

## Improvement of WMS Functionality, Aiming to Minimize Processing Time of Jobs in Grid Computing

V. Pilkauskas, R. Plestys, G. Vilutis, D. Sandonavičius

Department of Computer Network, Kaunas University of Technology,  
Studentų g. 50, Kaunas, Lithuania, e-mail: rimantas.plestys@ktu.lt

**crossref** <http://dx.doi.org/10.5755/j01.eee.113.7.624>

### Introduction

At present, the ERT (Estimate Response Time) algorithm used in GRID computing distributes resources well only when there is a flow of identical jobs. However, Work Management Service (WMS) operates poorly when different types of jobs appear in GRID: many jobs sent to GRID network are lost due to the expiry of a Proxy or MyProxy certificate. The validity of a job certificate ends when a job is queued in a cluster, whereas GRID network contains free clusters.

There are several types of GRID networks, depending on the application area: data, cooperation and high performance computing GRID networks. Jobs in GRID network are divided into five types. These types of jobs are described in detail in [1-3]. Most attention in this paper is paid to jobs that require many computing resources. Each job may be defined by List of Accessible Clusters (LAC), Working Node Amount ( $WNA^{min}$ ) necessary for job execution and minimal permissible average Central Processor Speed ( $\overline{CPS}_C^{min}$ ) in working nodes of a cluster. In order WMS would choose a suitable cluster, besides these basic requirements for job execution, it is also necessary to indicate the maximum permissible average time for job execution ( $J_C^{max}$ ) and the maximum permissible amount of jobs queued in a cluster ( $J_{NR}^{max}$ ).

Different WMS algorithms are necessary for different job types. The right choice of GRID resources is a challenging task. We propose a WMS algorithm, based on QoS (Quality of Service), which directs jobs to clusters, according to priorities set and reduces TTD (Total Time to Delivery) of a job, considering separate terms during management.

For evaluation of the proposed WMS algorithm, it is useful to use emulators or simulators, which would evaluate heterogeneous computing resources. Simulation

models Optorsim and ChicSim [4, 5], which are oriented not at high performance computing but at data GRID networks, are used most often in summary studies. Although GridSim and SimGrid algorithms [4] are introduced as the most universal simulation tools of GRID network but GridSim algorithm provides incomplete information about job losses despite the fact that GridSim algorithm supports all types of GRID networks. GSSIM [6] and GridSim tool [7] allows to change WMS algorithm but there is a lack of documentation about how this can be performed. Real time jobs can be generated and their runtime can be observed with the help of the emulator MicroGRID when job flows are small. Therefore, it has been decided to create WMS simulator described in this paper.

### Job execution stages

TTD of a job in GRID network may be divided into three stages. Then

$$TTD = T^{IN} + T^C + T^{OUT}, \quad (1)$$

where  $T^{IN}$  – time necessary for sending [8] a job to GRID network,  $T^C$  – time necessary for a job execution,  $T^{OUT}$  – time necessary for downloading computing results. Each of these terms may be divided into parts shown in Fig. 1.

Job sending time is subdivided as follows

$$T^{IN} = T_{DTR}^{UW} + T_{WMS}^{WN} + T_{DTR}^{WC} + T_Q^C, \quad (2)$$

where  $T_{DTR}^{UW}$  – time for sending a job to WMS;  $T_{WMS}^{WN}$  – time when job is queued in WMS and time when WMS is searching for a suitable cluster;  $T_{DTR}^{WC}$  – time taken by forwarding a job from WMS to a cluster;  $T_Q^C$  – time when job is waiting in queue of a chosen cluster

$$T_Q^C(t) = f(J_{NR}, \overline{J_C}), \quad (3)$$

where  $J_{NR}$  - a random number of jobs queued in a cluster and  $\overline{J_C}$  - the average time of the job execution in a cluster. Time of a job execution  $T^C$  is shown below

$$T^C = f(WNA, \overline{CPS_C}, \overline{CPAC}, \overline{SS_C}, \overline{RAM_C}, Q_C), \quad (4)$$

which depends on Working Node Amount  $WNA$  in a cluster, Central Processor Speed in working nodes of a cluster  $\overline{CPS_C}$ , the average amount of processors in a working node  $\overline{CPAC}$ , the average size of data storage medium of a cluster working node  $\overline{SS_C}$ , the average amount of the main memory of a cluster working node  $\overline{RAM_C}$  and the demand of computing resources necessary for a job execution  $Q_C$ .

