

*47<sup>th</sup> HPC User Forum  
Using HPC to Drive Economic and Scientific Competitiveness*

# *High Performance Computing at Moscow State University and more...*

*Prof. Vladimir Voevodin  
Deputy Director, Research Computing Center,  
Moscow State University*

*[voevodin@parallel.ru](mailto:voevodin@parallel.ru)*

*July, 9, 2012*

*HLRS / University of Stuttgart, Stuttgart, Germany*

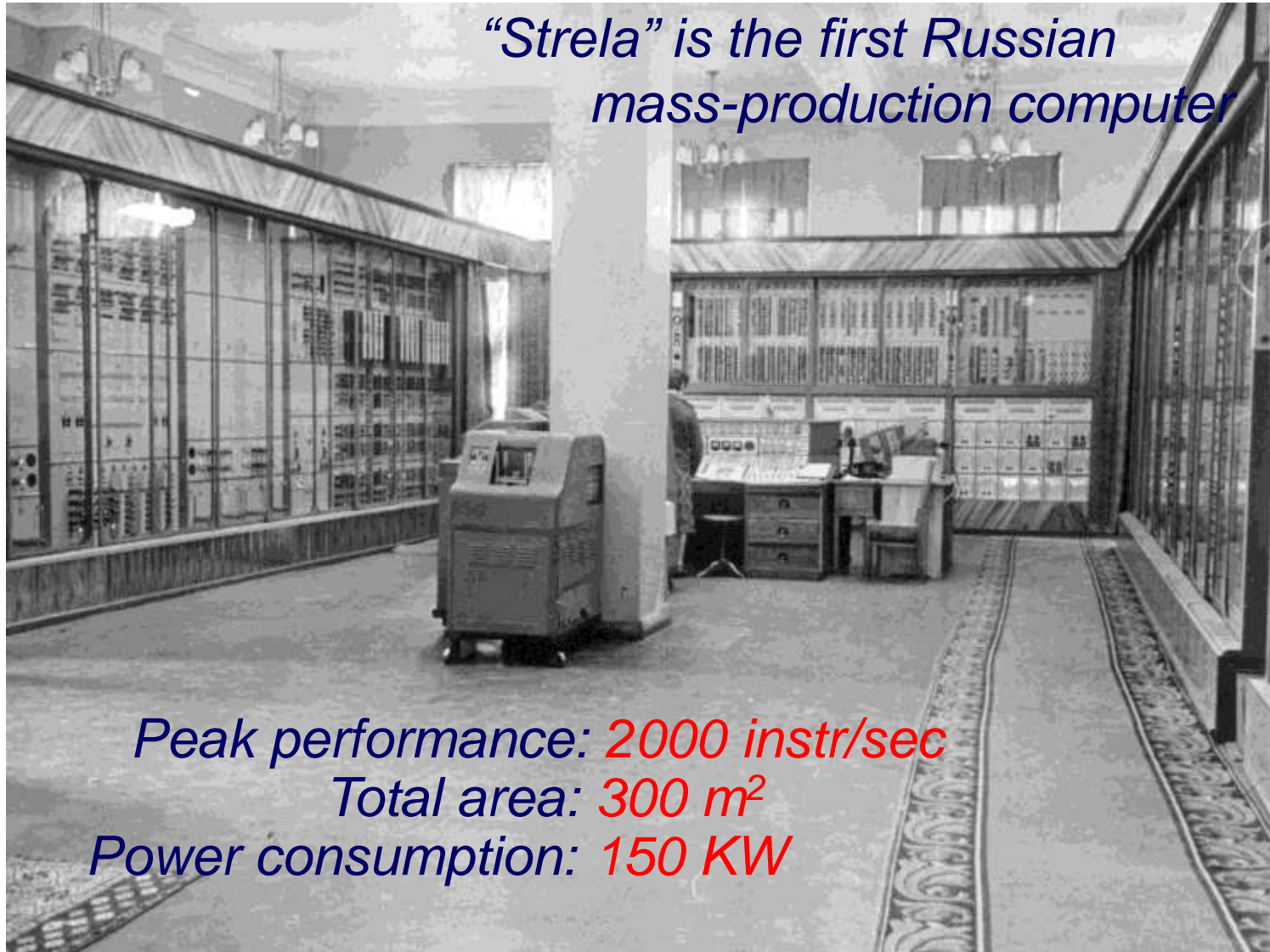


*Moscow State University*  
*1755 – 2012*

*40 Faculties*  
*350+ Departments*  
*5 major Research Institutes*

*More than 40 000 students,*  
*2500 full doctors, 6000 PhDs,*  
*1000+ full professors,*  
*5000 researchers.*

# Computing Center of MSU, 1956



*“Strela” is the first Russian mass-production computer*

*Peak performance: 2000 instr/sec*  
*Total area: 300 m<sup>2</sup>*  
*Power consumption: 150 KW*

*12 years ago ...*

*(24 CPUs, Intel P-III/500 MHz, SCI network, 8 m<sup>2</sup> , 12 Gflops)*



*... and now*

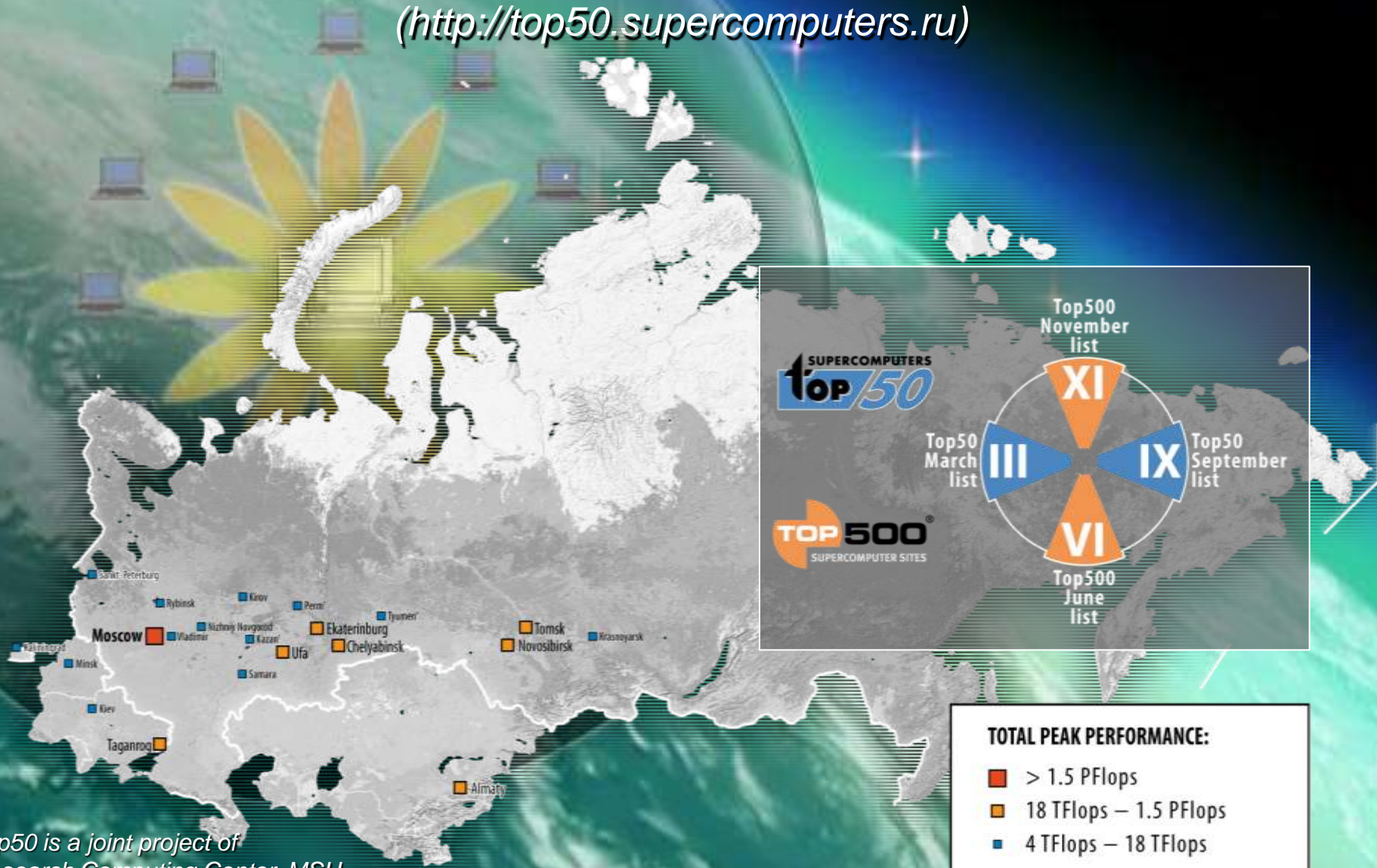
*(52.000+ Intel cores, 2.130 NVIDIA GPUs, QDR IB, 1200 m<sup>2</sup> , 1.7 Pflops)*



M.V.Lomonosov  
1711 – 1765

# Top50 supercomputers of CIS

(<http://top50.supercomputers.ru>)



Top50 is a joint project of  
Research Computing Center, MSU  
Joint Supercomputer Center, RAS

# Top50 Supercomputers: Sites/Cities



Research Computing Center, MSU  
Joint Supercomputer Center, RAS

# Moscow University Supercomputing Center

Today:

