Cost Optimization and Quality of Service Assurance in WAN-based Grid System

Marcin Markowski
Department of Systems and Computer Networks
Wrocław University of Technology
Wrocław, Poland
marcin.markowski@pwr.wroc.pl

***Abstract*—Grid systems become the common solutions supporting scientific researches, business analysis, entering even the world of entertainment. Modern grids should satisfy following requirements to be considered as useful and convenient tools: cost-effectiveness, performance, quality of service, reliability, security. The critical part of grid system is a communication network, responsible for proper and reliable data transfer between computing nodes. In case of grids where data transfer delays (QoS) and reliability are not crucial, public networks (usually in conjunction with VPN technology) are used. For better performance and stable communication parameters private wide area networks (WANs) are build and utilized in grid environments. WAN-based grid systems require specific designing methods in order to ensure the cost and QoS optimality. Some problems typical for wide area networks must be considered: topology of private WAN, capacity of channels, routing (flow assignment), allocation of computing nodes. In the paper the model of WAN-based grid system is presented and the cost minimizing and QoS ensuring algorithm is proposed. Both network optimization issues and distributed computing optimization issues are considered simultaneously. Algorithm may be useful as well for designing as optimizing of WAN-based grids.**

Keywords-grid networks; capacity and flow assignment problem; distributed computing; wide area networks; algorithm

I. INTRODUCTION

Computational power demanded for latest applications is incessantly increasing. Huge amounts of data are collected and analyzed for different purposes: scientific (i.e., biomedical simulation), business (statistics, trends analyzing), security (pattern, i.e., malware recognition) and other. Since computational requirements often exceed the possibilities of single host, then distributed solutions like grid computing and cloud computing become more and more useful and popular. With distributed technologies, usually unused computing resources may be better utilized. Many hosts in public networks use only few percents of processing power (examples may be DNS servers or even production servers after usual work hours). In data centers, it is common that servers use only 20%-30% of processing power on average [1]. Other hosts and clusters, dedicated for computing purposes, are not loaded all the time, some services are less utilized during nights and the computing power is being wasted. Grid technology allows utilizing spare computing resources distributed in remote locations, and computations may be done in parallel on many physical systems.

Modern grids should satisfy following requirements to be considered as useful and convenient tools: latency, bandwidth, reliability, fault-tolerance, jitter control and security [7]. Then, the critical part of grid system is a communication network, responsible for proper and reliable data transfer between computing nodes. Communication between nodes may be based on public network. In this case it is rather difficult to ensure minimal data transfer delays (QoS) and reliability, because public lines are shared with other users and are susceptible to overloading and security attacks. For better performance and stable communication parameters private wide area networks (WANs) are build and utilized for grid environments.

In grid optimization issues, that may be found in the literature, different task scheduling problems are considered [2, 3, 4], denoting that the structure, capacity and flow routes in the communication network are given. Such solutions are not able to ensure QoS and reliability in the network layer. There is a lack of solutions for simultaneously optimizing of task scheduling and network parameters (i.e., topology, capacity and flow). The problem of WAN-based grid optimization, considered in the paper, represents more complex approach, taking into account the structure of the communication network. The problem consist in task assignment to grid nodes simultaneously with network capacity and flow routes assignment. The combined optimization criterion includes the computing cost and the network cost. The optimization parameters are: task assignment to nodes of grid, location of grid management centre, capacities of channels and flow routes.

The paper is organized as follows. In Section III the optimization problem has been formulated and explained in details. The mathematical model of WAN-based grid, with decision variables and constraints has been presented in Section IV. Approximate algorithm for considered problem is proposed in Section V. Section VI concludes the paper.

II. RELATED WORK

Issues in designing and optimizing of grid systems are well known in literature [2-6]. Itami *et al* [3] relates to real-time distributed systems, where the rapid and reliable communication between nodes is critical. The authors proposed event-triggered distributed object models and developed a distributed computing environment. Sterritt *et al.*

[4] consider the problem of autonomic computing. Autonomic computing refers to self-managing and self-configuring distributed and grid systems that are able to grow and increase they complexity without or with only the little involvement of administrator.

