Paper

New Developments in a Two-criteria Approach to Dynamic Power Management in Energy-aware Computer Networks

Andrzej Karbowski[1,2] and Przemysław Jaskóła[2]

[1] *Institute of Control and Computation Engineering, Warsaw University of Technology, Warsaw, Poland*
[2] *Research and Academic Computer Network (NASK), Warsaw, Poland*

Abstract—In the paper authors continue the development of a model of dynamic power management in energy-aware computer networks, where two criteria: energy consumption and the quality of service are considered. This approach is appropriate when the routing problem with fixed demands is inadmissible. The formulation introducing edge indices is modified and tests on problems of different sizes are performed.

Keywords—data intensive computing, energy-aware network, energy-aware routing, dynamic power management, MIQP problems, traffic engineering.

1. Introduction

The methods for increasing energy efficiency of computer networks gained much attention last years. The reason is, that we are witnessing a rise of energy costs, customer increase, more on-demand services using cloud architectures, mobile Internet, a diffusion of broadband access and a growing number of services offered by Internet service providers. The growth of the energy consumption by network infrastructure may be well illustrated by the overall energy requirements of European Internet operators. In 2005 they needed 14 TWh, in 2010 – 21 TWh, and the forecast for year 2020 is 36 TWh [1].

At the same time the capacity surplus becomes a standard in almost all networks. Consequently, so-called green network technologies are quickly becoming a high-priority issue for the Internet [1], [2].

Efforts to reduce power consumption in telecommunication networks follow in two mutually related directions – design of a more efficient equipment and development of energy-aware network control strategies and protocols. Initial efforts were aimed at assessment of energy characteristics of network equipment and building elementary models [3]. However, it is possible to save even more energy by employing network-wide solutions.

The authors' earlier paper [4] presented a model of energy-aware router, an architecture of a control framework and two-criteria formulation of a network-wide energy saving optimization problem. A broad review of literature is presented there, hence it won't be repeated in this paper.

A two-criteria optimal routing and bandwidth allocation problem, taking into account the energy component, for a completely different network and cost model was recently presented in [5].

The objective was the minimization of the weighted sum of two components: total power utilized by network components and end-to-end Quality of Service (QoS) expressed by a quadratic utility function. In this model some parts of the network can be shifted to low energy mode as a result of the optimization algorithms, where both paths and flow rates are decision variables. It exploits the fact, that Internet traffic used to be elastic in a large part, which means, that a quality of service is little aggravated by small deviations from assumed flow rate.

In this paper the model mentioned above is presented first in a modified version, as regards to the edges. In the authors' opinion the new version is more convenient, because there is no necessity to remember the parities of links labels. Then the results of tests of this optimization model on network problems of different sizes are presented.

2. A Two-criteria Routing Problem

A hierarchical network model proposed in [4] and basically adapted from [6], [7], considers every single communication port $p \in \{1,\ldots,P\}$ of every line card of a router $r \in \{1,\ldots,R\}$ connected to an edge $e \in \{1,\ldots,E\}$. We do not consider individual cards of the router, as it is in [6], [7], because they do not bring anything into the model except additional summations.

Directed links connecting pairs of ports by an edge are denoted by $l \in \{1,\ldots,L\}$. Any network component can operate in $k \in \{1,\ldots,K\}$ energy states, but two ports connected by an edge are in the same state. A demand $d \in \{1,\ldots,D\}$ is characterized by its source s_d, the destination t_d node (router), the maximum volume V_d and the actual flow rate $v_d \in [0,V_d]$.

The topology of the physical network is described by four matrices of binary indicators: g_{pr}, a_{lp}, b_{lp}, which indicate, whether, respectively: port p belongs to the router r, link l is incoming to the port p and link l is outgoing from the port p. If l is a link outgoing from the port p, the link $\tilde{l}$ denotes its partner link in the edge going in the opposite direction, that is:

$$b_{lp} = 1 \Longleftrightarrow a_{\tilde{l}p} = 1\,. \tag{1}$$

Let us introduce a vector valued function ζ, which for a given edge e determines two links forming it, in the increasing order, that is

$$\zeta(e) = \begin{bmatrix} \zeta_1(e) \\ \zeta_2(e) \end{bmatrix} = \begin{bmatrix} l \\ \tilde{l} \end{bmatrix},\; l < \tilde{l}. \tag{2}$$

The decision variables are two vectors of binary indicators x_p, z_r – whether the port p or router r is used for data transmission and two incidence matrices with elements: y_{ek} – whether the edge e is in the state k and u_{dl} – whether the demand d uses the link l and flow rates v_d.

