1. 固定夠大的離群值範圍，調整中央比例

range1(0.0, 10000.0);

range2(1000000.0, 1001000.0);

range3(100000000.0, 100010000.0);

一次50000000筆測資，逐筆進行查詢，以上重複100次取平均，batch 50000

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | #1 | #2 | #3 | avg |
| 1 | 3039635793 | 2446827354 | 2992283977 | 2826249041 |
| 0.99 | 5097959726 | 5576117456 | 5094744338 | 5256273840 |
| 0.98 | 6246057547 | 5272366402 | 4839776028 | 5452733326 |
| 0.97 | 5484442508 | 5736985536 | 5276483369 | 5499303804 |
| 0.96 | 5352685378 | 4852805076 | 5155784700 | 5120425051 |
| 0.95 | 5786555924 | 4851499238 | 5399721174 | 5345925445 |
| 0.94 | 5826281067 | 5039847234 | 5512476647 | 5459534983 |
| 0.93 | 5391011662 | 5070504562 | 5270848788 | 5244121671 |
| 0.92 | 5616522392 | 5669480806 | 5290534360 | 5525512519 |
| 0.91 | 4700264256 | 5708435353 | 5118610689 | 5525512519 |
| 0.90 | 5588854295 | 4721575485 | 4655182255 | 4988537345 |
| 0.8 | 5268536442 | 5446622633 | 4766483922 | 5160547666 |
| 0.7 | 5401417874 | 4855471037 | 4747598139 | 5001495683 |
| 0.6 | 5100737924 | 5118418260 | 4961041112 | 5060065765 |
| 0.5 | 4594921503 | 4595381663 | 4826850892 | 4672384686 |

2億data

range1(0.0, 10000.0);

range2(1000000.0, 1001000.0);

range3(100000000.0, 100010000.0);

|  |  |
| --- | --- |
|  | Test1 200M |
| 1 | 11861787907 |
| 0.99 | 20187296508 |
| 0.98 | 19819837762 |
| 0.97 | 27524320069 |
| 0.96 | 27592656384 |
| 0.95 | 22889190429 |
| 0.94 | 26917516689 |

1. 中央部分加入多層隨機篩選，避免太完美

range1(0.0, 10000.0);

range2\_1(1000000.0, 1000100.0);

range2\_2(1000100.0, 1000200.0);

range2\_3(1000200.0, 1000300.0);

range2\_4(1000300.0, 1000400.0);

range2\_5(1000400.0, 1000500.0);

range2\_1(1000500.0, 1000600.0);

range2\_2(1000600.0, 1000700.0);

range2\_3(1000700.0, 1000800.0);

range2\_4(1000800.0, 1000900.0);

range2\_5(1000900.0, 1001000.0);

range3(100000000.0, 100010000.0);

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | #1 | #2 | #3 | avg |
| 1 | 2949424915 | 2451335552 | 2553331238 | 2651363902 |
| 0.99 | 5539284051 | 5349872109 | 5252351171 | 5380502444 |
| 0.98 | 5828312115 | 5070927638 | 4596973071 | 5165404275 |
| 0.97 | 4883614667 | 5406240718 | 5753925732 | 5347927039 |
| 0.96 | 5363409408 | 5519692963 | 5218563587 | 5367221986 |
| 0.95 | 5396831185 | 4792406572 | 5488617583 | 5225951780 |
| 0.94 | 5173772649 | 5215564539 | 5951759361 | 5447032183 |
| 0.93 | 4924798836 | 5111687224 | 5951170334 | 5329218798 |
| 0.92 | 5124153055 | 5525037482 | 5525037482 | 5391409340 |
| 0.91 | 5851642335 | 5879459127 | 4919518476 | 5550206646 |
| 0.90 | 5803552725 | 5333573985 | 5505586627 | 5547571112 |
| 0.8 | 5462073407 | 5021995815 | 5046663375 | 5176910866 |
| 0.7 | 5660129587 | 5026371870 | 5288782842 | 5325094766 |
| 0.6 | 4691475898 | 5353676934 | 4998158485 | 5014437106 |
| 0.5 | 5590966406 | 5004709093 | 5131568246 | 5242414582 |

1. Bulk load中央data，離群值data用insert方式加入

range1(0.0, 10000.0);

range2(1000000.0, 1001000.0);

range3(100000000.0, 100010000.0);

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | #1 | #2 | #3 | avg |
| 1 | 2173091243 | 2671689898 | 2570231854 | 2471670998 |
| 0.99 | 2978632975 | 2681990297 | 2871076122 | 2843899798 |
| 0.98 | 2901326834 | 2770668224 | 2745953786 | 2805982948 |
| 0.97 | 2719447053 | 2909285651 | 2948291255 | 2859007986 |
| 0.96 | 3659661159 | 2960899885 | 2727557112 | 3116039385 |
| 0.95 | 3080626553 | 3047663776 | 3073222870 | 3067171066 |
| 0.94 | 2838573537 | 2460801728 | 2747846481 | 2682407249 |
| 0.93 | 2501888901 | 2547939708 | 2865480032 | 2638436214 |
| 0.92 | 2543792164 | 3025297545 | 2406080112 | 2658389940 |
| 0.91 | 2687009209 | 2522800518 | 3907777152 | 3039195626 |
| 0.90 | 3045461155 | 3345870716 | 3602917068 | 3331416313 |
| 0.8 | 4478866409 | 3894017503 | 4952001874 | 4441628595 |
| 0.7 | 5393037930  7496337194 | 5003603581  4842768314 | 7076440321  6469313143 | 5824360611 |
| 0.6 | 5285020629 | 5748418675 | 5781303470 | 5604914258 |
| 0.5 | 4203917694 | 3850567284 | 4735912416 | 4263465798 |

