Sorting Using Hadoop MapReduce

Chan-Ching Hsu

Purpose

The goal is to write a program using MapReduce that can sort a large data set quickly. Sorting is a common task on large data sets, and the input is a standard data set.

Dataset

The name of the data file ends with the number of records in the input, for example "gensort-out-500K" has 500 thousand input records. Since each record is 100 bytes, this file takes about 50MB.

Each dataset is made up of multiple lines, one line per record. Each record has a key of 10 characters, followed by a space, followed by a payload (which in turn has a sequence number plus some other data). The final output is going to be sorted by key, and the output is going to include the payload for the key. Example of the input data:

```
[cchsu@n0 ~]$ hdfs dfs -cat /class/s17419/lab4/gensort-out-5M | head -n 10
AsfAGHM5om  0000000000000000000000000000000000  0000222200002222000022220000222200002222000000001111
~sHd0jDv6X  0000000000000000000000000000000001  77779999444488885555CCCC777755555555BBBB6666644446666
uI^EYm8s=|  0000000000000000000000000000000002  CCCCFFFF7777999995555FFFF1111222299998888444 4DDDDFFFF
Q)JN)R9z-L  0000000000000000000000000000000003  FFFF1111000000000000666688888BBBB33333333AAAA1111CCCC
o4FoBkqERn  0000000000000000000000000000000004  7777AAAABBBBBBBB22224444444499995555BBBB11118888DDDD
*}-Wz1;TD-  0000000000000000000000000000000005  AAAA88883333BBBB888888884444777722227777999900002222
0fssx}~[oB  0000000000000000000000000000000006  FFFF9999777744444AAAA7777EEEEDDDDAAAAAAAA99998888BBBB
mz4VCN@a#"  0000000000000000000000000000000007  DDDDBBBB1111FFFF2222DDDDFFFFBBBBFFFF6666444477778888
my+=5r7(N|  0000000000000000000000000000000008  22226666CCCC66662222FFFF0000EEEE11118888444455559999
5HA\z%qt{%  0000000000000000000000000000000009  0000AAAA8888FFFF000088880000000022225555111 1FFFFEEEE
```

Algorithm description

The first Mapper reads in an input data with *TextInputFormat* and outputs with *SequenceFileOutputFormat* with compression. There is no reducer actually implemented for the first Mapper since we just want to convert the input into another format (*TextInputFormat*). In the second MapReduce job, we essentially apply *TotalOrderPartitioner* to partition with *RandomSampler* to attempt to evenly distribute keys among Reducers; *RandomSampler* uniformly samples keys for partitioning keys into desired groups among which keys are in order. For example, the keys assigned to the i-th group will be less than the groups after i.

When reading input in *SequenceFile*, one should use *ByteWritable* to identify a key for the Mapper. Using other class, such as *Text* or *LongWritable* will result in exception.

The run time on of the program running on the largest dataset was roughly 295 seconds on Cystorm.