A two criteria – i.e. reflecting energy costs F_{LNb} and QoS – mixed integer network problem of simultaneous optimal bandwidth allocation and routing may be formulated in the following way:

$$\begin{aligned} &\min_{\substack{x_p, y_{ek}, z_r, u_{dl}, v_d,\\ p\in\overline{1,P}, e\in\overline{1,E}, k\in\overline{1,K}\\ r\in\overline{1,R}, d\in\overline{1,D}, l\in\overline{1,L}}} \Bigg\{ F_{2C} = \alpha F_{LNb} + (1-\alpha)\sum_{d=1}^{D} Q_d(v_d) = \\ &= \alpha \left[\sum_{e=1}^{E}\sum_{k=1}^{K} \xi_{ek} y_{ek} + \sum_{p=1}^{P} W_p x_p + \sum_{r=1}^{R} T_r z_r \right] + \\ &\quad + (1-\alpha)\sum_{d=1}^{D} Q_d(v_d) \Bigg\}\,, \end{aligned} \tag{3}$$

subject to the constraints:

$$\forall_{\substack{d=1,\ldots,D,\\ p=1,\ldots,P}} \quad \sum_{l=1}^{L} a_{lp} u_{dl} \le x_p\,, \tag{4}$$

$$\forall_{\substack{d=1,\ldots,D,\\ p=1,\ldots,P}} \quad \sum_{l=1}^{L} b_{lp} u_{dl} \le x_p\,, \tag{5}$$

$$\forall_{\substack{r=1,\ldots,R,\\ p=1,\ldots,P}} \quad g_{pr} x_p \le z_r\,, \tag{6}$$

$$\forall_{e=1,\ldots,E} \quad \sum_{k=1}^{K} y_{ek} \le 1 \tag{7}$$

$$\begin{aligned} \forall_{\substack{d=1,\ldots,D,\\ r=1,\ldots,R}} \sum_{p=1}^{P}\sum_{l=1}^{L} g_{pr} a_{lp} u_{dl} - \sum_{p=1}^{P}\sum_{l=1}^{L} g_{pr} b_{lp} u_{dl} &= \\ = \begin{cases} -1 & r = s_d \\ 1 & r = t_d \\ 0 & \text{otherwise} \end{cases}&, \end{aligned} \tag{8}$$

$$\sum_{d=1}^{D} v_d u_{d\zeta_1(e)} \le \sum_{k=1}^{K} M_{ek} y_{ek},\; e = 1,\ldots,E\,, \tag{9}$$

$$\sum_{d=1}^{D} v_d u_{d\zeta_2(e)} \le \sum_{k=1}^{K} M_{ek} y_{ek},\; e = 1,\ldots,E\,, \tag{10}$$

$$x_p, z_r \in \{0,1\}\; p \in \overline{1,P}, r \in \overline{1,R}\,, \tag{11}$$

$$y_{ek}, u_{dl} \in \{0,1\}\; e \in \overline{1,E}, k \in \overline{1,K}, d \in \overline{1,D}, l \in \overline{1,L}\,, \tag{12}$$

$$0 \le v_d \le V_d,\; d \in \overline{1,D}\,, \tag{13}$$

where M_{ek} and ξ_{ek} are, respectively, the capacity and the power consumption of the edge e in the state k, and W_p and T_r are power cost coefficients of the port p and the router r. F_{LNb} is the total power consumption by network devices component of the objective function, Q_d is a QoS related component. The latter represents a penalty for not achieving the assumed flow rate V_d by the flow d. $Q_d(v_d)$ is a convex and continuous function, decreasing on interval $[0, V_d]$. It is reaching minimum (zero) at V_d, the point in which user expectations are fully satisfied. The convexity of $Q_d(v_d)$ is associated with the conviction, that small deviations from the nominal throughput $\Delta = V_d - v_d$ are neglected by network users, while large deviations are noticed and should be avoided. Moreover, since $Q_d(v_d)$ is monotonically decreasing, it assures that the slope of the curve becomes steeper, as the rate v_d approaches zero. Constraints (4)–(6) determine the number of ports and routers that are used for data transmission. The conditions (7) assure, that each edge can be in one energy-aware state. The constraints (8) are formulated according to 1st Kirchhoff's law applied to source, destination and transit nodes, and finally, the constraints (9), (10) assure, that the flow will not exceed the capacity of a given edge. They are expressed in a more natural and convenient in implementation way than in the paper [4], where a trick based on parities of link labels was used. Now they concern edges instead of links and it is not necessary for links l and $\tilde{l}$ of the same edge be of different parities.

