The Development of Hybrid IP Architecture for Solving the Problems of Heating Networks (using pipeline-parallel data processing technology)

Alexander Baklanov, Olga Baklanova, Svetlana Grigoryeva, Saule Kumargazhanova, Indira Sagynganova, Yuriy Vais

D. Serikbayev East Kazakhstan State Technical University
Faculty of Information Technology and Power Engineering
A. K. Protazanov Str. 69, 070004, Ust-Kamenogorsk, Kazakhstan
e-mail: {ABaklanov, OBaklanova, SGrigorieva, Saule.Kumargazhanova, ISagynganova, YuVais}@ektu.kz

György Györök

Óbuda University, Alba Regia Technical Faculty
Budai út 45, H-8000 Székesfehérvár, Hungary
gyorok.gyorgy@amk.uni-obuda.hu

Abstract: The paper covers organisation of pipelined processing of data received from heat stations. New software architecture was developed for processing data received from the heating network. Development of a new analytical information system based on pipeline data processing has allowed increased efficiency of work in heat supply systems. The mechanisms for data storage have been established, as well as work sequence and monitoring of heating networks. This architecture is based on a parallel-pipeline data processing system. The idea of such an approach of data processing was transferred from the system of organization of work of the central processor of a personal computer with processes and streams. A distinctive feature of our system is the ability to work with different databases. It can be adapted to various modern systems of data storage and communication in heating systems. The timely operational monitoring described in the article made it possible to change the modes of operation of heat points and the central heat point in real time, which affects the increase in the reduction of heat consumption. The real reduction made was approximately 9 percent. Hopefully, this allows that with the implementation of this approach at the industrial level heat gain and, respectively, energy will increase even more.

Keywords: pipelined data processing; heat stations; heating networks; database; information system

1 Introduction

In 2011, Hungary declared a national energy strategy until 2030 [1]. It included several targets for district heating, such as the connection of public institutions to district heating and an increasing share of renewable energy and waste in heat production. The strategy forecasts the decrease of the share of district heat consumption within the residential and tertiary sector from 12% to 10% by 2020, due to renovation and insulation of buildings.

Despite the fact that the use of renewable energy is in great demand [2], natural gas remains the main fuel used in the energy sector in Hungary. It plays a very important role in electricity production, particularly in cogeneration, as well as in district heating, as 78% of district heating is produced by natural gas. The majority of residential areas are connected to the natural gas network. Out of 4.3 million dwellings, 3.3 million are connected to the natural gas supply and 2.7 million (63%) are heated by natural gas through central or individual heating. As seen in the graph below, the share of district heating in the residential sector is 12% (Figure 1).

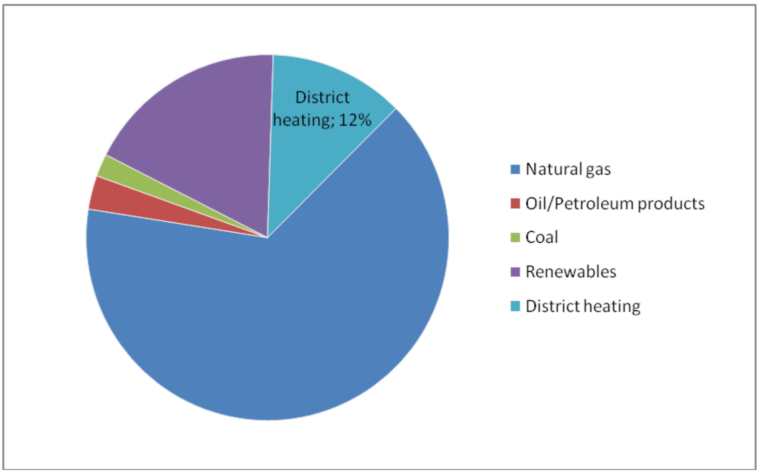


Figure 1

Share of energy source to satisfy heat demand in the residential sector

Before 2010, up to 60% of natural gas related to district heating was used in cogeneration, with the remaining part in heat-only boilers. Afterwards, CHP decreased by 35% until 2012 due to the end of the feed-in tariff in Hungary [3].

For the period 2014-2020, approximately 140 million EUR is to be spent to support the Environmental and Energy Efficiency Operational Programme for energy efficiency and renewable energy projects in district heating systems. According to the National Energy Strategy 2030, the share of energy use from renewables for residential and institutional heating will increase to 32% and the use of renewable energy will double by 2030. The best framework for renewable heat is district heating.

In the world's most developed countries the main part of the produced fuel and energy resources is spent on the production of electricity and heat, of low and medium potential. Therefore, an important task is to improve the schemes and equipment of energy-consuming plants of industrial enterprises. Each percentage of reduction in energy consumption (electricity and heat) in industry is currently equivalent to a national coal equivalent of fuel economy of about 4 million tons per year.

