

Алгоритмы DiamondTile для численного решения волновых уравнений на суперкомпьютерах с ускорителями на базе GPGPU

Андрей Закиров¹

Вадим Левченко²

¹ ООО «Кинтех Лаб», zakirovandrey@gmail.com

² Институт Прикладной Математики им.М.В.Келдыша РАН, vadimlevchenko@mail.ru

Численное моделирование различных оптических и устройств и материалов кроме комплексной модели среды и геометрии должно учитывать ограниченные размеры, дефекты, широкий спектр резонансных частот, а также высокий порядок точности. Наиболее подходящим в этом случае способом решения уравнений Максвелла является метод FDTD (Finite-Different Time-Domain) [1]. Множество задач при этом требуют больших вычислительных ресурсов.

В работе рассматривается трехмерное полноволновое моделирование с помощью метода FDTD распространения в среде электромагнитных или упругих волн. Метод FDTD обладает широкими возможностями в численном решении волновых уравнений. Мы рассматриваем в основном уравнения Максвелла и уравнения упругости. Следует отметить, что метод FDTD обладает высокой параллельностью и идеально подходит для реализации на GP-GPU платформах.

Существующие реализации метода FDTD (Finite-difference time domain), как коммерческих, так и с открытым исходным кодом, имеют крайне незначительную эффективность счета (относительно пиковой), что особенно важно при больших размерах счетной области. Это приводит к тому, что большое количество постановок ограничено сравнительно малыми размерами задач. Например, становится невозможным моделирование нано-оптических устройств в пространственно-временной области даже на массивно-параллельных многоядерных архитектурах. Принципиальным ограничением является послойная синхронизация на каждом шаге по времени всей счетной области, в результате чего темп вычислений ограничивается пропускной способностью памяти.

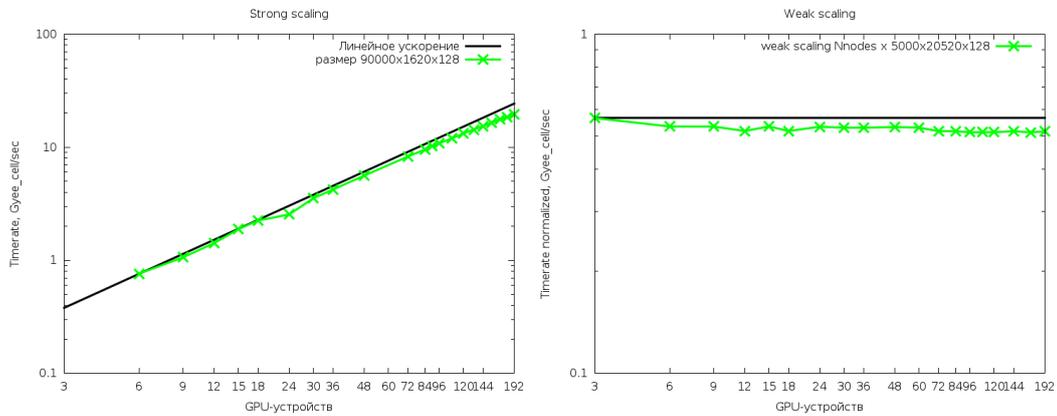
Для этих целей были разработаны алгоритмы DiamondTile (семейства LRnLA) [2] для архитектур с графическими процессорами общего назначения (GPGPU) с использованием технологии CUDA. Данный тип алгоритмов учитывает особенности иерархии памяти и параллельности GPGPU. Высокий уровень локализации данных, а также свойства асинхронности позволяют эффективно задействовать все уровни параллелизма компьютера. Вычислительная интенсивность (введенная в модели вычислений гоofline [3]) алгоритмов DiamondTile выше, чем у алгоритмов с пошаговой синхронизацией, и в результате становится возможным достичь эффективности $\sim 50\%$ относительно пиковой теоретической.

Алгоритмы LRnLA (Локально-рекурсивные нелокально-асинхронные) основываются на принципе «разделяй и властвуй» для рекурсивной декомпозиции задачи

на подзадачи на нескольких уровнях таким образом, чтобы данные каждой подзадачи располагались на более высоких уровнях иерархии памяти вычислителя. Это позволяет достичь пиковой производительности для задач с размерами данных соответствующих нижним уровням подсистемам памяти.

Использование суперкомпьютеров с тысячами современных GPU позволит провести 3D моделирование из первых принципов распространения света в области до 1000000 мкм^3 , то есть моделирование целиком некоторых оптических устройств, таких как оптические интегральные схемы или органические светодиоды. Такое моделирование может быть полезным при проектировании мобильных устройств и наноинженерии (нано-антенн, сенсоров, резонаторов).

Текущая реализация LRnLA на GPGPU имеет производительность до $3 \cdot 10^9$ ячеек ИИ в секунду на одном чипе GPU. Также проведенные расчеты показывают практически идеальное ускорение на системах до 200 GPU-устройств (вычислительный комплекс К-100).



Список литературы

- [1] *A. Taflov, S. C. Hagness.* Computational Electrodynamics: the Finite-Difference Time-Domain Method. — 3rd edition. — Norwood, MA: Artech House, 2005
- [2] *Perepelkina A.Yu., Levchenko V. D.* DiamondTorre Algorithm for High-Performance Wave Modeling // Keldysh Institute Preprints. 2015. No. 18.
- [3] *S. Williams, A. Waterman, and D. Patterson,* "Roofline: An insightful visual performance model for multicore architectures Commun. ACM , vol. 52, no. 4, pp. 65–76, Apr. 2009.