Grid optimizing solutions, that may be found in literature do not take into consideration simultaneous optimization of tasks assignment and network capacity. Usually overlay network, connecting grid nodes is being considered and bandwidth to each computing node is denoted. Such simplification allows to construct simpler algorithms, but the performance of solutions is strongly dependent on topology and capacity of communication network. Building grid networks on WAN base implies additional optimization issues. Designing of WANs involves such problem as: topology assignment, channels' capacities assignment and routing assignment (known also as flow assignment). In the classical capacity and flow assignment (CFA) problem, there are two design variables: channel capacities and flow routes (routing) [8, 9, 10]. The goal is to select those variables in order to minimize the criterion function, for example the total average delay per packet in wide area network or the capacity leasing cost [10]. Algorithm proposed in the paper allows to obtain better results, related to grid algorithms known in literature, because network capacity and network routing are considered and optimized. The advantages of such approach are better fixing of communication network to grid demands, less QoS and reliability inconveniences, finally lower costs of supporting the grid network.

Optimization problem in distributed computing systems, with different criterions and constraints set was the subject of our previous considerations [11]. The model of distributed computing system was proposed and optimization problem with the criterion combined of quality of service (indicated by average packet delay) and computing cost was solved. The WAN-based grid model, proposed in [11] is used in the paper in order to formulate an optimization problem and design an approximate algorithm.

III. WAN-BASED GRID OPTIMIZATION PROBLEM

Consider grid system build of certain number of computing nodes connected to nodes of the WAN network. In each time period (in example each second) the portion of data is generated and requires to be processed. It has to be divided into separable parts (usually called blocks or subtasks) an sent to proper computing nodes. We denote than computational outlay needed in each time period is similar and each generated computational task may be easily divided into subtasks. We denote that during processing of particular block computing nodes do not have to exchange any information with nodes processing other subtasks. After processing, result data must be sent to the main node and the final result is compiled from all subtasks. The main node, managing realization of computational tasks we call the computing management centre (CMC). Management centre divides tasks into blocks, transmits blocks to grid nodes, receives and collects result data for each block, and compiles final result. The grid node, in which CMC is located, also takes part in blocks processing. Blocks and results transferred between this node and CMC are not sent through the network. Each grid node communicates only with computational centre node.

Processing block of data is connected with specified processing cost. Cost may represent money (for buying computational power), resource utilization or other virtual costs. Costs are specified for each of the blocks. Blocks may differ in computational effort needed to process them, then processing costs may also differ. Management centre is located in one of the grid nodes, since it is source or destination of all data transmitted in the network, then the proper allocation of CMC has a critical impact on the quality of service in the network. Maintaining of CMC in the node generates also some maintaining cost in each time period. Minimizing of total processing cost and maintaining cost is one of the objectives of grid optimization.

The computational power resources (also called resource capacity [6]) of grid nodes are limited and determined for each node. It is denoted that the total resources capacity of grid is enough to presses all generated blocks.

For each block two parameters must be specified: size of the block and size of the results. Size of the block is the amount of data that must be sent from CMC to the computing node chosen to process the block. Size of results generated during block processing is the amount of data that size must be sent from processing node to CMC when processing is finished.

It is assumed that communication network for grid system is the packed switched network. Channel allocation (network topology) is given, capacity of each channel must be assigned in optimization process. Possibility of assignment of channels capacity is very important especially for CMC node, which generates and receives all data sent between nodes. We may ensure enough bandwidth to computing management centre node in order to operate all transmissions.

On the basis of above assumptions, considered problem we formulate as follows:

Given:

- number of grid nodes, number of channels of wide area network,
- for each node: computational capacity of node,
- for each channel the set of possible capacities and costs (i.e., cost-capacity function),
- list of candidate nodes for Computing Management Centre, for each candidate node the value of maintaining cost,
- number of blocks which must be processed in each time period,
- for each block: size of block, size of result data, computational outlay needed to process block, costs of processing at each computing node,
- maximal acceptable average delay in the network.

Minimize:

- linear combination of supporting cost of the network (capacity leasing cost) and the total computing cost (total cost of data processing and maintaining cost of management centre).

Over:

- CMC allocation,
- block allocation at computing nodes,
- channel capacities,
- flow routes (routing).

