High Performance Computing at Moscow State University and more…

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

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 m²
Power consumption: 150 KW
12 years ago ...

(24 CPUs, Intel P-III/500 MHz, SCI network, 8 m², 12 Gflops)
... and now

(52,000+ Intel cores, 2,130 NVIDIA GPUs, QDR IB, 1200 m², 1.7 Pflops)
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

http://top50.supercomputers.ru

TOTAL PEAK PERFORMANCE:
- > 1.5 PFlops
- 18 TFlops — 1.5 PFlops
- 4 TFlops — 18 TFlops

Research Computing Center, MSU
Joint Supercomputer Center, RAS
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, 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 m²
- Power Consumption (supercomputer): 2.7 MW
## MSU “Lomonosov” supercomputer, 2012

(node types)

<table>
<thead>
<tr>
<th>Node types</th>
<th>RAM per node</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 x Xeon 5570 2.93 GHz</td>
<td>12 GB</td>
<td>4160</td>
</tr>
<tr>
<td>2 x Xeon 5570 2.93 GHz</td>
<td>24 GB</td>
<td>260</td>
</tr>
<tr>
<td>2 x Xeon 5670 2.93 GHz</td>
<td>24 GB</td>
<td>640</td>
</tr>
<tr>
<td>2 x Xeon 5670 2.93 GHz</td>
<td>48 GB</td>
<td>40</td>
</tr>
<tr>
<td>2 x PowerXCell 8i 3.2 GHz</td>
<td>16 GB</td>
<td>30</td>
</tr>
<tr>
<td>2 x Xeon E5630 2.53 GHz, 2 x Tesla X2070</td>
<td>12 GB</td>
<td>777</td>
</tr>
<tr>
<td>2 x Xeon E5630 2.53 GHz, 2 x Tesla X2070</td>
<td>24 GB</td>
<td>288</td>
</tr>
<tr>
<td>4 x Xeon E7650 2.26 GHz</td>
<td>512 GB</td>
<td>4</td>
</tr>
</tbody>
</table>
### MSU Supercomputing Center
(users & organizations)

<table>
<thead>
<tr>
<th></th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>User groups, total:</td>
<td>241</td>
<td>369</td>
<td>545</td>
</tr>
<tr>
<td>including:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>from Moscow University:</td>
<td>155</td>
<td>241</td>
<td>359</td>
</tr>
<tr>
<td>from institutes of RAS:</td>
<td>53</td>
<td>77</td>
<td>110</td>
</tr>
<tr>
<td>from other organizations:</td>
<td>33</td>
<td>51</td>
<td>76</td>
</tr>
<tr>
<td>Faculties / Institutes of MSU:</td>
<td>15</td>
<td>21</td>
<td>24</td>
</tr>
<tr>
<td>Institutes of RAS:</td>
<td>20</td>
<td>28</td>
<td>35</td>
</tr>
<tr>
<td>Others:</td>
<td>19</td>
<td>24</td>
<td>34</td>
</tr>
</tbody>
</table>
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: \(1 \text{Pflop} \times 60\text{sec} \times 60\text{min} \times 24\text{hours} \times 365\text{days} = 31.5 \text{ZettaFlop per year}\)

What is in reality? 0.0...x% 

Why? Peculiarities of hardware, a complicated job-flow, poor data locality, a huge degree of parallelism in hardware, etc…
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 – HOlistic Performance System Analysis**

**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.
Efficiency and root cause analysis are the key points of the project
Efficiency of applications

Digest for job 189713

User: aoganov
Nodes: node-07-03,node-43-09
Command: vasp
CPU count: 16

Started: Wed May 2 18:40:14 2012
Queued: Wed May 2 18:28:42 2012

Wait time: 0 days 00:11:32
Run time: 0 days 00:11:30

CPU*Hours: 3.10666666666667

User CPU usage (%)

avg max min
Efficiency of applications

Digest for job 189515

User: larin
Nodes: node-11-01,node-30-01
Command: /home/larin/CRYSTAL/NaCl/slab/100_H2_Na/Na16a/Fixed_13_16/Fv134/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

CPU*Hours: 32.822222222222222

User CPU usage (%)

-25 -10 0 10 20 30 40 50 60 70 80 90 100

avg max min
Efficiency of applications

Digest for job 188606

- User: ulybyshev
- Nodes: node-02-02, node-08-05
- Started: Tue May 1 20:20:50 2012
- Ended: Wed May 2 18:33:34 2012
- CPU*Hours: 355.395555555556
- Command: /calculator constants_calc.txt confB1_400e-01M3 000e-02Nx20N40.txt1
- CPU count: 16
- Queued: Tue May 1 20:16:00 2012
- Wait time: 0 days 00:04:50
- Run time: 0 days 22:12:44

User CPU usage (%)

- avg, max, min line graph
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.
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. $A(i,j) = A(i,j) + B(i,k) \times C(k,j)$

---

In current IT-education? No.
GAUSS elimination: method and algorithm
(informational structure)

\[
\begin{align*}
&\text{do } i = n, 1, -1 \\
&s = 0 \\
&\text{do } j = i+1, n \\
&s = s + A(i,j) \cdot x(j) \\
&\text{end do} \\
&x(i) = (b(i) - s) / A(i,i) \\
&\text{end do}
\end{align*}
\]

In current IT-education? No.
**GAUSS elimination: method and algorithm**
*(informational structure)*

\[
s = s + A(i,j) \cdot x(j)
\]

\[
x(i) = \frac{b(i) - s}{A(i,i)}
\]

\[
\text{do } i = n, 1, -1
\]
\[
s = 0
\]
\[
\text{do } j = n, i+1, -1
\]
\[
s = s + A(i,j) \cdot x(j)
\]
\[
\text{end do}
\]
\[
x(i) = \frac{b(i) - s}{A(i,i)}
\]
\[
\text{end do}
\]

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

7,000 books were delivered in 43 universities in 2011.
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",
- ...
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).
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
High Performance Computing at Moscow State University and more…

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

voevodin@parallel.ru

July, 9, 2012
HLRS / University of Stuttgart, Stuttgart, Germany