Example output

gensort-out-50M

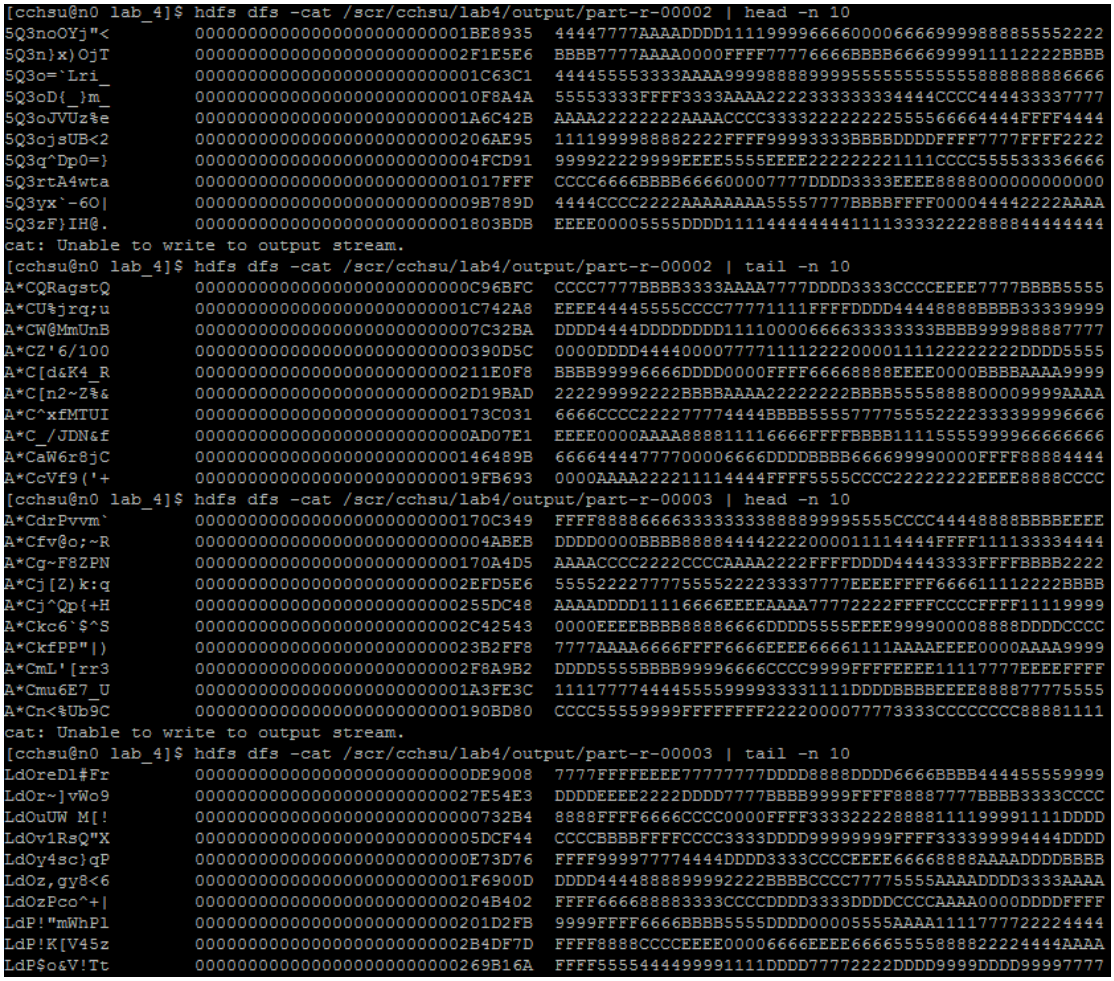
output file list

```
[cchsu@n0 lab_4]$ hdfs dfs -ls /scr/cchsu/lab4/output
Found 9 items
-rw-rw-r--   3 cchsu 419x          0 2017-02-09 17:48 /scr/cchsu/lab4/output/_SUCCESS
-rw-rw-r--   3 cchsu 419x  561305900 2017-02-09 17:48 /scr/cchsu/lab4/output/part-r-00000
-rw-rw-r--   3 cchsu 419x  571102200 2017-02-09 17:48 /scr/cchsu/lab4/output/part-r-00001
-rw-rw-r--   3 cchsu 419x  610220700 2017-02-09 17:48 /scr/cchsu/lab4/output/part-r-00002
-rw-rw-r--   3 cchsu 419x  611212900 2017-02-09 17:48 /scr/cchsu/lab4/output/part-r-00003
-rw-rw-r--   3 cchsu 419x  613859800 2017-02-09 17:48 /scr/cchsu/lab4/output/part-r-00004
-rw-rw-r--   3 cchsu 419x  564000200 2017-02-09 17:48 /scr/cchsu/lab4/output/part-r-00005
-rw-rw-r--   3 cchsu 419x  596407000 2017-02-09 17:48 /scr/cchsu/lab4/output/part-r-00006
-rw-rw-r--   3 cchsu 419x  871891300 2017-02-09 17:48 /scr/cchsu/lab4/output/part-r-00007
```

