Общие проблемы применения многопроцессорных систем

М.В.Якобовский

mail: lira@imamod.ru
web: http://lira.imamod.ru

2012

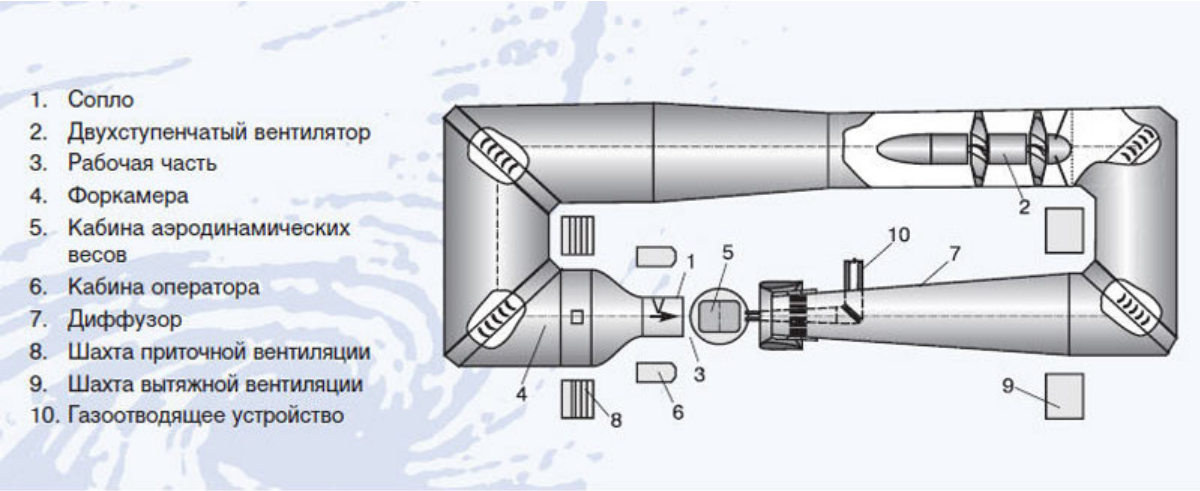
Особенности момента

- Потребность в суперкомпьютерах высока
- Эффективность использования суперкомпьютеров низка:
  - Использование каждого ядра последовательной программой составляет проценты и доли процентов
  - Обмены, синхронизация и другие дополнительные операции ещё снижают эффективность параллельной программы
- Есть минимальный объем вычислений на процессорное ядро, определяющий ***число используемых ядер***
- За счет ***многопроцессорности*** сложно ***сокращать*** *время* моделирования физического процесса, но можно повышать ***сложность*** решаемых задач, например за счет увеличения размеров изучаемых объектов

Дозвуковая аэродинамическая труба Т-104, ЦАГИ

- Скорость потока 10–120 м/с
- Диаметр сопла 7 м
- Длина рабочей части 13 м
- Мощность вентилятора **28.4 МВт**

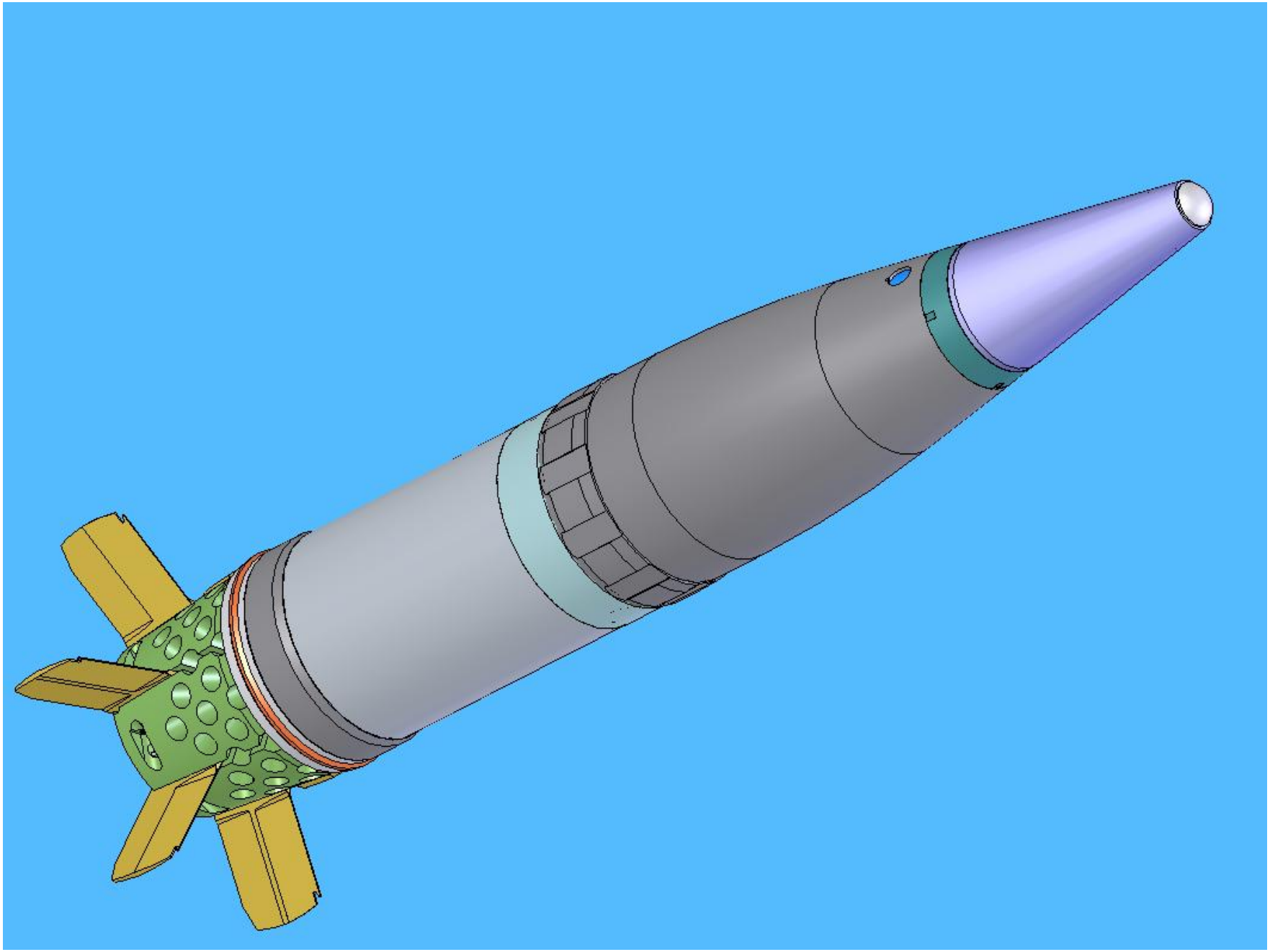
http://www.tsagi.ru/rus/base/t104

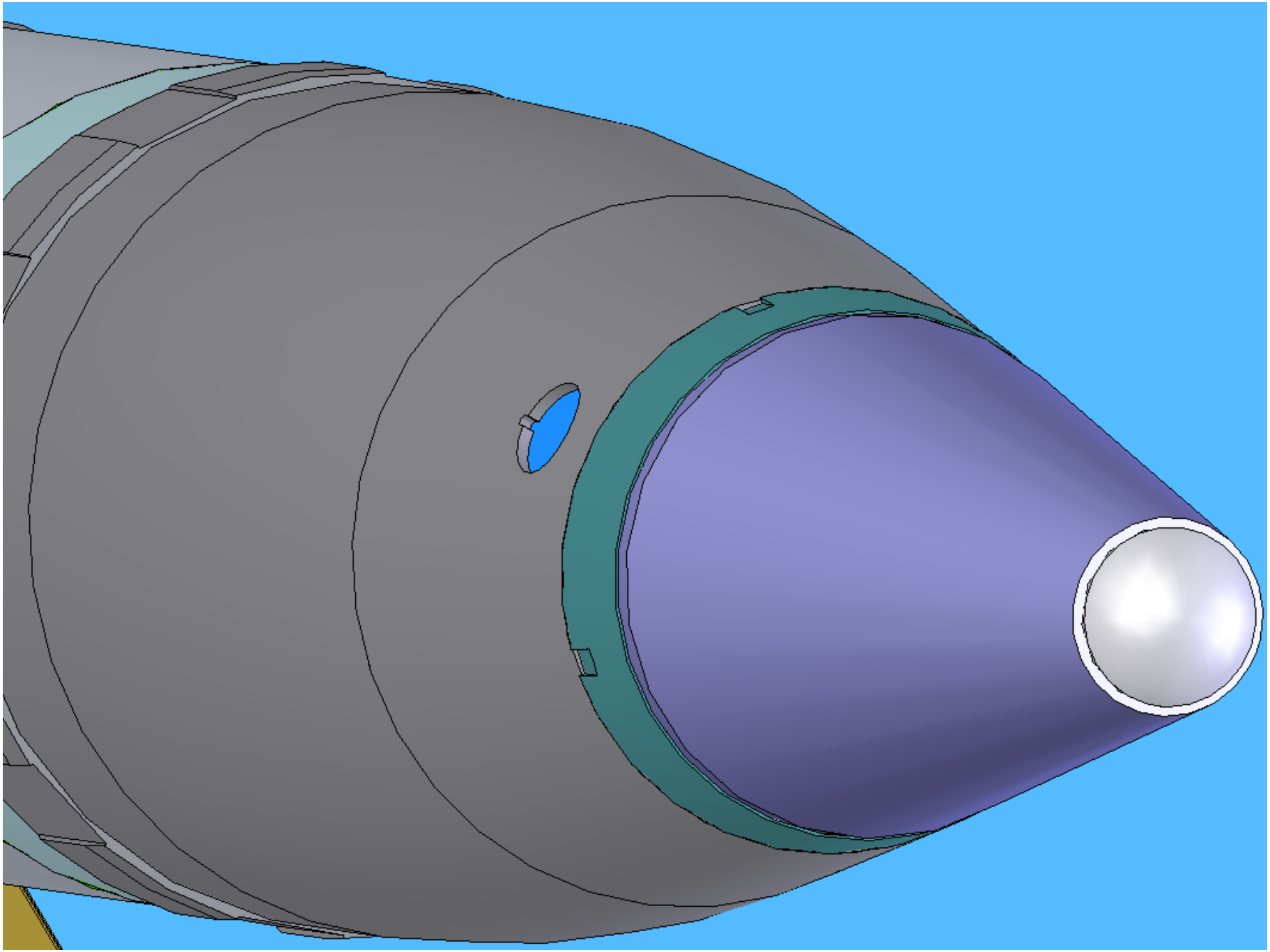


Суперкомпьютер СКИФ МГУ «ЧЕБЫШЁВ»

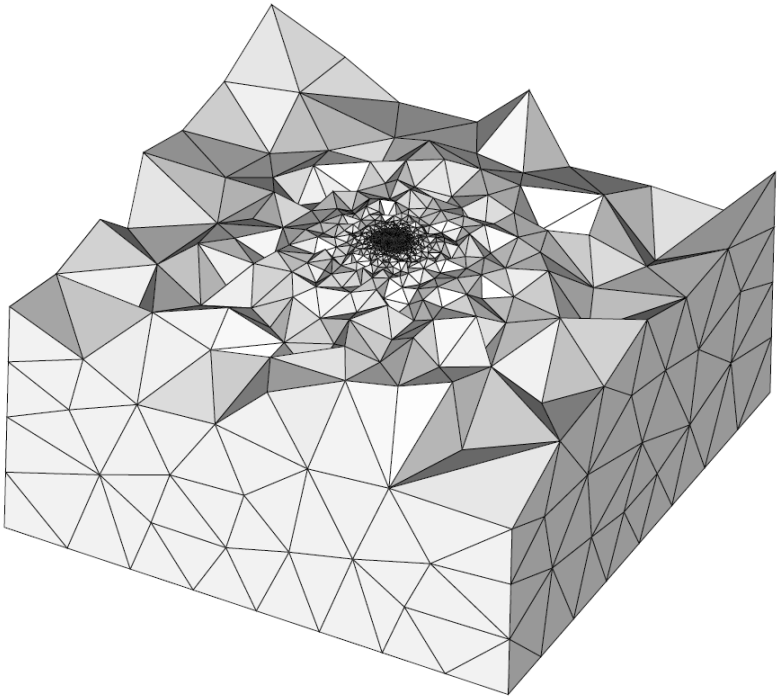
- Пиковая производительность 60 TFlop/s
- Мощность комплекса **0.72 МВт**

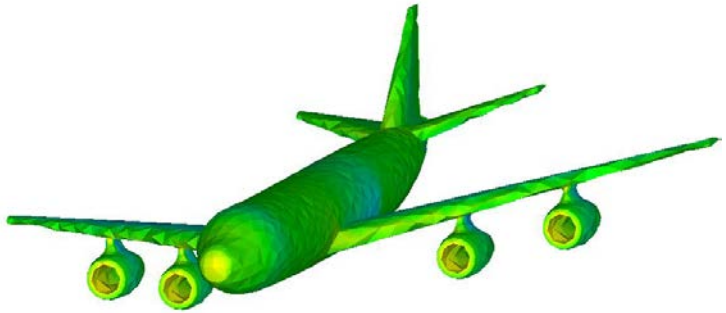
http://parallel.ru/cluster/skif_msu.html

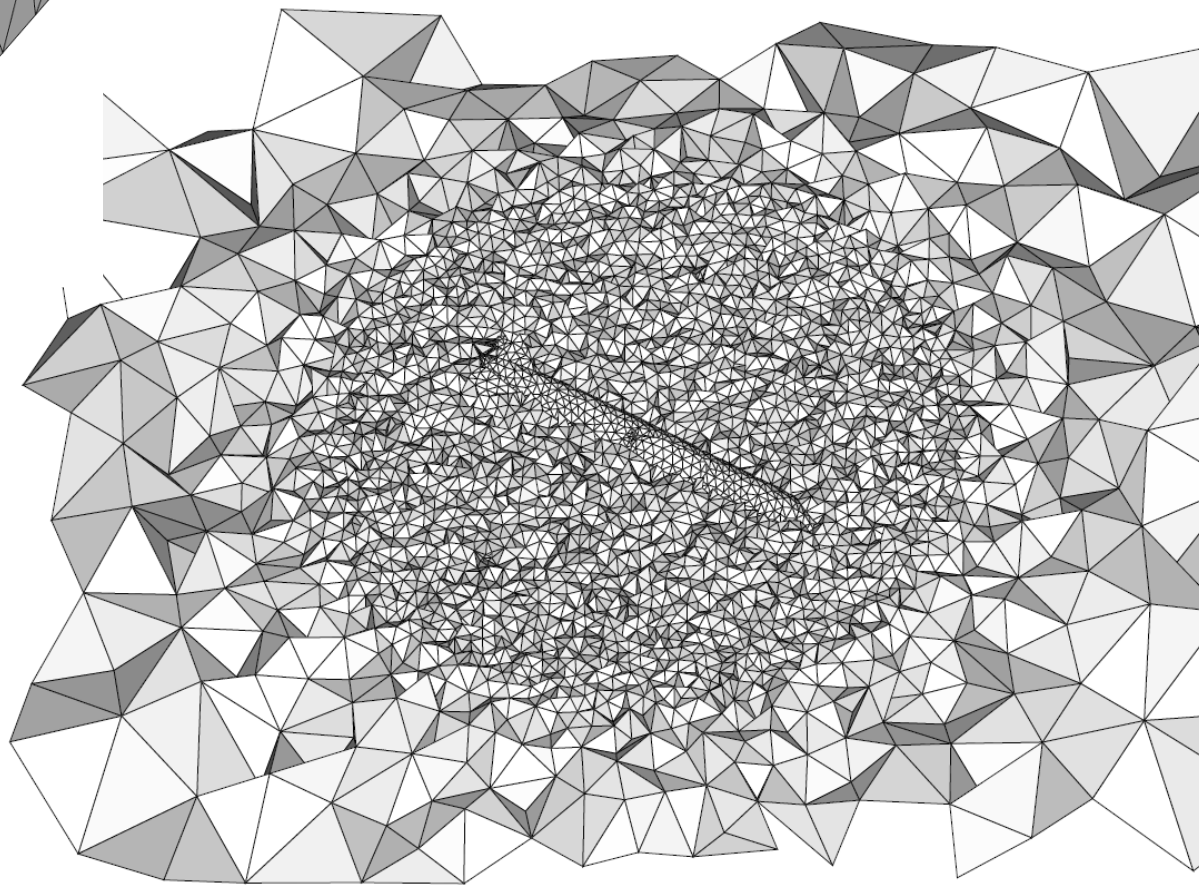




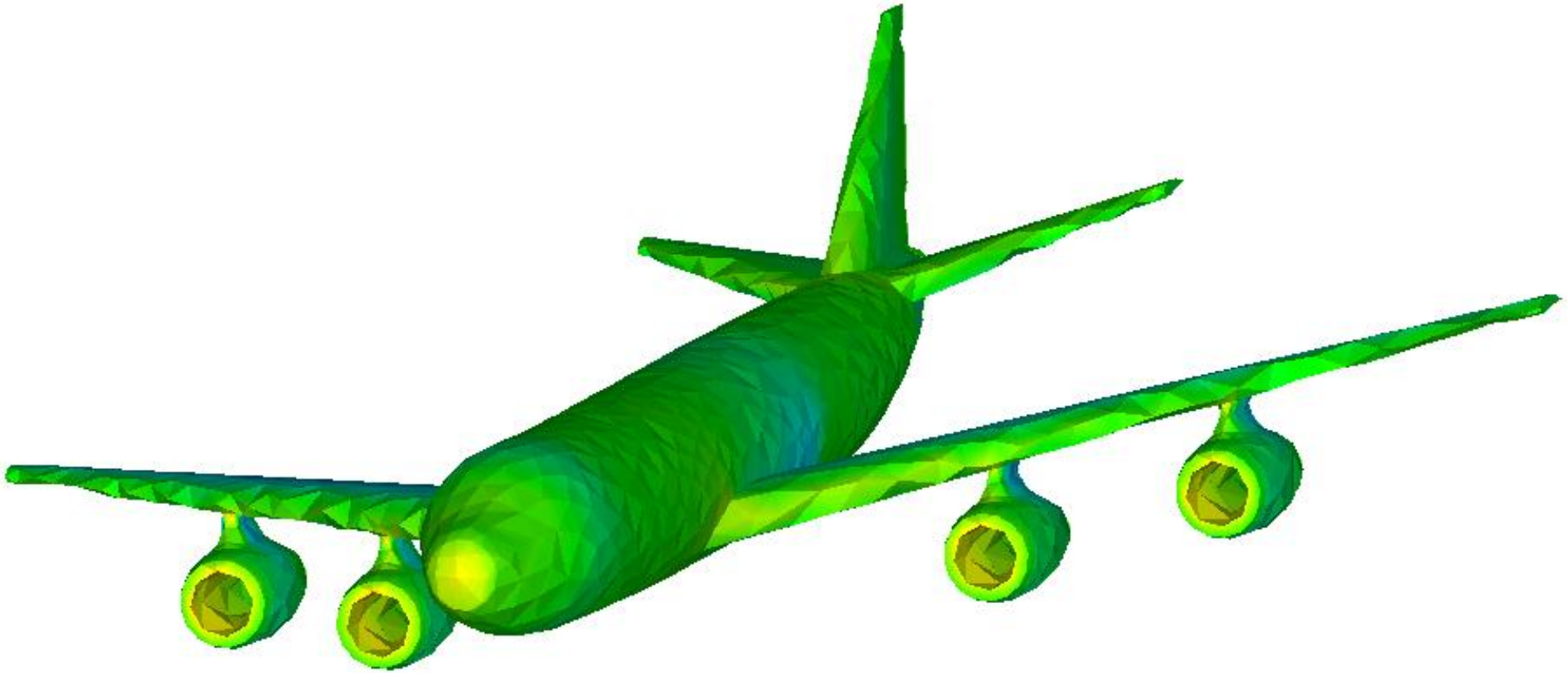
Большие сетки





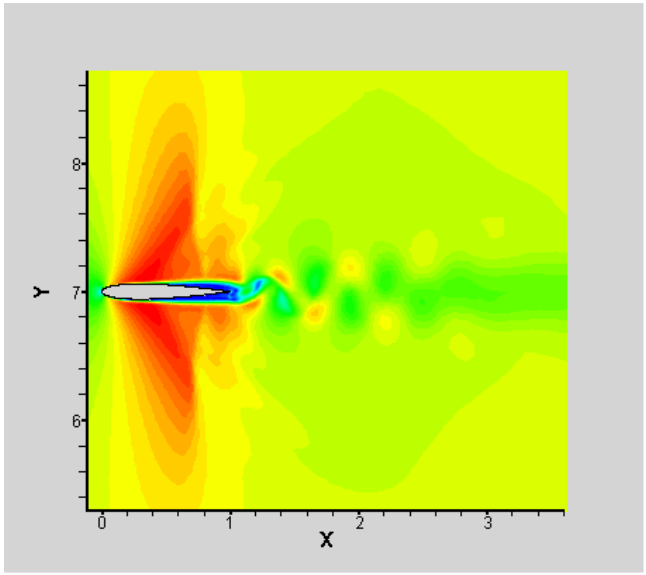


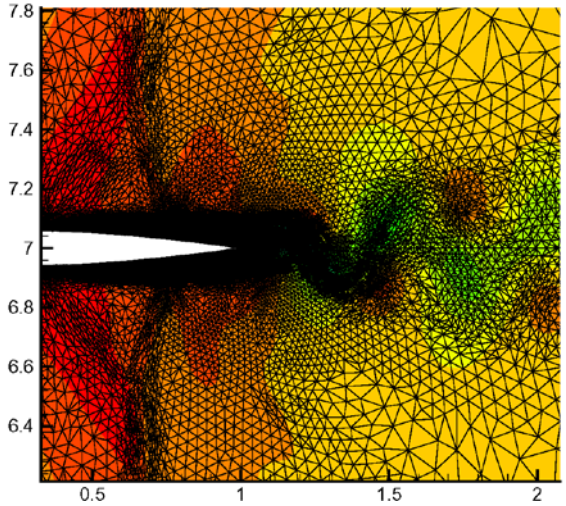
Тетраэдральные сетки 10^8 узлов



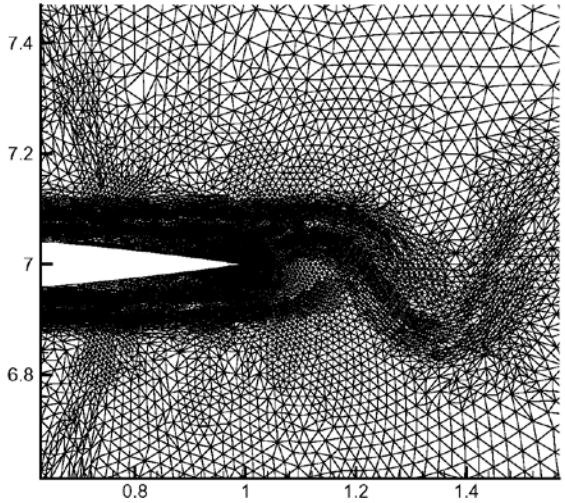
Использование адаптивной сетки

Обтекание профиля NACA0012
(M=0.85, Re=10^4)
под нулевым углом атаки:





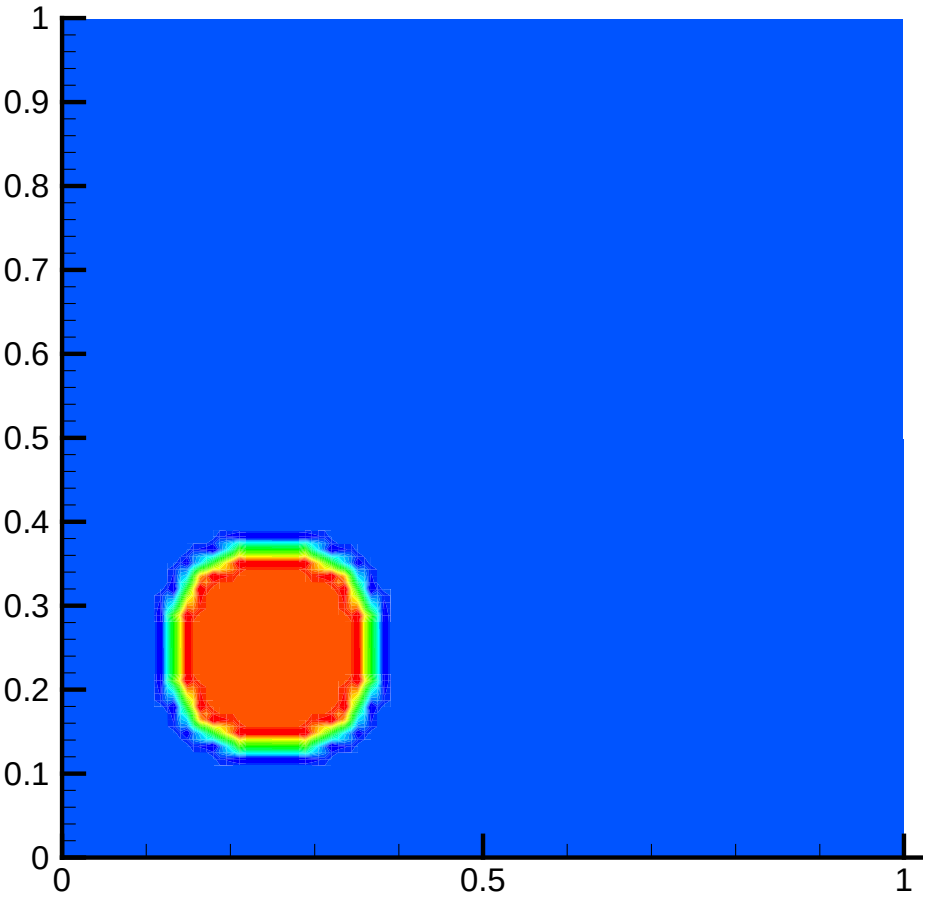
Поле продольной скорости

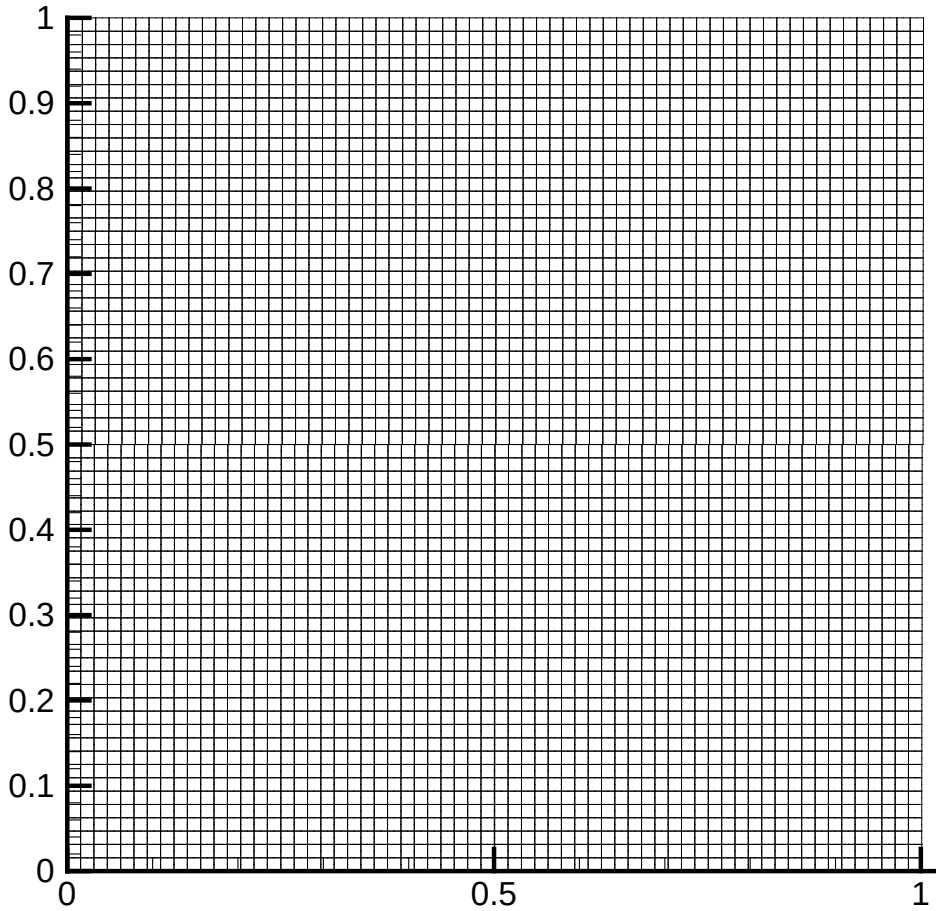


Фрагмент сетки

Равномерная сетка

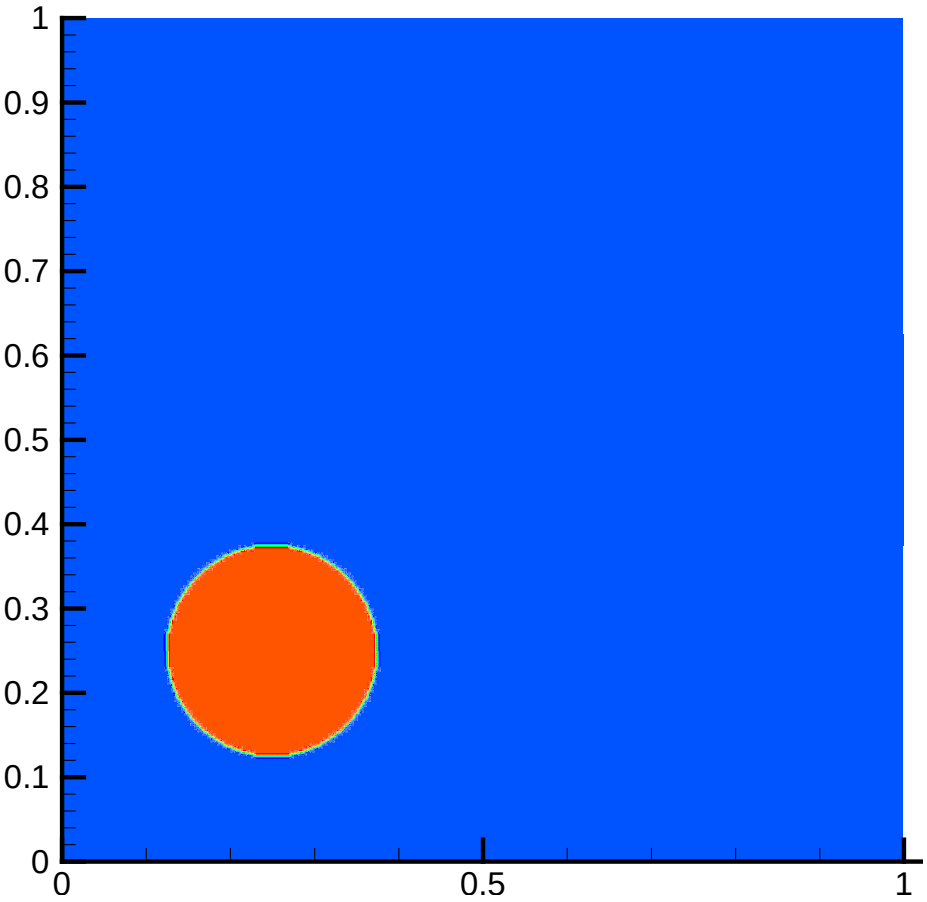
Слева – ??*круглое*?? пятно примеси

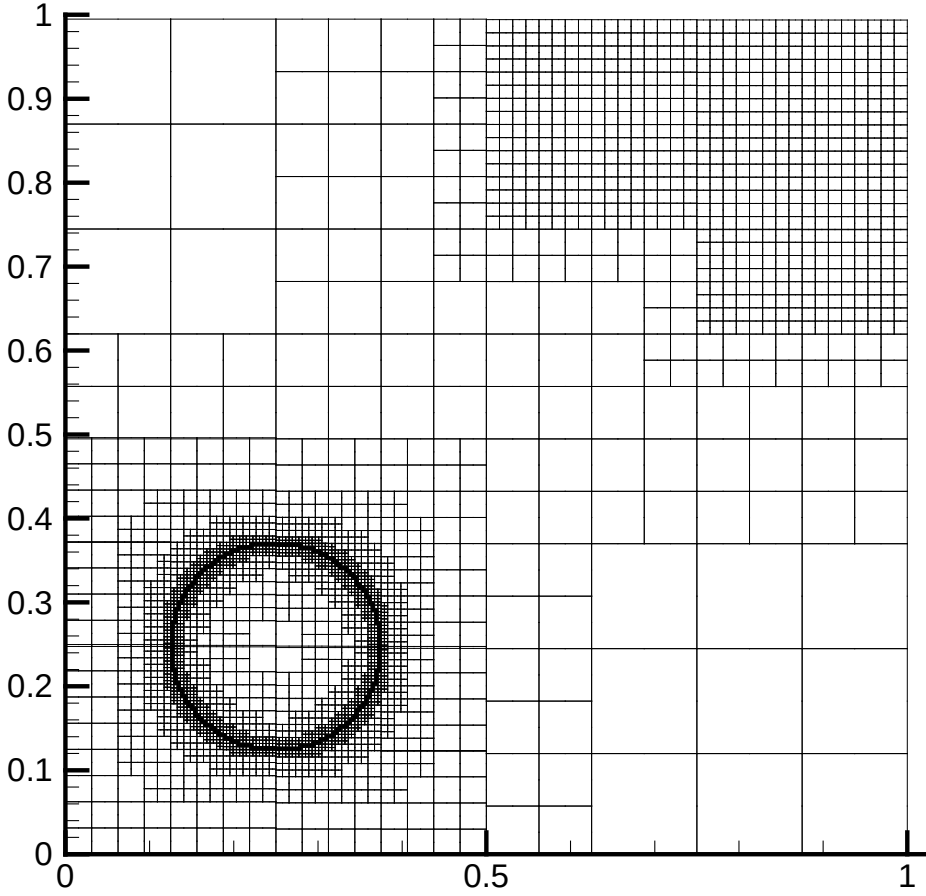




Адаптивная сетка

Слева – ***круглое*** пятно примеси

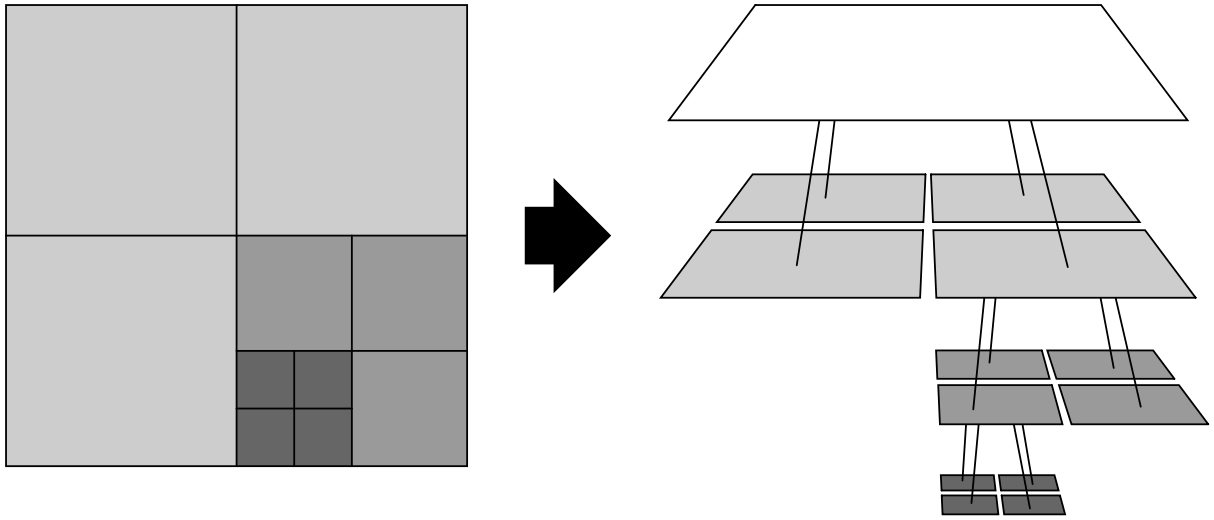




Адаптивные декартовы сетки

- Вначале сетка состоит из одной прямоугольной ячейки
- Каждая ячейка может быть разделена на четыре ячейки одинакового размера
- Если ячейки когда-то составляли одну ячейку, то они могут быть объединены обратно
- Каждая ячейка хранит величину, описывающую среднее значение неизвестной функции в пределах ячейки (метод конечных объёмов)

