Name: Tom Herschberg

Mentor: Dr. Howard Pritchard

LA-UR-20-25776

Enhancing the MPI Sessions Prototype for Use on Exa-Scale Systems

One of the new features to be included in the upcoming Message Passing Interface (MPI) 4.0 specification is MPI Sessions. A major goal of Sessions is to provide a more flexible way for applications to allocate and use MPI resources, and thereby potentially expand the application space which can make use of the high performance messaging capabilities provided by MPI implementations. Therefore, it is important to have a working MPI Sessions prototype that can be used to study the performance and behavior of the new MPI Sessions functionalities. Such a prototype has already been developed and implemented in Open-MPI, but it is currently only functional over a limited number of network stacks. In particular, the prototype cannot currently make use of the network stack (OFI libfabric) expected to be the interface of choice on DOE exa-scale systems such as the Argonne Aurora and Oak Ridge Frontier systems. My work this summer involved modifying the Sessions prototype to make it compatible with OFI LibFabric, which offers better performance on these exa-scale systems. With these modifications, the MPI Sessions prototype can be tested and studied on existing systems with LibFabric support including NERSC Cori, Argonne theta, and LANL Trinity.