



Учреждение Российской Академии Наук
Институт Прикладной Физики РАН



Нижегородский государственный университет
им. Н.И. Лобачевского
Национальный исследовательский университет



ДИНАМИЧЕСКАЯ БАЛАНСИРОВКА В КОДЕ PICADOR ДЛЯ МОДЕЛИРОВАНИЯ ПЛАЗМЫ

Бастраков С., Мееров И., Сурмин И.



Абрау-Дюрсо
25.09.2013

Содержание

- ❑ Постановка задачи
- ❑ Метод частиц в ячейках
- ❑ Параллельная вычислительная схема
- ❑ Балансировка нагрузки
- ❑ Результаты производительности



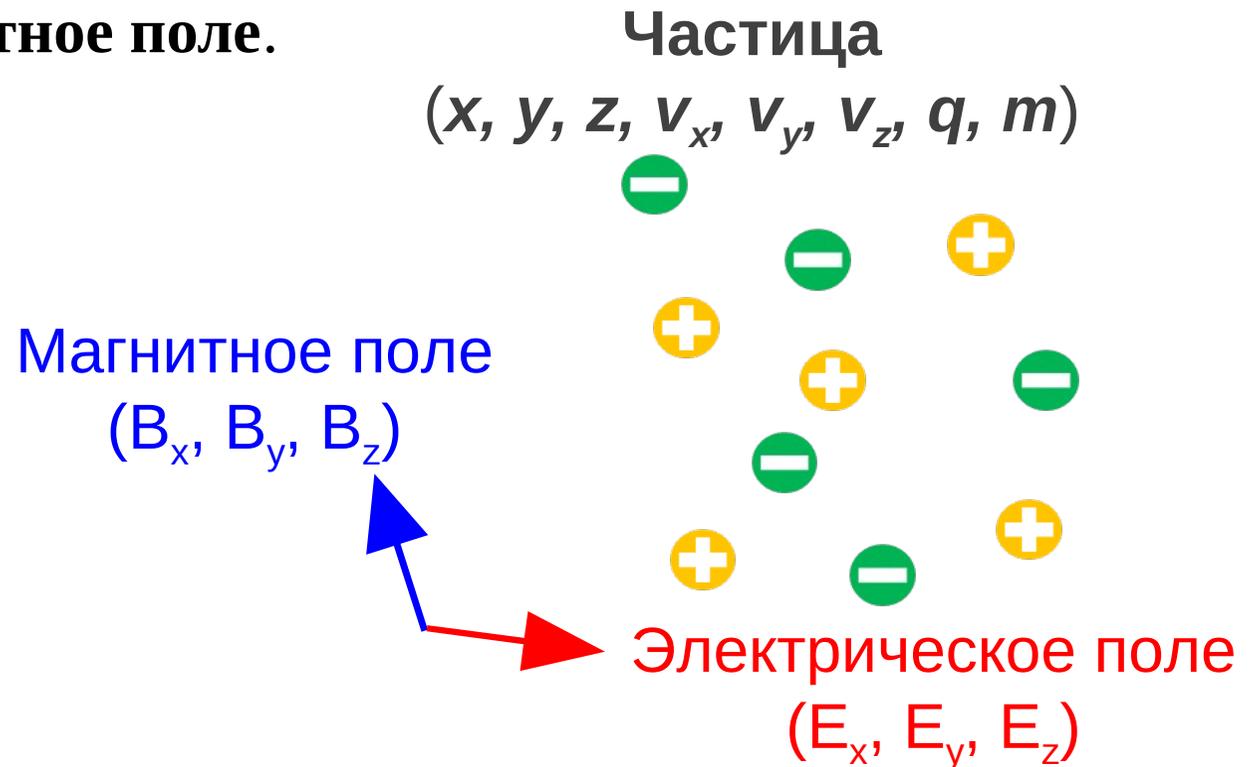
Постановка задачи

- ❑ Предметная область – *численное моделирование плазмы.*
- ❑ Примеры актуальных прикладных задач:
 - адронная терапия;
 - производство короткоживущих изотопов;
 - протонография;
 - генерация гамма-излучения;
- ❑ Один из основных методов моделирования – **метод частиц в ячейках** (*Particle-in-Cell, PIC*).
- ❑ Для решения многих прикладных задач в 3D требуется $\sim 10^8$ ячеек пространственной сетки, $\sim 10^9$ частиц, обычно используются $\sim 10^2$ – 10^4 вычислительных ядер.



