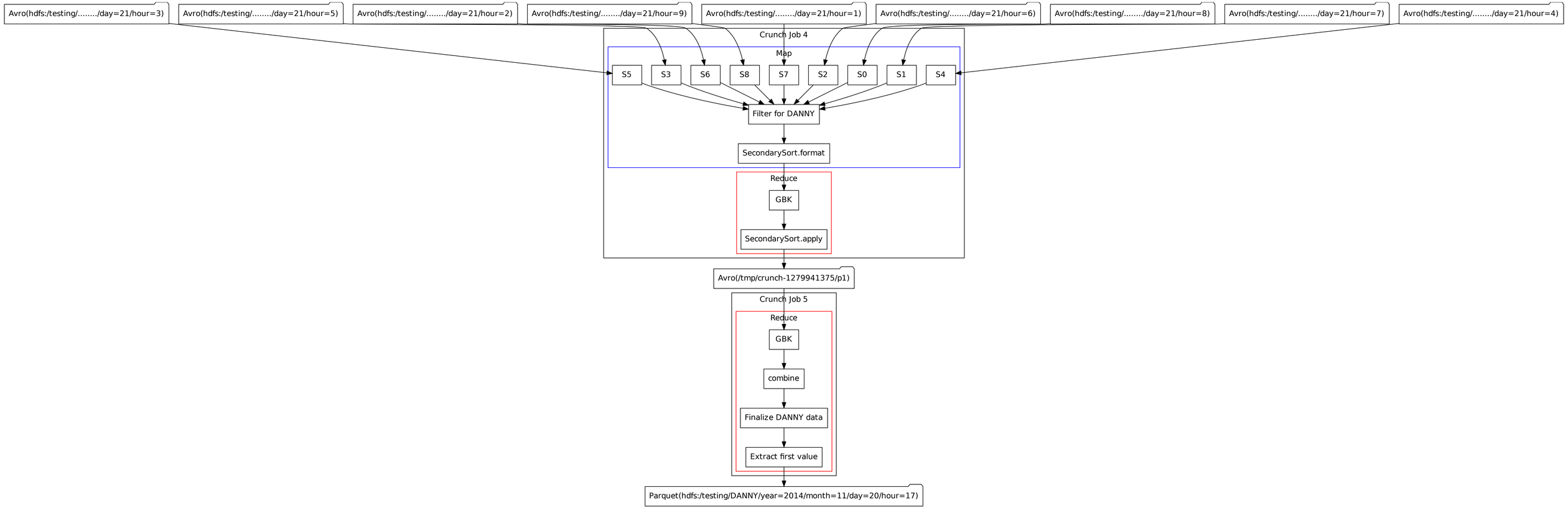
Avro(hdfs:/testing/...../day=21/hour=3)
Avro(hdfs:/testing/...../day=21/hour=5)
Avro(hdfs:/testing/...../day=21/hour=2)
Avro(hdfs:/testing/...../day=21/hour=9)
Avro(hdfs:/testing/...../day=21/hour=1)
Avro(hdfs:/testing/...../day=21/hour=6)
Avro(hdfs:/testing/...../day=21/hour=8)
Avro(hdfs:/testing/...../day=21/hour=7)
Avro(hdfs:/testing/...../day=21/hour=4)
Crunch Job 4
Map
S5
S3
S6
S8
S7
S2
S0
S1
S4
Filter for DANNY
SecondarySort.format
Reduce
GBK
SecondarySort.apply
Avro(/tmp/crunch-1279941375/p1)
Crunch Job 5
Reduce
GBK
combine
Finalize DANNY data
Extract first value
Parquet(hdfs:/testing/DANNY/year=2014/month=11/day=20/hour=17)