При данных предположениях сетку удобно хранить в виде четверичного дерева:

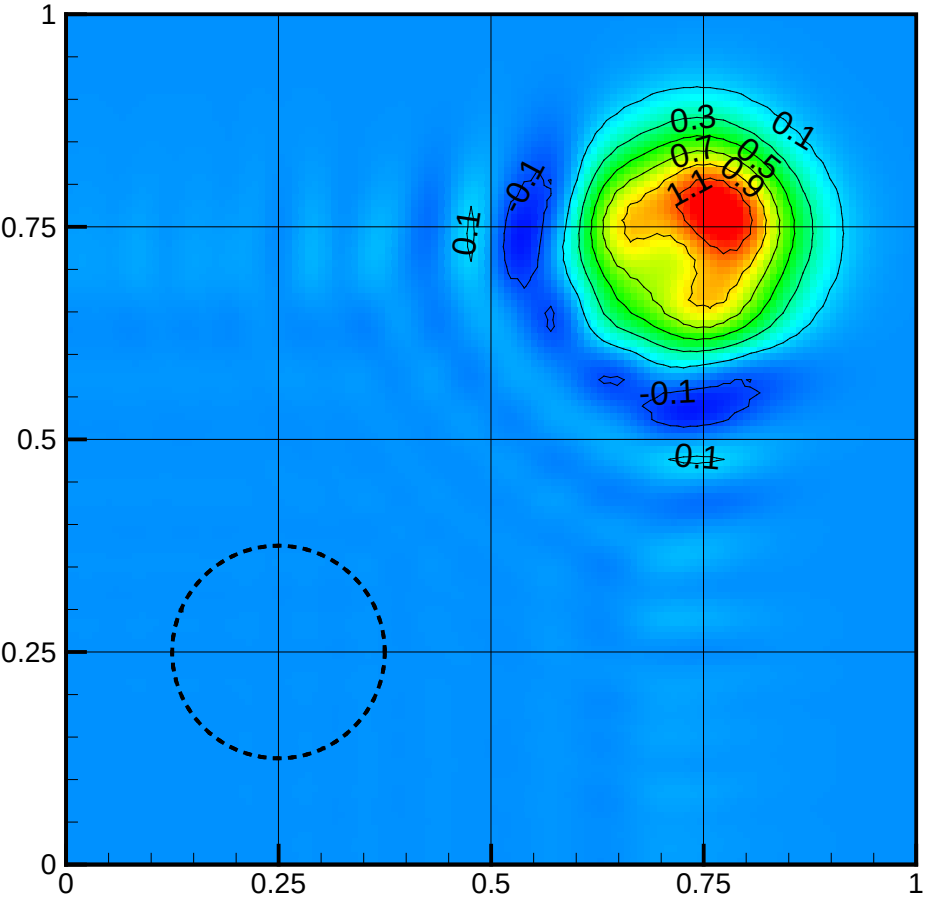


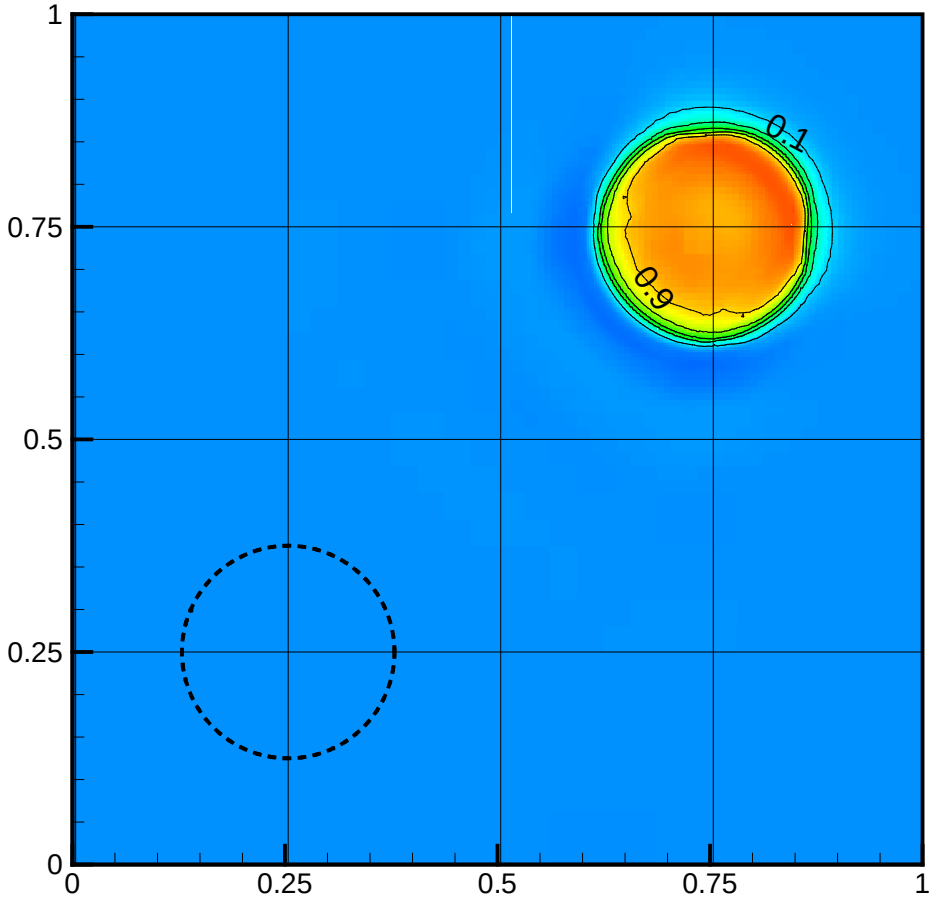
Дополнительные ограничения на размеры ячеек:

- Задан максимально допустимый размер ячеек
- Задан минимально допустимый размер ячеек
- Размеры соседних ячеек должны различаться не более, чем в 2 раза

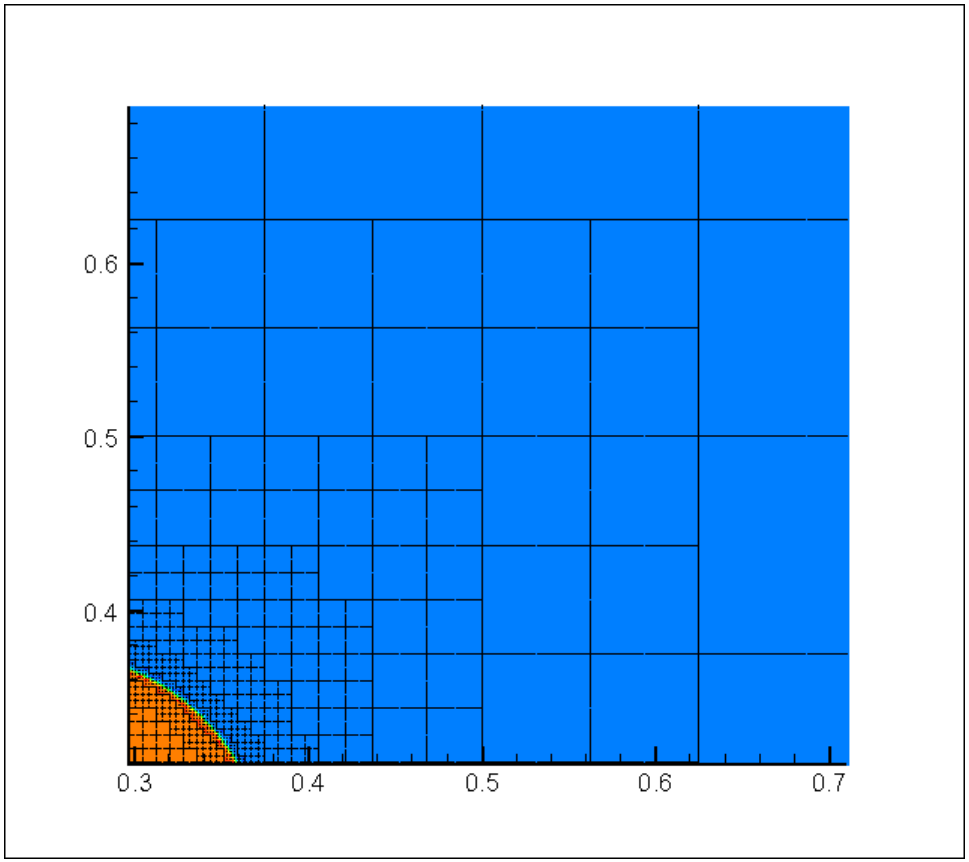
Сравнение с равномерной сеткой

На рисунках показаны результаты решения простейшей задачи переноса на равномерной (слева) и адаптивной (справа) сетках с одинаковым числом ячеек (4096 штук). Скорость переноса направлена под углом 45° к линиям сетки; начальное условие показано пунктиром

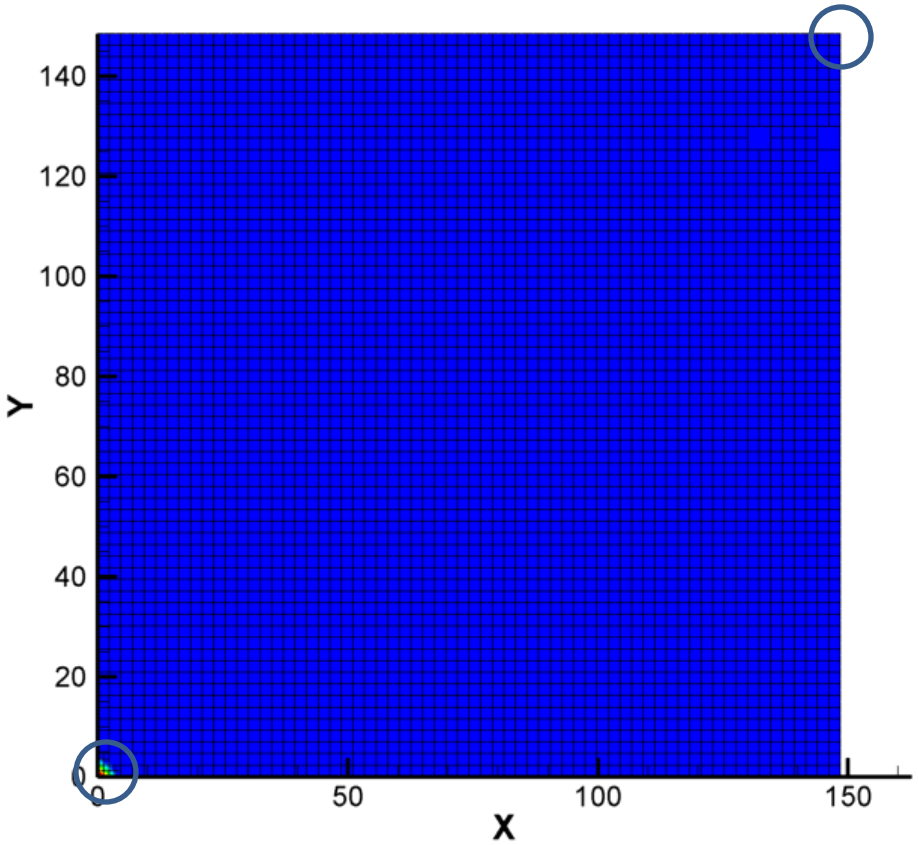




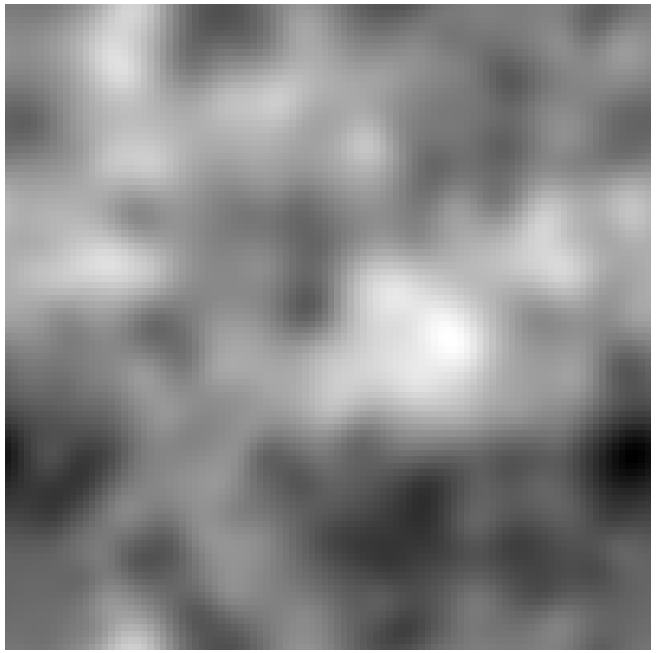
Адаптивная сетка



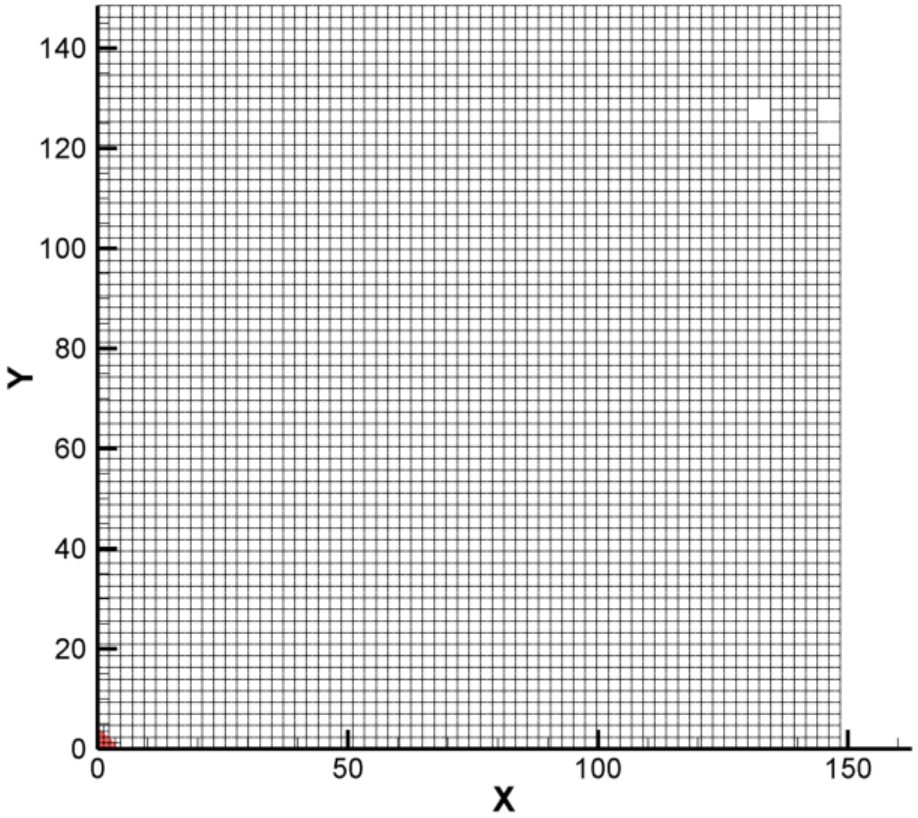
Решение двумерной задачи фильтрации нефтеводяной смеси в области с неоднородной проницаемостью



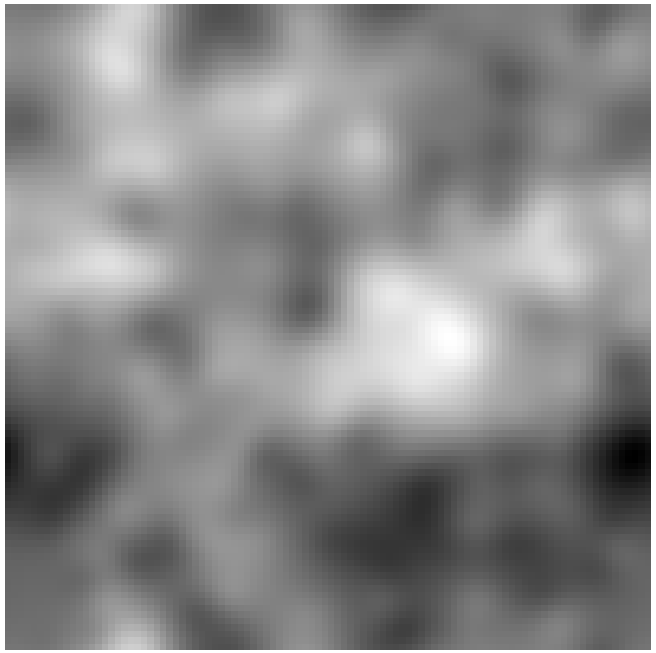
В юго-западном углу находится скважина, нагнетающая воду, в северо-восточном углу — добывающая скважина.
5-ти точечная схема
Поле проницаемости с разбросом значений на 4 порядка).



Решение двумерной задачи фильтрации нефтеводяной смеси в области с неоднородной проницаемостью



В юго-западном углу находится скважина, нагнетающая воду, в северо-восточном углу — добывающая скважина.
5-ти точечная схема
Поле проницаемости с разбросом значений на 4 порядка).

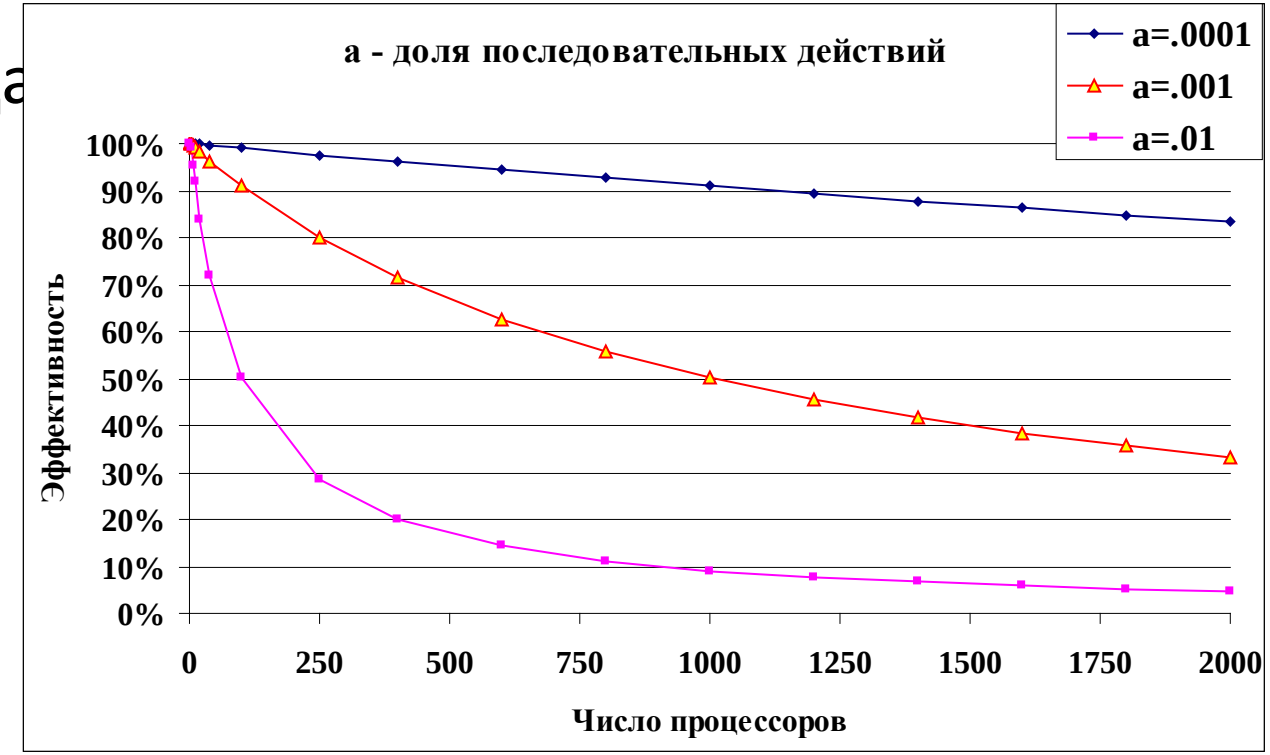


Ограничения

- Закон Амда

$$S(p) = \frac{1}{a + \frac{1-a}{p}}$$

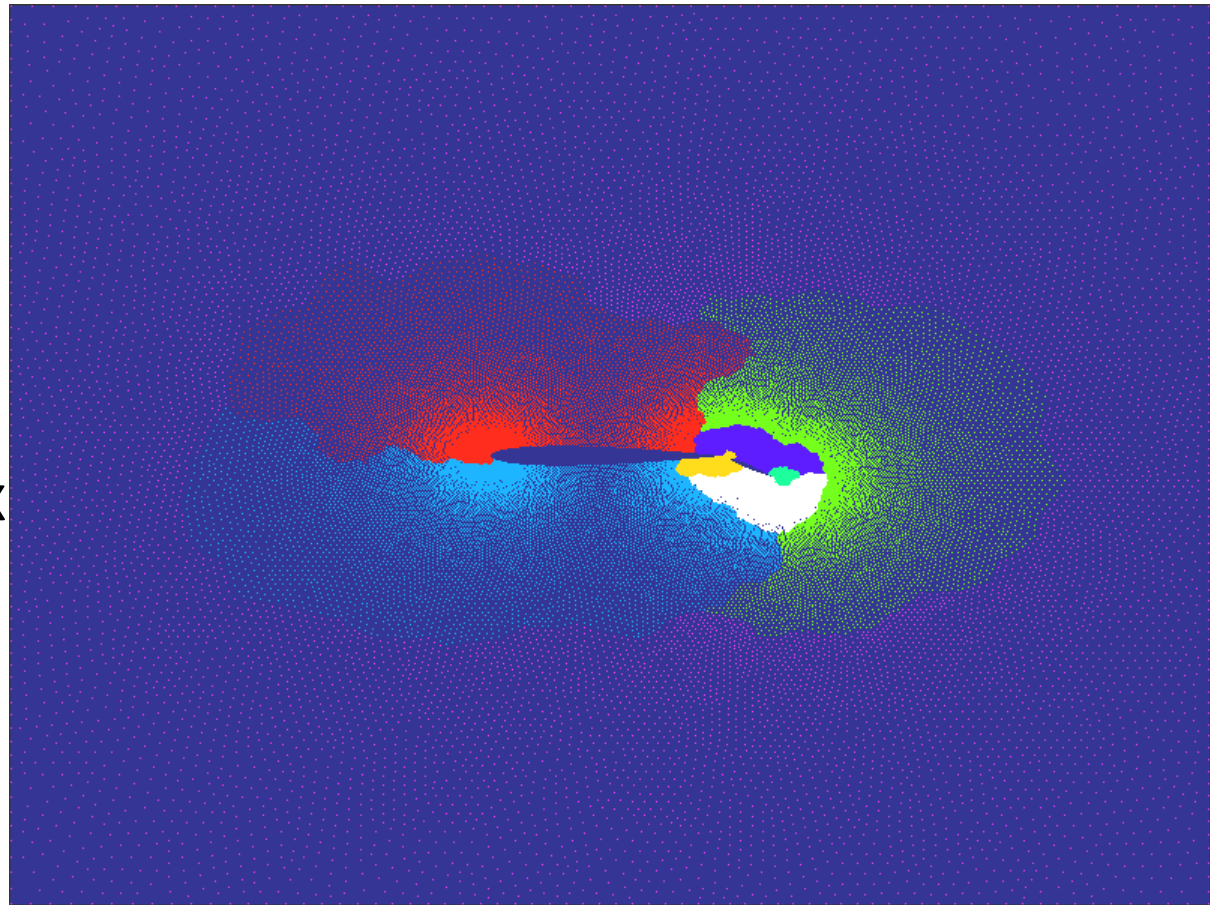
$$E(p) = \frac{1}{1 + a(p-1)}$$

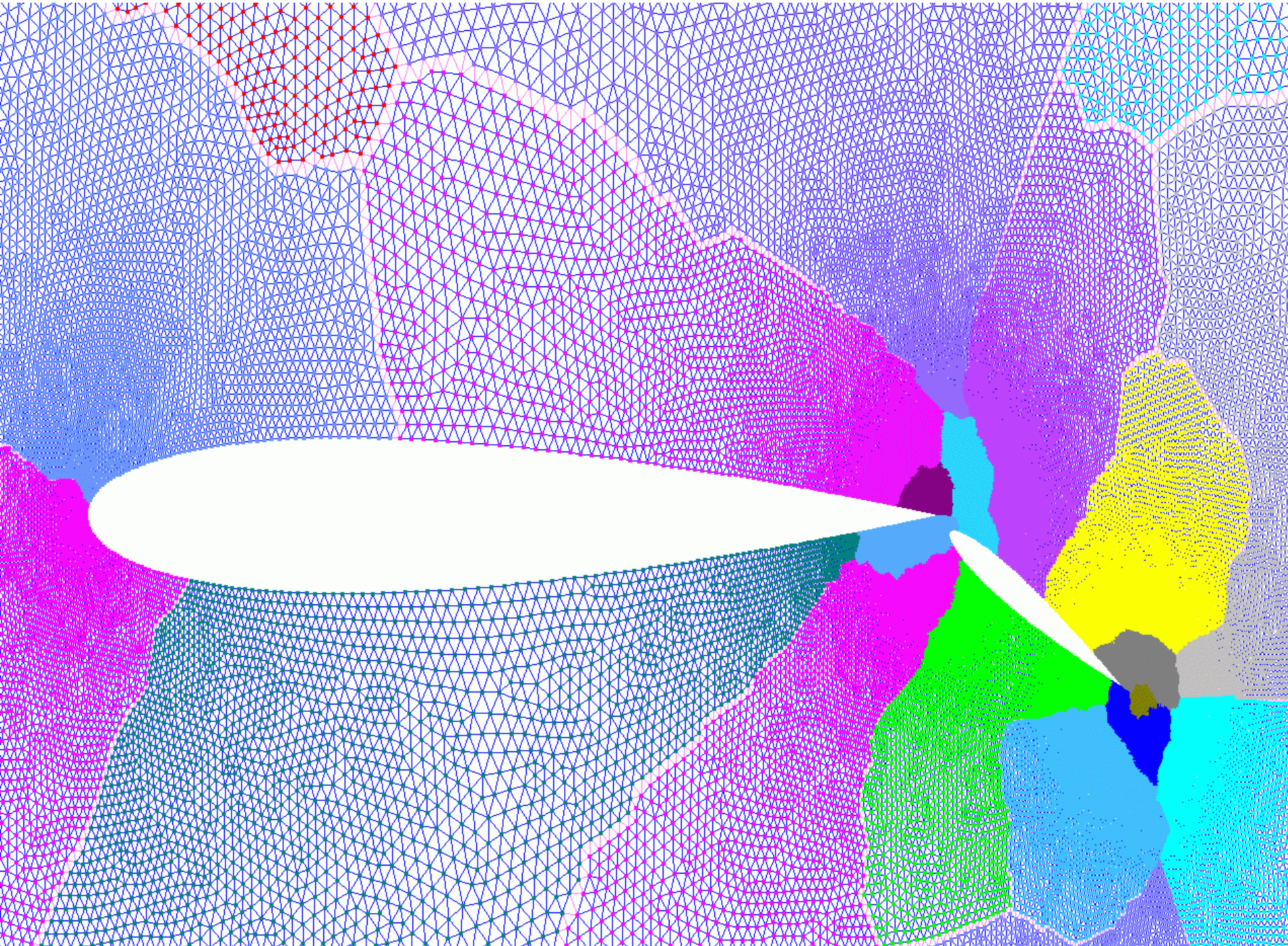


- Пакетный режим исполнения и отладки приложений
- Процедуры авторизованного доступа к удаленным системам
- Высокая динамика изменения конфигурации суперкомпьютеров
- Несоизмеримость ресурсов рабочей станции пользователя и суперкомпьютера

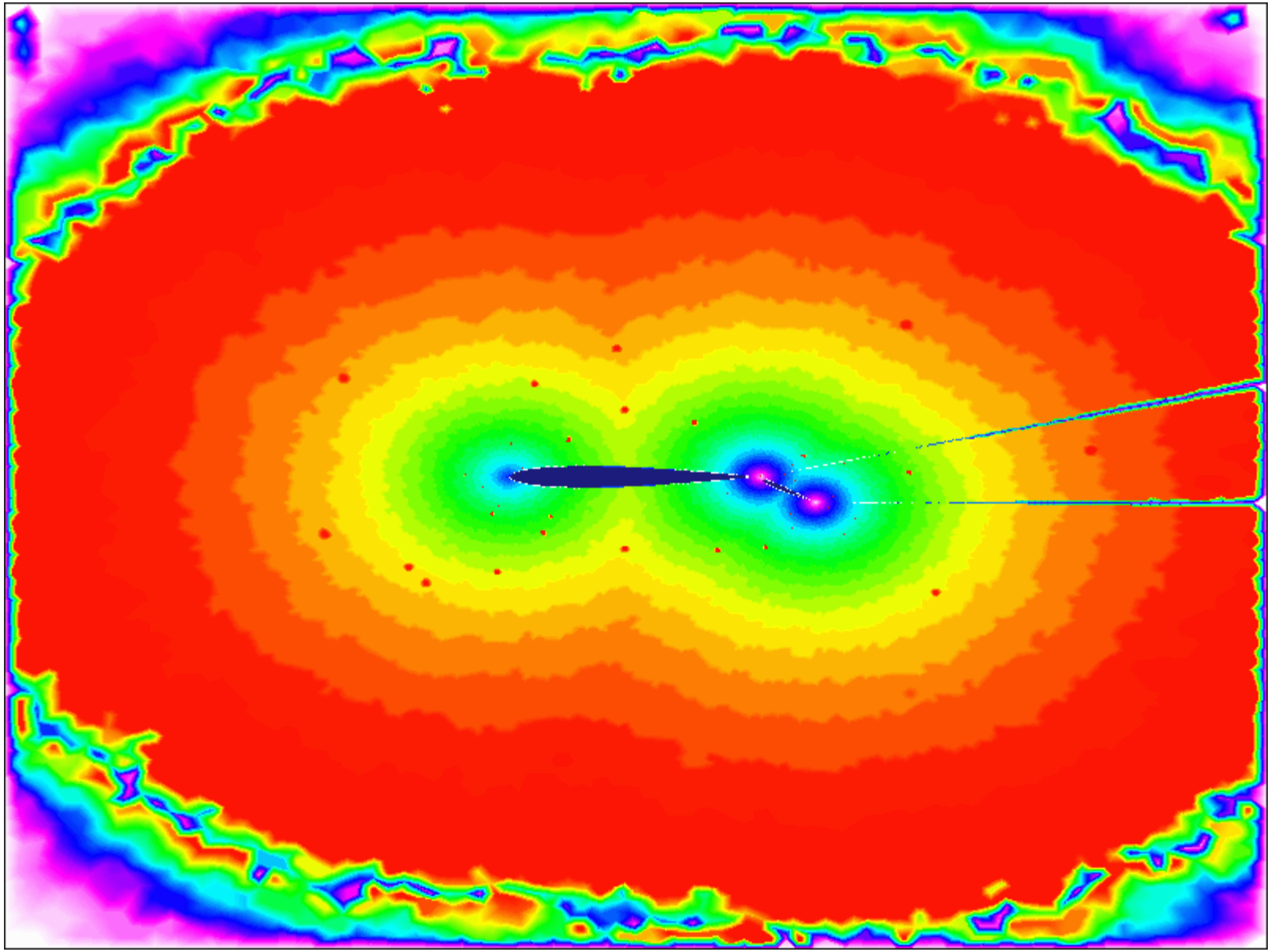
Статическая балансировка загрузки

- Критерии декомпозиции
- Инкрементный алгоритм декомпозиции
- Иерархическая обработка больших сеток

