ECE 6130 Big Data and Cloud Computing Spring 2019

Homework #4 Report

Name: Tianyu Yang

GW ID: G38878678

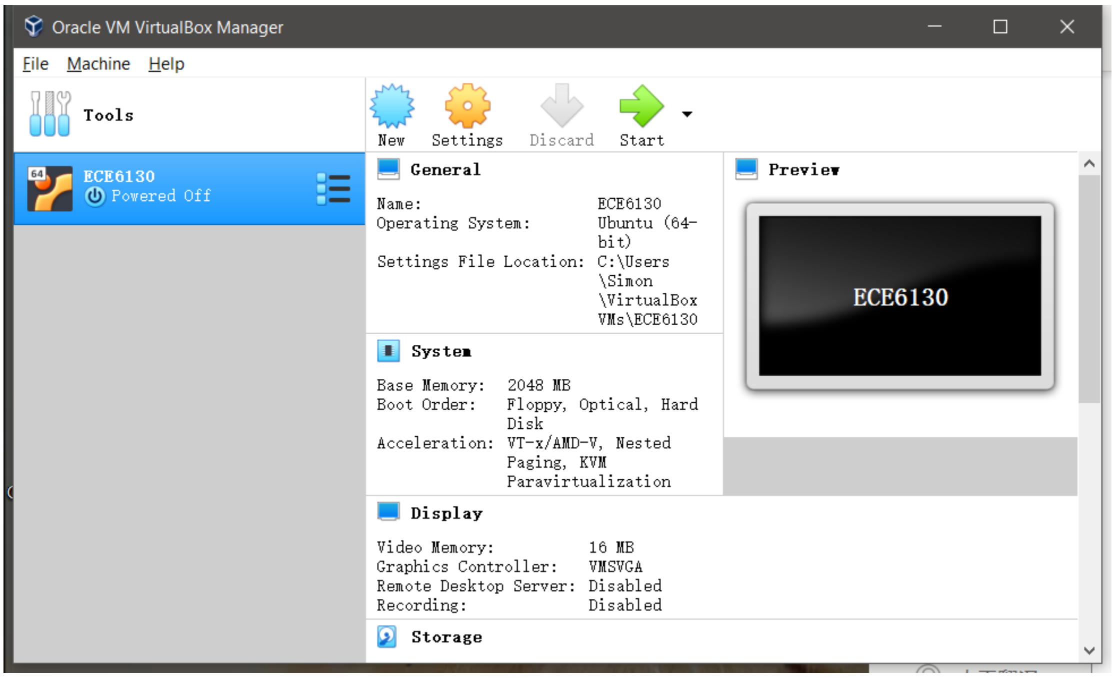
Date: 3/19/2019

Objective:

1. This homework is to use both virtual machine and physical machine to run BFS code
2. Record the configuration of virtual machine and physical machine.
3. Record the running time for both machine

Procedure 1: Computer configuration

1. In this project, I use Oracle VM VirtualBox Manager as the virtual machine. The configuration is shown as below:



```
General
Name:                    ECE6130
Operating System:        Ubuntu (64-bit)
Settings File Location:  C:\Users\Simon\VirtualBox VMs\ECE6130

System
Base Memory:   2048 MB
Boot Order:    Floppy, Optical, Hard Disk
Acceleration:  VT-x/AMD-V, Nested Paging, KVM Paravirtualization

Display
Video Memory:           16 MB
Graphics Controller:    VMSVGA
Remote Desktop Server:  Disabled
Recording:              Disabled

Storage
```

The Operation System is ubuntu (64-bit);

The Base Memory is 2048 MB;

The Video Memory is 16 MB with the Graphics Controller VMSVGA;

The Storage is Normal 20.00GB;

2. To settle the physical machine, I use USB to boot my computer in a Linux system. The reason is that I use Linux system as OS for my virtual machine. As for

comparation with the same OS, I also need to use Linux as OS for my physical machine. However, I do not own a physical computer with Linux system. Therefore, I use USB to store the Ubuntu system as boot my computer in the USB. Therefore, I can use any of the computer as physical machine with Ubuntu OS. The configuration of the computer is shown as below:

CPU
Intel(R) Core(TM) i7-7700HQ CPU @ 2.80GHz
Memory
SK Hynix 16.0 GB
GPU
NVIDIA GeForce GTX 1060 6 GB
Hard Disk
1 TB + 256 GB