Subject to:

- channel capacity constraints
- multicommodity flow constraints
- computing capacity of nodes constraints
- QoS (delay) constraints

We assume that channels' capacities can be chosen from discrete sequence defined by ITU-T (International Telecommunication Union – Telecommunication Sector) recommendations. The formulated above problem is NP-complete, as more general that capacity and flow assignment problem [10, 12].

IV. Model of WAN-based Grid System

In this section, the mathematical model of the WAN-based grid systems is presented. Then the optimization criterion is proposed and optimization problem is formulated. Developing of the system model is necessary in order to implement optimization algorithms. The model is based on graph theory. Selection of channel capacity, CMC allocation and block-to-node allocation are modeled with binary decision variables. Important constraint, connected to decision variables are also introduced. Model of WAN-based distributed system was developed and presented in our previous work [11]. Model presented in this section is an adaptation of the previous one.

A. Variables and Parameters

Let n be the number of computation nodes of the considered grid system. Let b be the number of channels of the wide area network connecting grid nodes. For each channel i there is the set $C^i = \{c_1^i, ..., c_{s(i)}^i\}$ of alternative values of capacities. Let $D^i = \{d_1^i, ..., d_{s(i)}^i\}$ be the set of leasing costs corresponding to channel capacities from the set C^i. Let x_j^i be the discrete decision variable, connected with capacity choice for channel i:

$$x_j^i = \begin{cases} 1, & \text{if the capacity } c_j^i \text{ is assigned to channel } i \\ 0, & \text{otherwise} \end{cases}$$

Exactly one capacity from the set C^i must be chosen for channel i, then the following condition must be satisfied [11]:

$$\sum_{j=1}^{s(i)} x_j^i = 1 \quad \text{for} \quad i = 1,..,b \tag{1}$$

Let X_r be the set of all variables x_j^i which are equal to one. r is the number of iteration, since in successive chapters we propose approximate iterative algorithm.

Let y_h be the discrete decision variable, connected with allocation of management node:

$$y_h = \begin{cases} 1, & \text{if the management centre is located in node } h \\ 0, & \text{otherwise} \end{cases}$$

Let Y_r be the set of all variables y_h which are equal to one. Since the management centre is located in one node only, the following condition must be satisfied [13]:

$$\sum_{h=1}^{n} y_h = 1 \tag{2}$$

Let u_h be the cost of maintaining of the management centre in the node h.

Let t denotes the number of blocks, which must be proceeded in the considered time period. The following given data are connected with each block:

- v_l is the size of block l. In each time period the data of size v_l must be sent from management node to the proper computing node;
- w_l is the size of data generated as result of proceeding block l. Data of size w_l must be sent from computing node to the management node at each time period;
- p_l is the computational outlay needed to proceed block l – measured in [instructions];
- q_m^l is the cost of proceeding block l in the computing node m, $Q^l = \{q_1^l, ..., q_n^l\}$ is the set of proceeding costs of block l in all computing nodes.

Let z_m^l be the discrete decision variable, connected with grid node choice for proceeding block l:

$$z_m^l = \begin{cases} 1, & \text{if block } l \text{ is proceeded in grid node } m \\ 0, & \text{otherwise} \end{cases}$$

Let Z_r be the set of all variables z_m^l which are equal to one. In [11], we have proposed the following condition, that must be satisfy in order to ensure that each of the blocks is proceeded exactly in one node:

$$\sum_{l=1}^{t} \sum_{m=1}^{n} z_m^l = t \tag{3}$$

Each of the grid nodes has a limited computing power. Let e_m be the computing power (also called computational capacity) of node m, given in [instructions per second] (IPS). In order to guarantee that the optimization problem has a solution, we propose the following condition which must be satisfied for given data:

$$\sum_{m=1}^{n} e_m > \sum_{l=1}^{t} p_l \tag{4}$$

Let r_{mk} be the average traffic rate sent from node m to node k in each time period. In packet switched networks the flow between nodes is realized as a multicommodity flow.

