Scalable Shared Memory Parallel Programming: Will One Size Fit All?

Dr. Barbara Chapman

University of Houston

In recent years, there has been much emphasis on improving the productivity of high-end parallel programmers. Efforts to design very large-scale platforms have focused on global address space machines that are capable of concurrently executing many thousands of threads. As a result, new higher level shared memory programming models have been proposed that are intended to reduce the programming effort and directly exploit the capabilities of such systems.

OpenMP was developed to be an industry standard for shared memory parallel programming and it is widely available. In controlled experiments, it has been shown that OpenMP raises the productivity of programmers of SMP platforms. Although most deployments have targeted SMPs with a modest number of CPUs, it has also been successfully deployed on large-scale distributed shared memory systems such as NASA’s Columbia. The consortium that maintains OpenMP is engaged in a revision of this API that, among other things, aims to increase its usability for large shared memory platforms. But it must also be useful for application development on the multicore, massively multithreaded chips that will soon be widely available. In this presentation we describe some of the features that have been proposed for inclusion into OpenMP that address scalability issues, and compare it with other global address space language design efforts. We also discuss the tension between the need for OpenMP to enable programming in the mass market and the desire to have a single API for both high end and chip parallel programming.