

Исследование эффективности различных вариантов реализации коллективной операции Allgather на гибридном вычислительном комплексе МВС-Экспресс

Д.Л. Аверичева

ФГУП «НИИ «Квант»

При решении многих вычислительных задач на многопроцессорных системах часто используется несколько стандартных коллективных операций. Одной из базовых операций является операция Allgather. К примеру, Allgather используется при умножении матрицы на вектор, а, значит, во многих вычислительных задачах. Целью данного исследования является разработка эффективной реализации коллективной операции Allgather на гибридном вычислительном комплексе МВС-Экспресс. В МВС-Экспресс реализована программно-аппаратная поддержка общей памяти, основным средством разработки параллельных программ является библиотека односторонних коммуникаций shmem-express. Мотивация работы состояла в том, что на машинах с аппаратной поддержкой общей памяти реализация коллективных операций с использованием односторонних коммуникаций оказывалась эффективней реализации библиотеки MPI.

Гибридный вычислительный комплекс МВС-Экспресс состоит из 8-ми узлов, объединенных полносвязным коммутатором PCI-Express. В состав каждого вычислительного узла входят 2 четырехядерных процессора Opteron с частотой 2,6 ГГц и оперативной памятью 16ГБ.

Библиотека shmem-express обеспечивает взаимодействие независимых процессов, каждый – со своей локальной памятью, занумерованных по порядку от нуля, и в этом она похожа на MPI. Отличие от MPI состоит в том, что обмены данными между процессами являются односторонними. Для чтения/записи данных в память удаленного узла требуется знать локальный адрес на удаленном узле и номер узла.

Разработка эффективных алгоритмов на MPI подразумевает использование «рукопожатий», что при небольшой длине и большом количестве отправляемых сообщений значительно снижает производительность. При использовании библиотеки shmem необходимости применения двусторонних коммуникаций нет, записать или прочитать данные из памяти можно без ведома удаленного процесса. Однако, для разработки эффективных программ на shmem необходимо как можно реже прибегать к общей синхронизации процессов, например, при помощи введения неблокирующих операций можно значительно повысить производительность. Кроме того, прирост производительности дает также использование дополнительных функций, обеспечивающих прямой доступ к небольшим областям памяти удаленных процессов.

Литература

1. R.Thakur, R.Rabenseifner, W.Gropp. Optimization of Collective Communication Operations in MPICH
2. Лацис А.О. Вычислительная система МВС-Экспресс, http://www.kiam.ru/MVS/research/mvs_express.html.
3. Горбунов В.С., Лацис А.О., Иванов А.Н. О построении суперкомпьютеров на основе интерфейса PCI-EXPRESS //Материалы международной научно-технической конференции СКТ-2010, 27 сентября — 2 октября 2010, Дивноморское, Россия. — Таганрог.
4. Соколов А.А. Особенности модели параллельного программирования для архитектур с сетью МВС-Экспресс. XXXIX Международная научно-практическая конференция «НЕДЕЛЯ НАУКИ СПбГПУ», декабрь 2010.