

Pedagogical: Visualizing MPI Data Movement and Performance with Tau

When visualizing MPI primitive collective operations and data movement, usually explicit images and animations have to be made to capture generic transformations on data ownership and explicit timing must be instrumented into the code to track performance. Beyond being tedious to make by hand, such visualizations also don't capture specifics of what data is owned by which nodes, and how much thereof. Tools such as Tau and Vampir allow for profiling and automatic data collection of MPI and OpenMP programs execution, performance, and data movement– and also support visualization of said information. For my final project, I would like to use the profiling and tracing data produced by Tau to automatically generate visualizations of data movement under the standard MPI collectives in C++ programs.

The purpose of this project is pedagogical: we all understand the MPI collectives, their performance, and how they move data. What I aim to accomplish is going beyond a lecture style consumption of collectives and their performance (bandwidth, latency, etc) and instead show how we can profile and visualize the execution of an MPI program using Tau, and also derive our own visualizations from the data produced as Tau profiles a program's execution. Specifically, I would like to implement the SUMMA matrix multiplication algorithm in MPI and demonstrate how we can observe data movement and performance beyond using explicit IO operations in the code that we write. This will allow us to visually confirm and analyze performance and actions in a way that we can't with explicit code instrumentation or by viewing generalized animations.

Some useful resources for me:

[Van de Geijn SUMMA Paper](#)

[Tau Overview on Functionality](#)

[Tau User Guide/Install Instructions](#)

In class, I plan to:

- Demo the general SUMMA code, performance metrics we can get with `MPI_WTime()`
- Show the same information extracted and visualized with Tau
- Go through analysis of performance (speedup, efficiency, data movement) using Tau
- Demonstrate visualizing data ownership and collectives through the profiling information collected by Tau

I plan to install Tau and run the MPI programs on my personal computer for demonstration, to avoid any sort of new software installation with the DEAC cluster and so that I can have a reliable source of compute power to actually build the project as I go about it