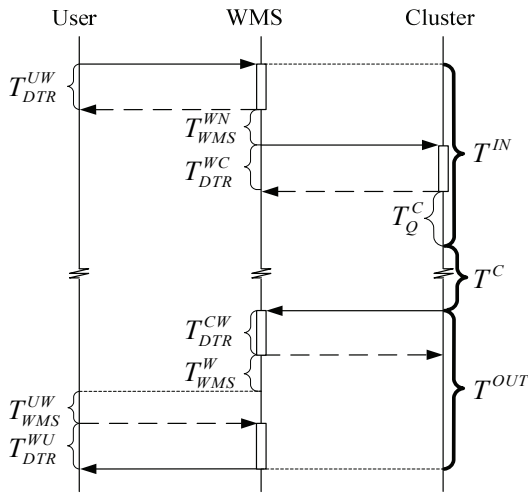


Fig. 1. Time states of a job in GRID network

Time  $T^{OUT}$  may be divided into terms

$$T^{OUT}(t) = T_{DTR}^{CW} + T_{WMS}^W + T_{WMS}^{UW} + T_{DTR}^{WU}, \quad (5)$$

where  $T_{DTR}^{CW}$  - time, when computing results are transferred from a cluster to WMS;  $T_{WMS}^W$  - time, when WMS checks information if all necessary data are received from a cluster;  $T_{WMS}^{UW}$  - time, during which a user learns about a job execution from WMS;  $T_{DTR}^{WU}$  - time necessary for transferring computing results from WMS to a user.

Terms of time of the job execution  $T_{DTR}^{UW}$ ,  $T_{DTR}^{WC}$ ,  $T_{DTR}^{CW}$  and  $T_{DTR}^{WU}$  influence on  $TTD$  is little. According to [9], time for the choice of a cluster  $T_{WMS}^{WN}$  may be shortened from 90 to 10 seconds. Time  $T_{WMS}^{UW}$  depends only on user's activeness in sending queries about the states of the jobs he sent to WMS.

Essential terms  $T_Q^C$  and  $T^C$  depend on the algorithm of the selection of WMS resources. Time  $T_Q^C$  depends on

a random number of jobs queued in a cluster ( $J_{NR}$ ) and the average time for the execution of a job in a cluster ( $\overline{J_C}$ ). This data can be received from specific monitoring systems. Data  $J_{NR}$  is received by methods used in GRID.  $\overline{J_C}$  is computed taking data statistically from a specific monitoring system. It is not enough to use only the parameter  $J_{NR}$  (a case of ERT algorithm) when choosing a cluster. We suggest evaluating the waiting time  $T_Q^C$  but not the queue length in a cluster. The cluster that has much longer queue may execute a job faster. The quantity of the product  $J_{NR} \times \overline{J_C}$  determines this.

### Selecting a cluster for a job execution

The cluster in GRID can be characterized by  $WNA$ ,  $\overline{CPS_C}$ ,  $\overline{CPAC}$ ,  $\overline{RAM_C}$ ,  $\overline{SS_C}$  and job execution parameters  $J_{NR}$ ,  $\overline{J_C}$  and the speed of a job transfer to a cluster  $DTR_C$ . It is possible to reject clusters that do not satisfy minimal requirements for a job execution (Eg.  $WNA$ ,  $\overline{CPS_C}$  or  $J_{NR}$ ). Execution parameters, calculated from previous jobs of all clusters and the present state of each cluster, are evaluated when choosing a suitable cluster according to parameters of a job.

