Computing Cost and Accounting Challenges for Octoshell Management System⋆

Yulia Belkina[1][0000−0003−1227−7827] and Dmitry Nikitenko[2][0000−0002−2864−7995]

[1] Lomonosov Moscow State University, Moscow, Russia
[2] Research Computing Center of Lomonosov Moscow State University, Moscow, Russia
yulia.belkina11@gmail.com,dan@parallel.ru

Abstract. Nowadays, the growing popularity of commercial use of high performance computing resources is explained by an increasing diversity of operations being eased through complex calculations. No application area could be even imagined without the advantages of digital modeling and other benefits of various scales of computing. All the activities can be done in cheaper ways obtaining better and faster outcomes together with competitive advantage gain. One of the questions that can arise for every supercomputer holder is the cost of spent computing resources for the job or for the whole project. This paper discusses general questions of computing costs and the prototype of a specialized accounting module for the Octoshell system which is based on proposed methodology of core-hour net cost calculation with subsequent price formation by implementing comparative and cost approaches which may be applicable to computing resources from a lab scale up to a petascale HPC center.

Keywords: Application run cost · High performance computing · Supercomputers · HPC center and operation cost of supercomputer

1   Introduction

In the modern world, every production, every small, medium-sized, and large businesses carry out calculations using computer technologies including high performance computing – HPC. Any type of calculations, especially complex calculations, require time and money, and similarly to manufacturing the product, calculations are not costless. Having an initiative to use supercomputers large productions have a choice to purchase the whole HPC system of a reasonable scale, starting from a powerful server, and up to a top-level supercomputer, or to rent computing resources. Reasons to use HPC technology in production vary: for many manufacturers supercomputers allow to approach their goals faster and at the lower cost, however, not everyone knows the actual cost of using such

computing resources. This study develops a methodology of a CPU core-hour net cost calculation which can be applied to both small clusters and large HPC systems. This methodology includes cost approach in order to account for an obvious initiative to compensate the initial and present costs of supercomputer, and comparison approach for an accurate price formation.

Diversification of the application areas in complex calculations imply growing demand for supercomputer resources. Analysis of the Fortune500 list [1] of the most profitable companies in the US demonstrates the wide range of economic activities, positive outcomes of which can be achieved quicker and in a more efficient way by implementing supercomputing technologies.

The first five Fortune500 corporations which use high-performance computing in order to improve their competence are: Walmart (retail) , ExxonMobil (oil and gas), Berkshire Hathaway (holding), Apple (technology), UnitedHealth Group (health care). All of the top five companies in Fortune500 list use HPC technology in different areas of the economy.
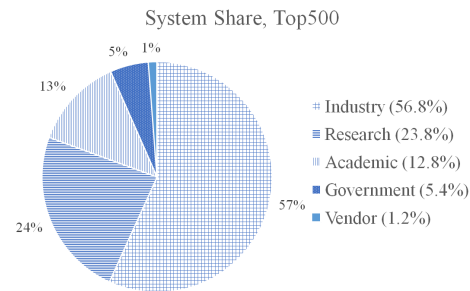

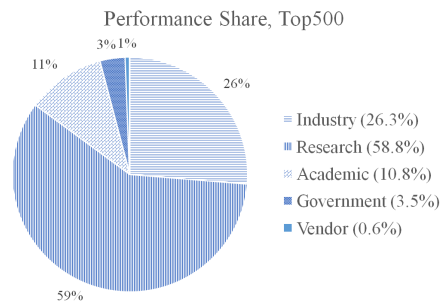**Fig. 1.** Segments System Share in the Top500 #51 list, June, 2018 [3]

Fig. 2. Segments Performance Share in the Top500 #51 list, June, 2018 [3]

Analogically most of the Fortune500 companies use high performance computing in order to increase their competitiveness and achieve higher production standards which confirms that supercomputer use gives competitive advantage.

Additionally, by looking into the Top500 rating of the most productive supercomputers of the world [2] it can be concluded that these HPC systems are being used in the full scope of economic activities. The pie-charts Fig. 1 and Fig. 2 demonstrate the variety of scientific spheres and economic activities calculated by world's Top500 supercomputers.