“Lomonosov” supercomputer:  
1.7 Pflops

SKIF MSU “Chebyshev” supercomputer:  
60 Tflops

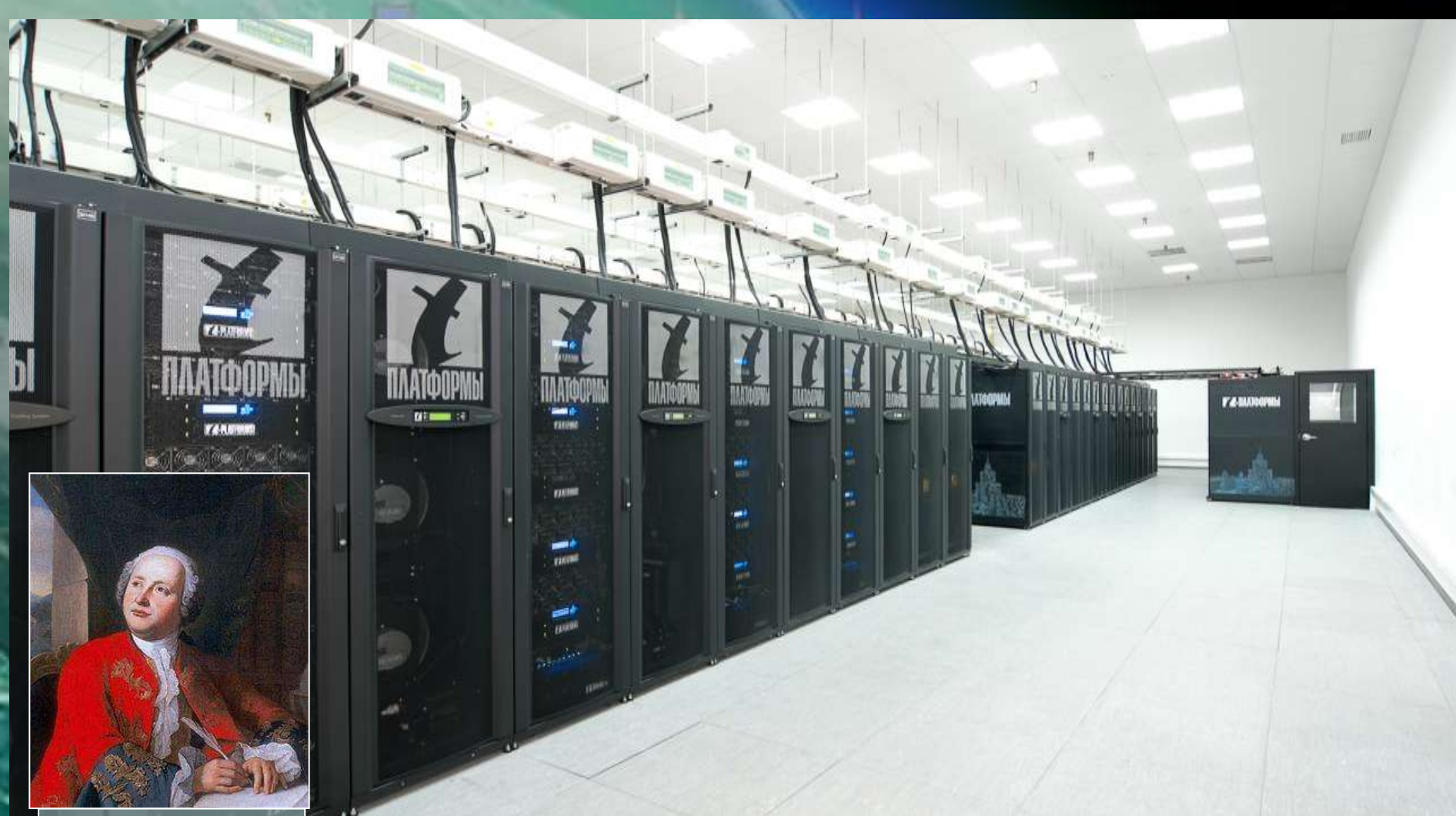
IBM Blue Gene/P supercomputer:  
27 Tflops

Hewlett-Packard GPU-supercomputer:  
26 Tflops





# MSU "Lomonosov" supercomputer



M.V.Lomonosov  
1711 – 1765

# MSU "Lomonosov" supercomputer



# MSU "Lomonosov" supercomputer



# MSU "Lomonosov" supercomputer



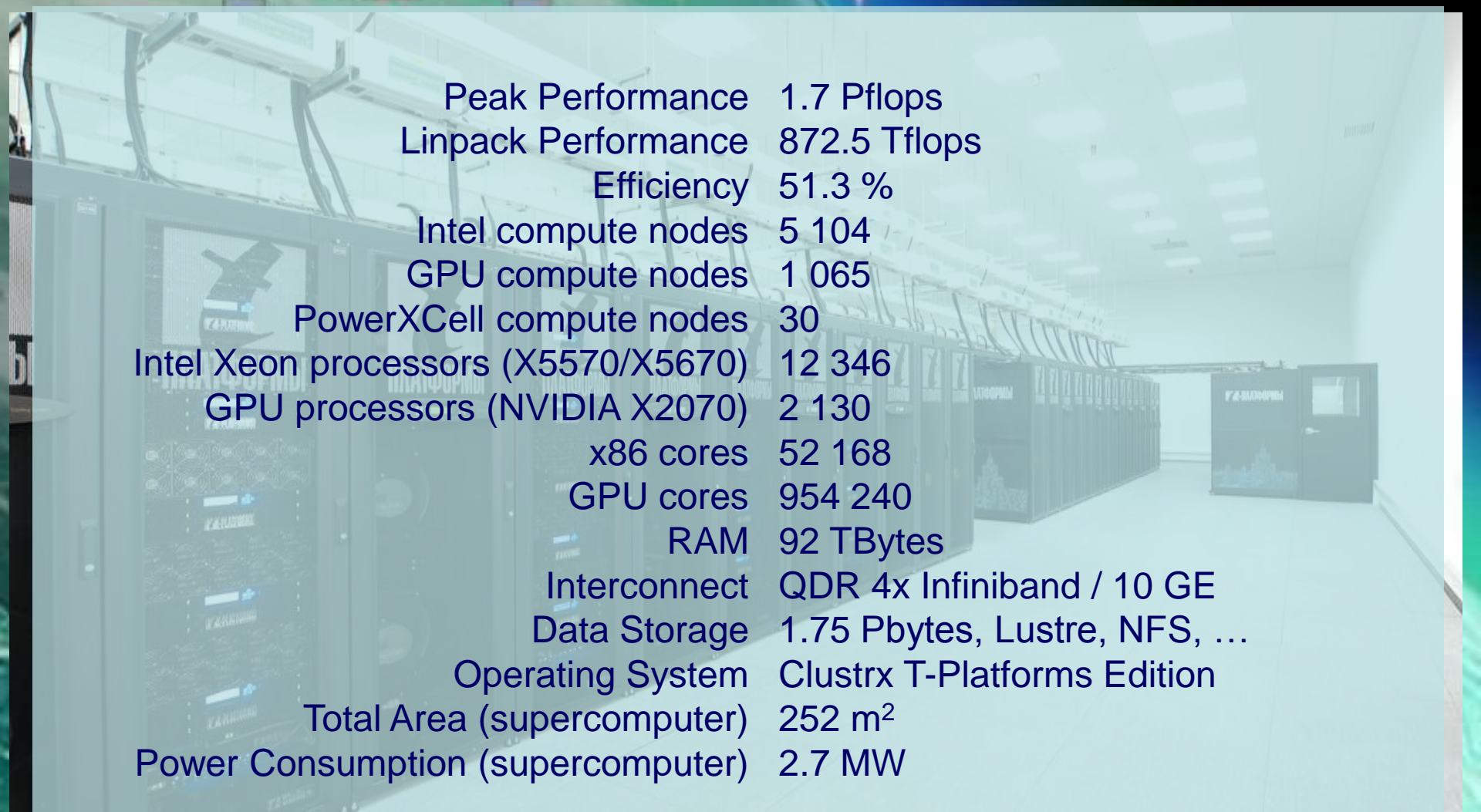
# MSU "Lomonosov" supercomputer



# MSU "Lomonosov" supercomputer



# MSU “Lomonosov” supercomputer, 2012



Peak Performance	1.7 Pflops
Linpack Performance	872.5 Tflops
Efficiency	51.3 %
Intel compute nodes	5 104
GPU compute nodes	1 065
PowerXCell compute nodes	30
Intel Xeon processors (X5570/X5670)	12 346
GPU processors (NVIDIA X2070)	2 130
x86 cores	52 168
GPU cores	954 240
RAM	92 TBytes
Interconnect	QDR 4x Infiniband / 10 GE
Data Storage	1.75 Pbytes, Lustre, NFS, ...
Operating System	Clustrx T-Platforms Edition
Total Area (supercomputer)	252 m <sup>2</sup>
Power Consumption (supercomputer)	2.7 MW

# MSU “Lomonosov” supercomputer, 2012

(node types)

Node types	RAM per node	Quantity
2 x Xeon 5570 2.93 GHz	12 GB	4160
2 x Xeon 5570 2.93 GHz	24 GB	260
2 x Xeon 5670 2.93 GHz	24 GB	640
2 x Xeon 5670 2.93 GHz	48 GB	40
2 x PowerXCell 8i 3.2 GHz	16 GB	30
2 x Xeon E5630 2.53 GHz, 2 x Tesla X2070	12 GB	777
2 x Xeon E5630 2.53 GHz, 2 x Tesla X2070	24 GB	288
4 x Xeon E7650 2.26 GHz	512 GB	4



# MSU Supercomputing Center

(users & organizations)

	2009	2010	2011
<i>User groups, total:</i>	241	369	545
<i>including:</i>			
<i>from Moscow University:</i>	155	241	359
<i>from institutes of RAS:</i>	53	77	110
<i>from other organizations:</i>	33	51	76
<i>Faculties / Institutes of MSU:</i>	15	21	24
<i>Institutes of RAS:</i>	20	28	35
<i>Others:</i>	19	24	34

# *MSU Supercomputing Center*

*(users & organizations)*

*Diversity of users/groups/applications implies two serious questions:*

- efficiency,*
- education.*

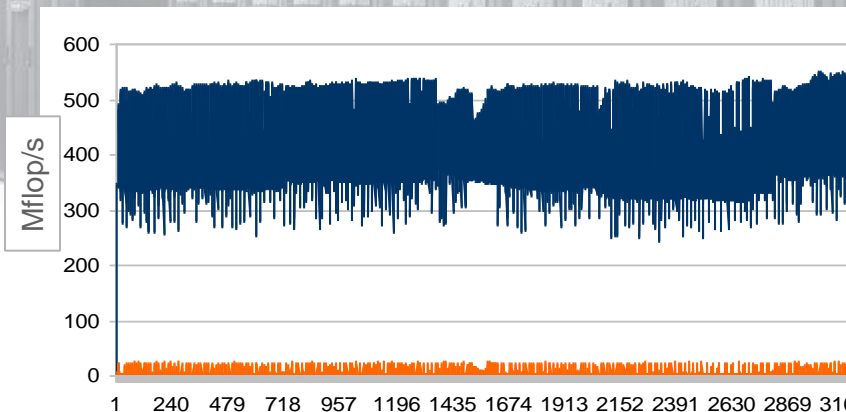
# Efficiency, efficiency, efficiency...

