MPI at Exascale: Challenges for Data Structures and Algorithms

William Gropp

University of Illinois at Urbana-Champaign, USA

Petascale computing is here and MPI continues to succeed as a effective, scalable programming model, despite previous predictions that MPI could not work at the Petascale. As the high-performance computing community considers Exascale systems, will MPI continue to be suitable? Already, a number of studies have looked at the behavior of MPI as the number of tasks approaches one million, and have identified scaling problems in the current state-of-the-art. In addition, the MPI Forum has begun to consider the scalability of the *definition* of some of the routines in MPI, based on the sizes of their input arguments. Concerns have also been raised about other routines, based on the scalability of current implementations.

In this talk, I will look at some of the challenges of implementing MPI efficiently on an Exascale system and identify places where new implementation approaches, for both data structures and algorithms, may allow MPI to scale effectively to exascale systems. New hardware capabilities expected on the next generation of Petascale systems as well as on Exascale systems may also offer new options for the efficient implementation of MPI. These results also suggest fundamental limitations in aspects of the current MPI programming model that can guide the evolution of MPI.