Keldysh Institute of Applied Mathematics (KIAM) RAS, Moscow, Russia

Algorithms in the parallel partitioning tool GridSpiderPar for large mesh decomposition

Evdokia N. Golovchenko,

Marina A. Kornilina,

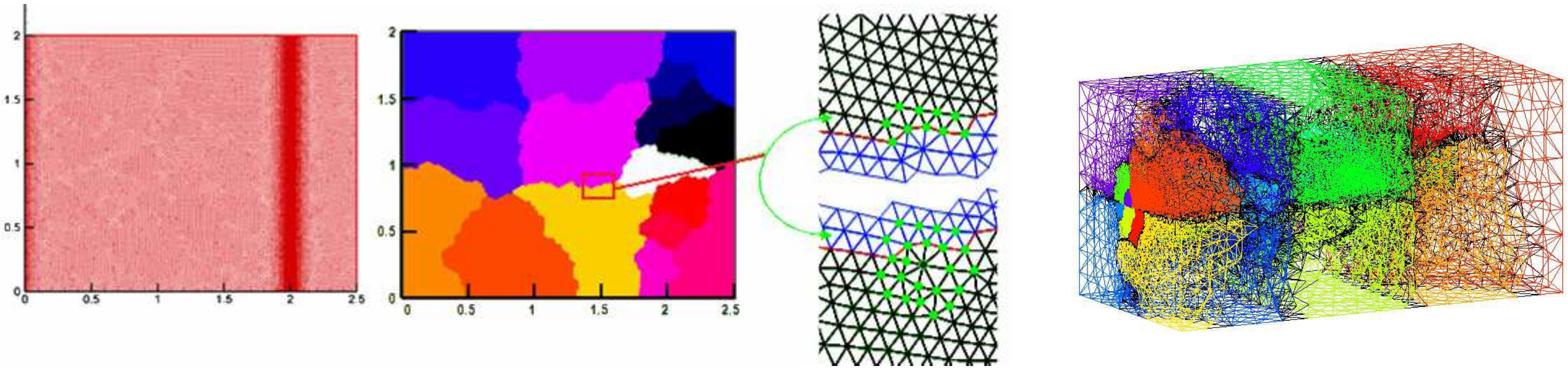
Mikhail V. Yakobovskiy

Decomposition

- **parallel mesh-based numerical simulations in continuum mechanics, electrodynamics and other PDE's problems on distributed memory systems**

↓

Geometric parallelism



Efficient processors usage

balanced mesh distribution among processors

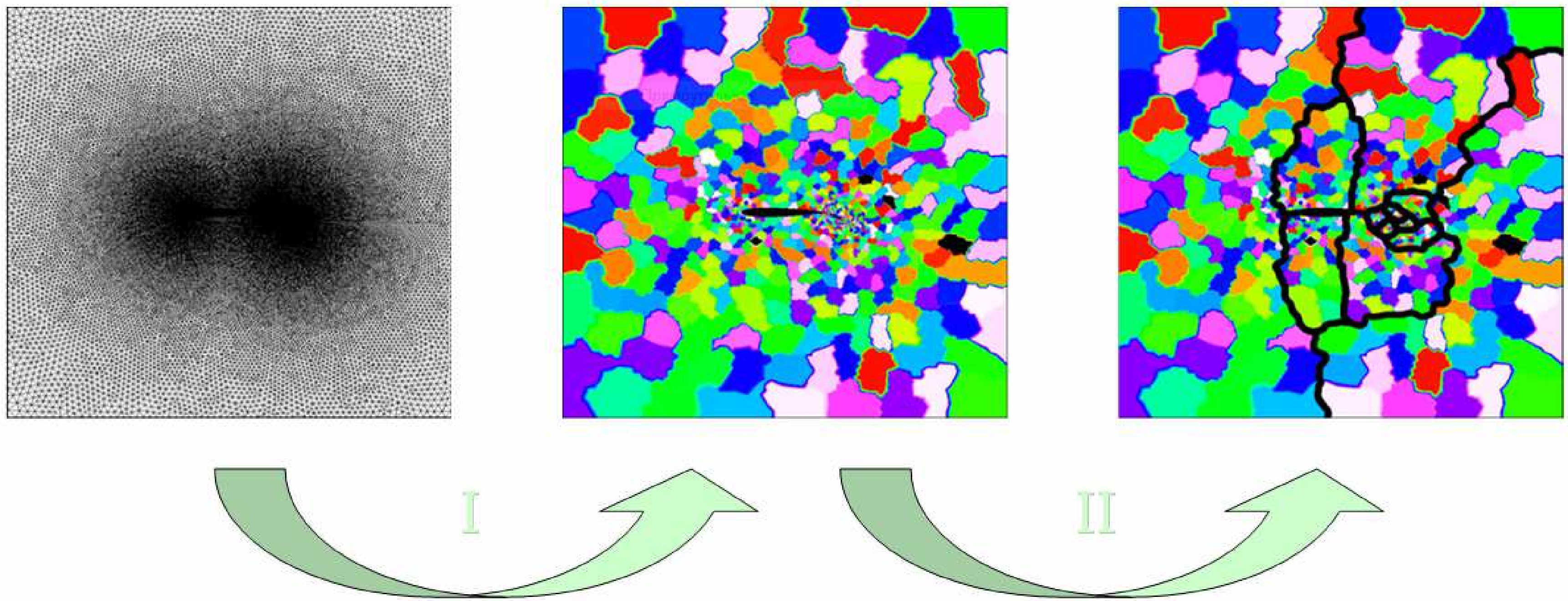
reducing interprocessor communications

Use of partitions into microdomains

forming of subdomains from microdomains

domain decomposition methods (Schwarz method)

