# FlowVision Scalability on Supercomputers with Angara Interconnect

**V. S. Akimov,[1, *] D. P. Silaev,[1, **] A. A. Aksenov,[2, ***]**
**S. V. Zhluktov,[2, ****] D. V. Savitskiy,[2, *****] and A. S. Simonov[3, ******]**

(Submitted by V. V. Voevodin)

[1] *Numerical Engineering Platform LLC, 125373, Moscow, Skolkovo Innovation Center, Nobelja str. 7*
[2] *Joint Institute for High Temperatures RAS, 125412, Moscow, Izhorskaya str. 13*
[3] *Scientific Research Centre for Electronic Computer Technology (NICEVT) JSC, 117587, Varshavskoe sh., 125, Moscow,*

Received February 11, 2018

**Abstract**—Scalability of computations in the FlowVision CFD software on the Angara-C1 cluster equipped with the Angara interconnect is studied. Different test problems with 260 thousand, 5.5 million and 26.8 million computational cells are considered. Computations in FlowVision are performed using a new solver of linear systems based on the algebraic multigrid (AMG) method. It is shown that the the special FlowVision's Dynamic balancing technology significantly improves performance of computations if features of the problem lead to the non-uniform loading of CPUs. The Angara-C1 cluster demonstrates excellent performance and scalability characteristics comparable with its analogues based on the 4x FDR Infiniband interconnect.

**2010 Mathematical Subject Classification:** 12345, 54321

Keywords and phrases: *scalability*, *FlowVision*, *CFD*, *gasdynamics*, *cluster*, *supercomputer*, *interconnect*, *Angara*

## 1. INTRODUCTION

Supercomputers are permanently being developed: the number of the processor cores grows, the memory capacity increases, the interconnect is being improved. The hardware and software engineers try to provide maximum performance and efficiency of multi-processor clusters [1, 2].

Modern cluster systems most often use Mellanox Infiniband and Intel OmniPath commercial interconnects. Interconnect solutions like IBM BlueGene/Q, Tofu are manufactured for particular series of supercomputers and are not supplied separately. Also, European interconnect developments, first of all, Bull Exascale Interconnect (BXI) [3] and Extoll [4] are of interest.

Angara is the first Russian high-speed interconnect [5, 6] based on router ASIC. The router ASIC was developed by JSC NICEVT and is manufactured using 65 nm technology. Angara interconnect supports multi-dimensional torus topology (from 1D- to 4D-torus), remote direct memory access (RDMA), GPUDirect technology and all the standard programming tools (MPI library, OpenMP technology, SHMEM library, TCP/IP protocol stack). Angara interconnect is compatible with x86, Elbrus, ARM processors as well as with GPU- and FPGA-based accelerators. Currently two computer clusters use the Angara interconnect: the 32-node Desmos hybrid cluster installed at JIHT RAS (4D-torus 4x2x2x2 topology) [7] and 36-node Angara-C1 cluster, installed at JSC NICEVT (3D- torus

* E-mail: mailto:akimov@flowvision.ru
** E-mail: silaev@tesis.com.ru
*** E-mail: andrey@tesis.com.ru
**** E-mail: sz@flowvision.ru
***** E-mail: dmvlsav@yandex.ru
****** E-mail: simonov@nicevt.ru

4x3x3 topology). Results of the Angara-C1 cluster evaluative testing using OSU micro-benchmarks, Intel MPI Benchmarks and NAS Parallel Benchmarks are presented in the article [8].

Multi-processor calculations are required to solve modern industrial CFD (Computational Fluid Dynamics) problems. Performance is mostly determined by the capability of a multi-processor system to speed up calculations by using a big number of cores and processors. This capability is called the scalability of computations. It strongly depends on the numerical methods and algorithms implemented in the code, since they affect the efficiency of using memory and interconnect.

The evaluation results of the FlowVision software scalability are presented in the given article. Three fluid mechanic problems were solved. The calculations have been performed on the Angara-C1 cluster with Angara interconnect. The scalability obtained on this supercomputer is compared to the scalability achieved on supercomputers with Infiniband 4xFDR interconnect.

## 2. FLOWVISION SOFTWARE

FlowVision is a multipurpose software designed for simulation of 3D flows of liquids and gases in and around technical and natural objects. It contains different instruments for visualization of the flows using the technologies of computer graphics [9]. The simulated flows can be steady and unsteady, compressible and incompressible. The FlowVision solver is based on the methods of computational fluid dynamics (CFD) developed within the finite volume approach. The finite volume approach assumes partitioning the computational region into small contacting volumes (computational cells). The computational grid built by the FlowVision grid generator is Cartesian. Hence, the computational cells are hexahedrons. The grid generator provides the possibility to locally split the cells in the regions where more accurate resolution of the geometry or the gradients of the sough-for variables is required. Such an adaptation can be performed in a volumetric object, on a surface and in the entire computational domain according to the specified criteria of adaptation to the developing solution (e. g., to the pressure gradient). The adaptation rules are specified in the FlowVision interface. The adaptation itself is performed automatically. The continuity, momentum, energy and other equations are integrated numerically on the locally adapted grid using the CFD methods.