Heating is Europe's largest end-use of energy. This accounts for approximately 50% of total final energy consumption. District heating (DH) systems provide heating for a wide range of customers, from residential to agricultural, including commercial, public and industrial customers.

There are about 7000 DH systems in Europe, which are currently providing more than 10% of total European needs in heat energy with an annual turnover of 25-30 billion EUR (556 TW*h). Market penetration in district heating is distributed unevenly, in some countries it's nearing zero, yet in others it's up to 70%.

The advantages of centralised heating and centralised cooling are most obvious in regions with high energy requirements. In the European Union about 73% of the population lives in cities, expecting growth to about 80% by 2030. Currently, 69% of all primary energy requirements are concentrated in urban regions [4].

At present, substations for domestic buildings and offices are mostly homemade. By agreeing with coordination of a number of functions at heating substations the district heating sector could produce standard heating stations with significantly reduced homemade portions. By reaching agreement on central heating substations, the industry will be able to produce safer, more environmentally friendly, cost-competitive equipment.

For the organization of effective monitoring and management of modern heating systems, we have developed a conveyor - parallel processing of these heat points. The implementation of such technology was carried out with the help of an information-analytical system that takes into account the experience of software development described in papers [5]. A standard IP network was used to connect all the heating units to the central heating unit.

2 Organization of an Informational and Analytical System

Some of the priority problems are the development and implementation of control and management primarily in large thermal networks; improvement of auto-regulation and protection devices, development of methods and devices for determining the places of coolant leakage before opening the channel. An important task is to improve the schemes and equipment of industrial heat-consuming plants from the point of view of the most rational combination of technological and energy processes and optimization of energy consumption.

Currently, the processing of these heat points is carried out by surveying the heat points in real-time. Because the service is using the server [6], the performance of which is often unable to cope with the amount of information, the incoming data is not fully accountable and does not allow proper control of the heat supply.

Consequently, the use of modern management technologies of heat stations united into a single network will significantly save electricity energy and more accurately distribute heating in residential and industrial premises.

A heat station (HS) is a set of equipment located in a separate, or in the same room with consumers which includes the elements of thermal power plants and makes it possible to connect these installations to a heating system, to control heat consumption modes, to convert and control heating agent parameters as well as heating agent distribution by the type of consumption.

Heating equipment capacity, heat consumption control, the distribution of heating agent by consumption type (heating, hot water, ventilation and air conditioning) is performed through a heat station; the parameters of the heating agent are adjusted and changed.

Heat stations are mandatory both in residential, industrial premises and warehouse facilities. Maintenance of heat stations depends on their type.

Heat station functionality:

- heating agent parameters control and optimization;
- converting heating agent type;
- heating systems protection, reducing the risk of an emergency;
- heating agent distribution in heating systems, water supply and ventilation systems;
- control over the heating agent and heat consumption, as well as providing the necessary consumption of heating agent (this trait is impacted by heat loss, characteristics of the object in conjunction with the specified parameters);
- turning heating systems on and off;
- reduction of heat loss.

There are signs of transition, resulting from the reform of housing and utility services, to a payment system where an owner or a tenant of property will pay for actually consumed heat. Prices for heat, provided by heating enterprises, are currently estimated for the maximum possible consumption. To solve this problem, we need objective testing methods of the amount of heat consumed. Thus, we need a system combining the metering functionality of heat consumption and heat supply regulation, in order to save energy and, at the same time, ensure comfortable conditions in the premises [7]. Based on what is mentioned above, we can conclude that we need such data integration, which includes combining data from different sources and providing users with data in a uniform manner. This process is essential for commercial issues (when two similar companies need to merge their databases), as well as scientific ones (e.g.: combining research results from various bioinformation repositories). The role of data integration increases with the increase in the amount and need for data sharing.

In contrast to the common approaches [8], where data from different databases is converted into a single database, we propose parallel work with data from different databases [9], in order to solve the problems of heat network information resources integration.

The proposed integrated information system architecture is a hybrid model that combines elements of "client-server" architecture and using a computing cluster with parallel distributed heterogeneous information processing (Figure 2).

The information system architecture has four levels.

