### 自定义数据类型

* 1. 要实现writable接口
  2. 读写顺序要一致
  3. 构造方法如果进行了重写，要显示定义无参的构造方法
  4. 重写toString方法

### 自定义分区

数据分发的策略

如何实现自定义分区？

前提要保证相同key的数据会发送到同一个reduce中，或者说是分到同一个分区中

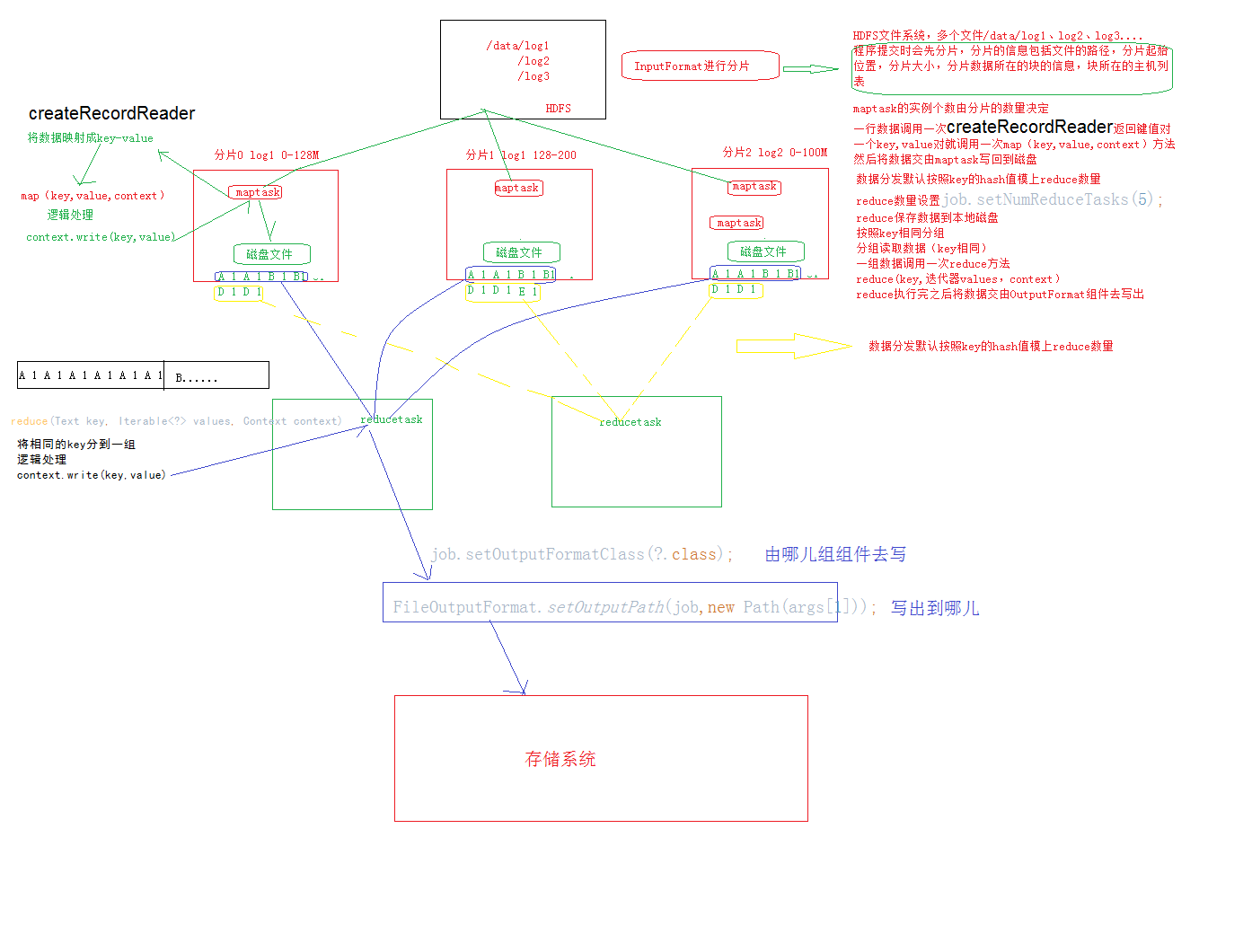
Partitioner<key,value> ，这里的key和value的数据类型和map输出的数据类型保持一致

继承Patitioner这个类，重写分区方法

然后在job中设置使用我们自定义的分区方法进行数据分发

### 使用MR建立索引

### Mr流程回顾



### MR的本地运行

1. 将org包放入java目录下
2. job.setJar("d:/word.jar");
   1. 打包将对应的包名改为word.jar
   2. 将word.jar放入对应的目录中
3. 将log4j文件放在resources
4. 设置运行参数，



1. MR的全流程

