|  |  |
| --- | --- |
|  | **2017** |
|  | Shalaj Shukla |

|  |
| --- |
| **[Different file formats]** |
|  |

Table of Contents

[Plain Text Format without Compression 2](#_Toc483910352)

[Load data to Plain Text format 2](#_Toc483910353)

[Data Size 4](#_Toc483910354)

[Select some columns 5](#_Toc483910355)

[Check Max Date 8](#_Toc483910356)

[User wise Max Date 12](#_Toc483910357)

[Select All Columns 16](#_Toc483910358)

[Select All Columns with filter Condition 18](#_Toc483910359)

[Plain Text Format Gzip compression 23](#_Toc483910360)

[Load data to Plain Text format 23](#_Toc483910361)

[Data Size 25](#_Toc483910362)

[Select some columns 26](#_Toc483910363)

[Check Max Date 28](#_Toc483910364)

[User wise Max Date 29](#_Toc483910365)

[Select All Columns 31](#_Toc483910366)

[Select All Columns with filter Condition 34](#_Toc483910367)

[Plain Text Format Bzip2 compression 37](#_Toc483910368)

[Load data to Plain Text format 37](#_Toc483910369)

[Data Size 39](#_Toc483910370)

[Select some columns 40](#_Toc483910371)

[Check Max Date 43](#_Toc483910372)

[User wise Max Date 46](#_Toc483910373)

[Select All Columns 50](#_Toc483910374)

[Select All Columns with filter Condition 53](#_Toc483910375)

[RC File format Snappy compression 58](#_Toc483910376)

[Load data to RC Table 59](#_Toc483910377)

[Data Size 62](#_Toc483910378)

[Select some columns 63](#_Toc483910379)

[Check Max Date 66](#_Toc483910380)

[User wise Max Date 68](#_Toc483910381)

[Select All Columns 70](#_Toc483910382)

[Select All Columns with filter Condition 73](#_Toc483910383)

[RC File format Gzip compression 75](#_Toc483910384)

[Load data to RC Table 75](#_Toc483910385)

[Data Size 77](#_Toc483910386)

[Select some columns 77](#_Toc483910387)

[Check Max Date 78](#_Toc483910388)

[User wise Max Date 79](#_Toc483910389)

[Select All Columns 81](#_Toc483910390)

[Select All Columns with filter Condition 84](#_Toc483910391)

[ORC File format Zlib Compression 86](#_Toc483910392)

[Load data to ORC Table 86](#_Toc483910393)

[Data Size 88](#_Toc483910394)

[Select some columns 90](#_Toc483910395)

[Check Max Date 91](#_Toc483910396)

[User wise Max Date 93](#_Toc483910397)

[Select All Columns 95](#_Toc483910398)

[Select All Columns with filter Condition 97](#_Toc483910399)

[ORC File format snappy Compression 99](#_Toc483910400)

[Load data to ORC Table 99](#_Toc483910401)

[Data Size 101](#_Toc483910402)

[Select some columns 101](#_Toc483910403)

[Check Max Date 102](#_Toc483910404)

[User wise Max Date 104](#_Toc483910405)

[Select All Columns 106](#_Toc483910406)

[Select All Columns with filter Condition 109](#_Toc483910407)

[ORC File format NONE Compression 111](#_Toc483910408)

[Load data to ORC Table 111](#_Toc483910409)

[Data Size 113](#_Toc483910410)

[Select some columns 113](#_Toc483910411)

[Check Max Date 114](#_Toc483910412)

[User wise Max Date 116](#_Toc483910413)

[Select All Columns 117](#_Toc483910414)

[Select All Columns with filter Condition 120](#_Toc483910415)

[Parquet File format Default Compression 122](#_Toc483910416)

[Load data to Parquet Table 122](#_Toc483910417)

[Data Size 124](#_Toc483910418)

[Select some columns 124](#_Toc483910419)

[Check Max Date 127](#_Toc483910420)

[User wise Max Date 129](#_Toc483910421)

[Select All Columns 130](#_Toc483910422)

[Select All Columns with filter Condition 134](#_Toc483910423)

[Parquet File format Gzip Compression 136](#_Toc483910424)

[Load data to Parquet Table 136](#_Toc483910425)

[Data Size 138](#_Toc483910426)

[Select some columns 138](#_Toc483910427)

[Check Max Date 139](#_Toc483910428)

[User wise Max Date 142](#_Toc483910429)

[Select All Columns 143](#_Toc483910430)

[Select All Columns with filter Condition 146](#_Toc483910431)

[Parquet File format Snappy Compression 148](#_Toc483910432)

[Load data to Parquet Table 148](#_Toc483910433)

[Data Size 150](#_Toc483910434)

[Select some columns 150](#_Toc483910435)

[Check Max Date 152](#_Toc483910436)

[User wise Max Date 154](#_Toc483910437)

[Select All Columns 155](#_Toc483910438)

[Select All Columns with filter Condition 158](#_Toc483910439)

[Avro File format default compression 162](#_Toc483910440)

[Load data to Avro Table 162](#_Toc483910441)

[Data Size 164](#_Toc483910442)

[Select some columns 165](#_Toc483910443)

[Check Max Date 166](#_Toc483910444)

[User wise Max Date 168](#_Toc483910445)

[Select All Columns 170](#_Toc483910446)

[Select All Columns with filter Condition 172](#_Toc483910447)

[Avro File format snappy compression 175](#_Toc483910448)

[Load data to Avro Table 175](#_Toc483910449)

[Data Size 178](#_Toc483910450)

[Select some columns 178](#_Toc483910451)

[Check Max Date 179](#_Toc483910452)

[User wise Max Date 181](#_Toc483910453)

[Select All Columns 183](#_Toc483910454)

[Select All Columns with filter Condition 186](#_Toc483910455)

[Sequence File format Snappy Compression 188](#_Toc483910456)

[Load data to Sequence Table 188](#_Toc483910457)

[Data Size 190](#_Toc483910458)

[Select some columns 190](#_Toc483910459)

[Check Max Date 192](#_Toc483910460)

[User wise Max Date 194](#_Toc483910461)

[Select All Columns 196](#_Toc483910462)

[Select All Columns with filter Condition 198](#_Toc483910463)

[Sequence File format Gzip Compression 201](#_Toc483910464)

[Load data to Sequence Table 201](#_Toc483910465)

[Data Size 203](#_Toc483910466)

[Select some columns 203](#_Toc483910467)

[Check Max Date 204](#_Toc483910468)

[User wise Max Date 206](#_Toc483910469)

[Select All Columns 207](#_Toc483910470)

[Select All Columns with filter Condition 210](#_Toc483910471)

[Sequence File format Bzip2 Compression 213](#_Toc483910472)

[Load data to Sequence Table 213](#_Toc483910473)

[Data Size 215](#_Toc483910474)

[Select some columns 215](#_Toc483910475)

[Check Max Date 217](#_Toc483910476)

[User wise Max Date 219](#_Toc483910477)

[Select All Columns 222](#_Toc483910478)

[Select All Columns with filter Condition 224](#_Toc483910479)

[Summary 227](#_Toc483910480)

[Conclusion 228](#_Toc483910481)

# Plain Text Format without Compression

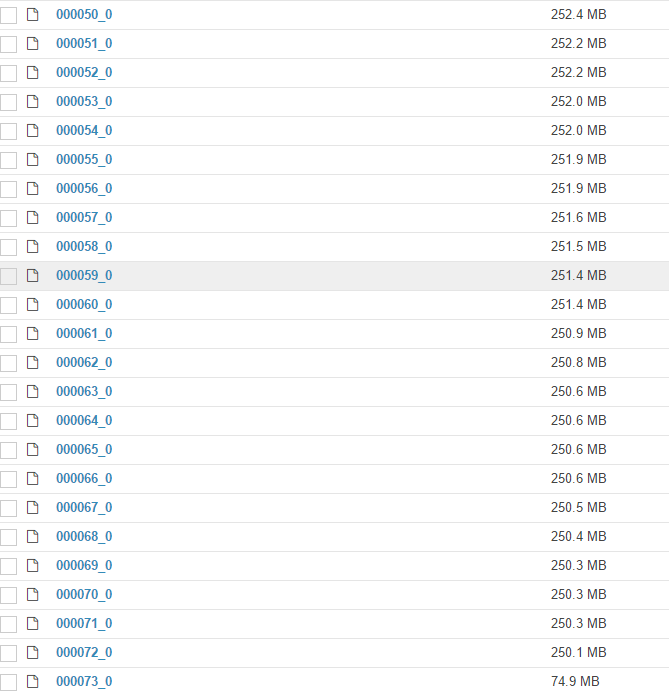
|  |
| --- |
| create table postimport\_text like postimportwithoutpart; |

## Load data to Plain Text format

|  |
| --- |
| hive> insert overwrite table postimport\_text select \* from postimportwithoutpart;  Query ID = shalaj\_20170405070404\_329dbfb2-74b0-472a-86bb-63a758e2b3e9  Total jobs = 3  Launching Job 1 out of 3  Number of reduce tasks is set to 0 since there's no reduce operator  Starting Job = job\_1490094830773\_0126, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0126/  Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0126  Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  2017-04-05 07:04:23,147 Stage-1 map = 0%, reduce = 0%  2017-04-05 07:04:35,208 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 40.05 sec  2017-04-05 07:05:35,695 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 81.63 sec  2017-04-05 07:06:06,714 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 96.16 sec  2017-04-05 07:06:11,542 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 99.01 sec  2017-04-05 07:06:15,395 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 100.7 sec  2017-04-05 07:06:44,132 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 113.12 sec  2017-04-05 07:06:47,752 Stage-1 map = 6%, reduce = 0%, Cumulative CPU 114.92 sec  2017-04-05 07:06:54,209 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 120.44 sec  2017-04-05 07:06:56,609 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 121.62 sec  2017-04-05 07:07:01,656 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 126.22 sec  2017-04-05 07:07:02,874 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 129.03 sec  2017-04-05 07:07:04,079 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 132.9 sec  2017-04-05 07:07:09,142 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 137.57 sec  2017-04-05 07:07:10,339 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 141.33 sec  2017-04-05 07:07:21,263 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 160.18 sec  2017-04-05 07:07:27,671 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 172.29 sec  2017-04-05 07:07:37,597 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 177.2 sec  2017-04-05 07:07:39,863 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 178.12 sec  2017-04-05 07:08:40,403 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 219.56 sec  2017-04-05 07:08:54,169 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 226.69 sec  2017-04-05 07:09:00,274 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 229.71 sec  2017-04-05 07:09:02,694 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 230.85 sec  2017-04-05 07:09:14,990 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 239.72 sec  2017-04-05 07:09:29,746 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 246.53 sec  2017-04-05 07:09:38,189 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 250.22 sec  2017-04-05 07:09:41,827 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 251.99 sec  2017-04-05 07:09:44,449 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 257.51 sec  2017-04-05 07:09:45,648 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 260.98 sec  2017-04-05 07:09:54,212 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 276.66 sec  2017-04-05 07:09:59,495 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 286.67 sec  2017-04-05 07:10:05,540 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 289.7 sec  2017-04-05 07:10:09,422 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 300.72 sec  2017-04-05 07:10:22,428 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 312.79 sec  2017-04-05 07:10:23,574 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 313.44 sec  2017-04-05 07:11:05,495 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 343.56 sec  2017-04-05 07:11:36,179 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 358.18 sec  2017-04-05 07:11:41,048 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 360.87 sec  2017-04-05 07:11:43,466 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 362.14 sec  2017-04-05 07:12:04,475 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 374.65 sec  2017-04-05 07:12:15,496 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 382.26 sec  2017-04-05 07:12:20,251 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 384.63 sec  2017-04-05 07:12:22,646 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 385.71 sec  2017-04-05 07:12:25,022 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 387.08 sec  2017-04-05 07:12:28,574 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 393.92 sec  2017-04-05 07:12:34,564 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 396.1 sec  2017-04-05 07:12:36,004 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 396.81 sec  2017-04-05 07:12:53,803 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 421.73 sec  2017-04-05 07:12:59,843 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 424.15 sec  2017-04-05 07:13:15,391 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 430.6 sec  2017-04-05 07:13:29,381 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 445.11 sec  2017-04-05 07:14:02,245 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 461.76 sec  2017-04-05 07:14:11,788 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 471.39 sec  2017-04-05 07:14:26,630 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 481.1 sec  2017-04-05 07:14:40,893 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 492.99 sec  2017-04-05 07:14:48,307 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 496.4 sec  2017-04-05 07:14:56,692 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 503.05 sec  2017-04-05 07:15:00,446 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 507.56 sec  2017-04-05 07:15:01,630 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 507.83 sec  2017-04-05 07:15:08,805 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 511.37 sec  2017-04-05 07:15:10,008 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 514.45 sec  2017-04-05 07:15:17,281 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 529.92 sec  2017-04-05 07:15:18,486 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 530.91 sec  2017-04-05 07:15:19,662 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 533.74 sec  2017-04-05 07:15:29,472 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 547.5 sec  2017-04-05 07:15:30,677 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 552.77 sec  2017-04-05 07:16:03,646 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 569.71 sec  2017-04-05 07:16:17,417 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 575.62 sec  2017-04-05 07:16:34,815 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 586.3 sec  2017-04-05 07:16:39,858 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 594.75 sec  2017-04-05 07:16:48,623 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 604.23 sec  2017-04-05 07:16:49,811 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 604.57 sec  2017-04-05 07:16:52,471 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 605.97 sec  2017-04-05 07:16:57,516 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 614.51 sec  2017-04-05 07:17:15,406 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 629.37 sec  2017-04-05 07:17:24,020 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 634.08 sec  2017-04-05 07:17:39,694 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 641.95 sec  2017-04-05 07:17:40,892 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 642.32 sec  2017-04-05 07:17:55,221 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 652.2 sec  2017-04-05 07:18:00,064 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 657.33 sec  2017-04-05 07:18:08,832 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 670.76 sec  2017-04-05 07:18:16,724 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 677.34 sec  2017-04-05 07:18:20,333 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 678.95 sec  2017-04-05 07:18:34,327 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 697.54 sec  2017-04-05 07:18:44,191 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 702.58 sec  2017-04-05 07:18:58,973 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 715.65 sec  2017-04-05 07:19:05,071 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 719.01 sec  2017-04-05 07:19:09,770 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 721.16 sec  2017-04-05 07:19:23,096 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 727.51 sec  2017-04-05 07:19:26,665 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 729.01 sec  2017-04-05 07:19:33,596 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 732.14 sec  2017-04-05 07:19:45,819 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 737.21 sec  2017-04-05 07:19:57,500 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 740.76 sec  2017-04-05 07:20:02,828 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 742.22 sec  2017-04-05 07:20:19,704 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 747.21 sec  2017-04-05 07:20:20,776 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 750.36 sec  2017-04-05 07:20:54,191 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 759.1 sec  2017-04-05 07:20:57,379 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 760.09 sec  2017-04-05 07:20:59,554 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 760.76 sec  MapReduce Total cumulative CPU time: 12 minutes 40 seconds 760 msec  Ended Job = job\_1490094830773\_0126  Stage-4 is selected by condition resolver.  Stage-3 is filtered out by condition resolver.  Stage-5 is filtered out by condition resolver.  Moving data to: hdfs://mac55:8020/user/hive/warehouse/esplus.db/postimport\_text/.hive-staging\_hive\_2017-04-05\_07-04-12\_159\_5061700321952921654-1/-ext-10000  Loading data to table esplus.postimport\_text  Table esplus.postimport\_text stats: [numFiles=74, numRows=42286538, totalSize=19481051559, rawDataSize=19438765021]  MapReduce Jobs Launched:  Stage-Stage-1: Map: 74 Cumulative CPU: 760.91 sec HDFS Read: 19486300793 HDFS Write: 19481058144 SUCCESS  Total MapReduce CPU Time Spent: 12 minutes 40 seconds 910 msec  OK  Time taken: 1013.319 seconds |

## Data Size

|  |
| --- |
| hive> dfs -du -s -h /user/hive/warehouse/esplus.db/postimport\_text;  18.1 G 54.4 G /user/hive/warehouse/esplus.db/postimport\_text |



Normal text file format has 74 files each around 250 MB accept last one

## Select some columns

|  |
| --- |
| select productname,username from postimport\_text;  INFO : Compiling command(queryId=hive\_20170404090101\_18fe947f-3ef5-4c6f-b6da-32460a5dfae2): select productname,username from postimport\_text  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:productname, type:string, comment:null), FieldSchema(name:username, type:string, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170404090101\_18fe947f-3ef5-4c6f-b6da-32460a5dfae2); Time taken: 0.105 seconds  INFO : Executing command(queryId=hive\_20170404090101\_18fe947f-3ef5-4c6f-b6da-32460a5dfae2): select productname,username from postimport\_text  INFO : Query ID = hive\_20170404090101\_18fe947f-3ef5-4c6f-b6da-32460a5dfae2  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:74  INFO : Submitting tokens for job: job\_1490094830773\_0107  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0107/  INFO : Starting Job = job\_1490094830773\_0107, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0107/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0107  INFO : Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  INFO : 2017-04-04 09:01:32,258 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-04 09:01:42,778 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 7.73 sec  INFO : 2017-04-04 09:01:44,826 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 26.68 sec  INFO : 2017-04-04 09:01:45,850 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 37.38 sec  INFO : 2017-04-04 09:01:46,871 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 39.31 sec  INFO : 2017-04-04 09:01:48,916 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 41.17 sec  INFO : 2017-04-04 09:01:49,938 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 43.66 sec  INFO : 2017-04-04 09:01:53,027 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 46.02 sec  INFO : 2017-04-04 09:01:55,075 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 52.53 sec  INFO : 2017-04-04 09:02:01,212 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 61.57 sec  INFO : 2017-04-04 09:02:02,232 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 63.55 sec  INFO : 2017-04-04 09:02:05,296 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 67.55 sec  INFO : 2017-04-04 09:02:08,399 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 69.03 sec  INFO : 2017-04-04 09:02:12,504 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 73.23 sec  INFO : 2017-04-04 09:02:13,524 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 76.48 sec  INFO : 2017-04-04 09:02:17,605 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 98.12 sec  INFO : 2017-04-04 09:02:18,626 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 99.48 sec  INFO : 2017-04-04 09:02:19,646 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 100.21 sec  INFO : 2017-04-04 09:02:20,670 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 102.82 sec  INFO : 2017-04-04 09:02:21,698 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 104.16 sec  INFO : 2017-04-04 09:02:23,741 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 105.2 sec  INFO : 2017-04-04 09:02:24,760 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 106.48 sec  INFO : 2017-04-04 09:02:29,870 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 111.87 sec  INFO : 2017-04-04 09:02:31,909 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 119.41 sec  INFO : 2017-04-04 09:02:34,971 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 125.41 sec  INFO : 2017-04-04 09:02:35,990 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 126.67 sec  INFO : 2017-04-04 09:02:38,041 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 127.57 sec  INFO : 2017-04-04 09:02:39,069 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 128.61 sec  INFO : 2017-04-04 09:02:42,138 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 132.4 sec  INFO : 2017-04-04 09:02:44,182 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 135.43 sec  INFO : 2017-04-04 09:02:45,205 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 137.82 sec  INFO : 2017-04-04 09:02:49,286 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 156.78 sec  INFO : 2017-04-04 09:02:50,311 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 157.26 sec  INFO : 2017-04-04 09:02:52,358 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 164.15 sec  INFO : 2017-04-04 09:02:53,377 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 166.56 sec  INFO : 2017-04-04 09:02:55,421 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 168.18 sec  INFO : 2017-04-04 09:02:57,475 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 172.83 sec  INFO : 2017-04-04 09:02:58,497 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 173.6 sec  INFO : 2017-04-04 09:02:59,525 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 174.51 sec  INFO : 2017-04-04 09:03:00,747 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 183.89 sec  INFO : 2017-04-04 09:03:01,766 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 186.78 sec  INFO : 2017-04-04 09:03:02,786 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 187.07 sec  INFO : 2017-04-04 09:03:04,828 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 194.83 sec  INFO : 2017-04-04 09:03:06,867 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 197.98 sec  INFO : 2017-04-04 09:03:09,932 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 199.72 sec  INFO : 2017-04-04 09:03:11,976 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 209.58 sec  INFO : 2017-04-04 09:03:13,011 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 210.54 sec  INFO : 2017-04-04 09:03:14,032 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 211.8 sec  INFO : 2017-04-04 09:03:16,078 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 218.08 sec  INFO : 2017-04-04 09:03:18,115 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 218.91 sec  INFO : 2017-04-04 09:03:23,211 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 229.44 sec  INFO : 2017-04-04 09:03:26,267 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 232.79 sec  INFO : 2017-04-04 09:03:30,346 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 239.62 sec  INFO : 2017-04-04 09:03:31,367 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 240.08 sec  INFO : 2017-04-04 09:03:33,403 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 241.68 sec  INFO : 2017-04-04 09:03:35,442 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 245.29 sec  INFO : 2017-04-04 09:03:38,515 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 248.8 sec  INFO : 2017-04-04 09:03:42,592 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 251.2 sec  INFO : 2017-04-04 09:03:43,611 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 254.06 sec  INFO : 2017-04-04 09:03:45,649 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 257.03 sec  INFO : 2017-04-04 09:03:49,726 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 264.28 sec  INFO : 2017-04-04 09:03:51,763 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 265.59 sec  INFO : 2017-04-04 09:03:52,781 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 266.16 sec  INFO : 2017-04-04 09:03:55,839 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 267.64 sec  INFO : 2017-04-04 09:03:57,875 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 268.65 sec  INFO : 2017-04-04 09:03:59,912 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 269.34 sec  INFO : 2017-04-04 09:04:02,968 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 270.55 sec  INFO : 2017-04-04 09:04:07,044 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 271.79 sec  INFO : 2017-04-04 09:04:14,185 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 272.97 sec  INFO : 2017-04-04 09:04:26,418 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 274.61 sec  INFO : MapReduce Total cumulative CPU time: 4 minutes 34 seconds 610 msec  INFO : Ended Job = job\_1490094830773\_0107  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 74 Cumulative CPU: 274.61 sec HDFS Read: 19486196675 HDFS Write: 2439779252 SUCCESS  INFO : Total MapReduce CPU Time Spent: 4 minutes 34 seconds 610 msec  INFO : Completed executing command(queryId=hive\_20170404090101\_18fe947f-3ef5-4c6f-b6da-32460a5dfae2); Time taken: 181.884 seconds  INFO : OK |

**System Counter**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Maps Total** | **Reduces Total** | **Total** |
| File Bytes Read | 0 | 0 | 0 |
| File Bytes Written | 17799802 | 0 | 17799802 |
| File Large Read Ops | 0 | 0 | 0 |
| File Read Ops | 0 | 0 | 0 |
| File Write Ops | 0 | 0 | 0 |
| Hdfs Bytes Read | 19486196675 | 0 | 19486196675 |
| Hdfs Bytes Written | 2439779252 | 0 | 2439779252 |
| Hdfs Large Read Ops | 0 | 0 | 0 |
| Hdfs Read Ops | 445 | 0 | 445 |
| Hdfs Write Ops | 148 | 0 | 148 |

## Check Max Date

|  |
| --- |
| select max(date) from postimport\_text;  INFO : Compiling command(queryId=hive\_20170404100808\_798a9e18-8afa-48bb-8fcb-905a5567d05d): select max(date) from postimport\_text  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:\_c0, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170404100808\_798a9e18-8afa-48bb-8fcb-905a5567d05d); Time taken: 0.102 seconds  INFO : Executing command(queryId=hive\_20170404100808\_798a9e18-8afa-48bb-8fcb-905a5567d05d): select max(date) from postimport\_text  INFO : Query ID = hive\_20170404100808\_798a9e18-8afa-48bb-8fcb-905a5567d05d  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks determined at compile time: 1  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:74  INFO : Submitting tokens for job: job\_1490094830773\_0116  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0116/  INFO : Starting Job = job\_1490094830773\_0116, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0116/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0116  INFO : Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 1  INFO : 2017-04-04 10:08:43,160 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-04 10:08:54,723 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 7.07 sec  INFO : 2017-04-04 10:08:56,776 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 35.27 sec  INFO : 2017-04-04 10:08:57,799 Stage-1 map = 6%, reduce = 0%, Cumulative CPU 38.76 sec  INFO : 2017-04-04 10:08:58,824 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 39.98 sec  INFO : 2017-04-04 10:09:01,903 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 44.42 sec  INFO : 2017-04-04 10:09:03,950 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 46.99 sec  INFO : 2017-04-04 10:09:05,007 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 48.16 sec  INFO : 2017-04-04 10:09:08,157 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 53.37 sec  INFO : 2017-04-04 10:09:09,194 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 55.69 sec  INFO : 2017-04-04 10:09:10,226 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 56.04 sec  INFO : 2017-04-04 10:09:11,247 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 57.3 sec  INFO : 2017-04-04 10:09:12,271 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 66.23 sec  INFO : 2017-04-04 10:09:15,340 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 74.27 sec  INFO : 2017-04-04 10:09:17,390 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 84.21 sec  INFO : 2017-04-04 10:09:18,416 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 86.2 sec  INFO : 2017-04-04 10:09:19,438 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 88.77 sec  INFO : 2017-04-04 10:09:20,473 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 93.87 sec  INFO : 2017-04-04 10:09:22,520 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 100.79 sec  INFO : 2017-04-04 10:09:23,543 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 102.28 sec  INFO : 2017-04-04 10:09:24,587 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 104.92 sec  INFO : 2017-04-04 10:09:28,773 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 110.34 sec  INFO : 2017-04-04 10:09:32,861 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 123.73 sec  INFO : 2017-04-04 10:09:33,882 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 123.94 sec  INFO : 2017-04-04 10:09:34,902 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 133.0 sec  INFO : 2017-04-04 10:09:36,944 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 136.95 sec  INFO : 2017-04-04 10:09:37,966 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 146.45 sec  INFO : 2017-04-04 10:09:41,031 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 148.42 sec  INFO : 2017-04-04 10:09:42,054 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 148.88 sec  INFO : 2017-04-04 10:09:43,099 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 152.62 sec  INFO : 2017-04-04 10:09:46,242 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 154.93 sec  INFO : 2017-04-04 10:09:47,296 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 162.18 sec  INFO : 2017-04-04 10:09:48,318 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 171.14 sec  INFO : 2017-04-04 10:09:49,339 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 172.4 sec  INFO : 2017-04-04 10:09:50,360 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 176.01 sec  INFO : 2017-04-04 10:09:51,387 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 179.4 sec  INFO : 2017-04-04 10:09:52,416 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 182.36 sec  INFO : 2017-04-04 10:09:53,442 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 185.87 sec  INFO : 2017-04-04 10:09:54,465 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 187.7 sec  INFO : 2017-04-04 10:09:57,590 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 190.5 sec  INFO : 2017-04-04 10:09:59,680 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 192.94 sec  INFO : 2017-04-04 10:10:01,739 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 204.65 sec  INFO : 2017-04-04 10:10:03,786 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 211.29 sec  INFO : 2017-04-04 10:10:04,806 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 216.1 sec  INFO : 2017-04-04 10:10:07,868 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 220.53 sec  INFO : 2017-04-04 10:10:08,889 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 226.18 sec  INFO : 2017-04-04 10:10:09,911 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 228.68 sec  INFO : 2017-04-04 10:10:10,932 Stage-1 map = 93%, reduce = 29%, Cumulative CPU 229.5 sec  INFO : 2017-04-04 10:10:13,997 Stage-1 map = 96%, reduce = 30%, Cumulative CPU 231.79 sec  INFO : 2017-04-04 10:10:17,062 Stage-1 map = 96%, reduce = 32%, Cumulative CPU 232.27 sec  INFO : 2017-04-04 10:10:35,467 Stage-1 map = 97%, reduce = 32%, Cumulative CPU 235.34 sec  INFO : 2017-04-04 10:10:36,490 Stage-1 map = 98%, reduce = 32%, Cumulative CPU 235.49 sec  INFO : 2017-04-04 10:10:38,551 Stage-1 map = 99%, reduce = 32%, Cumulative CPU 236.45 sec  INFO : 2017-04-04 10:10:40,595 Stage-1 map = 99%, reduce = 33%, Cumulative CPU 236.82 sec  INFO : 2017-04-04 10:10:48,787 Stage-1 map = 100%, reduce = 33%, Cumulative CPU 237.85 sec  INFO : 2017-04-04 10:10:49,812 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 238.51 sec  INFO : MapReduce Total cumulative CPU time: 3 minutes 58 seconds 510 msec  INFO : Ended Job = job\_1490094830773\_0116  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 74 Reduce: 1 Cumulative CPU: 238.51 sec HDFS Read: 19486246917 HDFS Write: 11 SUCCESS  INFO : Total MapReduce CPU Time Spent: 3 minutes 58 seconds 510 msec  INFO : Completed executing command(queryId=hive\_20170404100808\_798a9e18-8afa-48bb-8fcb-905a5567d05d); Time taken: 136.185 seconds  INFO : OK |

Normal text file took around 136 seconds to fetch max date

## User wise Max Date

|  |
| --- |
| select username,max(date) from postimport\_text group by username;  INFO : Compiling command(queryId=hive\_20170405050606\_a1c5a327-88a6-438f-950f-a0d12d2e792c): select username,max(date) from postimport\_text  group by username  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:username, type:string, comment:null), FieldSchema(name:\_c1, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170405050606\_a1c5a327-88a6-438f-950f-a0d12d2e792c); Time taken: 0.109 seconds  INFO : Executing command(queryId=hive\_20170405050606\_a1c5a327-88a6-438f-950f-a0d12d2e792c): select username,max(date) from postimport\_text  group by username  INFO : Query ID = hive\_20170405050606\_a1c5a327-88a6-438f-950f-a0d12d2e792c  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks not specified. Estimated from input data size: 291  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:74  INFO : Submitting tokens for job: job\_1490094830773\_0117  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0117/  INFO : Starting Job = job\_1490094830773\_0117, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0117/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0117  INFO : Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 291  INFO : 2017-04-05 05:06:34,941 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-05 05:06:47,569 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 26.94 sec  INFO : 2017-04-05 05:06:50,637 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 41.67 sec  INFO : 2017-04-05 05:06:51,660 Stage-1 map = 6%, reduce = 0%, Cumulative CPU 42.65 sec  INFO : 2017-04-05 05:06:53,585 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 48.28 sec  INFO : 2017-04-05 05:06:54,617 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 52.73 sec  INFO : 2017-04-05 05:06:56,705 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 53.41 sec  INFO : 2017-04-05 05:06:58,802 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 55.87 sec  INFO : 2017-04-05 05:07:01,917 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 59.01 sec  INFO : 2017-04-05 05:07:03,961 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 62.08 sec  INFO : 2017-04-05 05:07:04,982 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 71.74 sec  INFO : 2017-04-05 05:07:07,025 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 74.46 sec  INFO : 2017-04-05 05:07:08,048 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 81.66 sec  INFO : 2017-04-05 05:07:10,091 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 97.6 sec  INFO : 2017-04-05 05:07:11,119 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 99.17 sec  INFO : 2017-04-05 05:07:12,183 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 100.31 sec  INFO : 2017-04-05 05:07:13,214 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 110.77 sec  INFO : 2017-04-05 05:07:14,289 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 112.52 sec  INFO : 2017-04-05 05:07:15,323 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 113.98 sec  INFO : 2017-04-05 05:07:16,350 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 114.77 sec  INFO : 2017-04-05 05:07:20,455 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 120.28 sec  INFO : 2017-04-05 05:07:25,562 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 142.92 sec  INFO : 2017-04-05 05:07:26,582 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 147.73 sec  INFO : 2017-04-05 05:07:27,607 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 156.61 sec  INFO : 2017-04-05 05:07:28,628 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 158.27 sec  INFO : 2017-04-05 05:07:29,660 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 160.86 sec  INFO : 2017-04-05 05:07:30,692 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 162.49 sec  INFO : 2017-04-05 05:07:31,735 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 165.51 sec  INFO : 2017-04-05 05:07:32,773 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 167.01 sec  INFO : 2017-04-05 05:07:33,798 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 170.59 sec  INFO : 2017-04-05 05:07:34,836 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 171.7 sec  INFO : 2017-04-05 05:07:36,911 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 173.85 sec  INFO : 2017-04-05 05:07:37,936 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 176.34 sec  INFO : 2017-04-05 05:07:39,984 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 182.77 sec  INFO : 2017-04-05 05:07:41,049 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 183.73 sec  INFO : 2017-04-05 05:07:45,159 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 203.43 sec  INFO : 2017-04-05 05:07:48,220 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 217.42 sec  INFO : 2017-04-05 05:07:49,246 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 220.23 sec  INFO : 2017-04-05 05:07:50,267 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 223.24 sec  INFO : 2017-04-05 05:07:51,292 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 226.8 sec  INFO : 2017-04-05 05:07:54,361 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 231.21 sec  INFO : 2017-04-05 05:07:55,381 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 233.61 sec  INFO : 2017-04-05 05:07:56,407 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 234.65 sec  INFO : 2017-04-05 05:07:57,427 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 235.66 sec  INFO : 2017-04-05 05:07:58,452 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 238.89 sec  INFO : 2017-04-05 05:07:59,476 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 239.69 sec  INFO : 2017-04-05 05:08:08,667 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 262.27 sec  INFO : 2017-04-05 05:08:09,691 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 264.23 sec  INFO : 2017-04-05 05:08:10,712 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 267.98 sec  INFO : 2017-04-05 05:08:11,745 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 268.98 sec  INFO : 2017-04-05 05:08:12,775 Stage-1 map = 94%, reduce = 1%, Cumulative CPU 269.69 sec  INFO : 2017-04-05 05:08:13,807 Stage-1 map = 96%, reduce = 1%, Cumulative CPU 271.8 sec  INFO : 2017-04-05 05:08:17,159 Stage-1 map = 97%, reduce = 1%, Cumulative CPU 276.29 sec  INFO : 2017-04-05 05:08:19,235 Stage-1 map = 100%, reduce = 1%, Cumulative CPU 278.89 sec  INFO : 2017-04-05 05:08:20,291 Stage-1 map = 100%, reduce = 2%, Cumulative CPU 281.89 sec  INFO : 2017-04-05 05:08:21,346 Stage-1 map = 100%, reduce = 4%, Cumulative CPU 290.88 sec  INFO : 2017-04-05 05:08:24,516 Stage-1 map = 100%, reduce = 5%, Cumulative CPU 296.43 sec  INFO : 2017-04-05 05:08:25,556 Stage-1 map = 100%, reduce = 6%, Cumulative CPU 298.73 sec  INFO : 2017-04-05 05:08:26,584 Stage-1 map = 100%, reduce = 9%, Cumulative CPU 309.08 sec  INFO : 2017-04-05 05:08:29,703 Stage-1 map = 100%, reduce = 10%, Cumulative CPU 315.34 sec  INFO : 2017-04-05 05:08:30,760 Stage-1 map = 100%, reduce = 11%, Cumulative CPU 316.47 sec  INFO : 2017-04-05 05:08:32,812 Stage-1 map = 100%, reduce = 14%, Cumulative CPU 327.9 sec  INFO : 2017-04-05 05:08:34,883 Stage-1 map = 100%, reduce = 15%, Cumulative CPU 331.88 sec  INFO : 2017-04-05 05:08:35,916 Stage-1 map = 100%, reduce = 16%, Cumulative CPU 335.86 sec  INFO : 2017-04-05 05:08:37,988 Stage-1 map = 100%, reduce = 17%, Cumulative CPU 340.75 sec  INFO : 2017-04-05 05:08:39,022 Stage-1 map = 100%, reduce = 19%, Cumulative CPU 348.35 sec  INFO : 2017-04-05 05:08:40,063 Stage-1 map = 100%, reduce = 20%, Cumulative CPU 349.54 sec  INFO : 2017-04-05 05:08:41,117 Stage-1 map = 100%, reduce = 21%, Cumulative CPU 354.77 sec  INFO : 2017-04-05 05:08:42,170 Stage-1 map = 100%, reduce = 22%, Cumulative CPU 356.9 sec  INFO : 2017-04-05 05:08:44,278 Stage-1 map = 100%, reduce = 23%, Cumulative CPU 361.98 sec  INFO : 2017-04-05 05:08:45,314 Stage-1 map = 100%, reduce = 25%, Cumulative CPU 370.98 sec  INFO : 2017-04-05 05:08:46,362 Stage-1 map = 100%, reduce = 26%, Cumulative CPU 373.45 sec  INFO : 2017-04-05 05:08:48,466 Stage-1 map = 100%, reduce = 27%, Cumulative CPU 377.34 sec  INFO : 2017-04-05 05:08:50,567 Stage-1 map = 100%, reduce = 30%, Cumulative CPU 387.56 sec  INFO : 2017-04-05 05:08:51,611 Stage-1 map = 100%, reduce = 31%, Cumulative CPU 392.5 sec  INFO : 2017-04-05 05:08:53,708 Stage-1 map = 100%, reduce = 32%, Cumulative CPU 393.65 sec  INFO : 2017-04-05 05:08:54,757 Stage-1 map = 100%, reduce = 33%, Cumulative CPU 397.08 sec  INFO : 2017-04-05 05:08:56,818 Stage-1 map = 100%, reduce = 36%, Cumulative CPU 408.62 sec  INFO : 2017-04-05 05:08:58,934 Stage-1 map = 100%, reduce = 37%, Cumulative CPU 412.15 sec  INFO : 2017-04-05 05:08:59,989 Stage-1 map = 100%, reduce = 38%, Cumulative CPU 415.56 sec  INFO : 2017-04-05 05:09:01,024 Stage-1 map = 100%, reduce = 39%, Cumulative CPU 419.47 sec  INFO : 2017-04-05 05:09:03,086 Stage-1 map = 100%, reduce = 40%, Cumulative CPU 424.62 sec  INFO : 2017-04-05 05:09:04,132 Stage-1 map = 100%, reduce = 42%, Cumulative CPU 431.75 sec  INFO : 2017-04-05 05:09:05,173 Stage-1 map = 100%, reduce = 43%, Cumulative CPU 435.24 sec  INFO : 2017-04-05 05:09:06,232 Stage-1 map = 100%, reduce = 44%, Cumulative CPU 439.01 sec  INFO : 2017-04-05 05:09:08,329 Stage-1 map = 100%, reduce = 45%, Cumulative CPU 442.77 sec  INFO : 2017-04-05 05:09:09,383 Stage-1 map = 100%, reduce = 46%, Cumulative CPU 446.32 sec  INFO : 2017-04-05 05:09:10,413 Stage-1 map = 100%, reduce = 47%, Cumulative CPU 451.56 sec  INFO : 2017-04-05 05:09:11,444 Stage-1 map = 100%, reduce = 49%, Cumulative CPU 456.92 sec  INFO : 2017-04-05 05:09:14,587 Stage-1 map = 100%, reduce = 51%, Cumulative CPU 463.49 sec  INFO : 2017-04-05 05:09:15,623 Stage-1 map = 100%, reduce = 53%, Cumulative CPU 470.66 sec  INFO : 2017-04-05 05:09:16,653 Stage-1 map = 100%, reduce = 54%, Cumulative CPU 474.49 sec  INFO : 2017-04-05 05:09:19,790 Stage-1 map = 100%, reduce = 55%, Cumulative CPU 478.82 sec  INFO : 2017-04-05 05:09:20,850 Stage-1 map = 100%, reduce = 57%, Cumulative CPU 487.8 sec  INFO : 2017-04-05 05:09:21,890 Stage-1 map = 100%, reduce = 58%, Cumulative CPU 492.79 sec  INFO : 2017-04-05 05:09:22,925 Stage-1 map = 100%, reduce = 59%, Cumulative CPU 495.34 sec  INFO : 2017-04-05 05:09:25,026 Stage-1 map = 100%, reduce = 60%, Cumulative CPU 498.33 sec  INFO : 2017-04-05 05:09:26,080 Stage-1 map = 100%, reduce = 61%, Cumulative CPU 502.12 sec  INFO : 2017-04-05 05:09:27,207 Stage-1 map = 100%, reduce = 63%, Cumulative CPU 508.38 sec  INFO : 2017-04-05 05:09:28,232 Stage-1 map = 100%, reduce = 64%, Cumulative CPU 512.18 sec  INFO : 2017-04-05 05:09:29,270 Stage-1 map = 100%, reduce = 65%, Cumulative CPU 516.53 sec  INFO : 2017-04-05 05:09:31,349 Stage-1 map = 100%, reduce = 66%, Cumulative CPU 522.96 sec  INFO : 2017-04-05 05:09:32,409 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 526.83 sec  INFO : 2017-04-05 05:09:33,449 Stage-1 map = 100%, reduce = 68%, Cumulative CPU 530.54 sec  INFO : 2017-04-05 05:09:34,473 Stage-1 map = 100%, reduce = 69%, Cumulative CPU 533.04 sec  INFO : 2017-04-05 05:09:36,546 Stage-1 map = 100%, reduce = 72%, Cumulative CPU 542.92 sec  INFO : 2017-04-05 05:09:37,578 Stage-1 map = 100%, reduce = 73%, Cumulative CPU 546.65 sec  INFO : 2017-04-05 05:09:39,670 Stage-1 map = 100%, reduce = 74%, Cumulative CPU 550.21 sec  INFO : 2017-04-05 05:09:40,729 Stage-1 map = 100%, reduce = 75%, Cumulative CPU 554.04 sec  INFO : 2017-04-05 05:09:41,772 Stage-1 map = 100%, reduce = 77%, Cumulative CPU 559.94 sec  INFO : 2017-04-05 05:09:42,819 Stage-1 map = 100%, reduce = 78%, Cumulative CPU 564.71 sec  INFO : 2017-04-05 05:09:45,953 Stage-1 map = 100%, reduce = 80%, Cumulative CPU 572.42 sec  INFO : 2017-04-05 05:09:47,001 Stage-1 map = 100%, reduce = 81%, Cumulative CPU 575.08 sec  INFO : 2017-04-05 05:09:48,044 Stage-1 map = 100%, reduce = 82%, Cumulative CPU 581.54 sec  INFO : 2017-04-05 05:09:49,067 Stage-1 map = 100%, reduce = 84%, Cumulative CPU 585.69 sec  INFO : 2017-04-05 05:09:51,164 Stage-1 map = 100%, reduce = 85%, Cumulative CPU 590.88 sec  INFO : 2017-04-05 05:09:52,219 Stage-1 map = 100%, reduce = 86%, Cumulative CPU 593.23 sec  INFO : 2017-04-05 05:09:53,288 Stage-1 map = 100%, reduce = 87%, Cumulative CPU 598.1 sec  INFO : 2017-04-05 05:09:54,334 Stage-1 map = 100%, reduce = 88%, Cumulative CPU 601.73 sec  INFO : 2017-04-05 05:09:55,360 Stage-1 map = 100%, reduce = 89%, Cumulative CPU 605.74 sec  INFO : 2017-04-05 05:09:56,434 Stage-1 map = 100%, reduce = 90%, Cumulative CPU 608.37 sec  INFO : 2017-04-05 05:09:57,500 Stage-1 map = 100%, reduce = 91%, Cumulative CPU 612.18 sec  INFO : 2017-04-05 05:09:58,542 Stage-1 map = 100%, reduce = 92%, Cumulative CPU 617.13 sec  INFO : 2017-04-05 05:09:59,585 Stage-1 map = 100%, reduce = 93%, Cumulative CPU 620.9 sec  INFO : 2017-04-05 05:10:01,655 Stage-1 map = 100%, reduce = 95%, Cumulative CPU 626.28 sec  INFO : 2017-04-05 05:10:03,732 Stage-1 map = 100%, reduce = 96%, Cumulative CPU 630.43 sec  INFO : 2017-04-05 05:10:04,770 Stage-1 map = 100%, reduce = 97%, Cumulative CPU 637.84 sec  INFO : 2017-04-05 05:10:05,806 Stage-1 map = 100%, reduce = 99%, Cumulative CPU 644.52 sec  INFO : 2017-04-05 05:10:07,861 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 648.25 sec  INFO : MapReduce Total cumulative CPU time: 10 minutes 48 seconds 250 msec  INFO : Ended Job = job\_1490094830773\_0117  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 74 Reduce: 291 Cumulative CPU: 648.25 sec HDFS Read: 19487375914 HDFS Write: 171192 SUCCESS  INFO : Total MapReduce CPU Time Spent: 10 minutes 48 seconds 250 msec  INFO : Completed executing command(queryId=hive\_20170405050606\_a1c5a327-88a6-438f-950f-a0d12d2e792c); Time taken: 220.655 seconds  INFO : OK |

## Select All Columns

|  |
| --- |
| select \* from postimport\_text;  INFO : Compiling command(queryId=hive\_20170405094646\_7865337e-d1da-4bde-8654-4848377e7737): select \* from postimport\_text  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_text.username, type:string, comment:null), FieldSchema(name:postimport\_text.machinename, type:string, comment:null), FieldSchema(name:postimport\_text.machineid, type:string, comment:null), FieldSchema(name:postimport\_text.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_text.domainname, type:string, comment:null), FieldSchema(name:postimport\_text.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_text.modulename, type:string, comment:null), FieldSchema(name:postimport\_text.filename, type:string, comment:null), FieldSchema(name:postimport\_text.productname, type:string, comment:null), FieldSchema(name:postimport\_text.companyname, type:string, comment:null), FieldSchema(name:postimport\_text.description, type:string, comment:null), FieldSchema(name:postimport\_text.text, type:string, comment:null), FieldSchema(name:postimport\_text.activetime, type:double, comment:null), FieldSchema(name:postimport\_text.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_text.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_text.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170405094646\_7865337e-d1da-4bde-8654-4848377e7737); Time taken: 0.11 seconds  INFO : Executing command(queryId=hive\_20170405094646\_7865337e-d1da-4bde-8654-4848377e7737): select \* from postimport\_text  INFO : Query ID = hive\_20170405094646\_7865337e-d1da-4bde-8654-4848377e7737  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:74  INFO : Submitting tokens for job: job\_1490094830773\_0152  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0152/  INFO : Starting Job = job\_1490094830773\_0152, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0152/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0152  INFO : Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  INFO : 2017-04-05 09:46:40,953 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-05 09:46:58,629 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 64.93 sec  INFO : 2017-04-05 09:47:01,704 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 72.33 sec  INFO : 2017-04-05 09:47:04,767 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 76.56 sec  INFO : 2017-04-05 09:47:05,789 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 77.45 sec  INFO : 2017-04-05 09:47:43,558 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 103.65 sec  INFO : 2017-04-05 09:47:55,795 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 111.35 sec  INFO : 2017-04-05 09:48:47,780 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 140.5 sec  INFO : 2017-04-05 09:48:53,894 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 143.41 sec  INFO : 2017-04-05 09:48:55,931 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 144.16 sec  INFO : 2017-04-05 09:49:02,054 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 156.74 sec  INFO : 2017-04-05 09:49:08,171 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 168.72 sec  INFO : 2017-04-05 09:49:20,405 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 181.48 sec  INFO : 2017-04-05 09:49:23,463 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 183.52 sec  INFO : 2017-04-05 09:49:47,930 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 203.47 sec  INFO : 2017-04-05 09:49:48,949 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 204.02 sec  INFO : 2017-04-05 09:49:50,986 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 205.16 sec  INFO : 2017-04-05 09:49:56,084 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 208.13 sec  INFO : 2017-04-05 09:50:02,213 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 214.59 sec  INFO : 2017-04-05 09:50:04,250 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 215.5 sec  INFO : 2017-04-05 09:50:11,390 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 234.54 sec  INFO : 2017-04-05 09:50:49,084 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 258.2 sec  INFO : 2017-04-05 09:51:02,327 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 268.47 sec  INFO : 2017-04-05 09:51:03,347 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 268.69 sec  INFO : 2017-04-05 09:51:22,711 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 291.33 sec  INFO : 2017-04-05 09:51:28,828 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 294.9 sec  INFO : 2017-04-05 09:51:44,120 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 308.96 sec  INFO : 2017-04-05 09:51:47,180 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 310.62 sec  INFO : 2017-04-05 09:51:53,289 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 317.15 sec  INFO : 2017-04-05 09:51:57,363 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 319.74 sec  INFO : 2017-04-05 09:52:05,534 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 333.72 sec  INFO : 2017-04-05 09:52:06,560 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 334.24 sec  INFO : 2017-04-05 09:52:15,731 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 341.81 sec  INFO : 2017-04-05 09:52:24,921 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 354.08 sec  INFO : 2017-04-05 09:52:26,958 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 357.3 sec  INFO : 2017-04-05 09:52:51,414 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 372.92 sec  INFO : 2017-04-05 09:53:12,800 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 386.54 sec  INFO : 2017-04-05 09:53:32,142 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 398.53 sec  INFO : 2017-04-05 09:53:47,413 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 411.92 sec  INFO : 2017-04-05 09:54:30,172 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 432.12 sec  INFO : 2017-04-05 09:54:38,316 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 436.15 sec  INFO : 2017-04-05 09:54:46,464 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 442.95 sec  INFO : 2017-04-05 09:54:51,552 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 447.66 sec  INFO : 2017-04-05 09:54:55,630 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 452.37 sec  INFO : 2017-04-05 09:55:01,743 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 455.14 sec  INFO : 2017-04-05 09:55:03,781 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 462.77 sec  INFO : 2017-04-05 09:55:06,850 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 469.68 sec  INFO : 2017-04-05 09:55:38,425 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 492.5 sec  INFO : 2017-04-05 09:55:47,587 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 496.43 sec  INFO : 2017-04-05 09:56:06,941 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 517.74 sec  INFO : 2017-04-05 09:56:09,997 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 519.1 sec  INFO : 2017-04-05 09:56:24,256 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 531.35 sec  INFO : 2017-04-05 09:56:56,841 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 553.71 sec  INFO : 2017-04-05 09:57:17,207 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 564.24 sec  INFO : 2017-04-05 09:57:22,304 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 572.18 sec  INFO : 2017-04-05 09:57:29,431 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 575.75 sec  INFO : 2017-04-05 09:57:31,468 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 576.52 sec  INFO : 2017-04-05 09:57:32,489 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 577.13 sec  INFO : 2017-04-05 09:57:36,578 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 587.06 sec  INFO : 2017-04-05 09:57:42,706 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 592.91 sec  INFO : 2017-04-05 09:57:47,805 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 599.59 sec  INFO : 2017-04-05 09:57:48,827 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 600.11 sec  INFO : 2017-04-05 09:57:57,994 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 610.61 sec  INFO : 2017-04-05 09:58:01,049 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 615.11 sec  INFO : 2017-04-05 09:58:09,196 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 619.38 sec  INFO : 2017-04-05 09:58:21,421 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 631.57 sec  INFO : 2017-04-05 09:58:37,715 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 649.65 sec  INFO : 2017-04-05 09:59:17,427 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 671.95 sec  INFO : 2017-04-05 09:59:33,719 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 681.02 sec  INFO : 2017-04-05 09:59:54,087 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 692.71 sec  INFO : 2017-04-05 10:00:08,542 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 703.25 sec  INFO : 2017-04-05 10:00:14,654 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 709.48 sec  INFO : 2017-04-05 10:00:16,691 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 710.71 sec  INFO : 2017-04-05 10:00:23,819 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 717.17 sec  INFO : 2017-04-05 10:00:29,934 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 723.42 sec  INFO : 2017-04-05 10:00:37,067 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 733.01 sec  INFO : 2017-04-05 10:00:38,086 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 733.86 sec  INFO : 2017-04-05 10:00:40,122 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 741.31 sec  INFO : 2017-04-05 10:00:41,142 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 742.12 sec  INFO : 2017-04-05 10:00:43,178 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 745.75 sec  INFO : 2017-04-05 10:00:46,232 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 747.74 sec  INFO : 2017-04-05 10:00:51,350 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 758.2 sec  INFO : 2017-04-05 10:00:56,438 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 762.95 sec  INFO : 2017-04-05 10:01:07,633 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 768.48 sec  INFO : 2017-04-05 10:01:22,903 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 775.93 sec  INFO : 2017-04-05 10:01:42,247 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 785.05 sec  INFO : 2017-04-05 10:02:09,734 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 794.52 sec  INFO : 2017-04-05 10:02:16,859 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 796.98 sec  INFO : 2017-04-05 10:02:18,895 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 797.75 sec  INFO : 2017-04-05 10:02:21,949 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 798.66 sec  INFO : 2017-04-05 10:03:05,721 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 812.92 sec  INFO : 2017-04-05 10:03:07,758 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 813.67 sec  INFO : 2017-04-05 10:03:10,826 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 814.41 sec  INFO : MapReduce Total cumulative CPU time: 13 minutes 34 seconds 410 msec  INFO : Ended Job = job\_1490094830773\_0152  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 74 Cumulative CPU: 814.41 sec HDFS Read: 19486269850 HDFS Write: 19713140055 SUCCESS  INFO : Total MapReduce CPU Time Spent: 13 minutes 34 seconds 410 msec  INFO : Completed executing command(queryId=hive\_20170405094646\_7865337e-d1da-4bde-8654-4848377e7737); Time taken: 994.86 seconds INFO : OK |

## Select All Columns with filter Condition

|  |
| --- |
| hive> insert into dummy select \* from postimport\_text where username = '670063006100750074006F007500730065007200';  Query ID = shalaj\_20170406065959\_f2e1c369-51df-4c95-a8ed-225693d9aebf  Total jobs = 3  Launching Job 1 out of 3  Number of reduce tasks is set to 0 since there's no reduce operator  Starting Job = job\_1490094830773\_0181, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0181/  Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0181  Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  2017-04-06 06:59:20,571 Stage-1 map = 0%, reduce = 0%  2017-04-06 06:59:30,970 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 6.7 sec  2017-04-06 06:59:33,012 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 21.67 sec  2017-04-06 06:59:34,034 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 24.0 sec  2017-04-06 06:59:36,083 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 34.49 sec  2017-04-06 06:59:38,130 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 35.72 sec  2017-04-06 06:59:39,152 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 39.12 sec  2017-04-06 06:59:41,198 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 41.64 sec  2017-04-06 06:59:44,270 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 46.98 sec  2017-04-06 06:59:47,333 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 56.6 sec  2017-04-06 06:59:48,357 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 58.32 sec  2017-04-06 06:59:49,376 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 59.44 sec  2017-04-06 06:59:50,414 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 60.19 sec  2017-04-06 06:59:51,455 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 62.18 sec  2017-04-06 06:59:52,491 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 63.27 sec  2017-04-06 06:59:57,612 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 71.17 sec  2017-04-06 07:00:01,695 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 89.75 sec  2017-04-06 07:00:02,719 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 93.6 sec  2017-04-06 07:00:03,739 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 95.32 sec  2017-04-06 07:00:04,759 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 97.61 sec  2017-04-06 07:00:05,783 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 98.84 sec  2017-04-06 07:00:06,805 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 101.39 sec  2017-04-06 07:00:07,826 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 103.55 sec  2017-04-06 07:00:08,850 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 105.98 sec  2017-04-06 07:00:10,892 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 108.48 sec  2017-04-06 07:00:11,920 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 109.84 sec  2017-04-06 07:00:12,955 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 111.02 sec  2017-04-06 07:00:13,997 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 114.72 sec  2017-04-06 07:00:18,132 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 120.13 sec  2017-04-06 07:00:21,191 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 135.31 sec  2017-04-06 07:00:23,232 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 142.17 sec  2017-04-06 07:00:24,252 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 145.86 sec  2017-04-06 07:00:25,273 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 146.27 sec  2017-04-06 07:00:26,292 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 147.36 sec  2017-04-06 07:00:28,339 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 152.75 sec  2017-04-06 07:00:30,399 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 154.91 sec  2017-04-06 07:00:31,425 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 155.47 sec  2017-04-06 07:00:32,446 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 157.27 sec  2017-04-06 07:00:34,483 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 166.55 sec  2017-04-06 07:00:36,520 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 169.86 sec  2017-04-06 07:00:38,564 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 174.65 sec  2017-04-06 07:00:39,597 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 181.05 sec  2017-04-06 07:00:41,668 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 187.61 sec  2017-04-06 07:00:42,690 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 193.27 sec  2017-04-06 07:00:44,730 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 195.29 sec  2017-04-06 07:00:45,748 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 196.47 sec  2017-04-06 07:00:46,767 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 201.63 sec  2017-04-06 07:00:47,787 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 202.69 sec  2017-04-06 07:00:48,804 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 207.33 sec  2017-04-06 07:00:50,844 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 209.39 sec  2017-04-06 07:00:51,873 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 212.13 sec  2017-04-06 07:00:52,894 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 213.43 sec  2017-04-06 07:00:53,924 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 215.52 sec  2017-04-06 07:00:54,945 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 217.81 sec  2017-04-06 07:00:55,971 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 223.04 sec  2017-04-06 07:00:56,991 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 224.76 sec  2017-04-06 07:00:58,012 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 228.41 sec  2017-04-06 07:00:59,034 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 230.86 sec  2017-04-06 07:01:00,052 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 233.79 sec  2017-04-06 07:01:01,072 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 236.72 sec  2017-04-06 07:01:07,187 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 243.61 sec  MapReduce Total cumulative CPU time: 4 minutes 3 seconds 610 msec  Ended Job = job\_1490094830773\_0181  Stage-4 is filtered out by condition resolver.  Stage-3 is selected by condition resolver.  Stage-5 is filtered out by condition resolver.  Launching Job 3 out of 3  Number of reduce tasks is set to 0 since there's no reduce operator  Starting Job = job\_1490094830773\_0182, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0182/  Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0182  Hadoop job information for Stage-3: number of mappers: 74; number of reducers: 0  2017-04-06 07:01:18,896 Stage-3 map = 0%, reduce = 0%  2017-04-06 07:01:26,202 Stage-3 map = 7%, reduce = 0%, Cumulative CPU 8.37 sec  2017-04-06 07:01:27,228 Stage-3 map = 20%, reduce = 0%, Cumulative CPU 25.05 sec  2017-04-06 07:01:31,401 Stage-3 map = 24%, reduce = 0%, Cumulative CPU 29.95 sec  2017-04-06 07:01:33,497 Stage-3 map = 41%, reduce = 0%, Cumulative CPU 48.07 sec  2017-04-06 07:01:37,679 Stage-3 map = 45%, reduce = 0%, Cumulative CPU 52.48 sec  2017-04-06 07:01:39,762 Stage-3 map = 55%, reduce = 0%, Cumulative CPU 63.21 sec  2017-04-06 07:01:40,791 Stage-3 map = 61%, reduce = 0%, Cumulative CPU 68.96 sec  2017-04-06 07:01:41,823 Stage-3 map = 62%, reduce = 0%, Cumulative CPU 70.37 sec  2017-04-06 07:01:42,859 Stage-3 map = 65%, reduce = 0%, Cumulative CPU 72.94 sec  2017-04-06 07:01:44,941 Stage-3 map = 66%, reduce = 0%, Cumulative CPU 74.22 sec  2017-04-06 07:01:45,967 Stage-3 map = 78%, reduce = 0%, Cumulative CPU 85.96 sec  2017-04-06 07:01:47,013 Stage-3 map = 82%, reduce = 0%, Cumulative CPU 89.98 sec  2017-04-06 07:01:48,051 Stage-3 map = 85%, reduce = 0%, Cumulative CPU 92.41 sec  2017-04-06 07:01:49,084 Stage-3 map = 86%, reduce = 0%, Cumulative CPU 93.6 sec  2017-04-06 07:01:50,131 Stage-3 map = 88%, reduce = 0%, Cumulative CPU 94.63 sec  2017-04-06 07:01:51,182 Stage-3 map = 96%, reduce = 0%, Cumulative CPU 100.96 sec  2017-04-06 07:01:52,202 Stage-3 map = 99%, reduce = 0%, Cumulative CPU 103.28 sec  2017-04-06 07:01:53,223 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 104.19 sec  MapReduce Total cumulative CPU time: 1 minutes 44 seconds 190 msec  Ended Job = job\_1490094830773\_0182  Loading data to table esplus.dummy  Table esplus.dummy stats: [numFiles=74, numRows=86652, totalSize=5232192, rawDataSize=40448328]  MapReduce Jobs Launched:  Stage-Stage-1: Map: 74 Cumulative CPU: 243.61 sec HDFS Read: 19486323574 HDFS Write: 5237635 SUCCESS  Stage-Stage-3: Map: 74 Cumulative CPU: 104.19 sec HDFS Read: 5440576 HDFS Write: 5232192 SUCCESS  Total MapReduce CPU Time Spent: 5 minutes 47 seconds 800 msec  OK  Time taken: 166.443 seconds |

# Plain Text Format Gzip compression

|  |
| --- |
| SET hive.exec.compress.output=true;  SET mapred.output.compression.type=BLOCK;  set mapred.output.compress=true;  set mapred.output.compression.codec=org.apache.hadoop.io.compress.GzipCodec;  set io.compression.codecs=org.apache.hadoop.io.compress.GzipCodec; |

## Load data to Plain Text format

|  |
| --- |
| hive> insert overwrite table postimport\_text select \* from postimportwithoutpart;  Query ID = shalaj\_20170411083939\_462fffc3-9860-4d4b-85eb-1258f19b213c  Total jobs = 3  Launching Job 1 out of 3  Number of reduce tasks is set to 0 since there's no reduce operator  Starting Job = job\_1490094830773\_0405, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0405/  Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0405  Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  2017-04-11 08:39:40,633 Stage-1 map = 0%, reduce = 0%  2017-04-11 08:39:57,331 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 106.34 sec  2017-04-11 08:39:59,795 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 125.95 sec  2017-04-11 08:40:03,869 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 151.9 sec  2017-04-11 08:40:05,061 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 161.72 sec  2017-04-11 08:40:07,002 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 167.07 sec  2017-04-11 08:40:08,176 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 176.72 sec  2017-04-11 08:40:10,621 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 184.11 sec  2017-04-11 08:40:11,828 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 189.22 sec  2017-04-11 08:40:14,284 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 204.14 sec  2017-04-11 08:40:15,525 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 206.26 sec  2017-04-11 08:40:22,833 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 252.6 sec  2017-04-11 08:40:28,346 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 273.2 sec  2017-04-11 08:40:32,977 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 287.94 sec  2017-04-11 08:40:38,717 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 318.89 sec  2017-04-11 08:40:41,432 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 342.23 sec  2017-04-11 08:40:44,069 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 357.51 sec  2017-04-11 08:40:45,267 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 357.56 sec  2017-04-11 08:40:46,461 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 362.92 sec  2017-04-11 08:40:47,670 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 376.04 sec  2017-04-11 08:40:50,108 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 389.19 sec  2017-04-11 08:40:52,471 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 400.04 sec  2017-04-11 08:40:53,570 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 407.9 sec  2017-04-11 08:40:59,463 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 426.15 sec  2017-04-11 08:41:00,634 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 426.3 sec  2017-04-11 08:41:02,143 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 432.01 sec  2017-04-11 08:41:03,236 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 434.66 sec  2017-04-11 08:41:10,094 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 474.54 sec  2017-04-11 08:41:11,290 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 490.66 sec  2017-04-11 08:41:17,332 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 530.0 sec  2017-04-11 08:41:18,421 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 535.2 sec  2017-04-11 08:41:19,725 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 538.32 sec  2017-04-11 08:41:23,741 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 569.44 sec  2017-04-11 08:41:28,462 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 598.31 sec  2017-04-11 08:41:31,014 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 614.19 sec  2017-04-11 08:41:32,249 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 620.37 sec  2017-04-11 08:41:34,677 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 639.84 sec  2017-04-11 08:41:40,147 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 658.57 sec  2017-04-11 08:41:43,902 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 669.13 sec  2017-04-11 08:41:48,608 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 686.74 sec  2017-04-11 08:41:49,787 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 690.85 sec  2017-04-11 08:41:51,954 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 700.76 sec  2017-04-11 08:41:53,026 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 703.41 sec  2017-04-11 08:41:54,327 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 706.1 sec  2017-04-11 08:41:57,839 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 721.33 sec  2017-04-11 08:41:58,922 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 726.35 sec  2017-04-11 08:42:05,506 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 785.08 sec  2017-04-11 08:42:10,593 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 827.44 sec  2017-04-11 08:42:11,794 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 830.43 sec  2017-04-11 08:42:14,165 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 841.52 sec  2017-04-11 08:42:16,525 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 858.11 sec  2017-04-11 08:42:17,717 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 862.2 sec  2017-04-11 08:42:18,887 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 870.16 sec  2017-04-11 08:42:21,208 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 886.91 sec  2017-04-11 08:42:22,374 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 891.38 sec  2017-04-11 08:42:23,568 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 900.87 sec  2017-04-11 08:42:25,909 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 916.54 sec  2017-04-11 08:42:27,036 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 926.8 sec  2017-04-11 08:42:28,337 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 939.33 sec  2017-04-11 08:42:29,410 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 948.11 sec  2017-04-11 08:42:31,531 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 951.38 sec  2017-04-11 08:42:32,614 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 963.12 sec  2017-04-11 08:42:35,843 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 974.35 sec  2017-04-11 08:42:38,301 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 981.67 sec  2017-04-11 08:42:46,474 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 1052.83 sec  2017-04-11 08:42:47,545 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 1062.86 sec  2017-04-11 08:42:51,265 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 1087.12 sec  2017-04-11 08:42:52,432 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 1090.64 sec  2017-04-11 08:42:57,760 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 1122.02 sec  2017-04-11 08:42:58,984 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 1125.85 sec  2017-04-11 08:43:00,155 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 1132.88 sec  2017-04-11 08:43:01,376 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 1138.29 sec  2017-04-11 08:43:03,730 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 1147.52 sec  2017-04-11 08:43:11,995 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 1170.87 sec  2017-04-11 08:43:13,186 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 1175.22 sec  2017-04-11 08:43:17,808 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 1182.48 sec  2017-04-11 08:43:18,977 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 1186.36 sec  2017-04-11 08:43:48,415 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 1202.2 sec  2017-04-11 08:43:53,909 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1205.42 sec  MapReduce Total cumulative CPU time: 20 minutes 5 seconds 420 msec  Ended Job = job\_1490094830773\_0405  Stage-4 is selected by condition resolver.  Stage-3 is filtered out by condition resolver.  Stage-5 is filtered out by condition resolver.  Moving data to: hdfs://mac55:8020/user/hive/warehouse/esplus.db/postimport\_text/.hive-staging\_hive\_2017-04-11\_08-39-30\_715\_2247388098871785895-1/-ext-10000  Loading data to table esplus.postimport\_text  Table esplus.postimport\_text stats: [numFiles=74, numRows=42286538, totalSize=3615765118, rawDataSize=19438765021]  MapReduce Jobs Launched:  Stage-Stage-1: Map: 74 Cumulative CPU: 1205.42 sec HDFS Read: 19486300719 HDFS Write: 3615771703 SUCCESS  Total MapReduce CPU Time Spent: 20 minutes 5 seconds 420 msec  OK  Time taken: 266.638 seconds |

