Stony Brook University
The Graduate School

Doctoral Defense Announcement

Abstract

(Title: Scheduling divisible loads for parallel video systems,

real-time distributed networks, and computational Grid)

By

(Name: Taeyoung Lim)

(Abstract of 350 words or less, or recital program)

It has increasingly been significant to provide performance guarantee in terms of computing and communication capability in a large heterogeneous network and Grids. In this dissertation, four scheduling mechanisms considering communication delay, the effect of communication on computation, multiple date source, and deadline are proposed.

First, a simultaneous distribution and collection scheduling method via a multi-port switch in parallel video processing system is proposed. For the proposed mechanism, we analyze the video encoding time and derive a closed-form solution. The results show that the total encoding time is significantly faster than a previous method, Parallel Interlaced (PI), based on a bus network as well as achieve scalability in terms of the number of processors.

Secondly, a scheduling method considering the interference of communication on computation is proposed in detail and analyzed here from the perspective of divisible load theory using sequential distribution in heterogeneous networks and grids. Each processor is divided into two virtual processors with different computing speeds according to the degree of overlapping communication. These two virtual processors are used to obtain one virtual processor with equivalent computing speed. The interference aware scheduling via divisible load theory (DLT) provides more realistic and concrete modeling features.

Third, a multi-source scheduling scheme through network partitioning is proposed in a very large heterogeneous network. Each partition is evaluated by running an optimal single-source sequential load distribution scheduling method whenever network partitioning is repeated. After every evaluation, new partitions could be constructed by transferring sink nodes from one partition to another or rearranging the sequence of sink nodes receiving load from the source node within a partition.

Finally, a method scheduling deadline constrained jobs via divisible load theory (DLT) is proposed. In doing so, computing and communication capability should be guaranteed and each workload be estimated through schedulability test before submission. In this paper, scheduling heuristics via network partitioning are proposed for large scale heterogeneous Grid/cluster systems. The minimum number of nodes is obtained from a homogeneous model are used in schedulability test in the heterogeneous system.

Date: 11/28/2007 **Program**: Electrical and Computer Engineering

Time: 11:00 AM **Dissertation Advisor**: Thomas G. Robertazzi

Place: (Engineering Bldg, Room 301)