Since, we denote that in the considered networks only proceeding block and results of proceedings block are sent (there is no other network traffic) then r_{mk} consists of packed exchanging between computing nodes and management node only. r_{mk} we calculate as follows:

$$r_{mk} = \sum_{l=1}^{t} \left(z_m^l y_k w_l + z_k^l y_m v_l \right)$$

The triple of sets (X_r, Y_r, Z_r) is called a selection. Let $\Re$ be the family of all selections. The selection (X_r, Y_r, Z_r) defines the unique wide area network and distributed computer system, because:

- X_r determines the values of capacities for channels of the WAN,
- Y_r determines the allocation of CMC at the node of WAN.
- Z_r determines the blocks' allocation at the grid nodes of distributed grid.

B. Criteria and Constraints

Let $T(X_r, Y_r, Z_r)$ be the minimal average delay per packet in the wide area network in which values of channel capacities are given by set X_r and traffic requirements are given by sets Y_r and Z_r (depend on management node allocation and assigning of block to computing nodes). $T(X_r, Y_r, Z_r)$ can be obtained solving a multicommodity flow problem in the network [10, 12]:

$$T(X_r, Y_r, Z_r) = \min_{\underline{f}} \frac{1}{\gamma} \sum_{x_j^i \in X_r} \frac{f_i}{x_j^i c_j^i - f_i}$$

subject to:

- $\underline{f}$ is a multicommodity flow satisfying the traffic requirements r_{mk} given by Y_r and Z_r ,
- $f_i \le x_j^i c_j^i$ for every $x_j^i \in X_r$
- $f = [f_1, ..., f_b]$ is the vector of multicommodity flow, f_i is the total average bit rate on channel i , and γ is the total packet rate generated and sent through the network by computational nodes and management node.

Let $A(Y_r, Z_r)$ is the computing cost, composed of proceeding costs of block at computational nodes and cost of maintaining of management node. We propose to calculate it as follows:

$$A(Y_r, Z_r) = \sum_{h=1}^{n} y_h u_h + \sum_{l=1}^{t} \sum_{m=1}^{n} q_m^l$$

Let $B(X_r)$ be the regular leasing cost of channel capacities, given with the formula:

$$B(X_r) = \sum_{x_j^i \in X} x_j^i d_j^i$$

Then, the we propose following objective function for the channel capacities, routing assignment and location of management centre and task scheduling problem:

$$OBJ(X_r, Y_r, Z_r) = \sigma A(Y_r, Z_r) + B(X_r)$$

Computing cost may not be the money but also computational power, CPU utilization or other abstract cost. Network maintaining cost is usually the money.

Let $T^{\max}$ be the maximal acceptable average packet delay in the network. $T^{\max}$ defines the level of quality of service (QoS) in the grid network. Quality of service in the network become degraded when average packet delay exceeds $T^{\max}$.

C. Problem Formulation

Above definitions allow us to formulate the WAN-based grid optimization problem:

$$\min_{(X_r, Y_r, Z_r)} OBJ(X_r, Y_r, Z_r) \quad (5)$$

Subject to:

$$(X_r, Y_r) \in \Re \quad (6)$$

$$T(X_r, Y_r, Z_r) \le T^{\max} \quad (7)$$

$$\sum_{l=1}^{t} z_m^l p_l < e_m \text{ for each } m = 1,..,n \quad (8)$$

V. Approximate Algorithm

The problem (5)-(8) is NP-complete, as more general than classical CFA problem, which is NP-complete [12, 13]. NP-completeness is defined as the set of decision problems that can be solved in polynomial time on a nondeterministic Turing machine (Nondeterministic-Polynomial time). It means that complexity of the problem increases very quickly as the size of the problem (for example number of possible values of capacity for each channel) grows.

In the paper, an approximate algorithm for problem (5)-(8) is proposed. Unlike exact algorithms, approximate ones usually are able to find suboptimal solutions, not far from optimal. An advantage of approximate algorithms is the computational time – finding optimal solution for NP-complete problems takes very long time. Some exact algorithms for different WAN optimization problems were proposed by Markowski and Kasprzak [13, 14, 15].