Industrial share in both of the pie-charts takes over a half and exceeds academic and research shares taken together. This demonstrates the wide prevalence of Top500 supercomputers exploitation in the industrial production which can be explained by the competitive advantage gained through high performance computing.

Industrial share in both of the pie-charts takes over a half and exceeds academic and research shares taken together. This demonstrates the wide prevalence of Top500 supercomputers exploitation in the industrial production which can be explained by the competitive advantage gained through high performance computing.
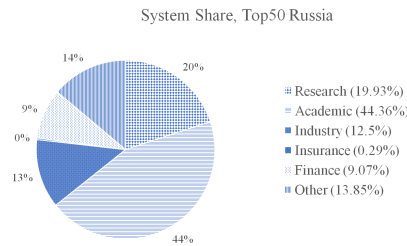

Fig. 3. Areas of application in the Top50 #28 list, April, 2018 [6]
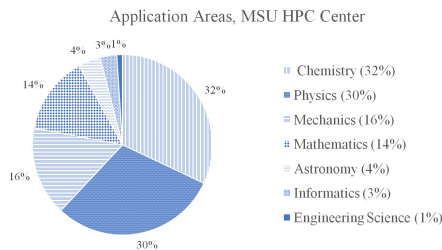

Fig. 4. Application areas of the MSU supercomputer center

The regional rating with its focus on the most powerful supercomputers in Russia – the Top50 [4, 5] is less rich with industrial popularity in application areas due to a big proportion of these HPC systems being installed in educational institutions and research centers, thus, the main focus of complex calculations on these HPC systems is fundamental science (Fig. 3).

The first place in Top50 is taken by the most powerful supercomputer system in Russia "Lomonosov-2", which is also currently rated 63rd in the world Top500 list. Its peak performance is estimated to be 4.946 PFlops while according to the LINPACK test "Lomonosov's-2" performance equals 2.478 PFlops. The areas of exploitation of computing resources of this supercomputer significantly varies, the pie chart below (Fig. 4) demonstrates the proportions of scientific spheres which "Lomonosov-2" is being used.

The reason of this supercomputer's availability for rent to commercial organizations is that this system was fully financed by the Lomonosov Moscow State University, where "Lomonosov-2" is installed.

Estimation of a job cost for a transparent demonstration of the importance of securing high efficiency of computing resources utilization is the main motive of the developed core-hour cost calculation method, which is presented in this paper.

2   The Goal of Research

In order to develop the methodology of estimation of computing resources net cost it is crucial to accurately identify what actually is the main computing resource and define the costs of exploitation of the system as a whole. The key computing resource is a core of central processor unit which is used to completing a job. Calculation of the net cost of this resource use for one hour – core-hour – is the main focus of the study, the cost of exploitation of a CPU core must include expenses on things like energy and amortization, salaries of stuff and rental cost of the building where the system operates, etc.

3   The Main Components of the HPC System

First, to have a proper understanding of the subject of this study the main resources of supercomputer are outlined:

- Computing resources (compute nodes which contain CPUs, having several CPU cores, and accelerators, memory, etc);
- Data storage (disk arrays, tape or network storage, etc.);
- Interconnect (communication network, transport network, service network);
- Infrastructure (power supply, cooling, fire extinguishing, and other engineering systems).

The exploitation of any of the resource is not possible without the use of at least one CPU core, thus, the basic computing resource is a core of central processing

unit, however, in practice it is not always possible to lease one particular core of a node as in framework of one node different jobs can interfere between each other causing competition for the resources and therefore there is a negative impact on the efficiency of calculations. The calculation of a net cost of CPU core-hour is considered essential for following computation of other resources costs as the latter is done proportionally to the cost of core-hour in case when it is required to estimate the cost of large number of different jobs, running in the system.