What we may say about efficiency of supercomputing centers?

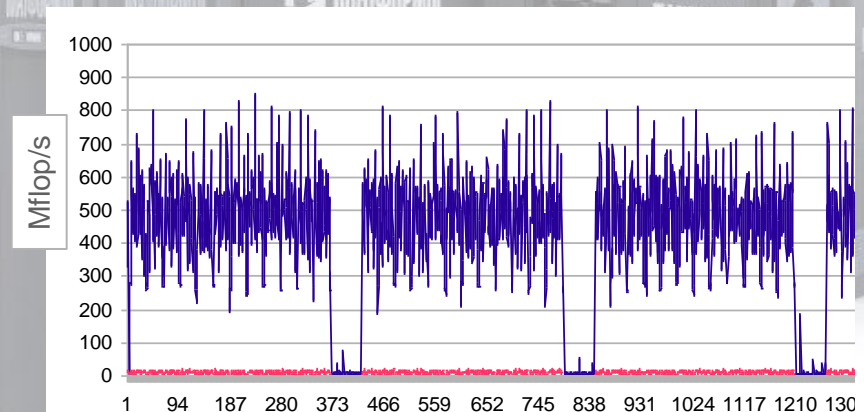
1 Pflops system :

Expected:  $1\text{Pflop} * 60\text{sec} * 60\text{min} * 24\text{hours} * 365\text{days} = 31,5 \text{ ZettaFlop per year}$

What is in reality? 0,0..x%



Drug design, 3.5% efficiency  
(serial code)



Climate modeling, 4% efficiency  
(serial code)

Why? Peculiarities of hardware, a complicated job-flow, poor data locality, a huge degree of parallelism in hardware, etc...



# HOLISTIC Performance System Analysis



RF part: **LAPTA** Project

Moscow State University Research Computing Center

# HOPSA-RU

*RF coordinator: Vladimir Voevodin*

Efficiency and root cause analysis are the key points of the project

# HOPSA project

ICT EU-Russia Coordinated Project (FP7-2011-EU-Russia)

**HOPSA** project – **HO**listic **P**erformance **S**ystem **A**nalysis



*EU partners:*

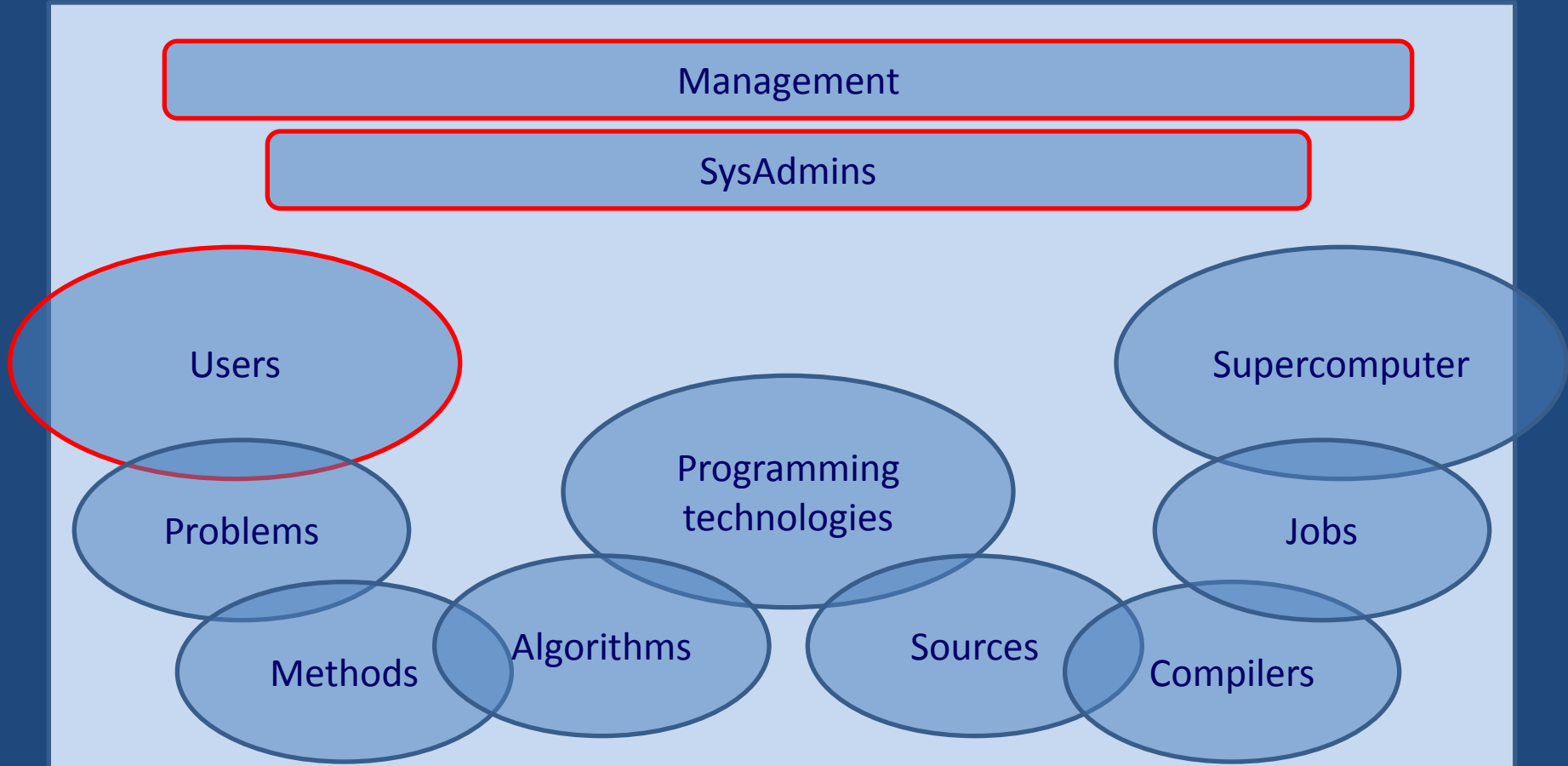
- *Forschungszentrum Juelich GmbH (EU coordinator);*
- *Rogue Wave Software AB;*
- *Barcelona Supercomputing Center;*
- *German Research School for Simulation Sciences;*
- *Technical University Dresden.*



*Russian partners:*

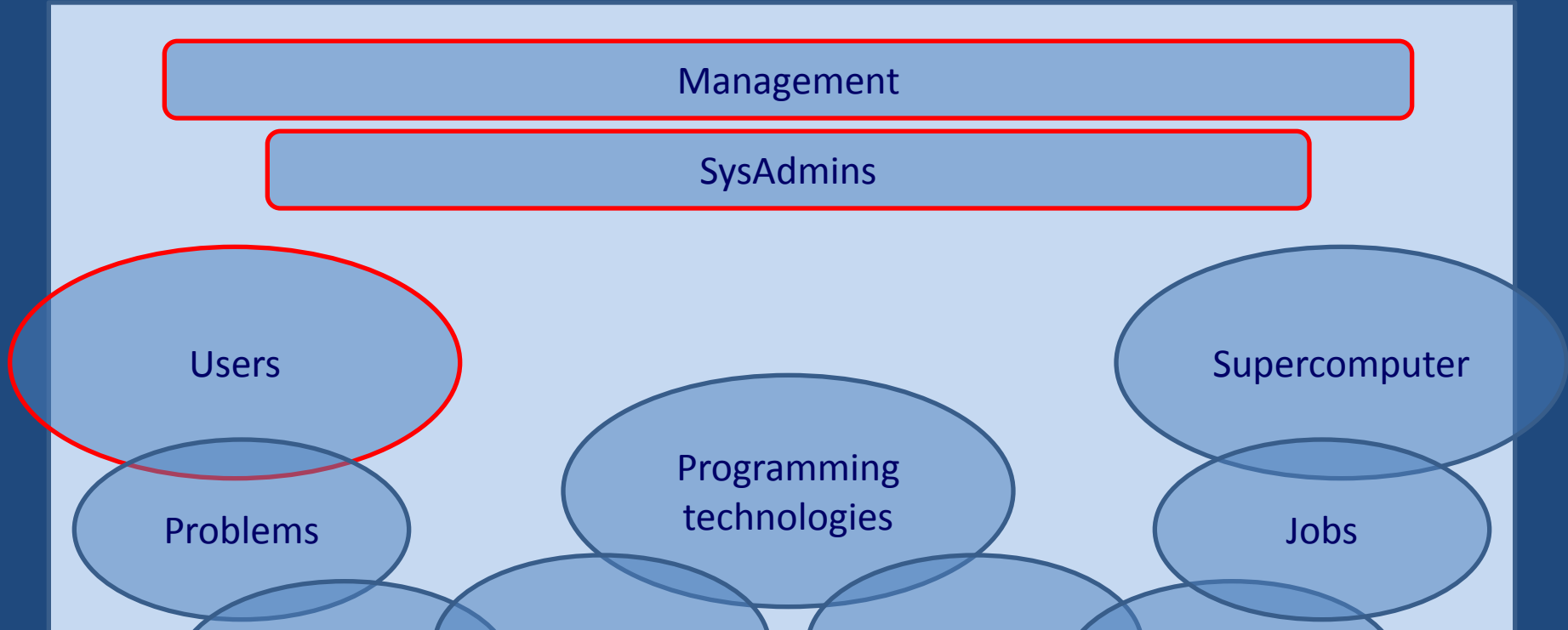
- *Research Computing Center, Moscow State University (Russian coordinator);*
- *T-Platforms;*
- *Joint Supercomputer Center, Russian Academy of Sciences;*
- *Scientific Research Institute of Multiprocessor Computer Systems, Southern Federal University.*