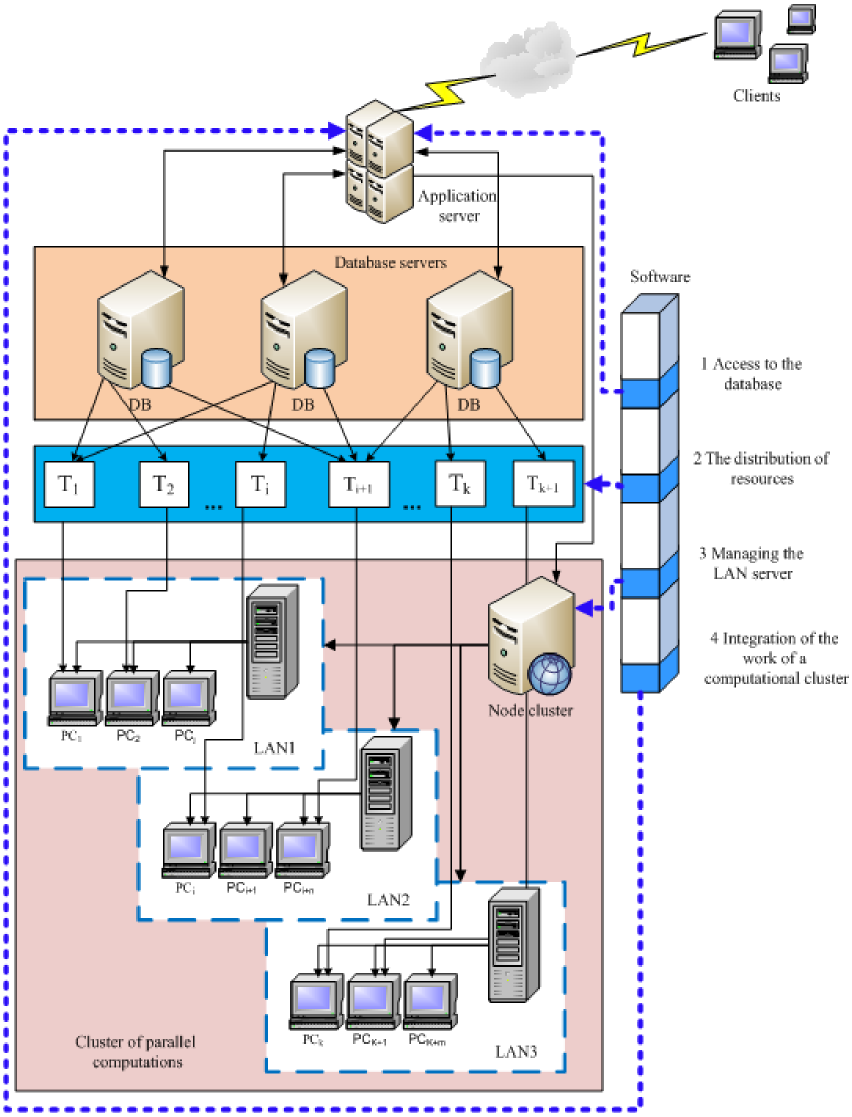
The first level is a "client-server" information system foundation, ensuring parallel operation management of a cluster of computers as well as interaction with the client (user) stations.

The second level is the level of work with the data. Key elements of this level are database servers that operate with databases having different data formats or different platforms based on different storage mechanisms such as relational databases and object-oriented databases.

The third level is the level of solved informational tasks, which specifies the algorithm and sorts out records used in the information system from the databases available at the second level.

The fourth level is the level of implementation of parallel data processing; it is based on a computer cluster which, in our case, allows parallel computing coordinated by the cluster node server which performs dispatch functions for distribution of tasks among the workstations in the cluster in accordance with the instructions of the application server of the information system.

In recent years the world experiences a rapid introduction of computational clusters – local networks, with nodes of workstations or personal computers specially collected to be used as a multiprocessor computing system (supercomputer).



LAN – local area network; DB – databases; Ti – The i-th task

Figure 2
Architecture of an integrated information system

World experience of computational cluster development represents a considerable number of examples from a modest 20-30 node cluster in academic or scientific laboratories at universities to giant computer systems consisting of 1000 – 2000 workstations created in the framework of special projects.

To build computational clusters the following are usually used: public computers based on Intel or AMD processors, a standard Ethernet network technology or the Fast Ethernet, open source Windows operating system [10] and the communication library MPI [11] which implements the connection between the branches of the parallel computing process. Thus, today computing clusters have become a public and relatively cheap alternative to traditional supercomputers. In many classes of tasks and with a sufficiently large number of nodes these cluster systems achieve performance comparable to supercomputer [8, 10].

A set of required software is determined by the objectives of the cluster: a stable multi-user and multi-tasking mode and support for parallel programming techniques.

Schematically, the interaction of software implementation of the information system with the hardware component is also shown in Figure 2.

Initially, information on IP-addresses of database servers, a cluster node, and local area network servers included in a parallel computing cluster is loaded in information system memory modules. These addresses are required to access the database and coordinate parallel operation of a computer cluster workstations. Access to databases is the next step; at that, the information system provides the possibility to use multiple technologies of interaction with databases, namely, ODBC, ADONet and Microsoft Jet, which allows working with relational databases. To work with object-oriented databases (OODB) the information system program code contains a module of access code to databases with a specially designed class, its objects are records in the used OODB.

The module of the information system associated with database interaction is shown in Figure 3.