large mesh storage



Serial partitioning tools

METIS, Jostle, Scotch, Chaco, Party

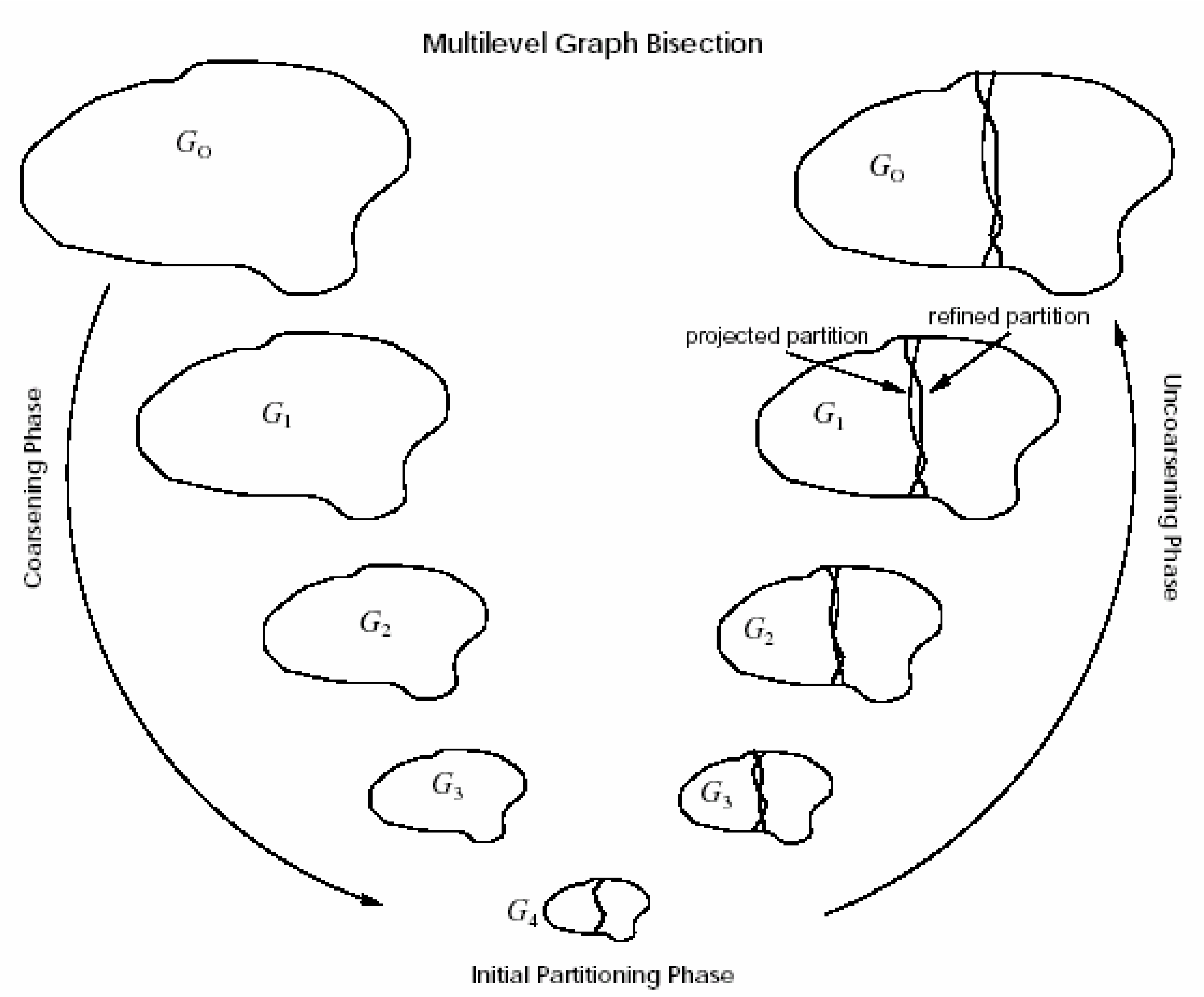
Parallel partitioning tools

ParMETIS, Jostle, PT-Scotch, Zoltan

Research area

- **unstructured meshes with up to 10^9 elements**

Multilevel algorithm of graph partitioning



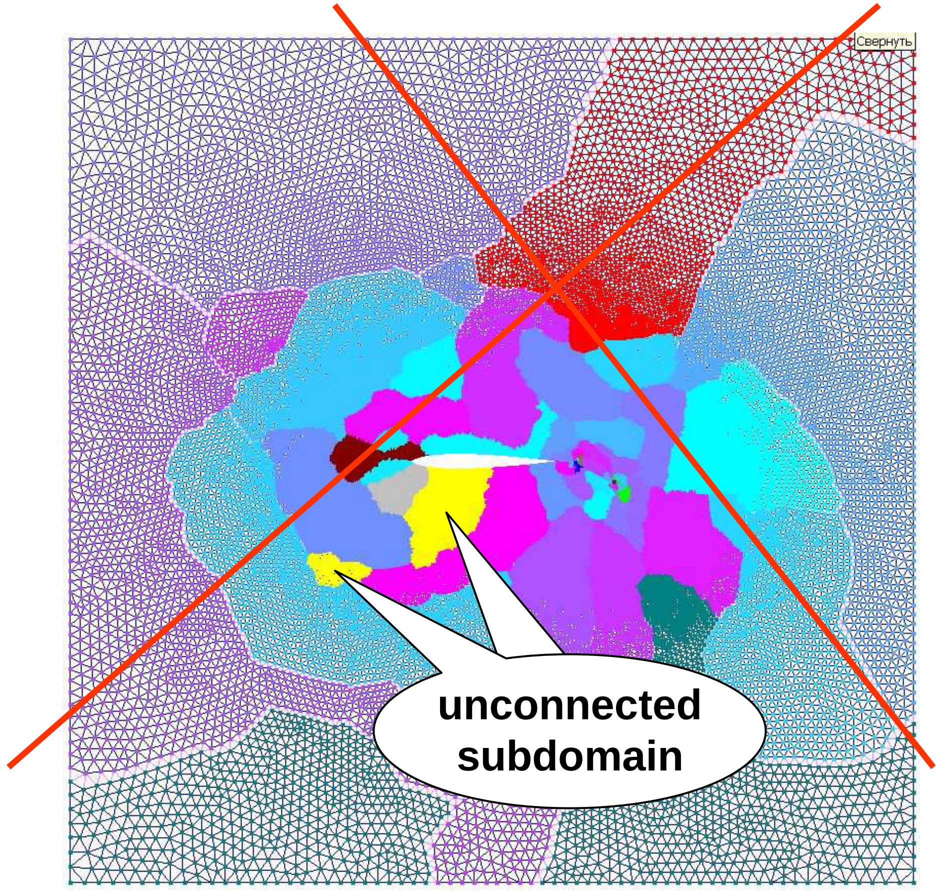
Shortcomings of present graph partitioning methods

- forming of unconnected subdomains
- generation of strongly imbalanced partitions

  (ParMETIS: number of vertices in some subdomains can be two times larger than in the others)
- can’t always make partitions into large number of microdomains

Connectivity is important:

- **iterative linear system solving methods**
- **mesh data compression**
- **subdomain composition algorithm[1]**
- **TIM-2D code parallelizing method[2]**



[1] Ilyushin A.I., Kolmakov A.A., Menshov I.S. Constructing parallel numerical model by means of the composition of computational objects // Mathematical Models and Computer Simulations. 2012. Vol. 4. Issue 1. 118-128.

[2] A. A. Voropinov. Data decomposition for TIM-2D code parallelizing method and its quality evaluation criteria // Bulletin of the South Ural State University. Series «Mathematical modelling, programming & computer software». 2009. Issue 4. №37(170). 40-50.

What’s new:
Partitioning tool GridSpiderPar

- **parallel incremental algorithm of graph partitioning**
- **parallel geometric algorithm of mesh partitioning**

Algorithms

❑ **make partitions of unstructured meshes with up to 10^9 elements into large number of microdomains**

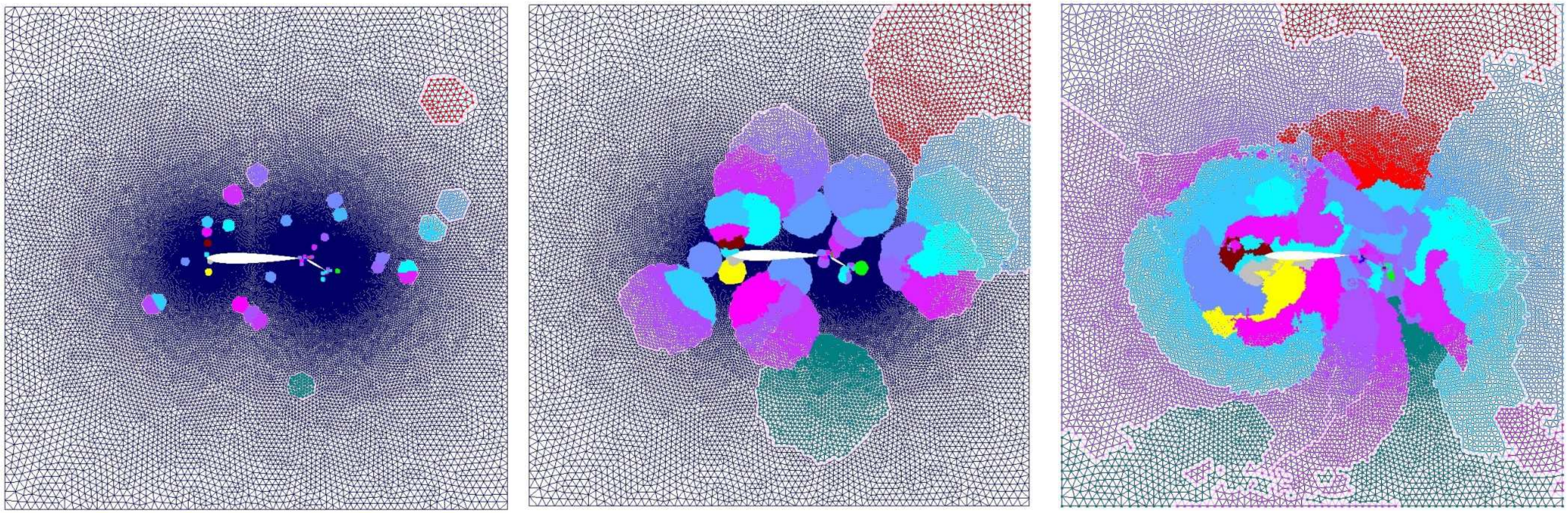
❑ **criteria:**

- **generation of balanced partitions**
- **forming of connected subdomains**
- **reducing edge-cut**

Incremental algorithm of graph partitioning

(M. Yakobovskiy, 2005, KIAM RAS)

- **incremental growth of subdomains**
- **diffusion of border vertices between subdomains**

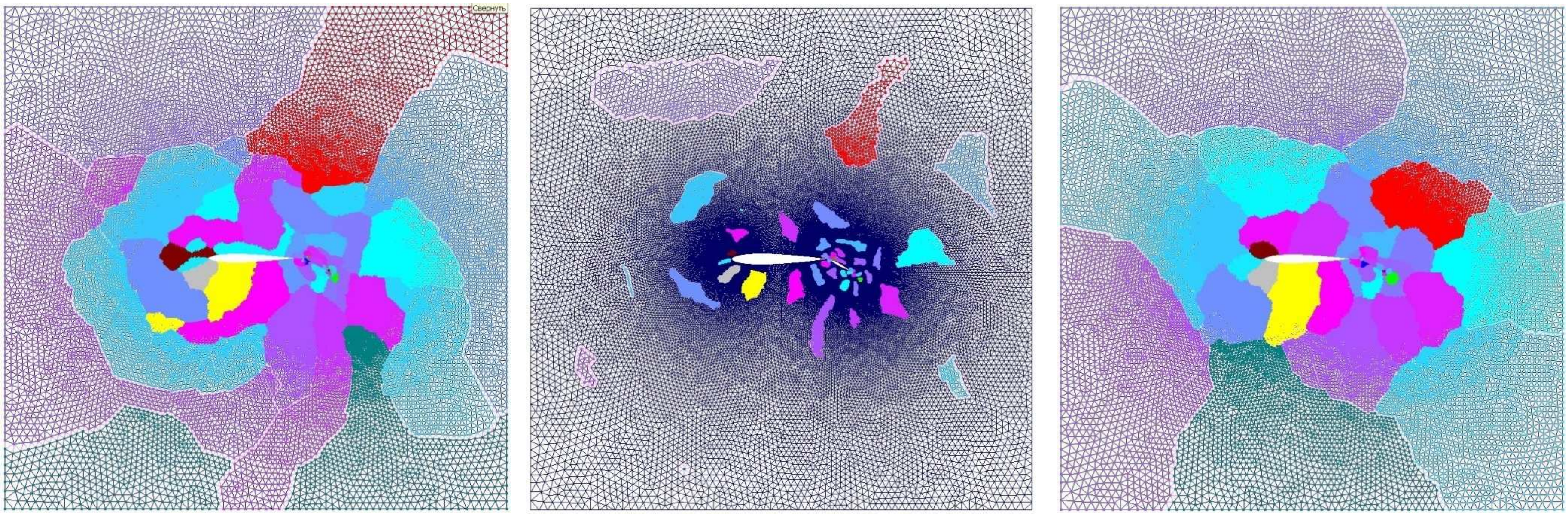


Example: mesh around an airfoil with a flap

Incremental algorithm

- **local refinement of subdomains**
- **subdomain quality control**
- **release some part of the vertices in bad subdomains**