Implicit numerical schemes implemented in FlowVision assume solving sparse systems of linear algebraic equations (SLAE). The systems must be solved with high accuracy on the computers with distributed memory. Solving SLAE consumes big portions of entire computing time and RAM. The choice of effective method for solving SLAE is an important problem, since it can reduce the simulation time. Three different SLAE solvers are implemented in FlowVision: algebraic multi-grid method (AMG) with aggregative coarsening, AMG with selective coarsening, and TParFBSS. The latter includes preconditioning with incomplete triangular factorization and iterative scheme of the Krylov type. Particular SLAE solver is selected with use of technology AST (Aggregative AMG – Selective AMG – TParFBSS). The following strategy of calculations at one time step is implemented in FlowVision. First, Aggregative AMG works. If it manages the SLAE, the program proceeds to the next time step. If not, Selective AMG (more expensive solver) continues solution of the system. If it does not manage the SLAE, TParFBSS (the most expensive solver) calculates the system obtained from Selective AMG. If Selective AMG or TParFBSS have been used at the given time step, the next 10 steps are performed with the same solver. At the 11th step, again the cheapest solver (Aggregative AMG) is tried. In all the calculations presented in the current paper, Aggregative AMG was used.

In order to reduce the time of solving a hydro- or gas-dynamic problem, it is necessary to parallelize computations according to the architecture of the used hardware. State-of-the-art clusters have architecture with distributed memory: on the one hand, they have a set of computational nodes, which exchange data by interconnect. On the other hand, each node represents a multi-processor (multi-socket) server with common access to RAM. Within one socket, the access to RAM is usually uniform (UMA), whereas the access to the memory in a neighbor socket is non-uniform (NUMA). Therefore, the hybrid approach to parallelizing computations is implemented in FlowVision. It combines the advantages of parallelization between processors using MPI with parallelization between threads [10] – see Figure 1. Uniform access to the memory within one processor makes advantageous parallelization between threads. Parallelization between processors requires using MPI library. However, sometimes using more than one MPI process per processor may increase the efficiency of computations. This depends on the architecture of the interconnect (for instance, processors contain large numbers of cores)

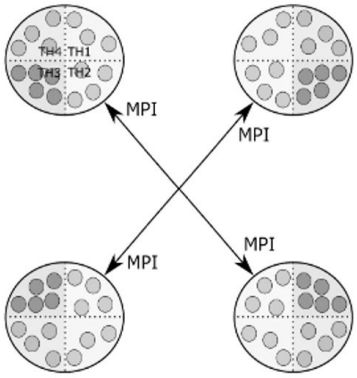



**Figure 1.** Hybrid parallel architecture of FlowVision

and the problem characteristic. Thus, FlowVision provides flexible possibility for specifying required combinations of MPI-processes and threads.

The spectrum of the problems solved by FlowVision is very wide. As a consequence, a lot of possibilities occur for distribution of the computational cells over processors. Many industrial hydro- and gas-dynamic problems imply creation of complex geometrical models. Many problems require good resolution of the gradients of physical quantities by computational grid. A computational domain may contain big non-computed volumes, e.g., inner parts of solid bodies. It is often necessary to fine the grid near solid surfaces. A boundary-layer grid can be built at the surface of a solid body. A solid body can move through the computational domain, causing dynamic rebuilding the grid during calculations. The computational domain can be bounded by a free surface changing its positions at each time step. This also causes dynamic rebuilding the grid. The amount of computations in the cells near a solid body, near a free surface, and in the boundary-layer grid differs from that in the ordinary cells. Therefore, regular balancing of computations assuming distribution of equal number of cells over processes is insufficient. In order to solve the problem, special tool called Dynamic balancing has been developed and implemented in FlowVision. It enables essential scaling of multiprocessor calculations by means of redistribution of cells between MPI-processes. It should be emphasized that the tool does not simply balance the number of cells processed by each MPI-process. That could be done in advance. The tool balances the time spent by the processes. This can be done only during calculations.

## 3. TEST PROBLEMS AND RUN OPTIONS

The mentioned Angara-C1 cluster and two other supercomputers were used in the scalability evaluation. Table 1 provides the configuration details of the systems.

### *3.1. Test problems*

Three fluid dynamic problems with different number of cells, kind of physical process, with/without adaption have been selected for the evaluation. The characteristics of the problems are given in Table 2.

The air flow in a cavity is a well-known test case M219 Cavity case [11], see Figure 2. It represents the group of external problems characterized by simple geometrical shape of an object. Typically, computational grids in such problems are locally fine-grained (Figure 2a). However, the distribution of the computational cells over the MPI-processes can be performed in such a way that the difference of the processor utilization does not exceed 20%.

On the contrary, the Airplane flow problem (Figure 3) may produce essential non-uniformity in the processor utilization (more than 50%) due to non-uniformity of the computational grid, since the geometry is complex (Figure 3a), the grid is adapted to the airplane surface (Figure 3b), and the boundary layer grid is used in these particular calculations.

Problem Mixer (Figure 4) is characterized by relatively small number of computational cells. It provides perfect uniformity in loading of processors.