## Data Size

|  |
| --- |
| hive> dfs -du -s -h /user/hive/warehouse/esplus.db/postimport\_text;  3.4 G 10.1 G /user/hive/warehouse/esplus.db/postimport\_text |



Created 74 gz files

## Select some columns

|  |
| --- |
| select productname,username from postimport\_text;  INFO : Compiling command(queryId=hive\_20170411085353\_ff84f897-5378-4859-9769-af76f9fadb49): select productname,username from postimport\_text  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:productname, type:string, comment:null), FieldSchema(name:username, type:string, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411085353\_ff84f897-5378-4859-9769-af76f9fadb49); Time taken: 0.101 seconds  INFO : Executing command(queryId=hive\_20170411085353\_ff84f897-5378-4859-9769-af76f9fadb49): select productname,username from postimport\_text  INFO : Query ID = hive\_20170411085353\_ff84f897-5378-4859-9769-af76f9fadb49  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:74  INFO : Submitting tokens for job: job\_1490094830773\_0406  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0406/  INFO : Starting Job = job\_1490094830773\_0406, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0406/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0406  INFO : Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  INFO : 2017-04-11 08:53:26,185 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 08:53:35,682 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 14.91 sec  INFO : 2017-04-11 08:53:36,877 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 34.69 sec  INFO : 2017-04-11 08:53:38,032 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 49.69 sec  INFO : 2017-04-11 08:53:39,212 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 74.82 sec  INFO : 2017-04-11 08:53:43,452 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 79.66 sec  INFO : 2017-04-11 08:53:44,909 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 89.18 sec  INFO : 2017-04-11 08:53:46,220 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 98.83 sec  INFO : 2017-04-11 08:53:47,499 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 108.27 sec  INFO : 2017-04-11 08:53:49,793 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 136.76 sec  INFO : 2017-04-11 08:53:50,946 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 146.13 sec  INFO : 2017-04-11 08:53:52,531 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 151.93 sec  INFO : 2017-04-11 08:53:54,013 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 161.53 sec  INFO : 2017-04-11 08:53:55,373 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 166.45 sec  INFO : 2017-04-11 08:53:56,595 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 175.79 sec  INFO : 2017-04-11 08:53:57,878 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 180.37 sec  INFO : 2017-04-11 08:53:59,058 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 185.19 sec  INFO : 2017-04-11 08:54:00,290 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 189.91 sec  INFO : 2017-04-11 08:54:01,446 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 227.51 sec  INFO : 2017-04-11 08:54:04,196 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 236.93 sec  INFO : 2017-04-11 08:54:05,731 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 241.54 sec  INFO : 2017-04-11 08:54:07,059 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 250.6 sec  INFO : 2017-04-11 08:54:08,363 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 255.02 sec  INFO : 2017-04-11 08:54:09,701 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 264.72 sec  INFO : 2017-04-11 08:54:12,143 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 301.84 sec  INFO : 2017-04-11 08:54:13,315 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 305.93 sec  INFO : 2017-04-11 08:54:15,802 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 324.77 sec  INFO : 2017-04-11 08:54:16,946 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 331.41 sec  INFO : 2017-04-11 08:54:19,228 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 336.08 sec  INFO : 2017-04-11 08:54:20,408 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 345.25 sec  INFO : 2017-04-11 08:54:21,547 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 349.65 sec  INFO : MapReduce Total cumulative CPU time: 5 minutes 49 seconds 650 msec  INFO : Ended Job = job\_1490094830773\_0406  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 74 Cumulative CPU: 349.65 sec HDFS Read: 3616047502 HDFS Write: 233687022 SUCCESS  INFO : Total MapReduce CPU Time Spent: 5 minutes 49 seconds 650 msec  INFO : Completed executing command(queryId=hive\_20170411085353\_ff84f897-5378-4859-9769-af76f9fadb49); Time taken: 64.919 seconds  INFO : OK |

## Check Max Date

|  |
| --- |
| select max(date) from postimport\_text;  INFO : Compiling command(queryId=hive\_20170411085454\_8d121f9a-a129-411b-85c9-93796440fbc1): select max(date) from postimport\_text  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:\_c0, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411085454\_8d121f9a-a129-411b-85c9-93796440fbc1); Time taken: 0.113 seconds  INFO : Executing command(queryId=hive\_20170411085454\_8d121f9a-a129-411b-85c9-93796440fbc1): select max(date) from postimport\_text  INFO : Query ID = hive\_20170411085454\_8d121f9a-a129-411b-85c9-93796440fbc1  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks determined at compile time: 1  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:74  INFO : Submitting tokens for job: job\_1490094830773\_0407  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0407/  INFO : Starting Job = job\_1490094830773\_0407, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0407/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0407  INFO : Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 1  INFO : 2017-04-11 08:54:49,204 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 08:54:59,449 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 7.82 sec  INFO : 2017-04-11 08:55:00,601 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 35.04 sec  INFO : 2017-04-11 08:55:01,797 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 54.62 sec  INFO : 2017-04-11 08:55:03,776 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 58.91 sec  INFO : 2017-04-11 08:55:09,042 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 79.02 sec  INFO : 2017-04-11 08:55:10,415 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 86.75 sec  INFO : 2017-04-11 08:55:11,636 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 98.39 sec  INFO : 2017-04-11 08:55:12,973 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 106.03 sec  INFO : 2017-04-11 08:55:14,235 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 117.79 sec  INFO : 2017-04-11 08:55:18,366 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 133.53 sec  INFO : 2017-04-11 08:55:19,536 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 141.45 sec  INFO : 2017-04-11 08:55:20,749 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 149.15 sec  INFO : 2017-04-11 08:55:22,017 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 169.17 sec  INFO : 2017-04-11 08:55:24,017 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 176.83 sec  INFO : 2017-04-11 08:55:26,584 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 180.65 sec  INFO : 2017-04-11 08:55:27,843 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 192.5 sec  INFO : 2017-04-11 08:55:29,000 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 196.44 sec  INFO : 2017-04-11 08:55:30,415 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 223.12 sec  INFO : 2017-04-11 08:55:31,698 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 226.96 sec  INFO : 2017-04-11 08:55:33,215 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 234.53 sec  INFO : 2017-04-11 08:55:35,964 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 250.04 sec  INFO : 2017-04-11 08:55:37,455 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 253.9 sec  INFO : 2017-04-11 08:55:38,616 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 265.49 sec  INFO : 2017-04-11 08:55:39,769 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 286.71 sec  INFO : 2017-04-11 08:55:40,914 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 287.8 sec  INFO : MapReduce Total cumulative CPU time: 4 minutes 47 seconds 800 msec  INFO : Ended Job = job\_1490094830773\_0407  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 74 Reduce: 1 Cumulative CPU: 287.8 sec HDFS Read: 3616097164 HDFS Write: 19 SUCCESS  INFO : Total MapReduce CPU Time Spent: 4 minutes 47 seconds 800 msec  INFO : Completed executing command(queryId=hive\_20170411085454\_8d121f9a-a129-411b-85c9-93796440fbc1); Time taken: 62.306 seconds  INFO : OK |

## User wise Max Date

|  |
| --- |
| select username,max(date) from postimport\_text group by username;  INFO : Compiling command(queryId=hive\_20170411085757\_cf0c56ec-a26d-4d96-bfd1-df9ddfe698fc): select username,max(date) from postimport\_text group by username  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:username, type:string, comment:null), FieldSchema(name:\_c1, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411085757\_cf0c56ec-a26d-4d96-bfd1-df9ddfe698fc); Time taken: 0.123 seconds  INFO : Executing command(queryId=hive\_20170411085757\_cf0c56ec-a26d-4d96-bfd1-df9ddfe698fc): select username,max(date) from postimport\_text group by username  INFO : Query ID = hive\_20170411085757\_cf0c56ec-a26d-4d96-bfd1-df9ddfe698fc  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks not specified. Estimated from input data size: 54  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:74  INFO : Submitting tokens for job: job\_1490094830773\_0408  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0408/  INFO : Starting Job = job\_1490094830773\_0408, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0408/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0408  INFO : Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 54  INFO : 2017-04-11 08:57:35,810 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 08:57:45,850 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 30.66 sec  INFO : 2017-04-11 08:57:47,040 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 44.09 sec  INFO : 2017-04-11 08:57:48,204 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 57.37 sec  INFO : 2017-04-11 08:57:49,489 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 66.26 sec  INFO : 2017-04-11 08:57:55,457 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 83.45 sec  INFO : 2017-04-11 08:57:56,832 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 105.6 sec  INFO : 2017-04-11 08:57:58,069 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 118.95 sec  INFO : 2017-04-11 08:57:59,314 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 127.65 sec  INFO : 2017-04-11 08:58:00,611 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 132.06 sec  INFO : 2017-04-11 08:58:04,530 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 162.26 sec  INFO : 2017-04-11 08:58:06,041 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 171.05 sec  INFO : 2017-04-11 08:58:08,481 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 188.93 sec  INFO : 2017-04-11 08:58:11,505 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 197.56 sec  INFO : 2017-04-11 08:58:12,848 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 210.92 sec  INFO : 2017-04-11 08:58:14,027 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 228.26 sec  INFO : 2017-04-11 08:58:16,748 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 237.09 sec  INFO : 2017-04-11 08:58:17,918 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 245.69 sec  INFO : 2017-04-11 08:58:19,315 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 250.05 sec  INFO : 2017-04-11 08:58:20,668 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 271.74 sec  INFO : 2017-04-11 08:58:22,210 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 276.04 sec  INFO : 2017-04-11 08:58:23,477 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 293.07 sec  INFO : 2017-04-11 08:58:26,059 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 297.6 sec  INFO : 2017-04-11 08:58:27,460 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 309.13 sec  INFO : 2017-04-11 08:58:28,659 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 322.29 sec  INFO : 2017-04-11 08:58:29,943 Stage-1 map = 100%, reduce = 15%, Cumulative CPU 334.52 sec  INFO : 2017-04-11 08:58:32,765 Stage-1 map = 100%, reduce = 19%, Cumulative CPU 337.51 sec  INFO : 2017-04-11 08:58:33,936 Stage-1 map = 100%, reduce = 26%, Cumulative CPU 343.9 sec  INFO : 2017-04-11 08:58:35,185 Stage-1 map = 100%, reduce = 31%, Cumulative CPU 348.31 sec  INFO : 2017-04-11 08:58:36,602 Stage-1 map = 100%, reduce = 43%, Cumulative CPU 357.12 sec  INFO : 2017-04-11 08:58:39,340 Stage-1 map = 100%, reduce = 46%, Cumulative CPU 360.12 sec  INFO : 2017-04-11 08:58:40,557 Stage-1 map = 100%, reduce = 56%, Cumulative CPU 367.61 sec  INFO : 2017-04-11 08:58:42,393 Stage-1 map = 100%, reduce = 70%, Cumulative CPU 379.82 sec  INFO : 2017-04-11 08:58:43,663 Stage-1 map = 100%, reduce = 72%, Cumulative CPU 381.19 sec  INFO : 2017-04-11 08:58:45,053 Stage-1 map = 100%, reduce = 74%, Cumulative CPU 382.74 sec  INFO : 2017-04-11 08:58:46,406 Stage-1 map = 100%, reduce = 81%, Cumulative CPU 388.99 sec  INFO : 2017-04-11 08:58:47,754 Stage-1 map = 100%, reduce = 94%, Cumulative CPU 399.45 sec  INFO : 2017-04-11 08:58:48,896 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 403.64 sec  INFO : MapReduce Total cumulative CPU time: 6 minutes 43 seconds 640 msec  INFO : Ended Job = job\_1490094830773\_0408  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 74 Reduce: 54 Cumulative CPU: 403.64 sec HDFS Read: 3616313582 HDFS Write: 35573 SUCCESS  INFO : Total MapReduce CPU Time Spent: 6 minutes 43 seconds 640 msec  INFO : Completed executing command(queryId=hive\_20170411085757\_cf0c56ec-a26d-4d96-bfd1-df9ddfe698fc); Time taken: 85.79 seconds  INFO : OK |

## Select All Columns

|  |
| --- |
| select \* from postimport\_text;  INFO : Compiling command(queryId=hive\_20170411085959\_cd23f8ae-b896-473c-8600-cef4126702bb): select \* from postimport\_text  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_text.username, type:string, comment:null), FieldSchema(name:postimport\_text.machinename, type:string, comment:null), FieldSchema(name:postimport\_text.machineid, type:string, comment:null), FieldSchema(name:postimport\_text.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_text.domainname, type:string, comment:null), FieldSchema(name:postimport\_text.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_text.modulename, type:string, comment:null), FieldSchema(name:postimport\_text.filename, type:string, comment:null), FieldSchema(name:postimport\_text.productname, type:string, comment:null), FieldSchema(name:postimport\_text.companyname, type:string, comment:null), FieldSchema(name:postimport\_text.description, type:string, comment:null), FieldSchema(name:postimport\_text.text, type:string, comment:null), FieldSchema(name:postimport\_text.activetime, type:double, comment:null), FieldSchema(name:postimport\_text.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_text.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_text.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411085959\_cd23f8ae-b896-473c-8600-cef4126702bb); Time taken: 0.119 seconds  INFO : Executing command(queryId=hive\_20170411085959\_cd23f8ae-b896-473c-8600-cef4126702bb): select \* from postimport\_text  INFO : Query ID = hive\_20170411085959\_cd23f8ae-b896-473c-8600-cef4126702bb  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:74  INFO : Submitting tokens for job: job\_1490094830773\_0409  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0409/  INFO : Starting Job = job\_1490094830773\_0409, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0409/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0409  INFO : Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  INFO : 2017-04-11 08:59:46,375 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 08:59:57,281 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 42.62 sec  INFO : 2017-04-11 08:59:58,580 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 62.62 sec  INFO : 2017-04-11 08:59:59,891 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 93.68 sec  INFO : 2017-04-11 09:00:01,309 Stage-1 map = 6%, reduce = 0%, Cumulative CPU 114.03 sec  INFO : 2017-04-11 09:00:02,532 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 127.47 sec  INFO : 2017-04-11 09:00:03,725 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 141.87 sec  INFO : 2017-04-11 09:00:05,169 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 148.94 sec  INFO : 2017-04-11 09:00:06,339 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 172.03 sec  INFO : 2017-04-11 09:00:07,784 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 185.49 sec  INFO : 2017-04-11 09:00:09,001 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 193.99 sec  INFO : 2017-04-11 09:00:10,247 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 211.17 sec  INFO : 2017-04-11 09:00:12,736 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 222.32 sec  INFO : 2017-04-11 09:00:13,988 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 226.24 sec  INFO : 2017-04-11 09:00:16,691 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 236.12 sec  INFO : 2017-04-11 09:00:19,527 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 269.5 sec  INFO : 2017-04-11 09:00:22,056 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 289.4 sec  INFO : 2017-04-11 09:00:23,445 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 305.42 sec  INFO : 2017-04-11 09:00:24,609 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 318.0 sec  INFO : 2017-04-11 09:00:25,783 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 323.67 sec  INFO : 2017-04-11 09:00:28,371 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 344.77 sec  INFO : 2017-04-11 09:00:30,990 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 362.68 sec  INFO : 2017-04-11 09:00:33,563 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 373.04 sec  INFO : 2017-04-11 09:00:35,922 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 382.99 sec  INFO : 2017-04-11 09:00:38,420 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 402.87 sec  INFO : 2017-04-11 09:00:40,976 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 426.9 sec  INFO : 2017-04-11 09:00:42,135 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 431.93 sec  INFO : 2017-04-11 09:00:43,360 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 444.59 sec  INFO : 2017-04-11 09:00:45,857 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 459.13 sec  INFO : 2017-04-11 09:00:47,021 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 484.15 sec  INFO : 2017-04-11 09:00:48,351 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 487.93 sec  INFO : 2017-04-11 09:00:50,659 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 499.49 sec  INFO : 2017-04-11 09:00:53,169 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 511.78 sec  INFO : 2017-04-11 09:00:55,658 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 518.82 sec  INFO : 2017-04-11 09:00:59,283 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 545.93 sec  INFO : 2017-04-11 09:01:02,949 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 558.09 sec  INFO : 2017-04-11 09:01:05,425 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 574.07 sec  INFO : 2017-04-11 09:01:09,049 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 584.77 sec  INFO : 2017-04-11 09:01:11,602 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 595.22 sec  INFO : 2017-04-11 09:01:14,185 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 601.4 sec  INFO : 2017-04-11 09:01:17,863 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 625.35 sec  INFO : 2017-04-11 09:01:19,069 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 640.16 sec  INFO : 2017-04-11 09:01:20,425 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 650.89 sec  INFO : 2017-04-11 09:01:22,743 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 661.74 sec  INFO : 2017-04-11 09:01:24,081 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 674.62 sec  INFO : 2017-04-11 09:01:26,425 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 687.14 sec  INFO : 2017-04-11 09:01:27,781 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 695.31 sec  INFO : 2017-04-11 09:01:30,109 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 707.45 sec  INFO : 2017-04-11 09:01:31,451 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 714.34 sec  INFO : 2017-04-11 09:01:33,780 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 728.39 sec  INFO : 2017-04-11 09:01:36,299 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 743.48 sec  INFO : 2017-04-11 09:01:38,877 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 752.99 sec  INFO : 2017-04-11 09:01:42,698 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 766.51 sec  INFO : 2017-04-11 09:01:46,671 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 783.27 sec  INFO : 2017-04-11 09:01:47,852 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 793.77 sec  INFO : 2017-04-11 09:01:50,348 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 820.86 sec  INFO : 2017-04-11 09:01:52,862 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 841.32 sec  INFO : 2017-04-11 09:01:54,021 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 859.47 sec  INFO : 2017-04-11 09:01:56,579 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 878.84 sec  INFO : 2017-04-11 09:01:58,891 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 889.66 sec  INFO : 2017-04-11 09:02:00,227 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 898.0 sec  INFO : 2017-04-11 09:02:01,556 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 911.9 sec  INFO : 2017-04-11 09:02:04,223 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 929.02 sec  INFO : 2017-04-11 09:02:05,441 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 935.84 sec  INFO : 2017-04-11 09:02:07,984 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 945.32 sec  INFO : 2017-04-11 09:02:09,180 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 962.52 sec  INFO : 2017-04-11 09:02:10,531 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 975.63 sec  INFO : 2017-04-11 09:02:12,898 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 995.61 sec  INFO : 2017-04-11 09:02:14,259 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 1005.65 sec  INFO : 2017-04-11 09:02:15,434 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 1011.42 sec  INFO : 2017-04-11 09:02:17,950 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 1032.99 sec  INFO : 2017-04-11 09:02:19,112 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 1042.16 sec  INFO : 2017-04-11 09:02:21,683 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 1058.4 sec  INFO : 2017-04-11 09:02:22,840 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 1074.2 sec  INFO : 2017-04-11 09:02:25,544 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 1084.24 sec  INFO : 2017-04-11 09:02:27,964 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 1102.42 sec  INFO : 2017-04-11 09:02:29,344 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 1114.05 sec  INFO : 2017-04-11 09:02:31,760 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 1135.55 sec  INFO : 2017-04-11 09:02:34,375 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 1147.57 sec  INFO : 2017-04-11 09:02:36,961 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 1165.42 sec  INFO : 2017-04-11 09:02:38,194 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 1172.07 sec  INFO : 2017-04-11 09:02:39,536 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 1182.05 sec  INFO : 2017-04-11 09:02:40,701 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 1195.07 sec  INFO : 2017-04-11 09:02:43,194 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 1212.5 sec  INFO : 2017-04-11 09:02:45,512 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 1215.69 sec  INFO : 2017-04-11 09:02:49,153 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 1230.0 sec  INFO : 2017-04-11 09:02:51,643 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 1233.88 sec  INFO : 2017-04-11 09:02:55,305 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 1248.6 sec  INFO : 2017-04-11 09:03:01,376 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1256.81 sec  INFO : MapReduce Total cumulative CPU time: 20 minutes 56 seconds 810 msec  INFO : Ended Job = job\_1490094830773\_0409  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 74 Cumulative CPU: 1256.81 sec HDFS Read: 3616119356 HDFS Write: 3634005039 SUCCESS  INFO : Total MapReduce CPU Time Spent: 20 minutes 56 seconds 810 msec  INFO : Completed executing command(queryId=hive\_20170411085959\_cd23f8ae-b896-473c-8600-cef4126702bb); Time taken: 208.037 seconds  INFO : OK |

## Select All Columns with filter Condition

|  |
| --- |
| select \* from postimport\_text where username = '670063006100750074006F007500730065007200';  INFO : Compiling command(queryId=hive\_20170411090606\_5d10129c-f12d-420a-b667-ba2ae48ccd1d): select \* from postimport\_text where username = '670063006100750074006F007500730065007200'  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_text.username, type:string, comment:null), FieldSchema(name:postimport\_text.machinename, type:string, comment:null), FieldSchema(name:postimport\_text.machineid, type:string, comment:null), FieldSchema(name:postimport\_text.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_text.domainname, type:string, comment:null), FieldSchema(name:postimport\_text.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_text.modulename, type:string, comment:null), FieldSchema(name:postimport\_text.filename, type:string, comment:null), FieldSchema(name:postimport\_text.productname, type:string, comment:null), FieldSchema(name:postimport\_text.companyname, type:string, comment:null), FieldSchema(name:postimport\_text.description, type:string, comment:null), FieldSchema(name:postimport\_text.text, type:string, comment:null), FieldSchema(name:postimport\_text.activetime, type:double, comment:null), FieldSchema(name:postimport\_text.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_text.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_text.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411090606\_5d10129c-f12d-420a-b667-ba2ae48ccd1d); Time taken: 0.498 seconds  INFO : Executing command(queryId=hive\_20170411090606\_5d10129c-f12d-420a-b667-ba2ae48ccd1d): select \* from postimport\_text where username = '670063006100750074006F007500730065007200'  INFO : Query ID = hive\_20170411090606\_5d10129c-f12d-420a-b667-ba2ae48ccd1d  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:74  INFO : Submitting tokens for job: job\_1490094830773\_0411  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0411/  INFO : Starting Job = job\_1490094830773\_0411, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0411/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0411  INFO : Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  INFO : 2017-04-11 09:06:59,296 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 09:07:07,932 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 3.99 sec  INFO : 2017-04-11 09:07:08,980 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 12.31 sec  INFO : 2017-04-11 09:07:10,017 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 39.82 sec  INFO : 2017-04-11 09:07:11,075 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 58.97 sec  INFO : 2017-04-11 09:07:16,394 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 63.27 sec  INFO : 2017-04-11 09:07:17,423 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 71.19 sec  INFO : 2017-04-11 09:07:18,454 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 78.69 sec  INFO : 2017-04-11 09:07:19,487 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 93.21 sec  INFO : 2017-04-11 09:07:20,528 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 112.83 sec  INFO : 2017-04-11 09:07:21,558 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 116.8 sec  INFO : 2017-04-11 09:07:23,704 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 120.69 sec  INFO : 2017-04-11 09:07:25,837 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 128.7 sec  INFO : 2017-04-11 09:07:26,896 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 132.66 sec  INFO : 2017-04-11 09:07:28,996 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 158.45 sec  INFO : 2017-04-11 09:07:30,040 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 162.11 sec  INFO : 2017-04-11 09:07:31,079 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 177.77 sec  INFO : 2017-04-11 09:07:33,203 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 181.46 sec  INFO : 2017-04-11 09:07:34,297 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 185.42 sec  INFO : 2017-04-11 09:07:35,355 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 189.46 sec  INFO : 2017-04-11 09:07:38,477 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 216.96 sec  INFO : 2017-04-11 09:07:39,527 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 224.9 sec  INFO : 2017-04-11 09:07:40,572 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 240.43 sec  INFO : 2017-04-11 09:07:44,678 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 256.81 sec  INFO : 2017-04-11 09:07:46,731 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 282.51 sec  INFO : MapReduce Total cumulative CPU time: 4 minutes 42 seconds 510 msec  INFO : Ended Job = job\_1490094830773\_0411  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 74 Cumulative CPU: 282.51 sec HDFS Read: 3616153470 HDFS Write: 2118870 SUCCESS  INFO : Total MapReduce CPU Time Spent: 4 minutes 42 seconds 510 msec  INFO : Completed executing command(queryId=hive\_20170411090606\_5d10129c-f12d-420a-b667-ba2ae48ccd1d); Time taken: 57.103 seconds  INFO : OK |

# Plain Text Format Bzip2 compression

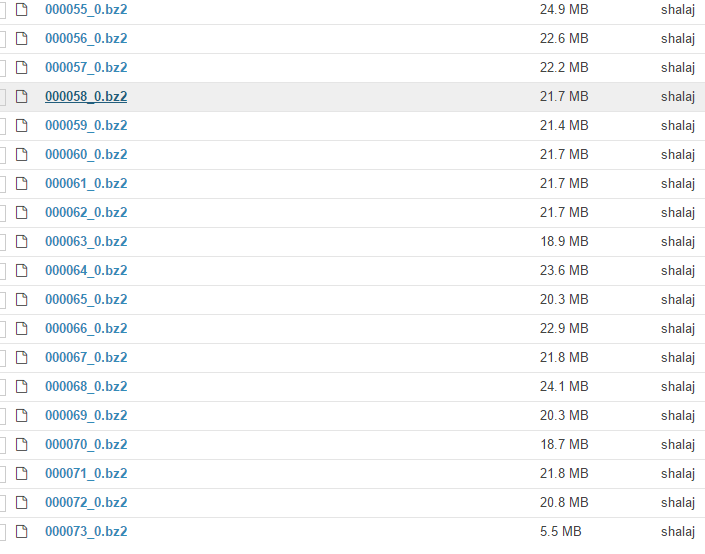
|  |
| --- |
| SET hive.exec.compress.output=true;  SET mapred.output.compression.type=BLOCK;  set mapred.output.compress=true;  set mapred.output.compression.codec=org.apache.hadoop.io.compress.BZip2Codec;  set io.compression.codecs=org.apache.hadoop.io.compress.BZip2Codec; |

## Load data to Plain Text format

|  |
| --- |
| hive> insert overwrite table postimport\_text select \* from postimportwithoutpart;  hive> insert overwrite table postimport\_text select \* from postimportwithoutpart;  Query ID = shalaj\_20170411090909\_874fc2bb-693d-4bb7-a8da-677fe45dc1e0  Total jobs = 3  Launching Job 1 out of 3  Number of reduce tasks is set to 0 since there's no reduce operator  Starting Job = job\_1490094830773\_0412, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0412/  Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0412  Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  2017-04-11 09:09:30,813 Stage-1 map = 0%, reduce = 0%  2017-04-11 09:09:57,041 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 257.55 sec  2017-04-11 09:09:58,090 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 282.37 sec  2017-04-11 09:09:59,134 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 290.81 sec  2017-04-11 09:10:01,231 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 329.98 sec  2017-04-11 09:10:03,563 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 355.09 sec  2017-04-11 09:10:04,629 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 374.39 sec  2017-04-11 09:10:12,012 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 468.88 sec  2017-04-11 09:10:14,082 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 505.32 sec  2017-04-11 09:10:15,124 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 509.27 sec  2017-04-11 09:10:17,209 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 533.22 sec  2017-04-11 09:10:18,250 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 534.57 sec  2017-04-11 09:10:21,371 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 550.45 sec  2017-04-11 09:10:37,105 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 766.94 sec  2017-04-11 09:10:39,222 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 792.65 sec  2017-04-11 09:10:41,320 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 823.91 sec  2017-04-11 09:10:42,372 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 834.46 sec  2017-04-11 09:10:46,556 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 898.12 sec  2017-04-11 09:10:47,616 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 915.07 sec  2017-04-11 09:10:50,742 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 957.9 sec  2017-04-11 09:10:52,813 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 982.92 sec  2017-04-11 09:10:56,994 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 1036.46 sec  2017-04-11 09:10:58,033 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 1048.89 sec  2017-04-11 09:11:02,178 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 1081.46 sec  2017-04-11 09:11:05,282 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 1119.6 sec  2017-04-11 09:11:06,328 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 1134.55 sec  2017-04-11 09:11:08,403 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 1149.12 sec  2017-04-11 09:11:09,462 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 1173.38 sec  2017-04-11 09:11:18,945 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 1284.91 sec  2017-04-11 09:11:25,222 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 1384.65 sec  2017-04-11 09:11:29,612 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 1436.18 sec  2017-04-11 09:11:31,744 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 1470.4 sec  2017-04-11 09:11:32,807 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 1480.73 sec  2017-04-11 09:11:34,886 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 1514.56 sec  2017-04-11 09:11:36,974 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 1537.08 sec  2017-04-11 09:11:38,023 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 1548.66 sec  2017-04-11 09:11:39,061 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 1553.67 sec  2017-04-11 09:11:41,195 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 1584.01 sec  2017-04-11 09:11:45,365 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 1632.09 sec  2017-04-11 09:11:49,510 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 1684.92 sec  2017-04-11 09:11:50,539 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 1699.26 sec  2017-04-11 09:11:53,648 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 1737.43 sec  2017-04-11 09:11:55,707 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 1766.59 sec  2017-04-11 09:11:56,742 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 1772.38 sec  2017-04-11 09:11:58,810 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 1792.0 sec  2017-04-11 09:11:59,840 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 1808.06 sec  2017-04-11 09:12:07,089 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 1910.21 sec  2017-04-11 09:12:08,139 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 1916.03 sec  2017-04-11 09:12:09,177 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 1934.17 sec  2017-04-11 09:12:12,283 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 1979.62 sec  2017-04-11 09:12:13,344 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 1993.1 sec  2017-04-11 09:12:14,388 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 1998.12 sec  2017-04-11 09:12:18,612 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 2045.07 sec  2017-04-11 09:12:20,762 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 2075.68 sec  2017-04-11 09:12:22,957 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 2111.38 sec  2017-04-11 09:12:24,042 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 2127.45 sec  2017-04-11 09:12:26,210 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 2148.54 sec  2017-04-11 09:12:29,650 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 2179.77 sec  2017-04-11 09:12:34,047 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 2246.33 sec  2017-04-11 09:12:36,450 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 2275.93 sec  2017-04-11 09:12:39,146 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 2314.13 sec  2017-04-11 09:12:40,351 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 2322.58 sec  2017-04-11 09:12:42,574 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 2349.03 sec  2017-04-11 09:12:43,669 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 2367.09 sec  2017-04-11 09:12:47,216 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 2410.63 sec  2017-04-11 09:12:48,309 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 2422.54 sec  2017-04-11 09:12:50,494 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 2448.08 sec  2017-04-11 09:12:52,662 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 2470.98 sec  2017-04-11 09:12:53,746 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 2502.17 sec  2017-04-11 09:12:54,818 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 2509.51 sec  2017-04-11 09:12:58,998 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 2552.05 sec  2017-04-11 09:13:01,076 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 2575.9 sec  2017-04-11 09:13:04,180 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 2606.44 sec  2017-04-11 09:13:06,232 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 2628.25 sec  2017-04-11 09:13:07,258 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 2636.38 sec  2017-04-11 09:13:09,316 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 2651.23 sec  2017-04-11 09:13:14,469 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 2676.41 sec  2017-04-11 09:13:17,537 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 2691.31 sec  2017-04-11 09:13:22,651 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 2709.56 sec  2017-04-11 09:13:23,668 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2711.5 sec  MapReduce Total cumulative CPU time: 45 minutes 11 seconds 500 msec  Ended Job = job\_1490094830773\_0412  Stage-4 is selected by condition resolver.  Stage-3 is filtered out by condition resolver.  Stage-5 is filtered out by condition resolver.  Moving data to: hdfs://mac55:8020/user/hive/warehouse/esplus.db/postimport\_text/.hive-staging\_hive\_2017-04-11\_09-09-20\_990\_377647128038960892-1/-ext-10000  Loading data to table esplus.postimport\_text  Table esplus.postimport\_text stats: [numFiles=74, numRows=42286538, totalSize=1732828647, rawDataSize=19438765021]  MapReduce Jobs Launched:  Stage-Stage-1: Map: 74 Cumulative CPU: 2711.5 sec HDFS Read: 19486300571 HDFS Write: 1732835232 SUCCESS  Total MapReduce CPU Time Spent: 45 minutes 11 seconds 500 msec  OK  Time taken: 246.877 seconds |

## Data Size

|  |
| --- |
| hive> dfs -du -s -h /user/hive/warehouse/esplus.db/postimport\_text;  1.6 G 4.8 G /user/hive/warehouse/esplus.db/postimport\_text |



Created 74 bz2 files

## Select some columns

|  |
| --- |
| select productname,username from postimport\_text;  INFO : Compiling command(queryId=hive\_20170411091616\_0b8f8953-bdca-4c61-81c4-cdbb41439dc0): select productname,username from postimport\_text  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:productname, type:string, comment:null), FieldSchema(name:username, type:string, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411091616\_0b8f8953-bdca-4c61-81c4-cdbb41439dc0); Time taken: 0.13 seconds  INFO : Executing command(queryId=hive\_20170411091616\_0b8f8953-bdca-4c61-81c4-cdbb41439dc0): select productname,username from postimport\_text  INFO : Query ID = hive\_20170411091616\_0b8f8953-bdca-4c61-81c4-cdbb41439dc0  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:74  INFO : Submitting tokens for job: job\_1490094830773\_0413  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0413/  INFO : Starting Job = job\_1490094830773\_0413, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0413/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0413  INFO : Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  INFO : 2017-04-11 09:16:13,501 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 09:16:25,055 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 18.0 sec  INFO : 2017-04-11 09:16:26,095 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 63.77 sec  INFO : 2017-04-11 09:16:28,175 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 94.97 sec  INFO : 2017-04-11 09:16:29,216 Stage-1 map = 6%, reduce = 0%, Cumulative CPU 118.33 sec  INFO : 2017-04-11 09:16:31,286 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 138.64 sec  INFO : 2017-04-11 09:16:32,327 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 161.83 sec  INFO : 2017-04-11 09:16:33,369 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 179.39 sec  INFO : 2017-04-11 09:16:35,448 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 205.74 sec  INFO : 2017-04-11 09:16:36,489 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 223.43 sec  INFO : 2017-04-11 09:16:37,530 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 238.08 sec  INFO : 2017-04-11 09:16:38,570 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 251.01 sec  INFO : 2017-04-11 09:16:39,608 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 268.67 sec  INFO : 2017-04-11 09:16:40,647 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 280.97 sec  INFO : 2017-04-11 09:16:41,688 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 292.58 sec  INFO : 2017-04-11 09:16:43,766 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 318.78 sec  INFO : 2017-04-11 09:16:44,797 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 332.93 sec  INFO : 2017-04-11 09:16:45,825 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 345.96 sec  INFO : 2017-04-11 09:16:48,969 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 361.85 sec  INFO : 2017-04-11 09:16:53,113 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 386.19 sec  INFO : 2017-04-11 09:16:55,187 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 424.79 sec  INFO : 2017-04-11 09:16:56,226 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 433.46 sec  INFO : 2017-04-11 09:16:57,260 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 456.65 sec  INFO : 2017-04-11 09:16:58,301 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 470.9 sec  INFO : 2017-04-11 09:16:59,338 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 488.9 sec  INFO : 2017-04-11 09:17:00,375 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 506.35 sec  INFO : 2017-04-11 09:17:01,415 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 521.07 sec  INFO : 2017-04-11 09:17:02,451 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 535.82 sec  INFO : 2017-04-11 09:17:03,490 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 553.19 sec  INFO : 2017-04-11 09:17:04,524 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 566.49 sec  INFO : 2017-04-11 09:17:05,567 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 581.1 sec  INFO : 2017-04-11 09:17:06,609 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 601.5 sec  INFO : 2017-04-11 09:17:08,683 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 624.23 sec  INFO : 2017-04-11 09:17:09,737 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 638.58 sec  INFO : 2017-04-11 09:17:10,772 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 647.26 sec  INFO : 2017-04-11 09:17:11,816 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 661.53 sec  INFO : 2017-04-11 09:17:12,855 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 676.97 sec  INFO : 2017-04-11 09:17:14,931 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 697.39 sec  INFO : 2017-04-11 09:17:16,986 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 716.88 sec  INFO : 2017-04-11 09:17:20,118 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 735.72 sec  INFO : 2017-04-11 09:17:22,208 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 749.34 sec  INFO : 2017-04-11 09:17:23,257 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 781.63 sec  INFO : 2017-04-11 09:17:25,330 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 797.38 sec  INFO : 2017-04-11 09:17:26,366 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 817.44 sec  INFO : 2017-04-11 09:17:27,403 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 831.9 sec  INFO : 2017-04-11 09:17:28,438 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 849.22 sec  INFO : 2017-04-11 09:17:29,476 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 866.82 sec  INFO : 2017-04-11 09:17:30,511 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 883.44 sec  INFO : 2017-04-11 09:17:32,582 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 915.49 sec  INFO : 2017-04-11 09:17:34,664 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 937.86 sec  INFO : 2017-04-11 09:17:35,734 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 955.13 sec  INFO : 2017-04-11 09:17:36,779 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 964.04 sec  INFO : 2017-04-11 09:17:38,848 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 990.1 sec  INFO : 2017-04-11 09:17:39,885 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 1008.99 sec  INFO : 2017-04-11 09:17:40,922 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 1024.92 sec  INFO : 2017-04-11 09:17:41,949 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 1029.13 sec  INFO : 2017-04-11 09:17:42,986 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 1045.6 sec  INFO : 2017-04-11 09:17:45,065 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 1064.1 sec  INFO : 2017-04-11 09:17:47,150 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 1085.6 sec  INFO : 2017-04-11 09:17:48,175 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 1097.54 sec  INFO : 2017-04-11 09:17:51,278 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 1130.88 sec  INFO : 2017-04-11 09:17:52,316 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 1151.69 sec  INFO : 2017-04-11 09:17:55,426 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 1182.56 sec  INFO : 2017-04-11 09:17:56,454 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 1183.76 sec  INFO : 2017-04-11 09:17:57,526 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 1201.11 sec  INFO : 2017-04-11 09:17:58,566 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 1232.76 sec  INFO : 2017-04-11 09:17:59,608 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 1242.0 sec  INFO : 2017-04-11 09:18:00,648 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 1256.41 sec  INFO : 2017-04-11 09:18:01,682 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 1270.92 sec  INFO : 2017-04-11 09:18:03,753 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 1291.34 sec  INFO : 2017-04-11 09:18:04,794 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 1305.98 sec  INFO : 2017-04-11 09:18:06,862 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 1329.45 sec  INFO : 2017-04-11 09:18:07,902 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 1343.62 sec  INFO : 2017-04-11 09:18:08,926 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 1370.32 sec  INFO : 2017-04-11 09:18:09,980 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 1378.64 sec  INFO : 2017-04-11 09:18:11,034 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 1390.3 sec  INFO : 2017-04-11 09:18:12,074 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 1403.92 sec  INFO : 2017-04-11 09:18:14,166 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 1422.29 sec  INFO : 2017-04-11 09:18:15,206 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 1435.53 sec  INFO : 2017-04-11 09:18:17,257 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 1452.27 sec  INFO : 2017-04-11 09:18:18,299 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 1476.65 sec  INFO : 2017-04-11 09:18:20,395 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 1501.9 sec  INFO : 2017-04-11 09:18:21,440 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 1513.73 sec  INFO : 2017-04-11 09:18:23,545 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 1527.83 sec  INFO : 2017-04-11 09:18:25,662 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 1548.19 sec  INFO : 2017-04-11 09:18:26,696 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 1565.78 sec  INFO : 2017-04-11 09:18:27,739 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 1577.63 sec  INFO : 2017-04-11 09:18:28,773 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 1592.42 sec  INFO : 2017-04-11 09:18:30,845 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 1610.08 sec  INFO : 2017-04-11 09:18:31,879 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 1624.78 sec  INFO : 2017-04-11 09:18:33,950 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 1640.75 sec  INFO : 2017-04-11 09:18:34,998 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 1655.72 sec  INFO : 2017-04-11 09:18:36,033 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 1661.73 sec  INFO : 2017-04-11 09:18:38,095 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 1675.78 sec  INFO : 2017-04-11 09:18:44,253 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1692.09 sec  INFO : MapReduce Total cumulative CPU time: 28 minutes 12 seconds 90 msec  INFO : Ended Job = job\_1490094830773\_0413  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 74 Cumulative CPU: 1692.09 sec HDFS Read: 1733792645 HDFS Write: 233687022 SUCCESS  INFO : Total MapReduce CPU Time Spent: 28 minutes 12 seconds 90 msec  INFO : Completed executing command(queryId=hive\_20170411091616\_0b8f8953-bdca-4c61-81c4-cdbb41439dc0); Time taken: 157.395 seconds  INFO : OK |

## Check Max Date

|  |
| --- |
| select max(date) from postimport\_text;  INFO : Compiling command(queryId=hive\_20170411091919\_38e94161-5ac3-4f74-85bf-8f64eb4995b8): select max(date) from postimport\_text  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:\_c0, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411091919\_38e94161-5ac3-4f74-85bf-8f64eb4995b8); Time taken: 0.137 seconds  INFO : Executing command(queryId=hive\_20170411091919\_38e94161-5ac3-4f74-85bf-8f64eb4995b8): select max(date) from postimport\_text  INFO : Query ID = hive\_20170411091919\_38e94161-5ac3-4f74-85bf-8f64eb4995b8  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks determined at compile time: 1  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:74  INFO : Submitting tokens for job: job\_1490094830773\_0414  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0414/  INFO : Starting Job = job\_1490094830773\_0414, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0414/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0414  INFO : Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 1  INFO : 2017-04-11 09:19:18,104 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 09:19:29,591 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 19.26 sec  INFO : 2017-04-11 09:19:30,629 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 63.49 sec  INFO : 2017-04-11 09:19:31,663 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 86.46 sec  INFO : 2017-04-11 09:19:33,742 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 118.42 sec  INFO : 2017-04-11 09:19:35,813 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 142.31 sec  INFO : 2017-04-11 09:19:36,843 Stage-1 map = 6%, reduce = 0%, Cumulative CPU 162.43 sec  INFO : 2017-04-11 09:19:38,913 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 186.2 sec  INFO : 2017-04-11 09:19:39,950 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 206.51 sec  INFO : 2017-04-11 09:19:40,982 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 227.09 sec  INFO : 2017-04-11 09:19:43,059 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 250.31 sec  INFO : 2017-04-11 09:19:44,093 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 267.78 sec  INFO : 2017-04-11 09:19:45,132 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 280.44 sec  INFO : 2017-04-11 09:19:46,165 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 293.51 sec  INFO : 2017-04-11 09:19:48,226 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 316.48 sec  INFO : 2017-04-11 09:19:49,253 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 330.49 sec  INFO : 2017-04-11 09:19:50,287 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 352.83 sec  INFO : 2017-04-11 09:19:51,355 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 354.64 sec  INFO : 2017-04-11 09:19:53,436 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 357.57 sec  INFO : 2017-04-11 09:19:58,640 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 384.87 sec  INFO : 2017-04-11 09:20:00,706 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 415.32 sec  INFO : 2017-04-11 09:20:02,770 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 472.61 sec  INFO : 2017-04-11 09:20:03,811 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 478.55 sec  INFO : 2017-04-11 09:20:05,881 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 518.61 sec  INFO : 2017-04-11 09:20:06,915 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 525.84 sec  INFO : 2017-04-11 09:20:08,984 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 560.66 sec  INFO : 2017-04-11 09:20:11,048 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 580.67 sec  INFO : 2017-04-11 09:20:12,083 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 600.91 sec  INFO : 2017-04-11 09:20:13,119 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 606.73 sec  INFO : 2017-04-11 09:20:15,184 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 636.77 sec  INFO : 2017-04-11 09:20:16,220 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 649.02 sec  INFO : 2017-04-11 09:20:17,266 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 663.06 sec  INFO : 2017-04-11 09:20:18,304 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 675.49 sec  INFO : 2017-04-11 09:20:19,331 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 682.53 sec  INFO : 2017-04-11 09:20:20,382 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 694.3 sec  INFO : 2017-04-11 09:20:21,428 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 712.6 sec  INFO : 2017-04-11 09:20:22,483 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 719.71 sec  INFO : 2017-04-11 09:20:23,531 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 722.72 sec  INFO : 2017-04-11 09:20:28,718 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 758.43 sec  INFO : 2017-04-11 09:20:29,751 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 784.74 sec  INFO : 2017-04-11 09:20:32,856 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 824.29 sec  INFO : 2017-04-11 09:20:33,889 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 858.14 sec  INFO : 2017-04-11 09:20:35,958 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 878.8 sec  INFO : 2017-04-11 09:20:36,991 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 899.28 sec  INFO : 2017-04-11 09:20:40,099 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 934.21 sec  INFO : 2017-04-11 09:20:42,191 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 969.5 sec  INFO : 2017-04-11 09:20:43,224 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 974.98 sec  INFO : 2017-04-11 09:20:44,255 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 980.22 sec  INFO : 2017-04-11 09:20:45,287 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 1014.23 sec  INFO : 2017-04-11 09:20:48,394 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 1051.78 sec  INFO : 2017-04-11 09:20:49,427 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 1055.76 sec  INFO : 2017-04-11 09:20:51,539 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 1076.19 sec  INFO : 2017-04-11 09:20:52,579 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 1082.74 sec  INFO : 2017-04-11 09:20:56,733 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 1110.29 sec  INFO : 2017-04-11 09:20:57,769 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 1119.07 sec  INFO : 2017-04-11 09:20:58,800 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 1151.75 sec  INFO : 2017-04-11 09:20:59,874 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 1163.6 sec  INFO : 2017-04-11 09:21:01,949 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 1184.22 sec  INFO : 2017-04-11 09:21:02,981 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 1207.48 sec  INFO : 2017-04-11 09:21:05,055 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 1234.07 sec  INFO : 2017-04-11 09:21:06,091 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 1250.02 sec  INFO : 2017-04-11 09:21:08,156 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 1270.12 sec  INFO : 2017-04-11 09:21:10,221 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 1300.34 sec  INFO : 2017-04-11 09:21:11,252 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 1317.38 sec  INFO : 2017-04-11 09:21:13,328 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 1341.27 sec  INFO : 2017-04-11 09:21:14,362 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 1358.83 sec  INFO : 2017-04-11 09:21:15,387 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 1381.43 sec  INFO : 2017-04-11 09:21:16,412 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 1389.24 sec  INFO : 2017-04-11 09:21:17,456 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 1400.1 sec  INFO : 2017-04-11 09:21:18,524 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 1417.62 sec  INFO : 2017-04-11 09:21:20,620 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 1430.05 sec  INFO : 2017-04-11 09:21:22,692 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 1444.27 sec  INFO : 2017-04-11 09:21:24,751 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 1457.41 sec  INFO : 2017-04-11 09:21:27,855 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 1506.2 sec  INFO : 2017-04-11 09:21:29,937 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 1521.3 sec  INFO : 2017-04-11 09:21:32,008 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 1547.89 sec  INFO : 2017-04-11 09:21:33,047 Stage-1 map = 92%, reduce = 29%, Cumulative CPU 1558.5 sec  INFO : 2017-04-11 09:21:35,110 Stage-1 map = 93%, reduce = 29%, Cumulative CPU 1576.25 sec  INFO : 2017-04-11 09:21:38,211 Stage-1 map = 94%, reduce = 29%, Cumulative CPU 1602.92 sec  INFO : 2017-04-11 09:21:39,247 Stage-1 map = 96%, reduce = 29%, Cumulative CPU 1615.4 sec  INFO : 2017-04-11 09:21:41,314 Stage-1 map = 96%, reduce = 31%, Cumulative CPU 1624.2 sec  INFO : 2017-04-11 09:21:42,346 Stage-1 map = 98%, reduce = 31%, Cumulative CPU 1633.15 sec  INFO : 2017-04-11 09:21:43,368 Stage-1 map = 99%, reduce = 31%, Cumulative CPU 1639.63 sec  INFO : 2017-04-11 09:21:44,391 Stage-1 map = 100%, reduce = 32%, Cumulative CPU 1643.4 sec  INFO : 2017-04-11 09:21:45,440 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 1644.2 sec  INFO : MapReduce Total cumulative CPU time: 27 minutes 24 seconds 200 msec  INFO : Ended Job = job\_1490094830773\_0414  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 74 Reduce: 1 Cumulative CPU: 1644.2 sec HDFS Read: 1733842444 HDFS Write: 19 SUCCESS  INFO : Total MapReduce CPU Time Spent: 27 minutes 24 seconds 200 msec  INFO : Completed executing command(queryId=hive\_20170411091919\_38e94161-5ac3-4f74-85bf-8f64eb4995b8); Time taken: 154.203 seconds  INFO : OK |

## User wise Max Date

|  |
| --- |
| select username,max(date) from postimport\_text group by username;  INFO : Compiling command(queryId=hive\_20170411092323\_461ed8d6-c6a0-40b5-9f89-929258039d71): select username,max(date) from postimport\_text group by username  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:username, type:string, comment:null), FieldSchema(name:\_c1, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411092323\_461ed8d6-c6a0-40b5-9f89-929258039d71); Time taken: 0.108 seconds  INFO : Executing command(queryId=hive\_20170411092323\_461ed8d6-c6a0-40b5-9f89-929258039d71): select username,max(date) from postimport\_text group by username  INFO : Query ID = hive\_20170411092323\_461ed8d6-c6a0-40b5-9f89-929258039d71  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks not specified. Estimated from input data size: 26  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:74  INFO : Submitting tokens for job: job\_1490094830773\_0415  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0415/  INFO : Starting Job = job\_1490094830773\_0415, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0415/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0415  INFO : Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 26  INFO : 2017-04-11 09:24:06,604 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 09:24:17,116 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 18.78 sec  INFO : 2017-04-11 09:24:19,186 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 63.35 sec  INFO : 2017-04-11 09:24:20,228 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 95.54 sec  INFO : 2017-04-11 09:24:22,289 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 118.59 sec  INFO : 2017-04-11 09:24:23,326 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 139.39 sec  INFO : 2017-04-11 09:24:25,399 Stage-1 map = 6%, reduce = 0%, Cumulative CPU 162.22 sec  INFO : 2017-04-11 09:24:27,470 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 202.97 sec  INFO : 2017-04-11 09:24:28,504 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 211.67 sec  INFO : 2017-04-11 09:24:30,573 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 251.16 sec  INFO : 2017-04-11 09:24:33,673 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 291.81 sec  INFO : 2017-04-11 09:24:34,698 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 304.23 sec  INFO : 2017-04-11 09:24:36,771 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 330.22 sec  INFO : 2017-04-11 09:24:37,810 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 343.32 sec  INFO : 2017-04-11 09:24:38,845 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 352.0 sec  INFO : 2017-04-11 09:24:39,901 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 369.39 sec  INFO : 2017-04-11 09:24:43,044 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 378.73 sec  INFO : 2017-04-11 09:24:49,262 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 436.02 sec  INFO : 2017-04-11 09:24:51,327 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 471.44 sec  INFO : 2017-04-11 09:24:52,364 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 494.9 sec  INFO : 2017-04-11 09:24:54,430 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 518.25 sec  INFO : 2017-04-11 09:24:55,466 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 538.83 sec  INFO : 2017-04-11 09:24:56,505 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 552.21 sec  INFO : 2017-04-11 09:24:58,571 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 581.21 sec  INFO : 2017-04-11 09:24:59,604 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 599.5 sec  INFO : 2017-04-11 09:25:00,627 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 605.38 sec  INFO : 2017-04-11 09:25:02,696 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 633.96 sec  INFO : 2017-04-11 09:25:04,766 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 654.28 sec  INFO : 2017-04-11 09:25:05,789 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 679.49 sec  INFO : 2017-04-11 09:25:06,815 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 693.14 sec  INFO : 2017-04-11 09:25:07,850 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 695.43 sec  INFO : 2017-04-11 09:25:08,886 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 728.67 sec  INFO : 2017-04-11 09:25:09,920 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 729.89 sec  INFO : 2017-04-11 09:25:10,968 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 735.25 sec  INFO : 2017-04-11 09:25:12,017 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 746.79 sec  INFO : 2017-04-11 09:25:18,250 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 794.07 sec  INFO : 2017-04-11 09:25:19,275 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 802.94 sec  INFO : 2017-04-11 09:25:21,339 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 841.06 sec  INFO : 2017-04-11 09:25:22,375 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 867.12 sec  INFO : 2017-04-11 09:25:23,413 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 874.78 sec  INFO : 2017-04-11 09:25:24,445 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 896.39 sec  INFO : 2017-04-11 09:25:25,479 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 910.77 sec  INFO : 2017-04-11 09:25:27,563 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 930.97 sec  INFO : 2017-04-11 09:25:29,632 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 954.31 sec  INFO : 2017-04-11 09:25:31,704 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 986.11 sec  INFO : 2017-04-11 09:25:33,779 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 1009.42 sec  INFO : 2017-04-11 09:25:34,823 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 1037.73 sec  INFO : 2017-04-11 09:25:35,862 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 1051.57 sec  INFO : 2017-04-11 09:25:36,902 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 1067.01 sec  INFO : 2017-04-11 09:25:37,939 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 1081.33 sec  INFO : 2017-04-11 09:25:38,964 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 1084.8 sec  INFO : 2017-04-11 09:25:41,090 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 1107.48 sec  INFO : 2017-04-11 09:25:44,188 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 1123.97 sec  INFO : 2017-04-11 09:25:47,284 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 1150.68 sec  INFO : 2017-04-11 09:25:48,316 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 1171.7 sec  INFO : 2017-04-11 09:25:49,347 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 1193.44 sec  INFO : 2017-04-11 09:25:51,416 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 1219.07 sec  INFO : 2017-04-11 09:25:53,488 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 1242.71 sec  INFO : 2017-04-11 09:25:55,553 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 1271.79 sec  INFO : 2017-04-11 09:25:57,622 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 1297.87 sec  INFO : 2017-04-11 09:25:58,654 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 1321.24 sec  INFO : 2017-04-11 09:25:59,685 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 1333.6 sec  INFO : 2017-04-11 09:26:01,753 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 1367.93 sec  INFO : 2017-04-11 09:26:03,809 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 1388.22 sec  INFO : 2017-04-11 09:26:05,885 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 1412.17 sec  INFO : 2017-04-11 09:26:06,936 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 1428.55 sec  INFO : 2017-04-11 09:26:07,973 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 1450.45 sec  INFO : 2017-04-11 09:26:10,051 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 1460.9 sec  INFO : 2017-04-11 09:26:11,115 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 1473.42 sec  INFO : 2017-04-11 09:26:12,176 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 1479.22 sec  INFO : 2017-04-11 09:26:13,217 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 1487.23 sec  INFO : 2017-04-11 09:26:14,245 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 1498.46 sec  INFO : 2017-04-11 09:26:17,342 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 1524.06 sec  INFO : 2017-04-11 09:26:20,433 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 1564.59 sec  INFO : 2017-04-11 09:26:21,459 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 1573.34 sec  INFO : 2017-04-11 09:26:22,495 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 1586.68 sec  INFO : 2017-04-11 09:26:23,525 Stage-1 map = 91%, reduce = 2%, Cumulative CPU 1599.02 sec  INFO : 2017-04-11 09:26:24,561 Stage-1 map = 91%, reduce = 4%, Cumulative CPU 1611.27 sec  INFO : 2017-04-11 09:26:25,586 Stage-1 map = 92%, reduce = 4%, Cumulative CPU 1620.12 sec  INFO : 2017-04-11 09:26:27,656 Stage-1 map = 93%, reduce = 4%, Cumulative CPU 1642.19 sec  INFO : 2017-04-11 09:26:28,692 Stage-1 map = 94%, reduce = 4%, Cumulative CPU 1650.8 sec  INFO : 2017-04-11 09:26:30,766 Stage-1 map = 94%, reduce = 6%, Cumulative CPU 1668.41 sec  INFO : 2017-04-11 09:26:31,797 Stage-1 map = 95%, reduce = 6%, Cumulative CPU 1677.69 sec  INFO : 2017-04-11 09:26:32,823 Stage-1 map = 97%, reduce = 6%, Cumulative CPU 1688.71 sec  INFO : 2017-04-11 09:26:34,892 Stage-1 map = 99%, reduce = 6%, Cumulative CPU 1700.25 sec  INFO : 2017-04-11 09:26:35,935 Stage-1 map = 100%, reduce = 6%, Cumulative CPU 1703.14 sec  INFO : 2017-04-11 09:26:36,979 Stage-1 map = 100%, reduce = 18%, Cumulative CPU 1708.78 sec  INFO : 2017-04-11 09:26:38,015 Stage-1 map = 100%, reduce = 35%, Cumulative CPU 1716.34 sec  INFO : 2017-04-11 09:26:39,056 Stage-1 map = 100%, reduce = 42%, Cumulative CPU 1719.67 sec  INFO : 2017-04-11 09:26:40,089 Stage-1 map = 100%, reduce = 46%, Cumulative CPU 1721.37 sec  INFO : 2017-04-11 09:26:41,115 Stage-1 map = 100%, reduce = 58%, Cumulative CPU 1726.39 sec  INFO : 2017-04-11 09:26:42,137 Stage-1 map = 100%, reduce = 65%, Cumulative CPU 1729.48 sec  INFO : 2017-04-11 09:26:43,158 Stage-1 map = 100%, reduce = 88%, Cumulative CPU 1739.57 sec  INFO : 2017-04-11 09:26:44,179 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 1744.39 sec  INFO : MapReduce Total cumulative CPU time: 29 minutes 4 seconds 390 msec  INFO : Ended Job = job\_1490094830773\_0415  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 74 Reduce: 26 Cumulative CPU: 1744.39 sec HDFS Read: 1733950525 HDFS Write: 31725 SUCCESS  INFO : Total MapReduce CPU Time Spent: 29 minutes 4 seconds 390 msec  INFO : Completed executing command(queryId=hive\_20170411092323\_461ed8d6-c6a0-40b5-9f89-929258039d71); Time taken: 166.282 seconds  INFO : OK |

