

Международная конференция
МАТЕМАТИЧЕСКОЕ
МОДЕЛИРОВАНИЕ
И ВЫЧИСЛИТЕЛЬНАЯ ФИЗИКА

Объединённый институт ядерных исследований
Лаборатория информационных технологий
г. Дубна, Россия
8–12 июля, 2013 г.

Труды конференции

The International Conference
MATHEMATICAL MODELING
AND COMPUTATIONAL PHYSICS

Joint Institute for Nuclear Research
Laboratory of Information Technologies
Dubna, Russia
July 8–12, 2013

Proceedings of the Conference

International Program Committee

Chairman: B. N. Chetverushkin (KIAM RAS, Moscow, Russia).

Vice-chairman: A. L. Afendikov (KIAM RAS, Moscow, Russia).

Vice-chairman: V. V. Ivanov (JINR).

Members:

Gh. Adam (JINR).

I. Antoniou (Aristotile Univ. of Thessaloniki, Greece).

S. A. Abramov (CC RAS, Moscow, Russia).

A. Bot (INCDTIM, Cluj-Napoca, Romania).

S. Budnyam (Mongolian Development Institute, Ulan-Bator, Mongolia).

J. Buša (TU, Košice, Slovakia).

S. Dimova (FMI SU, Sofia, Bulgaria).

M. Dulea (IFIN-HH, Bucharest, Romania).

V. Friese (GSI, Darmstadt, Germany).

V. P. Gerdt (JINR).

U. H. E. Hansmann (USA).

M. Hnatich (JINR).

J. Honkonen (National Defence University, Helsinki, Finland).

Chin-Kun Hu (IP Acad. Sinica, Taipei, Taiwan).

V. G. Kadyshevsky (JINR).

I. V. Kisel (Goethe Univ., Frankfurt am Main, Germany).

P. Kopčanský (IEP SAS, Košice, Slovakia).

V. V. Korenkov (JINR).

G. Yu. Kryuchkyan (YSU, Yerevan, Armenia).

N. A. Kudryashov (MEPhI, Moscow, Russia).

V. D. Lakhno (IMPB RAS, Pushchino, Russia).

V. S. Melezhik (JINR).

O. A. Mornev (ITEB RAS, Pushchino, Russia).

S. V. Polyakov (KIAM RAS, Moscow, Russia).

R. V. Polozov (ITEB, Pushchino, Russia).

I. V. Puzynin (JINR).

H. Safouhi (Alberta Univ., Edmonton, Canada).

S. Scott (Queen's Univ., Belfast, Northern Ireland).

G. Semanišin (Šafárik Univ., Košice, Slovakia).

L. A. Sevastianov (PFU, Moscow, Russia).

Yu. H. Shoukouryan (NAS RA, Yerevan, Armenia).

V. V. Voevodin (RCC MSU, Moscow, Russia).

T. Zhanlav (National University of Mongolia, Ulan-Bator, Mongolia).

P. Zinterhof (Salzburg Univ., Austria).

P. V. Zrelov (JINR).

Распределённые и параллельные вычисления в науке и технике

УДК 519.684.6 519.85 PACS 02.70.-с 07.05.-т

Метод фильтра Калмана для реконструкции траекторий заряженных частиц в эксперименте СВМ и его параллельная реализация на многоядерном сервере ЛИТ ОИЯИ

Т. О. Аблязимов*, М. В. Зызак[†],
В. В. Иванов*, П. И. Кисель*

** Лаборатория информационных технологий
Объединённый институт ядерных исследований*

ул. Жолио-Кюри, д. 6, Дубна, Московская область, Россия, 141980

*[†] Франкфуртский университет имени Иоганна Вольфганга Гёте
Грунебургплац, д. 1, Франкфурт-на-Майне, Германия, 60323*