When conditions  $V_{nj}^M \leq V_{kj}^{MR}$  and  $V_{ni}^D \geq V_{ki}^{DR}$  ( $i = \overline{1, I}$ ,  $j = \overline{1, J}$ ) are satisfied (where  $n$  - a number of a cluster, ( $n = \overline{1, N}$ );  $k$  - job number ( $k = \overline{1, K}$ );  $V_{ni}^D$  -  $n$ -th cluster's  $i$ -th parameter value, where  $D$  marks a parameter, and when this parameter value increases, qualitative estimate increases;  $V_{nj}^M$  -  $n$ -th cluster's  $j$ -th parameter value, where  $M$  marks a parameter, and when this parameter value increases, qualitative estimate decreases;  $V_{ki}^{DR}$  and  $V_{kj}^{MR}$  -  $i$ -th and  $j$ -th parameters' limit values set by a user, which indicate requirements set for  $k$ -th job execution), then cluster suitability for a job may be determined by evaluating the quality of GRID cluster  $Q_{nk}$  by

$$Q_{nk} = \frac{\sum_{i=1}^I \left( \omega_{ki}^D \cdot \frac{V_{ni}^D}{V_i^{D \max}} \right) + \sum_{j=1}^J \left( \omega_{kj}^M \cdot \frac{V_j^{M \max} - V_{nj}^M}{V_j^{M \max} - V_j^{M \min}} \right)}{\sum_{i=1}^I \omega_{ki}^D + \sum_{j=1}^J \omega_{kj}^M}, \quad (6)$$

where  $V_i^{D \max} = \max_{n=1, \dots, N} V_{ni}^D$  - the highest value of parameter  $i$ -th in GRID network.  $V_j^{M \max} = \max_{n=1, \dots, N} V_{nj}^M$  is the highest and  $V_j^{M \min} = \min_{n=1, \dots, N} V_{nj}^M$  is the lowest value of parameter  $j$ -th in GRID network.  $\omega_{ki}^D$  and  $\omega_{kj}^M$  - weights of  $i$ -th and  $j$ -th parameters, set by  $k$ -th job.

Otherwise, when conditions  $V_{nj}^M > V_{kj}^{MR}$  or



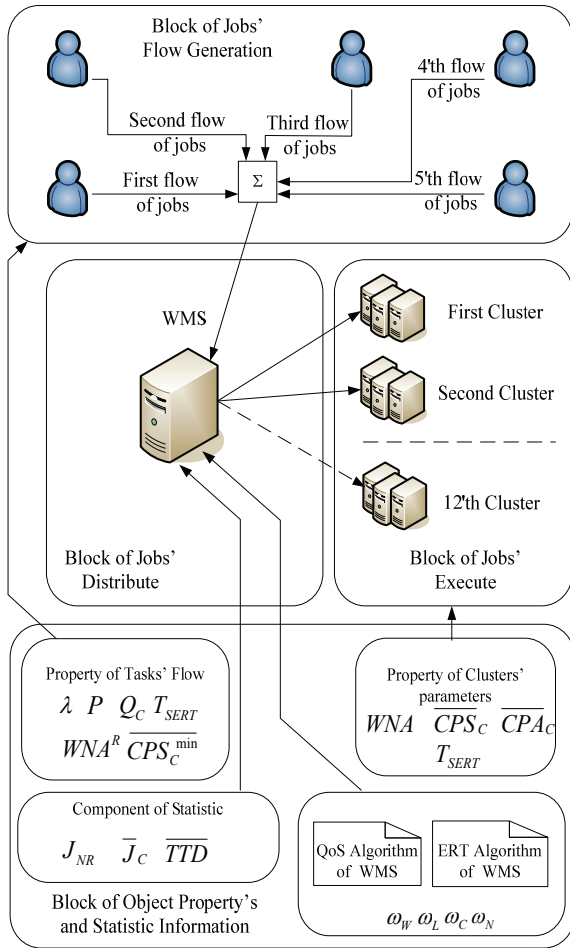


Fig. 3. Simulation Model Diagram

The following factors were taken into consideration when generating the job flow in the model:

- Intensity of job sending, irrespective of the job type. 5 job types that need different amounts of GRID resources: the minimal processor speed; the amount of free processors that will be necessary for a job execution; the expiry of Proxy or MyProxy certificate (12 hours/96 hours).