## Select All Columns

|  |
| --- |
| select \* from postimport\_text;  INFO : Compiling command(queryId=hive\_20170411092727\_ccf75124-570a-441d-919e-714c454e4f77): select \* from postimport\_text  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_text.username, type:string, comment:null), FieldSchema(name:postimport\_text.machinename, type:string, comment:null), FieldSchema(name:postimport\_text.machineid, type:string, comment:null), FieldSchema(name:postimport\_text.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_text.domainname, type:string, comment:null), FieldSchema(name:postimport\_text.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_text.modulename, type:string, comment:null), FieldSchema(name:postimport\_text.filename, type:string, comment:null), FieldSchema(name:postimport\_text.productname, type:string, comment:null), FieldSchema(name:postimport\_text.companyname, type:string, comment:null), FieldSchema(name:postimport\_text.description, type:string, comment:null), FieldSchema(name:postimport\_text.text, type:string, comment:null), FieldSchema(name:postimport\_text.activetime, type:double, comment:null), FieldSchema(name:postimport\_text.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_text.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_text.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411092727\_ccf75124-570a-441d-919e-714c454e4f77); Time taken: 0.102 seconds  INFO : Executing command(queryId=hive\_20170411092727\_ccf75124-570a-441d-919e-714c454e4f77): select \* from postimport\_text  INFO : Query ID = hive\_20170411092727\_ccf75124-570a-441d-919e-714c454e4f77  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:74  INFO : Submitting tokens for job: job\_1490094830773\_0416  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0416/  INFO : Starting Job = job\_1490094830773\_0416, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0416/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0416  INFO : Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  INFO : 2017-04-11 09:27:46,973 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 09:27:59,476 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 76.31 sec  INFO : 2017-04-11 09:28:00,528 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 87.79 sec  INFO : 2017-04-11 09:28:02,593 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 127.85 sec  INFO : 2017-04-11 09:28:04,658 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 139.89 sec  INFO : 2017-04-11 09:28:05,693 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 170.85 sec  INFO : 2017-04-11 09:28:08,796 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 217.15 sec  INFO : 2017-04-11 09:28:11,905 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 260.13 sec  INFO : 2017-04-11 09:28:13,979 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 272.11 sec  INFO : 2017-04-11 09:28:15,017 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 303.14 sec  INFO : 2017-04-11 09:28:18,114 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 346.32 sec  INFO : 2017-04-11 09:28:21,225 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 389.97 sec  INFO : 2017-04-11 09:28:22,261 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 396.04 sec  INFO : 2017-04-11 09:28:24,325 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 432.76 sec  INFO : 2017-04-11 09:28:26,405 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 451.07 sec  INFO : 2017-04-11 09:28:28,464 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 479.44 sec  INFO : 2017-04-11 09:28:30,521 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 515.27 sec  INFO : 2017-04-11 09:28:32,592 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 532.31 sec  INFO : 2017-04-11 09:28:34,639 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 558.5 sec  INFO : 2017-04-11 09:28:39,822 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 588.39 sec  INFO : 2017-04-11 09:28:45,008 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 626.74 sec  INFO : 2017-04-11 09:28:47,075 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 679.08 sec  INFO : 2017-04-11 09:28:49,132 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 701.51 sec  INFO : 2017-04-11 09:28:51,205 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 733.07 sec  INFO : 2017-04-11 09:28:53,253 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 757.07 sec  INFO : 2017-04-11 09:28:55,301 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 785.75 sec  INFO : 2017-04-11 09:28:57,345 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 809.47 sec  INFO : 2017-04-11 09:29:00,429 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 840.95 sec  INFO : 2017-04-11 09:29:02,475 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 858.65 sec  INFO : 2017-04-11 09:29:04,524 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 881.87 sec  INFO : 2017-04-11 09:29:06,569 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 905.38 sec  INFO : 2017-04-11 09:29:09,647 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 935.3 sec  INFO : 2017-04-11 09:29:11,706 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 952.39 sec  INFO : 2017-04-11 09:29:13,760 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 981.08 sec  INFO : 2017-04-11 09:29:16,824 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 1009.4 sec  INFO : 2017-04-11 09:29:19,898 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 1035.42 sec  INFO : 2017-04-11 09:29:22,958 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 1053.45 sec  INFO : 2017-04-11 09:29:26,036 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 1082.14 sec  INFO : 2017-04-11 09:29:29,112 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 1104.44 sec  INFO : 2017-04-11 09:29:33,231 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 1131.7 sec  INFO : 2017-04-11 09:29:36,312 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 1169.0 sec  INFO : 2017-04-11 09:29:39,393 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 1194.48 sec  INFO : 2017-04-11 09:29:42,525 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 1230.8 sec  INFO : 2017-04-11 09:29:45,656 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 1269.72 sec  INFO : 2017-04-11 09:29:47,742 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 1286.25 sec  INFO : 2017-04-11 09:29:49,817 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 1314.27 sec  INFO : 2017-04-11 09:29:51,875 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 1348.54 sec  INFO : 2017-04-11 09:29:53,936 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 1370.74 sec  INFO : 2017-04-11 09:29:56,004 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 1410.83 sec  INFO : 2017-04-11 09:29:57,039 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 1422.21 sec  INFO : 2017-04-11 09:29:59,103 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 1450.98 sec  INFO : 2017-04-11 09:30:01,171 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 1481.32 sec  INFO : 2017-04-11 09:30:03,241 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 1500.74 sec  INFO : 2017-04-11 09:30:04,288 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 1525.21 sec  INFO : 2017-04-11 09:30:07,389 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 1567.41 sec  INFO : 2017-04-11 09:30:09,458 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 1582.14 sec  INFO : 2017-04-11 09:30:12,566 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 1613.4 sec  INFO : 2017-04-11 09:30:13,611 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 1636.0 sec  INFO : 2017-04-11 09:30:16,760 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 1681.51 sec  INFO : 2017-04-11 09:30:17,798 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 1695.04 sec  INFO : 2017-04-11 09:30:19,860 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 1721.82 sec  INFO : 2017-04-11 09:30:22,960 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 1755.22 sec  INFO : 2017-04-11 09:30:26,065 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 1788.09 sec  INFO : 2017-04-11 09:30:28,127 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 1808.35 sec  INFO : 2017-04-11 09:30:31,237 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 1839.87 sec  INFO : 2017-04-11 09:30:33,302 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 1870.49 sec  INFO : 2017-04-11 09:30:35,369 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 1895.83 sec  INFO : 2017-04-11 09:30:37,431 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 1934.55 sec  INFO : 2017-04-11 09:30:39,488 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 1966.0 sec  INFO : 2017-04-11 09:30:40,515 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 1978.15 sec  INFO : 2017-04-11 09:30:42,583 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 2007.68 sec  INFO : 2017-04-11 09:30:44,646 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 2033.17 sec  INFO : 2017-04-11 09:30:45,693 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 2049.18 sec  INFO : 2017-04-11 09:30:48,782 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 2079.38 sec  INFO : 2017-04-11 09:30:50,846 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 2102.36 sec  INFO : 2017-04-11 09:30:52,898 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 2124.73 sec  INFO : 2017-04-11 09:30:55,986 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 2174.5 sec  INFO : 2017-04-11 09:30:57,010 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 2182.93 sec  INFO : 2017-04-11 09:30:59,082 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 2222.56 sec  INFO : 2017-04-11 09:31:01,137 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 2243.04 sec  INFO : 2017-04-11 09:31:02,163 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 2262.6 sec  INFO : 2017-04-11 09:31:05,239 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 2297.62 sec  INFO : 2017-04-11 09:31:07,344 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 2307.98 sec  INFO : 2017-04-11 09:31:10,428 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 2332.76 sec  INFO : 2017-04-11 09:31:13,509 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 2363.48 sec  INFO : 2017-04-11 09:31:16,594 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 2398.51 sec  INFO : 2017-04-11 09:31:17,635 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 2422.01 sec  INFO : 2017-04-11 09:31:20,715 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 2459.68 sec  INFO : 2017-04-11 09:31:21,741 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 2474.86 sec  INFO : 2017-04-11 09:31:23,803 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 2506.11 sec  INFO : 2017-04-11 09:31:24,825 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 2517.89 sec  INFO : 2017-04-11 09:31:25,848 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 2533.49 sec  INFO : 2017-04-11 09:31:27,894 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 2555.34 sec  INFO : 2017-04-11 09:31:29,939 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 2569.46 sec  INFO : 2017-04-11 09:31:33,010 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 2595.97 sec  INFO : 2017-04-11 09:31:36,077 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 2620.12 sec  INFO : 2017-04-11 09:31:38,123 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 2636.59 sec  INFO : 2017-04-11 09:31:41,187 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 2658.33 sec  INFO : 2017-04-11 09:31:48,363 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2684.94 sec  INFO : MapReduce Total cumulative CPU time: 44 minutes 44 seconds 940 msec  INFO : Ended Job = job\_1490094830773\_0416  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 74 Cumulative CPU: 2684.94 sec HDFS Read: 1733864425 HDFS Write: 3634005039 SUCCESS  INFO : Total MapReduce CPU Time Spent: 44 minutes 44 seconds 940 msec  INFO : Completed executing command(queryId=hive\_20170411092727\_ccf75124-570a-441d-919e-714c454e4f77); Time taken: 248.012 seconds  INFO : OK |

## Select All Columns with filter Condition

|  |
| --- |
| select \* from postimport\_text where username = '670063006100750074006F007500730065007200';  INFO : Compiling command(queryId=hive\_20170411093939\_e3af89e0-134e-4f3a-a159-ee33072c1228): select \* from postimport\_text where username = '670063006100750074006F007500730065007200'  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_text.username, type:string, comment:null), FieldSchema(name:postimport\_text.machinename, type:string, comment:null), FieldSchema(name:postimport\_text.machineid, type:string, comment:null), FieldSchema(name:postimport\_text.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_text.domainname, type:string, comment:null), FieldSchema(name:postimport\_text.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_text.modulename, type:string, comment:null), FieldSchema(name:postimport\_text.filename, type:string, comment:null), FieldSchema(name:postimport\_text.productname, type:string, comment:null), FieldSchema(name:postimport\_text.companyname, type:string, comment:null), FieldSchema(name:postimport\_text.description, type:string, comment:null), FieldSchema(name:postimport\_text.text, type:string, comment:null), FieldSchema(name:postimport\_text.activetime, type:double, comment:null), FieldSchema(name:postimport\_text.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_text.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_text.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411093939\_e3af89e0-134e-4f3a-a159-ee33072c1228); Time taken: 0.094 seconds  INFO : Executing command(queryId=hive\_20170411093939\_e3af89e0-134e-4f3a-a159-ee33072c1228): select \* from postimport\_text where username = '670063006100750074006F007500730065007200'  INFO : Query ID = hive\_20170411093939\_e3af89e0-134e-4f3a-a159-ee33072c1228  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:74  INFO : Submitting tokens for job: job\_1490094830773\_0417  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0417/  INFO : Starting Job = job\_1490094830773\_0417, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0417/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0417  INFO : Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  INFO : 2017-04-11 09:39:20,816 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 09:39:32,265 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 18.69 sec  INFO : 2017-04-11 09:39:33,317 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 41.37 sec  INFO : 2017-04-11 09:39:34,345 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 64.28 sec  INFO : 2017-04-11 09:39:35,377 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 94.88 sec  INFO : 2017-04-11 09:39:36,413 Stage-1 map = 6%, reduce = 0%, Cumulative CPU 109.33 sec  INFO : 2017-04-11 09:39:38,479 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 138.36 sec  INFO : 2017-04-11 09:39:39,517 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 161.68 sec  INFO : 2017-04-11 09:39:41,584 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 182.3 sec  INFO : 2017-04-11 09:39:42,617 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 205.65 sec  INFO : 2017-04-11 09:39:44,691 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 225.45 sec  INFO : 2017-04-11 09:39:45,726 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 248.73 sec  INFO : 2017-04-11 09:39:47,800 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 266.7 sec  INFO : 2017-04-11 09:39:48,836 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 291.49 sec  INFO : 2017-04-11 09:39:49,864 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 299.12 sec  INFO : 2017-04-11 09:39:51,948 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 323.75 sec  INFO : 2017-04-11 09:39:52,986 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 339.13 sec  INFO : 2017-04-11 09:39:56,104 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 354.17 sec  INFO : 2017-04-11 09:39:59,221 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 374.79 sec  INFO : 2017-04-11 09:40:01,287 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 397.24 sec  INFO : 2017-04-11 09:40:02,322 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 414.57 sec  INFO : 2017-04-11 09:40:04,379 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 439.91 sec  INFO : 2017-04-11 09:40:05,417 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 462.17 sec  INFO : 2017-04-11 09:40:07,493 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 493.97 sec  INFO : 2017-04-11 09:40:08,530 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 511.73 sec  INFO : 2017-04-11 09:40:10,603 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 539.44 sec  INFO : 2017-04-11 09:40:11,655 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 553.86 sec  INFO : 2017-04-11 09:40:12,694 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 568.51 sec  INFO : 2017-04-11 09:40:13,730 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 580.89 sec  INFO : 2017-04-11 09:40:14,761 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 595.91 sec  INFO : 2017-04-11 09:40:15,787 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 607.66 sec  INFO : 2017-04-11 09:40:17,855 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 630.97 sec  INFO : 2017-04-11 09:40:18,889 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 648.6 sec  INFO : 2017-04-11 09:40:19,931 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 660.52 sec  INFO : 2017-04-11 09:40:20,971 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 673.32 sec  INFO : 2017-04-11 09:40:22,003 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 688.2 sec  INFO : 2017-04-11 09:40:25,134 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 714.25 sec  INFO : 2017-04-11 09:40:27,216 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 725.35 sec  INFO : 2017-04-11 09:40:30,297 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 758.45 sec  INFO : 2017-04-11 09:40:31,323 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 775.7 sec  INFO : 2017-04-11 09:40:32,359 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 775.7 sec  INFO : 2017-04-11 09:40:33,389 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 804.35 sec  INFO : 2017-04-11 09:40:34,420 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 821.07 sec  INFO : 2017-04-11 09:40:35,452 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 828.72 sec  INFO : 2017-04-11 09:40:36,489 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 855.02 sec  INFO : 2017-04-11 09:40:37,520 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 869.38 sec  INFO : 2017-04-11 09:40:39,585 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 898.53 sec  INFO : 2017-04-11 09:40:40,621 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 907.31 sec  INFO : 2017-04-11 09:40:42,687 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 933.64 sec  INFO : 2017-04-11 09:40:43,723 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 942.16 sec  INFO : 2017-04-11 09:40:45,789 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 973.81 sec  INFO : 2017-04-11 09:40:47,847 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 1002.91 sec  INFO : 2017-04-11 09:40:48,895 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 1022.5 sec  INFO : 2017-04-11 09:40:50,969 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 1036.0 sec  INFO : 2017-04-11 09:40:54,060 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 1067.85 sec  INFO : 2017-04-11 09:40:57,169 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 1095.88 sec  INFO : 2017-04-11 09:40:59,222 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 1114.9 sec  INFO : 2017-04-11 09:41:00,249 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 1132.72 sec  INFO : 2017-04-11 09:41:02,312 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 1155.5 sec  INFO : 2017-04-11 09:41:03,343 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 1172.1 sec  INFO : 2017-04-11 09:41:04,373 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 1180.87 sec  INFO : 2017-04-11 09:41:05,408 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 1198.38 sec  INFO : 2017-04-11 09:41:06,438 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 1213.05 sec  INFO : 2017-04-11 09:41:08,506 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 1239.29 sec  INFO : 2017-04-11 09:41:09,538 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 1250.08 sec  INFO : 2017-04-11 09:41:10,570 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 1265.07 sec  INFO : 2017-04-11 09:41:11,602 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 1285.7 sec  INFO : 2017-04-11 09:41:12,635 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 1297.15 sec  INFO : 2017-04-11 09:41:13,668 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 1308.86 sec  INFO : 2017-04-11 09:41:14,689 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 1332.61 sec  INFO : 2017-04-11 09:41:15,712 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 1345.97 sec  INFO : 2017-04-11 09:41:16,749 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 1360.95 sec  INFO : 2017-04-11 09:41:17,791 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 1369.86 sec  INFO : 2017-04-11 09:41:19,879 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 1383.01 sec  INFO : 2017-04-11 09:41:21,952 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 1398.14 sec  INFO : 2017-04-11 09:41:24,019 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 1419.47 sec  INFO : 2017-04-11 09:41:25,050 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 1432.76 sec  INFO : 2017-04-11 09:41:27,116 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 1454.83 sec  INFO : 2017-04-11 09:41:28,144 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 1475.7 sec  INFO : 2017-04-11 09:41:29,180 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 1485.6 sec  INFO : 2017-04-11 09:41:31,260 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 1512.64 sec  INFO : 2017-04-11 09:41:33,325 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 1525.81 sec  INFO : 2017-04-11 09:41:34,358 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 1540.73 sec  INFO : 2017-04-11 09:41:35,394 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 1549.91 sec  INFO : 2017-04-11 09:41:37,456 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 1567.5 sec  INFO : 2017-04-11 09:41:38,494 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 1579.05 sec  INFO : 2017-04-11 09:41:39,527 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 1586.26 sec  INFO : 2017-04-11 09:41:42,613 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1603.79 sec  INFO : MapReduce Total cumulative CPU time: 26 minutes 43 seconds 790 msec  INFO : Ended Job = job\_1490094830773\_0417  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 74 Cumulative CPU: 1603.79 sec HDFS Read: 1733899353 HDFS Write: 2118870 SUCCESS  INFO : Total MapReduce CPU Time Spent: 26 minutes 43 seconds 790 msec  INFO : Completed executing command(queryId=hive\_20170411093939\_e3af89e0-134e-4f3a-a159-ee33072c1228); Time taken: 146.791 seconds  INFO : OK |

# RC File format Snappy compression

The RCFile (Record Columnar File) is a data storage structure that determines how to minimize the space required for relational data in HDFS (Hadoop Distributed File System). It does this by changing the format of the data using the MapReduce framework. The RCFile combines multiple functions such as data storage formatting, data compression, and data access optimization.

It is able to meet all the four requirements of data storage:

(1) Fast data storing,

(2) Improved query processing,

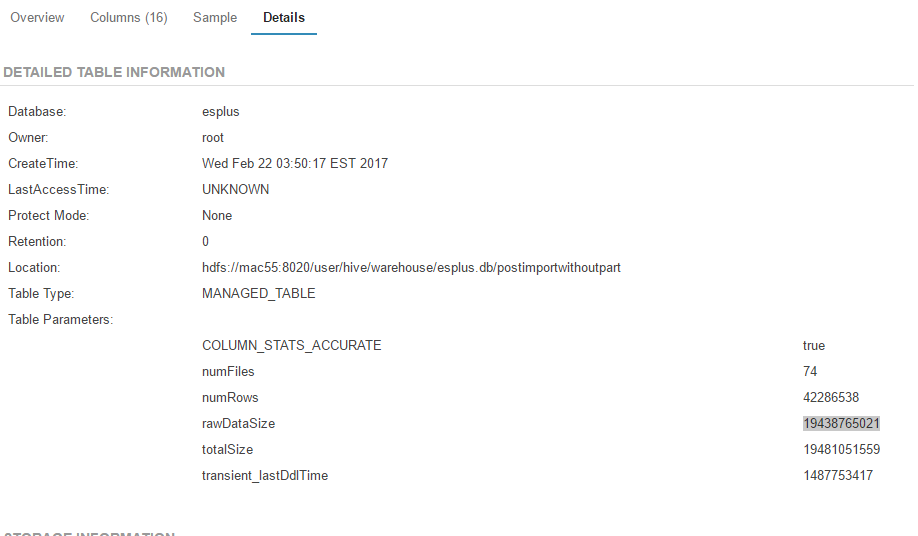
(3) Optimized storage space utilization

(4) Dynamic data access patterns.

The RCFile format can partition the data both horizontally and vertically. This allows it to fetch only the specific fields that are required for analysis, thereby eliminating the standard time needed to analyse the whole table in a database. The overall data size reduction can be as large as 14% of the original data format.

We have a simple text table namely **postimport\_text** whose size is around 18 GB

|  |
| --- |
| hive> create table postimport\_RC like postimport\_text STORED AS RCFILE;  OK  Time taken: 0.148 seconds |
| hive> show create table postimport\_RC;  OK  CREATE TABLE `postimport\_RC`(  `username` string,  `machinename` string,  `machineid` string,  `ipaddress` string,  `domainname` string,  `pcmlogeventid` string,  `modulename` string,  `filename` string,  `productname` string,  `companyname` string,  `description` string,  `text` string,  `activetime` double,  `datetime` timestamp,  `timeslot` timestamp,  `date` date)  ROW FORMAT SERDE  'org.apache.hadoop.hive.serde2.columnar.ColumnarSerDe'  STORED AS INPUTFORMAT  'org.apache.hadoop.hive.ql.io.RCFileInputFormat'  OUTPUTFORMAT  'org.apache.hadoop.hive.ql.io.RCFileOutputFormat'  LOCATION  'hdfs://mac55:8020/user/hive/warehouse/esplus.db/postimport\_rc'  TBLPROPERTIES (  'transient\_lastDdlTime'='1491308502')  Time taken: 0.386 seconds, Fetched: 27 row(s) |

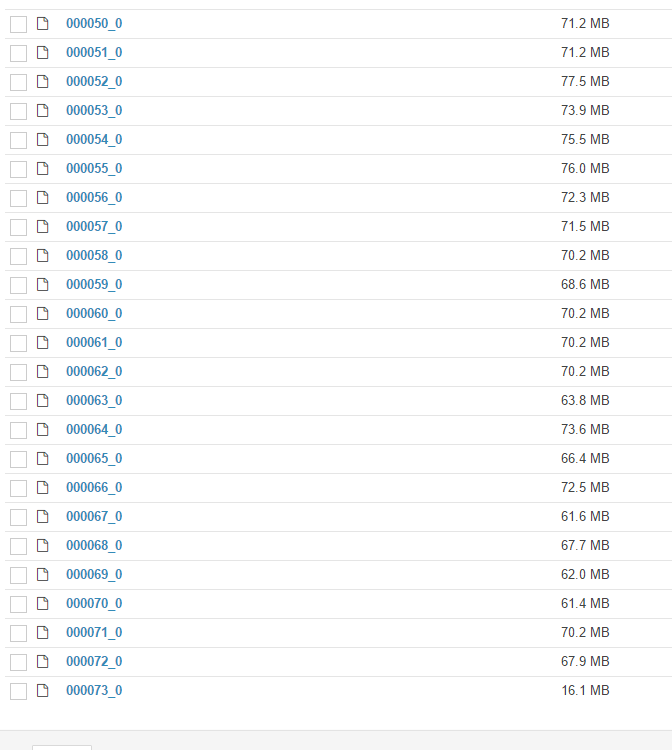


## Load data to RC Table

|  |
| --- |
| SET hive.exec.compress.output=true;  SET mapred.output.compression.type=BLOCK;  SET mapred.output.compression.codec=org.apache.hadoop.io.compress.SnappyCodec; |
| hive> INSERT OVERWRITE TABLE postimport\_RC select \* from postimport\_text;  Query ID = shalaj\_20170411013030\_e7ed9529-5559-489d-b9aa-ddadd6b3b5f3  Total jobs = 1  Launching Job 1 out of 1  Number of reduce tasks is set to 0 since there's no reduce operator  Starting Job = job\_1490094830773\_0307, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0307/  Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0307  Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  2017-04-11 01:30:35,209 Stage-1 map = 0%, reduce = 0%  2017-04-11 01:30:50,343 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 66.91 sec  2017-04-11 01:30:51,587 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 75.8 sec  2017-04-11 01:30:59,249 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 89.04 sec  2017-04-11 01:31:03,974 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 97.35 sec  2017-04-11 01:31:05,175 Stage-1 map = 6%, reduce = 0%, Cumulative CPU 98.13 sec  2017-04-11 01:31:06,357 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 101.08 sec  2017-04-11 01:31:09,977 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 104.15 sec  2017-04-11 01:31:15,093 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 115.39 sec  2017-04-11 01:31:16,547 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 123.5 sec  2017-04-11 01:31:17,741 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 127.65 sec  2017-04-11 01:31:18,942 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 131.05 sec  2017-04-11 01:31:20,369 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 137.45 sec  2017-04-11 01:31:32,215 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 155.63 sec  2017-04-11 01:31:43,035 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 170.27 sec  2017-04-11 01:31:51,778 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 185.16 sec  2017-04-11 01:31:53,012 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 187.86 sec  2017-04-11 01:31:56,874 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 191.14 sec  2017-04-11 01:31:58,041 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 192.41 sec  2017-04-11 01:32:00,418 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 197.73 sec  2017-04-11 01:32:05,762 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 208.85 sec  2017-04-11 01:32:06,954 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 214.36 sec  2017-04-11 01:32:11,972 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 222.37 sec  2017-04-11 01:32:22,843 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 248.12 sec  2017-04-11 01:32:24,046 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 249.45 sec  2017-04-11 01:32:28,089 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 254.44 sec  2017-04-11 01:32:29,294 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 255.36 sec  2017-04-11 01:32:36,731 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 267.79 sec  2017-04-11 01:32:43,066 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 279.3 sec  2017-04-11 01:32:46,894 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 290.69 sec  2017-04-11 01:32:48,079 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 291.41 sec  2017-04-11 01:32:49,278 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 294.69 sec  2017-04-11 01:32:50,757 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 296.02 sec  2017-04-11 01:32:51,940 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 300.9 sec  2017-04-11 01:32:54,640 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 303.19 sec  2017-04-11 01:32:55,776 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 307.88 sec  2017-04-11 01:32:56,906 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 312.23 sec  2017-04-11 01:32:59,341 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 314.77 sec  2017-04-11 01:33:10,416 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 348.6 sec  2017-04-11 01:33:11,611 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 350.94 sec  2017-04-11 01:33:17,777 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 361.62 sec  2017-04-11 01:33:18,959 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 367.87 sec  2017-04-11 01:33:23,897 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 382.84 sec  2017-04-11 01:33:26,561 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 388.47 sec  2017-04-11 01:33:27,766 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 390.11 sec  2017-04-11 01:33:34,963 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 404.93 sec  2017-04-11 01:33:38,498 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 412.08 sec  2017-04-11 01:33:42,083 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 419.16 sec  2017-04-11 01:33:43,274 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 420.21 sec  2017-04-11 01:33:48,040 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 436.62 sec  2017-04-11 01:33:50,614 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 447.44 sec  2017-04-11 01:33:52,826 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 450.89 sec  2017-04-11 01:33:55,123 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 458.78 sec  2017-04-11 01:33:56,285 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 460.01 sec  2017-04-11 01:33:57,434 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 460.21 sec  2017-04-11 01:34:01,940 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 465.53 sec  2017-04-11 01:34:07,988 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 498.35 sec  2017-04-11 01:34:12,828 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 516.94 sec  2017-04-11 01:34:17,654 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 528.89 sec  2017-04-11 01:34:21,258 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 533.92 sec  2017-04-11 01:34:29,211 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 553.69 sec  2017-04-11 01:34:38,251 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 567.74 sec  2017-04-11 01:34:39,433 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 571.35 sec  2017-04-11 01:34:42,475 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 574.61 sec  2017-04-11 01:34:44,929 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 580.17 sec  2017-04-11 01:34:46,128 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 584.88 sec  2017-04-11 01:34:48,486 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 587.15 sec  2017-04-11 01:34:51,991 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 590.52 sec  2017-04-11 01:34:55,554 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 602.16 sec  2017-04-11 01:34:57,939 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 605.24 sec  2017-04-11 01:34:59,404 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 609.77 sec  2017-04-11 01:35:01,857 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 611.54 sec  2017-04-11 01:35:13,151 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 635.92 sec  2017-04-11 01:35:19,159 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 641.7 sec  2017-04-11 01:35:20,370 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 643.36 sec  2017-04-11 01:35:23,920 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 646.24 sec  2017-04-11 01:35:25,098 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 646.99 sec  2017-04-11 01:35:29,413 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 651.58 sec  2017-04-11 01:35:30,600 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 652.45 sec  2017-04-11 01:35:35,366 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 656.25 sec  2017-04-11 01:35:39,139 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 658.75 sec  2017-04-11 01:35:46,234 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 663.83 sec  2017-04-11 01:35:47,445 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 665.3 sec  2017-04-11 01:35:59,584 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 668.02 sec  MapReduce Total cumulative CPU time: 11 minutes 8 seconds 20 msec  Ended Job = job\_1490094830773\_0307  Stage-4 is selected by condition resolver.  Stage-3 is filtered out by condition resolver.  Stage-5 is filtered out by condition resolver.  Moving data to: hdfs://mac55:8020/user/hive/warehouse/esplus.db/postimport\_rc/.hive-staging\_hive\_2017-04-11\_01-30-25\_502\_8515484391939762968-1/-ext-10000  Loading data to table esplus.postimport\_rc  Table esplus.postimport\_rc stats: [numFiles=74, numRows=42286538, totalSize=5502430802, rawDataSize=18804466951]  MapReduce Jobs Launched:  Stage-Stage-1: Map: 74 Cumulative CPU: 668.02 sec HDFS Read: 19486334082 HDFS Write: 5502437239 SUCCESS  Total MapReduce CPU Time Spent: 11 minutes 8 seconds 20 msec  OK  Time taken: 338.984 seconds |

## Data Size

|  |
| --- |
| hive> dfs -du -s -h /user/hive/warehouse/esplus.db/postimport\_rc;  5.1 G 15.4 G /user/hive/warehouse/esplus.db/postimport\_rc |



It creates same number of files **(total 74 files)** as text file format though each file size is lesser than the text file format

## Select some columns

SET hive.exec.compress.output=true;

|  |
| --- |
| select productname,username from postimport\_rc;  INFO : Compiling command(queryId=hive\_20170411014040\_8ba188cf-3d87-4b31-8e24-a100c0d27d5f): select productname,username from postimport\_rc  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:productname, type:string, comment:null), FieldSchema(name:username, type:string, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411014040\_8ba188cf-3d87-4b31-8e24-a100c0d27d5f); Time taken: 0.117 seconds  INFO : Executing command(queryId=hive\_20170411014040\_8ba188cf-3d87-4b31-8e24-a100c0d27d5f): select productname,username from postimport\_rc  INFO : Query ID = hive\_20170411014040\_8ba188cf-3d87-4b31-8e24-a100c0d27d5f  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:19  INFO : Submitting tokens for job: job\_1490094830773\_0308  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0308/  INFO : Starting Job = job\_1490094830773\_0308, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0308/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0308  INFO : Hadoop job information for Stage-1: number of mappers: 19; number of reducers: 0  INFO : 2017-04-11 01:40:33,944 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 01:40:45,341 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 10.47 sec  INFO : 2017-04-11 01:40:46,367 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 42.08 sec  INFO : 2017-04-11 01:40:48,416 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 56.74 sec  INFO : 2017-04-11 01:40:49,443 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 68.97 sec  INFO : 2017-04-11 01:40:50,471 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 72.06 sec  INFO : 2017-04-11 01:40:51,501 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 75.99 sec  INFO : 2017-04-11 01:40:52,526 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 83.09 sec  INFO : 2017-04-11 01:40:53,549 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 84.75 sec  INFO : 2017-04-11 01:40:55,607 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 90.64 sec  INFO : 2017-04-11 01:40:56,631 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 93.81 sec  INFO : 2017-04-11 01:40:57,658 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 97.64 sec  INFO : 2017-04-11 01:40:58,684 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 104.9 sec  INFO : 2017-04-11 01:40:59,707 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 109.86 sec  INFO : 2017-04-11 01:41:00,730 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 116.34 sec  INFO : 2017-04-11 01:41:01,750 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 117.83 sec  INFO : 2017-04-11 01:41:02,773 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 124.52 sec  INFO : 2017-04-11 01:41:03,792 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 128.45 sec  INFO : 2017-04-11 01:41:05,834 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 131.96 sec  INFO : 2017-04-11 01:41:06,856 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 132.82 sec  INFO : 2017-04-11 01:41:08,900 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 137.56 sec  INFO : 2017-04-11 01:41:10,942 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 139.42 sec  INFO : 2017-04-11 01:41:11,963 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 142.38 sec  INFO : 2017-04-11 01:41:12,983 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 142.84 sec  INFO : MapReduce Total cumulative CPU time: 2 minutes 22 seconds 840 msec  INFO : Ended Job = job\_1490094830773\_0308  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 19 Cumulative CPU: 142.84 sec HDFS Read: 908462915 HDFS Write: 233696021 SUCCESS  INFO : Total MapReduce CPU Time Spent: 2 minutes 22 seconds 840 msec  INFO : Completed executing command(queryId=hive\_20170411014040\_8ba188cf-3d87-4b31-8e24-a100c0d27d5f); Time taken: 46.076 seconds  INFO : OK |

**System Counter**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Maps Total** | **Reduces Total** | **Total** |
| Hdfs Write Ops | 38 | 0 | 38 |
| Hdfs Read Ops | 224 | 0 | 224 |
| Hdfs Large Read Ops | 0 | 0 | 0 |
| Hdfs Bytes Written | 2439779252 | 0 | 2439779252 |
| Hdfs Bytes Read | 908471850 | 0 | 908471850 |
| File Write Ops | 0 | 0 | 0 |
| File Read Ops | 0 | 0 | 0 |
| File Large Read Ops | 0 | 0 | 0 |
| File Bytes Written | 4569756 | 0 | 4569756 |
| File Bytes Read | 0 | 0 | 0 |

We can clearly see the difference in hdfs byte read

In normal table it reads full table **19486196675(18.1GB)** while in RC table it reads only **908471850 (.8 GB)** though full table size is 5 GB, basically RC file Format read only selected columns

## Check Max Date

|  |
| --- |
| select max(date) from postimport\_rc;  INFO : Compiling command(queryId=hive\_20170411022929\_ea2cc1b2-2802-44e9-ac13-e2c49b9f2fd2): select max(date) from postimport\_rc  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:\_c0, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411022929\_ea2cc1b2-2802-44e9-ac13-e2c49b9f2fd2); Time taken: 0.102 seconds  INFO : Executing command(queryId=hive\_20170411022929\_ea2cc1b2-2802-44e9-ac13-e2c49b9f2fd2): select max(date) from postimport\_rc  INFO : Query ID = hive\_20170411022929\_ea2cc1b2-2802-44e9-ac13-e2c49b9f2fd2  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks determined at compile time: 1  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:19  INFO : Submitting tokens for job: job\_1490094830773\_0313  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0313/  INFO : Starting Job = job\_1490094830773\_0313, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0313/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0313  INFO : Hadoop job information for Stage-1: number of mappers: 19; number of reducers: 1  INFO : 2017-04-11 02:29:49,765 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 02:30:07,787 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 34.46 sec  INFO : 2017-04-11 02:30:08,810 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 37.4 sec  INFO : 2017-04-11 02:30:10,853 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 40.68 sec  INFO : 2017-04-11 02:30:11,873 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 41.8 sec  INFO : 2017-04-11 02:30:13,917 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 44.65 sec  INFO : 2017-04-11 02:30:14,939 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 46.29 sec  INFO : 2017-04-11 02:30:16,979 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 47.8 sec  INFO : 2017-04-11 02:30:17,999 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 48.64 sec  INFO : 2017-04-11 02:30:21,060 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 53.25 sec  INFO : 2017-04-11 02:30:23,100 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 56.19 sec  INFO : 2017-04-11 02:30:24,122 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 56.45 sec  INFO : 2017-04-11 02:30:25,142 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 56.9 sec  INFO : 2017-04-11 02:30:26,168 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 62.78 sec  INFO : 2017-04-11 02:30:27,189 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 67.33 sec  INFO : 2017-04-11 02:30:28,212 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 67.79 sec  INFO : 2017-04-11 02:30:29,234 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 68.76 sec  INFO : 2017-04-11 02:30:30,257 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 69.72 sec  INFO : 2017-04-11 02:30:31,279 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 70.8 sec  INFO : 2017-04-11 02:30:32,300 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 71.67 sec  INFO : 2017-04-11 02:30:33,321 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 73.26 sec  INFO : 2017-04-11 02:30:37,408 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 74.21 sec  INFO : MapReduce Total cumulative CPU time: 1 minutes 14 seconds 210 msec  INFO : Ended Job = job\_1490094830773\_0313  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 19 Reduce: 1 Cumulative CPU: 74.21 sec HDFS Read: 329900402 HDFS Write: 19 SUCCESS  INFO : Total MapReduce CPU Time Spent: 1 minutes 14 seconds 210 msec  INFO : Completed executing command(queryId=hive\_20170411022929\_ea2cc1b2-2802-44e9-ac13-e2c49b9f2fd2); Time taken: 57.372 seconds  INFO : OK |

## User wise Max Date

|  |
| --- |
| select username,max(date) from postimport\_rc group by username;  INFO : Compiling command(queryId=hive\_20170411023131\_c73bc270-cccb-450a-8967-ef30a7c31474): select username,max(date) from postimport\_rc group by username  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:username, type:string, comment:null), FieldSchema(name:\_c1, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411023131\_c73bc270-cccb-450a-8967-ef30a7c31474); Time taken: 0.099 seconds  INFO : Executing command(queryId=hive\_20170411023131\_c73bc270-cccb-450a-8967-ef30a7c31474): select username,max(date) from postimport\_rc group by username  INFO : Query ID = hive\_20170411023131\_c73bc270-cccb-450a-8967-ef30a7c31474  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks not specified. Estimated from input data size: 82  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:19  INFO : Submitting tokens for job: job\_1490094830773\_0314  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0314/  INFO : Starting Job = job\_1490094830773\_0314, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0314/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0314  INFO : Hadoop job information for Stage-1: number of mappers: 19; number of reducers: 82  INFO : 2017-04-11 02:31:25,249 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 02:31:37,699 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 25.77 sec  INFO : 2017-04-11 02:31:38,723 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 55.08 sec  INFO : 2017-04-11 02:31:40,770 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 56.03 sec  INFO : 2017-04-11 02:31:42,829 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 60.63 sec  INFO : 2017-04-11 02:31:43,851 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 61.61 sec  INFO : 2017-04-11 02:31:46,923 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 68.0 sec  INFO : 2017-04-11 02:31:49,993 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 74.05 sec  INFO : 2017-04-11 02:31:51,018 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 74.93 sec  INFO : 2017-04-11 02:31:52,039 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 75.81 sec  INFO : 2017-04-11 02:31:53,060 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 76.79 sec  INFO : 2017-04-11 02:31:55,110 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 80.62 sec  INFO : 2017-04-11 02:31:56,132 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 81.56 sec  INFO : 2017-04-11 02:31:57,179 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 83.82 sec  INFO : 2017-04-11 02:31:58,211 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 88.2 sec  INFO : 2017-04-11 02:32:05,564 Stage-1 map = 100%, reduce = 6%, Cumulative CPU 94.01 sec  INFO : 2017-04-11 02:32:06,591 Stage-1 map = 100%, reduce = 15%, Cumulative CPU 101.63 sec  INFO : 2017-04-11 02:32:08,701 Stage-1 map = 100%, reduce = 18%, Cumulative CPU 107.07 sec  INFO : 2017-04-11 02:32:10,768 Stage-1 map = 100%, reduce = 20%, Cumulative CPU 108.32 sec  INFO : 2017-04-11 02:32:11,816 Stage-1 map = 100%, reduce = 28%, Cumulative CPU 116.25 sec  INFO : 2017-04-11 02:32:12,842 Stage-1 map = 100%, reduce = 33%, Cumulative CPU 121.2 sec  INFO : 2017-04-11 02:32:14,923 Stage-1 map = 100%, reduce = 37%, Cumulative CPU 125.32 sec  INFO : 2017-04-11 02:32:15,970 Stage-1 map = 100%, reduce = 38%, Cumulative CPU 126.56 sec  INFO : 2017-04-11 02:32:16,997 Stage-1 map = 100%, reduce = 43%, Cumulative CPU 129.95 sec  INFO : 2017-04-11 02:32:19,066 Stage-1 map = 100%, reduce = 51%, Cumulative CPU 140.99 sec  INFO : 2017-04-11 02:32:20,094 Stage-1 map = 100%, reduce = 55%, Cumulative CPU 145.13 sec  INFO : 2017-04-11 02:32:21,136 Stage-1 map = 100%, reduce = 56%, Cumulative CPU 146.46 sec  INFO : 2017-04-11 02:32:24,284 Stage-1 map = 100%, reduce = 65%, Cumulative CPU 154.79 sec  INFO : 2017-04-11 02:32:25,348 Stage-1 map = 100%, reduce = 71%, Cumulative CPU 161.38 sec  INFO : 2017-04-11 02:32:26,387 Stage-1 map = 100%, reduce = 73%, Cumulative CPU 163.96 sec  INFO : 2017-04-11 02:32:27,440 Stage-1 map = 100%, reduce = 74%, Cumulative CPU 165.18 sec  INFO : 2017-04-11 02:32:29,528 Stage-1 map = 100%, reduce = 76%, Cumulative CPU 166.43 sec  INFO : 2017-04-11 02:32:30,563 Stage-1 map = 100%, reduce = 85%, Cumulative CPU 176.41 sec  INFO : 2017-04-11 02:32:31,586 Stage-1 map = 100%, reduce = 90%, Cumulative CPU 181.78 sec  INFO : 2017-04-11 02:32:32,611 Stage-1 map = 100%, reduce = 93%, Cumulative CPU 184.46 sec  INFO : 2017-04-11 02:32:33,635 Stage-1 map = 100%, reduce = 94%, Cumulative CPU 185.71 sec  INFO : 2017-04-11 02:32:34,660 Stage-1 map = 100%, reduce = 96%, Cumulative CPU 187.91 sec  INFO : 2017-04-11 02:32:35,679 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 191.32 sec  INFO : MapReduce Total cumulative CPU time: 3 minutes 11 seconds 320 msec  INFO : Ended Job = job\_1490094830773\_0314  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 19 Reduce: 82 Cumulative CPU: 191.32 sec HDFS Read: 794544020 HDFS Write: 38508 SUCCESS  INFO : Total MapReduce CPU Time Spent: 3 minutes 11 seconds 320 msec  INFO : Completed executing command(queryId=hive\_20170411023131\_c73bc270-cccb-450a-8967-ef30a7c31474); Time taken: 79.277 seconds  INFO : OK |

## Select All Columns

|  |
| --- |
| select \* from postimport\_rc  INFO : Compiling command(queryId=hive\_20170411023333\_7fd69c68-1bf9-4c59-baf3-e11d0b441a93): select \* from postimport\_rc  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_rc.username, type:string, comment:null), FieldSchema(name:postimport\_rc.machinename, type:string, comment:null), FieldSchema(name:postimport\_rc.machineid, type:string, comment:null), FieldSchema(name:postimport\_rc.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_rc.domainname, type:string, comment:null), FieldSchema(name:postimport\_rc.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_rc.modulename, type:string, comment:null), FieldSchema(name:postimport\_rc.filename, type:string, comment:null), FieldSchema(name:postimport\_rc.productname, type:string, comment:null), FieldSchema(name:postimport\_rc.companyname, type:string, comment:null), FieldSchema(name:postimport\_rc.description, type:string, comment:null), FieldSchema(name:postimport\_rc.text, type:string, comment:null), FieldSchema(name:postimport\_rc.activetime, type:double, comment:null), FieldSchema(name:postimport\_rc.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_rc.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_rc.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411023333\_7fd69c68-1bf9-4c59-baf3-e11d0b441a93); Time taken: 0.107 seconds  INFO : Executing command(queryId=hive\_20170411023333\_7fd69c68-1bf9-4c59-baf3-e11d0b441a93): select \* from postimport\_rc  INFO : Query ID = hive\_20170411023333\_7fd69c68-1bf9-4c59-baf3-e11d0b441a93  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:19  INFO : Submitting tokens for job: job\_1490094830773\_0315  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0315/  INFO : Starting Job = job\_1490094830773\_0315, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0315/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0315  INFO : Hadoop job information for Stage-1: number of mappers: 19; number of reducers: 0  INFO : 2017-04-11 02:33:32,619 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 02:33:53,344 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 132.28 sec  INFO : 2017-04-11 02:33:57,450 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 164.28 sec  INFO : 2017-04-11 02:34:08,723 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 238.46 sec  INFO : 2017-04-11 02:34:09,747 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 248.15 sec  INFO : 2017-04-11 02:34:13,844 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 275.0 sec  INFO : 2017-04-11 02:34:14,873 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 283.23 sec  INFO : 2017-04-11 02:34:16,921 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 296.55 sec  INFO : 2017-04-11 02:34:17,945 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 303.05 sec  INFO : 2017-04-11 02:34:19,999 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 314.85 sec  INFO : 2017-04-11 02:34:21,024 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 319.81 sec  INFO : 2017-04-11 02:34:22,051 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 325.47 sec  INFO : 2017-04-11 02:34:23,076 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 334.13 sec  INFO : 2017-04-11 02:34:27,169 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 357.52 sec  INFO : 2017-04-11 02:34:28,196 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 362.42 sec  INFO : 2017-04-11 02:34:31,269 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 382.26 sec  INFO : 2017-04-11 02:34:35,359 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 401.28 sec  INFO : 2017-04-11 02:34:36,381 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 405.26 sec  INFO : 2017-04-11 02:34:38,447 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 413.36 sec  INFO : 2017-04-11 02:34:41,527 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 429.83 sec  INFO : 2017-04-11 02:34:47,651 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 460.03 sec  INFO : 2017-04-11 02:34:50,711 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 472.29 sec  INFO : 2017-04-11 02:34:54,790 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 489.59 sec  INFO : 2017-04-11 02:34:59,892 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 509.19 sec  INFO : 2017-04-11 02:35:10,092 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 550.71 sec  INFO : 2017-04-11 02:35:12,130 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 561.46 sec  INFO : 2017-04-11 02:35:13,151 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 563.8 sec  INFO : 2017-04-11 02:35:19,262 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 587.59 sec  INFO : 2017-04-11 02:35:20,280 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 589.65 sec  INFO : 2017-04-11 02:35:21,299 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 598.2 sec  INFO : 2017-04-11 02:35:24,359 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 610.35 sec  INFO : 2017-04-11 02:35:30,473 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 633.69 sec  INFO : 2017-04-11 02:35:33,529 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 646.63 sec  INFO : 2017-04-11 02:35:38,627 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 662.86 sec  INFO : 2017-04-11 02:35:39,648 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 669.78 sec  INFO : 2017-04-11 02:35:44,747 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 686.75 sec  INFO : 2017-04-11 02:35:50,866 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 712.53 sec  INFO : 2017-04-11 02:35:51,884 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 722.77 sec  INFO : 2017-04-11 02:35:54,945 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 742.87 sec  INFO : 2017-04-11 02:35:56,984 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 750.01 sec  INFO : 2017-04-11 02:35:58,002 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 759.64 sec  INFO : 2017-04-11 02:36:01,069 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 782.6 sec  INFO : 2017-04-11 02:36:04,146 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 803.51 sec  INFO : 2017-04-11 02:36:06,194 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 810.94 sec  INFO : 2017-04-11 02:36:07,214 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 821.58 sec  INFO : 2017-04-11 02:36:10,285 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 835.21 sec  INFO : 2017-04-11 02:36:12,338 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 841.84 sec  INFO : 2017-04-11 02:36:13,356 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 848.27 sec  INFO : 2017-04-11 02:36:18,459 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 870.86 sec  INFO : 2017-04-11 02:36:20,500 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 883.1 sec  INFO : 2017-04-11 02:36:22,540 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 892.56 sec  INFO : 2017-04-11 02:36:23,562 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 900.96 sec  INFO : 2017-04-11 02:36:26,624 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 917.75 sec  INFO : 2017-04-11 02:36:28,665 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 927.65 sec  INFO : 2017-04-11 02:36:31,737 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 940.98 sec  INFO : 2017-04-11 02:36:34,808 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 953.49 sec  INFO : 2017-04-11 02:36:35,825 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 960.32 sec  INFO : 2017-04-11 02:36:36,851 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 961.06 sec  INFO : 2017-04-11 02:36:37,869 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 966.9 sec  INFO : 2017-04-11 02:36:41,944 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 978.58 sec  INFO : 2017-04-11 02:36:46,022 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 989.49 sec  INFO : 2017-04-11 02:36:47,041 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 994.63 sec  INFO : 2017-04-11 02:36:53,208 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1000.26 sec  INFO : MapReduce Total cumulative CPU time: 16 minutes 40 seconds 260 msec  INFO : Ended Job = job\_1490094830773\_0315  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 19 Cumulative CPU: 1000.26 sec HDFS Read: 5502529086 HDFS Write: 3633971126 SUCCESS  INFO : Total MapReduce CPU Time Spent: 16 minutes 40 seconds 260 msec  INFO : Completed executing command(queryId=hive\_20170411023333\_7fd69c68-1bf9-4c59-baf3-e11d0b441a93); Time taken: 205.725 seconds  INFO : OK |

## Select All Columns with filter Condition

|  |
| --- |
| select \* from postimport\_rc where username = '670063006100750074006F007500730065007200'  INFO : Compiling command(queryId=hive\_20170411023737\_ecdceab5-f92c-4f8f-bdf7-b65919d0f800): select \* from postimport\_rc where username = '670063006100750074006F007500730065007200'  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_rc.username, type:string, comment:null), FieldSchema(name:postimport\_rc.machinename, type:string, comment:null), FieldSchema(name:postimport\_rc.machineid, type:string, comment:null), FieldSchema(name:postimport\_rc.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_rc.domainname, type:string, comment:null), FieldSchema(name:postimport\_rc.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_rc.modulename, type:string, comment:null), FieldSchema(name:postimport\_rc.filename, type:string, comment:null), FieldSchema(name:postimport\_rc.productname, type:string, comment:null), FieldSchema(name:postimport\_rc.companyname, type:string, comment:null), FieldSchema(name:postimport\_rc.description, type:string, comment:null), FieldSchema(name:postimport\_rc.text, type:string, comment:null), FieldSchema(name:postimport\_rc.activetime, type:double, comment:null), FieldSchema(name:postimport\_rc.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_rc.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_rc.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411023737\_ecdceab5-f92c-4f8f-bdf7-b65919d0f800); Time taken: 0.142 seconds  INFO : Executing command(queryId=hive\_20170411023737\_ecdceab5-f92c-4f8f-bdf7-b65919d0f800): select \* from postimport\_rc where username = '670063006100750074006F007500730065007200'  INFO : Query ID = hive\_20170411023737\_ecdceab5-f92c-4f8f-bdf7-b65919d0f800  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:19  INFO : Submitting tokens for job: job\_1490094830773\_0316  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0316/  INFO : Starting Job = job\_1490094830773\_0316, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0316/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0316  INFO : Hadoop job information for Stage-1: number of mappers: 19; number of reducers: 0  INFO : 2017-04-11 02:37:39,379 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 02:37:49,966 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 5.14 sec  INFO : 2017-04-11 02:37:50,988 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 17.24 sec  INFO : 2017-04-11 02:37:52,009 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 25.93 sec  INFO : 2017-04-11 02:37:53,033 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 36.6 sec  INFO : 2017-04-11 02:37:54,054 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 40.99 sec  INFO : 2017-04-11 02:37:55,077 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 42.52 sec  INFO : 2017-04-11 02:37:56,099 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 46.61 sec  INFO : 2017-04-11 02:37:57,321 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 47.62 sec  INFO : 2017-04-11 02:37:58,341 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 49.37 sec  INFO : 2017-04-11 02:37:59,363 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 54.89 sec  INFO : 2017-04-11 02:38:01,409 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 56.95 sec  INFO : 2017-04-11 02:38:02,433 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 64.06 sec  INFO : 2017-04-11 02:38:03,455 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 67.56 sec  INFO : 2017-04-11 02:38:04,479 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 68.72 sec  INFO : 2017-04-11 02:38:05,502 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 70.67 sec  INFO : 2017-04-11 02:38:06,523 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 72.29 sec  INFO : 2017-04-11 02:38:08,566 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 78.47 sec  INFO : 2017-04-11 02:38:09,586 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 81.91 sec  INFO : 2017-04-11 02:38:10,610 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 85.44 sec  INFO : 2017-04-11 02:38:11,629 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 86.68 sec  INFO : 2017-04-11 02:38:12,656 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 99.37 sec  INFO : 2017-04-11 02:38:13,706 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 101.45 sec  INFO : 2017-04-11 02:38:14,740 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 102.79 sec  INFO : MapReduce Total cumulative CPU time: 1 minutes 42 seconds 790 msec  INFO : Ended Job = job\_1490094830773\_0316  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 19 Cumulative CPU: 102.79 sec HDFS Read: 5502538054 HDFS Write: 2082971 SUCCESS  INFO : Total MapReduce CPU Time Spent: 1 minutes 42 seconds 790 msec  INFO : Completed executing command(queryId=hive\_20170411023737\_ecdceab5-f92c-4f8f-bdf7-b65919d0f800); Time taken: 43.421 seconds  INFO : OK |

# RC File format Gzip compression

## Load data to RC Table

|  |
| --- |
| SET hive.exec.compress.output=true;  SET mapred.output.compression.type=BLOCK;  SET mapred.output.compression.codec=org.apache.hadoop.io.compress.GzipCodec; |
| hive> INSERT OVERWRITE TABLE postimport\_RC select \* from postimport\_text;  Query ID = shalaj\_20170411030101\_fefc29fb-90c9-4f25-a686-2b92a03d2938  Total jobs = 1  Launching Job 1 out of 1  Number of reduce tasks is set to 0 since there's no reduce operator  Starting Job = job\_1490094830773\_0320, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0320/  Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0320  Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  2017-04-11 03:02:09,106 Stage-1 map = 0%, reduce = 0%  2017-04-11 03:02:26,533 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 102.82 sec  2017-04-11 03:02:31,516 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 139.42 sec  2017-04-11 03:02:34,066 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 157.61 sec  2017-04-11 03:02:35,223 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 161.77 sec  2017-04-11 03:02:37,683 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 178.6 sec  2017-04-11 03:02:41,185 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 190.74 sec  2017-04-11 03:02:43,355 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 204.32 sec  2017-04-11 03:02:46,842 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 216.83 sec  2017-04-11 03:02:49,264 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 230.67 sec  2017-04-11 03:02:50,355 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 235.56 sec  2017-04-11 03:02:53,726 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 246.27 sec  2017-04-11 03:02:54,904 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 251.16 sec  2017-04-11 03:02:59,460 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 286.56 sec  2017-04-11 03:03:00,935 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 302.18 sec  2017-04-11 03:03:04,354 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 345.66 sec  2017-04-11 03:03:05,420 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 345.94 sec  2017-04-11 03:03:06,699 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 367.34 sec  2017-04-11 03:03:08,821 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 380.55 sec  2017-04-11 03:03:10,120 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 397.27 sec  2017-04-11 03:03:12,537 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 411.88 sec  2017-04-11 03:03:13,604 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 423.34 sec  2017-04-11 03:03:14,696 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 425.99 sec  2017-04-11 03:03:15,770 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 437.6 sec  2017-04-11 03:03:19,314 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 457.38 sec  2017-04-11 03:03:21,880 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 464.09 sec  2017-04-11 03:03:25,439 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 507.1 sec  2017-04-11 03:03:26,589 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 525.2 sec  2017-04-11 03:03:27,711 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 530.74 sec  2017-04-11 03:03:28,800 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 545.48 sec  2017-04-11 03:03:31,203 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 558.41 sec  2017-04-11 03:03:33,553 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 570.71 sec  2017-04-11 03:03:34,677 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 575.72 sec  2017-04-11 03:03:35,732 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 581.31 sec  2017-04-11 03:03:37,927 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 593.43 sec  2017-04-11 03:03:44,398 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 620.66 sec  2017-04-11 03:03:46,809 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 665.36 sec  2017-04-11 03:03:47,887 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 673.72 sec  2017-04-11 03:03:51,141 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 690.19 sec  2017-04-11 03:03:52,230 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 700.19 sec  2017-04-11 03:03:53,309 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 705.21 sec  2017-04-11 03:03:55,542 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 716.83 sec  2017-04-11 03:03:58,305 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 729.78 sec  2017-04-11 03:03:59,636 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 757.35 sec  2017-04-11 03:04:02,075 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 773.25 sec  2017-04-11 03:04:04,242 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 778.83 sec  2017-04-11 03:04:07,526 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 798.65 sec  2017-04-11 03:04:09,660 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 807.25 sec  2017-04-11 03:04:10,696 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 813.42 sec  2017-04-11 03:04:15,857 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 856.45 sec  2017-04-11 03:04:16,882 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 859.71 sec  2017-04-11 03:04:19,351 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 878.48 sec  2017-04-11 03:04:20,402 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 881.63 sec  2017-04-11 03:04:23,675 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 912.74 sec  2017-04-11 03:04:24,734 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 923.84 sec  2017-04-11 03:04:26,017 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 932.72 sec  2017-04-11 03:04:28,376 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 960.73 sec  2017-04-11 03:04:29,446 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 963.74 sec  2017-04-11 03:04:30,742 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 972.11 sec  2017-04-11 03:04:32,029 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 985.3 sec  2017-04-11 03:04:33,331 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 991.0 sec  2017-04-11 03:04:34,423 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 996.67 sec  2017-04-11 03:04:36,795 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 1012.26 sec  2017-04-11 03:04:40,039 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 1020.93 sec  2017-04-11 03:04:44,304 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 1053.58 sec  2017-04-11 03:04:45,348 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 1055.99 sec  2017-04-11 03:04:47,446 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 1065.05 sec  2017-04-11 03:04:49,494 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1068.0 sec  MapReduce Total cumulative CPU time: 17 minutes 48 seconds 0 msec  Ended Job = job\_1490094830773\_0320  Stage-4 is selected by condition resolver.  Stage-3 is filtered out by condition resolver.  Stage-5 is filtered out by condition resolver.  Moving data to: hdfs://mac55:8020/user/hive/warehouse/esplus.db/postimport\_rc/.hive-staging\_hive\_2017-04-11\_03-01-58\_858\_8974695620913795424-1/-ext-10000  Loading data to table esplus.postimport\_rc  Table esplus.postimport\_rc stats: [numFiles=74, numRows=42286538, totalSize=2257185069, rawDataSize=18804466951]  MapReduce Jobs Launched:  Stage-Stage-1: Map: 74 Cumulative CPU: 1068.0 sec HDFS Read: 19486334674 HDFS Write: 2257191506 SUCCESS  Total MapReduce CPU Time Spent: 17 minutes 48 seconds 0 msec  OK  Time taken: 175.146 seconds |

## Data Size

|  |
| --- |
| hive> dfs -du -s -h /user/hive/warehouse/esplus.db/postimport\_rc;  2.1 G 6.3 G /user/hive/warehouse/esplus.db/postimport\_rc |

## Select some columns

SET hive.exec.compress.output=true;

|  |
| --- |
| select productname,username from postimport\_rc;  INFO : Compiling command(queryId=hive\_20170411030707\_e7704882-da73-456f-bfc6-2620464ebbc3): select productname,username from postimport\_rc  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:productname, type:string, comment:null), FieldSchema(name:username, type:string, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411030707\_e7704882-da73-456f-bfc6-2620464ebbc3); Time taken: 0.114 seconds  INFO : Executing command(queryId=hive\_20170411030707\_e7704882-da73-456f-bfc6-2620464ebbc3): select productname,username from postimport\_rc  INFO : Query ID = hive\_20170411030707\_e7704882-da73-456f-bfc6-2620464ebbc3  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:9  INFO : Submitting tokens for job: job\_1490094830773\_0321  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0321/  INFO : Starting Job = job\_1490094830773\_0321, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0321/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0321  INFO : Hadoop job information for Stage-1: number of mappers: 9; number of reducers: 0  INFO : 2017-04-11 03:07:36,444 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 03:07:45,696 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 4.96 sec  INFO : 2017-04-11 03:07:46,725 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 37.46 sec  INFO : 2017-04-11 03:07:48,774 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 49.4 sec  INFO : 2017-04-11 03:07:49,801 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 70.49 sec  INFO : 2017-04-11 03:07:51,852 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 76.0 sec  INFO : 2017-04-11 03:07:52,878 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 93.16 sec  INFO : 2017-04-11 03:07:54,926 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 101.12 sec  INFO : 2017-04-11 03:07:55,949 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 115.13 sec  INFO : 2017-04-11 03:07:56,969 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 116.15 sec  INFO : 2017-04-11 03:07:57,991 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 121.51 sec  INFO : 2017-04-11 03:07:59,012 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 125.92 sec  INFO : 2017-04-11 03:08:00,034 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 129.08 sec  INFO : MapReduce Total cumulative CPU time: 2 minutes 9 seconds 80 msec  INFO : Ended Job = job\_1490094830773\_0321  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 9 Cumulative CPU: 129.08 sec HDFS Read: 385825907 HDFS Write: 233679118 SUCCESS  INFO : Total MapReduce CPU Time Spent: 2 minutes 9 seconds 80 msec  INFO : Completed executing command(queryId=hive\_20170411030707\_e7704882-da73-456f-bfc6-2620464ebbc3); Time taken: 31.122 seconds  INFO : OK |

## Check Max Date

|  |
| --- |
| select max(date) from postimport\_rc;  INFO : Compiling command(queryId=hive\_20170411030909\_94a6e1dd-39ef-48ce-a24d-a5c95db2df8a): select max(date) from postimport\_rc  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:\_c0, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411030909\_94a6e1dd-39ef-48ce-a24d-a5c95db2df8a); Time taken: 0.127 seconds  INFO : Executing command(queryId=hive\_20170411030909\_94a6e1dd-39ef-48ce-a24d-a5c95db2df8a): select max(date) from postimport\_rc  INFO : Query ID = hive\_20170411030909\_94a6e1dd-39ef-48ce-a24d-a5c95db2df8a  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks determined at compile time: 1  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:9  INFO : Submitting tokens for job: job\_1490094830773\_0322  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0322/  INFO : Starting Job = job\_1490094830773\_0322, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0322/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0322  INFO : Hadoop job information for Stage-1: number of mappers: 9; number of reducers: 1  INFO : 2017-04-11 03:09:04,812 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 03:09:13,104 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 2.93 sec  INFO : 2017-04-11 03:09:14,130 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 9.23 sec  INFO : 2017-04-11 03:09:15,153 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 33.43 sec  INFO : 2017-04-11 03:09:16,174 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 35.35 sec  INFO : 2017-04-11 03:09:18,257 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 44.1 sec  INFO : 2017-04-11 03:09:21,325 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 47.6 sec  INFO : 2017-04-11 03:09:24,390 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 50.97 sec  INFO : 2017-04-11 03:09:27,477 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 54.64 sec  INFO : 2017-04-11 03:09:32,624 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 55.58 sec  INFO : MapReduce Total cumulative CPU time: 55 seconds 580 msec  INFO : Ended Job = job\_1490094830773\_0322  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 9 Reduce: 1 Cumulative CPU: 55.58 sec HDFS Read: 214339923 HDFS Write: 19 SUCCESS  INFO : Total MapReduce CPU Time Spent: 55 seconds 580 msec  INFO : Completed executing command(queryId=hive\_20170411030909\_94a6e1dd-39ef-48ce-a24d-a5c95db2df8a); Time taken: 32.797 seconds  INFO : OK |

## User wise Max Date

|  |
| --- |
| select username,max(date) from postimport\_rc group by username;  INFO : Compiling command(queryId=hive\_20170411031010\_c4ed20fc-1a84-480f-b674-d68ec3b04b80): select username,max(date) from postimport\_rc group by username  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:username, type:string, comment:null), FieldSchema(name:\_c1, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411031010\_c4ed20fc-1a84-480f-b674-d68ec3b04b80); Time taken: 0.1 seconds  INFO : Executing command(queryId=hive\_20170411031010\_c4ed20fc-1a84-480f-b674-d68ec3b04b80): select username,max(date) from postimport\_rc group by username  INFO : Query ID = hive\_20170411031010\_c4ed20fc-1a84-480f-b674-d68ec3b04b80  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks not specified. Estimated from input data size: 34  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:9  INFO : Submitting tokens for job: job\_1490094830773\_0323  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0323/  INFO : Starting Job = job\_1490094830773\_0323, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0323/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0323  INFO : Hadoop job information for Stage-1: number of mappers: 9; number of reducers: 34  INFO : 2017-04-11 03:10:06,420 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 03:10:14,658 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 3.61 sec  INFO : 2017-04-11 03:10:16,711 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 34.56 sec  INFO : 2017-04-11 03:10:17,733 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 35.6 sec  INFO : 2017-04-11 03:10:18,758 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 40.54 sec  INFO : 2017-04-11 03:10:19,787 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 42.71 sec  INFO : 2017-04-11 03:10:20,827 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 43.93 sec  INFO : 2017-04-11 03:10:21,849 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 59.26 sec  INFO : 2017-04-11 03:10:24,916 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 68.0 sec  INFO : 2017-04-11 03:10:25,937 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 68.76 sec  INFO : 2017-04-11 03:10:32,219 Stage-1 map = 100%, reduce = 12%, Cumulative CPU 74.14 sec  INFO : 2017-04-11 03:10:33,243 Stage-1 map = 100%, reduce = 41%, Cumulative CPU 86.77 sec  INFO : 2017-04-11 03:10:34,275 Stage-1 map = 100%, reduce = 44%, Cumulative CPU 88.06 sec  INFO : 2017-04-11 03:10:36,387 Stage-1 map = 100%, reduce = 50%, Cumulative CPU 90.71 sec  INFO : 2017-04-11 03:10:37,432 Stage-1 map = 100%, reduce = 53%, Cumulative CPU 91.89 sec  INFO : 2017-04-11 03:10:38,474 Stage-1 map = 100%, reduce = 62%, Cumulative CPU 95.5 sec  INFO : 2017-04-11 03:10:39,518 Stage-1 map = 100%, reduce = 68%, Cumulative CPU 97.84 sec  INFO : 2017-04-11 03:10:40,551 Stage-1 map = 100%, reduce = 88%, Cumulative CPU 106.88 sec  INFO : 2017-04-11 03:10:41,580 Stage-1 map = 100%, reduce = 94%, Cumulative CPU 109.62 sec  INFO : 2017-04-11 03:10:42,633 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 112.01 sec  INFO : MapReduce Total cumulative CPU time: 1 minutes 52 seconds 10 msec  INFO : Ended Job = job\_1490094830773\_0323  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 9 Reduce: 34 Cumulative CPU: 112.01 sec HDFS Read: 328623329 HDFS Write: 33077 SUCCESS  INFO : Total MapReduce CPU Time Spent: 1 minutes 52 seconds 10 msec  INFO : Completed executing command(queryId=hive\_20170411031010\_c4ed20fc-1a84-480f-b674-d68ec3b04b80); Time taken: 41.184 seconds  INFO : OK |