**Table 1.** Supercomputer configuration details

| Supercomputer | Angara-C1 (partition A) | Lomonosov-2 (compute partition) | Z cluster |
|---|---|---|---|
| Processor | Intel Xeon E5-2630, 2.30 GHz | Intel Xeon E5-2697v3, 2.6 GHz | Intel Xeon E5-2670, 2.6 GHz |
| Number of sockets | 2 | 1 | 2 |
| Number of physical cores in processor | 6 | 14 | 8 |
| Number of logical cores in processor when Hyper-Threading (HT) is used | 12 | 28 | HT is disabled |
| Cache-memory, MB | 15 | 35 | 20 |
| Maximum memory bandwidth, GB/s | 42.6 | 68 | 51.2 |
| Memory size per node, GB | 64 | 64 | 64 |
| Interconnect | Angara (3D-torus 4x3x3) | Mellanox FDR InfiniBand (56 Gbit/s) | FDR InfiniBand (56 Gbit/s) |
| MPI implementation | MPICH 3.0.4 | OpenMPI 1.8.4 | Intel MPI 5.1 |

**Table 2.** Test problems

| Test problems | M219 Cavity case, Figure 2 | Airplane flow, Figure 3 | Mixer, Figure 4 |
|---|---|---|---|
| Grid | 3D | 3D | 3D |
| Physical processes | Heat transfer, motion | Heat transfer, motion, turbulence | Heat transfer, motion, turbulence |
| Number of cells | 5.5 million | 26.8 million | 260 thousand |
| Boundary layer grid | None | None | None |
| Adaptaion | in local volume | at solid surface | None |

### *3.2. Mathematical model*

The mathematical model used in the calculations includes the following equations. The continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla (\rho \mathbf{V}) = 0.$$

Here $t$ is time, $\rho$ is fluid density, $\mathbf{V}$ is flow velocity. The momentum equation:

$$\frac{\partial \rho \mathbf{V}}{\partial t} + \nabla (\rho \mathbf{V} \otimes \mathbf{V}) = -\nabla p + \nabla \cdot \hat{\tau},$$

$$\hat{\tau} = (\mu + \mu_t) \left( 2\hat{\mathbf{S}} - \frac{2}{3} (\nabla \cdot \mathbf{V}) \hat{\mathbf{I}} \right), \quad S_{ij} = \frac{1}{2} \left( \frac{\partial V_i}{\partial x_j} + \frac{\partial V_j}{\partial x_i} \right), \quad \mu_t = \rho C_\mu k^2 / \varepsilon..$$

Here $p$ is static pressure, $\hat{\tau}$ is shear stress tensor, $\mu$ is dynamic coefficient of molecular viscosity, $\mu_t$ is dynamic coefficient of turbulent viscosity (in the 1st test case it is zero), $\hat{\mathbf{S}}$ is deformation rate tensor, $\hat{\mathbf{I}}$

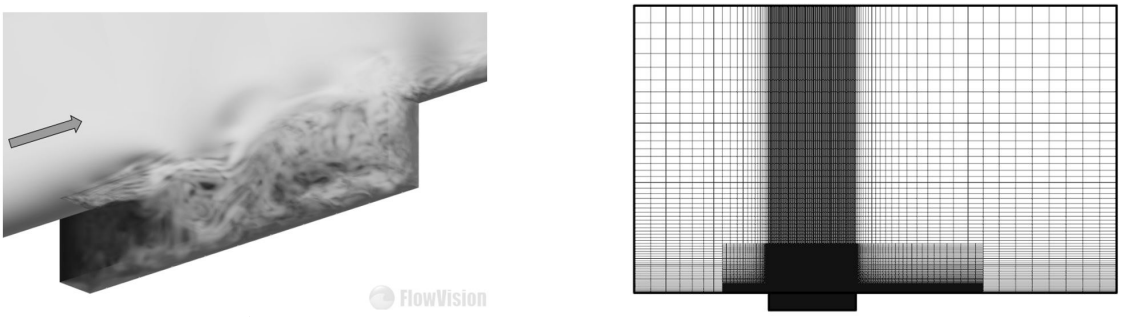



**Figure 2.** M219 Cavity case: a — volumetric visualization of velocity; b — computational grid

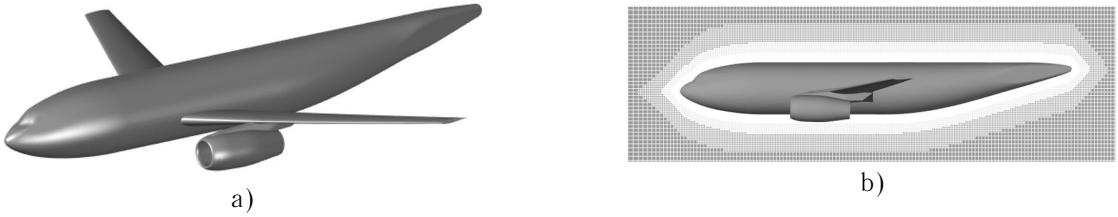



**Figure 3.** Airplane flow: a — airplane geometry; b — computational grid

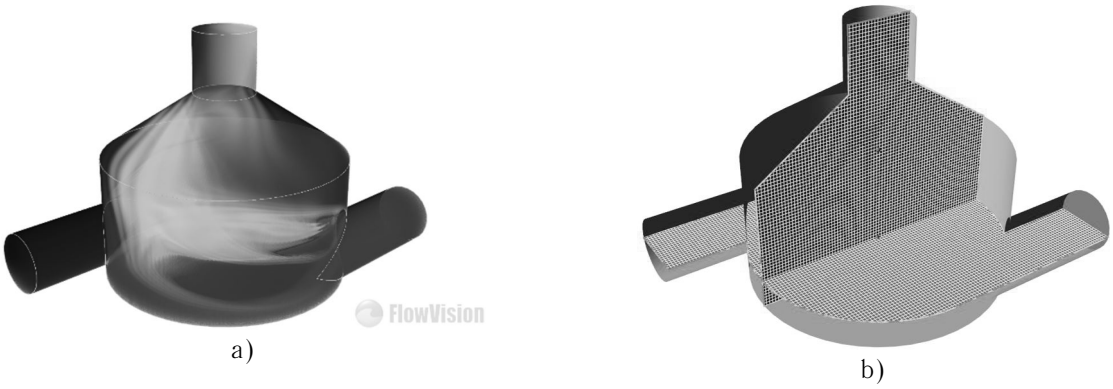



**Figure 4.** Mixing of hot and cold water flows in a mixer: a — volumetric visualization of temperature distribution; b — computational grid