4   System Expenses

Both small and large computer systems have similar costs and expenses structure. In order to define the overall cost of using HPC system it is necessary to identify cost item list:

– Energy costs;
– Rental cost;
– Amortization deductions;
– Salaries (wage) of staff members;
– Insurance contributions;
– License costs.

Most of the expenses of supercomputer system use is required on energy costs, usually it varies depending on air temperature and number of resources in exploitation. Rental cost causes the expenses to increase significantly due to the big area being occupied by supercomputer's infrastructure. Amortization includes depreciation deductions and contribute a weighty part to overall expenses because of high costs of supercomputer equipment and its relatively short life time. Usually the life time of a supercomputer system is considered to be equal to five years, however, in case of "Lomonosov-2" the life time already exceeds five years and it will serve for longer time because the system keeps to be upgraded. Expenses on salaries of staff members include wages of working personnel who support the work of HPC system and technical support from the side of vendor, for instance, maintenance and repairing of the system components which are not insured. Insurance contributions are covering expenses on insurance of some components of the system or the system as a whole. Licenses are purchased for program development and completion.

5   Methodology of the Main System Costs Calculation

In order to develop the methodology of calculating net cost of using a core of central processing unit in one hour, the overall yearly expenses on the use of HPC system are required. For simplicity, the parameters are introduced: E – energy costs per year;
R – amortization per year;
S – rental cost per year;

s – salaries of staff per year;

I – insurance per year;

L – licenses per year. The formula for net cost of the yearly system exploitation is:

$$Z = E + R + S + s + I + L \qquad (1)$$

where Z – overall expenses of an operating HPC system per year.

Because such parameters as expenditure on rental cost of buildings occupied by system infrastructure and salaries of staff members depend on the rate of rent and indexation respectively, and the parameter of energy costs varies depending on electricity tariffs, those factors need to be accounted for in order to calculate the net cost of the system in the future period. Indexes t and t-1 are introduced to show the time period: t for the present year and t-1 for the past period. Therefore, the formula for a concrete period of time t is written as:

$$Z = E_t + R_t + S + s_t + L + I_t \qquad (2)$$

5.1   Energy Costs

Expenses on the energy consumption of supercomputer system exploitation in one year will depend on possible increase in system load which can be caused by more upgrades, moreover, costs will vary according to electricity tariffs changing. When the energy consumption is expected to remain unchanged compared to the previous period because the system is put on balance, the formula for yearly cost of energy in the future period is:

$$E_t = k * E_{t-1} \qquad (3)$$

where k – expected change in electricity tariffs.

5.2   Amortization

Amortization is calculated as a sum of amortization variables such as furniture amortization (index F), amortization of special supercomputer equipment (index E), and amortization of the building occupied by system infrastructure (index B):

$$R_t = R^F + R^E + R^B \qquad (4)$$

Such a division between those three types of amortization is caused by life time variation between furniture, supercomputer equipment, and buildings. Each of these amortization variables are calculated as:

$$R^i = P_0^i * AN^i / 100\% \qquad (5)$$

where i – index standing for either of the components (F, E, or B, and AN – amortization norm, computed as:

$$AN^i = 100\% / T^i \qquad (6)$$

where T^i – life time of infrastructure component i.

5.3 Salaries of Staff Members

In case when the system load is expected to remain constant in the following period, the expenses on salaries of working personnel are likely to depend on general wage change in the country, therefore, the formula for calculation yearly cost of working stuff in the next year is:

$$s_t = n * s_{t-1} \tag{7}$$

where n – the expected overall wage increase in the state.

5.4 Rental Cost of Occupied by HPC System Building

In this paper the building rental cost estimation is computed via sales comparison approach associating the object of estimation, in case of "Lomonosov-2", this is the sum of all the areas occupied by the HPC system, to objects-analogs which are buildings sharing similar infrastructure and location as the object of estimation. Different adjustments to the comparables are:

- Location adjustment.
- Bargain adjustment.
- Transport accessibility adjustment.
- Location of the premises in the building adjustment.
- Space adjustment.
- Separate entry adjustment.

Accounting for all the adjustments the method gives an approximate estimate of the rental cost of one square meter per year, multiplication of this number gives the overall yearly rental cost.

