A DATA INTENSIVE COMPUTATION ON
A CLUSTER*

Parallel Elementwise Processing

Zoltán Horváth, Zoltán Hernyák, Tamás Közsik, Máté Tejfel, Attila Ulbert
Eötvös Loránd University, Budapest, Hungary
{hz,aroan,kto,mate},mormota}@inf.elte.hu

Abstract We have investigated the applicability of PC clusters with terabyte disk-
servers for data-intensive parallel computing. We used parallel element-
wise processing as our testcase. We have searched for the optimal value
for parameters of the algorithm running on our hardware environment.
The performance of several communication frameworks has been tested,
such as C/PVM, C/MPI, Distributed Haskell and socket interface in C.
On large inputs with heavy operations our implementations showed con-
siderable speedups.

Keywords: parallel elementwise processing, cluster, PVM, MPI, socket, Haskell

Introduction

Elementwise processable functions form a class of many often-used
operations such as merging ordered sequences, computing the union of
sets and updating a database. Elementwise processability means that
the output can be obtained by processing each single data item from
the input one after the other. Both the input and the output can be
multi-dimensional.

A well known example is the higher order function map, which has
as its first argument a function from type a to type b and applies it
elementwise to the elements of a data structure, of which elements having
type a. (Example: map inc [2,5,1] = [3,6,2]).

We generalised this concept and developed a parallel algorithm to
solve the general problem [1]. In this paper we searched for the optimal

*Supported by DemoGrid OMFB 01548 and OTKA T037742

45