is unit tensor. The equation for turbulent energy (in the 2nd and the 3rd test cases):

$$\frac{\partial(\rho k)}{\partial t} + \nabla(\rho \mathbf{V} k) = \nabla\left(\left(\mu + \frac{\mu_t}{\sigma_k}\right)\nabla k\right) + P_k + G_k - \rho\varepsilon\left(1 + \xi\left(\max\left(M_t^2, M_{t0}^2\right) - M_{t0}^2\right)\right), \quad (1)$$

$$P_k = \mu_t\left(S - \frac{2}{3}(\nabla\cdot\mathbf{V})^2\right) - \frac{2}{3}\rho(\nabla\cdot\mathbf{V})\,k, \quad G_k = \mu_t\frac{\beta}{\mathrm{Pr}_t}\mathbf{g}\cdot\nabla T,$$

$$S = 2\sum_{i,j} S_{ij}S_{ij}, S_{ij} = \frac{1}{2}\left(\frac{\partial V_i}{\partial x_j} + \frac{\partial V_j}{\partial x_i}\right).$$

Here $x_i$ is a Cartesian coordinate, $V_i$ is the $i$ th component of velocity, $k$ is the turbulent energy, $\varepsilon$ is its dissipation rate, $\beta$ is the coefficient of fluid thermal expansion, $a$ is the speed of sound. The equation for the dissipation rate of turbulent energy (in the 2nd and the 3rd test cases):

$$\frac{\partial(\rho\varepsilon)}{\partial t}+\nabla(\rho\mathbf{V}\varepsilon)=\nabla\left(\left(\mu+\frac{\mu_t}{\sigma_\varepsilon}\right)\nabla\varepsilon\right)+\frac{\varepsilon}{k}\left(C_{\varepsilon1}\left(P_k+G_k\right)-C_{\varepsilon2}\rho\varepsilon\right). \tag{2}$$

Equations (1) and (2) define the FlowVision version of the standard turbulence model [12]. In this model, $C_\mu=0.09$ is a fixed model constant. The default values of the other model constants are as follows:

$$\sigma_k=1,\quad \sigma_\varepsilon=1.3,\quad C_{\varepsilon1}=1.44,\quad C_{\varepsilon2}=1.92,\quad \xi=1.5,\quad M_{t0}=0.25. \tag{3}$$

Constants $\xi$ and $M_{t0}$ define the compressibility sub-model. Values $\xi=1.5,\quad M_{t0}=0.25$ define the Wilcox compressibility model [13]. Values $\xi=1$, $M_{t0}=0$ define the Sarkar compressibility model [13]. Note that a user can specify arbitrary values of constants (3) in the FlowVision interface. The energy equation written through total enthalpy (in the 1st and 2nd test cases):

$$\frac{\partial(\rho H)}{\partial t}+\nabla(\rho\mathbf{V}H)=\frac{\partial p}{\partial t}+\rho\mathbf{V}\cdot\mathbf{g}-\nabla\cdot\mathbf{J}_q+\nabla\cdot(\hat{\tau}\cdot\mathbf{V}), \tag{4}$$

$$H=h+\mathbf{V}^2/2,\quad h(T)=h_0(298.15)+\int\limits_{298.15}^{T}C_p(T)\,dT.$$

Here $H$ is total enthalpy, $h$ is thermodynamic enthalpy, $h_0(298.15)$ is the enthalpy of formation of the working fluid at 298.15 K, $C_p(T)$ is the specific heat of the working fluid at constant pressure, $T$ is absolute static temperature. The energy equation written through thermodynamic enthalpy (in the 3rd test case):

$$\frac{\partial(\rho h)}{\partial t}+\nabla(\rho\mathbf{V}h)=\frac{\partial P}{\partial t}+\mathbf{V}\cdot\nabla P-\nabla\cdot\mathbf{J}_q+\sum_{i,j=1}^{3}\mu\left(2S_{ij}-\frac{2}{3}(\nabla\cdot V)\,\delta_{ij}\right)S_{ij}+\rho\varepsilon. \tag{5}$$

In equations (4) and (5) $\mathbf{J}_q$ is specific heat flux. It is defined by the following expression:

$$\mathbf{J}_q=-\left(\lambda+\frac{\mu_t C_p}{\mathrm{Pr}_t}\right)\nabla T.$$

Here $\lambda$ is the coefficient of molecular thermal conductivity, $\mathrm{Pr}_t$ is the turbulent Prandtl number (in the 1st test case it is infinity). In the 1st and 2nd test cases, the system of the governing equations is completed by the ideal gas state equation. Computations are performed on relatively coarse grids. Therefore, in the 2nd and the 3rd test cases FlowVision model of wall functions [14] is used in the cells adjacent to the solid boundaries. The summary of equations is shown in Table 3.

