Proc. of the Intl. Conf. on Advances In Computing, Electronics and Electrical Technology - CEET 2014.

ISBN: 978-1-63248-005-7 doi: 10.15224/ 978-1-63248-005-7-37

Prediction of performance of cluster computing system with different configurations of nodes

[Daulet Akhmedov, Suleimen Yelubayev, Farida Abdoldina, Timur Bopeyev, Daulet Muratov]

***Abstract* — This article considers results of performance test of experimental model of cluster hybrid computing system based on GPU accelerators in Linpack. It was created six variants of configuration of cluster system. These variants differ from each other by count of nodes, GPU accelerators in each node and amount of random access memory.**

It was carried out qualitative analysis of test results using approximation methods. As a result it was created simulation model for determination of dependence of real performance (Tflops) of cluster system on count GPU accelerators in the node. The simulation model allows to estimate optimal value of cost-performance ratio in dependence on characteristics of electronic components of cluster computing system.

***Keywords*—parallel computing, high-performance computing, cluster computing system, GPU accelerator, CUDA technology.**

I. Introduction

Experimental model of hybrid cluster computing system (HCCS) based on Nvidia Tesla GPU was developed as part of the budget project "Develop a hybrid cluster computing system based on GPU" of the Ministry of Education and Science of the Republic of Kazakhstan [1]. Peak performance of the experimental model of cluster computing hybrid system is 16 teraflops of single precision and 8 teraflops of double precision.

Following results were obtained in the process of development of experimental model of HCCS based on GPU:

- principles of construction the HCCS on the base of GPU were determined;
- preliminary specifications for creation the HCCS on the base of GPU were developed;
- preliminary design for HCCS was carried out. Engineering solutions for HCCS were developed;
- experimental model of HCCS on the base of GPU was manufactured;
- tests of experimental model of HCCS on the base of GPU was carried out. These tests confirmed the correctness of selection the engineering solutions and allowed to specify the certain technical specifications.

D. Akhmedov, S. Yelubayev, F. Abdoldina, T. Bopeyev, D. Muratov

The Institute of Space Technique and Technologies
Republic of Kazakhstan

Then it is planned to develop the experimental-industrial model of HCCS having the performance of 21 Tfloops of double precision and 41 TFlops of single precision.

II. Experimental model of HCCS on the base of GPU

Experimental model of HCCS was constructed in accordance with the cluster technology Beowulf. Scalability is a main feature of this cluster, i.e. the possibility of increasing the number of nodes of the system with a proportional increase of performance. The number of nodes of experimental model of HCCS is limited by the type of selected network switch Mellanox InfiniScale IV Switch, which allows to create a system containing 2-8 nodes.

Six different configurations of hybrid cluster computing system was developed for testing its performance:

- 1-st variant contains 2 nodes, each node contains 2 GPU;
- 2-nd variant contains 2 nodes, each node contains 3 GPU;
- 3-d variant contains 2 nodes, each node contains 4 GPU;
- 4-th variant contains 3 nodes, each node contains 2 GPU;
- 5-th variant contains 3 nodes, each node contains 3 GPU;
- 6-th variant contains 3 nodes, each node contains 4 GPU.

The configuration of each node of experimental model of HCCS containes the following components: two CPUs Intel Xeon E5-2620 2.0 GHz; motherboard MB Supermicro X9DRG-QF; GPU Nvidia Tesla M2090, amount of GPU varied from 2 to 4; RAM DDR3 of 32Gb in each node; network adapter Mellanox ConnectX-2 VPI.

Experimental model of HCCS operates under control of operating system Linux RHEL 6.1 with open source. Technology “Message Passing Interface” (OpenMPI ver. 1.6) was used for the distribution of data between the nodes. Specialized software that was used: parallel computing platform and models of programming for speed up the solution of scientific and engineering problems on the GPU - NVIDIA CUDA 5.0; library of mathematical application programs Intel ® Math Kernel Library (Intel ® MKL 10), manager of distributed resources for computing clusters - TORQUE. Cuda

Proc. of the Intl. Conf. on Advances In Computing, Electronics and Electrical Technology - CEET 2014.