Proposed algorithm starts with assignment of an acceptable solution of the problem. Acceptable solution is the selection (X_r, Y_r, Z_r) and flow vector $\underline{f}$, satisfying conditions (6) - (8). For the considered problem, finding acceptable solution consist in allocating of computing management centre, assignment tasks to computing nodes in order to satisfy requirement (8) and find solution for CFA problem satisfying (7). In case that the problem has no solution (i.e., it is impossible to build the network satisfying flow demands with given budget restriction), it is discovered during this stage and algorithm finishes. The second phase of an algorithm is optimization phase, while sub-optimal solution is being found.

A. Initial Solution

Three main tasks appear during initial phase: choosing allocation for computing management centre, tasks allocation over the computing nodes, capacities of channels and flow assignment.

Few strategies for choosing the node for CMC allocation in distributed computing systems were proposed in [11]:

- Maximal computing power of candidate node
  The grid node maintaining CMC also takes part in blocks processing. Moreover, block processed in CMC node and results of them are not sent through the network. Locating CMC in the node of maximal computing power allows to minimize the data transfer in the network. To valuate quality of node, according to this criterion, we use the value of computational capacity e_m.
- Maximal capacity of node
  Since all blocks (except those processed by CMC node) must be send to grid nodes and results must be send back, links adjacent to CMC node must ensure enough capacity.
  Let P_m be the sum of capacities of all channels adjacent to node m:

$$P_m = \sum_{x_j^i \in X_r} x_j^i c_j^i p_m(i) \tag{9}$$

  where

$$p_m(i) = \begin{cases} 1, & \text{if } i\text{-th channel is adjacent to } m\text{-th node} \\ 0, & \text{otherwise} \end{cases}$$

  To valuate quality of node, according to this criterion, we use the value of P_m.
- Location of node
  In order to minimize the traffic in the whole network, it is beneficial to locate the CMC in the centre of the grid network.
  Let v_{gh} be the distance, in hops, between node g and node h. It means that v_{gh} is the minimal number of channels between nodes g and h. Let V_g be the distance between node g and all other nodes of distributed grid, defined as follows:

$$V_g = \sum_{h=1}^{n} v_{gh} \tag{10}$$

  Choosing the node for allocating CMC we should choose such nodes g, for which the value of V_g is minimal.

Another aim of an initial phase of an algorithm is task allocation to the computing nodes. We propose two strategies for initial task allocation:

- Regular tasks distribution approach.
  In this strategy, tasks are being allocated to the grid nodes in proportion to the computational power of each grid node.
- Regular traffic distribution approach.
  Task are being allocated to grid nodes in proportion to the capacity of node (the sum of capacities of all channels adjacent to the node). The constraint of the node computational power (4) must be satisfied.

Finally, capacity and flow (CFA) assignment problem for the initial phase may be solved using in order to satisfy requirement (6). We propose simple method, since optimal solution of CFA problem is not necessary on this phase. We start from maximal capacity for each channel of the network. Then, we minimize the capacities as long as (6) is satisfied.

B. Suboptimal solution

We start from the initial selection obtained in first phase. Then, in consecutive iterations we try to improve the solution by changing CMC allocation node, tasks allocation at grid nodes, routing and channel capacities. Algorithm finishes when there is no possibility of improving present solution.

To get the best choice we have to test all possible pairs of variables $y_h \in Y_r, y_m$ or $x_j^i \in X_r$, x_s^i or $z_m^l \in Z_r, z_h^l$ using a local optimization criterion. Because of the different nature of the variables denoting channel capacity choice, block assignment and the computing centre allocation, we have to formulate three different criteria – one for each of decision variables x_j^i, y_h and z_m^l.

Proposition 1. If the selection (X_t, Y_t) is obtained from the selection (X_r, Y_r) by complementing the variable $x_j^i \in X_r$ where $j < s(i)$ by another variable $x_s^i \in X_t$ then only the channel capacity change is being considered. We propose the following local optimization criterion on variables x_j^i where $j < s(i)$:

$$\Delta_{js}^i = \begin{cases} Q(X_r, Y_r, Z_r) - d_j^i + d_s^i & \text{for } c_s^i > f_i \\ \infty & \text{for } c_s^i \le f_i \end{cases}$$