## Select All Columns

|  |
| --- |
| select \* from postimport\_rc  INFO : Compiling command(queryId=hive\_20170411034949\_835dde0b-7a0b-4ac3-9167-5da337455607): select \* from postimport\_rc  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_rc.username, type:string, comment:null), FieldSchema(name:postimport\_rc.machinename, type:string, comment:null), FieldSchema(name:postimport\_rc.machineid, type:string, comment:null), FieldSchema(name:postimport\_rc.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_rc.domainname, type:string, comment:null), FieldSchema(name:postimport\_rc.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_rc.modulename, type:string, comment:null), FieldSchema(name:postimport\_rc.filename, type:string, comment:null), FieldSchema(name:postimport\_rc.productname, type:string, comment:null), FieldSchema(name:postimport\_rc.companyname, type:string, comment:null), FieldSchema(name:postimport\_rc.description, type:string, comment:null), FieldSchema(name:postimport\_rc.text, type:string, comment:null), FieldSchema(name:postimport\_rc.activetime, type:double, comment:null), FieldSchema(name:postimport\_rc.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_rc.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_rc.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411034949\_835dde0b-7a0b-4ac3-9167-5da337455607); Time taken: 0.098 seconds  INFO : Executing command(queryId=hive\_20170411034949\_835dde0b-7a0b-4ac3-9167-5da337455607): select \* from postimport\_rc  INFO : Query ID = hive\_20170411034949\_835dde0b-7a0b-4ac3-9167-5da337455607  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:9  INFO : Submitting tokens for job: job\_1490094830773\_0327  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0327/  INFO : Starting Job = job\_1490094830773\_0327, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0327/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0327  INFO : Hadoop job information for Stage-1: number of mappers: 9; number of reducers: 0  INFO : 2017-04-11 03:49:30,462 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 03:49:50,219 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 111.69 sec  INFO : 2017-04-11 03:49:51,242 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 121.04 sec  INFO : 2017-04-11 03:49:56,368 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 156.81 sec  INFO : 2017-04-11 03:50:00,473 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 183.85 sec  INFO : 2017-04-11 03:50:02,523 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 199.11 sec  INFO : 2017-04-11 03:50:03,550 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 205.76 sec  INFO : 2017-04-11 03:50:05,600 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 220.3 sec  INFO : 2017-04-11 03:50:11,749 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 263.47 sec  INFO : 2017-04-11 03:50:17,883 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 300.16 sec  INFO : 2017-04-11 03:50:18,905 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 306.87 sec  INFO : 2017-04-11 03:50:20,950 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 317.7 sec  INFO : 2017-04-11 03:50:21,973 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 324.29 sec  INFO : 2017-04-11 03:50:24,017 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 334.25 sec  INFO : 2017-04-11 03:50:27,092 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 350.76 sec  INFO : 2017-04-11 03:50:30,163 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 365.73 sec  INFO : 2017-04-11 03:50:37,331 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 403.99 sec  INFO : 2017-04-11 03:50:39,389 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 413.85 sec  INFO : 2017-04-11 03:50:40,410 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 420.74 sec  INFO : 2017-04-11 03:50:41,433 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 425.56 sec  INFO : 2017-04-11 03:50:50,628 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 474.09 sec  INFO : 2017-04-11 03:50:55,728 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 494.19 sec  INFO : 2017-04-11 03:51:01,847 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 518.49 sec  INFO : 2017-04-11 03:51:02,866 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 524.16 sec  INFO : 2017-04-11 03:51:10,006 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 553.04 sec  INFO : 2017-04-11 03:51:12,047 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 564.06 sec  INFO : 2017-04-11 03:51:14,086 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 570.88 sec  INFO : 2017-04-11 03:51:15,106 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 578.51 sec  INFO : 2017-04-11 03:51:20,200 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 598.24 sec  INFO : 2017-04-11 03:51:21,222 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 605.0 sec  INFO : 2017-04-11 03:51:27,351 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 629.09 sec  INFO : 2017-04-11 03:51:34,493 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 653.52 sec  INFO : 2017-04-11 03:51:35,512 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 657.9 sec  INFO : 2017-04-11 03:51:38,576 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 671.92 sec  INFO : 2017-04-11 03:51:44,692 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 695.37 sec  INFO : 2017-04-11 03:51:50,815 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 721.24 sec  INFO : 2017-04-11 03:51:53,883 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 733.86 sec  INFO : 2017-04-11 03:51:56,946 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 746.52 sec  INFO : 2017-04-11 03:51:58,990 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 750.13 sec  INFO : 2017-04-11 03:52:01,036 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 761.03 sec  INFO : 2017-04-11 03:52:02,055 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 766.03 sec  INFO : 2017-04-11 03:52:09,208 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 798.25 sec  INFO : 2017-04-11 03:52:18,400 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 835.74 sec  INFO : 2017-04-11 03:52:21,468 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 848.1 sec  INFO : 2017-04-11 03:52:26,570 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 863.6 sec  INFO : 2017-04-11 03:52:29,636 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 874.6 sec  INFO : 2017-04-11 03:52:30,656 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 883.59 sec  INFO : 2017-04-11 03:52:33,734 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 895.99 sec  INFO : 2017-04-11 03:52:35,782 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 899.74 sec  INFO : 2017-04-11 03:52:39,872 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 926.97 sec  INFO : 2017-04-11 03:52:42,944 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 942.97 sec  INFO : 2017-04-11 03:52:44,989 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 952.77 sec  INFO : 2017-04-11 03:52:47,030 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 957.06 sec  INFO : 2017-04-11 03:52:48,050 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 962.54 sec  INFO : 2017-04-11 03:52:54,167 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 977.13 sec  INFO : 2017-04-11 03:52:58,246 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 984.08 sec  INFO : 2017-04-11 03:53:05,376 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 991.94 sec  INFO : MapReduce Total cumulative CPU time: 16 minutes 31 seconds 940 msec  INFO : Ended Job = job\_1490094830773\_0327  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 9 Cumulative CPU: 991.94 sec HDFS Read: 2257235049 HDFS Write: 3633966185 SUCCESS  INFO : Total MapReduce CPU Time Spent: 16 minutes 31 seconds 940 msec  INFO : Completed executing command(queryId=hive\_20170411034949\_835dde0b-7a0b-4ac3-9167-5da337455607); Time taken: 224.2 seconds  INFO : OK |

## Select All Columns with filter Condition

|  |
| --- |
| select \* from postimport\_rc where username = '670063006100750074006F007500730065007200';  INFO : Compiling command(queryId=hive\_20170411035656\_163f863e-716e-484b-8baf-a6b9694a912d): select \* from postimport\_rc where username = '670063006100750074006F007500730065007200'  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_rc.username, type:string, comment:null), FieldSchema(name:postimport\_rc.machinename, type:string, comment:null), FieldSchema(name:postimport\_rc.machineid, type:string, comment:null), FieldSchema(name:postimport\_rc.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_rc.domainname, type:string, comment:null), FieldSchema(name:postimport\_rc.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_rc.modulename, type:string, comment:null), FieldSchema(name:postimport\_rc.filename, type:string, comment:null), FieldSchema(name:postimport\_rc.productname, type:string, comment:null), FieldSchema(name:postimport\_rc.companyname, type:string, comment:null), FieldSchema(name:postimport\_rc.description, type:string, comment:null), FieldSchema(name:postimport\_rc.text, type:string, comment:null), FieldSchema(name:postimport\_rc.activetime, type:double, comment:null), FieldSchema(name:postimport\_rc.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_rc.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_rc.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411035656\_163f863e-716e-484b-8baf-a6b9694a912d); Time taken: 0.1 seconds  INFO : Executing command(queryId=hive\_20170411035656\_163f863e-716e-484b-8baf-a6b9694a912d): select \* from postimport\_rc where username = '670063006100750074006F007500730065007200'  INFO : Query ID = hive\_20170411035656\_163f863e-716e-484b-8baf-a6b9694a912d  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:9  INFO : Submitting tokens for job: job\_1490094830773\_0328  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0328/  INFO : Starting Job = job\_1490094830773\_0328, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0328/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0328  INFO : Hadoop job information for Stage-1: number of mappers: 9; number of reducers: 0  INFO : 2017-04-11 03:56:47,383 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 03:56:56,724 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 3.99 sec  INFO : 2017-04-11 03:56:57,750 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 38.34 sec  INFO : 2017-04-11 03:56:59,798 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 58.07 sec  INFO : 2017-04-11 03:57:00,819 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 69.8 sec  INFO : 2017-04-11 03:57:01,842 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 71.76 sec  INFO : 2017-04-11 03:57:02,868 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 79.06 sec  INFO : 2017-04-11 03:57:04,927 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 80.19 sec  INFO : 2017-04-11 03:57:05,948 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 84.01 sec  INFO : MapReduce Total cumulative CPU time: 1 minutes 24 seconds 10 msec  INFO : Ended Job = job\_1490094830773\_0328  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 9 Cumulative CPU: 84.01 sec HDFS Read: 2257239297 HDFS Write: 2071791 SUCCESS  INFO : Total MapReduce CPU Time Spent: 1 minutes 24 seconds 10 msec  INFO : Completed executing command(queryId=hive\_20170411035656\_163f863e-716e-484b-8baf-a6b9694a912d); Time taken: 26.082 seconds  INFO : OK |

# ORC File format Zlib Compression

|  |
| --- |
| create table postimport\_orc like postimport\_text stored as ORC tblproperties ("orc.compress"="ZLIB"); |

|  |
| --- |
| SET hive.exec.compress.output=true; |

## Load data to ORC Table

|  |
| --- |
| hive> insert overwrite table postimport\_orc select \* from postimport\_text;  Query ID = shalaj\_20170411042020\_053980d1-9c56-4ce1-bbe4-adc03617b924  Total jobs = 1  Launching Job 1 out of 1  Number of reduce tasks is set to 0 since there's no reduce operator  Starting Job = job\_1490094830773\_0329, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0329/  Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0329  Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  2017-04-11 04:20:47,465 Stage-1 map = 0%, reduce = 0%  2017-04-11 04:21:01,576 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 66.46 sec  2017-04-11 04:21:02,599 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 75.67 sec  2017-04-11 04:21:04,754 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 92.99 sec  2017-04-11 04:21:05,781 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 99.43 sec  2017-04-11 04:21:06,807 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 105.21 sec  2017-04-11 04:21:07,832 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 118.78 sec  2017-04-11 04:21:11,000 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 140.26 sec  2017-04-11 04:21:12,089 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 145.22 sec  2017-04-11 04:21:15,284 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 156.16 sec  2017-04-11 04:21:16,467 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 163.73 sec  2017-04-11 04:21:18,802 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 169.94 sec  2017-04-11 04:21:19,885 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 173.44 sec  2017-04-11 04:21:20,995 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 183.61 sec  2017-04-11 04:21:27,357 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 195.01 sec  2017-04-11 04:21:31,483 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 256.45 sec  2017-04-11 04:21:32,503 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 267.56 sec  2017-04-11 04:21:33,823 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 281.3 sec  2017-04-11 04:21:34,845 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 289.55 sec  2017-04-11 04:21:35,868 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 295.3 sec  2017-04-11 04:21:37,920 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 314.16 sec  2017-04-11 04:21:40,009 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 327.42 sec  2017-04-11 04:21:41,497 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 333.8 sec  2017-04-11 04:21:43,557 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 336.9 sec  2017-04-11 04:21:45,738 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 341.51 sec  2017-04-11 04:21:48,229 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 359.72 sec  2017-04-11 04:21:49,459 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 368.04 sec  2017-04-11 04:21:50,750 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 377.48 sec  2017-04-11 04:21:54,343 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 387.66 sec  2017-04-11 04:21:56,940 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 389.86 sec  2017-04-11 04:22:01,532 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 412.95 sec  2017-04-11 04:22:03,600 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 421.58 sec  2017-04-11 04:22:04,635 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 434.58 sec  2017-04-11 04:22:07,706 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 459.14 sec  2017-04-11 04:22:09,747 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 481.31 sec  2017-04-11 04:22:10,770 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 485.99 sec  2017-04-11 04:22:13,955 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 500.59 sec  2017-04-11 04:22:15,004 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 506.76 sec  2017-04-11 04:22:18,243 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 517.29 sec  2017-04-11 04:22:20,302 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 538.18 sec  2017-04-11 04:22:22,489 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 548.12 sec  2017-04-11 04:22:24,647 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 561.62 sec  2017-04-11 04:22:29,249 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 577.02 sec  2017-04-11 04:22:33,440 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 604.68 sec  2017-04-11 04:22:35,516 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 614.07 sec  2017-04-11 04:22:36,541 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 627.75 sec  2017-04-11 04:22:37,568 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 638.63 sec  2017-04-11 04:22:38,588 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 645.77 sec  2017-04-11 04:22:39,620 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 652.98 sec  2017-04-11 04:22:40,642 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 677.49 sec  2017-04-11 04:22:41,697 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 682.37 sec  2017-04-11 04:22:42,717 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 689.66 sec  2017-04-11 04:22:43,752 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 703.11 sec  2017-04-11 04:22:46,815 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 721.83 sec  2017-04-11 04:22:47,835 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 723.85 sec  2017-04-11 04:22:48,855 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 727.02 sec  2017-04-11 04:22:50,993 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 732.38 sec  2017-04-11 04:22:52,116 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 738.42 sec  2017-04-11 04:22:54,445 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 751.69 sec  2017-04-11 04:22:55,530 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 754.63 sec  2017-04-11 04:22:56,608 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 759.74 sec  2017-04-11 04:22:59,835 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 788.07 sec  2017-04-11 04:23:01,889 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 800.92 sec  2017-04-11 04:23:02,912 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 805.31 sec  2017-04-11 04:23:03,932 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 809.07 sec  2017-04-11 04:23:04,959 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 827.16 sec  2017-04-11 04:23:06,061 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 832.77 sec  2017-04-11 04:23:07,222 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 836.0 sec  2017-04-11 04:23:08,441 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 846.38 sec  2017-04-11 04:23:10,769 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 850.05 sec  2017-04-11 04:23:32,207 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 861.78 sec  MapReduce Total cumulative CPU time: 14 minutes 21 seconds 780 msec  Ended Job = job\_1490094830773\_0329  Stage-4 is filtered out by condition resolver.  Stage-3 is selected by condition resolver.  Stage-5 is filtered out by condition resolver.  Starting Job = job\_1490094830773\_0330, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0330/  Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0330  Hadoop job information for Stage-3: number of mappers: 5; number of reducers: 0  2017-04-11 04:23:40,714 Stage-3 map = 0%, reduce = 0%  2017-04-11 04:23:49,147 Stage-3 map = 20%, reduce = 0%, Cumulative CPU 2.6 sec  2017-04-11 04:23:51,304 Stage-3 map = 26%, reduce = 0%, Cumulative CPU 5.51 sec  2017-04-11 04:23:52,374 Stage-3 map = 42%, reduce = 0%, Cumulative CPU 9.4 sec  2017-04-11 04:23:54,519 Stage-3 map = 53%, reduce = 0%, Cumulative CPU 10.36 sec  2017-04-11 04:23:57,738 Stage-3 map = 57%, reduce = 0%, Cumulative CPU 10.92 sec  2017-04-11 04:23:59,907 Stage-3 map = 63%, reduce = 0%, Cumulative CPU 11.42 sec  2017-04-11 04:24:00,971 Stage-3 map = 64%, reduce = 0%, Cumulative CPU 11.67 sec  2017-04-11 04:24:03,125 Stage-3 map = 67%, reduce = 0%, Cumulative CPU 12.14 sec  2017-04-11 04:24:04,206 Stage-3 map = 71%, reduce = 0%, Cumulative CPU 12.29 sec  2017-04-11 04:24:06,356 Stage-3 map = 73%, reduce = 0%, Cumulative CPU 12.46 sec  2017-04-11 04:24:09,589 Stage-3 map = 77%, reduce = 0%, Cumulative CPU 13.26 sec  2017-04-11 04:24:11,729 Stage-3 map = 80%, reduce = 0%, Cumulative CPU 13.52 sec  2017-04-11 04:24:12,789 Stage-3 map = 86%, reduce = 0%, Cumulative CPU 13.79 sec  2017-04-11 04:24:14,912 Stage-3 map = 87%, reduce = 0%, Cumulative CPU 13.9 sec  2017-04-11 04:24:15,969 Stage-3 map = 90%, reduce = 0%, Cumulative CPU 14.06 sec  2017-04-11 04:24:18,079 Stage-3 map = 92%, reduce = 0%, Cumulative CPU 14.32 sec  2017-04-11 04:24:21,257 Stage-3 map = 95%, reduce = 0%, Cumulative CPU 14.55 sec  2017-04-11 04:24:24,423 Stage-3 map = 96%, reduce = 0%, Cumulative CPU 14.77 sec  2017-04-11 04:24:27,574 Stage-3 map = 99%, reduce = 0%, Cumulative CPU 14.97 sec  2017-04-11 04:24:28,593 Stage-3 map = 100%, reduce = 0%, Cumulative CPU 15.07 sec  MapReduce Total cumulative CPU time: 15 seconds 70 msec  Ended Job = job\_1490094830773\_0330  Loading data to table esplus.postimport\_orc  Table esplus.postimport\_orc stats: [numFiles=5, numRows=42286538, totalSize=1230693597, rawDataSize=64936217951]  MapReduce Jobs Launched:  Stage-Stage-1: Map: 74 Cumulative CPU: 862.35 sec HDFS Read: 19486328458 HDFS Write: 1215540720 SUCCESS  Stage-Stage-3: Map: 5 Cumulative CPU: 15.07 sec HDFS Read: 1217868200 HDFS Write: 1230693597 SUCCESS  Total MapReduce CPU Time Spent: 14 minutes 37 seconds 420 msec  OK  Time taken: 235.644 seconds  hive> |

## Data Size

|  |
| --- |
| hive> dfs -du -s -h /user/hive/warehouse/esplus.db/postimport\_orc;  1.1 G 3.4 G /user/hive/warehouse/esplus.db/postimport\_orc |

 It creates lesser files than RC format, a single file as the output of each task, which reduces the NameNode's load

## Select some columns

SET hive.exec.compress.output=true;

|  |
| --- |
| select productname,username from postimport\_orc;  INFO : Compiling command(queryId=hive\_20170411044242\_b42b9ce8-55eb-498d-8d1d-139f54203c4d): select productname,username from postimport\_orc  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:productname, type:string, comment:null), FieldSchema(name:username, type:string, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411044242\_b42b9ce8-55eb-498d-8d1d-139f54203c4d); Time taken: 0.099 seconds  INFO : Executing command(queryId=hive\_20170411044242\_b42b9ce8-55eb-498d-8d1d-139f54203c4d): select productname,username from postimport\_orc  INFO : Query ID = hive\_20170411044242\_b42b9ce8-55eb-498d-8d1d-139f54203c4d  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:5  INFO : Submitting tokens for job: job\_1490094830773\_0333  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0333/  INFO : Starting Job = job\_1490094830773\_0333, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0333/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0333  INFO : Hadoop job information for Stage-1: number of mappers: 5; number of reducers: 0  INFO : 2017-04-11 04:42:23,834 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 04:42:34,065 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 20.19 sec  INFO : 2017-04-11 04:42:39,180 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 50.41 sec  INFO : 2017-04-11 04:42:40,203 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 56.51 sec  INFO : 2017-04-11 04:42:41,227 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 59.5 sec  INFO : 2017-04-11 04:42:45,320 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 80.48 sec  INFO : 2017-04-11 04:42:47,365 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 90.93 sec  INFO : 2017-04-11 04:42:48,387 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 95.71 sec  INFO : 2017-04-11 04:42:49,414 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 97.08 sec  INFO : MapReduce Total cumulative CPU time: 1 minutes 37 seconds 80 msec  INFO : Ended Job = job\_1490094830773\_0333  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 5 Cumulative CPU: 97.08 sec HDFS Read: 58815202 HDFS Write: 233679762 SUCCESS  INFO : Total MapReduce CPU Time Spent: 1 minutes 37 seconds 80 msec  INFO : Completed executing command(queryId=hive\_20170411044242\_b42b9ce8-55eb-498d-8d1d-139f54203c4d); Time taken: 32.097 seconds  INFO : OK |

**System Counter**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Maps Total** | **Reduces Total** | **Total** |
| File Bytes Read | 0 | 0 | 0 |
| File Bytes Written | 1202565 | 0 | 1202565 |
| File Large Read Ops | 0 | 0 | 0 |
| File Read Ops | 0 | 0 | 0 |
| File Write Ops | 0 | 0 | 0 |
| Hdfs Bytes Read | 58722627 | 0 | 58722627 |
| Hdfs Bytes Written | 2439779252 | 0 | 2439779252 |
| Hdfs Large Read Ops | 0 | 0 | 0 |
| Hdfs Read Ops | 70 | 0 | 70 |
| Hdfs Write Ops | 10 | 0 | 10 |

Here it reads 56 MB of data, RC Format reads .8 GB while normal Text format reads all 18 GB of data

## Check Max Date

|  |
| --- |
| select max(date) from postimport\_orc;  INFO : Compiling command(queryId=hive\_20170411044343\_8c7d8dae-25bd-4687-980b-7b1ca8d89a9a): select max(date) from postimport\_orc  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:\_c0, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411044343\_8c7d8dae-25bd-4687-980b-7b1ca8d89a9a); Time taken: 0.143 seconds  INFO : Executing command(queryId=hive\_20170411044343\_8c7d8dae-25bd-4687-980b-7b1ca8d89a9a): select max(date) from postimport\_orc  INFO : Query ID = hive\_20170411044343\_8c7d8dae-25bd-4687-980b-7b1ca8d89a9a  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks determined at compile time: 1  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:5  INFO : Submitting tokens for job: job\_1490094830773\_0334  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0334/  INFO : Starting Job = job\_1490094830773\_0334, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0334/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0334  INFO : Hadoop job information for Stage-1: number of mappers: 5; number of reducers: 1  INFO : 2017-04-11 04:43:14,630 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 04:43:19,754 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 4.3 sec  INFO : 2017-04-11 04:43:21,829 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 5.37 sec  INFO : 2017-04-11 04:43:24,918 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 6.34 sec  INFO : MapReduce Total cumulative CPU time: 6 seconds 340 msec  INFO : Ended Job = job\_1490094830773\_0334  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 5 Reduce: 1 Cumulative CPU: 6.34 sec HDFS Read: 771879 HDFS Write: 19 SUCCESS  INFO : Total MapReduce CPU Time Spent: 6 seconds 340 msec  INFO : Completed executing command(queryId=hive\_20170411044343\_8c7d8dae-25bd-4687-980b-7b1ca8d89a9a); Time taken: 15.388 seconds  INFO : OK |

## User wise Max Date

|  |
| --- |
| select username,max(date) from postimport\_orc group by username;  INFO : Compiling command(queryId=hive\_20170411044343\_ec845cfd-5245-48ea-9709-cd520c92e50a): select username,max(date) from postimport\_orc group by username  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:username, type:string, comment:null), FieldSchema(name:\_c1, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411044343\_ec845cfd-5245-48ea-9709-cd520c92e50a); Time taken: 0.107 seconds  INFO : Executing command(queryId=hive\_20170411044343\_ec845cfd-5245-48ea-9709-cd520c92e50a): select username,max(date) from postimport\_orc group by username  INFO : Query ID = hive\_20170411044343\_ec845cfd-5245-48ea-9709-cd520c92e50a  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks not specified. Estimated from input data size: 19  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:5  INFO : Submitting tokens for job: job\_1490094830773\_0335  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0335/  INFO : Starting Job = job\_1490094830773\_0335, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0335/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0335  INFO : Hadoop job information for Stage-1: number of mappers: 5; number of reducers: 19  INFO : 2017-04-11 04:43:54,038 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 04:44:00,261 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 11.02 sec  INFO : 2017-04-11 04:44:02,313 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 13.83 sec  INFO : 2017-04-11 04:44:07,587 Stage-1 map = 100%, reduce = 5%, Cumulative CPU 15.27 sec  INFO : 2017-04-11 04:44:08,618 Stage-1 map = 100%, reduce = 79%, Cumulative CPU 33.1 sec  INFO : 2017-04-11 04:44:11,730 Stage-1 map = 100%, reduce = 84%, Cumulative CPU 34.13 sec  INFO : 2017-04-11 04:44:13,799 Stage-1 map = 100%, reduce = 95%, Cumulative CPU 36.66 sec  INFO : 2017-04-11 04:44:14,822 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 39.76 sec  INFO : MapReduce Total cumulative CPU time: 39 seconds 760 msec  INFO : Ended Job = job\_1490094830773\_0335  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 5 Reduce: 19 Cumulative CPU: 39.76 sec HDFS Read: 24516090 HDFS Write: 30462 SUCCESS  INFO : Total MapReduce CPU Time Spent: 39 seconds 760 msec  INFO : Completed executing command(queryId=hive\_20170411044343\_ec845cfd-5245-48ea-9709-cd520c92e50a); Time taken: 25.749 seconds  INFO : OK |

## Select All Columns

|  |
| --- |
| select \* from postimport\_orc  INFO : Compiling command(queryId=hive\_20170411044545\_b3a33dd8-40a9-4d36-b276-3ae3e951b24b): select \* from postimport\_orc  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_orc.username, type:string, comment:null), FieldSchema(name:postimport\_orc.machinename, type:string, comment:null), FieldSchema(name:postimport\_orc.machineid, type:string, comment:null), FieldSchema(name:postimport\_orc.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_orc.domainname, type:string, comment:null), FieldSchema(name:postimport\_orc.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_orc.modulename, type:string, comment:null), FieldSchema(name:postimport\_orc.filename, type:string, comment:null), FieldSchema(name:postimport\_orc.productname, type:string, comment:null), FieldSchema(name:postimport\_orc.companyname, type:string, comment:null), FieldSchema(name:postimport\_orc.description, type:string, comment:null), FieldSchema(name:postimport\_orc.text, type:string, comment:null), FieldSchema(name:postimport\_orc.activetime, type:double, comment:null), FieldSchema(name:postimport\_orc.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_orc.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_orc.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411044545\_b3a33dd8-40a9-4d36-b276-3ae3e951b24b); Time taken: 0.104 seconds  INFO : Executing command(queryId=hive\_20170411044545\_b3a33dd8-40a9-4d36-b276-3ae3e951b24b): select \* from postimport\_orc  INFO : Query ID = hive\_20170411044545\_b3a33dd8-40a9-4d36-b276-3ae3e951b24b  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:5  INFO : Submitting tokens for job: job\_1490094830773\_0336  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0336/  INFO : Starting Job = job\_1490094830773\_0336, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0336/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0336  INFO : Hadoop job information for Stage-1: number of mappers: 5; number of reducers: 0  INFO : 2017-04-11 04:45:19,568 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 04:45:37,979 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 67.26 sec  INFO : 2017-04-11 04:46:38,230 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 270.36 sec  INFO : 2017-04-11 04:47:02,715 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 343.01 sec  INFO : 2017-04-11 04:48:02,879 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 537.65 sec  INFO : 2017-04-11 04:48:08,993 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 561.75 sec  INFO : 2017-04-11 04:48:18,168 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 598.15 sec  INFO : 2017-04-11 04:48:20,206 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 600.17 sec  INFO : 2017-04-11 04:48:22,244 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 611.97 sec  INFO : 2017-04-11 04:49:22,404 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 779.41 sec  INFO : 2017-04-11 04:49:39,764 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 818.38 sec  INFO : 2017-04-11 04:49:44,867 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 828.88 sec  INFO : 2017-04-11 04:49:53,078 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 837.34 sec  INFO : MapReduce Total cumulative CPU time: 13 minutes 57 seconds 340 msec  INFO : Ended Job = job\_1490094830773\_0336  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 5 Cumulative CPU: 837.34 sec HDFS Read: 1214311369 HDFS Write: 3633963607 SUCCESS  INFO : Total MapReduce CPU Time Spent: 13 minutes 57 seconds 340 msec  INFO : Completed executing command(queryId=hive\_20170411044545\_b3a33dd8-40a9-4d36-b276-3ae3e951b24b); Time taken: 282.546 seconds  INFO : OK |

## Select All Columns with filter Condition

|  |
| --- |
| select \* from postimport\_orc where username = '670063006100750074006F007500730065007200'  INFO : Compiling command(queryId=hive\_20170411045151\_cea8f40e-e66a-4508-b0e4-02ddaa449159): select \* from postimport\_orc where username = '670063006100750074006F007500730065007200'  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_orc.username, type:string, comment:null), FieldSchema(name:postimport\_orc.machinename, type:string, comment:null), FieldSchema(name:postimport\_orc.machineid, type:string, comment:null), FieldSchema(name:postimport\_orc.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_orc.domainname, type:string, comment:null), FieldSchema(name:postimport\_orc.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_orc.modulename, type:string, comment:null), FieldSchema(name:postimport\_orc.filename, type:string, comment:null), FieldSchema(name:postimport\_orc.productname, type:string, comment:null), FieldSchema(name:postimport\_orc.companyname, type:string, comment:null), FieldSchema(name:postimport\_orc.description, type:string, comment:null), FieldSchema(name:postimport\_orc.text, type:string, comment:null), FieldSchema(name:postimport\_orc.activetime, type:double, comment:null), FieldSchema(name:postimport\_orc.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_orc.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_orc.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411045151\_cea8f40e-e66a-4508-b0e4-02ddaa449159); Time taken: 0.182 seconds  INFO : Executing command(queryId=hive\_20170411045151\_cea8f40e-e66a-4508-b0e4-02ddaa449159): select \* from postimport\_orc where username = '670063006100750074006F007500730065007200'  INFO : Query ID = hive\_20170411045151\_cea8f40e-e66a-4508-b0e4-02ddaa449159  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:5  INFO : Submitting tokens for job: job\_1490094830773\_0337  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0337/  INFO : Starting Job = job\_1490094830773\_0337, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0337/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0337  INFO : Hadoop job information for Stage-1: number of mappers: 5; number of reducers: 0  INFO : 2017-04-11 04:51:56,110 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 04:52:06,373 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 26.46 sec  INFO : 2017-04-11 04:52:07,395 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 29.39 sec  INFO : 2017-04-11 04:52:08,434 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 36.5 sec  INFO : 2017-04-11 04:52:09,456 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 38.84 sec  INFO : 2017-04-11 04:52:18,762 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 41.31 sec  INFO : MapReduce Total cumulative CPU time: 41 seconds 310 msec  INFO : Ended Job = job\_1490094830773\_0337  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 5 Cumulative CPU: 41.31 sec HDFS Read: 1215615878 HDFS Write: 2070791 SUCCESS  INFO : Total MapReduce CPU Time Spent: 41 seconds 310 msec  INFO : Completed executing command(queryId=hive\_20170411045151\_cea8f40e-e66a-4508-b0e4-02ddaa449159); Time taken: 30.267 seconds  INFO : OK |

# ORC File format snappy Compression

|  |
| --- |
| create table postimport\_orc like postimport\_text stored as ORC tblproperties ("orc.compress"="SNAPPY"); |

|  |
| --- |
| SET hive.exec.compress.output=true; |

## Load data to ORC Table

|  |
| --- |
| hive> insert overwrite table postimport\_orc select \* from postimport\_text;  Query ID = shalaj\_20170411050808\_fa41329b-c6a2-4ce0-9d32-f5f8ed9d3717  Total jobs = 1  Launching Job 1 out of 1  Number of reduce tasks is set to 0 since there's no reduce operator  Starting Job = job\_1490094830773\_0344, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0344/  Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0344  Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  2017-04-11 05:09:04,404 Stage-1 map = 0%, reduce = 0%  2017-04-11 05:09:19,000 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 78.97 sec  2017-04-11 05:09:20,121 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 87.47 sec  2017-04-11 05:09:21,243 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 95.83 sec  2017-04-11 05:09:25,382 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 120.99 sec  2017-04-11 05:09:26,673 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 128.98 sec  2017-04-11 05:09:27,748 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 141.13 sec  2017-04-11 05:09:30,052 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 147.76 sec  2017-04-11 05:09:31,295 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 154.21 sec  2017-04-11 05:09:36,487 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 171.04 sec  2017-04-11 05:09:37,665 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 176.55 sec  2017-04-11 05:09:42,765 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 181.05 sec  2017-04-11 05:09:49,689 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 198.7 sec  2017-04-11 05:09:52,967 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 202.82 sec  2017-04-11 05:09:56,205 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 220.91 sec  2017-04-11 05:10:01,515 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 248.88 sec  2017-04-11 05:10:02,535 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 260.87 sec  2017-04-11 05:10:03,603 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 267.83 sec  2017-04-11 05:10:05,841 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 301.84 sec  2017-04-11 05:10:07,895 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 320.24 sec  2017-04-11 05:10:08,930 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 326.74 sec  2017-04-11 05:10:10,022 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 333.78 sec  2017-04-11 05:10:11,059 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 344.43 sec  2017-04-11 05:10:12,179 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 347.58 sec  2017-04-11 05:10:14,509 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 358.15 sec  2017-04-11 05:10:15,716 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 360.55 sec  2017-04-11 05:10:17,182 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 371.07 sec  2017-04-11 05:10:19,499 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 383.32 sec  2017-04-11 05:10:21,804 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 398.23 sec  2017-04-11 05:10:23,934 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 410.0 sec  2017-04-11 05:10:24,996 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 419.45 sec  2017-04-11 05:10:27,399 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 426.94 sec  2017-04-11 05:10:28,494 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 440.82 sec  2017-04-11 05:10:30,658 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 453.74 sec  2017-04-11 05:10:31,744 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 460.24 sec  2017-04-11 05:10:32,812 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 463.56 sec  2017-04-11 05:10:33,873 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 475.06 sec  2017-04-11 05:10:34,931 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 479.07 sec  2017-04-11 05:10:36,216 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 495.74 sec  2017-04-11 05:10:37,279 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 501.41 sec  2017-04-11 05:10:38,305 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 509.49 sec  2017-04-11 05:10:41,489 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 523.87 sec  2017-04-11 05:10:46,971 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 548.31 sec  2017-04-11 05:10:48,104 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 565.13 sec  2017-04-11 05:10:49,313 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 580.01 sec  2017-04-11 05:10:51,744 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 595.49 sec  2017-04-11 05:10:54,165 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 603.71 sec  2017-04-11 05:10:55,410 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 608.64 sec  2017-04-11 05:10:57,761 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 619.92 sec  2017-04-11 05:11:02,172 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 639.5 sec  2017-04-11 05:11:03,240 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 645.58 sec  2017-04-11 05:11:05,657 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 662.99 sec  2017-04-11 05:11:08,005 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 682.68 sec  2017-04-11 05:11:09,054 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 688.65 sec  2017-04-11 05:11:11,099 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 697.32 sec  2017-04-11 05:11:14,531 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 713.3 sec  2017-04-11 05:11:15,591 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 720.59 sec  2017-04-11 05:11:17,722 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 732.18 sec  2017-04-11 05:11:21,855 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 744.29 sec  2017-04-11 05:11:22,874 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 752.14 sec  2017-04-11 05:11:26,201 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 779.64 sec  2017-04-11 05:11:28,599 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 794.25 sec  2017-04-11 05:11:29,787 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 801.13 sec  2017-04-11 05:11:30,998 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 814.02 sec  2017-04-11 05:11:32,242 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 822.21 sec  2017-04-11 05:11:34,754 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 828.01 sec  2017-04-11 05:11:42,072 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 843.48 sec  2017-04-11 05:11:47,883 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 850.34 sec  2017-04-11 05:11:56,701 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 858.43 sec  MapReduce Total cumulative CPU time: 14 minutes 18 seconds 430 msec  Ended Job = job\_1490094830773\_0344  Stage-4 is selected by condition resolver.  Stage-3 is filtered out by condition resolver.  Stage-5 is filtered out by condition resolver.  Moving data to: hdfs://mac55:8020/user/hive/warehouse/esplus.db/postimport\_orc/.hive-staging\_hive\_2017-04-11\_05-08-55\_078\_5006362339089172464-1/-ext-10000  Loading data to table esplus.postimport\_orc  Table esplus.postimport\_orc stats: [numFiles=74, numRows=42286538, totalSize=2003238022, rawDataSize=64936217951]  MapReduce Jobs Launched:  Stage-Stage-1: Map: 74 Cumulative CPU: 858.58 sec HDFS Read: 19486328606 HDFS Write: 2003244534 SUCCESS  Total MapReduce CPU Time Spent: 14 minutes 18 seconds 580 msec  OK  Time taken: 186.12 seconds |

## Data Size

|  |
| --- |
| hive> dfs -du -s -h /user/hive/warehouse/esplus.db/postimport\_orc;  1.9 G 5.6 G /user/hive/warehouse/esplus.db/postimport\_orc |

## Select some columns

SET hive.exec.compress.output=true;

|  |
| --- |
| select productname,username from postimport\_orc;  INFO : Compiling command(queryId=hive\_20170411051212\_dcb79af3-9f95-410b-9a7d-1b7f3f7edd8b): select productname,username from postimport\_orc  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:productname, type:string, comment:null), FieldSchema(name:username, type:string, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411051212\_dcb79af3-9f95-410b-9a7d-1b7f3f7edd8b); Time taken: 0.104 seconds  INFO : Executing command(queryId=hive\_20170411051212\_dcb79af3-9f95-410b-9a7d-1b7f3f7edd8b): select productname,username from postimport\_orc  INFO : Query ID = hive\_20170411051212\_dcb79af3-9f95-410b-9a7d-1b7f3f7edd8b  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:8  INFO : Submitting tokens for job: job\_1490094830773\_0345  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0345/  INFO : Starting Job = job\_1490094830773\_0345, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0345/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0345  INFO : Hadoop job information for Stage-1: number of mappers: 8; number of reducers: 0  INFO : 2017-04-11 05:12:42,331 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 05:12:52,603 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 18.91 sec  INFO : 2017-04-11 05:12:53,630 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 37.81 sec  INFO : 2017-04-11 05:12:54,651 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 50.92 sec  INFO : 2017-04-11 05:12:55,676 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 58.85 sec  INFO : 2017-04-11 05:12:56,700 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 63.85 sec  INFO : 2017-04-11 05:12:57,722 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 69.75 sec  INFO : 2017-04-11 05:12:58,744 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 76.42 sec  INFO : 2017-04-11 05:12:59,764 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 80.62 sec  INFO : 2017-04-11 05:13:00,786 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 86.56 sec  INFO : 2017-04-11 05:13:01,807 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 92.79 sec  INFO : 2017-04-11 05:13:02,830 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 98.87 sec  INFO : 2017-04-11 05:13:03,851 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 101.05 sec  INFO : 2017-04-11 05:13:04,873 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 104.34 sec  INFO : 2017-04-11 05:13:05,896 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 105.38 sec  INFO : MapReduce Total cumulative CPU time: 1 minutes 45 seconds 380 msec  INFO : Ended Job = job\_1490094830773\_0345  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 8 Cumulative CPU: 105.38 sec HDFS Read: 95120704 HDFS Write: 233688883 SUCCESS  INFO : Total MapReduce CPU Time Spent: 1 minutes 45 seconds 380 msec  INFO : Completed executing command(queryId=hive\_20170411051212\_dcb79af3-9f95-410b-9a7d-1b7f3f7edd8b); Time taken: 28.702 seconds  INFO : OK |

## Check Max Date

|  |
| --- |
| select max(date) from postimport\_orc;  INFO : Compiling command(queryId=hive\_20170411051313\_d43d9c91-fa53-48f0-9a2e-687fffe9c018): select max(date) from postimport\_orc  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:\_c0, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411051313\_d43d9c91-fa53-48f0-9a2e-687fffe9c018); Time taken: 0.103 seconds  INFO : Executing command(queryId=hive\_20170411051313\_d43d9c91-fa53-48f0-9a2e-687fffe9c018): select max(date) from postimport\_orc  INFO : Query ID = hive\_20170411051313\_d43d9c91-fa53-48f0-9a2e-687fffe9c018  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks determined at compile time: 1  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:8  INFO : Submitting tokens for job: job\_1490094830773\_0346  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0346/  INFO : Starting Job = job\_1490094830773\_0346, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0346/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0346  INFO : Hadoop job information for Stage-1: number of mappers: 8; number of reducers: 1  INFO : 2017-04-11 05:13:51,011 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 05:13:55,157 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 0.93 sec  INFO : 2017-04-11 05:13:56,179 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 6.64 sec  INFO : 2017-04-11 05:13:58,244 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 8.98 sec  INFO : 2017-04-11 05:14:02,402 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 9.95 sec  INFO : MapReduce Total cumulative CPU time: 9 seconds 950 msec  INFO : Ended Job = job\_1490094830773\_0346  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 8 Reduce: 1 Cumulative CPU: 9.95 sec HDFS Read: 5332032 HDFS Write: 19 SUCCESS  INFO : Total MapReduce CPU Time Spent: 9 seconds 950 msec  INFO : Completed executing command(queryId=hive\_20170411051313\_d43d9c91-fa53-48f0-9a2e-687fffe9c018); Time taken: 21.181 seconds  INFO : OK |

## User wise Max Date

|  |
| --- |
| select username,max(date) from postimport\_orc group by username;  INFO : Compiling command(queryId=hive\_20170411051414\_ac8a61c5-c09b-418d-80e5-b2841e7a3627): select username,max(date) from postimport\_orc group by username  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:username, type:string, comment:null), FieldSchema(name:\_c1, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411051414\_ac8a61c5-c09b-418d-80e5-b2841e7a3627); Time taken: 0.133 seconds  INFO : Executing command(queryId=hive\_20170411051414\_ac8a61c5-c09b-418d-80e5-b2841e7a3627): select username,max(date) from postimport\_orc group by username  INFO : Query ID = hive\_20170411051414\_ac8a61c5-c09b-418d-80e5-b2841e7a3627  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks not specified. Estimated from input data size: 30  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:8  INFO : Submitting tokens for job: job\_1490094830773\_0347  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0347/  INFO : Starting Job = job\_1490094830773\_0347, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0347/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0347  INFO : Hadoop job information for Stage-1: number of mappers: 8; number of reducers: 30  INFO : 2017-04-11 05:14:25,650 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 05:14:31,848 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 15.01 sec  INFO : 2017-04-11 05:14:32,871 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 19.08 sec  INFO : 2017-04-11 05:14:40,202 Stage-1 map = 100%, reduce = 20%, Cumulative CPU 26.64 sec  INFO : 2017-04-11 05:14:41,229 Stage-1 map = 100%, reduce = 50%, Cumulative CPU 38.26 sec  INFO : 2017-04-11 05:14:45,406 Stage-1 map = 100%, reduce = 60%, Cumulative CPU 42.6 sec  INFO : 2017-04-11 05:14:46,440 Stage-1 map = 100%, reduce = 80%, Cumulative CPU 49.77 sec  INFO : 2017-04-11 05:14:47,468 Stage-1 map = 100%, reduce = 87%, Cumulative CPU 52.29 sec  INFO : 2017-04-11 05:14:48,488 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 57.42 sec  INFO : MapReduce Total cumulative CPU time: 57 seconds 420 msec  INFO : Ended Job = job\_1490094830773\_0347  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 8 Reduce: 30 Cumulative CPU: 57.42 sec HDFS Read: 50484744 HDFS Write: 32473 SUCCESS  INFO : Total MapReduce CPU Time Spent: 57 seconds 420 msec  INFO : Completed executing command(queryId=hive\_20170411051414\_ac8a61c5-c09b-418d-80e5-b2841e7a3627); Time taken: 29.488 seconds  INFO : OK |

## Select All Columns

|  |
| --- |
| select \* from postimport\_orc  INFO : Compiling command(queryId=hive\_20170411051515\_19976cc0-d30f-42f7-ba5f-afa5e12182c3): select \* from postimport\_orc  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_orc.username, type:string, comment:null), FieldSchema(name:postimport\_orc.machinename, type:string, comment:null), FieldSchema(name:postimport\_orc.machineid, type:string, comment:null), FieldSchema(name:postimport\_orc.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_orc.domainname, type:string, comment:null), FieldSchema(name:postimport\_orc.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_orc.modulename, type:string, comment:null), FieldSchema(name:postimport\_orc.filename, type:string, comment:null), FieldSchema(name:postimport\_orc.productname, type:string, comment:null), FieldSchema(name:postimport\_orc.companyname, type:string, comment:null), FieldSchema(name:postimport\_orc.description, type:string, comment:null), FieldSchema(name:postimport\_orc.text, type:string, comment:null), FieldSchema(name:postimport\_orc.activetime, type:double, comment:null), FieldSchema(name:postimport\_orc.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_orc.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_orc.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411051515\_19976cc0-d30f-42f7-ba5f-afa5e12182c3); Time taken: 0.085 seconds  INFO : Executing command(queryId=hive\_20170411051515\_19976cc0-d30f-42f7-ba5f-afa5e12182c3): select \* from postimport\_orc  INFO : Query ID = hive\_20170411051515\_19976cc0-d30f-42f7-ba5f-afa5e12182c3  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:8  INFO : Submitting tokens for job: job\_1490094830773\_0348  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0348/  INFO : Starting Job = job\_1490094830773\_0348, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0348/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0348  INFO : Hadoop job information for Stage-1: number of mappers: 8; number of reducers: 0  INFO : 2017-04-11 05:15:14,058 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 05:15:30,456 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 94.49 sec  INFO : 2017-04-11 05:15:33,552 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 115.69 sec  INFO : 2017-04-11 05:15:35,605 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 121.71 sec  INFO : 2017-04-11 05:15:36,626 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 137.17 sec  INFO : 2017-04-11 05:15:39,698 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 159.66 sec  INFO : 2017-04-11 05:15:45,838 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 201.51 sec  INFO : 2017-04-11 05:15:47,883 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 207.4 sec  INFO : 2017-04-11 05:15:48,907 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 223.12 sec  INFO : 2017-04-11 05:15:55,041 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 263.55 sec  INFO : 2017-04-11 05:15:57,084 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 271.12 sec  INFO : 2017-04-11 05:15:58,107 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 280.07 sec  INFO : 2017-04-11 05:16:00,152 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 291.09 sec  INFO : 2017-04-11 05:16:01,172 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 297.1 sec  INFO : 2017-04-11 05:16:03,214 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 308.27 sec  INFO : 2017-04-11 05:16:07,298 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 331.51 sec  INFO : 2017-04-11 05:16:09,341 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 340.87 sec  INFO : 2017-04-11 05:16:11,385 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 353.2 sec  INFO : 2017-04-11 05:16:13,427 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 368.34 sec  INFO : 2017-04-11 05:16:18,534 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 403.25 sec  INFO : 2017-04-11 05:16:20,577 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 410.11 sec  INFO : 2017-04-11 05:16:21,601 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 423.86 sec  INFO : 2017-04-11 05:16:24,665 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 444.08 sec  INFO : 2017-04-11 05:16:29,770 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 463.26 sec  INFO : 2017-04-11 05:16:30,792 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 474.58 sec  INFO : 2017-04-11 05:16:33,857 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 489.99 sec  INFO : 2017-04-11 05:16:36,923 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 506.22 sec  INFO : 2017-04-11 05:16:38,964 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 511.39 sec  INFO : 2017-04-11 05:16:39,985 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 522.79 sec  INFO : 2017-04-11 05:16:43,050 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 539.25 sec  INFO : 2017-04-11 05:16:45,093 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 544.94 sec  INFO : 2017-04-11 05:16:49,177 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 577.08 sec  INFO : 2017-04-11 05:16:51,221 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 582.87 sec  INFO : 2017-04-11 05:16:52,243 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 594.84 sec  INFO : 2017-04-11 05:16:55,308 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 613.47 sec  INFO : 2017-04-11 05:16:57,350 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 619.53 sec  INFO : 2017-04-11 05:16:58,371 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 631.7 sec  INFO : 2017-04-11 05:17:03,477 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 661.3 sec  INFO : 2017-04-11 05:17:04,497 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 668.11 sec  INFO : 2017-04-11 05:17:06,540 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 679.16 sec  INFO : 2017-04-11 05:17:09,605 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 695.97 sec  INFO : 2017-04-11 05:17:10,625 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 703.79 sec  INFO : 2017-04-11 05:17:12,665 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 714.04 sec  INFO : 2017-04-11 05:17:18,794 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 746.9 sec  INFO : 2017-04-11 05:17:24,923 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 784.8 sec  INFO : 2017-04-11 05:17:27,988 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 803.87 sec  INFO : 2017-04-11 05:17:29,009 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 804.72 sec  INFO : 2017-04-11 05:17:31,054 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 819.3 sec  INFO : 2017-04-11 05:17:36,154 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 835.91 sec  INFO : 2017-04-11 05:17:37,173 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 844.03 sec  INFO : 2017-04-11 05:17:39,210 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 848.1 sec  INFO : 2017-04-11 05:17:46,339 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 856.79 sec  INFO : 2017-04-11 05:17:55,507 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 866.34 sec  INFO : MapReduce Total cumulative CPU time: 14 minutes 26 seconds 340 msec  INFO : Ended Job = job\_1490094830773\_0348  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 8 Cumulative CPU: 866.34 sec HDFS Read: 2002752522 HDFS Write: 3633968737 SUCCESS  INFO : Total MapReduce CPU Time Spent: 14 minutes 26 seconds 340 msec  INFO : Completed executing command(queryId=hive\_20170411051515\_19976cc0-d30f-42f7-ba5f-afa5e12182c3); Time taken: 167.884 seconds  INFO : OK |

## Select All Columns with filter Condition

|  |
| --- |
| select \* from postimport\_orc where username = '670063006100750074006F007500730065007200';  INFO : Compiling command(queryId=hive\_20170411051818\_455c43c8-df12-4ccf-8a20-c2c2ccffd9fd): select \* from postimport\_orc where username = '670063006100750074006F007500730065007200'  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_orc.username, type:string, comment:null), FieldSchema(name:postimport\_orc.machinename, type:string, comment:null), FieldSchema(name:postimport\_orc.machineid, type:string, comment:null), FieldSchema(name:postimport\_orc.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_orc.domainname, type:string, comment:null), FieldSchema(name:postimport\_orc.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_orc.modulename, type:string, comment:null), FieldSchema(name:postimport\_orc.filename, type:string, comment:null), FieldSchema(name:postimport\_orc.productname, type:string, comment:null), FieldSchema(name:postimport\_orc.companyname, type:string, comment:null), FieldSchema(name:postimport\_orc.description, type:string, comment:null), FieldSchema(name:postimport\_orc.text, type:string, comment:null), FieldSchema(name:postimport\_orc.activetime, type:double, comment:null), FieldSchema(name:postimport\_orc.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_orc.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_orc.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411051818\_455c43c8-df12-4ccf-8a20-c2c2ccffd9fd); Time taken: 0.096 seconds  INFO : Executing command(queryId=hive\_20170411051818\_455c43c8-df12-4ccf-8a20-c2c2ccffd9fd): select \* from postimport\_orc where username = '670063006100750074006F007500730065007200'  INFO : Query ID = hive\_20170411051818\_455c43c8-df12-4ccf-8a20-c2c2ccffd9fd  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:8  INFO : Submitting tokens for job: job\_1490094830773\_0349  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0349/  INFO : Starting Job = job\_1490094830773\_0349, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0349/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0349  INFO : Hadoop job information for Stage-1: number of mappers: 8; number of reducers: 0  INFO : 2017-04-11 05:18:21,599 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 05:18:28,858 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 3.75 sec  INFO : 2017-04-11 05:18:29,882 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 9.66 sec  INFO : 2017-04-11 05:18:31,940 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 33.6 sec  INFO : 2017-04-11 05:18:32,961 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 42.79 sec  INFO : 2017-04-11 05:18:33,985 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 49.49 sec  INFO : MapReduce Total cumulative CPU time: 49 seconds 490 msec  INFO : Ended Job = job\_1490094830773\_0349  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 8 Cumulative CPU: 49.49 sec HDFS Read: 2004368826 HDFS Write: 2074373 SUCCESS  INFO : Total MapReduce CPU Time Spent: 49 seconds 490 msec  INFO : Completed executing command(queryId=hive\_20170411051818\_455c43c8-df12-4ccf-8a20-c2c2ccffd9fd); Time taken: 20.417 seconds  INFO : OK |

# ORC File format NONE Compression

|  |
| --- |
| create table postimport\_orc like postimport\_text stored as ORC tblproperties ("orc.compress"="NONE"); |

|  |
| --- |
| SET hive.exec.compress.output=true; |

## Load data to ORC Table

|  |
| --- |
| hive> insert overwrite table postimport\_orc select \* from postimport\_text;  Query ID = shalaj\_20170411045353\_5a871a27-7a51-4c4f-80f6-c6402d9bb647  Total jobs = 1  Launching Job 1 out of 1  Number of reduce tasks is set to 0 since there's no reduce operator  Starting Job = job\_1490094830773\_0338, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0338/  Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0338  Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  2017-04-11 04:53:33,145 Stage-1 map = 0%, reduce = 0%  2017-04-11 04:53:47,966 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 70.14 sec  2017-04-11 04:53:48,988 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 76.29 sec  2017-04-11 04:53:57,625 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 120.23 sec  2017-04-11 04:53:58,689 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 140.99 sec  2017-04-11 04:54:00,771 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 145.08 sec  2017-04-11 04:54:01,851 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 149.06 sec  2017-04-11 04:54:06,714 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 166.68 sec  2017-04-11 04:54:10,569 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 175.51 sec  2017-04-11 04:54:12,914 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 181.43 sec  2017-04-11 04:54:20,564 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 198.73 sec  2017-04-11 04:54:24,106 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 200.2 sec  2017-04-11 04:54:33,630 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 212.24 sec  2017-04-11 04:54:35,987 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 215.82 sec  2017-04-11 04:54:37,167 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 220.49 sec  2017-04-11 04:54:43,132 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 222.04 sec  2017-04-11 04:54:44,213 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 229.92 sec  2017-04-11 04:54:45,283 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 231.48 sec  2017-04-11 04:54:48,827 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 242.19 sec  2017-04-11 04:54:49,902 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 252.76 sec  2017-04-11 04:54:52,043 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 264.25 sec  2017-04-11 04:54:54,167 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 273.71 sec  2017-04-11 04:54:56,213 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 288.46 sec  2017-04-11 04:54:58,257 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 308.84 sec  2017-04-11 04:54:59,281 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 316.19 sec  2017-04-11 04:55:01,366 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 328.69 sec  2017-04-11 04:55:04,622 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 351.17 sec  2017-04-11 04:55:08,119 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 387.99 sec  2017-04-11 04:55:09,341 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 391.6 sec  2017-04-11 04:55:10,604 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 396.92 sec  2017-04-11 04:55:14,082 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 406.05 sec  2017-04-11 04:55:19,952 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 415.85 sec  2017-04-11 04:55:22,230 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 421.96 sec  2017-04-11 04:55:23,411 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 439.98 sec  2017-04-11 04:55:25,708 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 447.07 sec  2017-04-11 04:55:26,804 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 455.83 sec  2017-04-11 04:55:33,140 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 458.91 sec  2017-04-11 04:55:34,235 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 473.21 sec  2017-04-11 04:55:36,369 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 483.44 sec  2017-04-11 04:55:40,895 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 509.05 sec  2017-04-11 04:55:44,546 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 523.54 sec  2017-04-11 04:55:46,876 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 531.13 sec  2017-04-11 04:55:54,854 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 551.65 sec  2017-04-11 04:55:56,046 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 552.93 sec  2017-04-11 04:55:57,201 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 569.86 sec  2017-04-11 04:55:58,595 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 572.09 sec  2017-04-11 04:55:59,815 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 576.14 sec  2017-04-11 04:56:01,031 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 581.82 sec  2017-04-11 04:56:04,546 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 594.53 sec  2017-04-11 04:56:06,867 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 596.36 sec  2017-04-11 04:56:12,655 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 614.28 sec  2017-04-11 04:56:13,815 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 615.83 sec  2017-04-11 04:56:16,124 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 626.64 sec  2017-04-11 04:56:18,462 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 640.4 sec  2017-04-11 04:56:20,756 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 661.57 sec  2017-04-11 04:56:21,858 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 669.28 sec  2017-04-11 04:56:22,933 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 671.74 sec  2017-04-11 04:56:25,046 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 679.81 sec  2017-04-11 04:56:26,105 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 684.14 sec  2017-04-11 04:56:27,169 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 696.1 sec  2017-04-11 04:56:29,321 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 702.2 sec  2017-04-11 04:56:36,074 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 726.33 sec  2017-04-11 04:56:38,538 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 760.95 sec  2017-04-11 04:56:39,762 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 771.08 sec  2017-04-11 04:56:43,299 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 784.2 sec  2017-04-11 04:56:44,462 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 785.97 sec  2017-04-11 04:56:45,592 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 790.69 sec  2017-04-11 04:56:54,200 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 809.84 sec  2017-04-11 04:56:56,604 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 819.31 sec  2017-04-11 04:57:01,427 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 825.41 sec  2017-04-11 04:57:03,773 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 827.54 sec  2017-04-11 04:57:04,944 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 830.06 sec  2017-04-11 04:57:09,592 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 843.89 sec  2017-04-11 04:57:15,622 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 853.14 sec  2017-04-11 04:57:16,798 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 854.69 sec  2017-04-11 04:57:22,527 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 858.1 sec  2017-04-11 04:57:23,682 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 859.72 sec  2017-04-11 04:57:29,397 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 862.87 sec  2017-04-11 04:57:39,106 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 867.22 sec  2017-04-11 04:57:44,581 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 868.51 sec  MapReduce Total cumulative CPU time: 14 minutes 28 seconds 510 msec  Ended Job = job\_1490094830773\_0338  Stage-4 is selected by condition resolver.  Stage-3 is filtered out by condition resolver.  Stage-5 is filtered out by condition resolver.  Moving data to: hdfs://mac55:8020/user/hive/warehouse/esplus.db/postimport\_orc/.hive-staging\_hive\_2017-04-11\_04-53-25\_133\_321565157132704632-1/-ext-10000  Loading data to table esplus.postimport\_orc  Table esplus.postimport\_orc stats: [numFiles=74, numRows=42286538, totalSize=4088480566, rawDataSize=64936217951]  MapReduce Jobs Launched:  Stage-Stage-1: Map: 74 Cumulative CPU: 868.51 sec HDFS Read: 19486328310 HDFS Write: 4088487078 SUCCESS  Total MapReduce CPU Time Spent: 14 minutes 28 seconds 510 msec  OK  Time taken: 264.138 seconds |

## Data Size

|  |
| --- |
| hive> dfs -du -s -h /user/hive/warehouse/esplus.db/postimport\_orc;  3.8 G 11.4 G /user/hive/warehouse/esplus.db/postimport\_orc\_none |

## Select some columns

SET hive.exec.compress.output=true;

|  |
| --- |
| select productname,username from postimport\_orc;  INFO : Compiling command(queryId=hive\_20170411045959\_64175dbf-9a38-4184-b969-a6e63d964fd6): select productname,username from postimport\_orc  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:productname, type:string, comment:null), FieldSchema(name:username, type:string, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411045959\_64175dbf-9a38-4184-b969-a6e63d964fd6); Time taken: 0.093 seconds  INFO : Executing command(queryId=hive\_20170411045959\_64175dbf-9a38-4184-b969-a6e63d964fd6): select productname,username from postimport\_orc  INFO : Query ID = hive\_20170411045959\_64175dbf-9a38-4184-b969-a6e63d964fd6  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:15  INFO : Submitting tokens for job: job\_1490094830773\_0339  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0339/  INFO : Starting Job = job\_1490094830773\_0339, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0339/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0339  INFO : Hadoop job information for Stage-1: number of mappers: 15; number of reducers: 0  INFO : 2017-04-11 04:59:57,717 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 05:00:09,183 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 16.62 sec  INFO : 2017-04-11 05:00:10,219 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 58.74 sec  INFO : 2017-04-11 05:00:11,246 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 71.63 sec  INFO : 2017-04-11 05:00:12,273 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 87.25 sec  INFO : 2017-04-11 05:00:13,296 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 103.79 sec  INFO : 2017-04-11 05:00:14,320 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 111.96 sec  INFO : 2017-04-11 05:00:15,341 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 116.64 sec  INFO : 2017-04-11 05:00:16,364 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 118.08 sec  INFO : MapReduce Total cumulative CPU time: 1 minutes 58 seconds 80 msec  INFO : Ended Job = job\_1490094830773\_0339  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 15 Cumulative CPU: 118.08 sec HDFS Read: 180749699 HDFS Write: 233697266 SUCCESS  INFO : Total MapReduce CPU Time Spent: 1 minutes 58 seconds 80 msec  INFO : Completed executing command(queryId=hive\_20170411045959\_64175dbf-9a38-4184-b969-a6e63d964fd6); Time taken: 25.009 seconds  INFO : OK |

## Check Max Date

|  |
| --- |
| select max(date) from postimport\_orc;  INFO : Compiling command(queryId=hive\_20170411050101\_41ba4455-b62f-4a2d-8243-34e59364bf60): select max(date) from postimport\_orc  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:\_c0, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411050101\_41ba4455-b62f-4a2d-8243-34e59364bf60); Time taken: 0.132 seconds  INFO : Executing command(queryId=hive\_20170411050101\_41ba4455-b62f-4a2d-8243-34e59364bf60): select max(date) from postimport\_orc  INFO : Query ID = hive\_20170411050101\_41ba4455-b62f-4a2d-8243-34e59364bf60  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks determined at compile time: 1  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:15  INFO : Submitting tokens for job: job\_1490094830773\_0340  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0340/  INFO : Starting Job = job\_1490094830773\_0340, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0340/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0340  INFO : Hadoop job information for Stage-1: number of mappers: 15; number of reducers: 1  INFO : 2017-04-11 05:01:41,646 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 05:01:47,983 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 4.24 sec  INFO : 2017-04-11 05:01:49,202 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 7.47 sec  INFO : 2017-04-11 05:01:50,229 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 15.53 sec  INFO : 2017-04-11 05:01:51,250 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 16.63 sec  INFO : 2017-04-11 05:01:55,383 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 17.6 sec  INFO : MapReduce Total cumulative CPU time: 17 seconds 600 msec  INFO : Ended Job = job\_1490094830773\_0340  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 15 Reduce: 1 Cumulative CPU: 17.6 sec HDFS Read: 86052601 HDFS Write: 19 SUCCESS  INFO : Total MapReduce CPU Time Spent: 17 seconds 600 msec  INFO : Completed executing command(queryId=hive\_20170411050101\_41ba4455-b62f-4a2d-8243-34e59364bf60); Time taken: 22.393 seconds  INFO : OK |

## User wise Max Date

|  |
| --- |
| select username,max(date) from postimport\_orc group by username;  INFO : Compiling command(queryId=hive\_20170411050202\_d41b3807-0243-4fc6-b5aa-981d84391b63): select username,max(date) from postimport\_orc group by username  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:username, type:string, comment:null), FieldSchema(name:\_c1, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411050202\_d41b3807-0243-4fc6-b5aa-981d84391b63); Time taken: 0.148 seconds  INFO : Executing command(queryId=hive\_20170411050202\_d41b3807-0243-4fc6-b5aa-981d84391b63): select username,max(date) from postimport\_orc group by username  INFO : Query ID = hive\_20170411050202\_d41b3807-0243-4fc6-b5aa-981d84391b63  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks not specified. Estimated from input data size: 61  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:15  INFO : Submitting tokens for job: job\_1490094830773\_0341  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0341/  INFO : Starting Job = job\_1490094830773\_0341, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0341/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0341  INFO : Hadoop job information for Stage-1: number of mappers: 15; number of reducers: 61  INFO : 2017-04-11 05:02:36,120 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 05:02:43,455 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 6.11 sec  INFO : 2017-04-11 05:02:44,496 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 14.29 sec  INFO : 2017-04-11 05:02:45,523 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 22.72 sec  INFO : 2017-04-11 05:02:47,600 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 28.81 sec  INFO : 2017-04-11 05:02:48,648 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 30.86 sec  INFO : 2017-04-11 05:02:54,945 Stage-1 map = 100%, reduce = 11%, Cumulative CPU 39.86 sec  INFO : 2017-04-11 05:02:55,972 Stage-1 map = 100%, reduce = 25%, Cumulative CPU 50.2 sec  INFO : 2017-04-11 05:02:59,100 Stage-1 map = 100%, reduce = 26%, Cumulative CPU 51.61 sec  INFO : 2017-04-11 05:03:00,143 Stage-1 map = 100%, reduce = 30%, Cumulative CPU 54.38 sec  INFO : 2017-04-11 05:03:01,185 Stage-1 map = 100%, reduce = 39%, Cumulative CPU 61.57 sec  INFO : 2017-04-11 05:03:02,230 Stage-1 map = 100%, reduce = 44%, Cumulative CPU 65.54 sec  INFO : 2017-04-11 05:03:03,266 Stage-1 map = 100%, reduce = 49%, Cumulative CPU 69.66 sec  INFO : 2017-04-11 05:03:04,345 Stage-1 map = 100%, reduce = 52%, Cumulative CPU 72.51 sec  INFO : 2017-04-11 05:03:05,388 Stage-1 map = 100%, reduce = 54%, Cumulative CPU 74.04 sec  INFO : 2017-04-11 05:03:07,476 Stage-1 map = 100%, reduce = 66%, Cumulative CPU 82.48 sec  INFO : 2017-04-11 05:03:08,516 Stage-1 map = 100%, reduce = 70%, Cumulative CPU 86.5 sec  INFO : 2017-04-11 05:03:09,543 Stage-1 map = 100%, reduce = 77%, Cumulative CPU 92.06 sec  INFO : 2017-04-11 05:03:10,580 Stage-1 map = 100%, reduce = 79%, Cumulative CPU 93.28 sec  INFO : 2017-04-11 05:03:12,644 Stage-1 map = 100%, reduce = 84%, Cumulative CPU 96.71 sec  INFO : 2017-04-11 05:03:13,669 Stage-1 map = 100%, reduce = 98%, Cumulative CPU 107.89 sec  INFO : 2017-04-11 05:03:14,695 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 109.13 sec  INFO : MapReduce Total cumulative CPU time: 1 minutes 49 seconds 130 msec  INFO : Ended Job = job\_1490094830773\_0341  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 15 Reduce: 61 Cumulative CPU: 109.13 sec HDFS Read: 184325659 HDFS Write: 36370 SUCCESS  INFO : Total MapReduce CPU Time Spent: 1 minutes 49 seconds 130 msec  INFO : Completed executing command(queryId=hive\_20170411050202\_d41b3807-0243-4fc6-b5aa-981d84391b63); Time taken: 45.292 seconds  INFO : OK |

