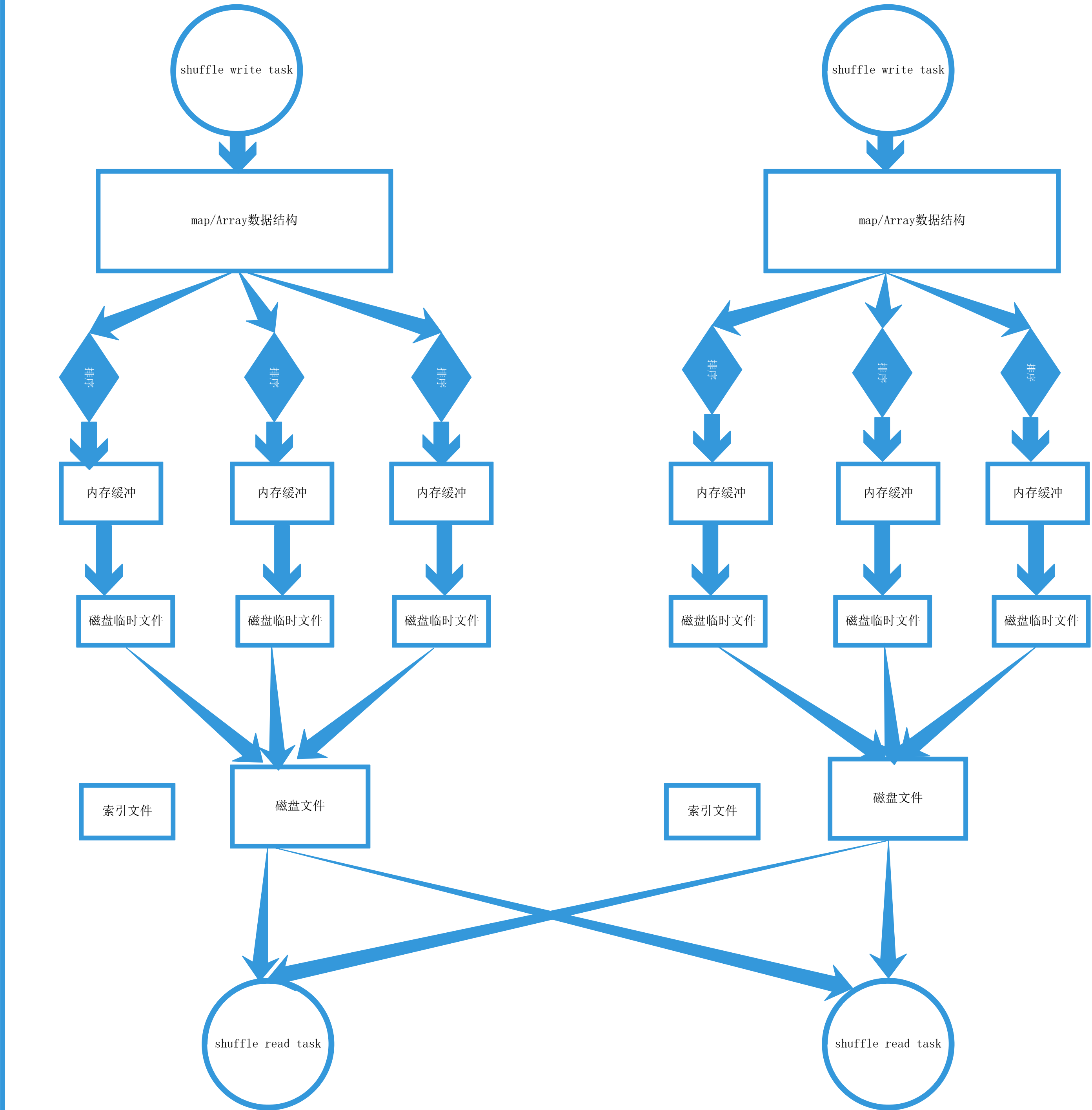


普通的sort shuffle 机制
数据在内存中先放入一个Map或者Array中 然后达到一定的阈值之后写入到内存缓冲中，数据在落入内存缓冲前有一个排序过程 排序之后按批次写入到内存缓冲中 当内存缓冲满了之后一次性写入到一个临时文件中 当shuffle writer task完成后 将所有的临时文件读取出来 merge成一个数据文件和一个索引文件 索引文件中保存的是下游shuffle read task 需要读取的startoffset和endoffset



ByPass机制的sort shuffle
task中的数据会将数据的key进行hash 然后发送到对应的内存缓冲中 等缓冲满了之后溢写到磁盘形成临时文件 当task执行完之后 将临时文件读取出来merge成一个磁盘文件和一个索引文件 当shuffle read task的数量小于 spark.shuffle.sort.bypassMergeThreshold（默认是200个）并且不是聚合类的shuffle算子比如reduceByKey时 就是ByPass机制的shuffle 否则就是普通机制的shuffle