part-r-00000 and part-r-00001

```
[cchsu@n0 lab_4]$ hdfs dfs -cat /scr/cchsu/lab4/output/part-r-00000 | head -n 10
    "O!uve        0000000000000000000000000001228D4  777788880000222244444DDDDDDDDEEEE00000000CCCC7777DDDD
    ^3CO],        00000000000000000000000000158C5C5  5555AAAA9999EEEE888822229999CCCCDDDD6666555544442222
   !&S3/]]        00000000000000000000000002145D78  8888BBBBDDDD1111CCCC55556666BBBB1111EEEEDDDD22229999
   !,=U#,9        00000000000000000000000019072E3  33332222FFFFBBBB0000FFFFAAAA666655553333DDDD3333CCCC
   #'^cl'~        00000000000000000000000002EDC5C8  8888AAAA11114444FFFF77773333EEEE44440000FFFF99999999
   &&!epOj        00000000000000000000000001CBB42E  2222BBBB2222FFFFFFFFFFFCCCC555566666666777700003333
   'AJ;hL;        0000000000000000000000001FBD93E  FFFFFFFF66669999FFFF444466662222333300000BBBB33333333
   (VcS$Pz        00000000000000000000000000A9F516  8888AAAA7777BBBB7777AAAA6666DDDDEEEE9999BBBBBBBBBBBB
   *>MS1.E        00000000000000000000000000281643F  5555DDDD111199998888EEEEBBBB555555551111BBBBCCCC0000
   ,K4a-:v        00000000000000000000000001B8132  5555EEEE888899994444FFFF1111CCCCEEEE1111EEEE6666FFFF
cat: Unable to write to output stream.
[cchsu@n0 lab_4]$ hdfs dfs -cat /scr/cchsu/lab4/output/part-r-00000 | tail -n 10
*^q6/($sZP        00000000000000000000000000779761  555555555555DDDD777755559999AAAA000099995555EEEE6666
*^q6G1Z~&]        00000000000000000000000000131BFAC  5555FFFF3333AAAA2222444477772222DDDD0000CCCC44445555
*^q86A*)gd        0000000000000000000000000002552351  0000DDDD77773333111166666BBBB1111111133335555777766666
*^q::**n*u        00000000000000000000000000F7B410  9999999999999FFFF999900008888AAAA1111DDDD9999BBBB1111
*^q;@"|8Zg        00000000000000000000000000028C028A  DDDD222277770000AAAA22229999BBBB777788884444FFFF7777
*^q;|,SIBq        00000000000000000000000001A0563C  4444AAAA888855550000111166666CCCC4444BBBB0000777755555
*^q<e?M0Y)        00000000000000000000000000011845C0  2222BBBBDDDDDDDD6666DDDDDDDD8888CCCC99993333444411111
*^q?v3+Nu4        00000000000000000000000000011437C9  33330000BBBB9999AAAA66663333EEEE888877778888333EEEE
*^qCrfg*f[        00000000000000000000000000FF9701  99995555888888888888222BBBB9999AAAAFFFF777744446666
*^qCro=9Dm        00000000000000000000000002A26E2B  BBBBAAAAABBBB3333EEEE9999AAAAEEEE66669999AAAAFFFF4444