The CPU is Inter(R) Core (TM) i7-7700HQ with frequency 2.80GHz;

The Memory is SK Hynix 16.0 GB;

The GPU is NVIDIA GeForce GTX 1060 6GB;

The Hard Disk is 1TB with HDD and 256GB SDD;

Procedure 2: BFS code review

The BFS code used to run on the machine is the simple code on Graph 500. http://graph500.org/?page_id=47. The related code for the algorithm of BFS is in bfs_reference.c

```
void make_graph_data_structure(const tuple_graph* const tg) {
    int i,j,k;
    convert_graph_to_oned_csr(tg, &g);
    column=g.column;
    rowstarts=g.rowstarts;

    visited_size = (g.nlocalverts + ulong_bits - 1) / ulong_bits;
    aml_register_handler(visithndl,1);
    q1 = xmalloc(g.nlocalverts*sizeof(int)); //100% of vertexes
```

```c
    q2 = xmalloc(g.nlocalverts*sizeof(int));
    for(i=0;i<g.nlocalverts;i++) q1[i]=0,q2[i]=0; //touch memory
    visited = xmalloc(visited_size*sizeof(unsigned long));
}

void run_bfs(int64_t root, int64_t* pred) {
    int64_t nvisited;
    long sum;
    unsigned int i,j,k,lvl=1;
    pred_glob=pred;
    aml_register_handler(visithndl,1);

    CLEAN_VISITED();

    qc=0; sum=1; q2c=0;

    nvisited=1;
    if(VERTEX_OWNER(root) == rank) {
        pred[VERTEX_LOCAL(root)]=root;
        SET_VISITED(root);
        q1[0]=VERTEX_LOCAL(root);
        qc=1;
    }

    // While there are vertices in current level
    while(sum) {
#ifdef DEBUGSTATS
        double t0=aml_time();
        nbytes_sent=0; nbytes_rcvd=0;
#endif
        //for all vertices in current level send visit AMs to all neighbours
        for(i=0;i<qc;i++)
            for(j=rowstarts[q1[i]];j<rowstarts[q1[i]+1];j++)
                send_visit(COLUMN(j),q1[i]);
        aml_barrier();

        qc=q2c;int *tmp=q1;q1=q2;q2=tmp;
        sum=qc;
        aml_long_allsum(&sum);

        nvisited+=sum;

        q2c=0;
#ifdef DEBUGSTATS
```

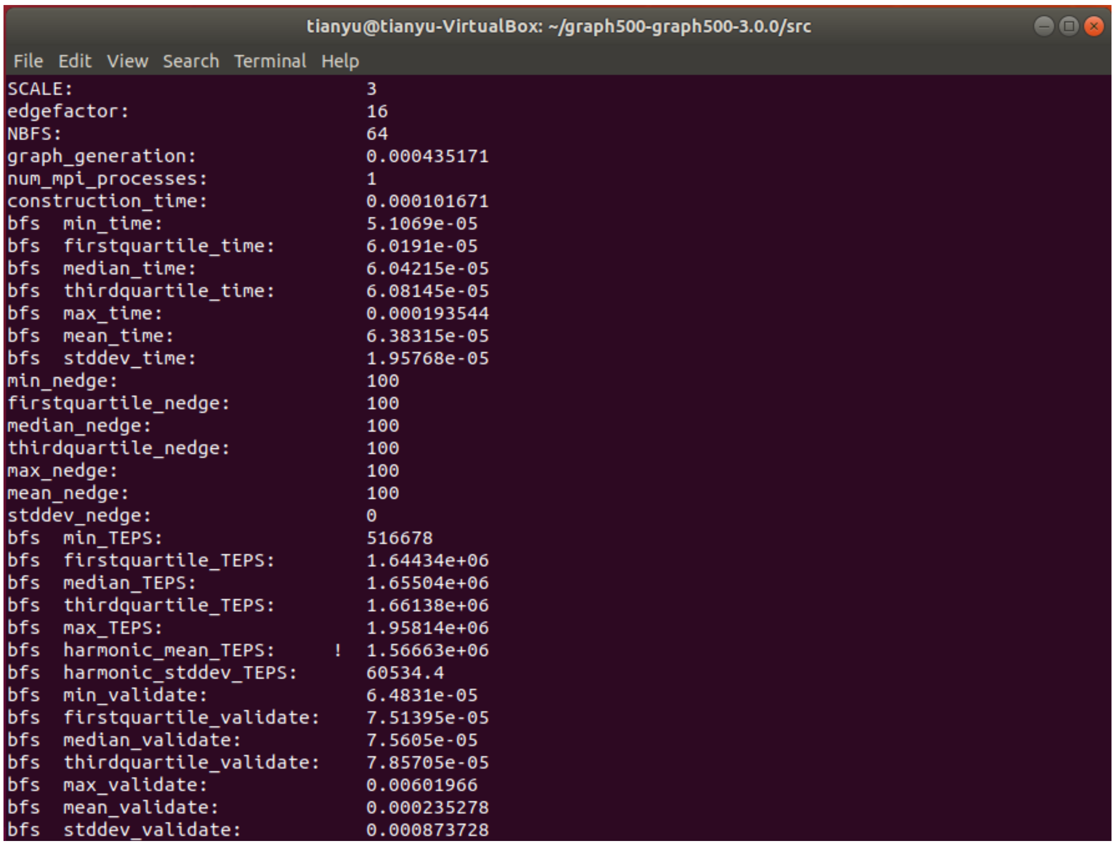
```
        aml_long_allsum(&nbytes_sent);
        t0-=aml_time();
        if(!my_pe()) printf (" --lvl%d : %lld(%lld,%3.2f) visited in %5.2fs, 
network
aggr %5.2fGb/s\n",lvl++,sum,nvisited,((double)nvisited/(double)g.notisolated)*1
00.0,-t0,-(double)nbytes_sent*8.0/(1.e9*t0));
#endif
    }
    aml_barrier();


}
```

