Duke University Human Capital / Synthetic Data Project
Server Performance Comparison
1/24/2018, T. Balmat

## BACKGROUND

In December of 2016, while testing new servers pap36 and pap46 before being placed into service, it was observed that certain R scripts that had been in use required longer to complete on pap46 than on pap1 (pap46 is intended to replace pap1). Table 1 lists completion times of select algorithms executed 12/4/2017 on pap1 and pap46 using randomly generated data sets of varying size that resemble authentic data used in research. Figure 1 plots completion times by number of observations in a data set. Since December, various execution time comparisons have been made between pap1, pap2, pap36, and pap46 with concentration on pap1 and pap46, since long periods of dedicated access are possible with these. In general, identical algorithms with similar data sets have tended to complete more quickly on pap1 than on pap46. Pap36 is intended to replace pap2 and is shared concurrently by researches and an active SQL Server instance. It is important to note that aspects of current research to be published, particularly with the Synthetic Data Project, compare and report computational efficiency of various algorithms, making consistency of results important.

Table 1: Comparison of pap1 and pap46 execution times using research models. December 2017. n indicates observations, Method indicates algorithm, and the "t" columns contain time to completion (in seconds) by server. Servers were dedicated during execution.

| | n | Method | t.pap1 | t.pap46 |
|---|---|---|---|---|
| 1 | 10000 | feXTX | 1.50 | 1.79 |
| 2 | 25000 | feXTX | 1.50 | 2.06 |
| 3 | 50000 | feXTX | 1.75 | 2.39 |
| 4 | 100000 | feXTX | 2.34 | 2.90 |
| 5 | 150000 | feXTX | 2.68 | 3.50 |
| 6 | 250000 | feXTX | 3.77 | 5.25 |
| 7 | 500000 | feXTX | 5.63 | 7.78 |
| 8 | 1000000 | feXTX | 13.25 | 17.64 |
| 9 | 1500000 | feXTX | 22.94 | 26.67 |
| 10 | 2000000 | feXTX | 47.56 | 54.87 |
| 11 | 5000000 | feXTX | 241.93 | 332.27 |
| 12 | 10000000 | feXTX | 1396.29 | 1725.81 |
| 13 | 25000000 | feXTX | 5928.08 | 8024.80 |
| 14 | 10000 | Matrix1 | 0.05 | 0.11 |
| 15 | 25000 | Matrix1 | 0.12 | 0.20 |
| 16 | 50000 | Matrix1 | 0.27 | 0.35 |
| 17 | 100000 | Matrix1 | 0.56 | 0.92 |
| 18 | 150000 | Matrix1 | 1.04 | 1.58 |
| 19 | 250000 | Matrix1 | 1.92 | 3.13 |
| 20 | 500000 | Matrix1 | 3.83 | 6.31 |
| 21 | 1000000 | Matrix1 | 8.62 | 13.38 |
| 22 | 1500000 | Matrix1 | 14.65 | 18.54 |
| 23 | 2000000 | Matrix1 | 23.99 | 33.57 |
| 24 | 5000000 | Matrix1 | 362.98 | 433.91 |
| 25 | 10000000 | Matrix1 | 5251.48 | 5981.92 |
| 26 | 25000000 | Matrix1 | 9411.06 | 10303.39 |
| 27 | 10000 | Matrix2 | 0.08 | 0.11 |
| 28 | 25000 | Matrix2 | 0.15 | 0.20 |
| 29 | 50000 | Matrix2 | 0.30 | 0.32 |
| 30 | 100000 | Matrix2 | 0.67 | 0.87 |
| 31 | 150000 | Matrix2 | 0.86 | 1.16 |
| 32 | 250000 | Matrix2 | 1.42 | 2.16 |
| 33 | 500000 | Matrix2 | 2.69 | 4.10 |
| 34 | 1000000 | Matrix2 | 5.68 | 8.80 |
| 35 | 1500000 | Matrix2 | 9.30 | 12.31 |
| 36 | 2000000 | Matrix2 | 14.10 | 19.86 |
| 37 | 5000000 | Matrix2 | 63.07 | 73.23 |
| 38 | 10000000 | Matrix2 | 375.82 | 480.18 |
| 39 | 25000000 | Matrix2 | 2092.43 | 2932.28 |