$$T_{k+1} = \mathbf{A}T_k \setminus T_k \setminus T_{k-1}, \quad T_0 = \phi$$



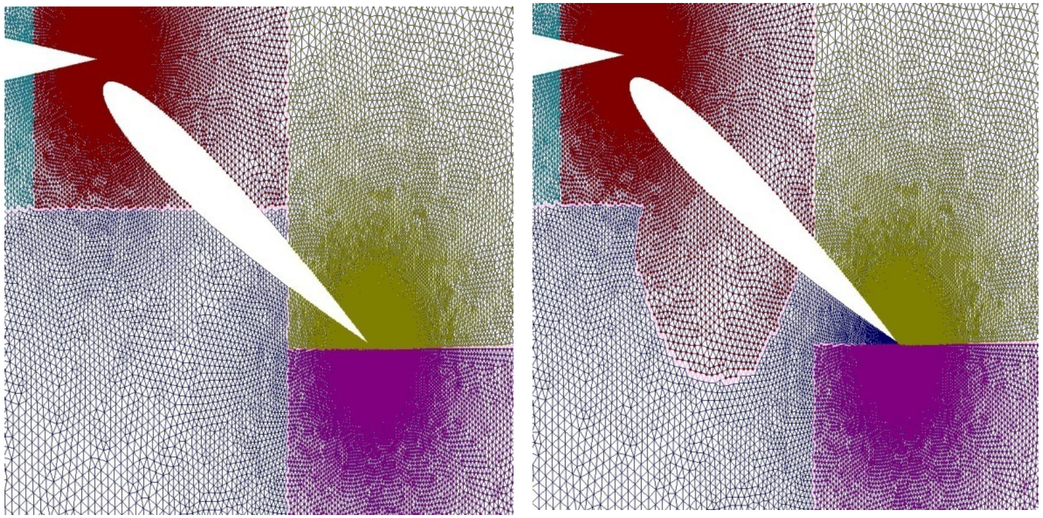
Example: mesh around an airfoil with a flap

Incremental algorithm of graph partitioning: Distinctions

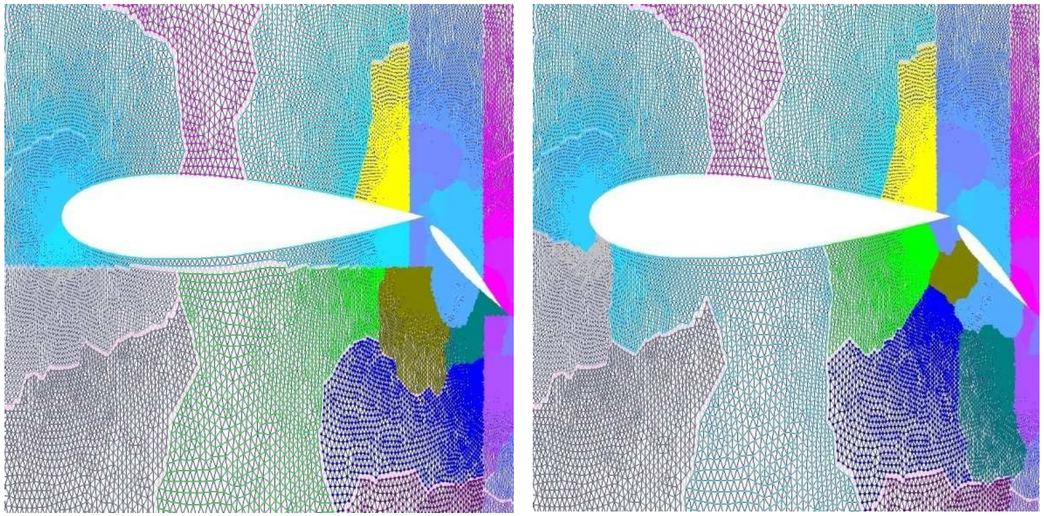
- **it is not based on multilevel approach**
- **it has some features similar to bubble growing and diffusion algorithms**
- **the bubble growing algorithm doesn’t guarantee that resulting partitions will be balanced**
- **difference from diffusion algorithms: it releases some part of the vertices in subdomains and then grows new subdomains**
- **new criterion for subdomain quality control (layers continuity)**

Parallel incremental algorithm of graph partitioning

- **geometric distribution of vertices among processors**
- **redistribution of small groups of vertices**



- **local partitioning**
- **collecting groups of bad subdomains and its repartitioning**



Example: mesh around an airfoil with a flap

Parallel incremental algorithm of graph partitioning:

Distinctions

- working with groups of subdomains of poor quality
- trying to decrease edge-cut in incremental growth of subdomains
- number of bad subdomains and edge-cut are taken into account in criterion of subdomains quality control

Parallel incremental algorithm of graph partitioning:

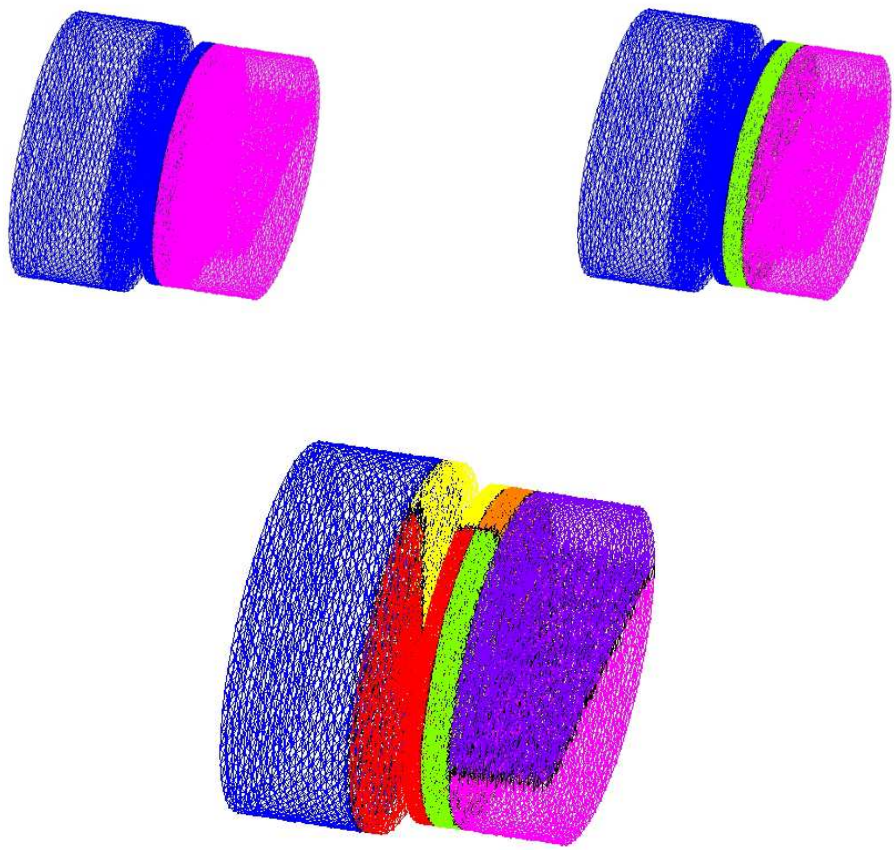
Advantages

- **is aimed at forming of connected subdomains**
- **balance of partitions is better than that made by other graph partitioning methods**

(5% (60%) → 0.05%)

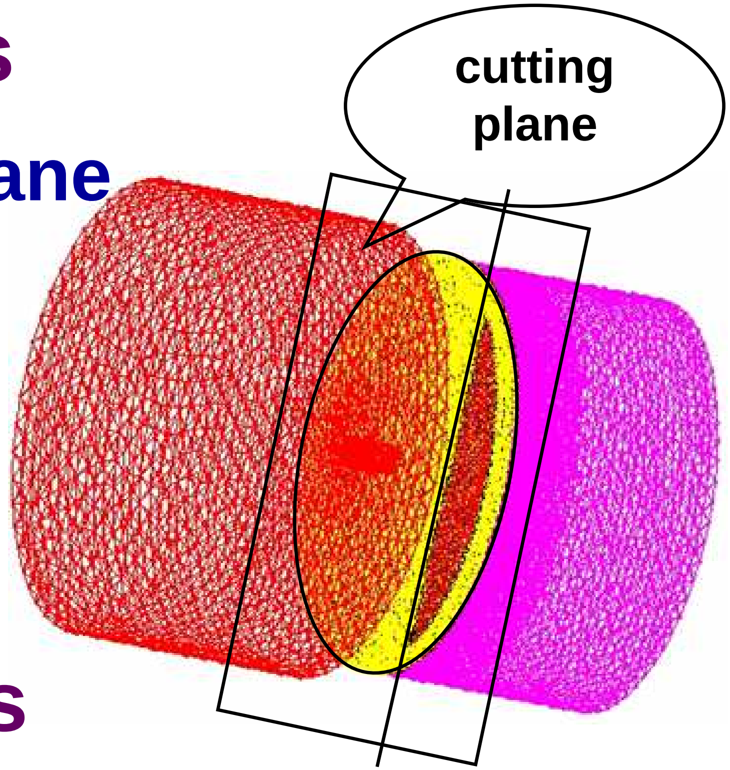
Parallel geometric algorithm of mesh partitioning

- **recursive coordinate bisection**



Parallel geometric algorithm of mesh partitioning:

Distinctions



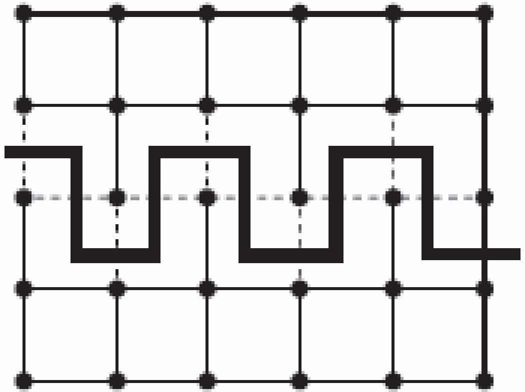
- making cuts of the cutting plane along other coordinate axes
- sorting only coordinates of vertices close to the cutting plane in local recursive coordinate bisection