# Who cares about efficiency?



*Users, management, sysadmins: work at different scope, have different rights, make different decisions.*

# Who cares about efficiency?

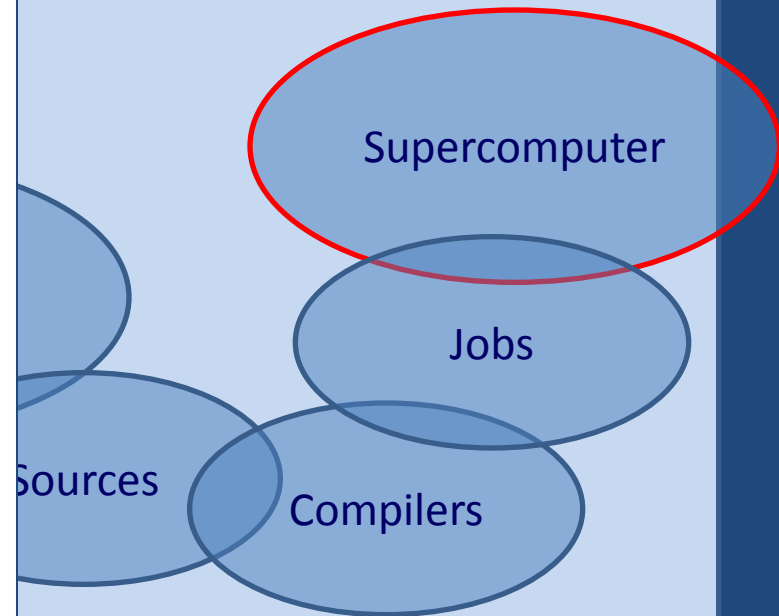


*Users, management, sysadmins: work at different scope,  
have different rights, make different decisions.  
Goal of the project is to provide a **total control over HW/SW**  
and applications for the target groups.*

# Holistic monitoring and analysis

*CPU usage (summary, and per-core) – user, system, irq, io, idle;*  
*Performance counters;*  
*Swap usage;*  
*Memory usage;*  
*Interconnect usage;*  
*Network errors;*  
*Disk usage;*  
*Filesystem usage;*  
*Network filesystem usage;*  
*Hardware alarms (ECC, SMART, etc);*  
*CPU and motherboard temperatures;*  
*FAN speeds;*  
*Voltages;*  
*Network switches errors;*  
*Cooling subsystem data;*  
*Power subsystem data;*

...



*ClustrX & LAPTA*

Efficiency and root cause analysis are the key points of the project



# Efficiency of applications

## Digest for job 189713

User:	aoganov	Command:	vasp
Nodes:	node-07-03,node-43-09	CPU count:	16
Started:	Wed May 2 18:40:14 2012	Queued:	Wed May 2 18:28:42 2012
Ended:	Wed May 2 18:51:53 2012	Wait time:	0 days 00:11:32
CPU*Hours:	3.10666666666667	Run time:	0 days 00:11:39

analyze\_cpu\_tpl7\_cpu\_1335970373476.csv

User CPU usage (%)



Highcharts.com

# Digest for job 13014

User: coctic                      Command: ./xhpl  
Nodes: cn02,cn15                CPU count: 6  
Started: Fri Apr 6 18:35:12 2012    Queued: Fri Apr 6 18:35:11 2012  
Ended: Fri Apr 6 22:35:14 2012    Wait time: 0 days 00:00:01  
CPU\*Hours: 24.0033333333333      Run time: 0 days 04:00:02

snl@vnc: /usr/bin/ld2\_0su\_1333737340001.dsu

## CPU usage for job



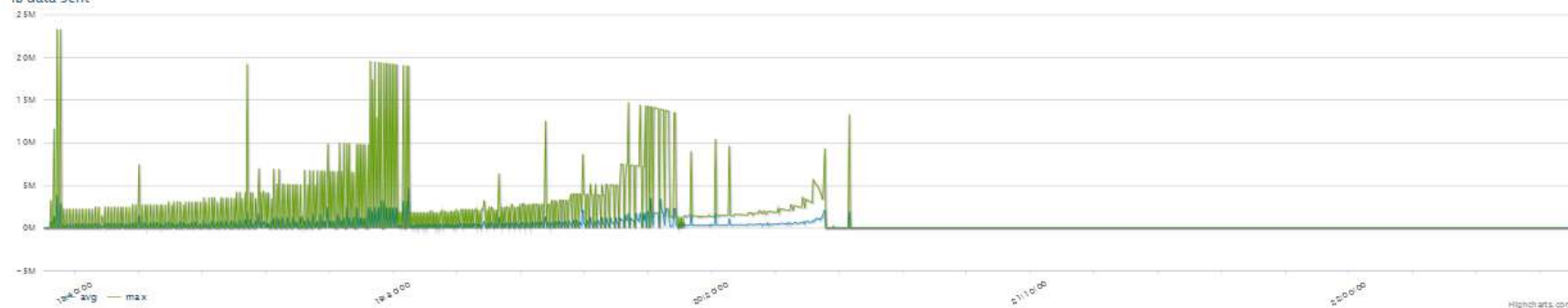
snl@vnc: /usr/bin/ld2\_0su\_1333737340001.dsu

## Free memory for job



snl@vnc: /usr/bin/ld2\_0su\_1333737340001.dsu

## IB data sent



# Efficiency of applications

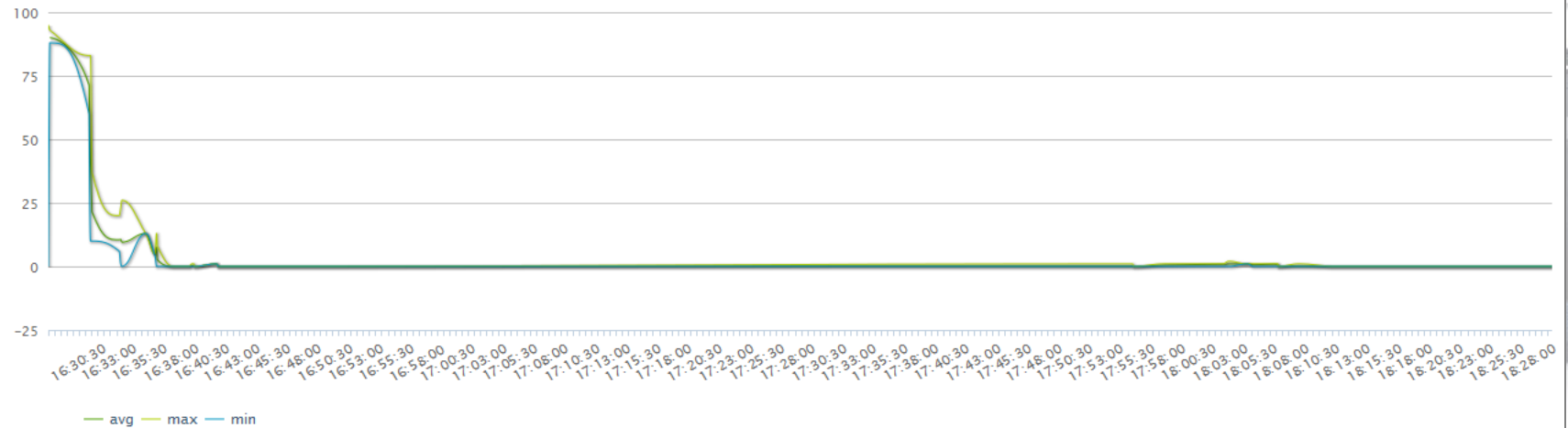
## Digest for job 189515

User: larin  
Nodes: node-11-01,node-30-01  
Started: Wed May 2 16:27:59 2012  
Ended: Wed May 2 18:31:04 2012  
CPU\*Hours: 32.822222222222

Command: /home/larin/CRYSTAL/NaCl/slab/100\_H2\_Na/Na/16a/Fixed\_13\_16/Fi/134/pcrystal  
CPU count: 16  
Queued: Wed May 2 16:16:22 2012  
Wait time: 0 days 00:11:37  
Run time: 0 days 02:03:05