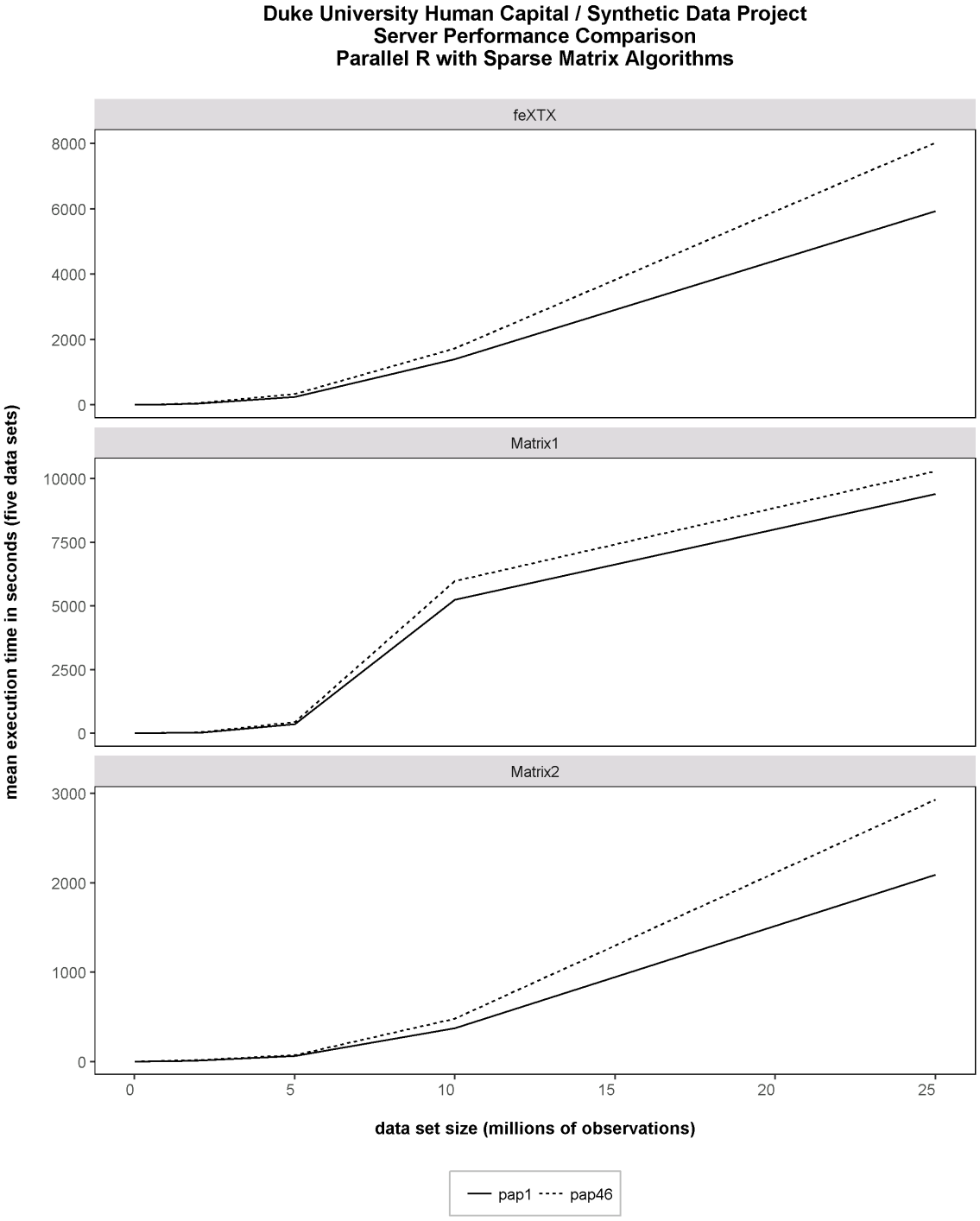



Figure 1: Comparison of pap1 and pap46 execution times using research models. December 2017. Servers were dedicated during execution.

CURRENT TESTING

Beginning around 1/15/2018 additional testing was begun using algorithms with similar computations, but different implementation from those used in December. While conducting tests, total CPU utilization generally appeared lower on pap1 and pap2 than on pap36 and pap46. Variability in CPU utilization also appeared greater on pap1 and pap2, with a 20% inter-day deviation observed on pap1 while executing identical algorithms with similar data. Observations were made after confirming that no other user processes were active. Table 2 compares execution time by server and figure 2 plots execution time by observation count. In these results, pap46 appears to out perform pap1. The data used in January are similar to that used in December and, although solution methods differ, the computational load is very similar, making the change in pap46 performance relative to pap1 unexpected.