Procedure 3: Running time

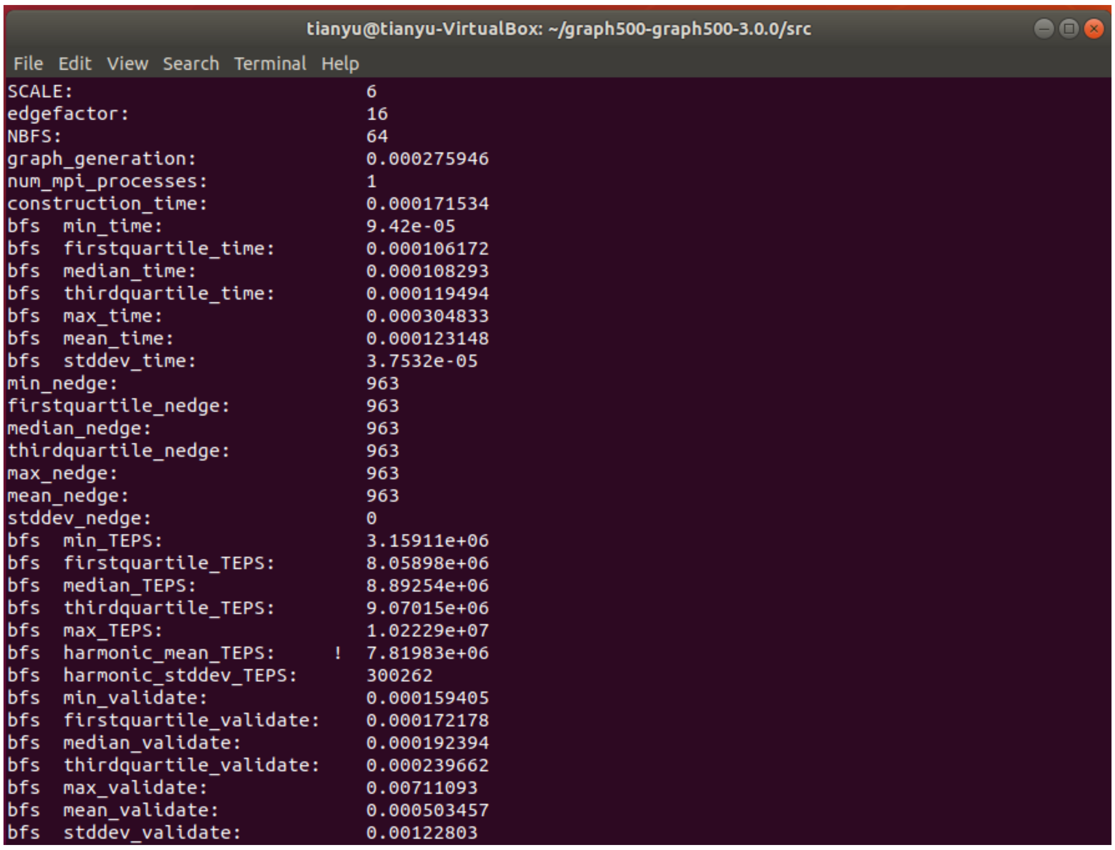
The running time for the virtual machine

