

Исследование производительности пакета молекулярной динамики GROMACS на различных архитектурах

А.В. Швецов^{1,2}, Е.П. Петухов²

ФГБУ Петербургский институт ядерной физики им. Б.П. Константинова НИЦ Курчатовский Институт¹,
Санкт-Петербургский политехнический университет Петра Великого²

В данной работе рассмотрены аспекты использования программного пакета GROMACS [1] на различных архитектурах современных вычислительных систем таких, как стандартные кластеры, кластеры с GPU, а так же массивно параллельная система РСК Петастрим. Для тестирования производительности использовался набор реальных систем различного размера и состава (от 140к атомов до 2.8M атомов). По результатам тестов приведён набор рекомендаций, позволяющий достичь оптимальной производительности для систем произвольного размера.

Ключевые слова: молекулярная динамика, GROMACS, высокопроизводительные вычисления, GPU, Intel Xeon Phi

1. Введение

В настоящее время одним из основных потребителей времени на высокопроизводительных вычислительных системах являются задачи, связанные с биофизикой и материаловедением. Большой процент таких задач связан с расчётами молекулярной динамики. Одним из наиболее распространённых программных пакетов для расчётов молекулярной динамики является GROMACS. Таким образом исследование производительности пакета GROMACS в зависимости от заданных параметров является актуальной задачей, нацеленной на получение набора рекомендаций для получения максимальной производительности данного программного пакета в зависимости от заданной системы.

2. Описание программного пакета GROMACS

Программный пакет GROMACS, предназначенный для решения задач молекулярной динамики, является одним из наиболее распространённых приложений на современных суперкомпьютерах. Это обусловлено не только актуальностью данного подхода к решению задач биофизики, но и высокой параллельной эффективностью решателя. На сегодняшний день вычислительное ядро GROMACS адаптировано и оптимизировано для множества микроархитектур, в частности, поддерживаются современные расширения наборов команд x86_64 (для Intel и AMD), архитектуры ppc, arm и другие. Реализована поддержка ускорителей вычислений на базе GPGPU (технология CUDA начиная с версии 4.6, а так же технология OpenCL начиная с версии 5.1).

Начиная с версии GROMACS (5.0) проведена существенная реструктуризация кода для создания платформонезависимого слоя инструкций, который упрощает переход на различные архитектуры наборов команд (АНК). Новую АНК можно добавить путем переопределения макросов в файле заголовков. Начиная с версии 5.0 была добавлена поддержка сопроцессора Intel Xeon Phi поколения KNC (работа в нативном режиме). А начиная с версии 2016 реализована так же поддержка архитектуры Xeon Phi поколения KNL.

3. Методика тестирования производительности

В качестве критерия производительности выступает скорость решения задачи, выражающая количество нс траектории молекулярной динамики вычисленных за сутки (нс/сутки) а так же параметр средней несбалансированности системы. Для тестирования использовался набор уже уравновешенных систем, что позволяло рассчитывать достаточно короткие траектории для оценки средней скорости решения. Рассматривались различные соотношения MPI процессов и OpenMP потоков.

Запуск GROMACS производился с помощью командного файла (скрипта), устанавливающего необходимые переменные окружения и параметры задачи. Сравниваются результаты производительности для разных комбинаций потоков и процессов при одинаковом числе использованных процессоров, графических ускорителей NVIDIA K40, результаты для разных количеств сопроцессоров, показывающих масштабируемость задачи и решателя.

4. Выводы и рекомендации по использованию GROMACS

Как показали проведённые тесты на различных системах программный пакет GROMACS [1] имеет хорошую масштабируемость как на обычных узлах, так и на узлах содержащих ускорители Nvidia K40x. При этом для достижения схожих параметров производительности (нс/день) для заданной системы в среднем требуется в 2.5-3 раза (что есть хороший результат, так как программный код GROMACS исполняемый на узле имеет высокую степень оптимизации для современных процессоров) меньше узлов с двумя ускорителями K40x чем обычных узлах.

Предпочтительная конфигурация для запуска на вычислительных узлах с ускорителями K40x это 2 MPI ранка по 28 openmp потоков. Это позволяет наиболее эффективным образом использовать ресурсы данного сегмента. При этом при подборе числа узлов стоит исходить из того, что бы на каждый вычислительный поток CPU приходилось не менее 800 частиц.

Предпочтительная конфигурация для запуска на вычислительных узлах без GPU это 4 MPI ранка по 14 openmp потоков. При этом для оценки числа узлов стоит исходить из того, что бы на каждый вычислительный поток приходилось на менее 200 частиц.

GROMACS способен масштабироваться на системах на базе процессоров Intel Xeon Phi в нативном режиме вплоть до 50 частиц на поток, однако при использовании таких систем, очень важно правильно подобрать соотношение количества MPI ранков и OpenMP потоков. Наиболее перспективным вариантом оказалось использование соотношение 20 MPI ранков и 12 OpenMP потоков на карту.

Литература

1. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX* 1-2 (2015) 19 – 25.

Performance analysis of GROMACS molecular dynamics package on different architectures

A.V. Shvetsov^{1,2}, E.P. Petukhov¹

Peter The Great St. Petersburg Polytechnic University¹, Petersburg Nuclear Physics Institute NRC "Kurchatov Institute"²

This paper discusses aspects of the application of software package GROMACS [1] on different architectures of modern computing systems such as standard clusters, clusters with the GPU, as well as massively parallel RSC Petastream system with processors Xeon Phi co-processors. For performance analysis we used a set of real systems of different sizes and composition (from 140K atoms to 2.8M atoms). According to the test set results a set of recommendations is presented to achieve optimal performance for systems arbitrary size.

Keywords: Molecular Dynamics, GROMACS, HPC, GPU, Intel Xeon Phi

References

1. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX* 1–2 (2015) 19 – 25.