The above model exploits the fact, that Internet traffic used to be elastic in a large part, which means, that the QoS is only a little aggravated by small deviations from the assumed flow rate. The combined routing and rate control problem has to be solved, which leads to the solution feasible in terms of the formulated model, even when the traffic demand is greater than the capacity offered by the network. Moreover, in some cases a minor reduction of flow rates, which is accepted by the comprehensive model taking into account the elasticity of a demand, may allow to accommodate the traffic in a smaller number of links, thus allowing for further great reduction of power consumption.

The parameter $\alpha \in [0,1]$ is a scalarizing weight coefficient, which can be altered to emphasize any of the objectives.

In general, the formulation (3)–(13) has some drawbacks: it defines a mixed-integer nonlinear programming problem with nonconvex, bilinear link capacity constraints (9),(10). At present the leading solvers – e.g., CPLEX, Gurobi – can solve efficiently convex quadratic mixed-integer quadratically constrained problems MIQCP, with positive semidefinite matrices of constraints quadratic forms, which is not the case of (9), (10) constraints. The general nonlinear, mixed-integer, non-convex solvers are very slow.

Fortunately, the problem (3)–(13) can be quite easily transformed to the form accepted by fast mixed-integer solvers, what is described in Section 3.

3. Elimination of the Nonlinearity from Constraints

From the QoS components of the objective function F_{2C} (3) it is usually expected, that they assure so-called proportional-fairness of the allocations of the bandwidth, when the network is subject to a congestion [8]. Quadratic functions may be used to achieve it [9] (unfortunately, linear – not), so the objective function F_{2C} can be quadratic and convex.

The only problem that still remains to solve is nonconvex nonlinearity of the constraints (9), (10). It can be eliminated by a transformation found in [10].

It consists in the introduction of auxiliary variables $w_{dl} = v_d u_{dl}$, $d \in \overline{1,D}$, $l \in \overline{1,L}$ (denoting the part of a traffic rate in the link l assigned to the flow d) and the substitution of these inequalities with subsequent set of linear inequalities:

$$\forall_{e=1,\dots,E} \quad \sum_{d=1}^{D} w_{d\zeta_1(e)} \le \sum_{k=1}^{K} M_{ek} y_{ek}\,, \tag{14}$$

$$\forall_{e=1,\dots,E} \quad \sum_{d=1}^{D} w_{d\zeta_2(e)} \le \sum_{k=1}^{K} M_{ek} y_{ek}\,, \tag{15}$$

$$\forall_{\substack{d=1,\dots,D,\\ l=1,\dots,L}} \quad w_{dl} \le V_d u_{dl}\,, \tag{16}$$

$$\forall_{\substack{d=1,\dots,D,\\ l=1,\dots,L}} \quad w_{dl} \le v_d\,, \tag{17}$$

$$\forall_{\substack{d=1,\dots,D,\\ l=1,\dots,L}} \quad w_{dl} \ge v_d - V_d(1-u_{dl})\,, \tag{18}$$

$$\forall_{\substack{d=1,\dots,D,\\ l=1,\dots,L}} \quad w_{dl} \ge 0\,. \tag{19}$$

4. The Final Formulation of the Problem

Summing up, the final formulation of presented two criteria energy-aware integrated routing and flow control problem is as follows:

$$\min_{\substack{x_p, y_{ek}, z_r, u_{dl}, v_d,\\ p\in\overline{1,P}, e\in\overline{1,E}, k\in\overline{1,K}\\ r\in\overline{1,R}, d\in\overline{1,D}, l\in\overline{1,L}}} \left\{ F_{2C} = \alpha F_{LNb} + (1-\alpha)\sum_{d=1}^{D} Q_d(v_d) = \right.$$

$$= \alpha \left[\sum_{e=1}^{E} \sum_{k=1}^{K} \xi_{ek} y_{ek} + \sum_{p=1}^{P} W_p x_p + \sum_{r=1}^{R} T_r z_r \right] +$$

$$\left. +(1-\alpha)\sum_{d=1}^{D} Q_d(v_d) \right\}, \tag{20}$$

subject to the constraints:

$$\forall_{\substack{d=1,\dots,D,\\ p=1,\dots,P}} \quad \sum_{l=1}^{L} a_{lp} u_{dl} \le x_p\,, \tag{21}$$