Table 2: Comparison of pap1, pap2, pap36, and pap46 execution times using research models. January 2018. n indicates observations, Problem indicates the problem and algorithm used, and the "t" columns contain time to completion (in seconds) by server. Servers were dedicated during execution.

| | n | Problem | t.pap1 | t.pap46 | t.pap2 | t.pap36 |
|---|---|---|---|---|---|---|
| 1 | 1000000 | tData | 3.44 | 2.65 | 2.63 | 3.01 |
| 2 | 5000000 | tData | 14.80 | 13.11 | 12.65 | 14.56 |
| 3 | 10000000 | tData | 32.99 | 26.31 | 25.90 | 27.49 |
| 4 | 25000000 | tData | 79.14 | 71.33 | 72.06 | 79.28 |
| 5 | 1000000 | tXTX | 0.36 | 0.35 | 0.35 | 0.39 |
| 6 | 5000000 | tXTX | 1.85 | 1.76 | 1.71 | 1.91 |
| 7 | 10000000 | tXTX | 4.97 | 4.41 | 4.41 | 4.99 |
| 8 | 25000000 | tXTX | 31.91 | 26.04 | 27.61 | 32.45 |
| 9 | 1000000 | tChol | 0.48 | 0.27 | 0.28 | 0.27 |
| 10 | 5000000 | tChol | 0.81 | 0.59 | 0.65 | 0.62 |
| 11 | 10000000 | tChol | 2.08 | 1.35 | 1.49 | 1.45 |
| 12 | 25000000 | tChol | 108.22 | 71.58 | 86.96 | 77.46 |
| 13 | 1000000 | tBeta | 0.01 | 0.01 | 0.01 | 0.01 |
| 14 | 5000000 | tBeta | 0.12 | 0.09 | 0.04 | 0.13 |
| 15 | 10000000 | tBeta | 0.58 | 0.49 | 0.28 | 0.59 |
| 16 | 25000000 | tBeta | 15.68 | 13.31 | 10.83 | 16.06 |
| 17 | 1000000 | tvBeta | 0.51 | 0.37 | 0.36 | 0.34 |
| 18 | 5000000 | tvBeta | 1.34 | 0.98 | 1.01 | 0.98 |
| 19 | 10000000 | tvBeta | 4.21 | 2.46 | 2.66 | 2.69 |
| 20 | 25000000 | tvBeta | 207.83 | 133.06 | 155.83 | 151.22 |