5.5 Cost of Insurance

Because of deterioration and depreciation of the system, cost of insured details of it will steadily decrease over years. In the paper, the proposed solution to account for insurance is to calculate it as a direct ratio to residual value.

5.6 Cost of Licenses and Software

License is an essential part required for running a job on supercomputer. Because it was purchased once and for all, the cost of it will remain unchanged.

6   Method of Core-Hour Net Cost Calculation

As the CPU core is an elementary computing resource of the supercomputer, it is required to estimate expenses on one hour exploitation of one core of central processing unit. This is done via cost approach where yearly expenses on the

overall system exploitation (Z) are divided by the product of the total numbers of cores in supercomputer and of the total number of hours in one year:

$$C = Z/(N * T')\tag{8}$$

where C – CPU core-hour net cost; N – total number of CPU cores; T' – total number of hours in one year. The presented formula is a method of net cost of core-hour calculation for the next year, thus, in theory, all CPU cores can be in exploitation for the whole year, therefore Z here is calculated using formula (2). This would be a lower bound estimation where total number of cores and all hours of the year are accounted for. In reality, number of cores used is expected to be smaller whereas time would decrease due to idle time caused by system maintenance or other factors.

7   Method of the Cost of Application Run Calculation

Every project requires a number and variety of job runs, and different time and various number of computing resources for its jobs completion. There are different partitions for the job runs and there is usually a limited TEST partition for test runs available. After the program (job) has successfully completed the test runs and thus considered to be fully developed, its execution can be conducted on a regular partition where the number of resources and the available time increase. The method of calculation of the cost of the project is simply a sum of the costs of each job of this project:

$$W = \sum_{j}(N_j{`} * t_j * C)\tag{9}$$

where W – total cost of running a project; N_j – number of CPU cores used for computing job j; t_j – number of hours required for computing job j.

Therefore, the calculation of the cost of all projects in the system is done through summing up the costs of all the projects:

$$V = \sum_{p}(W_p)\tag{10}$$

where index p stands for a particular project. This formula can expand when needed to calculate the cost of the whole supercomputer center containing more than one HPC systems. The net cost of a CPU core-hour will differ between the supercomputers, thus parameter defining total cost of running a project will become:

$$W^S = \sum_{S} N_j^{S`} * t_j^S * C^S)\tag{11}$$

where W^S – total cost of running a project in system S; $N_j^{S`}$ – number of CPU cores used for computing job j in system S; t_j^S – number of hours required for computing job j in system S; C^S – net cost of CPU core-hour in system S.

The cost of all the projects in all of the HPC systems of the supercomputer complex is computed as:

$$X = \sum_S V^S \tag{12}$$

where X – the cost of the total number of the projects run in the entire supercomputer complex; V^S – the cost of the total number of the projects run in particular system S, calculated as:

$$V^S = \sum_{p,S} W_p^S \tag{13}$$

8   Evaluation of the Presented Methodology

Formation of the price of the use of computing resources of supercomputer "Lomonosov-2" in order to lease those resources to researchers and commercial organizations for partial compensation of system expenses is done via two approaches: cost approach and comparative approach. Cost approach allows to estimate the cost of computing resources use by defining (accounting for) the system expenses such as development of the system, its cost and installation. Comparative approach allows to form the price via comparisons made between the object of estimation and its analogs. The actual price is calculated via both of the approaches.

Because "Lomonosov-2", being upgraded regularly, has already been running for a few years with a full queue of jobs, the system load is expected to remain unchanged, meaning both energy consumption and the cost working personnel are not expected to differ in the close future.

The separate data on the initial costs of either furniture, supercomputer equipment, or buildings is not provided, however the initial cost of the whole system is, with an estimated life time of seven years. Therefore, using formulas (4) and (5) the amortization deductions are calculated for the whole system in just two steps.

Rental cost of building occupied by HPC system was calculated with the use of all presented above adjustments.