analyze cpu tpl7 cpu 1335969124568.csv

### User CPU usage (%)



Highcharts.com

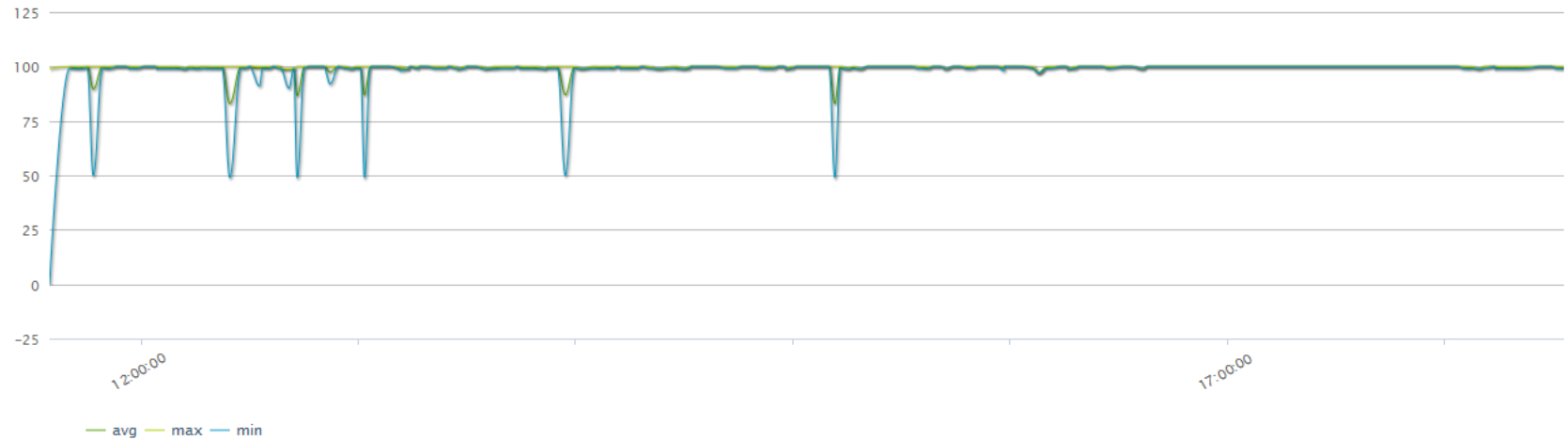
# Efficiency of applications

## Digest for job 188606

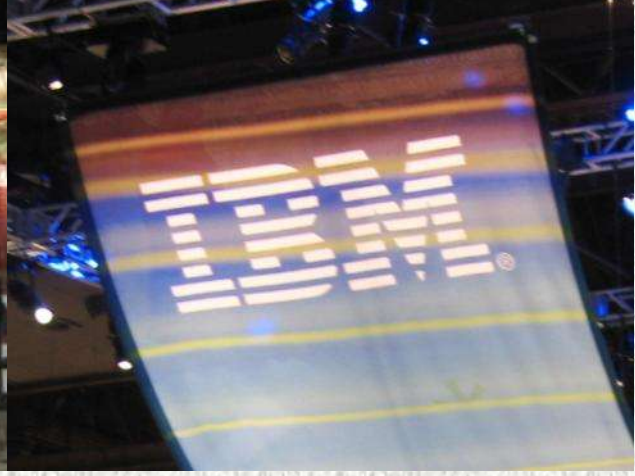
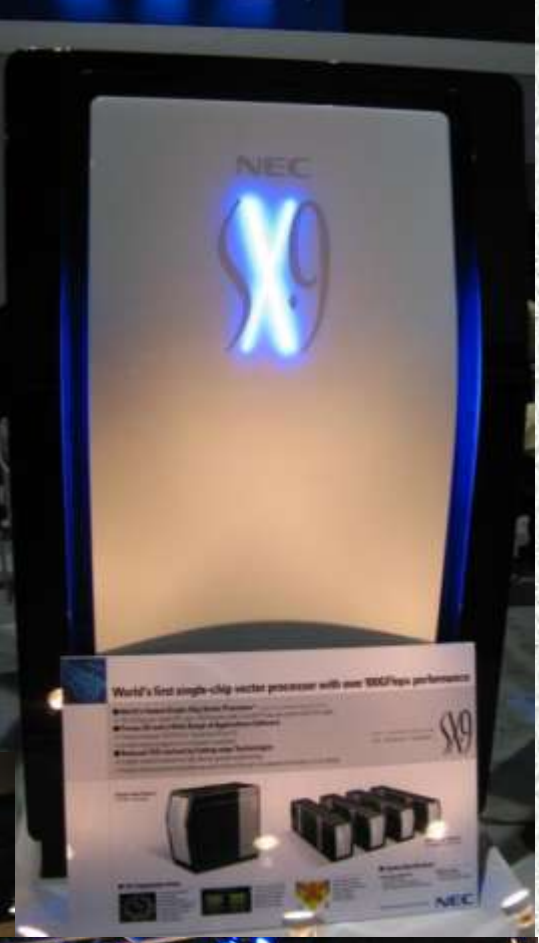
User:	ulybyshev	Command:	./calculator constants_calc.txt confB1.400e-01M3.000e-02Nx20Nt40.txt1
Nodes:	node-02-02,node-08-05	CPU count:	16
Started:	Tue May 1 20:20:50 2012	Queued:	Tue May 1 20:16:00 2012
Ended:	Wed May 2 18:33:34 2012	Wait time:	0 days 00:04:50
CPU*Hours:	355.395555555556	Run time:	0 days 22:12:44

analyze\_cpu\_tpl7\_cpu\_1335969274609.csv

### User CPU usage (%)



Highcharts.com

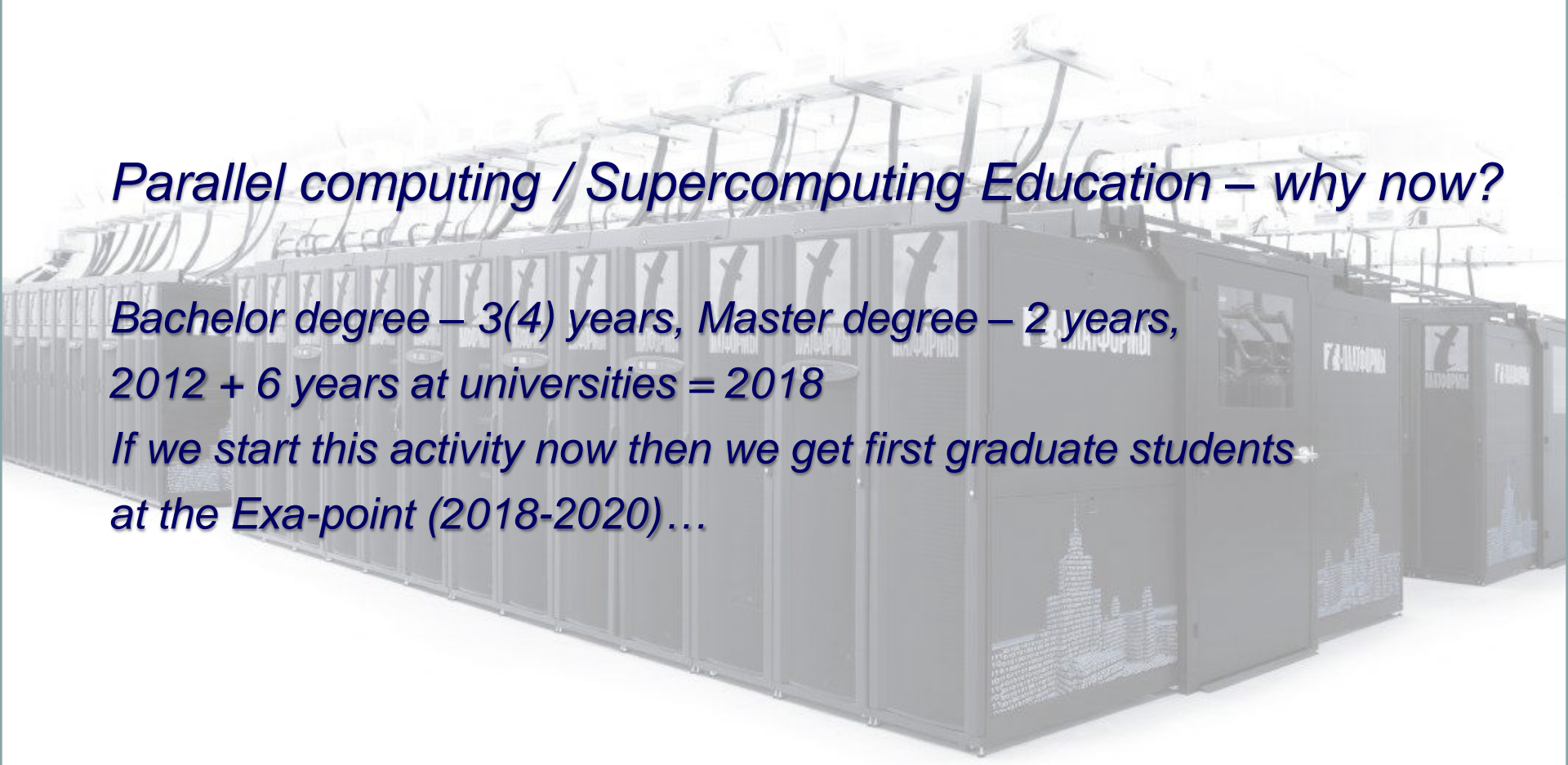


# *Education. Why now?*

*Parallel computing / Supercomputing Education – why now?*

*Bachelor degree – 3(4) years, Master degree – 2 years,  
2012 + 6 years at universities = 2018*

*If we start this activity now then we get first graduate students  
at the Exa-point (2018-2020)...*



# *Education. Why now?*

## *Parallel computing / Supercomputing Education – why now?*

*Bachelor degree – 3(4) years, Master degree – 2 years,  
2012 + 6 years at universities = 2018*

*If we start this activity now then we get first graduate students  
at the Exa-point (2018-2020):*

- Supercomputers – billions cores*
- Laptops – thousands cores*
- Mobile devices – dozens/hundreds cores*

*It is time to think about Parallel Computing...*

# Simple questions ?

(ask your students...)

- *What are potential bottlenecks/problems in a parallel code?*
- *What is parallel complexity of an algorithm? Why do we need to know a critical path of an informational graph?*
- *How to construct a communication free algorithm for a particular problem?*
- *How to detect and describe potential parallelism of an algorithm? How to extract potential parallelism from a code?*
- *How to estimate data locality in my application?*
- *How to estimate scalability of an algorithm and/or application? How to improve scalability of an application?*
- *How to express my problem in terms of Google's MapReduce model?*
- *How to solve a problem in a Condor environment?*
- *What parallel programming technology should I use for SMP/GPU/FPGA/vector/cluster/heterogeneous/distributed...*
- ...