The module links to the NET Framework software platform. NET Framework provides for a variety of ways to operate databases. The .NET Framework platform has its own technology for data access - ADO.NET (ActiveX Data Object for .NET). ADO.NET includes managed classes allowing .NET application to connect to databases, as well as operate data and control standalone data. ADO.NET technology allows to operate data with Microsoft SQL Server, Microsoft Access, Microsoft Excel, Microsoft Outlook, Microsoft Exchange, Oracle, OLE DB, ODBC, XML in a standalone mode with DataSet objects [12]. DataSet objects allow to extract copies of interconnected local data tables from MS Access. Afterwards, the module operates on DataSet contents, without a need for an active connection to the data source, while also allowing to send modified data back for processing with a corresponding data adapter.

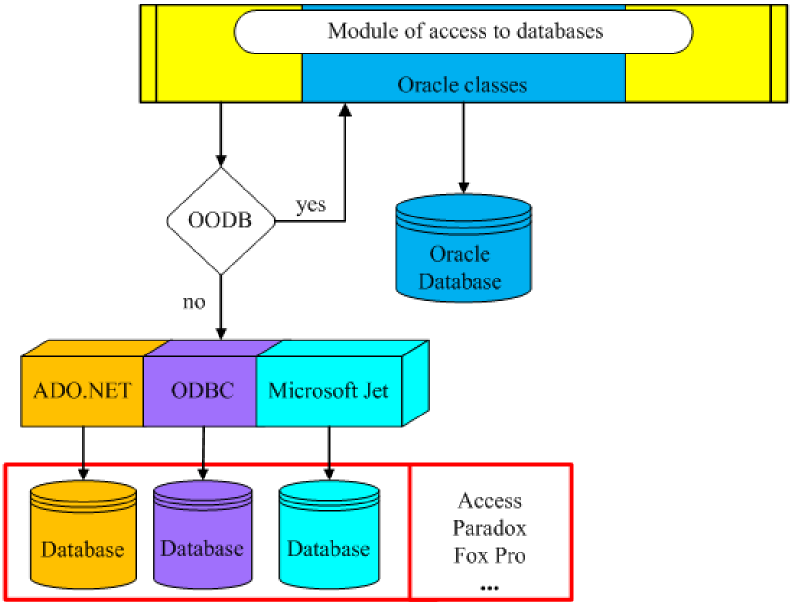


Figure 3
Scheme of the module of interaction with databases

The outline of interaction with the database through ADO.NET is provided in Figure 4.

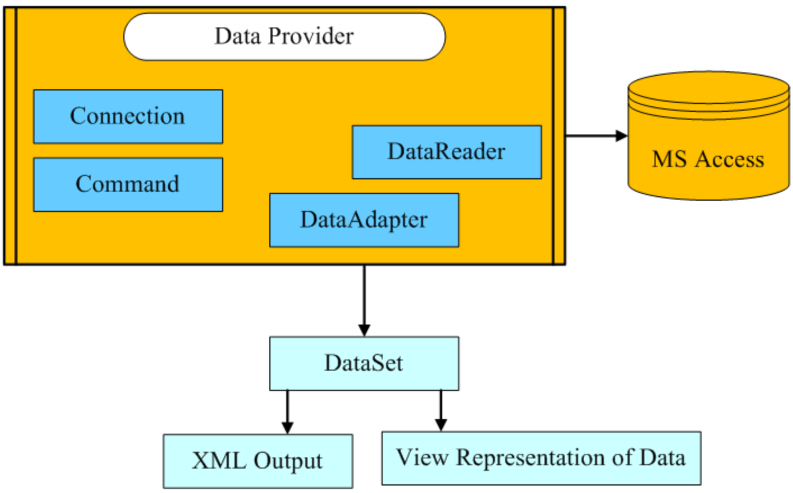


Figure 4
The outline of interaction with the database through ADO.NET

For database operation, including working with Paradox, the module implements an open interface for database access ODBC (Open DataBase Connectivity) [13]. ODBC allows our module to interact with various databases with no need to worry about the intricacies of interacting with multiple sources.

Database operation through ODBC API is carried out in the following manner. First, the connection with database is established. The scheme of access to Paradox database through ODBC interface is shown in Figure 5. In the system the Microsoft Driver Manager (odbc32.dll) interacts with odbcint.ini and odbc.ini. For the operation of Microsoft Driver Manager, which allows loading drivers, ODBC Administrator (odbcad32.exe, odbccp32.dll and odbccp32.cpl) is used. After ODBC Administrator is loaded, we can set the database name, load drivers, modify data, etc. The system has a corresponding driver for operating the databse. Afterwards, the query is run and, after getting the data, the connection is closed.