Data on the cost of system or its details insurance is not provided, information about the licenses costs is not given as well, however, even without those parameters the approximate cost of system expenses was equal to over 525 million RUB per year of exploitation.

Using formula (8), the net cost of one of 23,424 CPU cores of "Lomonosov-2" is approximately 2.56 RUB per hour, this estimate is the result of the cost approach. Therefore, if looking only at the cost approach results, then in order to break even the one hour use of one CPU core, this price should be paid. Required by an entrepreneur percentage might be added to the overall expenses (Z) in order to get an entrepreneurial profit.

Comparative approach compares the CPU core-hour cost of the object of estimation, in this case it is "Lomonosov's-2" core-hour, to the price of core-hour estimated by the two other supercomputer centers in the Russian Federation: Tomsk State University and Joint Supercomputer Center of the Russian Academy of Science. With equal to 0.5 weights for both approaches, where entrepreneurial profit is not included in the cost approach, the cost of core-hour is calculated to be around 2.44 RUB.

9   Enriching Octoshell Management System with the Developed Methods of Computing Cost Calculation

The Octoshell HPC center management system provides a wide range of tools both for system holders and administrators, and for regular users. Among others, there is a tool that allows controlling the amount of utilized resources by a specified account or a project [7, 8].

At the same time every HPC system in Octoshell has some properties, and the cost of CPUh, core-hour, or a node-hour could be added to these properties, based on the methods presented above.

This would easily allow creating another format of presenting the executed job summary, making a step towards the accounting module. This format could extend the existing list of jobs with states and total amount of used resources with value of these spent resources.

The prototype has been developed and an example is given below.

10   Example of Application Cost Calculations

In order to demonstrate the importance of the developed methodology, an example of one project run in HPC system "Lomonosov-2" is introduced. The table below (Table 1) shows different partitions of the project such as "test", "compute", and "low_io", the states of the job runs is listed in the second column for every partition. The only useful for sure are the completed jobs, those under the rest states are probably inefficient. The third column represents the number of jobs for each state and each partition, while the fourth represents the amount of CPU core-hours used for the total sum of those jobs.

First of all, it is obvious from the table that the number of unsuccessful job runs which were either failed, canceled, unfinished due to the timeout or due to the node fail, proportionally is impressive. As such, in partition "low_io", the approximate percentage of uncompleted jobs is 33.83%, in "compute" section this number rises to 49.58% - almost a half of the jobs were unsuccessful, while in partition "test" the percentage of uncompleted jobs was even higher: 61.42%. This extremely high rate of unsuccessful job runs demonstrates the need for careful consideration of debugging the program at the early stages, as the system load is less efficient when has a high number of uncompleted jobs.

Table 1. Example of a project resource utilization by the number of job runs and distribution by states and partitions

Partition	State	Count	Cores*Hours
low_io	Total	61,327	2,081,460.06
	Completed	40,581	968,530.01
	Failed	19,985	62,352.50
	Cancelled	551	711,285.72
	Timeout	194	297,321.65
	Running	10	39,900.00
	Node_fail	6	2,097.19
compute	Total	361	759,124.73
	Completed	182	405,695.01
	Cancelled	111	195,049.57
	Failed	51	18,890.46
	Timeout	15	117,376.22
	Running	1	998.36
	Node_fail	1	21,115.11
test	Total	438	970.35
	Completed	169	218.86
	Failed	150	74.32
	Cancelled	88	186.02
	Timeout	31	491.14

Table 1 also shows the amount of CPU core-hours used for all those jobs, these numbers, however, does not tell much to a common user, who considers supercomputer being capable of carrying out any computing for free.

The calculations are not free in any case, and it doesn't depend on if it is a commercial project, or a basic research. Somebody always pays for that.

Implementation of the developed in the presented methodology formulas results into a different table (Table 2) showing costs of the jobs from the first table.

The right column of Table 2, defining cost of the job, is calculated similarly to formula (8), where both number of the cores (N') and time (t) are given by the column "Core*Hours", and the cost of the core-hour is taken as was estimated in the approbation part of the paper: 2.44 RUB.