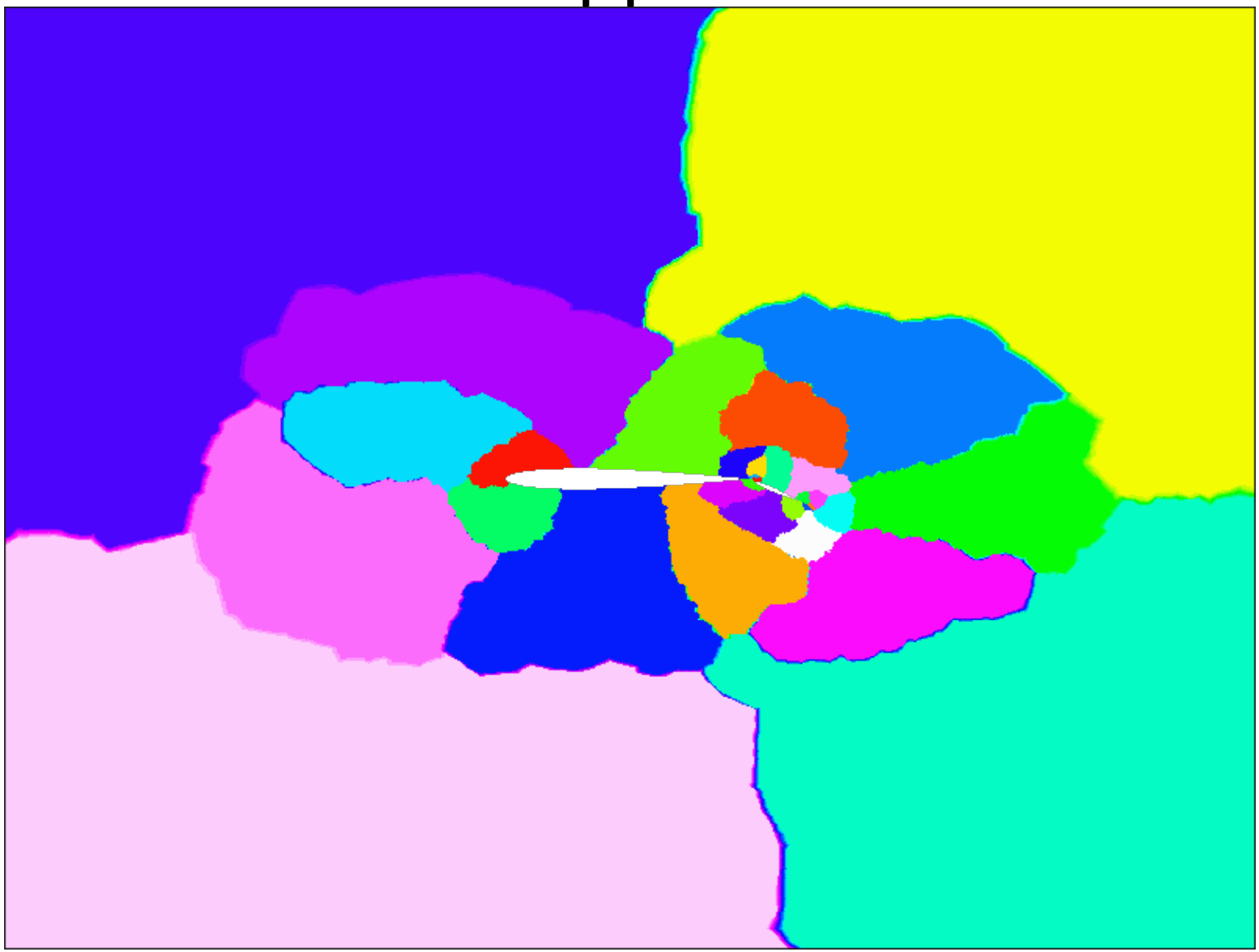




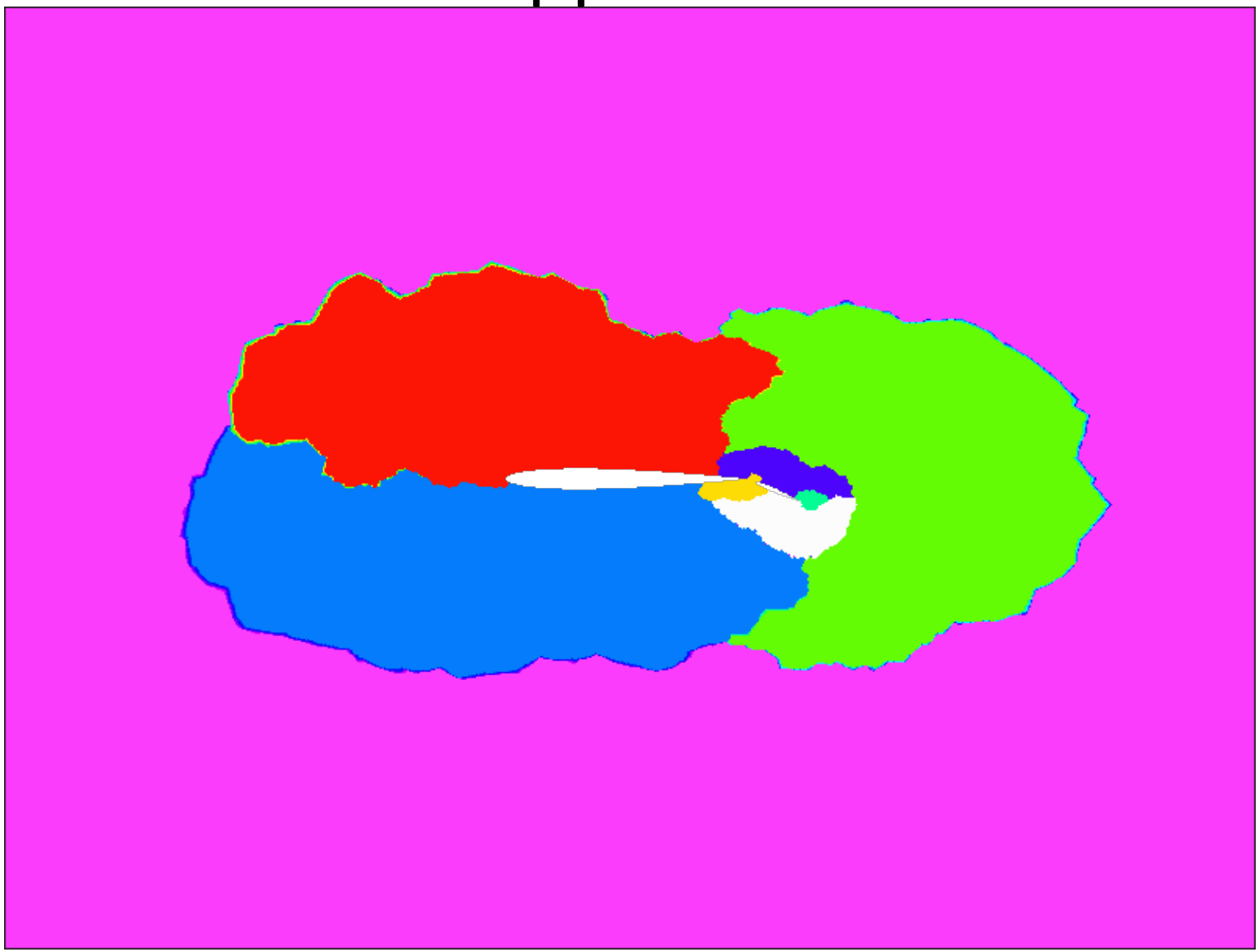
Простое разбиение на 32 домена



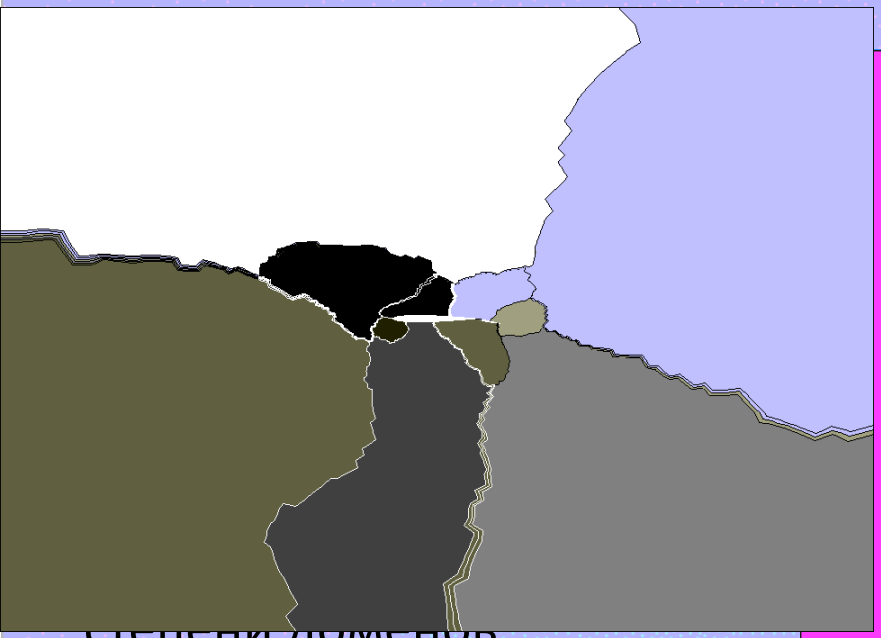
Рациональное разбиение на 32 домена

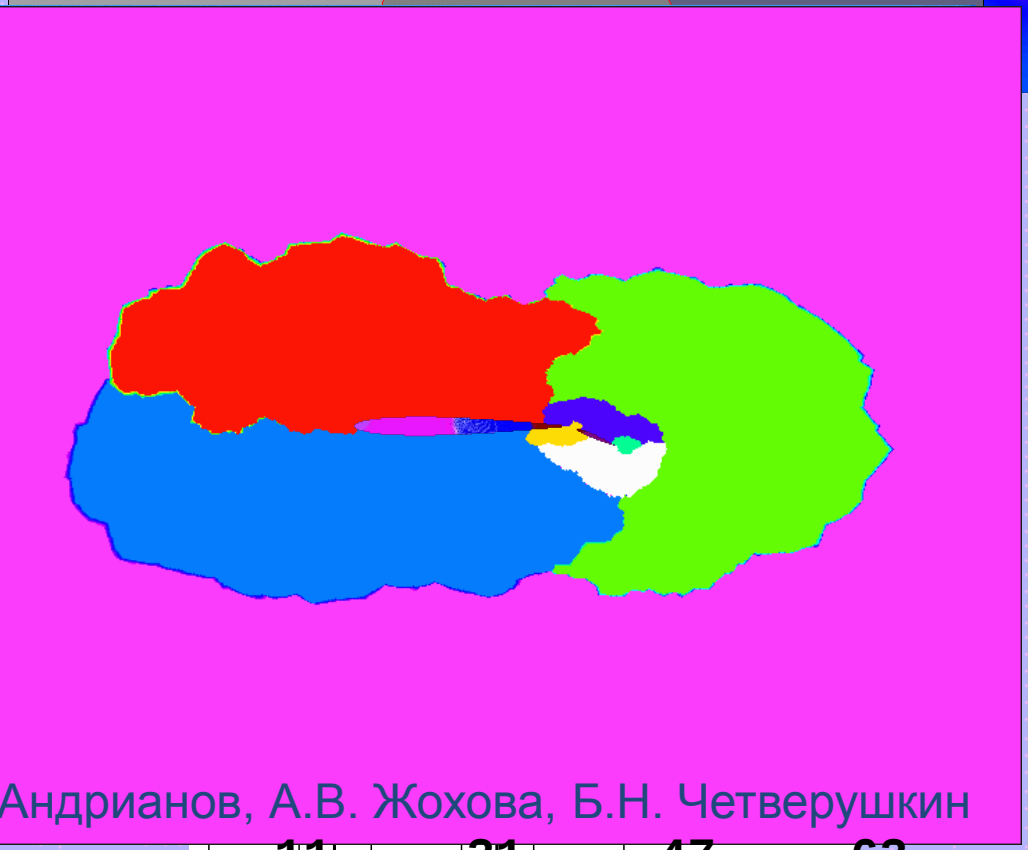


Рациональное разбиение на 8 доменов



Критерии декомпозиции графов





степени доменов

- Обеспечение связности доменов
- Обеспечение связности множества внутренних узлов доменов

А.Н. Андрианов, А.В. Жохова, Б.Н. Четверушкин

Процессоров	11	31	47	63
New_sort	13.59	5.59	4.38	4.16
METIS	13.61	11.00	11.10	10.56

Чему равно 25/4 ?

6.25

25/4=

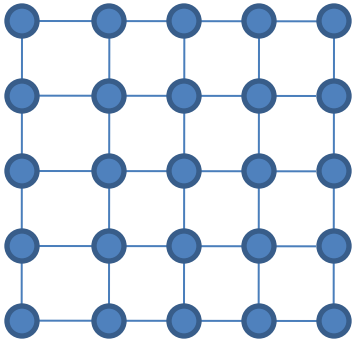
25/4=

6 9

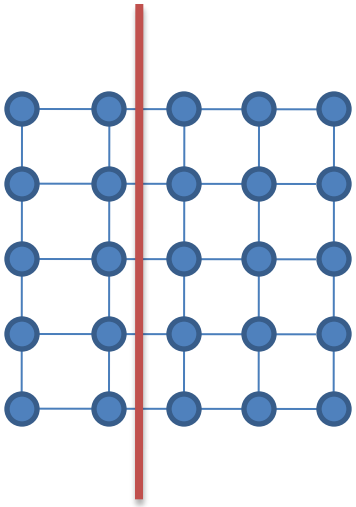
4

25/4 = 4 ? 6 ? 9

- Разрезать решетку 5 x 5 на 4 части

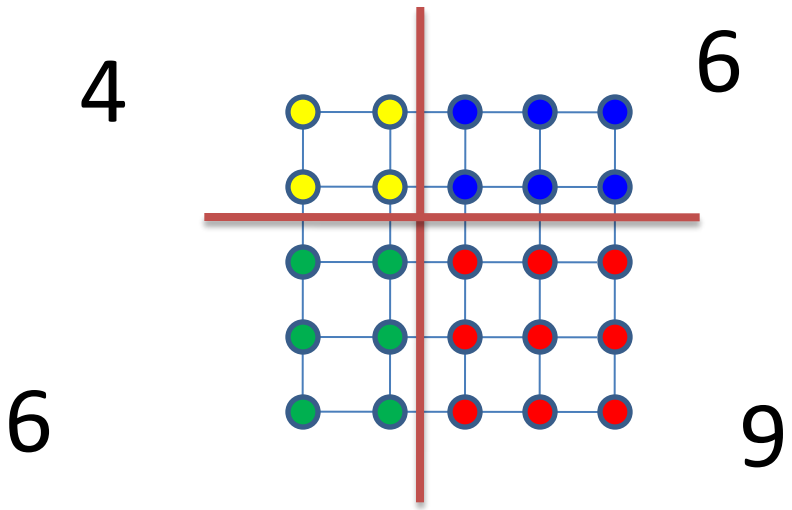


Декомпозиция сетки из 25 узлов на 4 части



25/4 = 4 ? 6 ? 9

- Декомпозиция решетки 5 x 5 на 4 домена



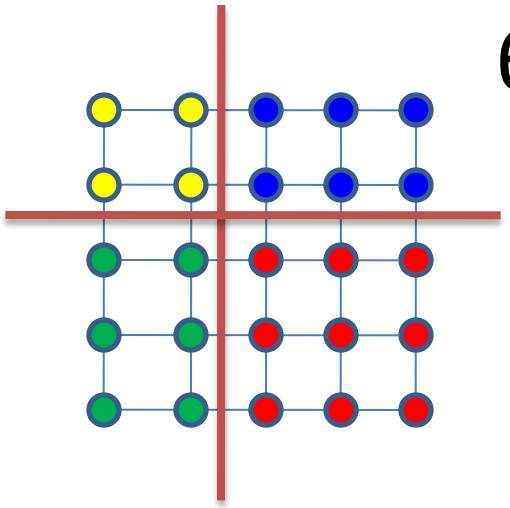
- Дисбаланс 9/4=2.25

25/4 = 4 ? 6 ? 9

- Декомпозиция решетки 5 x 5 на 4 домена

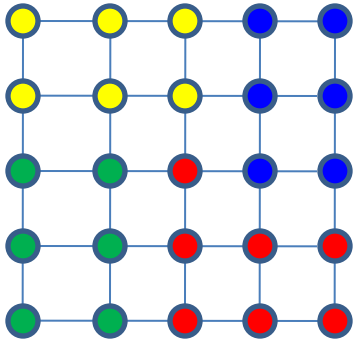
4

6

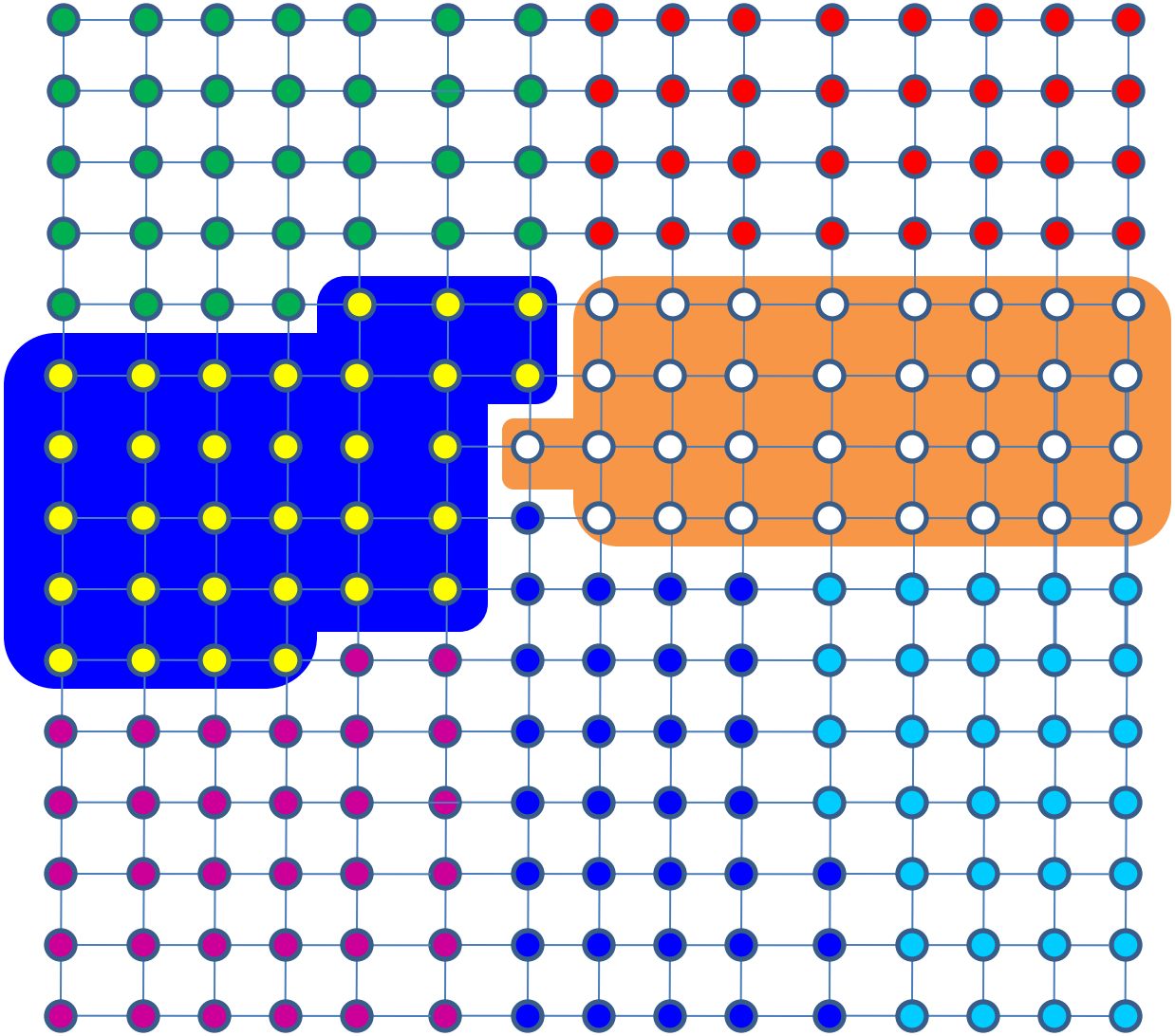


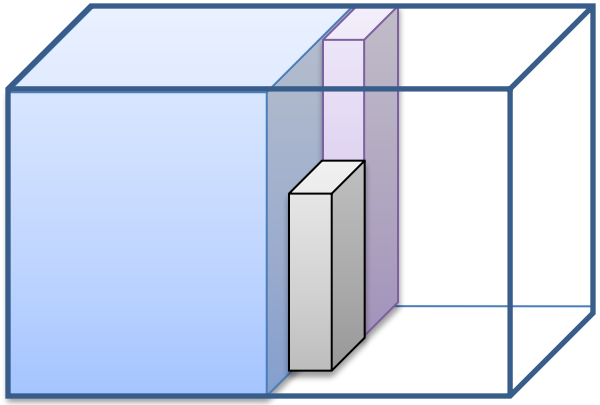
6

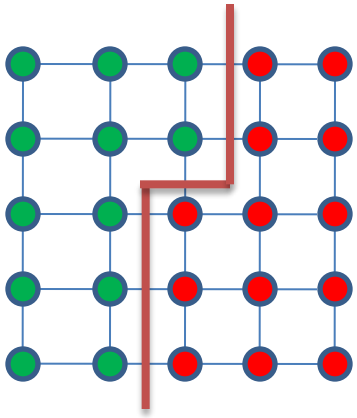
9



Декомпозиция сетки 25x25 на 7 частей



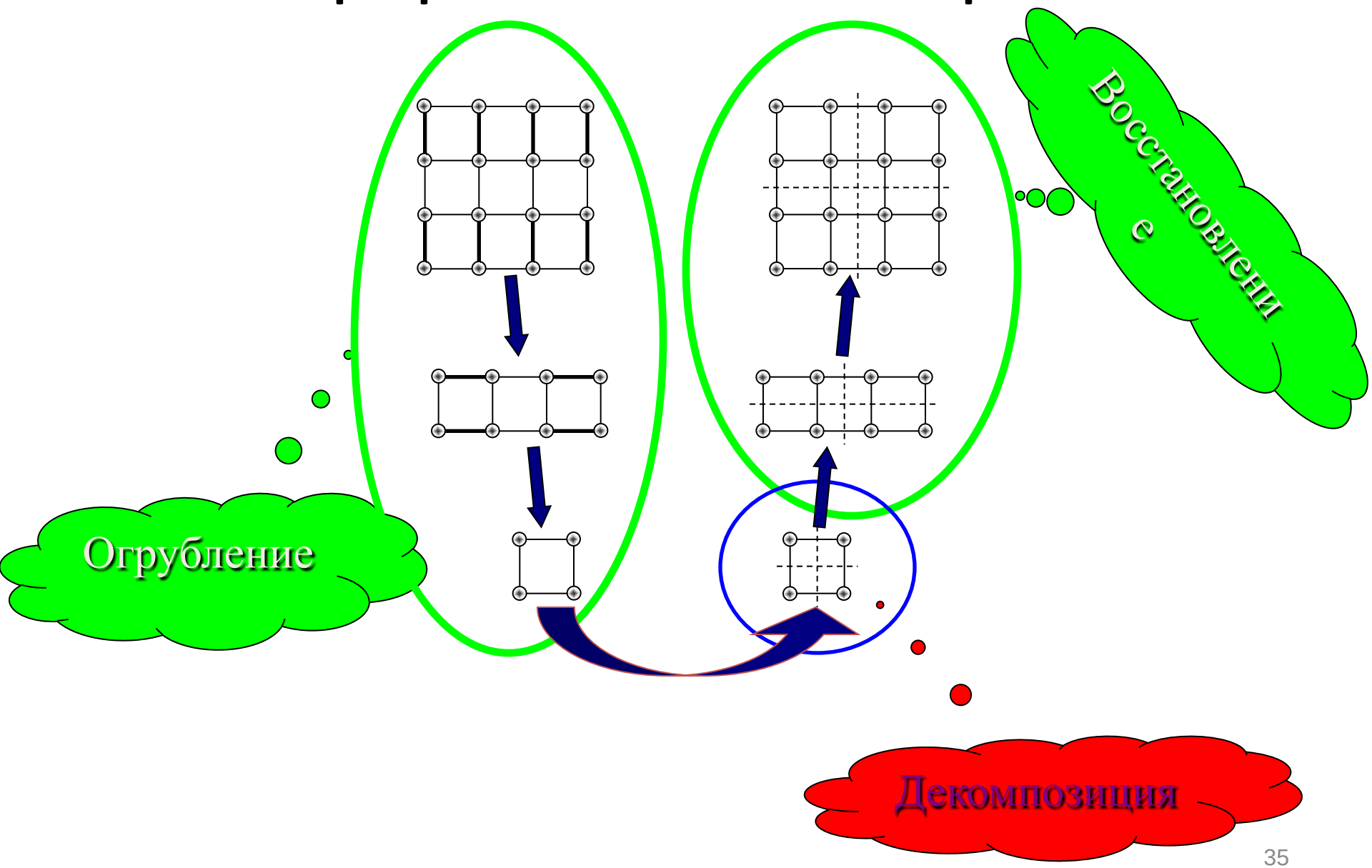




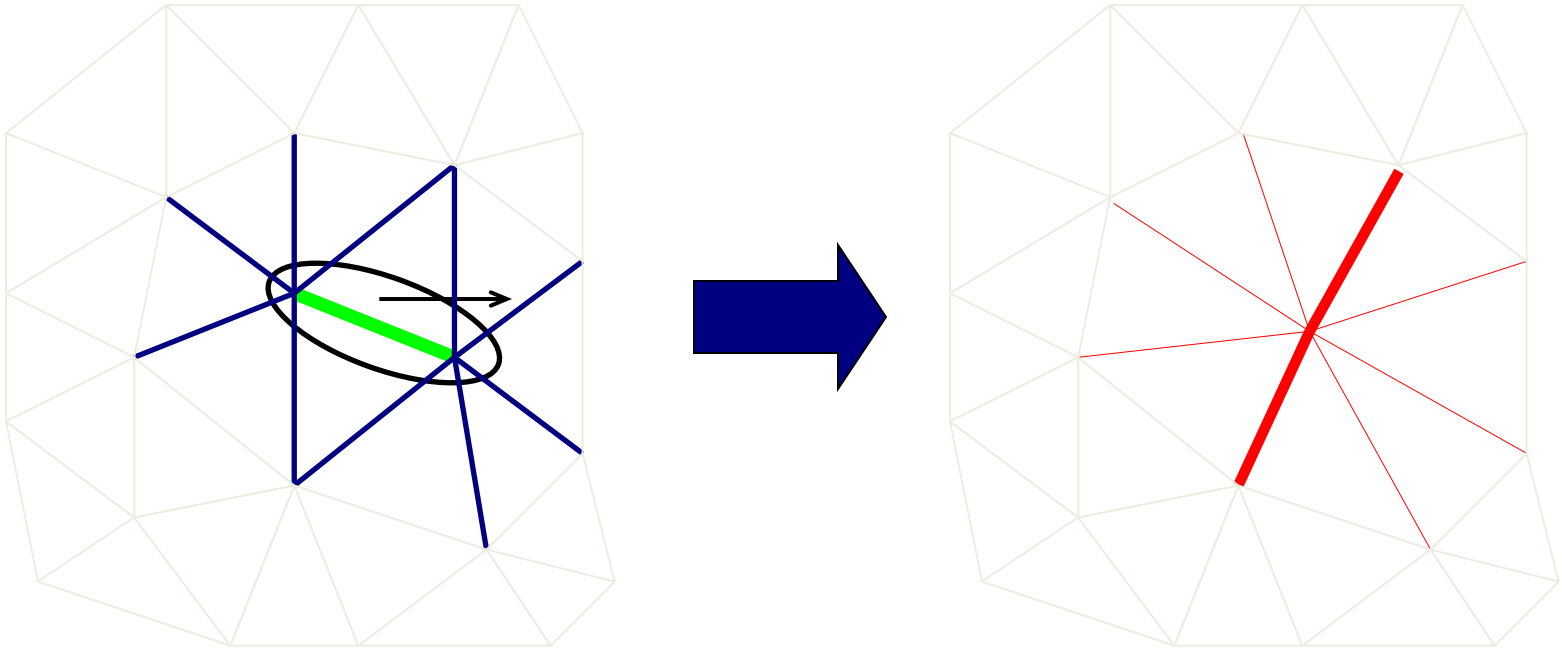
Пакеты декомпозиции графов

Chaco	Bruce Hendrickson Robert Leland
ParMETIS	George Karypis Vipin Kumar
PARTY	Robert Prais, et al.
JOSTLE	Chris Walshaw, et al.
SCOTCH	Francois Pellegrini

Иерархический алгоритм



Огрубление графа

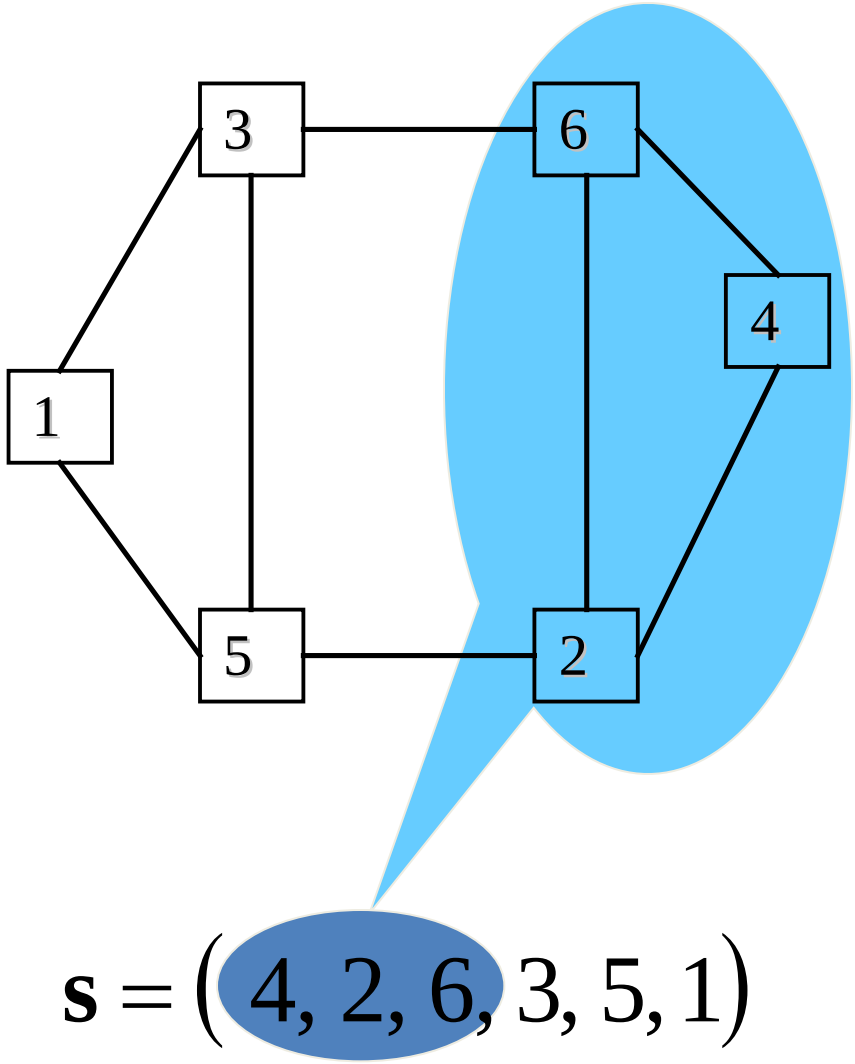


Спектральный метод

$$A = \begin{pmatrix} -2 & 0 & 1 & 0 & 1 & 0 \\ 0 & -3 & 0 & 1 & 1 & 1 \\ 1 & 0 & -3 & 0 & 1 & 1 \\ 0 & 1 & 0 & -2 & 0 & 1 \\ 1 & 1 & 1 & 0 & -3 & 0 \\ 0 & 1 & 1 & 1 & 0 & -3 \end{pmatrix}$$