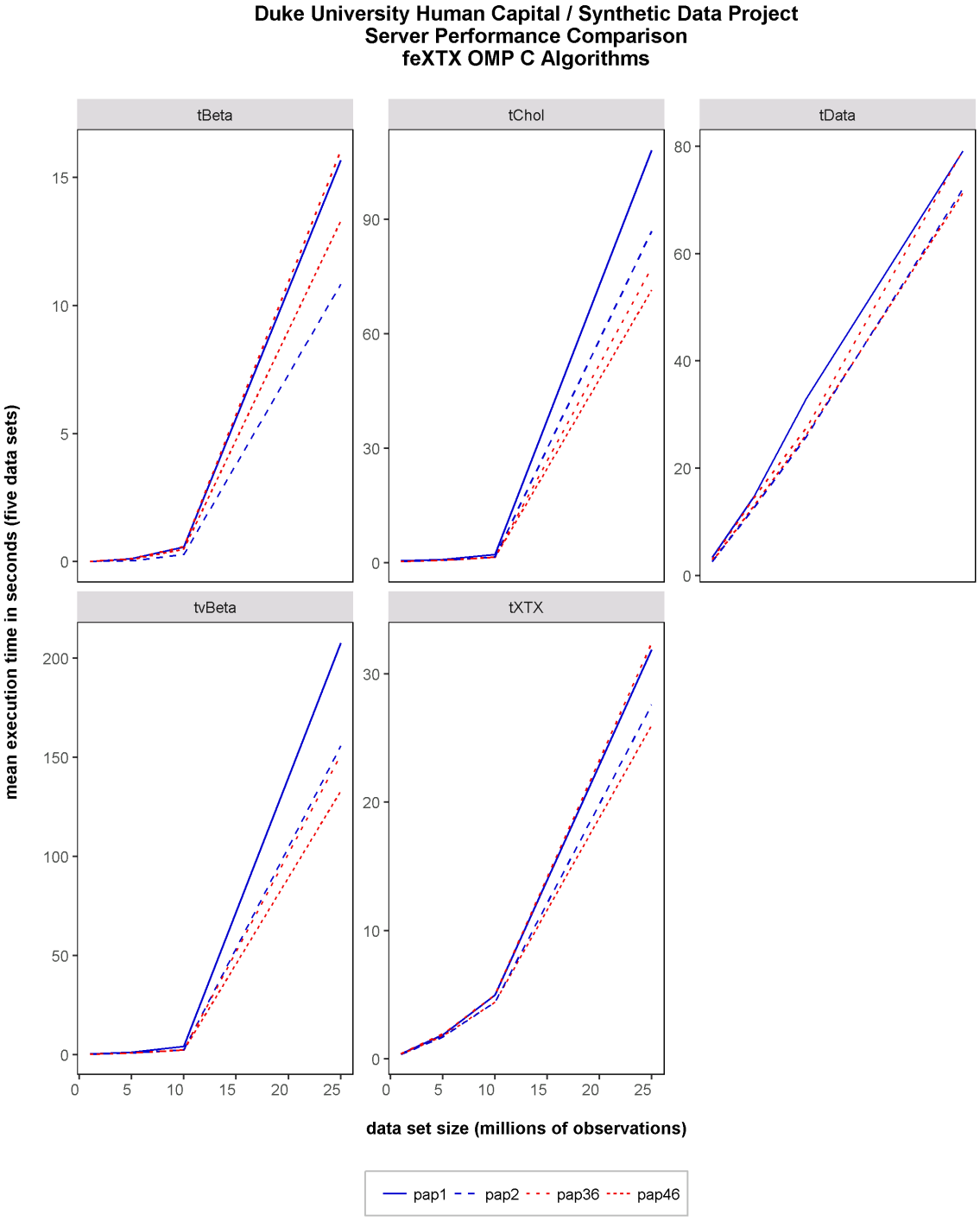



Figure 2: Comparison of pap1, pap2, pap36, and pap46 execution times using research models. January 2018. Servers were dedicated during execution.

The scripts used for the previous test involve custom, parallelize C functions using the open MP library. It should be noted that C functions were compiled on a separate Windows host (none of the servers evaluated here) and shared copies of the executable are used by all servers. To test performance of more standard R functions, models were solved using compatible algorithms from the Matrix and MatrixModels packages. Table 3 and figure 3 compare execution time between servers. As with other January results, pap46 appears to outperform pap1.

Table 3: Comparison of pap1 and pap46 execution times using Matrix package. January 2018. n indicates observations, Problem indicates the problem and algorithm used, and the "t" columns contain time to completion (in seconds) by server. Servers were dedicated during execution.

| | n | Problem | t.pap1 | t.pap46 |
|---|---|---|---|---|
| 1 | 1000000 | tData | 3.08 | 2.95 |
| 2 | 5000000 | tData | 14.79 | 13.01 |
| 3 | 10000000 | tData | 33.02 | 25.89 |
| 4 | 25000000 | tData | 83.56 | 71.40 |
| 5 | 1000000 | tX | 4.06 | 2.77 |
| 6 | 5000000 | tX | 25.19 | 17.22 |
| 7 | 10000000 | tX | 65.29 | 43.03 |
| 8 | 25000000 | tX | 218.94 | 154.93 |
| 9 | 1000000 | tXTX | 2.16 | 1.53 |
| 10 | 5000000 | tXTX | 19.00 | 10.73 |
| 11 | 10000000 | tXTX | 44.93 | 28.23 |
| 12 | 25000000 | tXTX | 183.27 | 114.02 |
| 13 | 1000000 | tChol | 0.24 | 0.18 |
| 14 | 5000000 | tChol | 1.66 | 1.11 |
| 15 | 10000000 | tChol | 14.64 | 10.10 |
| 16 | 25000000 | tChol | 1270.49 | 1089.93 |