## Select All Columns

|  |
| --- |
| select \* from postimport\_orc  INFO : Compiling command(queryId=hive\_20170411050303\_a83d5798-190a-4c34-9747-ec90a5ed490a): select \* from postimport\_orc  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_orc.username, type:string, comment:null), FieldSchema(name:postimport\_orc.machinename, type:string, comment:null), FieldSchema(name:postimport\_orc.machineid, type:string, comment:null), FieldSchema(name:postimport\_orc.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_orc.domainname, type:string, comment:null), FieldSchema(name:postimport\_orc.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_orc.modulename, type:string, comment:null), FieldSchema(name:postimport\_orc.filename, type:string, comment:null), FieldSchema(name:postimport\_orc.productname, type:string, comment:null), FieldSchema(name:postimport\_orc.companyname, type:string, comment:null), FieldSchema(name:postimport\_orc.description, type:string, comment:null), FieldSchema(name:postimport\_orc.text, type:string, comment:null), FieldSchema(name:postimport\_orc.activetime, type:double, comment:null), FieldSchema(name:postimport\_orc.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_orc.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_orc.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411050303\_a83d5798-190a-4c34-9747-ec90a5ed490a); Time taken: 0.092 seconds  INFO : Executing command(queryId=hive\_20170411050303\_a83d5798-190a-4c34-9747-ec90a5ed490a): select \* from postimport\_orc  INFO : Query ID = hive\_20170411050303\_a83d5798-190a-4c34-9747-ec90a5ed490a  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:15  INFO : Submitting tokens for job: job\_1490094830773\_0342  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0342/  INFO : Starting Job = job\_1490094830773\_0342, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0342/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0342  INFO : Hadoop job information for Stage-1: number of mappers: 15; number of reducers: 0  INFO : 2017-04-11 05:03:35,141 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 05:03:57,988 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 168.56 sec  INFO : 2017-04-11 05:04:00,056 Stage-1 map = 6%, reduce = 0%, Cumulative CPU 187.52 sec  INFO : 2017-04-11 05:04:03,161 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 217.24 sec  INFO : 2017-04-11 05:04:06,270 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 248.14 sec  INFO : 2017-04-11 05:04:07,299 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 260.22 sec  INFO : 2017-04-11 05:04:08,328 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 267.89 sec  INFO : 2017-04-11 05:04:10,413 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 290.48 sec  INFO : 2017-04-11 05:04:11,449 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 299.42 sec  INFO : 2017-04-11 05:04:12,488 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 311.99 sec  INFO : 2017-04-11 05:04:13,518 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 323.35 sec  INFO : 2017-04-11 05:04:14,543 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 333.58 sec  INFO : 2017-04-11 05:04:15,587 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 342.01 sec  INFO : 2017-04-11 05:04:17,655 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 364.62 sec  INFO : 2017-04-11 05:04:20,772 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 399.59 sec  INFO : 2017-04-11 05:04:22,831 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 408.6 sec  INFO : 2017-04-11 05:04:23,853 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 422.47 sec  INFO : 2017-04-11 05:04:25,911 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 431.98 sec  INFO : 2017-04-11 05:04:26,934 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 443.75 sec  INFO : 2017-04-11 05:04:30,020 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 465.78 sec  INFO : 2017-04-11 05:04:33,110 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 487.3 sec  INFO : 2017-04-11 05:04:34,134 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 488.98 sec  INFO : 2017-04-11 05:04:35,158 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 495.29 sec  INFO : 2017-04-11 05:04:36,183 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 508.78 sec  INFO : 2017-04-11 05:04:38,227 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 519.26 sec  INFO : 2017-04-11 05:04:39,258 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 530.68 sec  INFO : 2017-04-11 05:04:43,367 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 559.02 sec  INFO : 2017-04-11 05:04:44,391 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 567.46 sec  INFO : 2017-04-11 05:04:48,488 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 589.41 sec  INFO : 2017-04-11 05:04:50,531 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 598.23 sec  INFO : 2017-04-11 05:04:55,693 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 618.29 sec  INFO : 2017-04-11 05:05:00,792 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 638.78 sec  INFO : 2017-04-11 05:05:03,852 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 651.03 sec  INFO : 2017-04-11 05:05:05,890 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 660.79 sec  INFO : 2017-04-11 05:05:08,950 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 673.75 sec  INFO : 2017-04-11 05:05:15,073 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 700.98 sec  INFO : 2017-04-11 05:05:17,112 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 704.88 sec  INFO : 2017-04-11 05:05:18,133 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 714.88 sec  INFO : 2017-04-11 05:05:21,189 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 728.58 sec  INFO : 2017-04-11 05:05:22,209 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 730.05 sec  INFO : 2017-04-11 05:05:27,313 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 753.87 sec  INFO : 2017-04-11 05:05:35,483 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 786.56 sec  INFO : 2017-04-11 05:05:38,542 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 804.09 sec  INFO : 2017-04-11 05:05:39,565 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 809.89 sec  INFO : 2017-04-11 05:05:42,624 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 824.93 sec  INFO : 2017-04-11 05:05:46,702 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 838.91 sec  INFO : 2017-04-11 05:05:47,721 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 841.97 sec  INFO : 2017-04-11 05:05:50,780 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 852.8 sec  INFO : 2017-04-11 05:05:52,823 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 860.43 sec  INFO : 2017-04-11 05:05:55,880 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 870.48 sec  INFO : 2017-04-11 05:05:57,919 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 875.57 sec  INFO : 2017-04-11 05:06:00,981 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 880.81 sec  INFO : 2017-04-11 05:06:06,074 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 894.82 sec  INFO : 2017-04-11 05:06:10,187 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 901.17 sec  INFO : 2017-04-11 05:06:15,352 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 907.11 sec  INFO : MapReduce Total cumulative CPU time: 15 minutes 7 seconds 110 msec  INFO : Ended Job = job\_1490094830773\_0342  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 15 Cumulative CPU: 907.11 sec HDFS Read: 4084515359 HDFS Write: 3633966325 SUCCESS  INFO : Total MapReduce CPU Time Spent: 15 minutes 7 seconds 110 msec  INFO : Completed executing command(queryId=hive\_20170411050303\_a83d5798-190a-4c34-9747-ec90a5ed490a); Time taken: 166.84 seconds  INFO : OK |

## Select All Columns with filter Condition

|  |
| --- |
| select \* from postimport\_orc where username = '670063006100750074006F007500730065007200';  INFO : Compiling command(queryId=hive\_20170411050606\_0b7f47b6-47cb-444e-9405-9ee4a092c2cf): select \* from postimport\_orc where username = '670063006100750074006F007500730065007200'  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_orc.username, type:string, comment:null), FieldSchema(name:postimport\_orc.machinename, type:string, comment:null), FieldSchema(name:postimport\_orc.machineid, type:string, comment:null), FieldSchema(name:postimport\_orc.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_orc.domainname, type:string, comment:null), FieldSchema(name:postimport\_orc.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_orc.modulename, type:string, comment:null), FieldSchema(name:postimport\_orc.filename, type:string, comment:null), FieldSchema(name:postimport\_orc.productname, type:string, comment:null), FieldSchema(name:postimport\_orc.companyname, type:string, comment:null), FieldSchema(name:postimport\_orc.description, type:string, comment:null), FieldSchema(name:postimport\_orc.text, type:string, comment:null), FieldSchema(name:postimport\_orc.activetime, type:double, comment:null), FieldSchema(name:postimport\_orc.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_orc.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_orc.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411050606\_0b7f47b6-47cb-444e-9405-9ee4a092c2cf); Time taken: 0.096 seconds  INFO : Executing command(queryId=hive\_20170411050606\_0b7f47b6-47cb-444e-9405-9ee4a092c2cf): select \* from postimport\_orc where username = '670063006100750074006F007500730065007200'  INFO : Query ID = hive\_20170411050606\_0b7f47b6-47cb-444e-9405-9ee4a092c2cf  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:15  INFO : Submitting tokens for job: job\_1490094830773\_0343  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0343/  INFO : Starting Job = job\_1490094830773\_0343, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0343/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0343  INFO : Hadoop job information for Stage-1: number of mappers: 15; number of reducers: 0  INFO : 2017-04-11 05:07:07,302 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 05:07:20,305 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 12.04 sec  INFO : 2017-04-11 05:07:21,681 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 25.32 sec  INFO : 2017-04-11 05:07:22,730 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 38.33 sec  INFO : 2017-04-11 05:07:23,758 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 47.07 sec  INFO : 2017-04-11 05:07:24,780 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 47.75 sec  INFO : 2017-04-11 05:07:25,806 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 56.6 sec  INFO : 2017-04-11 05:07:26,827 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 60.14 sec  INFO : 2017-04-11 05:07:27,859 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 61.66 sec  INFO : 2017-04-11 05:07:28,881 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 65.1 sec  INFO : 2017-04-11 05:07:29,909 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 65.49 sec  INFO : 2017-04-11 05:07:33,004 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 66.72 sec  INFO : 2017-04-11 05:07:36,075 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 68.8 sec  INFO : 2017-04-11 05:07:37,126 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 69.71 sec  INFO : 2017-04-11 05:07:38,148 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 71.75 sec  INFO : MapReduce Total cumulative CPU time: 1 minutes 11 seconds 750 msec  INFO : Ended Job = job\_1490094830773\_0343  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 15 Cumulative CPU: 71.75 sec HDFS Read: 4089530712 HDFS Write: 2079044 SUCCESS  INFO : Total MapReduce CPU Time Spent: 1 minutes 11 seconds 750 msec  INFO : Completed executing command(queryId=hive\_20170411050606\_0b7f47b6-47cb-444e-9405-9ee4a092c2cf); Time taken: 41.737 seconds  INFO : OK |

# Parquet File format Default Compression

|  |
| --- |
| CREATE TABLE postimport\_parquet(  username string,  machinename string,  machineid string,  ipaddress string,  domainname string,  pcmlogeventid string,  modulename string,  filename string,  productname string,  companyname string,  description string,  text string,  activetime double,  datetime timestamp,  timeslot timestamp)  STORED AS PARQUET; |

|  |
| --- |
| SET hive.exec.compress.output=true; |

***Note: Parquet does not support date data type in hive 1.1 version.***

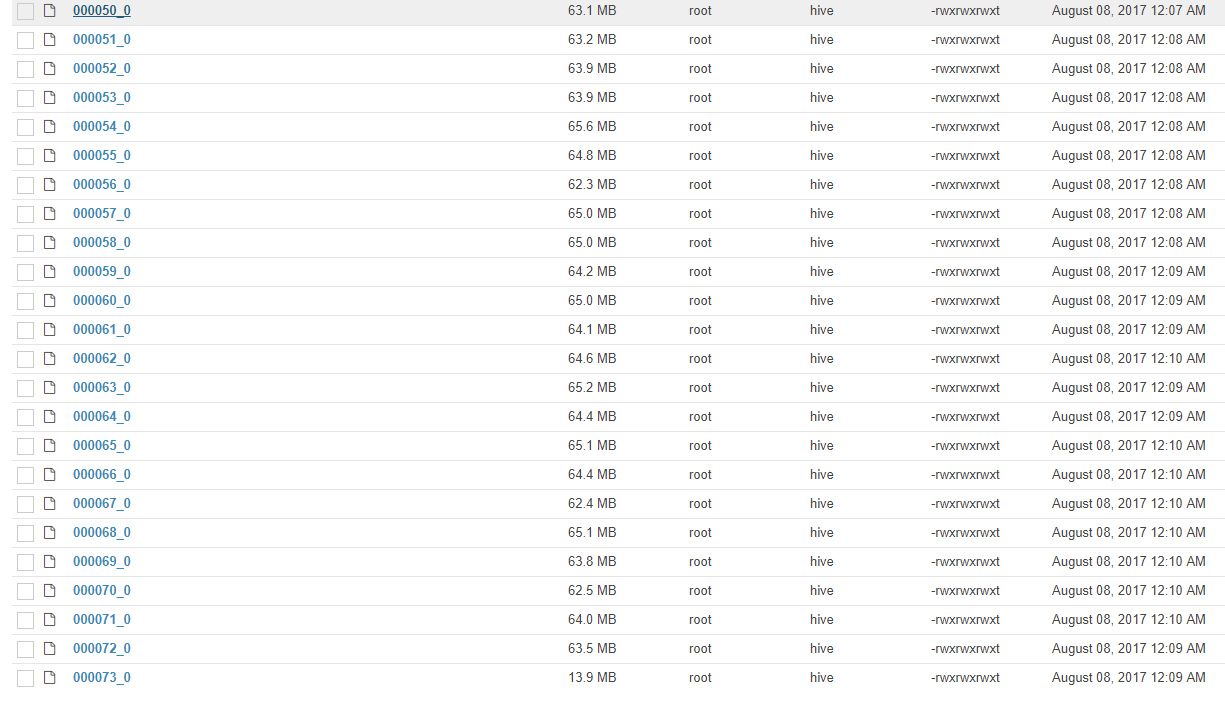
## Load data to Parquet Table

|  |
| --- |
| insert into postimport\_parquet select username,machinename,machineid,ipaddress,domainname,pcmlogeventid,modulename,filename,productname,companyname,description,text,activetime,datetime,timeslot from postimport\_text; |

|  |
| --- |
| hive> insert into postimport\_parquet select username,machinename,machineid,ipaddress,domainname,pcmlogeventid,modulename,filename,productname,companyname,description,text,activetime,datetime,timeslot from postimport\_text;  Query ID = shalaj\_20170411052020\_8b46c979-c7bf-4206-8d6b-1a050ea5ac6d  Total jobs = 3  Launching Job 1 out of 3  Number of reduce tasks is set to 0 since there's no reduce operator  Starting Job = job\_1490094830773\_0350, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0350/  Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0350  Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  2017-04-11 05:20:48,306 Stage-1 map = 0%, reduce = 0%  2017-04-11 05:21:03,016 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 79.04 sec  2017-04-11 05:21:04,051 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 84.12 sec  2017-04-11 05:21:06,094 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 101.42 sec  2017-04-11 05:21:08,135 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 107.69 sec  2017-04-11 05:21:09,203 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 117.0 sec  2017-04-11 05:21:10,277 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 119.71 sec  2017-04-11 05:21:16,872 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 140.45 sec  2017-04-11 05:21:18,123 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 142.76 sec  2017-04-11 05:21:22,045 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 156.53 sec  2017-04-11 05:21:23,250 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 169.5 sec  2017-04-11 05:21:28,010 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 176.14 sec  2017-04-11 05:21:36,391 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 191.87 sec  2017-04-11 05:21:37,572 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 198.48 sec  2017-04-11 05:21:39,951 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 201.58 sec  2017-04-11 05:21:42,371 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 205.48 sec  2017-04-11 05:22:13,191 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 245.45 sec  2017-04-11 05:22:14,250 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 251.9 sec  2017-04-11 05:22:16,394 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 255.62 sec  2017-04-11 05:22:23,298 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 282.62 sec  2017-04-11 05:22:24,466 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 288.73 sec  2017-04-11 05:22:25,720 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 296.36 sec  2017-04-11 05:22:26,942 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 311.66 sec  2017-04-11 05:22:29,301 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 329.14 sec  2017-04-11 05:22:30,464 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 336.06 sec  2017-04-11 05:22:31,623 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 338.84 sec  2017-04-11 05:22:32,796 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 341.74 sec  2017-04-11 05:22:33,967 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 345.43 sec  2017-04-11 05:22:44,562 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 381.71 sec  2017-04-11 05:22:45,730 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 401.63 sec  2017-04-11 05:22:46,904 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 404.98 sec  2017-04-11 05:22:49,186 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 427.18 sec  2017-04-11 05:22:51,489 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 447.74 sec  2017-04-11 05:22:52,778 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 449.2 sec  2017-04-11 05:22:53,952 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 451.46 sec  2017-04-11 05:22:58,645 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 457.87 sec  2017-04-11 05:22:59,739 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 466.05 sec  2017-04-11 05:23:02,190 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 479.2 sec  2017-04-11 05:23:03,263 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 485.28 sec  2017-04-11 05:23:05,544 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 505.74 sec  2017-04-11 05:23:09,094 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 519.63 sec  2017-04-11 05:23:11,340 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 522.88 sec  2017-04-11 05:23:15,737 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 547.42 sec  2017-04-11 05:23:16,837 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 551.05 sec  2017-04-11 05:23:20,455 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 558.82 sec  2017-04-11 05:23:21,671 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 562.68 sec  2017-04-11 05:23:24,206 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 563.38 sec  2017-04-11 05:23:31,442 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 583.73 sec  2017-04-11 05:23:34,972 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 596.52 sec  2017-04-11 05:23:37,582 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 599.5 sec  2017-04-11 05:23:41,141 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 603.07 sec  2017-04-11 05:23:44,697 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 610.88 sec  2017-04-11 05:23:48,481 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 620.82 sec  2017-04-11 05:23:49,664 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 624.78 sec  2017-04-11 05:23:55,686 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 639.89 sec  2017-04-11 05:23:58,316 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 647.77 sec  2017-04-11 05:24:00,573 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 665.02 sec  2017-04-11 05:24:01,668 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 671.3 sec  2017-04-11 05:24:02,750 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 679.19 sec  2017-04-11 05:24:04,923 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 682.08 sec  2017-04-11 05:24:06,234 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 688.41 sec  2017-04-11 05:24:08,405 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 694.46 sec  2017-04-11 05:24:10,848 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 720.14 sec  2017-04-11 05:24:12,149 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 722.6 sec  2017-04-11 05:24:15,759 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 736.48 sec  2017-04-11 05:24:19,530 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 758.43 sec  2017-04-11 05:24:22,390 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 770.65 sec  2017-04-11 05:24:24,882 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 774.24 sec  2017-04-11 05:24:26,082 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 782.34 sec  2017-04-11 05:24:28,569 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 790.29 sec  2017-04-11 05:25:00,021 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 806.25 sec  2017-04-11 05:25:02,346 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 810.31 sec  2017-04-11 05:25:17,727 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 822.84 sec  MapReduce Total cumulative CPU time: 13 minutes 42 seconds 840 msec  Ended Job = job\_1490094830773\_0350  Stage-4 is selected by condition resolver.  Stage-3 is filtered out by condition resolver.  Stage-5 is filtered out by condition resolver.  Moving data to: hdfs://mac55:8020/user/hive/warehouse/esplus.db/postimport\_parquet/.hive-staging\_hive\_2017-04-11\_05-20-39\_120\_2153481845016341438-1/-ext-10000  Loading data to table esplus.postimport\_parquet  Table esplus.postimport\_parquet stats: [numFiles=74, numRows=42286538, totalSize=4916157032, rawDataSize=634298070]  MapReduce Jobs Launched:  Stage-Stage-1: Map: 74 Cumulative CPU: 822.84 sec HDFS Read: 19486336080 HDFS Write: 4916163692 SUCCESS  Total MapReduce CPU Time Spent: 13 minutes 42 seconds 840 msec  OK  Time taken: 289.271 seconds |

## Data Size

|  |
| --- |
| hive> dfs -du -s -h /user/hive/warehouse/esplus.db/postimport\_parquet;  4.6 G 13.7 G /user/hive/warehouse/esplus.db/postimport\_parquet |



It creates more file than ORC

## Select some columns

SET hive.exec.compress.output=true;

|  |
| --- |
| select productname,username from postimport\_parquet;  INFO : Compiling command(queryId=hive\_20170411052727\_133452e9-ec8d-409e-9bdf-3c76b8405d83): select productname,username from postimport\_parquet  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:productname, type:string, comment:null), FieldSchema(name:username, type:string, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411052727\_133452e9-ec8d-409e-9bdf-3c76b8405d83); Time taken: 0.144 seconds  INFO : Executing command(queryId=hive\_20170411052727\_133452e9-ec8d-409e-9bdf-3c76b8405d83): select productname,username from postimport\_parquet  INFO : Query ID = hive\_20170411052727\_133452e9-ec8d-409e-9bdf-3c76b8405d83  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:19  INFO : Submitting tokens for job: job\_1490094830773\_0351  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0351/  INFO : Starting Job = job\_1490094830773\_0351, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0351/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0351  INFO : Hadoop job information for Stage-1: number of mappers: 19; number of reducers: 0  INFO : 2017-04-11 05:27:21,942 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 05:27:33,400 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 36.9 sec  INFO : 2017-04-11 05:27:34,427 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 77.25 sec  INFO : 2017-04-11 05:27:36,490 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 101.65 sec  INFO : 2017-04-11 05:27:37,518 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 112.59 sec  INFO : 2017-04-11 05:27:38,543 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 113.57 sec  INFO : 2017-04-11 05:27:39,569 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 116.2 sec  INFO : 2017-04-11 05:27:40,601 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 117.53 sec  INFO : 2017-04-11 05:27:43,666 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 121.43 sec  INFO : 2017-04-11 05:27:44,687 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 127.76 sec  INFO : 2017-04-11 05:27:45,708 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 135.87 sec  INFO : 2017-04-11 05:27:47,752 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 142.3 sec  INFO : 2017-04-11 05:27:48,776 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 143.21 sec  INFO : MapReduce Total cumulative CPU time: 2 minutes 23 seconds 210 msec  INFO : Ended Job = job\_1490094830773\_0351  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 19 Cumulative CPU: 143.21 sec HDFS Read: 119950838 HDFS Write: 233694992 SUCCESS  INFO : Total MapReduce CPU Time Spent: 2 minutes 23 seconds 210 msec  INFO : Completed executing command(queryId=hive\_20170411052727\_133452e9-ec8d-409e-9bdf-3c76b8405d83); Time taken: 34.516 seconds  INFO : OK |

## Check Max Date

|  |
| --- |
| select max(datetime) from postimport\_parquet;  INFO : Compiling command(queryId=hive\_20170411052828\_20c967e0-664d-4919-ba51-3320d699ef66): select max(datetime) from postimport\_parquet  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:\_c0, type:timestamp, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411052828\_20c967e0-664d-4919-ba51-3320d699ef66); Time taken: 0.097 seconds  INFO : Executing command(queryId=hive\_20170411052828\_20c967e0-664d-4919-ba51-3320d699ef66): select max(datetime) from postimport\_parquet  INFO : Query ID = hive\_20170411052828\_20c967e0-664d-4919-ba51-3320d699ef66  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks determined at compile time: 1  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:19  INFO : Submitting tokens for job: job\_1490094830773\_0352  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0352/  INFO : Starting Job = job\_1490094830773\_0352, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0352/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0352  INFO : Hadoop job information for Stage-1: number of mappers: 19; number of reducers: 1  INFO : 2017-04-11 05:28:23,429 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 05:28:35,280 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 41.53 sec  INFO : 2017-04-11 05:28:36,302 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 54.88 sec  INFO : 2017-04-11 05:28:37,517 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 65.35 sec  INFO : 2017-04-11 05:28:38,565 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 71.92 sec  INFO : 2017-04-11 05:28:39,601 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 74.92 sec  INFO : 2017-04-11 05:28:40,636 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 79.19 sec  INFO : 2017-04-11 05:28:41,669 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 82.48 sec  INFO : 2017-04-11 05:28:43,720 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 91.5 sec  INFO : 2017-04-11 05:28:44,745 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 96.4 sec  INFO : 2017-04-11 05:28:51,922 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 97.45 sec  INFO : MapReduce Total cumulative CPU time: 1 minutes 37 seconds 450 msec  INFO : Ended Job = job\_1490094830773\_0352  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 19 Reduce: 1 Cumulative CPU: 97.45 sec HDFS Read: 516820347 HDFS Write: 31 SUCCESS  INFO : Total MapReduce CPU Time Spent: 1 minutes 37 seconds 450 msec  INFO : Completed executing command(queryId=hive\_20170411052828\_20c967e0-664d-4919-ba51-3320d699ef66); Time taken: 35.166 seconds  INFO : OK |

## User wise Max Date

|  |
| --- |
| select username,max(datetime) from postimport\_parquet group by username;  INFO : Compiling command(queryId=hive\_20170411052929\_f9c29974-21b5-4409-bc86-548cad79f3e3): select username,max(datetime) from postimport\_parquet group by username  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:username, type:string, comment:null), FieldSchema(name:\_c1, type:timestamp, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411052929\_f9c29974-21b5-4409-bc86-548cad79f3e3); Time taken: 0.091 seconds  INFO : Executing command(queryId=hive\_20170411052929\_f9c29974-21b5-4409-bc86-548cad79f3e3): select username,max(datetime) from postimport\_parquet group by username  INFO : Query ID = hive\_20170411052929\_f9c29974-21b5-4409-bc86-548cad79f3e3  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks not specified. Estimated from input data size: 74  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:19  INFO : Submitting tokens for job: job\_1490094830773\_0353  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0353/  INFO : Starting Job = job\_1490094830773\_0353, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0353/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0353  INFO : Hadoop job information for Stage-1: number of mappers: 19; number of reducers: 74  INFO : 2017-04-11 05:29:27,009 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 05:29:38,432 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 21.49 sec  INFO : 2017-04-11 05:29:39,465 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 45.98 sec  INFO : 2017-04-11 05:29:40,498 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 64.27 sec  INFO : 2017-04-11 05:29:41,526 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 83.37 sec  INFO : 2017-04-11 05:29:42,554 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 86.48 sec  INFO : 2017-04-11 05:29:43,579 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 87.93 sec  INFO : 2017-04-11 05:29:44,607 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 91.29 sec  INFO : 2017-04-11 05:29:46,683 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 97.19 sec  INFO : 2017-04-11 05:29:47,727 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 109.01 sec  INFO : 2017-04-11 05:29:50,904 Stage-1 map = 100%, reduce = 1%, Cumulative CPU 110.53 sec  INFO : 2017-04-11 05:29:51,928 Stage-1 map = 100%, reduce = 19%, Cumulative CPU 126.36 sec  INFO : 2017-04-11 05:29:52,953 Stage-1 map = 100%, reduce = 20%, Cumulative CPU 127.59 sec  INFO : 2017-04-11 05:29:56,119 Stage-1 map = 100%, reduce = 23%, Cumulative CPU 130.31 sec  INFO : 2017-04-11 05:29:57,164 Stage-1 map = 100%, reduce = 31%, Cumulative CPU 137.25 sec  INFO : 2017-04-11 05:29:58,259 Stage-1 map = 100%, reduce = 36%, Cumulative CPU 141.97 sec  INFO : 2017-04-11 05:29:59,300 Stage-1 map = 100%, reduce = 41%, Cumulative CPU 145.86 sec  INFO : 2017-04-11 05:30:01,401 Stage-1 map = 100%, reduce = 42%, Cumulative CPU 147.16 sec  INFO : 2017-04-11 05:30:02,464 Stage-1 map = 100%, reduce = 45%, Cumulative CPU 149.95 sec  INFO : 2017-04-11 05:30:03,515 Stage-1 map = 100%, reduce = 54%, Cumulative CPU 158.3 sec  INFO : 2017-04-11 05:30:04,549 Stage-1 map = 100%, reduce = 58%, Cumulative CPU 162.36 sec  INFO : 2017-04-11 05:30:05,589 Stage-1 map = 100%, reduce = 61%, Cumulative CPU 164.83 sec  INFO : 2017-04-11 05:30:06,637 Stage-1 map = 100%, reduce = 62%, Cumulative CPU 166.15 sec  INFO : 2017-04-11 05:30:07,679 Stage-1 map = 100%, reduce = 65%, Cumulative CPU 168.67 sec  INFO : 2017-04-11 05:30:08,728 Stage-1 map = 100%, reduce = 66%, Cumulative CPU 169.93 sec  INFO : 2017-04-11 05:30:09,783 Stage-1 map = 100%, reduce = 76%, Cumulative CPU 178.08 sec  INFO : 2017-04-11 05:30:10,808 Stage-1 map = 100%, reduce = 78%, Cumulative CPU 180.58 sec  INFO : 2017-04-11 05:30:11,866 Stage-1 map = 100%, reduce = 82%, Cumulative CPU 184.49 sec  INFO : 2017-04-11 05:30:12,903 Stage-1 map = 100%, reduce = 85%, Cumulative CPU 187.05 sec  INFO : 2017-04-11 05:30:14,960 Stage-1 map = 100%, reduce = 89%, Cumulative CPU 190.8 sec  INFO : 2017-04-11 05:30:16,011 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 200.17 sec  INFO : MapReduce Total cumulative CPU time: 3 minutes 20 seconds 170 msec  INFO : Ended Job = job\_1490094830773\_0353  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 19 Reduce: 74 Cumulative CPU: 200.17 sec HDFS Read: 588361778 HDFS Write: 64445 SUCCESS  INFO : Total MapReduce CPU Time Spent: 3 minutes 20 seconds 170 msec  INFO : Completed executing command(queryId=hive\_20170411052929\_f9c29974-21b5-4409-bc86-548cad79f3e3); Time taken: 58.89 seconds  INFO : OK |

## Select All Columns

|  |
| --- |
| select \* from postimport\_parquet  INFO : Compiling command(queryId=hive\_20170411053030\_51591cf9-2911-47d1-83dc-f21883c2c5e1): select \* from postimport\_parquet  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_parquet.username, type:string, comment:null), FieldSchema(name:postimport\_parquet.machinename, type:string, comment:null), FieldSchema(name:postimport\_parquet.machineid, type:string, comment:null), FieldSchema(name:postimport\_parquet.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_parquet.domainname, type:string, comment:null), FieldSchema(name:postimport\_parquet.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_parquet.modulename, type:string, comment:null), FieldSchema(name:postimport\_parquet.filename, type:string, comment:null), FieldSchema(name:postimport\_parquet.productname, type:string, comment:null), FieldSchema(name:postimport\_parquet.companyname, type:string, comment:null), FieldSchema(name:postimport\_parquet.description, type:string, comment:null), FieldSchema(name:postimport\_parquet.text, type:string, comment:null), FieldSchema(name:postimport\_parquet.activetime, type:double, comment:null), FieldSchema(name:postimport\_parquet.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_parquet.timeslot, type:timestamp, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411053030\_51591cf9-2911-47d1-83dc-f21883c2c5e1); Time taken: 0.1 seconds  INFO : Executing command(queryId=hive\_20170411053030\_51591cf9-2911-47d1-83dc-f21883c2c5e1): select \* from postimport\_parquet  INFO : Query ID = hive\_20170411053030\_51591cf9-2911-47d1-83dc-f21883c2c5e1  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:19  INFO : Submitting tokens for job: job\_1490094830773\_0354  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0354/  INFO : Starting Job = job\_1490094830773\_0354, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0354/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0354  INFO : Hadoop job information for Stage-1: number of mappers: 19; number of reducers: 0  INFO : 2017-04-11 05:30:58,359 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 05:31:21,202 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 153.38 sec  INFO : 2017-04-11 05:31:26,400 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 194.51 sec  INFO : 2017-04-11 05:31:29,497 Stage-1 map = 6%, reduce = 0%, Cumulative CPU 219.33 sec  INFO : 2017-04-11 05:31:32,606 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 244.4 sec  INFO : 2017-04-11 05:31:34,677 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 253.96 sec  INFO : 2017-04-11 05:31:36,747 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 274.08 sec  INFO : 2017-04-11 05:31:38,798 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 293.42 sec  INFO : 2017-04-11 05:31:39,840 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 298.65 sec  INFO : 2017-04-11 05:31:40,864 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 310.76 sec  INFO : 2017-04-11 05:31:41,922 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 315.6 sec  INFO : 2017-04-11 05:31:43,983 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 333.75 sec  INFO : 2017-04-11 05:31:46,039 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 343.43 sec  INFO : 2017-04-11 05:31:50,151 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 379.97 sec  INFO : 2017-04-11 05:31:51,177 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 381.72 sec  INFO : 2017-04-11 05:31:53,229 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 398.4 sec  INFO : 2017-04-11 05:31:56,299 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 415.83 sec  INFO : 2017-04-11 05:32:01,430 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 439.65 sec  INFO : 2017-04-11 05:32:02,455 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 454.08 sec  INFO : 2017-04-11 05:32:05,538 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 470.81 sec  INFO : 2017-04-11 05:32:07,588 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 477.81 sec  INFO : 2017-04-11 05:32:08,613 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 485.37 sec  INFO : 2017-04-11 05:32:23,995 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 562.98 sec  INFO : 2017-04-11 05:32:25,027 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 567.86 sec  INFO : 2017-04-11 05:32:28,094 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 582.16 sec  INFO : 2017-04-11 05:32:30,147 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 593.26 sec  INFO : 2017-04-11 05:32:32,193 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 603.61 sec  INFO : 2017-04-11 05:32:33,214 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 608.68 sec  INFO : 2017-04-11 05:32:35,259 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 617.56 sec  INFO : 2017-04-11 05:32:36,277 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 623.73 sec  INFO : 2017-04-11 05:32:37,319 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 629.59 sec  INFO : 2017-04-11 05:32:40,381 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 643.94 sec  INFO : 2017-04-11 05:32:41,404 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 647.91 sec  INFO : 2017-04-11 05:32:47,569 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 677.65 sec  INFO : 2017-04-11 05:32:48,590 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 682.73 sec  INFO : 2017-04-11 05:32:49,616 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 688.96 sec  INFO : 2017-04-11 05:32:54,734 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 715.49 sec  INFO : 2017-04-11 05:32:57,789 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 726.06 sec  INFO : 2017-04-11 05:32:58,810 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 729.76 sec  INFO : 2017-04-11 05:33:04,926 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 743.04 sec  INFO : 2017-04-11 05:33:10,023 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 764.19 sec  INFO : 2017-04-11 05:33:14,102 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 778.36 sec  INFO : 2017-04-11 05:33:23,303 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 806.18 sec  INFO : 2017-04-11 05:33:25,343 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 809.62 sec  INFO : 2017-04-11 05:33:26,362 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 815.64 sec  INFO : 2017-04-11 05:33:32,484 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 835.89 sec  INFO : 2017-04-11 05:33:36,565 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 846.59 sec  INFO : 2017-04-11 05:33:37,583 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 848.68 sec  INFO : 2017-04-11 05:33:41,683 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 861.7 sec  INFO : 2017-04-11 05:33:44,742 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 871.08 sec  INFO : 2017-04-11 05:33:47,797 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 883.16 sec  INFO : 2017-04-11 05:33:49,837 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 890.17 sec  INFO : 2017-04-11 05:33:54,932 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 907.19 sec  INFO : 2017-04-11 05:33:55,955 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 912.12 sec  INFO : 2017-04-11 05:33:56,974 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 914.96 sec  INFO : 2017-04-11 05:34:00,031 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 923.02 sec  INFO : 2017-04-11 05:34:09,211 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 938.27 sec  INFO : 2017-04-11 05:34:18,391 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 947.92 sec  INFO : 2017-04-11 05:34:21,450 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 952.16 sec  INFO : 2017-04-11 05:34:30,633 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 961.19 sec  INFO : 2017-04-11 05:34:39,811 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 970.6 sec  INFO : MapReduce Total cumulative CPU time: 16 minutes 10 seconds 600 msec  INFO : Ended Job = job\_1490094830773\_0354  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 19 Cumulative CPU: 970.6 sec HDFS Read: 4916404308 HDFS Write: 3610418959 SUCCESS  INFO : Total MapReduce CPU Time Spent: 16 minutes 10 seconds 600 msec  INFO : Completed executing command(queryId=hive\_20170411053030\_51591cf9-2911-47d1-83dc-f21883c2c5e1); Time taken: 228.169 seconds  INFO : OK |

## Select All Columns with filter Condition

|  |
| --- |
| select \* from postimport\_parquet where username = '670063006100750074006F007500730065007200';  INFO : Compiling command(queryId=hive\_20170411053535\_3ee431a0-8e64-4b66-bc97-587ce25bb98c): select \* from postimport\_parquet where username = '670063006100750074006F007500730065007200'  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_parquet.username, type:string, comment:null), FieldSchema(name:postimport\_parquet.machinename, type:string, comment:null), FieldSchema(name:postimport\_parquet.machineid, type:string, comment:null), FieldSchema(name:postimport\_parquet.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_parquet.domainname, type:string, comment:null), FieldSchema(name:postimport\_parquet.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_parquet.modulename, type:string, comment:null), FieldSchema(name:postimport\_parquet.filename, type:string, comment:null), FieldSchema(name:postimport\_parquet.productname, type:string, comment:null), FieldSchema(name:postimport\_parquet.companyname, type:string, comment:null), FieldSchema(name:postimport\_parquet.description, type:string, comment:null), FieldSchema(name:postimport\_parquet.text, type:string, comment:null), FieldSchema(name:postimport\_parquet.activetime, type:double, comment:null), FieldSchema(name:postimport\_parquet.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_parquet.timeslot, type:timestamp, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411053535\_3ee431a0-8e64-4b66-bc97-587ce25bb98c); Time taken: 0.104 seconds  INFO : Executing command(queryId=hive\_20170411053535\_3ee431a0-8e64-4b66-bc97-587ce25bb98c): select \* from postimport\_parquet where username = '670063006100750074006F007500730065007200'  INFO : Query ID = hive\_20170411053535\_3ee431a0-8e64-4b66-bc97-587ce25bb98c  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:19  INFO : Submitting tokens for job: job\_1490094830773\_0355  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0355/  INFO : Starting Job = job\_1490094830773\_0355, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0355/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0355  INFO : Hadoop job information for Stage-1: number of mappers: 19; number of reducers: 0  INFO : 2017-04-11 05:35:12,282 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 05:35:27,931 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 38.17 sec  INFO : 2017-04-11 05:35:31,073 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 87.59 sec  INFO : 2017-04-11 05:35:32,095 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 99.93 sec  INFO : 2017-04-11 05:35:33,120 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 108.13 sec  INFO : 2017-04-11 05:35:34,173 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 110.42 sec  INFO : 2017-04-11 05:35:35,199 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 119.0 sec  INFO : 2017-04-11 05:35:36,227 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 120.93 sec  INFO : 2017-04-11 05:35:37,264 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 126.09 sec  INFO : 2017-04-11 05:35:38,285 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 128.6 sec  INFO : 2017-04-11 05:35:39,315 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 131.92 sec  INFO : 2017-04-11 05:35:40,362 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 138.05 sec  INFO : 2017-04-11 05:35:42,415 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 141.79 sec  INFO : 2017-04-11 05:35:43,436 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 151.04 sec  INFO : 2017-04-11 05:35:44,459 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 152.88 sec  INFO : 2017-04-11 05:35:47,533 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 159.36 sec  INFO : 2017-04-11 05:35:49,580 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 161.62 sec  INFO : 2017-04-11 05:35:51,647 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 170.37 sec  INFO : MapReduce Total cumulative CPU time: 2 minutes 50 seconds 370 msec  INFO : Ended Job = job\_1490094830773\_0355  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 19 Cumulative CPU: 170.37 sec HDFS Read: 4916413162 HDFS Write: 2065874 SUCCESS  INFO : Total MapReduce CPU Time Spent: 2 minutes 50 seconds 370 msec  INFO : Completed executing command(queryId=hive\_20170411053535\_3ee431a0-8e64-4b66-bc97-587ce25bb98c); Time taken: 51.371 seconds  INFO : OK |

# Parquet File format Gzip Compression

|  |
| --- |
| CREATE TABLE postimport\_parquet(  username string,  machinename string,  machineid string,  ipaddress string,  domainname string,  pcmlogeventid string,  modulename string,  filename string,  productname string,  companyname string,  description string,  text string,  activetime double,  datetime timestamp,  timeslot timestamp)  STORED AS PARQUET; |

***Note: Parquet does not support date data type in hive.***

## Load data to Parquet Table

|  |
| --- |
| hive> set parquet.compression=GZIP; |

|  |
| --- |
| hive> truncate table postimport\_parquet |

|  |
| --- |
| insert into postimport\_parquet select username,machinename,machineid,ipaddress,domainname,pcmlogeventid,modulename,filename,productname,companyname,description,text,activetime,datetime,timeslot from postimport\_text; |

|  |
| --- |
| hive> insert into postimport\_parquet select username,machinename,machineid,ipaddress,domainname,pcmlogeventid,modulename,filename,productname,companyname,description,text,activetime,datetime,timeslot from postimport\_text;  Query ID = shalaj\_20170411053838\_a5c7cc1d-d97a-4bfa-add0-6f6342e48b49  Total jobs = 3  Launching Job 1 out of 3  Number of reduce tasks is set to 0 since there's no reduce operator  Starting Job = job\_1490094830773\_0357, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0357/  Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0357  Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  2017-04-11 05:38:44,328 Stage-1 map = 0%, reduce = 0%  2017-04-11 05:38:58,897 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 63.19 sec  2017-04-11 05:38:59,934 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 72.07 sec  2017-04-11 05:39:01,982 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 104.77 sec  2017-04-11 05:39:04,112 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 124.98 sec  2017-04-11 05:39:05,184 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 131.33 sec  2017-04-11 05:39:06,259 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 145.16 sec  2017-04-11 05:39:07,334 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 151.18 sec  2017-04-11 05:39:13,457 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 184.84 sec  2017-04-11 05:39:15,775 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 192.38 sec  2017-04-11 05:39:16,977 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 195.24 sec  2017-04-11 05:39:19,354 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 210.72 sec  2017-04-11 05:39:26,061 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 227.77 sec  2017-04-11 05:39:27,173 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 229.49 sec  2017-04-11 05:39:32,406 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 272.34 sec  2017-04-11 05:39:34,491 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 298.78 sec  2017-04-11 05:39:36,532 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 325.03 sec  2017-04-11 05:39:37,582 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 331.24 sec  2017-04-11 05:39:38,608 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 346.57 sec  2017-04-11 05:39:39,629 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 353.58 sec  2017-04-11 05:39:40,654 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 355.45 sec  2017-04-11 05:39:41,675 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 365.04 sec  2017-04-11 05:39:42,700 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 371.44 sec  2017-04-11 05:39:44,843 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 385.5 sec  2017-04-11 05:39:48,461 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 406.9 sec  2017-04-11 05:39:50,784 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 416.22 sec  2017-04-11 05:39:52,039 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 420.38 sec  2017-04-11 05:39:54,675 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 424.64 sec  2017-04-11 05:40:01,944 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 437.34 sec  2017-04-11 05:40:05,107 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 490.43 sec  2017-04-11 05:40:06,380 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 495.68 sec  2017-04-11 05:40:08,819 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 502.53 sec  2017-04-11 05:40:12,214 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 533.18 sec  2017-04-11 05:40:14,269 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 542.11 sec  2017-04-11 05:40:15,316 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 548.0 sec  2017-04-11 05:40:18,427 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 565.61 sec  2017-04-11 05:40:20,478 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 592.8 sec  2017-04-11 05:40:22,526 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 608.41 sec  2017-04-11 05:40:24,628 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 625.7 sec  2017-04-11 05:40:26,877 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 639.65 sec  2017-04-11 05:40:28,046 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 660.37 sec  2017-04-11 05:40:30,382 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 676.37 sec  2017-04-11 05:40:31,583 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 684.74 sec  2017-04-11 05:40:32,741 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 689.59 sec  2017-04-11 05:40:35,076 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 699.68 sec  2017-04-11 05:40:40,672 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 720.02 sec  2017-04-11 05:40:42,720 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 733.15 sec  2017-04-11 05:40:44,770 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 755.84 sec  2017-04-11 05:40:48,115 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 780.97 sec  2017-04-11 05:40:49,183 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 787.17 sec  2017-04-11 05:40:50,211 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 793.97 sec  2017-04-11 05:40:51,242 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 807.06 sec  2017-04-11 05:40:54,366 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 824.23 sec  2017-04-11 05:40:55,407 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 830.79 sec  2017-04-11 05:40:58,514 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 850.5 sec  2017-04-11 05:41:00,649 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 873.83 sec  2017-04-11 05:41:01,929 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 890.09 sec  2017-04-11 05:41:03,001 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 898.2 sec  2017-04-11 05:41:04,073 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 904.58 sec  2017-04-11 05:41:05,118 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 914.39 sec  2017-04-11 05:41:07,248 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 919.92 sec  2017-04-11 05:41:09,306 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 926.39 sec  2017-04-11 05:41:11,429 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 930.97 sec  MapReduce Total cumulative CPU time: 15 minutes 30 seconds 970 msec  Ended Job = job\_1490094830773\_0357  Stage-4 is selected by condition resolver.  Stage-3 is filtered out by condition resolver.  Stage-5 is filtered out by condition resolver.  Moving data to: hdfs://mac55:8020/user/hive/warehouse/esplus.db/postimport\_parquet/.hive-staging\_hive\_2017-04-11\_05-38-35\_128\_4126610853887218044-1/-ext-10000  Loading data to table esplus.postimport\_parquet  Table esplus.postimport\_parquet stats: [numFiles=74, numRows=42286538, totalSize=1547810132, rawDataSize=634298070]  MapReduce Jobs Launched:  Stage-Stage-1: Map: 74 Cumulative CPU: 931.22 sec HDFS Read: 19486344812 HDFS Write: 1547816792 SUCCESS  Total MapReduce CPU Time Spent: 15 minutes 31 seconds 220 msec  OK  Time taken: 161.73 seconds |

## Data Size

|  |
| --- |
| hive> dfs -du -s -h /user/hive/warehouse/esplus.db/postimport\_parquet;  1.4 G 4.3 G /user/hive/warehouse/esplus.db/postimport\_parquet |

## Select some columns

SET hive.exec.compress.output=true;

|  |
| --- |
| select productname,username from postimport\_parquet;  INFO : Compiling command(queryId=hive\_20170411054444\_2ba5283e-c9d7-40d9-ab2e-08d56d3f7f39): select productname,username from postimport\_parquet  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:productname, type:string, comment:null), FieldSchema(name:username, type:string, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411054444\_2ba5283e-c9d7-40d9-ab2e-08d56d3f7f39); Time taken: 0.12 seconds  INFO : Executing command(queryId=hive\_20170411054444\_2ba5283e-c9d7-40d9-ab2e-08d56d3f7f39): select productname,username from postimport\_parquet  INFO : Query ID = hive\_20170411054444\_2ba5283e-c9d7-40d9-ab2e-08d56d3f7f39  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:6  INFO : Submitting tokens for job: job\_1490094830773\_0359  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0359/  INFO : Starting Job = job\_1490094830773\_0359, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0359/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0359  INFO : Hadoop job information for Stage-1: number of mappers: 6; number of reducers: 0  INFO : 2017-04-11 05:44:20,177 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 05:44:29,381 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 26.69 sec  INFO : 2017-04-11 05:44:32,445 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 38.68 sec  INFO : 2017-04-11 05:44:33,469 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 51.28 sec  INFO : 2017-04-11 05:44:35,512 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 62.59 sec  INFO : 2017-04-11 05:44:36,532 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 68.72 sec  INFO : 2017-04-11 05:44:38,575 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 79.87 sec  INFO : 2017-04-11 05:44:39,596 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 86.07 sec  INFO : 2017-04-11 05:44:40,617 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 86.48 sec  INFO : 2017-04-11 05:44:41,643 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 94.33 sec  INFO : 2017-04-11 05:44:42,664 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 102.41 sec  INFO : 2017-04-11 05:44:43,699 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 103.86 sec  INFO : 2017-04-11 05:44:44,722 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 109.6 sec  INFO : 2017-04-11 05:44:46,779 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 110.99 sec  INFO : 2017-04-11 05:44:47,819 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 113.21 sec  INFO : MapReduce Total cumulative CPU time: 1 minutes 53 seconds 210 msec  INFO : Ended Job = job\_1490094830773\_0359  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 6 Cumulative CPU: 113.21 sec HDFS Read: 63818264 HDFS Write: 233689434 SUCCESS  INFO : Total MapReduce CPU Time Spent: 1 minutes 53 seconds 210 msec  INFO : Completed executing command(queryId=hive\_20170411054444\_2ba5283e-c9d7-40d9-ab2e-08d56d3f7f39); Time taken: 37.26 seconds  INFO : OK |

## Check Max Date

|  |
| --- |
| select max(datetime) from postimport\_parquet;  INFO : Compiling command(queryId=hive\_20170411054545\_77d02e58-bc94-49db-9d2c-c1d26977e565): select max(datetime) from postimport\_parquet  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:\_c0, type:timestamp, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411054545\_77d02e58-bc94-49db-9d2c-c1d26977e565); Time taken: 0.253 seconds  INFO : Executing command(queryId=hive\_20170411054545\_77d02e58-bc94-49db-9d2c-c1d26977e565): select max(datetime) from postimport\_parquet  INFO : Query ID = hive\_20170411054545\_77d02e58-bc94-49db-9d2c-c1d26977e565  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks determined at compile time: 1  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:6  INFO : Submitting tokens for job: job\_1490094830773\_0360  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0360/  INFO : Starting Job = job\_1490094830773\_0360, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0360/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0360  INFO : Hadoop job information for Stage-1: number of mappers: 6; number of reducers: 1  INFO : 2017-04-11 05:45:57,887 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 05:46:07,108 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 7.18 sec  INFO : 2017-04-11 05:46:08,131 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 28.6 sec  INFO : 2017-04-11 05:46:09,153 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 31.2 sec  INFO : 2017-04-11 05:46:10,179 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 48.79 sec  INFO : 2017-04-11 05:46:11,203 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 51.42 sec  INFO : 2017-04-11 05:46:15,301 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 52.36 sec  INFO : MapReduce Total cumulative CPU time: 52 seconds 360 msec  INFO : Ended Job = job\_1490094830773\_0360  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 6 Reduce: 1 Cumulative CPU: 52.36 sec HDFS Read: 268644832 HDFS Write: 31 SUCCESS  INFO : Total MapReduce CPU Time Spent: 52 seconds 360 msec  INFO : Completed executing command(queryId=hive\_20170411054545\_77d02e58-bc94-49db-9d2c-c1d26977e565); Time taken: 24.164 seconds  INFO : OK |

## User wise Max Date

|  |
| --- |
| select username,max(datetime) from postimport\_parquet group by username;  INFO : Compiling command(queryId=hive\_20170411054646\_53af3b77-4922-44de-b49e-b5b5908c91df): select username,max(datetime) from postimport\_parquet group by username  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:username, type:string, comment:null), FieldSchema(name:\_c1, type:timestamp, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411054646\_53af3b77-4922-44de-b49e-b5b5908c91df); Time taken: 0.097 seconds  INFO : Executing command(queryId=hive\_20170411054646\_53af3b77-4922-44de-b49e-b5b5908c91df): select username,max(datetime) from postimport\_parquet group by username  INFO : Query ID = hive\_20170411054646\_53af3b77-4922-44de-b49e-b5b5908c91df  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks not specified. Estimated from input data size: 24  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:6  INFO : Submitting tokens for job: job\_1490094830773\_0361  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0361/  INFO : Starting Job = job\_1490094830773\_0361, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0361/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0361  INFO : Hadoop job information for Stage-1: number of mappers: 6; number of reducers: 24  INFO : 2017-04-11 05:46:49,460 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 05:46:59,705 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 31.0 sec  INFO : 2017-04-11 05:47:00,737 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 32.53 sec  INFO : 2017-04-11 05:47:01,758 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 52.01 sec  INFO : 2017-04-11 05:47:02,798 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 55.0 sec  INFO : 2017-04-11 05:47:04,843 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 61.95 sec  INFO : 2017-04-11 05:47:11,105 Stage-1 map = 100%, reduce = 29%, Cumulative CPU 70.84 sec  INFO : 2017-04-11 05:47:12,137 Stage-1 map = 100%, reduce = 63%, Cumulative CPU 80.73 sec  INFO : 2017-04-11 05:47:15,218 Stage-1 map = 100%, reduce = 79%, Cumulative CPU 85.62 sec  INFO : 2017-04-11 05:47:16,243 Stage-1 map = 100%, reduce = 83%, Cumulative CPU 86.85 sec  INFO : 2017-04-11 05:47:17,263 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 91.37 sec  INFO : MapReduce Total cumulative CPU time: 1 minutes 31 seconds 370 msec  INFO : Ended Job = job\_1490094830773\_0361  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 6 Reduce: 24 Cumulative CPU: 91.37 sec HDFS Read: 301299869 HDFS Write: 57422 SUCCESS  INFO : Total MapReduce CPU Time Spent: 1 minutes 31 seconds 370 msec  INFO : Completed executing command(queryId=hive\_20170411054646\_53af3b77-4922-44de-b49e-b5b5908c91df); Time taken: 35.62 seconds  INFO : OK |

## Select All Columns

|  |
| --- |
| select \* from postimport\_parquet  INFO : Compiling command(queryId=hive\_20170411054747\_d8707d63-c291-45dd-b5b9-8e03454b9ff4): select \* from postimport\_parquet  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_parquet.username, type:string, comment:null), FieldSchema(name:postimport\_parquet.machinename, type:string, comment:null), FieldSchema(name:postimport\_parquet.machineid, type:string, comment:null), FieldSchema(name:postimport\_parquet.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_parquet.domainname, type:string, comment:null), FieldSchema(name:postimport\_parquet.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_parquet.modulename, type:string, comment:null), FieldSchema(name:postimport\_parquet.filename, type:string, comment:null), FieldSchema(name:postimport\_parquet.productname, type:string, comment:null), FieldSchema(name:postimport\_parquet.companyname, type:string, comment:null), FieldSchema(name:postimport\_parquet.description, type:string, comment:null), FieldSchema(name:postimport\_parquet.text, type:string, comment:null), FieldSchema(name:postimport\_parquet.activetime, type:double, comment:null), FieldSchema(name:postimport\_parquet.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_parquet.timeslot, type:timestamp, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411054747\_d8707d63-c291-45dd-b5b9-8e03454b9ff4); Time taken: 0.106 seconds  INFO : Executing command(queryId=hive\_20170411054747\_d8707d63-c291-45dd-b5b9-8e03454b9ff4): select \* from postimport\_parquet  INFO : Query ID = hive\_20170411054747\_d8707d63-c291-45dd-b5b9-8e03454b9ff4  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:6  INFO : Submitting tokens for job: job\_1490094830773\_0362  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0362/  INFO : Starting Job = job\_1490094830773\_0362, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0362/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0362  INFO : Hadoop job information for Stage-1: number of mappers: 6; number of reducers: 0  INFO : 2017-04-11 05:47:37,968 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 05:47:57,413 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 86.37 sec  INFO : 2017-04-11 05:47:59,456 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 94.82 sec  INFO : 2017-04-11 05:48:00,478 Stage-1 map = 6%, reduce = 0%, Cumulative CPU 102.5 sec  INFO : 2017-04-11 05:48:03,539 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 118.42 sec  INFO : 2017-04-11 05:48:09,668 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 148.66 sec  INFO : 2017-04-11 05:48:11,710 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 157.03 sec  INFO : 2017-04-11 05:48:17,838 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 190.93 sec  INFO : 2017-04-11 05:48:18,858 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 195.39 sec  INFO : 2017-04-11 05:48:21,922 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 208.51 sec  INFO : 2017-04-11 05:48:23,965 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 219.23 sec  INFO : 2017-04-11 05:48:35,193 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 261.15 sec  INFO : 2017-04-11 05:48:36,218 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 269.72 sec  INFO : 2017-04-11 05:48:39,281 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 285.02 sec  INFO : 2017-04-11 05:48:42,346 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 300.05 sec  INFO : 2017-04-11 05:48:50,515 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 330.82 sec  INFO : 2017-04-11 05:48:51,537 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 342.89 sec  INFO : 2017-04-11 05:48:54,598 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 357.26 sec  INFO : 2017-04-11 05:48:57,667 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 371.5 sec  INFO : 2017-04-11 05:49:05,830 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 404.89 sec  INFO : 2017-04-11 05:49:06,850 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 413.87 sec  INFO : 2017-04-11 05:49:12,978 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 442.98 sec  INFO : 2017-04-11 05:49:15,021 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 450.81 sec  INFO : 2017-04-11 05:49:21,144 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 482.08 sec  INFO : 2017-04-11 05:49:25,225 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 503.06 sec  INFO : 2017-04-11 05:49:27,267 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 513.51 sec  INFO : 2017-04-11 05:49:30,331 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 533.78 sec  INFO : 2017-04-11 05:49:36,475 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 571.09 sec  INFO : 2017-04-11 05:49:37,494 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 574.05 sec  INFO : 2017-04-11 05:49:39,534 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 585.69 sec  INFO : 2017-04-11 05:49:45,647 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 614.66 sec  INFO : 2017-04-11 05:49:48,705 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 626.83 sec  INFO : 2017-04-11 05:49:54,820 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 651.86 sec  INFO : 2017-04-11 05:49:56,861 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 653.51 sec  INFO : 2017-04-11 05:50:00,937 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 677.58 sec  INFO : 2017-04-11 05:50:10,110 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 716.35 sec  INFO : 2017-04-11 05:50:13,171 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 729.56 sec  INFO : 2017-04-11 05:50:16,228 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 742.48 sec  INFO : 2017-04-11 05:50:18,266 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 746.05 sec  INFO : 2017-04-11 05:50:23,360 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 768.83 sec  INFO : 2017-04-11 05:50:24,379 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 774.25 sec  INFO : 2017-04-11 05:50:25,400 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 778.4 sec  INFO : 2017-04-11 05:50:30,496 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 795.03 sec  INFO : 2017-04-11 05:50:35,594 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 811.6 sec  INFO : 2017-04-11 05:50:36,611 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 819.31 sec  INFO : 2017-04-11 05:50:39,667 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 827.56 sec  INFO : 2017-04-11 05:50:48,841 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 849.22 sec  INFO : 2017-04-11 05:50:51,896 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 856.84 sec  INFO : 2017-04-11 05:50:54,956 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 864.5 sec  INFO : 2017-04-11 05:51:01,073 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 879.74 sec  INFO : 2017-04-11 05:51:04,136 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 888.82 sec  INFO : 2017-04-11 05:51:05,154 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 890.31 sec  INFO : 2017-04-11 05:51:12,294 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 898.99 sec  INFO : MapReduce Total cumulative CPU time: 14 minutes 58 seconds 990 msec  INFO : Ended Job = job\_1490094830773\_0362  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 6 Cumulative CPU: 898.99 sec HDFS Read: 1547994337 HDFS Write: 3610413326 SUCCESS  INFO : Total MapReduce CPU Time Spent: 14 minutes 58 seconds 990 msec  INFO : Completed executing command(queryId=hive\_20170411054747\_d8707d63-c291-45dd-b5b9-8e03454b9ff4); Time taken: 220.914 seconds  INFO : OK |

## Select All Columns with filter Condition

|  |
| --- |
| select \* from postimport\_parquet where username = '670063006100750074006F007500730065007200';   * [job\_1490094830773\_0363](http://mac55:8888/jobbrowser/jobs/job_1490094830773_0363)   INFO : Compiling command(queryId=hive\_20170411060404\_59143db8-da1e-4e43-969f-a119d4d2d038): select \* from postimport\_parquet where username = '670063006100750074006F007500730065007200'  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_parquet.username, type:string, comment:null), FieldSchema(name:postimport\_parquet.machinename, type:string, comment:null), FieldSchema(name:postimport\_parquet.machineid, type:string, comment:null), FieldSchema(name:postimport\_parquet.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_parquet.domainname, type:string, comment:null), FieldSchema(name:postimport\_parquet.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_parquet.modulename, type:string, comment:null), FieldSchema(name:postimport\_parquet.filename, type:string, comment:null), FieldSchema(name:postimport\_parquet.productname, type:string, comment:null), FieldSchema(name:postimport\_parquet.companyname, type:string, comment:null), FieldSchema(name:postimport\_parquet.description, type:string, comment:null), FieldSchema(name:postimport\_parquet.text, type:string, comment:null), FieldSchema(name:postimport\_parquet.activetime, type:double, comment:null), FieldSchema(name:postimport\_parquet.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_parquet.timeslot, type:timestamp, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411060404\_59143db8-da1e-4e43-969f-a119d4d2d038); Time taken: 0.112 seconds  INFO : Executing command(queryId=hive\_20170411060404\_59143db8-da1e-4e43-969f-a119d4d2d038): select \* from postimport\_parquet where username = '670063006100750074006F007500730065007200'  INFO : Query ID = hive\_20170411060404\_59143db8-da1e-4e43-969f-a119d4d2d038  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:6  INFO : Submitting tokens for job: job\_1490094830773\_0363  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0363/  INFO : Starting Job = job\_1490094830773\_0363, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0363/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0363  INFO : Hadoop job information for Stage-1: number of mappers: 6; number of reducers: 0  INFO : 2017-04-11 06:04:49,531 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 06:04:58,759 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 7.8 sec  INFO : 2017-04-11 06:04:59,786 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 30.51 sec  INFO : 2017-04-11 06:05:01,835 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 46.19 sec  INFO : 2017-04-11 06:05:02,856 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 57.09 sec  INFO : 2017-04-11 06:05:04,900 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 66.35 sec  INFO : 2017-04-11 06:05:05,925 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 75.98 sec  INFO : 2017-04-11 06:05:07,975 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 87.76 sec  INFO : 2017-04-11 06:05:08,998 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 94.17 sec  INFO : 2017-04-11 06:05:11,052 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 108.22 sec  INFO : 2017-04-11 06:05:12,097 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 108.62 sec  INFO : 2017-04-11 06:05:14,152 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 115.11 sec  INFO : 2017-04-11 06:05:15,187 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 116.61 sec  INFO : 2017-04-11 06:05:16,228 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 119.23 sec  INFO : MapReduce Total cumulative CPU time: 1 minutes 59 seconds 230 msec  INFO : Ended Job = job\_1490094830773\_0363  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 6 Cumulative CPU: 119.23 sec HDFS Read: 1547997085 HDFS Write: 2058649 SUCCESS  INFO : Total MapReduce CPU Time Spent: 1 minutes 59 seconds 230 msec  INFO : Completed executing command(queryId=hive\_20170411060404\_59143db8-da1e-4e43-969f-a119d4d2d038); Time taken: 31.779 seconds  INFO : OK |

# Parquet File format Snappy Compression

|  |
| --- |
| CREATE TABLE postimport\_parquet(  username string,  machinename string,  machineid string,  ipaddress string,  domainname string,  pcmlogeventid string,  modulename string,  filename string,  productname string,  companyname string,  description string,  text string,  activetime double,  datetime timestamp,  timeslot timestamp)  STORED AS PARQUET; |