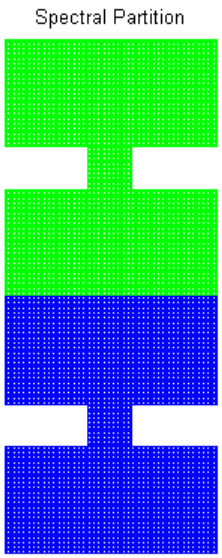
$$\lambda_2 = -1$$

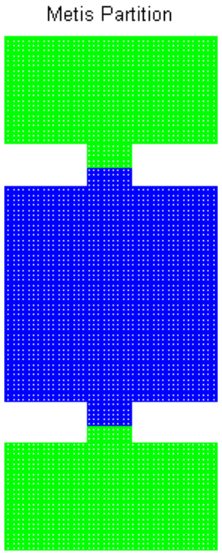
$$x_2 = (2, -1, 1, -2, 1, -1)$$



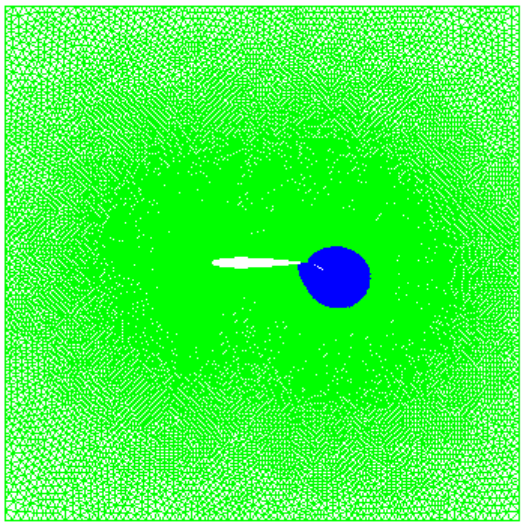
$$\mathbf{s} = (4, 2, 6, 3, 5, 1)$$

Метод спектральной бисекции



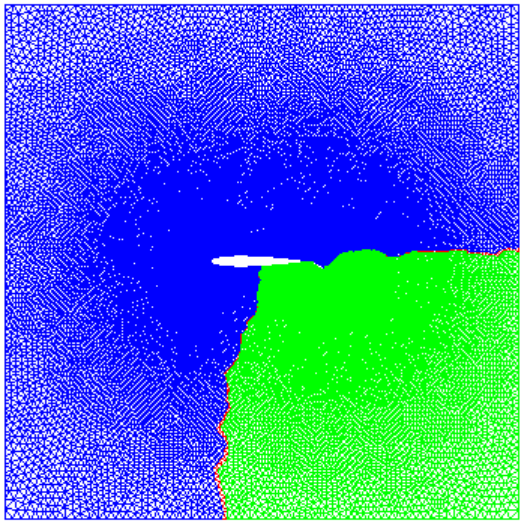


100 cut edges



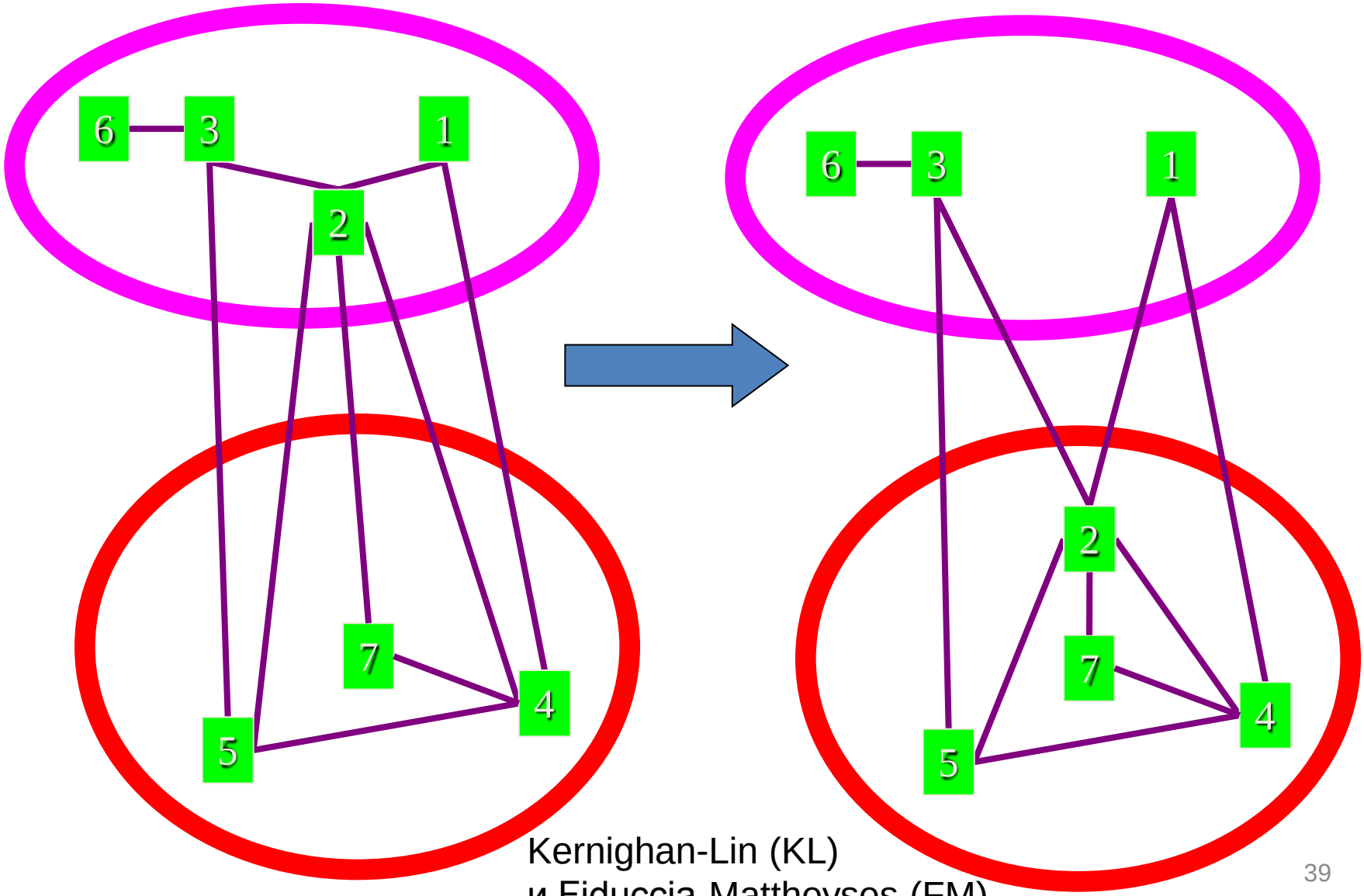
655 cut edges

42 cut edges



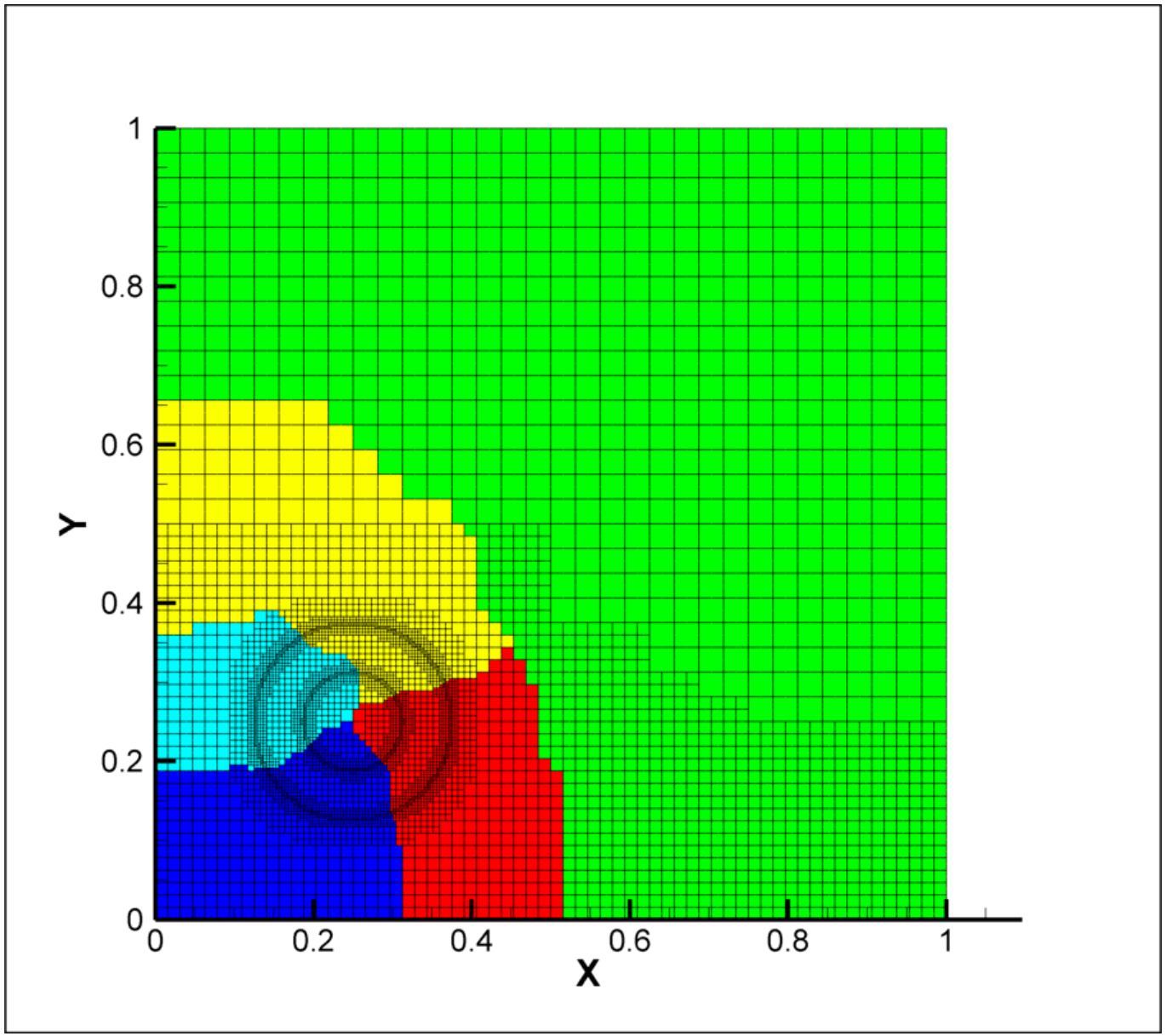
524 cut edges

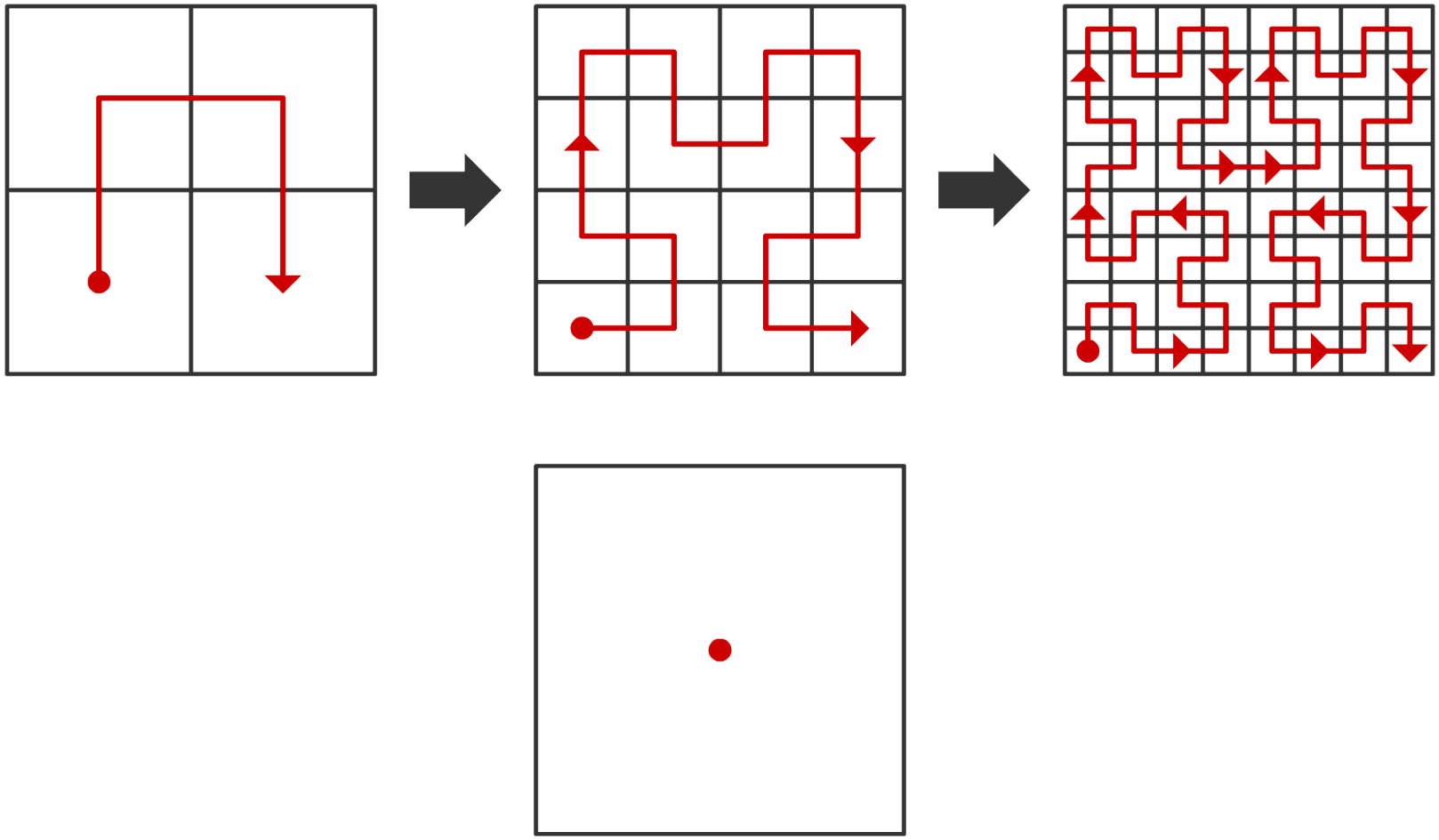
Локальное уточнение



Kernighan-Lin (KL)
и Fiduccia-Mattheyses (FM)

Декомпозиция пакетом Metis





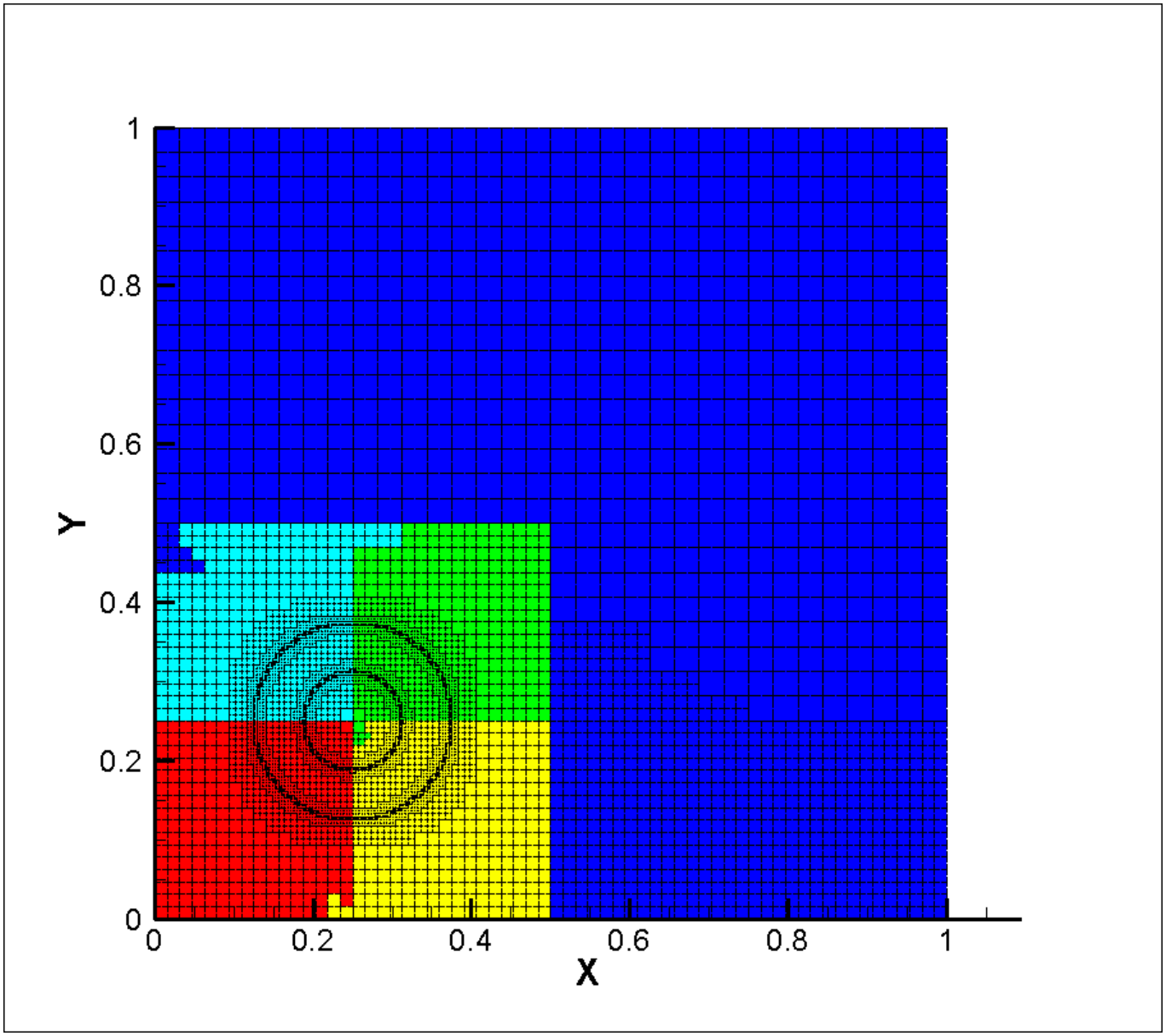
Hilbert-curve ordering

This ordering can be built by simple recursive procedure.

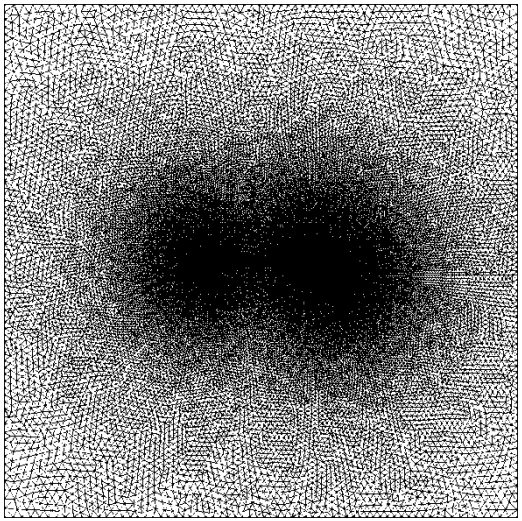
When mesh changes locally, Hilbert curve changes locally too.

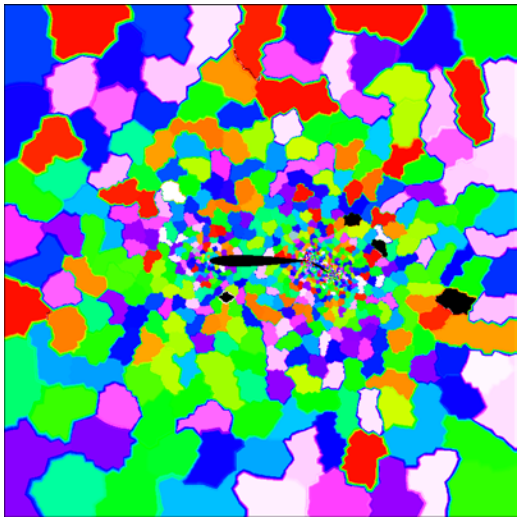
It cannot be used for parallel computations due to chain dependence of elements.

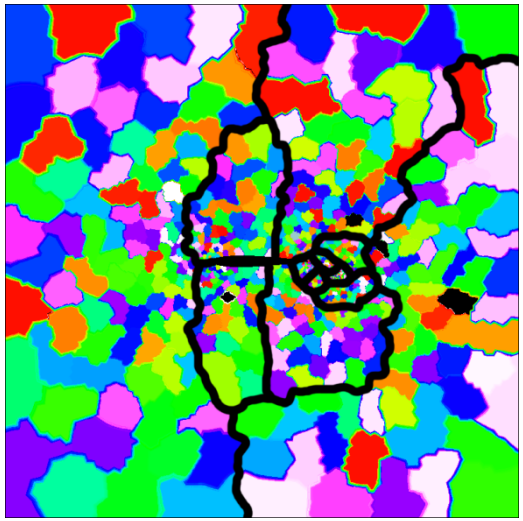
Декомпозиция по кривой Гильберта



Двухуровневое разбиение







I

II

Сетка предварительно разбивается на большое число *микродоменов,* образующих *макрограф*

Вершины макрографа распределяются по процессорам

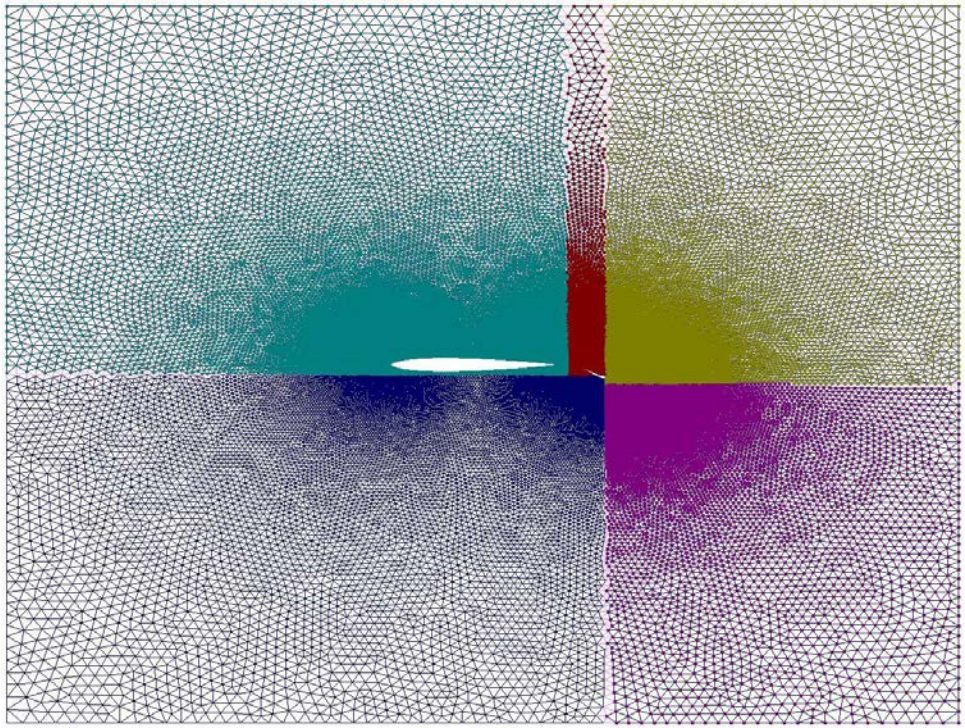
Разбиение тетраэдральной сетки, содержащей $2 \cdot 10^8$ узлов, на 125 процессорах

- вычисления производились на кластере СКИФ МГу (1250 4-хядерных процессоров, 60 TFlop/s)

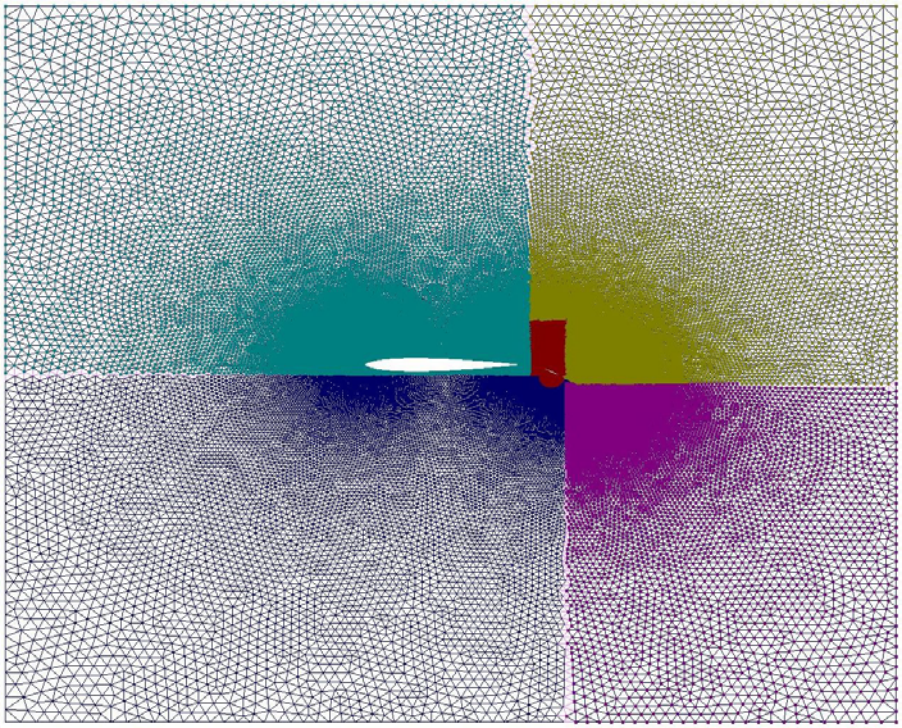
		геометрическая декомпозиция		ParMETIS	
число доменов		80 000		20 000	
время		21 сек.		10 сек.	
число вершин в домене		2612	2613	2 328	10 932
мин.	макс.				
среднее число связей с соседними доменами		14		14	
число некомпактных доменов		229		364	

Треугольная сетка из 75790 вершин (пространство вокруг крыла)

результат геометрической декомпозиции на 5 групп
(в дальнейшем каждый процессор считывает свою группу вершин)