[cchsu@n0 lab_4]$ hdfs dfs -cat /scr/cchsu/lab4/output/part-r-00001 | head -n 10
*^qE)r"[o         00000000000000000000000002F8D3AF  BBBB5555AAAABBBBEEEE2222EEEE3333FFFF00006666555500000
*^qE*~z4Q)        00000000000000000000000001DD0FF2  7777333322223333EEEE55556666BBBBEEEE6666AAAAAAAAFFFF
*^qE9ao#wl        00000000000000000000000000280DB12  5555666600001111AAAA1111BBBB1111FFFF111144440000FFFF
*^qEq-'jJ(        00000000000000000000000001C9561B  88887777AAAA22227777222266669999FFFF6666AAAA00004444
*^qFNS)cHb        00000000000000000000000000018068CB  77778888666688885555222266662222555 5BBBB666655554444
*^qJ;Ge4T,        00000000000000000000000012D2C1E  22224444222211112222333 3BBBB2222FFFF3333DDDDDDDD3333
*^qJDSr::c        00000000000000000000000000139F1C2  BBBBDDDD555533333333CCCCBBBB0000BBBBEEEEAAAA1111FFFF
*^qK'.wsp,        00000000000000000000000013263C7  EEEE000000004444000000004444 1111BBBB33338888BBBB8888
*^qKd:?2s"        00000000000000000000000000FE6C0  111144444444DDDD0000FFFFAAAAEEEE4444AAAAEEEE44441111
*^qLE[5#05        00000000000000000000000002DD27B1  777799990000EEEEDDDD9999111199999999AAAA888811116666
cat: Unable to write to output stream.
[cchsu@n0 lab_4]$ hdfs dfs -cat /scr/cchsu/lab4/output/part-r-00001 | tail -n 10
5Q3cNwxExD        00000000000000000000000000015D7EDD  7777AAAAAAAA66664444EEEE333355553333CCCC3333EEEEAAAA
5Q3dpMyJ7~        00000000000000000000000000BA4DBE  8888AAAA000033336666AAAA11117777DDDDAAAA0000BBBB3333
5Q3eu516&c        00000000000000000000000000D4D2F5  CCCCBBBBEEEE3333BBBB2222CCCC66667777BBBB0000999922222
5Q3f%7I}qN        00000000000000000000000019F29A1  3333333377776666AAAA22225555333399998888DDDDAAAA6666
5Q3fe#HL&0        0000000000000000000000034BB6F  22226666DDDD66667777BBBB999911110000222 2EEEE99990000
5Q3fyQu`;3        00000000000000000000000002656B06  FFFF9999AAAADDDD5555EEEE1111333311111111AAAA8888BBBB
5Q3k /-bPT        00000000000000000000000007EC068  444422222222DDDD0000EEEE77771111333355555EEEE77779999
5Q3lq!#VT^        00000000000000000000000002516A1A  555511115555000088888BBBB0000AAAAEEEE88883333AAAA7777
5Q3m(DupiP        0000000000000000000000000E30A88  22223333888833334444AAAA4444555544447777FFFFDDDD9999
5Q3nfvA6yy        00000000000000000000000000094BE53  BBBB11113333DDDD0000BBBB3333EEEECCCCBBBB2222CCCCCCCC
```