The following is evaluated while simulating the workload of the created job flow:

- WMS algorithm;
- Cluster capacity (the average amount of processors in clusters -  $CPA_C$ ; the amount of processors; central processor speed -  $CPS_C$ ; job queues in clusters  $\bar{J}_C$ );
- Job service algorithm in a cluster - FiFo.

## Experiment

The initial conditions set down during the experiment are the following: two identical job flows are formed; a time slot between jobs is distributed according to the normal law; all jobs in a network are identical and they may be executed on one cluster element (working node); job intensity – 6 jobs per hour; the experiment sample is 1,000,000 jobs; WMS may service up to 1,500 jobs at one time; the search time for a suitable cluster is from  $\frac{1}{2}$  to  $1\frac{1}{2}$

minute; when a job reaches a cluster with the expired proxy certificate, the waiting time for a valid certificate is 21 hours; proxy certificates are valid for 12/96/168 hours;

**Preliminary experiments.** The initial purpose of experiments is to determine the influence of  $\omega_i$  and  $\omega_j$  on the cluster choice. Every cluster's values, such as  $WNA$ ,  $CPS_C$ ,  $J_{NR}$  and  $\bar{J}_C$ , as well as time averages of job execution of the first and second flow  $\overline{TTD}^1$  and  $\overline{TTD}^2$  and the total time average of both flows  $\overline{TTD}^\Sigma$ , were evaluated.

Five experiments were carried out. The first four experiments were carried out when jobs of the first flow are directed to clusters according to WMS algorithm, evaluating  $Q_{nk}$ ; when both flows are directed according to the algorithm, evaluating queues, and when both flows are directed according to the algorithm, evaluating  $Q_{nk}$ . During every experiment a set of weight coefficients  $\omega_W$ ,  $\omega_L$ ,  $\omega_C$ ,  $\omega_N$  is chosen ( $\omega_W$  – is the weight coefficient of the parameter  $\bar{J}_C$ ;  $\omega_L$  – is the weight coefficient of the parameter  $J_{NR}$ ;  $\omega_C$  – is the weight coefficient of the parameter  $CPS_C$  and  $\omega_N$  – is the weight coefficient of the parameter  $WNA$ ). The set  $\omega_W$ ;  $\omega_L$ ;  $\omega_C$  will be chosen so that according to the proposed WMS algorithm, a cluster would be purposefully chosen in the case of every job sent.

$CPS$  influence is evaluated during the first experiment (Table 1, Experiment 1). Two clusters with the same amount of processors of different speed are taken. If a user chooses WMS algorithm, evaluating cluster QoS, then the time of the jobs he sent significantly gets better.

The influence of  $WNA$  parameter will be determined during the second experiment. In this case, two clusters that consist of identical servers are taken but the amount of the working nodes in clusters is different. Results improve in a similar way as during the first experiment. During the third experiment, the influence of the parameter  $\bar{J}_C$  is determined. Clusters identical to the second experiment are taken but coefficients  $\omega_W$ ;  $\omega_L$ ;  $\omega_C$  and  $\omega_N$  are chosen with the main attention paid to influence of queued parameter  $\bar{J}_C$  on choosing clusters for jobs according to WMS algorithm and evaluating QoS. Since we suggest to evaluate several parameters in the algorithm described in Fig. 3, not only  $\bar{J}_C$ , thus relatively small  $\overline{TTD}^2$  improvement is noticed. The conclusion may be drawn from the results of this experiment that it is better to choose clusters according to the parameter  $\bar{J}_C$  than according to  $J_{NR}$ . The influence of the parameter  $J_{NR}$  is determined during the fourth experiment. Identical clusters, enabling to eliminate the influence of parameters  $WNA$ ,  $CPS_C$  and  $\bar{J}_C$  are taken. In the case of choosing both WMS algorithms, identical results are achieved.

The obtained results enable to draw a conclusion that ERT algorithm, evaluating queues, serves well only when there are clusters with the same parameters in GRID network and identical jobs are sent to GRID network.