$$\forall_{\substack{d=1,\dots,D,\\ p=1,\dots,P}} \quad \sum_{l=1}^{L} b_{lp} u_{dl} \le x_p\,, \tag{22}$$

$$\forall_{\substack{r=1,\dots,R,\\ p=1,\dots,P}} \quad g_{pr} x_p \le z_r\,, \tag{23}$$

$$\forall_{e=1,\dots,E} \quad \sum_{k=1}^{K} y_{ek} \le 1\,, \tag{24}$$

$$\forall_{\substack{d=1,\dots,D,\\ r=1,\dots,R}} \sum_{p=1}^{P}\sum_{l=1}^{L} g_{pr} a_{lp} u_{dl} - \sum_{p=1}^{P}\sum_{l=1}^{L} g_{pr} b_{lp} u_{dl} = $$

$$= \begin{cases} -1 & r = s_d \\ 1 & r = t_d \\ 0 & \text{otherwise} \end{cases}, \tag{25}$$

$$\forall_{e=1,\dots,E} \quad \sum_{d=1}^{D} w_{d\zeta_1(e)} \le \sum_{k=1}^{K} M_{ek} y_{ek}\,, \tag{26}$$

$$\forall_{e=1,\dots,E} \quad \sum_{d=1}^{D} w_{d\zeta_2(e)} \le \sum_{k=1}^{K} M_{ek} y_{ek}\,, \tag{27}$$

$$\forall_{\substack{d=1,\dots,D,\\ l=1,\dots,L}} \quad w_{dl} \le V_d u_{dl}\,, \tag{28}$$

$$\forall_{\substack{d=1,\dots,D,\\ l=1,\dots,L}} \quad w_{dl} \le v_d\,, \tag{29}$$

$$\forall_{\substack{d=1,\dots,D,\\ l=1,\dots,L}} \quad w_{dl} \ge v_d - V_d(1-u_{dl})\,, \tag{30}$$

$$\forall_{\substack{d=1,\dots,D,\\ l=1,\dots,L}} \quad w_{dl} \ge 0\,, \tag{31}$$

$$x_p, z_r \in \{0,1\} \;\; p \in \overline{1,P}, r \in \overline{1,R}, \tag{32}$$

$$y_{ek}, u_{dl} \in \{0,1\} \;\; e \in \overline{1,E}, k \in \overline{1,K}, d \in \overline{1,D}, l \in \overline{1,L}, \tag{33}$$

$$0 \le v_d \le V_d,\; d \in \overline{1,D}. \tag{34}$$

When QoS components $Q_d(v_d)$ are quadratic and convex, the obtained mixed-integer quadratic problem can be solved by effective MILP/MIQP solvers, such as CPLEX, Gurobi.

5. Numerical Evaluation

The problem (20)-(34) was formulated, implemented and solved by means of the CPLEX solver for the test network. The topology of an example network is shown in Fig. 1. It has been inspired by an access/metropolitan segment of a telecom operator network, which was presented in [11]. Its size is however reduced compared with the original.

The access nodes, represented by circles indexed 1–3, are sources and destinations of traffic flows. Transit nodes (5–8) perform traffic switching, and a peering node, labeled T, provides access to the ISP transport network and the Internet.

In the presented example the number of routers $R = 9$, the number of edges $E = 14$ and the number of links and ports $L = P = 28$.

As the (penalty for not achieving) QoS functions we took:

$$Q_d(v_d) = \frac{1}{2}\left(v_d - V_d\right)^2. \tag{35}$$

It was possible for each edge to operate in $K = 5$ energy-aware states. The throughput of a given edge $e \in \overline{1,E}$ and the power consumption in energy-aware state $k \in \overline{1,K}$ were

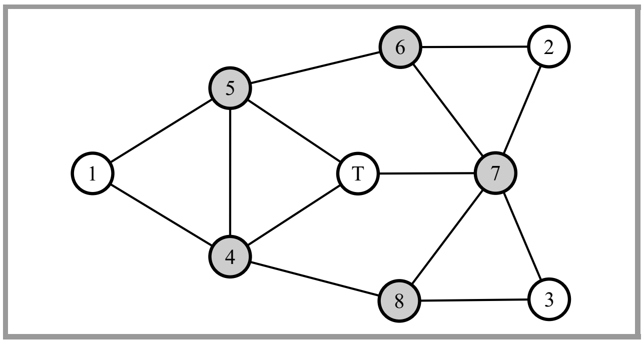

Fig. 1. Example network.