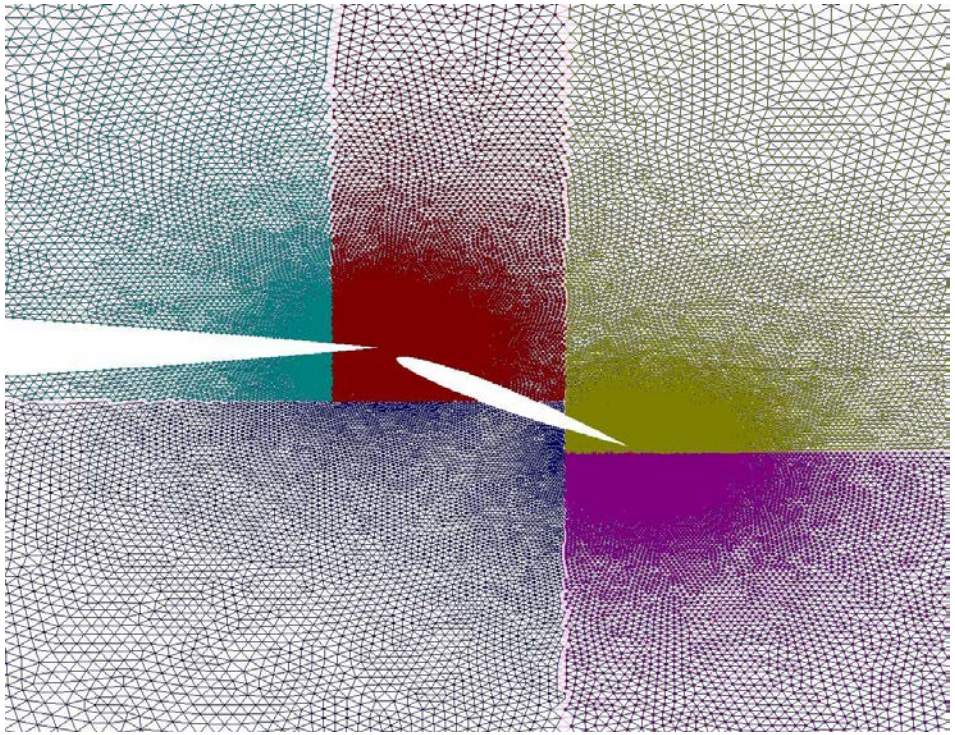


результат перераспределения малых блоков вершин

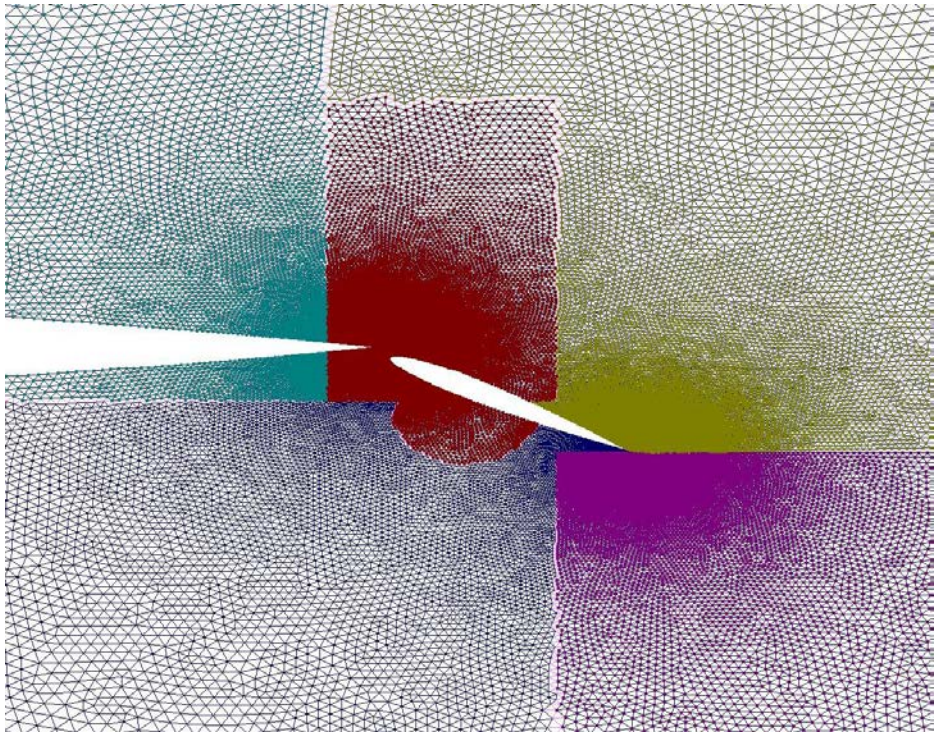


Фрагмент треугольной сетки из 75790 вершин

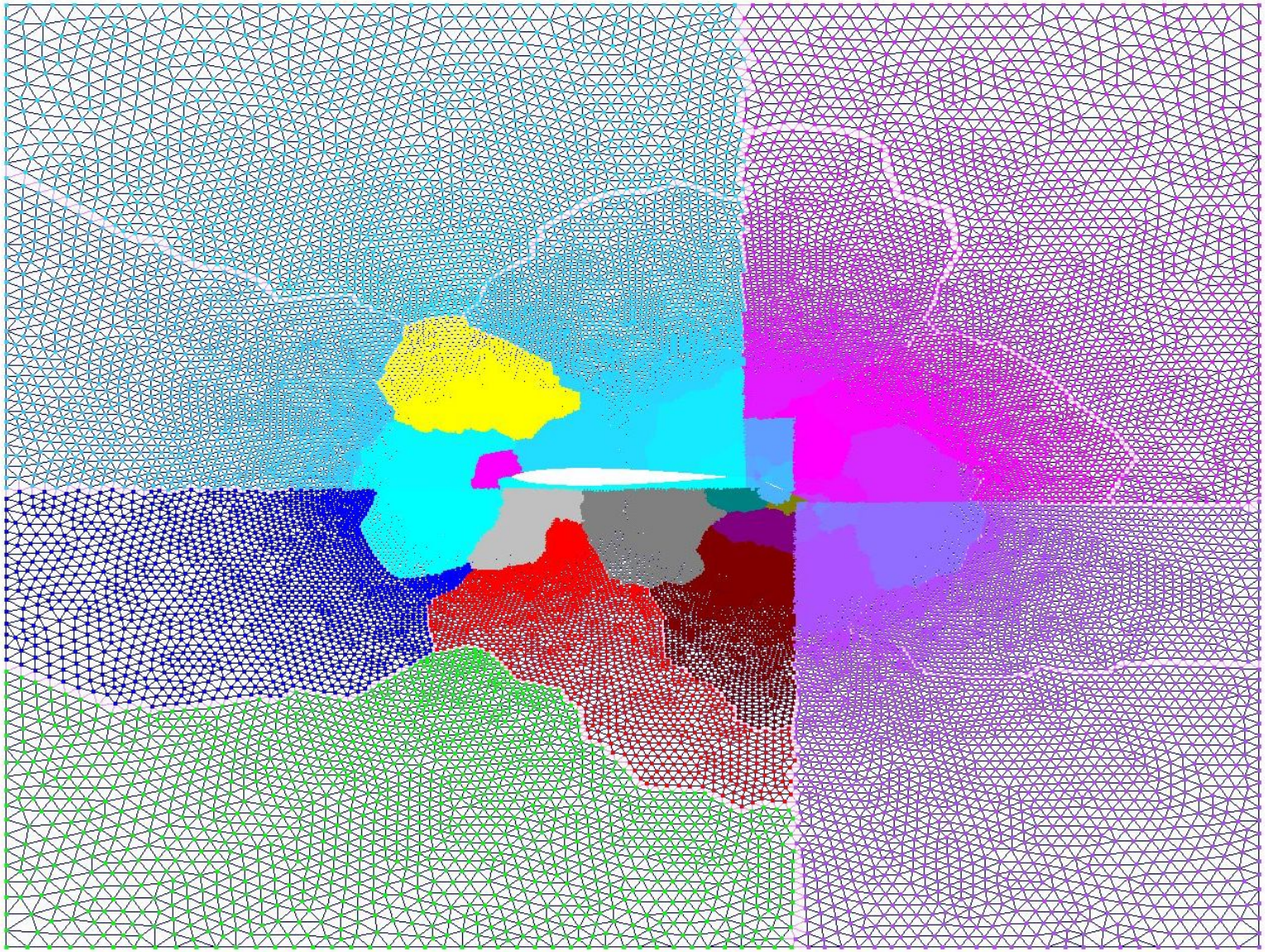
результат геометрической декомпозиции



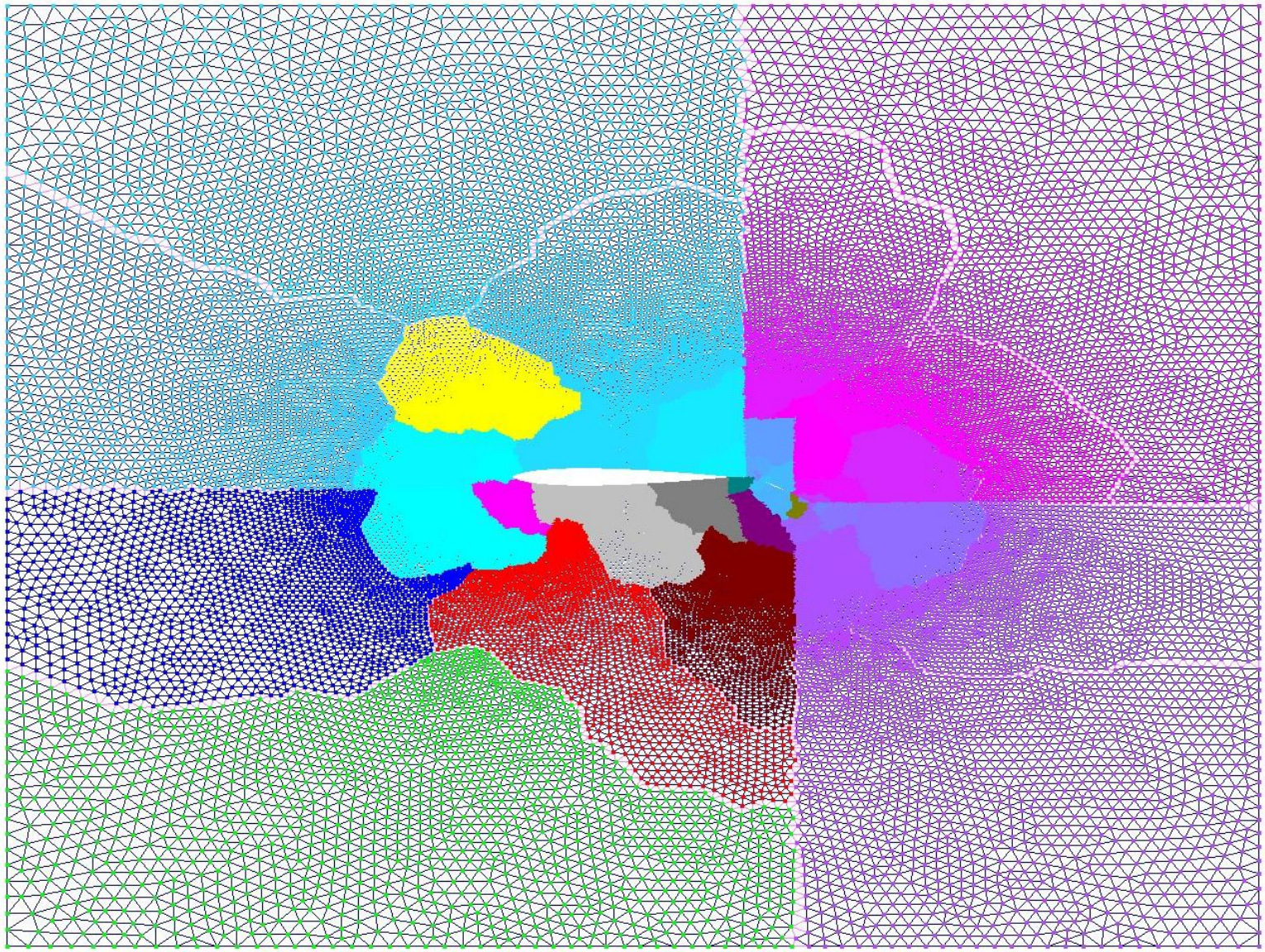
результат перераспределения малых блоков вершин



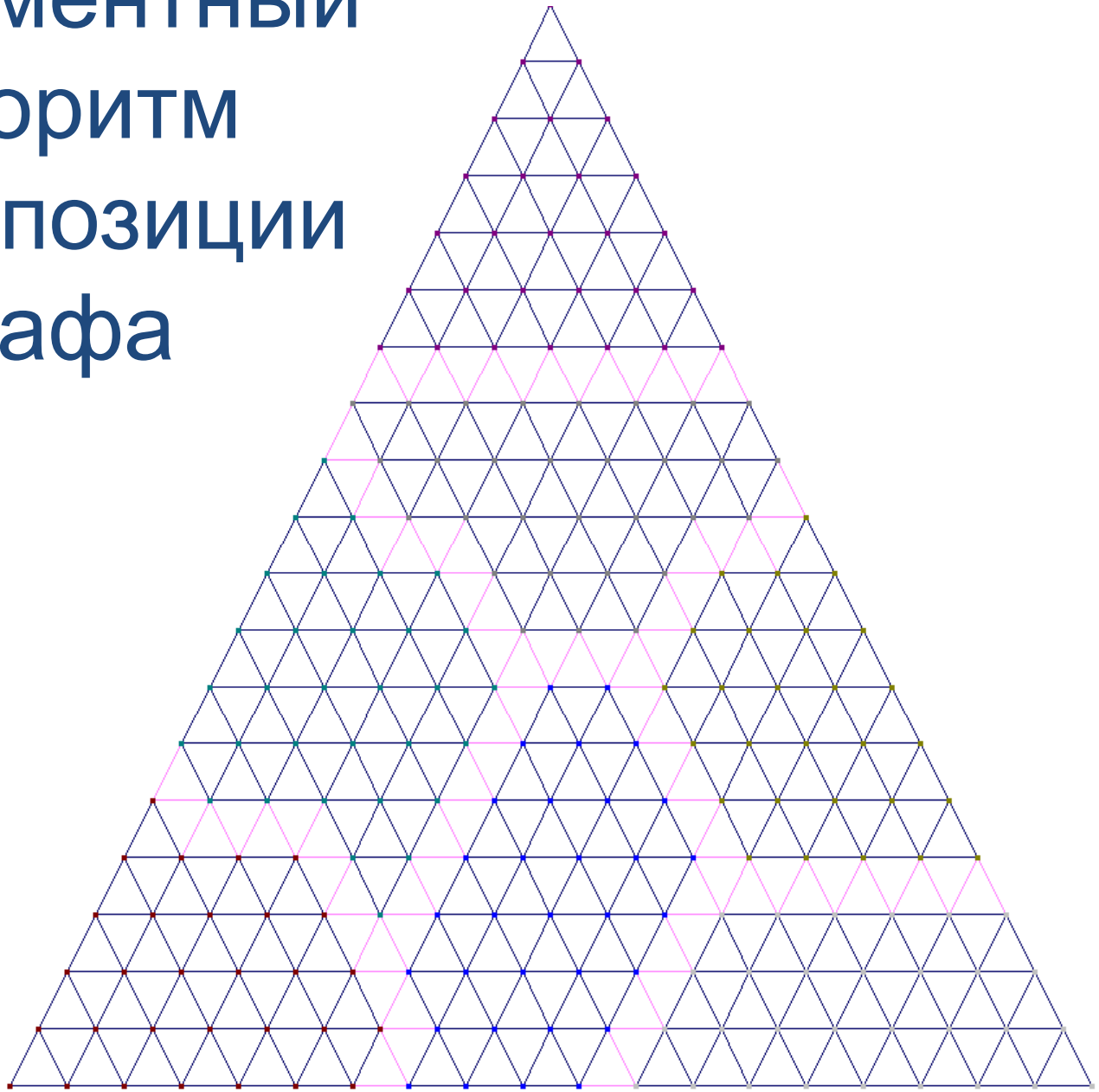
Результат локального разбиения сетки из 75790 вершин на 50 доменов на 5 процессорах



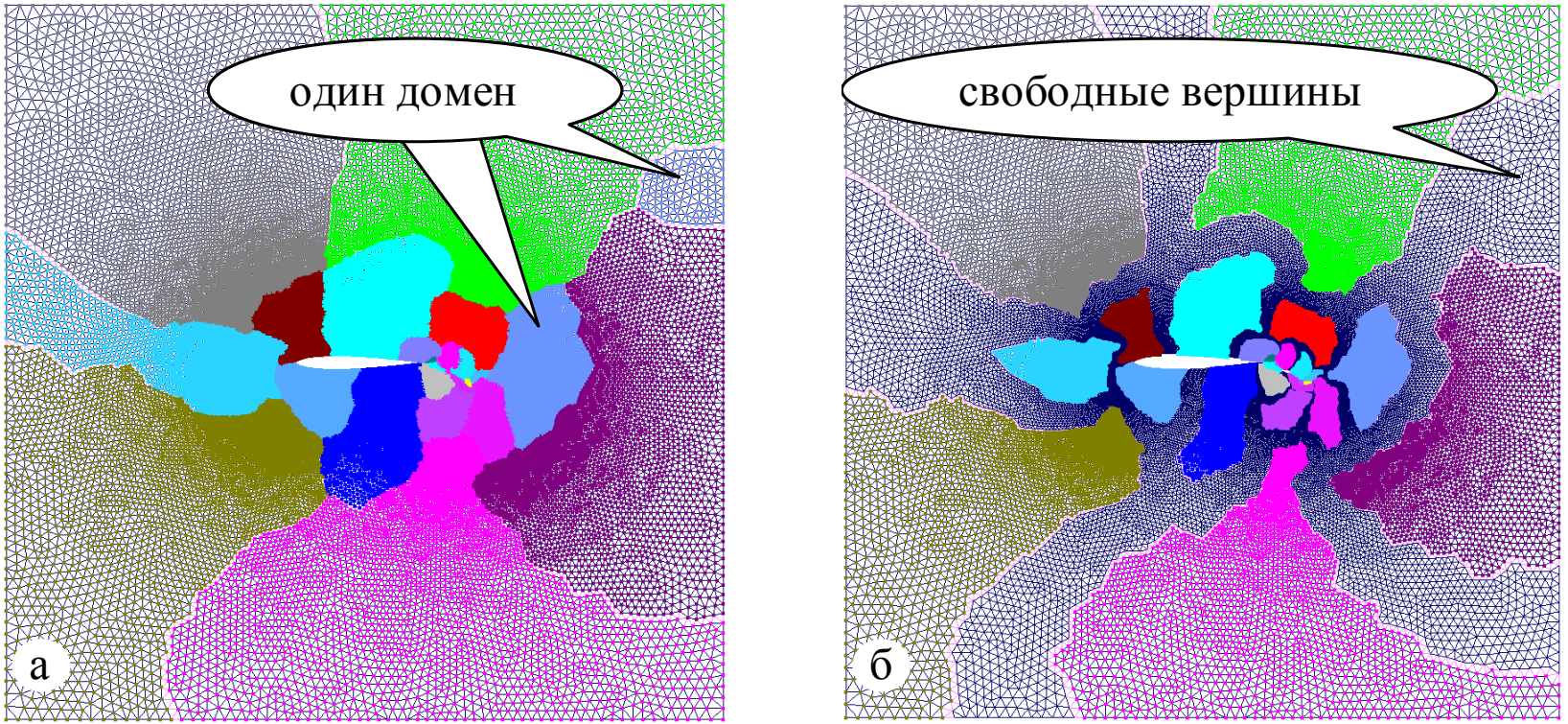
Результат сбора плохих групп доменов и их повторного разбиения



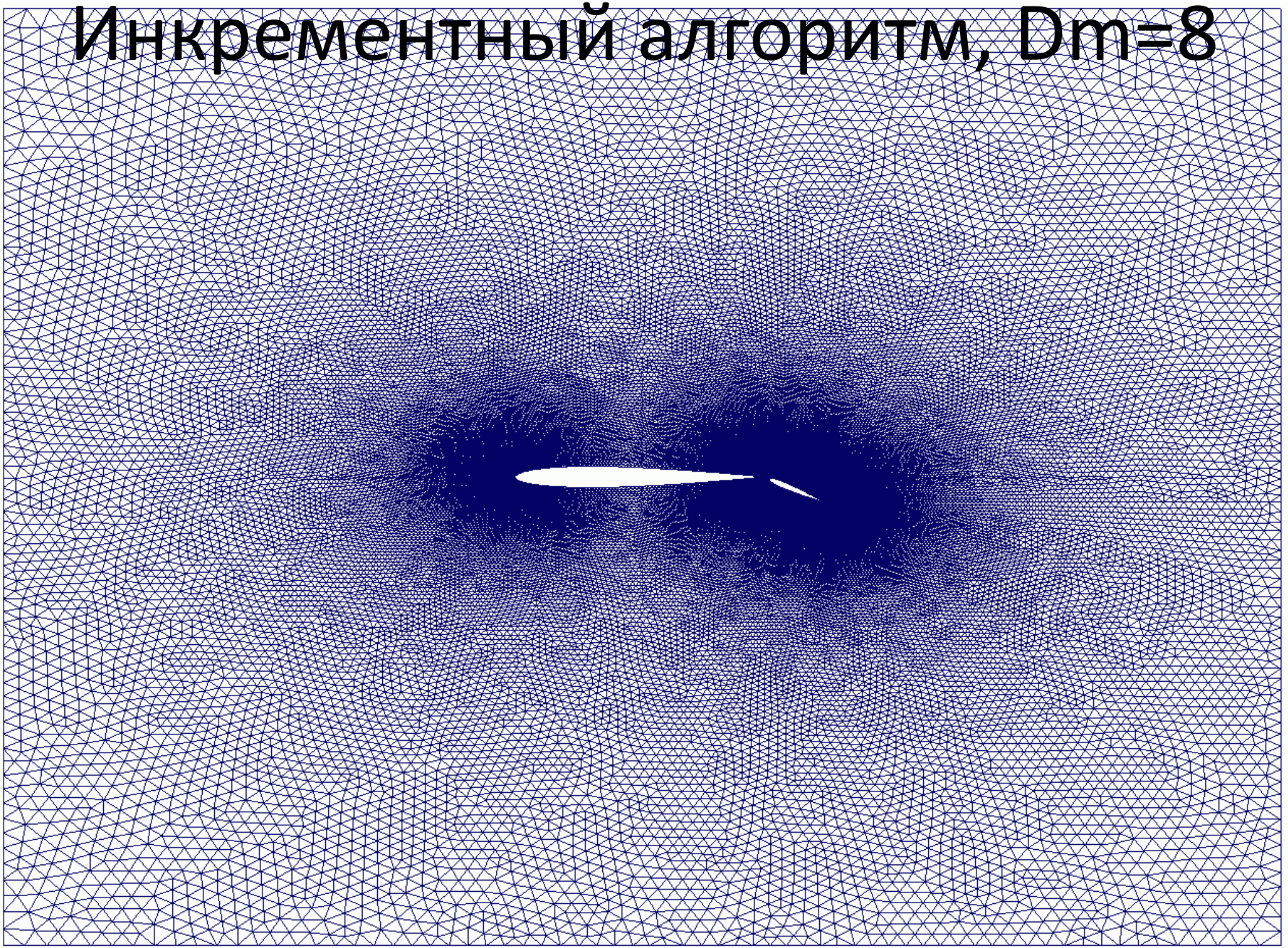
Инкрементный алгоритм декомпозиции графа



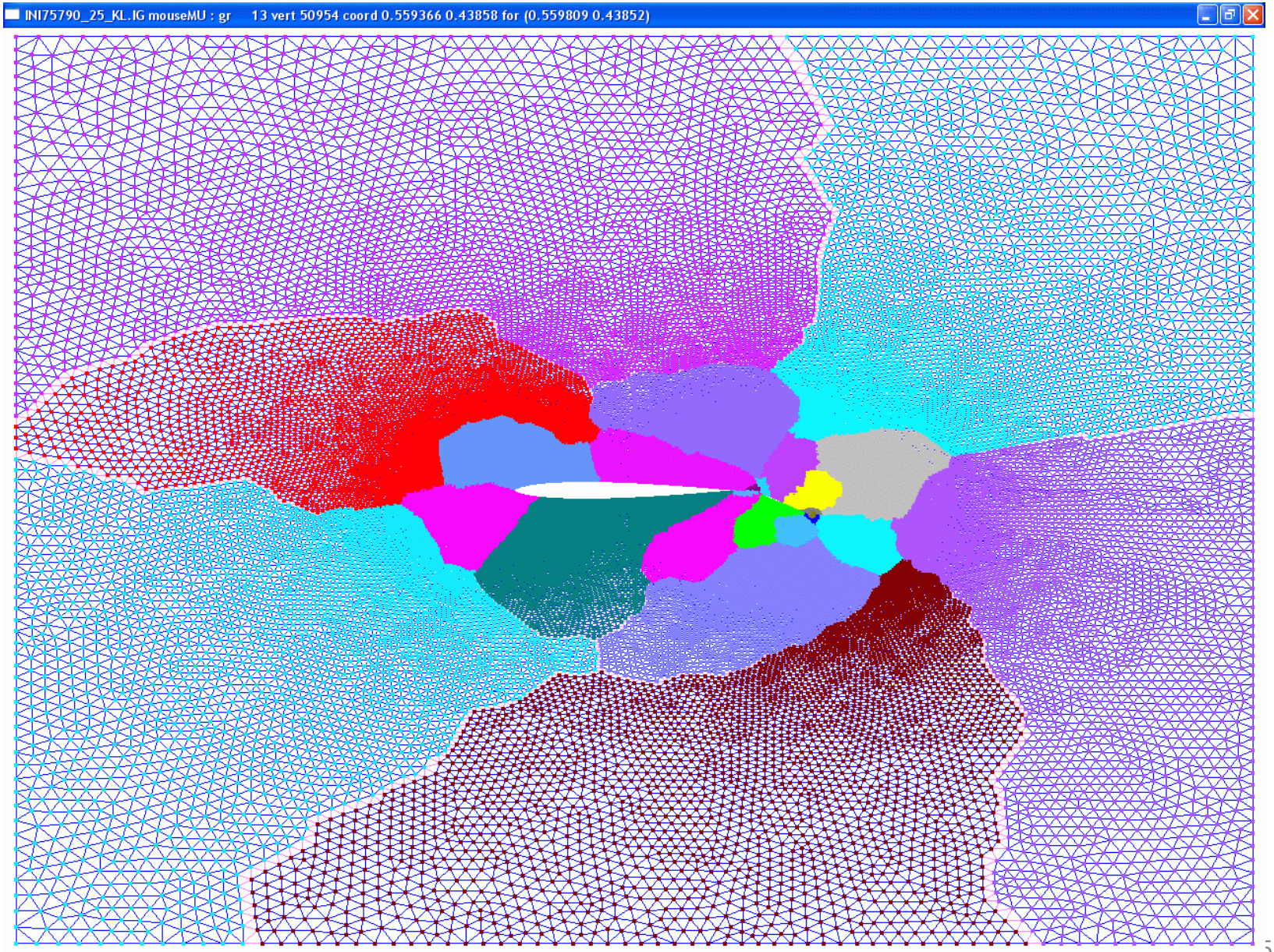
Редуцирование доменов



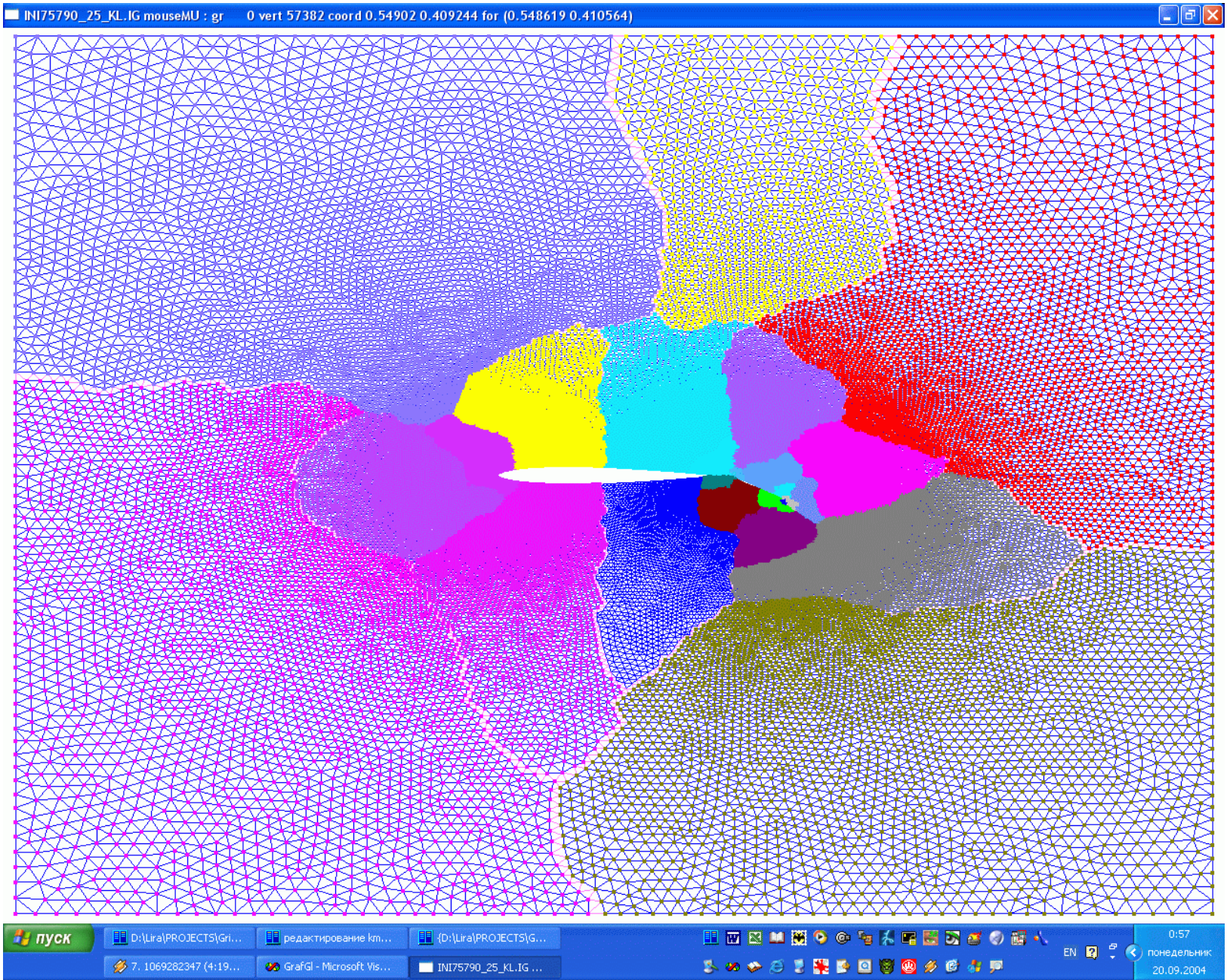
Инкрементный алгоритм, Dm=8



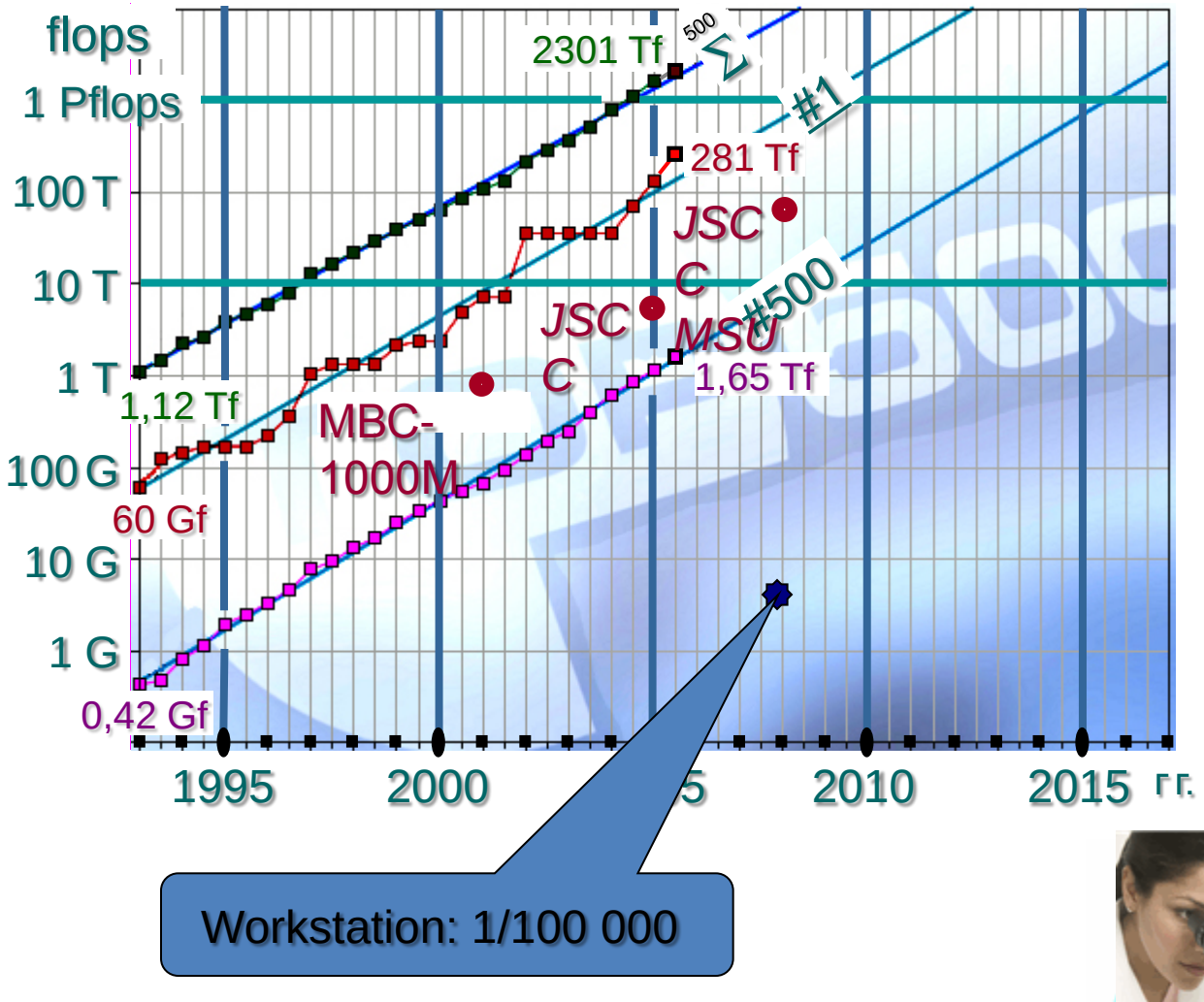
Инкрементный алгоритм, Dm=25



Kmetis, Dm=25

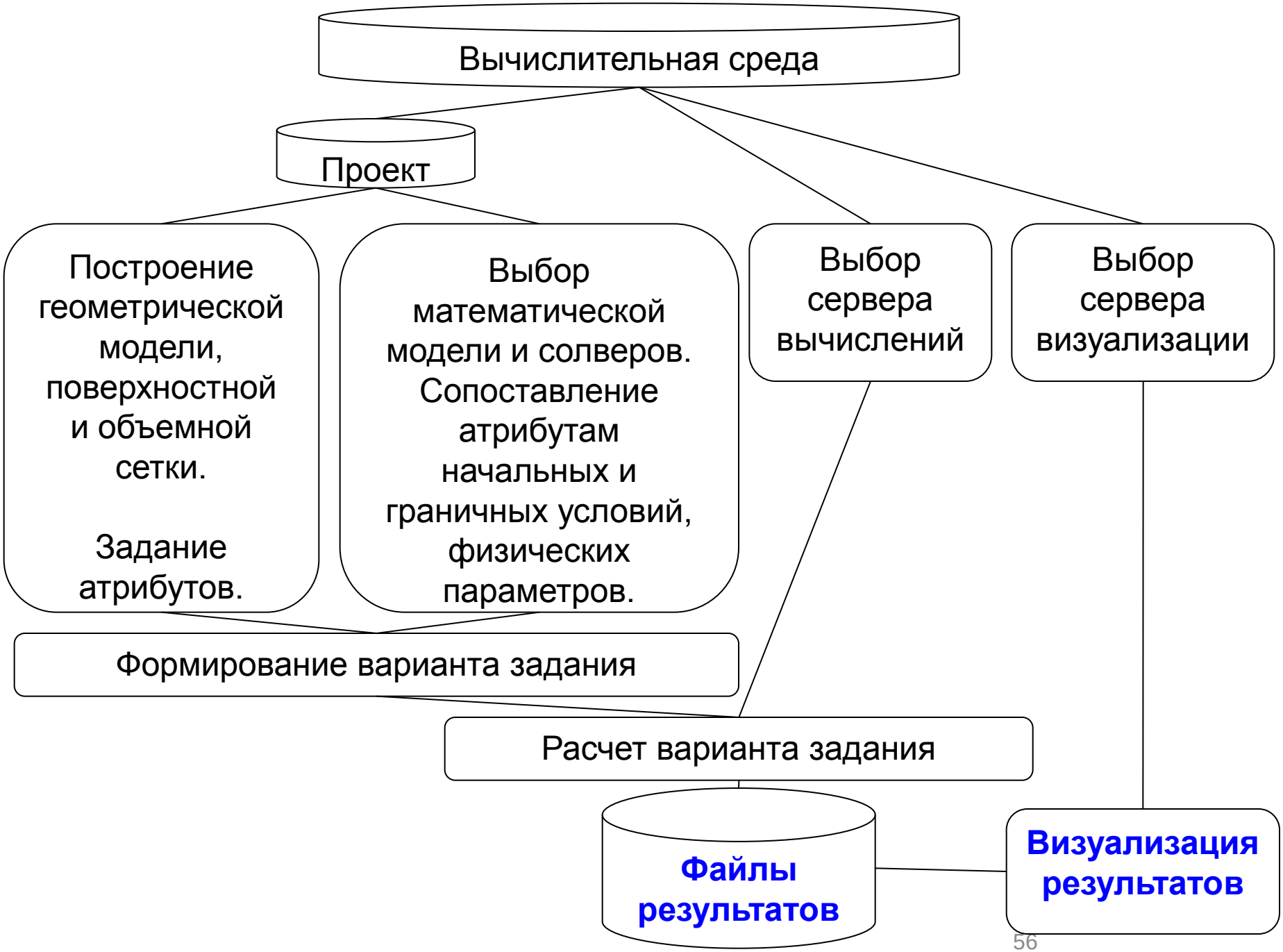

INI75790_25_KL.IG mouseMU : gr 0 vert 57382 coord 0.54902 0.409244 for (0.548619 0.410564)
пуск
D:\Lira\PROJECTS\Gri...
редактирование km...
{D:\Lira\PROJECTS\G...
7. 1069282347 (4:19...
GrafGl - Microsoft Vis...
INI75790_25_KL.IG ...
EN
0:57
понедельник
20.09.2004

Визуализация сеточных данных



- Оперативная память
- Кеш
- Операционные устройства
- Множественный доступ
- Бета-тестер

TOP 500


Вычислительная среда
Проект
Построение геометрической модели, поверхностной и объемной сетки.
Задание атрибутов.
Выбор математической модели и солверов. Сопоставление атрибутам начальных и граничных условий, физических параметров.
Выбор сервера вычислений
Выбор сервера визуализации
Формирование варианта задания
Расчет варианта задания
Файлы результатов
Визуализация результатов