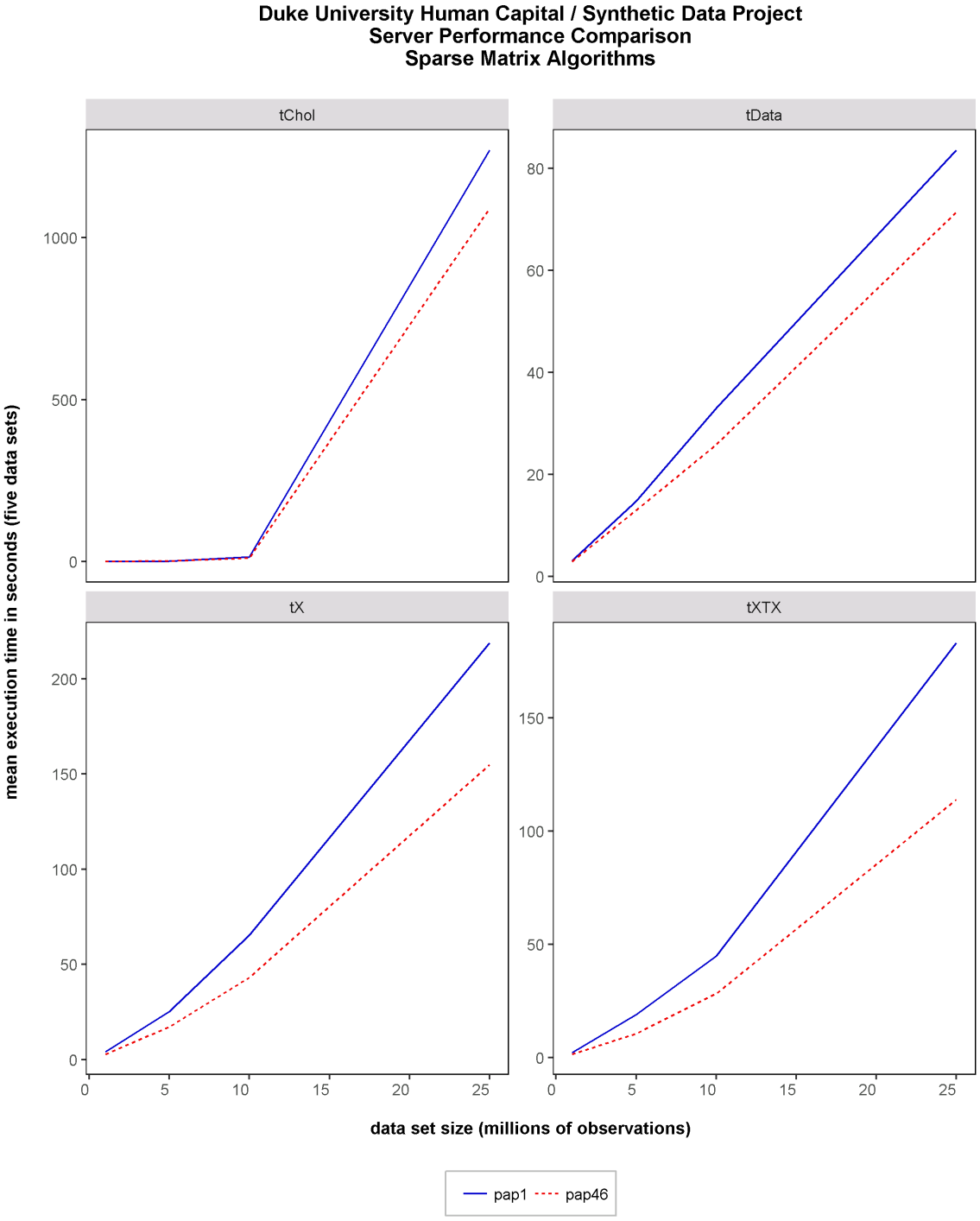
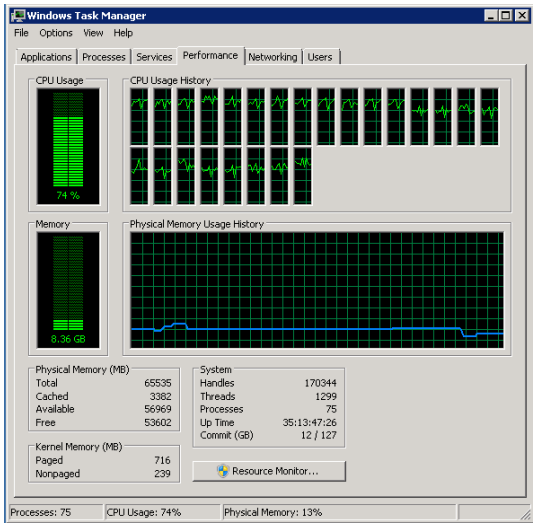