*How many software developers will be able to use easily these notions?*



# *To Discuss, to Think about...*

- *Supercomputing Education*
- *Parallel Computing Education*
- *Computational Science & Engineering Education*
- *IT Education*

## *Remarks:*

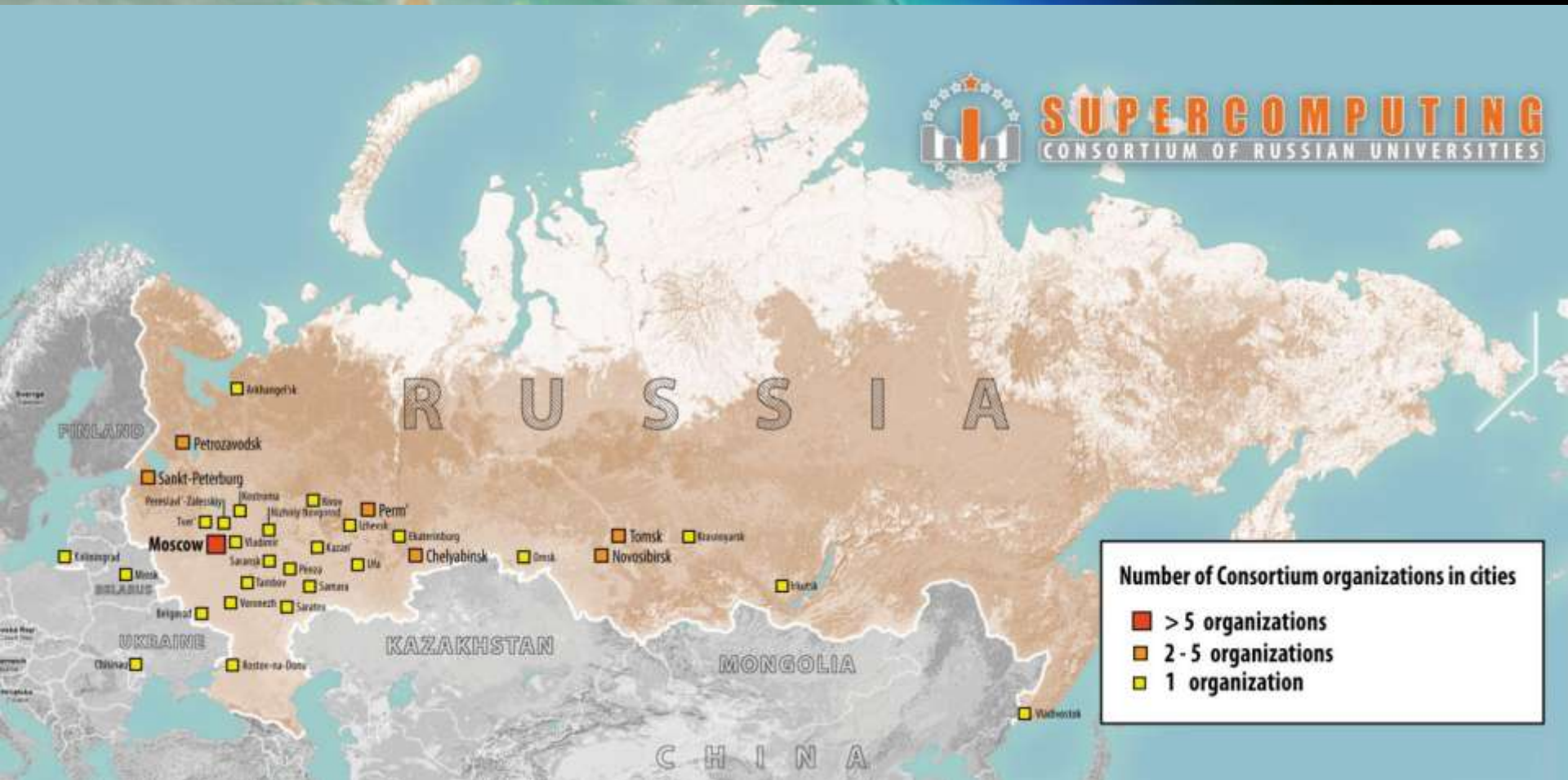
- *Supercomputing Today – Computing Tomorrow ...*
- *All our students will live in a “HyperParallel Computing World... How many students are ready for that?*
- *How many teachers are there in your countries which are able to teach Parallel Computing on a high level?..*

# *To Discuss, to Think about...*

- *Implementation: through national educational standards or other ways?*
- *Mass education (parallel computing) vs Individual (elite, supercomputing) education?*
- *Education or Training?*
- *Revolution or Evolution?*
- ...

- *Need for collaborative world-wide efforts.*

# Supercomputing Consortium of Russian Universities



2012: 50+ full and associated members



## *Project “Supercomputing Education”*

### *Commission for Modernization and Technological Development of Russia’s Economy*

*Duration: 2010-2012*

*Coordinator of the project: M.V.Lomonosov Moscow State University*

*Wide collaboration of universities:*

- Nizhny Novgorod State University*
- Tomsk State University*
- South Ural State University*
- St.Petersburg State University of IT, Mechanics and Optics*
- Southern Federal University*
- Far Eastern Federal University*
- Moscow Institute of Physics and Technology (State University)*
- members of Supercomputing Consortium of Russian Universities*

*More than 600 people from 63 universities were involved in the project in 2011.*

# *National System of Research and Education Centers on Supercomputing Technologies in Federal Districts of Russia*

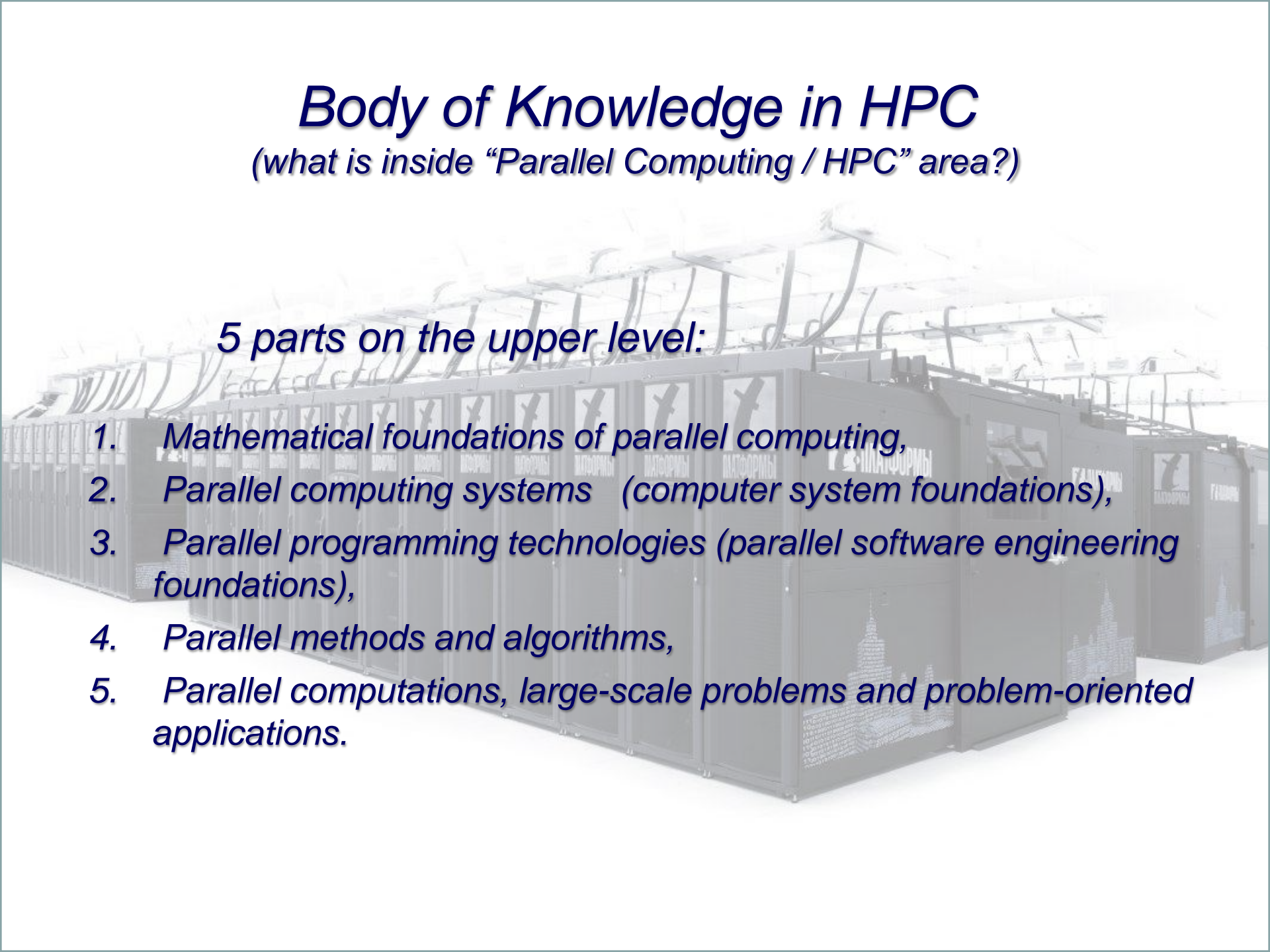


*8 centers were established in 7 federal districts of Russia  
during 2010-2011*

# *Body of Knowledge in HPC*

*(what is inside “Parallel Computing / HPC” area?)*

*5 parts on the upper level:*

- 1. Mathematical foundations of parallel computing,*
  - 2. Parallel computing systems (computer system foundations),*
  - 3. Parallel programming technologies (parallel software engineering foundations),*
  - 4. Parallel methods and algorithms,*
  - 5. Parallel computations, large-scale problems and problem-oriented applications.*
- 
- A large server rack in a data center, with the text overlaid on it. The rack is filled with server units, and the background shows a long aisle of similar racks. The text is in a dark blue, italicized font.

# Informational Structure is a Key Notion

(matrix multiplication as an example)