Advantages

- difference in numbers of vertices in resulting subdomains is no more than 1 vertex
- efficient memory usage (only coordinates are stored)

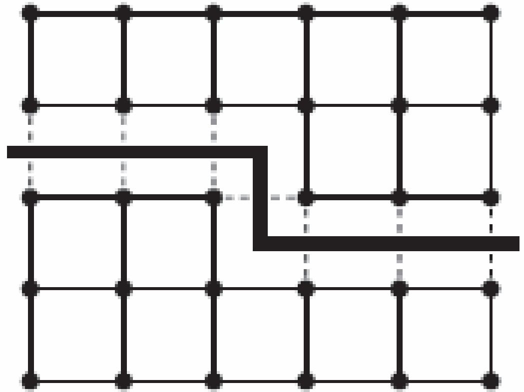
Edge-cut

11



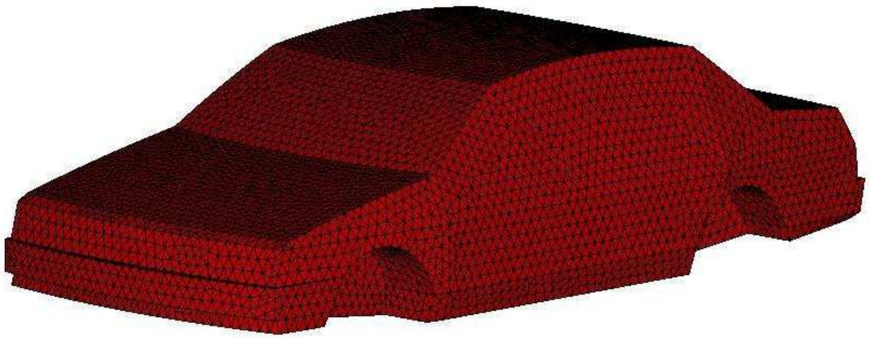
y[i-1]<=y[i]

7

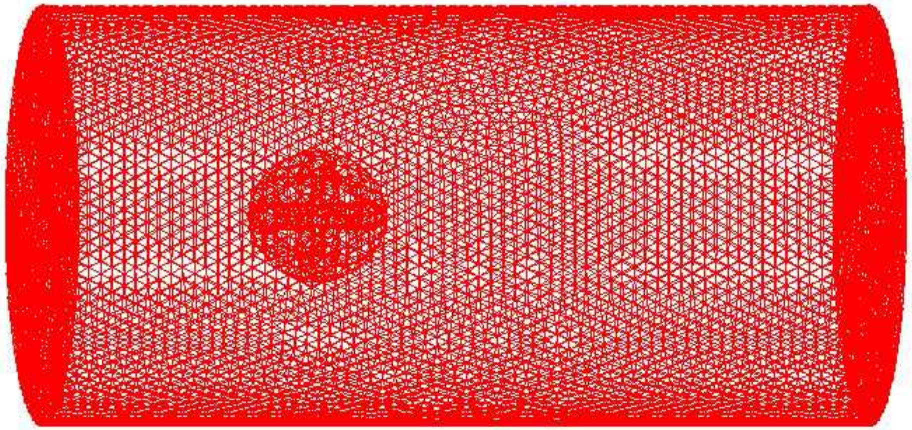


(y[i-1]<y[i])
или
((y[i-1]==y[i]) и (x[i-1]<x[i]))

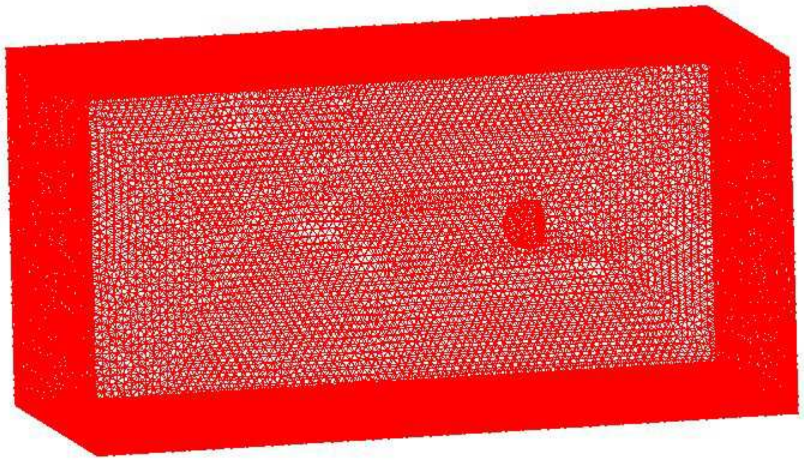
Tetrahedral meshes



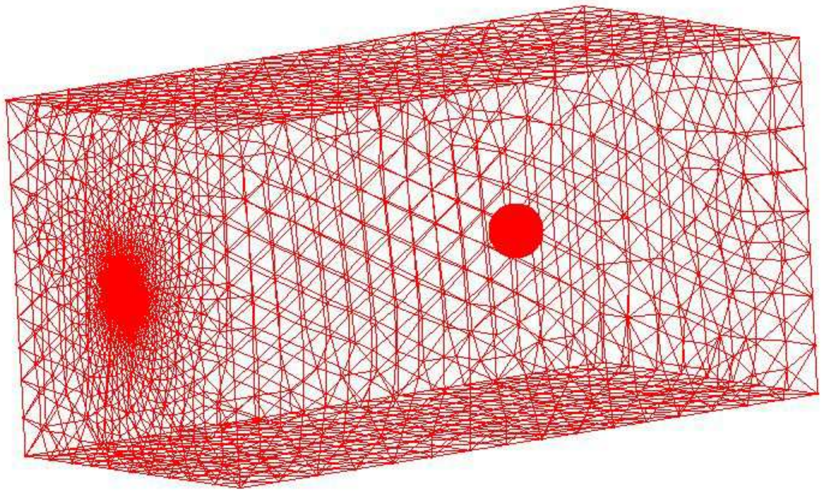
$2 \cdot 10^8$ vertices, $1.46 \cdot 10^9$ edges



$2.6 \cdot 10^8$ vertices, $1.8 \cdot 10^9$ edges



10^8 vertices, $7.7 \cdot 10^8$ edges



$2.8 \cdot 10^8$ vertices, $1.9 \cdot 10^9$ edges

Partitions into microdomains

Imbalance in 25600 microdomains, %

Methods	Mesh 1	Mesh 2	Mesh 3	Mesh 4
graph partitioning				
IncrDecomp	3,5	0,1	0,3	0,2
PartKway	53,4	59,8	58,6	64,3
PartGeomKway	48,7	50,4	62,4	56,5
PT-Scotch	8,3	8,3	8,3	8,3
geometric methods				
GeomDecomp	0,01	0,01	0,02	0,01
RCB	0,01	0,01	0,02	0,01

Partitions into microdomains

Number of unconnected microdomains in 25600

Methods	Mesh 1	Mesh 2	Mesh 3	Mesh 4
graph partitioning				
IncrDecomp	0	0	0	1
PartKway	69	35	37	29
PartGeomKway	67	34	28	37
PT-Scotch	7	0	2	4
geometric methods				
GeomDecomp	62	38	16	33
RCB	64	43	14	44

Partitions into subdomains

Imbalance in 512 subdomains, %

Methods	Mesh 1	Mesh 2	Mesh 3	Mesh 4
graph partitioning				
PartKway	12,9	20,6	17,6	28,4
PartGeomKway	31,1	35,7	44,2	51,4
PT-Scotch	4,9	1,7	2,8	2,9
geometric methods				
GeomDecomp	0	0	0	0
microdomain graph partitioning				
Simple average	5,3	5,4	3,7	5,1

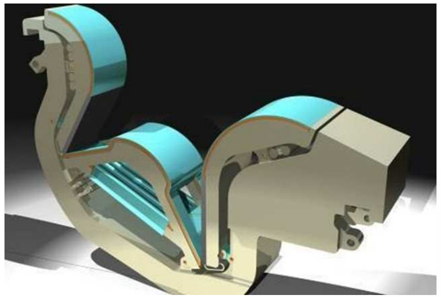
MARPLE3D code

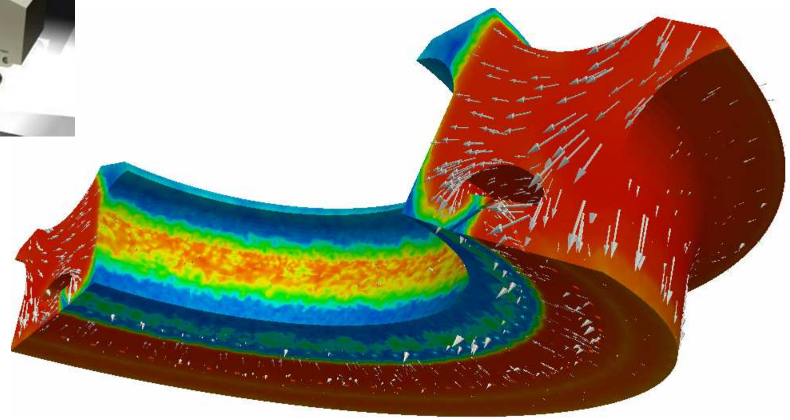
(KIAM RAS)

- Designed for multiphysics simulations in the field of radiative plasma dynamics

