

Исследование конкуренции за общие ресурсы при выполнении параллельных программ на кластерных системах

М.Р. Халиуллина, А.Л. Штангеев, А.В. Юлдашев

Узлы современных кластерных вычислительных систем строятся на базе многоядерных процессоров. При выполнении параллельных программ на узлах может находиться множество процессов (потоков), конкурирующих за общие ресурсы: кэш- и оперативную память, дисковую систему, а также каналы передачи данных, что негативно сказывается на эффективности выполнения программ. Для того чтобы минимизировать задержки, возникающие при использовании общих ресурсов, необходимо комплексно учитывать характеристики программ и архитектуру высокопроизводительного кластера при распределении задач на узлы вычислительной системы. Однако в существующие свободно-распространяемые и коммерческие системы пакетной обработки и планировщики не вложены алгоритмы планирования, необходимые для оптимального использования общих ресурсов многоядерных узлов.

На основе экспериментальных исследований [1-2] и результатов теории массового обслуживания нами разработана модель конкурентного использования канала передачи данных коммуникационной среды Infiniband MPI-процессами, выполняющимися на одном узле, при пересылке сообщений размером более 128 Кбайт. Модель базируется на модели Хокни и позволяет оценить время коммуникаций при наличии нескольких, конкурирующих за канал передачи данных MPI-процессов. В модели учитываются характеристики коммуникационной среды (латентность и пропускная способность), а также характеристики MPI-программ (количество пересылок и сетевых пакетов, передаваемых на одной итерации, среднее время вычислений на одной итерации и количество процессов, конкурирующих за канал передачи данных). Предложен метод назначения программ на узлы кластерной системы, который позволяет минимизировать время выполнения MPI-программ благодаря учету задержек, возникающих при конкуренции за каналы передачи данных. Проведено экспериментальное сравнение предложенного метода с некоторыми существующими (Best Fit и Least Utilized Node First) при назначении на узлы тестовых MPI-программ, реализующих коммуникации средствами функции MPI_Alltoall. Использование нового метода позволило сократить время выполнения программ на 6-20%.

Разработана модель конкурентного доступа к памяти для многоядерных SMP систем, которая позволяет оценить среднее время выборки из памяти при наличии конкуренции на различных уровнях иерархии памяти между потоками одной или нескольких многопоточных программ. Конкурентный доступ к иерархии памяти моделируется с помощью экспоненциальной сети массового обслуживания. В модели учитываются характеристики подсистемы памяти (средние времена обслуживания на каждом уровне иерархии) и характеристики многопоточных программ (среднее количество тактов обработки информации на одну выборку из памяти, уровни промахов в L1, L2 кэш, количество конкурирующих потоков). Предложены подходы для оценки ускорения многопоточных программ при монопольном выполнении на узле и в режиме пакетной обработки, с учетом задержек, возникающих при конкурентном доступе к памяти.

Планируется всеобъемлющее тестирование предложенных моделей на реальных приложениях, а также разработка новых алгоритмов планирования и их внедрение в рамках программного комплекса автоматизированных расчетов на кластерных системах, разрабатываемого в УГАТУ, в целях более эффективного использования ресурсов вычислительного кластера.

Литература

1. Халиуллина М.Р., Юлдашев А.В. Тестирование коммуникационной среды суперкомпьютера УГАТУ для решения задачи балансировки нагрузки // ПаВТ'2009: Тр. межд. науч. конф. Челябинск: ЮУрГУ, 2009. С. 826.
2. Юлдашев А.В. Балансировка нагрузки на основе сети в рамках программного комплекса автоматизированных расчетов на кластерных системах // Акт. пробл. в науке и техн. Том 1. Уфа: Изд-во «Диалог», 2009. С.573-577.