ISBN: 978-1-63248-005-7 doi: 10.15224/ 978-1-63248-005-7-37

Accelerated Linpack 2.0 v16 was selected as a performance testing tool, as a tool used for ranging of the fastest supercomputers in the world and CIS countries [2,3].

III. Testing the performance of experimental model of HCCS. Analysis of results

In this section we perform the analysis of test results and determine the optimal configuration of HCCS in terms of different criteria, depending on the amount of RAM and the number of GPU.

Experimental model of HCCS consisting of 2 nodes [4, 5]

Peak (theoretical) performance is determined by the sum of theoretical performance of all graphic processors used in the system. Actual performance of cluster system is determined by the test in Linpack.

Peak performance of the first variant of the experimental model of HCCS is 2640 GFlops, peak performance of the second variant is 3960 GFlops, peak performance of the third variant is 5280 gigaflops of double precision. Tests allowed to determine the dependence of actual performance of the cluster system on the amount of RAM.

Table 1 shows the test results for the first three variants, including the value of the actual performance, system performance and specific cost of 1 GFLOPS/sec, in other words, the value of price/performance ratio.

TABLE I. RESULTS OF TESTING THE 1-3 VARIANTS OF HCCS

Amount of RAM, Gb	Dimension of problem	1-st variant of HCCS	2-nd variant of HCCS	3-d variant of HCCS
		Real performance, Gflops	Real performance, Gflops	Real performance, Gflops
4.2	14 273	540.2	479.0	299.9
6.1	20 822	757.7	739.2	505.5
7.8	25 502	850.2	874.7	644.8
8.5	29 447	963.0	1 051.0	755.9
10.2	32 923	1 009.0	1 129.0	865.4
13.7	36 066	1 077.0	1 196.0	970.0
15.9	38 956	1 107.0	1 272.0	1 064.0
17.6	41 645	1 144.0	1 332.0	1 095.0
19	44 171	1 181.0	1 362.0	1 181.0
21.6	46 561	1 207.0	1 418.0	1 212.0
23.4	48 833	1 221.0	1 409.0	1 257.0
25.7	52 224	1 233.0	1 468.0	1 366.0
29.3	56 853	1 286.0	1 537.0	1 454.0
33.54	60 068	1 304.0	1 578.0	1 543.0
35.1	63 284	1 317.0	1 605.0	1 588.0
40.1	66 499	1 349.0	1 635.0	1 655.0
42.7	69 715	1 367.0	1 673.0	1 676.0
46.3	72 930	1 380.0	1 707.0	1 779.0
50.1	76 146	1 370.0	1 708.0	1 830.0
54.7	79 361	1 367.0	1 739.0	1 865.0
58.8	82 577	1 350.0	1 750.0	1 923.0

Preliminary analysis shows that specific cost of 1 Gflops falls with the increasing of efficiency of work of cluster. The cost in the table is specified in the Kazakh national currency, tenge. Tenge rate to dollar is 155.27 KZT per dollar at the time of writing this paper.

The table shows the amount of RAM used by the system for calculations (processes “Linpack “and “System”). The volume of RAM depends on the dimension of the Linpack problem.

We use the approximation methods for a more qualitative analysis of obtained results. Let's consider separately the dependence of actual performance, efficiency and specific cost on the amount of RAM. We use the polynomial approximation of second-order in all three cases.

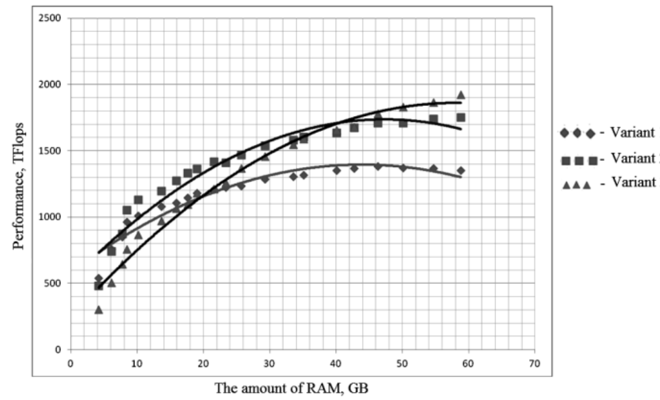

Figure 1. Dependence of performance of HCCS for 1-3 variants on the amount of RAM

Table data are showed with help of marker, results of approximation of table data are showed with help of smooth curves at the figure 1. As can be seen from Figure 1, the real data are approximated sufficiently accurate by obtained curves. Equation describing the dependence of real performance on the amount of RAM, has the following form:

$$y = ax^2 + bx + c \, , \qquad (1)$$

where: a=-0.4205; b=36.853; c=587.65.