**Table 1.** Influence of Weight Coefficients in Cluster Choice

Experiment No/ Cluster parameters/ QoS weight coefficients	$\overline{TTD}^1$ , min.	$\overline{TTD}^2$ , min.	$\overline{TTD}^\Sigma$ , min.	Algorithm
Experiment No. 1 $WNA_1=14; CPS_1=1$ GHz, $WNA_2=14; CPS_2=2$ GHz, $\omega_W=5, \omega_L=5, \omega_C=3, \omega_N=4$	497.14		403.69	WMS algorithm, evaluating queues
		310.24		WMS algorithm, evaluating QoS
	392.61	392.53	392.57	WMS algorithm, evaluating QoS
	434.73	436.9	435.815	WMS algorithm, evaluating queues
Experiment No. 2 $WNA_1=7; CPS_1=2$ GHz $WNA_2=14; CPS_2=2$ GHz $\omega_W=9, \omega_L=5, \omega_C=3$ and $\omega_N=4$	415.98		349.17	WMS algorithm, evaluating queues
		282.37		WMS algorithm, evaluating QoS
	348.21	350.96	349.58	WMS algorithm, evaluating QoS
	364.08	368.83	366.45	WMS algorithm, evaluating queues
Experiment No. 3 $WNA_1=7; CPS_1=2$ GHz $WNA_2=14; CPS_2=2$ GHz $\omega_W=10, \omega_L=1, \omega_C=0$ and $\omega_N=0$	408.61		343.31	WMS algorithm, evaluating queues
		278.01		WMS algorithm, evaluating QoS
	331.077	332.92	332.00	WMS algorithm, evaluating QoS
	379.4	378.07	378.73	WMS algorithm, evaluating queues
Experiment No. 4 $WNA_1=7; CPS_1=2$ GHz $WNA_2=14; CPS_2=2$ GHz $\omega_W=5, \omega_L=5, \omega_C=3$ and $\omega_N=4$	206.47		206.35	WMS algorithm, evaluating queues
		206.24		WMS algorithm, evaluating QoS
	206.88	206.87	206.87	WMS algorithm, evaluating QoS
	206.94	206.97	206.95	WMS algorithm, evaluating queues

During the fifth experiment, the job intensity was increased from 10 to 70 jobs per hour and efficiency of the resource search method, which has been described, in comparison with the standard method, was observed. During every simulation of GRID network, the total flow of 10, 0000 jobs, which consists of five job flows that are of the same intensity but of different parameters, is generated (Table 2).

**Table 2.** Parameters of Generated Job Flows

Job flow No.	1	2	3	4	5
Parameter					
Part of flow in total flow	1/5	1/5	1/5	1/5	1/5
$\overline{CPS}_C^{\min}$ , GHz	1	1	1,5	1	2
$WNA^R$ limits	35-40	10-20	15-25	5-15	35-45
$Q_C$ , min	30-30	40-40	50-70	50-50	10-30

Parameter values of different Working Node Amount Requirement ( $WNA^R$ ) and  $Q_C$  are taken from Table 2.  $Q_C$  is chosen from the interval indicated in the table. The resource  $Q_C$ , which is necessary for 1 GHz processor to execute a job, is presented in minutes.

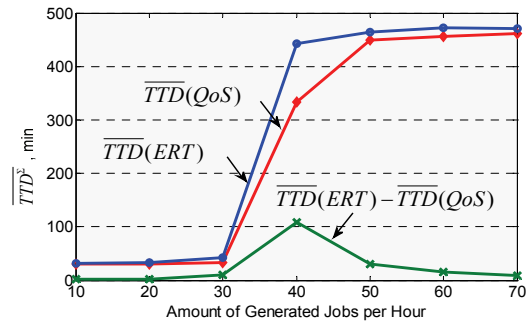
Parameters of different clusters  $\overline{CPS}$  and  $WNA$  are shown in Table 3.