Пучки тяжёлых ионов высокой интенсивности, которые будут предоставляться на будущих ускорителях FAIR (Facility for Antiproton and Ion Research), в сочетании с готовящимся экспериментом СВМ (Compressed Baryonic Matter) открывают превосходные возможности для изучения барионной материи при сверхвысоких плотностях и умеренных температурах в лабораторных условиях. Физическая программа СВМ нацелена на изучение структуры и поведения барионной материи при плотностях, сопоставимых с плотностями в центре нейтронных звёзд. Она включает в себя: 1) установление фазовой границы между адронной и партонной материями, 2) определение критической конечной точки, а также 3) поиск указаний на начало восстановления киральной симметрии при высоких чистых барионных плотностях. Нахождение параметров траекторий заряженных частиц является одной из основных задач эксперимента СВМ. Предполагается полная реконструкция событий в режиме реального времени, что требует развития быстрых алгоритмов, максимально использующих потенциал современных архитектур CPU и GPU. В настоящей работе приведены результаты анализа алгоритма реконструкции траекторий заряженных частиц на основе фильтра Калмана, реализованного с применением различных методов распараллеливания кода. Для работы использовался многоядерный сервер ЛИТ ОИЯИ с двумя CPU Intel Xeon X5660 и GPU NVidia GTX 480.

Ключевые слова: процессор, видеокарта, гибридная система, фильтр Калмана, фитирование треков, СВМ, OpenMP, OpenCL, NVIDIA, Intel.

1. Введение

В настоящее время в GSI (Дармштадт, Германия) на строящемся ускорительном комплексе FAIR (Facility for Antiproton and Ion Research) ведутся работы по созданию экспериментальной установки СВМ (Compressed Baryonic Matter). Физическая программа СВМ нацелена на всестороннее изучение фазовой диаграммы сильно взаимодействующей материи и уравнения состояния вещества при экстремально высоких плотностях барионной материи [1]. Для её реализации необходимы измерения распределений множественности в фазовом пространстве и потоков всех интересных с точки зрения физической программы частиц.

Быстрая реконструкция траекторий заряженных частиц с помощью детектора STS (Silicon Tracking System) является ключевой проблемой в задаче отбора полезных событий в эксперименте СВМ. Высокая множественность событий (до 1000 треков в каждом ядро-ядерном соударении), интенсивный фон (до 85% фоновых отсчётов STS детекторе), неоднородное магнитное поле и необходимость реконструкции всех событий в режиме реального времени (до 10^7 событий в 1

секунду) требуют не только развития новых подходов для решения рассматриваемой задачи, но и максимального использования потенциала современных многоядерных CPU/GPU архитектур.

Процедура реконструкции траекторий заряженных частиц, регистрируемых с помощью вершинного детектора СВМ, включает два последовательных этапа:

- 1) распознавание треков в условиях высокой множественности и плотности заряженных частиц, интенсивного фона и неоднородного магнитного поля,
- 2) восстановление параметров трека (место попадания в координатный детектор и направление трека) и импульса заряженной частицы.

В настоящей работе приведены результаты исследования производительности алгоритма реконструкции траекторий заряженных частиц, реализованного на основе фильтра Калмана. Сервер был оснащён двумя CPU Intel Xeon X5660 и GPU NVidia GTX 480. Для оценки алгоритма использовались различные подходы для распараллеливания и векторизации программного кода: заголовочные файлы, средства библиотеки Vc (Vector Classes) [2], программные среды OpenMP (Open Multi-Processing) [3] и OpenCL (Open Computing Language) [4].

2. Требования, предъявляемые к алгоритму

К алгоритму реконструкции параметров треков предъявляются два основных требования:

- 1) максимально возможная точность восстановления (x, y) -координат места попадания трека в конкретный координатный детектор, наклонов трека в этой точке и импульса частицы;
- 2) высокая скорость реконструкции треков в реальном времени эксперимента.

Первое играет ключевую роль, в частности, при реконструкции типа и места распада исследуемых в эксперименте наблюдаемых. В свою очередь, скорость выполнения алгоритма крайне важна для СВМ, поскольку планируется проводить реконструкцию событий в реальном времени эксперимента.

3. Точность восстановления треков методом фильтра Калмана

Для оценки точности реконструкции параметров трека используем такое понятие, как остаток ρ . При этом остаток ρ_x , в частности, для x -координаты трека в одном из детекторов STS-системы определяется как разность между величиной x_{mc} , полученной в результате Монте-Карло моделирования прохождения трека через детектор, и значением x_{reco} , реконструированным с помощью рассматриваемого алгоритма:

$$\rho_x = x_{reco} - x_{mc}. \quad (1)$$

В качестве характеристики надёжности реконструкции параметров трека используются нормированные остатки (пулы):

$$P(x) = \frac{\rho_x}{\sqrt{C_{xx}}}, \quad (2)$$

где C_{xx} — диагональный элемент, соответствующей ковариационной матрицы, полученной в результате реконструкции трека. В идеальном случае пулы должны быть распределены по закону Гаусса с единичной дисперсией.