The first conclusion to be made here is that computing can be highly expensive, total cost of all the partitions equals to 6,933,394.54 RUB, of it 3,353,643.07 RUB is the cost of completed jobs, leaving the uncompleted jobs with even higher cost of 3,579,751.47 RUB. This numbers would increase with a presence of data on insurance expenses and the costs of licenses, it is also going to become higher if the entrepreneurial profit is included in the price formation.

The second conclusion resulting from analysis of the table is that in order to decrease the costs of HPC system exploitation it is essential to accurately debug the program before running.

Table 2. Project resource utilization cost

Partition	State	Cores*Hours	Cost
low_io	Total	2,081,460.06	5,078,762.55 RUB
	Completed	968,530.01	2,363,213.22 RUB
	Failed	62,352.50	152,140.10 RUB
	Cancelled	711,285.72	1,735,537.16 RUB
	Timeout	297,321.65	725,464.83RUB
	Running	39,900.00	97,356.00 RUB
	Node_fai	2,097.19	5,117.14 RUB
compute	Total	759,124.73	1,852,264.34 RUB
	Completed	405,695.01	989,895.82 RUB
	Cancelled	195,049.57	475,920.95 RUB
	Failed	18,890.46	46,092.72 RUB
	Timeout	117,376.22	286,397.98 RUB
	Running	998.36	2,436.00 RUB
	Node_fai	21,115.11	51,520.87 RUB
test	Total	970.35	2,367.65 RUB
	Completed	218.86	534.02 RUB
	Failed	74.32	181.34 RUB
	Cancelled	186.02	453.89 RUB
	Timeout	491.14	1,198.38 RUB

11   Conclusions

The amount of CPU core-hours utilized by applications sometimes doesn't tell much to a common user, who considers supercomputer being capable of carrying out any computing for free. The calculations are not free in any case, and it doesn't depend on if it is a commercial project, or a pure scientific basic research. Somebody pays for the computing.

The developed methodology, described in this paper is applicable to most types of high performance computing systems either small or large ones, it also shows the methods of calculation the cost of the jobs running under a certain project, or in the entire supercomputer center even containing more than one HPC system.

Approbation of the methodology and the presented example demonstrate the real price of inefficiency and need in renovation of the HPC systems use, such as debugging programs at the early stages in order to decrease the cost of computing resources which will result in a more efficient system output.

References

1. The Fortune500 List. [Online]. Available: http://fortune.com/fortune500/ [Accessed: Aug. 9, 2018].
2. The Top500 list of the world's most powerful supercomputers. Available: https://www.top500.org/. [Accessed: Aug. 9, 2018].

3. Top500 Statistics. [Online]. Available: https://www.top500.org/statistics/list/. [Accessed: Aug. 9, 2018].
4. The Top50 list of supercomputers in the Russian Federation. [Online]. Available: http://top50.supercomputers.ru/?page=rating. [Accessed: Aug. 9, 2018].
5. D. Nikitenko and A. Zheltkov: The Top50 list vivification in the evolution of HPC rankings. In: Parallel Computational Technologies, vol. 753 of Communications in Computer and Information Science (CCIS), pp. 14–26, Springer International Publishing AG, New York, 2017. https://doi.org/10.1007/978-3-319-67035-5/_2
6. Top50 Statistics. [Online]. Available: http://top50.supercomputers.ru/?page=stat. [Accessed: Aug. 9, 2018].
7. Nikitenko, D.A., Voevodin, Vl.V., Zhumatiy, S.A.: Resolving frontier problems of mastering large-scale supercomputer complexes. In: ACM International Conference on Computing Frontiers (CF'16), May 16-18, 2016, Como, Italy. ACM New York, NY, USA. 349–352 (2016).
8. Nikitenko, D.A., Voevodin, Vl.V., Zhumatiy, S.A.: Octoshell: Large Supercomputer Complex Administration System. In: Russian Supercomputing Days International Conference, Moscow, Russia, September 28-29, 2015. CEUR Workshop Proceedings, vol. 1482, 69–83 (2015).