Визуализация

- Скалярные
- Векторные

- Стационарные
- Зависящие от времени

- Решетки
- Треугольные и тетраэдральные сетки

Этапы визуализации

Запись

Сетка *Сеточная функция*

Чтение

Формирование объектов виртуальной сцены

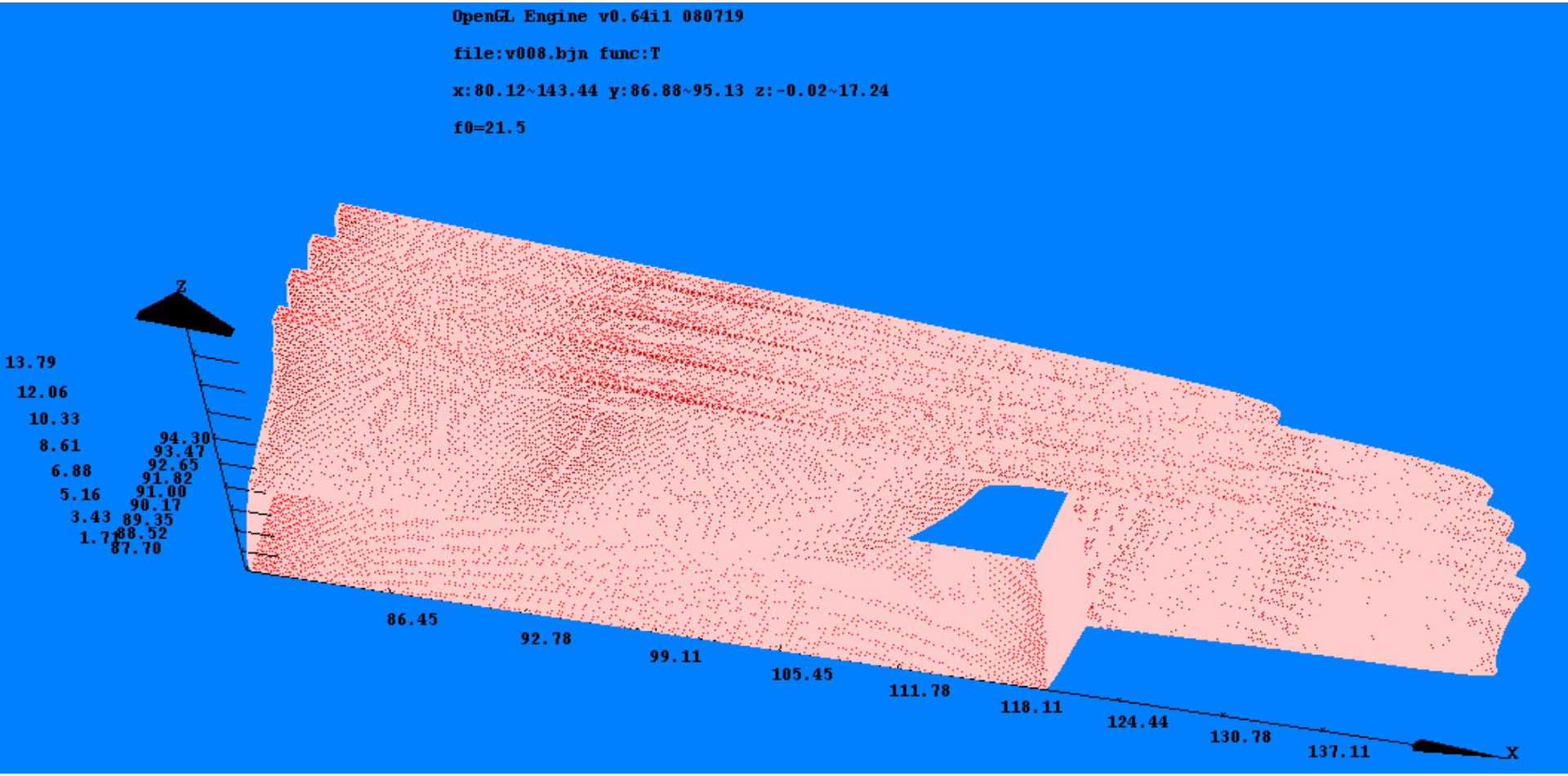
Отображение

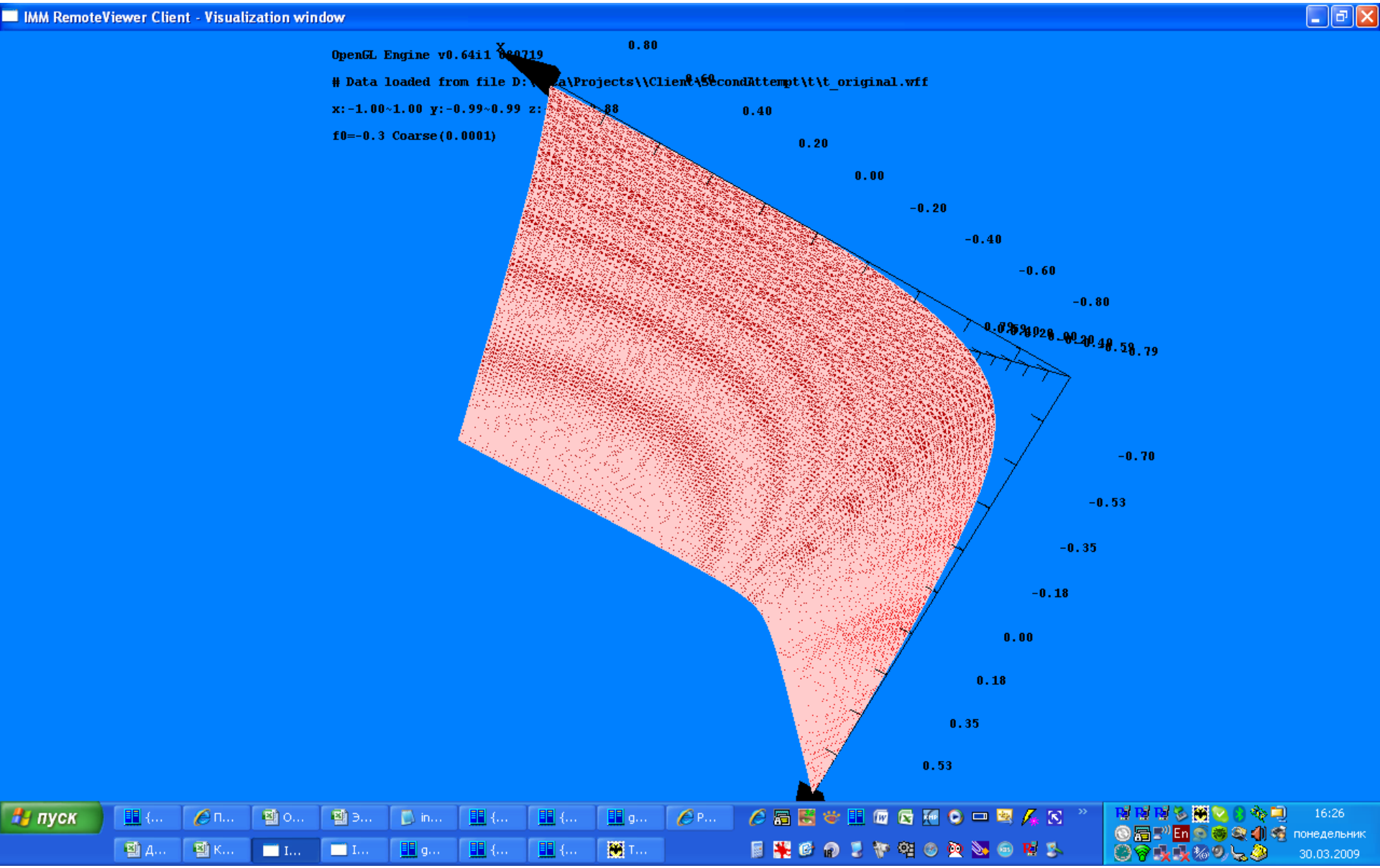
Методы

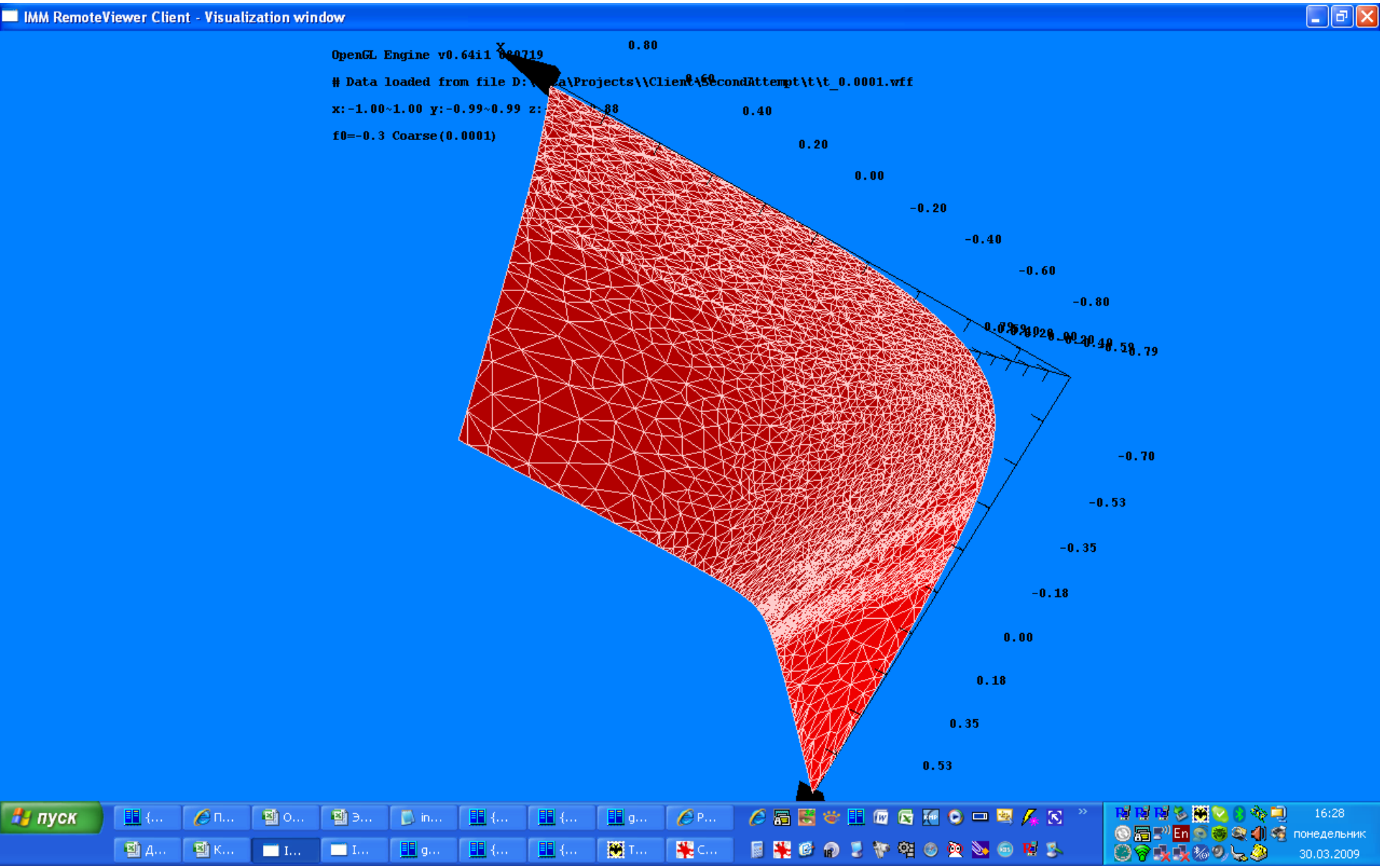
- Распределенное иерархическое хранение
- Декомпозиция
- Огрубление с контролируемой точностью
- Клиент-серверная технология

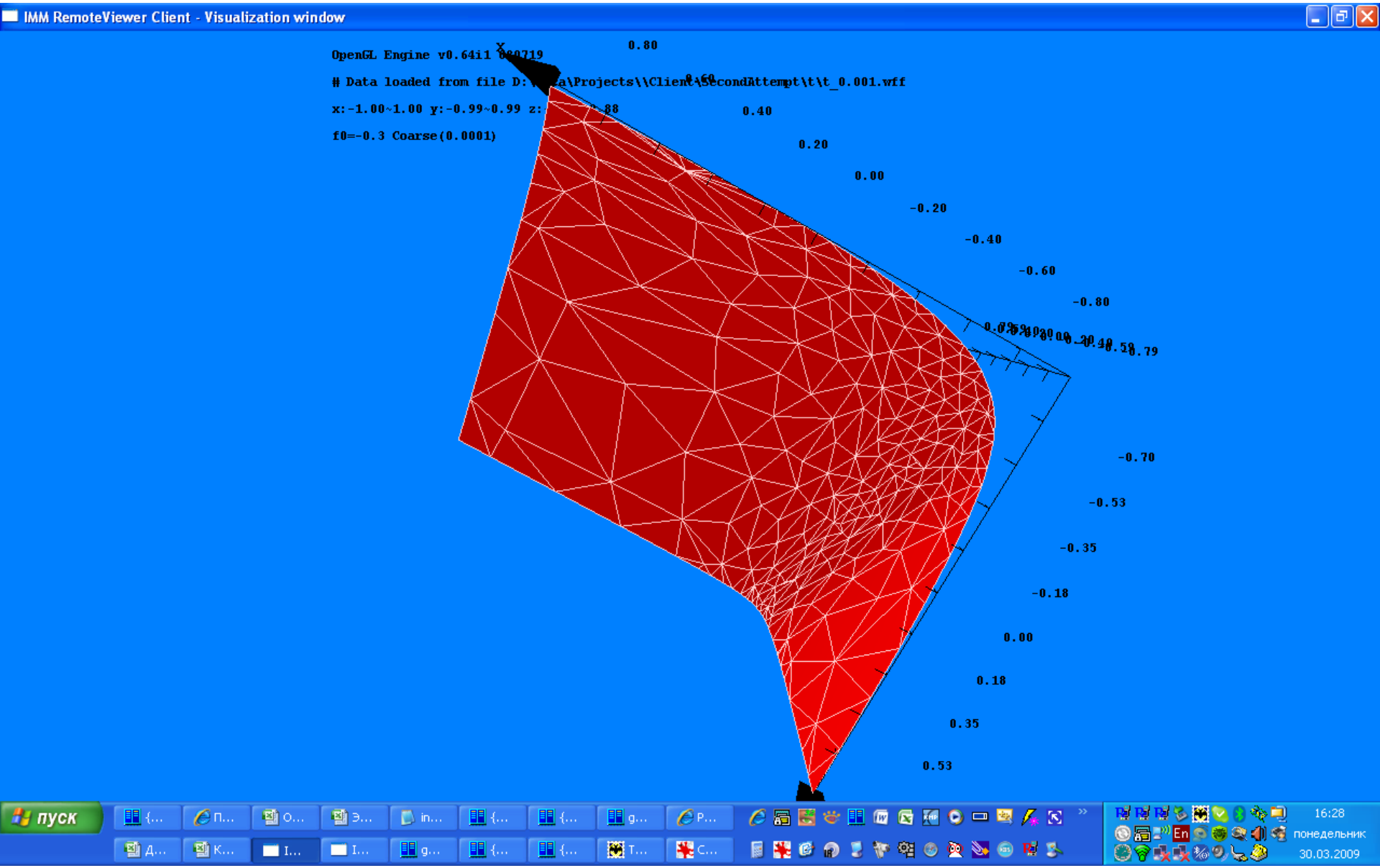
- Потоковая обработка
- Хранение образов

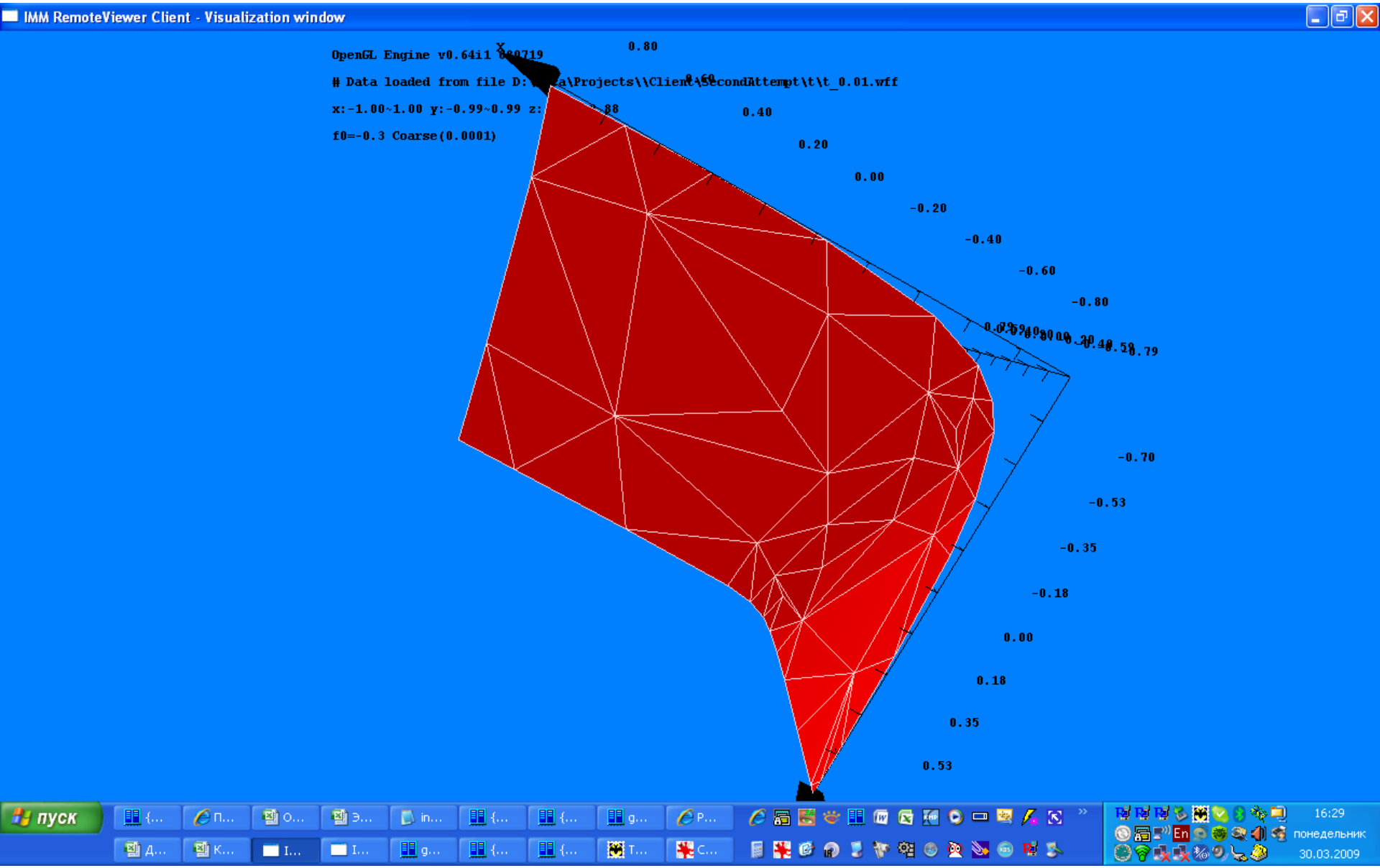
Визуализация сеточных данных

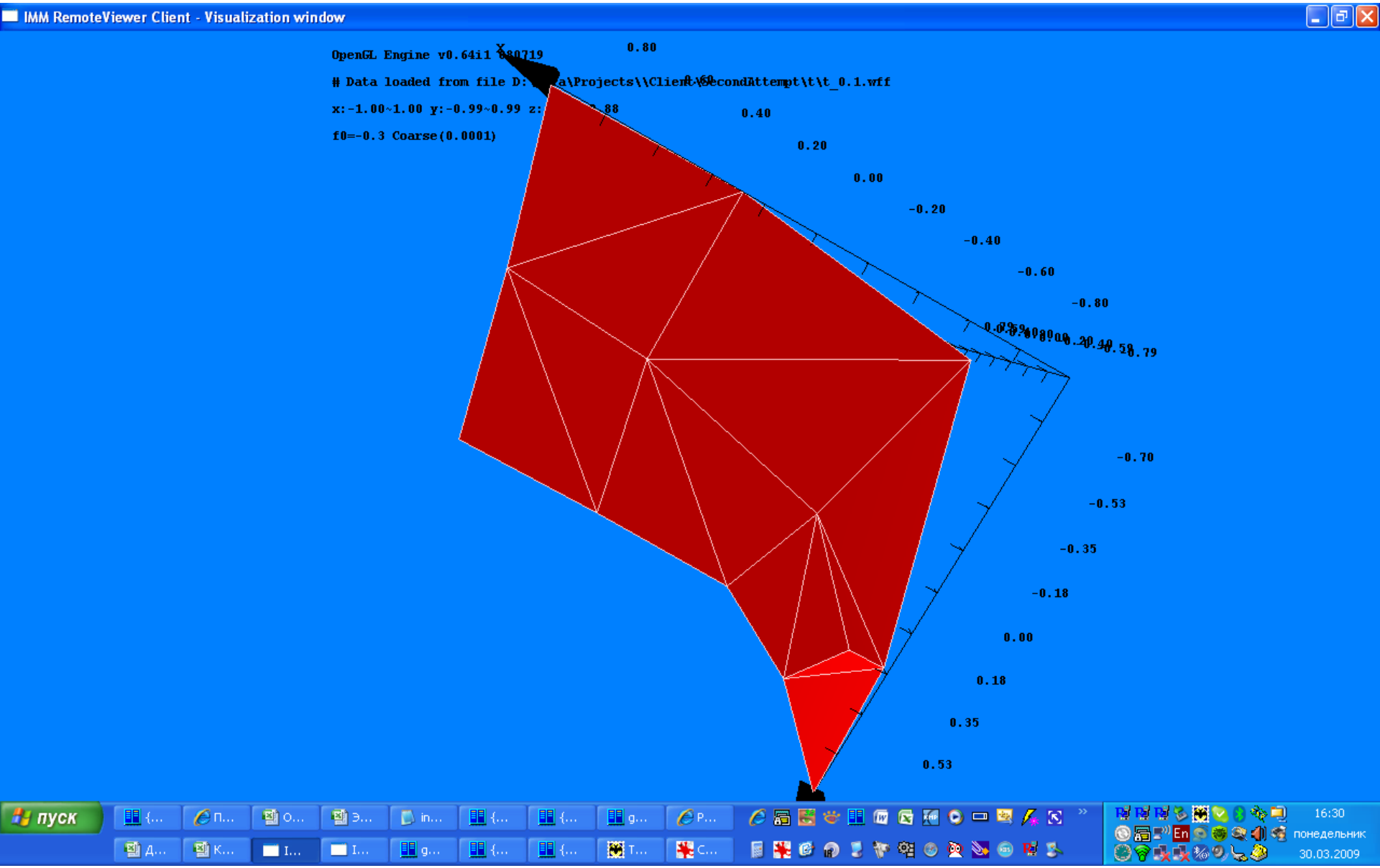


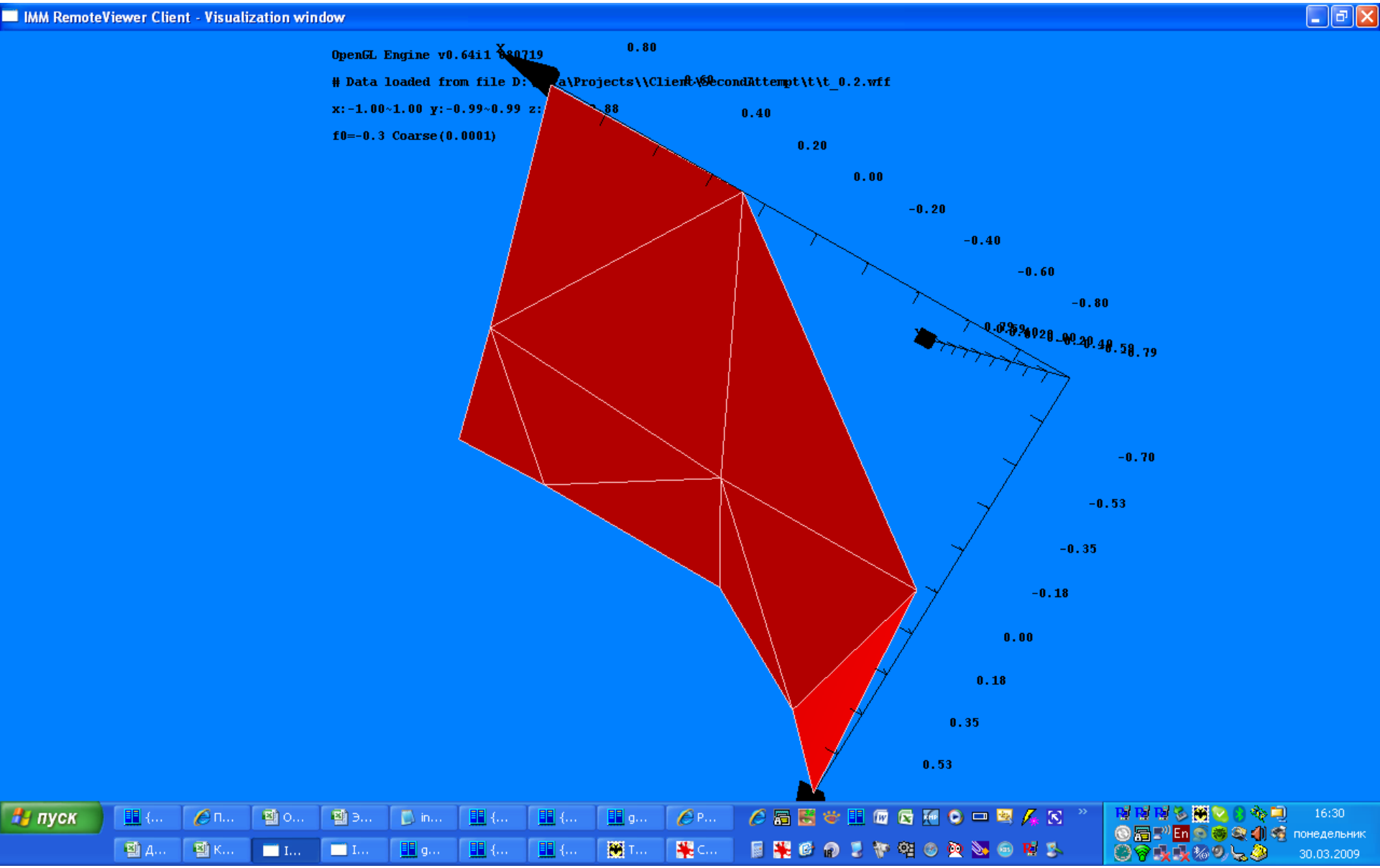

IMM RemoteViewer Client - Visualization window
OpenGL Engine v0.64i1
Data loaded from file D:
\Projects\\Client\SecondAttempt\t\t_original.wff
x:-1.00~1.00 y:-0.99~0.99 z:
f0=-0.3 Coarse(0.0001)
0.80
0.60
0.40
0.20
0.00
-0.20
-0.40
-0.60
-0.80
-0.70
-0.53
-0.35
-0.18
0.00
0.18
0.35
0.53
пуск
16:26
понедельник
30.03.2009

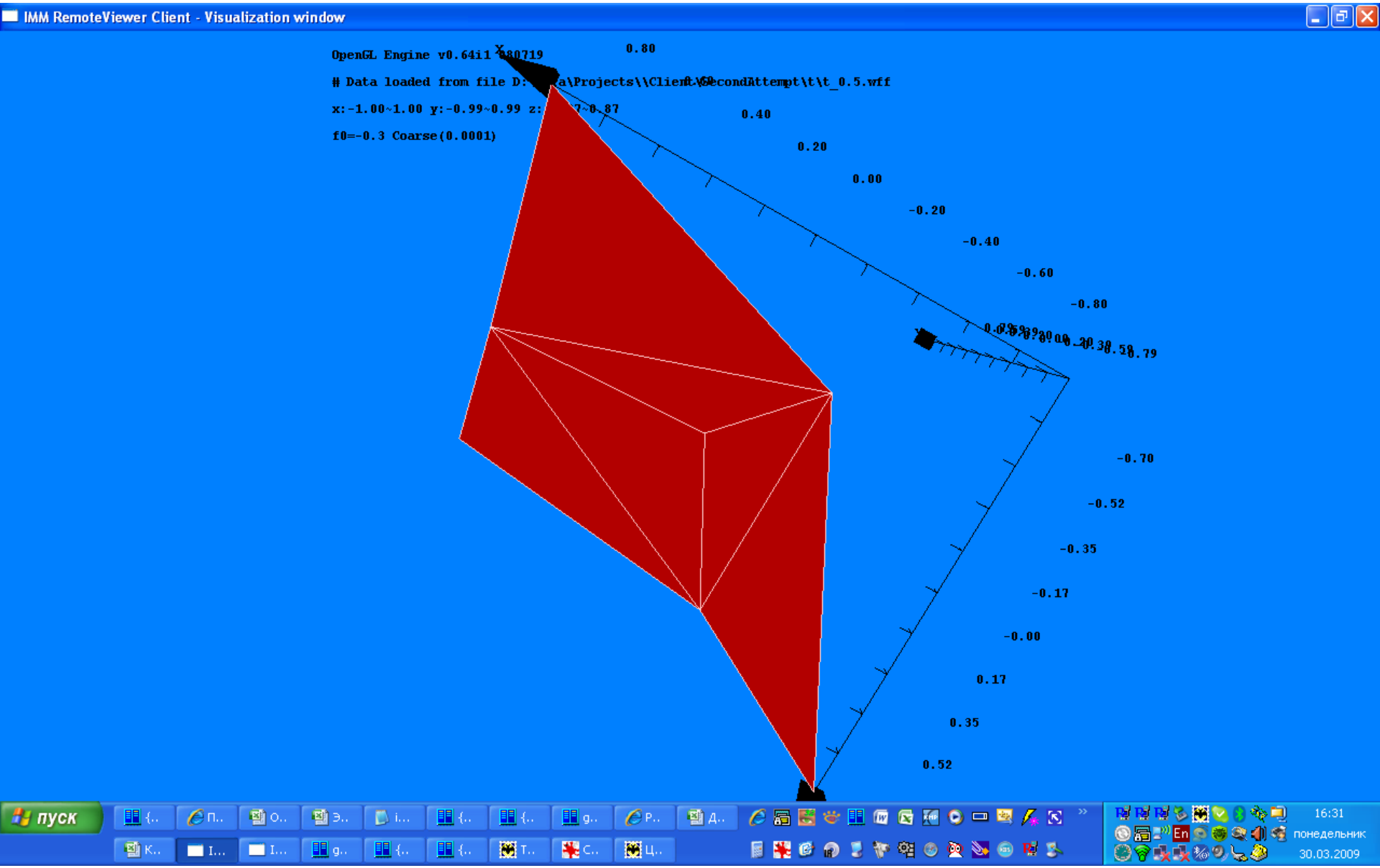

IMM RemoteViewer Client - Visualization window
OpenGL Engine v0.64i1
Data loaded from file D:
Projects\\Client\SecondAttempt\t\t_0.0001.wff
x:-1.00~1.00 y:-0.99~0.99 z:
f0=-0.3 Coarse(0.0001)
0.80
0.60
0.40
0.20
0.00
-0.20
-0.40
-0.60
-0.80
-0.70
-0.53
-0.35
-0.18
0.00
0.18
0.35
0.53
пуск
16:28
понедельник
30.03.2009