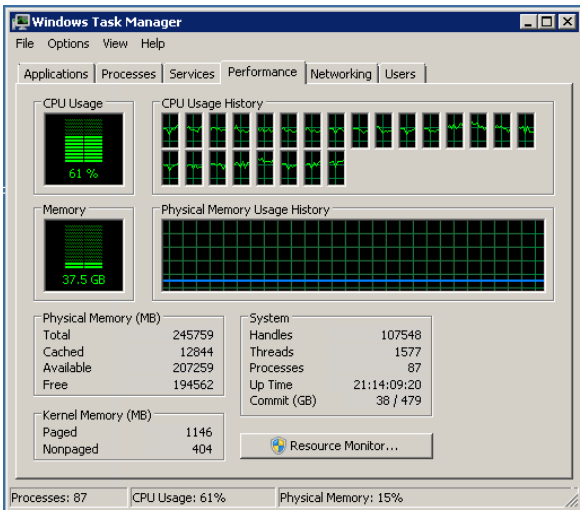

Figure 3: Comparison of pap1 and pap46 execution times using Matrix package. January 2018. Servers were dedicated during execution.

CPU UTILIZATION

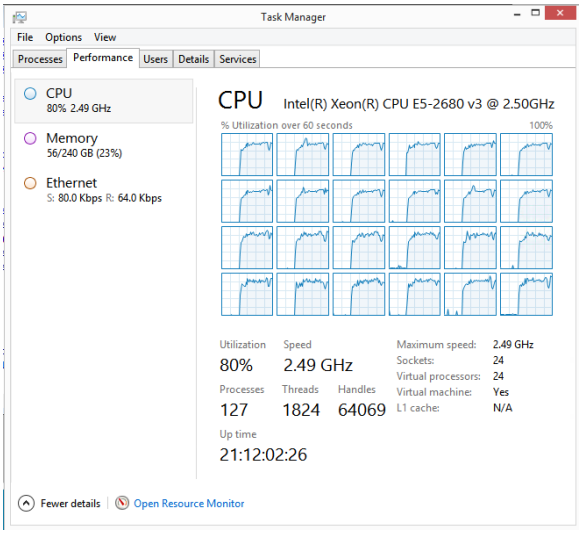
Two algorithms with high computational demand are the Cholesky decomposition and computation of a matrix inverse using a computed Cholesky decomposition. Custom parallelized C functions for these operations, using the open MP library, were executed on each server using a large random data set generated on each server using identical seed values. Figure 4 contains screen shots of the task manager performance tab for each server while computing the Cholesky decomposition. Although snapshots, images were taken at moments of customary utilization. Scripts were executed on 1/25/2018. Table 4 lists CPU utilization and time to compute the Cholesky decomposition. Figure 5 and table 5 show CPU utilization and time to compute the inverse of a large matrix using its Cholesky decomposition. From the tables and graphs, it appears that computation time is dependent on CPU utilization and, at the time of this test, pap46 experienced better CPU utilization than did the other servers.



