CS554 Project Ideas

MATRIX:Swift/M - Integrating Swift with MATRIX to Support Large-Scale Scientific Many-Task Computing Applications

Overview

Swift is a scripting language designed for composing application programs into parallel applications that can be executed on multicore processors, clusters, grids, clouds and supercomputers. Swift focuses on the issues that arise from Many-Task Computing (MTC) where many independently distributed computational tasks are executed, composited and coordinated concurrently. MATRIX is a distributed task execution framework for MTC applications, which applies the work stealing technique to achieve distributed load balancing. The workloads that MATRIX supports are extremely large number of independent and dependent jobs/tasks, and the tasks could be per-core, multi-core, and multi-node.

This project aims to integrate Swift with MATRIX to enable MATRIX running large-scale scientific MTC applications. Swift will serve in the middle as a data-driven workflow engine between the top applications and the runtime MATRIX framework. Applications will be directly forwarded to Swift, which then turns the application programs into many parallel or loosely coupled distributed jobs/tasks. These tasks are usually represented using Direct Acyclic Graph (DAG), where the vertices are the task programs and the edges are the dependencies among the tasks. These application DAGs are then going to be executed by the MATRIX framework. Students will pick several different applications that can be converted to different shapes of DAGs (e.g. Bag-of-tasks, fan-in or fan-out, and totally random) by Swift, and run MATRIX with these workloads.

Relevant Systems and Reading Material

Many-Task Computing:

> http://www.cs.iit.edu/~iraicu/research/publications/2009_PhD-UChicago_dissertation.pdf

Swift: http://www.ci.uchicago.edu/swift/main/

MATRIX: http://datasys.cs.iit.edu/projects/MATRIX/index.html

Preferred/Required Skills

Required: C/C++, sockets, multi-threading, Linux

Preferred: Scripting language

Parameters

Different applications for Swift, different scales to run MATRIX

Metrics

Throughput, latency, efficiency

Project Mentor

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