To find the maximum of this function, we equate the first derivative and determine the coordinates of the point of maximum and maximum of real performance respectively. As a result of calculations, we obtain:

$$x_{max}=43.8, \quad y_{max}=1395.1$$

i.e. the maximum performance of 1395.1 GFlops in the first variant is achieved when the amount of RAM is 43.8 GB. Further increase of RAM will not increase the performance. This is due to the fact that the load on the data transmission channels increases with the increase of dimension of the problem, that ultimately slows down the exchange of data between RAM and the GPU.

The same data processing were conducted for 2nd and 3d variants of HCCS. Summary result is presented in Table 2.

Proc. of the Intl. Conf. on Advances In Computing, Electronics and Electrical Technology - CEET 2014.

ISBN: 978-1-63248-005-7 doi: 10.15224/ 978-1-63248-005-7-37

TABLE II. RESULTS OF DETERMINING THE MAXIMUM REAL PERFORMANCE IN DEPENDENCE OF THE AMOUNT OF RAM

HCCS variant	**Coefficients of polynomial approximant of 2nd order**			**Maximum of performance y_{max}, Tflops**	**Value of amount of RAM at the point of maximum x_{max}, Gbyte**
	a	***b***	***c***		
1-st variant	-0.4205	36.853	587.65	1395.1	43.8
2-nd variant	-0.542	51.23	525.7	1736.3	47.3
3-d variant	-0.4773	55.574	243.21	1860.9	58.2

The dependence of the efficiency on the amount of RAM is considered.

TABLE III. RESULTS OF DETERMINING THE MAXIMUM EFFICIENCY IN LINPACK TEST IN DEPENDENCE ON THE AMOUNT OF RAM

HCCS variant	**Coefficients of polynomial approximant of 2nd order**			**Efficiency maximum in accordance with Linpack y_{max}, %**	**Value of amount of RAM at the point of maximum x_{max}, Gbyte**
	a	***b***	***c***		
1-st variant	-0.00016	0.013958	0.22259	53.1	43.62
2-d variant	-0.00014	0.012922	0.132875	43.0	46.15
3-d variant	-0.0009	0.010519	0.046248	35.3	58.338

The dependence of the specific cost of 1 GFlops on the amount of RAM is considered.

TABLE IV. RESULTS OF DETERMINING THE MINIMUM SPECIFIC COST OF 1GFLOPS OF HCCS IN DEPENDENCE ON THE AMOUNT OF RAM

HCCS variant	**Coefficients of polynomial approximant of 2nd order**			**Minimum specific cost of 1 Glops of HCCS, y_{min}, tenge**	**Value of amount of RAM at the point of minimum x_{min}, Gbyte**
	a	***b***	***c***		
1-st variant	0.2452	-22.317	1537.1	1029.1	45.5
2-d variant	0.3537	-35.45	2101.5	1213.3	50.1
3-d variant	0.9041	-92.97	3879.8	1489.7	51.4

Experimental model of HCCS consisting of 3 nodes

Peak performance of the fourth variant of experimental model of HCCS is 3960 GFlops, peak performance of the fifth variant is 5940 GFlops, peak performance of the sixth variant is 7920 Gflops of double precision. Dimension of the problem varied from 14 273 to 101 870, required amount of RAM ranged from 4.2 GB to 85.1 GB. Table 5 shows the test results.

Data processing was made for 4-th, 5-th and 6-th variants of test, that correspond to the group of variants of tests with 2 nodes. Results of processing are presented in the tables 6,7,8.