***Note: Parquet does not support date data type in hive.***

## Load data to Parquet Table

|  |
| --- |
| hive> set parquet.compression=SNAPPY; |

|  |
| --- |
| hive> truncate table postimport\_parquet; |

|  |
| --- |
| insert into postimport\_parquet select username,machinename,machineid,ipaddress,domainname,pcmlogeventid,modulename,filename,productname,companyname,description,text,activetime,datetime,timeslot from postimport\_text; |

|  |
| --- |
| hive> insert into postimport\_parquet select username,machinename,machineid,ipaddress,domainname,pcmlogeventid,modulename,filename,productname,companyname,description,text,activetime,datetime,timeslot from postimport\_text;  Query ID = shalaj\_20170411060808\_e05e6a75-b293-4f1f-b39e-209db52a8d9e  Total jobs = 3  Launching Job 1 out of 3  Number of reduce tasks is set to 0 since there's no reduce operator  Starting Job = job\_1490094830773\_0364, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0364/  Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0364  Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  2017-04-11 06:08:15,014 Stage-1 map = 0%, reduce = 0%  2017-04-11 06:08:29,753 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 57.9 sec  2017-04-11 06:08:30,782 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 68.28 sec  2017-04-11 06:08:32,868 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 88.37 sec  2017-04-11 06:08:35,956 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 111.29 sec  2017-04-11 06:08:37,238 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 119.95 sec  2017-04-11 06:08:41,686 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 128.69 sec  2017-04-11 06:08:42,907 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 133.81 sec  2017-04-11 06:08:45,259 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 158.16 sec  2017-04-11 06:08:46,474 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 161.96 sec  2017-04-11 06:08:48,875 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 171.05 sec  2017-04-11 06:08:51,272 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 178.06 sec  2017-04-11 06:08:53,628 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 190.48 sec  2017-04-11 06:08:55,829 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 196.9 sec  2017-04-11 06:08:58,013 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 201.46 sec  2017-04-11 06:09:00,460 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 202.4 sec  2017-04-11 06:09:03,915 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 232.49 sec  2017-04-11 06:09:04,972 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 245.42 sec  2017-04-11 06:09:06,043 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 247.94 sec  2017-04-11 06:09:09,256 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 269.55 sec  2017-04-11 06:09:10,292 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 276.56 sec  2017-04-11 06:09:13,736 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 292.28 sec  2017-04-11 06:09:19,407 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 329.96 sec  2017-04-11 06:09:20,555 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 336.85 sec  2017-04-11 06:09:21,714 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 352.29 sec  2017-04-11 06:09:22,917 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 359.22 sec  2017-04-11 06:09:26,493 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 372.05 sec  2017-04-11 06:09:27,675 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 377.69 sec  2017-04-11 06:09:28,881 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 387.7 sec  2017-04-11 06:09:32,160 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 410.11 sec  2017-04-11 06:09:33,213 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 418.04 sec  2017-04-11 06:09:35,435 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 423.97 sec  2017-04-11 06:09:36,575 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 427.34 sec  2017-04-11 06:09:44,271 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 449.82 sec  2017-04-11 06:09:45,345 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 452.07 sec  2017-04-11 06:09:49,635 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 504.06 sec  2017-04-11 06:09:51,705 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 512.13 sec  2017-04-11 06:09:52,734 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 545.44 sec  2017-04-11 06:09:54,888 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 560.52 sec  2017-04-11 06:09:55,958 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 565.3 sec  2017-04-11 06:09:57,040 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 568.57 sec  2017-04-11 06:09:58,118 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 574.11 sec  2017-04-11 06:10:00,271 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 593.8 sec  2017-04-11 06:10:01,384 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 595.89 sec  2017-04-11 06:10:02,518 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 598.48 sec  2017-04-11 06:10:05,776 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 618.99 sec  2017-04-11 06:10:06,826 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 630.2 sec  2017-04-11 06:10:09,005 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 640.64 sec  2017-04-11 06:10:10,131 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 647.93 sec  2017-04-11 06:10:14,721 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 669.59 sec  2017-04-11 06:10:19,369 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 697.05 sec  2017-04-11 06:10:22,705 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 728.16 sec  2017-04-11 06:10:23,855 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 736.49 sec  2017-04-11 06:10:24,937 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 739.05 sec  2017-04-11 06:10:26,050 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 744.66 sec  2017-04-11 06:10:27,202 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 750.83 sec  2017-04-11 06:10:28,398 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 752.68 sec  2017-04-11 06:10:29,668 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 758.04 sec  2017-04-11 06:10:30,877 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 760.29 sec  2017-04-11 06:10:33,280 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 764.28 sec  2017-04-11 06:10:34,462 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 781.29 sec  2017-04-11 06:10:37,089 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 789.87 sec  2017-04-11 06:10:38,527 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 796.3 sec  2017-04-11 06:10:39,725 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 801.92 sec  2017-04-11 06:10:43,512 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 810.29 sec  2017-04-11 06:10:45,889 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 813.85 sec  2017-04-11 06:11:03,563 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 823.34 sec  MapReduce Total cumulative CPU time: 13 minutes 43 seconds 340 msec  Ended Job = job\_1490094830773\_0364  Stage-4 is selected by condition resolver.  Stage-3 is filtered out by condition resolver.  Stage-5 is filtered out by condition resolver.  Moving data to: hdfs://mac55:8020/user/hive/warehouse/esplus.db/postimport\_parquet/.hive-staging\_hive\_2017-04-11\_06-08-04\_475\_5986319879173183809-1/-ext-10000  Loading data to table esplus.postimport\_parquet  Table esplus.postimport\_parquet stats: [numFiles=148, numRows=84573076, totalSize=4189068285, rawDataSize=1268596140]  MapReduce Jobs Launched:  Stage-Stage-1: Map: 74 Cumulative CPU: 823.34 sec HDFS Read: 19486345700 HDFS Write: 2641264813 SUCCESS  Total MapReduce CPU Time Spent: 13 minutes 43 seconds 340 msec  OK  Time taken: 183.337 seconds |

## Data Size

|  |
| --- |
| hive> dfs -du -s -h /user/hive/warehouse/esplus.db/postimport\_parquet;  2.5 G 7.4 G /user/hive/warehouse/esplus.db/postimport\_parquet |

## Select some columns

SET hive.exec.compress.output=true;

|  |
| --- |
| select productname,username from postimport\_parquet;  INFO : Compiling command(queryId=hive\_20170411062121\_23e4d696-67c0-41ae-9dd5-59896db8f3e9): select productname,username from postimport\_parquet  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:productname, type:string, comment:null), FieldSchema(name:username, type:string, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411062121\_23e4d696-67c0-41ae-9dd5-59896db8f3e9); Time taken: 0.106 seconds  INFO : Executing command(queryId=hive\_20170411062121\_23e4d696-67c0-41ae-9dd5-59896db8f3e9): select productname,username from postimport\_parquet  INFO : Query ID = hive\_20170411062121\_23e4d696-67c0-41ae-9dd5-59896db8f3e9  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:10  INFO : Submitting tokens for job: job\_1490094830773\_0368  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0368/  INFO : Starting Job = job\_1490094830773\_0368, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0368/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0368  INFO : Hadoop job information for Stage-1: number of mappers: 10; number of reducers: 0  INFO : 2017-04-11 06:21:35,249 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 06:21:45,668 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 13.21 sec  INFO : 2017-04-11 06:21:46,728 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 32.28 sec  INFO : 2017-04-11 06:21:47,769 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 47.69 sec  INFO : 2017-04-11 06:21:48,800 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 74.59 sec  INFO : 2017-04-11 06:21:50,875 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 84.76 sec  INFO : 2017-04-11 06:21:51,907 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 98.48 sec  INFO : 2017-04-11 06:21:52,942 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 100.69 sec  INFO : 2017-04-11 06:21:53,977 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 110.99 sec  INFO : 2017-04-11 06:21:55,018 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 114.93 sec  INFO : 2017-04-11 06:21:57,077 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 118.9 sec  INFO : 2017-04-11 06:22:00,165 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 123.67 sec  INFO : MapReduce Total cumulative CPU time: 2 minutes 3 seconds 670 msec  INFO : Ended Job = job\_1490094830773\_0368  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 10 Cumulative CPU: 123.67 sec HDFS Read: 90661293 HDFS Write: 233696381 SUCCESS  INFO : Total MapReduce CPU Time Spent: 2 minutes 3 seconds 670 msec  INFO : Completed executing command(queryId=hive\_20170411062121\_23e4d696-67c0-41ae-9dd5-59896db8f3e9); Time taken: 31.229 seconds  INFO : OK |

## Check Max Date

|  |
| --- |
| select max(datetime) from postimport\_parquet;  INFO : Compiling command(queryId=hive\_20170411062222\_84c006de-61b1-46c1-9b65-0bdd6b3c0f67): select max(datetime) from postimport\_parquet  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:\_c0, type:timestamp, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411062222\_84c006de-61b1-46c1-9b65-0bdd6b3c0f67); Time taken: 0.103 seconds  INFO : Executing command(queryId=hive\_20170411062222\_84c006de-61b1-46c1-9b65-0bdd6b3c0f67): select max(datetime) from postimport\_parquet  INFO : Query ID = hive\_20170411062222\_84c006de-61b1-46c1-9b65-0bdd6b3c0f67  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks determined at compile time: 1  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:10  INFO : Submitting tokens for job: job\_1490094830773\_0369  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0369/  INFO : Starting Job = job\_1490094830773\_0369, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0369/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0369  INFO : Hadoop job information for Stage-1: number of mappers: 10; number of reducers: 1  INFO : 2017-04-11 06:22:23,718 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 06:22:34,126 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 13.45 sec  INFO : 2017-04-11 06:22:35,161 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 24.82 sec  INFO : 2017-04-11 06:22:36,186 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 61.62 sec  INFO : 2017-04-11 06:22:39,305 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 65.2 sec  INFO : 2017-04-11 06:22:40,338 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 66.13 sec  INFO : MapReduce Total cumulative CPU time: 1 minutes 6 seconds 130 msec  INFO : Ended Job = job\_1490094830773\_0369  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 10 Reduce: 1 Cumulative CPU: 66.13 sec HDFS Read: 340159796 HDFS Write: 31 SUCCESS  INFO : Total MapReduce CPU Time Spent: 1 minutes 6 seconds 130 msec  INFO : Completed executing command(queryId=hive\_20170411062222\_84c006de-61b1-46c1-9b65-0bdd6b3c0f67); Time taken: 24.78 seconds  INFO : OK |

## User wise Max Date

|  |
| --- |
| select username,max(datetime) from postimport\_parquet group by username;  INFO : Compiling command(queryId=hive\_20170411062323\_2fdd34bd-249c-4418-a1e0-7ac1f5b42c81): select username,max(datetime) from postimport\_parquet group by username  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:username, type:string, comment:null), FieldSchema(name:\_c1, type:timestamp, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411062323\_2fdd34bd-249c-4418-a1e0-7ac1f5b42c81); Time taken: 0.11 seconds  INFO : Executing command(queryId=hive\_20170411062323\_2fdd34bd-249c-4418-a1e0-7ac1f5b42c81): select username,max(datetime) from postimport\_parquet group by username  INFO : Query ID = hive\_20170411062323\_2fdd34bd-249c-4418-a1e0-7ac1f5b42c81  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks not specified. Estimated from input data size: 40  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:10  INFO : Submitting tokens for job: job\_1490094830773\_0370  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0370/  INFO : Starting Job = job\_1490094830773\_0370, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0370/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0370  INFO : Hadoop job information for Stage-1: number of mappers: 10; number of reducers: 40  INFO : 2017-04-11 06:23:29,618 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 06:23:40,001 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 31.92 sec  INFO : 2017-04-11 06:23:41,025 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 72.43 sec  INFO : 2017-04-11 06:23:42,057 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 75.88 sec  INFO : 2017-04-11 06:23:43,084 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 79.61 sec  INFO : 2017-04-11 06:23:49,354 Stage-1 map = 100%, reduce = 23%, Cumulative CPU 89.33 sec  INFO : 2017-04-11 06:23:50,387 Stage-1 map = 100%, reduce = 38%, Cumulative CPU 98.44 sec  INFO : 2017-04-11 06:23:53,553 Stage-1 map = 100%, reduce = 40%, Cumulative CPU 99.84 sec  INFO : 2017-04-11 06:23:54,617 Stage-1 map = 100%, reduce = 45%, Cumulative CPU 102.6 sec  INFO : 2017-04-11 06:23:55,665 Stage-1 map = 100%, reduce = 63%, Cumulative CPU 110.86 sec  INFO : 2017-04-11 06:23:56,715 Stage-1 map = 100%, reduce = 65%, Cumulative CPU 112.0 sec  INFO : 2017-04-11 06:23:57,774 Stage-1 map = 100%, reduce = 73%, Cumulative CPU 116.28 sec  INFO : 2017-04-11 06:23:58,801 Stage-1 map = 100%, reduce = 75%, Cumulative CPU 117.93 sec  INFO : 2017-04-11 06:23:59,824 Stage-1 map = 100%, reduce = 83%, Cumulative CPU 121.98 sec  INFO : 2017-04-11 06:24:00,847 Stage-1 map = 100%, reduce = 90%, Cumulative CPU 125.37 sec  INFO : 2017-04-11 06:24:01,881 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 129.98 sec  INFO : MapReduce Total cumulative CPU time: 2 minutes 9 seconds 980 msec  INFO : Ended Job = job\_1490094830773\_0370  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 10 Reduce: 40 Cumulative CPU: 129.98 sec HDFS Read: 386420501 HDFS Write: 60468 SUCCESS  INFO : Total MapReduce CPU Time Spent: 2 minutes 9 seconds 980 msec  INFO : Completed executing command(queryId=hive\_20170411062323\_2fdd34bd-249c-4418-a1e0-7ac1f5b42c81); Time taken: 39.453 seconds  INFO : OK |

## Select All Columns

|  |
| --- |
| select \* from postimport\_parquet  INFO : Compiling command(queryId=hive\_20170411062424\_5168a262-824b-4e88-b61a-ce1edea1006b): select \* from postimport\_parquet  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_parquet.username, type:string, comment:null), FieldSchema(name:postimport\_parquet.machinename, type:string, comment:null), FieldSchema(name:postimport\_parquet.machineid, type:string, comment:null), FieldSchema(name:postimport\_parquet.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_parquet.domainname, type:string, comment:null), FieldSchema(name:postimport\_parquet.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_parquet.modulename, type:string, comment:null), FieldSchema(name:postimport\_parquet.filename, type:string, comment:null), FieldSchema(name:postimport\_parquet.productname, type:string, comment:null), FieldSchema(name:postimport\_parquet.companyname, type:string, comment:null), FieldSchema(name:postimport\_parquet.description, type:string, comment:null), FieldSchema(name:postimport\_parquet.text, type:string, comment:null), FieldSchema(name:postimport\_parquet.activetime, type:double, comment:null), FieldSchema(name:postimport\_parquet.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_parquet.timeslot, type:timestamp, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411062424\_5168a262-824b-4e88-b61a-ce1edea1006b); Time taken: 0.09 seconds  INFO : Executing command(queryId=hive\_20170411062424\_5168a262-824b-4e88-b61a-ce1edea1006b): select \* from postimport\_parquet  INFO : Query ID = hive\_20170411062424\_5168a262-824b-4e88-b61a-ce1edea1006b  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:10  INFO : Submitting tokens for job: job\_1490094830773\_0371  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0371/  INFO : Starting Job = job\_1490094830773\_0371, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0371/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0371  INFO : Hadoop job information for Stage-1: number of mappers: 10; number of reducers: 0  INFO : 2017-04-11 06:24:25,991 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 06:24:45,702 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 126.43 sec  INFO : 2017-04-11 06:24:48,822 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 151.32 sec  INFO : 2017-04-11 06:24:49,850 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 156.59 sec  INFO : 2017-04-11 06:24:50,893 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 171.17 sec  INFO : 2017-04-11 06:24:54,014 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 197.81 sec  INFO : 2017-04-11 06:24:55,055 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 208.85 sec  INFO : 2017-04-11 06:24:57,126 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 222.32 sec  INFO : 2017-04-11 06:25:00,226 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 247.79 sec  INFO : 2017-04-11 06:25:01,264 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 255.4 sec  INFO : 2017-04-11 06:25:04,377 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 280.97 sec  INFO : 2017-04-11 06:25:06,455 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 298.04 sec  INFO : 2017-04-11 06:25:09,577 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 320.09 sec  INFO : 2017-04-11 06:25:12,693 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 339.67 sec  INFO : 2017-04-11 06:25:15,773 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 355.79 sec  INFO : 2017-04-11 06:25:18,858 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 372.49 sec  INFO : 2017-04-11 06:25:19,916 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 375.75 sec  INFO : 2017-04-11 06:25:21,988 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 390.2 sec  INFO : 2017-04-11 06:25:30,236 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 426.94 sec  INFO : 2017-04-11 06:25:31,267 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 436.51 sec  INFO : 2017-04-11 06:25:33,362 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 441.6 sec  INFO : 2017-04-11 06:25:34,402 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 450.36 sec  INFO : 2017-04-11 06:25:37,529 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 463.75 sec  INFO : 2017-04-11 06:25:43,740 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 490.39 sec  INFO : 2017-04-11 06:25:45,814 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 499.21 sec  INFO : 2017-04-11 06:25:48,905 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 512.18 sec  INFO : 2017-04-11 06:25:55,097 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 542.89 sec  INFO : 2017-04-11 06:25:58,221 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 557.81 sec  INFO : 2017-04-11 06:26:00,293 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 562.93 sec  INFO : 2017-04-11 06:26:01,325 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 570.81 sec  INFO : 2017-04-11 06:26:03,385 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 577.01 sec  INFO : 2017-04-11 06:26:04,411 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 585.69 sec  INFO : 2017-04-11 06:26:16,811 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 635.51 sec  INFO : 2017-04-11 06:26:19,916 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 646.81 sec  INFO : 2017-04-11 06:26:23,008 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 657.68 sec  INFO : 2017-04-11 06:26:25,072 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 661.49 sec  INFO : 2017-04-11 06:26:29,149 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 680.76 sec  INFO : 2017-04-11 06:26:34,244 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 696.36 sec  INFO : 2017-04-11 06:26:37,306 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 707.1 sec  INFO : 2017-04-11 06:26:38,325 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 714.68 sec  INFO : 2017-04-11 06:26:41,383 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 724.59 sec  INFO : 2017-04-11 06:26:43,424 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 729.36 sec  INFO : 2017-04-11 06:26:44,447 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 735.72 sec  INFO : 2017-04-11 06:26:52,602 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 764.01 sec  INFO : 2017-04-11 06:26:58,717 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 789.28 sec  INFO : 2017-04-11 06:26:59,749 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 793.56 sec  INFO : 2017-04-11 06:27:01,798 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 806.05 sec  INFO : 2017-04-11 06:27:02,819 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 808.22 sec  INFO : 2017-04-11 06:27:06,902 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 824.9 sec  INFO : 2017-04-11 06:27:13,029 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 849.14 sec  INFO : 2017-04-11 06:27:14,052 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 854.34 sec  INFO : 2017-04-11 06:27:15,070 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 855.03 sec  INFO : 2017-04-11 06:27:17,111 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 866.92 sec  INFO : 2017-04-11 06:27:19,153 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 874.44 sec  INFO : 2017-04-11 06:27:27,314 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 910.17 sec  INFO : 2017-04-11 06:27:28,333 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 914.82 sec  INFO : 2017-04-11 06:27:29,354 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 919.26 sec  INFO : 2017-04-11 06:27:38,533 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 929.49 sec  INFO : MapReduce Total cumulative CPU time: 15 minutes 29 seconds 490 msec  INFO : Ended Job = job\_1490094830773\_0371  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 10 Cumulative CPU: 929.49 sec HDFS Read: 2641461844 HDFS Write: 3610416444 SUCCESS  INFO : Total MapReduce CPU Time Spent: 15 minutes 29 seconds 490 msec  INFO : Completed executing command(queryId=hive\_20170411062424\_5168a262-824b-4e88-b61a-ce1edea1006b); Time taken: 200.57 seconds  INFO : OK |

## Select All Columns with filter Condition

|  |
| --- |
| select \* from postimport\_parquet where username = '670063006100750074006F007500730065007200';  INFO : Compiling command(queryId=hive\_20170411062828\_12abce73-04d3-4ffa-8b42-39808cd00f51): select \* from postimport\_parquet where username = '670063006100750074006F007500730065007200'  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_parquet.username, type:string, comment:null), FieldSchema(name:postimport\_parquet.machinename, type:string, comment:null), FieldSchema(name:postimport\_parquet.machineid, type:string, comment:null), FieldSchema(name:postimport\_parquet.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_parquet.domainname, type:string, comment:null), FieldSchema(name:postimport\_parquet.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_parquet.modulename, type:string, comment:null), FieldSchema(name:postimport\_parquet.filename, type:string, comment:null), FieldSchema(name:postimport\_parquet.productname, type:string, comment:null), FieldSchema(name:postimport\_parquet.companyname, type:string, comment:null), FieldSchema(name:postimport\_parquet.description, type:string, comment:null), FieldSchema(name:postimport\_parquet.text, type:string, comment:null), FieldSchema(name:postimport\_parquet.activetime, type:double, comment:null), FieldSchema(name:postimport\_parquet.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_parquet.timeslot, type:timestamp, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411062828\_12abce73-04d3-4ffa-8b42-39808cd00f51); Time taken: 0.101 seconds  INFO : Executing command(queryId=hive\_20170411062828\_12abce73-04d3-4ffa-8b42-39808cd00f51): select \* from postimport\_parquet where username = '670063006100750074006F007500730065007200'  INFO : Query ID = hive\_20170411062828\_12abce73-04d3-4ffa-8b42-39808cd00f51  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:10  INFO : Submitting tokens for job: job\_1490094830773\_0372  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0372/  INFO : Starting Job = job\_1490094830773\_0372, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0372/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0372  INFO : Hadoop job information for Stage-1: number of mappers: 10; number of reducers: 0  INFO : 2017-04-11 06:28:19,157 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 06:28:29,479 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 18.91 sec  INFO : 2017-04-11 06:28:30,506 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 30.5 sec  INFO : 2017-04-11 06:28:32,594 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 44.28 sec  INFO : 2017-04-11 06:28:33,664 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 64.11 sec  INFO : 2017-04-11 06:28:35,715 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 78.95 sec  INFO : 2017-04-11 06:28:36,736 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 87.4 sec  INFO : 2017-04-11 06:28:37,765 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 92.7 sec  INFO : 2017-04-11 06:28:38,787 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 97.34 sec  INFO : 2017-04-11 06:28:39,837 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 98.58 sec  INFO : 2017-04-11 06:28:40,866 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 116.43 sec  INFO : 2017-04-11 06:28:43,967 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 123.43 sec  INFO : 2017-04-11 06:28:47,044 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 131.64 sec  INFO : 2017-04-11 06:28:49,117 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 133.25 sec  INFO : 2017-04-11 06:28:50,145 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 138.49 sec  INFO : 2017-04-11 06:28:51,170 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 139.88 sec  INFO : MapReduce Total cumulative CPU time: 2 minutes 19 seconds 880 msec  INFO : Ended Job = job\_1490094830773\_0372  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 10 Cumulative CPU: 139.88 sec HDFS Read: 2641466464 HDFS Write: 2062564 SUCCESS  INFO : Total MapReduce CPU Time Spent: 2 minutes 19 seconds 880 msec  INFO : Completed executing command(queryId=hive\_20170411062828\_12abce73-04d3-4ffa-8b42-39808cd00f51); Time taken: 40.857 seconds  INFO : OK |

# Avro File format default compression

|  |
| --- |
| CREATE TABLE postimport\_avro(  username string,  machinename string,  machineid string,  ipaddress string,  domainname string,  pcmlogeventid string,  modulename string,  filename string,  productname string,  companyname string,  description string,  text string,  activetime double,  date date)  STORED AS AVRO; |

***Note: AVRO does not support timestamp data type in hive.***

## Load data to Avro Table

In this case **hive.exec.compress.output** is true and **avro.output.codec** is not defined

|  |
| --- |
| insert into postimport\_avro select username,machinename,machineid,ipaddress,domainname,pcmlogeventid,modulename,filename,productname,companyname,description,text,activetime,date from postimport\_text; |

|  |
| --- |
| hive> insert into postimport\_avro select username,machinename,machineid,ipaddress,domainname,pcmlogeventid,modulename,filename,productname,companyname,description,text,activetime,date from postimport\_text;  Query ID = shalaj\_20170411063232\_c0b94667-ab0a-42fb-8e6c-9e7b4b3b2cf0  Total jobs = 3  Launching Job 1 out of 3  Number of reduce tasks is set to 0 since there's no reduce operator  Starting Job = job\_1490094830773\_0373, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0373/  Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0373  Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  2017-04-11 06:32:35,760 Stage-1 map = 0%, reduce = 0%  2017-04-11 06:32:50,390 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 73.39 sec  2017-04-11 06:32:55,753 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 92.59 sec  2017-04-11 06:32:59,602 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 104.47 sec  2017-04-11 06:33:01,994 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 112.11 sec  2017-04-11 06:33:05,850 Stage-1 map = 6%, reduce = 0%, Cumulative CPU 122.39 sec  2017-04-11 06:33:08,299 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 129.73 sec  2017-04-11 06:33:14,512 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 150.96 sec  2017-04-11 06:33:17,090 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 159.64 sec  2017-04-11 06:33:20,683 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 167.73 sec  2017-04-11 06:33:25,849 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 179.15 sec  2017-04-11 06:33:35,766 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 222.51 sec  2017-04-11 06:33:37,003 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 227.82 sec  2017-04-11 06:33:38,285 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 232.04 sec  2017-04-11 06:33:40,976 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 239.96 sec  2017-04-11 06:33:47,079 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 253.93 sec  2017-04-11 06:33:51,406 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 266.23 sec  2017-04-11 06:33:53,749 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 275.93 sec  2017-04-11 06:33:55,112 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 287.02 sec  2017-04-11 06:33:56,243 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 290.56 sec  2017-04-11 06:33:58,659 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 296.08 sec  2017-04-11 06:34:05,568 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 323.63 sec  2017-04-11 06:34:12,297 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 362.03 sec  2017-04-11 06:34:14,976 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 369.84 sec  2017-04-11 06:34:16,170 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 371.87 sec  2017-04-11 06:34:17,376 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 376.02 sec  2017-04-11 06:34:18,564 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 378.31 sec  2017-04-11 06:34:19,771 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 385.17 sec  2017-04-11 06:34:22,461 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 390.73 sec  2017-04-11 06:34:31,331 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 414.97 sec  2017-04-11 06:34:35,193 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 431.32 sec  2017-04-11 06:34:36,908 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 437.9 sec  2017-04-11 06:34:38,134 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 442.66 sec  2017-04-11 06:34:39,388 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 447.95 sec  2017-04-11 06:34:43,044 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 454.1 sec  2017-04-11 06:34:51,862 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 485.12 sec  2017-04-11 06:34:55,523 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 491.17 sec  2017-04-11 06:35:01,772 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 506.7 sec  2017-04-11 06:35:07,058 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 516.01 sec  2017-04-11 06:35:11,872 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 531.24 sec  2017-04-11 06:35:14,282 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 538.28 sec  2017-04-11 06:35:16,649 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 544.7 sec  2017-04-11 06:35:18,989 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 548.97 sec  2017-04-11 06:35:26,146 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 577.13 sec  2017-04-11 06:35:29,763 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 582.75 sec  2017-04-11 06:35:33,340 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 603.34 sec  2017-04-11 06:35:35,724 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 609.31 sec  2017-04-11 06:35:37,191 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 611.69 sec  2017-04-11 06:35:38,634 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 613.99 sec  2017-04-11 06:35:44,756 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 643.12 sec  2017-04-11 06:35:46,011 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 647.89 sec  2017-04-11 06:35:48,390 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 653.3 sec  2017-04-11 06:35:52,011 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 670.7 sec  2017-04-11 06:35:55,849 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 678.4 sec  2017-04-11 06:35:58,074 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 684.43 sec  2017-04-11 06:36:00,300 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 695.82 sec  2017-04-11 06:36:01,604 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 706.44 sec  2017-04-11 06:36:02,912 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 715.8 sec  2017-04-11 06:36:05,335 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 725.4 sec  2017-04-11 06:36:09,119 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 737.84 sec  2017-04-11 06:36:11,346 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 743.25 sec  2017-04-11 06:36:12,646 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 753.95 sec  2017-04-11 06:36:17,169 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 765.32 sec  2017-04-11 06:36:22,171 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 779.95 sec  2017-04-11 06:36:24,298 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 786.47 sec  2017-04-11 06:36:28,751 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 800.83 sec  2017-04-11 06:36:31,132 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 804.61 sec  2017-04-11 06:36:36,080 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 810.34 sec  2017-04-11 06:36:53,845 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 838.36 sec  2017-04-11 06:36:57,603 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 841.9 sec  2017-04-11 06:36:59,328 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 844.41 sec  2017-04-11 06:37:01,457 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 847.29 sec  2017-04-11 06:37:04,884 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 850.56 sec  2017-04-11 06:37:10,314 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 855.87 sec  2017-04-11 06:37:11,390 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 856.51 sec  2017-04-11 06:37:13,518 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 859.58 sec  2017-04-11 06:37:24,855 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 871.51 sec  2017-04-11 06:37:31,457 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 874.94 sec  2017-04-11 06:37:41,038 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 882.95 sec  2017-04-11 06:37:44,461 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 886.85 sec  2017-04-11 06:37:58,157 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 894.61 sec  MapReduce Total cumulative CPU time: 14 minutes 54 seconds 610 msec  Ended Job = job\_1490094830773\_0373  Stage-4 is selected by condition resolver.  Stage-3 is filtered out by condition resolver.  Stage-5 is filtered out by condition resolver.  Moving data to: hdfs://mac55:8020/user/hive/warehouse/esplus.db/postimport\_avro/.hive-staging\_hive\_2017-04-11\_06-32-23\_475\_2475310570530731429-1/-ext-10000  Loading data to table esplus.postimport\_avro  Table esplus.postimport\_avro stats: [numFiles=74, numRows=42286538, totalSize=4908000350, rawDataSize=0]  MapReduce Jobs Launched:  Stage-Stage-1: Map: 74 Cumulative CPU: 894.61 sec HDFS Read: 19486404012 HDFS Write: 4908005160 SUCCESS  Total MapReduce CPU Time Spent: 14 minutes 54 seconds 610 msec  OK  Time taken: 338.742 seconds |

## Data Size

|  |
| --- |
| hive> dfs -du -s -h /user/hive/warehouse/esplus.db/postimport\_avro;  4.6 G 13.7 G /user/hive/warehouse/esplus.db/postimport\_avro |

## Select some columns

SET hive.exec.compress.output=true;

|  |
| --- |
| select productname,username from postimport\_avro;  INFO : Compiling command(queryId=hive\_20170411063939\_2976909b-7497-48f2-b751-8de8666c4cff): select productname,username from postimport\_avro  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:productname, type:string, comment:null), FieldSchema(name:username, type:string, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411063939\_2976909b-7497-48f2-b751-8de8666c4cff); Time taken: 0.154 seconds  INFO : Executing command(queryId=hive\_20170411063939\_2976909b-7497-48f2-b751-8de8666c4cff): select productname,username from postimport\_avro  INFO : Query ID = hive\_20170411063939\_2976909b-7497-48f2-b751-8de8666c4cff  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:18  INFO : Submitting tokens for job: job\_1490094830773\_0374  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0374/  INFO : Starting Job = job\_1490094830773\_0374, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0374/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0374  INFO : Hadoop job information for Stage-1: number of mappers: 18; number of reducers: 0  INFO : 2017-04-11 06:39:38,665 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 06:39:55,383 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 138.85 sec  INFO : 2017-04-11 06:39:56,420 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 144.47 sec  INFO : 2017-04-11 06:39:57,462 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 173.36 sec  INFO : 2017-04-11 06:40:01,614 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 225.98 sec  INFO : 2017-04-11 06:40:03,692 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 259.27 sec  INFO : 2017-04-11 06:40:05,762 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 285.5 sec  INFO : 2017-04-11 06:40:06,813 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 302.35 sec  INFO : 2017-04-11 06:40:10,950 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 355.13 sec  INFO : 2017-04-11 06:40:11,979 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 377.72 sec  INFO : 2017-04-11 06:40:13,012 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 389.03 sec  INFO : 2017-04-11 06:40:15,103 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 421.04 sec  INFO : 2017-04-11 06:40:16,139 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 438.86 sec  INFO : 2017-04-11 06:40:18,215 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 464.49 sec  INFO : 2017-04-11 06:40:19,248 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 474.81 sec  INFO : 2017-04-11 06:40:20,273 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 484.66 sec  INFO : 2017-04-11 06:40:21,295 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 490.58 sec  INFO : 2017-04-11 06:40:22,318 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 491.34 sec  INFO : 2017-04-11 06:40:25,391 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 508.44 sec  INFO : 2017-04-11 06:40:30,545 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 524.78 sec  INFO : 2017-04-11 06:40:31,571 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 531.21 sec  INFO : 2017-04-11 06:40:34,646 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 540.51 sec  INFO : 2017-04-11 06:40:36,711 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 543.56 sec  INFO : 2017-04-11 06:40:37,734 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 549.6 sec  INFO : 2017-04-11 06:40:40,794 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 558.65 sec  INFO : 2017-04-11 06:40:43,895 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 567.94 sec  INFO : 2017-04-11 06:40:45,968 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 573.93 sec  INFO : 2017-04-11 06:40:50,079 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 587.55 sec  INFO : 2017-04-11 06:40:52,166 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 592.62 sec  INFO : 2017-04-11 06:40:53,204 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 593.35 sec  INFO : MapReduce Total cumulative CPU time: 9 minutes 53 seconds 350 msec  INFO : Ended Job = job\_1490094830773\_0374  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 18 Cumulative CPU: 593.35 sec HDFS Read: 4908681714 HDFS Write: 233693113 SUCCESS  INFO : Total MapReduce CPU Time Spent: 9 minutes 53 seconds 350 msec  INFO : Completed executing command(queryId=hive\_20170411063939\_2976909b-7497-48f2-b751-8de8666c4cff); Time taken: 81.219 seconds  INFO : OK |

## Check Max Date

|  |
| --- |
| select max(date) from postimport\_avro;  INFO : Compiling command(queryId=hive\_20170411064141\_de241bbb-7bc3-48d5-96e5-e99673e94e87): select max(date) from postimport\_avro  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:\_c0, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411064141\_de241bbb-7bc3-48d5-96e5-e99673e94e87); Time taken: 0.109 seconds  INFO : Executing command(queryId=hive\_20170411064141\_de241bbb-7bc3-48d5-96e5-e99673e94e87): select max(date) from postimport\_avro  INFO : Query ID = hive\_20170411064141\_de241bbb-7bc3-48d5-96e5-e99673e94e87  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks determined at compile time: 1  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:18  INFO : Submitting tokens for job: job\_1490094830773\_0375  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0375/  INFO : Starting Job = job\_1490094830773\_0375, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0375/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0375  INFO : Hadoop job information for Stage-1: number of mappers: 18; number of reducers: 1  INFO : 2017-04-11 06:41:56,416 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 06:42:11,113 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 98.76 sec  INFO : 2017-04-11 06:42:12,142 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 110.17 sec  INFO : 2017-04-11 06:42:13,171 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 130.42 sec  INFO : 2017-04-11 06:42:14,202 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 144.87 sec  INFO : 2017-04-11 06:42:15,238 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 153.46 sec  INFO : 2017-04-11 06:42:16,269 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 172.55 sec  INFO : 2017-04-11 06:42:17,305 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 190.1 sec  INFO : 2017-04-11 06:42:19,376 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 215.8 sec  INFO : 2017-04-11 06:42:20,408 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 236.14 sec  INFO : 2017-04-11 06:42:21,445 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 238.99 sec  INFO : 2017-04-11 06:42:22,478 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 259.37 sec  INFO : 2017-04-11 06:42:25,589 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 302.79 sec  INFO : 2017-04-11 06:42:26,623 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 325.78 sec  INFO : 2017-04-11 06:42:28,688 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 346.0 sec  INFO : 2017-04-11 06:42:29,715 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 372.07 sec  INFO : 2017-04-11 06:42:30,736 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 381.41 sec  INFO : 2017-04-11 06:42:31,756 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 386.95 sec  INFO : 2017-04-11 06:42:34,870 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 400.67 sec  INFO : 2017-04-11 06:42:40,046 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 412.85 sec  INFO : 2017-04-11 06:42:42,100 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 415.85 sec  INFO : 2017-04-11 06:42:46,242 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 438.0 sec  INFO : 2017-04-11 06:42:47,262 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 441.03 sec  INFO : 2017-04-11 06:42:52,372 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 456.47 sec  INFO : 2017-04-11 06:42:53,419 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 459.56 sec  INFO : 2017-04-11 06:42:54,446 Stage-1 map = 90%, reduce = 28%, Cumulative CPU 459.77 sec  INFO : 2017-04-11 06:42:57,532 Stage-1 map = 93%, reduce = 28%, Cumulative CPU 469.45 sec  INFO : 2017-04-11 06:42:58,553 Stage-1 map = 94%, reduce = 28%, Cumulative CPU 475.48 sec  INFO : 2017-04-11 06:43:00,599 Stage-1 map = 94%, reduce = 30%, Cumulative CPU 475.5 sec  INFO : 2017-04-11 06:43:03,677 Stage-1 map = 100%, reduce = 30%, Cumulative CPU 485.45 sec  INFO : 2017-04-11 06:43:05,725 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 486.25 sec  INFO : MapReduce Total cumulative CPU time: 8 minutes 6 seconds 250 msec  INFO : Ended Job = job\_1490094830773\_0375  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 18 Reduce: 1 Cumulative CPU: 486.25 sec HDFS Read: 4908693153 HDFS Write: 19 SUCCESS  INFO : Total MapReduce CPU Time Spent: 8 minutes 6 seconds 250 msec  INFO : Completed executing command(queryId=hive\_20170411064141\_de241bbb-7bc3-48d5-96e5-e99673e94e87); Time taken: 78.305 seconds  INFO : OK |

## User wise Max Date

|  |
| --- |
| select username,max(date) from postimport\_avro group by username;  INFO : Compiling command(queryId=hive\_20170411064343\_f2d1187f-2675-41cf-b02e-5cc48d8666fc): select username,max(date) from postimport\_avro group by username  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:username, type:string, comment:null), FieldSchema(name:\_c1, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411064343\_f2d1187f-2675-41cf-b02e-5cc48d8666fc); Time taken: 0.11 seconds  INFO : Executing command(queryId=hive\_20170411064343\_f2d1187f-2675-41cf-b02e-5cc48d8666fc): select username,max(date) from postimport\_avro group by username  INFO : Query ID = hive\_20170411064343\_f2d1187f-2675-41cf-b02e-5cc48d8666fc  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks not specified. Estimated from input data size: 74  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:18  INFO : Submitting tokens for job: job\_1490094830773\_0376  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0376/  INFO : Starting Job = job\_1490094830773\_0376, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0376/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0376  INFO : Hadoop job information for Stage-1: number of mappers: 18; number of reducers: 74  INFO : 2017-04-11 06:43:48,265 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 06:44:01,845 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 86.88 sec  INFO : 2017-04-11 06:44:02,909 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 97.56 sec  INFO : 2017-04-11 06:44:03,949 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 117.84 sec  INFO : 2017-04-11 06:44:06,018 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 149.8 sec  INFO : 2017-04-11 06:44:07,057 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 161.25 sec  INFO : 2017-04-11 06:44:08,091 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 176.93 sec  INFO : 2017-04-11 06:44:11,189 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 223.54 sec  INFO : 2017-04-11 06:44:12,238 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 235.27 sec  INFO : 2017-04-11 06:44:13,274 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 246.64 sec  INFO : 2017-04-11 06:44:14,311 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 267.19 sec  INFO : 2017-04-11 06:44:17,418 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 310.91 sec  INFO : 2017-04-11 06:44:18,452 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 322.45 sec  INFO : 2017-04-11 06:44:19,488 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 333.98 sec  INFO : 2017-04-11 06:44:20,524 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 354.64 sec  INFO : 2017-04-11 06:44:21,560 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 368.82 sec  INFO : 2017-04-11 06:44:23,635 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 398.88 sec  INFO : 2017-04-11 06:44:24,671 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 406.51 sec  INFO : 2017-04-11 06:44:25,702 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 410.02 sec  INFO : 2017-04-11 06:44:26,736 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 421.95 sec  INFO : 2017-04-11 06:44:27,763 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 423.74 sec  INFO : 2017-04-11 06:44:31,887 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 442.32 sec  INFO : 2017-04-11 06:44:33,944 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 455.37 sec  INFO : 2017-04-11 06:44:36,148 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 464.54 sec  INFO : 2017-04-11 06:44:39,309 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 473.64 sec  INFO : 2017-04-11 06:44:42,372 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 482.52 sec  INFO : 2017-04-11 06:44:45,444 Stage-1 map = 90%, reduce = 5%, Cumulative CPU 494.31 sec  INFO : 2017-04-11 06:44:48,513 Stage-1 map = 92%, reduce = 5%, Cumulative CPU 503.7 sec  INFO : 2017-04-11 06:44:52,606 Stage-1 map = 95%, reduce = 5%, Cumulative CPU 513.5 sec  INFO : 2017-04-11 06:44:53,639 Stage-1 map = 97%, reduce = 5%, Cumulative CPU 515.05 sec  INFO : 2017-04-11 06:44:54,676 Stage-1 map = 100%, reduce = 5%, Cumulative CPU 519.03 sec  INFO : 2017-04-11 06:44:55,696 Stage-1 map = 100%, reduce = 13%, Cumulative CPU 528.21 sec  INFO : 2017-04-11 06:44:56,738 Stage-1 map = 100%, reduce = 16%, Cumulative CPU 531.07 sec  INFO : 2017-04-11 06:45:00,939 Stage-1 map = 100%, reduce = 18%, Cumulative CPU 532.36 sec  INFO : 2017-04-11 06:45:03,015 Stage-1 map = 100%, reduce = 36%, Cumulative CPU 550.17 sec  INFO : 2017-04-11 06:45:07,179 Stage-1 map = 100%, reduce = 39%, Cumulative CPU 552.53 sec  INFO : 2017-04-11 06:45:08,219 Stage-1 map = 100%, reduce = 54%, Cumulative CPU 566.07 sec  INFO : 2017-04-11 06:45:09,254 Stage-1 map = 100%, reduce = 57%, Cumulative CPU 568.42 sec  INFO : 2017-04-11 06:45:11,365 Stage-1 map = 100%, reduce = 58%, Cumulative CPU 569.56 sec  INFO : 2017-04-11 06:45:12,402 Stage-1 map = 100%, reduce = 59%, Cumulative CPU 570.85 sec  INFO : 2017-04-11 06:45:13,431 Stage-1 map = 100%, reduce = 66%, Cumulative CPU 576.79 sec  INFO : 2017-04-11 06:45:14,470 Stage-1 map = 100%, reduce = 72%, Cumulative CPU 581.23 sec  INFO : 2017-04-11 06:45:15,491 Stage-1 map = 100%, reduce = 77%, Cumulative CPU 589.58 sec  INFO : 2017-04-11 06:45:16,541 Stage-1 map = 100%, reduce = 78%, Cumulative CPU 590.8 sec  INFO : 2017-04-11 06:45:17,590 Stage-1 map = 100%, reduce = 81%, Cumulative CPU 593.31 sec  INFO : 2017-04-11 06:45:18,637 Stage-1 map = 100%, reduce = 86%, Cumulative CPU 598.1 sec  INFO : 2017-04-11 06:45:19,671 Stage-1 map = 100%, reduce = 89%, Cumulative CPU 600.62 sec  INFO : 2017-04-11 06:45:20,724 Stage-1 map = 100%, reduce = 93%, Cumulative CPU 604.11 sec  INFO : 2017-04-11 06:45:21,749 Stage-1 map = 100%, reduce = 99%, Cumulative CPU 609.02 sec  INFO : 2017-04-11 06:45:22,771 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 610.19 sec  INFO : MapReduce Total cumulative CPU time: 10 minutes 10 seconds 190 msec  INFO : Ended Job = job\_1490094830773\_0376  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 18 Reduce: 74 Cumulative CPU: 610.19 sec HDFS Read: 4908972830 HDFS Write: 37574 SUCCESS  INFO : Total MapReduce CPU Time Spent: 10 minutes 10 seconds 190 msec  INFO : Completed executing command(queryId=hive\_20170411064343\_f2d1187f-2675-41cf-b02e-5cc48d8666fc); Time taken: 103.471 seconds  INFO : OK |

## Select All Columns

|  |
| --- |
| select \* from postimport\_avro  INFO : Compiling command(queryId=hive\_20170411064646\_6a8af1ee-4daa-48ae-91a5-62306eabb7f9): select \* from postimport\_avro  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_avro.username, type:string, comment:null), FieldSchema(name:postimport\_avro.machinename, type:string, comment:null), FieldSchema(name:postimport\_avro.machineid, type:string, comment:null), FieldSchema(name:postimport\_avro.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_avro.domainname, type:string, comment:null), FieldSchema(name:postimport\_avro.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_avro.modulename, type:string, comment:null), FieldSchema(name:postimport\_avro.filename, type:string, comment:null), FieldSchema(name:postimport\_avro.productname, type:string, comment:null), FieldSchema(name:postimport\_avro.companyname, type:string, comment:null), FieldSchema(name:postimport\_avro.description, type:string, comment:null), FieldSchema(name:postimport\_avro.text, type:string, comment:null), FieldSchema(name:postimport\_avro.activetime, type:double, comment:null), FieldSchema(name:postimport\_avro.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411064646\_6a8af1ee-4daa-48ae-91a5-62306eabb7f9); Time taken: 0.103 seconds  INFO : Executing command(queryId=hive\_20170411064646\_6a8af1ee-4daa-48ae-91a5-62306eabb7f9): select \* from postimport\_avro  INFO : Query ID = hive\_20170411064646\_6a8af1ee-4daa-48ae-91a5-62306eabb7f9  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:18  INFO : Submitting tokens for job: job\_1490094830773\_0377  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0377/  INFO : Starting Job = job\_1490094830773\_0377, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0377/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0377  INFO : Hadoop job information for Stage-1: number of mappers: 18; number of reducers: 0  INFO : 2017-04-11 06:46:08,881 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 06:46:35,742 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 208.38 sec  INFO : 2017-04-11 06:46:36,765 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 219.03 sec  INFO : 2017-04-11 06:46:40,861 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 248.4 sec  INFO : 2017-04-11 06:46:41,888 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 261.03 sec  INFO : 2017-04-11 06:46:50,085 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 335.57 sec  INFO : 2017-04-11 06:46:51,111 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 347.92 sec  INFO : 2017-04-11 06:46:52,133 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 354.97 sec  INFO : 2017-04-11 06:46:55,211 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 380.74 sec  INFO : 2017-04-11 06:46:56,232 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 389.29 sec  INFO : 2017-04-11 06:46:57,267 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 401.92 sec  INFO : 2017-04-11 06:46:58,299 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 407.4 sec  INFO : 2017-04-11 06:46:59,323 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 415.27 sec  INFO : 2017-04-11 06:47:00,355 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 429.36 sec  INFO : 2017-04-11 06:47:03,428 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 456.23 sec  INFO : 2017-04-11 06:47:05,473 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 468.93 sec  INFO : 2017-04-11 06:47:06,514 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 481.76 sec  INFO : 2017-04-11 06:47:17,755 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 558.68 sec  INFO : 2017-04-11 06:47:21,853 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 586.95 sec  INFO : 2017-04-11 06:47:26,988 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 616.68 sec  INFO : 2017-04-11 06:47:37,213 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 685.19 sec  INFO : 2017-04-11 06:47:39,256 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 695.5 sec  INFO : 2017-04-11 06:47:40,278 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 704.53 sec  INFO : 2017-04-11 06:47:43,342 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 723.35 sec  INFO : 2017-04-11 06:47:45,385 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 735.49 sec  INFO : 2017-04-11 06:47:47,424 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 743.41 sec  INFO : 2017-04-11 06:47:48,444 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 750.47 sec  INFO : 2017-04-11 06:47:52,556 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 773.28 sec  INFO : 2017-04-11 06:47:53,579 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 784.83 sec  INFO : 2017-04-11 06:47:54,618 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 788.75 sec  INFO : 2017-04-11 06:48:06,903 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 866.59 sec  INFO : 2017-04-11 06:48:07,932 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 870.82 sec  INFO : 2017-04-11 06:48:08,951 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 878.11 sec  INFO : 2017-04-11 06:48:15,103 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 914.68 sec  INFO : 2017-04-11 06:48:16,124 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 922.25 sec  INFO : 2017-04-11 06:48:22,258 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 957.59 sec  INFO : 2017-04-11 06:48:23,285 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 964.89 sec  INFO : 2017-04-11 06:48:25,347 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 974.91 sec  INFO : 2017-04-11 06:48:26,368 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 979.31 sec  INFO : 2017-04-11 06:48:27,392 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 985.09 sec  INFO : 2017-04-11 06:48:31,501 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 1015.56 sec  INFO : 2017-04-11 06:48:32,539 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 1021.17 sec  INFO : 2017-04-11 06:48:34,601 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 1036.99 sec  INFO : 2017-04-11 06:48:39,747 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 1082.08 sec  INFO : 2017-04-11 06:48:40,772 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 1091.15 sec  INFO : 2017-04-11 06:48:42,815 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 1106.88 sec  INFO : 2017-04-11 06:48:45,877 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 1132.74 sec  INFO : 2017-04-11 06:48:48,939 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 1154.13 sec  INFO : 2017-04-11 06:48:50,977 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 1165.47 sec  INFO : 2017-04-11 06:48:51,998 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 1172.34 sec  INFO : 2017-04-11 06:48:55,058 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 1190.19 sec  INFO : 2017-04-11 06:49:00,157 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 1205.63 sec  INFO : 2017-04-11 06:49:01,175 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 1209.82 sec  INFO : 2017-04-11 06:49:03,214 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 1214.81 sec  INFO : 2017-04-11 06:49:04,233 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 1218.85 sec  INFO : 2017-04-11 06:49:09,331 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 1228.06 sec  INFO : 2017-04-11 06:49:11,369 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 1232.01 sec  INFO : 2017-04-11 06:49:23,638 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1246.34 sec  INFO : MapReduce Total cumulative CPU time: 20 minutes 46 seconds 340 msec  INFO : Ended Job = job\_1490094830773\_0377  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 18 Cumulative CPU: 1246.34 sec HDFS Read: 4908696330 HDFS Write: 3166141692 SUCCESS  INFO : Total MapReduce CPU Time Spent: 20 minutes 46 seconds 340 msec  INFO : Completed executing command(queryId=hive\_20170411064646\_6a8af1ee-4daa-48ae-91a5-62306eabb7f9); Time taken: 200.149 seconds  INFO : OK |

## Select All Columns with filter Condition

|  |
| --- |
| select \* from postimport\_avro where username = '670063006100750074006F007500730065007200';  INFO : Compiling command(queryId=hive\_20170411064949\_b20d2107-e058-49cf-8a42-585093724f32): select \* from postimport\_avro where username = '670063006100750074006F007500730065007200'  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_avro.username, type:string, comment:null), FieldSchema(name:postimport\_avro.machinename, type:string, comment:null), FieldSchema(name:postimport\_avro.machineid, type:string, comment:null), FieldSchema(name:postimport\_avro.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_avro.domainname, type:string, comment:null), FieldSchema(name:postimport\_avro.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_avro.modulename, type:string, comment:null), FieldSchema(name:postimport\_avro.filename, type:string, comment:null), FieldSchema(name:postimport\_avro.productname, type:string, comment:null), FieldSchema(name:postimport\_avro.companyname, type:string, comment:null), FieldSchema(name:postimport\_avro.description, type:string, comment:null), FieldSchema(name:postimport\_avro.text, type:string, comment:null), FieldSchema(name:postimport\_avro.activetime, type:double, comment:null), FieldSchema(name:postimport\_avro.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411064949\_b20d2107-e058-49cf-8a42-585093724f32); Time taken: 0.099 seconds  INFO : Executing command(queryId=hive\_20170411064949\_b20d2107-e058-49cf-8a42-585093724f32): select \* from postimport\_avro where username = '670063006100750074006F007500730065007200'  INFO : Query ID = hive\_20170411064949\_b20d2107-e058-49cf-8a42-585093724f32  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:18  INFO : Submitting tokens for job: job\_1490094830773\_0378  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0378/  INFO : Starting Job = job\_1490094830773\_0378, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0378/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0378  INFO : Hadoop job information for Stage-1: number of mappers: 18; number of reducers: 0  INFO : 2017-04-11 06:50:00,249 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 06:50:14,858 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 95.38 sec  INFO : 2017-04-11 06:50:15,899 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 107.02 sec  INFO : 2017-04-11 06:50:16,937 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 118.33 sec  INFO : 2017-04-11 06:50:19,011 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 150.72 sec  INFO : 2017-04-11 06:50:20,042 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 162.05 sec  INFO : 2017-04-11 06:50:21,080 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 181.63 sec  INFO : 2017-04-11 06:50:22,113 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 193.4 sec  INFO : 2017-04-11 06:50:24,186 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 225.09 sec  INFO : 2017-04-11 06:50:25,225 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 236.78 sec  INFO : 2017-04-11 06:50:26,258 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 247.88 sec  INFO : 2017-04-11 06:50:27,298 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 268.62 sec  INFO : 2017-04-11 06:50:28,335 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 283.31 sec  INFO : 2017-04-11 06:50:30,394 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 312.83 sec  INFO : 2017-04-11 06:50:31,449 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 330.51 sec  INFO : 2017-04-11 06:50:32,489 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 335.91 sec  INFO : 2017-04-11 06:50:33,522 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 355.91 sec  INFO : 2017-04-11 06:50:34,557 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 377.31 sec  INFO : 2017-04-11 06:50:35,608 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 385.77 sec  INFO : 2017-04-11 06:50:37,656 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 402.3 sec  INFO : 2017-04-11 06:50:38,677 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 406.68 sec  INFO : 2017-04-11 06:50:39,702 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 409.85 sec  INFO : 2017-04-11 06:50:42,780 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 433.88 sec  INFO : 2017-04-11 06:50:45,844 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 443.37 sec  INFO : 2017-04-11 06:50:52,004 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 461.5 sec  INFO : 2017-04-11 06:50:58,196 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 480.09 sec  INFO : 2017-04-11 06:51:01,264 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 489.55 sec  INFO : 2017-04-11 06:51:02,313 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 490.81 sec  INFO : 2017-04-11 06:51:03,334 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 492.7 sec  INFO : MapReduce Total cumulative CPU time: 8 minutes 12 seconds 700 msec  INFO : Ended Job = job\_1490094830773\_0378  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 18 Cumulative CPU: 492.7 sec HDFS Read: 4908704718 HDFS Write: 1346686 SUCCESS  INFO : Total MapReduce CPU Time Spent: 8 minutes 12 seconds 700 msec  INFO : Completed executing command(queryId=hive\_20170411064949\_b20d2107-e058-49cf-8a42-585093724f32); Time taken: 69.397 seconds  INFO : OK |

# Avro File format snappy compression

|  |
| --- |
| CREATE TABLE postimport\_avro(  username string,  machinename string,  machineid string,  ipaddress string,  domainname string,  pcmlogeventid string,  modulename string,  filename string,  productname string,  companyname string,  description string,  text string,  activetime double,  date date)  STORED AS AVRO; |

***Note: AVRO does not support timestamp data type in hive.***

## Load data to Avro Table

|  |
| --- |
| SET hive.exec.compress.output=true;  SET avro.output.codec=snappy; |

|  |
| --- |
| hive> truncate table postimport\_avro |

|  |
| --- |
| insert into postimport\_avro select username,machinename,machineid,ipaddress,domainname,pcmlogeventid,modulename,filename,productname,companyname,description,text,activetime,date from postimport\_text; |

|  |
| --- |
| hive> insert into postimport\_avro select username,machinename,machineid,ipaddress,domainname,pcmlogeventid,modulename,filename,productname,companyname,description,text,activetime,date from postimport\_text;  Query ID = shalaj\_20170411070101\_20767b38-dc59-48ff-878d-c0a1cb849bf5  Total jobs = 3  Launching Job 1 out of 3  Number of reduce tasks is set to 0 since there's no reduce operator  Starting Job = job\_1490094830773\_0380, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0380/  Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0380  Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  2017-04-11 07:01:50,004 Stage-1 map = 0%, reduce = 0%  2017-04-11 07:02:04,439 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 73.47 sec  2017-04-11 07:02:10,409 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 94.09 sec  2017-04-11 07:02:16,520 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 113.64 sec  2017-04-11 07:02:17,709 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 116.5 sec  2017-04-11 07:02:18,884 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 122.02 sec  2017-04-11 07:02:20,057 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 123.59 sec  2017-04-11 07:02:31,113 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 164.26 sec  2017-04-11 07:02:32,314 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 168.0 sec  2017-04-11 07:02:34,698 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 178.28 sec  2017-04-11 07:02:37,042 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 181.99 sec  2017-04-11 07:02:40,657 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 186.69 sec  2017-04-11 07:02:45,518 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 203.34 sec  2017-04-11 07:02:47,884 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 210.54 sec  2017-04-11 07:02:49,111 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 213.74 sec  2017-04-11 07:02:56,466 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 236.96 sec  2017-04-11 07:03:02,389 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 250.06 sec  2017-04-11 07:03:05,914 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 261.64 sec  2017-04-11 07:03:09,668 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 271.59 sec  2017-04-11 07:03:14,386 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 282.64 sec  2017-04-11 07:03:16,748 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 286.83 sec  2017-04-11 07:03:19,349 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 291.95 sec  2017-04-11 07:03:21,708 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 295.35 sec  2017-04-11 07:03:24,137 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 304.83 sec  2017-04-11 07:03:27,663 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 314.55 sec  2017-04-11 07:03:31,013 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 325.74 sec  2017-04-11 07:03:33,224 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 327.26 sec  2017-04-11 07:03:42,047 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 371.17 sec  2017-04-11 07:03:43,215 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 375.2 sec  2017-04-11 07:03:45,550 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 378.05 sec  2017-04-11 07:03:51,636 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 396.96 sec  2017-04-11 07:03:55,358 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 406.14 sec  2017-04-11 07:03:57,694 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 410.22 sec  2017-04-11 07:04:01,237 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 427.7 sec  2017-04-11 07:04:02,403 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 433.28 sec  2017-04-11 07:04:04,733 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 441.15 sec  2017-04-11 07:04:07,092 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 449.48 sec  2017-04-11 07:04:10,604 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 455.16 sec  2017-04-11 07:04:16,736 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 473.97 sec  2017-04-11 07:04:18,985 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 482.14 sec  2017-04-11 07:04:22,911 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 490.17 sec  2017-04-11 07:04:32,372 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 521.91 sec  2017-04-11 07:04:36,156 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 539.01 sec  2017-04-11 07:04:37,334 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 543.36 sec  2017-04-11 07:04:42,264 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 550.11 sec  2017-04-11 07:04:46,057 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 559.98 sec  2017-04-11 07:04:48,417 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 566.78 sec  2017-04-11 07:04:53,122 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 585.33 sec  2017-04-11 07:04:55,485 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 589.54 sec  2017-04-11 07:04:58,088 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 596.27 sec  2017-04-11 07:05:01,952 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 608.23 sec  2017-04-11 07:05:03,162 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 609.55 sec  2017-04-11 07:05:08,107 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 626.11 sec  2017-04-11 07:05:13,947 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 633.93 sec  2017-04-11 07:05:19,808 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 658.72 sec  2017-04-11 07:05:23,353 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 668.94 sec  2017-04-11 07:05:26,072 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 674.7 sec  2017-04-11 07:05:27,267 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 677.86 sec  2017-04-11 07:05:35,865 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 696.99 sec  2017-04-11 07:05:39,704 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 701.56 sec  2017-04-11 07:05:50,740 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 723.82 sec  2017-04-11 07:05:54,541 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 729.59 sec  2017-04-11 07:05:57,936 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 741.4 sec  2017-04-11 07:06:00,075 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 746.0 sec  2017-04-11 07:06:02,201 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 759.37 sec  2017-04-11 07:06:08,936 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 764.53 sec  2017-04-11 07:06:17,255 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 779.85 sec  2017-04-11 07:06:22,270 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 796.87 sec  2017-04-11 07:06:30,362 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 812.04 sec  2017-04-11 07:06:31,430 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 814.72 sec  2017-04-11 07:06:34,274 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 823.18 sec  2017-04-11 07:06:36,406 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 825.78 sec  2017-04-11 07:06:40,278 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 832.94 sec  2017-04-11 07:06:43,993 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 837.69 sec  2017-04-11 07:06:49,792 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 842.86 sec  2017-04-11 07:06:55,339 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 847.77 sec  2017-04-11 07:07:01,425 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 852.6 sec  2017-04-11 07:07:07,282 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 856.84 sec  2017-04-11 07:07:16,269 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 863.06 sec  2017-04-11 07:07:19,455 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 864.63 sec  2017-04-11 07:07:22,740 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 868.34 sec  2017-04-11 07:07:32,498 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 873.33 sec  MapReduce Total cumulative CPU time: 14 minutes 33 seconds 330 msec  Ended Job = job\_1490094830773\_0380  Stage-4 is selected by condition resolver.  Stage-3 is filtered out by condition resolver.  Stage-5 is filtered out by condition resolver.  Moving data to: hdfs://mac55:8020/user/hive/warehouse/esplus.db/postimport\_avro/.hive-staging\_hive\_2017-04-11\_07-01-40\_613\_2104146198195383143-1/-ext-10000  Loading data to table esplus.postimport\_avro  Table esplus.postimport\_avro stats: [numFiles=74, numRows=42286538, totalSize=6665613855, rawDataSize=0]  MapReduce Jobs Launched:  Stage-Stage-1: Map: 74 Cumulative CPU: 873.33 sec HDFS Read: 19486404752 HDFS Write: 6665618665 SUCCESS  Total MapReduce CPU Time Spent: 14 minutes 33 seconds 330 msec  OK  Time taken: 361.001 seconds |

## Data Size

|  |
| --- |
| hive> dfs -du -s -h /user/hive/warehouse/esplus.db/postimport\_avro;  6.2 G 18.6 G /user/hive/warehouse/esplus.db/postimport\_avro |

## Select some columns

SET hive.exec.compress.output=true;

|  |
| --- |
| select productname,username from postimport\_avro;  INFO : Compiling command(queryId=hive\_20170411070808\_b26eab9e-f653-4d52-bdb8-fa547dc594ae): select productname,username from postimport\_avro  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:productname, type:string, comment:null), FieldSchema(name:username, type:string, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411070808\_b26eab9e-f653-4d52-bdb8-fa547dc594ae); Time taken: 0.106 seconds  INFO : Executing command(queryId=hive\_20170411070808\_b26eab9e-f653-4d52-bdb8-fa547dc594ae): select productname,username from postimport\_avro  INFO : Query ID = hive\_20170411070808\_b26eab9e-f653-4d52-bdb8-fa547dc594ae  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:24  INFO : Submitting tokens for job: job\_1490094830773\_0381  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0381/  INFO : Starting Job = job\_1490094830773\_0381, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0381/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0381  INFO : Hadoop job information for Stage-1: number of mappers: 24; number of reducers: 0  INFO : 2017-04-11 07:08:47,685 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 07:09:01,810 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 94.61 sec  INFO : 2017-04-11 07:09:05,933 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 165.08 sec  INFO : 2017-04-11 07:09:06,965 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 170.71 sec  INFO : 2017-04-11 07:09:07,997 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 182.77 sec  INFO : 2017-04-11 07:09:11,092 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 225.25 sec  INFO : 2017-04-11 07:09:12,137 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 256.37 sec  INFO : 2017-04-11 07:09:15,233 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 302.22 sec  INFO : 2017-04-11 07:09:16,263 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 305.24 sec  INFO : 2017-04-11 07:09:17,293 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 310.35 sec  INFO : 2017-04-11 07:09:18,325 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 339.58 sec  INFO : 2017-04-11 07:09:19,349 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 346.49 sec  INFO : 2017-04-11 07:09:20,390 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 348.55 sec  INFO : 2017-04-11 07:09:21,422 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 349.69 sec  INFO : 2017-04-11 07:09:27,595 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 394.15 sec  INFO : 2017-04-11 07:09:28,624 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 419.29 sec  INFO : 2017-04-11 07:09:29,649 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 422.47 sec  INFO : 2017-04-11 07:09:30,679 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 435.34 sec  INFO : 2017-04-11 07:09:31,703 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 450.99 sec  INFO : 2017-04-11 07:09:33,754 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 466.41 sec  INFO : 2017-04-11 07:09:34,780 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 478.01 sec  INFO : 2017-04-11 07:09:36,830 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 492.88 sec  INFO : 2017-04-11 07:09:37,860 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 505.11 sec  INFO : 2017-04-11 07:09:39,909 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 520.31 sec  INFO : 2017-04-11 07:09:40,931 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 532.69 sec  INFO : 2017-04-11 07:09:42,978 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 548.38 sec  INFO : 2017-04-11 07:09:45,019 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 552.62 sec  INFO : MapReduce Total cumulative CPU time: 9 minutes 12 seconds 620 msec  INFO : Ended Job = job\_1490094830773\_0381  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 24 Cumulative CPU: 552.62 sec HDFS Read: 6666317977 HDFS Write: 233695598 SUCCESS  INFO : Total MapReduce CPU Time Spent: 9 minutes 12 seconds 620 msec  INFO : Completed executing command(queryId=hive\_20170411070808\_b26eab9e-f653-4d52-bdb8-fa547dc594ae); Time taken: 65.855 seconds  INFO : OK |