When the scale = 3 and the edge value is 100:


tianyu@tianyu-VirtualBox: ~/graph500-graph500-3.0.0/src
File Edit View Search Terminal Help

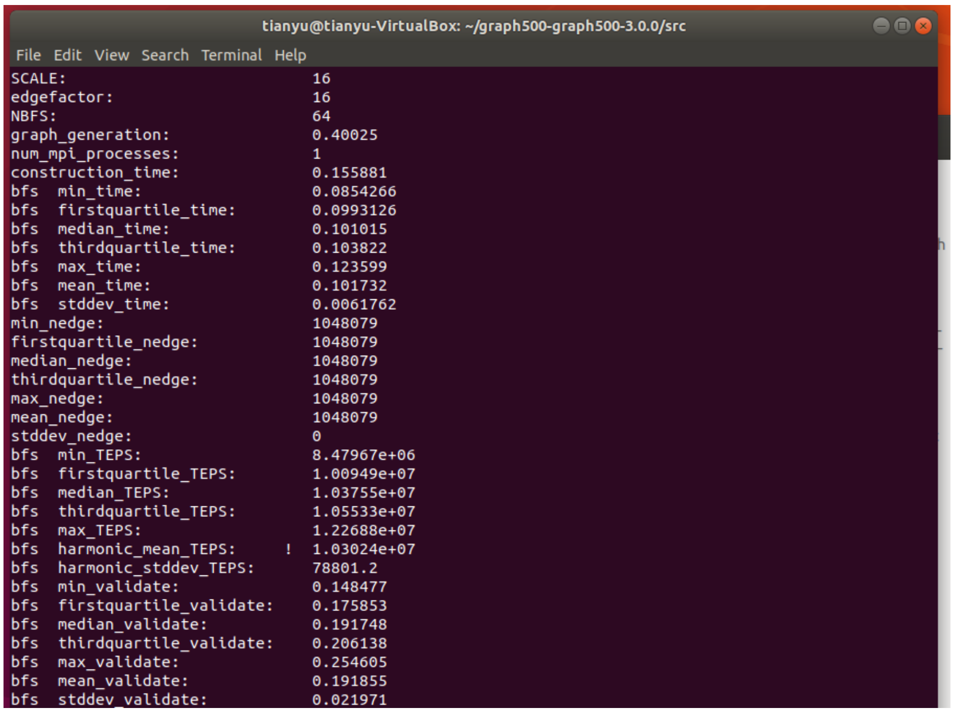
```
SCALE:                          3
edgefactor:                     16
NBFS:                           64
graph_generation:               0.000435171
num_mpi_processes:              1
construction_time:              0.000101671
bfs  min_time:                  5.1069e-05
bfs  firstquartile_time:        6.0191e-05
bfs  median_time:               6.04215e-05
bfs  thirdquartile_time:        6.08145e-05
bfs  max_time:                  0.000193544
bfs  mean_time:                 6.38315e-05
bfs  stddev_time:               1.95768e-05
min_nedge:                      100
firstquartile_nedge:            100
median_nedge:                   100
thirdquartile_nedge:            100
max_nedge:                      100
mean_nedge:                     100
stddev_nedge:                   0
bfs  min_TEPS:                  516678
bfs  firstquartile_TEPS:        1.64434e+06
bfs  median_TEPS:               1.65504e+06
bfs  thirdquartile_TEPS:        1.66138e+06
bfs  max_TEPS:                  1.95814e+06
bfs  harmonic_mean_TEPS:     !  1.56663e+06
bfs  harmonic_stddev_TEPS:      60534.4
bfs  min_validate:              6.4831e-05
bfs  firstquartile_validate:    7.51395e-05
bfs  median_validate:           7.5605e-05
bfs  thirdquartile_validate:    7.85705e-05
bfs  max_validate:              0.00601966
bfs  mean_validate:             0.000235278
bfs  stddev_validate:           0.000873728
```