**Table 3.** Summary of equations for the test cases

| Test case | Equations solved |
|---|---|
| M219 Cavity case | (3.3.2), (**??**), (4), $\mu_t=0$, $\mathrm{Pr}_t=\infty$ |
| Airplane flow | (3.3.2) – (4) |
| Mixer | (3.3.2) – (3), (5) |

### *3.3. Run options*

In order to obtain the required results, several runs of the test problems were performed with different start options. For each run the time of computation of a control time step was registered. The start

options for a given problem is characterized by number of used computational nodes, number of MPI-processes assigned to each node and number of the threads per each MPI-process, see Table 4.

**Table 4.** Start options for parallel tasks

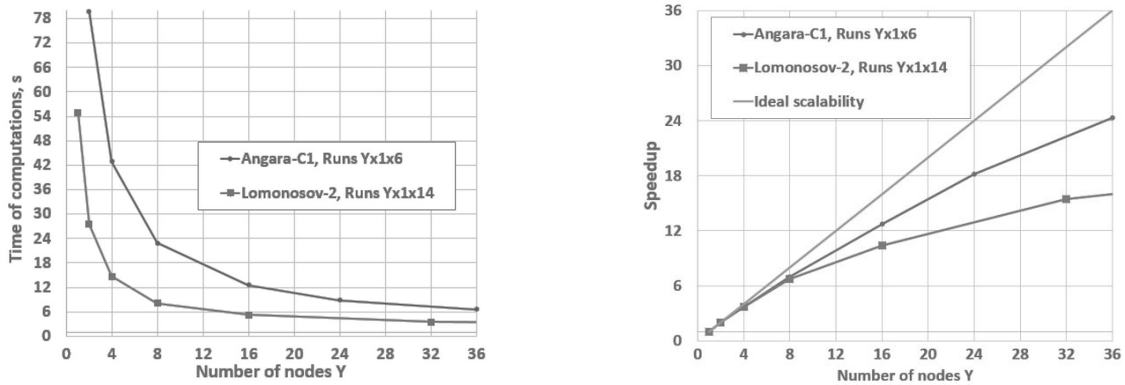
| | Number of nodes (nodes) | Number of MPI-processes per node (MPIs) | Number of threads per MPI Process (threads) |
|---|---|---|---|
| Parameter of the SLURM batch system or FlowVision | -N (parameter SLURM) | –ntasks-per-node (parameter SLURM) | Threads (parameter FlowVision) |
| Example | 24 | 2 | 6 |
| Notation (Nodes x MPIs x threads) | 24x2x6 | | |

## 4. EXPERIMENTAL RESULTS

### *4.1. Scalability of computations, single MPI-process per node*

At the first stage of the evaluation the M219 Cavity case problem was run with a single MPI-process per node and number of threads equals to the number of physical cores in a processor. The runs were performed on two supercomputers: Angara-C1 and Lomonosov-2. In all the regimes, the problem was calculated from time step 150 till time step 155. The computation time was registered for time step 155. It should be noted that up to the 150th time step the flow development was over and all the auxiliary operations (computational grid building, its adaptation, monitoring statistics, etc.) were terminated.

The dependency of computation time on number of nodes is shown in Figure 5a. Figure 5b demonstrates the scalability of computations relative to the run on a single node.



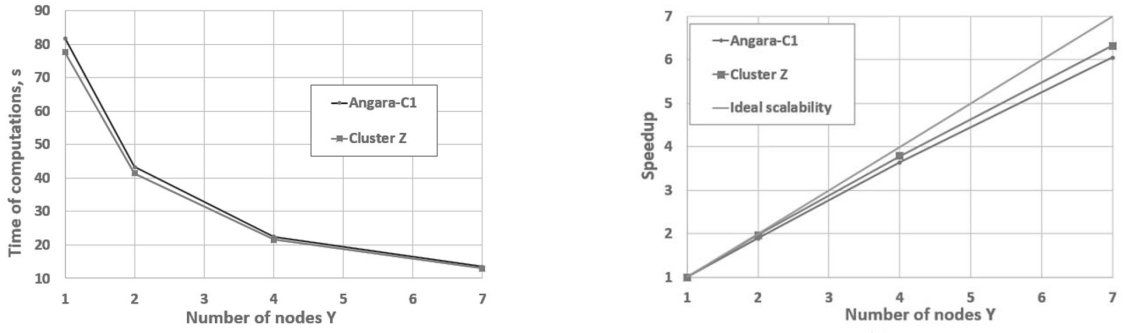

a) b)

**Figure 5.** Scalability of computations with one MPI-process per node: a – time of computations for control time step; b – speedup

One can see that the time of computations on the Lomonosov-2 supercomputer is much less than that on Angara-C1 due to the later processors – see Table 1. This intensifies the growth of the relative expenses on the MPI-exchanges with the number of the used nodes. Besides that, there are 2.3 times more cores in processors on Lomonosov-2, whereas the maximum memory bandwidth is only 1.6 times wider. These facts explain why the scalability of calculations on Lomonosov-2 looks worse compared to those on Angara-C1.

### *4.2. Scalability of computations, two MPI-process per node*

After that, runs of the same problem were performed on Angara-C1 and Z clusters with two MPI-processes per node, that is, with one MPI-process per each physical processor. In order to make the comparison correct, 6 threads per MPI-process were used on both clusters, i. e., the runs were performed in regime Yx2x6. The results of the comparison of the computation time and speedup are shown in Figure 6.




a) b)

**Figure 6.** M219 Cavity case problem. Scalability of computations with two MPI-process per node: (startup options Yx2x6) a – time of computations for control time step; b – speedup

One can see that the clusters demonstrate practically identical performance and scalability of computations.