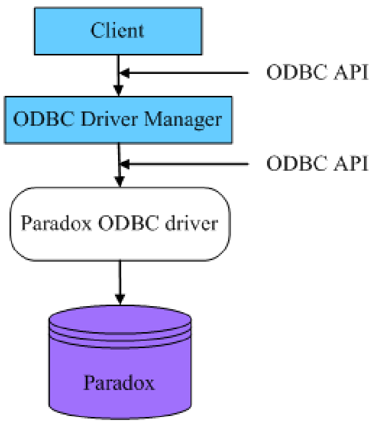


Figure 5
The scheme of access to Paradox database through ODBC interface

For operating Visual FoxPro databases the module makes use of the Microsoft JET Engine technology [14]. One of the three modules of Microsoft Jet Database Engine contains the ISAM drivers, DLL libraries allowing connections to ISAM [15] databases including Visual FoxPro. Another one of DAO modules implements the API. API allows to access JET databases through an arbitrary programming language, which the module of informational system uses for interaction with Visual FoxPro databases.

A part of module of an informational system for interaction with Oracle database was developed in an integrated development environment for database applications, PowerBuilder [16]. PowerBuilder was chosen because it uses the native interfaces for connection to Oracle and a patented technology for data operations – DataWindow. PBNI technology in PowerBuilder eliminates the flaw of long compilation time for analysis.

Switching of access modes of operating heat stations in the cluster is done by the program in accordance with the database format. At that, the module of work with relational databases uses standard technologies for accessing databases; when it works with an object-oriented database (DBMS Oracle in this case) it uses a special access code which creates mirror classes with the used OODB.

Such a mechanism of interaction with databases ensures independent operation of heat stations in the cluster.

The distribution of resources is as follows: the application server consistently produces the analysis of the forthcoming tasks (Figure 6), and then selects task fields associated with the use of a single database. The application server redirects this group of tasks to workstations in one of local area networks using the cluster node. Further, other groups of tasks are sequentially formed. The final distribution of parallel computation cluster resources is performed by the application server up to the last task. Moreover, if the task uses data stored in various databases that task is broken down into sub-tasks, and the distribution of these sub-tasks for workstations in the local area network is similar to the distribution of the tasks themselves.

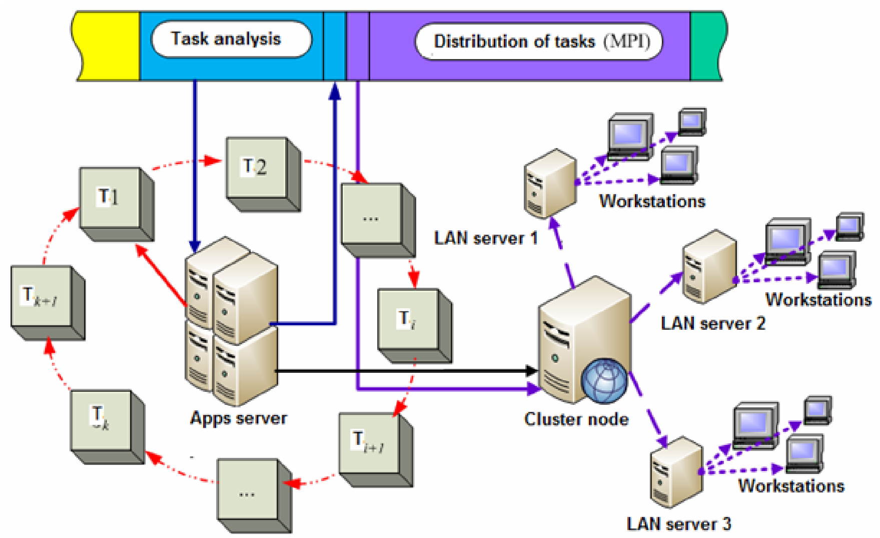


Figure 6
Scheme of distribution of information system resources

At the final stage of work, the program, using the MPI communication library [17], transfers control in accordance with the distributed tasks to the cluster workstations. Workstations independently process data from the appropriate storage and perform the calculations necessary to solve the corresponding task of the workstation (Figure 7).

The main application performs querying of all operating heat substations involved in the solution of the problem at specified intervals. The querying is carried out cyclically by searching all the working heat points; and in the case all the problems are solved, the last module is launched, integrating the results and providing documentation for the projects of construction, repair or reconstruction of heat networks.