Объект исследования

- Плазма – полностью или частично ионизованный газ.
- Две основные сущности:
 - ансамбль заряженных частиц (электроны, ионы);
 - электромагнитное поле.



Математическая модель

Уравнения движения частиц

$$\frac{\partial \mathbf{p}_i}{\partial t} = q_i \left(\mathbf{E}(\mathbf{r}_i) + \frac{1}{c} \mathbf{v}_i \times \mathbf{B}(\mathbf{r}_i) \right)$$

$$\frac{\partial \mathbf{r}_i}{\partial t} = \frac{\mathbf{p}_i}{m_i} \left(1 + \left(\frac{\mathbf{p}_i}{m_i c} \right)^2 \right)^{-1/2} = \mathbf{v}_i$$

$\mathbf{r}_i, \mathbf{v}_i$



Плотность тока и заряда

$$\mathbf{j}(\mathbf{r}) = \sum_i q_i \mathbf{v}_i \delta(\mathbf{r} - \mathbf{r}_i)$$

$$\rho(\mathbf{r}) = \sum_i q_i \delta(\mathbf{r} - \mathbf{r}_i)$$

\mathbf{E}, \mathbf{B}



Уравнения Максвелла

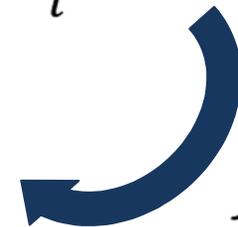
$$\nabla \times \mathbf{B} = \frac{4\pi}{c} \mathbf{j} + \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t}$$

$$\nabla \cdot \mathbf{B} = 0$$

$$\nabla \times \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}$$

$$\nabla \cdot \mathbf{E} = 4\pi \rho$$

\mathbf{j}, ρ



Метод частиц в ячейках

- Два разнородных набора данных:
 - набор частиц с положениями, скоростями, массами и зарядами;
 - электромагнитное поле и токи на пространственной сетке.
- Моделирование динамики во времени производится с некоторым фиксированным шагом.
- На каждой итерации по времени:
 - интерполяция сеточных значений электромагнитного поля, интегрирование уравнений движения;
 - вычисление плотности тока и заряда с помощью аппроксимации на сетке («взвешивание токов»);
 - интегрирование уравнений Максвелла.



