# CS475 – Project #3

# K-means: A Real Application Parallel Challenge

**Richard Phan**

phanri@oregonstate.edu

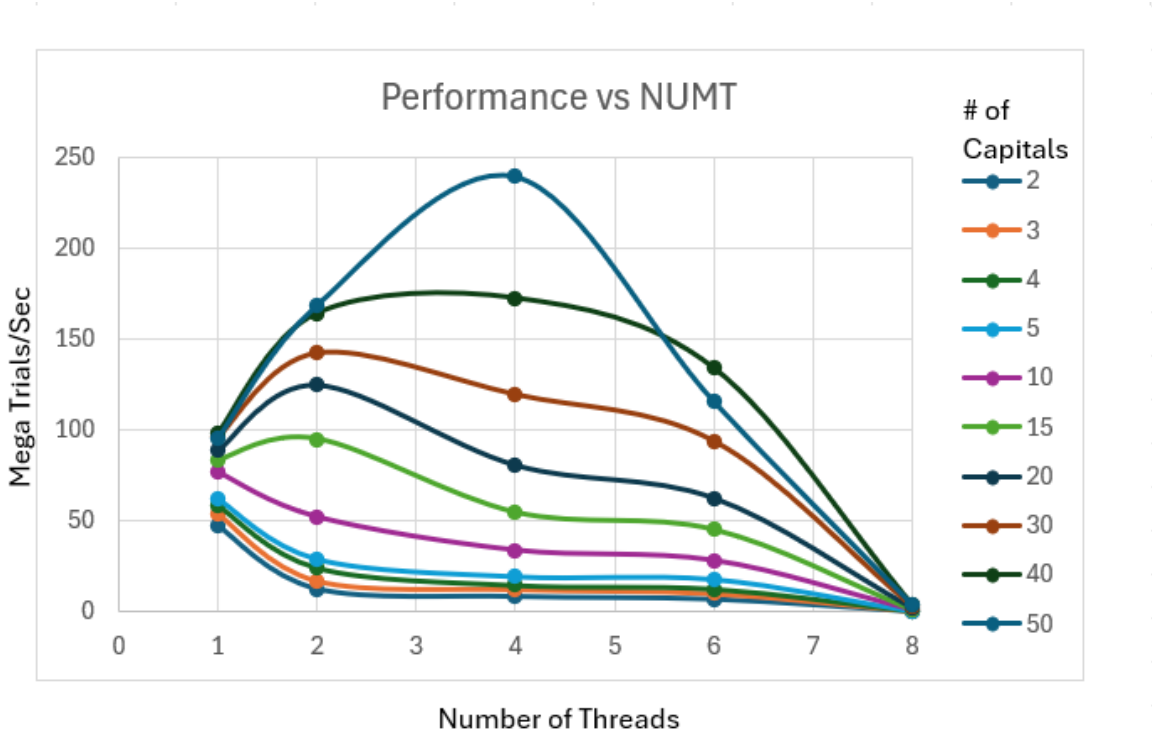
## Environment

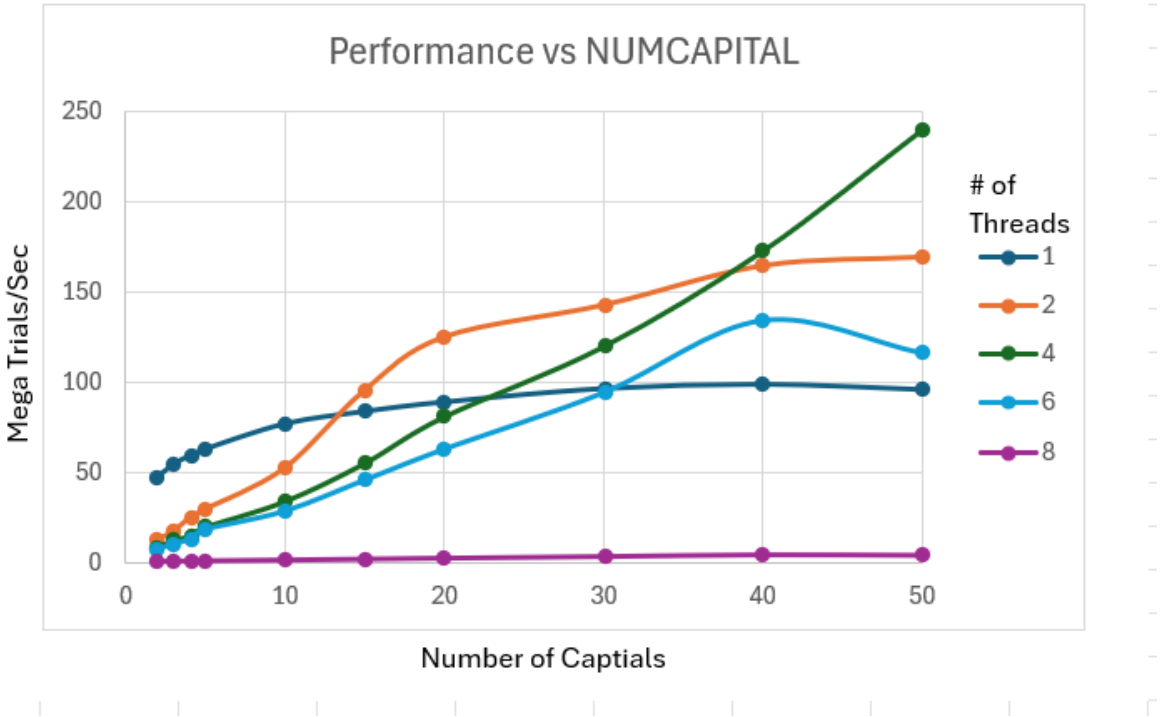
For this project, I used my personal desktop computer running the Windows operating system. To compile and run the program on the school's server, I used Windows PowerShell to connect to the Flip server via SSH. On Flip, I used the GNU g++ compiler to compile the C++ program with OpenMP support.