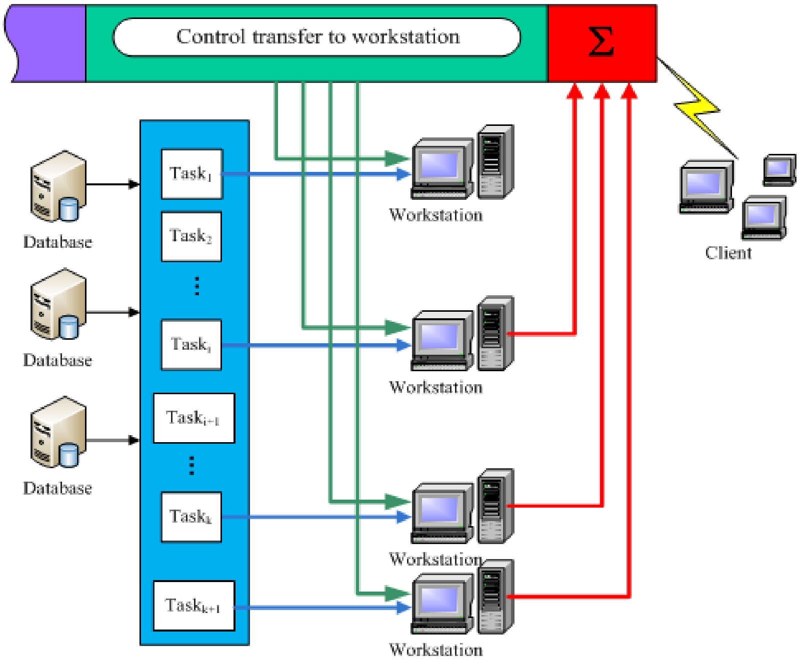


Figure 7

Scheme of IP management of workstations in a computational cluster

A client application that performs the functions of the automation unit on management of heating networks, uses received at the application server results ready for the formation of documentation on work of heating networks. In our case, it will allow to quickly process the data on thermal stations stored in different databases and at the same time to operate objects with remote access.

The organization of data protection was carried out taking into account the material of article [18, 19].

Based on the above approach, the authors developed an information system for servicing a typical heating station.

3 Implementation of Pipelined Data Processing in Heating Networks

Heat supply systems implement central, local and individual regulation.

Local regulation is applied at consumer entrance points and heat stations and and aims to adjust the mode of central regulation of heat consumption. The main factors causing the need for regulation of heat consumption for heating at entrance points can be categorised according to Table 1.

Table 1
The main factors causing the need for local regulation

Name of the factor	Description
1 Mismatch between the static characteristics of the heating system and the mode of central regulation of heat consumption.	Different values of calculated air temperatures in heated buildings (dwelling houses, schools, kindergartens, etc). Different values of calculated air temperatures of outside air (buildings of varying complexity). Mismatch of the heating surface of heating equipment installed in a building to the temperature regime of the heating network. Uneven cooling of the water in the pipes during the transport of heat carrier to various buildings.
2 Unequal dynamic characteristics of heated buildings, consumer heating systems and heating network sections from heat generator to the building.	Different thermal stability of buildings. Different dynamic characteristics of heating systems (radiator, panel, direct heating). Varied values of transport lag in the heating system (up to the building).
3 The influence on the regulating value of perturbations (temperature and heat carrier flow) acting between the heat source and the input into the building.	Operation of hot water supply installations. Inclusion of forced ventilation setups. Switching in the heating network.
4 Nonuniform nature of heat consumption.	Impossibility of implementation of central regulation according to the heating schedule in the entire range of heat demands.

The results of examination of the stated facts have shown that in the absence of local regulation in some cases there may be serious violations of the thermal regime of buildings.

Proceeding from the above, we have developed and introduced “TSmonitor”, an original software program (Analysis of parameters of a heat station), allowing to manage the system by heating network objects, calculate and accomplish other actions. The specified technique uses an informational model based on automation of a control system of heat stations with use of pipelined data processing [20, 21].

The goal of given software solution is to increase the effectiveness of control and metering of heat energy [22].

The software product represents a unified informational system in which the user can accomplish management in the optimal conditions. Informational system uses the ZULU geoinformation system, which was developed in Politerm Company (St. Petersburg, Russia), for map display [23]. The given geoinformational system gives an opportunity to draw maps with layers and use one's own object model (Activ X). The interface of the main working windows is shown in Figure 6.

For the experiment we have chosen 11 heat stations in Zyryanovsk in East Kasakhstan Oblast. A standard IP network was used to connect all the heating units to the central heating unit. In each heat station 11 parameters were registered with their minimum and maximum values, as well as the interval boundaries.

These are the following parameters (see Figure 8): the network water temperature in the supply pipeline, ^{o}C; the network water temperature in the return pipeline, ^{o}C; circulation temperature, ^{o}C; the network water flow in the supply pipeline, tons/hour; the network water flow in the return pipeline, tons/hour; the network water pressure in the supply pipeline, kgf/cm; the network water pressure in the return pipeline, kgf/cm; indoor temperature, ^{o}C; heat released, Gcal; voltage 1, V; voltage 2, V.

Besides that the following are displayed in real-time: date and time of connection and analysis; registered parameters and their value at the moment of connection.