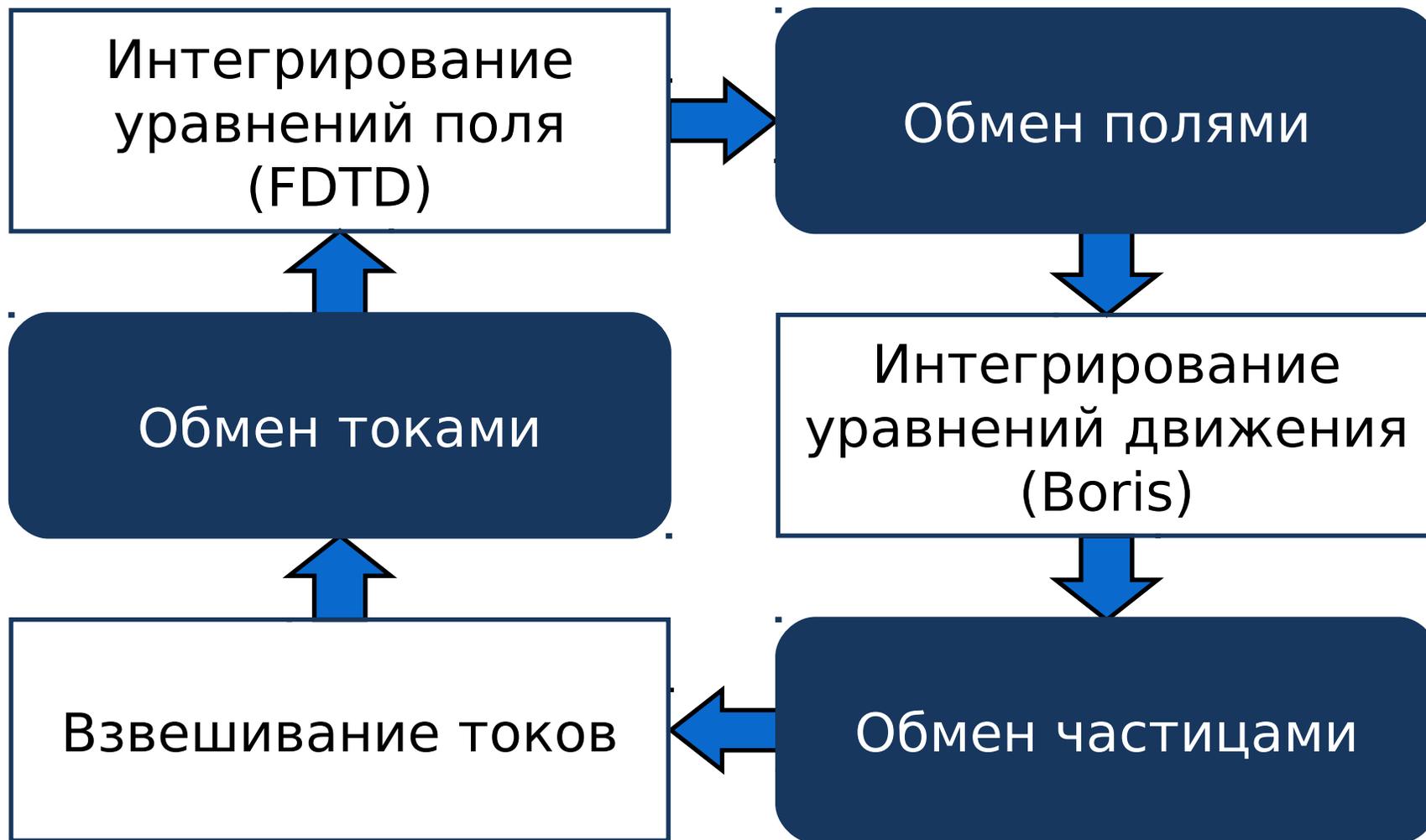
Параллельная вычислительная схема

Декомпозиция по территориальному принципу.

- ❑ Разбиение расчетной области на подобласти (*домены*), операции над которыми выполняются параллельно.
- ❑ Каждый домен содержит сеточные данные и набор частиц, попадающих в соответствующую подобласть.
- ❑ Перекрытие по данным с соседними доменами.
- ❑ Поддержание актуальности данных за счет обменов данными с соседними доменами на каждой итерации.



Вычислительный цикл



Программная реализация

- ❑ **PICADOR** – инструмент для моделирования плазмы на гетерогенных кластерных системах с несколькими CPU и GPU на узле.
- ❑ Реализация для кластера с использованием MPI. Каждый процесс использует для вычислений либо CPU (одно или несколько ядер), либо один GPU.
- ❑ 2 реализации вычислительного модуля (солвера) с общим интерфейсом:
 - для CPU на C с использованием OpenMP;
 - для GPU на C с использованием OpenCL.
- ❑ Равномерное статическое распределение работы между узлами кластера



Недостатки равномерного разбиения

Неэффективное использование вычислительных ресурсов:

- ❑ Вычислительная нагрузка распределена неравномерно между процессами,
- ❑ Некоторым процессам требуется большое количество памяти.

Балансировка нагрузки:

- ❑ Использование разбиения на домены различных размеров,
- ❑ Переразбиение расчетной области в процессе вычислений.



Статическое разбиение

- Прямолинейное разбиение на $m \times n \times l$ доменов:

$$P = (p_0, p_1 \dots, p_m), \quad Q = (q_0, q_1 \dots, q_m), \quad R = (r_0, r_1 \dots, r_l)$$

$$Domain(I, J, K) = \{ cells(i, j, k) \mid p_I, q_J, r_K \leq (i, j, k) < (p_{I+1}, q_{J+1}, r_{K+1}) \}$$

- Оценка вычислительной нагрузки:

$$DomainWorkload = nParticles + nCells$$

- Задача оптимизации:

$$\arg \min_{(P, Q, R)} \max_{i, j, k} DomainWorkload(i, j, k)$$

* D. N. Nicol. Rectilinear partitioning of irregular data parallel computations.
Journal of Parallel and Distributed Computing, 23: 119-134, 1994.

