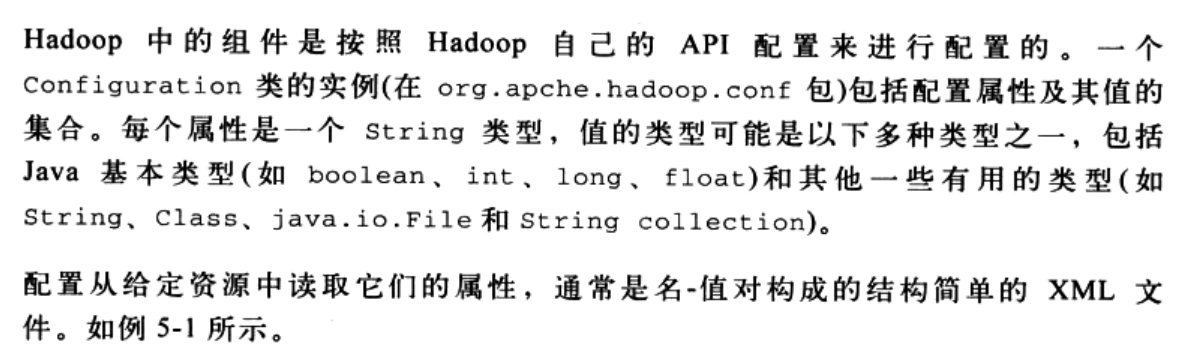
1. 序列化
2. writable接口



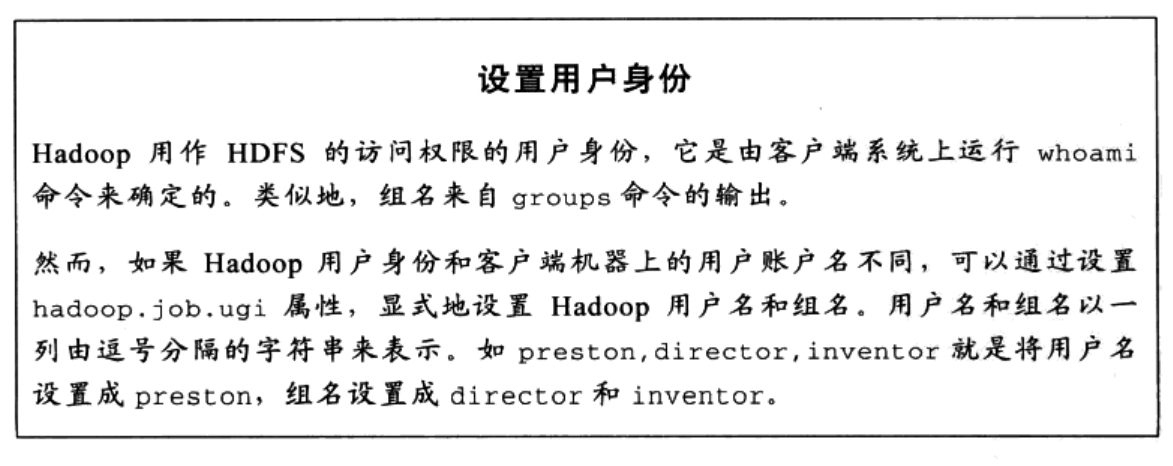
1. Writable类



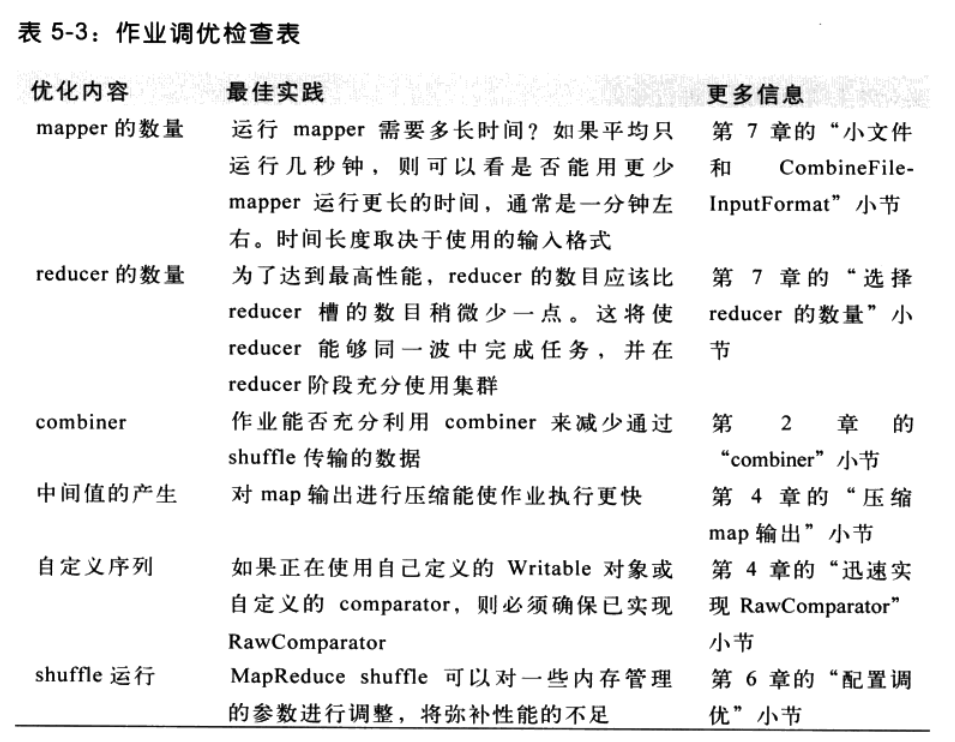
1. 基于文件的数据结构
2. Mapreduce应用开发
3. API配置



1. 配置开发环境



1. 编写单元测试
2. 本地运行测试数据
3. 在集群中运行
4. 作业调优



1. Mapreduce的工作流