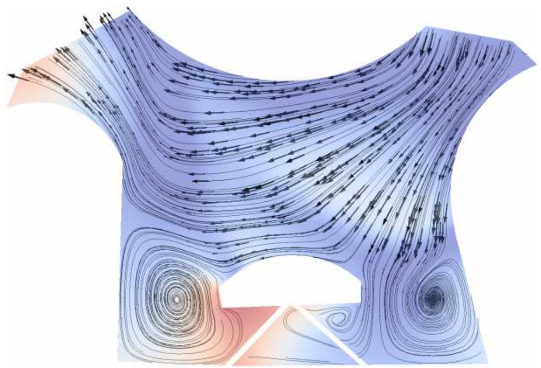
- **Testing of partitions obtained by tools GridSpiderPar, ParMETIS, Zoltan, and PT-Scotch was performed using simulations of the gas-dynamic problems**
- **Computational performance of the simulations with MARPLE3D code (KIAM RAS) run on different partitions was compared**

Model simulation of turbulent plasma flow in the ITER (future Tokamak) divertor

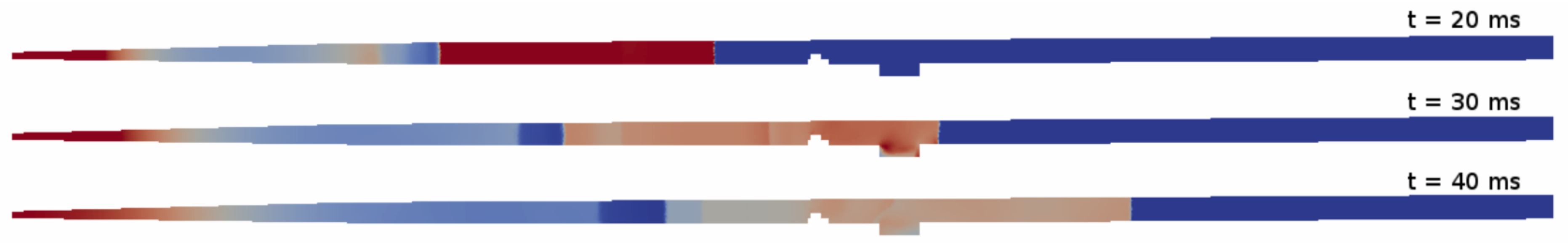




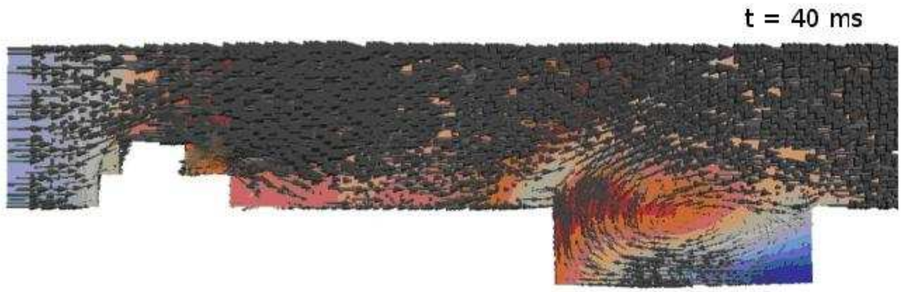
- complex hydrodynamics system including
- turbulence
- conductive&radiative heat transfer
- explicit and implicit schemes



Shock wave propagation in an extended structure (shock tube)

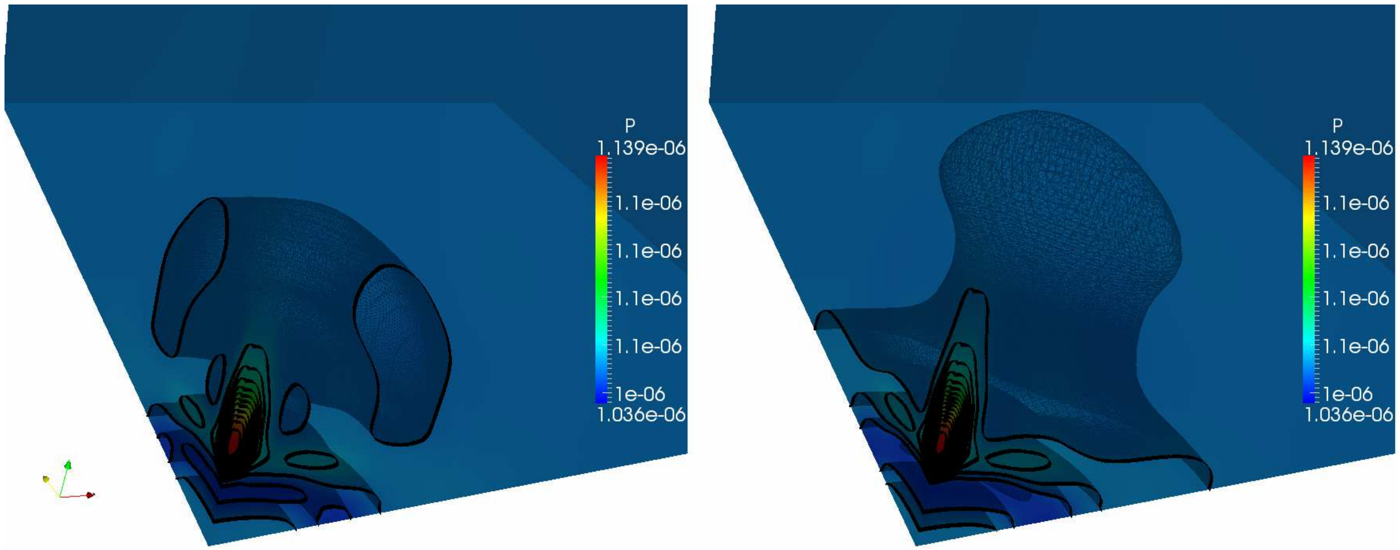


- complex hydrodynamics system including



- turbulence
- explicit and implicit schemes

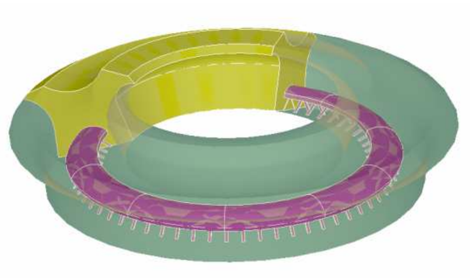
Near-earth explosion simulation

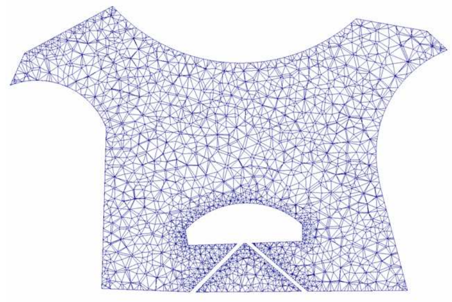


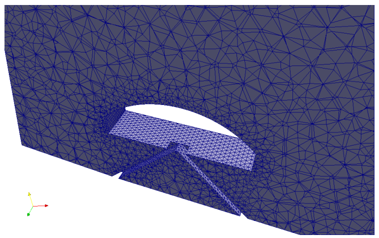
- **full hydrodynamics system including**
- **conductive heat transfer**
- **explicit and implicit schemes**

Test meshes

Tokamak divertor (divertor)

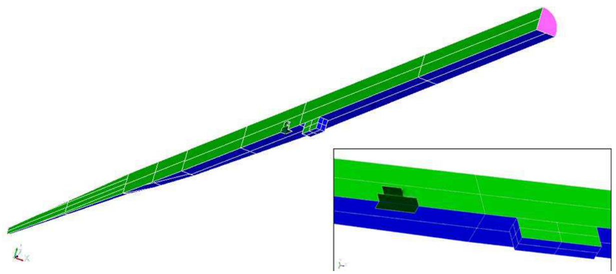


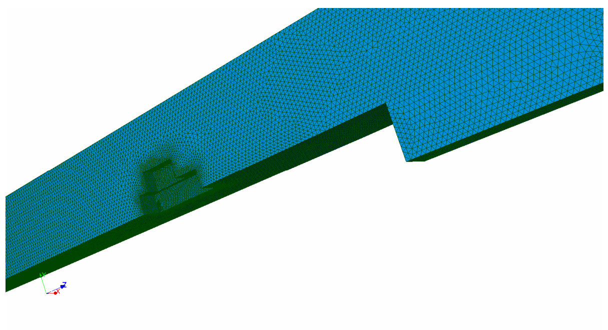




- 3D tetrahedral mesh (over 3 millions tetrahedrons)
- mesh refinement in the vicinity of small objects
- 256 subdomains

Shock tube (tube)

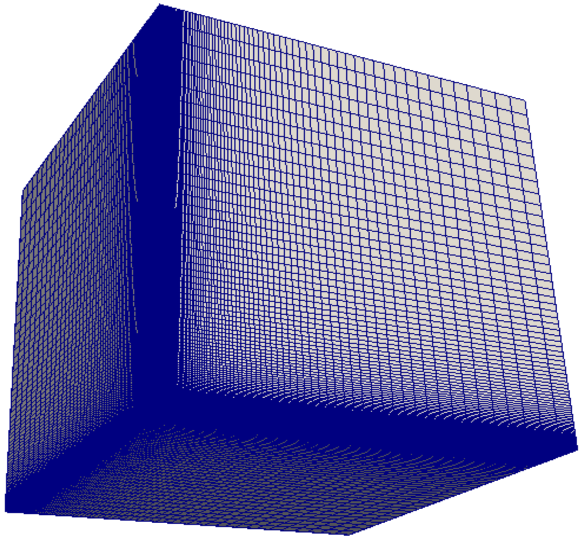




- 3D tetrahedral mesh (over 25 millions tetrahedrons)
- mesh refinement in the vicinity of small objects
- 4096 subdomains