TABLE V. TEST RESULTS FOR 4-6 VARIANTS OF HCCS

Amount of RAM, GB	**Dimension of problem**	**4-th variant of HCCS**	**5- th variant of HCCS**	**6- th variant of HCCS**
		Real performance, Gflops	**Real performance, Gflops**	**Real performance, Gflops**
4.2	14 273	431.8	229.2	433.6
6.1	20 822	823.7	444.9	735.5
7.8	25 502	952.0	622.2	871.5
8.5	29 447	1 076.0	768.5	1 069.0
10.2	32 923	1 166.0	888.9	1 205.0
13.7	36 066	1 263.0	974.9	1 282.0
15.9	38 956	1 305.0	1 158.0	1 335.0
17.6	41 645	1 383.0	1 214.0	1 471.0
19	44 171	1 412.0	1 316.0	1 525.0
21.6	46 561	1 437.0	1 386.0	1 593.0
23.4	48 833	1 492.0	1 482.0	1 625.0
25.7	52 224	1 529.0	1 553.0	1 733.0
29.3	56 853	1 599.0	1 706.0	1 825.0
33.54	60 068	1 625.0	1 811.0	1 903.0
35.1	63 284	1 668.0	1 930.0	1 933.0
40.1	66 499	1 704.0	1 934.0	2 031.0
42.7	69 715	1 746.0	1 997.0	2 110.0
46.3	72 930	1 732.0	2 017.0	2 196.0
50.1	76 146	1 765.0	2 115.0	2 283.0
54.7	79 361	1 784.0	2 158.0	2 365.0
58.8	82 577	1 808.0	2 176.0	2 385.0
62.4	85 793	1 807.0	2 278.0	2 429.0
66.8	89 008	1 816.0	2 271.0	2 437.0
71.3	92 224	1 799.0	2 295.0	2 519.0
75.3	95 439	1 860.0	2 390.0	2 569.0
80.9	98 655	**1 912.0**	2 396.0	2 608.0
85.1	101 870	1 897.0	**2 421.0**	**2 622.0**

TABLE VI. RESULTS OF DETERMINING THE MAXIMUM REAL PERFORMANCE IN DEPENDENCE OF THE AMOUNT OF RAM

HCCS variant	**Coefficients of polynomial approximant of 2nd order**			**Maximum of performance y_{max}, Tflops**	**Value of amount of RAM at the point of maximum x_{max}, Gbyte**
	a	***b***	***c***		
4-th variant	-0.289	37.034	705.84	1892.2	64.07
5-th variant	-0.4199	59.844	232.07	2364.8	71.25
6-th variant	-0.3377	52.009	552.14	2554.6	77.0

TABLE VII. RESULTS OF DETERMINING THE MAXIMUM EFFICIENCY IN DEPENDENCE ON THE AMOUNT OF RAM IN ACCORDANCE WITH THE LINPACK TEST

HCCS variant	**Coefficients of polynomial approximant of 2nd order**			**Maximum efficiency on Linpack y_{max}, %**	**Value of amount of RAM at the point of maximum x_{max}, Gbyte**
	a	***b***	***c***		
4-th variant	-0.00007	0.0094	0.1782	49.3	67.1
5- th variant	-0.00007	0.0101	0.0393	40.3	72.1
6- th variant	-0.00004	0.0066	0.0699	34.2	82.5

TABLE VIII. RESULTS OF DETERMINING THE MINIMUM SPECIFIC COST OF 1GFLOPS OF HCCS IN DEPENDENCE ON THE AMOUNT OF RAM

HCCS variant	**Coefficients of polynomial approximant of 2nd order**			**Minimum specific cost of 1 Glops of HCCS, y_{min}, tenge**	**Value of amount of RAM at the point of minimum x_{min}, Gbyte**
	a	***b***	***c***		
4-th variant	0.2527	-33.43	2742.9	1636.8	66.15
5- th variant	0.9166	-117.76	5603.8	1821.6	64.23
6- th variant	0.6728	-94.35	5655.2	2347.7	70.12

Proc. of the Intl. Conf. on Advances In Computing, Electronics and Electrical Technology - CEET 2014.

