Contribution ID: **227** Type: **poster**

A Multicore Communication Architecture for Distributed Petascale Computing

Thursday 26 March 2009 08:00 (20 minutes)

Distributed petascale computing involves analysis of massive data sets in a large-scale cluster computing environment. Its major concern is to efficiently and rapidly move the data sets to the computation and send results back to users or storage. However, the needed efficiency of data movement has hardly been achieved in practice. Present cluster operating systems usually are general-purpose operating systems, typically Linux or some other UNIX variant. UNIX was developed more than three decades ago, when computing systems were all single core. Computation intensive applications and timesharing were the major concerns. Though the UNIX OS family has evolved through the years, Unix network services are not well prepared for distributed petascale computing. The proliferation of multi-core architectures has added a new dimension of parallelism in computer systems. In this paper, we describe a Multi-core Communication Architecture (MCA) for the distributed petascale computing environment. Our goal is to design OS mechanisms that optimize network I/O operations for multi-core systems. In our proposed architecture, MCA vertically partitions CPU cores on a multi-core system, allocating cores for either computation or communication, respectively. Cores dedicated to communication perform "TCP Onloading."MCA will dynamically adjust core partitioning, based on detected system loads. CPU cores could be dynamically reassigned between communication and computation. Combined with Receive-Side Scaling and flow pinning technologies, MCA would perform flow scheduling to ensure interrupt- and connection-level affinity for TCP/IP processing.

Primary author: Dr WU, Wenji (Fermi National Accelerator Laboratory)

Co-authors: Dr CRAWFORD, Matt (Fermi National Accelerator Laboratory); Mr DEMAR, Phil (Fermi National Accelerator Laboratory); Prof. SUN, Xian-He (Illonis Institute of Technology)

Presenter: Dr WU, Wenji (Fermi National Accelerator Laboratory)

Session Classification: Poster session

Track Classification: Grid Middleware and Networking Technologies