Test meshes

Near-earth explosion (boom и boomL)



3D rectangular mesh
Over 61 millions cells for “boom”
Over 116 millions cells for “boomL”

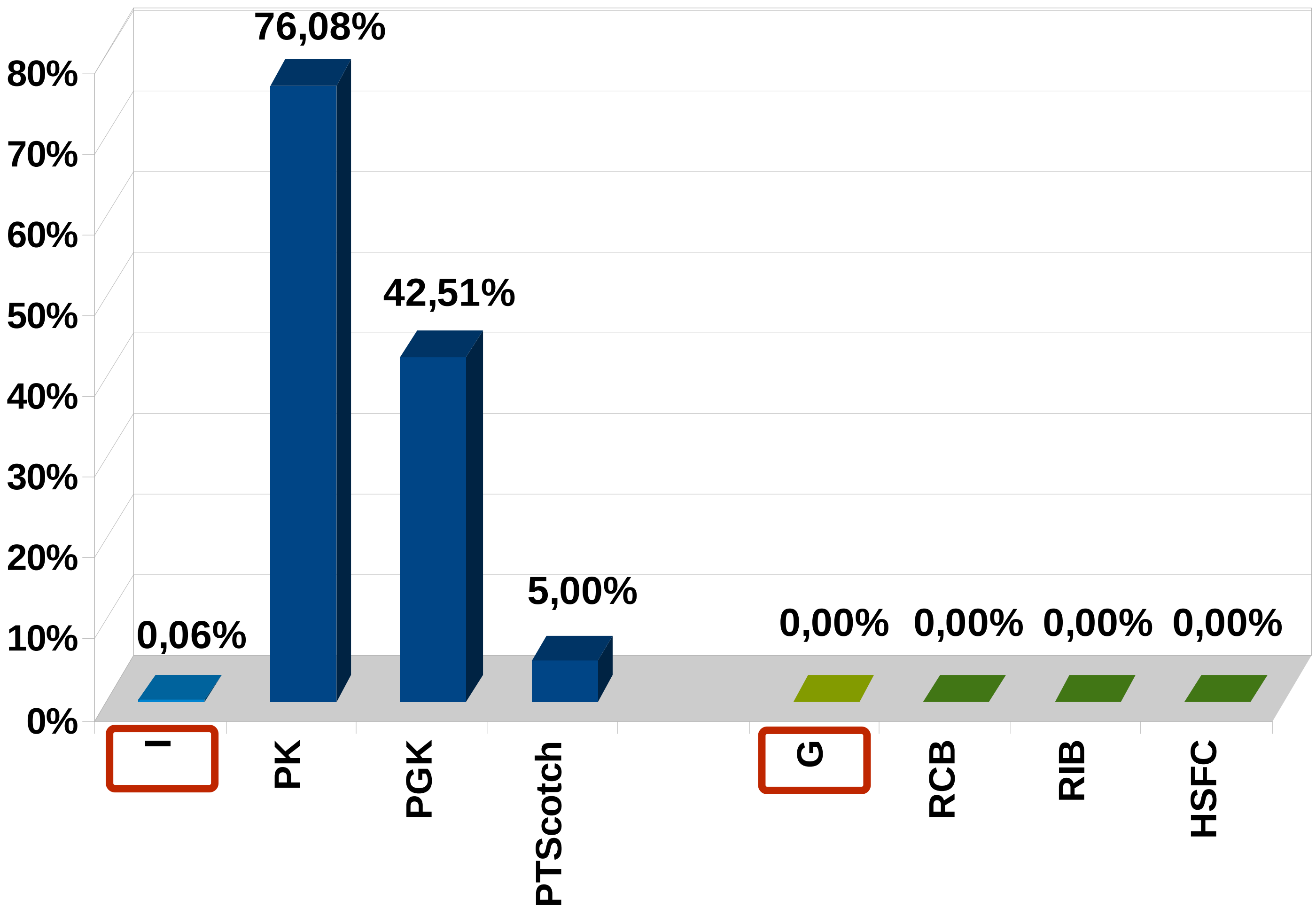
Parallelepipeds with different aspect ratio

boom:
- 4096 subdomains

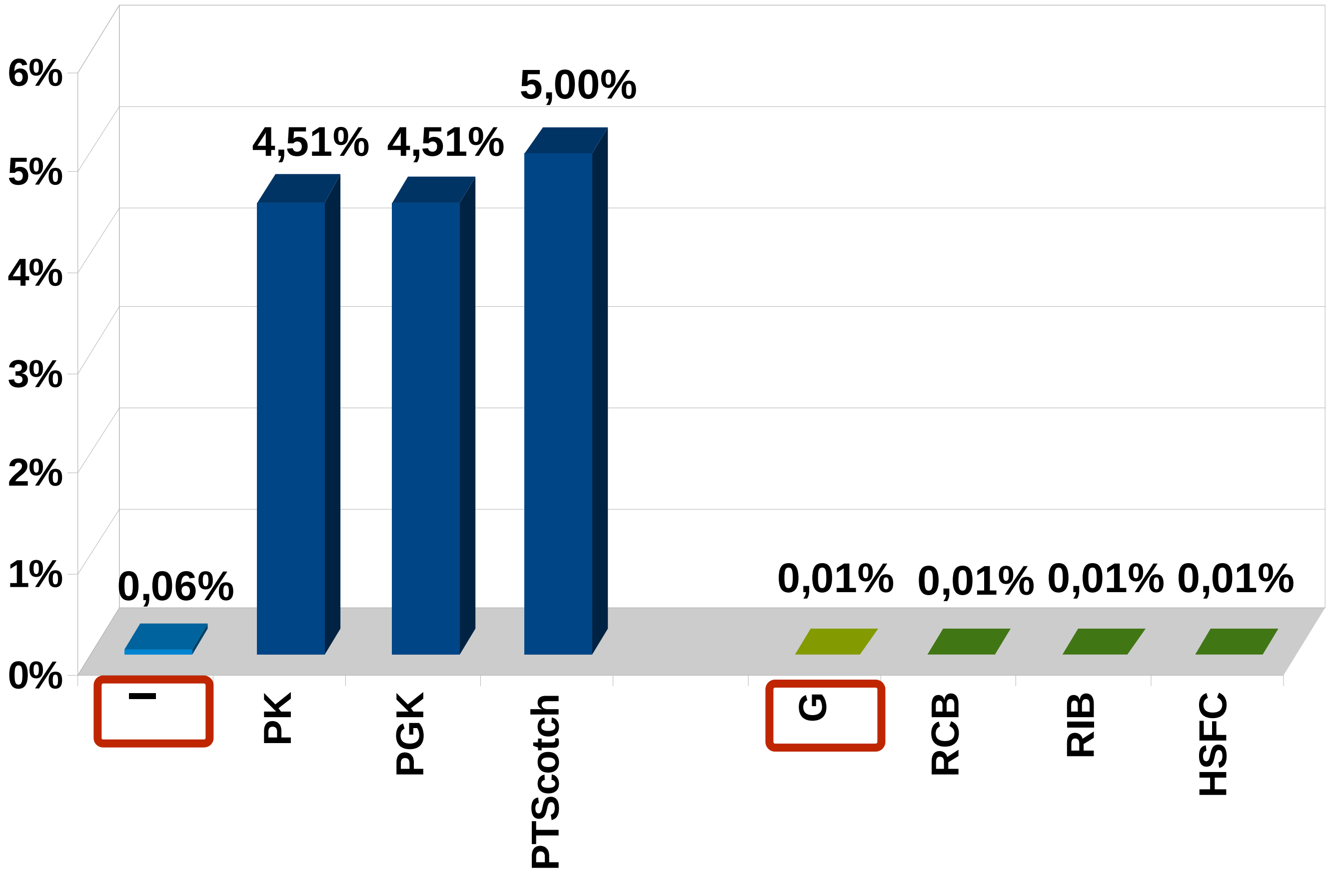
boomL:
- 10080 subdomains

- Dual graphs were constructed for each test mesh with number of vertices $2.8 \cdot 10^6$ - $1.2 \cdot 10^8$ and number of edges $2.3 \cdot 10^7$ - $1.0 \cdot 10^9$
- Computations were carried out on MVS-100K (227,94 TFlop/s), “Lomonosov" (1700 Tflop/s) and «Helios» (1524.1 TFlop/s)