**Table 3.** Cluster parameters

Cluster No	1	2	3	4	5	6	7
Parameter							
$\overline{CPS}_C$ , GHz	2	2	2	1,5	2	2	2
$WNA$	25	35	45	55	40	45	100

The experiment was carried out twice. In the first case, all jobs were divided according to the standard ERT algorithm of WMS. In the second case, clusters were chosen according to WMS QoS algorithm only for the jobs of the third flow. The value of the parameter  $\overline{TTD}^\Sigma$  was

recorded as the intensity of job sending (10; 20; 30; 40; 50; 60 and 70 jobs per hour) has been changed. Results are presented in Fig. 4.



**Fig. 4.** Advantage of QoS algorithm dependence on job sending intensity over ERT algorithm

Under present conditions, the advantage of the proposed method is gained only when the flow intensity equals 40 jobs per hour. It is possible to refuse 73 servers if the proposed algorithm is used and only jobs of the third flow are served; in this way energetic resources are saved or these servers are used for other purposes.

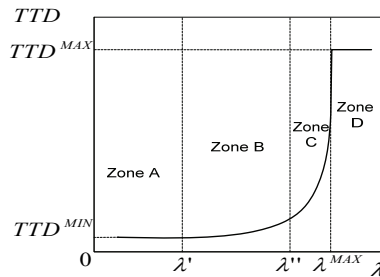
### Generalization of obtained results

WMS algorithm for job distribution, presented in Fig. 2, has meaning only when the intensity of sent jobs  $\lambda$  reaches a certain value.

The higher  $\lambda$ , the more important is to direct jobs suitably to clusters of GRID network. Four intensity zones may be distinguished (Fig. 5).

Thus, if jobs of at least one user are directed according to this WMS algorithm, then not only this user's jobs will be serviced faster but also all resources of GRID network will be loaded equally. This is determined by the characteristic of the proposed WMS algorithm, which evaluates  $T_Q^C$ . Since GRID resources will be used more, in

this case, the average Total Time to Delivery of jobs sent by all users  $\overline{TTD}$  will be much shorter.



**Fig. 5.** Service dependence of computing jobs on intensity of sent jobs

In the zone B ( $\lambda_1 < \lambda < \lambda_2$ ) at certain moments all clusters operate in full and there are job queues in them. However, there are moments when there are no queues in all clusters or in some of them (i.e.  $J_{NR} = 0$ ) and clusters are free. In case of choosing a cluster, it is purposeful to use WMS algorithm that chooses such cluster where the total time  $T_Q^C + T^C$  is the shortest. Due to different cluster configuration and capacity, there are cases when the total time of a job  $T_Q^C + T^C$  in one cluster will be shorter than the delivery time of the same job  $T^C$  in other cluster.

In the zone C ( $\lambda_2 < \lambda < \lambda^{MAX}$ ), all clusters are loaded fully during their operation time and time  $\overline{TTD}$  will depend little on the chosen WMS algorithm. In this case, if a concrete user uses the proposed WMS algorithm, his jobs in the total job flow will be serviced faster. However, in this case, there are some rejected jobs, which are rejected due to the expired proxy certificate that is valid for time  $T_{SERT}$ . In this case, time of some jobs is  $T^{IN} \geq T_{SERT}$ . Using the proposed WMS algorithm for choosing a job, the amount of the rejected jobs decreases much because in choosing a cluster time  $T_Q^C$  is evaluated. As a consequence, GRID network will service more jobs, although time  $\overline{TTD}$  will not practically change.

In zone D ( $\lambda > \lambda^{MAX}$ ),  $\overline{TTD} \rightarrow TTD^{MAX} = T_{SERT}$ . Usefulness of the proposed WMS algorithm remains the same as in the third zone.