## Performance Data Table

| | 1 | 2 | 4 | 6 | 8 |
|---|---|---|---|---|---|
| 2 | 47.104 | 12.594 | 8.421 | 6.944 | 0.222 |
| 3 | 54.348 | 17.078 | 12.341 | 10.037 | 0.327 |
| 4 | 58.965 | 24.323 | 14.665 | 12.628 | 0.434 |
| 5 | 62.775 | 29.323 | 19.673 | 18.131 | 0.548 |
| 10 | 76.768 | 52.203 | 33.915 | 28.307 | 1.097 |
| 15 | 83.555 | 95.137 | 54.977 | 45.495 | 1.644 |
| 20 | 88.743 | 124.824 | 80.788 | 62.691 | 2.225 |
| 30 | 96.167 | 142.484 | 119.686 | 94.113 | 3.252 |
| 40 | 98.525 | 164.331 | 172.684 | 134.272 | 4.341 |
| 50 | 95.604 | 169.074 | 239.841 | 116.015 | 4.021 |

## Graphs






## Commentary

From the graphs, I noticed that using more threads helped performance only when there were more cities to process. For small numbers of cities, using many threads actually made the program slower because of the extra overhead. As the number of cities increased, performance improved with more threads, especially from 2 to 4 threads. However, going beyond 4 threads didn't always help and sometimes made things worse. This shows that using too many threads can cause slowdowns, especially if the work isn't large enough to take advantage of them. In short, small tasks work best with fewer threads, and bigger tasks benefit from more threads but only up to a point.

## Extra Credit

| CapitalNum | Longitude | Latitude | CityName | State |
|---|---|---|---|---|
| 0 | 98.33 | 36.5 | OklahomaCity | OK |
| 1 | 83.1 | 42.38 | Detroit | MI |
| 2 | 117.76 | 34.06 | Pomona | CA |
| 3 | 71.02 | 42.08 | Brockton | MA |
| 4 | 77.02 | 38.9 | Washington | DC |
| 5 | 85.14 | 41.09 | FortWayne | IN |
| 6 | 117.46 | 34.11 | Fontana | CA |
| 7 | 92.33 | 38.95 | Columbia | MO |
| 8 | 105.15 | 39.83 | Arvada | CO |

| | | | | |
|---|---|---|---|---|
| 9 | 74.16 | 40.91 | Paterson | NJ |
| 10 | 94.15 | 30.08 | Beaumont | TX |
| 11 | 79.99 | 35.99 | HighPoint | NC |
| 12 | 82.47 | 27.97 | Tampa | FL |
| 13 | 97.75 | 30.3 | Austin | TX |
| 14 | 83.58 | 41.66 | Toledo | OH |
| 15 | 98.53 | 33.91 | WichitaFalls | TX |
| 16 | 112.31 | 33.79 | Peoria | AZ |
| 17 | 78.97 | 35.08 | Fayetteville | NC |
| 18 | 121.31 | 44.06 | Bend | OR |
| 19 | 82.35 | 29.68 | Gainesville | FL |
| 20 | 97.14 | 33.22 | Denton | TX |
| 21 | 104.77 | 40.41 | Greeley | CO |
| 22 | 76.3 | 36.68 | Chesapeake | VA |
| 23 | 96.98 | 33.05 | Lewisville | TX |
| 24 | 121.32 | 38.77 | Roseville | CA |
| 25 | 80.26 | 36.1 | Winston-Salem | NC |
| 26 | 98.25 | 26.22 | McAllen | TX |
| 27 | 89.97 | 35.11 | Memphis | TN |
| 28 | 111.86 | 33.68 | Scottsdale | AZ |
| 29 | 112.45 | 33.67 | Surprise | AZ |
| 30 | 84.51 | 39.14 | Cincinnati | OH |
| 31 | 97.51 | 35.47 | OklahomaCity | OK |
| 32 | 117.93 | 33.89 | Fullerton | CA |
| 33 | 80.07 | 32.92 | NorthCharleston | SC |
| 34 | 71.01 | 42.26 | Quincy | MA |
| 35 | 116.4 | 43.61 | Meridian | ID |
| 36 | 121.82 | 39.76 | Chico | CA |
| 37 | 117.43 | 47.67 | Spokane | WA |
| 38 | 121 | 37.64 | Modesto | CA |
| 39 | 95.69 | 39.03 | Topeka | KS |
| 40 | 122.36 | 37.95 | Richmond | CA |
| 41 | 96.05 | 41.26 | Omaha | NE |
| 42 | 121.38 | 38.41 | ElkGrove | CA |
| 43 | 98.12 | 29.7 | NewBraunfels | TX |
| 44 | 119.04 | 35.35 | Bakersfield | CA |
| 45 | 118.75 | 34.27 | SimiValley | CA |
| 46 | 104.77 | 40.41 | Greeley | CO |
| 47 | 72.54 | 42.12 | Springfield | MA |
| 48 | 112.09 | 33.57 | Phoenix | AZ |
| 49 | 117.13 | 33.49 | Temecula | CA |