ISBN: 978-1-63248-005-7 doi: 10.15224/ 978-1-63248-005-7-37

IV. Mathematical model of prediction of HCCS performance based on the results of numerical experiments

Let's derive the functions a(G), b(G), c(G) for each of the coefficients of equation (1) a, b, c respectively, where G is the number of GPU used in the nodes of HCCS. We use the information from the table 2 for this purpose, i.e. we approximate the required dependencies with help of the same polynomial of 2nd order. Polynomial of 2nd order is selected due to the fact that the dependence of the coefficients of equation (1) has a pronounced parabolic character (see Figure 2).

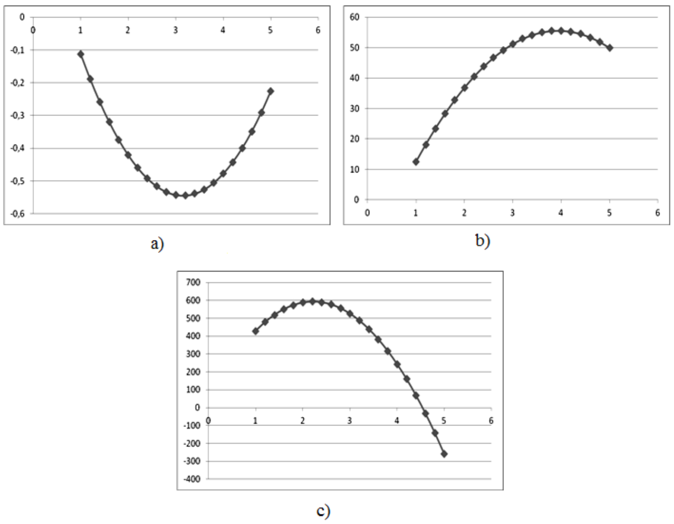

Figure 2. Dependencies of coefficients *a,b,c* of polynomial (1) on number of GPU

Approximating the numerical values of coefficients a, b, c on the base of data from table 2, we obtain the results of approximation presented in the following table (Table 9).

TABLE IX. RESULTS OF APPROXIMATION OF COEFFICIENTS OF EQUATION (1) IN DEPENDENCE ON THE NUMBER OF GPU

Variant of architecture	Coefficients of approximating polynomials of 2nd order					
	a		*b*		*c*	
2 nodes	a_1	0.0931	b_1	-5.0165	c_1	-110.27
	a_2	-0.587	b_2	39.459	c_2	489.4
	a_3	0.3811	b_3	-22.0	c_3	49.93

Thus, we obtained the functions *a (G), b (G), c (G)* in the following form:

$$a(G) = a_1G^2 + a_2G + a_3,$$
$$b(G) = b_1G^2 + b_2G + b_3, \quad (2)$$
$$c(G) = c_1G^2 + c_2G + c_3$$

where coefficients a1, a2, a3, b1, b2, b3, c1, c2, c3 are presented in the table 9.

Now we define the function F of two variables P and G, where P is the amount of RAM, and G is the number of GPU. For this purpose we use the equation (1). We substitute equations (2) in (1) and obtain:

$$F(P,G) = a(G)P^2 + b(G)P + c(G) =$$
$$(a_1G^2 + a_2G + a_3)P^2 + (b_1G^2 + b_2G + b_3)P + \quad (3)$$
$$+ (c_1G^2 + c_2G + c_3)$$

To check the correctness of obtained correlations let's calculate the performance of node variant with two GPU (*i.e. G = 2*) on the basis of (1) and (3).

First, define the coefficients *a(G), b(G), c(G)* using the equations (2) and coefficient values from Table 9:

$$a(G) = a_1G^2 + a_2G + a_3 =$$
$$0{,}0931G^2 - 0.587G + 0.3811 = -0.42053$$
$$b(G) = b_1G^2 + b_2G + b_3 = \quad (4)$$
$$-5.0165G^2 + 39.459G - 22.0 = 36.852$$
$$c(G) = c_1G^2 + c_2G + c_3 =$$
$$-110.27G^2 + 489.4 * G + 49.93 = 587.653$$

Results of calculations are presented in the table 10, and graphical results in the figure 3.