### *4.3. Effect of using Hyper-Threading (HT)*

Scalability results with variable number of threads for the same M219 Cavity case problem is shown in Figure 7. The problem was run on 24 nodes with two MPI-processes per node of the Angara-C1 cluster. One can see that use of all the logical HT cores increases performance only by 4.5%.

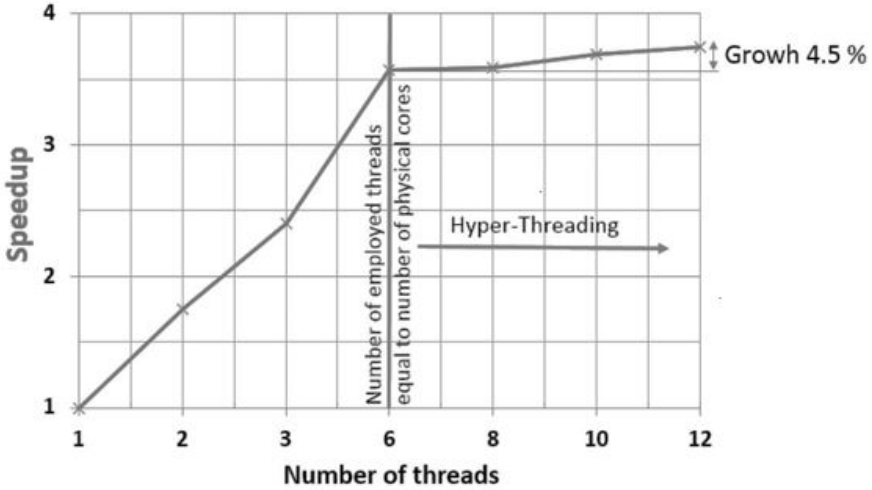



**Figure 7.** M219 Cavity case problem. Scalability of computations with different number of threads

In order to determine an optimal way of using logical cores HT, runs with different combinations of $MPIs$ and $threads$ have been performed (see Table 4). The total number of threads per node was kept constant ( $MPIs$ x $threads$ = 24). The number of nodes for these runs was also kept constant ( $Nodes$ = 24). Figure 8 demonstrates the performance speedup in these runs normalized to 24x2x6 run.

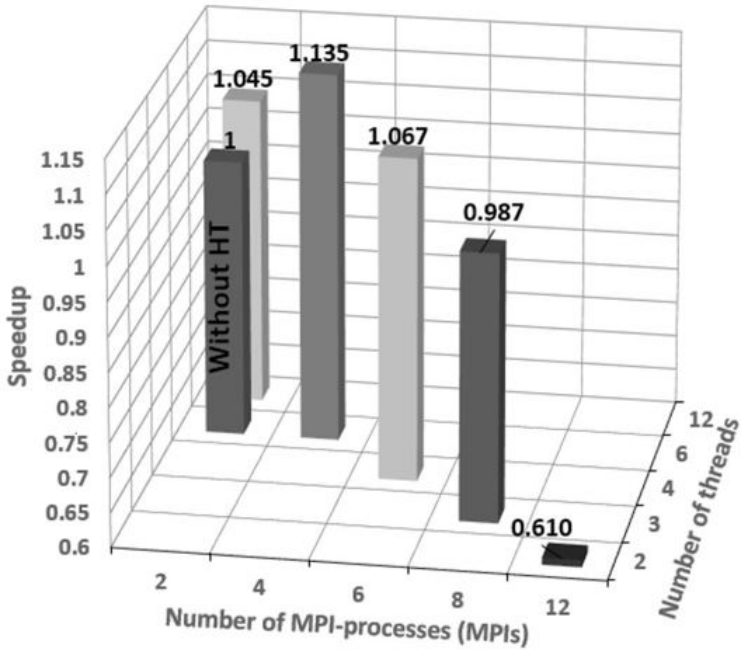
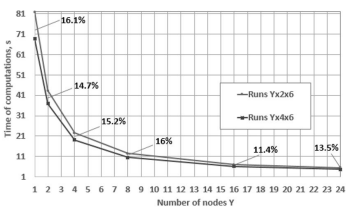


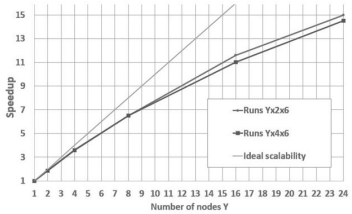

**Figure 8.** M219 Cavity case problem. Relative Performance speedup for different runs normalized to 24x2x6 run

One can see that the most effective is using all the HT logical cores due to doubling the number of the MPI-processes (24x4x6 run). The runs using more than 8 MPI-processes per node (i. e. 4 MPI-processes per physical node) have shown performance decrease.