Для оценки качества реконструкции треков с помощью рассматриваемого алгоритма нами использовались треки, найденные в STS-системе эксперимента СВМ методом клеточного автомата [5, 6]. Для этого было отобрано 20000 “длинных” первичных треков, координаты которых были зарегистрированы во всех STS-станциях. Качество реконструкции треков оказалось одинаковым для всех реализаций упомянутых выше средств векторизации и распараллеливания кода.

Результаты тестирования представлены на рис. 1. Поскольку распределения для x - и y -координат и соответствующих им наклонам абсолютно идентичны, то результаты приведены только для переменной x .

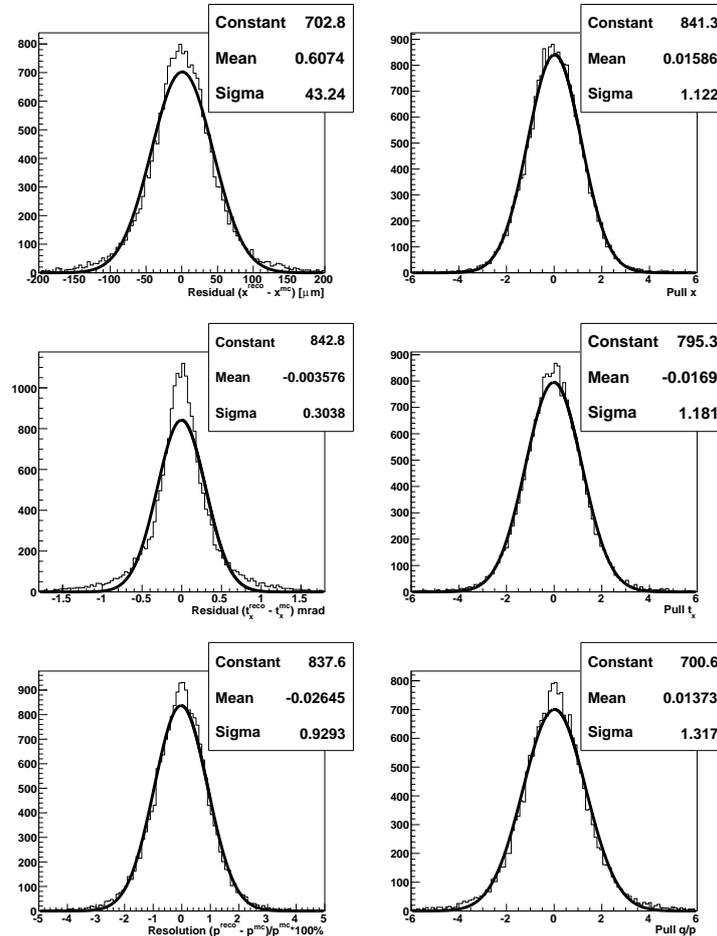


Рис. 1. Распределение остатков и пулов для параметров треков, реконструированных с помощью фильтра Калмана

Распределения пулов хорошо аппроксимируются нормальным законом с дисперсией, близкой к единице. Этот результат указывает на корректность процедуры фитирования.

4. Масштабируемость алгоритма на центральных процессорах

Для оценки производительности алгоритма использовался многоядерный сервер cuda.jinr.ru ЛИТ ОИЯИ, оснащённый двумя процессорами Intel Xeon X5660, каждый из которых содержит 6 физических ядер с частотой 2.8 ГГц. Используя технологию гиперпоточности, одновременно может быть запущено 24 потока.

Для распараллеливания алгоритма на ядра CPU использовались две программные среды: OpenMP и OpenCL.

На рис. 2 приведены графики масштабируемости алгоритма реконструкции треков заряженных частиц в зависимости от числа запущенных в среде OpenMP логических ядер при использовании для векторизации кода заголовочных файлов (слева) и библиотеки Vc (справа).