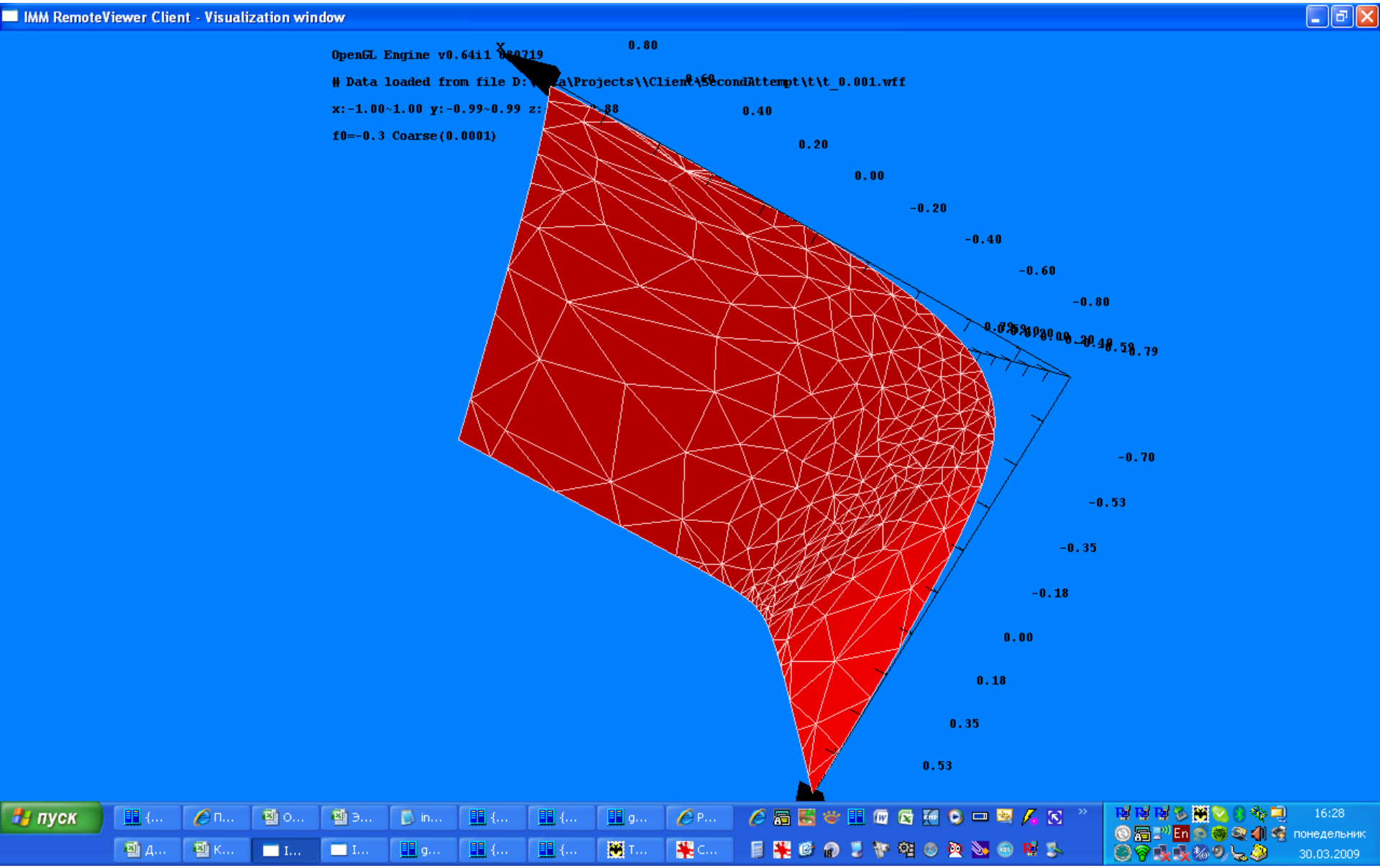

IMM RemoteViewer Client - Visualization window
OpenGL Engine v0.64i1
Data loaded from file D:
Projects\\Client\SecondAttempt\t\t_0.001.wff
x:-1.00~1.00 y:-0.99~0.99 z:
f0=-0.3 Coarse(0.0001)
X
0.80
0.60
0.40
0.20
0.00
-0.20
-0.40
-0.60
-0.80
-0.70
-0.53
-0.35
-0.18
0.00
0.18
0.35
0.53
пуск
16:28
понедельник
30.03.2009


IMM RemoteViewer Client - Visualization window
OpenGL Engine v0.64i1
Data loaded from file D:
Projects\\Client\SecondAttempt\t\t_0.01.wff
x:-1.00~1.00 y:-0.99~0.99 z:
f0=-0.3 Coarse(0.0001)
X
0.80
0.60
0.40
0.20
0.00
-0.20
-0.40
-0.60
-0.80
-0.70
-0.53
-0.35
-0.18
0.00
0.18
0.35
0.53
пуск
16:29
понедельник
30.03.2009


IMM RemoteViewer Client - Visualization window
OpenGL Engine v0.64i1
Data loaded from file D:
\Projects\\ClientSecondAttempt\t\t_0.1.wff
x:-1.00~1.00 y:-0.99~0.99 z:
f0=-0.3 Coarse(0.0001)
X
0.80
0.60
0.40
0.20
0.00
-0.20
-0.40
-0.60
-0.80
-0.70
-0.53
-0.35
-0.18
0.00
0.18
0.35
0.53
пуск
16:30
понедельник
30.03.2009


IMM RemoteViewer Client - Visualization window
OpenGL Engine v0.64i1
Data loaded from file D:
\Projects\\Client\SecondAttempt\t\t_0.2.wff
x:-1.00~1.00 y:-0.99~0.99 z:
f0=-0.3 Coarse(0.0001)
X
0.80
0.60
0.40
0.20
0.00
-0.20
-0.40
-0.60
-0.80
-0.70
-0.53
-0.35
-0.18
0.00
0.18
0.35
0.53
пуск
16:30
понедельник
30.03.2009


IMM RemoteViewer Client - Visualization window
OpenGL Engine v0.64i1 080719
Data loaded from file D:
a\Projects\\Client\SecondAttempt\t\t_0.5.wff
x:-1.00~1.00 y:-0.99~0.99 z:
7~0.87
f0=-0.3 Coarse(0.0001)
X
0.80
0.40
0.20
0.00
-0.20
-0.40
-0.60
-0.80
-0.70
-0.52
-0.35
-0.17
-0.00
0.17
0.35
0.52
пуск
16:31
понедельник
30.03.2009


IMM RemoteViewer Client - Visualization window
OpenGL Engine v0.64i1
Data loaded from file D:
Projects\\Client\SecondAttempt\t\t_0.001.wff
x:-1.00~1.00 y:-0.99~0.99 z:
f0=-0.3 Coarse(0.0001)
X
0.80
0.60
0.40
0.20
0.00
-0.20
-0.40
-0.60
-0.80
-0.70
-0.53
-0.35
-0.18
0.00
0.18
0.35
0.53
0.79
пуск
16:28
понедельник
30.03.2009

- Отображение

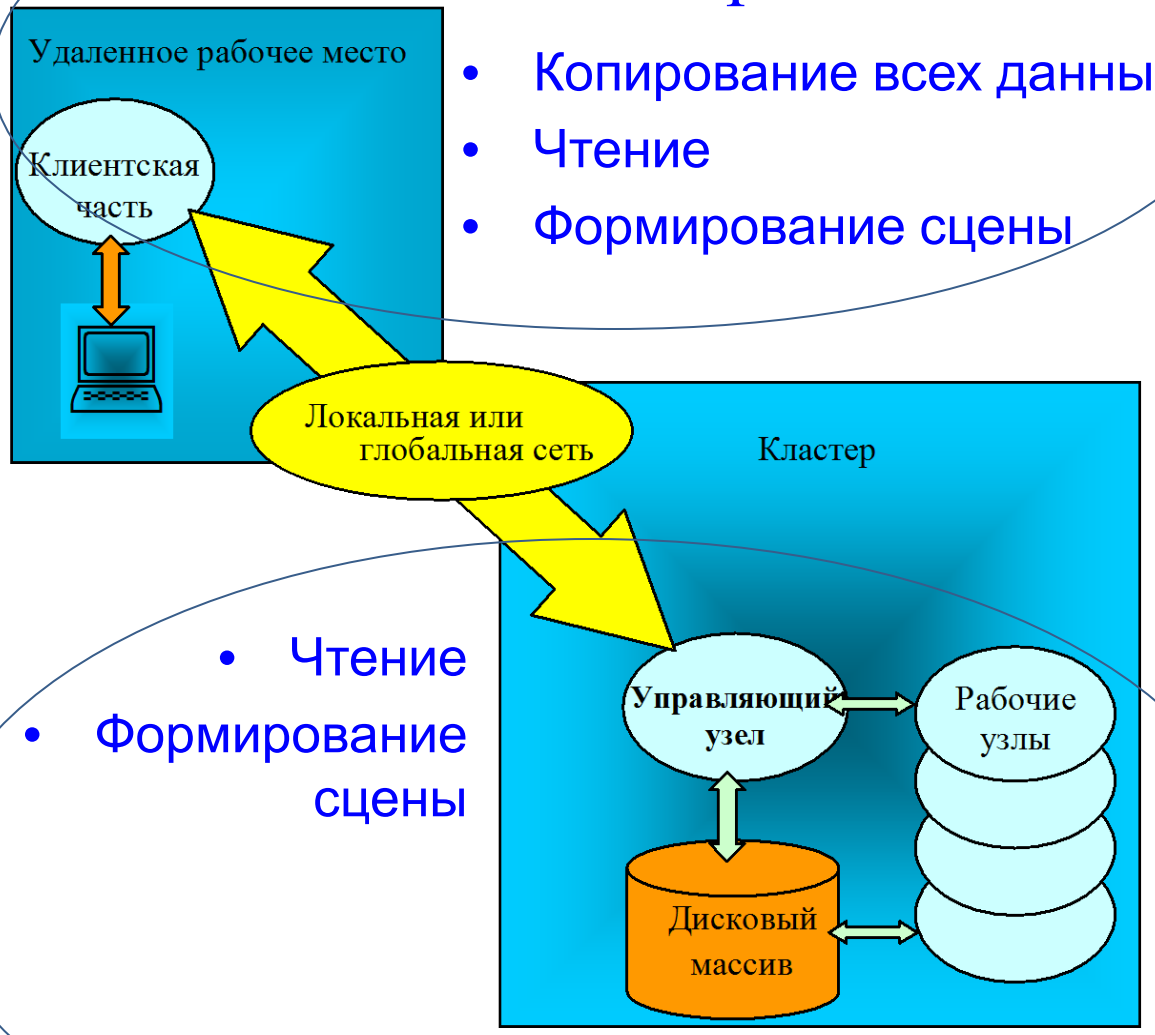
- Расчет
- Запись результатов

TecPlot
Origin

VISIT ParaView
EnSight OpenDX

Аппроксимация и потоковая обработка

- Копирование всех данных
- Чтение
- Формирование сцены

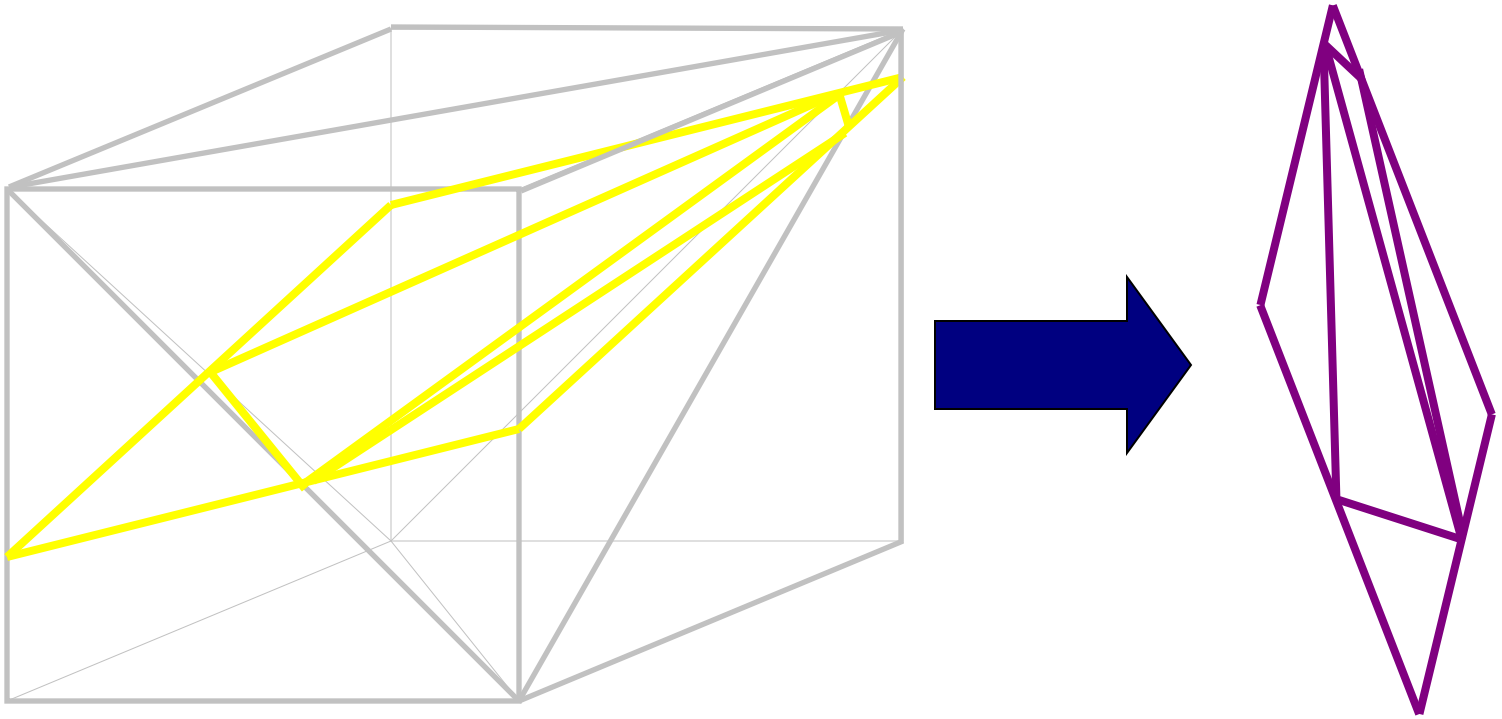


- Чтение
- Формирование сцены

Клиент-серверная технология

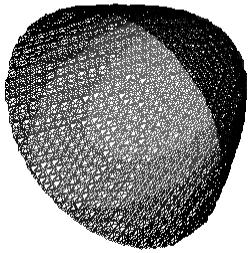
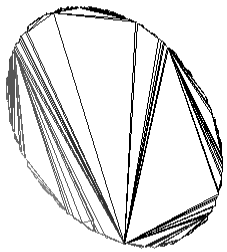
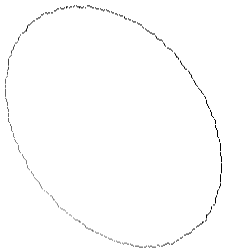
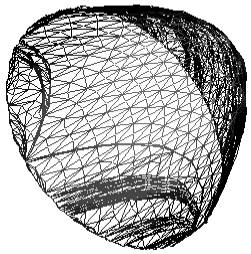
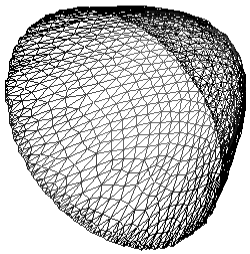
Сечение регулярной 3D сетки плоскостью

- В результате сечения регулярной кубической решетки получается фрагмент неструктурированной сетки



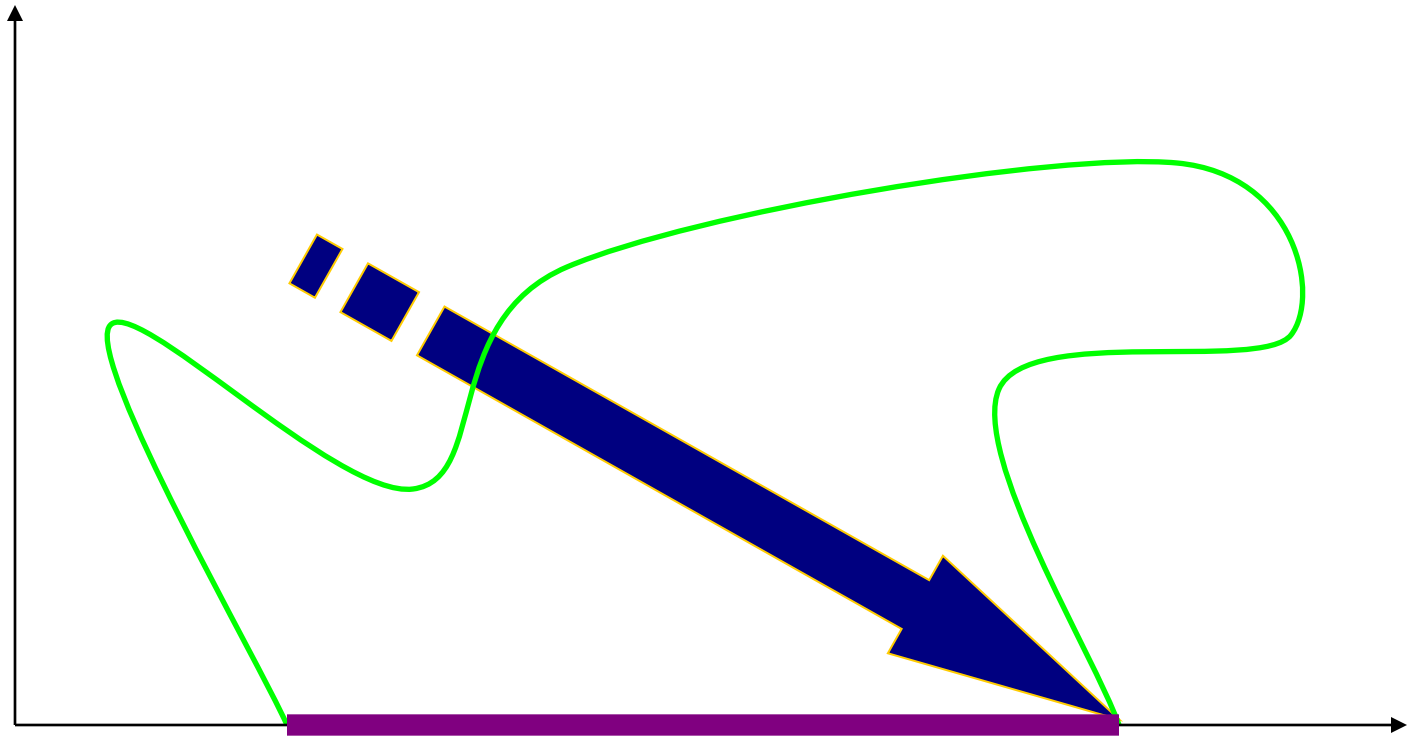
Аппроксимация триангулированных поверхностей