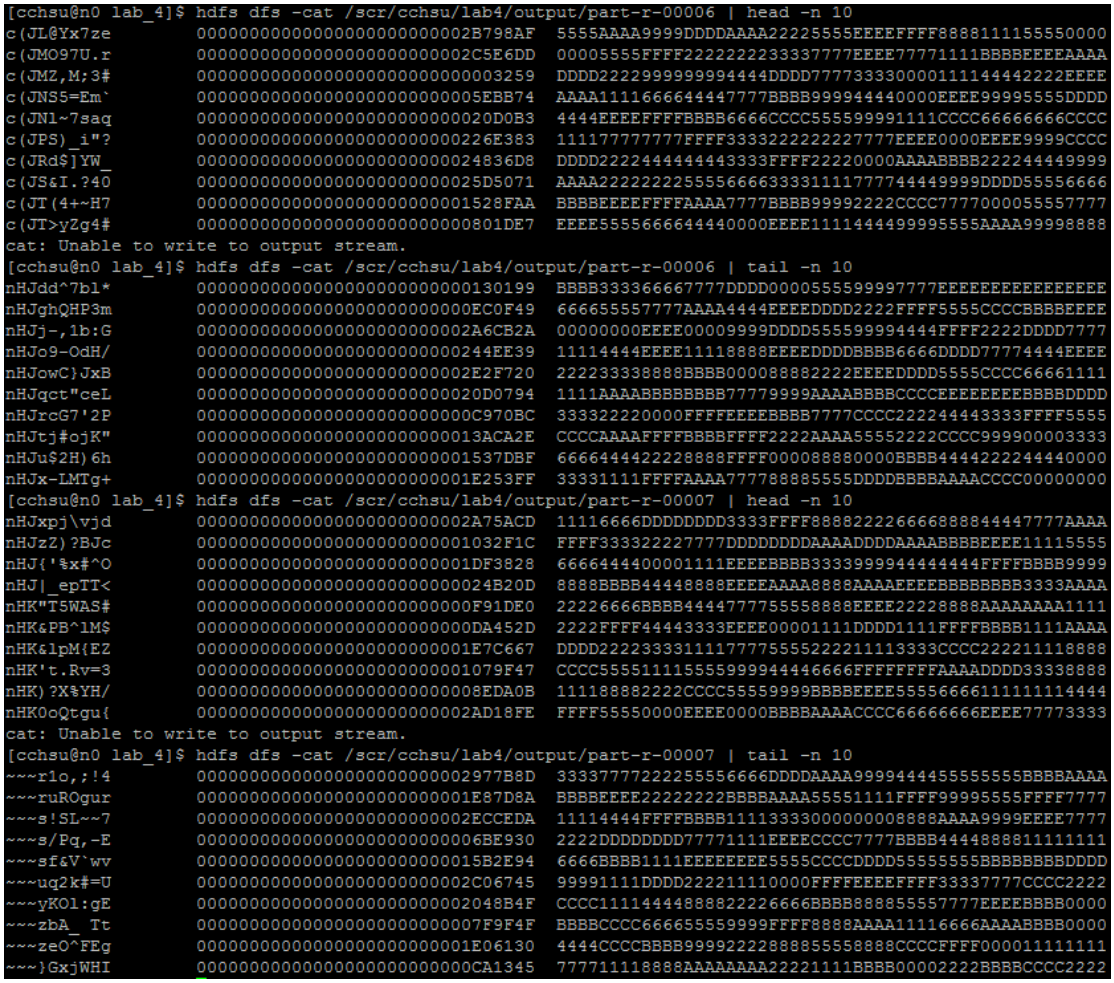
part-r-00002 and part-r-00003



part-r-00004 and part-r-00005

```
[cchsu@n0 lab_4]$ hdfs dfs -cat /scr/cchsu/lab4/output/part-r-00004 | head -n 10
LdP'RXP`,<        00000000000000000000000009A434B  99992222777711110000111177774444CCCCDDDD3333DDDD4444
LdP(G|`@dh        000000000000000000000000C13170   FFFFBBBB000066668888DDDD555577770000BBBBEEEEDDDD1111
LdP2;MC_R-        0000000000000000000000001BD9100  BBBB99992222FFFFEEEE5555EEEE000033338888BBBB00001111
LdP4cQJ]|W        00000000000000000000000001214C4E CCCC00003333CCCCCCCCCCCC9999BBBB00002222DDDD66663333
LdP5%CQe'M        000000000000000000000000795379   2222222233334444DDDDDDDDDDDCCCC7777AAAACCCC0000EEEE
LdP6iwApM*        0000000000000000000000064D4B7   2222DDDD8888FFFFAAAABBBBDDDDEEEE7777FFFFFFFCCCC8888
LdP7grbNaO        000000000000000000000003B5E9F   AAAA99997777DDDDEEEE00006666BBBB3333DDDD999966660000
LdP=iA'Jse        0000000000000000000000002EDE414 BBBB000099996666555566666666888833338888888833333DDDD
LdP>+N,;>`        00000000000000000000000140A875  6666CCCCBBBBEEEE22228888EEEEAAAA0000CCCCFFFF11112222
LdP>Bw*;3(        000000000000000000000000097DBC0 CCCCBBBB4444CCCCCCCCAAAAFFFFCCCC4444222255554444111
cat: Unable to write to output stream.
[cchsu@n0 lab_4]$ hdfs dfs -cat /scr/cchsu/lab4/output/part-r-00004 | tail -n 10
XCVbd1gI#4        0000000000000000000000000666FAD  0000FFFFDDDD88889999000066661111EEEE8888CCCC9999AAAA
XCVd8?2VDj        00000000000000000000000009957E9  999988884444EEEEBBBBEEEE222255557777CCCC11111111EEEE
XCVda2POVd        000000000000000000000000157323F  9999BBBBEEEEAAAA4444CCCC8888FFFFAAAA4444DDDDCCCC0000
XCVfP7O-w/        00000000000000000000000097DD8C   33332222AAAABBBB444499994444EEEE999955558888EEEE5555
XCVfi!PGJk        000000000000000000000000009452C3 5555AAAA4444CCCCCCCC7777555577772222BBBB33335555CCCC
XCVg_F{SDF        00000000000000000000000023EB102  11118888777777772222BBBB4444DDDD9999DDDD7777DDDDFFFF
XCVi+N<wy&        000000000000000000000000022BEF36 0000DDDD44442222FFFFEEEE33337777AAAA0000CCCC1111BBBB
XCVi:cgzKs        00000000000000000000000002A9A230 00008888DDDDDDDD6666EEEE3333FFFFDDDDAAAABBBB11111111
XCVkk'A+hr        000000000000000000000000BDF83F   55559999FFFFFFFF8888444433334444AAAA88887777CCCC0000
XCVko:Ef&`        00000000000000000000000047F9E2   CCCC3333777788881111444499994AAAA9999999922227777FFFF
[cchsu@n0 lab_4]$ hdfs dfs -cat /scr/cchsu/lab4/output/part-r-00005 | head -n 10