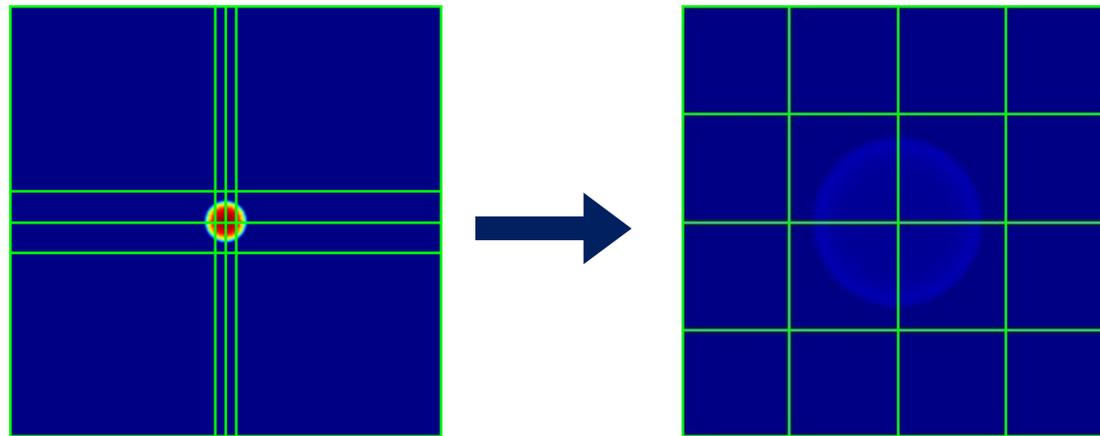


Динамическое разбиение

- ❑ Оценка качества разбиения – дисбаланс:

$$Imbalance = \frac{MaxDomainWorkload}{AverageDomainWorkload}$$

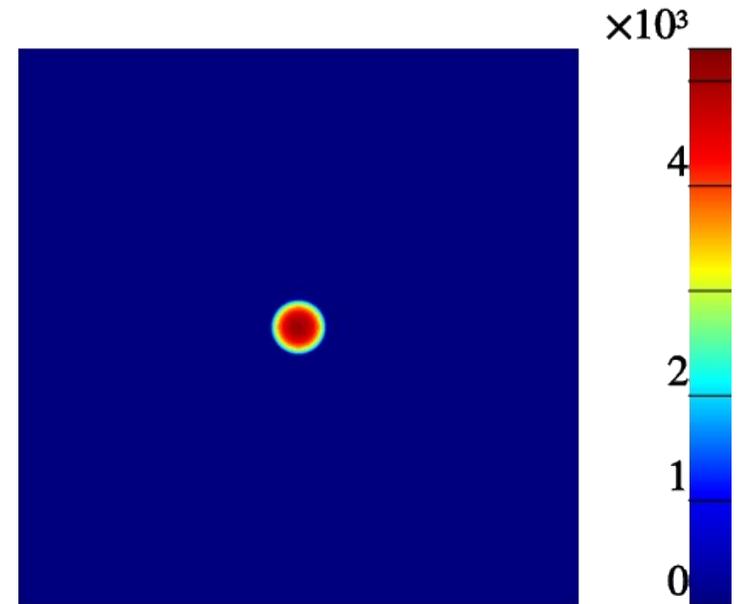
- ❑ Изменение распределения частиц со временем приводит к необходимости переразбиения (динамическая декомпозиция)