It is interesting to estimate advantage of combination $MPIs$ x $threads$ = 4x6 in using different number of nodes. Figure 9 allows to compare the time of calculations of time step and the speedup of computations in the runs with 2 and 4 MPI-processes per node. One can see that the time of calculations of time step is 11.4–16.1% lower in the case of using duplicated number of MPI-processes in the entire range of the number of nodes.




a)




b)

**Figure 9.** M219 Cavity case problem. Scalability of computations in runs with 2 and 4 MPI-process per node: (startup options Yx2x6) a – time of calculations of time step; b – speedup

### *4.4. Scalability of computations on a small grid. Limit of scalability*

Scalability of computations in solving a CFD problem has limits which become apparent at a certain number of computational cells per core. In order to determine these limits, the Mixer problem was solved on a relatively small number of cells – 260 000 (see Table 3). The time of computation of the 8th time step was registered in all the runs of this problem. Figure 10 demonstrates the scalability of computations on Angara-C1 and Lomonosov-2 clusters with increasing number of the used cores relative to the run on a core. Note that contrary to the runs submitted in Figure 5, 2 MPI-processes and 6 threads were used on each node in all the runs on both clusters ( $MPIs$ x $threads$ = 2x6).

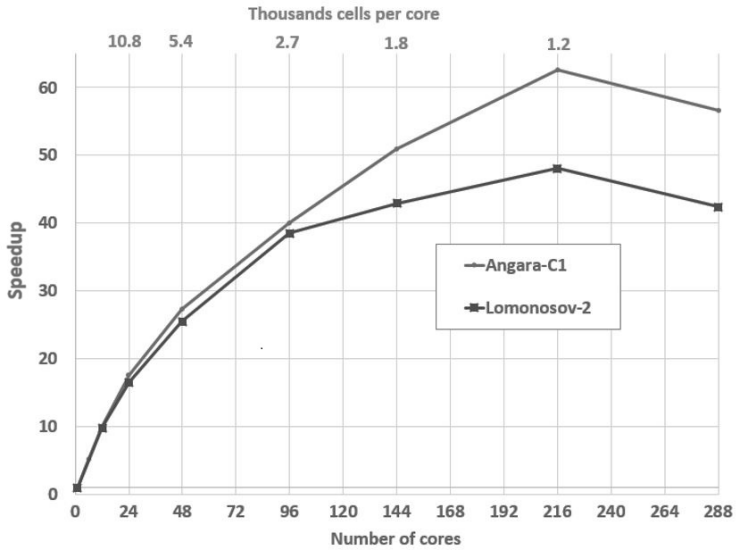



**Figure 10.** Mixer problem. Speedup of computations on a small grid

One can see that the both plots have maximum at 1200 cells per core. Angara-C1 demonstrates better scalability and better tolerance to small number of cells per core in the region where cash-memory effects may occur. It should be noted that usually at this number of cells per core, the time spent on the MPI-exchanges becomes comparable with the time of calculations of time step. Therefore, using a big number of cores in solving a small problem is unpractical from commercial point of view. The FlowVision users are advised to obtain a curve similar to the curves in Figure 10 when they are going to use less than 5000 cells per core in solving a CFD problem on a particular hardware.

### *4.5. Scalability of computations on an essentially non-uniform grid*

Above it was noticed that complexity of the geometrical model, local adaptation of the computational grid near solid surfaces and using the boundary-layer grid may cause substantial non-uniformity in the loading of the MPI-processes and, consequently, of the processors. In particular, the simulation of the airplane flow (see Table 2 and Figure 3) on the Angara-C1 cluster with use of 48 MPI-processes (24x2x6 run) produced the difference in the loading of processes about 300%. This difference indicates that most of time some processes wait of the other processes termination.

In order to estimate the Dynamic balancing technology, it was toggled on for 10 time steps (from 801 till 810). The data have been taken at the 811th time step, when the technology was toggled off and the computing time was not spent on the analysis of the non-uniformity of calculations. Figure 11 demonstrates the effect of the Dynamic balancing technology on the time of computation of one time step and on the parallel speedup. The speedup was measured relative to calculations on 4 nodes, since the RAM of less number of nodes is insufficient for solution of the problem discussed.

Figure 11 shows that using the dynamic balancing allows a user to reduce the time of computation of one time step and to improve the scalability when the number of the computational nodes exceeds 8.

As it was mentioned above, the process of dynamic balancing takes a certain processor time. Therefore, it is not recommended to keep this option toggled on during the entire process of calculations. Dynamic balancing option should be activated for 5–10 time steps when non-uniformity occurs in the process loading. If the grid is rebuilt in the course of calculations, it is advantageous to periodically toggle on and off this option, which is possible in FlowVision.

## 5. CONCLUSIONS

The scalability issues of the FlowVision software is evaluated in the paper. Three fluid-mechanic problems are solved on the Angara-C1 cluster equipped with the Angara interconnect (3D-torus topology). The sizes of the computational grids built for solving these problems are different. Different

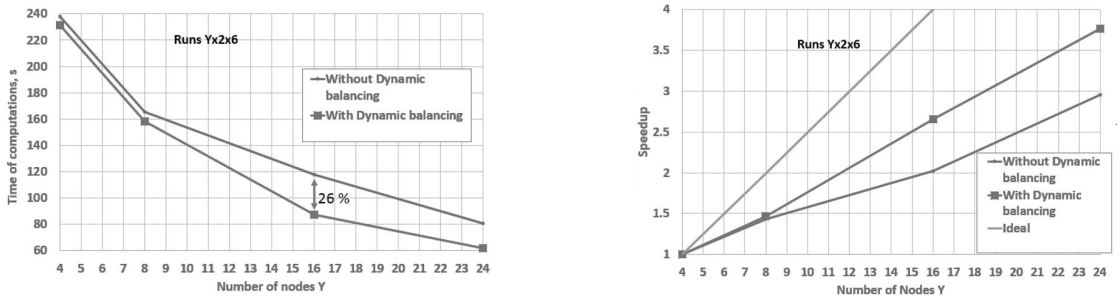



**Figure 11.** Airplane flow problem. Comparison of scalability of calculations on the Angara-C1 cluster with and without using Dynamic balancing technology: a — time of calculations of one time step; b — speedup

FlowVision capabilities are used in these simulations: turbulence models, grid adaptation in volumetric objects and in the vicinity of solid surfaces, boundary layer grid.