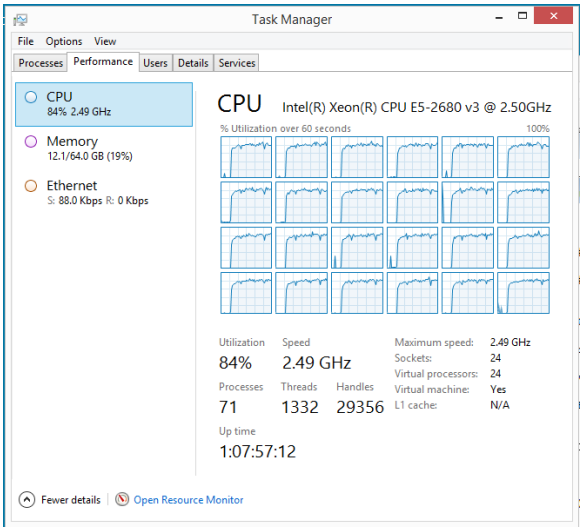

(a) pap1




(b) pap2




(c) pap36

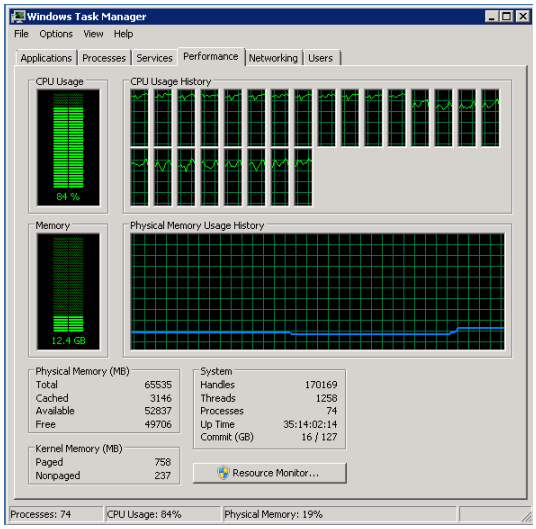



(d) pap46

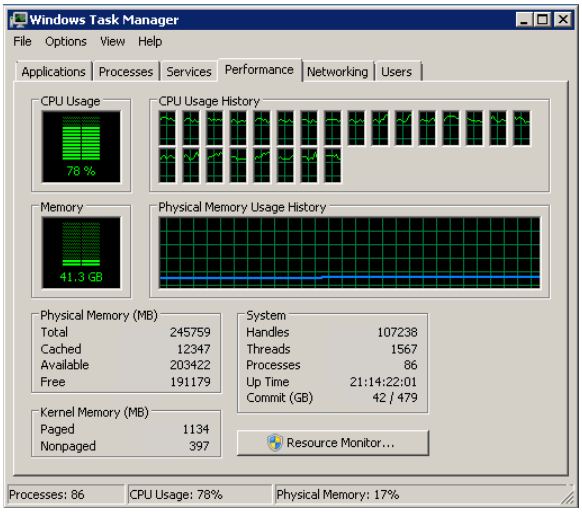
Figure 4: CPU utilization while computing Cholesky decomposition of a large matrix. 1/25/2018. Servers were dedicated during execution.

Table 4: CPU utilization and time to compute Cholesky decomposition of a large matrix. 1/25/2018. Servers were dedicated during execution.

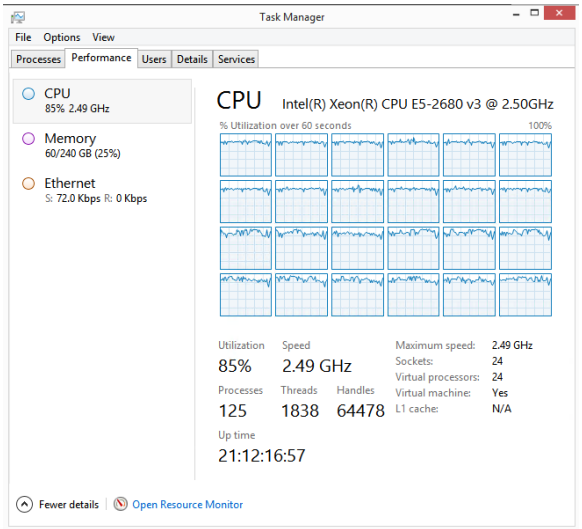
| Server | CPU Utilization | Computation Time (sec) |
|---|---|---|
| pap1 | 74% | 74.94 |
| pap2 | 61% | 105.09 |
| pap36 | 80% | 80.06 |
| pap46 | 84% | 70.20 |



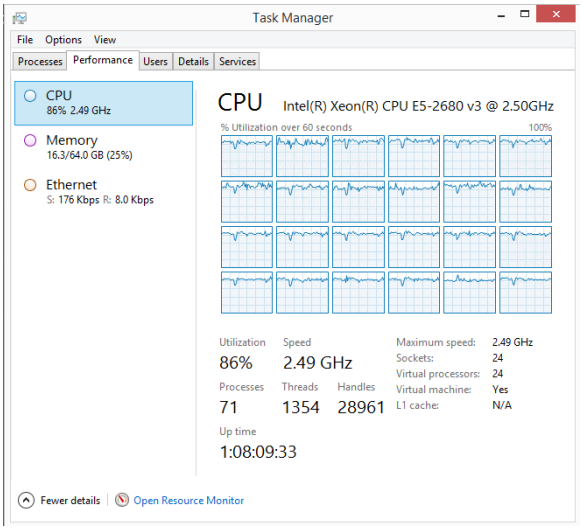

(a) pap1




(b) pap2




(c) pap36




(d) pap46

Figure 5: CPU utilization while computing matrix inverse using Cholesky decomposition. 1/25/2018. Servers were dedicated during execution.