## References

1. **Benkner S., Engelbrecht G.** A service-oriented Grid environment with on-demand QoS support // Proceedings of the 7th IEEE Int. Conference on Web Services (ICWS'2009). – Los Angeles, CA, USA, 2009. – P. 147–150.
2. **Chrisochoides N., Fedorov A., Kot A., Archip N., Black P., Clatz O., Golby A., Kikinis R., Warfield S. K.** Toward Real-Time Image Guided Neurosurgery Using Distributed and Grid Computing // Proceeding of the conference on Supercomputing. – ACM, New York, 2006. – P. 9–13.
3. **Van Ruler B.** The communication grid: an introduction of a model of four communication strategies // Public Relations Review June, 2004. – Vol. 30. – No. 2. – P. 123–143.
4. **Buyya R., Sulistio A.** Service and Utility Oriented Distributed Computing Systems: Challenges and Opportunities for Modeling and Simulation Communities // Proceedings of the 41st Annual Simulation Symposium (ANSS-41), 2008. – P. 68–81.
5. **Casanova H., Legrand A., Quinson M.** SimGrid: a Generic Framework for Large-Scale Distributed Experiments // Tenth International Conference on Computer Modeling and Simulation (UKSIM'2008), 2008. – P. 126–131. DOI: 10.1109/UKSIM.2008.28.
6. **Ungurean I.** Job Scheduling Algorithm based on Dynamic Management of Resources Provided by Grid Computing Systems // Electronics and Electrical Engineering. – Kaunas: Technologija, 2010. – No. 7(103). – P. 57–60.
7. **Sulistio A., Cibej U., Venugopal S., Robic B., Buyya R.** A toolkit for Modelling and Simulating Data Grids: An Extension to GridSim. In: Concurrency and Computation: Practice and Experience (CCPE'2007), 2007. – P. 1591–1609.
8. **Eidukas D., Valinevičius A., Vilutis G., Kilius Š., Vasylius T.** Duomenų perdavimo tinklo apkrautumo skaičiavimas // Elektronika ir Elektrotechnika. – Kaunas: Technologija, 2005. – Nr. 8(64). – P. 22–26.
9. **Elmroth E., Tordsson J.** A grid resource broker supporting advance reservations and benchmark-based resource // Lecture Notes in Computer Science. – Springer, New York, 2006. – Vol. 3732. – P. 1061–1070.

Received 2011 06 26

**V. Pilkauskas, R. Pleštys, G. Vilutis, D. Sandonavičius.** Improvement of WMS Functionality, Aiming to Minimize Processing Time of Jobs in Grid Computing // *Electronics and Electrical Engineering*. – Kaunas: Technologija, 2011. – No. 7(113). – P. 111–116.

WMS (Workload Management System) algorithm of functionality improvement, based on quality of a cluster and used in performing large scale computing in GRID network, is presented. Limitations for a minimal number of cluster nodes, a minimal number of central processors in each node, the largest number of jobs queued and the longest acceptable average time for job execution are defined in distributed jobs. The total time of time, when a job waits in queue, and time, when it is executed, is evaluated. Validity of the job certificate and a list of clusters are evaluated additionally. A simulator that evaluates the mentioned factors was created for researching WMS characteristics. Ill. 5, bibl. 9, tabl. 3 (in English; abstracts in English and Lithuanian).

**V. Pilkauskas, R. Pleštys, G. Vilutis, D. Sandonavičius.** WMS funkcionalumo gerinimas skaičiuojamuosiuose Grid tinkluose, siekiant minimizuoti užduočių vykdymo trukmę // *Elektronika ir elektrotechnika*. – Kaunas: Technologija, 2011. – Nr. 7(113). – P. 111–116.

Pristatomas WMS veikimo algoritmas, kurio veikimas pagrįstas resursų teikiamų paslaugų kokybe. Siunčiamuose į Grid tinklą užduotyse nurodomi apribojimai minimaliam klasterio darbinį mazgų kiekiui, kiekvieno darbinio mazgo minimaliam procesorių kiekiui, maksimalus užduočių eilės ilgis ir ilgiausias laikas, kurį užduotis gali užtrukti eilėje ir būti sprendžiama. Įvertinamas bendras užduoties stovėjimo eilėje ir vykdymo laikas. Papildomai įvertinamas ir sertifikato galiojimas bei prieinamų klasterių sąrašas. WMS charakteristikų tyrimui buvo sukurtas imitacinis modelis. Il. 5, bibl. 9, lent. 3 (anglų kalba; santraukos anglų ir lietuvių k.).