In the software an approach using pipelined data processing was implemented. Two threads are created. The first is responsible for the process of downloading data from the "device", that is, the cycle goes through the recorded thermal nodes (TN) and downloads the previously generated parameters to the database (generated by the DataGenerate module.exe). Thus, the process of downloading data from the devices is simulated, when the server alternately surveys each node and reads data from their database.

After the survey of each node there is a delay of several milliseconds, so that to delay the process of reading. The delay can be configured in the "Connection parameters – Settings" menu. The amount of records kept in the downloaded data storage table can also be configured; after reaching the maximum the recording will be done as if into "stack". Launching and stopping a thread takes place by pressing the buttons placed on the panel. The larger the number of nodes from which the reading is taken, the slower the update process for every node.

Figure 8
The main program window

The second thread is responsible for updating the parameters of the selected node on the map. It starts automatically when the program starts. When you select a TN on the map this stream in continuous mode refers to the database and downloads the latest records on this TN and updates the appropriate field.

Using the program allows to considerably increase the reliability, longevity, as well as work efficiency both of the heat stations and of the heating system in general. Thanks to the constant monitoring of the operation of heat stations in the pipeline mode, the importance of task pipelining techniques in automated control systems was proven.

This conclusion follows from the production development theory and data transformation theory: it makes sense to organize all the same type of task sequences in the pipeline plan; the presence of task pipelining mechanisms in an automated control system leads to the increase in the work efficiency of the system.

We have charted the heat consumption in residential area in Figure 9.

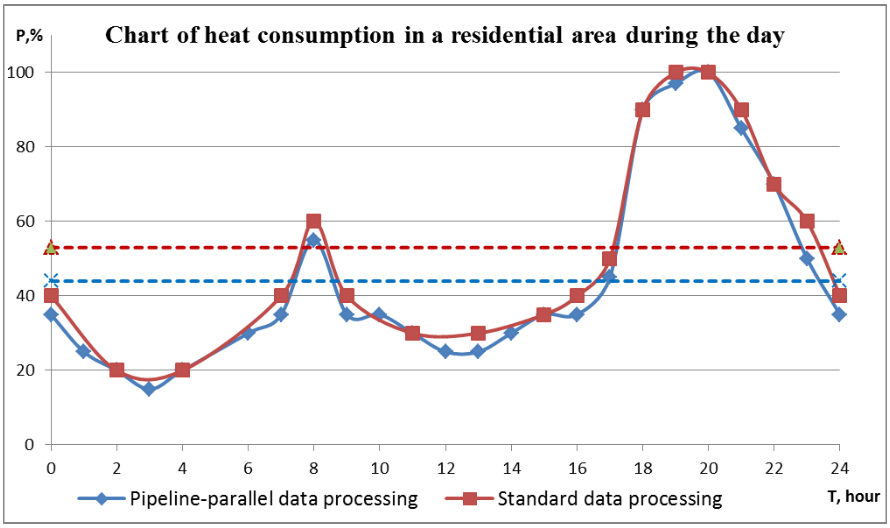


Figure 9
Heat consumption in residential area

We have made measurements of heat consumption during the day for the spring quarter. The data of heat consumption during the day was recorded while using pipeleined system (Chart 1) and while using a typical data processing system (Chart 2). It is seen that while using an informational and analytical system the heat consumption decreases by about 9%. So the average heat consumption when using pipelining data was 44% in relative units (red dotted line), whereas with the usual system it was 53% (blue dotted line).

Conclusion

The analysis of the measurements (Figure 9) showed that the use of parallel-conveyor data processing technology can significantly reduce heat and electricity consumption in the period when the heat consumption is close to or below the average level of heat consumption per day. So, in the period of time between 12-1400 hours the heat saving is 15%, and around 23 hours = 20%. It shows that the efficiency of our system increases in the period of time when there are no peak loads.

Thus, the hybrid information system that has been developed allows timely solving of complex problems of controlling heating networks operation, and regional heat stations in particular, without involving supercomputers but using a computer cluster developed on the basis of available technical support at JSC “Heating Networks”. Also, this approach can be used in image processing of minerals [24, 25].

In the future, it is planned to use intelligent mechanisms in the system [26] to improve data analysis, as we believe that this will eventually further reduce excessive heat consumption.