1. origin

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | #1 | #2 | #3 | avg |
| 1 | 2015215035 | 2252135925 | 2097257816 |  |
| 0.99 | 3545690822 | 3443654470 | 3494606885 |  |
| 0.98 | 3705831276 | 4102885991 | 3671539741 |  |
| 0.97 | 3655180225 | 3362503938 | 3814086834 |  |
| 0.96 | 3583813934 | 3640889058 | 3318546520 |  |
| 0.95 | 3802576396 | 3368727438 | 3460127419 |  |
| 0.94 | 3997196957 | 3543331136 | 3319264529 |  |
| 0.93 | 3803480141 | 3553202947 | 4070823228 |  |
| 0.92 | 4046198851 | 3416567279 | 4064040291 |  |
| 0.91 | 4323421931 | 4089157632 | 3938543635 |  |
| 0.90 | 5124822026 | 4397760168 | 3328465546 |  |
| 0.8 | 8107640930 | 7747771160 | 6683871201 |  |
| 0.7 | 7636870822 | 6597429857 | 6644367110 |  |
| 0.6 | 5490300881 | 5449450078 | 5362879909 |  |
| 0.5 | 4199285668 | 2935398186 | 4010202054 |  |

range1(0.0, 100.0);

range2(10000.0, 11000.0);

range3(1000000.0, 1000100.0);

1. 藍&灰載入資料的時間

lookup time:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | #1 | #2 | #3 | avg |
| Cont. | 1495792661 | 1501888151 | 1271207078 | 1422962630 |
| 1 | 2712185569 | 2450522289 | 3022676009 | 2728461289 |
| 0.9 | 3934907757 | 3958770610 | 3727375170 | 3873684512 |
| 0.8 | 6015372660 | 5233417887 | 5285721703 | 5511504083 |
| 0.7 | 3506304943 | 4513316544 | 3671641342 | 3897087610 |
| 0.6 | 6800850493 | 6527243303 | 4067819492 | 5798637763 |
| 0.5 | 4316154298 | 3882767653 | 4347465110 | 4182129020 |

Load time:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | #1 | #2 | #3 | avg |
| Cont. | 3427543000 | 3412548949 | 2913197630 | 3251096526 |
| 1 | 19196913652 | 20067495565 | 20239139136 | 19834516118 |
| 0.9 | 57463529384 | 62443517651 | 56661881812 | 58856309616 |
| 0.8 | 117654870635 | 117522072745 | 117057326038 | 1.17411E+11 |
| 0.7 | 173741116095 | 174710208285 | 177566537426 | 1.75339E+11 |
| 0.6 | 251817580748 | 251456863834 | 241703046755 | 2.48326E+11 |
| 0.5 | 308320858479 | 301010313875 | 312810493337 | 3.07381E+11 |

Expand&retrain times:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | #1 | #2 | #3 | avg |
| 1 | 0 | 0 | 0 |  |
| 0.9 | 19.68 | 17.88 | 17.22 |  |
| 0.8 | 17.18 | 18.31 | 17.79 |  |
| 0.7 | 18.12 | 18.31 | 17.72 |  |
| 0.6 | 19.72 | 19.73 | 19.96 |  |
| 0.5 | 19.80 | 20.27 | 20.52 |  |

藍純bulkload:

lookup time:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | #1 | #2 | #3 | avg |
| 1 | 3308814532 | 3964543893 | 3759670526 |  |
| 0.9 | 4855611993 | 5063226589 | 4315934350 |  |
| 0.8 | 3997963529 | 3949290131 | 3320577092 |  |
| 0.7 | 3472200286 | 3006672043 | 3098347657 |  |
| 0.6 | 2555988935 | 2877108942 | 2538518936 |  |
| 0.5 | 2328607166 | 2198957459 | 2785915817 |  |

Load time:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | #1 | #2 | #3 | avg |
| 1 | 21530435085 | 21788773458 | 23426633869 |  |
| 0.9 | 25699626613 | 26543875828 | 25855091288 |  |
| 0.8 | 22297267245 | 21477103489 | 21475352758 |  |
| 0.7 | 21501003748 | 21619816913 | 22234655717 |  |
| 0.6 | 16818703707 | 16066988451 | 16345758244 |  |
| 0.5 | 15328662142 | 15613232609 | 15567228982 |  |

跟前面觀測到的一樣，緩降

???

比起look up，insert花很久

0.4 0.3 0.2…

1. 拆散頭尾 避免密集(之前是拆散中間)
2. 不要bulkload全部insert
3. 最理想的情況:50M

1201392941

1192664846

1202495353

一開始的range起始位置差(平移)的影響

逐一插入vs bulk load再插入的快慢

Insert超過特定量後 紅藍 < 黃灰 why