Table 5: CPU utilization and time to compute inverse of matrix using Cholesky decomposition. 1/25/2018. Servers were dedicated during execution.

| Server | CPU Utilization | Computation Time (sec) |
|---|---|---|
| pap1 | 84% | 150.20 |
| pap2 | 78% | 152.32 |
| pap36 | 85% | 147.39 |
| pap46 | 86% | 128.77 |

QUESTIONS

- Has something changed in the configurations of pap1 and pap46 since December 2017?
- Why does pap46 appear to outperform the other three servers? Whatever the cause, and if it is robust to operating conditions and reboot, we want to capitalize on it!

FURTHER STEPS IN EVALUATING PERFORMANCE

- Repeat evaluation of December 2017 tests (parallel R and Sparse Matrix methods). Objective is to determine whether pap1 continues to outperform pap46. January results would suggest not.
- Repeat CPU utilization comparison, monitoring inter-day variation, especially after server reboot.

SERVER CONFIGURATION AS REPORTED BY R

Server: pap1 (configured in 2012)

```
R version 3.2.1 (2015-06-18)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows Server 2008 R2 x64 (build 7601) Service Pack 1

locale:
[1] LC_COLLATE=English_United States.1252
[2] LC_CTYPE=English_United States.1252
[3] LC_MONETARY=English_United States.1252
[4] LC_NUMERIC=C
[5] LC_TIME=English_United States.1252

attached base packages:
[1] parallel  stats     graphics  grDevices utils     datasets  methods
[8] base

other attached packages:
[1] MatrixModels_0.4-1 Matrix_1.2-1

loaded via a namespace (and not attached):
[1] grid_3.2.1      lattice_0.20-31
```

Server: pap2 (configured in 2012)

```
R version 3.2.1 (2015-06-18)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows Server 2008 R2 x64 (build 7601) Service Pack 1

locale:
[1] LC_COLLATE=English_United States.1252
[2] LC_CTYPE=English_United States.1252
[3] LC_MONETARY=English_United States.1252
[4] LC_NUMERIC=C
[5] LC_TIME=English_United States.1252

attached base packages:
[1] parallel  stats     graphics  grDevices utils     datasets  methods
[8] base

other attached packages:
[1] MatrixModels_0.4-1 Matrix_1.2-1

loaded via a namespace (and not attached):
[1] Rcpp_0.12.0     grid_3.2.1      lattice_0.20-31
```

Server: pap46 (replacement for pap1, configured in 2017)

```
R version 3.4.1 (2017-06-30)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows Server 2012 R2 x64 (build 9600)

Matrix products: default

locale:
[1] LC_COLLATE=English_United States.1252
[2] LC_CTYPE=English_United States.1252
[3] LC_MONETARY=English_United States.1252
[4] LC_NUMERIC=C
[5] LC_TIME=English_United States.1252

attached base packages:
[1] parallel  stats     graphics  grDevices utils     datasets  methods
[8] base

other attached packages:
[1] MatrixModels_0.4-1 Matrix_1.2-10

loaded via a namespace (and not attached):
[1] compiler_3.4.1  Rcpp_0.12.11    grid_3.4.1      lattice_0.20-35
```

Server: pap36 (replacement for pap2, configured in 2017)

```
R version 3.4.1 (2017-06-30)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows Server 2012 R2 x64 (build 9600)

Matrix products: default

locale:
[1] LC_COLLATE=English_United States.1252
[2] LC_CTYPE=English_United States.1252
[3] LC_MONETARY=English_United States.1252
```

```
[4] LC_NUMERIC=C
[5] LC_TIME=English_United States.1252

attached base packages:
[1] parallel  stats     graphics  grDevices utils     datasets  methods
[8] base

other attached packages:
[1] MatrixModels_0.4-1 Matrix_1.2-10

loaded via a namespace (and not attached):
[1] compiler_3.4.1  Rcpp_0.12.11    grid_3.4.1      lattice_0.20-35}
```

## R SCRIPTS USED FOR TESTING

```
\\ssri-nas-fe01.oit.duke.edu\ssri\OPM\Users\Current\tjb48\Analysis\FixedEffectsModel\
LargeModelSolutionComparison\pap1-pap46-Comparison\FixedEffectsRegression.r

\\ssri-nas-fe01.oit.duke.edu\ssri\OPM\Users\Current\tjb48\Analysis\FixedEffectsModel\
LargeModelSolutionComparison\SimulatedData.r (functions feSolution_MatrixSparseFitSolve(),
feSolution_MatrixSparseChol(), and feSolution_feXTX())
```