as follows: $M_{e1} = 2$, $\xi_{e1} = 20$; $M_{e2} = 4$, $\xi_{e2} = 40$; $M_{e3} = 6$, $\xi_{e3} = 60$; $M_{e4} = 8$, $\xi_{e4} = 80$; $M_{e5} = 10$, $\xi_{e5} = 100$. As the power cost coefficients we took for ports $W_1 = W_2 = \ldots = W_{28} = 90$ and for routers $T_1 \ldots T_3 = 1000$, $T_4 \ldots T_8 = 10000$, $T_9 = 3000$. The scalarizing coefficient was taken $\alpha = 10^{-3}$. This value was experimentally determined to ensure that from possible Pareto optimal solutions selected are these, which offer satisfactory QoS level in every run of the experiment.

In the first series of experiments a dependency between the number of binary variables and the time of computations have been examined. In subsequent experiments the number of traffic flows was increased, which resulted in a linear increase of the number of binary variables. The QoS function parameter V_d remained the same for every flow: $V_d = 10$. The computation time is presented in Fig. 2. It reveals a sharp increase in function of the problem dimension.

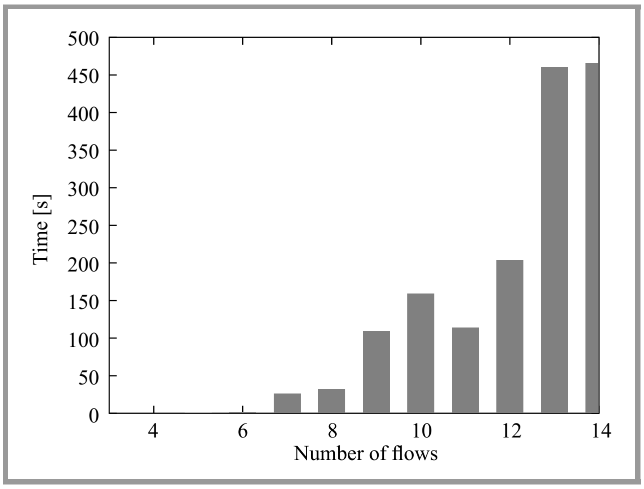

Fig. 2. Computation time as a function of number of traffic flows.

The applicability of this model depends on the chosen control scheme. In an on-line control the maximum acceptable computation time is closely related to the time between reconfigurations, which should be in the order of tens of minutes (see e.g. [11]). The observed trend indicates, that for bigger examples this limit can be exceeded.

The second series of experiments shows that the calculation time depends not only on the number of variables, but also on values of the parameters of the problem. In this case the number of flows was fixed to $D = 10$, while the parameter V_d was altered, which is presented in Fig. 3. Although the dimension of the problem has not changed, the computation time varies to the large extent between experiments.

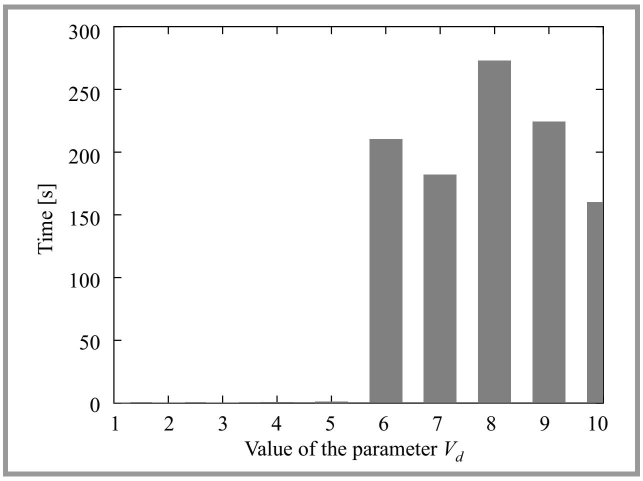

Fig. 3. Computation time as a function of a the parameter V_d.

All the results were obtained on PC with Intel Core i5-2430M CPU running at 2.4 GHz.

6. Conclusions

The authors modified the dynamic power management of energy-aware computer networks model presented in [4] to simplify it and make its implementation easier. It is still suggested to use a two-criteria model with flow rates as decision variables. When the QoS functions are quadratic and convex, it is possible to reformulate the problem in such a way, that the same standard solvers, e.g., CPLEX or Gurobi, can be used to find the solution. The resulting mixed-integer programming problem has more variables, but the additional ones are only real, not binary, what should not influence too much the time of calculations.

