

# A Graphical User Interface for Parallel Data Analysis

Atsushi Ogiya

*Graduate School of Engineering, CIMS, Hokkaido University*

*N11 W5, Kita-Ku,*

*Sapporo, 060-0811, Japan*

*atsushi@cims.hokudai.ac.jp*

Hiroyuki Minami, Masahiro Mizuta

*Center for Information and Multimedia Studies, Hokkaido University*

*N11 W5, Kita-Ku,*

*Sapporo, 060-0811, Japan*

*min@cims.hokudai.ac.jp, mizuta@cims.hokudai.ac.jp*

## 1. Introduction

Recently, it often takes unrealistic time to obtain a result when we analyze a lot of data, for example in the case of data mining. Then one of ways to solve this time-consuming problem is to use Parallel Virtual Machine(PVM). PVM allows us to compute quickly such as a super computer, whereas it requires much labor to use. In this paper, we would like to propose a Graphical User Interface for parallel data analysis with PVM.

## 2. PVM

PVM is a software package that was developed by Oak Ridge National Laboratory in the United States of America. It allows a collection of computers hooked together by a network to be used as a single large parallel computer. PVM realizes parallel computing with a mechanism, called “message passing”.

We developed some data analysis programs with PVM, such as clustering or Sliced Inverse Regression with projection pursuit(SIRpp), on Master-Slave model.

## 3. Problems of PVM

PVM is useful for parallel data analysis, but it has some operational problems. Because we operate PVM system with command line, we must type a PVM command with various options, for example to redirect output of a job to console. Moreover, PVM doesn't support command history. Consequently, we need to type the same PVM command repeatedly. Making “hostfile” is manual labor. “Hostfile” is a file containing a list of hostnames that consist of parallel virtual

machine. If we use PVM in a large amount of machines, listing hostnames of all machines must be a time-consuming operation. When we finish programming and preparing to execute it, we still have business to do. The another problem is to copy an executable file to all hosts. It is inefficient to copy it manually.

#### 4. Proposed GUI for PVM

To overcome these problems above, we have developed a GUI system for PVM. This system has many functions, including as follows:

- spawn tasks easily
- make a hostfile semi-automatically
- copy an executable file to all hosts with using hostfile

Figure 1 shows a snapshot of the proposed GUI system.

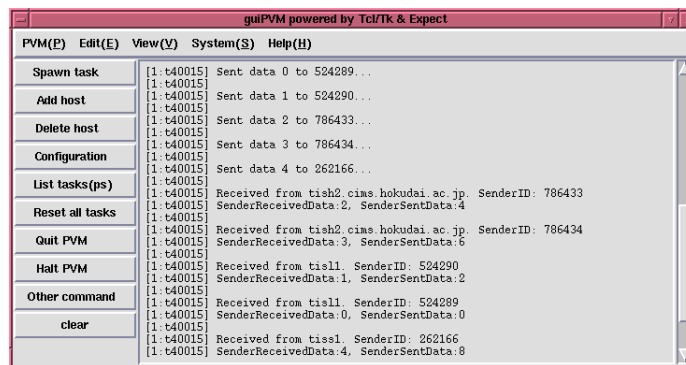


Figure 1: A snapshot of proposed GUI system

#### REFERENCES

<http://www.epm.ornl.gov/pvm/> *Oak Ridge National Laboratory.*

Kawane, M., Komiya, Y., Mizuta, M.(2000). Cluster Analysis with Parallel Virtual Machine. *Proceeding of the Tenth Japan and Korea Joint Conference of Statistics: Statistical Prediction And Computing, Beppu, Japan*, 101–106.

Matsuda, M., Minami, H., Mizuta, M.(2000). A Study on Parallel Computing for SIR with Projection Pursuit. *Proceeding of the Sixty-Eighth annual meeting of Japan Statistical Society, Sapporo, Japan*, 191–192.

#### RÉSUMÉ

Dans ce papier, nous aimerait proposer un Interface utilisateur Graphique pour Données Parallèle avec PVM.