- Алгоритмы синтеза

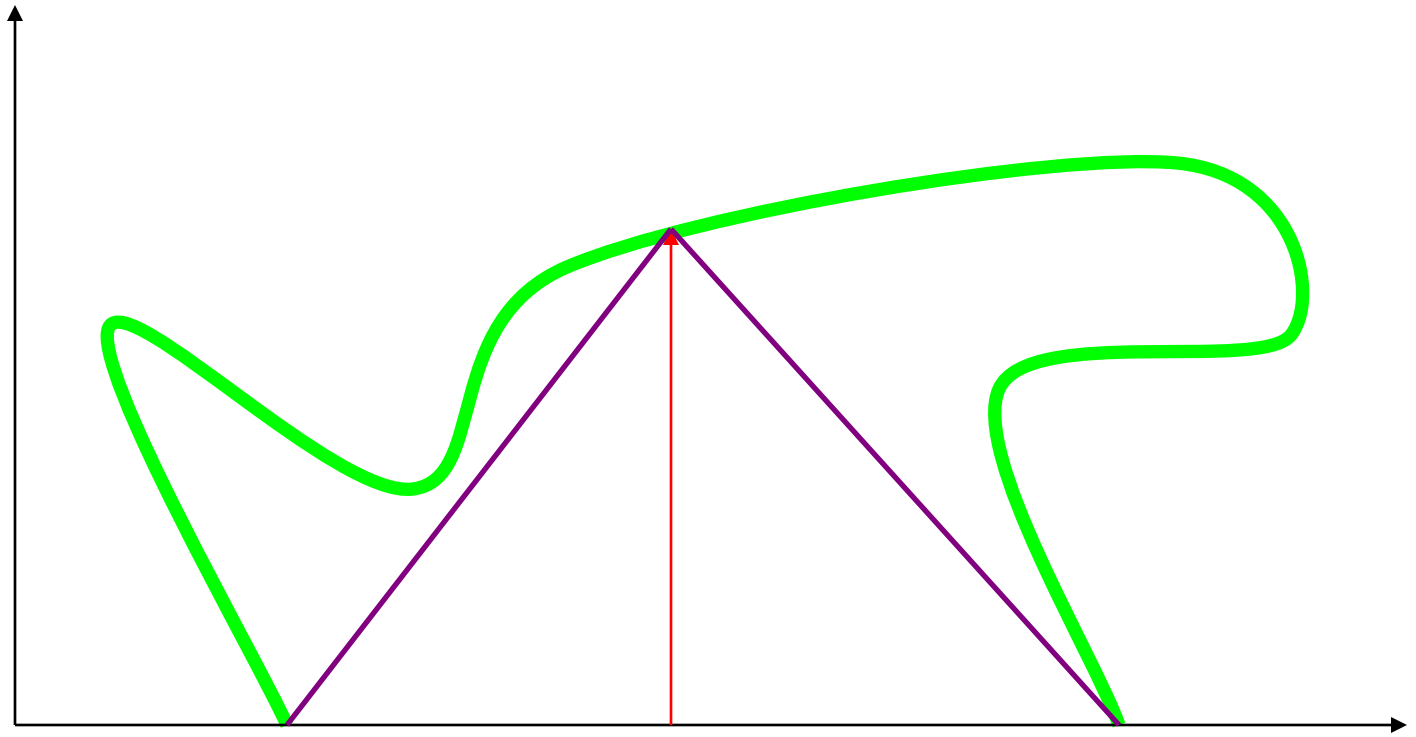






- Алгоритмы редуцирования

Начальная аппроксимация кривой

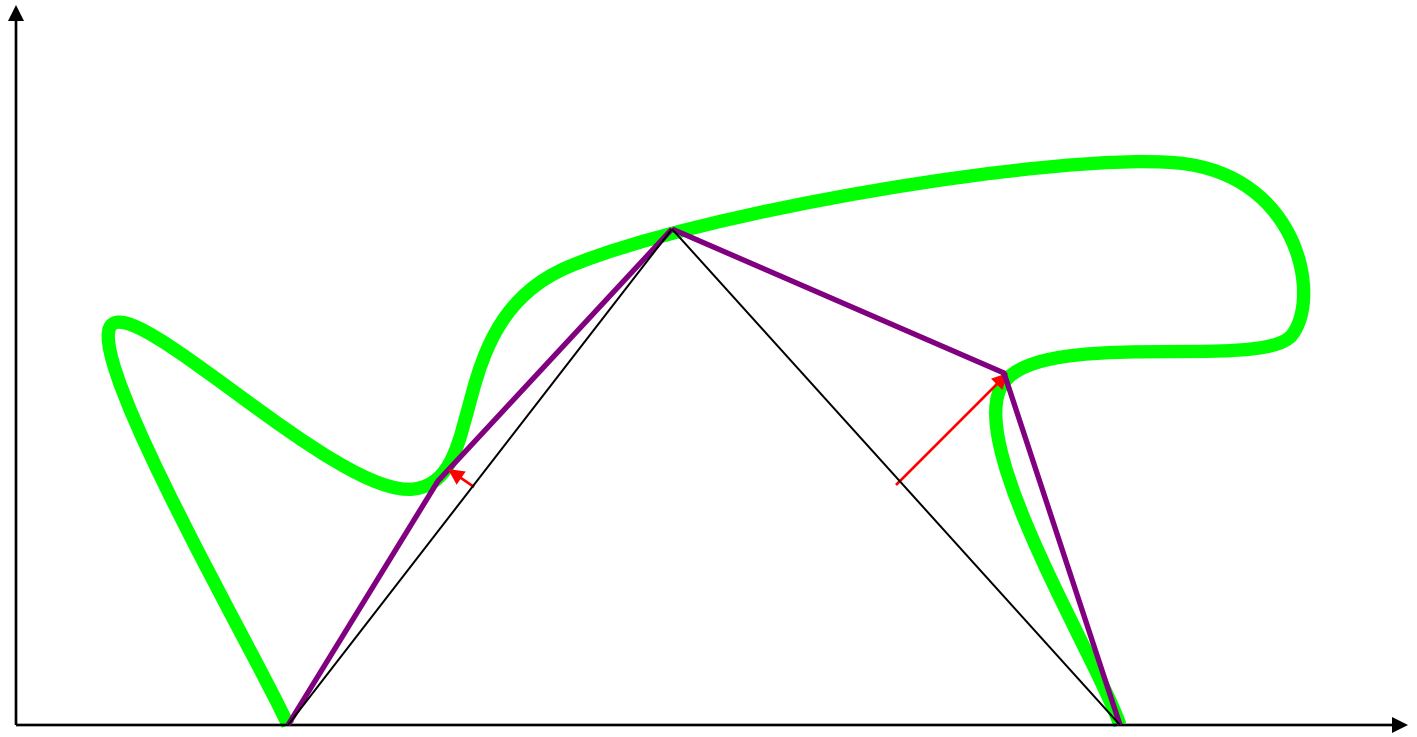


Аппроксимация кривой
этап 2



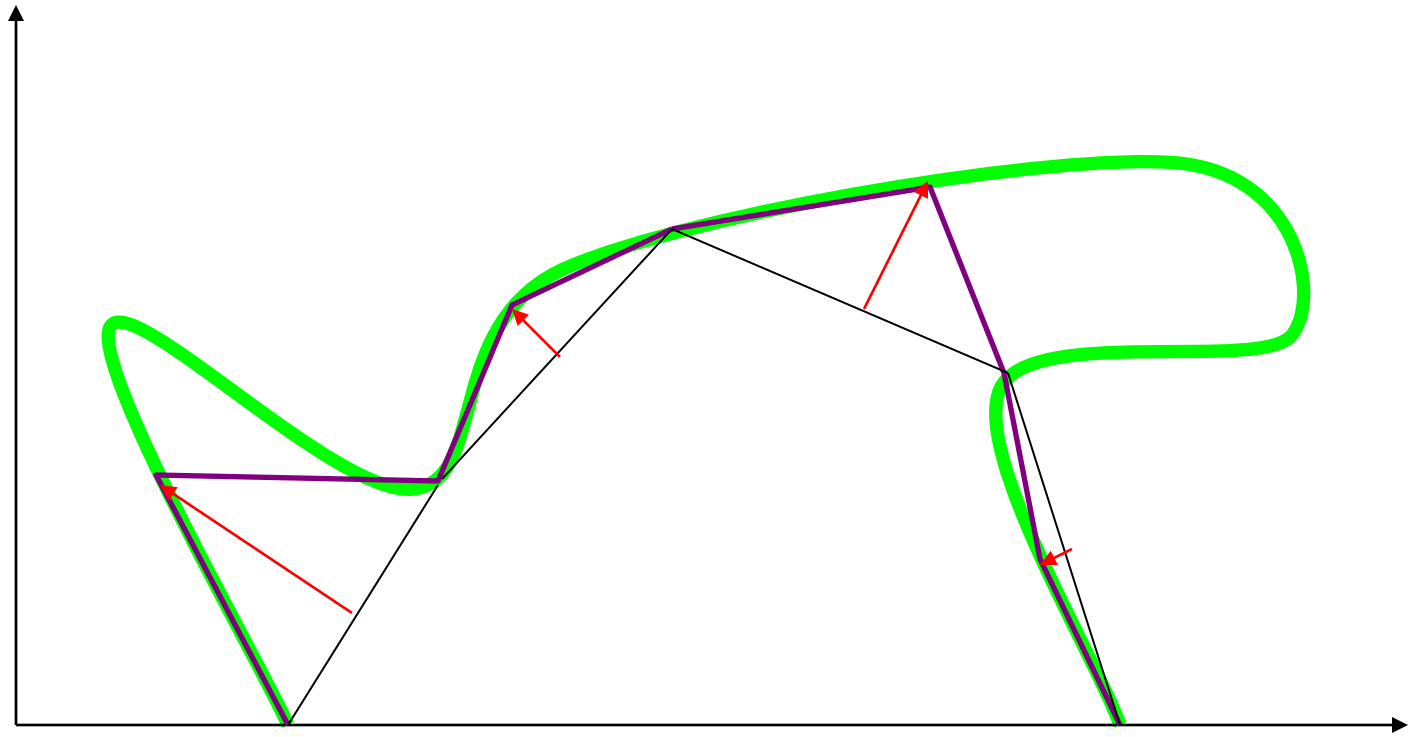
1 вектор

Аппроксимация кривой
этап 3



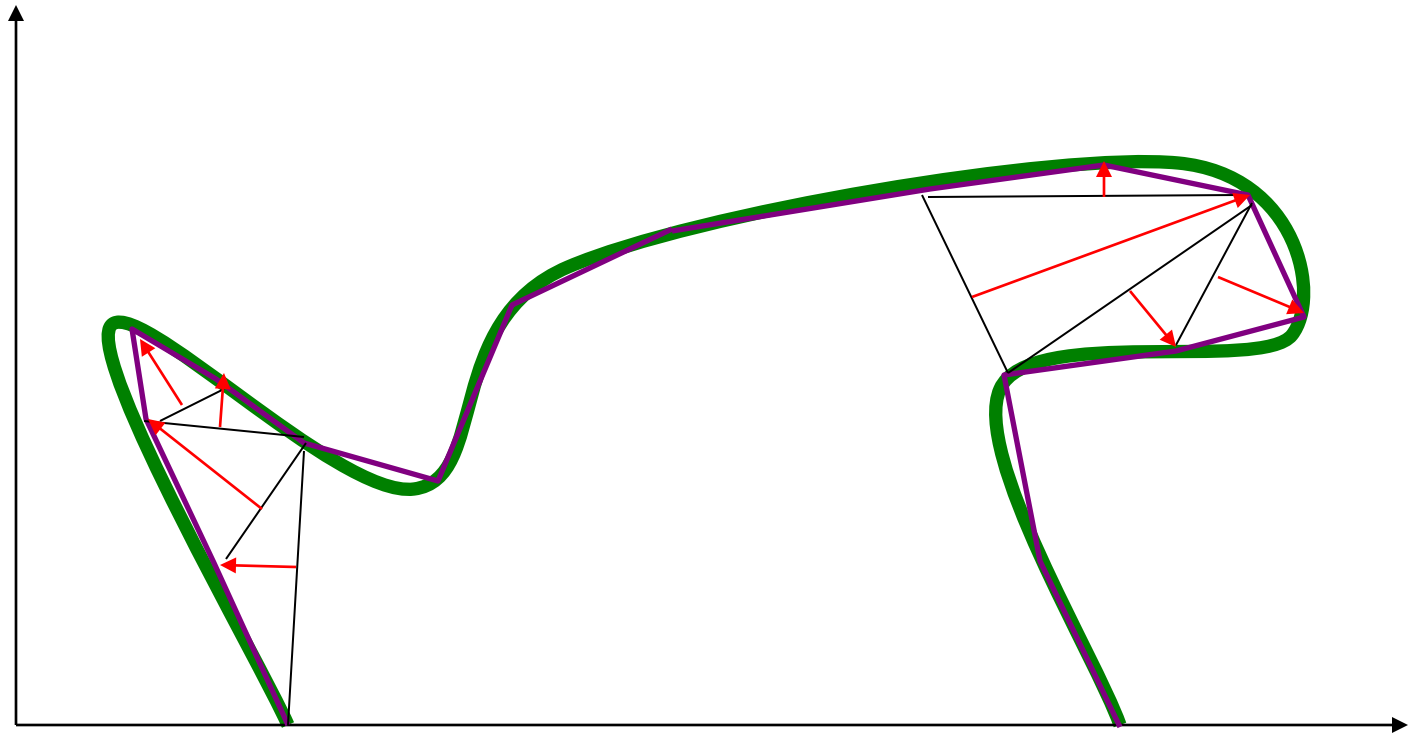
3 вектора

Аппроксимация кривой
этап 4



7 векторов

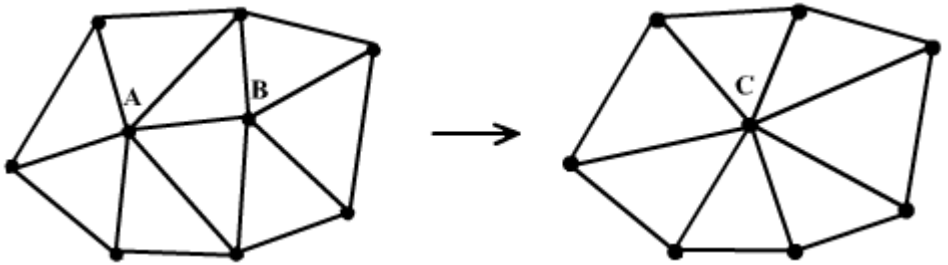
Аппроксимация кривой
этап 5



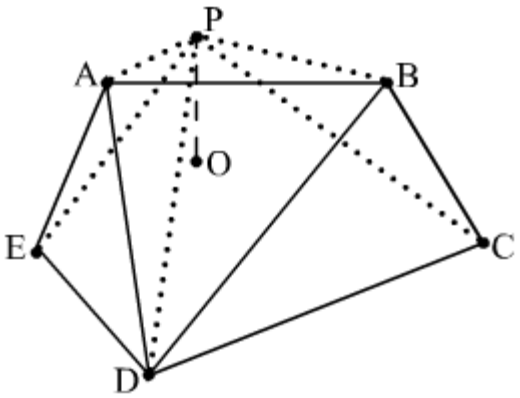
15 векторов

Методы редуцирования

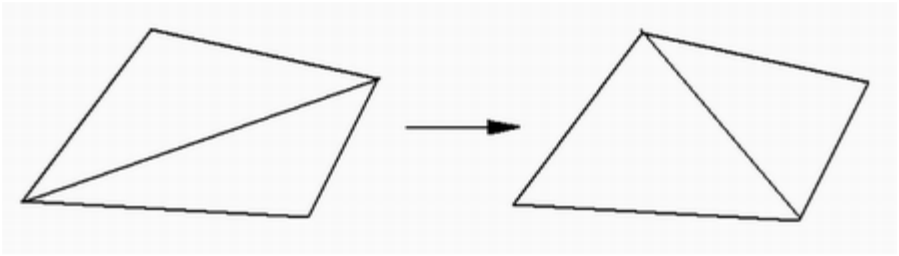
Удаление ребра



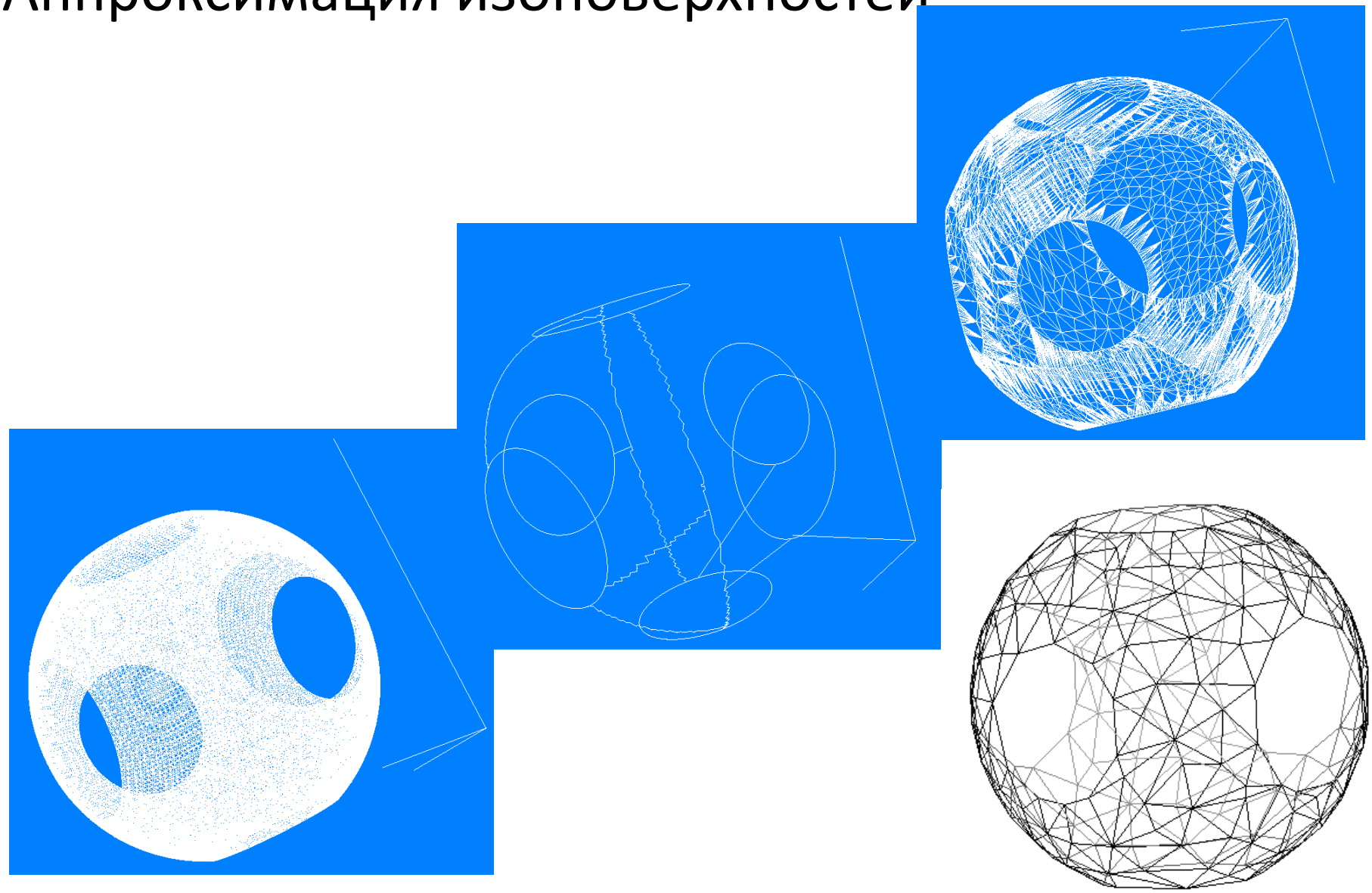
Удаление точки



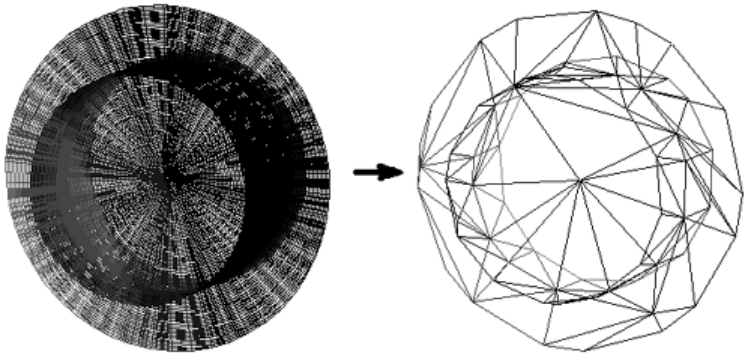
Уточнение топологии



Аппроксимация изоповерхностей

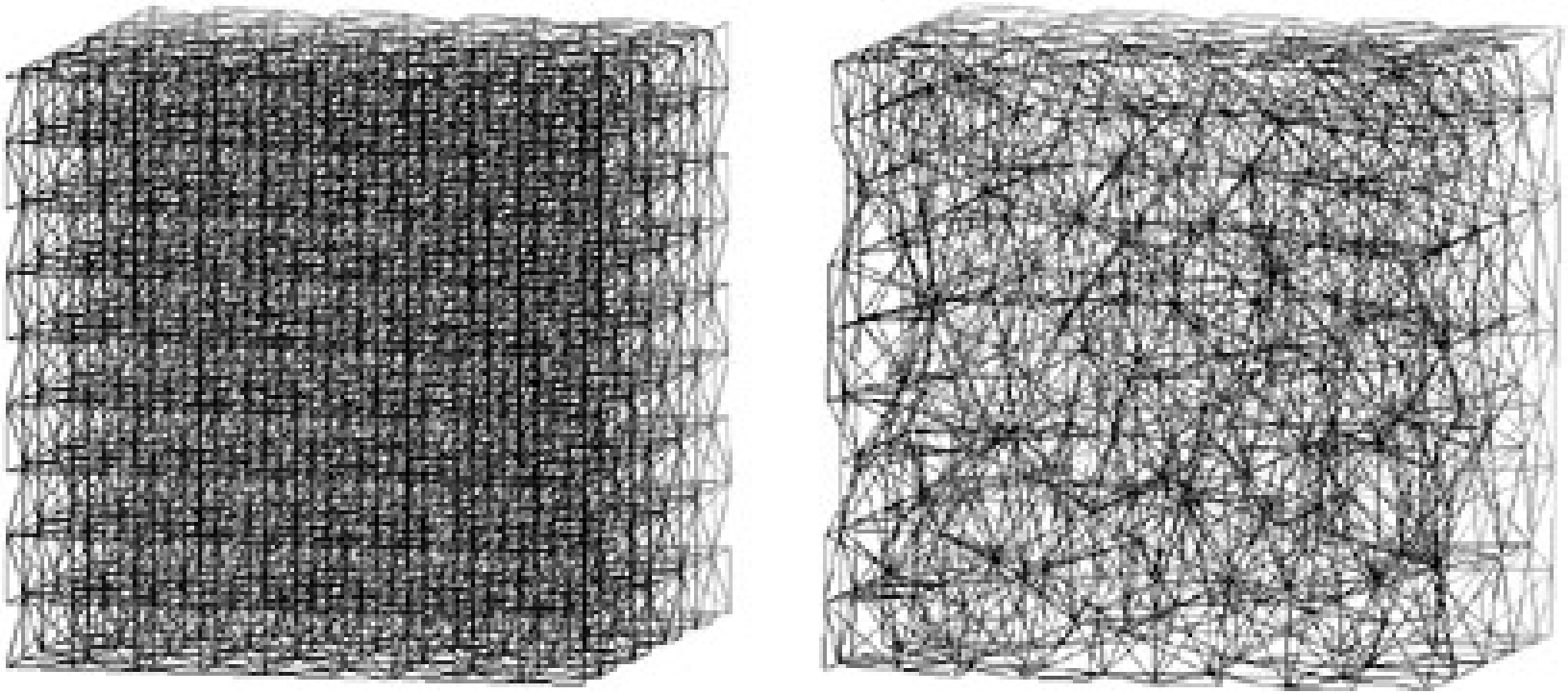


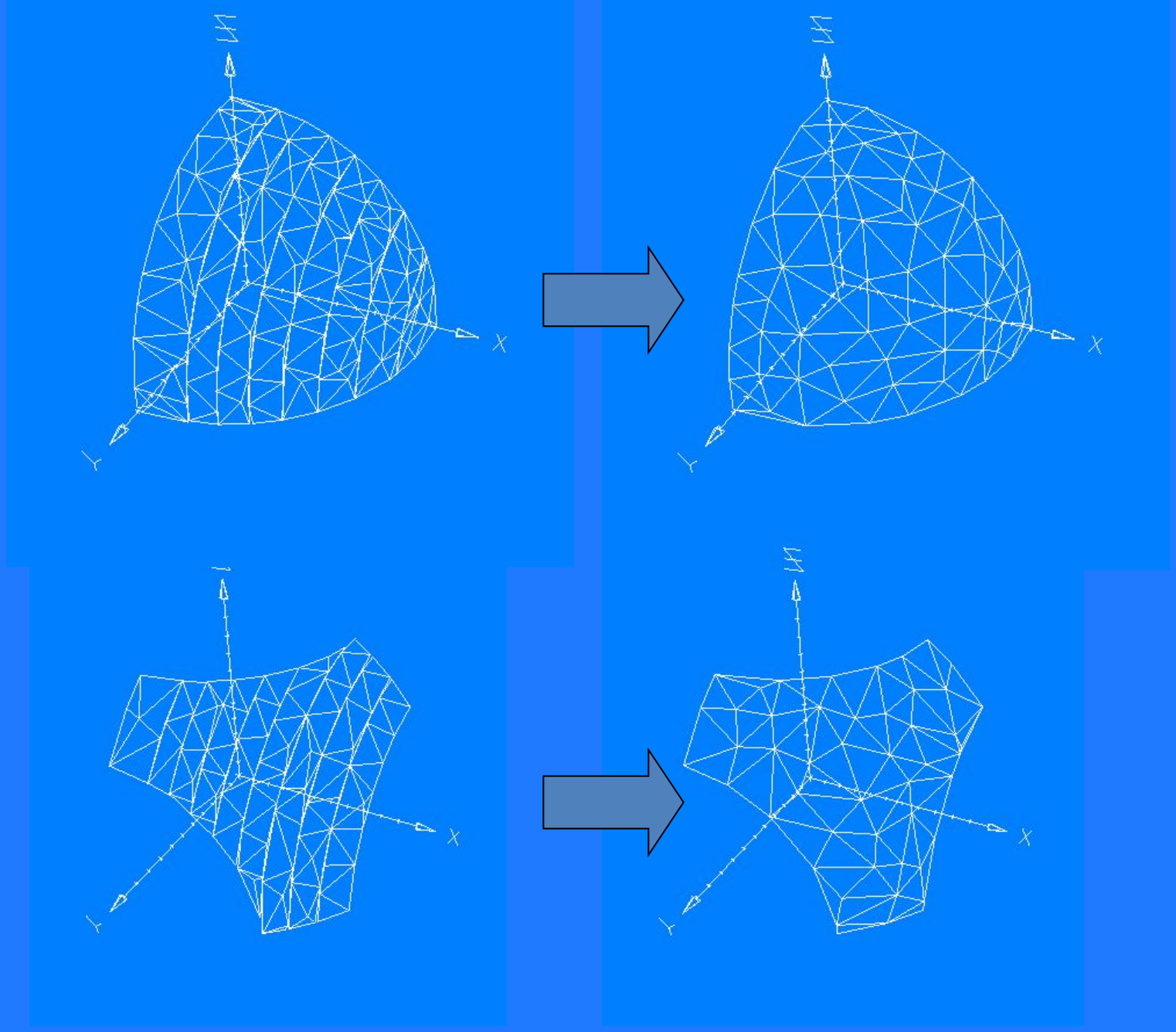
Плоскость, пересекающая цилиндр



Ошибка аппроксимации 5%

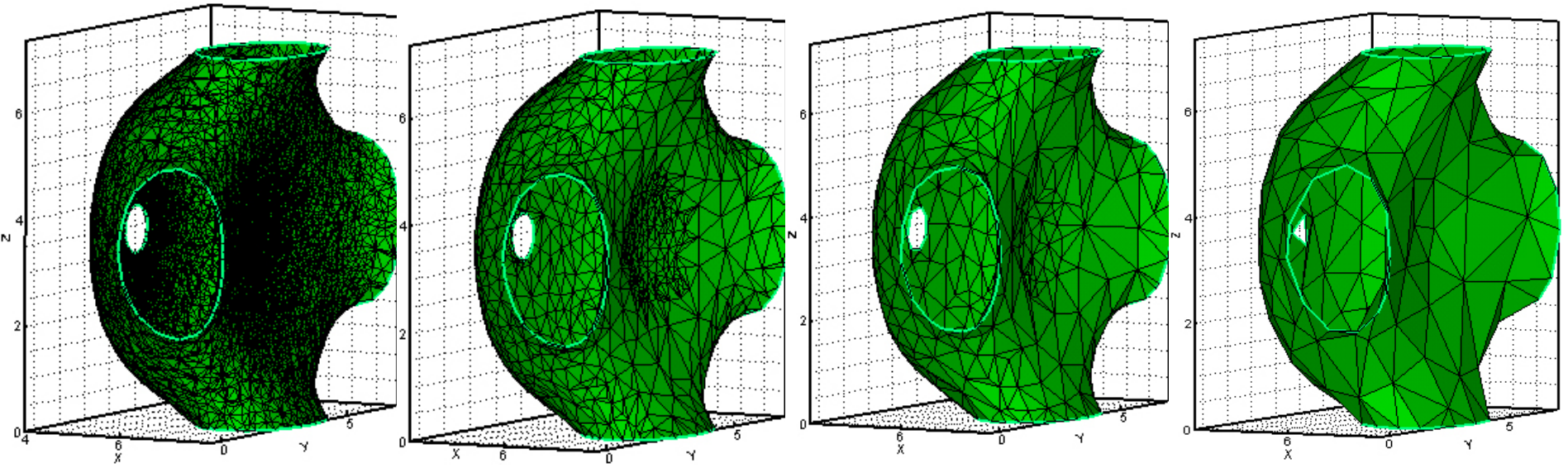
“Изоповерхность”





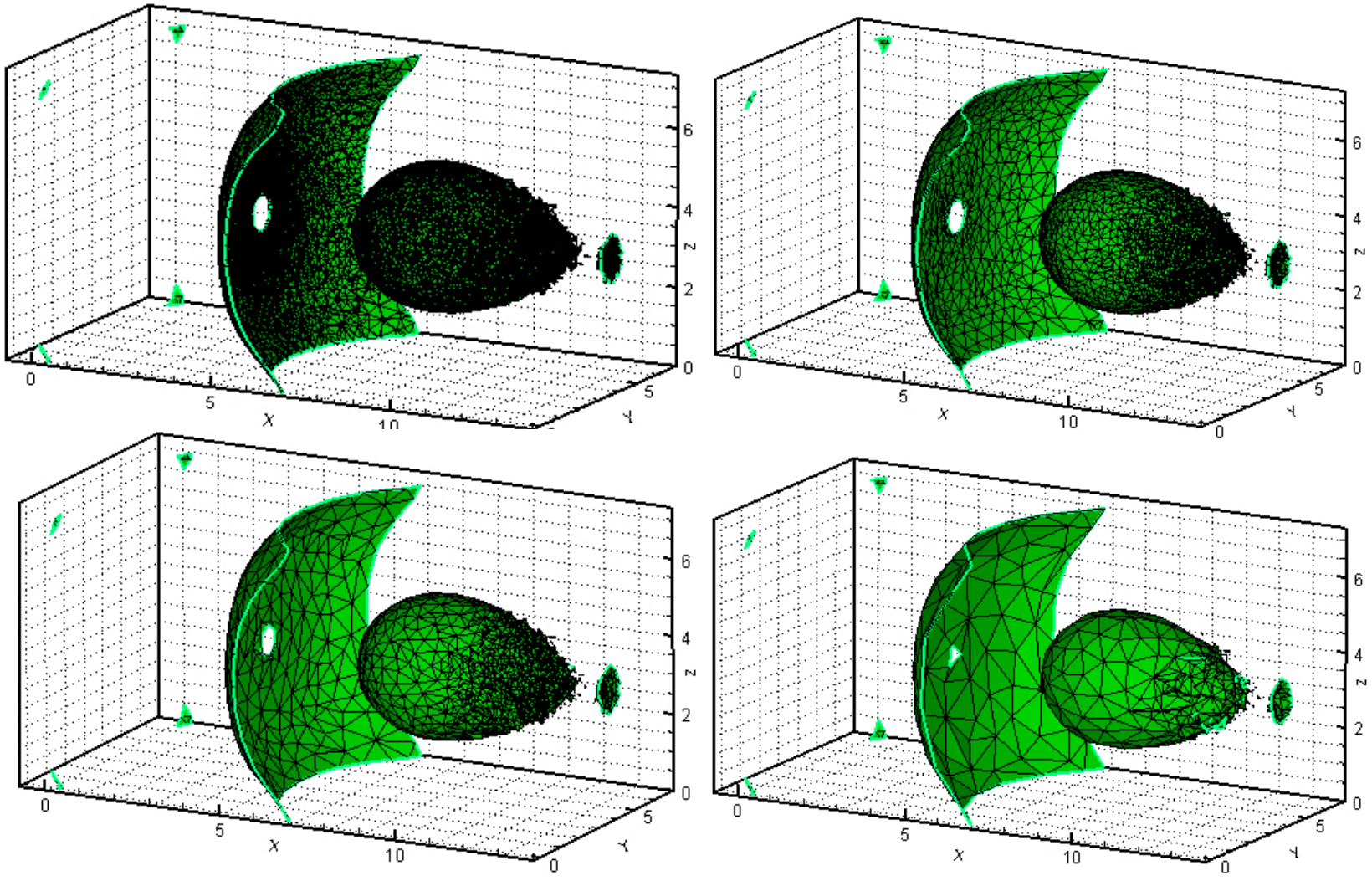
Многоуровневое огрубление больших сеток

Огрубление поверхностей



Ошибка	Количество точек	Количество треугольников	Коэффициент сжатия
0%	13800	27357	-
0,1%	1120	2117	12,9
0,2%	447	808	33,9
0,5%	175	304	90,0

Огрубление поверхностей

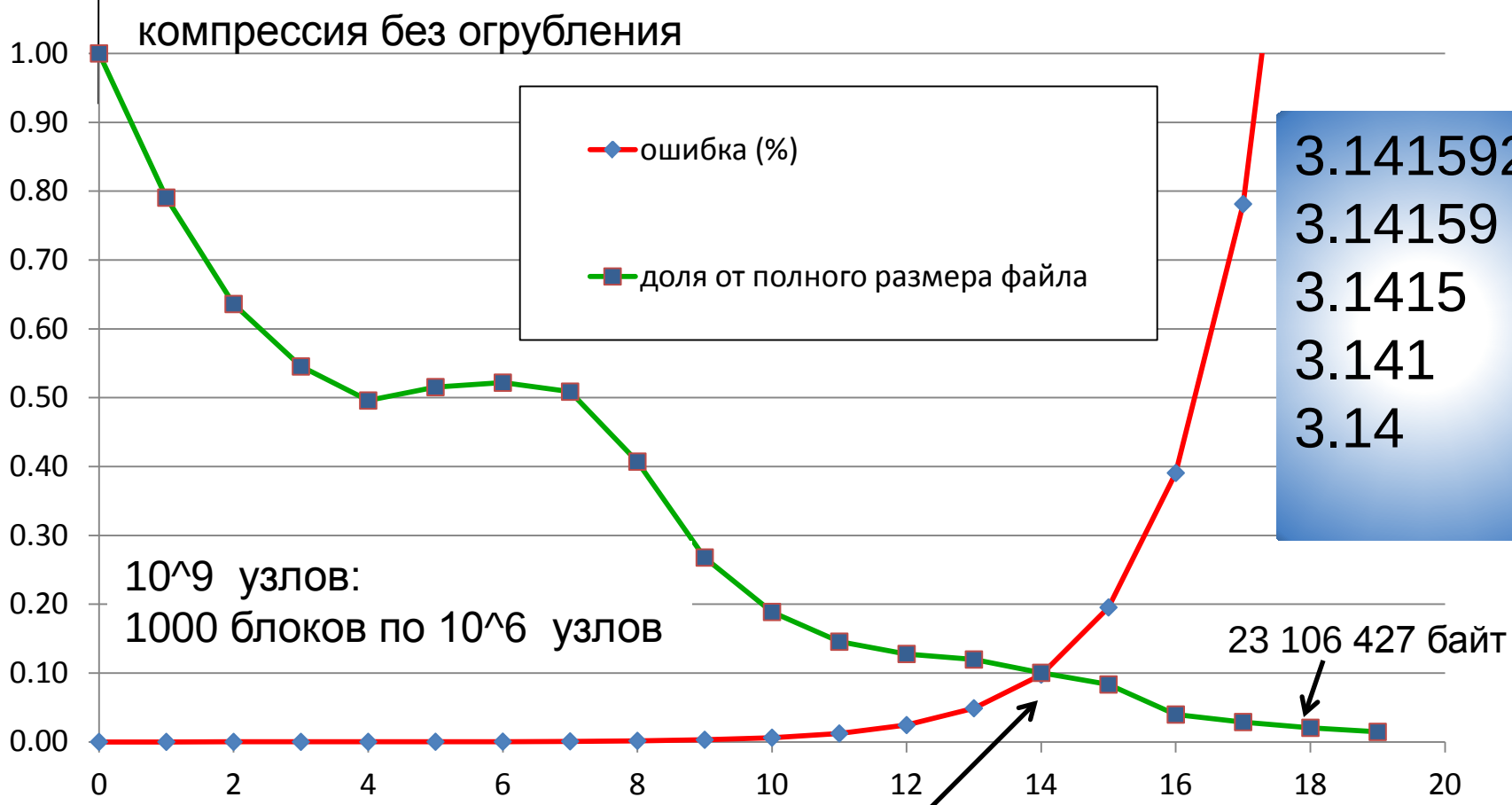


Хранение сеточных данных

Отсечение младших бит мантиссы

3.54 бинарный без компрессии без огрубления

$f=x^2+y^2+z^2$

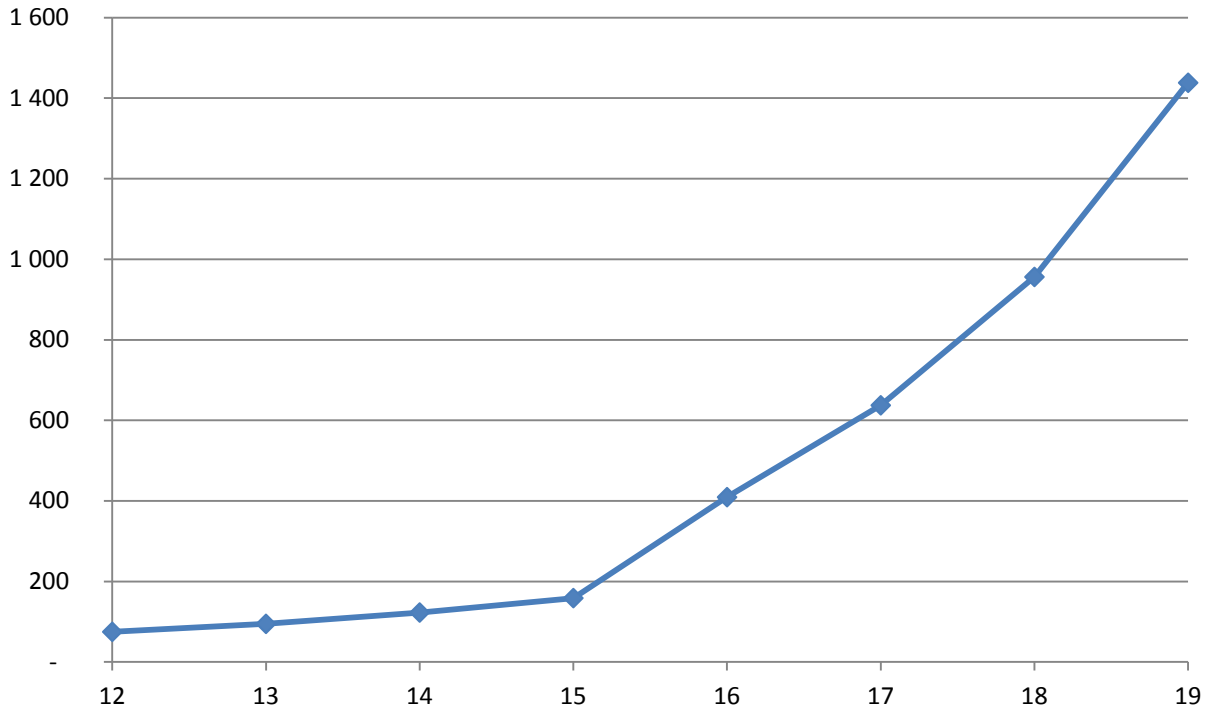


3.141592
3.14159
3.1415
3.141
3.14

10^9 узлов - 113 354 035 байт – *0.1% - 0.92 бита на узел*

Зависимость коэффициента сжатия от числа усеченных бит

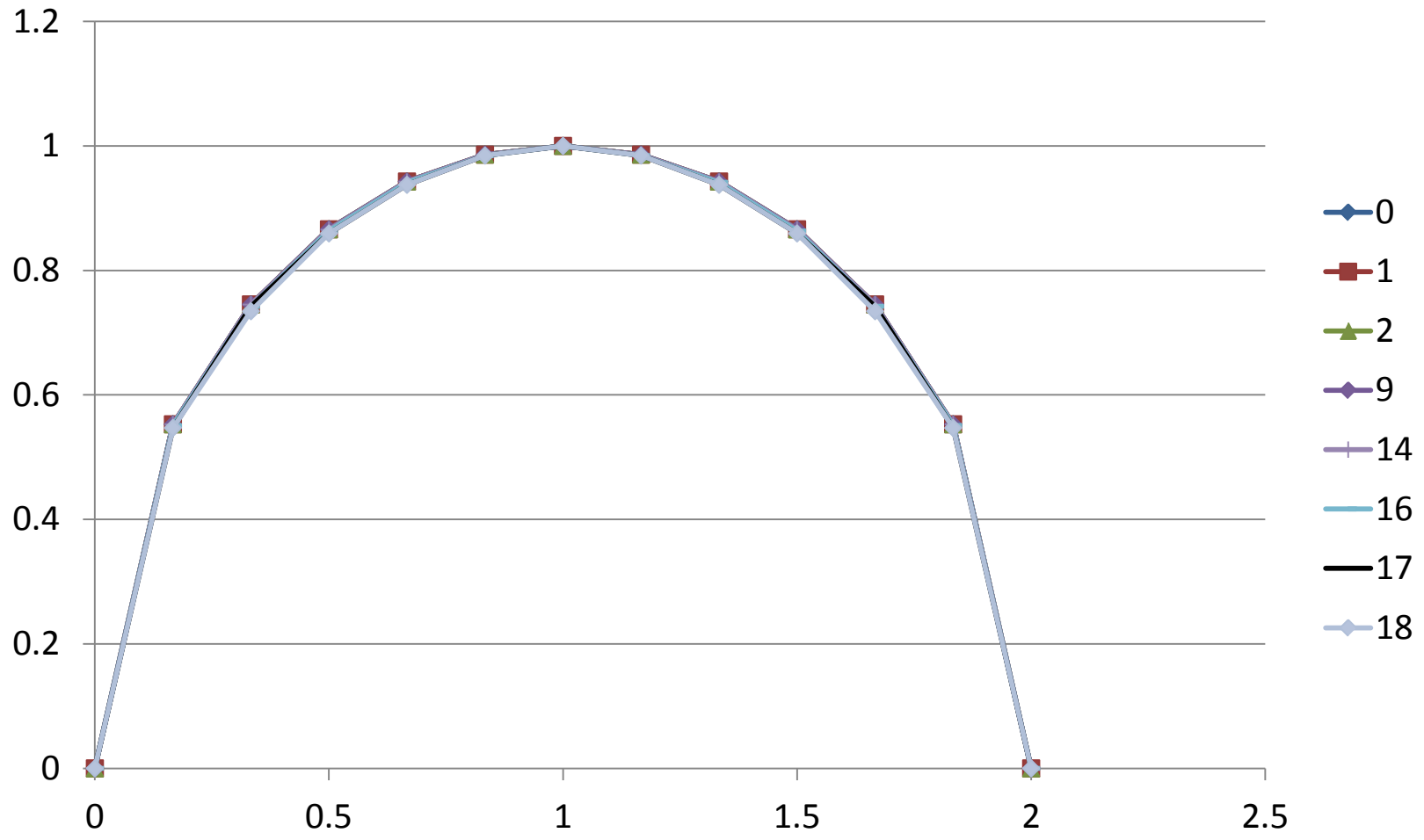
Сетка: 1000 x 3500 x 150 = 525 млн узлов




```
28 244 379   w101_reduced 12.bjn
22 340 718   w101_reduced 13.bjn
17 228 023   w101_reduced 14.bjn
13 339 249   w101_reduced 15.bjn
 5 171 208   w101_reduced 16.bjn
 3 321 150   w101_reduced 17.bjn
 2 213 949   w101_reduced 18.bjn
 1 471 818   w101_reduced 19.bjn

   793 457   w101grid.bjn
```

Огрубление данных



Якобовский М.В.

д.ф.-м.н.,

зав. сектором

«Программного обеспечения многопроцессорных систем и вычислительных сетей»

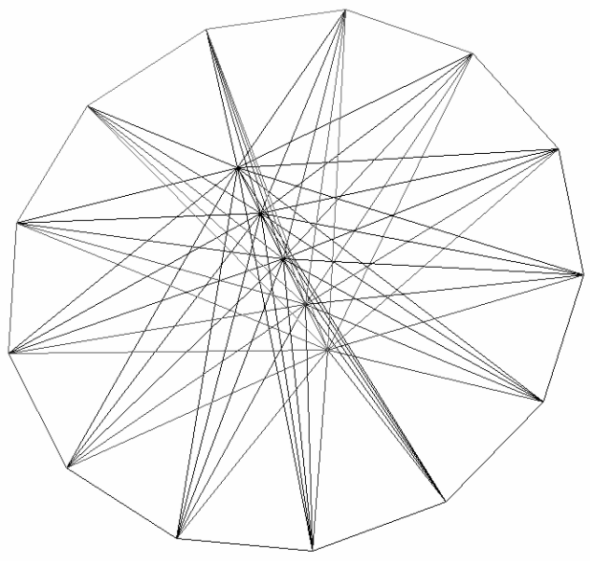
Института прикладной математики им. М.В.Келдыша Российской академии наук

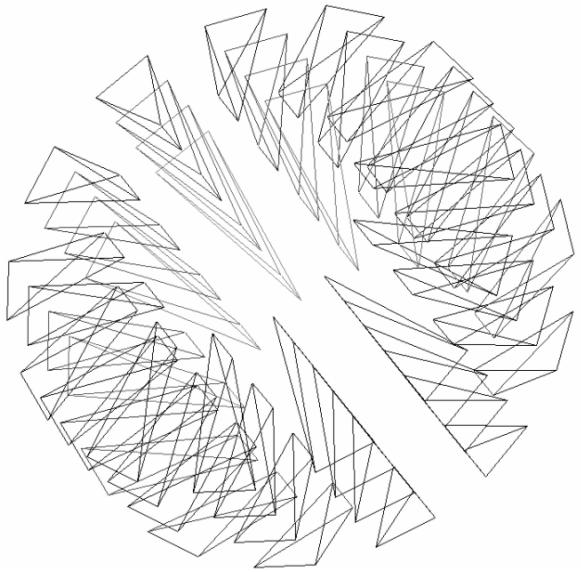
mail: lira@imamod.ru

http://lira.imamod.ru

Заполнение пространства пирамидами

- На каждую из 2n точек в среднем опирается 2n пирамид
- Число пирамид ~ n^2



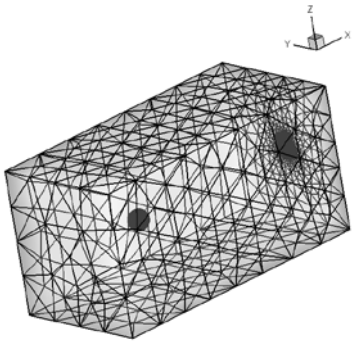


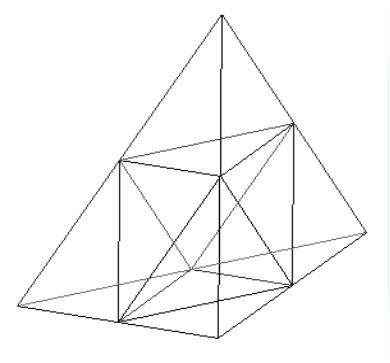
Зависимость объема хранимых данных от числа микродоменов

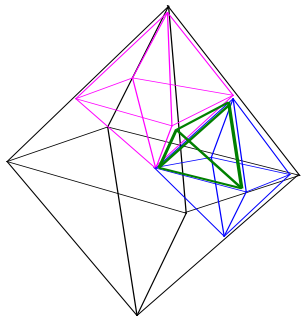
Число микродоменов	1	50	1000	1500	2000	2500	3000
Размер описания (МБ)	124	127	145	152	158	163	168

38 350 -> 2 356 196 узлов

219 034 * 8^2 -> 14 018 176 тетраэдров







На 35% больше чем 124