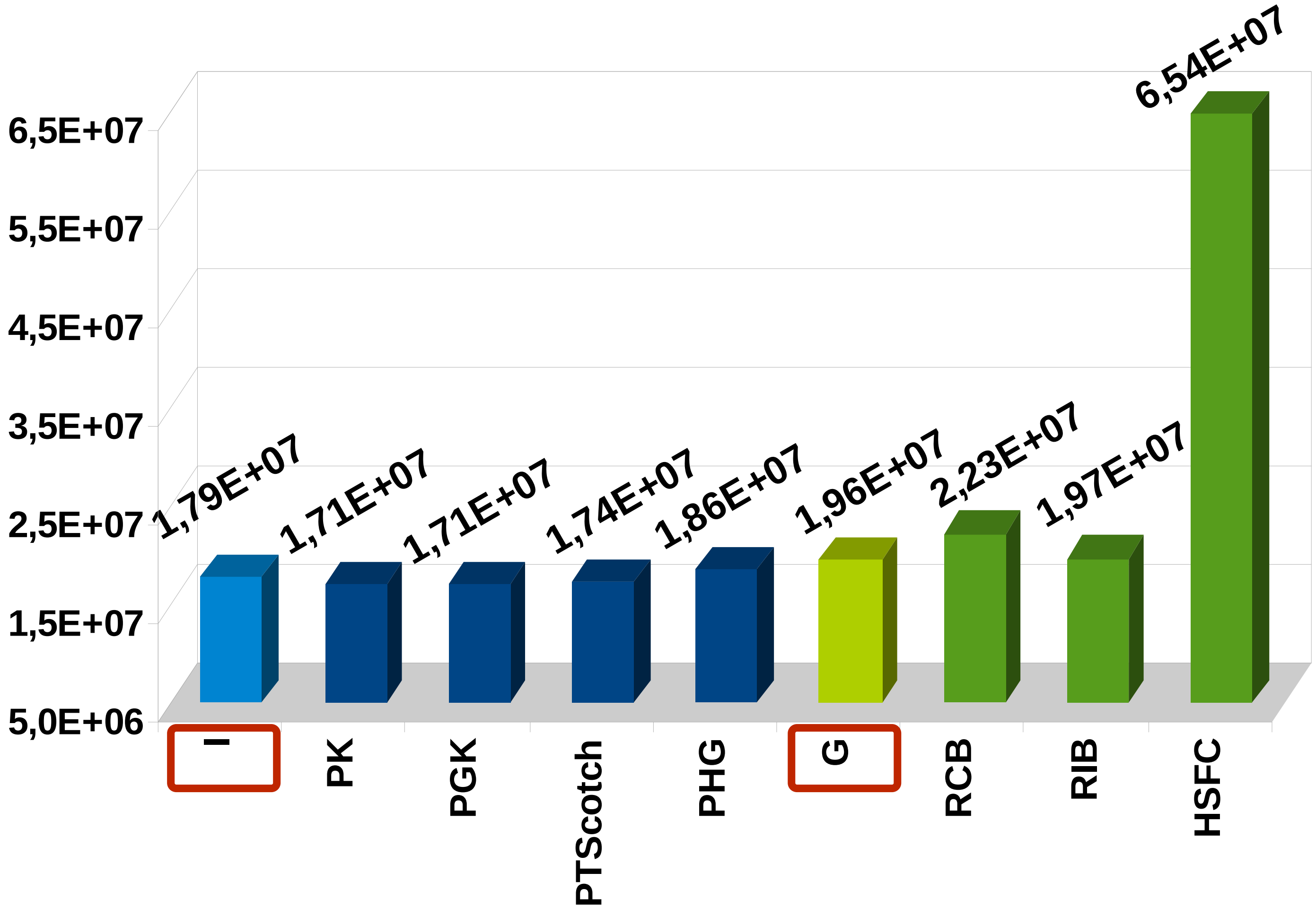
Imbalance in subdomains: lack of vertices (boom)



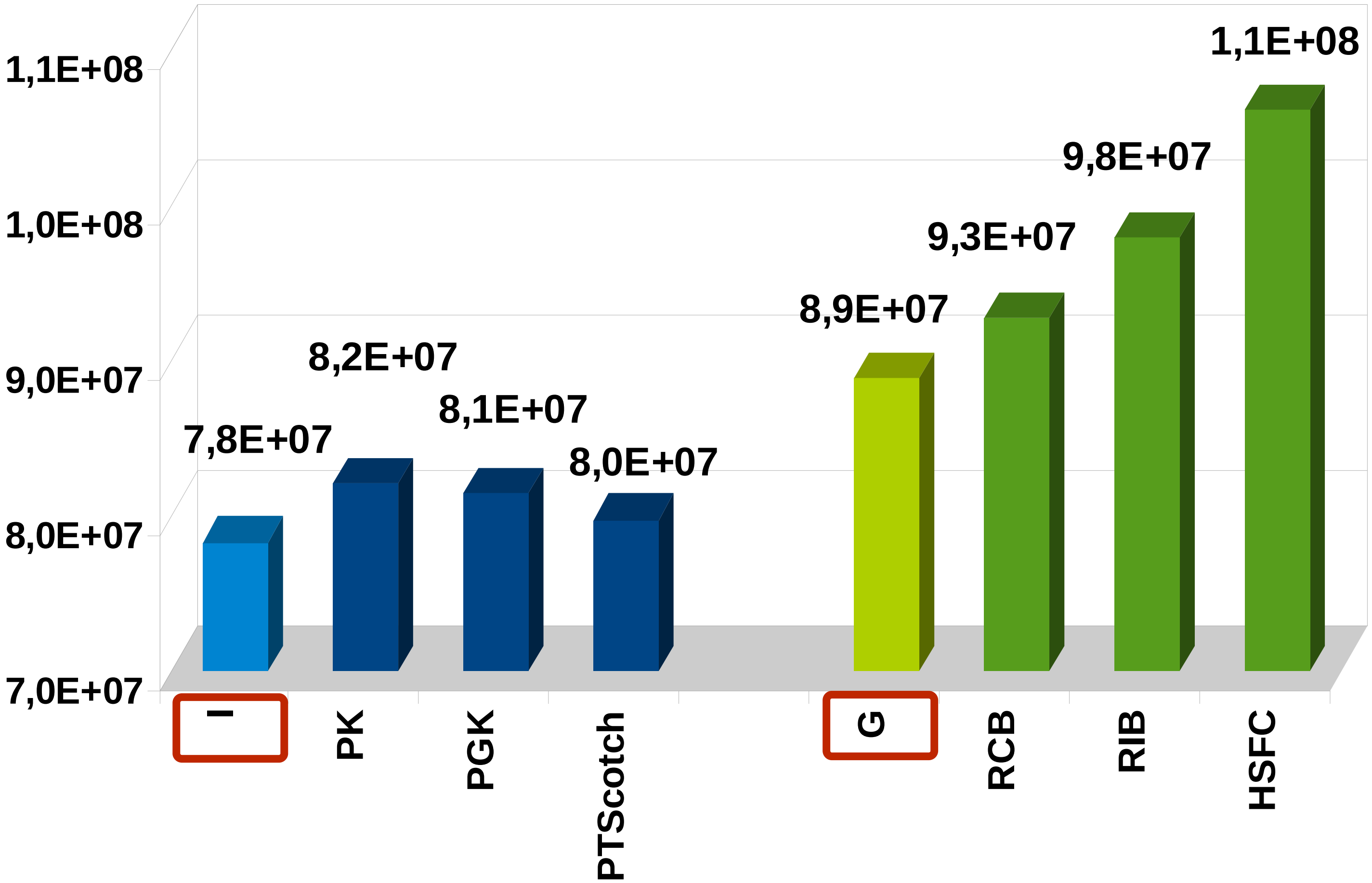
Imbalance in subdomains: overflow of vertices (boom)



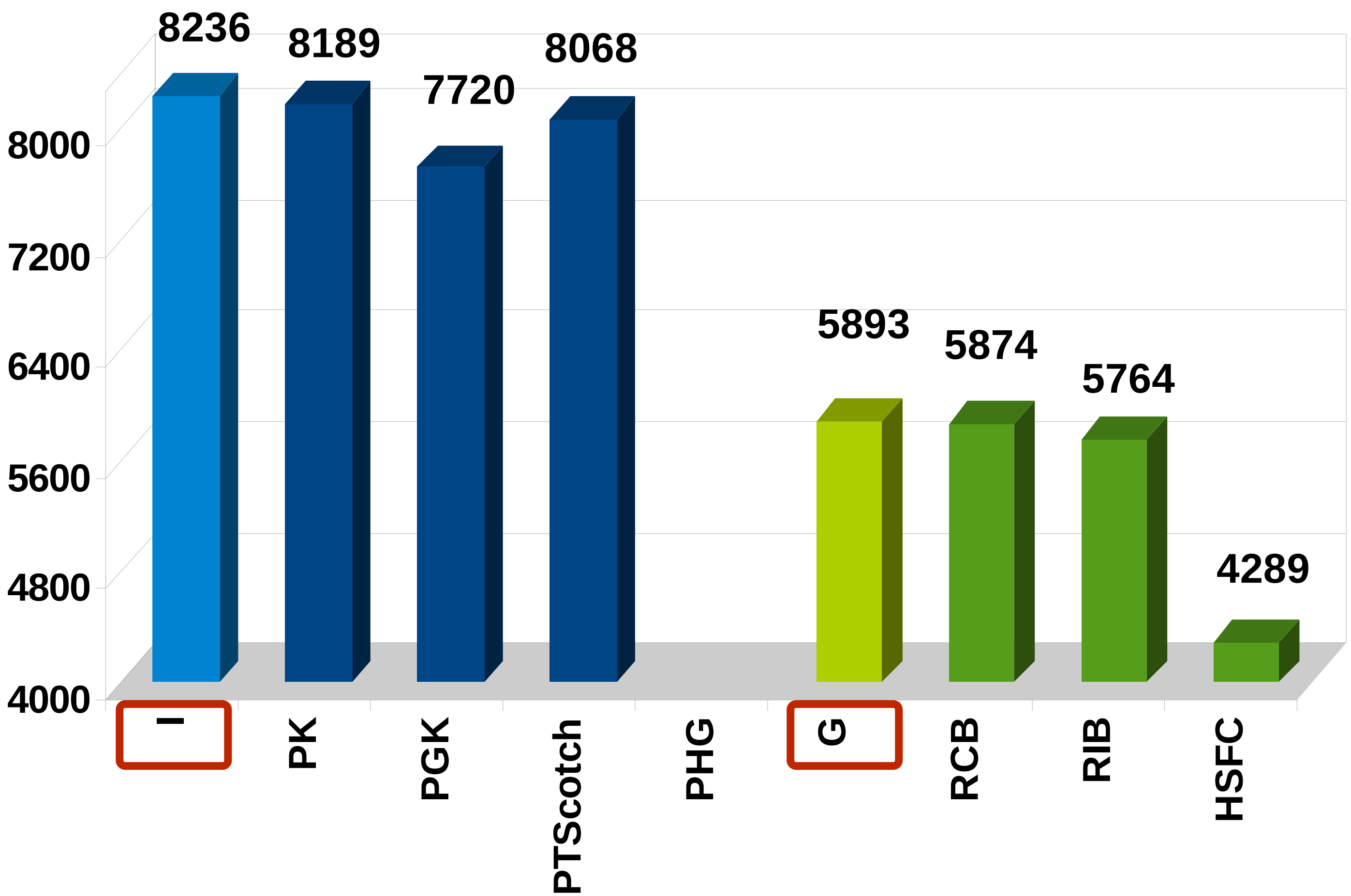
Cut edges
(tube)