When the scale = 6 and the edge value is 1K:


tianyu@tianyu-VirtualBox: ~/graph500-graph500-3.0.0/src
File Edit View Search Terminal Help

```
SCALE:                          6
edgefactor:                     16
NBFS:                           64
graph_generation:               0.000275946
num_mpi_processes:              1
construction_time:              0.000171534
bfs  min_time:                  9.42e-05
bfs  firstquartile_time:        0.000106172
bfs  median_time:               0.000108293
bfs  thirdquartile_time:        0.000119494
bfs  max_time:                  0.000304833
bfs  mean_time:                 0.000123148
bfs  stddev_time:               3.7532e-05
min_nedge:                      963
firstquartile_nedge:            963
median_nedge:                   963
thirdquartile_nedge:            963
max_nedge:                      963
mean_nedge:                     963
stddev_nedge:                   0
bfs  min_TEPS:                  3.15911e+06
bfs  firstquartile_TEPS:        8.05898e+06
bfs  median_TEPS:               8.89254e+06
bfs  thirdquartile_TEPS:        9.07015e+06
bfs  max_TEPS:                  1.02229e+07
bfs  harmonic_mean_TEPS:     !  7.81983e+06
bfs  harmonic_stddev_TEPS:      300262
bfs  min_validate:              0.000159405
bfs  firstquartile_validate:    0.000172178
bfs  median_validate:           0.000192394
bfs  thirdquartile_validate:    0.000239662
bfs  max_validate:              0.00711093
bfs  mean_validate:             0.000503457
bfs  stddev_validate:           0.00122803
```

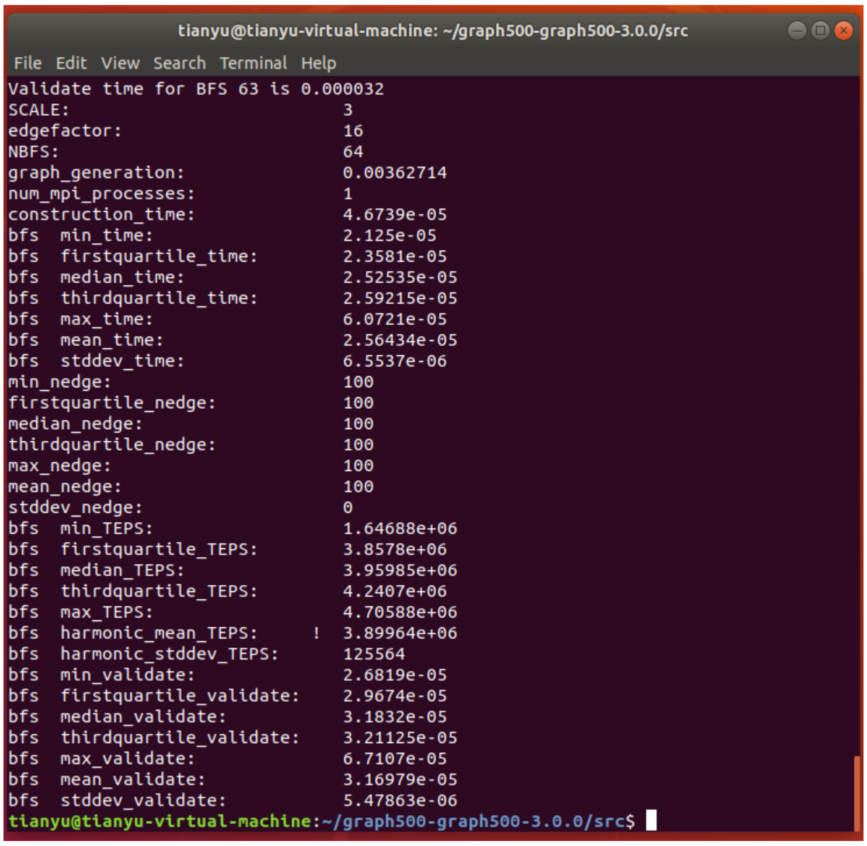
When the scale = 16 and the edge value is 1M:


tianyu@tianyu-VirtualBox: ~/graph500-graph500-3.0.0/src
File Edit View Search Terminal Help

```
SCALE:                          16
edgefactor:                     16
NBFS:                           64
graph_generation:               0.40025
num_mpi_processes:              1
construction_time:              0.155881
bfs  min_time:                  0.0854266
bfs  firstquartile_time:        0.0993126
bfs  median_time:               0.101015
bfs  thirdquartile_time:        0.103822
bfs  max_time:                  0.123599
bfs  mean_time:                 0.101732
bfs  stddev_time:               0.0061762
min_nedge:                      1048079
firstquartile_nedge:            1048079
median_nedge:                   1048079
thirdquartile_nedge:            1048079
max_nedge:                      1048079
mean_nedge:                     1048079
stddev_nedge:                   0
bfs  min_TEPS:                  8.47967e+06
bfs  firstquartile_TEPS:        1.00949e+07
bfs  median_TEPS:               1.03755e+07
bfs  thirdquartile_TEPS:        1.05533e+07
bfs  max_TEPS:                  1.22688e+07
bfs  harmonic_mean_TEPS:     !  1.03024e+07
bfs  harmonic_stddev_TEPS:      78801.2
bfs  min_validate:              0.148477
bfs  firstquartile_validate:    0.175853
bfs  median_validate:           0.191748
bfs  thirdquartile_validate:    0.206138
bfs  max_validate:              0.254605
bfs  mean_validate:             0.191855
bfs  stddev_validate:           0.021971
```

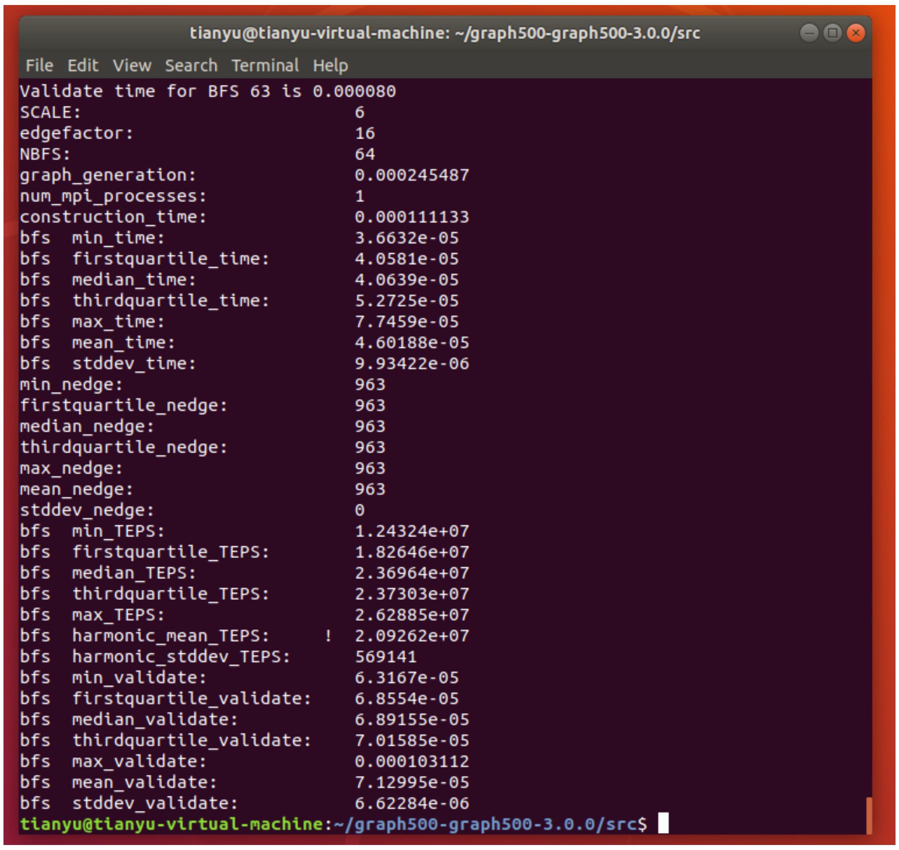
The running time for the physical machine

When the scale = 3 and the edge value is 100:


tianyu@tianyu-virtual-machine: ~/graph500-graph500-3.0.0/src
File Edit View Search Terminal Help