XCVl!'5YH"        0000000000000000000000001AF7F37  DDDD44448888EEEE2222FFFF55551111DDDDCCCC555544448888
XCVmK=*lJ$        00000000000000000000000001EAEE22 FFFFDDDDFFFFBBBB55553333FFFF00006666FFFFFFFF3333FFFF
XCVpP2d_:B        00000000000000000000000001CEF9FB BBBB11119999DDDD3333EEEE0000BBBB88888888000022224444
XCVq2?65}V        000000000000000000000000015EE09A DDDDEEEEFFFF6666111199999EEEE88886666BBBB777722227777
XCVq}}X:67        00000000000000000000000164A4C7   444477774444AAAA555544443333444BBBB99997777BBBB8888
XCVrA>80>s        00000000000000000000000002C5246B 0000DDDD3333AAAAEEEE33334444FFFFBBBB55550000BBBB4444
XCVtRP>py?        00000000000000000000000005C0AE4  555544442222AAAABBBB66669999977774444DDDD11112222DDDD
XCVt}+~~]q        000000000000000000000000009E337  DDDD2222444488888BBBBEEEE0000999990000EEEE111144448888
XCVu)j([3}        00000000000000000000000024B53E2  CCCC555577776666111111118888666655552222000077777FFFF
XCVx*,:;EY        000000000000000000000000095F395  4444DDDD5555FFFF999966660000CCCC8888CCCCAAAAFFFF2222
cat: Unable to write to output stream.
[cchsu@n0 lab_4]$ hdfs dfs -cat /scr/cchsu/lab4/output/part-r-00005 | tail -n 10
c(J=ZjAxUJ        0000000000000000000000000249BAB7 8888AAAADDDDEEEE555566666666FFFF111155559999CCCC8888
c(J>,JpY3=        000000000000000000000000009CA40F EEEE11116666444466664444777799994444999955555FFFF0000
c(J>Iv n'H        00000000000000000000000001D9EF43 99992222AAAACCCCAAAAFFFF3333CCCCDDDDDDDDEEEEDDDDCCCC
c(J>y#Y=OZ        000000000000000000000000001340373 22223333FFFF9999CCCCBBBB44448888DDDDEEEEFFFFAAAACCCC
c(JATY6#8K        00000000000000000000000001C50536 4444EEEE1111AAAAFFFF8888AAAABBBB00002222EEEE1111BBBB
c(JFiVaG:]        000000000000000000000000008342CE BBBBAAAA00004444AAAA666611119999EEEE44448888EEEE3333
c(JHX(:;z7        00000000000000000000000000E7DFEB 111100001111EEEE888844445555333311117777DDDD33334444
c(JJ!-?yd{        000000000000000000000000000945FB6 FFFFCCCC3333111100000CCCCBBBB22226666000055559999BBBB
c(JJJN6J0Q        00000000000000000000000000099CB71 44441111AAAACCCCFFFFAAAA222233334444DDDDAAAA55556666
c(JKb\XElf        000000000000000000000000001F82369 9999EEEEEEEECCCC11117777BBBB1111FFFF111111119999EEEE
```

part-r-00006 and part-r-00007



Resource

- https://hadoop.apache.org/docs/r2.4.1/api/org/apache/hadoop/mapred/Partitioner.html

Template Code

- Chapter 9 MapReduce Features: Total Sort, Hadoop – The Definitive Guide