**案例：找出每个月气温最高的2天**

1数据 tq.txt

|  |
| --- |
| 1949-10-01 14:21:02 34c  1949-10-01 19:21:02 38c  1949-10-02 14:01:02 36c  1950-01-01 11:21:02 32c  1950-10-01 12:21:02 37c  1951-12-01 12:21:02 23c  1950-10-02 12:21:02 41c  1950-10-03 12:21:02 27c  1951-07-01 12:21:02 45c  1951-07-02 12:21:02 46c  1951-07-03 12:21:03 47c |

2java打包到myTq.jar

3tq.txt，myTq.jar上传到/usr/local/software

4执行

|  |
| --- |
| cd /usr/local/software  hdfs dfs -mkdir -p /data/tq/input  hdfs dfs -put tq.txt /data/tq/input  hadoop jar myTq.jar hadoop02.MyTq |

5查看结果

|  |
| --- |
| hdfs dfs -get /data/tq/output/\* ./  cat part-r-00000 |

**mapReduce处理过程**

第一次map

|  |  |  |
| --- | --- | --- |
| 入参 | 1949-10-01 14:21:02 34c  1949-10-01 19:21:02 38c  1949-10-02 14:01:02 36c | |
| 出参 | key | value |
| {year:1949,month:10,day: 1,wd:34}  {year:1949,month:10,day: 1,wd:38}  {year:1949,month:10,day: 2,wd:36} | 34  38  36 |

第一次reduce

|  |  |  |  |
| --- | --- | --- | --- |
| 入参 | key | value | |
| {year:1949,month:10,day: 1,wd:34}  {year:1949,month:10,day: 1,wd:38}  {year:1949,month:10,day: 2,wd:36} | 34  38  36 | |
| 出参 | key | | value |
| 1949-10-1  1949-10-2 | | 34  38 |