

CS554 Project Ideas

MATRIX:HJLSim – Exploring HPC Hierarchical Job Launch and MTC Distributed Task Scheduling at Extreme Scales through Simulation

Overview

The majority of today's High Performance Computing (HPC) job launch is designed around a centralized paradigm, where a centralized server is in charge of all the compute resources. This architecture has scalability problems towards exascale computing. The solution to mitigate the issues is to distribute the whole system in either hierarchical or fully distributed architectures with multiple servers. Resource stealing is applied in a load balancing way among all the servers. In the hierarchical case, each server manages a partition of compute daemons; while in the fully distributed one, sever and compute daemon have one to one mapping, and they are collocated at the same compute node. The hierarchical case has less server state information to maintain but may still have scalability issues for each server, while the fully distributed one is the opposite. This project aims to find the suitable applications for both architectures. In theory, the hierarchical architecture is a good fit for HPC environment where the job size is usually big, while the fully distributed one is better for Many-Task Computing (MTC) applications where the jobs/tasks are usually small (per-core, multi-core or small number of nodes). We will find out the threshold of job size that can distinguish the HPC hierarchical and MTC fully distributed architectures under failures towards exascale through SimMatrix simulator.

Relevant Systems and Reading Material

SimMatrix:

Paper: http://datasys.cs.iit.edu/publications/2013_HPC13-SimMatrix.pdf

Source code: <http://datasys.cs.iit.edu/~kewang/software.html>

Resource Stealing:

Distributed Job Launch technical report is upon request

Key-value Store Simulation:

Paper: http://datasys.cs.iit.edu/publications/2013_SC13-KVS.pdf

Source code: Upon request

Methodology

Simulating both job launch architectures on top of SimMatrix simulator

Preferred/Required Skills

Required: Java, Simulation

Preferred: Discrete Event Simulation, HPC, MTC

Parameters

Hierarchical and fully distributed architectures, job size, different scales

Metrics

Efficiency, speedup

Project Mentor

Ke Wang, <http://datasys.cs.iit.edu/~kewang/>