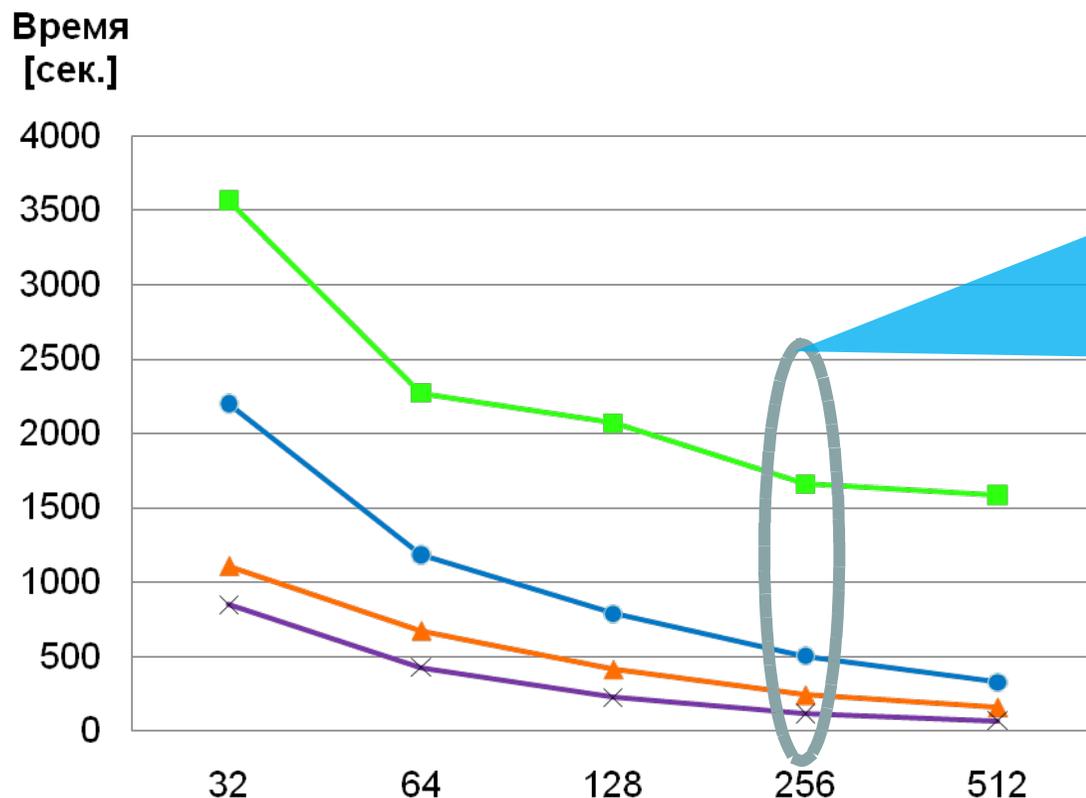
- ❑ Критерий переразбиения – снижение дисбаланса не менее, чем в заданное количество раз.

Эксперимент

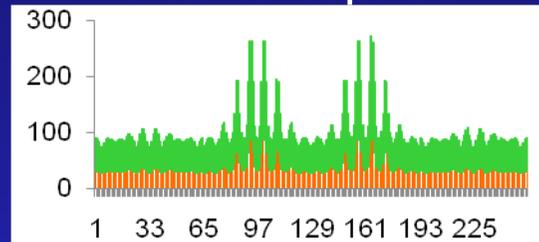
- ❑ Нагретые частицы равномерно распределены в шаре в центре расчетной области.
Сетка: 128x128x128, 40 млн. частиц, 1000 итераций по времени
- ❑ CPU, отдельный MPI-процесс на каждое ядро, OpenMP отключено.
- ❑ Вычислительная система:
 - «МВС-100К» в МСЦ РАН: на каждом узле 2 процессора Intel Xeon E5450 (3 ГГц), 8 ГБ RAM, Infiniband DDR, CentOS 5.6 x86_64.