## Check Max Date

|  |
| --- |
| select max(date) from postimport\_avro;  INFO : Compiling command(queryId=hive\_20170411071212\_819af152-08cb-4220-bba5-d75faeae2341): select max(date) from postimport\_avro  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:\_c0, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411071212\_819af152-08cb-4220-bba5-d75faeae2341); Time taken: 0.103 seconds  INFO : Executing command(queryId=hive\_20170411071212\_819af152-08cb-4220-bba5-d75faeae2341): select max(date) from postimport\_avro  INFO : Query ID = hive\_20170411071212\_819af152-08cb-4220-bba5-d75faeae2341  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks determined at compile time: 1  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:24  INFO : Submitting tokens for job: job\_1490094830773\_0382  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0382/  INFO : Starting Job = job\_1490094830773\_0382, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0382/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0382  INFO : Hadoop job information for Stage-1: number of mappers: 24; number of reducers: 1  INFO : 2017-04-11 07:12:35,371 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 07:12:52,010 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 118.96 sec  INFO : 2017-04-11 07:12:54,337 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 129.79 sec  INFO : 2017-04-11 07:12:55,363 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 170.35 sec  INFO : 2017-04-11 07:12:58,949 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 212.03 sec  INFO : 2017-04-11 07:13:01,306 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 250.51 sec  INFO : 2017-04-11 07:13:02,335 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 253.54 sec  INFO : 2017-04-11 07:13:03,361 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 260.82 sec  INFO : 2017-04-11 07:13:04,410 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 271.6 sec  INFO : 2017-04-11 07:13:06,517 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 278.85 sec  INFO : 2017-04-11 07:13:07,562 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 281.87 sec  INFO : 2017-04-11 07:13:15,866 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 332.41 sec  INFO : 2017-04-11 07:13:16,898 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 352.25 sec  INFO : 2017-04-11 07:13:17,929 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 358.1 sec  INFO : 2017-04-11 07:13:21,024 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 386.32 sec  INFO : 2017-04-11 07:13:22,050 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 389.28 sec  INFO : 2017-04-11 07:13:23,081 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 403.42 sec  INFO : 2017-04-11 07:13:24,109 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 412.04 sec  INFO : 2017-04-11 07:13:25,144 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 415.65 sec  INFO : 2017-04-11 07:13:26,178 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 430.83 sec  INFO : 2017-04-11 07:13:27,211 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 437.25 sec  INFO : 2017-04-11 07:13:28,230 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 443.39 sec  INFO : 2017-04-11 07:13:30,293 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 448.18 sec  INFO : 2017-04-11 07:13:32,338 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 449.16 sec  INFO : MapReduce Total cumulative CPU time: 7 minutes 29 seconds 160 msec  INFO : Ended Job = job\_1490094830773\_0382  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 24 Reduce: 1 Cumulative CPU: 449.16 sec HDFS Read: 6666332048 HDFS Write: 19 SUCCESS  INFO : Total MapReduce CPU Time Spent: 7 minutes 29 seconds 160 msec  INFO : Completed executing command(queryId=hive\_20170411071212\_819af152-08cb-4220-bba5-d75faeae2341); Time taken: 64.0 seconds  INFO : OK |

## User wise Max Date

|  |
| --- |
| select username,max(date) from postimport\_avro group by username;  INFO : Compiling command(queryId=hive\_20170411071313\_f90e5887-4f46-46f8-85dc-92f3bca5da1b): select username,max(date) from postimport\_avro group by username  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:username, type:string, comment:null), FieldSchema(name:\_c1, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411071313\_f90e5887-4f46-46f8-85dc-92f3bca5da1b); Time taken: 0.098 seconds  INFO : Executing command(queryId=hive\_20170411071313\_f90e5887-4f46-46f8-85dc-92f3bca5da1b): select username,max(date) from postimport\_avro group by username  INFO : Query ID = hive\_20170411071313\_f90e5887-4f46-46f8-85dc-92f3bca5da1b  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks not specified. Estimated from input data size: 100  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:24  INFO : Submitting tokens for job: job\_1490094830773\_0383  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0383/  INFO : Starting Job = job\_1490094830773\_0383, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0383/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0383  INFO : Hadoop job information for Stage-1: number of mappers: 24; number of reducers: 100  INFO : 2017-04-11 07:14:00,726 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 07:14:15,655 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 72.75 sec  INFO : 2017-04-11 07:14:16,829 Stage-1 map = 6%, reduce = 0%, Cumulative CPU 97.77 sec  INFO : 2017-04-11 07:14:18,096 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 112.25 sec  INFO : 2017-04-11 07:14:19,134 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 137.46 sec  INFO : 2017-04-11 07:14:20,438 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 146.32 sec  INFO : 2017-04-11 07:14:21,471 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 156.25 sec  INFO : 2017-04-11 07:14:22,678 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 180.99 sec  INFO : 2017-04-11 07:14:23,719 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 183.61 sec  INFO : 2017-04-11 07:14:24,921 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 209.51 sec  INFO : 2017-04-11 07:14:25,966 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 223.01 sec  INFO : 2017-04-11 07:14:27,175 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 239.26 sec  INFO : 2017-04-11 07:14:28,203 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 251.64 sec  INFO : 2017-04-11 07:14:29,339 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 263.7 sec  INFO : 2017-04-11 07:14:31,625 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 273.54 sec  INFO : 2017-04-11 07:14:32,658 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 278.19 sec  INFO : 2017-04-11 07:14:33,820 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 283.49 sec  INFO : 2017-04-11 07:14:36,013 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 294.89 sec  INFO : 2017-04-11 07:14:37,155 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 320.42 sec  INFO : 2017-04-11 07:14:38,185 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 325.45 sec  INFO : 2017-04-11 07:14:39,330 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 361.68 sec  INFO : 2017-04-11 07:14:40,362 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 367.71 sec  INFO : 2017-04-11 07:14:42,552 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 392.02 sec  INFO : 2017-04-11 07:14:43,700 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 395.01 sec  INFO : 2017-04-11 07:14:44,728 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 397.97 sec  INFO : 2017-04-11 07:14:45,920 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 421.28 sec  INFO : 2017-04-11 07:14:48,090 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 438.49 sec  INFO : 2017-04-11 07:14:49,130 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 448.03 sec  INFO : 2017-04-11 07:14:50,288 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 449.9 sec  INFO : 2017-04-11 07:14:51,316 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 464.55 sec  INFO : 2017-04-11 07:14:52,466 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 469.98 sec  INFO : 2017-04-11 07:14:53,486 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 471.65 sec  INFO : 2017-04-11 07:14:58,325 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 476.61 sec  INFO : 2017-04-11 07:14:59,662 Stage-1 map = 100%, reduce = 13%, Cumulative CPU 492.93 sec  INFO : 2017-04-11 07:15:00,839 Stage-1 map = 100%, reduce = 14%, Cumulative CPU 494.32 sec  INFO : 2017-04-11 07:15:03,298 Stage-1 map = 100%, reduce = 16%, Cumulative CPU 496.66 sec  INFO : 2017-04-11 07:15:04,339 Stage-1 map = 100%, reduce = 17%, Cumulative CPU 497.8 sec  INFO : 2017-04-11 07:15:05,534 Stage-1 map = 100%, reduce = 20%, Cumulative CPU 501.73 sec  INFO : 2017-04-11 07:15:06,563 Stage-1 map = 100%, reduce = 28%, Cumulative CPU 512.15 sec  INFO : 2017-04-11 07:15:07,700 Stage-1 map = 100%, reduce = 31%, Cumulative CPU 515.78 sec  INFO : 2017-04-11 07:15:10,216 Stage-1 map = 100%, reduce = 33%, Cumulative CPU 517.94 sec  INFO : 2017-04-11 07:15:11,252 Stage-1 map = 100%, reduce = 36%, Cumulative CPU 521.75 sec  INFO : 2017-04-11 07:15:12,532 Stage-1 map = 100%, reduce = 38%, Cumulative CPU 524.12 sec  INFO : 2017-04-11 07:15:13,561 Stage-1 map = 100%, reduce = 45%, Cumulative CPU 532.73 sec  INFO : 2017-04-11 07:15:14,782 Stage-1 map = 100%, reduce = 47%, Cumulative CPU 535.15 sec  INFO : 2017-04-11 07:15:15,833 Stage-1 map = 100%, reduce = 49%, Cumulative CPU 537.7 sec  INFO : 2017-04-11 07:15:17,249 Stage-1 map = 100%, reduce = 52%, Cumulative CPU 541.21 sec  INFO : 2017-04-11 07:15:18,288 Stage-1 map = 100%, reduce = 53%, Cumulative CPU 542.28 sec  INFO : 2017-04-11 07:15:19,521 Stage-1 map = 100%, reduce = 59%, Cumulative CPU 549.49 sec  INFO : 2017-04-11 07:15:20,558 Stage-1 map = 100%, reduce = 62%, Cumulative CPU 553.45 sec  INFO : 2017-04-11 07:15:21,985 Stage-1 map = 100%, reduce = 65%, Cumulative CPU 557.35 sec  INFO : 2017-04-11 07:15:23,057 Stage-1 map = 100%, reduce = 67%, Cumulative CPU 559.72 sec  INFO : 2017-04-11 07:15:24,392 Stage-1 map = 100%, reduce = 69%, Cumulative CPU 561.98 sec  INFO : 2017-04-11 07:15:25,573 Stage-1 map = 100%, reduce = 74%, Cumulative CPU 568.47 sec  INFO : 2017-04-11 07:15:26,603 Stage-1 map = 100%, reduce = 78%, Cumulative CPU 574.21 sec  INFO : 2017-04-11 07:15:27,765 Stage-1 map = 100%, reduce = 80%, Cumulative CPU 576.65 sec  INFO : 2017-04-11 07:15:28,805 Stage-1 map = 100%, reduce = 83%, Cumulative CPU 580.03 sec  INFO : 2017-04-11 07:15:30,293 Stage-1 map = 100%, reduce = 84%, Cumulative CPU 581.14 sec  INFO : 2017-04-11 07:15:31,333 Stage-1 map = 100%, reduce = 89%, Cumulative CPU 587.45 sec  INFO : 2017-04-11 07:15:32,572 Stage-1 map = 100%, reduce = 93%, Cumulative CPU 592.4 sec  INFO : 2017-04-11 07:15:33,591 Stage-1 map = 100%, reduce = 97%, Cumulative CPU 597.17 sec  INFO : 2017-04-11 07:15:34,734 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 600.45 sec  INFO : MapReduce Total cumulative CPU time: 10 minutes 0 seconds 450 msec  INFO : Ended Job = job\_1490094830773\_0383  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 24 Reduce: 100 Cumulative CPU: 600.45 sec HDFS Read: 6666711185 HDFS Write: 39950 SUCCESS  INFO : Total MapReduce CPU Time Spent: 10 minutes 0 seconds 450 msec  INFO : Completed executing command(queryId=hive\_20170411071313\_f90e5887-4f46-46f8-85dc-92f3bca5da1b); Time taken: 101.256 seconds  INFO : OK |

## Select All Columns

|  |
| --- |
| select \* from postimport\_avro  INFO : Compiling command(queryId=hive\_20170411071515\_ab231565-fac6-4106-89be-6d8cc6b65bf1): select \* from postimport\_avro  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_avro.username, type:string, comment:null), FieldSchema(name:postimport\_avro.machinename, type:string, comment:null), FieldSchema(name:postimport\_avro.machineid, type:string, comment:null), FieldSchema(name:postimport\_avro.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_avro.domainname, type:string, comment:null), FieldSchema(name:postimport\_avro.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_avro.modulename, type:string, comment:null), FieldSchema(name:postimport\_avro.filename, type:string, comment:null), FieldSchema(name:postimport\_avro.productname, type:string, comment:null), FieldSchema(name:postimport\_avro.companyname, type:string, comment:null), FieldSchema(name:postimport\_avro.description, type:string, comment:null), FieldSchema(name:postimport\_avro.text, type:string, comment:null), FieldSchema(name:postimport\_avro.activetime, type:double, comment:null), FieldSchema(name:postimport\_avro.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411071515\_ab231565-fac6-4106-89be-6d8cc6b65bf1); Time taken: 0.133 seconds  INFO : Executing command(queryId=hive\_20170411071515\_ab231565-fac6-4106-89be-6d8cc6b65bf1): select \* from postimport\_avro  INFO : Query ID = hive\_20170411071515\_ab231565-fac6-4106-89be-6d8cc6b65bf1  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:24  INFO : Submitting tokens for job: job\_1490094830773\_0384  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0384/  INFO : Starting Job = job\_1490094830773\_0384, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0384/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0384  INFO : Hadoop job information for Stage-1: number of mappers: 24; number of reducers: 0  INFO : 2017-04-11 07:16:02,873 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 07:16:26,667 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 188.68 sec  INFO : 2017-04-11 07:16:28,846 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 200.26 sec  INFO : 2017-04-11 07:16:29,879 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 215.36 sec  INFO : 2017-04-11 07:16:31,030 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 221.96 sec  INFO : 2017-04-11 07:16:33,207 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 241.05 sec  INFO : 2017-04-11 07:16:41,943 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 306.58 sec  INFO : 2017-04-11 07:16:42,964 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 312.88 sec  INFO : 2017-04-11 07:16:44,116 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 318.11 sec  INFO : 2017-04-11 07:16:46,289 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 342.45 sec  INFO : 2017-04-11 07:16:48,471 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 361.07 sec  INFO : 2017-04-11 07:16:51,681 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 386.09 sec  INFO : 2017-04-11 07:16:53,855 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 402.13 sec  INFO : 2017-04-11 07:16:55,011 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 410.46 sec  INFO : 2017-04-11 07:16:56,036 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 417.95 sec  INFO : 2017-04-11 07:16:57,188 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 433.66 sec  INFO : 2017-04-11 07:17:00,379 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 454.33 sec  INFO : 2017-04-11 07:17:03,702 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 471.56 sec  INFO : 2017-04-11 07:17:08,029 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 503.38 sec  INFO : 2017-04-11 07:17:17,727 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 568.59 sec  INFO : 2017-04-11 07:17:18,924 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 577.79 sec  INFO : 2017-04-11 07:17:25,462 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 615.33 sec  INFO : 2017-04-11 07:17:27,650 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 627.34 sec  INFO : 2017-04-11 07:17:29,829 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 642.76 sec  INFO : 2017-04-11 07:17:36,390 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 690.06 sec  INFO : 2017-04-11 07:17:39,602 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 715.85 sec  INFO : 2017-04-11 07:17:41,771 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 726.93 sec  INFO : 2017-04-11 07:17:45,088 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 746.24 sec  INFO : 2017-04-11 07:17:46,112 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 752.62 sec  INFO : 2017-04-11 07:17:51,644 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 779.03 sec  INFO : 2017-04-11 07:17:52,667 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 789.16 sec  INFO : 2017-04-11 07:17:54,830 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 799.58 sec  INFO : 2017-04-11 07:17:56,995 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 811.25 sec  INFO : 2017-04-11 07:18:00,325 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 827.74 sec  INFO : 2017-04-11 07:18:01,350 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 833.66 sec  INFO : 2017-04-11 07:18:14,344 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 895.12 sec  INFO : 2017-04-11 07:18:16,507 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 905.47 sec  INFO : 2017-04-11 07:18:30,630 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 970.33 sec  INFO : 2017-04-11 07:18:32,794 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 978.98 sec  INFO : 2017-04-11 07:18:33,817 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 987.63 sec  INFO : 2017-04-11 07:18:37,116 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 1001.76 sec  INFO : 2017-04-11 07:18:38,138 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 1003.42 sec  INFO : 2017-04-11 07:18:39,274 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 1009.14 sec  INFO : 2017-04-11 07:18:42,442 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 1021.21 sec  INFO : 2017-04-11 07:18:45,738 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 1033.55 sec  INFO : 2017-04-11 07:18:50,051 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 1049.87 sec  INFO : 2017-04-11 07:18:52,205 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 1059.38 sec  INFO : 2017-04-11 07:18:55,383 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 1066.93 sec  INFO : 2017-04-11 07:18:57,536 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 1075.79 sec  INFO : 2017-04-11 07:19:07,270 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 1107.44 sec  INFO : 2017-04-11 07:19:08,293 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 1110.28 sec  INFO : 2017-04-11 07:19:09,424 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 1113.61 sec  INFO : 2017-04-11 07:19:13,725 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 1127.85 sec  INFO : 2017-04-11 07:19:19,046 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 1149.31 sec  INFO : 2017-04-11 07:19:24,477 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 1164.2 sec  INFO : 2017-04-11 07:19:28,777 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 1175.88 sec  INFO : 2017-04-11 07:19:37,384 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 1204.63 sec  INFO : 2017-04-11 07:19:39,566 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 1206.81 sec  INFO : 2017-04-11 07:19:42,775 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1212.07 sec  INFO : MapReduce Total cumulative CPU time: 20 minutes 12 seconds 70 msec  INFO : Ended Job = job\_1490094830773\_0384  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 24 Cumulative CPU: 1212.07 sec HDFS Read: 6666337465 HDFS Write: 3166142895 SUCCESS  INFO : Total MapReduce CPU Time Spent: 20 minutes 12 seconds 70 msec  INFO : Completed executing command(queryId=hive\_20170411071515\_ab231565-fac6-4106-89be-6d8cc6b65bf1); Time taken: 230.673 seconds  INFO : OK |

## Select All Columns with filter Condition

|  |
| --- |
| select \* from postimport\_avro where username = '670063006100750074006F007500730065007200';  INFO : Compiling command(queryId=hive\_20170411072020\_b7445560-7447-4c6e-a944-0f3d511c0aef): select \* from postimport\_avro where username = '670063006100750074006F007500730065007200'  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_avro.username, type:string, comment:null), FieldSchema(name:postimport\_avro.machinename, type:string, comment:null), FieldSchema(name:postimport\_avro.machineid, type:string, comment:null), FieldSchema(name:postimport\_avro.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_avro.domainname, type:string, comment:null), FieldSchema(name:postimport\_avro.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_avro.modulename, type:string, comment:null), FieldSchema(name:postimport\_avro.filename, type:string, comment:null), FieldSchema(name:postimport\_avro.productname, type:string, comment:null), FieldSchema(name:postimport\_avro.companyname, type:string, comment:null), FieldSchema(name:postimport\_avro.description, type:string, comment:null), FieldSchema(name:postimport\_avro.text, type:string, comment:null), FieldSchema(name:postimport\_avro.activetime, type:double, comment:null), FieldSchema(name:postimport\_avro.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411072020\_b7445560-7447-4c6e-a944-0f3d511c0aef); Time taken: 0.095 seconds  INFO : Executing command(queryId=hive\_20170411072020\_b7445560-7447-4c6e-a944-0f3d511c0aef): select \* from postimport\_avro where username = '670063006100750074006F007500730065007200'  INFO : Query ID = hive\_20170411072020\_b7445560-7447-4c6e-a944-0f3d511c0aef  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:24  INFO : Submitting tokens for job: job\_1490094830773\_0385  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0385/  INFO : Starting Job = job\_1490094830773\_0385, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0385/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0385  INFO : Hadoop job information for Stage-1: number of mappers: 24; number of reducers: 0  INFO : 2017-04-11 07:20:35,578 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 07:20:49,678 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 96.03 sec  INFO : 2017-04-11 07:20:50,729 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 110.29 sec  INFO : 2017-04-11 07:20:51,999 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 118.42 sec  INFO : 2017-04-11 07:20:53,027 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 138.85 sec  INFO : 2017-04-11 07:20:56,537 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 192.94 sec  INFO : 2017-04-11 07:20:57,581 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 204.12 sec  INFO : 2017-04-11 07:20:58,779 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 223.5 sec  INFO : 2017-04-11 07:21:00,079 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 241.73 sec  INFO : 2017-04-11 07:21:01,113 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 252.5 sec  INFO : 2017-04-11 07:21:02,308 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 268.85 sec  INFO : 2017-04-11 07:21:03,338 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 276.33 sec  INFO : 2017-04-11 07:21:04,609 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 286.4 sec  INFO : 2017-04-11 07:21:12,370 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 321.78 sec  INFO : 2017-04-11 07:21:13,530 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 340.09 sec  INFO : 2017-04-11 07:21:14,555 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 355.07 sec  INFO : 2017-04-11 07:21:16,758 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 379.28 sec  INFO : 2017-04-11 07:21:18,970 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 388.25 sec  INFO : 2017-04-11 07:21:20,132 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 405.8 sec  INFO : 2017-04-11 07:21:21,158 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 413.57 sec  INFO : 2017-04-11 07:21:23,351 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 434.39 sec  INFO : 2017-04-11 07:21:24,509 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 434.92 sec  INFO : 2017-04-11 07:21:25,532 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 442.56 sec  INFO : 2017-04-11 07:21:26,665 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 452.46 sec  INFO : 2017-04-11 07:21:27,684 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 454.74 sec  INFO : MapReduce Total cumulative CPU time: 7 minutes 34 seconds 740 msec  INFO : Ended Job = job\_1490094830773\_0385  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 24 Cumulative CPU: 454.74 sec HDFS Read: 6666348649 HDFS Write: 1352027 SUCCESS  INFO : Total MapReduce CPU Time Spent: 7 minutes 34 seconds 740 msec  INFO : Completed executing command(queryId=hive\_20170411072020\_b7445560-7447-4c6e-a944-0f3d511c0aef); Time taken: 59.46 seconds  INFO : OK |

# Sequence File format Snappy Compression

SequenceFile is a flat file format consisting of binary key/value pairs. It is extensively used in Hadoop MapReduce as input/output formats, since it is splittable

|  |
| --- |
| create table postimport\_sequence like postimport\_text stored as sequencefile; |

|  |
| --- |
| SET hive.exec.compress.output=true;  SET mapred.output.compression.type=BLOCK;  SET mapred.output.compression.codec=org.apache.hadoop.io.compress.SnappyCodec;  SET io.seqfile.compression.type=BLOCK; |

## Load data to Sequence Table

|  |
| --- |
| hive> insert overwrite table postimport\_sequence select \* from postimport\_text;  Query ID = shalaj\_20170411072323\_10dd4837-4c1f-445f-bf8b-d2920100fe94  Total jobs = 3  Launching Job 1 out of 3  Number of reduce tasks is set to 0 since there's no reduce operator  Starting Job = job\_1490094830773\_0386, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0386/  Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0386  Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  2017-04-11 07:23:50,393 Stage-1 map = 0%, reduce = 0%  2017-04-11 07:24:05,271 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 63.49 sec  2017-04-11 07:24:11,340 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 75.24 sec  2017-04-11 07:24:12,559 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 76.65 sec  2017-04-11 07:24:17,850 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 84.59 sec  2017-04-11 07:24:22,693 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 94.29 sec  2017-04-11 07:24:25,086 Stage-1 map = 6%, reduce = 0%, Cumulative CPU 105.26 sec  2017-04-11 07:24:28,726 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 114.02 sec  2017-04-11 07:24:29,922 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 117.39 sec  2017-04-11 07:24:31,117 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 120.28 sec  2017-04-11 07:24:33,516 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 125.17 sec  2017-04-11 07:24:34,709 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 127.72 sec  2017-04-11 07:24:38,280 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 135.17 sec  2017-04-11 07:24:42,170 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 141.38 sec  2017-04-11 07:24:49,873 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 155.41 sec  2017-04-11 07:24:52,220 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 157.1 sec  2017-04-11 07:24:58,271 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 162.27 sec  2017-04-11 07:24:59,531 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 163.97 sec  2017-04-11 07:25:05,810 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 172.33 sec  2017-04-11 07:25:24,563 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 204.26 sec  2017-04-11 07:25:26,991 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 208.92 sec  2017-04-11 07:25:30,854 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 212.87 sec  2017-04-11 07:25:34,478 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 216.04 sec  2017-04-11 07:25:38,348 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 221.99 sec  2017-04-11 07:25:39,548 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 223.01 sec  2017-04-11 07:25:44,608 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 232.66 sec  2017-04-11 07:25:48,180 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 237.55 sec  2017-04-11 07:25:50,552 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 245.86 sec  2017-04-11 07:25:54,209 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 250.89 sec  2017-04-11 07:25:56,971 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 262.46 sec  2017-04-11 07:26:09,258 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 276.78 sec  2017-04-11 07:26:10,477 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 279.79 sec  2017-04-11 07:26:13,125 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 291.39 sec  2017-04-11 07:26:16,704 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 295.74 sec  2017-04-11 07:26:17,911 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 296.8 sec  2017-04-11 07:26:19,113 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 297.59 sec  2017-04-11 07:26:31,787 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 328.88 sec  2017-04-11 07:26:34,152 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 334.23 sec  2017-04-11 07:26:37,790 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 339.85 sec  2017-04-11 07:26:42,638 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 351.64 sec  2017-04-11 07:26:43,842 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 353.52 sec  2017-04-11 07:26:46,205 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 356.57 sec  2017-04-11 07:26:49,893 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 360.35 sec  2017-04-11 07:26:59,976 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 388.05 sec  2017-04-11 07:27:06,099 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 395.99 sec  2017-04-11 07:27:07,299 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 397.24 sec  2017-04-11 07:27:09,687 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 399.32 sec  2017-04-11 07:27:16,158 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 408.09 sec  2017-04-11 07:27:21,123 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 415.46 sec  2017-04-11 07:27:23,560 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 417.93 sec  2017-04-11 07:27:34,547 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 433.47 sec  2017-04-11 07:27:38,379 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 436.59 sec  2017-04-11 07:27:41,993 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 440.03 sec  2017-04-11 07:27:56,626 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 459.65 sec  2017-04-11 07:27:59,067 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 460.95 sec  2017-04-11 07:28:05,567 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 466.78 sec  2017-04-11 07:28:08,215 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 468.48 sec  2017-04-11 07:28:12,035 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 478.8 sec  2017-04-11 07:28:13,480 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 479.82 sec  2017-04-11 07:28:20,011 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 484.88 sec  2017-04-11 07:28:21,220 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 489.63 sec  2017-04-11 07:28:28,488 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 505.15 sec  2017-04-11 07:28:30,892 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 512.86 sec  2017-04-11 07:28:35,633 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 516.83 sec  2017-04-11 07:28:38,242 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 519.74 sec  2017-04-11 07:28:39,430 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 520.97 sec  2017-04-11 07:28:45,702 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 536.68 sec  2017-04-11 07:28:51,940 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 547.71 sec  2017-04-11 07:28:55,520 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 564.08 sec  2017-04-11 07:28:59,077 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 578.74 sec  2017-04-11 07:29:00,278 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 579.76 sec  2017-04-11 07:29:02,683 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 583.28 sec  2017-04-11 07:29:07,536 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 592.06 sec  2017-04-11 07:29:11,200 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 606.32 sec  2017-04-11 07:29:12,419 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 608.0 sec  2017-04-11 07:29:18,373 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 623.96 sec  2017-04-11 07:29:19,752 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 626.28 sec  2017-04-11 07:29:22,085 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 634.94 sec  2017-04-11 07:29:23,272 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 636.21 sec  2017-04-11 07:29:25,579 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 640.45 sec  2017-04-11 07:29:27,813 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 642.62 sec  2017-04-11 07:29:31,435 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 654.36 sec  2017-04-11 07:29:36,134 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 657.94 sec  2017-04-11 07:29:40,164 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 661.39 sec  2017-04-11 07:29:45,999 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 665.77 sec  2017-04-11 07:29:56,399 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 672.54 sec  2017-04-11 07:29:59,631 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 674.21 sec  2017-04-11 07:30:03,561 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 676.28 sec  2017-04-11 07:30:09,982 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 680.57 sec  2017-04-11 07:30:11,069 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 680.76 sec  2017-04-11 07:30:12,141 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 681.03 sec  2017-04-11 07:30:16,361 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 683.88 sec  MapReduce Total cumulative CPU time: 11 minutes 23 seconds 880 msec  Ended Job = job\_1490094830773\_0386  Stage-4 is selected by condition resolver.  Stage-3 is filtered out by condition resolver.  Stage-5 is filtered out by condition resolver.  Moving data to: hdfs://mac55:8020/user/hive/warehouse/esplus.db/postimport\_sequence/.hive-staging\_hive\_2017-04-11\_07-23-42\_106\_310343789209772010-1/-ext-10000  Loading data to table esplus.postimport\_sequence  Table esplus.postimport\_sequence stats: [numFiles=74, numRows=42286538, totalSize=7091472133, rawDataSize=19438765021]  MapReduce Jobs Launched:  Stage-Stage-1: Map: 74 Cumulative CPU: 683.88 sec HDFS Read: 19486320096 HDFS Write: 7091479014 SUCCESS  Total MapReduce CPU Time Spent: 11 minutes 23 seconds 880 msec  OK  Time taken: 399.632 seconds |

## Data Size

|  |
| --- |
| hive> dfs -du -s -h /user/hive/warehouse/esplus.db/postimport\_sequence;  6.6 G 19.8 G /user/hive/warehouse/esplus.db/postimport\_sequence |

## Select some columns

SET hive.exec.compress.output=true;

|  |
| --- |
| select productname,username from postimport\_sequence;  INFO : Compiling command(queryId=hive\_20170411073131\_41f99941-a664-4b32-aba6-2ee611a3e945): select productname,username from postimport\_sequence  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:productname, type:string, comment:null), FieldSchema(name:username, type:string, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411073131\_41f99941-a664-4b32-aba6-2ee611a3e945); Time taken: 0.103 seconds  INFO : Executing command(queryId=hive\_20170411073131\_41f99941-a664-4b32-aba6-2ee611a3e945): select productname,username from postimport\_sequence  INFO : Query ID = hive\_20170411073131\_41f99941-a664-4b32-aba6-2ee611a3e945  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:25  INFO : Submitting tokens for job: job\_1490094830773\_0387  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0387/  INFO : Starting Job = job\_1490094830773\_0387, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0387/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0387  INFO : Hadoop job information for Stage-1: number of mappers: 25; number of reducers: 0  INFO : 2017-04-11 07:31:46,749 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 07:31:57,541 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 15.2 sec  INFO : 2017-04-11 07:32:00,758 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 67.03 sec  INFO : 2017-04-11 07:32:01,908 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 80.18 sec  INFO : 2017-04-11 07:32:02,934 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 81.52 sec  INFO : 2017-04-11 07:32:04,097 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 92.02 sec  INFO : 2017-04-11 07:32:05,261 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 101.13 sec  INFO : 2017-04-11 07:32:06,284 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 105.67 sec  INFO : 2017-04-11 07:32:07,453 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 110.06 sec  INFO : 2017-04-11 07:32:08,478 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 117.12 sec  INFO : 2017-04-11 07:32:09,619 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 120.88 sec  INFO : 2017-04-11 07:32:10,640 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 128.37 sec  INFO : 2017-04-11 07:32:11,809 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 128.87 sec  INFO : 2017-04-11 07:32:12,852 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 137.22 sec  INFO : 2017-04-11 07:32:14,277 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 140.17 sec  INFO : 2017-04-11 07:32:15,325 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 147.03 sec  INFO : 2017-04-11 07:32:16,659 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 149.68 sec  INFO : 2017-04-11 07:32:17,691 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 155.5 sec  INFO : 2017-04-11 07:32:18,883 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 163.31 sec  INFO : 2017-04-11 07:32:19,905 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 165.34 sec  INFO : 2017-04-11 07:32:21,057 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 178.68 sec  INFO : 2017-04-11 07:32:22,079 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 193.84 sec  INFO : 2017-04-11 07:32:23,216 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 199.57 sec  INFO : 2017-04-11 07:32:24,244 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 208.11 sec  INFO : 2017-04-11 07:32:26,546 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 208.97 sec  INFO : 2017-04-11 07:32:27,569 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 214.55 sec  INFO : 2017-04-11 07:32:28,714 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 216.57 sec  INFO : 2017-04-11 07:32:36,199 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 221.98 sec  INFO : 2017-04-11 07:32:39,492 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 222.9 sec  INFO : 2017-04-11 07:32:45,953 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 224.77 sec  INFO : MapReduce Total cumulative CPU time: 3 minutes 44 seconds 770 msec  INFO : Ended Job = job\_1490094830773\_0387  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 25 Cumulative CPU: 224.77 sec HDFS Read: 7091585511 HDFS Write: 233695318 SUCCESS  INFO : Total MapReduce CPU Time Spent: 3 minutes 44 seconds 770 msec  INFO : Completed executing command(queryId=hive\_20170411073131\_41f99941-a664-4b32-aba6-2ee611a3e945); Time taken: 65.993 seconds  INFO : OK |

## Check Max Date

|  |
| --- |
| select max(date) from postimport\_sequence;  INFO : Compiling command(queryId=hive\_20170411073333\_2549e0f1-d9d0-4c48-921a-743e0c250411): select max(date) from postimport\_sequence  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:\_c0, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411073333\_2549e0f1-d9d0-4c48-921a-743e0c250411); Time taken: 0.105 seconds  INFO : Executing command(queryId=hive\_20170411073333\_2549e0f1-d9d0-4c48-921a-743e0c250411): select max(date) from postimport\_sequence  INFO : Query ID = hive\_20170411073333\_2549e0f1-d9d0-4c48-921a-743e0c250411  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks determined at compile time: 1  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:25  INFO : Submitting tokens for job: job\_1490094830773\_0388  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0388/  INFO : Starting Job = job\_1490094830773\_0388, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0388/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0388  INFO : Hadoop job information for Stage-1: number of mappers: 25; number of reducers: 1  INFO : 2017-04-11 07:33:12,166 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 07:33:25,163 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 24.67 sec  INFO : 2017-04-11 07:33:26,325 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 43.9 sec  INFO : 2017-04-11 07:33:28,492 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 52.43 sec  INFO : 2017-04-11 07:33:29,515 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 57.73 sec  INFO : 2017-04-11 07:33:30,657 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 58.35 sec  INFO : 2017-04-11 07:33:31,682 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 64.48 sec  INFO : 2017-04-11 07:33:32,837 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 68.61 sec  INFO : 2017-04-11 07:33:34,858 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 75.37 sec  INFO : 2017-04-11 07:33:35,883 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 78.31 sec  INFO : 2017-04-11 07:33:37,093 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 83.75 sec  INFO : 2017-04-11 07:33:39,590 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 88.57 sec  INFO : 2017-04-11 07:33:40,625 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 91.18 sec  INFO : 2017-04-11 07:33:41,993 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 93.44 sec  INFO : 2017-04-11 07:33:44,178 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 98.12 sec  INFO : 2017-04-11 07:33:45,348 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 100.67 sec  INFO : 2017-04-11 07:33:46,386 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 106.08 sec  INFO : 2017-04-11 07:33:47,540 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 114.03 sec  INFO : 2017-04-11 07:33:48,562 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 124.76 sec  INFO : 2017-04-11 07:33:49,708 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 126.32 sec  INFO : 2017-04-11 07:33:50,735 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 130.39 sec  INFO : 2017-04-11 07:33:51,884 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 134.42 sec  INFO : 2017-04-11 07:33:55,176 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 146.54 sec  INFO : 2017-04-11 07:33:58,375 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 150.23 sec  INFO : 2017-04-11 07:33:59,522 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 151.53 sec  INFO : 2017-04-11 07:34:00,546 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 153.9 sec  INFO : 2017-04-11 07:34:03,836 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 155.66 sec  INFO : 2017-04-11 07:34:04,998 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 156.11 sec  INFO : 2017-04-11 07:34:06,021 Stage-1 map = 100%, reduce = 32%, Cumulative CPU 157.51 sec  INFO : 2017-04-11 07:34:08,187 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 158.34 sec  INFO : MapReduce Total cumulative CPU time: 2 minutes 38 seconds 340 msec  INFO : Ended Job = job\_1490094830773\_0388  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 25 Reduce: 1 Cumulative CPU: 158.34 sec HDFS Read: 7091599224 HDFS Write: 19 SUCCESS  INFO : Total MapReduce CPU Time Spent: 2 minutes 38 seconds 340 msec  INFO : Completed executing command(queryId=hive\_20170411073333\_2549e0f1-d9d0-4c48-921a-743e0c250411); Time taken: 62.074 seconds  INFO : OK |

## User wise Max Date

|  |
| --- |
| select username,max(date) from postimport\_sequence group by username;  INFO : Compiling command(queryId=hive\_20170411073535\_bca79536-4fcc-4c81-9182-ac7123fb699a): select username,max(date) from postimport\_sequence group by username  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:username, type:string, comment:null), FieldSchema(name:\_c1, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411073535\_bca79536-4fcc-4c81-9182-ac7123fb699a); Time taken: 0.113 seconds  INFO : Executing command(queryId=hive\_20170411073535\_bca79536-4fcc-4c81-9182-ac7123fb699a): select username,max(date) from postimport\_sequence group by username  INFO : Query ID = hive\_20170411073535\_bca79536-4fcc-4c81-9182-ac7123fb699a  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks not specified. Estimated from input data size: 106  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:25  INFO : Submitting tokens for job: job\_1490094830773\_0389  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0389/  INFO : Starting Job = job\_1490094830773\_0389, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0389/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0389  INFO : Hadoop job information for Stage-1: number of mappers: 25; number of reducers: 106  INFO : 2017-04-11 07:35:20,652 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 07:35:32,463 Stage-1 map = 6%, reduce = 0%, Cumulative CPU 39.27 sec  INFO : 2017-04-11 07:35:34,640 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 59.19 sec  INFO : 2017-04-11 07:35:35,838 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 67.83 sec  INFO : 2017-04-11 07:35:36,865 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 70.32 sec  INFO : 2017-04-11 07:35:38,021 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 76.64 sec  INFO : 2017-04-11 07:35:39,060 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 84.43 sec  INFO : 2017-04-11 07:35:40,196 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 91.03 sec  INFO : 2017-04-11 07:35:41,237 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 92.27 sec  INFO : 2017-04-11 07:35:42,478 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 94.88 sec  INFO : 2017-04-11 07:35:43,814 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 97.64 sec  INFO : 2017-04-11 07:35:48,427 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 99.75 sec  INFO : 2017-04-11 07:35:49,456 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 113.12 sec  INFO : 2017-04-11 07:35:50,610 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 124.24 sec  INFO : 2017-04-11 07:35:51,648 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 144.58 sec  INFO : 2017-04-11 07:35:52,801 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 150.95 sec  INFO : 2017-04-11 07:35:53,945 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 153.74 sec  INFO : 2017-04-11 07:35:54,969 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 163.97 sec  INFO : 2017-04-11 07:35:56,108 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 165.82 sec  INFO : 2017-04-11 07:35:57,131 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 168.19 sec  INFO : 2017-04-11 07:35:58,401 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 173.74 sec  INFO : 2017-04-11 07:36:03,150 Stage-1 map = 100%, reduce = 6%, Cumulative CPU 181.11 sec  INFO : 2017-04-11 07:36:04,289 Stage-1 map = 100%, reduce = 14%, Cumulative CPU 192.52 sec  INFO : 2017-04-11 07:36:07,907 Stage-1 map = 100%, reduce = 17%, Cumulative CPU 196.49 sec  INFO : 2017-04-11 07:36:09,206 Stage-1 map = 100%, reduce = 21%, Cumulative CPU 201.02 sec  INFO : 2017-04-11 07:36:10,462 Stage-1 map = 100%, reduce = 27%, Cumulative CPU 209.92 sec  INFO : 2017-04-11 07:36:11,485 Stage-1 map = 100%, reduce = 28%, Cumulative CPU 211.27 sec  INFO : 2017-04-11 07:36:12,915 Stage-1 map = 100%, reduce = 30%, Cumulative CPU 213.94 sec  INFO : 2017-04-11 07:36:13,986 Stage-1 map = 100%, reduce = 31%, Cumulative CPU 215.36 sec  INFO : 2017-04-11 07:36:15,416 Stage-1 map = 100%, reduce = 35%, Cumulative CPU 219.93 sec  INFO : 2017-04-11 07:36:17,609 Stage-1 map = 100%, reduce = 42%, Cumulative CPU 230.04 sec  INFO : 2017-04-11 07:36:18,635 Stage-1 map = 100%, reduce = 45%, Cumulative CPU 233.92 sec  INFO : 2017-04-11 07:36:21,459 Stage-1 map = 100%, reduce = 49%, Cumulative CPU 238.24 sec  INFO : 2017-04-11 07:36:22,491 Stage-1 map = 100%, reduce = 51%, Cumulative CPU 240.76 sec  INFO : 2017-04-11 07:36:23,703 Stage-1 map = 100%, reduce = 56%, Cumulative CPU 247.0 sec  INFO : 2017-04-11 07:36:24,726 Stage-1 map = 100%, reduce = 59%, Cumulative CPU 252.38 sec  INFO : 2017-04-11 07:36:26,960 Stage-1 map = 100%, reduce = 60%, Cumulative CPU 253.49 sec  INFO : 2017-04-11 07:36:28,440 Stage-1 map = 100%, reduce = 65%, Cumulative CPU 259.67 sec  INFO : 2017-04-11 07:36:29,644 Stage-1 map = 100%, reduce = 70%, Cumulative CPU 265.58 sec  INFO : 2017-04-11 07:36:31,850 Stage-1 map = 100%, reduce = 75%, Cumulative CPU 271.89 sec  INFO : 2017-04-11 07:36:34,333 Stage-1 map = 100%, reduce = 81%, Cumulative CPU 280.39 sec  INFO : 2017-04-11 07:36:35,390 Stage-1 map = 100%, reduce = 83%, Cumulative CPU 282.7 sec  INFO : 2017-04-11 07:36:36,669 Stage-1 map = 100%, reduce = 85%, Cumulative CPU 285.12 sec  INFO : 2017-04-11 07:36:37,946 Stage-1 map = 100%, reduce = 86%, Cumulative CPU 286.38 sec  INFO : 2017-04-11 07:36:38,968 Stage-1 map = 100%, reduce = 92%, Cumulative CPU 295.64 sec  INFO : 2017-04-11 07:36:40,110 Stage-1 map = 100%, reduce = 93%, Cumulative CPU 296.83 sec  INFO : 2017-04-11 07:36:41,132 Stage-1 map = 100%, reduce = 98%, Cumulative CPU 303.0 sec  INFO : 2017-04-11 07:36:42,312 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 305.19 sec  INFO : MapReduce Total cumulative CPU time: 5 minutes 5 seconds 190 msec  INFO : Ended Job = job\_1490094830773\_0389  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 25 Reduce: 106 Cumulative CPU: 305.19 sec HDFS Read: 7092009550 HDFS Write: 40584 SUCCESS  INFO : Total MapReduce CPU Time Spent: 5 minutes 5 seconds 190 msec  INFO : Completed executing command(queryId=hive\_20170411073535\_bca79536-4fcc-4c81-9182-ac7123fb699a); Time taken: 92.388 seconds  INFO : OK |

## Select All Columns

|  |
| --- |
| select \* from postimport\_sequence  INFO : Compiling command(queryId=hive\_20170411073636\_665592ff-b2c9-45b2-abe4-3e957953f415): select \* from postimport\_sequence  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_sequence.username, type:string, comment:null), FieldSchema(name:postimport\_sequence.machinename, type:string, comment:null), FieldSchema(name:postimport\_sequence.machineid, type:string, comment:null), FieldSchema(name:postimport\_sequence.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_sequence.domainname, type:string, comment:null), FieldSchema(name:postimport\_sequence.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_sequence.modulename, type:string, comment:null), FieldSchema(name:postimport\_sequence.filename, type:string, comment:null), FieldSchema(name:postimport\_sequence.productname, type:string, comment:null), FieldSchema(name:postimport\_sequence.companyname, type:string, comment:null), FieldSchema(name:postimport\_sequence.description, type:string, comment:null), FieldSchema(name:postimport\_sequence.text, type:string, comment:null), FieldSchema(name:postimport\_sequence.activetime, type:double, comment:null), FieldSchema(name:postimport\_sequence.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_sequence.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_sequence.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411073636\_665592ff-b2c9-45b2-abe4-3e957953f415); Time taken: 0.1 seconds  INFO : Executing command(queryId=hive\_20170411073636\_665592ff-b2c9-45b2-abe4-3e957953f415): select \* from postimport\_sequence  INFO : Query ID = hive\_20170411073636\_665592ff-b2c9-45b2-abe4-3e957953f415  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:25  INFO : Submitting tokens for job: job\_1490094830773\_0390  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0390/  INFO : Starting Job = job\_1490094830773\_0390, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0390/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0390  INFO : Hadoop job information for Stage-1: number of mappers: 25; number of reducers: 0  INFO : 2017-04-11 07:37:06,668 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 07:37:29,053 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 166.39 sec  INFO : 2017-04-11 07:37:30,313 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 175.89 sec  INFO : 2017-04-11 07:37:31,341 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 185.27 sec  INFO : 2017-04-11 07:37:33,558 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 207.72 sec  INFO : 2017-04-11 07:37:41,698 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 268.02 sec  INFO : 2017-04-11 07:37:46,026 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 290.6 sec  INFO : 2017-04-11 07:37:50,363 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 305.17 sec  INFO : 2017-04-11 07:37:51,513 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 314.95 sec  INFO : 2017-04-11 07:37:53,680 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 320.76 sec  INFO : 2017-04-11 07:37:54,826 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 329.99 sec  INFO : 2017-04-11 07:37:58,011 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 344.39 sec  INFO : 2017-04-11 07:38:00,299 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 355.3 sec  INFO : 2017-04-11 07:38:03,483 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 370.92 sec  INFO : 2017-04-11 07:38:04,633 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 373.65 sec  INFO : 2017-04-11 07:38:15,652 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 437.86 sec  INFO : 2017-04-11 07:38:19,135 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 463.91 sec  INFO : 2017-04-11 07:38:21,320 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 476.19 sec  INFO : 2017-04-11 07:38:24,629 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 498.49 sec  INFO : 2017-04-11 07:38:26,815 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 502.11 sec  INFO : 2017-04-11 07:38:33,460 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 552.87 sec  INFO : 2017-04-11 07:38:36,787 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 576.43 sec  INFO : 2017-04-11 07:38:40,113 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 594.46 sec  INFO : 2017-04-11 07:38:42,296 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 607.66 sec  INFO : 2017-04-11 07:38:44,458 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 614.08 sec  INFO : 2017-04-11 07:38:45,621 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 628.11 sec  INFO : 2017-04-11 07:38:49,117 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 646.34 sec  INFO : 2017-04-11 07:38:57,871 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 704.89 sec  INFO : 2017-04-11 07:39:01,217 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 729.54 sec  INFO : 2017-04-11 07:39:03,392 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 743.47 sec  INFO : 2017-04-11 07:39:05,582 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 754.35 sec  INFO : 2017-04-11 07:39:06,728 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 766.28 sec  INFO : 2017-04-11 07:39:07,874 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 770.14 sec  INFO : 2017-04-11 07:39:10,087 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 782.83 sec  INFO : 2017-04-11 07:39:11,110 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 784.71 sec  INFO : 2017-04-11 07:39:13,654 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 798.15 sec  INFO : 2017-04-11 07:39:15,884 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 806.23 sec  INFO : 2017-04-11 07:39:19,092 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 832.34 sec  INFO : 2017-04-11 07:39:21,398 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 848.0 sec  INFO : 2017-04-11 07:39:24,604 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 863.77 sec  INFO : 2017-04-11 07:39:25,770 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 870.33 sec  INFO : 2017-04-11 07:39:29,095 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 883.97 sec  INFO : 2017-04-11 07:39:33,436 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 906.94 sec  INFO : 2017-04-11 07:39:34,579 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 916.04 sec  INFO : 2017-04-11 07:39:36,742 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 923.75 sec  INFO : 2017-04-11 07:39:41,069 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 945.24 sec  INFO : 2017-04-11 07:39:43,233 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 951.58 sec  INFO : 2017-04-11 07:39:44,373 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 957.96 sec  INFO : 2017-04-11 07:39:47,671 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 968.75 sec  INFO : 2017-04-11 07:39:48,690 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 972.21 sec  INFO : 2017-04-11 07:39:51,998 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 983.26 sec  INFO : 2017-04-11 07:40:01,796 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 1013.86 sec  INFO : 2017-04-11 07:40:04,069 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 1018.32 sec  INFO : 2017-04-11 07:40:07,246 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 1026.05 sec  INFO : 2017-04-11 07:40:11,690 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 1036.83 sec  INFO : 2017-04-11 07:40:17,185 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 1045.28 sec  INFO : 2017-04-11 07:40:21,506 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 1054.59 sec  INFO : 2017-04-11 07:40:30,261 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 1070.73 sec  INFO : 2017-04-11 07:40:31,280 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1074.41 sec  INFO : MapReduce Total cumulative CPU time: 17 minutes 54 seconds 410 msec  INFO : Ended Job = job\_1490094830773\_0390  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 25 Cumulative CPU: 1074.41 sec HDFS Read: 7091609611 HDFS Write: 3633977612 SUCCESS  INFO : Total MapReduce CPU Time Spent: 17 minutes 54 seconds 410 msec  INFO : Completed executing command(queryId=hive\_20170411073636\_665592ff-b2c9-45b2-abe4-3e957953f415); Time taken: 213.097 seconds  INFO : OK |

## Select All Columns with filter Condition

|  |
| --- |
| select \* from postimport\_sequence where username = '670063006100750074006F007500730065007200';  INFO : Compiling command(queryId=hive\_20170411074444\_9ec05def-53a3-49f5-a6d4-8f2ebbf4f4c8): select \* from postimport\_sequence where username = '670063006100750074006F007500730065007200'  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_sequence.username, type:string, comment:null), FieldSchema(name:postimport\_sequence.machinename, type:string, comment:null), FieldSchema(name:postimport\_sequence.machineid, type:string, comment:null), FieldSchema(name:postimport\_sequence.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_sequence.domainname, type:string, comment:null), FieldSchema(name:postimport\_sequence.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_sequence.modulename, type:string, comment:null), FieldSchema(name:postimport\_sequence.filename, type:string, comment:null), FieldSchema(name:postimport\_sequence.productname, type:string, comment:null), FieldSchema(name:postimport\_sequence.companyname, type:string, comment:null), FieldSchema(name:postimport\_sequence.description, type:string, comment:null), FieldSchema(name:postimport\_sequence.text, type:string, comment:null), FieldSchema(name:postimport\_sequence.activetime, type:double, comment:null), FieldSchema(name:postimport\_sequence.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_sequence.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_sequence.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411074444\_9ec05def-53a3-49f5-a6d4-8f2ebbf4f4c8); Time taken: 0.099 seconds  INFO : Executing command(queryId=hive\_20170411074444\_9ec05def-53a3-49f5-a6d4-8f2ebbf4f4c8): select \* from postimport\_sequence where username = '670063006100750074006F007500730065007200'  INFO : Query ID = hive\_20170411074444\_9ec05def-53a3-49f5-a6d4-8f2ebbf4f4c8  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:25  INFO : Submitting tokens for job: job\_1490094830773\_0391  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0391/  INFO : Starting Job = job\_1490094830773\_0391, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0391/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0391  INFO : Hadoop job information for Stage-1: number of mappers: 25; number of reducers: 0  INFO : 2017-04-11 07:44:32,590 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 07:44:44,060 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 11.66 sec  INFO : 2017-04-11 07:44:45,216 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 39.68 sec  INFO : 2017-04-11 07:44:46,240 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 51.09 sec  INFO : 2017-04-11 07:44:47,410 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 54.56 sec  INFO : 2017-04-11 07:44:48,552 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 63.24 sec  INFO : 2017-04-11 07:44:49,603 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 69.27 sec  INFO : 2017-04-11 07:44:50,759 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 72.68 sec  INFO : 2017-04-11 07:44:51,941 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 83.8 sec  INFO : 2017-04-11 07:44:52,966 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 88.93 sec  INFO : 2017-04-11 07:44:54,144 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 92.15 sec  INFO : 2017-04-11 07:44:58,522 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 96.58 sec  INFO : 2017-04-11 07:44:59,675 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 106.31 sec  INFO : 2017-04-11 07:45:00,701 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 116.51 sec  INFO : 2017-04-11 07:45:01,835 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 139.63 sec  INFO : 2017-04-11 07:45:02,974 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 146.6 sec  INFO : 2017-04-11 07:45:03,994 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 147.62 sec  INFO : 2017-04-11 07:45:05,140 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 149.36 sec  INFO : MapReduce Total cumulative CPU time: 2 minutes 29 seconds 360 msec  INFO : Ended Job = job\_1490094830773\_0391  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 25 Cumulative CPU: 149.36 sec HDFS Read: 7091621611 HDFS Write: 2086413 SUCCESS  INFO : Total MapReduce CPU Time Spent: 2 minutes 29 seconds 360 msec  INFO : Completed executing command(queryId=hive\_20170411074444\_9ec05def-53a3-49f5-a6d4-8f2ebbf4f4c8); Time taken: 40.242 seconds  INFO : OK |

# Sequence File format Gzip Compression

|  |
| --- |
| SET hive.exec.compress.output=true;  SET mapred.output.compression.type=BLOCK;  SET mapred.output.compression.codec=org.apache.hadoop.io.compress.GzipCodec;  SET io.seqfile.compression.type=BLOCK; |

## Load data to Sequence Table

|  |
| --- |
| hive> insert overwrite table postimport\_sequence select \* from postimport\_text;  Query ID = shalaj\_20170411074646\_fc11fb47-458f-4c42-b3d4-517bd69ada77  Total jobs = 3  Launching Job 1 out of 3  Number of reduce tasks is set to 0 since there's no reduce operator  Starting Job = job\_1490094830773\_0392, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0392/  Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0392  Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  2017-04-11 07:46:12,675 Stage-1 map = 0%, reduce = 0%  2017-04-11 07:46:32,220 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 110.77 sec  2017-04-11 07:46:34,658 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 126.18 sec  2017-04-11 07:46:35,889 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 129.5 sec  2017-04-11 07:46:38,547 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 145.46 sec  2017-04-11 07:46:39,701 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 146.78 sec  2017-04-11 07:46:40,904 Stage-1 map = 6%, reduce = 0%, Cumulative CPU 161.0 sec  2017-04-11 07:46:43,311 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 170.44 sec  2017-04-11 07:46:46,830 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 183.39 sec  2017-04-11 07:46:50,276 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 195.34 sec  2017-04-11 07:46:52,526 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 212.37 sec  2017-04-11 07:46:56,022 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 223.94 sec  2017-04-11 07:46:59,464 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 240.15 sec  2017-04-11 07:47:09,665 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 303.7 sec  2017-04-11 07:47:11,401 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 307.53 sec  2017-04-11 07:47:12,606 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 320.18 sec  2017-04-11 07:47:13,820 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 330.73 sec  2017-04-11 07:47:17,410 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 348.24 sec  2017-04-11 07:47:20,990 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 360.48 sec  2017-04-11 07:47:22,134 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 372.13 sec  2017-04-11 07:47:23,308 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 376.93 sec  2017-04-11 07:47:27,107 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 403.78 sec  2017-04-11 07:47:30,696 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 424.38 sec  2017-04-11 07:47:31,864 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 436.77 sec  2017-04-11 07:47:34,297 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 447.44 sec  2017-04-11 07:47:37,926 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 481.23 sec  2017-04-11 07:47:40,336 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 493.59 sec  2017-04-11 07:47:43,830 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 507.45 sec  2017-04-11 07:47:49,892 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 546.1 sec  2017-04-11 07:47:51,058 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 561.04 sec  2017-04-11 07:47:52,260 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 570.02 sec  2017-04-11 07:47:54,870 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 579.82 sec  2017-04-11 07:47:56,024 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 592.67 sec  2017-04-11 07:47:59,501 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 606.28 sec  2017-04-11 07:48:00,815 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 614.46 sec  2017-04-11 07:48:02,116 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 624.68 sec  2017-04-11 07:48:04,372 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 632.2 sec  2017-04-11 07:48:11,286 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 678.36 sec  2017-04-11 07:48:13,588 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 686.6 sec  2017-04-11 07:48:18,815 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 705.86 sec  2017-04-11 07:48:24,932 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 730.57 sec  2017-04-11 07:48:27,301 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 740.59 sec  2017-04-11 07:48:29,914 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 748.25 sec  2017-04-11 07:48:32,523 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 762.28 sec  2017-04-11 07:48:36,131 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 776.37 sec  2017-04-11 07:48:44,942 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 824.03 sec  2017-04-11 07:48:48,547 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 838.43 sec  2017-04-11 07:48:51,266 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 856.07 sec  2017-04-11 07:48:58,394 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 878.96 sec  2017-04-11 07:48:59,693 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 880.65 sec  2017-04-11 07:49:00,781 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 889.59 sec  2017-04-11 07:49:04,060 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 912.98 sec  2017-04-11 07:49:07,601 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 926.41 sec  2017-04-11 07:49:12,189 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 945.41 sec  2017-04-11 07:49:13,722 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 951.9 sec  2017-04-11 07:49:15,491 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 956.96 sec  2017-04-11 07:49:18,069 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 975.84 sec  2017-04-11 07:49:19,143 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 979.94 sec  2017-04-11 07:49:24,149 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 1010.95 sec  2017-04-11 07:49:25,543 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 1012.94 sec  2017-04-11 07:49:29,868 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 1039.72 sec  2017-04-11 07:49:31,380 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 1045.24 sec  2017-04-11 07:49:38,510 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 1069.72 sec  2017-04-11 07:49:39,582 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 1077.22 sec  2017-04-11 07:49:42,690 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 1095.32 sec  2017-04-11 07:49:44,860 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 1106.39 sec  2017-04-11 07:49:48,299 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 1140.1 sec  2017-04-11 07:49:49,589 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 1147.47 sec  2017-04-11 07:49:51,732 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 1162.01 sec  2017-04-11 07:49:53,878 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 1168.2 sec  2017-04-11 07:49:54,950 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 1172.71 sec  2017-04-11 07:49:56,247 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 1173.05 sec  2017-04-11 07:49:58,403 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 1181.38 sec  2017-04-11 07:50:06,070 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 1195.39 sec  2017-04-11 07:50:07,130 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 1195.96 sec  2017-04-11 07:50:17,666 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 1201.76 sec  2017-04-11 07:50:35,497 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1210.19 sec  MapReduce Total cumulative CPU time: 20 minutes 10 seconds 190 msec  Ended Job = job\_1490094830773\_0392  Stage-4 is selected by condition resolver.  Stage-3 is filtered out by condition resolver.  Stage-5 is filtered out by condition resolver.  Moving data to: hdfs://mac55:8020/user/hive/warehouse/esplus.db/postimport\_sequence/.hive-staging\_hive\_2017-04-11\_07-46-03\_344\_8738988929322067289-1/-ext-10000  Loading data to table esplus.postimport\_sequence  Table esplus.postimport\_sequence stats: [numFiles=74, numRows=42286538, totalSize=3803011895, rawDataSize=19438765021]  MapReduce Jobs Launched:  Stage-Stage-1: Map: 74 Cumulative CPU: 1210.19 sec HDFS Read: 19486320244 HDFS Write: 3803018776 SUCCESS  Total MapReduce CPU Time Spent: 20 minutes 10 seconds 190 msec  OK  Time taken: 280.252 seconds |

## Data Size

|  |
| --- |
| hive> dfs -du -s -h /user/hive/warehouse/esplus.db/postimport\_sequence  3.5 G 10.6 G /user/hive/warehouse/esplus.db/postimport\_sequence |

## Select some columns

SET hive.exec.compress.output=true;

|  |
| --- |
| select productname,username from postimport\_sequence;  INFO : Compiling command(queryId=hive\_20170411075151\_a1cc1b73-43cd-4908-8cec-b96e02d380f8): select productname,username from postimport\_sequence  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:productname, type:string, comment:null), FieldSchema(name:username, type:string, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411075151\_a1cc1b73-43cd-4908-8cec-b96e02d380f8); Time taken: 0.119 seconds  INFO : Executing command(queryId=hive\_20170411075151\_a1cc1b73-43cd-4908-8cec-b96e02d380f8): select productname,username from postimport\_sequence  INFO : Query ID = hive\_20170411075151\_a1cc1b73-43cd-4908-8cec-b96e02d380f8  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:15  INFO : Submitting tokens for job: job\_1490094830773\_0393  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0393/  INFO : Starting Job = job\_1490094830773\_0393, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0393/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0393  INFO : Hadoop job information for Stage-1: number of mappers: 15; number of reducers: 0  INFO : 2017-04-11 07:51:51,999 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 07:52:00,901 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 2.48 sec  INFO : 2017-04-11 07:52:03,154 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 21.72 sec  INFO : 2017-04-11 07:52:04,385 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 50.18 sec  INFO : 2017-04-11 07:52:05,424 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 60.52 sec  INFO : 2017-04-11 07:52:06,757 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 91.93 sec  INFO : 2017-04-11 07:52:07,795 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 112.4 sec  INFO : 2017-04-11 07:52:09,081 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 129.34 sec  INFO : 2017-04-11 07:52:10,299 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 141.07 sec  INFO : 2017-04-11 07:52:11,329 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 151.84 sec  INFO : 2017-04-11 07:52:12,627 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 171.0 sec  INFO : 2017-04-11 07:52:13,788 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 193.61 sec  INFO : 2017-04-11 07:52:14,816 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 200.05 sec  INFO : 2017-04-11 07:52:16,009 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 210.69 sec  INFO : 2017-04-11 07:52:17,170 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 220.39 sec  INFO : 2017-04-11 07:52:18,192 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 229.92 sec  INFO : 2017-04-11 07:52:19,328 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 236.2 sec  INFO : 2017-04-11 07:52:21,488 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 237.13 sec  INFO : 2017-04-11 07:52:22,624 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 239.61 sec  INFO : 2017-04-11 07:52:23,645 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 241.3 sec  INFO : MapReduce Total cumulative CPU time: 4 minutes 1 seconds 300 msec  INFO : Ended Job = job\_1490094830773\_0393  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 15 Cumulative CPU: 241.3 sec HDFS Read: 3803086225 HDFS Write: 233693474 SUCCESS  INFO : Total MapReduce CPU Time Spent: 4 minutes 1 seconds 300 msec  INFO : Completed executing command(queryId=hive\_20170411075151\_a1cc1b73-43cd-4908-8cec-b96e02d380f8); Time taken: 40.481 seconds  INFO : OK |

## Check Max Date

|  |
| --- |
| select max(date) from postimport\_sequence;  INFO : Compiling command(queryId=hive\_20170411075252\_d526450e-b275-4df2-9631-8996a2612e34): select max(date) from postimport\_sequence  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:\_c0, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411075252\_d526450e-b275-4df2-9631-8996a2612e34); Time taken: 0.121 seconds  INFO : Executing command(queryId=hive\_20170411075252\_d526450e-b275-4df2-9631-8996a2612e34): select max(date) from postimport\_sequence  INFO : Query ID = hive\_20170411075252\_d526450e-b275-4df2-9631-8996a2612e34  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks determined at compile time: 1  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:15  INFO : Submitting tokens for job: job\_1490094830773\_0394  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0394/  INFO : Starting Job = job\_1490094830773\_0394, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0394/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0394  INFO : Hadoop job information for Stage-1: number of mappers: 15; number of reducers: 1  INFO : 2017-04-11 07:53:06,415 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 07:53:18,553 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 16.89 sec  INFO : 2017-04-11 07:53:19,715 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 50.13 sec  INFO : 2017-04-11 07:53:20,752 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 67.43 sec  INFO : 2017-04-11 07:53:21,971 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 90.2 sec  INFO : 2017-04-11 07:53:23,250 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 103.83 sec  INFO : 2017-04-11 07:53:24,284 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 112.08 sec  INFO : 2017-04-11 07:53:25,548 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 128.9 sec  INFO : 2017-04-11 07:53:26,694 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 140.63 sec  INFO : 2017-04-11 07:53:27,715 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 150.76 sec  INFO : 2017-04-11 07:53:28,855 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 158.47 sec  INFO : 2017-04-11 07:53:32,181 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 163.19 sec  INFO : 2017-04-11 07:53:33,357 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 165.72 sec  INFO : 2017-04-11 07:53:34,380 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 169.27 sec  INFO : 2017-04-11 07:53:35,517 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 169.95 sec  INFO : 2017-04-11 07:53:36,702 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 172.3 sec  INFO : 2017-04-11 07:53:37,726 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 173.23 sec  INFO : MapReduce Total cumulative CPU time: 2 minutes 53 seconds 230 msec  INFO : Ended Job = job\_1490094830773\_0394  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 15 Reduce: 1 Cumulative CPU: 173.23 sec HDFS Read: 3803096028 HDFS Write: 19 SUCCESS  INFO : Total MapReduce CPU Time Spent: 2 minutes 53 seconds 230 msec  INFO : Completed executing command(queryId=hive\_20170411075252\_d526450e-b275-4df2-9631-8996a2612e34); Time taken: 39.042 seconds  INFO : OK |