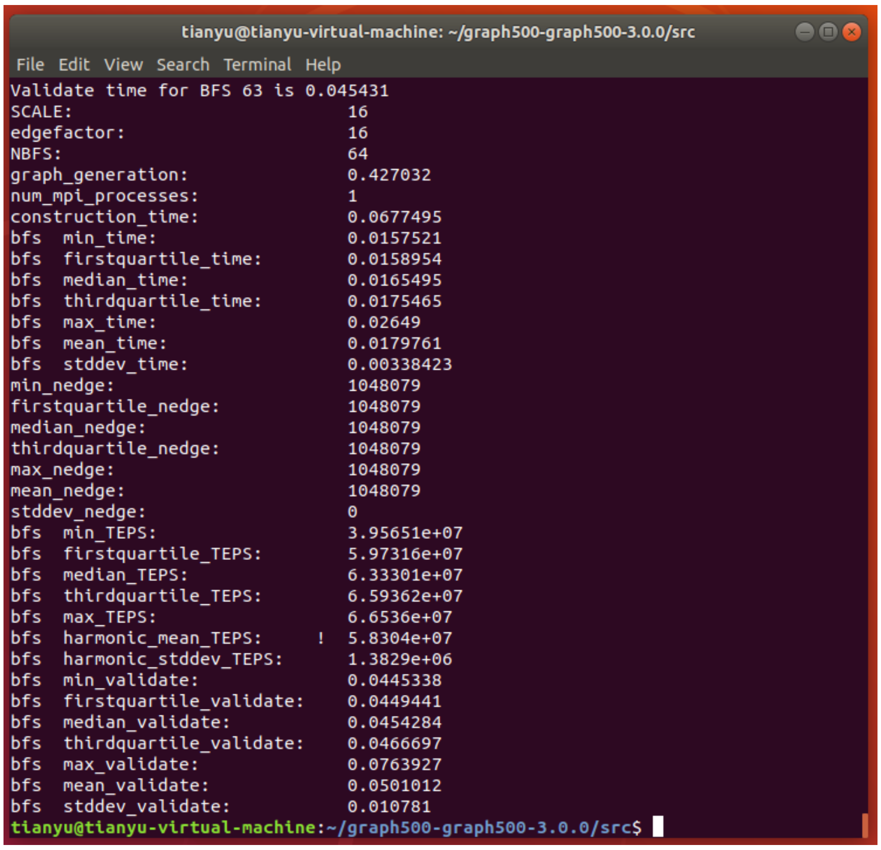
```
Validate time for BFS 63 is 0.000032
SCALE:                          3
edgefactor:                     16
NBFS:                           64
graph_generation:               0.00362714
num_mpi_processes:              1
construction_time:              4.6739e-05
bfs  min_time:                  2.125e-05
bfs  firstquartile_time:        2.3581e-05
bfs  median_time:               2.52535e-05
bfs  thirdquartile_time:        2.59215e-05
bfs  max_time:                  6.0721e-05
bfs  mean_time:                 2.56434e-05
bfs  stddev_time:               6.5537e-06
min_nedge:                      100
firstquartile_nedge:            100
median_nedge:                   100
thirdquartile_nedge:            100
max_nedge:                      100
mean_nedge:                     100
stddev_nedge:                   0
bfs  min_TEPS:                  1.64688e+06
bfs  firstquartile_TEPS:        3.8578e+06
bfs  median_TEPS:               3.95985e+06
bfs  thirdquartile_TEPS:        4.2407e+06
bfs  max_TEPS:                  4.70588e+06
bfs  harmonic_mean_TEPS:     !  3.89964e+06
bfs  harmonic_stddev_TEPS:      125564
bfs  min_validate:              2.6819e-05
bfs  firstquartile_validate:    2.9674e-05
bfs  median_validate:           3.1832e-05
bfs  thirdquartile_validate:    3.21125e-05
bfs  max_validate:              6.7107e-05
bfs  mean_validate:             3.16979e-05
bfs  stddev_validate:           5.47863e-06
tianyu@tianyu-virtual-machine:~/graph500-graph500-3.0.0/src$
```