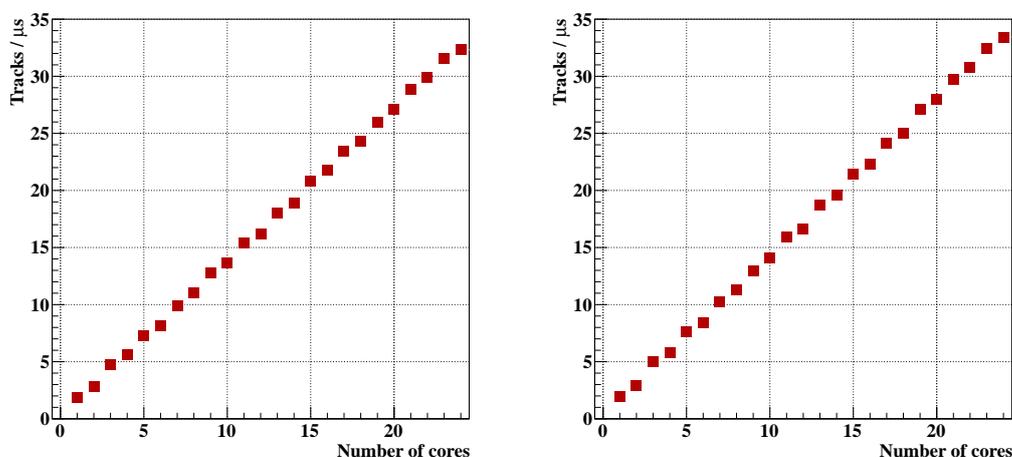


Рис. 2. Масштабируемость алгоритма реконструкции треков заряженных частиц в зависимости от числа запущенных в среде OpenMP логических ядер при использовании для векторизации кода заголовочных файлов (слева) и библиотеки Vc (справа)

Оба метода показали одинаковую производительность, позволив достичь скорости обработки 34 трека/мкс.

Использование OpenCL для CPU (рис. 3) также позволило добиться линейной масштабируемости. Результат немного скромнее (27 треков/мкс), а иной характер зависимости объясняется другим порядком подключения логических ядер нежели в случае с OpenMP.

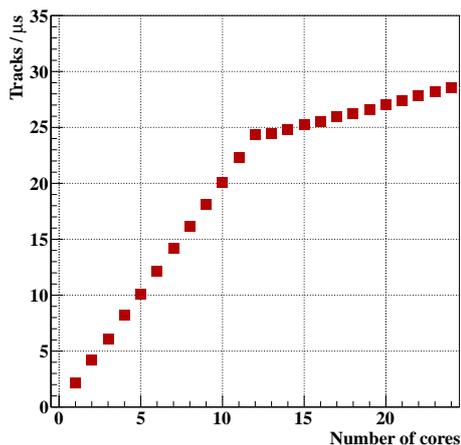


Рис. 3. Масштабируемость алгоритма реконструкции параметров треков заряженных частиц по отношению к числу логических ядер центрального процессора в среде OpenCL

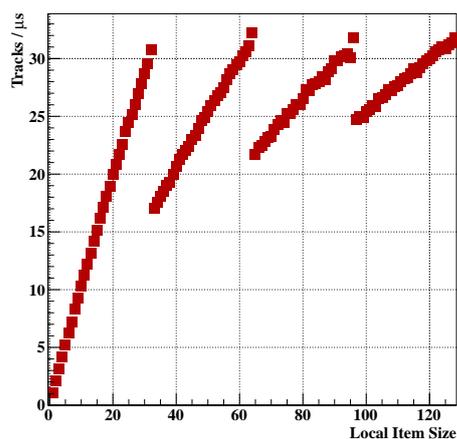


Рис. 4. Производительность графического процессора NVidia GTX 480 в зависимости от числа треков в рабочей группе

5. Производительность алгоритма на графических ускорителях и распределение задач по рабочим группам

Реализация алгоритма в среде OpenCL позволила запускать его также и на графических картах. Для тестирования использовалась NVidia GTX 480. Процессор этой карты содержит 448 ядер CUDA.

При запуске программы весь набор треков, подлежащих обработке, распределяется между рабочими группами. Каждая из них обрабатывается потоковым мультипроцессором (по 32 ядра в каждом).

На рис. 4 можно видеть, что максимальная скорость обработки достигается, если число треков в рабочей группе кратно числу ядер в потоковом мультипроцессоре и составляет 33 трека/мкс.