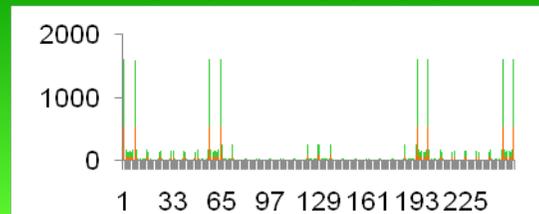
Анализ производительности



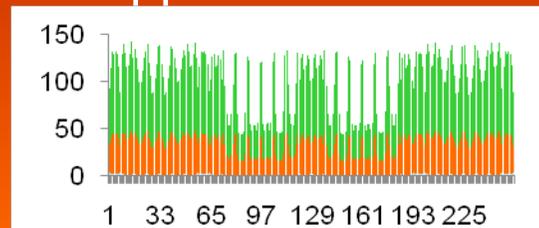
Равномерное



Статическое



Динамическое



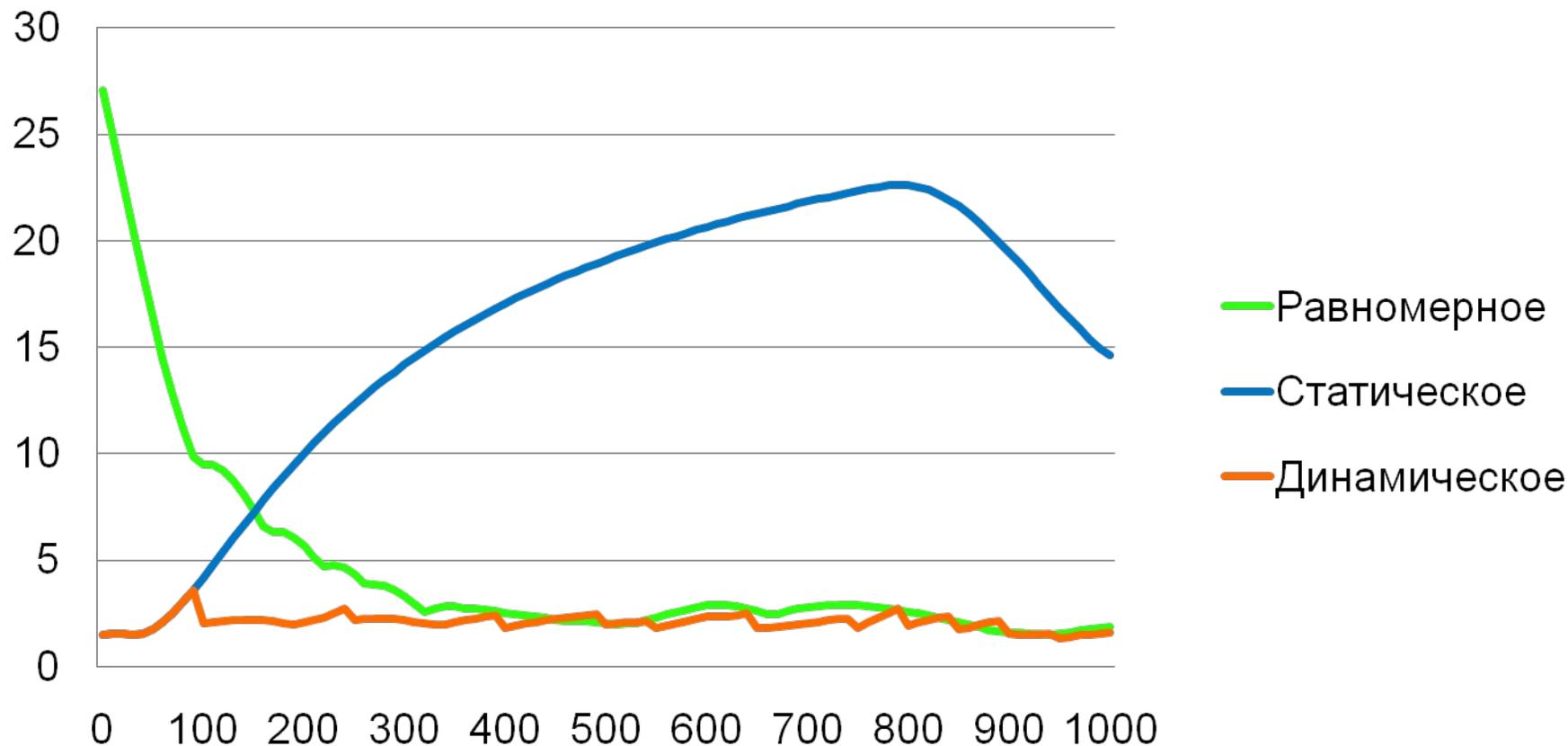
- Равномерное
- Статическое
- ▲ Динамическое
- × Идеальный баланс

- Интегрирование частиц
- Взвешивание токов
- Интегрирование полей



Изменение дисбаланса

Дисбаланс



Результаты

- ❑ Реализованы схемы статического и динамического прямолинейного разбиения для балансировки нагрузки.
- ❑ Данные схемы интегрированы в программный комплекс PICADOR
- ❑ Проведены эксперименты по анализу производительности
 - Динамическая схема обгоняет равномерную примерно в 2 раза.
 - Динамическая схема отстает от идеального случая в 1.5 раза.



Планы дальнейшего развития

- ❑ Автоматический подбор параметров балансировки.
- ❑ Реализация балансировки нагрузки между различными вычислительными устройствами (CPU, GPU, Intel Xeon Phi).



Вопросы

- ❑ Спасибо за внимание
- ❑ **email:** *i.surmin@gmail.com*