The performed numerical tests confirmed the appropriateness of the formulation for practical problems of small dimension. The computation times are however sensitive not only to the dimension of the problem, but also to its parameters, which indicates a need to develop effective heuristic solvers, which are more robust.

References

[1] D. G. Recupero, "Toward a Green Internet", *Science*, vol. 339, no. 6127, pp. 1533–1534, 2013 (doi: 10.1126/science.1235623).

[2] R. Bolla, R. Bruschi, F. Cucchietti, and F. Davoli, "Setting the Course for a Green Internet", *Science*, vol. 342, no. 6164, pp. 1316, 2013 (doi: 10.1126/science.342.6164.1316-a).

[3] R. Bolla, R. Bruschi, F. Davoli, and F. Cucchietti, "Energy efficiency in the Future Internet: A survey of existing approaches and trends in energy-aware fixed network infrastructures", *IEEE Commun. Surv. & Tutor.*, vol. 13, no. 2, pp. 223–244, 2011 (doi: 10.1109/SURV.2011.071410.00073).

[4] A. Karbowski and P. Jaskóła, “Two approaches to dynamic power management in energy-aware computer networks – methodological considerations”, *Proc. Federated Conf. Comp. Sci. & Inform. Sys. FedCSIS 2015* (One of the Best Papers Award), Łódź, Poland, 2015, pp. 1177–1182 (doi: 10.15439/2015F228).

[5] P. Jaskóła, P. Arabas, and A. Karbowski, “Simultaneous routing and flow rate optimization in energy-aware computer networks”, *Int. J. Appl. Mathem. & Comp. Sci.*, vol. 26, no. 1, pp. 231–243, 2016.

[6] E. Niewiadomska-Szynkiewicz, A. Sikora, P. Arabas, M. Kamola, K. Malinowski, P. Jaskóła, and M. Marks, “Network-wide power management in computer networks”, in *Proc. 22nd Int. Teletraff. Congr. Special. Seminar on Energy Efficient & Green Netw. SSEEGN 2013* (Best Paper Award), Christchurch, New Zealand, 2013, pp. 25–30.

[7] E. Niewiadomska-Szynkiewicz, A. Sikora, P. Arabas, M. Kamola, M. Mincer, and J. Kołodziej, “Dynamic power management in energy-aware computer networks and data intensive computing systems”, *Future Gener. Comp. Sys.*, vol. 37, pp. 284–296, 2014.

[8] F. Kelly, “Charging and rate control for elastic traffic”, *Eur. Trans. Telecommun.*, vol. 8, no. 1, pp. 33–37, 1997 (doi: 10.1002/ett.4460080106).

[9] C. Touati, E. Altman, and J. Galtier, “On fairness in bandwidth allocation”, *Tech. Rep. 4269*, Unité de recherche INRIA, Sophia Antipolis, France, 2001.

[10] J. Bisschop, *AIMMS Optimization Modeling*. Paragon Decision Technology B.V., 2007.

[11] A. P. Bianzino, L. Chiaraviglio, M. Mellia, and J. L. Rougier, “GRiDA: GReen Distributed Algorithm for energy-efficient IP backbone networks”, *Comp. Netw.*, vol. 56, no. 14, pp. 3219–3232, 2012.

Andrzej Karbowski received his Ph.D. in 1990 and D.Sc. in 2012 in Automatic Control and Robotics from the Warsaw University of Technology, Faculty of Electronics and Information Technology. Currently he is an Associate Professor at the Institute of Control and Computation Engineering of the Warsaw University of Technology and at the Research and Academic Computer Network (NASK). He is the editor and the co-author of two books (on parallel and distributed computing), the author and the co-author of two e-books (on grid computing and the optimal control synthesis) and over 120 journal and conference papers. His research interests concentrate on data networks management, optimal control in risk conditions, water systems management, decomposition and parallel implementation of numerical algorithms.

E-mail: A.Karbowski@elka.pw.edu.pl
Institute of Control and Computation Engineering
Warsaw University of Technology
Nowowiejska st 15/19
00-665 Warsaw, Poland

Research and Academic Computer Network (NASK)
Wąwozowa st 18
02-796 Warsaw, Poland

Przemysław Jaskóła received his M.Sc. in Automatic Control and Robotics from the Warsaw University of Technology, Poland, in 1999. He works as a research associate at the Research and Academic Computer Network (NASK). His current research interests focus on parallel and distributed computing, modeling and multicriteria optimization of computer networks.

E-mail: pjaskola@nask.pl
Research and Academic Computer Network (NASK)
Wąwozowa st 18
02-796 Warsaw, Poland