CS 525: Parallel Computing

Spring 1999

Problem Set 2: Programming in MPI

Due Date: Mar 4, 1999 in class

No late submissions will be accepted

1. Compute the values of t_s and t_w (in MegaBytes/s) for the IBM SP2. This is done by calling a barrier, ping-ponging a message between a pair of processors say 1000 times and then using it to compute the communication time. Plot the communication time for a range of message sizes (from 1 byte to 1MBytes). Use this plot to fit a line and determine the required values.

2. Write an MPI program to compute a parallel join of two relations. The join is computed as follows: The relations are assumed to be distributed across the processors with n/p records per processor. Each processor hashes each of its records to another processor by applying a hash function to a specified field. The record is then placed in a bucket meant for the processor. Once all the records have been placed in buckets, the buckets are transferred to respective processors. On being received, the processors compute a local join.

   For generating records, assume that each record in both relations have 5 fields, each of which is an integer from 0 - 99. To begin with, each processor can simply generate the n/p records by filling random numbers in the specified range. Processor 0 prompts the user to see which field the join must be computed on. The field number is broadcast to all the processors; which perform the hash.

   Plot two speedup curves for this; one for a fixed n and increasing p, and another for fixed p and increasing n.

3. Write an MPI program for computing all prime numbers in the range 2 to n using the sieve of eratosthenese. Use a block-cyclic distribution of the list across processors. Plot the speedup against n.