Complementing of variables x_j^i means that value of total average delay in the network changes, that may affect restriction (6). Then we propose criterion for estimating of the total average delay after complementing:

$$\Delta T_{js}^i = T(X_r, Y_r, Z_r) + \frac{\sigma}{\gamma}\left(\frac{f_i}{c_s^i - f_i} - \frac{f_i}{c_j^i - f_i}\right) \text{ for } c_s^i > f$$

Proposition 2. If the selection (X_t, Y_t) is obtained from the selection (X_r, Y_r) by complementing variable $y_h \in Y_r$ by another variable $y_m \in Y_t$ then only allocation of CMC is being changed. Then, the traffic requirements between nodes change, channels' capacities do not change and blocks' allocation at computing nodes do not change. Then, to evaluate the pair (y_h, y_m) we propose following criterion:

$$\delta_{hm}^{CMC} = \begin{cases} Q(X_r, Y_r, Z_r) - u_h + u_m \\ \qquad \text{if } \tilde{f}_i < x_j^i c_j^i \text{ for } x_j^i \in X_r \\ \qquad \text{and } P_m \geq \sum_{l=1}^{t} (w_l + v_l) \\ \infty \qquad \text{otherwise} \end{cases}$$

Value of total average delay, obtained in result of complementing is estimated as follows [11]:

$$\Delta T_{hm}^{CMC} = \frac{\sigma}{\gamma} \sum_{x_j^i \in X_r} \frac{\tilde{f}_i}{x_j^i c_j^i - \tilde{f}_i} \quad \text{if } \tilde{f}_i < x_j^i c_j^i \text{ for } x_j^i \in X_r$$

where

$$\tilde{f}_i = \sum_{h=1}^{n} \sum_{\pi_{hm}^a \in \Pi_{hm}} V_{hm}^a(i) \bar{f}_{hm}^a$$

$$\bar{f}_{hm}^a = \frac{m_{hm}^a}{m_{hm}} r_{mk}, \quad m_{ej} = \sum_{\pi_{ej}^a \in \Pi_{ej}} m_{ej}^a$$

$$m_{hm}^a = \min_{i \in \pi_{hm}^a} \left(x_j^i c_j^i \right) \text{ for } x_j^i \in X_r$$

Π_{hm}^a denotes the a-th path from node h to node m,

$$V_{hm}^a(i) = \begin{cases} 1 \text{ if } i\text{-th channel belong to path } \Pi_{hm}^a \\ 0 \text{ otherwise} \end{cases}$$

Π_{hm} denotes the set of all paths from node h to node m.

$\tilde{f}$ is the 'new' traffic flow in the network, after reallocating the CMC. After reallocation all routes in the network must be redirected from paths [*old CMC node; other nodes*] to paths [*new CMC node; other nodes*]. It is simply calculated as follows. For each node, we find all possible routes (paths) from that node to new CMC node. They are denoted by the set Π_{hm}. Then we allocate the traffic along all found routes, proportionally to they residual capacities.

Proposition 3. If the selection (X_t, Y_t) is obtained from the selection (X_r, Y_r) by complementing variable $z_m^l \in Z_r$ by another variable $z_h^l \in Z_t$ then the allocation of l-th block is being changed. So, the traffic requirements between nodes change, channels' capacities and CMC location do not change. We propose following criterion for evaluating the pair $\left(z_m^l, z_h^l\right)$:

$$\delta_{hm}^{l} = \begin{cases} Q(X_r, Y_r, Z_r) - q_m^l + q_h^l \\ \qquad \text{if } \tilde{f}_i < x_j^i c_j^i \text{ for } x_j^i \in X_r \\ \qquad \text{and } \sum_{l=1}^{t} z_m^l p_l < e_m \\ \infty \qquad \text{otherwise} \end{cases}$$