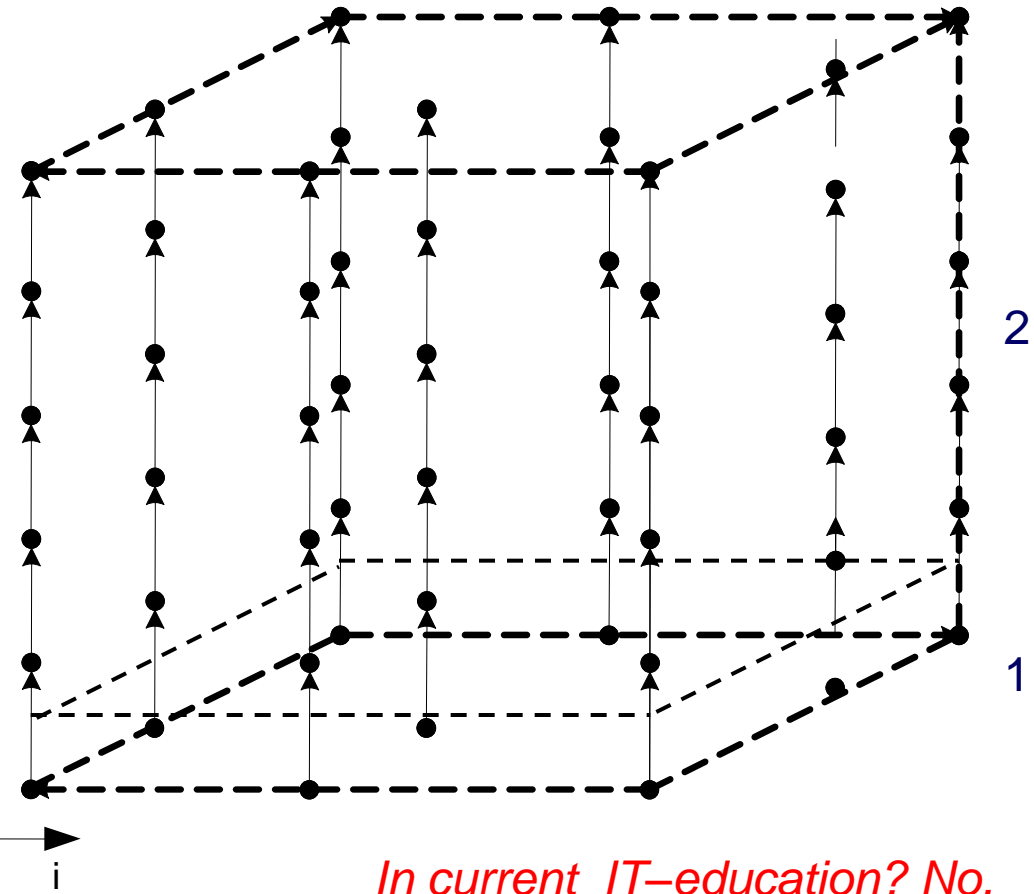
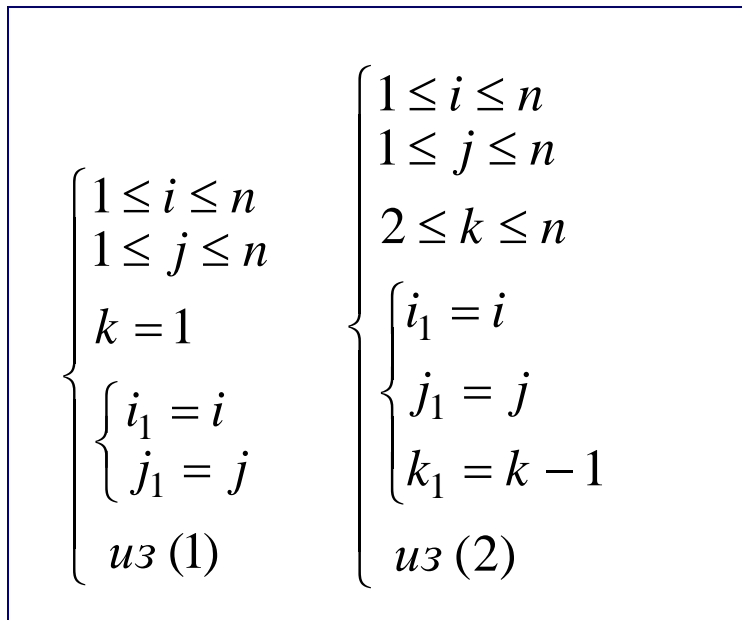
Do  $i = 1, n$

Do  $j = 1, n$

1  $A(i,j) = 0$

Do  $k = 1, n$

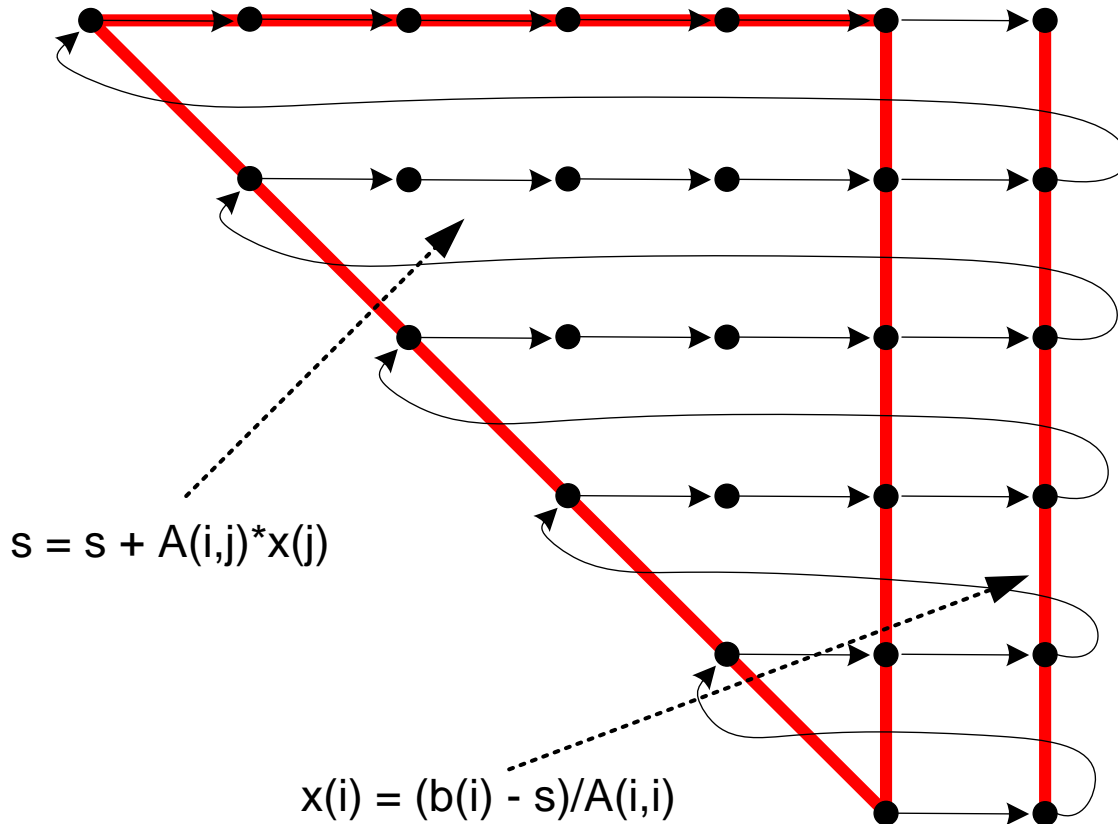
2  $A(i,j) = A(i,j) + B(i,k)*C(k,j)$



In current IT-education? No.

# GAUSS elimination: method and algorithm

(informational structure)



```
do i = n, 1, -1
  s = 0
  do j = i+1, n
    s = s + A(i,j)*x(j)
  end do
  x(i) = (b(i) - s)/A(i,i)
end do
```

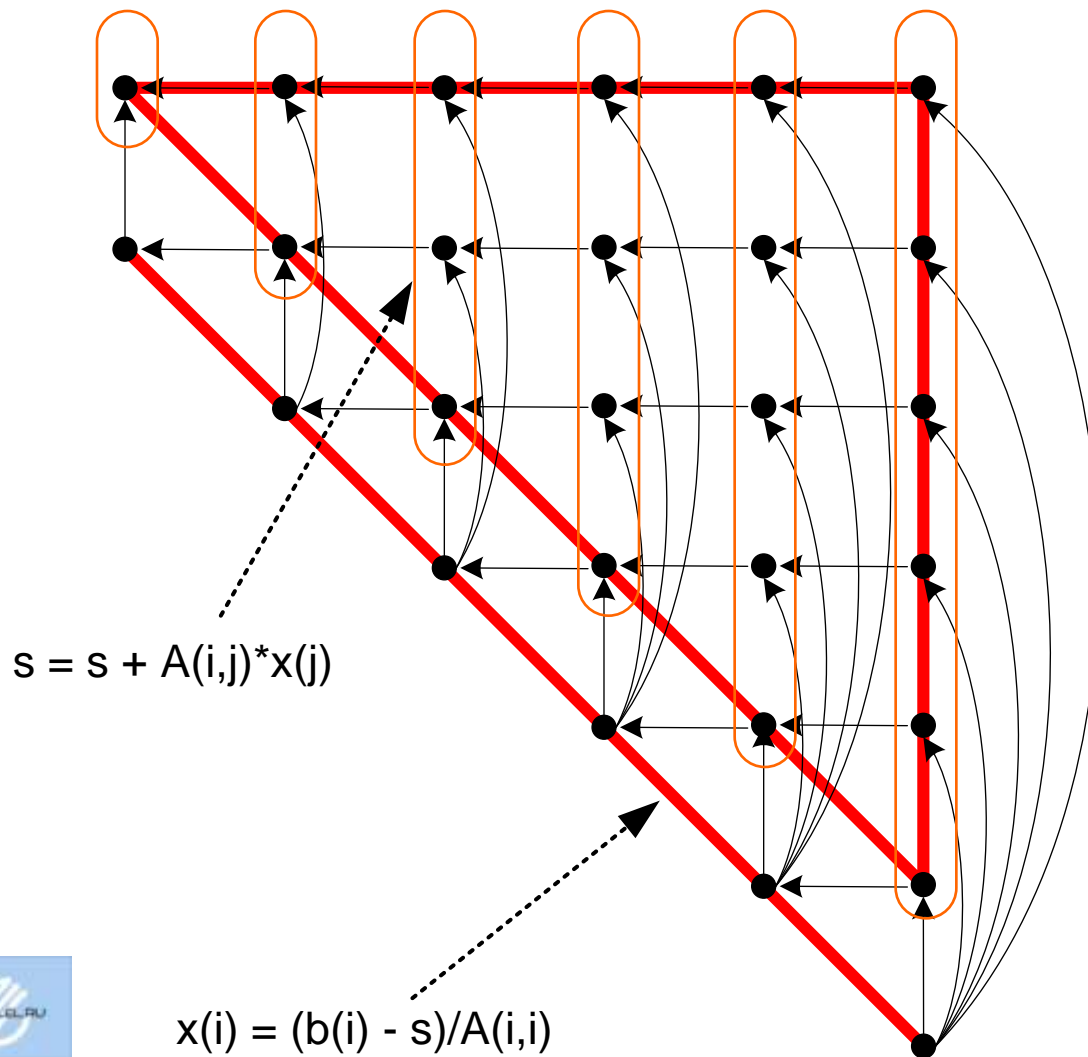


*In current IT-education? No.*



# GAUSS elimination: method and algorithm

(informational structure)



```
do i = n, 1, -1
  s = 0
  do j = n, i+1, -1
    s = s + A(i,j)*x(j)
  end do
  x(i) = (b(i) - s)/A(i,i)
end do
```

In current IT-education? No.