6. Заключение

В настоящей работе изучена и продемонстрирована возможность проведения быстрой реконструкции параметров траекторий заряженных частиц, регистрируемых системой координатных детекторов эксперимента CBM, на основе фильтра Калмана с использованием параллельных вычислений на многоядерном сервере ЛИТ ОИЯИ. К разрабатываемому алгоритму предъявлялись два основных требования: 1) высокая производительность алгоритма, 2) максимально возможная точность реконструкции нужных параметров трека (координата и направление) и частицы (импульс).

Первое требование было выполнено за счёт разработки параллельного алгоритма с использованием различных современных средств по распараллеливанию и векторизации кода, таких как заголовочные файлы, библиотека Vc, программные среды OpenMP и OpenCL. Указанные технологии позволяют запускать созданный программный код на высокопроизводительных многоядерных и гибридных системах, оснащённых векторными (SIMD) модулями и графическими ускорителями.

Решение по второму требованию было найдено в результате правильного выбора в качестве метода для реализации указанной задачи рекурсивного фильтра Калмана, а также за счёт применения ряда приближений, позволивших без потери точности вычислений обеспечить высокую надёжность и скорость обработки.

Разработанный алгоритм был протестирован на сервере cuda.jinr.ru ЛИТ ОИЯИ с двумя процессорами Intel Xeon X5660 и графической картой NVidia GTX 480. Все программные реализации показали линейную масштабируемость и высокую производительность (34 трека за 1 мкс на центральном процессоре и 33 трека за 1 мкс на графическом). Таким образом, уже в существующей комплектации сервер позволяет обрабатывать до 70-ти треков за 1 мкс.

Литература

1. The CBM Physics Book (1st Edition) / B. Friman, C. Höhne, S. Leupold et al. — New York: Springer, 2011. — Vol. 814 of *Lecture Notes in Physics*.
2. Vector Classes. — <http://gitorious.org/vc>.
3. OpenMP. — <http://openmp.org>.
4. OpenCL. — <http://www.khronos.org/opencl>.
5. *Kisel I.* Event Reconstruction in the CBM Experiment // *Nucl. Instr. and Meth.* — 2006. — No 1. — Pp. 85–88.
6. Performance Analysis of Cellular Automaton Algorithm to Solve the Track-Reconstruction Problem on a Multicore Server at the Laboratory of Information Technologies, Joint Institute of Nuclear Research / I. S. Kulakov, S. A. Baginyan, V. V. Ivanov, P. I. Kisel // *Physics of Particles and Nuclei Letters*. — 2013. — Vol. 10, No 2. — Pp. 162–170.

UDC 519.684.6 519.85 PACS 02.70.–c 07.05.-t

A Kalman Filter Method for the Charged Particles Trajectories Reconstruction in the CBM Experiment and Its Parallel Implementation at the JINR LIT Manycore Server

T. O. Ablyazimov*, M. V. Zyzak†, V. V. Ivanov*, P. I. Kisel*

* *Laboratory of Information Technologies
Joint Institute for Nuclear Research*

6, Joliot-Curie str., Dubna, Moscow region, Russia, 141980

† *Goethe-Universität Frankfurt am Main*

1, Grüneburgplatz, Frankfurt am Main, Russia, 60323

The high-intensity heavy-ion beams that will be produced by the accelerators at the Facility for Antiproton and Ion Research (FAIR), together with the Compressed Baryonic Matter (CBM) experiment, now in preparation, offer outstanding possibilities for studying baryonic matter at superhigh densities and moderate temperatures under laboratory conditions. The CBM physics program is aimed at studying the structure and behavior of baryonic matter at densities comparable to those in the center of neutron stars. The program includes 1) setting a phase boundary between the hadronic and partonic matter, 2) determining the critical end point, and 3) searching for indications of the origin of chiral symmetry reconstruction at high pure baryonic densities. The task of a charge particle trajectories reconstruction is one of the most important tasks of the CBM experiment. It assumes a full on-line event reconstruction, that requires development of fast algorithms, which utilize the potential of modern CPU and GPU architectures in the most efficient way. In the current work the results of analysis of the Kalman filter based track reconstruction algorithm, which is implemented using different parallelization approaches, are presented and discussed. For the analysis a manycore server with two Intel Xeon X5660 CPUs and a NVidia GTX 480 GPU at LIT, JINR was used.

Key words and phrases: CPU, GPU, heterogeneous system, Kalman filter, CBM, track fitting, OpenMP, OpenCL, NVIDIA, Intel.