Value of total average delay, obtained in result of complementing [12]:

$$\Delta T_{hm}^{l} = \frac{\sigma}{\gamma} \sum_{x_j^i \in X_r} \frac{\tilde{f}_i}{x_j^i c_j^i - \tilde{f}_i} \quad \text{if } \tilde{f}_i < x_j^i c_j^i \text{ for } x_j^i \in X_r,$$

where

$$\tilde{f}_i = f_i - f_{il}'' + \sum_{e=1}^{n} \sum_{\pi_{hm}^a \in \Pi_{hm}} V_{hm}^a(i) \bar{f}_{hm}^a$$

$$\bar{f}_{hm}^a = \frac{m_{hm}^a}{m_{hm}} r_{mk}, \quad m_{ej} = \sum_{\pi_{ej}^a \in \Pi_{ej}} m_{ej}^a$$

$$m_{hm}^a = \min_{i \in \pi_{hm}^a} \left(x_j^i c_j^i - f_i + f_{il}'' \right) \text{ for } x_j^i \in X_r$$

f_{il}'' is the part of the flow at i-th channel. It corresponds only to the packets connected with l-th block. Π_{hm}^a, Π_{hm} and $V_{hm}^a(i)$ are defined like previously.

Replacements of decision variables are made in order to obtain the distributed computing network with the possible least value of criterion function *OBJ*. We should choose such pairs $y_h \in Y_r, y_m$ or $x_j^i \in X_r$, x_s^i or $z_m^l \in Z_r, z_h^l$, for which the value of the criterion δ_{hm}^{CMC}, Δ_{js}^i or δ_{hm}^l is minimal and increase of value of average total delay $T(X_r, Y_r, Z_r)$ is minimal.

C. Calculation Scheme

<u>Initial Phase</u>

- <u>Step 1.1.</u> Choose node for computing management centre. Evaluate nodes using one of criteria defined in subsection A. Choose node with maximal computing power e_m, maximal capacity of node P_m or the node with the best location according to criterion (10). Also combined criteria may be used, in example e_m / V_m or $(e_m P_m) / V_m$.
- <u>Step 1.2.</u> Allocate tasks to the computing nodes, according to regular tasks distribution strategy or regular traffic distribution strategy.
- <u>Step 1.3.</u> Assign maximal possible capacity for each channel. Solve FA problem [8]. Calculate total average delay in the network. If $T(X_r, Y_r, Z_r) \leq T^{\max}$ then decide that problem (5)-(8) has no solution - algorithm finishes.
- <u>Step 1.4.</u> In consecutive steps, find the less utilized channel and decrease its capacity. Where there in no possibility for capacity reduction without violating requirement (6), then calculate value of objective function and remember it as $OBJ_{\min}$. Remember actual sets of decision variables as $(X_{\min}, Y_{\min}, Z_{\min})$.
  Go to optimization phase .

Optimization Phase

- Step 2.1. Perform $r = 0$. $(X_r, Y_r, Z_r) = (X_{\min}, Y_{\min}, Z_{\min})$.
- Step 2.2. Perform $r = r + 1$. Choose pair $y_h \in Y_{r-1}, y_m$ or $x_j^i \in X_{r-1}$, x_s^i or $z_m^l \in Z_{r-1}, z_h^l$, for which the value of the criterion δ_{hm}^{CMC}, Δ_{js}^i or δ_{hm}^l is minimal and, respectively, $\Delta T_{hm}^{CMC} \le T^{\max}$, $\Delta T_{js}^i \le T^{\max}$ or $\Delta T_{hm}^l \le T^{\max}$. If there is no such pair for which δ_{hm}^{CMC}, Δ_{js}^i or δ_{hm}^l is less than $OBJ_{\min}$ then stop, $(X_{\min}, Y_{\min}, Z_{\min})$ is sub-optimal solution of problem (5)-(8). Otherwise swap values of variables of chosen pair.
- Step 2.3. Solve the flow assignment problem in WAN, where traffic requirements are given by CMC allocation and blocks allocation at nodes and channels' capacities are given by set X_r. Calculate $OBJ_r(X_r, Y_r, Z_r)$. If $OBJ_r < OBJ_{\min}$ (better solution is found), then assign $(X_r, Y_r, Z_r) = (X_{\min}, Y_{\min}, Z_{\min})$, and $OBJ_{\min} = OBJ(X_{\min}, Y_{\min}, Z_{\min})$. Go to step 2.2.

D. Experiments and Analisis

Experiments conducted with proposed algorithm will validate the quality of approximate solutions and computational properties of an algorithm. Since the presented problem is NP-complete, there are no effective algorithms for finding optimal solutions - a brute force method may be used for small size problems but it takes exponential time. The proposed algorithm allows to find suboptimal solution in linear time.

