

Использование технологий MPI и OpenMP для организации обменов сообщениями в вычислительных системах с кластерной архитектурой*

Е.В. Аксенова, М.Л. Цымблер

1. Введение

Стандарт MPI предназначен для поддержки работы параллельных процессов в терминах передачи сообщений и ориентирован на многопроцессорные системы с распределенной памятью. Стандарт OpenMP используется для разработки параллельных приложений в модели общей памяти. Данные стандарты могут быть использованы совместно в рамках создания параллельного приложения для проведения научно-инженерных расчетов на кластерной вычислительной системе [1]. В данной работе предлагается использовать подход MPI+OpenMP для организации обменов сообщениями в однородных вычислительных кластерах.

2. Гибридная передача сообщений

Предлагаемая гибридная технология передачи сообщений может быть кратко описана следующим образом (без существенного ограничения общности мы можем полагать, что кластер состоит из однотипных узлов, имеющих архитектуру SMP).

Выполнение параллельного приложения на кластере осуществляется следующим образом. С помощью MPI-загрузчика программа запускается на одном процессоре каждого из SMP-узлов кластера. Наличие в исходном тексте данной программы директивы стандарта OpenMP `#pragma omp parallel` влечет за собой автоматическое создание нитей, количество которых совпадает с количеством процессоров в узле кластера. Одна из этих нитей (корневая) соответствует MPI-процессу, запущенному на данном процессоре узла. Остальные нити автоматически распределяются на оставшиеся незанятыми процессоры данного узла кластера. Таким образом, мы получаем выполняющиеся на кластере MPI-процессы и OpenMP-нити, которые могут рассматриваться нами как процессы исходного параллельного приложения.

Передача сообщений между процессами, которые принадлежат одному узлу кластера, реализуется на базе директив стандарта OpenMP, определяющих разделяемый нитями буфер данных. Передача сообщений между процессами, которые принадлежат различным узлам кластера, реализуется на базе коммуникационных функций стандарта MPI. Для реализации описанной технологии нами была разработана библиотека системных функций, названная MIX (Messaging Interface eXtension). Интерфейсы основных функций данной библиотеки приведены на Рис. 1.

<pre>// Отправить сообщение int MIX_Send(void*, int, int, int); // void* buf - буфер с сообщением // int length - длина сообщения // int dest - получатель сообщения // int tag - тег сообщения // Результат: 1 - успех, 0 - неудача</pre>	<pre>// Принять сообщение int MIX_Recv(void*, int, int, int); // void* buf - буфер с сообщением // int length - длина сообщения // int src - отправитель сообщения // int tag - тег сообщения // Результат: 1 - успех, 0 - неудача</pre>
--	--

Рис. 1. Интерфейсы основных функций библиотеки гибридной передачи сообщений

Литература

1. Крюков В.А. Разработка параллельных программ для вычислительных кластеров и сетей // Информационные технологии и вычислительные системы. –2003. –Вып. 1-2. –С. 42-61.

* Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект 06-07-89148) и Южно-Уральского государственного университета (проект 2006112).