The Ministry of Education and Science of the Russian Federation
Lobachevsky State University of Nizhni Novgorod
Computing Mathematics and Cybernetics faculty

The competitiveness enhancement program of the Lobachevsky State University of Nizhni Novgorod among the world's research and education centers

Strategic initiative “Achieving leading positions in the field of supercomputer technology and high-performance computing”

Parallel Programming
for Multiprocessor Distributed Memory Systems

03 Lecture
Derived Data Types, Communicators and Virtual Topologies

Brief description

Nizhni Novgorod
2014

03_LECTURE. DERIVED DATA TYPES, COMMUNICATORS AND VIRTUAL TOPOLOGIES

OBJECTIVES

An objective of the lecture is to study the derived data types as a way to transmit messages that have complex memory placement, the methods of groups and communicators creation including virtual topologies.

ABSTRACT

The usage of derived data types in MPI is discussed. General description is given. Methods of derived types constructing is considered. Additionally data packing/unpacking scheme as an alternative method of transmission of data with complex memory placement is presented. The group of processes and communicator operations are described. The virtual topologies implemented in MPI are considered.

BRIEF OVERVIEW

The first section is devoted to derived data types in MPI. The data necessary to be transmitted may not be located close to each other and may contain the values of different types. The data may be transmitted using several messages (this method will not be efficient because of accumulating the latencies of the number of executed data communication operations). The data can be packed into the format of a continuous vector or derived data type in MPI may be created to describe the placement of data in memory. The derived data type in MPI is the description of a set of the values of the predetermined MPI types; the described values are not necessarily located continuously in the memory. The type is set in MPI by means of the type map in the form of the sequential descriptions of values included into the type, each separate value is described by pointing to the type and the offset of the location address from a certain origin address.

There are four derived data types in MPI: continuous, vector, index, structure. In the first section necessary functions to create, register and unregister of each derived data type are considered. Index type creation is demonstrated on the example of the upper triangle matrix. At the end of section data packing and unpacking functions in MPI are discussed.

The second section focuses on the issues of process and communicator management. Processes are united into groups. The group may contain all the processes of a parallel program or a part of the available processes only. The same process may belong to several groups. The groups

of processes are formed to create communicators on their basis. The groups may be defined on the basis of the available groups only. Necessary functions are considered in the section. To create new communicators the two main methods are used: the duplication of the available communicator, the creation of a new communicator from the subset of the processes of the available communicator. Special simultaneous creation of several communicators is possible using the function MPI_Comm_split. The possibilities of MPI described in the section let to manage the number of processes participating in collective operations and to eliminate the mutual influence of different executed parallel program parts.

The last section of the lecture focuses on the virtual topologies in MPI. The topology may be presented as a graph, where the vertices are the system processors (processes), and the arcs correspond to the available communication links (channels). Initial communicator MPI_COMM_WORLD has a topology of complete graph. We may organize the logical presentation of any necessary virtual topology.

FOR STUDENTS

There are a number of sources, which provide information about MPI. First of all, this is the internet resource, which describes the standard MPI: http://www.mpiforum.org. One of the most widely used MPI realizations, the library MPICH, is presented on http://www.mpich.org/.

The following works may be recommended: Pacheco (1996), Snir, et al. (1996), Group, et al. (1999a). The description of the standard MPI-2 may be found in Group, et al. (1999b). The description of standard MPI-3 may be found at www.mpi-forum.org/docs/mpi-3.0/mpi30-report.pdf.

We may also recommend the work by Quinn (2004), which described a number of typical problems of parallel programming for the purpose of studying MPI.

REFERENCES

1. The internet resource, which describes the standard MPI: http://www.mpiforum.org.
2. One of the most widely used MPI realizations, the library MPICH, is presented on http://www.mpich.org.
3. Quinn, M.J. (2004). Parallel Programming in C with MPI and OpenMP. – New York, NY: McGraw-Hill.
4. Pacheco, P. (1996). Parallel Programming with MPI. - Morgan Kaufmann.
5. Snir, M., Otto, S., Huss-Lederman, S., Walker, D., Dongarra, J. (1996). MPI: The Complete Reference. – MIT Press, Boston, 1996.

6. Group, W., Lusk, E., Skjellum, A. (1999). Using MPI – 2nd Edition: Portable Parallel Programming with the Message Passing Interface (Scientific and Engineering Computation). – MIT Press.
7. Group, W., Lusk, E., Thakur, R. (1999). Using MPI-2: Advanced Features of the Message Passing Interface (Scientific and Engineering Computation). – MIT Press.

EXERCISES

1. Develop a sample program for each method of constructing the derived data types available in MPI.
2. Develop a sample program using data packing and unpacking functions. Carry out the experiments and compare the results to the results obtained in case of the use of the derived data types.
3. Develop the derived data types for the rows, columns and diagonals of matrices.
4. Develop a sample program for the Cartesian topology.
5. Develop a sample program for a graph topology.
6. Develop subprograms for creating a set of additional virtual topologies (a star, a tree, etc.).

TEST QUESTIONS

1. The derived data type in MPI is
   a. (+) The description of a set of the values of the predetermined in MPI type. The described values are not necessarily located continuously in the memory.
   b. The sequence of data value descriptions in the type. Each value is described using displacement from some base address.
   c. The description of data set of base algorithmic language types in terms of this language itself.
2. The type signature of the derived types in MPI is
   a. (+) The part of the type map, which contains only the types of values.
   b. The part of the type map, which contains only the displacements of values.
   c. The size of memory in bytes necessary for the storing of one element from the type.
3. The extent of the derived types in MPI is
   a. (+) The amount of memory in bytes that should be allocated for a derived type element.
   b. The number of bytes occupied by the data in a derived type element only.
   c. Displacement of the first byte of a derived type.

4. The derived type in the vector constructing method is
    a. (+) A number of blocks of the initial type elements with regular interval in the memory between blocks.
    b. Continuous sequence of elements of the initial type.
    c. A set of blocks of different sizes of the initial type elements and memory space between blocks may be different.
5. The derived type in the index constructing method is
    a. (+) A set of blocks of different sizes of the initial type elements and memory space between blocks may be different.
    b. A number of blocks of the initial type elements with regular interval in the memory between blocks.
    c. Continuous sequence of elements of the initial type.
6. What is the difference between vector and H-vector (index and H-index) derived type?
    a. (+) Intervals between blocks are set in bytes (not in the initial data type elements).
    b. H-vector (h-index) derived type may contain a sequence of elements of the initial data type with regular interval between them.
    c. H-vector (h-index) derived type may contain a sequence of elements of the initial data type with different interval between them.
7. The derived type in the structural constructing method is
    a. (+)The most general constructing way for creating the derived data type, when the corresponding type map is set directly.
    b. A number of blocks of the initial type elements with regular interval in the memory between blocks.
    c. Continuous sequence of elements of the initial type.
    d. A set of blocks of different sizes of the initial type elements and memory space between blocks may be different.
8. What virtual topologies are supported in MPI?
    a. (+) A rectangular grid with any dimension.
    b. A complete graph only.
    c. (+) A graph with any structure.
9. Cartesian topology is the presentation of a set of processes as
    a. (+) Rectangular grid.
    b. A complete graph.
    c. A graph with any structure.
10. What topology the initial communicator MPI_COMM_WORLD has?

a. Rectangular grid.
b. (+) A complete graph.
c. A graph with any structure.