The Organization of Distributed Systems by Means of PVM

Hrach V. Astsatryan

Institute for Informatics and Automation Problems of NAS RA and YSU

Abstract

The following task is beeing discussed in the article. Suppose a distributed system which consists of some computers using Unix operating system, which are servers and provide some services to users.. For managing of the distributed system the author offerse to use PVM programming environment.

There are tasks, solvement of which need large resources and time. The author offers to use the whole free capabalities of those servers, in the result of this could be solved that tasks.

Introduction

The distributed system is representing as a unity of independent computers, which represents for the user like a one connected unit.

The distributed systems have the following advantages from supercomputers:
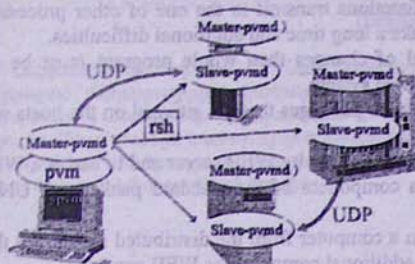
□ **Productivity**: By the Grosh's Law the speed of CPUs is proportional to the square of the cost price. When the microprocessors produced, this low stopped working, because by twice of price could be bought the same processor by more higher speed

□ **Effectivenice**: By collective a large amount of microprocessors' powers could be achieved a higher productivity, which physically couldn't be achieved by supercomputers

□ **Distributivness**: Compared to the supercomputers the distributed systems are really more distributed

□ **Prospectivness**: The spoiling of some components of distributed system will not essentially damage of the system productivity

On the other hand organization of distributed systems has her disadvantages compared to supercomputers:

□ **Software environment**: Software problems could be appeared depends on operating systems, programming languages and applied programs

□ **Connectivity**: The connectivity problems could depends on information loosing, development, changings or network's overloading

□ **Security**: For the organization of the distributed system security is a serious problem. The problem is this: to organize secure transmission of the messages between the components of distributed system

The distributed system will be enough productive, because in case if one of the system components failing, then the modules of current programs will run on the another components of the system.

The power of distributed system could be dynamic changed by means of adding or removing components of this system.

Parallel Virtual Machine (PVM) programming packet gives an opportunity the group of computers correspond like a one whole virtual machine (distributed system), which is intended to solve various kind of problems. The PVM gives an opportunity to collect computers of different kind of architecture like one whole unity.

The first version of PVM programming packet was issued at 1989 by Oak Ridge National laboratory.

The advantages of this software packet are:
- It is easy for the installation and using
- It is very popular. It is widely used Message Passing model
- It is enough to add that computer to the virtual distributed system by having an account on some computer which operating system is Unix
- The realization, managing and changing of distributed system is a dynamic

On the other hand PVM programming environment has the following disadvantages:
- PVM is not standard (it is not like MPI).
- For realization of some functions of message passing model it is unsatisfactory
- As the PVM is depends on the hardware of the computer, therefore it may be work more slowly than other similar programs.

The installation of PVM package is easy, which makes it possible to change rapidly the components of PVM, by adding or removing computers in PVM.
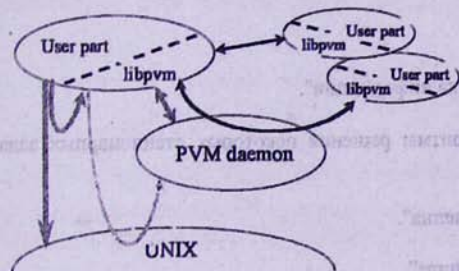
The main functions of PVM programming package are support of communication, administration of processes and the creation of software environment for users. The PVM consists from the following parts:
- From the programming libraries of functions. These functions are written by following languages: C and Fortran.
- Always working program (daemon), which named PVMD. This program correspond to automatically, without interacting with the user which accomplishes the tasks

As we mentioned a PVMD daemon is a working process in Unix system, which looks after the work of processes, which accomplishes from the users, and supports the connection between processes. On each component of the distributed system accomplishes only one daemon. Each daemon consists of detail information about distributed system components. The connections between the components of the distributed system are supported by daemons. The insertion of

the PVM package is essential. In case of daemon's absence the following disadvantages will affect.

□ The user program will occupy a most space and his structure will become more difficult.
□ If the program, which firstly began to work and took care the managing functions of PVM finishing his work, then some part of his functions transmit to the one of other processes. But accomplishment of such function will take a long time and additional difficulties.
□ If in the managing code will be any kind of changes then whole program must be re-compiled on the other components of distributed system.
□ The programming connections between software packages that are situated on the hosts will be more difficult.