References

[1] Hungary National Energy Strategy, Publisher: Ministry of National Development, 2012, 131 p. ISBN 978-963-89328-3-9

[2] Brożyna, J., Mentel, G., Szetela, B: Renevable Energy and Economic Development in the European Union, Acta Polytechnica Hungarica, Vol. 14, No. 7, 2017, pp. 11-34, DOI:10.12700/APH.14.7.2017.7.2

[3] Euroheat & Power, District Energy in Hungary. Link: https://www.euroheat.org/ knowledge-centre/district-energy-hungary

[4] European Committee for Standardization. Link: https://www.cen.eu/work/ areas/energy/heatingSubstations/Pages/default.aspx

[5] Györök, Gy., Kumargazhanova, S., Baklanov, A., Uvaliyeva, I., et al.: Development of the Information and Analytical System in the Control of Management of University Scientific and Educational Activities, Acta Polytechnica Hungarica, Vol. 15, No. 4, 2018, pp. 27-44, DOI: 10.12700/APH.15.4.2018.4.2

[6] Automated process control systems of heat points. Link: https://www.krug2000.ru/decisions/solutions_zkx/teplopunkt.html

[7] Kutateladze, S. S.: Research and development of the Siberian Branch of the Russian Academy of Sciences in the field of energy efficient technologies, SBRAS Institute of Thermophysics Novosibirsk, 2009, 405 p.

[8] Ádám, N., Balaz, A., Pietriková, E., Chovancová, E. and Feciľak, P.: The Impact of Data Representationson Hardware Based MLP Network Implementation, Acta Polytechnica Hungarica, Vol. 15, No. 2, 2018, pp. 69-88, DOI: 10.12700/APH.15.1.2018.2.4

[9] Rybakova, D., Kvasov, A., Baklanov, A.: Using pipeline data processing system to regulate operation of heat stations, Vestnik EKSTU, N.3, 2014, pp. 101-105

[10] Voevodin, V., Voevodin, Vl.: Parallel computing, BHV-St Petersburg, 2002, 609 p.

[11] Shpakovsky, G. I., Serikov, N. V.: Programming for multiprocessor systems in MPI standard. Mn .: BSU, 2002, 323 p.

[12] ADO.NET. Link:https://docs.microsoft.com/ru-ru/dotnet/framework/data/adonet/index+

[13] Ganeev, R. M.: Web-interface of ODBC databases, Publisher: Hotline - Telecom, 2003, 202 p.

[14] Haught, D., Ferguson, J.: Jet Database Engine Programmers Guide (Microsoft Professional Editions) 2nd edition, Publisher: Microsoft Press; 1997, 756 p.

[15] Informational portal Bauman National Library. Link: https://ru.bmstu.wiki/Microsoft_Jet_Database_Engine

[16] Brown, M., Armstrong, F.: PowerBuilder 9: Advanced Client/Server Development. Publisher: Sams, 2003, 840 p.

[17] Pacheco, P.: Parallel Programming with MPI. Publisher: Morgan Kaufmann; 1996, 500 p.

[18] Ádám, N., Madoš, B., Baláž, A. and Pavlik, T.: Artificial neural network based IDS, 15th International Symposium on Applied Machine Intelligence and Informatics, Herl'any, Slovakia, 2017, pp. 159-164

[19] Vokorokos, L., Baláž, A., Ádám, N.: Secure Web Server System Resources Utilization. Acta Polytechnica Hungarica, Vol. 12, No. 2, 2015, pp. 5-19, DOI: 10.12700/APH.12.2.2015.2.1

[20] Rybakova, D., Baklanov, A., Kvasov, A.: The conveyor organization of the tasks in the work of the ACS heat point, Energy and fuel resources of Kazakhstan, 2010, N3, pp. 87-92

[21] Grigoryeva, S., Baklanov, A., Gyorok, Gy.: Control of LED Lighting Equipment with Robustness Elements, Acta Polytechnica Hungarica, Vol. 13, No. 5, 2016, pp. 105-119

[22] Rybakova, D., Sygynganova, I., Kumargazhanova, S., Baklanov, A., Shvets, O.: Application of a CPU Streaming Technology to Work of the Computer with Data Coming from the Network on the Example of a Heating Station. 18th International Conference of Young Specialists on Micro/Nanotechnologies and Electron Devices (EDM), Erlagol, Russia, 2017, pp. 128-130

[23] Company Politerm. Link: https://www.politerm.com/en

[24] Baklanova, O., Baklanov, M.: Methods and Algorithms of Image Recognition for Mineral Rocks in the Mining Industry, Advances in Swarm Intelligence. ICSI 2016. Lecture Notes in Computer Science, Vol. 9713, Springer, 2016, pp. 253-263, DOI:10.1007/978-3-319-41009-8_27

[25] Baklanova, O., Shvets, O.: Cluster analysis methods for recognition of mineral rocks in the mining industry. 4th International Conference on Image Processing Theory, Tools and Applications (IPTA), Paris, France, 2014, pp. 273-277, DOI:10.1109/IPTA.2014.7001972

[26] Jocic, M., Pap, E., Szakál, A., Obradovic, D., Konjovic, Z.: Managing Big Data Using Fuzzy Sets by Directed Graph Node Similarity, Acta Polytechnica Hungarica, Vol. 14, No. 2, 2017, pp. 183-200, DOI: 10.12700/APH.14.2.2017.2.10