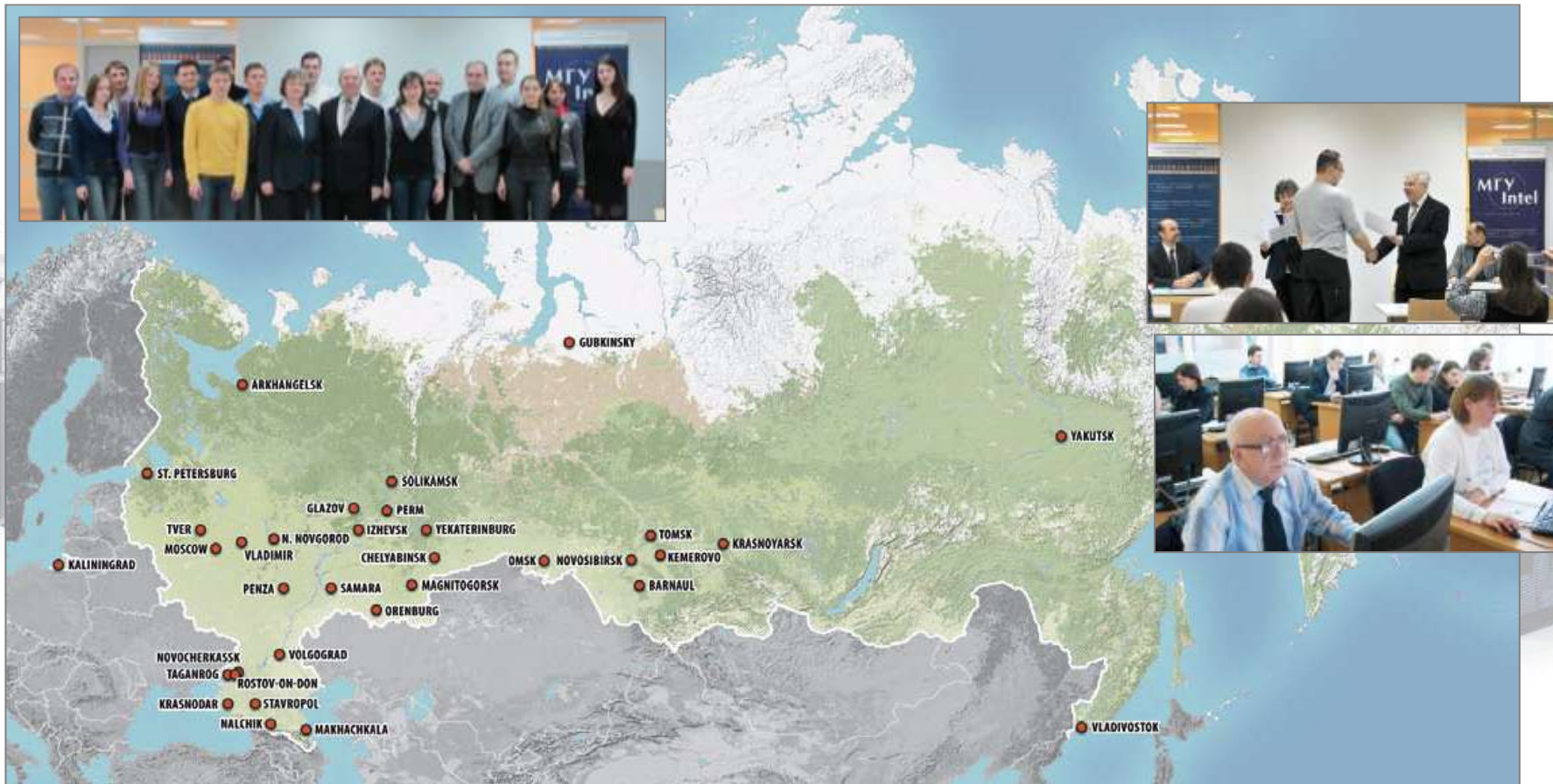


# *Entry-level Training on Supercomputing Technologies*



*1824 people passed trainings, 45 universities from 35 cities of Russia*

# Retraining Programs for Faculty Staff



**166** faculty staff passed trainings, **43** organisations, **29** cities,  
**8** education programs,  
All federal districts of Russia were presented.

# *Intensive Trainings in Special Groups*



**18** special groups of trainees were formed,  
**427** trainees successfully passed advanced retraining,  
**14** educational programs,  
All federal districts of Russia were presented.

# *IT-Companies + Research Institutes & Edu*

*(special group of students on Parallel Software Development)*

*October, 24 – November, 14, 2011*

*55 students of MSU (Math, Physics, Chemistry, Biology, ...)*

*Moscow State University in collaboration with:*

- *Intel*
- *T-Platforms*
- *NVIDIA*
- *TESIS*
- *IBM*
- *Center on Oil & Gas Research*
- *Keldysh Institute of Applied Mathematics, RAS*
- *Institute of Numerical Mathematics, RAS*

# Series of Books “Supercomputing Education”



“Computational Mathematics  
and Algorithm`s Structure”  
V.V. Voevodin

“High-Performance Computations  
for Multiprocessor  
Multi-Core Systems”  
V.P. Gergel

“Practical Course on Parallel  
Computing Techniques”  
A.V. Starchenko, E.A. Daniilkin,  
V.I. Laeva, S.A. Prokhanov

“Parallel Programming Tools in  
Shared-Memory Systems”  
K.V. Korniyakov, V.D. Kustikova,  
I.B. Meyerov, A.A. Sydnev,  
A.V. Sysoev, A.V. Shishkov

“Parallel Programming  
Technologies for New  
Architecture Processors”  
A.V. Linev, D.K. Bogolepov,  
S.I. Bastrakov

There are **25+** books in “Supercomputing Education” series.

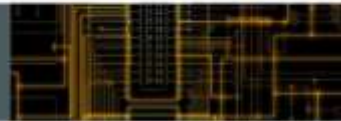
**7.000** books were delivered in **43** universities in 2011.

# Series of Books “Supercomputing Education”



Е. А. Гречников, С. В. Михайлов,  
Ю. В. Нестеренко, И. А. Поповян

Вычислительно сложные  
задачи теории чисел



В. Н. Лыкозов, А. В. Глазун, Д. В. Кузьмин,  
Е. В. Мортиков, В. М. Степаненко

Суперкомпьютерное  
моделирование в физике  
климатической системы

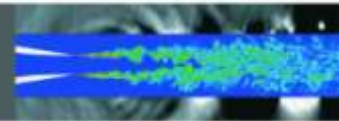


Параллельные вычисления  
на GPU: архитектура  
и программные модели



В. М. Головизнин, М. А. Зайцев,  
С. А. Карabasов, И. А. Короткин

Новые алгоритмы  
вычислительной гидродинамики  
для многопроцессорных  
вычислительных комплексов



А. С. Антонов

Технологии параллельного  
программирования  
MPI и OpenMP



“Computationally Intensive  
Problems of Numbers Theory”  
E.A. Grechnikov, S.V. Michailov,  
Y.V. Nesterenko, I.A. Popovyan

“Parallel Computing on GPU:  
Architecture and  
Programming Models”

“Parallel Programming  
Technologies MPI and OpenMP”  
A.S. Antonov

“Supercomputing Modeling in  
Climate System Physics”  
V.N. Lykosov, A.V. Glazunov,  
D.V. Kulyamin, E.V. Mortikov,  
V.M. Stepanenko

“New Computational Fluid  
Dynamics Algorithms for  
Parallel Computers”  
V.M. Goloviznin, M.A. Zaitsev,  
S.A. Karabasov, I.A. Korotkin

More than **30.000** books of the series will be delivered to **43** universities this year.

# Courses on Supercomputing Technologies

Development of new courses and extension of existing ones...

40+ courses covering all major parts of the **Body of Knowledge in SC**...

- "Parallel Computing",
- "High Performance Computing for Multiprocessing Multi-Core Systems",
- "Parallel Database Systems",
- "Practical Training on MPI and OpenMP",
- "Parallel Programming Tools for Shared Memory Systems",
- "Distributed Object Technologies",
- "Scientific Data Visualization on Supercomputers",
- "Natural Models of Parallel Computing",
- "Solution of Aero- and Hydrodynamic problems by Flow Vision",
- "Algorithms and Complexity Analysis",
- "History and Methodology of Parallel Programming",
- "Parallel Numerical Methods",
- "Parallel Computations in Tomography",
- "Final-Element Modeling with Distributed Computations",
- "Parallel Computing on CUDA and OpenCL Technologies",
- "Biological System Modeling on GPU",
- "High Performance Computing System: Architecture and Software",
- ...



# Summer Supercomputing Academy

at Moscow State University

June,25 – July,7

- *Plenary lectures by prominent scientists, academicians, CEO/CTO's from Russia and abroad,*
- *9 independent educational tracks,*
- *Trainings from Intel, IBM, NVIDIA, T-Platforms, Mellanox, RogueWave, Accelrys, ...*
- *120 attendees were selected (from students up to professors).*



# *Informatics Europe & HPC-Education*

*New working group within Informatics Europe  
(<http://informatics-europe.org/>):*

*“Parallel Computing (Supercomputing) Education in Europe: State-of-Art”  
- about 20 members from 10 countries.*

## *Nearest Goals:*

- to show the need for urgent changes in higher education in the area of computational sciences,*
- to compose a survey of the current landscape of parallel computing and supercomputing education in Europe with respect to different universities and countries,*
- to prepare a set of recommendations how to bring ideas of parallel computing and supercomputing into higher educational systems of European countries.*

*Join us! Write to [wg-hpc-edu@parallel.ru](mailto:wg-hpc-edu@parallel.ru)*



*47<sup>th</sup> HPC User Forum  
Using HPC to Drive Economic and Scientific Competitiveness*

# *High Performance Computing at Moscow State University and more...*

*Prof. Vladimir Voevodin  
Deputy Director, Research Computing Center,  
Moscow State University*

*[voevodin@parallel.ru](mailto:voevodin@parallel.ru)*

*July, 9, 2012*

*HLRS / University of Stuttgart, Stuttgart, Germany*