It is suggested to manage and use the distributed system by WEB server and to use as a WEB server Apache server, that is reliable and is a component of the standard package of UNIX operating system.

It is also suggested to install WEB server on a computer from the distributed system. In that case there will not occur the necessity to use an additional computer for WEB server.

Suppose that there are servers, which use UNIX operating system. For distributed system creation on the base of those servers will be more correct to use PVM, as it has the following advantages:

□ It was created for UNIX operating system
□ If one has access to UNIX operating system (has account), he can add that server into the distributed system.

The author has created a programming package for management and usage of distributed system. In that package the documents are written in HTML and DHTML programming languages. And the connection with the distributed system is made through script programs, and for them are used Java, C and Perl programming languages, that are very useful instruments for that kind of programs. The user in order to use the resources of the distributed system is applying to the corresponding address.

The page is written in the native language, that makes it possible for local users to use that without language barriers, and for safe use of the resources of the distributed system an identification is being defined for each user. The identification consists of username and password.

References

[1] А. Б. Бабичев, "Параллельная обработка информации".

[2] В. И. Лебедев, "Параллельные алгоритмы решения некоторых стационарных задач математической физики".

[3] Ю. Г. Дадаева, "Параллельные вычисления".

[4] G. Pffister, "Sizing Up Parallel Architectures".

[5] А. Танаев, "Программирование сокетов".

[6] Д. Арапов, "Можно ли превратить сеть в суперкомпьютер". Открытые системы #04.97.

[7] В. Крюков, "Операционные системы распределенных вычислительных систем", http://ergeal.newmail.ru/cs/krukov/index.htm#LECT1.

PVM ծրագրային փաթեթի օգնությամբ տարաբաշխված համակարգի կազմակերպումը

Հ. Վ. Ասցատրյան

Ամփոփում

Հոդվածում ֆիտարկվում է Unix օպերացիոն համակարգ օգտագործող սերվերներից բաղկացած տարաբաշխված համակարգ։ Կան խնդիրներ, որոնց լուծումը պահանջում է մեծ ռեսուրսներ և ժամանակ։ Առաջարկվում է օգտագործել այդ սերվերների "ազատ" հնարավորությունները լուծելու համար ազատ խնդիրները։ Առաջարկվում է որպես ծրագրային միջավայր օգտագործել PVM ծրագրային փաթեթը, որը ղեկավարումն իրականացվում է հեղինակի կողմից գրված ծրագրային միջավայրի կողմից։

Abstract

The idea and properties of measuring and Sollins-Lettree distance are used in history-an of analyzes variables space-dimenzionality had for selection of the most informative sadiks. The spedial computer program was written for realization of suggested methods. It was applied in the statistical analysis of data on geo-meteorological disease.

1 Definition of more informative characters

Let the characters of an object (patient) are random variables $X^{(i)}$ with values $x^{(i)}$, $i = 1, L$. Let $X = (x^{(1)}, x^{(2)}, ..., x^{(L)})$ is an observation with characters of different kinds, that is the vector X includes quantitative, qualitative and classification elements. If $X^{(i)}$ is quantitative character then $x^{(i)}$ is the center of i-th interval of whole diapason of admissible values $x^{(i)}$. If $X^{(i)}$ is a qualitative character then $x^{(i)} = v$ is the degree defining the quality of i-st character. If $X^{(i)}$ is a nominal character then $X^{(i)} = r$ is the index of the class which includes the object.

For unification of all characters analysis, we replace each $X^{(i)}$ by a R_i-dimensional
$$(y^{(i)}, X^{(i)}_1, ..., X^{(i)}_{R_i M_i}) \text{ vector, where}$$

$$X^{(i)}_j = \begin{cases} ... \end{cases}$$

The dimension of observed vector increases from L to $L^* = \sum R_i$.

In following analyzes we assume that all above described transformations already have been done, we denote L-dimensional system of characters by the same X and its value by x.

Let the quantity $z = z^{(1)}, ..., z^{(L*)}$ can be selected from the list of initial indices or can be said as their certain function $\varphi = \varphi$, where φ is the set of all admissible transformations. The presentation of each vector $X = (X^{(1)}, X^{(2)}, ..., X^{(L*)})$ by a M dimensional $1 \le M \le L - 1$, vector $Z = (Z^{(1)}, Z^{(2)}, ..., Z^{(M)})$, is conditioned by the following important aims and reasoning:

- a visual presentation of initial data,
- simplification of calculation.