TABLE X. RESULTS OF CALCULATION OF ACTUAL PERFORMANCE MADE WITH THE HELP OF FORMULAS (1) AND (3)

№	Amount of RAM, GB	Performance calculated with help of formula (1), Tflops	Performance calculated with help of formula (4), Tflops
1	4.2	540.2	573.2688
2	6.1	757.7	794.9067
3	7.8	850.2	847.4248
4	8.5	963	868.3275
5	10.2	1 009.00	917.3368
6	13.7	1 077.00	1010.412
7	15.9	1 107.00	1063.525
8	17.6	1 144.00	1101.715
9	19	1 181.00	1131.3
10	21.6	1 207.00	1181.771
11	23.4	1 221.00	1213.307
12	25.7	1 233.00	1249.548
13	29.3	1 286.00	1297.14
14	33.54	1 304.00	1338.899
15	35.1	1 317.00	1350.373
16	40.1	1 349.00	1373.043
17	42.7	1 367.00	1376.334
18	46.3	1 380.00	1371.294
19	50.1	1 370.00	1353.883
20	54.7	1 367.00	1316.19
21	58.8	1 350.00	1267.257

ISBN: 978-1-63248-005-7 doi: 10.15224/ 978-1-63248-005-7-37

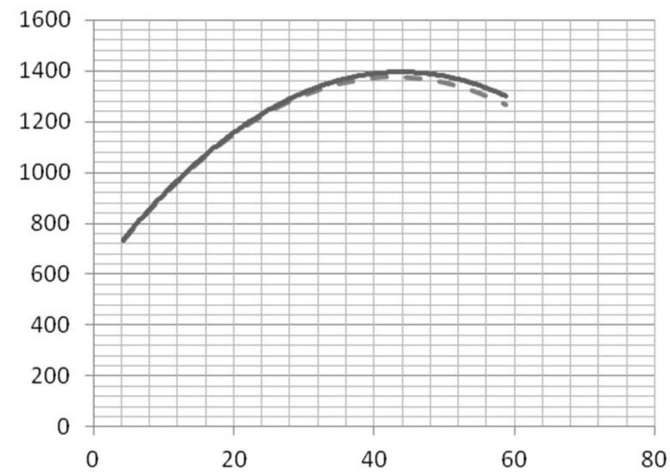

Figure 3. Diagrams of actual performance calculated with the help of formula (1) (solid line) and formula (3) (dot line)

V. Conclusion

Thus, it was obtained the formulas determining the dependence of the actual performance (TFlops) on the amount of RAM and the number of GPUs. It can form the basis for a computer simulation model which will allow «a priori» determine the optimum performance with the possibility of variation the amount of RAM and the number of GPUs, that can dramatically speed up the solution of the problems of definition the optimal parameters of HCCS architecture.

References

[1] NVidia Technical support, http://developer.nvidia.com

[2] Top500 Supercomputer sites, http://www.top500.org

[3] Computer performance tests and system software, http://www.parallel.ru/computers/benchmarks

[4] Akhmedov D. Determination of dependence of performance from specifications of separate components of the hybrid personal computing system based on GPU-processors / Akhmedov D., Yelubayev S., Abdoldina F., Bopeyev T., Muratov D., Povetkin R., Karataev A. // Proceedings 12th International Conference on Parallel Computing Technologies (PaCT 2013), –St. Petersburg, 2013,–P. 135-138.

[5] D. Akhmedov Research on performance dependence of cluster computing system based on GPU accelerators on architecture and number of cluster nodes / D. Akhmedov, S. Yelubayev, T. Bopeyev, F. Abdoldina, D. Muratov, R. Povetkin // Collection of scientific papers International conference " High Performance Computing 2013", – Kiev, 2013., –P. 9-13.

About Author (s):

Daulet Sh. Akhmedov
Doctor of technical science,
Director of
AALR “Institute of space technique and technology”

Suleimen A. Yelubayev
Head of laboratory of space systems simulation, AALR “Institute of space technique and technology”

Farida N. Abdoldina
candidate of technical science
Leading researcher of laboratory of space systems simulation, AALR “Institute of space technique and technology”

Timur M. Bopeyev
Senior scientist of laboratory of space systems simulation, AALR “Institute of space technique and technology”

Daulet M. Muratov
Leading software engineer of laboratory of space systems simulation, AALR “Institute of space technique and technology”