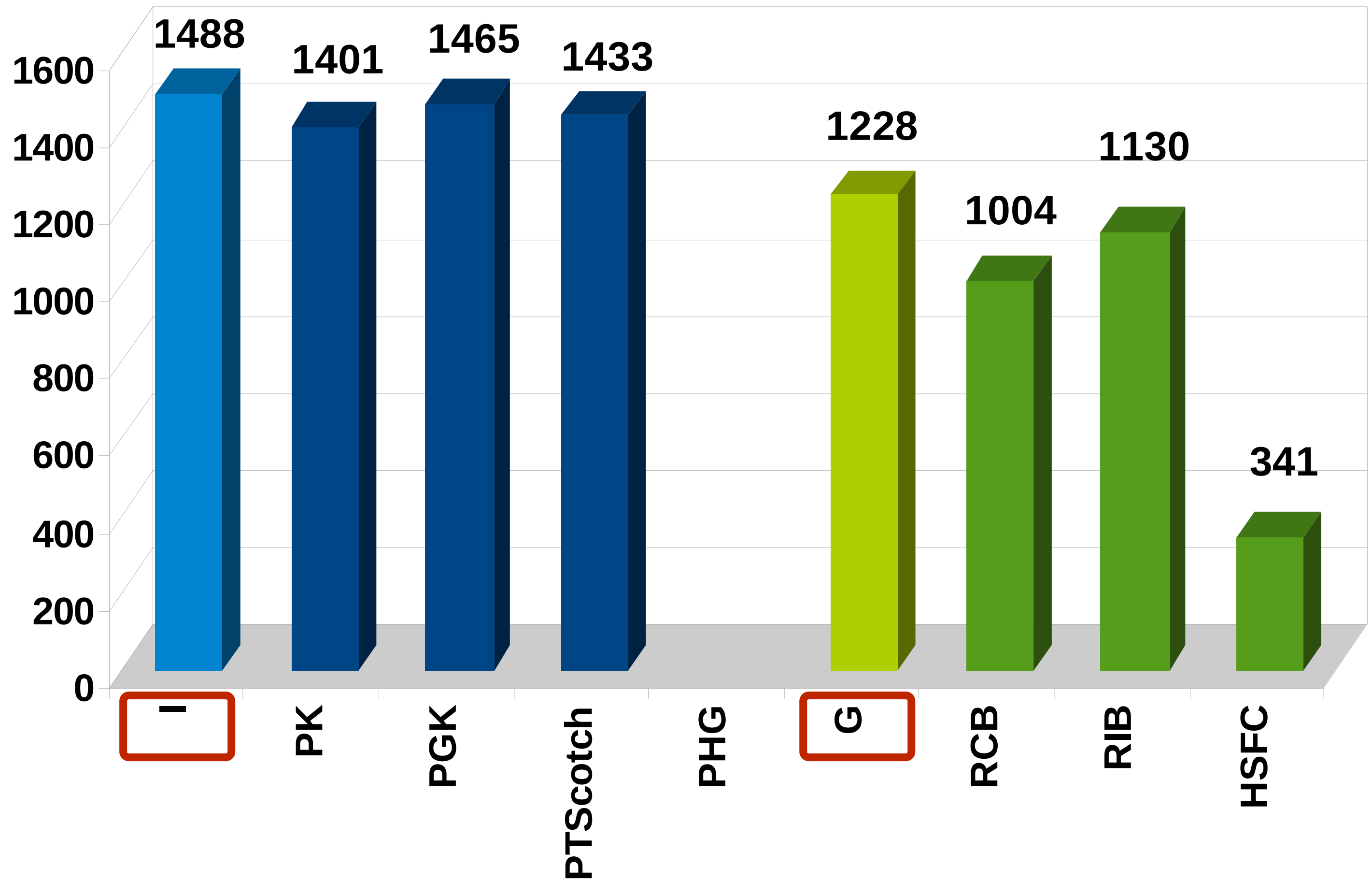
Cut edges (boomL)



Number of time steps (divertor)



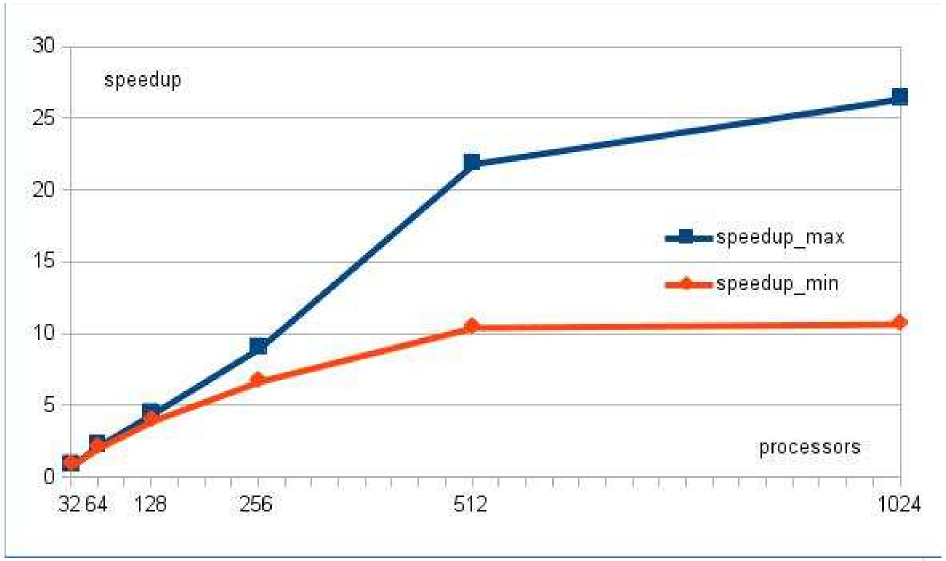
Number of time steps (tube)



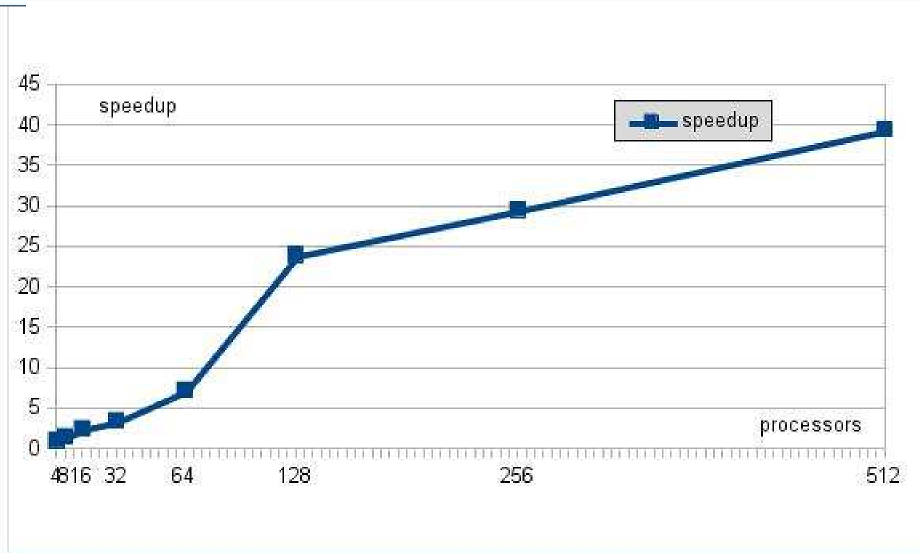
Testing of microdomain graph partitions on near-earth explosion simulation problem

Mesh info	Micro-domains	Micro-domains in sub-domain	Imbalance, %	Cut edges	Neighbou-ring subdomains (max.)	Uncon-nected sub-domains	Time steps
имя:	**3072**	**1**	**9,1**	**53 140 207**	**28**	0	**1107**
BoomL	24576	8	**62,5**	64 611 859	25	0	833
	49152	16	37,5	66 566 874	25	0	880
116 214 272	98304	32	18,7	68 841 339	23	0	949
hexahedrons	**196608**	**64**	**7,9**	**68 207 798**	**21**	0	**999**

Strong scaling



parallel incremental algorithm of graph partitioning



parallel geometric algorithm of mesh partitioning

partitioning of the hexahedral mesh with $1.47 \cdot 10^7$ cells into 1024 subdomains

Results

1. **Algorithms for parallel decomposition of large computational meshes (up to 10^9 elements) were devised: parallel incremental algorithm of graph partitioning and parallel geometric algorithm of mesh partitioning.**
2. **A partitioning tool GridSpiderPar was developed.**
3. **Different partitions into microdomains, microdomain graph partitions and partitions into subdomains of several meshes (10^8 vertices, 10^9 elements) obtained by means of the partitioning tool GridSpiderPar and the packages ParMETIS, Zoltan and PT-Scotch were compared. The results revealed advantages of the devised algorithms in the quality of the partitions.**

Results

4. **GridSpiderPar, ParMETIS, Zoltan, and PT-Scotch were compared via gas-dynamic problem simulations. Test studies demonstrate efficiency of the developed algorithms.**

5. **Testing of microdomain graph partitions on the near-earth explosion simulation problem revealed the potential of using this strategy for simulations.**

Thank You!