

Research and Application of Parallel Computing under Linux

Xuechun Wang

School of Information Engineering
Huanghe Science and Technology College
Zhengzhou, China
38856735@qq.com

Quanlu Zheng

School of Computer
Zhuhai College of Beijing Institute of Technology
Zhuhai, China
2268435023@qq.com

Abstract—Parallel computing is in parallel computer system for parallel processing of data and information, often also known as the high performance computing or super computing. The content of parallel computing were introduced, the realization of parallel computing and MPI parallel programming under Linux environment were described. The parallel algorithm based on divide and conquer method to solve rectangle placemen problem was designed and implemented with two processors. Finally, Through the performance testing and comparison, we verified the efficiency of parallel computing.

Keywords—parallel computing; Linux; MPI; placement

I. INTRODUCTION

Advantageous in tremendous numerical calculation and data processing capability, parallel computing can be widely applied in important projects with profound influence in national economy, national defense and technological development, such as oil exploration, earthquake prediction and forecast, climatic simulation and wide-range weather forecast, new weapon design, study and simulation of nuclear weapon, aerospacecraft, satellite image processing, celestial body, earth, and other industries.

The simplest one-dimensional packing problem can be regarded as NP-complete problem[1]. Parallel computing is indeed feasible in solving practical problems though theoretically, it can't solve the NP-complete problems comprehensively. The research on parallel packing algorithm is in accordance with the need of era development since parallelization is able to provide high-quality performance required by engineering and science, quick response in real-time control and essential fault-tolerant processing of high-reliability system.

II. INTRODUCTION TO PARALLEL COMPUTING

A. Concept of Parallel Computing

Parallel computing means to decompose one application into several subtasks on the parallel machine and allocate them to different processors that cooperate with each other and implement subtasks concurrently, thus achieving the goal of accelerating solving or expanding the scale of solution application[2]. It is necessary to equip with the

following three basic conditions to promote parallel computing successfully.

1) Parallel machine

The parallel machine shall include two or more processors, which are connected and communicated with each other through Internet.

2) Degree of parallelism

The application can be decomposed into several subtasks which can be performed concurrently. The progress of decomposing one application into several subtasks is called as design of parallel algorithm.

3) Parallel programming

The parallel algorithm will be achieved specifically and the parallel program will be compiled and operated on the parallel programming environment provided by parallel machine, thus realizing the goal of solving application issue concurrently.

B. The Research Content of Parallel Computing

According to the three essential conditions of parallel computing, its main research content can be divided into the following four aspects.

1) Extraction high-performance feature of parallel machine

It shall attach most emphasis on full understanding and extracting the high-performance feature of systematically structure of current parallel, proposing practical parallel computing model as well as evaluation method of parallel performance, guiding in parallel computing design and achieving parallel procedure.

2) Design and analyze parallel algorithm

High-efficient parallel algorithm can be designed by applying experts' discrete calculation method for solving various application problems, which can decompose the application problem into several subtasks that can be computed concurrently. In addition, the feasibility and efficiency of these algorithms were analyzed in detail.

3) Achieving parallel technology

It includes parallel procedure design and parallel performance optimization. Based on the parallel programming environment provided by parallel machine, for instance, the information transmission platform MPI or share and storage platform OpenMP were used to achieve parallel computing specifically and develop the parallel procedure of application solution. Meanwhile, with combined

consideration to the high-performance feature and practical application characteristics of parallel machine, the performance of parallel procedure was optimized continuously.

4) *Parallel application*

This is the ultimate purpose of parallel computing research. The procedure will be further developed to parallel application software for solving practical problems by verifying and confirming the correctness and efficiency of parallel procedure. Furthermore, the parallel computing and parallel procedure will be improved continuously with consideration to various problems in practical application. The algorithm and procedure without verification and affirmation through practical application can't be regarded as good algorithm and procedure.

In conclusion, the main objective of parallel computing is to solve a batch of great challenging science, engineering and business calculation tasks, aiming to satisfy the demand of continuous increasing application problems on speed and memory resources[3].

III. PROGRAMMING ENVIRONMENT OF PARALLEL COMPUTING

A. *Linux Operating System*

Linux operating system is the standard operating system applied by almost all current high-performance parallel machines (SMP, DSM, cluster, constellation, and MPP). The Linux operating systems manufactured by various parallel machine manufacturers provide general same basic interface for users though they have different achievement principles. Therefore, users can conveniently apply various parallel machines introduced above with a certain understanding on the Linux operating system.

B. *MPI Programming Model*

The global industry, government and research application department launched the standard user interface of parallel programming environment for information transmission together in June, 1994, which divided the parallel programming environment for information transmission into two parts: the first part is the standard interface instructions of all information transmission functions that composed the environment, which were made according to the different requirements of parallel application program on information transmission function without consideration to whether the function can be realized specifically. The second part is the detail achievement of these functions provided by various parallel machine manufacturers. Users can implement the program on a specific parallel machine supporting MPI parallel programming environment only with an understanding on the standard interface of MPI base function and design MIP parallel program. Generally speaking, the MPI system refers to the function base composed of these information transmission functions with standard interface instructions. Compared with other parallel programming environment, MPI possesses many advantages: (1) Good transportability and supported by all current parallel environment; (2) Good expandability and the most reliable

platform of current high-efficiency massive parallel computing (hundreds of processors); (3) Equipped with perfect asynchronous communication function; (4) Possesses accurate definition, thus providing essential conditions for the development of parallel software[4].