## User wise Max Date

|  |
| --- |
| select username,max(date) from postimport\_sequence group by username;  INFO : Compiling command(queryId=hive\_20170411075454\_2cc627e8-8a30-4a5a-b455-475ed75afb0c): select username,max(date) from postimport\_sequence group by username  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:username, type:string, comment:null), FieldSchema(name:\_c1, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411075454\_2cc627e8-8a30-4a5a-b455-475ed75afb0c); Time taken: 0.232 seconds  INFO : Executing command(queryId=hive\_20170411075454\_2cc627e8-8a30-4a5a-b455-475ed75afb0c): select username,max(date) from postimport\_sequence group by username  INFO : Query ID = hive\_20170411075454\_2cc627e8-8a30-4a5a-b455-475ed75afb0c  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks not specified. Estimated from input data size: 57  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:15  INFO : Submitting tokens for job: job\_1490094830773\_0395  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0395/  INFO : Starting Job = job\_1490094830773\_0395, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0395/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0395  INFO : Hadoop job information for Stage-1: number of mappers: 15; number of reducers: 57  INFO : 2017-04-11 07:54:49,548 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 07:54:58,300 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 2.75 sec  INFO : 2017-04-11 07:55:00,531 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 20.02 sec  INFO : 2017-04-11 07:55:01,908 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 43.62 sec  INFO : 2017-04-11 07:55:03,171 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 60.26 sec  INFO : 2017-04-11 07:55:04,206 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 67.24 sec  INFO : 2017-04-11 07:55:05,412 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 105.03 sec  INFO : 2017-04-11 07:55:06,648 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 112.07 sec  INFO : 2017-04-11 07:55:07,680 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 120.94 sec  INFO : 2017-04-11 07:55:08,917 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 139.25 sec  INFO : 2017-04-11 07:55:10,080 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 148.15 sec  INFO : 2017-04-11 07:55:11,103 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 156.6 sec  INFO : 2017-04-11 07:55:12,243 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 159.46 sec  INFO : 2017-04-11 07:55:14,422 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 164.2 sec  INFO : 2017-04-11 07:55:17,850 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 170.12 sec  INFO : 2017-04-11 07:55:21,180 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 176.39 sec  INFO : 2017-04-11 07:55:23,351 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 181.94 sec  INFO : 2017-04-11 07:55:24,489 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 183.9 sec  INFO : 2017-04-11 07:55:26,656 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 186.41 sec  INFO : 2017-04-11 07:55:28,040 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 187.84 sec  INFO : 2017-04-11 07:55:33,207 Stage-1 map = 100%, reduce = 19%, Cumulative CPU 202.27 sec  INFO : 2017-04-11 07:55:35,672 Stage-1 map = 100%, reduce = 26%, Cumulative CPU 206.95 sec  INFO : 2017-04-11 07:55:38,302 Stage-1 map = 100%, reduce = 33%, Cumulative CPU 211.64 sec  INFO : 2017-04-11 07:55:39,666 Stage-1 map = 100%, reduce = 40%, Cumulative CPU 216.62 sec  INFO : 2017-04-11 07:55:40,808 Stage-1 map = 100%, reduce = 46%, Cumulative CPU 220.59 sec  INFO : 2017-04-11 07:55:43,270 Stage-1 map = 100%, reduce = 53%, Cumulative CPU 225.49 sec  INFO : 2017-04-11 07:55:44,825 Stage-1 map = 100%, reduce = 65%, Cumulative CPU 234.39 sec  INFO : 2017-04-11 07:55:47,116 Stage-1 map = 100%, reduce = 72%, Cumulative CPU 239.64 sec  INFO : 2017-04-11 07:55:49,753 Stage-1 map = 100%, reduce = 84%, Cumulative CPU 247.9 sec  INFO : 2017-04-11 07:55:50,806 Stage-1 map = 100%, reduce = 91%, Cumulative CPU 252.32 sec  INFO : 2017-04-11 07:55:53,414 Stage-1 map = 100%, reduce = 93%, Cumulative CPU 253.38 sec  INFO : 2017-04-11 07:55:54,434 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 258.76 sec  INFO : MapReduce Total cumulative CPU time: 4 minutes 18 seconds 760 msec  INFO : Ended Job = job\_1490094830773\_0395  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 15 Reduce: 57 Cumulative CPU: 258.76 sec HDFS Read: 3803315176 HDFS Write: 36041 SUCCESS  INFO : Total MapReduce CPU Time Spent: 4 minutes 18 seconds 760 msec  INFO : Completed executing command(queryId=hive\_20170411075454\_2cc627e8-8a30-4a5a-b455-475ed75afb0c); Time taken: 76.483 seconds  INFO : OK |

## Select All Columns

|  |
| --- |
| select \* from postimport\_sequence  INFO : Compiling command(queryId=hive\_20170411075858\_e2bb7e6b-c0d8-4a20-aff3-2f5ddf5c06c8): select \* from postimport\_sequence  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_sequence.username, type:string, comment:null), FieldSchema(name:postimport\_sequence.machinename, type:string, comment:null), FieldSchema(name:postimport\_sequence.machineid, type:string, comment:null), FieldSchema(name:postimport\_sequence.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_sequence.domainname, type:string, comment:null), FieldSchema(name:postimport\_sequence.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_sequence.modulename, type:string, comment:null), FieldSchema(name:postimport\_sequence.filename, type:string, comment:null), FieldSchema(name:postimport\_sequence.productname, type:string, comment:null), FieldSchema(name:postimport\_sequence.companyname, type:string, comment:null), FieldSchema(name:postimport\_sequence.description, type:string, comment:null), FieldSchema(name:postimport\_sequence.text, type:string, comment:null), FieldSchema(name:postimport\_sequence.activetime, type:double, comment:null), FieldSchema(name:postimport\_sequence.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_sequence.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_sequence.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411075858\_e2bb7e6b-c0d8-4a20-aff3-2f5ddf5c06c8); Time taken: 0.105 seconds  INFO : Executing command(queryId=hive\_20170411075858\_e2bb7e6b-c0d8-4a20-aff3-2f5ddf5c06c8): select \* from postimport\_sequence  INFO : Query ID = hive\_20170411075858\_e2bb7e6b-c0d8-4a20-aff3-2f5ddf5c06c8  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:15  INFO : Submitting tokens for job: job\_1490094830773\_0396  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0396/  INFO : Starting Job = job\_1490094830773\_0396, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0396/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0396  INFO : Hadoop job information for Stage-1: number of mappers: 15; number of reducers: 0  INFO : 2017-04-11 07:58:47,465 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 07:58:59,986 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 48.35 sec  INFO : 2017-04-11 07:59:12,317 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 147.87 sec  INFO : 2017-04-11 07:59:20,117 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 190.49 sec  INFO : 2017-04-11 07:59:23,441 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 209.21 sec  INFO : 2017-04-11 07:59:27,923 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 230.47 sec  INFO : 2017-04-11 07:59:28,949 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 239.65 sec  INFO : 2017-04-11 07:59:31,291 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 246.53 sec  INFO : 2017-04-11 07:59:32,446 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 257.1 sec  INFO : 2017-04-11 07:59:34,629 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 265.99 sec  INFO : 2017-04-11 07:59:37,951 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 284.85 sec  INFO : 2017-04-11 07:59:39,099 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 290.89 sec  INFO : 2017-04-11 07:59:45,743 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 321.19 sec  INFO : 2017-04-11 07:59:48,035 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 332.05 sec  INFO : 2017-04-11 07:59:50,210 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 342.3 sec  INFO : 2017-04-11 07:59:51,357 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 347.48 sec  INFO : 2017-04-11 07:59:53,526 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 360.78 sec  INFO : 2017-04-11 07:59:54,675 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 364.14 sec  INFO : 2017-04-11 07:59:56,840 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 375.01 sec  INFO : 2017-04-11 07:59:57,987 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 378.33 sec  INFO : 2017-04-11 08:00:00,153 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 388.89 sec  INFO : 2017-04-11 08:00:02,321 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 398.56 sec  INFO : 2017-04-11 08:00:17,903 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 475.69 sec  INFO : 2017-04-11 08:00:19,057 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 478.37 sec  INFO : 2017-04-11 08:00:21,226 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 488.5 sec  INFO : 2017-04-11 08:00:23,531 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 497.4 sec  INFO : 2017-04-11 08:00:25,706 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 506.69 sec  INFO : 2017-04-11 08:00:27,873 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 518.71 sec  INFO : 2017-04-11 08:00:29,033 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 521.38 sec  INFO : 2017-04-11 08:00:33,502 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 543.33 sec  INFO : 2017-04-11 08:00:40,160 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 575.0 sec  INFO : 2017-04-11 08:00:42,455 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 585.16 sec  INFO : 2017-04-11 08:00:51,439 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 627.07 sec  INFO : 2017-04-11 08:00:52,457 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 629.93 sec  INFO : 2017-04-11 08:00:53,606 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 633.95 sec  INFO : 2017-04-11 08:00:54,764 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 641.9 sec  INFO : 2017-04-11 08:00:58,090 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 656.96 sec  INFO : 2017-04-11 08:01:00,265 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 664.54 sec  INFO : 2017-04-11 08:01:03,711 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 683.6 sec  INFO : 2017-04-11 08:01:10,342 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 712.68 sec  INFO : 2017-04-11 08:01:12,509 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 722.08 sec  INFO : 2017-04-11 08:01:15,835 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 739.17 sec  INFO : 2017-04-11 08:01:16,982 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 742.62 sec  INFO : 2017-04-11 08:01:18,141 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 745.22 sec  INFO : 2017-04-11 08:01:20,307 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 756.91 sec  INFO : 2017-04-11 08:01:21,473 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 761.14 sec  INFO : 2017-04-11 08:01:22,618 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 767.51 sec  INFO : 2017-04-11 08:01:28,121 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 790.36 sec  INFO : 2017-04-11 08:01:34,916 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 823.83 sec  INFO : 2017-04-11 08:01:40,414 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 852.38 sec  INFO : 2017-04-11 08:01:41,560 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 859.19 sec  INFO : 2017-04-11 08:01:43,731 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 866.64 sec  INFO : 2017-04-11 08:01:44,882 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 876.17 sec  INFO : 2017-04-11 08:01:46,034 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 882.14 sec  INFO : 2017-04-11 08:01:48,204 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 894.15 sec  INFO : 2017-04-11 08:01:50,522 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 908.06 sec  INFO : 2017-04-11 08:01:52,698 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 919.2 sec  INFO : 2017-04-11 08:01:53,874 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 931.71 sec  INFO : 2017-04-11 08:01:55,026 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 937.81 sec  INFO : 2017-04-11 08:01:57,191 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 950.02 sec  INFO : 2017-04-11 08:02:00,503 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 965.72 sec  INFO : 2017-04-11 08:02:06,110 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 996.88 sec  INFO : 2017-04-11 08:02:09,423 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 1009.17 sec  INFO : 2017-04-11 08:02:15,016 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 1033.49 sec  INFO : 2017-04-11 08:02:16,035 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 1034.93 sec  INFO : 2017-04-11 08:02:17,180 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 1036.38 sec  INFO : 2017-04-11 08:02:18,322 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 1045.65 sec  INFO : 2017-04-11 08:02:23,788 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 1061.93 sec  INFO : 2017-04-11 08:02:27,096 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 1073.43 sec  INFO : 2017-04-11 08:02:30,529 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 1080.56 sec  INFO : 2017-04-11 08:02:33,841 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 1086.89 sec  INFO : 2017-04-11 08:02:36,003 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 1091.17 sec  INFO : MapReduce Total cumulative CPU time: 18 minutes 11 seconds 170 msec  INFO : Ended Job = job\_1490094830773\_0396  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 15 Cumulative CPU: 1091.17 sec HDFS Read: 3803100865 HDFS Write: 3633975659 SUCCESS  INFO : Total MapReduce CPU Time Spent: 18 minutes 11 seconds 170 msec  INFO : Completed executing command(queryId=hive\_20170411075858\_e2bb7e6b-c0d8-4a20-aff3-2f5ddf5c06c8); Time taken: 236.055 seconds  INFO : OK |

## Select All Columns with filter Condition

|  |
| --- |
| select \* from postimport\_sequence where username = '670063006100750074006F007500730065007200'  INFO : Compiling command(queryId=hive\_20170411080303\_741b9156-ed60-4a2e-bc78-d054077849fc): select \* from postimport\_sequence where username = '670063006100750074006F007500730065007200'  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_sequence.username, type:string, comment:null), FieldSchema(name:postimport\_sequence.machinename, type:string, comment:null), FieldSchema(name:postimport\_sequence.machineid, type:string, comment:null), FieldSchema(name:postimport\_sequence.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_sequence.domainname, type:string, comment:null), FieldSchema(name:postimport\_sequence.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_sequence.modulename, type:string, comment:null), FieldSchema(name:postimport\_sequence.filename, type:string, comment:null), FieldSchema(name:postimport\_sequence.productname, type:string, comment:null), FieldSchema(name:postimport\_sequence.companyname, type:string, comment:null), FieldSchema(name:postimport\_sequence.description, type:string, comment:null), FieldSchema(name:postimport\_sequence.text, type:string, comment:null), FieldSchema(name:postimport\_sequence.activetime, type:double, comment:null), FieldSchema(name:postimport\_sequence.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_sequence.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_sequence.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411080303\_741b9156-ed60-4a2e-bc78-d054077849fc); Time taken: 0.236 seconds  INFO : Executing command(queryId=hive\_20170411080303\_741b9156-ed60-4a2e-bc78-d054077849fc): select \* from postimport\_sequence where username = '670063006100750074006F007500730065007200'  INFO : Query ID = hive\_20170411080303\_741b9156-ed60-4a2e-bc78-d054077849fc  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:15  INFO : Submitting tokens for job: job\_1490094830773\_0397  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0397/  INFO : Starting Job = job\_1490094830773\_0397, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0397/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0397  INFO : Hadoop job information for Stage-1: number of mappers: 15; number of reducers: 0  INFO : 2017-04-11 08:03:13,620 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 08:03:22,465 Stage-1 map = 7%, reduce = 0%, Cumulative CPU 2.2 sec  INFO : 2017-04-11 08:03:24,977 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 18.49 sec  INFO : 2017-04-11 08:03:26,359 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 57.31 sec  INFO : 2017-04-11 08:03:27,396 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 76.75 sec  INFO : 2017-04-11 08:03:28,677 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 94.08 sec  INFO : 2017-04-11 08:03:30,090 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 104.87 sec  INFO : 2017-04-11 08:03:31,278 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 122.61 sec  INFO : 2017-04-11 08:03:32,304 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 135.11 sec  INFO : 2017-04-11 08:03:33,452 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 146.79 sec  INFO : 2017-04-11 08:03:34,595 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 151.21 sec  INFO : 2017-04-11 08:03:36,790 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 162.04 sec  INFO : 2017-04-11 08:03:37,926 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 163.17 sec  INFO : MapReduce Total cumulative CPU time: 2 minutes 43 seconds 170 msec  INFO : Ended Job = job\_1490094830773\_0397  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 15 Cumulative CPU: 163.17 sec HDFS Read: 3803108005 HDFS Write: 2079806 SUCCESS  INFO : Total MapReduce CPU Time Spent: 2 minutes 43 seconds 170 msec  INFO : Completed executing command(queryId=hive\_20170411080303\_741b9156-ed60-4a2e-bc78-d054077849fc); Time taken: 30.566 seconds  INFO : OK |

# Sequence File format Bzip2 Compression

|  |
| --- |
| SET hive.exec.compress.output=true;  SET mapred.output.compression.type=BLOCK;  SET mapred.output.compression.codec=org.apache.hadoop.io.compress.BZip2Codec;  SET io.seqfile.compression.type=BLOCK; |

## Load data to Sequence Table

|  |
| --- |
| hive> insert overwrite table postimport\_sequence select \* from postimport\_text;  Query ID = shalaj\_20170411080505\_c38dcb93-94be-4615-acb1-db06233cf923  Total jobs = 3  Launching Job 1 out of 3  Number of reduce tasks is set to 0 since there's no reduce operator  Starting Job = job\_1490094830773\_0398, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0398/  Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0398  Hadoop job information for Stage-1: number of mappers: 74; number of reducers: 0  2017-04-11 08:05:14,380 Stage-1 map = 0%, reduce = 0%  2017-04-11 08:05:40,669 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 265.47 sec  2017-04-11 08:05:41,745 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 279.58 sec  2017-04-11 08:05:42,790 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 299.19 sec  2017-04-11 08:05:44,897 Stage-1 map = 6%, reduce = 0%, Cumulative CPU 322.11 sec  2017-04-11 08:05:45,952 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 342.78 sec  2017-04-11 08:05:49,117 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 388.07 sec  2017-04-11 08:05:55,459 Stage-1 map = 11%, reduce = 0%, Cumulative CPU 476.79 sec  2017-04-11 08:05:56,498 Stage-1 map = 12%, reduce = 0%, Cumulative CPU 488.86 sec  2017-04-11 08:05:59,672 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 522.15 sec  2017-04-11 08:06:00,706 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 544.67 sec  2017-04-11 08:06:01,735 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 546.39 sec  2017-04-11 08:06:02,779 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 549.72 sec  2017-04-11 08:06:03,812 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 557.39 sec  2017-04-11 08:06:20,809 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 749.31 sec  2017-04-11 08:06:21,841 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 772.12 sec  2017-04-11 08:06:27,276 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 840.54 sec  2017-04-11 08:06:28,318 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 869.16 sec  2017-04-11 08:06:29,350 Stage-1 map = 26%, reduce = 0%, Cumulative CPU 874.58 sec  2017-04-11 08:06:31,477 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 909.66 sec  2017-04-11 08:06:33,577 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 929.58 sec  2017-04-11 08:06:36,722 Stage-1 map = 31%, reduce = 0%, Cumulative CPU 975.53 sec  2017-04-11 08:06:37,784 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 999.81 sec  2017-04-11 08:06:44,042 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 1067.52 sec  2017-04-11 08:06:45,074 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 1080.53 sec  2017-04-11 08:06:47,133 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 1100.58 sec  2017-04-11 08:06:48,165 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 1120.42 sec  2017-04-11 08:06:49,200 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 1131.03 sec  2017-04-11 08:06:51,258 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 1140.59 sec  2017-04-11 08:06:53,325 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 1148.77 sec  2017-04-11 08:07:04,004 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 1297.63 sec  2017-04-11 08:07:11,275 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 1402.27 sec  2017-04-11 08:07:14,432 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 1445.14 sec  2017-04-11 08:07:15,466 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 1462.68 sec  2017-04-11 08:07:17,551 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 1487.65 sec  2017-04-11 08:07:18,590 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 1500.68 sec  2017-04-11 08:07:19,627 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 1509.26 sec  2017-04-11 08:07:21,720 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 1538.03 sec  2017-04-11 08:07:22,771 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 1557.91 sec  2017-04-11 08:07:26,938 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 1603.65 sec  2017-04-11 08:07:29,016 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 1635.74 sec  2017-04-11 08:07:31,082 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 1661.88 sec  2017-04-11 08:07:33,160 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 1676.91 sec  2017-04-11 08:07:34,188 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 1691.38 sec  2017-04-11 08:07:39,378 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 1745.37 sec  2017-04-11 08:07:43,553 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 1791.43 sec  2017-04-11 08:07:44,597 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 1808.76 sec  2017-04-11 08:07:50,854 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 1895.95 sec  2017-04-11 08:07:55,091 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 1954.81 sec  2017-04-11 08:07:57,196 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 1990.87 sec  2017-04-11 08:07:59,304 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 2008.02 sec  2017-04-11 08:08:00,345 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 2027.98 sec  2017-04-11 08:08:01,376 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 2033.23 sec  2017-04-11 08:08:03,501 Stage-1 map = 71%, reduce = 0%, Cumulative CPU 2063.81 sec  2017-04-11 08:08:06,687 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 2111.09 sec  2017-04-11 08:08:08,815 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 2135.2 sec  2017-04-11 08:08:13,024 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 2194.22 sec  2017-04-11 08:08:17,184 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 2230.8 sec  2017-04-11 08:08:18,240 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 2250.72 sec  2017-04-11 08:08:19,291 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 2253.91 sec  2017-04-11 08:08:20,337 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 2259.35 sec  2017-04-11 08:08:24,492 Stage-1 map = 82%, reduce = 0%, Cumulative CPU 2297.2 sec  2017-04-11 08:08:27,619 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 2344.96 sec  2017-04-11 08:08:33,871 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 2422.07 sec  2017-04-11 08:08:34,923 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 2432.47 sec  2017-04-11 08:08:39,083 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 2489.35 sec  2017-04-11 08:08:40,118 Stage-1 map = 88%, reduce = 0%, Cumulative CPU 2495.14 sec  2017-04-11 08:08:42,180 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 2518.57 sec  2017-04-11 08:08:44,249 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 2527.48 sec  2017-04-11 08:08:45,287 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 2547.82 sec  2017-04-11 08:08:48,415 Stage-1 map = 92%, reduce = 0%, Cumulative CPU 2580.27 sec  2017-04-11 08:08:50,508 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 2591.85 sec  2017-04-11 08:08:54,655 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 2638.56 sec  2017-04-11 08:08:55,680 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 2645.8 sec  2017-04-11 08:08:59,801 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 2678.36 sec  2017-04-11 08:09:00,829 Stage-1 map = 97%, reduce = 0%, Cumulative CPU 2685.86 sec  2017-04-11 08:09:01,851 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 2690.08 sec  2017-04-11 08:09:03,889 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 2696.38 sec  2017-04-11 08:09:04,911 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2700.84 sec  MapReduce Total cumulative CPU time: 45 minutes 0 seconds 840 msec  Ended Job = job\_1490094830773\_0398  Stage-4 is selected by condition resolver.  Stage-3 is filtered out by condition resolver.  Stage-5 is filtered out by condition resolver.  Moving data to: hdfs://mac55:8020/user/hive/warehouse/esplus.db/postimport\_sequence/.hive-staging\_hive\_2017-04-11\_08-05-04\_870\_3939172692376049850-1/-ext-10000  Loading data to table esplus.postimport\_sequence  Table esplus.postimport\_sequence stats: [numFiles=74, numRows=42286538, totalSize=1977982119, rawDataSize=19438765021]  MapReduce Jobs Launched:  Stage-Stage-1: Map: 74 Cumulative CPU: 2700.84 sec HDFS Read: 19486320244 HDFS Write: 1977989000 SUCCESS  Total MapReduce CPU Time Spent: 45 minutes 0 seconds 840 msec  OK  Time taken: 244.907 seconds |

## Data Size

|  |
| --- |
| hive> dfs -du -s -h /user/hive/warehouse/esplus.db/postimport\_sequence;  1.8 G 5.5 G /user/hive/warehouse/esplus.db/postimport\_sequence |

## Select some columns

SET hive.exec.compress.output=true;

|  |
| --- |
| select productname,username from postimport\_sequence;  INFO : Compiling command(queryId=hive\_20170411080909\_2ac776e2-568a-46cb-a20c-25c45b68963c): select productname,username from postimport\_sequence  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:productname, type:string, comment:null), FieldSchema(name:username, type:string, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411080909\_2ac776e2-568a-46cb-a20c-25c45b68963c); Time taken: 0.111 seconds  INFO : Executing command(queryId=hive\_20170411080909\_2ac776e2-568a-46cb-a20c-25c45b68963c): select productname,username from postimport\_sequence  INFO : Query ID = hive\_20170411080909\_2ac776e2-568a-46cb-a20c-25c45b68963c  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:8  INFO : Submitting tokens for job: job\_1490094830773\_0399  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0399/  INFO : Starting Job = job\_1490094830773\_0399, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0399/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0399  INFO : Hadoop job information for Stage-1: number of mappers: 8; number of reducers: 0  INFO : 2017-04-11 08:09:48,906 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 08:10:04,995 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 103.49 sec  INFO : 2017-04-11 08:10:08,335 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 128.01 sec  INFO : 2017-04-11 08:10:17,478 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 202.32 sec  INFO : 2017-04-11 08:10:19,661 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 220.42 sec  INFO : 2017-04-11 08:10:20,820 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 226.46 sec  INFO : 2017-04-11 08:10:26,470 Stage-1 map = 27%, reduce = 0%, Cumulative CPU 275.03 sec  INFO : 2017-04-11 08:10:28,786 Stage-1 map = 32%, reduce = 0%, Cumulative CPU 293.25 sec  INFO : 2017-04-11 08:10:29,809 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 300.13 sec  INFO : 2017-04-11 08:10:32,134 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 321.31 sec  INFO : 2017-04-11 08:10:35,477 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 342.52 sec  INFO : 2017-04-11 08:10:37,804 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 357.72 sec  INFO : 2017-04-11 08:10:41,130 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 382.02 sec  INFO : 2017-04-11 08:10:44,574 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 406.32 sec  INFO : 2017-04-11 08:10:46,741 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 421.49 sec  INFO : 2017-04-11 08:10:47,889 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 427.58 sec  INFO : 2017-04-11 08:10:50,192 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 448.83 sec  INFO : 2017-04-11 08:10:55,836 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 485.22 sec  INFO : 2017-04-11 08:10:56,992 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 491.29 sec  INFO : 2017-04-11 08:10:59,166 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 509.49 sec  INFO : 2017-04-11 08:11:02,613 Stage-1 map = 60%, reduce = 0%, Cumulative CPU 533.8 sec  INFO : 2017-04-11 08:11:08,229 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 573.37 sec  INFO : 2017-04-11 08:11:11,563 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 597.61 sec  INFO : 2017-04-11 08:11:13,878 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 612.79 sec  INFO : 2017-04-11 08:11:20,542 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 661.51 sec  INFO : 2017-04-11 08:11:22,881 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 673.67 sec  INFO : 2017-04-11 08:11:24,039 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 682.71 sec  INFO : 2017-04-11 08:11:26,223 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 700.95 sec  INFO : 2017-04-11 08:11:29,693 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 725.12 sec  INFO : 2017-04-11 08:11:31,862 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 734.18 sec  INFO : 2017-04-11 08:11:35,347 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 758.32 sec  INFO : 2017-04-11 08:11:36,501 Stage-1 map = 87%, reduce = 0%, Cumulative CPU 761.33 sec  INFO : 2017-04-11 08:11:39,837 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 780.63 sec  INFO : 2017-04-11 08:11:44,327 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 807.28 sec  INFO : 2017-04-11 08:11:47,802 Stage-1 map = 94%, reduce = 0%, Cumulative CPU 822.42 sec  INFO : 2017-04-11 08:11:51,156 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 834.53 sec  INFO : 2017-04-11 08:11:56,778 Stage-1 map = 96%, reduce = 0%, Cumulative CPU 858.78 sec  INFO : 2017-04-11 08:11:57,921 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 860.58 sec  INFO : 2017-04-11 08:11:59,066 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 862.96 sec  INFO : MapReduce Total cumulative CPU time: 14 minutes 22 seconds 960 msec  INFO : Ended Job = job\_1490094830773\_0399  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 8 Cumulative CPU: 862.96 sec HDFS Read: 1978029497 HDFS Write: 233691702 SUCCESS  INFO : Total MapReduce CPU Time Spent: 14 minutes 22 seconds 960 msec  INFO : Completed executing command(queryId=hive\_20170411080909\_2ac776e2-568a-46cb-a20c-25c45b68963c); Time taken: 138.292 seconds  INFO : OK |

## Check Max Date

|  |
| --- |
| select max(date) from postimport\_sequence;  INFO : Compiling command(queryId=hive\_20170411081414\_82c06f4a-a609-41c8-b499-c9cda8c4cda3): select max(date) from postimport\_sequence  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:\_c0, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411081414\_82c06f4a-a609-41c8-b499-c9cda8c4cda3); Time taken: 0.11 seconds  INFO : Executing command(queryId=hive\_20170411081414\_82c06f4a-a609-41c8-b499-c9cda8c4cda3): select max(date) from postimport\_sequence  INFO : Query ID = hive\_20170411081414\_82c06f4a-a609-41c8-b499-c9cda8c4cda3  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks determined at compile time: 1  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:8  INFO : Submitting tokens for job: job\_1490094830773\_0400  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0400/  INFO : Starting Job = job\_1490094830773\_0400, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0400/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0400  INFO : Hadoop job information for Stage-1: number of mappers: 8; number of reducers: 1  INFO : 2017-04-11 08:14:52,683 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 08:15:08,671 Stage-1 map = 5%, reduce = 0%, Cumulative CPU 89.72 sec  INFO : 2017-04-11 08:15:09,830 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 101.9 sec  INFO : 2017-04-11 08:15:17,824 Stage-1 map = 9%, reduce = 0%, Cumulative CPU 163.42 sec  INFO : 2017-04-11 08:15:20,145 Stage-1 map = 14%, reduce = 0%, Cumulative CPU 187.99 sec  INFO : 2017-04-11 08:15:21,300 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 194.06 sec  INFO : 2017-04-11 08:15:22,450 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 200.14 sec  INFO : 2017-04-11 08:15:29,241 Stage-1 map = 22%, reduce = 0%, Cumulative CPU 260.7 sec  INFO : 2017-04-11 08:15:30,397 Stage-1 map = 28%, reduce = 0%, Cumulative CPU 267.62 sec  INFO : 2017-04-11 08:15:33,727 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 294.86 sec  INFO : 2017-04-11 08:15:38,325 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 325.22 sec  INFO : 2017-04-11 08:15:42,808 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 358.57 sec  INFO : 2017-04-11 08:15:43,961 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 361.6 sec  INFO : 2017-04-11 08:15:45,117 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 367.67 sec  INFO : 2017-04-11 08:15:46,265 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 379.79 sec  INFO : 2017-04-11 08:15:47,289 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 388.88 sec  INFO : 2017-04-11 08:15:51,890 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 422.25 sec  INFO : 2017-04-11 08:15:55,227 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 443.53 sec  INFO : 2017-04-11 08:15:56,373 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 452.66 sec  INFO : 2017-04-11 08:15:57,524 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 458.71 sec  INFO : 2017-04-11 08:15:59,831 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 473.86 sec  INFO : 2017-04-11 08:16:04,306 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 507.23 sec  INFO : 2017-04-11 08:16:05,455 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 516.31 sec  INFO : 2017-04-11 08:16:06,609 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 522.35 sec  INFO : 2017-04-11 08:16:08,911 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 537.53 sec  INFO : 2017-04-11 08:16:12,239 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 558.7 sec  INFO : 2017-04-11 08:16:13,391 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 570.79 sec  INFO : 2017-04-11 08:16:16,841 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 591.98 sec  INFO : 2017-04-11 08:16:17,991 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 601.75 sec  INFO : 2017-04-11 08:16:19,014 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 613.9 sec  INFO : 2017-04-11 08:16:22,473 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 632.13 sec  INFO : 2017-04-11 08:16:23,625 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 638.17 sec  INFO : 2017-04-11 08:16:28,102 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 668.42 sec  INFO : 2017-04-11 08:16:29,256 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 674.52 sec  INFO : 2017-04-11 08:16:30,407 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 677.53 sec  INFO : 2017-04-11 08:16:31,556 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 688.59 sec  INFO : 2017-04-11 08:16:33,852 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 697.64 sec  INFO : 2017-04-11 08:16:34,875 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 706.33 sec  INFO : 2017-04-11 08:16:39,487 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 721.48 sec  INFO : 2017-04-11 08:16:40,637 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 730.58 sec  INFO : 2017-04-11 08:16:42,809 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 738.74 sec  INFO : 2017-04-11 08:16:47,385 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 748.33 sec  INFO : 2017-04-11 08:16:48,526 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 750.52 sec  INFO : 2017-04-11 08:16:51,844 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 751.47 sec  INFO : MapReduce Total cumulative CPU time: 12 minutes 31 seconds 470 msec  INFO : Ended Job = job\_1490094830773\_0400  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 8 Reduce: 1 Cumulative CPU: 751.47 sec HDFS Read: 1978036359 HDFS Write: 19 SUCCESS  INFO : Total MapReduce CPU Time Spent: 12 minutes 31 seconds 470 msec  INFO : Completed executing command(queryId=hive\_20170411081414\_82c06f4a-a609-41c8-b499-c9cda8c4cda3); Time taken: 129.195 seconds  INFO : OK |

## User wise Max Date

|  |
| --- |
| select username,max(date) from postimport\_sequence group by username;  INFO : Compiling command(queryId=hive\_20170411081717\_4d3fb689-e4d1-4289-9f04-d478a1a98de4): select username,max(date) from postimport\_sequence group by username  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:username, type:string, comment:null), FieldSchema(name:\_c1, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411081717\_4d3fb689-e4d1-4289-9f04-d478a1a98de4); Time taken: 0.247 seconds  INFO : Executing command(queryId=hive\_20170411081717\_4d3fb689-e4d1-4289-9f04-d478a1a98de4): select username,max(date) from postimport\_sequence group by username  INFO : Query ID = hive\_20170411081717\_4d3fb689-e4d1-4289-9f04-d478a1a98de4  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks not specified. Estimated from input data size: 30  INFO : In order to change the average load for a reducer (in bytes):  INFO : set hive.exec.reducers.bytes.per.reducer=<number>  INFO : In order to limit the maximum number of reducers:  INFO : set hive.exec.reducers.max=<number>  INFO : In order to set a constant number of reducers:  INFO : set mapreduce.job.reduces=<number>  INFO : number of splits:8  INFO : Submitting tokens for job: job\_1490094830773\_0401  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0401/  INFO : Starting Job = job\_1490094830773\_0401, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0401/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0401  INFO : Hadoop job information for Stage-1: number of mappers: 8; number of reducers: 30  INFO : 2017-04-11 08:17:57,438 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 08:18:13,578 Stage-1 map = 2%, reduce = 0%, Cumulative CPU 93.21 sec  INFO : 2017-04-11 08:18:15,902 Stage-1 map = 4%, reduce = 0%, Cumulative CPU 105.43 sec  INFO : 2017-04-11 08:18:17,061 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 120.7 sec  INFO : 2017-04-11 08:18:25,070 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 182.42 sec  INFO : 2017-04-11 08:18:26,229 Stage-1 map = 15%, reduce = 0%, Cumulative CPU 194.54 sec  INFO : 2017-04-11 08:18:27,388 Stage-1 map = 17%, reduce = 0%, Cumulative CPU 200.61 sec  INFO : 2017-04-11 08:18:34,218 Stage-1 map = 18%, reduce = 0%, Cumulative CPU 255.11 sec  INFO : 2017-04-11 08:18:36,532 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 273.26 sec  INFO : 2017-04-11 08:18:37,702 Stage-1 map = 24%, reduce = 0%, Cumulative CPU 288.41 sec  INFO : 2017-04-11 08:18:38,872 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 296.17 sec  INFO : 2017-04-11 08:18:40,031 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 302.23 sec  INFO : 2017-04-11 08:18:46,808 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 356.79 sec  INFO : 2017-04-11 08:18:49,107 Stage-1 map = 34%, reduce = 0%, Cumulative CPU 368.92 sec  INFO : 2017-04-11 08:18:50,258 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 378.05 sec  INFO : 2017-04-11 08:18:51,408 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 384.1 sec  INFO : 2017-04-11 08:18:52,557 Stage-1 map = 37%, reduce = 0%, Cumulative CPU 393.24 sec  INFO : 2017-04-11 08:18:55,892 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 420.61 sec  INFO : 2017-04-11 08:18:58,190 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 432.76 sec  INFO : 2017-04-11 08:19:00,488 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 447.96 sec  INFO : 2017-04-11 08:19:01,637 Stage-1 map = 42%, reduce = 0%, Cumulative CPU 456.99 sec  INFO : 2017-04-11 08:19:02,787 Stage-1 map = 43%, reduce = 0%, Cumulative CPU 463.05 sec  INFO : 2017-04-11 08:19:05,084 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 484.29 sec  INFO : 2017-04-11 08:19:09,550 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 511.59 sec  INFO : 2017-04-11 08:19:10,698 Stage-1 map = 46%, reduce = 0%, Cumulative CPU 520.71 sec  INFO : 2017-04-11 08:19:11,850 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 526.76 sec  INFO : 2017-04-11 08:19:13,002 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 532.83 sec  INFO : 2017-04-11 08:19:14,152 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 548.0 sec  INFO : 2017-04-11 08:19:19,777 Stage-1 map = 51%, reduce = 0%, Cumulative CPU 584.32 sec  INFO : 2017-04-11 08:19:22,072 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 596.41 sec  INFO : 2017-04-11 08:19:23,222 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 611.6 sec  INFO : 2017-04-11 08:19:25,519 Stage-1 map = 56%, reduce = 0%, Cumulative CPU 626.78 sec  INFO : 2017-04-11 08:19:27,816 Stage-1 map = 61%, reduce = 0%, Cumulative CPU 641.14 sec  INFO : 2017-04-11 08:19:32,291 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 671.34 sec  INFO : 2017-04-11 08:19:33,441 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 677.38 sec  INFO : 2017-04-11 08:19:34,591 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 683.43 sec  INFO : 2017-04-11 08:19:35,741 Stage-1 map = 66%, reduce = 0%, Cumulative CPU 689.46 sec  INFO : 2017-04-11 08:19:43,679 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 738.34 sec  INFO : 2017-04-11 08:19:44,828 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 744.37 sec  INFO : 2017-04-11 08:19:45,978 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 750.46 sec  INFO : 2017-04-11 08:19:47,131 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 759.53 sec  INFO : 2017-04-11 08:19:49,440 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 768.02 sec  INFO : 2017-04-11 08:19:53,912 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 791.4 sec  INFO : 2017-04-11 08:19:55,055 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 793.49 sec  INFO : 2017-04-11 08:20:00,504 Stage-1 map = 100%, reduce = 10%, Cumulative CPU 797.45 sec  INFO : 2017-04-11 08:20:01,713 Stage-1 map = 100%, reduce = 37%, Cumulative CPU 807.04 sec  INFO : 2017-04-11 08:20:02,878 Stage-1 map = 100%, reduce = 50%, Cumulative CPU 812.51 sec  INFO : 2017-04-11 08:20:05,432 Stage-1 map = 100%, reduce = 53%, Cumulative CPU 813.74 sec  INFO : 2017-04-11 08:20:06,739 Stage-1 map = 100%, reduce = 60%, Cumulative CPU 816.29 sec  INFO : 2017-04-11 08:20:07,987 Stage-1 map = 100%, reduce = 87%, Cumulative CPU 825.66 sec  INFO : 2017-04-11 08:20:09,153 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 831.17 sec  INFO : MapReduce Total cumulative CPU time: 13 minutes 51 seconds 170 msec  INFO : Ended Job = job\_1490094830773\_0401  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 8 Reduce: 30 Cumulative CPU: 831.17 sec HDFS Read: 1978149993 HDFS Write: 32473 SUCCESS  INFO : Total MapReduce CPU Time Spent: 13 minutes 51 seconds 170 msec  INFO : Completed executing command(queryId=hive\_20170411081717\_4d3fb689-e4d1-4289-9f04-d478a1a98de4); Time taken: 140.94 seconds  INFO : OK |

## Select All Columns

|  |
| --- |
| select \* from postimport\_sequence  INFO : Compiling command(queryId=hive\_20170411082020\_d56157fc-0b1c-4cd6-9812-d9ff4042f70c): select \* from postimport\_sequence  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_sequence.username, type:string, comment:null), FieldSchema(name:postimport\_sequence.machinename, type:string, comment:null), FieldSchema(name:postimport\_sequence.machineid, type:string, comment:null), FieldSchema(name:postimport\_sequence.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_sequence.domainname, type:string, comment:null), FieldSchema(name:postimport\_sequence.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_sequence.modulename, type:string, comment:null), FieldSchema(name:postimport\_sequence.filename, type:string, comment:null), FieldSchema(name:postimport\_sequence.productname, type:string, comment:null), FieldSchema(name:postimport\_sequence.companyname, type:string, comment:null), FieldSchema(name:postimport\_sequence.description, type:string, comment:null), FieldSchema(name:postimport\_sequence.text, type:string, comment:null), FieldSchema(name:postimport\_sequence.activetime, type:double, comment:null), FieldSchema(name:postimport\_sequence.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_sequence.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_sequence.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411082020\_d56157fc-0b1c-4cd6-9812-d9ff4042f70c); Time taken: 0.32 seconds  INFO : Executing command(queryId=hive\_20170411082020\_d56157fc-0b1c-4cd6-9812-d9ff4042f70c): select \* from postimport\_sequence  INFO : Query ID = hive\_20170411082020\_d56157fc-0b1c-4cd6-9812-d9ff4042f70c  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:8  INFO : Submitting tokens for job: job\_1490094830773\_0402  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0402/  INFO : Starting Job = job\_1490094830773\_0402, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0402/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0402  INFO : Hadoop job information for Stage-1: number of mappers: 8; number of reducers: 0  INFO : 2017-04-11 08:20:45,336 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 08:21:16,481 Stage-1 map = 1%, reduce = 0%, Cumulative CPU 225.29 sec  INFO : 2017-04-11 08:21:19,947 Stage-1 map = 8%, reduce = 0%, Cumulative CPU 249.88 sec  INFO : 2017-04-11 08:21:21,102 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 256.0 sec  INFO : 2017-04-11 08:21:22,256 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 274.45 sec  INFO : 2017-04-11 08:21:43,950 Stage-1 map = 16%, reduce = 0%, Cumulative CPU 446.53 sec  INFO : 2017-04-11 08:21:46,259 Stage-1 map = 19%, reduce = 0%, Cumulative CPU 464.89 sec  INFO : 2017-04-11 08:21:47,423 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 471.03 sec  INFO : 2017-04-11 08:21:48,580 Stage-1 map = 21%, reduce = 0%, Cumulative CPU 477.16 sec  INFO : 2017-04-11 08:21:50,892 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 501.67 sec  INFO : 2017-04-11 08:21:53,196 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 520.09 sec  INFO : 2017-04-11 08:22:10,411 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 664.31 sec  INFO : 2017-04-11 08:22:13,885 Stage-1 map = 33%, reduce = 0%, Cumulative CPU 691.72 sec  INFO : 2017-04-11 08:22:15,041 Stage-1 map = 35%, reduce = 0%, Cumulative CPU 699.31 sec  INFO : 2017-04-11 08:22:18,398 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 720.75 sec  INFO : 2017-04-11 08:22:19,553 Stage-1 map = 38%, reduce = 0%, Cumulative CPU 736.08 sec  INFO : 2017-04-11 08:22:23,029 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 757.6 sec  INFO : 2017-04-11 08:22:36,790 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 849.51 sec  INFO : 2017-04-11 08:22:37,947 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 864.79 sec  INFO : 2017-04-11 08:22:41,415 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 886.23 sec  INFO : 2017-04-11 08:22:46,042 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 913.8 sec  INFO : 2017-04-11 08:22:50,556 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 950.57 sec  INFO : 2017-04-11 08:22:52,867 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 972.07 sec  INFO : 2017-04-11 08:23:00,965 Stage-1 map = 49%, reduce = 0%, Cumulative CPU 1021.15 sec  INFO : 2017-04-11 08:23:02,122 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 1036.39 sec  INFO : 2017-04-11 08:23:07,779 Stage-1 map = 52%, reduce = 0%, Cumulative CPU 1076.24 sec  INFO : 2017-04-11 08:23:11,250 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 1100.65 sec  INFO : 2017-04-11 08:23:15,875 Stage-1 map = 54%, reduce = 0%, Cumulative CPU 1128.14 sec  INFO : 2017-04-11 08:23:20,509 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 1165.09 sec  INFO : 2017-04-11 08:23:22,688 Stage-1 map = 57%, reduce = 0%, Cumulative CPU 1180.42 sec  INFO : 2017-04-11 08:23:23,843 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 1186.53 sec  INFO : 2017-04-11 08:23:27,307 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 1214.12 sec  INFO : 2017-04-11 08:23:37,731 Stage-1 map = 62%, reduce = 0%, Cumulative CPU 1284.74 sec  INFO : 2017-04-11 08:23:42,226 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 1321.47 sec  INFO : 2017-04-11 08:23:44,538 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 1336.79 sec  INFO : 2017-04-11 08:23:50,318 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 1379.63 sec  INFO : 2017-04-11 08:23:54,810 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 1407.15 sec  INFO : 2017-04-11 08:24:05,213 Stage-1 map = 70%, reduce = 0%, Cumulative CPU 1486.67 sec  INFO : 2017-04-11 08:24:08,675 Stage-1 map = 72%, reduce = 0%, Cumulative CPU 1507.95 sec  INFO : 2017-04-11 08:24:09,833 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 1514.07 sec  INFO : 2017-04-11 08:24:20,080 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 1590.61 sec  INFO : 2017-04-11 08:24:21,231 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 1593.66 sec  INFO : 2017-04-11 08:24:22,392 Stage-1 map = 77%, reduce = 0%, Cumulative CPU 1599.79 sec  INFO : 2017-04-11 08:24:29,176 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 1654.86 sec  INFO : 2017-04-11 08:24:32,645 Stage-1 map = 80%, reduce = 0%, Cumulative CPU 1679.26 sec  INFO : 2017-04-11 08:24:37,255 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 1706.88 sec  INFO : 2017-04-11 08:24:47,502 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 1786.46 sec  INFO : 2017-04-11 08:24:48,659 Stage-1 map = 84%, reduce = 0%, Cumulative CPU 1792.58 sec  INFO : 2017-04-11 08:24:50,970 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 1808.65 sec  INFO : 2017-04-11 08:24:53,274 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 1823.92 sec  INFO : 2017-04-11 08:24:59,032 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 1854.58 sec  INFO : 2017-04-11 08:25:03,510 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 1888.16 sec  INFO : 2017-04-11 08:25:09,278 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 1920.29 sec  INFO : 2017-04-11 08:25:20,690 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 1981.55 sec  INFO : 2017-04-11 08:25:24,174 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 1996.98 sec  INFO : 2017-04-11 08:25:25,315 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 2002.57 sec  INFO : 2017-04-11 08:25:46,872 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 2047.2 sec  INFO : 2017-04-11 08:25:50,256 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2053.18 sec  INFO : MapReduce Total cumulative CPU time: 34 minutes 13 seconds 180 msec  INFO : Ended Job = job\_1490094830773\_0402  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 8 Cumulative CPU: 2053.18 sec HDFS Read: 1978037273 HDFS Write: 3633963344 SUCCESS  INFO : Total MapReduce CPU Time Spent: 34 minutes 13 seconds 180 msec  INFO : Completed executing command(queryId=hive\_20170411082020\_d56157fc-0b1c-4cd6-9812-d9ff4042f70c); Time taken: 314.03 seconds  INFO : OK |

## Select All Columns with filter Condition

|  |
| --- |
| select \* from postimport\_sequence where username = '670063006100750074006F007500730065007200';  INFO : Compiling command(queryId=hive\_20170411083030\_55e8a09a-52ba-4673-bddf-19b95c38f8aa): select \* from postimport\_sequence where username = '670063006100750074006F007500730065007200'  INFO : Semantic Analysis Completed  INFO : Returning Hive schema: Schema(fieldSchemas:[FieldSchema(name:postimport\_sequence.username, type:string, comment:null), FieldSchema(name:postimport\_sequence.machinename, type:string, comment:null), FieldSchema(name:postimport\_sequence.machineid, type:string, comment:null), FieldSchema(name:postimport\_sequence.ipaddress, type:string, comment:null), FieldSchema(name:postimport\_sequence.domainname, type:string, comment:null), FieldSchema(name:postimport\_sequence.pcmlogeventid, type:string, comment:null), FieldSchema(name:postimport\_sequence.modulename, type:string, comment:null), FieldSchema(name:postimport\_sequence.filename, type:string, comment:null), FieldSchema(name:postimport\_sequence.productname, type:string, comment:null), FieldSchema(name:postimport\_sequence.companyname, type:string, comment:null), FieldSchema(name:postimport\_sequence.description, type:string, comment:null), FieldSchema(name:postimport\_sequence.text, type:string, comment:null), FieldSchema(name:postimport\_sequence.activetime, type:double, comment:null), FieldSchema(name:postimport\_sequence.datetime, type:timestamp, comment:null), FieldSchema(name:postimport\_sequence.timeslot, type:timestamp, comment:null), FieldSchema(name:postimport\_sequence.date, type:date, comment:null)], properties:null)  INFO : Completed compiling command(queryId=hive\_20170411083030\_55e8a09a-52ba-4673-bddf-19b95c38f8aa); Time taken: 0.137 seconds  INFO : Executing command(queryId=hive\_20170411083030\_55e8a09a-52ba-4673-bddf-19b95c38f8aa): select \* from postimport\_sequence where username = '670063006100750074006F007500730065007200'  INFO : Query ID = hive\_20170411083030\_55e8a09a-52ba-4673-bddf-19b95c38f8aa  INFO : Total jobs = 1  INFO : Launching Job 1 out of 1  INFO : Starting task [Stage-1:MAPRED] in serial mode  INFO : Number of reduce tasks is set to 0 since there's no reduce operator  INFO : number of splits:8  INFO : Submitting tokens for job: job\_1490094830773\_0403  INFO : The url to track the job: http://mac55:8088/proxy/application\_1490094830773\_0403/  INFO : Starting Job = job\_1490094830773\_0403, Tracking URL = http://mac55:8088/proxy/application\_1490094830773\_0403/  INFO : Kill Command = /opt/cloudera/parcels/CDH-5.8.2-1.cdh5.8.2.p0.3/lib/hadoop/bin/hadoop job -kill job\_1490094830773\_0403  INFO : Hadoop job information for Stage-1: number of mappers: 8; number of reducers: 0  INFO : 2017-04-11 08:30:41,545 Stage-1 map = 0%, reduce = 0%  INFO : 2017-04-11 08:30:56,741 Stage-1 map = 3%, reduce = 0%, Cumulative CPU 83.09 sec  INFO : 2017-04-11 08:30:57,904 Stage-1 map = 10%, reduce = 0%, Cumulative CPU 95.44 sec  INFO : 2017-04-11 08:30:59,068 Stage-1 map = 13%, reduce = 0%, Cumulative CPU 101.61 sec  INFO : 2017-04-11 08:31:07,215 Stage-1 map = 20%, reduce = 0%, Cumulative CPU 169.94 sec  INFO : 2017-04-11 08:31:09,538 Stage-1 map = 23%, reduce = 0%, Cumulative CPU 195.65 sec  INFO : 2017-04-11 08:31:10,701 Stage-1 map = 25%, reduce = 0%, Cumulative CPU 201.73 sec  INFO : 2017-04-11 08:31:15,370 Stage-1 map = 29%, reduce = 0%, Cumulative CPU 245.26 sec  INFO : 2017-04-11 08:31:17,694 Stage-1 map = 30%, reduce = 0%, Cumulative CPU 253.15 sec  INFO : 2017-04-11 08:31:18,856 Stage-1 map = 36%, reduce = 0%, Cumulative CPU 268.37 sec  INFO : 2017-04-11 08:31:20,019 Stage-1 map = 39%, reduce = 0%, Cumulative CPU 274.42 sec  INFO : 2017-04-11 08:31:24,660 Stage-1 map = 40%, reduce = 0%, Cumulative CPU 311.39 sec  INFO : 2017-04-11 08:31:26,986 Stage-1 map = 41%, reduce = 0%, Cumulative CPU 323.51 sec  INFO : 2017-04-11 08:31:28,146 Stage-1 map = 44%, reduce = 0%, Cumulative CPU 333.1 sec  INFO : 2017-04-11 08:31:29,307 Stage-1 map = 45%, reduce = 0%, Cumulative CPU 339.1 sec  INFO : 2017-04-11 08:31:30,460 Stage-1 map = 47%, reduce = 0%, Cumulative CPU 354.22 sec  INFO : 2017-04-11 08:31:32,771 Stage-1 map = 48%, reduce = 0%, Cumulative CPU 360.45 sec  INFO : 2017-04-11 08:31:33,922 Stage-1 map = 50%, reduce = 0%, Cumulative CPU 375.58 sec  INFO : 2017-04-11 08:31:37,413 Stage-1 map = 53%, reduce = 0%, Cumulative CPU 397.22 sec  INFO : 2017-04-11 08:31:39,733 Stage-1 map = 55%, reduce = 0%, Cumulative CPU 418.7 sec  INFO : 2017-04-11 08:31:40,893 Stage-1 map = 58%, reduce = 0%, Cumulative CPU 424.76 sec  INFO : 2017-04-11 08:31:42,051 Stage-1 map = 59%, reduce = 0%, Cumulative CPU 430.84 sec  INFO : 2017-04-11 08:31:45,556 Stage-1 map = 63%, reduce = 0%, Cumulative CPU 461.27 sec  INFO : 2017-04-11 08:31:49,038 Stage-1 map = 64%, reduce = 0%, Cumulative CPU 482.51 sec  INFO : 2017-04-11 08:31:50,199 Stage-1 map = 65%, reduce = 0%, Cumulative CPU 488.59 sec  INFO : 2017-04-11 08:31:51,359 Stage-1 map = 67%, reduce = 0%, Cumulative CPU 497.7 sec  INFO : 2017-04-11 08:31:53,690 Stage-1 map = 68%, reduce = 0%, Cumulative CPU 509.83 sec  INFO : 2017-04-11 08:31:54,849 Stage-1 map = 69%, reduce = 0%, Cumulative CPU 525.08 sec  INFO : 2017-04-11 08:31:58,326 Stage-1 map = 73%, reduce = 0%, Cumulative CPU 546.37 sec  INFO : 2017-04-11 08:31:59,488 Stage-1 map = 74%, reduce = 0%, Cumulative CPU 552.42 sec  INFO : 2017-04-11 08:32:00,651 Stage-1 map = 75%, reduce = 0%, Cumulative CPU 567.57 sec  INFO : 2017-04-11 08:32:01,809 Stage-1 map = 76%, reduce = 0%, Cumulative CPU 573.63 sec  INFO : 2017-04-11 08:32:02,971 Stage-1 map = 78%, reduce = 0%, Cumulative CPU 576.65 sec  INFO : 2017-04-11 08:32:04,136 Stage-1 map = 79%, reduce = 0%, Cumulative CPU 589.41 sec  INFO : 2017-04-11 08:32:07,618 Stage-1 map = 81%, reduce = 0%, Cumulative CPU 607.63 sec  INFO : 2017-04-11 08:32:09,941 Stage-1 map = 83%, reduce = 0%, Cumulative CPU 626.21 sec  INFO : 2017-04-11 08:32:11,095 Stage-1 map = 85%, reduce = 0%, Cumulative CPU 632.26 sec  INFO : 2017-04-11 08:32:12,262 Stage-1 map = 86%, reduce = 0%, Cumulative CPU 638.29 sec  INFO : 2017-04-11 08:32:15,750 Stage-1 map = 89%, reduce = 0%, Cumulative CPU 662.48 sec  INFO : 2017-04-11 08:32:18,071 Stage-1 map = 90%, reduce = 0%, Cumulative CPU 674.59 sec  INFO : 2017-04-11 08:32:19,233 Stage-1 map = 91%, reduce = 0%, Cumulative CPU 681.2 sec  INFO : 2017-04-11 08:32:20,392 Stage-1 map = 93%, reduce = 0%, Cumulative CPU 687.31 sec  INFO : 2017-04-11 08:32:22,704 Stage-1 map = 95%, reduce = 0%, Cumulative CPU 698.03 sec  INFO : 2017-04-11 08:32:23,905 Stage-1 map = 98%, reduce = 0%, Cumulative CPU 706.91 sec  INFO : 2017-04-11 08:32:28,539 Stage-1 map = 99%, reduce = 0%, Cumulative CPU 715.36 sec  INFO : 2017-04-11 08:32:30,814 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 719.92 sec  INFO : MapReduce Total cumulative CPU time: 11 minutes 59 seconds 920 msec  INFO : Ended Job = job\_1490094830773\_0403  INFO : MapReduce Jobs Launched:  INFO : Stage-Stage-1: Map: 8 Cumulative CPU: 719.92 sec HDFS Read: 1978041049 HDFS Write: 2075382 SUCCESS  INFO : Total MapReduce CPU Time Spent: 11 minutes 59 seconds 920 msec  INFO : Completed executing command(queryId=hive\_20170411083030\_55e8a09a-52ba-4673-bddf-19b95c38f8aa); Time taken: 116.356 seconds  INFO : OK |

# Summary

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Load Query | Table Size | Select few Columns | Max Date | User wise Max  date | | Select All  Columns | Select All Columns with filter condition |
| File Format | **Time in Sec** | **Size in GB** | **Time in Sec** | **Time in Sec** | **Time in Sec** | **Time in Sec** | | **Time in Sec** |
| Text |  |  |  |  |  |  | |  |
| None Compression | [1013.319](#_Load_data_to) | [18.1](#_Select_some_columns_1) | [181.884](#_Select_some_columns_1) | [136](#_Check_Max_Date) | [220.655](#_User_wise_Max) | [994.86](#_Select_All_Columns) | | [166.443](#_Select_All_Columns_5) |
| Gzip Compression | [266.638](#_Load_data_to_2) | [3.4](#_Data_Size_10) | [64.919](#_Select_some_columns_17) | [62.306](#_Check_Max_Date_16) | [85.79](#_User_wise_Max_18) | [208.037](#_Select_All_Columns_7) | | [57.103](#_Select_All_Columns_29) |
| Bzip2 Compression | [246.877](#_Load_data_to_15) | [1.6](#_Data_Size_10) | [157.395](#_Select_some_columns_18) | [154.203](#_Check_Max_Date_17) | [166.282](#_User_wise_Max_19) | [248.012](#_Select_All_Columns_1) | | [146.791](#_Select_All_Columns_30) |
| RC |  |  |  |  |  |  | |  |
| Snappy Compression | [338.984](#_Load_data_to_7) | [5.1](#_Select_some_columns) | [46.076](#_Select_some_columns_15) | [57.372](#_Check_Max_Date_1) | [79.277](#_User_wise_Max_2) | [205.725](#_Select_All_Columns_1) | | [43.421](#_Select_All_Columns_6) |
| Gzip Compression | [175.146](#_Load_data_to_11) | [2.1](#_Data_Size_6) | [31.122](#_Select_some_columns_15) | [32.797](#_Check_Max_Date_11) | [41.184](#_User_wise_Max_11) | [224.2](#_Select_All_Columns_21) | | [26.082](#_Select_All_Columns_22) |
| ORC |  |  |  |  |  |  | |  |
| Zlib Compression | [235.644](#_Load_data_to_3) | [1.1](#_Select_some_columns_3) | [32.097](#_Select_some_columns_2) | [15.388](#_Check_Max_Date_3) | [25.749](#_User_wise_Max_15) | [282.546](#_Select_All_Columns_2) | | [30.267](#_Select_All_Columns_10) |
| Snappy Compression | [167.416](#_Load_data_to_4) | [1.9](#_Select_some_columns_4) | [28.702](#_Select_some_columns_16) | [21.181](#_Check_Max_Date_4) | [29.488](#_User_wise_Max_17) | [167.884](#_Select_All_Columns_4) | | [20.417](#_Select_All_Columns_8) |
| None Compression | [264.138](#_Load_data_to_5) | [3.8](#_Data_Size) | [25.009](#_Select_some_columns_5) | [22.393](#_Check_Max_Date_5) | [45.292](#_User_wise_Max_16) | [166.84](#_Select_All_Columns_4) | | [41.737](#_Select_All_Columns_9) |
| Parquet |  |  |  |  |  |  | |  |
| Default Compression | [338.457](#_Load_data_to_6) | [4.6](#_Data_Size_1) | [34.516](#_Select_some_columns_6) | [35.166](#_Check_Max_Date_2) | [58.89](#_User_wise_Max_6) | [228.169](#_Select_All_Columns_11) | | [51.371](#_Select_All_Columns_12) |
| Gzip Compression | [161.73](#_Load_data_to_9) | [1.4](#_Data_Size_4) | [37.26](#_Select_some_columns_9) | [24.164](#_Check_Max_Date_9) | [35.62](#_User_wise_Max_9) | [220.914](#_Select_All_Columns_17) | | [31.779](#_Select_All_Columns_18) |
| Snappy Compression | [183.337](#_Load_data_to_10) | [2.5](#_Data_Size_5) | [31.229](#_Select_some_columns_10) | [24.78](#_Check_Max_Date_15) | [39.453](#_User_wise_Max_10) | [200.57](#_Select_All_Columns_19) | | [40.857](#_Select_All_Columns_20) |
| Avro |  |  |  |  |  |  | |  |
| Default Compression | [338.742](#_Load_data_to_7) | [4.6](#_Data_Size_2) | [81.219](#_Select_some_columns_7) | [78.305](#_Check_Max_Date_7) | [103.471](#_User_wise_Max_7) | [200.149](#_Select_All_Columns_13) | | [69.397](#_Select_All_Columns_14) |
| Snappy Compression | [361.001](#_Load_data_to_8) | [6.2](#_Data_Size_3) | [65.855](#_Select_some_columns_8) | [64.0](#_Check_Max_Date_8) | [101.256](#_User_wise_Max_8) | [230.673](#_Select_All_Columns_15) | | [59.46](#_Select_All_Columns_16) |
| Sequence |  |  |  |  |  |  | |  |
| Snappy Compression | [399.632](#_Load_data_to_12) | [6.6](#_Data_Size_7) | [65.993](#_Select_some_columns_12) | [62.074](#_Check_Max_Date_12) | [92.388](#_User_wise_Max_14) | [213.097](#_Select_All_Columns_23) | | [40.242](#_Select_All_Columns_24) |
| Gzip Compression | [280.252](#_Load_data_to_13) | [3.5](#_Data_Size_8) | [40.481](#_Select_some_columns_13) | [39.042](#_Check_Max_Date_13) | [76.483](#_User_wise_Max_12) | [236.055](#_Select_All_Columns_25) | | [30.566](#_Select_All_Columns_26) |
| Bzip2 Compression | [244.907](#_Load_data_to_14) | [1.8](#_Data_Size_9) | [138.292](#_Select_some_columns_14) | [129.195](#_Check_Max_Date_14) | [140.94](#_User_wise_Max_13) | [314.03](#_Select_All_Columns_27) | | [116.356](#_Select_All_Columns_28) |

***Note: For parquet we dropped date column as it does not support date data type and for Avro we dropped timestamp columns as it does not support timestamp datatype***

***Note: ORC format not supported by Impala***

# Conclusion

* When you have tables with very large number of columns and you tend to use specific columns frequently, RC/ORC file format would be a good choice. Rather than reading the entire row of data you would just retrieve the required columns, thus saving time. The data is divided into groups of rows, which are then divided into groups of columns.
* RC/ORC format is not a good choice when you are dealing with all columns without filer criteria.
* If you are using filter in select query , ORC will give better performance , since ORC stores min and max value of each column in every file it can identify if value lies within the min and max range if it doesn’t fit under range it will skip row group at once.
* ORC is a good choice for aggregate function as each file contains aggregation like min, max, count, and sum for each column
* This benchmarks indicate that ORC files compress to be the smallest of all file formats in Hadoop
* Parquet file format is good choice when you are dealing with few columns in query.
* Avro file format is good choice when table schema frequently changes.