During the evaluation we obtain the following results:

1. The Angara-C1 cluster equipped with the Angara interconnect demonstrates good computational efficiency and scalability. The obtained results are highly competitive with the characteristics demonstrated by the modern clusters equipped with interconnect 4x FDR Infiniband.

2. Doubling the number of the MPI-processes per node by means of using logical nodes HT allows reducing the computational time by 11.4–16.1%.

3. If the problem is solved on the uniform grid containing 260 000 cells, the computation speedup is maximum when the calculations are performed on 216 cores, each processing 1200 computational cells. On the Angara-C1 cluster, these calculations proceed 60 times faster compared to those on one processor core.

4. In order to effectively use the hardware, it is recommended to load not less than 5000 computational cells onto each processor core.

5. Dynamic balancing technology implemented in the FlowVision software essentially augments the computational efficiency of a cluster, when the statement of a fluid mechanic problem produces a non-uniform loading of the processors.

## REFERENCES

1. V. V. Stegailov, N. D. Orekhov, and G. S. Smirnov, "HPC hardware efficiency for quantum and classical molecular dynamics", Int. Conf. on Parallel Computing Technologies. Springer, Cham (2015), 469–473.
2. G. S. Smirnov, and V. V. Stegailov, "Efficiency of classical molecular dynamics algorithms on supercomputers", Mathematical Models and Computer Simulations **8** (6), 734–743 (2016).
3. S. Derradji, "The BXI interconnect architecture", 23rd Annual Symposium on High-Performance Interconnects (HOTI), Santa Clara, CA, USA: IEEE (2015).
4. H. Frцning, "The On Achieving High Message Rates, 13th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing. Delft, The Netherlands: IEEE (2013).
5. A. S. Simonov, D. V. Makagon, I. A. Zhabin, A. N. Shcherbak, E. L. Syromyatnikov and D. A. Polyakov, "Pervoye pokoleniye vysokoskorostnoy kommunikatsionnoy seti «Angara»", Naukoyemkiye tekhnologii **15** (1), 21–28 (2014) [in Russian].
6. A. I. Slutskin, A. S. Simonov, I. A. Zhabin, D. V. Makagon and E. L. Syromyatnikov, "Razrabotka mezhuzlovoy kommunikatsionnoy seti «Angara» dlya perspektivnykh superkomp'yuterov", Uspekhi sovremennoy radioelektroniki **1**, 6–10 (2012) [in Russian].
7. V. Stegailov et al., "Early Performance Evaluation of the Hybrid Cluster with Torus Interconnect Aimed at Molecular-Dynamics Simulations. In: Wyrzykowski R., Dongarra J., Deelman E., Karczewski K. (eds) Parallel Processing and Applied Mathematics. PPAM 2017. Lecture Notes in Computer Science **10777**, 327–336. Springer, Cham (2018).

8. A. A. Agarkov, T. F. Ismagilov, D. V. Makagon, A. S. Semenov and A. S. Simonov, "Rezul'taty ocenochnogo testirovanija otechestvennoj vysokoskorostnoj kommunikacionnoj seti Angara" Superkomp'juternye dni v Rossii: Trudy mezhdunarodnoj konferencii (Moskva, 26–27 sentjabrja 2016) (Moscow, Publishing of Moscow State University) 626–639 (2016) [in Russian].
9. FlowVision – User's Guide of version 3.09.05, Official site FlowVision CFD. `https://flowvision.ru/webhelp/fven_30905/`. Accessed 2017.
10. G. B. Sushko and S. A Kharchenko, "Jeksperimental'noe issledovanie na SKIF MGU "Chebyshev" kombinirovannoj MPI+threads realizacii algoritma reshenija sistem linejnyh uravnenij, voznikajushhih vo FlowVision pri modelirovanii zadach vychislitel'noj gidrodinamiki". Parallelnye vychislitelnye tekhnologii (PaVT'2009): Trudy mezhdunarodnoj nauchnoj konferentsii (N. Novgorod, 30 marta – 3 aprelya 2009) (Chelyabinsk, Publishing of the South Ural State University) 316–324 (2009) [in Russian].
11. M. J. de C. Henshaw, "M219 cavity case: Verification and validation data for computational unsteady aerodynamics", Tech. Rep. RTO-TR-26,AC/323(AVT)TP/19. QinetiQ, UK, 453–472 (2002).
12. Yu. V. Fisher and A. E. Schelyaev, "Verifikatsija raschetnyh kharacteristic sverhzvukovyh turbulentnyh struj", Kompjuternye Issledovanija i Modelirovanije **9** (1), 21–35 (2017) [in Russian].
13. D. C. Wilcox, Turbulence modeling for CFD DCW Industries, Inc., 460 (1994).
14. S. V. Zhluktov and A. A. Aksenov, "Pristenochnye funktcii dlja vysoko-reynoldsovyh raschotov v programmnom komplekse FlowVision", Kompjuternye Issledovanija i Modelirovanije **7** (6), 1221–1239 (2015) [in Russian].