|  | 工程名 | 提供者 | 方法名 | 说明 | 参数 | 返回值 |
| --- | --- | --- | --- | --- | --- | --- |
|  | org.apache.hadoop.mapreduce.lib | FileInputFormat | getSplits(JobContext context) | 将输入划分为InputSplit | JobContext类型的contex | InputSplit对象 |
|  | org.apache.hadoop.mapred | IFile | Writer(Configuration conf, FSDataOutputStream out,          Class<K> keyClass, Class<V> valueClass,          CompressionCodec codec, Counters.Counter writesCounter) | 构造使用的支持压缩功能的输出流（类成员out，类成员rawOut保存了构造函数传入的out），相关的计数器，还有就是Key，Value的Serializer方法 | conf是配置参数，out是Writer的输出，keyClass 和valueClass 是输出的Key，Value的class属性，codec是对输出进行压缩的方法，参数writesCounter用于对输出字节数进行统计的Counters.Counter。 |  |
| append(K key, V value) | 主要过程是检查参数，然后将key和value序列化到DataOutputBuffer中，并获取序列化后的长度，最后把长度（2个）和DataOutputBuffer中的结果写到输出，并复位DataOutputBuffer和计数。 | key和value | void |
| append(DataInputBuffer key, DataInputBuffer value) |  | key和value | void |
|  | org.apache.hadoop.mapred | Task | run(JobConf job, TaskUmbilicalProtocol umbilical) | 执行Task |  | void |
| createRunner(TaskTracker tracker,TaskTracker.TaskInProgress tip) | 创建一个TaskRunner |  |  |
| isMapTask() | 判断是否是一个Map任务 |  | boolean型值 |
| done(TaskUmbilicalProtocol umbilical, TaskReporter reporter) | 用于做结束任务时的一些清理工作 |  | void |
|  | org.apache.hadoop.io | Writable | write(DataOutput out) | Mapreduce的输出接口 | DataOutput类型的out | void |
| readFields(DataInput in) | Mapreduce的输入接口 | DataInput类型的in | void |
|  | org.apache.hadoop.mapreduce | InputSplit | getLength() | 得到split的大小，以便可以按大小对splits进行排序 | 无 | long型数值 |
| getLocations() | 得到存放split数据的节点的名称 | 无 | String型数值 |
|  | org.apache.hadoop.mapreduce | RecordReader | nextKeyValue() | 读取下一个<key Value>对 | 无 | boolean型值 |
| getCurrentKey() | 读取当前的key值 | 无 | 返回当前的key值 |
| getCurrentValue() | 读取当前的Value值 | 无 | 返回当前的Value值 |
| getProgress() | 获取当前的读取进度 | 无 | 返回一个0和1之间的float值 |
| close() | 用于清理RecordReader | 无 | void |
|  | org.apache.hadoop.mapreduce | OutputFormat | getRecordWriter(TaskAttemptContext context ) | 得到给定任务的{@link RecordWriter} | TaskAttemptContext类型的context | 为输出的job返回一个{@link RecordWriter}值 |
| checkOutputSpecs | 检查输出是否已存在，以及其格式是否正确 | JobContext类型的contex | void |