When the scale = 6 and the edge value is 1K:


tianyu@tianyu-virtual-machine: ~/graph500-graph500-3.0.0/src
File Edit View Search Terminal Help

```
Validate time for BFS 63 is 0.000080
SCALE:                          6
edgefactor:                     16
NBFS:                           64
graph_generation:               0.000245487
num_mpi_processes:              1
construction_time:              0.000111133
bfs  min_time:                  3.6632e-05
bfs  firstquartile_time:        4.0581e-05
bfs  median_time:               4.0639e-05
bfs  thirdquartile_time:        5.2725e-05
bfs  max_time:                  7.7459e-05
bfs  mean_time:                 4.60188e-05
bfs  stddev_time:               9.93422e-06
min_nedge:                      963
firstquartile_nedge:            963
median_nedge:                   963
thirdquartile_nedge:            963
max_nedge:                      963
mean_nedge:                     963
stddev_nedge:                   0
bfs  min_TEPS:                  1.24324e+07
bfs  firstquartile_TEPS:        1.82646e+07
bfs  median_TEPS:               2.36964e+07
bfs  thirdquartile_TEPS:        2.37303e+07
bfs  max_TEPS:                  2.62885e+07
bfs  harmonic_mean_TEPS:     !  2.09262e+07
bfs  harmonic_stddev_TEPS:      569141
bfs  min_validate:              6.3167e-05
bfs  firstquartile_validate:    6.8554e-05
bfs  median_validate:           6.89155e-05
bfs  thirdquartile_validate:    7.01585e-05
bfs  max_validate:              0.000103112
bfs  mean_validate:             7.12995e-05
bfs  stddev_validate:           6.62284e-06
tianyu@tianyu-virtual-machine:~/graph500-graph500-3.0.0/src$
```

When the scale = 16 and the edge value is 1M:


tianyu@tianyu-virtual-machine: ~/graph500-graph500-3.0.0/src
File Edit View Search Terminal Help

```
Validate time for BFS 63 is 0.045431
SCALE:                          16
edgefactor:                     16
NBFS:                           64
graph_generation:               0.427032
num_mpi_processes:              1
construction_time:              0.0677495
bfs  min_time:                  0.0157521
bfs  firstquartile_time:        0.0158954
bfs  median_time:               0.0165495
bfs  thirdquartile_time:        0.0175465
bfs  max_time:                  0.02649
bfs  mean_time:                 0.0179761
bfs  stddev_time:               0.00338423
min_nedge:                      1048079
firstquartile_nedge:            1048079
median_nedge:                   1048079
thirdquartile_nedge:            1048079
max_nedge:                      1048079
mean_nedge:                     1048079
stddev_nedge:                   0
bfs  min_TEPS:                  3.95651e+07
bfs  firstquartile_TEPS:        5.97316e+07
bfs  median_TEPS:               6.33301e+07
bfs  thirdquartile_TEPS:        6.59362e+07
bfs  max_TEPS:                  6.6536e+07
bfs  harmonic_mean_TEPS:     !  5.8304e+07
bfs  harmonic_stddev_TEPS:      1.3829e+06
bfs  min_validate:              0.0445338
bfs  firstquartile_validate:    0.0449441
bfs  median_validate:           0.0454284
bfs  thirdquartile_validate:    0.0466697
bfs  max_validate:              0.0763927
bfs  mean_validate:             0.0501012
bfs  stddev_validate:           0.010781
tianyu@tianyu-virtual-machine:~/graph500-graph500-3.0.0/src$ 
```

Procedure 4: Result analysis

To compare with the result of the running time in virtual machine and physical, I make a table below for the comparation.

Edge=100	graph_generation	construction_time	bfs mean_time
Virtual Machine	0.000435171	0.000101671	0.000063815
Physical Machine	0.00362714	4.6739E-05	2.56434E-5
Comparation(less)	V	P	P
Edge=1K	graph_generation	construction_time	bfs mean_time
Virtual Machine	0.000275946	0.000171534	0.000123148
Physical Machine	0.000245487	0.000111133	4.60188E-5
Comparation	P	P	P
Edge=1M	graph_generation	construction_time	bfs mean_time
Virtual Machine	0.40025	0.155881	0.101732
Physical Machine	0.427032	0.0677495	0.0179761
Comparation	V	P	P

Therefore, with the comparation, we can conclude that the physical machine has the faster running time for BFS algorithm than the virtual machine. Though, there is natural error due to the limited number of edges (too small). With the comparation of computer configuration, we can get that when the machine has better performance CPU and Memory, the BFS algorithm can run faster on the machine.