C. *Building of Parallel Programming Environment*

Parallel processors: IBM System X3650 2CPU 8-core server, can provide the eight processors in parallel computing environment.

Linux operating system: Fedora 11 DISTRIBUTION

Parallel Programming Model Framework: MPICH installation package. MPICH is the free MPI system that the most widely used currently, it supports almost all Linux/UNIX and the Windows 9x, NT, 2000 and XP operating systems. Using MPICH can in a single computer or workstation build MPI debugging environment, the use of multiple processes to simulate the operation of MPI parallel program, also can be in SMP system or the cluster environment to establish the utility of parallel computing environment. In fact, it is run in the cluster system basically parallel environment.

The MPI programming environment to build steps are as follows.

1) *RSH configuration*

In order to installation and use of MPICH in the single computer, RSH service program and client program must be installed and enabled in Linux. If you use RedHat or Fedora Linux, you need to install the xinetd, rsh-server and RSH packets, and run the command:

```
/sbin/chkconfig xinetd on
/sbin/chkconfig rsh on
/sbin/chkconfig kshell on
```

2) *MPICH compile and installation*

In the simplest case, the compile and installation process only requires the following steps:

```
cd mpich-1.2.x
./configure --prefix=/usr/local
make
su
make install
```

3) *The MPI compiler*

In order to facilitate the user to compile and link MPI program, MPICH provides a set of Shell scripts, including the mpicc, mpicxx, mpif77, mpif90 etc, these script usage and Linux/UNIX common compiler uses exactly the same, is the difference between them at compile time will automatically add the MPICH header file path.

D. *MPICH program running*

IV. PARALLEL PLACEMENT ALGORITHM BASED ON DIVIDE AND CONQUER METHOD

A brief description of the rectangle placement problem as follows: known a rectangular container R_0 , and n blocks of rectangular block of R_1, R_2, \dots, R_n . Put these n rectangular blocks much as possible in the container R_0 , that is, to make the area of the container not utilization as small as possible,

and meet the following restrictions: (1) The edge of any rectangular block placed into the container to must parallel to the edge of the container; (2) Any two rectangular blocks placed in container are not embedded in each other.

According to the size of optional placement blocks relative to the current placement space of the feasible region, the proposed algorithm to sort the placement blocks, each priority put into the placement blocks of feasible region smaller, and be positioned in the upper left corner of the current placement space, and then the remaining placement space is decomposed into two smaller placement space (Figure 1), respectively, repeat the process until no suitable placement can be placed. The placement process is expressed with a binary tree structure, and the original placement space corresponds to the root node of the binary tree, two small placement spaces corresponding to the two sub-nodes for the corresponding node.

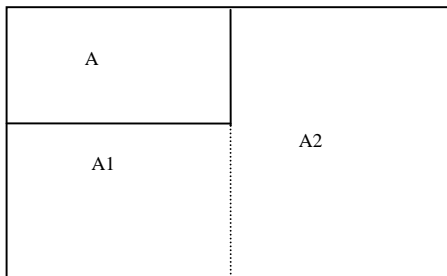


Figure 1. Space decomposition

Through the space decomposition as shown in Figure 1, the original placement problem is gradually break down into smaller sub-placement problem, then the problem is satisfactorily resolved. The algorithm uses two processors PE1 and PE2, the size of optional placement block is stored in shared memory, and sub-placement space is stored in an OPEN table of shared memory. Each time remove the dimension of two sub-placement spaces from the table header then sent to the two processor, local memory CLOSED table (if there is only one sub-placement space, send its size into CLOSED table of the local memory in PE1), respectively, to calculate the size of the feasible region and to sort the optional placement blocks, find the placement blocks respective to placement space for

positioning in the upper left corner of the current placement space, and then split the remaining placement space, respectively, the contents of the respective CLOSED table are instead of two sub-space obtain, at last put the contents of two CLOSED table into the OPEN table header, remove the content of CLOSED table, while the placement blocks placed are removed from the shared memory.

Repeat this process until the OPEN table is empty. If the two processors select a placement block, PE1 priority, while PE2 is recalculated.

Some simulation experiments for the parallel algorithm, the results obtained as same as the serial algorithm. Obviously, compared with the serial algorithm, the parallel algorithm does not reduce the computational complexity of the order, when n is gradually increasing, the speedup tends to 2. The computational complexity of the algorithm does not reduce, when the computing speed has increased by nearly 1 time.

V. CONCLUSION

Taking Linux environment as the platform background, the concept, study content and purpose of parallel computing were introduced in detail in this paper. Furthermore, the experimental environment for parallel computing with multi-processors was established by applying mature MPI programming model to test the rectangle placement problem, confirming the high efficiency of parallel computing.

ACKNOWLEDGMENT

This work was financially supported by the Zhengzhou Science and Technology Plan Foundation (20110346).

REFERENCES

- [1] Consortium Dowsland K.A., Dowsland W.B. Packing Problems. Eur. J.Opt. Res, 1992,56: 2~14.
- [2] Zhang Linbo, Chi xuebin,et.al.Introduction to Parallel Computing. Beijing:Tsinghua University press,2006.7.
- [3] Zhang Yu.Study on the Application of Linux Cluster for Parallel Computing. Guangzhou: Zhongshan University,2009
- [4] Michael J.Quinn.Parallel Programming in C with MPI and OpenMP. Beijing:Tsinghua University press,2004.