Important parameter is the quality of solutions calculated by an algorithm. It may be measured as the distance between optimal and suboptimal solution. Precise distance may be calculated for small size problem, when optimal solution is find with the brute force method.

VI. CONCLUSION

In the paper, the WAN-based grid optimization problem with combined cost function was formulated and an approximate algorithm was proposed. The considered problem is far more general than the similar problems presented in the literature, because network optimization (capacity and flow routes) is carried out simultaneously with block assignment and management centre allocation. Algorithms proposed so far for grid networks in the literature do not take into consideration network optimization problems, then performance of grid may be affected by network overloads.

Considering two different kinds of cost in criterion function is very important from practical point of view, since computing cost and supporting cost of the network are significant and carried regularly. Computing cost may not be the money but also computational power, CPU utilization or other. In some application computing cost may be more important and less important in other. With dual-cost criterion algorithm may be better fitted to different user demands – we may define the importance of each cost using proper parameter in criterion function. Since considered problem is NP-complete, the big advantage of proposed approximate solution in short computing time needed to find solution, even for grid composed of hundreds of computing nodes.

REFERENCES

[1] D. Meisner, B.T. Gold, and T.F. Wenisch, "PowerNap: Eliminating Server Idle Power", Proc. The 14th international Conference on Architectural Support for Programming Languages and Operating Systems, ACM, New York, 2009, pp. 205-216

[2] H. Attiya and J. Welch, Distributed Computing: Fundamentals, Simulations, and Advanced Topics, 2 ed., Wiley-Interscience, 2004

[3] Y. Itami, T. Ishigooka, and T. Yokoyama, "A Distributed Computing Environment for Embedded Control Systems with Time-Triggered and Event-Triggered Processing", Proc. The 14th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications, 2008, pp. 45–54.

[4] R. Sterritt and D. Bustard, "Towards an Autonomic Computing Environment", Proceedings of the 14th International Workshop on Database and Expert Systems Applications, 2003, pp. 699 – 703.

[5] F. Travostino, Grid Networks, J. Wiley & Sons, 2006.

[6] S. Demeyer, M. De Leenheer, J. Baert, M. Pickavet, and P. Demeester, "Ant colony optimization for the routing of jobs in optical grid networks", Journal Of Optical Networking, vol. 7, no. 2, pp. 160-172, 2008.

[7] M. Baker, R. Buyya, and D. Laforenza, Grids and Grid Technologies for Wide-Area Distributed Computing, Software — Practice and Experience, Hoboken, NJ: Wiley, 2002.

[8] L. Fratta, M. Gerla, and L. Kleinrock, "The Flow Deviation Method: An Approach to Store-and-Forward Communication Network Design", Networks, Vol. 3, 1973, pp. 97–133.

[9] M. P. Clark, Data networks, IP and the Internet: protocols, design and operation, John Wiley & Sons, 2003.

[10] M. Pioro and D. Medhi, "Routing, Flow, and Capacity Design in Communication and Computer Networks", Elsevier, Morgan Kaufmann Publishers, San Francisco, 2004.

[11] M. Markowski, "Resource and Task Allocation Algorithm for WAN-based Distributed Computing Environment", International Journal of Electronics and Telecommunications, Vol. 56, No 2, 2010, pp. 197-202

[12] A Kasprzak, Topological Design of the Wide Area Networks, Wroclaw University of Technology Press, Wroclaw, 2001.

[13] M. Markowski and A. Kasprzak, "The web replica allocation and topology assignment problem in wide area networks: algorithms and computational results", Lecture Notes in Computer Science, 3483, pp. 772-781, 2005.

[14] M. Markowski and A. Kasprzak, "The Three-Criteria Servers Replication and Topology Assignment Problem in Wide Area Networks", Lecture Notes in Computer Science, 3982, pp. 1119-1128, 2006.

[15] M. Markowski and A. Kasprzak, "An approximate algorithm for replica allocation problem in wide area networks", Proc. 3rd Polish-German Teletraffic Symp. PGTS 2004, VDE Verlag, Berlin, pp. 161-166, 2004.